# **Google data analytics**

# **Caso de estudio: Cyclistic**

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

De La Rivera Silva Gabriel

## **ASK (Preguntar)**

### **1. Declaración clara de la tarea de negocio**

El objetivo principal de este análisis es responder a la siguiente pregunta:

**¿Cómo utilizan de forma diferente las bicicletas los miembros anuales y los ciclistas ocasionales de Cyclistic?**

Esta información será clave para que el equipo de marketing de Cyclistic diseñe estrategias que puedan aumentar la conversión de ciclistas ocasionales a miembros anuales, mejorando la retención y aumentando los ingresos de la empresa.

### **2. Descripción de los interesados (stakeholders)**

* **Moreno**, el gerente de marketing, encargado de diseñar campañas para aumentar la membresía anual.
* **Clientes de Cyclistic**, especialmente los ciclistas ocasionales, quienes son el segmento objetivo para conversión.

### **3. Preguntas que guían el análisis**

* ¿Qué diferencias hay en el uso de bicicletas entre miembros anuales y ciclistas ocasionales?
* ¿En qué días de la semana y horarios utilizan más las bicicletas cada grupo?
* ¿Qué meses presentan mayor uso para cada segmento?
* ¿Qué patrones de uso pueden aprovecharse para mejorar la conversión de clientes ocasionales a miembros?

## **PREPARE (Preparar)**

### **1. Fuentes de datos utilizadas**

Para este análisis se utilizó el conjunto de datos históricos de viajes de Cyclistic, una empresa ficticia de bicicletas compartidas. Los datos corresponden a los últimos 12 meses de viajes registrados entre agosto de 2024 y julio de 2025.

Los archivos contienen registros de cada viaje con detalles como:

* Identificador del viaje
* Tipo de bicicleta
* Fecha y hora de inicio y fin del viaje
* Estación de inicio y fin
* Coordenadas geográficas
* Tipo de usuario (miembro anual o ciclista ocasional)

Estos datos son públicos y fueron obtenidos desde Google Cloud Storage, en formato CSV. No incluyen información personal identificable para proteger la privacidad de los usuarios.

### **2. Organización y estructura de los datos**

Cada archivo CSV corresponde a un mes y contiene miles de registros. Los archivos fueron descargados y luego combinados en una sola tabla para facilitar el análisis.

Las columnas principales usadas son:

| **Columna** | **Descripción** |
| --- | --- |
| ride\_id | Identificador único del viaje |
| rideable\_type | Tipo de bicicleta utilizada |
| started\_at | Fecha y hora de inicio del viaje |
| ended\_at | Fecha y hora de fin del viaje |
| start\_station\_name | Nombre estación de inicio |
| end\_station\_name | Nombre estación de destino |
| member\_casual | Tipo de usuario (miembro o casual) |

### **3. Consideraciones sobre calidad y credibilidad de los datos**

* Los datos provienen de una fuente pública y confiable.
* Se revisó la integridad y consistencia para detectar registros con datos erróneos, como viajes con duración negativa o estaciones vacías.
* Se respetó la privacidad al no usar datos personales.
* Los datos son representativos del comportamiento real de los usuarios de bicicletas compartidas en el periodo analizado.

### **4. Preparación previa a análisis**

* Los archivos CSV fueron subidos a Google Cloud Storage y luego importados a BigQuery para su manejo eficiente.
* Se combinaron los 12 archivos mensuales en una tabla única para facilitar consultas y análisis.

## **PROCESS (Procesar)**

Para preparar los datos para el análisis, utilicé Google BigQuery por su capacidad para manejar grandes volúmenes de datos de forma eficiente.

**Herramientas utilizadas:**

* Google Cloud Storage para almacenar los archivos CSV.
* Google BigQuery para importar, combinar, limpiar, transformar y consolidar los datos.

**Proceso de limpieza y manipulación:**

1. **Eliminación de registros inválidos:**
   1. Se eliminaron 279 registros correspondientes al mes de julio de 2024 que no formaban parte del periodo de análisis.
   2. Se eliminaron 132,441 registros con duración negativa o cero minutos, ya que no representan datos válidos para el análisis.
   3. Se descartaron 1,640,730 filas con estaciones de inicio o fin vacías o nulas, que impedían analizar rutas y patrones geográficos correctamente.
   4. Se eliminaron 144 registros con viajes mayores a 24 horas (1440 minutos), para evitar distorsiones en el análisis.
2. **Agregado de columnas calculadas:**
   1. Se añadió la duración del viaje en minutos (ride\_length\_minutes), calculada con la diferencia entre el tiempo de fin y el tiempo de inicio del viaje, redondeada a 2 decimales para mayor precisión.
   2. Se extrajo el día de la semana (day\_of\_week) de la fecha de inicio del viaje, para analizar patrones diarios.
   3. Se extrajo el mes (month) de la fecha de inicio, para analizar patrones mensuales.
3. **Creación de tabla consolidada para análisis:**
   1. Se creó una nueva tabla llamada bicicletas\_procesadas que contiene solo las columnas esenciales para el análisis, descartando datos redundantes o innecesarios, para optimizar la consulta y facilitar la interpretación de resultados.
4. **Verificación de integridad:**
   1. Se realizaron consultas para asegurar que los datos estén limpios y consistentes, verificando que no existan registros con valores negativos o nulos en columnas clave.

