

ANALYSIS USING TABLEAU

Table of Contents

Introduction.....	3
Data cleaning	3
Data exploration.....	3
Segment analysis.....	3
City analysis.....	5
Product analysis	7
Time analysis	9
Forecast trends	10
Identify patterns using correlation, trendlines, regression	11
Rating analysis	12
Correlation and regression	13
Group data.....	14
Data visualization using Dashboard.....	15
More Insights	16
Conclusion	17
Recommendations.....	17
References.....	18

Introduction

Sales, profit, revenue, and other variables are analysed for different product lines, cities, and customer types using Tableau. This tool is helpful in analysis, and creating visualization using graphs, charts, tables, and dashboards (Khatuwal & Puri, 2022).

Data cleaning

The process of fixing corrupted data, incorrect format, removal of duplicate information, or incomplete data can be considered as data cleaning (Ridzuan and Zainon, 2019). In the given dataset, there are no such issues as checked with data filtering, and manual screening. Hence, there is no need to apply any data-cleaning techniques.

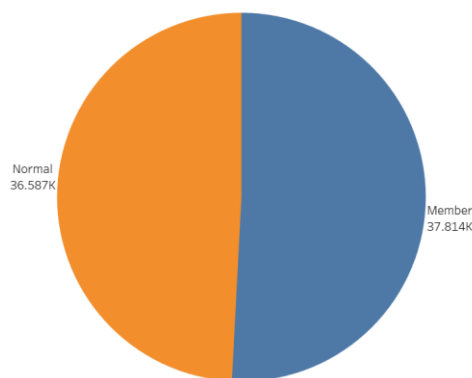
Data exploration

Data exploration is the analysis of data conducted by an analyst which can include statistical analysis and visualization (Qin *et al.*, 2020). For example, size, and quantity can be included in data exploration. However, the main visualization and statistical analysis are given in a particular section. The current data has 998 rows or datasets and 21 variables or columns. Among these, 12 variables are numerical variables such as revenue, total sales, profit, cogs or cost of goods. On the other hand, 9 variables are non-numerical which include customer type, date, time, city and others.

Segment analysis

Analysis of the customer segments in terms of different variables such as profit and revenue performance is done below.

Profit by Customer Segment



Customer Type

Member	37.814K
Normal	36.587K

Figure 1: Profit by customer segment

(Source: Self)

According to the above graph, over 36.58K of profit is generated by normal customers. Though members generate more profit which is over 37.8K. Let's look at the revenue from each customer segment.

Revenue by Customer Type



Customer Type

Member	56,613
Normal	55,482

Figure 2: Revenue by customer type

(Source: Self)

Members generate higher profits and revenue from them is also higher with a value of 56600 units of revenue. On the other hand, normal customers generate revenue of nearly 55500

units. Hence, members generate almost 1100 units more revenue than normal customers. This might be because; the members get better discounts so they are involved in more purchasing which eventually drives the sales of the company (Aguilar-Barrientos Villegas-Gomez, & Arias-Salazar, 2021).

City analysis

Analysis of profit, revenue, and shipping cost by cities are done here.

Profit and Quantity by City

City	Profit	Quantity
Montreal	25,976	2,017
Toronto	23,865	1,767
Vancouver	24,560	1,870

