Adventure Works Dashboard Report

1. Introduction

- **Purpose:** The Adventure Works 2019 Analysis Dashboard is designed to provide insights into the fictional Adventure Works company's operations. This dashboard helps users analyze key business metrics, track performance, and make data-driven decisions based on the provided dataset. The dashboard aims to enhance understanding of sales trends, purchasing patterns, manufacturing efficiency, and customers demographics.

- **Company Overview:** Adventure Works is a manufacturer and retailer of bicycles and related products. It operates in various regions, including North America, Europe, and Asia.

- **Scope:** The dashboard includes data from various departments of the Adventure Works company, including Sales, and production. It presents a unified view of critical business metrics and facilitates the exploration of trends, and comparisons.

2. Data Sources

- **Data Sources Overview:** The dashboard utilizes the Adventure Works 2019 sample database provided by Microsoft. This database includes multiple tables related to different business functions such as:
- Sales: Sales Order Header, Sales Order Detail, Customer, Sales Territory
- Purchasing: Purchase Order Header, Purchase Order Detail, Vendor
- Manufacturing: Product, Bill of Materials, Work Order
- Human Resources: Employee, Department, Job Candidate

3. Data Import and Initial Exploration

The dataset was imported using pd.read_csv(), pointing to the specified CSV file. Initial exploration revealed essential information about the data structure through dt.info(), including data types and non-null counts.

Key insights include:

- **Data Types**: The dataset contains various data types, primarily numerical and categorical.
- **Summary Statistics**: Using dt.describe(), basic statistics such as mean, median, and standard deviation for numerical columns were obtained.

- **Missing Values**: The analysis of null values with dt.isnull().sum() indicated any columns needing attention.
- **Duplicates**: A check for duplicates with dt.duplicated().sum() confirmed the uniqueness of records.

4. Data Cleaning

Data cleaning involved:

- Formatting string columns (Name, CountryRegionCode, Group) to title case, stripping whitespace, and removing special characters using .str.title().str.strip().str.replace().
- Converting OrderDate and ShipDate to datetime format for further analysis.
- Calculating fulfillment time as the difference between day and shipday.

5. Metrics and KPIs

Key Metrics:

- **Sales Metrics:**
  - Total Sales
  - Total COGS
  - Profit & Profit Margin
  - Total Freight
  - Total Tax
  - Sales by years, and quarter
  - Sales by country region & Territory
  - Sales by category & product name
- **Customers Metrics:**
  - Number of customers
  - Number of stores
  - Number of orders
  - Number of order quantity
  - Number of customers by gender
  - Number of customers by marital status
  - Number of customers by education
  - Number of customers by age group
  - Top 5 customers (total sales)

- **Products Metrics**
  - Number of products
  - Order quantity by category
  - Order quantity by subcategory
  - Order quantity by year and quarter
  - Order quantity by territory
  - Order quantity by product

Definitions:

- **Total Sales:** Sum of all sales transactions within a specified period.
- **Sales by Region:** Breakdown of sales figures by geographic regions.
- Total COGS: Total cost of goods sold.

6. Dashboard Layout

Navigation: The dashboard is organized into several sections, each focusing on a specific business function. Users can navigate between different tabs or sections using the menu or navigation pane.

Visualizations:

- **Charts:** Bar charts, line graphs, pie charts to display sales trends, performance comparisons, and breakdowns by category.
- **Tables:** Detailed tables showing granular data for in-depth analysis.
- **Maps:** Geographic maps to visualize sales and performance metrics by region.
- **Filters:** Dropdown menus and sliders for filtering data by date ranges, regions, or product categories.