Todo este proceso fue documentado y realizado mediante sentencias SQL en BigQuery, asegurando transparencia y reproducibilidad.

## **ANALYSIS (Análisis)**

### **1. Objetivo del análisis**

El objetivo principal de este análisis es identificar y entender las diferencias en el uso de bicicletas entre los miembros anuales y los ciclistas ocasionales de Cyclistic. Este conocimiento permitirá diseñar estrategias de marketing más efectivas para convertir ciclistas ocasionales en miembros anuales, mejorando la retención y aumentando los ingresos de la empresa.

### **2. Preguntas específicas abordadas**

* ¿Cuáles son las diferencias en la duración promedio de los viajes entre miembros anuales y ciclistas ocasionales?
* ¿Qué tipos de bicicletas prefieren cada grupo?
* ¿En qué días de la semana y horarios utilizan más las bicicletas cada segmento?
* ¿Qué meses presentan mayor uso para cada tipo de usuario?
* ¿Existen patrones de uso distintivos que puedan ayudar a enfocar las campañas de marketing para aumentar la conversión?

### **3. Metodología**

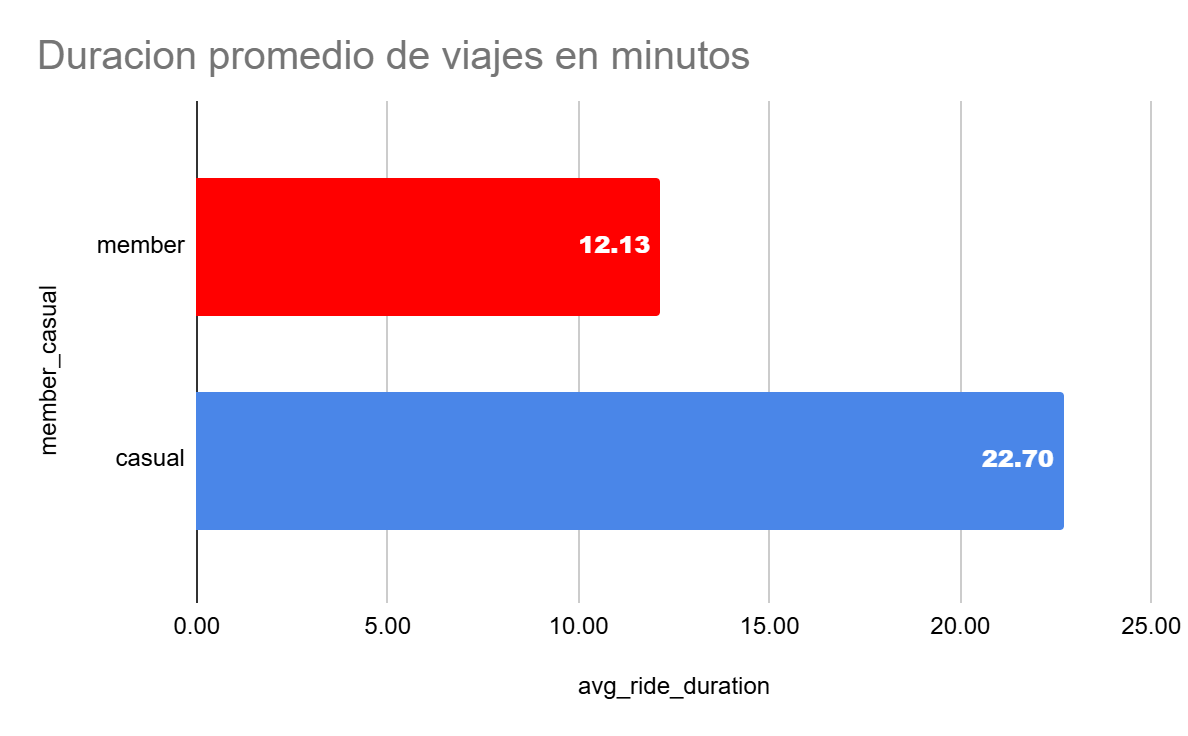
Para responder estas preguntas se realizaron consultas SQL en Google BigQuery que permitieron calcular métricas descriptivas agrupadas por tipo de usuario (member\_casual), día de la semana, tipo de bicicleta, hora del día y mes.

