

# Reporte de avance de caso práctico 1

## Elegimos el caso práctico de las bicicletas.

### Escenario

Eres un analista de datos junior que trabaja en el equipo de analistas de marketing de Cyclistic, una empresa de bicicletas compartidas de Chicago. La directora de marketing cree que el éxito futuro de la empresa depende de **maximizar la cantidad de membresías anuales**. Por lo tanto, tu equipo quiere entender qué diferencias existen en el uso de las bicicletas Cyclistic entre los ciclistas ocasionales y los miembros anuales. A través de estos conocimientos, tu equipo diseñará una nueva estrategia de marketing para convertir a los ciclistas ocasionales en miembros anuales. Sin embargo, antes de eso, los ejecutivos de Cyclistic deben aprobar tus recomendaciones; por eso, debes respaldar tu propuesta con una visión convincente de los datos y visualizaciones profesionales de los mismos.

### Personajes y equipos

- **Cyclistic:** Un programa de bicicletas compartidas que incluye
  - 5,800 bicicletas
  - y 600 estaciones.
- Cyclistic se destaca por ofrecer también
  - bicicletas reclinadas,
  - triciclos manuales
  - y bicicletas de carga
- que ofrecen un uso más inclusivo de las bicicletas compartidas para las personas con discapacidad y los ciclistas que no pueden utilizar una bicicleta estándar de dos ruedas. La mayoría de los ciclistas eligen las bicicletas tradicionales, alrededor de un 8% de los ciclistas usan las opciones asistidas. Los usuarios de Cyclistic son más propensos a utilizar la bicicleta para recreación, pero alrededor del 30% la utiliza para ir al trabajo cada día.
- **Lily Moreno:** La directora de marketing y tu gerente. Moreno es responsable del desarrollo de campañas e iniciativas para promover el programa de bicicletas compartidas. Las campañas pueden incluir correo electrónico, redes sociales y otros canales.
- **Equipo de análisis computacional de datos de marketing de Cyclistic:** Un equipo de analistas de datos que se encargan de recopilar, analizar e informar datos que ayudan a conducir la estrategia de marketing de Cyclistic. Te incorporaste a este equipo hace seis meses y te has dedicado no solo a conocer la misión y las metas de negocios de Cyclistic, sino

también a ver cómo puedes ayudar a Cyclistic a lograrlo, desde tu posición de analista de datos júnior.

- **Equipo ejecutivo de Cyclistic:** El equipo ejecutivo, sumamente detallista, decidirá si aprueba el programa de marketing recomendado.

## Acerca de la empresa

**En 2016, Cyclistic lanzó una exitosa oferta de bicicletas compartidas.** Desde entonces, el programa creció hasta alcanzar una flota de **5,824 bicicletas** georreferenciadas y bloqueadas en una red de **692 estaciones** en toda Chicago. Las bicicletas se pueden desbloquear desde una estación y devolverse en cualquier otra estación del sistema en cualquier momento.

Hasta ahora, la estrategia de marketing de Cyclistic se basaba en la construcción de un reconocimiento de marca general y en atraer a amplios segmentos de consumidores. Uno de los enfoques que ayudó a hacer esto posible fue la flexibilidad de sus **planes de precios: pases de un solo viaje, pases de un día completo y membresías anuales.** A los clientes que compran pases de un solo viaje o pases de un día completo se los llama ciclistas ocasionales. Los clientes que compran membresías anuales se llaman miembros de Cyclistic.

Los analistas financieros de Cyclistic llegaron a la conclusión de que los miembros anuales son mucho más rentables que los ciclistas ocasionales. Aunque la flexibilidad de precios ayuda a Cyclistic a atraer más clientes, Moreno cree que maximizar el número de miembros anuales será clave para el crecimiento futuro. En lugar de crear una campaña de marketing que apunte a todos los clientes nuevos, Moreno cree que hay muchas posibilidades de **convertir a los ciclistas ocasionales en miembros.** Ella señala que los ciclistas ocasionales ya conocen el programa de Cyclistic y han elegido a Cyclistic para sus necesidades de movilidad.

Moreno estableció una meta clara: Diseñar estrategias de marketing orientadas a convertir a los ciclistas ocasionales en miembros anuales. Sin embargo, para hacer eso, **el equipo de analistas de marketing necesita entender mejor cómo difieren los miembros anuales y los ciclistas ocasionales, por qué los ciclistas ocasionales comprarían una membresía y cómo los medios digitales podrían afectar sus tácticas de marketing.** Moreno y su equipo están interesados en analizar los datos históricos de viajes en bicicleta de Cyclistic para identificar tendencias.

## Preguntar

Tres preguntas guiarán el futuro programa de marketing:

1. ¿En qué se diferencian los socios anuales y los ciclistas ocasionales con respecto al uso de las bicicletas de Cyclistic?
2. ¿Por qué los ciclistas ocasionales comprarían membresías anuales de Cyclistic?
3. ¿Cómo puede usar Cyclistic los medios digitales para influenciar a los ciclistas ocasionales a convertirse en miembros?

Moreno te asignó la primera pregunta por responder: ¿En qué se diferencian los socios anuales y los ciclistas ocasionales con respecto al uso de las bicicletas de Cyclistic?

### Crearás un informe con los siguientes entregables:

1. Una instrucción clara de la tarea empresarial
2. Una descripción de todas las fuentes de datos utilizadas
3. Documentación de todas las limpiezas y manipulaciones de datos
4. Un resumen de tu análisis
5. Visualizaciones de respaldo y hallazgos clave
6. Las tres recomendaciones más importantes basadas en tu análisis

## Preguntar:

### Mapa de ruta de caso práctico - Preguntar

#### Preguntas orientativas

- ¿Cuál es el problema que intentas resolver?  
¿En qué se diferencian los socios anuales y los ciclistas ocasionales con respecto al uso de las bicicletas de Cyclistic?
- ¿Cómo tus conocimientos pueden impulsar las decisiones empresariales?

