Pandas Library Cheat Sheet

Importing Data

Actions	Description	Example Snippet
Import	Standard import statement to bring Pandas into the script.	`import pandas as pd`
Read_CSV	Reads a comma-separated values (CSV) file into DataFrame.	`df = pd.read_csv('file.csv')`
Read_Table	Reads a general delimited file into DataFrame.	`df = pd.read_table('file.txt')`
Read_Excel	Reads an Excel file into DataFrame.	`df = pd.read_excel('file.xlsx')`
Read_SQL	Reads SQL query or database table into DataFrame.	`df = pd.read_sql('SELECT * FROM table', conn)`
Read_JSON	Reads a JSON formatted string into DataFrame.	`df = pd.read_json('file.json')`
Read_HTML	Reads HTML tables into DataFrame.	`df = pd.read_html('url')`
Clipboard	Reads text from the clipboard into DataFrame.	`df = pd.read_clipboard()`

Exporting Data

Actions	Description	Example Snippet
To_CSV	Writes DataFrame to a comma-separated values (CSV) file.	`df.to_csv('file.csv')`
To_Excel	Writes DataFrame to an Excel file.	`df.to_excel('file.xlsx')`
To_SQL	Writes DataFrame to a SQL database.	`df.to_sql('table_name', conn)`
To_JSON	Writes DataFrame to a JSON formatted string.	`df.to_json('file.json')`
To_HTML	Writes DataFrame to HTML tables.	`df.to_html('file.html')`
To_Clipboard	Writes DataFrame to the clipboard.	`df.to_clipboard()`

Create Test Objects

Actions	Description	Example Snippet
Dataframe	Constructs a DataFrame object.	`df = pd.DataFrame(data)`
Series	Constructs a Series object.	`s = pd.Series(data)`
Index	Constructs an Index object.	`index = pd.Index(data)`

DataFrame Basics

Actions	Description	Example Snippet
Return Dimensions of a DataFrame	Gets shape of DataFrame.	`df.shape`
Read CSV file into a DataFrame	Reads CSV and returns DataFrame object.	`df = pd.read_csv('file.csv')`
Return the data type of each column	Returns data types of columns in DataFrame.	`df.dtypes`

Selecting DataFrame Values

Actions	Description	Example Snippet
Select the rank column from f500	Selects a specific column from DataFrame.	`f500['rank']`
Select the first 3 rows from f500	Slices the DataFrame.	`f500.head(3)`

ILOC / LOC

Actions	Description	Example Snippet
LOC	Access a group of rows and columns by labels.	`df.loc[row_index, 'column_name']`
ILOC	Access a group of rows and columns by integer index.	`df.iloc[row_index, col_index]`

coursera

Graphs

Actions	Description	Example Snippet
Generate a frequent table from a series object	Counts unique values in Series.	`series.value_counts()`
Generate a sorted frequency table from series object	Counts and sorts unique values.	`series.value_counts().sort_values()`
Generate a vertical bar plot from a series object	Plots bar chart from Series.	`series.plot.bar()`
Generate a horizontal bar plot from a series object	Plots horizontal bar chart from Series.	`series.plot.barh()`
Generate a line plot from a DataFrame object	Plots line chart from DataFrame.	`df.plot.line()`
Generate a scatter plot from a DataFrame object	Plots scatter chart.	`df.plot.scatter(x='col1', y='col2')`

Statistics

Actions	Description	Example Snippet
Describe	Generates descriptive statistics.	`df.describe()`
Mean	Computes mean of DataFrame.	`df.mean()`
Corr	Computes pairwise correlation of columns.	`df.corr()`
Count	Returns the number of non-NA/null observations.	`df.count()`
Max	Returns the maximum of DataFrame values.	`df.max()`
Min	Returns the minimum of DataFrame values.	`df.min()`
Median	Computes the median of DataFrame columns.	`df.median()`
STD	Computes the standard deviation of DataFrame columns.	`df.std()`

Data Cleaning

Actions	Description	Example Snippet
Columns	Access columns of DataFrame as attributes.	`df.columns`
Isnull	Detects missing values.	`df.isnull()`
Notnull	Detects non-missing values.	`df.notnull()`
Dropna	Removes missing values.	`df.dropna()`
Fillna	Fills missing values.	`df.fillna(value)`
Astype	Converts data type of a DataFrame column.	`df['col'] = df['col'].astype('int')`
Replace	Replaces values.	`df.replace(to_replace, value)`
Rename	Renames DataFrame columns.	`df.rename(columns={'old': 'new'})`
Set_index	Sets DataFrame index.	`df.set_index('col')`
Finding correlation	Computes pairwise correlation of columns.	`df.corr()`
Converting a column to datetime	Converts column to datetime format.	`df['date'] = pd.to_datetime(df['date'])`

Boolean Masks / Operators

Actions	Description	Example Snippet
Boolean Masks	Filters DataFrame based on a condition.	`df[df['column'] > value]`
Boolean Operators	Combines multiple conditions for filtering.	`df[(df['column'] > value) & (df['column'] < value2)]`