Case Study: How Does a Bike-Share Navigate Speedy Success?

Cyclistic bike-share analysis case study

Google Data Analytics Course Capstone Project

Case Study Roadmap

Ask Phase

The director of marketing team at Cyclistic bike share company believes the company's future success depends on maximizing the number of **annual memberships**. Therefore, your team wants to understand how **casual riders** and **annual members** use Cyclistic bikes differently.

From these insights, your team will design a new marketing strategy to convert casual riders into annual members.

But first, Cyclistic executives must approve your recommendations, so they must be backed up with compelling data insights and professional data visualizations.

Business Task

Analyzing the Cyclistic historical bike trip data to identify trends to design a new marketing strategy to convert Casual riders into annual Members.

Key Stakeholders

● **Cyclistic** : A bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike. The majority of riders opt for traditional bikes; about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day.

● Lily Moreno : The director of marketing and your manager. Moreno is responsible for the development of campaigns and initiatives to promote the bike-share program. These may include email, social media, and other channels.

● Cyclistic marketing analytics team : A team of data analysts who are responsible for collecting, analyzing, and reporting data that helps guide Cyclistic marketing strategy. You joined this team six months ago and have been busy learning about Cyclistic's mission and business goals — as well as how you, as a junior data analyst, can help Cyclistic achieve them.

●Cyclistic executive team: The notoriously detail-oriented executive team will decide whether to approve the recommended marketing program.

Prepare Phase

Data Source ,license , organization and location

Cyclistic's historical trip data to analyze and identify trends available through this link (https://divvy-tripdata.s3.amazonaws.com/index.html)
Downloaded the last recent 12 months data from December 2020 to November 2021
The datasets have a different name because Cyclistic is a fictional company.
The data has been made available by Motivate International Inc. under this license (https://ride.divvybikes.com/data-license-agreement).

Data shape description

Last year data are in 12 CSV and excel Sheets.
Files were compresses in zip file format ,Total 12 zipped files with total 577 MB
Each file size is ranging from 45 MB to above 100 MB (large fie sizes).
Each File Containing on average 250,000 Row and 13 column.

Process Phase

Documentation of cleaning and manipulation of data

Data already have been pre processed and cleaned by DivvyBikes Team and this is their steps

- First row contains column names.
- Trips that did not include a start or end date are excluded.
- Trips less than 1 minute in duration are excluded
- Trips greater than 24 hours in duration are excluded

My Actions on Data Cleaning in order to process it for analysis

- Remove Duplicates from ride _ id column as it should be unique .
- Remove All Records that the tripe End Time is before the trip Start Time .
- Remove of un needed data columns {'start_station_name‘ , 'start_station_id‘ , 'end_station_name', 'end_station_id', 'start_lat', 'start_lng', 'end_lat', 'end_lng'}
- Created new 4 columns to get the year, month, Day and Trip Duration by minute .
- Then grouped each month (file) by (Year – Month – Day of Week) with Customer Type to report (Count , Average and Sum of Tripe Duration Time in minutes .
- Then combined all table files through python scripti and created the final Data sheet for analysis and visualizing the insights

Tools used

- I used Excel To Explore data sheet and plan for required action.
- I used python to manipulate the data and Do cleaning actions on it as the data is large size
- Created a function in python to repeat same steps for all the 12 files and save the cleaning process steps
- Excel sheet to share data visualization
- Power point to present the project and analysis insights

Analysis Phase

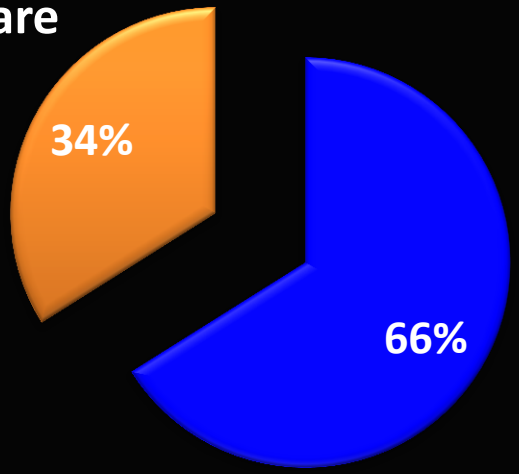
- Creating new column to calculate the trip duration.
- Time format was in in time delta format .
- Converted it to minutes .
- Grouped All Trip Duration by customer type (Casual – Annual Member)
- Grouped Trip Duration by customer type compared to Week Days.
- Grouped Trip Duration by customer type compared to Week Days and months.
- Get the average of trip duration , Count of Trips and Duration Sum .
- Below is the table shape for final analysis for each month.
- Then all tables are combined to start analysis then visualization.
- This has resulted in some interesting insights , that I will show in the coming Slides

