



MIDIENDO LA ACCESIBILIDAD: UN CASO DE ESTUDIO EN PROGRAMAS DE ESTILO DE
VIDA
MEASURING ACCESSIBILITY: A CASE STUDY IN LIFESTYLE PROGRAMS

SUN YIN HUANG HUANG

Proyecto de grado para optar al título de Magíster en Ciencia de los Datos y Analítica

Asesores

Juan Pablo Ospina Zapata
Nicolás Guarín-Zapata

UNIVERSIDAD EAFIT

MAESTRÍA EN CIENCIAS DE LOS DATOS Y LA ANALÍTICA

MEDELLÍN

2024

CONTENIDO

1.	INTRODUCCIÓN.....	8
1.1.	PLANTEAMIENTO DEL PROBLEMA.....	8
1.2.	JUSTIFICACIÓN.....	10
2.	OBJETIVOS.....	11
2.1.	GENERAL.....	11
2.2.	ESPECÍFICOS.....	11
3.	MARCO TEÓRICO.....	12
4.	DISEÑO METODOLÓGICO.....	17
4.1.	MEDIDA DE ACCESIBILIDAD GRAVITACIONAL.....	17
4.2.	CLÚSTERIZACIÓN K-MEANS.....	19
4.2.1.	Definición del número de clúster.....	19
4.2.2.	Implementación del modelo.....	20
4.2.2.1.	Inicialización de los centroides.....	20
4.2.2.2.	Asignación de puntos a los clústers.....	20
4.2.2.3.	Actualización de centroides.....	20
4.2.2.4.	Iteración hasta la convergencia.....	21
5.	DATOS.....	22
5.1.	DESCRIPCIÓN.....	22
5.2.	CASO DE ESTUDIO.....	22
5.3.	DESCRIPCIÓN DE LAS VARIABLES.....	22
5.3.1.	Puntos de origen.....	23
5.3.2.	Puntos de destino.....	23
5.3.3.	Red de transporte.....	24

5.4.	ANÁLISIS EXPLORATORIO.....	25
5.4.1.	Definición de la muestra.....	28
5.4.2.	Cálculo de las distancias: Matriz de orígenes y destinos.....	30
6.	RESULTADOS.....	32
6.1.	MEDIDAS DE ACCESIBILIDAD.....	32
6.1.1.	Atractividad.....	32
6.1.2.	Función de impedancia.....	34
6.2.	DESDE EL DESTINO.....	35
6.3.	DESDE EL ORIGEN.....	38
6.4.	SEGMENTACIÓN K-MEANS.....	41
7.	CONCLUSIONES.....	45
	REFERENCIAS.....	47
	ANEXOS.....	53
	Anexo A. Usuarios que participan en los puntos de destino.....	53
	Anexo B. Segmentación de los usuarios por variable.....	56

LISTA DE TABLAS

Tabla 1. Matriz de componentes y perspectivas en la medición de accesibilidad	13
Tabla 2. Puntos de destino.....	24
Tabla 3. Cantidad de cursos totales y cursos con subocupación por punto de destino.	26
Tabla 4. Cantidad de cupos totales y cupos utilizados por punto de destino.....	27
Tabla 5. Resumen de participaciones por programa.....	28
Tabla 6. Cupos totales de los programas en la muestra.	32

LISTA DE FIGURAS

Figura 1. Puntos de origen, destino y malla vial.....	25
Figura 2. Distribución de las edades por programa.....	30
Figura 3. Distribución espacial de los puntos de origen y destino por programa.....	31
Figura 4. Atractividad de los programas.	33
Figura 5. Función de impedancia de los programas.	34
Figura 6. Accesibilidad desde el destino del estado actual y potencial del programa de hidroaeróbicos.....	36
Figura 7. Accesibilidad desde el destino del estado actual y potencial del programa de yoga.....	37
Figura 8. Accesibilidad desde el origen del estado actual y potencial del programa de hidroaeróbicos.....	38
Figura 9. Accesibilidad desde el origen del estado actual y potencial del programa de yoga.....	39
Figura 10. Usuarios con la mayor diferencia entre la accesibilidad potencial y actual.....	40
Figura 11. Clúster con variables de ubicación.....	42
Figura 12. Clúster con variables de ubicación, sociodemográficas y del programa.....	43

RESUMEN

El tiempo de viaje dedicado al desplazamiento es un gran condicionante que determina el acceso a las oportunidades y servicios de ciudad, especialmente, aquellos relacionados con programas de estilo de vida que además tienen un número limitado de cupos. Bajo esta idea, el acceso a los servicios depende de una combinación de elementos entre el costo de desplazamiento y la disponibilidad de estos. Aunque el desplazamiento pueda implicar un menor consumo de tiempo, esto no garantiza su disponibilidad para usuarios reales o potenciales. Asimismo, aunque exista una gran oferta, el tiempo de desplazamiento se convierte en una gran barrera para su uso. El objetivo de este estudio es analizar las condiciones actuales y potenciales de acceso hacia una variedad de equipamientos y servicios relacionados con el estilo de vida, principalmente, los programas de hidroaeróbicos y yoga. Para lograr esto, se utilizarán medidas de accesibilidad y cobertura geográfica que tienen en cuenta la ubicación de los usuarios en el territorio, la distribución de las instalaciones, las redes de transporte público, así como, los costos expresados en tiempo o distancia y cupos óptimos. Adicionalmente, se complementa con un modelo K-Means para el análisis de los usuarios que utilizan cada programa para la identificación de públicos potenciales. Los resultados sugieren un patrón centro-periferia tanto desde el punto de vista del origen como del destino. Mientras que los mayores niveles de accesibilidad en hidroaeróbicos se concentran hacia el norte de Medellín, para yoga se desplaza hacia el centro y sur de la ciudad, lugares donde también se ubican la mayor cantidad de puntos de destino. Por otro lado, aunque no se encuentran patrones espaciales en cuanto a las variables sociodemográficas, los usuarios de hidroaeróbicos en comparación con yoga, que suelen tener un público más diverso, se caracterizan por ser de más edad y ser pensionados o independientes con una alta fidelización al programa.

Palabras claves: Accesibilidad, Programas de estilo de vida, Redes de transporte, Optimización de espacios, Segmentación de usuarios

ABSTRACT

The travel time dedicated to commuting is a major factor that determines access to city opportunities and services, especially those related to lifestyle programs that also have a limited number of places. Under this idea, access to services depends on a combination between the cost of travel and their availability. Although movement may imply less time consumption, this does not guarantee its availability for real or potential users. Likewise, although there is a great offer, travel time becomes a great barrier to its use. The objective of this study is to analyze the current and potential conditions of access to a variety of equipment and services related to lifestyle, mainly hydroaerobics and yoga programs. To achieve this, accessibility and geographic coverage measures will be used that consider the location of users in the territory, the distribution of facilities, public transportation networks, as well as costs expressed in time or distance and optimal quotas. Additionally, it is complemented with a K-Means model for the analysis of users who use each program to identify potential audiences. The results suggest a center-periphery pattern from both the point of view of origin and destination. While the highest levels of accessibility in hydroaerobics are concentrated towards the north of Medellín, for yoga it moves towards the center and south of the city, places where the largest number of destination points are also located. On the other hand, although no spatial patterns are found in terms of sociodemographic variables, users of hydroaerobics compared to yoga, who tend to have a more diverse audience, are characterized by being older and retired or independent with high loyalty to the program.

Keywords: Accessibility, Lifestyle programs, Transportation networks, Optimization, User segmentation

1. INTRODUCCIÓN

La noción convencional del tiempo libre o de ocio, arraigada usualmente al sentido común, se caracteriza por la concepción de un tiempo que no es trabajo y que a su vez se relaciona con diversión (Stebbins, 2018). No obstante, a lo largo de los últimos años se han presenciado cambios significativos en la cantidad, experiencia y uso de este, como consecuencia de la reducción de la jornada laboral y el aumento en la aceleración del ritmo de vida (Haller et al., 2013) asociado con niveles más altos de estrés y aburrimiento (Dattilo et al., 2012), lo que ha llevado a plantearse una definición más amplia de este concepto involucrando el bienestar y la calidad de vida de las personas (Fancourt et al., 2021).

En los últimos años se ha evidenciado un aumento significativo en la popularidad y oferta de programas para la gestión del tiempo libre, especialmente, aquellos dirigidos a un público adulto, lo que ha puesto en relieve la necesidad de examinar sistemáticamente dicha oferta (Dattilo et al., 2012). En particular, en la literatura se han señalado cuatro fuentes de aburrimiento que pueden presentarse en la cotidianidad: la falta de participación significativa, la escasez de estimulación mental, la percepción de tiempo lento y la falta de actividad física (Ragheb & Merydith, 2001), por otra parte, diversos estudios han confirmado que periodos llenos de actividades emocionantes, mejoran la percepción del tiempo en comparación con los momentos ocupados con actividades o experiencias monótonas asociadas al aburrimiento (Danckert & Allman, 2005). Los programas relacionados al estilo de vida surgen para solventar esta necesidad donde además se convierten en piezas claves para el desarrollo de hábitos saludables, tanto físicos como mentales, lo que, a su vez, se traduce en mecanismos efectivos para el aumento del bienestar y la calidad de vida (Wang et al., 2016).

1.1. PLANTEAMIENTO DEL PROBLEMA

El uso de programas relacionados con el estilo de vida es una práctica común en la sociedad que abarca una amplia gama de actividades como los ejercicios físicos, educación continua y entretenimiento cultural, sin embargo, su implementación se realiza dentro de períodos de tiempo reducidos y limitados, generalmente asociados con el tiempo libre disponible. Esta limitación plantea desafíos adicionales, especialmente en términos de accesibilidad y el costo de utilizarlos (asequibilidad). Por lo tanto, más allá de aspectos meramente financieros, el tiempo necesario y la distancia en el desplazamiento adquieren igualmente una gran relevancia. Bajo esta perspectiva, el costo de acceder a los programas, el cual puede ser expresado en términos de dinero, tiempo, distancia, entre otros, es un elemento que condiciona su uso (Wild et al., 2020).

Dentro de los costos asociados, el tiempo de desplazamiento, que relaciona la distancia y el consumo de tiempo, es una de las principales limitaciones a tener en cuenta debido a que no solamente se depende de la dedicación para el desarrollo del programa sino también del tiempo necesario para desplazarse hacia y desde el lugar donde se lleva a cabo la actividad —o desde el origen hasta el destino—. Adicionalmente, estas decisiones están relacionadas directamente con un medio de transporte que facilite su acceso y viabilidad al servicio. El término accesibilidad se puede definir como la facilidad o dificultad que tienen las personas para alcanzar las oportunidades o servicios (Wachs & Kumagai, 1973), proporcionando así una gran relevancia a la red de transporte debido a que justamente están diseñados para ayudar a las personas a participar en actividades distribuidas en el espacio y el tiempo (El-Geneidy & Levinson, 2006). Este tipo de planteamientos ha permitido el desarrollo de diversos conceptos de organización del sistema urbano, en términos de tiempos, espacios y redes de transporte. Adicionalmente, en los últimos años se ha observado un nuevo protagonismo por las tendencias relacionadas a los modos de transporte activos como los caminos peatonales (redes camineras) y las ciclo rutas (Gaglione et al., 2022). Es así, como han surgido diversos enfoques tales como la ciudad en 15 minutos, que se introdujo como una estrategia post-COVID para apoyar una recuperación más sostenible de la pandemia y desarrollar vecindarios completos, resilientes y sensibles al clima, cuyo objetivo es promover la descentralización de las funciones urbanas para garantizar un acceso justo a los servicios urbanos bajo el tiempo establecido (Khavarian-Garmsir et al., 2023).

Adicionalmente, estos programas suelen tener una oferta limitada en términos de horarios y espacios específicos, lo que está condicionado por la infraestructura requerida, además, de la competencia con otros programas que también requieren de esa infraestructura o que están dirigidas al mismo público objetivo. Esto plantea restricciones en la disponibilidad que se manifiestan en cupos limitados, lo cual también influye en su atractivo para los usuarios dado que una menor cantidad de cupos disponibles puede desmotivar la inscripción de usuarios a causa de la percepción de que estos posiblemente se ocupen en un menor tiempo. En ese mismo sentido, es frecuente que estos programas se enfrenten al desafío de equilibrar la oferta con la demanda de usuarios. Así, de un lado, puede que la oferta supere la demanda, lo que implica que los recursos sean subutilizados y se desaproveche su máximo potencial, pero también que se presente una demanda que sobrepase la oferta, generando que la cobertura de los espacios y servicios disponibles no sean suficientes para las personas interesadas.

Teniendo en cuenta lo anterior, este estudio se plantea responder ¿cómo hacer que los programas de estilo de vida sean más accesibles a las comunidades a las que sirve? En este contexto, la medida de accesibilidad se compone de tres elementos, (i) los puntos de origen asociados a la vivienda, (ii) los puntos de destinos asociados al lugar en el que se presta el curso y (iii) la malla vial correspondiente al nodo de transporte, de los cuales solo se podría tener incidencia en los puntos de destino que pueden ser estratégicamente dispuestos para mejorar la accesibilidad a través de cambios en los cursos ofrecidos, disposición de nuevas sedes, optimización del portafolio, modificaciones en los horarios, entre otros.

1.2. JUSTIFICACIÓN

Los resultados obtenidos a partir de esta investigación tienen el potencial de ofrecer valiosa información en un campo que se caracteriza por sus limitaciones, especialmente debido a la escasez de datos georreferenciados tanto del punto de origen como del punto de destino. En particular, Jiang (2017) menciona que la forma de medir la accesibilidad al transporte en diferentes áreas se ha limitado a la disponibilidad de datos y a la tecnología, no obstante, estos estudios han proporcionado una base sólida para la planificación y mejora de la prestación de servicios, lo que a su vez contribuye a la optimización de los espacios disponibles. Así mismo, estos hallazgos pueden ser útiles para ofrecer diferentes tipos de experiencias de aprendizaje de ocio para los adultos (Dattilo et al., 2012).

En consecuencia, este análisis buscar facilitar la toma de decisiones informadas en beneficio de la comunidad a la que sirven estos programas. Esta comprensión más profunda de los patrones de uso de programas y servicios, así como, de las necesidades, puede orientar la optimización del portafolio de servicios, además, de un mejor emparejamiento entre los puntos de origen y destino. Por ende, esto puede mejorar significativamente la participación de los usuarios al garantizar el ingreso a los programas que mejor se adapten a sus necesidades, condiciones e intereses bajo una premisa de mayor accesibilidad. Finalmente, el uso de los programas de estilo de vida ha demostrado ser eficaces en la promoción de hábitos saludables lo que ha destacado la necesidad de garantizar un acceso equitativo a estos recursos. Esta equidad en el acceso no solo es fundamental desde una perspectiva ética, sino que también es vital para maximizar el impacto positivo de dichos programas en la población, por lo que, la consideración y acción hacia un acceso equitativo se convierte en un componente esencial en la planificación y ejecución de estos.

2. OBJETIVOS

2.1. GENERAL

Analizar y evaluar el alcance y el acceso a los programas de estilo de vida, mediante la utilización de mediciones de accesibilidad, cobertura geográfica y segmentación de públicos.

2.2. ESPECÍFICOS

A lo largo de este trabajo, se cumplirán dos objetivos específicos, que se definen dentro del estudio como momentos, el estado actual y el estado potencial.

- En el *estado actual* se plantea analizar la accesibilidad de los programas a los usuarios que actualmente hacen uso.
- En el *estado potencial* se plantea analizar la accesibilidad y segmentación de los individuos potencialmente usuarios de los programas.

3. MARCO TEÓRICO

Diversos estudios mencionan que los tiempos de desplazamientos prolongados están relacionados con niveles más bajos de satisfacción con la vida y felicidad (Nie & Sousa-Poza, 2018; Lorenz, 2018) y el desarrollo de patrones de comportamiento que pueden contribuir en la disminución en las condiciones de salud física y mental debido a los altos niveles de fatiga y estrés (Gimenez-Nadal & Molina, 2019; Gottholmseder et al., 2009), así como con el desarrollo de enfermedades como la obesidad (Christian, 2012). De igual manera, los tiempos de viaje más largos han sido asociados negativamente con la dedicación a la actividad física (Hilbrecht et al., 2014), la duración y satisfacción con el ocio (Stutzer & Frey, 2008; Mouratidis, 2019) y la creación de restricciones que pueden impedir que los individuos realicen ciertas actividades o que las realicen con la frecuencia deseada (Crawford et al., 1991).

