

# GENEROWANIE WYKRESÓW NA PODSTAWIE DANYCH ODCZYTANYCH Z PLIKÓW

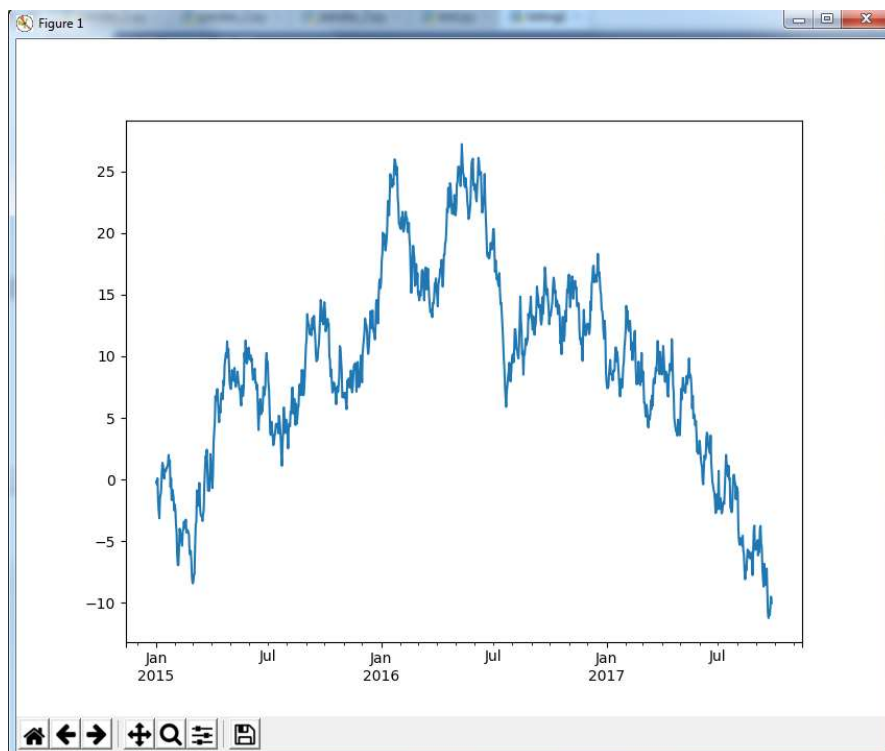
## 1. PRZYKŁADOWE WYKRESY

### Listing 1 – wykres liniowy na podstawie serii danych

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# korzystając z funkcji random oraz date_range możemy wygenerować szereg
# czasowy danych
ts = pd.Series(np.random.randn(1000), index=pd.date_range('1/1/2015',
periods=1000))
# funkcja biblioteki Pandas generująca skumulowaną sumę kolejnych elementów
ts = ts.cumsum()
print(ts)
ts.plot()
plt.show()
```

