

# SUPERMARKET PROMOTION

Group 3 DSB

# CONTENT PAGE

- Introduction
- Business & Scope
- Dataset
- Strategy Summary
- Data Model
- Revenue Prediction
- Clustering Customers
- Product Line Analysis
- Strategies
- Closed Loop Data Ecosystem
- Considerations For Future Work

# INTRODUCTION

- Myanmar supermarket operating in 3 cities – Yangon, Mandalay, Naypitaw.  
Yangon is the commercial capital of Myanmar, whereas Naypitaw is the capital.
- Supermarket has a current membership programme, of which 50% of transaction by registered members have enrolled.

## 6 product lines:

- Fashion Accessories
- Food & Beverages
- Electronic Accessories
- Sports and Travel
- Home and Lifestyle
- Health & Beauty

## 3 modes of payment type:

- Cash
- Credit Card
- e-Wallet



# BUSINESS AND SCOPE

Strategies for **customer retention / revenue growth** based on the following scope:

- Revenue from customers
- Revenue from payment mode
- Retention and upselling



# DATASET

Variables	Description
Invoice ID	Unique identifier for each invoice
Branch/City	Identify which city the transaction was made Yangon (A), Mandalay (B), Naypitaw (C)
Customer Type	If the customer is a member or normal customer
Gender	Male or Female
Product Line	6 product lines, as previously illustrated
Unit Price	Average price of each item per transaction (invoice)
Quantity	Number of item purchased
Date	Date of purchase
Time	Time of transaction
Payment	Payment Type



# **SUMMARY - STRATEGY**

- Revenue from customers – product line analysis
- Retention & Upselling - designing member coupon benefits to stimulate higher spending from the members
- Revenue from payment mode - deciding on payment method strategies (partnership with credit card companies, enhancement to e-wallet system, store shopping card system etc.)

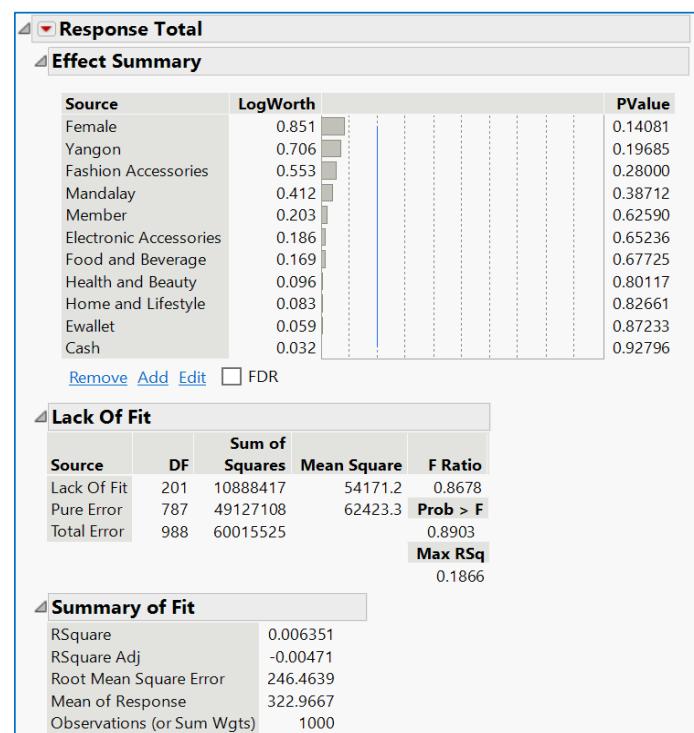


# REVENUE PREDICTION (1/2)

## Linear regression

- Regressed Supermarket revenue using regressors including gender, city, payment type, customer type, and product category.
- Created dummy variables of all the categorical variables to perform the analysis
- The results were inconclusive. The R-square value for this regression model was under 1%

## Linear regression using JMP Pro



# REVENUE PREDICTION (2/2)

## Forward Stepwise Logistic Regression

- Created a dummy variable for the revenue. Income over 500 was encoded as 1 and under 500 as 0. City, gender, payment type, and product type were used as input variables.
- The results were inconclusive.

### Forward Stepwise using Minimum AICc Indicator

Stepwise Fit for Over 500

Stepwise Regression Control

Stopping Rule: Minimum AICc      Enter All      Make Model

Direction: Forward      Remove All      Run Model

Rules: Combine

Go      Stop      Step

-LogLikelihood	p	RSquare	AICc	BIC
535.62592	14	0.0000	1073.26	1078.16

### Logistic regression using JMP Pro

Fit Group

Nominal Logistic Fit for Over 500  
Converged in Gradient, 4 iterations

Iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	-1.137e-13	0	2.3e-13	.
Full	535.6259			
Reduced	535.6259			

RSquare (U) -0.0000  
AICc 1073.26  
BIC 1078.16  
Observations (or Sum Wgts) 1000

