

03/12/2018

---

---

---

# Dash – Encuesta de Movilidad de Bogotá 2015

*Patrones y comportamientos de la movilidad en Bogotá D.C.*

---

---

---

Ana Milena Rodríguez - Daniel Rojas -  
Luis Francisco Ortiz - Nicolás Hernández

# Dash – Movilidad en Bogotá

## 1. Introducción

Bogotá DC es una de las ciudades en Colombia que presenta un rápido crecimiento de su población, lo que se traduce en una alta densidad de habitantes por kilómetro cuadrado. Según la Secretaría Distrital de Planeación, Bogotá tiene una densidad promedio de 215,26 habitantes por hectárea.

La ciudad tiene veinte localidades, de las cuales nueve concentran tasas de densidad poblacional mayores al promedio de toda la ciudad. Es un desafío conocer cómo se moviliza la población a lo largo del día y en qué zonas suelen pasar la mayor parte de su tiempo, de esa forma diseñar y planificar acciones que vayan acorde a ese flujo ciudadano siendo adaptables al tipo de transporte y trayectos más comunes.

Para este proyecto propusimos desarrollar una **herramienta de visualización que permitiera entender cómo los ciudadanos se mueven de un punto origen a un punto destino discriminado por factores que influyen en la caracterización de los trayectos** tales como: el medio de transporte, el tiempo de espera, el estrato, entre otros. Esto con el fin de entender mejor el día a día de los bogotanos y brindar un insumo que permita evidenciar con datos cómo la Secretaría de Movilidad puede adaptar los flujos vehiculares acorde a la hora, día y zonas de la ciudad.

Para orientar el ejercicio se plantean unas preguntas como punto de referencia:

- ¿Cuáles son los patrones en las líneas de deseo de los bogotanos al moverse en la ciudad durante una semana promedio?
- ¿Cuáles son las Zonas de Análisis de Transporte más concentradas y demandadas por los bogotanos?
- ¿Qué patrones se pueden evidenciar en el flujo de personas en su transporte diario según el medio de transporte que utiliza?

El objetivo del proyecto es desarrollar una herramienta visual que permita dar respuesta a las preguntas anteriormente expuestas.

## 2. Fases de construcción

Para aterrizar el ejercicio propuesto se trabajó en tres fases, con el fin trazar objetivos más claros y tangibles en menor tiempo.

- Fase I

En la primera fase trabajamos en la **adquisición, limpieza y exploración de datos**. Cargamos las bases directamente de la página de datos abiertos utilizando las llaves (ID) para cada base de datos. Continuamos con la exploración de las bases para elegir aquellas que iban a hacer parte del análisis, elegimos **viajes, encuesta y personas**. En cada una de las bases elegidas exploramos variable por variable para encontrar **datos vacíos, datos nulos, datos atípicos, categorías por estandarizar** y adicionalmente hacer la limpieza necesaria para dejar los datos como los necesitábamos para el análisis. Finalmente complementamos las bases con información del **Laboratorio Urbano de Bogotá** para encontrar las coordenadas asociadas a las variables **ZATs (Zonas de análisis de transporte)**.

Librerías: *Pandas, pickle, os, numpy, matplotlib, folium y seaborn*

- Fase II

En la segunda fase **aplicamos teoría de grafos**, como recurso analítico, para determinar patrones importantes en los comportamientos de los desplazamientos (ver sección 3 – Teoría de grafos). Esto nos permitió identificar factores clave en la visualización del tablero.

