

UberDrive- Project

1. Should we drop NA values completely from the original dataset? (Question-9) Ans: You should first copy the data in "df" object and then drop NA values from the "df".
2. In START* and STOP* columns, there is a value as R?walpindi and Rawalpindi. While analysing the unique values, these are treated as different values. Is it ok to proceed treating these as two different values? (Question-11) Ans: For this project, please treat R?walpindi and Rawalpindi as different entities. You will learn how to clean this type of issue in the future courses
3. Do we need to display the entire list or only one entry which has maximum value? (Question 15 & 16) Ans: You can choose to display the entire output in the descending order (Entry with the maximum value at the top) or only the most popular point
4. Do we need to display the entire list of most frequent routes? (Question-17) Ans: You may choose to display only the most frequent route
5. List the most frequent route taken by Uber drivers. What is the meaning of the most frequent route and an approach to solve this question? (Question-17) Ans: Most Frequent route means the route with most frequent START and STOP combination. Example Start Stop Hyderabad Delhi Delhi Chennai Chennai Bengaluru Hyderabad Bengaluru Delhi Chennai Delhi Bengaluru The expected answer: Frequent Route is Delhi -> Chennai
6. Plot a bar graph of Purposes vs Miles(Distance). In this question can I take the X-axis as Purposes as the title names are overlapping? (Question-19) Ans: Yes, you are free to choose either Purpose or Miles on X-axis
7. Display a dataframe of Purposes and the distance travelled for that particular Purpose. What is expected from this question? (Question-20) Ans: Display the total distance (using Miles variable) for each given purpose
8. What is the difference between a proportion and percentage? (Question-22) Ans: Proportion is basically a ratio and to if we multiply proportion with 100, we will have value in percentage