### Wykres 1

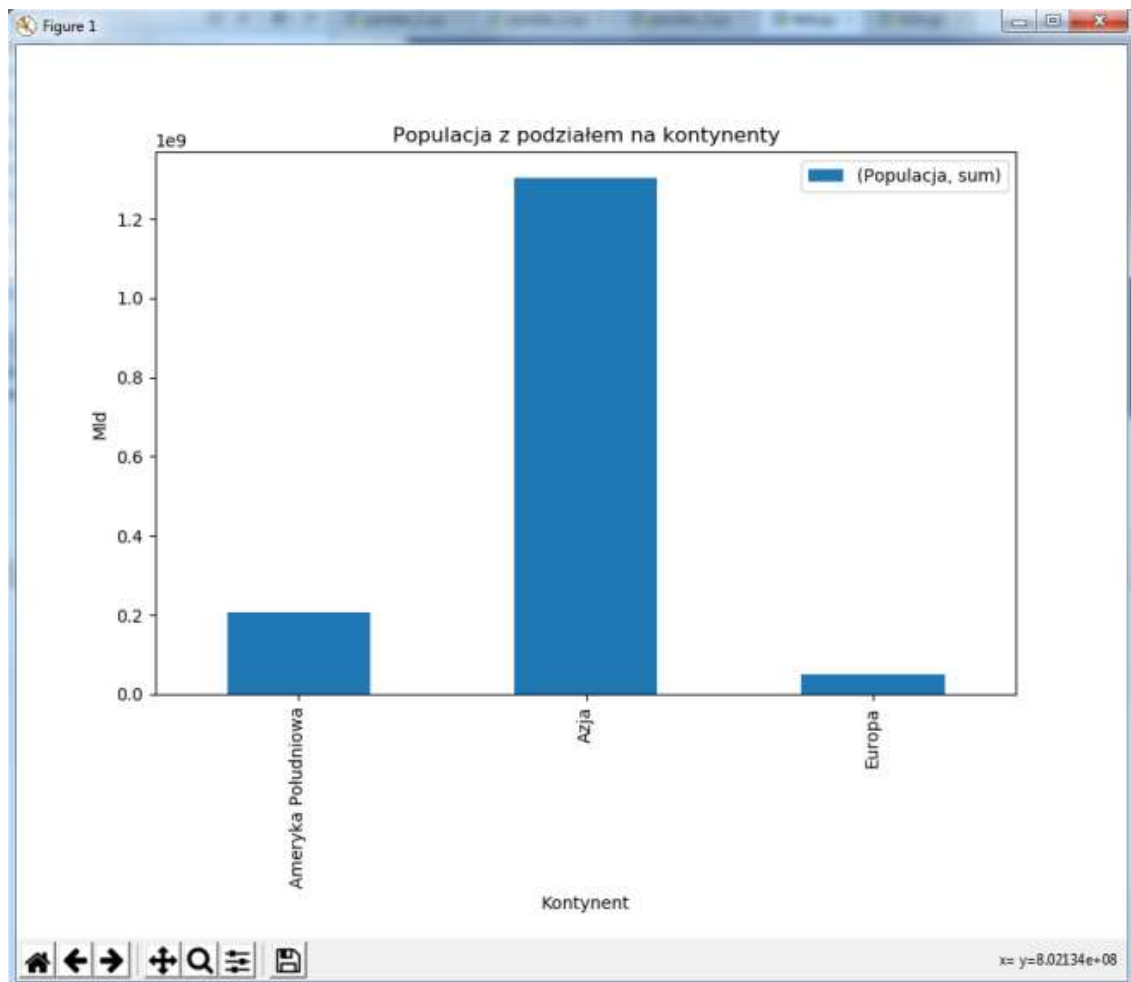


**Listing 2 – wykres kolumnowy z Pandas DataFrame**

```
import pandas as pd
import matplotlib.pyplot as plt

data = {'Kraj': ['Belgia', 'Indie', 'Brazylia', 'Polska'],
        'Stolica': ['Bruksela', 'New Delhi', 'Brasilia', 'Warszawa'],
        'Kontynent': ['Europa', 'Azja', 'Ameryka Południowa', 'Europa'],
        'Populacja': [11190846, 1303171035, 207847528, 38675467]}
df = pd.DataFrame(data, columns=['Kraj', 'Stolica', 'Kontynent',
                                 'Populacja'])
print(df)