Librerías: *networkx, collections, plotly, plotly.graph-objs, plotly.offline y spicy.stats*

- Fase III

La última fase fue la construcción del tablero. La elaboración se pensó de tal manera que permitiera al usuario manejar la herramienta de forma intuitiva y versátil, para que pudiera observar los comportamientos de la movilidad de acuerdo a sus inquietudes. Es por esto que decidimos incluir las siguientes variables en el tablero como características de filtro: **medio de transporte, motivo de viaje, estrato, hora valle, día hábil, edad y franja de inicio de viaje.**

Librerías: *dash, dash\_core\_components, dash\_html\_components, dash.dependencies y dash\_table\_experiments*

### 3. Descripción del tablero

El tablero está dividido en 4 módulos:

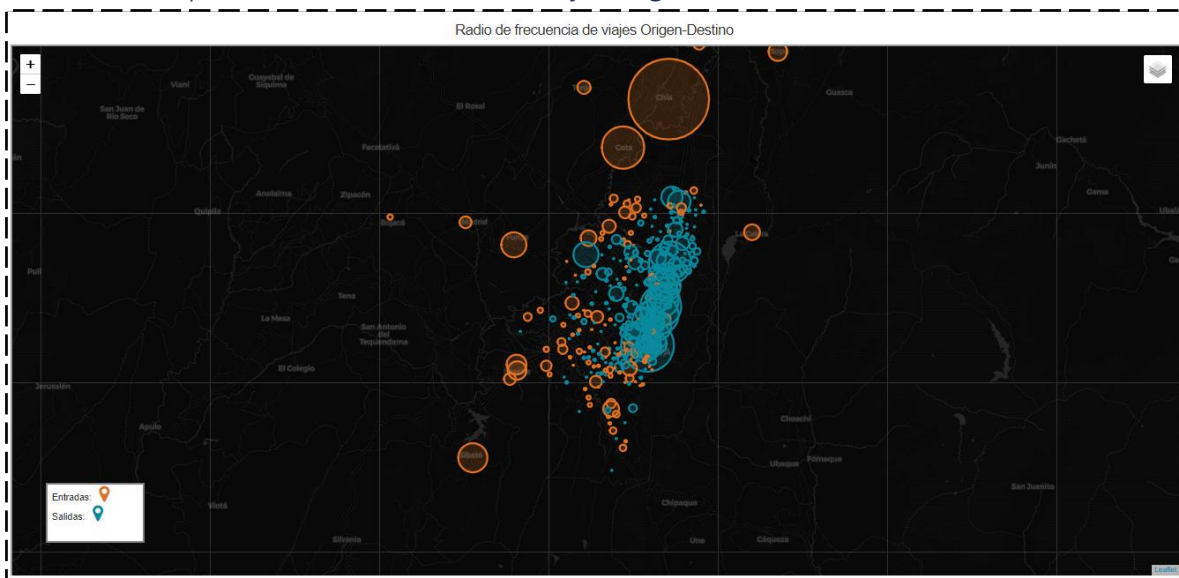
- Módulo de filtros

The screenshot displays the filter module of a dashboard. It includes several interactive components: a dropdown menu for 'Medio de transporte' with 'TPC-SITP' selected; a button-based selector for 'Motivo de viaje' with 'Estudiar' selected; a button-based selector for 'Estrato' with options 1 through 6; a dropdown menu for 'Hora Valle' with 'N' selected; a dropdown menu for 'Día hábil' with 'S' selected; a horizontal slider for 'Edad' ranging from 5 to 100 with a marker at 60; and a button-based selector for 'Franja de inicio de viaje' with 'Tarde' and 'Noche' selected.

El primer módulo permite al usuario escoger las variables de su interés para observar en el tablero. Son 7 variables en total con múltiples opciones y cuyos filtros se aplican en todos los módulos del tablero:

1. Medio de transporte: SITP / Transmilenio / Peatón / Especial / Auto / Moto / Taxi / Bicicleta / Ilegal / Alimentador / Intermunicipal / Otros
2. Motivo del viaje: Trámites / Volver a casa / Compras / Trabajar / Salud / Buscar o dejar algo / Buscar o dejar a alguien / Comer o tomar algo / Recreación / Buscar trabajo / Otros
3. Estrato: 1 - 6
4. Hora valle: Variable binaria (Si / No)
5. Día hábil: Variable binaria (Si/ No)
6. Edad: Quinquenios desde 5 años hasta 100 años
7. Franja inicio de viaje: Madrugada / Mañana / Tarde / Noche

