

Biblioteka Pandas, część 2.

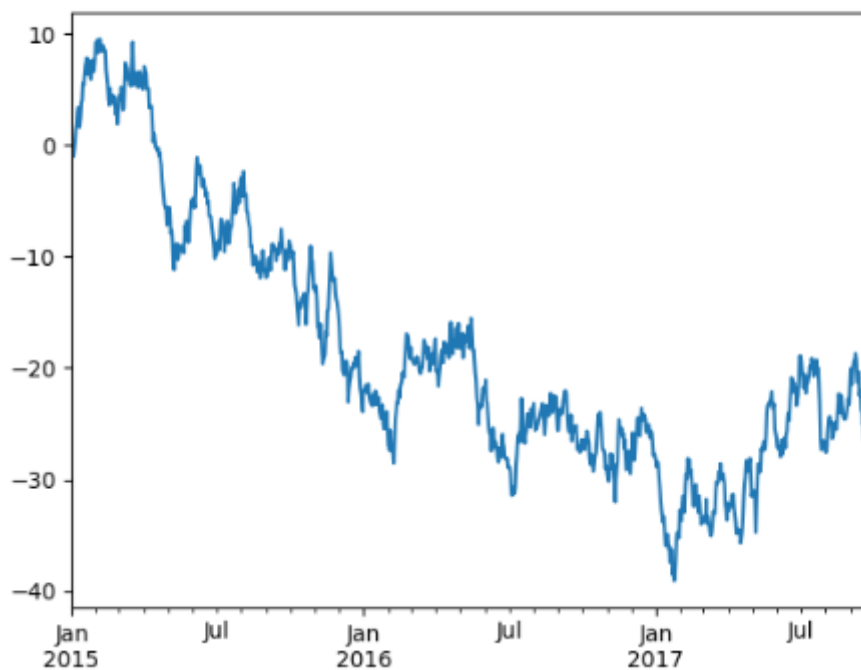
1. Pandaas i wykresy.

Listing 1 – wykres liniowy na podstawie serii danych

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

#korzystając z funkcji random oraz data_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



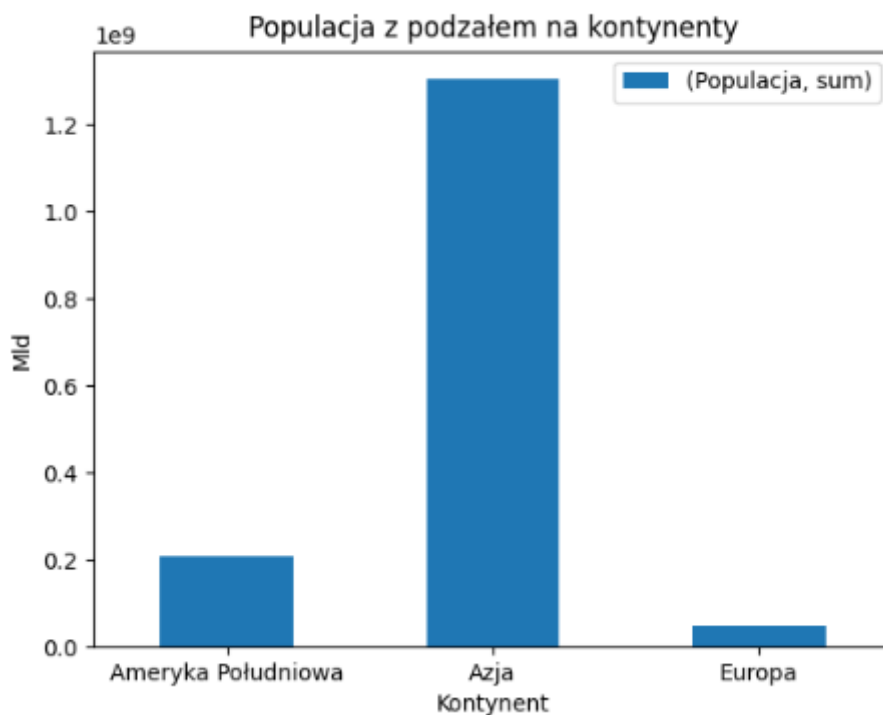
Listing 2 – wykres kolumnowy z Pandas DataFrame

```
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)
print(df)

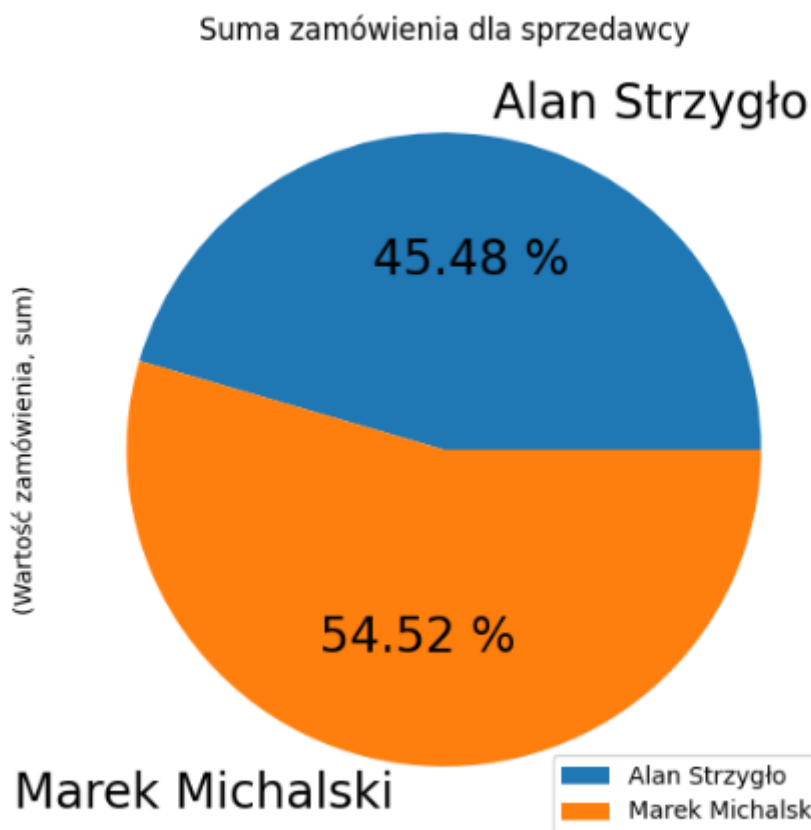
grupa = df.groupby(['Kontynent']).agg({'Populacja': ['sum']})
print(grupa)
wykres = grupa.plot.bar()
wykres.set_ylabel("Mld")
wykres.set_xlabel('Kontynent')
wykres.legend()
plt.title('Populacja z podziałem na kontynenty')
#zmiana kierunku tekstu etykiet słupków
plt.xticks(rotation=0)
plt.savefig('wykres.png')
plt.show()
```



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', header=0, sep=";", decimal=".")
print(df)
grupa = df.groupby(['Imię i nazwisko']).agg({'Wartość zamówienia': ["sum"]})
#wykres kolumnowy z wartościami procentowymi sformatowanymi z
dokładnością do 2 miejsc po przecinku
#figsize ustawia wielkość wykresu w calach, domyślnie [6.4,
4.8]
wykres = grupa.plot.pie(subplots=True, autopct='%0.2f %%',
fontSize=20, figsize=(6,6), legend=(0, 0))
plt.legend(loc="lower right")
plt.title('Suma zamówienia dla sprzedawcy')
plt.show()
```



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

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

#korzystając z funkcji random oraz data_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()
```

