



Implementación de Big Data en la planificación del transporte e infraestructura urbana

Encuesta de Movilidad Bogotá - Región 2023

Encuesta de Movilidad Bogotá, Región 2023

Grupo Cal y Mayor formó parte del consorcio que realizó “La Encuesta de Movilidad” para la Secretaría Distrital de Movilidad de Bogotá. Este proyecto incluyó la Encuesta Origen - Destino de Hogares (EODH) y la Encuesta Origen - Destino de Interceptación (EODI) para Bogotá y 19 municipios vecinos en su área de influencia, además de la actualización del modelo de transporte de cuatro etapas en el área de estudio.

A partir de los resultados de este estudio, se llevó a cabo una comparación entre los indicadores obtenidos mediante encuestas directas (método tradicional) y el uso de Big Data con información de la misma zona de estudio, provista por Kido Dynamics.

El objetivo del análisis fue identificar el potencial de este tipo de datos para complementar o reemplazar la información de campo obtenida con métodos tradicionales.



Zona de estudio y método

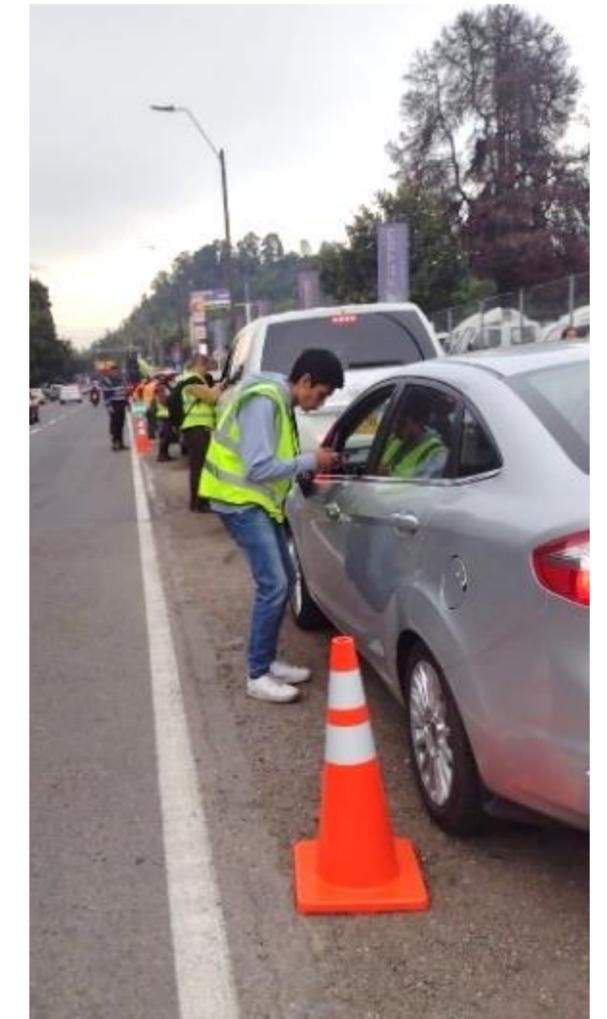
En el área de estudio habitan un total de 9,273,186 personas mayores de 5 años, de las cuales 7,451,690 residen en Bogotá (80.36%) y 1,821,496 en los 20 municipios del área metropolitana (19.64%): Bojacá, Cajicá, Chía, Choachí, Cota, El Rosal, La Calera, Facatativá, Funza, Gachancipá, Madrid, Mosquera, Sibaté, Soacha, Sopó, Subachoque, Tabio, Tenjo, Tocancipá y Zipaquirá.

Durante el desarrollo de la encuesta de Bogotá, se recolectó información primaria mediante la aplicación de un formulario físico previamente diseñado. Este proceso involucró visitas a los hogares de las personas y la interceptación de usuarios durante sus desplazamientos, lo que permitió indagar sobre los orígenes, destinos, motivos y otras características de sus viajes diarios.



Zona de estudio

Bogotá y los 20
municipios del área
metropolitana

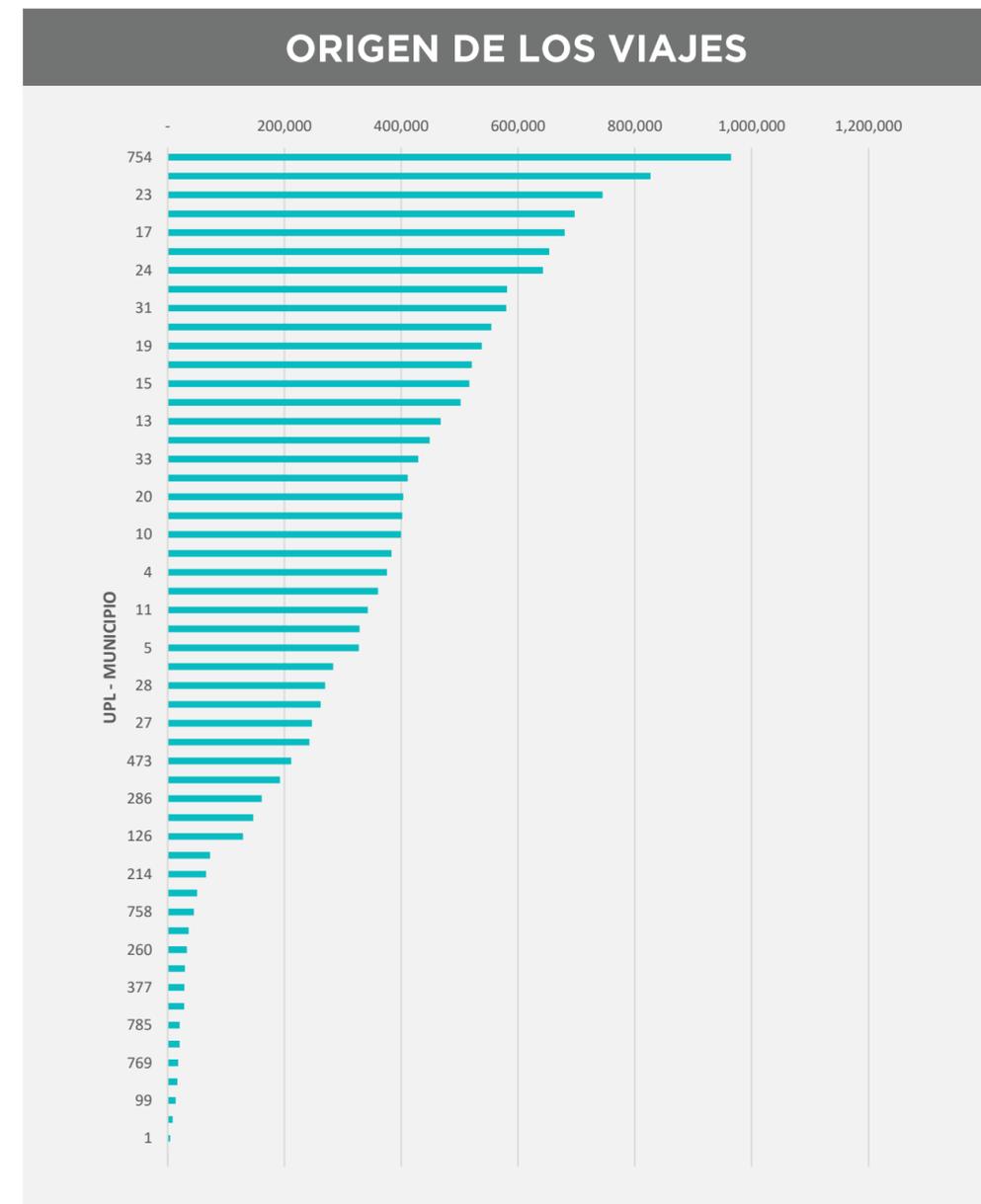
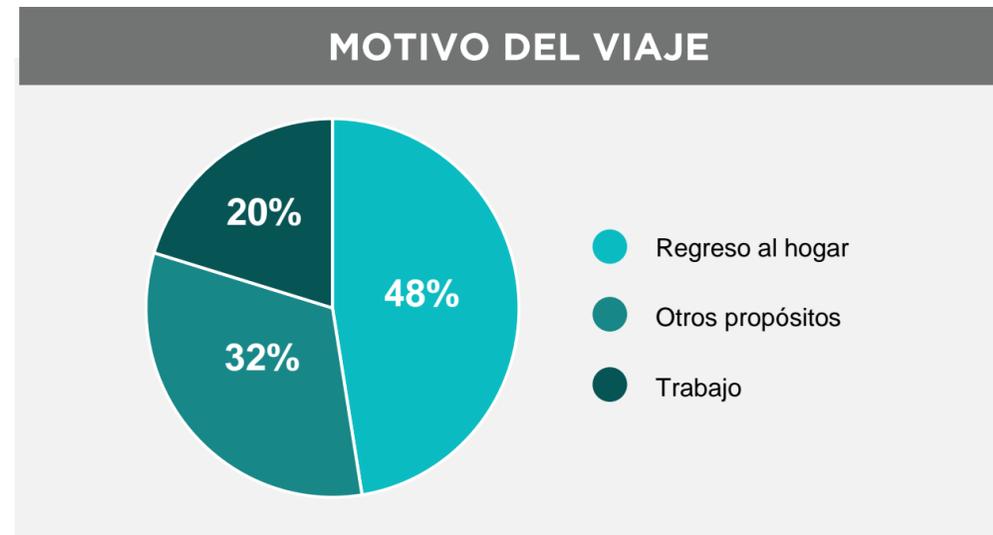