#### Tareas clave

1. Identificar la tarea empresarial
2. Considerar a los interesados clave

#### Entregable

Una instrucción clara de la tarea empresarial

## Preparar:

### Mapa de ruta de caso práctico - Preparar

#### Preguntas orientativas

- ¿Dónde se ubican tus datos?  
**En <https://divvy-tripdata.s3.amazonaws.com/index.html>**
- ¿Cómo están organizados los datos?  
**En una tabla con 13 columnas y 537114 filas**
- ¿Hay problemas con el sesgo o la credibilidad de estos datos? [¿Tus datos son confiables, originales, integrales, actuales y citados \(ROCCC\)?](#)  
**Si los datos provienen de la empresa Motivate International Inc.**
- ¿Cómo estás abordando la autorización, la privacidad, la seguridad y la accesibilidad?  
**Hay una licencia para la utilización de estos datos**
- ¿Cómo verificaste la integridad de los datos?  
**Se nos comunicó que son íntegros a excepción de los datos de tarjetas de crédito de los usuarios**
- ¿De qué manera te ayuda a responder tu pregunta?  
**Tendremos que observar los datos que nos pueden dar las pistas necesarias para observar las diferencias entre las personas con membresías y las personas eventuales.**
- ¿Existe algún problema con los datos?  
**Falta limpieza**

#### Tareas clave

1. Descarga los datos y almacénalos adecuadamente.  
**Se descargaron adecuadamente**
2. Identifica cómo están organizados.  
**Están en un solo archivo csv. Y están en una tabla vertical con 13 columnas y 537114 filas**
3. Ordena y filtra los datos.
4. Determina la credibilidad de los datos.

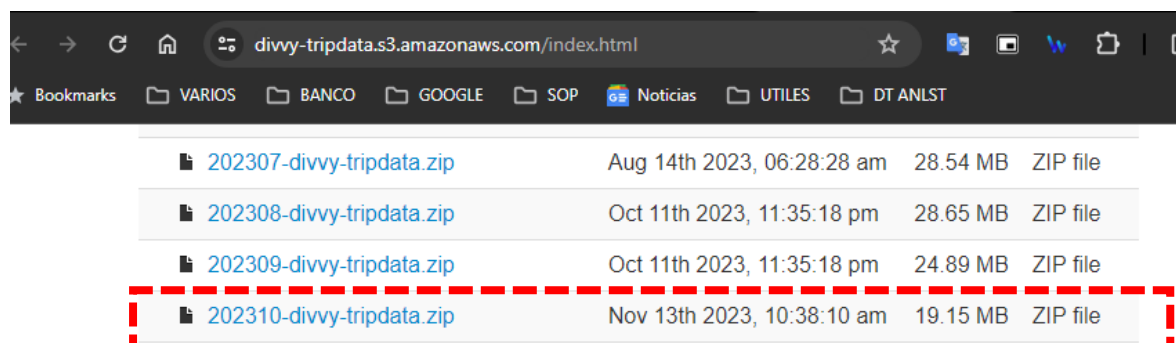
#### Entregable

Una descripción de todas las fuentes de datos utilizadas.

**La fuente es de la empresa Motivate International Inc. En <https://divvy-tripdata.s3.amazonaws.com/index.html> allí se encuentra todas las bases de datos de cada mes desde 2020.**

## Tomamos los datos de la base de datos en la pagina

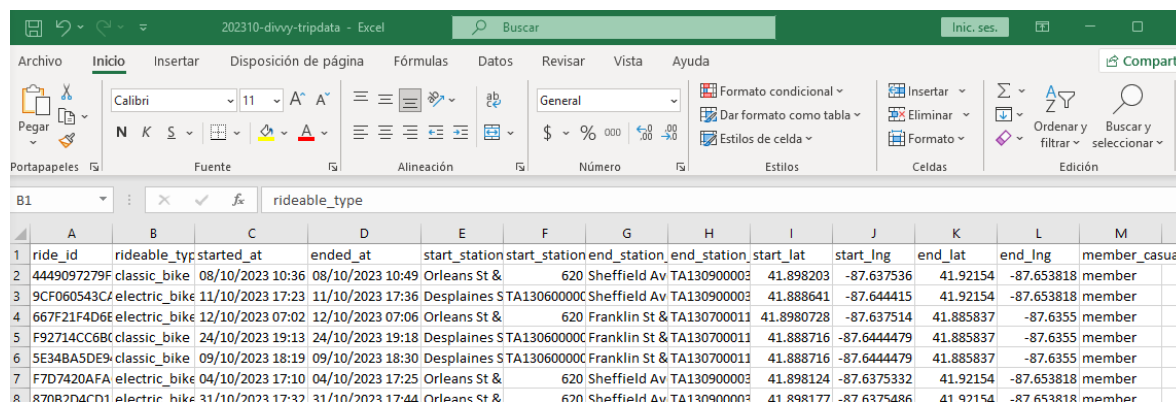
<https://divvy-tripdata.s3.amazonaws.com/index.html>



202307-divvy-tripdata.zip	Aug 14th 2023, 06:28:28 am	28.54 MB	ZIP file
202308-divvy-tripdata.zip	Oct 11th 2023, 11:35:18 pm	28.65 MB	ZIP file
202309-divvy-tripdata.zip	Oct 11th 2023, 11:35:18 pm	24.89 MB	ZIP file
202310-divvy-tripdata.zip	Nov 13th 2023, 10:38:10 am	19.15 MB	ZIP file

## Vamos a revisar la información otorgada por la base de datos

La información esta presentada en 13 columnas y 537114 filas de esta manera incluyendo los títulos.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ride_id	rideable_type	started_at	ended_at	start_station	start_station	end_station	end_station	start_lat	start_lng	end_lat	end_lng	member_casual
2	4449097279F	classic_bike	08/10/2023 10:36	08/10/2023 10:49	Orleans St &	620	Sheffield Av	TA130900003	41.898203	-87.637536	41.92154	-87.653818	member
3	9CF060543C7	electric_bike	11/10/2023 17:23	11/10/2023 17:36	Desplaines St	TA13060000C	Sheffield Av	TA130900003	41.888641	-87.644415	41.92154	-87.653818	member
4	667F21F4D6E	electric_bike	12/10/2023 07:02	12/10/2023 07:06	Orleans St &	620	Franklin St &	TA130700011	41.8980728	-87.637514	41.885837	-87.6355	member
5	F92714CC6B6	classic_bike	24/10/2023 19:13	24/10/2023 19:18	Desplaines St	TA13060000C	Franklin St &	TA130700011	41.888716	-87.6444479	41.885837	-87.6355	member
6	5E34BA5DE9	classic_bike	09/10/2023 18:19	09/10/2023 18:30	Desplaines St	TA13060000C	Franklin St &	TA130700011	41.888716	-87.6444479	41.885837	-87.6355	member
7	F7D7420AFA	electric_bike	04/10/2023 17:10	04/10/2023 17:25	Orleans St &	620	Sheffield Av	TA130900003	41.898124	-87.6375332	41.92154	-87.653818	member
8	870B2D4CD1	electric_bike	31/10/2023 17:32	31/10/2023 17:44	Orleans St &	620	Sheffield Av	TA130900003	41.898177	-87.6375486	41.92154	-87.653818	member

## Comenzaremos a revisar la información

### Analizamos la primera columna:

Vemos que la primera columna representa el id de viaje con la bicicleta, con la formula **=largo()** en una nueva columna **longitud\_id** podemos observar que el largo de id de viaje es alrededor de 16 pero vamos a filtrar los datos para observar si completamente todos los datos son de 16 caracteres.

