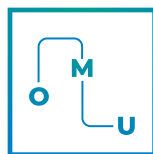


ACCESO A OPORTUNIDADES EN AMÉRICA LATINA

METODOLOGÍA PARA
LA CONSTRUCCIÓN
DE INDICADORES

AUTORES

VANOLI, CATALINA
ANAPOLSKY, SEBASTIÁN



Una iniciativa de

CAF BANCO DE DESARROLLO
DE AMÉRICA LATINA

BID
Banco Interamericano
de Desarrollo

Título: Acceso a Oportunidades en ciudades de América Latina:
Metodología para la construcción de indicadores

Depósito Legal: DC2023000021 **ISBN:** 978-980-422-292-4
Gerencia de Conocimiento | CAF -
Banco de Desarrollo de América Latina

Autoría de contenidos:
Catalina Vanoli (CAF) y Sebastián Anapolsky

Edición gráfica y corrección tipográfica
El Gato y La Caja

Créditos

El documento fue escrito por Catalina Vanoli, especialista en transporte y científica de datos trabajando en CAF, y Sebastián Anapolsky, consultor especialista en transporte y ciencia de datos. La librería de código abierto en python (pyomu) fue desarrollada por Sebastián Anapolsky.

Esta publicación forma parte de las actividades definidas dentro de dos iniciativas impulsadas por CAF: Observatorio de Movilidad Urbana y Ciudades con Futuro. El Observatorio de Movilidad Urbana —OMU— surgió en 2007 para cerrar la brecha de información sólida, confiable y actualizada de movilidad urbana en la región aportando evidencia robusta para la toma de decisiones. El programa Ciudades con Futuro o CcF se lanzó en 2014 con el objetivo de fomentar el desarrollo de líneas de trabajo estratégicas en la región.

Agradecimientos

Los autores agradecen a Nicolás Estupiñán, director de Análisis Sectorial de CAF, por apoyar este proyecto y por sus aportes sobre las implicaciones de política pública; Mariana Poskus y Pablo López, especialistas en desarrollo urbano de CAF, que impulsaron este proyecto desde el comienzo y contribuyeron con la metodología y análisis de casos; Nick Bristow y Diego Castelari, que colaboraron en el desarrollo metodológico y en la elaboración de indicadores; Guillermo Alves, economista en CAF por su revisión del documento y sus comentarios sobre la metodología y estructura del documento; y Lucila Berniell, Federico Juncosa y Walter Cont, economistas y especialistas de CAF por las discusiones técnicas sobre indicadores de acceso a empleo y a espacios verdes. Agradecimientos extensivos a las entidades públicas que dieron acceso a datos de ubicación de establecimientos y a las entidades que gratuitamente tienen dichos datos abiertos al público.

Este y otros libros sobre Movilidad Urbana y datos se encuentran en: scioteca.caf.com



Introducción al OMU

El Observatorio de Movilidad Urbana (OMU) para América Latina fue impulsado por el Banco de Desarrollo de América Latina (CAF) con la finalidad de fomentar prácticas que guíen decisiones políticas hacia sistemas urbanos de transporte que cumplan su rol económico, social y universal de manera eficiente, segura e inclusiva. Gracias a esfuerzos y cooperaciones locales e internacionales, el observatorio se ha convertido en un bien público regional y en una herramienta valiosa para el análisis de los sistemas urbanos de transporte. Esto ha permitido el desarrollo de diálogos en políticas sectoriales con gobiernos y ha enriquecido el proceso de toma de decisiones en el área de servicios de infraestructura de la región.

Los resultados de este trabajo forman parte de la tercera edición del OMU, desarrollado por CAF en conjunto con el Banco Interamericano de Desarrollo (BID).^[01] Los indicadores elaborados en este documento se enmarcan dentro de una matriz de 22 indicadores, bajo 4 pilares fundamentales: acceso universal, eficiencia y calidad, seguridad, y movilidad verde, y su objetivo es la evaluación y monitoreo de métricas relevantes para la movilidad urbana de las ciudades de la región.^[02]

[01]

El OMU fue creado por CAF en 2007. En 2010 lanzó su primer informe con información sobre movilidad y transporte de 14 ciudades de Latinoamérica. En 2015, la segunda edición, contó con datos relevantes para el sector de 29 ciudades y áreas metropolitanas de la región. La tercera edición del OMU, 2020-2022, se está desarrollando en conjunto con el BID, en donde se reformulan los objetivos para alinearlos con las agendas globales y se incorporan las tecnologías de información como herramientas de recolección de datos.

[02]

Ver página web del OMU:
www.omu-latam.org

Resumen

Las ciudades ofrecen el acceso a oportunidades de diferente tipo a su ciudadanía, como oportunidades de trabajo, de educación y salud, culturales y de recreación. El acceso equitativo y eficiente a estas oportunidades determina el nivel de productividad y calidad de vida en nuestras ciudades. El desarrollo de indicadores de accesibilidad permite mejorar el diseño de las políticas públicas, identificando inversiones e intervenciones urbanas con decisiones basadas en datos.

En un contexto de grandes avances tecnológicos, con acceso a nuevas fuentes de datos y herramientas para su análisis, la tercera edición del OMU presenta el desarrollo de una metodología que permite calcular una serie de métricas para monitorear la accesibilidad urbana en las ciudades utilizando datos públicos. A su vez, el OMU desarrolló una herramienta de código abierto que permite replicar este análisis en el futuro o para nuevas ciudades.

La metodología incluye el cálculo de indicadores de congestión, identificación de áreas de alta densidad de actividad, cálculo de los tiempos de viajes a distintas oportunidades urbanas (centralidades

urbanas, establecimientos educativos y de salud, etc), accesibilidad a espacios verdes y públicos y cobertura del transporte público. Estas métricas se calculan utilizando información censal y cartográfica de las ciudades y utilizan las aplicaciones de ruteo de viajes (Open Street Maps y Google Maps) para obtener información sobre los viajes. Esta herramienta permite también la construcción de un indicador de nivel socioeconómico que permite analizar a la población en forma más desagregada.

En la última parte del trabajo, se presenta el caso de estudio de la Región Metropolitana de Buenos Aires (RMBA) donde se muestran los indicadores que pueden ser elaborados con esta herramienta de código abierto.

La metodología es aplicable a diferentes contextos y puede ser replicable por especialistas técnicos de los gobiernos, la academia o analistas de ciudades. La librería `pyomu` es de código abierto y se puede acceder en el repositorio de GitHub del OMU. Se buscó que los resultados sean de fácil interpretación para poder ser presentados a autoridades y tomadores de decisión, acentuando la importancia de identificar brechas de acceso y diseñar las correspondientes intervenciones basadas en evidencia.

ACCESO A OPORTUNIDADES EN AMÉRICA LATINA

ÍNDICE

INTRODUCCIÓN 01 10

1.1 GUÍA DE LECTURA 17

MARCO CONCEPTUAL: ACCESIBILIDAD 02 18

2.1 COMPONENTES 23

2.2 METODOLOGÍAS DE ANÁLISIS 25

2.3 EL APORTE DE LAS NUEVAS TECNOLOGÍAS DE LA INFORMACIÓN 27

2.4 DESAFÍOS PARA LOS ANÁLISIS DE ACCESIBILIDAD 29

MODELO ANALÍTICO PARA MEDIR ACCESIBILIDAD 03 30

3.1 RESUMEN DE MODELO 35

3.2 PROCESAMIENTO DE DATOS DE ENTRADA 37

3.3 SIMULACIÓN DE VIAJES 44

3.4 INDICADORES DE ACCESIBILIDAD 53

3.5 VIRTUDES, LIMITACIONES Y MEJORAS 65

CASO DE ESTUDIO: REGIÓN METROPOLITANA DE BS.AS. 04 68

4.1 CONTEXTO 72

4.2 DATOS UTILIZADOS 73

4.3 INDICADORES 75

4.4 ANÁLISIS DEL CASO 85

CONCLUSIONES 05 86

GUÍA DE
LECTURA

{01}

01

GUÍA DE
LECTURA

PARTE 01

INTRO DUCCIÓN

- Breve recapitulación del proceso de urbanización en América Latina. Introducción a la accesibilidad e importancia de su análisis. Guía de lectura.

P. SECCIÓN
01

INTRODUCCIÓN

- | | | |
|-----|-----------------|----|
| 01. | Introducción | 13 |
| 1.1 | Guía de Lectura | 17 |

Introducción

América Latina ha experimentado un crecimiento acelerado de la población urbana durante el siglo XX. Este fenómeno fue advertido ya hacia 1976, cuando Herrera y Pecht (Herrera et al., 1976) señalaron especialmente la gran expansión de las áreas urbanas entre 1920 y 1960. En ese entonces, la llegada de inmigrantes y la movilidad desde áreas rurales hacia zonas urbanas iniciaron el proceso de urbanización de las ciudades en la región. Durante la segunda mitad del siglo pasado, las ciudades de los países en desarrollo experimentaron un proceso de urbanización tardío pero acelerado con respecto a países como Estados Unidos y a ciudades europeas, con incrementos en las tasas anuales de urbanización en América Latina superiores al 5% entre 1950 y 1970, y entre un 4% y un 5% en las décadas posteriores. Entre el año 2000 y 2010, el número estimado de áreas metropolitanas latinoamericanas de más de 100.000 habitantes aumentó un 55%, mientras que en Estados Unidos el aumento fue del 24% (Daude et al., 2017).

En 2012, América Latina fue considerada la región más urbanizada del mundo, con dos tercios de su población



residiendo en ciudades de más de 20.000 habitantes y alrededor de un 80% en zonas urbanas (CEPAL, 2012). Consecuentemente, la extensión del área urbana promedio en las ciudades de la región aumentó un 20% entre 2000 y 2010, pero la densidad poblacional promedio disminuyó levemente (cerca del 2%), según la Base de Extensión de Áreas Metropolitanas (BEAM) desarrollada por CAF y presentada en el informe RED de 2017 (Daude et al., 2017). Sin embargo, la adaptación a este fenómeno de procesos públicos como entidades, normas, regulaciones y planes no ocurrió a la misma velocidad, lo que permitió que la rápida urbanización y variación de la ocupación del suelo ocurriera sin planes ni supervisión del ordenamiento del territorio. Esto resultó en altos costos de urbanización, baja productividad de las ciudades, desigualdades sociales, aumentos de criminalidad, alta congestión vehicular, contaminación del aire, entre otros. A su vez, estos costos no recaen de manera equitativa en toda la ciudadanía (Daude et al., 2017).

La desigualdad territorial y el sistema de movilidad influyen en el nivel de accesibilidad de las personas a diferentes oportunidades. Las ciudades ofrecen empleo, servicios educativos, de salud,

culturales o recreativos, y las personas requieren moverse para realizar la mayoría de estas actividades relacionadas con su vida cotidiana. El nivel de accesibilidad de cada individuo determina la calidad y cantidad de actividades que este puede realizar. La disponibilidad de transporte público, la cercanía a las principales centralidades o la posesión de vehículo particular hacen que determinados niveles socioeconómicos cuenten con mejores posibilidades de acceder a los destinos con más y mejores oportunidades.

Un análisis de accesibilidad permite comprender cómo los individuos acceden a las oportunidades urbanas desde distintas zonas de la ciudad. En los Objetivos de Desarrollo Sostenible de la Organización de las Naciones Unidas (ONU), la accesibilidad surge como tema transversal, ya que un mejor acceso a los servicios y oportunidades esenciales para todos los ciudadanos promete un futuro más equitativo, sostenible y económicamente viable. Los análisis de este tipo permiten comprender las necesidades de infraestructura, servicios y planificación, priorizar intervenciones y definir mejores políticas territoriales y de movilidad.

Con esta motivación, en el marco del OMU, desarrollado por CAF en conjunto con BID, se presenta el siguiente modelo analítico para definir y calcular indicadores de accesibilidad urbana a centros de actividad y a distintos servicios básicos para distintos niveles socioeconómicos, con datos abiertos, y se describe el proceso de forma integral, desde la obtención de datos hasta la elaboración de gráficos y mapas para visualizar los resultados. Luego, se desarrolla una aplicación al caso de la Región Metropolitana de Buenos Aires (RMBA).

Guía de lectura

1.1

El objetivo de este documento es mostrar el desarrollo de una herramienta para medir la accesibilidad urbana en ciudades de América Latina. La metodología utiliza datos que son accesibles y de uso público en la mayoría de las ciudades. Los algoritmos para la elaboración de estos análisis se encuentran disponibles en una librería desarrollada en Python, que puede utilizarse para implementar esta metodología o para realizar otros análisis de accesibilidad en diferentes ciudades.^[03]

La parte 2 sirve como guía conceptual de la accesibilidad urbana con foco en la medición. Se revisan brevemente las definiciones existentes en la bibliografía y se adopta una de ellas para este documento. Se revisan también metodologías de análisis y las principales métricas de accesibilidad. Por último, se especifican los indicadores que se buscarán medir con la metodología propuesta en base a datos de acceso público, lo que permite una medición regional que cubre un conjunto de ciudades de América Latina.

En la parte 3 se detalla el modelo desarrollado para medir indicadores de accesibilidad en las ciudades seleccionadas de América Latina. Se presentan los requerimientos de información y las fuentes utilizadas, así como las definiciones metodológicas desarrolladas para poder llevar adelante una simulación de viajes, lo que permite realizar las mediciones de accesibilidad. El código utilizado es abierto, por lo que se incluyen referencias directas a las librerías de Python utilizadas.

Por último, la parte 4 presenta un caso de estudio con los resultados del modelo implementado, donde se muestran los indicadores con sus gráficos y mapas para la RMBA.

Finalmente, los anexos del documento son ricos en información metodológica, especialmente para las personas interesadas en replicar la metodología en otras ciudades y utilizar el paquete.

[03]

La librería `pyomu` es un paquete de código abierto que incluye los algoritmos desarrollados para implementar esta metodología. Se puede encontrar en <https://github.com/OMU-LATAM/> o se puede instalar el paquete en Python desde <https://pypi.org/project/pyomu/>

MARCO
CONCEPTUAL



ACCESI BILIDAD

- Definición de accesibilidad urbana y sus componentes. Perspectivas metodológicas para el análisis de accesibilidad. Beneficios de las nuevas tecnologías de la información para la toma y el análisis de datos. Desafíos metodológicos.

P.
SECCIÓN
02

ACCESIBILIDAD: MARCO CONCEPTUAL

02.	Accesibilidad: Marco conceptual	21
2.1	Componentes	23
F.1	Principales componentes de la accesibilidad territorial	23
2.2	Metodología de análisis	25
2.3	El aporte de las nuevas tecnologías de la información	27
2.4	Desafíos para los análisis de accesibilidad	29

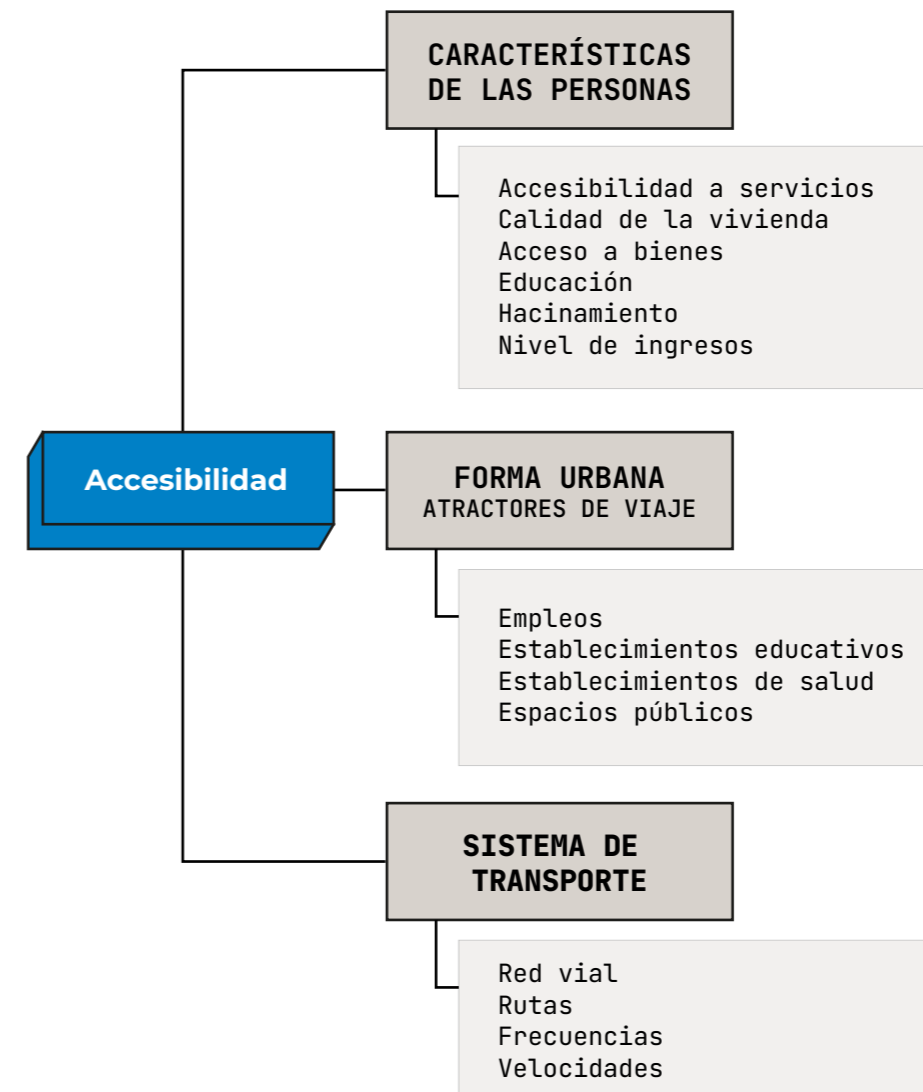
Accesibilidad: marco conceptual

La accesibilidad urbana es definida y operacionalizada de distintas formas y con distintos significados. A grandes rasgos, se define como la facilidad para acceder a bienes, servicios, actividades y destinos en una ciudad (Litman, 2003). Sin embargo, hay definiciones que encaran el problema desde diferentes perspectivas. Algunas hablan de “las oportunidades potenciales de interacción” (Hansen, 1959); la facilidad con la que se puede acceder a las actividades usando un modo particular de transporte (Dalvi y Martin, 1976); la libertad de los individuos de decidir participar o no de diferentes actividades (Burns, 1979); y los beneficios que produce un determinado sistema de usos de suelo y transporte (Ben-Akiva y Lerman, 1979). El Reporte de Economía y Desarrollo (RED) 2017 de CAF define la accesibilidad como “la capacidad de alcanzar las oportunidades que ofrece la ciudad”, y la considera la medida fundamental de bienestar en las ciudades. Además, la accesibilidad urbana se estudia como



Componentes

Los principales componentes que caracterizan la accesibilidad urbana incluyen la ubicación de los hogares, las características socioeconómicas, la forma urbana y distribución de actividades, y el sistema de transporte (ver figura 1).



La forma urbana está relacionada con la ubicación de los hogares y la relación que existe con los destinos laborales y de comercio, servicios primarios como salud y educación, y centros culturales y recreativos, lo que es determinante para comprender los niveles de accesibilidad de las personas. En este trabajo se desarrollan indicadores de accesibilidad para ciertos destinos de interés como la actividad económica, salud, educación y áreas verdes. Además, se utiliza el acceso al centro de la ciudad

uno de los factores que afectan el nivel de inclusión y productividad en las aglomeraciones urbanas, y se destaca su importancia en el crecimiento y desarrollo de las ciudades (Hernández y Hansz, 2021).

En este trabajo, se toman las definiciones del RED 2017 de CAF y de Hernández y Hansz para definir la accesibilidad urbana como “la capacidad de un espacio urbanizado y sus sistemas de proveer a todas las personas y comercios de oportunidades laborales, educativas, de salud, sociales y culturales de manera equitativa y eficiente”. La accesibilidad urbana es considerada fundamental para el crecimiento y desarrollo de las regiones.

Junto a los avances tecnológicos y la disponibilidad de nuevos datos, los nuevos estudios de accesibilidad urbana han despertado un renovado interés entre responsables políticos y tomadores de decisión. En este trabajo se explora el potencial de los datos abiertos y disponibles de uso público para construir indicadores de accesibilidad que en conjunto reflejen aspectos fundamentales de la capacidad de la infraestructura urbana y de transporte de cada ciudad para darle a sus ciudadanos el acceso a las oportunidades urbanas.

Metodologías de análisis

2.2

como un proxy de acceso a oportunidades, dada la gran variedad de actividades económicas y sociales que ocurren en los centros urbanos de las ciudades de la región.

Las características socioeconómicas como el ingreso, género, nivel educativo o edad determinan qué tipo de actividades o destinos son más relevantes, así como las posibilidades de realizar viajes a pie, en transporte público o en modos individuales como el auto, moto o bicicleta. La caracterización de la demanda es lo más desafiante desde el punto de vista de la información. En el caso de este estudio, se construye y considera un atributo de nivel socioeconómico (NSE) a partir de datos censales que permite caracterizar la demanda y poder ubicarla espacialmente.

Por último, el sistema de transporte, incluyendo sus calles, veredas y disponibilidad de transporte público termina de determinar las posibilidades que cada persona tiene de acceder a los distintos destinos. Esto último incluye tener en consideración toda la infraestructura física y virtual, así como los servicios públicos y privados, con sus reglas y normas, que coexisten en el espacio público de la ciudad para ofrecer distintas posibilidades de transporte. En este trabajo se tienen en cuenta los aspectos físicos de la infraestructura, que incluyen la malla vial y la red de transporte público, con sus frecuencias y recorridos. Además, se considera dentro del cálculo de tiempos el tráfico como medida de capacidad del sistema.

[04]

Otros autores, como Geurs y Van Wee, añaden un componente separado para las restricciones temporales tanto de las actividades a realizar en la ciudad como de la disponibilidad de los individuos (Geurs y van Wee, 2004). Sin embargo, en este documento se considera que el componente se encuentra inmerso en los tres componentes principales.

La interacción entre estos componentes determina la accesibilidad urbana.^[04] En este trabajo, se utilizan datos de libre acceso para caracterizar cada uno de estos componentes. Para identificar los principales atractores, como actividad, salud, educación y espacios verdes, se utiliza información generalmente disponible en sitios de datos abiertos gubernamentales y en Open Street Maps (OSM). Para la caracterización socioeconómica de las personas, se utilizan datos de los censos nacionales de población y viviendas. Finalmente, para caracterizar el sistema de transporte y obtener distancias y tiempos de viajes en la red vial, se utilizan las aplicaciones de ruteo que ya contienen esta información y permiten realizar consultas.

Si bien existe un consenso cada vez mayor sobre la importancia de medir la accesibilidad urbana como un insumo para la planificación y el desarrollo de políticas públicas, el uso de métricas de accesibilidad aún enfrenta importantes desafíos prácticos y metodológicos. Se pueden identificar cuatro perspectivas con las que se suelen realizar los análisis de accesibilidad (Geurs y van Wee, 2004):

01 → La primera, basada en mediciones centradas en la infraestructura de transporte, analiza principalmente el desempeño o nivel de servicio. Estas métricas incluyen niveles de congestión o velocidades promedio en la red vial.

02 → La segunda, más relacionada con la distribución espacial de las actividades, incluye indicadores como la cantidad de trabajos o centros educativos dentro de determinados rangos de tiempos (por ejemplo, el porcentaje de trabajos en un radio de 30 minutos de un lugar de origen).

03 → Una tercera analiza la accesibilidad en el nivel individual, es decir, las oportunidades que puede alcanzar una persona que reside en alguna parte de la ciudad, como la cantidad de actividades en las que un individuo puede participar en un determinado lapso de tiempo con sus posibilidades económicas, físicas y sociales.

04 → Por último, las mediciones basadas en los beneficios económicos o utilidades que las personas derivan del acceso a distintas actividades distribuidas espacialmente.

Muchos análisis de accesibilidad se han centrado tradicionalmente en una única perspectiva según el propósito con el cual hayan sido demandados, pero con el riesgo de perder la perspectiva integral de la accesibilidad. Por ejemplo, aquellos trabajos enfocados principalmente en el análisis del ahorro de tiempos de viaje y el alivio de la congestión tienden a no tener en cuenta otros aspectos más amplios relacionados con el acceso a oportunidades. Los nuevos enfoques de evaluación de cambios de usos de suelo y transporte intentan tener consideración sobre diferentes aspectos de la movilidad.^[05]

El modelo metodológico presentado en este documento incluye la oferta que brinda la infraestructura, la distribución espacial de las actividades y ciertas características a nivel individual. A partir de datos públicos y herramientas de uso abierto, se construyen los siguientes indicadores como parte del estudio de accesibilidad urbana de las ciudades del OMU:

[05]

Como ejemplo de un nuevo enfoque de accesibilidad, se encuentra la propuesta del Instituto de Transporte y Políticas de Desarrollo (ITDP), con once indicadores categorizados en tres dimensiones: proximidad al transporte, accesibilidad y características de la ciudad. Para más información, acceder a ITDP Online Indicators: <https://naindicators.itdp.org/>.

El aporte de las nuevas tecnologías de la información

01 → Índice de congestión: este indicador tiene en consideración las velocidades promedio a lo largo de un día hábil y se construye como el ratio entre la velocidad promedio en hora valle y la velocidad promedio en hora punta.

02 → Accesibilidad a actividades y al área central: esta serie de indicadores calcula los tiempos promedios de viaje y tiene en consideración las áreas de alta concentración de actividades (principales atractores) y el área central de una ciudad.

03 → Accesibilidad a educación, salud y espacios verdes: esta serie de indicadores calcula los tiempos de viaje hacia los distintos tipos de establecimientos.

04 → Indicador general de accesibilidad (clústeres de accesibilidad): construye un indicador de nivel de accesibilidad territorial teniendo en consideración los indicadores calculados anteriormente.

La disponibilidad de nuevas fuentes de información (conocidas como *big data*) y los avances en la ciencia de datos brindan nuevas oportunidades para el análisis de la accesibilidad. La escala, alcance y velocidad de los datos viene aumentando debido a la gran cantidad de dispositivos con sensores que se encuentran interconectados y que producen información en tiempo real con capacidad computacional de procesamiento de datos. La información que producen estos dispositivos, que se encuentran en autos, teléfonos inteligentes, puestos de peajes, relojes, aplicaciones, cámaras de seguridad, routers o sensores ambientales, permite conocer la ubicación, los recorridos y las velocidades de traslado. Con esta información, se pueden estimar orígenes y destinos de viajes, principales atractores y la distribución de la población en diferentes días y horarios. La disponibilidad de herramientas de código abierto para el análisis de datos, como por ejemplo Python, R o QGIS, vuelve además la capacidad analítica mucho más accesible (Ronan, 2017).

Asimismo, las instituciones públicas, privadas y los mismos usuarios están cada vez más dispuestos a compartir información. Las instituciones públicas y las empresas privadas (aunque a veces con un costo o restricción para el acceso) están llevando adelante procesos de apertura de sus datos a través de sus páginas web o API (Interfaces para la Programación de Aplicaciones),^[06] lo que promueve una mayor transparencia, responsabilidad y creación de valor. Como resultado, datos de censos, encuestas domiciliarias, cartografía y datos generados por distintos dispositivos, requeridos para la realización de análisis de accesibilidad, se encuentran cada vez más disponibles (Ronan, 2017). Esto se complementa con un aumento en la capacidad de los usuarios de compartir sus propios datos en forma voluntaria (*crowdsourced data*). Muchas aplicaciones que ofrecen información para el ruteo de viajes, servicios de delivery o reservas permiten a los usuarios compartir detalles sobre sus visitas a distintos lugares, como restaurantes, bares o negocios. Esto abre nuevas posibilidades para conocer el comportamiento y las preferencias de las personas en relación con las actividades que realizan en la ciudad.

Estos cambios permiten repensar los estudios tradicionales de accesibilidad y desarrollar nuevas metodologías e indicadores para comprender mejor el modo en

[06]

Las Interfaces de Programación de Aplicaciones (API en inglés) son protocolos que permiten la conectividad entre aplicaciones. De esta forma, la API de Google Maps o OSM permite realizar consultas en forma masiva a través de código de programación.

que los diferentes grupos experimentan la accesibilidad espacial en las ciudades. A continuación se presentan fuentes alternativas de datos para realizar análisis de movilidad que en la actualidad se encuentran disponibles. Se detallan algunos ejemplos para los distintos tipos de datos de fuentes alternativas (llamados en muchos casos *big data*). Muchas de las metodologías para analizar estas fuentes se están probando actualmente, y en el futuro será cada vez más común realizar análisis de movilidad utilizando fuentes alternativas de información.

