PANDAS SHEET

PANDAS

1. DataFrame Creation and Inspection

- `pd.DataFrame()`: Create a DataFrame from various data sources (lists, dictionaries, etc.).
- `pd.Series()`: Create a Series object.
- `df.head()`: Return the first n rows of the DataFrame.
- `df.tail()`: Return the last n rows of the DataFrame.
- `df.info()`: Summary of DataFrame including the index dtype and columns.
- `df.describe()`: Generate descriptive statistics of the DataFrame.
- `df.shape`: Return a tuple representing the dimensionality of the DataFrame.
- `df.columns`: Return the column labels of the DataFrame.
- `df.index`: Return the index (row labels) of the DataFrame.

PANDAS SHEET

- `df.loc[]`: Access a group of rows and columns by labels or a boolean array.
- `df.iloc[]`: Access a group of rows and columns by integer position.
- `df.at[]`: Access a single value for a row/column label.
- `df.iat[]`: Access a single value for a row/column position.
- `df.query()`: Query the DataFrame using a boolean expression.
- `df.filter()`: Filter the DataFrame by specific criteria.

3. Data Cleaning and Manipulation

- `df.drop()`: Remove rows or columns by label.
- `df.dropna()`: Remove missing values.
- `df.fillna()`: Fill missing values.
- `df.replace()`: Replace values with another value.
- `df.rename()`: Rename labels of rows or columns.
- `df.duplicated()`: Detect duplicate rows.

PANDAS SHEET

- `df.sort_values()`: Sort by the values along either axis.
- `df.sort_index()`: Sort by the DataFrame's index.
- `df.set_index()`: Set the DataFrame index using one or more columns.
- `df.reset_index()`: Reset the index of the DataFrame.

4. Data Aggregation and Grouping

- `df.groupby()`: Group DataFrame using a mapper or by a series of columns.
- `df.agg()`: Aggregate using one or more operations over the specified axis.
- `df.transform()`: Transform data by applying a function.
- `df.pivot_table()`: Create a pivot table.
- `df.crosstab()`: Compute a cross-tabulation of two or more factors.

5. Merging and Joining

- `pd.merge()`: Merge DataFrames based on one or more keys.
- `df.join()`: Join DataFrames based on the index or a key column.

PANDAS SHEET

- `pd.merge_asof()`: Perform an asof merge.

6. Data Visualization

- `df.plot()`: Plot DataFrame using matplotlib.
- `df.hist()`: Plot histograms.
- `df.boxplot()`: Generate a box plot.

7. Time Series

- `pd.to_datetime()`: Convert argument to datetime.
- `df.resample()`: Resample time series data.
- `df.shift()`: Shift the index by a specified number of periods.

8. Input/Output

- `pd.read_csv()`: Read a comma-separated values (csv) file into DataFrame.
- `pd.read_excel()`: Read an Excel file into DataFrame.
- `df.to_csv()`: Write DataFrame to a CSV file.

PANDAS SHEET

9. Statistical Functions

- `df.mean()`: Return the mean of the values.
- `df.median()`: Return the median of the values.
- `df.std()`: Return the standard deviation of the values.
- `df.sum()`: Return the sum of the values.
- `df.min()`: Return the minimum value.
- `df.max()`: Return the maximum value.

10. Other Functions

- `df.apply()`: Apply a function along an axis of the DataFrame.
- `df.applymap()`: Apply a function to each element of the DataFrame.
- `df.map()`: Map values using an input mapping function.
- `df.astype()`: Convert the DataFrame to a specified type.