

DATA SCIENCE ASSIGNMENT

Day - 25:

1. What is a common objective of data cleaning in a data science project?
2. Which imputation technique is suitable for filling in missing categorical data by replacing them with the most frequently observed category?
3. In data analysis, what does a box plot primarily visualize?
4. Write a Python function to calculate the correlation matrix for numerical columns in a pandas DataFrame.
5. How do you impute missing values in a time series using forward filling (ffill) in pandas?