REGISTROS ADMINISTRATIVOS: origen y destino de viajes, tiempos de viajes, condición de vehículos

- 01 → Tarjetas electrónicas de pagos de boletos
- 02 → Verificaciones vehiculares
- 03 → Registros de bicicletas públicas

CROWD-SOURCED DATA: información cualitativa de percepción del transporte, cortes y eventos

- 01 → Comentarios sobre transporte en redes sociales (Twitter, Instagram, Facebook)
- 02 → Comentarios sobre lugares
- 03 → Eventos en Waze

APLICACIONES EN TELEFONÍA CELULAR: movimiento, velocidad, duración de las estadías, principales atractores

- 01 → Registros de antenas
- 02 → Aplicaciones de ruteo (Google Maps)
- 03 → Aplicaciones con georreferenciación (Strava, Google Fit)

SENSORES: situación del tránsito en tiempo real

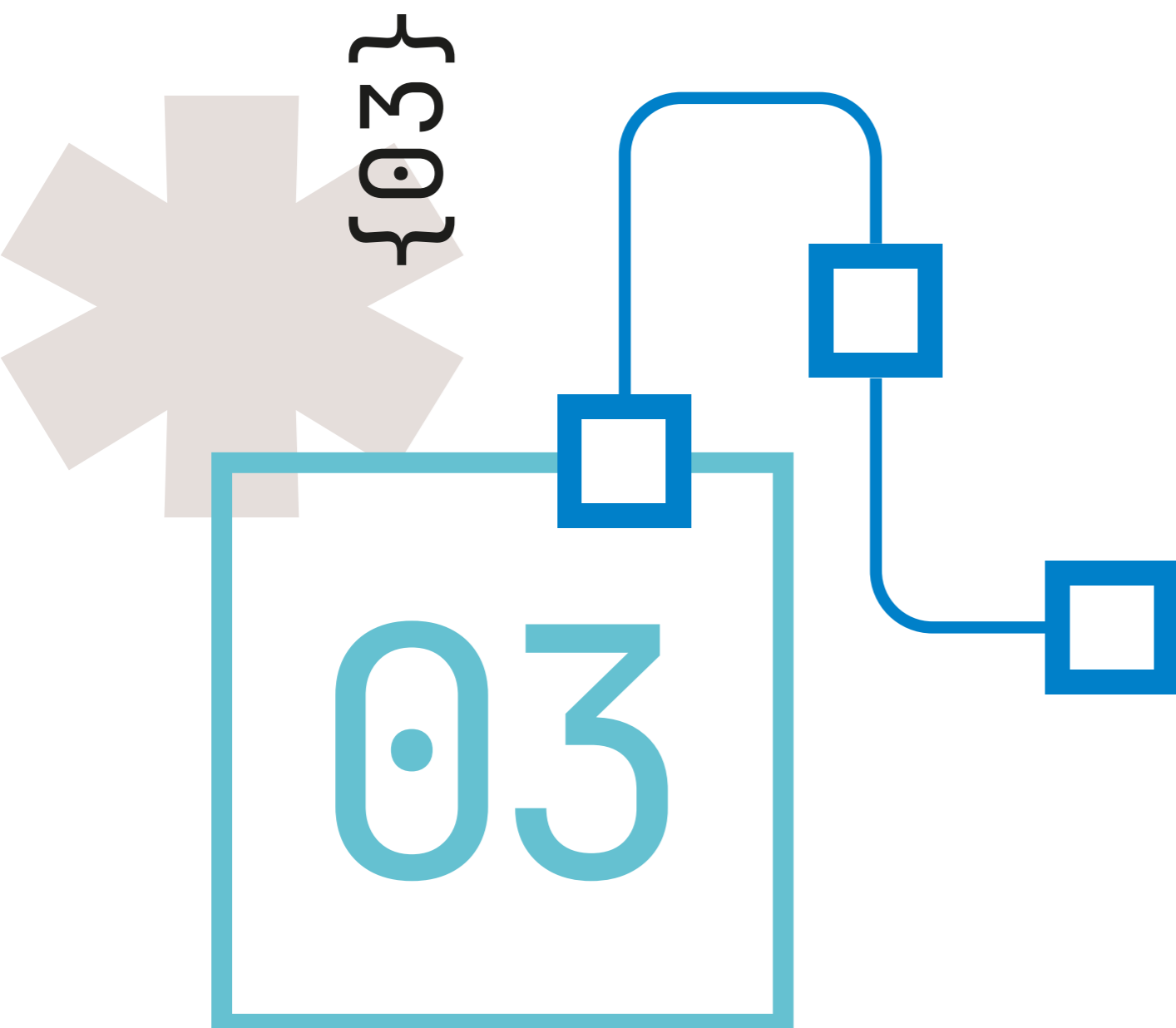
- 01 → Sensores de peaje
- 02 → Cámaras en vía pública
- 03 → Detectores de velocidad

Desafíos para los análisis de accesibilidad

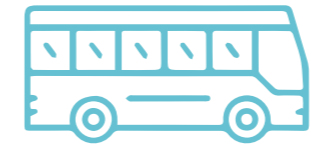
La disponibilidad de nuevos datos y herramientas permite evaluar la accesibilidad de maneras más sofisticadas y computar métricas desagregadas por actividad, propósito de viaje, modo, tiempo, edad, sexo, y grupo ocupacional (Ronan, 2017). El desafío es, entonces, construir indicadores que ayuden a comprender la accesibilidad en distintas ciudades y que, en forma adicional, sirvan para comparar sus sistemas de movilidad. A su vez, se debe procurar desarrollar indicadores que sean replicables y que estén contruidos con datos de acceso público. Preferentemente, estos indicadores no deberían utilizar índices que combinen información múltiple y sean difíciles de comprender, para promover así su transparencia y accionabilidad (Carruthers et al., 2005).

La construcción de un sistema de indicadores de accesibilidad debe servir para compartir los resultados no sólo con especialistas técnicos, sino con tomadores de decisión y otros miembros interesados de la comunidad. Por último, en el desarrollo de indicadores de accesibilidad es todavía una deuda pendiente incluir los factores subjetivos de los individuos (Ronan, 2017). Es sumamente importante contar con información sobre la percepción de los usuarios, sobre sus experiencias de viaje y sobre cómo estos factores subjetivos pueden alterar la selección de rutas, costos y tiempos de viaje, especialmente para quienes la experiencia de viaje es actualmente un impedimento o un costo alto que no se considera en los estudios tradicionales.

El modelo presentado a continuación desarrolla información valiosa para realizar estudios de accesibilidad en la región con datos disponibles de uso público y puede ser replicado en un conjunto de ciudades. Esta metodología utiliza información socioeconómica de los habitantes y construye un conjunto de indicadores de fácil interpretación. No se incluyen factores de percepción o aquellos aspectos cualitativos que pueden afectar las posibilidades reales de los y las habitantes de las áreas urbanas. Estos son una deuda pendiente del presente trabajo, y se analizará cómo podrán abordarse estos aspectos en el futuro.



PARA MEDIR
ACCESIBILIDAD



PARTE 03

MODELO ANALÍTICO

- Construcción de un modelo analítico para el análisis de la accesibilidad urbana. Indicadores y pasos operativos del modelo. Información y datos utilizados. Limitaciones y oportunidades de mejora.



MODELO ANALÍTICO PARA MEDIR ACCESIBILIDAD

3.1	Resumen del modelo	35
F.2	Resumen de pasos del modelo	35
3.2	Procesamiento de datos de entrada	37
3.2.1	Información socioeconómica y cartográfica	37
M.1	NSE en la RMBA en quintiles	40
3.2.2	Atractores de viajes	40
M.2	Densidad de actividad en la RMBA	42
3.2.3	Establecimientos específicos: salud, educación y espacios verdes	43
3.3	Simulación de viajes	44
3.3.1	Sistema de transporte	44
3.3.2	Aplicaciones de ruteo	45
3.3.3	Orígenes y destinos	49
M.3	Origen y destino para la simulación de viajes en el área metropolitana de Bogotá	50
3.3.4	Ejecución de la simulación	50
M.4	Red de hexágonos y ubicación de hospitales en la RMBA	52
3.4	Indicadores de accesibilidad	53
3.4.1	Índice de congestión	54
F.3	Tiempos y velocidades promedio de viajes en un día hábil en Bogotá, Colombia	55
3.4.2	Accesibilidad a principales centralidades	56
M.5	Tiempos de viaje a centros de actividad en Valle de Aburrá	57
M.6	Tiempos de viaje a México DF	58
3.4.3	Cobertura del sistema público de transporte	58
M.7	Distancia caminata en origen para viajes en transporte público, Montevideo	59
3.4.4	Accesibilidad a salud y educación	62
3.4.5	Accesibilidad a espacios públicos o áreas verdes	60
M.8	Espacios públicos Bogotá, Colombia	61
3.4.6	Indicadores de accesibilidad por NSE	62
F.4	Tiempos promedio de viaje en transporte público. Montevideo, Uruguay	62
3.4.7	Niveles de accesibilidad general	63
M.9	Clústeres de accesibilidad CDM	64
3.5	Virtudes, limitaciones y mejoras	65

Modelo analítico para medir accesibilidad

Dado que existen desafíos para analizar la accesibilidad en áreas urbanas, como se mencionó en el apartado anterior, esta sección presenta la construcción de un modelo analítico que permite analizar y elaborar indicadores de accesibilidad utilizando fuentes de datos abiertos de acceso público. Este modelo se va a aplicar en ciudades de América Latina con el principal objetivo de desarrollar una serie de indicadores para asesorar al sector público en la identificación de brechas de acceso y en la formulación de políticas públicas. Los resultados de indicadores de accesibilidad para al menos 22 ciudades de América Latina forman parte de la tercera edición del OMU y serán disponibilizados en la página web junto con el código de Python que los generó.

Esta sección desarrolla, en primer lugar, un resumen de los pasos operativos para llevar a cabo el

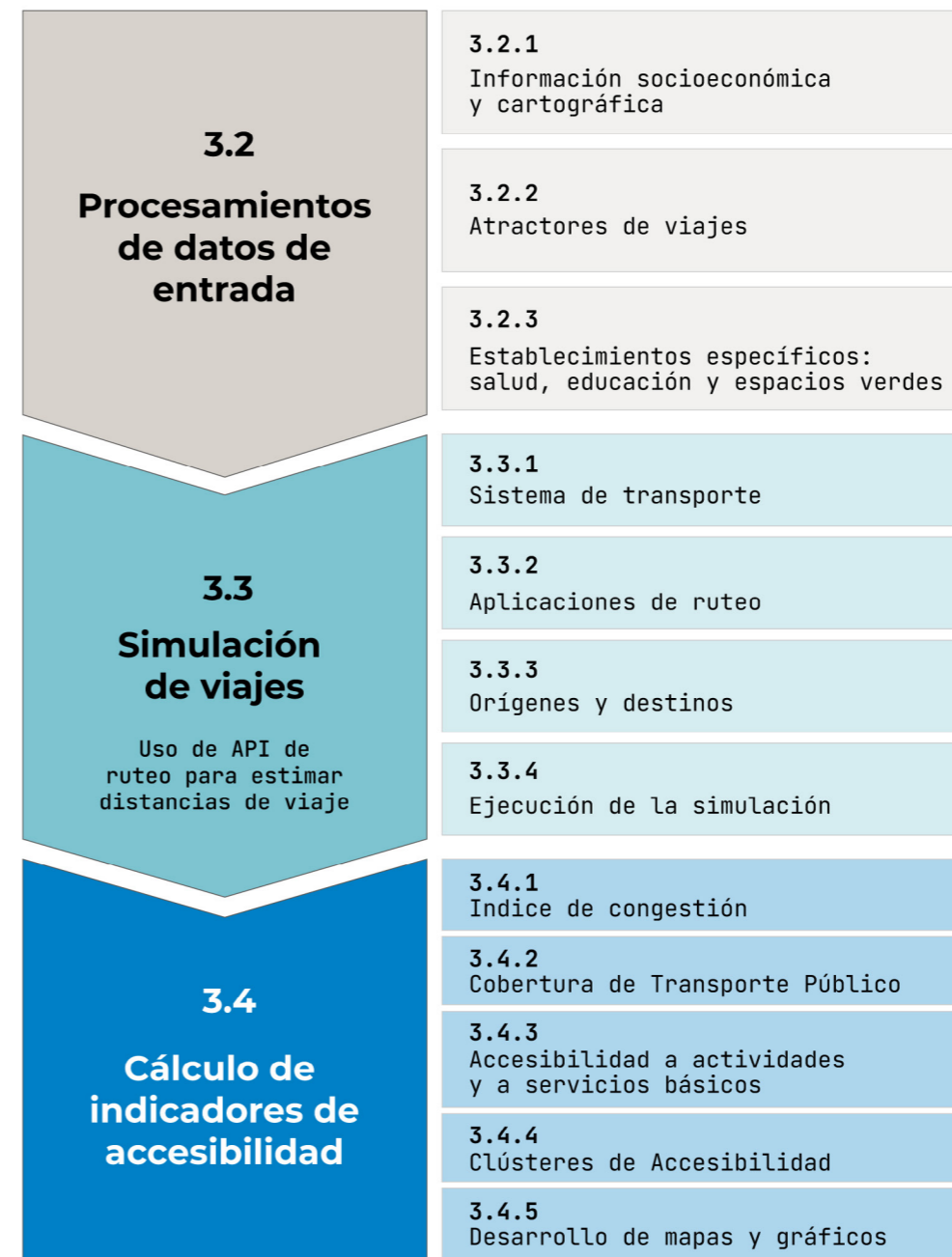
modelo propuesto. En segundo lugar, se detalla cada etapa del modelo. El proceso comienza con la recopilación y preprocesamiento de datos socioeconómicos y datos de ubicación de establecimientos para construir los elementos de entrada del modelo. Luego, se describe el procedimiento de simulación de viajes, de donde se extrae la información necesaria para calcular indicadores de congestión, accesibilidad, y cobertura. Al final, se presentan las virtudes, limitaciones y posibilidades de mejoras del modelo.

Resumen del modelo

3.1

A continuación se resumen los pasos operativos del modelo analítico para construir los indicadores y desarrollar las visualizaciones para contribuir con una herramienta ágil e interpretable sobre el estudio de la accesibilidad territorial en las ciudades.

Pasos del modelo



El primer paso consiste en obtener y procesar una serie de datos de entrada para generar orígenes y destinos de viajes. Para los orígenes, es posible identificar con datos demográficos y cartográficos cómo está distribuida la población y, seleccionando un conjunto de variables censales, construir un indicador de NSE para cada unidad censal. Para la homogeneización de la cartografía y eficiencia del algoritmo, se transforman las unidades censales espaciales a hexágonos de igual tamaño, a los cuales se les asigna la población (proporcional a su área) y el NSE, lo que asegura una completa representación del territorio. Los hexágonos con la información de la población y el NSE asignados se utilizarán como orígenes para los viajes.

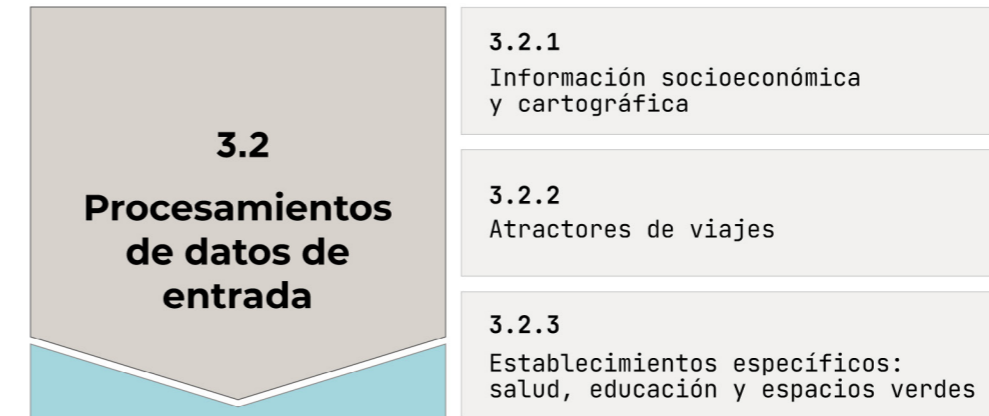
Por otro lado, para los destinos, se identifican puntos de destinos principales de los viajes realizando un análisis de la densidad de actividades en la ciudad. Para llevar adelante este análisis se obtienen los equipamientos urbanos de OSM, se realiza una ponderación por tipo de equipamiento y se identifican clústeres de concentración de actividad. Además, para realizar análisis de accesibilidad específicos, se recolecta la ubicación geográfica de equipamientos, como establecimientos educativos y de salud, y espacios públicos o verdes.

A partir de la información de los orígenes y los destinos, se realiza, como segundo paso, la simulación de los viajes consultando las API de las aplicaciones de ruteo. Las consultas en OSM permiten obtener distancias en la red, y las consultas en Google Maps permiten obtener los tiempos de viaje en distintos modos de transporte, incluyendo predicciones de tráfico para los viajes en automóvil y las transferencias en transporte público. Google permite obtener tiempos de viaje entre orígenes y destinos con estimaciones en base a información real de los usuarios de la aplicación.

Una vez construida una base de datos con los resultados de las simulaciones, el tercer paso consiste en la elaboración de los indicadores de accesibilidad, congestión y cobertura, y finalmente se desarrollan las visualizaciones que ayudarán a los y las tomadoras de decisiones a analizar la accesibilidad urbana en el territorio.

Procesamiento de datos de entrada

3.2



El primer paso requiere la recopilación de diferentes fuentes de datos útiles para representar dos de los componentes de la accesibilidad territorial: la forma urbana y las características de los habitantes. Para el componente que tiene en cuenta el sistema de transporte, se utilizarán las aplicaciones de ruteo de viaje que ya contienen la información necesaria.

A continuación, se resumen las distintas piezas de información y datos requeridos para la construcción del modelo analítico desarrollado en el estudio: información socioeconómica y cartográfica junto con la localización de equipamientos urbanos y servicios básicos, sumado a la información de la oferta de transporte.

3.2.1 Características de las personas: Población y nivel socioeconómico. Fuentes: censos de población y viviendas.

3.2.1 Forma urbana: Zonas de atracción de viajes. Fuentes: equipamientos urbanos (OSM), localización de servicios básicos (datos abiertos de ministerios de educación y salud).

3.2.1 Sistema de transporte: Oferta. Fuentes: aplicaciones de ruteo de viajes, información geoespacial de la red de transporte (OSM), datos del servicio de transporte público (GTFS).

Información socioeconómica y cartográfica

3.2.1

La información socioeconómica y la cartografía utilizada en este análisis proviene de los censos de población y viviendas elaborados por los institutos nacionales de

estadísticas de cada país. Se trabaja con la unidad geoestadística de desagregación más pequeña que ofrezca el censo en cada caso, como, por ejemplo, radios o secciones censales. Para caracterizar a los hogares y personas, se seleccionan una serie de datos de viviendas, hogares o personas de la información censal y se calcula un indicador de NSE para cada unidad geoestadística. ^[07] Para lograr una base cartográfica homogénea en toda la región, se utilizan divisiones espaciales hexagonales de igual tamaño sobre las que se distribuye la población y el NSE calculado con información censal en forma proporcional a la superficie de intersección.

Una vez procesados los datos censales y la cartografía, se construye un indicador de NSE que caracteriza cada unidad estadística. Para la construcción de este indicador, es necesario identificar variables que puedan caracterizar lo mejor posible al conjunto de los hogares y su población teniendo en consideración la disponibilidad de datos censales de cada país. En términos generales, se busca incluir indicadores relacionados con los siguientes aspectos:

- 01** → Materiales de la vivienda (piso, techo, paredes).
- 02** → Acceso a servicios (electricidad, agua, saneamiento).
- 03** → Bienes (automóvil).
- 04** → Acceso a medios de comunicación (teléfono fijo, internet).
- 05** → Hacinamiento habitacional (más de 2 personas por habitación).
- 06** → Educación.
- 07** → Nivel de ingresos.

Una vez seleccionadas las variables censales para cada ciudad de acuerdo a las características del censo de cada país, se trabaja con el nivel geoestadístico más pequeño que se encuentre disponible, se normalizan las variables como porcentaje de la población de cada unidad geoespacial

(por ejemplo, el porcentaje de viviendas con hacinamiento crítico en cada radio/sección), y se construye un indicador de NSE utilizando el método de componentes principales, una técnica estadística utilizada para reducir la dimensionalidad de las variables seleccionadas. Con el resultado de este proceso, se calcula el NSE en terciles y quintiles

La librería `pyomu` permite calcular el NSE para los polígonos censales, crear los hexágonos, y distribuir la población y el NSE a nivel de hexágono. La función requiere una serie de parámetros como la capa cartográfica, las variables que serán utilizadas para generar el indicador de NSE, la variable que contiene la población del polígono y el nivel de resolución de los hexágonos.

`pyomu.calculate_nse_in_hexagons()`

teniendo en consideración la población de cada unidad geoestadística). Como resultado de este proceso, cada radio o sección censal tiene asignado un NSE (Alto, Medio o Bajo, en tres niveles, o Alto, Medio Alto, Medio, Medio Bajo y Bajo, en cinco niveles). En el anexo se encuentra una descripción más detallada del proceso de construcción del NSE.

A efectos de normalizar la cartografía censal, trabajar con áreas homogéneas en toda la región y optimizar el algoritmo del modelo, se utiliza como división geográfica un sistema jerárquico de grillado en forma de hexágonos que cubre todo el territorio de la ciudad o área metropolitana. Este reticulado es globalmente uniforme y fue creado por Uber para optimizar sus procedimientos de asignación de precios y despachos de viajes, y para visualizar y explorar los datos geoespaciales en distintas resoluciones utilizando un sistema de indexación jerárquico. Uber liberó para uso público los códigos y librerías de programación que permiten generar, agregar y desagregar, y analizar con distintos algoritmos este sistema de hexágonos, los cuales fueron popularmente adoptados por expertos en datos geoespaciales. ^[07]

Como parte del preprocesamiento de datos y una vez realizado el proceso de componentes principales, para cada radio o sección censal se superpone esta capa geográfica con la grilla de hexágonos para asignarles en forma proporcional al área del hexágono la población y el resultado del proceso de componentes principales. ^[08] Teniendo en consideración que hay unidades geoestadísticas de gran tamaño, para la redistribución de la población se superpone en forma adicional una capa obtenida de OSM de parques, espacios públicos y otros establecimientos que suelen no tener población. De esta forma, se evita asignar población en hexágonos que no deberían tenerla, por ejemplo, en el caso de que haya un gran parque o espacio público. Por último, se reasignan los quintiles y terciles de los NSE calculados a cada hexágono.

El mapa 1 muestra la división geoespacial en hexágonos de la RMBA, con los cinco colores correspondientes a los quintiles socioeconómicos: bajo, medio-bajo, medio, medio-alto y alto. De esta manera, los viajes que se originan en cada unidad geográfica, sea censal o tipo hexágono, son caracterizados por un NSE.

[07]

Las unidades geoestadísticas son porciones de un territorio (representadas como polígonos en un mapa) que se utilizan para referir información estadística.

Bajo esta denominación se contemplan tanto las divisiones políticas (por ejemplo, provincias, departamentos, gobiernos locales) como aquellas unidades generadas por un instituto a fines operativos (por ejemplo, localidad censal, fracción censal, radio o sección censal).

[07]

<https://h3geo.org/#/>

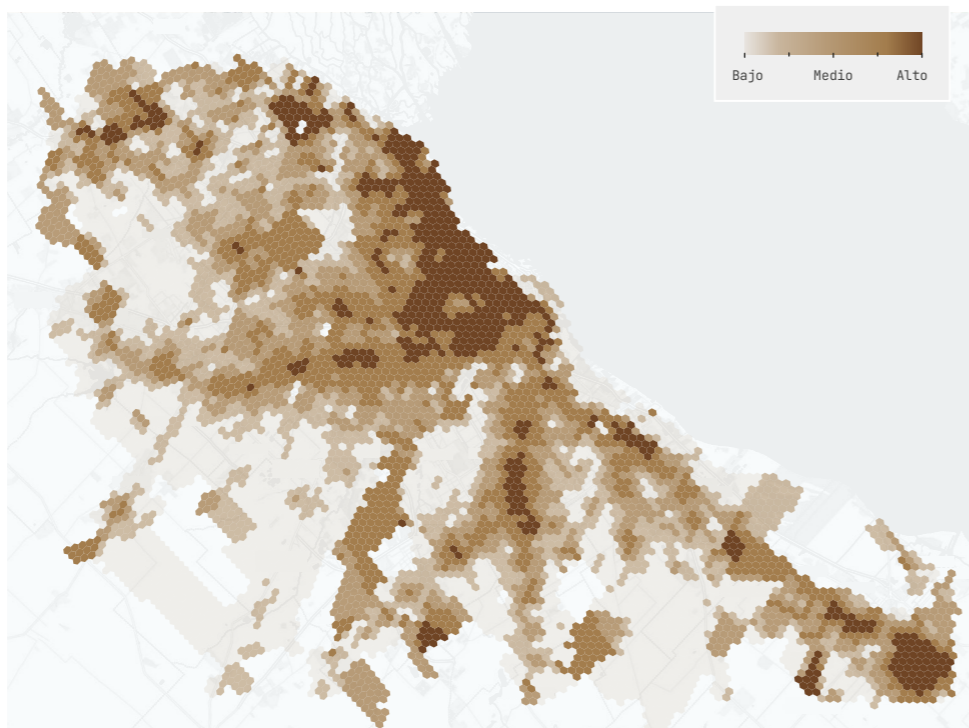
[08]

Cuando un hexágono contiene áreas de más de un radio o sección censal, se distribuye en forma proporcional con un promedio ponderado por población del análisis de componentes principales (PCA) de cada radio.

NSE EN LA
REGIÓN METRO-
POLITANA DE
BUENOS AIRES
EN QUINTILES

M.
01

Fuente: Elaboración propia en base a datos del censo 2010.



Atractores de viajes

3.2.2

Además de identificar las características socioeconómicas de la población que habita en las ciudades, es necesario conocer las actividades que desarrollan las personas y los lugares que frecuentan para poder estimar sus necesidades de movilidad de acuerdo a los principales atractores de viajes. Es por eso que otra pieza clave para el desarrollo de un modelo analítico de accesibilidad es la información de la demanda de transporte. A través de un mayor entendimiento de los patrones de movilidad de las personas, las elecciones de modos de transporte, los horarios de los viajes y los destinos, se puede explicar el rol del transporte en las actividades de la ciudad.

La falta de información sobre demanda de viajes en muchos casos es un desafío para la realización de este tipo de análisis. Existen diferentes fuentes de datos que permiten estimar esta información; las encuestas domiciliarias de movilidad son las más utilizadas tradicionalmente. Sin embargo, en los últimos años, se han estado investigando nuevos métodos de estimar la demanda de transporte utilizando datos alternativos a las encuestas tradicionales, como datos de

aplicaciones de telefonía celular. El principal problema de estos datos es que todavía son difíciles de obtener, ya que en general pertenecen a empresas privadas; además, las metodologías para interpretarlos aún están poco desarrolladas.

Uno de los objetivos de este modelo es la utilización de datos que sean accesibles para todas las ciudades, lo cual otorga escalabilidad al método. Por lo tanto, para incluir un estimador de demanda en el modelo con esta característica, se estima la densidad de actividad de las unidades geográficas utilizando datos de equipamientos urbanos de OSM, que son de acceso público. Se construye un indicador que permite identificar agrupamientos o clústeres a nivel territorial de alta densidad de equipamientos urbanos, asumiendo esto como un proxy de atracción de viajes en una ciudad. A continuación, se explica el procedimiento utilizado para generar los puntos de más densidad de actividad y la obtención de una serie de establecimientos urbanos específicos.

