Data science project report

Description of the project

a. What will the program do?

The program is splitting the customers to (n) groups according to the sum of total spending and their ages and printing a table displaying each customer name, age, total and the computed cluster number and generating association rules between items with minimum support and confidence taken from the user inputs.

b. What the input to the program will be

1. The number of points

2. The number of cluster

3. The min support

4. The min confidence

5. The dataset sheet

c. What the output from the program will be

1. The kmeans with elbow method

2. The clustring graph

3. The association graph

d.The full description of your dataset.

It’s a grocery store dataset has list of items with its count and rnd and it has a list of the customers names ,ages , cities and the way they pay (cash or credit)

e. Screenshots from your Project steps.

Text

Description automatically generated

Some imports to the libraries that I will use later in the project

Text

Description automatically generated

Asking the user to write the dataset file name and setting the data

Text

Description automatically generated

The kmeans method all it needs from the user the number of points

Text

Description automatically generated

The clustering method and its takes the number of points and the umber of clusters from the user

Text

Description automatically generated

The association rule that needs the min support and min confidence from the user and the end of the methods

Text

Description automatically generated

The front for the user that will ask the user which method he wants to run and it will keep looping till the user input number 4 that means to end the program .

f. Explain your results and insight by describing your plotted graphs.

Chart, line chart

Description automatically generated

Kmeans graph by the elbow method

Chart

Description automatically generated

The clustring graph

g. Discussing the Code (libraries used + attributes)

.Pandas library to read the data file

.Matplotlib.pyplot library to make the elbow graph for the kmeans and the clustering graph

.Sklearn.cluster library to set the kmeans clusters

.Mlxtend library to use the apriori and transactionEncoder to use them in the association rule