Fit Details

Measure	Training Definition
Entropy RSquare	-0.000 1-Loglike(model)/Loglike(0)
Generalized RSquare	-0.000 (1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.5356 $\sum -\log(p[j])/n$
RASE	0.4189 $\sqrt{\sum(y[i]-p[i])^2/n}$
Mean Abs Dev	0.3509 $\sum  y[i]-p[i] /n$
Misclassification Rate	0.2270 $\sum (p[i] \neq pMax)/n$
N	1000 n

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.225329	0.0754914	263.46	<.0001*

For log odds of 1/0

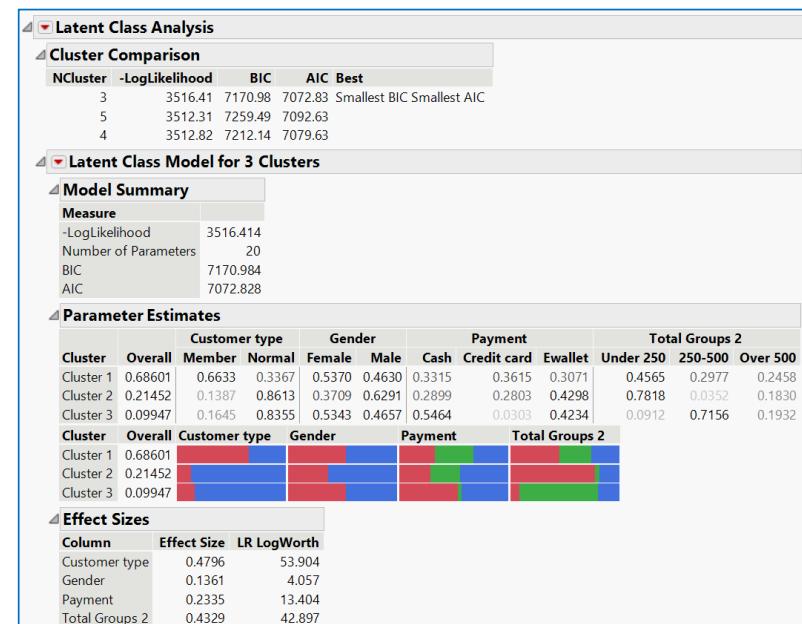


# CLUSTERING CUSTOMERS (1/2)

## A. Latent class analysis

- Latent class analysis (LCA) can be used when the attributes are mostly of categorical nature.
- The attributes considered were customer type (member, non-member of the supermarket), payment type (e-wallet, cash, credit card), gender, and transaction amount (Under 250, between 250 and 500, and over 500).
- Three clusters were derived for customer segmentation – Loyalists, Low Spenders, and Hot Targets

*LCA using JMP Pro*



# CLUSTERING CUSTOMERS (2/2)

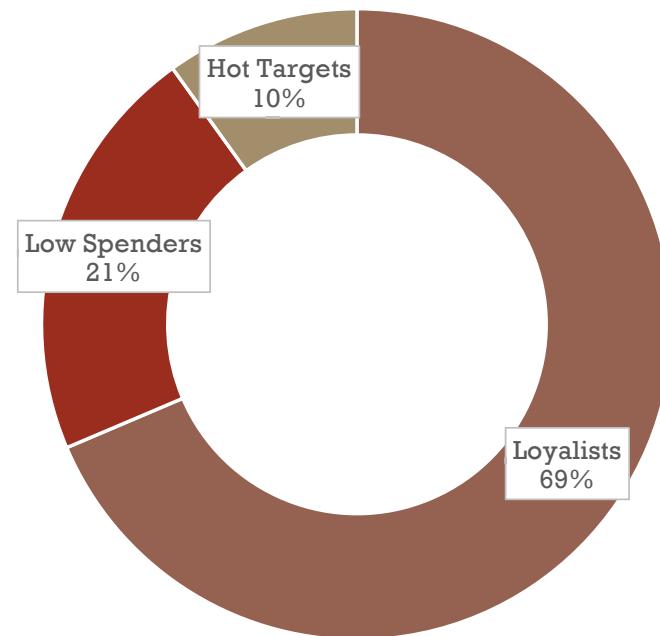
## Hot Targets

- Majority non- member group
- Preferred payment mode: cash, e-wallet
- 78% between 250 and 500

## Low Spenders

- Majority non- member group
- Male dominated
- 43% preferred E-wallet
- 78% spent under 250

## **Customer Clusters from LCA**



## Loyalists

- Majority member group
- Even split across gender, payment mode
- 46% spend under 250, rest over 250