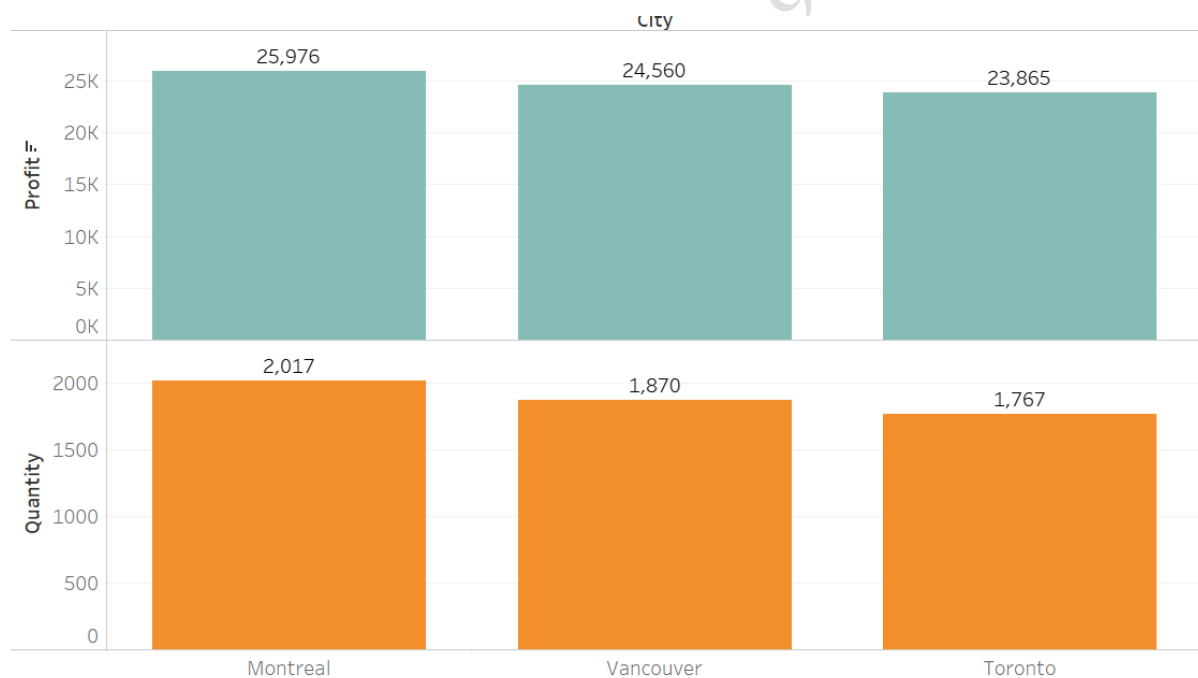


Figure 3: Profit and Quantity by City

(Source: Self)

The graph showed that Montreal is generating the highest profit of 25,976 units which is followed by Vancouver and Toronto with 24.56K and 23.865K respectively. Both of these cities also sell in low quantities compared to Montreal. 2017 units of Goods are sold from

Montreal in 2023 but only 1870 units and 1767 units of goods are sold from Vancouver and Toronto respectively. Now, the same cities are analysed as per revenue and sales generations.

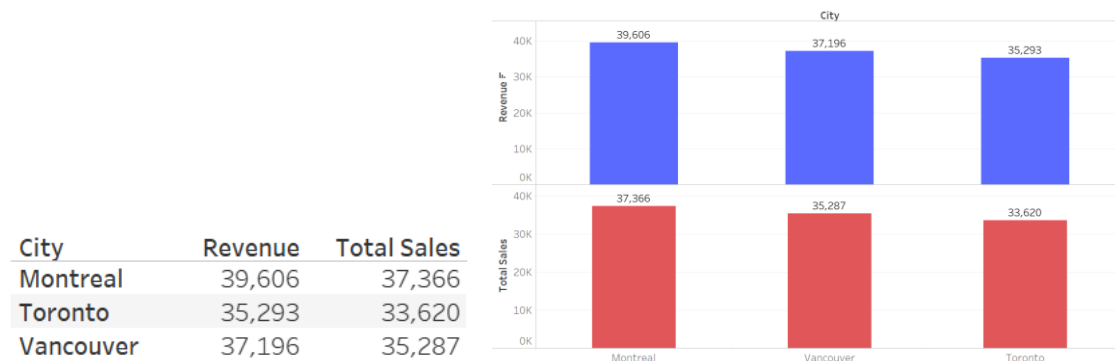


Figure 4: Revenue and Sales by Cities

(Source: Self)

Montreal generates over 39600 units of revenue which is the highest and its total sales value is also the highest with more than 37,366 units. In both cases, it is followed by Vancouver and Toronto respectively.

Now, the cities are analysed for shipping cost but each customer segment hence, it can be treated as another segment as well.

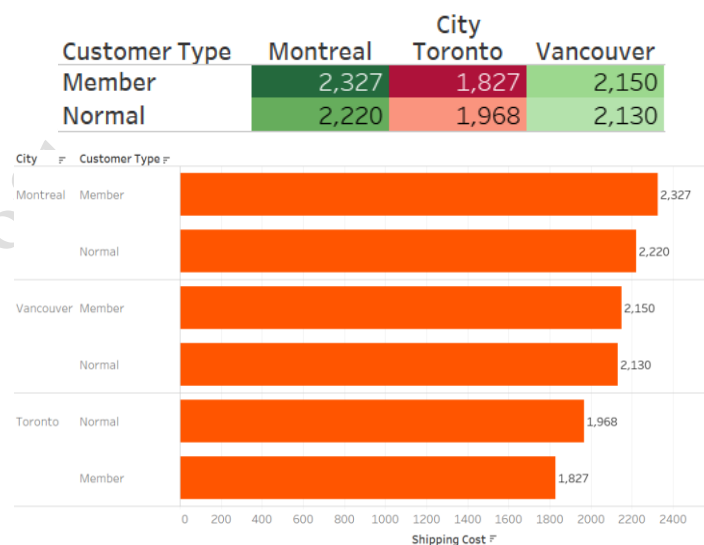


Figure 5: Shipping cost by cities for each customer segment

(Source: Self)

With higher profit and total sales, Montreal is the top city for the highest shipping cost. The shipping cost is 4547 units in Montreal. This is followed again by Vancouver with a total shipping cost of 4280 and 3795 for Toronto. However, it is high for members in Montreal and Vancouver, but lower for them in Toronto when compared with the Normal customer segment.

