

**UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE INGENIERÍA CIVIL**



TESIS

**“GENERACIÓN DE RUTAS ÓPTIMAS EN SISTEMAS DE
TRANSPORTE PÚBLICO MASIVO USANDO CIENCIA DE
DATOS CON BIG DATA Y MACHINE LEARNING”**

PARA OBTENER EL TÍTULO PROFESIONAL DE INGENIERO CIVIL

ELABORADO POR

BETZY SUEY TAPIA QUISPITUPA

ASESOR

DR.ING. JOSÉ CARLOS MATÍAS LEÓN

Lima - Perú

2021

© 2021, Universidad Nacional de Ingeniería. Todos los derechos reservados
“El autor autoriza a la UNI a reproducir de la Tesis en su totalidad o en parte, con fines estrictamente académicos”.

Tapia Quispitupa, Betzy Suey

btapiaq@uni.pe

961816495

DEDICATORIA

Esta tesis la dedico a la Universidad Nacional de Ingeniería, ya que gracias a esta institución me pude formar como profesional, en esta casa de estudios, no solo aprendí conocimientos, sino que también conocí grandes maestros y amigos de quienes aprendí mucho, gracias por todas las anécdotas que pase dentro de tus instalaciones, podría decir que pase los mejores momentos de mi en mi querida universidad.

Dedico esta tesis a todas aquellas personas que luchan cada día por tener un mejor país, sobre todo a aquellas personas que cada día se levantan para seguir sus sueños, que trabajan sin parar para llevar un pan a sus casas; por que gracias a los impuestos que el estado recauda de ellos, muchos estudiantes como yo, pueden educarse en instituciones prestigiosas como la UNI.

Finalmente dedico esta tesis a todas aquellas personas que tomen su tiempo en revisarla y a todas las personas que me apoyaron en este proceso, debido a que esta tesis fue elaborada en aproximadamente 2 años el cual implicó mucha dedicación y esfuerzo. Mi sueño es que esta tesis se pueda aplicar en la realidad, se que para lograrlo implicara muchos retos, tengo la esperanza que en un futuro cercano el sistema de transporte público urbano en la ciudad de Lima mejore.

AGRADECIMIENTO

Desde que inicié este proceso de la elaboración de tesis, aprendí a tener mucha paciencia, constancia, dedicación, sobre todo estuvo lleno de retos y de aprendizaje de nuevos conocimientos los cuales fueron muy satisfactorios. Al inicio quería realizar una tesis que solucionara el problema de transporte urbano, mientras evaluaba más el problema, me di cuenta, que es muy complejo solucionar ello, entonces decidí atacar parte del problema. Ya con el tema tesis alineado, ello implicaba tener conceptos de planificación en transporte, obtener la base de datos del Metropolitano, saber utilizar herramientas computacionales y programar, con los cuales al inicio no contaba, pero gracias a todas las personas que fueron apareciendo en mi vida durante este proceso pude conseguirlos.

Agradezco a mi asesor el Dr. Ing. José Carlos Matías León por ser el maestro que me guió durante este proceso de elaboración de tesis. Al MSc. Edward Santa María Davila quien a través del dictado de su curso "Economía del Transporte" me pude capacitar, al MSc. Leonardo Flores quien me facilitó una beca para capacitarse en el curso "Modelos de Simulación de Tránsito y Transporte".

Agradezco al Ing. Santiago Fernández por ayudarme que ponerme en contacto con la Gerencia de Programación de Protransporte. A los señores Fedor Daniel Morales Boluarte, Lenin Marín Bruno y Edwin Sencara por facilitarme los datos y los permisos para ingresar al Metropolitano

Agradezco al Centro de Tecnologías de la Información y Comunicaciones CTIC UNI por el apoyo incondicional que me brindó desde el día que ingresé a sus laboratorios. Gracias al MSc. Alonso Tenorio Trigos y MSc. Víctor Melchor Espinoza por facilitarme la beca en el "Especialización en Data Science for Business ". Al ingeniero PhD. Manuel Castillo Cara del laboratorio de Smart City por apoyarme con el tema de Machine Learning. Al MSc. Yuri Nuñez Medrano por permitirme ser parte de los investigadores del Laboratorio 4 Big Data y HPC. A los compañeros de este laboratorio Gabriel Vicente, Gianfranco Ferro, Enrique Peinado, Renzo Cerrón, por el apoyo y guía en la elaboración del código del

algoritmo. En este laboratorio conocí muchas amigos de las cuales aprendí mucho y por ello estoy muy agradecida.

Agradezco a mis amistades quienes me aconsejaban y sobre todo en cada momento me recordaban que tenía que terminar la tesis, ello me ayudó bastante a no rendirme y seguir adelante.

Agradezco a mi familia, por el apoyo incondicional, a mis padres y hermanos, por el apoyo constante.

Gracias a todas estas personas por haberme enseñando, ayudado y aconsejado.

ÍNDICE

	Pág.
RESUMEN	1
ABSTRACT	2
PRÓLOGO	3
LISTA DE TABLAS	4
LISTA DE FIGURAS	5
LISTA DE SÍMBOLO Y SIGLAS	8
CAPÍTULO I : INTRODUCCIÓN	9
1.1 GENERALIDADES	9
1.2 PLANTEAMIENTO DE LA REALIDAD PROBLEMÁTICA	11
1.3 OBJETIVOS	15
1.3.1 Objetivo General	15
1.3.2 Objetivos Específicos	15
1.4 HIPÓTESIS	15
1.4.1 Hipótesis General	15
1.4.2 Hipótesis Específicos	15
CAPÍTULO II : FUNDAMENTO TEÓRICO	16
2.1 ANTECEDENTES	16
2.2 MARCO CONCEPTUAL	19
2.2.1 Optimización	19
2.2.2 Optimización de Rutas de Transporte de Pasajeros	19
2.2.3 Transporte rápido masivo	20
2.2.4 Análisis del sistema de transporte	24
2.2.5 Modelo de las cuatro etapas (FSM)	26
2.2.6 Data Science	29
2.2.7 Big Data	31
2.2.8 Machine Learning	34
2.2.9 Algoritmo	36
2.2.10 Seudocódigo	36
2.3 INFORMACIÓN REFERENTE AL CASO DE ESTUDIO	37
2.3.1 Transporte urbano en Lima	37
2.3.2 Metropolitano	38
2.3.3 Simulación	41
CAPÍTULO III : ANÁLISIS METODOLÓGICO	42
3.1 DISEÑO METODOLÓGICO	43

3.1.1	Serie de Pasos	43
3.1.2	Herramientas	43
3.2	METODOLOGÍA PARA EL CASO DE ESTUDIO	48
	CAPÍTULO IV : TRATAMIENTO DE LA INFORMACIÓN	51
4.1	HERRAMIENTA 1: CIENCIA DE DATOS	51
4.1.1	Eta 1: Comprensión del Negocio	51
4.1.2	Eta 6: Implementación	53
4.2	HERRAMIENTA 2: BIG DATA	54
4.2.1	Paso 1: Recolección de datos	55
4.2.2	Paso 2: Almacenamiento de datos (lectura)	59
4.2.3	Paso 3: Data Analytics (Pre procesamiento)	63
4.2.4	Paso 4: Limpieza de los datos	73
4.2.5	Paso 5: Filtrado o selección de parámetros	74
4.2.6	Paso 6: Estructura de los datos	75
4.2.7	Paso 7: Análisis de los datos	77
4.3	HERRAMIENTA 3: MACHINE LEARNING	85
4.3.1	Definir el problema	86
4.3.2	Preparar los datos	86
4.3.3	Evaluar un conjunto de algoritmos	91
	CAPÍTULO V : MODELO DE GENERACIÓN DE RUTAS ÓPTIMAS	94
5.1	PASO 8: GENERACIÓN DE RUTAS	94
5.1.1	Algoritmo: Generación de rutas	94
5.1.2	Caso – Metropolitano	100
5.2	PASO 9: PROGRAMACIÓN DE RUTAS	102
5.2.1	Caso – Metropolitano	102
	CAPÍTULO VI : ESCENARIOS FUTUROS	104
6.1	PASO 10: VALIDACION DEL MODELO	104
6.1.1	Calibración	105
6.1.2	Macro simulación situación actual	106
6.1.3	Macro simulación situación propuesta	117
6.2	PASO 11: ANÁLISIS COMPARATIVO SITUACIÓN ACTUAL Y PROPUESTA	121
6.2.1	Rutas: Cantidad de viajes	121
6.2.2	Matriz de resultados	121
6.2.3	Perfil de Tiempo	124
6.2.4	Número de transferencias por estación	127

CONCLUSIONES	130
RECOMENDACIONES	132
REFERENCIAS BIBLIOGRAFICAS	133
ANEXOS	138

RESUMEN

Se quiere mejorar el Sistema de Transporte Público Masivo, mediante a la optimización de rutas. Se elabora la propuesta de modelo de "Generación de rutas óptimas" para un sistema de transporte donde se disponga la data de los usuarios y algunos parámetros de operación, del modelo se obtiene rutas óptimas en función de la demanda de los usuarios. Con el fin de comprobar la eficacia del modelo se toma como caso de estudio de un Sistema de Transporte Público Masivo de la ciudad de Lima, caso específico "Metropolitano" (Sistema de transporte público, urbano y masivo de Lima), ya que este sistema cuenta con una ruta que conecta extremos de la ciudad (distancia directa) y con vía exclusiva (tiempo óptimo sin obstáculos) y con la data de los usuarios, teniendo como resultado transportar a la mayor cantidad de usuarios en el menor tiempo posible.

El modelo que se propone se basa en generar rutas óptimas a través de la demanda de los usuarios; para ello se hará el uso de la Ingeniería de Transporte y herramientas de la Ciencia de la Computación como Ciencia de Datos, Big Data y Machine Learning, con la finalidad de usar los criterios de transporte urbano para procesar los datos, codificar el algoritmo y simular la generación de las rutas óptimas. A demás de ello cabe resaltar que la propuesta servirá para cualquier sistema de transporte que haga uso de tarjetas electrónicas o tenga el origen-destino de los usuarios. Al generar las rutas, en el "Metropolitano", con el algoritmo del modelo de "Generación de rutas óptimas", en la hora pico de 7 a 8am, se obtuvo como resultado 347 viajes lo que implica una mejora en 51 viajes comparado con los 398 viajes que usa el sistema del "Metropolitano". Al optimizar las rutas se reduce el Tiempo de espera en 66.6%, el Tiempo de transferencia en 83.44% y el Número de transferencia en 69.06%.

ABSTRACT

The objective is to improve the Mass Public Transportation System by means of route optimization. The proposal of a model of "Generation of optimal routes" for a transportation system where the data of the users and some operation parameters are available is elaborated, from the model optimal routes are obtained according to the demand of the users. In order to test the effectiveness of the model, a case study of a Mass Public Transportation System in the city of Lima is taken as a specific case "Metropolitano" (Lima's public, urban and mass transportation system), since this system has a route that connects ends of the city (direct distance) and with an exclusive route (optimal time without obstacles) and with the users' data, resulting in transporting the largest number of users in the shortest possible time.

The proposed model is based on generating optimal routes through user demand; for this purpose, Transportation Engineering and Computer Science tools such as Data Science, Big Data and Machine Learning will be used, with the purpose of using urban transportation criteria to process the data, codify the algorithm and simulate the generation of optimal routes. In addition, it should be noted that the proposal will serve for any transportation system that makes use of electronic cards or has the origin-destination of users. When generating the routes, in the "Metropolitano", with the algorithm of the "Generation of optimal routes" model, in the peak hour from 7 to 8am, 347 trips were obtained as a result, which implies an improvement of 51 trips compared to the 398 trips used by the "Metropolitano" system. By optimizing the routes, the waiting time was reduced by 66.6%, the transfer time by 83.44% and the number of transfers by 69.06%.

PRÓLOGO

El servicio de transporte público en el Perú es de mala calidad, uno de los problemas que hace que el servicio sea malo, es la programación de los buses a lo largo del recorrido de la ruta. Ante este problema se busca una propuesta de solución, que consiste en generar una programación de rutas en base a la demanda de los usuarios, para ello se desarrolló una serie de pasos, que están agrupados en los capítulos 4, 5 y 6. Para el desarrollo de estos pasos, como caso de estudio, se utilizó los datos de los usuarios del sistema de transporte masivo de los Buses de Tránsito Rápido (BRT) de la ciudad de Lima, caso específico “Metropolitano”.

El **primer capítulo** consta principalmente de la problemática, la hipótesis y los objetivos. El **segundo capítulo** explica los conceptos e información que son relevantes para el entendimiento de la tesis. El **tercer capítulo** se explica el diseño metodológico que se aplicó, es decir los pasos para obtener la generación de rutas óptimas, así como para el caso de estudio del Metropolitano. El **cuarto capítulo** explica cómo se realizó el tratamiento de la información haciendo uso de las herramientas de la información como “Ciencia de Datos”, “Big Data” y “Machine Learning”. El **quinto capítulo** se explica el algoritmo de “Modelo de generación de rutas óptimas” y como es aplicado en el caso de estudio. En el **sexto capítulo** se realizó un análisis comparativo del caso de estudio, es decir la situación “actual” y “propuesta” de las rutas generadas. La tesis culmina con las conclusiones y recomendaciones.

Se sugiere leer esta tesis de manera secuencial, debido a que está constituida principalmente por una serie de pasos relacionados.

LISTA DE TABLAS

	Pág.
Tabla 1 Modos de Transporte.....	12
Tabla 2 Porcentajes del transporte	12
Tabla 3 Intervención de la secuencia de pasos	50
Tabla 4 Muestra del archivo de la data brindada por PROTRANSPORTE.....	60
Tabla 5 Coordenadas de las Estaciones del Metropolitano.....	60
Tabla 6 Velocidades promedio de los BRT	61
Tabla 7 Capacidad máxima del BRT.....	61
Tabla 8 Despachos en Hora pico Mañana	61
Tabla 9 Tiempo por ciclo.....	62
Tabla 10 Salidas en la Hora pico Mañana	62
Tabla 11 Horarios de salida de buses.....	62
Tabla 12 Salidas en la Hora pico Mañana	62
Tabla 13 Cantidad de usuarios	63
Tabla 14 Días mínimos y máximos	64
Tabla 15 Promedios “Mono”	66
Tabla 16 Coordenadas de la estación Jr. De la Unión y Plaza Flores.....	71
Tabla 17 Limpieza de datos.....	73
Tabla 18 Validaciones filtradas	74
Tabla 19 Filtrado de datos	74
Tabla 20 Estructura de datos.....	76
Tabla 21 Cantidad de validaciones por minuto	82
Tabla 22 Cantidad de validaciones en Origen-Destino	83
Tabla 23 Tipo de datos para predicción	92
Tabla 24 Algoritmos de predicción con un día de datos.....	92
Tabla 25 Algoritmo BayesBayesNet según aumenta la data.....	92
Tabla 26 Algoritmo BayesBayesNet según aumenta los intervalos.....	93
Tabla 27 Formato de salida	99
Tabla 28 Resumen rutas del algoritmo de “Generación de rutas óptimas”.....	101
Tabla 29 Resumen de rutas de “Generación de rutas óptimas”	101
Tabla 30 Listado de algunas rutas Directas	102
Tabla 31 Listado de algunas rutas Regulares 1	102
Tabla 32 Listado de algunas rutas Regulares 2	103
Tabla 33 N° de salidas de bus situación actual.....	109
Tabla 34 N° de salidas de buses en “Situación Propuesta”.....	117
Tabla 35 Número de salida de buses ahorro	121
Tabla 36 Comparación de los tiempos de viaje.....	122
Tabla 37 Comparación de los tiempos de espera de transferencia	122
Tabla 38 Comparación de número de transferencia	123
Tabla 39 Coordenadas de las estaciones del Metropolitano	140
Tabla 40 Coordenadas para el modelo del Metropolitano	141
Tabla 41 Recolección de datos: capacidad máxima del bus	144
Tabla 42 Capacidad máxima del Bus.....	144
Tabla 43 Programación de las rutas directas con total de 89.....	150
Tabla 44 Programación de las rutas “Regulares 1” y “Regulares 2”.....	152
Tabla 45 Programación de las rutas “Regulares 2”	153

LISTA DE FIGURAS

	Pág.
Figura 1 Variación de los modos de transporte.	12
Figura 2 Transporte público vs Transporte público masivo	13
Figura 3 Cantidad de vehículos	13
Figura 4 Densidad de tráfico según el modo de transporte.	22
Figura 5 Red de Análisis del sistema de transporte	25
Figura 6 TSA y FSM	26
Figura 7 Modelo de las Cuatro Etapas	28
Figura 8 Campo interdisciplinario	30
Figura 9 Etapas y tareas de un proyecto de Ciencia de Datos (CRISP-DM)	30
Figura 10 IBM, características del Big Data 3Vs	31
Figura 11 Modelo de operación del Big Data.	33
Figura 12 Mapa conceptual de los aspectos claves del ML	35
Figura 13 Red básica del Metro de Lima y el Callao	37
Figura 14 Ruta de circulación del Metropolitano	38
Figura 15 Propuesta para la generación de rutas óptimas	42
Figura 16 Metodología: pasos a seguir	43
Figura 17 Esquema de herramientas	43
Figura 18 Etapas de la Ciencia de Datos	44
Figura 19 Intervención de la Ciencia de Datos en la secuencia de pasos	45
Figura 20 Procesos del Big Data	46
Figura 21 Intervención del Big Data en serie de pasos	46
Figura 22 Procesos de Machine Learning	46
Figura 23 Intervención del Machine Learning en la secuencia de pasos	47
Figura 24 Metodología: agrupamiento de pasos	48
Figura 25 lenguajes que intervendrán	49
Figura 26 Software que intervendrán	49
Figura 27 Metodología para generación de rutas del Metropolitano	50
Figura 28 Validación de un usuario	55
Figura 29 Ruta del Metropolitano	56
Figura 30 Ejemplo de obtención de coordenadas	56
Figura 31 Conductor de un bus del Metropolitano	57
Figura 32 Capacidad máxima en un bus Expreso 2	57
Figura 33 Validación de la salida de buses estación Naranjal	58
Figura 34 Vía troncal a hacer analizada	60
Figura 35 Tipo de validación	64
Figura 36 Parada: estaciones de los alimentadores	65
Figura 37 Parada: estaciones de la troncal	65
Figura 38 Tipo de validador	65
Figura 39 Perfil	66
Figura 40 Monto	66
Figura 41 Recaudación "Monto"	66
Figura 42 Porcentajes	67
Figura 43 Validaciones según hora	68
Figura 44 Validaciones en Troncal	69
Figura 45 Validaciones entre las 0:20:00 a 4:40:00 am madrugada	69
Figura 46 Validaciones entre las 4:40:00 a 0:20:00 am día	70
Figura 47 Matriz Origen-Destino del día 4/05/2018	78
Figura 48 Demanda en cada estación del día 4/05/2018	79
Figura 49 Fluencia de viajes del día 4/05/2018	80

Figura 50	Distribución en función del tiempo del día 4/05/208	81
Figura 51	Distribución en intervalo de 5 min del día 4/05/208	81
Figura 52	Distribución en intervalo de 1 min de hora pico	82
Figura 53	Mapa burbuja y mapa de calor del origen-destino	84
Figura 54	Pasos y proceso del Machine Learnig	85
Figura 55	Tipo de datos por defecto	86
Figura 56	Tipo de datos modificado según el atributo	86
Figura 57	Estadística descriptiva del atributo Fecha	87
Figura 58	Distribución de los atributos	87
Figura 59	Matriz de correlación	88
Figura 60	Delimitar valores extremos	88
Figura 61	Eliminar valores extremos	89
Figura 62	Selección de atributos	89
Figura 63	Troncal, eliminación del atributo “tipo de validación”	90
Figura 64	Eliminación del perfil “PL”	90
Figura 65	Precisión del algoritmo BayesBayesNet variando la data	93
Figura 66	Precisión del algoritmo BayesBayesNet variando los intervalos	93
Figura 67	Características del Sistema de transporte	106
Figura 68	Características de la via de la red del “Metropolitano”	106
Figura 69	Red del “Metropolitano”	107
Figura 70	Operador del “Metropolitano”.	108
Figura 71	Caracteristicas del bus	108
Figura 72	Rutas que circulan de 7 a 8am	109
Figura 73	Detalle Ruta Expreso 1 con dirección N-S	110
Figura 74	Matriz 38x38 Origen-Destino de 7 a 8 am	111
Figura 75	Lineas de deseo entre zonas 7 a 8 am	112
Figura 76	Asignación del volumen en la red 7 a 8 am	113
Figura 77	Distribución del volumen en la red 7 a 8 am	114
Figura 78	Asignacion para el calculo de matrices de resultado	115
Figura 79	Matriz de resultado, tiempo de viaje (2 JRT)	115
Figura 80	Detalle del recultado de caminos entre zonas	116
Figura 81	Grupo de rutas - Propuesta	117
Figura 82	Rutas de la Propuesta	118
Figura 83	Detalle de la ruta regular N°2 en dirección N-S	119
Figura 84	Matrices de resultados	119
Figura 85	Detalle del resultado de caminos entre zonas	120
Figura 86	Listado de matrices resultado, simulación actual	121
Figura 87	Listado de matrices resultado, simulación propuesta	122
Figura 88	Resumen de las matrices de resultado	123
Figura 89	Perfil de Tiempo - Actual	125
Figura 90	Perfil de Tiempo - Propuesta	126
Figura 91	Transferencia en estación-Actual	127
Figura 92	Transferencia en estación-Propuesta	128
Figura 93	N°de transferencias por estación	129
Figura 94	Comparación del N°de transferencias por estación	129
Figura 95	Datos brindados por Protransporte	138
Figura 96	Documento de entrega de datos	139
Figura 97	Red de transporte Metropolitano-Troncal	142
Figura 98	Primera parada del Expreso 2	143
Figura 99	Segunda parada del Expreso 2	143
Figura 100	Toma de datos en el Expreso 2	143
Figura 101	Tiempo de viaje (2 JRT) - Actual	154
Figura 102	Tiempo de viaje (2 JRT) - Propuesta	155

Figura 103	Tiempo de espera transferencia adaptado (4 TWTA) - Actual	156
Figura 104	Tiempo de espera transferencia adaptado (4 TWTA) - Propuesta	157
Figura 105	Número de transferencias (5 NTR) - Actual	158
Figura 106	Número de transferencias (5 NTR) - Propuesta	159

LISTA DE SÍMBOLO Y SIGLAS

A	: Alimentador
AATE	: Autoridad Autónoma del Tren Eléctrico
ATU	: Autoridad de Transporte Urbano
BRT	: Buses de Tránsito Rápido
COSAC	: Corredores Segregados de Buses de Alta Capacidad
CSV	: Valores Separados por Comas
ES	: Escolar
FSM	: Modelo de las cuatro etapas
GE	: General
GEI	: Gases de Efecto Invernadero
GIZ	: Sociedad Alemana de Cooperación Internacional
GTU	: Gerencia de transporte Urbano
JICA	: Agencia de Cooperación Internacional de Japón
LRT	: Tren de Tránsito Ligero
M	: Minusválido
ML	: Machine Learning
MML	: Municipalidad Metropolitana de Lima
MRT	: Transporte Rápido Masivo
MTC	: Ministerio de Transporte y Comunicaciones
N	: Normal
PL	: Pase Libre
PMTU	: Plan Maestro de Transporte Urbano
RFID	: Identificación por Radio Frecuencia
SIT	: Sistema de Transporte Integrado
T	: Troncal
TAZ	: Zona de Análisis de Tráfico
TSA	: Análisis del Sistema de Transporte
UN	: Universitario

CAPÍTULO I : INTRODUCCIÓN

1.1 GENERALIDADES

Lo que se ambiciona con esta tesis, es generar rutas óptimas basado en la demanda de los usuarios. Si se tendrían los orígenes destinos de todos los usuarios de toda una ciudad se podría saber cómo se desplazan en esta área, es por ello que la generación de las rutas se basa en los usuarios, así se podría generar nuevas rutas basadas en cómo se desplazan los ciudadanos, haciendo que estas rutas sean modificadas de acuerdo a como se mueve la ciudad, esto traería como consecuencia rutas directas donde los sistemas de transporte masivo llevan a los pasajeros de un lugar a otro como cuando se toma un taxi, esto implicaría tener un camino eficiente que radica en usar la menor distancia libre de congestión vehicular (tiempo óptimo). Una vez que se tiene el camino eficiente genérico (mínima distancia y tiempo), se pasa a generar rutas optimizadas para ese camino, eso quiere decir obtener la cantidad de buses y la hora en la que pasarán los buses por determinadas estaciones, generando una cantidad de rutas para una línea de transporte que hará uso del camino eficiente. Lo que se pretende en futuras investigaciones es predecir el comportamiento de la demanda de los usuarios así obtener rutas que varían según la estación del año (mes, día y hora).

Como se explicó en reglones anteriores para generar las rutas óptimas se necesitan de caminos eficientes, se buscó algún caso cercano que cumpla con estas características, teniendo como caso de estudio al Metropolitano de la ciudad de Lima. Se hace uso de los datos de los usuarios de este sistema para analizar y procesar con la finalidad de generar las rutas optimizadas; se mostrará los resultados del análisis con gráficos par un mejor entendimiento, así mismo se hace la simulación de este sistema para comparar el estado actual con la propuesta de las rutas óptimas. Finalmente se hace la evaluación en cuanto mejora el nivel de servicio del sistema, mejoras de los tiempos y ahorros en los costos operativos del sistema. Cabe resaltar que en esta tesis se utilizan dos ramas que son ciencia de la computación e ingeniería de transporte. Del procesamiento de los datos se podrá observar en los capítulos siguientes que el sistema de transporte del Metropolitano fue creado para la zona norte de la ciudad de Lima, la cantidad de usuarios de cada estación, horas pico y valle, etc.

Para el desarrollo de este trabajo de investigación se desarrollan seis capítulos los cuales son (1) Introducción, principalmente se desarrollará los objetivos e hipótesis de la tesis. (2) Fundamento teórico, se presentará los antecedentes, los conceptos a usar y los conceptos de los términos de los datos. (3) Análisis metodológico, aquí se planteará una metodología para la generación de las rutas, así como explicar el caso de estudio. (4) Procesamiento de información, se explicará a detalle los primeros siete pasos planteados en la metodología interviniendo los conceptos de Ciencia de Datos, Big Data y Machine Learning. (5) Optimización de rutas, se explica a detalle los pasos ocho y nueve, planteando el algoritmo para generar las rutas óptimas y la programación de los usuarios. (6) Escenarios futuros, se explicará a detalle los últimos dos pasos (diez y once), se hará una macro simulación en escenario actual y propuesto, a su vez se medirá algunos parámetros como nivel de servicio, tiempo de espera y gastos operativos.

Cabe mencionar que este problema se puede resolver de manera convencional con el Modelo Clásico del Transporte o Modelo de las Cuatro Etapas, no usamos este modelo ya que se basa en muestra de la población sin embargo usaremos los pasos de este modelo para implementarlo en el modelo que estamos planteando.

Con la finalidad de mejorar el modelo para más adelante se recomienda hacer las predicciones de las rutas, de acuerdo a la variabilidad de la estación del año, es decir, cuáles serán las rutas más concurridas en un determinado día de la semana, mes (vacaciones, escolaridad y feriados); hacer una plataforma web y aplicativo móvil para visualización de las estaciones más demandadas, afluencia de los pasajeros en las estaciones, rutas más concurridas, llegada de los buses, etc. Esta mejora del modelo se recomienda hacer cuando se tenga la data de una ciudad completa.

1.2 PLANTEAMIENTO DE LA REALIDAD PROBLEMÁTICA

La problemática que existe en la ciudad de Lima es el crecimiento desproporcionado y sin planeamiento territorial de la población; ello conlleva a uno de los problemas más grandes de esta ciudad que es el congestionamiento vehicular y el caos del transporte urbano. Se agrava el problema por la no existencia de un sistema de transporte integrado que articule el transporte público, transporte privado, ciclista y al peatón. Otro problema es que el tiempo de viaje sea mayor, la existencia de vehículos antiguos, la aparición de vehículos colectivos y el aumento de la contaminación del aire (mayor cantidad de emisiones de dióxido de carbono). Cabe resaltar que estos problemas son debidos a un mal ordenamiento territorial y mala gestión pública. (GIZ, 2018)

A continuación, se muestra datos de los problemas más relevantes, referidos al transporte.

Los tiempos de viaje promedio de las personas en la ciudad de Lima se han incrementado 21%, lo que representa pérdidas anuales en horas hombre y costos operativos de USD 500 MM. (Lanegra, Zevallos, Mora, & Apaella, 2016).

En el Perú, las emisiones de Gases de Efecto Invernadero (GEI) ascenderán aproximadamente a 140 MtCO₂ eq (gases emitidos en miles de toneladas equivalentes en CO₂), de los cuales aproximadamente el 10% corresponde al sector transporte con 14.9 MtCO₂ eq, lo que representa el 40% de las emisiones relacionadas con la energía (GIZ, 2014). Se estima que estas emisiones aumentarán en un 200% para el año 2050, sino se implementan medidas adecuadas. (GIZ, 2018)

Los estudios del Plan JICA (Agencia de Cooperación Internacional de Japón) en la región Metropolitana de Lima y Callao la cantidad de viajes realizados en el 2004 fue de 16.5 millones de viajes por día y en 2012 fue de 22.3 millones de viajes diarios (JICA, 2013). La GIZ (Sociedad Alemana de Cooperación Internacional) en el 2018 realizó un estudio similar, donde la cantidad de viajes diarios es de 22.8 millones (GIZ, 2018). Ambos estudios muestran que la participación de modo durante los años 2004, 2012, 2018 y 2028 es similar (Tabla 1), sin embargo, se presenta una distribución del modo de transporte de Bus (2012) a BRT y tren (2018).

Tabla 1 Modos de Transporte.

Año	2004		2012		2018		2028	
Población	7.496.342		8.481.415		9.684.302		10.911.223	
Modo	Nº de viajes	Modo Compartido	Nº de viajes	Modo Compartido	Nº de viajes	Modo Compartido	Nº de viajes	Modo Compartido
Caminar	4.208.000	24,8%	5.416.400	24,3%	5.549.696	24,3%	6.252.797	25,7%
Bicicleta	84.000	0,5%	77.300	0,3%	78.901	0,3%	88.897	0,4%
Motocicleta	30.000	0,2%	107.300	0,5%	109.641	0,5%	123.532	0,5%
Auto Privado	2.008.000	11,8%	3.445.400	15,4%	3.485.672	15,2%	5.159.646	21,2%
Mototaxi	600.000	3,5%	1.325.300	5,9%	1.357.708	5,9%	1.529.719	6,3%
Taxi	1.325.000	7,8%	591.300	2,7%	605.589	2,6%	682.312	2,8%
Colectivo	181.000	1,1%	333.000	1,5%	341.220	1,5%	384.450	1,6%
Bus	8.524.000	50,3%	10.664.000	47,8%	10.290.254	45,0%	9.200.986	37,8%
BRT	0	0,0%	274.000	1,2%	650.000	2,8%	732.350	3,0%
Tren	0	0,0%	74.000	0,3%	390.000	1,7%	160.000	0,7%
TOTAL	16.960.000	100,0%	22.308.000	100,0%	22.858.681	100,0%	24.314.689	100,0%

Fuente: Equipo de Estudio JICA, encuesta de viaje personal. (JICA, 2013)

Fuente: Equipo de Estudio GIZ, encuesta de viaje personal. (GIZ, 2018)

En la Figura 1 se muestra como varían los porcentajes, cantidad de viajes y población de los modos de transporte en base a los estudios del Plan JICA (2004 y 2013) y de la GIZ (2018), de ello se observa que el modo auto privado está en aumento, lo cual trae como consecuencia mayor incremento en el tráfico.

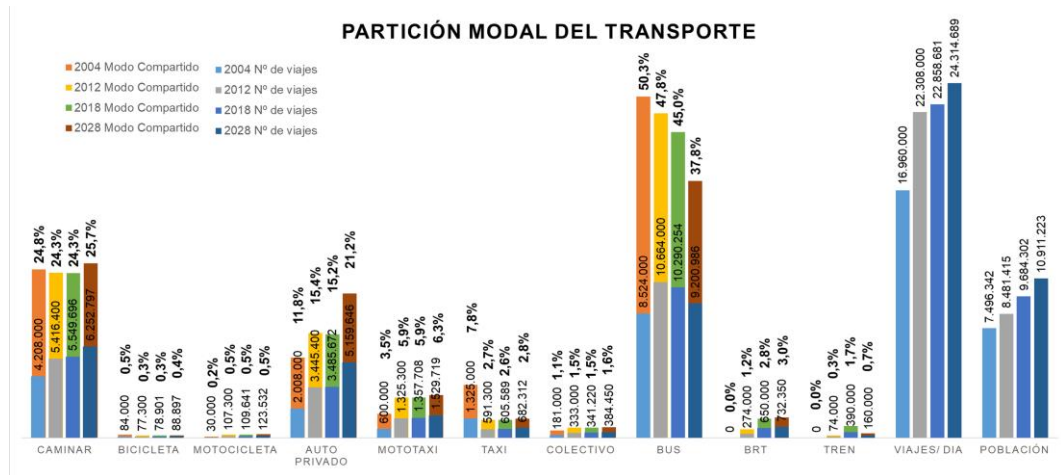


Figura 1 Variación de los modos de transporte.

Fuente: Elaboración propia con datos de (JICA, 2013) y (GIZ, 2018)

En la Tabla 2 se muestra un promedio aproximado de los porcentajes del transporte durante los años 2004, 2012, 2018 y 2028. A su vez se ve que el transporte público y transporte público masivo tiende a decrecer (Figura 2).

Tabla 2 Porcentajes del transporte

Transporte	No Motorizado	25%				Caminar
	Motorizado	75%	15%			Bicicleta
			Público	60%	10%	Motocicleta
						Auto Privado
					Público Masivo	50%
			Taxi			
			Colectivo			
Tren						

Fuente: Elaboración propia con datos de (JICA, 2013) y (GIZ, 2018)

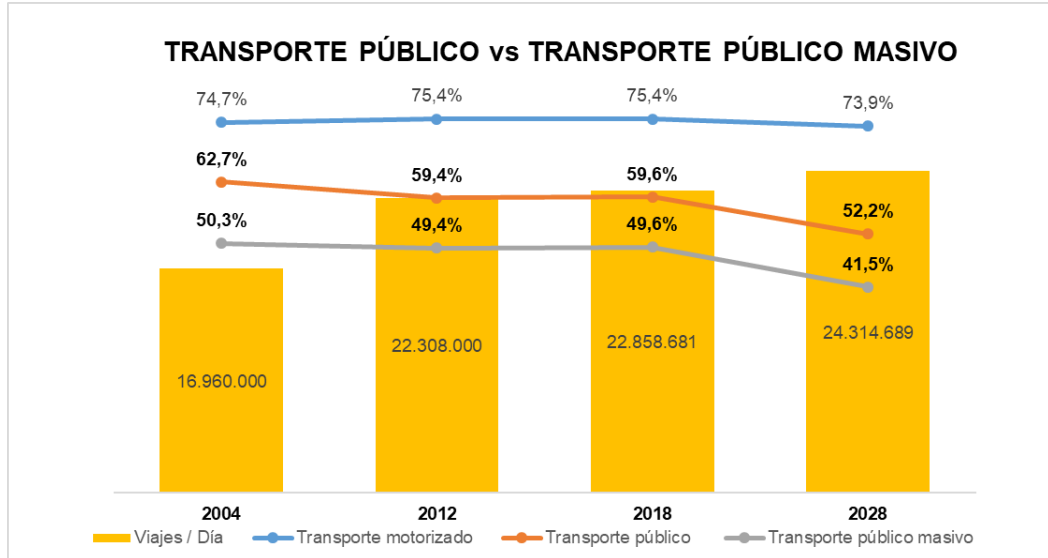


Figura 2 Transporte público vs Transporte público masivo
Fuente: Elaboración propia con datos de (JICA, 2013) y (GIZ, 2018)

Un informe elaborado por el Ministerio de Transportes y Comunicaciones (2010) que reúne una serie de datos relevantes sobre el transporte urbano metropolitano de Lima y Callao, refiere que según información de la Municipalidad Metropolitana de Lima se habían registrado hasta marzo de 2008 más de 600 rutas de transporte público (Lima 432 rutas urbanas y 55 rutas de interconexión y en provincia del Callao 117 rutas). De esta cantidad, el 73% fueron autorizadas por la provincia de Lima y el otro 27% por el Callao. De los datos actualizados de la Gerencia de transporte Urbano (GTU) existen 339 rutas y 15 167 vehículos autorizados para el transporte público, de estos últimos se tiene un padrón de 14 870 vehículos (GTU, 2019). En la Figura 3 se muestra la cantidad de vehículos que circulan en las rutas de Lima; el número de asientos es el promedio según el tipo de vehículo.

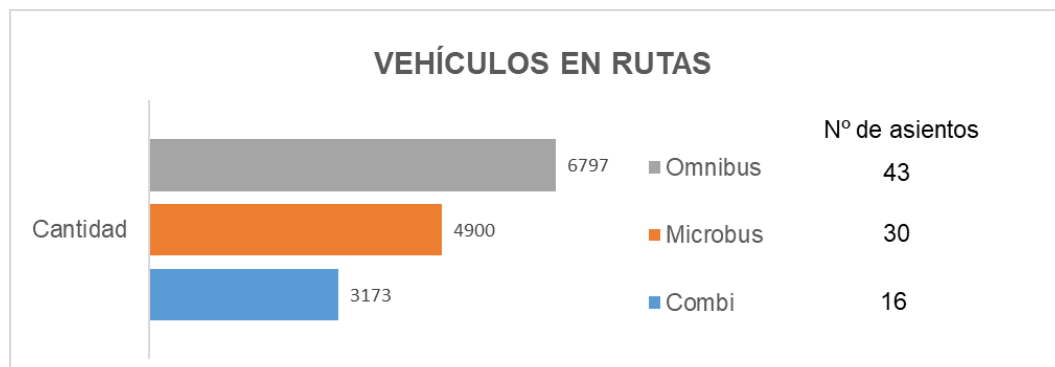


Figura 3 Cantidad de vehículos
Fuente: Elaboración propia con datos de (GTU, 2019b)

Uno de los sistemas de transporte masivo de alta capacidad más concurridos en la ciudad de Lima es el Metropolitano, el cual tiene una vía exclusiva; sin embargo, no es el más eficiente; debido a que hay dos principales problemas. (1) Rutas programadas erróneamente, puesto que el modelo que usan se basa en una muestra de los usuarios, ello trae como consecuencia, que en algunas estaciones concurren más usuarios y en otras menos, de igual forma que algunos buses estén llenos y otros vacíos. (2) Tiempo de espera, se da, ya que el usuario no puede transportarse debido a que el bus no se dirige a su destino o está en su capacidad máxima; actualmente el tiempo de espera en las estaciones son entre 10 a 45 min, esto trae como consecuencia, tiempo de viaje sea mayor. Estos problemas conllevan a reprogramar las rutas y minimizar el tiempo de viaje para disminuir los gastos operativos.

Con un plan estratégico, se puede planificar un buen sistema de transporte integrado y darle solución a la problemática que se tiene todos los días en la ciudad de Lima y Callao. El sistema de transporte integrado articula el transporte motorizado y no motorizado; ello implica una mejora en la ruta (optimizar ruta del transporte público), programación (salida y llegada transporte público), habilitar paraderos autorizados, establecer aforos (capacidad del vehículo y paraderos), regulación de las tarifas de los pasajes (sea por tiempo, distancia o zona), regular el transporte privado con políticas como establecer horarios de salidas y circulación por determinadas zonas, crear nueva infraestructura para los modos de transporte que se requiera. Con todo ello disminuirá la congestión vehicular y los tiempos de viajes dentro de una ciudad. Como se explicó anteriormente, el transporte motorizado representa el 75% de los viajes; ello implica que, si hace una mejora en el sistema, se estará solucionando en gran medida esta problemática. Es por ello que esta propuesta buscará contribuir a tener una base para un Sistema de Transporte Integrado (SIT), el cual consistirá en optimizar las rutas para el transporte público, tomando como caso de estudio a los Buses de Tránsito Rápido (BRT), caso específico del Metropolitano. La optimización de la ruta será tomando en cuenta la demanda de los usuarios, haciendo uso de una gran cantidad de datos para tener la ruta óptima, lo cual podrá ser aplicado para cualquier sistema de transporte. La propuesta podrá ser aplicada en cualquier sistema de transporte siempre en cuando se tenga la data de los usuarios.

1.3 OBJETIVOS

1.3.1 Objetivo General

El objetivo general consiste en proponer rutas óptimas para Sistema de Transporte Público Masivo en vía segregada utilizando la Ciencia de Datos (Data Science) con metodologías del Big Data y Machine Learning.

1.3.2 Objetivos Específicos

Los objetivos específicos son los siguientes:

- Pre procesamiento de los datos generales del Metropolitano (Data Analytics).
- Procesamiento con Big Data: Recopilar, leer, limpiar, filtrar (seleccionar), estructurar y analizar los datos de los usuarios del “Metropolitano”.
- Utilizar algoritmos de Machine Learning y plantear algoritmos para obtener las mejores rutas de acuerdo al origen-destino de los usuarios y mostrar la programación de las rutas.
- Elaborar una Macro simulación con la programación rutas actuales y propuestas para el Metropolitano.
- Comparar el sistema actual y propuesto del Metropolitano.

1.4 HIPÓTESIS

1.4.1 Hipótesis General

Empleando diversas metodologías y algoritmos para la generación de rutas óptimas, basándose en la demanda de los usuarios, se podrá llevar a la mayor cantidad de pasajeros en un menor tiempo, en un Sistema de Transporte Público Masivo.

1.4.2 Hipótesis Específicos

Las hipótesis específicas son:

- Empleando metodologías de procesamiento de información se podrá automatizar el análisis de los datos.
- Implementando un algoritmo creado para la generación de rutas el tiempo de espera de los usuarios en las estaciones disminuirá en un 30%.
- Usando la capacidad máxima del bus y con una buena programación de las rutas se podrá llevar a la mayor cantidad de usuarios en la hora pico del sistema.

CAPÍTULO II : FUNDAMENTO TEÓRICO

2.1 ANTECEDENTES

Los mejores sistemas de transporte público en el mundo son las ciudades de Hong Kong (China), Seúl (Corea del Sur), Londres (Inglaterra), Múnich (Alemania), Tokio (Japón), Taipéi (Taiwán), París (Francia), entre otras. Los sistemas se basan en una buena planificación del transporte integrado y en una movilidad urbana sostenible que se mejora constantemente.

En la mayoría de los países, por ejemplo: Estados Unidos, Canadá, Australia y Nueva Zelanda; el automóvil es el modo de transporte dominante. En comparación con estas ciudades dominantes de automóviles, las ciudades orientadas al tránsito, por ejemplo: Hong Kong, Curitiba, Tokio y Shanghái; tienen un sistema denso y con una avanzada red de transporte público; ofreciendo servicios de transporte público frecuentes, eficientes y confiables, que también tienen amplias coberturas en espacio y tiempo (Wong et al., 2018). Hong Kong es un ejemplo de ciudades orientadas al tránsito, evidenciado por el hecho de que casi el 90% de las personas usan el transporte público; teniendo una planificación y diseño de servicio que se adapten a todas las personas, es por ello que el gobierno propuso la visión de “Transporte para todos”, el esquema que plantean permite viajar en 4 modos de transporte (ferrocarril, autobús auspiciado, minibús verde y barco), con un tarifario de concesión de HK\$ 2 dólares Hongkonés equivalente a S/ 0.9 soles por viaje aproximadamente.(Yang, 2018)

Por otro lado, en América latina las ciudades de Santiago(chile), Curitiba (Brasil), Bogotá (Colombia), Rio de Janeiro (Brasil), Lima (Perú) estarían por encima del promedio del índice de movilidad urbana lo que indicaría que estas ciudades están en pleno crecimiento para alcanzar un buen sistema de Transporte Público.

En el Perú en los últimos 15 años se realizados esfuerzos para mejorar el sistema de transporte urbano teniendo como uno de los estudios más importantes el Plan JICA, el desarrollo y la construcción de las líneas del metro de lima, y la conformación de la Autoridad de Transporte Urbano (ATU).

La Agencia de cooperación Internacional del Japón (JICA), en el año 2004, realizó una serie de estudios de tránsito, incluyendo encuestas de viaje personal, realizando entrevistas a 38,000 hogares en el área metropolitana de Lima y Callao. Lo cual fue formulado en el 2005 para obtener el Plan Maestro de Transporte Urbano para el Área Metropolitana de Lima y Callao (PMTU – 2025). El PMTU-2025 propuso la futura red de transporte público consistente en 4 líneas de ferrovía y 15 corredores de buses troncales. En el del 2007 se estudió un sistema troncal de ómnibus para el corredor este-oeste a lo largo de la Av. Venezuela y la Av. Nicolás Ayllón. En el 2010 se inició un Sistema BRT (Bus Rapid Transit) con una longitud total de 27 km a lo largo de la Vía Expresa. El 23 de diciembre del 2010, la red básica del tren eléctrico de Lima y Callao de cinco (5) rutas fue formulada por el Ministerio de Transporte y Comunicaciones (MTC) y aprobada por el Decreto Presidencial (D.S. 059-2010-MTC). En enero del 2012, la Autoridad Autónoma del Tren Eléctrico (AATE) inauguró la primera línea del tren eléctrico en Lima. Para la actualización del PMTU-2025 del 2013 consideraron tres modos tales como Sistema de Bus Troncal, Sistema de Transporte de Media Capacidad y Sistema Ferroviario (elevado y subterráneo), mientras que en el plan inicial se basó en dos modos tales como el Sistema de Bus Troncal y el Sistema Ferroviario (elevado y a nivel). (JICA, 2013)

En la ciudad de Lima se han establecido proyectos tales como el sistema de corredores de alta capacidad (Metropolitano), el tren eléctrico y la implementación de los corredores de buses y alimentadores, cuyo sistema de cobro de los dos primeros son con tarjetas y sus rutas son directas, es decir unen distintos distritos en “poco tiempo”. Al igual que estos sistemas de transportes se están desarrollarán diversos sistemas de transporte masivos como las nuevas “líneas del tren de Lima” que serán similares al Metropolitano y la línea 1 del tren eléctrico.

Como se ve actualmente, se han estado implementado sistemas de transporte como los corredores y líneas de metro (1, 2, 3, 4, 5, 6 y 7) y se ha visto que la implementación de este sistema trajo como consecuencia cambios de rutas del transporte. Además, en los últimos meses se dio la implementación de nuevos paraderos, al inicio causó molestias a los usuarios, pero luego personas se han estado adaptando al nuevo sistema implementado por la Municipalidad.

Por último, a fines del 2018 se aprobó la creación de Autoridad de Transporte Urbano (ATU que tendrá como objetivo tener un sistema integrado de transporte de Lima y Callao que permita satisfacer las necesidades de traslado de los pobladores de las provincias de Lima y Callao de manera eficiente, sostenible, accesible, segura, ambientalmente limpia y de amplia cobertura. La ATU estará conformada por un Consejo Directivo: 2 miembros propuestos por el Ministerio de Transportes y Comunicaciones (MTC), 1 de los cuales lo preside, 1 miembro propuesto por el Ministerio de Economía y Finanzas, 1 miembro propuesto por el Ministerio de Vivienda, Construcción y Saneamiento, 4 miembros propuestos por las municipalidades provinciales existentes en el territorio, en proporción al número de habitantes y conforme a los mecanismos de designación que se establezca en el Reglamento de Organización y Funciones. (Congreso de la República, 2018).

Es por ello que tomando en cuenta las propuestas mencionadas y las experiencias de otros países como Colombia, Brasil, Japón, Francia, Inglaterra, Alemania, China etc. donde tratan que su ciudad sea un Smart City se planteará una propuesta donde regulará y ordenará las rutas de los Sistemas de Transporte Público, teniendo como caso de estudio el Metropolitano de la ciudad de Lima.

2.2 MARCO CONCEPTUAL

2.2.1 Optimización

La optimización consiste en la selección de una alternativa mejor, en algún sentido, que las demás alternativas posibles, (...). Resolver un problema de optimización consiste en encontrar el valor que deben tomar las variables para hacer óptima la función objetivo satisfaciendo el conjunto de restricciones. (Ramos, Sánchez, Ferrer, Barquín, y Linares, 2010)

Los métodos de optimización los podemos clasificar en: métodos clásicos (que son los algoritmos que habitualmente se explican en los libros de optimización) y métodos meta heurísticos (que aparecieron ligados a lo que se denominó inteligencia artificial e imitan fenómenos sencillos observados en la naturaleza). Dentro de los primeros se encuentra la optimización lineal, lineal entera mixta, no lineal, estocástica, dinámica, etc. En el segundo grupo se incluyen los algoritmos evolutivos (genéticos entre otros), el método del recocido simulado (Simulated Annealing), las búsquedas heurísticas (método tabú, búsqueda aleatoria, avariciosa, etc.) o los sistemas multi agente. De forma muy general y aproximada se puede decir que los métodos clásicos buscan y garantizan un óptimo local mientras que los métodos meta heurísticos tienen mecanismos específicos para alcanzar un óptimo global aunque no garantizan su alcance. (Ramos, Sánchez, Ferrer, Barquín, y Linares, 2010)

2.2.2 Optimización de Rutas de Transporte de Pasajeros

Las rutas de transporte de pasajeros consisten en la recogida de los clientes en puntos prefijados (paradas) distribuidas geográficamente a lo largo de un territorio y donde todos estos clientes tienen un punto de destino común prefijado. La ruta depende de dos parámetros básicamente: el espacio y el tiempo (Ruiz José, Aguado Andrés, 2012).

La optimización en el transporte se dará cuando se minimice la variable tiempo, distancia o costo del transporte, o cuando se minimiza cualquier combinación de éstas; para ello se puede usar distintos métodos de optimización ya sea por métodos clásicos (algoritmos convencionales) o métodos meta heurísticos (algoritmos de inteligencia artificial).

2.2.3 Transporte rápido masivo

Transporte Rápido Masivo (MRT), pertenece a un tipo de transporte público, este es un servicio de transporte de pasajeros, usualmente de ámbito local, que está disponible para cualquier persona que pague una tarifa prescrita. Generalmente, opera sobre carriles fijos específicos, o con uso separado y exclusivo de pistas comunes potenciales, según horarios establecidos, a lo largo de rutas designadas o líneas con paradas específicas, aunque los buses de transporte rápido y los tranvías, a veces, operan en tráfico mixto. Está diseñado para movilizar grandes números de personas, al mismo tiempo. Los ejemplos incluyen transporte público de tren pesado, transporte público de Tren de Tránsito Ligero (LRT) y Buses de Tránsito Rápido (BRT). (Lloyd, Karl, y Armin, 2006)

En seguida se presentan los tipos de Transporte rápido masivo:

2.2.3.1 *Transporte público de tren pesado*

Es un sistema de transporte público que usa trenes de alto rendimiento, con vagones de trenes potenciados eléctricamente, que operan en derechos de vía exclusivos, normalmente sin pasos a nivel, con estaciones de plataformas altas. (Transit Cooperative Research Program [TCRP], 2003)

2.2.3.2 *Metro*

Metro es el término internacional más común para transporte público de tren pesado subterráneo, aunque también aplicado para tren pesado elevado. En este módulo, usamos "Metro" para referirnos a sistemas de trenes pesados de ciudad separados de nivel. Son la forma de MRT más cara por kilómetro cuadrado, pero tienen la mayor capacidad teórica. (Lloyd, Karl, y Armin, 2006)

2.2.3.3 *Sistemas de trenes para ir al trabajo*

El tren para ir al trabajo o el tren suburbano es la porción de las operaciones de tren de pasajeros que circula dentro o entre las áreas urbanas y sus suburbios, pero difiere de los Metros y de LRT en que los vagones de pasajeros, generalmente, son más pesados, los promedios de viaje largos son generalmente más extensos, y las operaciones son llevadas a cabo en vías que son partes del sistema ferroviario en el área. (Lloyd, Karl, y Armin, 2006)

2.2.3.4 Transporte público de tren ligero

Un sistema de transporte público de tren ligero (LRT, del inglés: Light Rail Transit) es un sistema de tren eléctrico metropolitano, caracterizado por su habilidad para operar vagones únicos o trenes cortos, a lo largo de derechos de vía exclusivos, a nivel del suelo, o de estructuras aéreas, o en subterráneos, u, ocasionalmente, en las calles, y para tomar y descargar pasajeros a nivel de vía, o del suelo de los automóviles (Transit Cooperative Research Program [TCRP], 2003). Los sistemas LRT incluyen tranvías, aunque una diferencia principal es que los tranvías, a menudo, operan sin un derecho exclusivo de vía, en tráfico mixto. (Lloyd, Karl, y Armin, 2006)

2.2.3.5 Sistemas de Buses de Tránsito Rápido

Muchas ciudades han desarrollado variaciones sobre el tema de mejores servicios de buses, y la concepción reside en un conjunto de mejores prácticas, más que en una definición estricta. El BRT es una forma de transporte público orientado al consumidor que combina estaciones, vehículos, planificación, y elementos inteligentes de sistemas de transporte en un sistema integral con una identidad única.

El BRT, típicamente involucra corredores de vías de autobuses en carriles segregadas, ya sea a nivel o desnivel y tecnología de autobús modernizada. Sin embargo, aparte de las vías de autobuses segregadas, los sistemas BRT comúnmente también incluyen:

- Abordaje y descenso rápido
- Una eficiente recolección de tarifas
- Refugios y estaciones cómodas
- Tecnologías limpias de autobús
- Integración modal
- Sofisticada identidad de marketing
- Excelencia en el servicio al consumidor.

El BRT es más que simplemente la operación sobre carriles de autobuses exclusivas o vías de autobús. Según un reciente estudio de vías de autobús a nivel, sólo la mitad de las ciudades que tienen vías de autobús las han desarrollado como parte de un paquete sistemático y global de medidas, como parte de la red de transporte masivo de la ciudad que nosotros identificaríamos como un sistema BRT.

Mientras que los sistemas de BRT siempre incluyen alguna forma de exclusivo derecho de paso para los buses, las aplicaciones que nosotros consideramos, en su mayoría son: vías de autobús al nivel de la calle. Las vías de autobús elevadas o los túneles pueden necesitarse para atravesar algunos centros de ciudad, pero en muchas ciudades en desarrollo los fondos no van a estar disponibles para separaciones de nivel de envergadura. (Lloyd, Karl, y Armin, 2006)

La característica de un Transporte Rápido Masivo, son las siguientes:

Uso del espacio; similares consideraciones de un espacio eficiente se aplican a todos los modos MRT como en la Figura 4 (se muestra la cantidad de espacio requerido para transportar el mismo número de pasajeros: automóvil, bicicleta y autobús), aunque en la práctica surge como un asunto de políticas sólo respecto de los buses y algunas versiones de LRT, ya que los sistemas de trenes están ya totalmente segregados de otro tráfico. Los BRT y LRT, a menudo, involucran la redistribución del espacio de vías ya existentes a favor de modos más eficientes, mientras que los Metros están, por lo general, totalmente separados de nivel y no tienen impactos sobre la capacidad de vías, a menos que ellos sean elevados, en cuyo caso puede haber una pequeña reducción en la capacidad de vías. (Lloyd, Karl, y Armin, 2006)

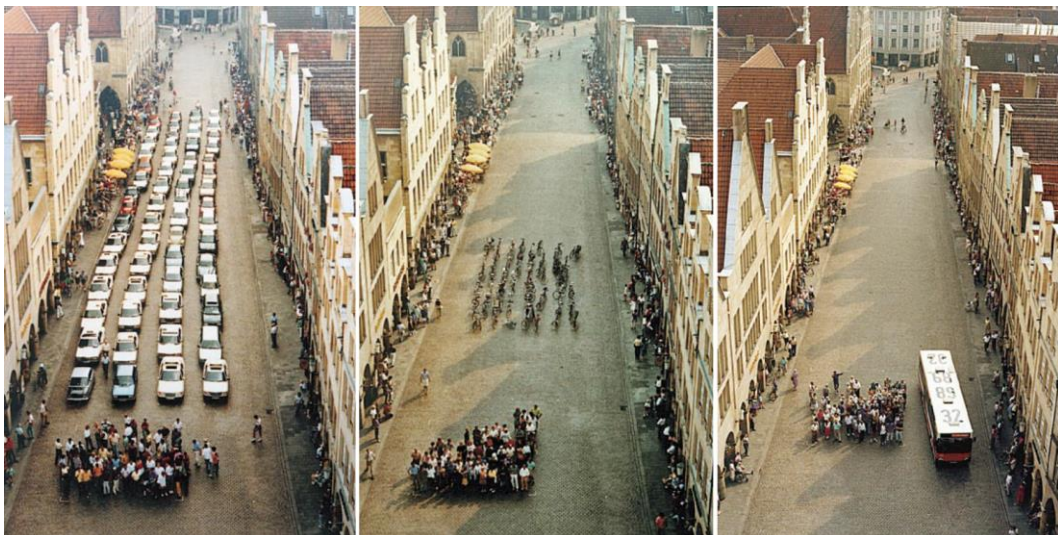


Figura 4 Densidad de tráfico según el modo de transporte.

Fuente: Opciones de transporte público masivo. Afiche desplegado en la Oficina de Planificación de la Ciudad de Münster, agosto de 2001. (Lloyd, Karl, y Armin, 2006)

Velocidad y capacidad de pasajeros; todas las formas de MRT operan con velocidades y capacidades de pasajeros relativamente altas, y el requerimiento básico de MRT en la ciudad en desarrollo es que éste transporte grandes cantidades de pasajeros rápidamente. Donde se apliquen Metros en Ciudades en Desarrollo, casi siempre, ellos son, de manera absoluta, el modo más rápido de MRT con una velocidad promedio de 40 a 50 km/h, mientras que LRT y los sistemas de BRT operan típicamente a velocidades promedio de entre 20 y 30 km/h. (Lloyd, Karl, y Armin, 2006)

Integración; todos los sistemas MRT requieren intercambios con otros elementos del sistema de transporte público, y también integración con otros modos de transporte, como lo son los automóviles, personas a pie y en bicicleta. Shanghai, por ejemplo, proporciona excelentes intercambios Metro/bicicleta y Metro/peatón y buenos intercambios Metro/autobús en algunas estaciones principales. El Metro de la Ciudad de México está físicamente integrado con el aeropuerto internacional y las estaciones principales de autobús. El sistema BRT de Curitiba incluye excelente integración con calles peatonizadas y paradas de taxis. El BRT de Sao Paulo se integra bien con el sistema de Metro. La escasa integración es una característica de algunos sistemas MRT basados en trenes que están rindiendo por debajo de su nivel, como es el caso de Kuala Lumpur y Manila. (Lloyd, Karl, y Armin, 2006)

Nivel de servicio; los sistemas MRT usualmente ofrecen un nivel de servicio superior en comparación con modos basados en caminos no segregados como es el caso de autobuses comunes, taxis, y el transporte público paralelo, o para transporte público.

