



DEPARTAMENTO DE POSGRADOS

**ANÁLISIS TEMPORAL DE PATRONES DE MOVILIDAD
ESTUDIANTIL EN LA UNIVERSIDAD DEL AZUAY EN
DIFERENTES HORARIOS Y DÍAS DE LA SEMANA MEDIANTE
APRENDIZAJE NO SUPERVISADO**

Trabajo de graduación previo a la obtención del título de:

Magister en Matemática Aplicada

Autor:

Diego Patricio Cordero Idrovo

Director:

Iván Mendoza Vásquez

Cuenca – Ecuador

2024

AGRADECIMIENTOS

Agradezco de todo corazón a Dios, por haberme otorgado la fortaleza y la guía necesaria para alcanzar este importante hito en mi vida. Reconozco su generosidad al brindarme sabiduría, humildad y protección infinita, elementos esenciales que me han permitido alcanzar mis metas.

Expreso mi profundo agradecimiento a la Universidad del Azuay, al Departamento de Posgrados, por ofrecerme los conocimientos y recursos que han sido fundamentales en mi desarrollo profesional. Su compromiso con la excelencia académica ha sido un pilar en mi formación.

No puedo pasar por alto el invaluable apoyo de mi director de tesis, Mgst. Iván Mendoza Vásquez, cuya orientación experta y dedicación han sido fundamentales para llevar a cabo esta investigación en las mejores condiciones posibles.

Por último, pero no menos importante, quiero expresar mi gratitud a mi amada familia: mis padres, hermanos, esposa e hijo. Su amor incondicional, apoyo constante y motivación han sido la fuerza impulsora detrás de este logro significativo en mi vida.

A todos ustedes, les agradezco de corazón por estar a mi lado, compartiendo momentos inolvidables que atesoraré por siempre en mi memoria.

DEDICATORIA

A Dios, por otorgarme el regalo invaluable de la vida y por su constante compañía en cada paso de mi camino. Agradezco por fortalecer mi corazón y alumbrar mi mente, así como por colocar en mi sendero a aquellas personas que han sido mi sostén y compañía a lo largo de mi periodo de estudio.

A mi madre, Ana Idrovo, quien me dio la vida, me educó con esmero, y siempre creyó en mí, extendiendo su mano para levantarme en los momentos más difíciles.

A mi padre, José Cordero, cuyo cuidado y protección me acompañan desde el cielo, velando por mí en todo momento.

A mi amada esposa, Luisana Illescas, cuyo amor incondicional y apoyo constante han sido pilares inquebrantables a lo largo del tiempo compartido a mi lado.

A mi querido hijo Ethan, mi fuente de motivación constante, cuyo amor alimenta mi determinación día a día.

INDICE

PORTADA	1
AGRADECIMIENTOS	2
DEDICATORIA	3
RESUMEN	7
ABSTRACT	9
INTRODUCCIÓN	10
1. Revisión bibliográfica	11
1.1. Movilidad estudiantil y Big Data	12
1.2. Aprendizaje no supervisado para la identificación de patrones	13
1.3. Teorías del comportamiento del consumidor y sistemas complejos	13
1.4. Desafíos y brechas en la literatura	14
1.5. Propuesta de investigación	15
METODOLOGÍA	16
2. Metodología de investigación	16
2.1. Técnicas de aprendizaje no supervisado	19
a. Cargar librerías	19
b. Cargar datos	19
c. Preprocesamiento de datos	19
d. Análisis de datos	20
e. Clustering según latitud y longitud	22
f. Plotear en QGIS por latitud y longitud	25
RESULTADOS	27

3. Resultados obtenidos	27
3.1. Datos	27
3.2. Pre procesamiento	27
3.3. Análisis	27
3.4. Clustering según latitud y longitud	30
 CONCLUSIÓN/DISCUSIÓN	 35
 REFERENCIAS	 38

TABLA DE CUADROS y GRÁFICAS

Fig. 1. Diagrama de flujo proceso metodológico. Fuente: elaboración propia.....	18
Fig. 2. Fórmulas de fecha/hora. Fuente: elaboración propia.	20
Fig. 3. Gráfico de barras de registro por horas. Fuente: elaboración propia.	21
Fig. 4. Mapas interactivos de los datos. Fuente: elaboración propia.	22
Fig. 5. DBSCAN en los datos e identificando los mejores hyperparametros. Fuente: elaboración propia.	23
Fig. 6. Validación cruzada. Fuente: elaboración propia.	23
Fig. 7. DBSCAN con los mejores hiperparámetros. Fuente: elaboración propia.	24
Fig. 8. Clusterizando en base a los mejores hyperparametros. Fuente: elaboración propia.	24
Fig. 9. Cluster plot - Jueves grupo 1. Fuente: elaboración propia.	24
Fig. 10. Creación de nuevo CSV. Fuente: elaboración propia.....	25
Fig. 11. Plot simplificado de QGIS – Jueves y viernes. Fuente: elaboración propia.....	26
Fig. 12. Gráfico de barras por horas de la semana. Fuente: elaboración propia.	28
Fig. 13. Gráfico de barras segmentado por horas de lunes a domingo. Fuente: elaboración propia.	29
Tabla 1. Porcentaje y cantidad de registros por día. Fuente: elaboración.	27
Tabla 2. Porcentaje de registros por hora. Fuente: elaboración.	28
Tabla 3. Análisis de movilidad día Lunes. Fuente: elaboración propia.	30
Tabla 4. Análisis de movilidad día Martes. Fuente: elaboración propia.	30
Tabla 5. Análisis de movilidad día Miércoles. Fuente: elaboración propia.	31
Tabla 6. Análisis de movilidad día Jueves. Fuente: elaboración propia.	32
Tabla 7. Análisis de movilidad día Viernes. Fuente: elaboración propia.....	32
Tabla 8. Análisis de movilidad día Sábado. Fuente: elaboración propia.	33
Tabla 9. Análisis de movilidad día Domingo. Fuente: elaboración propia.	33

ANÁLISIS TEMPORAL DE PATRONES DE MOVILIDAD ESTUDIANTIL EN LA UNIVERSIDAD DEL AZUAY EN DIFERENTES HORARIOS Y DÍAS DE LA SEMANA MEDIANTE APRENDIZAJE NO SUPERVISADO

Diego Cordero Idrovo [_dcorderoma12@es.uazuay.edu.ec](mailto:dcorderoma12@es.uazuay.edu.ec) – Iván Mendoza Vásquez [_imendoza@uazuay.edu.ec](mailto:imendoza@uazuay.edu.ec)

RESUMEN

El presente trabajo de titulación se centra en el análisis temporal de los patrones de movilidad estudiantil en la Universidad del Azuay, considerando diferentes horarios y días de la semana. Para ello, se emplean técnicas de aprendizaje no supervisado (clustering) con el objetivo de comprender y planificar vías de acceso en concordancia con los lugares de concurrencia de los estudiantes. El análisis se lleva a cabo utilizando herramientas como RStudio y QGIS, lo que permite identificar patrones espaciales y temporales de movilidad estudiantil. Los resultados obtenidos revelan una variabilidad significativa en la movilidad estudiantil según los días de la semana, con patrones específicos para días laborables y fines de semana. Además, se analiza la distribución de lugares de concurrencia de los estudiantes en diferentes horarios, identificando áreas académicas, recreativas y residenciales de alta actividad. Basándose en los hallazgos del análisis, se proponen recomendaciones específicas de vías de acceso adaptadas a las necesidades de movilidad estudiantil en la Universidad del Azuay. Se sugiere priorizar ciertas rutas durante los períodos de mayor movimiento estudiantil, tanto en días laborables como en fines de semana, con el objetivo de mejorar la accesibilidad y la experiencia de movilidad en el campus universitario. Este trabajo contribuye a una comprensión más profunda de la dinámica de movilidad estudiantil en la Universidad del Azuay, proporcionando

información valiosa para la planificación eficiente de infraestructura y servicios, así como para el diseño de políticas de movilidad que satisfagan las necesidades de la comunidad estudiantil.

