

Google Data Analytics Capstone Project: How Does a Bike-Share Navigate Speedy Success?



Introduction

The Cyclistic case study is part of the Google Professional Data Analytics Certification. In this case study, I play the role of a junior data analyst at Cyclistic. I will follow basic data analysis process steps to solve core business problems.

I'm a junior data analyst working on the marketing analyst team at Cyclistic, a bike-sharing company in Chicago. The marketing manager believes the company's future success will depend on maximizing annual memberships. Therefore, as a marketing team, he wants to understand how cyclistic bikes are used in different ways according to membership types. Here, my task will be to review and analyze 12 months of data, to obtain insights from this data through certain visualizations and present it to the team. I have analyzed very specific data and tried to support these insights with clear graphs so that my suggestions can be taken into account.

About the Company

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. Moreno has set a clear goal: Design marketing strategies aimed at converting casual riders into annual members. In order to do that, however, the marketing analyst team needs to better understand how annual members and casual riders differ, why casual riders would buy a membership, and how digital media could affect their marketing tactics. Moreno and her team are interested in analyzing the Cyclistic historical bike trip data to identify trends.

In this case-study, I will be drawing upon the 6 phases of the analysis process, namely Ask, Prepare, Process, Analyze, Share and Act.

I downloaded the datasets from the link below.

[The previous 12 months of Cyclistic trip data](#)

Phase 1 – ASK

Three questions will guide the future marketing program:

1. How do annual members and casual riders use Cyclistic bikes differently?
2. Why would casual riders buy Cyclistic annual memberships?
3. How can Cyclistic use digital media to influence casual riders to become members?

1.1 Business Task

Analyze data to gain insights into how users use Cyclistic's bikes by membership type and to identify trends based on Cyclistic's marketing strategy.

1.2 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.

1.3 A clear statement of the business task

The main purpose in this job role is to analyze Cyclistic data and gain insights for marketing trends based on user orientation and driving data and share them with key stakeholders. Specific and clear visualizations should be made and ready for sharing so that all stakeholders can understand the data.

Phase 2 – PREPARE

Guiding Questions

Q1: Where is your data located?

A1: The data is available at the address I gave the [link](#) above. The data I have downloaded from this address have been handled in a way to cover 12 months, January 2022 and December 2022.

Q2: How is the data organized?

A2: Many tools and programs can be used to edit data. I chose the R programming language for my work. In the **R programming language**, I combined 12 months of data into a single data frame, then added the necessary columns and removed the columns that would not be useful to us in our analysis.

Q3: Are there issues with bias or credibility in this data? Does your data ROCCC?

A3: Yes, the data appears to be ROCCC. There don't seem to be any issues with bias or credibility in the data, as it consists of trip data from Cyclistic's own bike riders and is made available under license by Motivate International Inc. Additionally, the data appears to be reliable, original, comprehensive, current, and cited, meeting the criteria for ROCCC.

Q4: How are you addressing licensing, privacy, security, and accessibility?

A4: It is important to adhere to any licensing restrictions when using publicly available data. As mentioned in the prompt, the data is made available under a specific license, and users should review and comply with any requirements or restrictions specified by the license. To address privacy and security concerns, any personally identifiable information should be removed or obscured from the data. Additionally, it is important to ensure that the data is accessible and usable by anyone who needs to work with it.

Q5: How did you verify the data's integrity?

A5: To verify the integrity of the data, it is important to check the data for any inconsistencies or errors. This may include checking for missing or incorrect data, identifying outliers or unusual values, or comparing data with other sources to ensure it is correct. I carried out this analysis with simple and fast methods in R language and as a result I checked that the data is consistent. I identified missing and incorrect data types, removed them and made them ready for analysis.

Q6: How does it help you answer your question?

A6: Cyclistic trip data helps answer the question of how annual members and casual cyclists use the bike share program differently. Analyzing the patterns in the data, my work to identify usage differences between the two groups and to understand how each group uses the bikes has helped me come up with answers that are very effective in answering marketing strategy questions.

Q7: Are there any problems with the data?

A7: I found that there are some deficiencies and incorrect data types in the data. I solved them using methods in R language and made them suitable for analysis.

Phase 3 – PROCESS

Guiding Questions:

Q1: What tools are you choosing and why?

A1: I used the R programming language to process the data and make it ready for analysis. Thanks to its simple syntax and fast usage, I both consolidated my knowledge and analyzed and organized the data. I visualized the data I edited in the R language in the Tableau program. Thus, I developed my skills in both programs.

Q2: Have you ensured your data's integrity?

A2: Yes, I managed to ensure the integrity of the data. I added 12 months of data to a single data frame, making it ready for a year's analysis.

Q3: Have you documented your cleaning process so you can review and share those results?

A3: Yes, I have documented these operations and analysis process using R markdown. You can reach the markdown in my Github account via [this link](#).

Phase 4 – Analyze

Q1: How should you organize your data to perform analysis on it?

A1: First of all, the driving data had to be grouped according to user types, hours, time zones of the day, days of the week, months and seasons, and I organized the data to be that way. As can be seen in the R Markdown document that I just shared during the process phase, time columns were created and the driving data of the users were classified according to the time type. In addition to examining the average and total driving times of the user types, the driving time lengths were also extracted from the data in the data frame, such as the end and start time, and included in the analysis, and important insights were obtained.

Phase 5 - SHARE

Below is a summary of my analysis to shed light on solutions to business problems:

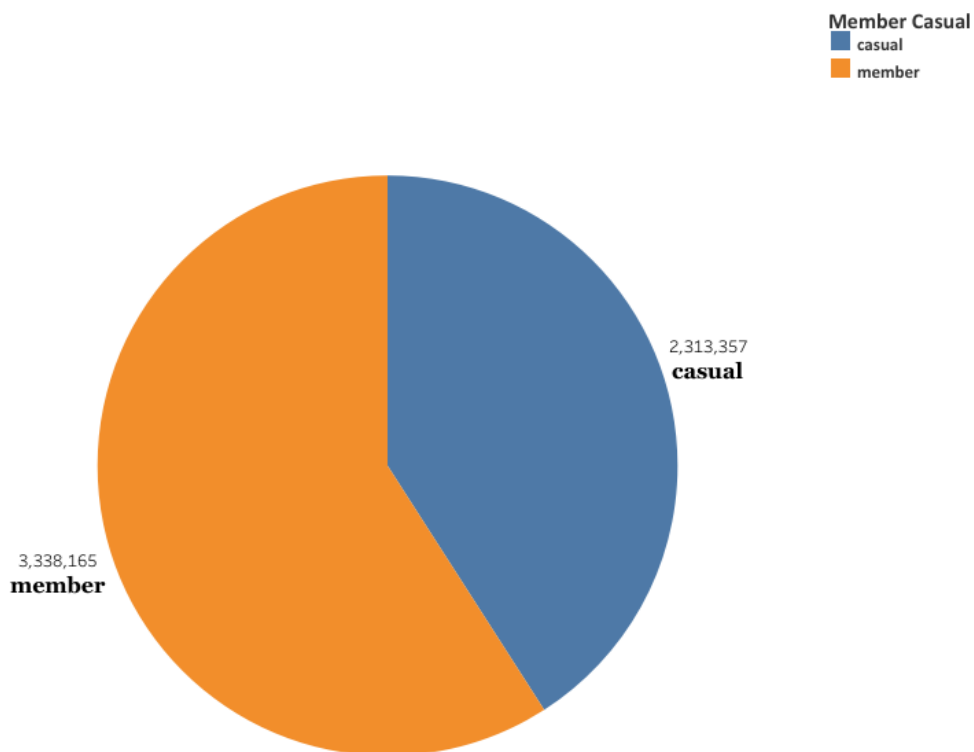
Total Number of Rides

In this data frame, we have a total of **5,651,522** rides. It took place between January 2022 and December 2022.

Total Rides by User Type

Of the total **5,651,522** rides, **2,313,357** were made by casual members and **3,338,165** by annual members.

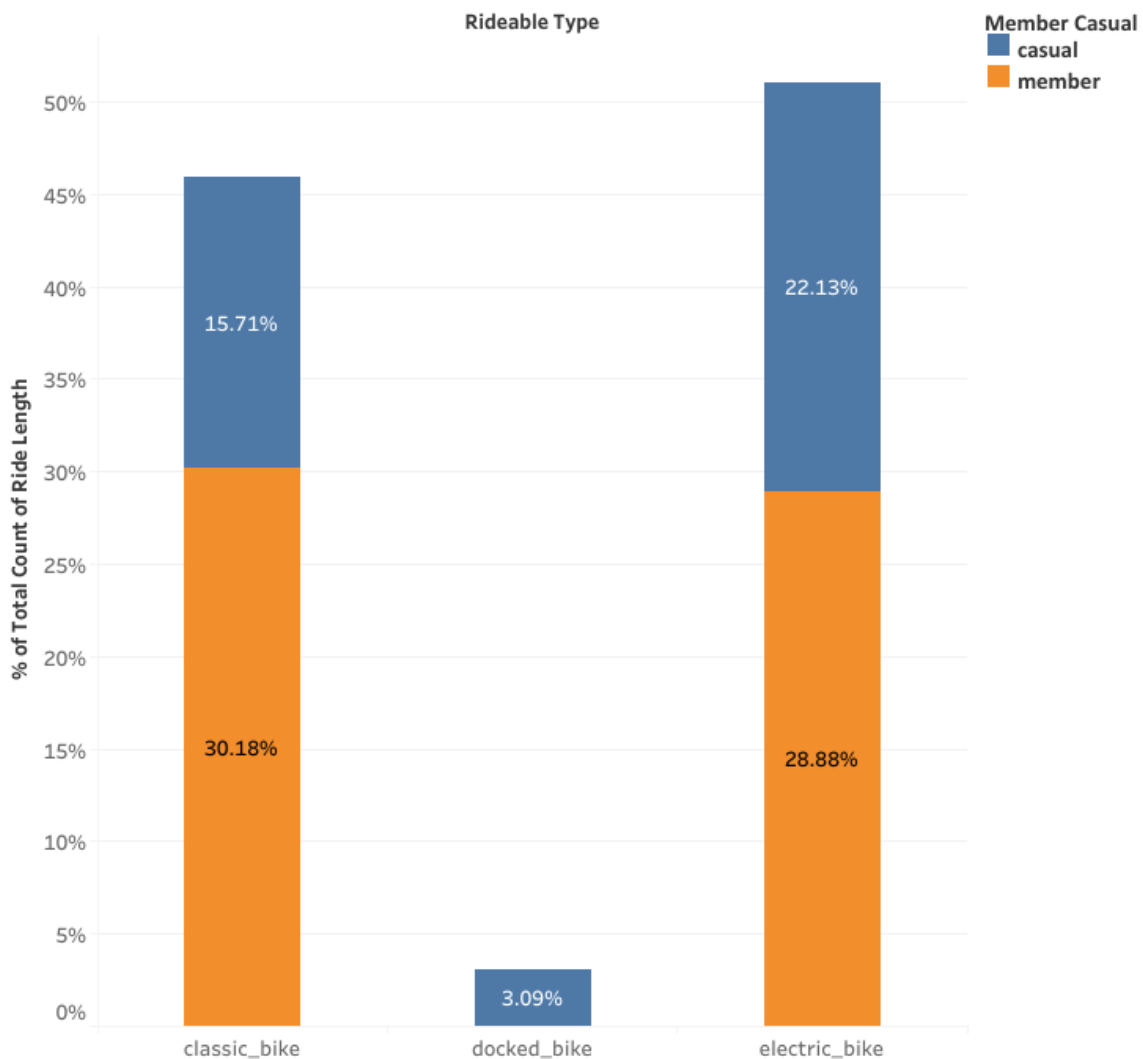
Total Rides By User Type



Total Rides by per Bike Type

When we group our data according to bike types; It was observed that **casual** members preferred the most electric bike type with **1,250,959** rides, while annual **members** preferred the classic bike type the most with **1,705,809** rides.