- Mapa radio de frecuencias de viajes origen-destino

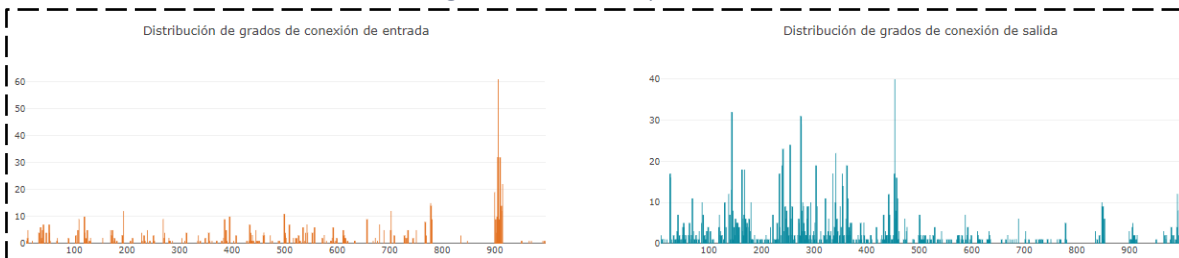


El segundo módulo del tablero muestra el mapa que identifica los nodos que mayoritariamente son de entrada y los que mayoritariamente son de salida de manera geolocalizada. Esto permite identificar la magnitud y el sentido en que se presentan los trayectos.

- Módulo de métricas

En el módulo de métricas se visualizan datos importantes para el entendimiento de las dinámicas en la movilización de Bogotá.

- Distribución de grados entrada y salida



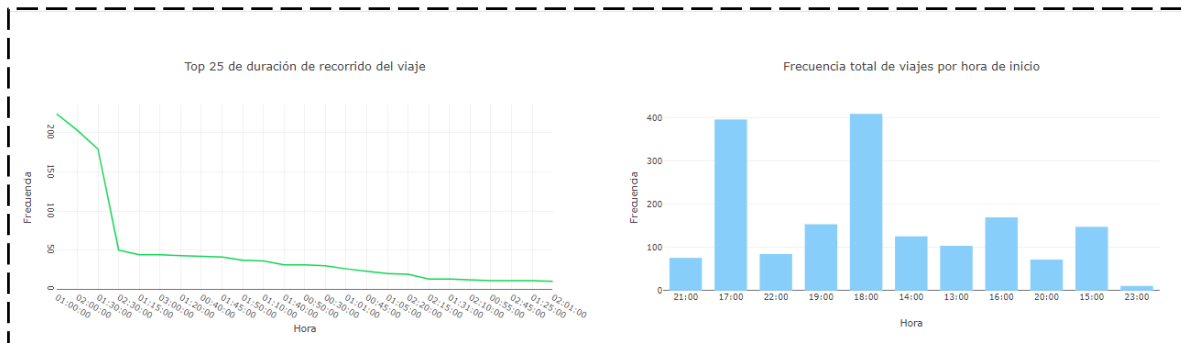
En estos histogramas se puede observar las **zonas de análisis de transporte** que tienen más entradas y más salidas. Estos histogramas reflejan el comportamiento del mapa de frecuencias.

- Analítica de grafos – tabla de métricas

ZAT	Barrio	Entradas	Salidas	Betweenness centrality	Closeness centrality	PageRank	Conexiones Totales
454	la catedral	0	40	0.015	0.582	0.004	40
144	chico	0	32	0.008	0.563	0.003	32
275	chapinero	0	31	0.013	0.576	0.004	31
276	la javeriana	0	29	0.017	0.591	0.005	29
255	chapinero	0	24	0.012	0.577	0.003	24
241	santanderito	0	23	0.003	0.528	0.002	23
342	sagrado corazon	0	22	0.003	0.535	0.002	22
363	santa isabel	0	19	0.01	0.566	0.003	19
304	betania	0	19	0.008	0.552	0.003	19
239	la esperanza	5	19	0.004	0.539	0.002	24
168	polo	0	18	0.003	0.538	0.002	18