Siguiendo la teoría de la demanda de viajes, cuya premisa menciona que el tiempo de viaje se percibe como un bien que las personas buscan preservar, se subraya la idea de que el ahorro en este tiempo se puede transformar (Recker et al., 2001). Por un lado, en algo de valor intrínseco como un mayor tiempo dedicado a realizar actividades de valor económico u otro tipo, en este caso, el desplazamiento se percibe como una inversión en tiempo y las decisiones sobre la movilidad se basan en la evaluación de los costos. Así, la reducción del tiempo de viaje se convierte en un aspecto clave no solo para mejorar la satisfacción en la percepción de estos viajes (Mouratidis et al., 2019), sino también en mejorar la satisfacción con las actividades de ocio (Mouratidis, 2019). Por otro lado, el ahorro del tiempo de viaje se puede transformar en el aumento del espacio de captura de ubicaciones alternativas de las instalaciones para tales actividades donde el ahorro de tiempo de viaje se mapea en una medida espacial de accesibilidad.

El concepto de accesibilidad surge a principios del siglo XX y ha tenido diversos desarrollos que muestran una amplia gama de mediciones y aplicaciones que abarcan diversas problemáticas y escalas espaciales (Harris, 2001). Gran parte de estos estudios se han concentrado en temas de urbanismo (Scott & Horner, 2024), planificación en los sistemas de transporte (Allen & Farber, 2020), acceso a las oportunidades laborales (Cheng & Bertolini, 2013) y acceso al sistema y servicios de salud (F. Wang, 2012). En la literatura se define como el potencial para alcanzar oportunidades distribuidas espacialmente y se considera como uno de los logros más destacados del desarrollo espacial, resultado de la combinación de un sistema de transporte y la distribución geográfica de las actividades (Páez et al., 2012). Adicionalmente, la accesibilidad ha sido identificada como un indicador potencial de exclusión social (Ribeiro et al., 2021), ya que las dificultades para alcanzar oportunidades desde un lugar determinado pueden reforzar dicha exclusión (Church et al., 2000; Kenyon et al., 2002). Asimismo, ha

permitido la identificación de grupos sociales y ubicaciones con bajos niveles de acceso a servicios e instalaciones que pueden limitar la capacidad de un individuo para participar plenamente en la sociedad (Achuthan et al., 2010).

Siguiendo a Geurs & Van Wee (2004) las medidas de accesibilidad se pueden componer de cuatro componentes, (i) el uso del suelo asociado a la cantidad, calidad y distribución espacial de las oportunidades, a la demanda de estas oportunidades y a la confrontación entre la oferta y demanda de estas oportunidades, (ii) el transporte asociado al sistema o red de transporte que utiliza un individuo para cubrir la distancia hasta la oportunidad, es decir, desde el punto de origen, usualmente relacionado a la ubicación del individuo, hasta el punto de destino que es el lugar donde se encuentra la oportunidad, (iii) la temporalidad asociada a las limitaciones temporales, es particular, la disponibilidad de las oportunidades y el tiempo disponible de los individuos y (iv) la componente del individuo que refleja sus necesidades y capacidades.

De igual manera, se pueden identificar cuatro perspectivas básicas sobre la medición de accesibilidad. En primer lugar, las medidas basadas en infraestructuras que buscan analizar el rendimiento o el nivel de servicio de la infraestructura de transporte (Mondschein & Taylor, 2017). En segundo lugar, las medidas basadas en la ubicación, que describen el nivel de accesibilidad a las oportunidades distribuidas espacialmente lo que implica mediciones de separación espacial entre los individuos y las actividades (Chen et al., 2024). De la misma manera, durante los últimos desarrollos se han propuesto medidas basadas en el individuo (Horner & Downs, 2014), que buscan analizar las actividades en las que un individuo puede participar en un momento determinado (Kwan et al., 2003). Y finalmente, las medidas basadas en la utilidad que analizan los beneficios, usualmente económicos, que los individuos obtienen del acceso a las oportunidades distribuidas espacialmente (Guzman et al., 2023).

Tabla 1. Matriz de componentes y perspectivas en la medición de accesibilidad.

Perspectiva	Componentes			
	Uso de la tierra	Transporte	Temporalidad	Individuo
Infraestructura	Velocidad en el desplazamiento, Congestión		Periodo de hora pico, Periodo de 24 horas	Estratificación basada en viajes
Ubicación	Tiempo de desplazamiento/Costos entre ubicaciones de las oportunidades	Cantidad y distribución espacial de la demanda y/u oferta de oportunidades	Tiempo de desplazamiento y costos asociados a la temporalidad	Estratificación de la población

Individuo	Tiempo de desplazamiento/Costos asociados entre ubicaciones de las oportunidades	Cantidad y distribución espacial de las oportunidades ofrecidas	Limitaciones temporales y tiempo disponible para las actividades	Accesibilidad analizada a nivel individual
Utilidad	Costos de desplazamiento entre ubicaciones de las oportunidades	Cantidad y distribución espacial de las oportunidades ofrecidas	Tiempo de desplazamiento y costos asociados a la temporalidad	Utilidad es derivada a nivel individual o de un grupo poblacional homogéneo

Fuente: Traducido y tomado de Geurs y Van Wee (2004)

En particular, en la literatura se han identificado dos grupos para las medidas basadas en la ubicación. Estos corresponden a las medidas de contorno y a las medidas gravitacionales (Geurs & Van Wee, 2004; El-Geneidy & Levinson, 2006). En el caso de las medidas de contorno, se cuenta el número de oportunidades potenciales que se pueden alcanzar dentro de un tiempo, distancia o costo de viaje determinado, y aunque se utiliza ampliamente debido a que tiende a ser más fácil de entender e interpretar, estas medidas no tienen en cuenta las elecciones reales de las personas en cuanto a los puntos de origen y destino, y ponderan por igual a las personas dentro del mismo intervalo de tiempo de viaje sin considerar el atractivo ni las dificultades para llegar (El-Geneidy & Levinson, 2006). Las medidas gravitacionales, por su parte, estiman la accesibilidad de las oportunidades de un punto de origen a todos los puntos de destino siguiendo un nodo de transporte y cumpliendo dos premisas, en primer lugar, que disminuya cuanto más alejadas estén las oportunidades del origen también denominada función de impedancia y que aumente con el aumento en la medida de las oportunidades o atractividad (Geurs & Ritsema van Eck, 2001).

Esto permite resultados que varían en función de la distancia y la importancia de los destinos, proporcionando una evaluación más matizada que considera tanto la proximidad como el atractivo de los destinos. La función de impedancia, al ponderar estas variables subraya la relevancia de los nodos de transporte que están diseñados para facilitar el acceso de las personas a diversas actividades distribuidas en el espacio y el tiempo (El-Geneidy & Levinson, 2006). Estos nodos no solo reducen los tiempos de desplazamiento, sino que también aumentan la conectividad entre diferentes oportunidades, mejorando la accesibilidad en general. No obstante, la reducción en los tiempos en el desplazamiento no constituye una garantía de que los servicios requeridos estén disponibles en la ubicación específica. Es importante reconocer que, en la prestación de servicios, especialmente aquellos programas relacionados con la

cultura, el entretenimiento y la educación, se tienen restricciones y limitaciones que pueden afectar significativamente la disponibilidad para el público tales como la existencia de un número limitado de cupos, la capacidad en la infraestructura y/o la experiencia en el servicio, que actúan como proxies de la atractividad. Siguiendo a Kelly et al., quienes realizaron una revisión sistemática, la falta de disponibilidad y/o acceso a los programas, se convierte en una barrera para la adopción y el mantenimiento de comportamientos saludables por parte de las personas de mediana edad (2016).

Por otro lado, estas medidas de accesibilidad pueden emplearse para identificar grupos de áreas con características de accesibilidad similares, lo que facilita la toma de decisiones orientadas a mejorar la accesibilidad (Geurs & Van Wee, 2004), sin embargo, este enfoque es insuficiente en la identificación adecuada de públicos potenciales. La elección en el uso de los programas de estilo de vida también se ven permeados por una serie de variables que incluyen factores socioeconómicos, preferencias personales, gustos individuales y comportamientos o hábitos (Van Acker, 2015), por lo que para tener un mejor panorama es necesario complementar las medidas de accesibilidad de los programas con información de los públicos que los utilizan.

La clústerización se refiere a un conjunto de modelos no supervisados utilizados en el análisis de datos, que permiten dividir una población heterogénea en grupos más pequeños y homogéneos (Naeem et al., 2023). Estos modelos no requieren un conjunto de datos etiquetados previamente, lo que significa que no dependen de resultados conocidos para aprender a realizar la segmentación. En lugar de esto, los modelos no supervisados identifican patrones y estructuras en los datos basándose únicamente en las características intrínsecas de los datos mismos (Ghahramani, 2004). Siguiendo a Jain (2010) este agrupamiento de los datos se ha utilizado usualmente para tres propósitos: (i) obtener información de la estructura subyacente de los datos para generar hipótesis, detectar anomalías e identificar características destacadas, (ii) para identificar el grado de similitud entre formas u organismos y (iii) como método para organizar los datos y resumirlos a través de clúster.

En la literatura se pueden categorizar los algoritmos de agrupamiento en dos grupos: jerárquicos y particionales. Los algoritmos de agrupamiento jerárquico dividen los objetos de datos en grupos en forma jerárquica, ya sea en un enfoque ascendente conocido como método aglomerativo o de arriba hacia abajo como método divisivo (Ikotun et al., 2023). En estos se destacan los de enlace único y enlace completo, que miden la distancia entre dos clústeres como la distancia mínima y máxima, respectivamente, entre cualquier par de puntos. En cambio, los algoritmos de agrupamiento particional encuentran todos los grupos simultáneamente como una partición de los datos y no imponen una estructura jerárquica (Jain, 2010). El algoritmo

particional más simple y popular es K-Means, a pesar de que su primera publicación se realizó hace aproximadamente 70 años. Este destaca como uno de los algoritmos de agrupación más utilizados dada su fácil implementación, simplicidad, eficiencia y éxito empírico aun cuando existen algunas falencias como la definición del valor óptimo de clúster con anticipación, la aleatorización en el centro de agrupamiento inicial y la sensibilidad a valores atípicos (Naeem et al., 2023).

Los algoritmos de K-Means tienen un proceso iterativo donde se comienza con la cantidad de clúster óptimos definidos previamente que se convierten en puntos iniciales llamados centroides. Cada punto del conjunto de datos se asigna al clúster cuyo centroide esté más cercano utilizando una métrica de distancia, usualmente, la euclidiana. Una vez que todos los puntos están asignados, se recalculan los centroides como el promedio de todos los puntos en cada clúster. Este proceso de asignación y recalculación de centroides se repite hasta que los centroides ya no cambian significativamente, es decir, hasta que las asignaciones de los puntos a los clústeres se estabilizan logrando una partición del conjunto de datos en la cantidad de clústeres óptimos donde los puntos dentro de cada clúster son más similares entre sí que con los puntos de otros clústeres (Ikotun et al., 2023). Aunque K-Means se ha utilizado en diversos dominios como el análisis de texto e imágenes (Et-taleby et al., 2020; (Mishra et al., 2015), predicciones (Oyelade et al., 2010), segmentación de clientes (Walters & Bekker, 2017), entre otros, en estudios urbanos suele ser una de las aplicaciones más frecuentes (Schmiedel et al., 2015; Naikoo et al., 2020).

Wang & Biljecki (2022) mencionan particularmente que los algoritmos de aprendizaje no supervisado, como el K-Means, demuestran una gran capacidad para manejar datos espaciales en estudios donde no solamente se tiene en cuenta la similitud o diferencia entre los datos, sino también las distancias espaciales relativas entre los puntos. Adicionalmente, han evidenciado en la literatura un enfoque en el que se combina estos métodos no supervisados con otros métodos. Estos flujos de trabajo colaborativos revelan que el aprendizaje no supervisado se puede aplicar no solo de forma independiente, sino también para optimizar y complementar los métodos existentes.

4. DISEÑO METODOLÓGICO

En esta sección se detallan las metodologías utilizadas para analizar la accesibilidad y la segmentación de los usuarios en relación con los programas ofertados. La subsección 4.1 presenta la medida de accesibilidad gravitacional, que se centra en la evaluación de la atractividad y la impedancia entre los puntos de origen y destino. Esta medida es crucial para entender cómo se distribuyen las oportunidades y los costos de desplazamiento en el espacio geográfico. Sin embargo, esta perspectiva puede ser insuficiente para identificar de manera precisa a los públicos potenciales debido a que la elección del uso de los programas también está influenciada por otras variables.

Para abordar esta limitación, la subsección 4.2 introduce la técnica de clústerización K-Means. Este método permite agrupar usuarios con características de accesibilidad similares y a su vez, considerar otras variables sociodemográficas y de uso de los programas. Al combinar estas dos metodologías, se busca una comprensión más completa y precisa de la distribución de oportunidades y la segmentación de los usuarios, mejorando así la identificación de públicos potenciales y la planificación de los programas.

4.1. MEDIDA DE ACCESIBILIDAD GRAVITACIONAL

Para la estimación de una medida de accesibilidad se requieren tres elementos básicos: (i) los puntos de origen, (ii) los puntos de destino y (iii) la red de transporte. Los puntos de origen representan las ubicaciones desde donde se inicia el desplazamiento, usualmente definidos como el hogar de los usuarios. En cambio, los puntos de destino son las ubicaciones de las oportunidades, como lo son las empresas, las tiendas o los cursos de los programas. La red de transporte es el sistema que conecta estos puntos, permitiendo medir el costo del desplazamiento. Con estos elementos se estima la distancia entre los puntos de origen y destino a través de la red de transporte que es la base en el proceso del cálculo de la accesibilidad.

En el caso específico de la medida clásica de la accesibilidad bajo la perspectiva de la ubicación con el componente del uso de la tierra se contemplan otros dos elementos: el primero relacionado con la atractividad y el segundo asociado a la función de impedancia.

El componente de atractividad suele medirse como el número de oportunidades en los destinos. Por ejemplo, al medir la accesibilidad a los puestos de trabajo, el valor de atracción puede ser el número de puestos de trabajo en los distintos destinos, mientras que para los centros comerciales este puede ser el número de tiendas. Para efectos de

este ejercicio, la atractividad de cada punto de destino se representa como la suma de los cupos disponibles de cada curso asociado a los programas:

$$g(W_{jk})f = \sum_k W_{jk}$$

Y la función de impedancia que representa el costo de superar la separación espacial entre los puntos de origen y destino (Chen et al., 2024) es medida como

$$f(c_{ij}^p) = e^{-\beta l_{ij}^2}$$

Estos dos componentes se pueden implementar de diferentes maneras para generar indicadores desde la perspectiva del punto de origen o demanda de los usuarios, considerando los viajes potenciales, así como desde la perspectiva del destino u oferta de los programas (Páez et al., 2012).

La siguiente ecuación tiene el punto de vista del origen i , hacia las oportunidades de tipo k y desde la perspectiva del individuo p . Esta medida de accesibilidad es una función del número de oportunidades (cupos disponibles) W de tipo k en el destino j y el costo de moverse entre i y j tal como lo percibe la persona p . La función de impedancia $f(c_{ij}^p)$ define el área de influencia alrededor de la ubicación i y, c_{ij} es el costo medido sobre una red de transporte.

$$A_{ik}^p = \sum_j g(W_{jk})f(c_{ij}^p)$$

Por otro lado, las medidas de accesibilidad que toman la perspectiva del destino pueden definirse como:

$$M_{kj}^p = \sum_j g(P_{ik}^p) f(c_{ij}^p)$$

Donde se resume el tamaño del segmento de la población p con acceso a la oportunidad de tipo k en el destino j . Esta medida es función del tamaño P del segmento de la población p en i , que debe ser atendido por la oportunidad k y el costo de viajar entre i y j tal como lo perciben los miembros del segmento de población p , y esto asociado a cada programa.

Ahora bien, estas funciones por su parte se calculan para los dos momentos: el estado actual que corresponde al cálculo de la accesibilidad para la sede donde actualmente hace uso del servicio y el estado potencial donde se tiene en cuenta la posibilidad de utilizar cualquiera de las sedes que ofertan el programa bajo el supuesto que de que no

hay limitaciones para acceder a ellos, no obstante, estos resultados se pueden quedar limitados frente a la identificación de públicos potenciales hacia las sedes y programas que mejor se adapten a sus necesidades, condiciones e intereses sin dejar de lado la premisa de mayor accesibilidad.

4.2. CLÚSTERIZACIÓN K-MEANS

Para abordar este desafío, se propone la aplicación de un algoritmo de K-Means en dos versiones, el primero relacionado únicamente con las variables de la ubicación del usuario, latitud y longitud, y el segundo adicionando variables sociodemográficas y el programa y la sede utilizada. Antes de aplicar el modelo K-Means, se realizó la normalización de las variables numéricas continuas utilizando la técnica de Min-Max Scaling y la codificación de variables categóricas como el género, la tarifa, el tipo de titular, la sede y el programa.

