Cyclistic Year Report

Youi

2024-04-19

Introduction

Scenario

Cyclistic is a bike-share program that features more than 5,800 bicycles and 600 docking stations. Finance analysts have concluded that annual members are much more profitable than casual riders. Hence, the director of marketing has set the goal:Design marketing strategies aimed at converting casual riders into annual members. In order to do that, the marketing analyst team needs to better understand how annual members and casual riders differ. Raw data of Cyclistic trip can be found here.

Descriptive analysis

Summary

The previous 12 months of Cyclistic trip data, June 2021 till May 2022, has 5,860,776 rows. 646 rows travel time (variable duration) are smaller than 0. After removing those data, 5,860,130 rows are left.

NA values

Only variable start_station_id, start_station_name, end_station_id, end_station_name, end_lat, end_lng have NA values. First four variables can be substituted by position variable(lat and lng). Last two variables' NA is not big portion to the whole data.

Data summary

Name	Piped data
Number of rows	5860130
Number of columns	28
_____	
Column type frequency:	
character	5
difftime	3
factor	3
numeric	12
POSIXct	5
_____	
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
ride_id	0	1.00	16	16	0	5860130	0
start_station_name	823145	0.86	3	53	0	1105	0
start_station_id	823142	0.86	3	44	0	1063	0
end_station_name	877927	0.85	9	53	0	1112	0
end_station_id	877927	0.85	3	44	0	1068	0

Variable type: difftime

skim_variable	n_missing	complete_rate	min	max	median	n_unique
started_hms	0	1	0 secs	86399 secs	55914 secs	86370
ended_hms	0	1	0 secs	86399 secs	56760 secs	86383
duration	0	1	1 secs	3356649 secs	681 secs	24839

Variable type: factor

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
Bikes	0	1	FALSE	3	cla: 3217479, ele: 2362211, doc: 274440
Users	0	1	FALSE	2	mem: 3300629, cas: 2559501
wday	0	1	TRUE	7	Sat: 987105, Sun: 864748, Fri: 819735, Tue: 811755

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
start_lat	0	1	41.90	0.05	41.64	41.88	41.90	41.93	45.64
start_lng	0	1	-87.65	0.03	-87.84	-87.66	-87.64	-87.63	-73.80
end_lat	5036	1	41.90	0.05	41.39	41.88	41.90	41.93	42.17
end_lng	5036	1	-87.65	0.03	-88.97	-87.66	-87.64	-87.63	-87.49
year	0	1	2021.26	0.44	2021.00	2021.00	2021.00	2022.00	2022.00
month	0	1	7.25	2.58	1.00	5.00	7.00	9.00	12.00
day	0	1	15.54	8.69	1.00	8.00	15.00	23.00	31.00
hour	0	1	14.21	5.07	0.00	11.00	15.00	18.00	23.00
start_lng_rd	0	1	-87.65	0.03	-87.84	-87.66	-87.64	-87.63	-73.80
start_lat_rd	0	1	41.90	0.05	41.64	41.88	41.90	41.93	45.64
end_lng_rd	5036	1	-87.65	0.03	-88.97	-87.66	-87.64	-87.63	-87.49
end_lat_rd	5036	1	41.90	0.05	41.39	41.88	41.90	41.93	42.17

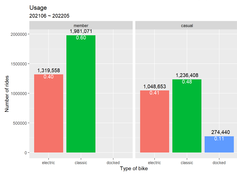
Variable type: POSIXct

skim_variable	n_missing	complete_rate	min	max	median	n_unique
started_at	0	1	2021-06-01 00:00:38	2022-05-31 23:59:56	2021-09-23 17:33:16	4896398
ended_at	0	1	2021-06-01 00:06:22	2022-06-02 11:35:01	2021-09-23 17:49:24	4893051
year_month	0	1	2021-06-01 00:00:00	2022-05-01 00:00:00	2021-09-01 00:00:00	12
started_round	0	1	2021-06-01 00:01:00	2022-06-01 00:00:00	2021-09-23 17:33:00	461901
ended_round	0	1	2021-06-01 00:06:00	2022-06-02 11:35:00	2021-09-23 17:49:00	462441

Visualize the data

1.Usage

Check on the usage of the previous 12 months :
Member users used classic bikes more often than electrics. Casual users used classic slightly more often than electric bikes (48% and 41%).



Assumption in **bold**:

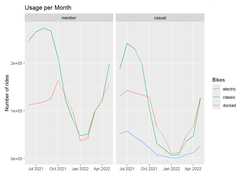
1. Members are limited to classic and electric usage.
2. Members commuted in **short distance**, using classic bike more often, classic bike is more **cheap**.
3. Casuals used **electric bike** to travel **far on weekends**.
4. Casuals rode bikes **for leisure**.

