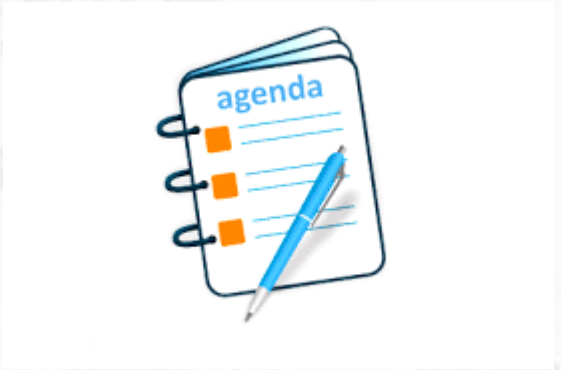
Business Analyst Career Program - Capstone Project

- By Priyanka Nadar

Agenda

skill academy

- 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



Data Exploration

skill academy

- First part of data exploration comprises of studying the data and understanding what it is trying to convey
- This is a sales data provided by the finance department for various products, classified on the basis of segments which sells these for countries from Europe and North America
- The first step is data cleaning i.e. removing extra data/outliers to avoid redundancy and to derive proper results
- We have an excel data comprising of 700 entries providing ample amount of information which will help us to gather insights as we proceed further

Data Cleaning

skill academy

- The amount in various columns are specified in dollars ($), changed the datatype to ‘General’ to ensure consistency
- Two columns ‘Sales’ and ‘Gross Sales’ have same values, hence removed the latter
- Added a new column ‘Discount %’, calculated the same using the formula: Discount Provided / Sales * 100
- There’s a column named ‘Date’, then ‘Month Number’, ‘Month Name’ and ‘Year’ are extra columns. Removed the extra columns. Changed the ‘Date’ column type to ‘Custom – dd/mm/YYYY’ to avoid inconsistency with the timelines in further analysis

Data Cleaning

skill academy

- Before and After result after cleaning looks like the sample data snippet below
- Save the file as CSV

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Segment	Country	Product	Discount Band	Units Sold	Manufacturing	Sale Price	Gross Sales	Discounts	Sales	COGS	Profit	Date	Month Number	Month Name	Year
2	Government	Canada	Carretera	None	1618.5	$ 3.00	$ 20.00	$ 32,370.00	$ -	$ 32,370.00	$ 16,185.00	$ 16,185.00	1/1/2014	1	January	2014
3	Government	Germany	Carretera	None	1321	$ 3.00	$ 20.00	$ 26,420.00	$ -	$ 26,420.00	$ 13,210.00	$ 13,210.00	1/1/2014	1	January	2014

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Segment	Country	Product	Discount Band	Units Sold	Manufacturing Pri	Sale Price	Discounts	Discount	Sales	COGS	Profit	Date
86	Government	Mexico	Velo	Low	362	120	7	25.34	1.01	2508.66	1810	698.66	01/05/2014
90	Government	Germany	VTT	Low	263	250	7	18.41	1.01	1822.59	1315	507.59	01/03/2014

Statistical Analysis using Excel

skill academy

- Statistical analysis has been done on each products and the no. of units sold
- On checking data in the below image, deviation from mean value is greater for the product ‘Velo’

Row Labels	Sum of Units Sold	Average of Units Sold	Standard Deviation of Units Sold	Variance of Units Sold
Amarilla	155315	1652.287234	853.2126243	727971.7822
Carretera	146846	1578.989247	869.7091186	756393.951
Montana	154198	1658.043011	796.591375	634557.8188
Paseo	338239.5	1674.45297	876.7778479	768739.3945
Velo	162424.5	1490.133028	923.203901	852305.4428
VTT	168783	1548.46789	862.9550766	744691.4642
Grand Total	**1125806**	**1608.294286**	**867.4278591**	**752431.0907**

Statistical Analysis using Excel

skill academy

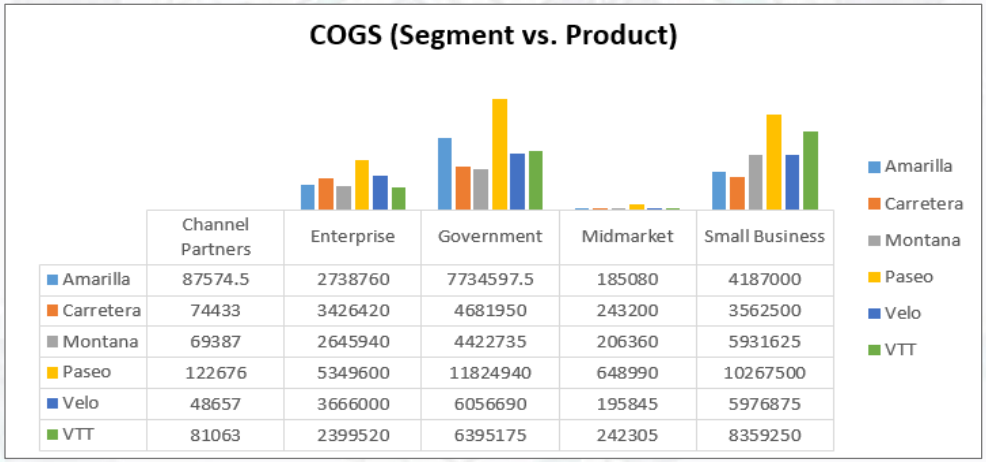
- Total, average, maximum and median values calculated for various parameters
- More data can be found in worksheet named '**KPIs**' from the cleaned dataset
- From the mode calculation, it is understood that for most products, the segments in various countries don't offer any sort of discount
- Also, from the manufacturing price's mode, we can infer that we have most entries for the product 'Paseo'

Total Discount	Average Discount	Maximum Discount
9205248.24	13150.35463	149677.5
Mode of Discount %	**Mode of Manufacturing Price**	**Median of Sales**
0	10	35540.2

Graphical Analysis using Excel

skill academy

- Graphical analysis and appropriate visuals helps to understand the numbers in a prompt manner
- Referring the image below, we can infer that the COGS(Cost of Goods Sold) for the product ‘Paseo’ in ‘Government’ segment is the highest