Sea $X = \{x_i, i = 1, \dots, n\}$ el conjunto de n -dimensiones que se agruparan en K clústers, $C = \{c_k, k = 1, \dots, K\}$, y sea μ_k la media de c_k .

4.2.1. Definición del número de clúster

Para determinar el número óptimo de clústeres, se emplearon dos métodos complementarios: el método del codo y la puntuación de la silueta. El método del codo evalúa la suma de los errores cuadrático (SSE) dentro de los clústers para diferentes valores de k , donde k hace referencia a los posibles clúster definidos en este ejercicio como un posible valor entre 1 y la cantidad de sedes disponibles. El punto donde la SSE comienza a disminuir a un ritmo más lento se considera el número óptimo de clústeres.

$$SSE = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2$$

Mientras que la puntuación de la Silueta mide la calidad del clústering mediante la evaluación de qué tan similares son los puntos dentro de un mismo clúster en comparación con puntos en otros clústers. La puntuación de la silueta varía entre -1 y 1, donde valores cercanos a 1 indican una mejor cohesión dentro del clúster.

$$s(x_i) = \frac{b(x_i) - a(x_i)}{\max \{a(x_i), b(x_i)\}}$$

donde $a(x_i)$ es la distancia media entre el punto x_i y todos los demás puntos del mismo clúster, y $b(x_i)$ es la distancia media entre el punto x_i y todos los puntos del clúster más cercano.

4.2.2. Implementación del modelo

El algoritmo K-Means requiere tres parámetros especificados antes de su anticipación: el número óptimo de clústers identificado en el proceso anterior, la inicialización del grupo y la métrica de distancia (Jain, 2010). En el caso de la métrica de distancia en este ejercicio se utiliza la métrica euclidiana para calcular la distancia entre los puntos y los centros de los grupos.

4.2.2.1. Inicialización de los centroides

El algoritmo K-Means comienza con la selección inicial de K centroides. En esta implementación, se utilizó el método de inicialización K-Means++ asociada a la librería KMeans de scikit-learn (Pedregosa et al., 2011) para mejorar la calidad de los clústers iniciales y acelerar la convergencia. Este método selecciona puntos que están lo más alejados posible entre sí, siguiendo un procedimiento probabilístico:

$$P(x_i) = \frac{D(x_i)^2}{\sum_{j=1}^n D(x_j)^2}$$

donde $D(x_i)^2$ es la distancia entre el punto x_i y el centroide más cercano previamente seleccionado.

4.2.2.2. Asignación de puntos a los clústers

Cada punto x_i se asigna al clúster cuyo centroide esté más cercano, minimizando así la distancia euclidiana entre el punto y el centroide. Matemáticamente, para un punto x_i y el conjunto de centroides $\{\mu_1, \mu_2, \dots, \mu_k\}$, la asignación se realiza mediante:

$$c_i = \arg \min_k \|x_i - \mu_k\|^2$$

Donde c_i es el índice del clúster asignado al punto x_i .

4.2.2.3. Actualización de centroides

Después de asignar todos los puntos a sus respectivos clústers, los centroides se recalculan como la media aritmética de todos los puntos asignados a cada clúster. Para el clúster c_k , el nuevo centroide μ_k se calcula como:

$$\mu_k = \frac{1}{|c_k|} \sum_{x_i \in c_k} x_i$$

donde $|c_k|$ es el número de puntos en el clúster c_k .

4.2.2.4. Iteración hasta la convergencia

El proceso de asignación de puntos y actualización de centroides se repite iterativamente hasta que los centroides convergen o se alcance un número máximo de iteraciones. La convergencia se define cuando los centroides no cambian significativamente entre iteraciones consecutivas. El objetivo de K-Means es minimizar la suma del error al cuadrado en todos los K grupos, dada por:

$$J(C) = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2$$

Esta función cuantifica la compacidad de los clústeres, con el objetivo primordial de minimizar la varianza dentro de cada clúster. Al hacerlo, se busca reducir la suma de las distancias al cuadrado entre los puntos de datos y sus centroides correspondientes, promoviendo una agrupación más homogénea y cohesionada de los datos.

5. DATOS

5.1. DESCRIPCIÓN

Las Cajas de Compensación Familiar (CCF) surgen como entidades fundamentales en el contexto de la política social colombiana establecidas mediante el Decreto 118 de 1957. Este decreto, promulgado por la Junta Militar de Gobierno, sentó las bases para la implementación del subsidio familiar. En su esencia, estas medidas intentaron reflejar el compromiso gubernamental de abordar las necesidades de las clases socioeconómicamente desfavorecidas y fomentar su desarrollo integral (Proyectamos Colombia S.A.S, 2020).

Desde entonces, las CCF han asumido el rol de gestionar y distribuir el subsidio familiar, financiado a través del 4% aplicado a las nóminas empresariales. Esta asignación se traduce en un apoyo tangible para aquellos de ingresos medianos y bajos, proporcionando beneficios en forma de dinero, bienes o servicios. No obstante, también tienen el papel de emprender políticas sociales que contribuyan al progreso y mejoría de la calidad de vida de las poblaciones beneficiadas en programas de salud, alimentación, estilo de vida, educación integral y continua, vivienda, crédito de fomento para industria familiares, recreación social, mercadeo de productos, entre otros.

5.2. CASO DE ESTUDIO

Entre los servicios que ofrecen las CCF asociados programas de estilo de vida, se han destacado dos en particular por su gran afluencia de usuarios y las condiciones particulares que presentan. Por un lado, el programa de hidroaeróbicos que se ha destacado como un curso altamente demandado por un público relativamente homogéneo, caracterizado por su interés en actividades físicas acuáticas. Sin embargo, su implementación requiere condiciones específicas de infraestructura como lo es la piscina. Por otro lado, el yoga emerge como una opción altamente ofertada, atrayendo a un público diverso y que presenta pocas barreras para su implementación, ya que puede ser adaptado a diferentes espacios y requiere un equipamiento mínimo.

5.3. DESCRIPCIÓN DE LAS VARIABLES

Los datos analizados provienen de los afiliados a la Caja de Compensación que han participado en los programas de estilo de vida de hidroaeróbicos y yoga dentro del Área Metropolitana del Valle de Aburrá durante el 2022 y 2023. Estos datos se dividen en tres categorías: información del afiliado, que hace referencia a los puntos de origen, detalles del evento o curso asociado a los programas, que hace referencia a los puntos de destino y la malla vial, que hace referencia a la red de transporte.

5.3.1. Puntos de origen

La primera comprende los datos individualizados de cada usuario de los servicios, donde se incluye las variables socioeconómicas de edad, género: femenino o masculino, categoría de afiliación vinculada al salario que perciben: A si ganan 1 salario mínimo mensual legal vigente (SMMLV), B si ganan entre 1 y 4 SMMLV y C si ganan más de 4 SMMLV, tipo de afiliado que corresponde a titular, si es la persona que cotiza a la CCF o beneficiario, si es familiar del titular que dependa económicamente y que a su vez cumpla con las edades dispuestas, en esta última opción, además se pueden dividir cónyuges, hijos, hijastros, madres, padres, hermanos y sobrinos del titular. También se encuentra información del tipo de titular, esto es si es dependiente, independiente, pensionado o facultativo y, en especial, la ubicación geográfica del hogar, identificada mediante coordenadas de latitud y longitud, que representan el punto de origen. Adicionalmente, los puntos de origen se posicionan espacialmente dentro de los polígonos que relacionan a los municipios del Área Metropolitana, o comuna en el caso de Medellín, logrando con ello la obtención de una nueva variable relacionada a la ubicación.

5.3.2. Puntos de destino

En un segundo lugar, se encuentra la información sobre los eventos o cursos de cada programa. Estos corresponden al programa que hace uso: hidroaeróbicos o yoga, y también al curso específico que realizó, en particular cada programa puede ofrecer diversos cursos que corresponden a distintas duraciones y temáticas. En hidroaeróbicos no varía la temática sino principalmente la duración, en este se presenta, el maratón y el ritual que son cursos de menor duración y compromiso, y los cursos completos con una mayor dedicación horaria. En cambio, en yoga se presentan no solamente cursos con diversas duraciones sino con distintos propósitos, tales como yoga digestivo, yoga en silla, yoga para madres gestantes, yoga para ser feliz y aqua yoga que se puede ver como una combinación de la temática o contenido de yoga, pero la infraestructura acuática de hidroaeróbicos. Adicionalmente, también se presenta el número de cupos óptimos disponibles, la fecha de inicio y fin del evento y finalmente la sede donde se desarrolla el evento, representado mediante coordenadas de latitud y longitud, que representan el punto de destino. Para este estudio, se tuvieron presente 22 sedes distribuidos como se muestran en la Figura 1.

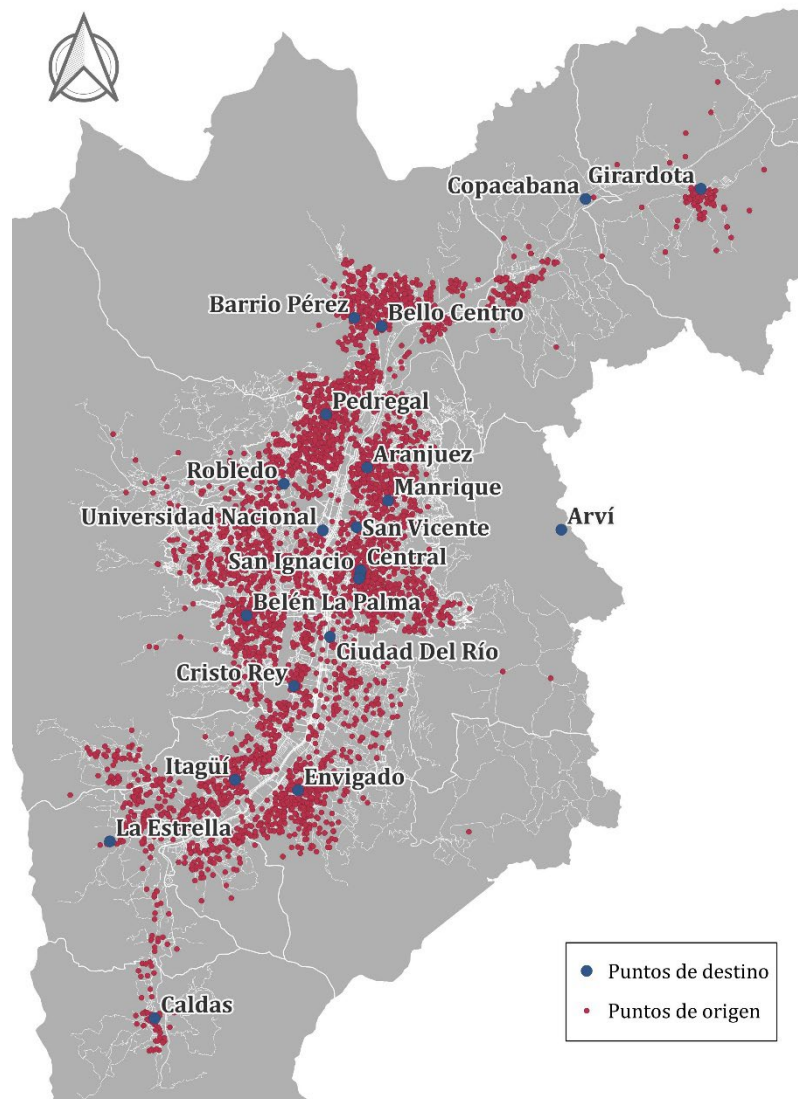
Tabla 2. Puntos de destino.

Sedes	
Aranjuez	Central
Arví	Copacabana
Barrio Pérez	Envigado
Bello Centro	Girardota
Caldas	Itagüí
Belén La Palma	La Estrella
Ciudad Del Río	Manrique
San Ignacio	Molinos
Centro	Universidad Nacional
Cristo Rey	Pedregal
Robledo	San Vicente

5.3.3. Red de transporte

En último lugar, se tiene la información de la malla vial proveniente de Open Street Map. Esta corresponde a una malla vial multimodal que se caracteriza por su capacidad para acomodar diversos medios de transporte, como automóviles, bicicletas, peatones y sistemas de transporte público colectivo. En este modelo, cada punto representa un nodo, que corresponde a un cruce o intersección en la red vial, mientras que las líneas conectan estos nodos y representan los tramos de calle que los unen. Esta representación facilita la visualización de la conectividad y la identificación de rutas potenciales entre diferentes puntos dentro del sistema vial, no obstante, presenta limitaciones en cuanto a la dirección de las vías, ya que cada una de ellas presenta los dos sentidos.

Figura 1. Puntos de origen, destino y malla vial.



5.4. ANÁLISIS EXPLORATORIO

Para los años de estudio, 2022 y 2023, hubo un total de 103.327 usuarios que participaron en los programas de hidroaeróbicos y yoga, distribuidos en 4.969 cursos. Estos cursos se ofrecieron en 22 sedes, de las cuales 2 no pertenecen a la infraestructura propia de la organización, Universidad Nacional y San Vicente.

En el caso de hidroaeróbicos se presentaron 75.535 usuarios en 3.066 cursos distribuidos en 11 sedes para un total de 81.917 cupos disponibles. De estos cursos, 779 (25,4%) no cumplieron con los cupos óptimos dispuestos logrando una ocupación

del 92,21%. Se destaca que, del total de las sedes, únicamente en la Universidad Nacional se presentó un porcentaje de subutilización por encima del 50%, en este, 22 de los 32 cursos ofrecidos tenían un número de cupos efectuados por debajo de los cupos óptimos, lo que representa un 68,8% de los cursos totales. En cambio, la sede de Cristo Rey presentó una ocupación del 100% en sus 29 cursos.

Por otro lado, para yoga se ofertaron 1.903 cursos en las 22 sedes con una participación de 27.792 usuarios y una disponibilidad de 38.708 cupos. De estos cursos, 1.363 (71,6%) no cumplieron con el cupo óptimo, lo que corresponde a una ocupación del 71,80%. En este caso, se presentan dos sedes, Cristo Rey y Los Molinos, con una ocupación cercana al 50% de los cupos. En el resto de las sedes, se observaron porcentajes de subutilización mayores, por lo que en la mayoría de los cursos ofrecidos se presentó una ocupación por debajo de su capacidad máxima. Adicionalmente, en ninguno de los cursos ofertados en las sedes externas a la organización se cumplió la meta óptima. Adicionalmente, es importante destacar que ‘aqua yoga’ es uno de los cursos con mayor ocupación. De los 38.708 cupos totales, 2.142 (5,53%) corresponden a este curso y de estos, el 90,57% fueron cupos efectuados, en comparación con el 70,7% de utilización de los cupos en los demás.

Tabla 3. Cantidad de cursos totales y cursos con subocupación por punto de destino.

Sede	Hidroaeróbicos			Yoga		
	Total cursos	Número de cursos sub-utilizados	% de cursos sub-utilizados	Total cursos	Número de cursos sub-utilizados	% de cursos sub-utilizados
Aranjuez	410	85	20,73%	75	48	64,00%
Arví				5	4	80,00%
Barrio Pérez	372	63	16,94%	15	9	60,00%
Bello Centro				202	139	68,81%
Caldas				29	28	96,55%
Belén La Palma				7	7	100,00%
Ciudad Del Río				173	126	72,83%
San Ignacio	361	122	33,80%	257	185	71,98%
Centro				1	1	100,00%
Cristo Rey	29	0	0,00%	34	17	50,00%
Robledo				10	10	100,00%
Central				112	97	86,61%
Copacabana	237	116	48,95%	67	43	64,18%
Envigado	485	50	10,31%	177	117	66,10%
Girardota	248	103	41,53%	55	51	92,73%
Itagüí	310	70	22,58%	152	135	88,82%
La Estrella	206	58	28,16%	65	52	80,00%
Manrique				10	10	100,00%

Molinos				332	180	54,22%
Universidad Nacional	32	22	68,75%	20	20	100,00%
Pedregal	376	90	23,94%	103	82	79,61%
San Vicente				2	2	100,00%
Total	3.066	779	25,41%	1.903	1.363	71,62%

En las 11 sedes que se ofrecen el programa de hidroaeróbicos, también se ofrecen los cursos de yoga, esto es porque, mientras que para hidroaeróbicos es necesario una infraestructura asociada a una piscina, en yoga, no se suele necesitar requerimientos específicos, además, de un lugar con suficiente espacio para los movimientos. En cuanto a estas sedes, Cristo Rey, es la sede con el nivel de ocupación más cercano al cupo óptimo en ambos programas, aunque también es la sede con menor disponibilidad, después le sigue Envigado y Barrio Pérez, por su parte, la sede de la Universidad Nacional fue la que presentó mayor subutilización seguido de Girardota y Copacabana.

