# Data analytics in Python Programming

## Week 10A:Data visualisation

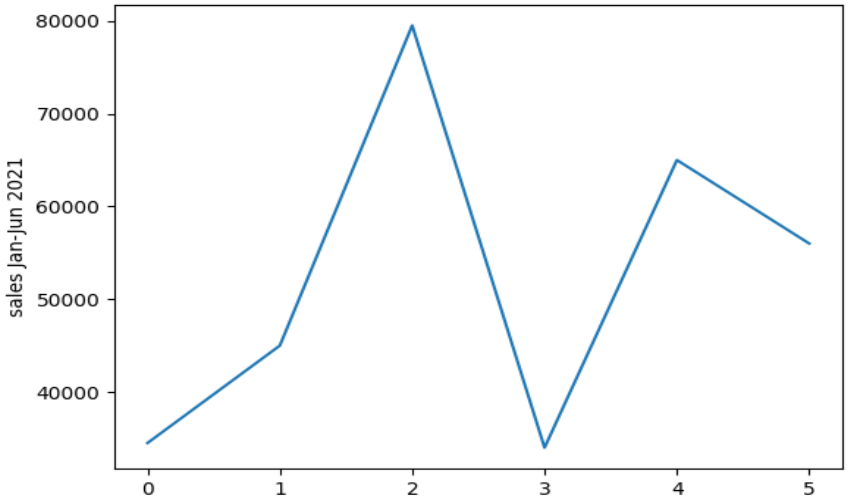
# Matplotlib: https://matplotlib.org/

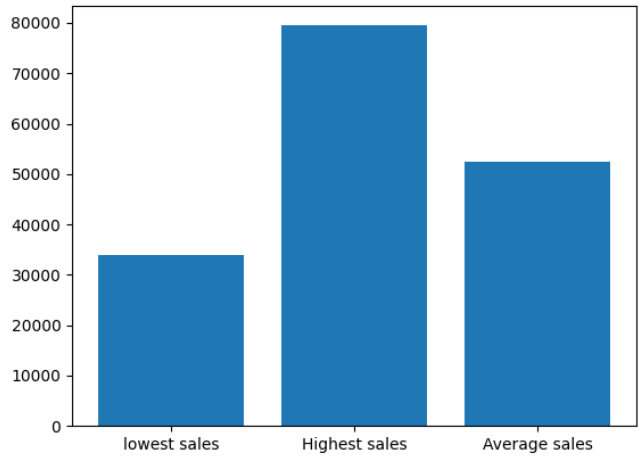
It is a cross-platform, data visualization and graphical plotting library for Python and its numerical extension NumPy.

| Month | Sales |
|---|---|
| 1 | 34500.00 |
| 2 | 45000.00 |
| 3 | 79500.00 |
| 4 | 34000.00 |
| 5 | 65000.00 |
| 6 | 56000.00 |

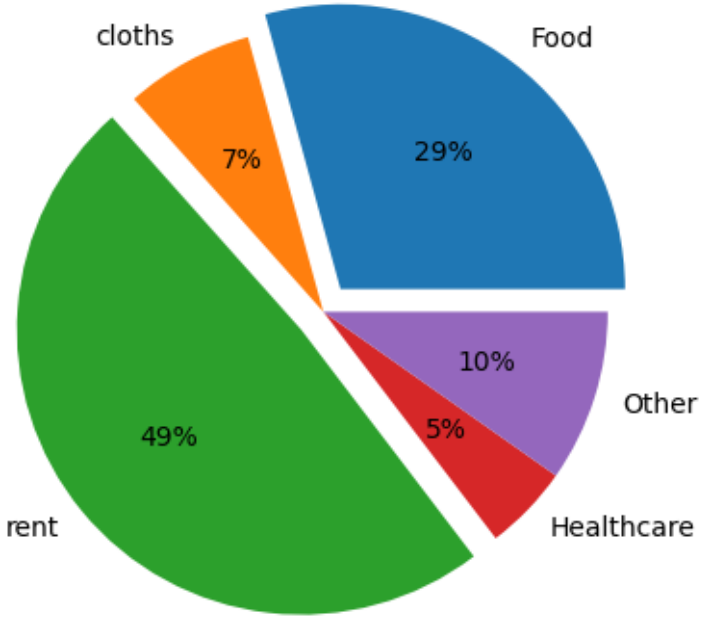
<untitled> * | numpymat.py *

```
plots some lines in a plotting area,
decorates the plot with labels"""
import numpy as np
import matplotlib.pyplot as mp
sales = np.array([34500.00,45000.00,79500.00,34000.00,65000.00,56000.00])
mp.plot(sales)
mp.ylabel("sales Jan-Jun 2021")
mp.show()
lowest = np.amin(sales)
highest = np.amax(sales)
avSales = np.average(sales)
mp.bar(["lowest sales","Highest sales", "Average sales"],[lowest,highest,avSales])
mp.show()
```



```python
import numpy as np
import matplotlib.pyplot as mp
xnames = ["Food","cloths","rent","Healthcare","Other"]
ypoints = np.array([300.00,75.00,500.00,50.00,100.00])
myexplode = [0.1, 0, 0.1, 0, 0] #splitting from pie
#mycolors = ["black", "hotpink", "b", "#4CAF50","red"]
mp.pie(ypoints, labels = xnames, explode = myexplode,autopct='%1.0f%%')
#mp.legend() #adding legends
mp.show()
```