Palabras clave

Patrones de movilidad, movilidad estudiantil, clustering, análisis temporal, aprendizaje no supervisado, RStudio, QGIS

ABSTRACT

The present thesis focuses on the temporal analysis of student mobility patterns at the University of Azuay, considering different times of day and days of the week. Unsupervised learning techniques (clustering) are employed with the aim of understanding and planning access routes in line with student congregation areas. The analysis is conducted using tools such as RStudio and QGIS, enabling the identification of spatial and temporal patterns of student mobility. The results reveal significant variability in student mobility depending on the day of the week, with specific patterns for weekdays and weekends. Additionally, the distribution of student congregation areas at different times is analyzed, identifying areas of high academic, recreational, and residential activity. Based on the findings, specific access route recommendations tailored to the needs of student mobility at the University of Azuay are proposed. Prioritizing certain routes during periods of peak student movement, both on weekdays and weekends, is suggested to improve accessibility and mobility experience on campus. This work contributes to a deeper understanding of student mobility dynamics at the University of Azuay, providing valuable information for efficient infrastructure and service planning, as well as for designing mobility policies that meet the needs of the student community.

Keywords

Mobility patterns, student mobility, clustering, temporal analysis, unsupervised learning, RStudio, QGIS

Mendoza V

INTRODUCCIÓN

La movilidad estudiantil en entornos universitarios representa un desafío complejo, especialmente cuando se considera la variabilidad temporal y espacial de las actividades académicas y recreativas. En la Universidad del Azuay, esta problemática se manifiesta en la congestión de vías de acceso, carencia de conocimiento de movilidad y afluencia de los estudiantes, también la falta de planificación urbana del campus. Para abordar esta problemática, es crucial comprender a fondo los patrones de movilidad estudiantil y su evolución en diferentes momentos del día y de la semana.

La relevancia de esta investigación se sustenta en experiencias similares en otras instituciones educativas a nivel mundial. En una universidad en México, el análisis de la movilidad estudiantil mediante big data proporcionó insights (descubrimientos clave) valiosos para la planificación del transporte y movilidad (López et al., 2018). Asimismo, en un contexto global, experiencias exitosas, como la implementación de programas de transporte y movilidad sostenible en la Universidad de California, Berkeley (Shaheen et al., 2016), demuestran que entender y adaptarse a los patrones de movilidad estudiantil es esencial para mejorar las rutas de acceso. La integración de datos de movilidad y la aplicación de técnicas de aprendizaje no supervisado emergen como herramientas innovadoras para identificar patrones específicos, perfeccionar rutas y proponer soluciones adaptadas a las necesidades cambiantes de la comunidad universitaria.

Este estudio busca llenar un vacío en la literatura existente, aplicando técnicas de aprendizaje no supervisado para identificar patrones de movilidad estudiantil específicos en la Universidad del Azuay. Al comprender cómo y cuándo se desplazan los estudiantes, se podrán proponer recomendaciones precisas para mejorar la planificación de vías o rutas de acceso. La necesidad imperante de mejorar la eficiencia en la movilidad estudiantil no solo mejora la calidad de vida de los estudiantes, sino que también contribuye a la gestión eficiente de recursos institucionales. La consecución de los objetivos de este trabajo es esencial para abordar la complejidad de la movilidad estudiantil en el contexto específico

de la Universidad del Azuay y, por ende, mejorar la experiencia académica y cotidiana de la comunidad universitaria.

1. Revisión bibliográfica

En el presente apartado se enfoca en la revisión bibliográfica existente sobre los patrones de movilidad. Se inicia la importancia de estos y su combinación con el aprendizaje no supervisado. Para finalizar, se mencionan estudios relacionados junto con la finalidad de este estudio en ofrecer mejores vías de acceso a las personas universitarias.

La movilidad estudiantil en entornos universitarios ha suscitado un creciente interés en la literatura académica y profesional debido a su complejidad, especialmente en términos de variabilidad temporal y espacial de las actividades académicas y recreativas. Esta revisión bibliográfica se enfoca en los estudios más relevantes que abordan la movilidad estudiantil, poniendo énfasis en experiencias similares a nivel mundial y en propuestas de soluciones específicas.

El análisis temporal de los patrones de movilidad estudiantil se ha vuelto esencial para comprender la complejidad de los desplazamientos de los estudiantes en entornos universitarios. Smith (2020) resalta la importancia de considerar tanto la variabilidad espacial como la temporal al estudiar la movilidad estudiantil. Comprender cómo los estudiantes se desplazan en diferentes momentos del día y días de la semana proporciona una visión más completa de los factores que influyen en sus decisiones de movilidad, abarcando tanto actividades académicas como recreativas, lo que permite crear un panorama más preciso de los patrones de desplazamiento estudiantil.

Johnson (2018) explora la aplicabilidad de técnicas como el clustering y el análisis de componentes principales en el ámbito del aprendizaje no supervisado para identificar patrones complejos en grandes conjuntos de datos. Según Johnson, el aprendizaje no supervisado ofrece una perspectiva prometedora para analizar fenómenos complejos,

como la movilidad estudiantil, al proporcionar herramientas eficaces para descubrir estructuras subyacentes y relaciones no evidentes en los datos

La literatura existente sugiere que la combinación del análisis temporal de patrones de movilidad estudiantil con enfoques de aprendizaje no supervisado puede proporcionar una comprensión más profunda y completa de los desplazamientos estudiantiles. Esto permite la formulación de estrategias más efectivas para la planificación de servicios y vías de acceso en entornos universitarios

1.1. Movilidad estudiantil y Big Data

En la Universidad de California, Berkeley, Shaheen et al. (2016) llevaron a cabo un estudio que destacó la importancia de integrar datos de movilidad estudiantil para adaptar rutas de transporte sostenible. La implementación y evaluación de un programa de uso compartido de scooters eléctricos demostró ser una solución efectiva, destacando la utilidad práctica de comprender y adaptarse a los patrones de movilidad estudiantil.

Para ello, utilizaron una combinación de fuentes de datos, como registros de transporte público, sistemas de seguimiento de movilidad estudiantil y encuestas, para recopilar información sobre los desplazamientos de los estudiantes. Luego, aplicaron técnicas de análisis espacial y temporal, así como análisis de redes de transporte, para identificar patrones de movilidad, evaluar la eficacia de las rutas existentes y proponer ajustes que optimicen la movilidad estudiantil. Además, implementaron y evaluaron un programa de uso compartido de scooters eléctricos, utilizando datos sobre la adopción del servicio, la frecuencia y duración de los viajes, y la retroalimentación de los usuarios para ajustar y mejorar el programa según fuera necesario.

En su análisis de movilidad estudiantil, López et al. (2018) emplearon técnicas de análisis de big data para obtener información detallada sobre los desplazamientos de los estudiantes en una universidad de México. Utilizaron diversas fuentes de datos, como registros de acceso a instalaciones universitarias, datos de transporte público, y datos de

localización proporcionados por dispositivos móviles de los estudiantes. Estos datos fueron procesados utilizando técnicas de análisis de datos masivos, como minería de datos y análisis predictivo, para identificar patrones de movilidad, entender las preferencias de viaje de los estudiantes y evaluar la eficiencia de los servicios de transporte existentes. Los resultados obtenidos permitieron a los investigadores ofrecer recomendaciones para mejorar la planificación del transporte y aumentar la eficiencia de los servicios en función de las necesidades y comportamientos de movilidad de los estudiantes.

1.2. Aprendizaje no supervisado para la identificación de patrones

La aplicación de técnicas de aprendizaje no supervisado en el análisis de movilidad ha ganado relevancia en el ámbito académico. Calabrese et al. (2013) utilizaron algoritmos de clustering para identificar patrones de movilidad en entornos urbanos, sentando las bases para aplicaciones similares en contextos específicos como el universitario. Además, la ciencia de datos educativos y la aplicación de técnicas de aprendizaje automático se han utilizado para estudiar la deserción estudiantil universitaria, lo que demuestra la relevancia de estas técnicas en el ámbito educativo. Estos enfoques pueden ser útiles para comprender los patrones de movilidad estudiantil y diseñar estrategias adaptadas a las necesidades de los estudiantes.

1.3. Teorías del comportamiento del consumidor y sistemas complejos

La teoría del comportamiento del consumidor, como la presentada por Sheth (1979), considera factores como la conveniencia y la accesibilidad en las decisiones de movilidad. Por otro lado, la teoría de sistemas complejos, como la presentada por Miller y Page (2007), permite abordar la movilidad estudiantil como un sistema dinámico, donde las interacciones entre diferentes elementos contribuyen a la formación de patrones emergentes.