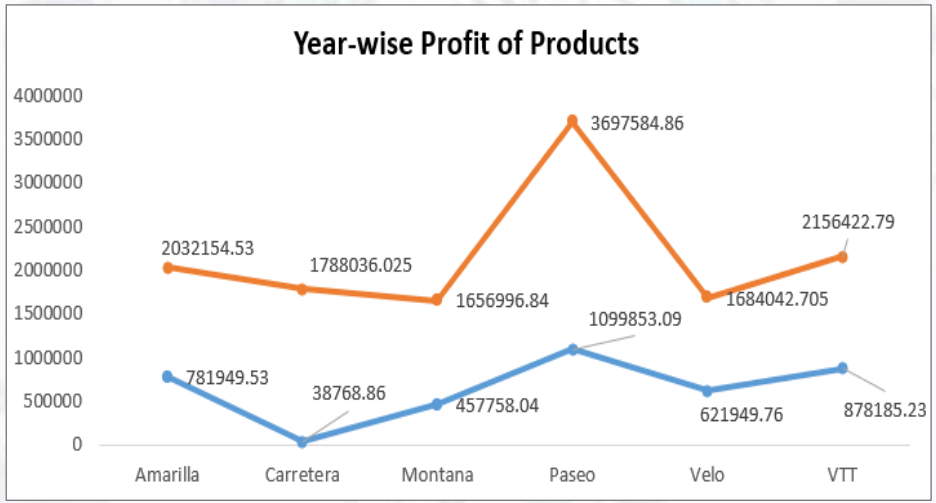


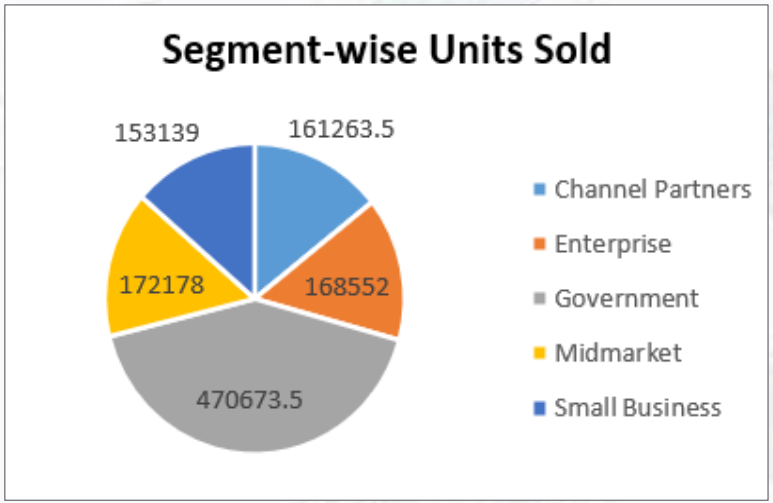
	Channel Partners	Enterprise	Government	Midmarket	Small Business
Amarilla	87574.5	2738760	7734597.5	185080	4187000
Carretera	74433	3426420	4681950	243200	3562500
Montana	69387	2645940	4422735	206360	5931625
Paseo	122676	5349600	11824940	648990	10267500
Velo	48657	3666000	6056690	195845	5976875
VTT	81063	2399520	6395175	242305	8359250

Graphical Analysis using Excel

skill academy

- Below mentioned two visuals suggest that the profit incurred by sales of the product ‘Paseo’ is highest for both the years (2013 & 2014), and the most sales is done by the ‘Government’ segment, followed by ‘Midmarket’



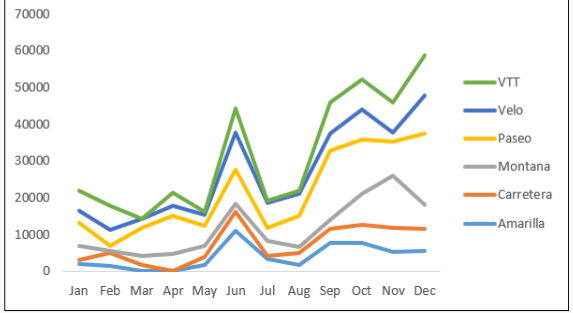


Overall Dashboard

skill academy

Financial Data Analysis Dashboard for Western Countries

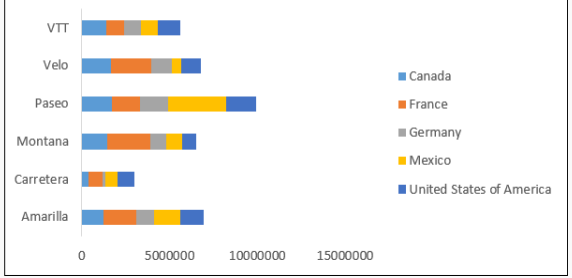
Units Sold for Products (Month-wise)



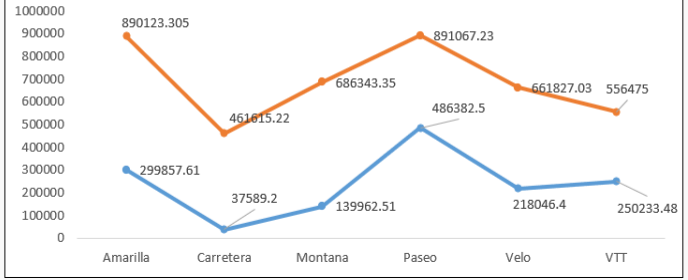
COGS (Segment vs. Product)

	Channel Partners	Enterprise	Government	Midmarket	Small Business
Amarilla	8412	555180	3723975	33115	1429250
Carretera	31200	291240	892135	81550	1178500
Montana	36204	1293720	1701715	32120	2646125
Paseo	28905	1723920	2899965	235650	3679500
Velo	14988	1301160	1922090	90810	2579125
VTT	36804	105240	888320	52240	3707750

Total Sales (Country vs. Product)



