**ISyE 6501**

**Introduction to Analytics Modeling**

Course Project

This project should be done individually.

The web sites <https://www.sas.com/en_us/customers.html>, <https://www.ibm.com/casestudies/search?search>, and <https://www.informs.org/Impact/O.R.-Analytics-Success-Stories> (among others) contain brief overviews of some major Analytics success stories. In this course project, your job is to think carefully about what analytics models and data might have been required.

1. Browse the short overviews of the projects. Read a bunch of them – they’re really interesting. But don’t try to read them all unless you have a lot of spare time; there are lots!
2. Pick a project for which you think at least three different Analytics models might have been combined to create the solution.
3. Think carefully and critically about what models might be used to create the solution, how they would be combined, what specific data might be needed to use the models, how it might be collected, and how often it might need to be refreshed and the models re-run. DO NOT find a description online (or elsewhere) of what the company or organization actually did. I want this project to be about your ideas, not about reading what someone else did.
4. Write a short report describing your answers to (3).

=====================================================================================

**China CITIC Bank**

Customer segmentation boosts credit card activations 40 percent: <https://www.sas.com/en_us/customers/citic.html>

**Models used**: In this case study of credit card sales, models used primarily will be clustering, multi-armed bandit a/b testing, and cost optimization.

**Analysis Walkthrough, Combination**:

* **Clustering** – Pipeline would consist of a clustering model for customer segmentation based on their profile to identify similar types of people. Within each group, a/b testing could be implemented based on different kinds marketing methods ranging from simple electronic direct mailers to offers and discounts.
* **Exploitation vs Exploration Testing** – Depending on the company or country policy on allowed marketing frequency and limited pool of customers in each segment, it will be a balance between exploitation vs exploration for the a/b tests. In this scenario, a multi-armed model would help to reduce profit wastage on the tests.
* **Cost Optimization** – Within each customer segment and between marketing methods, we will then be able to evaluate what marketing methods appeal to which type of customers. Along with segment count, marketing costs for each method, and potential credit card activation rates, we will then be able to run an optimization model to minimize marketing costs by only targeting and matching responsive customer segments to the most effective methods, while maximizing activation rates.

**Data required:**

* For segmentation, customer data is key. We would be looking at:
  + Customer spending behavior (on debit or credit card)
    - Average monthly spending pattern
    - Spend on retail categories
  + Customer credit score (cash flow)
  + Net worth in account
  + Demographic information (if allowed in the country)
  + Customer lifecycle (with the bank)
  + Holding which types of cards (miles, points, rebates)
  + Supplementary card holder
  + Supplementary spending behavior
  + Customer salary estimate
* For a/b testing, we would be interested in:
  + Customer segment from clustering
  + Customer sign-up rate for each customer segment after marketing methods
* For cost optimization, variables required:
  + Marketing costs for each method
  + Maximum customer pool for each advertising run
  + Profit by estimated activation rates for each segment and methods applied

**Data collection:**

* **Customer Data**
  + Customer data and predictor variables could be engineered from existing transactional data in the bank. Spending behavior can be aggregated from monthly card spend or transfers to/from savings account.
  + Customer credit data can be requested from the Credit Bureau (if in Singapore).
  + Demographic information should be provided by the customer on signup or any products.
  + Salary estimate could be derived from a monthly constant source of deposit, but might have too many missing data if customers don’t use accounts for salary deposit.
* **Marketing Costs**
  + Depending on marketing methods, coupons, discounts, or giveaways can be calculated internally or deducted from sponsorship costs.
  + Electronic direct mailers are usually free if managed internally.
  + Bulk SMS costs from partnered telcos.

**Refresh frequency:**

While it might take some time for a customer’s spending habits to change, customer preference and marketing effectiveness might be volatile. If marketing campaigns were to run monthly, I would suggest a **quarterly** **refresh** of the models, to churn new customer segments and to test marketing effectiveness of different methods.

However, for cost optimization, monthly campaigns should be listed out and optimized on a monthly basis based on number of available customers in a particular segment accounting for monthly marketing contact limit and business budget. Process might be iterative and new refresh cycle can be improved with more data.