

ANALISIS DEL TRASPORTE PUBLICO BRT VIA
DEL TERMINAL AEREO DE BOGOTA

CARLO MAXIMILIANO TOVAR RODRIGUEZ

TRABAJO DE GRADO PARA OPTAR POR EL TÍTULO DE INGENIERO
INDUSTRIAL

UNIVERSIDAD SANTO TOMÁS
DIVISIÓN DE INGENIERÍAS
FACULTAD DE INGENIERÍA INDUSTRIAL
BOGOTÁ
2019

ANÁLISIS DEL TRASPORTE PÚBLICO BRT VIA
DEL TERMINAL AÉREO DE BOGOTÁ

CARLO MAXIMILIANO TOVAR RODRIGUEZ

TRABAJO DE GRADO PARA OPTAR POR EL TÍTULO DE INGENIERO
INDUSTRIAL

DIRECTOR
LUIS MANUEL PULIDO RICO
INGENIERO INDUSTRIAL

UNIVERSIDAD SANTO TOMÁS
DIVISIÓN DE INGENIERÍAS
FACULTAD DE INGENIERÍA INDUSTRIAL
BOGOTÁ
2019

TABLA DE CONTENIDO

TABLA DE CONTENIDO	4
AGRADECIMIENTOS	7
RESUMEN	8
ABSTRACT	8
LÍNEA DE INVESTIGACIÓN	9
0. INTRODUCCIÓN	9
1. PLANTEAMIENTO DEL PROBLEMA	10
2. PREGUNTA DE INVESTIGACIÓN	14
3. JUSTIFICACIÓN	14
4. OBJETIVOS	15
4.1. Objetivo general	15
4.2. Objetivos específicos	15
5. ALCANCE	16
6. GLOSARIO	17
7. MARCO HISTÓRICO	19
8. MARCO TEÓRICO	22
8.1. El problema del transporte	22
8.2. Metodología general	24
8.2.1. Visiones metodológicas	25
8.2.2. Actividades de planeación estratégico	26
8.2.3. Actividades de planeación tácticas y operativas	29
8.3. Inteligencia artificial	30
8.3.1. Machine learning	30
8.3.2. Deep learning	32
8.3.3. Regresión estadística y fundamentos de redes neuronales	33
8.3.4. Función de activación	34
8.4. Enfoque de ruteo	35
8.5. Otros enfoques	38
8.6. Análisis de datos atípicos	39
9. MARCO METODOLÓGICO	45
9.1. Tipo de estudio:	45
10. Método de investigación	45
10.1. Etapa 1. Análisis de metodologías:	45
10.2. Etapa 2. Datos de entrada y planteamiento:	45
10.3. Etapa 3. Análisis exploratorio y comportamiento de la demanda	45
10.4. Etapa 4. Propuesta cuantitativa de respuesta a la demanda	46
10.5. Etapa 5. Prototipo y validación	46
11. DESARROLLO	47
11.1. Etapa 1. Análisis de metodologías	47
11.1.1. Metodología Formulación Analítica.	47
11.1.2. Metodología de optimización Heurística de autor	47
11.1.3. Metodología de optimización heurística VRP	48
11.1.4. Metodología de inteligencia artificial - Algoritmos genéticos	48

11.1.5.	Metodología de inteligencia artificial - Redes Neuronales.	48
11.1.6.	Proceso analítico jerárquico, AHP.	48
11.1.7.	Selección de la metodología.	53
11.2.	Etapa 2. Datos de entrada y planteamiento del sistema de transporte de pasajeros.	54
11.2.1.	Información externa y suposiciones.	54
11.2.2.	Planteamiento modelo.	55
11.3.	Etapa 3. Análisis exploratorio y comportamiento de la demanda.	59
11.3.1.	Media.	59
11.3.2.	Estadística descriptiva agrupación por meses.	62
11.3.3.	Estadística descriptiva por días de la semana.	63
11.3.4.	Estadística descriptiva por horas del día.	64
11.3.5.	Estadística descriptiva por estaciones de la troncal.	65
11.3.6.	Distribución de los datos.	66
11.3.7.	Variabilidad.	67
11.3.8.	Conclusiones. Estadística Descriptiva.	68
11.4.	Pruebas estadísticas detalladas de la demanda.	69
11.4.1.	Pruebas de homogeneidad.	69
11.4.2.	Pruebas de independencia.	72
11.4.3.	Clúster del comportamiento de la demanda.	80
11.4.4.	Resultados de la estructuración.	88
12.	Etapa 4. Propuesta cuantitativa de respuesta a la demanda.	89
12.1.	Variables.	89
12.1.1.	Matriz de ruta.	92
12.1.2.	Matriz de distancias.	92
12.1.3.	Matriz de tiempos.	93
12.1.4.	Matriz de origen destino:	94
12.1.5.	Elección del usuario.	96
12.2.	Pronóstico de la demanda.	99
12.2.1.	Horizonte de pronóstico:	99
12.2.2.	Métodos de pronóstico.	100
12.3.	Componentes del pronóstico.	102
12.3.1.	Componente Ciclicidad.	102
12.3.2.	Componentes De Aleatoriedad Modelo De Clústeres.	102
12.3.3.	Componente Estacionalidad y tendencia.	119
12.4.	Estructura del modelo.	123
12.4.1.	Metodología de pronóstico.	124
12.4.2.	Error del pronóstico.	125
13.	Etapa 5. Prototipo y validación.	126
13.1.	Alternativas de asignación.	126
13.2.	Escenarios.	128
14.	CONCLUSIONES.	133
15.	BIBLIOGRAFÍA.	136
	ANEXOS.	147
	TABLAS.	148

FIGURAS.....151

AGRADECIMIENTOS

Agradezco a mi familia con todo su cariño y a la comunidad Tomasina por brindarme un espacio propicio para soñar, reconozco a cada uno de los maestros de la facultad quienes brindaron grandes enseñanzas las cuales motivaron este proyecto de investigación. Resalto el acompañamiento metodológico del ingeniero industrial Luis Manuel Pulido Rico, profesor de logística y optimización de la Universidad Santo Tomas, quien apporto su amplia experiencia y conocimiento. Reconozco al ingeniero Oscar Eduardo Diaz Olariaga profesor en ingeniería civil e ingeniería aeronáutica en la Universidad Santo Tomas por su labor en investigación interdisciplinaria y por brindar al proyecto insumos fundamentales para este estudio. También agradezco al ingeniero Stephen P. Mattingly profesor asociado de la universidad de Texas en Arlington por compartir su visión del transporte público.

RESUMEN

Este trabajo es un estudio del caso Transmilenio, el cual tiene como objetivo principal el acercamiento metodológico para describir el funcionamiento y evaluar el rendimiento de la red de transporte público, específicamente en la “Troncal Portal Dorado”; toma como base los datos históricos del sistema implementado en la ciudad de Bogotá, recolectados el año 2017.

El sistema estudiado consiste en el tránsito de vehículos tipo bus, los cuales se desplazan por un carril exclusivo. La Investigación está basada en el análisis de sección delimitada del sistema, con un enfoque a la respuesta de la demanda y plantea la posibilidad de comprender momentos típicos o atípicos, orientados hacia la eficiencia del sistema.

Como herramienta de apoyo se desarrolló un aplicativo en Excel el cual permite evaluar diferentes escenarios, en función del costo para el operador y el costo para el usuario, según el planteamiento propuesto.

Palabras clave: bus, demanda, vía exclusiva, descripción, transporte.

ABSTRACT

This work is a study of the TRANSMILENIO case, whose main objective is the methodological approach to describe the operation and evaluate the performance of the public transport network, specifically in the “Troncal Portal El Dorado”; It is based on the historical data of the system implemented in Bogotá city, collected in 2017.

The system studied consists of the transit of bus-type vehicles, which move along an exclusive lane. Research is based on the delimited section analysis of the system, with a focus on the user's level response, and raises the possibility of understanding typical or atypical moments, oriented towards the efficiency of the system.

As a support tool, an application was developed in Excel which allows evaluating different scenarios, depending on the cost to the operator and the cost to the user, according to the proposed approach.

Key words: bus, demand, exclusive lane, description, transport.

LÍNEA DE INVESTIGACIÓN

La línea de investigación de este trabajo está enfocada al mejoramiento de procesos, relacionados con la optimización de operaciones, la toma de decisiones y estrategias de distribución en transporte de pasajeros.

0. INTRODUCCIÓN

Mediante un análisis, el presente trabajo investiga el rendimiento en varios aspectos del sistema de transporte de pasajeros de la ciudad de Bogotá “TRANSMILENIO”, como uno de los principales factores de infraestructura que definen la competitividad de esta ciudad, y así lo indica el reporte de competitividad realizado por el foro económico mundial (2018). Como el transporte es un reto logístico que no es propio de un solo sector de la sociedad (*Eriksson Linnea, 2016*), este adquiere gran importancia porque se identifica como una clave para alcanzar entornos propicios para el desarrollo económico, en el cual participan distintos grupos de interés (*BARFOD Michael Bruhn, 2018*).

Para realizar el análisis, se entiende que es necesario definir un alcance y una metodología o combinación de técnicas presentes en la literatura, así pues, distintas alternativas son recopiladas y expuestas como posibles caminos a seguir en esta investigación. Con base en lo anterior, se selecciona una o la combinación de varias metodologías que pretenden describir mediante una acotación el sistema de estaciones de buses; porque para encontrar relaciones uniformes se requiere evidenciar el efecto de eventos atípicos que alteran el funcionamiento esperado. Para el caso en particular, la sección tema de estudio corresponde a la línea que conecta el centro histórico de la ciudad, universidades y el terminal aéreo de Bogotá, para ello se recurre al uso de datos reales de la demanda del año 2017, registros que provienen del operador.

Teniendo en cuenta que el sistema de transporte de pasajeros responde a la demanda, en este escrito se realiza el análisis estadístico que permite describir su comportamiento general, delimitado a una sección denominada Portal el dorado. En segundo término, basado en los resultados anteriores, se procede a generalizar el comportamiento de la demanda en grupos con tendencias similares, los cuales permiten establecer el comportamiento de una estación en un día específico y establecer posibles predicciones de comportamientos normales y no normales por medio de metodologías de pronósticos y técnicas de inteligencia artificial.

Al identificar las variables del sistema por medio de funciones objetivo, y escoger una metodología (Optimización de partículas, Particle Swarm Optimización PSO), es posible describir el comportamiento de la troncal, lo que permite diseñar un aplicativo en lenguaje Visual Basic – Excel para evaluar distintos escenarios.

1. PLANTEAMIENTO DEL PROBLEMA

Teniendo en cuenta la necesidad de comprender una de las principales infraestructuras de transporte de pasajeros, que influye en el desarrollo de la ciudad de Bogotá, se busca analizar cómo funciona la operación típica de una solución de “Transporte Masivo Rápido (TMR), concretamente tipo “Bus Rapid Transit” (BRT), cuya característica principal son las vías o carriles exclusivos para buses. Se toma como caso de estudio la sección denominada Troncal Portal Dorado, que hace parte del sistema Transmilenio solución de transporte masivo implementada en la ciudad de Bogotá, Colombia.

Ilustración 1. Estación Sistema BRT en la ciudad de Bogotá.



Nota. La imagen muestra la vía exclusiva y la infraestructura del sistema para uso de buses del BRT, Imagen tomada por Amayo, O. (2011). Fotografía Transmilenio. Recuperado de: <https://flic.kr/p/awBR9x>

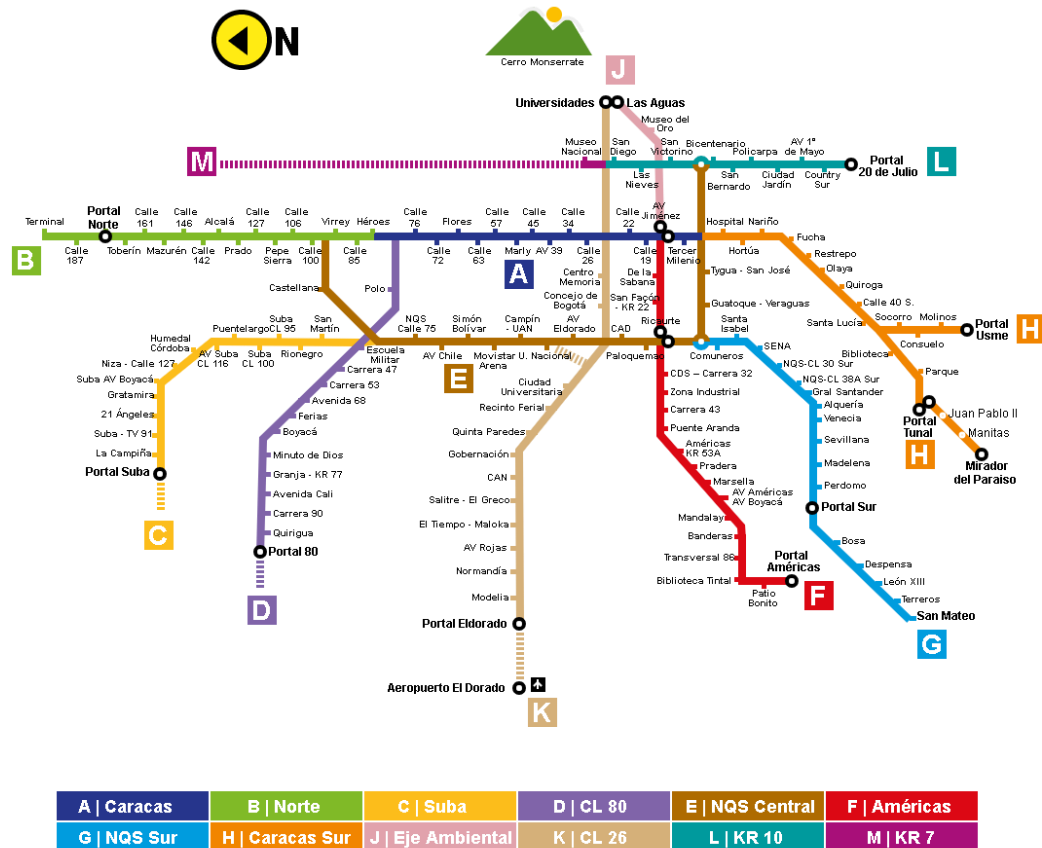
Este análisis está basado en la recolección de datos sobre el ingreso de pasajeros en el sistema del año 2017, y pretende establecer si la operación del día está relacionada con las variaciones de la demanda de acuerdo con la estación y el periodo en el que se analiza. Con la premisa de resolver el siguiente interrogante: ¿Cómo analizar el rendimiento de una sección del sistema de transporte de pasajeros BRT en la ciudad de Bogotá, con base en el comportamiento de los elementos y variables que conforman una troncal?

El sistema integrado de transporte de la ciudad de Bogotá cuenta con diferentes troncales en las que operan distintas rutas, la sección del portal el dorado está conformada por catorce estaciones las cuales son numeradas a continuación y correspondiente al trayecto de la vía exclusiva que corre a lo largo de la calle 26, y opera entre las carreras 3 y 103.

- | | |
|------------------------------|----------------------------|
| 1. Portal Eldorado | 8. Gobernación |
| 2. Modelia | 9. Quinta Paredes |
| 3. Normandía | 10. Corferias |
| 4. Av. Rojas | 11. Ciudad Universitaria |
| 5. El Tiempo | 12. Plaza de la Democracia |
| 6. Salitre El Greco | 13. Centro Memoria |
| 7. CAN Centro Administrativo | 14. Universidades |

Nacional

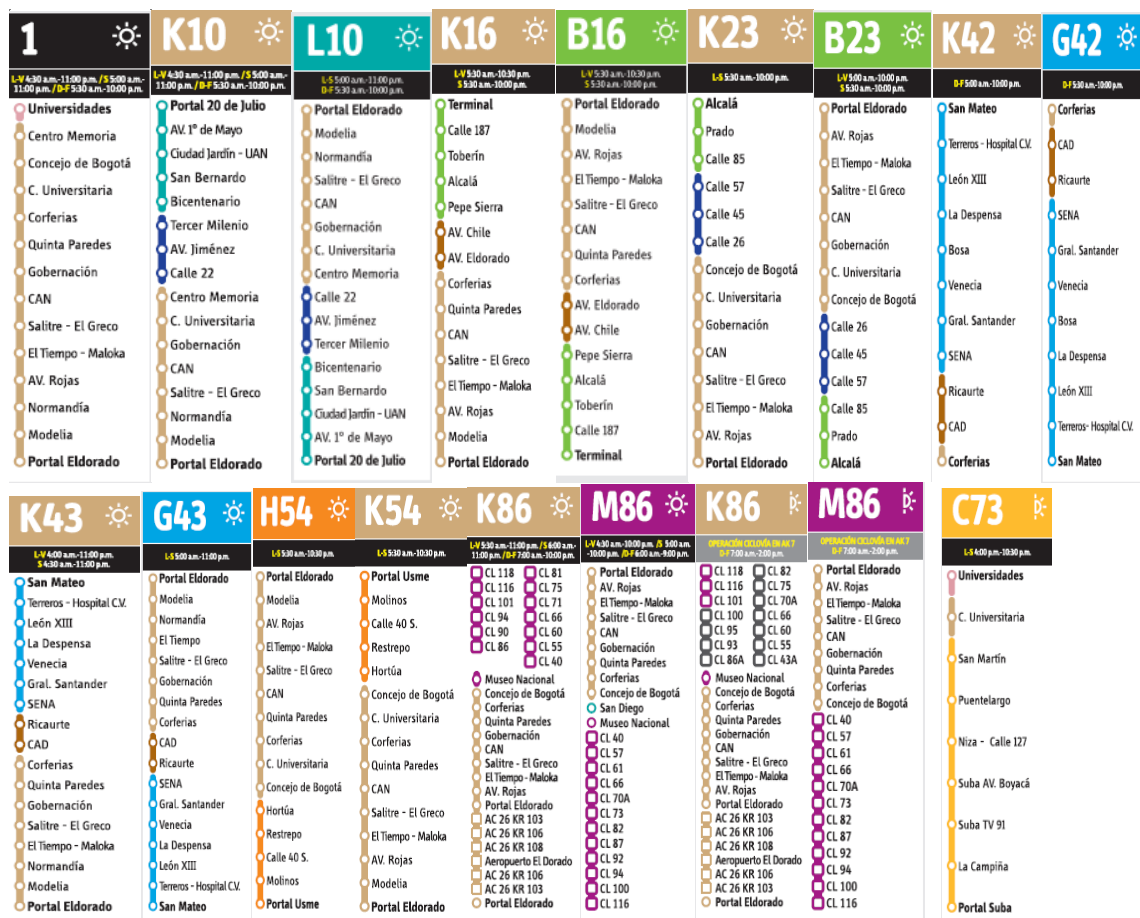
Figura 1. Mapa Sistema Transmilenio.



Nota. El mapa ilustra las diferentes troncales que componen el sistema de buses BRT, en cual se encuentra la troncal portal el dorada señalada con la letra K, Quintero, N. 2019. Mapa Interactivo de Transmilenio recuperado de: https://www.transmilenio.gov.co/publicaciones/150402/publicacionesmapa_interactivo_de_transmilenio/

En este segmento funcionan las rutas diseñadas por el operador que conectan la red de estaciones en sentido oriente-occidente y occidente oriente, programadas para funcionar en determinado horario y con paradas establecidas. A continuación, se relacionan figuras del sistema Transmilenio (2019) que describen las rutas operativas o posibles elecciones que tiene un usuario. Diagramas que muestran el nombre de la ruta, el horario en el cual funciona, un color característico que indica cual es el sector de la ciudad en el que opera y las estaciones en las que realiza paradas.

Figura 2. Diagramas de paradas de las rutas en la troncal Portal El Dorado.



Nota. Los diagramas están relacionados con la configuración del sistema Transmilenio para el año 2019, permiten observar que una ruta puede operar en más de una sección del sistema. Adaptado de “Mi Plan de viaje de Transmilenio 15 de octubre 2019”. Recuperado de: <https://www.transmilenio.gov.co/publicaciones/149036/mi-plan-viaje-de-transmilenio/>

El sistema de estaciones es conectado por vehículos que responden a niveles de demanda, que involucran variables endógenas y exógenas a la operación. El sistema analizado cuenta con diferentes tipos de vehículos, que se relacionan en la [Tabla 1. Tipología vehículos que funcionan en las troncales](#). Solo se tienen en cuenta aquellos que transitan en la sección analizada, los cuales se encargan de movilizar a los pasajeros en la red de estaciones.

Tabla 1. *Tipología Vehículos Que Funcionan En Las Troncales.*

Tipo de Vehículo	Descripción	Capacidad Número de pasajeros	Imagen
Dual	Buses que están designados para compartir secciones de vía compartida con carros particulares.	80	
Articulado	Cuenta con una articulación gris en la mitad del bus, llamada fuele.	160	
Biarticulado	Está compuesto por dos articulaciones, lo cual significa que la capacidad del vehículo es mayor. Tienen una longitud de 27.2 metros, lo cual lo convierte en el bus biarticulado más largo del mundo.”	250 y 270	

Nota. Descripción y capacidad de los vehículos que funcionan en las troncales del sistema Transmilenio. En el interior y exterior del bus se pueden observar tableros electrónicos que indican número de ruta, próximas paradas y destino (Información obtenida de la página oficial del sistema Transmilenio).

Debido a la complejidad del sistema y las diferentes variables involucradas, se propone estudiar dicha complejidad por medio de una medida basada en registros históricos año 2017. Siguiendo la anterior premisa, la cantidad de pasajeros que entran en una estación por día será implementada como herramienta para generalizar el comportamiento de un punto del sistema, lo cual a su vez permite analizar la sección en un tiempo específico.

2. PREGUNTA DE INVESTIGACIÓN.

¿Cómo analizar el rendimiento de una sección del sistema de transporte de pasajeros BRT en la ciudad de Bogotá, con base en el comportamiento de los elementos y variables que conforman una troncal?

3. JUSTIFICACIÓN

- **Académica:** Este trabajo busca ser una aplicación de los saberes aprendidos a lo largo del pregrado de Ingeniería Industrial de la Universidad Santo Tomás, siguiendo el principio de aportar al avance cultural de la sociedad mediante el conocimiento.
- **Competitividad:** El transporte es uno de los pilares del desarrollo de una sociedad, y por tal motivo es una actividad social que abre nuevas perspectivas para entender el desarrollo de las actividades sociales y comerciales de la ciudad de Bogotá.
- **Eficiencia:** A partir del uso de los recursos público-privados de un sistema, esta investigación apoya la correcta utilización de los recursos de la ciudad, cosa que permite un entendimiento más detallado del comportamiento generando y apoya el desarrollo de bases para la toma de decisiones que posibiliten el ahorro frente la sobreutilización y la subutilización de los recursos estudiados en torno a la integración de los modos de transporte, los objetivos de gobierno, los grupos sociales y las instituciones (*Santos, Behrendt, & Teytelboym, 2010*).
- **Servicio:** La ciudad de Bogotá tiene un amplio potencial de crecimiento, lo cual implica que la calidad de vida debe estar acorde con su potencial de desarrollo. Por este motivo, las visiones complementarias pueden aportar a la mejora de los niveles de servicio del sistema de transporte y promover así un mayor bienestar para los usuarios.
- **Ambiente:** En la actualidad, las actividades de transporte funcionan gracias a distintas fuentes energéticas, como los combustibles fósiles y la energía eléctrica, acompañadas de una gran red de infraestructura limitada. Políticas basadas en la academia, adecuadas para el manejo de dichos recursos también aportan a iniciativas proactivas para prevenir el deterioro del entorno urbano y del medio ambiente (*Rogge, van der Hurk, Larsen, & Sauer, 2018*), y promover el crecimiento amigable con el medio ambiente, en busca de una sostenibilidad multidimensional (*Santos & Ribeiro, 2013*).

4. OBJETIVOS

4.1. Objetivo general

Describir el comportamiento de un segmento del sistema de transporte masivo Transmilenio mediante el uso de una metodología o combinación de metodologías orientadas hacia la evaluación del rendimiento de un sistema de transporte de pasajeros caracterizado por el uso de buses que transitan por carriles exclusivos.

4.2. Objetivos específicos

1. Recopilar y analizar las cinco principales alternativas metodológicas basadas en variables usadas para el estudio del rendimiento de sistemas de transporte de pasajeros.
2. Comparar las diferentes alternativas para el estudio del rendimiento de sistemas de transporte de pasajeros entre sí.
3. Describir el sistema de buses BRT de la ciudad de Bogotá, “Transmilenio”, en términos de la demanda de pasajeros.
4. Simplificar en términos de las variables el funcionamiento de un segmento del sistema de buses de tránsito rápido BRT de la ciudad de Bogotá, Transmilenio, y específicamente el funcionamiento del “Portal El Dorado”.
5. Evaluar estado del sistema mediante metodología elegida.

5. ALCANCE

Con este trabajo se pretende establecer el comportamiento y el rendimiento de un segmento del sistema de transporte público de pasajeros de la ciudad de Bogotá, delimitado al tipo de servicio troncal, entendida una troncal como un corredor exclusivo para el uso de vehículos tipo bus y expuesto en un segmento que comprende 14 estaciones que se conectan mediante una vía exclusiva a lo largo de la calle 26 de la ciudad, grupo que se denomina “Troncal portal Dorado”, el cual opera entre las carreras 3 y 103, siendo las calles aquellas vías que conectan la ciudad en sentido norte-sur y las carreras aquellas vías que conectan la ciudad en el sentido oriente-occidente, basado en la selección de metodologías y la elaboración de un aplicativo que simplifique el manejo de la información y permita analizar fácilmente la operación.

El documento tiene un enfoque aplicativo desde dos puntos de vista: el primero para la representación de un sistema y el segundo la detección de datos atípicos. La revisión utiliza el comportamiento de los pasajeros que se concentran en un punto geográfico de la ciudad y en el reconocimiento de momentos que afectan la normalidad del funcionamiento de la red de estaciones que toma como fuente principal un set de datos recolectados por el sistema de Transmilenio del año 2017.

En este trabajo no se tiene en cuenta otras infraestructuras del sistema integrado de transporte, como semaforización, sincronización entre vehículos y tarjeta inteligente, ni tampoco otros servicios de la red, como servicios alimentadores compuestos por buses urbanos, ciclo rutas y otros modos de transporte, ni la complejidad de interacción con los vehículos particulares.

6. GLOSARIO

Las definiciones hechas a continuación están orientadas hacia temática de transporte.

Algoritmo: Conjunto ordenado y finito de operaciones que permite hallar la solución de un problema.

Anomalía: Como una observación o grupo de observaciones con notables inconsistencias respecto del grupo; otras definiciones vinculan el concepto con patrones en los datos que no se ajustan a un comportamiento normal definido.

Calidad: Es el rendimiento general medido o percibido del servicio de tránsito desde el punto de vista de los pasajeros; el concepto puede estar enfocado hacia la disponibilidad del tráfico, es decir, si el servicio es una opción para un viaje determinado, y en segundo lugar hacia el confort y la conveniencia.

Capacidad: En términos de sistemas de transporte, y refleja el máximo número de vehículos, personas, o ambos, que pueden viajar pasando por una ubicación en particular, en un tiempo determinado y bajo condiciones específicas. Existen siguientes cuatro conceptos de capacidad:

Capacidad diseñada: refleja el número de personas que pueden ser atendidas con una calidad de servicio específica.

Capacidad máxima: refleja el mayor número de personas o vehículos que pueden ser atendidos bajo una circunstancia; en el caso de las personas, se refiere a la cantidad de personas que pueden caber físicamente en un vehículo.

Capacidad de personas: expresa cuántas personas pueden pasar por un punto en una hora y se mide en personas por hora.

Capacidad de vehículo: Es la cantidad de personas puede transportar un vehículo frente a un nivel de carga de pasajeros diseñada. la capacidad de pasajeros se mide en personas por vehículo.

Densidad de vehículos: número de vehículos que circulan en un instante de tiempo determinado por la vía (Callejo, 2009).

Heurística: definida como aproximaciones mediante algoritmos para proporcionar resultados óptimos (Ng, Lee, Chan, & Lv, 2018). O según la real academia de la lengua española una manera de buscar la solución de un problema mediante la aplicación métodos no rigurosos, como por tanteo y reglas empíricas.

Horizonte de pronóstico: Se define como el rango de tiempo por delante para el cual se lleva a cabo la predicción de tráfico.

Inteligencia artificial: La inteligencia artificial (IA) es un área amplia de la informática que hace que las máquinas funcionen como un cerebro humano. Se utiliza para abordar problemas que son difíciles de aclarar con el uso de computadores tradicionales (Abduljabbar, Dia, Liyanage, & Bagloee, 2019).

Machine learning: El machine learning o aprendizaje automático es el campo de la ciencia de la computación que trata de la simulación del comportamiento inteligente en computadoras (Bagheri, Akbari, & Mirbagheri, 2019).

Principio de Parsimonia: Método que expone que dadas las mismas condiciones la explicación que suele resultar correcta es la más sencilla.

Pronóstico: Una secuencia plausible de eventos que conducen a una situación futura, sin la necesidad de incluir declaraciones sobre la probabilidad de esos hechos. Un análisis de la situación actual y una conexión entre los desarrollos futuros y la situación actual (Malone, Verroen, Korver, & Heyma, 2001. En transporte se realiza una descripción integrada del futuro, que presta atención a los desarrollos en todos los factores que afectan la demanda en los viajes.

TMR: Un espectro de modos de transporte público urbano que utilizan carretera fija específica o de manera exclusiva y separada, una pista potencialmente utilizable por otros modos (como los metros, los ferrocarriles suburbanos, el transporte liviano rápido y las vías segregadas para autobús). (Departamento de Transporte y Desarrollo Urbano del Banco Mundial, 2002).

Troncal: Corredor exclusivo para el uso de vehículos tipo bus (Transmilenio S.A., 2019).

7. MARCO HISTÓRICO

En el desarrollo urbano la movilidad desempeña un papel fundamental debido a que puede favorecer la competitividad, el desarrollo comercial y la actividad económica (Anguita Rodríguez, Duarte Monedero, & Flores Ureba, 2014). Por tal motivo esta revisión se concentra en el transporte urbano de pasajeros con el enfoque de crear un contexto orientado al uso de sistemas de buses que transitan por carriles exclusivos, buscando así describir, de lo general a lo particular, la solución presente en la ciudad de Bogotá, Colombia.

El transporte en una ciudad puede entenderse desde diferentes visiones, como los efectos y relación entre el uso del suelo y la aplicación de diferentes tecnologías, o también como la controversia entre la planeación pública, el mercado, la infraestructura y el diseño de los sistemas implementados, lo cual, a su vez, tienen en cuenta la estructura y configuración de las ciudades, que responde a variables espaciales y sociales (Murata, Delgado Campos, & Suárez Lastra, 2017), donde la eficiencia y la eficacia del transporte representan un reto trascendental en relación con temas de gestión público-privada relacionados en el artículo Observatorio de costes y financiación del transporte urbano colectivo: un programa de investigación, como los siguientes: (Sánchez Toledano, Carrasco Díaz, & Sánchez Toledano, 2014)

- Financiación (subvenciones y tarifas).
- Calidad del servicio (puntualidad y frecuencia).
- Innovación y tecnología (sistemas de información al usuario y sistemas de apoyo a la explotación). - Oferta del servicio (número de viajeros, diseño de la red, percepción del usuario del servicio).
- Costes (personal, combustibles).
- Necesidades de información para la gestión (absentismo, productividad del personal, costes, rentabilidad y rendimiento de la flota).

Por lo tanto, un factor para tener en cuenta es el crecimiento de la población en las áreas urbanas, que ha expandido el rol de los sistemas de transporte para satisfacer las necesidades de movilidad (Rizelioğlu & Arslan, 2019). En respuesta a ello, diferentes tipos sistemas son diseñados e implementados, entre ellos el TMR, que es el centro de esta revisión y es entendido como un espectro de modos de transporte público urbano que utilizan carretera fija específica o, de manera exclusiva y separada, una pista potencialmente utilizable por otros modos (como los metros, los ferrocarriles suburbanos, el transporte liviano rápido y las vías segregadas para autobús). (Departamento de Transporte y Desarrollo Urbano del Banco Mundial., 2002).

Para entender mejor las características del sistema a estudiar, en la bibliografía se relacionan fuentes sobre el origen histórico de este tipo de transporte, ya que el desarrollo de este sistema a través del tiempo no es el objeto de este estudio. A

continuación, describimos de una forma general los efectos de éste en el mundo y en Colombia.

Históricamente el uso de buses en el mundo ha tenido un enorme desarrollo, que permitió la aparición del modelo BRT, que abarca experiencias en Suramérica, Norteamérica, sudeste asiático, China, Australia y, ahora, progresivamente en África, India y Europa. Sistemas que tuvieron un auge bajo el desarrollo de tecnologías guiadas y no guiadas, en las cuales las no guiadas dependen de un conductor. Gracias a las primeras implementaciones, ellas inspiraron el desarrollo del modelo y las infraestructuras más complejos, entre los cuales resaltan las implementaciones hechas en las ciudades brasileñas que promovieron el sistema en la región. (Deng & Nelson, 2011).

Tomando como ejemplo a China, que presenta grandes retos por su excesiva concentración de la población, se ve la implementación de alternativas como los sistemas de transporte basados en rieles, pero debido a los costos de capital de éstos, la movilidad se complementa con los sistemas BRT, e incentivados por las lecciones aprendidas de 44 sistemas BRT alrededor del mundo, en la actualidad 13 ciudades chinas cuentan con este tipo de solución (Deng, Ma, & Wang, 2013). De esta forma se puede apreciar la acogida que tiene el sistema de transporte por características, así como su versatilidad y bajos costos de inversión.

Por otra parte, el contexto latinoamericano no es ajeno a las problemáticas del transporte; por ejemplo, la ciudad de Santo Domingo, capital de República Dominicana, tiene una extensión de 755 km², extensión denominada como la Gran Santo Domingo, encierra el distrito nacional, que ocupa 228 km², donde habitan, según el último censo, 10.433 personas por kilómetro cuadrado. Esta ciudad, con un crecimiento poco ordenado, se enfrenta a cuellos de botella vehiculares causados por el crecimiento de su parque automotor a un ritmo de setenta mil vehículos anuales.

Tal superficie debe responder a la demanda de tres millones de viajes diarios, que en su mayoría son respondidos por un sistema público con 2.000 vehículos capaces de satisfacer un promedio de 170 puestos. Pero la falta de inversión y coordinación de su transporte ha incrementado la problemática de la movilidad, que se convierte en un problema multidimensional que afecta directamente el bienestar de los ciudadanos. En ciudades como esta se percibe la necesidad de implementar una solución de transporte que se inclina hacia el uso de rieles para trenes metropolitanos para responder a la demanda propia del movimiento de pasajeros de la ciudad (Polanco, 2017).

En Colombia, como se evidencia en el libro Gestión de grandes proyectos urbanos en espacios metropolitanos (p. 36,37 y 40), Bogotá, en el lapso de un siglo, pasó de 100.000 habitantes a más de 7 millones, pero la urbanización acelerada no significa metropolización, idea que se caracterizó en el siglo XX por un modelo de

evolución del poblamiento y de diferenciación espacial, lo cual implicó la necesidad de idear megaproyectos que absorbieran la necesidad de las áreas metropolitanas y municipales, lo cual planteó la necesidad de ejecutar grandes cambios a nivel estatal, y de priorizar algunas necesidades, como el transporte. Bogotá se ha enfrentado al reto del transporte desde finales del siglo XIX, cuando el crecimiento de la ciudad exigió el uso de vehículos, desde el caballo, el tranvía y los rieles hasta el modelo hoy predominante; el sistema de buses BRT, que crece a la par del uso de vehículos privados, según una publicación oficial de Transmilenio (2013).

En 1967 funcionaban en Bogotá 2.679 autobuses urbanos, que transportaban en promedio 1.629.254 pasajeros por día. En esa época, cuando la ciudad alojaba a un poco más de un millón de habitantes y tenía 8.000 hectáreas de extensión, el servicio era relativamente razonable y cómodo. Pero a medida que la ciudad creció y llegó a alcanzar más de cinco millones de habitantes y una extensión superior a 30.000 hectáreas, no solo aumentó sustancialmente el parque automotor, hasta alcanzar a más de 20.000 vehículos, sino que se multiplicó la anarquía, la ineficiencia, la incomodidad y la contaminación, debido, entre otros factores, a que la innovación y la diversidad urbana de las ciudades atraen a más personas (Clark, 2002; Albouy, 2008 y 2010; Carlino y Saiz, 2008). La solución escogida por la administración de la ciudad fue el BRT Transmilenio, que ha provocado grandes cambios en la ciudad de Bogotá y afectado directamente el ordenamiento y el uso del suelo de la ciudad (Liliana, 2010).

8. MARCO TEÓRICO

8.1. El problema del transporte

En búsqueda de alternativas para abordar los retos de un sistema de transporte con las características propias del caso de estudio, se realizó una revisión bibliográfica en fuentes académicas sobre los conceptos más importantes y las metodologías más utilizadas orientadas hacia estrategias eficientes de operación de una red de transporte de pasajeros que funcionan por medio de buses o sistemas similares e Iniciativas que buscan incrementar la capacidad, la velocidad y la fiabilidad, orientadas a alcanzar una mejor calidad del servicio, evidenciada en factores como la reducción de tiempos de espera para los pasajeros (Liu, Liu, & Jia, 2019), tomando como punto de partida que la planeación del servicio de pasajeros consiste en la administración de líneas y frecuencias que implican las estipulaciones de horarios y la programación de vehículos y de tripulaciones, entre otros factores que responden a la demanda de pasajeros (Sâmara Visentini, Bassi, Borenstein, & Kummer Neto, 2017). Eso se relaciona con el uso de modelos para predecir de qué manera los viajeros escogen las rutas, desde el origen hasta su destino (Cepeda, Cominetti, & Florian, 2006).

Mediante bases de datos académicas, como SCOPUS, Science Direct, EBSCO Host y Web of science, entre otras, se analizaron las formas más comunes que se han implementado para afrontar ese reto, mediante distintos criterios, tales como el análisis del tránsito, la investigación de operaciones en el transporte y el análisis de sistemas de transporte, priorizando el estudio de infraestructuras con buses. Así se determinó que el tema del transporte aborda diversidad de campos, pero el tipo de planteamientos predominantes están relacionados con ciencias de la computación y la ingeniería, resaltando la presencia de la ciencia de la toma de decisiones y el modelado de sistemas.

La programación y el diseño de rutas influyen directamente sobre variables como la calidad del servicio, según el artículo "Transit network design and scheduling: a global review (2008)", en el que también se relaciona la planeación del tránsito en una secuencia de cinco pasos: 1) el diseño de rutas, 2) la configuración de frecuencias, 3) los horarios, 4) la programación de vehículos, y 5) la programación de tripulación. El primer paso corresponde a la planeación estratégica, que incluye el diseño de rutas, el segundo y tercer paso corresponden a la planeación táctica, y, por último, el cuarto y quinto se destinan a las actividades de la operación (Ng, Lee, Chan, & Lv, 2018). Las anteriores actividades de planeación, con sus respectivas entradas y salidas, pueden verse en la Tabla 2. Procesos de planeación del tránsito, que permite al estudio de diferentes problemas y relaciona técnicas de solución.

Tabla 2. *Procesos De Planeación Del Tránsito.*

Numero	Entradas independientes	Actividad	Salida
1	Datos de demanda	Diseño de la red	cambios en las rutas
	Datos de distribución Indicadores de desempeño de las rutas		Nuevas rutas Estrategias operativas
2	Subsidios disponibles	Configuraciones de las frecuencias	Frecuencias de servicio
	Buses disponibles Políticas de servicio Actual patrocinio		
3	Demanda por tiempo del día Tiempos para primeros y últimos viajes Tiempo de trayecto	Desarrollo de horarios	Tiempos de salida Tiempos de llegada
4	Tiempos de trayecto asociados al ingreso Tiempos de abordaje Restricciones de tiempo Estructura de costos	Programación de buses	Horarios de bus
5	Reglas de conductores o flota Costos de operación	Programación de conductores o flota	Horarios de conductores

Nota. Traducción de Inputs y outputs de los **procesos** para la planeación del tránsito, adaptado de: Review on meta-heuristics approaches for airside operation research. Ng, K. K. H., Lee, C. K. M., Chan, F. T. S., & Lv, Y. (2018). Disponible en: [doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.asoc.2018.02.013](https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.asoc.2018.02.013)

Para los pasos relacionados en la tabla anterior se deben tener en cuenta aspectos como:

- Posibles clientes frecuentes, relacionados con el área de cubrimiento del sistema.
- Matriz origen destino OD.
- La relación entre distintas variables objeto de cualquier iniciativa, ya que la orientación de cualquier estudio puede estar enfocada, por ejemplo, hacia la satisfacción del cliente o el costo de la operación, entre muchos otros enfoques.

8.2. Metodología general

Para efectos de análisis y optimización basados en trabajos anteriores se plantea la implementación de diferentes metodologías o algoritmos que tienen una función objetivo, sometida a restricciones, las que relacionan variables de tiempo de viaje, distancias y costos, con los propósitos más frecuentes relacionados con aumentar las ganancias del sistema, la clasificación, agrupación o clustering de comportamientos de la demanda y la planeación de paradas o administración operativa (Lyu et al., 2019), se relaciona el concepto de heurística definido como aproximaciones mediante algoritmos para proporcionar resultados óptimos, (Ng, Lee, Chan, & Lv, 2018).