Encuestas directas

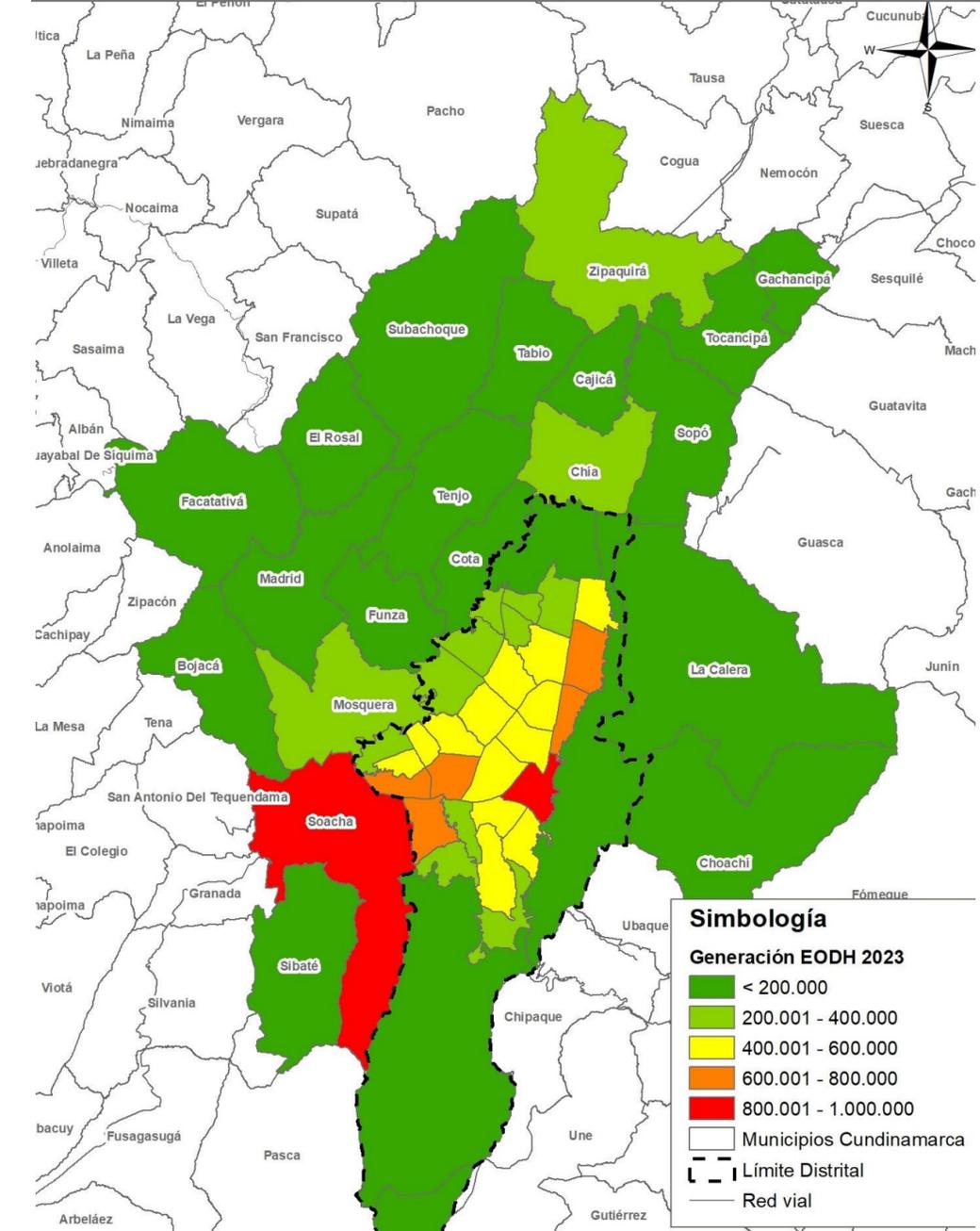
Resultados

Los resultados obtenidos por medio de la aplicación de más de 220 mil encuestas permitieron establecer que en el área de estudio se realizan diariamente 16,723,044 viajes. Asimismo, se determinó que el 48% de estos viajes diarios corresponde al motivo "regreso al hogar", seguido por el 32% con otros propósitos y el 20% relacionado con el trabajo.

A nivel espacial, este método identificó que los principales orígenes de los viajes diarios dentro del área de estudio se concentran en el municipio de Soacha (754) y las UPL 22 (Restrepo), 23 (Centro Histórico), 16 (Edén) y 17 (Bosa), que acumulan el 23% de la producción diaria.



El siguiente mapa, se muestra la **distribución espacial de los viajes diarios en la zona analizada**, obtenida por medio de los métodos convencionales de encuestas domiciliarias y de interceptación, observando alta concentración al sur y centro de la zona estudiada.