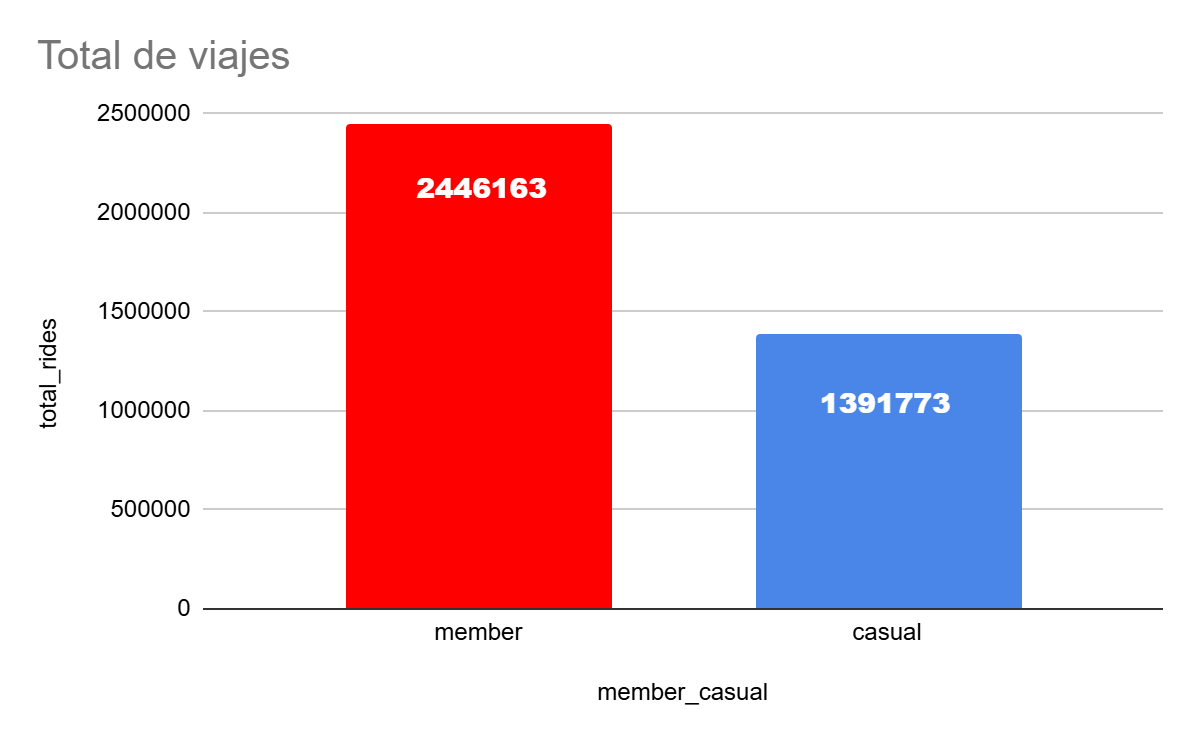
Los datos se limpiaron previamente eliminando registros inválidos (duración negativa o cero, estaciones vacías, viajes mayores a 24 horas).

Los resultados de las consultas fueron exportados a Google Sheets para su análisis exploratorio y visualización, facilitando la identificación de patrones y tendencias relevantes para el negocio.

