

Transitabilidad CABA

Elaboración de un índice de transitabilidad que represente el nivel de tránsito peatonal en la ciudad.

# Objetivo del proyecto

Favorecer el entendimiento de la circulación peatonal en la ciudad y facilitar la toma de decisiones de políticas públicas, tanto a nivel estratégico como operativo.

# Resultado esperado

Asignación de un valor al tránsito peatonal en cada una de las calles de la Ciudad Autónoma de Buenos Aires, desagregado por año, cuatrimestre, mes, día de semana y hora. La escala es entre 0 y 1, con el 0 representando un nivel muy bajo de tránsito peatonal y el 1 un nivel muy elevado.

# Resumen del procedimiento

El procedimiento realizado obtiene un índice entre 0 y 1 calculado a partir de los índices SUBE (80%) y Relevamiento de Usos del Suelo (20%). Por un lado, a partir de la ubicación de las transacciones SUBE, se tomaron supuestos según el modo de transporte y se estimó el tránsito peatonal de los diferentes tramos de calle. Por otra parte, el relevamiento de suelos permitió identificar atractores de tránsito peatonal y complementar el índice SUBE.

# Fuentes de datos

Las fuentes de datos utilizadas fueron:

* Callejero, USIG y BA Data.
* Parcelas, BA Data.
* Base de datos SUBE, provista por la empresa Nación Servicios S. A.
* Relevamiento de Usos del Suelo (RUS) 2017, BA Data.
* Calendario de fechas normalizado, con información adicional de día de semana y si se trata de un feriado o no.

# Etapas

## 1. Selección de los datos

Este Índice se elaboró integrando los datos del sistema SUBE que registra el uso del transporte público por parte de los individuos y del Relevamiento de Usos del Suelo (RUS) realizado en 2017 por el entonces Ministerio de Desarrollo Urbano y Transporte (MDUyT) para detectar la distribución espacial de comercios, oficinas, viviendas y otros usos atractores de personas.

## 2. Preparación de los datos

Para llevar adelante el análisis fue necesario preparar los datos disponibles.

Callejero

Para la elaboración del índice, se decidió utilizar como unidad de análisis territorial la traza del callejero de la Ciudad de Buenos Aires. Para ello, fueron filtradas del proceso aquellos segmentos de calle que no poseen tránsito peatonal, como autopistas, pasos bajo nivel y puentes vehiculares.

Relevamiento de Usos de Suelo (RUS)

Por otra parte, se agruparon las 84 ramas y 174 subramas que presenta el RUS en las siguientes 9 categorías:

1. Residencial
2. Comercial
3. Gastronomía y hotelería
4. Productivo (fábricas y establecimientos con fines productivos)
5. Oficinas
6. Sitios de interés y entretenimiento (museos, bibliotecas, cines, teatros, etc.)
7. Servicios (bancos, administración pública, etc.)
8. Educación
9. Salud

Datos SUBE

La base de datos es provista mensualmente por Nación Servicios, y está compuesta por varias tablas. Las dos más importantes son la de transacciones y la de marcaciones GPS. Al recibir nuevos datos, se ejecuta un proceso de ETL[[1]](#footnote-0) armado en la DGCINFO, y se incrementa la base de datos histórica. Como insumo para el índice de transitabilidad se utiliza el resultado de este ETL, donde cada transacción posee su ubicación espacial.

### Tabla auxiliar de distancias

El armado de esta tabla auxiliar permite vincular las transacciones con los diferentes segmentos de calles. Las transacciones SUBE son realizadas en determinada latitud-longitud, y esta tabla permite asociar espacialmente cada transacción con los tramos de calle de su entorno.

Se tomaron hipótesis sobre las distancias de tránsito peatonal según el modo utilizado. Las distancias escogidas por modo son:

* Tren: 600 metros.
* Subte: 500 metros
* Colectivo: 300 metros.

Para obtener el mejor resultado posible de estas hipótesis, se calcularon las distancias dentro de la red peatonal de calles, y no simplemente la distancia euclídea o “distancia a vuelo de pájaro”. Esto permite mejorar la representación del comportamiento, pues utiliza la topología de las calles y captura la existencia de barreras geográficas.