Big Data

KIDO DYNAMICS

Los datos de telefonía (24 horas de los días 15, 16 y 17 de agosto de 2023) utilizados para el estudio fueron proporcionados por Kido Dynamics. Fundada por un equipo español de científicos e ingenieros, Kido Dynamics es una empresa de Big Data y Machine Learning especializada en estudios de movilidad de la población mediante el análisis avanzado de los metadatos registrados por redes móviles.

La información se basa en el uso de datos de telefonía móvil, específicamente los conocidos como CDRs (Call Detail Records). Estos registros son almacenados para la facturación del servicio y contienen datos sobre llamadas telefónicas, mensajes SMS y conexiones a Internet. Cada registro proporciona información espacio-temporal del dispositivo, relacionando su posición con una zona de cobertura de la antena "celda". Dado que la información detallada de las celdas no siempre está disponible, se utiliza un nivel superior, la BTS (Base Transceiver Station), que agrupa varias celdas y cuya localización suele ser más accesible.

Los datos son registros anonimizados que incluyen llamadas de voz, tráfico de datos y registran la actividad de los usuarios. Los registros de voz incluyen llamadas y mensajes SMS, mientras que los registros de tráfico de datos contienen conexiones a Internet. Cada llamada proporciona dos localizaciones del usuario: una al inicio y otra al finalizar la llamada. En cambio, los mensajes SMS y el tráfico de datos solo proporcionan una localización por registro.

Un operador de telefonía también puede incluir información socio-demográfica sobre los usuarios, como edad, género, código postal de residencia, factura y tipo de contrato. Sin embargo, esta información no fue relevante para el cálculo de las matrices del estudio.

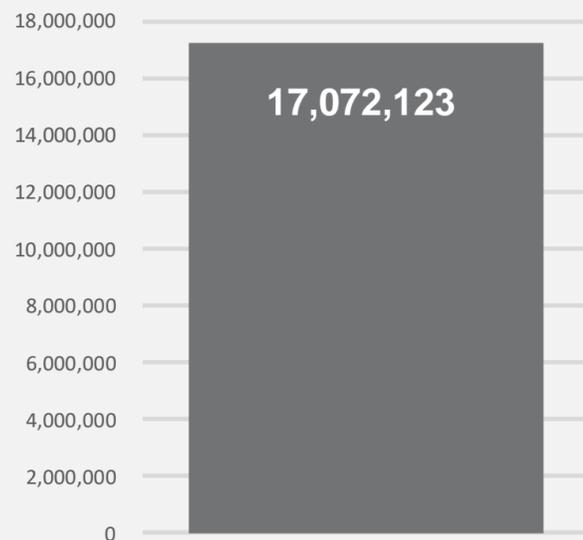


Big Data Resultados

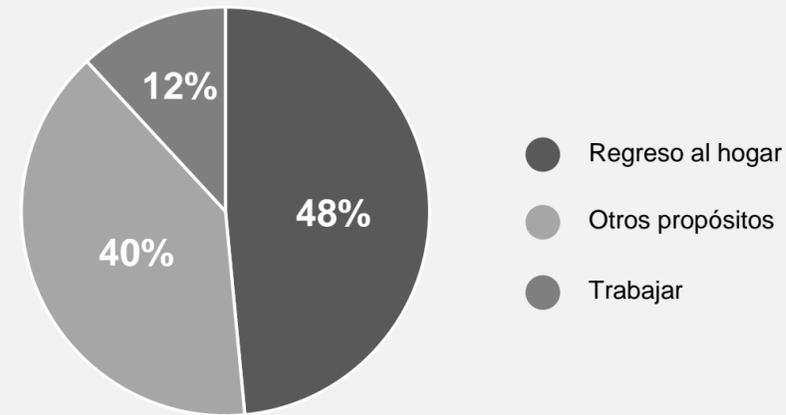
Los resultados obtenidos mediante el Big Data de telefonía celular indicaron que en el área de estudio se realizan diariamente 17,072,123 viajes. En cuanto a la distribución de los motivos de los viajes, esta metodología permitió establecer que el 48% se realiza por motivo de "regreso al hogar", seguido por el 40% con "otros" propósitos y el 12% con destino "a trabajar".

A nivel espacial, permitió identificar que los principales orígenes de los viajes diarios dentro del área de estudio se concentran en las UPL 22 (Restrepo), 30 (Salitre), 24 (Chapinero), 23 (Centro Histórico) y el municipio de Soacha (754), acumulando el 22% de la producción diaria.

