Working with Data in Python Cheat Sheet

Reading and writing files

```
Package/Method Description
                               Syntax: r (reading) w (writing) a (appending) + (updating: read/write) b (binary, otherwise text)
                  Different
                  modes to
File opening
                  open files
modes
                                  1. Examples: with open("data.txt", "r") as file: content = file.read() print(content) with open("output.txt", "w") as file:
                  for specific
                  operations.
                               Copied!
                               Syntax:
                                  1. 1
                                  2. 2
                                  1. file.readlines() # reads all lines as a list
                                  readline() # reads the next line as a string
                                  3. file.read() # reads the entire file content as a string
                  Different
                                Copied!
                  methods to
File reading
                  read file
                               Example:
methods
                  content in
                  various
                                  1. 1
                  ways.
                                  2. 2
                                 3. 3
4. 4
                                     with open("data.txt", "r") as file:
                                  1.
                                         lines = file.readlines()
next_line = file.readline()
content = file.read()
                                  2.
                                  3.
                                  4.
                               Copied!
                               Syntax:
                                  1. 1
                                  1. file.write(content) # writes a string to the file
                                  2. file.writelines(lines) # writes a list of strings to the file
                  Different
                               Copied!
                  write
File writing
                  methods to
                               Example:
methods
                  write
                  content to a
                  file.
                                  3. 3

    lines = ["Hello\n", "World\n"]
    with open("output.txt", "w") as file:

                                         file.writelines(lines)
                               Copied!
                               Syntax:
                                  1. 1
                                  1. for line in file: # Code to process each line
                  Iterates
                               Copied!
                  through
                  each line in
Iterating over
                               Example:
                  the file
lines
                  using a
                  'loop'.
                                  1. with open("data.txt", "r") as file:
                                  2. for line in file: print(line)
                               Copied!
Open() and
                  Opens a
                               Syntax:
close()
                  file,
                                 1. 1
2. 2
                  performs
                  operations,
                  and
                                  1. file = open(filename, mode) # Code that uses the file
                  explicitly
                                  2. file.close()
                  closes the
                                Copied!
                  file using
                  the close()
                               Example:
                  method.
                                  1. 1
                                  2. 2
                                  1. file = open("data.txt", "r")
```

about:blank 1/5

2. content = file.read()

```
3. file.close()
                               Copied!
                               Syntax:
                                 1. 1
                                 1. with open(filename, mode) as file: # Code that uses the file
                  Opens a file
                  using a with Copied!
                  block,
with open()
                  ensuring
                               Example:
                  automatic
                                 1. 1
                  file closure
                                 2. 2
                  after usage.
                                 1. with open("data.txt", "r") as file:
                                 2. content = file.read()
                               Copied!
Pandas
Package/Method
                                                                                                           Syntax and Code Example
                                        Description
                                                                          Syntax: dataframe_name = pd.read_csv("filename.csv") Example: df =
                  Reads data from a `.CSV` file and creates a DataFrame.
.read_csv()
                                                                          pd.read_csv("data.csv")
                                                                          Syntax:
                                                                             1. 1
                                                                             1. dataframe_name = pd.read_excel("filename.xlsx")
                                                                           Copied!
.read excel()
                  Reads data from an Excel file and creates a DataFrame.
                                                                          Example:
                                                                             1. 1
                                                                             1. df = pd.read_excel("data.xlsx")
                                                                           Copied!
                                                                          Syntax:
                                                                             1. 1
                                                                             1. dataframe_name.to_csv("output.csv", index=False)
                                                                          Copied!
                  Writes DataFrame to a CSV file.
.to_csv()
                                                                          Example:
                                                                             1. 1
                                                                             1. df.to_csv("output.csv", index=False)
                                                                          Copied!
                                                                          Syntax:
                                                                             1. 1
                                                                             1. dataframe_name["column_name"] # Accesses single column
2. dataframe_name[["column1", "column2"]] # Accesses multiple columns
                                                                          Copied!
Access Columns  Accesses a specific column using [] in the DataFrame.
                                                                          Example:
                                                                             1. 1
2. 2

    df["age"]
    df[["name", "age"]]

                                                                           Copied!
                                                                          Syntax:
                                                                             1. 1

    dataframe_name.describe()

                                                                          Copied!
                  Generates statistics summary of numeric columns in the
describe()
                  DataFrame.
                                                                          Example:
                                                                             1. 1
                                                                             1. df.describe()
                                                                          Copied!
```

