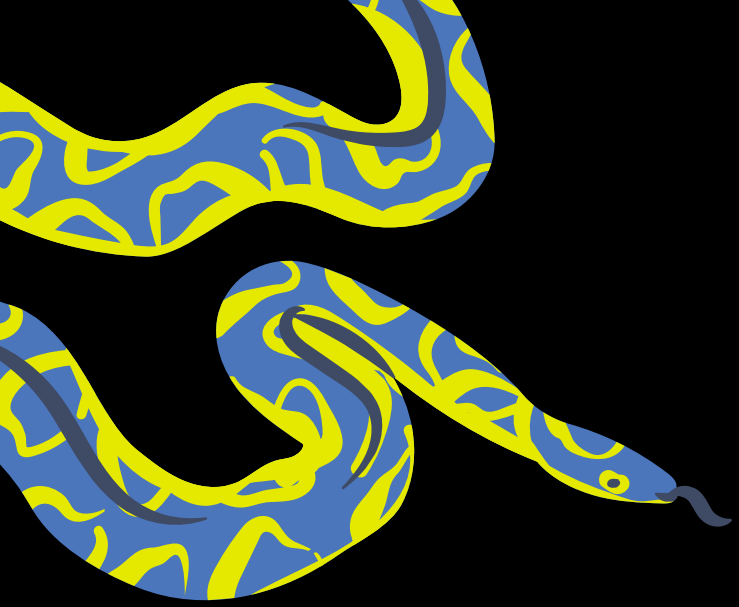
SciByteHub



Python

04.06

SciByteHub

RECAP - import

Importowanie bibliotek do kodu:

```
import biblioteka
```

Importowanie biblioteki jako alias:

```
import biblioteka as alias
```

SciByteHub

RECAP - pandas

Wczytanie pliku .csv:

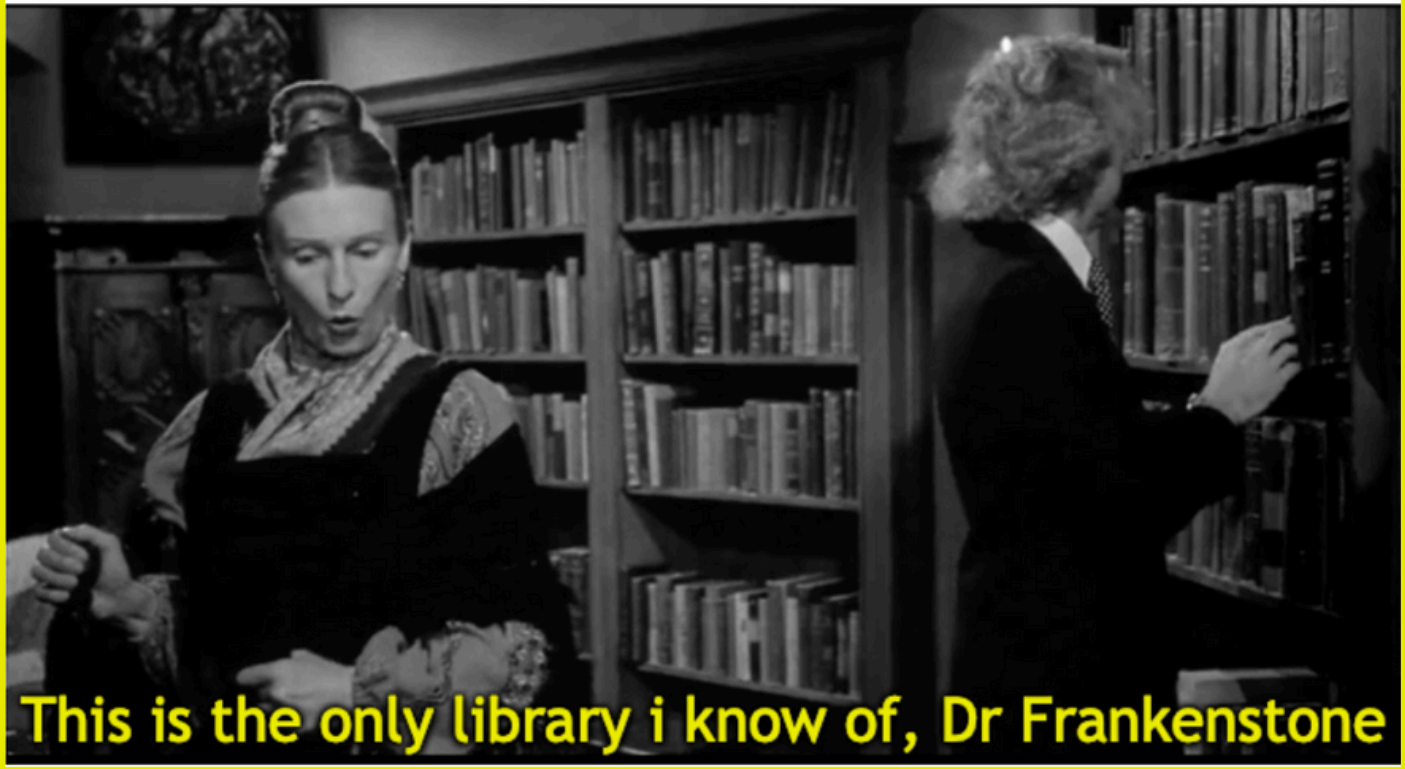
```
data = pd.read_csv('nazwa_pliku.csv', header=None)
```

Tworzenie ramki danych:

```
df = pd.DataFrame(dataset)
```

SciByteHub

When you ask a Data Scientist why he is importing Pandas to read a text file



SciByteHub

Pliki .txt

```
with open('data.txt', 'r', encoding='utf-8') as file:
    data = file.read()
```

