

Business Analyst Career Program - Capstone Project

- By Md Faisal
- skfaisal819@gmail.com



- Data Exploration
- Statistical Analysis using Excel
- Graphical Analysis using Excel
- Insert the given data into the SQL server
- Import the Data from the SQL Database into PowerBI
- Interactive Dashboard by using visualization tools
- Conclusion and Inferences
- Endnotes



Insert the relevant screenshots and give brief summary about the Data

Summary :-

The dataset in question comprises sales data segmented by country, product, and discount band. Key columns include 'Segment', 'Country', 'Product', 'Discount Band', and 'Units Sold', which detail the segmentation and quantity of products sold. Financial metrics are extensively covered with columns for 'Manufacturing Price', 'Sale Price', 'Gross Sales', 'Discounts', 'Sales', 'COGS' (Cost of Goods Sold), and 'Profit', allowing for a comprehensive analysis of revenue, expenses, and profitability. Additionally, the dataset includes temporal information with columns like 'Date', 'Month Number', 'Month Name', and 'Year', enabling trend analysis over time. This dataset is structured to facilitate a detailed examination of sales performance across different markets and product categories, assess the impact of discount strategies, and track financial outcomes on a monthly and yearly basis. By exploring this data in Excel, one can derive insights into market segmentation, pricing effectiveness, discount impacts, and overall business performance across various regions and periods.

Statistical Analysis using Excel

Here are some of the key statistical analysis tools in Excel along with their formulas:

DISCRPTIVE STATISTICS	SALE PRICE	GROSS SALES	SALES	COGS	PROFIT	UNIT SOLD	DISCOUNTS
MEAN	118.428571	182759.4	169609.1	145475.2114	24133.86	1608.294	13150.35
MEDIAN	20	37980	35540.2	22506.25	9242.2	1542.5	2585.25
MODE	20	37050	32670	17430	0	727	0
ST. DEVIATION	136.775515	254262.3	236726.3	203865.5061	42760.63	867.4279	22962.93
VARIANCE	18707.5414	6.46E+10	5.6E+10	41561144585	1.83E+09	752431.1	5.27E+08
PERCENTILE.EXC	300.00	282675	262072.5	246202.5	22662	2232.375	16042.78

INFERRENTIAL STATISTICS	SALE PRICE VS PROFIT	UNIT SOLD VS SALES	DISCOUNTS VS COGS	DICOUNTS VS SALE PRICE	GROSS SALES VS SALES
T.TEST	1.7557E-46	5.38E-64	5.2313E-64	2.25E-44	0.316756
CORRELATION	0.65049466	0.326914	0.782929941	0.641689	0.998174

Statistical Analysis using Excel

USING ANNOVA IN B/W GROSS SALES AND SALE PRICE

Anova: Single Factor

--	--	--	--	--	--	--

SUMMARY

Groups	Count	Sum	Average	Variance		
20	699	82880	118.569385	18720.44		
32370	699	1.28E+08	182974.576	6.47E+10		

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1.17E+13	1	1.1686E+13	361.1832	8.31E-72	3.848129
Within Groups	4.52E+13	1396	3.2355E+10			
Total	5.69E+13	1397				