Product analysis

Revenue and quantity as per products are analysed in this section.

Product	Quantity	Revenue
Bedding Set	454	8,528
Decorative Candel..	370	7,579
Designer Handb..	577	11,775
Fresh Juices	402	7,932
Healthy Snacks	402	7,966
Home Decor	545	10,447
Organic Tea	375	7,584
Portable Power ..	556	11,088
Smartphone Cha..	482	9,722
Sunglasses	537	11,276
Wireless Headph..	431	8,550
Wristwatch	523	9,648

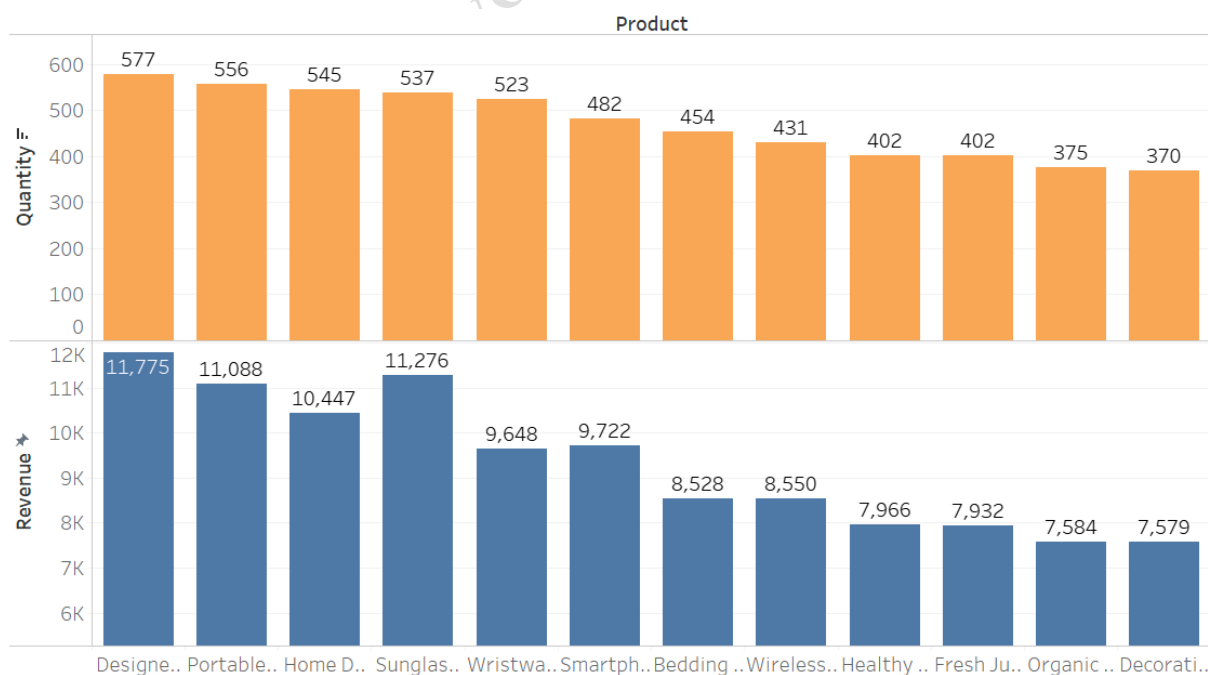


Figure 6: Quantity sold and revenue generated as per product

(Source: Self)

Hence, a designer handbag is sold most by quantity (577 units) and generated a revenue of more than 11,770 which is the highest. This is also a fashion product, which is the most selling product line as given in Figure 7. The second highest-selling product is the portable power bank from the electronics product line by quantity. However, by revenue, Sunglasses are the second highest with a value of 11,276 units. Decorative candles from the home and lifestyle category are worse-selling products. However, food and beverage is the worst-performing product line by sales. The number of products is huge to analyse, but it can be simplified with analysis of product line which is also a segment analysis of products.

Product Line	COGS	Total Sales
Electronic Acces..	2,217	27,744
Fashion	2,386	30,957
Food and Bevera..	1,846	22,352
Home and Lifest..	2,173	25,220