El servicio superior es evidente, por ejemplo, con:

- Terminales e intercambios.
- Limpieza.
- Una sofisticada imagen de marketing.
- Información para los pasajeros.
- Controles climatizados.
- Integración modal.
- Integración con captadores principales de viaje.

Históricamente los sistemas basados en trenes han funcionado mejor en cuanto a indicadores de “nivel de servicio”, aunque los éxitos recientes en BRT están desafiando esas concepciones tradicionales. (Lloyd, Karl, y Armin, 2006)

2.2.4 Análisis del sistema de transporte

El Análisis del Sistema de Transporte (TSA, Transportation Systems Analysis) es la interacción de macro variables que son, sistema de transporte, sistema de actividades y la estructura de flujo.

Sistema de transporte (T); son los elementos de infraestructura y servicio (equipo, operación y gobernanza) de transporte; por ejemplo: vías, tipo de vehículo, equipos para operar un sistema, reglas de operación, vehículos en un determinado horario, facilidades para un vehículo de carga, horarios de estacionamiento. El T, están representados a través de gráficos de red definidos por enlaces y nodos. (1) Enlaces, secciones homogéneas de un solo sentido de la infraestructura o servicio de transporte. (2) Nodos, puntos finales de enlace, típicamente intersecciones o puntos que representan cambios en los atributos enlace. Tanto los enlaces como los nodos tienen atributos asociados como la longitud, la velocidad y la capacidad de los enlaces y las prohibiciones y penalizaciones para los nodos.(McNally, 2008)

El sistema de actividades (A); son las distribuciones espaciales, la actividad demográfica y / o económica que se produce en los usos de la tierra; por ejemplo, dónde se emplaza la población de una ciudad, las características de las actividades en una región, lugar de actividades (centros comerciales, escuela, cines, etc.), niveles de ingreso según el modo de transporte, lugar fuentes productivas de la ciudad, lugares de acopio, etc. El A, están representados por los datos socioeconómicos, demográficos u otras actividades convenientes para la Zona de Análisis de Tráfico (TAZ).(McNally, 2008)

En la estructura de flujo (F); es el resultado de la interacción entre el sistema de transporte y el sistema de actividades, lo que nos interesa entender es cuál es el origen de un viaje, cuál es el destino de un viaje, cuál es el camino, cuál es la ruta que toma ese viaje, cuál es la cantidad de viajes que vamos a ver en cada uno de los elementos del sistema de transporte y cuál es la posición. Por ejemplo, en una calle, cuántos vehículos y camiones transitan por hora, cuál es la tasa de ocupación de los vehículos, etc. El sistema de actividad (A) está interconectado con el sistema de transporte (T) a través de conectores de centroide que son enlaces abstractos que conectan los centroides de TAZ con puntos de acceso realistas en la red física (generalmente, a mitad de bloque y no en nodos).(McNally, 2008)

A mismo existe una relación de las macro variables cuando hablamos del sistema de transporte, infraestructura, equipos y operación, sistema de actividades y la estructura de flujo. En la Figura 5 se observa un diagrama donde están la relación entre cada macro variable. La relación tipo 1 indica que F depende diariamente tanto de T como A; la relación tipo 2 en el largo plazo F afecta A, a través del nivel de servicio y el consumo de recursos; y la relación tipo 3 a mediano y largo plazo F provoca cambios en T, a partir de los niveles de servicio y flujos. (Giesen, Delgado, & De Dios Ortuzar, 2019)

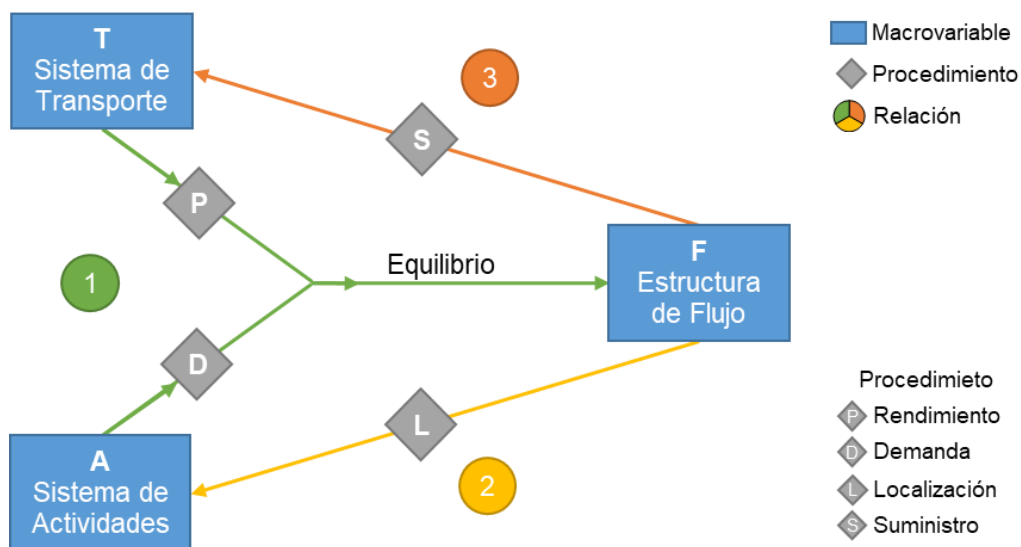


Figura 5 Red de Análisis del sistema de transporte

Fuente: *Elaboración y traducción propia en base a (Manheim, 1979), (Florian, Gaudry, & Lardinois, 1988), (McNally, 2008) y (Giesen et al., 2019)*

El T, A; sirven como insumos exógenos a los procedimientos de rendimiento (P) y procedimientos de demanda (D), respectivamente. Son tales P y D los que conforman el Modelo de Cuatro Etapas (FSM) básico. Si bien se requiere algún tipo de procedimiento de localización (L), típicamente se ha ejecutado de manera independiente del FSM y rara vez se ha integrado de alguna manera formal dentro del procedimiento de equilibrio básico. Del mismo modo, los procedimientos de suministro formal (S) son prácticamente inexistentes. Florian caracteriza el análisis formal como la participación de la elección de la perspectiva del análisis y la definición de los procedimientos, y por lo tanto las variables, se deben especificarse de forma endógena o exógena. La importancia crítica para este enfoque es la comprensión de las unidades de análisis para estos procedimientos, definidas espacial y temporalmente. (Florian et al., 1988)

2.2.5 Modelo de las cuatro etapas (FSM)

El modelo del transporte se desarrolló en los EE. UU, durante la era del desarrollo de la posguerra y el crecimiento económico. La aplicación se dio en la década del 50 en el Estudio de Transporte del Área de Chicago con el Modelo de las Cuatro Etapas (FSM, The Four Step Model). En el año de 1960 se institucionalizó el FSM y a partir de ello hubo mejoras a través del “Programa de Mejora del Modelo” de viajes y la mejora del modelo convencional con investigaciones. Al inicio el FSM se basó en la demanda de los viajes, durante el pasar de los años, el modelo mejoró hasta convertirse en el enfoque basado en actividades. (McNally, 2008)

El modelo de cuatro pasos es la herramienta principal para pronosticar la demanda futura y el rendimiento de un sistema de transporte; generalmente definido en una escala regional o subregional (las escalas más pequeñas a menudo aplican modelos simplificados). Para las redes elementales, las funciones de demanda directa (puede estimarse), funciones de rendimiento de enlace estándar y la enumeración de rutas, pueden proporcionar los flujos deseados. Para cualquier aplicación regional realista, se requiere modelo alternativo debido a la complejidad de la red. El FSM se desarrolló para enfrentar esta complejidad al formular el proceso como un modelo secuencial de cuatro etapas (Figura 6). La aplicación de modelos de previsión de viajes es un proceso continuo. El período requerido para la recopilación de datos, la estimación del modelo y los ejercicios subsecuentes de pronóstico pueden tomar años, durante los cuales las actividades y los sistemas de transporte cambian al igual que las políticas de interés, que a menudo requieren nuevos esfuerzos de recopilación de datos y un nuevo esfuerzo de modelado. (McNally, 2008)

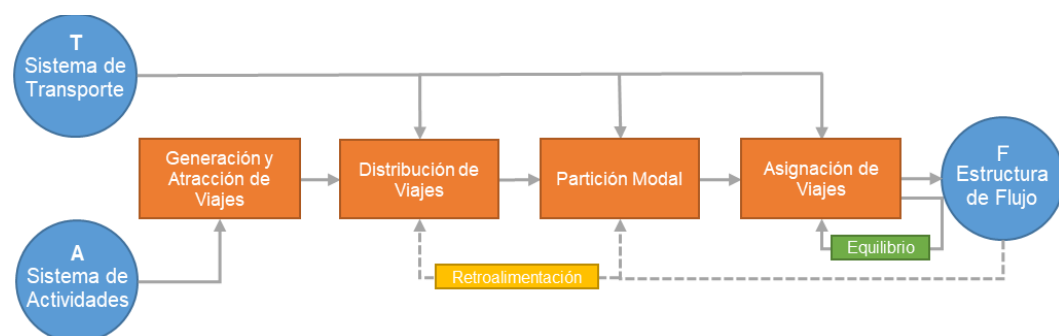


Figura 6 TSA y FSM

Fuente: Elaboración y traducción propia en base a (McNally, 2008)

En el Análisis del Sistema de Transporte (TSA) se discute el problema, la definición del área de estudio, se aplica los requisitos de los datos y el modelo; es por ello que el FSM es una aplicación del TSA, a su vez el FSM proporciona un mecanismo para determinar los flujos de equilibrio en el TSA como se ilustra en la Figura 5. Para un mejor entendimiento, se explicará cómo define el área de estudio y como se recolectan los datos en un TSA.

Área de estudio; una línea de cordón define el área de estudio, el área dentro del cordón se compone de zonas de análisis del tráfico (TAZ) está sujeta al modelado y análisis. La interacción con áreas fuera del cordón se define a través de estaciones externas (ESS) que sirven efectivamente como puertas para los viajes en, fuera de, y a través de la zona de estudio.

Datos; son recolectados por encuestas, estos datos se utilizan para validar la representatividad de la muestra; así desarrollar y estimar la generación de viajes, distribuir los viajes y la elección de modos, para realizar estudios de movimiento en el tiempo. El FSM tiene datos importantes de la demanda, para definir los sistemas de actividades (A) y transporte (T). La encuesta de hogares está conformada por tres tipos de datos. (1) Datos socioeconómicos a nivel de hogar y persona, por lo general, incluyen los ingresos y el número de miembros del hogar, trabajadores y automóviles. (2) Datos de viaje de actividad, generalmente incluyen para cada actividad realizada en un período de 24 horas como tipo de actividad, la ubicación, hora de inicio, duración, existencia de viaje, el modo de transporte, la hora de salida y hora de llegada. (3) Datos del vehículo del hogar. (McNally, 2008)

Las etapas del modelo de FSM son la generación y atracción de viajes, distribución de viajes, partición modal y asignación de viajes. (ver Figura 7)

2.2.5.1 Generación y atracción de viajes

Es contabilizar la cantidad de viajes atraídos o generados de una respectiva zona; denominaremos a O_i viajes que se originan en "i" y a D_j viajes que son atraídos a la zona "j". Estos dependen de las características socio económicas y de uso de suelo. Por ejemplo: en hogares con un ingreso adquisitivo más alto deberían generar más viajes que hogares más pobres. Hogares con tenencia de automóvil deberían, también, generar más viajes que aquellos hogares que no lo tienen. Si tenemos un uso de suelo de tipo comercial, uno debería esperar que a esas zonas llegara mucha gente atraída para hacer labores de compra. (Giesen et al., 2019)

2.2.5.2 Distribución de viajes

Es contabilizar la cantidad de viajes que se originan entre dos pares de zonas específicas; denominaremos a V_{ij} viajes entre las zonas “i” y zona “j”. Estos dependerán del nivel de servicio que hay entre las zonas. Por ejemplo: si se tiene dos zonas muy alejadas o separadas por un río, la cantidad de viajes entre las zonas deberá ser menor si se tuviera otras dos zonas mejor conectas. Por lo general para realizar la distribución de los viajes se usa el Modelo de Gravedad.(Giesen et al., 2019)

2.2.5.3 Partición modal

Es contabilizar la cantidad de viajes realizados según el modo ya sea caminata, bicicleta, auto, bus, BRT, metro, etc.; denominaremos a V_{ij}^m viajes entre la zona “i” y zona “j” usando el modo “m”. Estos dependerán de las características de los modos. Por ejemplo, influirá en el modo a usar la tarifa, el tiempo de viaje, la comodidad, etc. Por lo general para la partición modal se usa el modelo Logit.(Giesen et al., 2019)

2.2.5.4 Asignación de viajes

Es contabilizar la cantidad de viajes realizados por una determinada ruta; denominaremos a V_{ij}^{mr} viajes entre la zona “i” y la zona “j” usando el modo “m” que elegirá una ruta “r”. Estos dependerán de las características de las redes de transporte. Por ejemplo, si en la red existe congestión, características de la calle, etc. Por lo general la asignación de las rutas se realiza con el algoritmo de Frank-Wolfe.(Giesen et al., 2019)



Figura 7 Modelo de las Cuatro Etapas
Fuente: Elaboración propia en base a (Giesen et al., 2019).

2.2.6 Data Science

La Ciencia de Datos (Data Science) es una disciplina que recientemente ha surgido, es por ello que se mencionará las definiciones más relevantes:

La Ciencia de Datos, es un campo interdisciplinario sobre procesos y sistemas para extraer conocimiento o ideas de grandes volúmenes de datos en varias formas (datos estructurado o no estructurado) y traducirlo en acción (Liu, 2015). El procesamiento e interpretación de los datos, se refiere a responder preguntas del mundo real (Hicks & Irizarry, 2018). La Ciencia de Datos abarca un conjunto de principios, definiciones de problemas, algoritmos y procesos para extraer patrones no obvios y útiles de grandes conjuntos de datos (Kelleher & Tierney, 2018). En la práctica, la Ciencia de Datos ha evolucionado como un campo interdisciplinario que integra enfoques de análisis de datos como estadísticas, minería de datos (data mining) y el análisis predictivo, e incorpora avances en computación escalable y gestión de datos (Berman et al., 2018). Conceptos como la inferencia, el modelado y la visualización de datos, son una parte integral de la caja de herramientas del científico de datos moderno (Hicks & Irizarry, 2018).

La Ciencia de Datos está estrechamente relacionado con los campos de la minería de datos y el aprendizaje automático (Machine Learning), pero con un alcance más amplio. El crecimiento en el uso de la Ciencia de Datos en nuestras sociedades está impulsado por la aparición de Big Data y las redes sociales, la aceleración de la capacidad informática, la reducción masiva en el costo de la memoria de la computadora y el desarrollo de métodos más potentes para el análisis y modelado de datos, como el aprendizaje profundo (Deep Learning). Juntos, estos factores significan que nunca ha sido más fácil para las organizaciones recopilar, almacenar y procesar datos. Sin embargo, existe algunas formas en que la Ciencia de Datos puede afectar su vida diaria, es decir, como cuando incluya determinar que anuncios se le presentan en línea; qué películas, libros y recomendaciones de amigos; qué correos electrónicos se filtran en su carpeta de spam; qué ofertas recibe cuando renueva su servicio de telefonía celular; el costo de la prima de seguro de salud; la secuencia y la sincronización de los semáforos en su área; cómo se diseñaron los medicamentos que pueda necesitar; y qué lugares de tu ciudad están dirigidos la policía. (Kelleher & Tierney, 2018)

De lo mencionado anteriormente se infiere que la Ciencia de Datos es un campo interdisciplinario que integra campos de la estadística (análisis) y matemática (algoritmos), ciencia de la computación (procesamiento) y dominio del conocimiento del tema a tratar (experiencia) como de muestra en la Figura 8. El conjunto de habilidades para un científico de datos son: tener experiencia en el campo, ética de los datos y regulación, gestión de datos y base de datos, Big Data y HPC, visualización de datos, estadística y probabilidades, aprendizaje automático (ML) y comunicación. (Kelleher & Tierney, 2018)

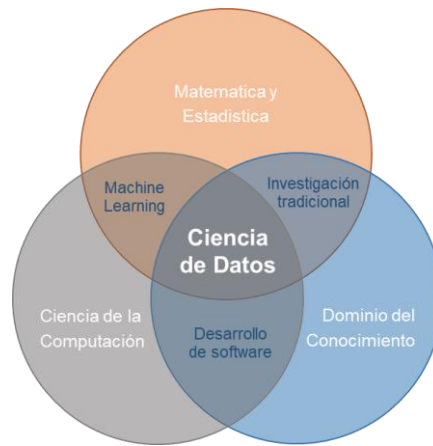


Figura 8 Campo interdisciplinario
Fuente: Elaboración y traducción propia en base a (Prieto Espinosa, 2018).

Las etapas de un proyecto de Ciencia de Datos son: comprensión del negocio (campo en que se esta desarrollando), comprensión de datos, preparación de datos, modelado, evaluación y despliegue. (Kelleher & Tierney, 2018)

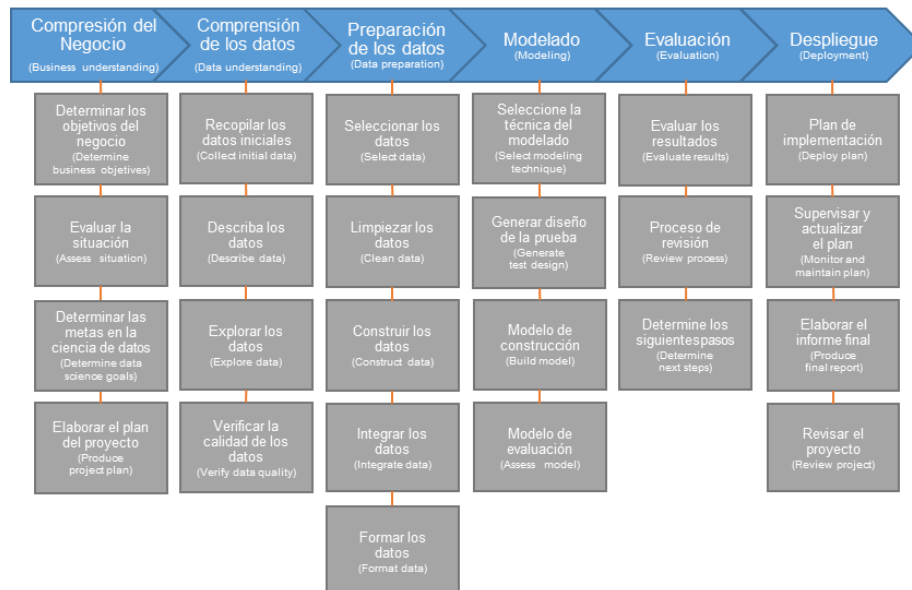


Figura 9 Etapas y tareas de un proyecto de Ciencia de Datos (CRISP-DM)
Fuente: Traducción propia en base a (Kelleher & Tierney, 2018).

2.2.7 Big Data

Zikopoulos y Eaton (2011) definen Big Data como una colección de grandes conjuntos de datos que pueden no ser procesados usando herramientas de administración de bases de datos tradicionales. Manyika et al. (2011) define Big Data como "conjuntos de datos cuyo tamaño está más allá de la capacidad de las herramientas típicas de software de base de datos para capturar, almacenar, administrar y analizar" (Chong y Shi, 2015). El Big Data a menudo se define en términos de las tres Vs: el volumen extremo de datos, la variedad de tipos de datos y la velocidad a la que se deben procesar los datos (Kelleher & Tierney, 2018). En comparación con los enfoques tradicionales, los análisis de datos de gran tamaño abordan grandes cantidades de datos heterogéneos y no estructurados. (Chong & Shi, 2015)

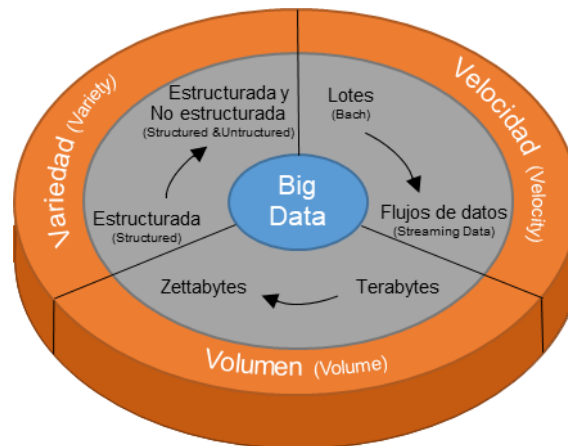


Figura 10 IBM, características del Big Data 3Vs

Fuente: Traducción propia en base a (Paul C. Zikopoulos, Chris Eaton, Dirk deRoos, Thomas Deutsch, & George Lapis, 2011).

Existen tecnologías de código abierto que pertenecen a la organización Apache, como Hadoop, Storm, Spark y Flink, los cuales son plataformas que sirven para adquirir y almacenar grandes volúmenes de datos de forma eficiente y mucho menos costosa comparada con una plataforma tradicional. Hadoop, está diseñado principalmente para el procesamiento por lotes de datos (batch, para datos estáticos, procesamiento no se requieren inmediatamente). Storm, diseñado para procesar flujos de datos (streams, procesa a medida que ingresa al sistema, es decir en tiempo real). Spark es híbrido (lotes y flujos), diseñado por procesamiento por lotes con algunas capacidades de procesamiento de flujos. Flink es híbrido (lotes y flujos), diseñado para procesamiento por flujos también se puede usar para procesamiento por lotes. (Kelleher & Tierney, 2018)

A continuación, presentaremos cuatro procesos de la arquitectura del Big Data

2.2.7.1 Adquirir el Big Data

La adquisición de datos consiste en la recopilación, transmisión y preprocesamiento de datos (X. Chen & Lin, 2014). Existen muchos tipos de métodos de recopilación de datos, como la recopilación de datos basada en sensores, la minería de datos en la Web (correos electrónicos, aplicaciones móviles y medios sociales), etc. Debido al rápido desarrollo de la tecnología de sensores, como Internet de las Cosas (IoT) y la tecnología de identificación por radiofrecuencia (RFID). (Chong & Shi, 2015)

2.2.7.2 Almacenar el Big Data

Infraestructura; es el hardware que se necesita para la construcción del sistema de almacenamiento del Big Data.

Gestión del almacenamiento de los datos; también es parte del procesamiento, contiene tres niveles principales: sistemas de archivos, tecnología de bases de datos y modelos de programación. (Chong & Shi, 2015)

2.2.7.3 Procesar el Big Data

Identifican cuatro requisitos clave del procesamiento de grandes datos: carga rápida de datos, procesamiento rápido de consultas, utilización altamente eficiente del espacio de almacenamiento y gran adaptabilidad a patrones dinámicos de carga de trabajo. (Chong & Shi, 2015)

La mayoría de estas arquitecturas se pueden clasificar en tres tipos de modelos: **Modelos de procesamiento MapReduce y MapReduce-relacionados;** son modelo de programación para la computación a gran escala que utiliza un gran número de clusters para lograr el procesamiento automático en paralelo y distribuido. (Chong & Shi, 2015)

Modelo de procesamiento de grafos; gran parte del análisis del Big data han adoptado modelos de procesamiento de grafos para explorar la relación de patrones - reglas entre objetos interconectados. (Chong & Shi, 2015)

Modelo de procesamiento de flujos; aplicaciones como: eventos, monitoreo de operaciones y los sistemas de análisis en tiempo real, requieren que el sistema tenga la capacidad de procesar continua y eficientemente los volúmenes de los flujos del Big Data en tiempo real. (Chong & Shi, 2015)

2.2.7.4 Analizar el Big Data

Es extraer información de múltiples fuentes para su uso posterior, como la toma de decisiones y la predicción. Existen tres tipos de análisis: (1) Descriptivo, utiliza datos históricos para resumir lo sucedido. (2) Predictivo, utiliza varias técnicas para predecir tendencias futuras. (3) Prescriptivo, se centra en la recomendación y la asistencia a la toma de decisiones, es la unión de las dos anteriores. (Chong & Shi, 2015)

Por otro lado el sistema Benchmark del Big Data, sirve para medir los rendimientos, se pueden clasificar en (1) Componentes, evalúa el rendimiento de las componentes del sistema (Hadoop, TeraSort, etc.) (2) Sistemas, evalúa el rendimiento del sistema completo (TPC-E , BigBench, etc.) (Chong & Shi, 2015)

A continuación, se muestra el proceso operativo, la secuencia del flujo de datos y el ecosistema del lago de datos del Big Data basado en Hadoop en la Figura 11. (Wu, Chen, Wu, & Lytras, 2018)

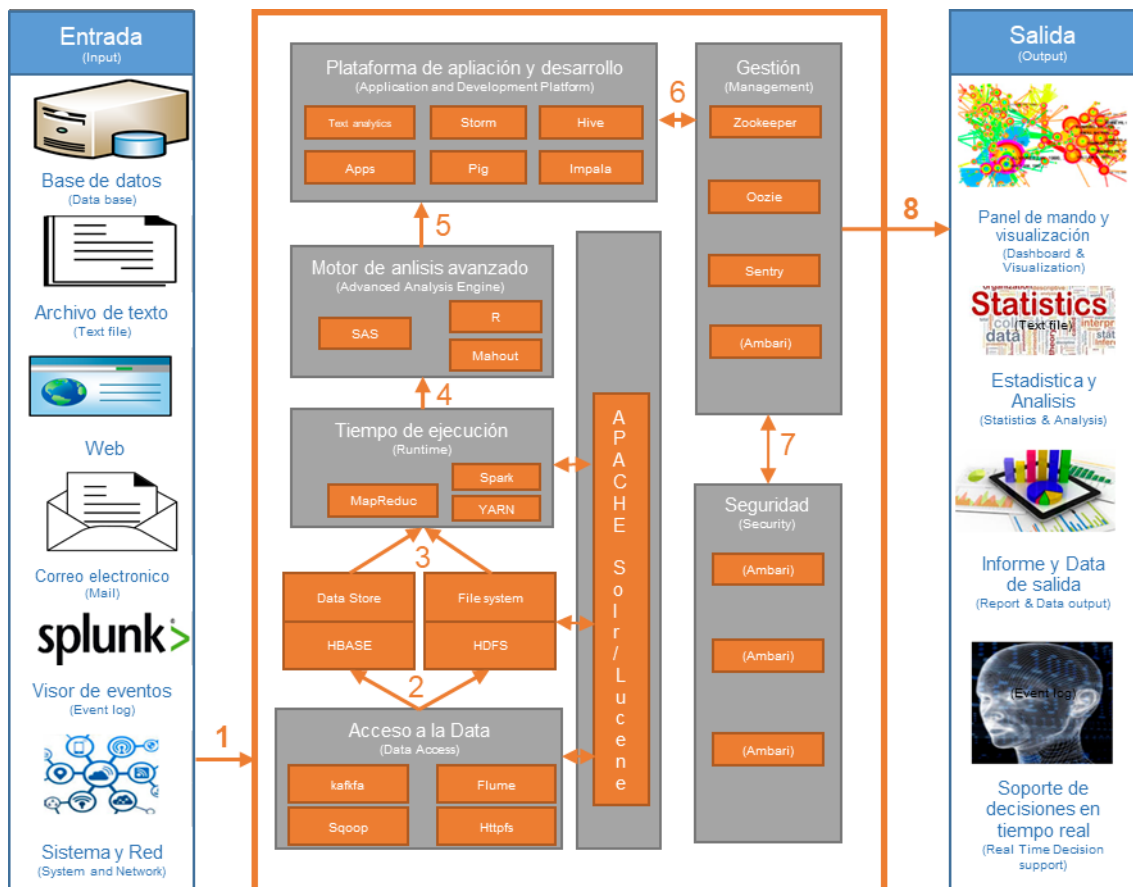


Figura 11 Modelo de operación del Big Data.
Fuente : Traducción propia en base a (Wu et al., 2018).

2.2.8 Machine Learning

Machine Learning (ML, aprendizaje automático, aprendizaje maquina); es el subcampo de la Ciencia de la Computación que se ha desarrollado a partir de los campos del reconocimiento de patrones, la inteligencia artificial y la teoría del aprendizaje computacional. El propósito de esta disciplina es el desarrollo de programas, que son capaces de aprender de los datos previamente vistos, a través de parámetros sintonizables (normalmente matrices de valores de doble precisión), que están diseñados para ser ajustados automáticamente para mejorar las predicciones resultantes. De esta manera, los ordenadores pueden predecir un comportamiento, generalizando la estructura subyacente de los datos, en lugar de simplemente almacenar (o recuperar) los valores como los sistemas de base de datos habituales. Por esta razón, el aprendizaje automático se asocia con la estática computacional, que también intenta predecir un comportamiento basado en datos anteriores. (Isoni & Cervellin, 2016)

El aprendizaje automático principalmente comprende el descubrimiento de modelos, patrones y regularidades en los datos. El ML puede ser visto desde dos enfoques, los simbólicos y los estadísticos. Los primeros trabajan aprendizaje inductivo de descripciones simbólicas, mientras que los segundos se centran en los métodos de reconocimiento de patrones o en la estadística. En los últimos años, el uso del Machine Learning (aprendizaje máquina) se ha extendido con rapidez, se ven aplicaciones en dominios como detección de fraudes, sistemas de recomendación, detección de spam, predicciones financieras, comercio y mercadeo, los motores de búsqueda, el reconocimiento óptico de caracteres y la visión por ordenador, entre otros. (Hernández-Leal et al., 2017)

Cualquier problema de aprendizaje comienza con un conjunto de datos de n muestras, que se utilizan para predecir las propiedades de los futuros datos desconocidos. Cada muestra se compone típicamente de más de un valor, por lo que es un vector. Los componentes de este vector se llaman características. Por ejemplo, imagine predecir el precio de un coche de segunda mano basado en sus características: año de fabricación, color, tamaño del motor, etc. Cada coche i en el conjunto de datos será un vector de las características $x(i)$ que corresponde a su color, tamaño del motor, y muchos otros. En este caso, también hay una variable objetivo (o etiqueta) asociada a cada coche i , $y(i)$ que

es el precio del coche de segunda mano. Un ejemplo de entrenamiento está formado por un par $(x(i), y(i))$ y por lo tanto el conjunto completo de N puntos de datos utilizados para aprender se llama conjunto de datos de entrenamiento $\{(x(i), y(i)); i=1, \dots, N\}$. El símbolo x indicará el espacio de los valores de la característica (entrada), y el espacio de los valores objetivo (salida). El algoritmo de ML elegido para resolver el problema se describirá mediante un modelo matemático, con algunos parámetros para sintonizar en el conjunto de entrenamiento. Una vez finalizada la fase de formación, el rendimiento de la predicción se evalúa utilizando otros dos conjuntos: los conjuntos de validación y los conjuntos de pruebas. El conjunto de validación se utiliza para elegir, entre varios modelos, el que arroja los mejores resultados, mientras que el conjunto de pruebas se utiliza normalmente para determinar la precisión real del modelo elegido. Normalmente, el conjunto de datos se divide en un 50% de conjunto de formación, un 25% de conjunto de validación y un 25% de conjunto de pruebas. (Isoni & Cervellin, 2016)

Los problemas de aprendizaje se pueden dividir principalmente en dos categorías: (1) Aprendizaje no supervisado, el objetivo habitual es encontrar ejemplos similares dentro de los datos, utilizando generalmente algoritmos de agrupamiento. (2) Aprendizaje supervisado, la tarea consiste en inferir los parámetros para predecir los valores objetivo de los datos de prueba utilizando generalmente algoritmos de clasificación y regresión. (Isoni & Cervellin, 2016)

El siguiente mapa conceptual representa los aspectos clave y la semántica del aprendizaje automático (Gollapudi, 2016)

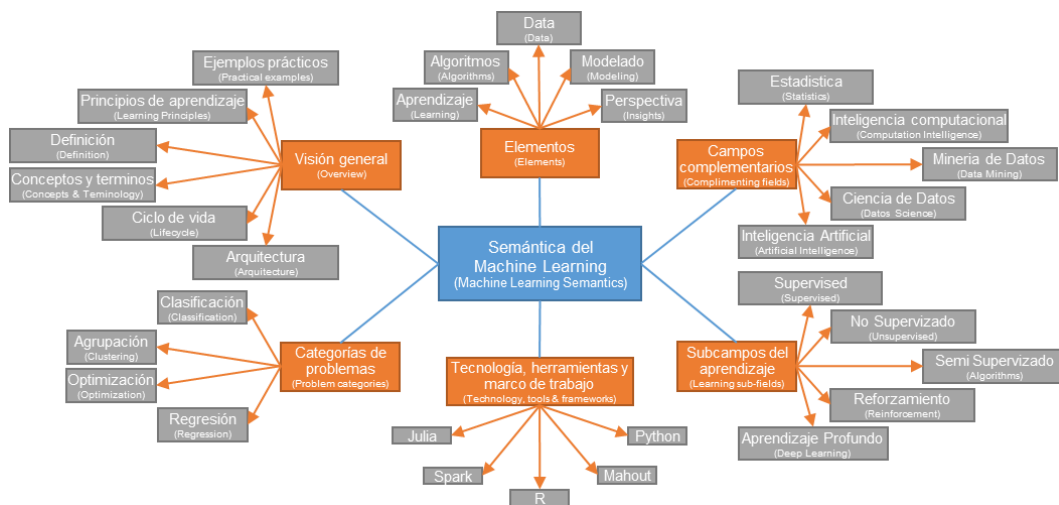


Figura 12 Mapa conceptual de los aspectos claves del ML
Fuente: Traducción propia en base a (Gollapudi, 2016).

2.2.9 Algoritmo

Un algoritmo es un conjunto finito de reglas bien definidas en su lógica de control que permiten la solución de un problema en una cantidad finita de tiempo (Mancilla Herrera, Ebratt Gómez, & Capacho Portilla, 2014). Dada una entrada de un conjunto de entradas posibles, nos permite obtener una salida, requerimos que la salida de una entrada legal sea única, es decir, si hay una salida para una entrada en particular, entonces decimos que el algoritmo se puede aplicar a esta entrada y procesarla para dar la salida correspondiente (Alsuwaiyel, 2016). En la resolución del problema con las reglas mencionadas, el algoritmo realiza un conjunto de pasos cuya ejecución para dar solución del problema puede ser ejecutada manualmente, mecánicamente o utilizando una máquina de procesamiento electrónico de datos (Mancilla Herrera et al., 2014).

Al estar conformado por un conjunto finito de reglas, el algoritmo tiene:

1. Una(s) regla(s) de entrada(s) al algoritmo o punto inicial del algoritmo.
2. Una o un conjunto de reglas intermedias (reglas de proceso o de cálculos).
3. Un conjunto o una regla de finalización del algoritmo, que aseguran su terminación, punto en el cual se obtiene una respuesta útil derivada de la ejecución del conjunto de reglas del algoritmo.
4. Exactitud, en el proceso lógico de organización de las reglas a fin de dar respuesta al usuario, o lo que es equivalente.

2.2.10 Seudocódigo

Es un pseudo lenguaje intermedio entre el lenguaje natural y el lenguaje de programación seleccionado, considerándose por tanto un lenguaje de pseudo programación. Un conjunto de sintaxis, lenguaje natural, símbolos, términos y otras características de programación de alto nivel; para describir a un algoritmo. Permite al diseñador centrarse en la lógica del algoritmo sin distraerse con los detalles de la sintaxis del lenguaje en el que será programado. El pseudocódigo debe describir lo que se está haciendo en el diseño, no cómo se va a implementar exactamente. Es independiente del lenguaje de programación en el que se va a utilizar y fácil de expresar el pseudocódigo en cualquier lenguaje de programación. (Doersing, 2016)

2.3 INFORMACIÓN REFERENTE AL CASO DE ESTUDIO

2.3.1 Transporte urbano en Lima

2.3.1.1 Autoridad Autónoma del Servicio Eléctrico de Transporte Masivo de Lima (AATE)

Es la encargada de liderar la consolidación de la red básica del Metro de Lima y Callao. Se encuentra adscrita al Vice Ministerio de Transportes del Ministerio de Transportes y Comunicaciones. La AATE tiene las competencias sobre Red Básica del Metro de Lima - Sistema Eléctrico de Transporte Masivo de Lima que incluye seis líneas, de las que la Línea 1 está en operación, la Línea 2 dispone de contrato de concesión (construcción y operación), las líneas 3 y 4 se encuentran en fase de preinversión. (GIZ, 2016)



Figura 13 Red básica del Metro de Lima y el Callao
Fuente: Página web del Metro de Lima y el Callao

2.3.1.2 Gerencia de Transporte urbano (GTU)

Es la entidad encargada de planificar, regular y gestionar el tránsito urbano de pasajeros, otorgando las concesiones, autorizaciones y permisos de operación para la prestación de las distintas modalidades de servicios públicos de transporte de pasajeros. Siendo responsable del servicio de transporte regular. Está bajo jurisdicción de la Municipalidad Metropolitana de Lima. (GIZ, 2016)

2.3.1.3 Instituto Metropolitano Protransporte de Lima

Es un organismo público descentralizado de la Municipalidad Metropolitana de Lima (MML), que tiene a su cargo el sistema de Corredores Segregados de Buses de Alta Capacidad (COSAC), dentro de los cuales se encuentra el Metropolitano (COSAC 1) y los Corredores Complementarios. Esta bajo jurisdicción de la Municipalidad Metropolitana de Lima. (GIZ, 2016)

2.3.2 Metropolitano

El Metropolitano es el primer transporte público, urbano y masivo de Lima que atiende alrededor de 700 mil viajes diarios que logran conectar la ciudad de Norte a Sur uniendo 18 distritos, que se extenderá hasta el distrito de Carabayllo. El sistema está compuesto por Buses de Tránsito Rápido (BRT) de alta capacidad, corredores exclusivos y estaciones; también armoniza rutas, tipo de servicio, frecuencia de viajes y programación horaria; cuyo un sistema de pago es con tarjetas. Cuenta con 35 estaciones intermedias, 1 estación central subterránea y 2 terminales de transferencias. (Metropolitano, 2019)

2.3.2.1 Rutas

El actual sistema de Metropolitano está compuesto por una ruta troncal y dos rutas alimentadoras.

Troncal; es una vía exclusiva que conecta 12 distritos en la ciudad de Lima, desde el distrito de Comas - zona Norte hasta el distrito de Chorrillos - zona Sur con un recorrido de 27km y que será ampliado 10.2km hacia el Norte. Cuenta con 4 servicios regulares (efectúan paradas por todas las estaciones) y 11 servicios expresos (efectúan paradas por algunas estaciones). En la troncal circulan 300 buses de longitud entre 18 a 19 metros de largo, con capacidad de trasladar 160 pasajeros. (Metropolitano, 2019)

Alimentador; son vías no exclusivas que tienen por finalidad acercar a los usuarios del Metropolitano hacia los terminales para realizar la conexión con la ruta troncal. Cuenta con 2 rutas alimentadoras frecuentes (17 servicios en el alimentador norte, 4 servicios en el alimentador sur) y una especial (3 servicios en el alimentador especial). (Metropolitano, 2019)



Figura 14 Ruta de circulación del Metropolitano
Fuente: Pagina web de INVERMET

2.3.2.2 Operadores

Los operadores tienen acceso y derecho a la utilización de la infraestructura del sistema para la explotación del servicio y el cobro de la contraprestación.

Operación del servicio, la operación del servicio del Metropolitano está concesionada a 3 empresas, las cuales son: Lima Bus Internacional S.A., Lima Vía Express, Consorcio Perú Masivo y Transvial Lima S.A.C.; cuyo contrato de concesión es de 12 años. (GIZ, 2016)

Unidad de recaudo, el sistema de recaudación se da a través de las tarjetas del Metropolitano, los usuarios ingresan al Metropolitano pasando las tarjetas RFID en un validador y girando un torniquete, el validador captura algunos datos de los usuarios. La unidad de recaudo está concesionada a la empresa ACS Solutions Perú S.A.; cuyo contrato de concesión es de 14 años (GIZ, 2016).

Gasocentro, la concesión para los gasocentros es la empresa PGN GASUR S.A.C. cuyo contrato de concesión es de 12 años. (GIZ, 2016)

2.3.2.3 Tipos de perfil de usuarios

Perfil General, tiene el tarifario general de S/2.5, puede ser adquirida por cualquier ciudadano peruano o extranjero; la tarjeta puede tener datos personales si el usuario lo desea, con el beneficio de ser bloqueada por pérdida o robo, transfiriendo el saldo a una nueva tarjeta. La tarjeta es de color amarillo. Los usuarios con discapacidad pueden personalizar su tarjeta para que pueda realizar el pago en el validador para sillas de ruedas y ascensores.

Perfil Universitario, tiene una tarifa preferencial de S/1.25 ruta troncal y S/0.25 ruta alimentadora (válida de lunes a sábados), es adquirida por estudiantes universitarios y técnicos hasta que caduque el carnet universitario; es personal e intransferible. La tarjeta es de color verde. (Metropolitano, 2019)

Perfil Escolar, tiene una tarifa preferencial de S/1.25 ruta troncal y S/0.25 ruta alimentadora; es adquirida por escolares (válida de lunes a viernes de marzo a diciembre); es personal e intransferible. La tarjeta es de color naranja.

Perfil Pase Libre, tiene una tarifa preferencial de S/0 ruta troncal y S/0 ruta alimentadora; son facilitadas en cada estación por el personal de cada estación en el caso de la ruta troncal y por los conductores en el caso de las rutas alimentadoras, son utilizadas por la Policía Nacional del Perú, Cuerpo de Bomberos Voluntarios del Perú y trabajadores del Metropolitano. (Metropolitano, 2019)

2.3.2.4 Tarifarios según tipo de tarjeta

Existen tres tipos de tarifarios

Tarifa General Representa el costo por el uso exclusivo de la vía, ruta troncal cuyo precio es de S/2.50. (Metropolitano, 2019)

Tarifa Preferencial Representa el cobro diferenciado por un servicio, y está determinado según el perfil del usuario con uso exclusivo de la ruta troncal cuyo precio es de S/1.25 y con uso exclusivo de la ruta alimentadora cuyo precio es de S/0.25. (Metropolitano, 2019)

Tarifa Integrada Representa el cobro de un servicio mixto, que integra dos o más modalidades de transportarse de un punto a otro, sin salir del sistema con Uso de la Ruta Alimentadora + Troncal + Alimentadora o Troncal + Alimentador (viceversa) cuyo precio es de S/2.50. Para hacer uso de la Tarifa Integrada, el viaje debe ser de Sur a Norte o de Norte a Sur. Si se viaja con dos o más personas, las validaciones deben de realizarse en el mismo validador. El tiempo de vigencia para activar su uso es de 122 minutos. (Metropolitano, 2019)

2.3.2.5 Términos

A continuación, se explicarán algún término relevante que se usarán más adelante.

Servicios; son rutas, dentro de las rutas de circulación troncal y de los alimentadores.

Usuario; son los pasajeros que circulan dentro del Metropolitano

Terminales; son los puntos o lugares de ingreso, salida y/o conexión de pasajeros, con facilidades para que los usuarios realicen intercambio entre buses troncales y buses alimentadores y/u otros medios de transporte. Cuentan con mecanismos de control, así como con facilidades para la venta y validación de las tarjetas. (Metropolitano, 2019)

Estaciones; son los puntos o lugares de ingreso, salida y/o conexión de pasajeros al Metropolitano, ubicadas en los corredores o ejes viales troncales, con mecanismos de control, así como facilidades para la venta y validación de las tarjetas. (Metropolitano, 2019)

Tarjeta Inteligentes Sin Contacto; se utilizan en el Metropolitano, son tarjetas con tecnología tipo RFID (Radio Frequency Identification, Identificación por Radio Frecuencia) tiene como objetivo identificar el validador de una forma rápida, con poca transmisión de información y en un entorno cercano, son similares como los identificadores de código de barras. (Rodríguez et al., 2006)

Validador de tarjetas; es el lector de Tarjetas Inteligentes RFID, a través de este aparato reconoce principalmente el pago del servicio de transporte.

Torniquete de ingreso; también llamado molinete, es una especie de barrera física, que tras verificar la autorización del validador permite o impide el acceso al Metropolitano.

Archivo tipo CSV; los archivos CSV (Comma Separated Values, Valores Separados por Comas) son un tipo de documento en formato abierto sencillo para representar datos en forma de tabla, en las que las columnas se separan por comas u otro carácter similar. Los datos de los usuarios que son captados por los validadores son guardados en un archivo tipo csv para cada usuario.

2.3.3 Simulación

Existen diversos softwares de simulación en transporte, a continuación, se menciona el software de simulación que se usará en esta tesis. (PTV GROUP, 2019)

VISUM, realiza el análisis de tránsito, planificación y gestión de datos basada en SIG. Modela de manera consistente a todos los usuarios de las vías públicas y sus interacciones y se ha convertido en un estándar reconocido en el campo de la planificación de transporte. Sirve para modelar redes de transporte y demanda de viajes, analizar los cambios del tránsito previstos, planificar servicios de transporte público y desarrollar avanzadas estrategias y soluciones de transporte. Las principales funciones que usaremos del software son:

Cálculos de demanda

- Modelo de 4 pasos.
- Modelo de demanda basado en rutas (Tour based).
- Modelo de demanda con distribución simultánea y cálculo de elección de modo (EVA).
- Modelo de demanda anidado.

Métodos y operaciones de asignación de Transporte Público (PuT)

- Asignación basada en tablas de horarios (Timetable based).
- Asignación basada en despachos (Headway based).
- Bloqueo de línea.
- Modelos de tarifas estándar, hasta sistemas complejos de precios en un modelo.
- Cálculo de costes de línea.

CAPÍTULO III : ANÁLISIS METODOLÓGICO

Con el objetivo de generar rutas óptimas existen dos formas de obtenerlas: el primero, recurriendo a los métodos tradicionales del transporte; el segundo, es combinando los criterios del método tradicional y las herramientas tecnológicas de procesamiento de datos reales, basado en Ciencia de Datos y Big Data. Esta segunda, es la que se desarrollará, a la vez se puede ejecutar de dos formas distintas: (01) usando algoritmos existentes para generar rutas; y (02) generando un propio algoritmo. En el caso de la presente tesis se propondrá un propio algoritmo para la generación de las rutas y se usará un software para analizar que algoritmo existente se adecua mejor a la data, basado en Machine Learning, para la predicción de las rutas; ver Figura 15. Luego de obtener las rutas óptimas, se procederá a identificar las mejoras comparativas del “modelo en funcionamiento actual” contra el “modelo planteado de generación de rutas”.

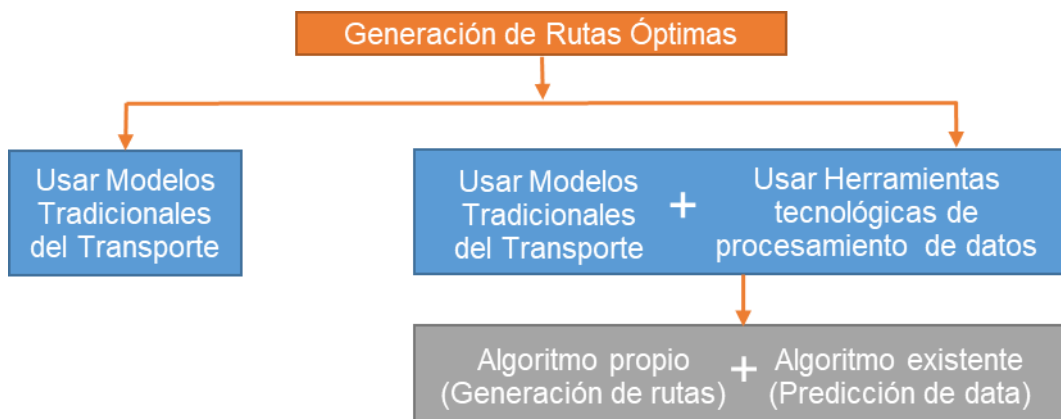


Figura 15 Propuesta para la generación de rutas óptimas
Fuente: Elaboración propia

El modelo para la generación de rutas usa criterios de Modelos del Transporte como el Análisis del Sistema de Transporte (TSA) y Método de las 4 etapas (FSM); y a su vez utiliza herramientas de tecnologías de procesamientos de datos como Data Science (Ciencia de Datos), Big Data (Grandes Datos) y Machine Learning (Aprendizaje automático). Para ello se usó distintos lenguajes de programación como R y Python, así como softwares Weka y Visum.

En seguida, se presenta el diseño metodológico para ser aplicado a cualquier caso de estudio de transporte público que tenga los datos de origen destino de todos los usuarios y luego se presentará el diseño metodológico aplicado específicamente al caso de estudio del Metropolitano (BRT).

3.1 DISEÑO METODOLÓGICO

En este subcapítulo, se explicará los pasos a seguir para generar rutas óptimas en transportes públicos masivos; los pasos están basados en las herramientas que se usaran, a su vez se mencionará como los leguajes de programación, softwares y framework contribuyen directa o indirectamente en la serie de pasos.

3.1.1 Serie de Pasos

Serán el eje guía para el desarrollo del modelo de generación de rutas.

- Paso 1: Recolección de datos
- Paso 2: Almacenamiento de datos
- Paso 3: Data Analytics (pre procesamiento)
- Paso 4: Limpieza de datos
- Paso 5: Filtros o Selección de parámetros
- Paso 6: Estructura de los datos
- Paso 7: Análisis de los datos
- Paso 8: Generación de rutas
- Paso 9: Programación de rutas
- Paso 10: Macro Simulación
- Paso 11: Comparación de sistema actual y propuesto

Esquematizando la serie de los pasos que se debe seguir:



Figura 16 Metodología: pasos a seguir
 Fuente: Elaboración propia.

3.1.2 Herramientas

Serán el soporte de desarrollo a lo largo de la serie de pasos.

- Herramienta 1: Ciencia de Datos
- Herramienta 2: Big Data
- Herramienta 3: Machine Learning

Las herramientas de la Ciencia de la Computación que se utilizarán se interrelacionan entre ellas, como se muestra en el siguiente esquema.

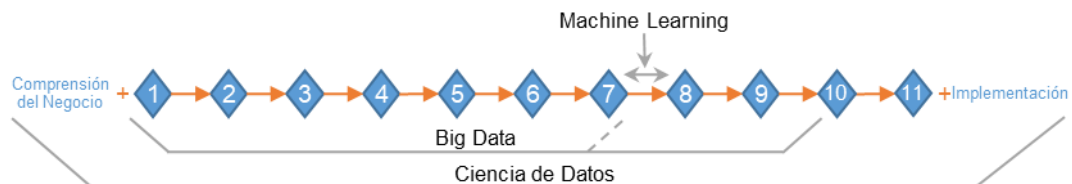


Figura 17 Esquema de herramientas
 Fuente: Elaboración propia

A continuación, se explica más a detalle las herramientas donde intervendrán la serie de pasos.

3.1.2.1 Herramienta 1: Ciencia de Datos

Esta herramienta abarca todos los pasos, esto es debido a que la Ciencia de Datos es la intersección de la de la estadística y matemática, ciencia de la computación, y el dominio en el conocimiento de la especialidad que se va a desarrollar; ver Figura 18. En el marco teórico, la Ciencia de Datos está conformada por 6 etapas, cada etapa está compuesta por tareas, es por ello que mencionaremos que pasos intervienen en cada etapa.

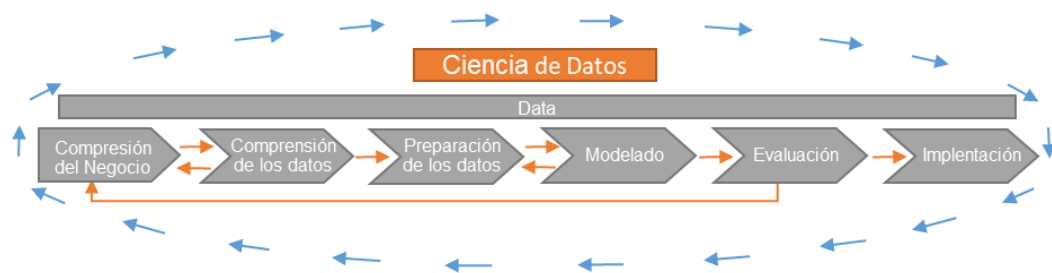


Figura 18 Etapas de la Ciencia de Datos
Fuente: Elaboración propia

I) Etapa 1: Comprensión del negocio, son los objetivos y requisitos del proyecto desde el punto de vista empresarial en este caso se refiere al sector de transporte público masivo, esta etapa no está implementada en dentro de ningún paso, debido a que, el objetivo es la generación de rutas. Dentro de esta etapa están las siguientes tareas:

- 1) Determinar el objetivo
- 2) Evaluar la situación
- 3) Determinar las metas
- 4) Elaborar un plan

II) Etapa 2: Comprensión de los datos, esta etapa se encuentra del paso 1 al paso 3. En esta etapa también interviene la Herramienta 2: Big Data. Las tareas de esta etapa son las siguientes:

- 1) Recopilar los datos, es el Paso 1
- 2) Describir los datos, es parte del Paso 2
- 3) Explorar los datos, es el Paso 3
- 4) Verificar la calidad de los datos, es parte del Paso 3

III) Preparación de los datos, esta etapa se encuentra del paso 4 al paso 6. En esta etapa también interviene la Herramienta 2: Big Data. Las tareas de esta etapa son las siguientes:

- 1) Seleccionar los datos, es el Paso 5
- 2) Limpiar los datos, es el Paso 4
- 3) Construir los datos, es parte del Paso 6
- 4) Integrar los datos, es parte del Paso 6
- 5) Formato de los datos, es parte del Paso 6

IV) Modelado, esta etapa se encuentra del paso 8 al paso 10. En esta etapa también interviene la Herramienta 3: Machine Learning. Las tareas de esta etapa son las siguientes:

- 1) Seleccionar la técnica del modelado, es parte del Paso 8
- 2) Generar el diseño de prueba, es parte del Paso 8
- 3) Modelo de construcción, es parte del Paso 9
- 4) Modelo de evaluación, es parte del Paso 10

V) Evaluación, esta etapa se encuentra el paso 11. Las tareas de esta etapa son las siguientes:

- 1) Evaluar los resultados, es parte del Paso 11
- 2) Proceso de revisión, es parte del Paso 11
- 3) Determinar los siguientes pasos, es parte del Paso 11

VI) Implementación, es explicar cómo se deberá implementar el modelo, esta etapa al igual que la etapa 1: comprensión del negocio no está incluida dentro de los pasos. Las tareas de esta etapa son las siguientes:

- 1) Plan de implementación
- 2) Monitoreo y mantenimiento del Plan
- 3) Elaborar el informe final
- 4) Revisión del proyecto

Esquematisando las etapas y tareas de la Ciencia de Datos intervienen en la serie de pasos, planteado para el modelo de generación de rutas.



Figura 19 Intervención de la Ciencia de Datos en la secuencia de pasos
Fuente: Elaboración propia

3.1.2.2 Herramienta 2: Big Data

Esta herramienta abarca del paso 1 al 9, principalmente debido a que en esta serie de pasos se hará uso de grandes volúmenes de datos y el código estará paralelizado. En el marco teórico, el Big Data está conformada por 4 procesos, es por ello que mencionaremos que pasos intervienen en cada proceso.

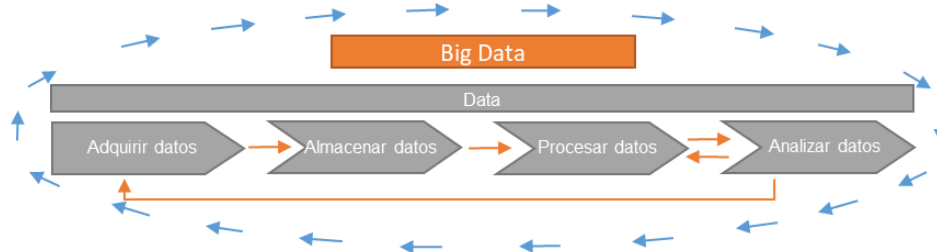


Figura 20 Procesos del Big Data
Fuente: Elaboración propia

- I) Adquirir los datos, a este proceso también se le puede llamar ingesta o recopilación de datos, en este proceso se encuentra el paso 1.
- II) Almacenar los datos, en este proceso se encuentra el paso 2.
- III) Procesar los datos, en este proceso se encuentra del paso 3 al paso 6, que son el pre procesamiento, limpieza selección y la estructura de los datos respectivamente.
- IV) Analizar los datos, en este proceso intervienen los pasos 7, 8 y 9, ya que en estos pasos se analiza la data de entrada, los resultados del modelo de “Generación de rutas óptimas”, así tener un buen modelo.

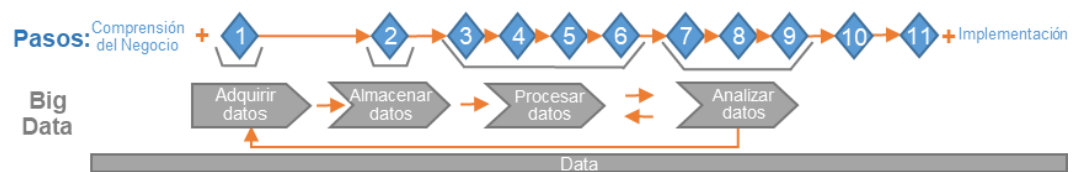


Figura 21 Intervención del Big Data en serie de pasos
Fuente: Elaboración propia

3.1.2.3 Herramienta 3: Machine Learning

La herramienta Machine Learning interviene directamente entre los pasos 7 y paso 8, lo cual servirá para hacer la predicción de la generación de rutas.

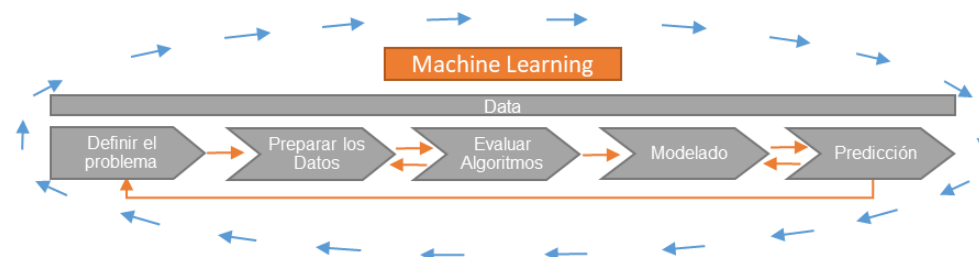


Figura 22 Procesos de Machine Learning
Fuente: Elaboración propia

Esta herramienta consta de 5 procesos en los cuales se mencionarán los pasos que intervendrán en cada proceso.