Year-wise Profit of Products



Segment

- Channel Partners
- Enterprise
- Government
- Midmarket
- Small Business

Years

- 2013
- 2014

Discount Band

- High
- Low
- Medium
- None

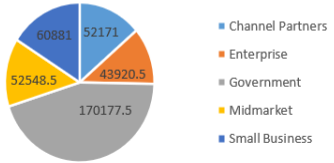
Country

- Canada
- France
- Germany
- Mexico
- United States of America

Product

- Amarilla
- Carretera
- Montana

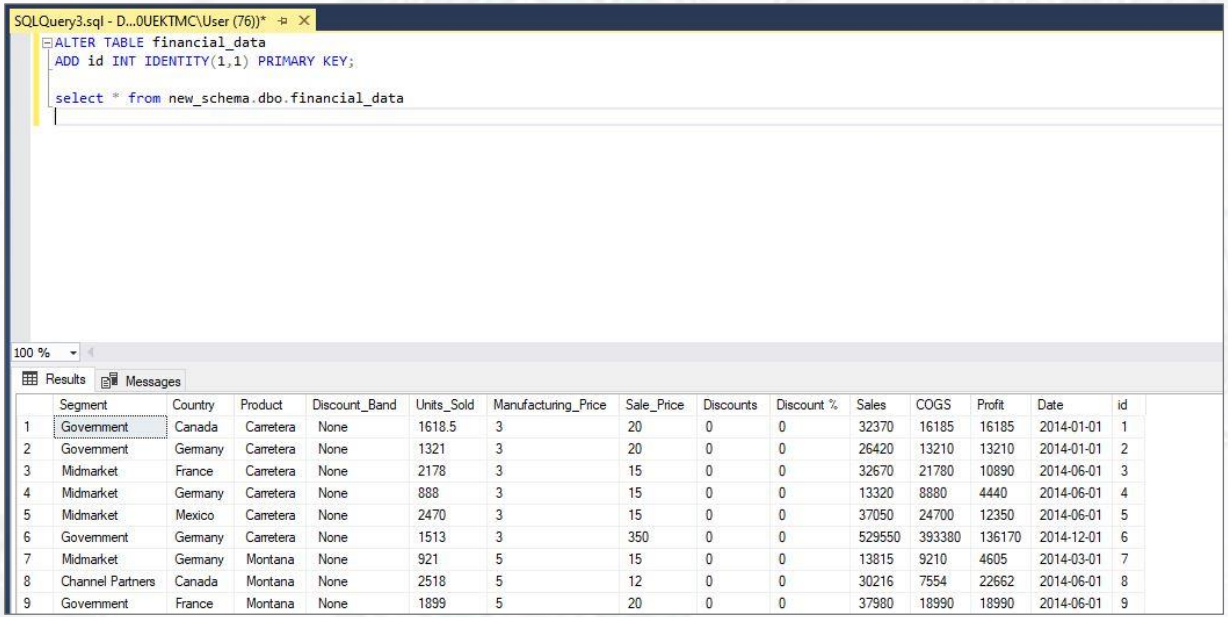
Segment-wise Units Sold



Insert the given data into the SQL server

skill academy

- Added a new column id (Primary Key), please find the attached screenshot

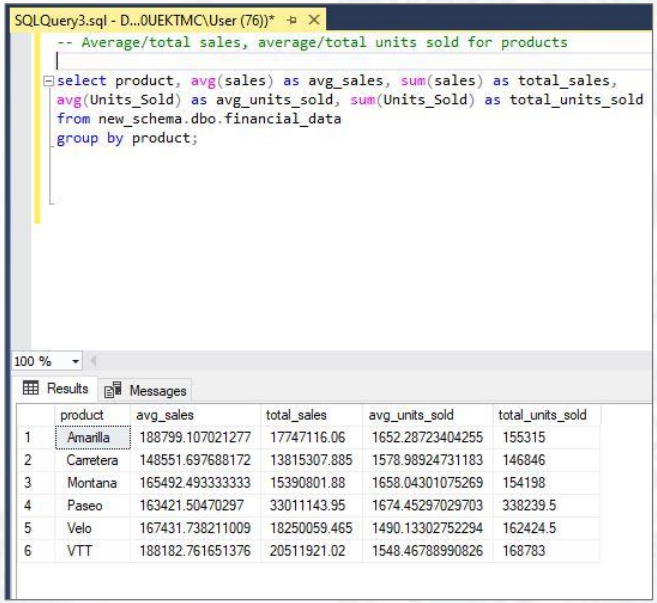


	Segment	Country	Product	Discount_Band	Units_Sold	Manufacturing_Price	Sale_Price	Discounts	Discount %	Sales	COGS	Profit	Date	id
1	Government	Canada	Carretera	None	1618.5	3	20	0	0	32370	16185	16185	2014-01-01	1
2	Government	Germany	Carretera	None	1321	3	20	0	0	26420	13210	13210	2014-01-01	2
3	Midmarket	France	Carretera	None	2178	3	15	0	0	32670	21780	10890	2014-06-01	3
4	Midmarket	Germany	Carretera	None	888	3	15	0	0	13320	8880	4440	2014-06-01	4
5	Midmarket	Mexico	Carretera	None	2470	3	15	0	0	37050	24700	12350	2014-06-01	5
6	Government	Germany	Carretera	None	1513	3	350	0	0	529550	393380	136170	2014-12-01	6
7	Midmarket	Germany	Montana	None	921	5	15	0	0	13815	9210	4605	2014-03-01	7
8	Channel Partners	Canada	Montana	None	2518	5	12	0	0	30216	7554	22662	2014-06-01	8
9	Government	France	Montana	None	1899	5	20	0	0	37980	18990	18990	2014-06-01	9

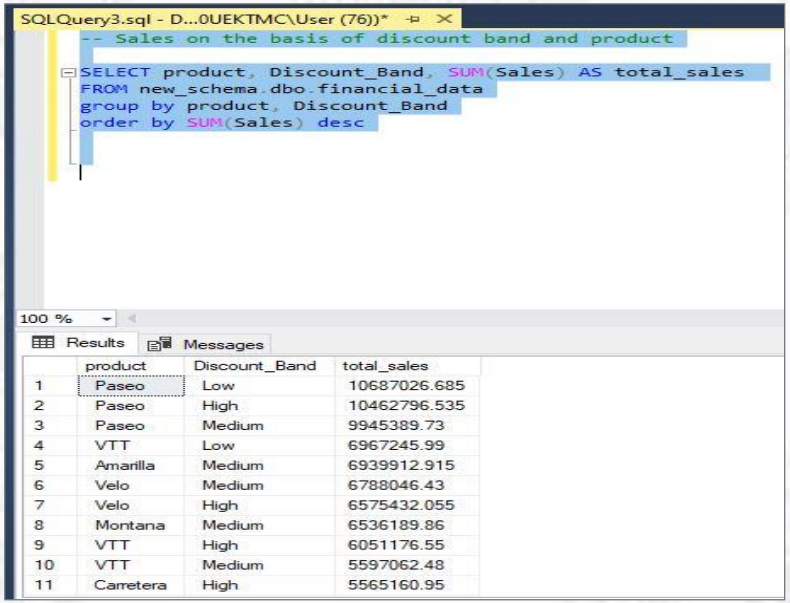
Insert the given data into the SQL server

skill academy

- Please find the attached screenshots



	product	avg_sales	total_sales	avg_units_sold	total_units_sold
1	Amarilla	188799.107021277	17747116.06	1652.28723404255	155315
2	Carretera	148551.697688172	13815307.885	1578.98924731183	146846
3	Montana	165492.493333333	15390801.88	1658.04301075269	154198
4	Paseo	163421.50470297	33011143.95	1674.45297029703	338239.5
5	Velo	167431.738211009	18250059.465	1490.13302752294	162424.5
6	VTT	188182.761651376	20511921.02	1548.46788990826	168783



	product	Discount_Band	total_sales
1	Paseo	Low	10687026.685
2	Paseo	High	10462796.535
3	Paseo	Medium	9945389.73
4	VTT	Low	6967245.99
5	Amarilla	Medium	6939912.915
6	Velo	Medium	6788046.43
7	Velo	High	6575432.055
8	Montana	Medium	6536189.86
9	VTT	High	6051176.55
10	VTT	Medium	5597062.48
11	Carretera	High	5565160.95