Tabla 4. Cantidad de cupos totales y cupos utilizados por punto de destino.

Sede	Hidroaeróbicos			Yoga		
	Total de cupos	Total de cupos utilizados	Porcentaje	Total de cupos	Total de cupos utilizados	Porcentaje
Aranjuez	10.949	10.604	96,85%	1.677	1.224	72,99%
Arví				160	82	51,25%
Barrio Pérez	10.091	9.578	94,92%	220	171	77,73%
Bello Centro				3.648	2.802	76,81%
Caldas				675	288	42,67%
Belén La Palma				160	66	41,25%
Ciudad Del Río				3.910	2.777	71,02%
San Ignacio	10.313	8.558	82,98%	5.652	4.252	75,23%
Centro				20	5	25,00%
Cristo Rey	464	464	100,00%	605	478	79,01%
Robledo				232	136	58,62%
Central				2.401	1.671	69,60%
Copacabana	6.661	5.726	85,96%	1.060	839	79,15%
Envigado	13.115	12.815	97,71%	3.746	2.930	78,22%
Girardota	6.718	6.077	90,46%	1.125	696	61,87%
Itagüí	6.862	6.387	93,08%	3.542	2.123	59,94%
La Estrella	5.805	5.244	90,34%	1.345	908	67,51%
Manrique				228	89	39,04%
Molinos				5.648	4.682	82,90%
Universidad Nacional	864	548	63,43%	490	182	37,14%
Pedregal	10.075	9.534	94,63%	2.119	1.377	64,98%
San Vicente				45	14	31,11%
Total	81.917	75.535	92,21%	38.708	27.792	71,80%

Ahora bien, la población usuaria puede aprovechar los servicios de manera recurrente y por distintas duraciones, gracias a la amplia variedad de cursos disponibles de estos programas. Esta diversidad en la oferta permite a los usuarios adaptar su participación a sus necesidades específicas tales como el tiempo disponible. En ese sentido, las 103.327 participaciones iniciales hacen referencia a 26.293 usuarios únicos, de los cuales, 14.981 pertenecen al programa de hidroaeróbicos y 11.312 a yoga. Esto quiere decir que mientras las personas usuarias de hidroaeróbicos tienen una participación mediana de tres eventos, yoga solamente tiene una cifra de una participación. Por otro lado, los usuarios también pueden hacer uso de ambos servicios aunque en este ejercicio no representa un porcentaje significativo, 1.692 de usuarios (6,44%).

De los usuarios únicos aproximadamente el 44% de la población hizo uso una única vez. En el caso de hidroaeróbicos es de 5.054 usuarios y en yoga de 6.136 usuarios, sin embargo, aun cuando existe un alto nivel de repitencia, las sedes que recurren suelen mantenerse. En hidroaeróbicos, el 90,16% de los usuarios que repiten el curso, lo hacen en la misma sede, mientras que para yoga es el 75,73%. Esto nos permite concluir que, aunque en hidroaeróbicos, en términos absolutos representa casi 3 veces la cantidad de yoga, en usuarios únicos de un solo uso, este último lo sobrepasa.

Tabla 5. Resumen de participaciones por programa.

Programa	Participaciones totales	Usuarios únicos	Usuarios con un único uso	Repitencia
Hidroaeróbicos	75.535	14.981	6.136	3
Yoga	27.792	11.312	5.054	1

Bajo esta perspectiva, surgen varios desafíos adicionales en donde se hace necesario identificar a la población donde la accesibilidad espacial se convierte en una de las principales razones por las cuales accede al servicio. En ese orden de ideas, con el propósito de controlar otras variables fuera del control del estudio y obtener resultados adecuados, se selecciona una muestra de esta población siguiendo el objetivo del estudio.

5.4.1. Definición de la muestra

En primer lugar, se seleccionan a los usuarios únicos que participaron una única vez en estos programas para controlar la preferencia por el programa asociado a una sede particular. En este caso, como se mencionó anteriormente, los usuarios que suelen repetir los cursos se mantienen en la misma sede, esto nos provee una dificultad ya que probablemente estas personas no definen su participación en el programa por la accesibilidad espacial sino por otras razones de preferencia.

De igual manera, la amplia diversidad en los cursos de los programas presenta otro reto en la definición de la atractividad de las sedes. Por un lado, los cupos óptimos asociados a los cursos de corta duración suelen ser diferentes a los de los cursos de larga duración, en el caso de hidroaeróbicos, un promedio de 31 cupos frente a 27, y en yoga, de 21 frente a 20. Por el otro lado, el contenido del curso se convierte en otra variable para tener en cuenta en la elección del servicio debido a la existencia de un portafolio más amplio en los de corta duración, a diferencia de los de larga duración que incluyen temáticas más generales. En este sentido, es así como para este estudio se excluyen los cursos de corta duración, es decir, cursos de aproximadamente 10 días, calculando una proxy de la duración como la diferencia entre el día de inicio y fin del evento con el propósito de homogenizar los programas.

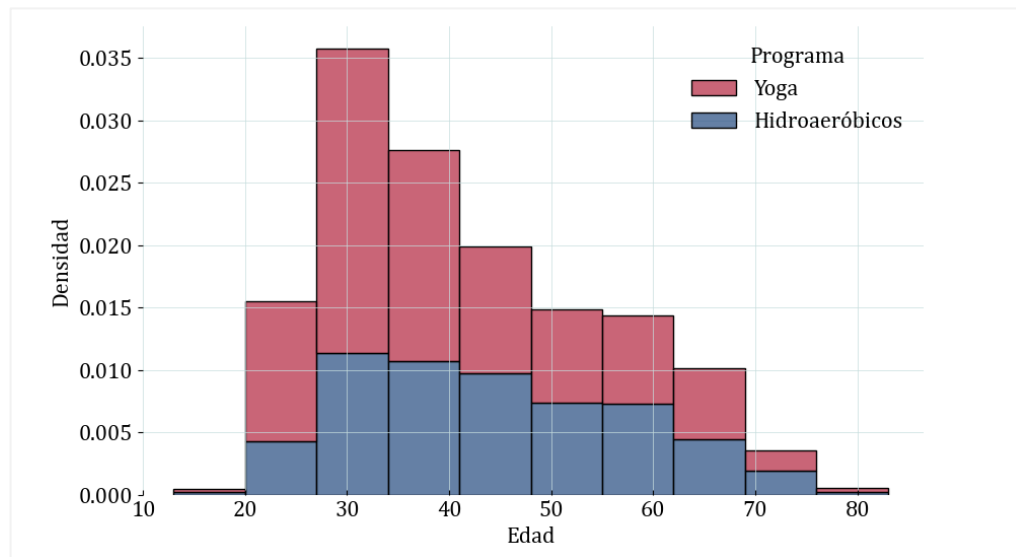
Una vez homogenizada las condiciones de los cursos y sus usos, se procede a homogeneizar la población que asiste en términos no solo del tiempo disponible sino de los recursos que tienen a disposición. Dentro de este desarrollo no se tiene presente el valor del curso, que depende además de la categoría a la que pertenecen, por lo que se seleccionan a los titulares, quienes son los que cotizan a la CCF, es decir, perciben un salario, y que corresponden al 64% de la muestra anterior. Los siguientes rubros con mayor participación serían los cónyuges con un 13,76% y los hijos o hijastros con un 6,85%. Por otro lado, aunque en esta muestra de titulares se encuentra un 14,28% pensionada, es decir, que posiblemente no se encuentre laborando y, por lo tanto, tendría una mayor disponibilidad de tiempo y otras condiciones salariales, es de interés conocer a este público en particular, ya que estos programas suelen ser llamativos para esta población. El porcentaje restante corresponde a dependientes con un 77,74% e independientes con un 7,98%.

Finalmente, aunque para cada usuario se tiene información de latitud y longitud asociado al punto de origen, en estos se encuentran algunas dificultades para georreferenciarlos de manera adecuada. De esta manera, una vez ubicados en la red de transporte, se descartan aquellos que se encuentren ubicados por fuera del Área Metropolitana del Valle de Aburrá o por encima de 1 kilómetro de distancia de la malla vial.

Esto da como resultado, una submuestra de 5.202 observaciones, un 19,78% de los usuarios únicos originales, de los cuales el 40,43% (2.103) son de hidroaeróbicos y el 59,57% restante (3.099) son de yoga. En cuanto a los programas, corresponde a 2.042 cursos con 47.319 cupos disponibles en 20 sedes donde desaparecen Arví y San Vicente. En el Anexo A se encuentra la distribución espacial de los usuarios por sede.

Esta población se compone principalmente por mujeres (86,39%) con una participación mayor en yoga por 3 puntos porcentuales por encima de hidroaeróbicos, 87,99% y 84,02%, respectivamente, cuyas edades rondan alrededor de los 38 años, siendo hidroaeróbicos el programa con usuarios de mayor edad, 42 años frente a los 36 años de yoga, y con salarios cercanos a 1 SMMLV, el 62,82% se encuentra clasificado en la categoría A, el 25,99% en la categoría B y el 9,67% en la C, el restante 1,52% hace alusión a la categoría D, es decir, usuarios que al momento de utilizar el servicio no se encontraban como afiliados de la CCF. Entre los programas, yoga cuenta con una mayor proporción de usuarios en las categorías B y C, un 40,95% contra el 27,87% de hidroaeróbicos. En cuanto al tipo de titular, el 77,9% corresponde a dependientes, 8% a independientes y el 14% restante a pensionados donde los usuarios dependientes de yoga sobrepasan en 4 puntos porcentuales a los usuarios de hidroaeróbicos y los usuarios pensionados de hidroaeróbicos en 6 puntos porcentuales a los de yoga. En el Anexo B, se encuentra la distribución espacial de los usuarios por estas variables.

Figura 2. Distribución de las edades por programa.



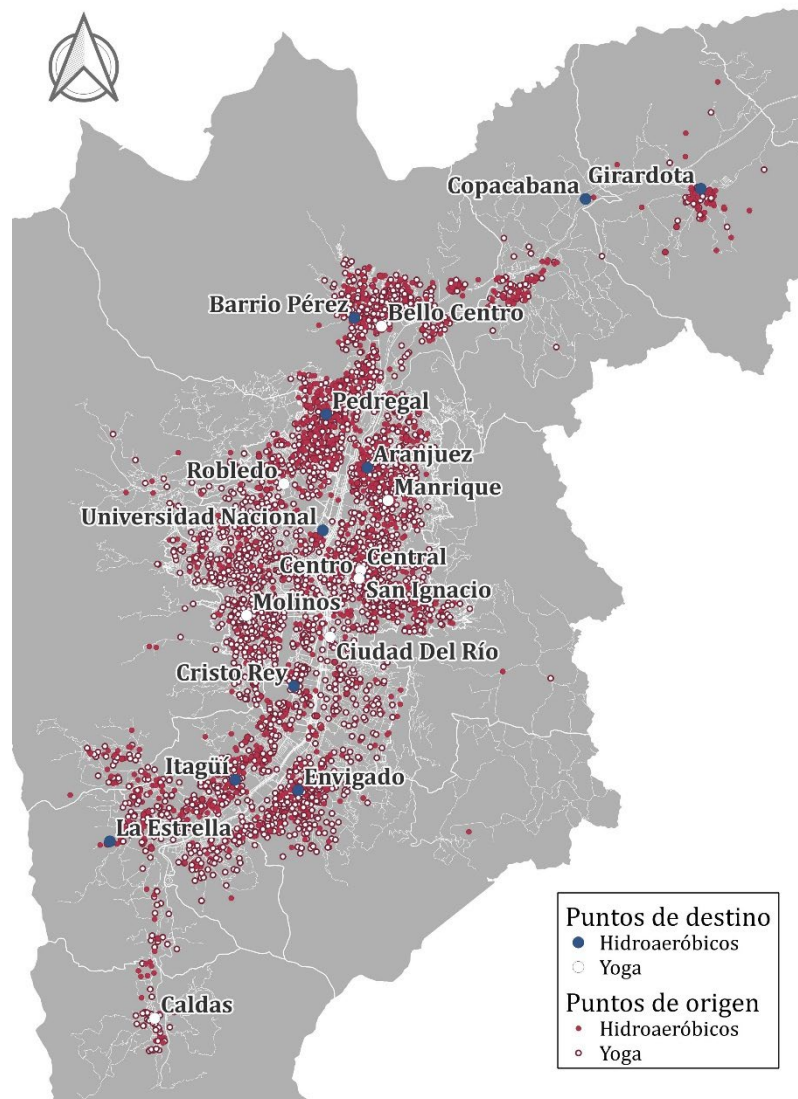
5.4.2. Cálculo de las distancias: Matriz de orígenes y destinos

Para el cálculo de las distancias, en primer lugar, se identifica el nodo más cercano del punto de origen a la malla vial. Esto arroja la distancia que un usuario necesitaría recorrer para llegar a la red, independientemente de cuál sea su medio de transporte. Esto arroja una distancia mediana de 30,13 m, una distancia mínima de 0,049 m y una distancia máxima de 952,43 m. Una vez conectados a la malla vial, se procede a calcular la matriz de orígenes y destinos que se compone por las distancias de todos los puntos de origen a cada uno de los puntos de destino donde se ofrecen los programas. En el

caso de hidroaeróbicos este está limitado a 11 sedes mientras que en yoga se encuentran las 20 sedes del estudio.

En promedio, los usuarios recorren 5.281 m para asistir a la sede en la que actualmente están haciendo uso del programa desde su punto de origen. De estos, el 42,23% está asociado además a la sede más cercana de cada afiliado. Mientras que aproximadamente la mitad de los usuarios de hidroaeróbicos participa en la sede más cercana (52,83%), solo el 35,04% de los usuarios de yoga lo hace.

Figura 3. Distribución espacial de los puntos de origen y destino por programa.



Nota: Los puntos de destino en color azul que hacen referencia a las sedes que ofertan el programa de hidroaeróbicos también incluyen dentro de sus servicios los cursos de yoga. En cambio, los puntos de color blanco hacen referencia a las sedes que únicamente ofertan el programa de yoga.

6. RESULTADOS

6.1. MEDIDAS DE ACCESIBILIDAD

6.1.1. Atractividad

Evalúe la atractividad, medido en términos de cupos disponibles, de los programas de hidroaeróbicos y yogas ofertados en las diferentes sedes presentados en la Tabla 6. En este se identifica un total de 27.322 cupos habilitados para hidroaeróbicos y 19.997 en yoga durante la temporalidad de estudio. Esto nos indica una mayor disponibilidad y, por lo tanto, un mayor nivel de atractividad para los programas de hidroaeróbicos en comparación con los de yoga aun cuando este primero tiene la mitad de la infraestructura disponible. Adicionalmente, todas las sedes que ofertan el programa de hidroaeróbicos son intensivas, es decir, proporcionan una mayor cantidad de cupos, exceptuando San Ignacio que si presenta el doble de cupos en yoga.

Tabla 6. Cupos totales de los programas en la muestra.