- I) Definir el problema, es parte de la comprensión del negocio, debido a que en este proceso se define que se va a predecir.
- II) Preparar los datos, la forma de preparar los datos siguen la metodología del paso 3 al 7.
- V) Evaluar los algoritmos, este proceso está entre el paso 7 y 8
- VI) Modelado, este proceso según el concepto de Ciencia de Datos es el paso 8, para el modelo de Machine Learning estará dentro del paso 7 y 8
- VII) Predicción, este proceso según el concepto de Ciencia de Datos son los pasos 9 y 10, para el modelo de Machine Learning estará dentro del paso 7 y 8

Se usará un software para analizar que algoritmos se adecúan mejor a la data del transporte para luego realizar el modelo de “Generación de rutas óptimas”.

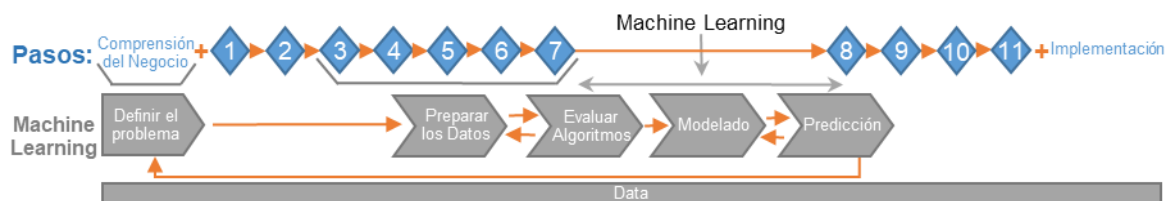


Figura 23 Intervención del Machine Learning en la secuencia de pasos
Fuente: Elaboración propia

3.1.2.4 Otras Herramientas

Los pasos y las herramientas son apoyados por los lenguajes de programación, softwares y framework; las cuales facilitaran el desarrollo del modelo de “Generación de rutas óptimas”.

Lenguajes de programación: para el tratamiento de los datos se puede usar lenguajes de programación interpretados y/o compilados. Intervienen directa e indirectamente desde el paso 1 al paso 10.

Softwares: contribuirán a facilitar el desarrollo de algunos pasos. Pueden intervenir para el tratamiento de los datos, análisis de los algoritmos, hacer simulaciones, entre otros. Intervienen directa e indirectamente desde el paso 1 al paso 10.

Framework (Marco de Trabajo): debido a que el framework son facilitadores para la construcción de modelos; ayudará al tratamiento de los datos (recopilar, almacenar, procesar y analizar) para realizar Big Data, es decir desde el paso 1 al paso 9.

3.2 METODOLOGÍA PARA EL CASO DE ESTUDIO

Se aplicará el diseño metodológico descrito en el capítulo anterior para este caso de estudio del Metropolitano, utilizando la serie de pasos, herramientas, lenguajes de programación y softwares.

Pasos: En los capítulos IV, V y VI se detallará la descripción y desarrollo de cada paso. Se muestra en el siguiente esquema (Figura 24) el agrupamiento de la serie de pasos, de acuerdo al nivel de complejidad, donde por ejemplo el paso 1 (Recolección de datos) y el paso 2 (Almacenamiento de los datos) están agrupados en uno solo, esto es debido a que estos dos procesos no requieren tanto detalle de desarrollo en esta tesis, sin embargo, todos los pasos son explicados.

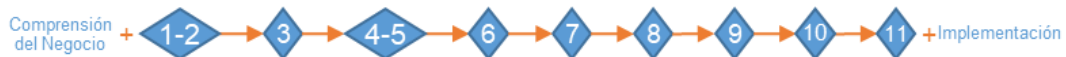


Figura 24 Metodología: agrupamiento de pasos
Fuente: Elaboración propia

Herramientas:

Herramienta 1: Ciencia de Datos, se desarrollan las etapas del 2, 3, 4 y 5 que son comprensión de los datos, preparación de los datos, modelado y evaluación respectivamente; la etapa 1: comprensión del negocio y la etapa 6: implementación se desarrollara de manera general.

Herramienta 2: Big Data, se desarrollarán los 4 procesos que son adquisición, almacenamiento, procesamiento y análisis. Los procesos de adquisición y almacenamiento son los pasos 1 y 2, los cuales serán descritos y no desarrollados, debido a que la data de los usuarios es recolectado y almacenado por la empresa Operadores ACS Solutions Perú S.A. y facilitada por Instituto Metropolitano Protransporte de Lima. El procesamiento y análisis son desarrollados en el capítulo IV (son los pasos del 3 al 7) y parte del V (pasos 8 y 9).

Herramienta 3: Machine Learning, se desarrollará parte del modelo de predicción de rutas, es decir se analizar con ayuda del software Weka el algoritmo que mejor se adecua a la data para que en futuras investigaciones se desarrolle el modelo de predicción de rutas. Esta herramienta la desarrollaremos entre los pasos 7 y 8.

Lenguajes de programación: se usará librerías de tratamiento de datos para un fácil manejo y control de la gran cantidad de datos:

- Lenguaje R: Se utilizará para el paso 3, lo cual es el pre procesamiento de los datos o Data Analytics.
- Lenguaje Python: Se utilizará para los pasos del 3 al 9, los cuales son limpieza de datos, filtros o selección de parámetros, estructura de los datos, análisis de los datos, generación de rutas y programación de rutas.

Esquemmatizando se muestra la intervención de los lenguajes en los pasos.



Figura 25 lenguajes que intervendrán
Fuente: Elaboración propia

Softwares: contribuirán a facilitar el desarrollo de algunos pasos.

- Weka: se utilizará entre el paso 8 y 9, servirá para la obtención del algoritmo adecuado a la data que será utilizado para la predicción de las rutas.
- Visum: se utilizará en los pasos 10 y 11, servirá para comparar el modelo de funcionamiento actual del caso de estudio y para la simulación del modelo planteado.

Esquemmatizando se muestra la intervención de los softwares en los pasos.

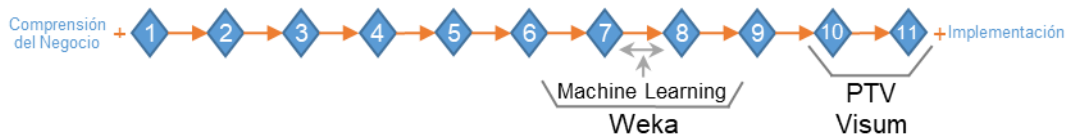


Figura 26 Software que intervendrán
Fuente: Elaboración propia

El diseño metodológico del caso de estudio para la generación de rutas se basó en la serie de pasos junto con el desarrollo general de la comprensión y la implementación del negocio, utilizando herramientas, lenguajes de programación y softwares. A continuación, se hace un esquema de la secuencia del diseño metodológico para este caso de estudio (Ver Figura 28).

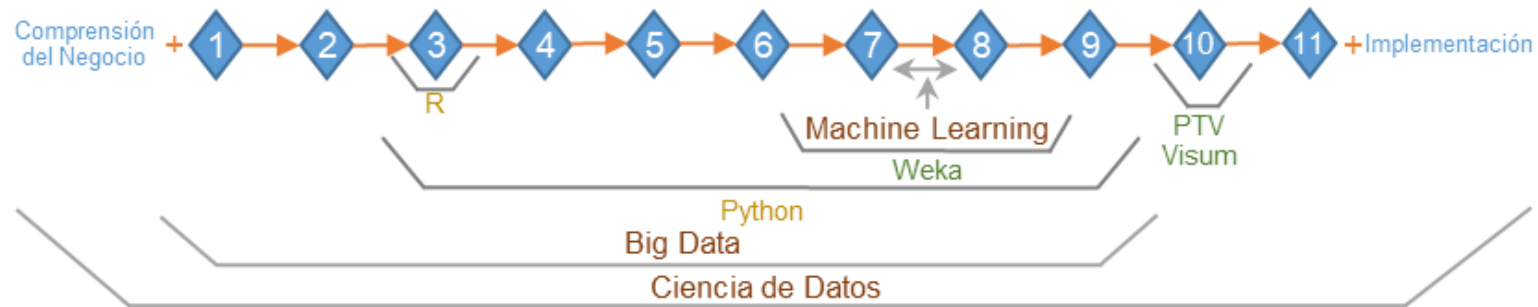


Figura 27 Metodología para generación de rutas del Metropolitano
Fuente: Elaboración propia

En la Tabla 3 siguiente se muestra como interviene los lenguajes de programación, los softwares y las herramientas que intervendrán en los pasos del modelo de “Generación de rutas óptimas”.

Tabla 3 Intervención de la secuencia de pasos

		Serie de pasos										
		1	2	3	4	5	6	7	8	9	10	11
Herramientas	Data Science	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Big Data	✓	✓	✓	✓	✓	✓	✓	✓			
	Machine Learning			✓	✓	✓	✓	✓				
Lenguaje de Programación	R			✓								
	Python			✓	✓	✓	✓	✓	✓			
Software	Weka			✓	✓	✓	✓	✓				
	Visum										✓	

LEYENDA:	
Serie de pasos	
1	Recolección de datos
2	Almacenamiento de datos
3	Data Analytics (pre procesamiento)
4	Limpieza de datos
5	Filtros o Selección de parámetros
6	Estructura de los datos
7	Análisis de los datos
8	Generación de rutas
9	Programación de rutas
10	Macro Simulación
11	Comparación de sistema actual y propuesto

Fuente: Elaboración propia.

CAPÍTULO IV: TRATAMIENTO DE LA INFORMACIÓN

En este capítulo se desarrolla el tratamiento de la información de los datos para ello utilizaremos la herramienta 1: Ciencia de Datos, la herramienta 2: Big Data y la herramienta 3: Machine Learning; en las dos primeras se usó los lenguajes de programación R y Python, en la tercera herramienta se usó el software Weka.

4.1 HERRAMIENTA 1: CIENCIA DE DATOS

Con esta herramienta se ve desde el inicio hasta el final del desarrollo de esta tesis, puesto que intervienen el campo de la Ciencia de la Computación, Estadística y Matemática, y el especialista en Transporte. Para ello en este subcapítulo se explicará de manera general la etapa 1: comprensión del negocio y la etapa 6: implementación del negocio, las etapas intermedias serán descritas en los otros capítulos siguiendo la serie de pasos.

4.1.1 Etapa 1: Comprensión del Negocio

La comprensión del negocio dependerá a qué tipo de empresa está dirigido la Ciencia de Datos. Para esta etapa se supondrá que el cliente es Protransporte (Metropolitano), es decir la comprensión del negocio está dirigida a una entidad de Sistemas de Buses de Tránsito Rápido (BRT), este pertenece al Transporte Público Masivo. Las tareas de esta etapa son las siguientes:

4.1.1.1 Determinar el objetivo

Los objetivos del negocio a nivel empresarial para este caso de estudio son los siguientes:

- Mejorar las rutas del sistema del Metropolitano manteniendo la cantidad de buses que circulan en la red.
- Mejorar el tiempo de viaje de los usuarios manteniendo los costos operativos del sistema.

4.1.1.2 Evaluar la situación

Los recursos, Protransporte cuenta con una base de datos de los usuarios, cámaras de monitoreo en todas las estaciones, los tiempos de los semáforos a lo largo de todo su recorrido los BRT cuentan con GPS y la programación de las salidas de los buses desde los terminales.

Las restricciones, Protransporte no cuenta con un área legal especializada en seguridad informática para el tema de los datos, ni con un área especializada en Ciencias de datos.

Las suposiciones, la base de datos de los usuarios son datos de ingreso, es por ello que se hará algunas suposiciones para el modelo.

4.1.1.3 Determinar las metas de la Ciencia de Datos

Las metas de la Ciencia de Datos para este caso de estudio es el siguiente:

- Mejorar las rutas en base a la demanda de usuarios.
- Disminuir el tiempo de espera de los usuarios en un 30% en las estaciones del Metropolitano.

4.1.1.4 Elaborar un plan

El plan se basará principalmente en generar rutas en base a la demanda de los usuarios disminuyendo el tiempo espera en las estaciones, así, si se tiene las rutas de los buses bien programados los usuarios podrán saber a qué hora llega y parte el bus de cada estación. Estas rutas se basarán en la data de los mismos usuarios ya que por lo general la mayoría de los usuarios realizan las mismas actividades diariamente. Para cumplir con este plan se debe tener en cuenta lo siguiente

- Elaborar un modelo “Generación de rutas óptimas”, este modelo debe de basarse en la demanda de los usuarios, para ello se realizará una serie de pasos los cuales servirán para tratar la data, luego se elaborará un algoritmo que genere las rutas de modo que sea óptimo comparado con la situación actual. Los tiempos de espera se disminuirá de manera forzada, es decir se colocará el parámetro de tiempo de espera, como entrada en el algoritmo de modo que el operador defina cuanto será el tiempo máximo de espera de los usuarios en el sistema.
- Predicción de las rutas óptimas: una vez que se tenga el modelo de “Generación de rutas óptimas”, para obtener rutas de un día que no se tenga data se ingresa los datos predichos con la herramienta de Machine Learning. Es decir, se podrá predecir las rutas para épocas de vacaciones, feriados, clases, entre semana, domingo, por horarios, etc. En esta tesis solo se obtendrá el mejor algoritmo de predicción.

Para lo planteado se tiene que apoyar en las herramientas del Big Data y Machine Learning, debido a que se utiliza gran volumen de datos y se desea predecir las rutas.

4.1.2 Etapa 6: Implementación

En un supuesto que se realice una implementación de los dos modelos, es decir la generación y predicción de rutas óptimas, la implementación se haría directamente en el área de programación de salida de los buses del Metropolitano. Esta etapa consta de las siguientes tareas.

4.1.2.1 *Plan de implementación*

La forma de implementar los modelos es automatizando del paso 1 al paso 9, de tal modo que tenga como entrada la recolección de los datos en tiempo real para luego tener como salida la programación de las salidas de los buses. Los usuarios solo verían las mejoras en los tiempos de viajes durante el todo el año, ya que con el modelo de predicción la programación de las rutas varía según la fecha en que se encuentre, es decir en feriados, época de escolaridad, horas pico, horas valle, etc.; esto implicara que los usuarios tengan conocimiento de los horarios de llegada de los buses en cada estación. También lo recomendable es elaborar un aplicativo móvil donde se mostraría las variaciones de las programaciones de los buses, rutas de viajes recomendados, estaciones con mayor cantidad de usuarios, etc.

4.1.2.2 *Monitoreo y mantenimiento del Plan*

Cada periodo de tiempo se debe calibrar y validar los modelos debido a que todos los días del año los modelo tendrán salidas y entradas con variaciones según las fechas y horarios, por ejemplo, en los días laborales tiene comportamiento diferente a los días feriados.

4.1.2.3 *Elaborar el informe final*

Una vez culminado los modelos se debería enviar un informe a Protransporte con los resultados y conclusiones para que puedan ser implementados.

4.1.2.4 *Revisión del proyecto*

Cuando se culmine los modelos es bueno documentar lo que se hizo bien o mal para futuras investigaciones, así como entrevistar a las personas involucradas en el desarrollo de los modelos.

4.2 HERRAMIENTA 2: BIG DATA

La herramienta del Big Data la utilizamos en el desarrollo de esta tesis, debido a que la metodología planteada está orientada a casos de transporte con gran cantidad de datos, variedad de datos teniendo en cuenta velocidad del procesamiento de los datos. En el marco teórico el Big Data debe cumplir las características de volumen, variedad y velocidad.

El Big Data para ser aplicado tiene procesos los cuales serán desarrollados en la serie de pasos.

Proceso 1: Adquisición de los datos

- Paso 1: Recolección

Proceso 2: Almacenamiento de los datos

- Paso 2: Almacenamiento

Proceso 3: Procesamiento de los datos

- Paso 3: Pre procesamiento
- Paso 4: Limpieza
- Paso 5: Filtro o selección
- Paso 6: Estructura

Proceso 4: Análisis de los datos

- Paso 7: Análisis

Para el caso de estudio del Metropolitano la cantidad de los datos que se está usando no cumple la característica del Big Data, debido a que los datos son solo de un mes y se usará los datos solo de la Troncal, es por ello que no cumple las características de volumen y variedad. El modelo se desarrollará, simulando que la data es de varios años (Volumen) y con otras fuentes de información como datos de coordenadas de los paraderos, velocidad de los buses, tiempo de los semáforos entre otros que serán explicados más adelante (variedad). El modelo de generación de rutas cumple con la característica de velocidad de procesamiento. Haciendo esas consideraciones el modelo cumple con las características de Big Data.

A continuación, se desarrollará los pasos del 1 al 7 los cuales forman parte de los procesos del Big Data, dentro de cada paso se hace una descripción y se explica la aplicación al caso de estudio.

4.2.1 Paso 1: Recolección de datos

4.2.1.1 Descripción

La recolección de los datos puede ser de cualquier sistema de transporte ya sea a través de las tarjetas (sistema RFID / NFC) o del GPS (cualquier dispositivo electrónico que cuenta con Sistema de Posicionamiento Global) que capture el origen y el destino del usuario. Otras fuentes de datos adicionales son capturar la ruta del recorrido del bus, las coordenadas de las estaciones o paraderos, los tiempos de los semáforos a lo largo del recorrido, la velocidad y la capacidad de cada bus según el tipo de transporte que se esté analizando (BRT, tren, buses, etc.)

4.2.1.2 Caso Metropolitano

Los datos que se usaran son los siguientes:

Validaciones de los usuarios; los datos fueron recolectados en todas las estaciones a través de las tarjetas (Sistema RFID) cuando los usuarios ingresan al Metropolitano, los cuales pasan por un validador, cuando giran el torniquete; así el sistema captura los datos de ingreso de cada uno de los usuarios. Esta data es recolectada por la empresa Operadores ACS Solutions Perú S.A (ver Figura 28).



Figura 28 Validación de un usuario
Fuente: Pagina web de Metropolitano (Metropolitano, 2019)

Ruta de recorrido; el sistema analizado consta de rutas Troncal y Alimentador como se muestra en la Figura 29, la ruta Troncal inicia en la estación Naranjal y termina en la estación Matellini. Esta información fue recolectada de la página web del Metropolitano.



Figura 29 Ruta del Metropolitano
 Fuente: Pagina web de Metropolitano (Metropolitano, 2019)

Coordenadas de las estaciones; se recopiló de forma manual, es decir se buscó las 38 estaciones de la ruta Troncal del Metropolitano en la página del Open OpenStreetMap. En la Figura 30 se muestra un ejemplo de la toma de datos de la estación Naranjal, en este caso la ubicación (-11,9827434; -770786044) es del punto de color naranja.

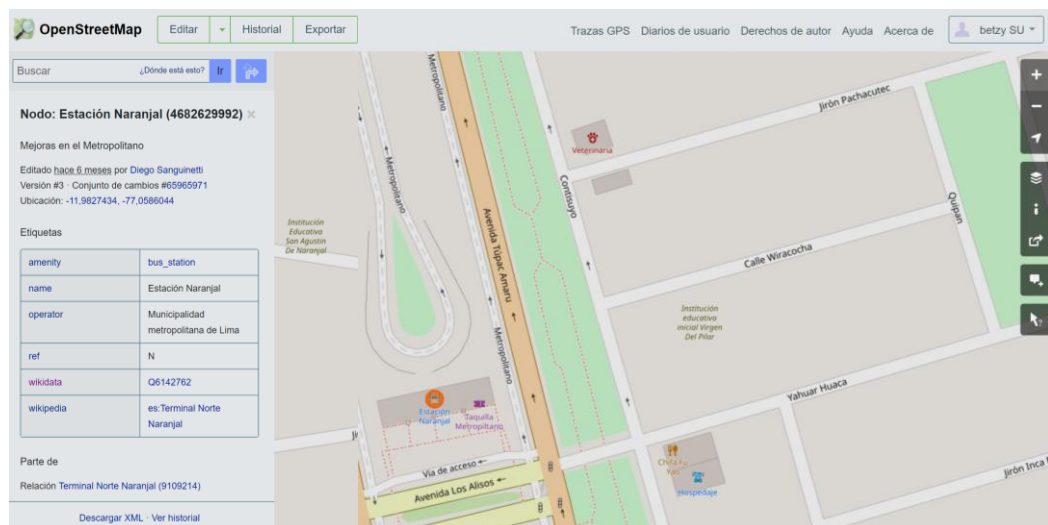


Figura 30 Ejemplo de obtención de coordenadas
 Fuente: Nodo: Estacion Naranjal (OpenStreetMap, 2019)

Velocidad de los buses; los datos promedios fueron brindados por Protransporte, también se verifico en campo con la finalidad de calibrar.



Figura 31 Conductor de un bus del Metropolitano
Fuente: Pgina web de Metropolitano (Metropolitano, 2019)

Capacidad de los buses; esta información fue brindada por el área de programación de Protransporte y validada en campo. De la información brindada por Protransporte la capacidad máxima del bus es 160 personas. Para realizar la calibración y validación se tomará en cuenta lo siguiente, los datos serán recopilados en la hora pico de 7:00 a 8:30am en la estación más congestionada que es la estación Naranjal, se tomó en el servicio Expreso 2 debido a que este es el más demandado de tal manera que se tiene que asegurar que el bus irá en su capacidad máxima. Ver Figura 32 y ANEXO 4.



Figura 32 Capacidad máxima en un bus Expreso 2
Fuente: Fotografía propia tomada el 17/10/2019 08:33:00 am

Programación de las salidas de los buses; esta información fue brindada por el área de programación de Protransporte y validada en campo. Para validar estos datos, se verifico el listado de salida de buses del Patio Norte en la estación Naranjal, los cuales, si coincidió. Para verificar los tiempos de viajes se subió a 3 buses. Esta información servirá para calibrar los modelos de simulación.

En la Figura 33 se observa el expreso 2 en el horario de 7 a 8 de la mañana, se verifica con la placa del bus, el nombre de la ruta y la hora en la estación Naranjal la salida del bus del Patio Norte del Metropolitano.



Figura 33 Validación de la salida de buses estación Naranjal
Fuente: Fotografía propia tomada el 17/10/2019 07:30:00 am

Los datos brindados de manera formal por Protransporte fueron las validaciones de los usuarios, los otros datos antes mencionados fueron acotados verbalmente por la gerencia del área de programación de Protransporte. Ante ello se resalta la valides de los datos de los usuarios, ya que es importante para los resultados del modelo de “Generación de rutas óptimas”, porque tiene como entrada estos datos.

4.2.2 Paso 2: Almacenamiento de datos (lectura)

4.2.2.1 Descripción

Los datos son guardados en servidores, estos pueden ser de empresas como Google Cloud Platform, Amazon Web Services o servidores propios de almacenamiento. En este paso también se hace una lectura de los datos, es decir la descripción y el tipo de los datos.

4.2.2.2 Caso Metropolitano

Todos los datos recolectados son de tipo estructurado, es decir se puede visualizar en filas y columnas. A continuación, se hará la descripción de los datos:

Validaciones de los usuarios

La data se almacena en los servidores de la empresa Operadores ACS Solutions Perú S.A., esta data es almacenada en tiempo real, de todas las validaciones realizadas durante el día de todas las estaciones. La data fue facilitada por Instituto Metropolitano Protransporte de Lima donde contiene información de todos los días del mes de mayo de 2018, que tiene 31 archivos (un archivo por día) los cuales contiene alrededor de 18 millones de validaciones con un tamaño de almacenamiento de 2 GB. Las validaciones del ingreso de cada usuario son almacenadas en archivo tipo “.dat” y luego los archivos de cada validación son agrupados por día en un archivo tipo “.csv” (ver ANEXO 1). Cada archivo tipo “.csv” tienen las siguientes variables: Fecha, validador, secuencia, archivo, tarjeta, perfil, tipo de validador, tipo de validación, parada y monto que serán descritas a continuación:

1. Fecha; indica que el día, mes, año y hora en que se realizó la validación.
2. Validador; contiene la información del código del lector de tarjetas o validador de tarjetas, cada lector tiene un código de identificación.
3. Secuencia; es la numeración contabilizada en el validador, cada día en cada validador comienza desde uno, es decir hace un conteo de la cantidad de usuarios que ingresan por un validador determinado.
4. Archivo; es el nombre del archivo “.dat”, este archivo es el registro de cada validación que fue extraído al usuario.
5. Tarjeta; es el código de identificación de la tarjeta del usuario.
6. Perfil; es el tipo de usuario, para el registro contiene al perfil General (GE), Universitario (UN), Escolar (ES) y Pase Libre (PL).

7. Tipo de validador; es la condición de la persona, existen dos, el usuario Normal (N) y el usuario Minusválido (M).
8. Tipo de validación; Vía en la que el usuario se moviliza; en este sistema son dos vías: Troncal (T) y el Alimentador (A).
9. Parada; es la estación en la que el usuario ingresa al sistema del Metropolitano.
10. Monto; es el costo del usuario, esta variable depende del perfil del usuario.

Tabla 4 Muestra del archivo de la data brindada por PROTRANSPORTE.

FECHA	VALIDADOR	SECUENCIA	ARCHIVO	TARJETA	PERFIL	TIPO_VALIDADOR	TIPO_VALIDACION	PARADA	MONTO
04/05/2018 7:50:11	4040	1217	2018\05\04\S_072_001_001_38535_06112_20180504_074300.DAT	340009838	GE	N	T	Uni	2,5
04/05/2018 18:53:58	12047	474	2018\05\04\S_072_001_011_38392_07526_20180504_120000.DAT	965593341	UN	N	A	AS-08	0,5

Fuente: Elaboración propia.

Ruta de recorrido

La vía Troncal comprende 38 estaciones de las cuales 2 son terminales, una central y las demás son rutas intermedias. Las vías alimentadoras que parten de los terminales y de la estación central.

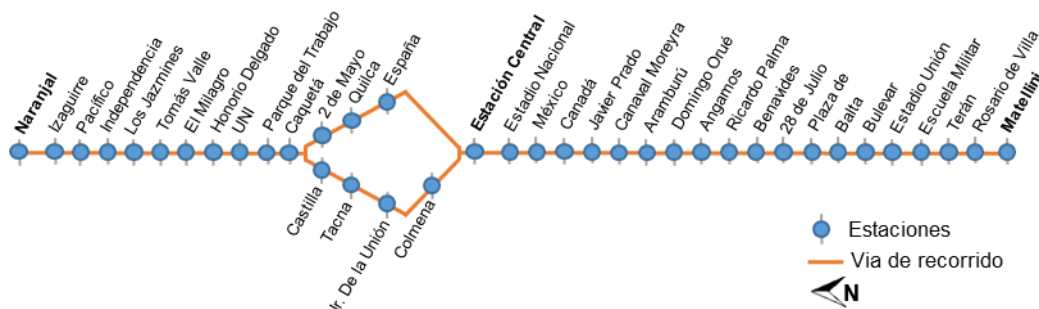


Figura 34 Vía troncal a hacer analizada

Fuente: Elaboración propia.

Coordenadas de las estaciones

Las coordenadas dependerán de la ubicación en la que se encuentren en el OpenStreetMap, es decir en los terminales las coordenadas se encuentran en una de las puertas de ingreso y en las otras estaciones las coordenadas se encuentran en el centro de gravedad de la edificación. Los 38 puntos geográficos están expresados en coordenadas angulares de latitud y longitud. Las coordenadas reales de las estaciones se ven en el ANEXO 2

Tabla 5 Coordenadas de las Estaciones del Metropolitano

COORDENADAS DEL METROPOLITANO		
Estación	Latitud	Longitud
Estación Naranjal	-11,9827434	-77,0586044
Izaquirre	-11,9895485	-77,0570156
...
Matellini	-12,1782643	-77,0104387

Fuente: Elaboración propia.

Velocidad de los buses

Los datos de velocidad promedio brindada por Protransporte es de 20km/h en los servicios regulares y 33km/h en los expresos. En la Tabla 6 se observa los datos de las velocidades, estos datos servirán para la calibración y validación del modelo de “Generación de rutas óptimas”.

Tabla 6 Velocidades promedio de los BRT

Velocidad máxima de bus	60km/h
Velocidad promedio de circulación	30km/h
Intervalo de velocidad de circulación (Protransporte)	[20km/h, 33km/h]
Intervalo de velocidad de circulación (Campo)	[20km/h, 45km/h]

Fuente: Elaboración propia

Capacidad de los buses

Los datos para la calibración y validación son recolectados en el servicio expreso 2, obteniendo como resultado que durante el intervalo de tiempo que la capacidad máxima es 160 personas, como se muestra en la Tabla 7 y en el ANEXO 4.

Tabla 7 Capacidad máxima del BRT

Capacidad máxima del bus tomada en el Expreso 2 de N-S						
Nº	Hora	ID BUS	Puerta delantera	Puerta del centro	Puerta trasera y Resevados	Resultado
44	8:02:45	21045	59	35	35	129
45	8:04:07	21003	72	46	42	160
46	8:05:30	21014	52	38	20	110

Fuente: Elaboración propia

Programación de las salidas de los buses

Los datos brindados por el área de programación, son los despachos que se realizan en hora pico de la mañana entre las 7 a 8 am (Ver Tabla 8), en esta información indica que la cantidad de salidas o despachos son de 398 en la hora pico

Tabla 8 Despachos en Hora pico Mañana

Estación de salida	Cantidad de buses	Ruta
Naranjal	285	A,D,SX,XN, Exp.2, Exp.3
Izaguirre	26	Exp.6
Tomas Valle	35	Exp.7
Matellini	52	Exp.1, Exp.C
Total	398	

Fuente: Información facilitada por Protransporte

Se observa en la Tabla 9 los tiempos por ciclos que duran los viajes, es decir es el tiempo que tarda en dar una vuelta de recorrido, esta información también fue brindada por la gerencia de programación de Protransporte.

Tabla 9 Tiempo por ciclo

Viaje	Tiempo por ciclo
Chinchaysuyo	160
Los Alisos – El Angel	150
Honorio Delgado	135
Habich – Caquetá acceso sur	160
Plaza Castilla	120
Angaraes – Roosevelt	120
Av. Alfonso Ugarte desde Quilca hasta Bolivia	140
España	160
Plaza de flores hasta Matellini	130
Ampliación sur	120

Fuente: Información facilitada por Protransporte

Los lugares donde se realizan los giros de los recorridos de las rutas se observan en la Tabla 10.

Tabla 10 Salidas en la Hora pico Mañana

Nº	Giros
1	Pasando la estación flores en Miraflores
2	Pasando la estación Tomas Valle en Independencia
3	Altura la estación Naranjal, en al Av. Alisos
4	Estación Naranjal, en la Av. Chinchaisuyo

Fuente: Información facilitada por Protransporte

Una muestra del listado de horarios de salida de los buses se puede observar en la Tabla 11. La cantidad de buses destinadas a cubrir las rutas de aprecia en la Tabla 12, el Metropolitano cuenta con 300 buses hasta el año 2019, los cuales están destinados a ciertas rutas.

Tabla 11 Horarios de salida de buses

Nº	Bus	Ruta	Salida progr.	Serv.
1	11022	C	04:22:00	1104
2	11016	C	04:32:00	1107
3	T010	Exp2	04:38:00	2200
⋮		⋮	⋮	⋮
300	41049	X6	07:39:00	2612

Fuente: Elaboración propia

Tabla 12 Salidas en la Hora pico Mañana

Buses	Cantidad
Expreso 1	36
Expreso 2	48
Expreso 3	32
Expreso 6	18
Expreso 7	20
Expreso 9	4
Regular A	17
Regular C	42
Regular D	18
Super Expreso	18
Super Expreso Norte	30
R1- Lampa	6
Exp financ.	11
Total	300

Fuente: Elaboración propia

4.2.3 Paso 3: Data Analytics (Pre procesamiento)

4.2.3.1 Descripción

En este paso se hace un análisis general de la data recolectada, es decir tener una idea de todos los parámetros que tiene las diferentes fuentes de datos, dentro de este paso se encuentran dos tareas (1) Exploración: se realiza un análisis de todos los parámetros de los datos mostrando la estadística más relevante para el caso de estudio. (2) Verificación: Se hará una verificación de los datos sean coherentes, si existen valores que faltan, los intervalos de los datos, tipo de datos, si existen datos distorsionados “ruido”, datos que no serán necesarios para nuestros objetivos, etc. El pre procesamiento se realiza para seleccionar que parámetros serán usados en el modelo de generación de rutas y predicción de rutas.

4.2.3.2 Caso Metropolitano

Con los datos del Metropolitano se realizará el pre procesamiento para luego ser usados en el modelo, para ello hará la exploración y verificación de los siguientes datos:

Validaciones de los usuarios

Exploración: Se explora de manera general todos los atributos de los datos con la finalidad de tener un mejor entendimiento de estos. Para ello se analizará los atributos “fecha”, “tipo de validación”, “Parada”, “tipo de validador”, “perfil” y “monto” ya que son estos los que nos darán mayor información.

Del pre procesamiento de la data se encontró que hubo 18 720 760 entradas o validaciones al Metropolitano durante el mes mayo del 2018, el promedio de entradas diarias es de 60 3895.5, entre la vías troncales y alimentadoras. Desglosando el promedio de validaciones durante la semana se observa que el promedio de lunes a viernes es de 710 657.9, el de sábado es de 499 657 y los domingos y feriados con un valor de 217 531.8; los cuales se observan en la Tabla 13.

Tabla 13 Cantidad de usuarios

USUARIOS	PROMEDIO DIARIO	PROMEDIO DE LUNES A VIERNES	PROMEDIO SABADOS	PROMEDIO FERIADO Y DOMINGOS	MES
Alimentador	189184.2	217490.7	162180.5	86238.6	5864711
Troncal	414711.3	493167.1	337476.5	131293.2	12856049
Total	603871.5	710657.9	499657.0	217531.8	18720760

Fuente: Elaboración propia

La Tabla 14 muestra los días mínimos y máximos del mes de mayo del 2018, es decir los días donde existe menor y mayor cantidad de usuarios; para esto, se buscó los mínimos y máximos dando como resultado el día con menor demanda el día domingo 13 de mayo del 2018, y el día con mayor demanda el día 9 de mayo del 2018. Luego también se buscó los días con mayor y menor demanda dependiendo de la ruta, dando como resultado que el día con menor demanda en ambas rutas es el día domingo 13 de mayo del 2018, en la ruta alimentador el día con mayor demanda es el día miércoles 2 de mayo del 2018 y en la ruta troncal el día con mayor demanda es el día viernes 4 de mayo del 2018.

Tabla 14 Días mínimos y máximos

USUARIOS	MÍNIMO DIARIO	DÍA MÍNIMO	MÁXIMO DIARIO	DÍA MÁXIMO
Alimentador	83643	Domingo 13-05-2018	221968	Miércoles 02-05-2018
Troncal	109814	Domingo 13-05-2019	499808	Viernes 04-05-2018
Fecha	193457	Domingo 13-05-2019	718343	Miercoles 09-05-2018

Fuente: Elaboración propia

De la Figura 35 a la Figura 40 se muestra como los atributos “tipo de validación”, “Parada”, “tipo perfil” y “monto” varían en función de la “fecha”. En estas figuras se observa que el comportamiento en bloque de lunes a domingo se repite durante el mes. Así como los días de lunes a viernes son los de mayor demanda y el de menor demanda es el domingo. El día feriado (1 de mayo) tiene un comportamiento similar al de un día domingo.

En la Figura 35 se visualiza el atributo “tipo de validación”, que está conformada por la ruta troncal y alimentador, además se observa que la ruta troncal tiene mayor demanda comparado con la ruta alimentadora.

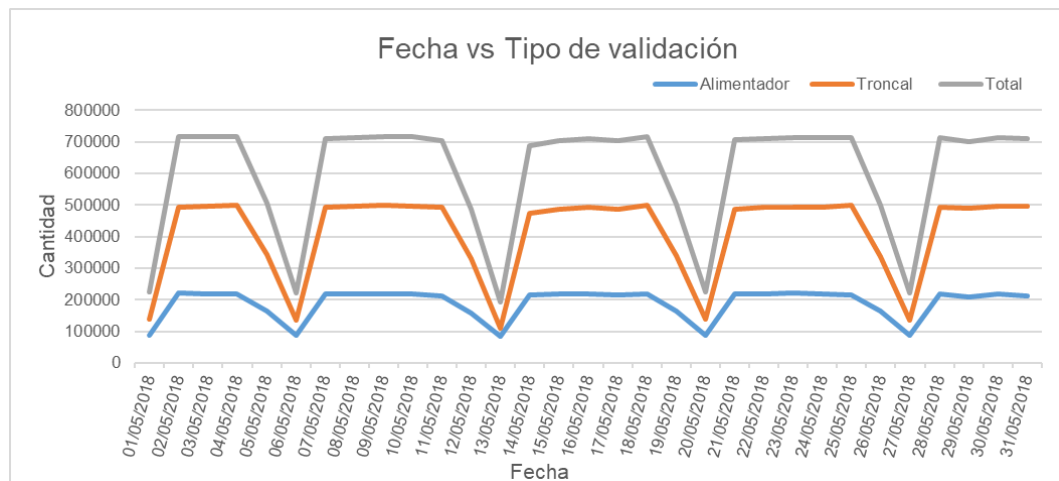
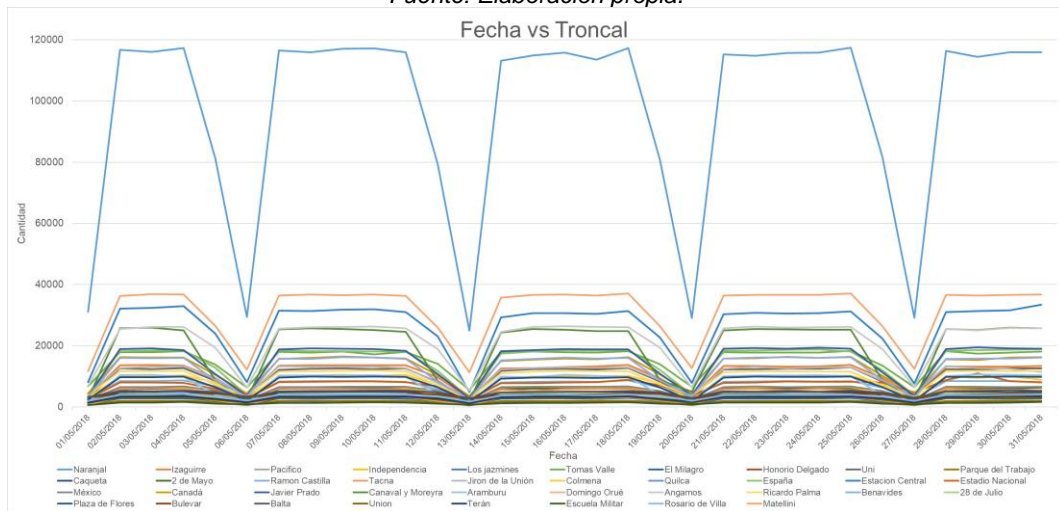
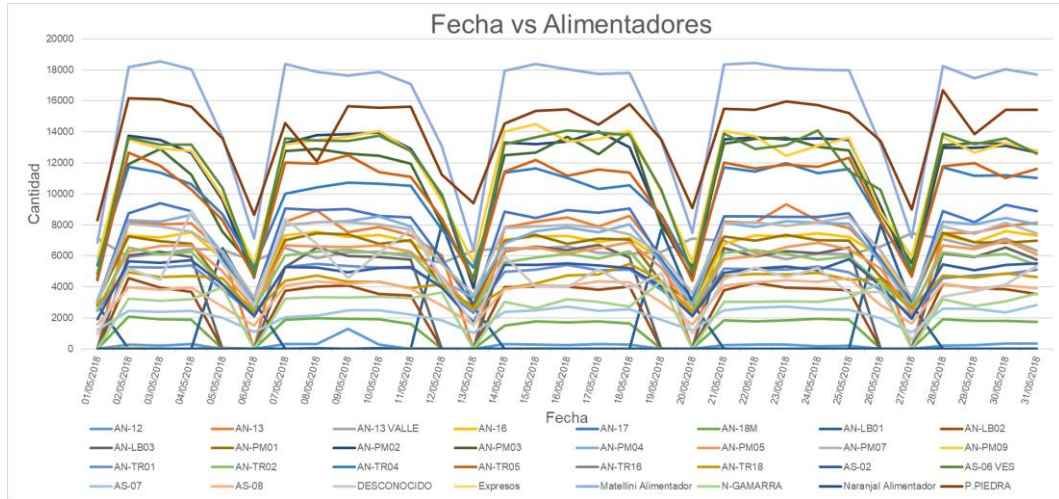


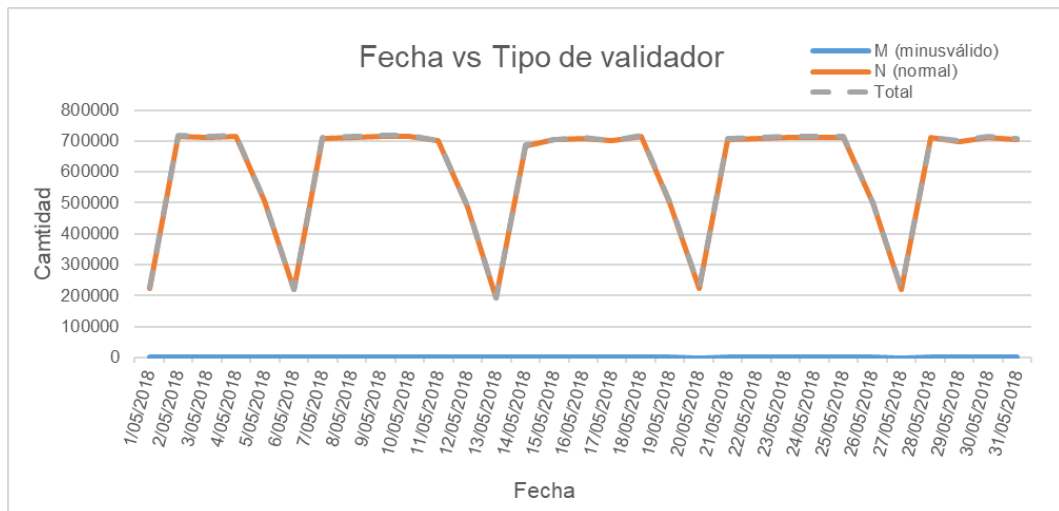
Figura 35 Tipo de validación

Fuente: Elaboración propia.

El atributo “parada” se separó según la ruta alimentador y troncal, en la Figura 36 y Figura 37 se ve que en la ruta alimentador la estación con mayor demanda es la estación “Matellini alimentador” y en la ruta troncal es la estación “Naranjal”.



La cantidad de validaciones en el atributo "tipo validador" minusválido (M) están entre 1031 y 2886, por eso no se visualiza de manera correcta en la Figura 38.



En la Figura 39 se muestra al atributo “perfil”, el comportamiento de los perfiles tipo escolar (ES), general (GE), pase libre (PL) y universitario (UN) en un mes.

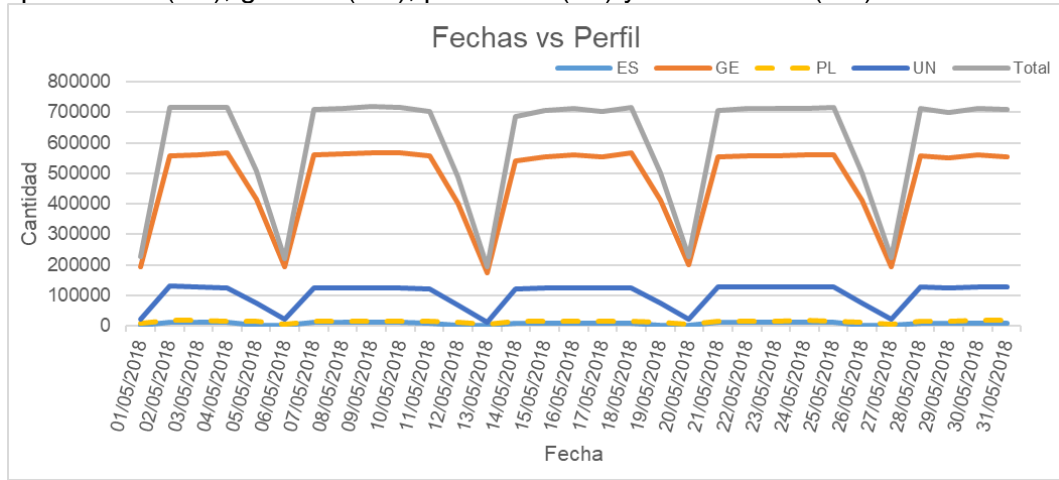


Figura 39 Perfil

Fuente: Elaboración propia.

En el atributo “monto”, los pasajes varían desde 0 hasta 2.5 soles, el monto con mayor recaudación es de 2.5 (Figura 40); el promedio diario recaudado de lunes a viernes bordea los 1.18 millones de soles (Figura 41 y Tabla 15). Del analisis se calculó que el pasaje promedio diario por persona es de S/1.6 y que el monto no recaudado por los usuarios cuyo perfil es PL asiende a S/22 508 por día (14221usuarios), el cual representa el 2.35% de los usuarios totales en un día.

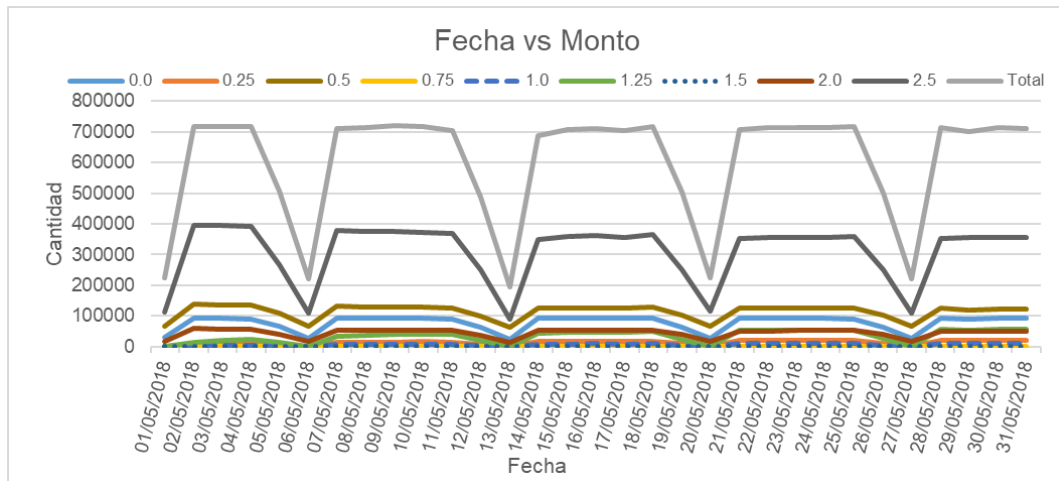


Figura 40 Monto

Fuente: Elaboración propia

Tabla 15 Promedios “Mono”

Monto	
Mínimo diario	S/ 286.336,5
Máximo diario	S/ 1.182.234,5
Promedio L-V	S/ 1.097.424,7
Promedio S	S/ 776.845,6
Promedio D y feriado	S/ 336.037,8
Promedio diario	S/ 933.255,3
Recaudación del mes	S/ 28.930.913,8

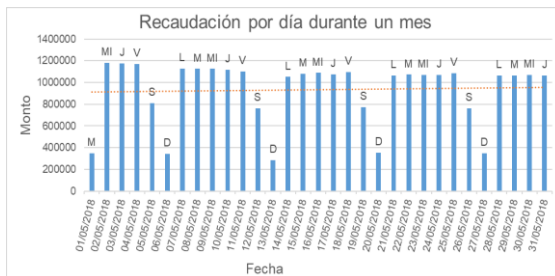


Figura 41 Recaudación “Monto”

Fuente: Elaboración propia.

Fuente: Elaboración propia.

Se elaboro un diagrama con los porcentajes que representa cada uno de los atributos con respecto del total de la data, por ejemplo, el atributo “tipo validador” se reparte en la ruta alimentador con un 31.33% y la ruta troncal con un 68.67% del total (Ver Figura 42).

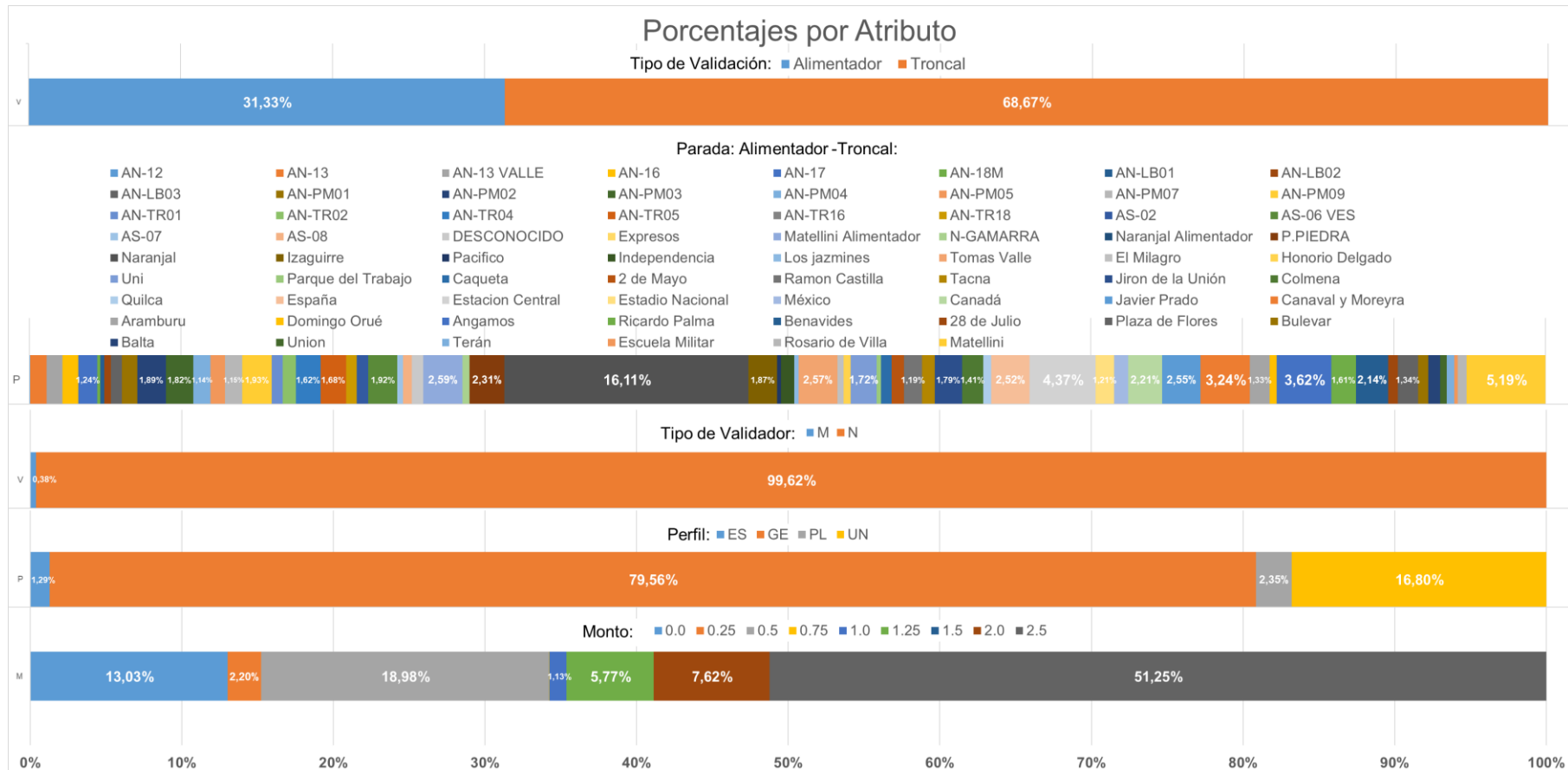


Figura 42 Porcentajes
Fuente: Elaboración propia

Verificación: Servirá para tener los criterios cuando se realice la limpieza de los datos, así el modelo dará resultados confiables.

Del análisis realizado en la exploración se observa lo siguiente

1. El parámetro “Secuencia” depende del “Validador” donde se cuenta las “Tarjetas”, el “Validador” depende de la “Parada” en el que se encuentre; el “Archivo” depende de la “Parada” donde se extrajo la información. De lo mencionado el parámetro que representa a “Secuencia”, “Validador”, “Archivo” es la “Parada” debido a que todos estos parámetros indican las coordenadas de la estación.
2. El horario de atención del Metropolitano esta entre las 5:00:00 am a 23:00:00 pm; como se observa en la Figura 43 existen datos fuera del horario de atención.

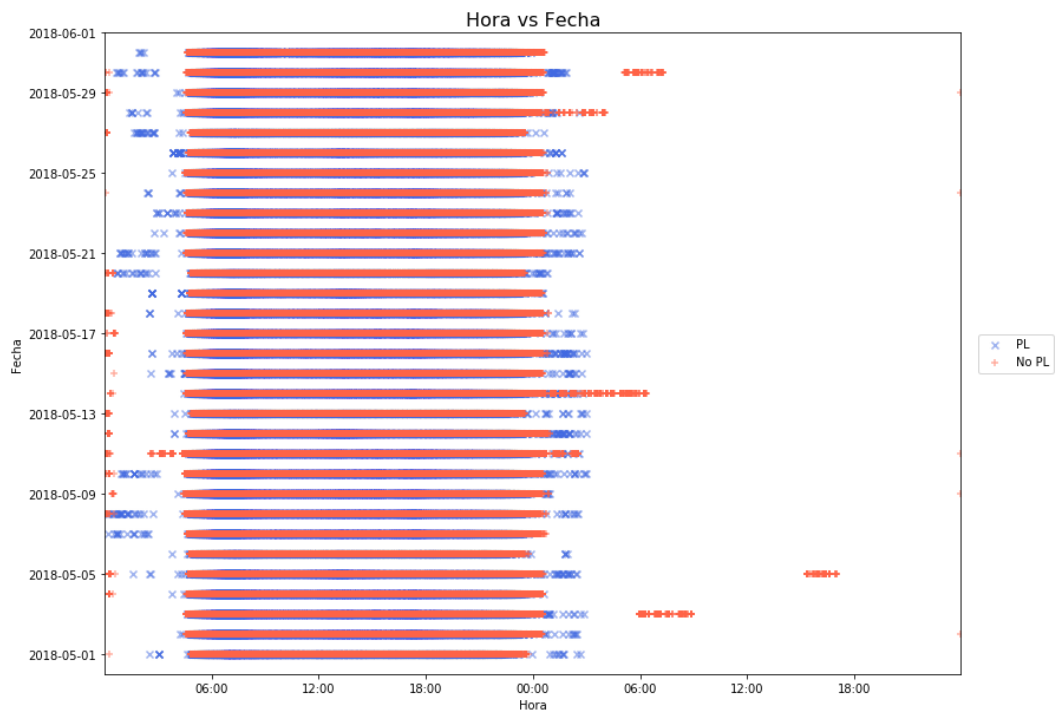


Figura 43 Validaciones según hora
 Fuente: Elaboración propia.

3. Para obtener un rango de trabajo adecuado, se realiza un filtrado por perfil tipo PL y no PL, debido a que el perfil tipo PL es usado por los trabajadores para hacer ingresar a las personas que no pagan pasaje, así como les sirve a los trabajadores para movilizarse dentro de las estaciones o para verificar si los validadores están funcionando correctamente. Como se observa el Figura 44, los valores de los validadores que pertenece a la vía Troncal y que es no es PL tienen

como valor mínimo a 04:43:00 y como valor máximo a 00:15:00 indicadas por el color verde.

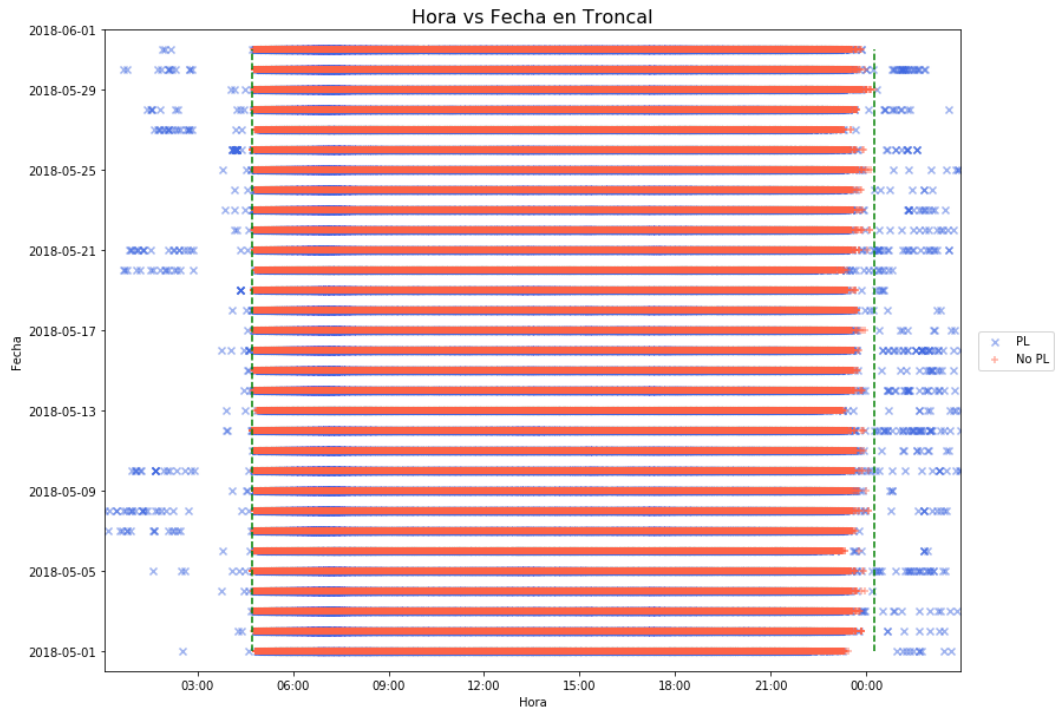


Figura 44 Validaciones en Troncal
 Fuente: Elaboración propia.

- En las Figura 45 y Figura 46 se ve que existen datos fuera del horario de atención. Para tener una mejor visualización de los datos en las Figura 45 y Figura 46 se separa los datos entre las 00:20:00 a 4:40:00 de la madrugada y las 4:40:00 a 00:20:00 del día. Este análisis servirá para el paso de limpieza de datos ya que si no se elimina los datos de la madrugada generaría "ruido" lo cual haría que resultados sean erróneos.

