Topic: Pandas

Name: Ritesh Meshram

1. What is the purpose of the pandas library in Python? What are the available data structures in pandas?

Answer: The pandas library is a popular data manipulation library in Python used to handle and analyze tabular data. It provides easy-to-use data structures for data analysis such as Series (1D), DataFrame (2D), and Panel (3D). The available data structures in pandas are Series, DataFrame, and Panel.

2. What are the notable features provided by the pandas library?

Answer: Some notable features provided by the pandas library are easy handling of

missing data, merging and joining datasets, grouping and aggregation, time-series analysis, data filtering, pivoting, reshaping, and plotting.

3. What are the differences between a series and a dataframe in pandas?

Answer: A series is a one-dimensional labeled array that can hold any data type while a DataFrame is a two-dimensional labeled data structure with columns of potentially different data types. In other words, a series is a single column of a DataFrame.

4. What are the various methods for creating a series in pandas?

Answer: There are several methods for creating a series in pandas, such as creating a series from a list, dictionary, scalar value, ndarray, and file.

5. How can a dataframe be created in different ways?

Answer: A DataFrame can be created in different ways, such as creating a DataFrame from a dictionary, list of dictionaries, ndarray, CSV file, Excel file, SQL database, and HTML file.

6. Can you provide some examples of statistical functions available in the pandas library for Python?

Answer: The pandas library provides various statistical functions such as mean, median, mode, standard deviation, variance, skewness, kurtosis, correlation, covariance, and percentile.

7. How can you sort a dataframe in pandas?

Answer: A DataFrame in pandas can be sorted using the sort_values() method. You can sort by one or more columns and specify the ascending or descending order.

8. What is the process of setting an index for a pandas dataframe?

Answer: The process of setting an index for a pandas DataFrame involves selecting a column as the index to label the rows. You can set an index using the set_index() method or by passing the index_col parameter when reading a CSV file.

9. How do you add a new row to an existing pandas dataframe?

Answer: To add a new row to an existing pandas DataFrame, you can create a new DataFrame and concatenate it with the original using the concat() method or use

the append() method to append a row to the original DataFrame.

10. What is the procedure for adding a new column to a pandas dataframe?

Answer: To add a new column to a pandas DataFrame, you can assign a new column name to the DataFrame and populate it with values using the indexing operator or the assign() method.

11. How can you convert a numpy array into a dataframe using pandas?

Answer: To convert a numpy array into a DataFrame using pandas, you can pass the numpy array and column names as arguments to the DataFrame() constructor.

12. What is the method to delete a row from a pandas dataframe?

Answer: To delete a row from a pandas DataFrame, you can use the drop() method and specify the index or label of the row to be removed.