grupa = df.groupby(['Kontynent']).agg({'Populacja': ['sum']})
print(grupa)
wykres = grupa.plot.bar()
wykres.set_ylabel('Mld')
wykres.set_xlabel('Kontynent')
wykres.legend()
plt.title('Populacja z podziałem na kontyenty')
plt.show()
```

**Wykres 2**

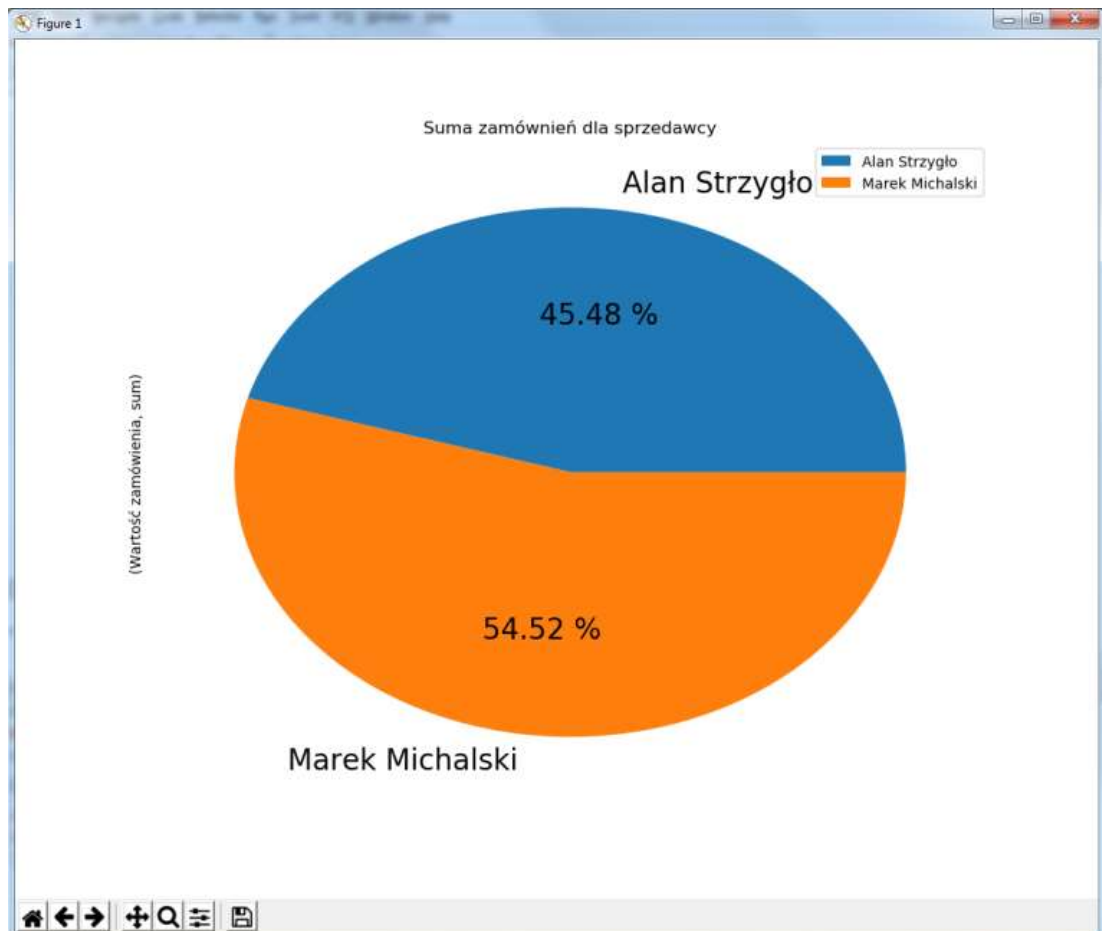
**Listing 3 – wczytanie danych z pliku i wyświetlenie zgrupowanych wartości**

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('dane.csv', delimiter=';')
grupa = df.groupby(['Imię i nazwisko']).agg({'Wartość zamówienia':['sum']})
# wykres kołowy z wartościami procentowymi sformatowanymi z dokładnością do
# 2 miejsc po przecinku
wykres = grupa.plot.pie(subplots=True, autopct='%0.2f %%', fontsize=20,
    figsize=(6, 6))
plt.title('Suma zamówień dla sprzedawcy')
plt.show()
```

zawartość pliku 'dane.csv'