			mean		Count		sum	
			Trip_Minutes		Trip_Minutes		Trip_Minutes	
		member_casual	casual	member	casual	member	casual	member
year	month	Day_of_Week						
2021	1	Friday	23	12	2,838	12,643	65,134	155,862
		Monday	20	13	2,088	11,135	42,788	147,123
		Saturday	32	14	4,003	12,319	127,140	166,989
		Sunday	29	13	2,864	8,872	83,668	118,498
		Thursday	22	12	2,348	11,966	50,844	146,657
		Tuesday	22	12	1,882	10,569	42,074	127,201
		Wednesday	26	13	2,094	11,207	53,720	150,768

Share Phase

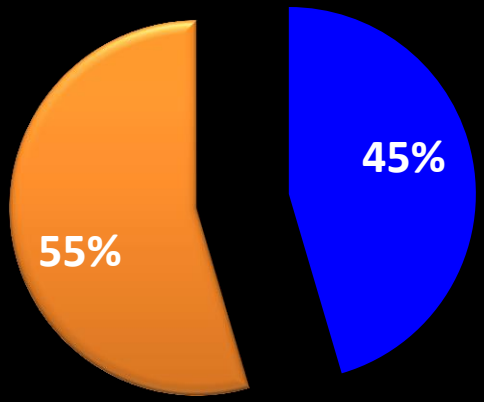
In the coming slides I will share the insights from Cyclistic Data Analysis

Total Trips Time Share



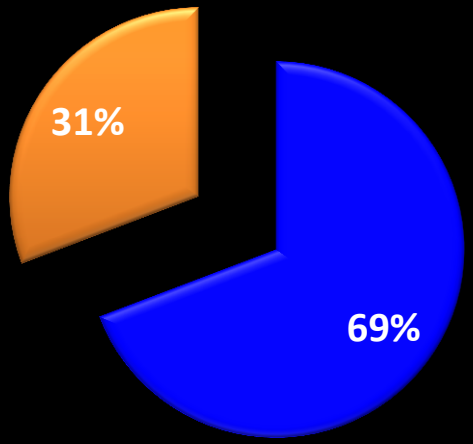
Casual Trips Duration Sum Member Trips Duration Sum

Trips Count Share



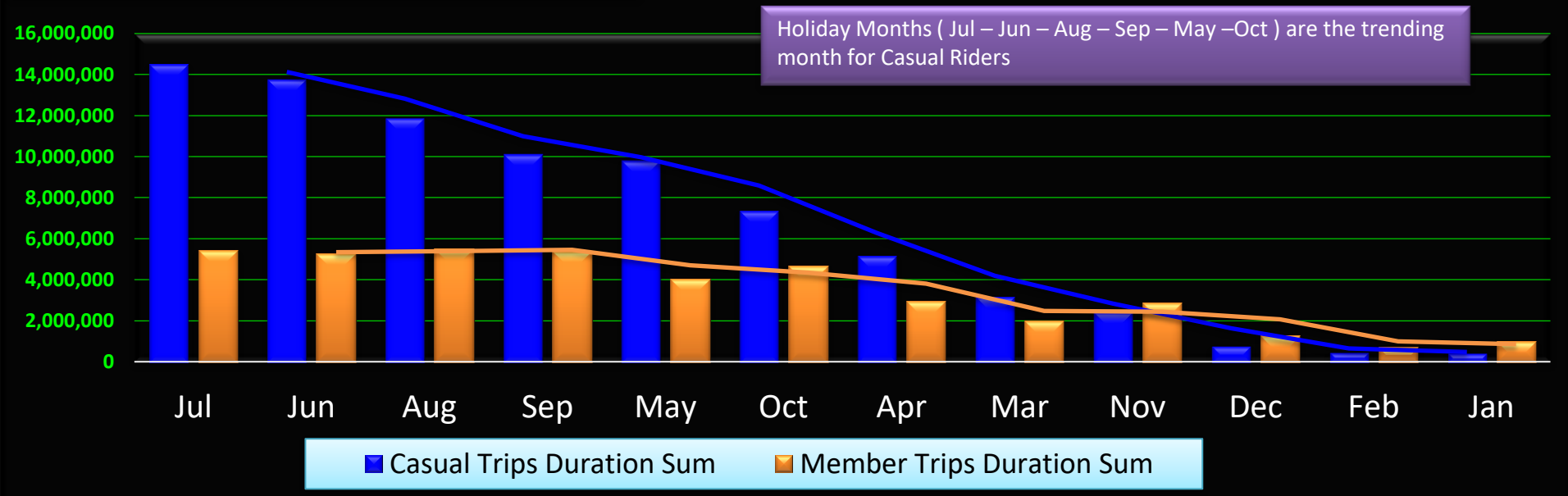
Casual Trips Count Sum Member Trips Count Sum