Total Rides by per Bike Type

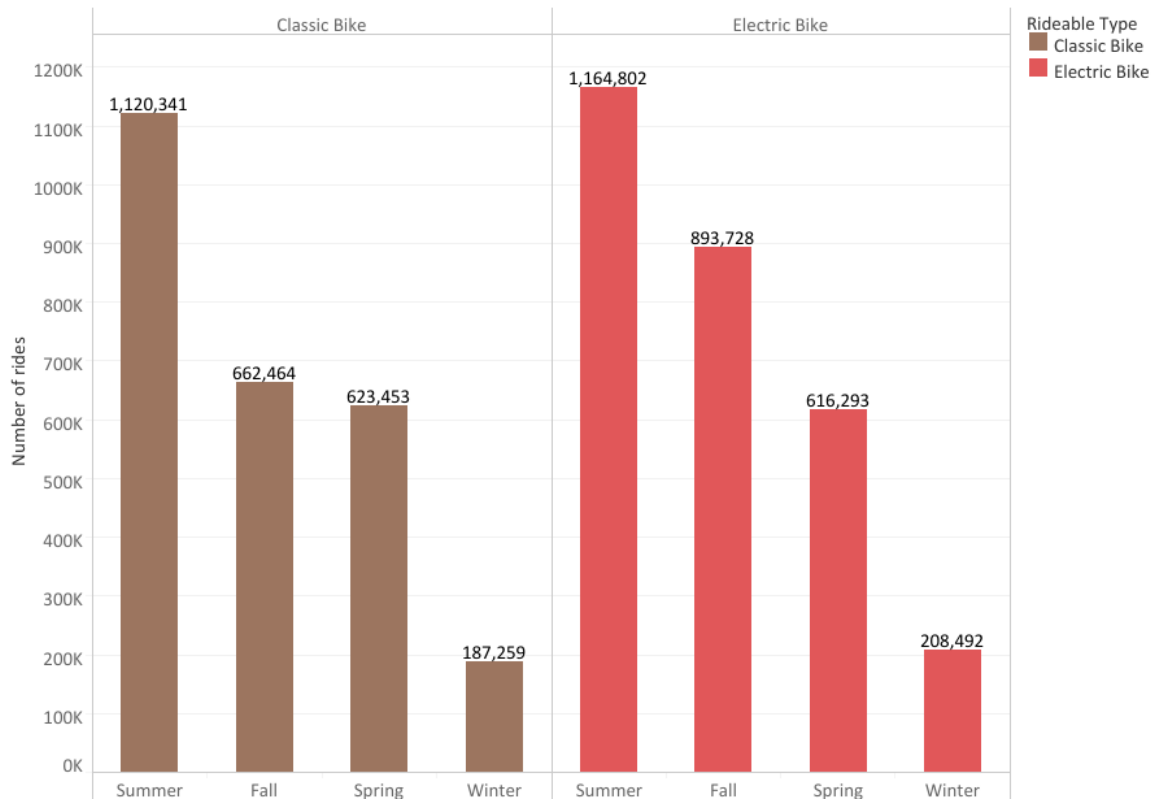


When these data are examined, electric bicycles are preferred by users (both **casual** and annual **members**) at a rate of **51%** of the total ride. It was seen that the second most used type of bicycle after the electric bicycle is the classical bicycle at a rate of **45%**. It has been clearly seen that the docked bike type is used only by casual members.

Usage Of Bicycle Types According to Seasons

When the use of bicycle types according to the seasons is examined, it can be said that the electric bicycle is used more often than the classical bicycle. In the summer, the difference is very small. In both types of bicycles, the least usage occurred in the winter season.

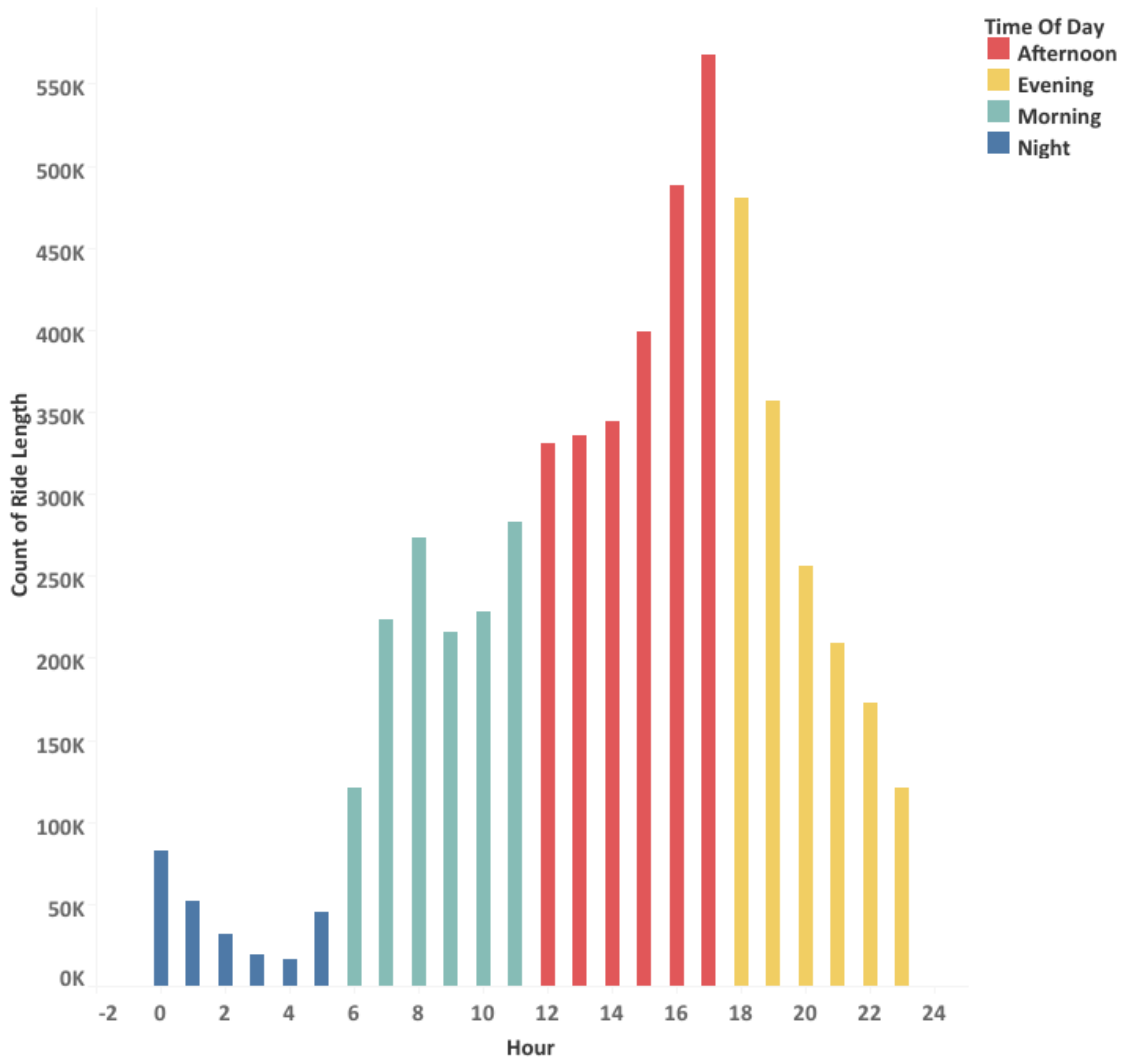
Seasonal use of bicycle types



Riding by the Hours

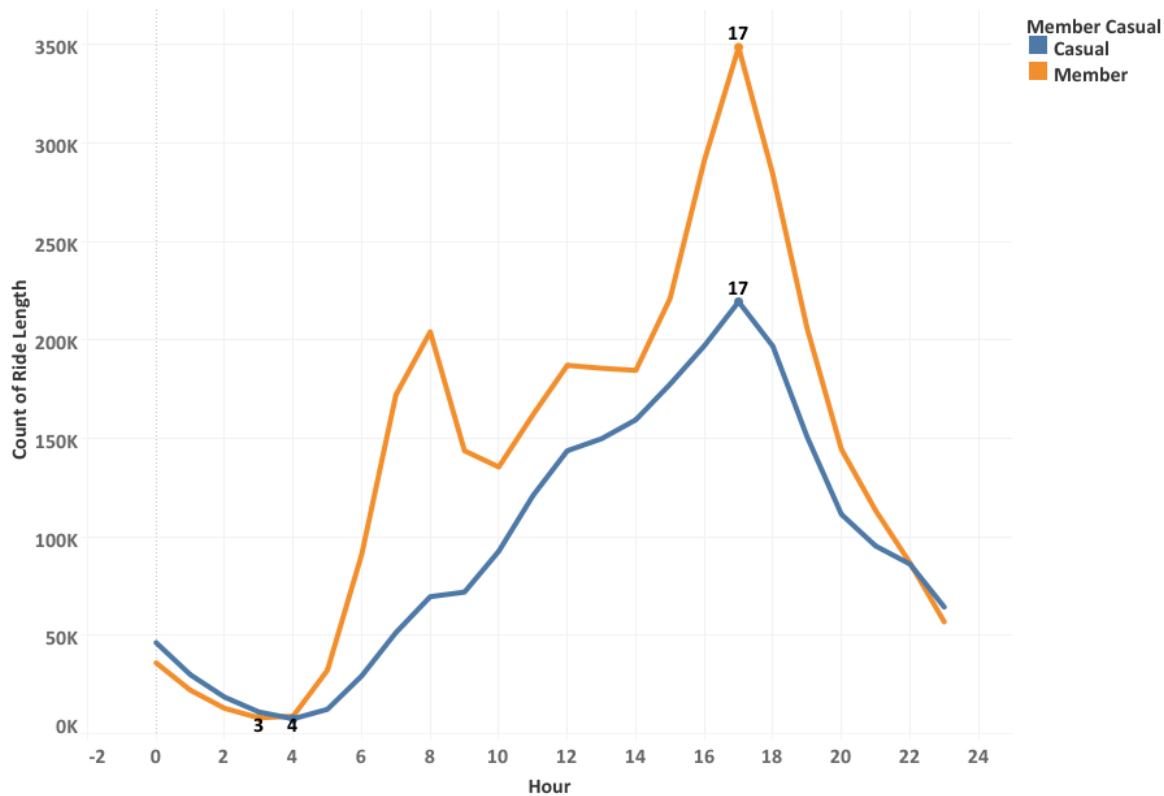
When we examine our driving data according to the hours of the day, it is seen that there is a serious increase in the number of driving in the afternoon. The number of driving in the afternoon and evening hours was higher than the other time periods.

Riding by the Hours



When the number of driving hours of the day according to the user type is examined, it is clearly seen that the annual and casual members drive more especially in the afternoon and evening hours.

Riding hours by the Member Type

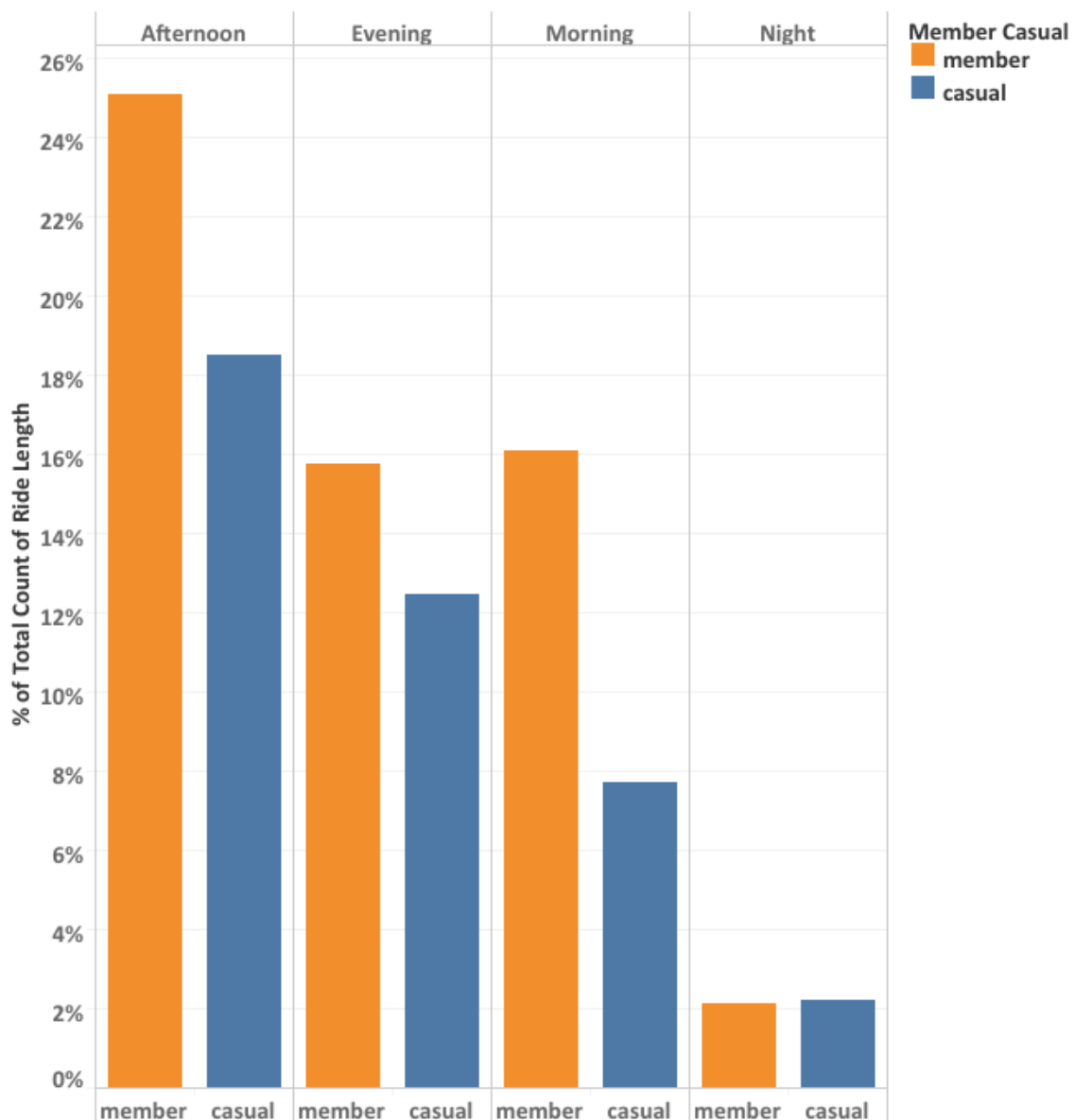


Riding by Day

The driving behaviors of the users in certain time periods of the day were examined. According to these reviews, it has been observed that annual and casual members have a higher driving percentage in the Afternoon compared to other time periods. While the evening time period has the second highest percentage of driving, the least driving occurred at night.