Como orígenes o fuentes de transacciones se tomó una grilla de puntos de latitud-longitud de 3 decimales. Esto resultó en una matriz de unos 34.000 puntos que luego fue recortada con la forma de la ciudad. El resultado fue un total de unos 20.000 puntos con una separación de 110 m entre puntos contiguos.

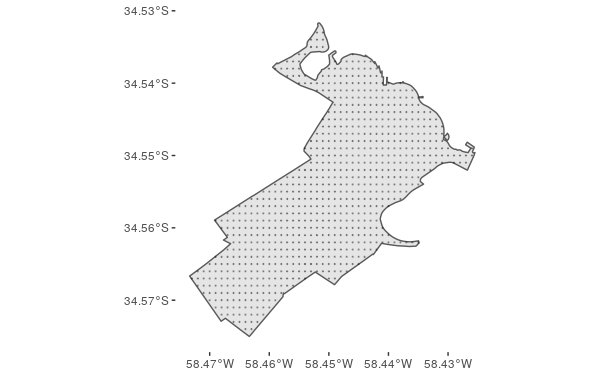


Figura N° 1. Grilla del barrio de Belgrano donde pueden verse los puntos cada 110 m (0.001 en grados), utilizados como fuentes de transacciones.

Para la obtención de la tabla, se crearon buffers[[2]](#footnote-1) circulares de 600 m para cada punto de la grilla (~20.000 puntos, cada uno con un *id\_grilla*) y se realizó una selección espacial de centroides de tramos del callejero (o *id\_calle*). La cantidad de tramos del callejero es del orden de 30.000, y en promedio cada *id\_grilla* seleccionó 163 *id\_calle*. Con esto se conformó el universo de los pares de puntos para rutear, es decir para obtener la distancia entre ellos, pero sujeta a recorridos válidos dentro de la red de calles.

El ruteo se realizó con un servidor local de OSRM, que luego de optimizar el código, realizó más de 3 millones de ruteos en unas 4 horas totales.

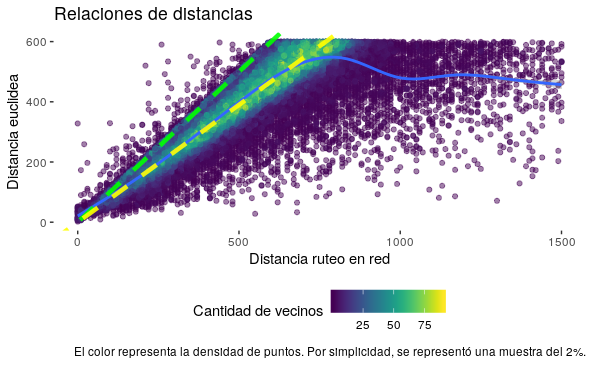


Figura N° 2. Relación de las distancias obtenidas. La recta verde posee pendiente 1, mientras que la recta amarilla pendiente 0.8. Esto muestra que para muchos puntos la distancia en la red de calles es un 25% mayor. También abundan casos donde la distancia en red es mucho más grande que la euclídea.

Como puede verse en esta imagen, en términos generales la distancia en la red de calles es mayor a la distancia euclídea. Una vez realizado el ruteo, se aplica el criterio de distancia escogido para los diferentes modos.

La tabla resultante contiene contiene el id\_calle, id\_grilla, el modo, y la distancia sobre la red entre el punto de la grilla y el centroide del tramo de calle.

A modo de ejemplo y evaluación se analizó el entorno del Palacio Lezama.



Figura N° 3. Entorno del Palacio Lezama. Si bien el buffer de distancia euclídea selecciona calles en un radio de 600 m, aquellas calles resaltadas en rojo se encuentran a más de 600 m de caminata, por lo que son filtradas para los análisis posteriores.

A su vez, a partir de la distancia y el modo se calcularon 2 ponderadores. Los ponderadores se utilizan para modelar la relación entre el punto de la grilla y el tramo de calle. Se entiende que a mayor distancia es más improbable que quien haya realizado la transacción hubiese utilizado ese tramo de calle. De este modo, los ponderadores disminuyen a medida que la distancia crece.