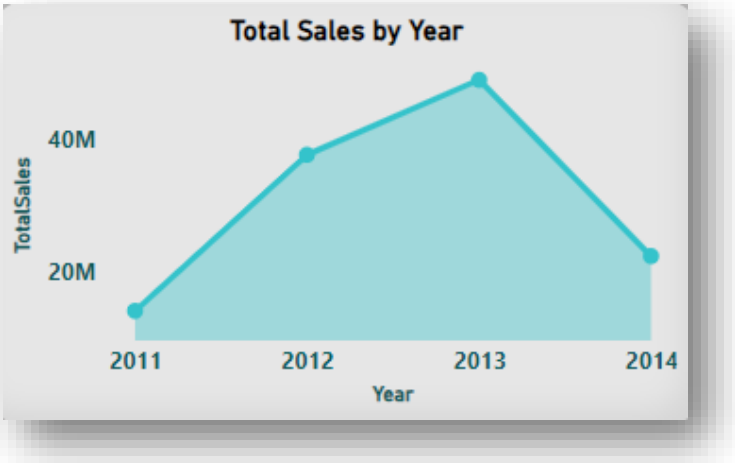
7. Analysis and Insights

- **Key Performance Indicators:**

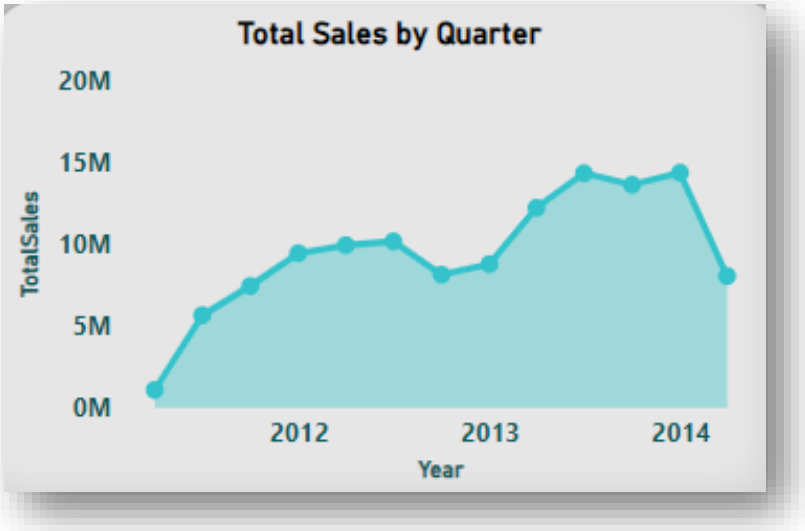
Total sales: 123.22 M	Total tax: 10.19 M	No. of orders: 31 K
Total COGS: 100.47 M	Total freight: 3.18 M	Total order quantity: 275 K
Profit: 22.74 M	No. of customers: 20 K	No. of products: 504
Profit margin: 18.46 %	No. of stores: 702	

➢ **Insights:**

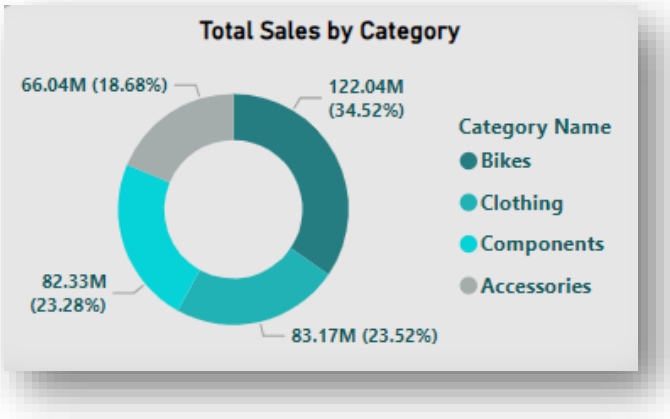
- In the year 2013, the company witnessed remarkable sales, reaching a total of 48,963,887.96. Following closely behind, the sales in 2012 amounted to 37,675,700.31. Conversely, the lowest sales were recorded in 2011. Through careful analysis, it was observed that the sales performance exhibited an upward trend, steadily improving as the years progressed until reaching 2014.

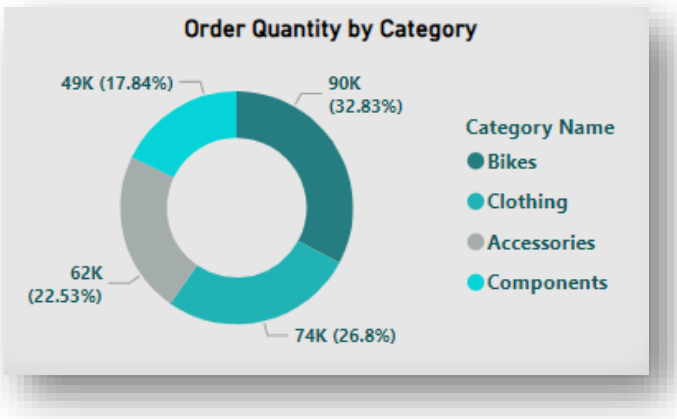


- Sales performance exhibited a significant decline in quarter 2 of 2011, followed by fluctuating sales and a decline in quarter 4 of 2012. However, sales in 2013 demonstrated consistent growth with minimal declines across the fourth quarter. In 2014, there were instances of declining sales observed in the second quarter.

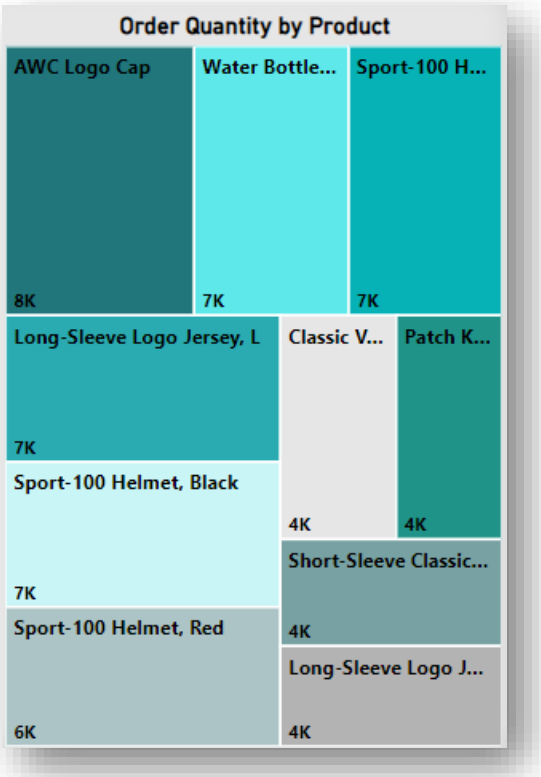


- Components emerged as the product category with the lowest number of orders, and the lowest sales. Conversely, bikes recorded the highest sales, indicating a significant contribution to sales. Notably, the order quantity for bikes is 90268, suggesting that customers tend to spend this amount on products within the bike category.

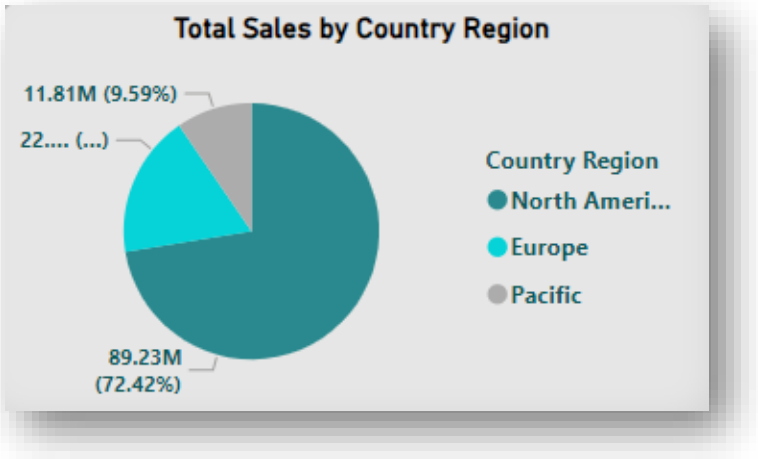




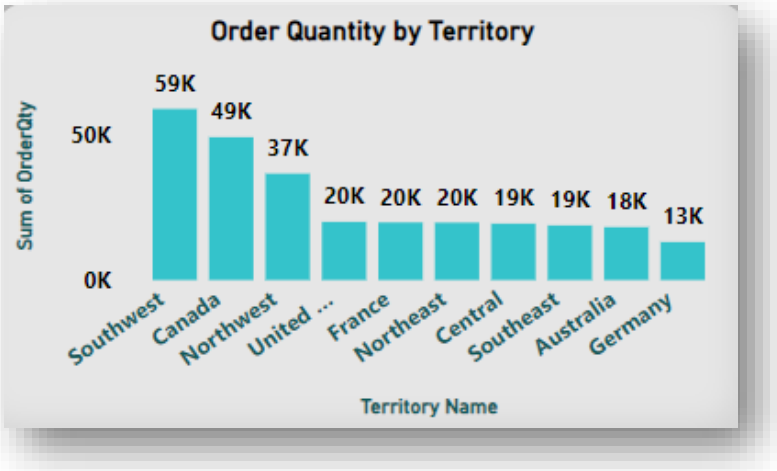
- The product "AWC Logo Cap" achieved the highest number of units sold, totaling 8311. Following closely behind is the "Water Bottle", totaling 6815, and the "Sport 100 Helmet" in the blue color, totaling 6743.



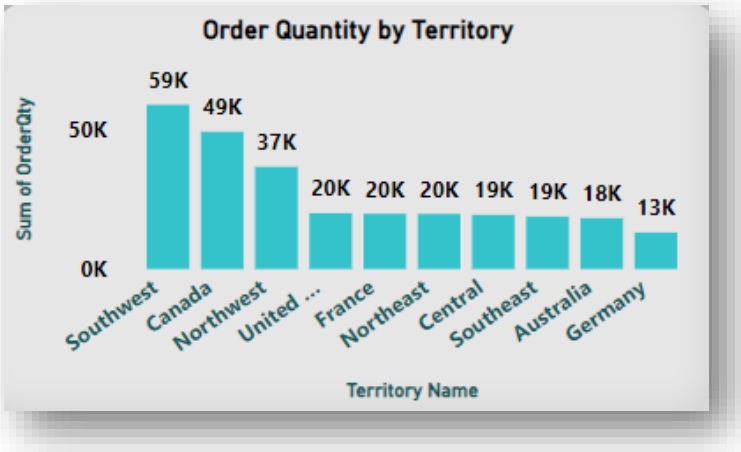
- Among the regions, North America recorded the highest sales, amounting to a total of 89,228,792.39. Europe followed with sales of 22,173,617.63, while the Pacific region generated a total revenue of 11,814,376.10.

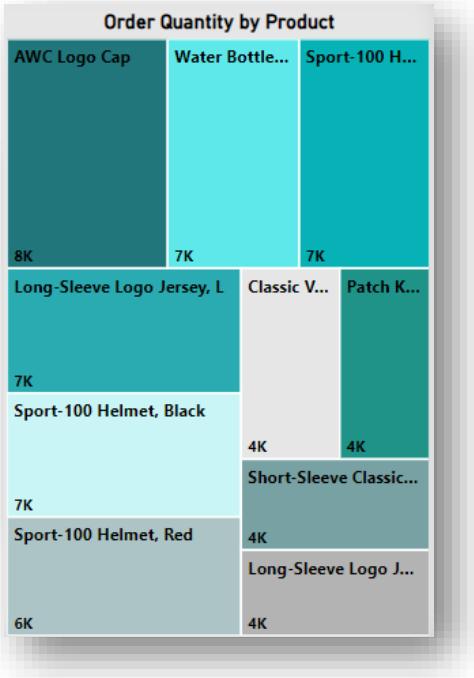


- The analysis showed that the southwest region of the US merged as the leader in sales, surpassing other regions, despite not having the highest number of orders. The region's impressive sales revenue can be attributed to its large customer base, which has likely contributed to the overall success. In contrast, Australia had the highest number of orders but struggled in terms of sales. This can be attributed to the relatively low order quantity within the region. Despite a high volume of orders, the quantity per order was not sufficient to drive substantial sales.

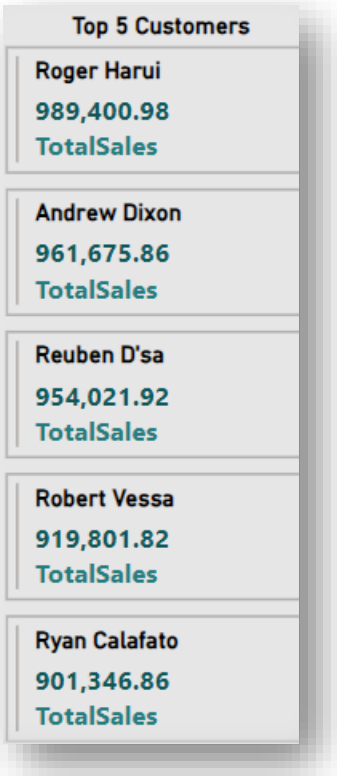


- The most popular products across the regions varied, the AWC Logo caps ranked as the most ordered item among customers in seven regions. However, in specific regions such as the northwest, Australia, and Germany, the water bottle gained prominence as the most popular and frequently ordered product.





- Among the customers, Roger Harui, Andrew Dixon, Reuben D'sa, Robert Vessa, and Ryan Calafato displayed a strong purchasing power, having spent 4,726,247 across 61 orders (9287 order quantity).



8. Interactive Elements

Filters and Controls:

- **Date Filters:** Allows users to select specific time periods for analysis.

User Actions:

- **Drill-down:** Click on summary metrics to explore detailed views.
- **Export Data:** Option to export filtered data or reports in various formats (e.g., Excel, PDF).

9. Sales Insights by SQL

➢ **Key Metrics:**

1. **Number of Orders:**

   Total distinct sales orders

```sql
--Number of Orders
select count(distinct salesorderid) number_of_orders
from Sales.SalesOrderDetail
```

number_of_orders
31465

2. **Number of Products:**

   Total distinct products available

```sql
--Number of Products
select count(distinct Name) number_of_products
from Production.Product
```

number_of_products
504

3. **Total Freight Costs:**

   Overall freight costs incurred

```sql
--Total Freight
select sum(Freight) total_freight
from sales.SalesOrderHeader
```

total_freight
3183430.2518

4. **Total Tax Amount:**

   Tax amount collected

```sql
--Total Tax amount
select sum (TaxAmt) total_tax
from Sales.SalesOrderHeader
```

total_tax
10186974.4602

5. **Number of Customers:**

Total distinct customers

```
--Number of Customers
select count(distinct CustomerID) number_of_customers
from Sales.Customer
```

number_of_customers
19820

6. **Number of Stores:**

Total distinct stores operating

```
--Number of stores
select count(distinct StoreID) number_of_stores
from Sales.Customer
```

number_of_stores
701

- **Sales Performance:**

1. **Total Sales by Year**:

The total sales revenue was summarized by year, showing the trend of sales over time.

```
--Total sales by year
select sum(TotalDue) total_sales, year(OrderDate) year
from sales.SalesOrderHeader h
group by year(OrderDate)
order by sum(TotalDue) desc
```

total_sales	year
48965887.9632	2013
37675700.312	2012
22419498.3157	2014
14155699.525	2011

2. **Total Sales by Territory**:

Sales performance across different territories:

```
--Total Sales by territory
select t.Name territortory_name, sum(h.totaldue) total_sales
from Sales.SalesTerritory t join Sales.SalesOrderHeader h
on t.TerritoryID = h.TerritoryID
group by t.Name
order by sum(h.totaldue) desc
```

territortory_name	total_sales
Southwest	27150594.5893
Canada	18398929.188
Northwest	18061660.371
Australia	11814376.0952
Central	8913299.2473
Southeast	8884099.3669
United Kingdom	8574048.7082
France	8119749.346
Northeast	7820209.6285
Germany	5479819.5755

3. **Number of Order Quantity by Year**:

Analysis of order quantities year-over-year indicates trends in customer purchasing behavior.

```
--Total Sales by territory
select t.Name territortory_name, sum(h.totaldue) total_sales
from Sales.SalesTerritory t join Sales.SalesOrderHeader h
on t.TerritoryID = h.TerritoryID
group by t.Name
order by sum(h.totaldue) desc
```

territortory_name	total_sales
Southwest	27150594.5893
Canada	18398929.188
Northwest	18061660.371
Australia	11814376.0952
Central	8913299.2473
Southeast	8884099.3669
United Kingdom	8574048.7082
France	8119749.346
Northeast	7820209.6285
Germany	5479819.5755

4. **Total Number of Orders by Year:**

The results help identify trends in customer purchasing behavior and seasonal patterns.

```
--Number of Order by Year
select count(salesorderid) number_of_orders, year(orderdate) year
from sales.SalesOrderHeader
group by year(orderdate)
```

number_of_orders	year
14182	2013
11761	2014
1607	2011
3915	2012

➢ Product Insights

1. **Total Products in Each Subcategory**:

Breakdown of total products available by subcategory.

```
--Find the total number of products in each subcategory:
select count (*)total_number_of_products,s.Name subcategory_name
from Production.Product p
join Production.ProductSubcategory s
on s.ProductSubcategoryID=p.ProductSubcategoryID
group by s.Name
```

total_number_of_products	subcategory_name
32	Mountain Bikes
43	Road Bikes
22	Touring Bikes
8	Handlebars
3	Bottom Brackets
2	Brakes
1	Chains
3	Cranksets
2	Derailleurs
3	Forks
3	Headsets
28	Mountain Frames

total_number_of_products	subcategory_name
7	Pedals
33	Road Frames
9	Saddles
18	Touring Frames
14	Wheels
3	Bib-Shorts
1	Caps
6	Gloves
8	Jerseys
7	Shorts
4	Socks
3	Tights
3	Vests

total_number_of_products	subcategory_name
3	Vests
1	Bike Racks
1	Bike Stands
3	Bottles and Cages
1	Cleaners
1	Fenders
3	Helmets
1	Hydration Packs
3	Lights
1	Locks
1	Panniers
2	Pumps
11	Tires and Tubes

2. **Total Orders in Each Category by Territory**:

   Analysis of orders categorized by territory and product category.

```
--Territory Name , category name , Total orders in each category
select t.Name as TerritoryName ,c.Name as categoryname ,count(d.OrderQty) as Totalorders
from Sales.SalesTerritory t
join Sales.SalesOrderHeader h
on t.TerritoryID=h.TerritoryID
join sales.SalesOrderDetail d
on h.SalesOrderID=d.SalesOrderID
join Production.Product p
on p.ProductID=d.ProductID
join Production.ProductSubcategory s
on s.ProductSubcategoryID= p.ProductSubcategoryID
join Production.ProductCategory c
on s.ProductCategoryID=c.ProductCategoryID
group by t.Name ,c.Name
```

TerritoryName	categoryname	Totalorders	TerritoryName	categoryname	Totalorders	TerritoryName	categoryname	Totalorders
Central	Clothing	1148	Canada	Accessories	6392	United Kingd...	Clothing	1750
Southwest	Bikes	8605	Northwest	Clothing	2873	Germany	Bikes	2255
Southwest	Clothing	4455	Germany	Components	548	France	Components	1225
Northwest	Bikes	5187	United Kingd...	Components	1099	United Kingd...	Accessories	4392
Australia	Bikes	5241	Southeast	Accessories	498	Canada	Clothing	3872
Northeast	Components	1806	Australia	Clothing	2152	Central	Components	1615
Southwest	Accessories	8386	Northwest	Accessories	6176	Canada	Bikes	5366
Southeast	Bikes	2560	France	Bikes	2618	Germany	Accessories	3519
Central	Bikes	2569	Canada	Components	3434	United Kingd...	Bikes	3185
France	Clothing	1577	Southwest	Components	4198	Southeast	Components	1657
Central	Accessories	500	Germany	Clothing	1206	Australia	Accessories	7178
France	Accessories	3668	Northeast	Bikes	2445	Northeast	Accessories	485
Southeast	Clothing	1261	Northeast	Clothing	1100	Northwest	Components	2629

40	Australia	Components	487

➢ Customer Demographics

1. **Number of Customers by Education Level**:

   Customer base segmented by education level

```
--number of customers by education
select count(distinct CustomerID)number_of_customers, Education
from Sales.Customer c join Person.Person p
on c.PersonID = p.BusinessEntityID
join Sales.vPersonDemographics vdemo
on vdemo.BusinessEntityID = p.BusinessEntityID
where Education is not null
group by Education
```

	number_of_customers	Education
1	3294	High School
2	1581	Partial High School
3	5064	Partial College
4	3189	Graduate Degree
5	5356	Bachelors

➢ **Yearly Comparison Analysis**

1. **Sales Decline in 2014**:

Investigation into the reasons behind lower sales in 2014 compared to 2013 revealed a drop in order quantities and unit prices.

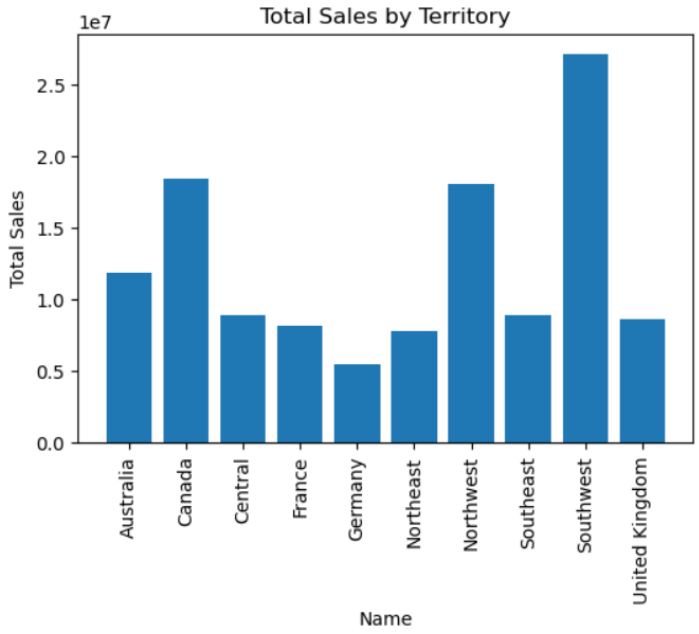
```
--18--Why is 2014 Sales lower than 2013

select count(sod.orderqty) order_quantity,sum(sod.UnitPrice) unit_price, year(soh.ModifiedDate) year
from sales.SalesOrderDetail sod join sales.SalesOrderHeader soh
on sod.SalesOrderID=soh.SalesOrderID
group by year(soh.ModifiedDate)
order by sum(sod.UnitPrice) desc
```

	order_quantity	unit_price	year
1	55100	21017032.8096	2013
2	20016	14548943.8875	2012
3	40526	13776951.6371	2014
4	5675	7080819.2805	2011

10. Sales Analysis by Python

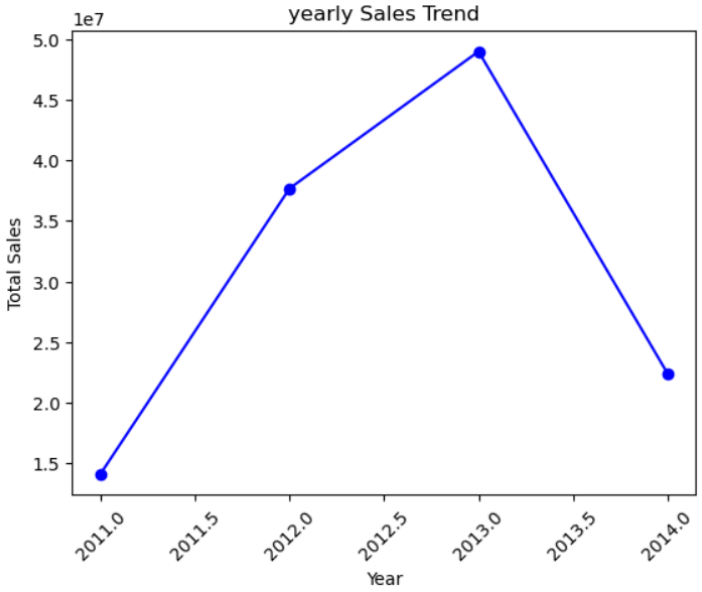
1. **Total Sales by Territory**:
    - Aggregated total sales using groupby on Name, producing a bar chart to visualize total sales across different territories.

```
#Total Sales by Territory
total_sales_by_territory = dt.groupby('Name')['TotalDue'].sum().reset_index()
plt.figure(figsize=(6,6))
plt.bar(total_sales_by_territory['Name'], total_sales_by_territory['TotalDue'])
plt.title('Total Sales by Territory')
plt.xlabel('Name')
plt.ylabel('Total Sales')
plt.xticks(rotation=90)
```



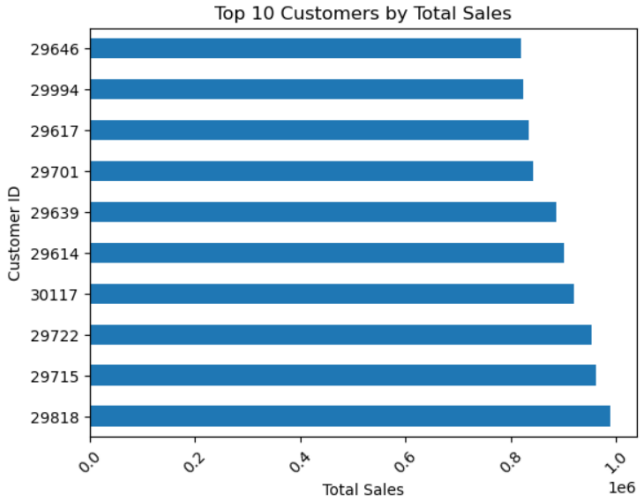
2. **Sales Trend Over Time**:
    - Analyzed yearly sales trends by grouping data by year and plotting total sales, indicating growth or decline trends over the years.

```
#Sales Trend Over Time
sales_trend =dt.groupby('year')['TotalDue'].sum().reset_index()
plt.plot(sales_trend['year'], sales_trend['TotalDue'], marker='o', color='b')
plt.title('yearly Sales Trend')
plt.xlabel('Year')
plt.ylabel('Total Sales')
plt.xticks(rotation=45)
```



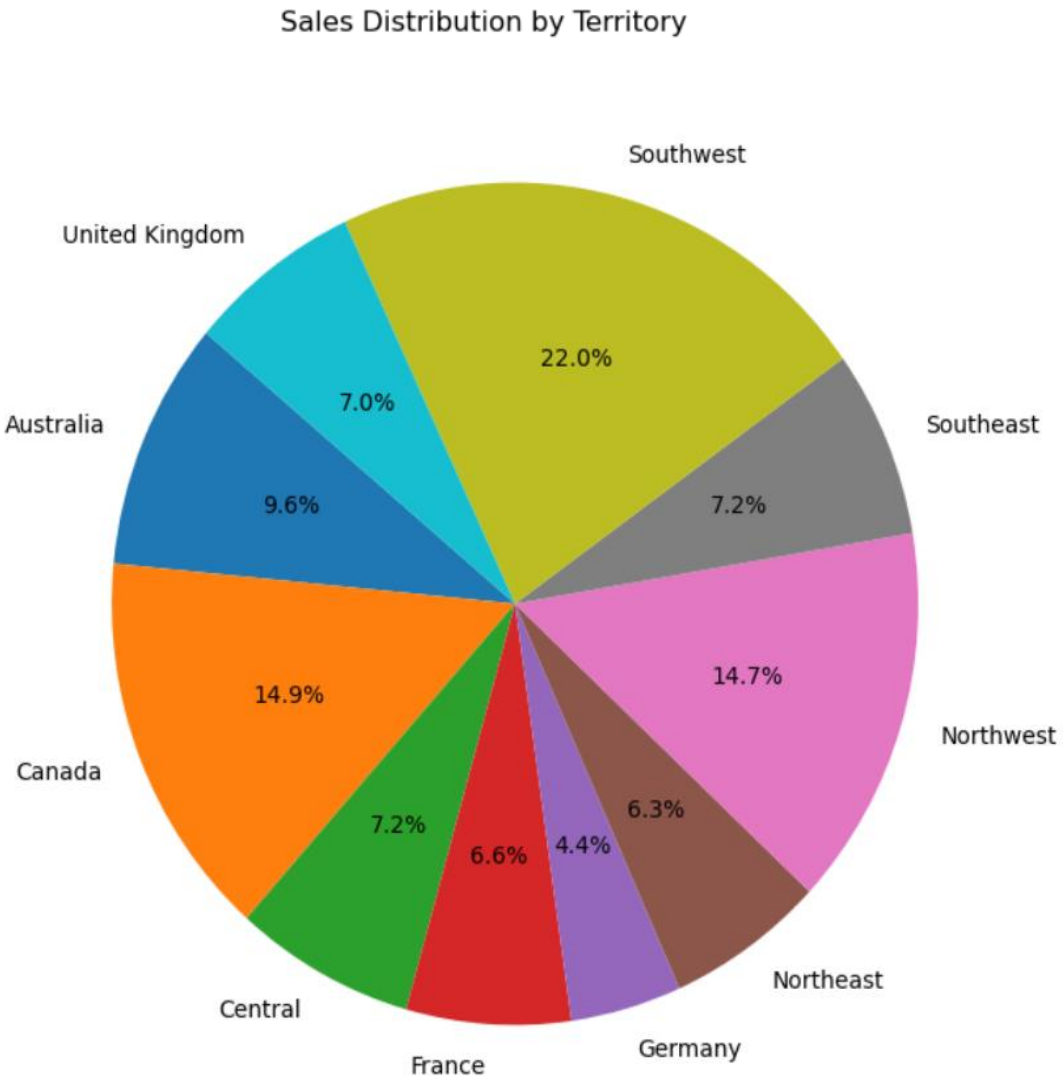
3. **Top Customers**:
    - Identified the top 10 customers by total sales, visualized through a horizontal bar chart.

```
#top 10 customers by customerID
top_customers=dt.groupby('CustomerID')['TotalDue'].sum().nlargest(10)
top_customers.plot(kind='barh')
plt.title('Top 10 Customers by Total Sales')
plt.xlabel('Total Sales')
plt.ylabel('Customer ID')
plt.xticks(rotation=45)
```



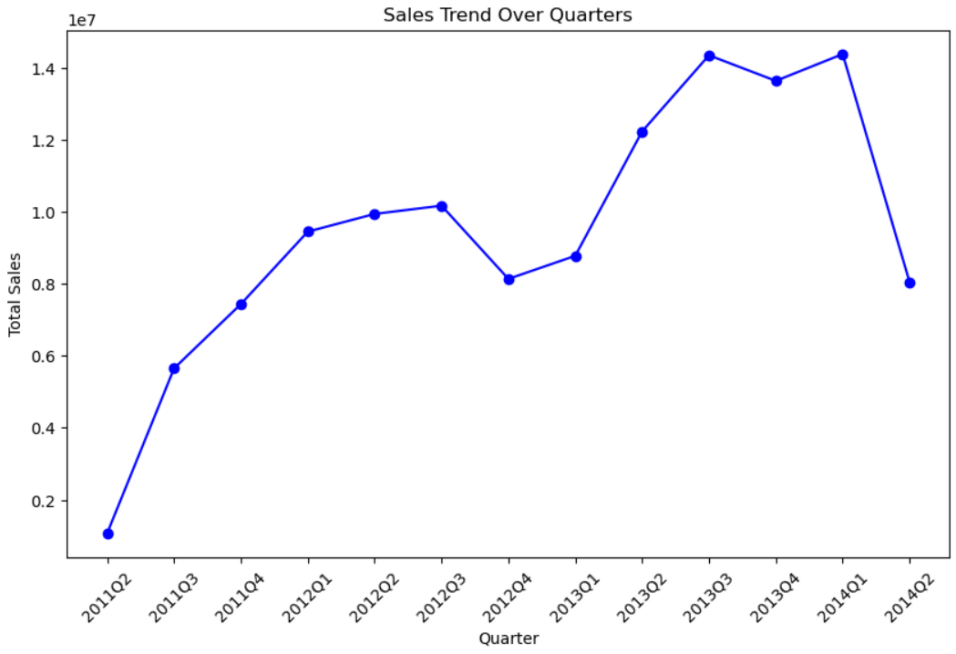
4. **Sales Distribution**:
    - Visualized sales distribution by territory using a pie chart, showing the proportion of total sales per territory.

```
#Sales Distribution
sales_distribution = dt.groupby('Name')['TotalDue'].sum()
plt.figure(figsize=(8, 8))
plt.pie(sales_distribution, labels=sales_distribution.index, autopct='%1.1f%%', startangle=140)
plt.title('Sales Distribution by Territory')
```



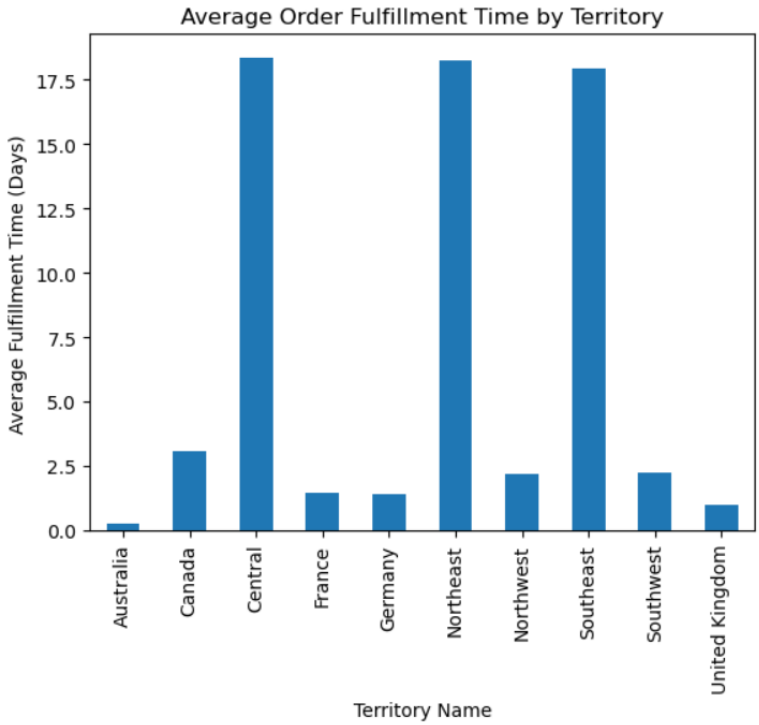
5. **Order Fulfillment Analysis**:
    - Analyzed average fulfillment time per territory, displayed in a bar chart, highlighting efficiency across different regions.

```
#Sales Trend Over quarter
sales_trend_quarterly =dt.groupby('Quarter')['TotalDue'].sum()
plt.figure(figsize=(10, 6))
plt.plot(sales_trend_quarterly.index.astype(str), sales_trend_quarterly, marker='o', linestyle='-', color='b')
plt.title('Sales Trend Over Quarters')
plt.xlabel('Quarter')
plt.ylabel('Total Sales')
plt.xticks(rotation=45)
```



6. **Quarterly Sales Trend**:
    - Sales were further analyzed on a quarterly basis, showing trends through line plots for a more granular view of sales performance.

```
#Order Fulfillment Analysis
fulfillment_analysis =dt.groupby('Name')['Fulfillment Time'].mean()
fulfillment_analysis.plot(kind='bar')
plt.title('Average Order Fulfillment Time by Territory')
plt.xlabel('Territory Name')
plt.ylabel('Average Fulfillment Time (Days)')
```



11. Technical Details

- SQL
- Power bi (Dax, Power query, Data modeling)
- Python

12. Appendices

Glossary

- COGS = order quantity * cost
- Growth profit = Total sales – Total COGS

13. References

Documentation for SQL Server and Adventure Works 2019 database.