It is noteworthy that casual members have a higher percentage of driving at night than annual members.

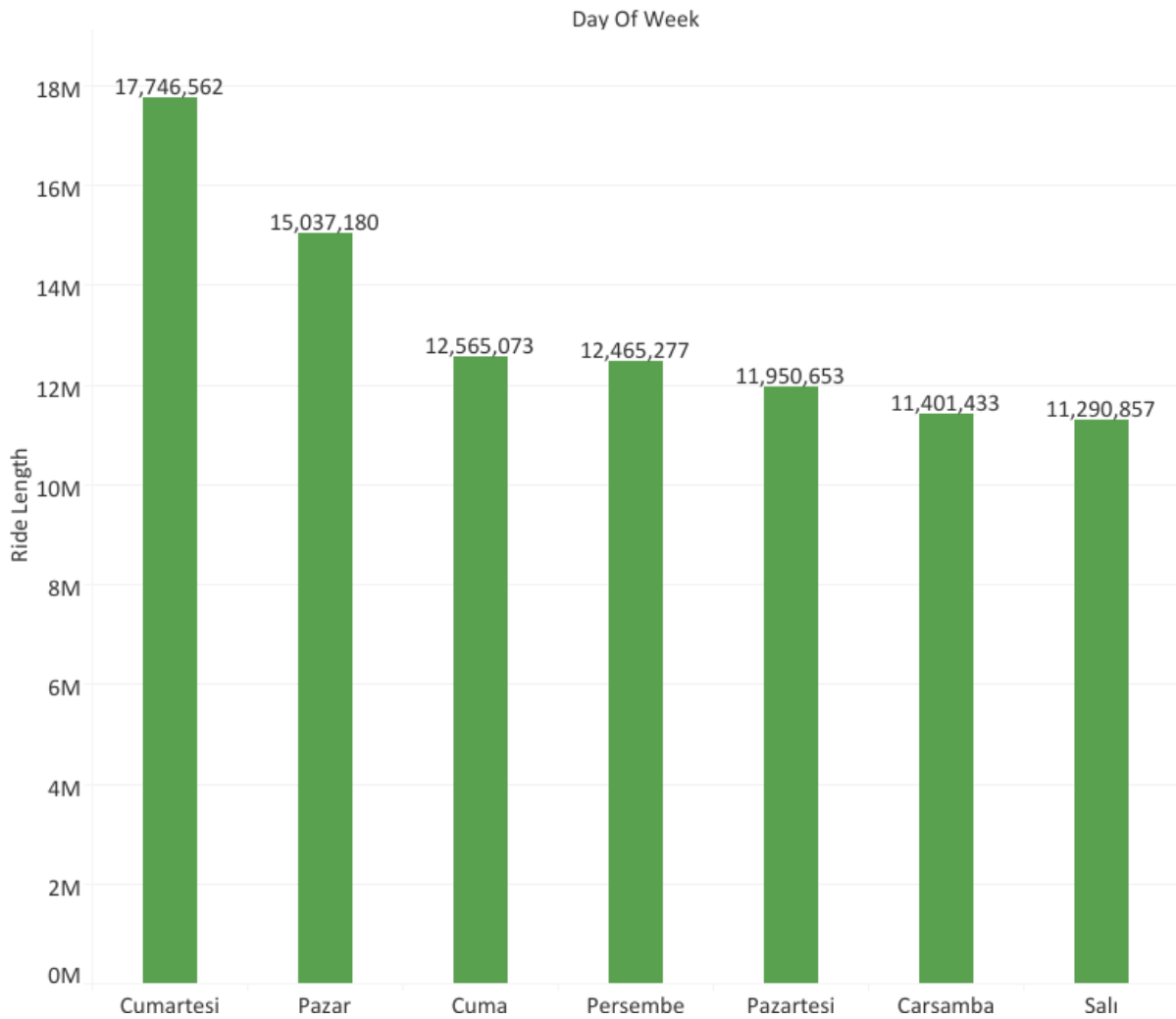
Riding by Day



Riding by the Weekday

When we examine according to the days of the week, it has been observed that most weekend driving takes place. It is assumed that the most important reason for this is that the members prefer to tour by bike as a weekend activity, since they have weekend holidays.

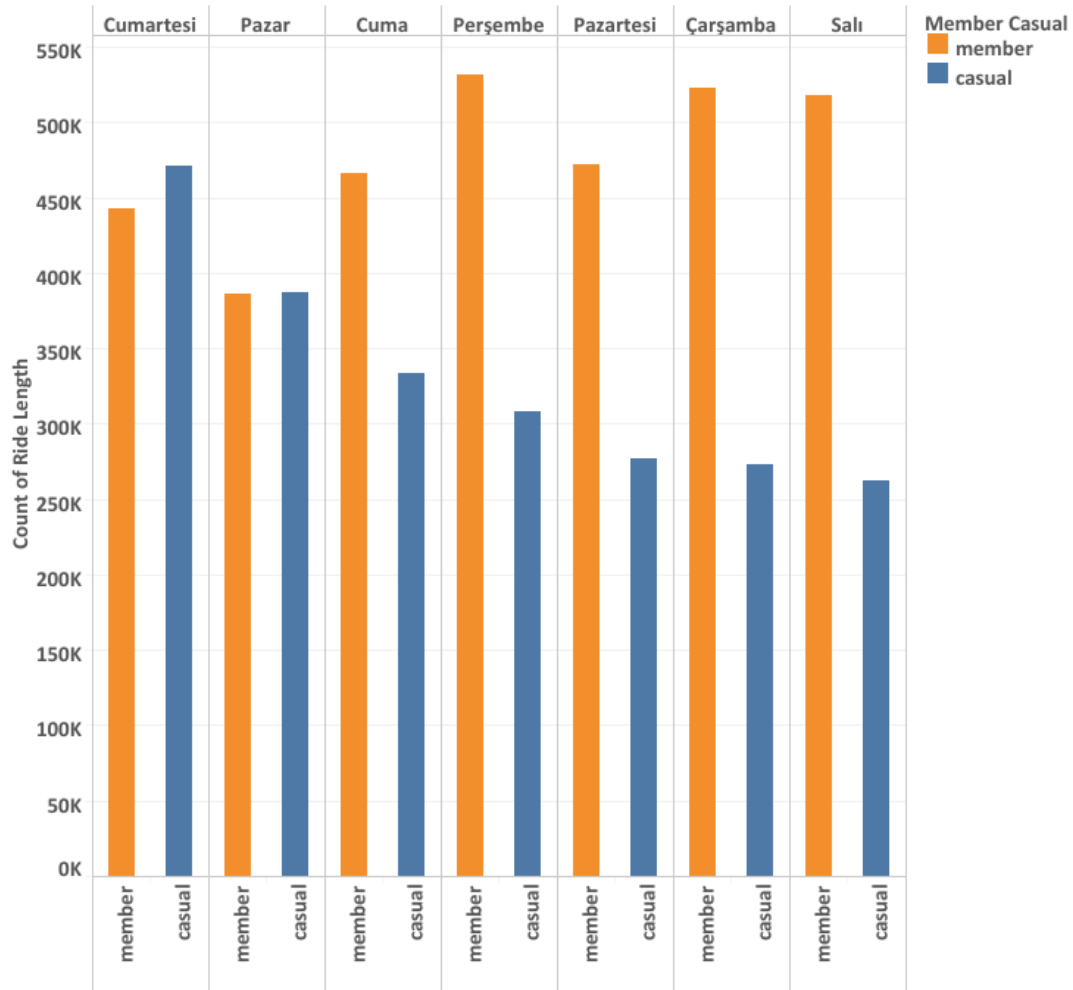
Total Rides by Weekday



Riding weekday by the Member Type

If we divide this analysis according to member types, it is observed that casual members reach the highest number of rides on Saturdays, while annual members reach the highest number of rides on Thursdays. Annual members made more rides on all other days in total number of rides except Saturday.