Para ello, se trabajó con los equipamientos urbanos como proxy de actividad. Los equipamientos urbanos incluyen desde comercios de distinto tipo hasta edificios públicos o de servicios, como hospitales o escuelas, entre otros. La ventaja de elegir equipamientos urbanos para identificar áreas de alta densidad de actividad es que estos equipamientos pueden obtenerse georreferenciados desde OSM, lo que vuelve este método escalable y útil para una gran cantidad de ciudades en América Latina. Otros métodos más tradicionales, como el uso de encuestas de movilidad, requieren por un lado la disponibilidad de encuestas, y por el otro, la elaboración del modelo clásico de transporte, lo que resulta en un método poco escalable.

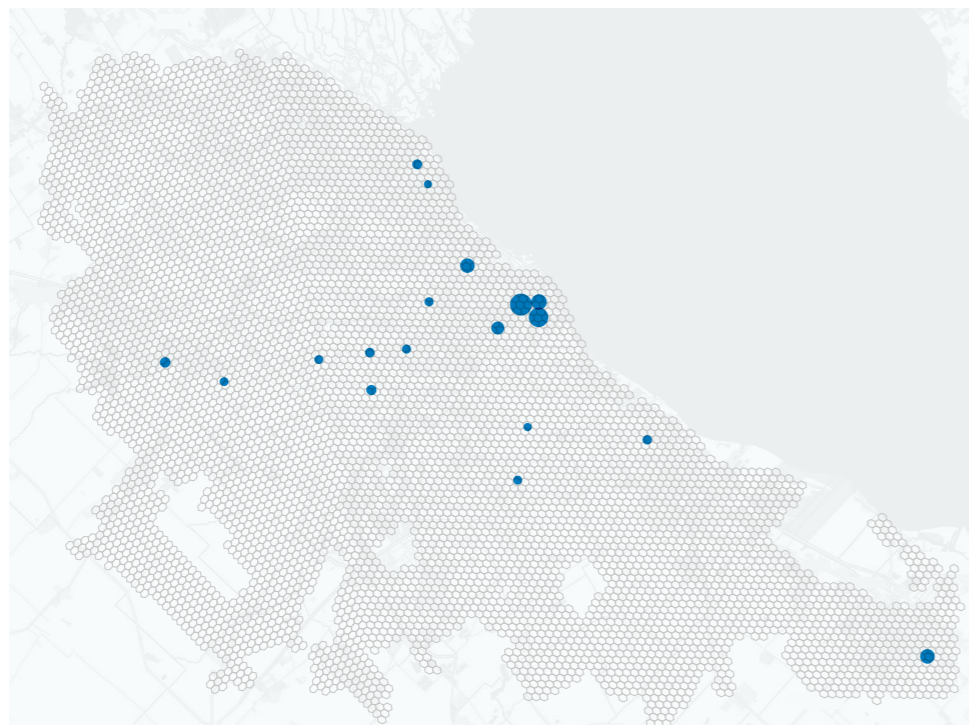
A través de la librería `pyomu` se puede ejecutar una función que identifica las áreas o puntos de alta densidad de actividad. Estos puntos se identifican según la densidad de establecimientos, que pueden ser datos de entrada o descargarse desde OSM, y sus pesos específicos. Esta función requiere una serie de parámetros, como la capa cartográfica para identificar la ciudad o área metropolitana, cantidad de clústeres a identificar y equipamientos a descargar de OSM. La ponderación según el tipo de establecimientos mediante los pesos específicos tiene valores predeterminados pero es flexible para que sean parametrizables por el usuario.

`pyomu.calculate_activity_density()`

Una vez obtenidos los equipamientos de distinto tipo, se pueden identificar aquellos más relacionados con el trabajo, salud, educación u otras actividades que sirvan como atractores de viajes en la ciudad. La alta concentración de este tipo de equipamientos permite suponer un área donde hay alta densidad de actividades. Por este motivo, se considera en este análisis que la densidad de equipamientos explica gran parte del uso del espacio y que, según el tipo de establecimiento, la

densidad y la identificación de ciertos tipos de establecimientos son variables que pueden usarse para identificar áreas con alta demanda de viajes.

Para poder identificar áreas de alta densidad de actividad en las ciudades, se trabajó con la información georreferenciada de equipamientos que se encuentra en OSM y que puede descargarse para su utilización. Una vez descargados los equipamientos, se desarrolló una metodología para identificar áreas de alta densidad de actividad. Esta metodología se desarrolla con más detalle en el anexo. En el siguiente mapa se pueden visualizar los clústeres de densidad de actividades para la RMBA.



CLÚSTERES DE DENSIDAD DE ACTIVIDAD EN LA REGIÓN METROPOLITANA DE BUENOS AIRES

M. 02

Fuente: Elaboración propia.

3.2.3

Establecimientos específicos: salud, educación y espacios verdes

Además de identificar las áreas de alta densidad de actividad con los equipamientos obtenidos de OSM, se realizó un análisis específico para cierto tipo de establecimientos que incluyen los educativos, de salud, y parques o espacios públicos. La información sobre establecimientos educativos y de salud suele estar disponible como capas geográficas en sitios de acceso público de los ministerios de educación y salud de municipios o departamentos, provincias o estados, o estados nacionales. Cuando estas no estaban disponibles, fueron solicitadas a las secretarías de movilidad que trabajan en conjunto con el OMU. Los establecimientos educativos son categorizados según su nivel (inicial, primario, secundario) y según sean públicos o privados; los establecimientos de salud, según el nivel de atención (atención primaria y hospitales).

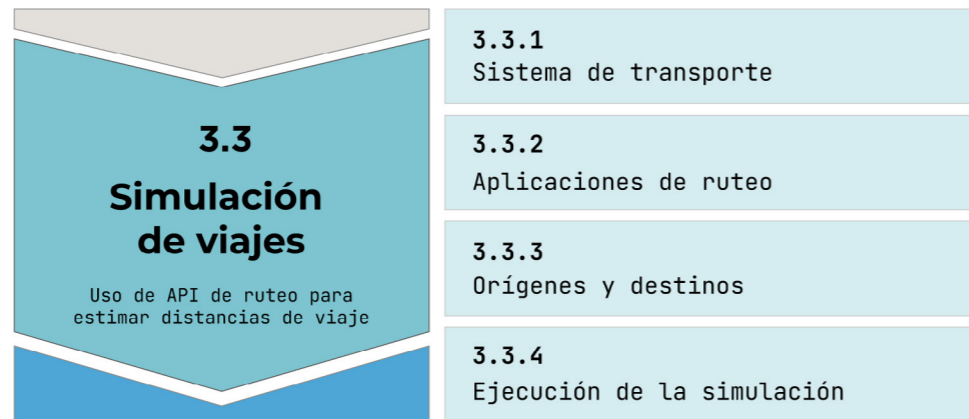
Los espacios verdes o públicos en algunos casos están disponibles en sitios de datos abiertos gubernamentales, pero para este análisis se obtuvieron de la información disponible en OSM. En OSM existen dos tipos de establecimientos que contienen información sobre parques, espacios públicos y reservas naturales. Esta información se obtiene en una capa de polígonos bajo las categorías “ocio” y “naturaleza” (en inglés “leisure” y “natural”). A su vez, estos polígonos están subcategorizados según su función (por ejemplo, parques, lagos, campos de golf o clubes municipales, entre otras categorías).^[10] De esta información se identifican espacios públicos o verdes de acceso público, que son los utilizados para el análisis de accesibilidad a espacios verdes o públicos. A su vez, como se mencionó en la sección anterior, estos datos son utilizados para la imputación de población de los polígonos censales hacia los hexágonos.

[10]

La lista completa se puede ver en https://wiki.openstreetmap.org/wiki/Map_features#Leisure.

Simulación de viajes

3.3



Es cada vez más común que con la disponibilidad de fuentes de datos alternativos se puedan recrear situaciones modelando el mundo real y obtener así información a través de simulaciones. De esta forma, podemos suplementar la necesidad de realizar observaciones en el mundo real que resultan complejas y costosas. En este trabajo se utilizan aplicaciones de ruteo de viajes para realizar simulaciones de viajes en distintas ciudades de América Latina, lo que permite obtener datos para el análisis de accesibilidad.

La simulación de viajes aquí propuesta selecciona una serie de puntos de origen que representan todo el territorio de la ciudad y una serie de puntos que representan destinos de viajes. Los destinos son áreas identificadas como de alta atracción o alta densidad de actividad, como se mencionó anteriormente. Las simulaciones se realizan a través de aplicaciones de ruteo de viajes que contienen información detallada de la oferta del transporte, y en algunos casos, información real del tráfico y frecuencias del sistema de transporte. Se utilizan las aplicaciones de ruteo de Google Maps y de Open Streets a través de sus API de acceso público.

Sistema de transporte

3.3.1

Para estudiar la accesibilidad, este modelo analítico requiere tener conocimiento de la oferta de transporte. Esto implica tener información sobre tres componentes del sistema: la infraestructura, que incluye obras físicas, como calles, veredas, ciclovías, autopistas; los vehículos que operan sobre esta infraestructura,

como automóviles, colectivos, metros o trenes, incluyendo información de las estaciones y paradas; y las condiciones y reglas de operación que determinan la funcionalidad del sistema, como frecuencias, horas de operación, costos, etc. Contar con información de la oferta del transporte permite realizar simulaciones entre distintos orígenes y destinos, lo que permite estimar tiempos de viajes, distancias, distintas alternativas de viaje y transferencias entre modos de transporte público.

OFERTA DE TRANSPORTE

01 → Infraestructura: calles, ciclovías, autopistas, paradas de transporte público.

02 → Vehículos: buses, metro, tren, bicicleta, caminata.

03 → Parámetros operativos: rutas, frecuencias, costos.

Aplicaciones de ruteo

3.3.2

Dado que es difícil recolectar toda la información sobre los sistemas de transporte para todas las ciudades de América Latina, las aplicaciones de ruteo nos ofrecen una alternativa con información de alta calidad. Estas aplicaciones permiten realizar consultas de viajes entre orígenes y destinos con la posibilidad de obtener información sobre los viajes y sus modos, incluyendo las alternativas de viajes en distintos modos y estimaciones muy precisas de los tiempos y distancias de los viajes. Estas aplicaciones funcionan gracias a que existen protocolos que permiten ordenar y compartir las piezas de información que completan el mapa de la oferta de transporte. En el cuadro 1, se detalla particularmente el protocolo diseñado por Google para la información de transporte público.

Cuadro 1. Protocolo GTFS

El protocolo más utilizado es el GTFS (General Transport Feed Specification o Especificación de Alimentación de Tránsito General), que define un formato común para toda la información de oferta de transporte público, como rutas, horarios, paradas, etc. Los GTFS permiten que los gobiernos y empresas de transporte público de todo el mundo publiquen sus datos de oferta de transporte y que los desarrolladores de aplicaciones puedan utilizar esta información de manera interoperable. Esta homogeneización de la información permite a las aplicaciones brindar herramientas para la planificación de viajes, teniendo en consideración aspectos particulares de la información de tránsito, como paradas, rutas, viajes y otros datos relacionados con los servicios de transporte público. Algunas aplicaciones incluso utilizan información que recolectan en tiempo real para realizar estimaciones de tráfico que puedan afectar los tiempos de viaje.

Archivos que deben contener los GTFS

Específicamente, un GTFS se refiere a una carpeta comprimida con archivos en formato .txt. Algunos de estos archivos son obligatorios para que el GTFS sea funcional y otros, opcionales, proporcionan aun más información, que puede enriquecer la aplicación que lo utiliza.

Los archivos obligatorios contienen información sobre las empresas de transporte público, las paradas y estaciones, las rutas y viajes, así como los horarios de llegada de los vehículos a cada estación según el día de la semana y del año.

Como ejemplo, Google Maps comparte con sus usuarios una descripción general de los datos de GTFS, una guía para construir tus propios feeds GTFS, una herramienta de validación del formato, y una guía para incluir un feed en Google Maps. Esta información permite planificar viajes en transporte público en la aplicación y obtener principalmente tiempos y distancias de caminata, de espera y de viaje.

La utilización de estas aplicaciones de rutas o planificación de viajes ofrece una oportunidad para acceder a la información de la oferta de transporte en una ciudad en forma sencilla y a un costo relativamente bajo. A su vez, muchos de estos servicios ofrecen la posibilidad de realizar consultas a través de API, lo que permite estimar viajes en forma masiva a través de códigos de programación. El cuadro 2 muestra diferentes aplicaciones y sus características para realizar este tipo de análisis.

Cuadro 2. Aplicaciones de ruteo para consultas de viajes

La utilización de aplicaciones para la consulta de viajes resulta conveniente para este tipo de estudios ya que incluye algoritmos de búsqueda sofisticados sobre la red de transporte en sus diferentes modos. Asimismo, algunas de estas aplicaciones incorporan predicción de tráfico ofreciendo mejores estimaciones de acuerdo con el día y horario de la consulta. En muchos casos, estas aplicaciones ofrecen interfaces (API) para realizar consultas en forma masiva. Si bien el mercado de este tipo de aplicaciones está evolucionando en forma acelerada y surgen nuevas aplicaciones todo el tiempo, para la elaboración de este estudio se analizaron las siguientes:

Google Maps, Directions API: permite realizar consultas de viajes desde puntos geográficos en la ciudad en distintos modos de transporte, incluyendo viajes en automóvil, en transporte público y a pie. En todos los casos devuelve las diferentes alternativas de viaje en cada modo y se puede obtener un estimado de tiempo de viaje, distancia, costo y un desagregado de las etapas del viaje. Google incorpora predicción de tráfico. El uso de la aplicación requiere un registro de usuario y tiene un costo por cada consulta, con un monto de consultas gratis mensuales.

<https://developers.google.com/Maps/documentation/directions/usage-and-billing>

Here (<https://wego.here.com/>): Here.com ofrece servicios similares a los de Google Maps. Permite realizar consultas de viajes entre dos puntos en la ciudad con predicción de tráfico. Al igual que la API de Google Maps, Here requiere una suscripción y tiene un costo, con un límite de consultas gratis en forma mensual.

<https://developer.here.com/>

OpenStreetMaps (OSM): OSM es un proyecto colaborativo para crear mapas editables y libres, y permite consultas en forma gratuita. Si bien la aplicación ofrece un servicio de búsqueda punto a punto en la ciudad, tiene la desventaja de no ofrecer la búsqueda en transporte público y no tiene predicción de tráfico.

<https://www.openstreetmap.org/>

La simulación de viajes se puede obtener a través de la librería pandana, que resulta muy eficiente para la realización de consultas de viajes en gran escala, o bien puede usarse la librería osmnx, que permite obtener los establecimientos para cada ciudad así como las distancias teniendo en cuenta la red vial.

<https://udst.github.io/pandana/>

Para este modelo, las simulaciones de viajes se realizan a través de las aplicaciones de OSM y Google Maps. Estas aplicaciones permiten realizar consultas de viajes en forma masiva a través de sus API. Por un lado, se hace uso de la simulación utilizando la API de OSM, ya que permite calcular distancias entre los orígenes y destinos asignados sobre la red vial de forma gratuita. Por el otro, se realizan las simulaciones en Google Maps, que tiene un costo por consulta, y que devuelve distancias y tiempos de viaje en los distintos modos y etapas del viaje, con la distinción de que estos resultados incluyen las predicciones de tráfico, y ofrecen de esta forma mayor precisión en la estimación de tiempos según el día y el horario de consulta.

Luego de la simulación, se obtiene una serie de variables que describen la ruta seleccionada por el algoritmo de Google Maps como la más eficiente, dadas las condiciones impuestas. Las variables de salida son los tiempos de viaje en transporte público y en automóvil, información sobre distancias y tiempos de acceso a las paradas, tiempos estimados para cada etapa del viaje (como el tiempo de espera del bus), información de las transferencias e incluso el precio del viaje en transporte público para algunas ciudades. Es importante considerar que las simulaciones en automóvil no incluyen el tiempo destinado a buscar estacionamiento, que debe sumarse al tiempo total para el análisis, ya que los tiempos de acceso están incluidos en el transporte público. Este punto puede ser despreciable en zonas o ciudades con poca congestión o una elevada oferta de estacionamiento sobre la vía pública. Sin embargo, en áreas de alta congestión y gran uso del vehículo privado, el tiempo para buscar aparcamiento puede llegar a ser muy significativo.

A continuación se enumeran las piezas de información que devuelve la aplicación Google Maps para cada viaje simulado en transporte público y en automóvil.

Transporte público:

- 01 → Tiempo total del viaje
- 02 → Modo de transporte sugerido en cada etapa del viaje
- 03 → Tiempo de viaje en cada etapa
- 04 → Distancia a la parada en cada etapa
- 05 → Tiempo estimado de espera

Automóvil:

- 01 → Tiempo estimado con predicción de tráfico
- 02 → Distancia total del viaje

Para explorar más en detalle lo que devuelve la consulta, ver <https://developers.google.com/maps/documentation/directions/get-directions#DirectionsResponse>

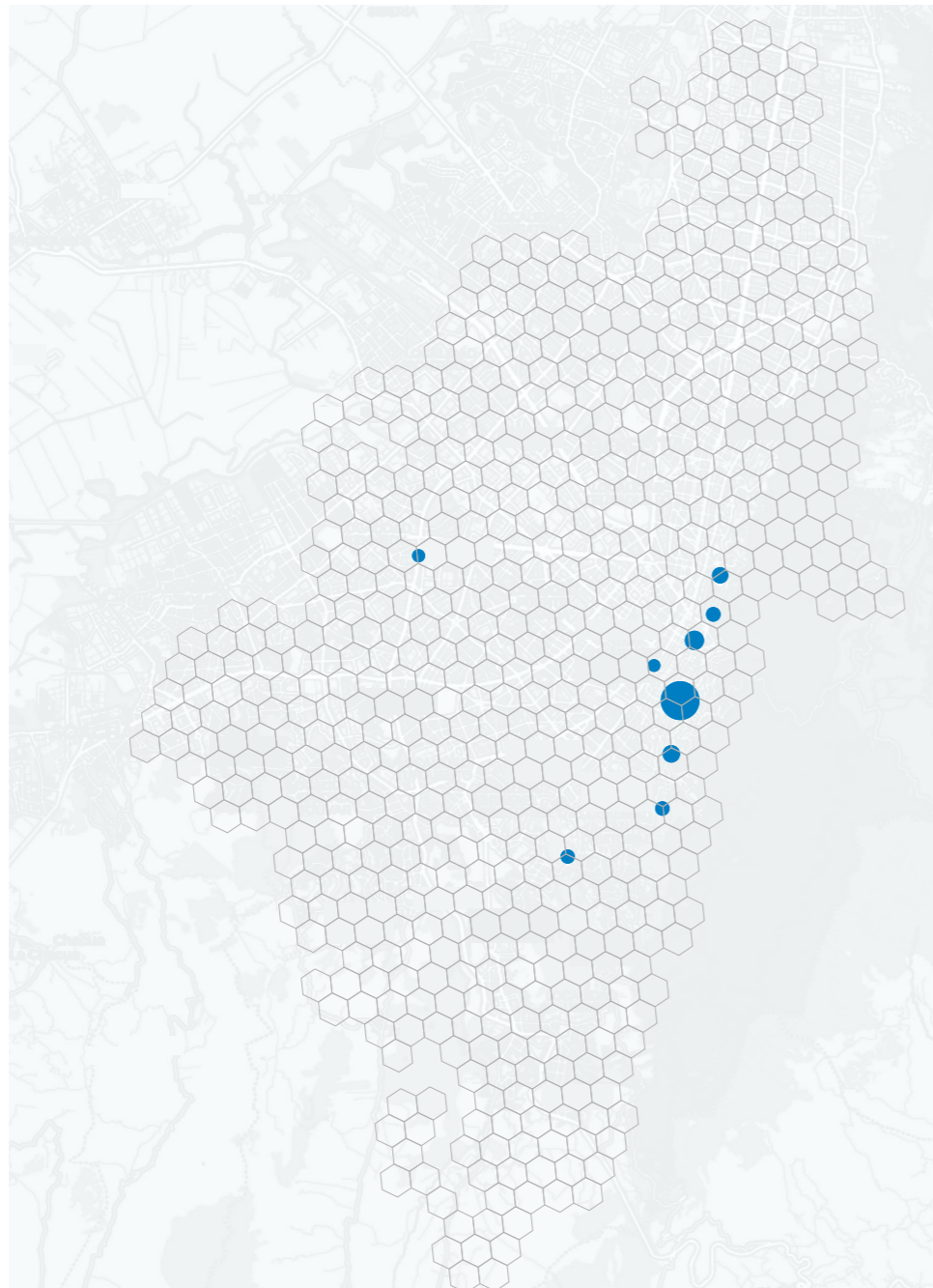
Orígenes y destinos

3.3.3

Los viajes que se simulan se conforman con orígenes y destinos seleccionados para cada ciudad. Los puntos de origen se seleccionan de tal manera que representen todo el territorio del aglomerado en cuestión. Esos puntos están representados por el centroide de cada uno de los hexágonos. Para la selección de destinos, se toman los puntos seleccionados en el análisis de demanda, donde se tienen en consideración la densidad de actividades calculada con los equipamientos obtenidos de OSM.

En una segunda etapa del análisis, se realizan simulaciones a establecimientos específicos, como los educativos, de salud, y parques o espacios públicos. El siguiente mapa muestra los datos utilizados para la simulación de viajes en el caso del área metropolitana de Bogotá. Se pueden observar los hexágonos, cuyos centroides representan los orígenes de los viajes; los destinos son identificados con el algoritmo de densidad de actividad, donde el tamaño del círculo representa la atracción relativa de cada centro de actividad. En el caso de Bogotá, se pueden observar 822 hexágonos con sus centroides y 16 puntos de destino, que son las áreas identificadas como de alta densidad de actividad o mayor atracción de via-

jes. Con esta información, se construyen 13.152 pares de orígenes-destino, y sus viajes se simulan en OSM para el cálculo de distancias y en Google Maps para obtener estimaciones de tiempos a los establecimientos más cercanos.



PUNTOS DE ORIGEN Y DESTINO DE VIAJES PARA LA SIMULACIÓN DE VIAJES EN EL ÁREA METROPOLITANA DE BOGOTÁ

M.
03

Fuente: Elaboración propia en base a datos de OSM y hexágonos de la librería H3.

Ejecución de la simulación

3.3.4

Para obtener las piezas de información necesarias para la construcción de indicadores y visualizaciones relevantes sobre la accesibilidad territorial, se realizan esta serie de simulaciones a través de las API de las aplicaciones de ruteo mencionadas anteriormente. Para el cálculo de distancias en la red, se realizan consultas en OSM. Para el cálculo de los tiempos de viaje en transporte público y en automóvil con predicción de tráfico, la consulta se realiza en Google Maps. Los tiempos de viaje a pie se estiman de acuerdo a la distancia de cada viaje sobre la red.

Los pasos para la ejecución de las simulaciones son los siguientes:

- 01 → Se construye una matriz de origen y destino de viajes entre todos los orígenes (centroides de los hexágonos que representan todo el territorio) y todos los destinos (áreas identificadas como de alta densidad de actividad).
- 02 → Con el propósito de identificar horas punta y valle y calcular velocidades promedio a lo largo de un día completo, se simulan viajes en automóvil a cada hora de un día hábil de una muestra de alrededor de 300 viajes de la matriz de orígenes y destinos.
- 03 → Para la hora punta de la mañana, identificada en el proceso anterior, se simulan los viajes en automóvil y en transporte público para la matriz completa de orígenes (centroides de hexágonos) y destinos (centros de alta densidad de actividad) de un día hábil típico (generalmente un miércoles o jueves) y en el horario punta de la mañana.
- 04 → Para el cálculo de tiempos de viajes a servicios específicos como salud y educación y a espacios verdes, primero se realizan simulaciones de las distancias en OSM y luego se calculan los tiempos de viaje en Google Maps sólo a aquellos establecimientos que se encuentren más cercanos.

El resultado de estas consultas constituye la base de datos que se utiliza para la construcción de indicadores.

El mapa 4 describe los orígenes y destinos de una simulación a establecimientos específicos en la RMBA. Los orígenes corresponden a los centroides de todos los hexágonos que cubren la RMBA, y los destinos son las ubicaciones de los hospitales.

Indicadores de accesibilidad

3.4 Cálculo de indicadores de accesibilidad	3.4.1 Índice de congestión
	3.4.2 Cobertura de Transporte Público
	3.4.3 Accesibilidad a actividades y a servicios básicos
	3.4.4 Clústeres de Accesibilidad
	3.4.5 Desarrollo de mapas y gráficos

Una vez construida la base de datos con la simulación de los viajes para cada ciudad, se procede a la construcción de una serie de indicadores de accesibilidad urbana.

Índice de congestión

A partir de los datos obtenidos de la simulación de la muestra de viajes en un día completo, se calculan los tiempos y velocidades promedio a lo largo de un día completo. Con la identificación de las velocidades en horas punta y valle, se calcula el índice de congestión.

Accesibilidad a centralidades y establecimientos

Con la simulación de los viajes de la matriz de orígenes y destinos para la hora punta de la mañana, se calculan los tiempos promedios a los principales destinos para construir el indicador de accesibilidad a actividades y los tiempos de viaje al área central.

La simulación de viajes hacia los establecimientos seleccionados (salud, educación, espacios verdes o públicos) permite identificar el más cercano para cada tipo de establecimiento y el tiempo de viaje hasta estos en transporte público o caminando. A partir de esto, se calcula la cantidad de establecimientos de

La librería `pyomu` permite realizar la simulación de los viajes para obtener tiempos y distancias, y el cálculo de indicadores.

La siguiente función permite calcular una serie de indicadores, como el índice de congestión y la identificación de horas punta y valle, entre otros, en una muestra de viajes para un día completo. La función requiere principalmente de la capa que identifica los orígenes y la capa de destinos. Para el cálculo de tiempos de viaje en Google Maps, se requiere como parámetro la clave de la API de Google Maps, que tiene que ser generada por el usuario.

`pyomu.calculate_od_matrix_all_day()`

Otra serie de funciones permite calcular tiempos y distancias de viaje para diferentes tipos de atractores:

`pyomu.access.measure_distances_osm()`

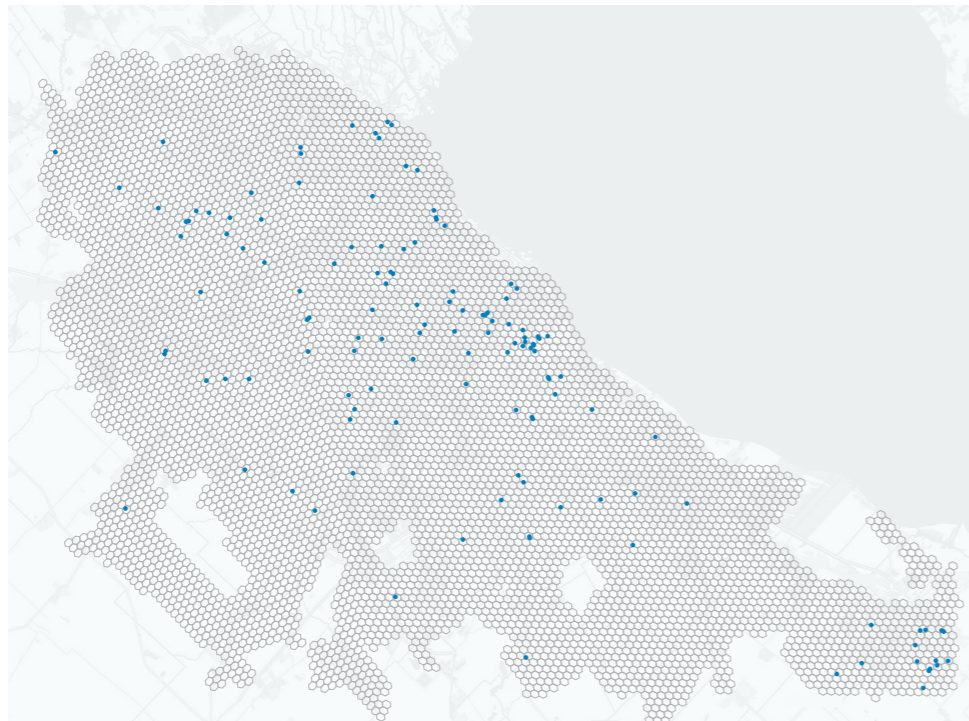
`pyomu.access.trips_qmaps_from_matrix()`

`pyomu.access.distances_to equipments()`

Existe a su vez una serie de funciones para la impresión y generación de indicadores:

`pyomu.vizuals.print_time_distance()`

Estas funciones requieren de la capa de orígenes y de destinos y otros parámetros especificados en la documentación de la librería.