Trip Duration Average Share

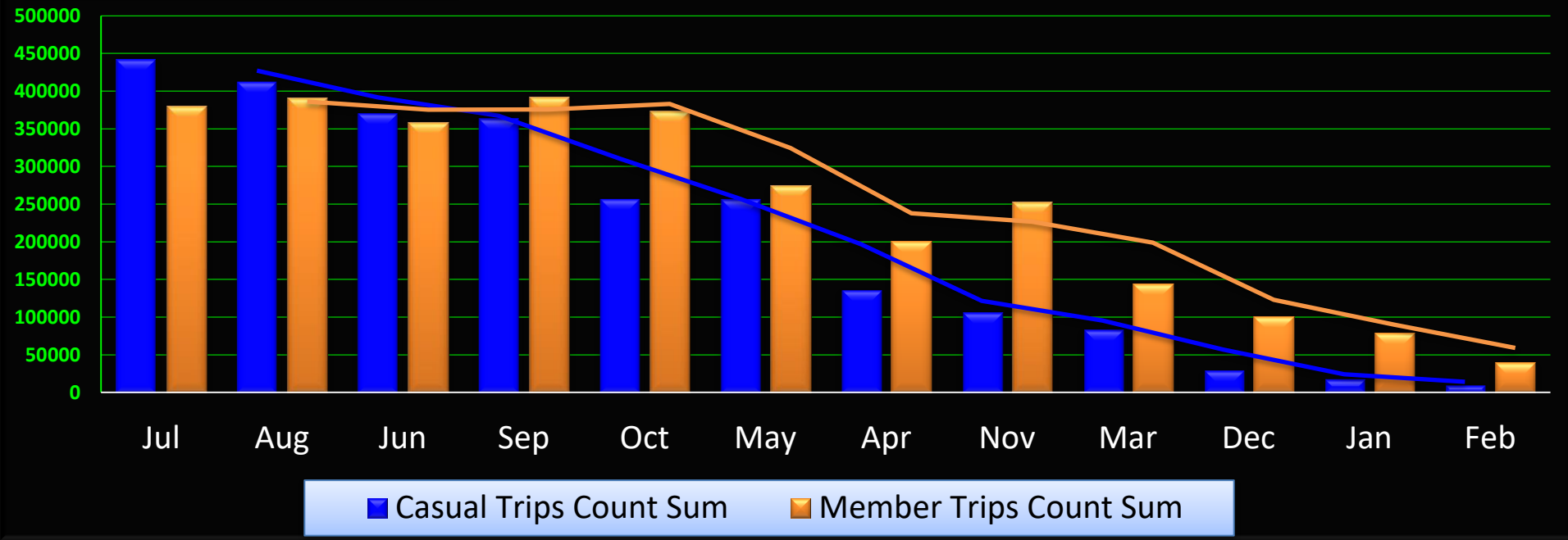


Casual Trips Duration Average Member Trips Duration Average

