Cyclistic Case Study Documentation Draft

Introduction

This project has been carried out following the five steps of data analysis:

- Ask
- Prepare
- Process
- Analyse
- Share

N.B the last step of data analysis process has not been included as this case study is only meant for portfolio showcase

The following documentation contain a detail explanation of our data analysis has been carried out through each process.

Ask

What is the business task?

The business task is to identify new marketing strategies aimed at converting casual riders into annual members.

What is the main question of our analysis?

How do annual members and casual riders use Cyclistic bikes differently?

This question has been specifically assigned by the director of marketing, Lily Moreno.

Prepare

The type of data used for this project: Cyclistic's historical trip data-2023. To note that the only available data is from the month of January to July.

The data is organised in:

- ride_id
- rideable_type (type of bicycle used by the client)
- started_at (date and time of the ride start)
- ended_at (date and time of the ride end)
- start_station_name (name of the station the client picked the bike)
- start_station_id (id of the start station)
- end_station_name (name of the station the client left their bike)
- end_station_id (id of the end station)
- start_lat (starting latitude)
- start_lng (starting longitude)
- end_lat (ending latitude)
- end_lng (ending longitude)
- member_casual (if the client is an annual or casual rider)

There are no issues with bias or credibility in this data.

The data has been made available by Motivate International Inc. under this license.

Process

The software used for this project is Microsoft Excel.

The following contains a changelog of the data used for this project.

#Changelog

This file contains all the data of Cyclistic of each month in 2023.

#Added

4. Added four columns:
- started_time: to extract the start time from column started_at. Column format: Time 37:30:55
- ended_time: to extract the end time from column ended_at. Column format: Time 37:30:55

- ride_length: to calculate the length of each ride by subtracting the column “started_at” from the column “ended_at”. Column format: Time 37:30:55

- day_of_week: to calculate the day of the week that each ride started. Column format: Number

#Changed

1. The file has been changed from .csv to .xlxs file
2. Changed data from text to column. Each column is in a general format except started_at, ended_at, start_lat, start_lng, end_lat and end_lng which are in a text format.
3. Fixed column width with auto width function

#Deprecated

N/A

#Removed

N/A

#Fixed

N/A

#Security

N/A

N.B This case did not require any additional cleaning or data addition. However, in a real-life circumstance I would have follow these steps before adding or manipulating any addition data:

- check for duplicated data

- remove outdated data

- fill incomplete data

- check for inaccurate or incorrect data by fixing misspellings, inconsistent capitalization and incorrect punctuation and other typos

- check for inconsistent data and formatting

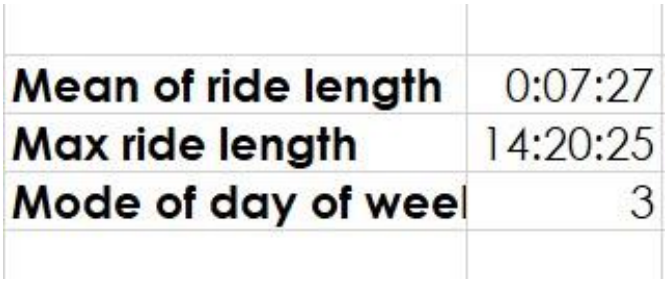
- remove inconsistent formatting

- check for duplicates once again for

Analyse

In the spreadsheet of each month of the year 2023, it has been calculated:

- the mean of ride time
- the max ride times
- the most recurring day of the week

