PHARMA DATA ANALYSIS

Presented by:
Ajay Yogesh Varma
ajayvarma2003@gmail.com
https://www.linkedin.com/in/ajay-varma-b427b7225/

Data Description

Dataset Overview:

- **Rows:** 254,082
- **Columns:** City, Country, Latitude, Longitude, Channel, Sub-channel, Product Name, Product Class, Quantity, Price, Sales, Month, Year, Name of Sales Rep, Manager, Sales Team.

Key Columns:

1. **City:** Represents the city where the sales transaction occurred.
2. **Country:** Indicates the country associated with the sales data.
3. **Channel:** Describes the distribution channel through which the product was sold.
4. **Sub-channel:** Provides additional granularity within distribution channels.
5. **Product Name:** Identifies the name of the product sold.
6. **Product Class:** Classifies the product into specific categories.
7. **Quantity:** Indicates the quantity of products sold in each transaction.
8. **Price:** Represents the unit price of the product.
9. **Sales:** Total sales amount for each transaction.
10. **Month:** Specifies the month in which the sales transaction occurred.
11. **Year:** Indicates the year of the sales transaction.
12. **Name of Sales Rep:** Identifies the sales representative associated with the transaction.
13. **Manager:** Represents the manager overseeing the sales representative.
14. **Sales Team:** Specifies the sales team to which the sales representative belongs.

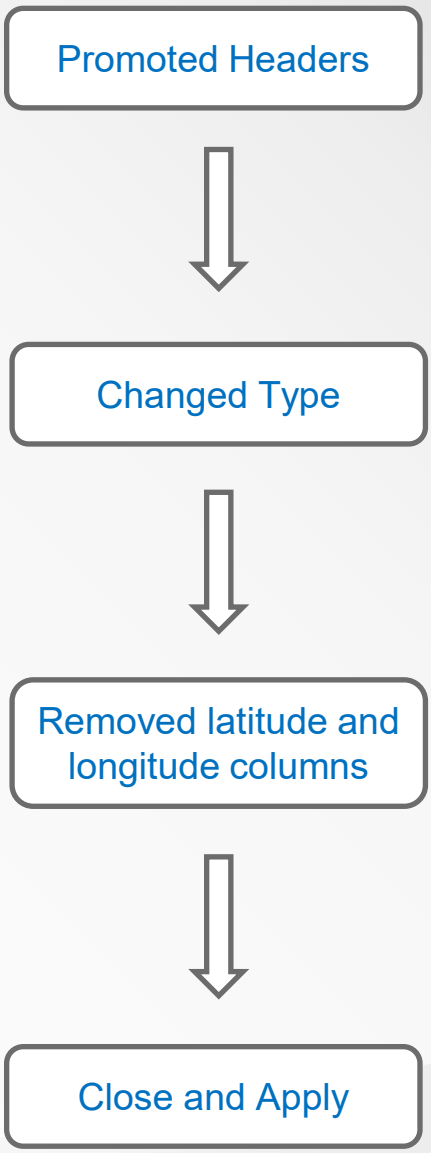
Dataset Characteristics:

- **Scope:** Sales data across various cities and countries.
- **Granularity:** Transaction-level data capturing individual sales.
- **Temporal Dimension:** Monthly sales data across multiple years.
- **Hierarchy:** Hierarchical structure with information on products, sales reps, and teams.

Data Transformation Overview:

Objective:
To ensure that the raw data is cleaned, structured, and organized in a way that facilitates meaningful analysis.

Steps Performed:



Task 1: Schema Design

The main data after transformation is saved as Pharma_data and all the next analysis performed on this data.

Objective:

To create a Power BI data model with appropriate tables and relationships, considering key columns such as Customer Name, City, and others.

```
1 Sales = SUMMARIZE( Pharma_data,Pharma_data[City], Pharma_data[Country], Pharma_data[Channel], Pharma_data[Sub-channel], Pharma_data[Product Name], Pharma_data[Product Class],Pharma_data[Month], Pharma_data[Year], Pharma_data[Name of Sales Rep],Pharma_data[Manager], Pharma_data[Quantity],Pharma_data[Price],Pharma_data[Sales])
```

City	Country	Channel	Sub-channel	Product Name	Product Class	Month	Year	Name of Sales Rep	Manager	Quantity	Price	Sales
Torgau	Germany	Hospital	Private	Adalatamine	Antiseptics	January	2017	Jessica Smith	Britanny Bold	10	794	7940
MÃ¶nchengladbach	Germany	Hospital	Private	Menogine	Antiseptics	February	2017	Jessica Smith	Britanny Bold	10	742	7420
Frankenthal	Germany	Hospital	Private	Ribabyclor	Antiseptics	February	2017	Jessica Smith	Britanny Bold	10	469	4690
Bietigheim-Bissingen	Germany	Hospital	Private	Lansostral Acgestin	Antiseptics	February	2017	Jessica Smith	Britanny Bold	10	627	6270
JÃ¼lich	Germany	Hospital	Private	Magnemunex	Antiseptics	February	2017	Jessica Smith	Britanny Bold	10	689	6890
Waldshut-Tiengen	Germany	Hospital	Private	Novakyn Tracprox	Antiseptics	March	2017	Jessica Smith	Britanny Bold	10	288	2880
Friedrichsdorf	Germany	Hospital	Private	Adriafinil Ehtymara	Antiseptics	March	2017	Jessica Smith	Britanny Bold	10	535	5350
Gummersbach	Germany	Hospital	Private	Travoloride	Antiseptics	May	2017	Jessica Smith	Britanny Bold	10	785	7850
Andernach	Germany	Hospital	Private	Panpion	Antiseptics	May	2017	Jessica Smith	Britanny Bold	10	692	6920
Bad Oldesloe	Germany	Hospital	Private	Malaxolol Aeroprosyn	Antiseptics	May	2017	Jessica Smith	Britanny Bold	10	502	5020
Werl	Germany	Hospital	Private	Zonitonin	Antiseptics	May	2017	Jessica Smith	Britanny Bold	10	518	5180
Husum	Germany	Hospital	Private	Atrabicin Alkerotec	Antiseptics	June	2017	Jessica Smith	Britanny Bold	10	249	2490
Wittenberg	Germany	Hospital	Private	Dantocept Ferurenone	Antiseptics	July	2017	Jessica Smith	Britanny Bold	10	307	3070
Geretsried	Germany	Hospital	Private	Lansostral Acgestin	Antiseptics	July	2017	Jessica Smith	Britanny Bold	10	627	6270
Solingen	Germany	Hospital	Private	Benzoprox	Antiseptics	July	2017	Jessica Smith	Britanny Bold	10	284	2840
Husum	Germany	Hospital	Private	Abtasol	Antiseptics	August	2017	Jessica Smith	Britanny Bold	10	754	7540
Bayreuth	Germany	Hospital	Private	Biotasol Siloderm	Antiseptics	August	2017	Jessica Smith	Britanny Bold	10	373	3730
Aichach	Germany	Hospital	Private	Alisteride Pemidizem	Antiseptics	August	2017	Jessica Smith	Britanny Bold	10	35	350
Aachen	Germany	Hospital	Private	Factolamide	Antiseptics	October	2017	Jessica Smith	Britanny Bold	10	345	3450
Starnberg	Germany	Hospital	Private	Proderal	Antiseptics	November	2017	Jessica Smith	Britanny Bold	10	34	340
Bayreuth	Germany	Hospital	Private	Diaxolol	Antiseptics	November	2017	Jessica Smith	Britanny Bold	10	765	7650
Ravensburg	Germany	Hospital	Private	Alisteride Pemidizem	Antiseptics	December	2017	Jessica Smith	Britanny Bold	10	35	350
Frankenthal	Germany	Hospital	Private	Effidomide Evofribrate	Antiseptics	January	2018	Jessica Smith	Britanny Bold	10	612	6120
Siegburg	Germany	Hospital	Private	Adriafinil Ehtymara	Antiseptics	February	2018	Jessica Smith	Britanny Bold	10	535	5350
Freital	Germany	Hospital	Private	Ceretosine	Antiseptics	February	2018	Jessica Smith	Britanny Bold	10	343	3430
Marktoberdorf	Germany	Hospital	Private	Choriogestrel	Antiseptics	February	2018	Jessica Smith	Britanny Bold	10	347	3470
Neckarsulm	Germany	Hospital	Private	Diaxolol	Antiseptics	March	2018	Jessica Smith	Britanny Bold	10	765	7650

able: Sales (2,54,060 rows)

Data

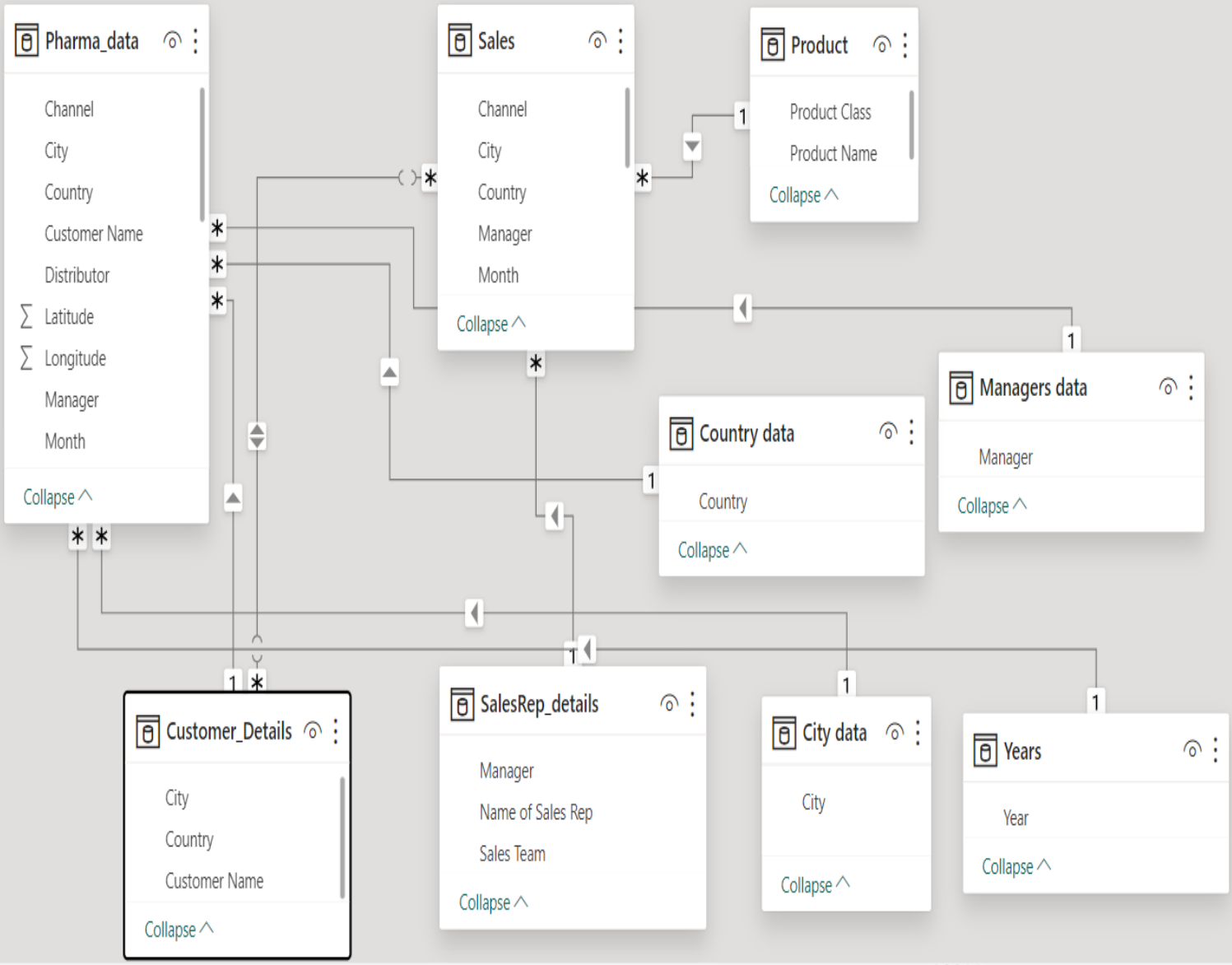
Search

- City data
- Country data
- Customer_Details
- Managers data
- Pharma_data
- Product
- Sales
- SalesRep_details
- Years

Necessary tables created to established relationship

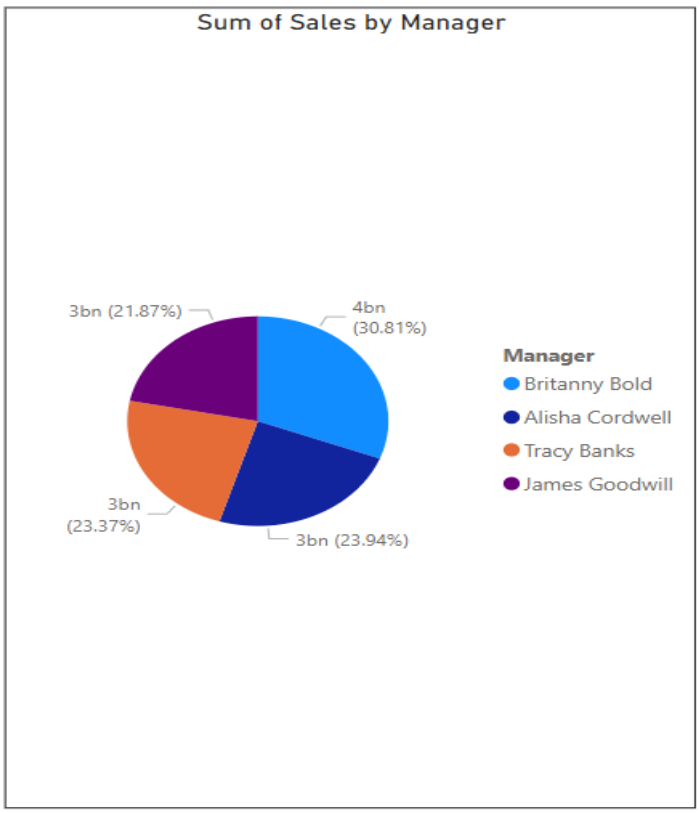
Task 2: Relationships

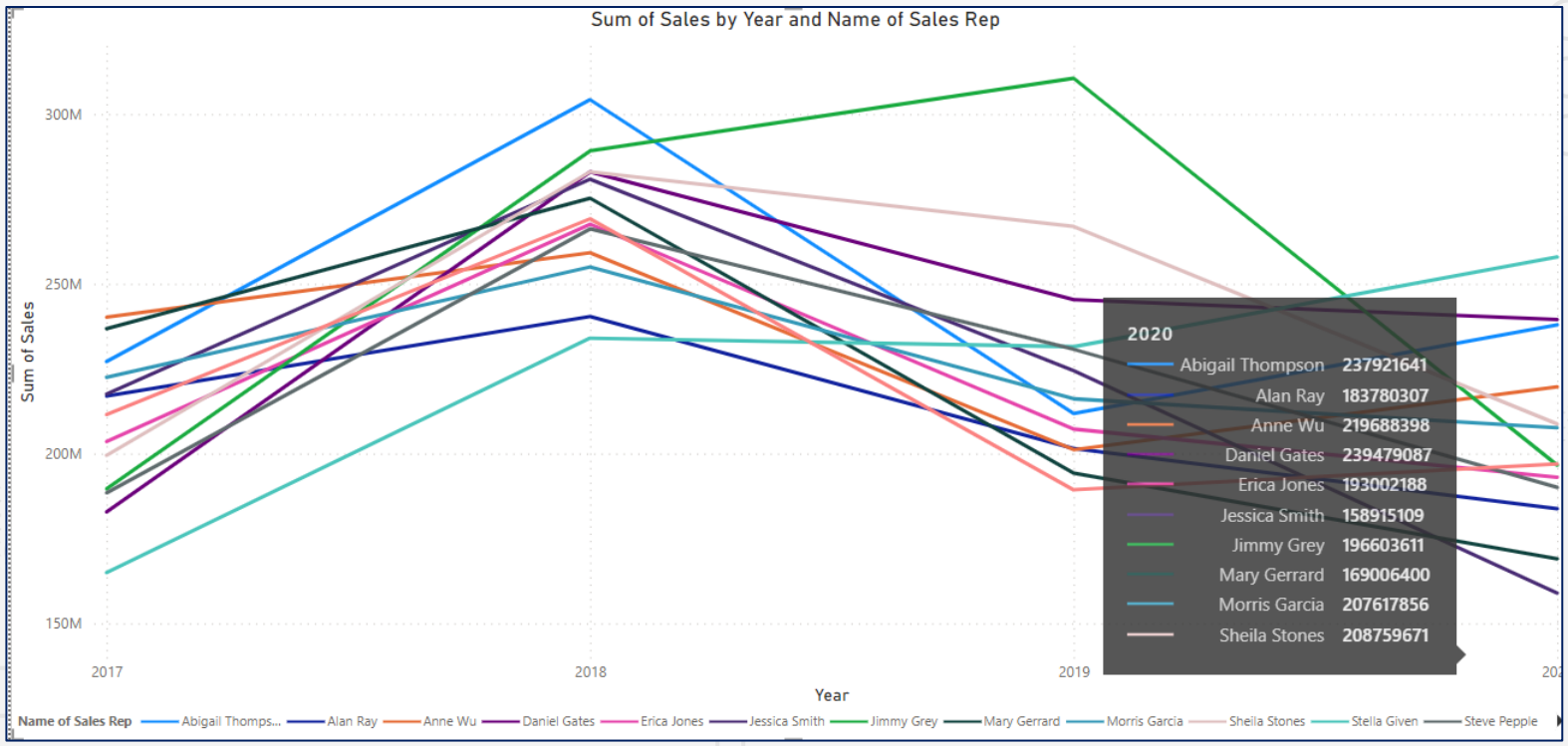
Objective:
To establish necessary relationships between the tables in Power BI data model.



Task 3: Role-Playing Dimensions

Objective:
To demonstrate how I handle role-playing dimensions for "Sales Rep" and "Manager" in your Power BI data model.





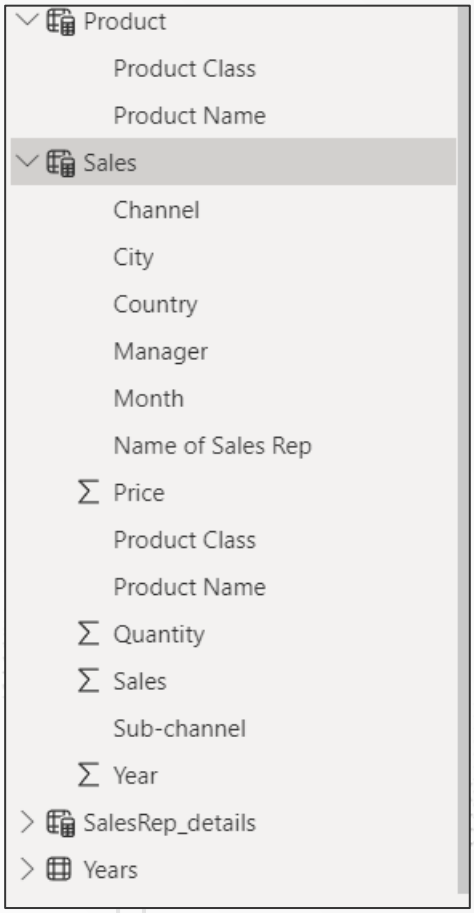
Task 4: Schemas

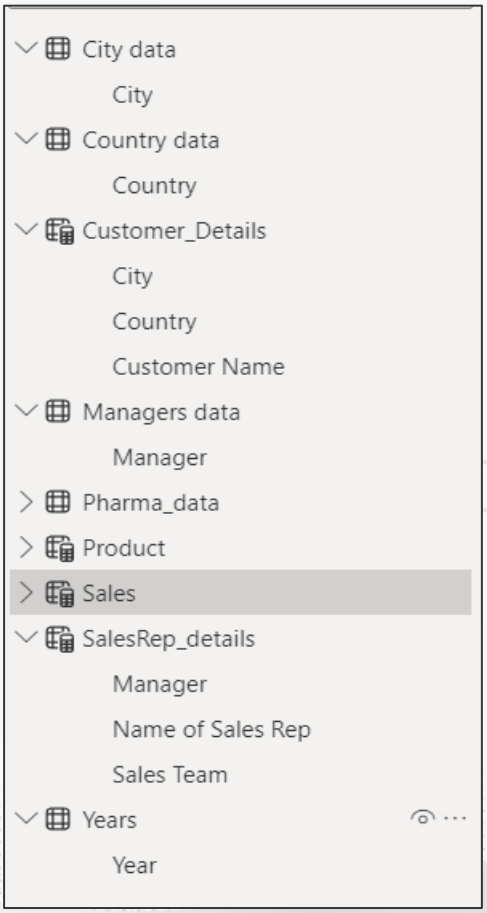
Objective:
To explain how I built a star schema based on my data and elaborate on how this schema design optimizes report performance.

As we delve into the architecture of our Power BI data model, the star schema emerges as a key design choice. The star schema revolves around a central fact table—in our case, Pharma_data—surrounded by related dimension tables: Customers, Products, and Sales Reps.

The fact table, Pharma_data, contains transactional data, and the dimension tables provide additional context to this data. This design not only simplifies the organization of our dataset but also brings forth several advantages.

As we delve into the architecture of our Power BI data model, the star schema emerges as a key design choice. The star schema revolves around a central fact table—in our case, Pharma_data—surrounded by related dimension tables: Customers, Products, and Sales Reps. Refer slide no. 5.





Task 5: Row-Level Security

Objective:

To demonstrate the implementation of row-level security in your Power BI data model and illustrate its impact on specific sales teams.

In our scenario, we've implemented row-level security to restrict access for a particular sales team. This is particularly crucial for maintaining confidentiality and controlling who can view certain pieces of information.

Task 6: Calculated Columns vs. Measures

Objective:

To calculate the total sales for each product both as a calculated column and as a measure, and compare the results.

```
1 TotalCost = 'Pharma_data'[Quantity] * 'Pharma_data'[Price]
```

ub-channel	Product Name	Product Class	Quantity	Price	Sales	Month	Year	Name of Sales Rep	Manager	Sales Team	TotalCost	New Column from Example
tail	Spiroderall	Mood Stabilizers	10	653	6530	January	2017	Anne Wu	Britanny Bold	Delta	6530	Geislingen an der Steige Germany
tail	Pazofenac	Mood Stabilizers	10	123	1230	January	2017	Anne Wu	Britanny Bold	Delta	1230	LÃ¼neburg Germany
tail	Hepavice	Mood Stabilizers	10	140	1400	February	2017	Anne Wu	Britanny Bold	Delta	1400	Springe Germany
tail	Lovapur	Mood Stabilizers	10	551	5510	February	2017	Anne Wu	Britanny Bold	Delta	5510	Heidenheim Germany
tail	Presetron	Mood Stabilizers	10	216	2160	March	2017	Anne Wu	Britanny Bold	Delta	2160	Achern Germany
tail	Primatate Univatol	Mood Stabilizers	10	172	1720	March	2017	Anne Wu	Britanny Bold	Delta	1720	Bensheim Germany
tail	Angioparin Brimosumab	Mood Stabilizers	10	682	6820	April	2017	Anne Wu	Britanny Bold	Delta	6820	WeiÃŸenfels Germany
tail	Alarudin Azarolac	Mood Stabilizers	10	567	5670	April	2017	Anne Wu	Britanny Bold	Delta	5670	Bad Kissingen Germany
tail	Topipizole	Mood Stabilizers	10	368	3680	April	2017	Anne Wu	Britanny Bold	Delta	3680	Schleswig Germany
tail	Lovapur	Mood Stabilizers	10	551	5510	May	2017	Anne Wu	Britanny Bold	Delta	5510	Bad Kissingen Germany
tail	Factofribrate	Mood Stabilizers	10	99	990	May	2017	Anne Wu	Britanny Bold	Delta	990	UnterschleiÃŸheim Germany
tail	Trazozaprine	Mood Stabilizers	10	708	7080	June	2017	Anne Wu	Britanny Bold	Delta	7080	Herrenberg Germany
tail	Adriacaine	Mood Stabilizers	10	361	3610	July	2017	Anne Wu	Britanny Bold	Delta	3610	Herzogenrath Germany
tail	Clinolan Pitoletra	Mood Stabilizers	10	130	1300	July	2017	Anne Wu	Britanny Bold	Delta	1300	Weingarten Germany
tail	Lovavatol Azelavarix	Mood Stabilizers	10	106	1060	July	2017	Anne Wu	Britanny Bold	Delta	1060	WolfenbÃ¼ttel Germany
tail	Feruprazole	Mood Stabilizers	10	115	1150	August	2017	Anne Wu	Britanny Bold	Delta	1150	Northeim Germany
tail	Hepavice	Mood Stabilizers	10	140	1400	August	2017	Anne Wu	Britanny Bold	Delta	1400	Schwandorf Germany
tail	Feruprazole	Mood Stabilizers	10	115	1150	August	2017	Anne Wu	Britanny Bold	Delta	1150	Kevelaer Germany
tail	Solasteride	Mood Stabilizers	10	520	5200	September	2017	Anne Wu	Britanny Bold	Delta	5200	Leipzig Germany
tail	Cephovelam	Mood Stabilizers	10	542	5420	November	2017	Anne Wu	Britanny Bold	Delta	5420	Herzogenrath Germany
tail	Rositroban	Mood Stabilizers	10	481	4810	November	2017	Anne Wu	Britanny Bold	Delta	4810	Hof Germany
tail	Ergomethate	Mood Stabilizers	10	604	6040	January	2018	Anne Wu	Britanny Bold	Delta	6040	Pulheim Germany
tail	Pazofenac	Mood Stabilizers	10	123	1230	January	2018	Anne Wu	Britanny Bold	Delta	1230	Meerbusch Germany
tail	Liomulin	Mood Stabilizers	10	490	4900	February	2018	Anne Wu	Britanny Bold	Delta	4900	Friedrichshafen Germany
tail	Trazobalamin	Mood Stabilizers	10	762	7620	February	2018	Anne Wu	Britanny Bold	Delta	7620	Freiberg Germany
tail	Exotropin Empizine	Mood Stabilizers	10	785	7850	April	2018	Anne Wu	Britanny Bold	Delta	7850	Stralsund Germany
tail	Hepavice	Mood Stabilizers	10	140	1400	April	2018	Anne Wu	Britanny Bold	Delta	1400	Bad Reichenhall Germany

```
1 TotalSales_Measure = SUMX('Pharma_data', 'Pharma_data'[Quantity] * 'Pharma_data'[Price])
```

ub-channel	Product Name	Product Class	Quantity	Price	Sales	Month	Year	Name of Sales Rep	Manager	Sales Team	TotalCost	New Column from Example
tail	Spiroderall	Mood Stabilizers	10	653	6530	January	2017	Anne Wu	Britanny Bold	Delta	6530	Geislingen an der Steige Germany
tail	Pazofenac	Mood Stabilizers	10	123	1230	January	2017	Anne Wu	Britanny Bold	Delta	1230	LÃ¼neburg Germany
tail	Hepavice	Mood Stabilizers	10	140	1400	February	2017	Anne Wu	Britanny Bold	Delta	1400	Springe Germany
tail	Lovapur	Mood Stabilizers	10	551	5510	February	2017	Anne Wu	Britanny Bold	Delta	5510	Heidenheim Germany
tail	Presetron	Mood Stabilizers	10	216	2160	March	2017	Anne Wu	Britanny Bold	Delta	2160	Achern Germany
tail	Primatate Univatol	Mood Stabilizers	10	172	1720	March	2017	Anne Wu	Britanny Bold	Delta	1720	Bensheim Germany
tail	Angioparin Brimosumab	Mood Stabilizers	10	682	6820	April	2017	Anne Wu	Britanny Bold	Delta	6820	WeiÃŸenfels Germany
tail	Alarudin Azarolac	Mood Stabilizers	10	567	5670	April	2017	Anne Wu	Britanny Bold	Delta	5670	Bad Kissingen Germany
tail	Topipizole	Mood Stabilizers	10	368	3680	April	2017	Anne Wu	Britanny Bold	Delta	3680	Schleswig Germany
tail	Lovapur	Mood Stabilizers	10	551	5510	May	2017	Anne Wu	Britanny Bold	Delta	5510	Bad Kissingen Germany
tail	Factofribrate	Mood Stabilizers	10	99	990	May	2017	Anne Wu	Britanny Bold	Delta	990	UnterschleiÃŸheim Germany
tail	Trazozaprine	Mood Stabilizers	10	708	7080	June	2017	Anne Wu	Britanny Bold	Delta	7080	Herrenberg Germany
tail	Adriacaine	Mood Stabilizers	10	361	3610	July	2017	Anne Wu	Britanny Bold	Delta	3610	Herzogenrath Germany
tail	Clinolan Pitoletra	Mood Stabilizers	10	130	1300	July	2017	Anne Wu	Britanny Bold	Delta	1300	Weingarten Germany
tail	Lovavatol Azelavarix	Mood Stabilizers	10	106	1060	July	2017	Anne Wu	Britanny Bold	Delta	1060	WolfenbÃ¼ttel Germany
tail	Feruprazole	Mood Stabilizers	10	115	1150	August	2017	Anne Wu	Britanny Bold	Delta	1150	Northeim Germany
tail	Hepavice	Mood Stabilizers	10	140	1400	August	2017	Anne Wu	Britanny Bold	Delta	1400	Schwandorf Germany
tail	Feruprazole	Mood Stabilizers	10	115	1150	August	2017	Anne Wu	Britanny Bold	Delta	1150	Kevelaer Germany
tail	Solasteride	Mood Stabilizers	10	520	5200	September	2017	Anne Wu	Britanny Bold	Delta	5200	Leipzig Germany
tail	Cephovelam	Mood Stabilizers	10	542	5420	November	2017	Anne Wu	Britanny Bold	Delta	5420	Herzogenrath Germany

Data

Search

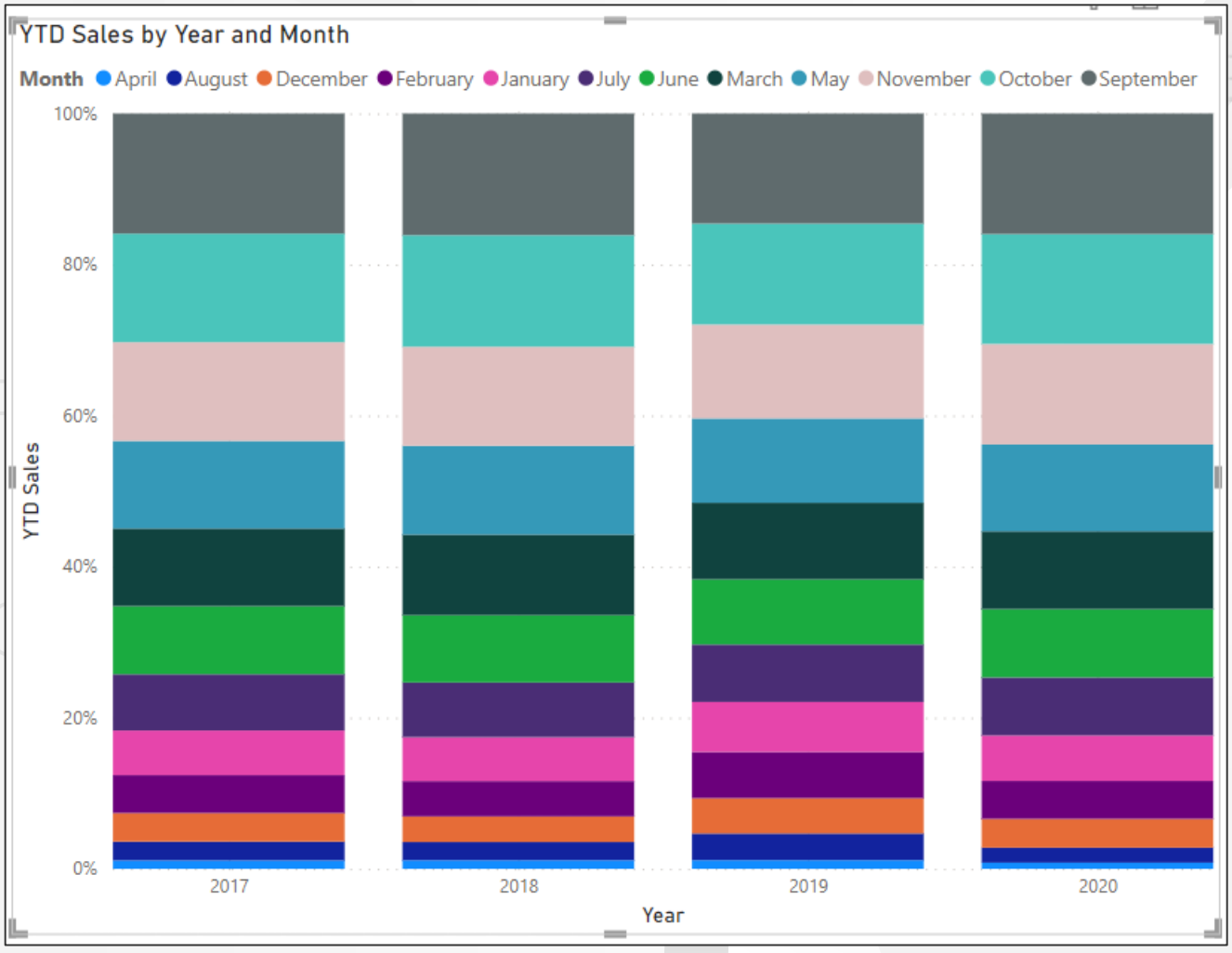
Month
Name of Sales Rep
New Column from Example
Price
Product Class
Product Name
Product Rank
Quantity
Sales
Sales Team
Sub-channel
Sub-channel Sales
Total Quantity by Sales Rep
TotalCost
TotalSales_Measure
Year
YoYGrowth

Task 7: Time Intelligence - YTD Sales

Objective:

To use DAX to create a measure that calculates the year-to-date (YTD) sales for each month.

Month	2017	2018	2019	2020
April	188800584	246410406	229413671	135409908
August	422635259	524580129	709669002	329777539
December	643394801	739462296	942750244	636061700
February	852390690	1014534905	1231505533	835367741
January	1004262874	1268465157	1329169609	1006092374
July	1264751650	1575464058	1530882305	1279428829
June	1531420041	1939462508	1735883225	1507794507
March	1746187997	2323181543	2038168233	1715824156
May	1969822796	2557634053	2246157080	1914935215
November	2221890306	2854471640	2501271665	2219718088
October	2441229131	3214306723	2688902224	2424561980
September	2701480741	3506897354	2930937133	2659672415
Total				

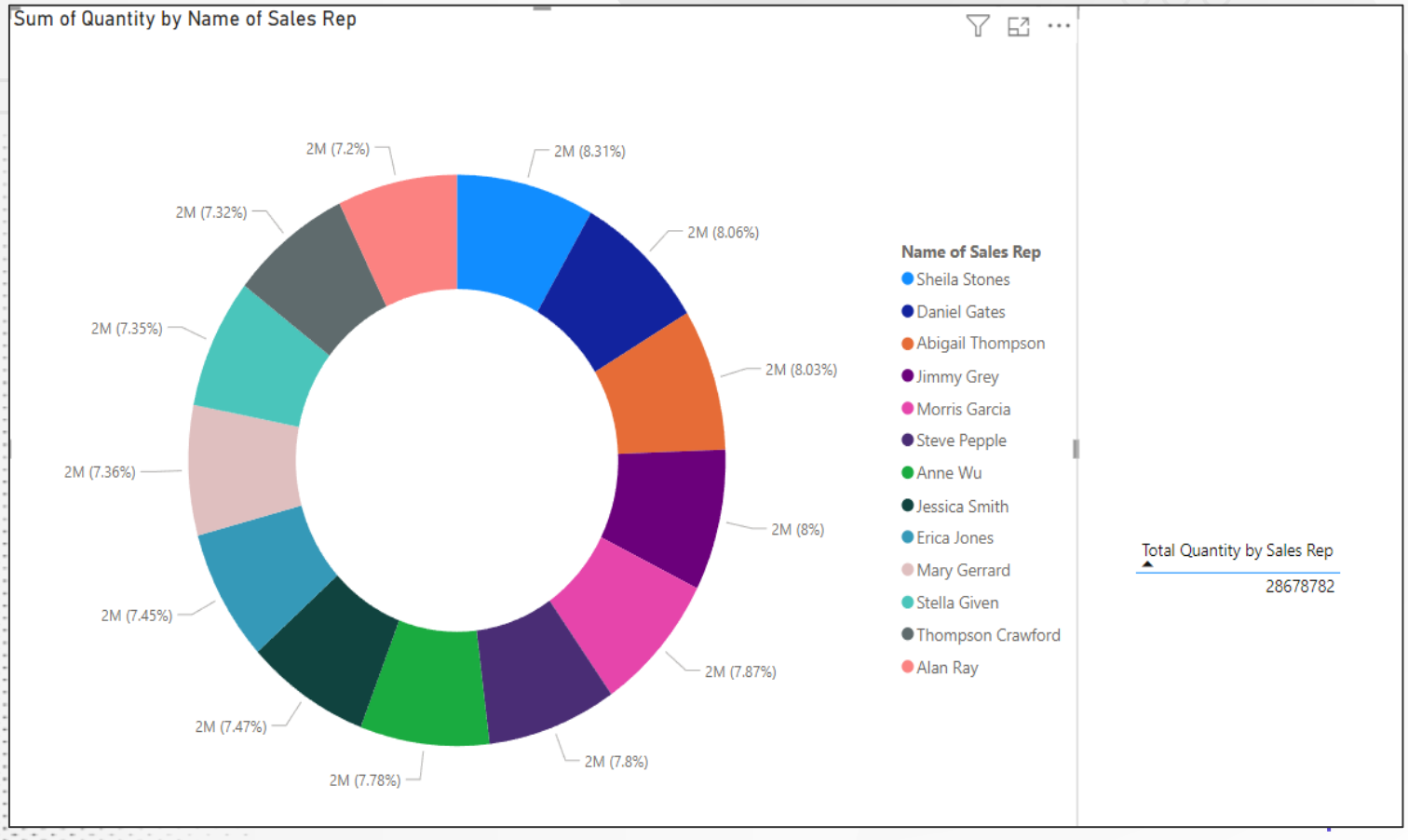


Task 8: Filter Context vs. Row Context - Total Quantity Sold by Each Sales Rep

Objective:
To write a DAX calculation that shows the total quantity sold by each sales rep, and explain how filter and row contexts apply.

I'll showcase the results by visualizing the Total Quantity Sold measure. This will give us a clear understanding of how each sales rep contributes to the overall quantity sold.

```
1  Total Quantity by Sales Rep =
2  CALCULATE(
3      SUM('Pharma_data'[Quantity]),
4      ALLEXCEPT('Pharma_data', 'Pharma_data'[Name of Sales Rep])
```

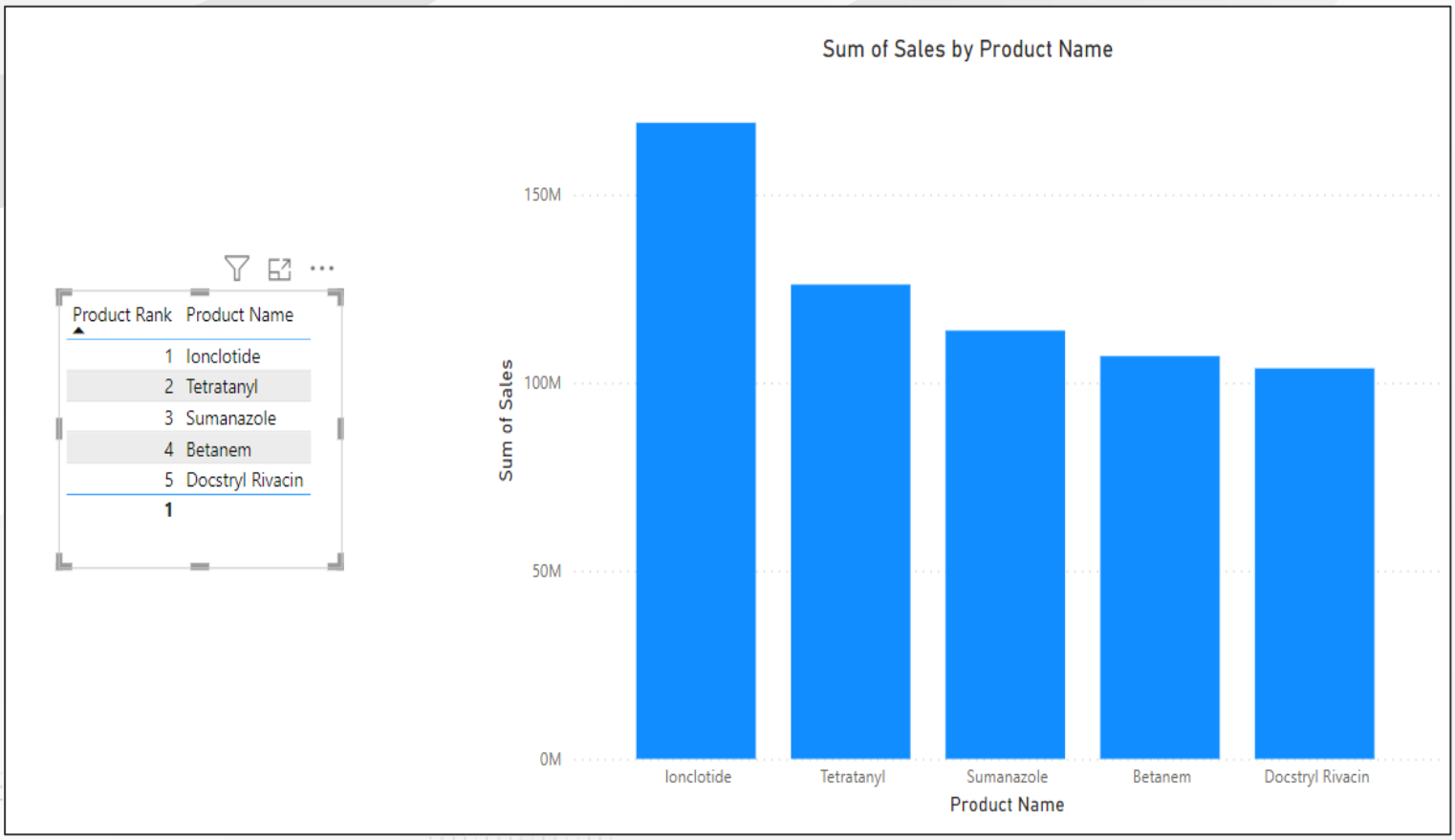


Task 9: Ranking - Top 5 Products by Sales

Objective:
To create a DAX measure that ranks products by sales and display the top 5 products by rank in a visual.

This visual will provide us with a clear view of the top 5 products by sales rank, offering insights into the high-performing products within our dataset.

```
Product Rank = RANKX(ALL('Pharma_data'[Product Name]), CALCULATE(SUM('Pharma_data'[Sales])))
```

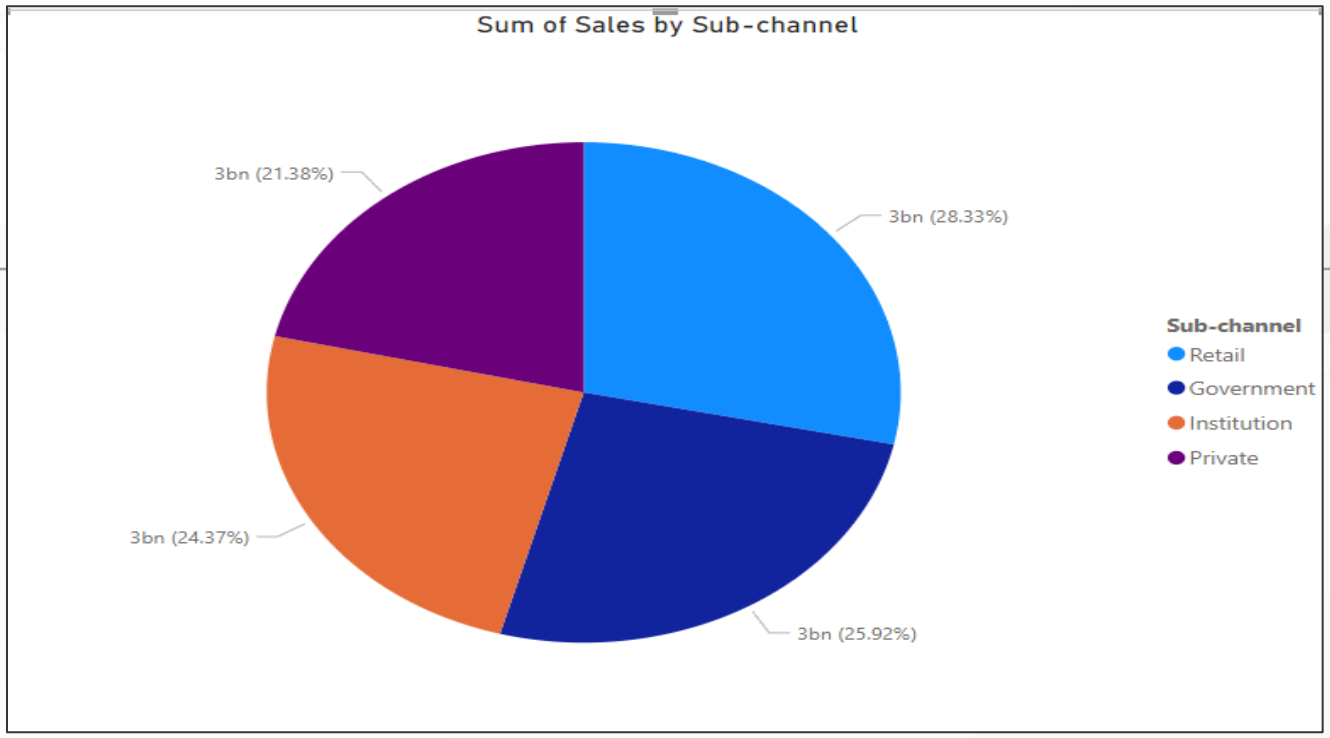


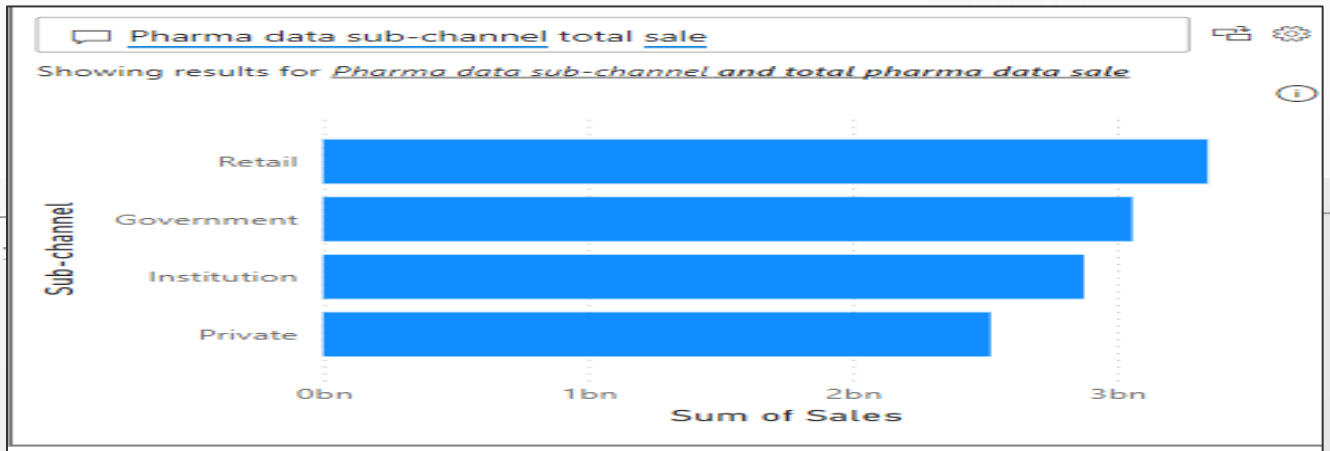
Task 10: Parent-Child Hierarchies - Summarizing Sales at Subcategory Level

Objective:

To create a DAX measure that summarizes sales at the subcategory level within a parent-child hierarchy.

```
Sub-channel Sales =
CALCULATE(
    SUM('Pharma_data'[Sales]),
    ALLEXCEPT('Pharma_data', 'Pharma_data'[Channel], 'Pharma_data'[Sub-channel])
)
```





Sub-channel Sales
11798987643

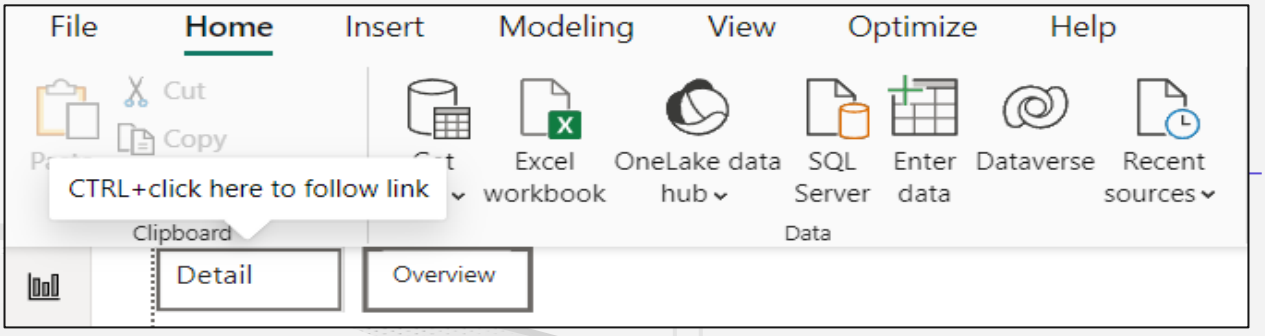
Task 11: Drill-Through - Summary to Detailed Data

Objective:
To build a report that allows users to drill through from a summary to detailed data, for example, starting from a map and drilling through to a table of individual sales for a specific city.

Now, let's explore the powerful feature of drill-through in Power BI, which allows users to seamlessly transition from a summary view to detailed data. In our report, I've identified a summary visual, such as a map showcasing sales by city.
I've created a dedicated drill-through page where users can delve into specific details based on their selections. For instance, clicking on a city in the summary map will drill through to a table displaying individual sales for that particular city.
I'll right-click on a city in the summary map and initiate the drill-through process.

Task 12: Bookmarks and Buttons - Navigating Between Report Pages

Objective:
To create a report with bookmarks and buttons, enabling users to navigate seamlessly between different pages or states within the report.

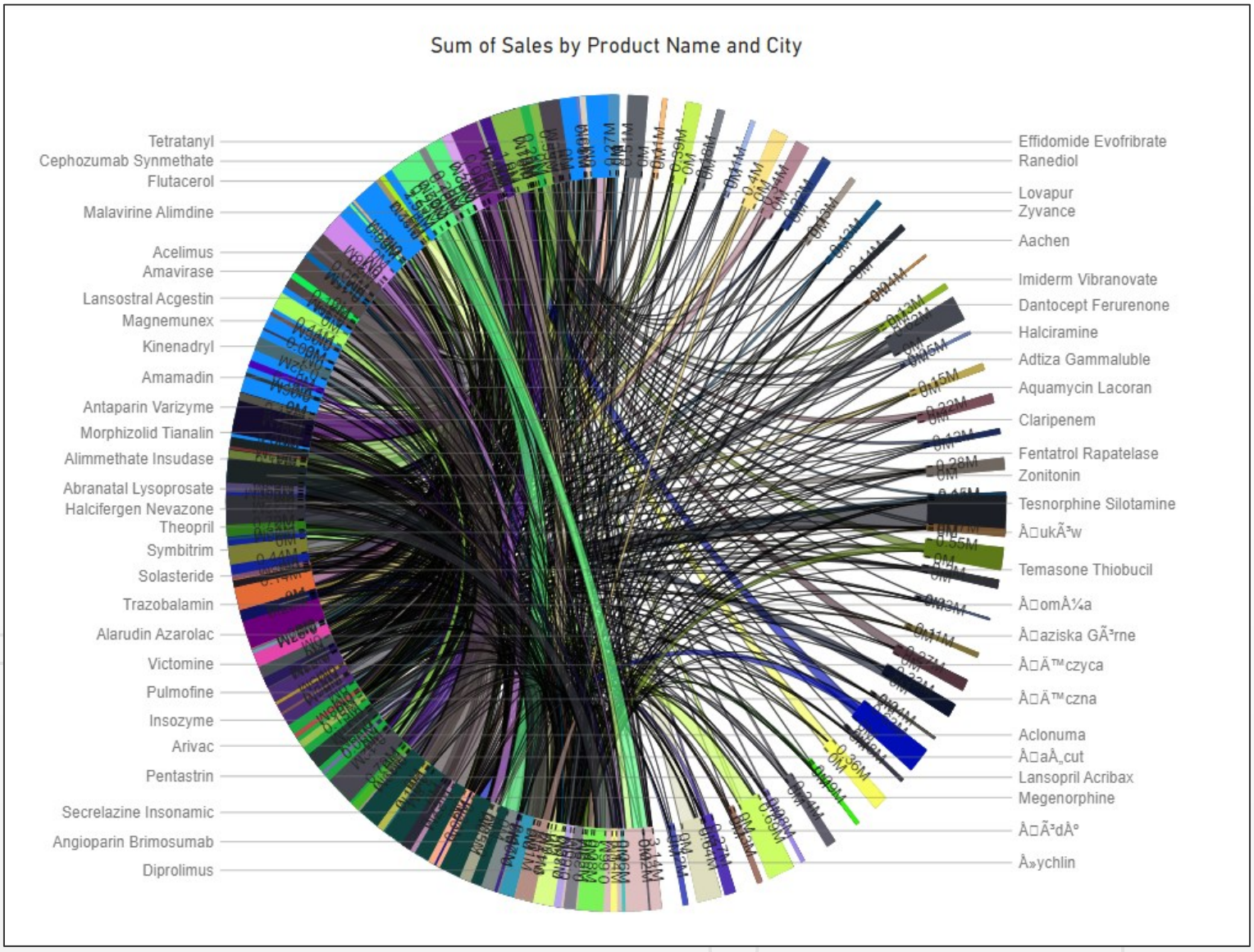


Task 13: Custom Visuals - Unique Visualization of Sales Data

Objective:

To use a custom visual in report to showcase sales data in a unique and engaging way.

Chord Chart



Task 14: Conditional Formatting - Color Change Based on Sales Target

Objective:

To apply conditional formatting to a measure in Power BI, causing it to change color when sales exceed a certain target value.

Let's add a layer of visual clarity to our report by applying conditional formatting to a key measure. I've identified the sales target measure, and by utilizing the conditional formatting options in Power BI, I've configured it to dynamically change color based on sales performance against the target.

Month	Sum of Quantity	Sum of Sales
January	1608717	674191145
April	1972091	800034569
May	2138724	865187215
October	2346474	971648359
February	2360568	972129828
December	2339283	975007112
September	2537671	1029987585
July	2478666	1042536828
June	2643066	1064033439
March	2665619	1108801648
November	2701803	1108802555
August	2886100	1186627360
Total	**28678782**	**11798987643**

Task 15: Calculated Columns - Total Cost Calculation

Objective:
To add a calculated column to the data model that calculates the total cost of each product based on the quantity and price.

Now, let's enrich our data model by adding a calculated column that calculates the total cost of each product. In the 'Pharma_data' table, I've created a calculated column named 'Total Cost,' which is a straightforward multiplication of the quantity and price columns.

```
TotalCost = 'Pharma_data'[Quantity] * 'Pharma_data'[Price]
```

TotalCost
6530
1230
1400
5510
2160
1720
6820
5670
3680
5510
990
7080
3610
1300
1060
1150
1400
1150
5200
5420
4810
6040
1230

Task 16: New Column from Example - City Categorization into Regions

Objective:
To add a new column to the data model using the "New Column from Example" feature, categorizing cities into regions based on a predefined mapping.

City	Country	New Column from Example
Geislingen an der Steige	Germany	Geislingen an der Steige Germany
LÃ¼neburg	Germany	LÃ¼neburg Germany
Springe	Germany	Springe Germany
Heidenheim	Germany	Heidenheim Germany
Achern	Germany	Achern Germany
Bensheim	Germany	Bensheim Germany
WeiÃŸenfels	Germany	WeiÃŸenfels Germany
Bad Kissingen	Germany	Bad Kissingen Germany
Schleswig	Germany	Schleswig Germany
Bad Kissingen	Germany	Bad Kissingen Germany
UnterschleiÃŸheim	Germany	UnterschleiÃŸheim Germany
Herrenberg	Germany	Herrenberg Germany

Task 17: Time-Based Calculations - Year-Over-Year (YoY) Growth

Objective:
To create a measure that calculates the year-over-year (YoY) growth in sales for each month in Power BI.

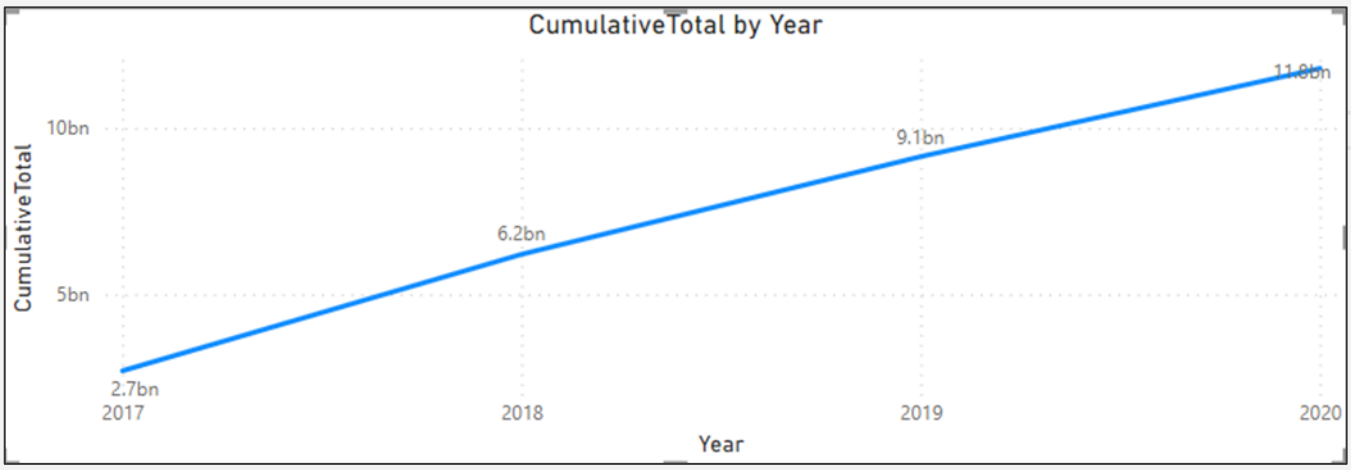
Now, let's delve into the time-based analysis of our sales data by calculating the year-over-year (YoY) growth. In the 'Pharma_data' table, I've created a measure named 'YoY Growth' that evaluates the difference in sales between the current month and the same month in the previous year.

```
YoYGrowth =
    VAR CurrentMonthSales = SUM('Pharma_data'[Sales])
    VAR PreviousYearSales = CALCULATE(
        SUM('Pharma_data'[Sales]),
        FILTER(
            ALL('Pharma_data'),
            'Pharma_data'[Month] = VALUES('Pharma_data'[Month]) &&
            'Pharma_data'[Year] = VALUES('Pharma_data'[Year]) - 1
        )
    )
    RETURN
        IF(ISBLANK(PreviousYearSales), BLANK(), (CurrentMonthSales - PreviousYearSales) / PreviousYearSales)
```

Task 18: Cumulative Total - Line Chart Visualization

Objective:
To create a measure that calculates the cumulative total of sales over time and visualize it in a line chart in Power BI.

```
CumulativeTotal =
    CALCULATE(
        SUM('Pharma_data'[Sales]),
        FILTER(
            ALL('Pharma_data'),
            'Pharma_data'[Month] <= MAX('Pharma_data'[Month]) &&
            'Pharma_data'[Year] <= MAX('Pharma_data'[Year])
        )
    )
```



Thanks!

Do you have any questions?
ajayvarma2003@gmail.com

+91 8600289210

https://www.linkedin.com/in/ajay-varma-b427b7225/