SciByteHub

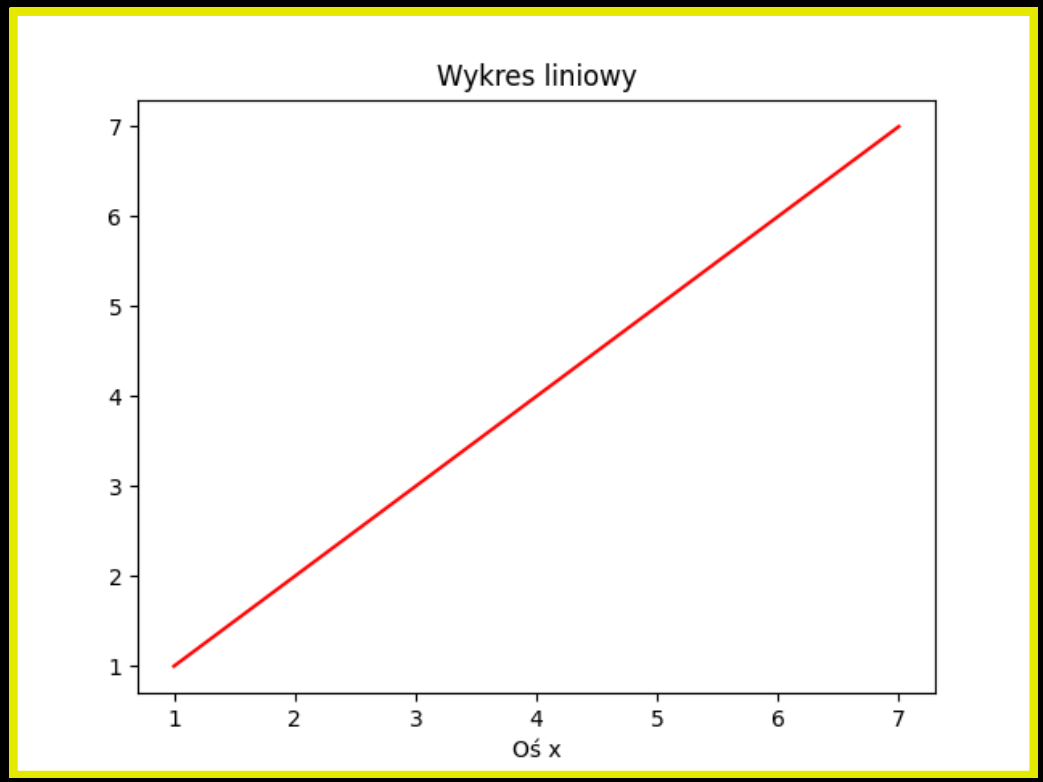
Biblioteki do wykresów

```
import matplotlib.pyplot as plt

import seaborn as sns
```

SciByteHub

Wykres liniowy

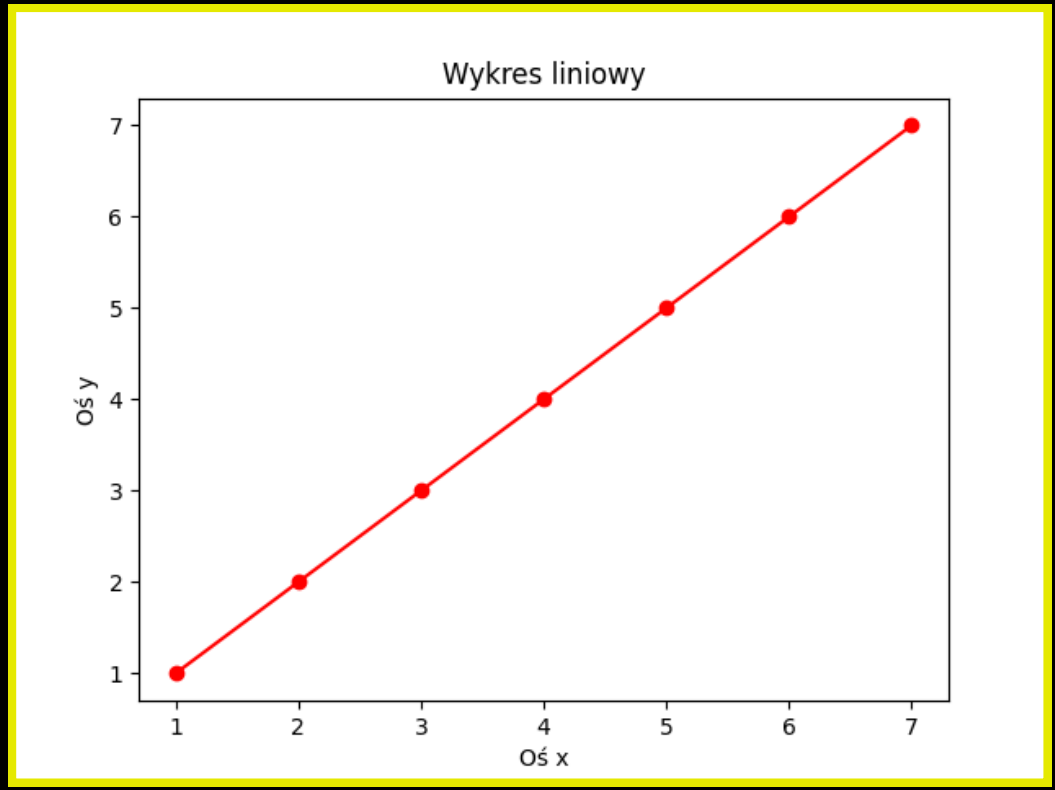
```
import matplotlib.pyplot as plt
x=[1,2,3,4,5,6,7]
y=[1,2,3,4,5,6,7]
plt.plot(x, y, color='red')
plt.title('Wykres liniowy')
plt.xlabel('Oś x')
plt.ylabel('Oś y')
plt.show()
```