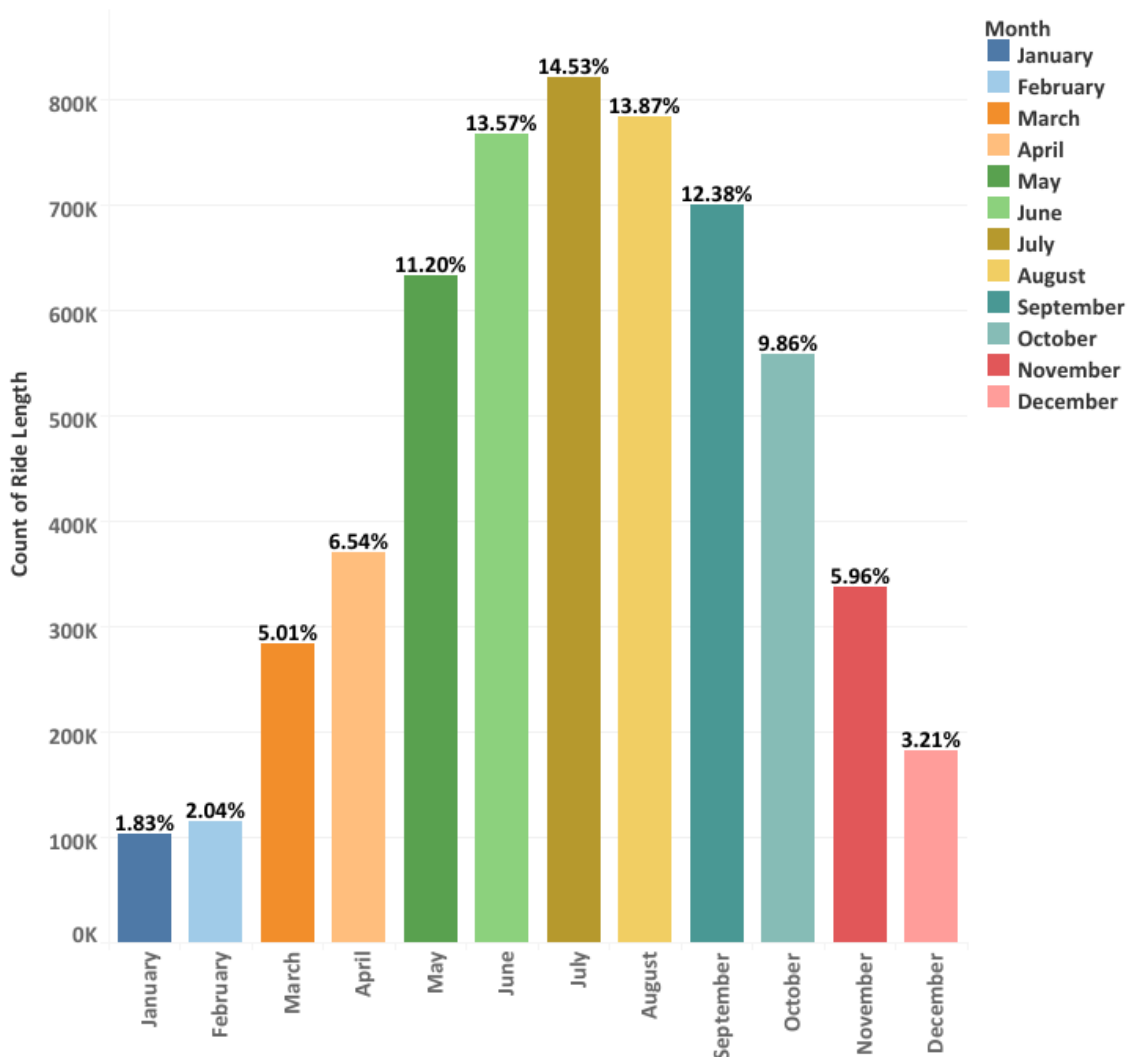
Riding weekday by the Member Type



Rides by the Month

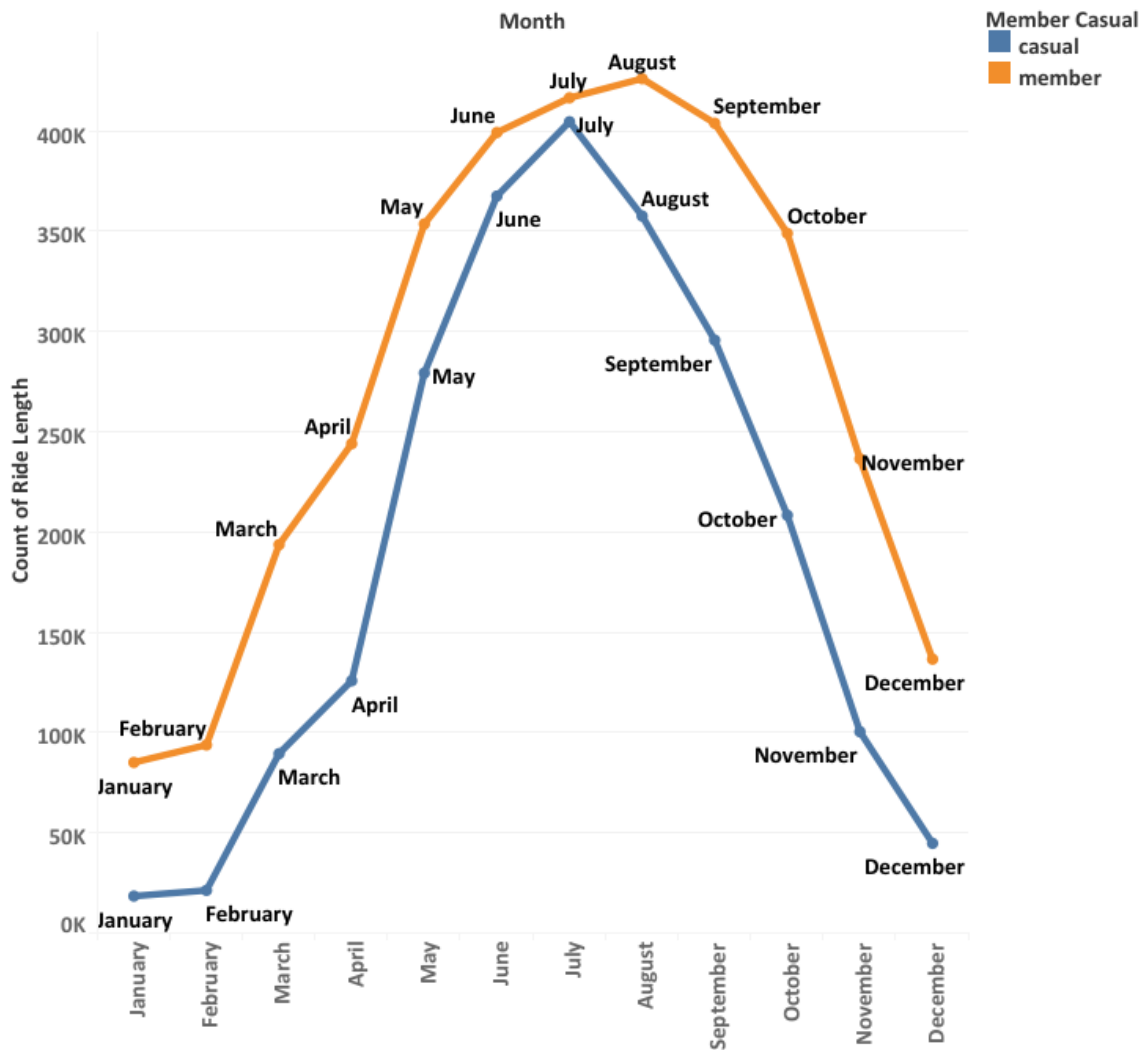
When we group our total number of rides by months, we observe that the highest number of rides occurs in July with 14.53 percent. Again, with 13.87% and 13.57 % of our total driving numbers, August and June were observed as the months with the most driving. It was clearly seen that December, January and February had the lowest driving percentages.

Total Rides by the Month



If we group these data according to the member types;

Total Rides Month by the Member Type



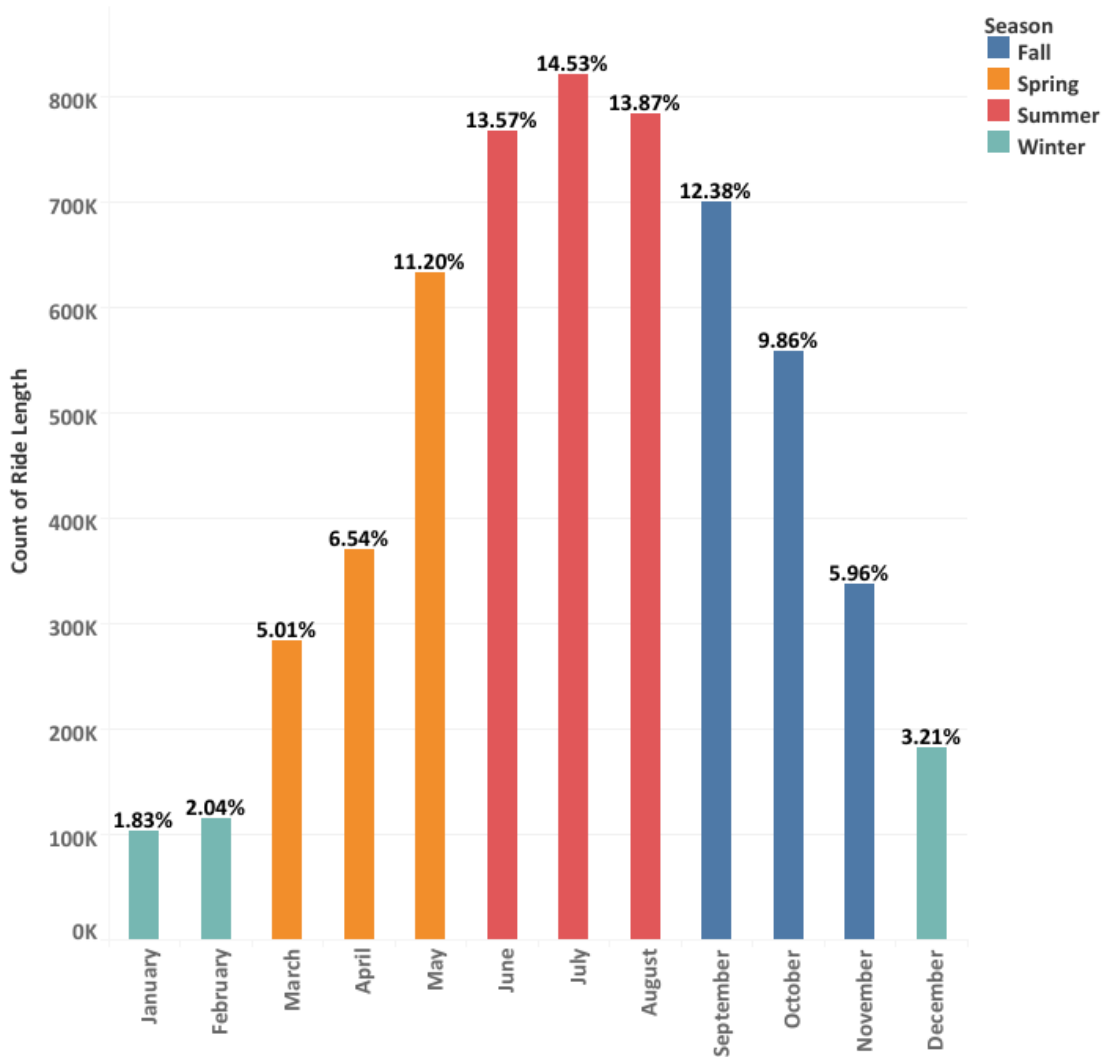
As can be seen from the line graph above, the behavior of the membership types in the total number of rides according to the months is quite similar to each other. In both membership types, while more rides are made in the summer months, this number drops considerably in the winter months.

After this analysis, it would be very accurate to analyze these driving types according to the seasons.

Rides by The Seasons

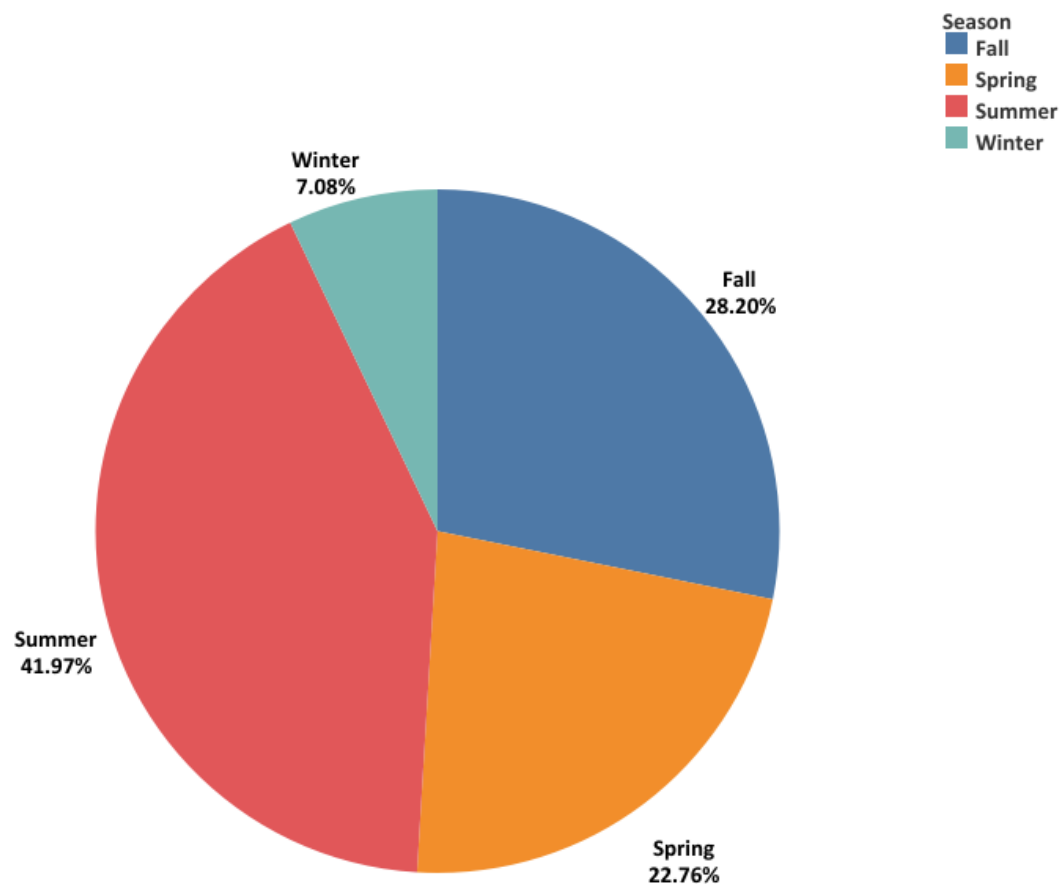
The driving data according to the seasons, which is a big picture of the analysis graph according to the months given above, is examined in this section.