Vinculando la aplicación de ciencias de toma de decisiones, se encontró que existe una variedad de alternativas que pueden clasificarse según el tipo de modelo utilizado, clasificación que se aprecia en la Tabla 3. Clasificación Modelos matemáticos.

Tabla 3. *Clasificación De Modelos Matemáticos*

N Relación	Tipos	Descripción
1 Tiempo	Estáticos	Permanecen invariables dentro de un periodo previamente establecido.
	Dinámicos	Presentan cambios dentro de tal periodo con el transcurrir del tiempo.
2 Naturaleza de las variables	Determinístico	Se tiene certeza de los valores de los parámetros.
	Estocástico	Los parámetros usados para caracterizar el modelo son variables aleatorias que tienen comportamientos estimados, pero no se conoce con certidumbre previamente cuál será el valor que tomen.
3 Grados de certidumbre	Continuos	Variables que pueden tomar valores continuos.
	Enteros	Variables que solo pueden tomar valores enteros.
4 Tipo de ecuaciones utilizadas	Lineales	Ecuaciones de un solo grado representados por una recta.
	No lineales	

Nota. Tabla extraída de Universidad Santo Tomas a distancia investigación de operaciones disponible en:
http://soda.ustadistancia.edu.co/enlinea/invoperaciones2/enteros__continuos.html.

Como dato complementario, según las bases de datos utilizada, las metodologías más usadas, según la clasificación de la tabla anterior sería recomienda la opción número 2, naturaleza de variables, tipo determinístico, en los cuales los parámetros del sistema y las perturbaciones de la línea de bus se consideran y satisfacen una distribución probabilística aleatoria dada (Li, Liu, Yang, & Gao, 2019).

8.2.1. Visiones metodológicas

Problemas como la asignación de personal y el tamaño de la flota han recibido amplio tratamiento, contando con algoritmos eficientes para la resolución de esta clase de planteamientos, mientras que los planteamientos relacionados con la optimización de rutas y las frecuencias de los buses poseen diferentes niveles de complejidad relacionados con la no linealidad, la no convexidad y los múltiples objetivos que dificultan su formulación (Ramírez, Hernández, & Varela Hugo, 2015). Según los principales autores, las mayores dificultades de ese el planteamiento de diseño y optimización de rutas en la literatura, son:

Tabla 4. Dificultades Diseño Y Optimización de rutas.

Numero	Actividad	Descripción
1	Formulación del problema	Variables de decisión. Función objetivo.
2	No linealidad y no convexidad del problema	Relacionado al comportamiento de los datos y los modelos.
3	Naturaleza combinatoria del problema	No permiten realizar una enumeración completa de las soluciones factibles. Uso de variables discretas en la descripción del recorrido.
4	Múltiples Objetivos	Existen diferentes soluciones no dominadas, que se obtienen como efecto de la renuncia o variación de cualidades que afectan el comportamiento de otras cualidades y por lo tanto del equilibrio del sistema, relacionado con los bienes de los usuarios y el de los operadores.
5	Configuración de rutas	Disposición especial.

Nota. Adaptación de Ramirez, S., Hernandez, L., & Varela Hugo. (2015). Diseño de rutas de transporte público.

8.2.2. Actividades de planeación estratégico.

La metodología principal para el diseño de una red consiste en la elección de un modelo que involucra distintas variables. A continuación, se presentan algunas propuestas de autores que plantearon distintas formas de tratar el problema.

Tabla 5. *Modelos Para El Diseño Estratégico*

Año	Nombre	Tipo Herramienta	Objetivo	Descripción y Características	Variables
2000	Van Nes & P.H.L. Bovy	Formulación analítica	Cuantificar el espacio optimo entre líneas y entre paradas	1. Discretización de un área en km^2 que cuente con presencia de autobuses. 2. Los costos son asociados a los tiempos de trayecto, espera y llegada a destino que de suponer cero transferencias será contante. 3. El costo del operador es proporcional a la densidad de vehículos. 4. Los ingresos del operador son la suma de la tarifa de pasajeros y la subvención multiplicada por el volumen de pasajeros que acceden a la línea. 5. Se plantean 6 funciones generales según los objetivos para el sistema.	1. Costos asumidos por cada uno de los actores. 2. Ingreso del operador. 3. Demanda por línea. 4. Transferencias. 5. Tarifa. 6. Flujo
1986	Ceder & Wilson	Modelo discreto de optimización	Minimizar el exceso de tiempo de viaje en el proceso de carga y de carga en las paradas.	Planteamiento a: - Restricción relativa máximo tiempo de viaje por cada Origen-Destino. Planteamiento b: Agrega el tiempo de espera de los pasajeros, costes de operación y de capital El algoritmo busca encontrar el grupo de rutas factibles, en el que cada terminal es tratada de manera separada.	Planteamiento a: 1. Tiempo viaje ruta directa. 2. Tiempos de transferencia. 3. Máximo tiempo OD expresado como Porcentaje del mínimo 4. Tiempo de una ruta directa Planteamiento b: 5. tiempos de espera pasajeros. 6. costos operación. 7. costos de capital. 8. Rutas factibles

1998	Ceder & israelí	Modelo discreto de optimización	Min. tiempos de transferencia y tamaño de flota. (Multiobjetivo).	Función multiobjetivo, incluye la asignación fija de la matriz de demanda y frecuencias mínimos para cada ruta. 1. Se enumeran el conjunto de rutas factibles (Ceder y Wilson (1986)) 2. Rutas directas entre OD. 3. Mínimo de rutas obtenido por procedimiento heurístico. 4. Asignación de flujo y frecuencias son calculadas según la hora punta. 5. Tamaño de flota Stern & Ceder (1983). 6. Nuevas rutas y eficiencias para la toma de decisiones.	1. Tiempo total de recorrido. 2. Tiempo de espera Phi y transferencia. 3. Tiempo en el que no hubo asientos disponibles 4. Pesos para variables. 5. Frecuencias Mínimas por rutas. 6. Rutas factibles C y W (1986). 7. Rutas directas. 8. Flujo. 9. Frecuencias
				Restricciones: 1. Frecuencia factible 2. Factor de carga. 3. Tamaño de la flota.	
1991	Baaj & Mahmassani	Modelo discreto de optimización	Función Objetivo: Min. tiempos de transferencia y tamaño de flota.	1. El costo total es la combinación del costo para el pasajero y para el operador. Donde en el primero se reduce el tiempo de viaje y para el segundo es la minimización de la cantidad de vehículos. 2. Caracterizado por la generalización de zonas en Nodos como centroides agrupadas de orígenes y destinos. 3. Ruta definida como conjunto de nodos. 4. Cada ruta tiene ida y retorno. 5. Distancia entre nodos considerada como el recorrido por el camino mínimo. Fundamento para los Algoritmos TRUST: Frecuencias y planificación táctica. RGA: Planificación estratégica. Restricciones 1. Frecuencia factible 2. Factor de carga. 3. Tamaño de la flota.	1. Cantidad de nodos 2. Demandas entre nodos. 2. Tiempo de ciclo en la ruta. 3. Tiempo de viaje total 4. Tiempo de espera entre paradas. 5. Frecuencia de los auto buses. 6. Pesos relativos componentes de cote. 7. Tiempo de viaje entre nodos. 8. Tiempo de transferencia. 9. Frecuencia mínima. 10. Cantidad de autobuses operando. 11. Tamaño flota disponible. 12. Factor de carga. 11. Máximo flujo por arco. 12. Capacidad de pasajeros sentados. 11. Conjunto de rutas para una solución. 12. Factores de conversión y pesos relativos de la función objetivo.

2000	Nagamchai & Lovell	Modelo discreto de optimización	Minimizar los tiempos de transferencia y el tamaño de la flota (detallado)	Criterio multiobjetivo. Permite calcular frecuencia de rutas. Restricciones: Factor de carga	1. Costo de la flota. 2. Costo viaje pasajero. 3. Costo espera usuario. 4. Cantidad de Nodos. 5. Cantidad de rutas. 6. Costo por hora de operación autobús. 7. Largo de ruta. 8. Variable binaria de uso de arco. 9. Coeficientes valor subjetivo tiempos de viajes vs tiempo de espera. 10. Espaciamiento temporal 11. Velocidad de buses.
2002	Gruttner, Pinninghoff, Teudela y Díaz	Modelo discreto de optimización	Maximizar el beneficio del operador y minimizar el costo del usuario por línea. Restricciones: Distribución de acceso y egreso (origen y destino)	Alternativa basada en los métodos Logit es decir realizar una regresión para encontrar el resultado de la variable. 1. Cálculo de utilidades o línea. 2. No tiene en cuenta la frecuencia de los buses. 3. No tiene en cuenta el tamaño de la flota.	1. Coeficientes de conversión. 2. Valores Subjetivos del tiempo. 3. Rutas válidas. 4. Afluencia total de viajes. 5. Tarifa por línea. 6. Costo unitario por operación por (km). 7. Tiempo de viaje. 8. Tiempo de acceso a línea. 9. Tiempo de espera. 10. Número de viajes entre OD. 10. Pesos relativos de los tiempos de acceso y espera

Nota. Adaptado de: Callejo, M; Estrada M. (2009). Optimización del diseño de líneas de autobús aplicación a Donostia San Sebastián.

8.2.3. Actividades de planeación tácticas y operativas.

Existen distintos lineamientos de trabajo que plantean los retos de configuración de frecuencias y estipulación de horarios, tal como lo muestra la Tabla que aparece a continuación y que se relaciona con el trabajo de destacados de autores.

Tabla 6. *Modelos Para El Estudio Táctico Y Operacional*

	Autores	Línea de trabajo
1	Furth and Willson (1981) Gallo and Miele (2001) Avila Torres et al. (2017)	Optimización simultanea de frecuencias y horarios
2	Sun et al. (2015) Wang et al. (2015) Yu al. (2017)	Optimización separada de frecuencias y horarios Resolución Secuencial.
3	Ceder et al. (2001) Cevallos y Zhao (2006) Wei y Sun (2017) Gkiotsalitis and Maleskar (2018)	Sincronización de horarios para minimizar los costos asociados a los tiempos de espera de los pasajeros.
4	Kang et al. (2019)	Pasajeros varados en último servicio
5	Eranki (2004) Ibarra Rojas y Rios Solis (2012)	Horarios flexibles y márgenes de tiempo en la sincronización de horarios.
6	Adamski y Turnau (1998) ZHao et al. (2006)	Periodo eficiente de baja actividad.
7	Ceder et al. (2013) Ceder (2011)	Nivel de carga uniforme mediante control de tiempos de despacho.
8	Daduna y VoB (1995) Shafahi y Khani (2010)	Sincronización de transferencias y control de despacho.
9	Wu et al. (2015)	Utilización de tiempos de demoras determinísticas.
10	Hickman (2001)	Tiempos estocásticos de viaje.
11	Hans et al (2015)	Distribuciones de probabilidad en los tiempos de viaje entre estaciones.

Nota. Adaptado de Gkiotsalitis, K., & Alesiani, F. (2019). Robust timetable optimization for bus lines subject to resource and regulatory constraints

8.3. Inteligencia artificial

Las primeras técnicas utilizadas fueron: Knowledge-based system (KBS), que consiste en la postulación de reglas basadas en el conocimiento de los seres humanos, y Artificial Neural Network Systems (ANN's), que construye una red de conexiones neuronales en varias capas, proceso que continuó con esfuerzos por disminuir el error de predicción mediante el entrenamiento de ANNs.

La aplicación del método gradiente descendente sentó las bases para el desarrollo del machine learning como una subcategoría de la inteligencia artificial que permite a los computadores abordar complejos problemas (Abduljabbar, Dia, Liyanage, & Bagloee, 2019), los cuales buscan:

1. Toma de decisiones, planeación y administración en respuesta a un crecimiento continuo de la demanda con la oferta vial limitada, lo cual mejora las herramientas de pronóstico del volumen de tráfico, las condiciones de tráfico y los incidentes.
2. Mejora y desarrollo de modos de transporte público sustentable.
3. Conexión de vehículos autónomos con el objetivo de buscar mayores eficiencias y reducir la accidentalidad.

8.3.1. Machine learning

El machine learning o aprendizaje automático es el campo de la ciencia de la computación que trata de la simulación de comportamiento inteligente en las computadoras (Bagheri, Akbari, & Mirbagheri, 2019), que suele subdividirse en varias ramas:

1. Métodos de modelado, como redes neuronales artificiales, Fuzzy Logic o lógica difusa, programación genética, modelos de árbol y modelos inteligentes híbridos.
2. Métodos de optimización, en donde se usan también algoritmos genéticos (GA), y Enjambre de partículas, conocido como Particle Swarm Optimization (PSO).
3. Métodos de clustering.
4. Otras metodologías, como la detección de imágenes y las selecciones variables.

A continuación, se relacionan en la Tabla 7 los modelos generales de inteligencia artificial que convergen hacia el machine learning, Las técnicas anteriores son complementadas por técnicas supervisadas y no supervisadas.

Tabla 7. *Fundamentos IA*

Nombre	Descripción
(KBS) Knowledge-based system	Conocimiento presentado por humanos.
(ANNs) Neural network systems	Usado para la identificación de imágenes.
(CNN) Convolutional Neural Network	Aplicadas para el reconocimiento de escritura y la corrección de textos, usualmente conocidas como deep learning debido a la gran cantidad de capas ocultas.
(RNN) Recurrent neural network.	

Nota. Adaptado de Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. Sustainability (Switzerland), 11(1) doi: 10.3390/su11010189

Tabla 8. *Otras Técnicas IA.*

Categoría	Nombre
Técnicas supervisadas	(SVM) Support vector machine (PNN) Probabilística Neural Network (RBN) Radial Basis Network
Técnicas no supervisadas	K-Nearest Neighbors and Decision Tree. Greedy layer-wise Clúster análisis

Nota. Adaptado de Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. Sustainability (Switzerland), 11(1) doi: 10.3390/su11010189

Para la implementación de las anteriores técnicas se han desarrollado algoritmos que permiten la resolución de planteamientos de pronóstico y clasificación. Algunos de los más populares se presentan a continuación:

Tabla 9. *Algoritmos genéticos*

Nombre	Descripción
(SA) Simulated Annealing	Basado en el proceso de recocido de un metal.
(ACO) Ant Colony Optimiser	Basado en el comportamiento de las hormigas (parte de swarm intelligence systems).
(AIS) An artificial Immune system	Basado en el sistema inmune humano.
(BCO) Bee Colony Optimization	Basado en el comportamiento de las abejas (parte de swarm intelligence systems).
(FLM) Fuzzy Logic Model	Usado para la optimización de caminos cortos comparado con Logistic Regression Model (LRM).
(ABSE) Agent-Based Software Engineering	Usado con un enfoque dinámico para identificar la trayectoria más corta mediante la información de multicriterios y multiescenarios.

Nota. Adaptado de Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. Sustainability (Switzerland), 11(1) doi:10.3390/su11010189

8.3.2. Deep learning

El Deep learning es un derivado del Machine learning, que en el marco de las relaciones entre vehículos inteligentes y sistemas de transporte inteligentes y las redes de aprendizaje profundo o Deep Neural Networks, han permitido una mayor flexibilidad en cuanto el análisis de las relaciones no lineales entre las variables predictivas y las salidas de un sistema, que funciona en gran parte gracias a una mayor capacidad de procesamiento computacional, la disponibilidad de mayor información que permite combatir los sobreajustes y el desarrollo de algoritmos de optimización que funcionan como una aproximación estocástica de una variable objetivo (Kraus, Feuerriegel, & Oztekin, 2019).

Esta categoría ofrece otra clasificación, como técnicas paramétricas y técnicas no paramétricas, donde el primer grupo ofrece bajos errores durante la predicción, pero en redes con ruido y perturbaciones pueden resultar en aproximaciones con baja exactitud.

Por otro lado, las aproximaciones no paramétricas, que pueden considerarse como Deep learning, han tomado gran importancia para los investigadores en la aplicación de la clasificación de datos, procesamiento del lenguaje natural (NLP), o natural language processing, y el reconocimiento de objetos según su desempeño.

Tabla 10. *Técnicas para el análisis del tráfico*

Clasificación	Sigla	Nombre
Paramétricas		Series de tiempo
	KF	Kalman Filtering
No paramétricas	(CNN)	Convolution Neural Network
	(DBN)	Deep Belief Network
	(RNN)	Recurrent Neural Network
	(AE)	Autoencoder

Nota. Tomado de Salvador, M. M., Budka, M., & Quay, T. (2018). Automatic transport network matching using deep learning doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trpro.2018.09.053

Bajo este enfoque, los algoritmos también pueden ser clasificados como uní variables y multivariantes; el enfoque univariante observa las variables de tráfico de un solo sitio, mientras que el enfoque multivariante tiene múltiples sitios de entrada y salida (Liu, Liu, & Jia, 2019).

8.3.3. Regresión estadística y fundamentos de redes neuronales

Para describir el comportamiento de un conjunto de datos, la mayoría de las técnicas utiliza el concepto de regresión, donde se busca una función que describa mejor las relaciones existentes:

“En un análisis de regresión simple existe una variable de respuesta o dependiente (y) y una variable explicativa o independiente (x). El propósito es obtener una función sencilla de la variable explicativa, que sea capaz de describir lo más ajustadamente posible la variación de la variable dependiente. La variable explicativa puede estar formada por un vector de una sola característica o puede ser un conjunto de n características, atributos o dimensiones (regresión múltiple)” (Pereira González Augusto, 2010).

Expresado en una forma matemática una regresión lineal sigue la estructura:

$$\text{Ecuación 1} \quad \hat{y}_i = X_i w = w_0$$

Concepto que puede extenderse para una regresión múltiple, en el cual existe más de una variable independiente que describe la variable dependiente:

$$\text{Ecuación 2 } \hat{y}_i = X_i w = w_0 + w_1 x_{i,1} + w_2 x_{i,2} + \dots + w_n x_{i,n}$$

El anterior principio es el fundamento para la construcción de una red neuronal, como los modelos predictivos, que utilizan las características de los datos de entrada ($X \in \mathbb{R}$) para predecir una variable $y \in Y$, donde el valor de la variable dependiente de una función $y = F(x, w)$, que incluye un parámetro adicional w , donde dichos parámetros son aproximados mediante una técnica de optimización que se relaciona al error entre valores observados y valores predichos.

8.3.4. Función de activación

La función de activación es conocida por realizar una combinación lineal o no lineal, que facilita la identificación de patrones en las variables (Kraus, Feuerriegel, & Oztekin, 2019), a continuación, se relacionan las tres funciones más populares.

La función sigmoide transforma los datos en valores que se encuentran entre 0 y 1:

$$\text{Ecuación 3 } \textit{Función Sigmoide} = \frac{1}{1+e^{-x}}$$

La función tangente hiperbólica transforma los datos en valores entre -1 y 1; también existen aplicaciones con otras funciones trigonométricas, pero son menos utilizadas, tal como la siguiente ecuación:

$$\text{Ecuación 4 } \textit{Tangente Hiperbolica} = \textit{Tanh} = \frac{(e^{2x}-1)}{(e^{2x}+1)}$$

La función conocida como regresión logística genera valores positivos:

$$\text{Ecuación 5 } \textit{Relu} = \sigma(x) = \max(0, x) \in [0, \infty)$$

Otra función es destacada por su aplicación en la clasificación de clases, que recibe como entrada una capa de agrupamiento con promedios globales y genera una distribución de probabilidades (Zhou et al., 2019):

$$\text{Ecuación 6 } \textit{Softmax} = \frac{e^{(W_{Lc} * h_L + b_{Lc})}}{\sum_{i=1}^n e^{(W_{Li} * h_L + b_{Li})}}$$

8.4. Enfoque de ruteo.

La técnica VRP (Vehicle route problem) consiste, en forma general, en identificar un set de rutas para vehículos, que se relacionan con distintas variables que explican una red de clientes o puntos de consumo. Esta metodología tiene diferentes variaciones, relacionadas en la Tabla adjunta.

Tabla 11. *Clasificación de enfoques VRP*

Nombre	Descripción
Multidepósito	Minimizar la flota de vehículos, la suma total del tiempo de transporte y la demanda total de artículos que deben ser entregados desde varios puntos.
Periódico	Minimizar la flota de vehículos y la suma total del tiempo de transporte necesario para suplir a todos los clientes. La planeación se hace para determinado periodo.
Capacitado	Minimizar la flota de vehículos, la suma total del tiempo de transporte y la demanda total de artículos para una flota mixta de vehículos de entrega con una capacidad uniforme que debe atender una demanda de clientes, conocida para un solo artículo, con un mismo depósito y un costo de transporte mínimo.
Estocástico	Minimizar la flota de vehículos y la suma total del tiempo de transporte necesario para atender a todos los clientes en cada recorrido. Se consideran aleatorios sus demandas y el tiempo de servicio y/o transporte.
Backhauls (Retornos fuera de ruta)	El objetivo es encontrar un conjunto de rutas que minimice la distancia total transportada. Los clientes pueden demandar o regresar algunos artículos. Se debe tener en cuenta que el vehículo no sobrepase su capacidad.
Entrega y reparto	Minimizar la flota de vehículos y la suma total del tiempo de transporte, con la restricción de que cada vehículo debe tener la capacidad suficiente para transportar artículos que vayan a ser repartidos y uno de ellos recogido en los clientes para regresarlos al depósito.
Ventanas de tiempo	Minimizar la flota de vehículos, la suma total del tiempo de transporte y el tiempo de espera necesarios para atender a los clientes en una hora determinada.

Nota. Adaptación de tipos de enfoques. VRP Bermeo Muñoz, E., & Calderón Sotero, J. (2009). Diseño de un modelo de optimización de rutas de transporte. El hombre y la máquina (32), 52-67.

Tabla 12. *Fundamentos VRP*

Nombre		Descripción	Características
TSP	Agente viajero /Travelling Salesman Problem	Agente vendedor que debe visitar cierta cantidad de ciudades en un solo viaje.	1 origen Origen = Destino
TPS		Variación del modelo TSP con mayores variables y restricciones.	*1 Origen *Origen = Destino *Mayor Cantidad de Variables y *restricciones
m-TSP		Variación de TSP, que consiste en m agentes o vehículos en el que cada cliente debe ser visitado una vez.	*1 Origen * Origen = Destino + m agentes * m rutas que deben visitar p clientes.
PTSP	TSP probabilístico	El objetivo es encontrar el mínimo costo de recorrido esperado, atreves de un conjunto de nodos con probabilidades asociadas la presencia o no de consumidores.	* Origen = Destino * m agentes * m rutas que deben visitar p clientes. * Demanda Determinística
VRP	Vehicle Routing Problem	Problema de ruteo de vehículos donde la capacidad de la flota se convierte en restrictiva para la formulación.	* Restricción del tamaño de la flota
CVRP	Capacited VRP	Minimiza el costo total para surtir a todos los consumidores. Mínimo costo definido como la suma de los costos de los arcos pertenecientes a los recorridos, , cada centro de consumo es visitado por exactamente un vehículo y la suma de las demandas de los centros de consumo visitados no excede la capacidad del vehículo	* 1 centro de Distribución que responde a * Una Demanda Determinística. * Cada recorrido visita el centro de distribución. * Cada vehículo visita un cliente. * Una ruta no puede superar el vehículo

Nota. Fundamento VRP adaptado de Linda Rocha, González Elsa, & Orjuela Javier. Una revisión al estado del arte del problema de ruteo de vehículos: Ev. Evolución Histórica Y Métodos De Solución. En: Ingeniería, 16(2), 35-55.

Se realiza a continuación una ampliación de la metodología VRP por presentar condiciones favorables para ser utilizada como herramienta de análisis de sistema.

Relacionando dos grandes grupos de variaciones de la metodología que se diferencia según las restricciones que tienen en cuenta. El primer grupo en la Tabla 6, caracterizado por las premisas de homogeneidad en uno o varias variables analizadas en el planeamiento matemático.

Tabla 13. *Variantes VRP CVRP homogéneo*

Nombre		Descripción	Características
DVRP	<i>Distance VRP</i>	Distancia homogénea	Restricción de máxima longitud Longitud de arcos no puede exceder la máxima Longitud de ruta
DCVRP	<i>Distance and capacite VRP</i>	Distancia y capacidad homogénea	Restricción de DVRP y restricción de capacidad
VRPB	<i>VRP with backhauls</i>	Con retornos	Los clientes pueden demandar y devolver mercancía El tamaño de la mercancía es uniforme. Los centros de despacho deben ser atendidos antes que los centros de consumo
VRPBTW	<i>VRP with backhauls Time Windows</i>	Con retornos y ventanas de tiempo	Combinación del concepto de ventana de tiempo y devoluciones
VRPTD	<i>VRP Dead lines</i>	Plazo final	Tiempo que debe cumplirse para atender a un cliente o consumidor
VRPTW	<i>VRP Time Windows</i>	Ventanas rígidas de tiempo	Time window. Tiempo de recorrido más tiempo de atención para atender un cliente.
VRPMTW	<i>VRP Multi Time Windows</i>	Ventanas de tiempo múltiples	Cada cliente tiene una ventana de tiempo (Maaiké Hoogeboom, Dullaert Wout, Lai David, & Vigo Daniele, 2018)
VRPSTW	<i>VRP Soft Time Windows</i>	Tiempo flexible de entrega	Las entregas pueden ser fuera de ventanas de tiempo (Bouchra, Btissam, & Mohammad, 2018)
SDVRP	<i>Split Delivery VRP</i>	Entregas fraccionadas	EL cliente puede ser atendido por distintos vehículos

Nota. Variantes VRP y CVRP Homogéneo adaptado de Linda Rocha, González Elsa, & Orjuela Javier. Una revisión al estado del arte del problema de ruteo de vehículos: Ev. Evolución Histórica Y Métodos De Solución. . En: Ingeniería, 16(2), 35-55.

El segundo grupo diferenciado por la complejidad que se agrega al estudio de las variables alisadas en los modelos.

Tabla 14. *Variantes VRP CVRP Heterogéneo*

Nombre	Descripción	Características
--------	-------------	-----------------

VRPHF		En él los costos y capacidades de los vehículos varían	*Cantidad de vehículos ilimitada * Se decide sobre las rutas y la composición de la flota de vehículos a utilizar.
PVRP	<i>VRP Periodic</i>		El modelo se extiende a un período de planeación de M días.
Multi-Trip VRP	<i>VRP Multiviaje</i>	Varias rutas en el mismo periodo de planeación	Los horarios diarios de los conductores deban estar diseñados para una flota de vehículos fija.
MCVRP	VRP Multi Compartment	el CVRP se puede considerar como VRP con capacidad 1 y el MCVRP se considera como capacidad mayor a 1	Consiste en transportar más de una cantidad de objetos a la vez
MOVRP	Multi Objective VRP	Diferentes objetivos que buscan el análisis de diferentes variables	Ruta (costo, beneficio, etc.), nodos y arcos (ventanas de tiempo, satisfacción del cliente, etc.) y recursos (mantenimiento de flota de vehículos, especificaciones de producto, etc.)
SVRP	Stochastic VRP	Uno o varios componentes de la formulación son aleatorios.	Problema de Ruteo de Vehículos estocástico- puede tener variantes como: *VRPUD, cuando la demanda de cada consumidor es una variable aleatoria. *VRPSTT cuando los tiempos de servicio y tiempos de viaje son variables aleatorias * SVRP-SN, en el cual cada consumidor tiene una probabilidad p de presentarse

Nota. Variantes VRP y CVRP Homogéneo adaptado de Linda Rocha, González Elsa, & Orjuela Javier. Una revisión al estado del arte del problema de ruteo de vehículos: Ev. Evolución Histórica Y Métodos De Solución. En: Ingeniería, 16(2), 35-55.

8.5. Otros enfoques

Programación: Según Sun, Jin, Lee, Axhausen, & Erath, (2014), desde otro punto de vista el estudio de un sistema de transporte puede ser estudiado como:

- El diseño de la programación de la red “Network timetable design problem” es una metodología que suele buscar como objetivo minimizar el costo de transferencia a través de la coordinación y sincronización del tránsito entre diferentes rutas del problema estudiado en redes de buses y metros.

- Diseño de horarios de una sola vía, o “Single track timetable design problem.”, que difiere del anterior en el sentido que busca el ajuste de la frecuencia del servicio incluyendo las variables de beneficio por pasajero y ahorro de tiempos.

Satisfacción del pasajero: La cuantificación de la satisfacción de un pasajero es compleja, debido a que existen muchos atributos que pueden ser percibidos positiva o negativamente por un individuo. A pesar de ello, puede ser expresada matemáticamente como una función de algunas variables explicativas, como los tiempos de espera, el tiempo de abordaje, las sillas disponibles, el tiempo pasado dentro del vehículo, el tiempo total del trayecto y el tiempo del transbordo, entre otras (Shang, Huang, & Wu, 2019).

Paradas de la red: Uno de los factores que contribuyen al control coordinado es la ubicación entre las paradas de un sistema de transporte, lo cual permite una disminución de los retrasos de la red (Liu, Li, & Ni, 2017).

Estrategias: Los métodos de optimización tienen como objetivo mejorar el flujo de los vehículos en la red de vías, lo cual reviste dos tipos de estrategias: las comportamentales, que consisten en la configuración de luces, señales, etc. La estrategia que ejemplifica este tipo puede ser la sincronización de una señal, como las luces de tránsito, como alternativa para incrementar la velocidad y la capacidad de un sistema (Nesmachnow et al., 2019). Por otro lado, un segundo grupo de medidas está relacionado con la infraestructura, que consiste en agregar nuevos carriles o ampliar las carreteras, que tienen un costo mayor.

Comportamiento con otros vehículos: Políticas como la priorización del transporte público en las ciudades y la estipulación de reglas para los conductores pueden mejorar el comportamiento del flujo de vehículos, vistos sus efectos en la gestión de cambios de carril y espacios entre los vehículos (Cheng, Huang, & Zhang, 2019).

Contexto multimodal: la complejidad que tiene el estudio del comportamiento del pasajero incluye la decisión a escoger entre distintos modos de transporte, que pueden hacer parte de una misma red (Bingfeng, Ming, Xiaobao, & Ziyu, 2017).

Condiciones externas: estudios de la correlación entre las condiciones extremas del ambiente; por ejemplo, referentes a las temperaturas muy bajas y las muy altas (Miao, Welch, & Sriraj, 2019).

8.6. Análisis de datos atípicos

Uno de los temas más importantes a tratar en esta investigación son los datos atípicos, también conocidos como Outliers, definidos como datos extremos que se

salen de los patrones generales de la distribución de las variables. La importancia de tratar correctamente estos datos radica en que los valores perdidos y los valores atípicos pueden cambiar drásticamente los resultados de un análisis (Sang & Jong, 2017).

Otras definiciones pueden ser abarcadas bajo el término anomalía, definida como una observación o grupo de observaciones con notables inconsistencias respecto del grupo; otras definiciones vinculan el concepto con patrones en los datos que no se ajustan a un comportamiento normal definido, por lo cual es claro que el concepto de normalidad reviste uno de los principales pasos para construir una solución para la detección de anomalías (Fernandes, Rodrigues, Joel J. P. C., Carvalho, Al-Muhtadi, & Proença, 2019), Las anomalías pueden ser clasificadas en dos grupos, desde un enfoque de ciencias de la computación:

Tabla 15. *Tipos de anomalías*

Clasificación	Tipo de variable	Descripción
	Puntuales	Desviación individual de un dato.
	Colectivas	Cuando una colección de datos similares presenta un comportamiento anormal con respecto al resto de datos.
Basadas en la naturaleza:	Contextuales	Para su consideración, del contexto dependen: dos sets de atributos deben definir el contexto, uno relacionado con los datos y otro de relacionado con el comportamiento; por ejemplo, cuando se relaciona un conjunto de medidas y, por otro lado, condiciones horarias que afecten el valor de las medidas, como ocurre en el caso de los picos horarios.

Basadas en aspectos causales.		Los conocidos como de configuración errónea son problemas no maliciosos, que pueden ocurrir en un sistema de red, en la mayoría de los casos, por: Fallas de hardware, errores de software Errores humanos Fallas del servidor Cortes de energía Errores de configuración Congestión del tráfico Configuración inadecuada de recursos Cambios en el comportamiento de la red causados por la imposición de límites de velocidad Implementación de nuevos equipos
	Operacional	
	Flash crowd	Relacionado con el rápido crecimiento del flujo de tráfico en un sistema. Puede estar vinculado con promociones o eventos que afecten la demanda.
	Medidas:	Relacionadas con la metodología de recolección de datos.
	Abuso de una red o ataques malintencionados	Conjunto de acciones maliciosas que buscan interrumpir, negar, degradar o destruir información o el servicio general.

Nota. Adaptación y clasificación de tipos de anomalías: Fernandes, G. Rodrigues, J. J. P. C., Carvalho, L. F., Al-Muhtadi, J. F., & Proença, M. L. (2019). A comprehensive survey on network anomaly detection. *Telecommunication Systems*, 70(3), 447–489. <https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1007/s11235-018-0475-8>

Un factor importante para tener en cuenta en esta revisión es que los procesos de recolección de datos se caracterizan por la incertidumbre espaciotemporal, que depende de: los dispositivos e instrumentos usados, la distribución geográfica de las medidas, las cuales pueden ser irregulares o heterogéneas, y los tiempos entre medidas (Haosheng Huang , Yi Chengb , Robert Weibela), otras variables que pueden generar datos atípicos o datos perdidos que son manejados de diferentes formas.

Ubicando con un parámetro de búsqueda la palabra outlier, mediante la herramienta de análisis de información Scopus Dashboard se obtuvo un análisis que presenta la presencia de este tema en distintos campos científicos, especialmente relacionados con el tratamiento de datos encontrando en 113.505 documentos la presencia de la palabra de búsqueda, donde la producción de documentos es liderada por las ciencias de la computación con un 19.1%, ingeniería 12.6% y otros artículos relacionados con diferentes temáticas.

De esta forma, en la revisión se priorizaron las aplicaciones relacionadas con ingeniería y ciencia de la computación, con el criterio de una mayor aplicación y variedad de posibles metodologías, resaltando que el crecimiento de los documentos producidos con un enfoque similar al de este trabajo se inicia cerca del año 2.000 y crece cada vez más, hasta alcanzar, para el año 2018, 12.500 documentos producidos anualmente, de los cuales cerca de 3.500 aproximadamente están catalogados como ingeniería y 5.000 como ciencias de la computación. Con base en este gran volumen de información, se decide definir cuáles son las tendencias y técnicas más utilizadas.

Desde un punto de vista general, se encontró que este tipo de enfoque se relaciona directamente con la siguiente afirmación: "La extracción de información útil en el procesamiento de datos y su transformación en conocimiento efectivo para el soporte de decisiones es el objetivo de la minería de datos" (Quiñones Marcos, 2019).

Específicamente, las formas más comunes utilizadas para detectar los datos atípicos consisten en pruebas que asumen la normalidad de los datos y miden la distancia entre un dato y el centro de los datos basados en diferentes medidas. Por ejemplo, teniendo en cuenta la media y la desviación estándar, método que no siempre ofrece resultados apropiados debido a que este criterio es sensible a la presencia de datos atípicos, por lo cual se han desarrollado diferentes criterios como la utilización de la mediana y los rangos de cuartiles en conjunto con métodos gráficos como el Boxplot, o diagrama de caja y bigotes (Sang & Jong, 2017), alternativas sencillas que buscan mejorar la veracidad de los análisis y son la base para otras metodologías que pretenden abordar este tipo de reto.

El machine learning también es aplicado para el análisis de los datos atípicos, como los siguientes:

Aprendizaje no supervisado: utilizado para la estimación de grupos similares de observaciones o clustering. Algunas de las técnicas más usadas son: K-means, Fuzzy C. means, Mixture Models (MM), Window Based y otros, como Heuristic Smoothing Method, Hierarchical Approach, Component Wise, Hidden markov model.

Aprendizaje supervisado: caso de entranmiento con profesor y que utiliza información global, la salida comuada por la red se compara con la salida deseada, y los mediante pesos la red se modifica en el sentido de reducir el error. (Quesada, Graciani, Bonal, & Díaz-Mata, 1994)

Aprendizaje reforzado: Relacionado con la minimización o maximización de una función, para reducir los impactos de los datos atípicos se emplean tres técnicas:

1. Recorte: reemplazarlos por la media.
2. Winsorization: reducir los pesos de los datos atípicos.
3. Métodos de estimación robustos, que consisten en una variedad de metodologías estadísticas.

RNN: La idea general surgida detrás de este modelo consiste en una transformación que se realiza de forma repetitiva en cada nuevo dato, mediante un bucle, y que a su vez almacena un estado oculto que se pasa como entrada adicional a la siguiente iteración, también considerado como desplegar la información en el tiempo (Automatic Transport Networ Matching Using Deep learning)

Pronostico

El pronosticar se utiliza como insumo en todos los tipos control empresarial, dentro y fuera de la función de operaciones. Existen dos tipos de métodos cuantitativos para pronóstico: series de tiempo y causales (Loreto, Jimenez, & Rey, 2011).. En general, los métodos cuantitativos hacen uso de un modelo subyacente para llegar a un pronóstico. La lógica básica de todos los métodos cuantitativos para ser pronósticos es que mediante las series datos del tiempo pasado o y los modelos patrones causales de para datos se puede hacer indicadores confiables para predecir el futuro

Pronósticos por Series de Tiempo.

Los métodos por series de tiempo se utilizan para análisis detallados de los patrones de la demanda en el pasado, a lo largo del tiempo y para proyectar estos patrones a futuro. Una de las suposiciones básicas de todos los métodos por series tiempo es que la tendencia, estacionalidad, ciclo y en error. pueden interpretarse como s componentes.

Los componentes asociados a una serie de tiempo son; la tendencia, la estacionalidad la ciclicidad y la aleatoriedad.

Promedio móvil

El método más simple para el pronóstico por series de tiempo es el método del promedio móvil. En este método patrones de estacionalidad, se supone que la serie de tiempo tiene sólo un componente de nivel y un componente aleatorio. No se presupone la presencia de tendencias demanda.

Suavización exponencial:

Se basa en la idea., muy simple, de que es posible calcular un promedio nuevo a partir de demanda más recientemente observada. El pronóstico de suavización exponencial simple, el cual es recomendable para comportamientos de demanda

aleatorios o nivelados, donde se procura excluir el impacto de los elementos irregulares históricos, mediante un enfoque en períodos de demanda reciente, este presenta una ventaja sobre el modelo de promedio móvil ponderado, porque no requiere de una gran cantidad de períodos y de ponderaciones para lograr óptimos resultados; y un pronóstico de promedio simple (método estándar sencillo), ideal para patrones de demanda aleatorios o nivelados sin elementos estacionales o de tendencia.

Proceso analítico jerárquico, AHP o “Analytic Hierachy Process”.

AHP es una herramienta fiable para facilitar los procesos de toma de decisiones sistemáticas y lógicas, y determinar la importancia de una serie de criterios y sub criterios. Se aplica extensamente a los campos de la construcción, tales como la asignación de recursos, el diseño del proyecto, la planificación para el desarrollo urbano, gestión de mantenimiento, la evaluación de políticas, etc. (Lee & Chan, 2008). se compone de ocho pasos principales:

- Identificar el problema: La decisión problema se descompone en las partes más importantes, que resulte en los escenarios o alternativas a ser seleccionadas.
- Determinar que el problema puede ser resuelto por AHP
- Estructurar el problema: Un enfoque de nivel superior se descompone en criterios y a su vez en un segundo nivel y de ser necesario en sub criterios en un tercer nivel.
- Determinar las escalas de los AHP, -se solicita la opinión de expertos.
- Determinar las prioridades relativas sobre los diferentes criterios en cada nivel por medio de una comparación por pares.
- Calificar las alternativas en función de los criterios.
- Coeficiente de consistencia: Se evalúa la calificación de la importancia relativa de los criterios. AHP no exige consistencia perfecta, pero un juicio sólo se considera aceptable cuando la consistencia es de 0,10.

9. MARCO METODOLÓGICO

9.1. Tipo de estudio:

El estudio corresponde a una combinación entre un estudio exploratorio y un estudio descriptivo, ya que empieza con la recolección bibliográfica, concierne al estado actual del conocimiento sobre la descripción orientada hacia de optimización de sistemas de transporte

10. Método de investigación

La investigación se realiza mediante tres grandes fases, siguiendo el objetivo principal. En primer lugar se realiza una limpieza y un análisis de la información recolectada en una base de datos que recopila los concernientes registros de la demanda del sistema durante el año 2017; en segundo lugar, estos datos se clasifican mediante una estadística descriptiva, pruebas de homogeneidad e independencia, lo cual permite obtener las piezas fundamentales para la propuesta de un modelo de comportamiento normal o típico; en tercer lugar, se desarrolla un análisis del modelo que busca describir de una forma general la dinámica de la operación. Para lo anterior se describen las etapas que conforman este proceso.

10.1. Etapa 1. Análisis de metodologías:

Mediante el proceso analítico jerárquico se desarrolla un análisis de las metodologías disponibles en la literatura relacionadas al análisis táctico y operacional de un sistema de transporte, que responden a planteamientos estratégicos y estructurales.

