BUSA3020 - Assignment 3

Assignment Points: 100
Assignment Weight: 20%
Due Date: Friday of Week 11 at 11:59pm

About This Assignment

Customer segmentation is the process of dividing customers into groups based on common characteristics so companies can market to each group effectively and appropriately. It can be employed by all types of business, regardless of size, industry and whether they sell online or in person. For example, a small business selling guitars might decide to promote lower-priced products to younger guitarists and higher-priced premium guitars to older musicians based on segment knowledge which tells them that younger musicians have less disposable income than their older counterparts.

Task

You are employed by a large supermarket chain to perform customer segmentation analysis. In particular, you are provided with a dataset on 2000 customers that has been collected through loyalty cards that customers use at checkout, and includes variables such as customer age, gender, annual income, etc.

Your task is to analyse the data in Python using relevant techniques covered in lectures, and identify customer segments given the dataset. You will then collect your results in tables and diagrams which you will include in a report of **up to 1000 words** (not counting tables and figures). Make sure your report is informative, intuitive, and easy to understand by the management, who may not have technical expertise.

Your report should include the following headings:

1. **Introduction** (10 marks)
   - Motivate the problem and what you are trying to achomplish in this report
2. **Exploratory Data Analysis** (20 marks)
   - Perform basic exploratory data analysis, present and comment on summary statistics, etc.
3. **Customer Segmentation** (40 marks)
   - Apply clustering analysis we covered in lectures
   - Choose an optimal number of customer segments/clusters
   - Interpret identified clusters in terms customer profiles, e.g. there may be a cluster of university educated females living in a mid-sized city
   - Use two different clustering techniques
   - How do the clusters identified by the two techniques compare?
4. **Recommendations** (20 marks)
   - Provide some suggestions regarding marketing techniques that may be developed for different clusters of customers
5. **Conclusion** (10 marks)
   - A brief summary of what you have done in the report

Submission

You will need to submit two files:

1. One Jupyter Notebook containing Python code used in your analysis. Make sure your code is clean and easy to understand by the markers. Include comments.
2. One PDF document containing your report to be submitted to the management. The report should be self-contained and include all relevant tables/graphs that you generated using the submitted Jupyter notebook. You can write the report in Word, and then save it as a PDF file which will be submitted.