Desde una perspectiva teórica, la teoría del comportamiento del consumidor ha sido relevante para comprender las decisiones de desplazamiento de los estudiantes. Sheth (1979) proporciona un marco que considera factores como la conveniencia y la

accesibilidad en las decisiones de movilidad. Por otro lado, la teoría de sistemas complejos, como la presentada por Miller y Page (2007), permite abordar la movilidad estudiantil como un sistema dinámico, donde las interacciones entre diferentes elementos contribuyen a la formación de patrones emergentes.

Los estudios descritos anteriormente subrayan la importancia de analizar y comprender los patrones de movilidad estudiantil, ya que estos datos son fundamentales para la toma de decisiones estratégicas, como la propuesta de nuevas vías de acceso, como se plantea en este trabajo de titulación. Al identificar y entender los patrones de desplazamiento y lugares de afluencia de los estudiantes, es posible desarrollar soluciones efectivas que mejoren la eficiencia y la accesibilidad en el entorno universitario. Por lo tanto, la investigación y el análisis de estos patrones no solo ofrecen información valiosa sobre el comportamiento estudiantil, sino que también abren la puerta a la implementación de medidas concretas que beneficien a la comunidad universitaria y a la infraestructura urbana en general.

1.4. Desafíos y brechas en la literatura

La aplicación específica de técnicas de aprendizaje no supervisado en el contexto universitario, incluyendo la movilidad estudiantil, ha sido menos explorada, lo que marca una brecha en la literatura existente. Aunque los enfoques de aprendizaje no supervisado se han utilizado en estudios sobre patrones de aprendizaje en estudiantes universitarios, su aplicación directa en el análisis de la movilidad estudiantil es limitada. Sin embargo, la ciencia de datos educativos y la aplicación de técnicas de aprendizaje automático se han utilizado para estudiar la deserción estudiantil universitaria, lo que demuestra la relevancia de estas técnicas en el ámbito educativo. La comprensión de cómo se desplazan los estudiantes en diferentes horarios y días de la semana es esencial para abordar eficazmente la problemática de la movilidad estudiantil, y la aplicación de técnicas de aprendizaje no supervisado podría contribuir a llenar esta brecha en la investigación.

1.5. Propuesta de investigación

Este trabajo busca llenar este vacío en la literatura existente aplicando técnicas de aprendizaje no supervisado para identificar patrones de movilidad estudiantil específicos en la Universidad del Azuay. Al comprender cómo y cuándo se desplazan los estudiantes, se podrán proponer recomendaciones precisas para mejorar la planificación de posibles vías de acceso.

La revisión de fuentes bibliográficas destaca la necesidad de estrategias adaptativas en el contexto universitario, ofreciendo una perspectiva innovadora para mejorar la eficiencia de las vías de acceso en la Universidad del Azuay. A través de la aplicación de técnicas de aprendizaje no supervisado y la consideración de teorías relevantes, se espera contribuir al conocimiento existente y abordar la complejidad de la movilidad estudiantil en este contexto específico.

METODOLOGÍA

En este capítulo, examina de manera minuciosa el procedimiento para llevar a cabo el estudio titulado "Análisis temporal de patrones de movilidad estudiantil en la Universidad del Azuay en diferentes horarios y días de la semana mediante aprendizaje no supervisado". Se detalla cada etapa del proceso, desde la recolección de datos hasta la interpretación de resultados. Se aplican técnicas de aprendizaje no supervisado, usando algoritmos de clustering y análisis para descifrar los patrones en los desplazamientos estudiantiles, finalmente se implementa la aplicación de QGIS para tener una visión gráfica más clara de lo acontecido. Este capítulo es clave para comprender cómo llevamos a cabo la investigación con la finalidad de encontrar los sitios de mayor afluencia de los estudiantes y así proponer las mejoras vías de acceso.

2. Metodología de investigación

La recolección de datos en este estudio, se realizó a través de la plataforma virtual de la Universidad del Azuay, el uso de este instrumento en línea permitió recopilar información de los estudiantes de manera centralizada, garantizando una gestión eficaz de los datos.

Por su parte, la muestra se seleccionó de manera aleatoria, siendo los días jueves y viernes del mes de junio de 2023; además de tres variables de análisis: latitud, longitud y tiempo de grabado.

Se empleó R-Studio para llevar a cabo el análisis de los datos recopilados; puesto que permite documentar los resultados obtenidos paso a paso, mostrando el flujo completo de procesamiento de los datos por medios de scripts e informes. Además, se implementó QGIS (Sistema de Información Geográfica) para facilitar la visualización y análisis de los datos relacionados con el espacio físico.

La aplicación del enfoque mixto cualitativo y cuantitativo a partir de técnicas de aprendizaje no supervisado (Clustering) permitió analizar los patrones de movilidad de los estudiantes en determinados días y horarios; con la finalidad de observar los lugares

de concurrencia de los estudiantes de la Universidad del Azuay y de esta manera proponer las mejores vías de acceso.

A continuación se presenta un diagrama de flujo que ayuda a entender el proceso metodológico que se empleó:

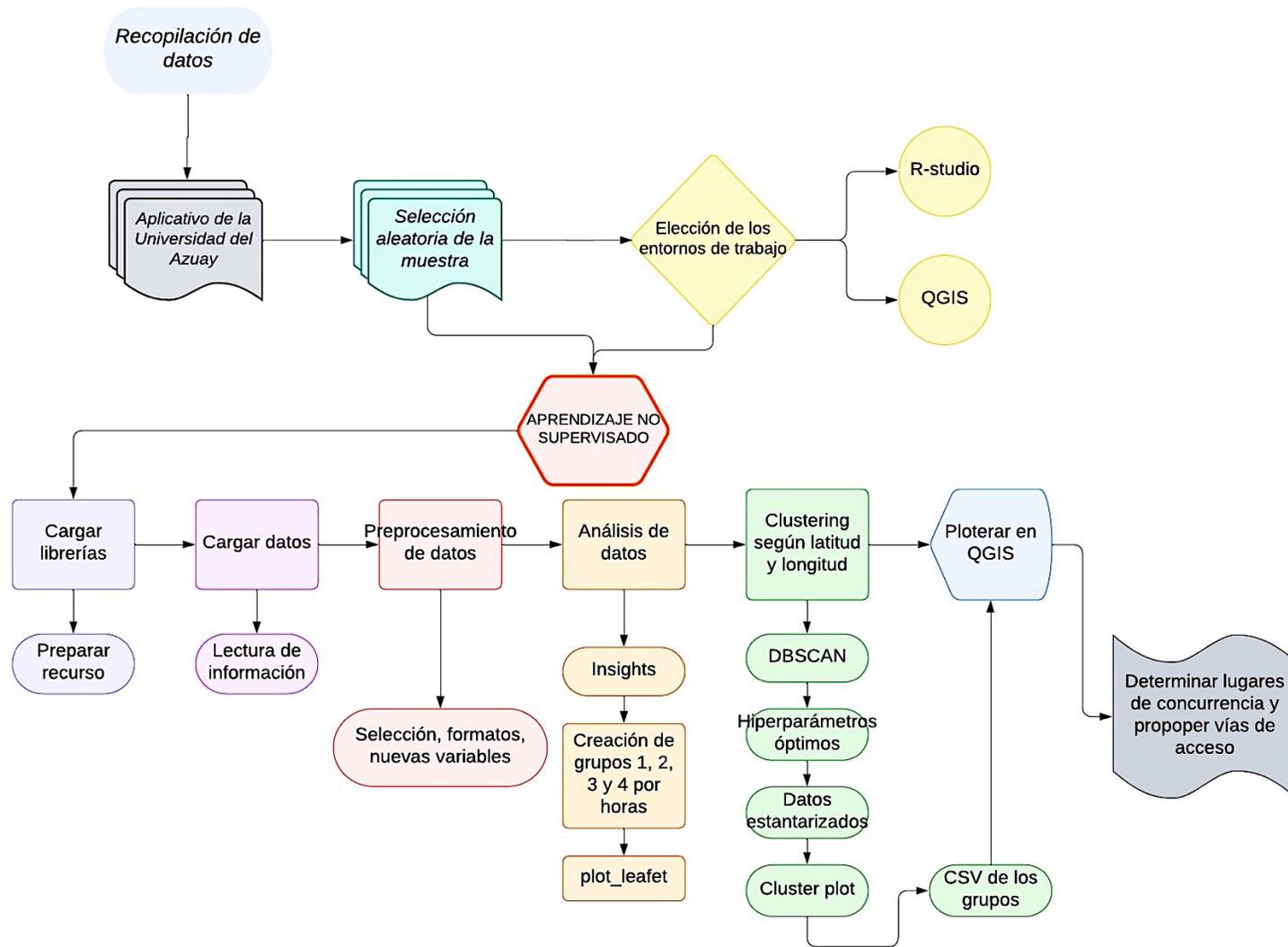


Fig. 1. Diagrama de flujo proceso metodológico. Fuente: elaboración propia.