USING ANNOVA IN B/W SALES AND COGS

Anova: Single Factor

--	--	--	--	--	--	--

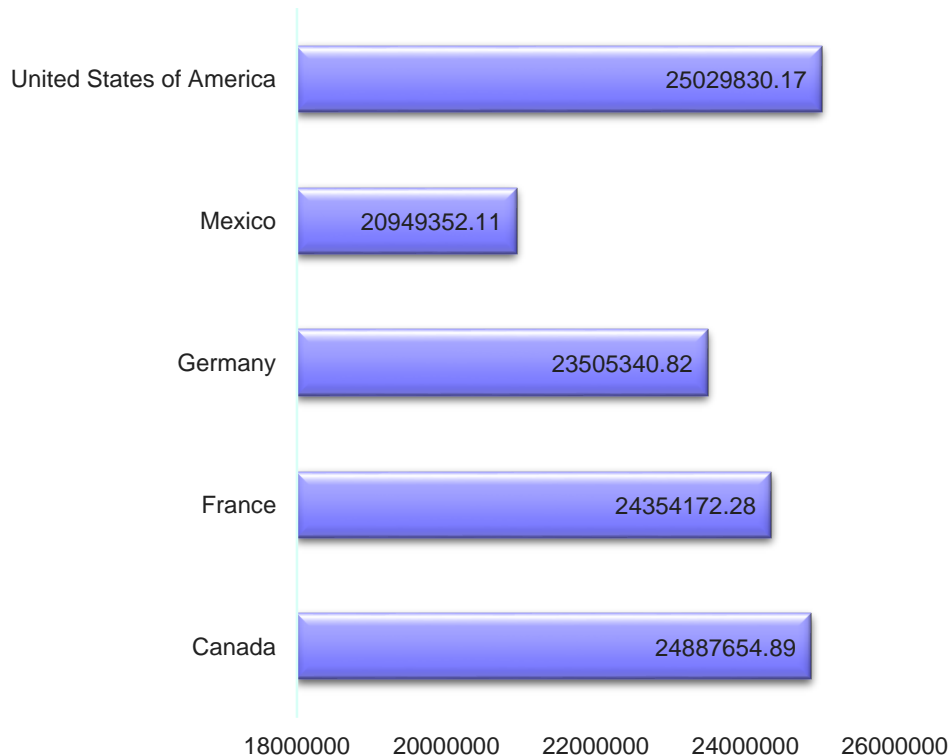
SUMMARY

Groups	Count	Sum	Average	Variance		
32370	699	1.19E+08	169805.4	5.61E+10		
16185	699	1.02E+08	145660.2	4.16E+10		

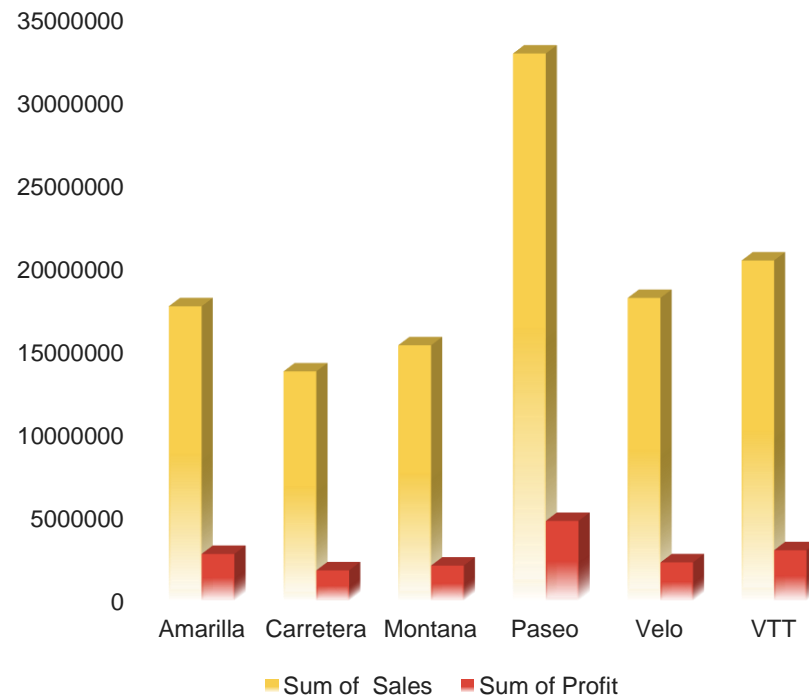
ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2.04E+11	1	2.04E+11	4.171505	0.041297	3.848129
Within Groups	6.82E+13	1396	4.88E+10			
Total	6.84E+13	1397				