A la hora de analizar las transacciones, se observó que un único ponderador era insuficiente, por lo que fue desdoblado en dos para diferentes tipos de transacciones: inicios de viaje y transbordos. El ponderador para transbordo decrece aún más con la distancia, dado que se supone que quien realiza combinaciones procura una distancia de caminata reducida entre tramo y tramo. Por su parte, la red de transporte está diseñada para propiciar centros de transbordo donde esa distancia es minimizada.

| Modo | Distancia | Ponderador para inicios de viaje | Ponderador para transbordo |
| --- | --- | --- | --- |
| Colectivo | Menor a 100 m  Entre 100 m y 200 m  Entre 200 m y 300 m | 1  0.7  0.4 | 1  0.5  0.2 |
| Subte | Menor a 100 m  Entre 100 m y 200 m  Entre 200 m y 400 m  Entre 400 m y 500 m | 1  0.7  0.4  0.2 | 1  0.4  0.2  0.05 |
| Tren | Menor a 100 m  Entre 100 m y 200 m  Entre 200 m y 400 m  Entre 400 m y 600 m | 1  0.7  0.4  0.2 | 1  0.3  0.2  0.05 |

Tabla N° 1. Ponderadores para transacciones de inicio de viaje y transbordo.

## 3. Desarrollo

### Transformación de los datos SUBE

Debido a que la identificación del tramo de calle de cada transacción de SUBE no es factible de hacerse directamente, se redujo la precisión de los valores de latitud y longitud a 3 decimales.

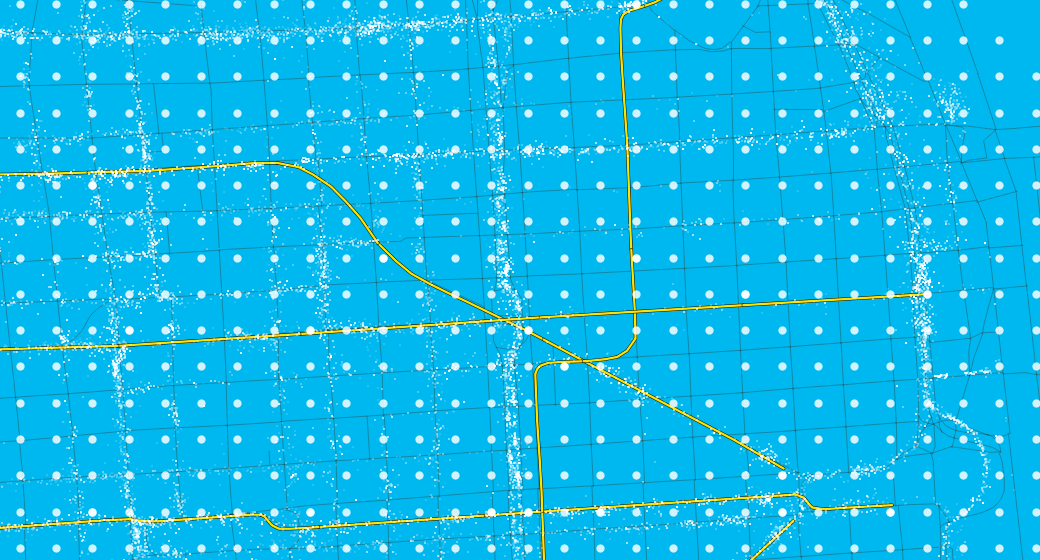
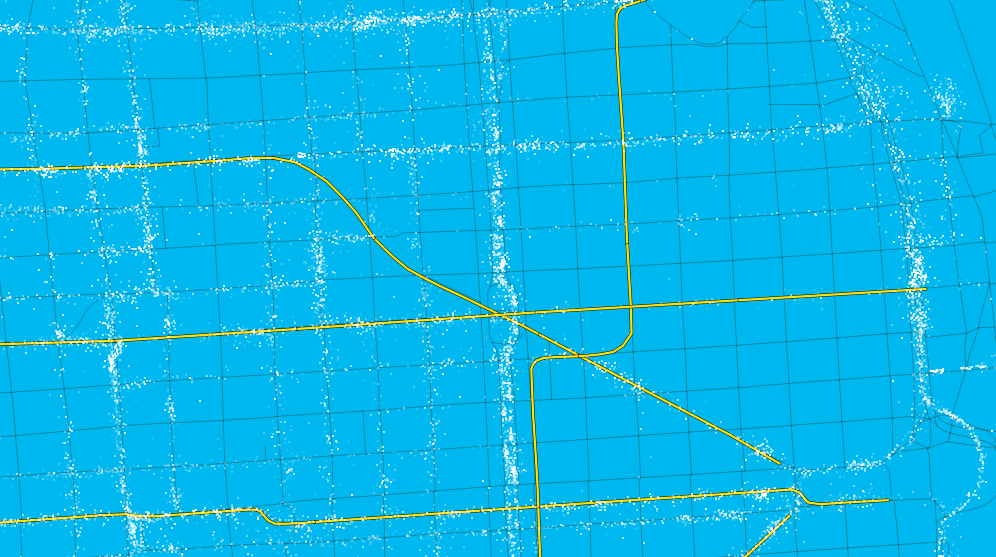