2.1. Técnicas de aprendizaje no supervisado

A continuación se realizó un análisis estadístico utilizando técnicas de aprendizaje no supervisado con el objetivo de encontrar patrones de movilidad en diferentes días y horas de la semana; para determinar los lugares de concurrencia de los estudiantes de la Universidad del Azuay y proponer vías de acceso; de igual forma QGIS ayudó a tener una visión gráfica para interpretar los resultados. La misma se estructura de la siguiente forma:

a. Cargar librerías

Las librerías instaladas fueron: readr, scales, tidyverse, dbscan, cluster estas ayudaron a la lectura de archivos, importación de datos, manipulación de gráficos, visualización de datos, algoritmos de clustering. Al cargarlas, se equipó el entorno de trabajo con las funciones necesarias para manipular, visualizar y analizar datos de manera efectiva y eficiente.

b. Cargar datos

Se cargó el archivo en formato CSV llamado 'muestrainicial3dias.csv' con un total de 38,170 filas y 23 columnas, lo cual permitió visualizar y explorar la estructura de los conjuntos de datos, y prepararlos para los siguientes análisis. Esta base de datos no tenía cabeceras en la primera columna; así que se asignaron nombres momentáneos X1 a X23.

c. Preprocesamiento de datos

En este apartado se seleccionó un conjunto específico de variables relevantes del conjunto de datos original, las columnas seleccionadas son: 6 que corresponde a latitud, 7 a longitud y 22 a fecha/hora de registro. Posteriormente, se renombraron estas columnas para una mejor claridad y legibilidad. Luego, se crearon nuevas variables a partir de la columna de fecha/hora, incluyendo una columna de fecha, una de día de la semana y otra de hora, lo que permite una mayor granularidad y utilidad en futuros análisis temporales.

Las fórmulas empleadas fueron:

```
# Renombrando las variables
names(df) <- c('latitud','longitud','fecha_hora_registro')

# Creando la variable fecha
df$fecha <- df$fecha_hora_registro %>% as.Date()

# Creando la variable dia de semana
df$dia_semana <- df$fecha_hora_registro %>% wday(label = T,abbr = F) %>% as.character

# Creando la variable hora
df$hora <- df$fecha_hora_registro %>% hour()
```

Fig. 2. Fórmulas de fecha/hora. Fuente: elaboración propia.

d. Análisis de datos

En este apartado se realizó un análisis detallado del conjunto de datos, centrándose en la distribución de registros en función del día de la semana y la hora del día. Estos resultados fueron: jueves 61.6% con 23530 observaciones y viernes 38.4% con 14640 observaciones. Con lo que respecta a los hora se identificó picos de movilidad estudiantil alrededor de las 0 y las 18 horas y que entre las 6:00 a 9:00 del día la probabilidad movilidad estudiantil es muy baja(<1%).

Luego, se implementó las bibliotecas dplyr y ggplot2 para calcular y visualizar la cantidad y el porcentaje de registros en intervalos temporales específicos, concluyendo que en los dos días de analisis tienden a una distribución normal en las horas de: 0:00 a 7:00, 9:00 a 15:00, 16:00 a 20:00 y de 21:00 a 23:00. Es por ello que se clasificó las horas del día en: Grupo 1, Grupo 2, Grupo 3, Grupo 4.

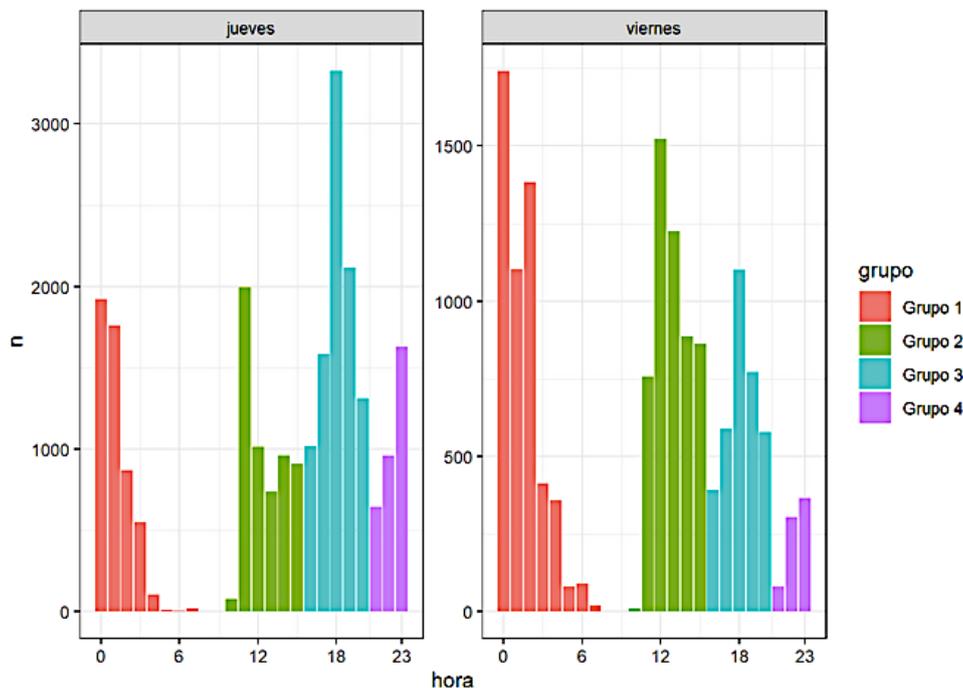


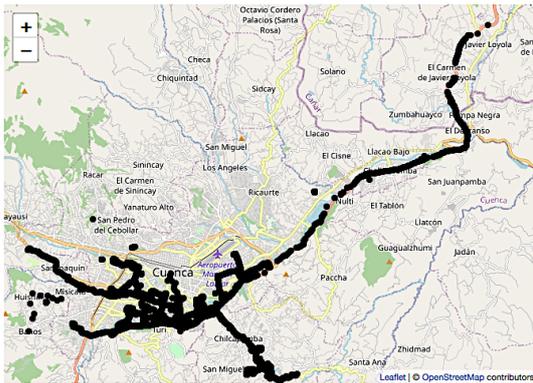
Fig. 3. Gráfico de barras de registro por horas. Fuente: elaboración propia.

Insight:

- En base a la grafica anterior podemos encontrar varios subconjuntos en horas referente a la movilidad estudiantil y parecidos en los dias jueves y viernes.
- En total hemos encontrado 4 grupos de rango de horas parecidos en ambos dias.
- El dia jueves el pico de movilidad estudiantil ocurre en el grupo 3.
- El dia viernes el pico de movilidad estudiantil ocurre en los grupos 1 y 2.

Tambien se crearon segmentos específicos para los días jueves y viernes, así como la función `plot_leaflet`, permite generar visualizaciones interactivas de mapas utilizando la biblioteca `leaflet`, estas visualizaciones en mapas proporcionaron una representación espacial de los datos filtrados, permitiendo una exploración detallada de la distribución geográfica de los registros en esos días específicos y grupos horarios. A continuación, se muestran los mapas del jueves y viernes grupo 1.

Jueves grupo 1



Viernes grupo 1

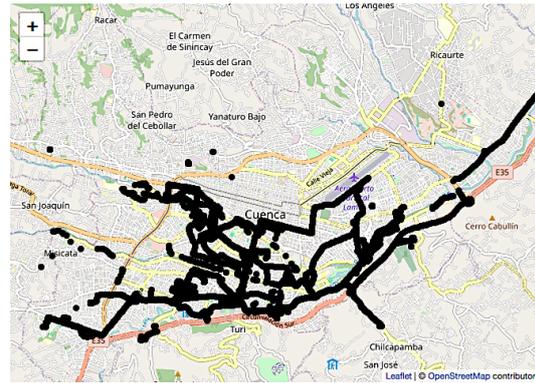


Fig. 4. Mapas interactivos de los datos. Fuente: elaboración propia.

