

Retail Sales Analysis Project using



Prepared By :- Karan Singal
20th Feb 2020

Under the guidance of :- Ar Kar Min

Presentation Outline

01

Introduction & Objectives

02

Methodology & Study of Framework

03

Descriptive Analysis

04

Results & Conclusion



Introduction

- The data is from Canada's retail Industry sales of 1 year from May 2007 to April 2008 for more than 1200 customers nearly 30,000 Transactions
- Goal is to predict the yearly sales as per Quantity and price
- Presentation using SAS 9.4 Environment

Business Objectives



- Analyze the total sale in Canada's different Province and focus more in sales by Canadian's spending pattern
- To identify customer's frequent source of spending channel and try to boost the sale
- To infer the category in which customer's are spending more money

Methodology

- Identifying the Dataset
- Data Cleaning:- Missing values and sorting
- Statistical Procedures:- Proc Contents, means, freq
- Descriptive Analysis & Visualization
- Conclusion

Dataset-1

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
6	Category	Char	1	\$1.	\$1.
1	Customer_ID	Num	8	BEST12.	BEST32.
2	Item_Code	Char	8	\$8.	\$8.
5	Item_Description	Char	50	\$50.	\$50.
4	Order_Date	Num	8	DATETIME	ANYDTDTM40.
8	Quantity	Num	8	BEST12.	BEST32.
3	Source	Char	7	\$7.	\$7.
7	price	Num	8	BEST12.	BEST32.

Dataset-2

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
11	Category_code	Char	4	\$4.	\$4.
3	City	Char	13	\$13.	\$13.
2	Customer_Number	Num	8	BEST12.	BEST32.
10	Item_Description	Char	47	\$47.	\$47.
9	Item_Num	Num	8	BEST12.	BEST32.
6	Order_First_Time	Char	1	\$1.	\$1.
1	Order_Number	Char	8	\$8.	\$8.
5	Postal_Code	Char	6	\$6.	\$6.
4	Prov	Char	2	\$2.	\$2.
12	Quantity	Num	8	BEST12.	BEST32.
8	Sales_amount	Char	10	\$10.	\$10.
7	Source	Char	7	\$7.	\$7.



Details	
Number of Transactions	34,288
Order Year	May-2007 to April-2008
Demographic Information	Not provided
Unique Value	Customer_ID
Important variables	Order Date, Quantity, & Category

Dataset-1	
Number of Transactions	34,288
Order Year	May-2007 to April-2008
Demographic Information	Not provided
Unique Value	Customer_ID
Important variables	Order Date, Quantity, & Category

Dataset-2	
Number of Transactions	1,297
Order Year	May-2007 to April-2008
Demographic Information	City, Province and Postal codes
Unique Value	Customer_Number
Important variables	Locations, First time buyers

