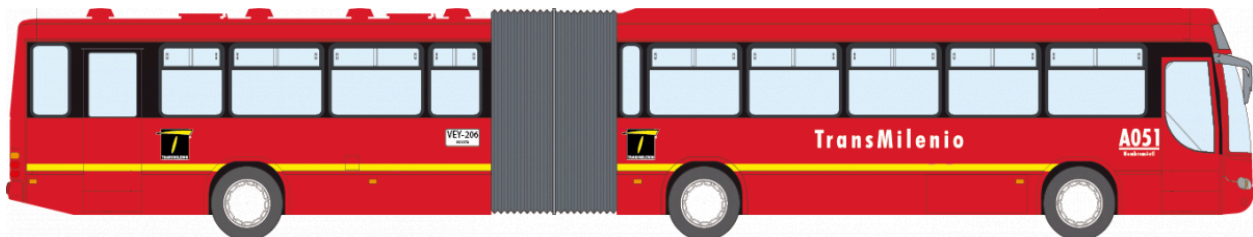


Transmilenio

Análisis Sistema Troncal

Noviembre 2022



DOCENTE

MARIA ISABEL SERRANO GÓMEZ - Pontificia Universidad Javeriana

AUTORES

MARIA PAULA LÓPEZ - mariap.lopezr@javeriana.edu.co

ESTEBAN CASTIBLANCO - je.castiblanco@javeriana.edu.co

DANIEL SANDOVAL - daniel_sandoval@javeriana.edu.co

CONTENIDO

INTRODUCCIÓN	2
OBJETIVOS	2
CONTEXTO	3
INDICADORES	4
DISEÑOS	6
HERRAMIENTAS	6
ANÁLISIS DE RESULTADOS	8
CONCLUSIONES	15
LECCIONES APRENDIDAS	15
REFERENCIAS	17

INTRODUCCIÓN

“El sector transporte en Colombia, está regulado por el Ministerio de Transporte, que es el organismo encargado de formular y adoptar las políticas, planes, programas, proyectos y regulaciones económicas del transporte, el tránsito y la infraestructura.” (Findeter 2021) . En particular, el transporte de la capital del país es el resultado de la articulación de los diferentes sistemas de transporte. El sistema de transporte de Bogotá es el más grande de Colombia y se realizan cerca de 12,2 millones de viajes diarios. De ellos el 42% se realiza en el Sistema Integrado de Transporte de Bogotá, el 26% en TransMilenio, el 22% en automóviles privados, el 12% a pie, el 3% en taxi, el 3% en motocicleta y el 2% en bicicleta. Desde la implementación del sistema de transporte masivo en el año 2000 no solo se han dado alternativas de movilidad para toda la ciudad, sino que también se han reducido las dificultades para personas con algún tipo de discapacidad. Actualmente se cuenta con paradas fijas con una distancia razonable entre unas y otras, los vehículos paran frente a las plataformas y el tiempo de espera para bajar del vehículo es suficiente

En conjunto con lo anterior, el propósito de esta investigación es reconocer los principales datos que brinda la empresa Transmilenio y llevarlos a un análisis que permita comprender el comportamiento de las personas dentro del transporte.

OBJETIVOS

A partir del uso de bases de datos del mes de Septiembre, se pretende lo siguiente:

1. Identificar las estadísticas de los principales indicadores.
2. Identificar los patrones de afluencia de las personas por Día y franja horaria.
3. Reconocer el movimiento de las personas en las estaciones, en principio agrupando la información por troncales.
4. Relacionar la distribución de las personas y el valor de la transacción con los dos anteriores objetivos.
5. Encontrar el Origen y Destino de las personas, y relacionarlos.
6. Evaluar la relación entre la cantidad de personas que recibe una estación y el número de estaciones con las que conecta, a partir del uso de un grafo.

