1. What is NumPy?

NumPy is a Python library used for numerical computing, providing support for arrays, matrices, and a wide range of mathematical functions to operate on these data structures efficiently.

1. How is vstack() different from hstack() in NumPy?

vstack() stacks arrays vertically, combining them along the vertical axis (rows), while hstack() stacks arrays horizontally, combining them along the horizontal axis (columns).

1. List the advantages NumPy Arrays have over (nested) Python lists.

NumPy arrays are more memory efficient, provide faster mathematical operations, support vectorized operations, and have additional functionalities such as broadcasting, slicing, and reshaping.

1. How do you convert a Pandas DataFrame to a NumPy array?

You can convert a Pandas DataFrame to a NumPy array using the values attribute of the DataFrame. For example: numpy\_array = dataframe.values.

1. What are the different types of data structures in Pandas?

The main data structures in Pandas are Series (one-dimensional labeled arrays) and DataFrame (two-dimensional labeled data structure with columns of potentially different types).

1. What are the most important features of The Pandas library?

Some of the most important features of Pandas include data manipulation and analysis tools, powerful indexing and slicing capabilities, support for handling missing data, integration with other libraries, and flexible reshaping and pivoting functions.

1. How do you get the frequency count of the unique items in a series?

You can get the frequency count of unique items in a Series using the value\_counts() method. For example: series.value\_counts().

1. What are the different ways of creating DataFrame in Pandas? Explain with examples.

DataFrame in Pandas can be created from dictionaries, lists, NumPy arrays, or other DataFrames.

For example:

From dictionary:

data = {'Column1': [value1, value2], 'Column2': [value3, value4]}

df = pd.DataFrame(data)

From list of lists:

data = [[value1, value2], [value3, value4]]

df = pd.DataFrame(data, columns=['Column1', 'Column2'])

1. How are loc and iloc different in Pandas?

loc is label-based indexing, used for selecting rows and columns by label names, while iloc is integer-based indexing, used for selecting rows and columns by integer indices.

1. How does the groupby() method works in Pandas?

The groupby() method in Pandas is used to split the DataFrame into groups based on specified criteria, such as a column name, and then applies a function (such as mean, sum, count) to each group independently, and finally combines the results into a new DataFrame.