SciByteHub

Wykres liniowy

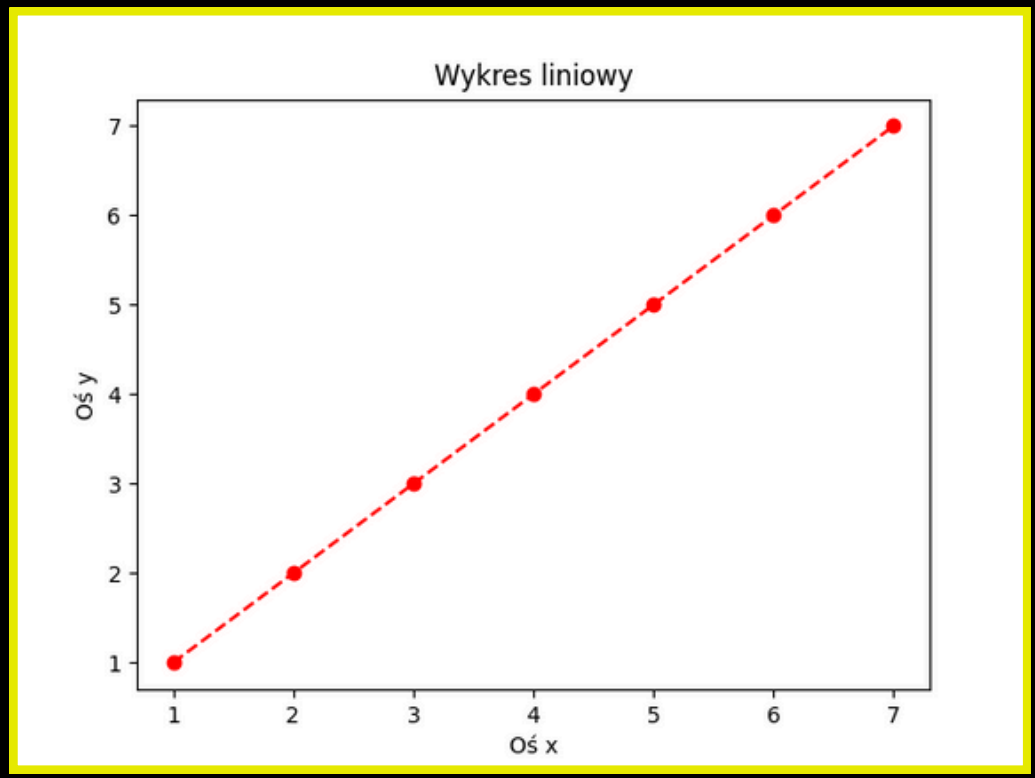
```
import matplotlib.pyplot as plt
x=[1,2,3,4,5,6,7]
y=[1,2,3,4,5,6,7]
plt.plot(x, y, color='red',
marker='o')
plt.title('Wykres liniowy')
plt.xlabel('Oś x')
plt.ylabel('Oś y')
plt.show()
```



SciByteHub

Wykres liniowy

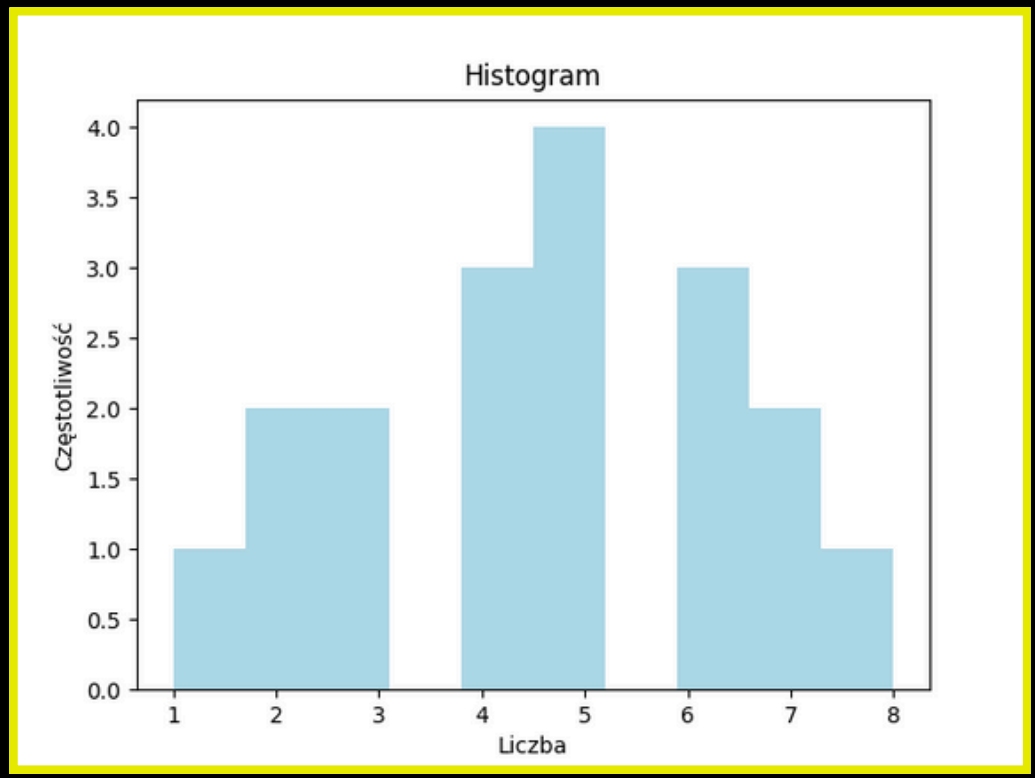
```
import matplotlib.pyplot as plt
x=[1,2,3,4,5,6,7]
y=[1,2,3,4,5,6,7]
plt.plot(x, y, color='red',
marker='o', linestyle='--')
plt.title('Wykres liniowy')
plt.xlabel('Oś x')
plt.ylabel('Oś y')
plt.show()
```