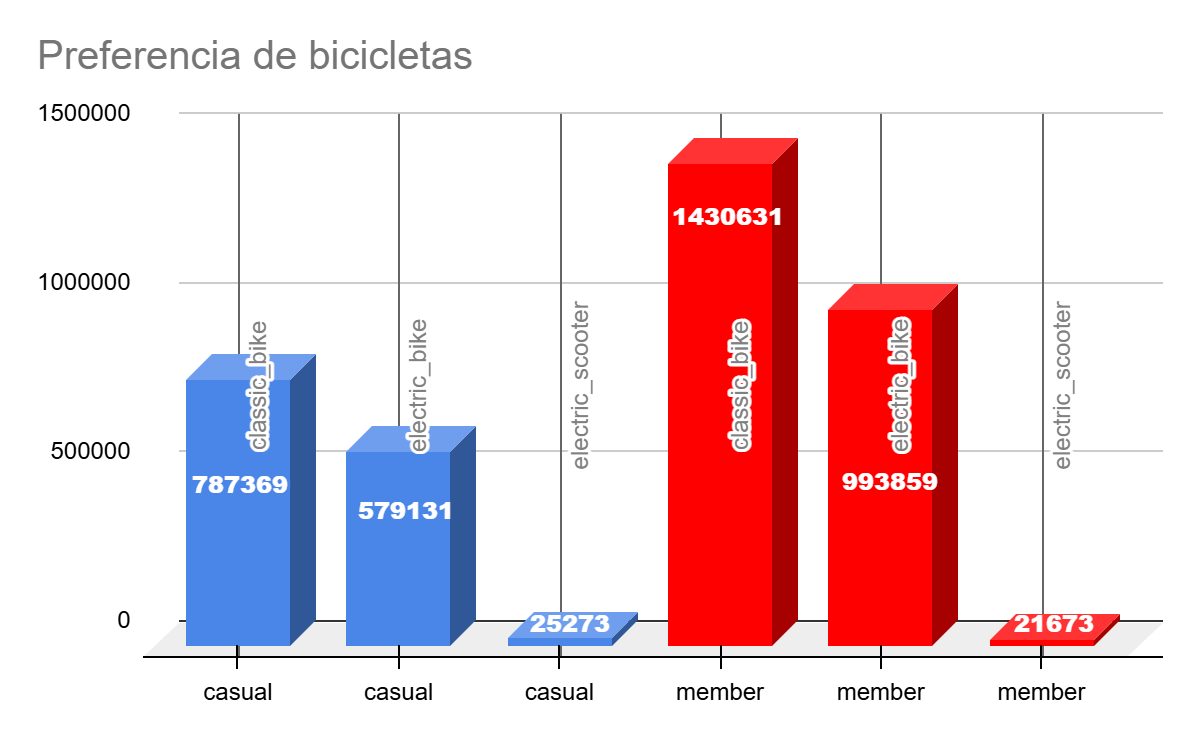
### **4. Resultados clave y hallazgos**

* La duración promedio de los viajes es de **12.13 minutos para miembros anuales** y **22.70 minutos para ciclistas casuales**, indicando que los miembros hacen viajes más cortos.