	A	B	
1	longitud_id	ride_id	
2	16	4449097279F8BBE7	4449
3	16	9CF060543CA7B439	9CF0
4	16	667F21F4D6BDE69C	667F
5	16	F92714CC6B019B96	F927
6	16	5E34BA5DE945A9CC	5E34
7	16	F7D7420AFAC53CD9	F7D7
8	16	870B2D4CD112D7B7	870B
9	16	D9179D36E32D456C	D917
10	16	F8E131281F722FEF	F8E1
11	16	91938B71748FA405	9193
12	16	191BFA255C1820FC	191B

Ordenar de menor a mayor

Ordenar de mayor a menor

Ordenar por color

Vista de Hoja

Borrar filtro de "longitud\_id"

Filtrar por color

Filtros de número

☒ (Seleccionar todo)
☒ 14
☒ 15
☒ 16
☒ 17
☒ 18
☒ 19
☒ 20

4449  
9CF0  
667F  
F927  
5E34  
F7D7  
870B  
D917  
F8E1  
9193  
191B  
52E0  
0552  
3A06  
0A57  
F914  
2E53  
E03F  
D917

Y vemos que no, que al momento de capturar la información se generaron errores puesto a que no son todos los id de la misma longitud. Pero afortunadamente esta columna no es de tanta importancia para nuestra tarea empresarial.

## Analizamos la segunda columna:

Vemos que aquí se muestra el tipo de bicicleta utilizada.

B		
rideable_type	status	
classic_bike	08	
electric_bike	11	
electric_bike	12	
classic_bike	24	
classic_bike	09	
electric_bike	04	
electric_bike	31	
classic_bike	02	
classic_bike	17	

Ordenar de A a Z

Ordenar de Z a A

Ordenar por color

Vista de Hoja

Borrar filtro de "rideable\_type"

Filtrar por color

Filtros de texto

☒ (Seleccionar todo)
☒ classic\_bike
☒ electric\_bike

Y al filtrar vemos que si corresponde a los únicos dos tipos de bicicleta utilizada (eléctrica y la clásica) por lo tanto vemos que la información en esta columna es correcta.

## Analizamos la tercera y cuarta columna:

Al ver estas dos columnas nos damos cuenta que tienen un tipo de fecha y hora y que todas las fechas corresponden al año 2023 y al mes de octubre. Lo que indica que posiblemente estén bien los datos.

started_at	ended_at
08/10/2023 10:36	08/10/2023 10:49
11/10/2023 17:23	11/10/2023 17:36
12/10/2023 07:02	12/10/2023 07:06

Pero haremos un análisis de duración de los viajes para ver que si corresponden y que no se ve algo extraño. Así que agregaremos una fila de ride\_duration en la que restaremos los minutos entre la primera fecha y la segunda.

=SI(C2<D2,D2-C2,SI(C2>=D2,"00:00:00"))					
C	D	E	F	G	
started_at	ended_at	ride_duration	start_status	start_status	
08/10/2023 10:36	08/10/2023 10:49	=SI(C2<D2,D2-C2,SI(C2>=D2,"00:00:00"))			

Le ponemos unas condicionantes para que solo me de la resta de el tiempo inicial del viaje menos el final, y que si fuera inverso el tiempo o igual a cero me ponga cero, puesto a que corresponderia a que el viaje fue cancelado antes de iniciar o se cancelo al momento de que iba a iniciar.

Nos da algo así

ride_duration	00/01/1900 00:12
---------------	------------------

Por lo tanto cambiaremos el tipo de formato a solo hora así nos da solo el tiempo de la duración del viaje.

Obtener y transformar datos

Consultas y conexiones

E2

✕

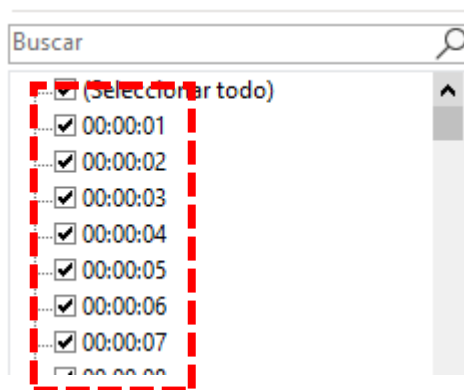
✓

fx

=SI(C2<D2,D2-C2,SI(C2>=D2,"00:00:00"))

	A	B	C	D	E
1	ride_id	rideable_type	started_at	ended_at	ride_duration
2	4449097279F8BBE7	classic_bike	08/10/2023 10:36	08/10/2023 10:49	00:12:53
3	9CF060543CA7B439	electric_bike	11/10/2023 17:23	11/10/2023 17:36	00:12:09
4	667F21F4D6BDE69C	electric_bike	12/10/2023 07:02	12/10/2023 07:06	00:04:20
5	F92714CC6R019R96	classic_bike	24/10/2023 19:13	24/10/2023 19:18	00:05:26

Filtramos y observamos que los tiempos sean congruentes.



Se ve bien la información, todavía podríamos suponer que los viajes que tuvieron pocos segundos de duración quiere decir que el viaje fue cancelado casi al instante de haber iniciado pero sería suponer cosas antes de tener saber el motivo verdadero de la duración corta del viaje.