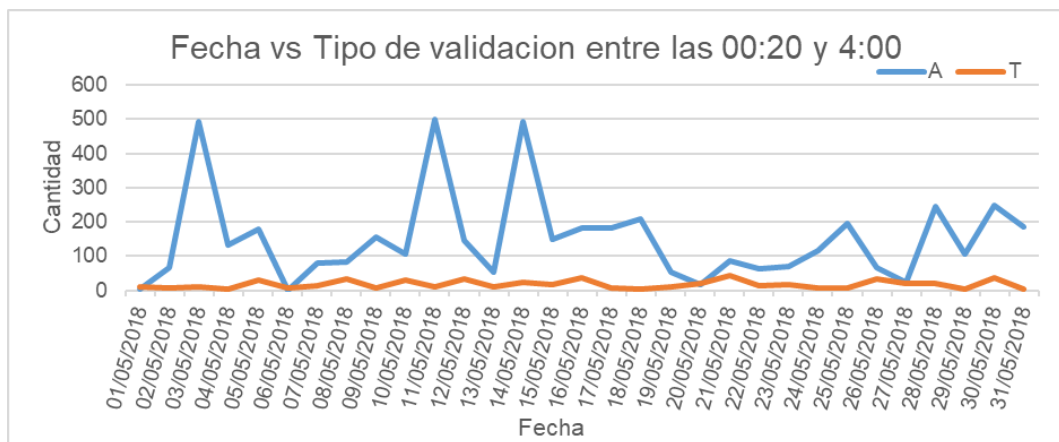


Figura 45 Validaciones entre las 0:20:00 a 4:40:00 am madrugada
 Fuente: Elaboración propia.

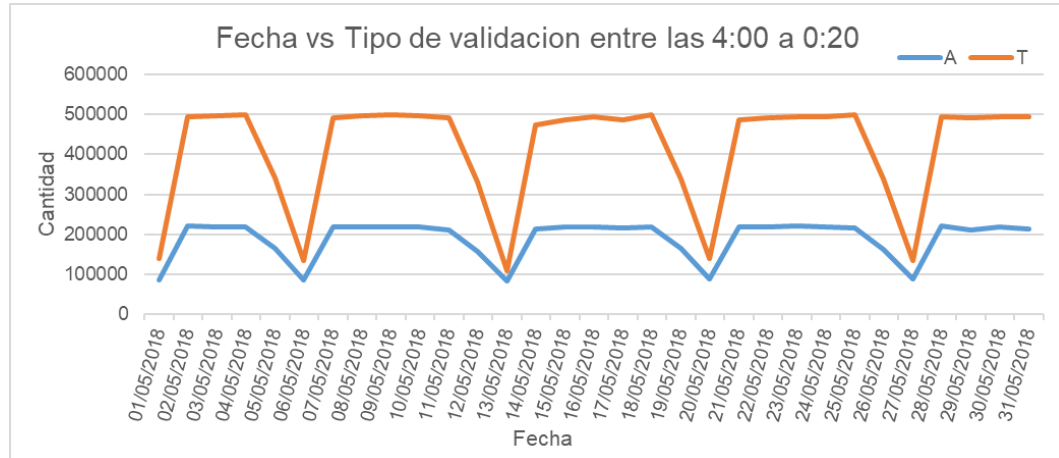


Figura 46 Validaciones entre las 4:40:00 a 0:20:00 am día
Fuente: Elaboración propia.

- Los atributos “Fecha”, “Tarjeta”, “Parada”, “Tipo de validador”, “Tipo de validación”, “Perfil” y “Monto” son los más significativos para ser analizados; sin embargo, para enfocarnos en el objetivo que es la generación de rutas se utilizara solo los parámetros “Fecha”, “Tarjeta”, “Parada” con “Tipo de validación” solo en la troncal “T” serán necesarios para el modelo de generación de rutas. Así mismo de acuerdo al pre análisis la data se centra mejor entre las 4:40 am de la mañana hasta las 00:20:00 de la noche realizadas en el día que corresponda, estableciendo como nuevo horario de atención.

Ruta de recorrido

El modelo de “Generación de rutas óptimas” no solo implica las validaciones de los usuarios, sino que también de otras fuentes ya mencionadas, es por ello que se analizara la vía Troncal y no la vía Alimentadoras, ya que el nivel de complejidad aumentaría. Otra característica muy importante para desarrollar el modelo es que el recorrido de la vía sea exclusivo, el cual hace que el tiempo de viaje solo dependa del tiempo espera de los usuarios dentro de las estaciones, en tal sentido la vía troncal es la que tiene vía exclusiva. La ruta recorrida de la red troncal se puede ver en la Figura 34.

Coordenadas de las estaciones

De las 38 coordenadas, las estaciones Jirón de la Unión y Plaza de Flores cuentan con dos edificaciones de ingreso, para el modelo se considera que cada estación tiene una edificación de ingreso y salida de usuarios, es por ello que las coordenadas de las edificaciones de estas estaciones se promediarán para cada estación.

En la Tabla 16 se muestran las coordenadas de las estaciones Jirón de la Unión y Plaza de Flores, obteniendo los valores promedio con los cuales se hacen los cálculos. El listado de todas las coordenadas de las 38 estaciones está en el ANEXO 2.

Tabla 16 Coordenadas de la estación Jr. De la Unión y Plaza Flores

Estación		Latitud	Longitud
Jirón de la Unión	Estación 1	-12,04847	-77,03408
	Estación 2	-12,04934	-77,03269
	Promedio	-12,04891	-77,03338
Plaza Flores	Estación 1	-12,13609	-77,01876
	Estación 2	-12,13571	-77,01870
	Promedio	-12,13590	-77,01873

Fuente: Elaboración propia.

Velocidad de los buses

Los BRT tienen como velocidades promedio entre 20 y 30 km/h, los datos por Protransporte indican que los servicios regulares circulan con una velocidad promedio de 20 km/h y los servicios expresos con 33 km/h, y los valores tomados en campo varían entre 20 a 45km/h. Estos valores tienen un rango que dependen del tramo en el que está circulando el bus, el valor ingresado en el modelo de "Generación de rutas óptimas" debe ser una velocidad que refleje la realidad. Para obtener la velocidad estándar que se ingresó al modelo se realizó la Macro simulación situación actual sin considerar los semáforos, el cual se puede observar en la página 106, luego de varias corridas se calibró la velocidad, dando como mejor valor 30km/h lo que se asemeja al comportamiento de realidad.

Capacidad de los buses

El modelo utilizará como capacidad máxima de 160 pasajeros, es decir se considera como un valor máximo para todos los buses, este dato se basa en la información brindada por Protransporte y fue validada en campo, esta validación de capacidad máxima del bus se puede ver en el ANEXO 4

Programación de las salidas de los buses

Los datos brindados por Protransporte y los datos recolectados en campo sirvió para comparar la programación de salidas de buses del Metropolitano. Los horarios de salidas brindados fueron validados, es decir la información se verificó de manera correcta. Los horarios de salida de los buses se utilizaron en la simulación de la situación actual del Metropolitano, estas salidas de buses son comparadas con los horarios de salidas del modelo de "Generación de rutas óptimas".

A continuación, se muestra un resumen de los datos que intervinieron en este paso 3 de pre procesamiento:

- Validaciones de los usuarios, requiere de procesamiento de los datos con la finalidad obtener los orígenes y destinos de los usuarios para utilizarlo en el modelo de generación de rutas.
- Ruta de recorrido, se utilizará la vía Troncal
- Coordenadas de las estaciones, se utilizará las coordenadas adaptadas.
- Velocidad de los buses, en el modelo usará la velocidad de 30 km/h
- Capacidad de los buses, el modelo usará como capacidad máxima 160 usuarios.
- Programación de las salidas de los buses, los datos de los horarios de salida de los buses del Metropolitano están listos para ser comparados con el modelo de generación de rutas.

Una vez realizado el pre procesamiento se redefine las metas de la Ciencia de Datos:

- Mejorar las rutas en base a la demanda de usuarios de tal manera que se satisfaga a la mayor cantidad usuarios
- Disminuir el tiempo de espera de los usuarios como mínimo en un 50% en las estaciones del Metropolitano.

Para lograr estas metas, el modelo de “Generación de rutas óptimas” debe obtener las rutas basándose en los orígenes destinos más demandados, de tal manera que satisfaga a la mayor cantidad de usuarios, es decir el modelo debe tener como entrada a los orígenes destinos de los usuarios. Así mismo el modelo debe controlar la variable tiempo de espera de los usuarios en las estaciones, lo que implica que el usuario tiene un tiempo máximo de espera en cada estación para poder embarcarse. Además para que el modelo logre cumplir estas metas debe de considerar las variables de la red de transporte, es decir la velocidad con la circula el bus, la capacidad del bus y los paraderos. La velocidad del bus implica el tiempo de viaje de los usuarios, con la capacidad del bus ayudara a la eficiencia de la ruta y la cantidad de buses a utilizar, los paraderos y los orígenes destino involucra a las paradas que realizara una ruta a lo largo de su recorrido.

4.2.4 Paso 4: Limpieza de los datos

4.2.4.1 Descripción

En este paso se elimina o separa los datos alterados de acuerdo con el pre procesamiento de los datos.

4.2.4.2 Caso Metropolitano

De acuerdo con el pre procesamiento de los datos visto en el paso anterior, los datos de las validaciones de las tarjetas se debe procesar debido a que aún no se cuenta con los orígenes y destinos de los usuarios que son necesarios para el modelo, sin embargo las otras fuentes de datos como ruta de recorrido, coordenadas de las estaciones, tiempo de semáforo, velocidad de los buses, capacidad de los buses y programación de las salidas de los buses ya están listas para ser utilizadas en el modelo de generación de rutas. Es por ello que en los pasos 4, 5, 6 y 7 se procesará los datos de la validación de las tarjetas.

Para realizar la limpieza de los datos de las validaciones se debe tener las siguientes consideraciones:

1. En el pre procesamiento, de los datos de validaciones, se vio que existe data inválida durante la noche, esto es debido a que estas validaciones son realizadas por el personal de servicio del Metropolitano. Se ve que las validaciones están fuera del intervalo de atención al cliente (horario de atención del metropolitano de 5 am a 11 pm)
2. Del análisis de dispersión de los datos al inicio y final del horario de atención al cliente que se realizó en el pre procesamiento, en la Figura 43 se ve que la mayor cantidad de datos está en el intervalo de las 04:40:00 am de la mañana a las 00:20:00 am del día siguiente (noche).

En base a las consideraciones antes mencionadas se eliminará los datos que están en el intervalo de 00:20:00 a 04:40:00 según la fecha del archivo. La data del mes de mayo tiene 18720760 (100%) filas o instancias, realizando la limpieza de los datos queda 18715521 (99.97%) instancias.

Tabla 17 Limpieza de datos

	Data inicial	Data eliminanda	Data restante
Porcentaje	100.00%	0.03%	99.97%
Validaciones	18720760	5239	18715521

Fuente: Elaboración propia.

4.2.5 Paso 5: Filtrado o selección de parámetros

4.2.5.1 Descripción

Se seleccionan los parámetros y el intervalo de los datos que se usaran en el modelo; ello se realiza con la finalidad que el modelo utilice datos coherentes y válidos.

4.2.5.2 Caso Metropolitano

Para el filtrado se seleccionará los datos en un intervalo de tiempo del día y como se explicó en los pasos anteriores el caso de estudio se limitará a vía Troncal. Es importante mencionar que la selección, se realizará a todos los archivos del mes. La forma como se seleccionará las validaciones son las siguientes:

1. Filtrar el parámetro “tipo de validación” y solo seleccionar las validaciones que tienen como vía la Troncal “T”.
2. Filtrar las tarjetas diferentes al tipo de perfil “PL”, debido a que estas tarjetas permanecen estáticas en cada estación no se podrá saber el Origen – Destino de estas tarjetas.
3. Filtrar los atributos “fecha”, “tarjeta” y “parada”, ya que intervienen directamente en el modelo de “Generación de rutas óptimas”.

En el proceso de filtrado se separó 6 297 290 de la data restante después de realizar la limpieza de datos, la data separada son validaciones de la ruta alimentador y el tipo de perfil tipo “PL”.

Tabla 18 Validaciones filtradas

	Validaciones	Porcentaje
Data inicial	18720760	100.00%
Data eliminanda por limpieza	5239	0.03%
Data separada por filtrado	6297290	33.64%
Data filtrada	12423470	66.36%

Fuente: Elaboración propia.

Al filtrar, la data queda reducida a 12 423 470 (66.36%) instancias. Quedando a estructurar los parámetros “fecha”, “tarjetas”, “perfil” y “parada” en el intervalo de 04:40:00 am de la mañana del día x y 00:19:59 am de la madrugada del día x+1 con validaciones de usuarios que utilizan la vía Troncal.

Tabla 19 Filtrado de datos

FECHA	TARJETA	PERFIL	TIPO_VALIDACION	PARADA
Del 1 al 31 de mayo del 2018	12 423 470 validaciones	GE, ES, UN	Troncal	38 estaciones

Fuente: Elaboración propia.

4.2.6 Paso 6: Estructura de los datos

4.2.6.1 Descripción

En este paso lo que se hace es armar o estructurar los datos con los parámetros filtrados que se realizó en el paso 5; la estructura depende de lo que se quiere obtener de los datos. En este paso también se debe integrar la fuente de datos que tengan los mismos parámetros y colocarlo en el mismo formato, por ejemplo, si se tiene los orígenes destinos del tren, BRT y metro se tiene que integrar todos estos datos con el mismo tipo de formato.

4.2.6.2 Caso Metropolitano

En el caso del Metropolitano los datos de los validadores solo capturan los orígenes y no de orígenes-destinos de los usuarios, para el modelo se va a necesitar los orígenes-destinos de los usuarios, para lograr ello se hará las siguientes suposiciones:

- La 1er validación se considera como origen, la 2da validación como el destino formando así primer par de origen-destino realizada por el usuario m; luego la 2da validación se considera como origen, la 3da validación como el destino formando así el segundo par de origen-destino realizada por el usuario "m"; así sucesivamente hasta la n enésima validación considerando como origen y 1er validación se considera como destino del enésimo par de origen destino.
- Cuando las tarjetas que son validadas más de una vez en la misma estación es el un intervalo de tiempo pequeño, es decir cuando con una misma tarjeta ingresan más de una persona a la estación; los pares orígenes destinos se generan dependiendo de la cantidad de veces que marco en la primera validación y las siguientes validaciones. los pares orígenes destinos de generan de la siguiente manera: En la primera validación tiene "x" marcaciones, en la segunda validación "y", en el primer par origen destino se genera "z" viajes, donde "z" es el menor valor entre "x" y "y", con "z" pares orígenes destinos para la estructura de los datos. Por ejemplo, si la tarjeta tiene 5 marcaciones en la primera validación y 2 marcaciones en la segunda validación, se considera solo dos pares orígenes destinos y no cinco.

La manera en la que se estructura el origen-destino de los viajes con los parámetros seleccionados es de la siguiente manera:

1. La estructuración se elabora con los datos limpiados y filtrados.
2. Se procede a identificar los usuarios que tienen el mismo código en el parámetro “tarjeta” para luego armar la estructura haciendo la siguiente programación.
 - Se tomará en consideración las validaciones entre los horarios de 04:40:00 am de la mañana y 00:19:59 de la noche. Se considera al “archivo del día i” al archivo generado el día “i” el cual está compuesto por un intervalo desde las 04:40:00 am hasta 23:59:59 que pertenece al día “i” y el otro intervalo desde las 0:00:00 hasta las 0:19:59 pertenecen al día “i+1”.
 - Se genera una matriz con los siguientes campos “tarjeta”, “fecha”, “hora” y “parada”, es decir el parámetro “fecha” se convierte en dos parámetros que son “fecha” y “hora”
 - Luego se procede a buscar los códigos de identificación (ID) del parámetro “tarjetas” que se repiten durante el “archivo del día i”
 - Se considera n la cantidad de validaciones (cantidad de viaje) y se considera como primer viaje a la validación j como origen y la validación j+1 como destino, segundo viaje a la validación j+1 como origen y validación j+2 destino y de la misma manera se hace para los siguientes viajes teniendo como viaje n a la validación j+ (n-1) de origen y validación j como destino.
 - Teniendo la consideración anterior se elabora una matriz con los siguientes campos de “tarjeta”, “fecha” de la validación origen, “hora” de la validación origen, “parada origen” y “parada destino”.
3. La data estructurada tendrá 4 parámetros “tarjeta”, “fecha”, “hora”, “parada origen” y “parada destino”.

En la Tabla 20 se observa como quedan las validaciones, después de ser estructuradas, teniendo como nuevos parámetros, “fecha”, “hora”, “origen” y “destino”, también se observa el comportamiento de la “tarjeta” 3058796891.

Tabla 20 Estructura de datos

Fecha	Hora	Tarjeta	Origen	Destino
4/05/2018	05:01:01	3058796891	Izaguirre	Estacion Central
4/05/2018	05:28:40	3058796891	Estacion Central	Matellini
4/05/2018	16:32:30	3058796891	Matellini	Izaguirre

Fuente: Elaboración propia.

4.2.7 Paso 7: Análisis de los datos

4.2.7.1 Descripción

El análisis es un paso similar al pre procesamiento, pero con los datos listos para ser utilizados en el modelo de generación de rutas. Este paso contribuirá en el obtener unos resultados más precisos y no de manera general como en el pre procesamiento, en otras palabras, es un análisis direccionado para el modelo.

4.2.7.2 Caso Metropolitano

Para realizar el análisis basado en el origen-destino de los usuarios evaluaremos lo siguiente: matriz origen destino, demanda por estación, fluencia de los viajes y distribución en función del tiempo de la data de un día. Después este análisis se elabora la distribución de la data en la hora pico, demanda por estación, flujo de los viajes y una gráfica de origen destino respecto de la estación naranjal. Todos estos análisis ayudan a formular el modelo de “Generación de rutas óptimas”, los resultados de este modelo deben de concordar con este análisis, así como los orígenes y destinos de; así ver el comportamiento.

1. La matriz origen-destino de un día, es el resultado de los pares orígenes-destinos de los usuarios formulado a partir de paso 6 (estructura de datos). En la Figura 47 se observa los datos del día 4 de mayo del 2018, luego de estructurar la data se obtuvo 304 705 validaciones durante ese día, debido a la forma como se estructuró los orígenes destino la matriz del día tiene la característica que la suma de la fila “n” es igual a suma de la comuna “n”, con una diagonal igual a cero debido ya que los orígenes son diferentes a los destinos. La matriz se ha dividido en 4 intervalos así se visualiza que orígenes-destinos son más demandados, el primer intervalo [0,100] está sin colorear, el segundo intervalo <100,1000] de color verde, el tercer intervalo <1000,5000] de color amarillo y el cuarto intervalo <5000,10000] de color rojo. Por ejemplo en el caso del primer intervalo la cantidad de usuarios que viajan desde la estación el “Independencia” a la estación “Milagro” es cero, en el segundo intervalo desde la estación “Tomas Valle” a “Matellini” existen 662 viajes, en la tercer intervalo de la estación “Uni” a “Canaval y Moreyra” hay 1 108 viajes y en el cuarto intervalo “Naranjal” a “Estación Central” se encuentra 8 288 viajes; estos intervalos sirve para tener un criterio de las paradas de las rutas de los buses, es decir tiene un recorrido de ruta considerando la mayor demanda de viajes generando rutas directas.

La estación Naranjal es la que tiene varios pares orígenes destino de color amarillo a rojo, lo que indica una mayor interacción de esta estación con las demás estaciones. También se observa interacción de las estaciones “Matellini”, Estación Central” “Canaval y Moreyra” y “Angamos”, ya que los pares orígenes destinos están coloreados desde el color verde hasta el rojo esto se ve reflejado en el histograma de la Figura 48. Por otro lado, que la interacción entre dos estaciones sea cero (“Independencia” hacia “Milagro”), no quiere decir que en la estación de origen o de destino el flujo de personas durante el día sea cero, por ejemplo, en la estación “Independencia” la cantidad de flujo de personas es de 3410 durante el día, esto es debido a la interacción con las otras estaciones.

MATRIZ DE ESTACIONES		DESTINO																																					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
ORIGEN	1 Naranjal	0	60	8	18	11	128	44	162	434	175	273	1190	857	4004	580	1117	3228	2385	8288	1457	1445	4878	5227	7314	3506	486	7735	3520	4819	913	1626	679	500	318	160	318	494	3230
	2 Izaquire	68	0	1	7	1	20	7	15	29	16	29	62	64	321	77	151	328	279	742	159	167	374	618	947	303	78	838	322	576	119	230	76	62	46	18	48	56	377
	3 Pacifico	22	2	0	2	1	5	1	5	8	5	5	18	12	84	7	15	36	47	143	32	21	82	51	64	27	8	88	40	67	14	26	10	12	4	5	6	9	65
	4 Independencia	17	4	3	0	3	5	0	3	11	5	12	40	45	288	34	65	116	122	426	85	78	210	242	323	91	26	398	121	221	31	90	28	18	12	4	28	24	181
	5 Los jazmines	13	2	1	10	0	6	0	0	4	5	5	21	15	87	12	18	59	55	122	22	24	64	69	81	24	9	74	26	35	14	17	7	5	2	2	3	7	53
	6 Tomas Valle	197	26	1	6	4	0	7	2	17	12	18	44	57	328	62	128	384	238	830	204	171	445	867	1515	303	116	1259	429	631	208	354	125	87	64	32	68	113	662
	7 El Milagro	49	8	1	0	0	7	0	1	4	1	5	16	19	77	11	22	75	49	150	55	27	79	112	169	50	14	145	73	110	32	43	17	13	5	8	8	7	115
	8 Honorio Delgado	191	13	4	3	0	8	3	0	13	0	4	11	29	153	12	43	74	47	204	31	42	107	126	114	52	18	171	50	75	16	62	19	22	11	3	18	8	181
	9 Uni	465	28	9	8	5	23	4	9	0	2	8	34	25	216	26	90	162	110	444	92	100	389	526	1108	222	94	932	345	629	131	307	100	88	36	22	61	74	707
	10 Parque del Trabajo	181	12	5	5	5	9	1	0	3	0	5	4	3	27	3	14	29	27	70	14	21	42	41	71	24	13	78	29	42	18	35	17	20	15	13	17	21	98
	11 Caqueta	280	33	6	10	6	19	5	5	8	3	0	2	2	27	8	20	41	35	109	20	25	85	137	322	86	33	267	104	153	63	135	39	37	34	13	29	44	204
	12 2 de Mayo	1212	63	20	37	21	51	17	14	34	4	7	0	2	4	11	3	2	5	115	12	23	57	109	245	59	39	195	70	117	39	96	77	40	38	6	39	30	483
	13 Quilca	850	61	13	42	17	56	17	27	23	4	1	4	0	12	3	6	8	48	11	20	43	61	115	27	21	112	49	73	23	55	43	27	28	7	33	30	280	
	14 España	3987	324	85	286	81	324	80	149	227	29	25	11	10	0	7	16	35	19	147	14	35	107	235	508	120	92	444	155	275	84	345	170	159	114	51	133	133	1429
	15 Ramon Castilla	568	65	6	32	14	61	5	13	26	2	9	9	3	6	0	10	64	65	146	37	82	178	221	489	145	100	383	175	267	93	142	140	121	74	35	85	85	888
	16 Tacna	1148	152	15	61	16	142	24	39	87	14	21	5	1	11	15	0	16	14	71	19	15	64	74	168	49	31	149	89	129	43	109	49	53	39	23	27	36	349
	17 Iiron de la Unión	3252	327	39	117	60	403	74	74	177	25	47	8	3	39	60	20	0	13	199	44	37	150	177	271	117	88	345	237	294	101	262	130	113	75	21	92	74	797
	18 Colmena	274	287	48	118	55	245	52	54	115	23	28	5	5	19	56	11	9	0	159	19	21	67	106	262	106	62	273	175	172	65	165	119	107	60	36	80	76	513
	19 Estacion Central	8746	778	143	444	121	927	146	186	475	65	100	86	35	120	164	63	137	143	0	130	116	296	436	669	211	133	710	371	487	182	657	274	241	234	79	260	241	2461
	20 Estadio Nacional	1398	157	30	79	24	203	50	31	102	11	19	16	6	16	42	18	50	28	161	0	7	35	52	83	37	39	140	108	46	125	119	74	101	29	95	108	1198	
	21 México	1413	160	23	74	23	165	27	44	97	23	29	27	21	34	65	15	39	21	124	8	0	12	31	68	31	18	75	79	91	32	79	68	39	53	14	57	52	562
	22 Canadá	4792	378	76	207	65	428	81	113	400	43	78	64	44	107	178	66	146	68	357	22	9	0	36	128	57	25	112	139	152	81	139	78	83	79	37	95	105	800
	23 Javier Prado	5117	612	55	238	65	841	113	127	516	41	138	106	63	229	216	81	208	113	530	50	31	43	0	80	47	37	172	148	184	79	248	94	83	75	33	66	100	931
	24 Canaval y Moreyra	7109	940	70	314	84	1515	175	131	1100	74	324	271	119	520	482	192	316	263	776	87	59	105	67	0	23	26	186	193	349	97	415	91	87	70	17	100	121	1425
	25 Aramburu	3492	299	26	92	26	303	50	54	221	29	86	64	33	117	145	52	130	108	222	34	28	61	42	19	0	3	26	25	42	32	52	25	32	25	9	33	51	526
	26 Domingo Orúe	479	72	7	26	9	116	15	18	88	14	35	38	19	90	99	31	98	63	134	44	21	24	42	35	4	0	9	12	23	15	14	19	15	10	9	16	18	187
	27 Angamos	7668	846	87	411	78	1257	148	175	956	79	266	186	113	455	378	148	353	259	762	134	79	102	168	173	29	2	0	22	43	32	79	54	49	71	24	67	110	726
	28 Ricardo Palma	3502	309	40	124	23	427	74	51	343	28	108	78	40	145	168	102	244	172	404	90	75	131	150	172	27	12	29	0	21	22	27	22	24	24	6	27	54	478
	29 Benavides	4825	581	74	226	41	646	107	82	619	43	157	120	74	275	265	122	310	179	491	113	83	145	182	313	50	21	45	11	0	5	25	25	19	17	7	22	52	439
	30 28 de Julio	875	114	14	30	15	210	33	17	132	19	64	44	24	88	90	46	107	61	197	47	31	83	71	107	24	18	32	14	13	0	11	4	7	13	0	7	20	166
	31 Plaza de Flores	1620	224	28	93	18	356	43	68	300	36	142	100	54	350	138	118	277	170	653	128	81	136	241	422	50	18	76	23	25	10	0	9	5	4	6	21	135	
	32 Balsa	666	80	9	27	7	115	20	19	98	15	35	81	48	169	140	53	125	105	271	124	64	75	92	91	29	17	49	19	23	4	20	0	10	6	7	15	38	212
	33 Bulevar	500	68	12	19	5	86	16	22	83	20	38	36	27	165	119	50	103	108	243	76	36	90	96	79	28	16	48	23	21	9	8	10	0	6	4	10	49	184
	34 Unión	315	47	6	13	3	64	4	11	35	19	33	39	30	114	72	38	77	61	246	95	53	79	79	71	26	10	68	25	20	12	8	1	2	0	2	1	28	105
	35 Escuela Militar	154	17	4	3	2	33	9	5	19	11	15	7	6	54	35	23	20	35	77	33	16	35	31	16	11	8	24	6	7	1	5	5	5	4	0	4	42	57
	36 Terán	323	52	5	26	4	63	8	20	56	17	25	43	28	130	82	29	88	83	270	94	56	105	68	106	33	15	66	24	19	5	8	11	9	7	3	0	4	16
	37 Rosario de Villa	501	59	9	25	6	115	8	7	76	24	44	28	30	128	83	30	75	83	240	112	48	94	102	117	50	17	107	49	48	20	20	37	47	32	35	5	0	49
	38 Matellini	3216	371	66	177	54	632	109	185	695	99	211	478	267	1436	859	344	797	506	2496	1189	566	807	935	1458	536	205	729	481	450	157	144	187	178	101	61	22	25	0
Total	304705	71585	7631	1049	3410	973	10014	1577	1938	7631	1032	2449	3396	2233	10445	4804	3367	8362	6147	21067	4937	3793	9868	11910	18293	6614	1968	16589	7773	10811	2848	6183	2978	2493	1908	839	2001	2560	21229

Figura 47 Matriz Origen-Destino del día 4/05/2018
Fuente: Elaboración propia.

2. La demanda por estación, es resultado del análisis de la cantidad de validaciones que hay en un determinado intervalo de tiempo, día, semanas, mes o año en cada una de las estaciones del Metropolitano, así determinar las estaciones más congestionadas. En la Figura 48 en base a la matriz origen destino del día 4 de mayo se calculó la cantidad de validaciones que interactúan en cada estación, estos valores se pueden observar en la parte inferior de la matriz origen destino. Por ejemplo, la cantidad de que interactúan en la estación “Naranjal” con las demás estaciones es de 71 585 validaciones. En la figura también se observa las 5 estaciones más concurridas, por ejemplo, la quinta estación más concurrida es la estación “Angamos” con 16 589” validaciones.

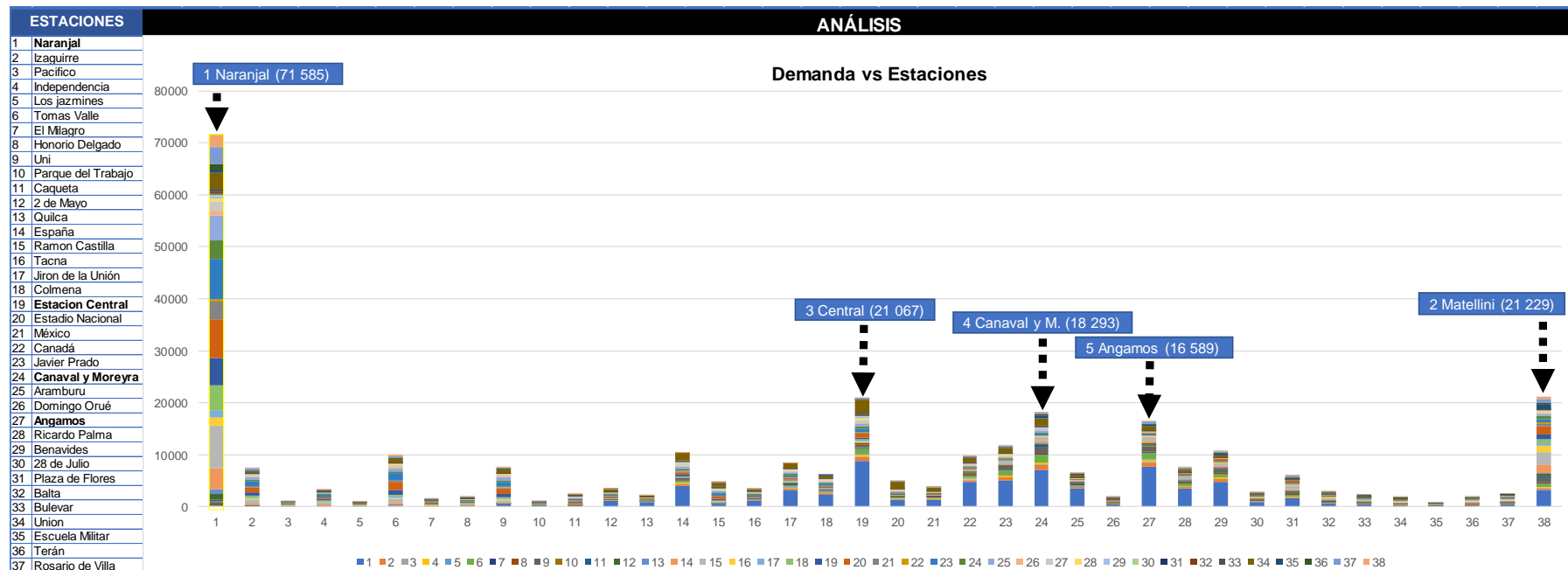


Figura 48 Demanda en cada estación del día 4/05/2018
 Fuente: Elaboración propia.

3. La distribución y fluencia de viajes, es el resultado de la matriz origen-destino de las validaciones del Metropolitano. En la Figura 49 existe mayor interacción de la estación “Naranjal” con las demás estaciones, también queda reflejado que los usuarios circulan de los extremos de la ciudad hacia el centro (desde “2 mayo” a “Plaza de flores”), sin embargo, es más del norte hacia el centro.

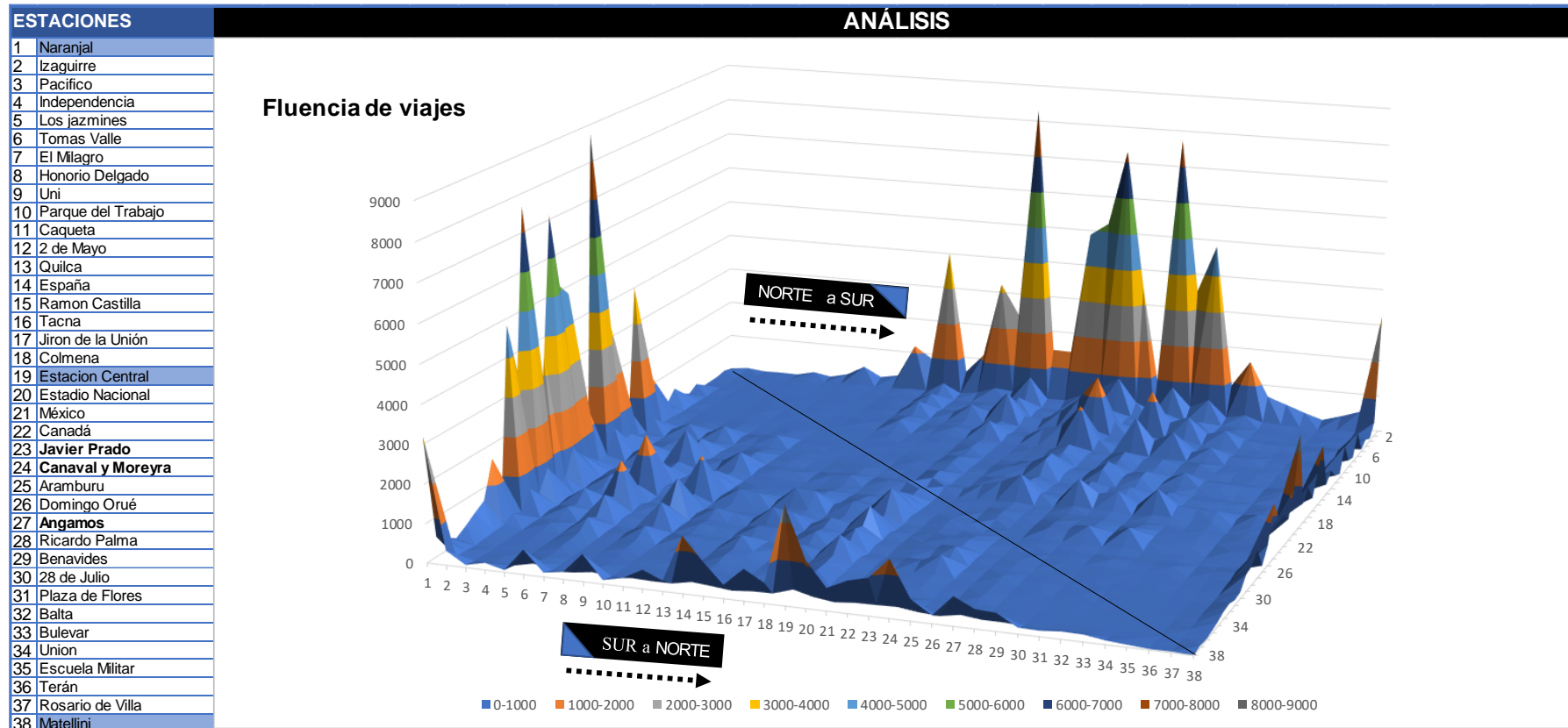


Figura 49 Fluencia de viajes del día 4/05/2018
Fuente: Elaboración propia

4. Luego del análisis de la matriz origen destino, se analizó el flujo de los usuarios durante el día, ello servirá para determinar el horario en el que se realizara la simulación, para esto se elaboró la Figura 50 donde se muestra la cantidad de validaciones en función del tiempo, el conteo comienza el horario durante las 4:20:00 y las 00:19:59 del día 4 de mayo del 2018, en este gráfico se ve que el horario pico está en la mañana en el intervalo [7:00:00, 8:00:00> con 37 514 validaciones, el horario pico en la noche está en el intervalo [18:00:00, 19:00:00> con 31 137 validaciones.

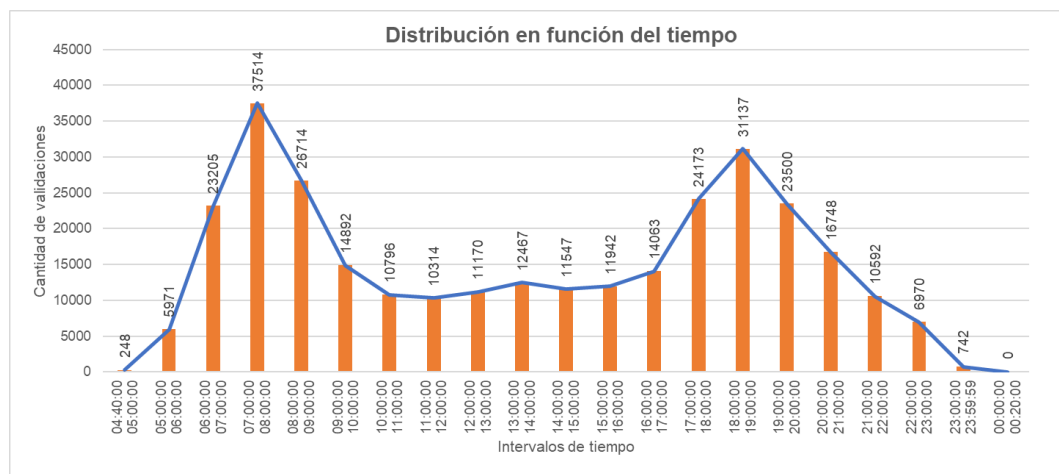


Figura 50 Distribución en función del tiempo del día 4/05/2018
Fuente: Elaboración propia.

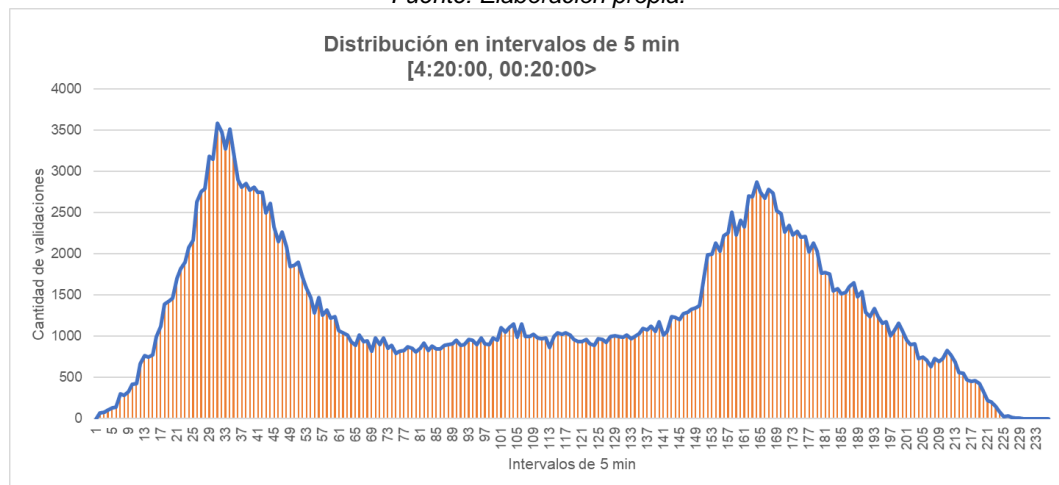


Figura 51 Distribución en intervalo de 5 min del día 4/05/2018
Fuente: Elaboración propia.

En las Figura 50 y Figura 51 se ve que las distribuciones, durante el día tiene 2 picos, a su vez que el intervalo de mayor demanda es en [7:00:00, 8:00:00]; en la Figura 50 el pico más alto se da en el intervalo de [7:10:00 a 7:15:00> con 3580 validaciones.

La cantidad de validaciones que existe en el intervalo [4:20:00, 0:20:00] es de 304 705 y la cantidad de validaciones en el intervalo [7:00:00,8:00:00] es de 37 518, esto quiere decir que la cantidad de validaciones en la hora pico representa el 12.31% del total validaciones del día 4 de mayo del 2018.

5. Teniendo el intervalo en la hora pico de [7:00:00,8:00:00] con 371518 validaciones, se analiza como es el comportamiento de estas con la Figura 52, teniendo como valor mínimo de 470 en el intervalo [7:59:00,8:00:00], valor máximo 748 en el intervalo [7:11:00,7:12:00], y promedio 625.3 validaciones por minuto; este último valor representa el 1.66% de las validaciones del intervalo de 7 a 8am.

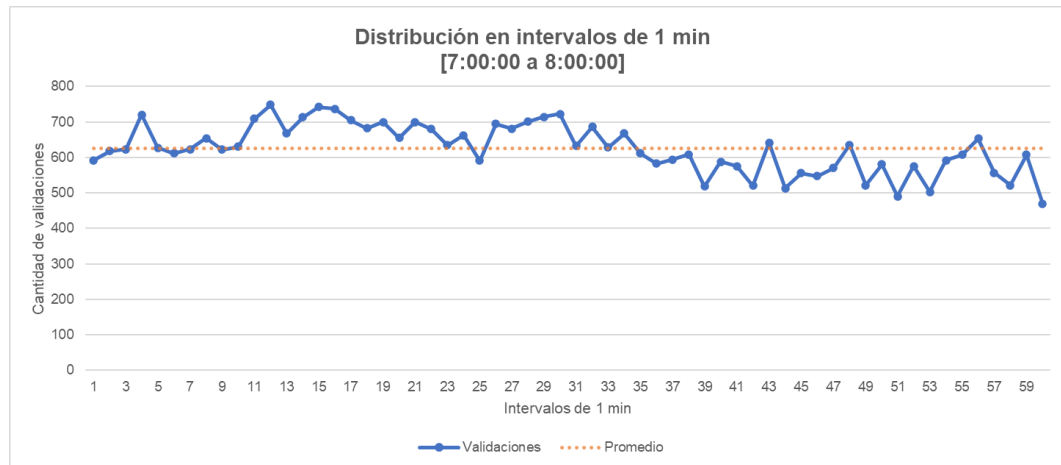


Figura 52 Distribución en intervalo de 1 min de hora pico

Fuente: Elaboración propia.

La cantidad de validaciones en la troncal sin ser tratados es de 59646 y las validaciones después de ser tratada es de 37518, esto quiere decir que, si consideramos los valores reales, la cantidad de validaciones que ingresan en la hora pico por minuto es de 994.1.

Tabla 21 Cantidad de validaciones por minuto

Estado de data	Validaciones	Porcentaje	Promedio en 1min
Data sin tratar	59646	100.0%	994.1
Data tratada	37518	62.9%	625.3

Fuente: Elaboración propia.

6. En la Tabla 22 se tiene los valores de los orígenes y destinos de las validaciones de la hora pico [7:00:00,8:00:00] en base a la matriz origen destino de la Figura 74, estos valores fueron convertidos a porcentajes con el objetivo de tener un mejor análisis. En el origen se ve que la estación más demandada es la estación “Naranjal”, en el destino la estación más demanda es “Canaval y Moreyra”.

Tabla 22 Cantidad de validaciones en Origen-Destino

N°	Estación	Origen		Destino	
		Valor	%	Valor	%
1	Naranjal	18051	48.11	1096	2.92
2	Izaguirre	1837	4.90	271	0.72
3	Pacifico	159	0.42	33	0.09
4	Independencia	695	1.85	108	0.29
5	Los jazmines	208	0.55	10	0.03
6	Tomas Valle	2478	6.60	245	0.65
7	El Milagro	368	0.98	42	0.11
8	Honorio Delgado	150	0.40	206	0.55
9	Uni	1339	3.57	452	1.20
10	Parque del Trabajo	129	0.34	91	0.24
11	Caqueta	455	1.21	123	0.33
12	2 de Mayo	314	0.84	523	1.39
13	Quilca	139	0.37	430	1.15
14	España	600	1.60	2097	5.59
15	Ramon Castilla	702	1.87	329	0.88
16	Tacna	163	0.43	359	0.96
17	Jiron de la Unión	71	0.19	1614	4.30
18	Colmena	105	0.28	865	2.31
19	Estacion Central	547	1.46	3617	9.64
20	Estadio Nacional	108	0.29	1082	2.88
21	México	212	0.57	691	1.84
22	Canadá	375	1.00	2199	5.86
23	Javier Prado	312	0.83	2811	7.49
24	Canaval y Moreyra	254	0.68	5168	13.77
25	Aramburu	97	0.26	1770	4.72
26	Domingo Orué	102	0.27	340	0.91
27	Angamos	694	1.85	3527	9.40
28	Ricardo Palma	210	0.56	1476	3.93
29	Benavides	363	0.97	2326	6.20
30	28 de Julio	107	0.29	439	1.17
31	Plaza de Flores	886	2.36	696	1.86
32	Balta	195	0.52	309	0.82
33	Bulevar	243	0.65	277	0.74
34	Union	255	0.68	195	0.52
35	Escuela Militar	46	0.12	111	0.30
36	Terán	364	0.97	144	0.38
37	Rosario de Villa	323	0.86	181	0.48
38	Matellini	3862	10.29	1265	3.37

Fuente: Elaboración propia.

En la Figura 53 en la parte izquierda se muestra el mapa de burbujas y mapa de calor de los orígenes basado en los porcentajes; en la parte derecha se muestra el mapa de burbujas y el mapa de calor del destino basado en los porcentajes. En los mapas de burbujas se puede ver las estaciones con mayor demanda y menor según el tamaño color de la burbuja, donde amarillo intenso indica porcentajes cercanos a 50%, rojos cercanos a 25% y el verde azulado 0%. En los mapas de calor se puede ver como se distribuyen los usuarios desde sus orígenes hacia sus destinos, donde el color verde azulado indica baja intensidad y el amarillo alta intensidad.

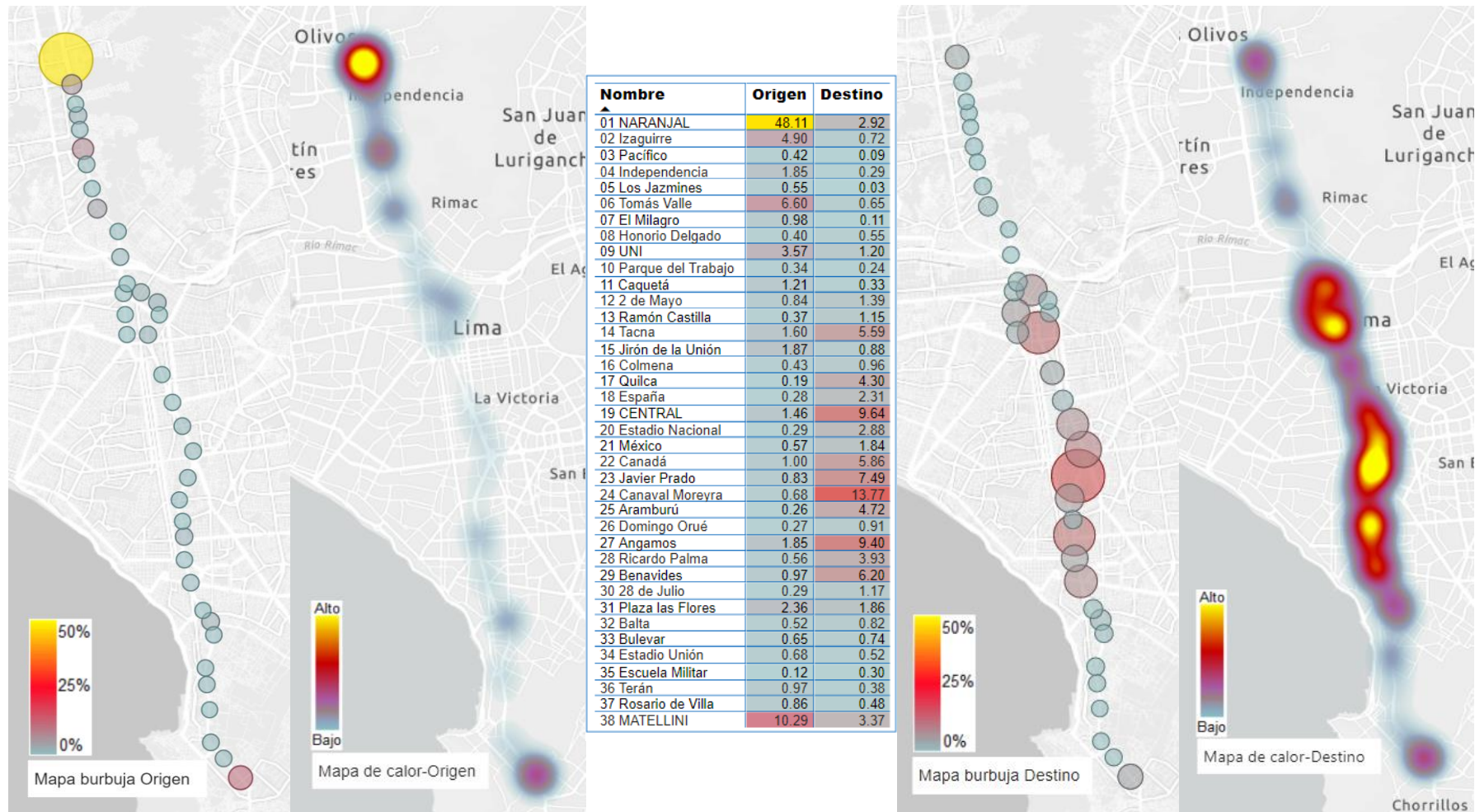


Figura 53 Mapa burbuja y mapa de calor del origen-destino
Fuente: Elaboración propia.

4.3 HERRAMIENTA 3: MACHINE LEARNING

La herramienta de Machine Learning servirá para predecir los datos que serán ingresados al modelo de “Generación de rutas óptimas”, es decir predecir los datos de los usuarios de un día cualquiera. En esta tesis se desarrollará parte del modelo de predicción teniendo como objetivo obtener un algoritmo que mejor se adecue a la fuente de datos utilizando el software Weka, este último sirve para realizar modelos predictivos automatizados.

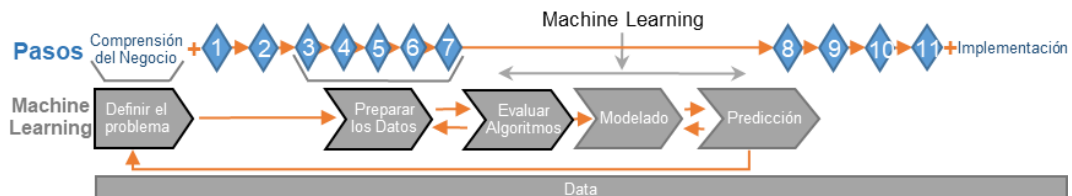


Figura 54 Pasos y proceso del Machine Learning
Fuente: Elaboración propia

El proceso de aprendizaje en Machine Learning se tiene en cuenta los siguientes pasos.

1. Definir el problema, se define el objetivo de la predicción
2. Preparar los datos, se realiza el pre análisis, pre procesamiento y análisis de los datos con la finalidad de tener unos datos que haga que la predicción de buenos resultados
3. Evaluar un conjunto de algoritmos, mediante un test estadístico se verificará que algoritmos tienen un mejor comportamiento con el conjunto de datos
4. Mejorar los resultados realizando un fine-tuning (modificar hiper parámetros y reentrenar) y ensambles, se optimizará los hiper parámetros de los algoritmos escogidos, con la finalidad de escoger el mejor modelo para la predicción
5. Modelo escogido y resultados, cuando se tenga el modelo adecuado para la predicción se evalúa en el modelo.

El objetivo de utilizar la herramienta de Machine Learning, será obtener el mejor algoritmo de predicción según la fuente de datos que se tiene, es por ello que de los pasos del proceso de aprendizaje de Machine Learning se llegara hasta el paso 3, es decir evaluar el conjunto de algoritmos que se encuentran en el software Weka. Realizar los dos últimos pasos implica mayor investigación, lo que implicaría realizar otra tesis.

4.3.1 Definir el problema

El objetivo de la predicción es la “hora” en la que una “Tarjeta” en una determinada “Parada” realiza un ingreso al sistema de Transporte Público, por lo tanto, el atributo clase o de salida es la “Fecha”, es decir el atributo “Fecha” está compuesta por el día y la hora de ingreso de la validación al sistema. Para encontrar la mejor combinación de atributos que determinen la predicción de la hora de ingreso de cada validación, se evaluó las combinaciones de los atributos y los algoritmos que están dentro del software Weka.

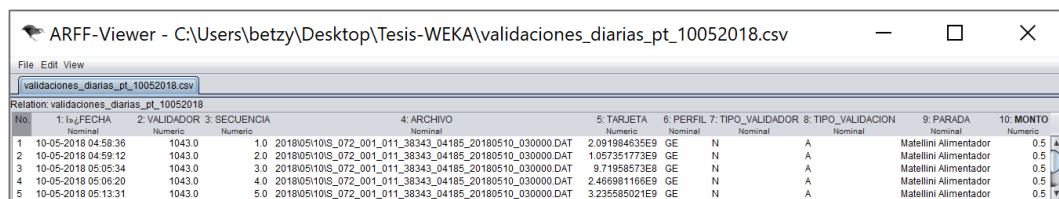
4.3.2 Preparar los datos

La preparación del conjunto de datos (31 archivos) servirán para la fase de predicción. Se usara como ejemplo al archivo del día 10/05/2018 para que se muestre los resultados. La preparación de los datos se dividirá en tres etapas.

4.3.2.1 Pre análisis

Se procede a identificar el tipo de datos de cada atributo del dataset, para ello se realiza la siguiente secuencia:

1. Abrir el archivo tipo csv
2. Visualizar como el software lee el tipo de datos
3. Abrir el archivo y cambiar el tipo de datos según corresponda
4. Guardar el archivo en tipo arff (tipo de archivo de Weka)

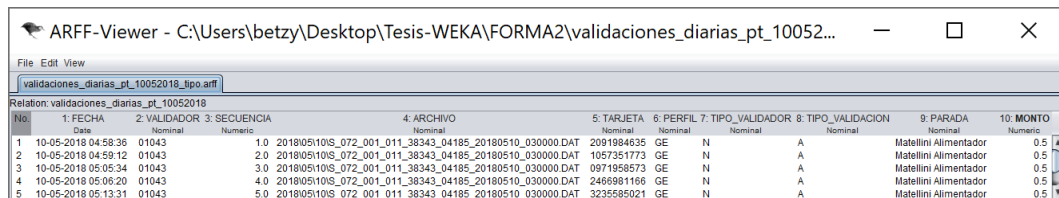


Nº	1: FECHA	2: VALIDADOR	3: SECUENCIA	4: ARCHIVO	5: TARJETA	6: PERFIL	7: TIPO_VALIDADOR	8: TIPO_VALIDACION	9: PARADA	10: MONTO
	Nominal	Numerico	Numerico	Nominal	Numerico	Nominal	Nominal	Nominal	Nominal	Numerico
1	10-05-2018 04:58:36	1043.0	1.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	2.091984635E9	GE	N	A	Matellini Alimentador	0.5
2	10-05-2018 04:59:12	1043.0	2.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	1.057351773E9	GE	N	A	Matellini Alimentador	0.5
3	10-05-2018 05:05:34	1043.0	3.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	9.71858573E8	GE	N	A	Matellini Alimentador	0.5
4	10-05-2018 05:06:20	1043.0	4.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	2.468981166E9	GE	N	A	Matellini Alimentador	0.5
5	10-05-2018 05:13:31	1043.0	5.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	3.235582021E9	GE	N	A	Matellini Alimentador	0.5

Figura 55 Tipo de datos por defecto

Fuente: Elaboración propia

En Figura 56 se muestra los atributos con su respectivo tipo de datos: fecha (date), validador (nominal), secuencia (numérico), archivo (nominal), tarjeta (nominal), perfil (nominal), tipo de validador (nominal), tipo de validación (nominal), parada (nominal) y monto (numérico).



Nº	1: FECHA	2: VALIDADOR	3: SECUENCIA	4: ARCHIVO	5: TARJETA	6: PERFIL	7: TIPO_VALIDADOR	8: TIPO_VALIDACION	9: PARADA	10: MONTO
	Date	Nominal	Numerico	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Numerico
1	10-05-2018 04:58:36	01043	1.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	2091984635	GE	N	A	Matellini Alimentador	0.5
2	10-05-2018 04:59:12	01043	2.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	1057351773	GE	N	A	Matellini Alimentador	0.5
3	10-05-2018 05:05:34	01043	3.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	971858573	GE	N	A	Matellini Alimentador	0.5
4	10-05-2018 05:06:20	01043	4.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	2468981166	GE	N	A	Matellini Alimentador	0.5
5	10-05-2018 05:13:31	01043	5.0	20180510IS_072_001_011_38343_04185_20180510_030000.DAT	3235582021	GE	N	A	Matellini Alimentador	0.5

Figura 56 Tipo de datos modificado según el atributo

Fuente: Elaboración propia

Visualizando la estadística descriptiva, se observa en la Figura 57 que el archivo del día 10/05/2018 contiene 716813 validaciones (instancias) y 10 atributos; el atributo “Fecha” es de tipo de dato Date, con 68205 valores distintos y 2472 valores únicos, teniendo como valor mínimo de 10-05-2018 00:01:52 equivalente a 1525928512000, valor máximo de 11-05-2018 02:59:47 equivalente a 1526025587000, media de 10-05-2018 13:46:21.6 equivalente a 1525977981621.957 y una desviación estándar de 05:14:17 equivalente a 18857381.25.

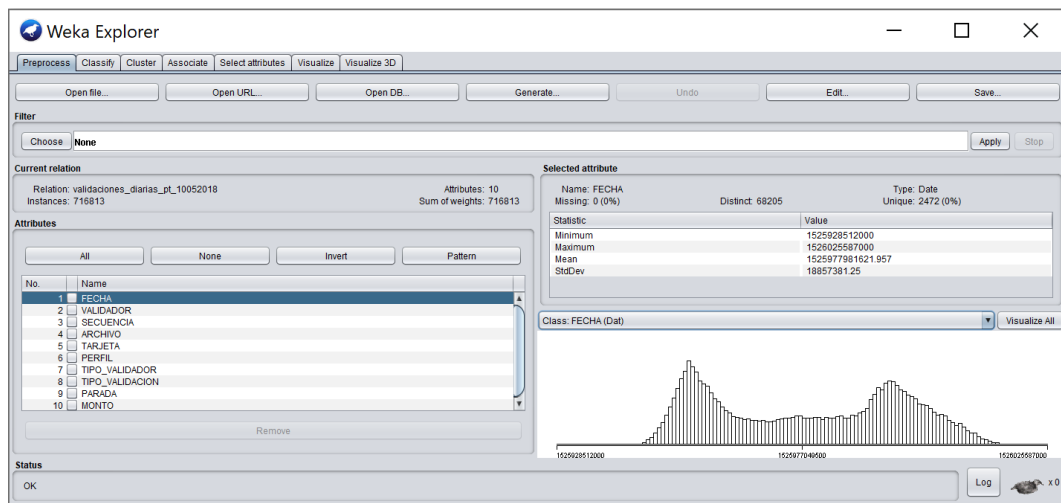


Figura 57 Estadística descriptiva del atributo Fecha

Fuente: Elaboración propia

Visualizando la distribución de los atributos, en la Figura 58 los atributos no presentan distribución de tipo gaussiana, así mismo los atributos “tipo validador” y “perfil” dentro de sus clases contienen cantidades diferenciadas por ejemplo en el atributo “tipo de validador” existen 714077 (99,6%) validadores tipo “N” y 2736 (0.4%) validadores tipo “M”.