Insert the given data into the SQL server

skill academy

- Please find the attached screenshots

SQLQuery3.sql - D...0UEKTMC\User (76))*

```
select distinct product, Manufacturing_Price
from new_schema.dbo.financial_data
```

100 %

Results | Messages

	product	Manufacturing_Price
1	Amarilla	260
2	Carretera	3
3	Montana	5
4	Paseo	10
5	Velo	120
6	VTT	250

SQLQuery3.sql - D...0UEKTMC\User (76))*

```
-- display only those entries whose COGS < 10000 using subquery display most recent 10 rows

select top 10 * from new_schema.dbo.financial_data where id IN
(select id from new_schema.dbo.financial_data where COGS<10000)
order by id desc;
```

100 %

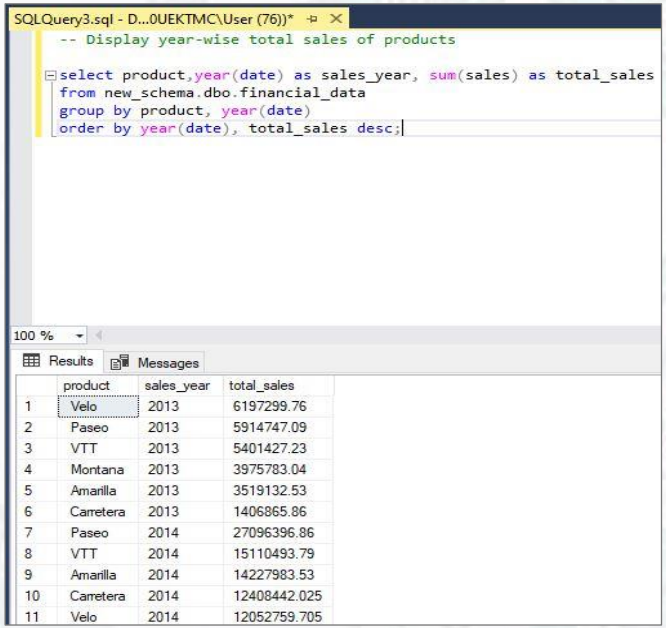
Results | Messages

	Segment	Country	Product	Discount_Band	Units_Sold	Manufacturing_Price	Sale_Price	Discounts	Discount %	Sales	COGS	Profit	Date	id
1	Channel Partners	United States of America	VTT	High	1806	250	12	3250.8	17.65	18421.2	5418	13003.2	2014-05-01	700
2	Government	Canada	Paseo	High	723	10	7	759.15	17.65	4301.85	3615	686.85	2014-04-01	699
3	Government	Mexico	Montana	High	1368	5	7	1436.4	17.65	8139.6	6840	1299.6	2014-02-01	698
4	Government	France	VTT	High	293	250	20	879	17.65	4981	2930	2051	2014-12-01	695
5	Government	United States of America	VTT	High	267	250	20	801	17.65	4539	2670	1869	2013-10-01	691
6	Midmarket	Germany	VTT	High	492	250	15	1107	17.65	6273	4920	1353	2014-07-01	690
7	Government	Canada	VTT	High	865.5	250	20	2596.5	17.65	14713.5	8655	6058.5	2014-07-01	689
8	Channel Partners	United States of America	Velo	High	914	120	12	1645.2	17.65	9322.8	2742	6580.8	2014-12-01	688
9	Channel Partners	Mexico	Velo	High	500	120	12	900	17.65	5100	1500	3600	2014-03-01	683
10	Government	France	Paseo	High	293	10	20	879	17.65	4981	2930	2051	2014-12-01	682

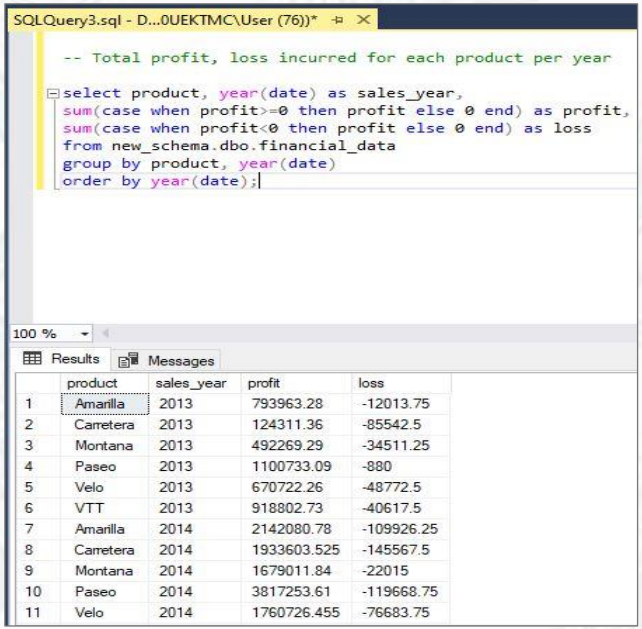
Insert the given data into the SQL server

skill academy

- Please find the attached screenshots



	product	sales_year	total_sales
1	Velo	2013	6197299.76
2	Paseo	2013	5914747.09
3	VTT	2013	5401427.23
4	Montana	2013	3975783.04
5	Amarilla	2013	3519132.53
6	Carretera	2013	1406865.86
7	Paseo	2014	27096396.86
8	VTT	2014	15110493.79
9	Amarilla	2014	14227983.53
10	Carretera	2014	12408442.025
11	Velo	2014	12052759.705