CONTEXTO

La empresa Transmilenio S.A. tiene la misión de “Gestionar el desarrollo del Sistema Integrado de Transporte Público en Bogotá, en el marco del Plan Maestro de Movilidad, con estándares de calidad, dignidad y comodidad, con criterios de intermodalidad, con sostenibilidad financiera y ambiental, con una orientación hacia el mejoramiento de la calidad de vida de los usuarios y procurando la integración de la ciudad con la Región.” A pesar de que el servicio más conocido de esta empresa es el de Transmilenio, también existen servicios troncales de otros tres tipos: ruta fácil, expresos y alimentadores. Los servicios troncales circulan por corredores exclusivos iniciando y terminando su recorrido en los Portales o Estaciones de Cabecera. Los vehículos sólo se pueden detener a dejar y recoger pasajeros en las estaciones, conforme a una planeación previa y a un riguroso control en tiempo real. En el corredor troncal sólo opera el servicio de transporte masivo TransMilenio, es decir, está prohibida la circulación de vehículos de transporte colectivo. Las vías para servicios troncales corresponden a los carriles centrales de las principales avenidas de la ciudad. Estos carriles exclusivos se acondicionan especialmente para soportar el paso de los buses y se separan físicamente de los carriles de uso mixto, disponibles para circulación de vehículos particulares, camiones, taxis, etc. Con el fin de ordenar el tránsito y darle velocidad al Sistema TransMilenio, se construyen estaciones diseñadas especialmente para facilitar el acceso rápido y cómodo de los pasajeros. Las estaciones son los únicos puntos de parada de los servicios troncales para recoger y dejar pasajeros; son espacios cerrados y cubiertos, construidas en aluminio, concreto, acero y vidrio soplado, con taquillas a la entrada, y con acceso seguro para los usuarios a través de semáforos, puentes o túneles peatonales.

Actualmente el sistema cuenta con 1.114,4 km de vía troncal en operación, 11 troncales en operación, 138 estaciones, 9 portales y 53 patios garajes. Adicionalmente el Sistema troncal tiene 2.364 buses y 2.685 conductores en alimentación y 5.118 conductores en troncal. El componente zonal cuenta con 7.048 buses zonales, 2400 Kilómetros de cobertura, 38 patio talleres. Se tienen 655 buses eléctricos que son operados por 15.256 conductores. En cuanto a la infraestructura del Sistema, 3.500 puntos de acceso, 18 ascensores, 21 baños, 1.158 torniquetes y 3.668 puntos de recarga externa. Además el Sistema tiene a su servicio 22 cicloparqueaderos con 6.059 puestos en total. Espacios

seguros, cómodos y de fácil acceso para todas las personas que utilizan la bicicleta como una alternativa formal de transporte para la movilidad diaria.

En este informe solo se relacionan las estaciones.

INDICADORES

A partir de los conjuntos de datos tomados de los datos abiertos de la Alcaldía de Bogotá y del Transmilenio¹, identificamos las siguientes variables, las cuales cobran relevancia durante los análisis.

Datos Generales

Título Variable: Línea

Descripción: Nombre de la ruta principal a la que pertenece una estación de Transmilenio

Valores Unicos: 14

Título Variable: Nombre de Estación

Descripción: El sistema integrado de Transmilenio cuenta con estaciones distribuidas a lo largo de toda Bogotá. Cada estación tiene un identificador.

Valores Unicos: 158

Título Variable: Coordenadas Estación

Tipo: Geográfica

Descripción: Ubicación de cada una de las 158 estaciones de Transmilenio.

Título Variable: Número de Vagones | Estación

Descripción: Según el tamaño de la estación, cada estación cuenta con cierto número de vagones

Título Variable: Rutas

Descripción: Matriz de relación de los viajes que se pueden tomar desde cada estación.

Título Variable: Identificador Tarjeta

¹ Ver links de Bases de Datos en la Bibliografía

Descripción: Para ingresar al Sistema integrado de Transmilenio cada persona debe contar con una tarjeta que tiene un identificador único.

Título Variable: Tipo Tarjeta (**Nombre Perfil**)

Descripción: Una persona puede elegir personalizar su tarjeta y de esta forma recibir beneficios especiales en el pago de la tarifa de ingreso. Dentro de los tipos encontramos Anónimo, Adulto Mayor, Apoyo Ciudadano (Sisben) entre otros, con 7 valores únicos.

Datos Registros Ingresos de Usuarios:

Para poder hacer uso del sistema de Transmilenio cada usuario cuenta con una tarjeta con la cual pasa por un proceso de validación apenas entra al sistema integrado de Transmilenio. Este proceso nos brinda los siguientes datos de importancia:

Título Variable: ID Tarjeta de Ingreso

Descripción: Identificador único de la tarjeta que se usó para ingresar al sistema

Título Variable: Fecha Clearing

Tipo: Tiempo

Descripción: Hora y Día en que se registra el ingreso de la persona en el sistema integrado de Transmilenio

Título Variable: Estación de Entrada

Descripción: Nombre de la estación en la que se entró al sistema.

Datos Registros Salidas de Usuarios:

Para salir del sistema integrado de Transmilenio las personas deben pasar por un torniquete. Esto nos permite obtener los siguientes datos con un nivel detalle de intervalos de 15 minutos:

Título Variable: Estación de Salida

Descripción: Nombre de la estación en la que se salió del sistema.

Título Variable: Intervalo de Tiempo y fecha

Descripción: Contiene la información del momento a que hacen referencia los datos.