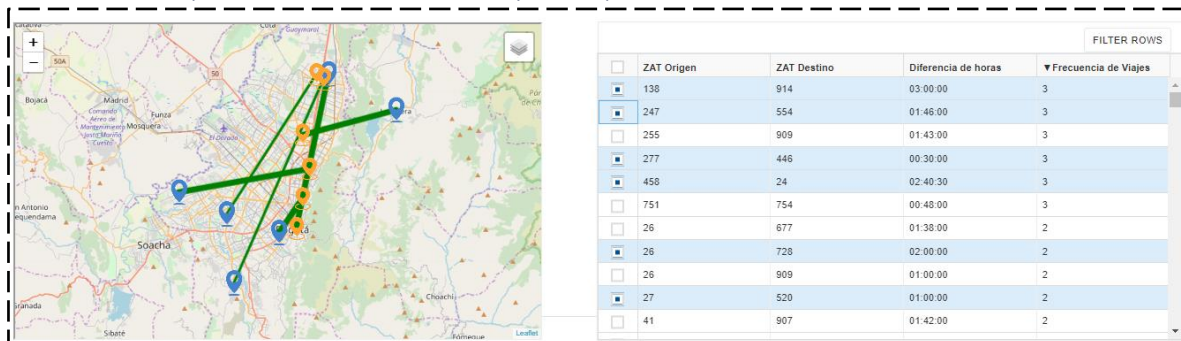
La tabla de métricas de grafos muestra los siguientes indicadores por cada ZAT: **Betweenness centrality**, **closeness centrality** y **Page Rank** (en la sección 4 se explica detalladamente qué significan estos indicadores y cómo se construyen). Adicionalmente se muestran **las entradas, las salidas y las conexiones totales de cada ZAT**.

- Duración y horarios de los viajes



También es posible entender la importancia de los desplazamientos por rango horario. La cantidad de trayectos varía según la hora, no hay el mismo flujo de transporte en horas valle que en horas pico.

- Mapa de visualización de trayectos por ruta seleccionada



El cuarto y último módulo del Dash permite al usuario filtrar por trayecto para observar trayectos específicos. Además, se puede seleccionar los tiempos de desplazamiento y las frecuencias de viajes en orden ascendente o descendente.

#### 4. Teoría de grafos

- Betweenness centrality

Es la medida que indica la frecuencia con que un nodo puede actuar como puente entre los diferentes componentes de la red.

$$C_{BET}(i) = \sum_{j,k} \frac{b_{jik}}{b_{jk}}$$

Donde:  $b_{jk}$  es el número de caminos más cortos desde nodo  $j$  hasta el nodo  $k$  y  $b_{jik}$  el número de caminos más cortos desde  $j$  hasta  $k$  que pasan por el nodo  $i$ .

- Closeness centrality

Closeness centrality es la medida de centralidad en una red. De esta manera, el nodo con la medida closeness centrality más alta es el nodo más cercano a todos los otros nodos que componen la red.

Esta medida tiene gran importancia en la red, dado que **permite diferenciar aquellos nodos que actúan como conectores dentro de la red urbana de la ciudad**. Es decir, los lugares que actúan no como destinos sino como zonas de tránsito y que reciben mayor concentración de trayectos.

El cálculo del closeness centrality se da de la siguiente manera:

$$C(x) = \frac{N}{\sum_y d(y; x)}$$

Donde:  $N$  es el número de nodos en el grafo, dividido por la sumatoria de la longitud de las rutas más cortas entre el nodo que se está analizando y todo el resto de los nodos del grafo.

