Final Project Report

DATA609 – Mathematical Modeling Techniques for Data Analytics

Author: Bethany Poulin & Leland Randles

May 31st, 2020

# 

Contents

[Project Proposal 3](#_Toc41823896)

[Merchant Pricing 3](#_Toc41823897)

[Company Overview 3](#_Toc41823898)

[Dataset 5](#_Toc41823899)

[Methods 5](#_Toc41823900)

[Final Model 6](#_Toc41823901)

[Results 7](#_Toc41823902)

[Conclusions 7](#_Toc41823903)

# Project Proposal

Credit card processing is a vital part of the new world economy and responsible for the vast majority of both online and brick & mortar purchases in today’s marketplace. The margins for processors are slight and competition is fierce forcing providers to rely on both volume and strategic means of carving out new revenue streams to grow.

This analysis uses real-world, deidentified data from a processing provider’s rate and fee structure, along with the current rates, regulations and requirements for each of the four major global regions to optimize future earnings through a combination of controlled rate and fee retooling.

The resulting schedule of rates and fees would mean a 1.6% increase in revenue for 2020, (with purchase volume and number of transactions held constant) while maintaining all our corporate policies, a host of contractual agreements as well as respecting regulatory requirements in all countries and regions of operation.

# Merchant Pricing

Merchants are priced via a combination of a percentage amount charged on the “volume” (amount) of the transaction and a per-transaction fee. For example, if a customer goes into a retailer and buy an item for $100, the “volume” of the transaction is $100.00, and it is a single transaction. So if the pricing for the merchant is 1.95% plus $0.15 per transaction, then the merchant would be charged $2.10 for the transaction:

(1.95% \* 100.00) + (1 \* 0.15) = $1.95 + $0.15 = $2.10

In this case, $2.10 would go to the credit card processing company, and $97.90 would go to the retailer. This, of course, is all blind to the consumer. They just know they spent $100.00 on the item they purchased.

# Company Overview

Our dataset came from an S&P500 credit card processing company. The company has four regional headquarters: The United States, Canada, Europe and Asia-Pacific. Within Asia-Pacific, there are multiple countries within which the company operates.

The company offers virtually every card brand that exists, but for the purposes of the this project, the data was limited to four majors worldwide card brands: Visa, Mastercard, American Express and Discover.

Lastly, the company divides its markets into four main lines of business: Direct, Enterprise, Integrated and Wholesale.

The region, country, and card brand can also restrict the ways in which merchants can be repriced. In some regions or countries, credit card processing is heavily regulated, and to reprice merchants, you must get approval for increases outside of safe harbors. The card brands also impose limits on what credit card processors are allowed to do as part of the right to use their brand.

Of course, the company also has internal policies regarding price increases that are based on competitive approach, concerns about churn, and the like. For this dataset, churn thresholds were set for each line of business. The churn threshold is the point at which a price increase causes merchants to leave at a greater rate than they can be replaced.

The tables below summarize all of the regions, countries, card brands and lines of business the company is involved with, along with pricing limits and churn thresholds:

|  |  |  |
| --- | --- | --- |
| **Region** | **Country** | **Repricing Restrictions** |
| Asia-Pacific | Hong Kong | No restrictions |
| Asia-Pacific | India | The max price increase by percentage is 1% + inflation rate (2019 rate was 4.54%, so 5.54%) |
| Asia-Pacific | Macau | No restrictions |
| Asia-Pacific | Malaysia | The max price increase by percentage is 20% |
| Asia-Pacific | Maldives | No restrictions |
| Asia-Pacific | Philippines | No price increases for competitive reasons |
| Asia-Pacific | Singapore | The max price increase by percentage is 15% |
| Asia-Pacific | Sri Lanka | The max price increase by percentage is 1% + inflation rate (2019 rate was 5.6%, so 6.6%) |
| Canada | Canada (CAD) | The max price increase by percentage is 10% |
| Canada | Canada (USD) | The max price increase by percentage is 10% |
| Europe | United Kingdom | The max price increase by percentage is 6% |
| USA | United States | No restrictions |

|  |  |
| --- | --- |
| **Card Brands** | **Repricing Restrictions** |
| Visa | Cannot raise the percentage rate a merchant is paying by more than 100 basis points |
| MasterCard | Cannot raise the percentage rate a merchant is paying by more than 100 basis points |
| Amex | Cannot raise total percentage rate above 500 basis points |
| Discover | Only per-item rate increase are allowed |

|  |  |  |
| --- | --- | --- |
| **Lines of Business** | **Price Sensitivity / Churn Threshold** | **Description** |
| Direct | Low / 10.0% | SMEs (small to medium enterprises) sold directly by internal sales staff |
| Enterprise | High / 4.5% | Large enterprises sold directly by internal sales staff |
| Integrated | Low / 11.5% | Payment services are integrated into software sold by 3rd party vendor |
| Wholesale | n/a | Payment services sold to resellers |

# Dataset

Our dataset consisted of 638,074 rows for 277,498 merchants. A merchant usually accepts multiple card brands, so each row of the dataset represents a specific merchant / card brand combination. Each merchant is tagged with a region, country and line of business.