SALES BY COUNTRY

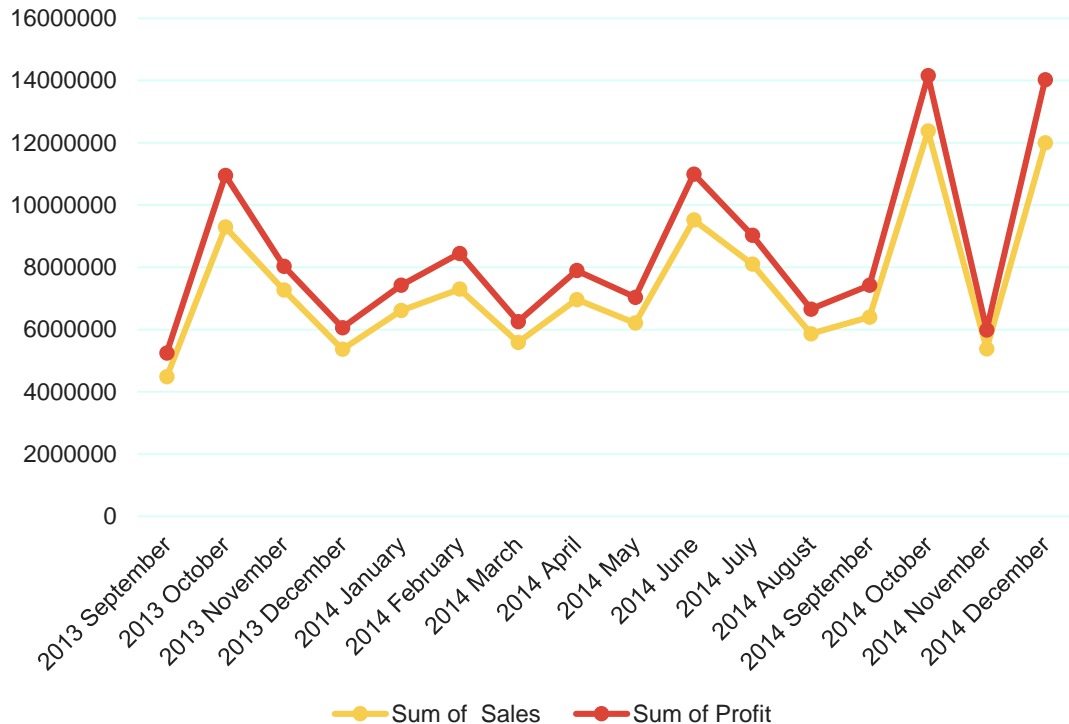


SALES AND PROFIT BY PRODUCT

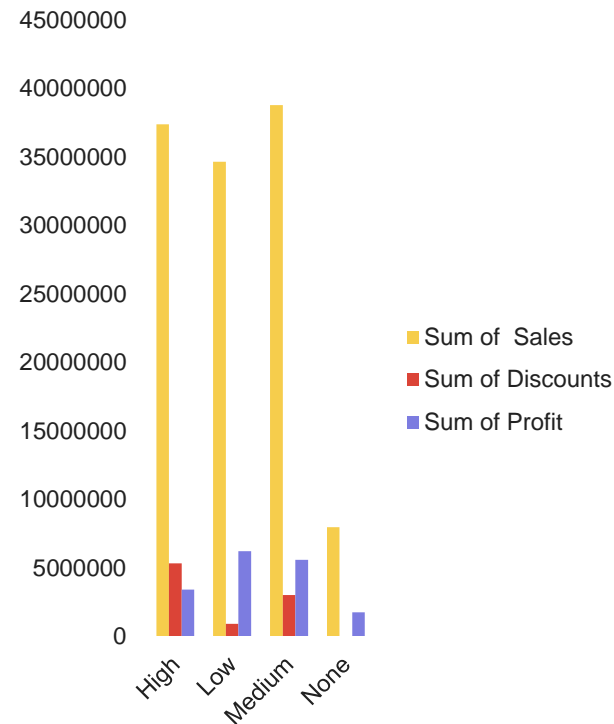


Graphical Analysis using Excel

TREND ANALYSIS BY MONTH AND YEAR



DISCOUNT IMPACT ANALYSIS



SALES BY COUNTRY

- ❑ The bar chart reveals which countries generate the highest and lowest sales. For example, countries like the USA and Germany may show significantly higher sales compared to smaller markets.