Figura N° 4. Simplificación de la localización de transacciones SUBE.

A su vez, para cada transacción se calcularon las siguientes variables:

1. Día de la semana: Fue obtenido a partir del calendario armado. Se trata de 8 categorías, una correspondiente a cada día de la semana y la octava correspondiente a los días feriados.
2. Hora: Se unificó el campo de hora y minuto a horas enteras.
3. Año
4. Cuatrimestre: Fue obtenido a partir del calendario armado.

El siguiente paso fue contar la cantidad de transacciones para los niveles de agrupación definidos que son:

1. Tiempo: año, cuatrimestre, mes, día de semana, hora
2. Espacio: latitud, longitud
3. Modo: colectivo, subte o tren

A continuación, se transfirió este conteo de transacciones a los diferentes tramos de callejero dentro de la zona de influencia, teniendo en consideración el tipo de transacción y la distancia entre la transacción y el tramo de calle.

Posteriormente, para cada tramo de calle se sumaron los aportes de cada punto de la grilla de transacciones, y se calculó el promedio para la agrupación deseada.

Finalmente, se escaló el resultado entre 0 y 1 con el método de máximos y mínimos, para todos valores de la tabla. Realizarlo de esta manera habilita la comparación del índice entre todos los días, horarios y períodos en cada una de las calles de la Ciudad.

A modo de ejemplo, se calculó el índice categorizado en 7 niveles, utilizando el método de rupturas naturales de Jenks[[3]](#footnote-2).

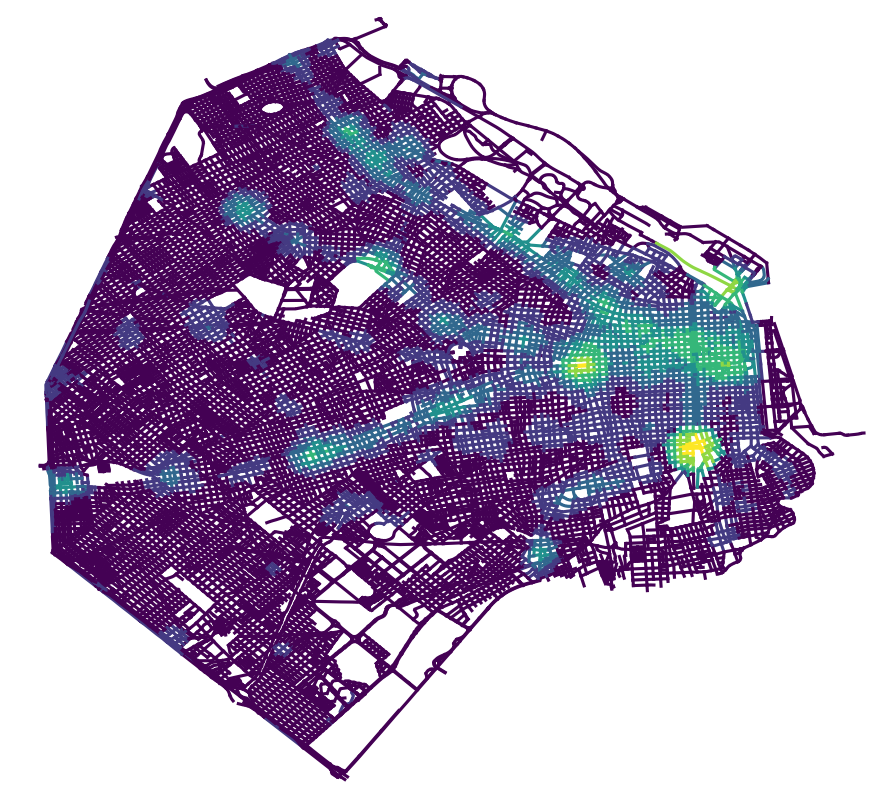


Figura N° 6. Representación del índice SUBE para un día miércoles de 12 a 16 horas.