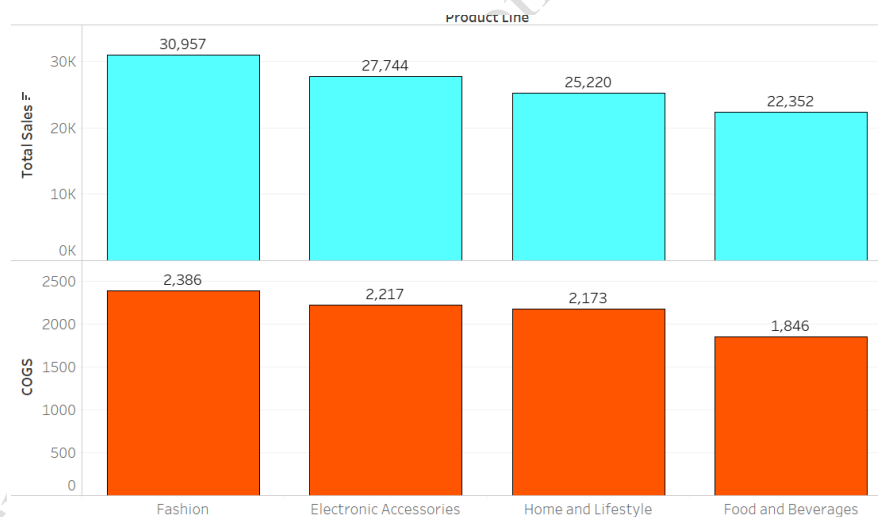


Figure 7: Total sales and COGS by product line

(Source: Self)

Fashion products generate the highest sales of over 30950 units but the cost of goods sold (COGS) is also the highest which is 2386 units. This is followed by electronic accessories with over 27,700 of total sales and COGS of 2,217 units. Food and beverages generate the lowest sales value of more than 22350 units. However, its COGS is also the lowest which is

only 1846 units. In between there is sales value of home and lifestyle which is 25,220 units and its COGS is 2,173 units.

Time analysis

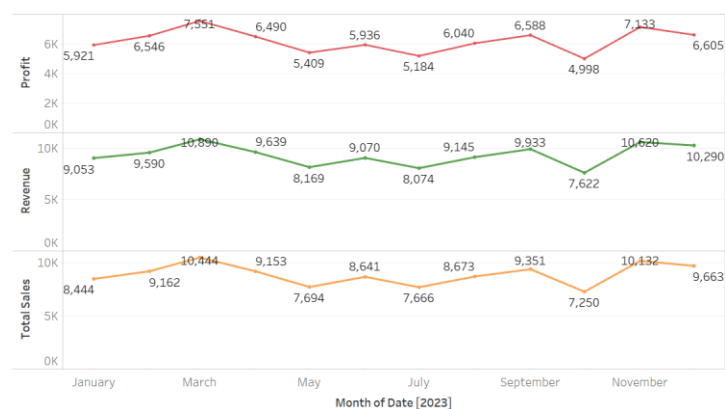
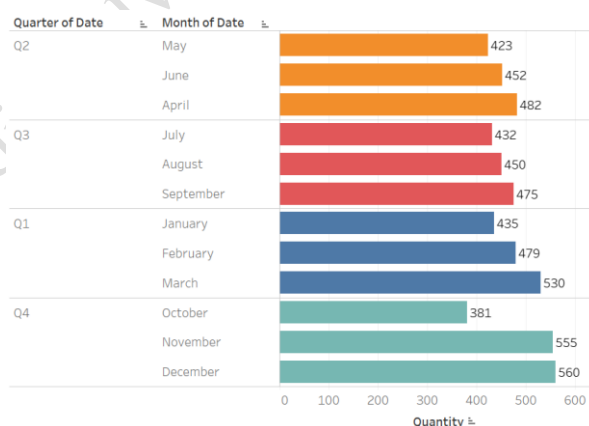


Figure 8: Performance by month

(Source: Self)

March 2023 is the best performing month for highest profit of 7551 units, revenue of 10890 units and 10444 units of total sales. This is followed by November 2023. Now, the sales quantity for each quarter is analysed



Q1			Q2			Q3			Q4		
Janua..	Febru..	March	April	May	June	July	August	Septe..	Octob..	Nove..	Dece..
435	479	530	482	423	452	432	450	475	381	555	560

Figure 9: Sales Quantity by Quarter

(Source: Self)

Hence, December of Q4 is the best-performing and May of Q2 is the worst-performing quarter-months by quantity.

Forecast trends

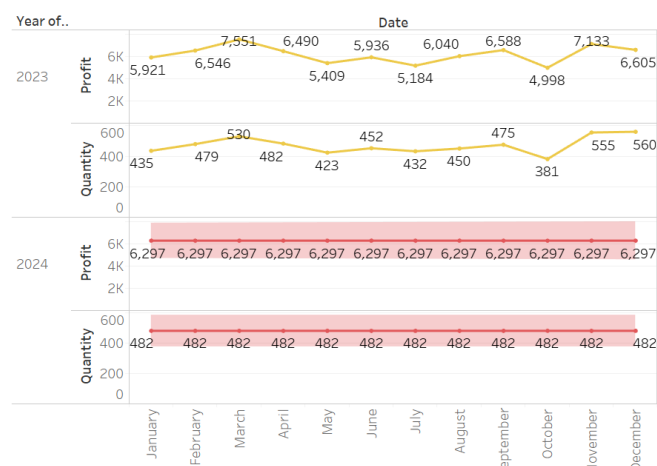


Figure 10: Profit and Quantity Forecasting

(Source: Self)

The forecasted profit in each month of 2024 is $6,297 \pm 1589$ units. Moreover, 482 ± 104 goods will be sold each month in 2024 as forecasted. Forecasted range of profit could be from 4708 to 7886. Similarly, quantity could be varied from 378 units to 586 units.