Total Rides by the Season



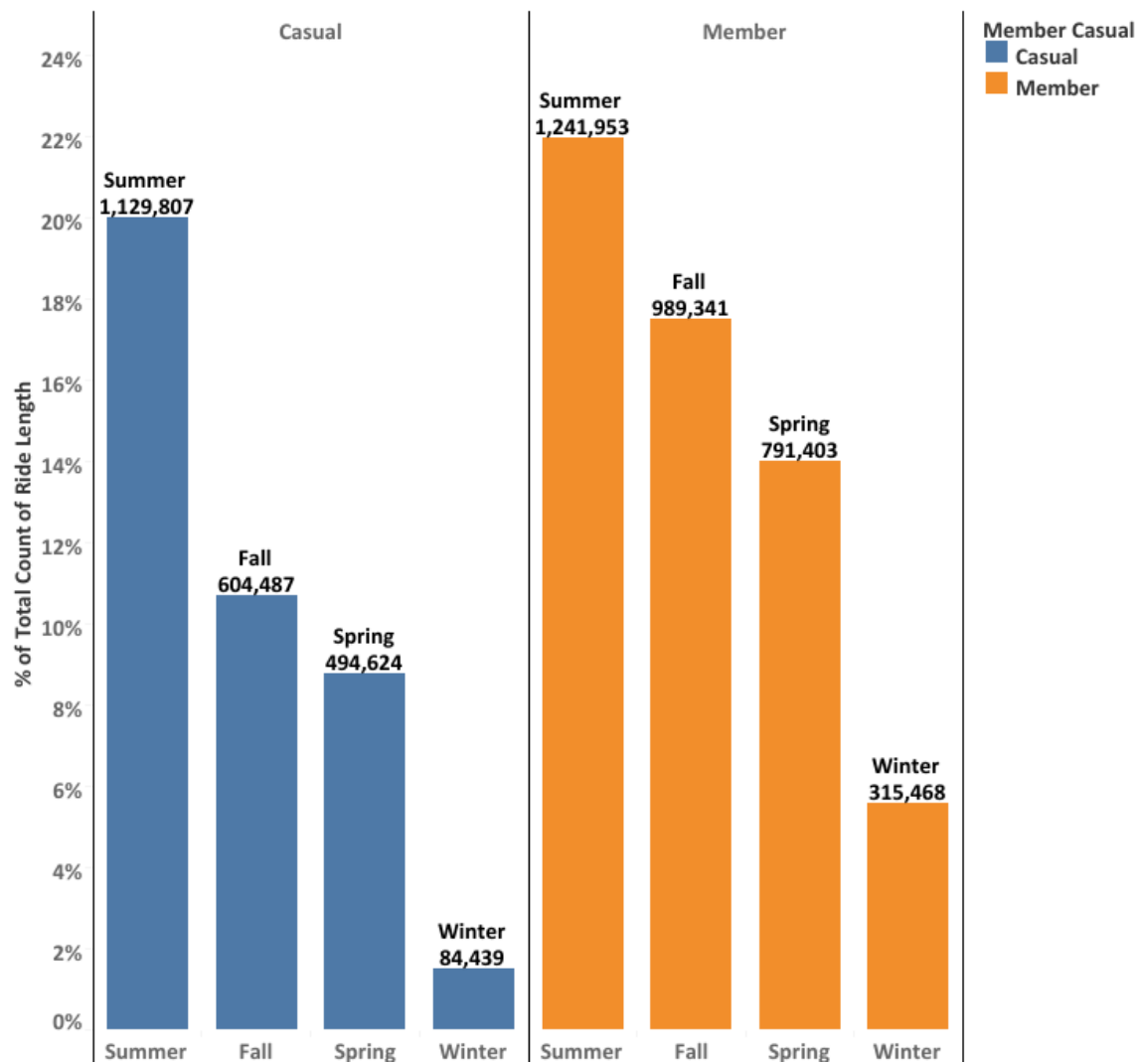
When the graph given above is examined, it is observed that while the most driving is in the summer season, a close number of driving is observed between the fall season and the spring season. The significant decrease in the number of rides in the winter season has statistically revealed that the drivers do not prefer bicycles when the weather and road conditions deteriorate.

Display of total driving percentages by season



Looking at the pie chart given above, summer months stand out as the season with the highest number of rides with 41.97% in total driving numbers, while fall ranks second with 28.20%. While the spring season is the third season with the most driving, as we just mentioned, the winter season is the season with the least driving.

Comparison of total number of rides by seasons of user types



The graphic above shows the comparison of the total number of seasonal rides according to the member types. As we mentioned above, summer and spring seasons are at their peak in both member types.

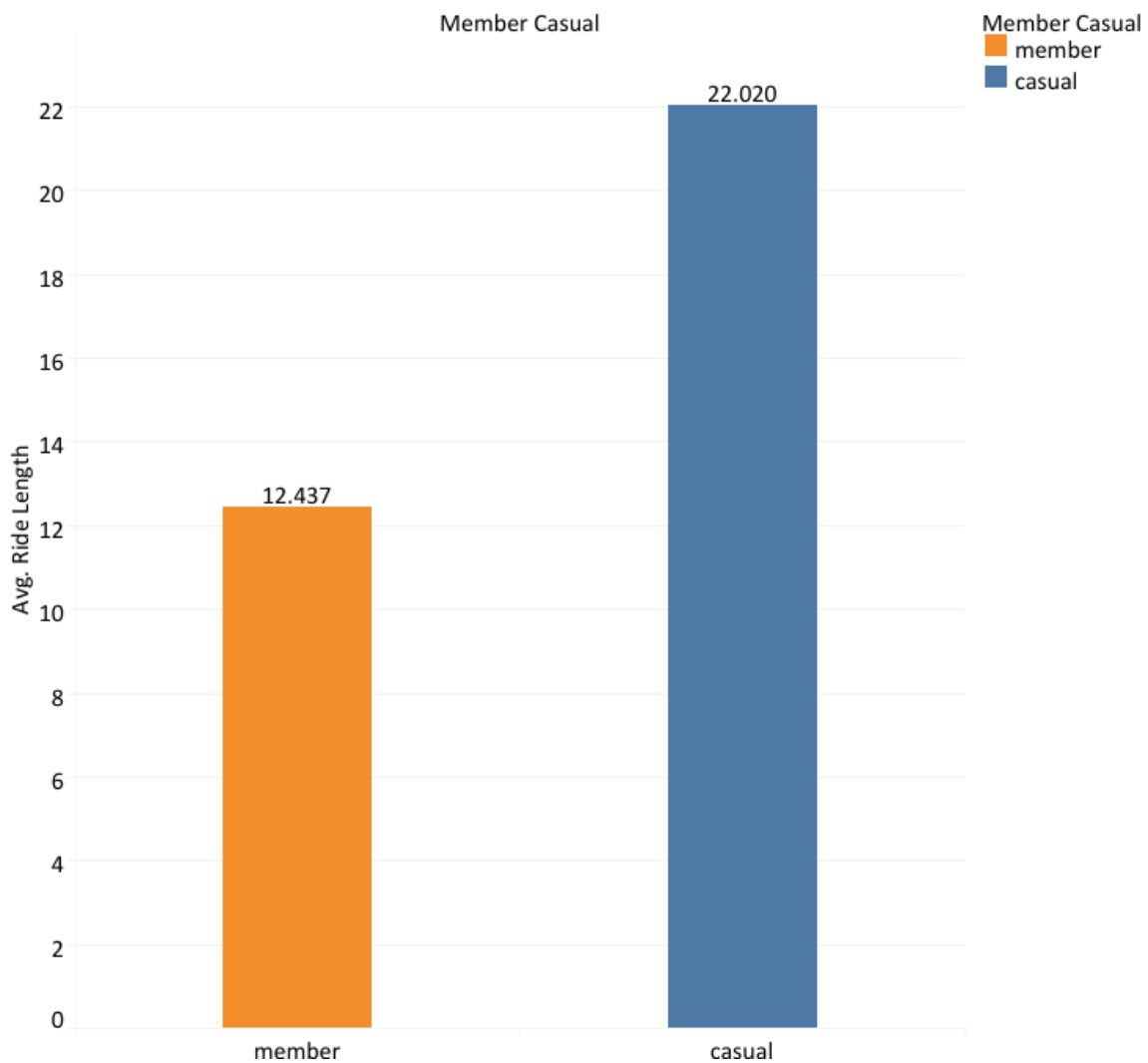
Analysis of Average Riding Times

Average Ride Length

The average ride length is defined as the average of the numbers (ride length) obtained by subtracting the start time from the ride finish time. With the analysis made, the average driving length was measured as **16.36** minutes.

When the user types are examined, it is seen that casual members have an average driving length of 22 minutes, and annual members have an average driving length of 12.4 minutes.

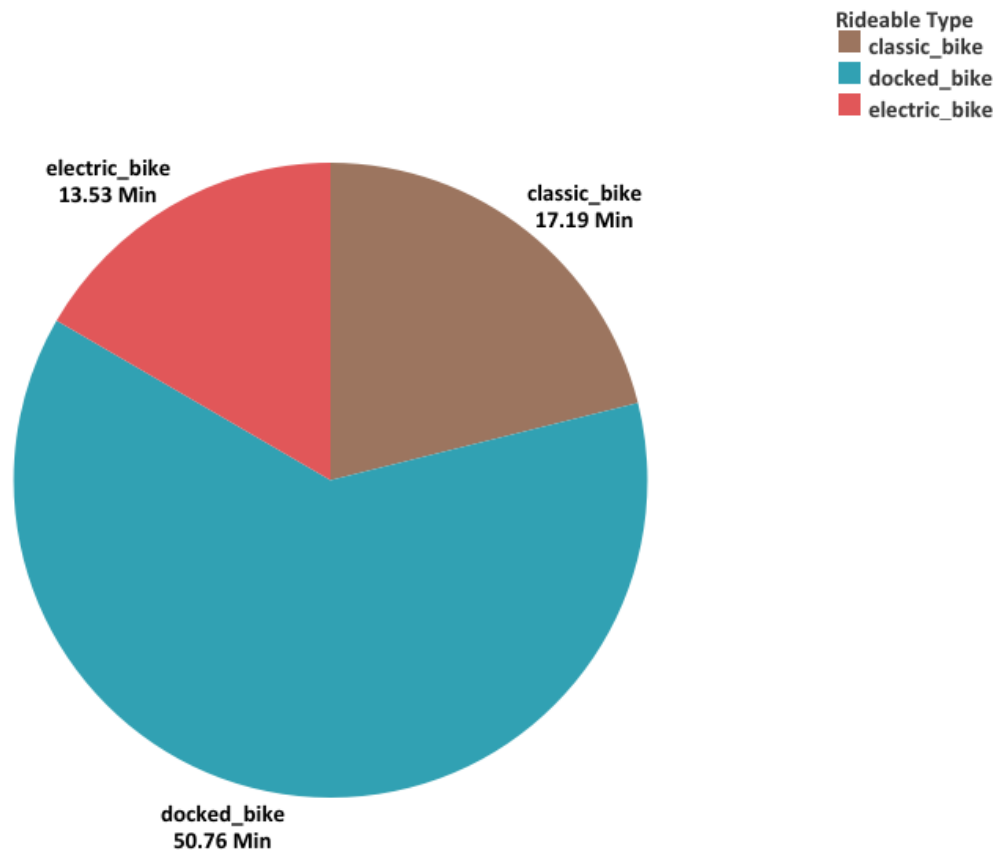
Average Ride Length by User Type



Average ride length by the Bike Type

When we compare the average riding times according to the bicycle type, it is seen that the docked bike type is seen as the bicycle type with the highest average riding time with 50.76 minutes, while the classical bicycle type has an average of 17.19 minutes and the electric bicycle has an average of 13.53 minutes.