Research Framework

Independent
Variables

Customer ID

Order Date

Quantity

Category

Order First Time

Dependent Variables

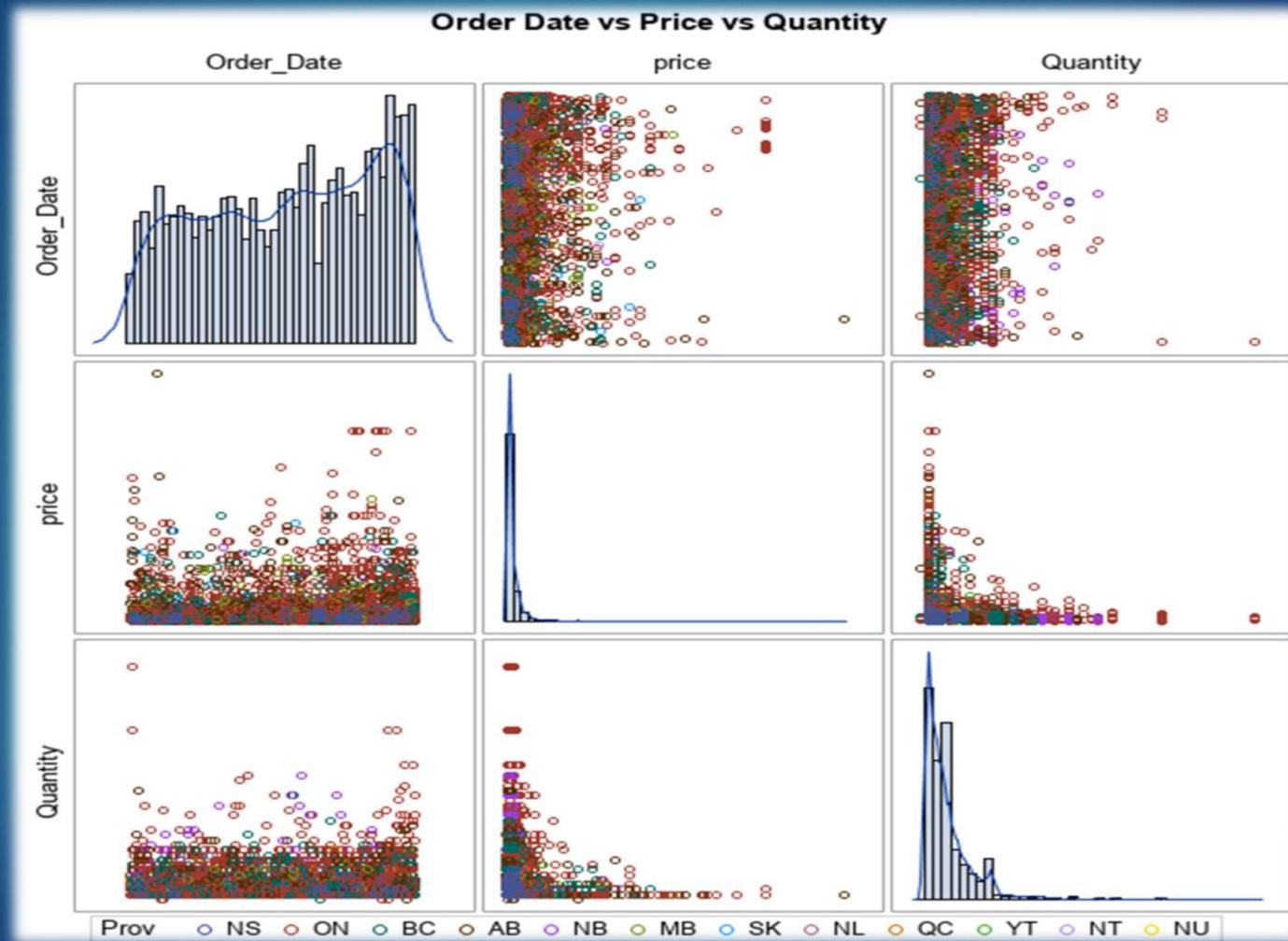
Province

Sales Price

Preparing the Data