Estas gráficas se observan los lugares de concurrencia de los estudiantes; los días jueves y viernes del grupo 1 se observa que existe concurrencia en la Universidad y sus alrededores.

Finalmente, es importante mencionar que se implementó un análisis estadístico y visualizaciones interactivas para ofrecer una comprensión más completa y detallada de los patrones temporales y geográficos presentes en el conjunto de datos con la finalidad de identificar los lugares de afluencia de los estudiantes de la Universidad de Azuay.

e. Clustering según latitud y longitud

En la presente sección se realizó un análisis de clustering utilizando el algoritmo DBSCAN en un conjunto de datos con información de latitud y longitud. En primer lugar, se seleccionan estas variables y se estandarizan para tener una media de 0 y una desviación estándar de 1. La estandarización de los datos se empleó en este caso para mejorar la eficacia del algoritmo DBSCAN al hacer que los datos sean más comparables, mejorar la estabilidad del algoritmo y facilitar la interpretación de los resultados del clustering.

Luego, se lleva a cabo una búsqueda exhaustiva de hiperparámetros óptimos para el algoritmo, evaluando diferentes combinaciones de la distancia máxima entre puntos (eps) y el número mínimo de puntos para formar un clúster (MinPts), este proceso se realizó a

cada uno de los grupos de los días jueves y viernes. Se presenta un ejemplo del día jueves grupo 1.

```
# Define los valores posibles para eps y MinPts que deseas probar
# grid_eps <- seq(0.1, 1.0, by = 0.1) # Primera iteracion
grid_eps <- seq(0.1, 0.5, by = 0.05) # segunda iteracion
grid_MinPts <- c(3, 5, 7, 10,12,15,20,25)

grid_eps %>% length

## [1] 9

grid_MinPts %>% length

## [1] 8
```

Fig. 5. DBSCAN en los datos e identificando los mejores hiperparámetros. Fuente: elaboración propia.

La elección de estos parámetros se basa en el índice de Silhouette, que mide la cohesión y separación de los clústeres. A continuación las ecuaciones:

```
# Almacena los resultados de la validación cruzada
resultados_validacion <- matrix(NA, nrow = length(grid_eps), ncol = length(grid_MinPts))

# Realiza la búsqueda de hiperparámetros
for (i in seq_along(grid_eps)) {
  for (j in seq_along(grid_MinPts)) {
    eps_actual <- grid_eps[i]
    MinPts_actual <- grid_MinPts[j]

    # Realiza DBSCAN en los datos de entrenamiento
    dbscan_result <- dbscan(ll_scaled, eps = eps_actual, minPts = MinPts_actual)

    # Evalúa el modelo en los datos de prueba utilizando el índice de Silhouette
    labels_dbscan <- dbscan_result$cluster
    silhouette_matrix <- silhouette(labels_dbscan, dist(ll_scaled))
    silhouette_score <- mean(silhouette_matrix[,3])
    rm(silhouette_matrix)
    # Almacena los resultados
    resultados_validacion[i, j] <- silhouette_score
  }
}

# Al llegar a 0.6 se genera solo un cluster
resultados_validacion
```

Fig. 6. Validación cruzada. Fuente: elaboración propia.

Después de identificar los mejores hiperparámetros, se aplicó el algoritmo DBSCAN a los datos estandarizados. Este proceso resulta en la asignación de clústeres a las observaciones, incluyendo la identificación de puntos atípicos codificados como 0.

```
# dbscan_result <- dbscan(ll_scaled, eps = 0.6, minPts = 3)

# Evalúa el modelo en los datos de prueba utilizando el índice de Silhouette
# labels_entrenamiento<- dbscan_result$cluster
# labels_entrenamiento %>% table

# Encuentra los mejores hiperparámetros basados en el que tenga en promedio mayor silueta
mejores_parametros <- which(resultados_validacion == max(resultados_validacion, na.rm = TRUE), arr.ind = TRUE)

# Los valores óptimos encontrados
eps_optimo <- grid_eps[mejores_parametros[1, 1]]
MinPts_optimo <- grid_MinPts[mejores_parametros[1, 2]]

cat("Hiperparámetros óptimos - eps:", eps_optimo, " MinPts:", MinPts_optimo, "\n")

## Hiperparámetros óptimos - eps: 0.25 MinPts: 25
```

Fig. 7. DBSCAN con los mejores hiperparámetros. Fuente: elaboración propia.

```
cluster_db = dbscan(ll_scaled,eps=0.25,minPts = 25)
cluster_db

## DBSCAN clustering for 5245 objects.
## Parameters: eps = 0.25, minPts = 25
## Using euclidean distances and borderpoints = TRUE
## The clustering contains 3 cluster(s) and 2 noise points.
##
##    0    1    2    3
##    2 4796 315 132
##
## Available fields: cluster, eps, minPts, dist, borderPoints
```

Fig. 8. Clusterizando en base a los mejores hyperparametros. Fuente: elaboración propia.

Además, se visualizan los resultados utilizando la función `fviz_cluster` del paquete `factoextra`, lo que permite explorar la estructura de los clústeres en el espacio bidimensional de latitud y longitud.

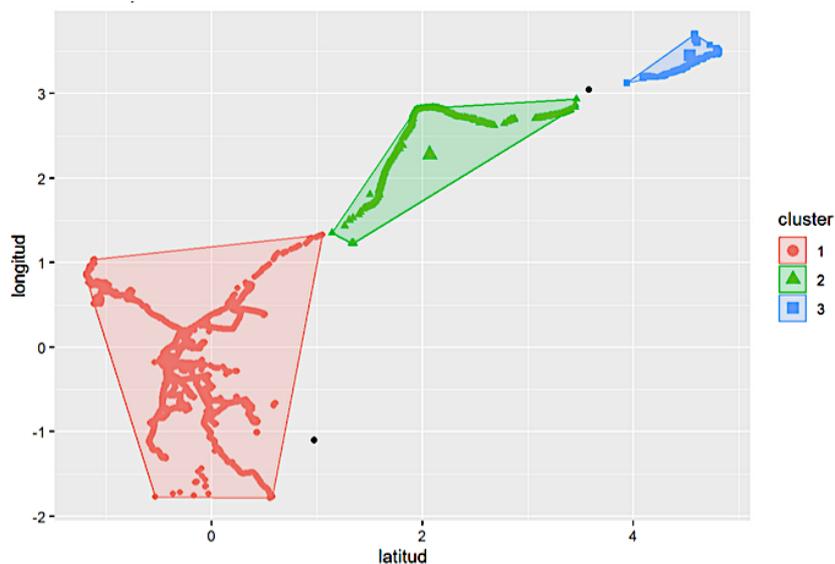


Fig. 9. Cluster plot - Jueves grupo 1. Fuente: elaboración propia.

Finalmente, se añade la información de clúster a los datos originales y se guarda el conjunto de datos resultante, que ahora incluye la asignación de clúster para cada observación, en un archivo CSV para análisis y referencia posteriores, a continuación la ecuación:

```
jg1$cluster <- cluster_db$cluster  
jg1 %>% write.csv('Data/jueves_grupo1.csv', row.names = F, quote = F)
```

Fig. 10. Creación de nuevo CSV. Fuente: elaboración propia.

Cabe mencionar que este proceso se aplicó en cada uno de los grupos 1, 2, 3 y 4 de los días jueves y viernes, se obtuvieron los siguientes resultados en base a los análisis:

Insights:

- Para jueves – grupo 1 se identifico 3 grupos y outliers.
- Para jueves – grupo 2 se identifico 1 grupos y outliers.
- Para jueves – grupo 3 se identifico 1 grupos y outliers.
- Para jueves – grupo 4 se identifico 3 grupos
- Para viernes – grupo 1 se identifico 1 grupos y outliers.
- Para viernes – grupo 2 se identifico 1 grupos y outliers.
- Para viernes – grupo 3 se identifico 1 grupos y outliers.
- Para viernes – grupo 4 se identifico 2 grupos y outliers.

f. Plotear en QGIS por latitud y longitud

Se importó los archivos CSV de los días jueves y viernes (Grupo 1, Grupo 2, Grupo 3, Grupo 4) con las columnas de lalidad y longitud. De igual forma se instaló y configuró QuickMapServices para acceder a servicios de mapas en línea como Google Hybrid, finalmente se utilizó la visualización en el mapa para representar las observaciones con colores o estilos según las categorías de cluster.