# PRODUCT LINE ANALYSIS



IMPROVE SALES



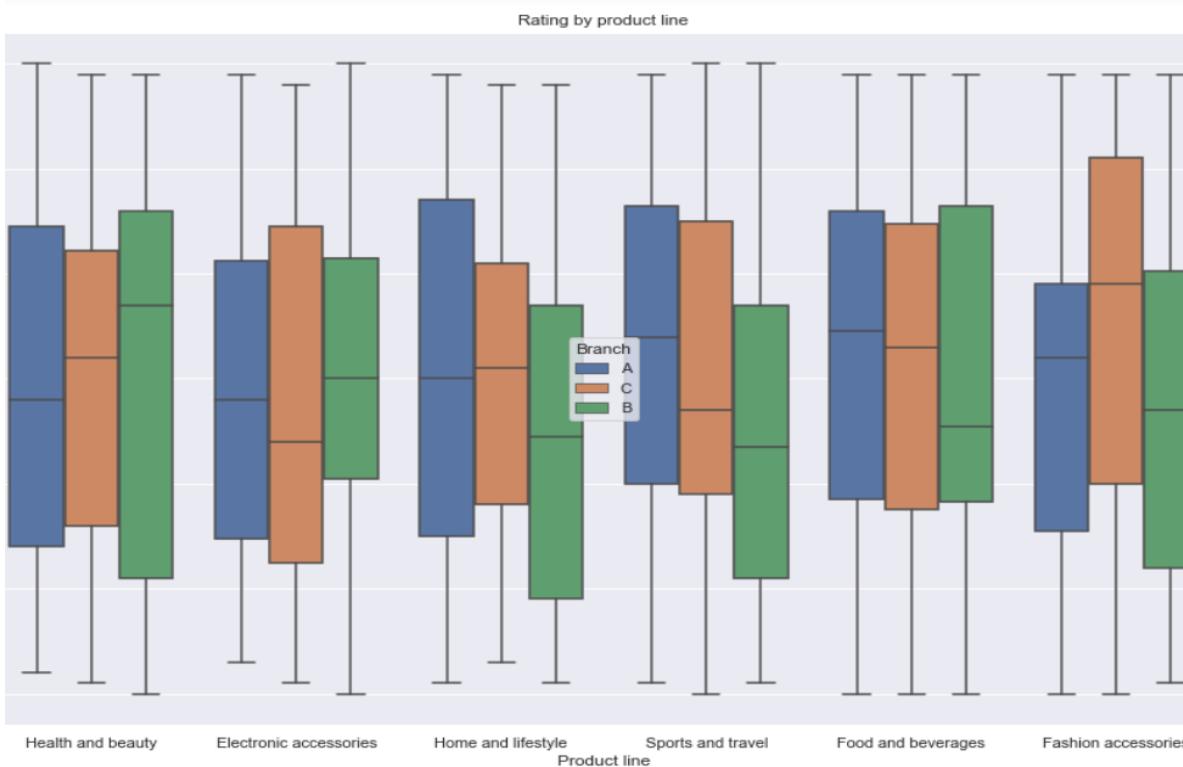
STOCK IN BRANCHES



DONE USING  
PYTHON  
VISUALIZATION  
TOOL



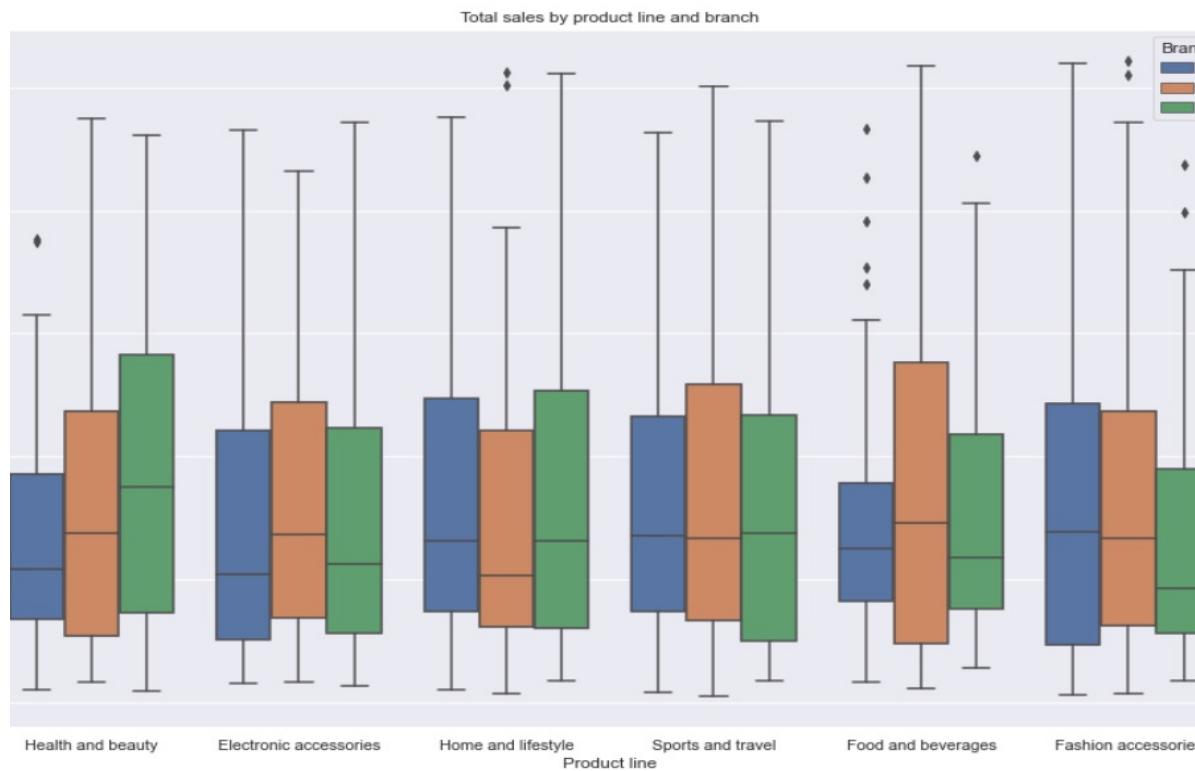
# PRODUCT LINE ANALYSIS - RATING



Insight 1 : Branch B has a low average rating in 4 product lines: Home & lifestyle, Sports & Travel, Food & beverages and lastly, Fashion Accessories



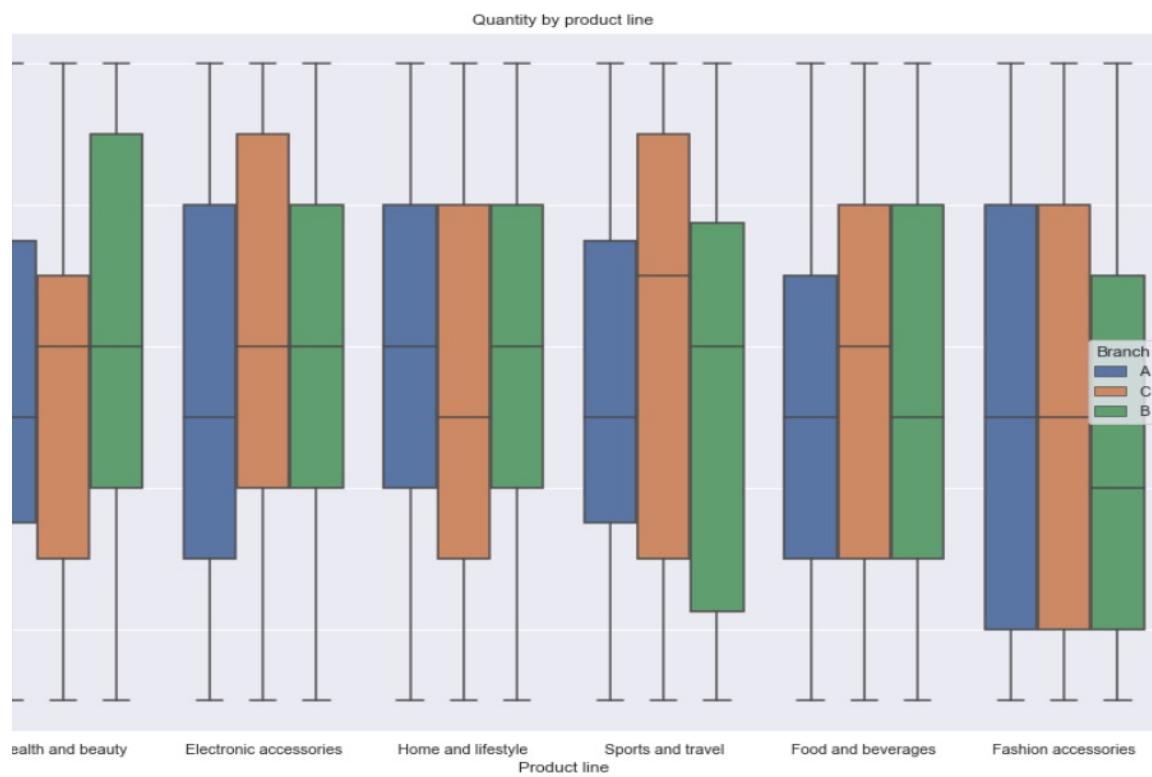
# PRODUCT LINE ANALYSIS - SALES



Insight 2: Highest and lowest average sales for each branch by product line



# PRODUCT LINE ANALYSIS - QUANTITY



Insight 3: Highest and lowest average quantity by product lines and branches



# RECOMMENDATIONS FOR PRODUCT LINE ANALYSIS

- Rating:

Branch B has a low average rating in 4 product lines:

- provide promotions for these four product lines.
- Advertising for these product lines

- Sales:

- Improve sales of required product lines for specific branches.
- Introduce discounts and improve the reach for these products through relevant advertising channels



# RECOMMENDATIONS FOR PRODUCT LINE ANALYSIS

- Quantity:
  - Branch C:
    - Increase stock of all product lines except Home and lifestyle
    - Promote Home and lifestyle products through strategies like combining products from this line with products of popular lines for a discounted price
  - A similar strategy can be applied for branch A – Electronics Accessories
  - For branch B
    - Increase stock for Health and beauty as well as Electronic Accessories product lines



# STRATEGY – COUPONS FOR MEMBERS

- Aim: to stimulate higher spending per receipt among the members
- Questions:
  - How to make members feel the benefit of the coupons and increase our profits at the same time?
  - What will make the members want to increase spending and use the coupon?
    - What is the minimum spend required?
    - How much should the face value of coupon be?
  - What is our profit increase, net of coupon cost?