SALES AND PROFIT BY PRODUCT

- ❑ The clustered column chart highlights which products are the most and least profitable. High sales do not always correlate with high profit, indicating the need for strategic focus on high-margin products.

TREND ANALYSIS BY MONTH AND YEAR

- ❑ The line chart displays seasonal trends and annual growth. For instance, there might be noticeable sales peaks during specific months like December, suggesting a holiday effect

DISCOUNT IMPACT ANALYSIS

- ❑ The bar chart shows the relationship between discount levels and their effect on sales and profit. A higher discount band might increase sales but at the expense of profit margins.

Insert the given data into the SQL server

Insert the relevant screenshots verifying that the data is successfully inserted into a SQL DB. Run a basic Queries on Data.

SCHEMAS

Filter objects

faisaldb

sys

- Load Spatial Data
- Set as Default Schema
- Filter to This Schema
- Schema Inspector
- Table Data Import Wizard**
- Copy to Clipboard
- Send to SQL Editor
- Create Schema...
- Alter Schema...
- Drop Schema...
- Search Table Data...
- Refresh All

Table Data Import

Select file to import

Table Data Import allows you to easily import CSV, JSON datatables. You can also create destination table on the fly.

File Path: Browse...

Open

Organize New folder

Name Date modified Type

- Last week (1)
 - Western Countries Financial Data 2 01-06-2024 18:41 Microsoft Excel
- Last month (1)
 - Telegram Desktop Type: Microsoft Excel Comma Separated Values Size: 72.5 KB Date Modified: 01-06-2024 18:41
- Earlier this year (1)
 - Tableau Assignment (Major) - Mohd Am... 28-03-2024 11:51 File folder

File name: Western Countries Financial Data 2 Comma Separated Values (*.csv)

Open Cancel

capstone

Limit to 100 rows

```

1 SELECT * FROM `western countries financial data 2`
2
3 /*Total Sales by Country
4 SELECT Country, SUM(Sales) AS Total_Sales
5 FROM `western countries financial data 2`

```

Result Grid

Segment	Country	Product	Discount Band	Units Sold	Manufacturing Price	Sale Price	Gross Sales	Discounts	Sales	COGS	Prof
Government	Canada	Carretera	None	1618.5	3	20	32370	0	32370	16185	1618
Government	Germany	Carretera	None	1321	3	20	26420	0	26420	13210	1321
Midmarket	France	Carretera	None	2178	3	15	32670	0	32670	21780	1089
Midmarket	Germany	Carretera	None	888	3	15	13320	0	13320	8880	4440
Midmarket	Mexico	Carretera	None	2470	3	15	37050	0	37050	24700	1235
Government	Germany	Carretera	None	1513	3	350	529550	0	529550	393380	1361
Midmarket	Germany	Montana	None	921	5	15	13815	0	13815	9210	4605

Running Queries in SQL server

Insert the relevant screenshots verifying that the data is successfully inserted into a SQL DB. Run a Queries on Data.

capstone*

```

3  /*Total Sales by Country*/
4  • SELECT Country, SUM(Sales) AS Total_Sales
5    FROM `western countries financial data 2`
6    GROUP BY Country
7    ORDER BY Total_Sales DESC;

```

Result Grid

Country	Total_Sales
United States of America	25029834
Canada	24887660
France	24354182
Germany	23505352
Mexico	20949357

capstone*

```

9  /*Total Sales and Profit by Product*/
10 • SELECT Product, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit
11    FROM `western countries financial data 2`
12    GROUP BY Product
13    ORDER BY Total_Sales DESC;