### Analizamos la quinta, sexta, séptima y octava columna:

Observamos que la quinta columna es el nombre de la estación donde comienza la bicicleta y la sexta columna es el id de la estación, la séptima es el nombre de la estación de donde termina y la octava es el id de donde termina el viaje.

	start_station_name	start_station_id	end_station_name	end_station_id
3	Orleans St & Chest	620	Sheffield Ave & Webst	TA130900003
3	Desplaines St & Kir	TA1306000003	Sheffield Ave & Webst	TA130900003
0	Orleans St & Chest	620	Franklin St & Lake St	TA130700011
5	Desplaines St & Kir	TA1306000003	Franklin St & Lake St	TA130700011
0	Desplaines St & Kir	TA1306000003	Franklin St & Lake St	TA130700011
2	Orleans St & Chest	620	Sheffield Ave & Webst	TA130900003
0	Orleans St & Chest	620	Sheffield Ave & Webst	TA130900003

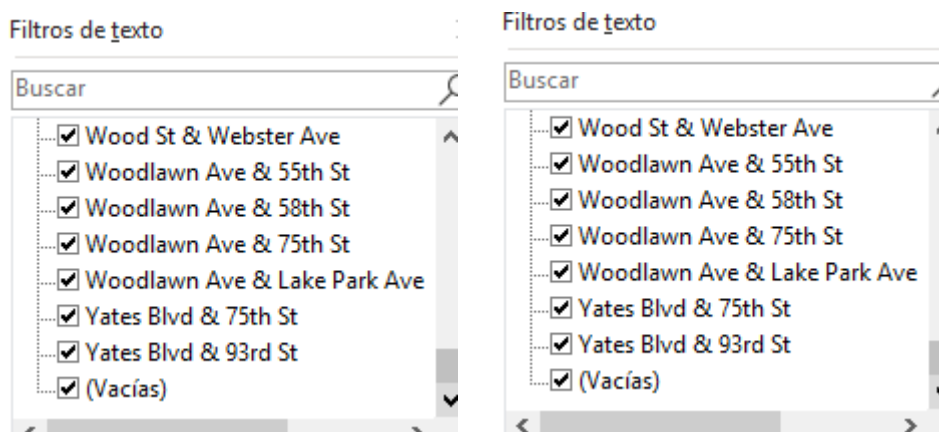
A simple vista se observa que la columna sexta tiene diferentes tipos de id de estación y vamos a visualizar la información a ver si corresponde el nombre de estación y el número de id.

	F	G		
	start_station_name	start_station_id	end_station_name	end_station_id
53	Orleans St & Chest	620	Sh	
20	Orleans St & Chest	620	Fr	
22	Orleans St & Chest	620	Sh	
00	Orleans St & Chest	620	Sh	
22	Orleans St & Chest	620	Sh	

	start_station_name	start_station_id	end_station_name	end_station_id
30	Western Ave & Wir	13068	L	
39	Western Ave & Wir	13068	S	
51	Western Ave & Wir	13068	P	
39	Western Ave & Wir	13068	C	
34	Western Ave & Wir	13068	L	



Hicimos varios filtrados en diferentes estaciones y verificamos que el nombre de estacion si coincide con el id de estacion, aunque se vean los id de diferentes tipos de id. Todo esto al igual que en los nombres de estacion del fin de viaje y el id de la estacion.



Aunque tambien vemos que existen celdas vacias en las estaciones de inicio y las de finalizacion de viaje.

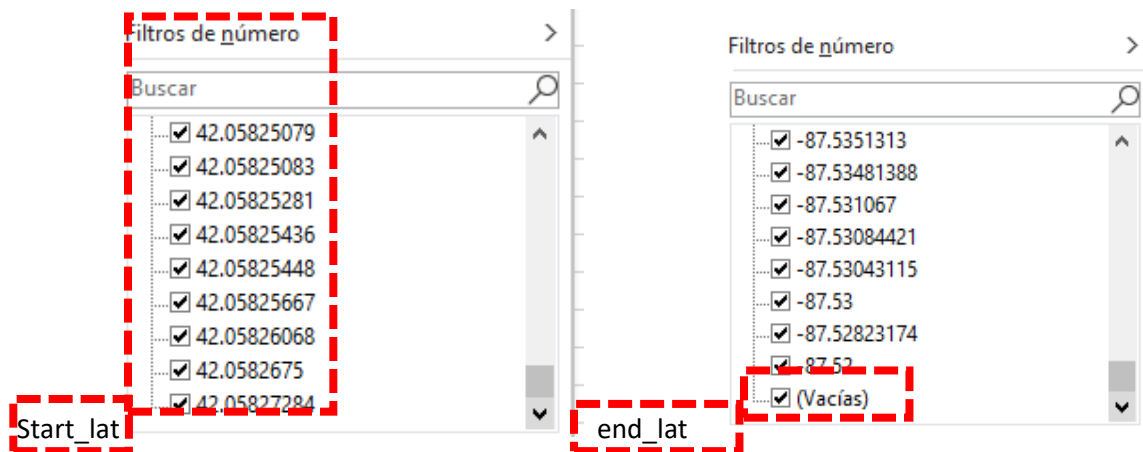
### Analizamos la novena, decima, undécima y duodécima columna:

La novena es la latitud de inicio y la decima es la longitud del inicio de estación.

La undecima es la latitud de fin de viaje y la duodécima es la longitud del fin de viaje de estación.

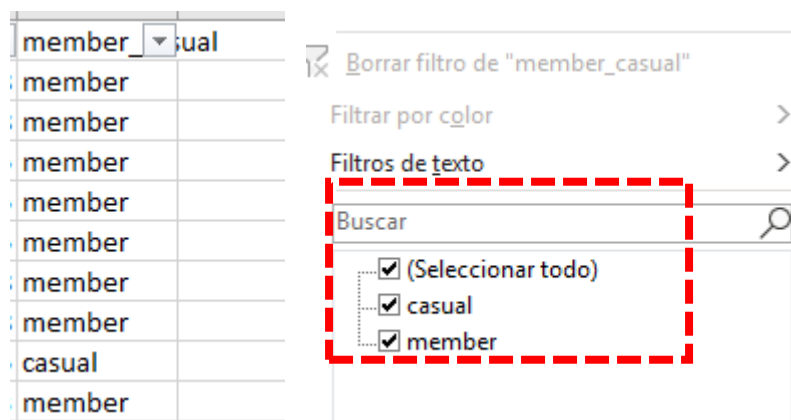
J	K	L	M	
start_lat ▼	start_lng ▼	end_lat ▼	end_lng ▼	n
41.898203	-87.637536	41.92154	-87.653818	n
41.888641	-87.644415	41.92154	-87.653818	n
41.8980728	-87.637514	41.885837	-87.6355	n

Si hacemos varios filtros podemos ver que no existen en start\_lat y start\_lng celdas vacias pero en las coordenadas de fin de viaje si hay celdas vacias. Y creo que es por que al momento de contratar el viaje o pudo haber sido cancelado o pudo haber finalizado en el mismo lugar que inicio el viaje y por illo tanto no colocaron las coordenadas de finalización.



