## **Basic Conceptual Questions**

#### 1. What is a library in Python?

➤ A library is a collection of pre-written code that provides functions and methods for specific tasks.

#### 2. Why do we use matplotlib?

➤ We use matplotlib to visualize data using different types of plots and graphs.

#### 3. What is a DataFrame?

➤ A DataFrame is a 2D labeled data structure like a table in pandas.

#### 4. What is the difference between DataFrame, Excel, and SQL table?

➤ DataFrame is an in-memory table in Python; Excel is a spreadsheet software; SQL table is stored in a database.

#### 5. What is the difference between a bar plot and a histogram?

➤ Bar plot shows categorical data; histogram shows frequency of continuous data.

#### 6. Can we set marker color in line plot?

➤ Yes, using markerfacecolor and markeredgecolor in plot().

#### 7. What is pandas used for?

➤ It's used for data manipulation and analysis.

#### 8. Difference between Series and DataFrame?

➤ Series is 1D: DataFrame is 2D with rows and columns.

#### 9. What is the use of .head() in pandas?

➤ It displays the first few rows of the DataFrame.

#### 10. What is nan?

➤ NaN stands for "Not a Number", used to represent missing values.

# 🔁 Tricky or Comparison-Based Questions

#### 11. Difference between .loc[] and .iloc[]?

➤ .loc[] is label-based indexing; .iloc[] is position-based indexing.

#### 12. Difference between mean() and median()?

➤ Mean is the average; median is the middle value.

#### 13. Difference between max() and idxmax()?

➤ max() returns max value; idxmax() returns index of max value.

- 14. Difference between apply () and map ()?
  - ➤ apply() works on rows/columns; map() works on individual elements of Series.
- 15. Difference between list and array in Python?
  - Lists can store mixed types; arrays require same data type.
- 16. Can we create plots without importing matplotlib.pyplot?
  - ➤ No, pyplot is essential for plotting with matplotlib.
- 17. What is the use of groupby () in pandas?
  - ➤ It groups data based on a column for aggregation.
- 18. What is vectorization in pandas?
  - ➤ Applying operations to entire arrays without loops.
- 19. How to check data types of all columns?
  - ➤ Use df.dtypes.
- 20. What does .describe() do in pandas?
  - ➤ It gives statistical summary like mean, std, min, max, etc.

# 🚺 Visualization-Focused Questions

- 21. What is a histogram used for?
  - ➤ To show frequency distribution of continuous data.
- 22. When do we use pie charts?
  - ➤ To show proportion of parts to a whole.
- 23. Can bar plots be horizontal?
  - ➤ Yes, using barh().
- 24. What is subplot() in matplotlib?
  - ➤ It is used to display multiple plots in one figure.
- 25. What does autoput do in pie chart?
  - ➤ It formats percentage display on the pie chart.
- 26. What is the use of legend()?
  - ➤ It adds a legend (label info) to the plot.
- 27. What is the role of xlabel() and ylabel()?
  - ➤ They label the x and y axes respectively.
- 28. What does title() do in a plot?
  - ➤ Sets the title for the plot.

- 29. What is plt. show()?
  - ➤ It displays the plot on the screen.
- 30. What does color do in a plot?
  - ➤ It sets the color of lines or bars in the plot.

# Core Data Questions

- 31. What does value\_counts() do?
  - ➤ It returns frequency counts of unique values.
- 32. How do you find null values?
  - ➤ Using df.isnull().sum().
- 33. How to drop missing values?
  - ➤ Use df.dropna().
- 34. How to fill missing values?
  - ➤ Use df.fillna(value).
- 35. What is data wrangling?
  - ➤ It's the process of cleaning and organizing raw data.
- 36. Can pandas read Excel files?
  - ➤ Yes, using read excel().
- 37. How to merge two DataFrames?
  - ➤ Use pd.merge(df1, df2).
- 38. How to concatenate DataFrames vertically?
  - ➤ Use pd.concat([df1, df2]).
- 39. How to rename a column?
  - ➤ Use df.rename(columns={'old':'new'}).
- 40. What is the shape of DataFrame?
  - ➤ It gives the tuple (rows, columns).

# > Bonus Tricky Questions

- 41. Is pandas built on top of NumPy?
  - ➤ Yes, it uses NumPy internally.

### 42. Can a DataFrame contain different data types in columns?

➤ Yes, each column can have different data types.

#### 43. Can a DataFrame contain lists as values?

➤ Yes, but it's not preferred for analysis.

#### 44. Can we sort a DataFrame?

➤ Yes, using sort values().

# 45. Can you plot directly from a DataFrame?

➤ Yes, using df.plot().

#### 46. How to reset index in a DataFrame?

➤ Use df.reset\_index().

# 47. Can you filter rows based on condition?

➤ Yes, using boolean indexing like df[df['age'] > 20].

## 48. Can you apply function to a column?

➤ Yes, using df['col'].apply(func).

## 49. What is slicing in pandas?

➤ Extracting subset of rows or columns.

## 50. Can you remove duplicates in pandas?

➤ Yes, using df.drop\_duplicates().