



DATA ANALYSIS & VISUALIZATION

HW 1: Supermarket Transaction - Business Intelligence and Big Data Analytics

Group 3

- Abhilash Basuru Yethesh Kumar
- Abhinaya Krishnamachari
- Annapoorna Chandrashekar Kadur
- Darshan Dalvi
- Frida Lin (Rong-Chian)
- Gokulramanan Soundararajan
- Harjeet Singhs Jaggi
- Pradnya Prabhudesai
- Supriya Nanjundaswamy

Introduction

Supermarket transaction data provides a lot of insights, because it contains information about customers and products purchased. If we further analyze trends and patterns in the buying habits of customers, we can then target customers more effectively and ,in turn, create actionable strategies accordingly.

Analyzing Supermarket transaction data allows us to have a complete landscape of customer base. Supermarket Transaction data is collected in order to create customer profiles, which can be used to promote product to target customers. Information inferred from this data can not only be used on advertising and building rewards or loyalty program but also be used to create specific shopper incentive programs that are aimed at a targeted group.

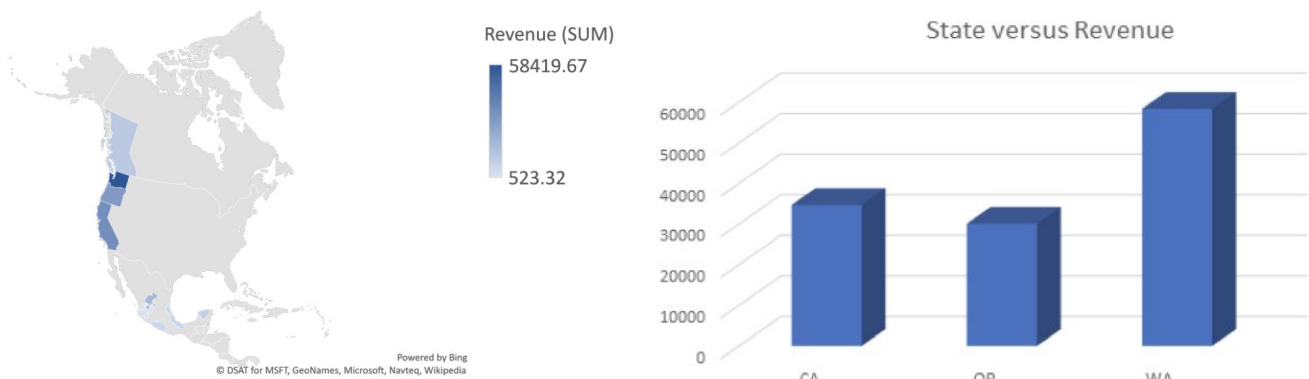
Our marketing/sales strategies were created on the following aspects:

- Create Customer Profiles and Find out Premium Customers
- Examine cities where sales decreased in 2008-2009
- Analyze profit of product purchased

The Aspects for increasing the revenue

1. Location Centric:

We decided to examine the performances of where we invest most -- the states which have largest number of cities with stores -- to see whether there's anything we can do to improve the performances.



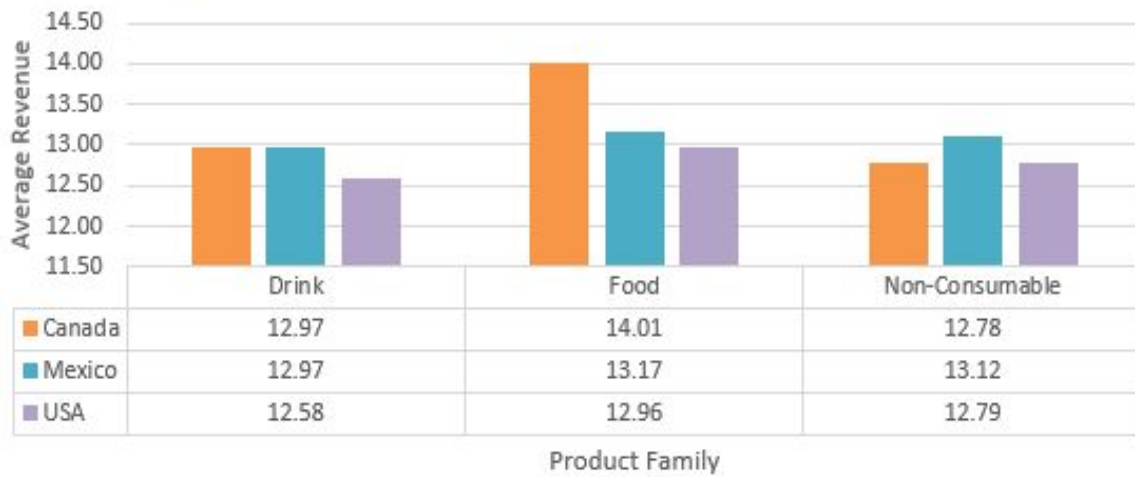
	CA	OR	WA
Number of cities with stores	4	2	7
Distinct customers	1651	879	1392
Total Revenue	34,730.41	30,138.44	58,419.67
Rev. generated per customer	21.04	34.29	41.97
Population estimated in 2009	36,961,664	3,825,657	6,664,195

