

Lesson05 ライブラリ

この時間の目的

ライブラリを有効に使う事で、

- ・プログラム外部からのデータの読み込み
- ・配列データの展開
- ・統計量の計算

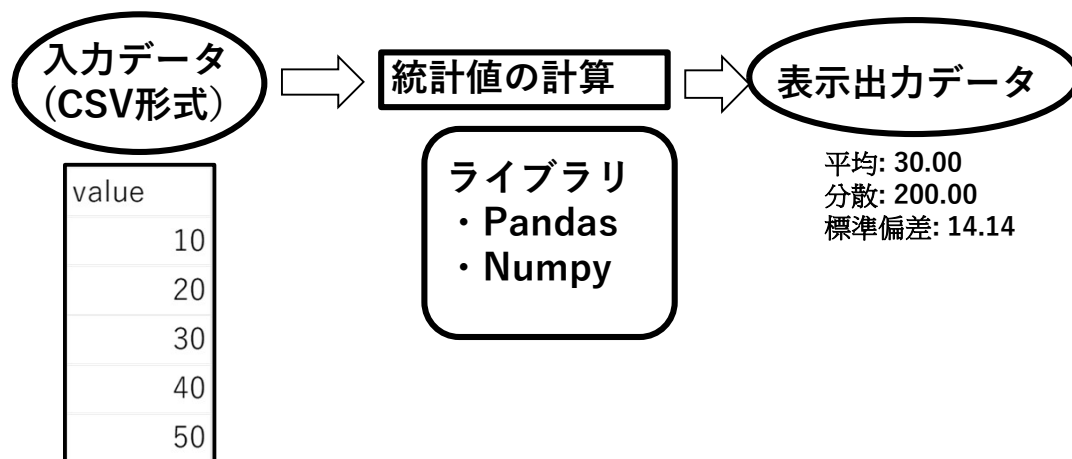
などの一連のデータ処理ができるようになる。

注意！今回実習のプログラムとデータファイルの格納先（フォルダー）はすべて次に変更してください。

C:ドライブ→Winpython64-3.12.4.1→notebooks

1

データの流れ



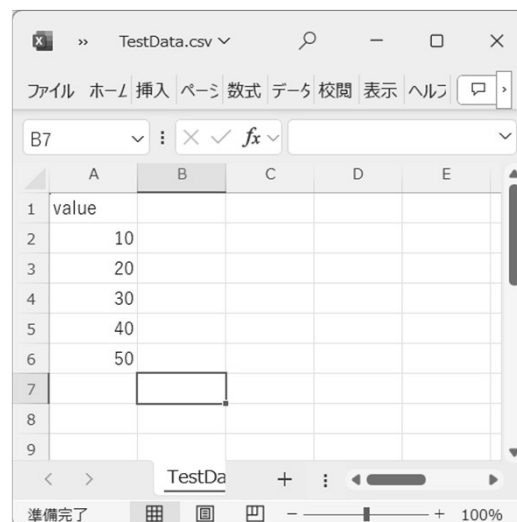
2

CSVとは？ (Wikipedia)

(準備) エクセルでTestData.csv を作しましょう

comma-separated values (略称: CSV) は、テキストデータをいくつかのフィールド (項目) に分け、区切り文字であるカンマ「,」で区切ったデータ形式。拡張子は .csv、MIMEタイプは text/csv。

「comma-separated variables」とも言う。日本語では広く普及した訳語はないが、「カンマ区切り」「コンマ区切り」などとも呼ばれる。Microsoft Excelの日本語版では「CSV (カンマ区切り)」としている



3

```

01 05_csvtest.py
02 import pandas as pd # データ取扱パッケージ
03 import numpy as np # 数値計算パッケージ
04
05 # CSVファイルからデータを読み込みます
06 data = pd.read_csv('TestData.csv')
07
08 # データを配列に変換します
09 # ここでは、データの'value'という列にデータが入っていると仮定しています
10 # それらを配列'values'に変換します
11 values = data['value'].to_numpy()
12
13 # 試しに読み込んだデータ5件を表示してみます
14 print("読込結果:", values[0], values[1], values[2], values[3], values[4]) #test
15
16 mean = np.mean(values) # 平均を計算
17 variance = np.var(values) # 分散を計算
18 std_dev = np.std(values) # 標準偏差を計算
19
20 print(f"平均: {mean:.2f}")
21 print(f"分散: {variance:.2f}")
22 print(f"標準偏差: {std_dev:.2f}")
23

```

4

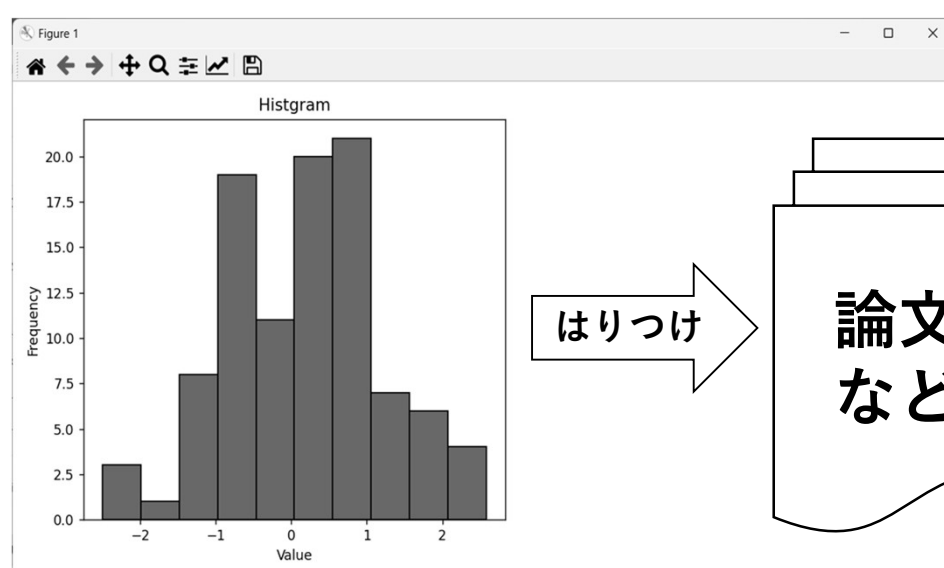
このプログラムは、CSVファイルからデータを読み込み、そのデータの統計量（平均、分散、標準偏差）を計算して表示します。

また、PandasやNumpyなどの強力なライブラリを駆使し、プログラム内部のデータ配列に展開しています。

データの解析や可視化の前処理として非常に基本的で重要な操作を行っています。

5

グラフ表示を追加することも簡単にできます。
（必要に応じて口頭説明します）



6