RED DE HEXÁGONOS Y UBICACIÓN DE HOSPITALES EN LA REGIÓN METROPOLITANA DE BUENOS AIRES

M.
04

Fuente:Elaboración propia en base a capa geográfica de hospitales y hexágonos de la librería H3.

cada tipo según rangos de distancia, y se construyen así los indicadores de accesibilidad a educación, salud y espacios verdes.

Con la distancia promedio a las paradas de transporte público, se calcula el indicador de cobertura. Este indicador identifica la distancia promedio para acceder al transporte público en una ciudad.

Accesibilidad general

Para cada punto de origen, en este caso cada hexágono, se desarrolla el indicador de accesibilidad (clústeres de accesibilidad), que condensa la información brindada por una selección de los indicadores mencionados. El indicador de accesibilidad se presenta en una escala de cuatro niveles.

Como parte del proceso, a cada viaje se le asigna el NSE correspondiente según el hexágono donde se origina el viaje, lo que permite la desagregación de los indicadores por NSE. A continuación, se brinda mayor detalle de cómo se construyen los indicadores en este modelo.

Índice de congestión

3.4.1

Con el objetivo de visualizar la variabilidad de tiempos de viaje y velocidades que existen durante el día en una ciudad debido al tráfico, se promedian los tiempos de viaje obtenidos de la consulta a través de la muestra simulada en cada hora del día. Se utiliza el tiempo de viaje con predicción de tráfico, lo que permite identificar tiempos promedios a lo largo de un día completo tanto para automóvil

como para el transporte público. La velocidad de cada viaje se calcula utilizando la distancia sobre el tiempo de viaje en automóvil con la predicción de tráfico y se promedia para cada hora del día. Luego, se utilizan las velocidades promedio en hora valle y punta para calcular un índice de congestión. El índice se calcula como el ratio entre la velocidad en hora valle y la velocidad en hora punta.

Fórmula 1 Índice de congestión

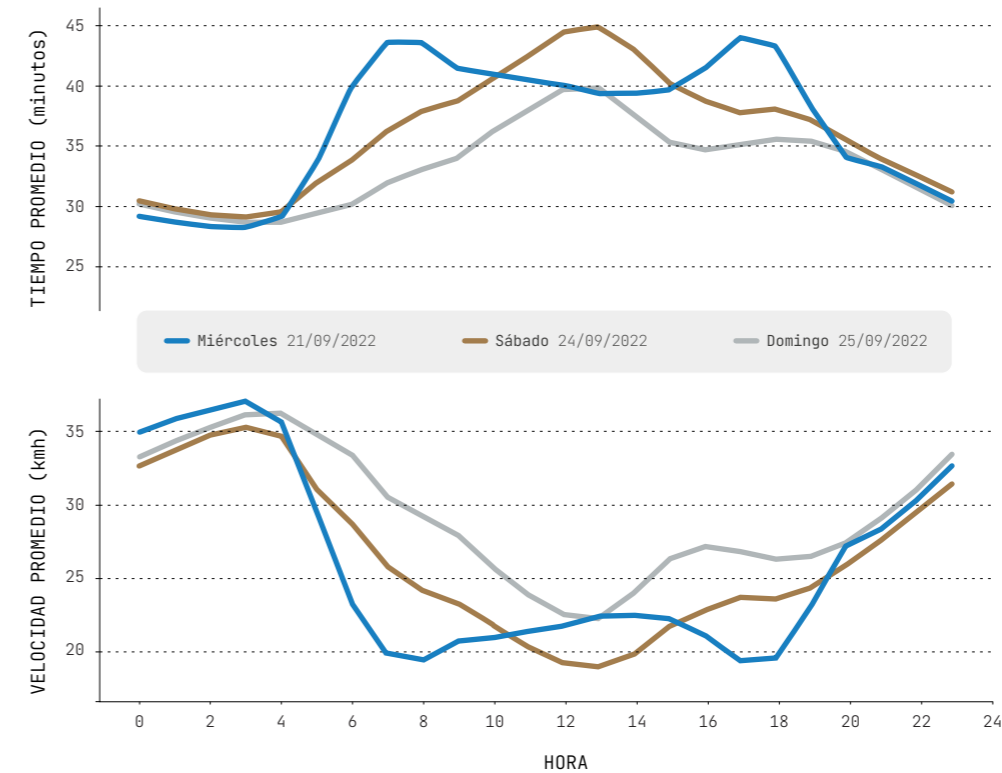
$$IC = V_{hv} / V_{hp}$$

IC: índice de congestión

V_{hv}: velocidad identificada como centro de la hora valle

V_{hp}: velocidad identificada como centro de la hora pico

En la figura 3 se observan los perfiles de tiempos y velocidades de viaje obtenidos de un día típico en automóvil en el área metropolitana de Bogotá, en donde se pueden identificar los horarios pico y valle de tráfico, lo que da como resultado un índice de congestión de 1,9 en día hábil.



TIEMPOS Y VELOCIDADES PROMEDIO DE VIAJES EN UN DÍA HÁBIL EN BOGOTÁ, COLOMBIA

F. 03

Fuente: Elaboración propia.

Accesibilidad a principales centralidades

3.4.2

La accesibilidad a las principales centralidades se calcula a través de la accesibilidad a centros de actividad ubicados en distintos lugares de la ciudad y la accesibilidad al área central.

La accesibilidad a actividades se analiza a través del tiempo promedio de viaje a centros de actividad en tres modos: transporte público, en automóvil y a pie, en un día

hábil típico a la hora punta de la mañana. Para cada punto de origen se promedian los tiempos de viajes hacia los puntos identificados como de alta densidad de actividad. Dado que cada destino tiene un nivel de atracción diferente, para el cálculo del promedio se pondera utilizando este factor. Esta ponderación le asigna pesos a los distintos destinos, y como resultado, un viaje hacia el área central tendrá una ponderación mayor que un viaje a una subcentralidad de menor importancia y con menor densidad de actividad.

Fórmula 2 Cálculo de tiempos promedio ponderados

$$TV_{i,m} = \sum_d TV_{i,m,d} \times p_d / P_d$$

TV: tiempo de viaje

i: hexágono i

m: modo (automóvil, transporte público, a pie)

d: destinos identificados por el método de densidad de actividades

pd: ponderador de cada destino en función de la densidad de sus actividades

Pd: suma de las ponderaciones de todos los destinos

La fórmula 2 expresa cómo calcular el tiempo de viaje promedio (TV) para un hexágono (i), en un modo (m), que puede ser automóvil, transporte público o caminata.

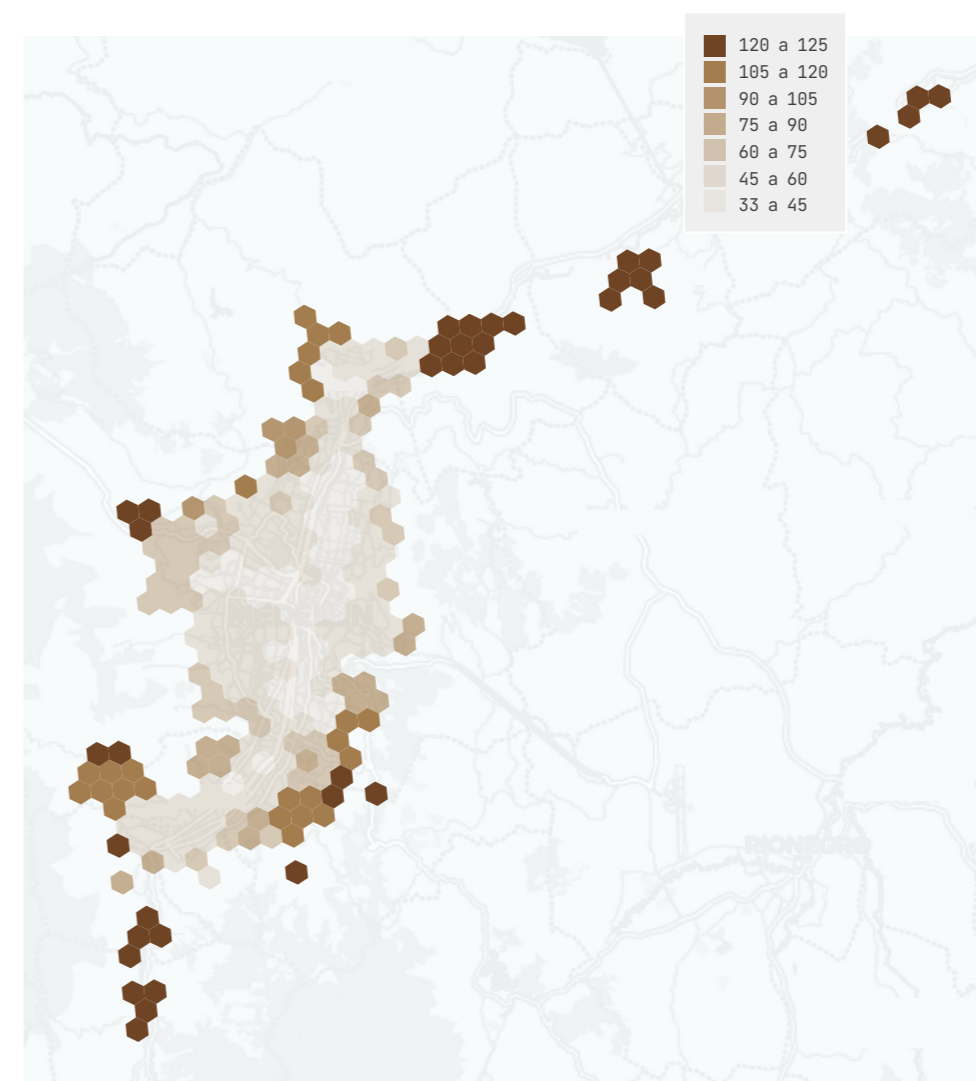
En el mapa 5 se puede observar el cálculo de los tiempos de viaje a los destinos atractores del área metropolitana del Valle Aburrá. Se muestran los hexágonos en una escala de colores que representan el tiempo de viaje promedio de cada unidad.

En las ciudades latinoamericanas, es usual que el área central concentre una gran variedad de actividades relevantes para la ciudadanía, como negocios comerciales, oficinas, edificios públicos, o establecimientos de cultos. Por eso, es común visualizar mapas con isócronas calculadas con el destino al centro. Este modelo incluye un indicador específico para los viajes al centro de la ciudad. ^[11]

Si bien es común en estudios similares medir un indicador de tiempos de viajes al área central, los autores consideran que es más interesante el indicador de accesibilidad a actividades, que mide tiempos promedio a los destinos de alta densidad de actividad, ya que en las grandes ciudades suele haber más de una centralidad principal, lo cual este indicador tiene en consideración.

El tiempo de viaje al centro es el tiempo de viaje en transporte público, en automóvil y a pie desde cada origen hacia el destino identificado con el área central: el punto con mayor atracción de actividad, resultado del análisis de la densidad de actividad.

En el mapa 6 se muestran los tiempos de viaje al centro en la ciudad de México DF. En el mapa de la izquierda, se ven los tiempos de viaje al centro en transporte público, y en el mapa de la derecha, en automóvil.



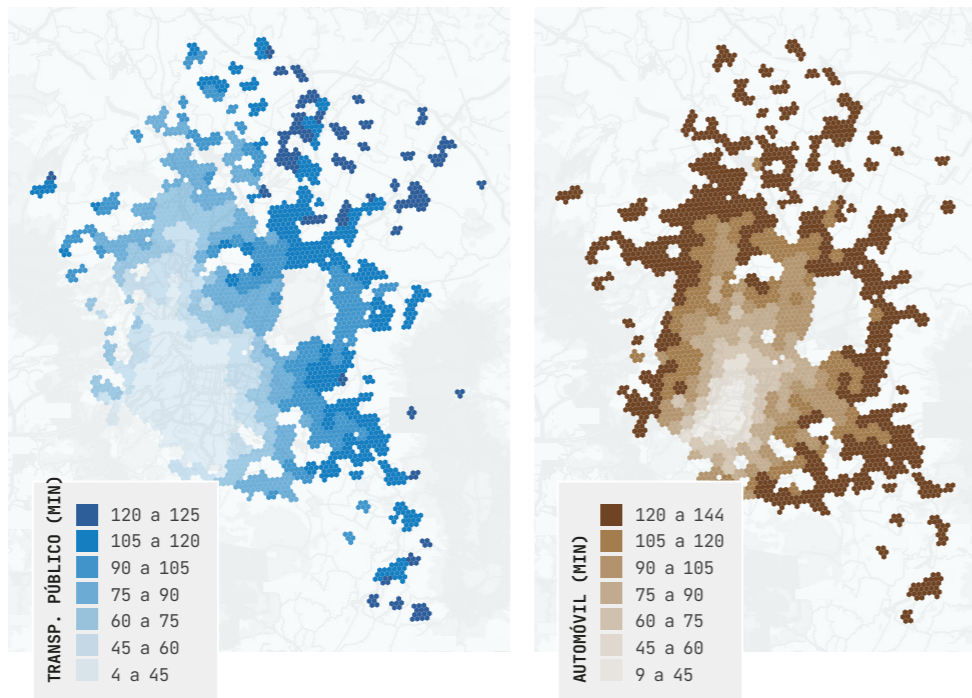
ACCESIBILIDAD A ACTIVIDADES: TIEMPOS DE VIAJE PROMEDIO EN TRANSPORTE PÚBLICO A CENTROS DE ACTIVIDAD EN EL ÁREA METROPOLITANA DEL VALLE ABURRÁ

M.
05

Fuente: Elaboración propia.

TIEMPOS DE VIAJE AL CENTRO DE MÉXICO DF. **M. 06**

Fuente: Elaboración propia.



Cobertura del sistema público de transporte

3.4.3

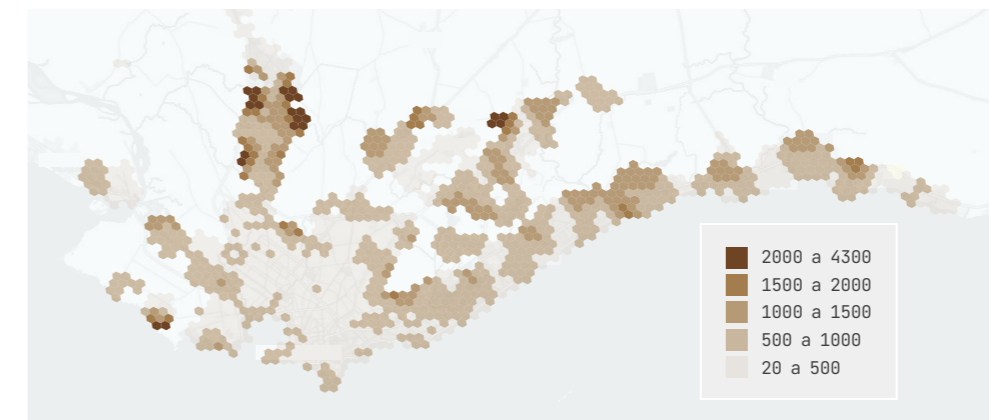
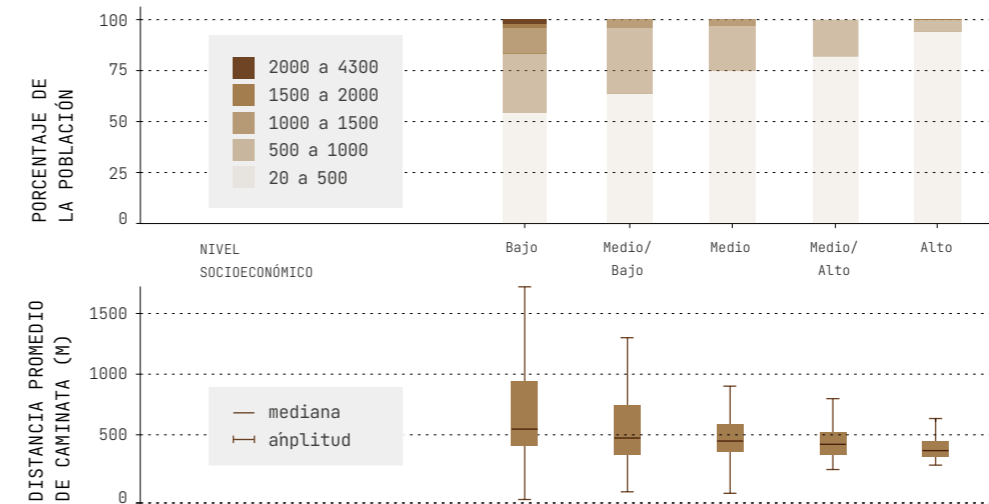
Como indicador de cobertura del transporte público, se calcula la distancia promedio de acceso caminando al transporte público. Utilizando los resultados de la serie de simulaciones de viajes a los principales centros de actividad identificados, para este indicador se promedia la distancia de acceso al transporte público en origen de estos viajes simulados. Se utiliza el promedio ponderado, donde el tiempo de viaje a cada destino tiene asignado un peso o ponderador construido a partir del análisis de la densidad de actividad.

Además del valor de la distancia promedio al transporte público que se le asigna a cada hexágono, se calcula el porcentaje de la población que habita en los hexágonos con una distancia promedio al transporte público menor a 1000 m. De esta manera, se puede conocer, a nivel de la ciudad, cuál es el porcentaje de personas que tiene acceso al transporte público en menos de 1000 m. Finalmente, dicho porcentaje puede ser desagregado en NSE, lo que permite brindar a las personas tomadoras de decisiones información del acceso al transporte público por estrato socioeconómico.

El siguiente mapa muestra las distancias promedio de caminata a partir de las simulaciones en transporte público para los hexágonos del área metropolitana de Montevideo. Sobre el mapa, dos gráficos de barras muestran la agregación de estos valores por NSE, en donde se puede observar la distancia promedio por NSE y los porcentajes de población por NSE según rango de distancia. [12]

[12]

Debido a que esta variable es muy sensible a la distancia, se realizó un ajuste en la variable (*spatial lag*) que promedia los valores vecinos de una ubicación con el objetivo de suavizar los resultados. Este ajuste lo realiza la librería *pyomu* y cuenta con una serie de opciones para su parametrización.



Fuente: Elaboración propia.

Accesibilidad a salud y educación

3.4.4

La accesibilidad a los servicios básicos de salud y educación se analiza mediante los tiempos de viaje a estos establecimientos y la cantidad a los que se puede acceder en distintos rangos de distancia. Para la construcción de los mapas de tiempos de viaje, se identifica para cada origen el establecimiento más cercano y se toma el

tiempo de viaje en cada modo. Este paso se repite para cada tipo de establecimiento, por ejemplo, en el caso de establecimientos educativos, se identifica el más cercano para escuelas iniciales públicas y privadas, primarias públicas y privadas, y secundarias públicas y privadas. Esto va a depender en gran parte de la disponibilidad de información. Además, se identifica la cantidad de establecimientos de cada tipo que se encuentran dentro de una determinada distancia, que se calcula teniendo en cuenta la red vial de OSM.

En este modelo y específicamente en el proyecto enmarcado en el OMU, se construyeron los indicadores de tiempos de viaje y cantidad de oferta según distancia a los siguientes tipos de establecimientos, y se calcularon para las ciudades del OMU según la disponibilidad de información:

- 01 → Educación pública: inicial, primaria, secundaria
- 02 → Educación privada: inicial, primaria, secundaria
- 03 → Salud: atención primaria, hospitales

Accesibilidad a espacios públicos o áreas verdes

3.4.5

El cálculo de áreas verdes per cápita en un rango de distancia se puede hacer a través de la librería `pyomu`. La función requiere de la capa de hexágonos, los rangos de distancia a calcular y una capa de los polígonos identificados como espacios verdes o públicos. En caso de no ser enviada como parámetro, esta última capa se obtiene automáticamente a través de OSM.

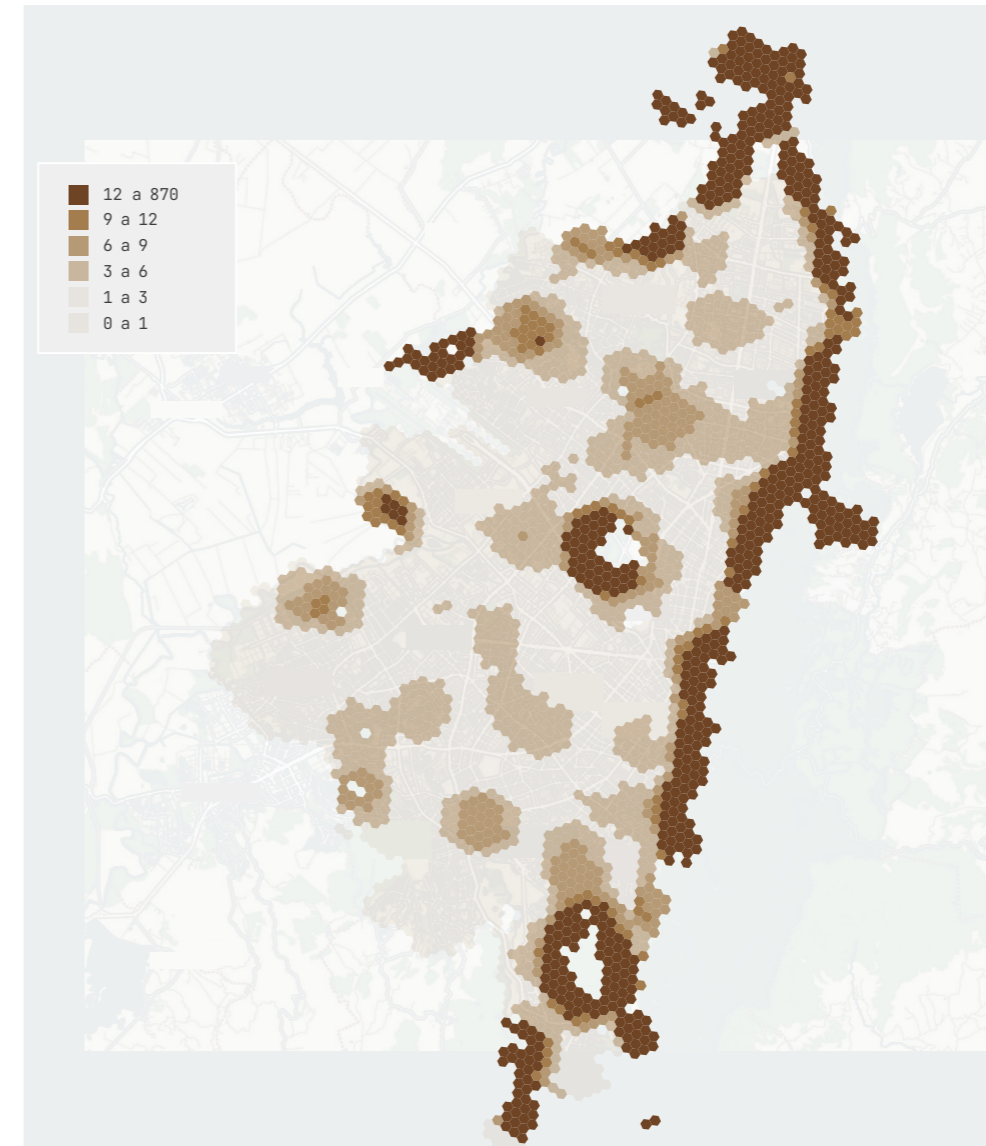
`pyomu.access.calculate_green_space()`

Un indicador de estas características permite identificar no solo las áreas verdes disponibles por habitante, sino que tiene en consideración la cercanía y la densidad de población, ya que la oferta de áreas verdes y espacio público se divide por las personas que habitan en un rango de distancia. Las áreas públicas o verdes se obtienen de OSM según lo descrito en la sección 3.2.3. A continuación se especifica con mayor detalle cómo se construye este indicador.

[13]

La superficie de espacios verdes y públicos para cada hexágono se logra de la superposición de la capa de hexágonos con la capa obtenida en OSM.

Para el cálculo de accesibilidad a espacios públicos o áreas verdes, se construye un indicador de espacios públicos o verdes disponibles en un rango de distancia y de cantidad de población que puede hacer uso del espacio público o verde en las distancias establecidas.



ESPACIOS PÚBLICOS O ÁREAS VERDES PER CÁPITA (m²) EN UN RANGO DE 1200m. BOGOTÁ, COLOMBIA. **M. 08**

Fuente: Elaboración propia.

Primero se calcula la oferta de áreas verdes y espacios públicos per cápita que ofrece cada hexágono j (OAV_j) dividiendo la superficie de espacios verdes y públicos de cada hexágono^[13] por la población en un rango de distancia ($d1$), por ejemplo, 1000 m. Luego, para conocer el área de espacios verdes y públicos per cápita al que puede acceder una persona en el hexágono i ($AVpCi$), se realiza la sumatoria de la oferta OAV_j de los hexágonos j que están a una distancia (d) del hexágono i , por ejemplo, 1200 m.

La fórmula 3 expresa el cálculo del acceso a áreas verdes y espacios públicos.

Fórmula 3 Cálculo de área verde/espacio público per cápita

PASO 1: Oferta por hexágono considerando la densidad poblacional

$$OAV_j = AV_j / P_{d1}$$

AV_j : área verde o espacio público en el hexágono j
 P_{d1} : Población total en un rango de distancia $d1$
 $d1$: distancia en la que se calcula la oferta de los espacios verdes o públicos
 OAV_j : Oferta de espacios verdes y públicos per cápita del hexágono j

PASO 2: Acceso a AV per cápita en cada hexágono i

$$AVpC_i = \sum_{j \in d} OAV_j$$

d : distancia en la que se suma la oferta de áreas verdes o espacios públicos
 $AVpCi$: superficie de áreas verdes y espacios públicos per cápita

Indicadores de accesibilidad por NSE

3.4.6

Uno de los objetivos de este modelo es construir herramientas para estudiar las desigualdades que se presentan en las oportunidades y los costos que perciben las personas con distintos NSE. Se utilizan los NSE calculados para cada unidad censal o hexágono derivados del método mencionado en la sección 3.2.1. y detallado en el anexo.

Fórmula 4. Cálculo de tiempos de viaje por NSE

$$TV_{nse,m} = \sum_i TV_{i,m} \times h_i / H_{T,nse} \quad \forall nse_i \in nse$$

TV: tiempo de viaje

nse: nivel socioeconómico (bajo, medio-bajo, medio, medio-alto y alto)

m: modo (automóvil, transporte público, a pie)

i: hexágono i. Sumatoria en todos los hexágonos.

TV_{i,m}: tiempo de viaje promedio en el modo m del hexágono i. (ver fórmula 1)

h_i: población del hexágono i

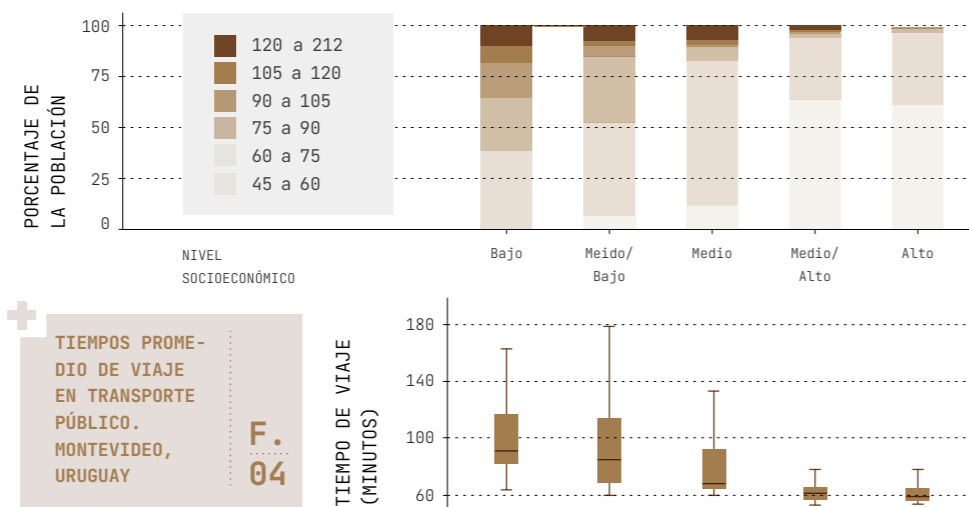
H_{T,nse}: población total correspondiente al nivel socioeconómico (nse).

nse_i: nivel socioeconómico asociado al hexágono i

Para la construcción de los indicadores de accesibilidad y de cobertura, se promedian los tiempos de viaje en transporte público, en automóvil y a pie para cada NSE, y se obtienen indicadores a nivel de la ciudad o área metropolitana desagregados. Se realiza un promedio ponderado para cada NSE teniendo en consideración la población asignada a cada hexágono. De esta forma, los viajes desde hexágonos que se encuentran en áreas con más población tienen un mayor peso.