Oregon could be the most profitable state

Washington state has more stores and less distinct customers than California, and each customer in WA generated almost doubled the amount of revenue generated by customers in CA. We can infer that the customers in WA are more loyal to the stores because of convenience (several stores can be found), so the money was worth investing.

California has four stores, compared to Oregon, which has only two stores, but it only brought four thousand more revenue to CA. Supermarkets in OR are obviously the highest revenue generators across the United States. Opening more stores in OR can be one of the considerations. Relatively poor performance in CA may be fixed by optimizing online shopping experience and home delivery service.

We look into different countries, as the graph shown below:



The graph shows the average revenue generated per transaction in Canada and Mexico are higher than that in USA. This result gave us the idea that, with further analysis of location and estimation of ROI, there could be some unmet needs in Canada and Mexico since people there tend to spend more at one time. Therefore opening up more stores in these two countries may be a good choice.

Other than that, We also look into unfavorable and surprising changes in performance of cities:



married: blue/ single: orange

[1] Decrease in Revenue



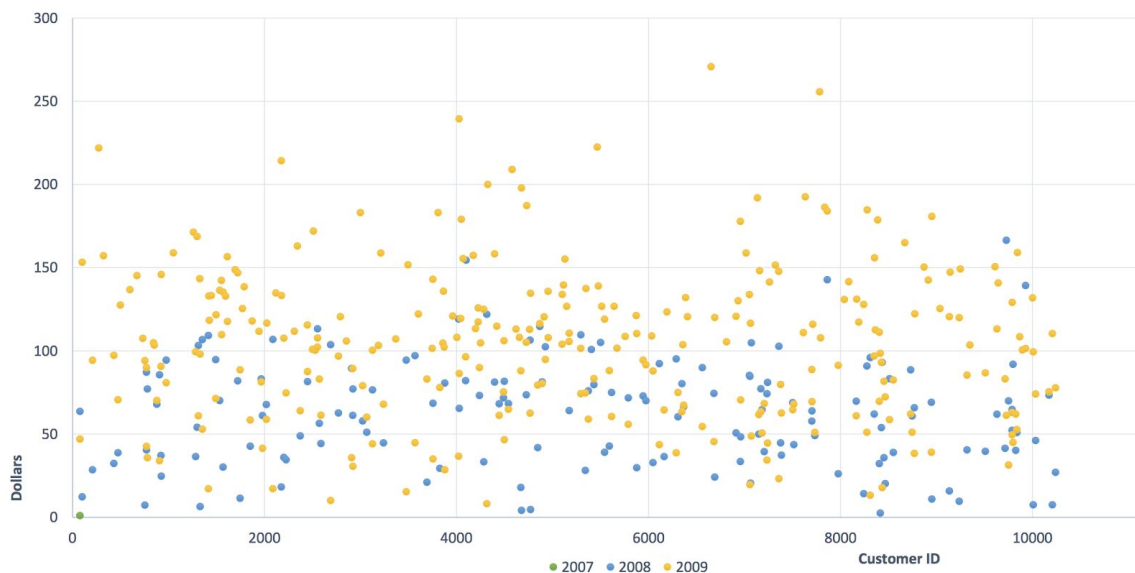
[2] Booming Sales

Cities Los Angeles and Salem seen decrease in revenue generated by both single and married people in year 2009. In Los Angeles, revenue from single people decreased 20% in 2009, whereas revenue from married people dropped 14% in Salem. Based on these results, further investigation can be done to analyse decrease in sales with the help of more data.