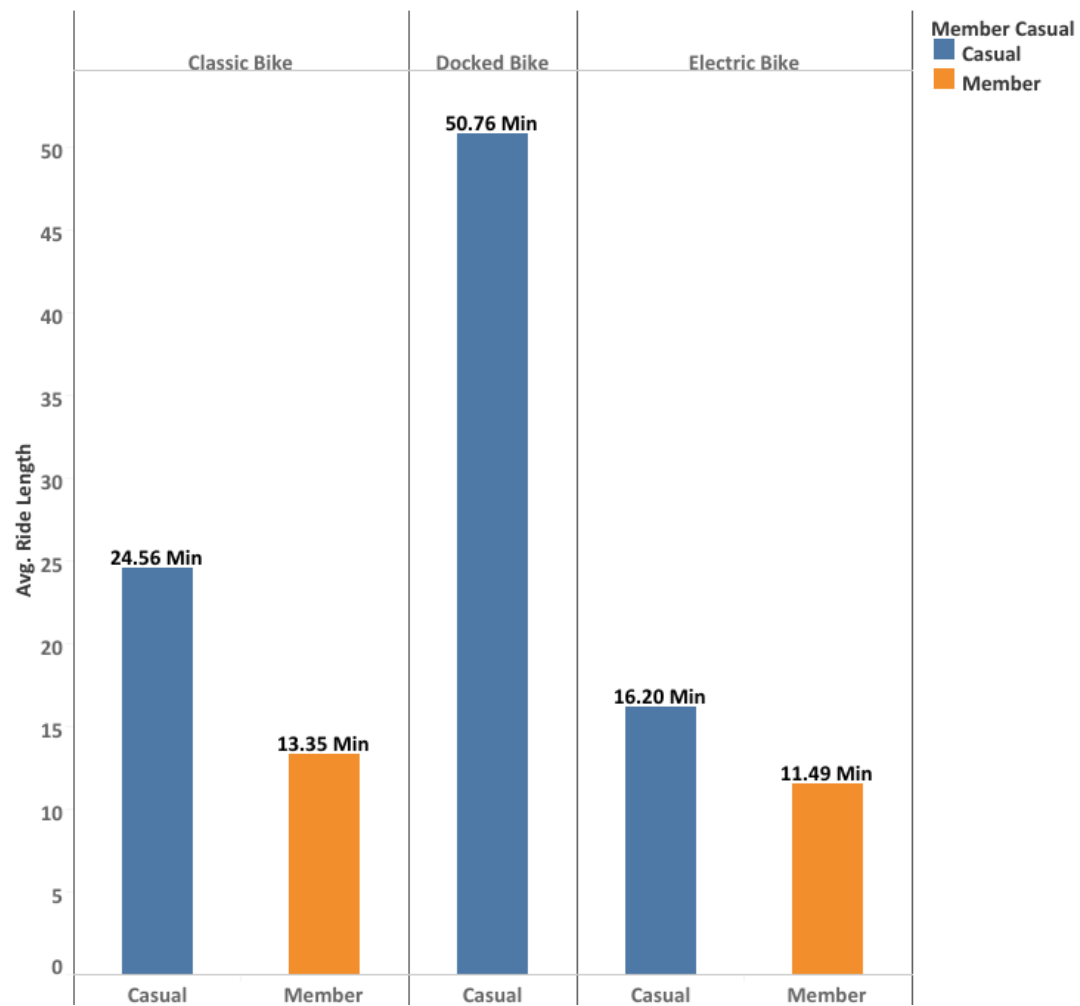
Average ride length by the Bike Type



Average ride length bike type by the Member Type

When we compare the average riding times according to user type and bike type, it has been observed that Casual members use docked bikes the most. With an average of 50.76 minutes, all docked bike rides belong to casual members. In the classic bicycle type, casual members have an average of 24.56 minutes of riding time and more usage than annual members with 13.35 minutes of riding time. In the electric bicycle type, casual members have an average driving time of 16.20 minutes, while annual members have an average of 11.49 minutes.

Average ride length bike type by the Member casual

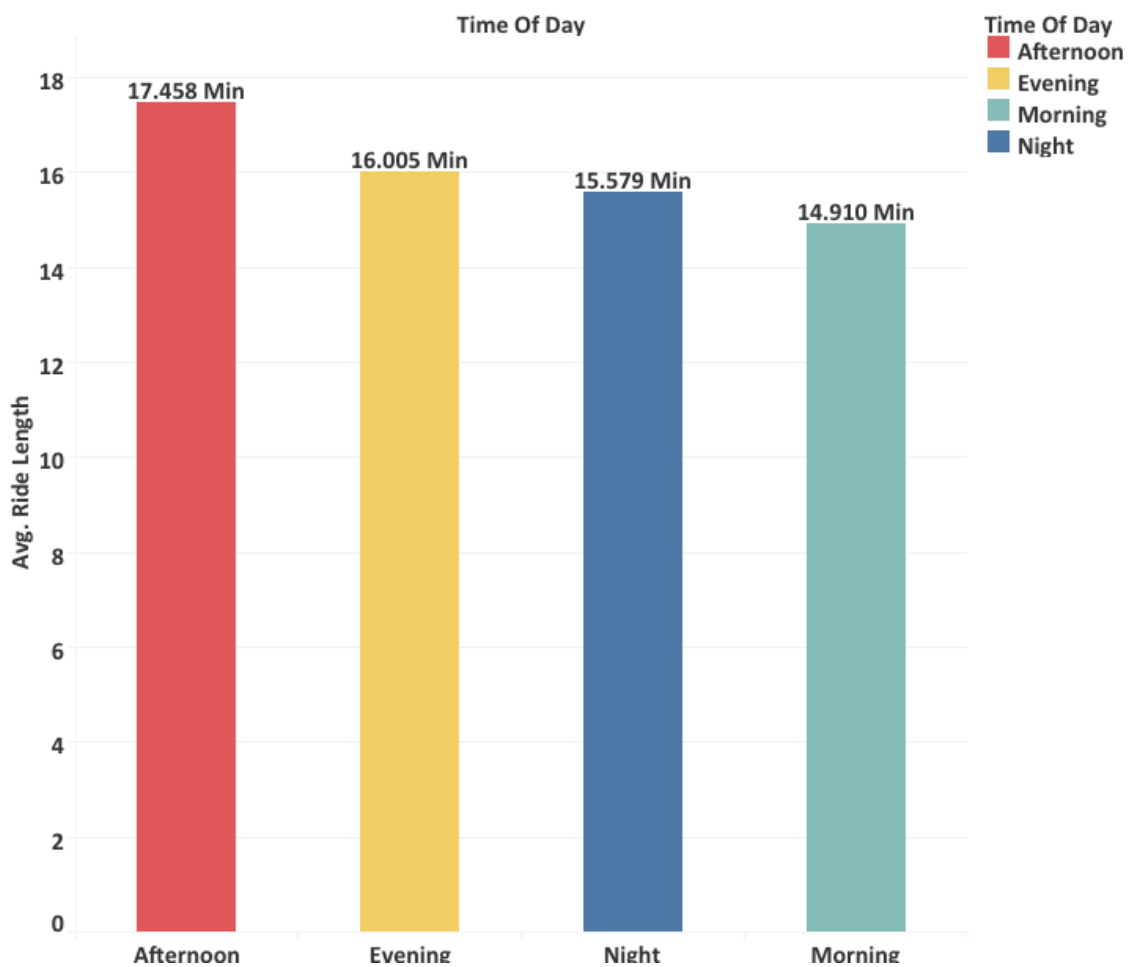


When the average driving times are examined according to the time zones of the day, there is an average driving time of 17,458 minutes for the afternoon. While there is an average driving time of 16,005 minutes in the evening, the time zone with the least average driving time is morning.

Average riding times according to time zones of the day

When the average driving times are analyzed according to user types and time zones of the day, it is observed that casual members have 23.76 minutes of average driving time in the afternoon than annual members with an average of 12.81 minutes. Casual members have the most average driving time in the morning after the afternoon, while annual members have the most average driving time in the evening and night hours.

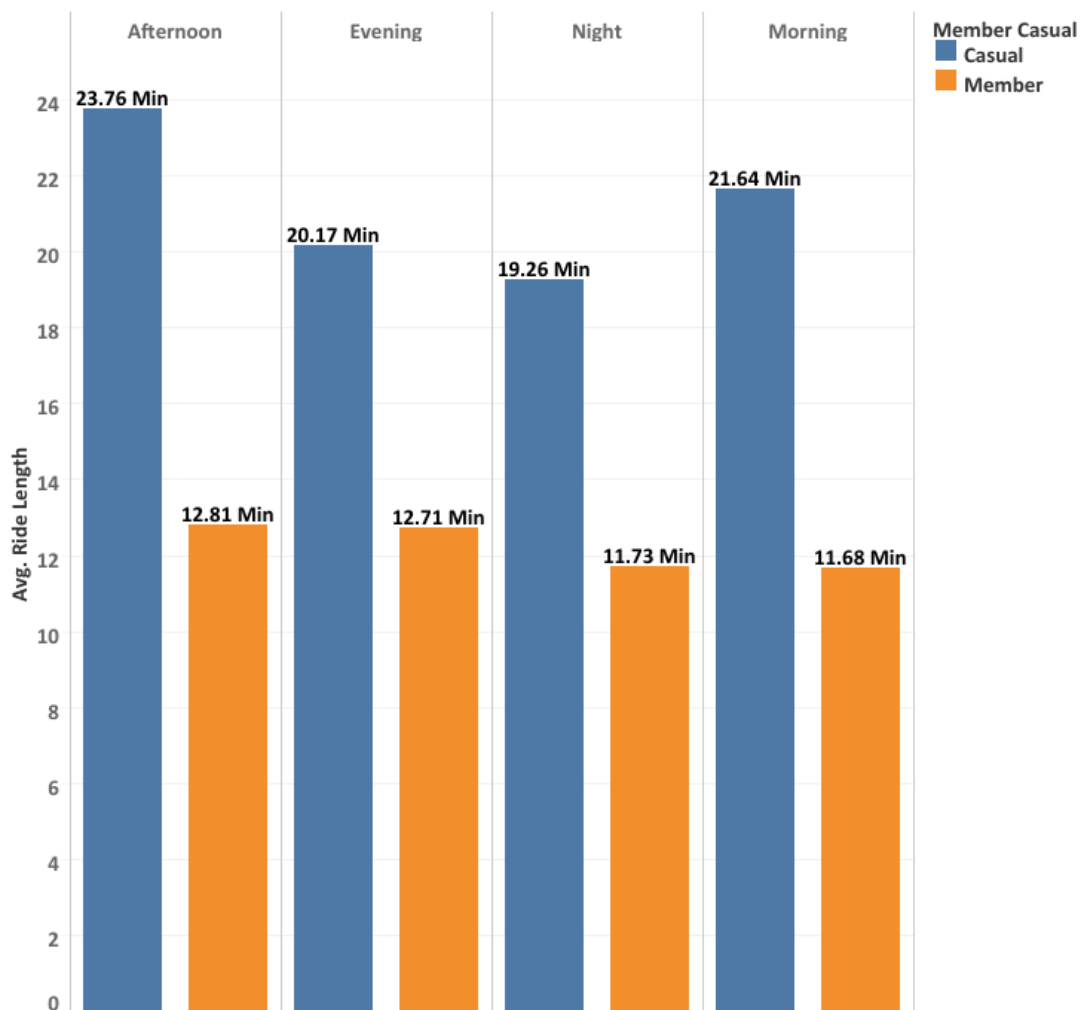
Average driving times according to time zones of the day



Average riding times by user types and time zones of the day

When the average driving times of the member types according to the days are examined, casual members and annual members have the most average driving times on weekends. While the average driving time is generally higher in casual members, it has been observed that annual members have an average of 13 to 11 minutes of driving time.

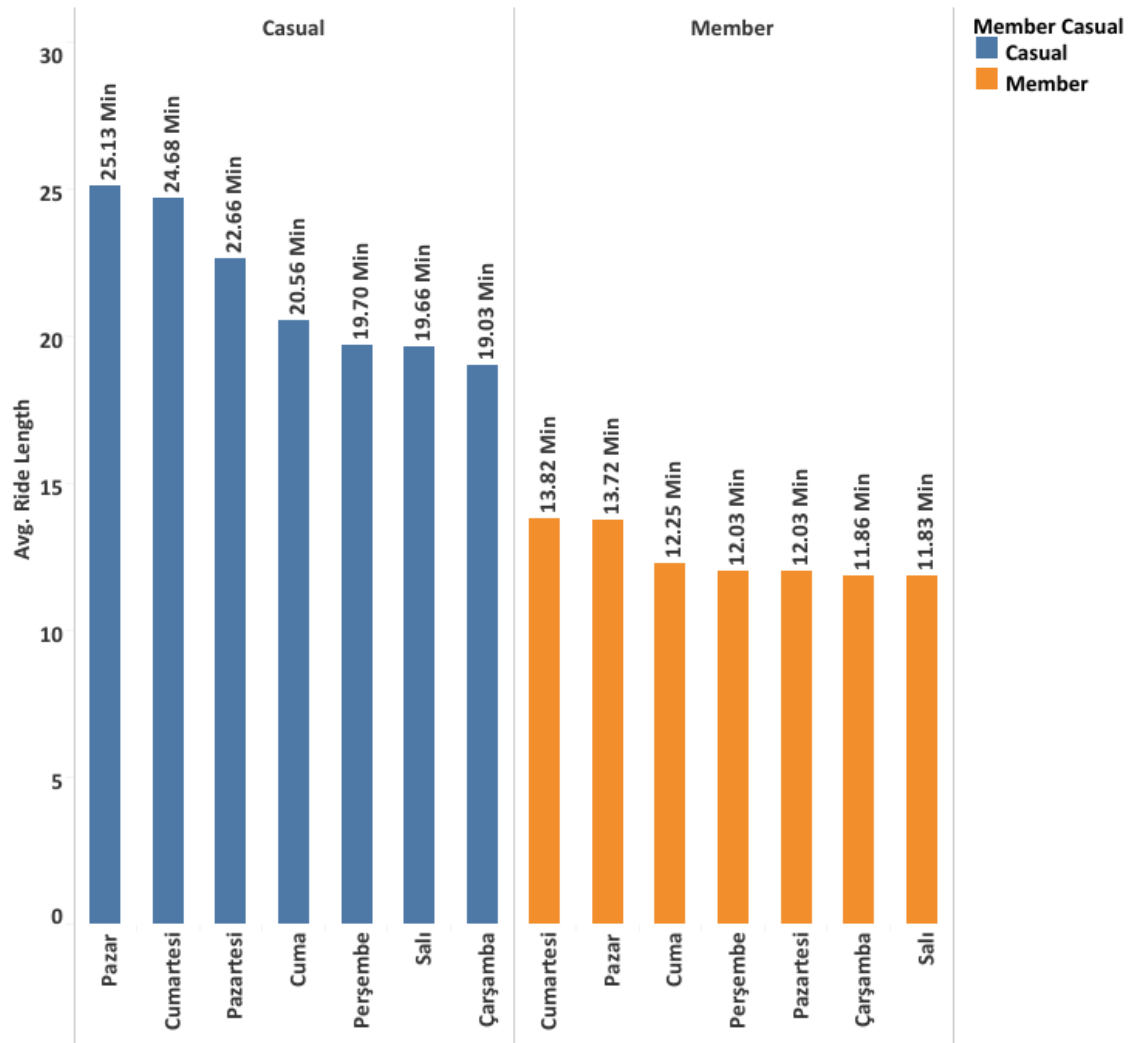
Display of average driving times by user types and time zones of the day



Average riding times of member types by days

When the average driving times are analyzed according to the days, Sunday and Saturday are at the top. With an average of 15,956 minutes on Monday, the first working day, it is the day with the highest average driving time after the weekend.