10.2. Etapa 2. Datos de entrada y planteamiento:

Consiste en la organización de los datos de una forma estándar que permite el tratamiento de una forma sencilla. En este paso se incluye la elección de variables que describen mejor el problema a tratar y cumplen con el alcance del proyecto.

10.3. Etapa 3. Análisis exploratorio y comportamiento de la demanda

Con ayuda de las herramientas de Microsoft Excel y el software IBM SPSS Statics 25, se realiza una estadística descriptiva, medidas de tendencia central y pruebas de homogeneidad e independencia sobre el conjunto de registros, con el objetivo de determinar si los datos se comportan de alguna forma en particular e identificar posibles relaciones que permitan comprender integralmente su comportamiento y disminuir la complejidad.

Basado en el análisis de los datos y teniendo en cuenta los pasos anteriores, en compañía de fuentes externas, se realizará un aprendizaje semi-supervisado mediante la identificación de tendencias que pueden corresponder a eventos

exógenos al sistema y que se relacionan con eventos pasados no típicos que hayan influido en el comportamiento normal del sistema.

10.4. Etapa 4. Propuesta cuantitativa de respuesta a la demanda

Teniendo en cuenta los resultados anteriores, se propone un grupo de variables que suelen utilizarse en las metodologías estudiadas, con el fin de realizar un modelo, el cuál usa la metodología elegida o la combinación de metodologías para resolver el número de vehículos necesario para la operación y otras variables relevantes para la evaluación del rendimiento.

10.5. Etapa 5. Prototipo y validación

Fundamentado en la naturaleza multiobjetivo del problema, el modelo describe diferentes escenarios de la operación basado en el costo del operador, y el tiempo de viaje, esto permite al usuario del aplicativo analizar el rendimiento de la troncal.

11. DESARROLLO

11.1. Etapa 1. Análisis de metodologías.

En la revisión bibliográfica se identifican diferentes metodologías relacionadas con el análisis estratégico, táctico y operacional de un sistema de transporte, las cuales utilizan buses. Cada metodología tiene distintas variaciones que buscan el estudio de una o distintas variables según el autor, lo cual permite apreciar la diversidad de acercamientos desarrollados en el pasado.

Es importante tener en cuenta que, dentro de un proceso de análisis y diseño de un sistema de transporte, la etapa estratégica del análisis determina las variables básicas y la estructura del problema; lo cual determina el manejo de las dos siguientes etapas, la táctica y la operacional, enfoque de este trabajo, cinco alternativas metodológicas seleccionadas aplicables al estudio del sistema de transporte son:

1. Formulación Analítica.
2. Modelo de optimización – Solución por heurística de autor.
3. Modelo de optimización – VRP (Vehicle Route Problem) o red de cliente o punto de consumo.
4. Modelo de optimización – Solución por algoritmos genéticos.
5. Redes Neuronales.

11.1.1. Metodología Formulación Analítica.

Secciona el problema en partes más pequeñas, las cuales desarrolla mediante ecuaciones presentes en la literatura para acercarse a un valor exacto lo cual provee una visión general del sistema o de la variable a analizar.

Este tipo de metodologías tiene limitaciones debido a que tiende a simplificar los conceptos, y brinda aproximaciones que responden a un sistema ideal que presenta variaciones significativas en relación con el comportamiento real.

11.1.2. Metodología de optimización Heurística de autor.

Bajo esta metodología la mayor parte de los autores plantean una forma heurística para desarrollar los análisis de los sistemas de transporte, buscando valores óptimos para las distintas variables inmersas en cada etapa. Prefiriendo los modelos discretos que desarrollan una estrategia la cual busca alcanzar un óptimo, y utiliza un conjunto de ecuaciones objetivo y restricciones para representar el sistema.

La característica principal de esta metodología es que la solución al problema planteado resulta en valores que se encuentran dentro de un conjunto finito,

propiedad que es ideal para el tratamiento de variables enteras como el número de pasajeros o el número de vehículos.

11.1.3. Metodología de optimización heurística VRP

Un planteamiento heurístico especializado para la logística de vehículos relaciona las distintas variables bajo funciones objetivo, que encuentra un conjunto de rutas y la cantidad de vehículos a utilizar para dar respuesta a la demanda.

Este método está orientado hacia un centro de distribución y puntos de consumo, lo cual dificulta su aplicación en el sistema de transporte de pasajeros estudiado, ya que las rutas deben realizar el mismo recorrido. Presenta complejidad para su adaptación al sistema de transporte de pasajeros.

11.1.4. Metodología de inteligencia artificial - Algoritmos genéticos.

Utiliza algoritmos que permiten conocer soluciones a un planteamiento objetivo inspirado en procesos complejos como el comportamiento animal, en el caso de la optimización de partículas permite explorar las posibles soluciones en

11.1.5. Metodología de inteligencia artificial - Redes Neuronales.

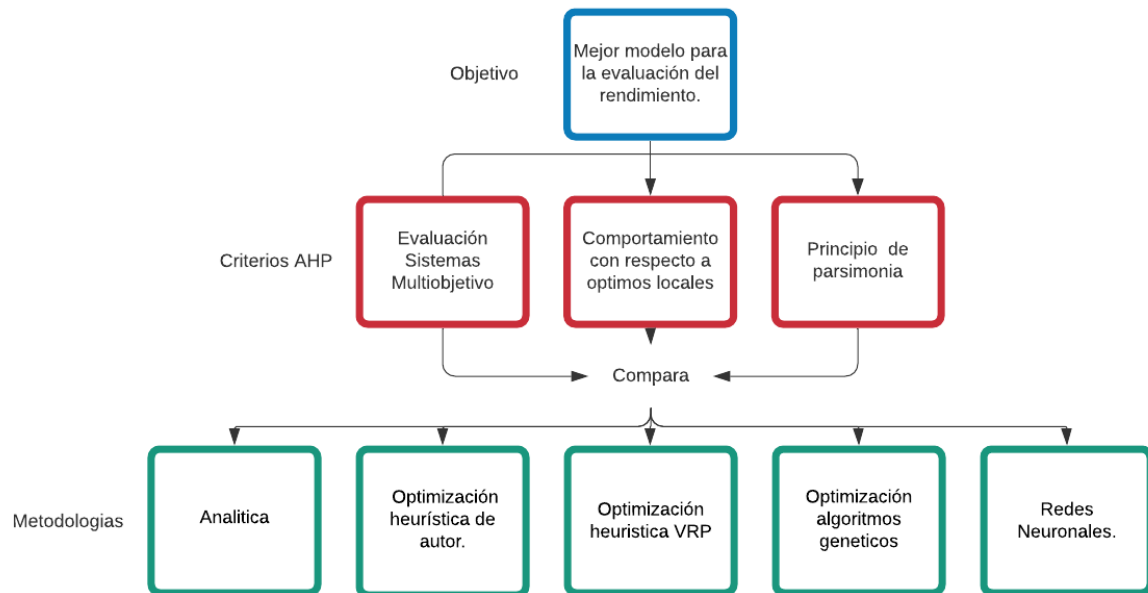
Identifica patrones del comportamiento de los datos inspirado en la forma de pensar del ser humano que con base en los datos de entrada y el entrenamiento previo logra generalizar reglas para el conjunto de datos.

11.1.6. Proceso analítico jerárquico, AHP.

Para seleccionar la metodología a utilizar, se aplica el método analítico jerárquico o AHP desarrollado a continuación.

- a) Definición del problema: ¿Cuál Metodología es la mejor opción para describir el sistema de transporte de pasajeros, que evalúe el rendimiento?
- b) Se pueden comparar las metodologías por AHP, debido a que se permite realizar comparaciones pareadas de sus características.
- c) Criterios para evaluar el rendimiento del sistema:
 - Evaluación de sistemas multiobjetivo.
 - Comportamiento con respecto a óptimos locales.
 - Principio de parsimonia.

Figura 3. Criterio para la comparación de metodologías por la técnica AHP



Nota. El diagrama muestra los criterios para tener en cuenta para la selección del método más apropiado para aplicar en la etapa táctica y operacional del estudio, Autoría propia.

d) Escala de medición.

Tabla 16. Escala de 9 puntos para la comparación pareada en AHP

Intensidad	Definición	Explicación
1	Igual importancia	Contribuyen igual
3	Importancia moderada	Juicio débil que prefiere un criterio o alternativa sobre otro.
5	Importancia fuerte	Juicio fuerte que prefiere un criterio o alternativa sobre otra.
7	Importancia muy fuerte	El criterio o alternativa es mucho más preferida que la otra.
9	Importancia absoluta	Existe evidencia que afirma que un criterio o alternativa es preferida sobre otra.

Nota. Escala de 1 a 9 que permite jerarquizar los criterios para elegir una de las alternativas, adaptada de Lee, G., & Chan, E. (2008). The analytic hierarchy process (AHP) approach for assessment of urban renewal proposals. *Social Indicators Research*, 89(1), 155. doi:10.1007/s11205-007-9228-x

e) Jerarquización de los criterios.

Tabla 17. *Jerarquización de criterios AHP*

	Matriz de criterios		
	Multiobjetivo	Óptimos locales	Principio de Parsimonia.
Multiobjetivo	1/1	3/1	7/1
Óptimos locales	1/3	1/1	3/1
Principio de parsimonia.	1/7	1/3	1 /1

Nota. Comparación pareada en escala de 1 a 9 según criterio del investigador.

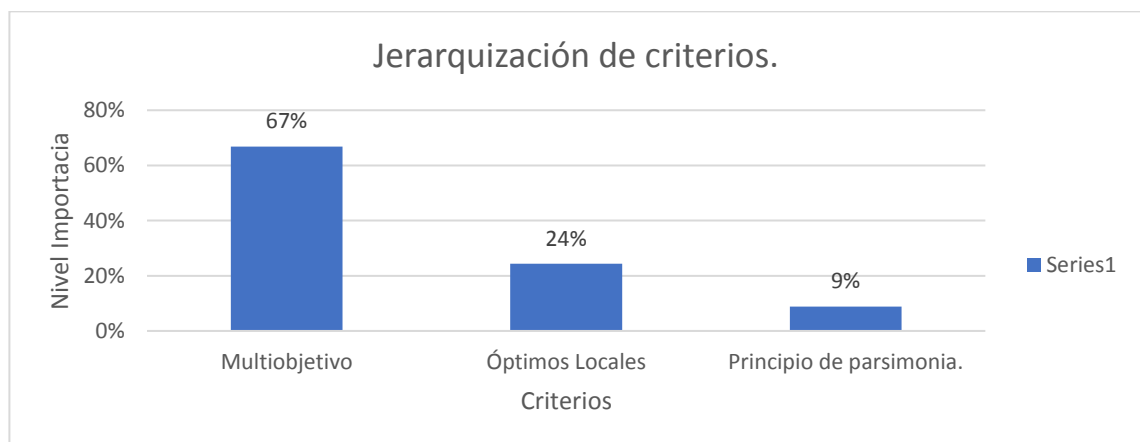
Esta matriz cumple con el criterio de consistencia para una matriz tres por tres como lo muestra la tabla a continuación, calculado por la razón entre el índice de consistencia IC y el índice de consistencia aleatoria.

Tabla 18. *Nivel de consistencia.*

Razones	Valor
IC	0,01
IA	0,66
RC	0,01

Nota. El nivel de consistencia aceptado para una matriz de 3x3 debe se acepta si el RC es menor a 0,1

Figura 4. *Jerarquización de criterios AHP*



Nota. La figura evidencia que el criterio con mayor peso corresponde al multiobjetivo para la selección de una metodología, representado con el 67%.

f) Jerarquización de metodologías.

Tabla 19. *Jerarquización en función del criterio multiobjetivo*

	Analítico	Heurística Autores	Heurística VRP	Algoritmo genético	Red Neuronal
Analítico	1/1	1/5	1/5	1/3	1/5
Heurística					
Autores	5/1	1/1	1/1	1/3	1/5
Heurística VRP	5/1	1/1	1/1	1/3	1/5
Algoritmo genético	3/1	3/1	3/1	1/1	1/5
Red Neuronal	5/1	5/1	5/1	5/1	1/1

Tabla 20. *Nivel de consistencia de la matriz de Jerarquización en función del criterio multiobjetivo.*

Razones	Valor
IC	0,59
IA	0,99
RC	0,60

Nota. El nivel de consistencia aceptado para una matriz de 4x4 debe ser aceptado si el RC es menor a 0,9

Tabla 21. *Jerarquización en función del criterio Óptimos locales.*

	Analítico	Heurística Autores	Heurística VRP	Algoritmo genético	Red Neuronal
Analítico	1/1	3/1	3/1	3/1	3/1
Heurística					
Autores	1/3	1/1	1/1	1/3	1/5
Heurística VRP	1/3	1/1	1/1	1/3	1/5
Algoritmo genético	1/3	3/1	3/1	1/1	1/5
Red Neuronal	1/3	5/1	5/1	5/1	1/1

Tabla 22. *Nivel de consistencia de la matriz de Jerarquización en función del criterio Óptimos locales.*

Razones	Valor
IC	0,59
IA	0,99
RC	0,60

Nota. El nivel de consistencia aceptado para una matriz de 4x4 debe se acepta si el RC es menor a 0,9

Tabla 23. *Jerarquización en función del criterio Óptimos locales.*

	Analítico	Heurística Autores	Heurística VRP	Algoritmo genético	Red Neuronal
Analítico	1/1	3/1	3/1	3/1	7/1
Heurística Autores	1/3	1/1	1/1	3/1	7/1
Heurística VRP	1/3	1/1	1/1	3/1	7/1
Algoritmo genético	1/3	1/3	1/3	1/1	5/1
Red Neuronal	1/7	1/7	1/7	1/5	1/1

Tabla 24. *Nivel de consistencia de la matriz de Jerarquización en función del criterio Óptimos locales.*

Razones	Valor
IC	0,59
IA	0,99
RC	0,60

Nota. El nivel de consistencia aceptado para una matriz de 4x4 debe se acepta si el RC es menor a 0,9

Tabla 25. *Resultados Jerarquización por los diferentes criterios del método AHP*

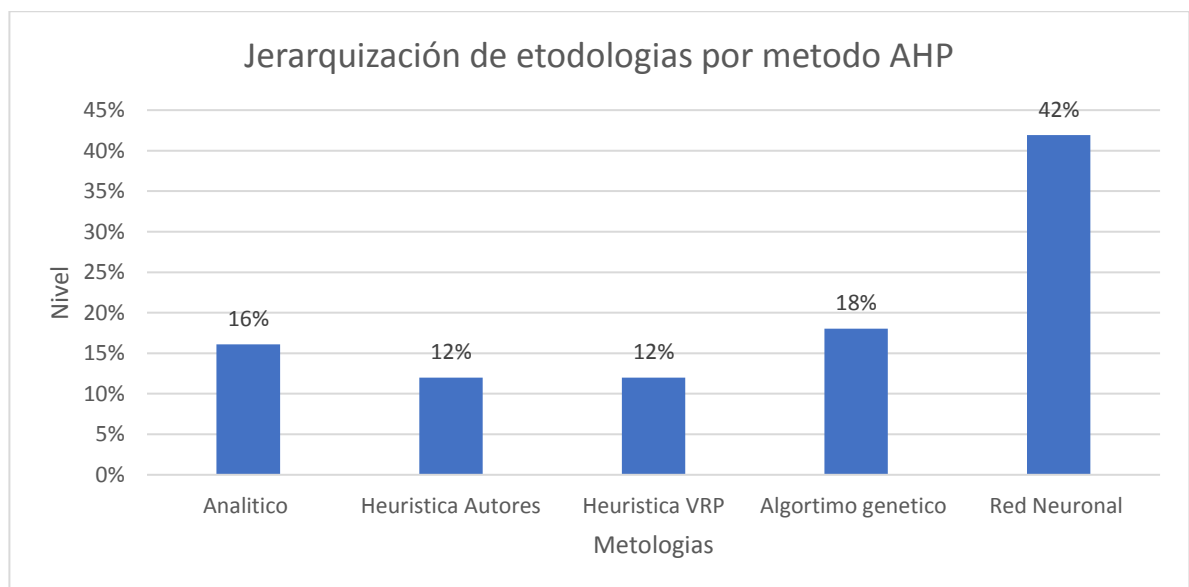
	Criterio multiobjetivo	Criterio Óptimos Locales	Principio de parsimonia
Analítico	5%	37%	42%
Heurística Autores	12%	7%	21%
Heurística VRP	12%	7%	21%
Algoritmo genético	20%	15%	11%
Red Neuronal	50%	33%	4%

Nota. En la tabla se relaciona la jerarquización de los métodos en función de los tres criterios de comparación.

De acuerdo con la anterior comparación se observa que los métodos de inteligencia artificial, redes artificiales y algoritmos genéticos logran abordar mucho mejor la complejidad multiobjetivo. También que las metodologías que son menos afectadas por los óptimos locales son las metodologías analíticas y teniendo en cuenta la facilidad de su aplicación o principio de parsimonia los métodos analíticos son las mejores.

Por tanto, la metodología AHP permite establecer la metodología más apropiada según los tres criterios unidos, obteniendo:

Figura 5. Jerarquización por metodología AHP



Nota. La metodología más recomendada para el estudio del sistema es las redes neuronales con 42% seguida de los algoritmos genéticos y la formulación analítica con un 16% bajo el criterio del investigador.

11.1.7. Selección de la metodología.

Teniendo en cuenta la jerarquización de los criterios, y las metodologías comparadas en los puntos anteriores, se establece que la mejor opción corresponde a la aplicación de redes neuronales, pero esta opción implica un alto grado de complejidad ya que el criterio de parsimonia, implementación sencilla, es muy bajo. Adicionalmente requiere un nivel de conocimiento especializado y herramientas computacionales más robustas. Motivo por el cual se considera aplicar las opciones que ocupan el segundo lugar con un 18% de recomendación, Algoritmo genético, y el tercer lugar con 16% de recomendación correspondiente a la formulación analítica.

11.2. Etapa 2. Datos de entrada y planteamiento del sistema de transporte de pasajeros.

11.2.1. Información externa y suposiciones.

Como factores complementarios, la información a continuación soporta el planteamiento del sistema de estudio, con base en comunicados oficiales de la empresa Transmilenio S. A., el observatorio de Transmilenio de la Universidad de los Andes y la Cámara de Comercio de Bogotá.

- a) Existen horas del día en las cuales el sistema presenta un comportamiento que no es relevante para el objetivo del estudio. En estos periodos de tiempo la mayor parte de las rutas no ha empezado su operación o ya finalizaron su jornada, Por esa razón, a partir de este momento se limita el horario estudiado.
 - Horario de inicio: 4:00 a. m.
 - Horario de fin: 11:00 p. m.
- b) Existen picos y valles en la tendencia del sistema los cuales esta relacionados con la cantidad de pasajeros por día que entran en una estación; son horas en las cuales su comportamiento es característico.
 - El primer pico se presenta antes o después de las 7 a.m.
 - El segundo pico se acerca hacia las 5 p.m.
 - Las horas valle se presentan en horarios en los cuales la cantidad de pasajeros disminuye considerablemente, contrariamente en lo que ocurre en las horas pico mencionadas.
- c) De conformidad con las cifras emitidas oficialmente, en el mismo periodo del año 2017, hay un comportamiento característico que depende de variables aleatorias y exógenas al sistema, lo cual implica una clasificación del comportamiento de la demanda según fechas y la estación que se analiza, lo que permite agrupar los estudios de la siguiente forma:
 - Agrupación por mes.
 - Agrupación por días.
 - Agrupación por horas.
- d) Se encontró que los meses de octubre y noviembre tienen datos faltantes en la fuente de información utilizada para el estudio, por lo tanto, los periodos incompletos no representan el comportamiento del sistema, por esta razón los datos asociados a estas fechas no serán tenidos en cuenta.

- e) El mes de diciembre presenta características únicas que pueden afectar el análisis normal del sistema y debido a que distorsiona la continuidad del periodo analizado, no será tenido en cuenta debido a que asumen un comportamiento no conocido.

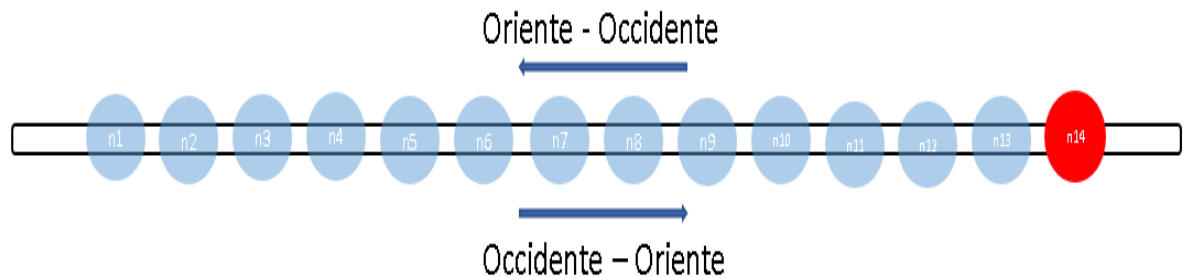
11.2.2. Planteamiento modelo

El sistema de doble sentido presenta el comportamiento del servicio de movimiento de pasajeros de occidente a oriente y de oriente a occidente. Está representado por nodos o estaciones conectados por una vía continua donde circulan buses, y adicionalmente se utiliza un nodo final que acumula la demanda de los puntos que salen del área estudiada.

La estructura propuesta acumula los pasajeros que tienen como destino el final de la ruta de todos los servicios en el nodo catorce al oriente de la ciudad el cual corresponde a la estación universidades.

Dicho planteamiento permite que la operación sea descrita en los diferentes puntos del sistema, los cuales responden a la demanda de pasajeros, el momento temporal y las variables relacionadas a la operación de los buses.

Figura 6. *Diagrama sistema de estaciones*



Nota. El diagrama explica de una forma sencilla la configuración del sistema estudiado en la sección escogida para el estudio. Diagrama de autoría propia.

Distancia.

Por medio de la herramienta de mapas ArcGis, se obtuvo la distancia promedio entre las estaciones, como la medida entre el centro de una estación y la estación inmediatamente próxima, de esta forma se calcula que la sección estudiada corresponde a 21,58 kilómetros de vía exclusiva desde la estación Portal el dorado hasta la estación centro memoria.

Se separa una segunda sección que corresponde al final de la vía, el trayecto entre la estación centro memoria y la estación universidades con 2,23 km, debido a que tiene una distancia diferente con respecto a las otras estaciones, además este el final de la ruta es propuesto como un nodo que tendrá un comportamiento distinto para este estudio.

Por lo anterior la distancia promedio del trayecto entra una estación otra es de 0,83 km, sin tener en cuenta el tramo entre la estación centro memoria y la estación Universidades.

Tabla 26. *Distancias entre Estaciones*

Trayecto	Estación 1	Estación 2	Distancia (km)
1	Portal Eldorado	Modelia	0,88
2	Modelia	Normandía	0,81
3	Normandía	Av. Rojas	0,93
4	Av. Rojas	El Tiempo	0,63
5	El Tiempo	Salitre Greco	0,82
6	Salitre Greco	CAN	0,52
7	CAN	Gobernación	0,52
8	Gobernación	Quinta Paredes	0,7
9	Quinta Paredes	Recinto ferial	0,57
10	Recinto Ferial	Ciudad Universitaria	0,78
11	Ciudad Universitaria	Concejo de Bogotá	0,65
12	Consejo Electoral	Centro de Memoria	0,75
13	Centro de Memoria	Universidades	2,23
Total. Distancias sin tener en cuenta el trayecto 13			8,56
Total. Distancia de todos los trayectos			10,79

Nota. Las distancias corresponden a una vía lineal que conecta consecutivamente las estaciones.

Velocidad de los vehículos.

Según el programa de gestión de la velocidad realizado por la alcaldía de Bogotá para el año 2019, los límites de velocidad que respetan un nivel de seguridad para los diferentes actores presentes en la ciudad dependen del tipo de vía analizada.

De acuerdo con las características de las vías clasificadas como corredores arteriales, vías que están diseñadas para el uso de vehículos, no segregadas para el uso exclusivo de un solo modo de transporte y que no cuentan con espacios adyacentes libres, se recomienda una velocidad entre 40 y 50 km/las vías conocidas como autopistas, que cuentan con zonas adyacentes libres de objetos, la velocidad puede alcanzada puede ser entre 60 km/h y 70 km/h.

Tabla 27. *Tipo de vías*

Tipo de vías	Características	Límite de velocidad máximo (km/h)
Corredores Arteriales	Presencia de otros modos de transporte como ciclistas y peatones.	40-50
Autopistas	Vía segregada para el uso de vehículos motorizados Zonas adyacentes libre de objetos	>60

Nota. Datos adaptados de: Programa de gestión de la velocidad 2019, Alcaldía mayor de Bogotá, 2019

Por otra parte, en las fuentes oficiales de información sobre la velocidad del sistema BRT Transmilenio no se relaciona con exactitud la velocidad promedio con la que opera el servicio troncal, pero de acuerdo con la informe rendición de cuentas realizado por Transmilenio en el 2017, la velocidad promedio de una troncal ha alcanzado valores cercanos a los 30 km/h en horas de mayor demanda en diferentes troncales.

Tiempo de viaje.

De acuerdo con el observatorio de movilidad de Bogotá para el 2017, el tiempo de viaje promedio en la ciudad de Bogotá es de 57 minutos teniendo en cuenta los diferentes modos de transporte en una vía compartida, al discriminar los viajes en el modo Transmilenio los tiempos de viaje rondan los 78 minutos, pero debido a que la sección analizada tiene una distancia más pequeña el tiempo de viaje promedio puede ser mucho menor.

El tiempo de un usuario en el sistema será medido de acuerdo con los diferentes tiempos que componen un viaje, el tiempo de espera en la estación, el tiempo de transferencia en el caso que sea pertinente, el tiempo de abordaje y el tiempo de recorrido.

Entradas de pasajeros al sistema.

Los datos corresponden a muestras del número de pasajeros que accedieron a una estación específica de la red durante los meses de febrero, marzo, abril, mayo, junio, julio, agosto, septiembre, del año 2017, conjunto de datos segmentado por las variables que se muestran en la siguiente tabla.

Tabla 28. Descripción de datos

Nombre de variable	Descripción	Tipo de variable
Fecha	Corresponde a una variable vinculada a una escala de tiempo que describe en qué día exacto ocurre el hecho.	Numérica, ordinal y temporal.
Mes	Concierne al periodo específico al cual pertenece un registro. Puede tomar los valores desde enero hasta diciembre y ser resumido mediante una codificación numérica.	Cadena, ordinal y temporal.
Día de la semana	Relacionado con la variable “fecha”, asigna a qué día de la semana corresponde determinado dato. Los posibles valores que puede tomar son: lunes, martes, miércoles, jueves, viernes, sábado y domingo.	Cadena, ordinal y temporal.
Hora	Relaciona en una escala de 24 horas el momento de un registro de un dato de demanda.	Numérica, ordinal y temporal.
Estación	Corresponde a un nodo dentro del sistema, que relaciona un punto geográfico de la ciudad identificado con el nombre de un espacio referente a las estaciones analizadas. Para facilitar el uso de herramientas de análisis mediante software esta variable es representada por un número.	Cadena, nominal y numérica.
Entradas al sistema	Relaciona la cantidad de pasajeros que accedieron a una estación en el tiempo correspondiente.	Numérica, ordinal y escalar.

Nota. Mediante la anterior descripción se observa que los datos pueden ser organizados, agrupados y analizados de formas diferentes, siguiendo variables de tiempo y espacio en torno a la demanda.

11.3. Etapa 3. Análisis exploratorio y comportamiento de la demanda

se le aplican medidas de tendencia central y variabilidad, utilizando Excel y el software IBM SPSS, con el objetivo de describir el comportamiento del sistema de transporte de pasajeros.

11.3.1. Media.

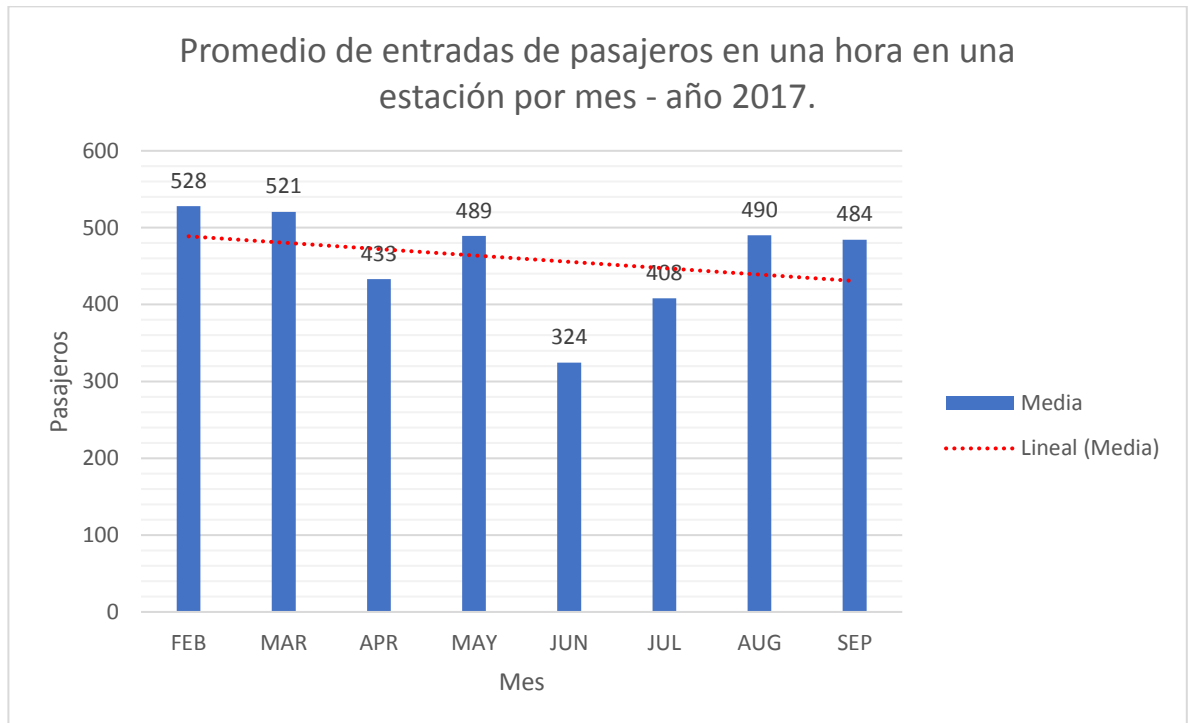
A continuación, se representan los gráficos de la media.

se determina que el promedio de entradas en toda la troncal en una hora es 459 pasajeros.

Media agrupada por meses:

Se observa que el mayor promedio de pasajeros por hora de cada mes es de 528, que corresponden al mes de febrero y el menor promedio de pasajeros por hora de cada mes es de 324 y corresponde al mes de junio, lo cual se relaciona con las dinámicas de la ciudad de Bogotá.

Figura 7. Comportamiento de las entradas al sistema agrupadas por meses

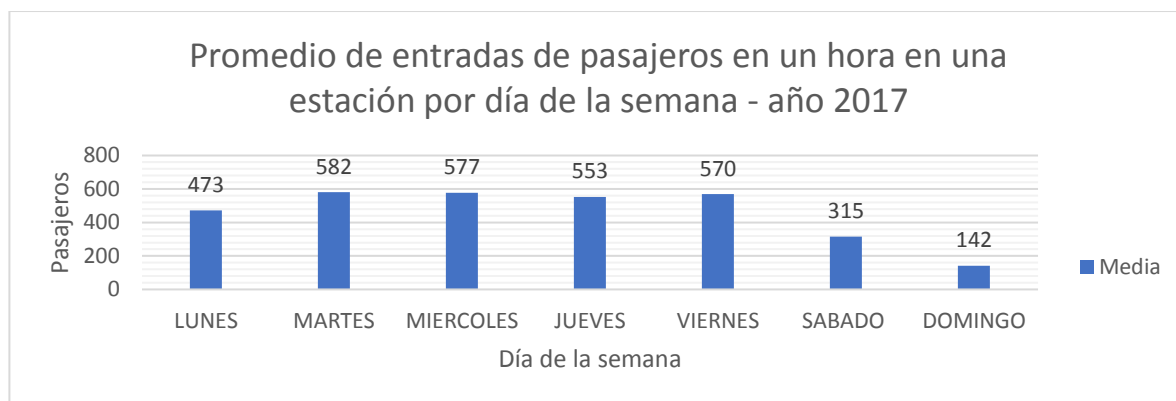


Nota. Promedio de entradas por hora en un mes, medidas en número de pasajeros.

Media por agrupación de día de la semana.

Se observa que el mayor promedio de pasajeros por hora según el día de la semana es de 582 que corresponde al día martes y el menor valor del promedio de pasajeros por hora según día de la semana es de 142 pasajeros corresponde al día domingo.

Figura 8. Comportamiento de las entradas al sistema agrupadas por hora por semana

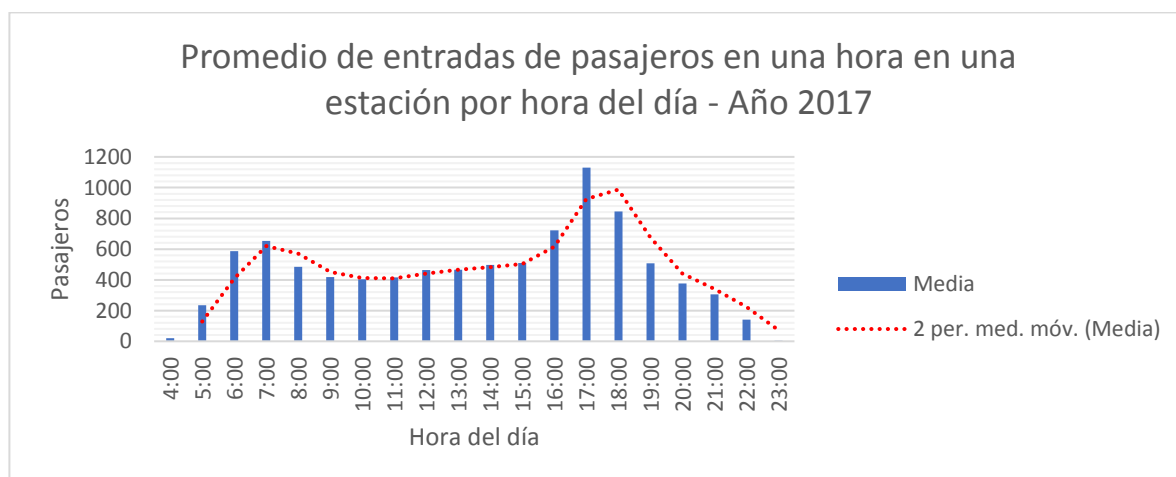


Nota. Promedio de entradas por hora en una semana, medidas en número de pasajeros.

Media por agrupación por horas.

Se observa que el mayor promedio de pasajeros por hora según la hora del día es de 1131 y corresponde a las 5 pm y el menor promedio de pasajeros por hora según la hora del día es de 4 am y corresponde a las 11 pm.

Figura 9. Comportamiento de las entradas al sistema agrupadas por horas del día



Nota. Promedio de entradas en una hora, medidas en número de pasajeros según el momento del día.

Media por agrupación por estaciones.

Se observa que el mayor promedio de pasajeros por hora por estación es de 2393, que corresponde a la estación portal el Dorado y el menor valor promedio por hora por estación es de 134 que corresponde a la estación plaza de la democracia.

Figura 10. Comportamiento de las entradas al sistema agrupadas por horas del día



Nota. Promedio de entradas en una hora, medidas en número de pasajeros según la estación.

11.3.2. Estadística descriptiva agrupación por meses.

El valor de la mediana indica la existencia de valores extremos ya que a diferencia de la media es más estable, también se observa que las muestras agrupadas, tienen una gran variación. Por otra parte, la distribución sigue una asimetría positiva, lo cual indica que los valores tienden hacia valores inferiores.

Tabla 29. *Estadística descriptiva de la demanda por mes.*

MES	Suma	% de suma total	% de N total	Mínimo	Máximo	Rango	Media	Error estándar de la media	Mediana	N	Desviación	Varianza	Curtosis	Asimetría	Coficiente variación
FEB	4139778	13,3%	11,6%	0	8932	8932	528,03	10,206	265,00	7840	903,681	816639,346	27,686	4,619	1,71141038
MAR	4518164	14,5%	12,8%	0	8249	8249	520,53	9,547	260,00	8680	889,471	791158,123	27,308	4,595	1,7087927
APR	3637658	11,7%	12,4%	0	7744	7744	433,05	8,486	205,00	8400	777,779	604939,930	30,855	4,828	1,79602983
MAY	4245626	13,6%	12,8%	0	7839	7839	489,13	8,994	240,00	8680	837,965	702185,219	28,996	4,707	1,7131833
JUN	2724078	8,8%	12,4%	0	6201	6201	324,30	6,527	160,00	8400	598,246	357898,721	34,277	5,185	1,84475978
JUL	3542128	11,4%	12,8%	0	7331	7331	408,08	7,813	204,00	8680	727,954	529916,799	32,347	5,003	1,78385405
AUG	4254053	13,7%	12,8%	0	8127	8127	490,10	9,126	245,00	8680	850,225	722883,277	30,085	4,808	1,73480601
SEP	4069274	13,1%	12,4%	0	8118	8118	484,44	9,027	250,00	8400	827,354	684513,983	30,608	4,832	1,70786491
Total	31130759	100,0%	100,0%	0	8932	8932	459,43	3,108	228,00	67760	808,973	654437,454	31,007	4,860	1,76083135

Agrupación por meses

a. Los datos contienen valores negativos y positivos y, posiblemente valores cero.

Nota. Medidas de estadística descriptiva que muestra el comportamiento general del sistema agrupado por meses.

11.3.3. Estadística descriptiva por días de la semana.

Se observan cuatro días con valores altos en la media; martes miércoles, jueves y viernes, que caracterizan el comportamiento de la demanda en la semana, con valores inferiores en los lunes y los fines de semana. Igualmente presenta gran variación con respecto a la media por las medidas de desviación estándar y varianza evidente en los coeficientes de variación, y también se observa que el comportamiento de los datos por meses tiende a tener valores inferiores, a causa de las medidas de curtosis.

Tabla 30. *Estadística descriptiva de la demanda por día de la semana*

DAY	Suma	% de suma total	% de N total	Mínimo	Máximo	Rango	Media	Error estándar de la media	Mediana	N	Desviación	Varianza	Curtosis	Asimetría	Coeficiente variación
SUN	1351568	4,3%	14,0%	0	1858	1858	141,97	2,786	62,00	9520	271,858	73906,810	12,271	3,541	1,91487873
MON	4500594	14,5%	14,0%	0	7888	7888	472,75	8,646	232,00	9520	843,620	711695,444	29,536	4,806	1,78449035
TUE	5538217	17,8%	14,0%	0	8249	8249	581,75	9,497	318,00	9520	926,664	858706,486	24,433	4,416	1,59290308
WED	5652019	18,2%	14,5%	0	8187	8187	576,74	9,357	316,00	9800	926,278	857991,605	24,977	4,458	1,60606819
THU	5421498	17,4%	14,5%	0	8932	8932	553,21	9,137	297,00	9800	904,546	818202,836	26,505	4,563	1,63507344
FRI	5581664	17,9%	14,5%	0	8028	8028	569,56	9,027	320,00	9800	893,620	798555,896	23,922	4,359	1,56897147
SAT	3085199	9,9%	14,5%	0	4072	4072	314,82	5,393	165,00	9800	533,917	285067,284	15,033	3,750	1,69596382
Total	31130759	100,0%	100,0%	0	8932	8932	459,43	3,108	228,00	67760	808,973	654437,454	31,007	4,860	1,76083135

Agrupación por tipo de día. a. Los datos contienen valores negativos y positivos y, posiblemente, valores cero.

11.3.4. Estadística descriptiva por horas del día.

Se observan valores pequeños en las primeras horas del día y en las últimas, y la entrada de pasajeros alcanza valores máximos en las 7:00 a.m. y las 5:00 p.m., cuando el máximo de la media alcanza los 1113 pasajeros, como lo muestra la tabla a continuación.