Título Variable: Cantidad de Salidas

Descripción: Número que nos indica cuántas personas registraron los torniquetes saliendo.

Nota: Estos son los datos principales que tuvimos en cuenta durante los análisis, hay disponibles datos de otros detalles más específicos como Saldo en las Tarjetas entre otros, los cuales ignoramos para mayor facilidad de comprensión para el lector.

DISEÑOS

- Debido a que las estaciones de Transmilenio existen en un lugar físico, se planteó usar gráficos que mostrarán su posición en el mapa y una vez se tuviese esto se pudiesen apreciar otros datos importantes de estas, como la relación que tienen con otras estaciones o la ruta a la que pertenecen.
- Se observó que gran cantidad de los datos utilizados están relacionados con información de espacio-tiempo. Por lo que se decidió que las gráficas de líneas son importantes para mostrar las tendencias que muestran los datos.
- Mediante gráficos de barras se pueden comparar las estaciones con respecto a alguna variable en específico.

Nota: Por favor ver la descripción de cada gráfica para obtener información más a profundidad de cada gráfica.

HERRAMIENTAS

Visualización:

Power Bi: Servicio de análisis de datos de Microsoft orientado a proporcionar visualizaciones interactivas y capacidades de inteligencia empresarial. Nos permite crear desde gráficos simples como Líneas de Tiempo y Gráficos de Barras hasta otros más complejos como mapas. Elegimos Power Bi porque cuenta con una interfaz fácil y práctica de usar. Además también cuenta con un robusto sistema de manipulación de datos.

Gephi: Es un software open-source de análisis de redes y visualización. En nuestro caso es de utilidad porque podemos relacionar las estaciones de Transmilenio a un sistema de redes y mostrar de esta manera relaciones entre las estaciones. También permite ubicar los puntos de la red geográficamente.

Manipulación de Datos:

Excel: Hoja de cálculo desarrollada por Microsoft que permite la manipulación de datos mediante una interfaz gráfica, práctica y rápida. Actualmente Excel es un estándar en muchas industrias. Aparte de sus funciones de cálculo también cuenta con la posibilidad de realizar gráficas.

Python: Lenguaje de alto nivel de programación interpretado cuya filosofía hace hincapié en la legibilidad de su código. Es ampliamente usado en el sector de las ciencias de datos gracias a la facilidad de uso y la comunidad que ha creado paquetes eficientes para el manejo de datos como **Pandas**.

Pandas: Librería de Python especializada en la manipulación y el análisis de datos. Ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales, es como el Excel de Python.

Almacenamiento y Cómputo:

Google Drive: Google Drive es un servicio de alojamiento de archivos de Google

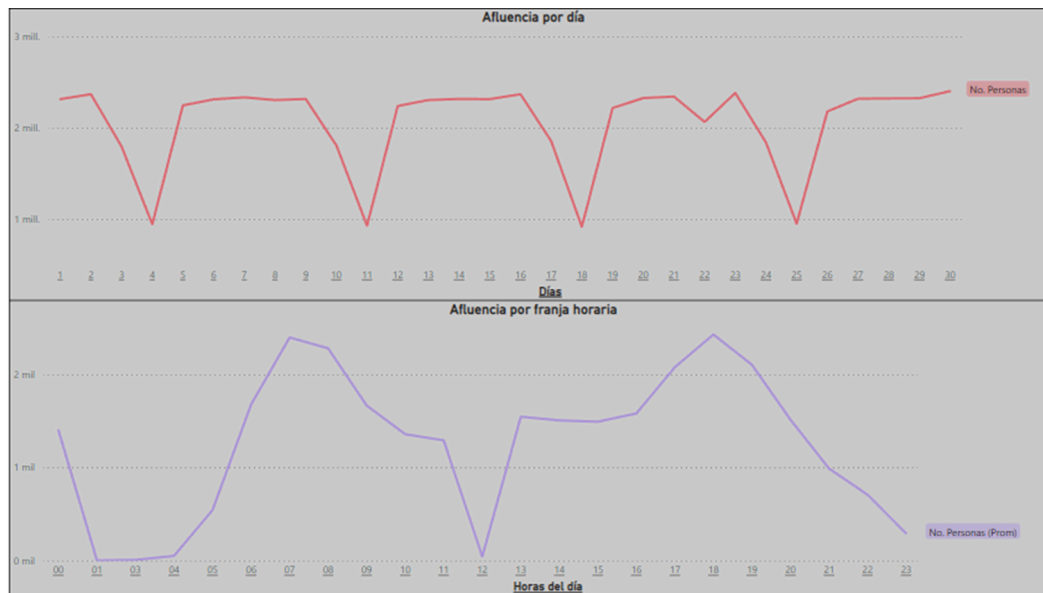
Jupyter Notebook: Formato de archivo para ejecutar código de python por partes e ir visualizando el resultado de cada parte en el navegador. También permite añadir imágenes y descripciones con uso de markdown.