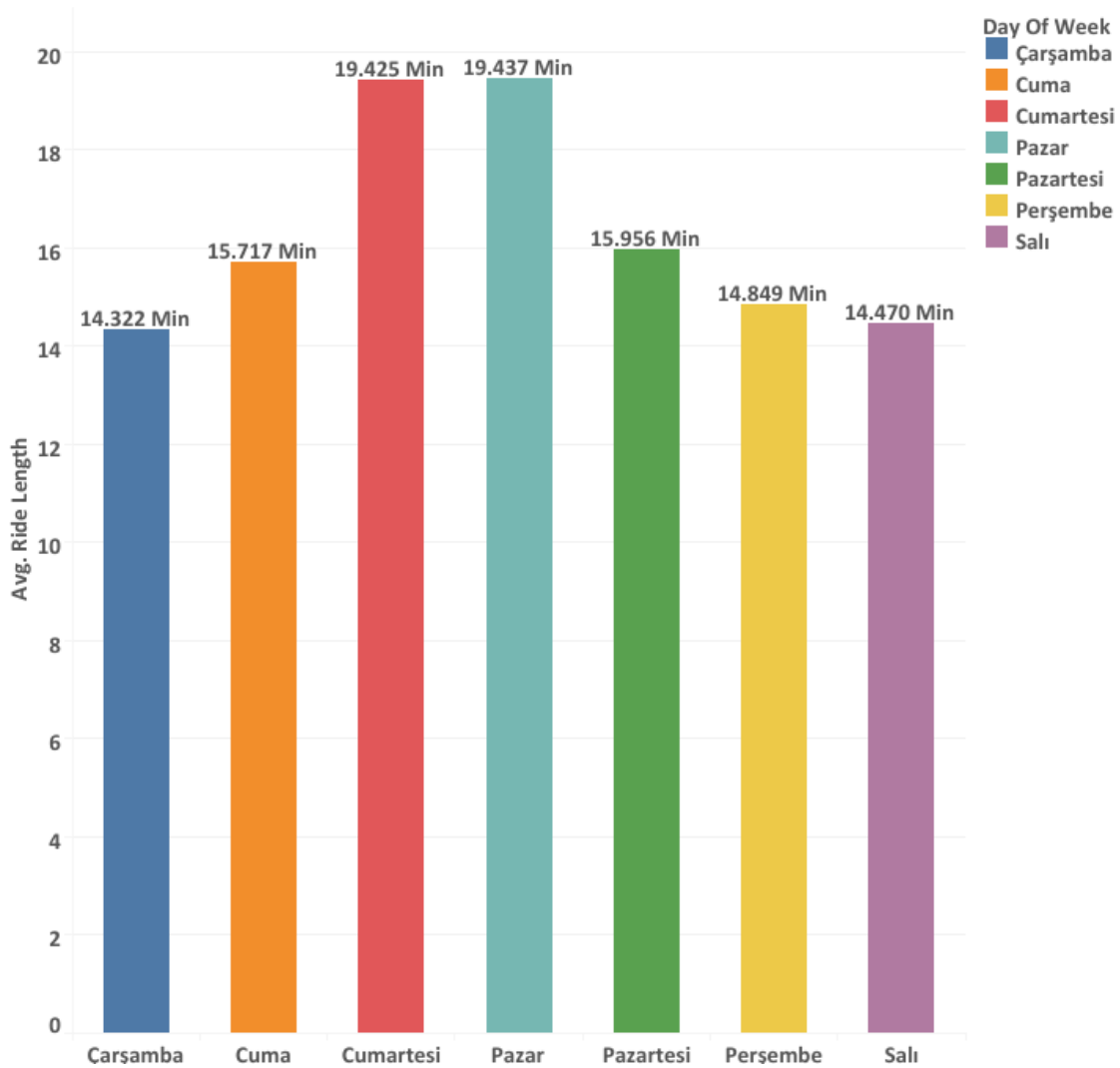
Average riding times of member types by days



Average riding times by Day

When the average driving times are analyzed according to the days, Sunday and Saturday are at the top. With an average of 15,956 minutes on Monday, the first working day, it is the day with the highest average driving time after the weekend.

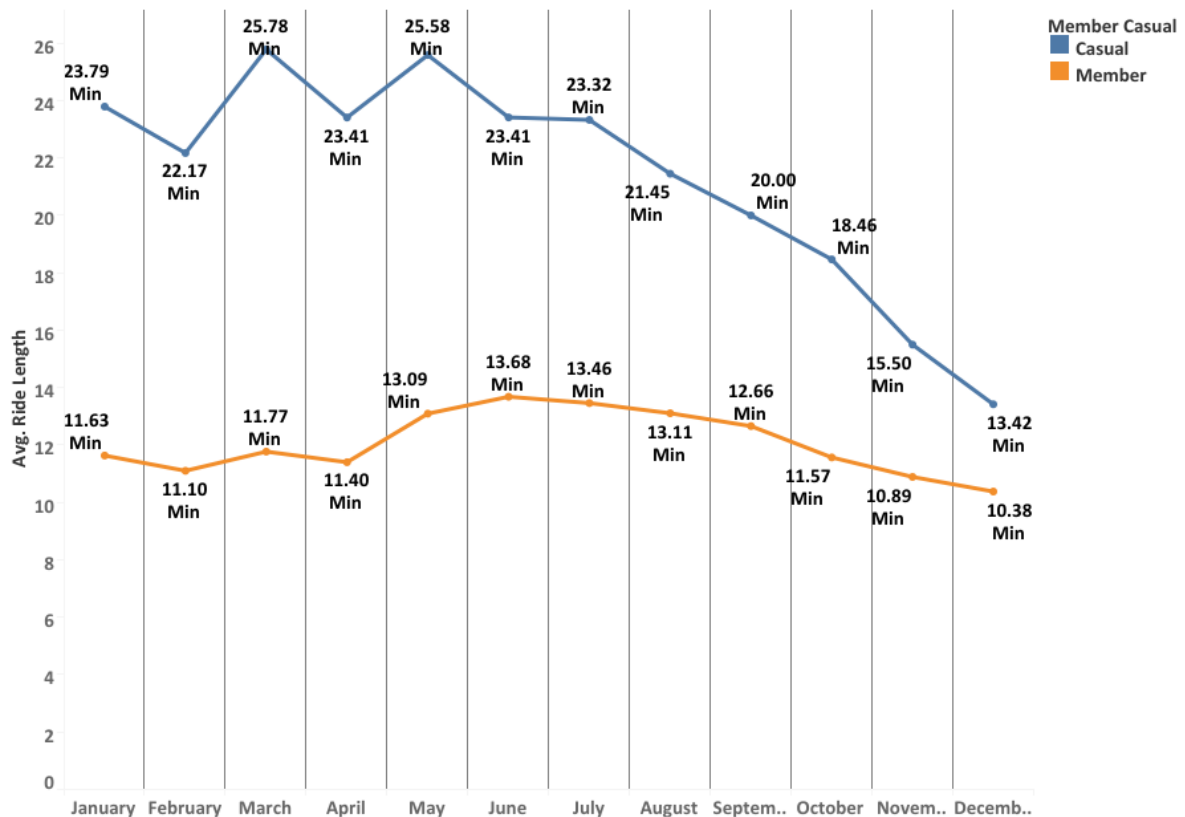
Average riding times by day



Average driving times of user types by Month

When the average driving times of the user types according to the months are examined, it has been observed that the months with the most average driving time for casual members are March, April, May and June. With an average driving time of 25.78 minutes in March, casual members had the most average driving time in this month. Annual members, on the other hand, have the most average driving time in June with 13.68 minutes.

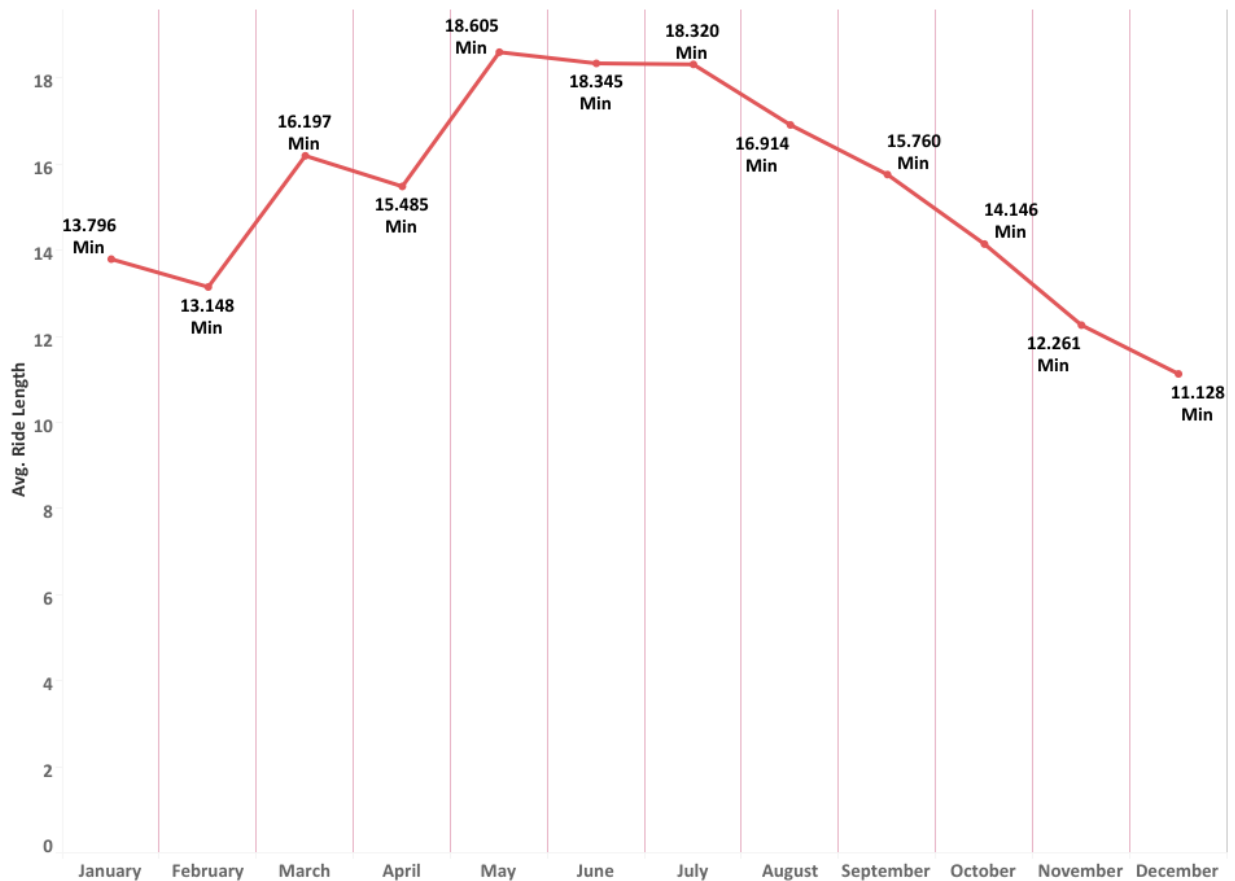
Average riding times of user types by month



Average riding times by Month

When the average driving times are analyzed by months, May, June and July are the months with the highest average driving time. While it was observed that December had the lowest average driving time, it was clearly determined that the average driving time decreased even more in winter.