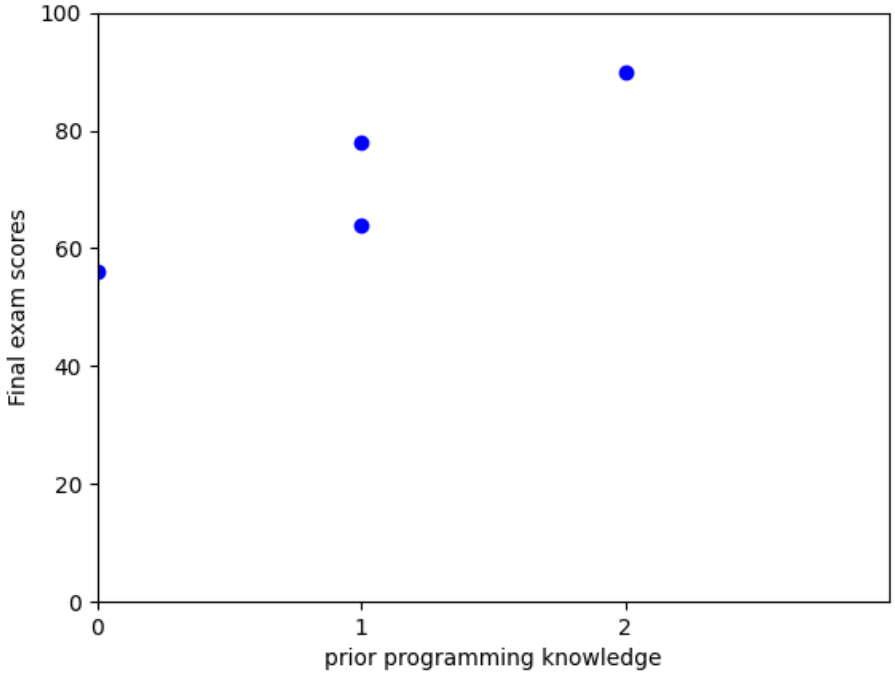

**https://www.w3schools.com/python/matplotlib_pie_charts.asp**

| Id | PPK | FE |
|---|---|---|
| 1 | 0 | 72 |
| 2 | 1 | 69 |
| 3 | 2 | 95 |
| .. | .. | .. |
| .. | .. | .. |

Student Id, prior programming knowledge level(PPK) and final exam scores(FE)

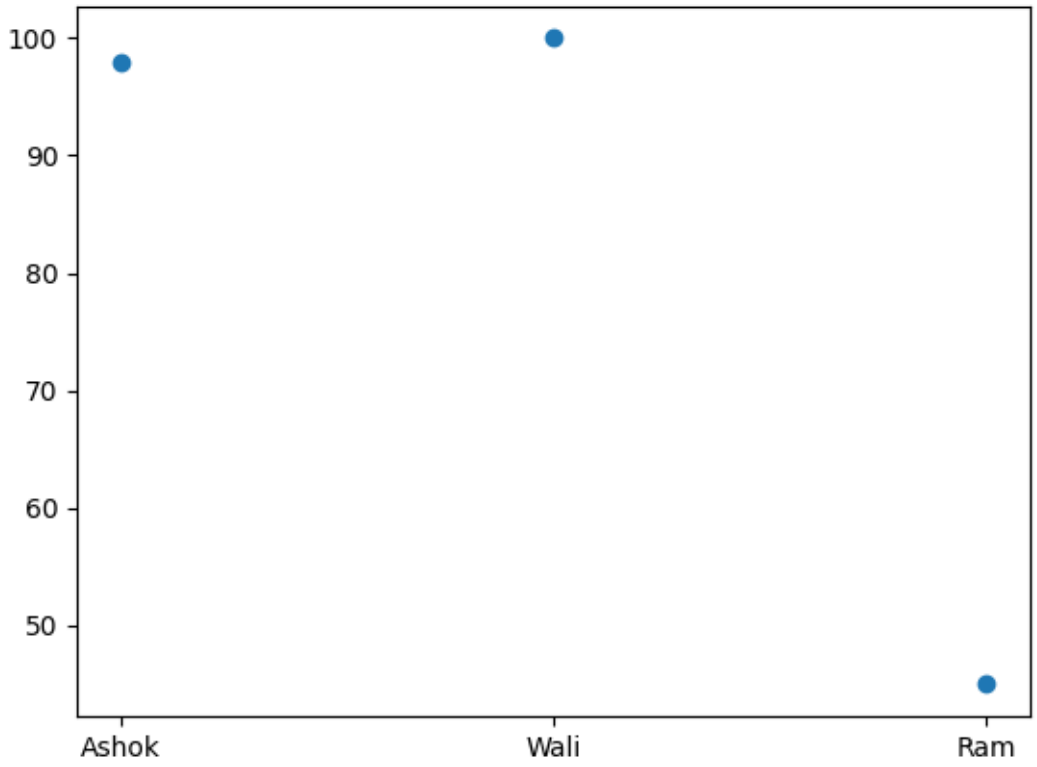
sales.py numpymat.py

```
1 import matplotlib.pyplot as mp
2 mp.plot(["0", "1", "2", "1"], [56, 78, 90, 64], 'bo') # 'bo'- blue circle
3 mp.axis([0, 3,0,100]) # x axis min, max y axis min, max
4 mp.ylabel("Final exam scores")
5 mp.xlabel("prior programming knowledge")
6 mp.show()
7
```