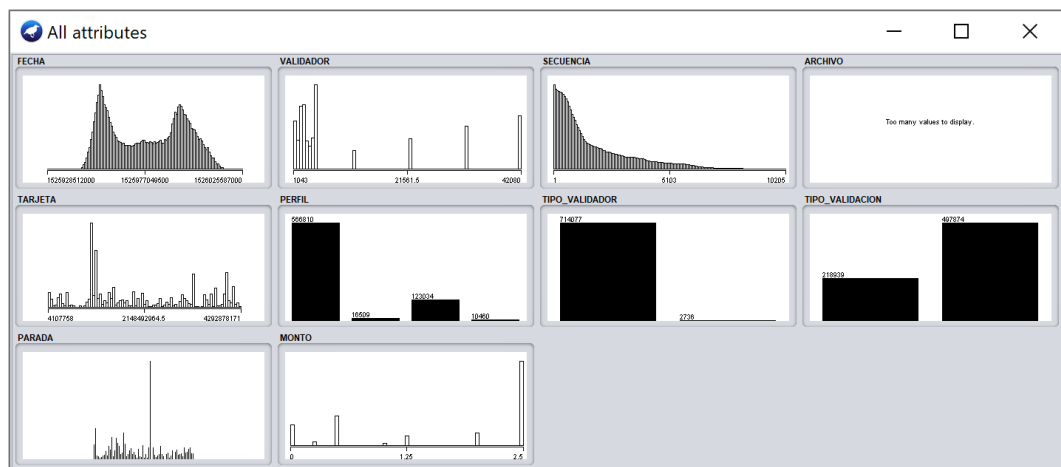
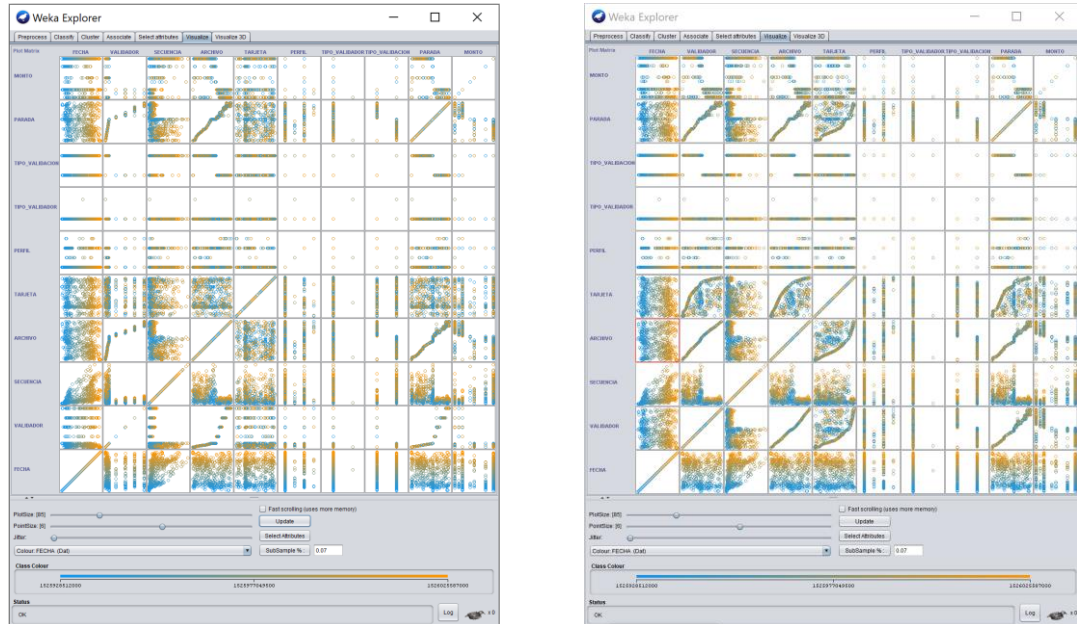


Figura 58 Distribución de los atributos

Fuente: Elaboración propia

Visualizando la matriz de correlación en función al atributo clase o de salida “fecha”, en la Figura 59 existe correlación entre los atributos tarjeta-parada y tarjeta-archivo.



a) Sin modificación de tipo de datos b) Con modificación de tipo de datos

Figura 59 Matriz de correlación
Fuente: Elaboración propia.

En el conjunto de datos no existen valores faltantes (missing), sin embargo, existen valores outliers, es decir existen valores que no están dentro del horario de atención del Metropolitano, es por ello que se tomara en cuenta en el rango de las 4:40am a 12:20am para el análisis del atributo “fecha” (ver Figura 57), la no eliminación de estos valores outliers genera ruido para el modelo predictivo. Para ello se realizará lo siguiente:

1. Delimitar los valores fuera del rango con NaN con el filtro Numeric Cleaner. Donde las 10-05-2018 4:40:00 equivale a 1525945200000 y 11-05-2018 00:20:00 equivale a 1526016000000

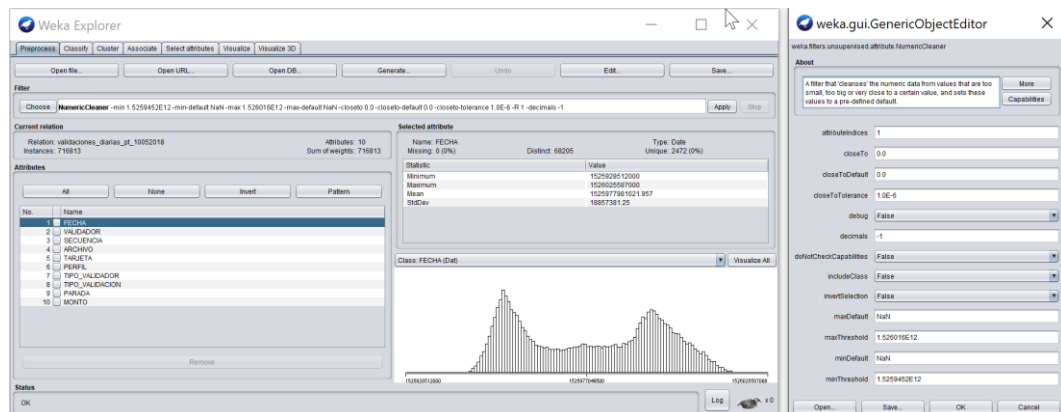


Figura 60 Delimitar valores extremos
Fuente: Elaboración propia

2. Eliminar los valores NaN con el filtro Remove with Values

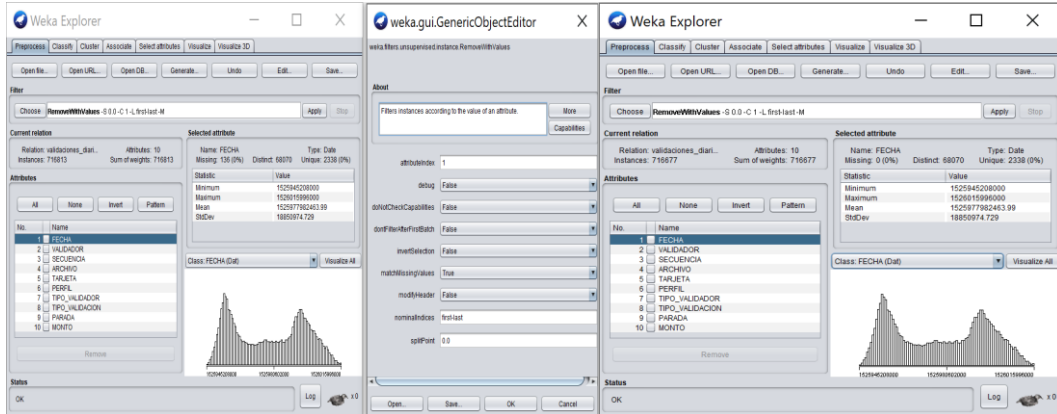
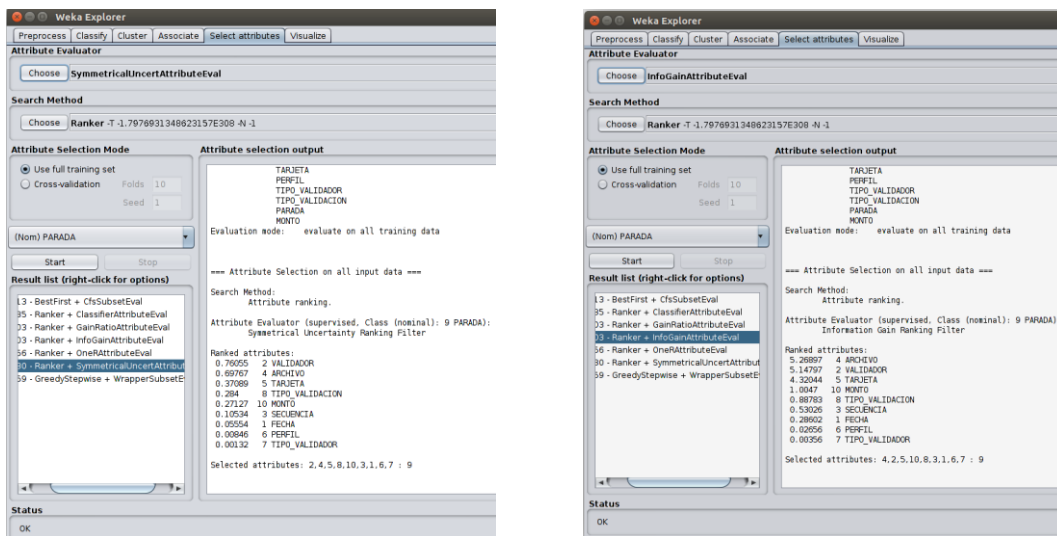


Figura 61 Eliminar valores extremos
Fuente: Elaboración propia

4.3.2.2 Análisis

Los datos deben ser entendidos antes de ser modelados, para ello se hará uso de la pestaña selección, que tendrá como finalidad seleccionar los atributos más significativos para tener una buena predicción.

La selección de mejores atributos implica probar la relación que existe entre ellos, así saber si existe una dependencia directa o que varias variables indiquen lo mismo. Como se ve en la Figura 62, el atributo “validador” esta correlacionado con el atributo “secuencia”, el atributo “parada” está altamente correlacionada con los atributos “archivo” y “validador” debido a que brindan la misma información, que es la ubicación de la “parada”. Por lo tanto, se eliminará los atributos “validadores”, “secuencia” y “archivo”.



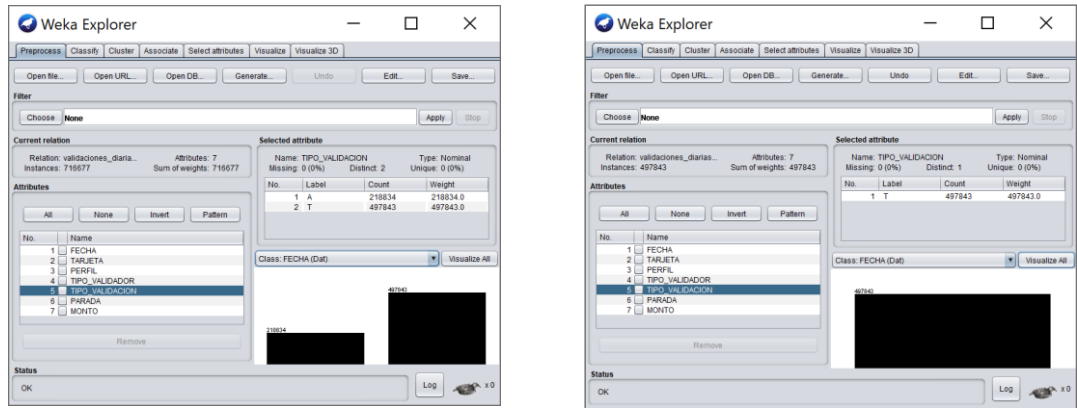
a) Ranker + Symetrical Uncertainty

b) Ranker + InfoGain

Figura 62 Selección de atributos
Fuente: Elaboración propia

En la Figura 63 se muestra los algoritmos que uso para la selección de atributos son el evaluador Symetrical Uncertainty, InfoGain y OneR con el buscador Ranker, los cuales indican resultados similares.

Debido a que la predicción de las rutas se evaluara en la ruta Troncal, se elimina el tipo de validación que contenga a los Alimentadores, al hacer ello se procede a eliminar el atributo “tipo de validación”.



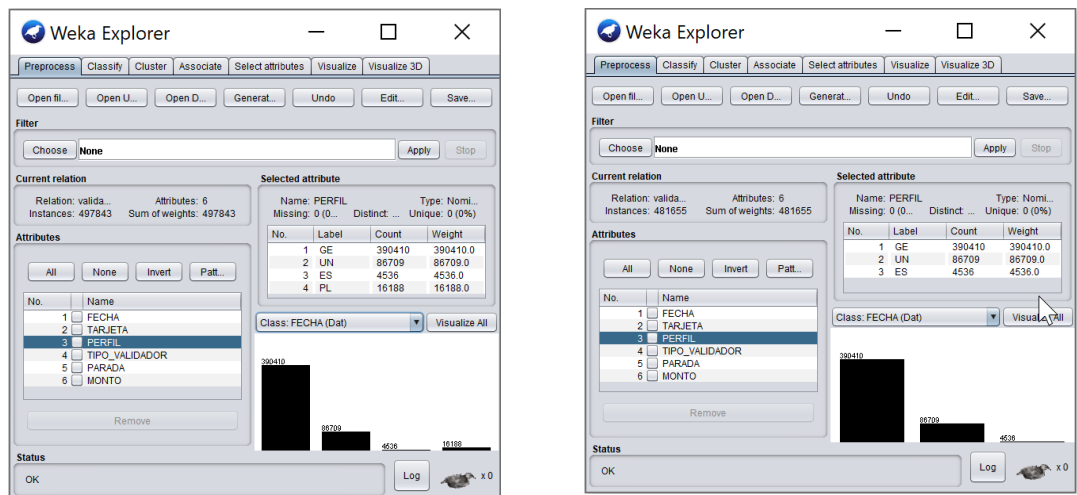
a) Filtrado de atributo “Tipo de validación”

b) Tipo de validador “Troncal”

Figura 63 Troncal, eliminación del atributo “tipo de validación”

Fuente: Elaboración propia

Debido a que el tipo de perfil PL son tarjetas que están en determinadas estaciones y no pertenecen a usuarios del Metropolitano, generará ruido cuando se realice la predicción, es por ello que se eliminara este tipo de perfil



a) Filtrado de atributo “Perfil”

b) Perfil “GE, UN y ES”

Figura 64 Eliminación del perfil “PL”

Fuente: Elaboración propia

4.3.3 Evaluar un conjunto de algoritmos

La evaluación del conjunto de algoritmos implica a realizar varias pruebas experimentales al conjunto de datos, el cual trae como consecuencia realizar modificaciones a los atributos para que los algoritmos puedan tener un mejor comportamiento. Así mismo para la evaluación de los algoritmos se usará datos de los días de la semana sin considerar sábados, domingos y feriados, así tener una predicción con mejores resultados

Después de experimentar con los datos de un día, evaluando el comportamiento con las diferentes combinaciones del tipo y forma del atributo que mejora la precisión del algoritmo, de estos resultados la mejor forma y tipo de atributos a usar son los que están representados en la Tabla 23. La manera como cambia el tipo y la forma de los atributos fue de la siguiente manera:

1. El atributo “tarjeta” del tipo de datos pasa de Numérico a Nominal manteniendo la forma del dato.
2. El atributo “perfil” se cambió la forma del dato, es decir, el perfil tipo general (GE) paso a “1”, el universitario (UN) a “2” y el tipo de perfil escolar (ES) a “3”, mantenido el tipo de dato Nominal.
3. El atributo “tipo de validador” cambio la forma del dato, es decir, paso de validador normal (N) a “1” y de minusválido (M) a “2”; manteniendo el tipo de dato Nominal.
4. El atributo “parada” mantiene el tipo de dato Nominal, cambiando en la forma del dato, es decir las 38 estaciones pasan a ser representadas del 1 al 38, por ejemplo, la estación Naranjal es representado por el “1” así sucesivamente manteniendo el orden ver ANEXO 2 en la Tabla 39.
5. El atributo “monto” cambia el tipo de dato de Numérico a Nominal, manteniendo la forma del dato.
6. El atributo “fecha” se convierte en dos tipos de atributo que son “fecha “y “hora”; el atributo “fecha” es el día al que pertenece cada validación sin considerar los sábados, domingo y feriados, pasando tipo de dato Nominal representados del 1 al 22.
7. El atributo “hora”, la hora de cada validación se pasa a segundos manteniendo el tipo de dato Date, agrupados en intervalos de tiempo (datos discretizados)

Tabla 23 Tipo de datos para predicción

Atributo	Datos originales		Datos para Predicción	
	Tipo	Forma	Tipo	Forma
Tarjeta	Numérico	343718030	Nominal	343718030
Perfil	Nominal	GE,UN,ES	Nominal	1,2,3
Tipo Validador	Nominal	N,M	Nominal	1,2
Parada	Nominal	Naranjal	Nominal	1,2,...,38
Monto	Numérico	2.5	Nominal	2,5
Fecha	Date	10/05/2018 4:40	Nominal	1,2,...,22
Hora	***	***	Date Discretizado	[17187,17396]

Fuente: Elaboración propia

Al realizar las pruebas con un día de data, con el tipo de datos para la predicción que se ve Tabla 23; en la Tabla 24 se obtuvo los algoritmos que mejor se comportan usando la métrica de exactitud (Accuracy) con los datos de un día. Es importante mencionar que se probó con todos los algoritmos del software Weka, en la siguiente tabla se muestra alguno de esos resultados.

Tabla 24 Algoritmos de predicción con un día de datos

Algoritmo	Exactitud
Bayes. Bayes Net	21.84%
Bayes.Naive.Bayes	20.84%
Bayes.Naive.MultiText	11.72%
Bayes.Naive.BayesUpdable	20.84%
Lazy.IBK	28.17%
Lazy.Kstar	22.54%
Lazy.KWL	14.72%
Meta.AdaBoostM.	14.72%
Meta.IterativeClassifierOptimizer	17.10%
Meta.LogitBootst	17.19%

Fuente: Elaboración propia

De los algoritmos en la Tabla 24, según se aumentaba la data con más días, se observó que el algoritmo que mejor se comportaba es el algoritmo de BayesBayesNet. En la Tabla 25 se muestra cómo se comporta el algoritmo BayesBayesNet en la medida que se aumenta los días, con una tendencia a seguir aumentando la exactitud cuando se aumente la data (ver Figura 65). Para tener unos mejores resultados y ver la tendencia correcta se debe probar con más data, es decir por lo menos de un año.

Tabla 25 Algoritmo BayesBayesNet según aumenta la data

Días	Exactitud
1	21.84%
5	37.17%
22	41.90%

Fuente: Elaboración propia

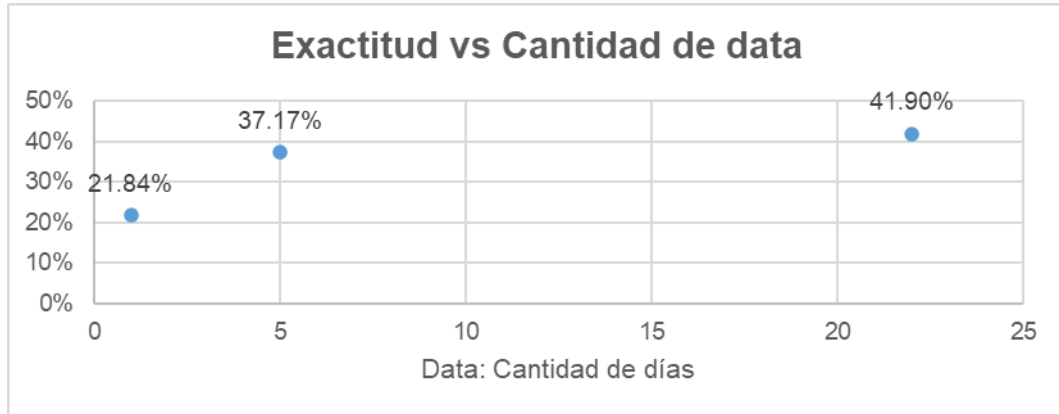


Figura 65 Precisión del algoritmo BayesBayesNet variando la data
Fuente: Elaboración propia

Se varió la cantidad de intervalos del atributo "Hora", así se evaluó en cuanto mejora la exactitud del algoritmo BayesBayesNet, ello se puede observar en la Tabla 26. En la Figura 66 se observa que no sigue una tendencia, esto es debido a que la data durante el día tiene picos y valle, es por eso que la mejor cantidad de intervalos se encontró en el valor óptimo, es decir esto se llega cuando la data tiene 10 intervalos llegando a una precisión de 51.76%.

Tabla 26 Algoritmo BayesBayesNet según aumenta los intervalos

Intervalos	Precisión
5	46.97%
10	51.76%
15	22.42%
20	41.90%
40	21.39%

Fuente: Elaboración propia

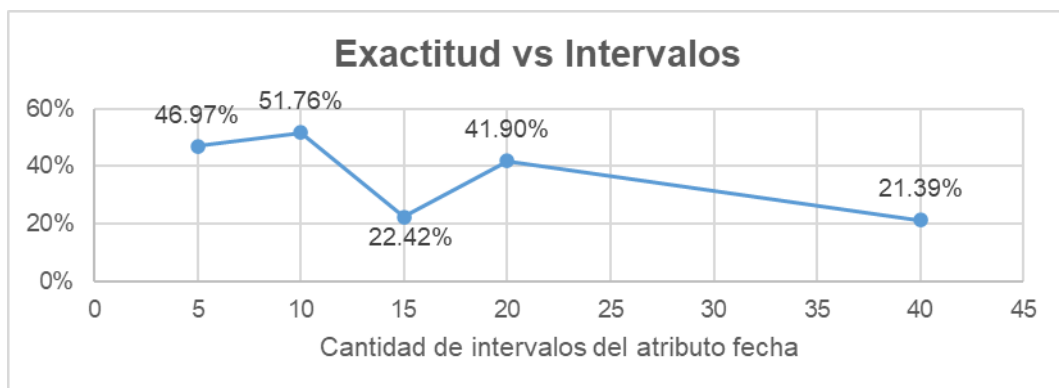


Figura 66 Precisión del algoritmo BayesBayesNet variando los intervalos
Fuente: Elaboración propia

De la evolución de los algoritmos de predicción se ve que el mejor algoritmo para este tipo de datos es el de BayesBayesNet, este es un algoritmo de clasificación, también se observó que mientras se aumenta la cantidad de datos la precisión mejora y que la cantidad de intervalos debe ser la óptima ya que influye en la exactitud del algoritmo.

CAPÍTULO V : MODELO DE GENERACIÓN DE RUTAS ÓPTIMAS

En este capítulo se desarrolla el paso 8 y 9, el primero consiste en elaborar el algoritmo de “Generación de rutas óptimas”, elaborar el seudocódigo (ver en ANEXO 5) para generar el código y ponerlo a prueba en el caso de estudio. En el paso 9 se hace la programación de las rutas del caso de estudio, es decir se analiza los resultados del algoritmo de “Generación de rutas óptimas”, así obtener los horarios de salidas de los buses. Este modelo utilizo los datos del caso de estudio, generando rutas para la hora pico de 7 a 8 am para luego compararlo con las rutas del Metropolitano

5.1 PASO 8: GENERACIÓN DE RUTAS

El paso 8 consiste implica elaborar un algoritmo y un seudocódigo para la programación. El algoritmo está basado en la demanda y comportamiento de los usuarios dentro de un sistema de Transporte Público Masivo. En el caso de estudio se pasa el seudocódigo a código usando el lenguaje de programación Python. El algoritmo de “Generación de rutas óptimas” requiere de los pasos anteriores, es por ello que se enlista los datos adicionales que se utilizó.

- Origen-destino de los usuarios, es el inicio y fin del viaje de un usuario
- Ruta recorrida, son las vías que existen dentro del sistema analizado
- Coordenadas de las estaciones, son los paraderos
- Programación de semáforos, tiempo de color rojo del ciclo semafórico.
- Velocidad de los buses, es una velocidad estándar del transporte
- Capacidad de los buses, es la capacidad máxima del transporte

5.1.1 Algoritmo: Generación de rutas

5.1.1.1 Descripción

En una red de transporte público masivo, que cuenta con una vía exclusiva de recorrido que circula en ambos sentidos, a lo largo recorrido cuenta con “N” estaciones y con “S” semáforos. Los usuarios pueden esperar en los embarques como máximo “Tmax.esp” antes de ser trasportados. Se requiere generar rutas que lleven a la mayor cantidad de usuarios en el menor tiempo posible de un origen “x” a un destino “y”. Se tiene los siguientes datos: coordenadas de las estaciones, velocidad de los buses, capacidad máxima de los buses, validación de los orígenes destinos de los usuarios y programación de semáforos.

5.1.1.2 Desarrollo

La generación de rutas constará de dos: rutas directas y rutas regulares dando como prioridad a las rutas directa debido a que se desea llevar a los usuarios sin pasar por todas las estaciones, las rutas regulares pasan por todas las estaciones. A sí mismo para la generación de las rutas se debe formar paquetes tanto directo como regulares según sea la ruta.

Entrada

1. Validaciones de los usuarios, cuyos parámetros son “fecha”, “hora”, “tarjeta”, “origen” y “destino”.
2. Coordenadas de las estaciones, cuyos parámetros son “latitud_E” y “longitud_E”.
3. La velocidad del bus, capacidad máxima del bus y tiempo de máxima espera de la validación.
4. Programación de los semáforos, cuyos parámetros “hora semáforo rojo”, “ID”, “latitud_Se”, “longitud_Se”, “tiempo de semáforo rojo” y “sentido de vía”

Constantes y variables

1. Constantes
 1. Velocidad del bus
 2. Cantidad de estaciones: N
 3. Tiempo del torniquete hasta el embarque de la validación: Ttor_emb.V
 4. Tiempo de máxima espera de la validación: Tmax.esp.V
 5. Capacidad máxima del bus: Cmaxb
 6. Tiempo en que el bus se detiene en una estación: Tdet.B
2. Variables
 1. “Latitud_E” y “longitud_E” de la estación
 2. “Fecha” de la validación
 3. “Hora” de la validación
 4. “Tarjeta” de validación
 5. “Origen” de la validación
 6. “Destino” de la validación
 7. “Sentido” de recorrido de la validación
 8. Hora de embarque de la validación: Hemb.V
 9. Hora de máxima salida de la validación: Hmax.S.V
 10. Hora de máxima salida del paquete: HSP

11. Hora de llegada del bus: HLB
12. Hora de salida del bus: HSB
13. Tiempo de la estación x a la estación y: Tx_y
14. Cantidad de semáforos rojos entre dos estaciones: sem.R
15. "Hora de semáforo rojo"
16. "ID" del semáforo
17. "Latitud_sem" y "longitud_sem" del semáforo
18. "Tiempo de semáforos rojos"
19. "Sentido de vía" del semáforo
20. Hora de bus en el semáforo rojo: HBsem

Inicio

1. Sea una "fecha" D/M/A
2. Identificar el sentido dirección de viaje de los pares orígenes destinos, es decir si la validación va de izquierda-derecha o derecha-izquierda (norte-sur o sur-norte).
3. Ordenar la variable "Hora" de menor a mayor, hacer $Hemb.V=Hora+Ttor_emb.V$ y $Hmax.S.V=Hemb.V+Tmax.esp.V$ de cada validación.
4. Generación de rutas directas
 - 4.1. Paquete directo:
 - 4.1.1. Agrupar los registros que tengan el mismo par origen destino, según la hora de llegada de cada validación (menor a mayor).
 - 4.1.2. Sea el par "origen x" y "destino y"; con "m" una validación que inicia en 1 y termina en la última validación del par "origen x" y "destino y".
 - 4.1.3. Identificar la "hora de máxima salida" de la validación "m" y compararla con la "hora de embarque" de las validaciones consecutivas:
 - 4.1.4. Sea "n" la cantidad de validaciones contadas, n inicia en 1 y termina cuando llega a la "capacidad máxima del bus -1"
 - 4.1.5. Si la "hora de máxima salida" de la "validación m" es menor que la "hora de embarque" de la "validación m+n", la "validación m" se guarda para que ingrese en la lista de viajes regulares. Luego volver al 4.1.2 haciendo que la "validación m+1" sea la "validación m"

4.1.6. Si la “hora máxima de salida” de la “validación m” es mayor igual que la “hora de embarque” de la “validación m+n”, la “cantidad de validaciones” contadas es n+1, volver al paso 4.1.4, haciendo que n sea n+1, cuando n+1=capacidad máxima del bus se forma un paquete para el viaje directo, con “hora salida del paquete” es igual a “hora de máxima salida” de la “validación m” y luego volver al paso 4.1.2 haciendo que la “validación m+n+1” sea la “validación m”

4.2. Ruta directa:

4.2.1. El bus llega a la estación origen “x”, recogiendo al “paquete directo” en un tiempo “Tdet.B”; $HLBx=HSPx$ y $HSBx=HLBx+Tdet.B$

4.2.2. Luego parte a la estación destino “y” en un tiempo “Tx_y”, $HLBy=HSBx+Tx_y$; el Tx_y se calcula de la diferencia de “coordenadas de las estaciones x e y” dividido entre la “velocidad del bus”.

4.2.3. En el destino “y” desembarca al “paquete directo” en un tiempo “Tdet.bus”, $HSBy=HLBy+Tdet.B$

4.2.4. Siendo x-y un par de la ruta directa

4.2.5. Si HSBy sea cercano a HSPy volver a 4.2.1, haciendo que la estación “y” sea la nueva estación de origen “x”. Sino unir todos los pares de la ruta directa, siendo la ruta x-y-...-z

4.2.6. Teniendo como salida Ruta x-y-...z, con hora de inicio de recorrido del bus HLBx y hora de fin de recorrido del bus HSBz, en cada estación con su respectiva hora de llegada y salida del bus.

5. Generación de rutas regulares

5.1. Paquete regular:

5.1.1. De las validaciones guardadas de los paquetes directos, ordenar según la “hora” de cada validación (menor a mayor).

5.1.2. Sea el par “origen x” y “destino desconocido”; con “m” una validación que inicia en 1 termina en la última validación del par “origen x” y “destino desconocido”.

5.1.3. Identificar la “hora de máxima salida” de la validación “m” y compararla con la “hora de embarque” de las validaciones consecutivas:

- 5.1.4. Sea n la cantidad de validaciones contadas, n inicia en 1 y termina cuando llega a la capacidad máxima del paquete -1
- 5.1.5. Si la “hora de máxima salida” la “validación m ” es menor que la “hora de embarque” de la “validación $m+n$ ”, las validaciones desde la “validación m ” hasta la “validación $m+n-1$ ” viajaran en una ruta regular, formando un paquete, con “hora salida del paquete” es igual a “hora de máxima salida” de la “validación m ”. Luego volver al 5.1.2 haciendo que la “validación $m+1$ ” sea la “validación m ”
- 5.1.6. Si la “hora máxima de salida” de la “validación m ” es mayor igual que la “hora de embarque” de la “validación $m+n$ ”, la cantidad de validaciones contadas es $n+1$, volver al paso 5.14, haciendo que n sea $n+1$; cuando $n+1$ =capacidad del paquete se forma un paquete para el viaje regular, con “hora salida del paquete” es igual a “hora de máxima salida” de la “validación m ” y luego volver al paso 5.1.2 haciendo que “registro $m+n+1$ ” sea el “registro m ”

5.2. Ruta regular:

- 5.2.1. El bus llega a la estación “origen x ”, recogiendo al “paquete regular” en un tiempo “ $T_{det.B}$ ”; $HLB_x = HSP_x$ y $HSB_x = HLB_x + T_{det.B}$. El paquete regular tendrá como “capacidad” a “cantidad de validaciones que podrán embarcar”, para la primera parada de la ruta la “capacidad” es la “capacidad máxima del bus”.
- 5.2.2. Sea y que inicia en $x+1$ hasta N
- 5.2.3. Pregunta si existe algún par con “origen menor igual a x ” y “destino y ”. Si existe, contar cuantas validaciones desembarcan en la estación y , ir a 5.2.4. Sino existen validaciones volver a 5.2.3 haciendo $y=y+1$
- 5.2.4. Luego parte a la estación “destino y ” en un tiempo “ $T_{x,y}$ ”, $HLB_y = HSB_x + T_{x,y}$; el $T_{x,y}$ se calcula de la diferencia de “coordenadas de las estaciones x e y ” dividido entre la “velocidad del bus”.
- 5.2.5. En el “destino y ” desembarcan parte de las validaciones del “paquete regular” en un tiempo “ $T_{det.bus}$ ”, $HSB_y = HLB_y + T_{det.B}$
- 5.2.6. La “cantidad de validaciones” que queda en el bus es “paquete regular” menos la “cantidad que desembarcan”, por lo tanto, la “cantidad de validaciones que podrán embarcar” es la “capacidad

máxima del bus” menos la “cantidad de validaciones que queda en el bus”.

- 5.2.7. Siendo x-y un par de la ruta regular
- 5.2.8. Si HSBy sea cercano a HSPy volver a 5.2.1, haciendo que la estación “y” sea la nueva estación de origen “x”, la capacidad de máxima del paquete será la cantidad de validaciones que podrán embarcar. Sino unir todos los pares de la ruta directa, siendo la ruta x-y-...-z
- 5.2.9. Teniendo como salida Ruta x-y-...-z, con hora de inicio de recorrido del bus HLBx y hora de fin de recorrido del bus HSBz, en cada estación con su respectiva hora de llegada y salida del bus

6. Programación de las rutas

- 6.1. Ordenar las rutas según la hora de inicio de recorrido del bus (menor a mayor). Sea la Ruta x-y-...-z
- 6.2. Sea que entre la “estación x” y la “estación y”:
- 6.3. Sino existen semáforos en rojo
- 6.4. continuar con las horas de las rutas. Volver a 6.3 con “origen x” y “destino ...” hasta que se termine la ruta
- 6.5. Si existen sem.R semáforos en rojos. A la hora salida del bus de la “estación x” se le adiciona: el “tiempo de los SR semáforos rojos”, el tiempo entre las estaciones y el tiempo detenimiento del bus en “y”; siendo la suma de ellos la nueva hora de salida del bus de la “estación y”. El semáforo rojo se considera si la “hora de semáforo rojo” es igual a la “hora del bus en el semáforo rojo”, donde “hora del bus en el semáforo rojo” es igual a la “hora en una parada anterior” más el “tiempo entre la parada anterior y el semáforo rojo”, esta se calcula con la diferencia de “coordenadas” dividido entre la “velocidad del bus”. Volver a 6.3 con “origen x” y “destino ...” hasta que se termine la ruta
- 6.6. La programación indica la hora de salida de los buses de la estación x, con el horario de llegada a las estaciones.

Salida: El algoritmo da como resultado las rutas generadas, con la hora en que llega del bus a cada parada de la ruta.

Tabla 27 Formato de salida

RUTA	x	y	...	z
HLB	HLBx	HLBy	HLB...	HLBz

Fuente: Elaboración propia.

5.1.2 Caso – Metropolitano

En este caso de estudio se aplica el algoritmo elaborado “Generación de rutas óptimas”. Se considero que no existen semáforos, debido a que implica mayor tiempo de estudio, para ello se realizó un ajuste en la velocidad de los buses de modo que se asemeje a la realidad.

A continuación, se muestra los datos de entrada que fueron ingresados al algoritmo:

- Ruta de recorrido: vía Troncal con la ramificación en el centro: ANEXO 3
- Coordenadas de las estaciones: ANEXO 2
- Validaciones de usuarios: ANEXO 1
- Cantidad de estaciones: N=38
- Velocidad de los buses: V=30 km/h
- Capacidad máxima de los buses: Cmax=160 usuarios
- Tiempo del torniquete al embarque, de la validación: Ttor_emb.Val=30seg
- Tiempo de máxima espera, de la validación: Tmax.esp.Val=15 min
- Tiempo extra de espera de las validaciones=1 min
- Tiempo en que el bus se detiene en una estación: Tdet.B=30seg

En el caso de estudio, se generó las rutas con los datos de los usuarios que se encuentran en el intervalo de 7 a 8 am del día 4 de mayo 2018, que cuenta con 37518 validaciones. Todos los datos fueron ingresados en el modelo de “Generación de rutas óptimas” sin considerar los semáforos, obteniendo 218 rutas que se puede ver de manera detallada en el ANEXO 5, estas rutas están divididas en tres grupos “Directas”, “Regular 1” y “Regulares 2”

En la Tabla 28 se muestra un resumen de algunas de las rutas óptimas del grupo “Regulares 1”, resultado del algoritmo de “Generación de rutas óptimas”, por ejemplo en el primer bus del grupo “Regulares 1” con nombre N°1_N_S, se dirigen en la dirección de Norte a Sur, inicia su viaje a las 7:03:46 de la estación los “Jazmines”, realizando paradas por las estaciones intermedias como se ve en la columna 1 para luego culminar su viaje a las 7:41:34 en la estación Matellini. En la Tabla 28 solo se muestra un resumen del resultado del modelo de generación de rutas óptimas, el modelo además de estos datos mostrados indica la hora de llegada a cada estación.

Tabla 28 Resumen rutas del algoritmo de “Generación de rutas óptimas”

N°01_N_S	N°02_N_S	N°03_N_S	N°04_N_S	N°59_N_S
07:03:46	07:06:23	07:09:02	07:11:23	07:29:43
Los jazmines	Naranjal	El Milagro	Los jazmines	Tacna
El Milagro	El Milagro	Honorio Delgado	El Milagro	Jiron de la Unión
Honorio Delgado	Honorio Delgado	Caqueta	Honorio Delgado	Colmena
Uni	Uni	2 de Mayo	Parque del Trabajo	México
Parque del Trabajo	Parque del Trabajo	Quilca	Caqueta	Estacion Central
Caqueta	Caqueta	España	Ramon Castilla	Estadio Nacional
2 de Mayo	Ramon Castilla	Estacion Central	Tacna	México
Quilca	Jiron de la Unión	Estadio Nacional	Jiron de la Unión	Canadá
España	Colmena	México	Colmena	Javier Prado
Estacion Central	Estacion Central	Canadá	Estacion Central	Canaval y Moreyra
Estadio Nacional	Estadio Nacional	Javier Prado	Estadio Nacional	Aramburu
México	México	Canaval y Moreyra	México	Domingo Orué
Canadá	Canadá	Aramburu	Canadá	Angamos
Javier Prado	Javier Prado	Domingo Orué	Javier Prado	Ricardo Palma
Canaval y Moreyra	Canaval y Moreyra	Angamos	Canaval y Moreyra	Benavides
Aramburu	Aramburu	Ricardo Palma	Aramburu	28 de Julio
Domingo Orué	Domingo Orué	Benavides	Domingo Orué	Plaza de Flores
Angamos	Angamos	28 de Julio	Angamos	Balta
Ricardo Palma	Ricardo Palma	Plaza de Flores	Ricardo Palma	Bulevar
Benavides	Benavides	Balta	Benavides	Matellini
28 de Julio	28 de Julio	Bulevar	28 de Julio	
Plaza de Flores	Plaza de Flores	Union	Plaza de Flores	
Balta	Balta	Escuela Militar	Balta	
Bulevar	Bulevar	Terán	Bulevar	
Union	Union	Rosario de Villa	Union	
Escuela Militar	Escuela Militar	Matellini	Escuela Militar	
Terán	Terán		Terán	
Rosario de Villa	Rosario de Villa		Rosario de Villa	
Matellini	Matellini		Matellini	
07:41:34	07:43:27	07:42:41	07:47:29	07:58:28

Fuente: Elaboración propia.

En la Tabla 29 , al igual que en la Tabla 28 se observa la hora de inicio y la hora llegada del recorrido de algunos buses. En esta Tabla 29 se muestra algunas rutas del grupo “Regulares 2”, se puede ver que estas rutas que son más cortas en recorrido comparadas con grupo de “Regulares 1”, esto es debido a que el algoritmo usa las validaciones restantes para generar estas rutas, es decir, primero se genera las rutas “Directas”, luego las “Regulares 1” y finalmente las “Regulares 2”.

Tabla 29 Resumen de rutas de “Generación de rutas óptimas”

N°125_S_N	N°126_S_N	N°127_S_N	N°128_S_N	N°129_S_N
08:00:13	07:17:09	07:46:31	07:09:32	07:58:45
Plaza de Flores	Angamos	Domingo Orué	Ramon Castilla	Ramon Castilla
Jiron de la Unión	Canadá	España	Naranjal	Naranjal
Naranjal	España	Naranjal		
	Quilca			
	2 de Mayo			
	Naranjal			
08:33:34	07:45:34	08:14:05	07:22:27	08:11:40

Fuente: Elaboración propia.

5.2 PASO 9: PROGRAMACIÓN DE RUTAS

En la programación de las rutas se hará un análisis del resultado del algoritmo de generación de rutas con la finalidad de agrupar algunas rutas o verificar el recorrido de la estas. Con este paso se tendrá el listado de horarios de salidas de los buses para el sistema.

5.2.1 Caso – Metropolitano

En este caso de estudio las rutas se pueden ver de manera detallada en el ANEXO 5, estas rutas se agrupo en 3 las cuales son:

Rutas directas; después de obtener los resultados del algoritmo se observa que las rutas directas están formadas por rutas que circulan de una estación a otra, haciendo que esta formación de ruta este formadas por dos puntos de paradas, esto es debido a que el algoritmo se basa en la demanda de los usuarios; es decir el algoritmo los escoge a los usuarios que viajaran de un lugar a otro con la capacidad máxima del bus 160 y con un tiempo de espera en el embarque menor a 15 min. En este grupo como resultado del algoritmo se obtuvo 89 salidas de buses en la hora pico de 7 a 8 am. En la Tabla 30 se observa un grupo de rutas directas, donde HLB es hora de llegada del bus y HSB la hora de salida de una estación.

Tabla 30 Listado de algunas rutas Directas

Nombre	Estación	HLB	HSB	Estación	HLB
L_1_14_T6_N_S	Naranjal	07:07:36	07:08:06	España	07:23:45
	Naranjal	07:14:57	07:15:27	España	07:31:06
	Naranjal	07:20:43	07:21:13	España	07:36:52
	Naranjal	07:28:01	07:28:31	España	07:44:10
	Naranjal	07:35:45	07:36:15	España	07:51:54
	Naranjal	07:47:03	07:47:33	España	08:03:12

Fuente: Elaboración propia.

Rutas regular 1; estas rutas están conformadas por rutas que embarcan y desembarcan teniendo en cuenta que el usuario tenga como un tiempo de espera en el embarque menor a 15 min y que el bus transporte como mínimo en todo su recorrido a 160 personas. El total de las rutas “Regulares 1” son 113, en la Tabla 31 se muestra algunas rutas.

Tabla 31 Listado de algunas rutas Regulares 1

Nombre	Recorrido	N°Paradas	Trasportados	Estación	HLB	...	Estación	HLB
N°01_N_S	00:37:49	29	312	Los jazmines	07:03:46		Matellini	07:41:34
N°02_N_S	00:37:04	29	285	Naranjal	07:06:23		Matellini	07:43:27
N°03_N_S	00:33:39	26	243	El Milagro	07:09:02		Matellini	07:42:41
N°04_N_S	00:36:07	29	269	Los jazmines	07:11:23		Matellini	07:47:29
N°59_N_S	00:28:45	20	160	Tacna	07:29:43		Matellini	07:58:28

Fuente: Elaboración propia.

Rutas regular 2; estas están conformadas por las validaciones restantes que cumplan la condición que el tiempo de espera en el embarque sea menor a 15 min. La mayoría de estas rutas no son tan eficientes debido a que se realizan con las validaciones restantes lo que indica que una ruta incluso puede solo transportar a un solo usuario. El total de las rutas “Regulares 2” son 16, en la Tabla 32 se muestra algunas rutas.

Tabla 32 Listado de algunas rutas Regulares 2

Nombre	Recorrido	N°Paradas	Trasportados	Estación	HLB	...	Estación	HLB
N°125_S_N	00:33:21	3	5	Plaza de Flores	08:00:13		Naranjal	08:33:34
N°126_S_N	00:28:25	6	18	Angamos	07:17:09		Naranjal	07:45:34
N°127_S_N	00:27:35	3	5	Domingo Orué	07:46:31		Naranjal	08:14:05
N°128_S_N	00:12:55	2	1	Ramon Castilla	07:09:32		Naranjal	07:22:27
N°129_S_N	00:12:55	2	1	Ramon Castilla	07:58:45		Naranjal	08:11:40

Fuente: Elaboración propia

CAPÍTULO VI: ESCENARIOS FUTUROS

6.1 PASO 10: VALIDACION DEL MODELO

A fin de validar, que las rutas del modelo de “Generación de rutas óptimas” son mejores que las rutas establecidas por el operador del “Metropolitano”. Se realizó una macrosimulación, en ambas simulaciones solo se cambiará los horarios de salida de los buses, manteniendo el resto de las condiciones iguales.

El programa de apoyo para realizar la simulación será el software PTV Visum en la versión PTV Visum Expert Thesis 2020 y PTV Visum Student 2021. Existen varios modelos de simulación macroscópica; para este tipo de trabajo es mejor usar como modelo de asignación al modelo “Basado en horarios” (Timetable-Based), debido a que este modelo considera como entrada a los horarios de salidas de buses.

El modelo de asignación “Basado en horarios” (PTV-VISUM, 2020), considera las siguientes condiciones:

- Los usuarios conocen todos los servicios en su red, y toda su programación (horarios de salida, tiempos de viaje, horarios de llegada)
- Los usuarios analizan esta programación y deciden su estrategia antes de iniciar el viaje
- Los usuarios asumen servicios puntuales
- Se lleva a cabo un análisis detallado de las transferencias
- Permite considerar todos los efectos de una programación detallada

Con la simulación se podrá comparar algunos parámetros, los cuales permitirán mostrar los beneficios del modelo de “Generación de Rutas óptimas” planteado en la presente tesis.

La simulación requiere los siguientes datos de entrada:

- Red de circulación de los buses del transporte público
- Velocidad de los buses
- Características de los buses
- Operadores de la red de transporte
- Rutas de los buses
- Cantidad de buses por Ruta
- Horarios de salida de los buses
- Matriz origen-destino de la red

6.1.1 Calibración

Se tomó datos en campo los cuales servirán para calibrar la simulación, así tener un modelado más acorde a la realidad tanto para la situación actual como la propuesta. Los datos fueron tomados en la hora pico de 7 a 8 am en el Expreso 2, los cuales son los siguientes:

- Tiempo de espera de los usuarios.
- Velocidad de viaje del bus.
- Capacidad máxima del bus.
- Cantidad de buses que salen del terminal.
- Tiempo de parada del bus por estación.

Estos datos se podrán ver en los Paso 2: Almacenamiento de datos (lectura) y ANEXO 4, tanto de los datos de campo como los datos brindados por el Metropolitano. El tiempo de espera máximo de un usuario en la hora pico es 45 min, este dato fue tomado en campo, el tiempo máximo de espera no ingresa en el software Visum, este valor se utilizó para determinar el tiempo máximo de espera que tendrá el modelo de generación de rutas, este valor fue reducido a la tercera parte en el modelo, es decir para modelo de generación de rutas el tiempo máximo de espera es 15min y para el sistema actual es de 45min. La velocidad del bus en campo y por datos brindados varía entre 20km/h a 45km/h se calibro realizando el cálculo del tiempo promedio de viaje entre los terminales dividiéndolo entre la distancia de ambas terminales, obteniendo 30km/h, el cual fue validado con la simulación de la situación actual comparándolo con los tiempos de viaje tomados en campo. La capacidad máxima del bus fue validada en campo teniendo como valor 160 con 44 asientos. La cantidad de buses que circulan en la "situación actual" en la hora pico también fue validado en campo el cual es de 398 viajes. El tiempo de parada en cada por estación fue tomada en campo dando como resultado 30 segundos en promedio.

En cuanto a la red del sistema, en las amabas simulaciones, es decir en las situación actual y situación propuesta se consideró las mismas condiciones de infraestructura; en cuanto a condiciones de operación se consideró el mismo tipo de buses, velocidad y tiempo de parada, en cuanto a demanda se utilizó la misma matriz origen- destino de 7 a 8 am. Lo único que varía en las simulaciones es los horarios de salidas y recorrido de ruta de los buses.

6.1.2 Macro simulación situación actual

Para elaborar la siguiente simulación se seguirá la siguiente secuencia

1. Ingresar las características de la red del “Metropolitano” colocando el sistema de transporte que se simulara, la capacidad de la vía, cantidad de carriles, velocidad máxima de circulación y tipo de vehículo que circulan.

La consideración que se tomará en la simulación es que no se ingresará los datos de los semáforos, ya que esto implica mayor toma de datos en toda la red, para solucionar ello se calibro la velocidad a 30km/h, así el tiempo de viaje se mantendrá similar a la realidad. Esta reducción se ha logrado luego de verificar los tiempos de viajes de una estación a otra, comparándola con la simulación.

En la Figura 67 se observa la creación del sistema de transporte en la simulación, para este caso se usó el sistema de transporte público (PuT) con la interacción de peatones (PuTWalk) en las estaciones.

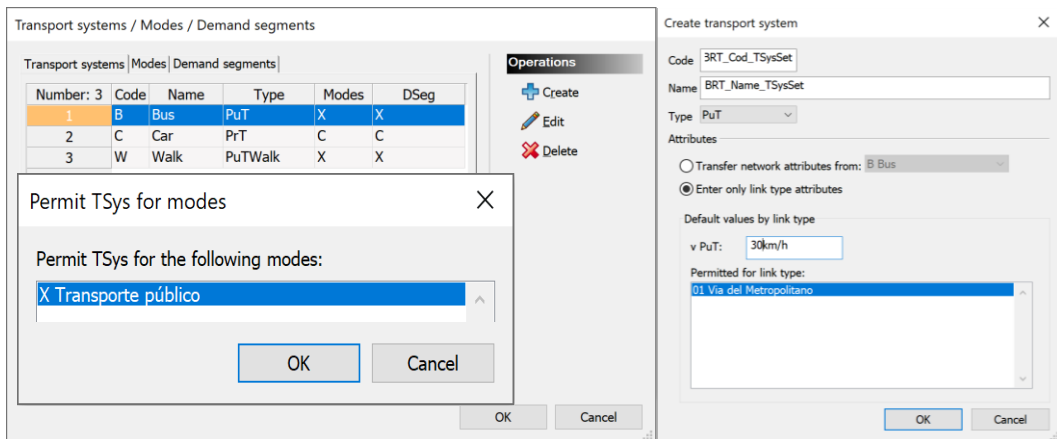


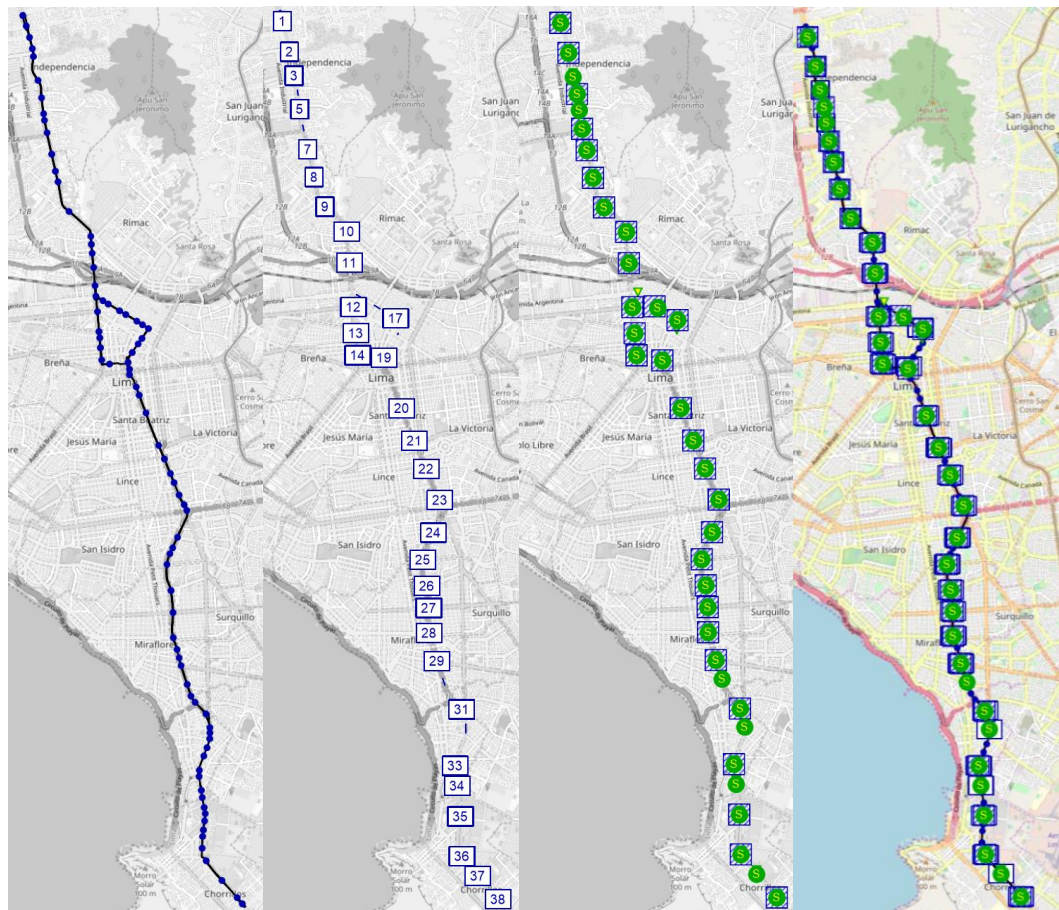
Figura 67 Características del Sistema de transporte
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

Revisando la infraestructura de la red de transporte en la mayoría de sus vías tiene 2 carriles, es por ello que se consideró este valor como número de carriles (NumLanes), se consideró el valor por defecto de 4km/h como velocidad promedio de peatón (VDef_PuTSysW). Ver Figura 68

Number	No	Name	Strict	Rank	TSysSet	NumLanes	VDef_PuTSys(BRT_Cod_TSysSet)	VDef_PuTSys(W)
1	1	Via del Metropolitano	<input checked="" type="checkbox"/>	1	B,BRT_Cod_TSysSet,W	2	30km/h	4km/h

Figura 68 Características de la via de la red del “Metropolitano”
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

- Dibujar la red del Metropolitano, es decir por donde circularan los buses colocando el sentido de las vías, nodos, enlaces zonas, conectores, los giros y paraderos como se muestra en la Figura 69. Se consideró que cada estación es una zona, debido a que en las estaciones se da los ingresos y salidas de los usuarios, es decir los orígenes y destinos de los usuarios.



69.a)Nodos y enlaces 69.b)Zonas 69.c)Paraderos 69.d) Red

Figura 69 Red del "Metropolitano"
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

En la Figura 69.a se indica que los puntos azules son los nodos, las rayas son los enlaces, en la Figura 69.b los cuadrados con números están las 38 zonas, en la Figura 69.c los cuadrados con círculos verdes indican las estaciones y en la Figura 69.d se puede ver toda la red del Metropolitano.

En el proceso de dibujar la red de transporte se tiene que revisar constantemente que los nodos, sentido de la vía, enlaces, giros, los paraderos y las zonas; así como revisar que los conectores estén vinculados con los paraderos, ya que si no se revisa implicaría resultados incorrectos.

3. Operador, características del tipo de bus, líneas, rutas recorridas de las líneas, cantidad de buses por ruta de línea, horarios de las salidas de los buses, tiempos de demora por estación.

En la Figura 70, se consideró un solo operador para la red de simulación, con un solo tipo de bus, así mismo se indica que en esta tesis no se consideró los datos de los costos en la simulación, ya que se quiere comparar las rutas de la situación actual con la situación propuesta.

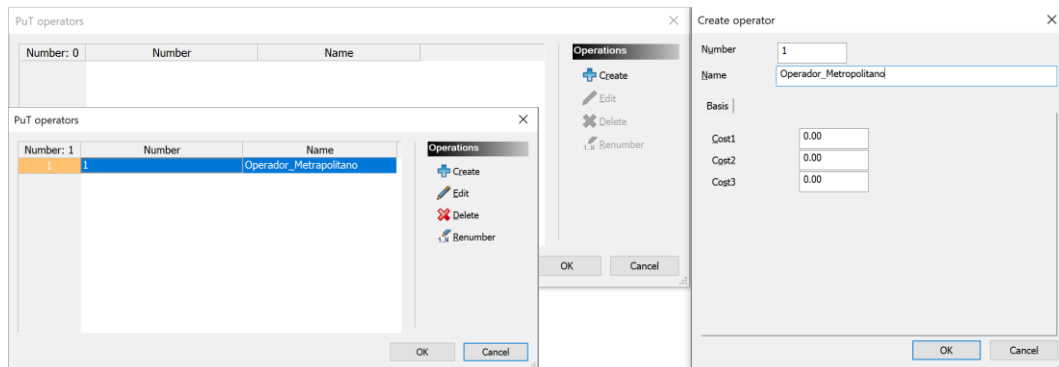


Figura 70 Operador del “Metropolitano”.
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

En la Figura 71 se muestra el tipo de bus que se utilizó en la simulación, con una capacidad total de 160 y número de asientos 44, asignándole al sistema de transporte publico BRT.

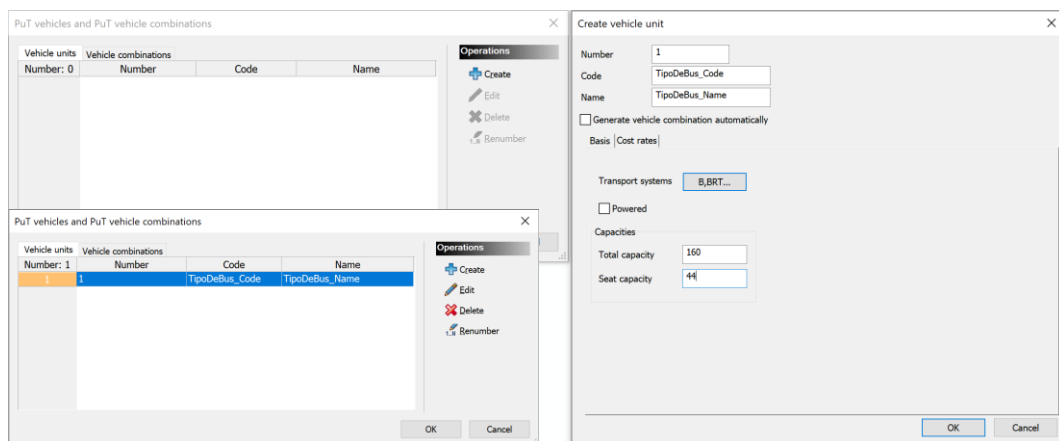


Figura 71 Características del bus
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

Como se mencionó anteriormente la simulación se realizó en el horario pico de la mañana de 7 a 8 am, en esta hora según los datos validados, la cantidad de viajes que realiza el operador del

Metropolitano es 398 despachos con 300 buses y 17 rutas de recorrido, todos ellos circulan en la ruta Troncal.