# STRATEGY – COUPONS FOR MEMBERS

- **Nature of coupon**

- Cash coupon (fixed face amount) to offset payment.
- Minimum spending required.
- Distributed to each member when they make a purchase.

- **Assumptions**

- Gross profit margin: 50%
  - $\text{Gross profit} = \text{revenue} * 50\%$
- Equivalent discount to stimulate spending: 20%
  - A member will increase his/her purchase to the minimum spend amount, if the coupon represents a 20% or more discount for the additional purchase amount.

Coupon face amount	100
Minimum spend required	1000
Current spending	700
Top-up amount; implied discount	$300; 100/300 = 33\% > 20\%, \text{will use coupon}$
Profit increase net of coupon cost	$300*50\% - 100 = 50$



# STRATEGY – COUPON FOR MEMBERS

```
def coupon_benefit(coupon_value, min_spend):
    required_discount = 0.2
    profit_margin = 0.5
    #to find the minimum original purchase amount such that using a voucher is considered worthwhile
    purchase_min = min_spend - 10/2*coupon_value

    profit_inc = 0
    for i in l:
        if i >= purchase_min:
            new_purchase = max(min_spend, i)
            profit_inc = profit_inc + (new_purchase - i)*profit_margin - coupon_value

    return profit_inc

import pandas as pd
sales = pd.read_csv('supermarket_sales - Sheet1.csv',encoding = "ISO-8859-1")

l = sales['Total'][sales["Customer type"] == 'Member']

for min_spend in [100,300,500,700,900,1000]:
    print(min_spend,':', coupon_benefit(100, min_spend))

100 : -48150.34824999998
300 : -28526.175
500 : 4952.451249999997
700 : 6921.0520000000015
900 : 6308.225000000002
1000 : 5984.734250000003
```

# STRATEGY – COUPONS FOR MEMBERS

- Things to monitor to improve the coupon strategy:
  - Coupon usage rate
  - Average spending from members during the coupon promotional period
  - Profits of the supermarket over time
  - Validity of assumptions (equivalent discount to stimulate spending; gross profit margin)
  - Revised spending habits from members, e.g. favoring certain product lines

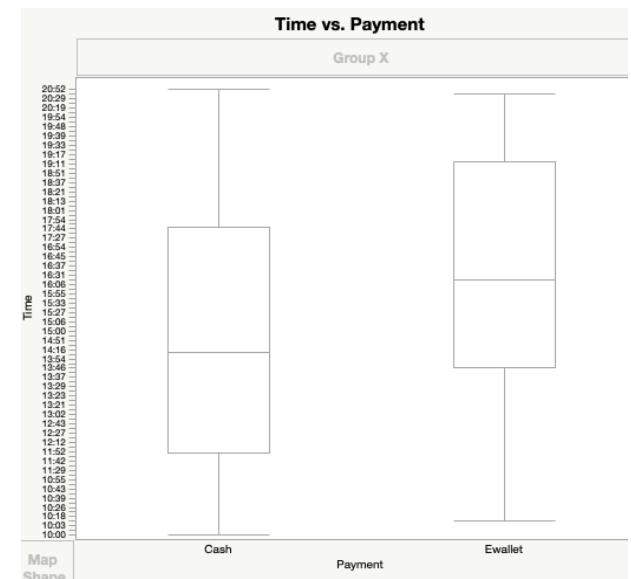
# OTHER SALES STRATEGIES - E-WALLET

- Enhancing e-Wallet system focused on “Hot Targets”
- Agenda: Increase e-Wallet spend amongst “Hot Target” customers as a foot in the door to build brand loyalty and grow wallet share
- Methodology: Product line & time specific promotions for e-wallet users. Target for 30% increase in revenue

**“swing the user”**

Time of Day	Product Line	% Weightage	% e-Wallet	Avg Unit Price	Ttl Incremental E-wallet Rev
Before 2pm	Fashion Accessories	11%	17%	59.1	1,260
After 2pm	Home and Lifestyle	12%	69%	63.7	1,527
After 2pm	F&B	12%	63%	59.8	1,435