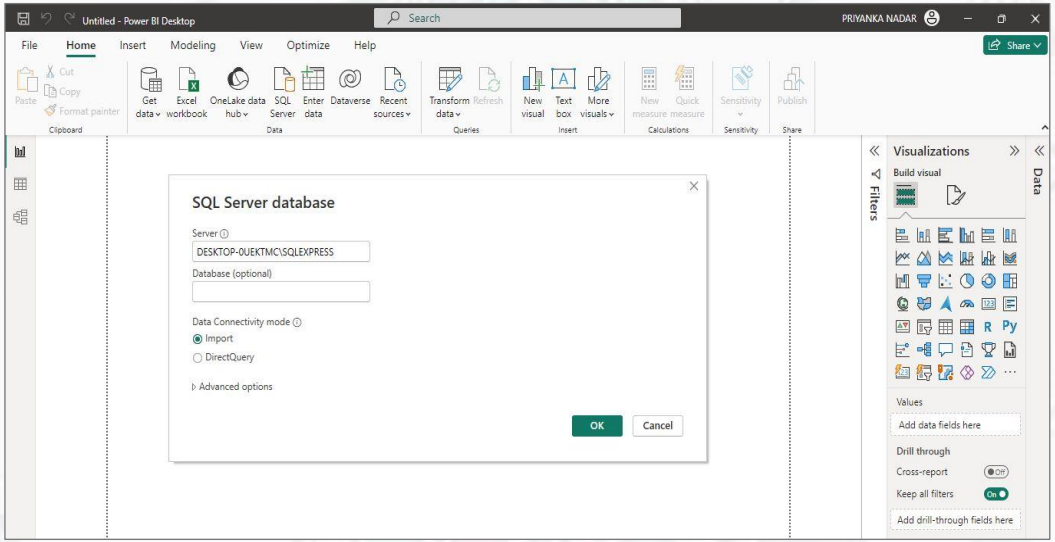


	product	sales_year	profit	loss
1	Amarilla	2013	793963.28	-12013.75
2	Carretera	2013	124311.36	-85542.5
3	Montana	2013	492269.29	-34511.25
4	Paseo	2013	1100733.09	-880
5	Velo	2013	670722.26	-48772.5
6	VTT	2013	918802.73	-40617.5
7	Amarilla	2014	2142080.78	-109926.25
8	Carretera	2014	1933603.525	-145567.5
9	Montana	2014	1679011.84	-22015
10	Paseo	2014	3817253.61	-119668.75
11	Velo	2014	1760726.455	-76683.75