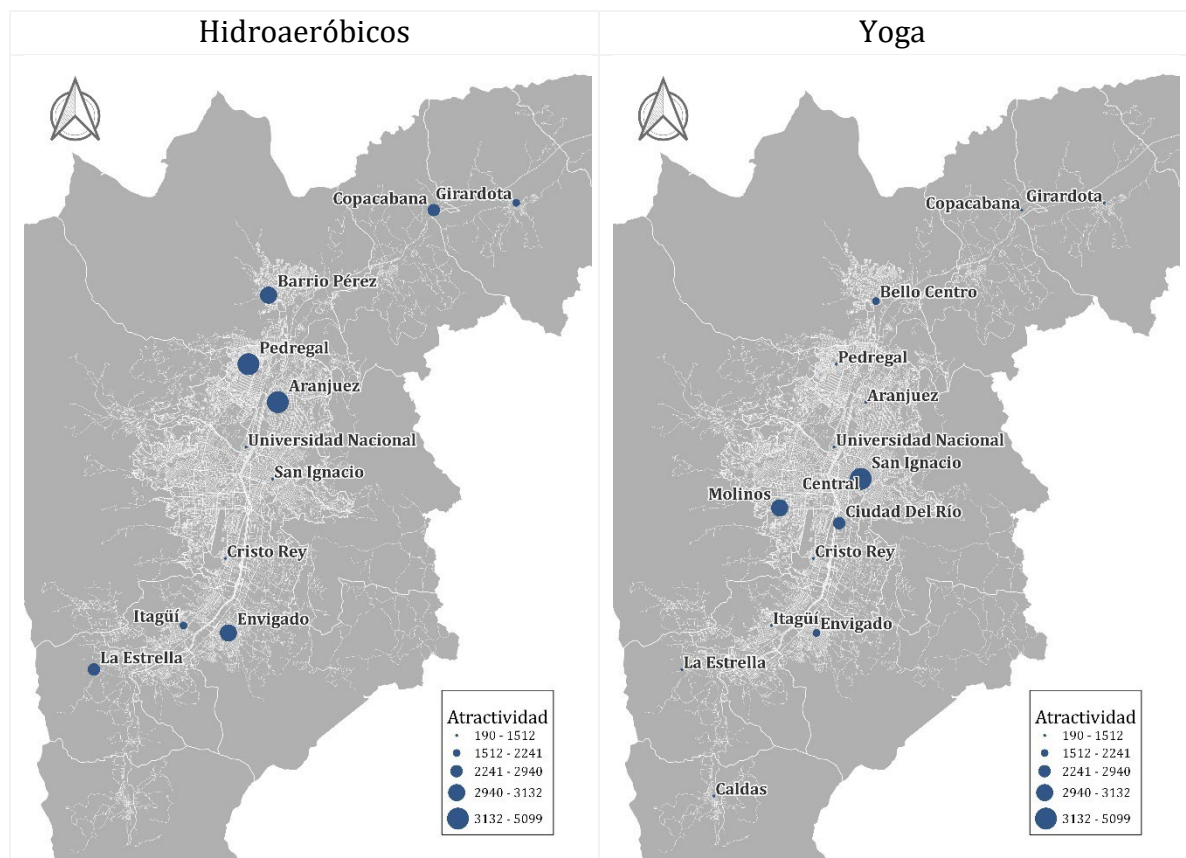
Sede	Programa	
	Hidroaeróbicos	Yoga
Aranjuez	3.950	1.087
Barrio Pérez	3.132	94
Bello Centro		1.776
Caldas		330
Belén La Palma		45
Ciudad Del Río		2.497
San Ignacio	1.512	3.219
Centro		20
Cristo Rey	190	362
Robledo		69
Central		1.011
Copacabana	2.940	360
Envigado	2.984	1.974
Girardota	2.241	669
Itagüí	1.971	1.320
La Estrella	2.655	752
Manrique		113
Molinos		3.094
Universidad Nacional	648	225
Pedregal	5.099	980
Total	27.322	19.997

Estos resultados muestran una tendencia clara hacia una mayor oferta en los programas de hidroaeróbicos en comparación con yoga que se ve influenciada a su vez por el interés de los mismos, donde como se observa en el apartado de análisis descriptivo, esta oferta viene acompañada de un nivel similar de demanda con un gran

porcentaje de cupos efectuados. No obstante, en yoga se encuentra una mayor cantidad de usuarios únicos a diferencia del programa de hidroaeróbicos que presenta un público menor, pero con un alto nivel de repitencia.

Por otro lado, esta atractividad revela ciertas particularidades o patrones en la distribución de los cupos disponibles entre las sedes representados en la Figura 4. Por ejemplo, la sede de Pedregal muestra la mayor oferta en hidroaeróbicos con 5.099 cupos, indicando una alta expectativa de demanda y posiblemente una mejor infraestructura que puede atraer a más usuarios. En contraste, la sede San Ignacio destaca en yoga con 3.219 cupos, lo que podría sugerir una oferta más diversificada o programas específicos que capturan el interés de los participantes en esa ubicación.

Figura 4. Atractividad de los programas.



Otras sedes como Aranjuez y Barrio Pérez también presentan cifras significativas de cupos en hidroaeróbicos, mientras que sedes como Envigado y Molinos tienen una oferta destacada en yoga. Esto podría indicar una variación geográfica en las preferencias de los usuarios, posiblemente influenciada por otros factores como la disponibilidad del tiempo o accesibilidad a las instalaciones.

En términos de demanda, aunque hay una menor cantidad de usuarios únicos en hidroaeróbicos, la alta repetición sugiere una fuerte lealtad y satisfacción con el programa. Por otro lado, yoga atrae a más usuarios nuevos, lo que puede reflejar una mayor curiosidad o tendencia a probar diferentes tipos de actividades físicas que a su vez puede ser explicada por la gran diversidad de cursos.

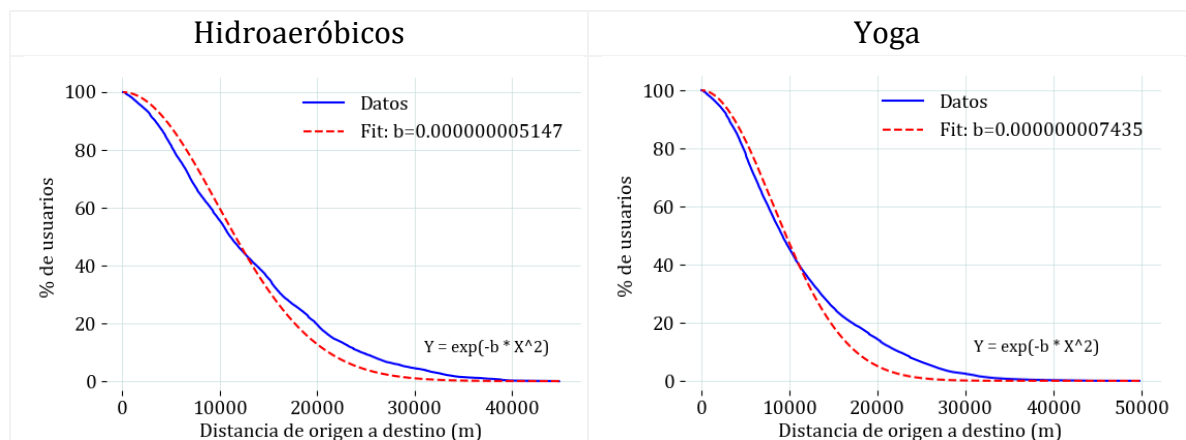
6.1.2. Función de impedancia

La Figura 5 muestra la función de impedancia para los programas donde en el eje X se representa la distancia desde el origen hasta los destinos en metros, mientras que en el eje Y se muestra el porcentaje de personas ubicadas a esa distancia de las sedes. La línea azul representa los datos observados y la línea roja punteada representa el ajuste de la función exponencial cuadrática.

Para ambos programas indican que la accesibilidad disminuye con la distancia, como era de esperarse. Sin embargo, la tasa de disminución es diferente para cada programa. Para los hidroaeróbicos, la accesibilidad disminuye a una tasa de $b = 0.00000005147$, mientras que, para yoga, la tasa es más alta con $b = 0.00000007435$. Esto sugiere que las personas están dispuestas a viajar distancias más largas para asistir a clases de hidroaeróbicos en comparación con yoga, lo cual corresponde con la menor cantidad de sedes que ofertan el servicio.

Además, los valores del coeficiente de determinación (R^2) para los ajustes indican un buen nivel de precisión en ambos modelos. Para hidroaeróbicos, el valor es de 96,76%, mientras que, para yoga, el valor es de 97,01%.

Figura 5. Función de impedancia de los programas.



6.2. DESDE EL DESTINO

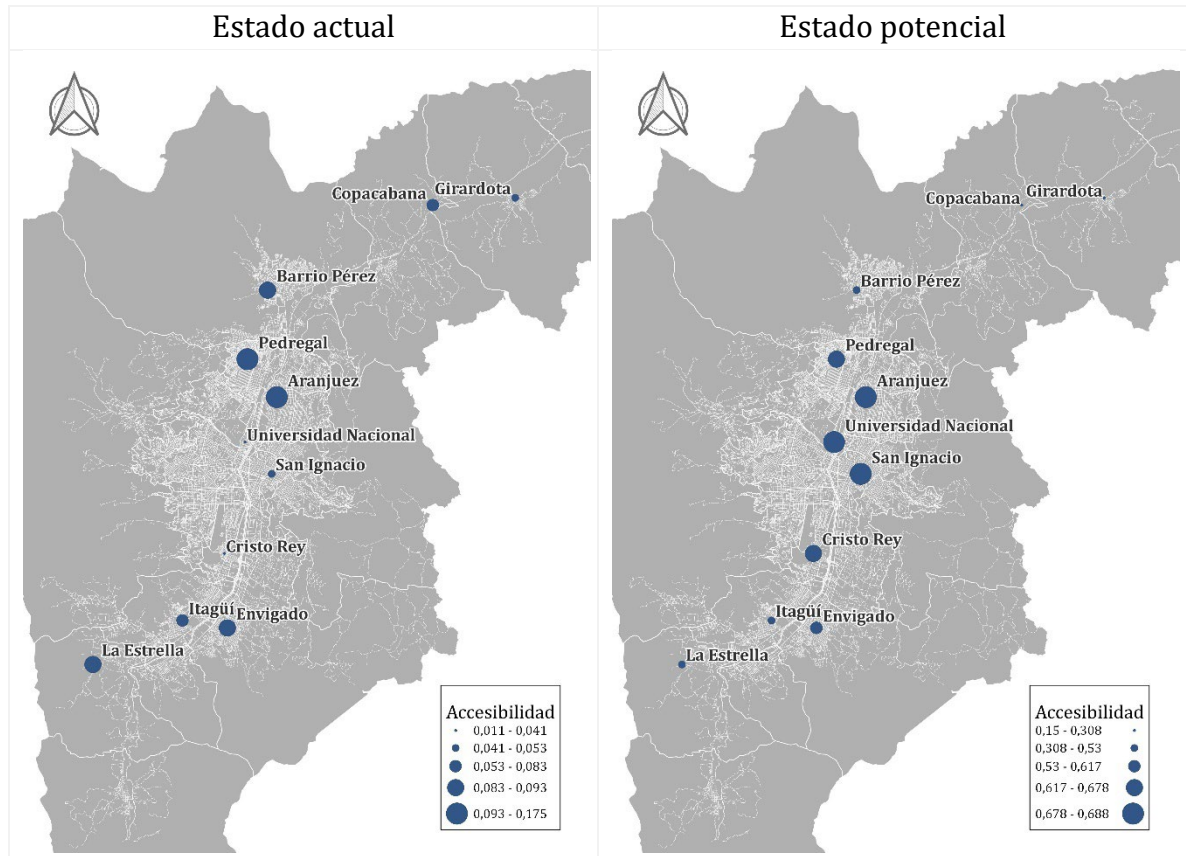
Para los programas de hidroaeróbicos, las sedes con una mayor disponibilidad de cupos son a su vez las de mayor accesibilidad. Esto se puede identificar en las sedes de Pedregal, Aranjuez y Barrio Pérez, que también son las tres sedes con más cupos disponibles, no obstante, aun cuando Pedregal cuenta con aproximadamente 1.000 cupos más que Aranjuez, Aranjuez se encuentra a distancias menores a sus usuarios dado su centralidad logrando una accesibilidad similar.

Existen otras sedes como San Ignacio, Cristo Rey y la Universidad Nacional, que muestran discrepancias significativas entre su gestión actual y su potencial accesibilidad. A pesar de estar ubicadas en áreas centrales y rodeadas por una población densa, esta disparidad sugiere que están enfrentando barreras que limitan su accesibilidad. En el caso de Cristo Rey, esta sede cuenta con el 100% de sus cupos disponibles efectuados, aunque la cantidad de cupos es la mínima entre todas las sedes. En cambio, en las otras dos, aun cuando existe una cantidad prudente de cupos disponibles, suelen ser las sedes con la menor proporción de cupos efectuados por lo que también se vuelve necesario la identificación de otras barreras que están impidiendo que los usuarios potenciales accedan a estos programas.

Por otro lado, las sedes ubicadas en los otros municipios que no son Medellín, como La Estrella, Envigado, Copacabana y Girardota, presentan una dinámica contraria. Aunque inicialmente muestran un estado potencial bajo debido a las mayores distancias a los usuarios en comparación con las sedes centrales, en la práctica exhiben altos niveles de accesibilidad y participación real. Este resultado sugiere que estas sedes logran abarcar la población que las rodea con un nivel adecuado de cupos disponibles. Esta observación resalta la importancia de no subestimar el potencial de las sedes ubicadas en municipios cercanos, ya que con la estrategia adecuada pueden superar las expectativas teóricas de accesibilidad y convertirse en puntos clave de participación.

En cuanto al programa de Yoga, en general, se revela una fuerte concordancia entre la accesibilidad teórica y la participación real. Gran parte de las sedes se encuentran ubicadas céntricamente y el nivel de accesibilidad no se encuentra alejado de sus valores potenciales a excepción de Robledo y la Universidad Nacional que también son sedes con bajos niveles de cupos efectuados frente a los cupos disponibles.

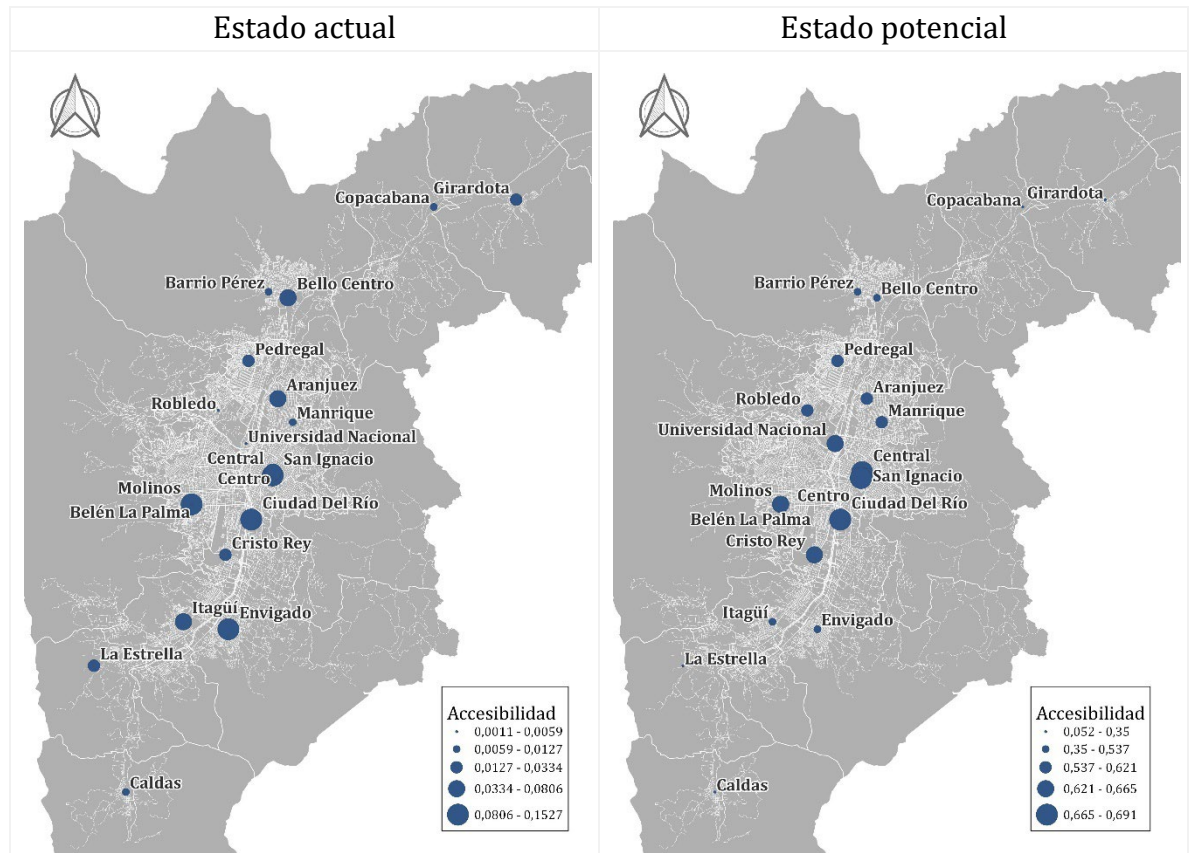
Figura 6. Accesibilidad desde el destino del estado actual y potencial del programa de hidroaeróbicos.



Dada la diversidad de cursos ofrecidos dentro del programa de yoga y aproximadamente el doble de sedes, el porcentaje de ocupación tiende a ser menor. Esta amplia variedad de opciones y ubicaciones proporciona a los usuarios múltiples alternativas para inscribirse, lo que distribuye la participación entre más sedes y cursos, resultando en una menor densidad de ocupación en cada una. La amplia oferta de cursos de yoga puede incluir diferentes temáticas, horarios y ubicaciones para facilitar el acceso a una audiencia más amplia, sin embargo, esta misma diversidad puede diluir la concentración de usuarios en cada curso o sede, dado que los participantes pueden elegir entre muchas opciones disponibles. Además, la mayor cantidad de sedes implica una mayor dispersión geográfica, lo que puede influir en la accesibilidad y la conveniencia percibida por los usuarios. Aunque esto mejora la accesibilidad teórica al programa al acercar las opciones a más personas, también puede llevar a que algunas sedes no alcancen su capacidad máxima debido a la dispersión de la demanda.

