CAPSTONE PROJECT – RETAIL ANALYTICS

Authored by: SHREEVIDYA CHANDRASEKAR

1.1 INTRODUCTION

What is retail analytics?

Retail analytics is the process through which retailers discover, gather, and make sense of the data generated throughout their departments.

What emerges from these processes is a **set of actionable insights** based on business trends, emerging patterns, and performance indicators. Using these insights, you can **improve vital aspects of your retail business**, such as customer experience, logistics, management, and even sales.

The Retail Database provides comprehensive insights into the retail business's performance. The dashboard will focus on sales, product, customer, and demographic analysis, aiming to facilitate data-driven decision-making, optimize sales strategies, and enhance customer experiences.

1.2 OBJECTIVE

The primary objective of this project is to leverage data and insights to enhance the overall performance and profitability of a retail business. This involves using analytical tools and techniques to gather, process, and interpret data related to various aspects of retail operations and customer behavior.

THE SPECIFIC OBJECTIVES OF RETAIL ANALYTICS

- Understanding Customer Behavior.
- Optimizing Inventory Management.
- Enhancing Marketing Strategy.
- Price Optimization.

1.3 THE PROCESS

a) **Data Acquisition from GitHub:**
Obtain the requisite dataset from a designated GitHub repository, containing essential information on university rankings, encompassing various countries and their performance across distinct ranking systems.

b) **Data Transformation and Enhancement:**
If necessary, execute data transformation procedures to ensure data quality and consistency. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential.

c) **Connecting with Tools:**
Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, Excel, and MySQL Workbench, facilitating seamless data integration and processing.

d) **Problem Statement Solution in Power BI:**
Utilize Power BI to delve into the specified problem statements. Employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions.

e) **Exploratory Data Analysis (EDA):**
Perform exploratory data analysis using either Excel or SQL Workbench, depending on the complexity of the analysis. Extract meaningful patterns, relationships, and trends from the data to inform subsequent decision-making.

f) **Creation of Visual and Insightful PowerPoint:**
Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights.

g) **Detailed Documentation:**
Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions, EDA insights, and PowerPoint visualizations.

1.4 ANALYSIS OVERVIEW

The Project Includes the following Analysis

- ❖ Analyzing the sales, customers, products gives an understanding about the trends in sales pattern, order volume and demographics.
- ❖ Based on the Analysis valuable insights can be drawn, which Will help in giving recommendations to the stakeholders.

Key Metrics of this project are

- ✓ The quality of the analysis
- ✓ The relevance of the insights
- ✓ The impact of the recommendations

This project is significant because it has the potential to improve the quality and competitiveness of retail markets. By understanding these factors, the offices can be established across various regions and countries globally.

1.5 SIGNIFICANCE OF RETAIL ANALYTICS

1) Data-Driven Decision Making: Retail analytics provides actionable insights based on data, enabling retailers to make informed decisions rather than relying on gut feelings or assumptions.
2) Personalization: By understanding customer preferences, retailers can offer personalized shopping experiences, leading to increased customer loyalty and higher conversion rates.
3) Inventory Optimization: Retail analytics helps retailers maintain the right level of inventory, reducing carrying costs and minimizing stockouts.
4) Operational Efficiency: Retailers can identify operational bottlenecks and inefficiencies, leading to streamlined processes and improved resource allocation.
5) Competitive Advantage: Retail analytics enables retailers to stay ahead of competitors by identifying emerging trends, adapting quickly to changes in the market, and capitalizing on new opportunities.
6) Enhanced Customer Experience: By understanding customer behaviour, retailers can create better shopping experiences, from personalized recommendations to smoother checkout processes.
7) ROI Measurement: Retail analytics helps assess the effectiveness of various strategies and investments, allowing retailers to allocate resources to initiatives with the highest returns.
8) Predictive Analytics: Retailers can use historical data to predict future trends, demand patterns, and customer behaviour, aiding in proactive decision-making.
9) In summary, retail analytics plays a crucial role in modern retail operations by providing insights that drive strategic decisions, improve customer experiences, and optimize various aspects of the business. It empowers retailers to navigate the dynamic and competitive retail landscape with data-driven precision.