On the other hand, We saw sudden booming sales in Vancouver, Hidalgo, and Acapulco. Supermarkets in these cities performed relatively better than most of the other cities. It is a promising sign that could lead the supermarket to achieve even greater revenue if the sales team further analyze there's unmet needs in these cities.

2. Customer Centric

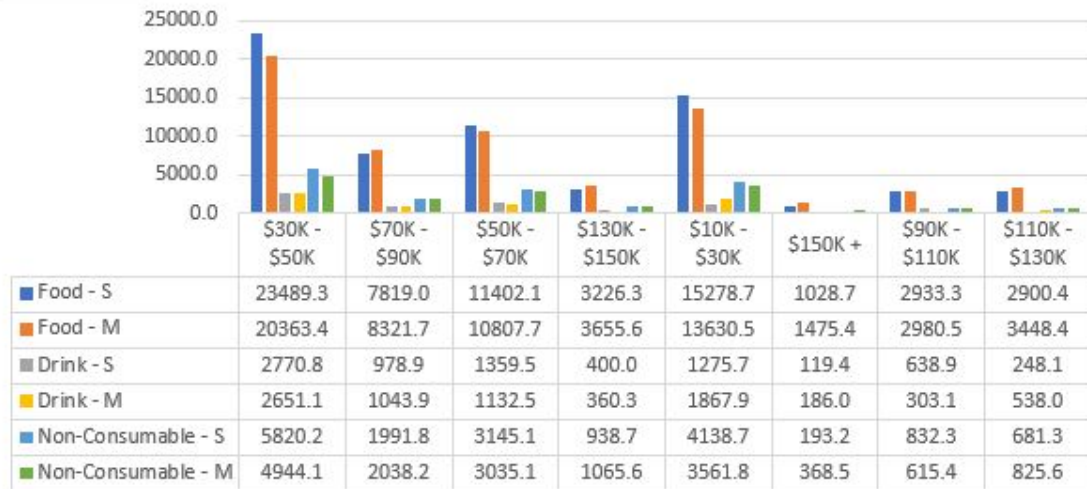
We aim at building consumer profiles and finding out the group of customers who come to store most frequently and therefore generate most revenue.



We define customers who spent more than \$100 in total as premium customers. They formed only Appx. 5.4% of total customer base yet generated Appx. 96.8% of the revenue. The average times they've purchased in the store is 4.2. We suggest to further analyze what product these people bought and offer incentives such as coupon based on their preferences accordingly.

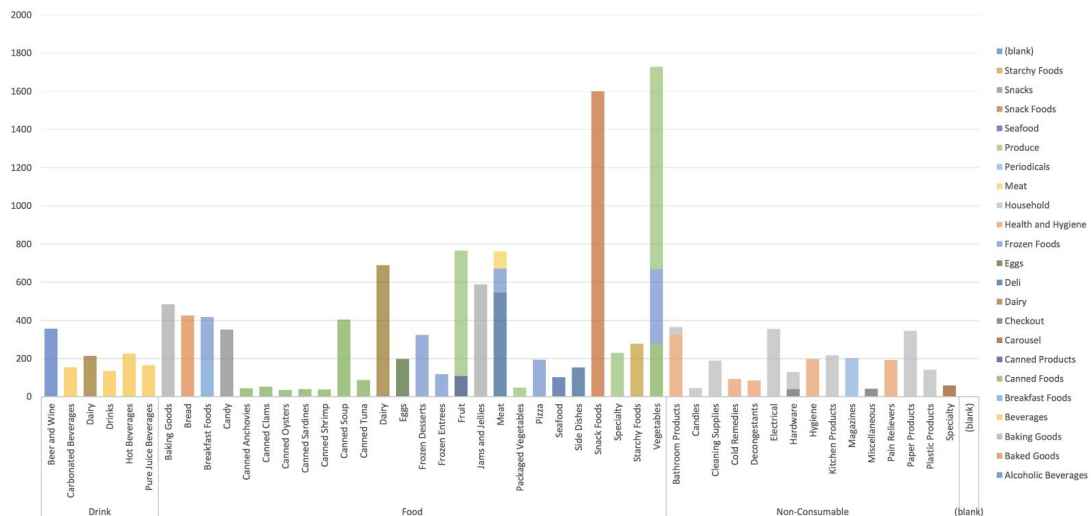
3. Product Centric

Product portfolio is a crucial part of supermarket transactions. The transaction data contains 3 product families comprised of 22 product departments which have around 50 products.