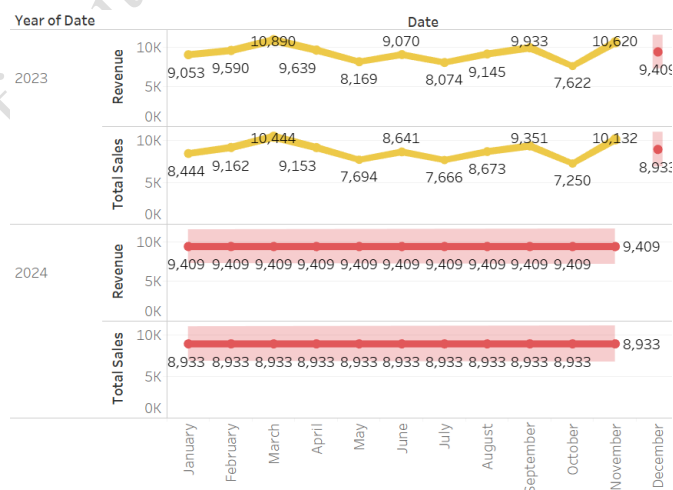


Figure 11: Forecasting Revenue and Total Sales

(Source: Self)

The forecasted revenue and total sales are 9409 and 8933 for each month of 2024 respectively. However, revenue could be $9409 - 2117 = 7292$ or $9409 + 2117 = 11,526$. Similarly, total sales forecasted range could be between 6865 (-2068) and 11001 (+2068).

Identify patterns using correlation, trendlines, regression

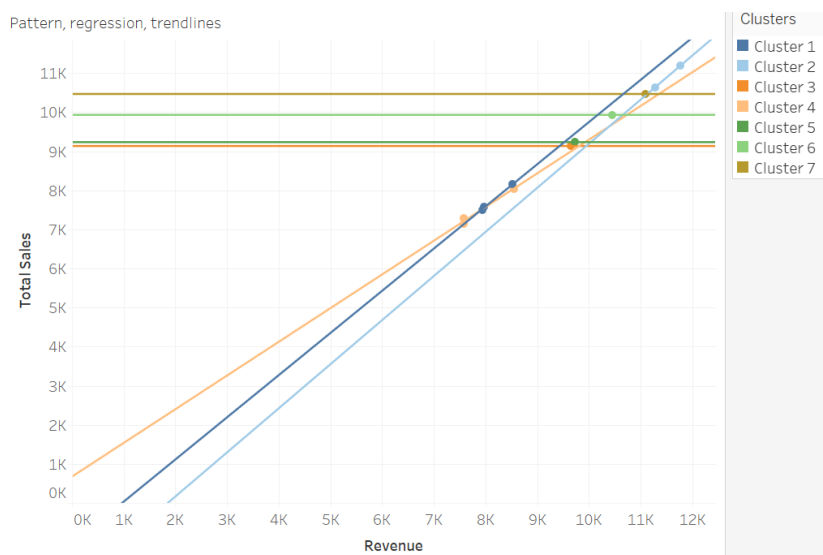


Figure 12: Pattern analysis with regression, and trendlines

(Source: Self)

This graph showed patterns for revenue, sales, quantity, rating, shipping cost, tax%, and unit price where revenue is the independent variable and total sales is the dependent variable.

Hovering over each data point could help to see the regression value of R-square, and regression equation. For cluster 1, the regression equation is:

$$\text{Total sales} = 1.07926 * \text{Revenue} - 1047.64$$

Hence, for this cluster, a 1-unit change in revenue results in 1.0796 units of change in total sales. The value of R-square is 0.994368 which means, that almost 99.44% variation in total sales is accurately evaluated by this cluster-1 model. With a p-value of $0.0478 < 0.05$, it is statistically significant.