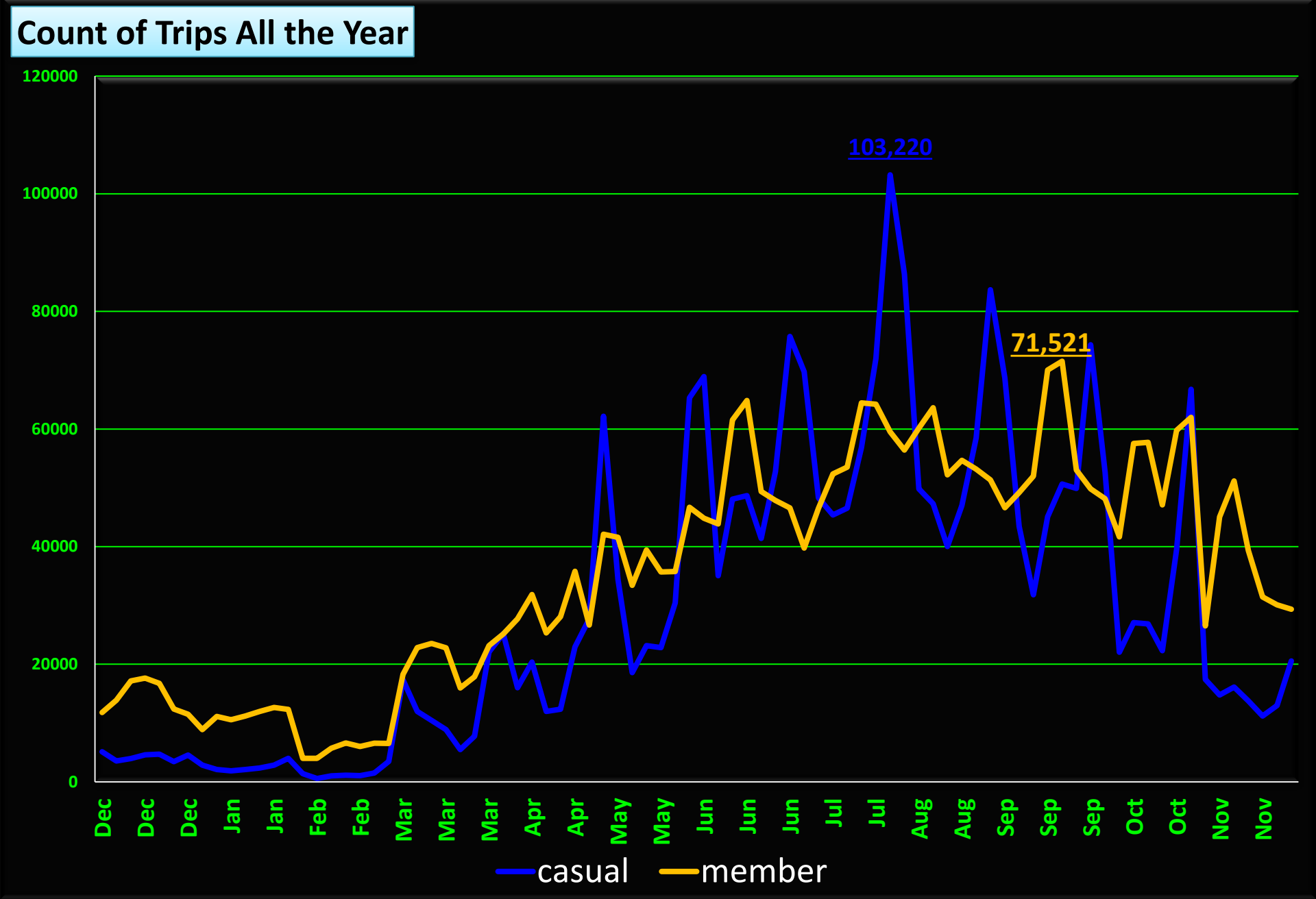
Monthly Tripe *Duration Sum* / minute



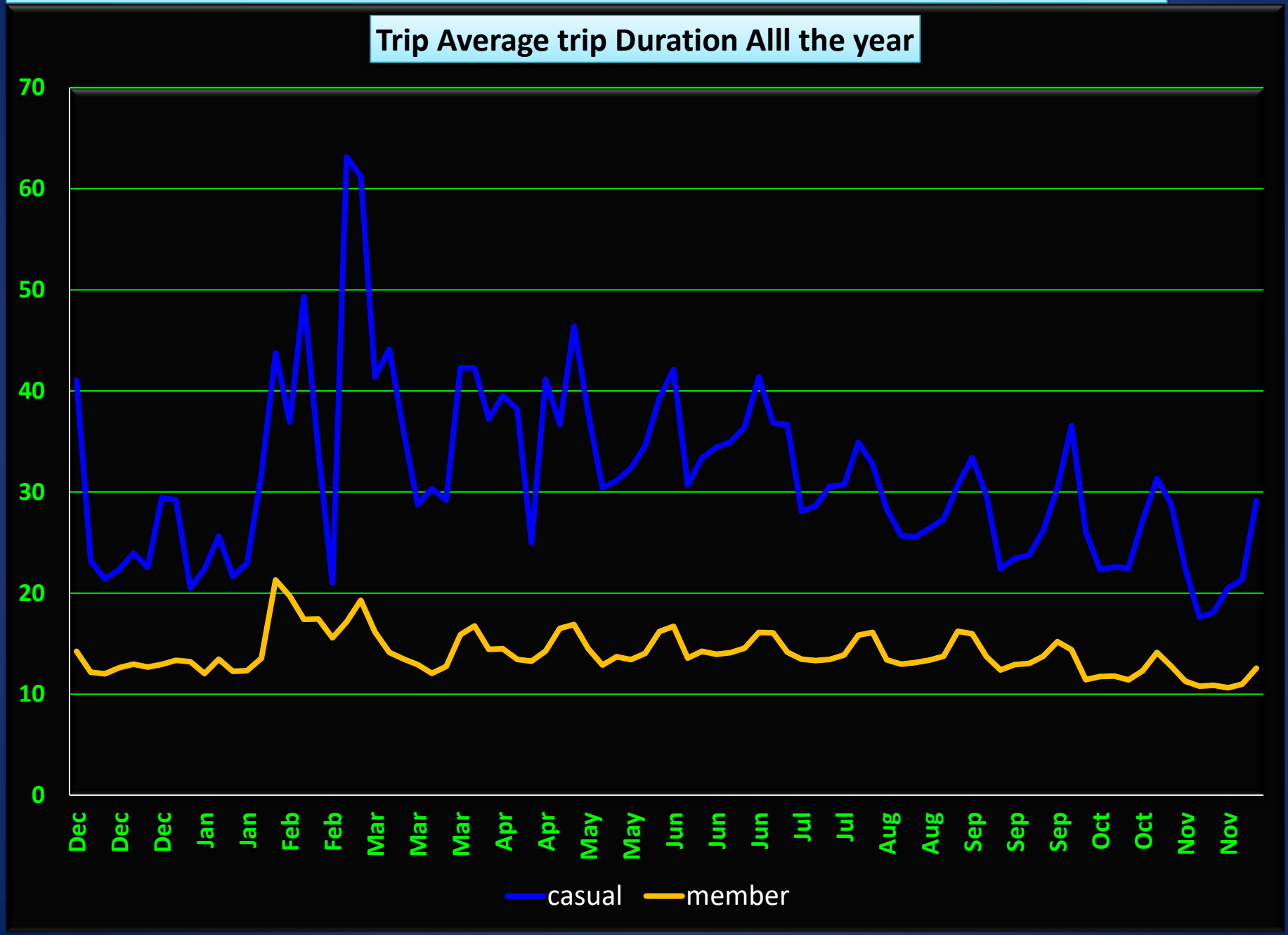
Monthly Total Trips Count



