Theoretical Questions

1) What is NumPy, and why is it widely used?

NumPy is a Python library for fast numerical computation, offering ndarrays, broadcasting, and linear algebra.

2) Broadcasting in NumPy

Allows operations between arrays of different shapes by virtually expanding smaller ones.

3) Pandas DataFrame

2D labeled data structure with rows and columns.

4) groupby() use

Splits, applies function, and combines results for grouped analysis.

5) Why Seaborn for visualization?

Provides statistical plots with attractive defaults and easy syntax.

6) NumPy vs Python lists

Arrays: homogeneous, efficient, support vectorized ops. Lists: heterogeneous, slower.

7) Heatmap

A color-coded matrix for correlations/values.

8) Vectorized operation

Applying operations element-wise over arrays without loops.

9) Matplotlib vs Plotly

Matplotlib: static, flexible; Plotly: interactive, web-ready.

10) Hierarchical indexing

Multi-level row/column indexes for complex data.

11) Seaborn pairplot()

Plots pairwise relationships in dataset.

12) describe() in Pandas

Gives summary stats of columns.

13) Handling missing data importance

Prevents errors and biased results.

14) Benefits of Plotly

Interactive, supports 3D and dashboards.

15) NumPy multidimensional arrays

Supports ndarrays efficiently.

16) Role of Bokeh

Interactive web visualizations.

17) apply() vs map() in Pandas

apply: on DataFrame/Series, rows/cols. map: element-wise on Series.

18) Advanced NumPy features

Broadcasting, FFT, linear algebra, random sampling, masked arrays.

19) Pandas time series

Has datetime index, resampling, shifting, rolling functions.

20) Pivot table role

Summarizes and reshapes data by aggregating values.

21) NumPy slicing faster

Contiguous memory storage vs list references.

22) Common Seaborn use cases

Distribution plots, categorical plots, regression, correlation heatmaps