**Análisis de las Tendencias de Uso de las Bicicletas de Capital Bikeshare en Washington D.C.**

Un grupo de personas en un parque de diversiones

Descripción generada automáticamente con confianza baja

By: Esteban Cid Solaun

Thebridgeschool

# Introducción

El uso de sistemas de bicicletas compartidas,ha experimentado un crecimiento notable en los últimos años en muchas ciudades, convirtiéndose en una alternativa ecológica y accesible para la movilidad urbana.

Sin embargo, la demanda y el uso de estos sistemas están influenciados por distintos factores como el clima, el tipo de usuario, la infraestructura de la ciudad, etc. Comprender cómo estos factores impactan el uso de bicicletas es crucial para optimizar la planificación de recursos, mejorar la satisfacción de los usuarios y maximizar la eficiencia del sistema.

Este trabajo tiene como objetivo analizar el comportamiento de los usuarios de Capital Bikeshare, una empresa que pone bicicletas de alquiler al servicio de los usuarios en la ciudad de Washington DC. El análisis se enfoca en tres aspectos principales:

1. **Impacto del Clima en el Uso de las Bicicletas:** Determinar cómo variables climáticas como la temperatura, la lluvia y el viento afectan tanto la cantidad de viajes como la duración de los mismos. Esta información puede orientar estrategias para adaptar el servicio en función de las condiciones climáticas, o identificar los mejores días para realizar labores de mantenimiento, por ejemplo.
2. **Patrones de Uso Semanales:** Examinar las variaciones en la frecuencia de uso de bicicletas a lo largo de la semana. Este análisis permite identificar los días de mayor y menor demanda, facilitando la asignación de recursos y posibles campañas de promoción específicas.
3. **Análisis de Rutas y Estaciones:** Analizar qué zonas son más concurridas dependiendo del día de la semana y del tipo de cliente. Estos conocimientos pueden mejorar la distribución de bicicletas y el diseño de estaciones, haciendo el sistema más accesible y eficiente, además de orientar en la elaboración de promociones específicas.

En resumen, este trabajo busca proporcionar una base de datos valiosa para la gestión y planificación de Capital Bikeshare, ofreciendo una visión integral de los factores que impulsan su uso y pueden servir para optimizar su operación en Washington D.C.

# Obtención de los datos

El dataset utilizado para este trabajo se descargó de [Kaggle](https://www.kaggle.com/datasets/taweilo/capital-bikeshare-dataset-202005202408). El conjunto de datos de Capital Bikeshare contiene información detallada sobre el uso diario de bicicletas compartidas en Washington D.C., abarcando desde mayo de 2020 hasta agosto de 2024. Incluye cuatro archivos principales:

* Datos de alquileres diarios: contiene el tipo de bicicleta, fechas y horas de inicio y fin, estaciones de inicio y destino, y el tipo de usuario (casual o miembro).
* Lista de estaciones: incluye el nombre e identificación de cada una
* Frecuencia de uso en cada estación: indica la cantidad de alquileres y devoluciones en cada estación diariamente
* Datos climáticos: incluye información climática muy detallada para cada día. La información utilizada en este análisis fue la relativa a temperatura, precipitaciones y viento.

Los dataframes se pueden descargar en este [enlace](https://wetransfer.com/downloads/ce2ac5b3eea91ba4a02ed9e64997876f20241111194857/aec4b6c216136628f1c066938510134320241111194913/04651b?t_exp=1731613737&t_lsid=b7166a3a-0aff-4b18-aa3a-e0d1662712a9&t_network=email&t_s=download_link&t_ts=1731354553). Tiene una validez de 3 días desde el 11/11/2024.

# Metodología

Este análisis se realizó íntegramente utilizando Python junto con distintas librerías útiles para el análisis de datos y la visualización, principalmente pandas, numpy, matplotlib, seaborn y folium.

## Filtrado y limpieza de los datos

El dataset original abarca desde mayo de 2020 hasta agosto de 2024. Con el fin de simplificar el procesamiento, los datos se acotaron al año 2023, y se pusieron los datos de fechas en formato datetime.

En el registro de alquileres diarios, había algunas entradas en las que faltaba el nombre e id de la estación. Al disponer de las coordenadas, y de la lista completa de estaciones, se podían rellenar esos datos. Sin embargo, al final se optó por eliminar esos registros ya que representaban un porcentaje muy pequeño respecto al total, y el tiempo de procesamiento de rellenar los nulos era muy elevado.

También se calculó la duración de cada trayecto, y se eliminaron los que no estaban comprendidos entre 2 y 120 minutos, ya que se consideró que podrían ser errores a la hora del desbloqueo o devolución de la bicicleta.

Al eliminar registros de la tabla de alquileres diarios, hubo que rehacer de nuevo la tabla de frecuencia de uso en cada estación, para que fuera en consonancia con los datos del estudio.

## Análisis general para aportar contexto

El análisis inicial ofrece una visión general del uso de bicicletas y las condiciones climáticas en Washington D.C., proporcionando un contexto fundamental que sustenta el análisis posterior.

A través de transformaciones de tablas y diversas visualizaciones de datos, como líneas de tendencia, histogramas y gráficos de barras y tarta, se examinaron varios aspectos clave: la cantidad promedio de alquileres diarios, la duración promedio de cada trayecto, y el porcentaje de alquileres clasificados por tipo de bicicleta y tipo de usuario.