# Methods

Data Grooming

The raw data included 8 variables (in addition to a unique merchant ID Number), the following table is an example of those variables.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **lob** | **region** | **country** | **card** | **rate** | **fee** | **volume** | **transactions** |
| Direct | Canada | Canada | Amex | 0.035 | 0 | 7.22 | 1 |
| Direct | Canada | Canada | Master Card | 0.015394 | 0 | 1104.1 | 8 |
| Direct | Europe | UK | Visa | 0.004548 | .0004 | 51427.59 | 3605 |
| Direct | Canada | Canada | Amex | 0.0175 | 0 | 40.27 | 5 |
| Direct | Canada | Canada | Amex | 0.0175 | 0 | 413.25 | 13 |
| Direct | Asia-Pacific | India | Visa | 0.007195 | 0 | 14286.41 | 207 |

In order to create the constraint equations, we used the volume and number of transactions to create an *average traction* *amount*. Using this average transaction amount we calculated the complete average transaction cost associated with the average transaction, which we call the *current* *price* by adding the *rate price (average transaction amount x rate) to the fee*.

These are the values which we used to calculate the proposed rates and fees to optimize the company returns using linear programming so they were appended to the data frame.

The next thing we needed to do was in corporate observation-specific minimum rates, maximum rates as well as minimum fees and maximum fees to use in our linear program. Although this could be done with a function and lookup tables however at scale this would be computationally heavy, and much easier to include the values in the data frame itself.

This was done using the card-based rate limits and the companies own limit of 10% when no limit was supplied. Fees were card-specific or ½ of the current fee as minimum and 2 times the current fee as maximum when no fee limits were supplied.

There were multiple price constraints for each merchant/card combination

* Country/Region Price Maximum
* Card Specific Price Maximum
* Company Price Maximum
* Churn Risk Price Maximum

These fees and rate limitations were extracted from individual tables for each price constraint, then prices were calculated and assigned to the data frame in separate columns.

All four of the price constraints are measured against the same equation in the linear program (above), so when setting up the program in R, the minimum of the all of the price constraints is used in the final configuration, as the model would select for the minimum of the four price maximums.

# Final Model

The Objective Function

Constraints

Price Constraint

Rate Constraints

Fee Constraints

After creating the data frame, adding the new rates, fees and prices to the data frame and creating constraints we built a function containing the following linear programming code.

## Linear Model Code

  obj = c(1, average\_transaction)

  mat = matrix(c(1, x2\_coeff, 1,0, 1,0,0,1,0,1),ncol =2, byrow=TRUE)

  dir = c('<=', '>=', '<=', '>=', '<=')

  rhs = c(min\_max\_price, min\_fee, max\_fee, min\_rate, max\_rate)

  output =lp("max", obj, mat, dir, rhs)

  new\_fee = output$solution[1]

  new\_rate = output$solution[2]

  new\_transaction\_price = new\_fee + new\_rate \* average\_transaction

# Results

The function was applied across 200,000 rows of the original data returning a data frame with the following variables

* Merchant
* Card
* Old\_Fee
* Old\_Rate
* Old\_Transaction\_Price
* Old\_Revenue
* New\_Fee
* New\_Rate
* New\_Transaction\_Price
* New\_Revenue

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Merchant** | **Card** | **Old Fee** | **Old Rate** | **Old Price** | **Old Revenue** | **New Fee** | **New Rate** | **New Price** | **New Revenue** |
| b12cc76c8e84f | Amex | 0 | 0.0175 | 0.154875 | 0.154875 | 0 | 0.01925 | 0.1703625 | 0.1703625 |
| 04a56fa3be44e6 | Amex | 0.0142 | 0 | 0.0142 | 0.1278 | 0.0071 | .000340075 | 0.01562 | 0.14058 |
| c63246892ad3dfd5 | Amex | 0.0569 | 0 | 0.0569 | 0.3414 | 0.02845 | .0000187 | 0.06259 | 0.37554 |
| d938af0bb6181f30 | Amex | 0.0085 | 0 | 0.0085 | 0.034 | 0.00425 | .000166043 | 0.00935 | 0.0374 |
| 54022b8597ed9bcb | MasterCard | 0 | 0.022423 | 0.29641154 | 178.1433354 | 0 | 0.0246653 | 0.326052694 | 195.957669 |

**Total Increase In Revenue :** $ 1,359,632

**Percent Increase in Revenue**: 1.611406%

The following two observations represent the minimum and maximum increases in total revenube by Country-Region, Business Line and Card.

Lowest Increase Country-Line-Card Combination

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Country** | **Business Line** | **Card** | **Projected Revenue** | **Total 2019** | **Percent Increase** |
| Canada | Enterprise | Discover | 2,322.01 | 2,314.04 | .34475 |
| United States | Integrated | AMEX | 1,082,676.24 | 3,661,641.85 | 11.49851 |

Highest Increase Country-Line-Card Combination

# Conclusions

Based on running a linear program for each row in the dataset, it appears the company could increase revenue by approximately $1.36 million without inducing damaging churn.

The analysis also gleans insights into what can be done to improve the analysis. Perhaps it would be better to separate the dataset into tranches and have a separate linear program for each tranche. If such tranches were created, it might make sense to further sub-stratify them by industry or market, as there is a great range in rates within the dataset, no matter how you break it down (by region, country, card brand, line of business, etc).

In this project we limited ourselves to linear programming, but this may be a multi-objective optimization problem that requires more than linear programming, including a fully-formed churn model, inclusion of costs, and a ranking method. In a real-world portfolio repricing initiative, there would be a revenue target, costs would be accounted for, and some ranking would be needed to list merchants from most favorable for repricing to those least favorable for repricing.