SciByteHub

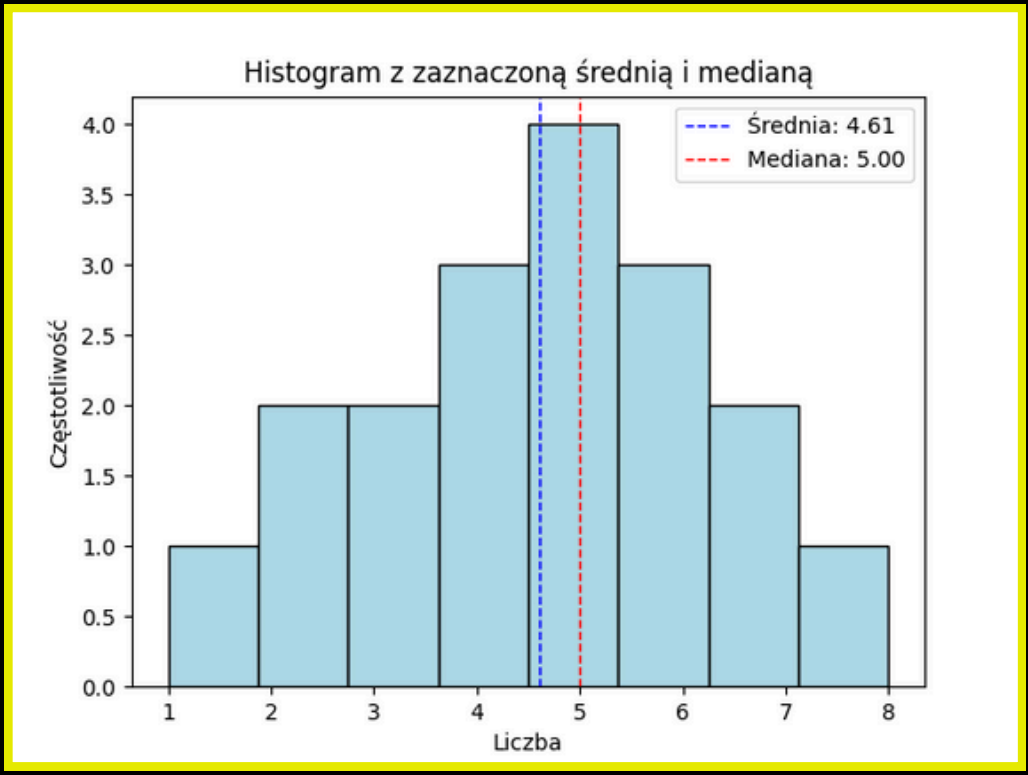
Histogram

```
import matplotlib.pyplot as plt
x=[1,2,2,3,3,4,4,4,5,5,
5,5,6,6,6,7,7,8]
plt.hist(x, color='lightblue')
plt.title('Histogram')
plt.xlabel('Liczba')
plt.ylabel('Częstotliwość')
plt.show()
```



SciByteHub

Histogram

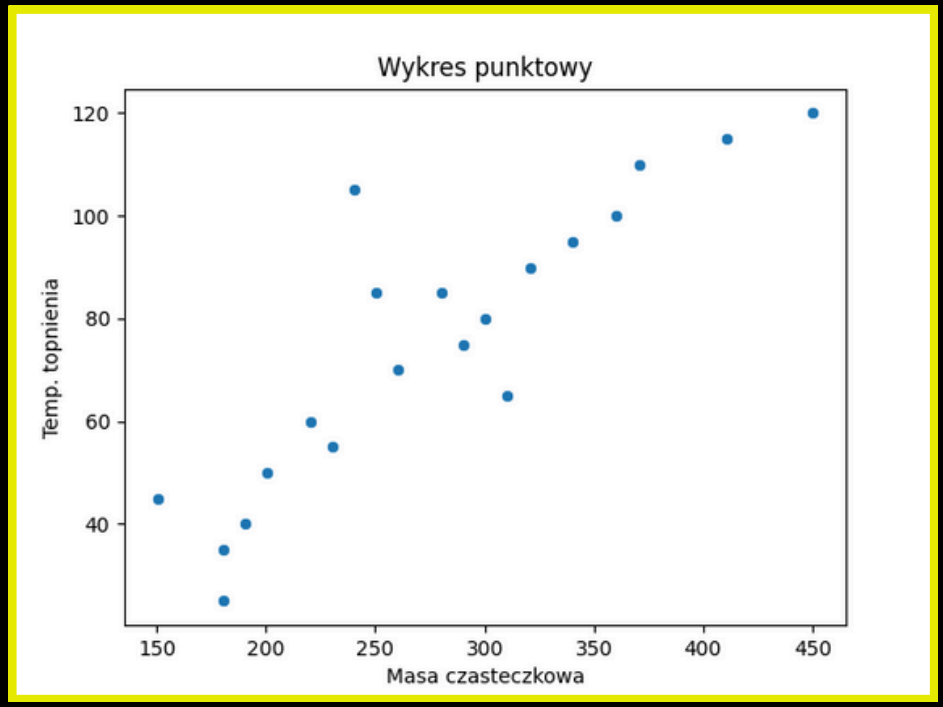



```
plt.hist(x, color='lightblue',
bins=8, edgecolor=’black’)
plt.axvline(mean,
color='blue', linestyle='--',
linewidth=1,
label=f'Średnia: {mean_x:.2f}')
plt.axvline(median,
color='red', linestyle='--',
linewidth=1,
label=f'Mediana: {median_x:.2f}')
plt.title('Histogram z
zaznaczoną średnią i medianą')
plt.xlabel('Liczba')
plt.ylabel('Częstotliwość')
plt.show()
```