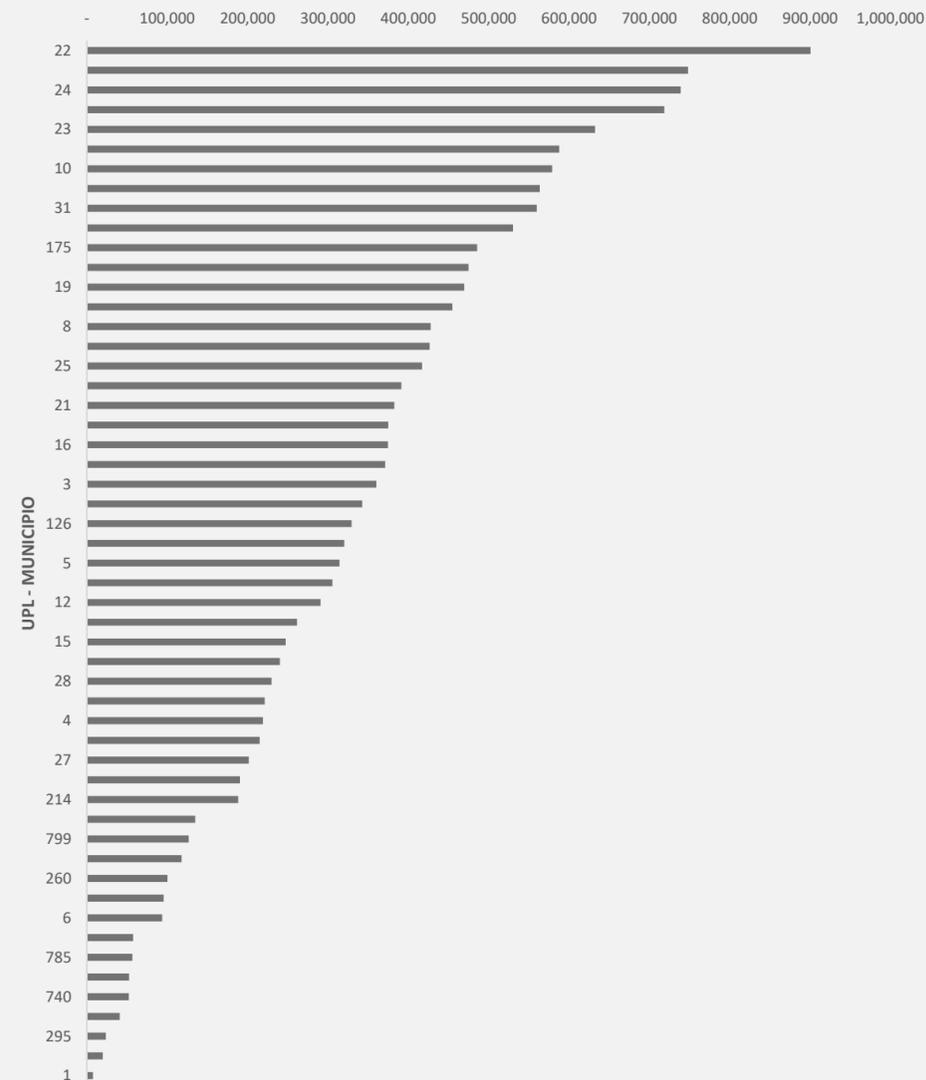
VIAJES DIARIOS



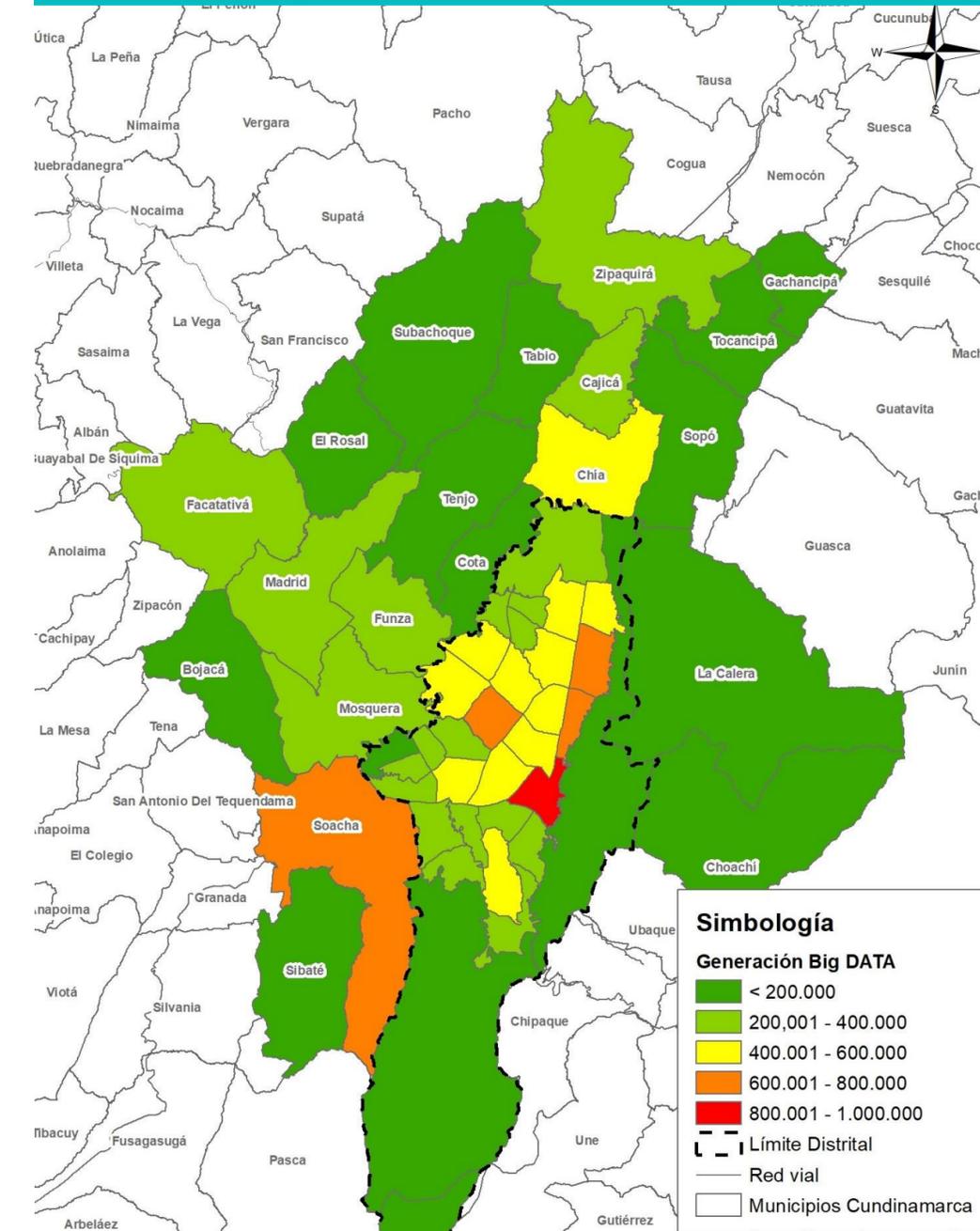
MOTIVO DEL VIAJE



ORIGEN DE LOS VIAJES



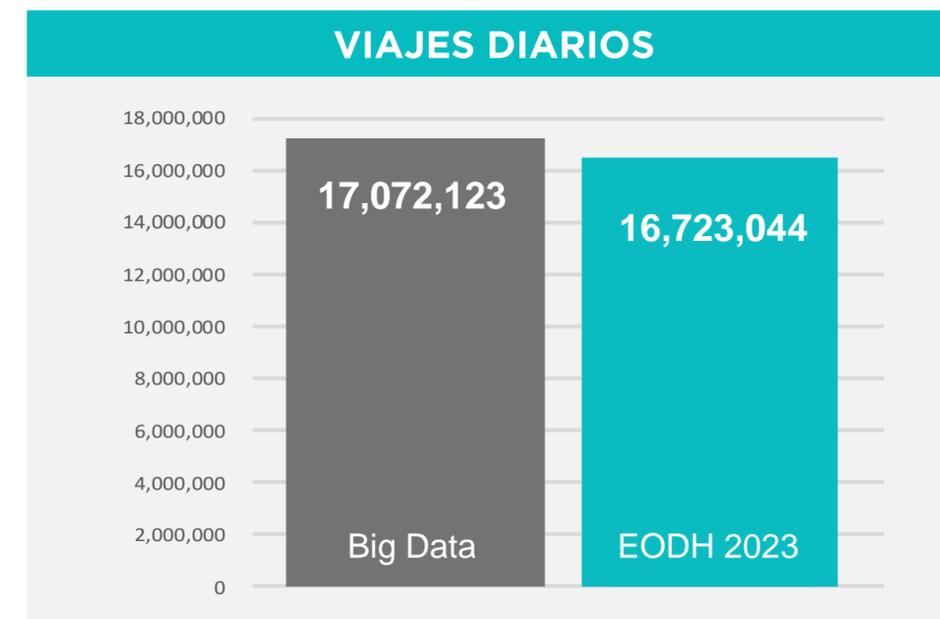
El siguiente mapa, muestra la **distribución espacial de los viajes diarios en la zona analizada**, obtenida por medio de este método alternativo, observando alta concentración al sur, occidente y centro de la zona estudiada.



