**Capstone Project**

**Document Skeleton**

By Chau Tran

# Problem statement

* **What is the problem or the opportunity that the project is investigating?**

My Capstone project is about predicting correct target customers who will take offers from direct marketing campaigns.

Direct marketing has been one of the efficient strategies for retail business to retain customer loyalty and to improve revenue against their competitors. With advanced technology nowadays, delivering outstanding and personalised shopping experiences are the key for retailers to encourage consumers coming back to their stores.

This project investigates different customer segmentations at Kroger, how their marketing campaigns impact on the overall engagement and forecasts the target groups for their future marketing campaigns.

* **Why is this problem valuable to address?**

A research done by Frederick Reichheld of Bain & Company shows an increasing customer retention rates by 5% increases profits by at least 25%.

## According to a Customer Retention Statistics in 2018 of Annexcloud (a marketing consulting firm), the financial impact of customer retention include:

1. Recruiting new customers costs [five times](http://www.destinationcrm.com/Articles/Web-Exclusives/Viewpoints/Listen-to-the-Voice-of-the-Customer-53239.aspx) as much as retaining current customers.
2. [20 percent](https://www.slideshare.net/custthermometer/22-customer-retention-stats) of customers of a given company are the source of 80 percent of the company’s profits.
3. It is [16 times](https://martech.zone/infographic-customer-retention-numbers/) as costly to build a long-term business relationship with a new customer than simply to cultivate the loyalty of an existing customers.

Korger – the retail studied in this project - has 2757 supermarkets and multi-department stores across 35 states, 240 carts per second, 11 million customers daily and 60 million household data. A lost of 1% customers daily results in 110,000 carts at the cost of average $3,190,000 sale revenue.

* **What is the current state (e.g. unsatisfied customers, lost revenue)?**
  + Retails use direct marketing to communicate with their customers and offer discount vouchers.
  + Highly competitive against other general retails.
  + Threats of loosing market shares with online retails.
  + Customers have various and flexible choices with their shopping activities and experiences on multiple offline / online channels.
* **What is the desired state?**
  + Maintain retention rate of loyal customers group.
* **Has this problem been addressed by other research projects? What were the outcomes?**
  + Yes, other research projects suggest to re-define retail customer experiences by using data and technology to understand customers’ behaviours better. Business should customise communications, services and products given to their customers at personal level using history data to improve customer satisfaction in overall and their life time value.

# Industry/ domain

* **What is the industry/ domain?**
  + General retail that sells foods and goods for household consumption.
* **What is the current state of this industry? (e.g. challenges from startups)**
  + Highly competitive against other general retails (Walmart).
  + High threats of loosing market shares with online retails (Instacart, Amazon), especially due to Covid19 pandemic.
  + Consumers have various and flexible choices with their shopping activities and experiences.
* **What is the overall industry value-chain?**
  + Inbound logistics: inventory (raw or finished) receiving from suppliers, storing and management.
  + Operations: processes of manufacturing raw materials to finished products; labelling and organising products in bundles that add more sale value.
  + Outbound logistics: distributing final products to different stores, shipping orders to end consumers.
  + **Marketing and Sales (main focus of this project)**: activities that increase the awareness of consumers, target appropriate groups using advertising and pricing strategies and maintain customer’s loyalty.
  + Services: customer service, refund, exchange and other communication with customers to enhance and maintain consumer experience with products and the stores.
* **What are the key concepts in the industry?**
  + Foods and goods for household consumption
  + In-store/ physical shopping experiences
* **Is the project relevant to other industries?**
  + Yes, this project can be applied to other types of in-store retails such as equipment, fashion, cosmetic, furniture, stationery, etc…

# Stakeholders

* **Who are the stakeholders?** 
  + Marketing Manager at Kroger retail
* **Why do they care about this problem?**
  + Direct marketing may come at a very high cost for business with large amount of customers. Taking Kroger retail as the studied business in this project. Korger has 2757 supermarkets and multi-department stores across 35 states, 240 carts per second, 11 million customers daily and 60 million household data.
  + Identifying the right target groups to send suitable marketing campaigns can help the business:
    - to maintain and improve customer retention rate.
    - to increase revenue
    - to save cost from sending marketing to the wrong customer segments
* **What are the stakeholders’ expectations?**
  + maintain customer retention rate of loyalty group.

# Business question

* **What is the main business question that needs to be answered?**

Can retail business forecast their target customers before they launch a direct marketing campaign?

* **What is the business value of answering this question? (quantify value and make necessary assumptions)**

It is vital to identify correct target customers as this allow marketers to focus on those most likely to purchase the products, limit the budgets to the groups of people with the highest profit potential and personalise shopping experiences for their customers.

* **What is the required accuracy? What are the implications of false positives or false negatives?**

The aim is to get accuracy of better than 60%. The cost of false positives will be the printing, posting and monitoring cost due to targeting the wrong customer groups**.** This cost can be minimised by using digital marketing. The cost of false negatives will be a lost in achieving upsell from value customers. False negatives are more important than false positives.

# Data question

* **What is the data question that needs to be answered?**

Can we predict conversion probability of retail customers who will receive direct marketing offers using customer demographic and transaction history data?

* **What is the data required to answer the question?**
  + Customer transaction history: customer id, transaction id, basket id, product id, transaction date time, transaction value, product details, promotion offer redemption.
  + Customer demographic data: age group, income group, gender, marital status
  + Marketing campaign data: campaign date, discount vouchers, who received campaign, who redeemed vouchers.

# Data

* **Where was the data sourced?**
  + The data was retrieved from Dunnhumby (a consultant company using data science) website:

https://www.dunnhumby.com/source-files

* **What is the volume and attributes of the data?**
  + 2,595,732 transaction records of 2500 customers with 276,484 baskets in 102 weeks.
  + Demographic details of 801 customers
  + 30 mail campaigns sent to 1584 customers
  + 1135 coupons records
  + 92,353 products
* **How reliable is the data?**
  + The data is simulated based on patterns of real data. There maybe some degree that data do not truly reflect the real business data. The average shopping basket size appeared small $30.
* **What is the quality of the raw data?**
  + The raw data are cleaned up with clear column names. It has a detailed description and a schema diagram.
* **How was this data generated?**

#### The data was generated by Dunnhumby with assistance of Raj Venkatesan - Research Professor of Business Administration, University of Virginia.

* **Is this data available on an ongoing basis?**
  + Yes, according to Kroger’s Fact Book 2019, Kroger are able to capture 97% transaction data on their customer loyalty card.

# References

<https://www.emarketer.com/content/challenges-and-solutions-for-customer-retention>

<https://news.microsoft.com/transform/kroger-smart-shelves-ditch-paper-drop-lights-delight-shoppers/>

<http://s1.q4cdn.com/137099145/files/doc_downloads/irw/fact_books/2019-Fact-Book.pdf>

<https://www.investopedia.com/ask/answers/050115/what-are-primary-activities-michael-porters-value-chain.asp>

<https://www.annexcloud.com/blog/21-surprising-customer-retention-statistics-2018/>

<https://hbr.org/2014/10/the-value-of-keeping-the-right-customers>