**Data Mining Group Project Proposal** (MScA 31008, Professor U. Pamuksuz)

**Team:** MScA Advisors

**Team members:** Alicia Zhang, Oleksiy Anokhin, Prafulla Ranjan Dash, Surendiran Rangaraj, Hanwen Serena Xu

**Suggested dataset:** Starbucks App Customer Reward Program (<https://www.kaggle.com/blacktile/starbucks-app-customer-reward-program-data>)

**Dataset overview**: This data set contains simulated data that mimics customer behavior on the Starbucks rewards mobile app. This data set is a simplified version of the real Starbucks app because the underlying simulator only has one product whereas Starbucks actually sells dozens of products. This transactional data showing user purchases made on the app including the timestamp of purchase and the amount of money spent on a purchase. It also has a record for each offer that a user receives as well as a record for when a user actually views the offer. There are also records for when a user completes an offer. The dataset includes one month of simulated customer data, including their purchasing habits, and interactions with promotional offers.

**Dataset structure:** 3 .json files, 4500 records, unbalanced data, some missing data, minor typos

**Client’s expectations:** our team was hired by Starbucks (Chicago) team to analyze and improve the current Starbucks Rewards program. Our client expects data-driven recommendations and new insights for the improvement of its marketing operations

**Business problem:** improvement of existing Starbucks Rewards system through analysis and segmentation of existing customers with a future application of learned lessons on potential customers

**Examples of business questions:**

* What were the main drivers of an effective offer on the Starbucks app?
* Which customers’ group is the most promising audience for a promo campaign?
* Can we classify customers per different segments and what are their unique features?
* Which promo offers receive the highest response rate and why?
* Is the current promo campaign effective? How can it be improved?

**The potential application of received knowledge:** our team considers this dataset as an excellent opportunity to use a wide variety of data mining algorithms. Combining the analysis of numeric and categorical variables, our team will use the acquired data mining algorithms and extract valuable insights. This topic allows us to create a great project for a future portfolio with retail, finance, text mining, health industry, etc).

Respectfully **Team MScA Advisors**