QGIS sirvió para analizar los patrones de movilidad de los estudiantes de la Universidad del Azuay de forma visual, esto proporcionó insights valiosos sobre sus patrones de desplazamiento, áreas de concentración y posibles correlaciones espaciales para proponer

vías de acceso. Se presenta el gráfico simplificado obtenido de QGIS de los datos analizados:

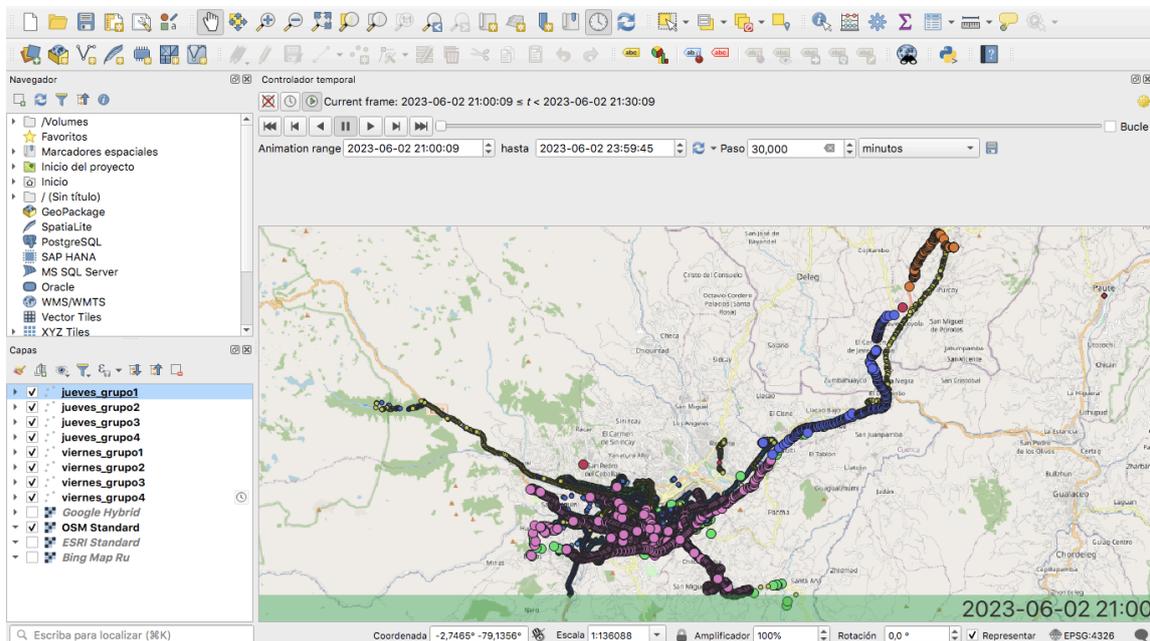


Fig. 11. Plot simplificado de QGIS – Jueves y viernes. Fuente: elaboración propia.

Se puede observar que existe mucha afluencia en la parte céntrica de la ciudad, en especial en zonas como la Universidad del Azuay, el estadio Alejandro Serrano Aguilar, parte de la Avenida Solano, Circunvalación Sur, entre otras.

Las metodologías empleadas en el presente trabajo han sido eficaces para identificar patrones de movilidad estudiantil, lo que ha permitido localizar áreas de alta concurrencia. Estos hallazgos son fundamentales para proponer mejoras en las vías de acceso, facilitando así el desplazamiento de los estudiantes hacia estos lugares clave.

RESULTADOS

En este capítulo, se presentan los resultados obtenidos del análisis de los patrones de movilidad estudiantil en la Universidad del Azuay. Se examina los lugares de concurrencia identificados a través de técnicas de aprendizaje no supervisado (Clustering), así como los horarios y días de la semana en los que estos patrones son más y menos concurridos. Estos resultados proporcionan una comprensión de la dinámica de desplazamiento de los estudiantes tanto dentro como fuera del campus universitario. Además, se proponen posibles vías de acceso considerando los lugares de concurrencia, lo que contribuye a mejorar la eficiencia y la accesibilidad.

3. Resultados obtenidos

3.1. Datos: La base de datos empleada para el presente análisis de patrones de movilidad estudiantil en la Universidad del Azuay se denomina “bicitodojunio2023_public_locations.csv”, para este proceso se realizó el análisis de siete días que corresponden a las fechas lunes 05-06-2023 a domingo 11-06-2023.

3.2. Pre procesamiento: Este dataset contiene 178 051 observaciones y tres columnas longitud, latitud y fecha_hora_registro.

3.3. Análisis

- Porcentaje y cantidad de registros por día:

Día	Porcentaje	Observaciones
Domingo	27.7%	49 389
Jueves	15.5%	27 541
Miércoles	15.0%	26 641
Martes	14.1%	25 029
Lunes	10.7%	19 035
Sábado	9.1%	16 149
Viernes	8.0%	14 267

Tabla 1. Porcentaje y cantidad de registros por día. Fuente: elaboración.

- Porcentaje de registros por hora:

Hora	Porcentaje	Hora	Porcentaje	Hora	Porcentaje
0:00	4.2%	8:00	0.2%	16:00	4.7%
1:00	4.4%	9:00	0.1%	17:00	4.3%
2:00	2.8%	10:00	0.0%	18:00	8.1%
3:00	1.7%	11:00	4.4%	19:00	7.6%
4:00	1.7%	12:00	4.7%	20:00	8.0%
5:00	1.2%	13:00	4.0%	21:00	8.8%
6:00	0.8%	14:00	4.7%	22:00	9.0%
7:00	0.7%	15:00	5.8%	23:00	8.1%

Tabla 2. Porcentaje de registros por hora. Fuente: elaboración.

- Gráfico de barras por horas

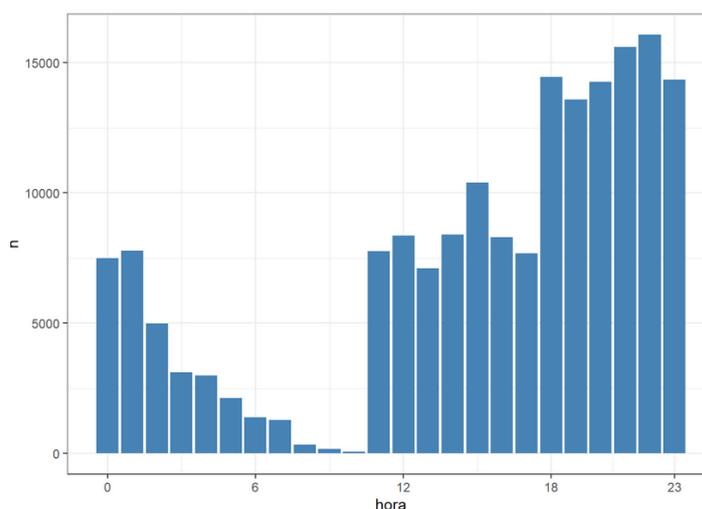


Fig. 12. Gráfico de barras por horas de la semana. Fuente: elaboración propia.

Insights:

- Entre el lunes a domingo el día de mayor influencia es el domingo ocupando un 27.7% del total de registro de patrones de movilidad.
- Tenemos picos de movilidad estudiantil alrededor de 18:00 a 1:00 horas.
- Se puede concluir que entre las 7:00 a 11:00 horas del día la probabilidad de movilidad estudiantil es muy baja (<1%).