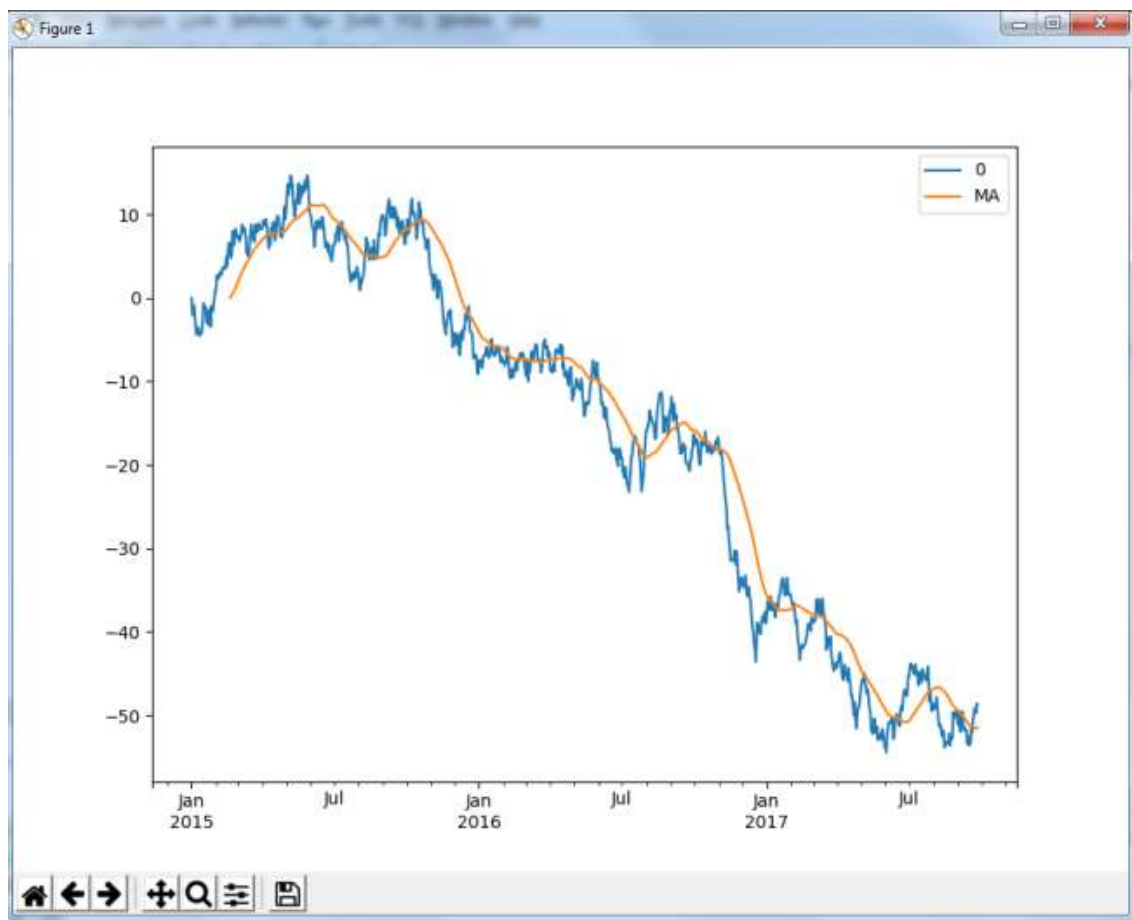
```
Imię i nazwisko;Data;Wartość zamówienia
Marek Michalski;2018-01-10;1200.0
Marek Michalski;2018-01-21;456.5
Alan Strzygło;2018-01-22;350.0
Marek Michalski;2018-01-23;800.0
Alan Strzygło;2018-01-24;349.0
Alan Strzygło;2018-01-26;1350.0
```

**Wykres 3**

**Listing 4 – zmodyfikowana wersja listingu 1 z dodatkowym wykresem średniej kroczącej**

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# korzystając z funkcji random oraz date_range możemy wygenerować szereg
# czasowy danych
ts = pd.Series(np.random.randn(1000), index=pd.date_range('1/1/2015',
periods=1000))
# funkcja biblioteki Pandas generująca skumulowaną sumę kolejnych elementów
ts = ts.cumsum()
# rzutowanie Series na DataFrame
df = pd.DataFrame(ts)
# dodanie nowej kolumny i wykorzystanie funkcji rolling do stworzenia
# kolejnych wartości średniej kroczącej
df['MA'] = df.rolling(window=50).mean()
df.plot()
plt.show()
```

**Wykres 4**

## 2. ZADANIA

Korzystając ze zbiorów danych z lekcji „Biblioteka Pandas” wykonaj poniższe zadania.

### Zadanie 1

Stwórz wykres liniowy, który wyświetli liczbę urodzonych dzieci dla każdego roku.

### Zadanie 2

Stwórz wykres słupkowy, który wyświetli liczbę urodzonych chłopców i dziewczynek z całego zbioru.

### Zadanie 3

Wykres kołowy z wartościami % ukazującymi ilość urodzonych chłopców i dziewczynek w ostatnich 5 latach z datasetu.

### Zadanie 4

Z repozytorium UCI (<http://archive.ics.uci.edu/ml/index.php>) pobierz dataset Iris i za pomocą wykresu punktowego (scattered) wyświetl wartość 2 wybranych cech tego datasetu. Dla każdego rodzaju kwiatu użyj innego koloru na wykresie. Przykład można znaleźć w galerii wykresów biblioteki matplotlib - link <https://matplotlib.org/stable/gallery/index.html>

### Zadanie 5

Wyświetl na pomocą wykresu słupkowego ilość złożonych zamówień przez poszczególnych sprzedawców (zbiór danych zamówienia.csv).