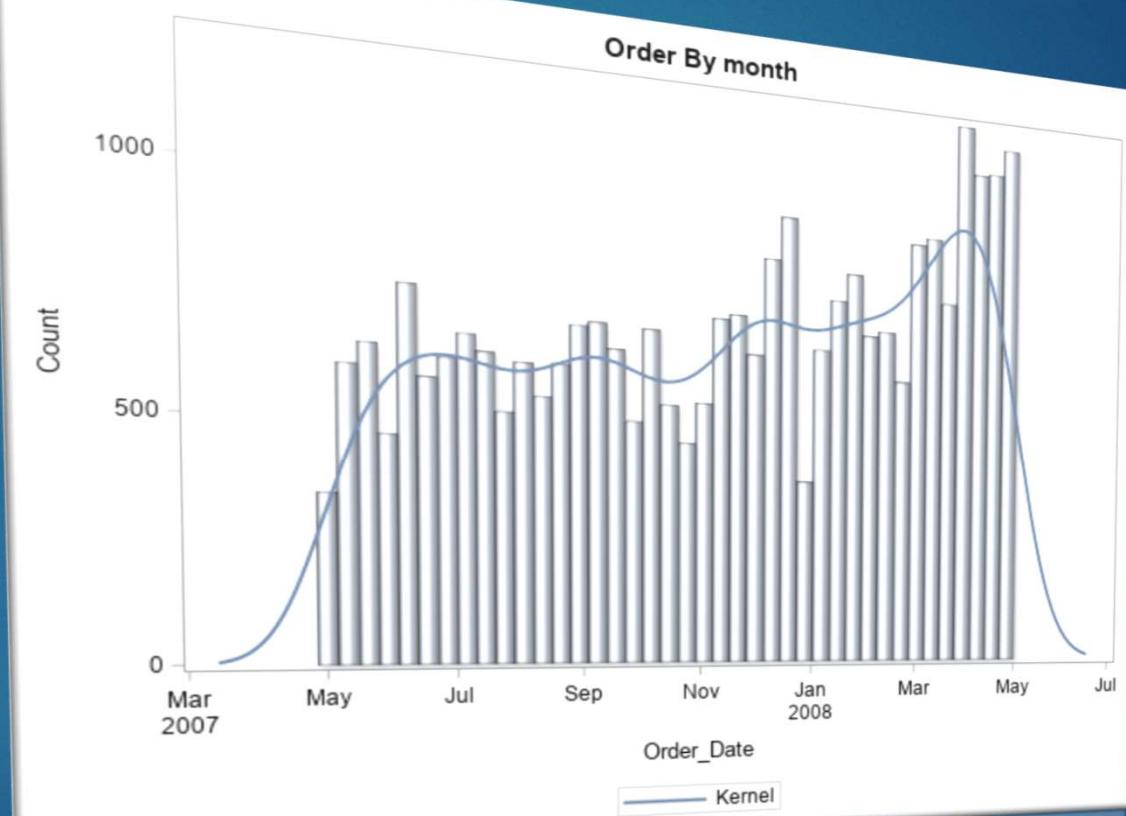
- Remove all the duplicate records
- Count missing values and remove unwanted records
- Merge both the datasets with unique values for more detailed analysis