SciByteHub

Wykres punktowy

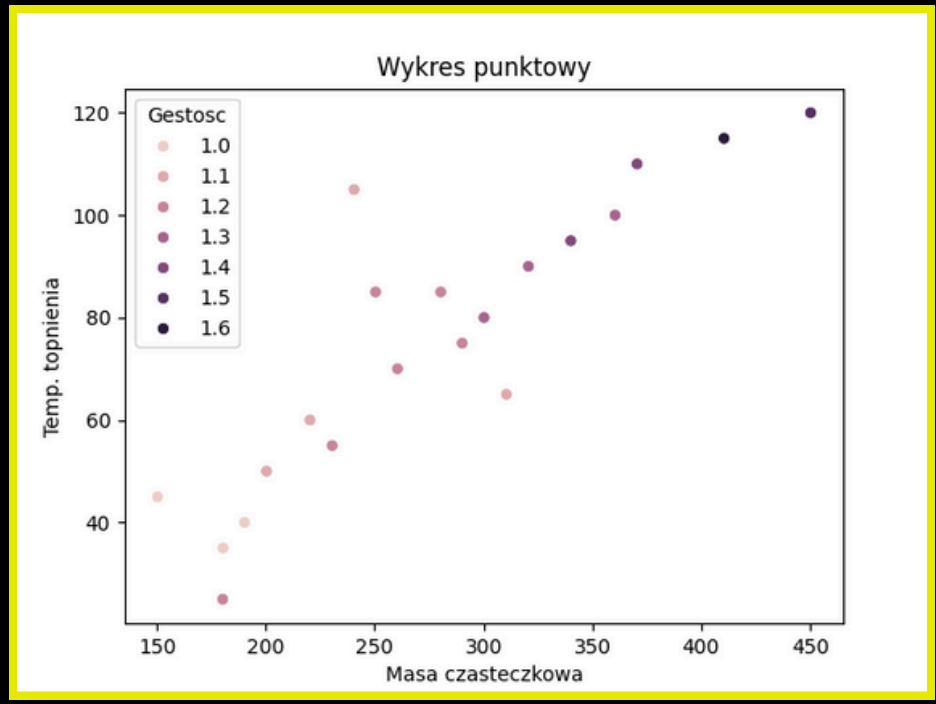
```
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv(“file.csv”)

sns.scatterplot(df,x='MW',y=’MP’)
plt.title('Wykres punktowy')
plt.xlabel('Masa czasteczkowa')
plt.ylabel('Temp. topnienia')
plt.show()
```



SciByteHub

Wykres punktowy

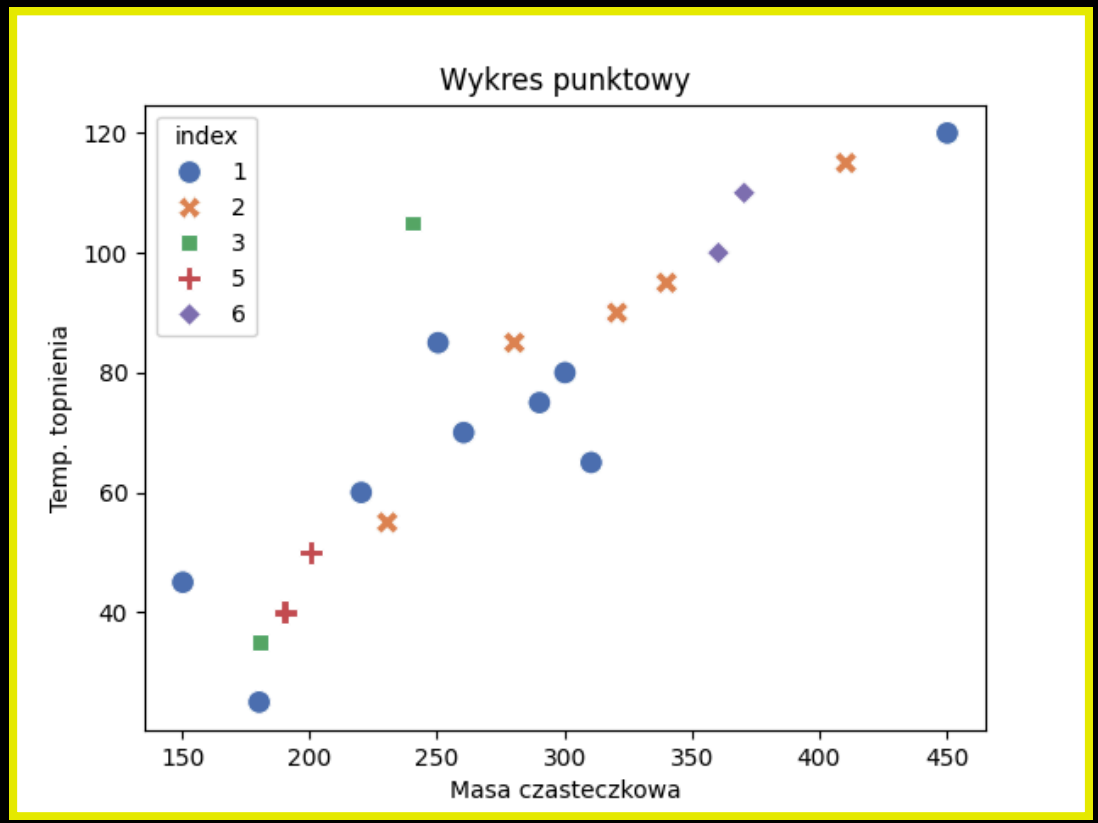
```
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv(“file.csv”)

sns.scatterplot(df,x='MW',y=’MP’,
hue=’Gestosc’, legend=”full”)
plt.title('Wykres punktowy')
plt.xlabel('Masa czasteczkowa')
plt.ylabel('Temp. topnienia')
plt.show()
```