```

Result Grid

Product	Total_Sales	Total_Profit
Paseo	3011153	4797448
VTT	20511928	3034610
Velo	18250066	2305995
Amarilla	17747118	2814105
Montana	15390807	2114760
Carretera	13815313	1826806

capstone*

```

15 /*Monthly Sales Trends*/
16 • SELECT Year, `Month Name`, SUM(Sales) AS Total_Sales
17    FROM `western countries financial data 2`
18    GROUP BY Year, `Month Number`, `Month Name`
19    ORDER BY Year, `Month Number`;

```

Result Grid

Year	Month Name	Total_Sales
2013	September	4483999
2013	October	9295620
2013	November	7267206
2013	December	5368443
2014	January	6607763
2014	February	7297532
2014	March	5586864
2014	April	6964776

capstone*

```

21 /*Sales and Profit by Discount Band*/
22 • SELECT `Discount Band`, SUM(Sales) AS Total_Sales, SUM(Profit) AS Total_Profit
23    FROM `western countries financial data 2`
24    GROUP BY `Discount Band`
25    ORDER BY Total_Sales DESC;

```

Result Grid

Discount Band	Total_Sales	Total_Profit
Medium	38780444	5579531
High	37372497	3388869
Low	34629789	6188868
None	7943655	1736456

capstone*

```

27 /*Units Sold vs. Profit by Segment*/
28 • SELECT Segment, SUM(`Units Sold`) AS Total_Units_Sold, SUM(Profit) AS Total_Profit
29    FROM `western countries financial data 2`
30    GROUP BY Segment
31    ORDER BY Total_Profit DESC;

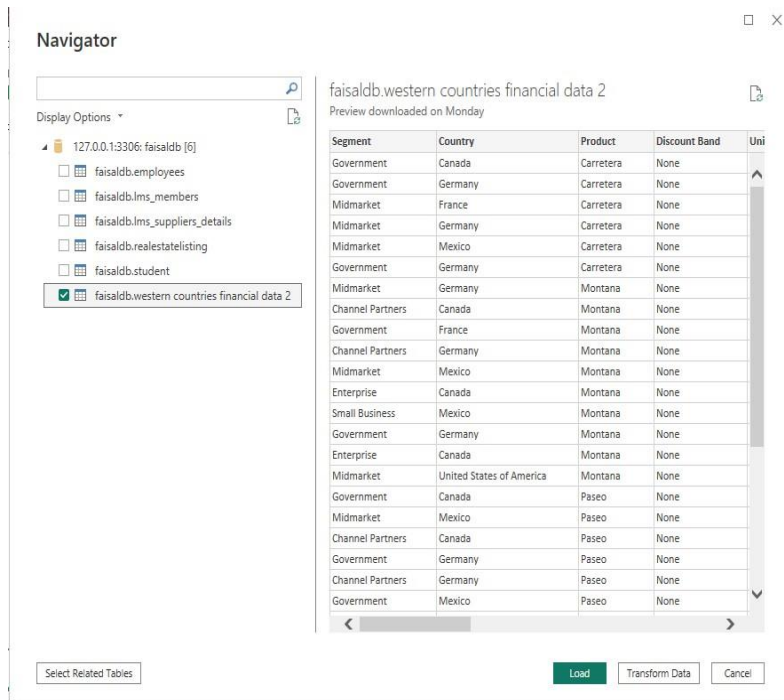
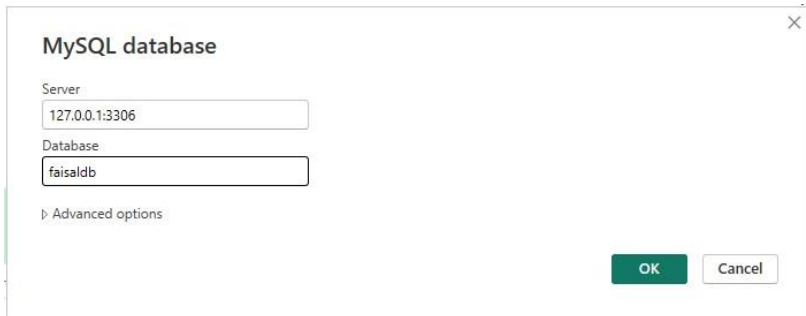
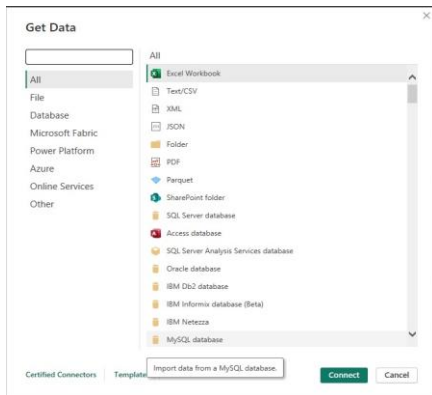
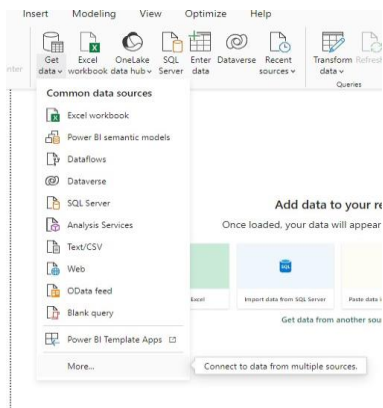
```