Por otro lado, la oferta diversificada también puede responder a una estrategia de adaptación a las necesidades específicas de diferentes segmentos de la población. Esto puede incluir la oferta de cursos especializados para principiantes, avanzados, aqua yoga, yoga prenatal, entre otros, lo que añade valor al programa, pero también contribuye a la distribución de los usuarios en una mayor cantidad de cursos específicos.

Figura 7. Accesibilidad desde el destino del estado actual y potencial del programa de yoga.

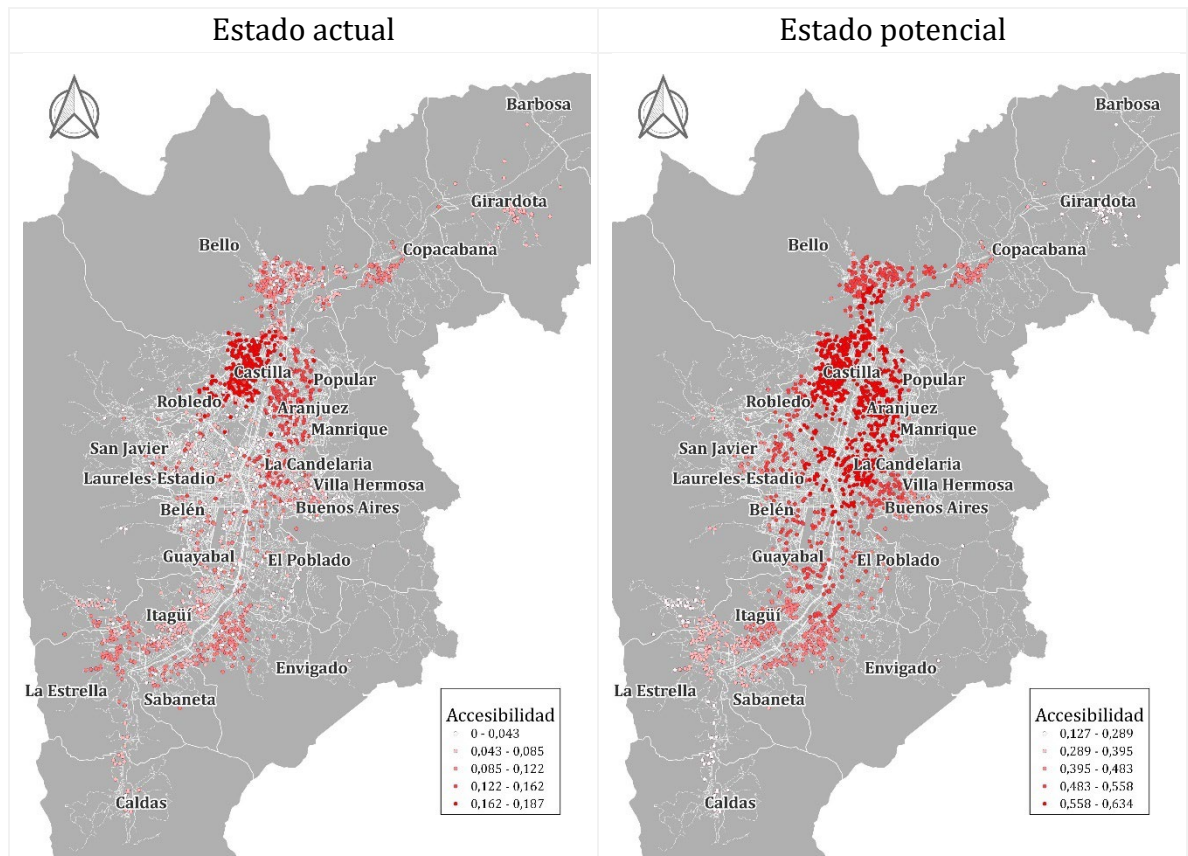


Adicionalmente, se identifican también sedes muy cercanas entre sí, lo que sugiere la posibilidad de estudiar la consolidación de un único portafolio para esas zonas. Esta estrategia podría optimizar la distribución de los recursos y aumentar la ocupación de los cursos al concentrar la demanda en menos sedes, mejorando así la eficiencia operativa y reduciendo costos. Estas situaciones plantean un desafío para la planificación y la gestión del programa de yoga donde es fundamental equilibrar la oferta de cursos y sedes con la demanda real para optimizar la ocupación y asegurar la sostenibilidad de cada curso y sede.

6.3. DESDE EL ORIGEN

En la Figura 8 y Figura 9, que hacen referencia a los usuarios de hidroaeróbicos y yoga, respectivamente, se muestra una clara concentración de puntos rojos en las zonas centrales asociadas igualmente a la cantidad de sedes que ofertan estos programas. En el caso de hidroaeróbicos se presenta una mayor accesibilidad para los usuarios que se encuentran ubicados alrededor de la comuna de Doce de Octubre, zona que está rodeada por cuatro sedes. En cambio, para yoga, la concentración se desplaza más hacia el centro con Laureles-Estadio, La Candelaria y Villa Hermosa, y sur con Belén y Guayabal, zonas que están rodeada por seis sedes.

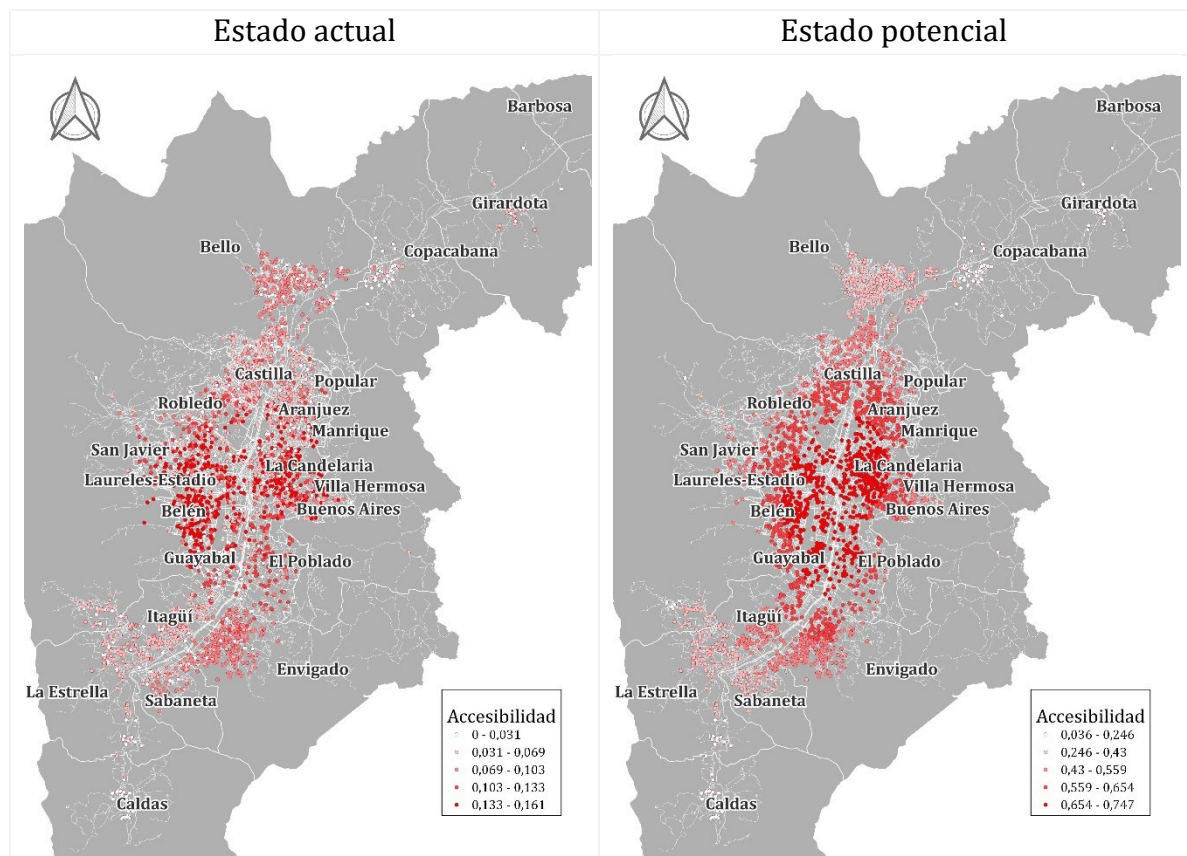
Figura 8. Accesibilidad desde el origen del estado actual y potencial del programa de hidroaeróbicos.



En contraste, las zonas periféricas presentan puntos más dispersos y de tonalidades más claras que indican una menor accesibilidad, lo cual está asociado a una menor densidad de servicios y redes transporte. La disparidad en la accesibilidad entre las áreas centrales y periféricas resalta la necesidad de estrategias de planificación que aborden estas desigualdades, promoviendo un desarrollo más equilibrado y equitativo.

Por otro lado, los corredores de puntos que se extienden desde las zonas centrales hacia las periferias son indicativos de las principales rutas de transporte que facilitan el acceso a estas áreas más alejadas. Estos corredores son esenciales para la conectividad regional y juegan un papel crucial en la integración de las zonas periféricas con el centro urbano. La identificación de estos corredores permite planificar mejoras en la infraestructura de transporte, optimizando la movilidad y reduciendo el tiempo de viaje entre diferentes partes de la región.

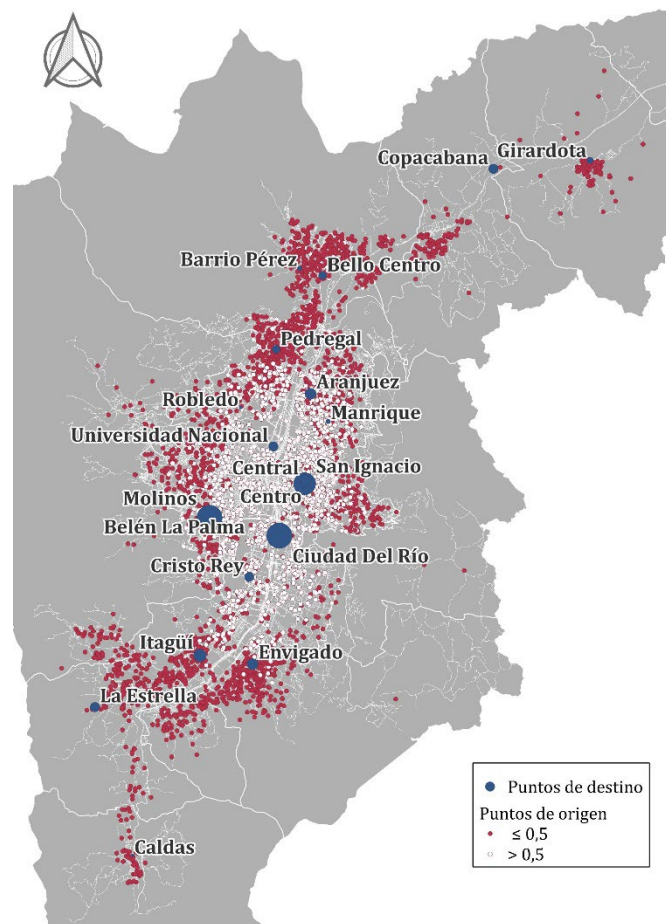
Figura 9. Accesibilidad desde el origen del estado actual y potencial del programa de yoga.



En vista de las diferencias entre la accesibilidad en el estado potencial y el estado actual, el 34,78% (1.809) de los usuarios superan los 0,5 puntos entre estos resultados, lo que corresponde a la población con mayores oportunidades de mejora en torno al uso de los programas. De estos, el 80,6% pertenece a los usuarios de los cursos de yoga. Esto conversa con lo anteriormente mencionado en la accesibilidad desde el punto de vista del destino donde no solamente las sedes presentan un menor nivel de ocupación, sino que también una mayor dispersión geográfica, y a su vez con los resultados del análisis descriptivo donde solo cerca del 35% de estos usuarios asiste a su sede más cercana.

En cuanto a las variables sociodemográficas, la edad mediana es un año menor al de la muestra total con un valor de 37 años y en la tarifa se presenta una mayor concentración en B y C que se ve influenciada por la cantidad de usuarios en yoga. Frente al género no se encuentran patrones diferenciadores en particular por la gran concentración de mujeres, pero con respecto al tipo de titular, si se encuentra una gran diferencia con al menos 4 puntos porcentuales por encima del promedio en torno a los dependientes, esto último sugiere que es probable que estas personas utilicen los servicios en las sedes más cercanas a su lugar de trabajo en lugar de las sedes cercanas a su hogar, supuesto que concuerda con lo representado en la Figura 10 donde existe un patrón centro-periferia, con los usuarios con mayor diferencia en accesibilidad en los alrededores del Área Metropolitana, pero las sedes utilizadas, Ciudad del Río, Molinos y San Ignacio, en zonas caracterizadas por alta concentración empresarial.

Figura 10. Usuarios con la mayor diferencia entre la accesibilidad potencial y actual.



Ahora bien, estas medidas de accesibilidad en el estado potencial suponen que todos los usuarios pueden acceder a estos programas. Este análisis teórico se basa en la

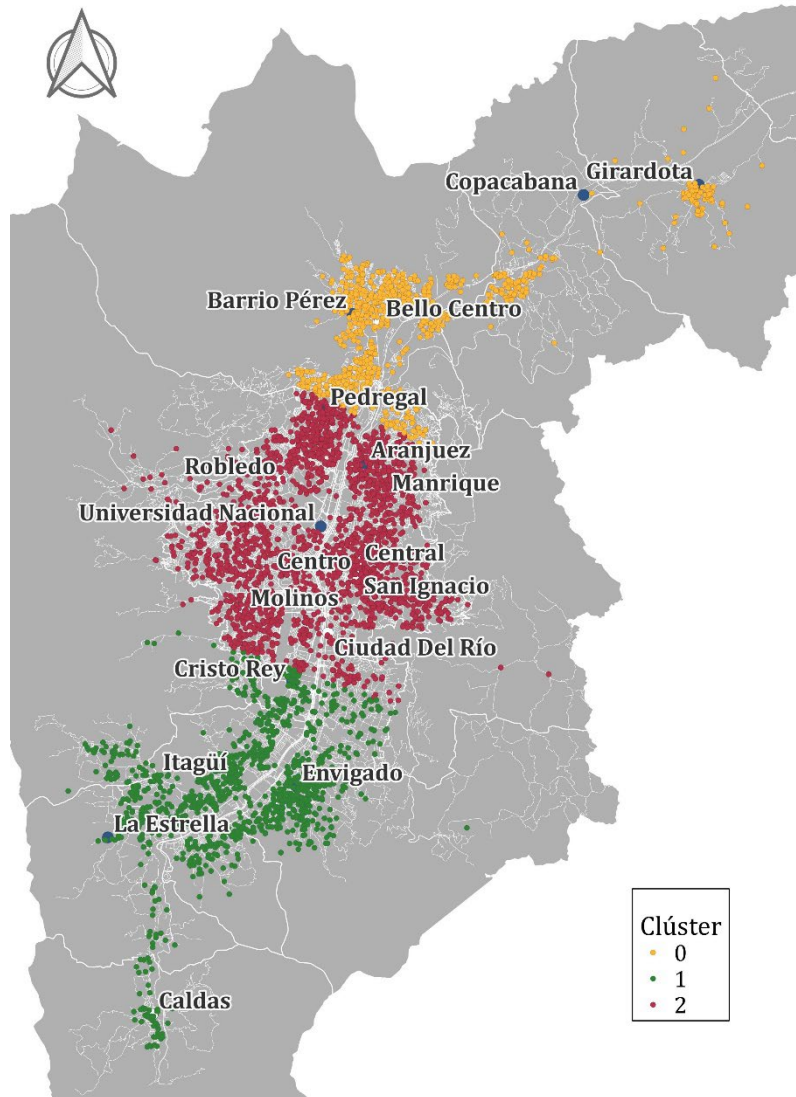
premisa de que la infraestructura y la ubicación de las sedes permiten que cualquier persona pueda asistir a los cursos, no obstante, en la realidad, aunque exista la posibilidad de utilizar estos servicios, no necesariamente se encuentra en el deseo de la persona su uso. Si bien las medidas de accesibilidad en el estado potencial proporcionan una valiosa visión de las oportunidades teóricas para la participación en los programas de yoga e hidroaeróbicos, es esencial considerar la realidad de uso para una planificación efectiva dado que la participación real está influenciada por una compleja interacción de factores personales y sociales.