### Transformación de los datos RUS

La unidad de análisis del Relevamiento de Usos de Suelo es cada parcela, que está georreferenciada en forma de punto por las coordenadas del centroide de la misma. En los casos donde existe más de un uso en una misma parcela, se cuentan con dos o más puntos con las mismas coordenadas. Por esto, luego de un tratamiento preliminar de datos se generó una nueva variable que cuenta la cantidad de usos de suelo para cada dupla (latitud, longitud), lo que equivale a decir para cada parcela.

A continuación, se realizó el cruce espacial de las parcelas y el callejero. Se identificaron para cada tramo de calle de la ciudad las parcelas correspondientes, que a su vez ya estaban vinculadas a sus usos de suelo.

Se construyeron también cinco nuevas variables:

1. Día de la semana
2. Hora: Se unificó el campo de hora y minuto a horas enteras.
3. Año
4. Cuatrimestre: Fue obtenido a partir del calendario armado.
5. Periodo: indica a qué cuatrimestre corresponde, 1 es 3C-2017 es 1, 1C-2018 es 2, etc.

El siguiente paso fue utilizar los siguientes niveles de agrupación definidos:

1. Modo
2. Periodo
3. Año
4. Cuatrimestre
5. Día semana
6. Latitud
7. Longitud

Para que los valores del ponderador fueran comparables con el índice SUBE, se construyó una ponderación de los valores según las variables descritas. Para esto, se tuvieron en cuenta cuáles son los horarios y días en los que cada uso presenta mayor actividad. A modo de ejemplo, para las oficinas se estimó mayor tránsito peatonal de lunes a viernes en los horarios de 8 h o 12 h y a las 18 h, mientras que en los locales gastronómicos el mayor tránsito peatonal fue los viernes y sábados en horario nocturno (20 h a 00 h).

El valor de transitabilidad RUS se calculó para cada tramo de calle, día y horario de la siguiente forma:

Finalmente, de igual forma que en el caso se SUBE, se realizó la normalización por el método de mínimos y máximos, de forma de obtener un índice con valores entre 0 y 1.

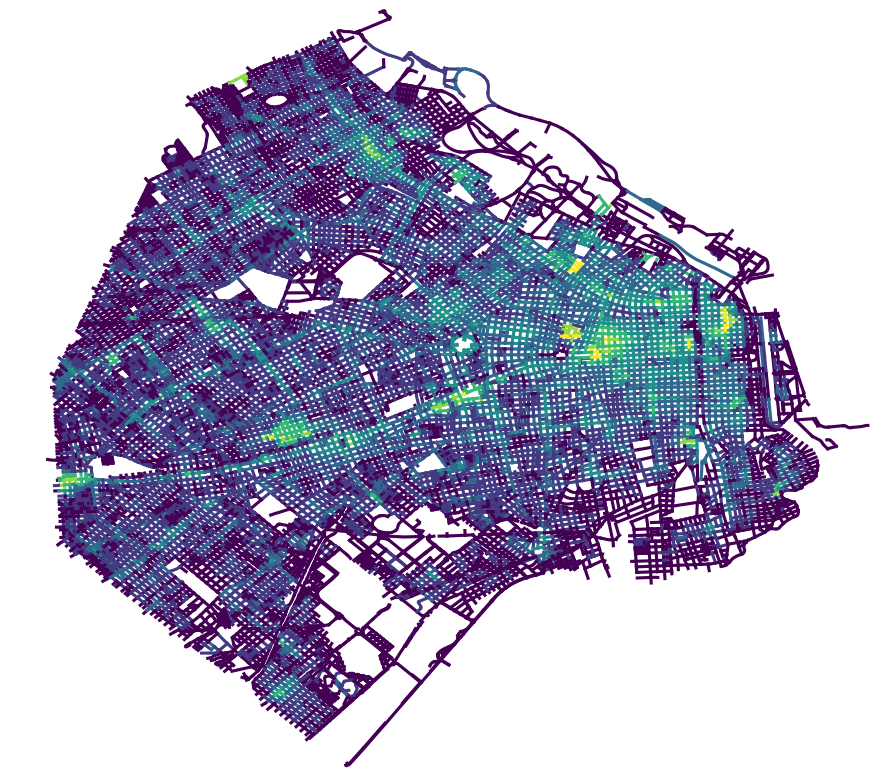


Figura N° 7. Representación del índice RUS para un día miércoles de 12 a 16 horas.