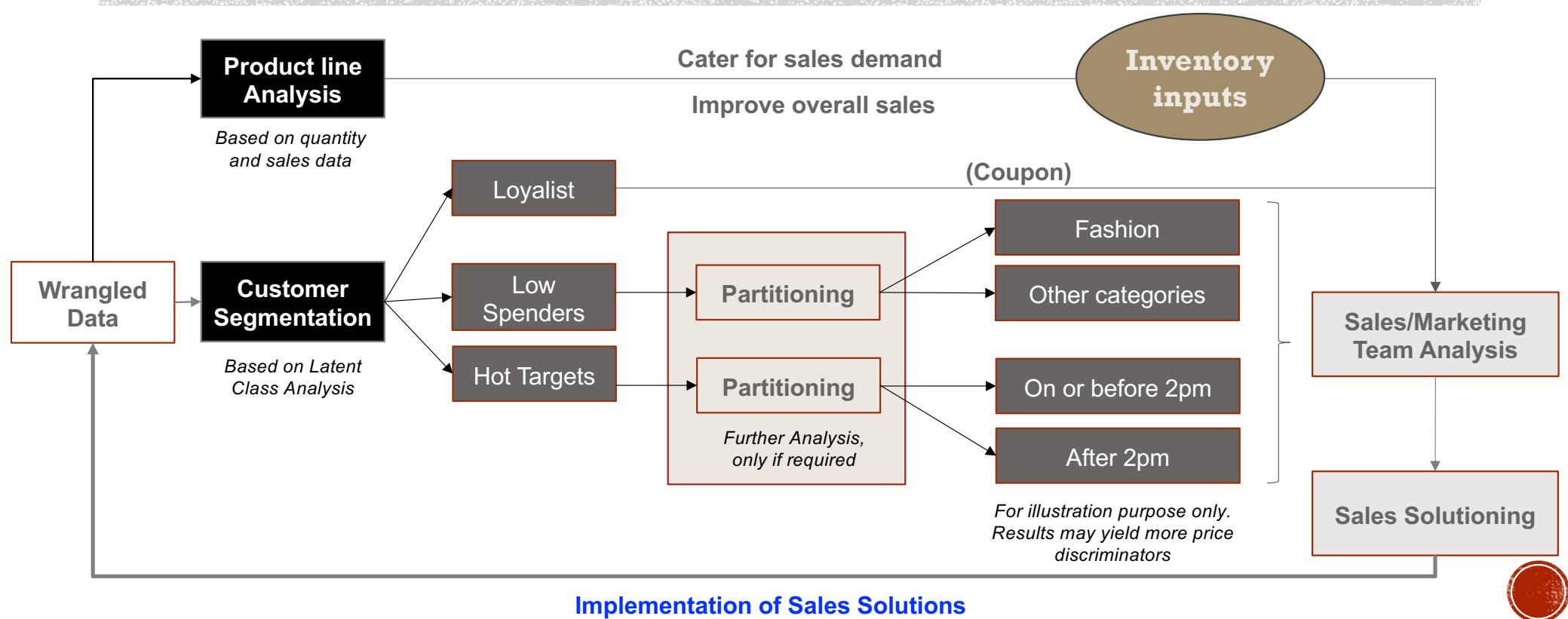
- Consideration: 14% upfront discount, or >14% rebate for return spend? Depends on assumption of return rate



Mode of Payment amongst “Hot Target”, by time of day



# CLOSED LOOP DATA ECOSYSTEM



# CONSIDERATIONS FOR FUTURE WORK

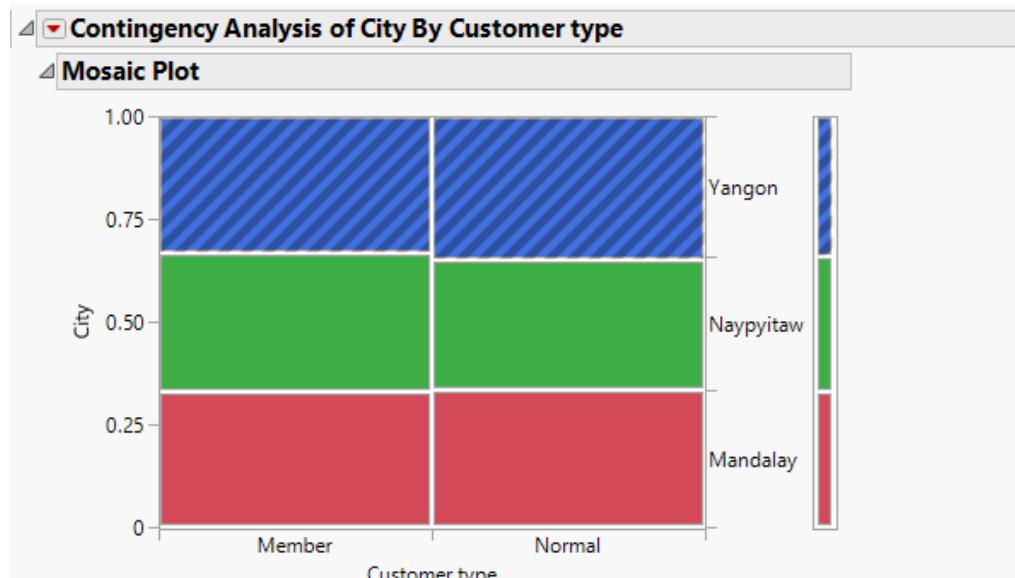
- Consider better data capture to contain the biases and multicollinearity among attributes.
- Consider capturing additional attributes such as frequency of purchase, actual products purchased, as well as additional demographic attributes of the customers.
- Additional attributes can help enhance customer clusters from the latent class analysis.
- Enhanced clustering can be used in the future to tailor coupons by customer's respective categories.