6.4. SEGMENTACIÓN K-MEANS

En la primera segmentación, se analizó únicamente las variables relacionadas a la ubicación del usuario, es decir, latitud y longitud, para crear los clústeres. En este se identifica claramente tres grupos asociados a las divisiones usuales del centro, norte y sur del Área Metropolitana como se muestra en la Figura 11. El primer clúster agrupa a los usuarios de municipios periféricos y zonas al norte de Medellín. Dentro de los límites de este clúster se encuentra cuatro sedes, Barrio Pérez, Bello Centro, Copacabana y Girardota. En el segundo, se agrupan a los usuarios que residen en los barrios centrales y del norte de Medellín, incluyendo áreas como Doce de Octubre, Robledo, Laureles-Estadio, La Candelaria y Villa Hermosa con la mayor cantidad de sedes disponibles, en particular la sede de Pedregal se encuentra muy al límite del primer clúster. Y el tercer clúster con los usuarios del sur de Medellín y los municipios de Itagüí, Envigado, La Estrella y Caldas, cuentan con cinco sedes donde la de Cristo Rey está localizada cerca del borde del segundo clúster.

En la segunda segmentación, a las variables de ubicación se les añadieron las variables sociodemográficas de los usuarios, tales como género, edad, tarifa y tipo de titular, y las variables correspondientes al servicio como la sede y programa que utilizan. El modelo arrojó dos clústers óptimos, determinados mediante el método del codo y la puntuación de silueta, los cuales se muestran en la Figura 12. El clúster 1 agrupa a 2.080 usuarios ubicados en los municipios periféricos de Copacabana y Girardota, así como en zonas densamente pobladas del norte de Medellín, incluyendo Barrio Pérez, Bello Centro y Aranjuez. Por otro lado, el clúster 0 incluye a 3.122 usuarios que se encuentran desde el centro hasta el sur de Medellín, abarcando barrios como Pedregal, Laureles, La Candelaria, y extendiéndose hasta Itagüí, Envigado y La Estrella.

Figura 11. Clúster con variables de ubicación.

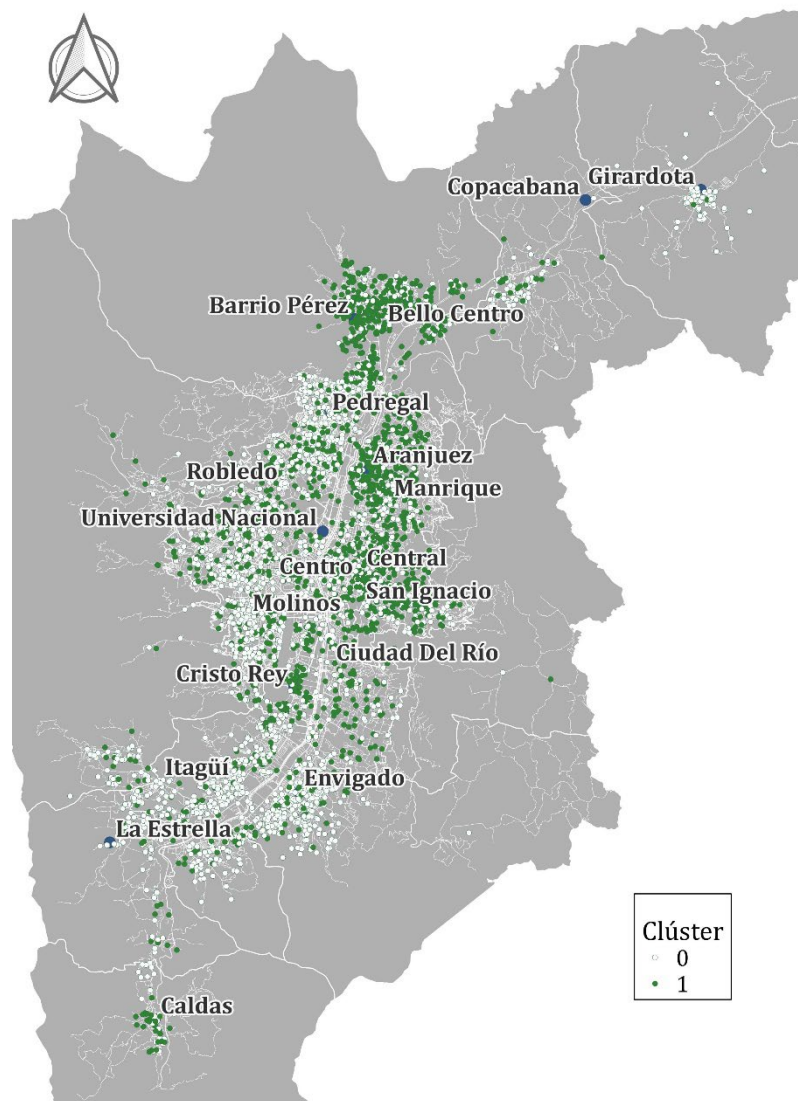


En términos sociodemográficos, los usuarios en el clúster 0 tienen una edad promedio de 39 años, en comparación con los 37 años del clúster 1. Además, el clúster 0 presenta una mayor proporción de usuarios con tarifas B y C, y una predominancia de pensionados e independientes. Estas diferencias indican que el clúster 0 tiende a incluir a una población ligeramente mayor y con una posible mejor situación económica, así como, una mayor representación de personas retiradas o con ocupaciones autónomas.

Aunque el número de usuarios en cada clúster es similar a la cantidad de usuarios de cada programa, no se observa un patrón generalizado donde se relaciona que todos los usuarios del mismo clúster pertenecen al mismo programa. En el clúster 0, el 52,34% de los usuarios pertenecen a programas de yoga, mientras que el 47,66% restante

participa en hidroaeróbicos, reflejando una distribución bastante equilibrada. En contraste, el clúster 1 muestra una mayor inclinación hacia yoga, con un 70,43% de los usuarios participando en este programa, frente al 29,57% que opta por hidroaeróbicos. No obstante, espacialmente si se puede notar una concordancia entre estas dos variables, por ejemplo, en Girardota, Itagüí, Envigado y La Estrella cuyas sedes ofrecen ambos programas, pero con mayor intensidad en los de hidroaeróbicos, se encuentra una gran concentración del clúster 0, en cambio, en Caldas y zonas céntricas de Medellín donde sus sedes ofrecen únicamente yoga se muestra una gran proporción del clúster 1.

Figura 12. Clúster con variables de ubicación, sociodemográficas y del programa.



Siguiendo lo anterior y analizando los datos desde la perspectiva de los programas, esta relación se identifica de manera más clara. En el caso de los usuarios de yoga, la segmentación es bastante equilibrada entre los dos clústers, con cifras cercanas al 50% en cada uno. Sin embargo, en el caso de los hidroaeróbicos, existe una clara predominancia en el clúster 0, con una proporción de 70-30 a favor de este clúster. Esta diferencia se alinea con el análisis descriptivo previo, donde se identificó que el público de hidroaeróbicos tiende a ser de mayor edad y en su mayoría pensionados o independientes. Esto quiere decir que, aunque en ambos programas se encuentran estos dos públicos, en yoga se identifica una población más homogénea, mientras que en hidroaeróbicos un público más específico.

7. CONCLUSIONES

Aunque hidroaeróbicos tiene una menor cantidad de usuarios únicos, la alta recurrencia sugiere una fuerte lealtad y satisfacción con el programa, lo que a su vez puede ser aprovechado para apalancar el uso de otros programas como aqua yoga. En cambio, los cursos de yoga muestran un nivel de atracción mayor entre los usuarios nuevos que puede explicarse por la mayor diversidad de temáticas y los pocos requerimientos necesarios para su desarrollo. Esta diferencia se ve reflejada en los patrones de desplazamiento, donde las personas parecen dispuestas a viajar distancias más largas para asistir a los cursos de hidroaeróbicos, lo cual también corresponde con la menor cantidad de sedes que ofrecen este servicio.

Por otro lado, aunque una mayor variedad en el portafolio y sedes que ofertan los programas proporciona múltiples alternativas a los usuarios y, en teoría, una mayor cantidad de cupos o distancias más cortas que mejoran la accesibilidad se identifica que esto puede llevar a una subutilización de la capacidad óptima, como es el caso de yoga. Esto resalta la urgencia de equilibrar la oferta de servicios con las necesidades del público, debido a que estos espacios subutilizados podrían ser aprovechados para otros propósitos. Además, también se observan sedes muy cercanas entre sí, que sugieren la posibilidad de consolidar un único portafolio para esas zonas, optimizando la distribución de recursos y aumentando la ocupación de los cursos.

De igual manera, la disparidad en la accesibilidad entre áreas centrales y periféricas resalta la necesidad de estrategias de planificación que aborden estas desigualdades, promoviendo un desarrollo más equilibrado y equitativo. Es necesario identificar corredores de alta accesibilidad que permitan planificar mejoras en la infraestructura de transporte, optimizando la movilidad y reduciendo el tiempo de viaje. Además, estos resultados invitan a no subestimar el potencial de las sedes ubicadas en zonas aledañas, ya que pueden superar las expectativas teóricas de accesibilidad y convertirse en puntos clave de participación.

Finalmente, (Levesque et al., 2013) menciona que la accesibilidad también puede considerarse como un concepto de cinco dimensiones: (i) proximidad, (ii) aceptabilidad, (iii) disponibilidad, (iv) asequibilidad y (v) adecuación. Aunque este estudio no aborda el desarrollo de las cinco dimensiones, la combinación de dos métodos procura justamente analizar un panorama más amplio en el entendimiento del uso de estos servicios. Al identificar y abordar este panorama, se pueden diseñar estrategias más efectivas que optimicen los recursos y maximicen la participación en los programas, no obstante, las variables referenciadas pueden quedar cortas en la definición de las preferencias de los usuarios por estos programas. Adicionalmente, el

ejercicio también está limitado únicamente a los puntos de origen asociados a los hogares, sin embargo, dada la disponibilidad del tiempo de ocio asociado usualmente a horarios después de la jornada laboral, la ubicación del lugar de trabajo puede arrojar otro tipo de análisis interesante que complementa la información. A pesar de estas limitaciones, este enfoque nos ha permitido no solo identificar patrones y tendencias, sino también proporcionar insumos valiosos para la toma de decisiones, mejorando la accesibilidad y efectividad de los programas, asegurando que se adapten a las características y necesidades específicas de los diferentes grupos de usuarios.

REFERENCIAS

- Achuthan, K., Titheridge, H., & Mackett, R. L. (2010). Mapping accessibility differences for the whole journey and for socially excluded groups of people. *Journal of Maps*, 6(1), 220-229. <https://doi.org/10.4113/jom.2010.1077>
- Allen, J., & Farber, S. (2020). Planning transport for social inclusion: An accessibility-activity participation approach. *Transportation Research Part D: Transport and Environment*, 78, 102212. <https://doi.org/10.1016/j.trd.2019.102212>
- Chen, P., Wang, W., Qian, C., Cao, M., & Yang, T. (2024). Gravity-based models for evaluating urban park accessibility: Why does localized selection of attractiveness factors and travel modes matter? *Environment and Planning B: Urban Analytics and City Science*, 51(4), 904-922. <https://doi.org/10.1177/23998083231206168>
- Cheng, J., & Bertolini, L. (2013). Measuring urban job accessibility with distance decay, competition and diversity. *Journal of Transport Geography*, 30, 100-109. <https://doi.org/10.1016/j.jtrangeo.2013.03.005>
- Christian, T. J. (2012). Trade-Offs Between Commuting Time and Health-Related Activities. *Journal of Urban Health*, 89(5), 746-757. <https://doi.org/10.1007/s11524-012-9678-6>
- Church, A., Frost, M., & Sullivan, K. (2000). Transport and social exclusion in London. *Transport Policy*, 7(3), 195-205. [https://doi.org/10.1016/S0967-070X\(00\)00024-X](https://doi.org/10.1016/S0967-070X(00)00024-X)
- Crawford, D. W., Jackson, E. L., & Godbey, G. (1991). A hierarchical model of leisure constraints. *Leisure sciences*, 13(4), 309-320.
- Danckert, J. A., & Allman, A.-A. A. (2005). Time flies when you're having fun: Temporal estimation and the experience of boredom. *Brain and Cognition*, 59(3), 236-245. <https://doi.org/10.1016/j.bandc.2005.07.002>
- Dattilo, A. E. L., Ewert, A., & Dattilo, J. (2012). Learning as leisure: Motivation and outcome in adult free time learning. *Journal of Park & Recreation Administration*, 30(1).

- El-Geneidy, A. M., & Levinson, D. M. (2006). Access to Destinations: Development of Accessibility Measures. *Report # 1 in the series Access to Destinations Study*, 1-125.
- Et-taleby, A., Boussetta, M., & Benslimane, M. (2020). Faults Detection for Photovoltaic Field Based on K-Means, Elbow, and Average Silhouette Techniques through the Segmentation of a Thermal Image. *International Journal of Photoenergy*, 2020, 1-7. <https://doi.org/10.1155/2020/6617597>
- Fancourt, D., Aughterson, H., Finn, S., Walker, E., & Steptoe, A. (2021). How leisure activities affect health: A narrative review and multi-level theoretical framework of mechanisms of action. *The Lancet Psychiatry*, 8(4), 329-339. [https://doi.org/10.1016/S2215-0366\(20\)30384-9](https://doi.org/10.1016/S2215-0366(20)30384-9)
- Gaglione, F., Gargiulo, C., Zucaro, F., & Cottrill, C. (2022). Urban accessibility in a 15-minute city: A measure in the city of Naples, Italy. *Transportation Research Procedia*, 60, 378-385. <https://doi.org/10.1016/j.trpro.2021.12.049>
- Geurs, K. T., & Ritsema van Eck, J. R. (2001). Accessibility measures: Review and applications. Evaluation of accessibility impacts of land-use transportation scenarios, and related social and economic impact. *RIVM rapport 408505006*.
- Geurs, K. T., & 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>
- Ghahramani, Z. (2004). Unsupervised learning. *In Summer school on machine learning (pp. 72-112)*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Gimenez-Nadal, J. I., & Molina, J. A. (2019). Daily feelings of US workers and commuting time. *Journal of Transport & Health*, 12, 21-33. <https://doi.org/10.1016/j.jth.2018.11.001>
- Gottholmseder, G., Nowotny, K., Pruckner, G. J., & Theurl, E. (2009). Stress perception and commuting. *Health Economics*, 18(5), 559-576. <https://doi.org/10.1002/hec.1389>

- Guzman, L. A., Cantillo-Garcia, V. A., Oviedo, D., & Arellana, J. (2023). How much is accessibility worth? Utility-based accessibility to evaluate transport policies. *Journal of Transport Geography*, *112*, 103683. <https://doi.org/10.1016/j.jtrangeo.2023.103683>
- Haller, M., Hadler, M., & Kaup, G. (2013). Leisure Time in Modern Societies: A New Source of Boredom and Stress? *Social Indicators Research*, *111*(2), 403-434. <https://doi.org/10.1007/s11205-012-0023-y>
- Harris, B. (2001). Accessibility: Concepts and applications. *Journal of Transportation and Statistics*, *4*(2/3), 15-30.
- Hilbrecht, M., Smale, B., & Mock, S. E. (2014). Highway to health? Commute time and well-being among Canadian adults. *World Leisure Journal*, *56*(2), 151-163. <https://doi.org/10.1080/16078055.2014.903723>
- Horner, M. W., & Downs, J. (2014). Integrating people and place: A density-based measure for assessing accessibility to opportunities. *Journal of Transport and Land Use*, *7*(2), 23-40. <https://doi.org/10.5198/jtlu.v7i2.417>
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*, 178-210. <https://doi.org/10.1016/j.ins.2022.11.139>
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651-666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Jiang, Y. (2017). *Urban Accessibility Measurement and Visualization—A Big Data Approach*.
- Kelly, S., Martin, S., Kuhn, I., Cowan, A., Brayne, C., & Lafortune, L. (2016). Barriers and Facilitators to the Uptake and Maintenance of Healthy Behaviours by People at Mid-Life: A Rapid Systematic Review. *PLOS ONE*, *11*(1), e0145074. <https://doi.org/10.1371/journal.pone.0145074>