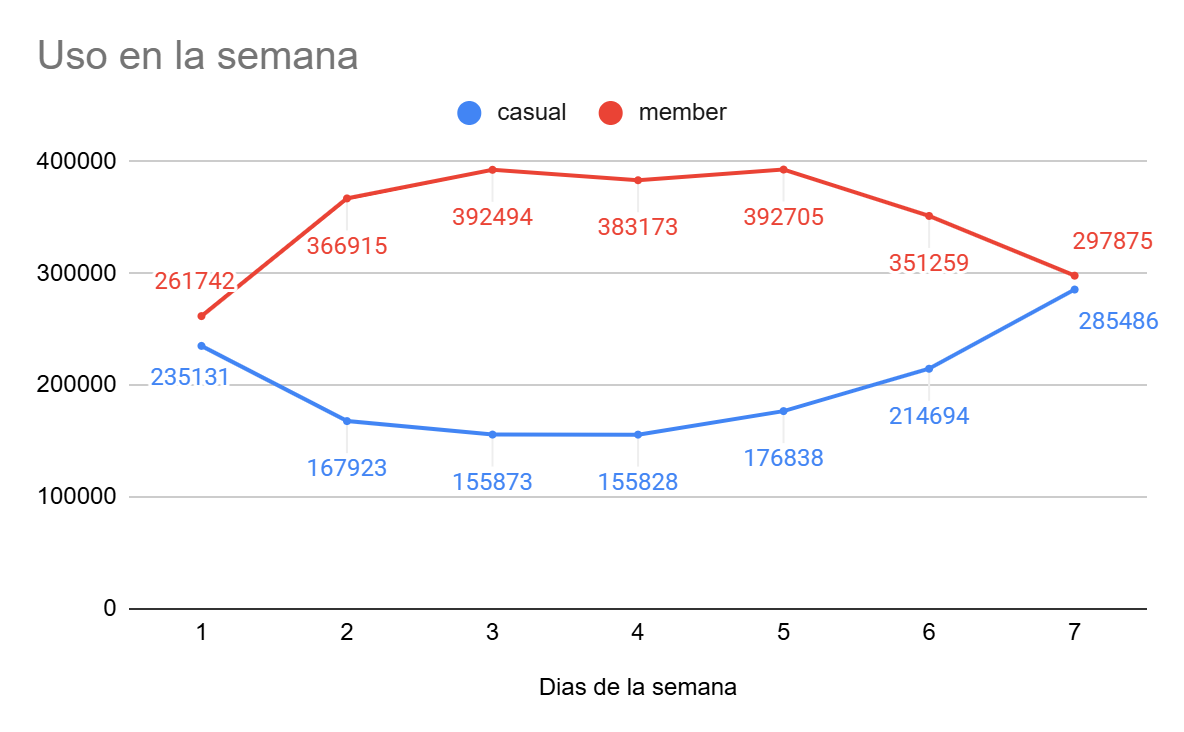


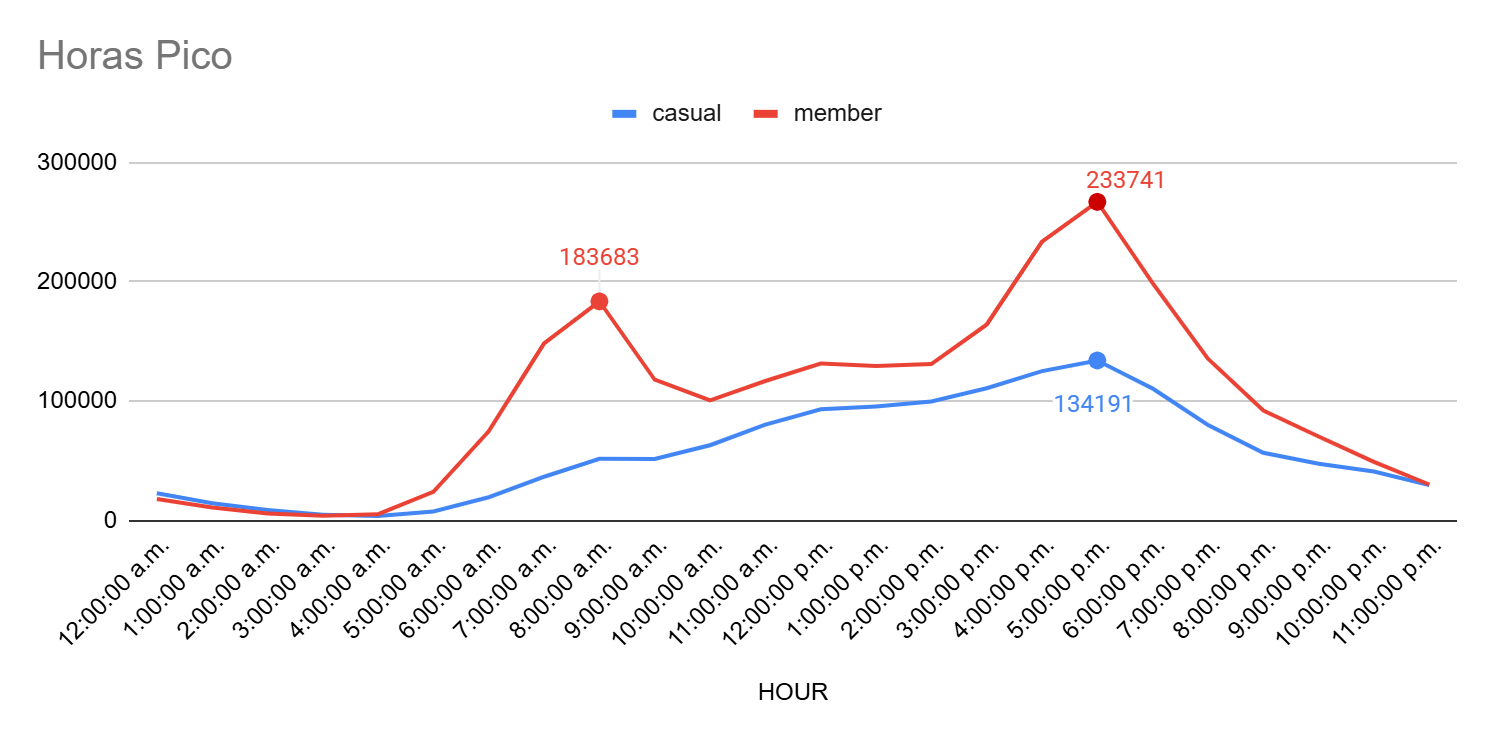


* Tanto miembros como ocasionales muestran preferencia por la bicicleta **clásica**.