La cantidad de validaciones de la data sin tratar en el horario pico de 7 a 8am es de 59 646, estas validaciones hicieron uso de los 398 viajes. Sin embargo, para comparar las simulaciones se utilizó la matriz origen destino con la data tratada, es decir la simulación se realizó con 37 518 validaciones el cual representa el 62.9% de la data sin tratar, lo que quiere decir que se tuvo que usar 250.3 viajes repartidas en 17 rutas de recorrido.

Tabla 33 N° de salidas de bus situación actual

Estado de data	Validaciones	Porcentaje	N°salidas de Buses
Data sin tratar	59646	100.0%	398
Data tratada	37518	62.9%	250.3

Fuente: Elaboración propia.

En este proceso en la Figura 72 en la parte izquierda se muestra cómo se ingresa la ruta Expreso 1 asignándole el sistema de transporte público BRT_ Cod_TSysSet, con el operador Metropolitano, esta ruta recorrido está en el sentido de Norte a Sur (ida); en la parte izquierda de muestra el listado de todas las rutas, el sentido y la cantidad de buses asignados para cada ruta recorrida. Al realizar este proceso se tiene que asignar a cada ruta recorrido el camino por donde circulara, los paraderos en los que embarca y desembarca.

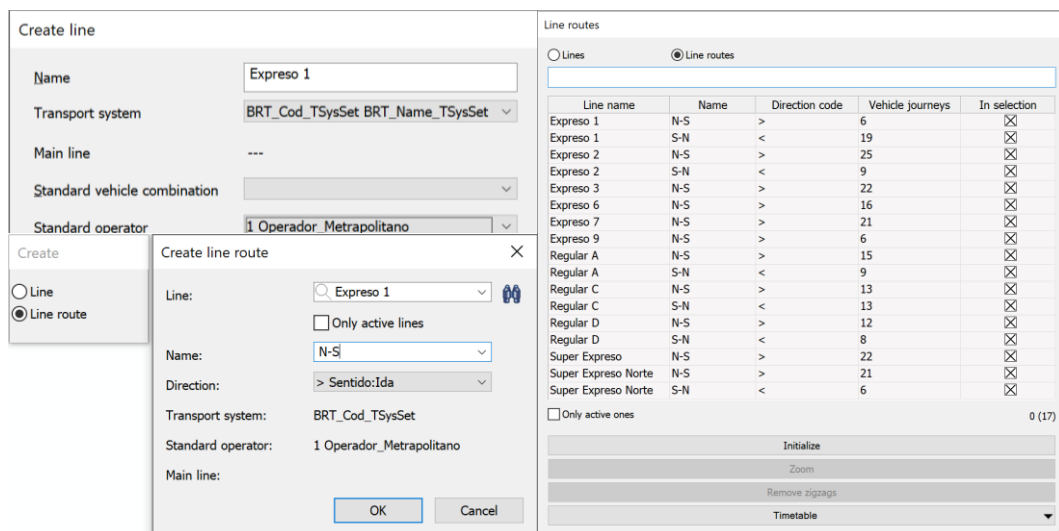


Figura 72 Rutas que circulan de 7 a 8am

Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

En Figura 73, se ingresó el tiempo que se demora en embarque y desembarque, según los datos validados este valor es de 30 segundos. También se muestra un ejemplo del recorrido de la ruta del “Expreso 1” el cual inicia en la “Estación central”, embarcando y desembarcando en algunas estaciones, terminando su recorrido en la estación “Matellini”, la distancia recorrida es de 14.34 km en un tiempo de 34 min.

Index	StopPointNo	StopPointName	IsRoutePoint	Accum.Length	Length	Info	Profile	point	Board	Alight	StopTime	Run time	Dep	Arr	AccumRunTime	CoupledVehJourneys
1	40381	Estación Central	<input checked="" type="checkbox"/>	0.000km	278km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0min		00:00:00	00:00:00	0h	<input type="checkbox"/>
9	40383	Estadio Nacional	<input checked="" type="checkbox"/>	1.278km	857km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	2min 34s	00:03:04	00:02:34	2min 34s	<input type="checkbox"/>
10	40306	México	<input checked="" type="checkbox"/>	2.136km	750km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						<input type="checkbox"/>
15	40385	Canada	<input checked="" type="checkbox"/>	2.886km	840km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						<input type="checkbox"/>
18	40312	Javier Prado	<input checked="" type="checkbox"/>	3.726km	877km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s		00:08:26	00:07:56	7min 56s	<input type="checkbox"/>
20	40128	Canaval y More	<input checked="" type="checkbox"/>	4.603km	721km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	1min 45s	00:10:41	00:10:11	10min 11s	<input type="checkbox"/>
23	40130	Aramburu	<input checked="" type="checkbox"/>	5.323km	661km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		3min 50s				<input type="checkbox"/>
24	40132	Domingo Orué	<input checked="" type="checkbox"/>	5.984km	541km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						<input type="checkbox"/>
25	40134	Angamos	<input checked="" type="checkbox"/>	6.525km	626km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s		00:15:01	00:14:31	14min 31s	<input type="checkbox"/>
26	40136	Ricardo Palma	<input checked="" type="checkbox"/>	7.150km	715km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		3min 39s				<input type="checkbox"/>
29	40138	Benavides	<input checked="" type="checkbox"/>	7.865km	869km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>						<input type="checkbox"/>
30	40140	28 de Julio	<input checked="" type="checkbox"/>	8.354km	869km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s		00:19:10	00:18:40	18min 40s	<input type="checkbox"/>
35	40318	Plaza de Flores	<input checked="" type="checkbox"/>	9.223km	474km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		2min 41s				<input type="checkbox"/>
38	40320	Balta	<input checked="" type="checkbox"/>	9.697km	964km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s		00:22:21	00:21:51	21min 51s	<input type="checkbox"/>
52	40321	Bulevar	<input checked="" type="checkbox"/>	10.661km	500km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	1min 56s	00:24:47	00:24:17	24min 17s	<input type="checkbox"/>
60	40323	Union	<input checked="" type="checkbox"/>	11.161km	761km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	1min 1s	00:26:18	00:25:48	25min 48s	<input type="checkbox"/>
70	40324	Escuela Militar	<input checked="" type="checkbox"/>	11.922km	038km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	1min 33s	00:28:21	00:27:51	27min 51s	<input type="checkbox"/>
88	40326	Terán	<input checked="" type="checkbox"/>	12.960km	619km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	2min 5s	00:30:56	00:30:26	30min 26s	<input type="checkbox"/>
90	40327	Rosario de Villa	<input checked="" type="checkbox"/>	13.580km	763km			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30s	1min 14s	00:32:40	00:32:10	32min 10s	<input type="checkbox"/>
103	40407	Matellini	<input checked="" type="checkbox"/>	14.342km				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0min	1min 32s	00:34:12	00:34:12	34min 12s	<input type="checkbox"/>

Figura 73 Detalle Ruta Expreso 1 con dirección N-S
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

4. Matriz origen-destino, una vez terminada la red de transporte público, se ingresa al software los datos de la demanda, para esta tesis se utilizó la matriz origen destino de los usuarios del Metropolitano, en el horario de 7 a 8 de la mañana con 37 518 validaciones.

En la Figura 74 se muestra la matriz origen destino de 38x38, en la parte izquierda están los nombres de las estaciones que se identifican con un número del 1 al 38, así mismo de observa que la suma de las filas y columnas de color verde, también se ve que la diagonal es de la matriz tiene valores iguales a cero, estos son debido a que no existen viajes en la misma estación. Por otro lado, en la matriz se ve que existe una mayor interacción de la estación de la estación Naranjal hacia el centro de la ciudad, este análisis ve mejor en la página 82 del paso 7 Análisis de datos.

La matriz origen destino sirve para la asignación de la demanda a red de transporte, también con esta matriz se calcula la matriz de resultados.

Matrix editor (Matrix '1 BRT_Actual_1h_Code')

38 x 38		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38				
	Name	aranj	aguirre	pacifico	de	Jazmin	V	Milagro	De	UNI	del	Taquet	Ma	Quilca	spaion	C	Tacna	Unico	merion	Cio	Na	México	anadier	Pral	y	Nambu	ingom	gmirdo	P	navid	de	Ju	de	F	Balta	uleva	Union	ela	IV	Terán	rio	del	atelli
	Sum	1096	271	33	108	10	245	42	206	452	91	123	523	430	2097	329	359	1614	865	3617	1082	691	2199	2811	5168	1770	340	3527	1476	2326	439	696	309	277	195	111	144	181	1265				
1	Naranjal	18051	0	9	0	1	0	6	0	29	99	32	51	336	245	1108	87	172	738	427	1764	353	321	1366	1513	2448	1109	112	2052	900	1367	188	331	102	89	56	17	49	77	497			
2	Izaguirre	1837	1	0	0	1	0	2	0	3	5	3	2	7	11	58	12	20	86	49	191	46	40	98	182	327	90	17	204	82	149	19	37	15	14	7	1	5	8	45			
3	Pacifico	159	1	0	0	0	0	0	0	0	3	0	1	2	0	17	1	2	6	10	16	6	1	13	9	17	7	1	16	7	11	1	0	0	0	0	0	0	10				
4	Independencia	695	1	0	0	0	0	0	0	0	0	0	3	7	9	68	3	8	22	15	84	18	19	57	65	90	23	6	79	29	38	6	16	5	3	1	1	2	1	16			
5	Los Jazmines	208	2	0	0	1	0	2	0	0	2	0	0	5	1	21	3	0	20	9	23	4	7	14	17	25	5	1	14	5	10	4	1	1	0	0	0	1	9				
6	Tomás Valle	2478	2	0	0	0	0	0	1	0	1	0	7	13	74	4	13	91	41	160	36	39	116	253	530	80	28	408	93	198	55	71	17	17	11	7	9	18	85				
7	El Milagro	368	1	0	0	0	0	0	0	0	0	0	2	3	13	0	2	13	12	33	14	5	20	31	66	17	3	45	16	38	8	9	3	0	1	1	2	0	10				
8	Honorio Delgado	150	1	0	0	0	0	0	0	0	0	0	0	1	20	0	0	10	6	23	2	3	10	17	17	6	3	11	5	5	1	2	1	2	1	0	0	0	3				
9	UNI	1339	2	0	0	1	0	0	0	0	0	0	3	5	24	0	10	27	15	40	20	12	91	125	316	64	26	177	63	134	19	37	11	12	8	2	8	12	75				
10	Parque del Trabajo	129	6	1	0	0	0	0	0	0	0	0	0	0	1	1	0	9	1	9	2	3	10	10	18	5	3	14	6	8	6	3	2	1	1	3	2	1	3				
11	Caquetá	455	14	1	0	1	0	0	0	0	0	0	0	0	4	0	1	5	4	22	4	5	19	34	97	22	5	59	25	41	7	22	3	12	3	2	6	2	35				
12	2 de Mayo	314	19	3	0	4	0	1	0	2	4	1	1	0	0	1	0	0	0	14	1	1	4	15	65	13	10	26	11	18	7	16	15	5	2	2	2	5	46				
13	Quilca	139	14	2	0	3	0	1	1	3	0	0	0	0	0	0	0	0	5	0	0	0	5	7	29	3	4	18	3	8	2	5	7	2	1	0	1	1	14				
14	España	600	62	29	3	15	0	11	7	19	15	0	2	0	1	0	0	2	1	10	2	3	10	32	111	19	14	41	15	38	7	27	12	5	10	3	7	8	59				
15	Ramon Castilla	702	8	3	1	3	0	2	0	1	1	0	0	2	0	0	0	0	2	10	4	7	36	50	147	43	26	74	30	61	24	22	20	18	4	9	8	8	78				
16	Tacna	163	13	1	0	1	0	1	0	5	2	0	0	0	0	0	0	0	0	0	0	1	5	8	41	8	5	16	8	10	6	8	1	4	3	3	1	1	11				
17	Jr.Union	71	6	2	0	0	0	2	1	0	2	0	1	0	0	0	0	0	0	0	3	1	0	5	3	7	6	3	5	3	2	2	0	2	1	3	1	1	7				
18	Colmena	105	5	2	0	3	0	1	0	0	3	0	0	0	0	0	0	0	0	0	2	0	1	1	3	21	8	3	7	3	5	7	4	0	2	2	4	3	4	11			
19	Estación Central	547	84	13	1	3	0	6	0	11	14	2	5	3	3	9	3	1	2	0	0	12	10	10	21	53	17	7	34	22	30	7	20	19	17	18	9	11	6	64			
20	Estadio Nacional	108	14	3	1	1	0	4	0	3	5	1	1	1	0	2	1	1	1	1	0	0	2	3	5	4	2	4	3	2	4	4	3	2	0	2	2	1	25				
21	México	212	36	10	3	5	0	10	2	7	9	3	2	2	1	5	1	4	2	3	1	0	1	2	4	7	4	9	5	15	2	8	6	6	3	1	3	2	27				
22	Canada	375	61	20	0	3	0	12	3	20	29	1	0	7	4	7	8	2	15	2	22	0	0	0	2	9	6	2	5	18	20	12	11	6	6	6	8	7	5	36			
23	Javier Prado	312	68	13	2	8	0	9	0	13	16	0	2	2	6	13	4	1	19	11	27	5	2	2	0	3	4	2	7	4	7	5	11	4	8	3	6	4	2	19			
24	Canaval y Moreyra	254	121	17	0	3	1	12	1	6	17	1	3	0	3	6	4	2	9	4	16	2	1	1	0	0	0	0	5	3	1	2	0	0	1	0	2	1	9				
25	Aramburu	97	32	8	0	0	2	1	0	1	3	0	4	0	0	3	3	2	10	7	7	0	0	1	0	0	0	0	1	0	0	1	1	0	3	0	1	0	6				
26	Domingo Orué	102	9	5	0	0	0	7	2	1	4	0	1	0	4	12	4	1	12	3	13	3	0	0	2	1	2	0	0	1	2	1	0	1	1	1	1	1	1	6			
27	Angamos	694	139	32	6	14	2	38	6	25	72	7	9	11	13	61	20	8	33	22	79	20	6	4	20	12	3	0	0	1	1	4	3	1	3	1	1	0	17				
28	Ricardo Palma	210	34	4	0	1	0	7	1	5	7	0	1	4	2	8	5	7	29	12	17	10	6	5	7	14	0	0	2	1	0	1	2	4	2	1	0	1	3	6			
29	Benavides	363	66	14	4	6	0	9	0	10	11	2	4	7	8	18	5	9	57	12	40	12	1	6	20	29	2	0	1	0	0	0	0	1	2	0	1	1	6				
30	28 de Julio	107	26	5	1	0	0	5	0	1	4	0	0	0	1	5	2	1	11	3	6	4	2	3	5	16	3	0	1	0	0	0	0	0	0	1	0	0	0	1			
31	Plaza de Flores	886	63	28	5	10	0	33	5	11	36	4	7	19	10	106	8	22	88	36	165	35	13	23	42	85	13	2	6	2	0	1	0	1	0	1	0	1	4				
32	Balta	195	10	1	1	0	0	2	0	0	4	0	2	4	6	17	9	3	12	6	26	20	12	6	12	16	5	0	4	0	1	1	1	0	0	1	2	2	1	8			
33	Bulevar	243	11	2	0	0	0	3	1	2	6	1	0	6	5	28	3	3	21	16	43	21	5	15	12	17	5	1	3	1	1	0	0	0	0	1	0	0	2	8			
34	Union	255	6	2	0	1	0	1	1	3	1	0	2	3	6	31	6	5	16	10	45	25	13	17	17	18	3	1	10	6	1	1	0	0	0	0	0	0	1	3			
35	Escuela Militar	46	0	2	0	0	0	2	0	0	1	1	0	0	1	3	0	1	2	3	9	7	1	3	3	4	1	0	1	0	0	0	0	0	0	0	0	0	0	1			
36	Terán	364	13	5	0	1	0	2	0	3	5	3	2	5	10	30	10	5	22	20	71	30	14	23	18	34	10	2	13	1	6	1	1	1	0	0	0	0	1	2			
37	Rosario de Villa	323	25	5	0	1	0	7	1	0	2	7	4	3	5	18	8	3	16	16	37	25	6	16	21	28	13	5	14	8	10	3	0	4	4	2	3	0	3				
38	Matellini	3862	120	29	5	17	5	46	10	21	70	21	13	75	54	313	106	54	206	87	582	337	131	181	230	448	144	42	147	95	83	29	21	41	41	35	16	2	5	0			

Figura 74 Matriz 38x38 Origen-Destino de 7 a 8 am
Fuente: Elaboración propia en PTV Visum Expert Thesis 2

5. Con los datos de la demanda, en la Figura 75 se observa las líneas de deseo entre las zonas; por ejemplo, de la zona 1 estación “Naranjal” hacia la zona 24 estación “Canaval y Moreyra” se realiza 2448 viajes y de zona 1 “Naranjal” a la zona 27 “Angamos” se realiza 2052 viajes estas están coloreadas de color naranja, esto quiere decir que en menos de 5 min se llena un bus de 160 pasajeros, es decir los pares orígenes destinos entre las 7 a 8 am son mayores a 1920.

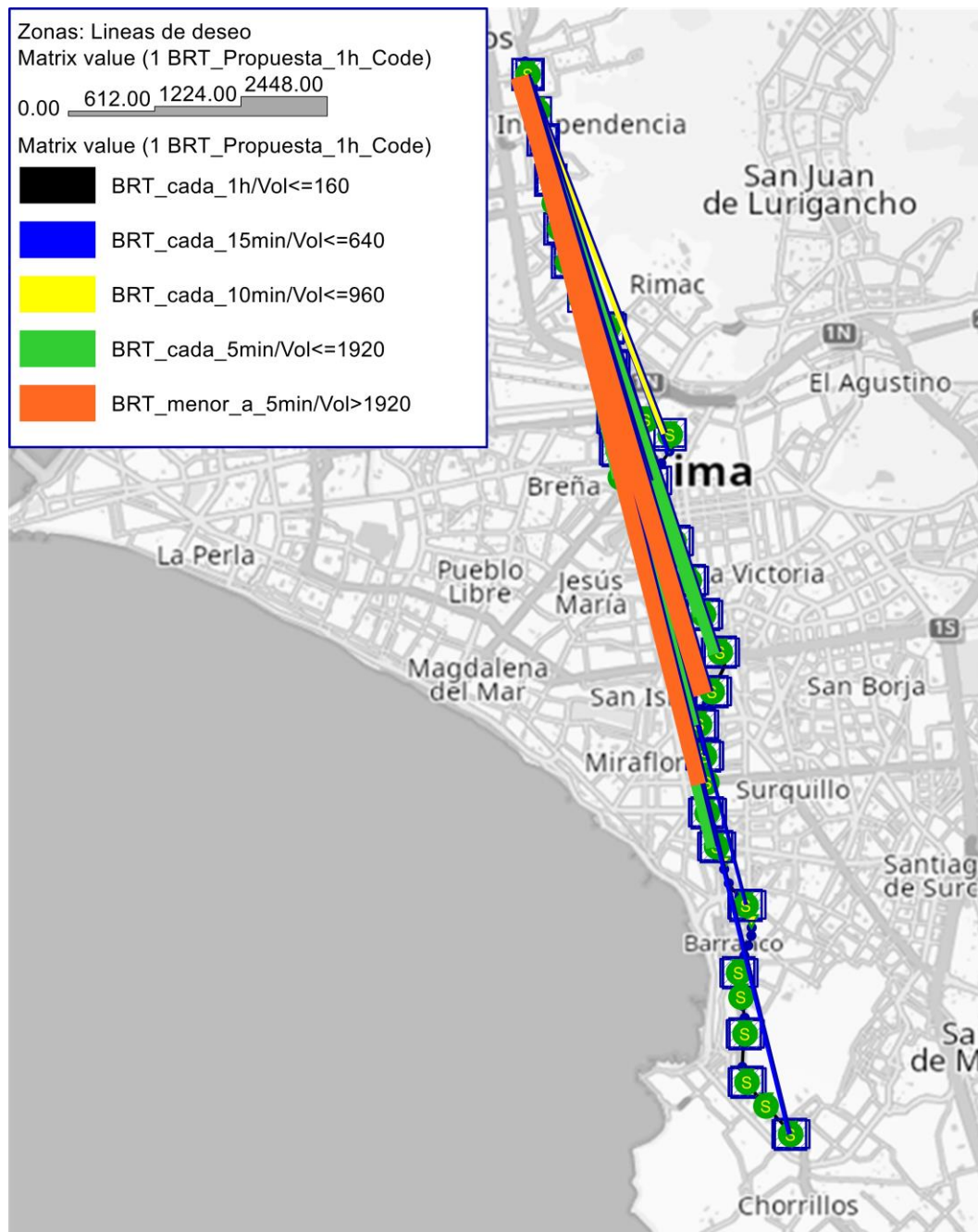


Figura 75 Líneas de deseo entre zonas 7 a 8 am
Fuente: Elaboración propia en PTV Visum Student 2021

6. Asignación de la demanda, se consideró que los usuarios tengan una distribución homogénea entre las 7 a 8 am como se muestra en la Figura 76.

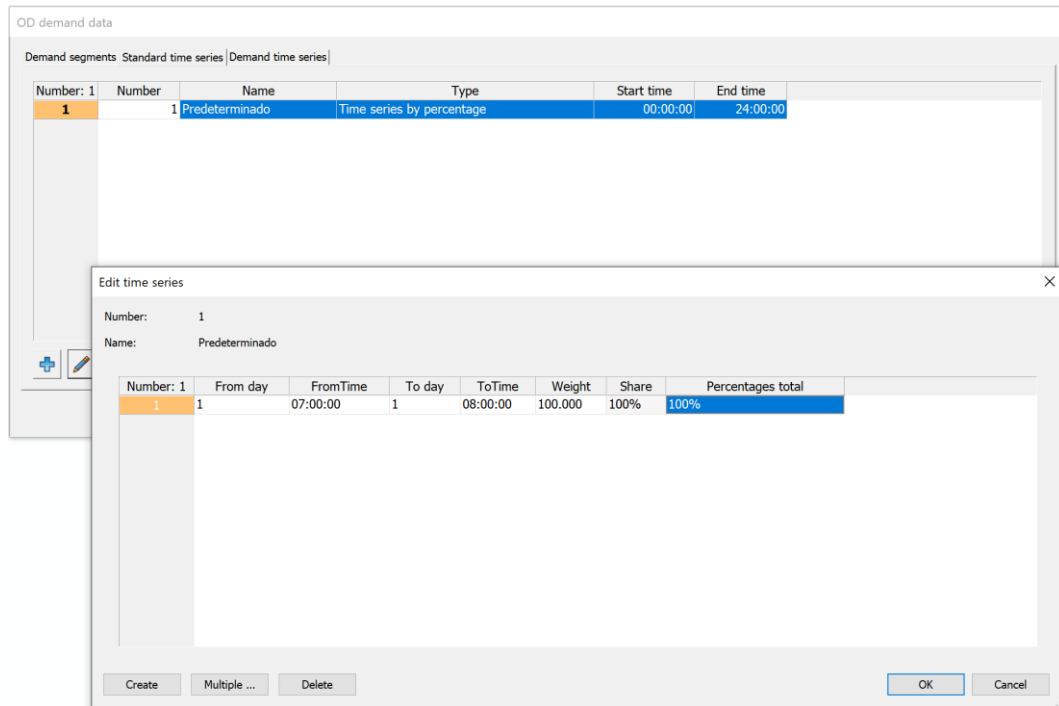


Figura 76 Asignación del volumen en la red 7 a 8 am
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

En el cálculo de los volúmenes de la demanda a lo largo de la red, el software se basa en la matriz de origen destino, es decir no depende del modelo de asignación, es decir la cantidad de buses y rutas. Por ejemplo, en la Figura 77 la mayor cantidad de flujo de personas se encuentra entre las estaciones “Caquetá” y “2 de mayo”, tramo coloreado de color rojo, lo que indica que entre las 7 a 8 am circulan más 25200, lo que quiere decir que en este tramo ingresan más de 7 personas por segundo, este tramo sería el tramo más congestionado de toda la red. En tramo de color amarillo inicia en la estación “Naranjal” y termina en la estación “Javier Prado” donde circulan entre 18000 a 25200, el cual es el tramo que abarca gran parte de la red. El tramo de color violeta claro es el tramo de menor flujo de personas donde circulan entre 0 a 3600 pasajeros.

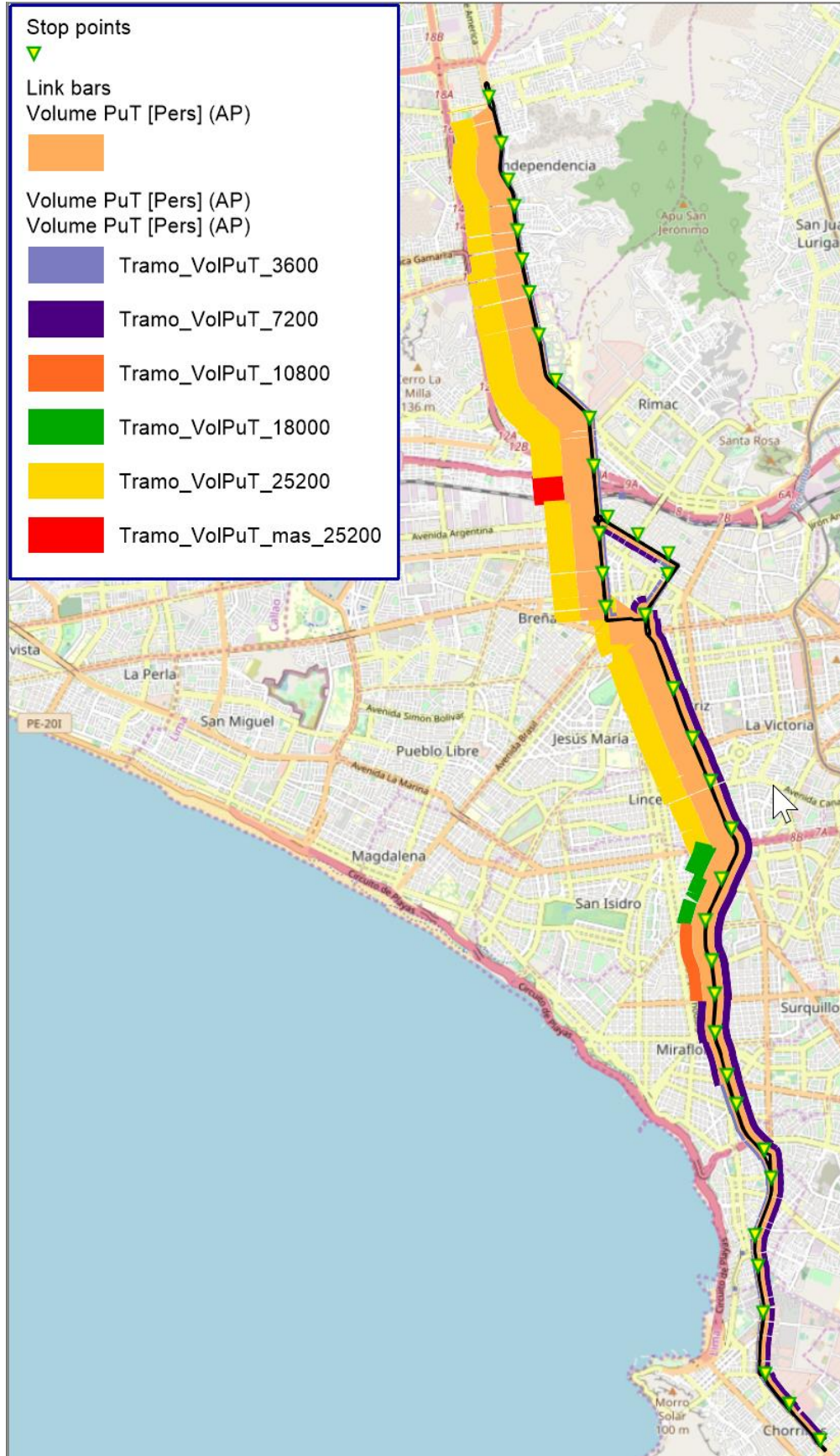


Figura 77 Distribución del volumen en la red 7 a 8 am
 Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

7. Asignación matriz de resultados, utilizaremos el modelo “Basado en Horarios” (Timetable-based), como se muestra en la Figura 78 Del cálculo de las matrices de resultados se evaluó los siguientes parámetros Tiempo de viaje (2 JRT), tiempo de transferencia adaptado (4 TWTA) y Número de transferencias (5 NTR) que servirá la comparación entre la simulación actual y la propuesta.

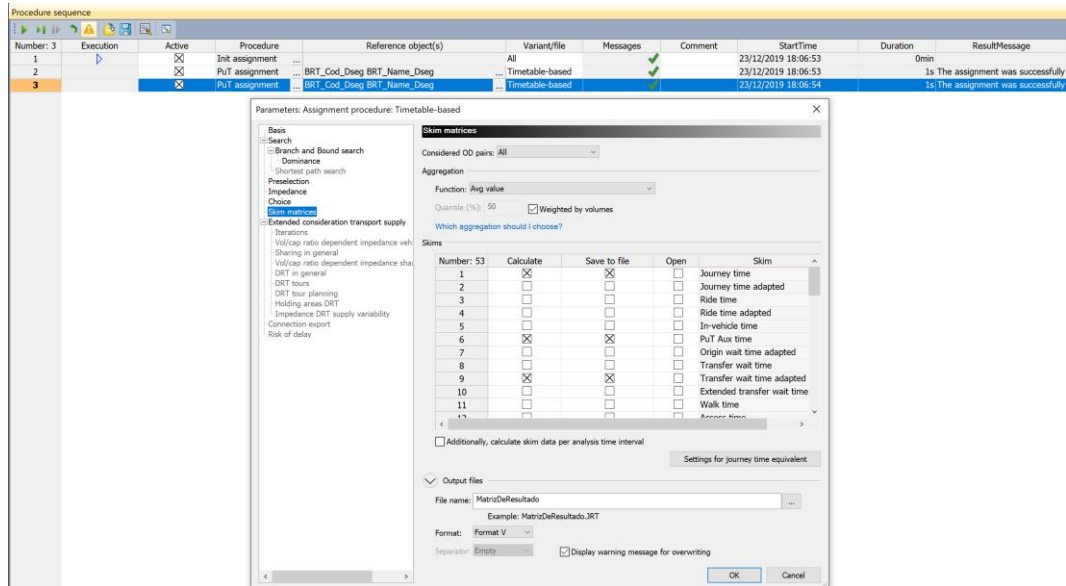


Figura 78 Asignación para el calculo de matrices de resultado
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

		1	2	3	4	5
	Name	Naranjal	Izaguirre	Pacifico	Independencia	Los Jazmines
	Sum	1061.9	992.1	1006.6	975.3	930.5
1	Naranjal	1009.6	0.0	2.1	3.9	5.3
2	Izaguirre	1078.5	2.1	0.0	1.5	2.8
3	Pacifico	1018.9	3.5	1.4	0.0	1.1
4	Independencia	972.9	4.5	2.5	1.1	0.0
5	Los Jazmines	932.3	5.3	3.3	2.0	1.0
6	Tomás Valle	889.8	6.4	4.3	3.0	2.0
7	El Milagro	874.1	7.5	5.5	4.2	3.2
8	Honorio Delgado	824.7	8.9	6.9	5.6	4.6
9	UNI	730.4	10.8	8.8	7.4	6.4
10	Parque del Trabajo	799.7	12.1	10.2	8.9	8.0

Figura 79 Matriz de resultado, tiempo de viaje (2 JRT)
Fuente: Elaboración propia en PTV Visum Expert Thesis 2020

8. Otro resultado importante cuando se ejecuta el procedimiento de asignación PuT, se puede ver el listado de caminos de una zona a otra, por ejemplo, en la Figura 80 se ve cantidad de caminos de la zona 1 a la zona 24 que son 9, de estos, en la figura esta seleccionado la ruta 8 donde indica el tiempo y la distancia que toma recorrer dicha ruta.

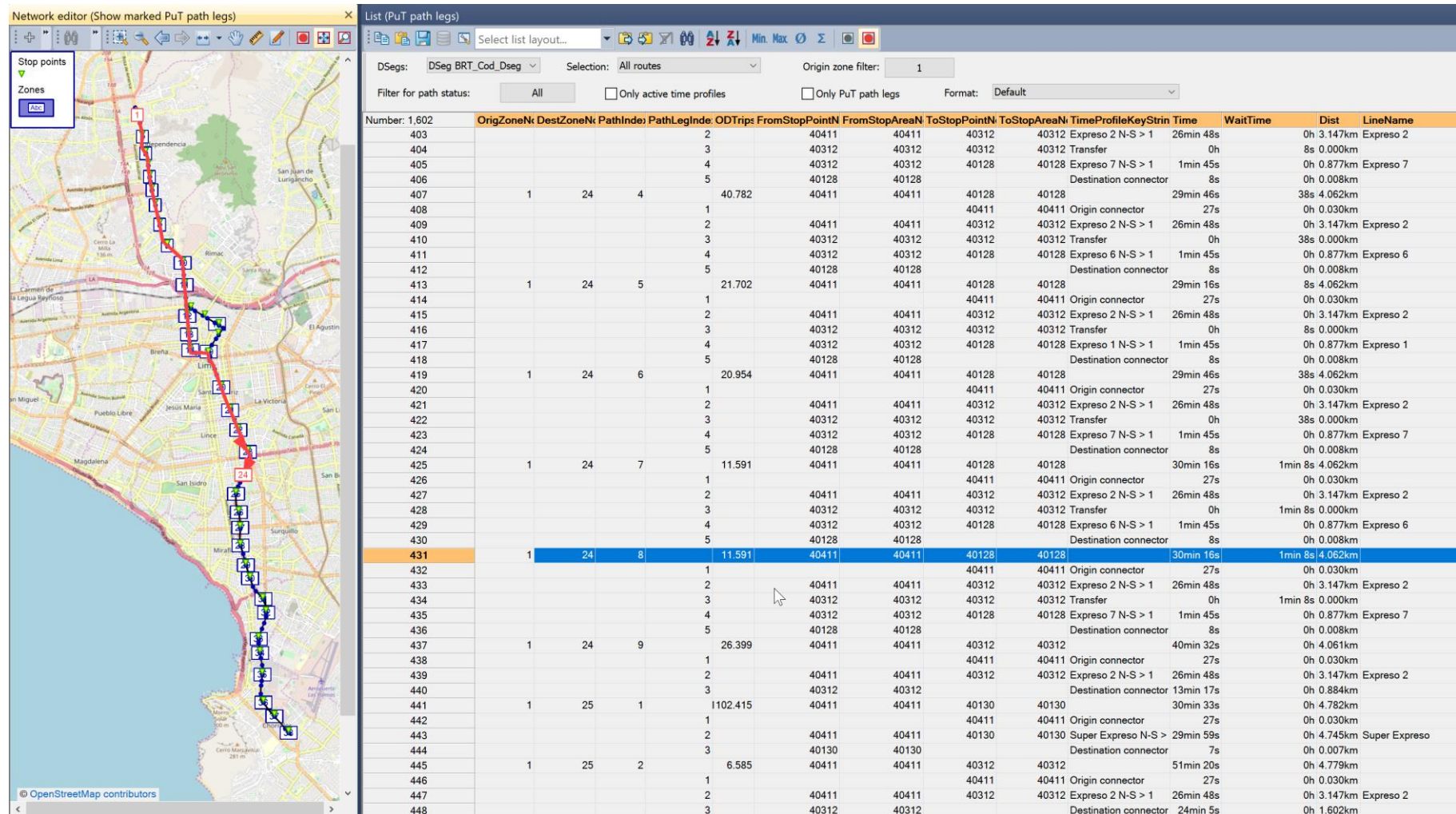


Figura 80 Detalle del resultado de caminos entre zonas
 Fuente: Elaboración propia en PTV Visum Expert Thesis 20

6.1.3 Macro simulación situación propuesta

La Macro simulación de la propuesta implica todos los pasos de la macro simulación actual (6.1.2). En la macro simulación propuesta usa la misma red transporte (nodos, enlaces, zonas, conectores y paraderos), características de bus, y la demanda con la matriz origen destino de la macro simulación actual, el cambio se da en las ruta y horarios de salida, el cual implica una matriz de resultado diferente.

1. Red de transporte público; las rutas ingresadas en la simulación propuesta son del “Modelo de generación de rutas óptimas”, el cual da como resultado 218 viajes, estos viajes están basados en la matriz origen destino de las 7 a 8 am del día 4 de mayo del 2018. La matriz origen destino es resultado de la data tratada el cual cuenta con 37 518 validaciones que según el modelo de “Generación de rutas óptima” necesita 218 viajes, sin embargo la cantidad de datos sin tratar es de 59 646 el cual necesita 346.6 viajes, las simulaciones en la situación actual y situación propuesta se elaboraron con la data sin tratar, es por eso que la cantidad de Número de Salidas de Buses ingresadas en la simulación de situación propuestas es de 218 viajes, que están agrupas en “directas”, “regulares 1” y “regulares 2”.

Tabla 34 N° de salidas de buses en “Situación Propuesta”

Estado de data	Validaciones	Porcentaje	N°salidas de Buses
Data sin tratar	59646	100.0%	346.6
Data tratada	37518	62.9%	218

Fuente: Elaboración propia.

Figura 81 Grupo de rutas - Propuesta

Fuente: Elaboración propia en PTV Visum Student 2021

En la Figura 81 se muestran los 3 grupos de rutas generadas por el modelo de “Generación de rutas óptimas”, la ruta “directa” son se agrupan a los bloques de 160 pasajeros de una estación a otra para transportarla de forma directa, es decir sin parar en las demás estaciones, la ruta “regular 1” tiene la condición de transportar en su recorrido transporte como mínimo a 160 pasajeros, en la ruta “regular 2” recoge a todos los usuarios restantes, es por eso que se puede decir que los viajes de la regular 2 no son óptimos. Estos detalles se pueden revisar en la página 99 de Programación de las rutas. En la Figura 82 se creó 143 recorridos de ruta con 218 viajes según al grupo que pertenece.

Line routes

Lines Line routes

Line name	Name	Direction code	Vehicle journeys	In selection	^
Directo	L_1_14_T6_N_S	>	6	<input checked="" type="checkbox"/>	
Directo	L_1_15_T4_N_S	>	4	<input checked="" type="checkbox"/>	
Directo	L_1_19_T11_N_S	>	11	<input checked="" type="checkbox"/>	
Directo	L_1_22_T8_N_S	>	8	<input checked="" type="checkbox"/>	
Directo	L_1_23_T9_N_S	>	9	<input checked="" type="checkbox"/>	
Directo	L_1_24_T15_N_S	>	15	<input checked="" type="checkbox"/>	
Directo	L_1_25_T6_N_S	>	6	<input checked="" type="checkbox"/>	
Directo	L_1_27_T12_N_S	>	12	<input checked="" type="checkbox"/>	
Directo	L_1_28_T5_N_S	>	5	<input checked="" type="checkbox"/>	
Directo	L_1_29_T8_N_S	>	8	<input checked="" type="checkbox"/>	
Directo	L_1_38_T1_N_S	>	1	<input checked="" type="checkbox"/>	
Directo	L_38_19_T2_S_N	<	2	<input checked="" type="checkbox"/>	
Directo	L_38_24_T1_S_N	<	1	<input checked="" type="checkbox"/>	
Directo	L_6_24_T1_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°01_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°02_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°03_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°04_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°05_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°06_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°07_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°08_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°09_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°10_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°11_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°12_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°13_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°14_N_S	>	1	<input checked="" type="checkbox"/>	
Regulares1	N°15_N_S	>	1	<input checked="" type="checkbox"/>	

Only active ones 0 (143)

Initialize

Zoom

Remove zigzags

Timetable

Figura 82 Rutas de la Propuesta
 Fuente: Elaboración propia en PTV Visum Student 2021

Al igual que la simulación de situación actual el tiempo de embarque y desembarque es de 30 segundos como se observa en la Figura 83.

Index	StopPointNo	StopPointName	IsRoutePoint	AccumLength	Length	Info	Board	Alight	StopTime	Run time	Dep	Arr	AccumRunTime	
1	40411	1_N_S Naranjal	☑	0.000km	0.779km	☑	=	☑	☐	0min	6min 30s	07:06:23	07:06:23	0h
11	40371	02_N_S Izaguirre	☑	0.779km	0.628km	☐								
13	40354	03_N_S Pacifico	☑	1.407km	0.443km	☐								
16	40353	04_N_S Independencia	☑	1.850km	0.399km	☐								
19	40351	05_N_S Los Jazmines	☑	2.249km	0.473km	☐								
24	40348	06_N_S Tomás Valle	☑	2.722km	0.541km	☑	=	☑	☑	30s	1min 25s	07:13:23	07:12:53	6min 30s
32	40346	07_N_S El Milagro	☑	3.262km	0.808km	☑	=	☑	☑	30s	1min 36s	07:15:18	07:14:48	8min 25s
35	40373	08_N_S Honorio Delgado	☑	3.961km	0.834km	☑	=	☑	☑	30s	1min 34s	07:17:24	07:16:54	10min 31s
40	40375	09_N_S UNI	☑	4.769km	0.772km	☑	=	☑	☑	30s	2min 3s	07:19:35	07:19:05	12min 42s
43	40345	10_N_S Parque del Trabajo	☑	5.603km	1.026km	☑	=	☑	☑	30s	1min 58s	07:21:39	07:21:09	14min 46s
50	40343	11_N_S Caquetá	☑	6.375km	0.475km	☑	=	☑	☑	30s	1min 58s	07:24:12	07:23:42	17min 19s
55	40409	15_N_S Ramon Castilla	☑	7.401km	0.490km	☑	=	☑	☑	30s	1min 16s	07:26:40	07:26:10	19min 47s
63	40366	16_N_S Tacna	☑	7.876km	0.629km	☑	=	☑	☑	30s	1min 36s	07:28:26	07:27:56	21min 33s
73	40389	17_N_S Jr. Union	☑	8.366km	0.794km	☑	=	☑	☑	30s	2min 34s	07:30:32	07:30:02	23min 39s
84	40391	18_N_S Colmena	☑	8.995km	1.278km	☑	=	☑	☑	30s	30s	07:33:36	07:33:06	26min 43s
101	40381	19_N_S Estación Central	☑	9.789km	0.857km	☑	=	☑	☑	30s	1min 43s	07:35:49	07:35:19	28min 56s
109	40383	20_N_S Estadio Nacional	☑	11.067km	0.750km	☑	=	☑	☑	30s	1min 29s	07:37:48	07:37:18	30min 55s
110	40306	21_N_S México	☑	11.925km	0.840km	☑	=	☑	☑	30s	1min 40s	07:39:58	07:39:28	33min 5s
115	40385	22_N_S Canada	☑	12.675km	0.877km	☑	=	☑	☑	30s	1min 45s	07:42:13	07:41:43	35min 20s
118	40312	23_N_S Javier Prado	☑	13.515km	0.721km	☑	=	☑	☑	30s	1min 26s	07:44:09	07:43:39	37min 16s
120	40128	24_N_S Canaval y Morey	☑	14.392km	0.661km	☑	=	☑	☑	30s	1min 19s	07:45:58	07:45:28	39min 5s
123	40130	25_N_S Aramburu	☑	15.112km	0.541km	☑	=	☑	☑	30s	1min 5s	07:47:33	07:47:03	40min 40s
124	40132	26_N_S Domingo Orué	☑	15.773km	0.626km	☑	=	☑	☑	30s	1min 15s	07:49:18	07:48:48	42min 25s
125	40134	27_N_S Angamos	☑	16.314km	0.715km	☑	=	☑	☑	30s	1min 25s	07:51:13	07:50:43	44min 20s
126	40136	28_N_S Ricardo Palma	☑	16.939km	0.474km	☑	=	☑	☑	30s	59s	07:52:42	07:52:12	45min 49s
129	40138	29_N_S Benavides	☑	17.654km	0.869km	☑	=	☑	☑	30s	1min 44s	07:54:56	07:54:26	48min 3s
130	40140	30_N_S 28 de Julio	☑	18.143km	0.474km	☑	=	☑	☑	30s	57s	07:56:23	07:55:53	49min 30s
135	40318	31_N_S Plaza de Flores	☑	19.012km	0.964km	☑	=	☑	☑	30s	1min 56s	07:58:49	07:58:19	51min 56s
138	40320	32_N_S Balta	☑	19.486km	0.500km	☑	=	☑	☑	30s	1min 1s	08:00:20	07:59:50	53min 27s
152	40321	33_N_S Bulevar	☑	20.450km	0.761km	☑	=	☑	☑	30s	1min 33s	08:02:23	08:01:53	55min 30s
160	40323	34_N_S Union	☑	20.950km	1.038km	☑	=	☑	☑	30s	2min 5s	08:04:58	08:04:28	58min 5s
170	40324	35_N_S Escuela Militar	☑	21.711km	0.619km	☑	=	☑	☑	30s	1min 14s	08:06:42	08:06:12	59min 49s
188	40326	36_N_S Terán	☑	22.749km	0.763km	☑	=	☑	☑	30s	1min 32s	08:08:14	08:08:14	1h 1min 51s
190	40327	37_N_S Rosario de Villa	☑	23.368km		☑	=	☑	☑	30s	1min 32s			
203	40407	38_N_S Matellini	☑	24.131km		☑	=	☑	☑	0min				

Figura 83 Detalle de la ruta regular N°2 en dirección N-S
Fuente: Elaboración propia en PTV Visum Student 2021

2. Asignación matriz de resultados, al igual que la macro simulación actual se utiliza el modelo “Basado en Horarios” (Timetable-based). De la asignación se tiene Tiempo de viaje (2 JRT), tiempo de transferencia adaptado (4 TWTA), Número de transferencias (5 NTR). Las matrices de resultados se comparan el ítem 6.2.2.

		1	2	3	4	5
	Name	Naranjal	Izaguirre	Pacifico	Independencia	Los Jazmines
	Sum	1119.32	1073.43	1026.93	984.73	1034.86
1	Naranjal	0.00	12.22	13.52	20.15	26.12
2	Izaguirre	1221.35	2.08	1.47	8.10	14.07
3	Pacifico	1032.18	3.83	1.42	0.00	6.87
4	Independencia	1063.60	5.23	2.82	1.07	0.00
5	Los Jazmines	890.41	7.27	4.46	2.46	0.97
6	Tomás Valle	912.52	7.58	5.16	3.40	1.98
7	El Milagro	825.41	9.26	6.84	5.08	3.66
8	Honorio Delgado	785.21	11.10	8.68	6.95	5.57
9	UNI	754.10	13.17	10.75	9.01	7.63
10	Parque del Trabajo	723.50	15.14	12.75	11.19	9.80

Figura 84 Matrices de resultados
Fuente: Elaboración propia en PTV Visum Student 2021

En la Figura 85 la cantidad de caminos de la zona 1 a la zona 24 son 18, de este, en el camino 8 se observa la ruta, cantidad de viaje, el tiempo y la distancia que toma recorrer dicho camino.

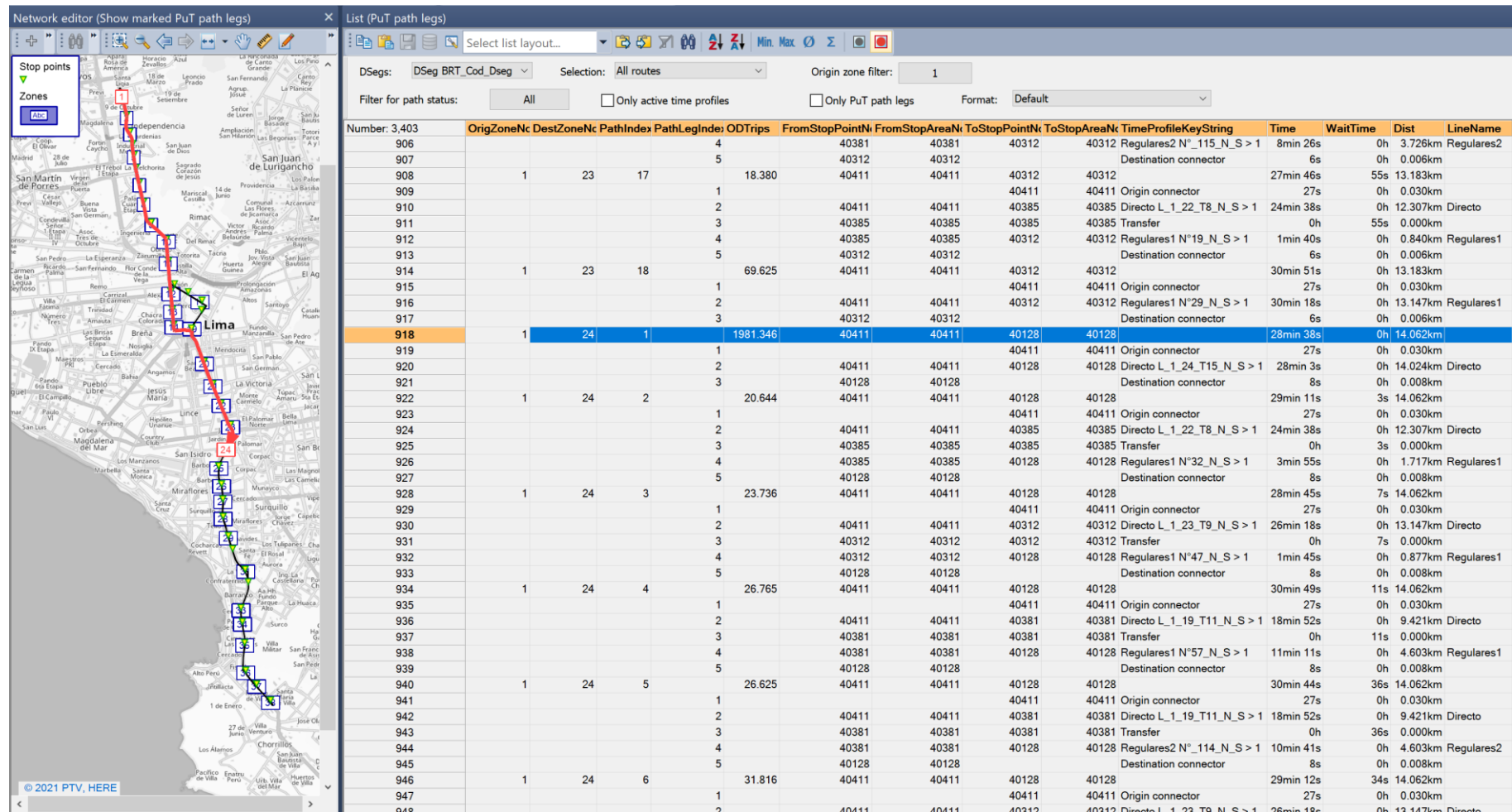


Figura 85 Detalle del resultado de caminos entre zonas
Fuente: Elaboración propia en PTV Visum Student 2021

6.2 PASO 11: ANÁLISIS COMPARATIVO SITUACIÓN ACTUAL Y PROPUESTA

Después de haber realizado la simulación del estado actual del Metropolitano y de la propuesta, se procede a comparar los resultados para ello se compara la cantidad de viajes, matriz de resultados, perfiles de tiempo y números de transferencias en los paraderos. Con la simulación principalmente se verá las ventajas en el número de transferencias y tiempo de transferencias.

6.2.1 Rutas: Cantidad de viajes

Los horarios de salida de la situación actual están basados en la data brindada por Protransporte, el cual indica que en la hora pico de 7 a 8 mañana se realizan 398 viajes. Luego del tratamiento de la información visto en el capítulo 4, del total de datos en la hora pico de 7 a 8 am del día 4 de mayo del 2018 queda reducida al 62.9%; con esta data se realizó la generación de rutas dando como resultado 218 viajes, lo que quiere decir que si la data estaría a un 100% la cantidad de viajes serían 346.6 viajes. Entonces para que las simulaciones tengan las mismas condiciones se usó en la situación actual 250 viajes que corresponde al 62.9% de 398 viajes, y 218 viajes en la situación propuesta. En la Tabla 35 también se puede observar que al hacer la diferencia de los viajes de Protransporte y el modelo de “Generación de rutas óptimas” con la data al 100%, hay un ahorro de 51 viajes en la hora pico de 7 a 8am.

Tabla 35 Número de salida de buses ahorro

Estado de data	Validaciones	Data al:	N° salida de Buses		Ahorro
			Actual	Propuesta	
			Protransporte	Modelo	
Data sin tratar	59646	100.0%	398	346.6	51.4
Data tratada	37518	62.9%	250.3	218	32.3

Fuente: Elaboración propia.

6.2.2 Matriz de resultados

En Figura 86 y Figura 87 se muestra los valores de la matriz resultados “tiempo de viaje (2 JRT)”, “tiempo de espera transferencia adaptado (4 TWTA)” y “número de transferencias (5 NTR)”, tanto de la simulación actual como simulación propuesta.

Number: 4	No	Code	Name	MatrixType	ObjectTypeRe	Sum	DSegCode	DataSourceType
1	1	BRT_Actual_1h_Code	BRT_Actual_1h_Name	OD demand	Zone	37518.000		Data
2	2	JRT	Journey time BRT_Cod_Dseg	Skim	Zone	31686.358	BRT_Cod_Dseg	Data
3	4	TWTA	Transfer wait time adapted BRT_Cod_Dseg	Skim	Zone	529.717	BRT_Cod_Dseg	Data
4	5	NTR	Number of transfers BRT_Cod_Dseg	Skim	Zone	456.821	BRT_Cod_Dseg	Data

Figura 86 Listado de matrices resultado, simulación actual

Fuente: Elaboración propia en PTV Visum Student 2021

Number.4	No	Code	Name	MatrixType	ObjectTypeRef	Sum	DSegCode	DataSourceType
1	1	BRT_Propuesta_1h_Code	BRT_Propuesta_1h_Name	OD demand	Zone	37518.000		Data
2	2	JRT	Journey time BRT_Cod_Ds	Skim	Zone	31258.812	BRT_Cod_Dseg	Data
3	4	TWTA	Transfer wait time adapted f	Skim	Zone	87.695	BRT_Cod_Dseg	Data
4	5	NTR	Number of transfers BRT_Ci	Skim	Zone	141.343	BRT_Cod_Dseg	Data

Figura 87 Listado de matrices resultado, simulación propuesta
Fuente: Elaboración propia en PTV Visum Student 2021

De acuerdo a simulación la matriz “tiempo de viaje (2 JRT)”, toma en cuenta el tiempo de viaje de una zona a otra en función de los horarios de salidas de los buses y la red de transporte, en ambas simulaciones los resultados son similares, sin embargo, en la simulación propuesta el tiempo de viaje de toda la red en menor a la simulación actual (ver Figura 88). Estas matrices tanto de simulación actual y propuesta se puede ver en el ANEXO 7, sacando un extracto de estas matrices, se observa en la Tabla 36 que de la zona1 a la zona 2 en la simulación actual, el tiempo de viaje es menor comparado con la simulación propuesta, esto es debido a que existen más rutas que paran en la zona 2 en la simulación actual; sin embargo el tiempo de viaje de la zona 1 a la zona 21 pasa lo contrario, esto es debido a que en la simulación propuesta existen más rutas que paran en esa zona 21.

Tabla 36 Comparación de los tiempos de viaje

Simulación	Zonas	2 Izaguirre		21 México	
		Demanda	Tiempo	Demanda	Tiempo
Actual	1 Naranjal	9	2.10 min	321	36.68 min
Propuesta	1 Naranjal		12.20 min		27.18 min

Fuente: Elaboración propia.

La matriz de resultado tiempo de espera de transferencia adaptado (4 TWTA), como se ve en la Figura 88, pasa de 529.72 min (simulación actual) a 87.69 min (simulación propuesta), el cual indica que el tiempo de espera de transferencia se reduce en un 83.44%. Por ejemplo, en la Tabla 37 se muestra un extracto de las matrices de resultados donde en ambos casos disminuye el de tiempo de espera de transferencia adaptado, es decir pasa de 1.22 min a 0.02 min y de 1.27min a 0.05 min. La matriz de tiempo de espera de transferencia adaptado está en el ANEXO 7.

Tabla 37 Comparación de los tiempos de espera de transferencia

Simulación	Zonas	21 México		27 Angamos	
		Demanda	Tiempo	Demanda	Tiempo
Actual	6 Tomas valle	39	1.22 min	408	1.27 min
Propuesta	6 Tomas valle		0.02 min		0.05 min

Fuente: Elaboración propia.

En las matrices de resultado de número de transferencias (5 NTR), se observa en la Figura 88 como disminuye de 456.82 (simulación actual) a 141.34 (simulación propuesta) transferencias en toda la red, lo que indica que disminuye en un 69.06% en la cantidad de transferencias. Las matrices de resultados se pueden ver en el ANEXO 7.

La Tabla 38 es un extracto de las matrices de número de transferencias, donde se observa que para ir de la zona 6 a la zona 21, así como, para ir de la zona 6 a la zona 27 existe una disminución en el número de transferencias.

Tabla 38 Comparación de número de transferencia

Simulación	Zonas	21 México		27 Angamos	
		Demanda	Transferencias	Demanda	Transferencias
Actual	6 Tomas valle	39	1.00	408	1.24
Propuesta	6 Tomas valle		0.13		0.22

Fuente: Elaboración propia.

En la Figura 88 se muestra un resumen de las matrices de resultado de toda la red, en este grafico se puede observar como la propuesta de rutas por el modelo de “Generación de rutas óptimas” es mejor frente a las rutas programadas por Protransporte (“actules”). Estos resultados son mas notorios en el Tiempo de espera de transferencia adoptado pasando de 100% a 16.56% y en el Número de transferencias pasando de 100% a 30.94%.