### Analizamos la décima tercera columna:

Podemos ver que todos los datos ingresados son strings o sea textos y no existen celdas vacías. Lo único que existe son las dos opciones viajero que es **miembro** o viajero **casual**.



### Procesar:

#### Mapa de ruta de caso práctico - Procesar

##### Preguntas orientativas

- ¿Qué herramientas eliges y por qué?

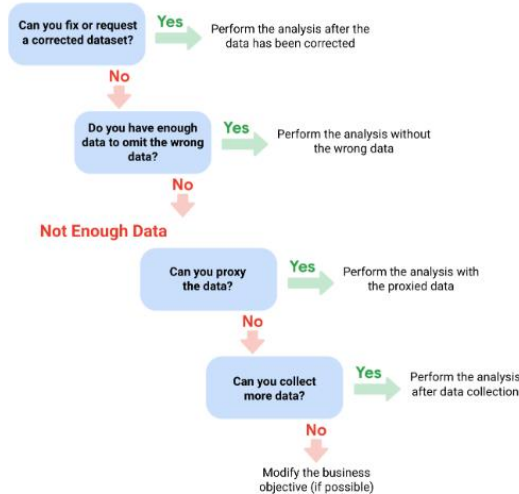
Para limpieza de datos la revisé en Excel para poder visualizar de manera rápida

- ¿Has garantizado la integridad de los datos?

Si se ha garantizado la integridad de los datos. Pues no se ha cambiado ningún tipo de dato solo agregamos una columna extra para poder tener idea de la duración de cada viaje y poder analizar tiempos.

- ¿Qué pasos seguiste para garantizar que tus datos están limpios?  
**Verifique cada una de las columnas y me di cuenta que las columnas que no están completas puedo omitir la información, y esto me permite que pueda continuar con mi análisis siguiendo el diagrama de toma de decisiones por errores en los datos o falta de datos.**
- ¿Cómo puedes verificar que tus datos están limpios y listos para analizar?  
**Con el diagrama de toma de decisiones.**

#### Data Errors



- ¿Documentaste tu proceso de limpieza para poder revisar y compartir estos resultados?  
**si**

#### Tareas clave

1. Verifica si hay errores en los datos.
2. Elige tus herramientas.
3. Transforma los datos para que puedas trabajar con ellos eficazmente.
4. Documenta el proceso de limpieza.

#### Entregable

Documentación de todas las limpiezas y manipulaciones de datos

**Agregaremos también que día de la semana se usa más las bicicletas:**

Retomaremos el analisis previo que se estaba haciendo con la duracion en cada uno de los viajes visualizaremos tambien con una columna nueva llamada "day\_of\_week" para colocar que dia de la semana comenzo cada viaje .

=DIASEM(C2)				
C	D	E	F	
started_at	ended_at	start day of week	ride_duration	s
08/10/2023 10:36	08/10/2023 10:49	1	0:12:53	(
11/10/2023 17:23	11/10/2023 17:36	4	0:12:09	I
12/10/2023 07:02	12/10/2023 07:06	5	0:04:20	(

## Analizar:

### Preguntas orientativas

- ¿Cómo deberías organizar tus datos para realizar un análisis?  
Pues si se planea realizar el análisis en SQL o R habrá que realizar una limpieza minuciosa para que la base de datos quede solamente en formato.csv y con solo una hoja.
- ¿Tus datos tienen el formato correcto?  
Si tienen el formato correcto desde el paso de la limpieza fui cambiando y analizando que el tipo de dato coincidiera con el tipo que debería de tener, y así mismo al ir analizando iba modificando los posibles cambios de tipo de datos que contenía cada resultado.
- ¿Qué sorpresas descubriste en los datos?  
Descubrí que hay viajes que tienen muchas horas de uso de bicicleta y otros viajes que casi no tienen uso de bicicleta.
- ¿Qué tendencias o relaciones encontraste en los datos?  
Encontré que la mayoría de los viajes se hacen en fin de semana el día sábado como primer lugar, en viernes el segundo lugar y jueves el 3er lugar.
- ¿Cómo te ayudarán estos conocimientos para responder a tus preguntas empresariales?  
Pues se extrajo toda la información relevante que se pudo obtener de los datos para poder identificar tendencias en cada uno de los tipos de miembros y los días y horarios de más uso.

### Tareas clave

1. Consolida tus datos para que sean útiles y accesibles.
2. Organiza y formatea tus datos.
3. Realiza cálculos.
4. Identifica tendencias y relaciones.

### Entregable

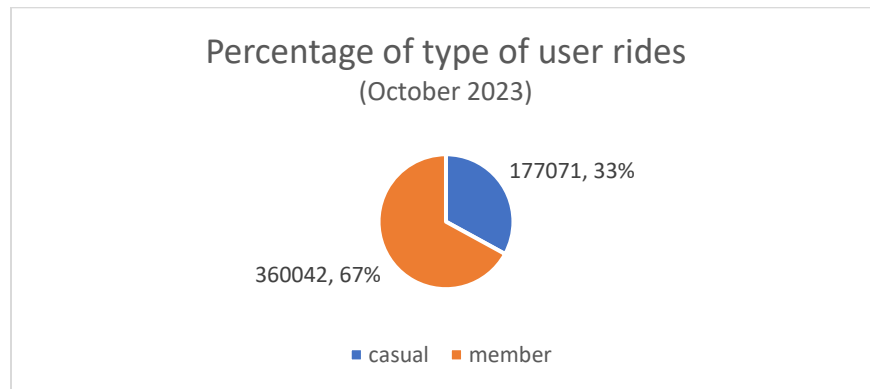
Un resumen de tu análisis

### Podemos relizar un analisis desde lo mas general, que nos muestra que:

Que en lo que respecta a numero de viajes los usuarios que son miembros son los que usan en mayor numero de veces en comparacion con los usuarios que son casuales.