La fórmula 4 expresa cómo calcular la accesibilidad por NSE, como el tiempo de viaje promedio (TV) de la ciudad o área metropolitana para un nivel

socioeconómico (NSE), en un modo (m), que puede ser automóvil, transporte público o caminata.



Fuente: Elaboración propia.

La figura 4 muestra los tiempos promedio de viaje según NSE. El gráfico de arriba muestra qué porcentaje de la población se encuentra en los distintos rangos de tiempos promedio de acceso para cada NSE; el gráfico de abajo, tiempos promedio según NSE.

Niveles de accesibilidad general

3.4.7

Para el cálculo de los clústeres de accesibilidad se identifican una serie de variables relevantes para evaluar la accesibilidad de los hogares. Con ellas, se realiza un proceso de clusterización (k-means) para clasificar los hexágonos en cuatro clases. El método de clusterización k-means es una técnica de aprendizaje automático no supervisado que permite la partición de una serie de observaciones en grupos relativamente homogéneos.

Cada uno de estos clústeres representa distintos niveles de accesibilidad según una serie de indicadores: distancia de acceso caminando al transporte público, distancia a los establecimientos principales (educación primaria y atención primaria de salud), cantidad de estos establecimientos en 2 km y tiempos de viaje al área central.

Los pasos para la generación de los clústeres de accesibilidad general son los siguientes:

01 → Selección de variables: se seleccionan las variables relacionadas con educación (inicial y primaria) y salud (atención primaria y hospitales). Se incluyen las variables de cantidad de establecimientos en un rango de 2 km y tiempos de viaje en colectivo al establecimiento más cercano. Se incluye también el tiempo de viaje al área central.

02 → Estandarización: se estandarizan las variables seleccionadas.

03 → K-means: se realiza el proceso de clusterización con k-means, especificando cuatro clústeres.

Virtudes, limitaciones y mejoras

Entre las virtudes del modelo se pueden mencionar la elaboración de una serie de mapas e indicadores sencillos y replicables que brindan información para tomar decisiones sobre movilidad y transporte. Los indicadores y mapas incluyen información sociodemográfica y tiempos de viaje, con diferentes alternativas de transporte y condiciones de tráfico en los días y horarios analizados, para buscar de este modo replicar lo más posible las condiciones reales de movilidad en cada ciudad.

En comparación con otros modelos que requieren una gran cantidad de información, con relevamientos que no siempre están disponibles y el uso de software más complejo, esta herramienta permite analizar a través de distintos mapas e indicadores los niveles de accesibilidad en una ciudad en un tiempo corto, con muy bajo costo ^[14] y sin mayor complejidad.

Hay algunas consideraciones a tener en cuenta sobre el modelo analítico presentado, sobre todo en relación a la disponibilidad y calidad de los datos de cada ciudad. En primer lugar, los datos sociodemográficos utilizados provienen de los últimos censos nacionales disponibles en cada ciudad, los cuales pueden encontrarse desactualizados en algunos casos. Al momento de realizar este documento, en algunos países esa información tiene hasta 10 años de antigüedad. Sin embargo, el modelo es un proceso automatizado y construido sobre código abierto, por lo que la actualización de dicha información es simple. En segundo lugar, la elaboración de los indicadores de transporte público mediante la simulación de viajes en Google Maps requiere que la plataforma tenga disponibles y actualizadas las rutas y frecuencias del transporte público de cada ciudad (los archivos GTFS). Se debe considerar que aun utilizando esta información, se trata de horarios programados sin considerar retrasos debido al tráfico vehicular, lo cual afecta a los sistemas que operan en tráfico mixto. La existencia de archivos GTFS en Google Maps es un desafío importante en las ciudades de la región, por la estandarización del calendario del servicio de transporte público, pero aún más en ciudades con servicios de transporte público informales o semiformales. Por último, para utilizar esta metodología se necesita que estén disponibles las capas georreferenciadas de establecimientos educativos y de salud y, para el cálculo de densidad de actividades, es necesario que los distintos equipamientos estén disponibles y actualizados en OSM. ^[15]

[14]

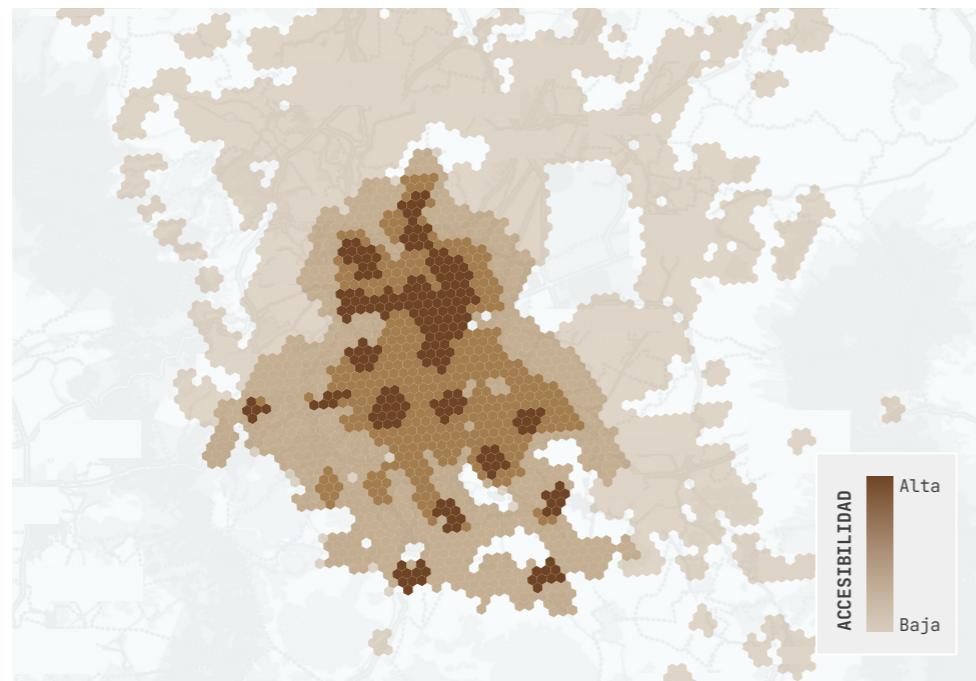
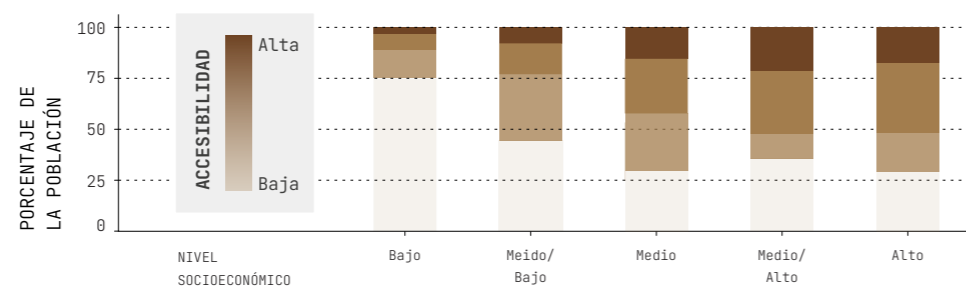
El costo por consulta en transporte público es USD 0,05 y en automóvil con predicción de tráfico es USD 0,1. El costo para las simulaciones de este modelo puede ir desde los USD 400 en una ciudad como Medellín, Colombia, a los USD 3500 para una ciudad con un área metropolitana extensa como Buenos Aires.

[15]

Existe información relativamente buena en OSM para las principales ciudades. Generalmente, la calidad y actualización es mejor para las áreas centrales y se encuentra más desactualizada en las áreas periféricas de las ciudades. La gran evolución que se está produciendo en el área de datos compartidos en los últimos años puede generar datos abiertos de mayor calidad en el corto plazo.

El mapa 9 muestra un ejemplo del resultado de los clústeres de accesibilidad identificados para el área metropolitana del Valle de México, en donde se observan las zonas de baja, media, media-alta y alta accesibilidad general.

Además, en base a los clústeres identificados y su nivel de accesibilidad, la siguiente figura muestra para cada NSE el porcentaje de la población según su nivel de accesibilidad.



CLÚSTERES DE ACCESIBILIDAD POR NSE. CIUDAD DE MÉXICO. F. 05

Fuente: Elaboración propia.

CLÚSTERES DE ACCESIBILIDAD CIUDAD DE MÉXICO. M. 09

Fuente: Elaboración propia.

Más allá de estas consideraciones, que pueden limitar el alcance de este modelo analítico, se considera que la disponibilidad y calidad de los datos está en un proceso de constante mejora. Tanto los datos de transporte público en plataformas como Google Maps como los datos sobre infraestructura urbana en OSM van mejorando su disposición y calidad con el tiempo, teniendo en consideración los grandes avances en esta materia que se produjeron en los últimos diez años. La perspectiva para los próximos años es que la disposición de datos masivos de acceso público será cada vez mayor y de mejor calidad, lo que definitivamente va a redundar en que estudios como este sean cada vez más robustos.

Ventajas:

- 01 → Replicable y escalable a otras ciudades
- 02 → Basado en datos públicos
- 03 → Condiciones de tráfico y tiempos de viaje reales
- 04 → Resultados interpretables
- 05 → Bajo costo para las simulaciones
- 06 → Código abierto colaborativo
- 07 → Automatización para actualización de resultados

Limitaciones:

- 01 → Los datos socioeconómicos pueden estar desactualizados (año de cada censo)
- 02 → Los indicadores de transporte público dependen de la existencia y actualización de GTFS de las ciudades en Google Maps

Existe evidencia en otros estudios de que los patrones de movilidad son diferentes según las características socioeconómicas, como niveles de ingreso, género o edad. Este modelo podría mejorar en la medida en que se puedan diferenciar mejor distintos grupos usuarios. Los datos de tarjetas electrónicas de pago de transporte público pueden ayudar a generar estimaciones de demanda diferenciadas, aunque se estaría sólo teniendo en cuenta el transporte público.^[16] En aquellas ciudades donde existen encuestas de movilidad y modelos de transporte actualizados, podrían estimarse los atractores de viajes según características socioeconómicas de las personas y así mejorar la precisión de los destinos de los viajes. Otro tipo de encuestas y relevamientos podrían informar también sobre otras características más cualitativas sobre

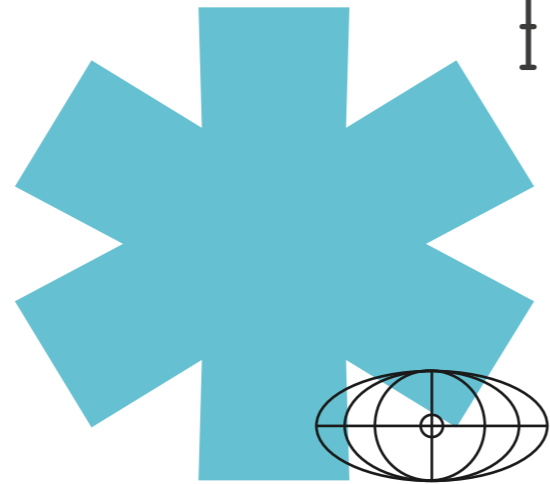
los usuarios, así como necesidades actualmente no satisfechas por el sistema de transporte disponible.

Una de las virtudes de este modelo analítico es que puede ser replicado para otros días de la semana o para distintas temporadas, por lo que se podría utilizar para analizar la movilidad durante los fines de semana o, en ciudades turísticas, podrían compararse la temporada baja con la temporada alta. Asimismo, podría realizarse este análisis a lo largo del tiempo para monitorear cambios en los niveles de accesibilidad. Por último, las funciones que existen en la librería abierta pyomu pueden utilizarse para otros fines, como estudios específicos de accesibilidad a un determinado establecimiento o para el análisis de determinados corredores de movilidad.

^[16] SUBE en Buenos Aires, Tullave en Bogotá, Tarjeta de Movilidad Integrada (MI) en México, etc.

04

REGIÓN
METROPOLITANA
DE BS. AS.



REGIÓN
METROPOLITANA
DE BS. AS.



{04}



REGIÓN
METROPOLITANA
DE BS. AS.



PARTE 04

CASO DE ESTUDIO

- Desarrollo del modelo de análisis de accesibilidad urbana para la Región Metropolitana de Buenos Aires.

P.
SECCIÓN
04

CASO DE ESTUDIO: REGIÓN METROPOLITANA DE BUENOS AIRES

4.1	Contexto	72
4.2	Datos utilizados	73
M.10	Origen y destino de viajes para la simulación de viajes en la RMBA	74
4.3	Indicadores	75
4.3.1	Indicadores de congestión	75
F.6	Tiempos y velocidades promedio de viaje	75
4.3.2	Indicadores de accesibilidad	
M.11	Tiempo promedio de viajes en transporte público. RMBA	76
F.8	Distribución de tiempos de viaje en transporte público y en automóvil	77
4.3.3	Indicadores de cobertura	
M.12	Distancia promedio de caminata en origen para viajes en transporte público	79
4.3.4	Accesibilidad a establecimientos de salud, educación y espacios verdes	80
M.13	Escuelas secundarias en 1500 m	80
F.9	Secundarios en un rango de 1500 m	80
M.14	Espacios verdes (m2) en 2000 m	81
F.10	Tiempo promedio de acceso a la atención primaria y hospitales por NSE	82
4.3.5	Clústeres de accesibilidad	
M.15	Clústeres de accesibilidad de la RMBA	82
F.11	Niveles de accesibilidad según NSE	83
		84
4.4	Análisis del caso	85

Caso de estudio: Región Metropolitana de Buenos Aires

En esta sección se presenta un caso de estudio. Se desarrolla el modelo presentado de accesibilidad territorial para la RMBA, en donde se llevan a cabo las simulaciones, se calculan los indicadores detallados anteriormente, se construyen las visualizaciones correspondientes y se analizan sus resultados.

Contexto

4.1

La RMBA está compuesta por 40 partidos de la provincia de Buenos Aires ^[17] y la Ciudad Autónoma de Buenos Aires (CABA). La RMBA se caracteriza por ser la concentración urbana de mayor densidad del país y la tercera megaciudad en América Latina, después del área metropolitana de Valle de México y la región metropolitana de San Pablo. Cuenta con una población de alrededor de 15 millones de habitantes, ^[18] lo que representa el 37% de la población total del país, y contribuye con cerca del 50% del PBI nacional. A su vez, la CABA, capital administrativa del país, se encuentra subdividida en 15 comunas, y allí viven 3 millones de personas.

La RMBA cuenta con una extensa red con diferentes modos de transporte. En primer lugar, un sistema ferroviario metropolitano, con ocho líneas que conectan la CABA con las áreas periurbanas, y que cuenta con unos 830 km de vía férrea y 272 estaciones. La CABA cuenta también con un sistema de subterráneo con seis líneas y 91 estaciones, que incluye un corredor de tren ligero de unos 52 km. Esta red es complementada por servicios de autobuses urbanos, con 342 líneas operadas por cerca de 200 empresas. El sistema de autobuses cubre unos 25.000 km, atendidos por 15.000 autobuses. Este sistema público también incluye cerca de 40.000 taxis, 6000 vehículos a pedido y 41.000 servicios de chárter (Muzzini et al., 2016). Además, tiene 285 km de ciclovías exclusivas para el uso de la bicicleta.

[17]

Partidos de la provincia de Buenos Aires: Almirante Brown, Avellaneda, Berazategui, Berisso, Brandsen, Campana, Cañuelas, Ensenada, Escobar, Esteban Echeverría, Exaltación de la Cruz, Ezeiza, Florencio Varela, General Las Heras, General Rodríguez, General San Martín, Hurlingham, Ituzzaingó, José C. Paz, La Matanza, Lanús, La Plata, Lomas de Zamora, Luján, Marcos Paz, Malvinas Argentinas, Moreno, Merlo, Morón, Pilar, Presidente Perón, Quilmes, San Fernando, San Isidro, San Miguel, San Vicente, Tigre, Tres de Febrero, Vicente López, y Zárate.

[18]

Población estimada al 1 de julio de 2021 (INDEC).

Datos utilizados

4.2

Como se mencionó en la parte 3, para construir las matrices de orígenes y destinos que se utilizan en las simulaciones, se deben obtener y procesar datos cartográficos y socioeconómicos. Para el caso de la RMBA, el modelo se basa en la cartografía del censo nacional de 2010, con el radio censal como unidad geográfica más desagregada (INDEC). ^[19] Las variables socioeconómicas del censo utilizadas para construir los NSE son las siguientes:

- 01 → Sin primario completo
- 02 → Sin secundario completo
- 03 → Sin universitario completo
- 04 → Hacinamiento
- 05 → Hacinamiento crítico
- 06 → Tiene teléfono
- 07 → Tiene celular
- 08 → Tiene computadora
- 09 → Tiene heladera
- 10 → Índice de calidad constructiva
- 11 → Índice de calidad de los materiales
- 12 → Índice de calidad de los servicios
- 13 → Tipo de vivienda

Para construir la cartografía normalizada, se utilizan los hexágonos H3 de resolución 8. En el mapa 10 se pueden visualizar los 6423 hexágonos utilizados para representar la RMBA (y los puntos de orígenes de los viajes), así como los 18 centros de actividad considerados como destinos atractores de viajes de la RMBA. Se observa que los principales centros de actividad se encuentran en el área central de la ciudad. Asimismo, se pueden ver otras subcentralidades, tanto en la ciudad como en el área metropolitana.

Para los establecimientos específicos de salud y educación, se obtiene la información de las páginas de datos públicos generados, guardados y publicados por el Gobierno de la Ciudad y de la provincia de Buenos Aires. ^[20]

[19]

Página web:
<https://www.indec.gov.ar/>

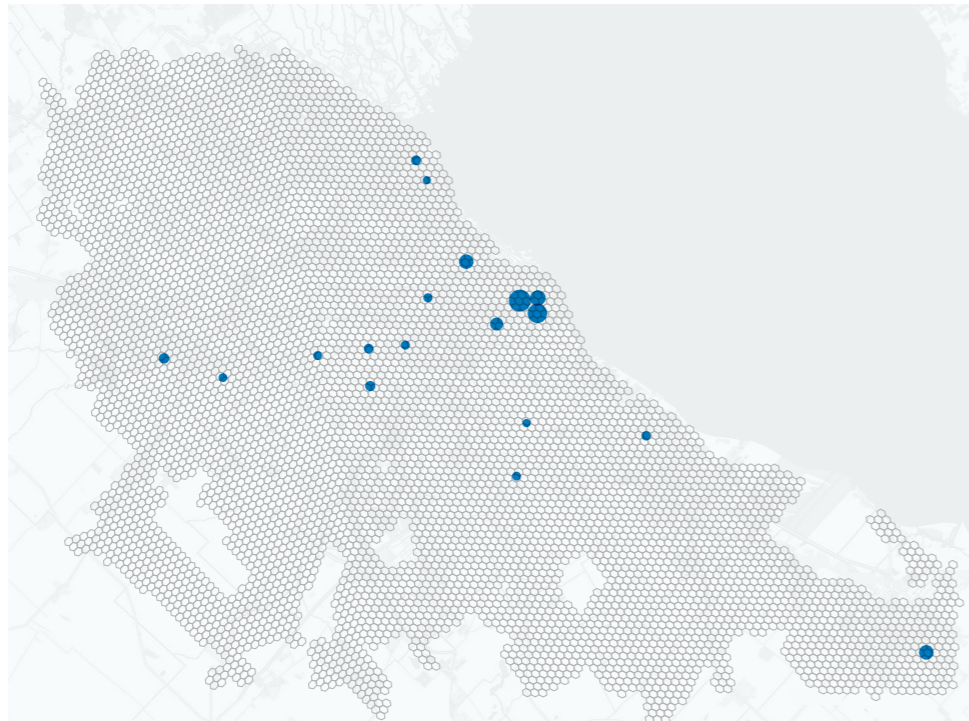
[20]

Escuelas y centros de salud desde <https://data.buenosaires.gob.ar/> y <https://catalogo.datos.gba.gov.ar/>.

A partir de estos datos, se realizan las simulaciones correspondientes. A continuación se presentan y discuten los resultados obtenidos.

PUNTOS DE ORIGEN Y DESTINO DE VIAJES PARA LA SIMULACIÓN DE VIAJES EN LA REGIÓN METROPOLITANA DE BUENOS AIRES

M. 10



Fuente: Elaboración propia.

Indicadores

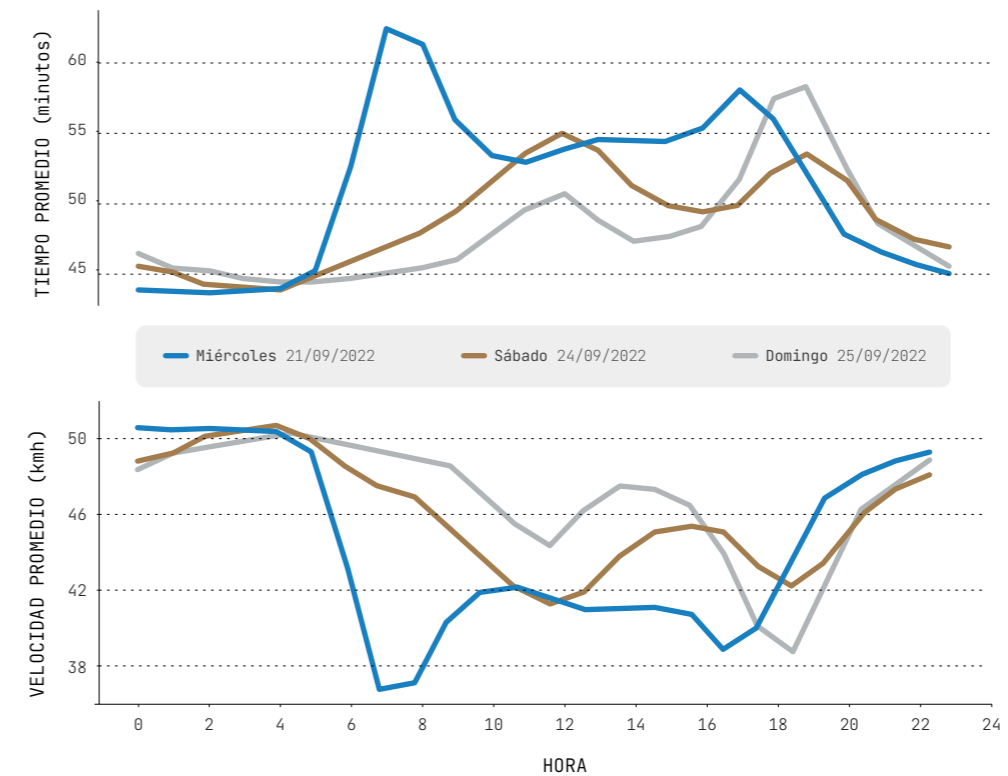
4.3

Indicadores de congestión

4.3.1

El cálculo de los tiempos de viaje es fundamental para comprender los niveles de accesibilidad y conocer en qué medida el transporte responde a las necesidades de la población. Dado que los tiempos de viaje van a ser mayores en aquellas zonas donde los niveles de accesibilidad son bajos, el análisis de tiempos permite identificar la desigualdad espacial en la ciudad.

Los tiempos de viaje varían a lo largo de un día de acuerdo con los niveles de congestión, por lo que se realiza un proceso de consulta de tiempos de viajes en diferentes horarios (figura 6). Este análisis permite identificar las horas punta y las horas valle de congestión y calcular las velocidades promedio a los diferentes horarios del día. En el caso de Buenos Aires, se observa una hora punta por la mañana entre las 7 y las 8, con una velocidad promedio de 37 km/h, y una hora punta por la tarde (menos pronunciada que la hora punta de la mañana) entre las 17 y las 18



TIEMPOS Y VELOCIDADES PROMEDIO DE VIAJE. ÍNDICE DE CONGESTIÓN DÍA HÁBIL: 1,4

F. 06

Fuente: Elaboración propia.

horas, con una velocidad promedio de 39 km/h. El índice de congestión calculado es de 1,4; la velocidad promedio en hora valle es de 50 km/h.

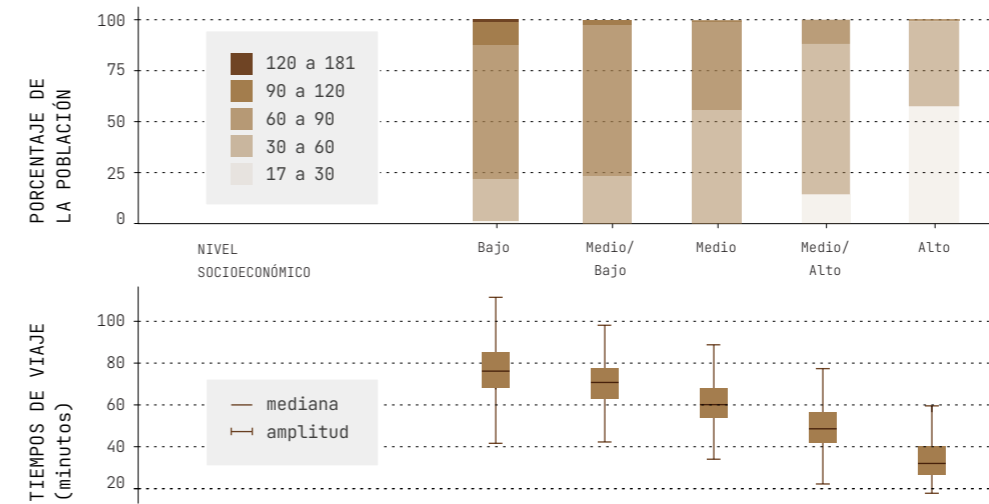
Indicadores de accesibilidad

4.3.2

Otro factor a analizar es el tiempo de viaje en transporte público, con lo cual se estudian los niveles de accesibilidad territorial de la población en la RMBA. El mapa 11 muestra el resultado de los tiempos de viaje en transporte público para la RMBA, donde se observan menores tiempos promedio en la Ciudad de Buenos Aires, en el primer cordón del área metropolitana y en los corredores norte, oeste y sur que coinciden con las líneas de ferrocarril. El segundo y tercer cordón, especialmente en las zonas más alejadas a los corredores ferroviarios, tienen mayores tiempos de viaje. La zona norte de la ciudad de Buenos Aires y el corredor norte del área metropolitana poseen los tiempos promedio de viajes en transporte público más bajos, seguidas por algunos corredores al sur y al oeste que coinciden con la red ferroviaria. Las zonas más periféricas del área metropolitana tienen tiempos promedio de viaje significativamente mayores. Esto a su vez está altamente asociado al NSE de los hogares. Se puede observar en la figura 7 que los NSE altos tienen tiempos promedio de viaje en transporte público significativamente menores que los NSE bajos.

Otro indicador de desigualdad de acceso es la distancia al área central, donde se concentra la mayor cantidad de empleo formal y registrado del sector privado (ver Alves et al., 2018). En la siguiente figura, se puede observar cómo los niveles medio-altos y altos tienen la posibilidad de realizar la mayoría de los viajes al centro en menos de 90 minutos, mientras que los niveles bajos y medio-bajos pueden realizar la mayoría de esos viajes en más de 90 minutos.

