Date: 6th Oct, 2017

REPORT

Data Analysis and Visualisation

ANALYSIS ON SUPERMARKET TRANSACTION

Submitted By: **Submitted To:**

Ajantika Paul   **Prof. Sam Adhikari**

                                                              RBS-Newark

Introduction

Without customers, a business cannot survive. Therefore, it is important to understand their needs and how to motivate them to buy products in order to increase the sales. One must be clear and well trained on how to deal with the competition in the market and on providing the best solutions to customers to attract them towards one's business.

Here, we have performed Data Predictive Analysis on Transaction Processing Data (TPS) of supermarkets to come up with the strategy to increase their revenue.

Categories which can increase revenue

Let's study various categories that can help the supermarkets to increase their revenue.

Analysis 1: Annual Income vs Revenue

Annual Income is the first parameter to increase the revenue of the supermarkets. The customers with higher income levels are likely to spend more hence increasing the sales. This will help supermarkets to increase their revenues.

$$y = 68.784x^3 - 817.92x^2 + 2638.6x - 2281.91$$
$$R^2 = 0.778$$
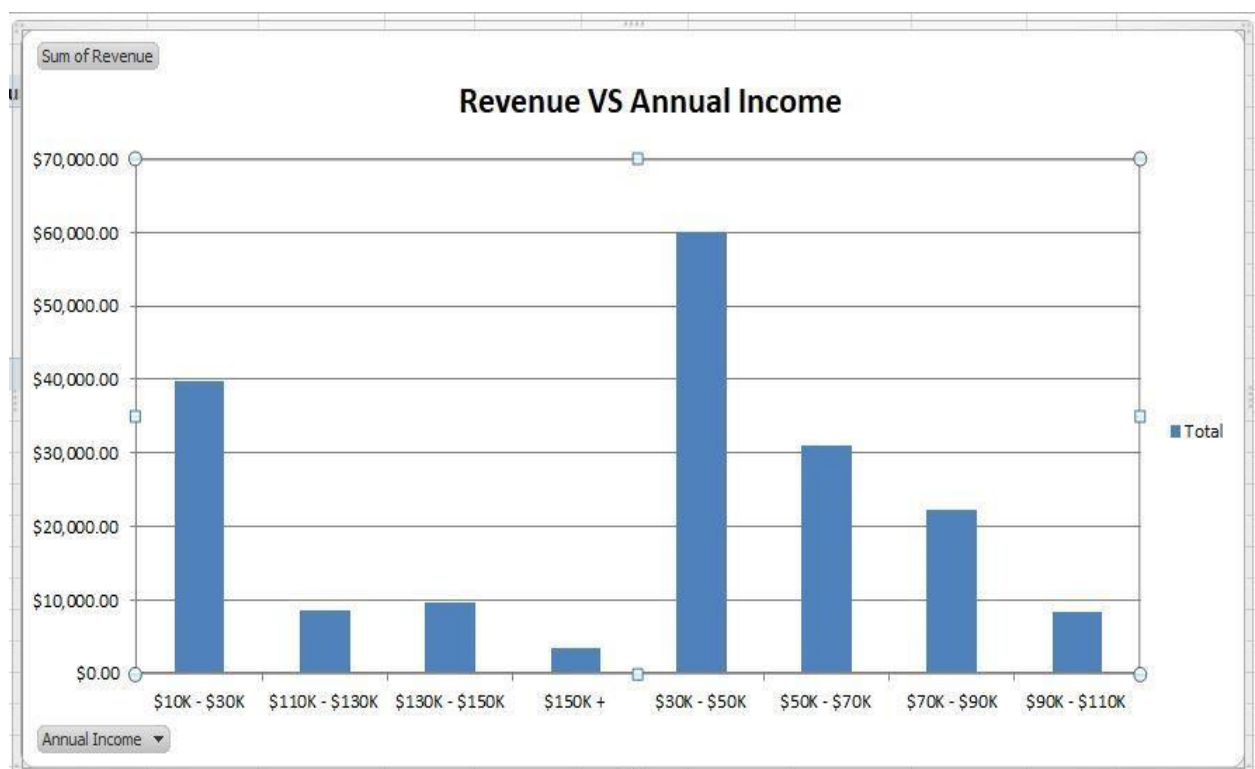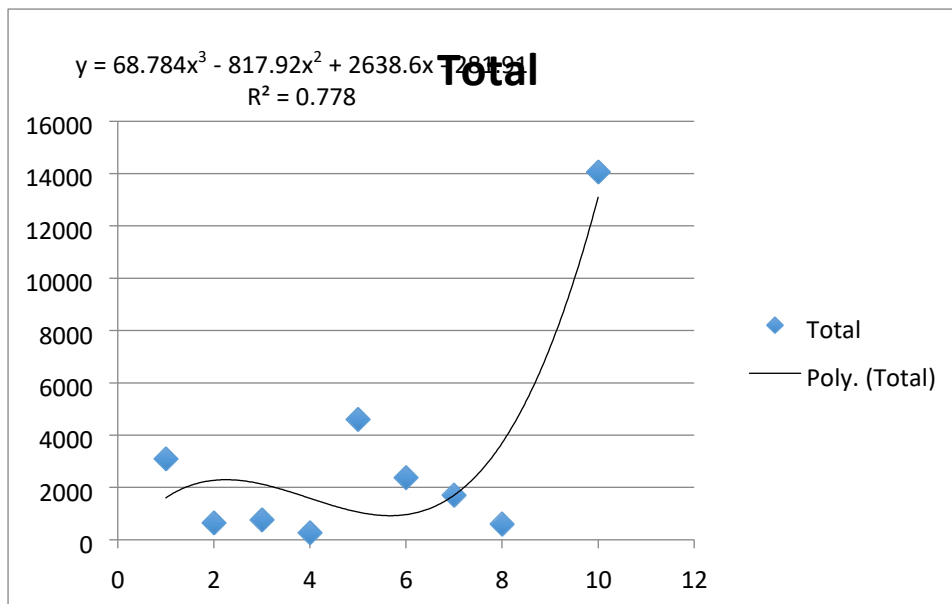
Observation:

- Highest spender in the supermarket falls in the Annual Income group of $30 - $50K
- This group is spending maximum on Vegetables.

Inference:

- We can offer benefit coupons to these buyers which they can utilize to buy other products at supermarket as well, this will increase their reach and thus cumulative revenues.
- These buyers being our regular customers can be offered Subscription plans for buying vegetables that will improve the fixed income of the store as well as help in customer retention.

Analysis 2: Number of Children vs Revenue

The other parameter could be the number of children in a family. The more number of children can lead to more expenditure further leading to increase in revenue for a supermarket.

$$y = -4.6019x^3 - 211.63x^2 + 1698.1x - 52.333$$
$$R^2 = 0.9258$$
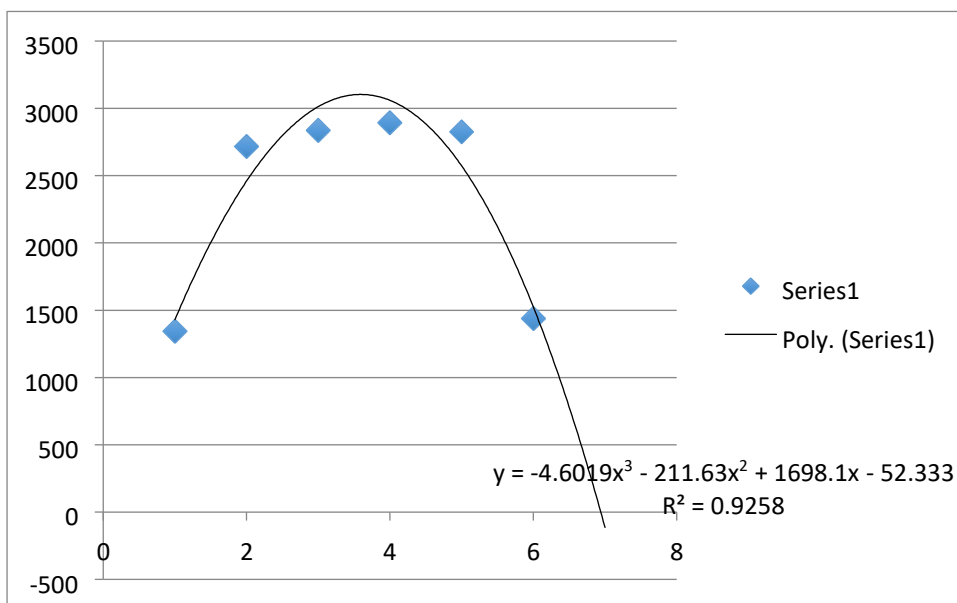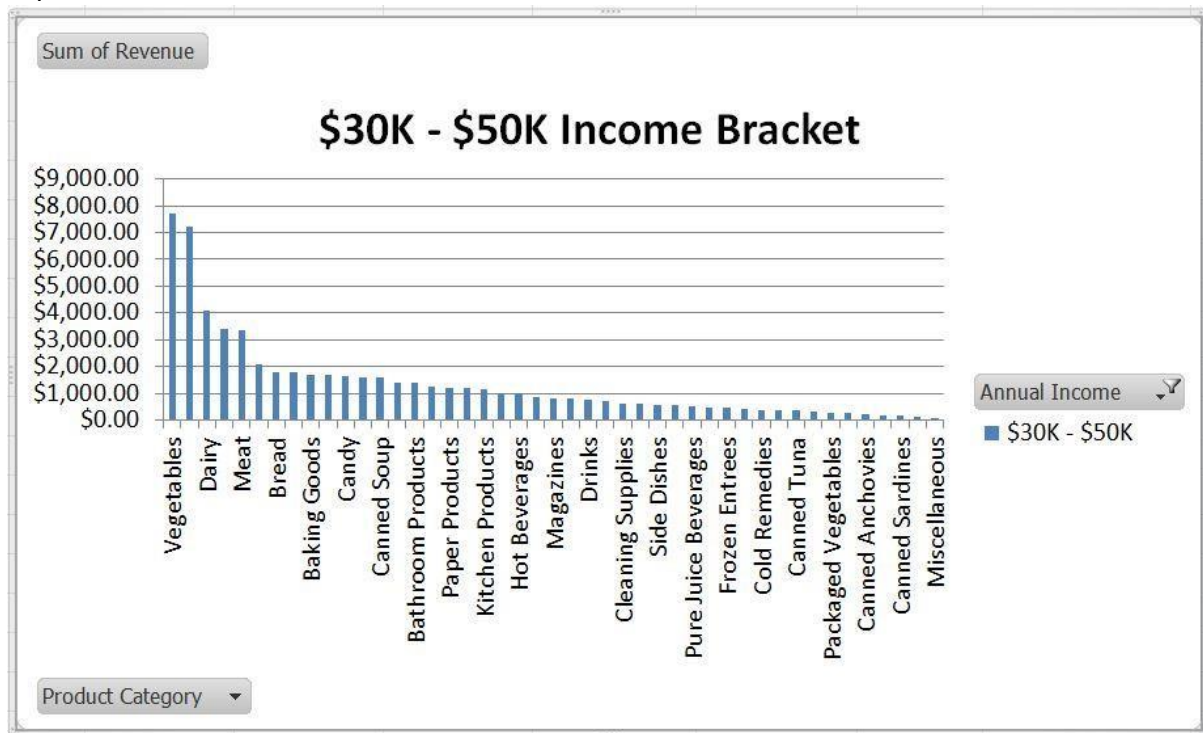
Observation:

- Buyers with children are giving more revenues to the supermarket than buyers with no children.
- A buyer with 3 and 4 children are making maximum purchases.

Inference:

- A buyer can be affected by the presence of their children in the store and this can disturb their spending potential. Thereby, we can introduce playing and gaming area at the supermarket where children can spend their time while the parents are shopping. This will also promote our supermarket as a family shopping place.