According to the graph above, We found out that 72% of the revenue is generated by Food Department in that 40% came from by people with annual income in range \$30K - \$50K and \$10K - \$30K. These people bring revenue in Food so personalized offers can be sent to this salary range customers. However, the revenue generated by Singles and by Married have little difference. We need not target particular group regarding this aspect.

Food items consisted the highest revenue among the product family. In this family, vegetables, snack foods and baking goods are among the highest revenue generators whereas canned products (fruits) are the least ones.



We suggest to either eliminate the canned products or try putting them on eye level around the checkout desk. Moreover, supermarkets can sell certain food products in the packs/bundle instead of single piece. This will be beneficial in following manner –

1. Attract customers with offers/discounts
2. Easy for customers to choose and pick
3. Optimize space utilization on shelf, warehouse and for transportation

Conclusion

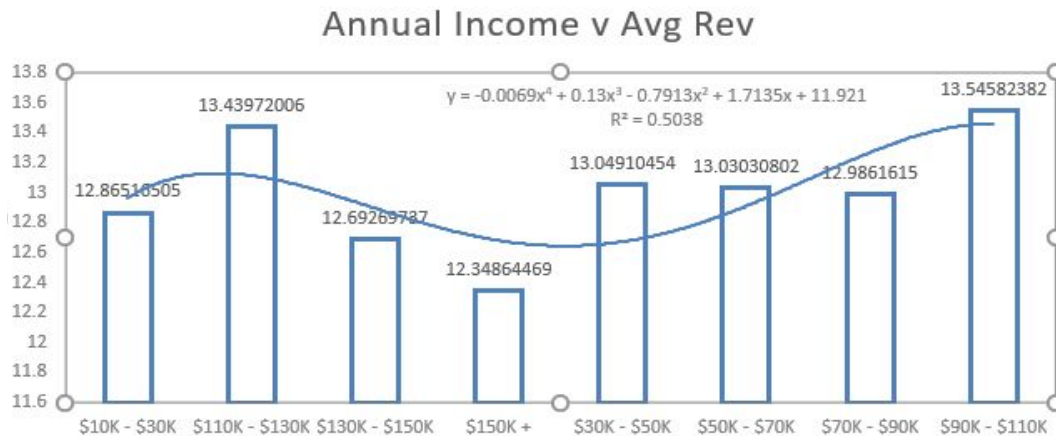
Based on our research and analysis we suggest the following changes in order to increase the supermarket's revenue:

- Open new stores in Oregon.
- Optimize online shopping experience and home delivery service in California.
- Focus on premium customers and develop suitable incentives individually.
- Certain food products can be sold in bundle instead of selling single piece.
- Eliminate or cut cost of canned food that brought little revenue.

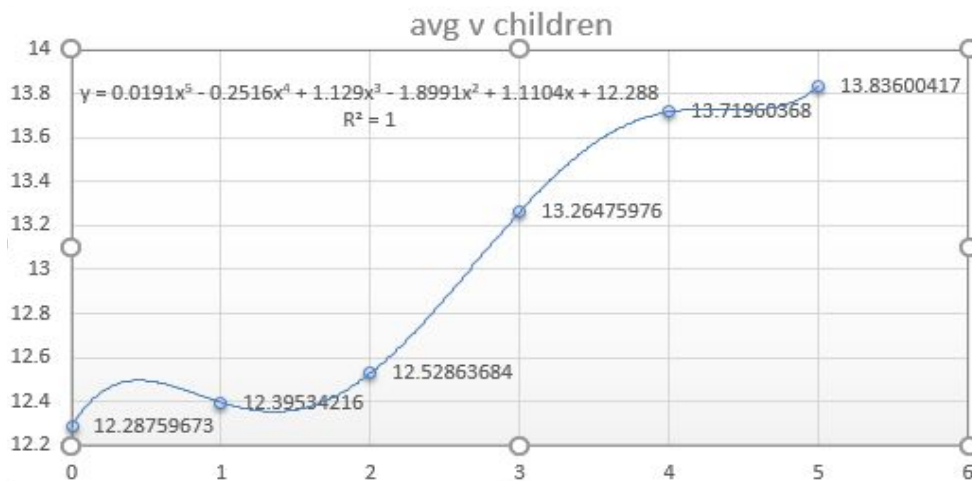
References

1. <http://www.gourmetads.com/>
2. <https://www.census.gov/>
3. <http://www.gofrugal.com/blog/2017/02/11/10-ways-to-boost-sales-in-your-supermarket/>
4. <https://www.quora.com>
5. <https://www.businessknowhow.com/marketing>
6. <https://toughnickel.com/industries>

Some Additional Information



From the above data, it can be seen that the average revenue is high from the customers who fall in the income range \$90k-\$110k and \$110k-\$130k. Since the number of customers who fall in that income range is less, further investigation can be done to find out if the population itself is less in those areas or if there can be any marketing strategies applied to improve the number of customers.