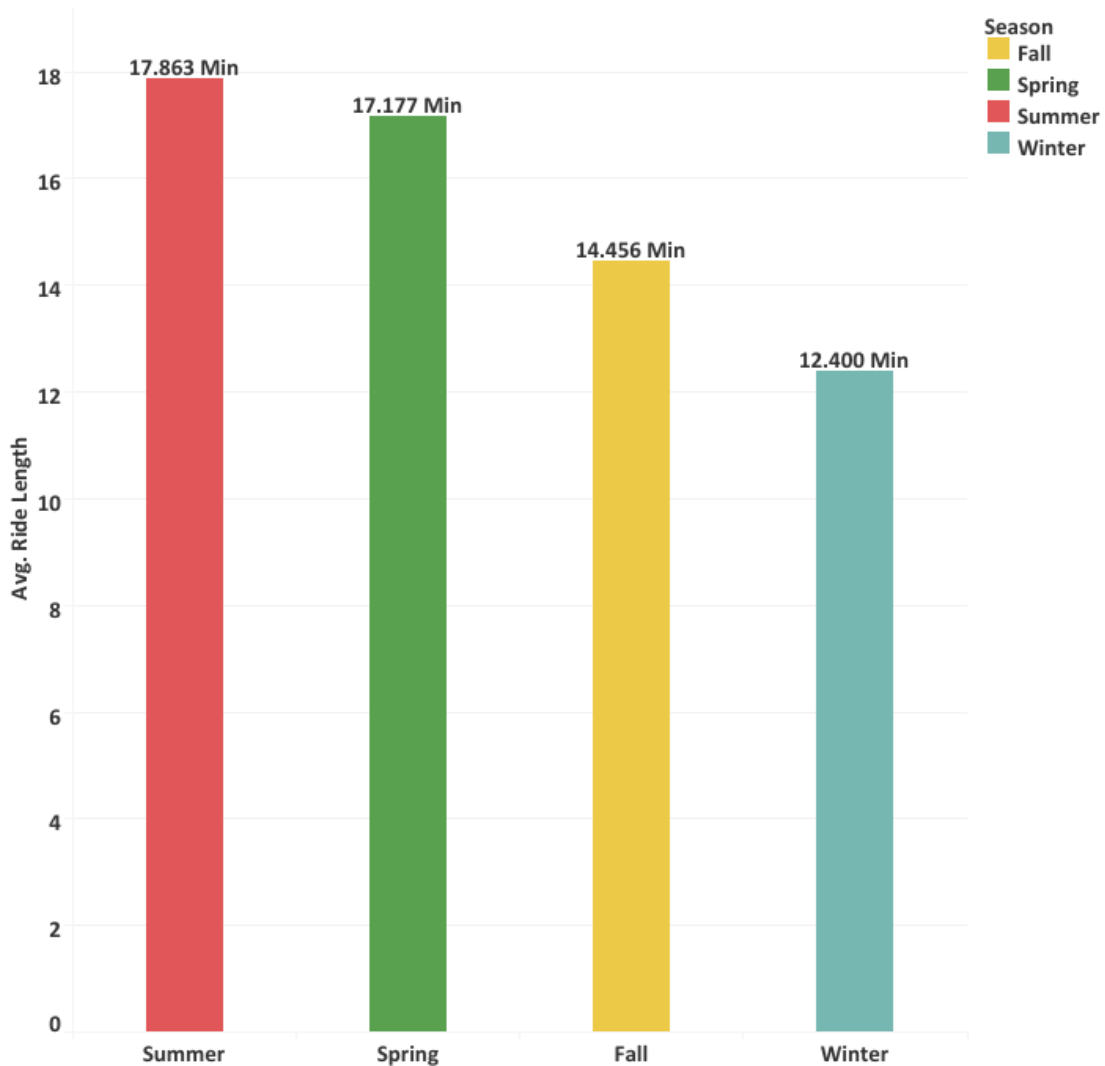
Average riding times by month



Average riding times by Season

When the average driving time according to the seasons is examined, the average driving time is 17,863 minutes at the most, while the summer season is followed by the spring months with an average driving time of 17,177 minutes. The lowest average driving time was 12.40 minutes in winter.

Average riding times by season

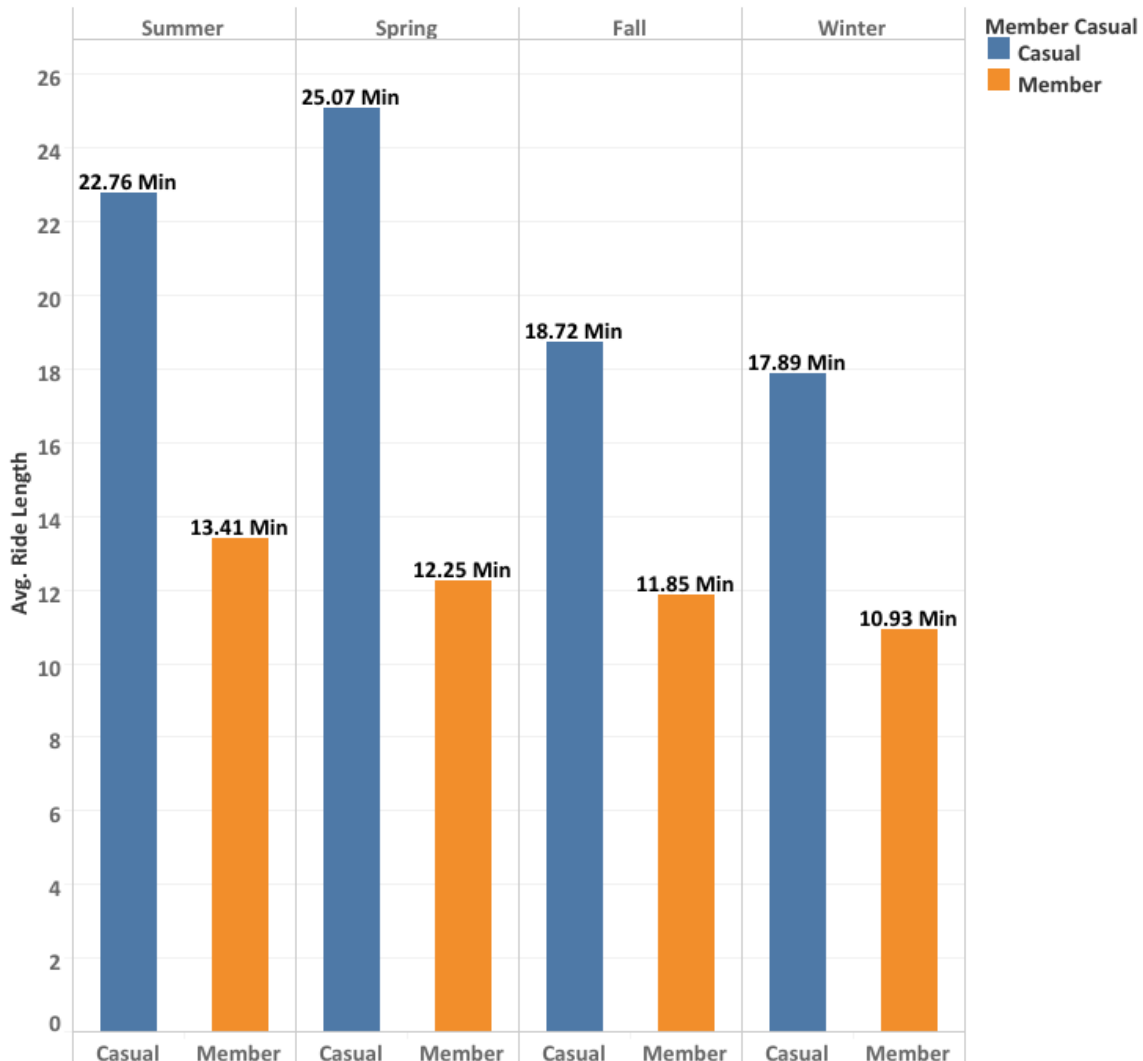


Average Riding Times Per Season By Member Type

When the average driving times in the seasons are examined according to the member types, it is observed that the casual members have the most average driving time in the spring season. Casual members have the highest average driving time in summer with 22.76 minutes

after the spring season, and the lowest average time in winter is 17.89 minutes. Annual members, on the other hand, had the most average driving time of 13.41 minutes in the summer season, while the average of 10.93 minutes in the winter season.

Average riding times per season by member type



Yes, graphics and analysis can be increased further, but we have completed the necessary investigations for us. Therefore, we can move on to the next section where we discuss how this data will affect our approach to current business problems.

As a result of all these analyzes and visualizations I have done, I combined the information I obtained in a presentation file and made it ready to share with my team. The presentation file can be accessed via [this link](#).

Phase 6 – ACT

Final Conclusion and Recommendation

My answers to the three questions that will guide the future marketing program are as follows, based on my analysis and visualizations:

Q1: How do annual members and casual riders use Cyclistic bikes differently?

A1:

- Our annual members use the Cyclistic bike sharing service frequently and consistently in terms of total ride and ride time. It has shown that users prefer bicycles to other types of vehicles on issues such as cost, environmental pollution and traffic to use for certain jobs.
- On the other hand, casual riders do not use the bike frequently and consistently. Therefore, it is understandable that they do not prefer to switch to the annual membership plan. Despite this, casual riders like to cycle inconsistently compared to annual members, with long rides on weekends that we think are for fun and relaxation.
- Annual members are very likely to opt for the annual membership plan, as their daily and weekly routines are set. The fact that users have school, work and other activities at certain time intervals seems to push them to use a certain means of transportation in a planned way. Due to this planned use, annual membership seems quite reasonable to meet the needs of such members.
- As a result, usage differences give us information about drivers' tendencies and behaviors. These differences vary according to the drivers' behaviors and perceptions of use.

Q2: Why would casual riders buy Cyclistic annual memberships?

A2: In the light of the analysis and visualizations I have made, the suggestions I can give to the marketing team on how casual riders can be converted into annual members are as follows:

- The advertising and marketing strategy to get casual riders to get their annual subscription plan should be around very specific issues. These issues are the cost of cycling, health, environmental health, traffic, etc. It should be noted that the benefits in these subjects are more than other tools.

- In the light of the analyzes I have made; it has been observed that the reason for the annual members to use the service stems from their instincts to stick to their planned and routines in their lives. For this reason, mobile applications can be developed to encourage casual riders to move on to a planned life, and advertisements can be made for users to create routines at certain times of the day and use Cyclistic bicycles as a means of travel to these routines.
- It is a known fact that time and money are very important in today's societies. For this reason, the transition to the annual plan can be encouraged by reducing the costs of the annual membership plans and by making special campaigns, sweepstakes, coupons and discounts for people who switch to the annual membership plan. Thanks to the agreements to be made with different establishments (food restaurants, cafes, book stores, entertainment centers, etc.), a system can be developed in which users can use the points they earn according to the number and frequency of driving in these establishments.
- A system where driving time and number can be tracked would be very useful. Thanks to the application to be integrated into smart devices, it can be connected to other mobile applications such as Google fit and Apple Health, and the data can be created in the minds of users to create the perception that they are doing an activity to protect their health with their daily driving goals.

Q3: How can Cyclistic use digital media to influence casual riders to become members?

A3:

- The impact of social media on marketing and campaign processes in recent times is an undeniable fact. For this reason, encouraging campaigns can be organized using social media platforms. YouTube, Facebook, Twitter, TikTok etc. With the community and activity groups to be created on social media platforms, it can be ensured that members can share in these groups. In these community and group events where existing members will register, activities such as cycling and collective trips to certain points can be organized and shared on social media accounts. Especially with the planned activities to be held on social media, effective communication can be ensured

between our current users and this communication can be met on social media and attract new users to our structure.

- Various agreements are made with the faces known to the public from social media phenomena, and the sharing of these people about the Cyclistic company in their accounts can cause a significant increase in the number of members.
- Especially in media sharing tools such as YouTube and Tik-tok, agreements with influencers who have a channel on travel will be very effective, and the owners of these channels will use their Cyclistic bikes on their trips and make special coupons and discount draws for people watching their videos. Historical, cultural, touristic, etc. Collective trips to places can be organized through these channels and shared on social media, thus attracting the attention of more users.