**RMIT Online – Data Science Strategy and Leadership**

**Consumer Analytics - Assignment 1**

**Hassan Ismail / Student ID: S3823277**

**Data Plan and Business Solution / ASOS case study**

**Business requirements:**

Decreasing retention rate within 15-24 age group in the last two years is impacting the profit margin and the market share, which is an issue due to the fact that a new customer acquisition cost is 6-7 times higher than the cost of an existing customer retention ([struto.co.uk](https://www.struto.co.uk/) - Ref#1).

**The objective** of this proposed initiative is to improve the customer experience and improve retention rates, in particular within the 15-24 age group.

**Possible solution:**

The proposed solution consists of two parallel approaches:

1. The first one targets the consumers (in 14-25 age group) who already decreased their spend or most likely to decrease their spend, then plan marketing campaigns and tailored offers/reward programs to keep them from looking elsewhere and increase the retention rates.
2. The second approach will be general for all consumers, which will also increase profit margin by increasing the average order value and help in retaining consumers by increasing the trustworthiness.

**How can the problem be solved:**

For the first approach: We need to identify the consumers (with decreasing repeat purchase ratio) who already decreases or most likely to decrease spent: by using a regression model based on the historical data we have for our individual consumers, and then different teams need to be involved in the next steps to initiate the proper marketing campaigns and the reward programs.

For the second one: we need to analyse our historical consumer data and develop a recommendation system to offer what consumers may like, this will help to increase cross-selling and the average order value by motivating shoppers to buy additional products, we also need to conduct a simple analysis on the products and select the most profitable one to promote them within the ecommerce platforms.

**Feasibility:**

For the regression model, we need available historical data for consumers which should contain features like ( age, gender, suburb, spend history (to calculate the monthly or annual spend), type of purchases, work and education) those feature will be the input of the regression model which will help to get the predicted change in the spend for each customers for the next months/year.

For the recommender system algorithm, we need historical for consumers which should contain features: age, gender, suburb, purchase history, work, education).

If we don't have the data then this approach will not work. If we do have the data then this approach may work depending on the consistency, quality, and volume of the data.

**Proposed timeline:** (detailed in appendices)

(1) Collect data for both approaches (1 week)

(2) Asses data feasibility of the proposed Regression Model and the Recommender System (1 week)

(3) Data preparation and pre-processing (2 week)

(4) Build and test the regression model and feature selections (2 weeks)

(5) Build and train the recommender engine algorithm (2 weeks)

(6) Present preliminary results and insights for feedback (1 day)

(7) Put the algorithm into operation/production (2 weeks)

(8) Ongoing testing and improving of ML/AI algorithms (BAU task)

**Success criteria:**

The regression model can accurately predict the consumers who are most likely to decrease RPR.

The recommender engine can lead to increment in AOV.

**Risks/Assumptions:**

We assume the marketing and other teams will be able to build on the information and suggestion provided and perform the required campaigns and programs to meet the objectives.

We assume data are sufficiently available in quantity and quality.

Unreliable data will result unreliable ML/AI algorithms.