- Page Rank

Es un algoritmo diseñado por Google para darle un ranking a las páginas web. El Page Rank se calcula así:

$$PR(A) = (1 - d) + d \sum_{i=1}^n \frac{PR(i)}{C(i)}$$

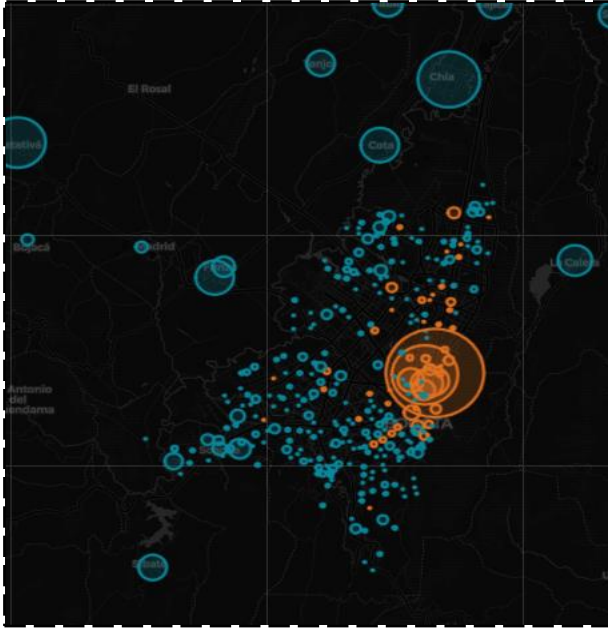
Donde:  $PR(A)$  es el Page Rank del nodo A,  $d$  es un factor de amortiguación que toma un valor entre 0 y 1,  $PR(i)$  son los valores del Page Rank que tiene cada uno de los nodos  $i$  que enlazan al nodo A y  $C(i)$  es el número total de conexiones salientes del nodo  $i$  sin importar hacia qué nodo salen.

El algoritmo busca priorizar aquellos nodos que permiten mayor número de conexiones de entrada, lo que ayuda mejorar la dinámica entre los nodos. Aplicado en la red de movilidad urbana de Bogotá, un nodo con un alto indicador de Page Rank permite mayores conexiones de entrada y por lo tanto se convierte en un nodo destino con alta concentración de llegadas.

## 5. Resultados

### Análisis de desplazamientos por motivo ESTUDIO

#### Franja: Mañana / madrugada – Medio: Transmilenio

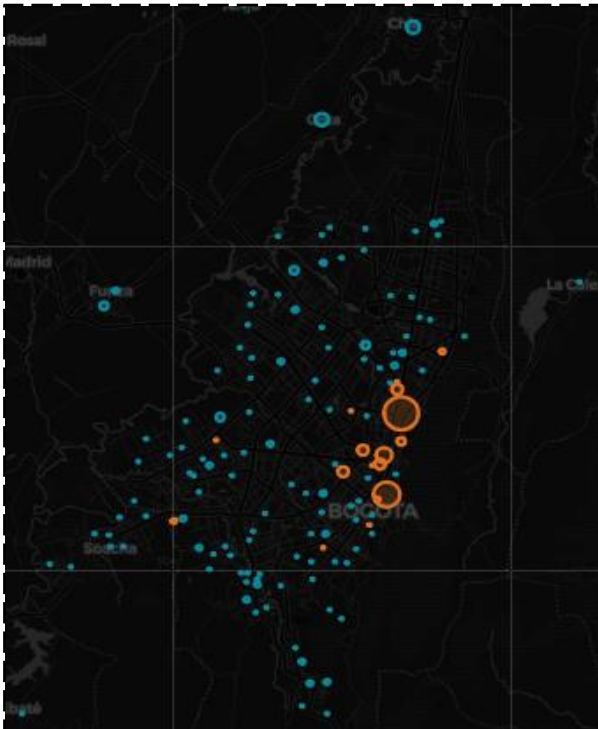


Los desplazamientos por motivo estudio en horas de la madrugada - mañana está marcados en su mayoría por trayectos con orígenes desde toda la ciudad, pero con un destino común que es el centro de la ciudad. En este sentido, estrategias de horarios diferenciados para centros educativos podría ser una opción para evitar concentraciones en puntos de la ciudad que actúen como nodos transitorios.

Al revisar las métricas de los grafos, **la javeriana es un nodo que toma mucha importancia dentro de la red**. Este nodo actúa como puente para conectar diferentes componentes, dado que tiene el índice **betweenness centrality más alto**.

Adicionalmente, es el nodo central principal, su **closeness centrality indica que la javeriana es el nodo más cercano al resto de nodos dentro de la red**.