Tabla 31. *Estadística descriptiva de la demanda por hora*

	Suma	% de suma total	% de N total	Mínimo	Máximo	Rango	Media	Error estándar de la media	Mediana	N	Desviación	Varianza	Curtosis	Asimetría	
4:00:00 a. m.	70613	0,2%	5,0%	0	419	419	20,84	1,197	3,00	3388	69,678	4855,057	15,783	4,129	3,34315017
5:00:00 a. m.	793054	2,5%	5,0%	0	4154	4154	234,08	11,035	68,00	3388	642,303	412552,739	15,501	4,077	2,74397644
6:00:00 a. m.	1988942	6,4%	5,0%	0	8854	8854	587,05	24,611	192,00	3388	1432,516	2052103,407	15,056	4,005	2,44017461
7:00:00 a. m.	2216053	7,1%	5,0%	0	8402	8402	654,09	25,793	235,00	3388	1501,312	2253937,984	14,435	3,936	2,29527243
8:00:00 a. m.	1644429	5,3%	5,0%	0	6053	6053	485,37	16,050	245,00	3388	934,197	872723,877	13,654	3,799	1,92471621
9:00:00 a. m.	1415821	4,5%	5,0%	0	4171	4171	417,89	10,953	258,00	3388	637,564	406487,821	11,544	3,477	1,52566373
10:00:00 a. m.	1367822	4,4%	5,0%	0	3081	3081	403,73	9,029	269,00	3388	525,561	276214,870	9,591	3,164	1,30177925
11:00:00 a. m.	1410372	4,5%	5,0%	0	2939	2939	416,28	8,435	272,00	3388	490,998	241078,916	7,831	2,850	1,17947663
12:00:00 p. m.	1572878	5,1%	5,0%	0	3532	3532	464,25	8,864	298,00	3388	515,926	266179,682	7,510	2,706	1,11131152
1:00:00 p. m.	1577151	5,1%	5,0%	0	3190	3190	465,51	9,156	299,00	3388	532,947	284032,985	7,118	2,651	1,14486562
2:00:00 p. m.	1686735	5,4%	5,0%	0	3574	3574	497,86	10,134	318,00	3388	589,849	347922,413	8,402	2,919	1,18478009
3:00:00 p. m.	1728995	5,6%	5,0%	0	3328	3328	510,33	9,889	316,00	3388	575,623	331341,935	6,222	2,542	1,12794485
4:00:00 p. m.	2444056	7,9%	5,0%	0	6057	6057	721,39	13,114	466,50	3388	763,323	582661,254	6,368	2,384	1,05813315
5:00:00 p. m.	3830264	12,3%	5,0%	0	8932	8932	1130,54	22,904	825,00	3388	1333,136	1777252,254	9,886	2,985	1,17920478
6:00:00 p. m.	2859935	9,2%	5,0%	22	5804	5782	844,14	16,965	597,00	3388	987,492	975139,728	8,861	2,883	1,16982437
7:00:00 p. m.	1718193	5,5%	5,0%	0	3385	3385	507,14	9,313	314,00	3388	542,056	293824,832	4,615	2,171	1,0688474
8:00:00 p. m.	1272123	4,1%	5,0%	0	2291	2291	375,48	7,442	196,50	3388	433,200	187662,159	3,373	1,976	1,15372592
9:00:00 p. m.	1038405	3,3%	5,0%	5	2474	2469	306,49	7,007	157,00	3388	407,842	166335,100	7,253	2,638	1,33066454
10:00:00 p. m.	480679	1,5%	5,0%	0	1338	1338	141,88	3,565	76,00	3388	207,487	43051,047	7,602	2,780	1,46244693
11:00:00 p. m.	14239	0,0%	5,0%	0	202	202	4,20	0,142	2,00	3388	8,268	68,364	132,777	8,353	1,96733546
	31130759	100,0%	100,0%	0	8932	8932	459,43	3,108	228,00	67760	808,973	654437,454	31,007	4,860	1,76083135

11.3.5. Estadística descriptiva por estaciones de la troncal.

Se observan tres rangos de ingresos de pasajeros: muy alto, como el correspondiente al portal El Dorado, un rango medio que se encuentra entre los 200 y los 600 pasajeros; por último, los valores bajos encontrados en las estaciones Quinta Paredes, la plaza Democracia y el Centro de Memoria Histórica.

Tabla 32. *Estadística descriptiva de la demanda por estaciones*

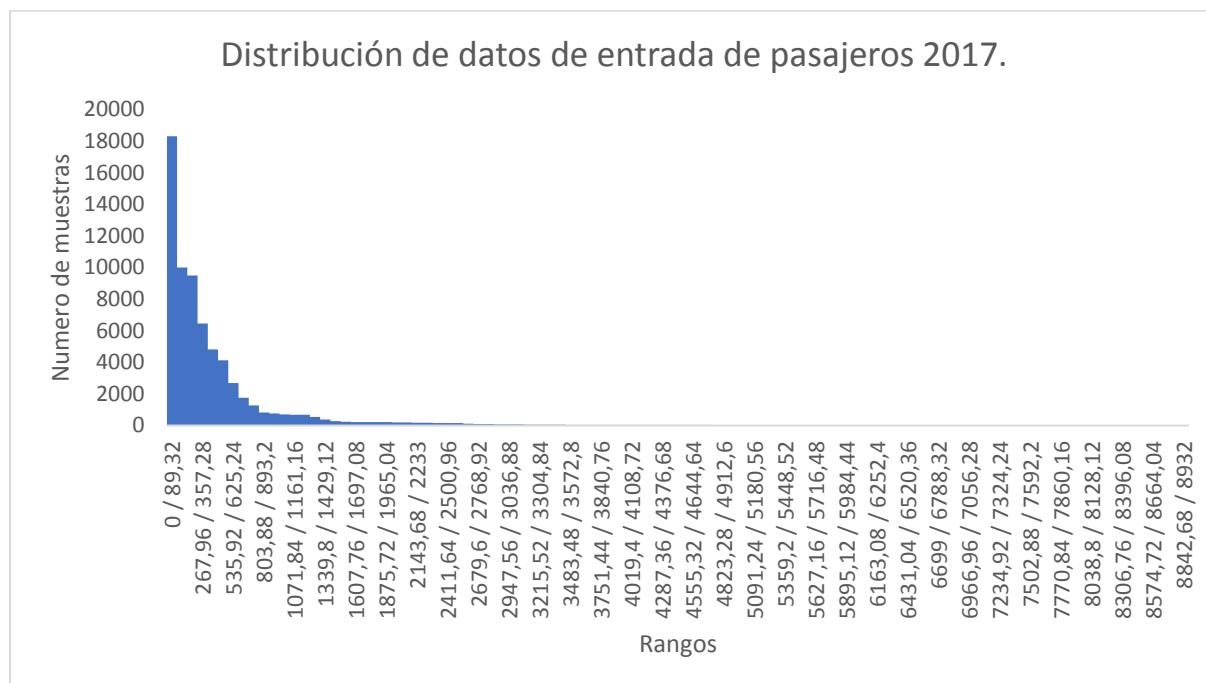
STATION		Suma	% de suma total	% de N total	Mínimo	Máximo	Rango	Media	Error estándar de la media	Mediana	N	. Desviación	Varianza	Curtosis	Asimetría	
(06000) Portal Eldorado	Portal	11583718	37,2%	7,1%	0	8932	8932	2393,33	26,873	2144,50	4840	1869,582	3495338,048	1,029	1,197	Coeficiente variación
(06001) Modelia	Modelia	1159438	3,7%	7,1%	0	1289	1289	239,55	3,335	186,00	4840	232,048	53846,282	3,428	1,730	
(06002) Normandía	Normandía	1278746	4,1%	7,1%	0	1068	1068	264,20	3,307	230,00	4840	230,042	52919,460	1,186	1,252	
(06100) Av. Rojas	Av. Rojas	1378088	4,4%	7,1%	0	2416	2416	284,73	3,749	239,00	4840	260,832	68033,424	3,950	1,609	
(06101) El Tiempo	El Tiempo	1980563	6,4%	7,1%	0	2083	2083	409,21	5,212	388,50	4840	362,592	131473,103	3,836	1,732	
(06102) Salitre El Greco	Salitre El Greco	1923095	6,2%	7,1%	0	1527	1527	397,33	4,472	363,50	4840	311,106	96786,709	0,542	0,873	
(06103) CAN	CAN	1749100	5,6%	7,1%	0	2040	2040	361,38	5,779	244,50	4840	402,062	161653,760	3,410	1,861	
(06104) Gobernación	Gobernación	1380009	4,4%	7,1%	0	2004	2004	285,13	5,228	160,00	4840	363,708	132283,229	5,320	2,208	
(06105) Quinta Paredes	Quinta Paredes	717352	2,3%	7,1%	0	1325	1325	148,21	2,089	128,00	4840	145,311	21115,251	5,180	1,941	
(06106) Corferias	Corferias	1389612	4,5%	7,1%	0	1867	1867	287,11	3,943	235,00	4840	274,306	75244,043	4,425	1,918	
(06107) Ciudad Universitaria	Ciudad Universitaria	1856875	6,0%	7,1%	0	1621	1621	383,65	4,766	328,00	4840	331,556	109929,374	0,491	0,999	
(06108) Plaza de la Democracia	Plaza de la Democracia	646293	2,1%	7,1%	0	722	722	133,53	1,860	104,00	4840	129,421	16749,813	3,091	1,602	
(06109) Centro Memoria	Centro Memoria	884026	2,8%	7,1%	0	1095	1095	182,65	2,464	151,00	4840	171,430	29388,354	4,912	1,942	
(06111) Universidades	Universidades	3203844	10,3%	7,1%	0	3180	3180	661,95	9,736	450,00	4840	677,344	458794,659	1,044	1,343	
Total		31130759	100,0%	100,0%	0	8932	8932	459,43	3,108	228,00	67760	808,973	654437,454	31,007	4,860	1,02325335
																1,76083135

Agrupación por estaciones, a. Los datos contienen valores negativos y positivos y, posiblemente valores cero.

11.3.6. Distribución de los datos

La mayor frecuencia absoluta es de 18000, muestras que se encuentran en el primer intervalo de clase cero (0) - 89.32 número de pasajeros entrando al sistema en una hora.

Figura 11. *Histograma del nivel de pasajeros que entran al sistema en una hora.*



Nota. La figura se observa que todas las muestras no están distribuidas como una distribución normal, histograma realizado con 100 rangos de frecuencia.

Tabla 33. *Estadística descriptiva de toda la muestra.*

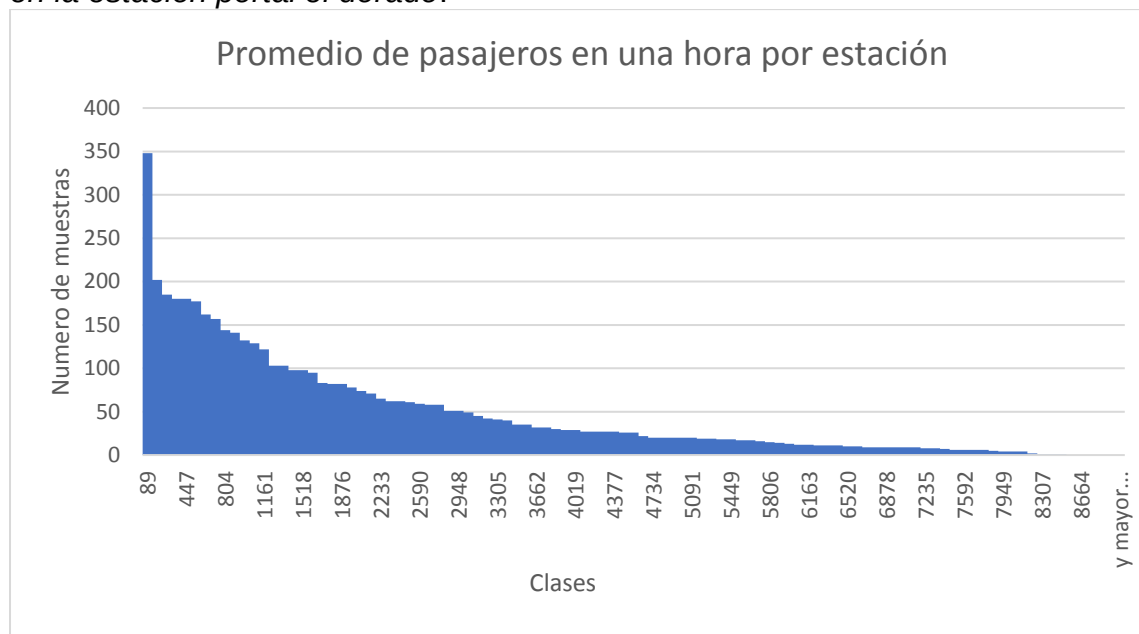
DEMANDA	
Media	459,426786
Error típico	3,1077593
Mediana	228
Moda	0
Desviación estándar	808,973086
Varianza de la muestra	654437,454
Curtosis	31,0066075
Coeficiente de asimetría	4,86025733
Rango	8932
Mínimo	0
Máximo	8932
Suma	31130759
Cuenta	67760

Nota. Las medidas indican que la variación de la muestra es muy grande, como lo muestra la desviación estándar con 808 y la varianza 654437.

11.3.7. Variabilidad

La anterior descripción no es suficiente, debido a que la variabilidad es muy alta, como se puede observar en las estadísticas descriptivas. Al tomar como ejemplo la estación Portal El Dorado, este análisis dice que el promedio de entradas a la estación es de 2393 pasajeros con una desviación de 1869, es decir que los datos se pueden alejar de la media hasta un 78%. Precisión que no es relevante para un sistema de transporte.

Figura 12. *Histograma del nivel de pasajeros que entran al sistema en una hora. en la estación portal el dorado.*



Nota. Histograma realizado con 100 rangos

Figura 13. *Estadística descriptiva entradas al sistema portal el dorado*

DEMANDA	
Media	2393,33017
Error típico	26,8733564
Mediana	2144,5
Moda	0
Desviación estándar	1869,58232
Varianza de la muestra	3495338,05
Curtosis	1,02910222
Coeficiente de asimetría	1,19692304
Rango	8932
Mínimo	0
Máximo	8932
Suma	11583718
Cuenta	4840

Nota. Tabla realizada con los registros de la estación Portal el dorado.

11.3.8. Conclusiones. Estadística Descriptiva

Se encuentra por medio de la estadística descriptiva que los datos de la troncal tienen una gran variabilidad, los datos de entrada de pasajeros difieren bastante del promedio, en otras palabras, los datos anteriores presentan una descripción muy sesgada, y, en consecuencia, con dicho planteamiento se estipula la hipótesis de que los datos tienen comportamientos diferentes, según el día de la semana, la hora del día y la estación relacionada.

Con el análisis particular de la muestra por las variables categóricas mes, día, hora y estación se encontró que:

1. El mes con mayor demanda es febrero, cuando alcanza los 529 pasajeros (más o menos 10 por hora y por estación), y el menor en el mes de junio, con 325 más o menos 7 pasajeros.
2. Agrupados por día de la semana, se observan diferentes niveles de demanda, es decir, que por lo menos existen entre 3 y 4 comportamientos distintos, según el día.
3. Se confirma la presencia de picos horarios en la agrupación por horas, tal como se planteaba en las premisas, basadas en información externa.
4. Todas las formas de agrupación muestran niveles altos de variabilidad.
5. Una mayor segmentación no disminuye la variabilidad del grupo.

Debido a la variabilidad de los datos no se recomienda continuar con esta forma de tratar los datos, ya que ignora las dinámicas del sistema. Se puede concluir que los datos tomados como un solo grupo no siguen una distribución normal.

11.4. Pruebas estadísticas detalladas de la demanda

Debido a que la forma anterior de analizar los datos presenta gran variabilidad, es necesario analizar las muestras desde otra perspectiva.

11.4.1. Pruebas de homogeneidad

Mediante estas pruebas se propone comprobar si las muestras corresponden a una misma población o, bajo otra forma de interpretación, conocer si los datos de una medición de variables segmentadas en varios grupos tienen una misma distribución o, por el contrario, cada distribución es diferente.

Debido a que las pruebas paramétricas exigen que los datos cumplan un mínimo de dos supuestos:

- Que sigan una distribución normal.
- Que tengan igualdad de varianzas

Se observa que en este caso el conjunto de datos no cumple con ninguno de los dos supuestos por lo cual se decide aplicar pruebas no paramétricas.

PRUEBA NO PARAMÉTRICA, ANOVA Kruskal Wallis y Medias:

Consiste en un contraste no paramétrico de la varianza, utilizado para comparar varias muestras al mismo tiempo, midiendo la tendencia central de las muestras. Esta prueba busca encontrar si se cumple alguna de las dos siguientes hipótesis:

- Hipótesis nula: las poblaciones testeadas tienden a la misma mediana.
- Hipótesis alternativa: se rechaza la hipótesis de medianas similares.

Pruebas de homogeneidad datos agrupados por estaciones.

Como lo prueba el test de Kruskal Wallis, se rechaza la hipótesis nula de homogeneidad, al igual que el test de la mediana, por lo cual podemos concluir que esta prueba rechaza la hipótesis de que el conjunto agrupado por estaciones sea homogéneo y siga una misma distribución.

Tabla 34. Homogeneidad Kruskal Wallis por estaciones
Estadísticos de prueba^{a,b}

	DEMANDA
H de Kruskal-Wallis	14988,500
Gl	13
Sig. asintótica	0,000

a. Prueba de Kruskal Wallis

b. Variable de agrupación: STATIONNUM

Tabla 35. Homogeneidad Medias por estaciones
Estadísticos de prueba

	DEMANDA
N	67760
Mediana	228,00
Chi-cuadrado	10856,288 ^b
gl	13
Sig. asintótica	0,000

a. Variable de agrupación: STATIONNUM

b. 0 casillas (,0%) han esperado frecuencias menores que 5. La frecuencia mínima de casilla esperada es 2412,7.

Pruebas de homogeneidad de datos agrupados por meses

Se realizó la misma prueba para los datos agrupados por meses. Tanto la prueba de Kruskal Wallis como la de medias rechazan la hipótesis nula de homogeneidad y que las muestras agrupadas por meses se comporten según una misma distribución.

Tabla 36.

*Homogeneidad
Kruskal Wallis por
meses*

Estadísticos de prueba ^{a,b}		DEMANDA
-		
H de Kruskal-Wallis		930,143
Gl		7
Sig. Asintótica		0,000
a. Prueba de Kruskal Wallis		
b. Variable de agrupación: MES1NUM		

Tabla 37.

*Homogeneidades medias
por meses.*

Estadísticos de prueba ^a		DEMANDA
-		
N		67760
Mediana		228,00
Chi-cuadrado		795,277 ^b
gl		7
Sig. asintótica		0,000
a. Variable de agrupación: MES1NUM		
b. 0 casillas (0,0%) han esperado frecuencias menores que 5. La frecuencia mínima de casilla esperada es 3908,2.		

Pruebas de homogeneidad. Datos agrupados por día de la semana

Prueba para los datos agrupados por día de la semana: tanto la prueba de Kruskal Wallis como la de medias rechazan la hipótesis nula de homogeneidad y que las muestras agrupadas por día de la semana se comporten según una misma distribución.

Tabla 38.

*Homogeneidad
Kruskal Wallis por
día de la semana.*

Estadísticos de prueba ^{a,b}		DEMANDA
-		
H de Kruskal-Wallis		9646,417
Gl		6
Sig. asintótica		0,000
a. Prueba de Kruskal Wallis		
b. Variable de agrupación: DAYNUM		

Tabla 39. *Homogeneidades medias por día de la semana*

Estadísticos de prueba ^a		DEMANDA
-		
N		67760
Mediana		228,00
Chi-cuadrado		8933,224 ^b
Gl		6
Sig. Asintótica		0,000
a. Variable de agrupación: DAYNUM		
b. 0 casillas (0,0%) han esperado frecuencias menores que 5. La frecuencia mínima de casilla esperada es 4745,7.		

Pruebas de homogeneidad. Datos agrupados por horas

Se realizó la misma prueba para los datos agrupados por horas. Tanto la prueba de Kruskal Wallis como la de medias rechazan la hipótesis nula de homogeneidad y que las muestras agrupadas por meses se comporten según una misma distribución.

Tabla 40. *Homogeneidad Kruskal Wallis por horas*

Estadísticos de prueba^{a,b}

	DEMANDA
H de Kruskal-Wallis	27025,450
Gl	19
Sig. Asintótica	0,000

a. Prueba de Kruskal Wallis

b. Variable de agrupación:
HORASNUM

Tabla 41. *Homogeneidades medias por horas*

Estadísticos de prueba^a

	DEMANDA
N	67760
Mediana	228,00
Chi-cuadrado	16447,615 ^b
Gl	19
Sig. Asintótica	0,000

a. Variable de agrupación:
HORASNUM

b. 0 casillas (0,0%) han esperado frecuencias menores que 5. La frecuencia mínima de casilla esperada es 1688,9.

11.4.2. Pruebas de independencia

Buscan determinar si existe asociación entre dos o más variables de una misma muestra y cuál es la fuerza de esta asociación. Las pruebas escogidas para analizar la independencia de la muestra fueron:

Coefficiente de correlación de Pearson: Mide la relación lineal entre dos variables. Con una hipótesis nula el coeficiente es igual a 0. En caso contrario existe una dependencia de tipo lineal.

Coefficiente de correlación de Spearman: Es uno de los métodos no paramétricos más utilizados, el cual asigna un valor por rangos de datos y, cuando hay varios datos con el mismo valor, se asigna un promedio de los rangos correspondientes. Su hipótesis es similar al coeficiente de Pearson, donde la hipótesis nula es que el coeficiente tiene un valor de 0. De lo contrario, presenta una correlación acompañada de un signo que indica el sentido de la relación entre las variables

Coefficiente de Kendall: Asigna valores a rangos de los valores de variables, con la misma hipótesis nula de que no existe relación si el valor del coeficiente es 0.

Análisis de independencia mediante coeficientes (por estaciones)

Se realizan las pruebas anteriormente descritas a las variables con respecto a la demanda para identificar si existe relación entre ellas, omitiendo la variable mes debido a que no tiene relevancia para el estudio porque el número de muestras disponibles es limitado.

Tabla 42. Coeficiente de Pearson Demanda vs Estaciones

Correlaciones			
		DEMANDA	STATIONNUM
DEMANDA	Correlación de	1	-,274**
	Pearson		
	Sig. (bilateral)		0,000
	Suma de cuadrados y productos vectoriales	44344110321,173	-60621141,500
	Covarianza	654438,677	-894,658
	N	67760	67760
STATIONNUM	Correlación de	-,274**	1
	Pearson		
	Sig. (bilateral)	0,000	
	Suma de cuadrados y productos vectoriales	-60621141,500	1101100,000
	Covarianza	-894,658	16,250
	N	67760	67760

** La correlación es significativa en el nivel 0,01 (bilateral).

Tabla 43. *Coeficiente de Kendall Demanda vs Estaciones*

Correlaciones			
		DEMANDA	STATIONNUM
Tau_b de Kendall	DEMANDA	Coeficiente de correlación de 1,000	-,127**
		Sig. (bilateral)	0,000
		N	67760
	STATIONNUM	Coeficiente de correlación de -,127**	1,000
		Sig. (bilateral)	0,000
		N	67760
Rho de Spearman	DEMANDA	Coeficiente de correlación de 1,000	-,184**
		Sig. (bilateral)	0,000
		N	67760
	STATIONNUM	Coeficiente de correlación de -,184**	1,000
		Sig. (bilateral)	0,000
		N	67760

** . La correlación es significativa en el nivel 0,01 (bilateral).

Los tres coeficientes muestran que existe una relación inversamente proporcional entre la demanda y las estaciones, es decir, que donde las estaciones 1 es Portal Dorado y la 14 Universidades, según el segmento analizado, a medida que aumenta en número de estación en sentido occidente oriente la demanda disminuye levemente.

Para confirmar esta relación se realizó la prueba coeficiente de correlación parcial es decir manteniendo controladas otras variables que pudiesen afectar la medida, Las cuales comprueban que persiste la relación inversa significante.

Tabla 44. *Correlación parcial demanda vs estaciones controlando la categoría mes*

Correlaciones				
Variables de control			DEMANDA	STATIONNUM
MES1NUM	DEMANDA	Correlación	1,000	-0,274
		Significación (bilateral)		0,000
		gl	0	67757
	STATIONNUM	Correlación	-0,274	1,000
		Significación (bilateral)	0,000	
		gl	67757	0

Tabla 45. *Correlación parcial demanda vs estaciones controlando la categoría hora del día*

Correlaciones				
Variables de control			DEMANDA	STATIONNUM
HORASNUM	DEMANDA	Correlación	1,000	-0,274
		co+J6:Z14ntrolando día de la semana		0,000
		gl	0	67757
	STATIONNUM	Correlación	-0,274	1,000
		Significación (bilateral)	0,000	
		gl	67757	0

Tabla 46. *Correlación parcial demanda vs estaciones controlando la categoría día de la semana*

Correlaciones				
Variables de control			DEMANDA	STATIONNUM
DAYNUM	DEMANDA	Correlación	1,000	-0,277
		Significación (bilateral)		0,000
		gl	0	67757
	STATIONNUM	Correlación	-0,277	1,000
		Significación (bilateral)	0,000	
		gl	67757	0

Análisis independencia mediante coeficientes. (por día de la semana)

Se observa cómo existe una relación inversamente proporcional; en otras palabras, a medida que aumenta el día de la semana, empezando en lunes y terminando en domingo, la demanda decrece de forma general.

Tabla 47. *Coeficiente de Pearson: demanda vs. día de la semana*

-	-	DEMANDA	DAYNUM
DEMANDA	Correlación de Pearson	1	-,135**
	Sig. (bilateral)		0,000
	Suma de cuadrados y productos vectoriales	44344110321,173	-14679935,339
	Covarianza	654438,677	-216,649
	N	67760	67760
DAYNUM	Correlación de Pearson	-,135**	1
	Sig. (bilateral)	0,000	
	Suma de cuadrados y productos vectoriales	-14679935,339	268235,372
	Covarianza	-216,649	3,959
	N	67760	67760

** . La correlación es significativa en el nivel 0,01 (bilateral).

Tabla 48. *Coeficientes de Kendall y Spearman: demanda vs. estaciones*

		DEMANDA	DAYNUM
Tau_b de Kendall	de DEMANDA	Coeficiente de correlación	de 1,000
		Sig. (bilateral)	-,183**
		N	0,000
		67760	67760
	DAYNUM	Coeficiente de correlación	de -,183**
		Sig. (bilateral)	0,000
Rho de Spearman	de DEMANDA	Coeficiente de correlación	de 1,000
		Sig. (bilateral)	-,249**
		N	0,000
		67760	67760
	DAYNUM	Coeficiente de correlación	de -,249**
		Sig. (bilateral)	0,000
		N	67760

** . La correlación es significativa en el nivel 0,01 (bilateral).

De la misma forma que se realizó con la anterior, se analiza la independencia manteniendo controladas otras variables, evitando de esta forma sesgar el estudio hacia una variable en específico, prueba que se relaciona en las tablas 33, 34 y 35, donde se muestra que la relación persiste incluso controlando las otras variables estudiadas.

Tabla 49. *Correlación parcial de demanda vs. día de la semana (controlando mes)*

Variables de control			DEMANDA	DAYNUM
MES1NUM	DEMANDA	Correlación	1.000	-0.135
		Significación (bilateral)		0.000
		GI	0	67757
	DAYNUM	Correlación	-0.135	1.000
		Significación (bilateral)	0.000	
		GI	67757	0

Tabla 50. *Correlación parcial de demanda vs. día de la semana (controlando estaciones).*

Variables de control			DEMANDA	DAYNUM
STATIONNUM	DEMANDA	Correlación	1.000	-0.140
		Significación (bilateral)		0.000
		GI	0	67757
	DAYNUM	Correlación	-0.140	1.000
		Significación (bilateral)	0.000	
		GI	67757	0

Tabla 51. *Correlación parcial de demanda vs. día de la semana (controlando horas del día)*

Variables de control			DEMANDA	DAYNUM
HORASNUM	DEMANDA	Correlación	1.000	-0.135
		Significación (bilateral)		0.000
		GI	0	67757
	DAYNUM	Correlación	-0.135	1.000
		Significación (bilateral)	0.000	
		gl	67757	0

Análisis independencia mediante coeficientes (por horas)

Las pruebas muestran que existe una relación no significativa entre la demanda y la hora del día, debido a que el valor está muy cercano a 0 o muy cerca del cumplimiento de la hipótesis nula si se toma un rango de confianza del 0,05 o del 0,01.

Tabla 52. *Coeficiente de Pearson (demanda vs. Hora)*

-		DEMANDA	HORASNUM
DEMANDA	Correlación de Pearson	1	0,006
	Sig. (bilateral)		0,111
	Suma de cuadrados y productos vectoriales	44344110321,173	1937805,500
	Covarianza	654438,677	28,598
	N	67760	67760
HORASNUM	Correlación de Pearson	0,006	1
	Sig. (bilateral)	0,111	
	Suma de cuadrados y productos vectoriales	1937805,500	2253020,000
	Covarianza	28,598	33,250
	N	67760	67760

Por el contrario, el coeficiente de Kendall muestra que sí existe una relación, ya que el nivel de significancia supera el 0,01.

Tabla 53. *Coeficiente de Kendall y Spearman (demanda vs. hora)*

		DEMANDA	HORASNUM
Tau_b de Kendall	de DEMANDA	Coeficiente de correlación	de 1,000
		Sig. (bilateral)	,049**
		N	0,000
			67760
	HORASNUM	Coeficiente de correlación	de ,049**
Rho de Spearman		Sig. (bilateral)	0,000
		N	67760
	de DEMANDA	Coeficiente de correlación	de 1,000
		Sig. (bilateral)	,052**
		N	0,000
			67760
	HORASNUM	Coeficiente de correlación	de ,052**
		Sig. (bilateral)	0,000
		N	67760
			67760

** . La correlación es significativa en el nivel 0,01 (bilateral).

Debido a que existe discrepancia entre ambos métodos, se realizan las pruebas de correlación parcial, evidenciando que las variables afectan el coeficiente de Kendall y Spearman, bajo el criterio parcial se evidencia una correlación leve no significativa que puede verse en las tablas que aparecen a continuación.

Tabla 54. *Correlación parcial de la demanda vs. la hora controlando estaciones.*

Variables de control			DEMANDA	HORASNUM
STATIONNUM	DEMANDA	Correlación	1,000	0,006
		Significación (bilateral)		0,097
	HORASNUM	gl	0	67757
		Correlación	0,006	1,000
		Significación (bilateral)	0,097	
		gl	67757	0

Tabla 55. *Correlación parcial demanda vs. hora controlando fecha*

Variables de control			DEMANDA	HORASNUM
DATE	DEMANDA	Correlación	1,000	0,006
		Significación (bilateral)		0,110
	HORASNUM	gl	0	67757
		Correlación	0,006	1,000
		Significación (bilateral)	0,110	
		gl	67757	0

Tabla 56. *Correlación parcial de demanda vs. hora controlando día de la semana*

Variables de control			DEMANDA	HORASNUM
DAYNUM	DEMANDA	Correlación	1,000	0,006
		Significación (bilateral)		0,107
	HORASNUM	gl	0	67757
		Correlación	0,006	1,000
		Significación (bilateral)	0,107	
		gl	67757	0

Tabla 57. *Correlación parcial demanda vs. hora controlando mes*

Variables de control			DEMANDA	HORASNUM
MES1NUM	DEMANDA	Correlación	1,000	0,006
		Significación (bilateral)		0,110
		gl	0	67757
	HORASNUM	Correlación	0,006	1,000
		Significación (bilateral)	0,110	
		gl	67757	0

11.4.3. Clúster del comportamiento de la demanda

Una vez determinado el comportamiento de la serie en general, se detectó la necesidad de tratar los datos de otra forma, por lo cual se realizó una comparación de medias mediante la prueba Anova de un solo factor y un análisis PostHoc.

Agrupación por estaciones

La prueba expone que los datos agrupados por estaciones no se pueden tomar como un solo grupo, por lo tanto, debe tomar este conjunto como uno solo no representa eficazmente el sistema.

Tabla 58. ANOVA. Análisis por estaciones

DEMANDA						
	Suma de cuadrados	gl	Media cuadrática	F	Sig.	
Entre grupos	20616037145,423	13	1585849011,186	4527,756	0,000	
Dentro de grupos	23728073175,752	67746	350250,541			
Total	44344110321,175	67759				

Nota. La prueba rechaza la hipótesis de que la media de la cantidad de pasajeros se comporta de una forma similar en todas las estaciones. Con base en la anterior prueba se procede a realizar un análisis detallado de posibles grupos para el tratamiento de los datos.

Tabla 59. Análisis posthoc. Demanda agrupada por estaciones

			Subconjunto para alfa = 0.05								
TATIONNUM		N	1	2	3	4	5	6	7	8	9
HSD Tukey ^a	Plaza de la Democracia	4840	133,53								
	Quinta Paredes	4840	148,21	148,21							
	Centro Memoria	4840		182,65							
	Modelia	4840			239,55						
	Normandia	4840			264,04	264,04					
	Av. Rojas	4840				284,73					
	Gobernación	4840				285,13					
	Corferias	4840				287,11					
	CAN	4840					361,38				
	Ciudad Universitaria	4840					383,65	383,65			
	Salitre Greco	4840					397,33	397,33			
	El Tiempo	4840						409,21			
	Universidades	4840							661,95		
	Portal El Dorado	4840									2393,33
	Sig.		0,995	0,194	0,746	0,817	0,142	0,685	1,000	1,000	
Tukey B ^a	Plaza de la Democracia	4840	133,53								

	Quinta Paredes	4840	148,21								
	Centro Memoria	4840		182,65							
	Modelia	4840			239,55						
	Normandia	4840			264,04	264,04					
	Av. Rojas	4840				284,73					
	Gobernación	4840				285,13					
	Corferias	4840				287,11					
	CAN	4840					361,38				
	Ciudad Universitaria	4840					383,65	383,65			
	Salitre Greco	4840						397,33			
	El Tiempo	4840						409,21			
	Universidades	4840							661,95		
	Portal El Dorado	4840								2393,33	
Duncan ^a	Plaza de la Democracia	4840	133,53								
	Quinta Paredes	4840	148,21								
	Centro Memoria	4840		182,65							
	Modelia	4840			239,55						
	Normandia	4840				264,04					
	Av. Rojas	4840				284,73					
	Gobernación	4840				285,13					
	Corferias	4840				287,11					
	CAN	4840					361,38				
	Ciudad Universitaria	4840					383,65	383,65			
	Salitre Greco	4840						397,33	397,33		
	El Tiempo	4840						409,21			
	Universidades	4840							661,95		
	Portal El Dorado	4840								2393,33	
	Sig.		0,222	1,000	1,000	0,080	0,064	0,255	0,324	1,000	1,000
Scheffe ^a	Plaza de la Democracia	4840	133,53								
	Quinta Paredes	4840	148,21								
	Centro Memoria	4840	182,65								
	Modelia	4840		239,55							
	Normandia	4840		264,04							
	Av. Rojas	4840		284,73							
	Gobernación	4840		285,13							
	Corferias	4840		287,11							
	CAN	4840			361,38						
	Ciudad Universitaria	4840			383,65						
	Salitre Greco	4840				397,33					
	El Tiempo	4840				409,21					
	Universidades	4840					661,95				
	Portal El Dorado	4840						2393,33			
	Sig.		0,215	0,270	0,260	1,000	1,000				
Waller-Duncan ^{a,b}	Plaza de la Democracia	4840	133,53								
	Quinta Paredes	4840	148,21								

Centro Memoria	4840	182,65			
Modelia	4840	239,55			
Normandía	4840	264,04	264,04		
Av. Rojas	4840		284,73		
Gobernación	4840		285,13		
Corferias	4840		287,11		
CAN	4840			361,38	
Ciudad Universitaria	4840			383,65	383,65
Salitre Greco	4840			397,33	
El Tiempo	4840			409,21	
Universidades	4840				661,95
Portal El Dorado	4840				2393,33

Se visualizan las medias para los grupos en los subconjuntos homogéneos.

a. Utiliza el tamaño de la muestra de la media armónica = 4840,000.

b. Razón de gravedad de error de Tipo 1/Tipo 2 = 100.

Escogiendo para tratar los datos los clústeres.

Tabla 60. *Clúster estación*

	1	2	3	4	5
Plaza de la Democracia	Modelia	CAN	Universidades	Portal El Dorado	
Quinta Paredes	Normandía	Ciudad Universitaria			
Centro Memoria	Av. Rojas	Salitre Greco			
	Gobernación	El Tiempo			

Nota. Grupos de estaciones que tienen comportamientos de medias similares.

Agrupación por día

La prueba Anova relacionada muestra que no sería recomendable analizar la demanda como un mismo grupo, bajo el criterio de día de la semana, por tal motivo se realiza una propuesta para el tratamiento de los datos.

Tabla 61. *ANOVA. Análisis por día de la semana*

DEMANDA					
	Suma de cuadrados	Gl	Media cuadrática	F	Sig.
Entre grupos	1648210627,832	6	274701771,305	435,917	0,000
Dentro de grupos	42695899693,343	67753	630169,877		
Total	44344110321,175	67759			

Nota. La prueba rechaza la hipótesis de que la media de la cantidad de pasajeros se comporta de forma similar en todos los días de la semana, por lo cual se busca una forma de tratar los datos basada en comportamientos similares.

Tabla 62. *Análisis PostHoc de la demanda, agrupada por días de la semana*

			Subconjunto para alfa = 0.05				
DAYNUM		N	1	2	3	4	5
HSD Tukey ^{a,b}	Sunday	9520	141,97				
	Saturday	9800		314,82			
	Monday	9520			472,75		
	Thursday	9800				553,21	
	Friday	9800				569,56	
	Wednesday	9800				576,65	
	Tuesday	9520				581,75	
	Sig.		1,000	1,000	1,000	0,159	
Tukey B ^{a,b}	Sunday	9520	141,97				
	Saturday	9800		314,82			
	Monday	9520			472,75		
	Thursday	9800				553,21	
	Friday	9800				569,56	
	Wednesday	9800				576,65	
	Tuesday	9520				581,75	
	Sig.		1,000	1,000	1,000	0,159	
Duncan ^{a,b}	Sunday	9520	141,97				
	Saturday	9800		314,82			
	Monday	9520			472,75		
	Thursday	9800				553,21	
	Friday	9800				569,56	569,56
	Wednesday	9800				576,65	576,65
	Tuesday	9520				581,75	581,75
	Sig.		1,000	1,000	1,000	0,051	0,317
Scheffe ^{a,b}	Sunday	9520	141,97				
	Saturday	9800		314,82			
	Monday	9520			472,75		
	Thursday	9800				553,21	
	Friday	9800				569,56	
	Wednesday	9800				576,65	
	Tuesday	9520				581,75	
	Sig.		1,000	1,000	1,000	0,396	
Waller-Duncan ^{a,b,c}	Sunday	9520	141,97				
	Saturday	9800		314,82			
	Monday	9520			472,75		
	Thursday	9800				553,21	
	Friday	9800				569,56	
	Wednesday	9800				576,65	
	Tuesday	9520				581,75	
	Sig.		1,000	1,000	1,000	0,396	

Se visualizan las medias para los grupos en los subconjuntos homogéneos.

a. Utiliza el tamaño de la muestra de la media armónica = 9678,008.

b. Los tamaños de grupo no son iguales. Se utiliza la media armónica de los tamaños de grupo. Los niveles de error de tipo I no están garantizados.

c. Razón de gravedad de error de Tipo 1/Tipo 2 = 100.

Eligiendo el menor número de grupos y buscando menor complejidad para futuros análisis, se elige la agrupación relacionada en la Tabla 47.

Tabla 63. *Clauster Día*

	1	2	3	4
Domingo	Sábado	Lunes	Martes Miércoles Jueves Viernes	

Nota. Los grupos aluden al comportamiento similar del valor medio de la cantidad de pasajeros con base en tres criterios: HSD Tukey, Tukey B^a, Scheffe y Waller-Duncan^a.

Agrupación por hora del día

Tabla 64. ANOVA – *Análisis por hora*

DEMANDA					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	4459576487,462	19	234714551,972	398,640	0,000
Dentro de grupos	39884533833,713	67740	588788,512	-	-
Total	44344110321,175	67759	-	-	-

Nota. La prueba rechaza la hipótesis nula es decir que el conjunto no se recomienda tratarlo como uno solo.

Debido al tamaño de las prueba Post Hoc del análisis por hora, solo se presenta a continuación el criterio de HSD Tukeya, parte del anexo Análisis estadístico.xls.

Tabla 65. *Análisis PostHoc de la demanda, agrupada por horas de la semana*

Subconjunto para alfa = 0.05														
HORAS	NUM	N	1	2	3	4	5	6	7	8	9	10	11	12
HSD Tukey ^a	11:00 pm	3388	4,20	-	-	-	-	-	-	-	-	-	-	-
	4:00 am	3388	20,84	-	-	-	-	-	-	-	-	-	-	-
	10:00 pm	3388	-	141,88	-	-	-	-	-	-	-	-	-	-
	5:00 am	3388	-	-	234,08	-	-	-	-	-	-	-	-	-
	9:00 p. m.	3388	-	-	-	306,49	-	-	-	-	-	-	-	-
	8:00 pm	3388	-	-	-	-	375,48	-	-	-	-	-	-	-
	10:00 am	3388	-	-	-	-	403,73	403,73	-	-	-	-	-	-
	11:00 am	3388	-	-	-	-	416,28	416,28	-	-	-	-	-	-

9:00 am	3388	-	-	-	-	417,89	417,89	-	-	-	-	-	-	-
12:00 pm	3388	-	-	-	-	-	464,25	464,25	-	-	-	-	-	-
1:00 pm	3388	-	-	-	-	-	465,51	465,51	-	-	-	-	-	-
8:00 am	3388	-	-	-	-	-	-	485,37	-	-	-	-	-	-
2:00 pm	3388	-	-	-	-	-	-	497,86	-	-	-	-	-	-
7:00 pm	3388	-	-	-	-	-	-	507,14	-	-	-	-	-	-
3:00 pm	3388	-	-	-	-	-	-	510,33	-	-	-	-	-	-
6:00 am	3388	-	-	-	-	-	-	-	587,05	-	-	-	-	-
7:00 am	3388	-	-	-	-	-	-	-	-	654,09	-	-	-	-
4:00 pm	3388	-	-	-	-	-	-	-	-	-	721,39	-	-	-
6:00 pm	3388	-	-	-	-	-	-	-	-	-	-	844,14	-	-
5:00 pm	3388	-	-	-	-	-	-	-	-	-	-	-	1130,30	-
Sig.		1,000	1,000	1,000	1,000	0,752	0,102	0,606	1,000	1,000	1,000	1,000	1,000	-

Nota. La agrupación horaria corresponde a una agrupación general la cual tiene en cuenta todas las muestras.

