یکی از کتابخانه‌های اصلی پایتون برای آماده‌سازی و پیش‌پردازش داده‌ها، پانداس (Pandas) است.

* این کتابخانه می‌تواند داده‌ها را با بهره‌گیری از ساختارهای Series و DataFrame که ارائه می‌کند،
  + یک series مشابه با آرایه یک‌بُعدی است.
  + ساختار داده دیتافریم مشابه یک جدول است.
* پانداس دارای ابزارهای گوناگونی برای انجام عملیات ورودی/خروجی است و می‌تواند داده‌ها را از فرمت‌های گوناگونی شامل MS Excel ،TSV ،CSV و دیگر موارد بخواند.

نصب پانداس

pip install pandas

نوع ساختار داده:

0 1

1 2

2 3

3 4

dtype: int64

import pandas as pd

series1 = pd.Series([1,2,3,4])

print(series1)

df = pd.DataFrame({

"Column1": [1, 4, 8, 7, 9],

"Column2": ['a', 'column', 'with', 'a', 'string'],

"Column3": [1.23, 23.5, 45.6, 32.1234, 89.453],

"Column4": [True, False, True, False, True]

})

items = [['Phone', 2000], ['TV', 1500], ['Radio', 800]]

Item Price

0 Phone 2000.0

1 TV 1500.0

2 Radio 800.0

df = pd.DataFrame(items, columns=['Item', 'Price'], dtype=float)

print(df)