From the above graph, it can be inferred that the average revenue is more from five children family. Whereas the number of 4 children family is more and hence the revenue generated by them is more. Festival offers or any such customized offers can be given to the four children family to increase revenue.

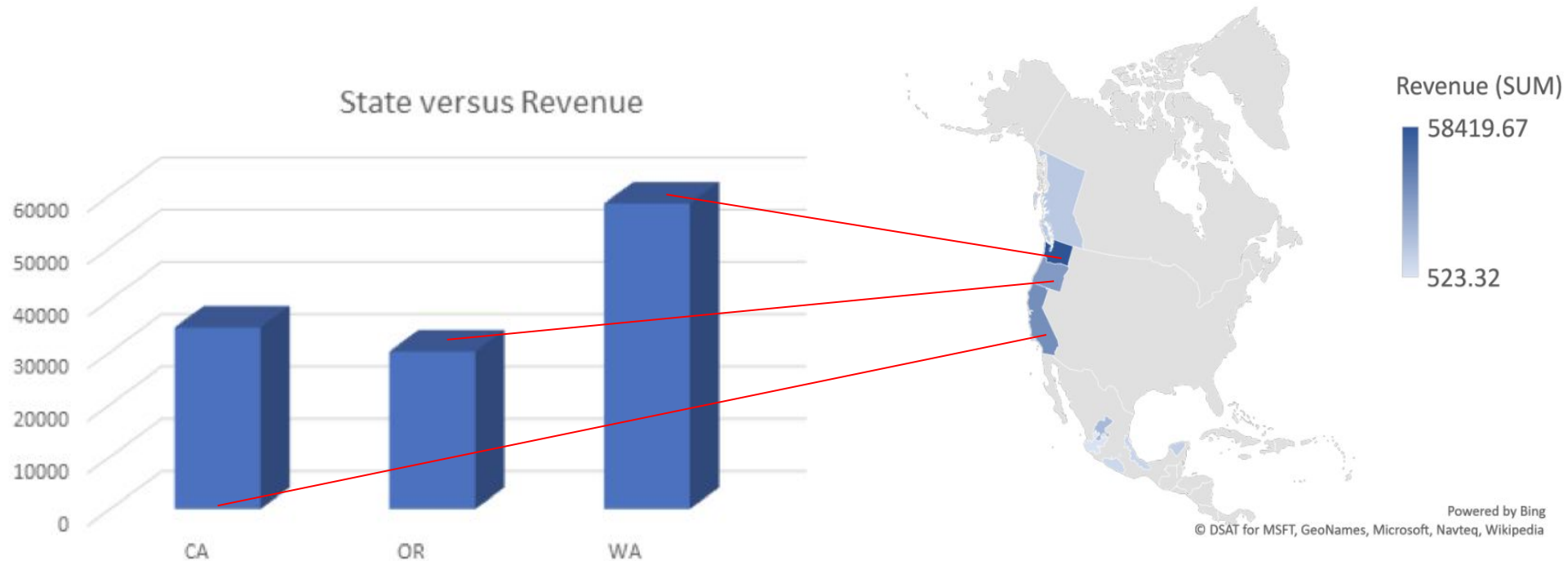


Data Analysis & Visualization

HW1: Supermarket Transaction Business Intelligence and Big Data analysis

- Abhilash Basuru Yethesh Kumar
- Abhinaya Krishnamachari
- Annapoorna Chandrashekar Kadur
- Darshan Dalvi
- Frida Lin (Rong-Chian)
- Gokulramanan Soundararajan
- Harjeet Singhs Jaggi
- Pradnya Prabhudesai
- Supriya Nanjundaswamy

Most Profitable State?