#### Franja: Tarde / noche – Medio: Transmilenio



Los desplazamientos en Transmilenio durante la tarde y la noche son en su mayoría hacia el centro de la ciudad. **Chapinero y la candelaria se consolidan como un punto de gran actividad académica, como se pudo observar también en la franja de la mañana**.

Se puede inferir que en horas de la noche los desplazamientos por motivo de estudio son relacionados a estudios técnicos, tecnológicos, pregrado y posgrado.

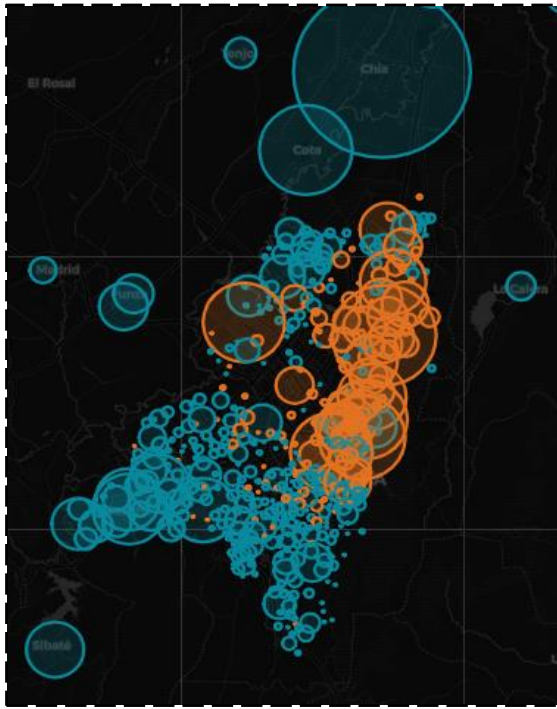
Los desplazamientos están dados principalmente en transporte público comparado con la cantidad que se da por otros medios de transporte con motivo estudio y en los horarios de la tarde – noche y para todos los estratos.



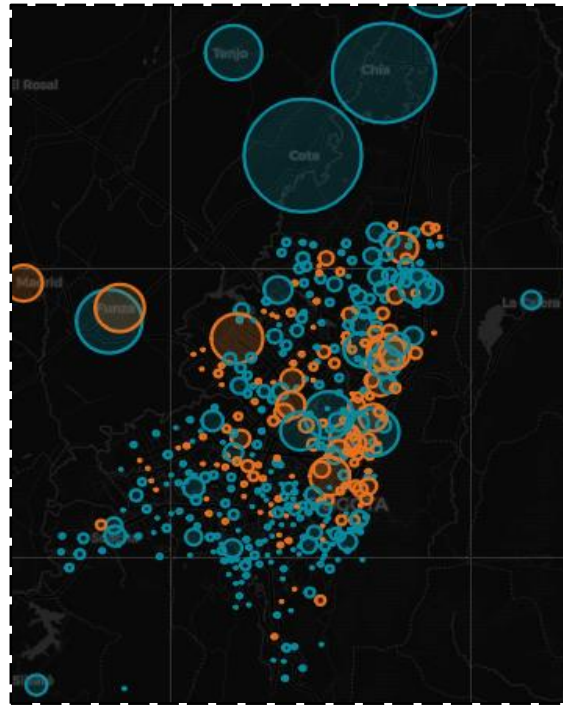
## Análisis de desplazamientos por motivo TRABAJO

Franja: Mañana / madrugada

**Medio de transporte Transmilenio**



**Medio de transporte auto**



Transmilenio

Los trayectos en **Transmilenio para ir a trabajar**, tienen una alta concentración de desplazamientos hacia el oriente de la ciudad. **Los nodos que actúan como puente dentro de la ciudad son: Javeriana, Chapinero, Santa Isabel, Candelaria y Restrepo (betweenness centrality)** y así mismo son los nodos de **mayor centralidad dentro de la red (closeness centrality)**. El mayor indicador de Page Rank dentro de la ciudad es para el centro de Bogotá y esto se puede evidenciar en el mapa, donde las entradas más significativas recaen sobre este nodo.

