Basics of Pandas

Function: pd.Series()

Description: Creates a one-dimensional labeled array

Example: pd.Series([1, 2, 3])

Function: pd.DataFrame()

Description: Creates a two-dimensional labeled data structure

Example: pd.DataFrame({'A': [1, 2]})

Data Input/Output (I/O)

Function: read_csv()

Description: Reads a CSV file into a DataFrame

Example: pd.read_csv('data.csv')

Function: to_csv()

Description: Writes a DataFrame to a CSV file

Example: df.to_csv('output.csv')

Function: read_excel()

Description: Reads an Excel file

Example: pd.read_excel('data.xlsx')

Function: to_excel()

Description: Writes to an Excel file

Example: df.to_excel('output.xlsx')

Data Inspection

Function: head()

Description: Returns first n rows

Example: df.head()

Function: info()

Description: Provides DataFrame structure

Example: df.info()

Function: describe()

Description: Statistical summary

Example: df.describe()

Function: shape

Description: Returns DataFrame shape

Example: df.shape

Selection and Indexing

Function: df['col']

Description: Selects a column

Example: df['Name']

Function: loc[]

Description: Label-based indexing

Example: df.loc[0]

Function: iloc[]

Description: Position-based indexing

Example: df.iloc[0]

Function: at[]

Description: Access scalar by label

Example: df.at[0, 'A']

Data Manipulation

Function: drop()

Description: Drops rows or columns

Example: df.drop('col', axis=1)

Function: rename()

Description: Renames columns/index

Example: df.rename(columns={'old': 'new'})

Function: astype()

Description: Casts to dtype

Example: df['col'].astype(int)

Missing Data Handling

Function: isnull()

Description: Detects missing values

Example: df.isnull()

Function: fillna()

Description: Fills missing values

Example: df.fillna(0)

Function: dropna()

Description: Removes missing values

Example: df.dropna()

Statistical Operations

Function: mean()

Description: Mean of values

Example: df.mean()

Function: median()

Description: Median of values

Example: df.median()

Function: std()

Description: Standard deviation

Example: df.std()

Function Application

Function: apply()

Description: Applies function to axis

Example: df['col'].apply(lambda x: x * 2)

Function: map()

Description: Maps values using dict/function

Example: df['col'].map({'A': 1})

Function: applymap()

Description: Element-wise function application

Example: df.applymap(lambda x: x*2)

Sorting and Ranking

Function: sort_values()

Description: Sorts by column

Example: df.sort_values(by='Age')

Function: sort_index()

Description: Sorts by index

Example: df.sort_index()

Function: rank()

Description: Ranks values

Example: df['score'].rank()

Combining DataFrames

Function: concat()

Description: Concatenates DataFrames

Example: pd.concat([df1, df2])

Function: merge()

Description: Merges DataFrames

Example: pd.merge(df1, df2, on='ID')

Function: join()

Description: Joins on index

Example: df1.join(df2)

Grouping and Aggregation

Function: groupby()

Description: Groups and aggregates

Example: df.groupby('A').mean()

Function: agg()

Description: Aggregate with functions

Example: df.agg({'col1': 'sum'})

Pivoting and Reshaping

Function: pivot()

Description: Reshape data

Example: df.pivot(index='A', columns='B', values='C')

Function: melt()

Description: Unpivots DataFrame

Example: pd.melt(df)

Date and Time

Function: to_datetime()

Description: Convert to datetime

Example: pd.to_datetime(df['date'])

Function: resample()

Description: Resample time series

Example: df.resample('M').mean()

String Handling

Function: str.lower()

Description: Lowercase strings

Example: df['name'].str.lower()

Function: str.contains()

Description: Checks for substring

Example: df['email'].str.contains('@')

Window Functions

Function: rolling()

Description: Rolling window operations

Example: df['col'].rolling(3).mean()

Function: expanding()

Description: Expanding window

Example: df['col'].expanding().mean()

Categorical Data

Function: astype('category')

Description: Convert to categorical

Example: df['col'] = df['col'].astype('category')

Function: cat.codes

Description: Category codes

Example: df['col'].cat.codes

Evaluation

Function: eval()

Description: Evaluate expression

Example: df.eval('total = A + B')

Function: query()

Description: Query DataFrame

Example: df.query('A > B')

MultiIndex

Function: set_index()

Description: Sets DataFrame index

Example: df.set_index(['A', 'B'])

Function: swaplevel()

Description: Swaps MultiIndex levels

Example: df.swaplevel()

Data Security/Integrity

Function: copy()

Description: Deep copy of DataFrame

Example: df2 = df.copy()

Function: equals()

Description: Checks equality

Example: df.equals(df2)

Advanced Options

Function: set_option()

Description: Sets display options

Example: pd.set_option('display.max_rows', 100)

Plotting

Function: plot()

Description: Plots data using Matplotlib

Example: df.plot(kind='line')