Tabla 66. *Clúster por mes*

1	2	3	4	5	6	7	8	9	10	11	12	13
11:00 pm	10:00 pm	5:00 am	9:00 p. m.	8:00 pm	12:00 pm	8:00 am	6:00 am	7:00 a. m.	4:00 pm	6:00 pm	5:00 pm	
4:00 am				10:00 am	1:00 pm	2:00 pm						
				11:00 am		7:00 pm						
				9:00 am		3:00 pm						

Agrupación por mes

Con base en la prueba Anova, que el set de datos no puede comportarse como un solo grupo.

Tabla 67. *ANOVA – Análisis por mes*

DEMANDA						
Entre grupos	Suma de cuadrados	Gl	Media cuadrática	F	Sig.	
	272450282,093	7	38921468,870	59,835	0,000	
Dentro de grupos	44071660039,082	67752	650485,005			
Total	44344110321,175	67759				

Nota. La prueba rechaza la hipótesis de que la media de la cantidad de pasajeros se comporta de una forma similar en todos los meses. Por este motivo se amplía la prueba de la siguiente forma (encontrando cuatro agrupaciones que tienen un comportamiento similar en valores de la media).

Tabla 68. *Análisis posthoc. Demanda agrupada por meses*

		Subconjunto para alfa = 0.05				
MES1NUM	N	1	2	3	4	5
HSD	June	8400	324,30			
Tukey ^{a,b}	July	8680		408,08		
	April	8400		433,05		
	September	8400			484,44	
	May	8680			489,03	
	August	8680			490,10	
	March	8680			520,53	520,53
	February	7840			528,03	
	Sig.		1,000	0,472	0,071	0,999
Tukey B ^{a,b}	June	8400	324,30			
	July	8680		408,08		
	April	8400		433,05		
	September	8400			484,44	
	May	8680			489,03	489,03
	August	8680			490,10	490,10
	March	8680			520,53	520,53
	February	7840				528,03
Duncan ^{a,b}	June	8400	324,30			
	July	8680		408,08		
	April	8400			433,05	
	September	8400				484,44
	May	8680				489,03
	August	8680				490,10
	March	8680				520,53
	February	7840				528,03
Scheffe ^{a,b}	Sig.		1,000	1,000	1,000	0,671
	June	8400	324,30			0,545
	July	8680		408,08		
	April	8400		433,05		
	September	8400			484,44	
	May	8680			489,03	
	August	8680			490,10	
	March	8680			520,53	
Waller-Duncan ^{a,b,c}	February	7840			528,03	
	Sig.		1,000	0,773	0,089	
	June	8400	324,30			
	July	8680		408,08		
	April	8400		433,05		
	September	8400			484,44	
	May	8680			489,03	489,03
	August	8680			490,10	490,10
	March	8680			520,53	520,53
	February	7840				528,03

Se visualizan las medias para los grupos en los subconjuntos homogéneos.

a. Utiliza el tamaño de la muestra de la media armónica = 8460,923.

b. Los tamaños de grupo no son iguales. Se utiliza la media armónica de los tamaños de grupo. Los niveles de error de tipo I no están garantizados.

c. Razón de gravedad de error de Tipo 1/Tipo 2 = 100.

Nota. En la Tabla 43 se muestran las diferentes posibilidades de agrupar el set de datos según diferentes criterios mes.

Tabla 69. *Clúster por mes*

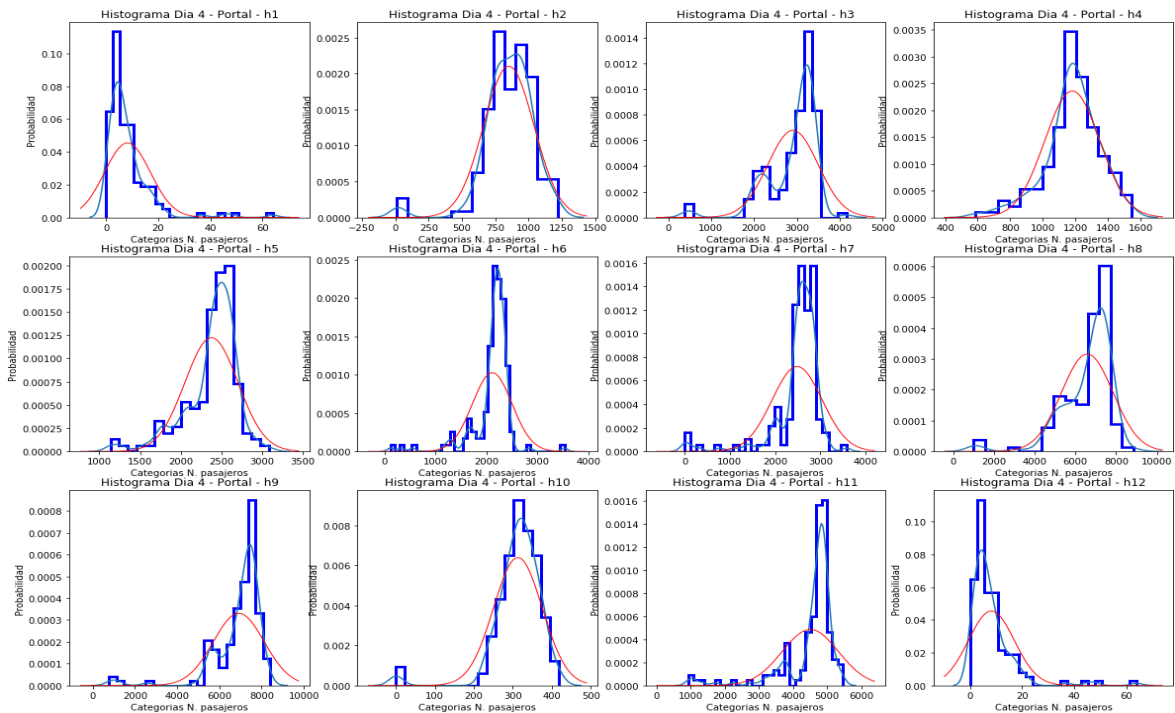
1	2	3
Junio	Abril Julio	Febrero Marzo Mayo Agosto Septiembre

Nota. Escogido por el criterio de Scheffe, bajo el principio de la mayor simplicidad, ya que brinda menor complejidad para el tratamiento de los datos.

11.4.4. Resultados de la estructuración

El análisis de la información de la cantidad de pasajeros presentes en las estaciones del segmento analizado gracias a la metodología descrita anteriormente pasa de 67760 muestra por hora a 9.598 registros diarios que pueden resumirse bajo una muestra consolidada de 690 registros, que permiten describir el comportamiento diario de la red en el segmento. Posibilita la comparación de un dato mediante 20 variantes distintas que corresponden a un día y estación específica, que a su vez pueden ser segmentados según las doce posibles agrupaciones horarias, en total 240 alternativas, al tener en cuenta la agrupación de horas propuesta por prueba Poshoc de hora, o como una alternativa más detallada un total de 400 alternativas, sin tener en cuenta la agrupación. Agrupaciones que tienen la capacidad de describir el sistema en cada punto y en cada posible momento, dentro del periodo comprendido entre febrero y septiembre del año 2017. Desde otra perspectiva el análisis permite evidenciar una descripción mucho más detallada del comportamiento de la demanda cuando se segmenta las muestra a un nivel horario, por ejemplo, como resultado de las agrupaciones que tienen en cuenta los 20 clúster y las 20 horas agrupadas en 12 grupos; se observa una menor variabilidad, y un ajuste mucho más cercano a una tendencia normal en el estudio de una estación como se muestra a continuación en las gráficas del portal el dorado para los martes, miércoles, jueves y viernes en una hora específica.

Figura 14. *Histogramas comportamiento de las entradas de pasajeros en la estación portal el dorado en el tipo de día 4. Agrupación criterios de día, estación y hora.*



12. Etapa 4. Propuesta cuantitativa de respuesta a la demanda.

De acuerdo con las características del sistema y las metodologías analizadas, se propone una propuesta, basada en las metodologías analizadas, la cual permite describir variables dependientes de la demanda, como el número de vehículos, la velocidad promedio en la vía, los tiempos de viaje para los usuarios y el costo para el operador.

12.1. Variables

El comportamiento de la trocal es descrito por las variables a continuación, variables que se desarrollan para un ejemplo específico y es la base del aplicativo.

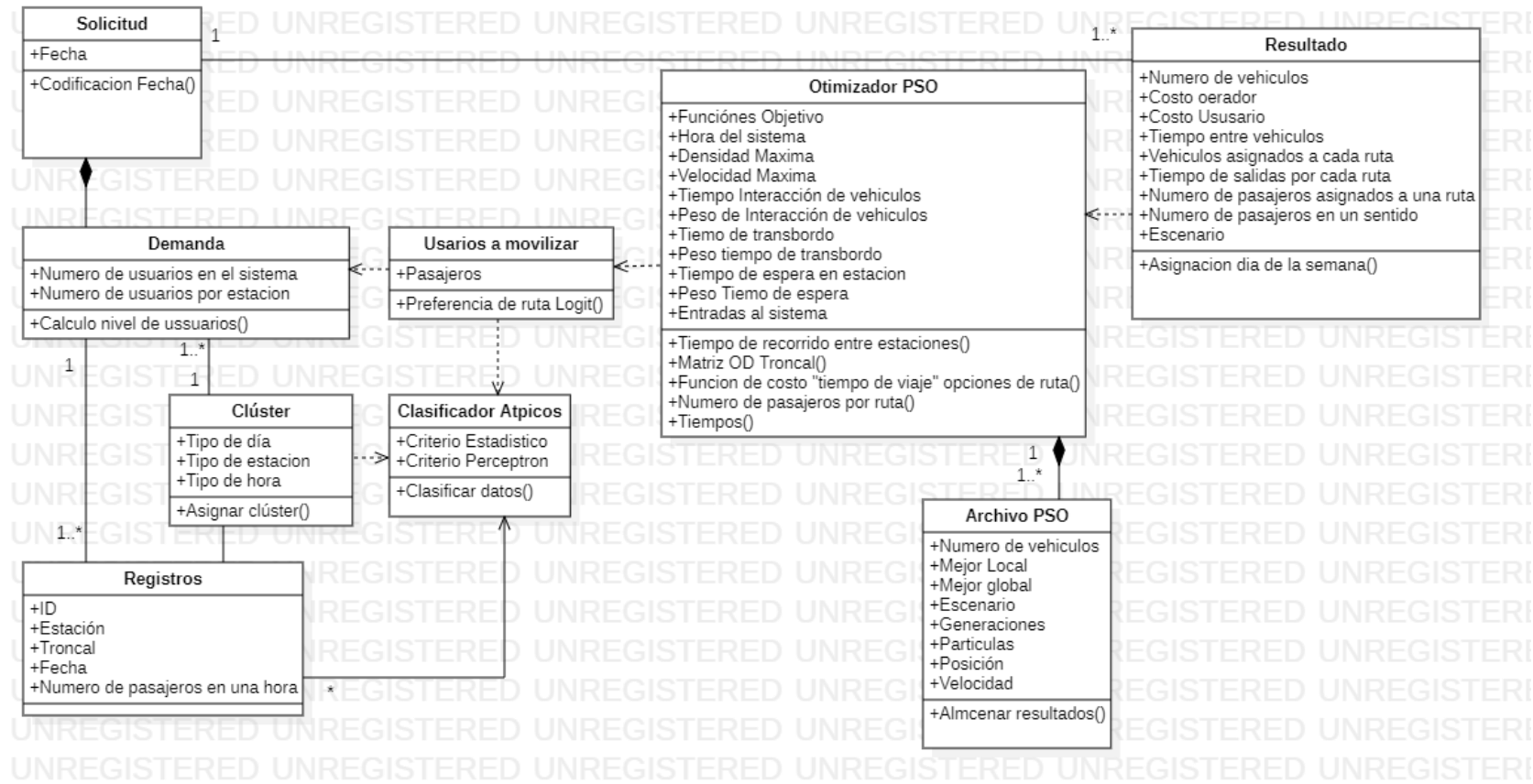
Tabla 70. Variables del modelo.

	Símbolo	Nombre	Unidades
Modelo analítico	P	Demanda de pasajeros	Numero entero
	Y_D	Clúster día	Numero entero
	Y_S	Clúster día de la estación	Numero entero
	Y_W	Clúster día de la semana	
	D	Fecha	dd/mm/aaaa
	H_t	Hora	h
	t_t	Tiempo de transbordo o transferencia	h
	t_{te}	Tiempo de espera	h
	t_a	Tiempo promedio de abordaje	h
	w_t	Peso tiempo de transferencia	
	w_e	Peso tiempo de espera	
	w_a	Peso tiempo de abordaje	
	tr_{ij}	Tiempo de recorrido dentro de vehículo	h
	L	Distancia entre estaciones	km
	V_p	Velocidad promedio de la vía.	km/h
	T	Tiempo total de viaje	h
	O	Numero de transferencias	Numero entero
	R	Ruta o servicio	Alfanumérico
	e	Preferencia de elección según modelo Logit.	% Probabilidad
	N_b	Cantidad de vehículos tipo bus	Numero entero

	N_{bex}	Cantidad de vehículos tipo bus necesarios no programados	<i>Numero entero</i>
	C	Capacidad de vehículo	<i>Numero entero</i>
	S	Número de estaciones	<i>Numero entero</i>
	I_t	Numero de intervalos de tiempo analizados	<i>Numero entero</i>
	Q_v	Flujo de la vía	<i>Km/h</i>
	k_v	Densidad de la vía	<i>Veh/Km</i>
	V_{max}	Velocidad Máxima	<i>Km/h</i>
	K_{max}	Densidad Máxima	<i>Km/ Km2</i>
	Fo	Factor de ocupación	<i>[]</i>
PSO Optimización.	Sw	Tamaño del enjambre PSO	<i>Partícula</i>
	Gn	Numero de generaciones o iteraciones del enjambre	<i>Numero entero</i>
	W	Peso de inercia	
	F	Factor de aprendizaje	
	a_i	Aleatorios	<i>Numero</i>
	XS_t	Posición actual de una partícula.	
	VS_t	Velocidad de la partícula	
	MS_l	Mejor local de la generación de partículas iteración PSO	
	MS_G	Mejor global de la optimización PSO	

Nota. La tabla relaciona las variables que describen la operación de la troncal estudiada.

Figura 15 Diagrama metodología propuesta aplicativo.



Nota. Diagrama de clases que explica cómo se evalúa el rendimiento de la troncal, autoría propia.

12.1.1. Matriz de ruta

Una matriz relata en qué nodos o estaciones se detiene un vehículo y en cuáles no lo hace. En ella se puede obtener el número de rutas disponibles en cada estación y el número de paradas que realiza cada una de ellas.

Tabla 71. Paradas por servicio

Estacion	K10/L10	K16/B16	K23/B23	k42/G42	K43/G43	K54/H54	K86/M86	C73	1	Numero de rutas
Portal El Dorado	1	1	1	0	1	1	1	0	1	7
Modelia	1	1	0	0	1	1	0	0	1	5
Normandia	1	0	0	0	1	0	0	0	1	3
Av. Rojas	0	1	1	0	0	1	1	0	1	5
El Tiempo	0	1	1	0	1	1	1	0	1	6
Salitre Greco	1	1	1	0	1	1	1	0	1	7
CAN	1	1	1	0	0	1	1	0	1	6
Gobernación	1	0	1	0	1	0	1	0	1	5
Quinta Paredes	0	1	0	0	1	1	1	0	1	5
Corferias	0	1	0	1	1	1	1	0	1	6
Ciudad Universitaria	1	0	1	0	0	1	0	1	1	5
Plaza de la Democracia	0	0	1	0	0	1	1	0	1	4
Centro Memoria	1	0	0	0	0	0	0	0	1	2
Numero de paradas	8	8	8	1	8	10	9	1	13	

Nota. Adaptación de los diagramas obtenidos del sistema de Transmilenio S.A.

12.1.2. Matriz de distancias

Relaciona la distancia que existe entre un punto y otro del sistema. Las distancias fueron tomadas utilizando los servicios de mapas online de ArcGis.

Tabla 72. Distancias de la red.

Nodo	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	0,88	1,69	2,62	3,25	4,07	4,59	5,11	5,81	6,38	7,16	7,81	8,56
2	0,88	0	0,81	1,74	2,37	3,19	3,71	4,23	4,93	5,5	6,28	6,93	7,68
3	1,69	0,81	0	0,93	1,56	2,38	2,9	3,42	4,12	4,69	5,47	6,12	6,87
4	2,62	1,74	0,93	0	0,63	1,45	1,97	2,49	3,19	3,76	4,54	5,19	5,94
5	3,25	2,37	1,56	0,63	0	0,82	1,34	1,86	2,56	3,13	3,91	4,56	5,31
6	4,07	3,19	2,38	1,45	0,82	0	0,52	1,04	1,74	2,31	3,09	3,74	4,49
7	4,59	3,71	2,9	1,97	1,34	0,52	0	0,52	1,22	1,79	2,57	3,22	3,97
8	5,11	4,23	3,42	2,49	1,86	1,04	0,52	0	0,7	1,27	2,05	2,7	3,45
9	5,81	4,93	4,12	3,19	2,56	1,74	1,22	0,7	0	0,57	1,35	2	2,75
10	6,38	5,5	4,69	3,76	3,13	2,31	1,79	1,27	0,57	0	0,78	1,43	2,18
11	7,16	6,28	5,47	4,54	3,91	3,09	2,57	2,05	1,35	0,78	0	0,65	1,4
12	7,81	6,93	6,12	5,19	4,56	3,74	3,22	2,7	2	1,43	0,65	0	0,75
13	8,56	7,68	6,87	5,94	5,31	4,49	3,97	3,45	2,75	2,18	1,4	0,75	0

Nota. Distancias tomadas en kilómetros: son el acumulado de las distancias de los nodos necesarios recorridos para movilizarse desde un origen hasta un destino.

12.1.3. Matriz de tiempos

De acuerdo con la distancia y la velocidad de los vehículos se calcula cuánto tiempo demoraría un vehículo en recorrer el segmento comprendido entre dos nodos, y a su vez calcular cuánto demora un vehículo en recorrer la red en un sentido, a una velocidad de 60 kilómetros por hora, en la siguiente matriz se relaciona un ejemplo que comprende tiempos multiplicados por una variable que representa el efecto de los semáforos y la interacción entre vehículos, entre otras variables no conocidas. En este caso el valor de la variable es 2.

$$\text{Ecuación 7 } T_{\text{Recorrido}} = \frac{\Delta \text{Distancia}}{\text{Velocidad de un vehiculo}} * \text{Variable}$$

Tabla 73. Tiempos de viaje de un vehículo a 60 km/h

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0,00	1,76	3,38	5,24	6,50	8,14	9,18	10,22	11,62	12,76	14,32	15,62	17,12
2	1,76	0,00	1,62	3,48	4,74	6,38	7,42	8,46	9,86	11,00	12,56	13,86	15,36
3	3,38	1,62	0,00	1,86	3,12	4,76	5,80	6,84	8,24	9,38	10,94	12,24	13,74
4	5,24	3,48	1,86	0,00	1,26	2,90	3,94	4,98	6,38	7,52	9,08	10,38	11,88
5	6,50	4,74	3,12	1,26	0,00	1,64	2,68	3,72	5,12	6,26	7,82	9,12	10,62
6	8,14	6,38	4,76	2,90	1,64	0,00	1,04	2,08	3,48	4,62	6,18	7,48	8,98
7	9,18	7,42	5,80	3,94	2,68	1,04	0,00	1,04	2,44	3,58	5,14	6,44	7,94
8	10,22	8,46	6,84	4,98	3,72	2,08	1,04	0,00	1,40	2,54	4,10	5,40	6,90
9	11,62	9,86	8,24	6,38	5,12	3,48	2,44	1,40	0,00	1,14	2,70	4,00	5,50
10	12,76	11,00	9,38	7,52	6,26	4,62	3,58	2,54	1,14	0,00	1,56	2,86	4,36
11	14,32	12,56	10,94	9,08	7,82	6,18	5,14	4,10	2,70	1,56	0,00	1,30	2,80
12	15,62	13,86	12,24	10,38	9,12	7,48	6,44	5,40	4,00	2,86	1,30	0,00	1,50
13	17,12	15,36	13,74	11,88	10,62	8,98	7,94	6,90	5,50	4,36	2,80	1,50	0,00
14	18,55	16,79	15,17	13,31	12,05	10,41	9,37	8,33	6,93	5,79	4,23	2,93	1,43

Nota. Tiempos de viaje en minutos calculados como la distancia dividida la velocidad máxima de referencia del sistema 60 km/h.

La anterior matriz ofrece una descripción superficial del comportamiento de los vehículos, ya que tiene en cuenta una velocidad constante y una variable que representa los tiempos no estimados en el estudio. El estimado del recorrido desde la estación Portal Dorado hasta la estación Centro de Memoria es de 17,12 minutos, a una velocidad constante de 60 km/hora.

12.1.4. Matriz de origen destino:

Consiste en una distribución uniforme de las entradas de pasajeros hacia los posibles 14 destinos en una hora específica, donde los extremos acumulan gran cantidad de pasajeros en origen y en destino; el nodo 14 es el extremo oriental de la red, que a su vez representa las estaciones no comprendidas dentro del segmento estudiado.

En este paso es necesario escoger una hora del día sobre la cual trabajar, como muestra a continuación el ejemplo a las 7:00 a.m.

Tabla 74. *Demanda del modelo a las 7:00 a.m.*

Numero	Estación	Entradas al sistema	Demanda Hora 7:00 a.m.	Pasajeros 7:00 en destino
1	Portal El Dorado	52147,60	3712	185,60
2	Modelia	5948,20	423	21,17
3	Normandía	5948,20	423	21,17
4	Av. Rojas	5948,20	423	21,17
5	El Tiempo	9011,60	641	32,07
6	Salitre Greco	9011,60	641	32,07
7	CAN	9011,60	641	32,07
8	Gobernación	5948,20	423	21,17
9	Quinta Paredes	3380,00	241	12,03
10	Corferias	5948,20	423	21,17
11	Ciudad Universitaria	9011,60	641	32,07
12	Plaza de la Democracia	3380,00	241	12,03
13	Centro Memoria	3380,00	241	12,03

Nota. Las entradas al sistema corresponden al pronóstico con el método suavización exponencial. Adicionalmente cada hora del día está relacionada con un porcentaje de la demanda total del día, en el caso de las 7:00 a.m. es el 7.12%

Tabla 75. *Matriz OD Red (número de pasajeros)*

Nota. La matriz de OD de la red, presenta los pasajeros que deben dejarse en una determinada estación. Muestra también la relación nula que existe entre un mismo nodo.

Por otra parte, permite saber cuántos pasajeros debe recoger los diferentes servicios, calculado como la suma horizontal de los valores en cada nodo. Como ejemplo resaltado en verde en la estación 12 se deberá tener una combinación de servicios que permitan recoger 184 pasajeros durante la hora.

12.1.5. Elección del usuario

Cuando un pasajero elige la ruta a tomar lo hace basado en distintos criterios, para este estudio se asume que un usuario decide qué servicio abordar según la utilidad que representa para cada trayecto; en otras palabras, cuando un usuario tiene la posibilidad de escoger entre distintos servicios que lo llevan a su destino, elige el que sea más rápido con preferencia de los viajes directos.

Medir dicha utilidad en este estudio incluye: el tiempo de espera de un vehículo, el tiempo de abordaje y el tiempo de transbordo. La elección se realiza mediante un modelo Logit, el cual ha presentado a la literatura una alternativa para representar, según una función de utilidad, la decisión entre una o muchas alternativas. La técnica apropiada consiste en un modelo Logit Multinomial, que compara las diferentes alternativas al mismo tiempo y asigna la probabilidad de ser escogida una de ellas.

La decisión del usuario será entonces representada por:

$$\text{Ecuación 8 Probabilidad de decision} = \frac{e^u}{\sum_{i=1}^n (e^{u_i})}$$

Donde el u será la suma de los distintos tiempos implícitos en un trayecto.

$$\text{Ecuación 9 } U = T_{Abordaje} * P + w_1 * T_{Espera} + w_2 * T_{Trasbordo} * P_{Trasbordo}$$

Donde w_1 y w_2 son los pesos que representan que tan importante es para el usuario el tiempo de espera y el tiempo de hacer un trasbordo respectivamente.

Obteniendo la matriz de utilidad para el usuario, en minutos.

Matriz que debe ser comparada con respecto otras matrices, utilizar la metodología Logit. se permite así comparar las 9 alternativas que tiene un usuario al mismo tiempo, Asignando una mayor probabilidad a las rutas que representan mayor utilidad para el usuario. Resaltado con amarillo las estaciones en las que ara la ruta y en un color más claro las que no.

Tabla 76. *Matriz de Utilidad ruta k10/L10*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	-0,1	-0,2	-0,3	-0,4	-0,4	-0,4	-0,5	-0,5	-0,6	-0,6	-0,7	-0,7	-0,7	-0,8
2	-0,2	-0,1	-0,2	-0,3	-0,3	-0,3	-0,4	-0,5	-0,6	-0,6	-0,6	-0,7	-0,7	-0,7
3	-0,3	-0,2	-0,1	-0,2	-0,3	-0,3	-0,3	-0,4	-0,5	-0,5	-0,5	-0,6	-0,6	-0,7
4	-0,4	-0,3	-0,3	-0,2	-0,2	-0,2	-0,3	-0,3	-0,4	-0,5	-0,5	-0,6	-0,6	-0,6
5	-0,4	-0,4	-0,3	-0,3	-0,2	-0,2	-0,3	-0,3	-0,4	-0,4	-0,4	-0,5	-0,5	-0,6
6	-0,4	-0,4	-0,3	-0,3	-0,2	-0,1	-0,2	-0,2	-0,3	-0,4	-0,4	0,0	-0,5	-0,5
7	-0,5	-0,4	-0,4	-0,3	-0,3	-0,2	-0,1	-0,2	-0,3	-0,3	-0,3	-0,4	-0,4	-0,5
8	-0,6	-0,5	-0,4	-0,4	-0,3	-0,3	-0,2	-0,1	-0,2	-0,3	-0,3	-0,4	-0,4	-0,4
9	-0,6	-0,6	-0,5	-0,5	-0,4	-0,4	-0,3	-0,3	-0,2	-0,2	-0,2	-0,3	-0,3	-0,4
10	-0,7	-0,6	-0,5	-0,5	-0,5	-0,4	-0,3	-0,3	-0,2	-0,2	-0,2	-0,3	-0,3	-0,4
11	-0,7	-0,6	-0,5	-0,5	-0,5	-0,4	-0,3	-0,3	-0,3	-0,2	-0,1	-0,2	-0,2	-0,3
12	-0,8	-0,7	-0,6	-0,6	-0,6	0,0	-0,4	-0,4	-0,3	-0,3	-0,3	-0,2	-0,2	-0,3
13	-0,8	-0,7	-0,6	-0,6	-0,6	-0,5	-0,4	-0,4	-0,3	-0,3	-0,3	-0,2	-0,1	-0,2
14	-0,8	-0,8	-0,7	-0,7	-0,6	-0,6	-0,5	-0,4	-0,4	-0,4	-0,3	-0,3	-0,2	-0,1

Nota. Tabla expresada en horas que incluye la acumulación de tiempo que representa tomar la opción en un recorrido dentro de la red. Propia autoría.

Asignación de la demanda a cada ruta

Conociendo la voluntad generalizada del usuario, se distribuye la demanda total en los siete servicios distintos, Las rutas C73 y K42/G2 , Tienen una costo muy alto para el usuario debido a que no paran en la mayoría de la estaciones en la sección estudiada por lo cual no son agregadas dentro de la distribución.

Tabla 77. *Distribución de la demanda.*

DEMANDA	11.937,00	100%	(Pasajeros)
K10/L10	1.753,00	14,69%	(Pasajeros)
K16/B16	1.709,00	14,32%	(Pasajeros)
K23/B23	1.743,00	14,60%	(Pasajeros)
K43/G43	1.720,00	14,41%	(Pasajeros)
K54/H54	1.689,00	14,15%	(Pasajeros)
K86/M86	1.709,00	14,32%	(Pasajeros)
Ruta facil	1.614,00	13,52%	(Pasajeros)

Nota. Distribución de la demanda de viajes en número de pasajeros.

Gracias a lo anterior se obtiene un matriz origen de destino individual. Dicha matriz se convierte en una herramienta que permite conocer cuántos pasajeros abordarán en una estación en cada recorrido y evaluar el tiempo que gastan los usuarios en general en usar el servicio.

Tabla 78. Matriz OD Ruta k10/L10

1752	OD Ruta k10/L10														968,12
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	0	27	27	25	26	27	27	27	26	27	28	27	29	221	544,57
2	14	0	3	3	3	3	3	3	3	3	3	3	3	14	49,92
3	14	3	0	3	3	3	3	3	3	3	3	3	4	15	47,61
4	12	3	3	0	3	3	3	3	3	3	3	3	3	14	43,80
5	20	5	5	5	0	5	5	5	5	5	5	5	5	21	59,17
6	21	5	5	5	5	0	5	5	5	5	5	7	5	21	56,45
7	21	5	5	5	5	5	0	5	5	5	5	5	5	21	50,28
8	14	3	3	3	3	3	3	0	3	3	3	3	3	14	30,25
9	8	2	2	2	2	2	2	2	0	2	2	2	2	8	15,38
10	14	3	3	3	3	3	3	3	3	0	3	3	3	13	22,93
11	22	5	5	5	5	5	5	5	5	5	0	5	5	21	30,54
12	8	2	2	2	2	3	2	2	2	2	2	0	2	8	9,46
13	8	2	2	2	2	2	2	2	2	2	2	2	0	8	7,77
14	161	20	21	20	20	20	20	20	19	19	20	19	19	0	
784	336	58	56	52	46	42	37	34	31	27	23	21	19		

Nota. Tabla en número de pasajeros calculado a una velocidad de la vía de 80 km/h, en azul relatados la cantidad de pasajeros que se deslazan en ambos sentidos.

Asignación de flota

Como punto de referencia, se debe satisfacer un número de puestos disponibles igual a la cantidad de pasajeros que entran al sistema, pero la cantidad de vehículos que realizaran una ruta en específico depende del volumen que requiere la ruta, por lo cual, a menor demanda en una ruta, menor será la cantidad de vehículos que requiere.

Como criterio para tener en cuenta, la capacidad de los vehículos varía dependiendo el factor de carga el cual se estipula según los objetivos del operador, en ese trabajo se maneja un factor de carga del 1 que corresponde a una ocupación de diseño. El resultado de este criterio permite de ser necesario cambiar este criterio donde los vehículos con una capacidad de 150 pasajeros aproximadamente podrían movilizar 225 pasajeros con un criterio 1.5, un servicio menos cómodo.

En el caso que los vehículos asignados para la configuración de la red y teniendo en cuenta el nivel de demanda no alcance a movilizar los pasajeros hacia su respectivo destino, se proponen vehículos adicionales que suplen la demanda por recoger, dichos buses extras seguirán el comportamiento de la ruta más lenta, la cual para en todos los nodos, y es penalizada en los indicadores.

12.2. Pronóstico de la demanda

12.2.1. Horizonte de pronóstico:

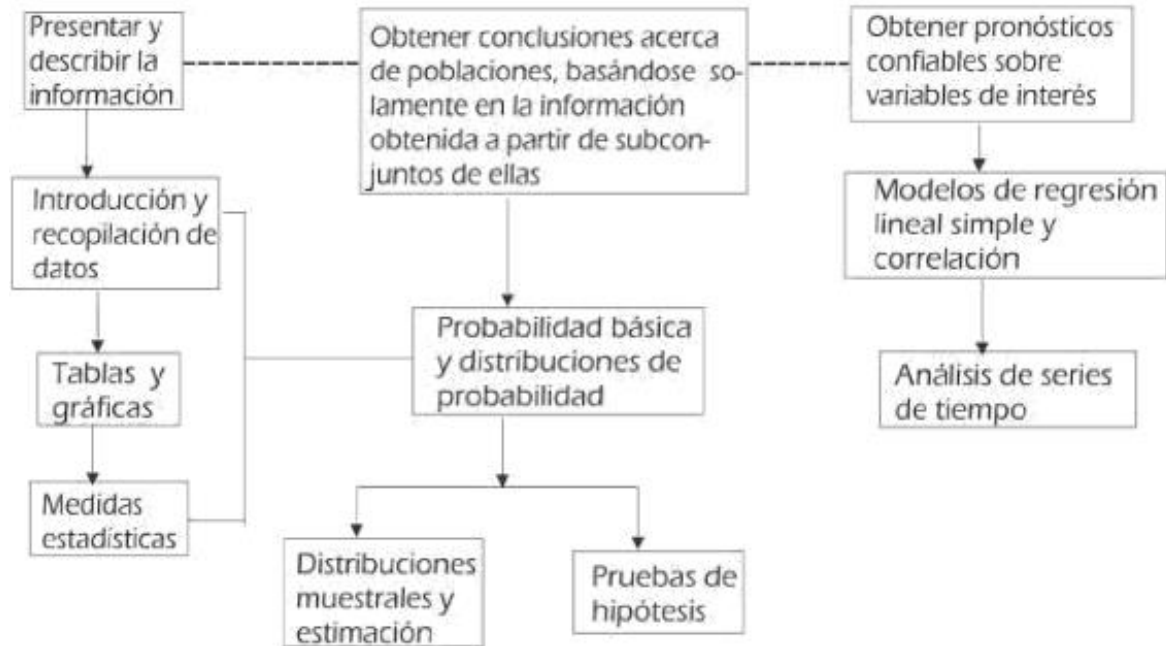
El pronóstico permite conocer hasta una semana después del periodo analizado, es decir siete días posteriores al 30 de septiembre de 2017. Y se restringe a un mínimo Justificado en que las muestras de entradas de pasajeros se encuentran recolectadas durante el periodo de ocho meses, debido a las agrupaciones realizadas el comportamiento de la troncal es descrita según la estación y el día de la semana, lo cual propone entre 34 y 35 posibles datos históricos.

¿Porque se decide usar pronósticos en frente a distribuciones de probabilidad?

Los pronósticos permiten describir como una variable cambia con respecto al tiempo y conocer el comportamiento de una variable en el futuro, en este caso, la demanda desempeña un papel fundamental para la descripción del modelo por lo cual un mejor entendimiento de sus niveles con respecto a las características propias de la troncal es útil.

Es claro para esta investigación que el tratamiento de datos sea del tipo numérico relacionado con variables discretas, porque surgen de un conteo basado en muestras de una hora. Y como dichos datos disponibles para el desarrollo de la investigación una vez observados, no exhiben un comportamiento de dependencia ideal, que pueda asimilarse a una función real, y tampoco brindan un comportamiento basado en la experimentación o información previsible, o que se pueda homologar a una distribución de probabilidad, el autor propone en el presente trabajo como hipótesis, no solo basarse en herramientas estadísticas, sino que en virtud de que lo que se encuentra, es un conjunto de datos registrados durante un periodo de tiempo “históricos”. Se relaciona a continuación un diagrama que es acorde a este planteamiento.

Figura 16. Esquema análisis descriptivo.



NOTA: Tomado del libro Estadística descriptiva y distribuciones de probabilidad Llinás S, H., & Rojas A., C. (2005). Estadística descriptiva y distribuciones de probabilidad;

Vale la pena explorar el camino descriptivo del análisis de datos asimilándolo a predicciones, lo que conlleva a que se proponga un componente de pronóstico que permita en primera instancia entender mejor las variables del sistema, como un todo, por cuando los pronósticos en general emplean técnicas para el análisis de datos y proyectan efectos futuros. (Mason, D. Lind, & Marchal, 2003)

12.2.2. Métodos de pronóstico.

Los datos históricos de la muestra organizados y filtrados con respecto a la fecha y el lugar en el que fue tomado un registro, complementado por los resultados del análisis de normalidad es la base para predecir un dato, mediante metodologías de pronóstico, se toma como referencia las predicciones realizadas por el método de promedio móvil y suavización exponencial.

1. **Regresión lineal:** utiliza la técnica de mínimos cuadrados para encontrar una recta que describe el comportamiento de la demanda.
2. **Pronóstico Intuitivo:** elije el dato del periodo inmediatamente anterior.

3. **Promedio simple:** Realiza la asignación con un valor igual al promedio de toda la muestra.
4. **Promedio móvil:** Realiza un promedio de n periodos anteriores, aclarando que para ese estudio se realiza un promedio de los tres periodos inmediatamente anteriores.
5. **Suavización exponencial:** Mediante una variable independiente esta metodología es sensible al error de predicción obtenido en periodos anteriores.

12.3. Componentes del pronóstico.

12.3.1. Componente Ciclicidad

Este componente es descartado debido a que el periodo estudiado no supera el año.

12.3.2. Componentes De Aleatoriedad Modelo De Clústeres

Agrupación por día de la semana

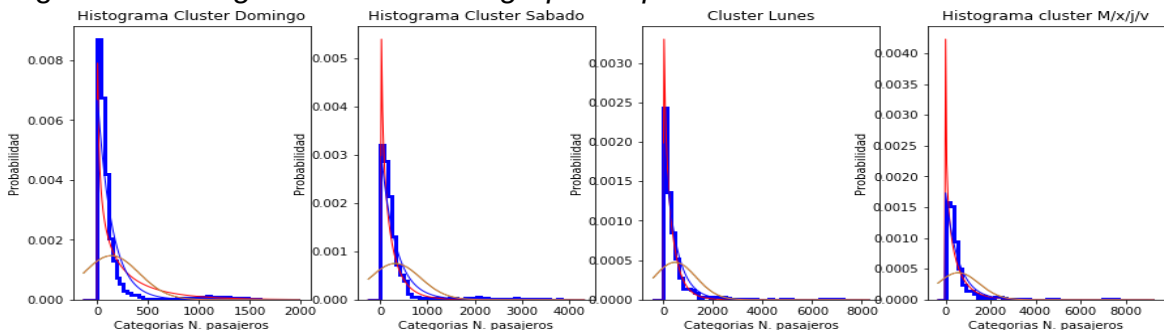
Para generalizar el comportamiento de una fecha se busca generalizar el sistema utilizando las agrupaciones realizadas en el paso anterior. Según la agrupación realizada por la metodología Anova, de un solo factor, se decidió tomar como muestra 20 grupos de datos que presentan el comportamiento total del sistema basado en la agrupación, manteniendo una temporalidad constante de las muestras. Para esta etapa la agrupación horaria no se tiene en cuenta. Los 20 grupos están determinados por la combinación de días y estaciones representativos, teniendo en cuenta que el comportamiento de cada mes es distinto.

Debido a que en una agrupación existe más de una opción en el momento de elegir qué grupo de estaciones o que día representa mejor el comportamiento general. Se decide escoger solo un representante del grupo, al aplicar un cálculo de cuán grande es la diferencia entre la media del grupo en valores positivos con respecto a los otros integrantes se elige el valor central que sirve como criterio para la elección, resaltando los representantes por cada opción.

Tabla 79. Clúster día semana con elección de día representativo.

1	2	3	4
Domingo	Sábado	Lunes	Martes Miércoles Jueves Viernes

.Figura 17. Histogramas entradas agrupadas por día de la semana



Nota. Los histogramas exponen la probabilidad que un dato de entrada de pasajeros alcance un valor determinado, el histograma 4 agrupa los días: martes, miércoles, jueves y viernes.

Agrupación por estaciones.