Analysis 3: Missing Product Categories

The next parameter to be considered is missing products in the supermarkets which can lead to decrease in revenue. Those missing products shall be introduced as per the customer's demands and this could help them in increasing their revenues.

Count of Units Sold Product Category	State or Province BC	CA	DF	Guerrero	Jalisco	OR	Veracruz	WA	Yucatan	Zacatecas	Grand Total
Cold Remedies	6	20	7	2		13	1	31	3	10	93
Decongestants	4	22	5	2		14	5	23	4	6	85
Bathroom Products	19	75	27	16		56	6	122	14	30	365
Canned Sardines	1	11	2			7	3	13	1	2	40
Plastic Products	7	24	8	6		28	6	40	10	12	141
Canned Shrimp	3	9	2			5	3	10	1	5	38
Candy	18	71	28	6		60	7	113	14	35	352
Cleaning Supplies	14	33	8	3		34	7	67	3	20	189
Canned Clams	4	9	3			11	2	18	3	3	53
Bread	26	87	20	14		74	9	121	27	47	425
Candles	3	14	2			6		18	1	1	45
Hot Beverages	10	46	13	3		33	10	82	9	20	226
Canned Oysters	1	8	2	1		4		11	3	5	35
Miscellaneous	2	7	2			10	3	16		2	42
Canned Anchovies	7	11	3	2		6	1	9	2	3	44
Packaged Vegetables	1	8	1	3		6	1	23	3	2	48
Drinks	6	33	8	1	1	14	2	48	6	16	135
Side Dishes	6	32	8	6	1	18	7	58	5	12	153
Frozen Entrees	8	20	6	2	1	22	1	42	4	12	118
Beer and Wine	24	73	22	20	1	62	16	97	10	31	356
Hygiene	12	31	12	1	1	40	10	71	5	14	197
Pain Relievers	10	35	5	4	1	21	11	84	6	15	192
Magazines	8	42	13	3	1	29	7	63	8	28	202
Pizza	13	37	14	3	1	21	7	71	10	17	194
Breakfast Foods	28	86	24	10	1	55	8	134	22	49	417

Total

$y = 5.0581x^3 - 156.09x^2 + 1299.7x - 1882.9$
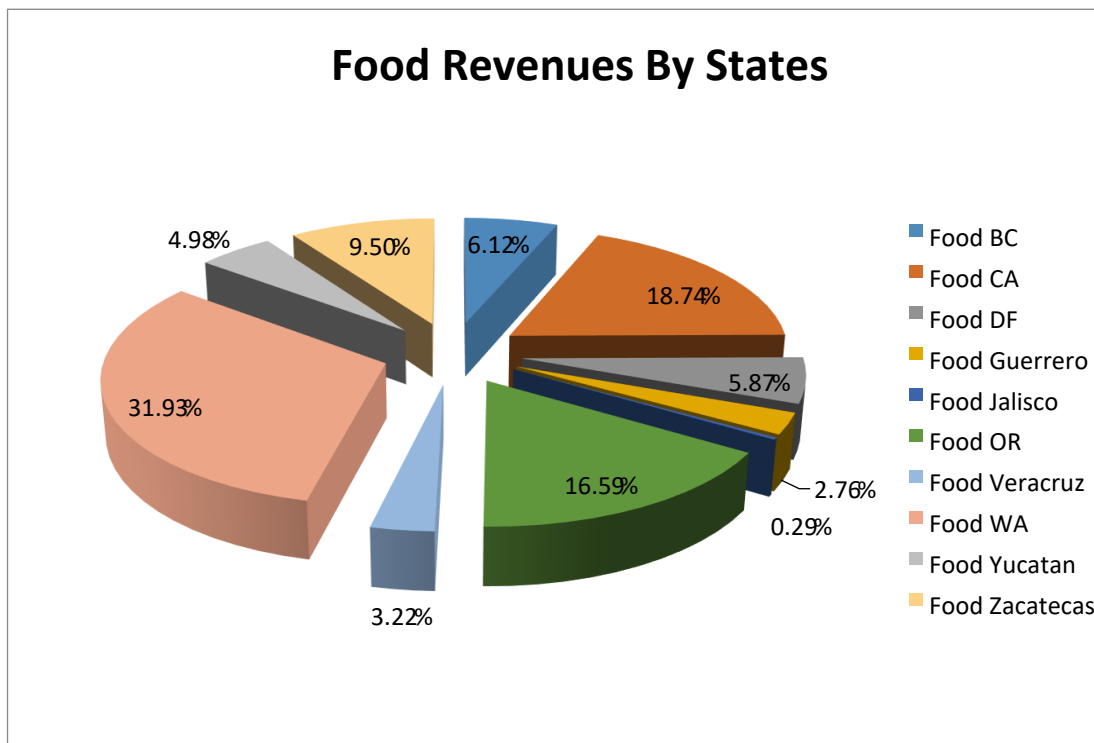$R^2 = 0.5936$