Google Colab: Aunque buena parte de los análisis los corrimos usando nuestros computadores personales, Google Colab nos permite acceder a máquinas virtuales con gran velocidad para correr Jupyter Notebooks.

ANÁLISIS DE RESULTADOS

1. Análisis Mensual - Septiembre 2022

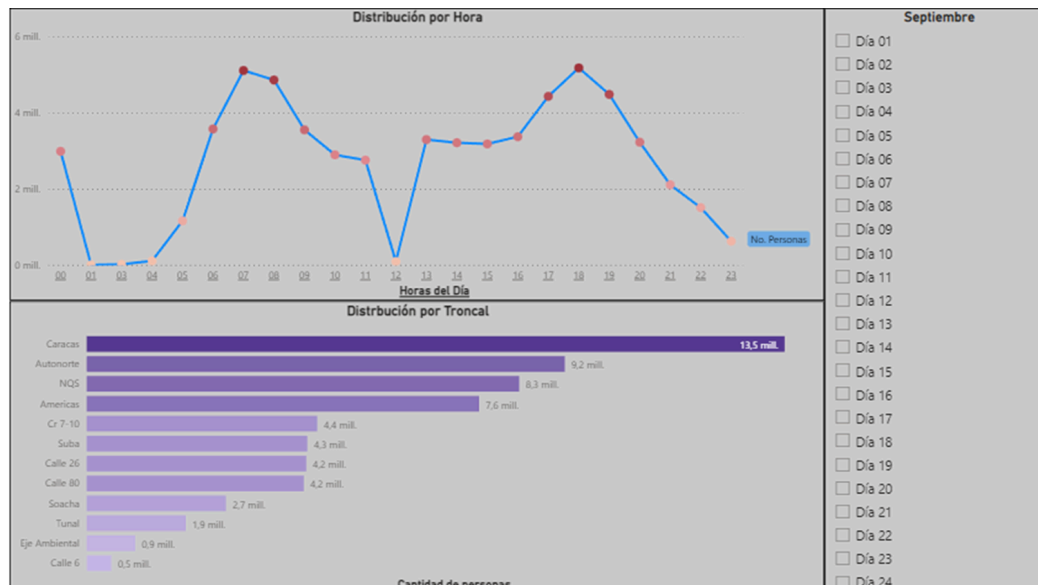
En primera instancia se identificaron los patrones de afluencia de las personas por Día y franja horaria a lo largo del mes de Septiembre. La gráfica de arriba en color rojo representa la afluencia por día, y la de abajo como un promedio por franja horaria.



Para resaltar:

- La afluencia por día muestra un comportamiento obvio por días de la semana, donde se presenta un comportamiento común entre lunes y viernes, y una baja los fines de semana.
- Más de 2 millones de personas usan Transmilenio entre semana.
- De igual manera con el análisis en el punto a, la afluencia por franja horaria identifica las horas pico de las personas, entre las 6-8 AM y las 5-7 PM.
- Asimismo un punto mínimo en la gráfica en las horas del mediodía donde las personas ya han llegado a su destino.

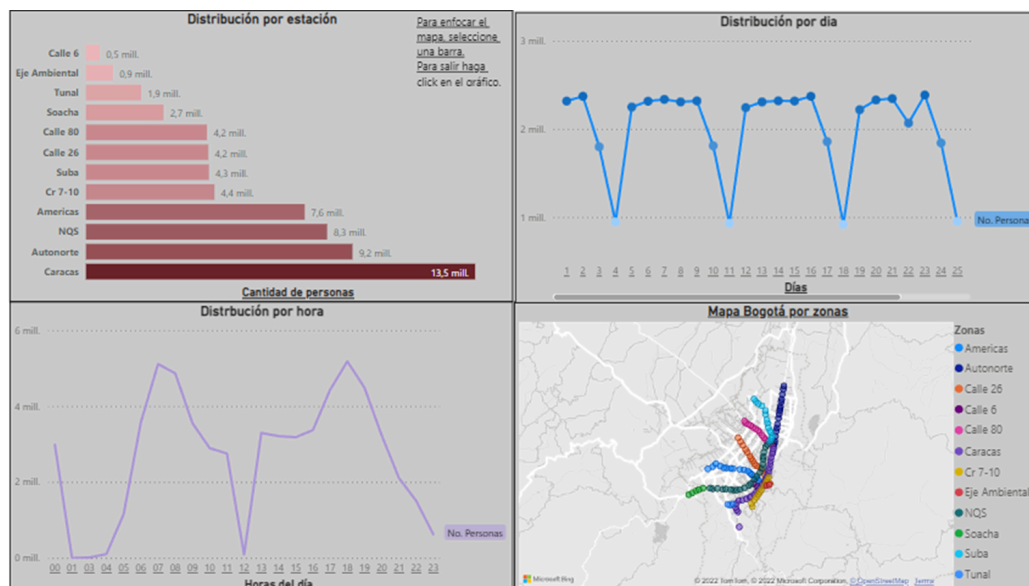
La siguiente gráfica muestra el movimiento de las personas en las estaciones, en principio agrupando la información por líneas troncales. Lo anterior en función a los 30 días en Septiembre y que identifica también la distribución por hora.