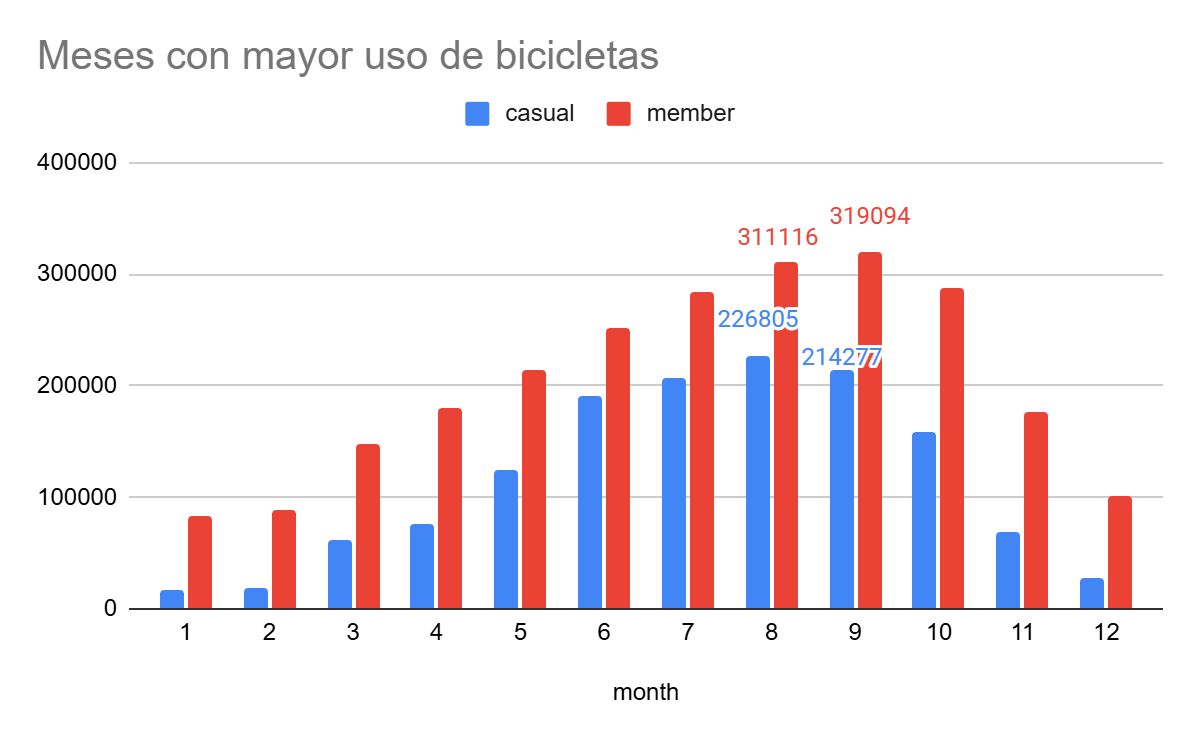


* El uso de bicicletas por parte de los miembros es mayor en días laborables, especialmente en las horas pico de la mañana y tarde , mientras que los ciclistas ocasionales usan más las bicicletas en fines de semana y durante horarios variados, con un pico notable a las 5pm .





* Los meses con mayor uso para ambos tipos de ciclistas son los mismos meses de agosto y septiembre.



### **6. Conclusiones preliminares**

Los resultados revelan que los miembros anuales y los ciclistas ocasionales tienen patrones de uso significativamente diferentes en términos de duración, horarios.Estas diferencias sugieren que las estrategias de marketing deben ser personalizadas para cada segmento.

Por ejemplo, las campañas para convertir ocasionales podrían enfocarse en promocionar beneficios para el uso frecuente durante los fines de semana o en ciertos horarios donde hay alta actividad ocasional. También se podrían resaltar las ventajas de algunos tipos de bicicleta para favorecer un mayor uso.

Estos hallazgos ofrecen una base sólida para que el equipo de marketing diseñe campañas digitales dirigidas a incrementar la conversión de ciclistas ocasionales en miembros anuales, apoyando así los objetivos de negocio de Cyclistic.

## **SHARE (Compartir)**

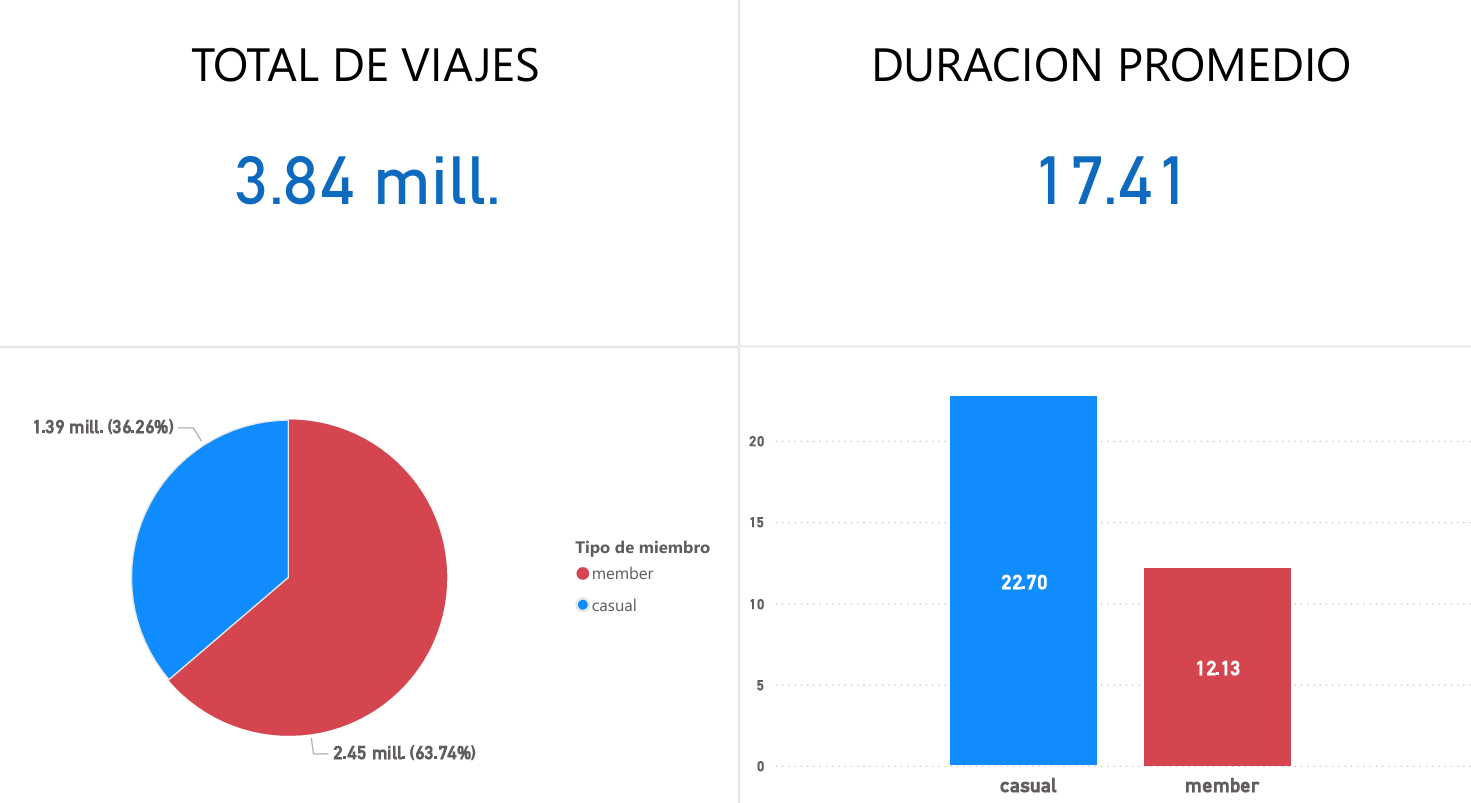
**1. Título del caso:** *Análisis del uso de bicicletas de Cyclistic: Miembros anuales vs. ciclistas ocasionales*