Además, se exploraron patrones climáticos característicos de la ciudad de Washington D.C., evaluando la temperatura promedio, la frecuencia e intensidad de las precipitaciones y las condiciones de viento.

## Impacto del clima en el uso del sistema

Se exploraron las relaciones entre variables climáticas (como la temperatura, la intensidad de la lluvia y la velocidad del viento) y el comportamiento de los usuarios en cuanto a la duración y frecuencia de los viajes.

Para facilitar el análisis, los días lluviosos se clasificaron en categorías según la intensidad de precipitación[[1]](#footnote-1):

* Lluvia ligera: 0-2 mm/h
* Lluvia moderada: 2-15 mm/h
* Lluvia fuerte: 15-30 mm/h
* Lluvia muy fuerte: 30-60 mm/h

De manera similar, se clasificó el viento en función de su velocidad:

* Brisa suave: 0-19 km/h
* Brisa moderada: 19-34 km/h
* Brisa fuerte: 34-49 km/h
* Viento fuerte: 49-60 km/h

Es importante señalar que el dataset original no proporcionaba las unidades de velocidad del viento ni de precipitación. Sin embargo, tras realizar búsquedas pertinentes en fuentes externas, se determinó que las unidades de velocidad se ajustaban mejor a km/h y las de precipitación a mm/h.

Para estudiar el impacto de cada variable climática, se emplearon gráficos de dispersión, boxplots y gráficos de barras, analizando cómo cada una de estas variables influía tanto en la cantidad de alquileres diarios como en la duración de los trayectos. Finalmente, se calculó la correlación entre cada variable climática y ambas métricas de uso (cantidad de alquileres y duración de los trayectos).

Este análisis permitió identificar el impacto de las condiciones climáticas en el comportamiento de los usuarios del sistema de bicicletas, revelando patrones que reflejan cómo el clima influye en la decisión de utilizar el servicio y en la duración de los trayectos.

## Cambios en los patrones de uso según el día de la semana y tipo de usuario

Por último, se estudiaron las variaciones en las rutas y en el uso de bicicletas según el día de la semana, diferenciando entre usuarios casuales y miembros.

Para ello, se empleó la librería Folium para crear mapas de calor en el área de Washington y sus alrededores, resaltando las zonas con mayor concentración de alquileres.

Primero, se visualizaron dos mapas: uno para los días entre semana y otro para los fines de semana. A pesar de ser bastante similares, se observaron diferencias sutiles en las áreas de mayor actividad, con un aumento significativo del uso en zonas turísticas o más agradables para pasear durante los fines de semana, especialmente a lo largo del National Mall, donde se encuentran el Capitolio, el Monumento a Lincoln y el Monumento a Washington.

Para confirmar estos patrones, se generaron mapas de calor separados para miembros y usuarios casuales. Los resultados mostraron que el mapa de los días entre semana coincidía notablemente con el de los miembros, mientras que el mapa de fines de semana era muy similar al de los usuarios casuales.

Además, se visualizaron las 300 rutas más comunes entre miembros y las 300 rutas más populares entre usuarios casuales, lo que permitió evidenciar aún más la diferencia entre desplazamientos de rutina, probablemente hacia el trabajo, y aquellos de ocio o turismo.

Este enfoque facilitó la identificación de patrones de uso que pueden informar decisiones para optimizar los recursos, distinguiendo entre viajes recreativos y desplazamientos habituales. Comprender cómo y cuándo cada tipo de usuario emplea el servicio permite ajustar la distribución y disponibilidad de bicicletas en función de la demanda, mejorando así la eficiencia y experiencia del sistema de bicicletas compartidas.

# Conclusiones

* **Sensibilidad al clima:** La temperatura y, en menor medida, las precipitaciones tienen un impacto significativo en el número de alquileres y la duración de los trayectos, mientras que el viento no es un factor determinante a la hora de coger la bici (al menos en una ciudad con los niveles de viento de Washington DC, en ciudades muy ventosas como por ejemplo Chicago quizás sea diferente).
* **Uso diversificado:** El sistema de bicicletas compartidas de Capital Bikeshare es ampliamente utilizado tanto para desplazamientos cotidianos como para actividades recreativas, atendiendo a una amplia gama de necesidades de movilidad. Se observan patrones de uso diferenciados entre los usuarios miembros y los usuarios casuales. Los miembros tienden a concentrar su actividad en áreas residenciales y zonas de oficinas como el Downtown, reflejando un uso frecuente como medio de transporte diario. En cambio, los usuarios casuales muestran una mayor actividad en áreas turísticas y parques, especialmente a lo largo del National Mall, lo que sugiere un uso recreativo o turístico del servicio. Esta distinción en los patrones permite adaptar estrategias de distribución y disponibilidad de bicicletas según el perfil del usuario y la ubicación, además de la elaboración de campañas de marketing y ofertas específicas según el tipo de uso
* **Diferentes patrones de uso según los distintos días:** Los días entre semana tienen un patrón de uso que coincide con el de los usuarios miembros, lo que reafirma la suposición de que éstos realizan un uso del servicio para transporte diario (ir al trabajo, recados, etc.). Los fines de semana el uso de bicicletas sigue un patrón similar al de los usuarios casuales. Las bicicletas se utilizan más para uso recreativo, como paseos o turismo.

1. https://www.eltiempo.es/noticias/precipitacion-cuando-es-poco-y-cuando-es-mucho [↑](#footnote-ref-1)