Para resaltar:

- Por distribución Troncal, la Caracas es el lugar en donde transitan más personas. En cuanto a la Autonorte y la NQS donde son las siguientes en cuanto a la afluencia de cantidad de personas no tienen un comportamiento predefinido, hay días en los que una tiene más movimiento que la otra.
- La distribución por hora mantiene el mismo comportamiento a lo largo de los días.

Por último, la siguiente gráfica muestra una distribución por hora y día, además de una visualización en mapa, en función a las líneas troncales en todo Septiembre.

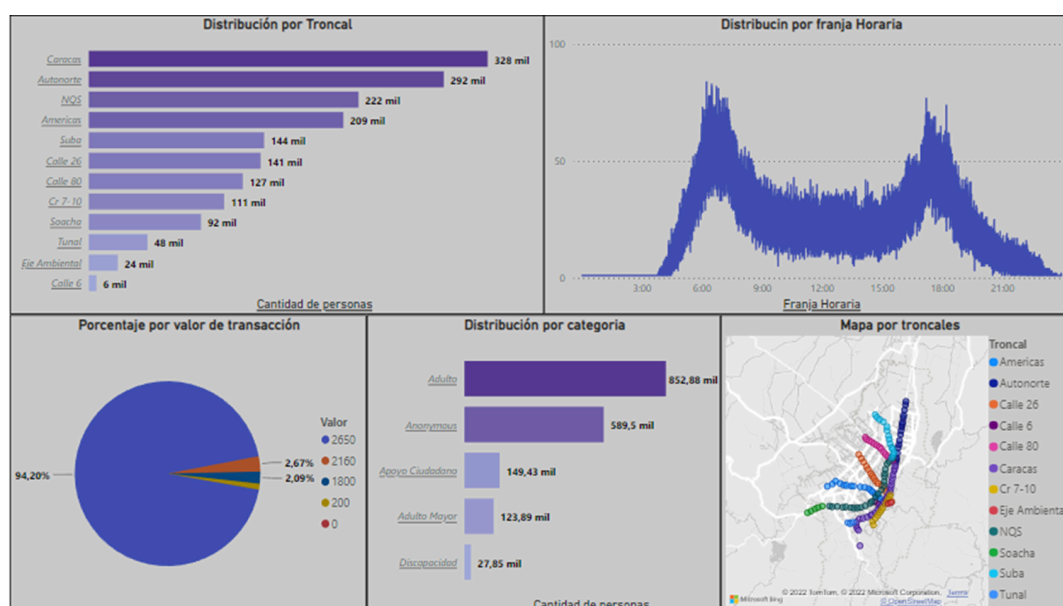


Para resaltar:

- Las troncales Tunal, Soacha y Américas presentan un pico de movimiento de personas en horas de la tarde noche, indicando la posibilidad de que alrededor de estas troncales trabajan una gran cantidad de personas, mas no que viven cerca.
- La distribución por día mantiene un mismo comportamiento.

2. Análisis Diario - 30 de Septiembre 2022

En primera instancia se identificó la distribución de las personas y el valor de la transacción en relación a las troncales y a una franja horaria.

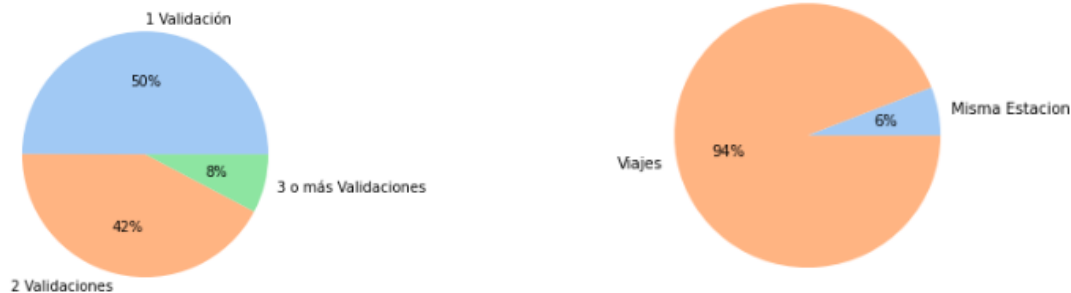


Para resaltar:

- Los valores de \$0 y \$200 identifican los transbordos entre SITP y Transmilenio que en su mayoría se ubican en la población adulta. El valor de \$1800 está guiado por el Incentivo del Sisbén. La Tarifa Preferencial para el Adulto Mayor es de 2160. Y el valor del pasaje en troncal sin descuento es de \$2.650 COP.
- Aunque los valores están dados por una distribución de ingresos y de edad, de acuerdo a los datos brindados, el valor por transacción no siempre coincide por su tarifa. Es decir, por ejemplo que hay adultos mayores que siguen pagando un pasaje de \$2.650.
- De acuerdo a la distribución por categoría como lo es una persona en situación de discapacidad, su franja horaria no es muy recurrente en las horas pico. Hay cantidades cercanas entre las 6 AM y las 6 PM.

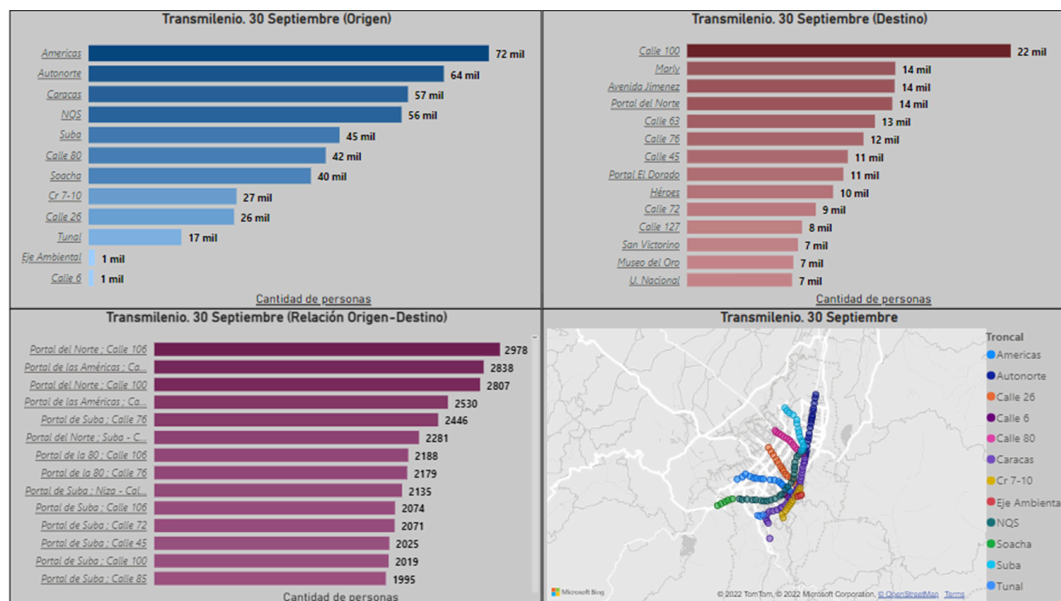
2. Análisis Viajes - 30 de Septiembre 2022

En las siguientes dos gráficas observamos que una muestra representativa del 42% son personas que usan la tarjeta 2 veces al día. De esas personas la mayoría (94%) usa la tarjeta en estaciones diferentes al día.



De aquí podemos realizar suposición (Hipótesis) de que la estación en la que primero usa la tarjeta es la estación en que inicia su ruta una persona y en la que hace la segunda validación es la estación en la que termina su ruta (Y la persona tiene la misma ruta pero al revés de regreso)

Nota: Las tarjetas de 1 sola validación en el futuro se puede relacionar con la información de validaciones en los buses que no se toman dentro de las estaciones de transmilenio.




Para resaltar:

- Las estaciones más influyentes se encuentran a los extremos de las rutas marcadas con color en la gráfica anterior.
- Un punto a resaltar es que las personas van a un conjunto mucho más reducido de estaciones al comenzar su viaje. Y de regreso el uso de más estaciones.
- El principal destino de las personas es la estación de la Calle 100 esto puede ser debido a que es zona de oficinas y médicos. Además sirve como límite entre las localidades de Usaquén con Chapinero, y Suba con Barrios Unidos.
- La estación de las Américas es la que más cuenta con
- La estación de las Américas se encuentra en una zona residencial donde muchas personas no cuentan con el beneficio de tener carro propio. Por lo que muchas personas inician su viaje en esta estación.
- La estación Autonorte también presenta gran cantidad de inicios de viajes posiblemente porque es la entrada a Bogotá de muchas personas que viven fuera de Bogotá ingresan por aquí.