Import the Data from the SQL Database into PowerBI

skill academy

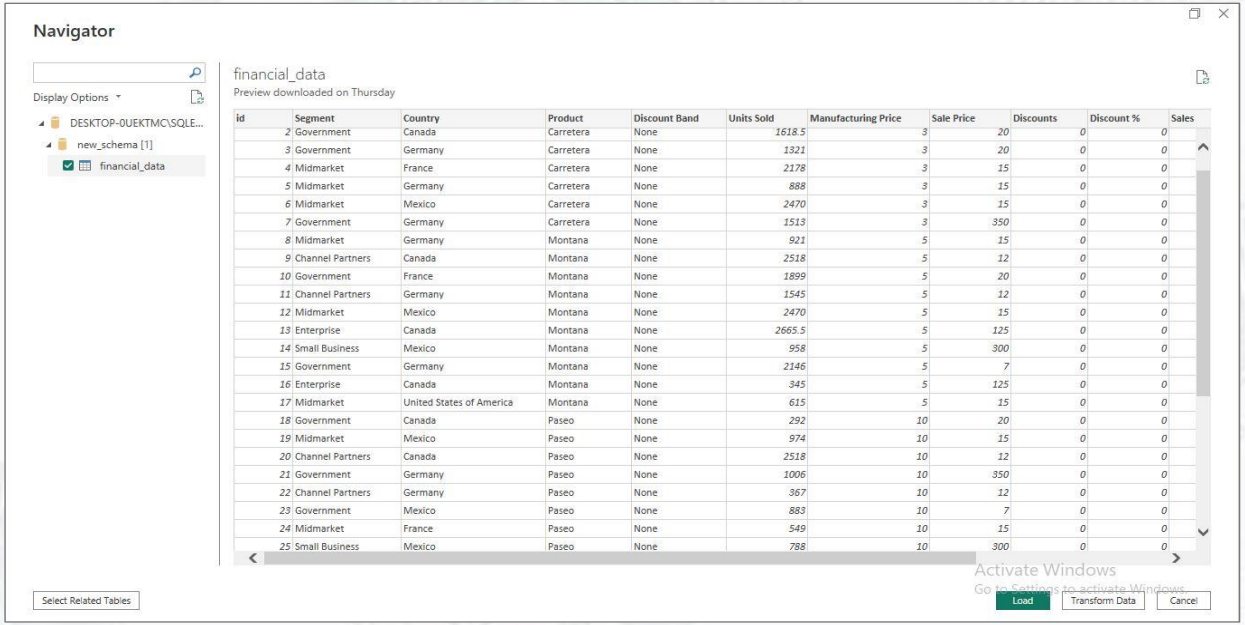
- Data Import from SQL



Import the Data from the SQL Database into PowerBI

skill academy

- Data Loading

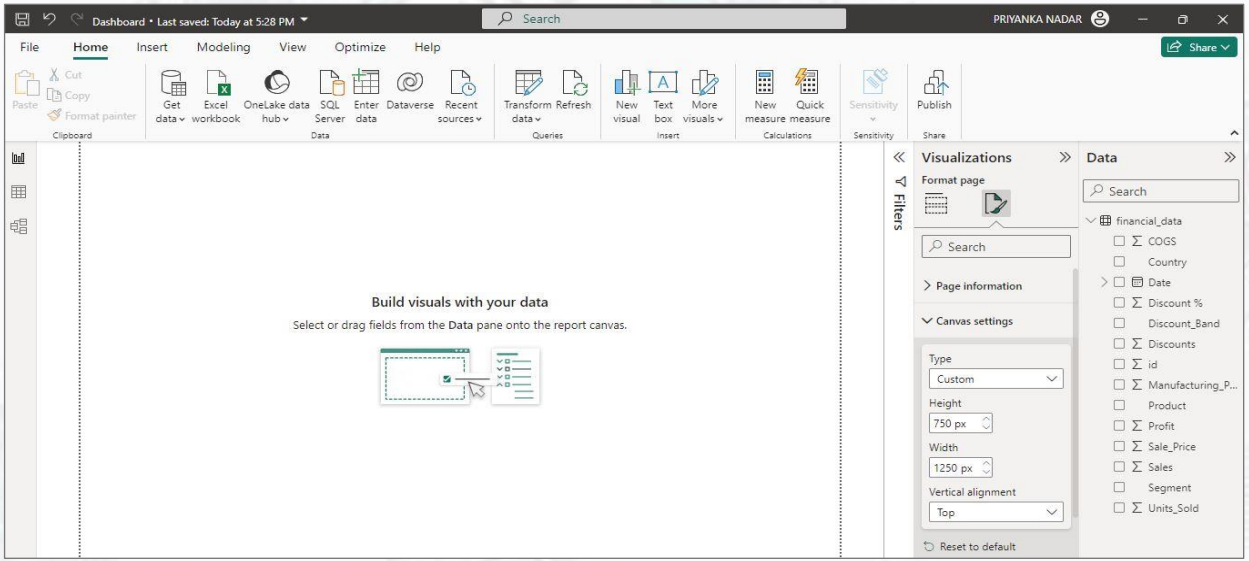


id	Segment	Country	Product	Discount Band	Units Sold	Manufacturing Price	Sale Price	Discounts	Discount %	Sales
2	Government	Canada	Carretera	None	1618.5	3	20	0	0	
3	Government	Germany	Carretera	None	1321	3	20	0	0	
4	Midmarket	France	Carretera	None	2178	3	15	0	0	
5	Midmarket	Germany	Carretera	None	888	3	15	0	0	
6	Midmarket	Mexico	Carretera	None	2470	3	15	0	0	
7	Government	Germany	Carretera	None	1513	3	350	0	0	
8	Midmarket	Germany	Montana	None	921	5	15	0	0	
9	Channel Partners	Canada	Montana	None	2518	5	12	0	0	
10	Government	France	Montana	None	1899	5	20	0	0	
11	Channel Partners	Germany	Montana	None	1545	5	12	0	0	
12	Midmarket	Mexico	Montana	None	2470	5	15	0	0	
13	Enterprise	Canada	Montana	None	2665.5	5	125	0	0	
14	Small Business	Mexico	Montana	None	958	5	300	0	0	
15	Government	Germany	Montana	None	2146	5	7	0	0	
16	Enterprise	Canada	Montana	None	345	5	125	0	0	
17	Midmarket	United States of America	Montana	None	615	5	15	0	0	
18	Government	Canada	Paseo	None	292	10	20	0	0	
19	Midmarket	Mexico	Paseo	None	974	10	15	0	0	
20	Channel Partners	Canada	Paseo	None	2518	10	12	0	0	
21	Government	Germany	Paseo	None	1006	10	350	0	0	
22	Channel Partners	Germany	Paseo	None	367	10	12	0	0	
23	Government	Mexico	Paseo	None	883	10	7	0	0	
24	Midmarket	France	Paseo	None	549	10	15	0	0	
25	Small Business	Mexico	Paseo	None	788	10	300	0	0	

Import the Data from the SQL Database into PowerBI

skill academy

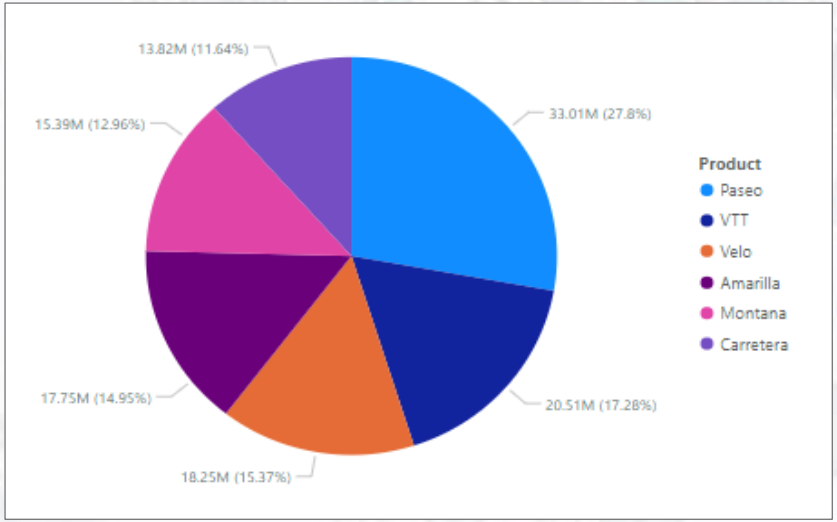
- Data Loaded Successfully
- Prepare viz, add filters, and assemble them all in a dashboard (canvas size is 1250 x 750 pixels)



Answer the questions by using visualization tools

skill academy

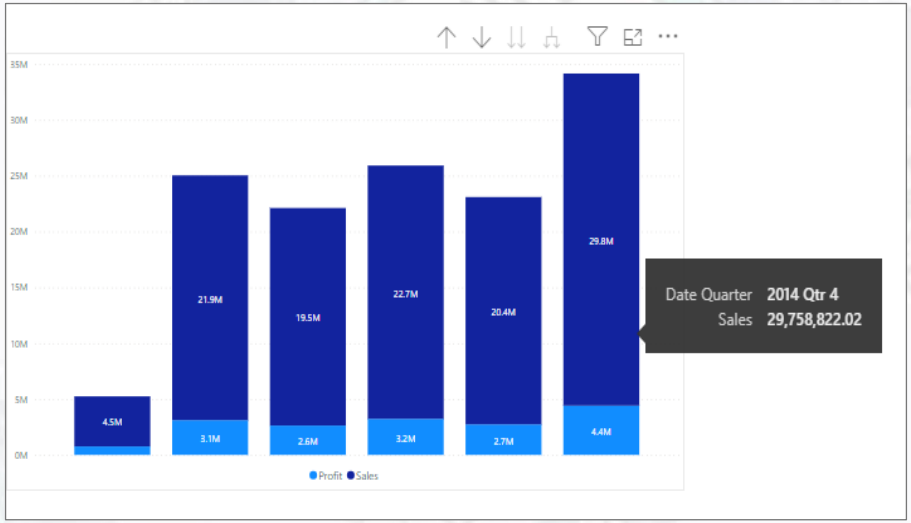
- Product wise Sales - given in the visualization below
- Bottom 3 products(inferring it as products with least sales) - Carretera, Montana and Amarilla



Answer the questions by using visualization tools

skill academy

- Profit and sales by quarter - calculated for all the quarters present according to the dates in dataset
- Yearly sales and profit



Year	Total Profit	Total Sales	Total Units Sold
2014	13,015,237.75	92,311,094.75	861,132.00
2013	3,878,464.51	26,415,255.51	264,674.00
Total	**16,893,702.26**	**118,726,350.26**	**1,125,806.00**