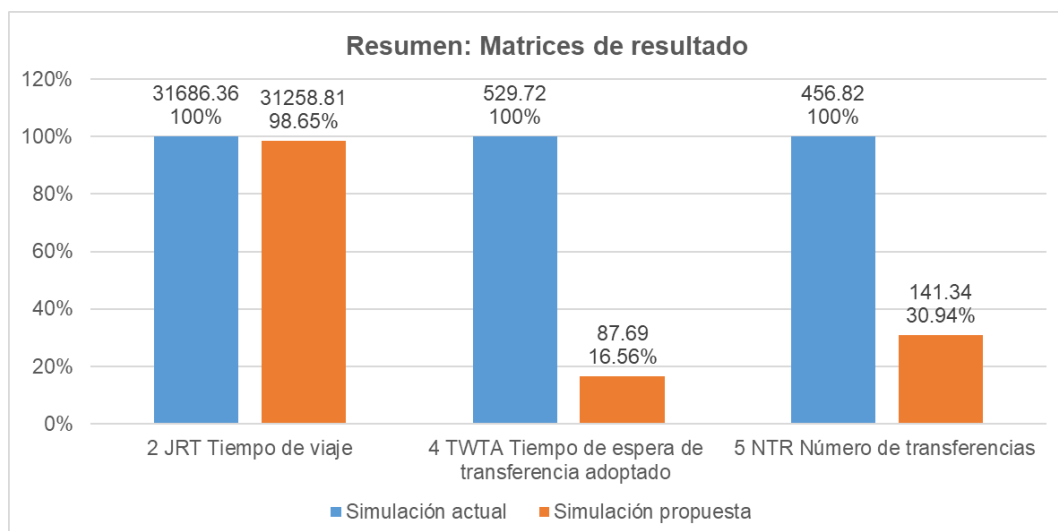


Figura 88 Resumen de las matrices de resultado
Fuente: Elaboración propia

6.2.3 Perfil de Tiempo

En la simulación actual los horarios de salida de los buses en las rutas establecidas han sido ingresadas de acuerdo a los datos brindados por “Protransporte” y validados en campo. En la simulación propuesta los horarios de salida son el resultado del modelo de “Generación de rutas óptimas” que son basadas en la demanda de los usuarios.

En las Figura 89 y Figura 90 , en el eje vertical se encuentran las estaciones en dirección de norte a sur (color negro), seguido de las estaciones de sur a norte (color naranja); en el eje horizontal se encuentra el tiempo, para ambas simulaciones los horarios de salida de los buses se dan de 7 a 8 de la mañana. El recorrido de cada bus según su ruta desde la hora de salida hasta la hora de llegada esta dibuja en el perfil de tiempo, donde los puntos negros indican las paradas que realizan los buses.

Analizando la matriz de resultados de Numero de Transferencias con los Perfiles de Tiempo, se explica por que el número de trasderencias en la simulación actual es mayor a la simulación propuesta, esto es debido a que en la simulación actual (Figura 89) los horarios de salidas de los buses se dan en bloques, por ejemplo inicio de ruta estación naranjal hacia estación central. Sin embargo en la simulación propuesta (Figura 90), los horarios de salida estan en funcion de la demanda de los usuarios, ello hace que exita menor número de transferencias.

En la simulación actual las paradas de los buses se distribuyen homogéneamente a lo largo de cada ruta, en cambio en la simulación propuesta se ve que las paradas se dan a partir de las estaciones con mayor demanda; por ejemplo en la dirección norte sur de simulación actual existe una distribución homogénea de las paradas desde la estación “Naranjal” (01_N_S) hacia la estación “Matellini” (38_N_S), en cambio en la simulación propuesta se ve una distribución homogénea a partir de la estación “Quilca” (13_N_S) hasta la estación Matellini (38_N_S).

En las rutas propuestas se observa que existen paradas después de haber trasportado a la mayor cantidad usuarios, esto es debido a que existe un pequeño grupo de usuarios que se transportan hasta las últimas estaciones.

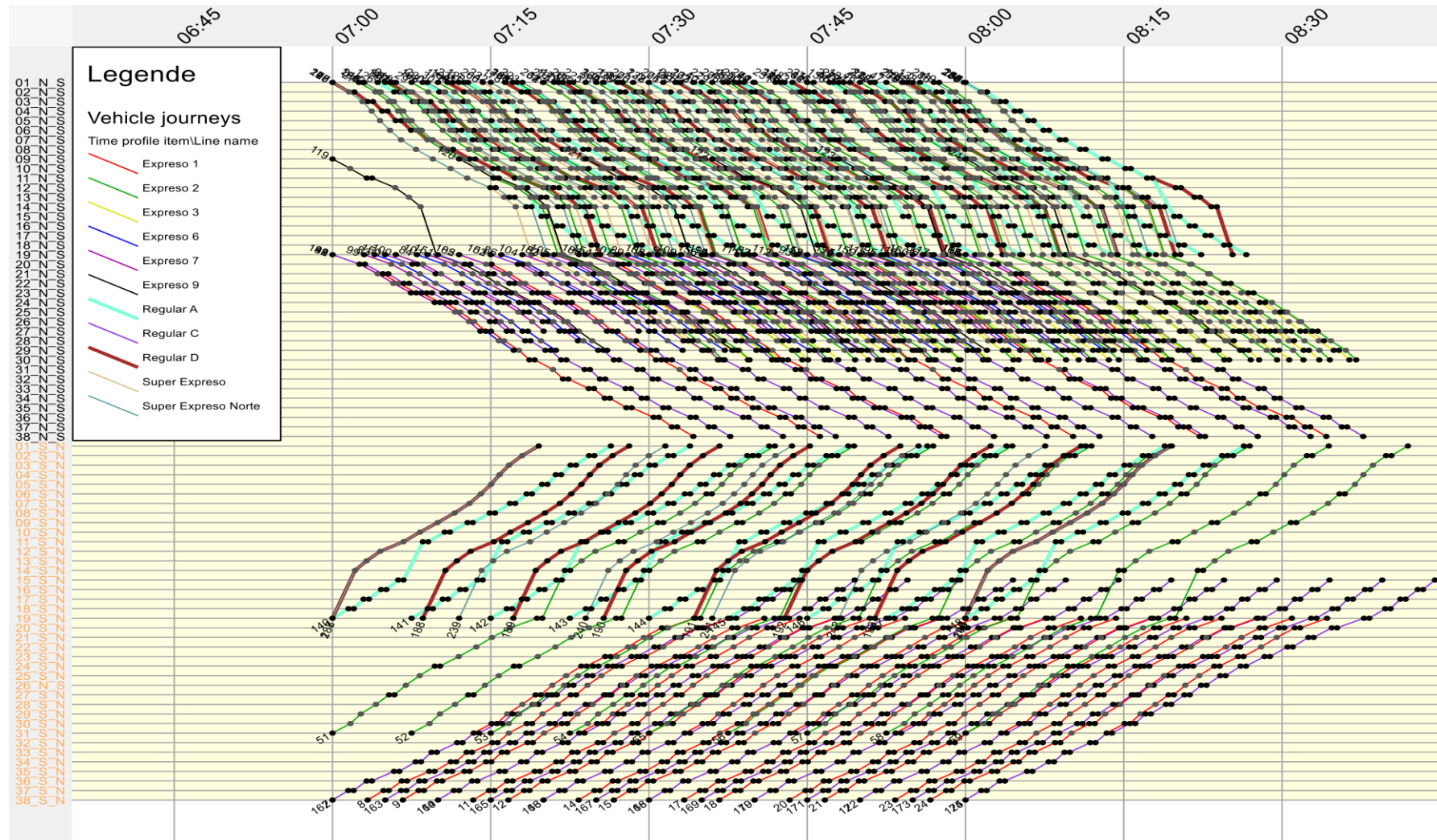


Figura 89 Perfil de Tiempo - Actual
 Fuente: Elaboración propia.

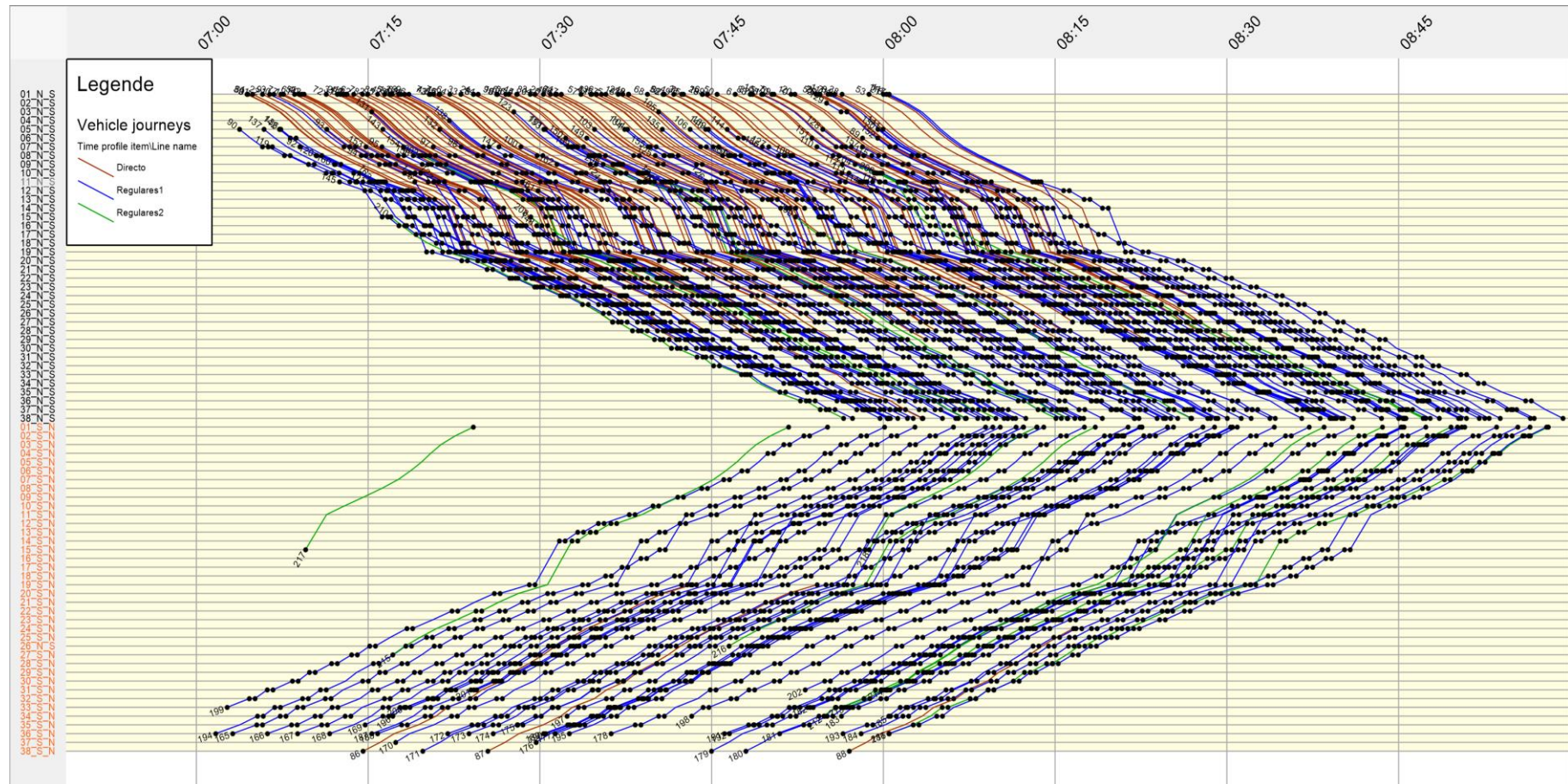


Figura 90 Perfil de Tiempo - Propuesta
 Fuente: Elaboración propia

6.2.4 Número de transferencias por estación

En las Figura 91y Figura 92 se muestra cómo se da el número de transferencias en cada estación tanto de la simulación actual como de la simulación propuesta respectivamente. Así mismo con estos valores se ha elaborado unos gráficos comparativos si analizar en cuanto mejora la situación propuesta respecto de la situación actual.

Number: 76	No	Code	Name	TypeNo	NumLines	PassBoard(AP)	PassAlign(AP)	PassOrigin(AP)	PassDestination(AP)	PassTransTotal(AP)	PassThroughStop(AP)	PassThroughNoStop(AP)
1	1	01_N_S	Naranjal	0	6	18051	0	18051	0	0	0	0
2	2	02_N_S	Izagirre	0	3	1859	32	1836	9	23	1604	16415
3	3	03_N_S	Pacifico	0	2	163	5	158	0	5	3458	16415
4	4	04_N_S	Independencia	0	3	695	3	694	2	1	3619	16415
5	5	05_N_S	Los Jazmines	0	2	208	3	205	0	3	4311	16415
6	6	06_N_S	Tomás Valle	0	3	2477	11	2476	10	1	4508	16415
7	7	07_N_S	El Milagro	0	2	342	4	338	0	4	6980	16415
8	8	08_N_S	Honorio Delgado	0	2	138	34	137	33	1	7288	16415
9	9	09_N_S	UNI	0	3	2418	1149	1378	109	1040	6276	16415
10	10	10_N_S	Parque del Trabajo	0	2	93	37	92	36	1	6782	18290
11	11	11_N_S	Caquetá	0	3	397	76	378	57	19	8675	16415
12	12	12_N_S	2 de Mayo	0	2	303	444	229	369	75	6639	14563
13	13	13_N_S	Quilca	0	2	73	288	73	288	0	6654	14563
14	14	14_N_S	España	0	2	170	1408	170	1408	0	5320	14563
15	15	15_N_S	Ramon Castilla	0	2	1716	1054	773	111	943	2787	0
16	16	16_N_S	Tacna	0	2	149	237	140	228	9	4266	0
17	17	17_N_S	Jr.Union	0	2	59	1023	57	1022	2	3392	0
18	18	18_N_S	Colmena	0	2	92	583	91	582	1	2868	0
19	19	19_N_S	Estación Central	0	7	4670	6593	744	2654	3939	1856	14563
20	20	20_N_S	Estadio Nacional	0	2	69	277	68	277	1	3417	17394
21	21	21_N_S	México	0	1	105	283	105	283	0	2865	17731
22	22	22_N_S	Canada	0	2	187	2125	159	2069	56	5368	13208
23	23	23_N_S	Javier Prado	0	5	465	2780	89	2405	376	5943	10040
24	24	24_N_S	Canaval y Moreyra	0	6	670	5209	24	4563	646	6466	4774
25	25	25_N_S	Aramburu	0	2	13	1602	13	1602	0	3075	7232
26	26	26_N_S	Domingo Orué	0	1	15	189	15	189	0	1739	8391
27	27	27_N_S	Angamos	0	6	775	4037	33	3295	742	4299	1809
28	28	28_N_S	Ricardo Palma	0	2	239	1563	22	1346	218	1642	3677
29	29	29_N_S	Benavides	0	4	352	2624	9	2281	343	1951	984
30	30	30_N_S	28 de Julio	0	4	26	690	3	667	23	2596	0
31	31	31_N_S	Plaza de Flores	0	1	8	406	8	406	0	1556	661
32	32	32_N_S	Balta	0	2	14	196	14	195	0	2029	0
33	33	33_N_S	Bulevar	0	2	11	232	11	232	0	1811	0
34	34	34_N_S	Union	0	2	4	158	4	158	0	1664	0
35	35	35_N_S	Escuela Militar	0	2	1	92	1	92	0	1576	0
36	36	36_N_S	Terán	0	2	3	142	3	142	0	1435	0
37	37	37_N_S	Rosario de Villa	0	2	3	176	3	176	0	1262	0
38	38	38_N_S	Matellini	0	2	0	1265	0	1265	0	0	0
39	41	01_S_N	Naranjal	0	4	0	1096	0	1096	0	0	0
40	42	02_S_N	Izagirre	0	3	1	274	1	274	0	983	112
41	43	03_S_N	Pacifico	0	2	1	21	1	21	0	480	888
42	44	04_S_N	Independencia	0	2	1	76	1	76	0	500	888
43	45	05_S_N	Los Jazmines	0	2	3	5	3	5	0	573	888
44	46	06_S_N	Tomás Valle	0	3	2	270	2	270	0	1353	112
45	47	07_S_N	El Milagro	0	2	1	33	1	33	0	645	1088
46	48	08_S_N	Honorio Delgado	0	2	1	125	1	125	0	676	1088
47	49	09_S_N	UNI	0	3	3	400	3	400	0	1775	112
48	50	10_S_N	Parque del Trabajo	0	2	7	26	7	26	0	895	1385
49	51	11_S_N	Caquetá	0	3	22	71	16	65	6	2172	112
50	52	12_S_N	2 de Mayo	0	1	35	107	35	107	0	761	1432
51	53	13_S_N	Quilca	0	1	24	43	24	43	0	844	1432
52	54	14_S_N	España	0	2	164	475	164	475	0	2043	112
53	55	15_S_N	Ramon Castilla	0	2	19	248	19	248	0	108	0
54	56	16_S_N	Tacna	0	2	23	131	23	131	0	333	0
55	57	17_S_N	Jr.Union	0	2	14	592	14	592	0	450	0
56	58	18_S_N	Colmena	0	2	14	275	14	275	0	1028	0
57	59	19_S_N	Estación Central	0	6	1186	2566	178	1572	1008	2746	0
58	60	20_S_N	Estadio Nacional	0	2	174	692	48	566	126	3199	1940
59	61	21_S_N	México	0	1	96	232	96	232	0	1408	4326
60	62	22_S_N	Canada	0	1	230	306	203	306	27	1410	4326
61	63	23_S_N	Javier Prado	0	2	193	409	193	409	0	3910	1940
62	64	24_S_N	Canaval y Moreyra	0	3	343	749	315	722	28	5916	0
63	65	25_S_N	Aramburu	0	1	46	204	46	204	0	1743	4877
64	66	26_N_S	Domingo Orué	0	1	70	53	70	53	0	1877	4877
65	67	27_S_N	Angamos	0	3	1301	734	770	202	532	5506	0
66	68	28_S_N	Ricardo Palma	0	1	138	113	138	113	0	1696	4406
67	69	29_S_N	Benavides	0	1	295	103	295	103	0	1513	4406
68	70	30_S_N	28 de Julio	0	2	123	65	95	36	29	5070	828
69	71	31_S_N	Plaza de Flores	0	2	1131	231	923	23	208	1318	3514
70	72	32_S_N	Balta	0	2	157	46	156	46	0	4907	0
71	73	33_S_N	Bulevar	0	2	225	45	225	45	0	4728	0
72	74	34_S_N	Union	0	2	248	37	248	37	0	4525	0
73	75	35_S_N	Escuela Militar	0	2	45	19	45	19	0	4517	0
74	76	36_S_N	Terán	0	2	361	2	361	2	0	4175	0
75	77	37_S_N	Rosario de Villa	0	2	320	5	320	5	0	3857	0
76	78	38_S_N	Matellini	0	2	3862	0	3862	0	0	0	0

Figura 91 Transferencia en estación-Actual

Fuente: Elaboración propia

Number	No	Code	Name	NumLines	PassBoard(AP)	PassAlign(AP)	PassOrigin(AP)	PassDestination(AP)	PassTransTotal(AP)	PassThroughStop(AP)	PassThroughNoStop(AP)
4	4	04_N_S	Independencia	1	223	0	223	0	0	0	19690
5	5	05_N_S	Los Jazmines	1	976	0	976	0	0	0	19913
6	6	06_N_S	Tomás Valle	2	1849	3	1849	3	0	37	20849
7	7	07_N_S	El Milagro	1	1173	139	1034	0	139	2517	20078
8	8	08_N_S	Honorio Delgado	2	255	142	150	36	106	4106	19521
9	9	09_N_S	UNI	2	1346	116	1336	106	10	1672	22094
10	10	10_N_S	Parque del Trabajo	2	196	114	119	36	78	5023	19976
11	11	11_N_S	Caquetá	2	716	350	437	72	279	7784	17061
12	12	12_N_S	2 de Mayo	1	319	404	277	362	42	5715	14784
13	13	13_N_S	Quilca	2	188	351	117	280	71	5887	14579
14	14	14_N_S	España	3	660	1642	435	1416	226	5578	13435
15	15	15_N_S	Ramon Castilla	2	712	150	685	123	27	3337	1170
16	16	16_N_S	Tacna	2	222	293	142	212	80	3612	1315
17	17	17_N_S	Jr.Union	3	279	1300	57	1079	222	3577	272
18	18	18_N_S	Colmena	2	90	521	88	517	4	3308	298
19	19	19_N_S	Estación Central	3	1495	3522	391	2444	1114	8380	11467
20	20	20_N_S	Estadio Nacional	2	95	561	67	533	28	9318	11463
21	21	21_N_S	México	2	106	478	103	475	3	7988	12409
22	22	22_N_S	Canada	3	618	2358	162	1901	456	8405	9741
23	23	23_N_S	Javier Prado	3	655	2968	89	2402	566	7581	8215
24	24	24_N_S	Canaval y Moreyra	3	814	5236	24	4446	790	5225	5225
25	25	25_N_S	Aramburu	3	292	1867	13	1588	279	5640	4522
26	26	26_N_S	Domingo Orué	2	16	265	16	265	1	5096	5092
27	27	27_N_S	Angamos	3	741	4135	33	3427	708	3979	2089
28	28	28_N_S	Ricardo Palma	3	334	1572	22	1261	312	3754	1483
29	29	29_N_S	Benavides	3	329	2639	10	2321	319	2648	283
30	30	30_N_S	28 de Julio	2	11	346	2	337	9	2617	297
31	31	31_N_S	Plaza de Flores	2	16	656	8	648	8	2077	193
32	32	32_N_S	Balta	2	15	258	14	257	1	1538	490
33	33	33_N_S	Bulevar	2	22	246	11	235	11	1315	482
34	34	34_N_S	Union	2	9	159	4	155	5	973	686
35	35	35_N_S	Escuela Militar	2	8	99	1	93	7	658	912
36	36	36_N_S	Terán	2	12	152	3	143	9	738	687
37	37	37_N_S	Rosario de Villa	2	14	185	3	174	11	684	568
38	38	38_N_S	Matellini	3	0	1266	0	1266	0	0	0
39	41	01_S_N	Naranjal	2	0	1096	0	1096	0	0	0
40	42	02_S_N	Izazuirre	2	1	262	1	262	0	776	319
41	43	03_S_N	Pacifico	2	1	32	1	31	0	678	678
42	44	04_S_N	Independencia	2	21	126	1	106	19	1044	323
43	45	05_S_N	Los Jazmines	2	3	7	3	7	0	330	1159
44	46	06_S_N	Tomás Valle	2	4	244	2	242	2	1276	217
45	47	07_S_N	El Milagro	2	1	42	1	42	0	1005	731
46	48	08_S_N	Honorio Delgado	2	1	173	1	173	0	1412	365
47	49	09_S_N	UNI	2	7	347	3	343	4	1726	217
48	50	10_S_N	Parque del Trabajo	2	7	55	7	55	0	1607	676
49	51	11_S_N	Caquetá	2	26	75	17	66	9	1700	612
50	52	12_S_N	2 de Mayo	2	35	154	35	154	0	1188	104
51	53	13_S_N	Quilca	2	26	138	24	136	2	1377	42
52	54	14_S_N	España	2	186	657	168	640	18	1372	0
53	55	15_S_N	Ramon Castilla	2	43	239	22	218	21	979	39
54	56	16_S_N	Tacna	2	20	128	20	128	0	1139	98
55	57	17_S_N	Jr.Union	2	20	566	14	560	6	1345	0
56	58	18_S_N	Colmena	2	15	254	14	254	1	1819	77
57	59	19_S_N	Estación Central	3	503	1688	158	1317	381	3481	195
58	60	20_S_N	Estadio Nacional	2	49	532	44	527	5	4091	1225
59	61	21_S_N	México	2	108	216	103	210	6	4251	1489
60	62	22_S_N	Canada	2	256	325	225	294	31	4678	1021
61	63	23_S_N	Javier Prado	2	222	396	215	389	7	4344	1458
62	64	24_S_N	Canaval y Moreyra	3	488	1007	233	752	255	4368	1341
63	65	25_S_N	Aramburu	2	84	207	81	204	3	4727	1906
64	66	26_S_N	Domingo Orué	2	227	125	155	53	72	4941	1673
65	67	27_S_N	Angamos	2	605	209	595	200	10	3166	2968
66	68	28_S_N	Ricardo Palma	2	212	138	189	115	23	4439	1692
67	69	29_S_N	Benavides	2	375	128	350	103	25	3951	1943
68	70	30_S_N	28 de Julio	2	181	112	105	36	76	3475	2366
69	71	31_S_N	Plaza de Flores	2	970	102	889	21	81	1577	3407
70	72	32_S_N	Balta	2	174	52	170	48	4	2337	2574
71	73	33_S_N	Bulevar	2	346	157	234	45	112	3154	1464
72	74	34_S_N	Union	2	335	123	249	37	86	2957	1483
73	75	35_S_N	Escuela Militar	1	79	55	45	19	36	1919	2564
74	76	36_S_N	Terán	2	676	0	676	0	0	0	3862
75	77	37_S_N	Rosario de Villa	1	383	0	383	0	0	0	3478
76	78	38_S_N	Matellini	2	3478	0	3478	0	0	0	0

Figura 92 Transferencia en estación-Propuesta
Fuente: Elaboración propia

En la Figura 93 se muestra una comparación de la cantidad de transferencias en cada estación, el color azul muestra la situación actual y el color naranja la situación propuesta a lo largo de las 38 estaciones. Los valores de la simulación actual son mucho mayores que en la simulación propuesta, esto se ve más reflejado en la dirección de Norte a Sur en especial en la estación central (zona 19) teniendo pico de 3939.3 en la situación actual y un pico 1113.5 en la

situación propuesta; por otro lado, en la dirección de Sur a Norte el pico más alto también se da en la estación central con 1008.3 en la situación actual y 308.7 en la situación propuesta.

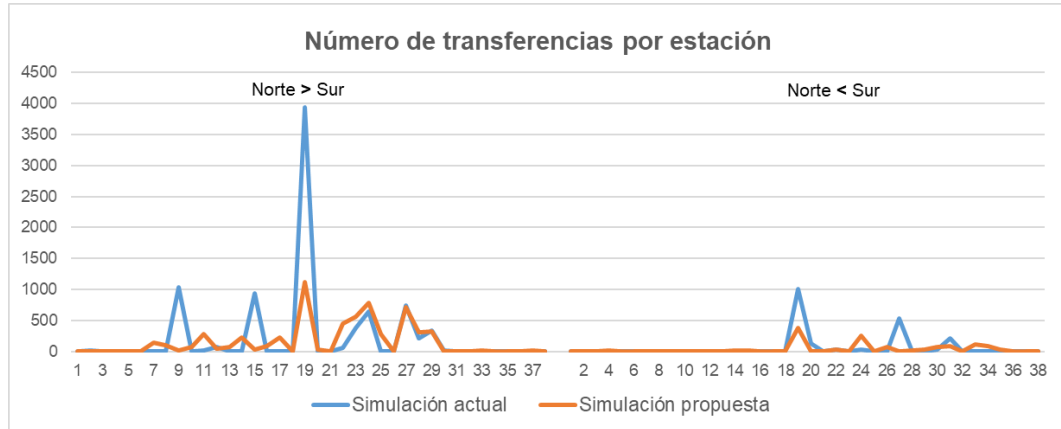


Figura 93 N° de transferencias por estación
Fuente: Elaboración propia.

En Figura 94 se hace una comparación de la situación actual con la propuesta, se ve que el número de transferencias en la situación actual es de 10433.3 y en la situación propuesta es de 7213.8, el ahorro de transferencia es de 3219.4, es decir se tiene una disminución del 44.6%. También se ve las transferencias en función de la dirección por ejemplo en la situación actual de los 10433.2 transferencias, el 81.2% son de transferencias debido a los viajes realizados de Norte a Sur y el 18.8% pertenece a los de Sur a Norte.

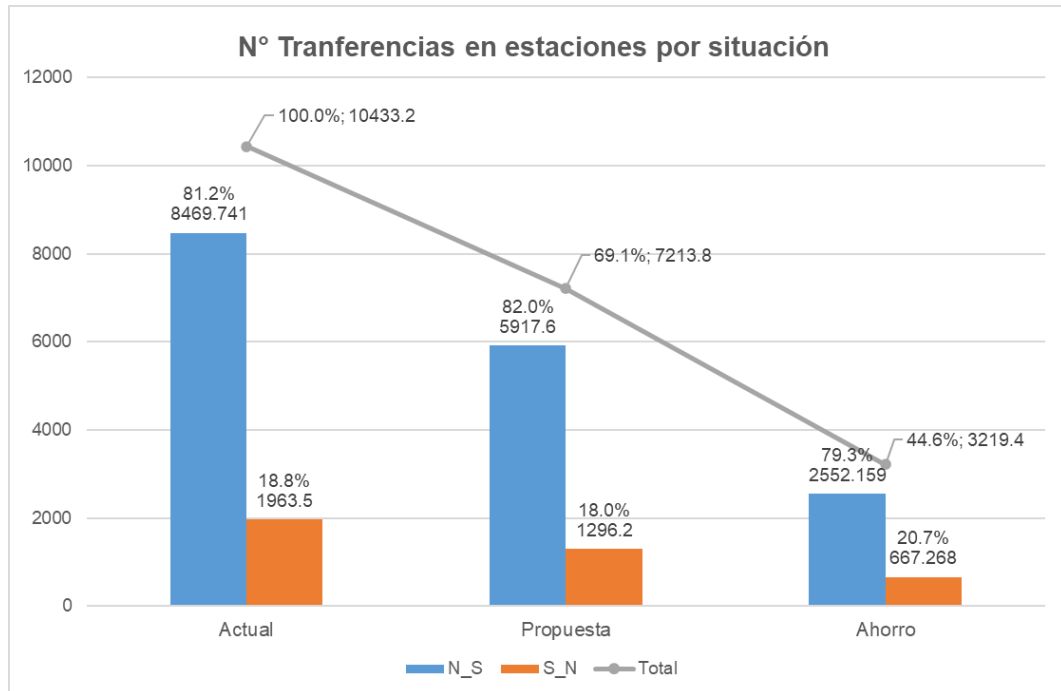


Figura 94 Comparación del N° de transferencias por estación
Fuente: Elaboración propia

CONCLUSIONES

1. En los datos analizados del Sistema de Transporte Masivo de Buses de Transporte Rápido “Metropolitano”, en la data sin procesar, es decir en el proceso de limpieza de datos, las validaciones con perfil PL (pases libres) representa el 2.35%, el cual es 14 421 validaciones diarias.
2. En el “Metropolitano”, el comportamiento es diferente de lunes a viernes que los días sábados y domingos-feriados. El modelo de Machine Learning tiene como datos de entrada a los atributos “tarjeta”, “perfil”, “tipo de validador”, “paradero”, “monto” y “fecha”; como salida o predicción la hora de ingreso al sistema de transporte; es por ello que para tener mejores resultados de exactitud (accuracy) se usó la data de lunes a viernes. Dando como mejor modelo de predicción para este tipo de datos, el Modelo de BayesBayesNet con una exactitud de 51.76%, usando 22 días de validaciones, aproximadamente 11 millones de validaciones y con 20 intervalos en el atributo “Hora”.
3. El modelo de “Generación de rutas óptimas” elaborado en esta tesis, tiene como datos de entrada del usuario la hora de ingreso, origen y destino, es por ello que este modelo puede ser usado en redes transporte con pocos o grandes datos, es decir puede ser aplicado en casos específicos o en toda una ciudad. Por lo tanto, el modelo de “Generación de rutas óptimas” puede ser aplicado en cualquier sistema de transporte.
4. En el caso de estudio, “Metropolitano”, al generar las rutas con el algoritmo de modelo de “Generación de rutas óptimas” elaborado en esta tesis, da como resultado 347 viajes en la hora pico de 7 a 8am; lo que implica una mejora en 51 viajes comparado con los 398 viajes que usa el sistema del metropolitano.
5. El modelo de “Generación de rutas óptimas” elaborado en esta tesis, tienes como datos de entrada del operador la velocidad promedio de los buses, coordenadas de los paraderos y el tiempo máximo de espera en las estacione de origen de un usuario. En el caso del “Metropolitano”, al usar el modelo de “Generación de rutas óptimas” se establece como tiempo máximo de espera 15 min, por otro lado, al realizar la toma de datos en campo el tiempo máximo de espera es de 45min; esto indica que al usar el modelo de “Generación de rutas óptimas” el tiempo de

espera disminuya en un 66.6%, es decir pase de 45 min a como máximo 15 min.

6. La simulación valida que el modelo de "Generación de rutas óptimas" es más eficiente en la distribución de salida de los buses, reduciendo en 83.44% en el tiempo de transferencia y 69.06% en el número de transferencias.

RECOMENDACIONES

1. Los usuarios con perfil tipo Pase Libre (PL), deben obtener sus propias tarjetas, así se podrá tener los orígenes destinos de estos usuarios. Este tipo de perfil también afecta en el modelo de predicción ya que genera errores (ruido).
2. En la data analizada del caso de estudio, se observó que el comportamiento del domingo es similar al de un día feriado, para verificar esta tendencia se debe de analizar data de al menos un año.
3. Para el desarrollo de Big Data similares al caso de estudio, se debe utilizar framework en tiempo real de manera paralela y distribuida (framework Flink).
4. Para usar las validaciones de los sábados y domingos-feriados en los modelos de predicción de Machine Learning, se sugiere usar la data de al menos de un año.
5. En la elaboración de casos estudios similares usando el modelo de “Generación de rutas óptimas”, se debe usar la data de todos los usuarios del sistema analizado.
6. El comportamiento de los usuarios en las estaciones más importantes se puede analizar de manera detallada mediante modelos de meso simulación.
7. El modelo de “Generación de rutas óptimas” se basa en la demanda de los usuarios, es decir es un modelo que usa los conceptos de movilidad urbana sostenible, ya que prioriza al usuario. Así mismo se debe hacer la integración con los otros sistemas de transporte no motorizados como son los ciclistas y peatones de manera que se fortalezca los conceptos de movilidad de transporte urbano sostenible, que ya de por si es de uso obligatorio en el país.
8. En esta tesis, en el caso de estudio se utilizó principalmente el origen destino de los usuarios, queda por analizar los otros parámetros de las validaciones por ejemplo los costos; así mismo se recomienda comparar este algoritmo de “Generación de rutas” con otros tipos de obras similares.

REFERENCIAS BIBLIOGRAFICAS

1. Agencia de cooperación Internacional del Japón [JICA]. (2013). *Encuesta de Recolección de Información Básica del Transporte Urbano en el Área Metropolitana de Lima y Callao*. Recuperado de [https://portal.mtc.gob.pe/estadisticas/files/estudios/01 - EI JR13012 Resumen.pdf](https://portal.mtc.gob.pe/estadisticas/files/estudios/01_EI_JR13012_Resumen.pdf)
2. Agencia de Cooperación Internacional del Japón [JICA]. (2004). *Plan Maestro de Transporte Urbano para el area Metropolitana de Lima y Callao en la República del Perú*. Recuperado de http://open_jicareport.jica.go.jp/pdf/11798261_01.pdf
3. Alsuwaiyel, M. H. (2016). *Algorithms: Design Techniques and Analysis* (2.^a ed.; New Jersey: World Scientific, Ed.). Recuperado de <http://educacioncreativa.org/OMINL/Libros/Libros Algoritmos/ALGORITHMMS DESIGN TECHNIQUES AND ANALYSIS.pdf>
4. Berman, F., Stodden, V., Szalay, A. S., Rutenbar, R., Hailpern, B., Christensen, H., ... Raghavan, P. (2018). Realizing the potential of data science. *Communications of the ACM*, 61(4), 67-72. <https://doi.org/10.1145/3188721>
5. Chong, D., & Shi, H. (2015). Big data analytics: a literature review. *Journal of Management Analytics*, 2(3), 175-201. <https://doi.org/10.1080/23270012.2015.1082449>
6. Congreso de la República. (2018, diciembre 28). Ley N° 30900: «Ley que Crea la Autoridad de Transporte Urbano para Lima y Callao». *El Peruano*, pp. 16-21. Recuperado de <https://busquedas.elperuano.pe/download/url/ley-que-crea-la-autoridad-de-transporte-urbano-para-lima-y-c-ley-n-30900-1727064-7>
7. Deutsche Gesellschaft für Internationale Zusammenarbeit [GIZ]. (2014). *TRANSPerú-Peru-Sustainable Urban Transport NAMA*. Recuperado de http://transferproject.org/wp-content/uploads/2014/04/Overview_PERU_TRANSPerú.pdf
8. Deutsche Gesellschaft für Internationale Zusammenarbeit [GIZ]. (2016). *Autoridad de Transporte Urbano para Lima y el Callao: Propuesta*

- conceptual*. Recuperado de <https://portal.mtc.gob.pe/estadisticas/Autoridad de Transporte Urbano para Lima y Callao-144dpi.pdf>
9. Deutsche Gesellschaft für Internationale Zusammenarbeit [GIZ]. (2018). *Sistemas de transporte eficientes, efectivos y sostenibles TRANSPerú*. Recuperado de www.bmu.bund.de
10. Doersing, N. (2016). *Pseudocode, the Pseudocode Programming Process, and Alternatives to the PPP*. Recuperado de <http://ps.informatik.uni-tuebingen.de/teaching/ss15/sct/StudentMaterial/09 - Noah Doersing - pseudocode.pdf>
11. Florian, M., Gaudry, M. J. I., & Lardinois, C. (1988). A two-dimensional framework for the understanding of transportation planning models. *Transportation Research B*, 22B, 411-419. [https://doi.org/10.1016/0191-2615\(88\)90022-7](https://doi.org/10.1016/0191-2615(88)90022-7)
12. Gerencia de Transporte Urbano [GTU]. (2019a). *Nombres de las Rutas de Transporte Público*. Recuperado de <http://www.gtu.munlima.gob.pe/sites/default/files/publicaciones/SRT/CRONOGRAMA DE ENTREGA DE TUC.pdf>
13. Gerencia de Transporte Urbano [GTU]. (2019b). *Padrón Total de Unidades de Transporte Público*. LIMA.
14. Giesen, R., Delgado, F., & De Dios Ortuzar, J. (2019). Análisis de Sistemas de Transporte. Recuperado 5 de junio de 2019, de Pontificia Universidad Católica de Chile | Coursera website: <https://www.coursera.org/learn/analisis-sistemas-de-transporte/lecture/ATqle/modelo-clasico-de-transporte-4-etapas>
15. Gollapudi, S. L. V. (2016). *Practical Machine Learning Practical* (UK:Packt Publishing, Ed.). Recuperado de www.PacktPub.com
16. Hernández-Leal, E. J., Duque-Méndez, N. D., Moreno-Cadavid, J., Hernández-Leal, E. J., Duque-Méndez, N. D., Moreno-Cadavid, J., & Big, ". (2017). Big Data: una exploración de investigaciones, tecnologías y casos de aplicación. *TecnoLógicas*, 20(39), 24. Recuperado de <http://www.scielo.org.co/pdf/teclo/v20n39/v20n39a02.pdf>
17. Hicks, S. C., & Irizarry, R. A. (2018). A Guide to Teaching Data Science. *American Statistician*, 72(4), 382-391.

<https://doi.org/10.1080/00031305.2017.1356747>

18. Isoni, A., & Cervellin, D. (2016). Introduction to Practical Machine Learning Using Python. En Packt Publishing. (Ed.), *Machine Learning for the Web* (p. 298). Recuperado de <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1295356&site=eds-live>
19. Kelleher, J. D., & Tierney, B. (2018). *Data Science* (The MIT Pr; A. Borsuk, Ed.). Recuperado de <http://search.ebscohost.com/login.aspx?direct=true&db=edsee&AN=edsee.8544210&site=eds-live>
20. Lanegra, I., Zevallos, P., Mora, J., & Apacla, K. (2016). *Gobernando con el Clima a Favor*. Recuperado de Libélula website: <http://libelula.com.pe/wp-content/uploads/2016/03/Gobernando-con-el-clima-a-favor-pantalla.pdf>
21. Liu, A. (2015). *Data Science and Data Scientist*. Recuperado de <http://www.researchmethods.org/DataScienceDataScientists.pdf>
22. Lloyd, W. [ITDP], & Karl, F. [GTZ]. (2006). *Opciones de Transporte Público Masivo* (Deutsche G; Manfred Breithaupt, Stefan Opitz, Karl Fjellstrom, & Jan Schwaab, Eds.). Recuperado de <http://www.sutp.org>
23. Mancilla Herrera, A., Ebratt Gómez, R., & Capacho Portilla, J. (2014). Marco conceptual de la Lógica de Programación. En Universidad del Norte (Ed.), *Diseño y construcción de algoritmos* (p. 440). Bogota.
24. Manheim, M. L. (1979). *Fundamentals of transportation systems analysis*. MIT Press.
25. Manyika, J., Chui, M., Brown, B., Bughin Jacques, Dobbs Richard, Roxburgh Charles, & Hung Angela. (2011). *Big data: The next frontier for innovation, competition, and productivity*. Recuperado de www.mckinsey.com/mgi.
26. McNally, M. G. (2008). The Four Step Model. En *Handbook of Transport Modeling* (2nd ed., p. 19). Recuperado de <http://www.its.uci.edu>
27. Metropolitano. (2019). Metropolitano. Recuperado 21 de junio de 2019, de PROTRANSPORTE website: <http://www.metropolitano.com.pe>
28. OpenStreetMap. (2019). Nodo: Estación Naranjal (4682629992). Recuperado 10 de julio de 2019, de <https://www.openstreetmap.org/node/4682629992>

29. Paul C. Zikopoulos, Chris Eaton, Dirk deRoos, Thomas Deutsch, & George Lapis. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data* (McGraw–Hill; Jeff Weeks, Ed.). Recuperado de <http://it.toolbox.com/blogs/db2luw>.
30. Prieto Espinosa, A. (2018). ¿Qué es la Ciencia de Datos, Minería de Datos y Big Data? Recuperado 12 de junio de 2019, de Universidad de Granada website: <https://www.youtube.com/watch?v=UiO0WYCgkTI>
31. PTV-VISUM. (2020, septiembre 10). Visum Fundamentals-Timetable-based assignment. Recuperado 25 de enero de 2021, de © 2020 PTV AG, Karlsruhe, Alemania website: https://cgi.ptvgroup.com/vision-help/VISUM_2021_ENG/Content/1_Benutzermodell_OeV/1_6_Fahrplanfeine_Umlegung.htm#1_benutzermodell_oev_2318507456_1027323
32. PTV GROUP. (2019). Visum. Recuperado 24 de junio de 2019, de <http://vision-traffic.ptvgroup.com/es/productos/ptv-visum/>
33. Ramos, A., Sánchez, P., Ferrer, J. M., Barquín, J., & Linares, P. (2010). *Modelos Matemáticos de Optimización* (Segunda; Universidad Pontificia Comillas, Ed.). Recuperado de https://www.gams.com/fileadmin/community/contrib/doc/modelado_en_gams.pdf
34. Rodríguez, C., García, M., Daniel, J., Frías, M., Sadot, M., & Fernández, A. (2006). *RFID: La tecnología de identificación por radiofrecuencia*. Recuperado de <https://repositorio.comillas.edu/jspui/bitstream/11531/5321/1/IIT-05-035A.pdf>
35. Ruiz José, Aguado Andrés, J. J. (2012). *Optimizan de Rutas de Transporte*. Recuperado de https://eprints.ucm.es/23027/1/Memoria_OptimizacionRutasTransporte.pdf
36. Transit Cooperative Research Program [TCRP]. (2003). *Transit Capacity and Quality of Service Manual-2nd Edition Kittelson & Associate*. Recuperado de <http://imenteraddod.com/wp-content/uploads/2017/07/460-Transit-Capacity-and-Quality-of-Service-Manual-National-Research-Council-0309087767-Transport.pdf>
37. Wu, S. M., Chen, T. chun, Wu, Y. J., & Lytras, M. (2018). Smart cities in Taiwan: A perspective on big data applications. *Sustainability (Switzerland)*,

10(1). <https://doi.org/10.3390/su10010106>

38. Yang, L. (2018). Modeling the mobility choices of older people in a transit-oriented city: Policy insights. *Habitat International*, 76, 10-18. <https://doi.org/10.1016/j.habitatint.2018.05.007>

ANEXOS**ANEXO 1: Validaciones de los usuarios del Metropolitano**

En la Figura 95 se muestra los archivos de las validaciones del mes de mayo del 2018 que fueron brindados por Protransporte, los cuales están en formato tipo CSV.
































Nombre	Fecha de modifica...	Tipo	Tamaño
 validaciones_diarias_pt_01052018	18/05/2018 13:00	Archivo de valores se...	26.323 KB
 validaciones_diarias_pt_02052018	18/05/2018 13:01	Archivo de valores se...	84.170 KB
 validaciones_diarias_pt_03052018	18/05/2018 13:02	Archivo de valores se...	84.090 KB
 validaciones_diarias_pt_04052018	26/10/2018 7:27	Archivo de valores se...	85.620 KB
 validaciones_diarias_pt_05052018	18/05/2018 13:03	Archivo de valores se...	59.475 KB
 validaciones_diarias_pt_06052018	18/05/2018 13:03	Archivo de valores se...	25.805 KB
 validaciones_diarias_pt_07052018	14/05/2018 8:01	Archivo de valores se...	83.523 KB
 validaciones_diarias_pt_08052018	14/05/2018 8:02	Archivo de valores se...	83.908 KB
 validaciones_diarias_pt_09052018	14/05/2018 8:03	Archivo de valores se...	84.375 KB
 validaciones_diarias_pt_10052018	14/05/2018 8:04	Archivo de valores se...	84.207 KB
 validaciones_diarias_pt_11052018	14/05/2018 8:05	Archivo de valores se...	82.557 KB
 validaciones_diarias_pt_12052018	14/05/2018 8:06	Archivo de valores se...	57.242 KB
 validaciones_diarias_pt_13052018	14/05/2018 8:06	Archivo de valores se...	22.574 KB
 validaciones_diarias_pt_14052018	18/05/2018 13:04	Archivo de valores se...	80.738 KB
 validaciones_diarias_pt_15052018	18/05/2018 13:05	Archivo de valores se...	82.825 KB
 validaciones_diarias_pt_16052018	18/05/2018 13:06	Archivo de valores se...	83.537 KB
 validaciones_diarias_pt_17052018	18/05/2018 8:16	Archivo de valores se...	82.515 KB
 validaciones_diarias_pt_18052018	21/05/2018 6:53	Archivo de valores se...	84.174 KB
 validaciones_diarias_pt_19052018	21/05/2018 6:50	Archivo de valores se...	58.886 KB
 validaciones_diarias_pt_20052018	21/05/2018 6:50	Archivo de valores se...	26.453 KB
 validaciones_diarias_pt_21052018	28/05/2018 8:05	Archivo de valores se...	83.057 KB
 validaciones_diarias_pt_22052018	28/05/2018 8:06	Archivo de valores se...	83.612 KB
 validaciones_diarias_pt_23052018	28/05/2018 8:07	Archivo de valores se...	83.798 KB
 validaciones_diarias_pt_24052018	28/05/2018 8:08	Archivo de valores se...	83.837 KB
 validaciones_diarias_pt_25052018	28/05/2018 8:09	Archivo de valores se...	83.979 KB
 validaciones_diarias_pt_26052018	28/05/2018 8:09	Archivo de valores se...	58.751 KB
 validaciones_diarias_pt_27052018	19/06/2018 12:37	Archivo de valores se...	25.987 KB
 validaciones_diarias_pt_28052018	04/06/2018 7:06	Archivo de valores se...	83.766 KB
 validaciones_diarias_pt_29052018	04/06/2018 7:07	Archivo de valores se...	82.411 KB
 validaciones_diarias_pt_30052018	19/06/2018 12:38	Archivo de valores se...	83.844 KB
 validaciones_diarias_pt_31052018	04/06/2018 7:09	Archivo de valores se...	83.284 KB

Figura 95 Datos brindados por Protransporte

Fuente: Elaboración propia

En la Figura 96 se muestra la documentación, donde responden a la solicitud que se le envió, respondiendo favorablemente. En el CD brindan las validaciones de los usuarios del mes de mayo del 2018.

 **Municipalidad de Lima**

PROTRANSPORTE
INSTITUTO METROPOLITANO PROTRANSPORTE DE LIMA

"Año del Diálogo y la Reconciliación Nacional"

Lima, 09 JUL. 2018

CARTA N° 395-2018-MML/IMPL/GOC

Egresada - UNI
BETZY TAPIA QUISPITUPA
Tesis FIC2018
Código 20121024B

Presente -

Ref: Oficio N° 001/TesistaFIC2018

De mi mayor consideración:

Me dirijo a usted, para saludarlo cordialmente y comunicarle en atención al documento de la referencia, mediante el cual solicita información y una reunión para realizar las consultas en relación a la operación y la programación del servicio del Metropolitano para el desarrollo de tesis respecto a Transporte Urbano – BRT.

Al respecto, de acuerdo a la reunión realizada el jueves 28 de julio del presente, se realizó las consultas del caso y se adjunta un CD con la información de validaciones del sistema para la elaboración de una matriz Origen – Destino, el cual deberá ser compartido con nuestra institución.

Sin otro particular, quedo de usted.

Atentamente,


EDWIN SANCHEZ SANCHEZ
Jefe de Unidad de Planeación y Programación

MUNICIPALIDAD METROPOLITANA DE LIMA
INSTITUTO METROPOLITANO PROTRANSPORTE DE LIMA
Gerencia de Operaciones del COSAC


FEDOR DANIEL MORALES BOLUARTE
GERENTE

HR: 7923

Instituto Metropolitano Protransporte de Lima / Jr. Cuzco 286 –Lima / Central: 428-3333

Figura 96 Documento de entrega de datos
Fuente: Documentación respondida por Protransporte.

ANEXO 2: Ruta y estaciones del Metropolitano vía Troncal

Los siguientes datos son las coordenadas del Metropolitano, son de 40 estaciones debido a que las estaciones Jiron de la Unión y Plaza de flores tiene 2 estaciones cada una. Ver Tabla 39.

Tabla 39 Coordenadas de las estaciones del Metropolitano

COORDENADAS DEL METROPOLITANO REAL			
Nº	Estación	Latitud	Longitud
1	Naranjal	-11,9827434	-77,0586044
2	Izaguirre	-11,9895485	-77,0570156
3	Pacífico	-11,9948804	-77,0559897
4	Independencia	-11,9979702	-77,0552882
5	Los Jazmines	-12,0018372	-77,0548151
6	Tomás Valle	-12,0071044	-77,0538935
7	El Milagro	-12,0112957	-77,0529446
8	Honorio Delgado	-12,0179012	-77,0514099
9	UNI	-12,0233577	-77,0499812
10	Parque del Trabajo	-12,0295698	-77,0442306
11	Caquetá	-12,0365329	-77,0436359
12	2 de Mayo	-12,0463973	-77,0427601
15	Ramón Castilla	-12,0438334	-77,0417289
16	Tacna	-12,0461285	-77,0379041
17	Jirón de la Unión	-12,0484706	-77,0340798
18	Jirón de la Unión	-12,0493408	-77,0326892
19	Colmena	-12,0522640	-77,0328736
13	Quilca	-12,0522011	-77,0422599
14	España	-12,0576357	-77,0417886
20	Estación Central	-12,0576764	-77,0359753
21	Estadio Nacional	-12,0685302	-77,0321386
22	México	-12,0761016	-77,0292237
23	Canadá	-12,0825564	-77,0265013
24	Javier Prado	-12,0893704	-77,0235320
25	Canaval Moreyra	-12,0966099	-77,0249873
26	Aramburú	-12,1026563	-77,0273261
27	Domingo Orué	-12,1084679	-77,0264329
28	Angamos	-12,1126430	-77,0260065
29	Ricardo Palma	-12,1189574	-77,0259296
30	Benavides	-12,1251887	-77,0241960
31	28 de Julio	-12,1293792	-77,0228415
32	Plaza las Flores	-12,1360918	-77,0187605
33	Plaza las Flores	-12,1357091	-77,0186979
34	Balta	-12,1393886	-77,0178131
35	Bulevar	-12,1483917	-77,0201145
36	Estadio Unión	-12,1529051	-77,0197335
37	Escuela Militar	-12,1597668	-77,0189703
38	Terán	-12,1686435	-77,0186235
39	Rosario de Villa	-12,1730247	-77,0151099
40	Matellini	-12,1782643	-77,0104387

Fuente:Elaboración propia.

La Tabla 40 son los valores de las coordenadas que serán utilizadas para el modelo de generación de ruta, se modificó la Tabla 39 sacando los promedios de las dos estaciones Jirón de la Unión y de las dos estaciones de Plaza de Flores, así obtener una sola estación que represente a cada una de ellas, ello se realiza para facilitar el modelado.

Tabla 40 Coordenadas para el modelo del Metropolitano

Nº	Estación	Latitud	Longitud
1	Naranjal	-11,9827434	-77,0586044
2	Izaguirre	-11,9895485	-77,0570156
3	Pacífico	-11,9948804	-77,0559897
4	Independencia	-11,9979702	-77,0552882
5	Los Jazmines	-12,0018372	-77,0548151
6	Tomás Valle	-12,0071044	-77,0538935
7	El Milagro	-12,0112957	-77,0529446
8	Honorio Delgado	-12,0179012	-77,0514099
9	UNI	-12,0233577	-77,0499812
10	Parque del Trabajo	-12,0295698	-77,0442306
11	Caquetá	-12,0365329	-77,0436359
12	2 de Mayo	-12,0463973	-77,0427601
15	Ramón Castilla	-12,0438334	-77,0417289
16	Tacna	-12,0461285	-77,0379041
17	Jirón de la Unión	-12,0489057	-77,0333845
18	Colmena	-12,0522640	-77,0328736
13	Quilca	-12,0522011	-77,0422599
14	España	-12,0576357	-77,0417886
19	Estación Central	-12,0576764	-77,0359753
20	Estadio Nacional	-12,0685302	-77,0321386
21	México	-12,0761016	-77,0292237
22	Canadá	-12,0825564	-77,0265013
23	Javier Prado	-12,0893704	-77,0235320
24	Canaval Moreyra	-12,0966099	-77,0249873
25	Aramburú	-12,1026563	-77,0273261
26	Domingo Orué	-12,1084679	-77,0264329
27	Angamos	-12,1126430	-77,0260065
28	Ricardo Palma	-12,1189574	-77,0259296
29	Benavides	-12,1251887	-77,0241960
30	28 de Julio	-12,1327355	-77,0208010
31	Plaza las Flores	-12,1357091	-77,0186979
32	Balta	-12,1393886	-77,0178131
33	Bulevar	-12,1483917	-77,0201145
34	Estadio Unión	-12,1529051	-77,0197335
35	Escuela Militar	-12,1597668	-77,0189703
36	Terán	-12,1686435	-77,0186235
37	Rosario de Villa	-12,1730247	-77,0151099
38	Matellini	-12,1782643	-77,0104387

Fuente:Elaboración propia.

ANEXO 3: Red de la Vía Troncal

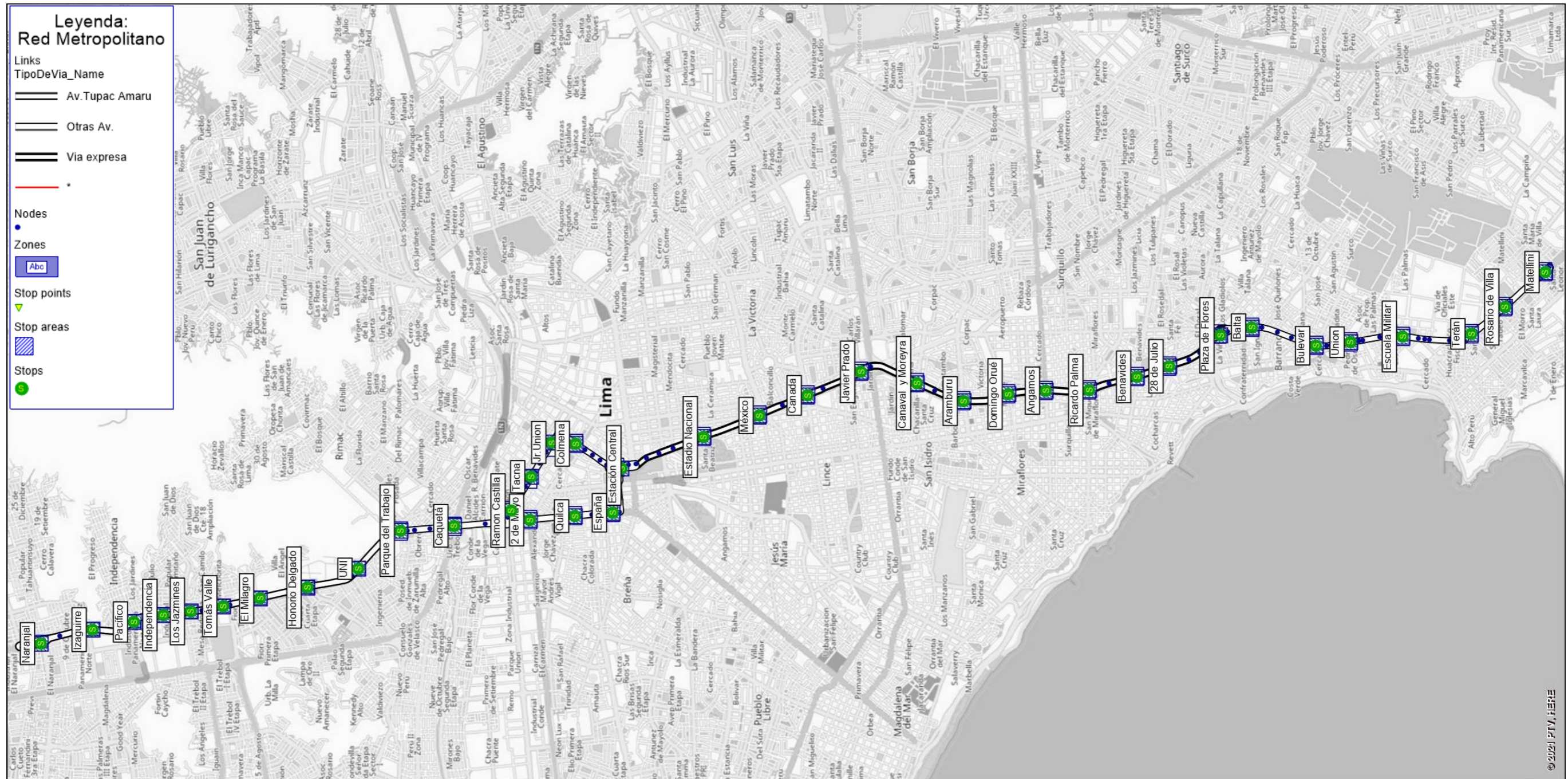


Figura 97 Red de transporte Metropolitano-Troncal
 Fuente: Elaboración propia

ANEXO 4: Capacidad máxima de un bus

Los datos fueron tomados el día 17/10/2019 entre las 7:00 y las 8:30am en el embarque 1 al Expreso 2 en dirección Norte a Sur en el terminal Naranjal; los conteos fueron tomados de manera manual; como se muestra en la Figura 98 y Figura 99 el bus realiza dos paradas, la primera para recoger a las personas que requieran asiento reservado, en la segunda parada recoge a los pasajeros restantes, en cada puerta se coloca a una persona para que realice el conteo de las personas que ingresan al bus. Ver Figura 100.