1.6 DATA DICTONARY:

TABLE: Customers - This table contains information about the retail company's customers

FIELDS: -

- Customer number: Unique identifier for each customer
- Customer Name : The company or shop name of the customer is provided
- Contact First Name and last name: Customer names are given
- Phone-Number: phone number of the customers is provided
- City: Name of the city
- State: Name of the state
- Postal Code: postal code corresponding to the country
- Country: Name of the country
- Sales Representative number: ID number of the Employee
- Credit Limit: Customer credit limit

TABLE: Employees - This table holds details about the company's employees

FIELDS: -

- Employee number: Unique identifier for each employee
- Employee First Name and last name: Employee names are given
- Extension: Employee extension code
- Email Address: Employee Email-ID
- Office code: Office where the employee works
- Reports to: the employee number of the person to whom the employee reports
- Job title: Designation of the employee

TABLE: Offices - This table stores information about the offices of the retail company

FIELDS: -

- Office code: Unique identifier for each office
- City: Place where offices are located
- Phone Number: contact of each Office
- Address: Address of the Office
- State: The state corresponding to each City where offices are located
- Country: Offices are located in various Countries
- Postal Code: Code corresponding to city, state and Country
- Territory: Offices spread across regions

TABLE: Orders -This table stores data related to customer orders

FIELDS: -

- Order number: Unique identifier for each order
- Order Date: The date in which the order was placed
- Required Date: The date in which the order is required
- Shipped Date: shipment date of the order
- Order Status: progression of the order
- Comments: Specifications of customers for different products
- Customer Number: This is to establish a link between customers and orders

TABLE: Order Details - This table contains information about the individual items included in each order

FIELDS: -

- Order number: This is to establish a link between orders and order details
- Product Code: This is linking to the Products table
- Quantity Ordered: The number of quantities ordered
- Price Each: Price per Item
- Order Line number: The Order in which the orders were placed

TABLE: Products - This table stores details about various products sold by the retail company

FIELDS: -

- Product Code: Unique identifier for each Product
- Product Name: Name of each Product
- Product Line: Categorization of the Products
- Product Scale: Size of the Product with respect to size of the Original Product
- Product Vendor: Manufacturer of each product
- Product Description: Description of each product
- Quantity in stock: The number of products that are in stock
- Buy Price:
- MSRP: Manufacturer's Suggested Retail Price

TABLE: Product Lines - This table is used to describe the different product lines available in the company's inventory

FIELDS: -

- ProductLineName: Name for each Product Category
- Text Description: Description for each Product Category
- HTML Description: For web-based content
- Image: To represent the Product Visually