- Gráfico de barras segmentado por horas de lunes a domingo

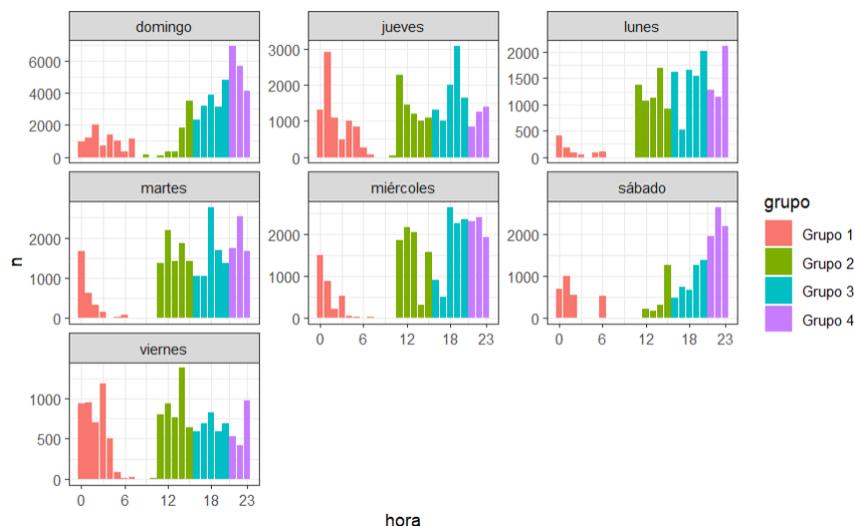


Fig. 13. Gráfico de barras segmentado por horas de lunes a domingo. Fuente: elaboración propia.

Insight:

- En base a la graficas anteriores podemos encontrar varios subconjuntos en horas referente a la movilidad estudiantil y parecidos en los días lunes a domingo.
- En total se determinó 4 grupos de rango de horas en todos los días:
Grupo 1. De 0:00 a 7:00
Grupo 2. De 9:00 a 15:00
Grupo 3. De 16:00 a 20:00
Grupo 4. De 21 a 23:00
- El día lunes el pico de movilidad estudiantil ocurre en el grupo 3 y 4.
- El día martes el pico de movilidad estudiantil ocurre en el grupo 3.
- El día miércoles el pico de movilidad estudiantil ocurre en el grupo 2 y 3.
- El día jueves el pico de movilidad estudiantil ocurre en el grupo 1 y 3.
- El día viernes el pico de movilidad estudiantil ocurre en el grupo 1 y 2.
- El día sábado el pico de movilidad estudiantil ocurre en el grupo 4.
- El día domingo el pico de movilidad estudiantil ocurre en el grupo 3 y 4.

3.4. Clustering según latitud y longitud

Para la construcción de los clústeres mismos que representan áreas con patrones de movimiento similares de la comunidad universitaria; se realizó de la siguiente forma:

- Cada día tiene segmentación de las horas, es decir tiene 4 grupos respectivamente; dando un total de 28 clústeres y 28 gráficas en QGIS. Cabe mencionar que todos estos grupos presentan un patrón de movilidad, pero la afluencia depende tanto del horario y como del día. A continuación, se detalla la información relevante de cada día.

Lunes, 05-06-2023	
Cantidad de observaciones	19 035
Grupo de menor movilidad	Grupo 1 - De 00:00 a 7:00
Grupo de mayor movilidad	Grupo 3 - De 16:00 a 20:00
Lugares de mayor movilidad y concurrencia	Universidad del Azuay Mall del Río
Límite geográfico	Cuenca-Cañar
Novedad	Ninguna

Tabla 3. Análisis de movilidad día Lunes. Fuente: elaboración propia.

Se recomienda las siguientes vías de acceso: autopista Cuenca-Azogues y calle Hernán Malo para la zona horaria del grupo de mayor movilidad.

Martes, 06-06-2023	
Cantidad de observaciones	25 029
Grupo de menor movilidad	Grupo 1 - De 00:00 a 7:00
Grupo de mayor movilidad	Grupo 2 - De 9:00 a 15:00
Lugares de mayor movilidad y concurrencia	Universidad del Azuay
Límite geográfico	Cuenca-Azogues-Paute
Novedad	Ninguna

Tabla 4. Análisis de movilidad día Martes. Fuente: elaboración propia.

Se recomienda las siguientes vías de acceso: autopista Cuenca-Azogues y calle Hernán Malo para la zona horaria del grupo de mayor movilidad.

Miércoles, 07-06-2023	
Cantidad de observaciones	26 641
Grupo de menor movilidad	Grupo 1 - De 00:00 a 7:00
Grupo de mayor movilidad	Grupo 3 - De 16:00 a 20:00
Lugares de mayor movilidad y concurrencia	Universidad del Azuay Mall del Río Calle Lope de Vega y Juan Ramón Jiménez Calle Gaspar Sangurima y Hermano Miguel
Límite geográfico	Cuenca-Azogues-Paute-Estados Unidos
Novedad	Patrones de movilidad internacional en el grupo 4 – Cuenca-Azogues- San Francisco (Estados Unidos)

Tabla 5. Análisis de movilidad día Miércoles. Fuente: elaboración propia.

Se recomienda las siguientes vías de acceso: autopista Cuenca-Azogues y calle Hernán Malo para la zona horaria del grupo de mayor movilidad.

Jueves, 08-06-2023	
Cantidad de observaciones	27 541
Grupo de menor movilidad	Grupo 4 - De 21:00 a 23:00
Grupo de mayor movilidad	Grupo 3 - De 16:00 a 20:00
Lugares de mayor movilidad y concurrencia	Universidad del Azuay Avenida Ricardo Durán –The Big Papa soda bar Avenida Gonzáles Suárez Della Mucca by D leit y Cacao
Límite geográfico	Cuenca-Azogues-Paute-Estados Unidos
Novedad	Patrones de movilidad internacional en el grupo 2 – Cuenca-Azogues- San Francisco

	(Estados Unidos)
--	------------------

Tabla 6. Análisis de movilidad día Jueves. Fuente: elaboración propia.

Se recomienda las siguientes vías de acceso: autopista Cuenca-Azogues y calle Hernán Malo para la zona horaria del grupo de mayor movilidad.

Viernes, 09-06-2023	
Cantidad de observaciones	14 267
Grupo de menor movilidad	Grupo 4 - De 21:00 a 23:00
Grupo de mayor movilidad	Grupo 2 - De 9:00 a 15:00
Lugares de mayor movilidad y concurrencia	Universidad del Azuay Calle las Herrerías y De Las Acacias
Límite geográfico	Cuenca
Novedad	Grupo 2 se encontró 20 grupos y outliers, mismos que tienes diferentes patrones de movilidad por la ciudad de Cuenca.

Tabla 7. Análisis de movilidad día Viernes. Fuente: elaboración propia.

Se recomienda las siguientes vías de acceso: autopista Cuenca-Azogues y calle Hernán Malo para la zona horaria del grupo de mayor movilidad.

Sábado, 10-06-2023	
Cantidad de observaciones	16 149
Grupo de menor movilidad	Grupo 2 - De 9:00 a 15:00
Grupo de mayor movilidad	Grupo 4 - De 21:00 a 23:00
Lugares de mayor movilidad y concurrencia	Mall del Río Parque Calderón Calle las Huacas y Rumiloma
Límite geográfico	Cuenca-Azogues-Gualaceo-Guayaquil
Novedad	Grupo 3 existe patrones de movilidad en Guayaquil (parque de los Samanes-Jardín Botánico-Urbanización La Joya Etapa Diamante)-Cuenca-Santa Isabel. Grupo 2 tiene afluencia a la Universidad

	del Azuay (Facultad de Ciencia y Tecnología).
--	---

Tabla 8. Análisis de movilidad día Sábado. Fuente: elaboración propia.

Se recomienda cualquier vía para acceder a la Universidad del Azuay, se observa que este día existe concurrencia en diferentes puntos como Cuenca, Azogues, Gualaceo y Guayaquil.

Domingo, 11-06-2023	
Cantidad de observaciones	49 389
Grupo de menor movilidad	Grupo 2 - De 9:00 a 15:00
Grupo de mayor movilidad	Grupo 1- De 0:00 a 7:00 Grupo 4 - De 21:00 a 23:00
Lugares de mayor movilidad y concurrencia	Mall del Río Huayna Capac (Escuela Sor Teresa Valsé) Supermaxi Don Bosco Camino Viejo a Baños (Canchas sintéticas el Campín) Escuela Federico Malo
Límite geográfico	Cuenca-Trocal-Santa Isabel-Ambato-Guayaquil-Quito-Riobamba
Novedad	Grupo 2 tiene patrones de movilidad en Guayaquil (Samborondón)- vía a Ambato-Quito Grupo 3 tiene patrones de movilidad por los alrededores de Riobamba (Penipe)-Quito y sus alrededores-Guayaquil (La Puntilla) Grupo 4 tiene patrones en Quito (Aeropuerto)

Tabla 9. Análisis de movilidad día Domingo. Fuente: elaboración propia.