Total

Poly. (Total)
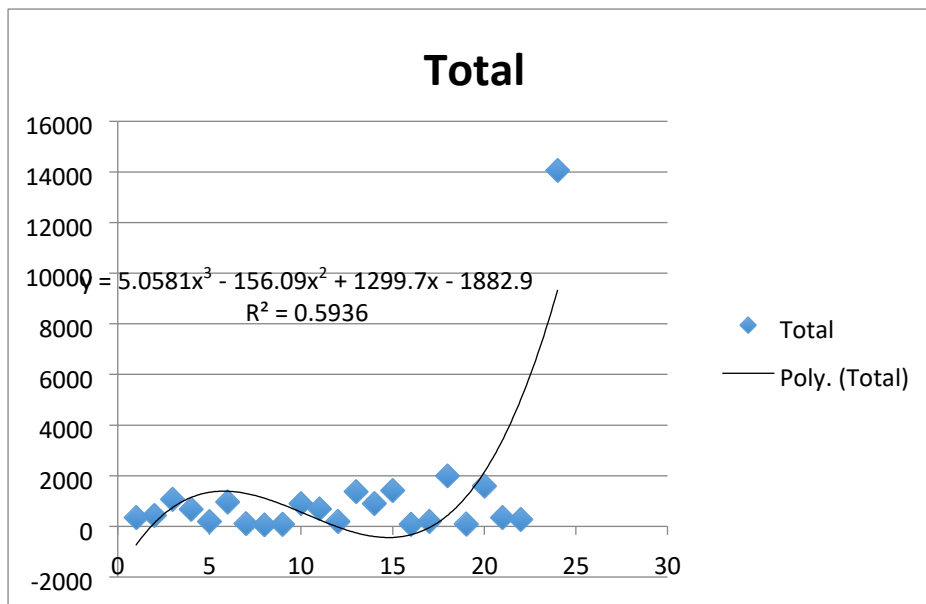
Observation:

- Some states/provinces like Jalisco and Guerrero are not having a few product categories for example Canned products, Bathroom products etc. in their stores.

Inference:

- Introducing these products in the stores can increase the overall revenue.

Food Revenues By States

Legend:
- Food BC
- Food CA
- Food DF
- Food Guerrero
- Food Jalisco
- Food OR
- Food Veracruz
- Food WA
- Food Yucatan
- Food Zacatecas

Values shown: 6.12%, 18.74%, 5.87%, 2.76%, 0.29%, 16.59%, 3.22%, 31.93%, 4.98%, 9.50%

