

UBER SUPPLY - DEMAND GAP

PROBLEM STATEMENT:

You may have some experience of travelling to and from the airport. Have you ever used Uber or any other cab service for this travel? Did you at any time face the problem of cancellation by the driver or non-availability of cars?

Well, if these are the problems faced by customers, these very issues also impact the business of Uber. If drivers cancel the request of riders or if cars are unavailable, Uber loses out on its revenue. Let's hear more about such problems that Uber faces during its operations.

The aim of the analysis is to identify the root cause of the problem (ie. cancellation and non-availability of cars) and recommended ways to improve the situation. As a result of your analysis, you should be able to present to the client the root cause and possible hypotheses of the problem and recommended ways to improve them

EXPECTED RESULTS:

1. Visually identify the most pressing problems for Uber.
 - Hint: Create plots to visualize the frequency of requests that get cancelled or show 'no cars available'; identify the most problematic types of requests (city to airport / airport to city etc.) and the time slots (early mornings, late evenings etc.) using plots
2. Find out the gap between supply and demand and show the same using plots.
 - Find the time slots when the highest gap exists
 - Find the types of requests (city-airport or airport-city) for which the gap is the most severe in the identified time slots.
3. What do you think is the reason for this issue for the supply-demand gap? Write the answer in less than 100 words. You may accompany the write-up with plot(s).
4. Recommend some ways to resolve the supply-demand gap.

Present the problem, the analyses and the recommendations using plots to the Chief Data Scientist in a well-formatted presentation (make sure to **submit a PDF version** of the PPT). Also, include **commented jupyter Notebook** in your submission. Please note that the assignment has to be done **completely in Python including the plots**.

EVALUATION CRITERIA:

These are the parameters will you be evaluated on.

Criteria	Meets expectations	Does not meet expectations
Data Cleaning and Manipulation (25 %)	<p>All data quality issues are correctly identified and reported.</p> <p>The data is converted to a clean format suitable for analysis in Python. New metrics are derived wherever required and are used for analysis.</p>	<p>Data quality issues are overlooked or are not identified correctly.</p> <p>The data is not converted to a clean format which is suitable for analysis or is not cleaned using commands in Python. New metrics are not derived or are not used for analysis.</p>
Data Analysis (45 %)	<p>The right problem is solved which is coherent with the needs of the business. The analysis has a clear structure and the flow is easy to understand.</p> <p>Realistic assumptions are made and proper reasons are given for all choices made.</p> <p>The time slots and problems are identified correctly with valid reasons.</p> <p>Univariate and segmented analysis are done correctly and successfully identify the problems.</p> <p>The demand and supply are defined properly and the numbers are correct.</p> <p>All relevant plots during the analysis are created. The choice of plots is correct, i.e. the plots clearly display the important insights. The reason for choosing certain plots, aesthetics and geometries etc. is mentioned in the comments.</p>	<p>The analyses do not address the right problem or deviate from the business objectives.</p> <p>The analysis lacks a clear structure and is not easy to follow.</p> <p>Realistic assumptions are not made wherever required or unrealistic ones are made.</p> <p>The time slots and problems are selected without any sound reasoning.</p> <p>Univariate and segmented analysis are not done correctly or are not able to identify the problems.</p> <p>The demand and supply numbers are incorrect/not defined meaningfully.</p> <p>All relevant plots are not created. The choice of plots is not ideal and the plots are either difficult to interpret or lack clarity or neatness. The reason for choosing certain plots, aesthetics and geometries etc. is not mentioned in the comments.</p> <p>Not all major issues are pointed out/many minor issues are pointed out.</p>

	All major issues are correctly pointed out.	
Presentation and Recommendations (20 %)	<p>The presentation has a clear structure, is not too long, and explains the most important results concisely.</p> <p>The recommendations to solve the problems are realistic, actionable and coherent with the analysis.</p> <p>If any assumptions are made, they are stated clearly.</p>	<p>The presentation lacks structure, is too long or does not put emphasis on the important observations.</p> <p>Contains unnecessary details or lacks the important ones.</p> <p>The recommendations to solve the problems are either unrealistic, non-actionable or incoherent with the analysis.</p> <p>Assumptions made, if any, are not stated clearly.</p>
Conciseness and Readability of the Code (10 %)	<p>The code is concise and syntactically correct.</p> <p>Wherever appropriate, built-in functions are used instead of writing long code (if-else statements, for loops).</p> <p>Custom functions are used to perform repetitive tasks.</p> <p>The code is readable with variables appropriately named and detailed comments are written wherever necessary.</p>	<p>The code has errors in the syntax.</p> <p>Long and complex codes used instead of shorter built-in functions.</p> <p>Custom functions are not used to perform repetitive tasks resulting in the same piece of code being repeated multiple times.</p> <p>Code readability is poor because of vaguely named variables or lack of comments wherever necessary.</p>