Usage per month

During the previous 12 months (June 2021 till May 2022), usage varied a lot with months. The total number of rides reached highest at around July and August. Fewer users used the bike, especially for casual member, during winter. The total number of rides reached lowest at January.

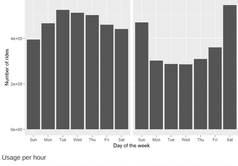
Member riders had higher classic bikes usage (266,818) before September 2021, almost twice the number of electric bikes (125,439). After October 2021, usage of both classic and electric bikes started to drop, and hit lowest point (48,093 and 37,157) at January 2022. Usage increased almost at the same speed of dropping and reached 197,971 and 156,472 for classic and electric bikes at May 2022.

Casual riders had higher classic bikes usage (241,489) before July 2021, almost twice the number of electric bikes (142,869). After October 2021, usage of both classic and electric bikes started to drop, and hit lowest point (6,974 and 10,585) at January 2022. Usage increased almost at the same speed of dropping and reached 126,076 and 127,931 for classic and electric bikes at May 2022.



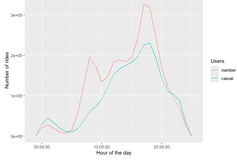
Usage per day

Relatively speaking, in member users group, they used bikes more frequently during weekdays, while casual users used more frequently on weekends. Besides, casual users have higher number of rides than member users' on weekends.



Usage per hour

There are two peaks for member users, around 8:00 and 17:00, while casual users' two peaks are around 1:15 and 17:00. Those who use bike around 1:15 use classic bike for long distance travel, long duration, like road trip.

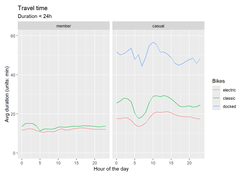


2.Travel time

Few users used bikes more than 24 hours.

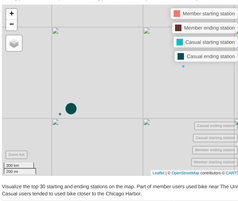
```
## # A tibble: 8 × 4
## # Groups:   Users, Bikes [5]
##   Users  Bikes    duration_less_than_24h     n
##   <fct>  <fct>                     <dbl> <int>
## 1 member electric                      1 1319558
## 2 member classic                       0     534
## 3 member classic                       1 1980537
## 4 casual electric                      1 1048853
## 5 casual classic                       0    1349
## 6 casual classic                       1 1234739
## 7 casual docked                        0    1413
## 8 casual docked                        1  273027
```

For every starting time, casual user's average travel time are more than 20 mins, member average travel time are less than 20mins.

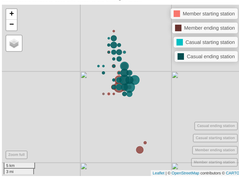


3.Position

Ending point is more far-spread than starting point. Casual user spread even more widely. The size of circle based on number of rides.



Visualize the top 30 starting and ending stations on the map. Part of member users used bike near The University of Chicago(Hyde Park district). Casual users tended to used bike closer to the Chicago Harbor.



Conclusion

The characteristic of each group using bikes can be summarized as below.

1. Members used bikes to commute **between** work(school) and home.
2. Members commuted in **short distance**, using classic bike more often.
3. Casuals used bikes to commute **from** work to home.
4. Casuals used **docked and electric bikes** to travel **far**.
5. Casuals rode bikes **for leisure**(on weekends).
6. Casuals bike ride's ending points are more far-spread than members'.
7. Casuals bike ride's **average duration**(15~30mins) are **longer** than member's(10~15mins).
8. There's **seasonal trend** in number of rides for both members and casuals users.

To enhance the biking experience for both members and casual users, the company could consider implementing several suggestions and actions. Firstly, providing incentives or discounts for members who use bikes more frequently could encourage greater adoption among this group. Additionally, offering maintenance services or bike-sharing programs specifically tailored to short-distance commuters could further incentivize bike usage among members. For casual users, introducing special weekend packages or promotions for docked bike rentals could attract more customers looking to explore farther destinations during their leisure time. Furthermore, organizing group biking events or tours on weekends could tap into the leisure aspect of biking and attract both members and casual riders alike. Lastly, analyzing seasonal trends in bike usage could inform strategic planning, such as adjusting bike availability or promotional efforts during peak riding seasons to maximize engagement. By implementing these suggestions, the company can promote biking as a convenient, enjoyable, and sustainable mode of transportation for all its users.