3. Análisis del sistema como una red



 Grafo estaciones Transmilenio.pdf

Para hacer el análisis del sistema de Transmilenio como una red se utilizó un enfoque de red completa donde los actores fueron las estaciones de Transmilenio y las relaciones las rutas (servicios troncales) que componen el sistema. Los enlaces se caracterizan por

representar relaciones no dirigidas y su peso está representado por el grosor del enlace. El grado de los nodos se representa mediante la dimensión del tamaño que cada uno toma.

Información proporcionada por el grafo:

Densidad: 0.037, Diámetro: 12.

La primera métrica representa qué tan cerca está el grafo de ser completo. Es decir, que tan bien conectados están los nodos entre sí.

Un grafo es completo cuando tiene todas las aristas posibles y su densidad toma un valor igual a uno. La densidad determina la proporción de aristas que el grafo posee. Entre mayor es su densidad, significa que está más cerca al número máximo de aristas posibles, es decir, a las que tendría si fuese un grafo completo. En este caso a pesar de que no es coherente ni posible pretender que el sistema de Transmilenio sea un grafo completo los valores del diámetro y la densidad proporcionan aspectos positivos y negativos sobre el sistema. Entre ellos, el diámetro indica que aunque es sistema está distribuido a lo largo de distintas zonas de la ciudad, si una persona quisiera desplazarse de un punto límite de la ciudad a otro podría tomar hasta 12 rutas diferentes para hacerlo. Esto representa un aspecto positivo puesto que muestra la variedad de opciones que tiene una persona para ir de un punto a otro, pero a su vez expone la complejidad de la red y la dificultad que puede representar el uso del sistema para un nuevo usuario. En cuanto a la densidad, aunque no se pretende ni el objetivo del sistema es que llegue a ser una red completa, una densidad tan baja es un indicador de que las estaciones podrían estar mejor conectadas entre sí para descongestionar el sistema y mejorar el tiempo de espera y/o viaje de los usuarios. La alternativa para esto es la creación de nuevas rutas.

Grado medio: 5.473, Grado medio con pesos: 16.405, Longitud media de camino: 4.5

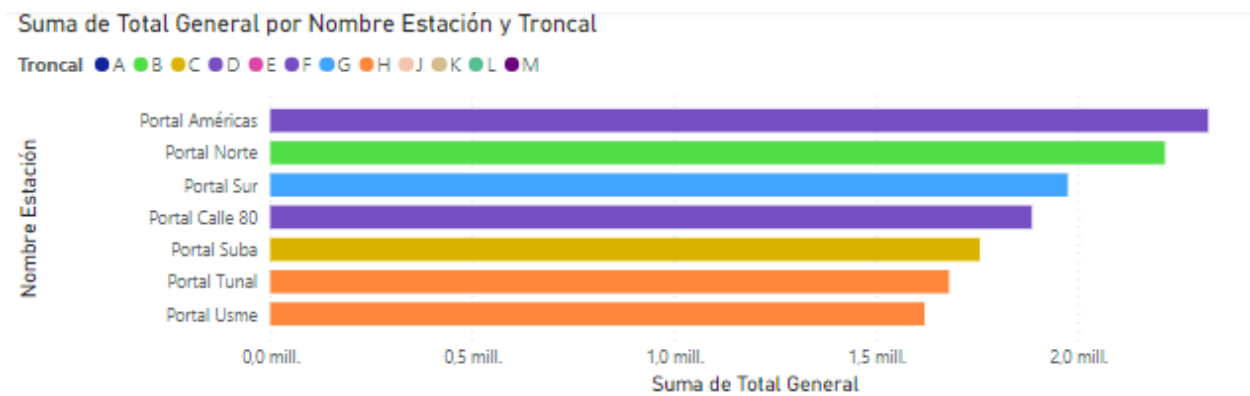
Estas métricas indican que en promedio, cada estación se conecta con otras 5 estaciones, pero si se toma en cuenta el número de veces que aparece un mismo enlace (dos estaciones pueden estar conectadas por rutas diferentes que representan un mismo enlace) se aumenta en más del triple la influencia promedio de las estaciones tienen en la red. La longitud media indica que la distancia media más corta entre dos estaciones es de 4.5 rutas.

Puentes

El grafo permite ver que las estaciones de Ricaurte y Avenida Jiménez tienen un rol de puente en la red ya que además de enlazar diferentes zonas de la ciudad, en ambas se

presenta que pertenecen a más de una troncal. Este hecho también explica el elevado número de enlaces que tienen respecto a estaciones vecinas o de sus mismas troncales.

