

DAB303 – Marketing Analytics– Project 2: Segmentation

Introduction

The purpose of the lab is to understand and gain insights from a retail dataset, by performing various exploratory data analysis, data visualization, and data modelling tasks.

Data:

The provided dataset, made available during the lab, contains all the information needed for the project.

Tasks:

1. **Download and load data file** – as described below:
 - Download the dataset (in .csv file format) from Blackboard.
2. **Use Python Tools for developing the desired model:**
 - Import the dataset as a Pandas Dataframe, followed by data pre-processing and data cleaning.
 - Perform exploratory data analysis (EDA):
 - o Data aggregation
 - o Data Visualization
 - o Advanced analysis, etc.
 - Advanced Machine Learning:
 - o Feature Engineering,
 - o Segmentation and Clustering,
 - o Predictive Analytics,
 - o Association Rule Mining,
 - o ... etc., as described in the Jupyter Notebook

You may use additional techniques which may not be listed above, provided that you can submit a rationale for why the technique is useful and an indication of what you hope to achieve.

3. **Report** – In a separate word document:
 - Record your observations with respect to the most important outputs of the Python code.

Submission – Deliverables

Submission will be done via Blackboard, and it will be group submission, including:

- One file per group (in .zip format):
 - o Jupyter Notebook/lab file (.ipynb)
 - o Exported Jupyter notebook in html (.html)
 - o Report (.pdf), and
 - o Presentation (.pptx)