

Final Project

The main goal for this project is to guide you to practice and apply what you have learned in a real-world task.

Rules:

1. Cheating is strictly prohibited, and if a match is found between any two groups, strict measures will be taken with both parties.
2. Each group should consist of 2 students as minimum to 6 students as maximum.
 - The group will be assigned to a certain number that will be announced on teams.
 - Discussion of the projects will be announced soon.
3. Each group must prepare a pdf report called "Project Report + Group No.", this report must contain the following:
 - Students' names, IDs, and Group number.
 - Explain the problem and briefly describe the role of each member. Note that: the problem description must answer the following questions:
 - a. What will the program do?
 - b. What the input to the program will be?
 - c. What the output from the program will be?
 - The full description of your dataset.
 - Screenshots from your Project steps.
 - Explain your results and insight by describing your plotted graphs.
 - Discussing every part in the Code (libraries used + attributes)
 - ✓ (Screenshot for code parts + Describing what it does)

Project Details:

We have attached Grocery (GRC) dataset that you can download from the project file.

Using this dataset, you are asked to design your software that includes the following items:

1. A user interface that allows customers to enter the required data and see results through out.
2. Use (R) to do the following tasks:
 - a. Assess and clean your data if needed.
 - b. Use a different type of Data Visualization tools for each of the following:
 - i. Compare cash and credit totals.
 - ii. Compare each age and sum of total spending.
 - iii. Show each city total spending and arrange it by total descending.
 - iv. Display the distribution of total spending.
 - c. Put all previous plots in one dashboard.
 - d. Split the customers to (n) groups -using one of the studied methods- according to:
 - i. The sum of total spending for each customer.
 - ii. And the age for each customer.
 - iii. Print a table displaying each customer name, age, total spending and the computed cluster number.
 - e. Generate association rules between items with minimum support and confidence taken from the user as inputs (State the algorithm used).

❖ **Also, you can use the following guidelines to help you with implementing your Program:**

❖ **Program user inputs**

Variable name	label	Notes	Validation
Dataset_Path	Dataset path	User should input the full path of the data file.	Required
Number_Of_Clusters	Numbers of clusters	To use in the clustering process.	Number between 2 and 4
Min_Support	Minimum Apriori support	To use in the Apriori Algorithm.	Number between 0.001 and 1
Min_Confidence	Minimum Apriori confidence	To use in the Apriori Algorithm.	Number between 0.001 and 1

Submission Details:

- ❖ You should submit the following on a form according to instructions:
 - ✓ R version of the code.
 - ✓ Project Report with Group number as a pdf version.

- ❖ You can use the following links to guide you with building R Interface
 - <https://cran.r-project.org/web/packages/ggplotgui/readme/README.html>
 - <https://rstudio.github.io/shinythemes/>
 - https://doserlab.com/files/for875/_book/shiny-interactive-web-apps-in-r

Good Luck 😊