**Los horarios de inicio de los viajes en su mayoría empiezan desde las 5 hasta las 7 de la mañana. Los tiempos de desplazamiento con mayor frecuencia son: 1 hora, 2 horas, 1 hora y media y 45 minutos.**

### Auto

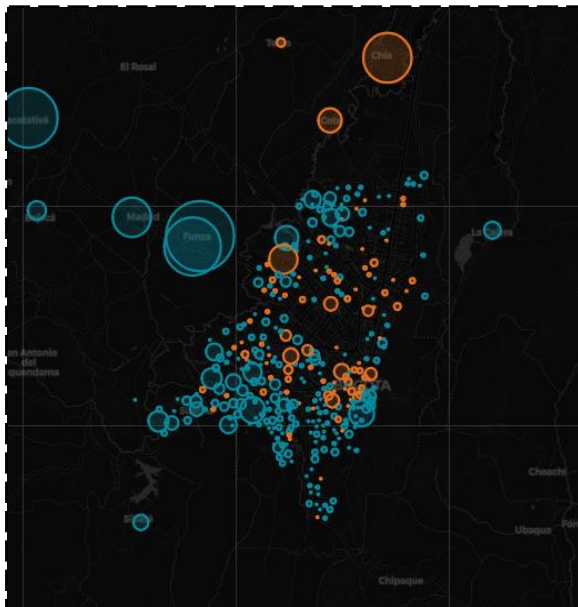
Los trayectos en auto para ir a trabajar ya no tienen una concentración tan marcada como lo tienen los trayectos en Transmilenio. **Los nodos de entrada y salida están repartidos por toda la ciudad.**

**Los horarios de inicio de los viajes en su mayoría están concentrados desde las 5 hasta las 9 de la mañana. Los tiempos de desplazamiento con mayor frecuencia son: 1 hora, media hora, 15 minutos y 10 minutos.**

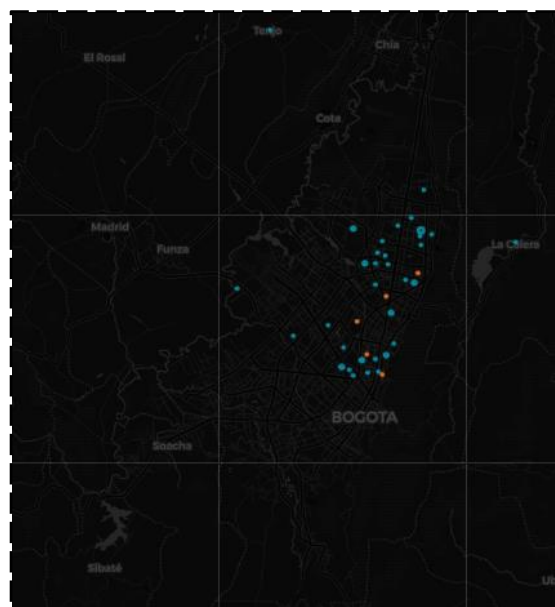
Franja: Mañana / madrugada – Medio: Moto



**Estratos 1, 2 y 3**



**Estratos 4, 5 y 6**



#### Estratos 1, 2 y 3

Desplazamientos periféricos, desde el exterior hacia el interior de la ciudad. Tienen un alto volumen en el flujo de desplazamientos. Los nodos centrales son: Las Nieves, Candelaria, Sierra Morena y Restrepo.