Rating analysis

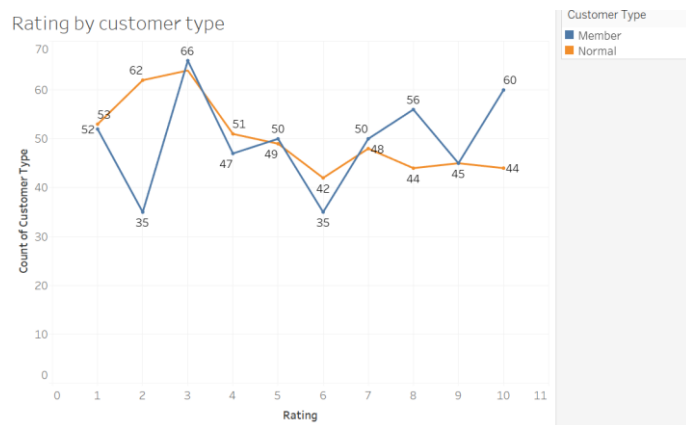


Figure 13: Rating by customer segment

(Source: Self)

Members provide more than 3, and 7 to 10 stars but normal customers provide more than 1 to 3 stars. 60 members provide 10-rating while only 44 non-members give that. In case of rating score of 1 2, 4, and 6, the normal customers rated more than the members. Each of them rated 9 for 45 times. For 3, 5, 7, 8, and 10, members provide more of these ratings than normal customers. Hence, the overall rating score of member are higher than non-members customers.

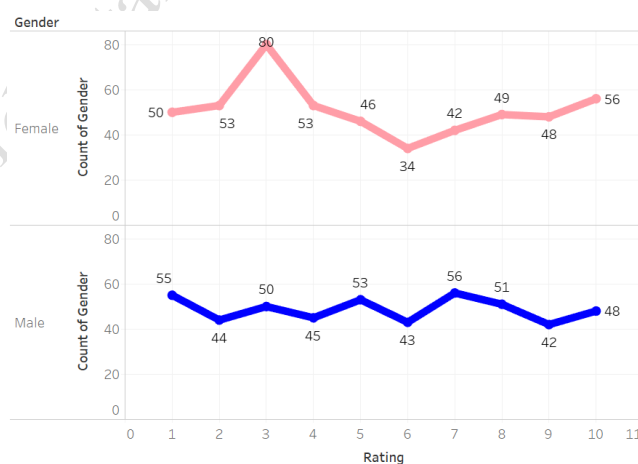


Figure 14: Rating by Gender

(Source: Self)

Most females rated 3 stars but ratings by males are more balanced from 1 to 10 stars. For example, males 42 to 56 males rated the products in each category which is a range of 14. Though, 34 females rated 6 stars which is the lowest and highest is 80 females while 3 stars. Hence, the range of female-rating is 46.

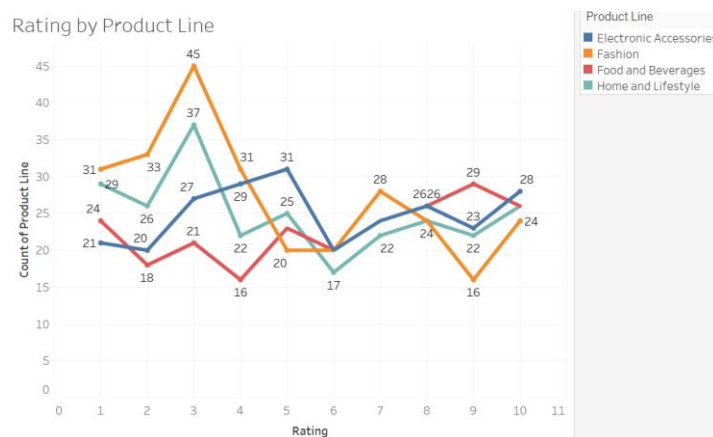
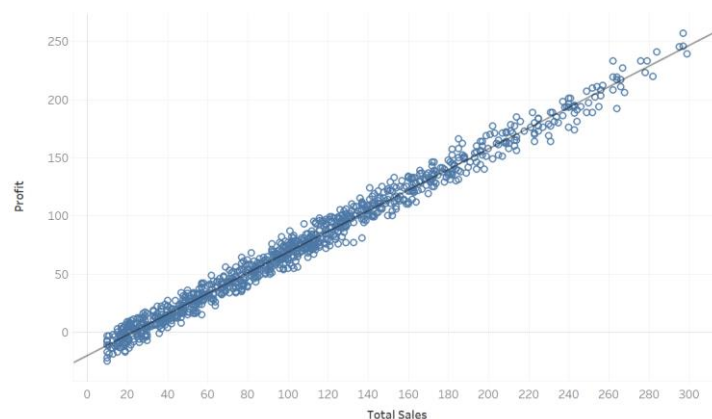


Figure 15: Rating by Product Line

(Source: Self)

Fashion products are highly rated between 1 and 4, 7, 8 and 10. For example, 28 people has rated fashion products 7 which is highest in this rating score. Similarly, food and beverages got most of the 9 rating which is from 29 people. Highest 10-rating is given to electronics accessories by 28 people.