Además de la información proporcionada por las métricas del grafo, esta visualización permitió evaluar si existe una relación entre la cantidad de personas que recibe una estación al mes y el número de estaciones con las que se conecta (una de las hipótesis que se tenía al principio del proyecto). Después de contrastar el tamaño de las aristas con la suma del total de ingresos durante un mes a cada estación que se representa con una gráfica de barras podría decirse que la hipótesis no es correcta. Sin embargo esto puede ser explicado por el hecho de que la información de los ingresos se toma únicamente del primer ingreso que realiza un usuario a una estación troncal, lo que imposibilita saber qué ruta tomó un usuario o un grupo de usuarios en concreto ni dónde finalizó su recorrido.



Para poder decir que la hipótesis es cierta, la información en el gráfico de barras tendría que mostrar las estaciones de Ricaurte y Avenida Jiménez en los primeros lugares pero esto no sucede.

CONCLUSIONES

- Como se mencionó en la introducción, cerca del 22% de los viajes diarios en Bogotá se realizan en Transmilenio (más de 2 millones de personas evidenciado en el análisis de este informe). Donde posiblemente en un día entre semana recaudan más de 2.650 millones de pesos.
- Transmilenio no tiene de referencia un diccionario para diferentes variables. Por lo que sus bases de datos no tiene muy buena concordancia y genera conflicto a la hora de realizar relaciones. Por ejemplo en las variables de estaciones, los nombres tienden a cambiar entre Mayúsculas, Minúsculas, actualización de texto entre otras.
- A pesar de que Transmilenio cuenta con grandes bases de datos, la información en ellas no permite realizar un análisis con resultados concluyentes puesto que no se recolectan los datos necesarios para ello. Entre los datos que permitirían esto estaría no solamente la hora y ubicación de ingreso sino también la ruta tomada por el usuario. Para esto sería necesario que los torniquetes no estuviesen ubicados en las estaciones sino en los buses de Transmilenio como es el caso de los buses zonales y duales. Dado que la posibilidad de implementar esto no es realista, una alternativa podría ser que se tuviese que utilizar la tarjeta tanto para el ingreso como para la salida de una estación.

LECCIONES APRENDIDAS

- Uso de la herramienta de análisis, manipulación y visualización de datos dentro de las que se encuentran Power Bi, Python, Pandas, Excel, y Gephi.
- Proceso de creación de un proyecto de visualización de datos, desde la lluvia de ideas, análisis exploratorio de datos, identificación de objetivos hasta llegar a conclusiones para presentarlas a una audiencia específica
- Proceso de limpieza de datos y manipulación de datos. Vital para obtener buenas gráficas.

- A la hora de realizar un proyecto de aplicación es fundamental tener entendimiento del negocio y conocer el contexto en el que este se desarrolla para poder establecer objetivos alcanzables y evitar incurrir en interpretaciones erróneas.
- Es importante tener conocimiento acerca de qué información contienen los conjuntos de datos para que los objetivos establecidos y preguntas puedan ser resueltos con la información que reposa allí.

REFERENCIAS

- Banca de Desarrollo Territorial, F. (2021, September 1). Estudio del Sector Transporte Colombiano. Repositorio Institucional Findeter. Retrieved September 17, 2022, from <https://repositorio.findeter.gov.co/handle/123456789/9708>
- habitat.aq.upm.es, B. C. F. S. (2002, June 26). Transmilenio: Sistema Integrado de Transporte Masivo (Bogotá, Colombia). Transmilenio: sistema integrado de transporte masivo (Bogotá, Colombia). Retrieved September 18, 2022, from <http://habitat.aq.upm.es/dubai/02/bp129.html#:~:text=Los%20objetivos%20principales%20con%20vistas,ancianos%2C%20ni%C3%B1os%2C%20discapacitados%20y%20embarazadas>
- TransMilenio. (2022, February 15). Historia de Transmilenio. TransMilenio. Retrieved September 17, 2022, from <https://www.transmilenio.gov.co/publicaciones/146028/historia-de-transmilenio/>
- TransMilenio. (2022, May 26). Servicios Troncales. TransMilenio. Retrieved September 18, 2022, from <https://www.transmilenio.gov.co/publicaciones/146189/servicios-troncales/>
- <https://datosabiertos-transmilenio.hub.arcgis.com/>
- https://console.cloud.google.com/storage/browser/validaciones_tmsa;tab=objects?prefix=&forceOnObjectsSortingFiltering=false
- https://storage.googleapis.com/validaciones_tmsa/index.html
- <https://datosabiertos.bogota.gov.co/organization/transmilenio>