Análisis comparativo

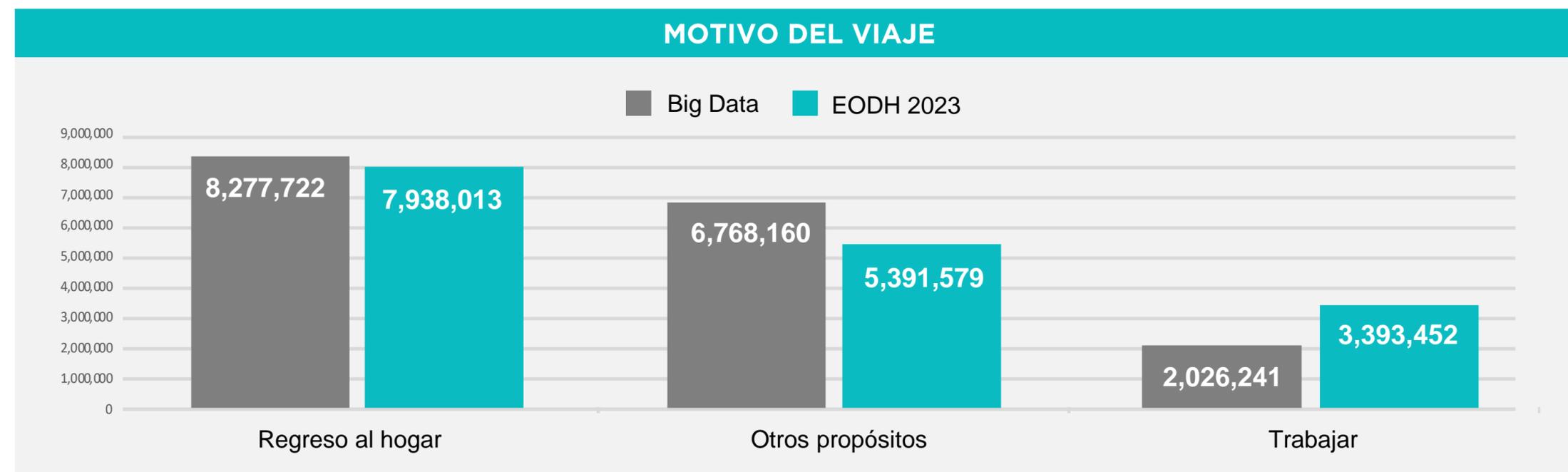
El ejercicio tuvo como objetivo principal, el realizar una comparación entre dos fuentes de información, con el fin de cuantificar las diferencias y aproximarse a la precisión de ambos métodos. En primer lugar, se comparó el total de viajes generados en la zona de estudio en un día típico, obtenidos mediante dos fuentes de información: EODI vs. Big Data.

A través de encuestas de interceptación y su correspondiente proceso de expansión de muestras, se estimó un valor de 16,7 millones de viajes diarios para Bogotá y los 20 municipios de estudio, mientras que la información de telefonía celular aproximó este dato a los 17 millones de viajes en la misma región. **En términos de viajes totales al día, la diferencia entre las dos fuentes de información fue de solamente el -2%.**



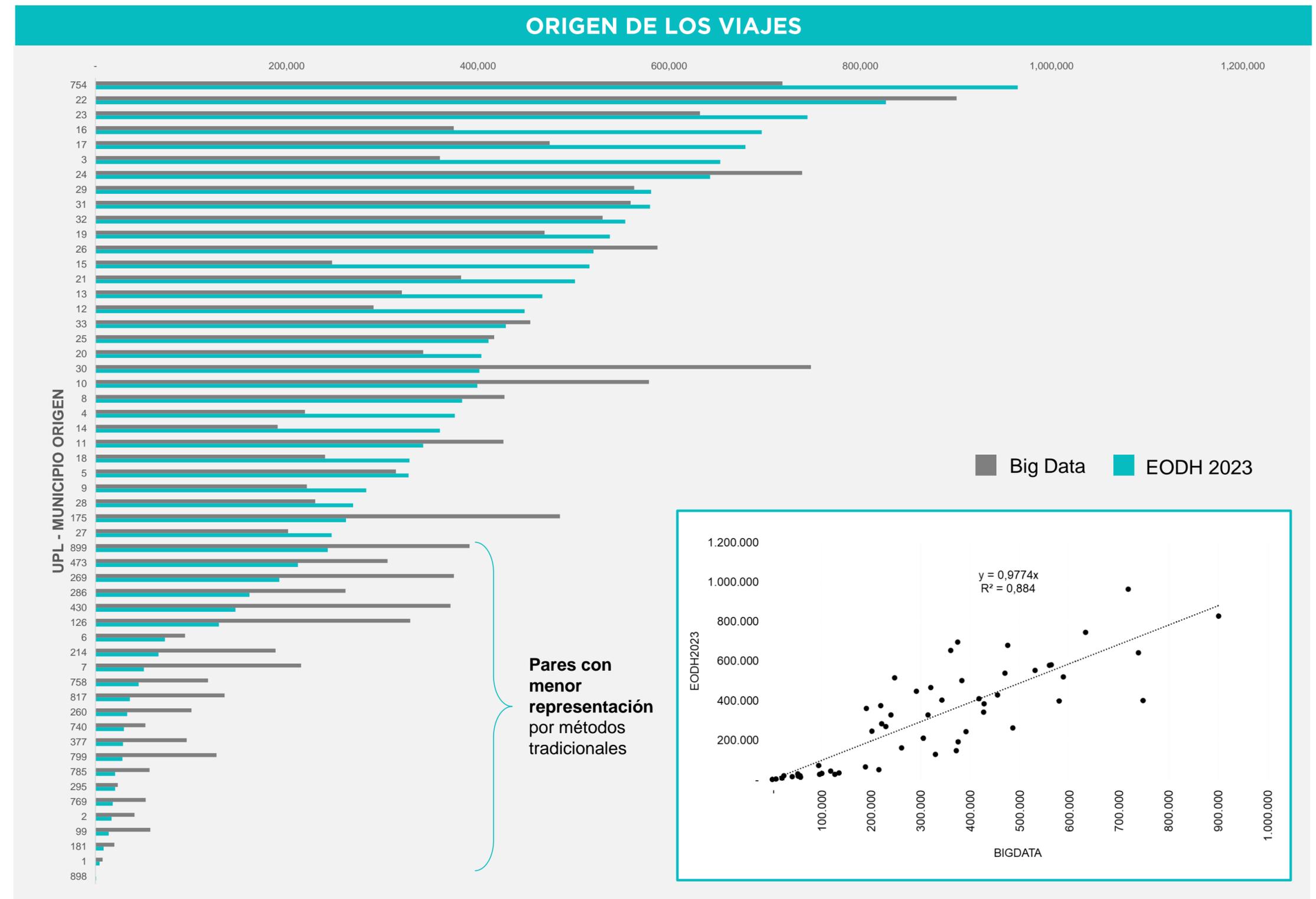
Otra comparación se realizó a nivel de motivo de los viajes diarios, donde la información de telefonía celular permitió establecer una aproximación a los motivos de los viajes bajo consideraciones relacionadas con las zonas de producción y atracción de los viajes. Las diferencias entre las fuentes de información son más acentuadas para un desglose mayor en la segmentación del mercado como el motivo (trabajo y otros), esto debido a las diferencias en los criterios de clasificación de viajes y el nivel de agregación de los motivos con datos de telefonía celular según el estado del arte de esta tecnología hasta la fecha.