### Otros análisis realizados

En el proceso de elaboración del índice se exploraron otras alternativas o índices posibles. Entre ellos:

* Un índice donde la distribución de transacciones se dividiera a todas las calles de su entorno por igual. Se descartó pues una misma persona puede caminar varios tramos de calle, por lo que no sería incorrecto asignar una misma transacción a varios tramos.
* Un índice donde la distribución de los puntos a los tramos de calle fuera la misma. Fue reemplazada por la versión actual y la utilización del ponderador.

## 4. Elaboración del índice compuesto

Una vez calculados ambos indicadores entre 0 y 1, se realizó la unión de ambos por la clave de agrupación (id\_calle, año, cuatrimestre, día de la seman, hora). Luego se calculó el índice compuesto ponderado donde se le dio mayor importancia al índice SUBE. Esta definición se realizó en base a considerar a esta fuente por ser más representativa de la movilidad de las personas. De esta forma:

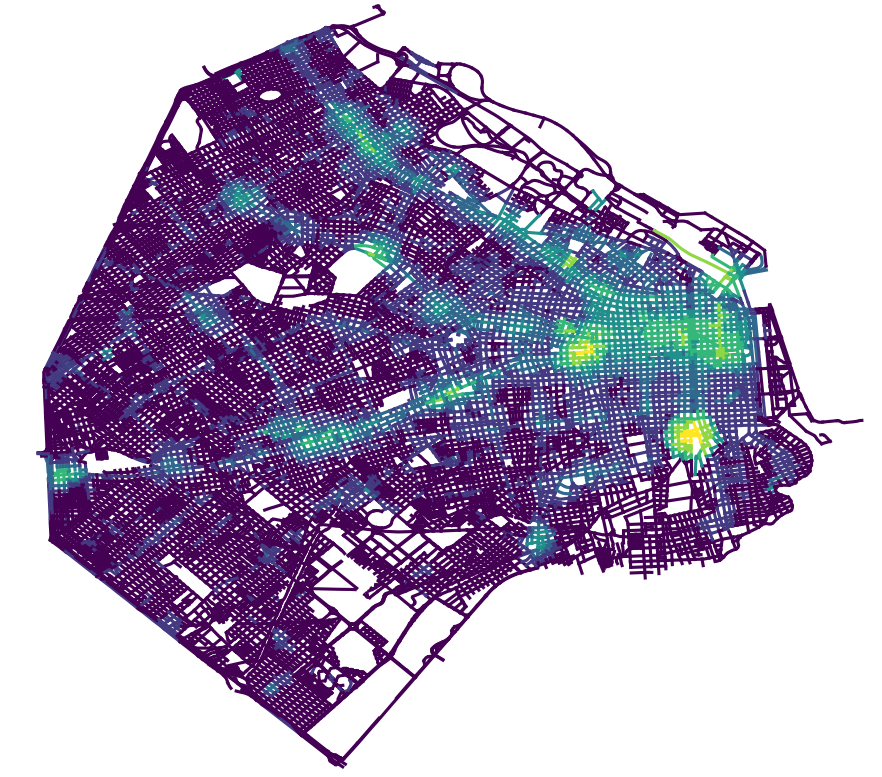


Figura N° 8. Representación del índice de transitabilidad para un día miércoles de 12 a 16 horas.

## 5. Visualización y casos de uso

### Visualización de resultados

A la hora de mostrar los resultados obtenidos, se analizaron distintas alternativas. Por una parte, se observó la distribución de valores del índice y se encontró que se trataba de una distribución fuertemente asimétrica, con muchos valores bajos en la izquierda y una larga cola hacia la derecha.