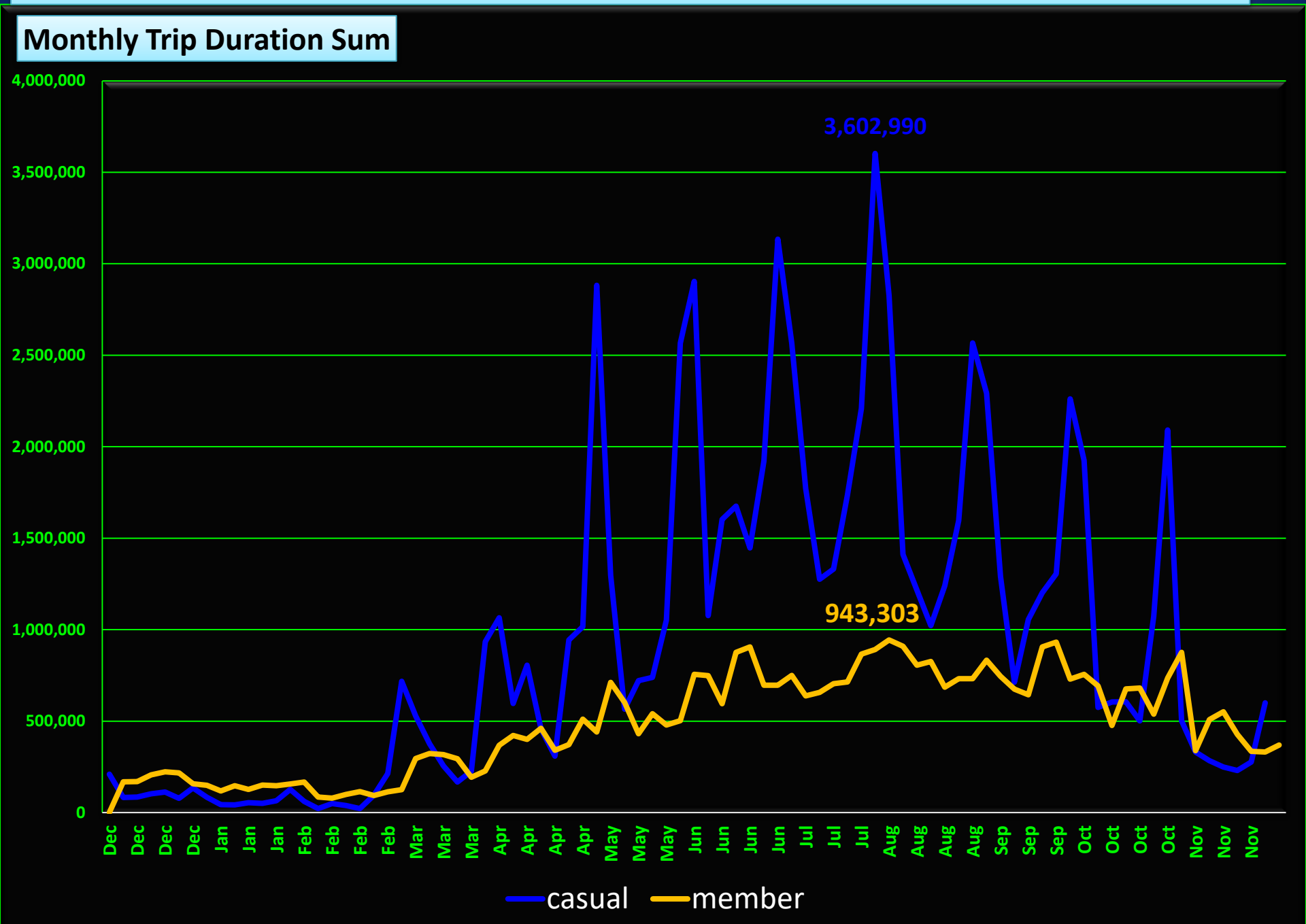
There is **casual Riders** peaks in number of trips. Specially from April to October season (**Holiday Season**).



