**Introduction**

In this case study, I will be working as a junior data analyst. As part of this will have to follow the analysis phases of ask, prepare, process, analyze, share and act. With this process, I need to come up with a solution for the business task with data insights for the business task with insights of the data.

**Company & Scenario overview:**

In 2016, Cyclistic launched a successful bike-share offering. Since then, the program has grown to a fleet of 5,824 bicycles that are geo-tracked and locked into a network of 692 stations across Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.

Until now, Cyclistic’s marketing strategy relied on building general awareness and appealing to broad consumer segments. One approach that helped make these things possible was the flexibility of its pricing plans: single-ride passes, full-day passes, and annual memberships. Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual memberships are Cyclistic members.

Cyclistic’s finance analysts have concluded that annual members are much more profitable than casual riders. Although the pricing flexibility helps Cyclistic attract more customers, Moreno believes that maximizing the number of annual members will be key to future growth. Rather than creating a marketing campaign that targets all-new customers, Moreno believes there is a very good chance to convert casual riders into members. She notes that casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

**As part of the data analyst process ASK:**

We want to look at the future growth of the business which is making more profit in the annual members of bike riders when compared to the casual riders.

We would like to convert all the casual riders to Cyclistic members. To find a solution to this problem I have to analyze how differently casual, and members use bike share based on the seasons, day of the week & locations.

Coming up with a solution to report to the executive team, Director of Marketing &Marketing analytics team.

**Prepare:**

Got data from the link <https://divvy-tripdata.s3.amazonaws.com/index.html> 12 months from May-2021 to Apr-2022. Data is open source vetted, and reliable as per the license link below:

https://ride.divvybikes.com/data-license-agreement

Downloaded files in CSV format which contains bike riders’ info with 13 variables of more than 200K records for each month. By looking at the data figured out some incomplete information of some columns info and had missing values in columns start\_station\_name, start\_staion\_id, end\_station\_name &end\_station\_id by filtering and sorting data. All the other columns had some level of credibility to use for analysis.

**Process:**

Saved all the 12 months in Excel format. Created column to find out day\_of\_week for each month for the further phase of analysis. Did some initial

Using SQL as part of the processing phase of Data Analytics. Merging data in SQL is easy compared to Excel.

Imported 12 months of excel data into SQL management studio.

Before merging data found that all the data types are consistent and accurate.

**Analyze**:

--checking tables rows from SQL aligned with excel files

select \* from dbo.jan

select \* from dbo.feb

select \* from dbo.mar

select \* from dbo.apr

select \* from dbo.may

select \* from dbo.june

select \* from dbo.july

select \* from dbo.aug

select \* from dbo.sep

select \* from dbo.oct

select \* from dbo.nov

select \* from dbo.dec

--INSERT INTO ONE DATAFRAME FROM 12 MONTHS OF DATA

SELECT \* INTO year FROM dbo.jan

SELECT \* FROM dbo.year

INSERT INTO year

SELECT \* FROM dbo.feb

INSERT INTO year

SELECT \* FROM dbo.mar

INSERT INTO year

SELECT \* FROM dbo.apr

INSERT INTO year

SELECT \* FROM dbo.may

INSERT INTO year

SELECT \* FROM dbo.june

INSERT INTO year

SELECT \* FROM dbo.july

INSERT INTO year

SELECT \* FROM dbo.aug

INSERT INTO year

SELECT \* FROM dbo.sep

INSERT INTO year

SELECT \* FROM dbo.oct

INSERT INTO year

SELECT \* FROM dbo.nov

INSERT INTO year

SELECT \* FROM dbo.dec

--checking for the accuracy of data per month compared with excel files.

SELECT month(started\_at)month, count(\*) from total group by month(started\_at)

order by month

--ADD Ride\_length1 total of seconds value

ALTER TABLE dbo.year

ADD ride\_length1 BIGINT NULL

--Getting ride\_length1 diff of srarted\_at,ended\_at

UPDATE dbo.year

SET ride\_length1 = DATEDIFF(second, started\_at,ended\_at)

--deleting 0 value from ride\_length1

DELETE FROM year

WHERE ride\_length1 <=0

--Total avg ride\_length as per member & casual

SELECT DISTINCT member\_casual, avg(ride\_length1) avg

FROM dbo.year

GROUP BY member\_casual

--Findout monthly longest ride\_length & average ride\_length as per member\_casual

SELECT DISTINCT month(started\_at) month, member\_casual, MAX(ride\_length1) max, avg(ride\_length1) avg

FROM year

GROUP BY month(started\_at), member\_casual

ORDER BY month

--Findout monthly highest no\_of\_rides as per member\_casual

SELECT DISTINCT month(started\_at) month, member\_casual, count(ride\_id) no\_of\_rides

FROM year

GROUP BY month(started\_at), member\_casual

ORDER BY month

**Share:**

Exported all these summary reports into Excel for further analysis.

Pivoted in excel for visualization.

**ACT:**

**Recommendations based on Key findings:**

* Offering monthly or quarterly memberships to casual riders instead of annual memberships.
* To encourage casual riders, provide discounted fares during the winter months.
* In comparison to weekends, give greater promotions over the weekdays to casual riders.
* Charge more for longer rides and encourage casual riders to become members.