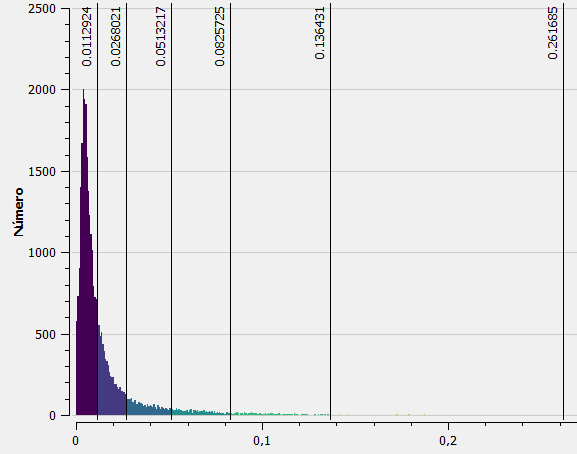


Figura N° 8. Histograma de valores del índice para un cuatrimestre, día y hora seleccionados.

Realizar la clasificación con *natural breaks* resultaba en matices o granularidad en los valores altos de la distribución, pero también en la agrupación de valores bajos en una misma categoría. De este modo, podía distinguirse diferencias de tránsito en las comunas más transitadas como las número 1, 2, 3, 13 y 14, y en los grandes centros de transporte, pero la mayor parte de la ciudad se representaba como un único valor bajo.

Esta visualización también implicaba que en el horario menos transitado, de medianoche a las 4 de la madrugada, toda la ciudad se mostrara de un único color.

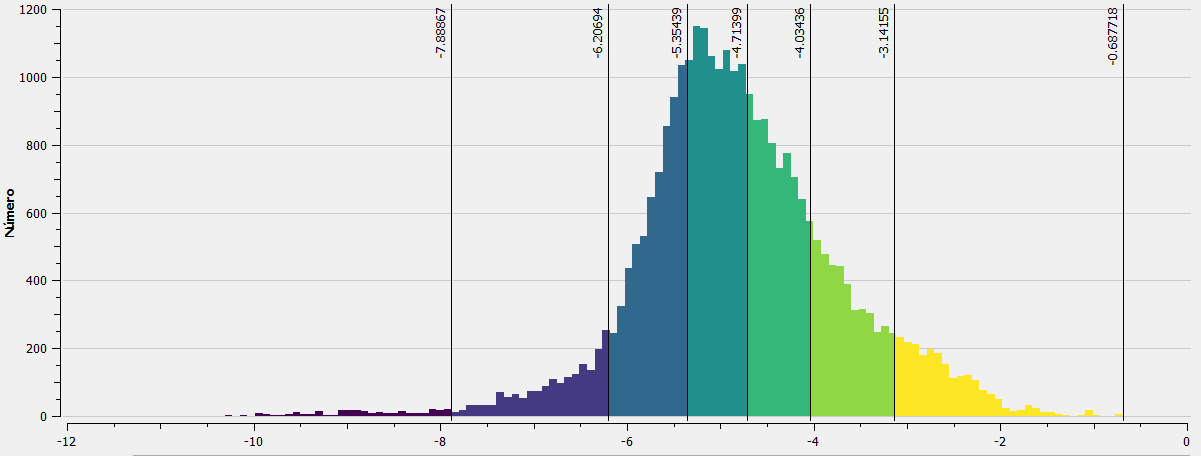


Figura N° 9. Histograma del logaritmo del índice para un cuatrimestre, día y hora seleccionado.

Por esta razón, se procedió a representar el índice aplicando una transformación logarítmica y posteriormente clasificarlo con el algoritmo de *Jenks natural breaks*. Esto permite comparar valores en áreas donde la circulación es menor e identificar patrones en áreas más amplias de la ciudad.