Correlation and regression



R-Squared:	0.984915
Standard error:	7.23832
p-value (significance):	< 0.0001
Individual trend lines:	
Panes	Line
Row	Column
p-value	DF
Coefficients	Term
Value	StdErr
t-value	p-value
Profit	Total Sales
< 0.0001	904
Total Sales	0.88865
0.0036578	242.949
< 0.0001	
intercept	-20.2428
0.463487	-43.675
< 0.0001	

Figure 16: Correlation

(Source: Self)

It showed profit is highly and positively correlated with total sales and the model explained the variation with 98.5% accuracy (R-square = 0.984915). hence, the correlation coefficient is 0.9924 approximately (square root of R-square). The regression equation is:

$$Profit = 0.88865 * Total Sales - 20.2428.$$

p-value < 0.0001 hence it is a statistically significant model.

Group data

Here, the rating data is grouped into four groups. "Very Low Rating" means ratings of 1 to 3. Similarly, 4-5, 6-7, and 8-10 are grouped as "Low Rating", "Medium Rating" and "High Rating". This helps to provide a clearer rating analysis than the previous one.

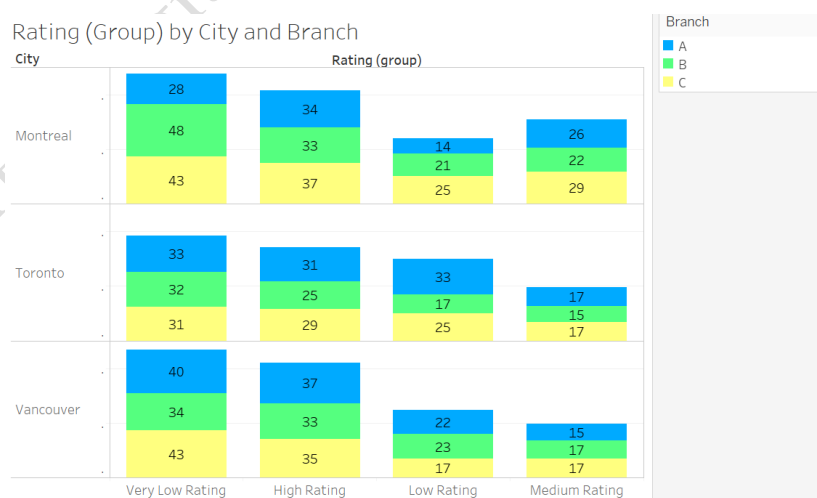


Figure 17: Rating group by city and branch

(Source: Self)

Hence, in each city and branch, a "Very Low Rating" is done more than any other group of ratings. However, a High Rating stood second. It can be improved.

Data visualization using Dashboard

Though, already data was visualization this section showed an interactive visualization of data using the dashboard.

Sales, Revenue and Profit Analysis

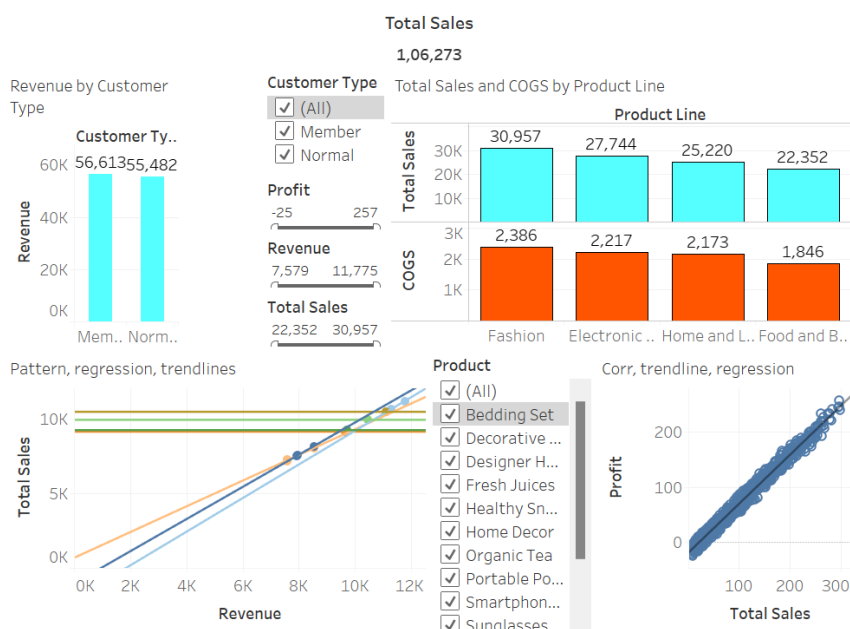


Figure 18: Dashboard visualization

(Source: Self)

It is an interactive dashboard, where the filters are linked with each graph and selecting any value from the filters can change others. Even, the “Revenue by Customer Type” graph is linked with other graphs as shown below.