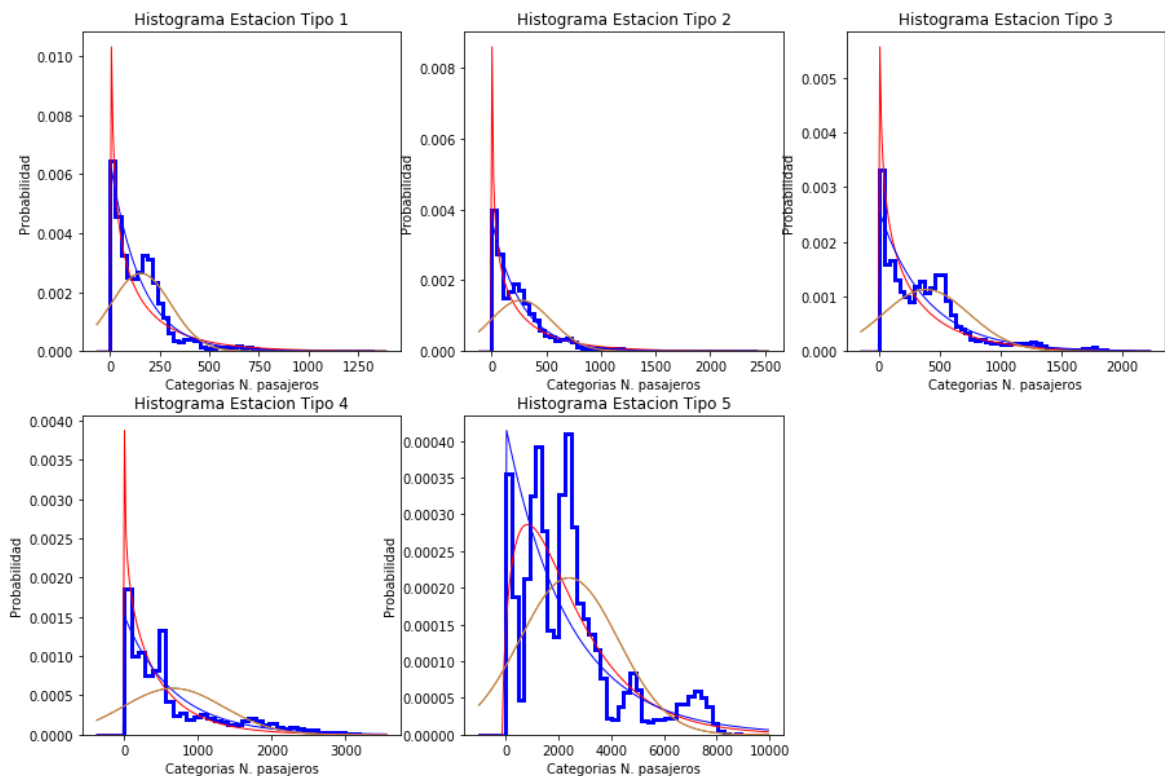
Para los tres clústeres de estaciones que requieren decisión se escogieron las de Quinta Paredes, Normandía y Ciudad Universitaria, mientras que para los otros dos solo se usaron los datos de las estaciones Universidades y Portal Dorado ya que tienen un comportamiento único

Tabla 80. *Clúster estaciones Transmilenio.*

1	2	3	4	5
Plaza de la Democracia	Modelia	CAN Ciudad Universitaria	Universidades	Portal El Dorado
Quinta Paredes	Normandía	Salitre Greco		
Centro Memoria	Av. Rojas Gobernación	El Tiempo		

Nota. Elección de estaciones representativas del grupo obtenido mediante ANOVA de un solo factor.

Figura 18. *Histogramas de datos agrupados por estaciones de la troncal El Dorado*



Nota. Los gráficos representan la distribución de los diferentes grupos de datos agrupados por estaciones.

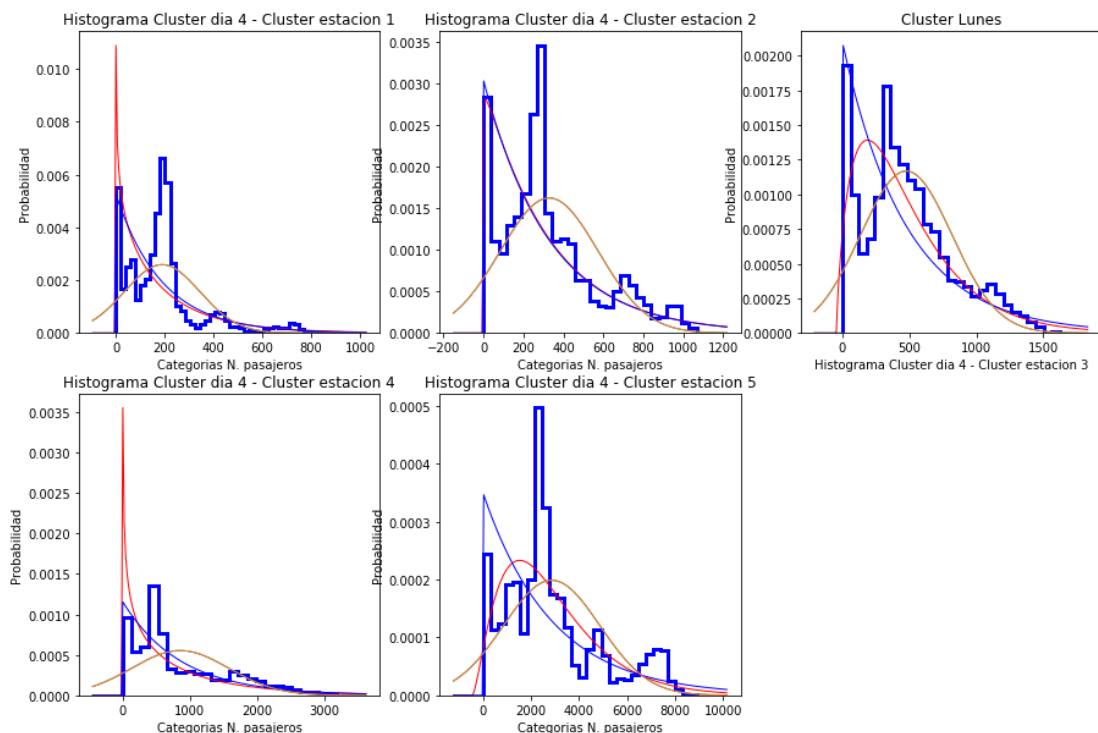
Como resultado de las anteriores elecciones el sistema podría simplificarse mediante las veinte posibles etiquetas que se le asignara a un dato de demanda, el cual representara en el modelo al grupo al que pertenece. Como ejemplo se analiza el grupo o clúster conformado por las muestras recolectadas en la estación portal dorado y el tipo de día cuatro que representa los días, martes, miércoles jueves y viernes, se observa un mayor ajuste a la distribución normal.

Tabla 81. Grupos análisis de demanda

Clúster	Quinta Paredes	Normandía	Ciudad Universitaria	Universidades	Portal El Dorado
Domingo	Domingo - Quinta Paredes	Domingo - Normandía	Domingo - Ciudad Universitaria	Domingo - Universidades	Domingo - Portal El Dorado
Sábado	Sábado - Quinta Paredes	Sábado - Normandía	Sábado - Ciudad Universitaria	Sábado - Universidades	Sábado - Portal El Dorado
Lunes	Lunes- Quinta Paredes	Lunes- Normandía	Lunes- Ciudad Universitaria	Lunes- Universidades	Lunes- Portal El Dorado
Viernes	Viernes- Quinta Paredes	Viernes- Normandía	Viernes- Ciudad Universitaria	Viernes- Universidades	Viernes- Portal El Dorado

Nota. 20 set de datos que representan el sistema en un lugar y un día a la semana.

Figura 19. Histogramas datos agrupados por el criterio de clúster dos criterios para las diferentes estaciones el día viernes.



Nota. El grafico presenta las 5 posibles agrupaciones dentro de la agrupación, conjuntos que respetan la clasificación de estaciones y días de la semana.

Datos atípicos

Se basa en las medidas obtenidas de la comparación entre los valores de media y mediana con respecto al dato de la demanda particular, lo cual posibilita la identificación de patrones que, de acuerdo con el criterio del investigador, pueden corresponder a disrupciones del sistema o hitos que deben ser confirmados mediante fuentes de información externa. Todo ello soportado en las 60 posibilidades de valores de media y mediana, con criterios bajo los cuales se estableció la normalidad para guiar la diferenciación de comportamientos entre típicos y no típicos.

Criterio 1. Normalidad de coincidencia con la ocurrencia de eventos

Gracias a fuentes externas, como calendarios de días festivos o feriados, y fuentes de información, como el tiempo.com noticias, se relacionaron eventos ocurridos en el año 2017, proceso en el cual se detectó 63 eventos que afectaron el comportamiento del sistema, de los cuales 22 corresponden a días conmemorativos y el resto a disrupciones basadas en la aglomeración de personas.

Eventos tipo 1. Festivos o feriados

- | | |
|----------------|---------------------------------------|
| 1. 2017-01-09 | Reyes magos |
| 2. 2017-03-20 | San José |
| 3. 2017-04-12 | Semana Santa |
| 4. 2017-04-13 | Semana Santa |
| 5. 2017-04-14 | Semana Santa |
| 6. 2017-04-15 | Semana Santa |
| 7. 2017-05-01 | Día del trabajo/ marchas |
| 8. 2017-05-29 | Día de la Asunción |
| 9. 2017-06-19 | Corpus Christi |
| 10. 2017-06-26 | Día del Sagrado Corazón. |
| 11. 2017-07-03 | San Pedro y San Pablo/ Rock al Parque |
| 12. 2017-07-20 | Día de la Independencia Patria |
| 13. 2017-08-07 | Batalla de Boyacá |
| 14. 2017-08-21 | Asunción de la Virgen |
| 15. 2017-10-16 | Día de la Raza |
| 16. 2017-11-06 | Día de todos los santos |
| 17. 2017-11-13 | Día de la independencia de Cartagena |

18.2017-12-08	Navidad
19.2017-12-25	Inmaculada Concepción de la Virgen
20.2017-08-27	Caminata de la Solidaridad
21.2017-04-09	Bogotazo
22.2017-05-08	Feria del libro

Eventos tipo 2. Disrupciones.

23.2017-01-12	Cierre del Portal Dorado
24.2017-01-17	Empleados de Transmilenio
25.2017-01-19	Empleados de Transmilenio
26.2017-02-02	Día sin carro
27.2017-02-27	Hasta el 4 de marzo cerrado el carril Sur /aguacero fuerte
28.2017-03-01	Cierre de calle
29.2017-04-17	Cierre de carril
30.2017-05-10	Paro de taxistas
31.2017-05-15	Marchas de maestros
32.2017-05-16	Marchas de maestros
33.2017-05-18	Marchas de maestros
34.2017-05-25	Marchas de maestros
35.2017-06-05	Marchas de maestros
36.2017-06-07	Marchas de maestros
37.2017-06-08	Marchas de maestros
38.2017-06-09	Marchas de maestros
39.2017-06-13	Marchas de maestros
40.2017-06-14	Marchas de maestros
41.2017-06-15	Marchas de maestros
42.2017-06-16	Marchas de maestros

43.2017-06-17	Marchas de maestros
44.2017-06-18	Cumpleaños del equipo de Millonarios
45.2017-07-01	Rock al parque
46.2017-07-02	Rock al parque
47.2017-07-07	Accidente en la Carrera Boyacá
48.2017-08-15	Sindicato de maestros de Bogotá protesta ante la Secretaría de Educación.
49.2017-08-31	Día cívico
50.2017-09-04	Trancones de la Calle 26
51.2017-09-06	Visita del Papa
52.2017-09-07	Visita del Papa
53.2017-09-08	Visita del Papa
54.2017-09-09	Visita del Papa
55.2017-09-10	Visita del Papa
56.2017-09-17	Disturbios
57.2017-09-27	Disturbios
58.2017-10-04	Manifestación estudiantil
59.2017-10-23	Paro de taxistas
60.2017-10-31	Traslado de avión en la calle 26
61.2017-04-07	Manifestación estudiantil

Eventos tipo 3. Eventos relacionados con fuerzas de la naturaleza

La complejidad que reviste el examen de la influencia de las fuerzas de la naturaleza como terremotos y lluvias torrenciales sobrepasa los objetivos de este estudio. Por tal motivo los eventos más significativos son catalogados dentro de la anterior clasificación.

Normalidad respecto del grupo específico

Los niveles de aceptación de normalidad se estipularon aproximadamente como valores por encima del promedio en la desviación individual con respecto la media, porcentaje obtenido como el cociente entre la diferencia de cada cantidad de pasajeros de cada registro con respecto a la media del grupo, sobre el valor de la media del grupo.

$$\text{Ecuación 10 } d_1 = \frac{d - \bar{x}_{\text{grupo}}}{\bar{x}_{\text{grupo}}}$$

Cabe aclarar que una variación de este criterio podría ser la desviación individual con respecto a la mediana.

$$\text{Ecuación 11 } d_2 = \frac{d - M_{e\text{grupo}}}{M_{e\text{grupo}}}$$

Con fundamento de la hipótesis de que los datos normales presentarán variaciones leves respecto del grupo, una gran variación con respecto a la media sugiere que el dato tiende hacia los valores extremos de la distribución, lo cual a su vez apunta hacia el posible análisis de los registros que tengan similares características.

Por ejemplo, si se extrae la información de un registro como el siguiente, que corresponde a un dato tomado el primero de mayo de 2017, éste será comparado en primer lugar con respecto a su grupo, que utiliza muestras con comportamientos similares. En este caso el clúster de mes número tres generaliza los meses de febrero, marzo, mayo, agosto y septiembre. El clúster del día de la semana solo tiene en cuenta los lunes y al mismo tiempo el tipo de estación se limita a las paradas Quinta Paredes, Plaza Democracia y Centro de Memoria.

Tabla 82. *Clúster estaciones Transmilenio.*

Identificación por Clúster				Información del grupo		Variaciones con respecto al grupo				Normalidad		
Fecha	Mes	Día	Estación	Demanda	Media	Mediana	V. Media	%	V. Mediana	%	Normal	Eventos relevantes
01/05/2017	3	3	1	681	3223,05	2555,00	-					Día del
							2542,04					
							762	-79%	-1874	-73%	1	Marchas

Así, se observa que el dato se aleja bastante de la posición de los otros registros tomados en la misma estación y en el mismo día de la semana. Pero para identificarlo como una anomalía, un análisis de la fecha en diferentes estaciones enriquece el criterio del investigador, es decir, que la visión de la red permite decidir si en ese día la disminución de la demanda fue de toda la troncal o fue un caso aislado.

Tabla 83. Clúster de estaciones de Transmilenio

Fecha	Mes	Día	Estación	Demanda	Media	Mediana	V. Media	%	V. Mediana	%	Norma	Eventos relevantes
01/05/2017	3	3	1	681	3223,05	2555,00	- 2542,047 62	-79%	-1874	-73%	1	Día del trabajo / Marchas
01/05/2017	3	3	2	1572	5655,10	4323,00	- 4083,095 24	-72%	-2751	-64%	1	Día del trabajo / Marchas
01/05/2017	3	3	3	4562	8684,57	6777,00	- 4122,571 43	-47%	-2215	-33%	1	Día del trabajo / Marchas
01/05/2017	3	3	4	2392	16066,48	12652,00	- 13674,47 62	-85%	-10260	-81%	1	Día del trabajo / Marchas
01/05/2017	3	3	5	18905	51163,05	41139,00	- 32258,04 76	-63%	-22234	-54%	1	Día del trabajo / Marchas

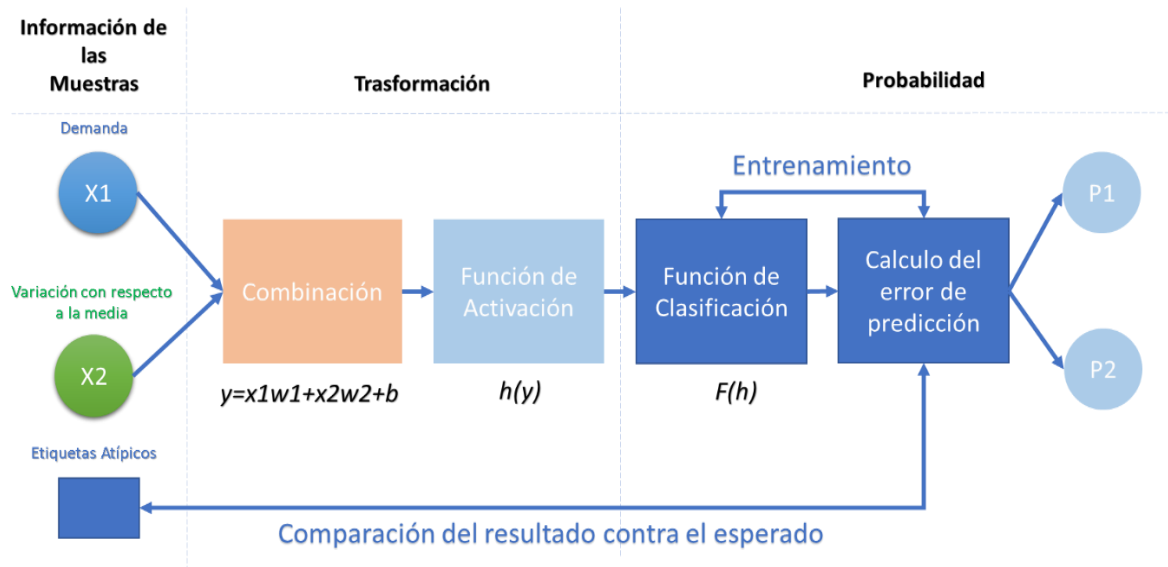
Nota. Las variaciones coinciden con un hecho histórico que ocurrió en la ciudad, por lo cual se confirma que durante esta fecha se presentó una interrupción en la red.

Criterio 2. El Perceptrón

Debido a que el análisis de cada dato individual puede estar sujeto a errores, una forma de clasificar mucho mejor los datos según sus propiedades puede ser el uso de una regresión multivariable que describa qué tan similar es el comportamiento individual respecto a los días festivos o a los días en los que ocurrió una interrupción.

Este análisis se realiza mediante un aprendizaje supervisado, que consiste en la implementación de un perceptrón, basado en los datos etiquetados como atípicos, cumple con la finalidad de calcular la probabilidad de que una muestra pertenezca al conjunto de datos típicos o al de datos atípicos.

Figura 20. Gráfico explicativo del modelo perceptrón para la clasificación de datos



Nota. El gráfico resume el proceso matemático implícito en el cálculo de las probabilidades de que un dato pertenezca al grupo de muestras normales o muestras atípicas.

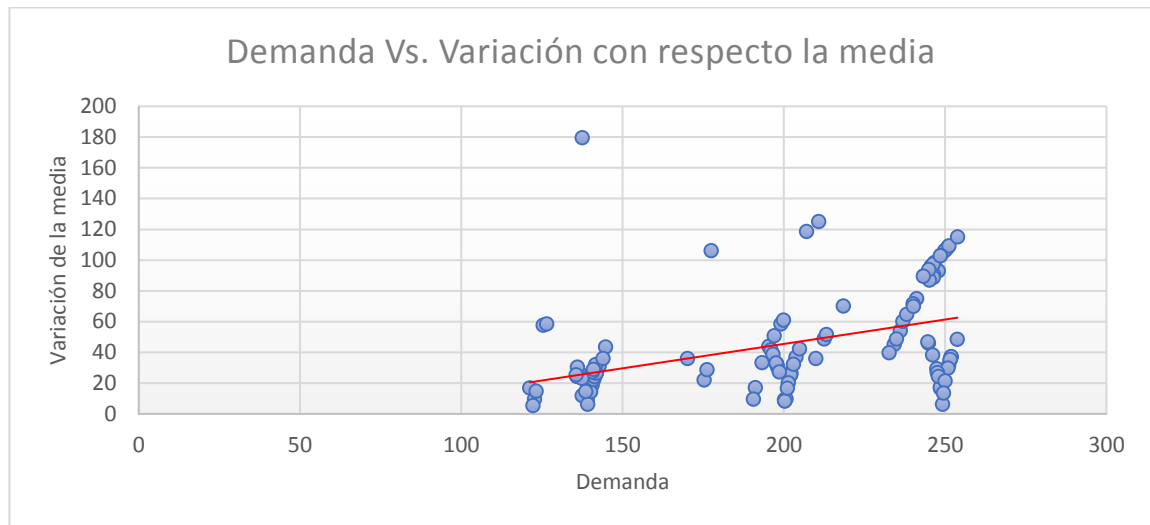
Para que esto sea posible los datos son transformados, tomando únicamente dos variables para analizar. En primer lugar, la variable x_1 corresponde al número diario de pasajeros que entran en la estación y la segunda variable a la desviación respecto de la media en términos absolutos, donde d corresponde a la demanda del registro y $\bar{x}$ a la media del grupo.

$$\text{Ecuación 12} \quad x_1 = d$$

$$\text{Ecuación 13} \quad x_2 = |d - \bar{x}_{grupo}|$$

Graficando las dos variables, se observa una distribución como la que se muestra a continuación:

Figura 21. *Clúster en estaciones de Transmilenio*

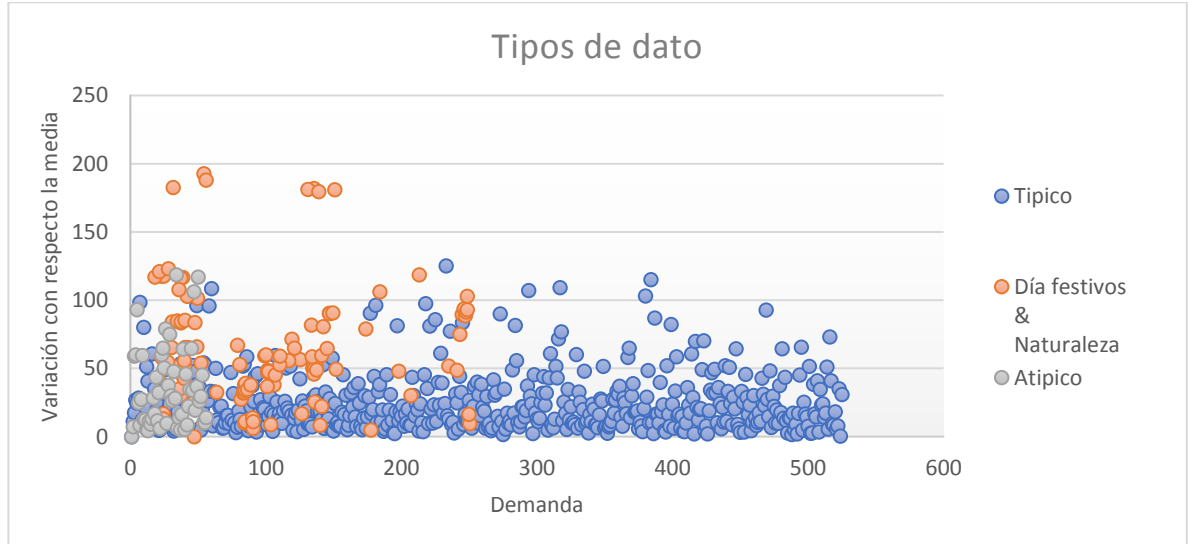


Nota. La grafica muestra la relación del valor absoluto de la variación con respecto a la media. Autoría propia

Este modelo tiene en cuenta que la diferencia con respecto a la media aumenta a medida que sube el volumen de la demanda, lo cual es útil para diferenciar los comportamientos entre las estaciones con baja demanda y aquellas que tienen altos niveles.

Un factor importante para tener en cuenta es que la muestra está conformada por tres tipos de registros, clasificados como típico, días festivos y atípico, tal como se observa en el gráfico que aparece a continuación, donde los colores de los tres grupos se visualizan dentro del campo muestra.

Figura 22. Tipos de datos en el campo muestra de la muestra de clúster



Nota. El gráfico muestra tres tipos de registros clasificados según los criterios de normalidad, y utiliza una transformación $X_1 = \sqrt{x_1}$ sobre las variables que disminuye la escala en ambos ejes, proporcionalmente.

Para saber a cuál grupo pertenecería un registro se plantea un perceptrón, que tiene cuatro pasos en primer lugar, una transformación lineal o suma ponderada de las dos variables de la forma:

$$\text{Ecuación 14 } Y = x_1 * w_1 + x_2 * w_2 + b$$

donde las x_i corresponden a las variables o características, w_i a pesos aleatorios y b_i al intercepto. En segundo lugar, al resultado de la transformación se le aplica una función de activación; en este caso se utiliza la función conocida como Leaky Relu, que tiene el objetivo de transformar una función, así:

$$\text{Ecuación 15 } Relu = f(Y) = \begin{cases} -0,01 * Y, & x < 0 \\ Y, & x \geq 0 \end{cases}$$

Y para este caso:

$$\text{Ecuación 16 } Relu = f(x) = \begin{cases} -0,01 * (x_1 * w_1 + x_2 * w_2 + b), & x < 0 \\ x_1 * w_1 + x_2 * w_2 + b, & x \geq 0 \end{cases}$$

El resultado de este procedimiento es denominado como h , una función que simplifica la identificación de patrones en el campo muestral. Esta capa también es conocida como una capa de entrada para una red neuronal.

Una vez activados los datos, se procede a analizarlos mediante la función Softmax, cuya utilidad es la posible interpretación de su resultado como una probabilidad. Se utilizan nuevos pesos determinado como w , y otros interceptos b .

$$\text{Ecuación 17 } \textit{Softmax} = f(x) = \frac{e^{(w_i \cdot h_i + b_i)}}{\sum_{i=0}^n e^{(w_i \cdot h_i + b_i)}}$$

La aplicación de redes neuronales suele utilizar métodos distintos para elegir el valor de los pesos y los correspondientes valores de intercepto, como el gradiente descendente para encontrar los pesos adecuados con base en el error de la predicción, el cual se obtiene como la diferencia entre el valor deseado y el valor predicho. Esta medición se aborda en la literatura desde distintas ecuaciones, como puede ser el error cuadrado medio:

$$\text{Ecuación 18 } ECM = \frac{1}{n} * \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

Una ecuación que también presenta buenos resultados es la ecuación de entropía cruzada binaria, usada en las regresiones logísticas.

$$\text{Ecuación 19 } error = -\frac{1}{n} * \sum_{i=1}^n y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i)$$

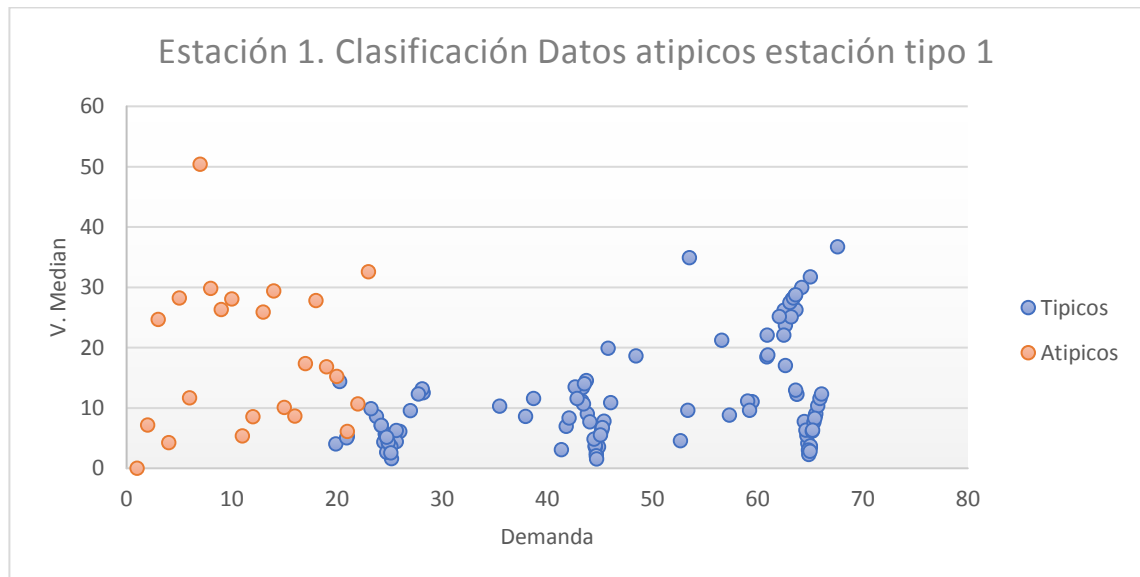
Esta ecuación logra castigar valores bajos cuando deberían ser altos y altos cuando deberían ser bajos. Para ambas ecuaciones el signo $\hat{y}_i$ designa la probabilidad calculada por el modelo, y_i el resultado deseado o el valor de la etiqueta asignada en pasos anteriores y n el número de datos analizados.

Para que el modelo de clasificación funcione como una neurona es fundamental que el entrenamiento de los pesos se haga de forma iterativa. Por esta razón la herramienta escogida para imitar este proceso fue “Solver” de Microsoft Excel, con la cual se eligen números entre -1 y 1, que varían según el criterio de error del modelo, teniendo presente que la efectividad de usar esta herramienta no se iguala con la arquitectura de modelos basados en técnicas bastante populares, como el gradiente descendente.

En este punto se propone diferenciar de algún modo el comportamiento de un dato de acuerdo con la posición en el campo muestral. Para facilitar dicho trabajo se divide el problema en varias partes: según la estación, para evitar la influencia creciente de la tendencia de la población, o según el tipo de dato a detectar: festivo o atípico.

Un ejemplo de este planteamiento se desarrolla a continuación, cuando se elige estudiar la estación tipo 1 y los días atípicos, y se busca diferenciar el comportamiento de los dos grupos, tal como se muestra en el gráfico.

Figura 23. *Campo muestral del tipo de estación 1, Datos normales y atípicos.*

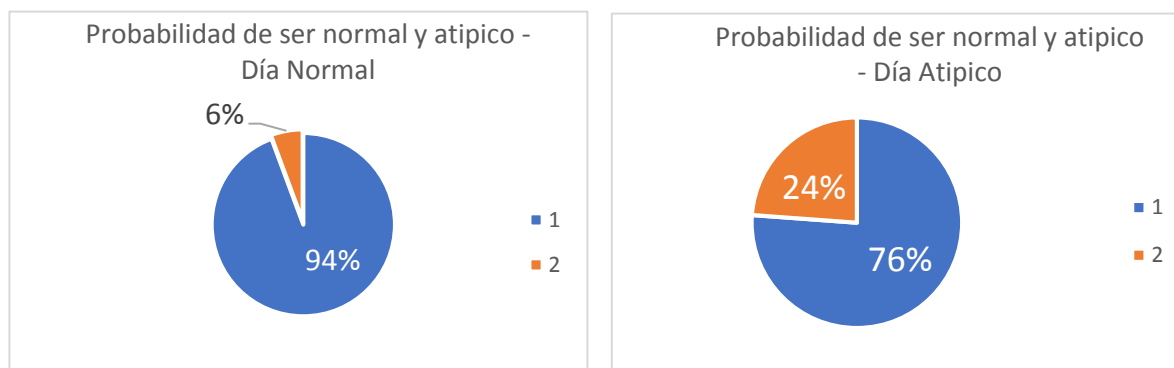


Nota. El gráfico solo incluye los datos que corresponden al tipo de Estación 1 y muestra los datos considerados como normales y los datos considerados como atípicos, y no incluye los días festivos.

En la etapa de entrenamiento se divide el grupo en dos subgrupos, con datos para entrenamiento y datos de prueba. Los datos de entrenamiento son divididos nuevamente entre el 80% para un entrenamiento inicial y el 20% para reentrenar la red buscando aumentar el asertividad en la detección, predicción que se mide en un grupo de muestras que no ha sido utilizado anteriormente.

El resultado de esta implementación agrega un criterio para que un dato sea catalogado como atípico o que su comportamiento corresponde a un día festivo, lo cual permite tener una medida de probabilidad que cuenta con el soporte de la comparación de las medidas de tendencia central, como lo muestra el siguiente gráfico.

Figura 24. Ejemplo clasificación de datos normales y atípicos según la variación de la media



Nota. Las dos graficas muestran el resultado de la predicción del perceptrón, el porcentaje representa la probabilidad de que un registro pertenezca al grupo de muestra normales con el numero 1 o atípicos con el número 2.

Cuando un registro es etiquetado como normal se espera que la probabilidad de que sea normal resulte muy alta; por ese motivo. en el primer caso la probabilidad de que sea atípico será muy pequeña. A medida que la variación de la demanda de una muestra con respecto a la media aumenta, también lo hace la probabilidad de que sea un dato atípico.

A continuación, se muestran los resultados del ejemplo utilizado, que se muestra en con datos del tipo de Estación 1, con el objetivo de clasificar entre registros anormales y atípicos. El conjunto de entrenamiento se muestra en la siguiente tabla:

Tabla 84. Muestras en el entrenamiento del perceptrón.

Tipo de dato	Numero de muestras	Porcentaje
Típico	76	82%
Atípico	17	18%
Total	93	100%

Nota: los resultados se pueden interpretar mediante la comparación de la etiqueta contra la predicción mediante la siguiente tabla:

Tabla 85. *Error en el entrenamiento de la predicción.*

e^2	Entropía	Numero de predicciones
30,9%	42,9%	
2872,23%	3993,86%	5

Es decir que de los 17 datos atípicos el perceptrón detecto 5 con un criterio de que etiqueta un dato atípico si su probabilidad de ser atípicos supera el 30%.

Tabla 86. *Resultado del perceptrón en la etapa de entrenamiento*

	Normal	Atípico
Predicción Normal	73	12
Predicción Atípico	3	5

Nota. La tabla compara la clasificación previa de un registro y la predicción del perceptrón.

Como ejemplo de la interpretación de la anterior Tabla, la cantidad de predicciones que deberían ser normales y fueron clasificadas como normales es 73; por el contrario, las predicciones erróneas de datos que deberían ser normales y fueron catalogados como atípicos sumaron 3.

Medida de asertividad:

$$\text{Ecuación 20 } \textit{Exactitud} = \frac{V.\textit{Normales} + V.\textit{Atipicos}}{V.\textit{Normales} + V.\textit{Atipicos} + \textit{Falsos Atipicos} + \textit{Falsos Negativos}}$$

Aplicada la medida a los anteriores datos, el modelo predice con una exactitud del 83,87%:

$$\textit{Exactitud} = \frac{73 + 5}{93} = 0.8387$$

Etas mismas medidas utilizadas en la etapa de reentrenamiento utilizando el 20 % restante de los datos, que consiste en alterar los valores de la matriz de pesos e intercepto establecidos en el primer entrenamiento, arrojaron los resultados:

Tabla 87. *Muestras en el reentrenamiento del perceptrón*

Tipo de dato	Numero de muestras	Porcentaje
Típico	15	78,94%
Atípico	4	21,05%
Total	19	100%

Tabla 88. *Error y entropía en la etapa de reentrenamiento*

e^2	Entropía
32,8%	41%
622,93%	780,62%

Tabla 89. *Resultado del perceptrón en la etapa de reentrenamiento*

	Normal	Atípico
Predicción Normal	15	4
Predicción Atípico	0	0

La red predijo con un 79% de asertividad la probabilidad que un conjunto pertenezca a un conjunto en específico, pero tiene un inconveniente que no reconoció a ninguno de los atípicos como atípicos.

$$Exactitud = \frac{15 + 0}{19} = 0.78947$$

Todo el proceso anterior se compara en un conjunto de datos no utilizado, Tabla 90. *Muestras en el conjunto de evolución de la predicción*

Tipo de dato	Numero de muestras	Porcentaje
Típico	14	93,33%
Atípico	1	6,6%
Total	15	

Tabla 91. *Error y entropía final sin reentrenamiento y con reentrenamiento.*

Sin reentrenamiento			Con reentrenamiento		
e^2	F costo	Resultado	e^2	Entropía	Resultado
3,50%	5,47%		4,24%	7,20%	
52,49%	81,99%	1	63,59%	108,00%	1

Nota, Muestra el mismo nivel de predicción, pero el perceptrón con reentrenamiento presenta mayor error y mayor entropía.

$$Exactitud = \frac{15 + 0}{19} = 100$$

En el modelo sin reentrenamiento y con reentrenamiento para el único dato etiquetado como atípico presenta las probabilidades en la figura:



Nota, Las gráficas muestran que el reentrenamiento propuesto no genera una mejor aproximación

Figura 25. Probabilidades de pertenecer a al conjunto de datos normales o atípicos

12.3.3. Componente Estacionalidad y tendencia..

A continuación, se presenta una descripción de los componentes, se presenta un análisis grafico de las tendencias y el cálculo de la estacionalidad por la técnica de media móvil.

Se observa a continuación que la tendencia es similar para el promedio de pasajeros que entran al sistema en una hora, como para el total de pasajeros que entraron al sistema. La diferencia evidente es la escala de las unidades.

Figura 26. Tendencia y ciclicidad del promedio de entradas en una hora.

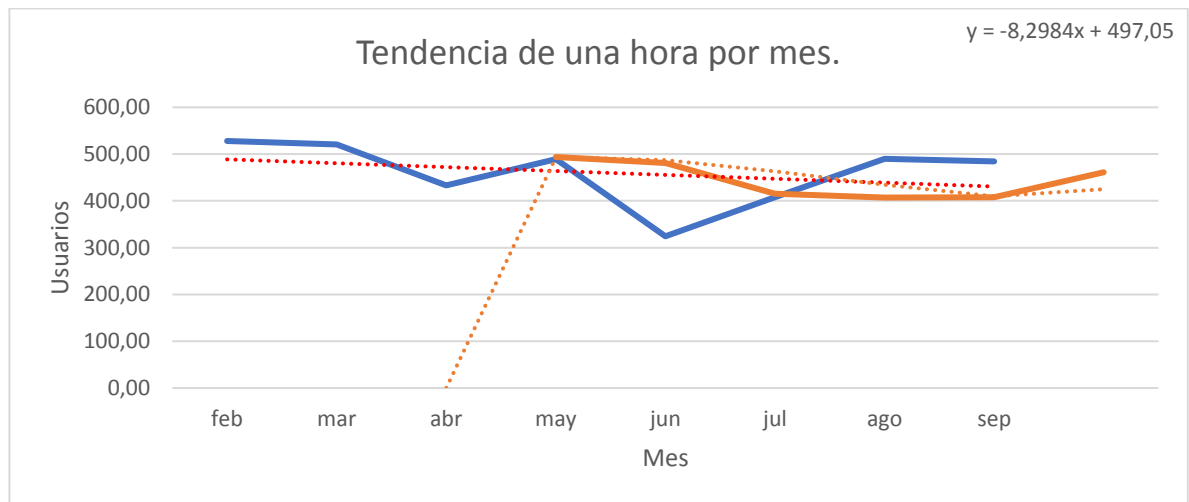
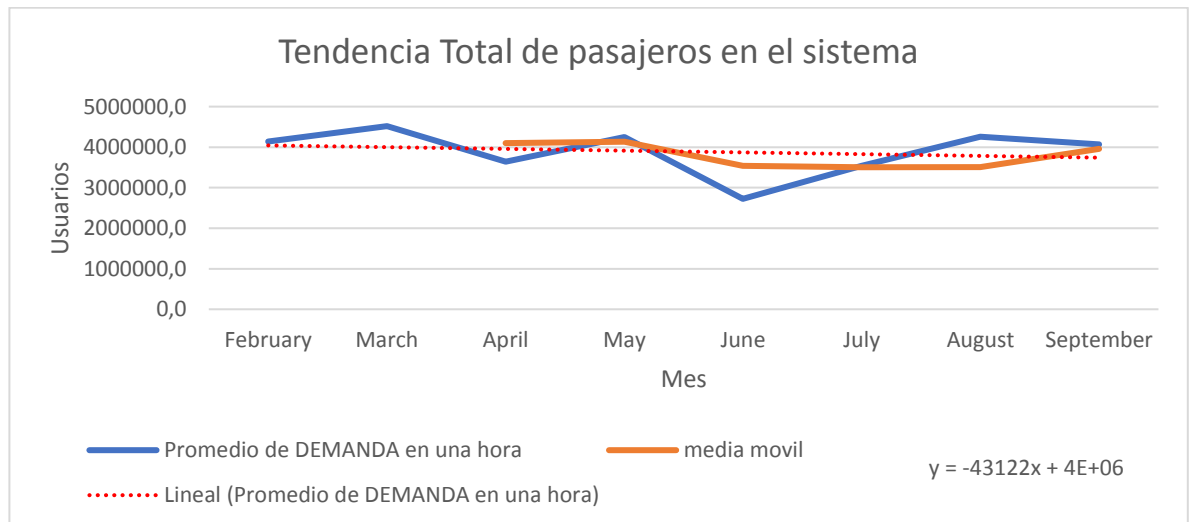
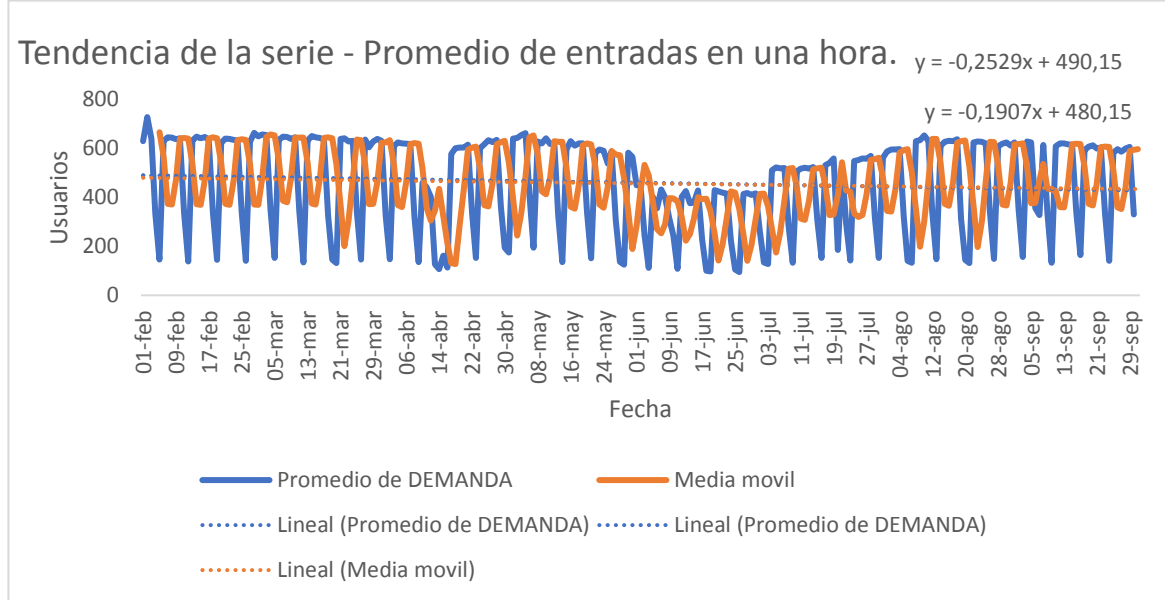


Figura 27. Tendencia y ciclicidad del total de pasajeros en el sistema.