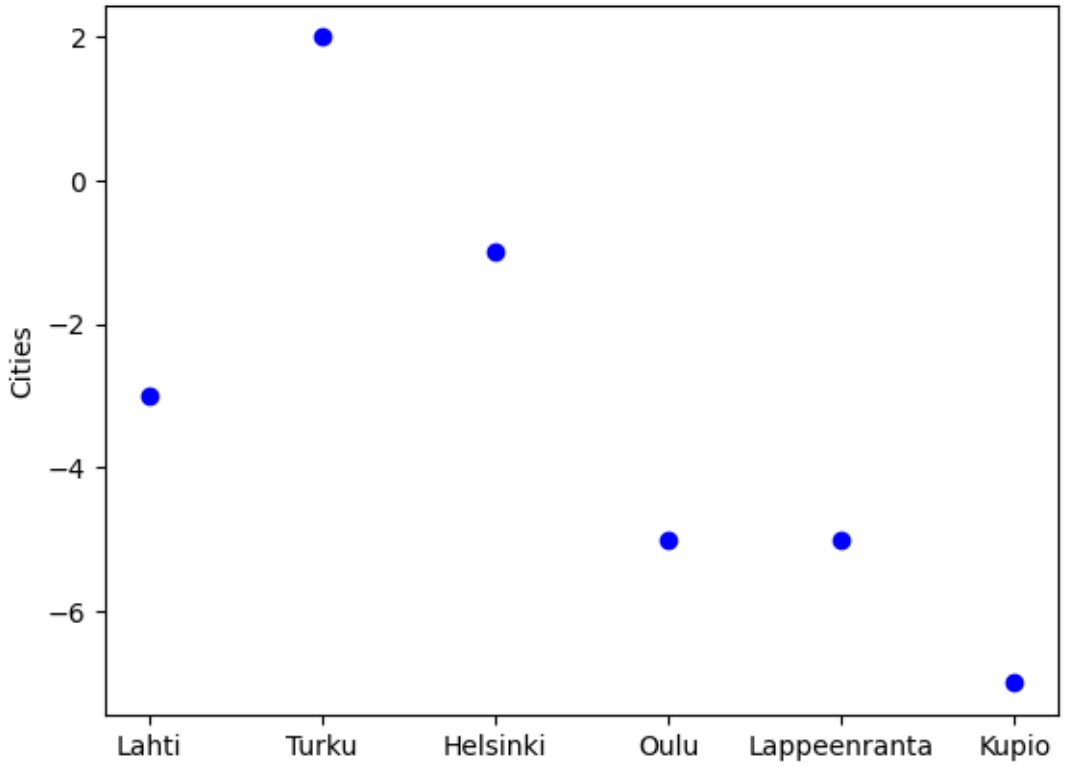
numpymat4.py × Q4sample.py ×

```
1  import matplotlib.pyplot as plt # library to plot graph
2  import numpy as np # it should be installed first
3  xpoints = np.array(["Ashok","Wali","Ram"]) # x points
4  ypoints = np.array([98,100,45]) # y points
5
6  plt.plot(xpoints, ypoints,'o') # o means filled circle
7  plt.show() # show the graph
8
```



numpymat8.py

```
1  import numpy as np
2  import matplotlib.pyplot as mp
3  #mp.style.use('classic')
4  f1 = open("weathercity.txt")
5  city ={}
6  for cities in f1:
7      c = cities.split(",")
8      city.update({c[0]:int(c[1].strip())})
9
10 xpoints = city.keys()
11 ypoints = city.values()
12 mp.plot(xpoints,ypoints, 'bo') # 'bo'- blue circle
13 mp.ylabel("Cities")
14 mp.show()
```



# Data visualization : from text files

numpymat7scores.py *

```
1  import numpy as np
2  import matplotlib.pyplot as mp
3  f1 = open("FEscores.txt")
4  list1 = list()
5  for i in f1:
6      list1.append(int(i.strip()))
7  print(list1)
8  mp.plot(list1, 'bo') # 'bo'- blue circle
9  mp.ylabel("Final exam scores")
10 #mp.xlabel("prior programming knowledge")
11 mp.show()
```

Ok. **How to report data from xls/csv or other type of files? And statistical calculations**

Here we go with Pandas