Result Grid

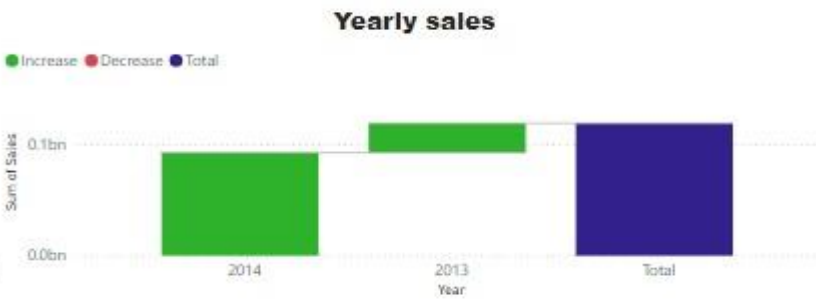
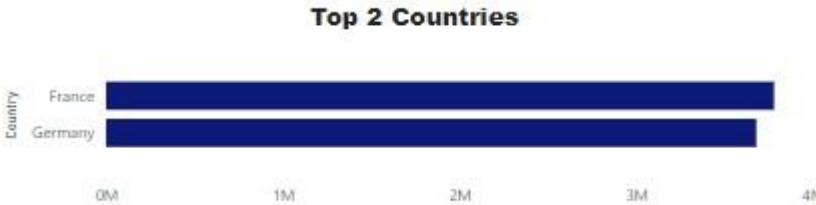
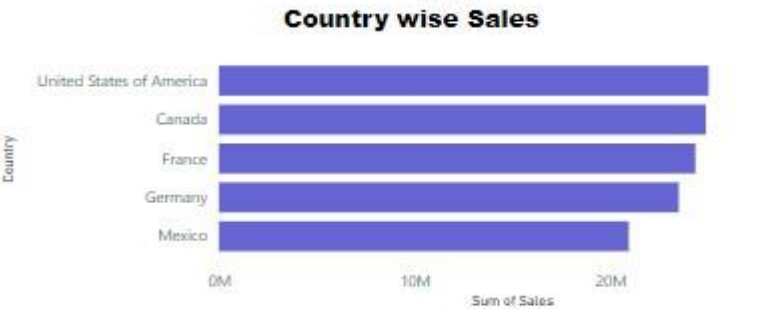
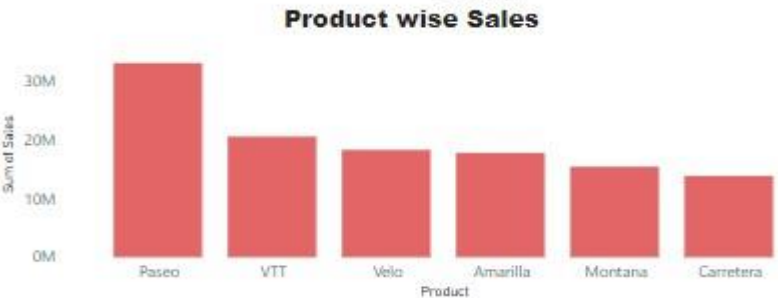
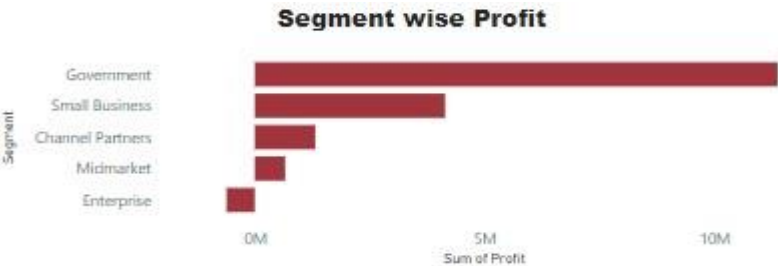
Segment	Total_Units_Sold	Total_Profit
Government	470673.5	11388197
Small Business	153139	4143170
Channel Partners	161263.5	1316803
Midmarket	172178	660103
Enterprise	168552	-614549