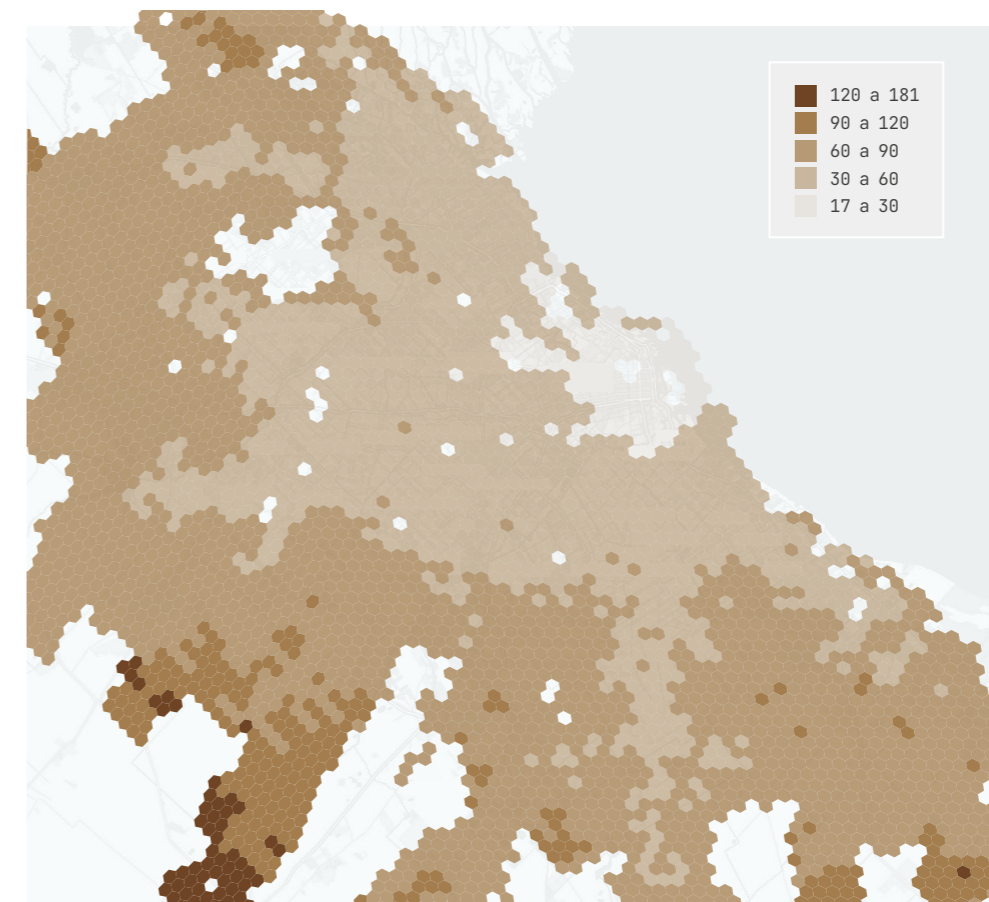
Un análisis más detallado de todos los componentes de un viaje demuestra que el sistema de transporte público tiene diferente calidad en zonas de diferente NSE. En todos los aspectos, la calidad del transporte público decrece para los NSE más bajos (tabla 1). Las personas de NSE más bajos, además de mayor tiempo de viaje en el transporte, tienen mayores tiempos de caminata, de espera



TIEMPO PROMEDIO DE VIAJES EN TRANSPORTE PÚBLICO POR NSE. RMBA

F. 07

Fuente: Elaboración propia.



TIEMPO PROMEDIO DE VIAJES EN TRANSPORTE PÚBLICO. RMBA

M. 11

Fuente: Elaboración propia.

[21]

Para acceder a los distintos centros de actividad en la RMBA, la población con NSE bajo tiene que viajar 1,8 veces más tiempo, caminar 1,4 veces más y hacer 1,5 más transferencias en los viajes en transporte público que las personas de NSE alto.

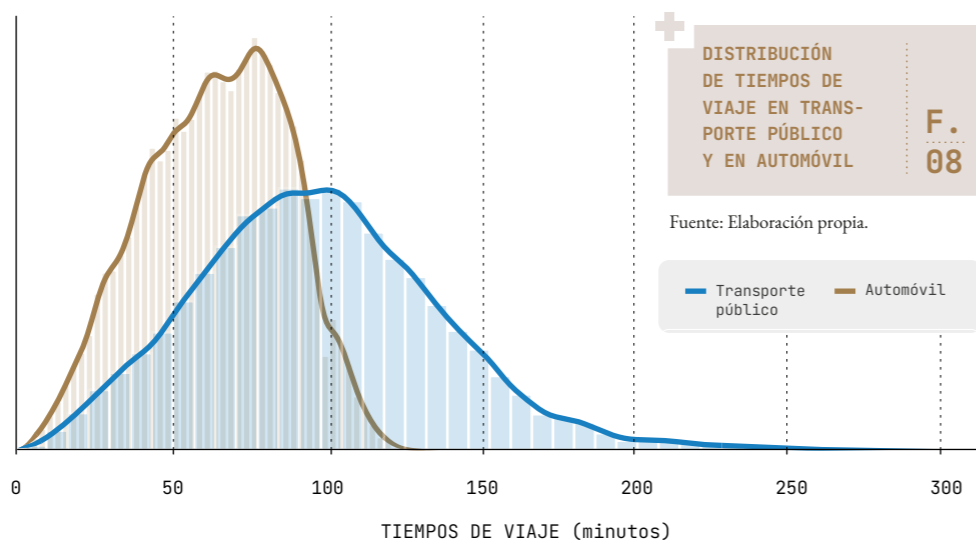
VIAJES EN TRANSPORTE PÚBLICO SEGÚN NSE. T. 01

Fuente: Elaboración propia.

y un promedio de más cantidad de transferencias en sus viajes que las personas de NSE más altos. [21] La siguiente tabla muestra con más detalle cómo la desigualdad en términos de NSE se corresponde en los tiempos de viaje total, los tiempos de viaje al centro, la caminata total y las transferencias promedio realizadas. El tiempo promedio de viaje se refiere a los viajes a todos los destinos atractores encontrados ponderados por su factor de nivel de actividad.

Nivel socioeconómico	Tiempo promedio de viaje	Tiempo de viaje al centro	Caminata total promedio(m)	Transferencias promedio
Alto	75	51	937	1.2
Medio-Alto	91	71	975	1.5
Medio	108	91	1082	1.8
Medio-Bajo	123	107	1225	1.9
Bajo	136	114	1336	1.8

Es interesante analizar las diferencias de viaje en los modos de transporte público y privado. La figura 8 muestra la distribución de tiempos de viaje en estos dos modos de transporte. Se puede observar que tanto los promedios como la dispersión son significativamente menores para los viajes en automóvil.



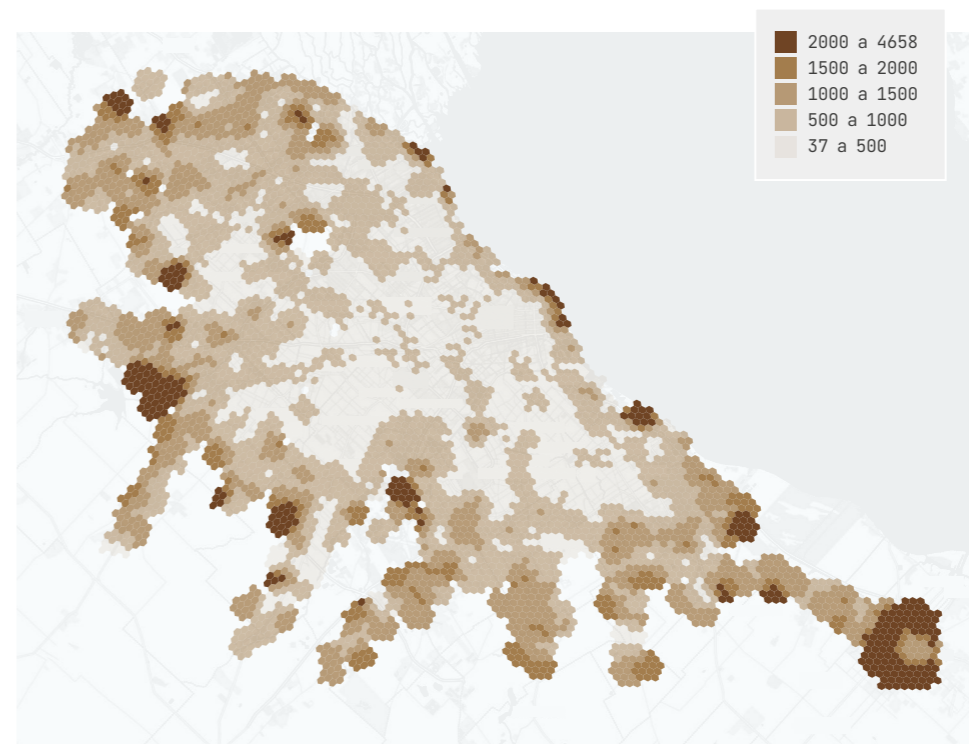
Indicadores de cobertura

4.3.3

En términos generales, se observa que la cobertura del sistema público de transporte es amplia y suministra servicios a toda el área metropolitana (ver mapa 12). [22]

[22]

La consulta de tiempos de viaje se realizó con fecha el 27 de junio de 2018.



La tabla 2 muestra que cerca del 74% de la población accede al transporte público caminando menos de 800 m y que solo el 8,4% de la población tiene que caminar más de 1500 m. Esto es gracias al extenso sistema de buses, subterráneos

Distancia a parada de transporte público	Población en 2010	Población (%)	Distancia promedio al centro (km)
0 a 500m	7.256.387	54	27
500m a 1000m	4.141.523	31	34
1000m a 1500m	1.250.441	9	41
1500m a 2000m	544.844	4	43
Más de 2000m	250.667	2	48
Total	13.443.862		39

DISTANCIA CAMINADA PARA ACCEDER AL TRANSPORTE PÚBLICO. RMBA. T. 02

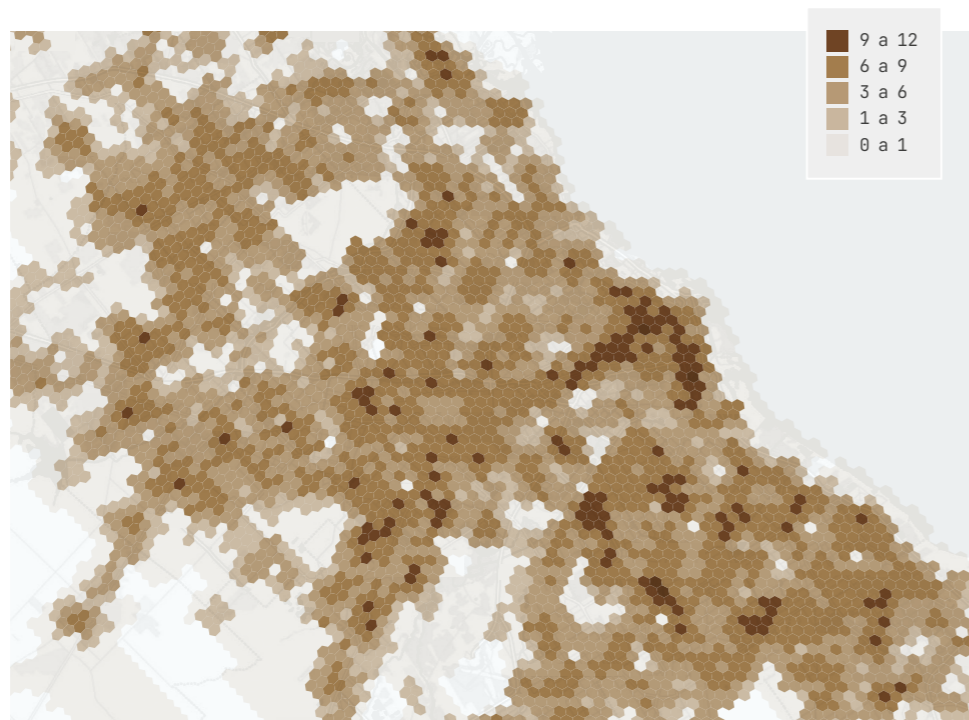
Fuente: Elaboración propia.

y ferrocarriles que sirven el área metropolitana. Se puede observar también que en las áreas periféricas (donde la distancia al área central es mayor) se requiere una caminata mayor para acceder al transporte público.

Accesibilidad a establecimientos de salud, educación y espacios verdes

4.3.4

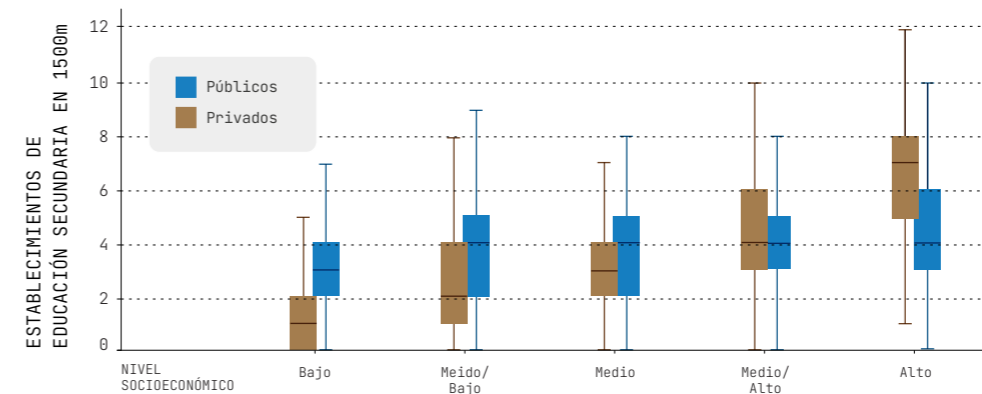
Este modelo permite también analizar el acceso a determinados tipos de establecimientos para identificar los niveles de accesibilidad territorial o por NSE a un tipo de equipamiento en particular. Como se menciona en la sección metodológica, en este trabajo se realizó el análisis para el acceso a los establecimientos educativos (inicial, primario y secundario, tanto públicos como privados), a los establecimientos de salud (hospitales y atención primaria), y a los espacios verdes.



CANTIDAD DE ESCUELAS SECUNDARIAS PÚBLICAS EN UN RANGO DE 1500m. M. 13

Fuente: Elaboración propia.

Para el caso de los establecimientos educativos y de salud, se calculan dos indicadores: tiempo de acceso en transporte público o caminando al equipamiento



ESTABLECIMIENTOS SECUNDARIOS EN UN RANGO DE 1500 M. F. 09

Fuente: Elaboración propia.

más cercano y cantidad de equipamientos en ciertos rangos de distancia (800, 1500 o 2000 m). Esto se calcula por separado para cada tipo de establecimiento. El primer indicador describe la cercanía a cada tipo de establecimiento, y el segundo, la oferta disponible, suponiendo las oportunidades de elección son mayores a mayor oferta. En el mapa 13, se observa la cantidad de escuelas secundarias públicas en un rango de 1500 m.

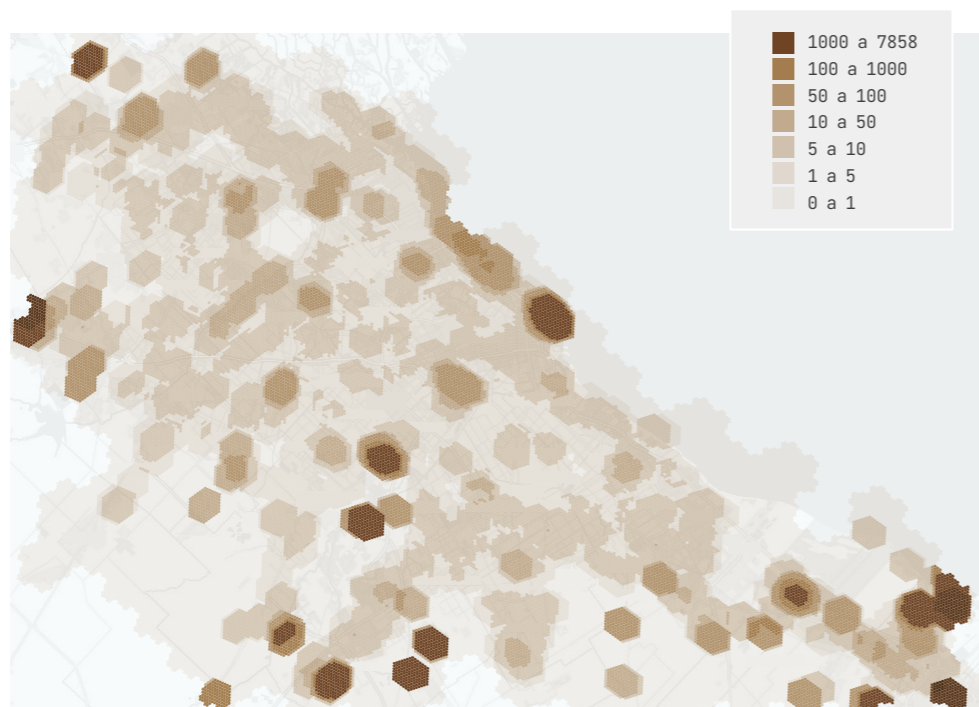
En el mapa 13 se muestra que la oferta de escuelas secundarias públicas está bien distribuida en el territorio, con una gran superficie del territorio donde hay al menos 1 o más escuelas en un rango de 1500 m. En la figura 9, se pueden observar diferencias en el acceso a escuelas públicas y privadas según NSE. Se puede observar que el acceso a escuelas secundarias privadas es considerablemente más bajo para los NSE más bajos.

Un análisis similar puede hacerse de los metros cuadrados de áreas verdes o espacios públicos per cápita en un rango de 2000 m, como se puede observar en el mapa 14, donde se observan las zonas de la ciudad con menos de 3 metros cuadrados per cápita, por ejemplo la zona sur de la CABA.

El análisis en relación al acceso a la salud difiere dependiendo si analizamos la accesibilidad a la atención primaria de la salud o a hospitales. Mientras que los tiempos de acceso a los centros de atención primaria son menores para las áreas de NSE bajo, el tiempo de acceso a hospitales es menor para las zonas de NSE alto. Esto se puede observar en la figura 10.

ESPACIOS VERDES O PÚBLICOS PER CÁPITA (M2) EN UN RANGO DE 2000m

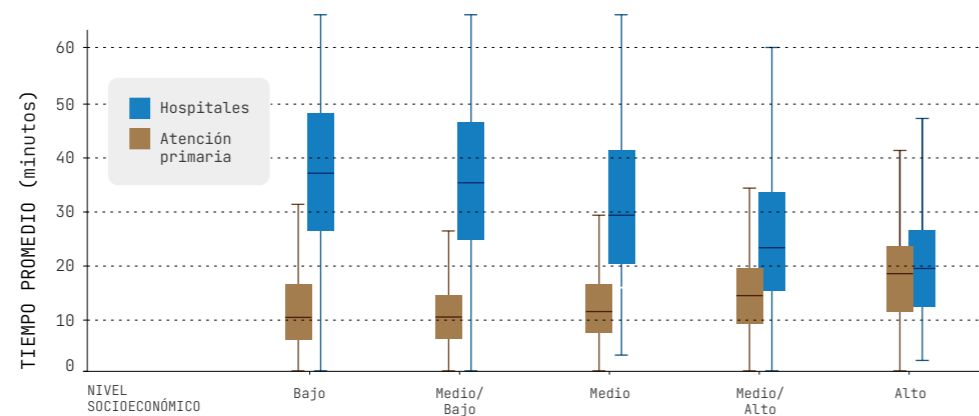
M. 14



Fuente: Elaboración propia.

TIEMPO PROMEDIO DE ACCESO A LA SALUD.

F. 10



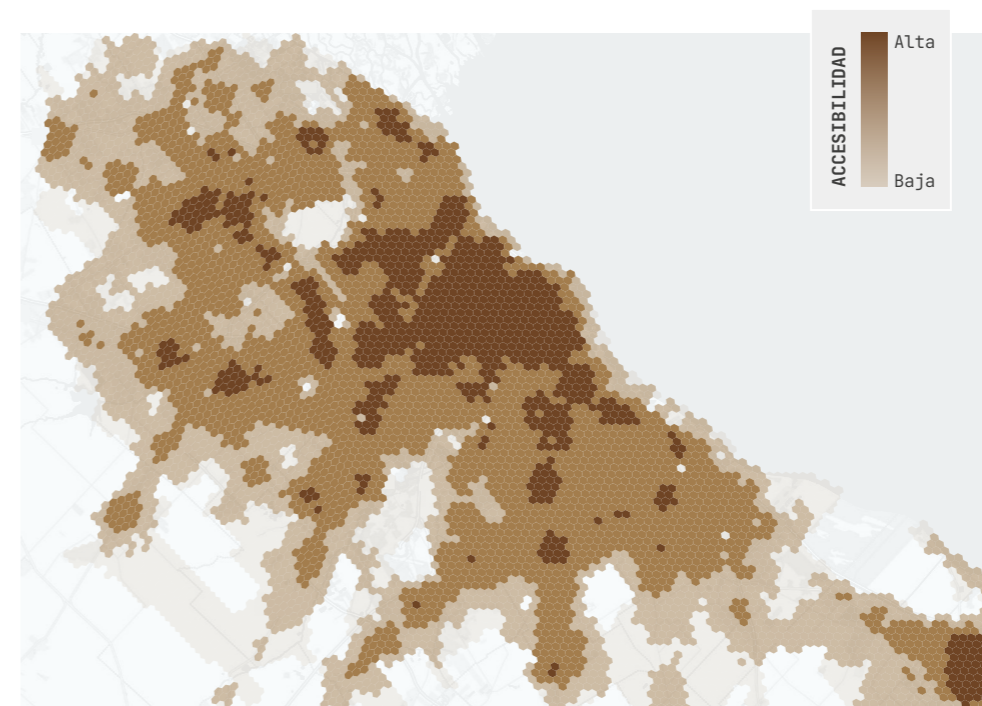
Fuente: Elaboración propia.

Clústeres de accesibilidad

4.3.5

La construcción de un índice más general utilizando distintas variables permite obtener un indicador agregado de accesibilidad. Para la construcción de este indicador, se calculan los clústeres de accesibilidad de acuerdo a lo explicitado en la sección metodológica. Este indicador tiene en consideración los distintos in-

dicadores que se construyen en el análisis y categoriza cada hexágono en uno de cuatro niveles de accesibilidad general. El mapa 15 muestra las distintas zonas del área metropolitana de acuerdo a su nivel de accesibilidad, donde se observa que la mayor parte de la ciudad de Buenos Aires está categorizada con un nivel de accesibilidad alto, con la excepción del sur de la ciudad, que está categorizado con un nivel intermedio-alto. La alta accesibilidad se extiende hacia el norte y el oeste de la ciudad (primer cordón del conurbano). Se pueden observar también altos niveles de accesibilidad en los alrededores de las estaciones de tren: al norte, el ferrocarril Mitre; al oeste, el ferrocarril San Martín y Sarmiento; y al Sur, el ferrocarril Roca, que se extiende hasta la ciudad de La Plata. Los niveles de accesibilidad más bajos se encuentran en el tercer cordón, especialmente en las zonas alejadas a los corredores de ferrocarril.



CLÚSTERES DE ACCESIBILIDAD DE LA RMBA

M. 15

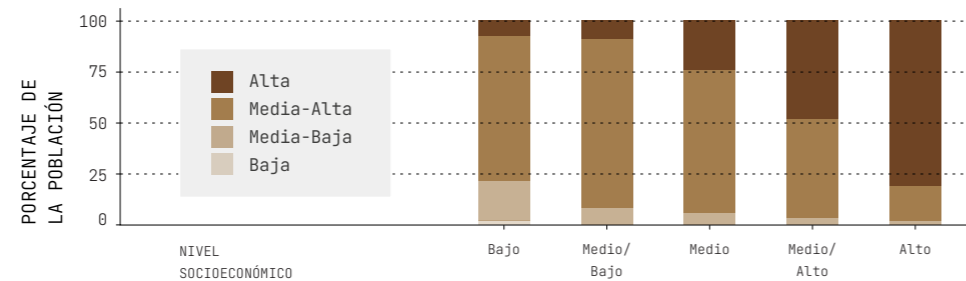
Fuente: Elaboración propia.

Análisis del caso

4.4

En la figura 11 se puede ver que los niveles más altos de accesibilidad corresponden a la población de NSE más alto. De cualquier forma, es interesante observar que solo el 20% de la población de NSE bajo tiene niveles de accesibilidad bajos o medio bajos, y que cerca del 80% tiene niveles de accesibilidad medio altos o altos.

NIVELES DE ACCESIBILIDAD TERRITORIAL GENERAL SEGÚN NSE. F. 11



Fuente: Elaboración propia.

El uso de la herramienta aplicada a la RMBA permite identificar patrones de desigualdad territorial. Si bien la cobertura del sistema de transporte es extensa y brinda servicio a la mayoría de la población que vive en el área metropolitana, la calidad del servicio varía significativamente según el NSE de las personas.

La accesibilidad medida en términos de tiempos de viaje, cantidad de transferencias y costos es relativamente alta para la CABA y para aquellas zonas del área metropolitana que se encuentran cercanas a los corredores del ferrocarril. De cualquier forma, se observa desigualdad en la accesibilidad entre quienes habitan en la CABA y quienes viven en provincia, siendo la accesibilidad en la ciudad considerablemente mayor.

Las personas que habitan en las zonas con NSE más bajos son quienes tienen los peores niveles de accesibilidad en transporte público. Los distintos indicadores comprueban este fenómeno, y se observa que la desigualdad no se explica solamente porque estos viven más alejados de las zonas céntricas de la ciudad, donde hay más oportunidades, sino también por otros componentes, como la calidad del transporte, la distancia a las paradas, tiempos de acceso y cantidad de transferencias. Los distintos mapas permiten también identificar aquellas zonas donde los niveles de accesibilidad son más bajos y podrían beneficiarse de intervenciones para mejorar las ofertas del transporte.

La comparación entre las alternativas de viajes más rápidas y las que requieren menos transferencia muestra que los tiempos de viaje mejoran cuando se realizan viajes combinando líneas o modos de transporte, pero el uso de la alternativa más rápida aumenta los costos del viaje y posiblemente impacte en otros factores como la seguridad y el confort. La reducción en los costos y mejoras en la seguridad y confort de las transferencias podría reducir los tiempos de viaje para aquellas personas que más utilizan el sistema de transporte público, como las personas de NSE más bajos, mujeres, menores de edad y adultos mayores.

Por último, se demostró que este modelo puede ser utilizado para realizar análisis de localización de establecimientos o actividades. El análisis de accesibilidad a los centros de atención primaria en la CABA indica las zonas que tienen mejor accesibilidad a estos centros y los sectores socioeconómicos beneficiados. Estos resultados son valiosos para construir nuevos centros y asegurar la provisión de un servicio a grupos específicos de la población.

Conclu
siones



{05}

05



PARTE 05

CONCLU SIONES

- Resumen del modelo analítico y oportunidades que presenta. Beneficios del uso del modelo. Resultados para la Región Metropolitana de Buenos Aires.

Conclusiones

Este trabajo presenta un modelo analítico para calcular indicadores de accesibilidad urbana a través de métodos digitales e información de uso público. Su objetivo es contribuir con una herramienta ágil y económica para estudiar los niveles de accesibilidad territorial en una ciudad, y así comprender las necesidades, priorizar intervenciones y definir mejores políticas territoriales y de movilidad.

En el marco del OMU, desarrollado por CAF en conjunto con el BID, se implementa este modelo en 29 ciudades de América Latina para calcular indicadores de cobertura de transporte público, congestión, y accesibilidad a actividades y a servicios, enmarcados en una matriz de 22 indicadores de movilidad sostenible cuyo objetivo es la evaluación y monitoreo de métricas relevantes para la movilidad urbana de las ciudades en la región.^[22]

[22]

Sitio web: <http://omu-latam.org>

Con la finalidad de comprender mejor la desigualdad territorial en las ciudades a través del monitoreo de métricas de accesibilidad, este documento presenta los pasos metodológicos para calcular indica-

dores de acceso, cobertura y congestión junto con sus elementos de visualización (gráficos y mapas). La metodología incluye un método que usa información socioeconómica y cartográfica de los censos de población y viviendas para identificar la ubicación de los hogares y construir un indicador de NSE que permite obtener indicadores desagregados por esta característica. Para la construcción del modelo de accesibilidad, se realiza una simulación de viajes utilizando aplicaciones de ruteo de viajes (OSM y Google Maps) y se implementan métodos de aprendizaje automático no supervisado para obtener centros de actividad basados en información pública sobre equipamiento urbano (extraída de la plataforma OSM) y construir un indicador agregado de accesibilidad.