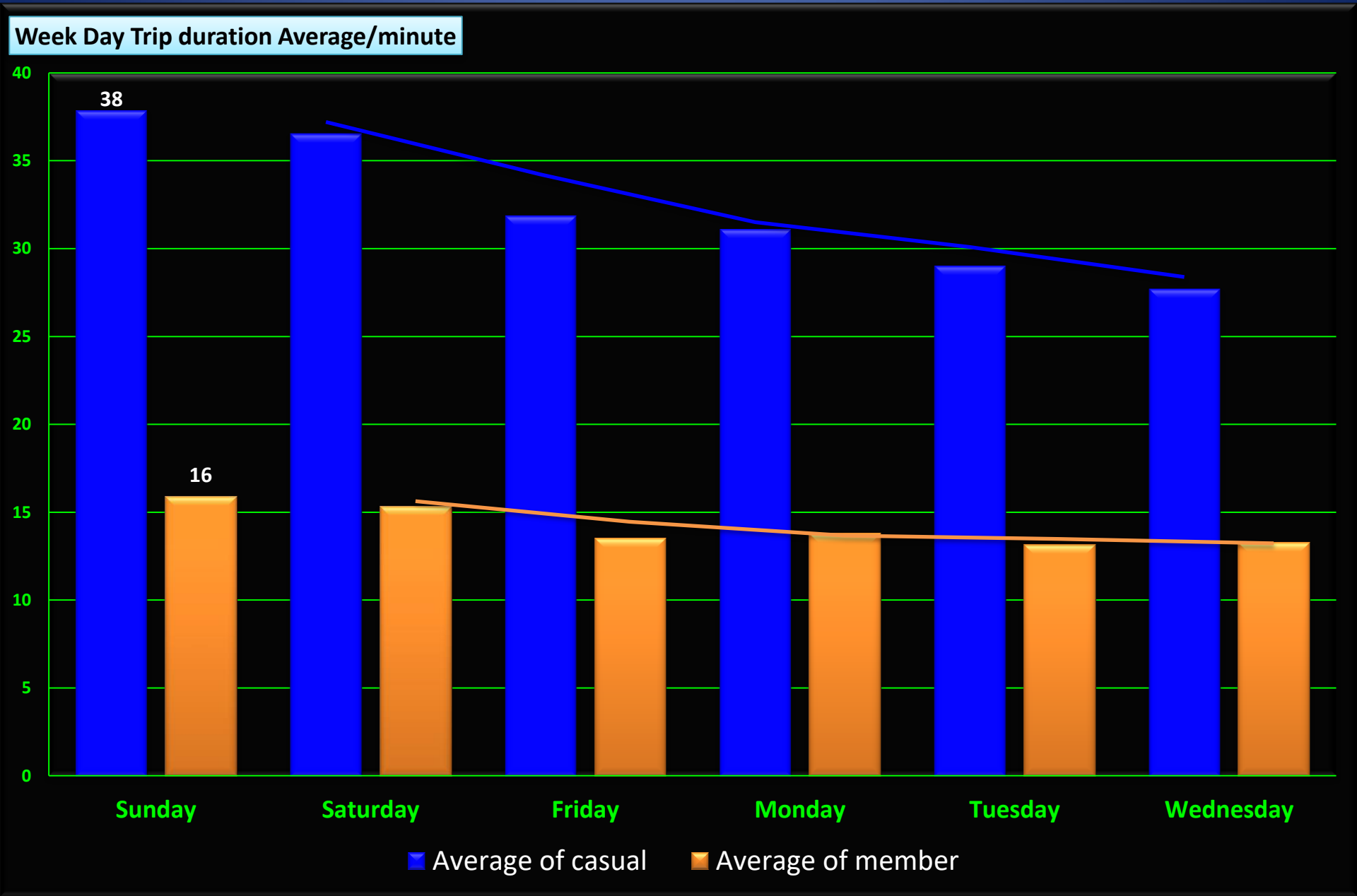
Casual trip duration Average is higher than Members trip duration All year days