# APPENDIX



# CORRELATION BETWEEN CATEGORICAL VARIABLES: CHI-SQUARED TEST



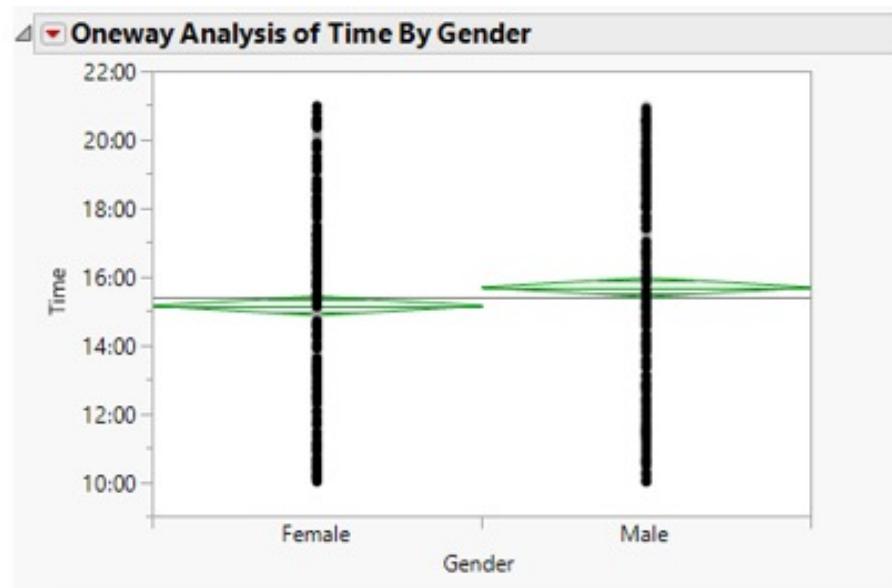
Tests			
N	DF	-LogLike	RSquare (U)
1000	2	0.20943070	0.0002
Test	ChiSquare	Prob > ChiSq	
Likelihood Ratio	0.419	0.8110	
Pearson	0.419	0.8111	

$H_0$ : The 2 categorical variables are correlated. Not Rejected

$H_1$ : The 2 categorical variables are not correlated.



# CORRELATION BETWEEN CATEGORICAL VARIABLES: ONE WAY ANOVA



Analysis of Variance

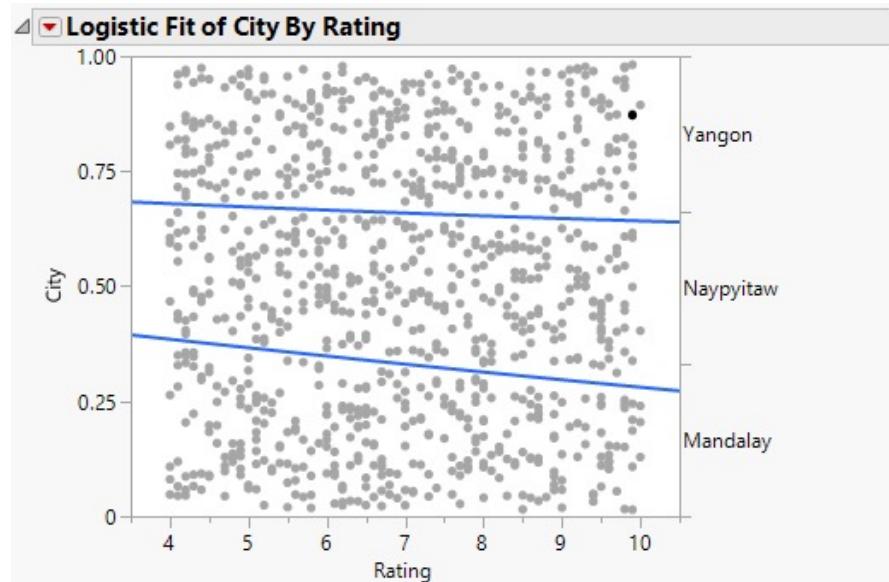
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Gender	1	947752302	947752302	7.2215	0.0073*
Error	998	1.3098e+11	131241089		
C. Total	999	1.3193e+11			

$H_0$ : The numerical means of the categorical variables' groups are correlated. **Not Rejected**

$H_1$ : The numerical means of the categorical variables' groups are not correlated.