SciByteHub

Wykres punktowy

```
sns.scatterplot(df,x='MW',y='MP',hue='index',style='index',
legend='full', palette='deep', s=100)

plt.title('Wykres
 punktowy')
plt.xlabel('Masa
 czasteczkowa')
plt.ylabel('Temp.
 topnienia')

plt.show()
```



Heatmap

df.iloc[:,1:8]

	Molecular_Weight	Boiling_Point	Melting_Point	Solubility_Water	pKa	LogP	Density
0	180.16	100	25	0.050	4.75	2.1	1.2
1	300.22	150	80	0.010	3.15	4.5	1.3
2	450.30	250	120	0.002	7.20	5.7	1.5
3	220.25	120	60	0.080	6.50	3.8	1.1
4	150.20	110	45	0.030	5.80	3.0	1.0
5	340.10	200	95	0.004	3.90	4.1	1.4
6	410.50	220	115	0.010	4.50	5.0	1.6
7	280.34	180	85	0.060	6.00	3.5	1.2
8	200.30	130	50	0.040	4.20	2.8	1.1

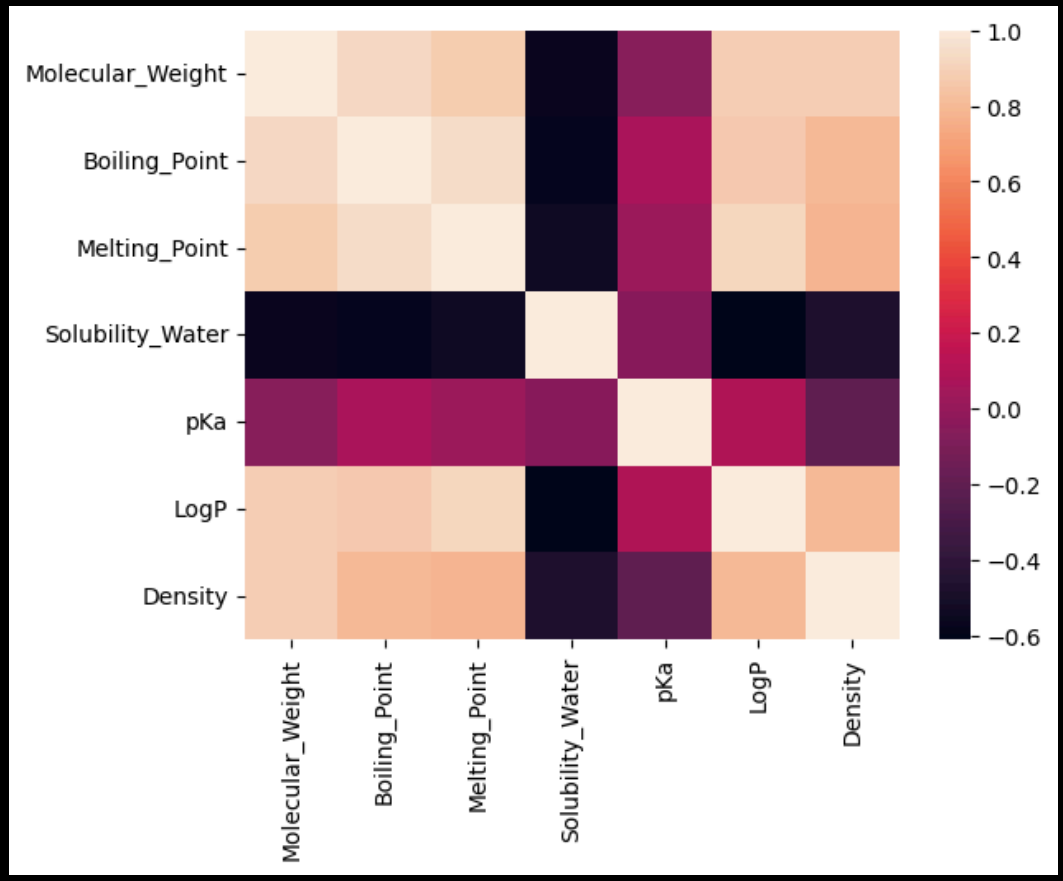
df.iloc[:,1:8].corr()

	Molecular_Weight	Boiling_Point	Melting_Point	Solubility_Water	pKa	LogP	Density
Molecular_Weight	1.000000	0.930384	0.879701	-0.566668	-0.054103	0.888637	0.891415
Boiling_Point	0.930384	1.000000	0.938799	-0.579716	0.063836	0.865053	0.792812
Melting_Point	0.879701	0.938799	1.000000	-0.543586	0.014824	0.912810	0.779392
Solubility_Water	-0.566668	-0.579716	-0.543586	1.000000	-0.046570	-0.611935	-0.465506
pKa	-0.054103	0.063836	0.014824	-0.046570	1.000000	0.083548	-0.197807
LogP	0.888637	0.865053	0.912810	-0.611935	0.083548	1.000000	0.796161
Density	0.891415	0.792812	0.779392	-0.465506	-0.197807	0.796161	1.000000