El método se construye en Python (un lenguaje de programación de uso libre con una gran comunidad de usuarios), utiliza paquetes digitales de código abierto, se alimenta de datos de acceso público, e incorpora herramientas que contienen información sobre las condiciones reales del tráfico. Estas características dotan al método de una facilidad de replicabilidad en otras ciudades y regiones, ya que tiene bajo costo, es flexible y ágil, y no requiere de permisos de entidades gubernamentales para

utilizar los datos de entrada. La automatización de este método hace que sea fácilmente replicable cuando la información se actualiza o cuando los tomadores de decisión precisan indicadores actualizados en diferentes días, horarios o épocas del año.

En la actualidad, hay ciudades donde la información es incompleta, con censos no actualizados y redes de transporte no digitalizadas (por ejemplo, cuando las redes de transporte informal no están disponibles). De cualquier forma, los rápidos avances que se vienen desarrollando en temas de digitalización y apertura de datos brindan confianza en que la disponibilidad y calidad de la información va a seguir mejorando en el corto y mediano plazo. Se estima que con mejor calidad de información estos métodos brindarán información más robusta y para una mayor cantidad de ciudades.

El caso de estudio aquí presentado para la RMBA muestra cómo se aplica la metodología para un caso particular. Se pueden observar los resultados para analizar la accesibilidad a través de distintos indicadores, mapas y gráficos en una ciudad. El análisis muestra que la ciudad tiene una buena cobertura de transporte, donde cerca del 74% de la población

camina en promedio menos de 800 m para acceder al sistema de transporte público, y cerca del 91% de la población, menos de 1500 m. Asimismo, todos los indicadores muestran diferencias importantes según los NSE, donde los NSE más bajos tienen tiempos promedio de viaje considerablemente mayores que los NSE más altos. El tiempo promedio de viaje en transporte público para el área metropolitana es más elevado y con una mayor dispersión, mientras que la distribución de tiempos de viaje en vehículo particular tiene una media menor y menos dispersión. Los mapas muestran las zonas de mejor accesibilidad general, donde se observa una mayor accesibilidad en la CABA y en algunos corredores del área metropolitana que tienden a coincidir con el trazado ferroviario. Otros mapas y figuras permiten analizar la accesibilidad a equipamientos particulares, como establecimientos educativos, de salud y espacios verdes o públicos.

Anexo

PARTE 06

Construcción del indicador de NSE

6.1

Con el objetivo de medir y etiquetar de manera consistente el NSE en un conjunto diverso de países y ciudades de América Latina, se desarrolló una metodología para la construcción de un indicador compuesto para resumir en una sola variable la medición de varios subindicadores. La necesidad de realizar este tipo de análisis se basa en que el NSE es un indicador que no puede observarse directamente y debe inferirse a través de un conjunto de variables.

Este anexo presenta la construcción de un indicador compuesto y la selección de las variables para el análisis. El principal objetivo de esta metodología es que pueda implementarse con datos disponibles en todos los países de la región que pertenecen al OMU y que la metodología sea replicable de la manera más sencilla posible.

El código utilizado para la realización de este análisis se encuentra disponible en el sitio del OMU en formato de código abierto para su reutilización.

Indicador compuesto

6.1.1

Los indicadores compuestos son un tipo particular de indicadores que resumen en una sola variable la medición de varios subindicadores. El cálculo de un indicador de NSE entra en esta categoría dado que no es posible realizar una medición directa del NSE de un hogar.

El desafío de medir un indicador de NSE se basa, por un lado, en la disponibilidad de información, y por otro, en la selección de variables apropiadas para identificar las diferencias entre los distintos grupos sociales. Respecto a la disponibilidad de información, la principal fuente con la que cuentan los países

son los censos de población y vivienda, relevamientos que se realizan generalmente cada 10 años en la mayoría de los países y tienen la característica de cubrir todo el territorio.

El segundo desafío está relacionado con la selección de las variables apropiadas. Dado que no existe una sola variable que pueda definir el NSE del hogar y que se necesita un conjunto de variables, la selección de cuáles se van a incluir requiere de una revisión y análisis de la información censal de cada país.

Para dar un ejemplo, variables relacionadas con las necesidades básicas insatisfechas (como hogares sin acceso a un baño) nos permiten identificar bien situaciones de vulnerabilidad en los hogares respecto a la situación de los sectores medios y altos, pero no nos permiten identificar diferencias entre los sectores medios y altos, dado que ambos grupos suelen tener esta necesidad básica resulta en igual medida. Otras variables, como nivel educativo o acceso a ciertos bienes de consumo, nos pueden permitir realizar la diferenciación entre los sectores medios y altos. El objetivo entonces es identificar un conjunto de variables que permita diferenciar a los distintos grupos sociales entre sí.

A su vez, cada variable a incluirse debe tener en consideración el contexto en el que se realiza el análisis, dado que, por ejemplo, una variable puede ser relevante para una determinada ciudad (por ejemplo, el acceso a cloacas), pero no para otra (como una ciudad en la que la cobertura del servicio es nula o total independientemente del NSE del hogar). Se debe tener la misma consideración respecto al momento histórico en que se realiza el análisis, ya que la incidencia de las variables puede ir cambiando a lo largo del tiempo.

Análisis de variables

6.1.2

La selección de las variables que conformarán el indicador de NSE tiene algunas complejidades. En primer lugar, es necesario encontrar fuentes de indicadores que estén disponibles en toda la región a una escala geográfica adecuada, por lo que se utilizaron los censos nacionales, que son la fuente de datos más consistente

disponible a nivel territorial. En segundo lugar, es necesario identificar indicadores que puedan caracterizar lo mejor posible al conjunto de la población, que se puedan aplicar a hogares con distinta composición, incluyendo, por ejemplo, a niños y ancianos, u hombres y mujeres en edad laboral. En tercer lugar, algunos indicadores, como la ocupación, el nivel de ingresos, el acceso a servicios o el nivel educativo, pueden ofrecer resultados segmentados. Por ejemplo, las evaluaciones basadas en los niveles de ocupación tienden a enfocarse principalmente en hombres en edad laboral, por lo que puede ser un indicador deficiente para otros grupos. En este mismo sentido, el nivel de educación puede saldar la brecha de género, pero también verse afectado por la edad, ya que, por ejemplo, no incluye a la población joven que aún no obtuvo un nivel de educación formal. Finalmente, los datos sobre ingreso son menos accesibles y pueden ser una métrica engañosa en áreas con economías informales significativas.

Arias y De Vos (Arias y De Vos, 1996) describieron un método para crear un indicador socioeconómico a partir de seis variables asociadas a la vivienda que pueden ser igualmente relevantes para todos los grupos sin diferencias por edad o sexo:

- 01 → Material de la pared exterior
- 02 → Material del piso
- 03 → Material del techo
- 04 → Acceso a energía eléctrica
- 05 → Conexión al alcantarillado
- 06 → Acceso a agua potable

Si bien no desarrollaron una escala estándar, plantearon que estos indicadores pueden ser adecuados tanto para la evaluación del NSE al interior de un grupo (es decir, al interior de ciudades individuales), como entre grupos (en la comparación entre ciudades).

En esta misma línea, la ONU plantea la relevancia de indicadores de NSE similares, ya que “el tipo y la calidad de los hogares en los que se alojan las personas (el espacio, el grado de hacinamiento, las instalaciones, los alrededores, el transporte disponible) afectan sus actividades personales y el cuidado familiar e influyen en su actividad económica, en sus condiciones de salud y en sus relaciones sociales”(ONU, 1989; citado en Arias y De Vos, 1996).

Otro estudio, realizado por Vyas y Kumaranayake (Vyas y Kumaranayake, 2006), adoptó un enfoque similar para evaluar el NSE en Brasil y Etiopía, y amplió los indicadores utilizados incluyendo la propiedad de activos duraderos como radios, televisores, refrigeradores, automóviles y bicicletas. También incluyó el acceso a la red telefónica y una medida de hacinamiento vinculada a la cantidad de habitaciones disponibles para dormir.

Es importante considerar que los censos de los distintos países difieren en las variables que se recolectan, por lo que es necesario realizar el análisis de acuerdo a la disponibilidad de información en cada país o ciudad a analizar. Por otro lado, hay ciertas variables que pueden ser relevantes para algunas ciudades o en algún momento histórico en particular y que con el tiempo pueden dejar de ser relevantes. Un ejemplo podría ser el acceso a internet o a telefonía celular, dado que en algún momento representaba un bien de lujo y hoy está mucho más ampliamente disponible para el conjunto de la población.

ACCESO A LOS DATOS

Si bien los datos de los censos están generalmente disponibles como datos abiertos, existen desafíos al trabajar en distintos países. Dentro de los principales desafíos, se encuentran:

- 01 → Los sitios web de los institutos nacionales de estadísticas no siempre disponen de datos abiertos.
- 02 → Los datos e información geográfica son publicados por diversos departamentos gubernamentales, por lo que no están centralizados en un mismo sitio ni con variables o formatos similares.
- 03 → Los datos e información geográfica no se publican.
- 04 → Los datos se publican a través de plataformas separadas o de empresas privadas (por ejemplo, ESRI).^[23]
- 05 → Los datos se publican en sitios web con poca navegabilidad, enlaces rotos, etc.

^[23]
<https://www.esri.com/>

A su vez, existen diferencias en la agregación de los datos de acuerdo con los estándares adoptados en cada país y la antigüedad del relevamiento. Mientras algunos países presentan sus datos en una escala geográfica apropiada y proveen cartografía de soporte para la información censal, otros publican los datos separados de la información geográfica (y sin estar asociados por variables comunes a una capa cartográfica). Esto genera problemas al momento de querer integrar la cartografía con la información censal. En los casos en los que la información geográfica se encuentra separada de los datos, el proceso para unirlos puede encontrarse con los siguientes problemas:

- 01 → Que los datos se encuentren agregados a un nivel más alto que el requerido para el análisis, por ejemplo, a nivel municipal o departamental.
- 02 → Que los datos estén disponibles a nivel de microdatos (a nivel persona) y no existan las variables que permitan agregar al mismo nivel en el que se encuentra la cartografía.
- 03 → Que la combinación de datos e información geográfica no pueda realizarse porque alguno de los dos está desactualizado.
- 04 → Que las subdivisiones geográficas estén publicadas como archivos individuales que cubren todo el país.
- 05 → Que las subdivisiones geográficas estén publicadas en diferentes archivos cubriendo pequeñas áreas.
- 06 → Que existan errores en las geometrías publicadas que dificulten el procesamiento.
- 07 → Que los datos se encuentren incompletos, por ejemplo, debido al procesamiento en curso de un censo reciente.

A los fines de este trabajo, fue posible saldar estas limitaciones y procesar los datos y la información geográfica para crear un indicador de NSE para cada país analizado. Sin embargo, fue necesario ajustar los indicadores en cada caso de acuerdo con la información censal disponible. Los indicadores utilizados para cada país se enumeran en el apéndice, pero en términos generales incluyen indicadores del tipo:

- 01 → Materiales de la vivienda (paredes, techo, etc.).
- 02 → Acceso a servicios (electricidad, agua, saneamiento, etc.).

- 03 → Bienes (auto, computadora, etc.).
- 04 → Acceso a medios de comunicación (teléfono fijo, internet).
- 05 → Hacinamiento habitacional (más de 2 personas por habitación).
- 06 → Educación (nivel más alto alcanzado).
- 07 → Nivel de ingresos.

Sin embargo, no todas las variables seleccionadas se encontraban disponibles en todos los países o eran relevantes para todos los países o ciudades de la misma forma. Por este motivo, se seleccionó un subconjunto de variables para cada caso en particular a partir del análisis de la distribución y la correlación entre variables, pero siempre considerando que se incluyan en lo posible las características arriba listadas.

SELECCIÓN DE VARIABLES

Un paso fundamental para la utilización de este método es la selección de variables (subindicadores) que sean apropiados para calcular el indicador compuesto. Es importante considerar que las variables seleccionadas tienen que servir para medir diferencias entre distintos grupos sociales. Por un lado, tenemos que identificar una serie de hipótesis respecto a las variables a analizar, y por el otro, analizar el comportamiento de la variable en la fuente de datos disponible.

Para dar un ejemplo, podemos tener la hipótesis de que los hogares sin acceso a baño y con hacinamiento crítico corresponden a hogares muy vulnerables. Esto es válido desde el punto de vista teórico, pero si nos encontramos en los datos que la mayoría de los hogares (por un ejemplo, un 99% de los casos) tiene acceso a baño, desde el punto de vista práctico esta variable no va a resultar de utilidad. Si por el contrario, el análisis del indicador de hacinamiento nos muestra que hay diferentes niveles de hacinamiento en las diferentes divisiones geográficas, podemos concluir que este indicador nos va a servir para identificar niveles de vulnerabilidad.

Una vez que se seleccionan una o más variables que permiten identificar los hogares más vulnerables, se requiere hacer el mismo ejercicio para seleccionar variables que segmenten grupos medios y altos. Variables como el nivel de edu-

cación o el acceso a determinados bienes pueden ayudar ya que las divisiones censales con altos porcentajes de tenencia de autos o altos porcentajes de población con estudios universitarios corresponden a hogares con NSE más altos.

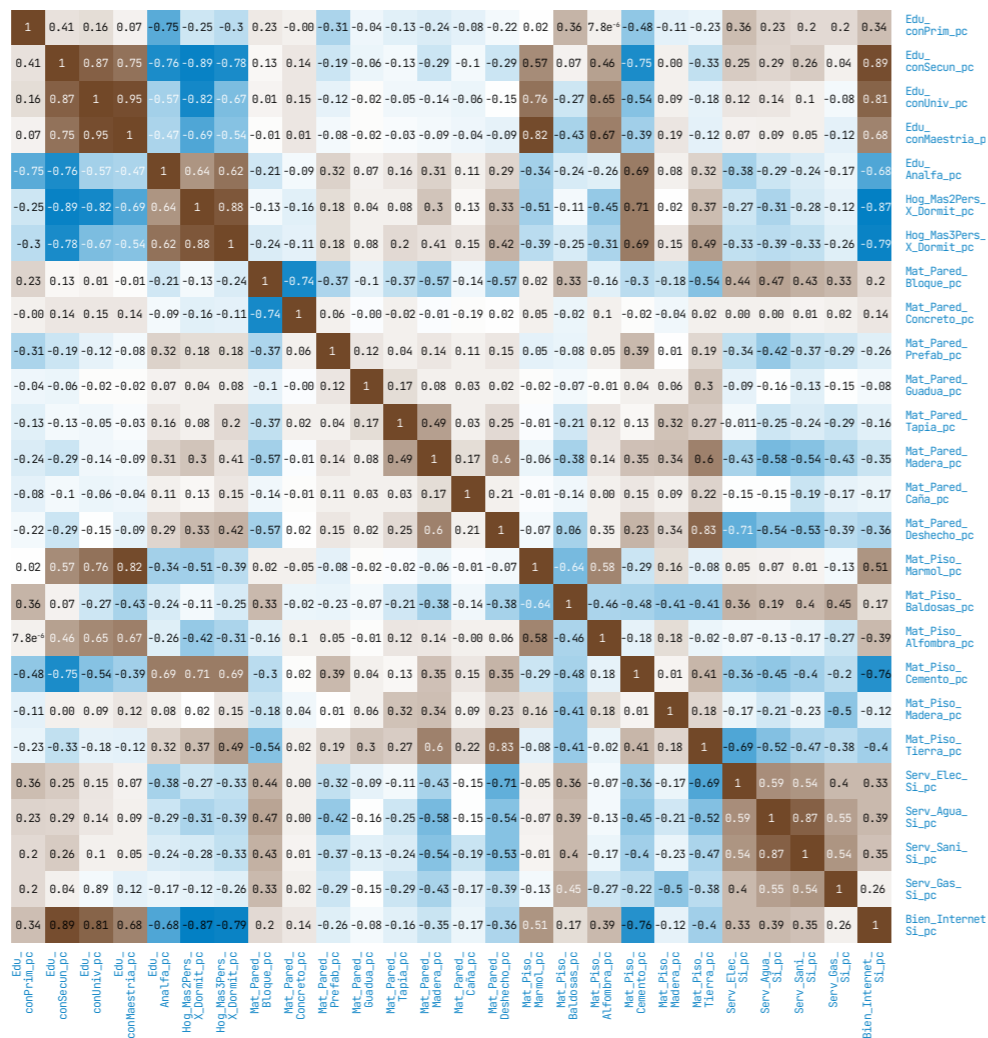
Este proceso va a requerir en una primera instancia la selección de las variables a utilizar (teniendo en cuenta nuestras hipótesis de cómo afectan al NSE) y un análisis tanto de la distribución de las variables como de su correlación con las demás. La visualización de la correlación entre todas las variables seleccionadas para la construcción del indicador de NSE permite identificar variables que tienen poca relación con otras en el conjunto.

La figura 12 a continuación presenta la correlación entre las variables utilizadas para la ciudad de Bogotá. La escala de colores divergente muestra los distintos niveles de correlación, donde un color más intenso representa una correlación más fuerte y un color más claro una correlación más débil (coeficiente de Pearson cercano a cero). Las celdas con colores del celeste claro al azul tienen valores de correlación menores a 0 y mayores a -1, es decir, una correlación negativa entre variables. Por ejemplo, variables como hogares con 2 o 3 personas por dormitorio (Hog_mas2Pers_X_Dormit, Hog_mas3Pers_X_Dormit) muestran una correlación negativa con las variables de educación secundaria, universitaria y de posgrado (Educ_conSecun, Educ_conUniv, Educ_conMaestria). En cambio, los colores tierra representan correlaciones positivas, valores mayores a 0, como el acceso a internet (Bien_Internet_Si) y las variables de educación.

CORRELACIÓN ENTRE INDICADORES PARA LA CIUDAD DE BOGOTÁ

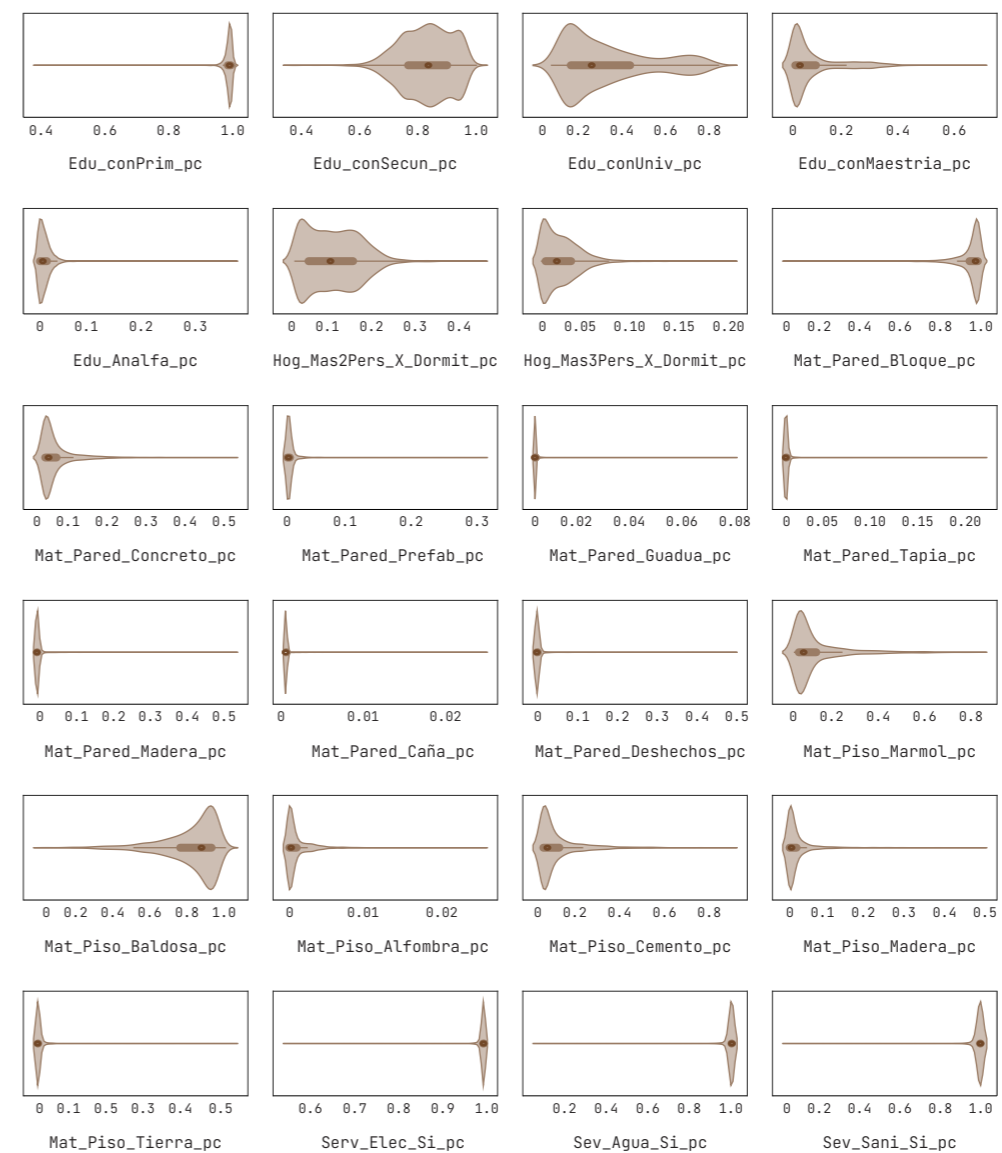
F. 12

Fuente: Elaboración propia.



DISTRIBUCIÓN DE DENSIDAD O DE FRECUENCIA DE VARIABLES

Un análisis de la distribución de las distintas variables permite identificar cuáles son las más apropiadas para incluir en el análisis. En los casos en los que las variables presentan poca variación (por ejemplo, si un bajo porcentaje de las casas en todas las zonas tiene paredes hechas de caña, por lo que no ayudan a distinguir entre casos), conviene eliminar estas del análisis. El gráfico a continuación presenta ejemplos de estas distribuciones y muestra cómo algunas



VISUALIZACIÓN DE LA DISTRIBUCIÓN DE VARIABLES DEL CENSO DE BOGOTÁ

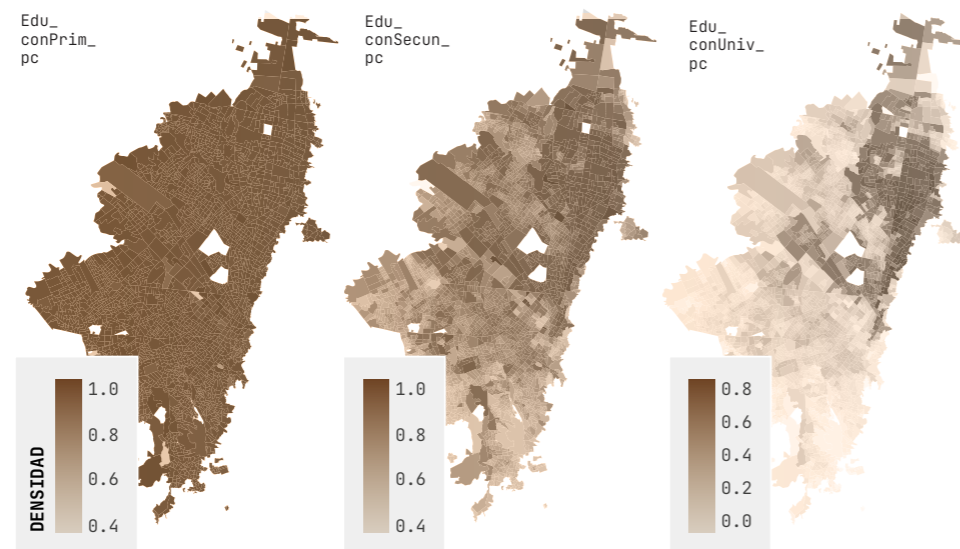
F. 13

Fuente: Elaboración propia.

variables presentan mucha variación (y por lo tanto, otorgan mucha información), mientras que otras, no. Por ejemplo, por un lado, Bien_Internet_Si y Educ_conUniv muestran una amplia variación, mientras que Serv_Electr_Si, Mat_Pared_Caña presentan una dispersión muy baja.

DISTRIBUCIÓN ESPACIAL DE VARIABLES

Por otro lado, observar la distribución espacial de cada variable individualmente puede resaltar los niveles de variación y servir para verificar la distribución espacial de NSE y su correlación con otros indicadores. En el siguiente ejemplo, todas las personas de la ciudad de Bogotá, en todas las zonas, tienen una educación al menos de nivel primario (Edu_conPrim, a la izquierda en el siguiente ejemplo), pero no así de un nivel superior (universitaria, Edu_conUniv, a la derecha).



VISUALIZACIÓN DE LA DISTRIBUCIÓN ESPACIAL DE UNA MUESTRA DE VARIABLES EN LA CIUDAD DE BOGOTÁ. M. 16

Fuente: Elaboración propia.

Método para la reducción de dimensiones de variables

Debido a que para la construcción de este indicador no se quiere limitar demasiado el número de variables, se siguió el ejemplo de Vyas y Kumaranayake (Vyas y Kumaranayake, 2006) y se recurrió a las técnicas de aprendizaje automático no supervisado para extraer patrones de datos multidimensionales y reducir los muchos indicadores originales a un solo indicador compuesto. En base a lo expuesto, se evaluaron dos enfoques: por un lado, el análisis de componentes principales (PCA), y por otro, el agrupamiento de k-medias (k-means clustering).

El enfoque PCA fue utilizado por Vyas y Kumaranayake, y también ha sido descrito por Reades, De Souza y Hubbard (Reades et al., 2019), mientras que el enfoque de agrupamiento k-means siguió de cerca el trabajo de Spielman y Singleton (Spielman y Singleton, 2015) y de Arribas-Bel (Arribas-Bel, 2019).

Componentes principales (o PCA, por sus siglas en inglés) es una técnica de reducción del total de las variables utilizadas a un número específico de componentes principales. Esta es una técnica estadística multivariable utilizada para reducir el número de variables en un conjunto de datos a un número más pequeño de dimensiones (Vyas y Kumaranayake, 2006). Las dimensiones se ordenan por la cantidad de varianza original que describen, por lo que la técnica es útil para reducir la dimensionalidad de un conjunto de datos.

Este análisis es más efectivo cuando la correlación entre variables es alta, es decir, cuando hay una relación lineal entre las variables. El PCA identifica esta relación y la expresa como una nueva dimensión, efectivamente reduciendo las dos dimensiones originales a una. El proceso es análogo en dimensiones superiores.

Los primeros componentes principales describen la mayor parte de la varianza de los datos (que será mayor cuanto más correlacionadas estuvieran las variables originales), y se espera, como señalan Vyas y Kumaranayake (2006), que el primer componente principal sea la medida del estado económico. Estos componentes de bajo orden a veces contienen el aspecto más importante de la información, y los demás componentes se pueden ignorar (Peres-Neto et al., 2005). Por lo tanto,

para la construcción de este indicador de NSE se selecciona el primer componente principal, por lo que es importante analizar el porcentaje de la varianza explicada por este componente de acuerdo a la variables seleccionadas para el análisis.

Para este enfoque no es importante si los indicadores individuales (por ejemplo, porcentaje de personas con o sin automóvil) contribuyen positiva o negativamente al puntaje socioeconómico, lo que importa es que haya una buena distribución de valor que no sufra aglomeraciones o truncamiento, y que estos contribuyan a un primer componente principal con máxima varianza.

Como se mencionó anteriormente, debido a que el resultado de PCA es una variable continua, es posible usar una ponderación adicional para ajustar los puntos de ruptura al clasificar los resultados en grupos. En los casos de estudio de este trabajo, se dividieron los resultados en quintiles y terciles ponderados por población. La ponderación por población permite obtener clases o grupos de NSE con poblaciones más o menos homogéneas.

El k-medias es otra técnica de aprendizaje no supervisado que permite realizar una agrupación de observaciones de una cierta cantidad de grupos. Con esta técnica, cada observación (división censal, en este caso) va a pertenecer al grupo cuyo valor medio sea más cercano. El principal problema de utilizar este método es que, si bien podemos separar distintos grupos según las características de las variables, no tenemos forma de asignar un orden a estas clasificaciones. Por otro lado, al ser un método de clasificación, podemos tener como resultado clases que representan cantidades de población muy diferentes entre sí.