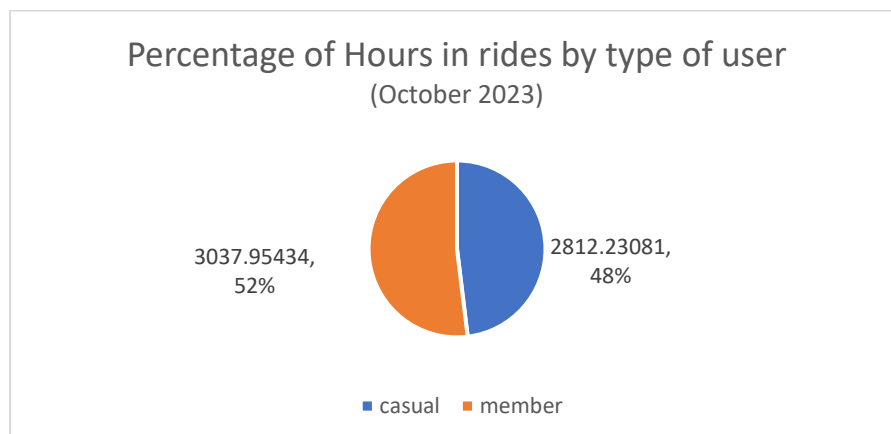
#### Porcentaje de viajes entre miembros y casuales

Etiquetas de fila	Cuenta de ride_id	perc_of_rides
casual	177071	33%
member	360042	67%
<b>Total general</b>	<b>537113</b>	100%



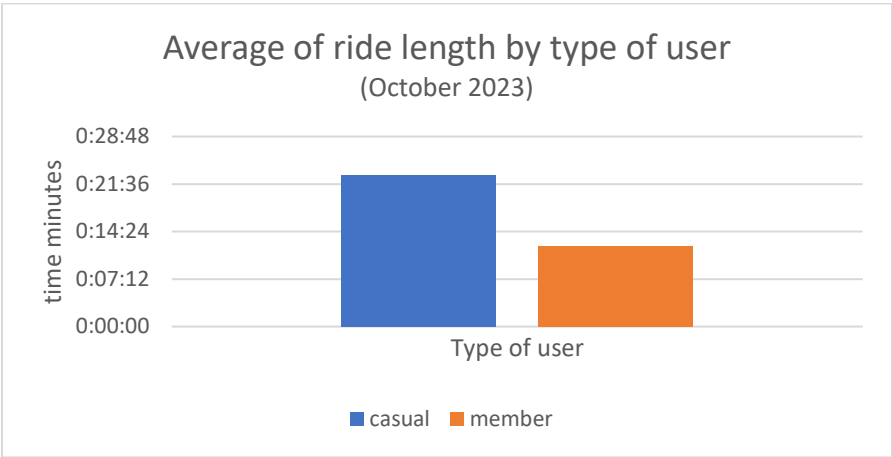
Pero a pesar de eso podemos observar que si calculamos el numero de horas que se usa la bicicleta por cada tipo de usuario vemos que los usuarios casuales son los que emplean el mayor tiempo de uso de la bicicleta pero en una cantidad de horas muy similares a las de los miembros. Lo que podria significar que si cobramos por tiempo usado, tenemos un mayor ingreso por la gran cantidad de horas usadas en relacion a los pocos viajes que se realizaron por parte de los usuarios casuales.

Etiquetas de fila	Suma de ride_duration	perc_of_hours_of_all_rides
casual	2812.23081	48%
member	3037.95434	52%
<b>Total general</b>	<b>5850.18515</b>	100%



De alli que nos llevo a realizar un analisis de cual es el tiempo promedio de duracion de los viajes de usuarios que son miembros o casuales y obserbamos que la mayoría de los viajes casuales son de mayor duracion que los miembros un aproximado de 10 minutos.

Etiquetas de fila	Promedio de ride_duration
casual	0:22:53
member	0:12:09
Total general	0:15:41

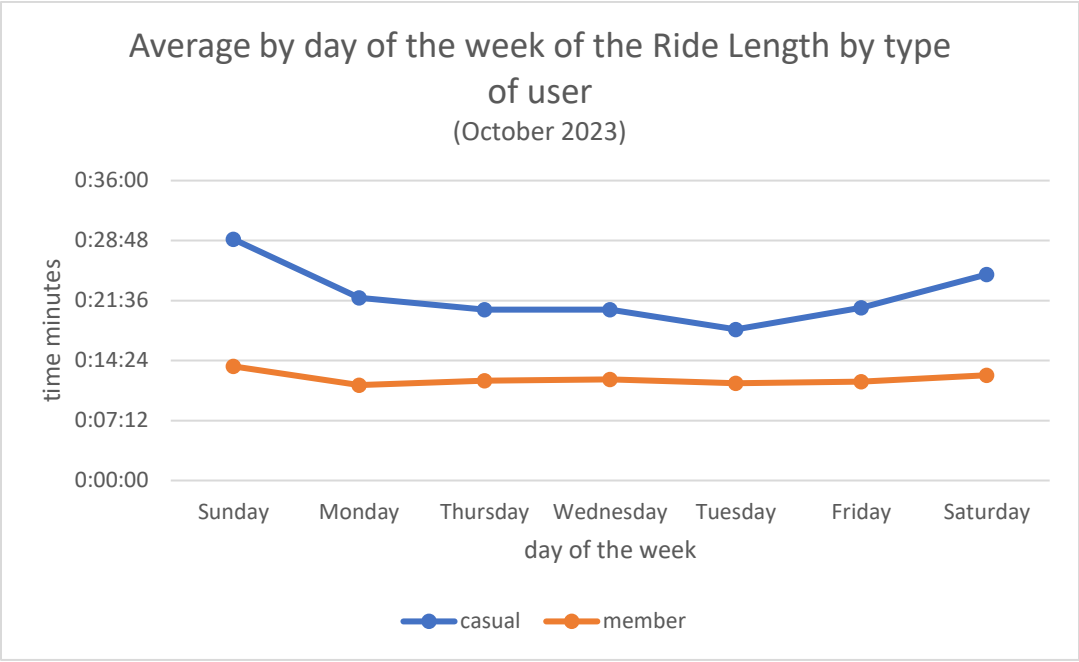


Y decidimos revisar a que se debe este aumento de tiempo de tiempo en esos viajes y realizamos un analisis por dia de la semana para poder encontrar como se comporta la duracion de los viajes en cada dia de la semana.

Y el resultado fue que los casuales tardan mas tiempo en realizar un viaje en los dias del fin de semana, que seguramente se debe al uso de esparcimiento o de turistas que quieren usar las bicicletas para realizar sus recorridos.