SciByteHub

Heatmap

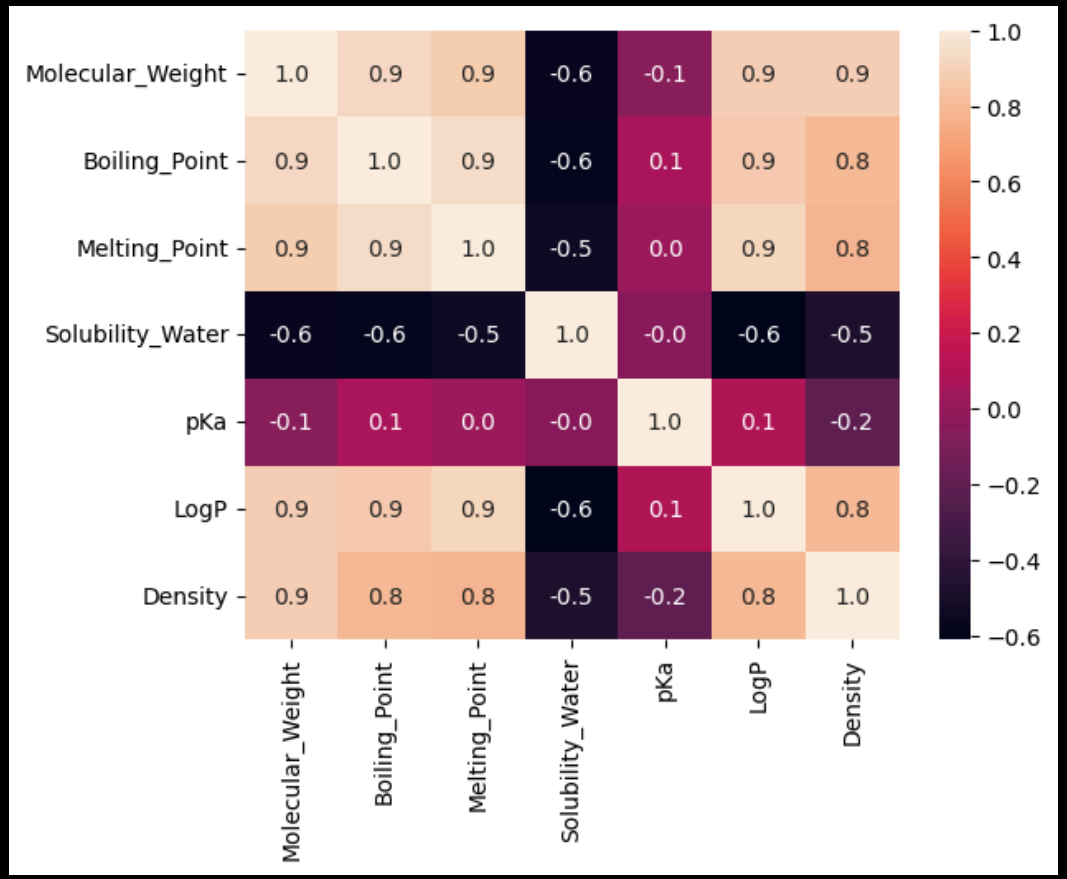
```
c = df.iloc[:,1:8]
sns.heatmap(c)

plt.show()
```



SciByteHub

Heatmap

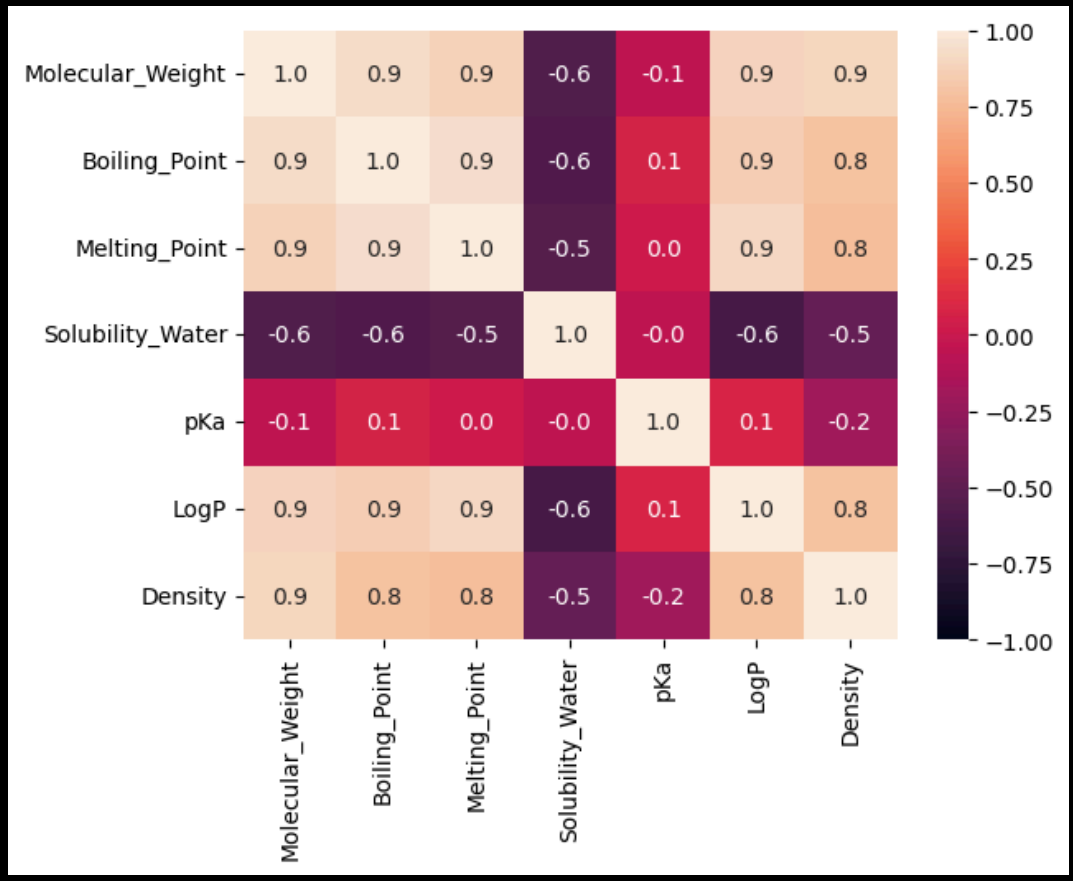
```
c = df.iloc[:,1:8]
sns.heatmap(c,
 annot=True, fmt=".1f")

plt.show()
```