Scatter plot

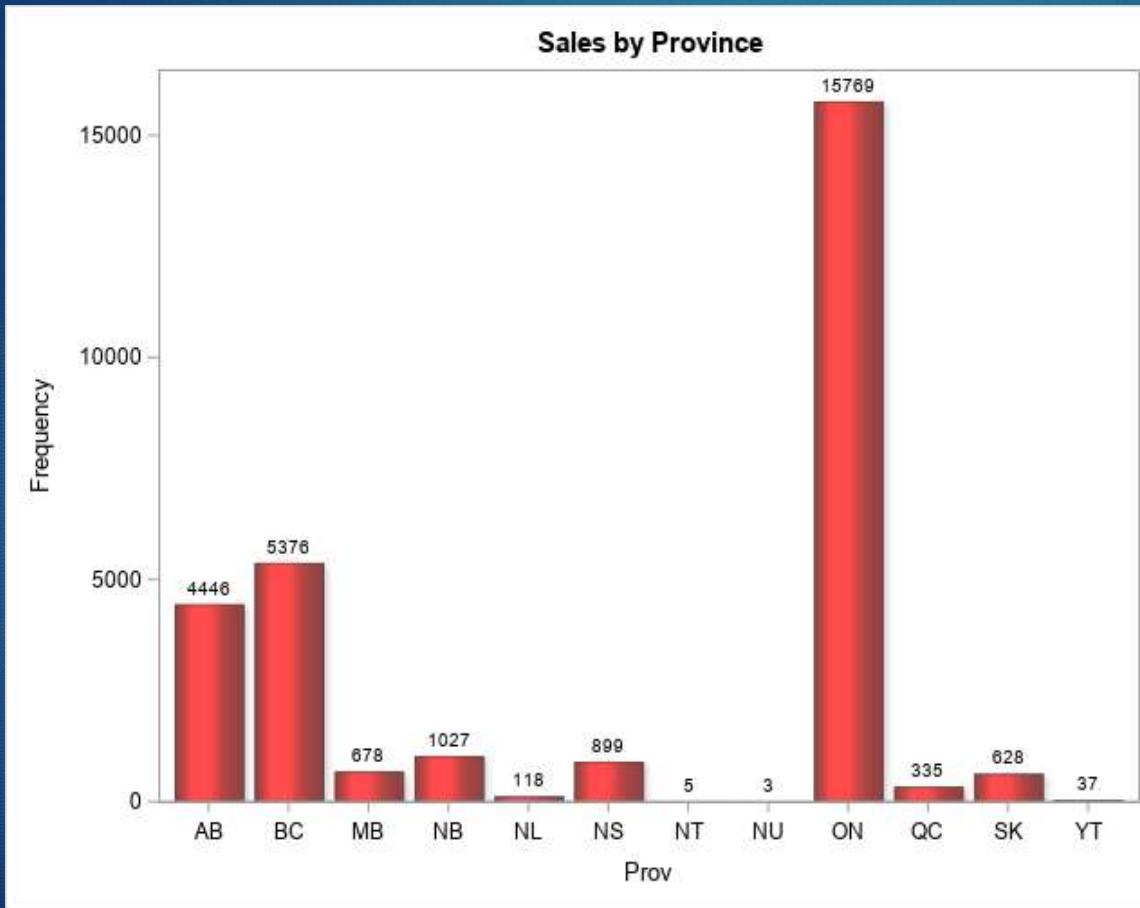


10

Sales by each month Groups



Number of Sales by Province



53.8 %

Ontario

18.3 %

British Columbia

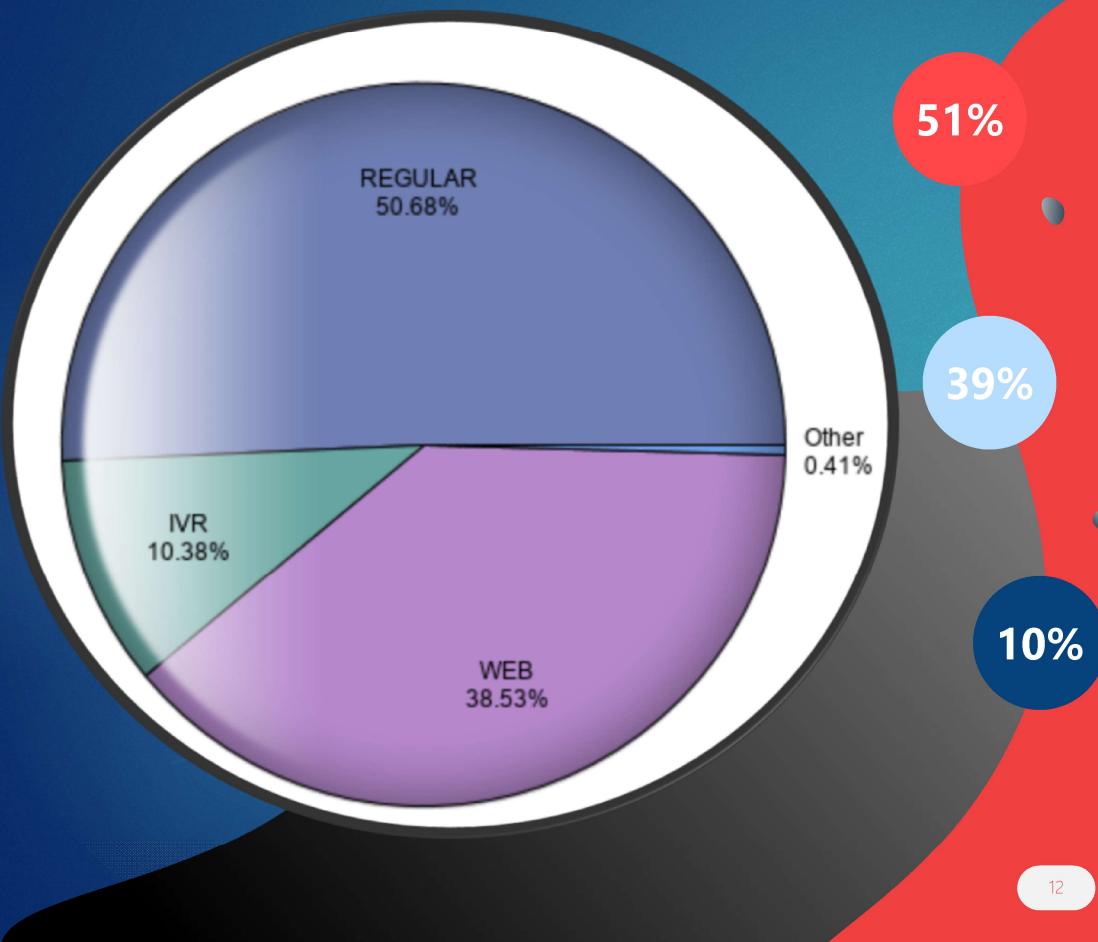
15.2 %

Alberta

12.7 %

Others

Total Sales by Source



Regular

In Canada majority buyers wants to do sopping in traditional way, they like to visit stores for shopping