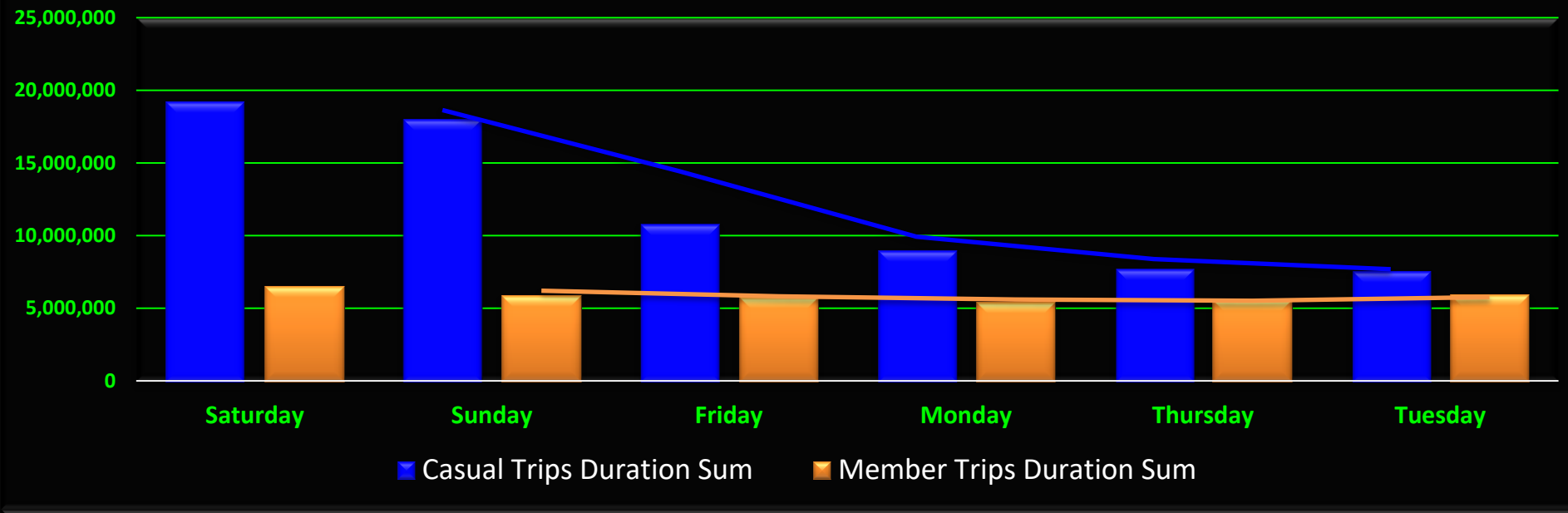
Casual trip duration Sum is higher than Members trip duration most of year days (Holiday Months)



