***MIT Academy of Engineering***

***School of Chemical engineering***

***Name : Gauraang Prashant Dhamal***

***Roll No : CC-66 PRN:202401050066***

Question: Explain all pandas operations.

Solution : Pandas is a Python library providing data structures and tools for data analysis. Its core components are Series (one-dimensional labeled array) and DataFrame (two-dimensional table with labeled rows and columns) Here's a breakdown of common Pandas operations:

* **Data Input/Output:**
  + read\_csv(), read\_excel(), read\_json(): Read data from various file formats into a DataFrame.
  + to\_csv(), to\_excel(), to\_json(): Write DataFrame to various file formats.
* **Data Selection and Indexing:**
  + []: Access columns by label.
  + .loc[]: Access rows and columns by label.
  + .iloc[]: Access rows and columns by integer position.
  + .at[]: Access a single value by label.
  + .iat[]: Access a single value by integer position.
* **Data Manipulation:**
  + Adding columns: df['new\_column'] = ...
  + Deleting columns: del df['column'] or df.drop('column', axis=1)
  + Adding rows: df.loc['new\_row'] = ...
  + Deleting rows: df.drop('row\_label')
  + Renaming columns/rows: df.rename(columns={'old': 'new'}, index={'old': 'new'})
* **Data Cleaning:**
  + dropna(): Remove rows or columns with missing values.
  + fillna(): Fill missing values with a specified value or method.
  + replace(): Replace specific values.
  + duplicated(): Identify duplicate rows.
  + drop\_duplicates(): Remove duplicate rows.
* **Data Transformation:**
  + groupby(): Group data based on one or more columns for aggregation.
  + sort\_values(): Sort data by one or more columns.
  + pivot\_table(): Create a pivot table from a DataFrame.
  + melt(): Unpivot a DataFrame from wide to long format.
  + apply(): Apply a function to rows or columns.
* **Data Aggregation:**
  + count(): Count non-null values.
  + sum(): Sum of values.
  + mean(): Mean of values.
  + median(): Median of values.
  + min(): Minimum value.
  + max(): Maximum value.
  + std(): Standard deviation of values.
  + var(): Variance of values.
* **Combining Data:**
  + concat(): Concatenate DataFrames along rows or columns.
  + merge(): Join DataFrames based on common columns.
* **Mathematical Operations:**
  + Arithmetic operations (+, -, \\*, /, \\*\\*) are performed element-wise, aligning on index and columns.
* **Boolean Operations:**
  + Comparison operators (==, !=, <, >, <=, >=) are performed element-wise.
  + Logical operators (&, |, ~) can be used for boolean indexing.
* **Other operations:**
  + all(): Returns True if all elements are True, otherwise returns False.
  + any(): Returns True if any element is True, otherwise returns False.
  + isin(): Check if elements in a Series or DataFrame are contained in another sequence.
  + unique(): Returns unique values in a Series.
  + nunique(): Returns the number of unique values in a Series.
  + value\_counts(): Returns the frequency of each unique value in a Series.

Pandas operations are designed to be efficient and flexible, enabling users to perform complex data analysis tasks with ease.