Para una mejor comprensión de la tendencia a continuación se presenta la serie completa.

Figura 28. Tendencia y ciclicidad del promedio de pasajeros en una hora.



La serie completa agrupada por mes tiene los índices de estacionalidad como se muestra en la tabla a continuación.

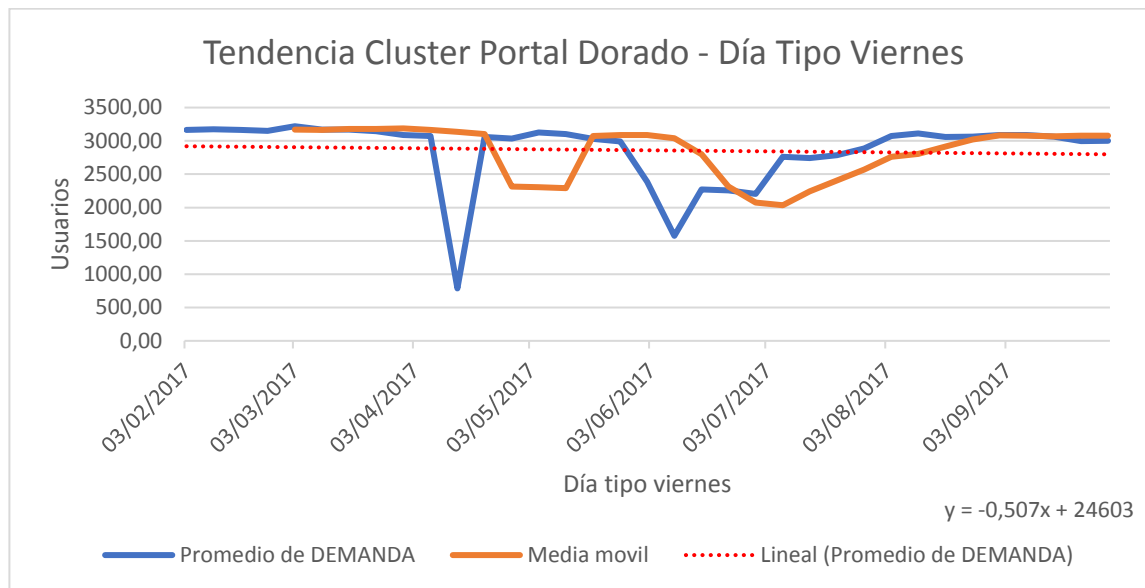
Tabla 92. Estacionalidad de las muestras sin agrupación.

Etiquetas de fila	lunes	martes	miércoles	jueves	viernes	sábado	domingo
February	5,48	7,00	5,16	4,11	3,50	2,38	0,96
March	4,42	7,30	8,62	6,26	5,15	2,11	0,98
April	8,22	7,65	5,99	3,81	3,33	2,77	1,76
May	4,58	9,05	9,04	4,91	4,17	2,25	1,08
June	3,36	7,46	7,76	6,69	5,10	2,21	1,11
July	6,15	7,31	6,94	4,30	4,27	3,02	1,60
August	3,29	9,60	9,64	6,82	4,31	2,12	1,01
September	5,60	6,70	5,34	4,10	5,29	2,99	1,12
Total	41,09	62,08	58,48	41,01	35,12	19,84	9,61
media	5,14	7,76	7,31	5,13	4,39	2,48	1,20
Ajuste	4,31	6,50	6,13	4,30	3,68	2,08	1,01
indice	4,31	6,51	6,13	4,30	3,68	2,08	1,01

Este modo de tratar los datos presenta una estacionalidad muy alta lo cual afecta los pronósticos directamente y evita el cálculo de valores acertados de la demanda. en el ejemplo del día tipo viernes existiría una fluctuación de tres veces el valor.

A un nivel mayor de filtro en termino de clúster por mes, la demanda se comporta de la siguiente forma.

Figura 29. Tendencia Clúster Portal Dorado- Día tipo 4.



Con una estacionalidad un poco más alta como se observa en la tabla a continuación.

Tabla 93. Estacionalidad de la muestra agrupada por dos criterios

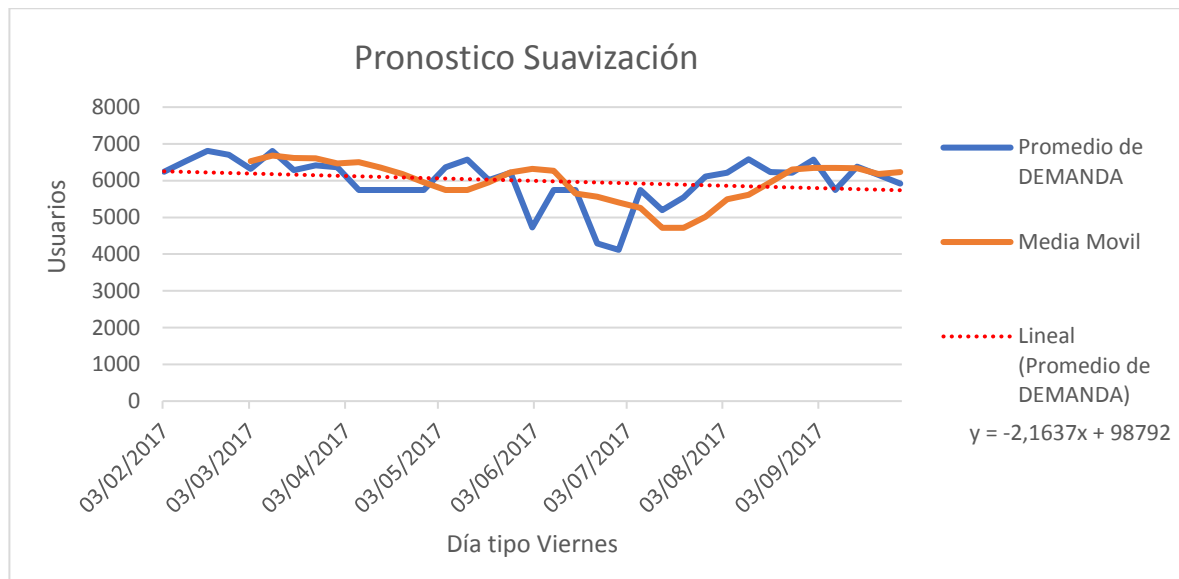
Mes	Valor estacional
February	4,01
March	5,03
April	6,64
May	3,65
June	6,17
July	3,44
August	3,79
September	5,05
Total general	37,77
Índice de estacionalidad	4,72

Se observa una estacionalidad muy grande de aproximadamente cuatro veces el valor real, si se realizase un pronóstico de esta forma el resultado no brinda una descripción del sistema

Mediante una agrupación mucho más al detalle, se analiza la tendencia y la estacionalidad del clúster Portal Dorado – los días tipo 4 en una hora específica 5:00 p.m.

Tendencia y Estacionalidad Cluster a las 5:00 pm

Figura 30. Tendencia y Estacionalidad Clúster a las 5 pm



Segmentación que tiene una estacionalidad mucho menor a las dos anteriores.

Tabla 94. Estacionalidad de la muestra agrupada por 3 criterios.

Etiquetas de fila	Valor estacional
February	1,02
March	1,04
April	1,12
May	0,90
June	1,33
July	0,87
August	0,89
September	0,97
Total	27,45
Estacionalidad	1,02

Lo anterior expone que el efecto de la estacionalidad sobre un pronóstico realizado a nivel de Clúster, estación tipo, día tipo en una hora específica, es afectado en aproximadamente en menos del 2% de un pasajero, (1,02) veces el valor predicho, bajo un modelo aditivo.

12.4. Estructura del modelo

Los datos aquí agrupados son el recurso para estudiar el comportamiento del sistema mediante nuevas medidas estadísticas que permiten la comprensión de las fluctuaciones de la demanda, la variable protagonista de este estudio.

De la misma forma como puede describirse una función mediante una o varias variables, las cuales, a su vez, pueden ser independientes o dependientes, la entrada de pasajeros en el sistema, entendida como la demanda para la operación, puede ser modelada desde el punto de vista matemático. Siguiendo este planteamiento, las principales variables independientes constituyen el momento y lugar en que fue tomado un registro. Que a su vez puede ser clasificado, gracias a las pruebas anteriores, bajo dos filtros:

- Clúster día de la semana o clúster mes.
- Clúster estación.

Cabe resaltar que la utilidad de los grupos o clústers consiste en disminuir la complejidad del estudio, toda vez que para representar un elemento de un grupo específico basta con utilizar una muestra referente a la información del representante escogido en una hora específica, electo por ser el valor central entre los niveles de demanda del grupo.

Tabla 95. *Estructura de la información.*

Fecha	Clúster	Demanda	Medidas	Normalidad	Eventos
Día, mes y año en los que fue recolectado el dato, teniendo en cuenta que corresponde a una agrupación daría.	Describe, mediante tres números, a cuál grupo, relacionado con un momento y un lugar pertenece un dato.	Indica el número de pasajeros que registraron en una estación.	Con respecto al grupo al que pertenece una muestra, se relacionan valores relevantes.	Criterio de normalidad del investigador, basado en la reacción del dato respecto del grupo.	Basados en fuentes externas, relacionan coincidencias entre la fecha y las noticias relacionadas con las interrupciones del sistema.

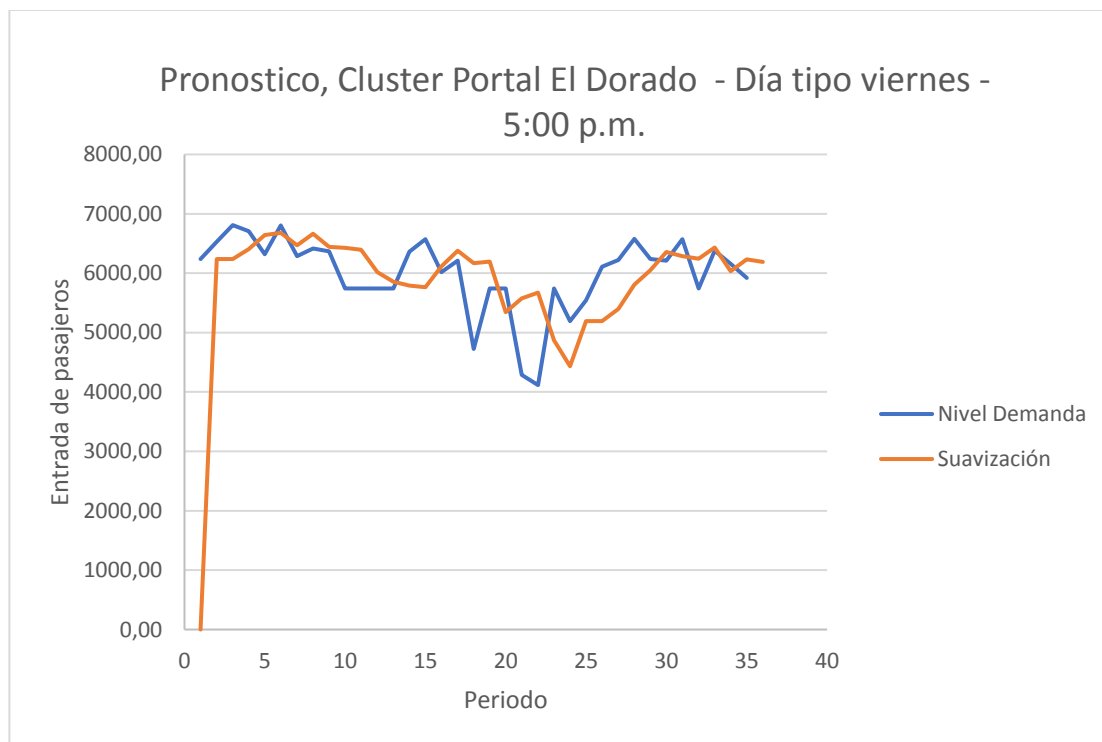
Nota, Propuesta para el tratamiento de los registros de entrada de pasajeros a la troncal estudiada.

12.4.1. Metodología de pronóstico.

Mediante suavización exponencial simple se aborda la predicción y descripción de los datos, método que debe ser aplicado a nivel de clúster segregado por horas. De esta forma se responde a la tendencia de la muestra, y debido a que la estacionalidad de la muestra segmentada en este nivel no supera el 3% en promedio no es tenida en cuenta.

Tomando como solicitud la fecha 29 de septiembre el 2017, se desea conocer cuántos pasajeros entraran al sistema a la 5:00 p.m. que dio como resultado 4615 pasajeros.

Figura 31. Pronóstico para la estación portal el dorado el viernes 29 de septiembre a las 5 pm.

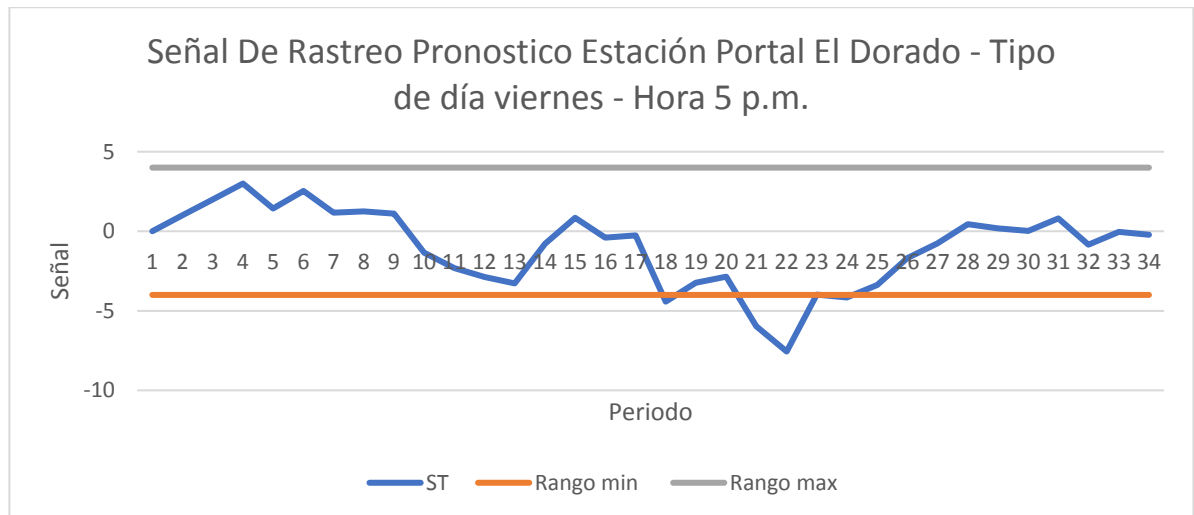


Nota. Se realizó el pronóstico mediante suavización exponencial con un Alpha de 0,37, para la fecha 29 de septiembre en la estación portal el dorado, un viernes.

12.4.2. Error del pronóstico

Cada metodología lleva consigo un error de predicción y una sensibilidad. Se presenta a continuación el ejemplo del cálculo del error para el resultado de la metodología de suavización exponencial.

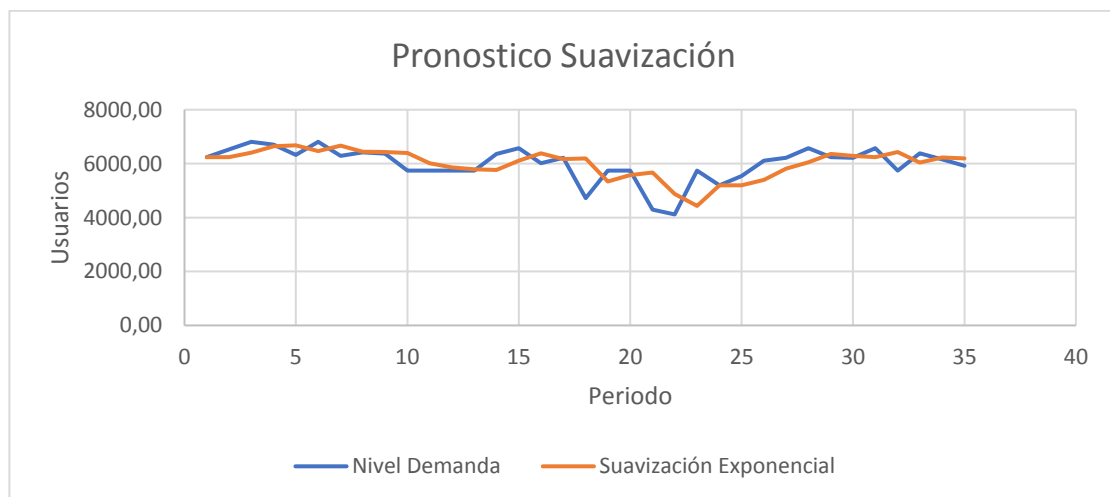
Figura 32. Señal de rastreo pronostico suavización exponencial alpha 0,37



Nota, La señal de rastreo permite conocer que en los periodos de 10 al 14 y del 18 al 23 es prudente usar un criterio para eliminar datos atípicos como el perceptrón. Confirmando con la base de día atípicos las fechas relacionadas a:

- a) Periodo 10 al 14: Corresponde a fechas especiales, semana santa.
- b) Periodos del 18 al 23 están relacionados con marchas en la ciudad.

Figura 33. Suavización Exponencial (Alpha 0,37)



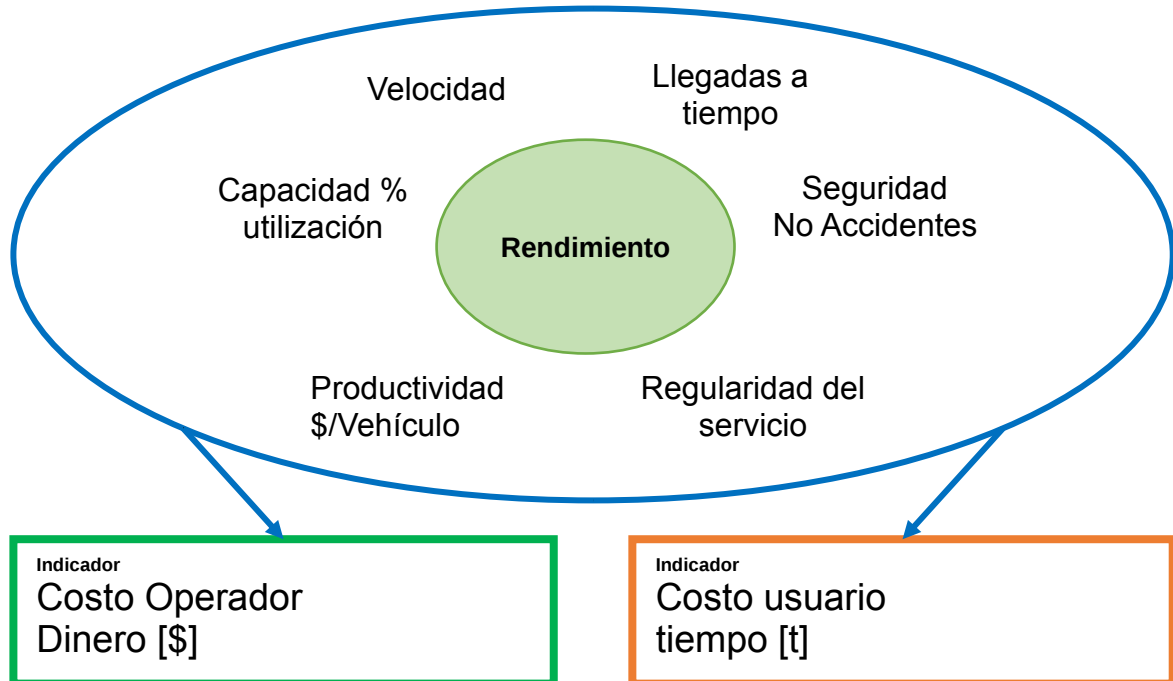
13. Etapa 5. Prototipo y validación.

A continuación, se presenta un ejemplo del funcionamiento del aplicativo, los datos relacionados no corresponden al ejemplo planteado.

13.1. Alternativas de asignación.

El entorno construido por los pasos anteriores sirve para evaluar el rendimiento del sistema, permite conocer cuál es el tiempo total que los pasajeros gastan en el sistema y el costo relacionado con el funcionamiento de los buses asignados a cada ruta y a la red. Está además sometido a otras mediciones, como el factor de ocupación de los vehículos y los excedentes o déficits que arroja una política escogida.

Figura 34. Concepto de rendimiento



Una alternativa será medida mediante dos funciones objetivo, que describen el comportamiento del sistema. Adaptado de: Molinero, A., & Sanchez, L. (1997). Transporte público: Plantación, diseño, operación y administración. Toluca, Mexico: Universidad Autónoma del Estado de México.

Función objetivo 1. Costo acumulado de todos los usuarios en sistema:

Función calculada como el número de vehículos más el número de vehículos no programados, necesarios para cumplir con la demanda de pasajeros por el costo de movilizar un bus.

$$\text{Ecuación 21. } T = \text{Costo de mover un vehiculo} * \text{Numero de vehiculos}$$

Función objetivo 2. Costo acumulado de todos los usuarios en sistema:

Es la suma de todo el tiempo de todos los usuarios que usaron el sistema durante cada hora. En el ejemplo del 1 de mayo a las 7.00 a. m., para movilizar 11.937 pasajeros se obtienen -8.785.14 horas de viaje, es decir que un pasajero gasta en promedio 44 minutos en su viaje, el signo negativo debido a que un menor tiempo de viaje representa un beneficio para el usuario.

$$T = \text{Tiempo de los usuarios en el sisema}$$

Tabla 96. *Tiempos acumulados de los pasajeros en el sistema.*

Ruta 1	Ruta 2	Ruta 3	Ruta 4	Ruta 5	Ruta 6	Ruta 7	EXTRAS PROGRAMA DAS	EXTRAS NO ROGRAMA DAS	Costo pasajero (h)
-1147,71	-1150,28	-1144,07	-1146,22	-1152,37	-1150,65	-1215,83	-677,976	0,000	- 8.785,14

Campo de solución.

Debido a que el costo del operado no necesariamente sigue la tendencia del costo para el usuario, existe más de una solución para satisfacer la demanda de viajes en el sistema, es allí donde se utiliza la metodología Optimización de partículas que es una técnica iterativa basada en los cardúmenes de peces o el comportamiento de las aves mediante las ecuaciones (Lin, Li, Du, Chen, & Ming, 2015).

La técnica recorre distintos tamaños de flota que responden a la demanda siguiendo las ecuaciones:

$$\text{Ecuación 22 } X_i(t + 1) = X_i + V_i(t + 1)$$

Donde X_i representa el tamaño de la flota y V_i al velocidad con la que los integrantes del grupo simulado se mueven hacia los tamaños con mejores resultados en las funciones objetivo.

La velocidad en cada iteración se modifica siguiendo las siguientes ecuaciones.

$$\text{Ecuación 23 } V_i(t + 1) = W V_i(t) + c_1 * r_1 (X_{P_{mejor}} - X_i(t)) + c_2 * r_2 (X_{G_{mejor}} - X_i(t))$$

O mediante el planteamiento multinomial que converge hacia un mínimo o máximo según el mejor global de la iteración anterior.

$$\text{Ecuación 24 } V_i(t + 1) = W V_i(t) + C_i * r_1 (X_{Gmejer} - X_i(t))$$

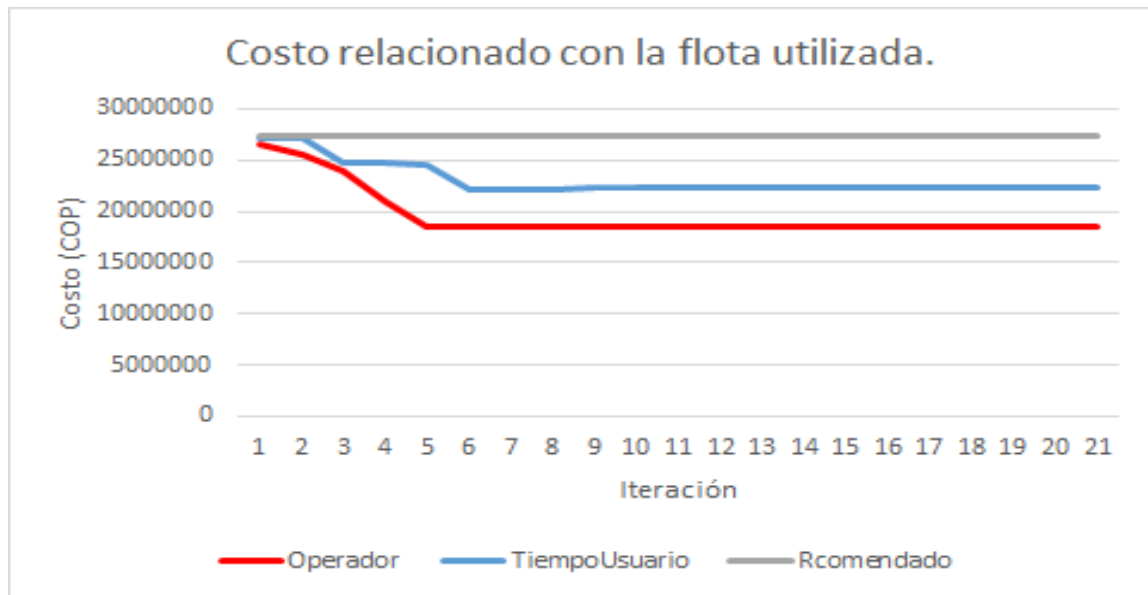
Donde W es una variable de peso de inercia, Las variables r_i numero aleatorios y c_1 pesos de las particulas. Esta metodología permite optimizar según las distintas funciones objetivo, según el costo del operador y según el nivel de servicio medido en tiempo gastado por los usuarios en el sistema.

13.2. Escenarios.

A continuación, se ve el comportamiento de la optimización según la función objetivo del tiempo de usuario usando 10 partículas con 20 iteraciones:

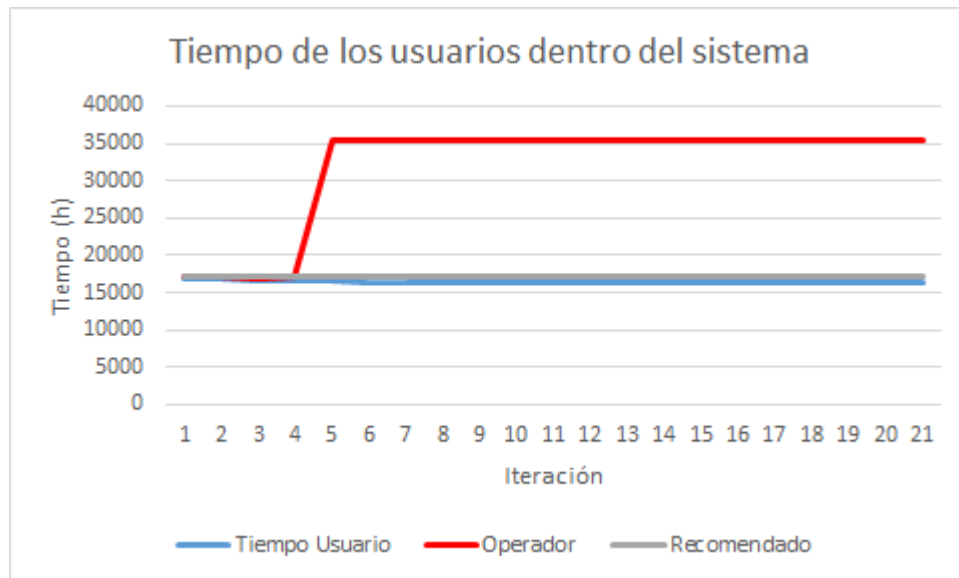
Aplicado al viernes 29 de septiembre de 2017, en el que se obtuvo el siguiente acercamiento que describe el comportamiento de la red para ese día para tres escenarios.

Figura 35. Optimización de la flota para la Troncal Portal El Dorado. Función de costo.



Nota. Se observa que el óptimo para el operador, en cuanto a costo, es más alto que cuando se opta por la función objetivo en función de tiempo acumulado por los usuarios en el sistema.

Figura 36. Optimización de la flota para la Troncal Portal El Dorado. Función tiempo acumulado en el sistema por los usuarios.



Nota. Se observa que el óptimo para el operador, en cuanto a tiempo, resulta en un mayor tiempo acumulado, en más del doble de tiempo promedio de viaje.

Tu Ruta

Principal

Red Día

Opciones de Asignación

Asignación manual de rutas

Pronostico Estación

Elige una hora:

5:00:00 p. m.

6:00:00 p. m.

Pasajeros a movilizar:

23828

Numero de vehiculos

128

Asignación

RECOMENDADO

COSTO DEL OPERADOR

SERVICIO

*Velocidad de la via

45,17

*Densidad en las estaciones

9,15

*Tiempo promedio de viaje

-0,69

*Factor de ocupacion de vehiculos

0,67

Auto

Occidente - Oriente

Ruta	N. Buses	Pasajeros	Tiempo de salidas
K10	10	1940	6
K16	10	1890	6
K23	10	1929	6
K43	10	1895	6
K54	10	1868	6
K86	10	1893	6
Ruta Facil	9	1784	6,6666666
Adicionales	2		30
RF + AD	11		5,4545454
	71	13199	0,85

20

13

60

Numero de estaciones para distribuir la demanda

Densidad Maxima por estación (veh)/(estación)

Velocidad maxima de la via (Km)

2

2

2

2

Peso interaccion entre vehiculos

Peso tiempo de transferencia

Peso tiempo de espera

Peso penalidad rutas extra

Oriente - Occidente

Ruta	N. Buses	Pasajeros	Tiempo de salidas
K10	8	1560	7,5
K16	8	1522	7,5
K23	8	1552	7,5
K43	8	1534	7,5
K54	8	1504	7,5
K86	8	1520	7,5
Ruta Facil	7	1437	8,5714285
Adicionales	3		20
RF + AD	10		6
	58	10629	

3

0,67

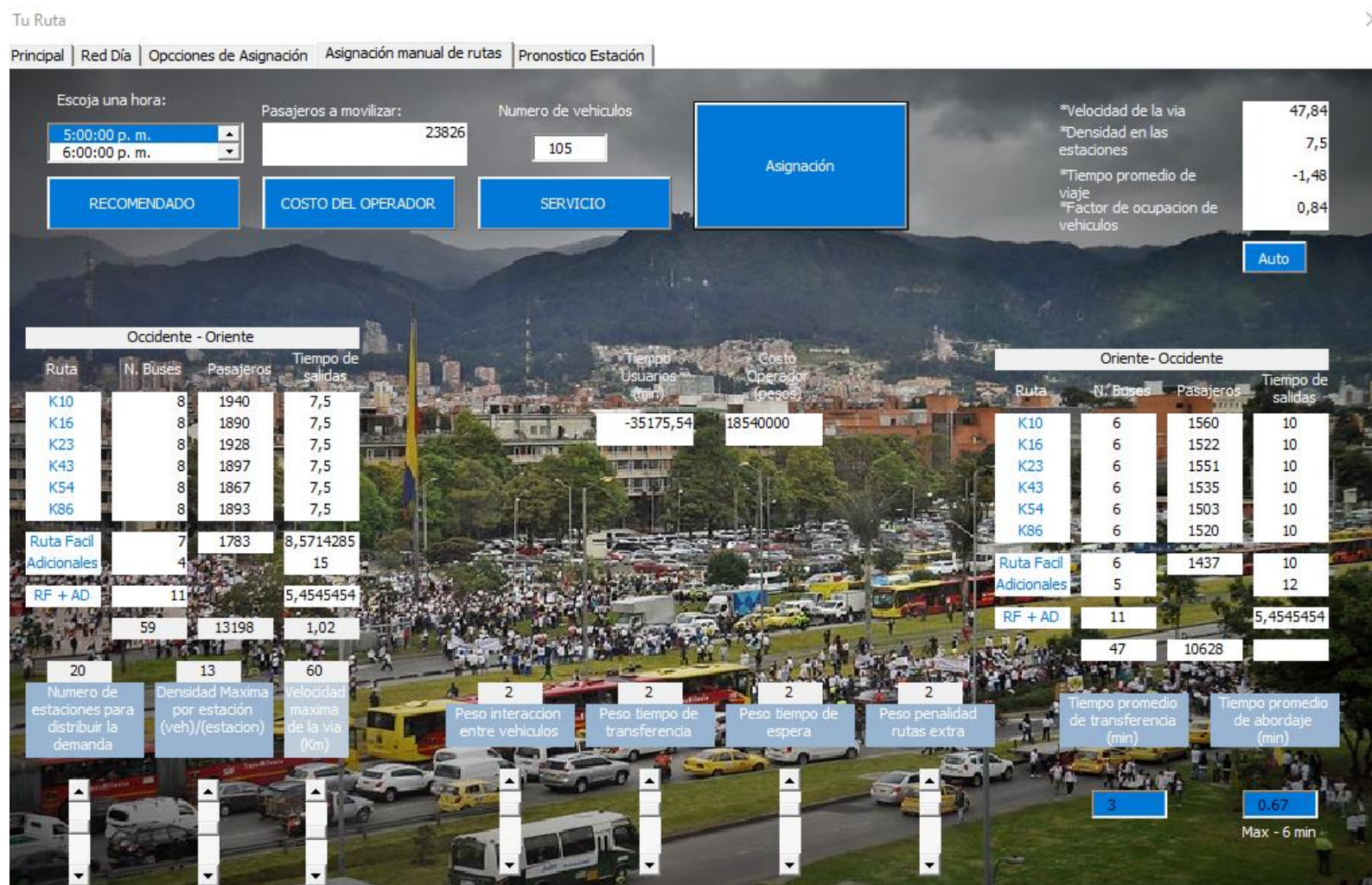
Tiempo promedio de transferencia (min)

Tiempo promedio de abordaje (min)

Max - 6 min

130

Figura 38. Escenario de optimización del tiempo de viaje



Nota, Unidades de tiempo medidas en minutos y unidades de velocidad medidas en km/h

[illegible]

132

14. CONCLUSIONES

En este trabajo se halló la existencia de tres niveles de análisis, estratégico, táctico y operacional, El texto tiene un enfoque táctico y operacional que involucra los niveles finales del diseño de una red de transporte: para este proceso fue vital la revisión de cinco metodologías, estas fueron evaluadas bajo el método AHP, después de este proceso de selección, se establece que las metodologías a aplicar son formulación analítica y algoritmo genético.

El texto describe un sistema ya implementado en la ciudad de Bogotá. Presenta el funcionamiento del sistema BRT generalizado mediante las variables principales, necesarias para abordar el problema de la evaluación del rendimiento y la estimación del tamaño de la flota en una hora, teniendo en cuenta el comportamiento de la demanda de pasajeros que entran al sistema en la troncal Portal El Dorado que hace parte del servicio masivo ‘Transmilenio’, a lo largo de la Av. Calle 26 de la ciudad de Bogotá.

Se encontró que la complejidad del sistema tiene un amplio campo de datos dispersos referentes al gran número de partes con comportamientos que se rigen por la temporalidad y al efecto de datos atípicos, lo cual es visible cuando se estudia la tendencia de todas las muestras que resultan en un comportamiento que no sigue una distribución normal.

Se complementa el proceso con pruebas no paramétricas, en primera instancia con pruebas de homogeneidad: estas pruebas rechazan la hipótesis de que el conjunto de muestras agrupado por estaciones sea homogéneo. Es decir que no se recomienda analizarlas como componentes de un mismo grupo; Como segunda instancia se hayo por medio de pruebas de independencia que los datos tienen relaciones relevantes con respecto a las variantes del día de la semana y las estaciones. Tales relaciones afectan el comportamiento de la cantidad de pasajeros que entran al sistema en un lugar específico de la red estudiada.

El manejo de dicha complejidad dada en la segunda instancia fue abordada mediante grupos de relevancia o clustering, donde los diferentes tipos de comportamiento del segmento fueron agrupados por etiquetas de tiempo y lugar, que comprenden la fecha de la muestra y la estación. De esta forma se tomó una muestra representativa del comportamiento de la red en cada punto a fin de construir un escenario general, el cual es aproximado mediante metodologías de pronóstico, pero presenta la posibilidad de ser estudiado mediante el análisis de las distribuciones de probabilidad que tienen características únicas ara cada caso en particular.

Una vez segmentados los datos en veinte posibles opciones o grupos que recopilan comportamientos similares para los registros agrupados de una forma diaria, se determinó un comportamiento típico y atípico: basado en las medidas de error, con respecto a las medidas de tendencia central y una recopilación de hechos históricos.

Esto permite a su vez construir un grupo de datos para la aplicación de un aprendizaje supervisado, utilizando principios de inteligencia artificial para la clasificación de datos típicos, atípicos y días festivos. Por un perceptrón o red neuronal en su más sencilla expresión, se encontró que debe tener diferenciación, previamente de los datos según el lugar o tiempo en el que fue tomada una muestra para encontrar relaciones matemáticas que puedan detectar comportamientos similares entre los registros y el tipo de disrupción que desea ser analizada (Atípicos y días festivos). En este punto del análisis la metodología de aprendizaje supervisado tiene un amplio potencial de mejora, pero complementa el análisis de los días que se comportan de una forma no típica, que son causados por días festivos o feriados y eventos externos como manifestaciones.

Utilizar la anterior información para la estimación de la demanda de entradas de pasajeros mediante los veinte grupos en una misma hora permite conocer cuántos pasajeros se esperan en el sistema, es decir que se conoce cuantas personas entran a una estación específica, en un día específico de la semana y en una hora específica. Teniendo en cuenta los procedimientos anteriores posibilita el estudio de un punto en el sistema y la construcción de matrices que describen el comportamiento de una red de nodos conectados por una vía, en términos de pasajeros que se transportan desde distintos orígenes hacia distintos destinos, lo cual responde a las restricciones del flujo de vehículos y a la densidad de vehículos, esto permite calcular la velocidad promedio del sistema.

Una vez estimados estos datos, es posible determinar los tiempos de viaje entre nodos, lo cual permite encontrar una función de utilidad desde la opinión del usuario o tiempo que gasta un usuario dentro del sistema, que a la vez posibilita la construcción de un modelo de decisión basado en la probabilidad, La metodología implementada que trata este análisis es un modelo Logit Multinomial que describe la relación que representa escoger entre una ruta y otra, de esta forma es posible determinar una estimación del número de pasajeros que escogen uno de los servicios que operan en la red.

Respecto a los diferentes enfoques de estudio se determinó que la metodología que permite utilizar los datos encontrados para estudiar un abanico de escenarios corresponde a los planteamientos hechos por medio de algoritmos; particularmente se resalta la flexibilidad de los algoritmos genéticos. Por tal motivo, la optimización de partículas fue escogida para explorar de una forma iterativa las posibles soluciones de un problema multiobjetivo.

Se escogieron las dos funciones más representativas las cuales hacen parte del planteamiento analítico: el costo para el usuario y el costo para el operador. Estas medidas permiten evaluar la asignación de determinado número de buses para responder a la demanda solicitada por los usuarios de los servicios disponibles en la red. En la validación del modelo se obtuvo que esta relación entre el costo del operador y el tiempo de viaje disminuye el número de vehículos necesarios cuando se busca disminuir el costo del operador, lo cual disminuye a su vez la calidad del servicio; por el contrario, pero se requieren más vehículos para ofrecer un viaje con un mayor nivel de servicio. El punto medio entre ambos factores significa un equilibrio para los dos grupos de interés, usuario y operador. Cabe resaltar que la complejidad del sistema depende de muchas más variables que afectan el funcionamiento y asignación de los vehículos, por lo cual escenarios que varíen los parámetros de entrada y diferentes algoritmos con mayor cantidad de funciones objetivo y mayor detalle tendrán diferentes aproximaciones.

Estas conclusiones están basadas en una política que utiliza los vehículos según la cantidad de pasajeros proyectados y que a su vez se basa en que los buses inician y terminan un recorrido en los extremos de la red estudiada. Política que puede ser mejorada con las optimizaciones de los puntos de inicio y fin de las rutas, un estudio más detallado de los recorridos que cada bus debe realizar, la creación o eliminación de servicios y frecuencias, y el control de la distribución de los pasajeros en el interior del sistema, entre otros factores que pueden ser implementados por el operador para incrementar el nivel de servicio para el usuario y disminuir los costos de operación.

15. BIBLIOGRAFÍA

- Abduljabbar, R., D  a, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. *Sustainability (Switzerland)*, 11(1) doi:10.3390/su11010189
- Anguita Rodr  guez, F., Duarte Monedero, B., & Flores Ureba, S. (2014). *Situaci  n actual del transporte p  blico urbano: La visi  n de las empresas operadoras* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.iedee.2013.10.003
- Aparicio, L. S. (2010). El impacto del transporte en el ordenamiento de la ciudad: El caso de Transmilenio en Bogot  . *Territorios: Revista De Estudios Regionales Y Urbanos*, (22), 33-64. Retrieved from
- Arias-Rojas, J., Jim  nez, J. F., & Montoya-Torres, J. (2012). Solving of school bus routing problem by ant colony optimization. *Revista. EIA*, (17), 193-207. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=80417513&lang=es&site=ehost-live>
- ARTEAGA ARREDONDO, I., MARCELA GARC  A, D., ERNESTO GUZM  N GUZM  N, C., & MARIO MAYORGA, J. (2017). De los puentes peatonales: Lecciones del brt de Bogot  . *Urbano*, 20(36), 104-114. doi:10.22320/07813607.2017.21.36.09
- Arbex, R. O., & da Cunha, C. B. (2015). Efficient transit network design and frequencies setting multi-objective optimization by alternating objective genetic algorithm doi:<https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trb.2015.06.014>
- Bagheri, M., Akbari, A., & Mirbagheri, S. A. (2019). Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: A critical review. *Process Safety & Environmental Protection: Transactions of the Institution of Chemical Engineers Part B*, 123, 229-252. doi:10.1016/j.psep.2019.01.013
- Baaj, M. H., & Mahmassani, H. S. (1995). *Hybrid route generation heuristic algorithm for the design of transit networks* doi:[https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/0968-090X\(94\)00011-S](https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/0968-090X(94)00011-S)
- Balya, M. I., & Kumar, R. (2017). Walking feeder mode services choice analysis for integration of bus rapid transit system: A case study. *International Journal for Traffic & Transport Engineering*, 7(4), 487-497. doi:10.7708/ijtte.2017.7(4).07
- BARFOD, M. B. (2018). Supporting sustainable transport appraisals using stakeholder involvement and mcda. *Transport (16484142)*, 33(4), 1052-1066. doi:10.3846/transport.2018.6596
- Bingfeng, S., Ming, Z., Xiaobao, Y., & Ziyu, G. (2017). *Bi-level programming model for exclusive bus lanes configuration in multimodal traffic network* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trpro.2017.05.449
- Boonkleaw, A., Suthikarnnarunai, N., & Srinon, R. (2010). Strategic planning for newspaper delivery problem using vehicle routing algorithm with time window

- (VRPTW). *Engineering Letters*, 18(2), 183-192. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=50994867&lang=es&site=ehost-live>
- Bouchra, B., Btissam, D., & Mohammad, C. (2018). Multiobjective local search based hybrid algorithm for vehicle routing problem with soft time windows. Paper presented at the *Big Data*, 312-325.
- Base de datos registros de entradas al sistema Transmilenio, (2017)
- BCallejo, M. A. (2009). Optimización del diseño de líneas de autobús. aplicación a donostia-san sebastián.
- Cascetta, E., & Coppola, P. (2016). *Assessment of schedule-based and frequency-based assignment models for strategic and operational planning of high-speed rail services* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.tra.2015.09.010
- Castaño, J. A., Diez, A. E., & Giraldo, D. P. (2018). Análisis del consumo energético de un sistema de transporte público de pasajeros con tracción eléctrica desde una perspectiva sistémica. *Información Tecnológica*, 29(1), 147-160. doi:10.4067/S0718-07642018000100016
- Cepeda, M., Cominetti, R., & Florian, M. (2006). *A frequency-based assignment model for congested transit networks with strict capacity constraints: Characterization and computation of equilibria* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trb.2005.05.006
- Cervero, R. (2013). Bus rapid transit (BRT): An efficient and competitive mode of public transport.
- Cesafsky, L. (2017). How to mend a fragmented city: A critique of 'infrastructural solidarity'. *International Journal of Urban & Regional Research*, 41(1), 145-161. doi:10.1111/1468-2427.12447
- Chen, Q. (2014). Global optimization for bus line timetable setting problem. *Discrete Dynamics in Nature & Society*, , 1-9. doi:10.1155/2014/636937
- Chen, X. (., Li, Z., Jiang, H., & Li, M. (2015). Investigations of interactions between bus rapid transit and general traffic flows. *Journal of Advanced Transportation*, 49(3), 326-340. doi:10.1002/atr.1268
- Chen, X., Hellinga, B., Chang, C., & Fu, L. (2015). Optimization of headways with stop-skipping control: A case study of bus rapid transit system. *Journal of Advanced Transportation*, 49(3), 385-401. doi:10.1002/atr.1278
- Cheng, H., Huang, H., & Zhang, K. (2019). *Study on the influence of bus front-end intrusion-free distance to the bus moving characteristics* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.matcom.2018.05.019
- A critical evaluation of the colombian model of mass or bus rapid transit systems. (2010). *Ingeniería Y Competitividad*, 12(2), 23-40. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=60988198&lang=es&site=ehost-live>

- Cui, S., Zhao, H., Chen, H., & Zhang, C. (2018). The mobile charging vehicle routing problem with time windows and recharging services. *Computational Intelligence & Neuroscience*, , 1-11. doi:10.1155/2018/5075916
- Davis, J. (2013). BRT: Transit mode of choice? *Planning*, 79(5), 28. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=f5h&AN=87711772&lang=es&site=ehost-live>
- Deng, T., Ma, M., & Wang, J. (2013). Evaluation of bus rapid transit implementation in china: Current performance and progress. *Journal of Urban Planning & Development*, 139(3), 226-234. doi:10.1061/(ASCE)UP.1943-5444.0000150
- Deng, T., & Nelson, J. (2011). Recent developments in bus rapid transit: A review of the literature. *Transport Reviews*, 31(1), 69-96. doi:10.1080/01441647.2010.492455
- Departamento de Transporte y Desarrollo Urbano del Banco Mundial. (2002). *Ciudades en movimiento*. ().
- De Medeiros, R., Rodríguez-Ramos, R., Guinovart-Díaz, R., Bravo-Castillero, J., Otero, J. A., & Tita, V. (2015). Numerical and analytical analyses for active fiber composite piezoelectric composite materials. *Journal of Intelligent Material Systems and Structures*, 26(1), 101-118. doi:10.1177/1045389X14521881
- Duarte, F., & Ultramari, C. (2012a). Making public transport and housing match: Accomplishments and failures of curitiba's BRT. *Journal of Urban Planning & Development*, 138(2), 183-194. doi:10.1061/(ASCE)UP.1943-5444.0000107
- Ebrahimnejad, A., & Verdegay, J. L. (2018). A new approach for solving fully intuitionistic fuzzy transportation problems. *Fuzzy Optimization & Decision Making*, 17(4), 447-474. doi:10.1007/s10700-017-9280-1
- Elver A. Bermeo Muñoz, & Jaime Hernán Calderón Sotero. (2009, Diseño de un modelo de optimización de rutas de transporte. *El Hombre Y La Máquina*, , 52-67. Retrieved from <http://www.redalyc.org/pdf/478/47811604005.pdf>
- Engelbrecht, L., & Bekker, J. (2012). *A decision support model for passenger capacity design of brt stations* South African Journal of Industrial Engineering (University of Pretoria). Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=76405947&lang=es&site=ehost-live>
- Eriksson, L. (2017). The role of organizational identities for policy integration processes - managing sustainable transport development. *Public Organization Review*, 17(4), 525-544. doi:10.1007/s11115-016-0348-0
- Feng, X., Zhu, X., Qian, X., Jie, Y., Ma, F., & Niu, X. (2019). *A new transit network design study in consideration of transfer time composition* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trd.2018.03.019
- Fernandes, G., Rodrigues, Joel J. P. C., Carvalho, L. F., Al-Muhtadi, J., & Proença, M. L. (2019). A comprehensive survey on network anomaly detection. *Telecommunication Systems*, 70(3), 447-489. doi:10.1007/s11235-018-0475-8

- GALINDRES GUANCHA, D., ARMANDO, SOTO MEJÍA, J., ADALBERTO, & ESTRADA MEJÍA, S. (2016). Asignación de frecuencias apropiadas a Través de un modelo multiobjetivo para un sistema brt. *Revista EIA*, 13(26), 141-152. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=122708552&lang=es&site=ehost-live>
- García Schilardi, M. E. (2013). Organización territorial urbana Y transporte Público masivo: Metrópolis del gran Mendoza. *Urbano*, 16(28), 50-57. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=110614025&lang=es&site=ehost-live>
- Gkiotsalitis, K., & Alesiani, F. (2019). Robust timetable optimization for bus lines subject to resource and regulatory constraints doi: <https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.tre.2019.05.016>
- Guihaire, V., & Hao, J. (2008). *Transit network design and scheduling: A global review* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.tre.2008.03.011
- HARAN, J. M. S. (2008). Estrategia de precios: Aplicación a un servicio ferroviario. *Revista Ingeniería Industrial*, 7(1), 49-58. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=35409350&lang=es&site=ehost-live>
- Hernández Ortiz, Y. A. (2016). Diseño de un sistema de ruteo de vehículos con múltiples depósitos en empresas de transporte de carga por carretera. Retrieved from <http://repository.udistrital.edu.co/bitstream/11349/3600/1/Hern%C3%A1ndezOrtizYimyAlexander2016.pdf>
- Hidalgo, D. (2005). Comparación de alternativas de transporte público masivo - una aproximación conceptual. *Revista De Ingeniería*, (21), 94-105. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=20001262&lang=es&site=ehost-live>
- Hidalgo, D., Lleras, G., & Hernández, E. (2013). *Methodology for calculating passenger capacity in bus rapid transit systems: Application to the Transmilenio system in Bogotá, colombia* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.retrec.2012.06.006
- Huo, Y., Li, W., Zhao, J., & Zhu, S. (2018). Modelling bus delay at bus stop. *Transport (16484142)*, 33(1), 12-21. doi:10.3846/16484142.2014.1003324
- Ingenia UdeA (Producer), & Santa, C. (Director). (2018, Oct 18,). *Regresión logística*. [Video/DVD] Universidad de Antioquia. Retrieved from https://www.youtube.com/watch?v=_TQg8LvO5hk
- Jafarian, E., Razmi, J., & Baki, M. F. (2018). *A flexible programming approach based on intuitionistic fuzzy optimization and geometric programming for solving multi-objective nonlinear programming problems* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.eswa.2017.10.030

- Jha, S. B., Jha, J. K., & Tiwari, M. K. (2019). *A multi-objective meta-heuristic approach for transit network design and frequency setting problem in a bus transit system* doi:<https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.cie.2019.02.025>
- Jiménez Narváez, M. L., & Soto Mejía, J. A. (2016). Utilización de tarjetas inteligentes para estimar matrices origen-destino. Aplicación al sistema Megabus , pereira. *Ciencia E Ingeniería Neogranadina*, 26(2), 73-93. doi:10.18359/rcin.1661
- Kathuria, A., Parida, M., & Sekhar, C. R. (2017). Route performance evaluation of a closed bus rapid transit system using GPS data. *Current Science (00113891)*, 112(8), 1642-1652. doi:10.18520/cs/v112/i08/1642-1652
- Kermanshahi, S., Shafahi, Y., & Bagherian, M. (2015). Application of a new rapid transit network design model to bus rapid transit network design: Case study isfahan metropolitan area. *Transport (16484142)*, 30(1), 93-102. doi:10.3846/16484142.2013.816772
- Kiani Mavi, R., Zarbakhshnia, N., & Khazraei, A. (2018). *Bus rapid transit (BRT): A simulation and multi criteria decision making (MCDM) approach* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.tranpol.2018.03.010
- Kihwan, N. A. M., & PARK, M. (2018). Improvement of an optimal bus scheduling model based on transit smart card data in seoul. *Transport (16484142)*, 33(4), 981-992. doi:10.3846/transport.2018.6045
- Kittelson & Associates, I., Parsons Brinckerhoff, KFH Group, I., & Texas A&M, T. I. (2013). *Transit capacity and quality of service manual (3rd ed.)*. Washington D.C.: Transportation Research Board.
- Lee, G., & Chan, E. (2008). The analytic hierarchy process (AHP) approach for assessment of urban renewal proposals. *Social Indicators Research*, 89(1), 155. doi:10.1007/s11205-007-9228-x
- Lee, W., & Li, J. (2018). An eco-driving advisory system for continuous signalized intersections by vehicular ad hoc network. *Journal of Advanced Transportation*, , 1-12. doi:10.1155/2018/5060481
- Leyva López, J. C. (2011). *Análisis multicriterio para la toma de decisiones métodos y aplicaciones*. México, D.F.: Plaza y Valdés, S.A. de C.V. Retrieved from <http://ebookcentral.proquest.com/lib/bibliotecaustasp/detail.action?docID=3217661>
- Levinson, H., Zimmerman, S., Clinger, J., Rutherford, S., Smith, R. L., Cracknell, J., & Soberman, R. (2003). *Bus rapid transit. Transit Cooperative Research Program Report, 90*
- Li, D., & Yoon, S. W. (2015). A novel fill-time window minimisation problem and adaptive parallel tabu search algorithm in mail-order pharmacy automation system. *International Journal of Production Research*, 53(14), 4189-4205. doi:10.1080/00207543.2014.985392

- Li, S., Liu, R., Yang, L., & Gao, Z. (2019). *Robust dynamic bus controls considering delay disturbances and passenger demand uncertainty* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trb.2019.03.019
- Li, X., Hu, S., Fan, W., & Deng, K. (2018). Modeling an enhanced ridesharing system with meet points and time windows. *PLoS ONE*, 13(5), 1-19. doi:10.1371/journal.pone.0195927
- Lin, Q., Li, J., Du, Z., Chen, J., & Ming, Z. (2015). A novel multi-objective particle swarm optimization with multiple search strategies doi:https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.ejor.2015.06.071
- Lin, Y., Huang, C., Liao, Y., & Yeh, C. (2017). System reliability for a multistate intermodal logistics network with time windows. *International Journal of Production Research*, 55(7), 1957-1969. doi:10.1080/00207543.2016.1247997
- Llinás S, H., & Rojas A., C. (2005). Estadística descriptiva y distribuciones de probabilidad
- Linda Rocha, González Elsa, & Orjuela Javier. Una revisión al estado del arte del problema de ruteo de vehículos: Ev. *Evolución Histórica Y Métodos De Solución*. . En: *Ingeniería*, 16(2), 3-55 .
- Liu, S., Luo, X., & Jin, P. J. (2018). Improving bus operations through integrated dynamic holding control and schedule optimization. *Journal of Advanced Transportation*, , 1-18. doi:10.1155/2018/9714046
- Liu, Y., Liu, Z., & Jia, R. (2019). *DeepPF: A deep learning based architecture for metro passenger flow prediction* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trc.2019.01.027
- Loreto, Jiménez Lillo, Daniela. Análisis y pronósticos de demanda para telefonía móvil, D. Universidad de Chile, 2011. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/bibliotecaustasp/detail.action?docID=3198312>. Created from bibliotecaustasp on 2019-12-13 12:44:49.
- Liu, Z., Li, K., & Ni, Y. (2017). *Optimization of bus stops layout under the conditions of coordinated control* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trpro.2017.05.190
- Lu, J., Feng, L., Yang, J., Hassan, M. M., Alelaiwi, A., & Humar, I. (2019). *Artificial agent: The fusion of artificial intelligence and a mobile agent for energy-efficient traffic control in wireless sensor networks* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.future.2018.12.024
- Lyu, Y., Chow, C., Lee, V. C. S., Ng, J. K. Y., Li, Y., & Zeng, J. (2019). *CB-planner: A bus line planning framework for customized bus systems* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trc.2019.02.006
- Maaïke Hoogeboom, Dullaert Wout, Lai David, & Vigo Daniele. (2018). Efficient neighborhood evaluations for the vehicle routing problem with multiple time windows. Retrieved from http://or.dei.unibo.it/sites/or.dei.unibo.it/files/VRPMTW_OnlineVersion.pdf

- Malone, K. M., Verroen, E., Korver, W., & Heyma, A. (2001). The scenario explorer for passenger transport: A strategic model for long-term travel demand forecasting. *Innovation: The European Journal of Social Sciences*, 14(4), 331-353. doi:10.1080/13511610120106138
- Márquez, L. G., & Monsalve, H. J. (2009). Modelo de demanda para un tren de pasajeros entre Tunja y Bogotá. *Revista De Ingeniería*, (30), 16-24. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=49050981&lang=es&site=ehost-live>
- Mauttone, A., Cancela, H., & Urquhart, M. Diseño y optimización de rutas y frecuencias en el transporte colectivo urbano, modelos y algoritmos.
- Mauttone, A., & Urquhart, M. E. (2009). A route set construction algorithm for the transit network design problem doi:<https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.cor.2008.09.014>
- Miao, Q., Welch, E. W., & Sriraj, P. S. (2019). Extreme weather, public transport ridership and moderating effect of bus stop shelters doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.jtrangeo.2018.11.007
- Molinero, A., & Sanchez, L. (1997). Transporte público: Plantación, diseño, operación y administración. Toluca, Mexico: Universidad Autónoma del Estado de México.
- Mulley, C., & Tsai, C. (. (2017). Impact of bus rapid transit on housing price and accessibility changes in sydney: A repeat sales approach. *International Journal of Sustainable Transportation*, 11(1), 3-10. doi:10.1080/15568318.2015.1106223
- Murata, M., Delgado Campos, J., & Suárez Lastra, M. (2017). ¿Por qué la gente no usa el metro? efectos del transporte en la zona metropolitana de la ciudad de México Este artículo forma parte de la tesis que masanori murata realizó en el posgrado de geografía de la unam para obtener el grado de doctor en geografía. doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.14350/rig.56661
- Nesmachnow, S., Massobrio, R., Arreche, E., Mumford, C., Olivera, A. C., Vidal, P. J., & Tchernykh, A. (2019). Traffic lights synchronization for bus rapid transit using a parallel evolutionary algorithm doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.ijtst.2018.10.002
- Ng, K. K. H., Lee, C. K. M., Chan, F. T. S., & Lv, Y. (2018). Review on meta-heuristics approaches for airside operation research doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.asoc.2018.02.013
- Orozco Palacio, S., & Velásquez Fernández, M. (2015). Uber: La plataforma tecnológica que rompe con los esquemas tradicionales de la regulación en materia de transporte individual de pasajeros en Colombia. *Revista De Derecho Comunicaciones Y Nuevas Tecnologías*, (14), 1-36. doi:10.15425/redecom.14.2015.04
- Orrego Cardozo, J. P., Ospina Toro, D., & Toro Ocampo, E. M. (2016). Solución al problema de ruteo de vehículos con capacidad limitada (CVRP) usando una técnica metaheurística. *Scientia Et Technica*, 21(3), 225-233. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=122337598&lang=es&site=ehost-live>

- Parra Arias, Z., & Padilla Magdaleno, I. (2008). Indicadores de evaluación en la operación del servicio interurbano de pasajeros por ómnibus. una aplicación: Organización económica estatal de transporte (OEET) "mártires de girón". *Transporte Desarrollo Y Medio Ambiente*, 28(2-3), 21-24. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=49765758&lang=es&site=ehost-live>
- Pérez Kaligari, E., & Guerrero Rueda, W. J. (2015). Métodos de Optimización para el problema de ruteo de Vehículos con inventarios Y ventanas de tiempo duros. *Revista Ingeniería Industrial*, 14(3), 31-49. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=115205202&lang=es&site=ehost-live>
- Pérez Polanco, M. (2017). La congestión vial y el transporte masivo de pasajeros en el gran santo domingo. *Archivos De Arquitectura Antillana*, (62), 68-73. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=vth&AN=126630746&lang=es&site=ehost-live>
- Quesada, F. J. G., Graciani, M. A. F., Bonal, M. T. L., & Díaz-Mata, M. A. (1994). Aprendizaje con redes neuronales artificiales. *Ensayos: Revista De La Facultad De Educación De Albacete*, (9), 169-180.
- Quiñones-Grueiro, M., Prieto-Moreno, A., Verde, C., & Llanes-Santiago, O. (2019). *Data-driven monitoring of multimode continuous processes: A review* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.chemolab.2019.03.012
- Qiao, J., Zhou, H., Yang, C., & Yang, S. (2019). A decomposition-based multiobjective evolutionary algorithm with angle-based adaptive penalty doi:https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.asoc.2018.10.028
- Ramírez, S., Hernández, L., & Varela Hugo. (2015). *Diseño de rutas de transporte publico*
- Ríos-Mesías, M., Ramos-López, G., & García-Mora, G. (2011). Características de un sistema eléctrico de transporte masivo con tranvías. *Ingeniería Y Universidad*, 15(1), 91-104. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=75236928&lang=es&site=ehost-live>
- Rizelioğlu, M., & Arslan, T. (2019). *A comparison of LRT with an imaginary BRT system in performance: Bursa example* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.cstp.2019.01.007
- Robles, D., Ñañez, P., & Quijano, N. (2009). Control y simulación de tráfico urbano en Colombia: Estado del arte. *Revista De Ingeniería*, (29), 59-69. Retrieved from <https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=44459634&lang=es&site=ehost-live>

- Rodríguez, D. A., & Targa, F. (2004). Value of accessibility to bogotá's bus rapid transit system. *Transport Reviews*, 24(5), 587-610. doi:10.1080/0144164042000195081
- Rodríguez, D., & Vergel, E. (2013). Sistemas de transporte público masivo tipo BRT (bus rapid transit) y desarrollo urbano en américa latina. *Land Lines*, 25(1), 16-24.
- Rogat, J., Dhar, S., Joshi, R., Mahadevia, D., & Mendoza, J. C. (2015). Sustainable transport: BRT experiences from mexico and india. *WIREs: Energy & Environment*, 4(6), 564-574. doi:10.1002/wene.162
- Rogge, M., van der Hurk, E., Larsen, A., & Sauer, D. U. (2018). *Electric bus fleet size and mix problem with optimization of charging infrastructure* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.apenergy.2017.11.051
- Rojas-Galeano, S., & Rodríguez Garzón, C. A. (2014). A discrete model of TransMilenio station occupation: Representation and algorithms. *Revista De Ingeniería*, (40), 27-32. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=98421880&lang=es&site=ehost-live>
- Sabogal Cardona, O. A., Hincapié Zea, J. D., Santa Chávez, J. J., & Escobar, J. W. (2015). Modelos de Regresión lineal para Estimación de tiempos de viaje en sistemas de transporte masivo. *Ciencia E Ingeniería Neogranadina*, 25(1), 77-89. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eih&AN=103735651&lang=es&site=ehost-live>
- Sâmara Visentini, M., Bassi, d. A., Borenstein, D., & Kummer Neto, A. F. (2017). Resolución integrada de problemas do planejamento do transporte público: Foco na tabela de horarios E no escalonamiento de vehículos com frota heterogênea. *Brazilian Journal of Management / Revista De Administracio Da UFSM*, 10(5), 828-843. doi:10.5902/1983465916249
- Sánchez Toledano, D., Carrasco Díaz, D., & Sánchez Toledano, J. (2014). *Observatorio de costes y financiación del transporte urbano colectivo: Un programa de investigación* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.iedee.2013.05.001
- Sánchez-Galván, F., Bautista-Santos, H., Garay-Rondero, C., Gibaja-Romero, D., & Mora-Castellanos, C. (2017). Optimización de costos de transporte bajo el enfoque de teoría de juegos. estudio de caso. *Nova Scientia*, 9(19), 185-210. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=126554541&lang=es&site=ehost-live>
- Sang, K. K., & Jong, H. K. (2017). Statistical data preparation: Management of missing values and outliers. *Korean Journal of Anesthesiology*, 70(4), 407-411. doi:10.4097/kjae.2017.70.4.407
- Santos, A. S., & Ribeiro, S. K. (2013). *The use of sustainability indicators in urban passenger transport during the decision-making process: The case of rio de janeiro, brazil* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.cosust.2013.04.010

- Santos, G., Behrendt, H., & Teytelboym, A. (2010). *Part II: Policy instruments for sustainable road transport* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.retrec.2010.03.002
- Shang, H., Huang, H., & Wu, W. (2019). *Bus timetabling considering passenger satisfaction: An empirical study in beijing* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.cie.2019.01.057
- Silva, P. P. B., & Zuluaga, A. E. (2016). Revisión del estado del arte del problema de ruteo de vehículos con recogida y entrega (VRPPD). *Ingeniería Y Desarrollo*, 34(2), 463-482. doi:10.14482/inde.33.2.6368
- Silva, R. (2019). ¿método numérico, analítico y experimental: Concurrentes o complementarios en la ingeniería? Retrieved from <https://www.esss.co/es/blog/metodo-numerico-analitico-y-experimental-concurrentes-o-complementarios-en-la-ingenieria/>
- Singh, H., Kathuria, A., Sekhar, C. R., & Parida, M. (2017). Estimating capacity of hybrid bus rapid transit corridor. *Current Science (00113891)*, 113(8), 1586-1592. doi:10.18520/cs/v113/i08/1586-1592
- Steinrucke, M. (2011). An approach to integrate production-transportation planning and scheduling in an aluminium supply chain network. *International Journal of Production Research*, 49(21), 6559-6583. doi:10.1080/00207543.2010.528461
- Sun, L., Jin, J. G., Lee, D., Axhausen, K. W., & Erath, A. (2014). *Demand-driven timetable design for metro services* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trc.2014.06.003
- Tatarenko, T. (2019). *Stochastic learning in multi-agent optimization: Communication and payoff-based approaches* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.automatica.2018.10.001
- Thomas Marcelo (Producer), & Thomas Marcelo (Director). (2017, Agu 2.). Rede neural artificial (ANN) com solver do excel - iris dataset. [Video/DVD] Retrieved from <https://www.youtube.com/watch?v=Px8IRt5tRuQ>
- Tong, L. (., Zhou, L., Liu, J., & Zhou, X. (2017). *Customized bus service design for jointly optimizing passenger-to-vehicle assignment and vehicle routing* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trc.2017.09.022
- Toro O., E. M., Escobar Z., A. H., & Granada E., M. (2016). Literature review on the vehicle routing problem in the green transportation context. *Revista Luna Azul*, (42), 362-387. doi:10.17151/luaz.2016.42.21
- Umlauf, T., Galicia, L. D., Cheu, R. L., & Horak, T. (2016). Ridership estimation procedure for a transit corridor with new bus rapid transit service. *Journal of Advanced Transportation*, 50(4), 473-488. doi:10.1002/atr.1355
- Van Nes, R., & Bovy, P. H. L. (2000). Importance of objectives in urban transit-network design. *Transportation Research Record*, 1735(1), 25-34. doi:10.3141/1735-04

- Velásquez, J. M. (2009). Análisis de factores que inciden en la demanda del sistema Transmilenio en Bogotá, Colombia. *Revista De Ingeniería*, (30), 56-64. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=49050986&lang=es&site=ehost-live>
- Virojboonkiate, N., Chanakitkarnchok, A., Vateekul, P., & Rojviboonchai, K. (2019). Public transport driver identification system using histogram of acceleration data. *Journal of Advanced Transportation*, , 1-15. doi:10.1155/2019/6372597
- Wang, Y., D'Ariano, A., Yin, J., Meng, L., Tang, T., & Ning, B. (2018). *Passenger demand oriented train scheduling and rolling stock circulation planning for an urban rail transit line* doi://doi.org/10.1016/j.trb.2018.10.006
- Wood, A. (2015). Multiple temporalities of policy circulation: Gradual, repetitive and delayed processes of BRT adoption in south african cities. *International Journal of Urban & Regional Research*, 39(3), 568-580. doi:10.1111/1468-2427.12216
- Wu, J., Song, R., Wang, Y., Chen, F., & Li, S. (2015). Modeling the coordinated operation between bus rapid transit and bus. *Mathematical Problems in Engineering*, 2015, 1-7. doi:10.1155/2015/709389
- Wu, Y., Zhu, Y., & Cao, T. (2018). Mathematical analysis and an exact solution combined with preprocessing method for resynchronizing of bus timetable problem. *Mathematical Problems in Engineering*, , 1-10. doi:10.1155/2018/7832638
- Xia, T., Zhang, Y., Braunack-Mayer, A., & Crabb, S. (2017). Public attitudes toward encouraging sustainable transportation: An Australian case study. *International Journal of Sustainable Transportation*, 11(8), 593-601. doi:10.1080/15568318.2017.1287316
- Xu, W., Li, Z., & Ling, Q. (2018). *Robust decentralized dynamic optimization at presence of malfunctioning agents* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.sigpro.2018.06.024
- Yan, Y., Liu, Z., Meng, Q., & Jiang, Y. (2013). Robust optimization model of bus transit network design with stochastic travel time. *Journal of Transportation Engineering*, 139(6), 625-634. doi:10.1061/(ASCE)TE.1943-5436.0000536
- Yang, M., Sun, G., Wang, W., Sun, X., Ding, J., & Han, J. (2018). Evaluation of the pre-detective signal priority for bus rapid transit: Coordinating the primary and secondary intersections. *Transport* (16484142), 33(1), 41-51. doi:10.3846/16484142.2015.1004556
- Zhang, F., & Liu, W. (2019). *Responsive bus dispatching strategy in a multi-modal and multi-directional transportation system: A doubly dynamical approach* doi://doi-org.crai-ustadigital.usantotomas.edu.co/10.1016/j.trc.2019.04.005
- 吴影辉, 唐加福, & 宫俊. (2015). 考虑随机行驶时间的单线路公交时刻表设计优化模型. *Journal of Northeastern University (Natural Science)*, 36(10), 1393-1407. doi:10.3969/j.issn.1005-3026.2015.10.006

ANEXOS

Anexo 1. Transcripción entrevista Doctor Stephen P. Mattingly sobre temáticas de transporte urbano, el futuro del transporte y el estado del transporte en la ciudad de Arlington, Texas.

Anexo 2. Análisis de metodologías AHP (Excel).

Anexo 3. Análisis Estadístico (Excel).

Anexo 4. Análisis Estacionalidad (Excel).

TABLAS

Tabla 1. Tipología Vehículos Que Funcionan En Las Troncales.	13
Tabla 2. Procesos De Planeación Del Tránsito.....	23
Tabla 3. Clasificación De Modelos Matemáticos.....	24
Tabla 4. Dificultades Diseño Y Optimización de rutas.	25
Tabla 5. Modelos Para El Diseño Estratégico.....	26
Tabla 6. Modelos Para El Estudio Táctico Y Operacional.....	29
Tabla 7. Fundamentos IA.....	31
Tabla 8. Otras Técnicas IA.	31
Tabla 9. Algoritmos genéticos.....	32
Tabla 10. Técnicas para el análisis del tráfico	33
Tabla 11. Clasificación de enfoques VRP	35
Tabla 12. Fundamentos VRP.....	36
Tabla 13. Variantes VRP CVRP homogéneo.....	37
Tabla 14. Variantes VRP CVRP Heterogéneo	37
Tabla 15. Tipos de anomalías.....	40
Tabla 16. Escala de 9 puntos para la comparación pareada en AHP.....	49
Tabla 17. Jerarquización de criterios AHP.....	50
Tabla 18. Nivel de consistencia.	50
Tabla 19. Jerarquización en función del criterio multiobjetivo	51
Tabla 20. Nivel de consistencia de la matriz de Jerarquización en función del criterio multiobjetivo.	51
Tabla 21. Jerarquización en función del criterio Óptimos locales.	51
Tabla 22. Nivel de consistencia de la matriz de Jerarquización en función del criterio Óptimos locales.....	52
Tabla 23. Jerarquización en función del criterio Óptimos locales.	52
Tabla 24. Nivel de consistencia de la matriz de Jerarquización en función del criterio Óptimos locales.....	52
Tabla 25. Resultados Jerarquización por los diferentes criterios del método AHP	52
Tabla 26. Distancias entre Estaciones.....	56
Tabla 27. Tipo de vías	57
Tabla 28. Descripción de datos.....	58
Tabla 29. Estadística descriptiva de la demanda por mes.....	62
Tabla 30. Estadística descriptiva de la demanda por día de la semana	63
Tabla 31. Estadística descriptiva de la demanda por hora	64
Tabla 32. Estadística descriptiva de la demanda por estaciones.....	65
Tabla 33. Estadística descriptiva de toda la muestra.....	66
Tabla 34. Homogeneidad Kruskal Wallis por estaciones	69
Tabla 35. Homogeneidad Medias por estaciones.....	69
Tabla 36. Homogeneidad Kruskal Wallis por meses.....	70
Tabla 37. Homogeneidades medias por meses.....	70
Tabla 38. Homogeneidad Kruskal Wallis por día de la semana.....	70
Tabla 39. Homogeneidades medias por día de la semana.....	70
Tabla 40. Homogeneidad Kruskal Wallis por horas	71

Tabla 41. Homogeneidades medias por horas	71
Tabla 42. Coeficiente de Pearson Demanda vs Estaciones	72
Tabla 43. Coeficiente de Kendall Demanda vs Estaciones.....	73
Tabla 44. Correlación parcial demanda vs estaciones controlando la categoría mes	74
Tabla 45. Correlación parcial demanda vs estaciones controlando la categoría hora del día.....	74
Tabla 46. Correlación parcial demanda vs estaciones controlando la categoría día de la semana	74
Tabla 47. Coeficiente de Pearson: demanda vs. día de la semana.....	75
Tabla 48. Coeficientes de Kendall y Spearman: demanda vs. estaciones.....	75
Tabla 49. Correlación parcial de demanda vs. día de la semana (controlando mes)	76
Tabla 50. Correlación parcial de demanda vs. día de la semana (controlando estaciones).....	76
Tabla 51. Correlación parcial de demanda vs. día de la semana (controlando horas del día).....	76
Tabla 52. Coeficiente de Pearson (demanda vs. Hora)	77
Tabla 53. Coeficiente de Kendall y Spearman (demanda vs. hora).....	77
Tabla 54. Correlación parcial de la demanda vs. la hora controlando estaciones.	78
Tabla 55. Correlación parcial demanda vs. hora controlando fecha	78
Tabla 56. Correlación parcial de demanda vs. hora controlando día de la semana	78
Tabla 57. Correlación parcial demanda vs. hora controlando mes	79
Tabla 58. ANOVA. Análisis por estaciones	80
Tabla 59. Análisis posthoc. Demanda agrupada por estaciones	80
Tabla 60. Clúster estación	82
Tabla 61. ANOVA. Análisis por día de la semana.....	82
Tabla 62. Análisis PostHoc de la demanda, agrupada por días de la semana	83
Tabla 63. Clúster Día	84
Tabla 64. ANOVA – Análisis por hora.....	84
Tabla 65. Análisis PostHoc de la demanda, agrupada por horas de la semana	84
Tabla 66. Clúster por mes.....	85
Tabla 67. ANOVA – Análisis por mes	85
Tabla 68. Análisis posthoc. Demanda agrupada por meses	86
Tabla 69. Clúster por mes.....	87
Tabla 70. Variables del modelo.....	89
Tabla 71. Paradas por servicio	92
Tabla 72. Distancias de la red.....	92
Tabla 73. Tiempos de viaje de un vehículo a 60 km/h	93
Tabla 74. Demanda del modelo a las 7:00 a.m.....	94
Tabla 75. Matriz OD Red (número de pasajeros)	95
Tabla 76. Matriz de Utilidad ruta k10/L10	97
Tabla 77. Distribución de la demanda.....	97
Tabla 78. Matriz OD Ruta k10/L10.....	98

<i>Tabla 79. Clúster día semana con elección de día representativo.....</i>	<i>102</i>
<i>Tabla 80. Clúster estaciones Transmilenio.</i>	<i>103</i>
<i>Tabla 81. Grupos análisis de demanda</i>	<i>104</i>
<i>Tabla 82. Clúster estaciones Transmilenio.</i>	<i>108</i>
<i>Tabla 83. Clúster de estaciones de Transmilenio</i>	<i>109</i>
<i>Tabla 84. Muestras en el entrenamiento del perceptrón.....</i>	<i>115</i>
<i>Tabla 85. Error en el entrenamiento de la predicción.</i>	<i>116</i>
<i>Tabla 86. Resultado del perceptrón en la etapa de entrenamiento.....</i>	<i>116</i>
<i>Tabla 87. Muestras en el reentrenamiento del perceptrón.....</i>	<i>117</i>
<i>Tabla 88. Error y entropía en la etapa de reentrenamiento</i>	<i>117</i>
<i>Tabla 89. Resultado del perceptrón en la etapa de reentrenamiento</i>	<i>117</i>
<i>Tabla 90. Muestras en el conjunto de evolución de la predicción.....</i>	<i>117</i>
<i>Tabla 91. Error y entropía final sin reentrenamiento y con reentrenamiento.</i>	<i>117</i>
<i>Tabla 92. Estacionalidad de las muestras sin agrupación.</i>	<i>120</i>
<i>Tabla 93. Estacionalidad de la muestra agrupada por dos criterios.....</i>	<i>121</i>
<i>Tabla 94. Estacionalidad de la muestra agrupada por 3 criterios.</i>	<i>122</i>
<i>Tabla 95. Estructura de la información.</i>	<i>123</i>
<i>Tabla 96. Tiempos acumulados de los pasajeros en el sistema.</i>	<i>127</i>

FIGURAS

<i>Figura 1. Mapa Sistema Transmilenio.</i>	<i>11</i>
<i>Figura 2. Diagramas de paradas de las rutas en la troncal Portal El Dorado.</i>	<i>12</i>
<i>Figura 3. Criterio para la comparación de metodologías por la técnica AHP.....</i>	<i>49</i>
<i>Figura 4. Jerarquización de criterios AHP.....</i>	<i>50</i>
<i>Figura 5. Jerarquización por metodología AHP</i>	<i>53</i>
<i>Figura 6. Diagrama sistema de estaciones.....</i>	<i>55</i>
<i>Figura 7. Comportamiento de las entradas al sistema agrupadas por meses</i>	<i>59</i>
<i>Figura 8. Comportamiento de las entradas al sistema agrupadas por hora por semana.....</i>	<i>60</i>
<i>Figura 9. Comportamiento de las entradas al sistema agrupadas por horas del día</i>	<i>60</i>
<i>Figura 10. Comportamiento de las entradas al sistema agrupadas por horas del día.....</i>	<i>61</i>
<i>Figura 11. Histograma del nivel de pasajeros que entran al sistema en una hora.</i>	<i>66</i>
<i>Figura 12. Histograma del nivel de pasajeros que entran al sistema en una hora.</i>	<i>67</i>
<i>Figura 13. Estadística descriptiva entradas al sistema portal el dorado</i>	<i>67</i>
<i>Figura 14. Histogramas comportamiento de las entradas de pasajeros en la estación portal el dorado en el tipo de día 4. Agrupación criterios de día, estación y hora.....</i>	<i>88</i>
<i>Figura 15 Diagrama metodología propuesta aplicativo.....</i>	<i>91</i>
<i>Figura 16. Esquema análisis descriptivo.....</i>	<i>100</i>
<i>.Figura 17. Histogramas entradas agrupadas por día de la semana</i>	<i>102</i>
<i>Figura 18. Histogramas de datos agrupados por estaciones de la troncal El Dorado</i>	<i>103</i>
<i>Figura 19. Histogramas datos agrupados por el criterio de clúster dos criterios para las diferentes estaciones el día viernes.</i>	<i>104</i>
<i>Figura 20. Gráfico explicativo del modelo perceptrón para la clasificación de datos</i>	<i>110</i>
<i>Figura 21. Clúster en estaciones de Transmilenio.....</i>	<i>111</i>
<i>Figura 22. Tipos de datos en el campo muestra de la muestra de clúster.....</i>	<i>112</i>
<i>Figura 23.Campo muestral del tipo de estación 1, Datos normales y atípicos.....</i>	<i>114</i>
<i>Figura 24.Ejemplo clasificación de datos normales y atípicos según la variación de la media</i>	<i>115</i>
<i>Figura 25. Probabilidades de pertenecer a al conjunto de datos normales o atípicos</i>	<i>118</i>
<i>Figura 26. Tendencia y ciclicidad del promedio de entradas en una hora.</i>	<i>119</i>
<i>Figura 27. Tendencia y ciclicidad del total de pasajeros en el sistema.....</i>	<i>119</i>
<i>Figura 28. Tendencia y ciclicidad del promedio de pasajeros en una hora.....</i>	<i>120</i>
<i>Figura 29. Tendencia Clúster Portal Dorado- Día tipo 4.</i>	<i>121</i>
<i>Figura 30. Tendencia y Estacionalidad Clúster a las 5 pm</i>	<i>122</i>
<i>Figura 31. Pronóstico para la estación portal el dorado el viernes 29 de septiembre a las 5 pm.</i>	<i>124</i>
<i>Figura 32.Señal de rastre pronostico suavización exponencial alpha 0,37.....</i>	<i>125</i>

<i>Figura 33. Suavización Exponencial (Alpha 0,37).....</i>	<i>125</i>
<i>Figura 34. Concepto de rendimiento.....</i>	<i>126</i>
<i>Figura 35. Optimización de la flota para la Troncal Portal El Dorado. Función de costo.</i>	<i>128</i>
<i>Figura 36. Optimización de la flota para la Troncal Portal El Dorado. Función tiempo acumulado en el sistema por los usuarios.</i>	<i>129</i>
<i>Figura 37. Optimización costo del operador.</i>	<i>130</i>
<i>Figura 38. Escenario de optimización del tiempo de viaje</i>	<i>131</i>
<i>Figura 39. Escenario recomendado</i>	<i>132</i>

ILUSTRACIONES

<i>Ilustración 1. Estación Sistema BRT en la ciudad de Bogotá.</i>	<i>10</i>
---	-----------