SciByteHub

Heatmap

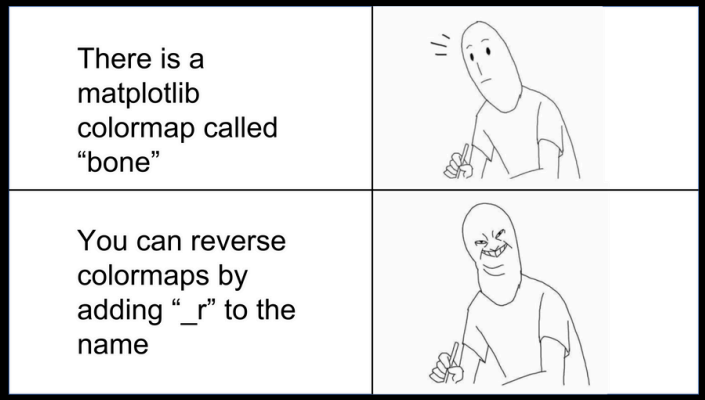
```
c = df.iloc[:,1:8]
sns.heatmap(c,
 annot=True,fmt=".1f",
 vmin=-1, vmax=1)

plt.show()
```



SciByteHub

Heatmap

```
c = df.iloc[:,1:8]
sns.heatmap(c,
 annot=True,fmt=".1f",
 vmin=-1, vmax=1,
 cmap = "crest")

plt.show()
```



SciByteHub

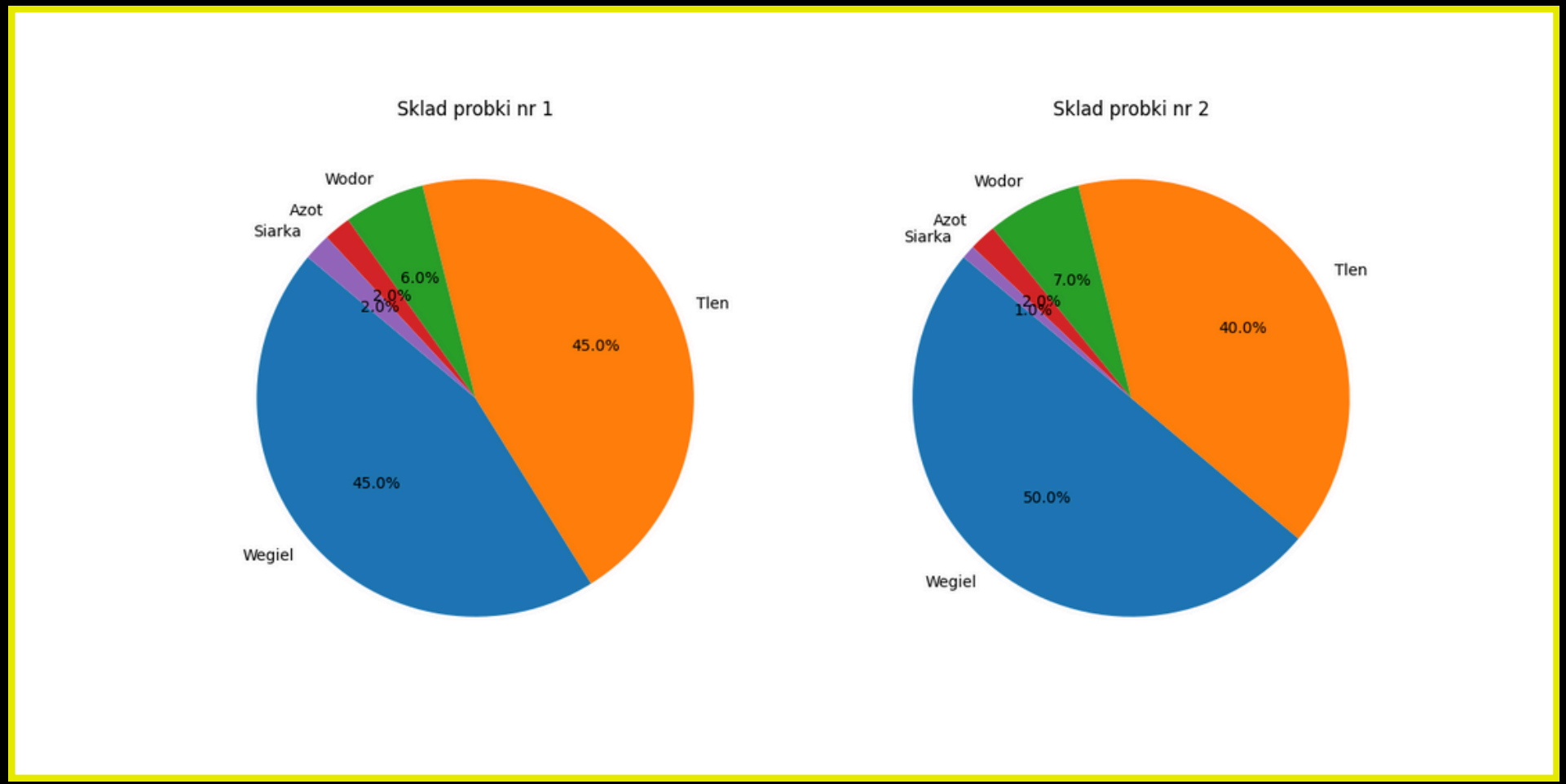
Wykres kołowy

```
fig,axes=plt.subplots(1,2,figsize=(14, 7))

axes[0].pie(data1['Procent'],labels=data1['Element'],
autopct='%1.1f%%', startangle=140)
axes[0].set_title('Sklad probki nr 1')
axes[1].pie(data2['Procent'],labels=data2['Element'],
autopct='%1.1f%%', startangle=140)
axes[1].set_title('Sklad probki nr 2')

plt.show()
```

SciByteHub

Wykres kołowy

```
fig,axes=plt.subplots(1,2,figsize=(14, 7))
```



SciByteHub

Zapisanie wykresu do pliku

```
plt.savefig('wykres.png')
```

SciByteHub