Los viajes con motivo “regreso al hogar”, la diferencia entre las fuentes analizadas fue de solo el -4%. Sin embargo, el motivo “trabajo” presentó mayores diferencias, alcanzando hasta el 67% entre las fuentes, mientras que “otros motivos” tuvieron diferencias del 20%, reiterando que esto puede ser ajustado en función de los criterios de clasificación y agregación de los datos de telefonía.

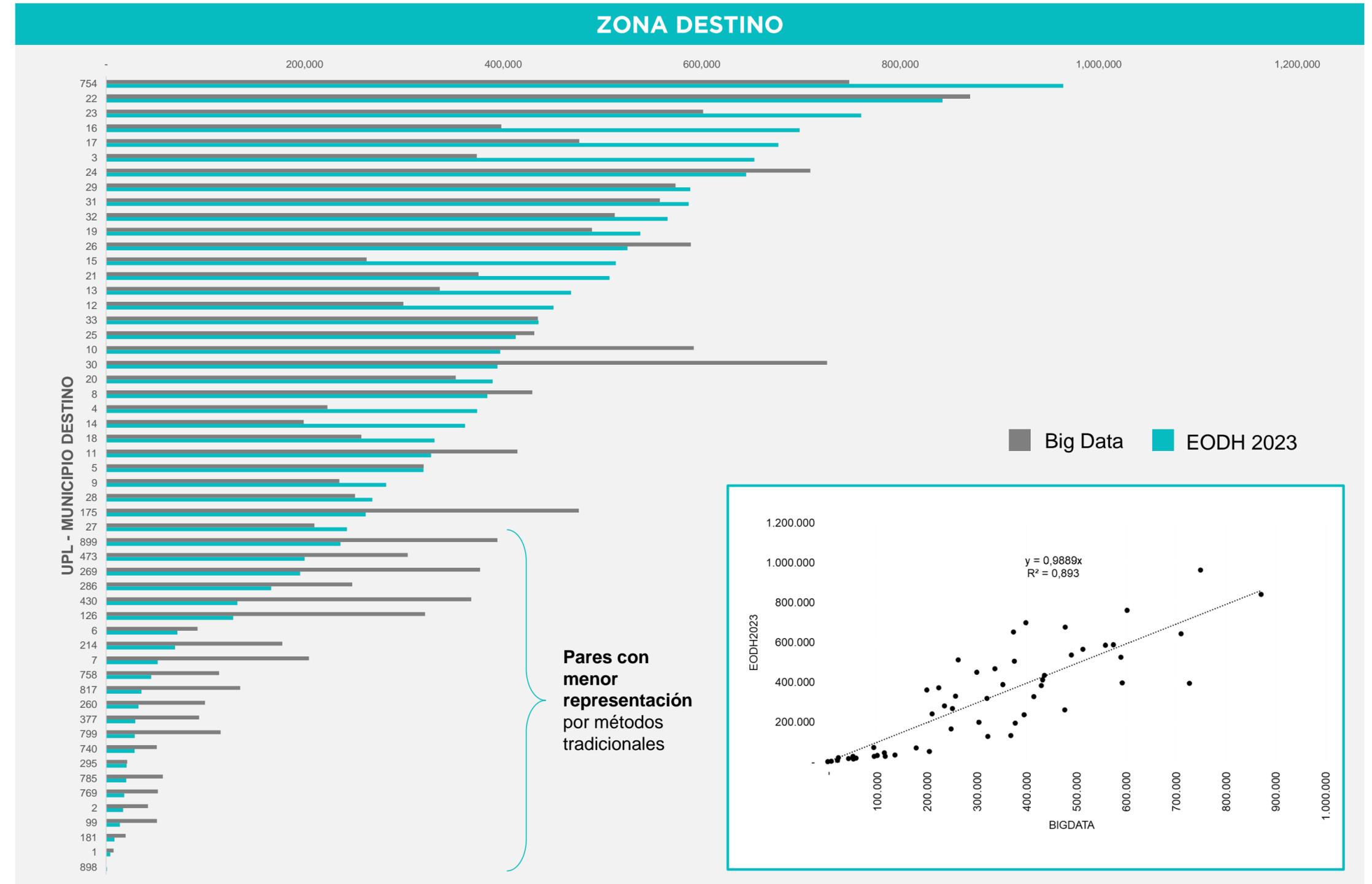


Al realizar la comparación a nivel de **zonas de origen de los viajes**, la distribución de los viajes diarios entre las dos fuentes de información por zona de origen mostró un patrón similar. La correlación entre los viajes producidos por las zonas analizadas fue significativamente alta entre las dos fuentes de información, con un R2 cercano a 1 y una pendiente de la función también cercana a 1.

Análisis comparativo



Similarmente, la distribución de los **viajes diarios por zona de destino** entre las dos fuentes de información mostró un patrón similar.



Análisis comparativo

Estos resultados demuestran una correlación fuerte entre las dos fuentes de información, lo que indica que ambas conducen a resultados comparativos en términos de patrones espaciales de producción y atracción de viajes. Sin embargo, se identificó una posible subestimación en los sectores de estratos bajos debido a un menor acceso a la telefonía celular, como en Soacha y Kennedy, donde la encuesta de hogares mostró una mayor concentración de viajes producidos.



Conclusiones

- Los resultados indican que la **información de Big Data es efectiva para representar los patrones espaciales y la concentración de los orígenes y destinos** de los viajes en la zona analizada durante el periodo de estudio (un día hábil).
- **Dado el carácter anónimo de los datos de telefonía se considera fundamental complementar por ahora esta información con métodos tradicionales para mayores niveles de segmentación del mercado por motivo, género, etc..**
- **Los datos de Big Data pueden reemplazar la aplicación de encuestas de interceptación y complementarse con aforos** y muestras pequeñas de hogares para fines de caracterización, lo que podría reducir la toma de información en un 30% a 50% en este tipo de estudios.
- En el caso de la Encuesta de Movilidad de Bogotá, el 95% de las matrices de modelación se basan en la encuesta de interceptación. La encuesta de hogares se utiliza para obtener tasas de generación de viajes a nivel del estrato socioeconómico de los hogares, lo que permitiría diseñar tamaños muestrales mucho más reducidos para caracterizar la población de estudio (aproximadamente el 20% de las 23 mil aplicadas).
- La Big Data a través de telefonía celular permite **caracterizar diferencias en los patrones de viajes a lo largo de semanas o meses** para mejores análisis de incertidumbre y **evaluaciones ex post** de eventos que modifiquen la dinámica de viajes en una zona de interés, además de brindar **mayor diversidad de pares origen-destino** en comparación con el método tradicional por sus limitaciones de muestreo.



Documento elaborado por nuestros consultores:



Miguel Andrés Castillo Rangel

Gerente de Planeación de Transporte
México y Centroamérica
macastillo@grupocalymayor.com

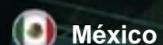


Juan Guillermo Ruiz Fonseca

Consultor Principal
Movilidad Urbana y Regional Suramérica
jrui@grupocalymayor.com



www.grupocalymayor.com



México

Canadá No. 110
Col. Parque San Andrés,
Coyoacán, C.P. 04040,
Ciudad de México
T. +52 (55) 5985 1050



Colombia

Carrera 19B # 84-17,
Oficina 704,
Bogotá
T. +57 (601) 718 0735



Perú

Calle Bolívar N° 472,
Edificio Business Club,
Torre 1 Oficina 902,
Miraflores Lima 18
T. +51 (1) 249 5703



Panamá

Edificio Ventura Office
Oficinas 303 y 304,
Vía Grecia, El Carmen,
Ciudad de Panamá
T. + (507) 381 3346 / 47



EE.UU.