**2. Objetivo:** Identificar diferencias en el uso de bicicletas entre miembros anuales y ciclistas ocasionales para guiar estrategias de marketing.

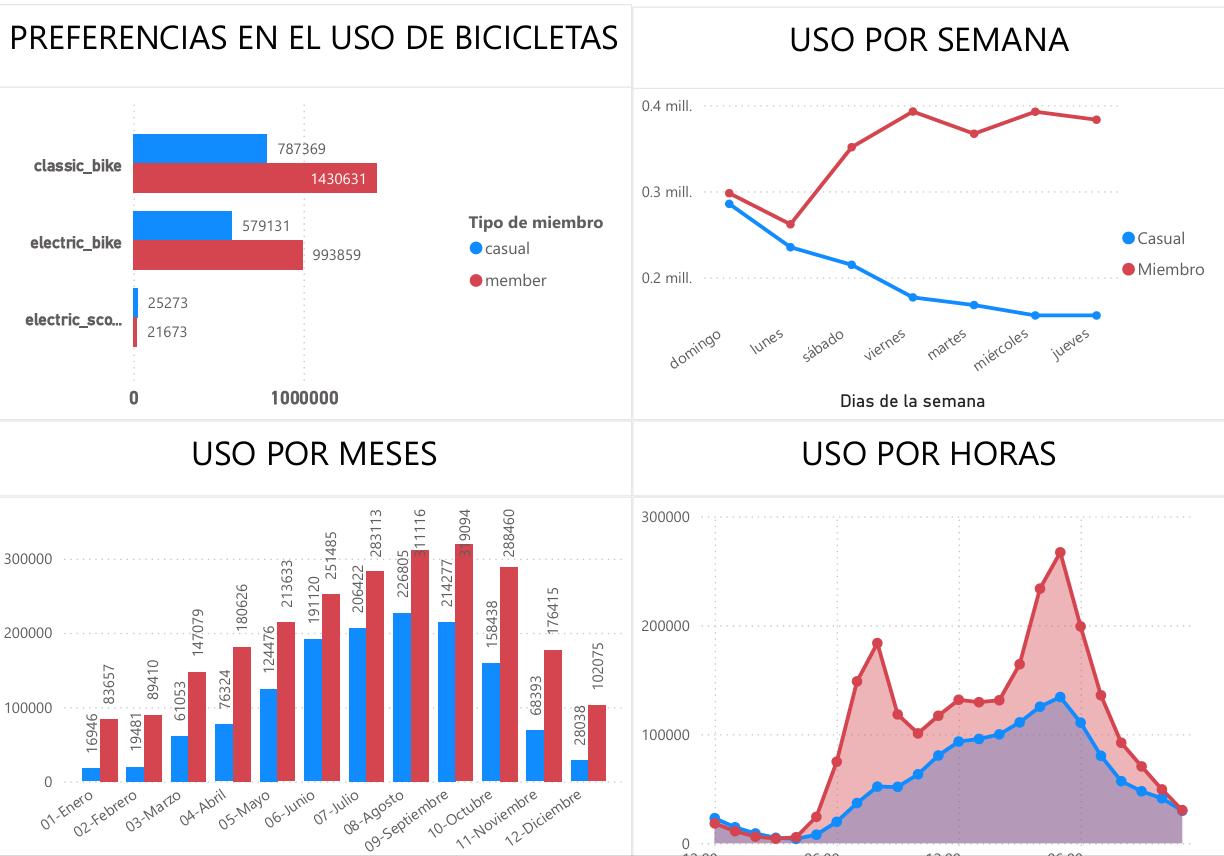
**3. Audiencia:**

* Moreno, gerente de marketing.
* Equipo de marketing y gerentes de producto de Cyclistic.