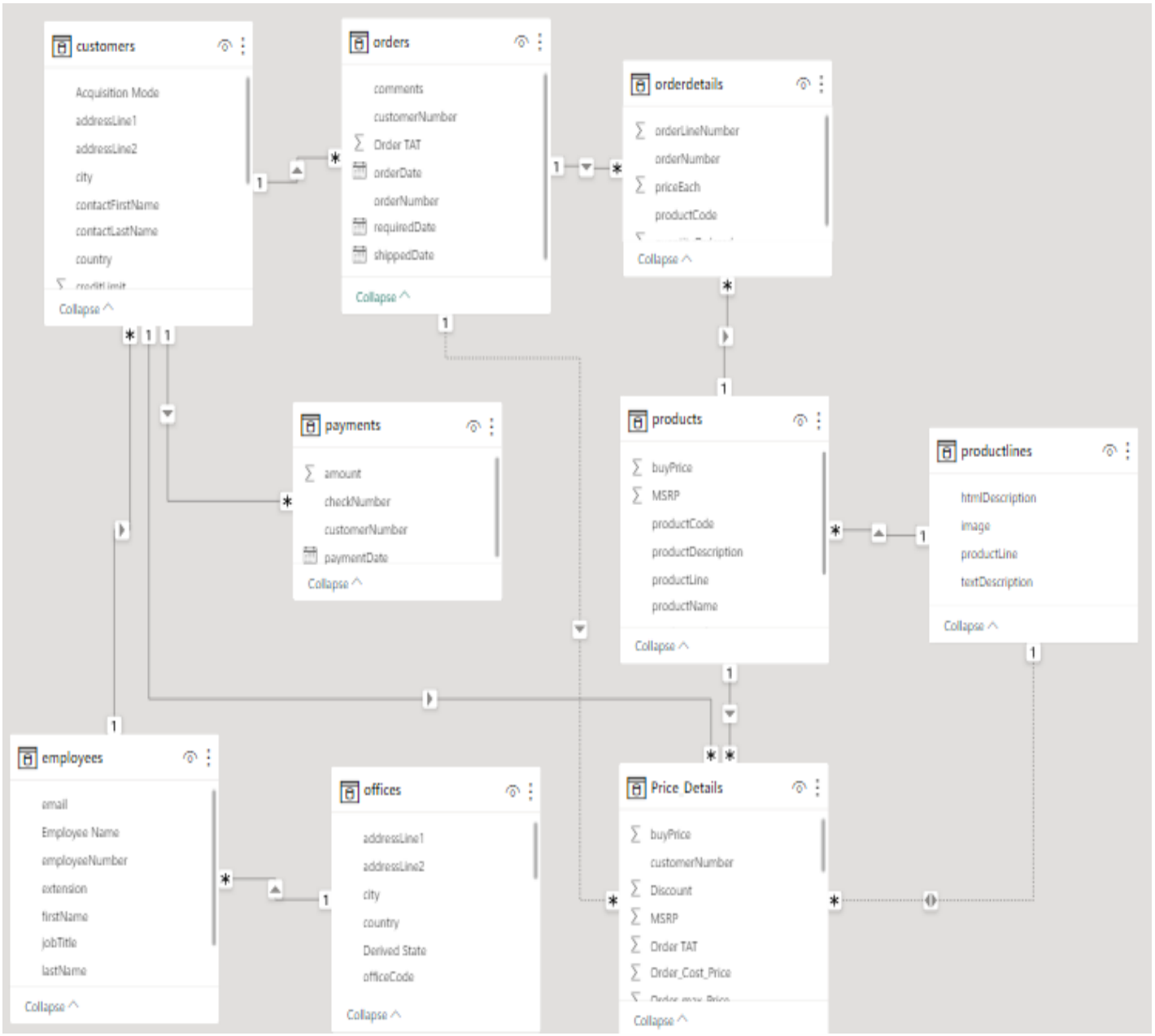
TABLE: Payments - This table stores details about payments made by customers

FIELDS: -

- Customer Number: For Linking Customers and Payments table
- Check Number: Check number for each payment
- Payment Date: The date in which the Payment has been made
- Payment Amount: The amount paid for the Order

The data dictionary provides a comprehensive overview of the tables and their respective fields in the dataset. It outlines the relationships between the tables, allowing for a better understanding of the data structure and facilitating the design and implementation of the Power BI Dashboard