Tollway Towers North,
Suite 870, 15770
North Dallas Parkway,
Dallas, Texas 75248
T. +1 (214) 245 5300



El Salvador

Calle a Santa Tecla
Residencial El Carmen
Pasaje Flor de Loto, Casa 9
Santa Tecla, La Libertad
T. +503 2278 1447

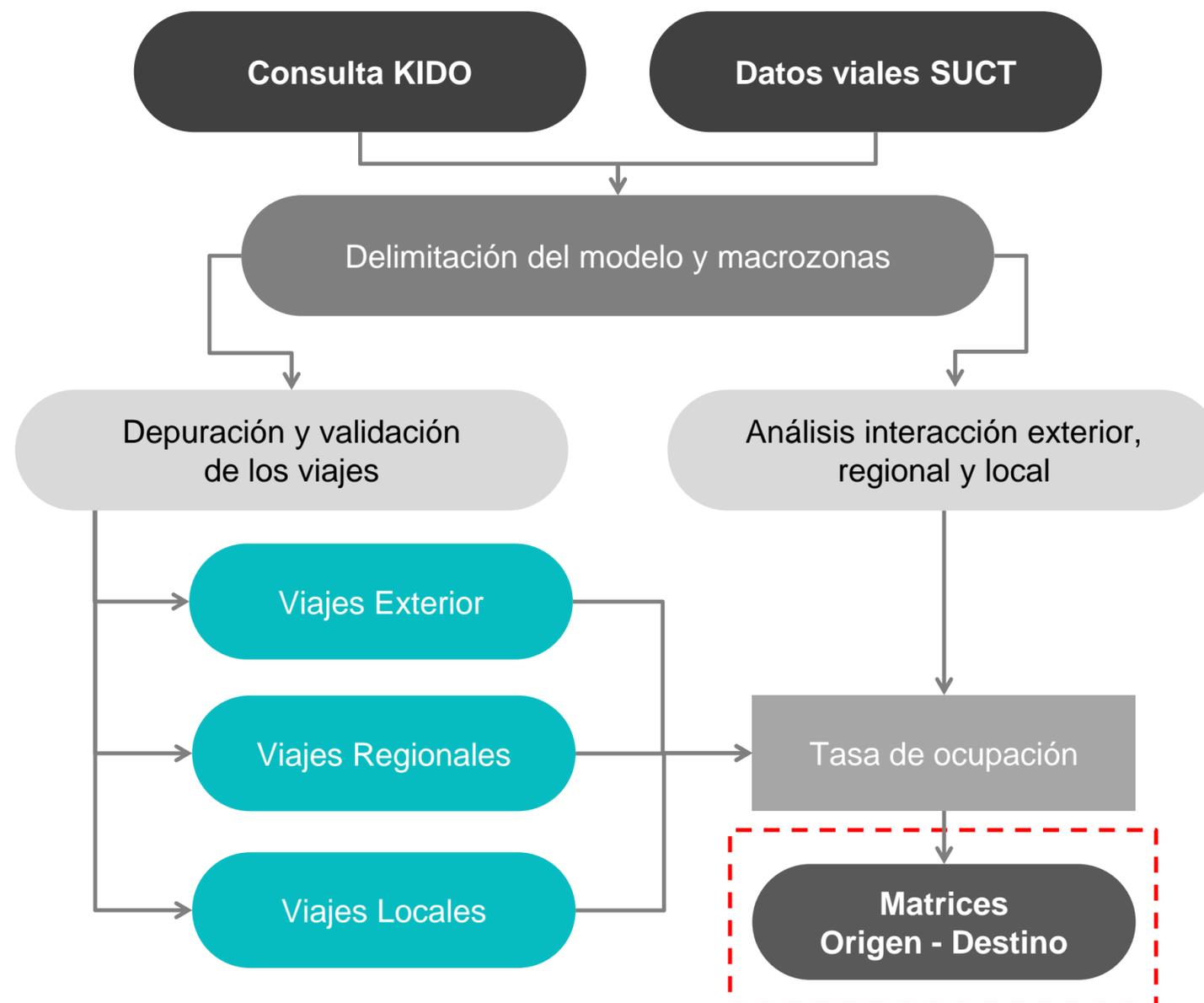


Bolivia

Urbanización Cañoto
Calle Los Mellizos N-4
entre radial 17 1/2 y Pitao
Santa Cruz de la Sierra
T. + 591 3321 3743

Anexo

Estrategia de complementariedad en la información con Big Data para estudios de carreteras



