



Exploratory Data Analysis Interview Questions

1. What is the lifecycle of a typical data science project?
2. Mention the two kinds of target variables for predictive modelling.
3. During the data preprocessing step, how should one treat missing/null values? How will you deal with them?
4. What is an outlier and how to identify them? Why should we treat outliers?
5. How can the data be normalized?
6. What is the Difference between Univariate, Bivariate, and Multivariate analysis?
7. What is a scatter plot? For what type of data is scatter plot usually used for?
8. What Are the Key Properties of a Dataset?
9. Define the term 'Data Wrangling' in Data Analytics.
10. What type of plot would you use if you need to demonstrate “relationship” between variables/parameters?
11. What is the significance of Exploratory Data Analysis(EDA)?
12. What are the different types of sampling techniques used by data analysts?
13. When will you use a histogram and when will you use a bar chart? Explain with an example.
14. How do you treat outliers in a dataset?
15. What is data cleansing and what are the best ways to practice data cleansing?
16. What type of data is box-plots usually used for? Why?
17. How is box plot different from histograms?
18. How is Box-Cox Transformation and when we set Lambda to 0 (in Scikit Learn), what kind of transformation is processed.
19. What are the different challenges one faces during data analysis?
20. What information could you gain from a box-plot?
21. _____ is a good choice to show how some quantity varies among some discrete set of items
22. What is the difference between Label Encoding and One Hot Encoding? How do you decide which one to use? Share some examples.

23. Why do we use power transformation methods on the dataset?
Explain with an example.

24. What is the difference between feature engineering and feature selection?

25. How do you transform skewed distribution to normal distribution?