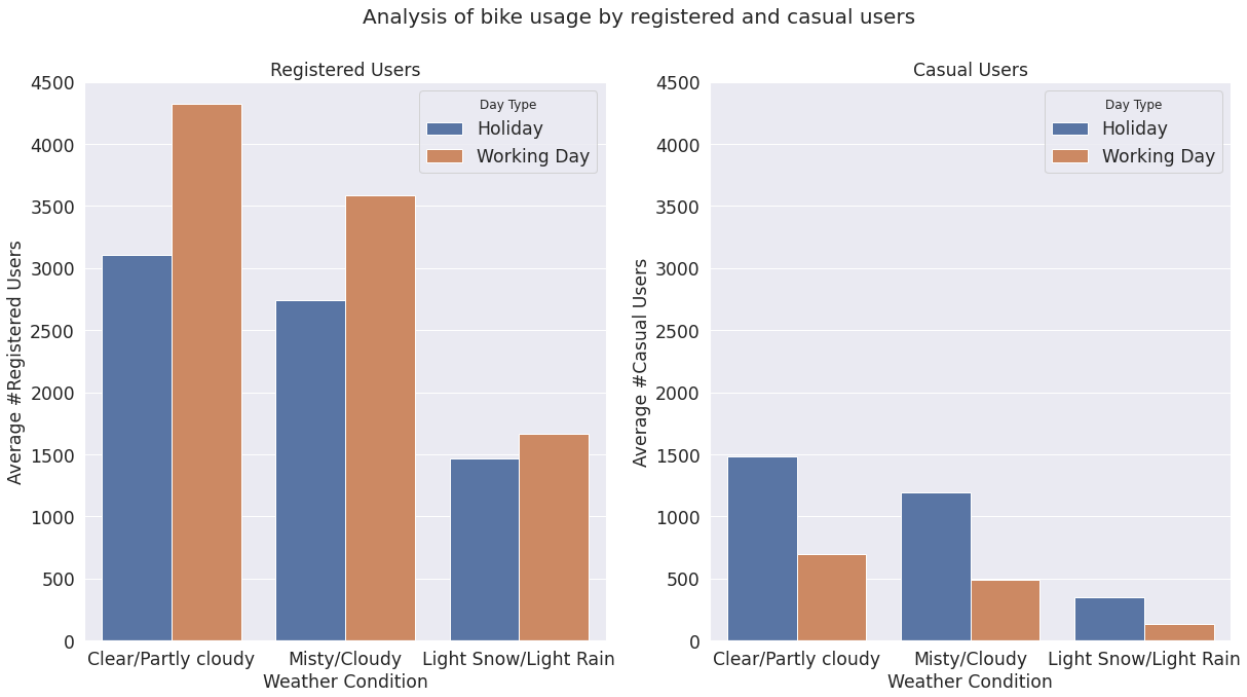
Data Visualization of Bike Sharing Dataset



Question

How does the demand for bikes from registered users and casual users vary under different conditions of weather and on working days and non-working days?

Motivation

I thought of the factors a user might consider before using a rental bike. Some users use it as part of their daily routine on working days like traveling to work or school while others use it mostly on weekends or holidays to relax. Most users don't prefer to use the bikes when the weather is not ideal. Based on these ideas, I tried to visualize the bike usage behavior of casual and registered users under different conditions of weather and whether the day is a working day or not.

Visualization

I used two **barplots**, one for registered users and another for casual users. In each plot, I plotted the average user count on working day and non-working day for different daily weather conditions. I used contrasting colors for user count on working day and non-working day and same colors across the two plots for easy comparison. I also used the same scale on the y-axis to illustrate that the demand gap from registered users and casual users.

Inferences

Firstly, as expected, the demand from both registered and casual users on both working and non-working days is higher when weather is clear(no rain or snow) than when there is chance of rain(cloudy) or actually raining or snowing. Secondly, the demand from registered users is clearly higher than casual users irrespective of the day being a working day or a holiday under a given condition of weather. More interestingly, it can be seen that the registered users are more likely to rent a bike on a working day while casual users are more likely to rent a bike on a holiday. With accurate weather forecast easily available, the bike rental company can use these insights to predict the demand on a given day and also to convert casual users to registered users by providing some incentives for registering on non-working days that could possibly make them more regular users of the service.

Cons: Temperature is found to have a strong correlation with the bike usage. This visualization does not consider the temperature of the day as a factor to assess the usage of bikes.