Business task :

Navigate speedy success of cyclistic company by attracting more riders and **design marketing strategies to convert casual riders into annual members**.

Data gathered for analysis :

The necessary data for the analysis was gathered from the link <https://divvy-tripdata.s3.amazonaws.com/index.html>. The data includes data for a time period of twelve months in the year 2021.

Processing data:

A new field called ride length was included in the data set to measure the usage of bikes by casual riders and annual members.

In the January month cyclistic dataset there were two records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the March month cyclistic dataset there were two records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the April month cyclistic dataset there were five records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the May month cyclistic dataset there were two records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the June month cyclistic dataset there were five records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the July month cyclistic dataset there were thirteen records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the August month cyclistic dataset there were twenty-nine records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the September month cyclistic dataset there were thirty-six records with negative ride length which was found using “arrange” function in R and was removed using “subset” function.

In the November month cyclistic dataset there were fifty-three records with negative ride length with a substantial duration and they were all converted to positive ride length with “absolute” function in R.

All the data sets were checked for misspellings, inconsistencies and typographical errors using the “group by” and “summarise ” functions in R.

The entire dataset was verified for uniqueness and each customer is completely unique.

Summary :

Total ride length in each month:-

As per the given data it is found that **casual riders** have a significant time difference in the ride length compared to **annual membership riders** and there was a steady usage of bikes by both the riders during the month of may,June,July,august and September. The usage of casual and annual membership riders peaked during the month of June and July.

In the month of July there were 67.6% of more **casual riders** compared to **annual membership riders.**

In the month of June there were 61.2% of more **casual riders** compared to **annual membership riders.**

The lowest ride length was recorded during the month of January and February, Hence in these two months marketing strategies can be altered to attract more customers in these two months.

Total count of casual and membership riders:

With the given data its concluded that both casual riders and membership riders have opted almost equal amount of subscriptions in the both varities(casual/membership) but **casual riders** are the ones who have a substantial ride length compared to **annual membership riders.**

Total ride length for rideable type:

Based on the different types of rideable type(classic bike, docked bike, electric bike) in the company,classic bike is the one which has the maximum ride length ,docked bike and electric bike have almost equal ride length over all the given months.

Count of rideable type:

As per the data classic bike is the most preferred rideable type and the least preferred is docked bike.

Average ride length for each month:

Comparing “average ride length in each month chart” and “total ride length in each month chart” it is found that that the duration in average ride length chart is very little compared to “total ride length in each month chart” this is due to a large number of customers get the subscription but do not ride the bike, their usage varies from no ride length to very little ride length

Overall bike usage:

Casual riders have almost 50% more ride length compared to annual membership riders. Since classic bike is the most preferred rideable type the company can advertise, educate and encourage customers to get their preferred type to increase revenue of the company.

Overall bike count for membership and casual riders:

The overall count of the casual riders is 2528977 that makes up 45.2% of the overall customers and annual membership riders add up to 3065992 that makes 54.8% of the overall customers.

Overall ride length for rideable type:

As per the given data classic bike has the maximum ride length and is customers preference. Classic bike is also the highest preferred even by count

Recommendations:

* Device marketing strategies during the month of January and February to attract more customers in this period.
* Even though the number of “casual riders” is less compared to the number of “annual membership members” the ride length of casual riders is considerably higher than annual membership riders, we can convert casual riders to annual membership riders by increasing the number of classic bike since it’s the most preferred and by advertising, educating and through exciting offers.
* The number and maintenance of docked bikes can be reduced since this is the least preferred, this would save time and unwanted use of resources like money and labour work.