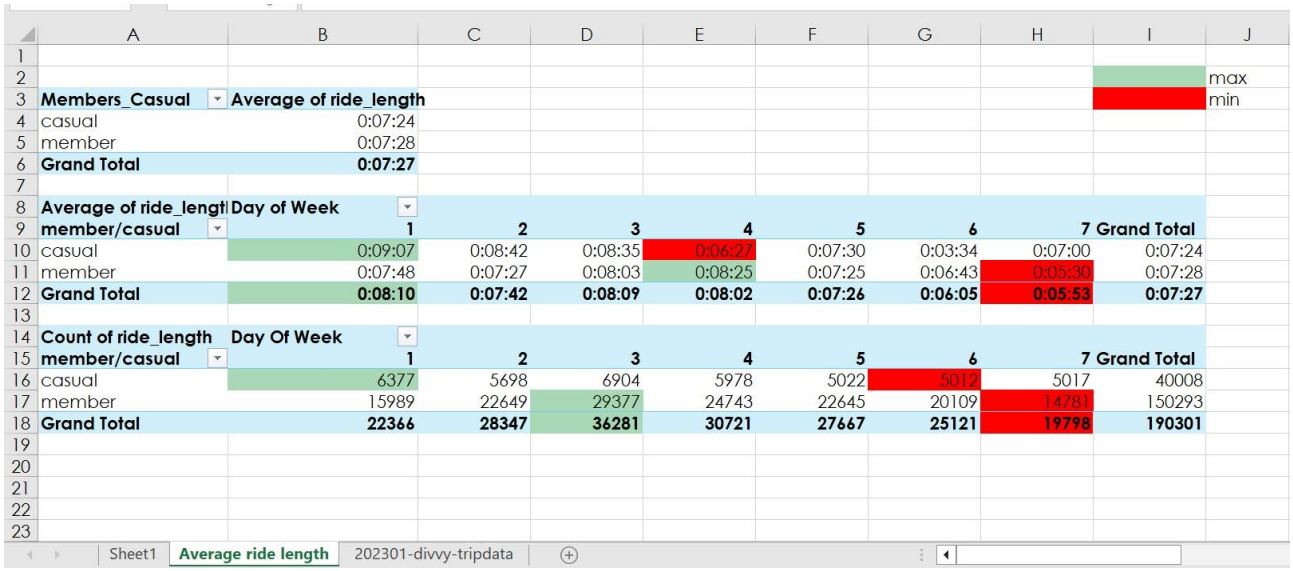


Mean of ride length	0:07:27
Max ride length	14:20:25
Mode of day of weel	3

Figure 1, mean of ride, max ride time and most recurring day of the month of January

In addition, a pivot table has been created in each spreadsheet to calculate the following:

- the average ride time for members and casual riders.
- the average ride time for riders by the day of week.
- the number of rides by day of the week.



	A	B	C	D	E	F	G	H	I	J
1										
2										max
3	Members_Casual	Average of ride_length								min
4	casual	0:07:24								
5	member	0:07:28								
6	Grand Total	0:07:27								
7										
8	Average of ride_lengt	Day of Week								
9	member/casual	1	2	3	4	5	6	7	Grand Total	
10	casual	0:09:07	0:08:42	0:08:35	0:06:27	0:07:30	0:03:34	0:07:00	0:07:24	
11	member	0:07:48	0:07:27	0:08:03	0:08:25	0:07:25	0:06:43	0:05:30	0:07:28	
12	Grand Total	0:08:10	0:07:42	0:08:09	0:08:02	0:07:26	0:06:05	0:05:53	0:07:27	
13										
14	Count of ride_length	Day Of Week								
15	member/casual	1	2	3	4	5	6	7	Grand Total	
16	casual	6377	5698	6904	5978	5022	5012	5017	40008	
17	member	15989	22649	29377	24743	22645	20109	14781	150293	
18	Grand Total	22366	28347	36281	30721	27667	25121	19798	190301	
19										
20										
21										
22										
23										

Sheet1 | Average ride length | 202301-divvy-tripdata

Figure 2, pivots table of the month of January

All analysis has been then merged into a single spreadsheet to calculate the followings:

- Average ride time per month
- Total rides per month
- Casual and annual members' ride time per month
- Total casual and annual rides per month

Months	Average ride_length
January	0:07:27
February	0:13:32
March	0:13:05
April	0:17:13
May	0:19:02
June	0:19:59
July	0:21:44
Highest	0:21:44
Lowest	0:07:27

Months	Total
January	190301
February	190445
March	258678
April	426590
May	604827
June	719618
July	767650
Highest	767650
Lowest	190301