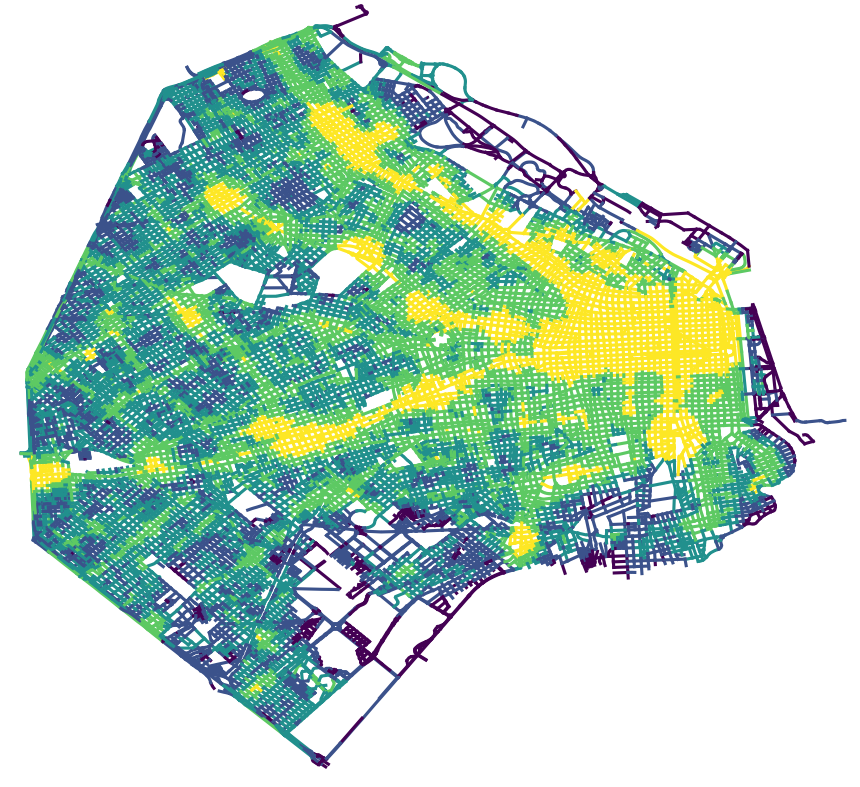
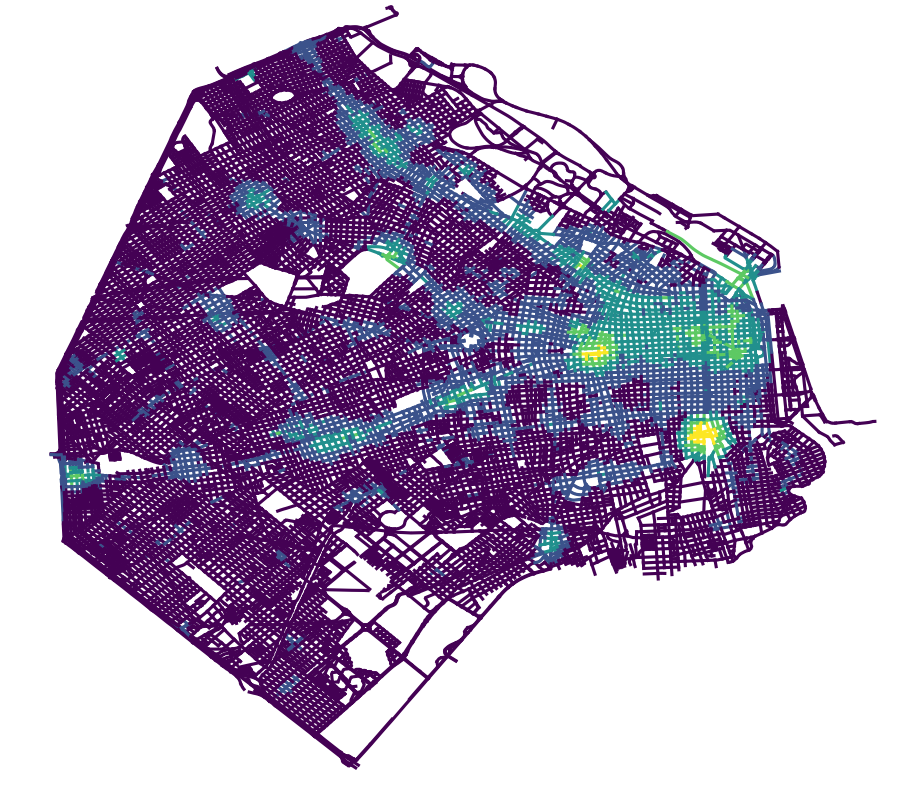


Figura N° 10. Representación del índice de transitabilidad para el día miércoles de 8 a 12 horas, sin realizar ninguna transformación (izquierda) y realizando transformación logarítmica (derecha).

1. Siglas de extracción, transformación y carga de datos. [↑](#footnote-ref-0)
2. Un buffer representa un área de influencia espacial, en este caso generado a partir de los puntos de grilla donde se iniciaron transacciones SUBE. [↑](#footnote-ref-1)
3. Conocido como *natural breaks*, del inglés, es una clasificación popular si se busca identificar y agrupar elementos similares. [↑](#footnote-ref-2)