Tambien nos dimos cuenta que los miembros mantienen el rango promedio de tiempo de uso de la bicicleta la mayoría de los días de la semana y que seguramente se debe a que emplean el uso de la bicicleta para transportarse a sus trabajos y en fin de semana alargan un poco el tiempo pero no se compara con los casuales.

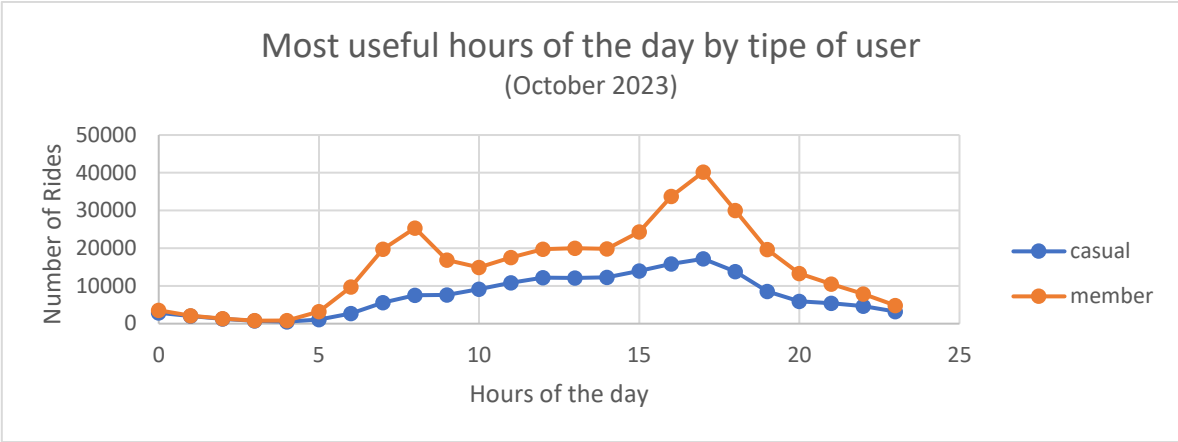
Etiquetas de fila	1	2	3	4	5	6	7	Total general
casual	0:28:57	0:21:56	0:20:30	0:20:31	0:18:09	0:20:45	0:24:44	0:22:53
member	0:13:42	0:11:27	0:12:00	0:12:08	0:11:42	0:11:52	0:12:37	0:12:09
Total general	0:20:32	0:14:29	0:14:23	0:14:34	0:13:30	0:14:43	0:17:36	0:15:41



Para seguir sustentando nuestros resultados seguimos analizando y observamos que efectivamente los usuarios miembros tienen un mayor numero de viajes cercanos a las 8:00hrs y a las 17:00hrs con lo cual validamos nuestra hipotesis.

Así mismo tambien visualizamos que los miembros casuales prefieren hacer ejercicio en las bicicletas a partir de las 6:00hrs y poco a poco ir incrementando su uso hasta las 17:00hrs que es la hora promedio que comienza a ponerse el sol y la mayoría de los usuarios casuales dejan de usar las bicicletas.

Total	member	casual	Tipo usuario/horas
6450	3574	2876	0
4140	2103	2037	1
2602	1314	1288	2
1557	790	767	3
1395	862	533	4
4319	3222	1097	5
12450	9729	2721	6
25317	19753	5564	7
32942	25368	7574	8
24496	16876	7620	9
24006	14893	9113	10
28390	17537	10853	11
32019	19788	12231	12
32098	20018	12080	13
32093	19834	12259	14
38297	24352	13945	15
49577	33714	15863	16
57387	40208	17179	17
43791	30012	13779	18
28149	19635	8514	19
19210	13269	5941	20
15894	10510	5384	21
12560	7895	4665	22
7974	4786	3188	23
537113	360042	17707	Total general

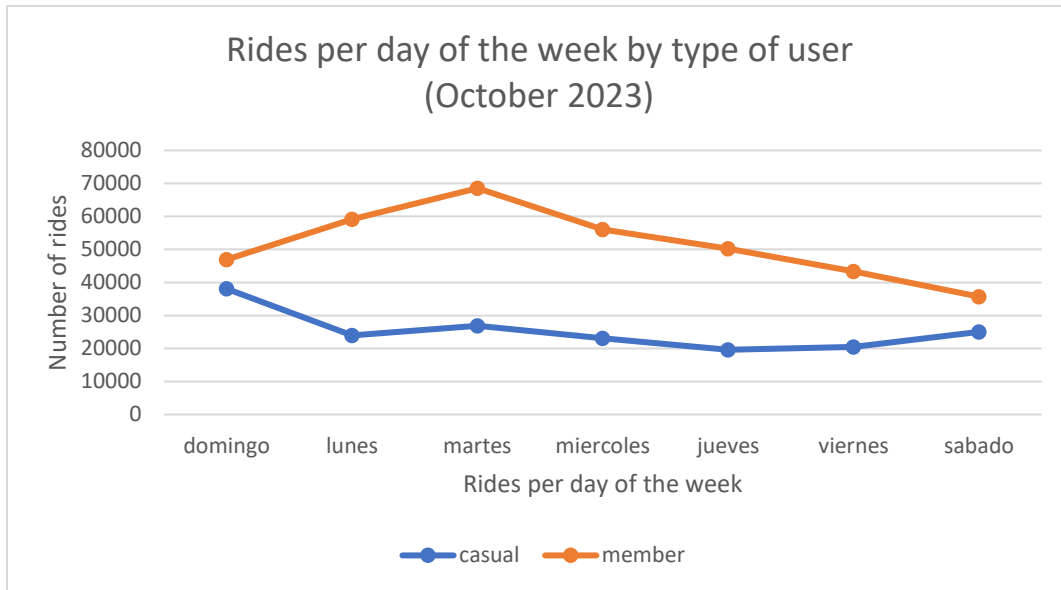


Analizando los mismos datos desde el punto de vista de la cantidad de viajes por día podemos observar que los miembros son los que tienen una frecuencia de uso mas constante por cantidad de viajes relizados por día como vimos en la primer figura, pero tambien vemos que los miembros tienden a tener un mayor uso de la bicicleta entre semana concentrando el mayor numero de viajes el martes.