Answer the questions by using visualization tools

skill academy

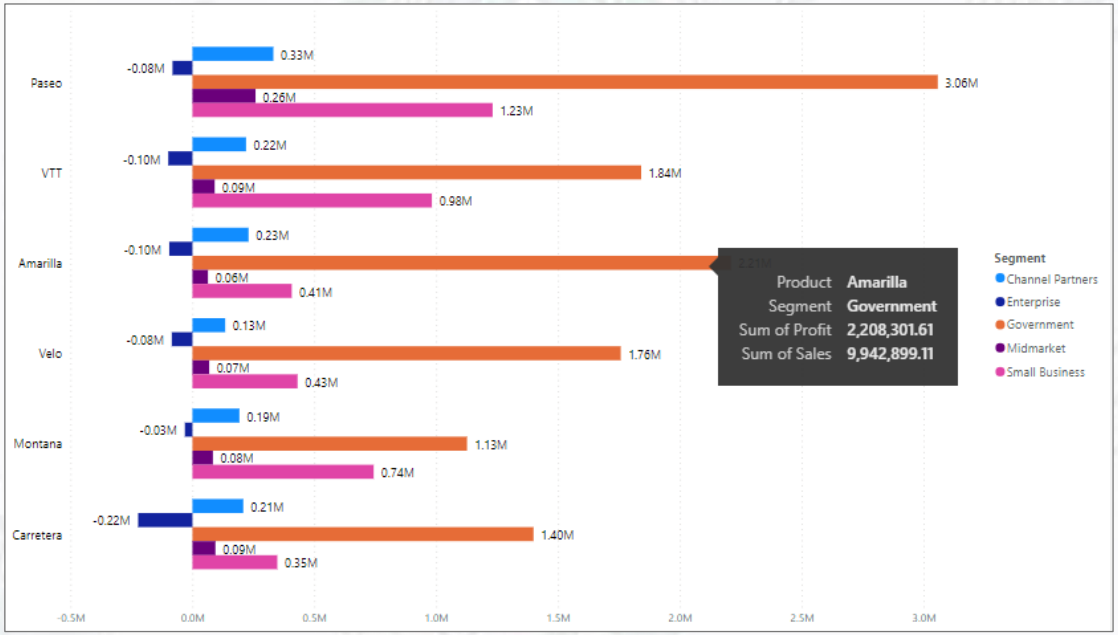
- Country-wise Sales
- Top 2 countries (in terms of sales - USA and Canada, profit – France and Germany, units sold – Canada and France)

Country	Sum of Sales	Sum of Units_Sold	Sum of Profit
United States of America	25,029,830.17	232,627.50	2,995,540.67
Canada	24,887,654.89	247,428.50	3,529,228.89
France	24,354,172.28	240,931.00	3,781,020.78
Germany	23,505,340.82	201,494.00	3,680,388.82
Mexico	20,949,352.11	203,325.00	2,907,523.11
Total	**118,726,350.26**	**1,125,806.00**	**16,893,702.26**

Answer the questions by using visualization tools

skill academy

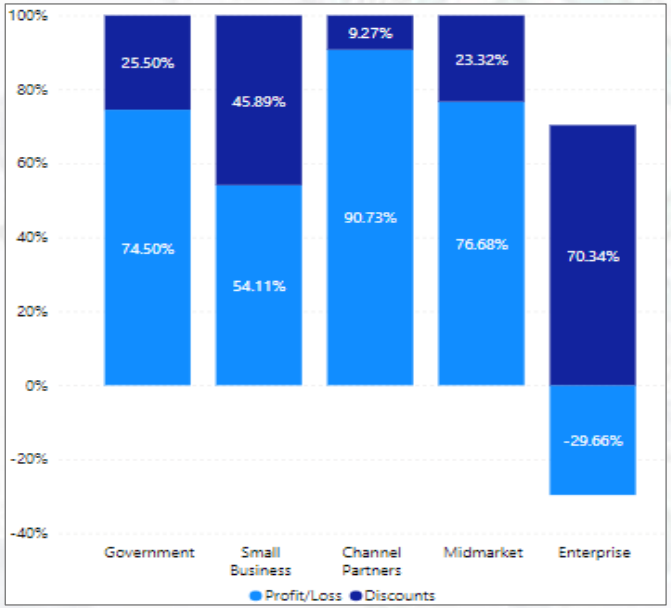
- Segment-wise and product-wise sales/profit

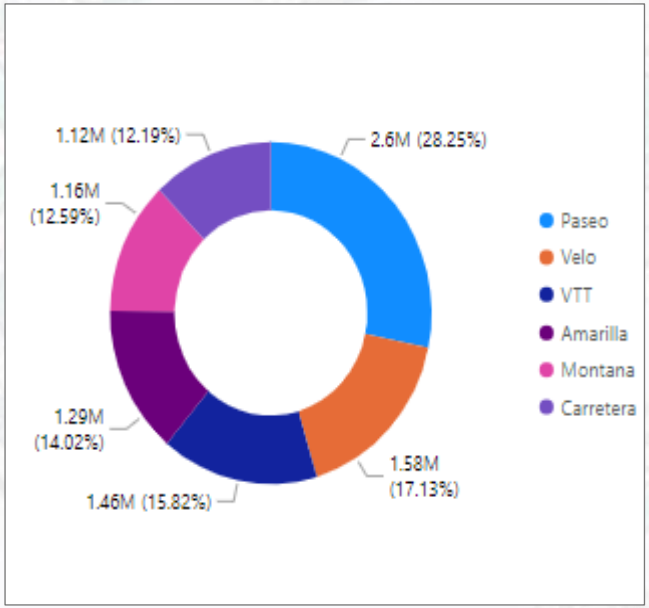


Answer the questions by using visualization tools

skill academy

- Segment-wise profit, discounts
- Product-wise discount





Overall Dashboard

skill academy

Pages

- Dashboard1
- Dashboard2
- Profit and Sales by Quarter
- Yearly Profit, Sales and Units Sold
- Profit/Loss, Sales by Segment and Product
- Country Details
- Sales by Product
- Discount and Profit/Loss by Segments
- Product-wise Discount

File | Export | Share | Chat in Teams | Explore this data | Get insights | Subscribe to report | Edit

Financial Data Analysis Dashboard (pt. 1)

Year-wise Profit, Sales and Units Sold

Year	Profit	Sales	Units Sold
2013	3,878,464.51	26,415,255.51	264,674.00
2014	13,015,237.75	92,311,094.75	861,132.00
Total	**16,893,702.26**	**118,726,350.26**	**1,125,806.00**

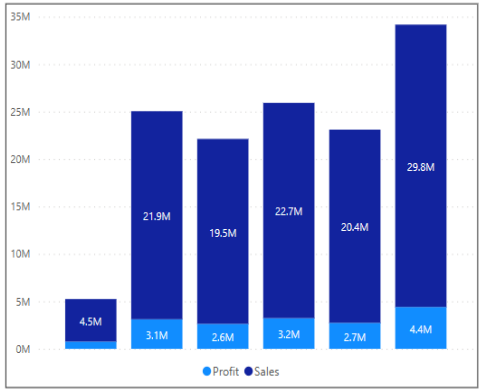
Country-wise Profit, Sales and Units Sold