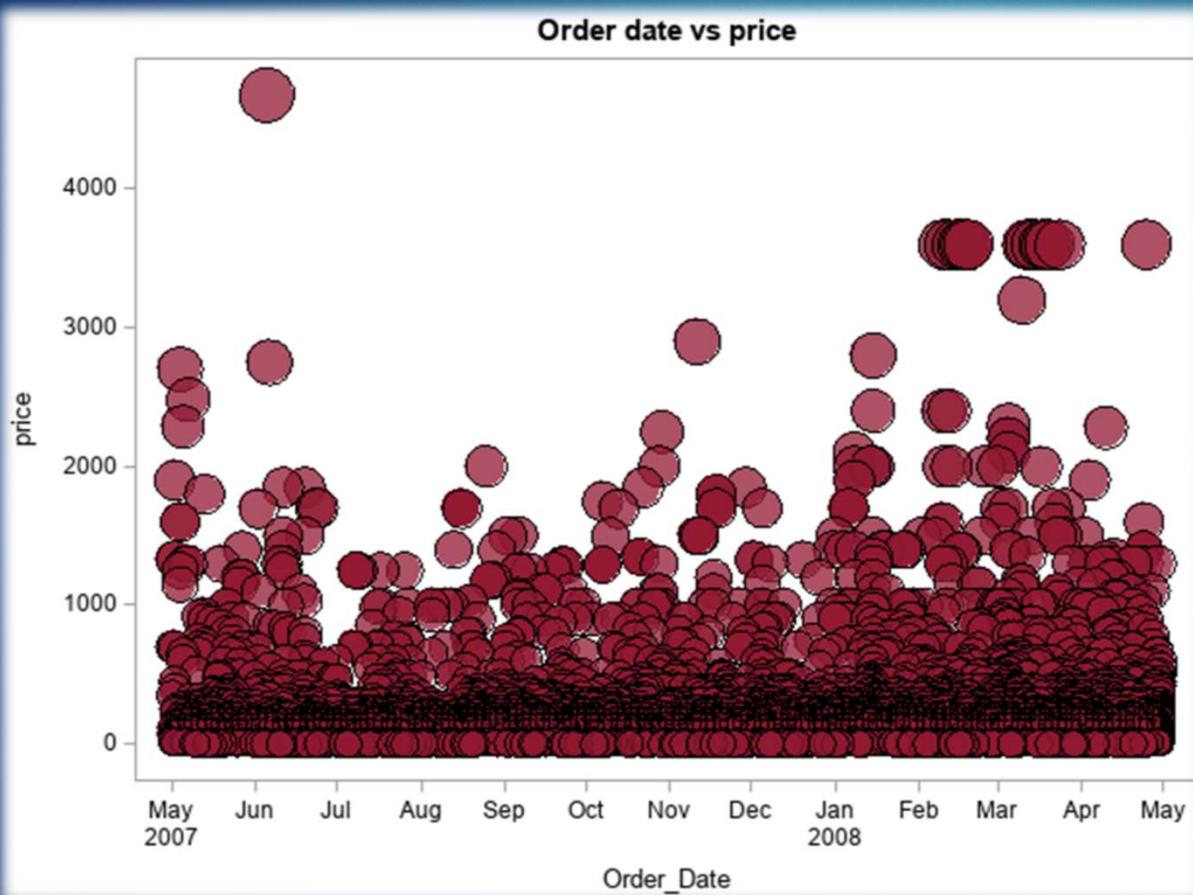
Web

However people are getting used to with online shopping, as the online website like Amazon and eBay were just in their initial stages

IVR

As people don't want to go outside in winter and internet sites were not that popular, Canadians still use telephone shopping and shops trough IVR

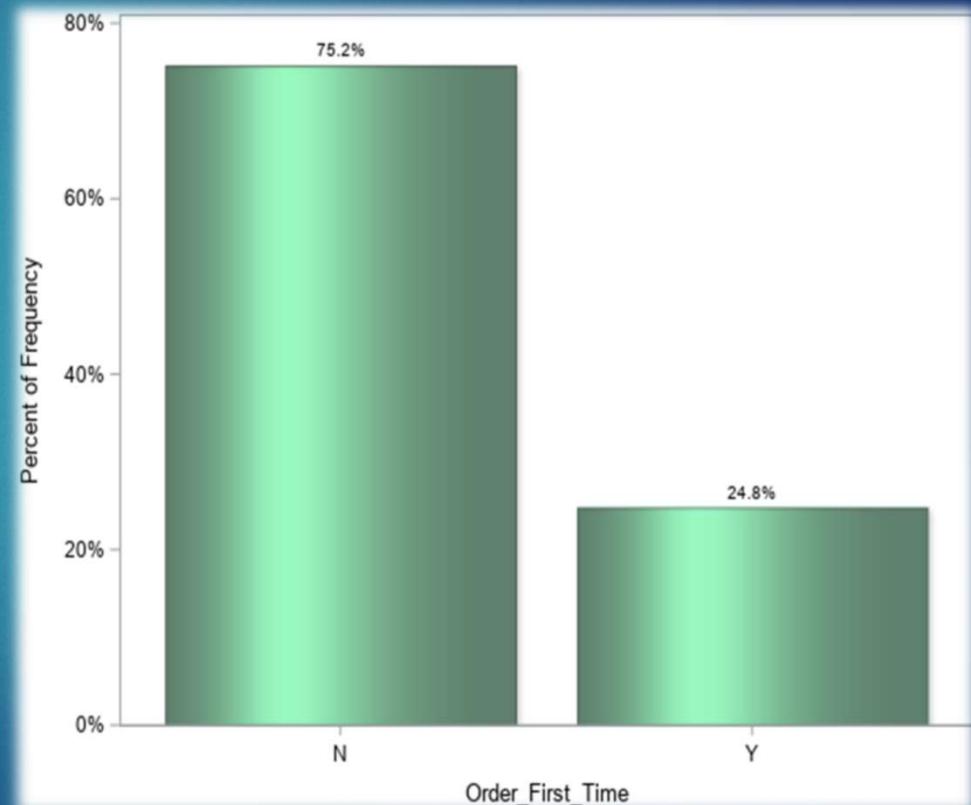
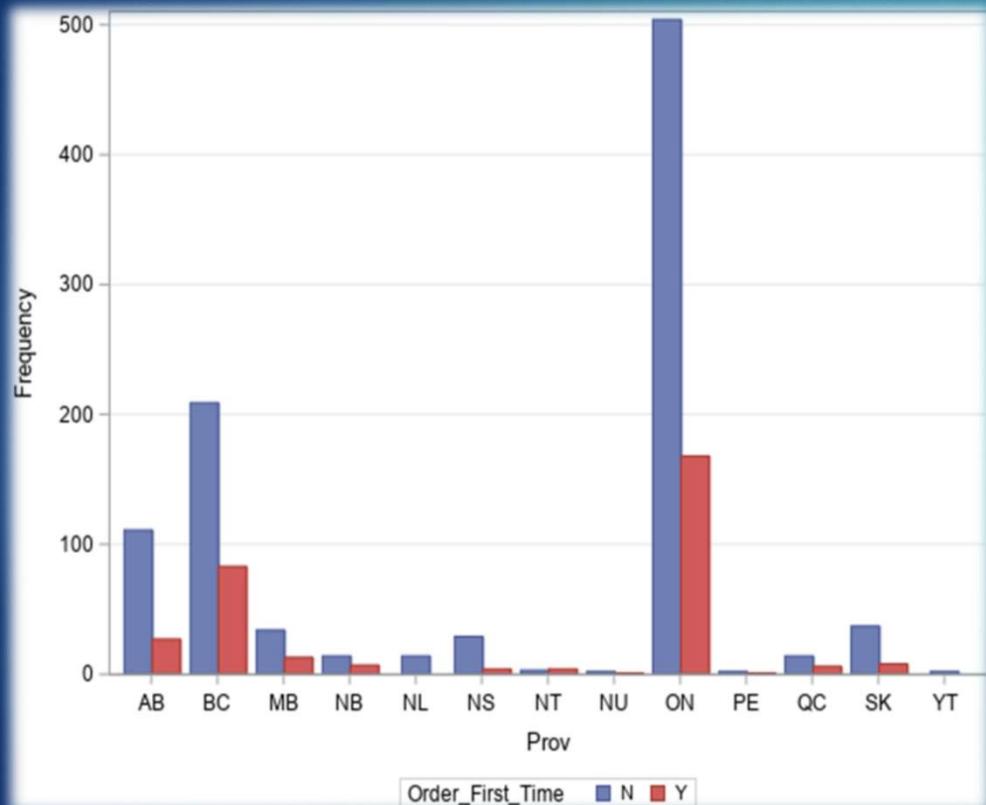
Other analysis



- Most of the transactions are below 500 \$
- From 2008 Feb 3700 \$ of transactions occurred which is for 18 caret gold neckless
- Transaction of 4600 is adjusted bill, we can not consider that transaction