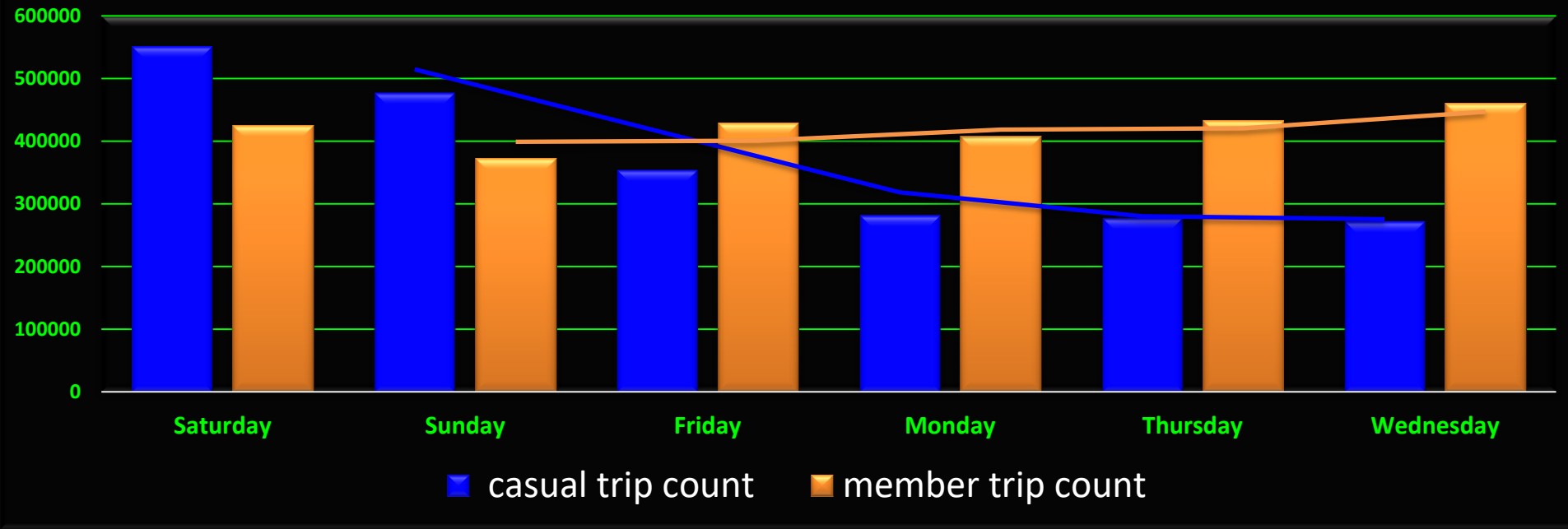
Weekly Casual trip duration Average is higher in Holiday Day (Sunday – Saturday



Week Days Trips Duration Sum /minute



week Day Trip Count Sum / Trip



Act Phase

Insights from Data Analysis on the behavior of casual Customers

- Casual Customers Share in Trips Duration last year is 66 % which is double the share of member 34%.
- which reflects that there is a good chance to Convert Some of casual Customers to Annual Members.
- The Average Trip Duration of Casual Customers is always higher than Annual Members.
- The Data Shows peaks of causal Customers Usage in Holyday Months and Days as shown in graph.

- **Based on this high level Analysis ,**
- **I See that causality increases in Holidays while Annual members are constant usage Customers.**
- **Also Casual Customer Usage without holidays are always higher than Annual Members.**

- **I recommend more Focus Casual Customers Behavior to be able to design a tailored membership for casual customers based on their needs.**

Thank You Analysis Done by
Ahmed Salah Mohamed Abo ElEneen
Data Analyst
Google Certified
+201064044443
wareed2015@gmail.COm