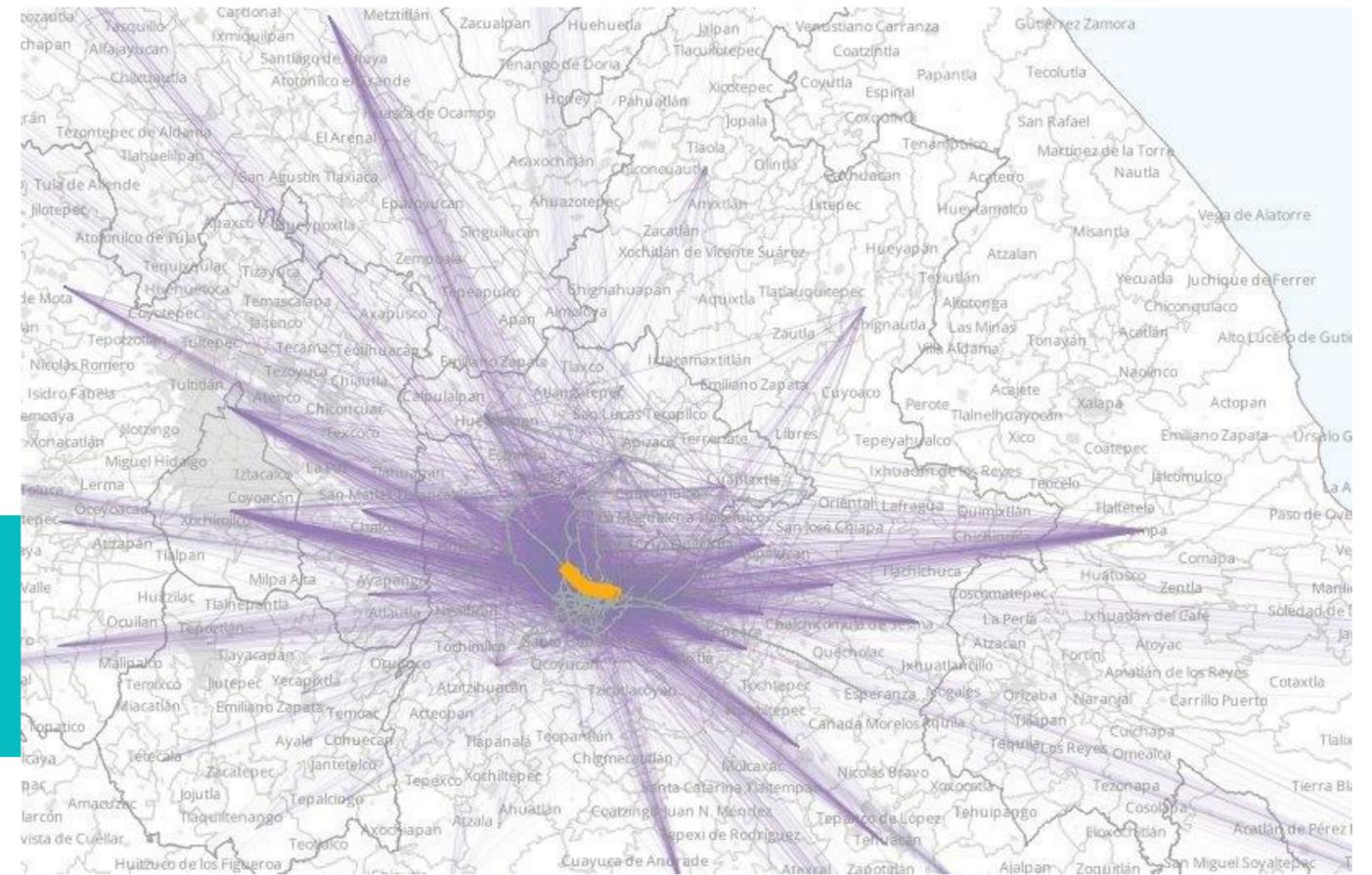
Consideraciones

- De la matriz KIDO, cada viaje representa un usuario que tiene un celular.
- Se realizan normalmente consultas de 7 días en la misma fecha de los trabajos de campo, por lo que es posible validar y complementar la información
- Se conocen el aforo y clasificación vehicular por Datos viales de SICT o aforos tomados en campo.
- Con la consulta de Big Data, se conoce la cantidad de viajes, pero no clasificación vial o el modo de transporte en el que se han realizado. Para ello se emplean los aforos y encuestas OD disponibles
- Se investigan las tasas de ocupación vehicular del banco de datos de Cal y Mayor y los publicados por las autoridades en la zona de influencia del proyecto.
- Con la tasa de ocupación determinada se puede obtener el valor de los viajes o usuarios para Datos viales y con la distribución vehicular de viajes o usuarios de datos viales y aplicarla a los observados en el Big Data.

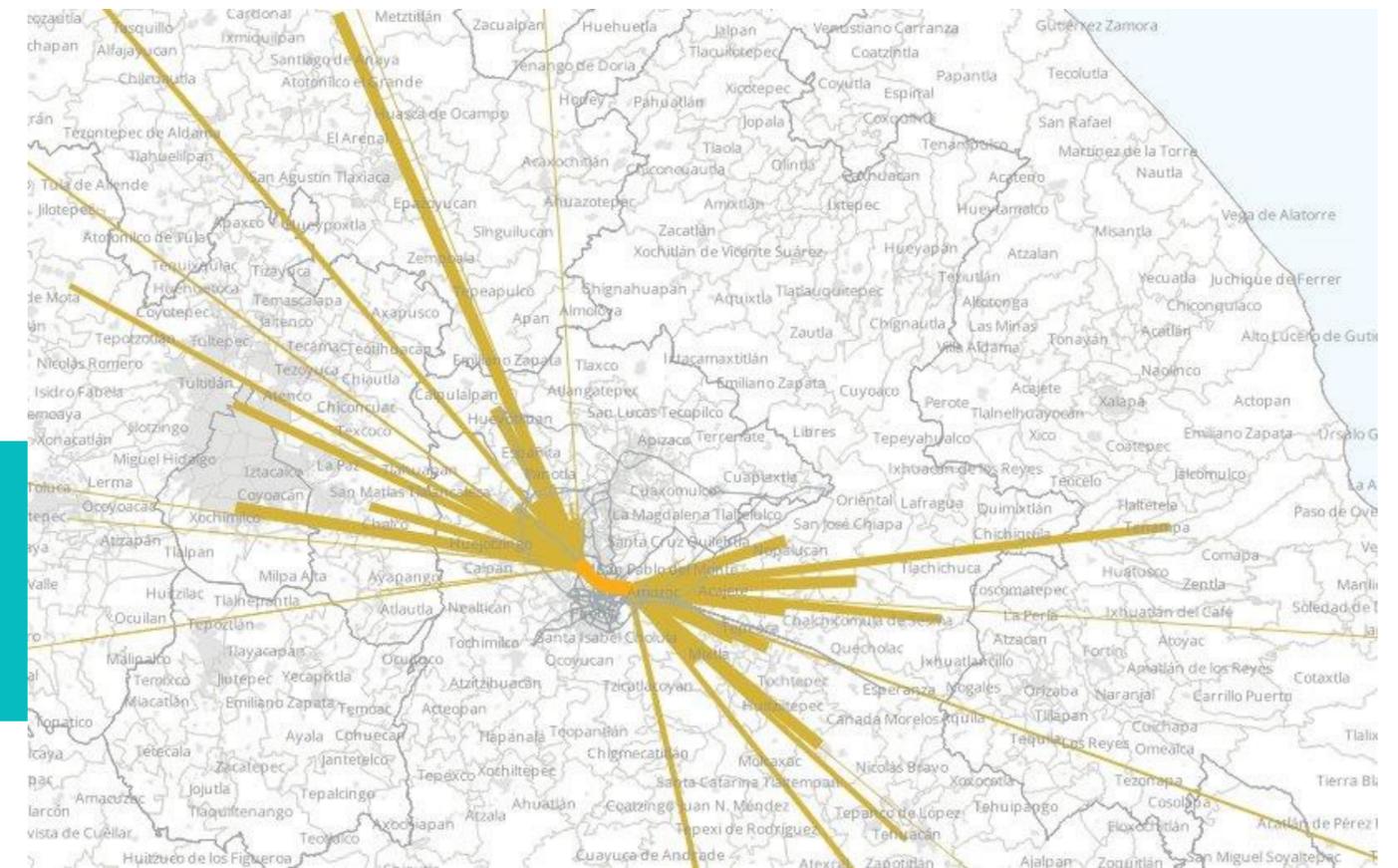


La información proveniente de Big Data permite tener una **mayor densidad y magnitud de viajes** con respecto a técnicas tradicionales

**PARES DE VIAJES
BIG DATA**



**PARES DE VIAJES
MÉTODOS TRADICIONALES**



**Comparativa
Big Data
con EOD
tradicional**



Conclusiones

- Usamos la información de Big Data principalmente para poder representar pares de viaje que **congestionen la red**.
- Identificar las **variaciones** en la magnitud y distribución de los viajes en **diferentes periodos del tiempo**.
- **Complementando** los métodos tradicionales de encuestas para **mayor certidumbre** en nuestras estimaciones.
- Además, en **análisis de visión de experto** tenemos un recurso valioso que complementa los aforos y matrices de nuestro banco de datos para orientar mejor nuestros análisis a ese nivel.
- **Próximos pasos:** clasificación por segmentos de mercado por IA, estimación de modelos PR, análisis de series de tiempo.