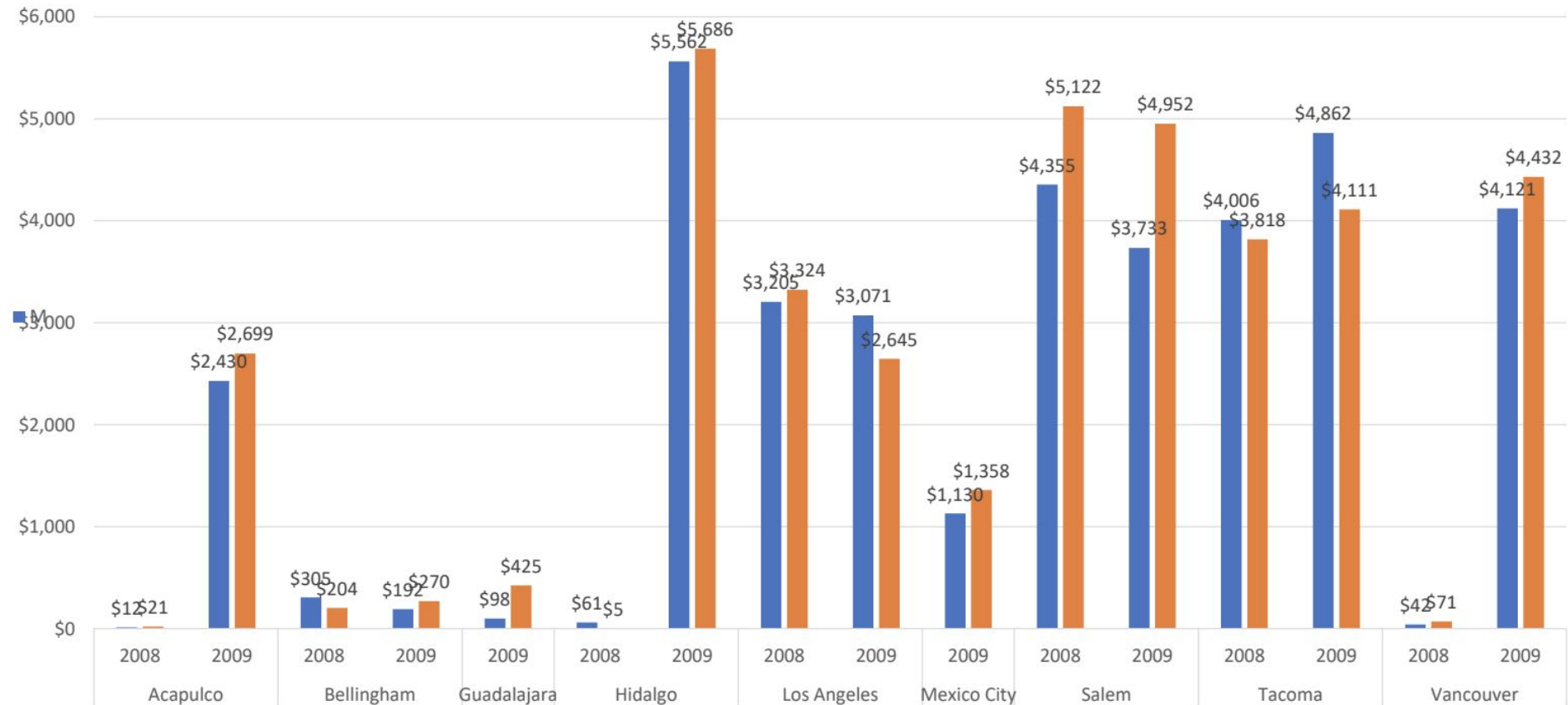


	CA	OR	WA
Number of cities with stores	4	2	7
Distinct customers	1651	879	1392
Total Revenue	34,730.41	30,138.44	58,419.67
Rev. generated per customers	21.04	34.29	41.97