Figure 3, average ride time and total rides of month of January

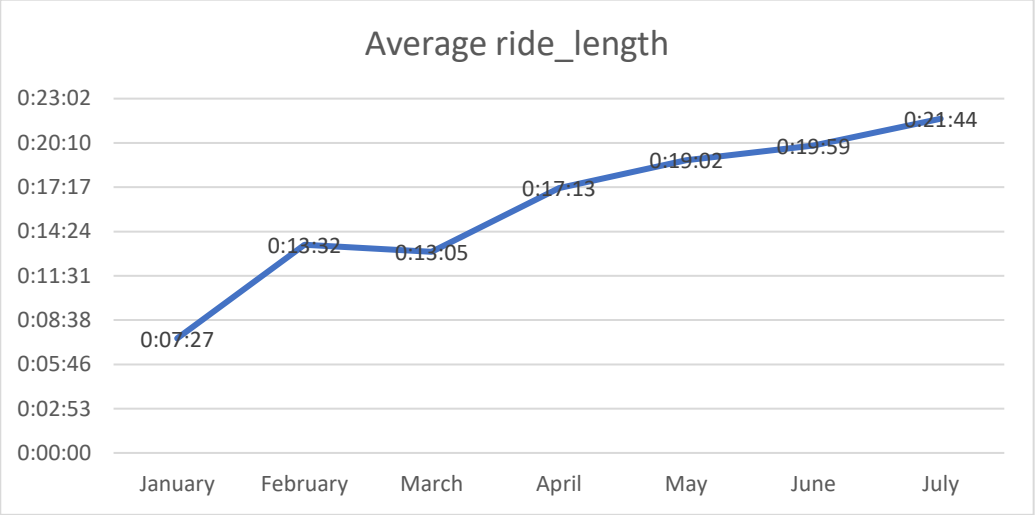
Months	Casual avg ride_length	Annual avg ride_length
January	0:07:24	0:07:28
February	0:23:12	0:10:43
March	0:21:25	0:10:27
April	0:27:40	0:11:42
May	0:28:31	0:13:02
June	0:29:24	0:19:59
July	0:32:20	0:13:41

Months	Total casual	Total annual
January	40008	150293
February	43016	147429
March	62201	196477
April	147285	279305
May	234181	370646
June	301230	418388
July	331358	436992

Figure 4, casual and annual ride time

Observations:

1) The lowest average time is in January and the highest average time is in July



2) The general most recurring day is Tuesday.

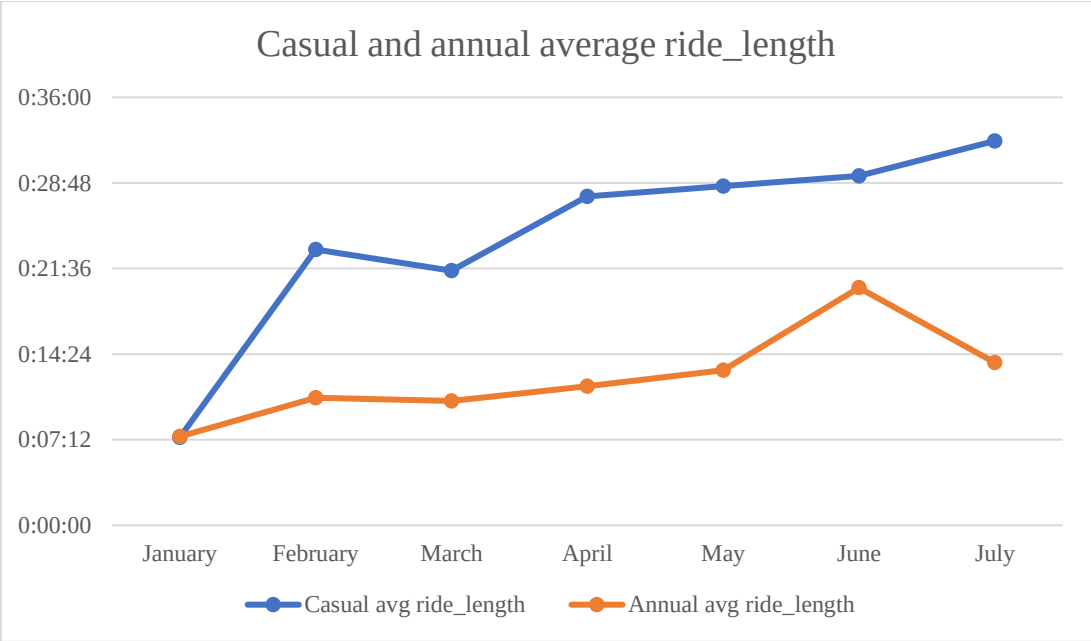
Months	Most recurring day
January	3
February	3
March	4
April	7
May	3
June	6
July	7

However, the most recurring day among casual riders is Saturday and Sunday. On the other hand, the most recurring day among annual riders is Tuesday.

