**LAB – 2: GETTING STARTED WITH PANDAS**

**PANDAS:**

Pandas is a Python software library for data analysis and manipulation widely used in data science and machine learning applications enabling cleansing, merging, and reshaping of data structures for big datasets, handling various file formats including Excel, JSON, Parquet, SQL database tables or queries, and CSV files. **Installation: pip install pandas**

Here is a brief description and syntax for the Pandas methods:

1. **Series()**:This method creates a new Series with the specified index and data.

Syntax: pd.Series(data, index=['a', 'b', 'c'])

1. **DataFrame()**: This method creates a new DataFrame with the specified data, where each row is a pandas Series.

Syntax: pd.DataFrame(data)

1. **read\_csv()**: This method reads a CSV file into a pandas DataFrame.

Syntax: pd.read\_csv('path/to/file.csv')

1. **loc()**: This method is used to select rows and columns of the DataFrame based on labels or index values.

Syntax: df.loc[label1, 'column1']

1. **head()**: This method returns the first N rows of the DataFrame by default, or the number specified as an argument.

Syntax: df.head(n=5)

1. **tail()**: This method returns the last N rows of the DataFrame by default, or the number specified as an argument.

Syntax: df.tail(n=5)

1. **sample()**: This method returns a random sample of the DataFrame of the specified size or type.

Syntax: df.sample(n=5)

1. **info()**: This method provides a summary of the DataFrame, including the index type, column types, and memory usage.

Syntax: df.info()

Lab program with all the above mentioned methods is given below: