

Retail Sales Analysis using SAS

(A Data Science project)

By

Karan Singal

Under the guidance of: - Ar Kar Min



Project Report Outline:-

<u>Index</u>	<u>Page</u>
1. Introduction	2-2
2. Business objectives	2-2
3. Methodology	2-2
4. Study Framework	3-3
5. Descriptive Analysis	4-8
6. Results	8-8
7. Conclusion	8-8

1. 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

2. 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 customers are spending more money

3. Business objectives

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

4. Study Framework

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.

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	Dataset-1	Dataset-2
Number of Transactions	34288	1297
Order Year	May-2007 to April-2008	May-2007 to April-2008
Demographic Information	Not provided	City. Province and Postal codes
Unique Value	Customer ID	Customer Number
Important variables	Order Date, Quantity, & Category	Locations, First time buyers

Independent Variables

Customer ID

Order Date

Quantity

Category

Order First Time

Dependent Variables

Province

Sales Price

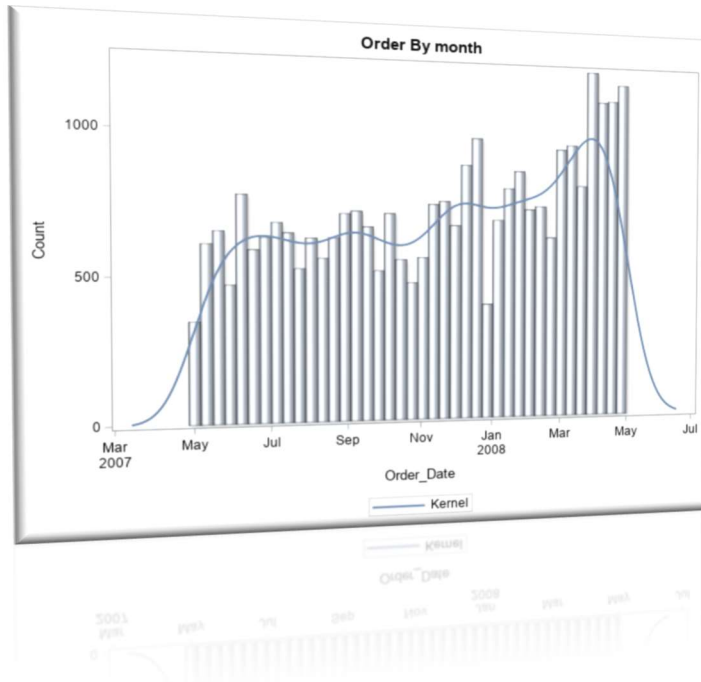
5. Descriptive Analysis

- Remove all the duplicate records
- Count missing values and remove unwanted records
- Merge both the datasets with unique values for more detailed analysis
- Drop unnecessary columns and prepare final Dataset



The overall Scatter plot gives a brief details about relationship with variables

Sales by each month groups



- As per the histogram we it can be predict that each month the sale goes up

Total Sales by Source

51%

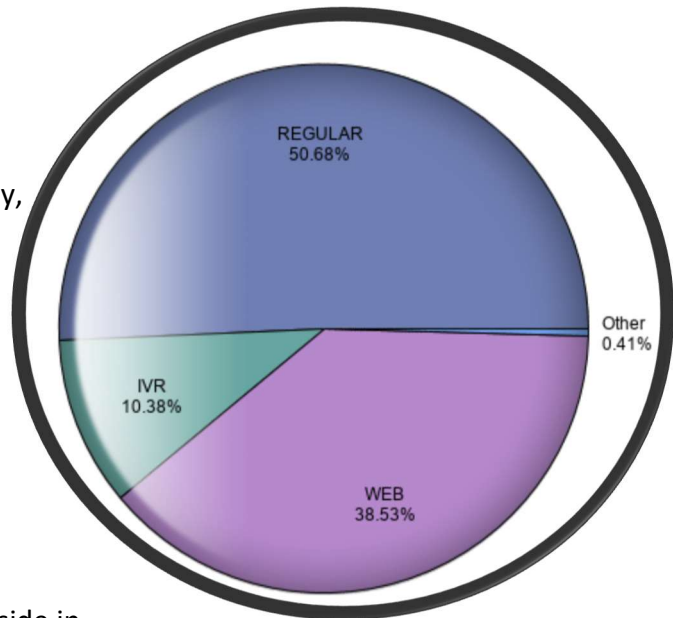
Regular :- Canada majority buyers wants to do sopping in traditional way, they like to visit stores for shopping

39%

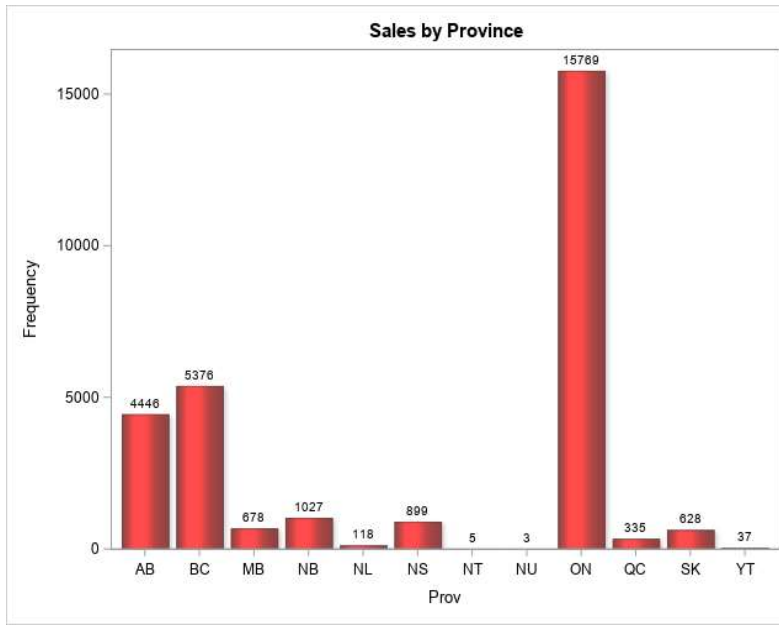
Web:- However people are getting used to with online shopping, as the online website like Amazon and eBay

10%

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



Number of Sales by Province



53.8 %

Ontario

18.3 %

British Columbia

15.2 %

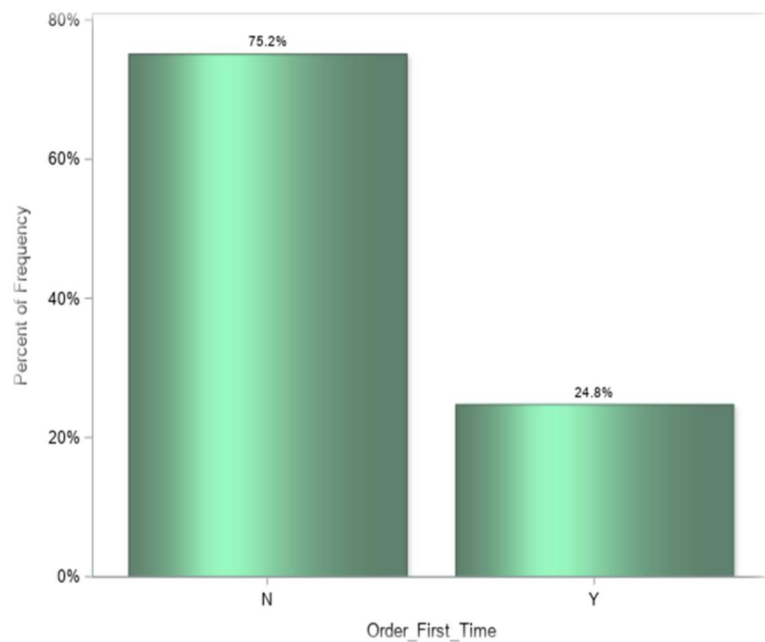
Alberta

12.7 %

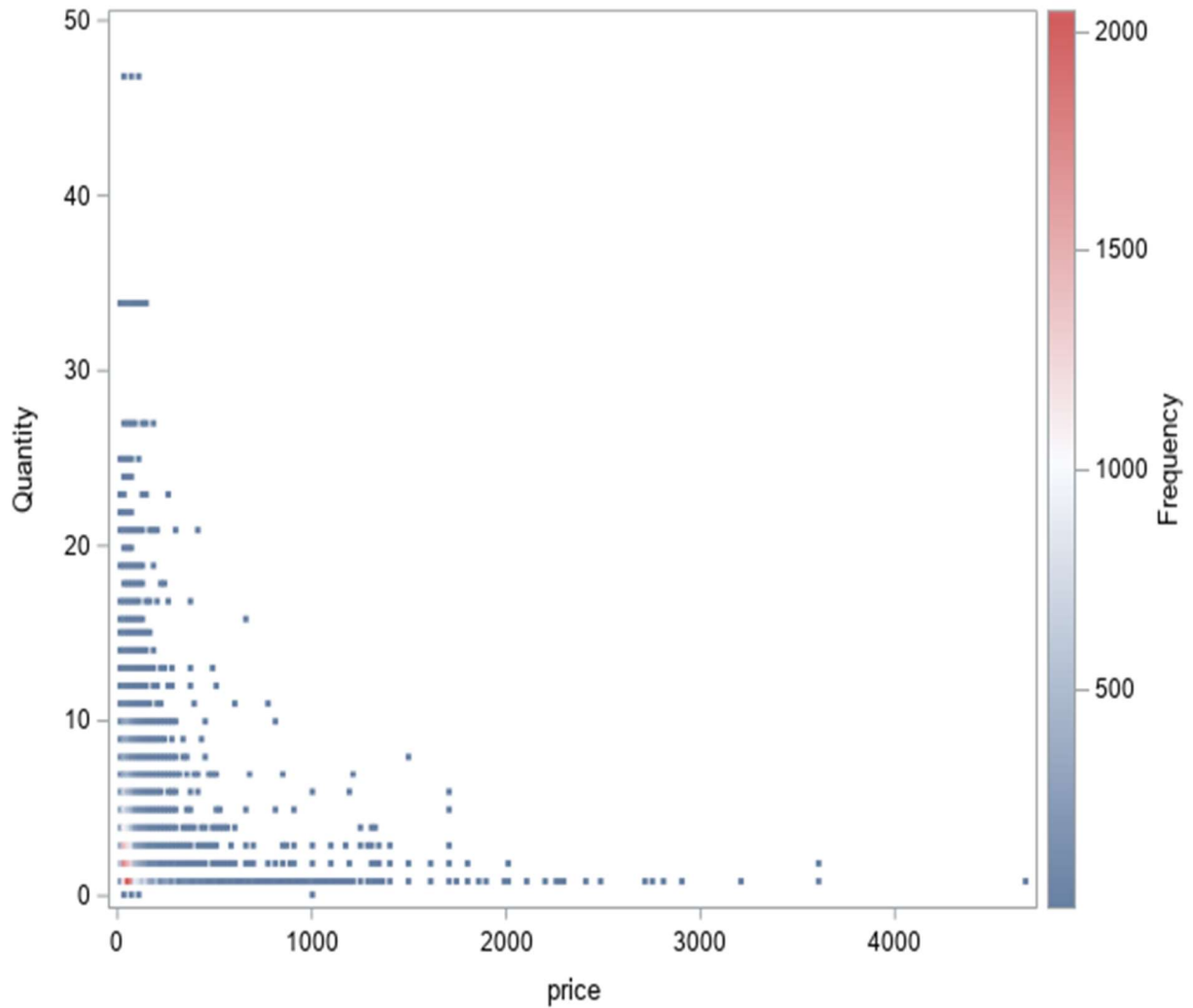
Others

First time Buyer detail by Province

- As per the analysis it is strongly recommended that new offers or gift coupons should be given for almost 25 % customers who never come back for shopping again



Heat map 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

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		Total
		N	Y	
	IVR	17 1.31 100.00 1.74	0 0.00 0.00 0.00	17 1.31
	REGULAR	769 59.29 75.17 78.87	254 19.58 24.83 78.88	1023 78.87
	WEB	189 14.57 73.54 19.38	68 5.24 26.46 21.12	257 19.81
	Total	975 75.17	322 24.83	1297 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

6. Conclusion :-

- 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

7. Recommendations:-

- Ontario is the cream market for retail sales, Hence to focus on this province will boost the sales
- In west coast 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
- A gift voucher or a couple should be offered to almost 25% buyers who don't come again for shopping