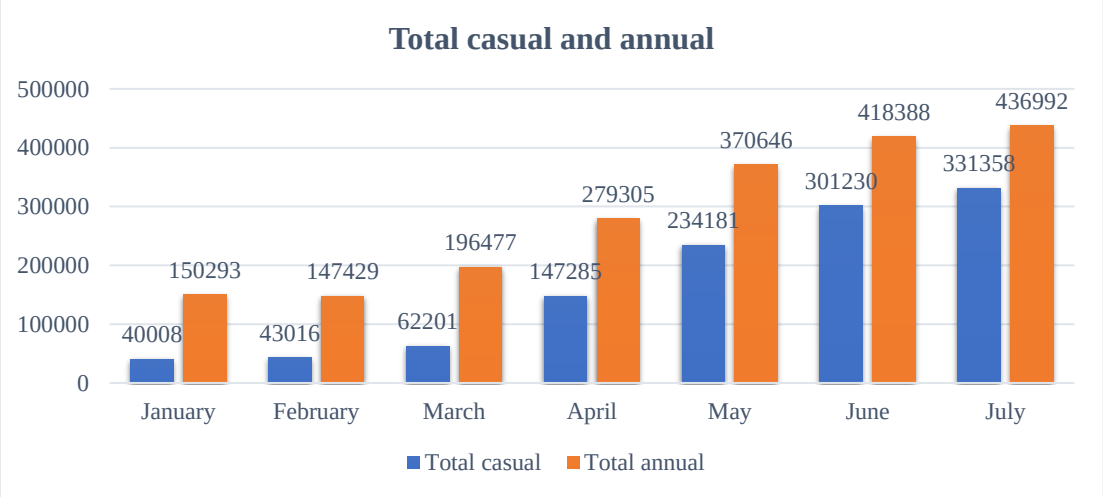
Months	Casual recurring day
January	1
February	1
March	4
April	7
May	1
June	7
July	7

Months	Annual recurring day
January	3
February	3
March	4
April	5
May	3
June	6
July	2

3) The average ride time among casual riders is higher than annua members



4) The total annual rides are higher than total casual rides



Conclusions:

1) Peak season is end of Spring till end of Summer
2) Casual riders prefer weekends while annual riders prefer weekdays
3) Casual riders prefer longer trips in comparison to annual members. However, annual riders are the one to use Cyclistic bikes more frequently.

Recommendations:

First recommendation

1) To make single ride passes available for only 20 minutes, the average casual ride time. After 20 mins, the rider will have to pay an additional 7 euro for every 10 mins passed which will be paid at the end station.
2) Increase of full day pass. For example, if the full day pass is currently at 35 euro then it should be changed to 42 euro
3) Introduction of weekends membership and instalments. Most casual riders prefer to ride only during weekends. So, to accommodate them, Cyclistic could benefit of this new flexible membership.

Second recommendation

Casual riders need to find value and benefits in converting to an annual member. Therefore, a second recommendation would be the increase of additional services to only members.

Benefits of memberships could be like the following

Weekend Membership	Annual Membership
Bike reservation up to two people	Bike reservation up to four people
Second priority access during peak times	First priority access during peak times
Available every week: Friday to Sunday	Available all year round
Renewal every week	Renewal every year and if the member refers our program to a new members, the client will get a 20% discount next year.

Third recommendation

Implementation of marketing advertisement during Spring- Summer as it is our peak season. The goal is to target as much as many casual riders to inform them about new services and memberships.

Share

A presentation of this project has been made through PowerPoint.