Rev. per Transaction in Countries

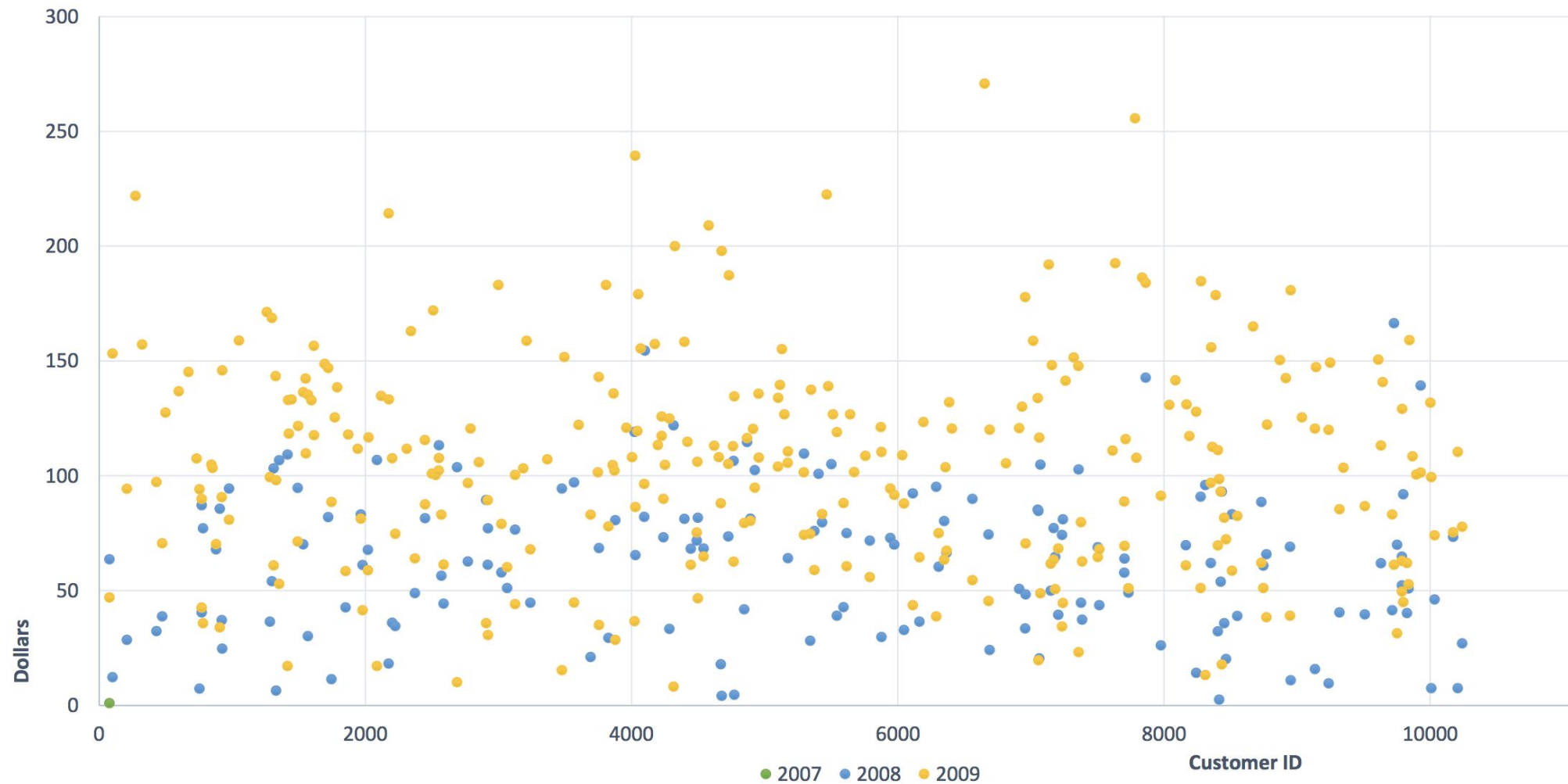


Trends Observed in Cities

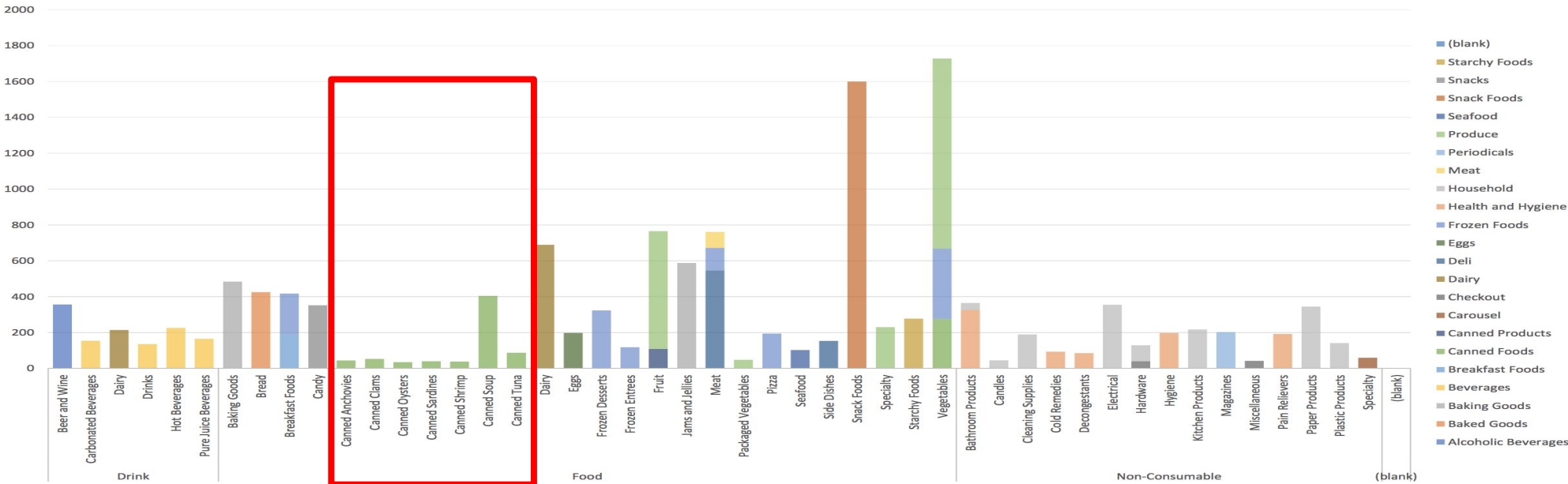