# CORRELATION BETWEEN CATEGORICAL AND NUMERICAL VARIABLES: CHI SQUARE TEST



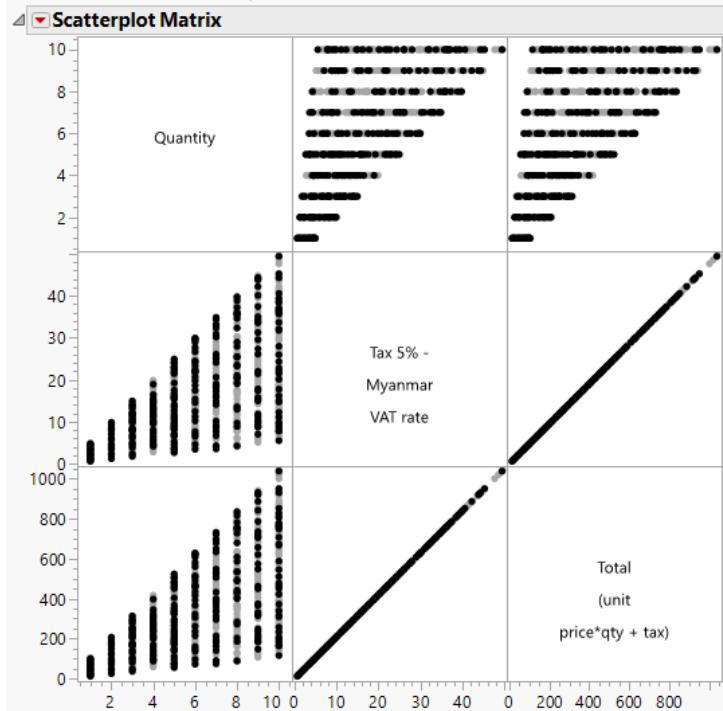
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	2.0764	2	4.152839	0.1254
Full	1096.4241			
Reduced	1098.5005			

$H_0$ : The pair of categorial and numerical variables are correlated. Not Rejected  
 $H_1$ : The pair of categorial and numerical variables are not correlated.



# CORRELATION BETWEEN THE NUMERICAL VARIABLES: MULTIVARIATE ANALYSIS



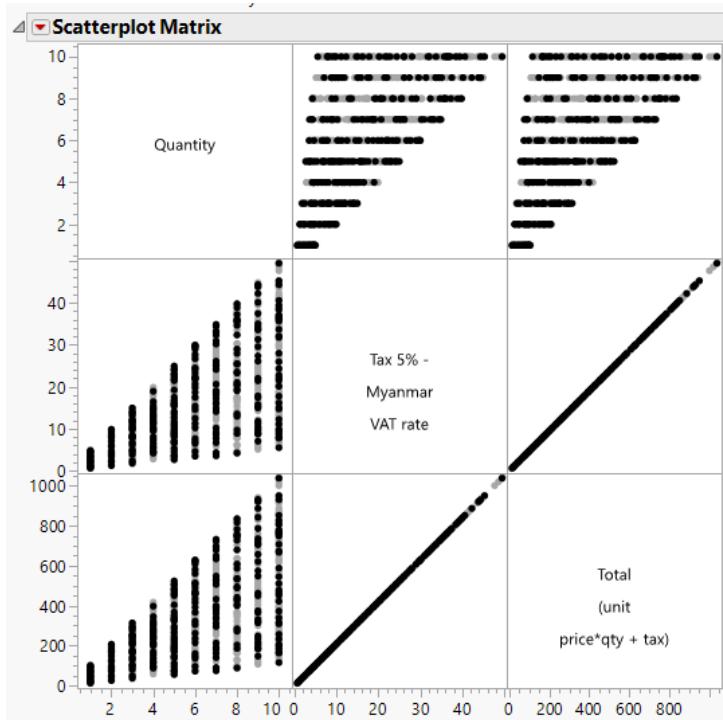
Pairwise Correlations

Variable	by Variable	Correlation
Total (unit price*qty + tax)	Tax 5% - Myanmar VAT rate	1.0000
Tax 5% - Myanmar VAT rate	Quantity	0.7055
Total (unit price*qty + tax)	Quantity	0.7055
Tax 5% - Myanmar VAT rate	Unit price	0.6340
Total (unit price*qty + tax)	Unit price	0.6340
Quantity	Unit price	0.0108
Time	Unit price	0.0076
Time	Tax 5% - Myanmar VAT rate	-0.0052
Time	Total (unit price*qty + tax)	-0.0052
Time	Quantity	-0.0086
Rating	Unit price	-0.0088
Rating	Quantity	-0.0158
Rating	Time	-0.0261
Rating	Tax 5% - Myanmar VAT rate	-0.0364
Rating	Total (unit price*qty + tax)	-0.0364

To remove Myanmar tax rate from analysis



# CORRELATION BETWEEN THE NUMERICAL VARIABLES: MULTIVARIATE ANALYSIS



Pairwise Correlations

Variable	by Variable	Correlation
Total (unit price*qty + tax)	Tax 5% - Myanmar VAT rate	1.0000
Tax 5% - Myanmar VAT rate	Quantity	0.7055
Total (unit price*qty + tax)	Quantity	0.7055
Tax 5% - Myanmar VAT rate	Unit price	0.6340
Total (unit price*qty + tax)	Unit price	0.6340
Quantity	Unit price	0.0108
Time	Unit price	0.0076
Time	Tax 5% - Myanmar VAT rate	-0.0052
Time	Total (unit price*qty + tax)	-0.0052
Time	Quantity	-0.0086
Rating	Unit price	-0.0088
Rating	Quantity	-0.0158
Rating	Time	-0.0261
Rating	Tax 5% - Myanmar VAT rate	-0.0364
Rating	Total (unit price*qty + tax)	-0.0364

To remove Myanmar tax rate from analysis