The fourth parameter is the availability of food items in the supermarket that helps in increasing the revenue of the supermarkets.
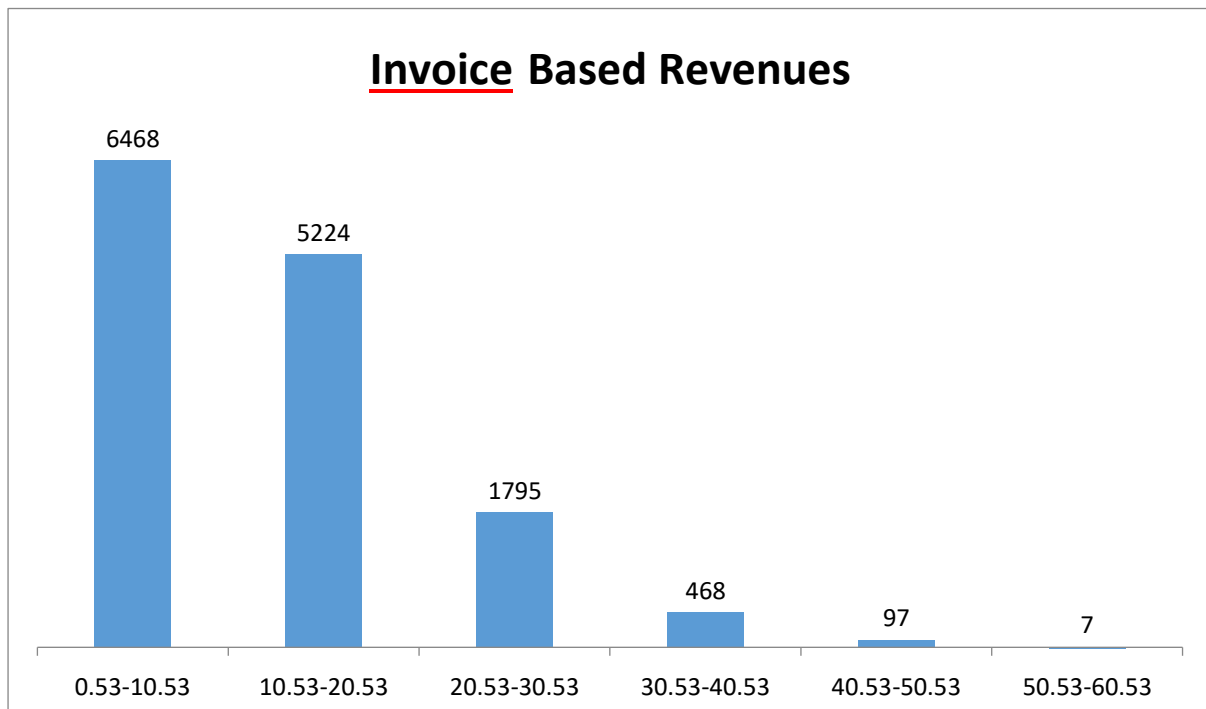
Observation:

- As per the statistics, food contributes maximum to the revenue of supermarkets.
- We can tie up with new vendors to compare the quality differences, by providing free samples to the customers.

Inference:

- By increasing the variety of food items can make the customers buy those items.
- The food item with the maximum share in the market can be made available in different brands which are not available with other stores so as to increase the revenue.
- Increasing the number of supermarket chains in a particular state will increase the overall revenue.

Invoice Based Revenues

Category	Value
0.53-10.53	6468
10.53-20.53	5224
20.53-30.53	1795
30.53-40.53	468
40.53-50.53	97
50.53-60.53	7

The Invoices are categorized on basis of amount of bill that customers have on average.

Inference:
- By offering discounts on high cost products, their sales/revenue can be increased.
- By implementing marketing strategies, we can increase sales of high cost products. Common platforms can be digital marketing, usage of social buttons, promo videos, reviews of products.
- If the supermarkets start recommending other brands, then permanent customers will be inclined to buy it.

Conclusion

By performing Data Predictive Analysis on Supermarket Transaction data, we have concluded that annual income is the most revenue generating factor. For analyzing the data, we have used Data Analytics Tools such as Pivot Table, Pivot Charts, Scatter Plots, Pie Charts and Bar Graphs in Excel. On the basis of the statistical data derived from the above-mentioned parameters, we can increase the revenue of the Supermarket by establishing multiple supermarket chains, offering discounts to our regular customers, by introducing new catchy products, by making the unavailable products available.