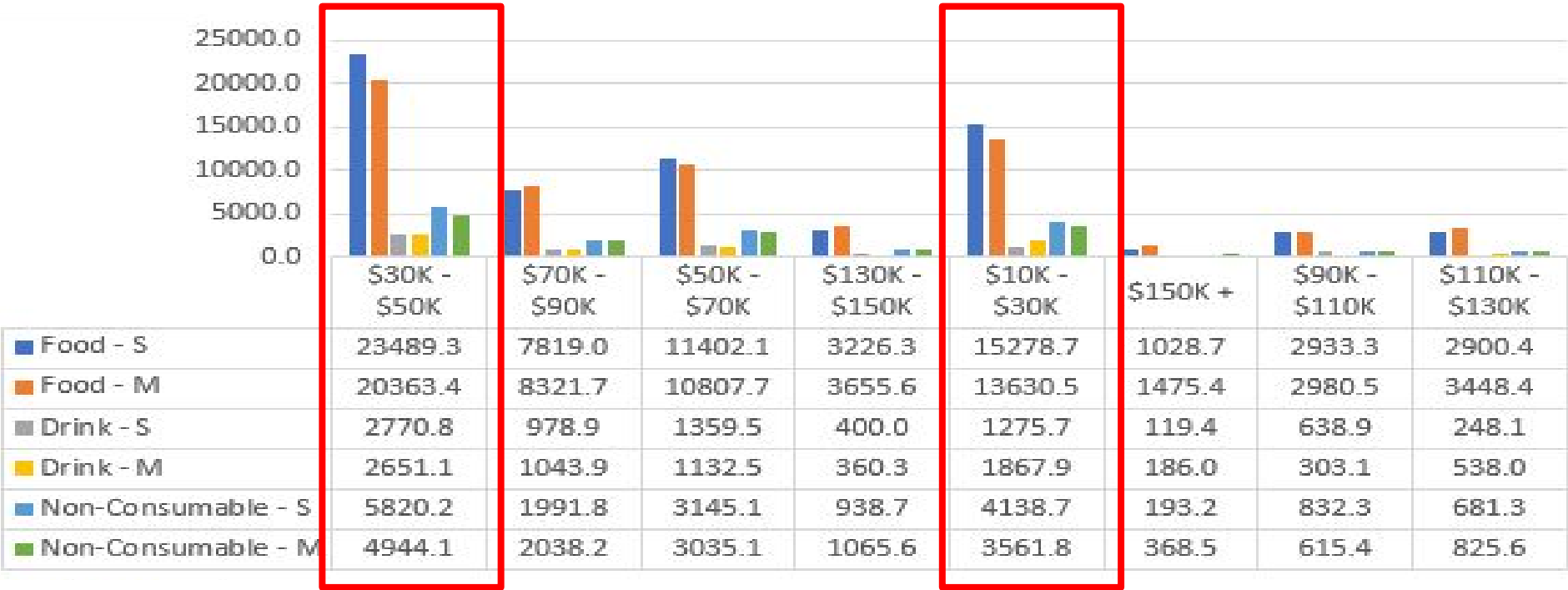


married: blue/ single: orange

Premium Customers



Product & Revenue



Thank You