**4. Hallazgos clave**

****

* Volumen total: 3.84 millones de viajes analizados (63.7% miembros, 36.3% ocasionales)
* Duración de viajes: Los ocasionales realizan viajes 87% más largos (22.70 min vs 12.13 min)

****

* Patrones semanales: Comportamientos opuestos - miembros dominan días laborables, ocasionales los fines de semana
* Patrones horarios: Miembros con picos claros de commute (8am y 5pm), ocasionales más distribuidos con pico vespertino
* Preferencias de bicicleta: Ambos prefieren bicicletas clásicas, pero ocasionales usan más bicicletas eléctricas proporcionalmente
* Estacionalidad: Agosto y septiembre como meses pico para ambos grupos

### **5.Análisis detallado**

### Diferencias en Duración de Viajes

Insight clave: Los ocasionales invierten casi el doble de tiempo por viaje

* Ocasionales: 22.70 minutos promedio - Uso recreativo/exploratorio
* Miembros: 12.13 minutos promedio - Uso utilitario/transporte

*Implicación*: Los ocasionales valoran la experiencia, no solo el transporte

### Preferencias de Tipo de Bicicleta

Bicicletas Clásicas dominan en ambos grupos:

* Miembros: 1,430,631 viajes vs Ocasionales: 787,369 viajes

### Patrones Semanales: Comportamientos Complementarios

Miembros - Patrón de transporte laboral:

* Uso constante alto Miercoles y Viernes
* Descenso progresivo hacia el fin de semana

Ocasionales - Patrón recreativo:

* Descenso dramático Lunes-Miércoles
* Pico en fin de semana, especialmente Sábado

### Estacionalidad: Oportunidades Estacionales

Meses de mayor actividad (ambos grupos):

1. Septiembre: 319,094 viajes (miembros) vs 214,417 (ocasionales)
2. Agosto: 311,116 viajes (miembros) vs 222,805 (ocasionales)

Temporada baja: Enero-Febrero presenta la mayor oportunidad para campañas de retención

### Patrones Horarios: Ventanas de Oportunidad

Miembros - Horarios de commute definidos:

* Pico matutino: 8:00 AM
* Pico vespertino masivo: 5:00 PM
* Caída drástica después de 6:00 PM

Ocasionales - Flexibilidad con concentración vespertina:

* Crecimiento gradual durante el día
* Pico único: 5:00 PM
* Mantienen actividad hasta más tarde

# ACT (Actuar)

### **1. ¿Cuál es tu conclusión final basada en tu análisis?**

Conclusión Principal: Los datos revelan que los miembros anuales y ciclistas ocasionales de Cyclistic representan dos mercados fundamentalmente diferentes con necesidades y comportamientos distintos. Los 1.39 millones de usuarios ocasionales (36.3% del total) no son simplemente "miembros potenciales", sino un segmento con propósito de uso completamente diferente: experiencial vs utilitario.

**2. Hallazgos Clave:**

Los ocasionales invierten 87% más tiempo por viaje (22.70 vs 12.13 minutos), indicando uso recreativo/exploratorio

Patrones temporales complementarios: miembros dominan días laborables, ocasionales los fines de semana

Convergencia estratégica a las 5:00 PM presenta la mayor oportunidad de conversión

Temporada alta común (Agosto-Septiembre) ofrece ventana natural para campañas intensivas

### **3. ¿Cómo podría tu equipo y negocio aplicar estos insights?**

#### Para el Equipo de Marketing

Segmentación Inmediata:

* "Flexitrabajadores": Ocasionales que usan en días laborables
* "Weekend Warriors": Uso exclusivo de fines de semana
* "Exploradores Eléctricos": Preferencia por tecnología premium

Campañas Dirigidas:

* "Momento 5PM": Notificaciones push diarias a las 4:30 PM
* "Temporada Alta": Promociones intensivas Agosto-Septiembre
* "Saturday Premium": Campañas específicas para fines de semana

### **4. ¿Hay datos adicionales que podrías usar para expandir estos hallazgos?**

#### Datos Internos Recomendados

Geolocalización Avanzada:

* Rutas populares por segmento: ¿Dónde van los ocasionales vs miembros?
* Distancia promedio: Correlación entre distancia y duración por tipo de usuario
* Estaciones de alta conversión: ¿Qué ubicaciones generan más conversiones?

## **5.Recomendaciones Basadas en el Análisis**

### Lanzar Campaña "Momento 5PM"

Qué: Campaña de notificaciones push dirigida a usuarios ocasionales a las 4:30 PM todos los días Por qué: Las 5:00 PM es el único momento donde ambos grupos convergen masivamente

### Crear Membresía "Weekend Premium"

Qué: Producto intermedio específico para usuarios de fin de semana como paso hacia membresía completa

Precio: 60% del costo de membresía anual

Beneficios: Acceso prioritario fines de semana, descuentos en viajes largos

### Implementar "Programa de Conversión Estacional"

Qué: Campaña intensiva durante Agosto-Septiembre cuando ambos grupos están en pico máximo Por qué: Estos meses representan el 20% del volumen total anual - momento de máximo engagement

Descuento del 30% en membresías durante temporada alta

Bonificación: "Usa todo Septiembre gratis al convertirte en miembro"