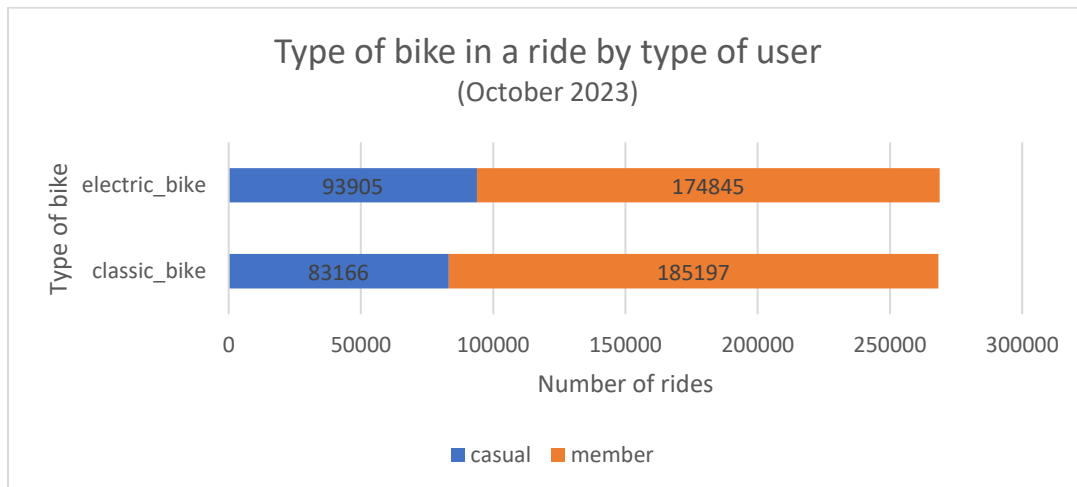
Y los miembros casuales de manera invertida concentran un mayor numero de viajes los fines de semana dominco como día pico y sabado como segundo día.

Cuenta de		perc_each_day_of_week
Etiquetas de fila	start_day_of_week	
casual	177071	
1	38078	21.5% domingo
2	24005	13.6% lunes
3	26837	15.2% martes
4	23049	13.0% miércoles
5	19616	11.1% jueves
6	20484	11.6% viernes
7	25002	14.1% sábado
member	360042	100.0% total casuales
1	46947	13.0% domingo
2	59180	16.4% lunes
3	68543	19.0% martes
4	56046	15.6% miércoles
5	50279	14.0% jueves
6	43338	12.0% viernes
7	35709	9.9% sábado
Total general	537113	100.0% total miembros





Por ultimo como un extra quisimos observar cual es la preferencia de los usuarios respecto al tipo de bicicleta electrica o la bicicleta clasica para ver si es pertinente mantener los dos tipos y la verdad es que se compensa el numero de viajes que se hacen con la electrica y la clasica por los dos tipos de usuarios.



## Compartir:

### Preguntas orientativas

- ¿Pudiste responder la pregunta en qué se diferencian los socios anuales y los ciclistas ocasionales de Cyclistic?

**La principal diferencia es que los usuarios casuales emplean la mayoría de sus viajes para la recreación y que los miembros para el traslado a su lugar de trabajo.**

- ¿Qué historia cuentan tus datos?

**Cuentan la historia de como existen 2 tipos de usuarios que se diferencian mucho el uno del otro y que cada uno tiene sus rutinas de uso de las bicicletas de manera distinta y complementaria la una de la otra.**

- ¿De qué manera tus hallazgos se relacionan con tu pregunta original?

**Se relacionan por que nos confirman las diferencias esenciales de los dos tipos de usuarios.**

- ¿Cuál es tu audiencia? ¿Cuál es la mejor manera de comunicarte con ella?

**Mi audiencia son los encargados de marketing. Hablar de la manera en que se llegó a obtener las conclusiones que obtuvimos y cuáles eran las principales diferencias entre los miembros y los usuarios que son casuales.**

- ¿La visualización de datos puede ayudarte a compartir tus hallazgos?

**Si las visualizaciones nos muestran con mucha claridad cuáles fueron los resultados que obtuvimos puesto que así comprendimos las diferencias básicas en cada uno de los tipos de usuarios.**

- ¿Tu representación es accesible para tu público?

**Si. Creo que es clara y concisa.**

### Tareas clave

1. Determina la mejor manera de compartir tus hallazgos.
2. Crea visualizaciones de datos efectivas.
3. Presenta tus hallazgos.
4. Garantiza que tu trabajo sea accesible.

### Entregable

Visualizaciones de respaldo y hallazgos clave

## Actuar:

### Preguntas orientativas

- ¿Cuál es tu conclusión en función de tu análisis?

De todos los resultados de nuestros análisis asumimos que la mayoría de los afiliados utilizan la bicicleta para ir al trabajo y volver a casa y su uso se mantiene a lo largo de cada día.

También que la mayoría de usuarios ocasionales utilizan la bicicleta para recreación, ya que la utilizan durante más tiempo los fines de semana en horarios que aumentan hasta las 17:00 horas. y en viajes más largos.

- ¿Cómo podrían tu equipo y tu empresa aplicar tus conclusiones?

La primera estrategia sería crear promociones para que sea más accesible para las personas que van a trabajar en bicicleta si se convierten en usuarios.

La segunda estrategia sería crear promociones para que las personas que usan los fines de semana bicicletas que les gusta recrear, puedan usarlas a menor costo los fines de semana si se convierten en usuarios.

- ¿Qué próximos pasos tú o los interesados podrían adoptar en función de tus hallazgos?

Tomar en cuenta más marketing que promueva el uso de la bicicleta los días entre semana a menor costo y así promover la fidelización de los usuarios que son miembros casuales.

- ¿Existen datos adicionales que podrías utilizar para ampliar tus hallazgos?

Actualmente no.

### Tareas clave

1. Crea tu portfolio.
2. Agrega tu caso práctico.
3. Practica presentando tu caso práctico a un amigo o familiar.

### Entregable

Las tres recomendaciones más importantes basadas en tu análisis

1. La primera estrategia sería crear promociones para que sea más accesible para las personas que van a trabajar en bicicleta si se convierten en usuarios.
2. La segunda estrategia sería crear promociones para que las personas que usan los fines de semana bicicletas que les gusta recrear, puedan usarlas a menor costo los fines de semana si se convierten en usuarios.
3. Tomar en cuenta más marketing que promueva el uso de la bicicleta los días entre semana a menor costo y así promover la fidelización de los usuarios que son miembros casuales.