

Case Study : Numpy & Pandas

Relevel
by Unacademy



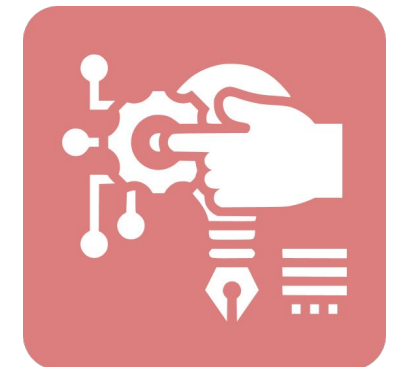
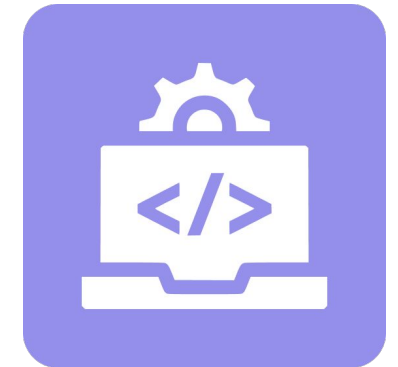
Questions for Uber Data Case Study

1. Import the dataset ('My Uber Drives - 2016.csv') into Jupyter Notebook.
2. Write a code to get the top 7 rows of the dataset.
3. Write a code to get the last 5 rows of the dataset.
4. Get the total number of rows and columns in the dataset.
5. Get the total number of elements in the dataset (Use size function).
6. Write a code to get the total number of NULL values across every column in the dataset.
7. Write a code to get the total number of Non-NULL values across every column in the dataset.
8. Write a code to get the entries having NULL values in the 'Purpose' column.
9. Write a code to get the entries having Non-NULL values in the 'Purpose' column using the Tilde (~) operator.
10. Write a code to get Non-NULL entries in the PURPOSE column.
11. Write a code to remove the * in every column name using the rename function.
12. Write a code to remove the * in every column name using the str.replace() function.
13. Write a code to remove the * in every column name using the lambda function.
14. Get the entries in the data where the START location is 'Fort Pierce'.
15. Get the entries in the data where the STOP location is 'Fort Pierce'.



Questions for Uber Data Case Study

16. Write a code to sort the entries in the data in descending order of the 'MILES' column.
17. Write a code to drop all the rows where there are NULL values in the STOP column.
18. Use describe() function to get the Statistical Properties about the numerical columns in the data.
19. Create a report in an html file using pandas profiling.
20. Get the unique and total number of unique values in the START and STOP column of the data.
21. Get the rides where we have the same START and STOP locations using a comparison operator (==).
22. Get the rides where we have the same START and STOP locations using a membership operator.
23. Use value_counts() function to demonstrate the proportion of different categorical values in the data.
24. Get the number of rides where START and STOP locations are the same.
25. Find the favorite starting point according the the total number of MILES covered. (Use groupby function).



Questions for Uber Data Case Study

26. Find the starting point for the ride where maximum miles are covered.
27. Check the data types of all the columns in the dataset.
28. Drop the 'unknown location' value from START and STOP column.
29. Find the most popular START-STOP pair according to the total number of rides covered.
30. Convert the datatypes of START_DATE and END_DATE columns to datetime.
31. Extract the month from START_DATE and try to get the proportion of rides of different months.
32. Find the average distance covered each month.
33. Extract the day from the START_DATE column.
34. Find the total miles covered per category per purpose.
35. Find the percentage of Business Miles covered and Personal miles covered.



Thank You