#### Estratos 4, 5 y 6

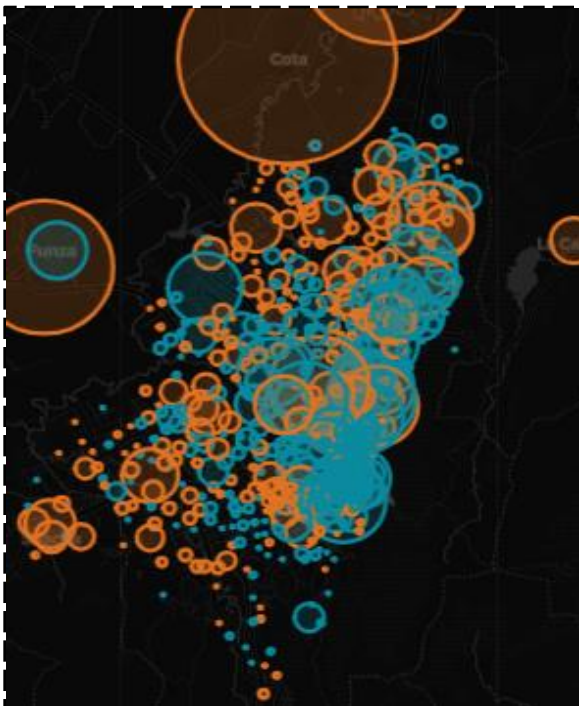
Desplazamientos concentrados en la zona nororiental de la ciudad. En cuanto a **volumen de desplazamiento**, las motos, tienen un **bajo nivel en los estratos 4, 5 y 6**. Los nodos que sirven como puentes de conexión son Galerías, Centro, Country, Sauzalito y Chapinero (Betweenness centrality), que a su vez son los nodos centrales de la red.



Franja: Tarde / Noche – Medio: Transmilenio

Los desplazamientos de regreso a casa muestran la verdadera dimensión del volumen de trayectos que realiza Transmilenio. Hay un patrón en dichos desplazamientos desde el oriente hacia el resto de la ciudad, mayoritariamente.

Los tiempos de desplazamiento son mayores, las duraciones que más frecuencia tienen los trayectos son 1 hora, 2 horas, 1 hora y media y 2 horas y media. La Mayoría de los viajes que se realizan en la tarde – noche son en la hora pico: 4, 5 y 6 pm.



Franja. Tarde / Noche – Medio: Auto

Los desplazamientos de regreso a casa no tienen un direccionamiento tan claro, como se evidencia en Transmilenio. Sin embargo los nodos del oriente de la ciudad siguen siendo los nodos de mayor conexiones de salida.

Los tiempos, según los datos de la encuesta de movilidad, son menores en los desplazamientos en auto. Las duraciones de mayor frecuencia son 1 hora, media hora, 1 hora y media y 10 minutos. Los horarios de salida de los viajes con más frecuencia siguen siendo los horarios de Transmilenio, sin embargo, los horarios anteriores, desde la 1 pm hasta las 3 pm, aumentan en proporción de viajes.

## 6. Conclusiones

La herramienta permite al usuario tener una clara percepción de las dinámicas de movilización en Bogotá. Tan solo encontrar el sentido en el que se presentan la mayoría de los trayectos, es el punto de partida para generar estrategias y toma de decisiones enfocadas en pro de la movilidad.

Finalmente, el objetivo principal, de la construcción del dash, se cumple en la medida que a través de esta herramienta se puede dar respuesta a las 3 preguntas base que se habían planteado.

- Se puede encontrar los patrones de las líneas de deseo en la movilización de los bogotanos, desagregando por la herramienta.
- Las métricas establecidas por teoría de grafos permiten identificar aquellas zonas de análisis de transporte más concentradas y demandadas por los bogotanos, además aquellos que actúan como nodos conectores y centrales
- Los patrones, según el medio de transporte, se pueden evidenciar fácilmente gracias a la forma en que fue construido el dash, de tal manera que se puede filtrar por diferentes características, una de esas es el medio de transporte.

Si se aplicara el uso de esta herramienta en el sector público para la construcción de políticas públicas existe un alto potencial para generar insumos que deriven en estudios más detallados sobre la aplicación y viabilidad de iniciativas como la instalación de parqueaderos para bicicletas (utilizando las métricas calculadas se puede conocer cuales puntos son los más estratégicos para construirlos) o también la posibilidad de dinámicamente identificar un grupo de estudio con características sociales como su estrato o su medio de transporte e ir validando su comportamiento a lo largo de diferentes franjas horarias.