# Objectives of the project

We want to identify the differences between casual riders and annual members and any trends from historical trip data to build a targeted marketing campaign to convert casual riders into annual members.

By understanding how casual riders and annual members use the bikes, we can identify any opportunities for casual riders who could potentially be convinced to join the membership.

## Stakeholders

* Executive team who will decide whether to approve our recommendations.
* Our manager, the director of marketing, Lily Moreno.
* My team and I who will collect, analyze and report the data.

# Prepare

The monthly data collected from Divvy Bikes are from June 2023 to May 2024. As the volume of data exceeds 5 million rows, I will use PostgreSQL to store and clean the entire dataset.

## Problems encountered when importing datetime values.

* Could not import datetime values
  1. I created the ‘started\_at’ and ‘ended\_at’ columns as **date** datatype. Turns out I was supposed to store them as **timestamp** datatype.
* Could not convert date datatype to timestamp datatype.

A black text on a white background

Description automatically generated

## Solution

I had to convert the columns to varchar datatype first before changing it to timestamp datatype.

 ;  
A close up of text

Description automatically generated

# Process

I will be keeping the null values for station IDs and names as those columns can still provide important information such as start and end time, type of bicycle used and membership type.

## Data checks

Check for no duplicate ride record --- passed

A close up of text

Description automatically generated

Check ride id for no null --- passed

A black text on a white background

Description automatically generated

Create VIEW for filtered data

A computer code with text

Description automatically generated

* Filter out bike trips that last less than 10 seconds as we assume the initial 10 seconds are for registering the bike. We are also excluding negative interval data.

## Data

Export data in batches according to month

A computer code with text

Description automatically generatedA screenshot of a computer

Description automatically generated

# Analyze

#### Question 1: Which group of riders ride longer and have more trips?

* Casual riders ride longer on average.
* Annual members have more trips in terms of volume.

A computer code with black text

Description automatically generatedA screenshot of a phone

Description automatically generated

* Members do have a greater number of trips but do not ride longer on average.

#### Question 2: What is the average trip duration between casual riders and annual members grouped by bicycle type?

A screenshot of a computer code

Description automatically generatedA screenshot of a computer

Description automatically generated

* Casual riders spent the most time on docked bikes at an average of 3 and a half hours.
* **Member riders do not have any trips with docked bikes.**
* Casual riders spend 20 minutes longer (> 100%) on classic bikes than annual members.
* Causal riders spend 3 minutes (> 30%) longer on electric bikes.

#### Question 3: Which bicycles do casual riders and annual members ride the most?

A computer code with text

Description automatically generatedA screenshot of a computer

Description automatically generated

* Electric bikes and classic bikes are the bicycles with the most number of trips for each membership group.
* Number of docked bike trips are very low.

Conclusion: Based on the data from question 2 and question 3, the marketing campaign should target casual riders who **ride electric bikes and classic bikes.**

#### Question 4: What are the days with the highest number of trips?

A computer code with text

Description automatically generated

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated

* Ridership among casual riders is higher closer to the end of the week.
* Ridership among members is higher during the weekdays

Conclusion: Members may be using the bikes for commuting to work as ridership numbers are lowest during the weekends. Therefore, our marketing campaign will be targeted towards **casual riders who commute to work during the weekdays.**

#### Question 5: Which month has the highest number of trips?

A computer code with text

Description automatically generated with medium confidence A screenshot of a calendar

Description automatically generated

* The **top 4 months with the most number of trips** are August, July, June and September.
  + These are closer towards the second half of the year. This could be **due to the summer season.**
* The bottom 4 months with the least number of trips are March, February, December and January
  + These are close to or are at the start of the year. This could be due to the Winter season.
* Ridership numbers start picking up after Winter.

### Summary

Based on what we have gathered from our analysis, the marketing strategy should target certain key points:

1. Target casual riders who ride the classic and electric bikes the most.
2. Target casual riders who use the bike to commute to work on weekdays (To mimic pattern of annual members).
3. Deploy the marketing campaign after the end of Winter.

# Share + Act

A close-up of a graph

Description automatically generated

Key Findings:

* Classic and electric bikes are the most used.
* Most rides take within 20 minutes to complete.
* Casual riders ride more as the week ends while annual members ride more when the week starts.
* Ridership begins to pick up rapidly in April while it starts to decline in August. (Spring to Summer)

Insights and Recommendations:

* Based on the information we have gathered; our marketing campaigns should be targeted towards working age adults who commute to work in the classic and electric bicycles. And it should be held at the end of Spring in May as ridership begin to increase rapidly.

Additional deliverables that would be helpful to include for further exploration:

* Cleaned data on the engagement rates of riders and the channels they were engaged from.
* Past yearly data to make y-o-y, q-o-q and m-o-m comparison.