Country	Profit	Sales	Units Sold
Mexico	2,907,523.11	20,949,352.11	203,325.00
Germany	3,680,388.82	23,505,340.82	201,494.00
France	3,781,020.78	24,354,172.28	240,931.00
Canada	3,529,228.89	24,887,654.89	247,428.50
United States of America	2,995,540.67	25,029,830.17	232,627.50
Total	**16,893,702.26**	**118,726,350.26**	**1,125,806.00**

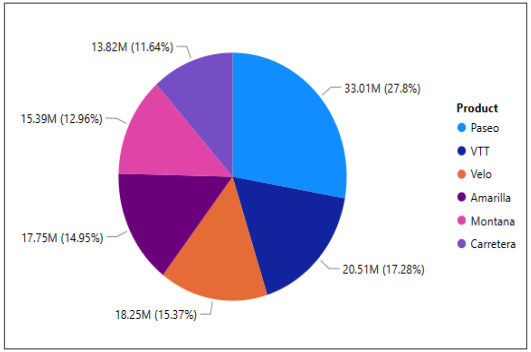
Product-wise Discount

Product	Total Discount
Carretera	1,122,212.62
Montana	1,159,032.62
Amarilla	1,290,163.44
VTT	1,456,612.48
Velo	1,576,709.04
Paseo	2,600,518.05

Quarterly Sales and Profit



Total Sales by Product



Filters

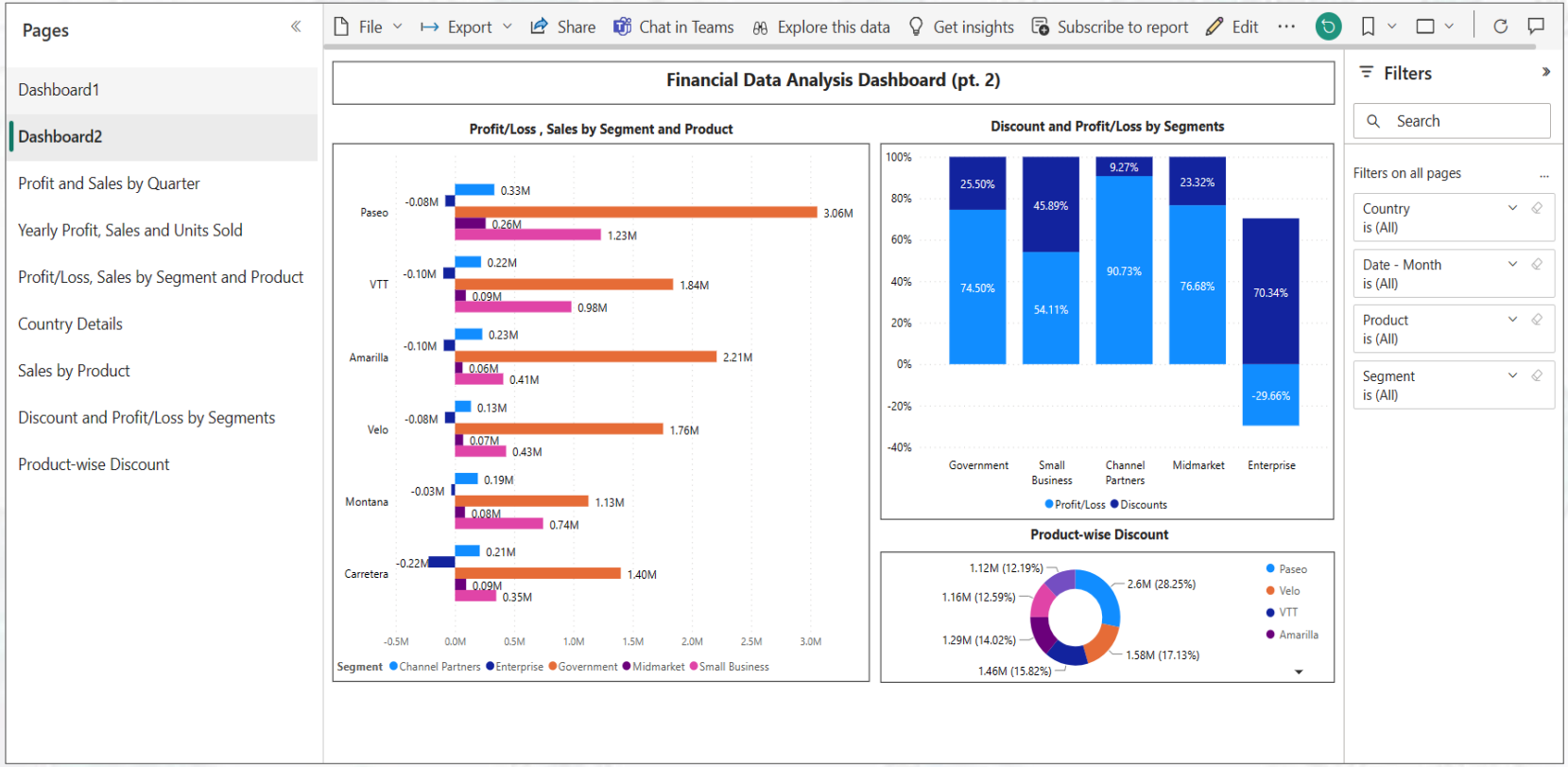
Search

Filters on all pages

- Country is (All)
- Date - Month is (All)
- Product is (All)
- Segment is (All)

Overall Dashboard

skill academy



Conclusion and Inferences

skill academy

- Paseo has the highest sales in the ‘Government’ sector
- Also, customers trust other products from this segment the most
- Germany has sold least number of units but in terms of profit, it has outgrown significantly
- With good marketing campaigns, sales can be increased
- USA currently looks like it has incurred loss even after higher sales, hence we need to analyze the bigger picture (supply chain, customer satisfaction, etc.) in this region
- As the sales are good, it has a scope for improvement
- Last quarter of 2014 has been a great time in terms of sales, kudos to the entire team
- ‘Enterprise’ segment seems to have incurred loss, probably because of the high discount rates offered in products
- But at the same time, Paseo, which has the highest number of sales, has offered most discounts to customers
- We need to study the customers’ interests in depth and also interact with the sales people to understand their point of view for further analysis

Endnotes

skill academy

Reference Links:-

Excel, PowerBI and SQL :
https://drive.google.com/drive/folders/1J9Gd3y_ub9JogvGBIdQqc9QFVdbnNyfn?usp=sharing

THANK YOU!