Sales, Revenue and Profit Analysis

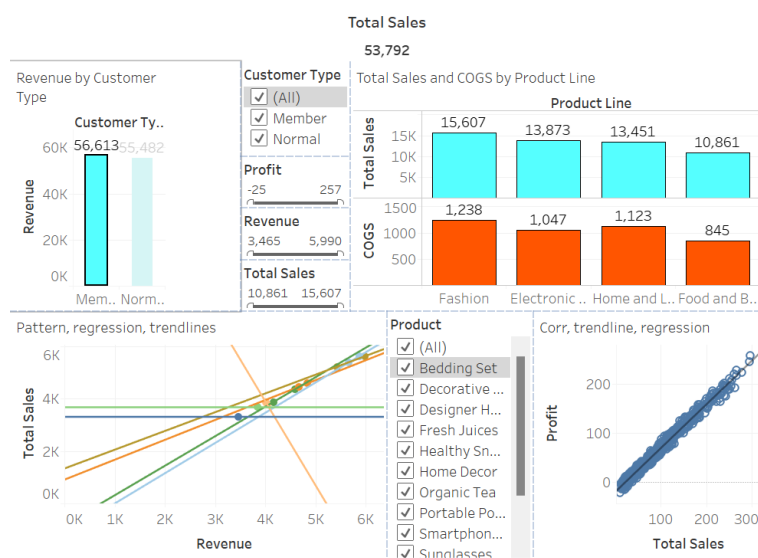


Figure 19: Interactive dashboard

(Source: Self)

More Insights

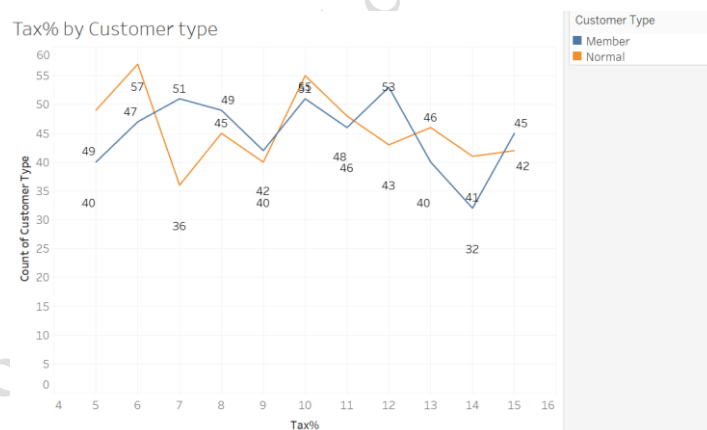


Figure 20: Tax% by customer type

(Source: Self)

This showed, most of the members and normal customers are paying 10% taxes. Though, most of the members are paying 6% to 10% tax normal customers are paying 10-12% of taxes.

Conclusion

Tableau helps to analyse sales, profit, and customer behaviour. However, it also helps beyond data analysis as the interactive dashboard helps to visualize data more effectively.

For example, Montreal is having also the highest shipping cost which is over 4500. This is again followed by Vancouver which is 4280 and then Toronto where the total shipping cost is 3795 units.

Recommendations

Special promotions and discounts could be offered like tax benefits to convert normal customers into members. It could help to increase their purchase frequency and bridge the revenue and sales gap with the members. Shipping strategies in Montreal and Vancouver and sales strategies in Toronto could be improved to reduce cost and increase profit respectively. Lastly, focus on the high-rated products category for customer satisfaction.

References

- Aguilar-Barrientos, S., Villegas-Gomez, J., & Arias-Salazar, A. (2021). Pricing and promotion: A literature review. *Aibi revista de investigación, administración e ingeniería*, 9(3), 59-65. Retrieved on 15th February 2024 from: https://revistas.udes.edu.co/aibi/article/download/pricing_and_promotion%3A_a_literature_review/2388
- Khatuwal, V. S., & Puri, D. (2022, April). Business Intelligence Tools for Dashboard Development. In *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 128-131). IEEE. Retrieved on 12th February 2024 from: <https://ieeexplore.ieee.org/abstract/document/9853086/>
- Qin, X., Luo, Y., Tang, N., & Li, G. (2020). Making data visualization more efficient and effective: a survey. *The VLDB Journal*, 29, 93-117. Retrieved on 14th February 2024 from: <https://m.kmkjmk.com/dbgroup.cs/ligl/papers/vldbj20-vizsurvey.pdf>
- Ridzuan, F., & Zainon, W. M. N. W. (2019). A review on data cleansing methods for big data. *Procedia Computer Science*, 161, 731-738. Retrieved on 13th February 2024 from: <https://www.sciencedirect.com/science/article/pii/S1877050919318885/pdf?md5=c8d975a00d9baaf0fdbcf1c527ccc96a&pid=1-s2.0-S1877050919318885-main.pdf>

Expertsofdigitalworld.com, +91 8766602532