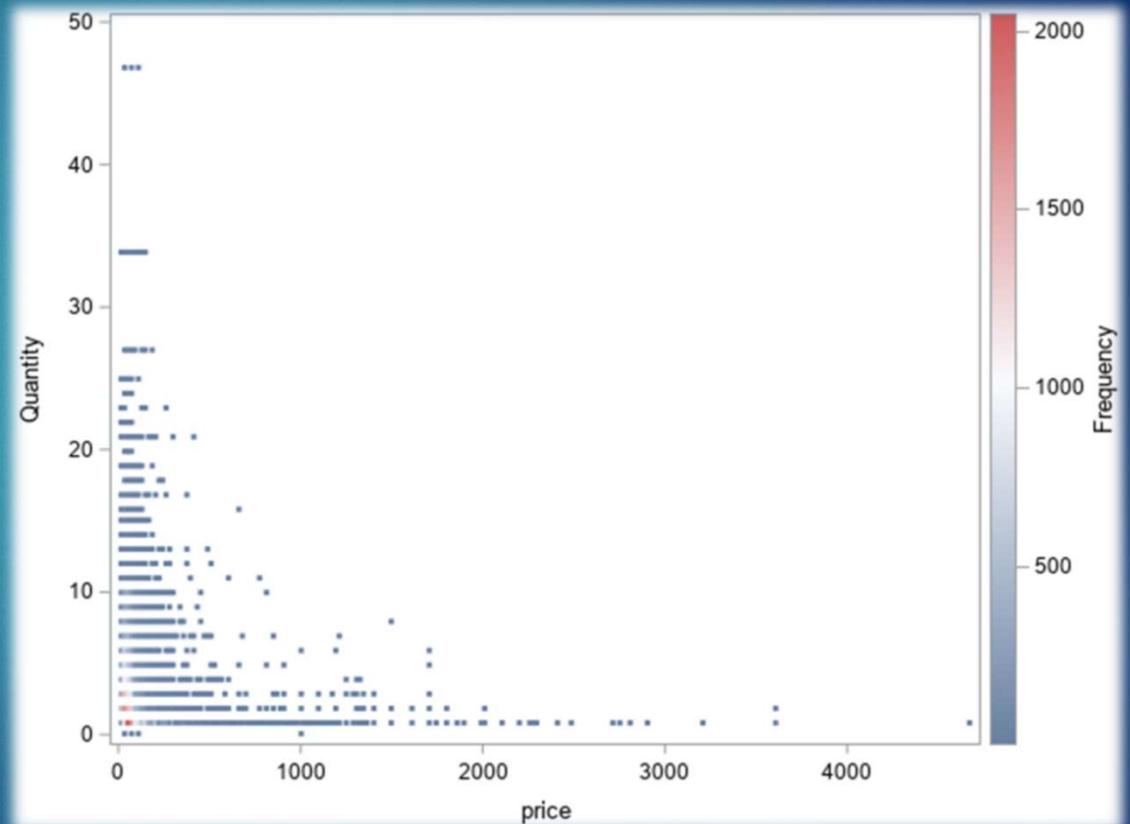
First time Buyer detail by Province

14

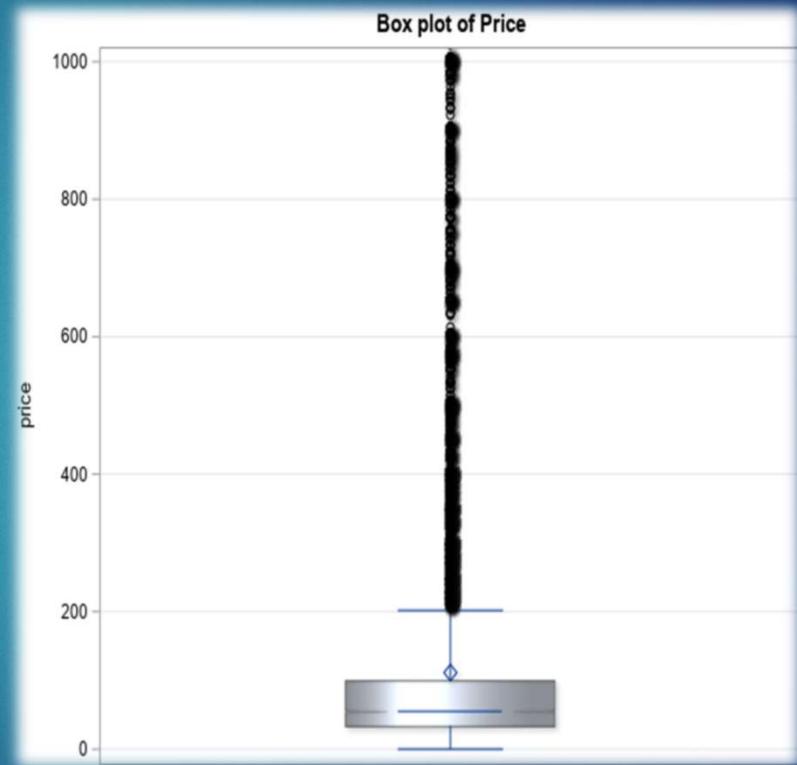
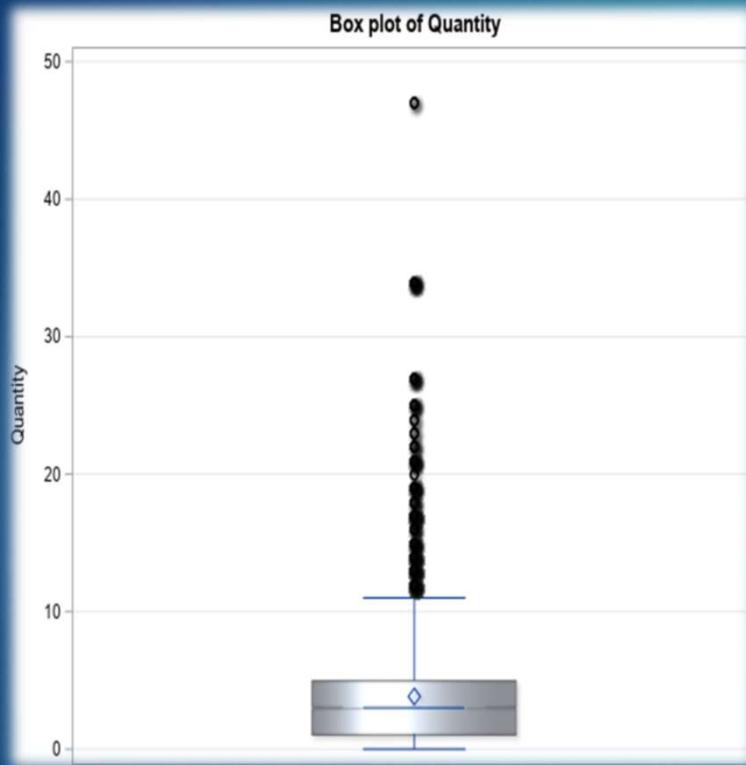


Heatmap of Quantity vs Price

- Most of the transactions quantity are below 10
- Majority of transactions are under 200 \$
- There is a direct relation of price and quantity, if price increases quantity is decreases



Box plot for Quantity and Price



Chi-square Statistics Of sources and First time buyers

Chi-Square Statistics for Source and Order first time or not

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of Source by Order_First_Time			
	Source	Order_First_Time		
		N	Y	Total
IVR	17	0	17	
	1.31	0.00	1.31	
	100.00	0.00		
	1.74	0.00		
REGULAR	769	254	1023	
	59.29	19.58	78.87	
	75.17	24.83		
	78.87	78.88		
	189	68	257	
WEB	14.57	5.24	19.81	
	73.54	26.46		
	19.38	21.12		
	975	322	1297	
Total	75.17	24.83	100.00	

Statistics for Table of Source by Order_First_Time

Statistic	DF	Value	Prob
Chi-Square	2	5.9814	0.0503
Likelihood Ratio Chi-Square	2	10.0645	0.0065
Mantel-Haenszel Chi-Square	1	1.6519	0.1987
Phi Coefficient		0.0679	
Contingency Coefficient		0.0678	
Cramer's V		0.0679	

Sample Size = 1297

Conclusions

- Sales goes up each month since last 1 year
- Ontario is the biggest market for retail sales
- People wants to go in store to purchase instead of web channels
- Canadians are spending more money in fashion & gems products

Recommendations

- Ontario is the cream market for retail sales, Hence to focus on this province will boost the sales
- In west cost BC and Alberta too are good revenue generated province, we can start new campaign to increase sales in those province like Ontario
- People are still not used to with online shopping so introduce more promotions and user friendly method for online shopping.

Thank You for your attention