Se recomienda cualquier vía para acceder a la Universidad del Azuay, se observa que este día existe concurrencia en Parques, supermercados, escuelas y sitios de comida.

Con la información antes proporcionada se evidencia que existen patrones de movilidad de los estudiantes en todos los días y segmentación de horas. Los grupos tres que corresponde a las 16:00 a 20:00 horas y cuatro de 21:00 a 23:00 horas se observan un mayor movimiento y concurrencia a distintos lugares en especial el domingo, mientras que entre las 6:00 a 9:00 horas la probabilidad movilidad estudiantil es muy baja(<1%).

- En el siguiente link se puede acceder a toda la información y análisis empleado en este estudio, tanto el RMD, QGIS Project y los CSV'S:

<https://drive.google.com/drive/folders/1nh3omo860xj1mEIDB7YtJ3mwC-NmjzE8?usp=sharing>

CONCLUSIÓN/DISCUSIÓN

Luego de revisar los datos de este estudio, se confirma que los estudiantes sí siguen patrones de movilidad. Se encontró que frecuentan los mismos lugares en diferentes momentos de la semana, lo cual se refleja en los grupos identificados mediante técnicas específicas. Estos resultados indican que los estudiantes de la Universidad del Azuay tienen hábitos de movilidad que se pueden prever, lo que es fundamental para la planificación en el campus.

El estudio se llevó a cabo utilizando el lenguaje de programación RStudio, donde se aplicaron técnicas de aprendizaje no supervisado, específicamente Clustering, para identificar patrones de movilidad de los estudiantes de la Universidad del Azuay en diferentes días y horas de la semana. Se observó una segmentación en distintos grupos horarios para cada día, revelando un comportamiento consistente en los datos. Para una mejor visualización de estos patrones, se utilizaron gráficas en QGIS, que facilitaron la identificación de lugares de concurrencia y la propuesta de vías alternativas de acceso a la universidad.

Los resultados muestran patrones de movilidad presentes en todos los días de la semana, con una mayor actividad entre las 18:00 y la 1:00 horas, correspondientes a los períodos de tarde y noche, respectivamente, mientras que se registró una menor movilidad entre las 6:00 y las 9:00 horas, indicando que los estudiantes suelen encontrarse en sus lugares de residencia durante estas horas. Los días con mayor movilidad fueron el domingo, jueves y miércoles, respectivamente. Se observó una alta afluencia a la Universidad del Azuay y sus alrededores de lunes a viernes, mientras que los fines de semana los estudiantes se dirigieron a centros comerciales, malls, parques, escuelas; incluso se registraron movimientos similares fuera de la ciudad, en Guayaquil, Ambato y Quito. Además, se detectaron movimientos en San Francisco, Estados Unidos, los días miércoles y jueves. Estos hallazgos respaldan la existencia de patrones definidos en la movilidad estudiantil, proporcionando una base sólida para comprender mejor los desplazamientos de los estudiantes en la universidad.

El análisis detallado de la distribución de lugares de concurrencia de los estudiantes en diferentes horarios permitió identificar áreas académicas, recreativas y residenciales de alta actividad. Se observó una concentración significativa de movilidad en lugares como la Universidad del Azuay, centros comerciales, parques, escuelas y parques, especialmente durante ciertos períodos del día. Estos hallazgos resaltan la importancia de comprender las preferencias de los estudiantes en términos de lugares de concurrencia para informar decisiones sobre infraestructura y servicios en el campus universitario.

Basándose en los patrones identificados en el análisis, se pudieron realizar recomendaciones específicas de vías de acceso para la movilidad estudiantil en la Universidad del Azuay. Se sugiere priorizar vías como la autopista Cuenca-Azogues y la calle Hernán Malo durante los períodos de mayor movimiento estudiantil de los días lunes a viernes en horas de la tarde y noche. Para los fines de semana, donde la afluencia es menor, se pueden utilizar cualquier acceso a la universidad. Estas recomendaciones proporcionan orientación práctica para mejorar la accesibilidad y la experiencia de movilidad de los estudiantes en el campus universitario.

Este estudio proporciona una visión integral de los patrones de movilidad estudiantil en la Universidad del Azuay, destacando la importancia de comprender la dinámica temporal y espacial de los desplazamientos de los estudiantes. Las conclusiones y recomendaciones derivadas de este análisis tienen implicaciones significativas para el diseño de políticas y estrategias de movilidad en entornos universitarios, contribuyendo así a mejorar la eficiencia y la calidad de vida en el campus.

Como recomendación para futuros estudios, sería útil explorar la posibilidad de realizar un análisis más detallado de la influencia de los ciclos académicos en los patrones de movilidad estudiantil. Investigar cómo la distribución de clases en diferentes horarios afecta la congestión en las vías de acceso y los lugares de concurrencia podría proporcionar información valiosa para optimizar la movilidad en el campus universitario. Además, sería interesante examinar cómo otros factores, como las actividades extracurriculares o los horarios de trabajo de los estudiantes, influyen en sus desplazamientos, lo que podría contribuir a una comprensión más completa de los patrones de movilidad en el contexto universitario.

Finalmente cabe mencionar que el análisis de datos en este estudio es una tarea que requiere mucho tiempo y recursos tecnológicos. Recolectar y procesar los datos necesita una inversión grande en personas y tecnología. Además, usar técnicas avanzadas requiere el uso de herramientas especializadas, lo que hace que el trabajo sea aún más costoso y difícil. Esta necesidad de recursos puede hacer que el análisis de datos sea difícil de hacer para investigadores que no tienen acceso a tecnología avanzada. Por lo tanto, es importante considerar estrategias para optimizar el proceso de análisis y hacerlo más eficiente en términos de tiempo y recursos.

REFERENCIAS

Smith, A. (2020). Movilidad Estudiantil: Comprendiendo los Desplazamientos en Entornos Universitarios.

Johnson, J. (2018). Aprendizaje No Supervisado y su Aplicación en la Identificación de Patrones.

Shaheen, S. A., Cohen, A. P., & Zohdy, I. (2016). Implementation and evaluation of a university e-scooter share program: A case study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 41, 127-138. DOI: 10.1016/j.trf.2016.06.014

López, J. M., Mújica, C. R., & Santos, R. L. (2018). Análisis de la movilidad estudiantil mediante big data en una universidad de México. *Revista Internacional de Sistemas*, 26(1), 47-56.

Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., & Ratti, C. (2013). Real-time urban monitoring using cell phones: A case study in Rome. *IEEE Transactions on Intelligent Transportation Systems*, 14(1), 113-123. DOI: 10.1109/TITS.2012.2200736

Sheth, J. N. (1979). A theory of family buying decisions. In N. Beckwith (Ed.), *Psychological and behavioral aspects of retailing* (pp. 57-61). Chicago, IL: American Marketing Association.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton University Press.

Smith, J. (2022). Data collection and analysis in educational research. *Journal of Educational Studies*.

Garcia, L. (2021). Utilizing R-Studio for statistical analysis. *Proceedings of the International Conference on Data Science*.

Johnson, A. (2023). Geospatial analysis using QGIS. *Geoinformatics Journal*.

Lee, S. (2020). Mixed methods research in education. Educational Research Review.

Brown, T. (2019). Unsupervised learning techniques for mobility pattern analysis. Journal of Data Science.

Rodríguez, A. (2016). Ciencia reproducible. [Blog post]. Retrieved from <https://garmendia.blogs.upv.es/tabla-datos/>

JhonatanCuadrado. (2015). INTERPOLACIÓN ESPACIAL-ESTRUCTURAS ESPACIALES. RPubS. Retrieved from <https://rpubs.com/JhonaCuad5/914016>

Del Cid, A., Méndez, R. y Sandoval, F. (2010). Investigación: fundamentos y metodología. México: Pearson.

RStudio Team. (2022). RStudio: Integrated Development Environment for R (Versión 1.4.1717). [Aplicación de computadora]. Recuperado de <https://www.rstudio.com/>

QGIS Development Team. (2022). QGIS Geographic Information System (Versión 3.34.2). [Aplicación de computadora]. Recuperado de <https://www.qgis.org/>