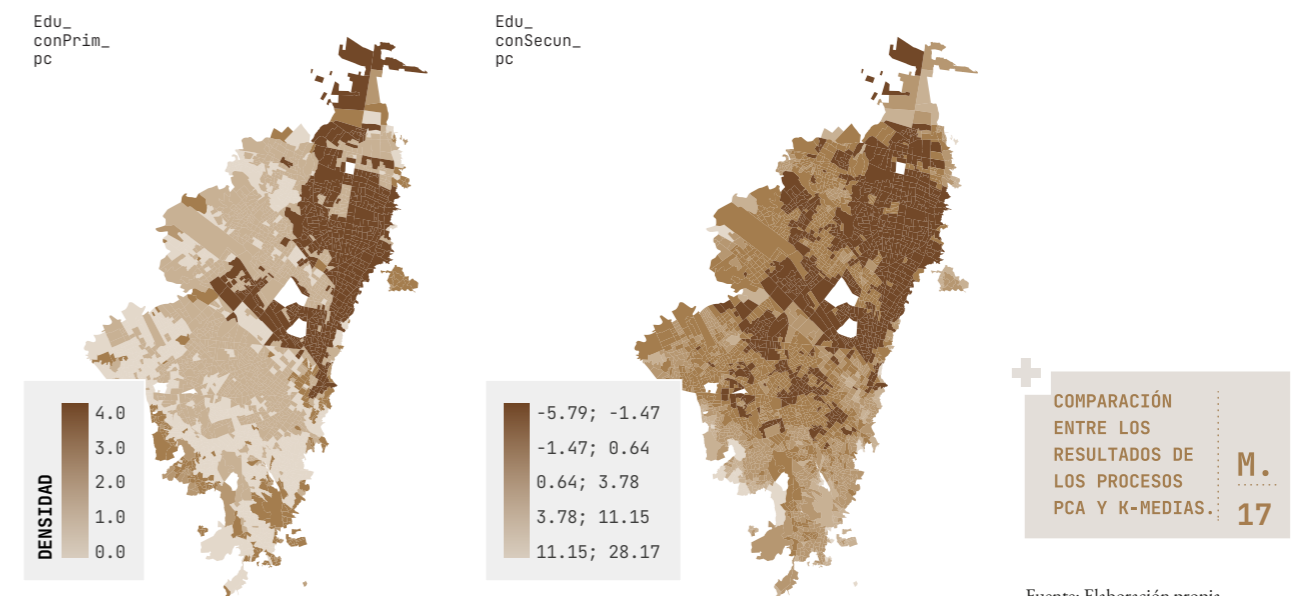
ESTANDARIZACIÓN DE LAS VARIABLES

Las técnicas de aprendizaje automático como el análisis de componentes principales requieren estandarización de las variables para evitar los sesgos de incluir valores absolutos de variables con escalas diferentes. Por ejemplo, si los materiales de la pared obtuvieron un rango de 0-6 mientras que el tipo de conexión a la alcantarilla solo se clasificó de 0-2, el material de la pared puede tener más peso en el resultado final. Para evitar esto, las variables categóricas se dividieron en variables binarias indivi-

duales, por ejemplo: “madera no o sí” (0-1), “ladrillo no o sí”, etc. Luego se evaluó el valor relativo de esos indicadores cuantificando el porcentaje de hogares en cada zona con ese indicador. El procedimiento que realiza el cálculo de componentes principales lleva a cabo la estandarización de cada una de las variables utilizadas.

COMPARACIÓN DE TÉCNICAS

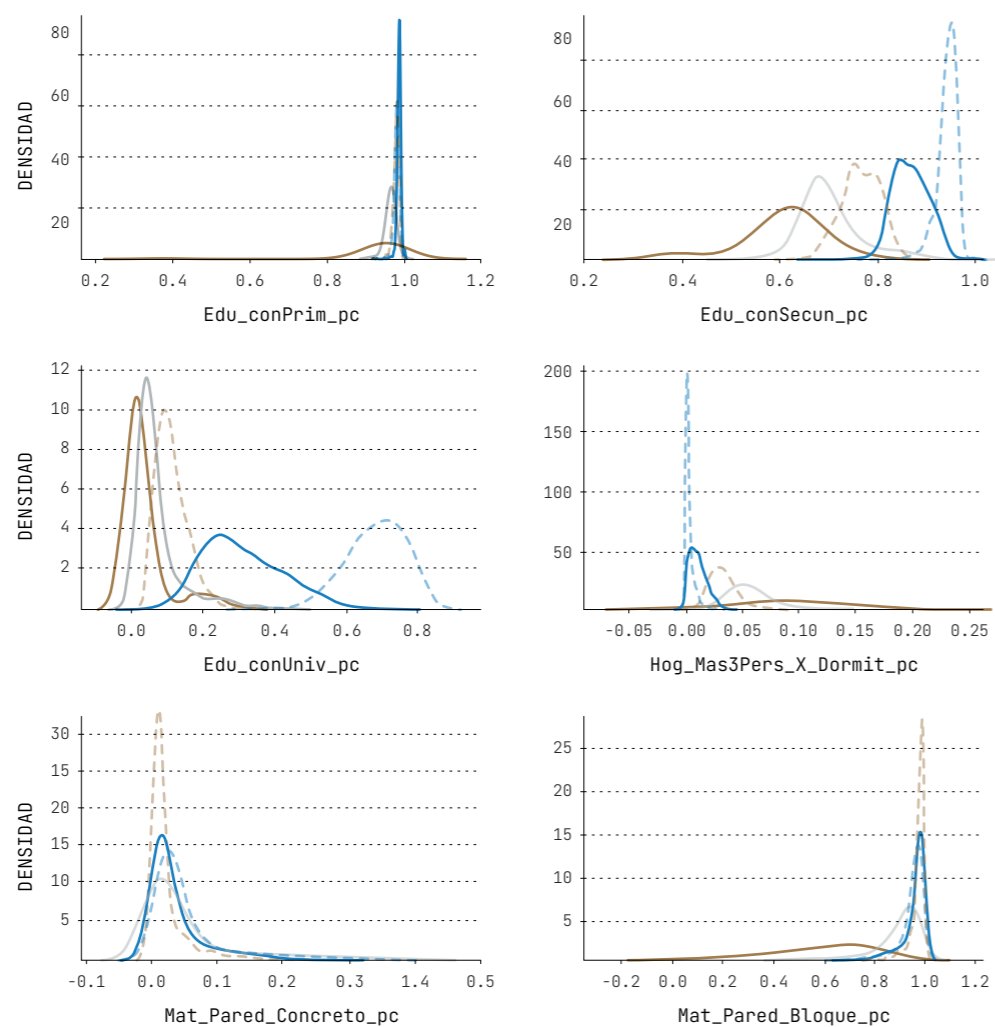
La diferencia clave entre las dos técnicas es que la agrupación k-medias (más conocida como *k-means*) recopila casos con métricas similares en grupos discretos, mientras que PCA calcula un nuevo indicador con una escala continua. Si bien la forma en que las técnicas agrupan los resultados difiere, la distribución espacial de los valores altos y bajos para los casos estudiados ha sido similar. En los mapas a continuación se presentan la distribución de k-medias (a la izquierda) y PCA (a la derecha):



Fuente: Elaboración propia.

Un beneficio que proporciona el agrupamiento k-medias es que es posible comparar la distribución de las variables componentes dentro de los grupos discretos. La siguiente imagen muestra la distribución de los indicadores originales dentro de los agrupamientos creados por k-means clustering. Se supone que cada agrupamiento (representado por diferentes líneas) representa un NSE distinto.

Fuente: Elaboración propia.



Una desventaja de la agrupación de k-medias es que los grupos pueden ser muy desiguales en tamaño, con un gran número de casos asignados a un grupo y pocos a los otros. Como los grupos son discretos, no es posible ajustar las divisiones entre los grupos. Por el contrario, debido a que el resultado de PCA es continuo, es posible dividir en grupos, por ejemplo, terciles o quintiles poblacionales, lo que permite obtener grupos más homogéneos.

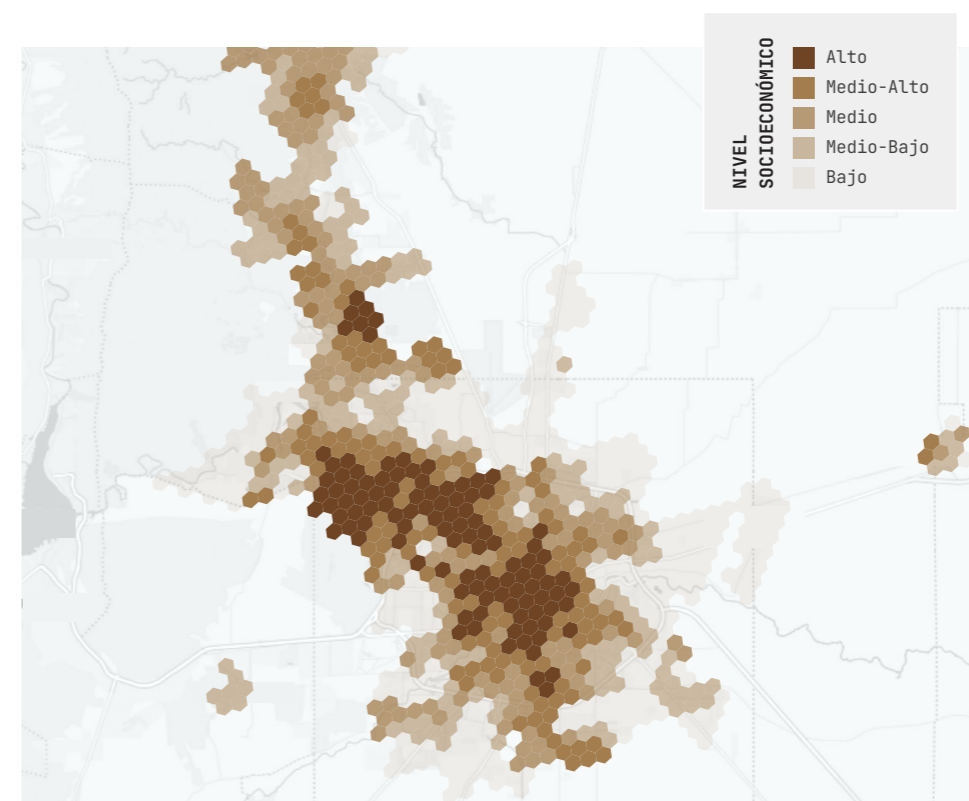
Por esta razón, en este estudio, se eligió trabajar con la técnica de PCA para poder pesar los grupos de diferentes NSE por población. Esto permitió crear cinco grupos con aproximadamente el mismo número de personas en cada uno, algo que no hubiera sido posible con la técnica k-means.

Construcción de indicador compuesto NSE

6.1.4

Luego de seleccionar las variables que deben ser incluidas en el análisis y de seleccionar el método para reducir dichas dimensiones, se procedió a aplicar este método para 33 ciudades de América Latina (de 9 países) en el marco de la tercera edición del OMU. El resultado de este procedimiento es la construcción de una única variable que caracteriza el NSE de un conjunto de hogares que conforman divisiones geográficas censales. Una vez asignado el indicador a las divisiones geográficas, estas se dividen en quintiles y terciles de NSE teniendo en cuenta la población de cada división geográfica.

A continuación se presenta como ejemplo el resultado para la ciudad de Córdoba en Argentina.

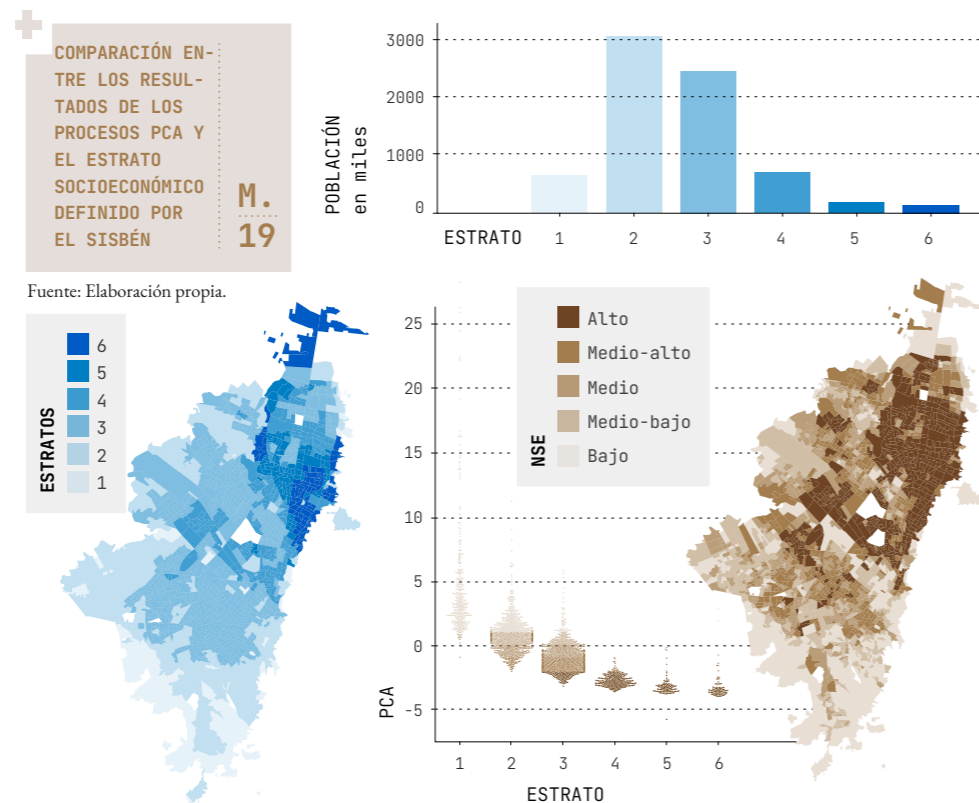


Fuente: Elaboración propia.

COMPARACIÓN CON OTROS INDICADORES DE NSE

Algunos países, como el caso de Colombia, publican su propio indicador de NSE. Esto permitió realizar una comparación de los resultados obtenidos con el indicador de NSE construido para este análisis.

El mapa a continuación muestra el ejemplo de Bogotá con los resultados obtenidos a partir de la utilización de ambos indicadores de NSE, el construido para este análisis (a la derecha) y el existente a nivel gubernamental (a la izquierda).^[24] Más allá de que los resultados y las agrupaciones evidencian diferencias debido al uso de metodologías distintas (el indicador de Colombia está más enfocado a identificar situaciones de vulnerabilidad), se puede identificar una amplia correlación entre las categorías y la distribución espacial de ambos métodos.



Este tipo de comparaciones con indicadores de NSE existentes permitió validar los resultados obtenidos por este estudio.

VENTAJAS Y LIMITACIONES

Contar con una herramienta de estas características para construir un indicador de NSE permite realizar análisis de accesibilidad teniendo en cuenta los distintos grupos poblacionales. Las ventajas de este método son varias: 1) utiliza datos que generalmente se encuentran disponibles a nivel país (censos de población y viviendas); 2) estos datos suelen tener un nivel de granularidad suficiente para que se pueda construir a nivel territorial (radios, secciones o manzanas censales); 3) es sencillo de implementar, ya que solo requiere poder identificar variables apropiadas para el análisis; 4) es fácilmente replicable para otras ciudades y el método se encuentra a disposición como código abierto en el sitio del OMU. Si bien pueden existir otros métodos más robustos, que incluyan un análisis más detallado de las variables y ponderaciones más complejas, estos son más difíciles de implementar, tienen un mayor requerimiento de datos que no siempre están disponibles y son menos replicables para todo el conjunto de ciudades que se están analizando.

Conclusiones

6.1.5

Contar con un indicador de NSE es de gran utilidad, ya que este sirve como insumo para una gran variedad de estudios de accesibilidad. El cálculo de este indicador permite asignar a todo el territorio un NSE que se puede utilizar para distintos análisis. Por ejemplo, puede utilizarse para realizar simulaciones donde se le asigne viajes a hogares con distintos NSE para analizar las diferencias de accesibilidad, o asignarse un NSE al origen de viajes obtenidos de transacciones de los sistemas de pago de boletos en transporte público o de telefonía celular. En cualquier análisis de movilidad, la posibilidad de incluir la dimensión relacionada con el NSE permite dar cuenta de las desigualdades territoriales y ofrece un análisis con más detalle para desarrollar políticas de transporte.

Identificación de destinos atractores de viajes (DBSCAN) 6.2

Para generar un indicador de demanda del transporte que sea escalable y replicable para el conjunto de ciudades, se desarrolló un indicador que calcula la densidad de equipamientos en los aglomerados. En consecuencia, la densidad de equipamientos es usada como un proxy de densidad de actividad y, por ende, como un proxy de atracción de viajes.

Datos de OSM 6.2.1

OSM es una plataforma abierta que permite la descarga de una importante variedad de equipamientos que están categorizados en distintos tipos, y van desde restaurantes, establecimientos de salud y educativos, bancos, parques o espacios públicos.^[25] Para este análisis se usan las categorías relacionadas con establecimientos (*Amenity*). Para la identificación de los parques y espacios públicos, se descargan las categorías de esparcimiento y espacios naturales (*Leisure* y *Natural*). La categoría de espacios aeroportuarios (*Airways*) se utiliza, junto con las de esparcimiento y espacios naturales, para la redistribución de la población desde los radios o secciones censales hacia los hexágonos.

Ponderación de equipamientos 6.2.2

Una vez descargados los equipamientos para cada ciudad desde OSM, se le asigna un factor de ponderación a cada tipo de establecimiento de acuerdo al nivel de atracción de viajes designado (por ejemplo, un hospital tiene un factor de atracción mayor al de una farmacia). A continuación se muestra un ejemplo del listado de equipamientos y la ponderación utilizada.

Amenity	fex	Amenity	fex
bank	1000	shelter	20
bus_station	1000	fire_station	10
hospital	1000	fuel	10
university	1000	parking	10
clinic	500	police	10
college	500	animal_shelter	0
food_court	500	bbq	0
school	500	bench	0
arts_centre	200	bicycle_rental	0
casino	200	bureau_de_change	0
childcare	200	car_rental	0
community_centre	200	car_wash	0
kindergarten	200	clock	0
library	200	dentist	0
atm	100	dojo	0
bar	100	drinking_water	0
biergarten	100	driving_school	0
cafe	100	events_venue	0
convenience	100	love_hotel	0
courthouse	100	marketplace	0
fast_food	100	motorcycle_parking	0
fountain	100	nightclub	0
internet_cafe	100	nursing_home	0
payment_centre	100	parking_entrance	0
place_of_worship	100	patrimonio_cultural	0
post_office	100	payment_terminal	0
pub	100	post_box	0
restaurant	100	prison	0
taxi	100	public_b	0
cinema	50	public_bookcase	0
exhibition_centre	50	recycling	0
ice_cream	50	security_booth	0
parking_space	50	social_facility	0
pharmacy	50	stock_exchange	0
social_centre	50	studio	0
theatre	50	telephone	0
toilets	50	townhall	0
veterinary	50	waste_basket	0
doctors	20		

LISTA DE EQUIPAMIENTOS CON EL FACTOR DE PONDERACIÓN DE ACTIVIDAD ASIGNADO (FEX). **T. 03**

Fuente: Elaboración propia.

^[25] En https://wiki.openstreetmap.org/wiki/Map_features se puede ver el listado de equipamientos y sus categorías.

DBSCAN

6.2.3

Utilizando la base de los establecimientos con su ponderador de actividad, se aplican técnicas de agrupamiento o *clustering* para identificar áreas de alta densidad de actividad. Se aplica el algoritmo DBSCAN, que permite conformar grupos o clústeres con alta densidad de equipamientos (pesando cada equipamiento con su factor de ponderación). Esta técnica de aprendizaje automático no supervisado permite la detección de patrones en forma automática basándose en la ubicación espacial de los puntos y la distancia entre ellos. Como resultado de este proceso, se detectan áreas con alta concentración de puntos y se identifican las áreas vacías o con escasos puntos. Los puntos que no forman parte de un clúster se etiquetan como ruido. Este tipo de algoritmos es ampliamente utilizado para realizar análisis espaciales.^[26]

[26]
Wang et al., NS-DBSCAN.

Se utilizaron los siguientes parámetros de agregación:

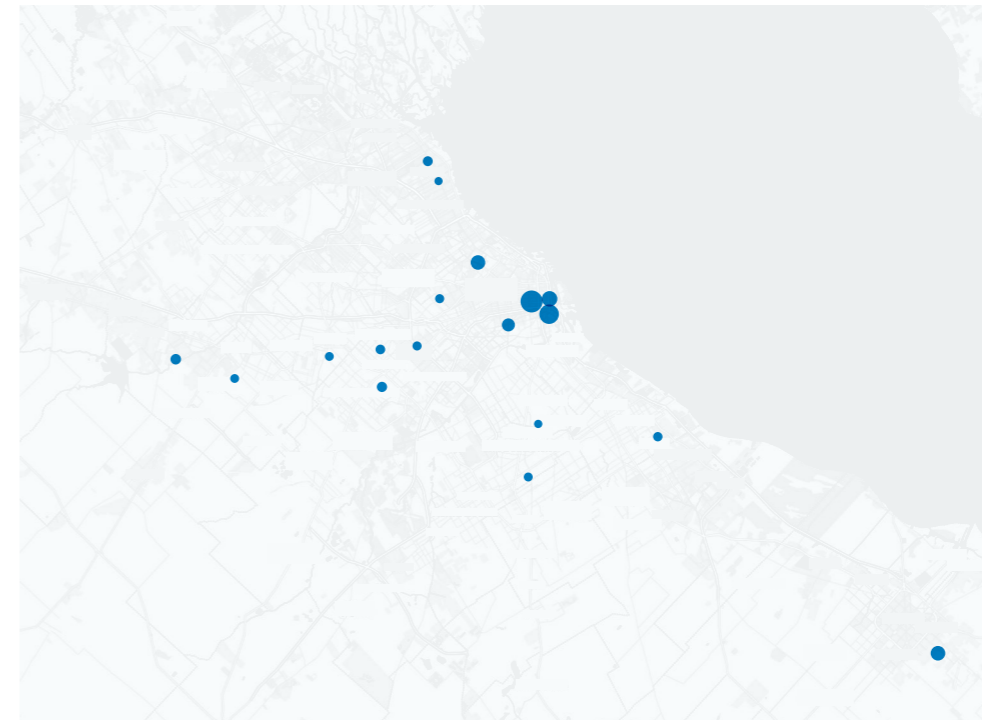
01 → Punto núcleo (o puntos directamente alcanzables): representa la distancia máxima que puede haber entre puntos para conformar un clúster. Para este ejercicio se realiza una serie de corridas de clusterización con los siguientes parámetros: 500, 1000, 1500 y 2000 m, que representan respectivamente los clústeres 1, 2, 3 y 4. Según el tipo de ciudad, se elige el más conveniente.

02 → Mínimo de puntos: es la cantidad mínima de puntos que tiene que haber para conformar un clúster. Esta cantidad mínima de puntos se va a establecer de acuerdo al tamaño de cada ciudad. La cantidad de puntos tiene en cuenta el factor de expansión de cada punto (equipamiento).

Clústeres de actividades

6.2.4

Una vez identificados los clústeres de alta densidad de actividad, estas ubicaciones geográficas son utilizadas como puntos de destino para realizar la simulación de los viajes. Dentro de los puntos identificados como áreas de alta densidad, hay unos con mayor densidad de actividad que otros. Para el análisis, se considera que los puntos con mayor densidad de actividad tienen mayor atracción de



CLÚSTERES DE DENSIDAD DE ACTIVIDAD EN LA RMBA
M.
20

Fuente: Elaboración propia.

viajes. Este factor de atracción va a ser utilizado para ponderar la construcción de indicadores relacionados con los tiempos de viaje.

El mapa resultante (20) muestra la densidad de actividades a través de la conformación de los distintos clústeres en la medida que cambian los hiperparámetros del modelo de clusterización. Los clústeres que se conforman en un radio de 500 m muestran una mayor densidad de actividad que los clústeres conformados en un radio de 2000 m.

La librería `pyomu` permite realizar un proceso de clustering de equipamientos urbanos. La función tiene la opción de traer equipamientos de Open Street Maps, les da un ponderador a cada tipo de establecimiento y realiza un proceso de clusterización para identificar las áreas de alta densidad de actividad.

`calculate_activity_density()`

Referencias

Alves, G., Berniell, L., V. C., de la Mata, D., Fernandez, D., Juncosa, C., y Rotondo, S. (2018). Distribución espacial del empleo formal en la ciudad autónoma de buenos aires: Un diagnóstico a partir de registros administrativos. <https://cafscioteca.azurewebsites.net/handle/123456789/1406>

Arias, E., y De Vos, S. (1996). Using housing items to indicate socioeconomic status: Latin America. *Social Indicators Research*, 38(1), 53–80. <https://doi.org/10.1007/BF00293786>

Arribas-Bel, D. (2019). A course on Geographic Data Science. *Journal of Open Source Education*, 2(16), 42. <https://doi.org/10.21105/jose.00042>

Carmines, E., y Zeller, R. (1979). *Reliability and Validity Assessment*. SAGE Publications, Inc. <https://doi.org/10.4135/9781412985642>

Carruthers, R. C., Dick, M. C., y Saurkar, A. (2005). Affordability of public transport in developing countries (N.o 33900; pp. 1-27). The World Bank. <http://documents.worldbank.org/curated/en/230991468153275100/Affordability-of-public-transport-in-developing-countries>

CEPAL, (Comisión Económica para América Latina y el Caribe). (2012). *Población, territorio y desarrollo sostenible*. CEPAL. <https://www.cepal.org/es/publicaciones/22425-poblacion-territorio-desarrollo-sostenible>

Daude, C., Fajardo, G., Brassiolo, P., Estrada, R., Goytia, C., Sanguinetti, P., Álvarez, F., y Vargas, J. (2017). RED 2017. Crecimiento urbano y acceso a oportunidades: Un desafío para América Latina. CAF. <http://scioteca.caf.com/handle/123456789/1090>

DeVellis, R. F. (1991). *Scale Development: Theory and Applications*. Sage Publications, Applied Social Research Methods Series, Vol. 26.

Geurs, K. T., y van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, 12(2), 127-140. <https://doi.org/10.1016/j.jtrangeo.2003.10.005>

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, 25(2), 73-76. <https://doi.org/10.1080/01944365908978307>

Hernández, D., y Hansz, M. (2021). Accesos a oportunidades para favorecer la inclusión. Aspectos conceptuales, indicadores y su medición. CAF. <https://cafscioteca.azurewebsites.net/handle/123456789/1692>

Herrera, L., Olivares, F., y Pecht, W. (1976). Crecimiento urbano de América Latina. <https://repositorio.cepal.org/handle/11362/8609>

Litman, T. (2003). Measuring transportation: Traffic, mobility and accessibility. *Social Research in Transport (SORT) Clearinghouse*, 73(10).

Muzzini, E., Eraso Puig, B., Anapolsky, S., Lonnberg, T., y Mora, V. (2016). *Leveraging the Potential of Argentine Cities: A Framework for Policy Action*. World Bank Publications. <https://openknowledge.worldbank.org/handle/10986/24185>

Peres-Neto, P. R., Jackson, D. A., y Somers, K. M. (2005). How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 49(4), 974–997. <https://doi.org/10.1016/j.csda.2004.06.015>

Reades, J., De Souza, P., y Hubbard, P. (2019). Understanding urban gentrification through machine learning. <https://journals.sagepub.com/doi/10.1177/0042098018789054>

Ronan. (2017, mayo 12). *Linking People and Places* [Text]. ITF. <https://www.itf-oecd.org/linking-people-and-places>

Spielman, S. E., y Singleton, A. (2015). Studying Neighborhoods Using Uncertain Data from the American Community Survey: A Contextual Approach. *Annals of the Association of American Geographers*, 105(5), 1003–1025. <https://doi.org/10.1080/00045608.2015.1052335>

Vyas, S., y Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, 21(6), 459–468. <https://doi.org/10.1093/heapol/czl029>



OBSERVATORIO
DE MOVILIDAD
URBANA

ESTE REPORTE
SE PUBLICÓ
EN ARGENTINA.
ENERO 2023



CAF
BANCO DE DESARROLLO
DE AMÉRICA LATINA

BID
Banco Interamericano
de Desarrollo