Import the Data from the SQL Database into PowerBI

Insert the relevant screenshots verifying that the data is successfully inserted into PowerBI and do the required data cleaning or creation of measured for data analysis



Interactive Dashboard by using visualization tools

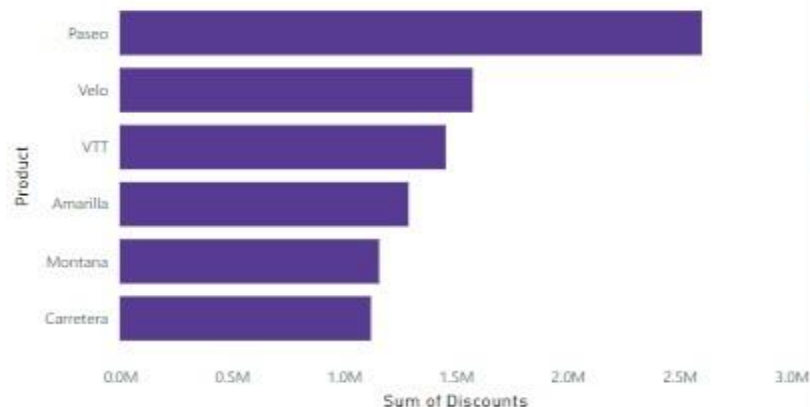


Interactive Dashboard by using visualization tools

Segment wise Product Wise Profit

Segment	Amarilla	Carretera	Montana	Paseo	Velo	VTT	Total
Channel Partners	230068	208405	192457	331840	134268	219765	1316803
Enterprise	-95154	-222714	-31095	-81738	-84764	-99084	-614549
Government	2208304	1398998	1126205	3057297	1756735	1840658	11388197
Midmarket	63606	94105	83879	258739	68653	91121	660103
Small Business	407281	348012	743314	1231310	431103	982150	4143170
Total	2814105	1826806	2114760	4797448	2305995	3034610	16893724