about:blank 2/5

```
Syntax:
                                                                                1. 1
                                                                                1. dataframe_name.drop(["column1", "column2"], axis=1, inplace=True)
2. dataframe_name.drop(index=[row1, row2], axis=0, inplace=True)
                                                                              Copied!
                  Removes specified rows or columns from the
drop()
                  DataFrame. axis=1 indicates columns. axis=0 indicates
                                                                             Example:
                  rows.
                                                                                2. 2
                                                                                1. df.drop(["age", "salary"], axis=1, inplace=True) # Will drop columns
2. df.drop(index=[5, 10], axis=0, inplace=True) # Will drop rows
                                                                             Copied!
                                                                             Syntax:
                                                                                1. 1

    dataframe_name.dropna(axis=0, inplace=True)

                                                                              Copied!
                  Removes rows with missing NaN values from the
dropna()
                  DataFrame. axis=0 indicates rows.
                                                                             Example:
                                                                                1. 1

    df.dropna(axis=0, inplace=True)

                                                                              Copied!
                                                                             Syntax:
                                                                                1. 1

    dataframe_name.duplicated()

                                                                             Copied!
                  Duplicate or repetitive values or records within a data
duplicated()
                                                                             Example:
                                                                                1. 1
                                                                                1. duplicate_rows = df[df.duplicated()]
                                                                              Copied!
                                                                             Syntax:
                                                                                1. filtered_df = dataframe_name[(Conditional_statements)]
                                                                              Copied!
                  Creates a new DataFrame with rows that meet specified
Filter Rows
                  conditions.
                                                                             Example:
                                                                                1. filtered_df = df[(df["age"] > 30) & (df["salary"] < 50000)</pre>
                                                                             Copied!
                                                                             Syntax:
                                                                                1. 1
                                                                                2. 2
                                                                                1. grouped = dataframe_name.groupby(by, axis=0, level=None, as_index=True,
                                                                                sort=True, group_keys=True, squeeze=False, observed=False, dropna=True)
                  Splits a DataFrame into groups based on specified
                                                                             Copied!
groupby()
                  criteria, enabling subsequent aggregation,
                  transformation, or analysis within each group.
                                                                             Example:
                                                                                1. 1
                                                                                1. grouped = df.groupby(["category", "region"]).agg({"sales": "sum"})
                                                                             Copied!
                  Displays the first n rows of the DataFrame.
head()
                                                                             Syntax:

    dataframe_name.head(n)

                                                                             Copied!
                                                                             Example:
```

1. 1

about:blank

```
1. df.head(5)
                                                                       Copied!
                                                                      Syntax:
                                                                         1. import pandas as pd
                                                                      Copied!
Import pandas
                 Imports the Pandas library with the alias pd.
                                                                      Example:
                                                                         1. 1
                                                                         1. import pandas as pd
                                                                       Copied!
                                                                      Syntax:
                                                                         1. 1

    dataframe_name.info()

                                                                      Copied!
                 Provides information about the DataFrame, including
info()
                 data types and memory usage.
                                                                      Example:
                                                                         1. 1
                                                                         1. df.info()
                                                                       Copied!
                                                                      Syntax:
                                                                         1. 1
                                                                         1. merged_df = pd.merge(df1, df2, on=["column1", "column2"])
                                                                       Copied!
                 Merges two DataFrames based on multiple common
merge()
                 columns.
                                                                      Example:
                                                                         1. 1
                                                                         1. merged_df = pd.merge(sales, products, on=["product_id", "category_id"])
                                                                       Copied!
                                                                      Syntax:
                                                                         1. 1

    print(df) # or just type df

                                                                      Copied!
print DataFrame Displays the content of the DataFrame.
                                                                      Example:
                                                                         2. 2

    print(df)
    df

                                                                      Copied!
                                                                      Syntax:
                                                                         1. dataframe_name["column_name"].replace(old_value, new_value, inplace=True)
                                                                       Copied!
replace()
                 Replaces specific values in a column with new values.
                                                                      Example:

    df["status"].replace("In Progress", "Active", inplace=True)

                                                                       Copied!
                 Displays the last n rows of the DataFrame.
tail()
                                                                      Syntax:
                                                                         1. 1
                                                                         1. dataframe_name.tail(n)
                                                                      Copied!
                                                                      Example:
                                                                         1. 1
```

about:blank 4/5

1. df.tail(5)

Copied!

```
Numpy
```

Package/Method Description Syntax and Code Example Syntax: 1. 1 1. import numpy as np Copied! Importing NumPy Imports the NumPy library. Example: 1. 1 1. import numpy as np Copied! Syntax: 1. 1 2. 2 1. array_1d = np.array([list1 values]) # 1D Array 2. array_2d = np.array([[list1 values], [list2 values]]) # 2D Array Copied! Creates a one or multi-dimensional array, np.array() Example: 1. array_1d = np.array([1, 2, 3]) # 1D Array 2. array_2d = np.array([[1, 2], [3, 4]]) # 2D Array Copied! Example: 2. 2 3. 3 - Calculates the mean of array elements - Calculates the sum of array elements Numpy Array Attributes - Finds the minimum value in the array np.mean(array) - Finds the maximum value in the array np.sum(array) - Computes dot product of two arrays np.min(array 4. np.max(array) 5. np.dot(array_1, array_2) Copied!



© IBM Corporation. All rights reserved.

about:blank 5/5