Figura 98 Primera parada del Expreso 2
Fuente: Fotografía propia tomada el 17/10/2019



Figura 99 Segunda parada del Expreso 2
Fuente: Fotografía propia tomada el 17/10/2019



Figura 100 Toma de datos en el Expreso 2
Fuente: Fotografía propia tomada el 17/10/2019

Entre las 7:00 y las 8:30 se contabilizó 64 buses, donde NSPC significa “no se puedo contar”, ver Tabla 41. Los resultados ver en la Tabla 42.

Tabla 41 Recolección de datos: capacidad máxima del bus

Capacidad máxima del bus tomada en el Expreso 2 de N-S						
Nº	Hora	ID BUS	Puerta delantera	Puerta del centro	Puerta trasera y Reservados	Resultado
1	7:00:00	11056	53	28	NSPC	NSPC
2		0103T016	NSPC	49	NSPC	NSPC
3		41061	NSPC	33	NSPC	NSPC
4		41030	58	33	50	141
5		21054	52	35	10	97
6		0103t052	61	37	NSPC	NSPC
7		0103T007	45	32	15	92
8		41056	56	38	NSPC	NSPC
9	7:08:00	21012	59	33	24	116
10		21010	53	33	15	101
11		0103t048	60	32	NSPC	NSPC
12		11012	59	36	8	103
13		21049	64	41	42	147
14	7:17:00	21051	66	48	28	142
15		11009	64	47	34	145
16		41078	51	48	29	128
17		0103t062	60	37	NSPC	NSPC
18		11046	53	44	18	115
19		21015	60	31	NSPC	NSPC
20		0103t069	51	40	15	106
21		41031	NSPC	38	60	NSPC
22		0103t014	66	37	32	135
23	7:30:00	0103t077	NSPC	42	32	NSPC
24		21001	NSPC	48	40	NSPC
25		0103t050	59	42	20	121
26		41014	60	42	NSPC	NSPC
27		21053	59	42	38	139
28		0103t030	55	44	NSPC	NSPC
29		42037	62	49	35	146
30		11017	52	40	NSPC	NSPC
31		41038	NSPC	38	19	NSPC
32		41067	61	47	51	159
33		21059	75	38	28	141
34		41063	79	42	21	142
35		0103t070	NSPC	48	NSPC	NSPC
36	7:52:00	21023	67	43	27	137
37		0103t063	67	48	24	139
38		21034	54	37	32	123
39		0104t020	69	41	38	148
40		41017	63	47	19	129
41		0103t059	66	42	NSPC	NSPC
42	8:00:00	21040	NSPC	40	33	NSPC
43		41019	NSPC	47	34	NSPC
44		21045	59	35	35	129
45		21003	72	46	42	160
46		21014	52	38	20	110
47		0103t066	52	44	NSPC	NSPC
48		21029	74	38	15	127
49		41046	57	38	NSPC	NSPC
50	8:11:00	0103t021	NSPC	42	18	NSPC
51		21028	63	40	NSPC	NSPC
52		0103t016	63	45	30	138
53		0103t007	66	40	21	127
54		41009	NSPC	42	49	NSPC
55		11056	63	46	37	146
56		41030	NSPC	48	39	NSPC
57		21010	68	47	20	135
58		41061	52	34	17	103
59		21054	69	42	37	148
60		11012	70	47	22	139
61		0103t052	56	31	NSPC	NSPC
62		41056	61	37	NSPC	NSPC
63		21012	58	37	24	119
64	8:30:00	21049	60	37	NSPC	NSPC

Leyenda NSPC : No se pudo contar

Fuente: Elaboración propia

Tabla 42 Capacidad máxima del Bus

Conteo	Capacidad
Mínimo	92
Promedio	129,8
Máximo	160

Fuente: Elaboración propia

ANEXO 5: Pseudocódigo

Se mostrará el pseudocódigo del algoritmo con la finalidad de hacer fácil la codificación de la programación.

Creación de funciones:Tiempo entre dos puntos del bus

Archivo con datos de coordenadas de estaciones

Coordenada w = (lat._w, long._w)

Coordenada z = (lat._z, long._z)

Velocidad del bus = V

FTw_z (lat._w, long._w, lat._z, long._z, velocidad del bus)

$$\text{Distancia } w_z = \sqrt{(\text{lat.}_z - \text{lat.}_w)^2 + (\text{long.}_z - \text{long.}_w)^2}$$

$$T_{w_z} = \frac{\text{Distancia } w_z}{\text{velocidad del bus}}$$

Retornar Tw_z

Paquete directo

validaciones con los parámetros “Fecha”, “Hora”, “Tarjeta”, “Origen”, “Destino”, “Sentido”, “Hemb.V”, “Hmax.S.V”, “HSP”, **PD son validaciones x_y con HSP y**

VPR validadores para la formación de paquetes regulares

FPD (capacidad máxima del bus, validaciones x_y)

Ordenar las validaciones según el parámetro “hora” de menor a mayor

M ← contar validadoresx_y

Para i ← 1 hasta M // cantidad de validaciones x_y

Para “validación m” ← 1 hasta M // m es una validación

Para n ← 1 hasta “capacidad máxima del bus - 1” // conteo para el bus

Si “Hmax.S.V”_m >= “Hemb.V”_m+n entonces “cantidad de validaciones” ← n+1

Si “cantidad de validaciones” = “Capacidad máxima del bus”

entonces:

PD ← validaciones desde m hasta Capacidad máxima del bus,

“HSP” ← “Hmax.S.V”_m // agregar este parámetro

NP ← i // Numero de paquetes

m ← m+1

Sino n ← n+1

Si “Hmax.S.V”_m < “Hemb.V”_m+n entonces “validación m” guardar en

Viajes regulares, m ← m+1 // vuelve a validación m

Retornar PD // un archivo con las validaciones

Archivos con paquetes directos

FAPD(capacidad máxima del bus, validaciones++++)

Para $x \leftarrow 1$ hasta N

Para $y \leftarrow x+1$ hasta N

validaciones $x_y \leftarrow$ seleccionar validaciones++con origen x y destino y

FPD (capacidad máxima del bus, validaciones x_y)=(1)PD x_y (2)paraPR x_y

Ps \leftarrow PD x_y // guardar paquetes en un archivo

Validaciones++++ \leftarrow paraPR x_y

Retornar Ps

Paquete regular

FPR (Capacidad máxima de embarque B, validaciones $x_desconocido$)

Ordenar las validaciones según el parámetro “hora” de menor a mayor

M \leftarrow contar validaciones $x_desconocido$

Para $i \leftarrow 1$ hasta M // cantidad de validaciones x_d

Para “validación m” $\leftarrow 1$ hasta M // m es una validación

Para $n \leftarrow 1$ hasta “capacidad del paquete - 1” // conteo para el bus

Si “Hmax.S.V” $_m \geq$ “Hemb.V” $_m+n$ entonces “cantidad de validaciones” $\leftarrow n+1$

Si “cantidad de validaciones” = “capacidad del paquete” entonces:

PR \leftarrow validaciones desde m hasta “capacidad del paquete”

“HSP” \leftarrow “Hora de máxima salida” $_m$, NP $\leftarrow i$; m $\leftarrow m+1$ // agregar

Sino n $\leftarrow n+1$

Si “Hmax.S.V” $_m <$ “Hemb.V” $_m+n$ entonces:

PR \leftarrow validaciones desde m hasta validación m+n-1,

m $\leftarrow m+1$ // vuelve a validación m

Retornar PR // un archivo con las validaciones

Archivos con paquetes regulares

FAPR(Capacidad máxima de embarque B, validaciones++++)

Para $x \leftarrow 1$ hasta N

validaciones $x_desconocido \leftarrow$ seleccionar validaciones +++ con origen x

FPR (Capacidad máxima de embarque B, validaciones $x_desconozc$) =PR x_d

Ps \leftarrow PR x_d //guardar paquetes en un archivo

Retornar Ps

Función generación de rutas directas y regulares

FGR (Paquete, capacidad máxima del B)

Val.B \leftarrow 0

Para z \leftarrow 1 hasta N-1

Para k \leftarrow 0 hasta N-1

N° Val.d_z \leftarrow contar Val.B con “origen (descocido \leq z)” y “destino z”

N° Val.B \leftarrow N° Val.B - N° Val.d_z

Capacidad máxima de embarque B \leftarrow capacidad máxima del B - N° Val.B

Si Val.d_z=0 y Capacidad máxima de embarque B \geq 1

Si k=0

PR \leftarrow seleccionar un Paquete con menor HSPz// Eliminar los P, en la Z

Validaciones++++ \leftarrow Val.AP-PR

Si PR \geq 1 // bus para

HLBz \leftarrow HSPz

HSBz \leftarrow HLBz+ Tdet.B

N° Val.B \leftarrow N° Val.B + PR

Ruta \leftarrow (z,HLBz,HSBz N° Val.B) // generar un archivo

HSBm \leftarrow HSBz; w \leftarrow z; k \leftarrow k+1; z \leftarrow z+1

Sino k \leftarrow 0; z \leftarrow z+1 // Sigue con N° Val.B \leftarrow 0 sigue buscando 1era parada

Sino // k $>$ 0

HLBz \leftarrow HSBw+FTw_z(lat._w, long._w, lat._z, long._z, velocidad del bus)

PR \leftarrow seleccionar el Paquete con HSPz=HLBz-t //condición

Validaciones++++ \leftarrow Val.AP-PR

Si PR \geq 1 // bus para

HSBz \leftarrow HLBz+ Tdet.B

N° Val.B \leftarrow N° Val.B + PR

Ruta \leftarrow (z) // generar un archivo con z,HLBz,HSBz N° Val.B

HSBm \leftarrow HSBz; w \leftarrow z; k \leftarrow k+1; z \leftarrow z+1

Si PR=0

HSBz \leftarrow HLBz; HSBw \leftarrow HSBz; w \leftarrow z; k \leftarrow k; z \leftarrow z+1 // bus no para

Si Val.d_z \geq 1 // bus para

HLBz \leftarrow HSBw+FTw_z(lat._w, long._w, lat._z, long._z, velocidad del bus)

PR \leftarrow seleccionar el Paquete con HSPz=HLBz-t

Validaciones++++ \leftarrow Val.AP-PR

HSBz \leftarrow HLBz+ Tdet.B

N° Val.B \leftarrow N° Val.B + PR

Ruta ← (z) // generar un archivo con z,HLBz,HSBz N° Val.B

HSBm←HSBz; w←z; k←k+1; z←z+1

Si Capacidad máxima de embarque B= capacidad máxima del B

FIN ← Z=N //ver si esto o imprimir los N°Val.B

Si Capacidad máxima de embarque B=0

HLBz←HSBw+FTw_z(lat._w, long._w, lat._z, long._z, velocidad del bus)

HSBz←HLBz; HSBw←HSBz; w←z; k←k; z←z+1 // bus no para

Retornar (1) Ruta: Z, HLB, HSB y N°ValB en cada parada (2) Paquete:
validadores sobrantes= validadores++++

Generación de rutas:

Llamar los siguientes datos

Archivo con validaciones// “Fecha”, “Hora”, “Tarjeta”, “Origen”, “Destino”

Capacidad máxima del bus← Cmaxb//160

Tiempo del torniquete hasta el embarque de la validación←Ttor_emb.V // 30seg

Tiempo de máxima espera de la validación←Tmax.esp.V //15min

Archivo con datos de coordenadas de estaciones

Velocidad del bus←V

Validaciones+

A las validaciones agregar otros parámetros:

Sentido ←Según el Origen-Destino de la ruta recorrido//N-S y S-N

Hemb.V_m←Hora_m+ Ttor_emb.V

Hmax.S.V_m←Hora_m+ Tmax.esp.V

Validaciones+ con los parámetros “Fecha”, “Hora”, “Tarjeta”, “Origen”, “Destino”,
“Sentido”, “Hemb.V”, “Hmax.S.V”

Sentidos

Validaciones++←Seleccionar las validaciones+ con sentido “ida” o “vuelta”

Generación de rutas directas

validaciones +++++← validaciones ++

Paquete← FAPD(capacidad máxima del bus, validaciones++++)

MP←contar validadores de Paquete

Mientras (MP>0)

FGR (Paquete, capacidad máxima del B) = (1) Ruta: HLB, HSB, N°ValB y Z en
cada parada (2) Paquete: validadores sobrantes

Paquete← validadores sobrantes

GR← Guardar las rutas en un archivo

Generación de rutas regulares

validaciones +++++← validaciones ++++

Paquete←FAPR (Capacidad máxima de embarque B, validaciones +++++)

M←contar validadores de Paquete

Mientras (MP>0)

FGR (Paquete, capacidad máxima del B) = (1) Ruta: HLB, HSB, N°ValB y Z en cada parada (2) Paquete: validadores sobrantes

Paquete← validadores sobrantes

GR← Guardar las rutas en un archivo

Programación de rutas

Llamar

Archivo de rutas GR

Archivo con programación de semáforos

Rutas con semáforos

Ordenar de menor a mayor las rutas según la hora de inicio de recorrido del bus

Elegir un sentido “ida” o “vuelta”

Para m←1 hasta la M// Cantidad de rutas en el archivo ruta m: a-b-c-...-q

Para estación←a hasta q

Sem← semáforos entre a y b

Nº sem←Contar Sem

Si Nº sem>=1// SR3, SR4, SR5

HSBp← HSBa

Para i←1 hasta Nº sem// me indica el semáforo

hora del bus en el semáforo rojo←HSBp+FTp_sem

Si hora del bus en el semáforo rojo= hora de semáforo rojo

TsemR← tiempo del semáforo rojo

HSBp← hora del bus en el semáforo rojo+ TsemR

i←i+1

Sino HSBp← hora del bus en el semáforo rojo; i←i+1

HLBb← HSBa suma de los TsemR; HSBb← HLBb+ Tdet.B

Sino a←b y b←c

Ruta: HLB, HSB, N°ValB y Z en cada parada// el HLB y el HSB están actualizados

ANEXO 6: Programación resultado del modelo “Generación de rutas óptimas”
En las siguientes tablas se muestra el resultado de los horarios de salida de todos los buses que fueron generados por el modelo “Generación de rutas óptimas”, de las cuales rutas “Directas” son 89, las rutas “Regulares 1” son 89 y rutas “Regulares 2” son 16

Tabla 43 Programación de las rutas directas con total de 89

Nombre	Estacion	HLB	HSB	Estacion	HLB
L_1_14_T6_N_S	Naranjal	07:07:36	07:08:06	España	07:23:45
	Naranjal	07:14:57	07:15:27	España	07:31:06
	Naranjal	07:20:43	07:21:13	España	07:36:52
	Naranjal	07:28:01	07:28:31	España	07:44:10
	Naranjal	07:35:45	07:36:15	España	07:51:54
	Naranjal	07:47:03	07:47:33	España	08:03:12
L_1_15_T4_N_S	Naranjal	07:13:04	07:13:34	Jiron de la Unión	07:28:26
	Naranjal	07:24:11	07:24:41	Jiron de la Unión	07:39:33
	Naranjal	07:37:44	07:38:14	Jiron de la Unión	07:53:06
	Naranjal	07:51:56	07:52:26	Jiron de la Unión	08:07:18
L_1_19_T11_N_S	Naranjal	07:04:54	07:05:24	Estacion Central	07:22:12
	Naranjal	07:09:24	07:09:54	Estacion Central	07:26:42
	Naranjal	07:13:09	07:13:39	Estacion Central	07:30:27
	Naranjal	07:17:05	07:17:35	Estacion Central	07:34:23
	Naranjal	07:21:16	07:21:46	Estacion Central	07:38:34
	Naranjal	07:26:17	07:26:47	Estacion Central	07:43:35
	Naranjal	07:31:52	07:32:22	Estacion Central	07:49:10
	Naranjal	07:36:50	07:37:20	Estacion Central	07:54:08
	Naranjal	07:44:12	07:44:42	Estacion Central	08:01:30
	Naranjal	07:52:16	07:52:46	Estacion Central	08:09:34
Naranjal	08:00:00	08:00:30	Estacion Central	08:17:18	
L_1_22_T8_N_S	Naranjal	07:05:46	07:06:16	Canadá	07:28:23
	Naranjal	07:12:34	07:13:04	Canadá	07:35:11
	Naranjal	07:18:11	07:18:41	Canadá	07:40:48
	Naranjal	07:24:12	07:24:42	Canadá	07:46:49
	Naranjal	07:30:18	07:30:48	Canadá	07:52:55
	Naranjal	07:37:14	07:37:44	Canadá	07:59:51
	Naranjal	07:44:10	07:44:40	Canadá	08:06:47
	Naranjal	07:54:27	07:54:57	Canadá	08:17:04
L_1_23_T9_N_S	Naranjal	07:06:45	07:07:15	Javier Prado	07:30:52
	Naranjal	07:12:18	07:12:48	Javier Prado	07:36:25
	Naranjal	07:17:39	07:18:09	Javier Prado	07:41:46
	Naranjal	07:23:06	07:23:36	Javier Prado	07:47:13
	Naranjal	07:29:34	07:30:04	Javier Prado	07:53:41
	Naranjal	07:34:56	07:35:26	Javier Prado	07:59:03
	Naranjal	07:42:22	07:42:52	Javier Prado	08:06:29
	Naranjal	07:48:34	07:49:04	Javier Prado	08:12:41
Naranjal	07:56:23	07:56:53	Javier Prado	08:20:30	

Nombre	Estacion	HLB	HSB	Estacion	HLB
L_1_24_T15_N_S	Naranjal	07:04:25	07:04:55	Canaval y Moreyra	07:30:01
	Naranjal	07:09:12	07:09:42	Canaval y Moreyra	07:34:48
	Naranjal	07:12:52	07:13:22	Canaval y Moreyra	07:38:28
	Naranjal	07:16:26	07:16:56	Canaval y Moreyra	07:42:02
	Naranjal	07:20:26	07:20:56	Canaval y Moreyra	07:46:02
	Naranjal	07:24:36	07:25:06	Canaval y Moreyra	07:50:12
	Naranjal	07:27:58	07:28:28	Canaval y Moreyra	07:53:34
	Naranjal	07:30:55	07:31:25	Canaval y Moreyra	07:56:31
	Naranjal	07:34:27	07:34:57	Canaval y Moreyra	08:00:03
	Naranjal	07:37:44	07:38:14	Canaval y Moreyra	08:03:20
	Naranjal	07:41:41	07:42:11	Canaval y Moreyra	08:07:17
	Naranjal	07:45:28	07:45:58	Canaval y Moreyra	08:11:04
	Naranjal	07:49:29	07:49:59	Canaval y Moreyra	08:15:05
	Naranjal	07:54:07	07:54:37	Canaval y Moreyra	08:19:43
	Naranjal	07:58:44	07:59:14	Canaval y Moreyra	08:24:20
L_1_25_T6_N_S	Naranjal	07:08:59	07:09:29	Aramburu	07:35:53
	Naranjal	07:16:53	07:17:23	Aramburu	07:43:47
	Naranjal	07:27:16	07:27:46	Aramburu	07:54:10
	Naranjal	07:33:35	07:34:05	Aramburu	08:00:29
	Naranjal	07:40:45	07:41:15	Aramburu	08:07:39
	Naranjal	07:50:26	07:50:56	Aramburu	08:17:20
L_1_27_T12_N_S	Naranjal	07:04:29	07:04:59	Angamos	07:33:24
	Naranjal	07:08:34	07:09:04	Angamos	07:37:29
	Naranjal	07:13:44	07:14:14	Angamos	07:42:39
	Naranjal	07:17:51	07:18:21	Angamos	07:46:46
	Naranjal	07:22:08	07:22:38	Angamos	07:51:03
	Naranjal	07:27:01	07:27:31	Angamos	07:55:56
	Naranjal	07:31:21	07:31:51	Angamos	08:00:16
	Naranjal	07:35:22	07:35:52	Angamos	08:04:17
	Naranjal	07:39:23	07:39:53	Angamos	08:08:18
	Naranjal	07:44:38	07:45:08	Angamos	08:13:33
	Naranjal	07:50:15	07:50:45	Angamos	08:19:10
	Naranjal	07:55:25	07:55:55	Angamos	08:24:20
L_1_28_T5_N_S	Naranjal	07:11:20	07:11:50	Ricardo Palma	07:41:30
	Naranjal	07:20:20	07:20:50	Ricardo Palma	07:50:30
	Naranjal	07:31:27	07:31:57	Ricardo Palma	08:01:37
	Naranjal	07:42:32	07:43:02	Ricardo Palma	08:12:42
	Naranjal	07:54:21	07:54:51	Ricardo Palma	08:24:31
L_1_29_T8_N_S	Naranjal	07:07:22	07:07:52	Benavides	07:38:50
	Naranjal	07:14:30	07:15:00	Benavides	07:45:58
	Naranjal	07:21:11	07:21:41	Benavides	07:52:39
	Naranjal	07:29:09	07:29:39	Benavides	08:00:37
	Naranjal	07:34:31	07:35:01	Benavides	08:05:59
	Naranjal	07:40:51	07:41:21	Benavides	08:12:19
	Naranjal	07:48:22	07:48:52	Benavides	08:19:50
	Naranjal	07:55:45	07:56:15	Benavides	08:27:13
L_1_38_T1_N_S	Naranjal	07:15:49	07:16:19	Matellini	07:58:53
L_6_24_T1_S_N	Tomas Valle	07:58:10	07:58:40	Canaval y Moreyra	08:18:47
L_38_24_T1_S_N	Matellini	07:57:01	07:57:31	Canaval y Moreyra	08:14:59
L_38_19_T2_S_N	Matellini	07:14:33	07:15:03	Estacion Central	07:40:48
	Matellini	07:25:27	07:25:57	Estacion Central	07:51:42

Fuente: Elaboración propia

Tabla 45 Programación de las rutas "Regulares 2".

Nombre	N°114_N_S	N°115_N_S	N°116_N_S	N°117_N_S	N°118_N_S	N°119_N_S	N°120_N_S	N°121_N_S	N°122_S_N	N°123_S_N	N°124_S_N	N°125_S_N	N°126_S_N	N°127_S_N	N°128_S_N	N°129_S_N
Recorrido	00:35:23	00:34:14	00:34:15	00:31:40	00:31:10	00:29:38	00:29:38	00:29:38	00:37:43	00:35:29	00:35:58	00:33:21	00:28:25	00:27:35	00:12:55	00:12:55
Paradas	23	24	5	23	3	11	4	14	26	28	14	3	6	3	2	2
Transp.	62	162	7	160	1	10	3	32	199	172	52	5	18	5	1	1
Estación	Honorio Delgado	Uni	Uni	Parque del Trabajo	Caqueta	Ramon Castilla	Ramon Castilla	Ramon Castilla	Terán	Union	Bulevar	Plaza de Flores	Angamos	Domingo Orué	Ramon Castilla	Ramon Castilla
HLB	07:19:40	07:57:31	07:35:31	07:59:29	07:40:45	07:29:12	07:52:19	07:16:59	08:00:29	07:54:48	07:56:49	08:00:13	07:17:09	07:46:31	07:09:32	07:58:45
HSB	07:20:10	07:58:01	07:36:01	07:59:59	07:41:15	07:29:42	07:52:49	07:17:29	08:00:59	07:55:18	07:57:19	08:00:43	07:17:39	07:47:01	07:10:02	07:59:15
Espera	00:00:00	00:53:48	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	00:00:46	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00
Suben	7	10	6	1	1	1	3	32	1	2	2	5	13	5	1	1
Estación	Caqueta	Parque del Trabajo	Caqueta	Tacna	España	Tacna	Jiron de la Unión	Canadá	Bulevar	Bulevar	Plaza de Flores	Jiron de la Unión	Canadá	España	Naranjal	Naranjal
	Quilca	Caqueta	Quilca	Jiron de la Unión	Matellini	Canadá	Colmena	Javier Prado	Balta	Plaza de Flores	Canadá	Naranjal	España	Naranjal		
	España	Ramon Castilla	España	Colmena		Canaval y Moreyra	Matellini	Canaval y Moreyra	Plaza de Flores	28 de Julio	México		Quilca			
	Estacion Central	Tacna	Matellini	Estacion Central		Aramburu		Angamos	Benavides	Benavides	Estacion Central		2 de Mayo			
	México	Jiron de la Unión	X	México		Angamos		Ricardo Palma	Ricardo Palma	Colmena			Naranjal			
	Canadá	Colmena		Canadá		28 de Julio		Benavides	Domingo Orué	Angamos	Jiron de la Unión					
	Javier Prado	Estacion Central		Javier Prado		Plaza de Flores		28 de Julio	Aramburu	Domingo Orué	Ramon Castilla					
	Canaval y Moreyra	Estadio Nacional		Canaval y Moreyra		Bulevar		Plaza de Flores	Canaval y Moreyra	Aramburu	Uni					
	Aramburu	Canadá		Aramburu		Escuela Militar		Balta	Canadá	Canaval y Moreyra	Honorio Delgado					
	Domingo Orué	Javier Prado		Domingo Orué		Matellini		Bulevar	México	Javier Prado	Tomas Valle					
	Angamos	Canaval y Moreyra		Angamos				Terán	Estadio Nacional	México	Independencia					
	Ricardo Palma	Aramburu		Ricardo Palma				Rosario de Villa	Estacion Central	Estadio Nacional	Izaguirre					
	Benavides	Domingo Orué		Benavides				Matellini	España	Estacion Central	Naranjal					
	28 de Julio	Angamos		28 de Julio					Quilca	Colmena						
	Plaza de Flores	Ricardo Palma		Plaza de Flores					2 de Mayo	Jiron de la Unión						
	Balta	Benavides		Balta					Caqueta	Tacna						
	Bulevar	28 de Julio		Bulevar					Parque del Trabajo	Ramon Castilla						
	Union	Plaza de Flores		Union					Uni	Caqueta						
	Escuela Militar	Balta		Escuela Militar					Honorio Delgado	Parque del Trabajo						
	Terán	Bulevar		Terán					El Milagro	Uni						
	Rosario de Villa	Union		Rosario de Villa					Tomas Valle	Honorio Delgado						
	Matellini	Rosario de Villa		Matellini					Los jazmines	El Milagro						
		Matellini							Independencia	Tomas Valle						
									Izaguirre	Independencia						
									Naranjal	Pacífico						
										Izaguirre						
										Naranjal						
	07:55:03	08:31:45	08:09:46	08:31:09	08:11:55	07:58:50	08:21:57	07:46:37	08:38:12	08:30:17	08:32:47	08:33:34	07:45:34	08:14:06	07:22:27	08:11:40

Fuente: Elaboración propia.

ANEXO 7: Matriz de resultados

Matrix editor (Matrix '2 JRT (BRT_Cod_Dseg BRT_Name_Dseg)')		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38		
	Name	Naranjajazguirre	Pacifico	Independencia	Jazmines	El Milagro	Del UNI	de Tacna	Quilca	Español	Cañaval y Aramburu	Jr. Union	Colmena	Cedro	Benavides	de Julio	Ba	Bulevar	Union	Escuela Militar	Terán	Rosario de Villa	Matellini																		
	Sum	1061.9	992.1	1006.6	975.3	930.5	857.5	890.8	846.1	753.4	840.3	743.0	782.0	827.4	762.6	718.3	712.9	722.0	670.0	602.2	672.3	651.2	646.7	624.5	614.2	699.1	734.5	691.9	753.8	736.2	792.5	902.6	906.5	941.7	985.2	1047.1	1130.9	1192.3	1268.2		
1	Naranjal	1009.6	0.0	2.1	3.9	5.3	6.5	8.0	9.5	11.4	13.5	15.7	17.8	15.5	17.3	18.9	20.3	22.4	24.6	25.3	21.8	32.4	36.7	25.2	27.4	28.8	30.7	38.7	32.8	34.9	35.8	37.3	47.8	51.0	43.5	45.1	47.1	49.7	51.5	53.5	
2	Izaguirre	1078.5	2.1	0.0	1.5	2.8	4.1	5.5	7.1	9.0	11.1	13.3	15.4	18.0	19.9	21.5	17.9	20.0	22.1	22.6	24.9	35.2	30.1	32.1	33.2	33.4	38.0	40.3	39.8	42.3	40.4	44.2	48.8	49.0	50.6	52.2	54.2	56.8	58.6	60.6	
3	Pacifico	1018.9	3.5	1.4	0.0	1.1	2.4	3.8	5.4	7.3	9.4	11.6	13.6	16.3	18.1	19.7	16.2	18.2	20.3	20.9	23.1	33.3	28.4	30.3	31.4	31.6	36.3	38.6	38.0	40.6	38.7	42.5	47.1	47.2	48.9	50.4	52.5	55.0	56.8	58.9	
4	Independencia	972.9	4.5	2.5	1.1	0.0	1.0	2.4	4.0	5.9	8.0	10.2	12.2	14.9	16.7	18.3	14.8	16.8	18.9	19.5	21.7	31.7	27.0	28.9	30.0	30.2	34.9	37.1	36.6	39.2	37.3	41.1	45.7	45.8	47.5	49.0	51.1	53.6	55.4	57.5	
5	Los Jazmines	932.3	5.3	3.3	2.0	1.0	0.0	1.1	2.7	4.6	6.7	8.9	10.9	13.6	15.4	17.0	13.5	15.5	17.6	18.2	20.4	30.2	25.7	27.6	28.6	28.9	33.6	35.8	35.2	37.9	35.9	39.8	44.4	44.5	46.2	47.7	49.8	52.3	54.1	56.2	
6	Tomás Valle	889.8	6.4	4.3	3.0	2.0	1.1	0.0	1.2	3.1	5.2	7.4	9.5	12.2	14.0	15.6	12.0	14.1	16.2	16.7	19.0	28.6	24.2	26.2	27.2	27.4	32.2	34.4	33.8	36.5	34.5	38.4	42.9	43.0	44.8	46.3	48.3	50.9	52.7	54.8	
7	El Milagro	874.1	7.5	5.5	4.2	3.2	2.4	1.3	0.0	1.6	3.7	5.9	7.9	10.6	12.4	14.0	10.5	12.5	14.6	15.2	17.4	26.7	22.6	24.6	25.3	29.1	34.4	36.0	31.8	34.9	37.1	41.7	45.6	45.6	43.2	44.7	46.8	49.3	51.1	53.2	
8	Honorio Delgado	824.7	8.9	6.9	5.6	4.6	3.8	2.8	1.6	0.0	1.7	3.9	6.0	8.7	10.5	12.1	8.5	10.6	12.7	13.3	15.5	24.5	20.7	22.7	25.0	26.3	32.4	34.1	31.8	33.2	33.5	38.9	43.5	43.6	41.3	42.8	44.8	47.4	49.2	51.3	
9	UNI	730.4	10.8	8.8	7.4	6.4	5.6	4.5	3.4	1.8	0.0	1.8	3.7	6.6	8.4	10.0	6.4	8.5	10.6	11.2	13.3	21.9	18.6	20.6	24.2	25.7	27.8	26.8	32.0	37.5	39.0	36.3	37.9	39.9	42.5	44.3	46.4	48.4	50.4	52.4	54.7
10	Parque del Trabajo	799.7	12.1	10.2	8.9	8.0	7.2	6.2	5.1	3.6	1.9	0.0	1.8	4.4	6.2	7.8	4.3	6.3	8.4	9.0	11.1	19.9	19.8	22.0	22.5	23.0	29.2	32.3	29.7	34.4	30.3	35.2	41.7	42.2	43.9	45.5	47.7	50.5	52.4	54.7	
11	Caquetá	724.1	14.0	12.1	10.6	9.6	8.9	7.9	6.8	5.3	3.6	1.7	0.0	2.4	4.2	5.8	2.2	4.3	6.4	7.0	9.1	18.0	19.5	21.6	22.1	15.9	23.8	28.3	26.3	30.2	23.0	28.5	36.3	37.9	40.0	41.6	43.8	46.5	48.4	50.6	
12	2 de Mayo	759.0	15.6	13.6	12.4	11.5	10.7	9.7	8.7	7.2	5.7	4.0	2.4	0.0	1.5	3.1	1.9	18.7	24.6	20.8	6.0	13.5	13.5	15.8	16.3	19.0	23.1	25.1	23.7	28.9	28.5	30.9	35.2	34.5	37.2	38.9	41.2	44.0	45.9	48.2	
13	Quilca	767.6	16.8	14.9	13.7	12.8	12.0	11.0	9.9	8.5	6.9	5.2	3.7	1.5	0.0	1.3	24.4	23.2	23.1	19.2	4.2	12.2	13.7	16.0	16.7	19.3	23.0	24.8	23.8	28.4	28.3	30.7	34.6	34.0	36.6	38.2	40.3	43.1	44.9	47.1	
14	España	732.3	18.5	16.4	14.8	13.8	13.3	12.4	11.0	9.6	8.2	6.3	4.8	2.6	1.3	0.0	21.8	20.7	19.7	17.5	2.6	10.8	13.2	15.1	15.8	18.0	21.3	23.1	22.1	26.4	26.1	28.4	32.0	31.5	34.0	35.5	37.6	40.2	42.0	44.0	
15	Ramon Castilla	715.0	19.9	17.5	15.8	14.5	13.2	11.8	10.2	8.3	6.2	4.0	2.0	1.9	0.0	1.7	3.8	4.4	7.4	9.5	11.8	13.7	15.8	18.1	20.1	21.9	23.1	25.2	26.7	28.3	31.1	32.0	34.5	36.0	38.0	40.6	42.4	44.4	46.4	48.4	
16	Tacna	717.4	21.9	19.6	17.9	16.6	15.3	13.9	12.3	10.4	8.4	6.2	4.1	1.9	0.0	2.9	3.5	6.5	8.7	10.9	12.9	14.9	17.2	19.3	21.1	22.2	24.3	25.9	27.5	30.2	31.2	33.6	35.1	37.2	39.7	41.5	43.6	45.6	47.6	49.6	
17	Jr. Union	732.0	24.2	21.9	20.2	18.8	17.6	16.2	14.6	12.7	10.6	8.4	6.4	2.4	0.0	2.9	0.0	2.6	5.6	7.8	10.0	12.0	14.1	16.3	18.4	20.2	21.4	23.5	25.0	26.6	29.3	30.3	32.8	34.3	36.3	38.9	40.6	42.7	44.7	46.7	
18	Colmena	667.1	24.5	22.2	20.5	19.2	17.9	16.5	14.9	13.0	11.0	8.8	6.7	1.9	0.0	2.7	4.8	7.0	9.0	11.1	13.4	15.4	17.2	18.4	20.5	22.1	23.7	26.4	27.3	29.8	31.3	33.3	35.9	37.7	39.7	41.7	43.7	45.7	47.7	49.7	
19	Estación Central	581.7	21.4	20.0	18.3	17.4	16.7	16.0	14.7	13.2	11.8	9.9	8.4	5.6	4.3	3.2	7.3	6.2	5.1	2.6	0.0	3.6	5.9	7.8	8.6	10.9	14.1	16.0	15.1	18.9	18.5	20.9	24.7	24.4	26.8	28.3	30.4	33.0	34.8	36.8	
20	Estadio Nacional	657.6	31.9	30.5	28.0	26.9	26.5	26.4	23.4	22.5	22.0	19.3	18.3	12.3	12.4	12.2	9.4	8.3	7.3	4.7	3.6	0.0	1.9	3.9	5.6	7.9	10.2	12.1	12.6	15.5	17.0	17.1	21.2	20.3	22.8	24.4	26.5	29.1	30.9	33.0	
21	México	626.4	28.0	26.5	25.7	24.8	23.7	22.3	21.7	20.2	17.9	19.9	12.5	14.2	14.3	11.7	10.5	9.5	7.0	5.9	1.9	0.0	1.7	3.9	6.1	8.0	9.9	11.3	13.2	15.0	16.4	19.1	20.1	22.6	24.1	26.1	28.7	30.5	32.6		
22	Canada	639.4	30.6	29.0	28.3	27.4	26.2	24.7	24.3	22.6	20.2	24.1	21.6	14.8	18.2	16.5	13.7	12.5	11.5	8.9	7.8	3.9	1.7	0.0	1.9	4.1	6.0	7.9	8.1	11.1	13.0	12.3	17.0	18.0	20.4	21.9	23.9	26.5	28.3	30.3	
23	Javier Prado	641.0	36.1	34.3	32.5	31.3	30.8	29.8	22.4	21.6	24.9	23.1	22.7	14.0	17.7	17.0	15.7	14.6	13.6	11.0	9.1	5.1	3.8	1.9	0.0	2.0	3.9	5.7	6.1	9.1	9.7	10.2	14.8	14.6	17.1	18.6	20.7	23.3	25.1	27.2	
24	Canaval y Moreyra	633.8	32.0	29.8	34.5	33.6	31.1	25.4	31.8	28.9	20.9	25.3	17.1	22.1	19.2	12.0	18.0	16.9	16.0	13.6	10.8	7.4	6.1	4.1	1.9	0.0	1.7	3.5	4.1	6.1	7.3	8.9	12.6	12.3	14.8	16.4	18.5	21.1	22.9	25.0	
25	Aramburu	752.3	40.5	38.6	41.5	40.6	38.2	34.4	38.1	35.6	29.9	32.6	27.0	28.2	26.7	22.0	20.2	18.9	17.9	15.5	14.6	10.4	8.1	6.1	3.9	1.6	0.0	1.6	3.1	4.0	6.8	8.2	10.7	11.9	14.3	15.8	17.8	20.4	22.2	24.3	
26	Domingo Orúe	795.8	44.5	42.5	44.4	43.4	41.5	38.1	40.3	38.4	33.4	36.3	30.6	29.9	29.9	25.3	21.8	20.7	19.7	17.1	16.2	12.2	9.9	7.9	5.7	3.4	1.6	0.0	1.3	3.1	5.0	6.5	8.9	10.3	12.7	14.3	16.4	19.0	20.8	22.9	
27	Angamos	691.9	36.4	34.2	39.4	38.5	35.6	29.8	37.3	33.8	25.3	30.1	21.5	27.3	23.6	16.4	23.2	22.1	21.3	18.9	15.2	11.7	11.5	9.5	6.5	4.1	3.2	1.4	0.0	1.4	2.9	4.1	7.3	7.6	10.2	11.7	13.8	16.4	18.2	20.3	
28	Ricardo Palma	805.2	43.1	41.1	45.8	44.9	42.0	36.9	43.2	40.0	32.5	36.9	29.2	33.5	30.3	24.1	25.4	24.2	23.2	20.7	19.5	15.5	13.2	11.3	9.1	6.9	4.9	3.0	1.5	0.0	1.6	3.1	5.6	6.8	9.3	10.9	13.0	15.7	17.5	19.7	
29	Benavides	865.0	47.2	45.2	49.3	48.3	46.1	40.7	46.1	43.4	36.0	41.8	33.7	35.7	34.9	28.3	27.1	26.0	25.0	22.4	21.3	17.4	15.2	13.2	11.0	8.7	6.9	5.0	3.5	1.7	0.0	1.2	3.6	4.9	7.4	9.0	11.0	13.7	15.5	17.6	
30	28 de Julio	825.0	46.6	45.0	45.2	44.1	42.9	40.4	40.1	38.9	35.6	38.0	34.0	30.8	32.3	28.5	28.1	27.0	26.0	23.5	19.9	15.9	16.7	14.7	10.8	8.5	8.4	6.5	4.2	3.3	1.4	0.0	2.2	3.1	5.6	7.1	9.1	11.7	13.5	15.6	
31	Plaza de Flores	835.9	42.7	40.3	46.8	46.0	42.2	36.0	45.0	41.0	31.4	37.1	27.7	35.2	30.1	22.6	30.6	29.5	28.5	26.2	21.5	21.1	19.1	17.2	14.9	10.8	10.8	9.1	6.3	5.7	3.8	2.4	0.0	1.3	3.8	5.4	7.5	10.2	12.0	14.2	
32	Ba																																								

Matrix editor (Matrix '2 JRT (BRT_Cod_Dseg BRT_Name_Dseg))

38 x 38	Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	
	Naranjal	1051.3	0.0	12.2	13.5	20.1	26.1	38.2	7.0	8.8	11.3	12.9	14.7	17.4	19.5	18.3	17.5	19.9	19.4	22.5	20.6	25.5	27.2	26.0	27.5	28.9	31.1	35.3	32.9	35.4	35.9	39.7	43.1	43.9	45.3	47.1	48.9	51.3	53.1	53.4
	Izaguirre	1221.3	2.1	0.0	1.5	8.1	14.1	22.0	15.5	15.7	18.5	11.3	14.2	16.6	19.2	21.0	25.2	26.8	29.2	29.8	25.8	28.1	30.4	32.7	35.2	37.7	39.7	41.7	43.4	45.5	47.5	49.4	51.5	52.3	55.3	59.8	62.5	66.1	62.2	64.0
	Pacifico	1032.2	3.8	1.4	0.0	6.9	12.8	13.8	4.0	5.6	7.6	9.1	11.5	14.3	16.2	17.9	17.4	19.5	21.3	22.4	21.5	23.7	25.9	28.2	30.5	32.9	34.9	36.8	38.7	40.6	42.6	44.8	46.3	46.8	49.9	52.6	55.6	58.4	57.6	58.4
	Independencia	1063.6	5.2	2.8	1.1	0.0	6.2	7.1	8.3	10.4	11.9	13.6	15.1	17.4	19.8	21.1	14.7	17.2	19.7	20.3	23.5	25.7	28.7	29.8	32.0	34.4	36.3	38.2	39.7	41.5	43.4	44.9	47.3	49.6	51.6	53.5	55.7	58.1	58.4	59.3
	Los Jazmines	890.4	7.3	4.5	2.5	1.0	0.0	1.1	2.2	4.0	5.7	7.6	9.1	11.5	13.8	15.1	11.5	13.6	15.7	16.3	18.6	20.7	23.0	24.9	27.1	29.4	31.3	33.4	34.6	36.4	38.3	39.7	42.2	44.0	46.3	48.4	50.2	53.0	53.3	53.5
	Tomás Valle	912.5	7.6	5.2	3.4	2.0	1.1	0.0	1.2	3.4	12.0	13.5	8.8	13.4	13.3	14.9	10.7	12.6	14.8	15.4	18.5	20.6	22.9	24.9	27.2	28.5	31.5	33.6	34.9	36.8	38.7	40.5	42.7	44.3	46.6	48.7	51.6	54.6	57.1	55.2
	El Milagro	825.4	9.3	6.8	5.1	3.7	2.8	1.3	0.0	1.6	3.7	5.5	7.3	9.8	11.7	13.2	9.8	11.3	13.6	14.1	16.1	18.3	20.6	22.3	24.5	26.7	28.6	30.5	32.0	33.8	35.7	37.2	39.6	40.9	43.5	45.1	47.3	49.3	51.5	51.5
	Honorio Delgado	785.2	11.1	8.7	7.0	5.6	4.7	3.2	1.6	0.0	1.7	3.7	5.4	8.1	9.9	11.4	8.0	9.6	11.7	12.3	14.4	16.6	18.9	20.6	22.8	25.1	27.0	28.8	30.4	32.2	34.1	35.6	38.0	39.2	41.7	43.2	45.6	47.9	49.4	49.9
	UNI	754.1	13.2	10.8	9.0	7.6	7.0	5.2	3.7	1.8	0.0	1.8	3.8	6.4	8.1	9.7	6.2	8.3	10.4	10.9	13.2	15.1	17.3	19.3	21.4	23.7	25.6	27.4	29.0	30.8	32.7	34.2	36.5	37.7	40.2	41.6	43.7	46.1	47.9	46.7
	Parque del Trabajo	723.5	15.5	12.8	11.2	9.8	9.4	7.4	5.9	3.9	1.9	0.0	1.8	4.4	6.2	8.0	4.2	6.2	8.3	8.9	11.1	13.3	15.5	17.4	19.7	22.0	23.9	25.7	27.3	29.1	31.0	32.5	35.0	35.9	38.6	39.8	42.2	44.7	46.3	47.1
	Caquetá	678.4	17.2	14.8	13.2	11.8	12.0	9.4	7.9	6.0	3.9	1.7	0.0	2.4	4.1	5.6	2.2	4.3	6.3	6.9	8.8	10.9	13.1	15.0	17.2	19.5	21.5	23.3	25.0	26.8	28.7	30.2	32.6	33.8	36.2	37.4	39.5	42.3	43.9	42.6
	2 de Mayo	700.8	17.8	17.6	16.1	14.6	14.9	12.1	10.7	8.7	6.6	4.4	2.4	0.0	1.5	3.0	19.5	18.3	23.7	20.3	5.9	8.0	10.3	12.2	14.3	16.6	18.5	20.4	21.9	23.6	25.5	27.0	29.4	30.6	32.9	34.4	36.4	38.9	40.2	41.4
	Quilca	688.4	19.7	19.3	17.8	16.3	16.9	13.9	12.4	10.4	8.4	6.2	4.1	1.5	0.0	1.3	24.0	22.7	22.6	18.5	4.2	6.3	8.5	10.4	12.6	14.9	16.8	18.6	20.1	21.9	23.8	25.3	27.7	28.9	31.2	32.7	34.7	37.2	38.6	38.1
	España	665.5	21.4	20.8	19.5	17.9	18.0	15.4	14.0	12.0	10.0	7.8	5.7	3.1	1.3	0.0	21.2	20.1	19.4	17.4	2.6	4.7	7.0	8.9	11.0	13.3	15.2	17.1	18.6	20.3	22.2	23.7	26.1	27.3	29.6	31.1	33.1	35.6	37.0	36.0
	Ramon Castilla	695.2	16.6	17.7	15.9	13.8	14.5	11.3	10.3	8.3	6.0	3.9	2.0	20.0	18.2	22.8	0.0	1.7	3.7	4.3	7.2	9.4	11.6	12.8	15.0	17.2	19.2	21.8	22.3	24.5	26.4	27.4	29.9	31.5	33.3	35.9	37.2	39.4	41.1	40.4
	Tacna	705.8	21.4	19.9	18.1	15.7	16.3	13.3	12.3	10.3	8.1	6.1	4.1	1.9	1.7	0.0	2.9	3.5	6.4	8.6	10.8	12.3	14.8	16.7	18.6	20.8	21.9	24.1	26.0	27.0	29.4	31.0	33.0	35.0	36.3	39.3	40.8	41.3		
	Jr. Union	733.4	23.1	21.6	20.3	18.1	18.4	15.7	14.5	12.6	10.5	8.3	6.4	24.7	22.7	20.9	4.0	2.9	0.0	2.6	5.6	7.8	10.0	11.8	13.9	16.2	18.1	20.2	21.7	23.6	25.5	27.0	29.5	30.8	33.1	34.6	36.7	39.1	40.8	40.0
	Colmena	671.4	24.2	22.5	20.6	18.5	18.8	16.0	14.9	12.9	10.8	8.7	6.7	21.0	19.2	17.5	4.3	3.2	2.2	0.0	2.7	4.8	7.1	9.0	11.2	13.5	15.5	17.3	18.9	20.8	22.7	24.2	26.7	27.9	30.3	31.6	33.9	36.3	37.9	37.1
	Estación Central	638.1	26.7	24.3	22.8	21.3	21.4	18.8	17.7	15.6	13.5	11.5	9.4	6.6	4.8	3.2	7.3	6.2	5.1	2.6	0.0	3.6	5.9	7.7	9.9	12.2	14.1	15.9	17.5	19.2	21.1	22.6	25.0	26.2	28.5	30.0	32.0	34.4	36.0	37.5
	Estadio Nacional	607.0	28.4	26.4	24.9	23.4	23.5	20.9	19.8	17.7	15.6	13.6	11.5	8.7	6.9	5.3	9.4	8.3	7.2	4.7	3.6	0.0	1.9	3.8	6.0	8.3	10.2	12.0	13.5	15.3	17.2	18.7	21.3	22.5	24.9	26.0	28.0	30.9	32.7	34.1
	México	619.1	31.5	29.1	27.9	26.1	26.5	23.6	22.4	20.4	18.2	16.1	14.0	11.2	9.3	7.6	11.6	10.5	9.5	6.9	5.9	1.9	0.0	1.7	3.8	6.1	8.0	9.9	11.4	13.2	15.2	16.7	19.1	20.3	22.7	23.9	25.9	28.6	30.4	31.9
	Canada	615.4	31.7	30.5	29.0	27.5	28.3	25.1	24.0	21.9	19.8	17.8	15.7	12.7	10.8	9.2	13.6	12.5	11.5	8.9	7.7	3.7	1.7	0.0	1.9	4.1	6.0	7.9	9.3	11.1	13.0	14.4	16.8	18.0	20.3	21.9	23.7	26.2	28.0	29.3
	Javier Prado	636.3	35.0	33.1	31.7	30.1	30.8	27.7	26.6	24.5	22.3	20.3	18.1	15.4	13.5	11.8	15.8	14.7	13.6	11.1	10.0	6.0	3.8	1.9	0.0	2.0	3.9	5.7	7.1	8.9	10.8	12.3	14.7	15.9	18.3	19.8	21.7	24.1	25.8	27.4
	Canaval y Moreyra	647.8	36.8	34.9	33.3	31.9	33.4	29.5	28.3	26.3	24.2	22.1	20.1	17.1	15.3	13.7	18.1	16.9	15.9	13.4	12.1	8.1	6.1	4.1	1.9	0.0	1.7	3.5	4.9	6.7	8.6	10.0	12.4	13.7	15.9	17.6	19.2	21.8	23.5	25.0
	Aramburu	676.1	39.4	37.3	35.7	34.3	35.2	31.9	30.6	28.6	26.5	24.5	22.4	19.6	17.8	16.1	20.0	18.9	17.8	15.3	14.3	10.3	8.1	6.1	3.9	1.6	0.0	1.6	3.1	4.9	6.8	8.1	10.6	12.0	14.0	15.7	17.3	20.2	22.1	23.6
	Domingo Orué	687.7	38.7	38.5	36.9	35.5	36.8	33.1	31.9	29.8	27.8	25.8	23.7	20.8	19.0	16.8	21.8	20.6	19.6	17.1	15.8	11.8	9.8	7.8	5.7	3.4	1.6	0.0	1.3	3.1	5.0	6.5	9.0	10.2	12.5	13.9	15.8	18.3	20.2	21.9
	Angamos	721.2	41.2	41.0	39.8	37.9	39.0	35.5	33.9	31.9	30.1	27.3	25.3	22.2	20.3	18.6	24.6	23.4	22.4	20.2	17.8	13.7	12.1	9.1	7.6	5.0	3.3	1.3	0.0	1.4	3.3	4.7	7.2	8.5	10.6	12.3	13.9	16.4	18.3	19.7
	Ricardo Palma	733.2	43.4	41.6	40.2	38.5	40.2	36.2	35.0	33.0	30.9	29.0	27.0	24.1	22.2	20.6	24.8	23.7	22.6	20.1	18.9	14.9	12.9	10.8	8.7	6.4	4.5	2.7	1.5	0.0	1.6	3.1	5.6	6.7	9.1	10.6	12.5	14.7	16.6	18.2
	Benavides	771.3	45.4	43.5	42.2	40.5	42.2	38.1	36.9	35.0	32.9	30.9	28.9	26.1	24.3	22.6	27.1	26.0	24.9	22.6	20.9	16.9	15.0	12.8	10.8	8.4	6.6	4.7	3.5	1.7	0.0	1.2	3.6	4.8	7.2	8.7	10.6	12.8	14.7	16.3
	28 de Julio	798.4	46.6	44.8	43.2	41.8	43.8	39.5	38.4	36.3	34.2	32.4	30.3	27.5	25.7	24.1	28.4	27.3	26.3	23.9	22.3	18.3	16.5	14.3	12.2	9.9	8.0	6.1	5.1	3.2	1.4	0.0	2.2	3.3	5.7	7.2	9.0	11.3	13.2	14.7
	Plaza de Flores	843.3	47.2	46.6	45.8	43.5	46.3	41.2	40.7	37.9	35.9	34.7	32.6	29.9	28.0	26.4	30.0	29.5	27.4	25.4	24.3	20.7	18.3	16.2	14.5	12.3	10.4	8.5	7.6	5.5	3.8	2.4	0.0	1.3	3.6	5.2	6.9	9.3	11.2	12.6
	Balta	874.6	49.3	48.3	46.7	45.0	47.5	42.8	41.6	39.5	37.5	35.8	33.7	31.3	29.4	27.7	32.2	31.1	29.9	27.7	25.7	21.7	19.7	17.7	15.5	13.2	11.2	9.4	8.7	6.3	4.5	3.4	1.6	0.0	2.0	3.5	5.5	7.5	9.5	11.0
	Bulevar	918.4	51.2	49.9	48.8	46.6	49.2	44.5	43.8	41.3	39.1	37.8	35.7	33.2	31.3	29.6	32.6	32.7	30.5	27.9	27.4	23.8	21.3	19.4	17.6	15.4	13.5	11.7	10.7											

Matrix editor (Matrix: 5 NTR (BRT_Cod_Dseg BRT_Name_Dseg))

38 x 38		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38			
	Name	Naranjal	Izaguirre	Pacifico	Independencia	Los Jazmines	Tomás Valle	El Milagro	Honorio Delgado	UNI	Parque del Trabajo	Caquetá	2 de Mayo	Quilca	España	Ramon Castilla	Tacna	Jr Union	Colmena	Estación Central	Estadio Nacional	México	Canada	Javier Prado	Canaval y Moreyra	Aramburu	Domingo Orué	Angamos	Ricardo Palma	Benavides	28 de Julio	Plaza de Flores	Balta	Bulevar	Union	Escuela Militar	Terán	Rosario de Villa	Matellini			
	Sum	12.8	12.4	14.4	14.4	13.5	12.2	16.0	14.5	12.1	11.4	8.2	18.6	12.6	8.8	5.4	5.5	4.9	4.2	1.3	8.5	12.8	12.3	11.9	9.9	12.8	12.8	14.0	15.0	10.7	13.5	13.4	13.4	15.3	15.3	15.4	15.5	15.6	15.7			
1	Naranjal	8.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		
2	Izaguirre	22.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	Pacifico	22.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	Independencia	22.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	Los Jazmines	22.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	Tomás Valle	22.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	El Milagro	20.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	Honorio Delgado	19.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	UNI	15.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	Parque del Trabajo	14.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	Caquetá	10.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	2 de Mayo	18.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	Quilca	14.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	España	9.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	Ramon Castilla	5.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
16	Tacna	5.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
17	Jr Union	4.5	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
18	Colmena	3.9	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
19	Estación Central	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	Estadio Nacional	8.5	0.5	0.5	0.6	0.6	0.6	0.5	0.6	0.6	0.5	0.6	0.5	0.7	0.7	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
21	México	13.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.8	0.6	1.0	0.8	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
22	Canada	13.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.7	0.5	1.0	0.7	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
23	Javier Prado	12.5	0.7	0.7	0.8	0.8	0.8	0.7	1.0	1.0	0.8	0.8	0.5	1.0	0.8	0.6	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
24	Canaval y Moreyra	5.2	0.1	0.1	0.4	0.4	0.3	0.1	0.7	0.4	0.1	0.4	0.1	0.6	0.2	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
25	Aramburu	6.8	0.3	0.3	0.5	0.5	0.5	0.3	0.7	0.6	0.3	0.5	0.2	0.6	0.4	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
26	Domingo Orué	8.9	0.5	0.5	0.7	0.7	0.6	0.5	0.8	0.7	0.6	0.6	0.4	0.8	0.6	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
27	Angamos	4.7	0.1	0.1	0.4	0.4	0.3	0.1	0.6	0.4	0.1	0.4	0.1	0.5	0.2	0.0	0.2	0.2	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
28	Ricardo Palma	7.2	0.4	0.4	0.5	0.5	0.5	0.3	0.7	0.6	0.3	0.5	0.3	0.6	0.5	0.3	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
29	Benavides	9.7	0.6	0.6	0.7	0.7	0.7	0.6	0.8	0.7	0.6	0.6	0.5	0.8	0.6	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
30	28 de Julio	11.3	0.7	0.7	0.8	0.8	0.8	0.7	0.9	0.8	0.7	0.7	0.6	0.9	0.6	0.5	0.3	0.3	0.3	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
31	Plaza de Flores	4.1	0.1	0.1	0.3	0.3	0.2	0.0	0.5	0.3	0.0	0.3	0.0	0.4	0.2	0.0	0.2	0.2	0.2	0.2	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
32	Balta	8.8	0.5	0.4	0.6	0.6	0.5	0.4	0.7	0.6	0.4	0.6	0.4	0.7	0.5	0.3	0.3	0.3	0.3	0.3	0.0	0.0	0.1	0.1																		

Matrix editor (Matrix '5 NTR (BRT_Cod_Dseg BRT_Name_Dseg)')

38 x 38	Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38					
	Sum	3.8	2.8	3.5	2.0	3.2	2.2	2.5	2.0	1.8	2.3	1.9	5.7	5.8	4.2	4.8	5.2	4.7	4.1	1.5	2.4	2.7	2.3	2.6	2.6	3.1	3.5	3.3	3.6	3.7	4.0	4.5	4.8	4.6	5.5	5.6	7.1	5.8	5.8					
1	Naranjal	11.6	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.2	0.1	0.3	0.1	0.4	0.5	0.2	0.2	0.2	0.4	0.6	0.2	0.4	0.4	0.6	0.7	0.7	0.8	0.8	0.8	0.8	0.7	0.7					
2	Izaguirre	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.2	0.4	0.4	0.4	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.4	0.3	0.2	0.1			
3	Pacifico	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.4	0.3	0.3	0.2		
4	Independencia	5.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.2	0.1	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.3	0.3	0.4	0.3	0.2	0.3	0.3	0.2	0.3	0.2	0.3		
5	Los Jazmines	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.2	0.2	0.3	0.3	0.3		
6	Tomás Valle	4.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.3		
7	El Milagro	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
8	Honorio Delgado	3.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
9	UNI	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
10	Parque del Trabajo	3.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
11	Caquetá	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2 de Mayo	5.5	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.5	0.3	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
13	Quilca	4.5	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.6	0.5	0.4	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
14	España	3.8	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.4	0.4	0.3	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
15	Ramon Castilla	5.1	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.1	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
16	Tacna	5.4	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	1.1	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
17	Jr. Union	4.6	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.6	0.7	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
18	Colmena	3.8	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.4	0.4	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
19	Estación Central	2.0	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
20	Estadio Nacional	2.1	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
21	México	2.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
22	Canada	1.8	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
23	Javier Prado	1.9	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
24	Canaval y Moreyra	1.8	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
25	Aramburu	1.8	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
26	Domingo Orué	1.8	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
27	Angamos	2.7	0.1	0.2	0.2	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
28	Ricardo Palma	1.5	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
29	Benavides	1.6	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.1	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
30	28 de Julio	2.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
31	Plaza de Flores	1.9	0.1	0.0	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
32	Balta	2.4	0.2	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.1	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		
33	Bulevar	2.8	0.2	0.1	0.1																																							