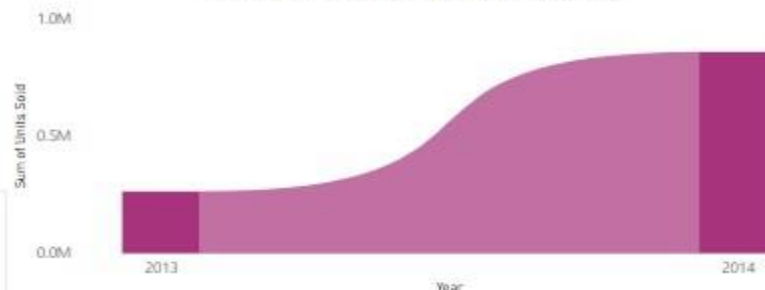
Product wise Discount



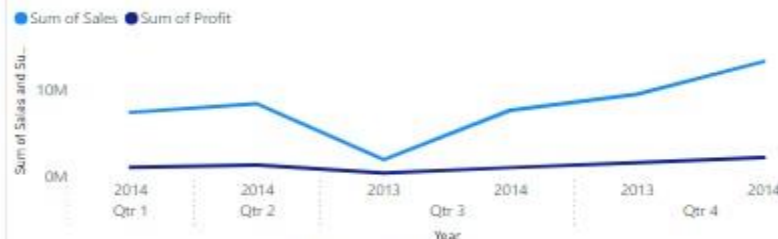
Segment Wise Product Sales

Segment	Amarilla	Carretera	Montana	Paseo	Velo	VTT	Total
Channel Partners	317642	282838	261844	454516	182925	300829	1800594
Enterprise	2643608	3203710	2614845	5267862	3581240	2300440	19611705
Government	9942901	6080948	5548940	14882236	7813425	8235833	52504283
Midmarket	248686	337305	290239	907729	264498	333426	2381883
Small Business	4594281	3910512	6674939	11498810	6407978	9341400	42427920
Total	17747118	13815313	15390807	33011153	18250066	20511928	118726385

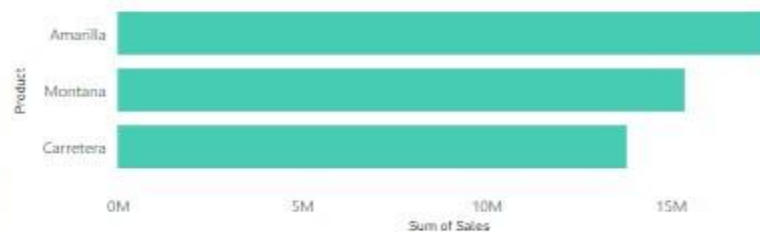
No. of units sold year over year



Profit and Sales by Qtr



Bottom 3 Products



Trends

1. Product Performance:

- Paseo is the top-selling and most discounted product.
- VTT and Montana also show strong sales but with lower discounts.

2. Segment Performance:

- Government and Small Business segments are highly profitable.
- Enterprise segment struggles with negative profits.

3. Geographical Insights:

- USA and Canada dominate sales, with the USA slightly ahead.
- Lower sales observed in countries such as Carretera, Montana, and Amarilla.

4. Yearly Performance:

- Significant growth in sales and units sold from 2013 to 2014.

Conclusions and Inferences

▪ Strong Product Sales:

- Paseo, VTT, and Montana are the top-performing products, driving substantial revenue.

▪ Profitability by Segment:

- Government segment's high profitability indicates strong demand and effective sales strategies.
- The negative profit in the Enterprise segment suggests inefficiencies or high costs.

▪ Geographic Sales Distribution:

- USA and Canada are key markets, showing significant contributions to overall sales.
- Countries with lower sales may require targeted marketing and sales strategies to improve performance.

▪ Annual Growth:

- The marked increase in profits and units sold from 2013 to 2014 highlights effective business strategies and market growth during this period.

Reference Links:-

Insert related links for Excel file, PowerBI file

Excel – <https://docs.google.com/spreadsheets/d/1QE4R1vom0FEMRSCWB-VvukedEQ3NmY4V/edit?usp=sharing&ouid=109362111683177474455&rtpof=true&sd=true>

Power Bi -

https://drive.google.com/file/d/1iQ_HEPMZRto9nYWwlh22cT8DHB5D8n0C/view?usp=drive_link