- Kenyon, S., Lyons, G., & Rafferty, J. (2002). Transport and social exclusion: Investigating the possibility of promoting inclusion through virtual mobility. *Journal of Transport Geography*, 10(3), 207-219. [https://doi.org/10.1016/S0966-6923\(02\)00012-1](https://doi.org/10.1016/S0966-6923(02)00012-1)
- Khavarian-Garmsir, A. R., Sharifi, A., Hajian Hossein Abadi, M., & Moradi, Z. (2023). From Garden City to 15-Minute City: A Historical Perspective and Critical Assessment. *Land*, 12(2), 512. <https://doi.org/10.3390/land12020512>
- Kwan, M.-P., Murray, A. T., O'Kelly, M. E., & Tiefelsdorf, M. (2003). Recent advances in accessibility research: Representation, methodology and applications. *Journal of Geographical Systems*, 5(1), 129-138. <https://doi.org/10.1007/s101090300107>
- Levesque, J.-F., Harris, M. F., & Russell, G. (2013). Patient-centred access to health care: Conceptualising access at the interface of health systems and populations. *International Journal for Equity in Health*, 12(1), 18. <https://doi.org/10.1186/1475-9276-12-18>
- Lorenz, O. (2018). Does commuting matter to subjective well-being? *Journal of Transport Geography*, 66, 180-199. <https://doi.org/10.1016/j.jtrangeo.2017.11.019>
- Mishra, R. K., Saini, K., & Bagri, S. (2015). Text document clustering on the basis of inter passage approach by using K-means. *International Conference on Computing, Communication & Automation*, 110-113. <https://doi.org/10.1109/CCAA.2015.7148354>
- Mondschein, A., & Taylor, B. D. (2017). Is traffic congestion overrated? Examining the highly variable effects of congestion on travel and accessibility. *Journal of Transport Geography*, 64, 65-76. <https://doi.org/10.1016/j.jtrangeo.2017.08.007>
- Mouratidis, K. (2019). Built environment and leisure satisfaction: The role of commute time, social interaction, and active travel. *Journal of Transport Geography*, 80, 102491. <https://doi.org/10.1016/j.jtrangeo.2019.102491>
- Mouratidis, K., Ettema, D., & Næss, P. (2019). Urban form, travel behavior, and travel satisfaction. *Transportation Research Part A: Policy and Practice*, 129, 306-320. <https://doi.org/10.1016/j.tra.2019.09.002>

- Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. *International Journal of Computing and Digital Systems*, 13(1), 911-921. <https://doi.org/10.12785/ijcnds/130172>
- Naikoo, M. W., Rihan, M., Ishtiaque, M., & Shahfahad. (2020). Analyses of land use land cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using landsat datasets. *Journal of Urban Management*, 9(3), 347-359. <https://doi.org/10.1016/j.jum.2020.05.004>
- Nie, P., & Sousa-Poza, A. (2018). Commute time and subjective well-being in urban China. *China Economic Review*, 48, 188-204. <https://doi.org/10.1016/j.chieco.2016.03.002>
- Oyelade, O. J., Oladipupo, O. O., & Obagbuwa, I. C. (2010). Application of k-Means Clustering algorithm for prediction of Students' Academic Performance. *arXiv preprint arXiv:1002.2425*.
- Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: Positive and normative implementations of various accessibility indicators. *Journal of Transport Geography*, 25, 141-153. <https://doi.org/10.1016/j.jtrangeo.2012.03.016>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Proyectamos Colombia S.A.S. (2020). *Estudio sobre el impacto de la cuota monetaria del subsidio familiar en la población afiliada beneficiaria. Contrato N° 034 del 29 de septiembre de 2020*. <https://www.ssf.gov.co/documents/20127/1095844/final+final+Cuota+Monetaria+Ajustado+20210206.pdf/42e03871-0e7c-1e1a-d5d1-de4de58ba5ad>
- Ragheb, M. G., & Merydith, S. P. (2001). Development and validation of a multidimensional scale measuring free time boredom. *Leisure Studies*, 20(1), 41-59.

Recker, W. W., Chen, C., & McNally, M. G. (2001). Measuring the impact of efficient household travel decisions on potential travel time savings and accessibility gains. *Transportation Research Part A: Policy and Practice*, 35(4), 339-369. [https://doi.org/10.1016/S0965-8564\(99\)00062-2](https://doi.org/10.1016/S0965-8564(99)00062-2)

Ribeiro, J., Fontes, T., Soares, C., & Borges, J. L. (2021). Accessibility as an indicator to estimate social exclusion in public transport. *Transportation Research Procedia*, 52, 740-747. <https://doi.org/10.1016/j.trpro.2021.01.019>

Schmiedel, I., Bergmeier, E., & Culmsee, H. (2015). Plant species richness patterns along a gradient of landscape modification intensity in Lower Saxony, Germany. *Landscape and Urban Planning*, 141, 41-51. <https://doi.org/10.1016/j.landurbplan.2015.03.009>

Scott, D. M., & Horner, M. W. (2024). *The role of urban form in shaping access to opportunities*.

Stebbins, R. A. (2018). Leisure as not work: A (far too) common definition in theory and research on free-time activities. *World Leisure Journal*, 60(4), 255-264. <https://doi.org/10.1080/16078055.2018.1517107>

Stutzer, A., & Frey, B. S. (2008). Stress that Doesn't Pay: The Commuting Paradox*. *The Scandinavian Journal of Economics*, 110(2), 339-366. <https://doi.org/10.1111/j.1467-9442.2008.00542.x>

Van Acker, V. (2015). Defining, Measuring, and Using the Lifestyle Concept in Modal Choice Research. *Transportation Research Record: Journal of the Transportation Research Board*, 2495(1), 74-82. <https://doi.org/10.3141/2495-08>

Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, 7(5), 437-456.

Walters, M., & Bekker, J. (2017). CUSTOMER SUPER-PROFILING DEMONSTRATOR TO ENABLE EFFICIENT TARGETING IN MARKETING CAMPAIGNS. *South African Journal of Industrial Engineering*, 28(3). <https://doi.org/10.7166/28-3-1846>

Wang, F. (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Annals of the Association of American Geographers*, 102(5), 1104-1112. <https://doi.org/10.1080/00045608.2012.657146>

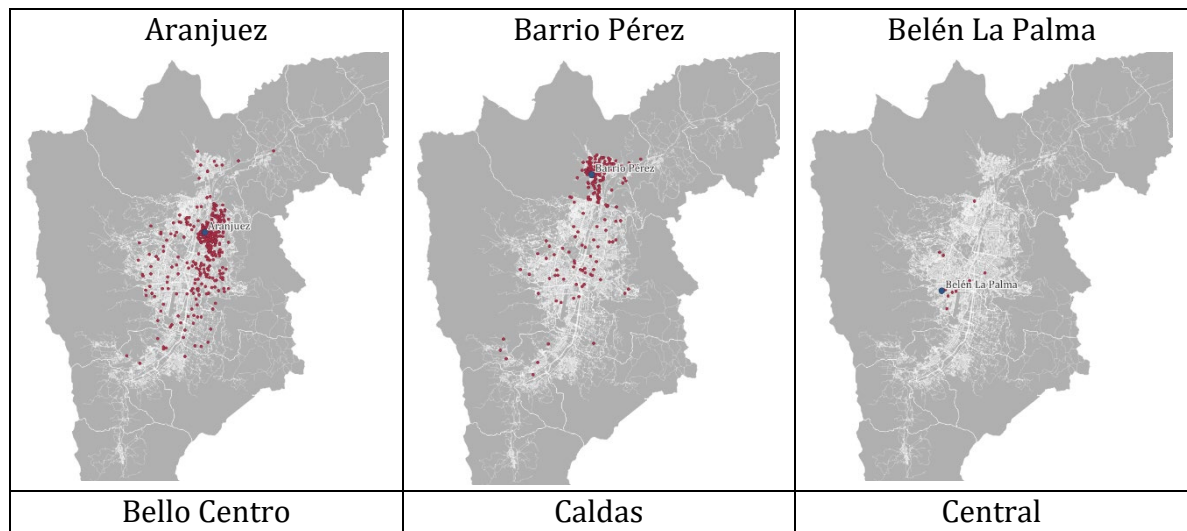
Wang, J., & Biljecki, F. (2022). Unsupervised machine learning in urban studies: A systematic review of applications. *Cities*, 129, 103925. <https://doi.org/10.1016/j.cities.2022.103925>

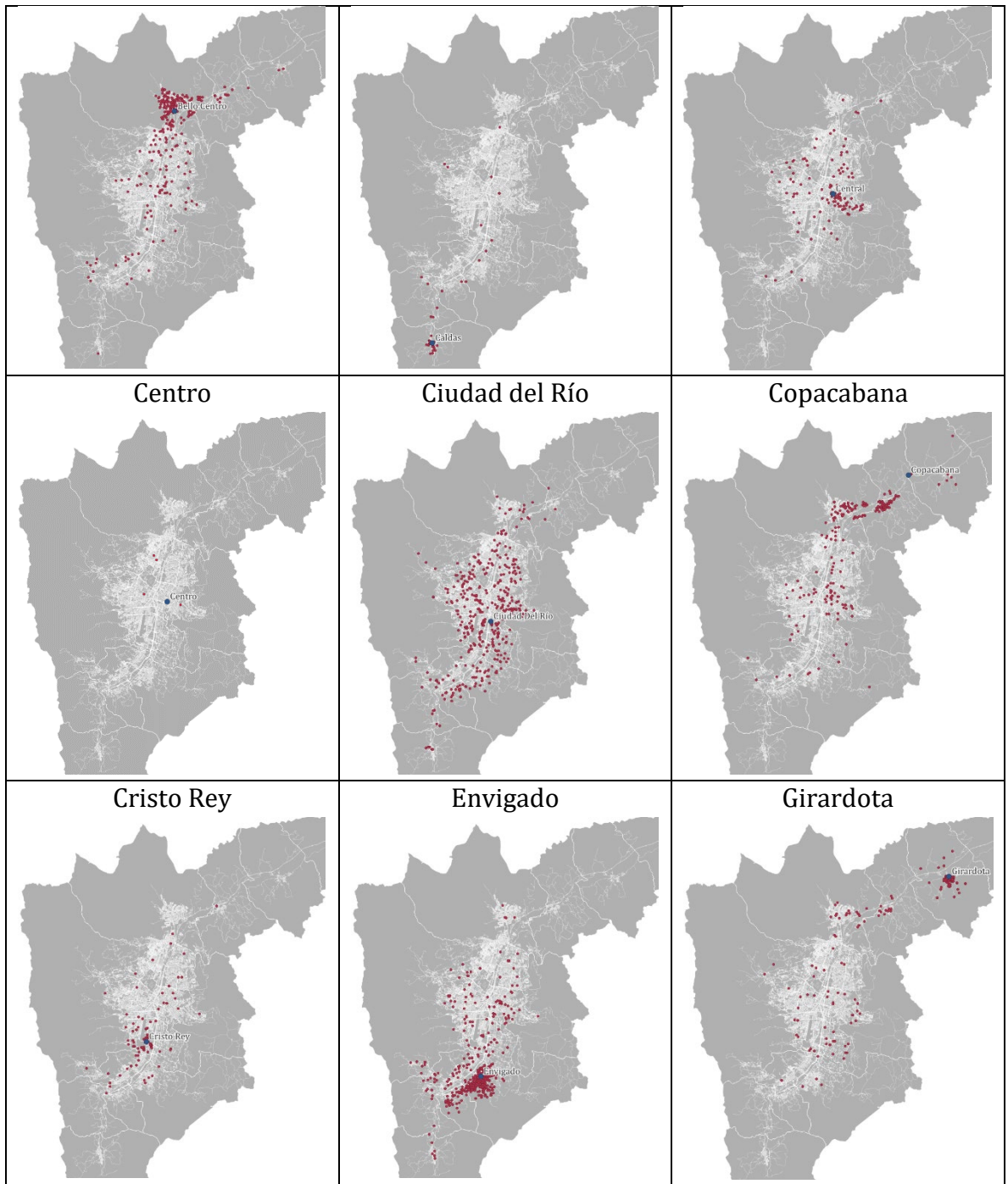
Wang, M. L., Lemon, S. C., Clausen, K., Whyte, J., & Rosal, M. C. (2016). Design and methods for a community-based intervention to reduce sugar-sweetened beverage consumption among youth: H2GO! study. *BMC Public Health*, 16(1), 1150. <https://doi.org/10.1186/s12889-016-3803-5>

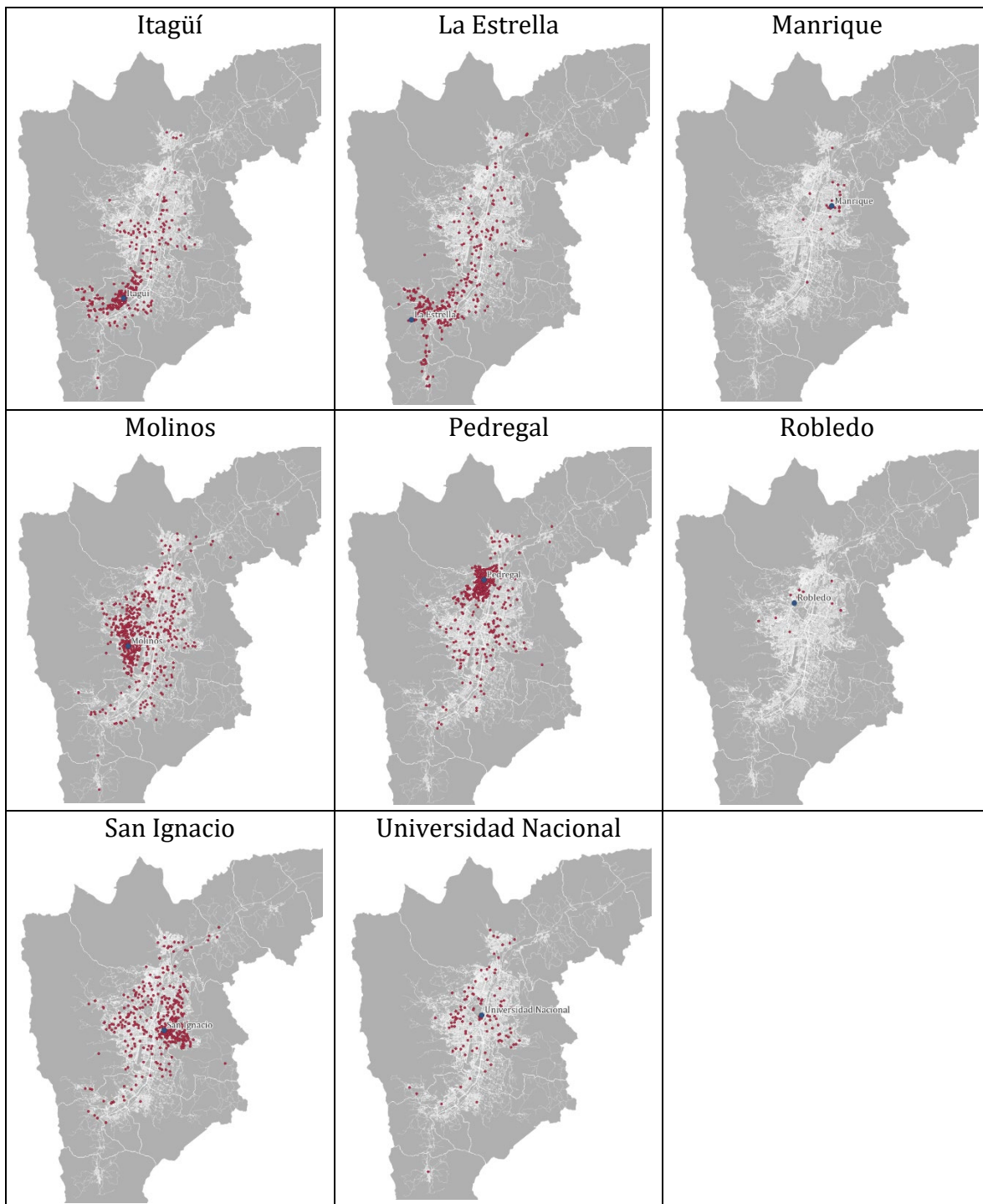
Wild, C. E. K., O’Sullivan, N. A., Lee, A. C., Cave, T. L., Willing, E. J., Cormack, D. M., Hofman, P. L., & Anderson, Y. C. (2020). Survey of Barriers and Facilitators to Engagement in a Multidisciplinary Healthy Lifestyles Program for Children. *Journal of Nutrition Education and Behavior*, 52(5), 528-534. <https://doi.org/10.1016/j.jneb.2019.10.010>

ANEXOS

ANEXO A. USUARIOS QUE PARTICIPAN EN LOS PUNTOS DE DESTINO







ANEXO B. SEGMENTACIÓN DE LOS USUARIOS POR VARIABLE

