ETL & visualization

2022-23 Practical 1, Module installation and warm-up code.

a) Installation of necessary modules using pip:

Install the following modules on your system: pandas, numpy, matplotlib, plotly, dash the rest we will install when necessary

b) Test matplotlib with an example:

Write the following code in a python file:

```
import matplotlib.pyplot as plt
a='single','married','complicated','don't know'
sizes=[40,30,15,20]
explode = (0, 0.1, 0, 0) #get that slice out
fig1,ax1=plt.subplots()
ax1.pie(sizes,explode = explode,labels=a,autopct='%1.1f%%',shadow=True,startangle=90)
ax1.axis('equal')
plt.show()
```

Run your code, a pie chart should appear.

Back to your code, eliminate the explode option from the ax1.pie(…)

Run again.

Eliminate shadow=True, check

Put back the explodes.

c) Let's try to make a graph reading from a file:

As you can see in the “sizes” is a python list of numbers. And ”labels” is a series of strings.

Now imagine that we have a text file containing the labels on the first line and the sizes on the second line. Strings and numbers are all separated by a coma with no space.

Write the program to read the data from the file and create the graph.

d) There is a graph called the ring or donut. Google how to turn your pie chart into a donut

e) Place the text of this practical sheet in a notepad text file and save it as text in uft-8 format with a new name.

Write a program that will display occurrences of the words, “of”,'the”,”to”,”that”, from the file with a donut chart. To write the occurrence program, use the classic version by dictionary. If you do not remember this code, google it!

f) Play with the graph to understand well several possibilities, try to put an index for example or put a name for graph.