I used a Jupyter Notebook to conduct my analysis. I first accessed the dataset, and then cleaned the dataset by

1. taking out a row with erroneous member's age.
2. Changing the datatype of several columns, including "member_birth_year", "start_station_id", "end_station_id", "bike_id", "start_time", "end_time", etc.

To enable more analysis, I created several columns:

1. "age_group" column to divide the members by different age groups: teenagers, 20s, 30, etc.
2. "day_of_week" column to reflect on which day did each ride happen.
3. "start_hour" and "end_hour" columns to extract the hour that each ride took place.

After these steps, I started analyzing the dataset. I'm most interested in two factors: the number of rides and the duration of rides, and how they could be influenced by different independent variables. First, I did univariate exploration, in which I mainly focus on how different factors contribute to the total number of rides. I analyzed the following variables:

1. The distribution of duration of ride
2. The distribution of member birth year
3. The busiest stations
4. The number of rides in each gender group
5. The number of rides for different users (customer vs. subscriber)
6. The number of rides on different days of week
7. The number of rides variation throughout the day

For bivariate exploration, I focus on the relationship between ride duration and various factors:

1. The distribution of ride duration and members' birth year
2. The average ride duration of different gender groups
3. The average ride duration of different types of users
4. The average ride duration of different age groups
5. The average ride duration throughout the day

For multivariate exploration, I investigate the relationship among these factors:

1. How the distribution of member birth year and ride duration varies in different gender groups
2. How the distribution of member birth year and ride duration varies in different user types
3. How the distribution of member birth year and ride duration varies in different user types x gender groups
4. How the distribution of member birth year and ride duration varies throughout the week