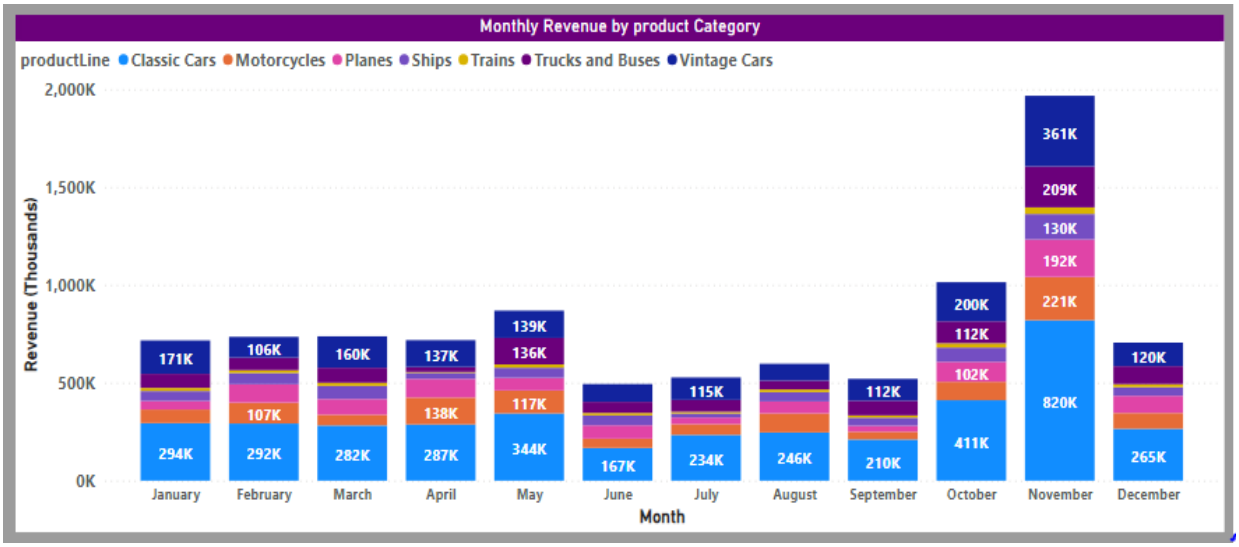
1.7 ER DIAGRAM



1.8 POWER BI PROBLEM STATEMENTS

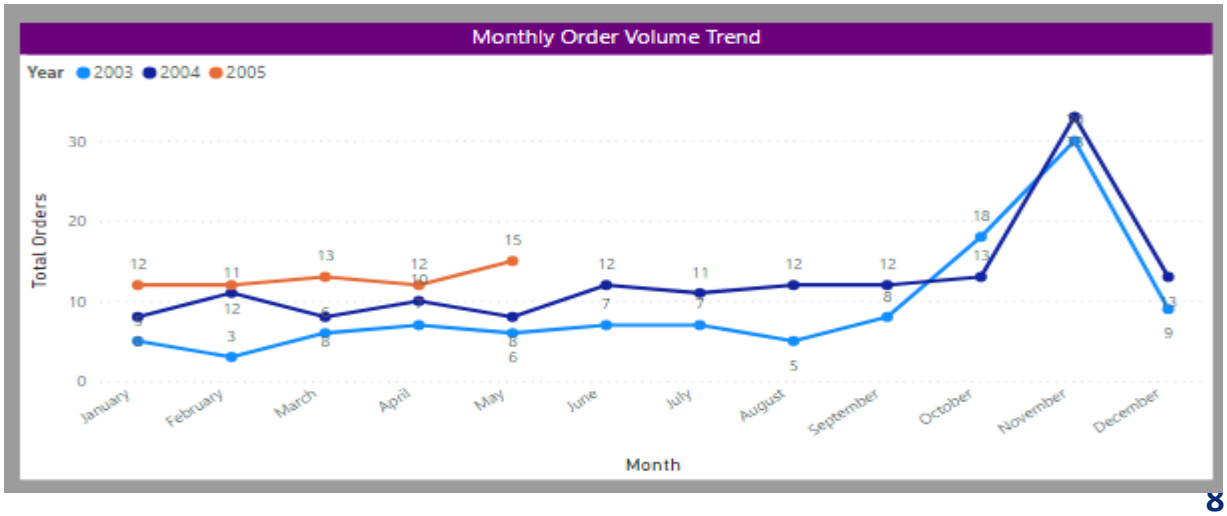
1) How does monthly revenue vary across different product categories?

- ❖ Analyzing the monthly revenue for different product categories it is been identified that classic cars and vintage cars are the two highly sold product lines compared to other product lines, that too especially in the month of May (Holiday season), October and November (start of festive season).
- ❖ The inference is, Monthly revenue is high in the holiday and Festive season.



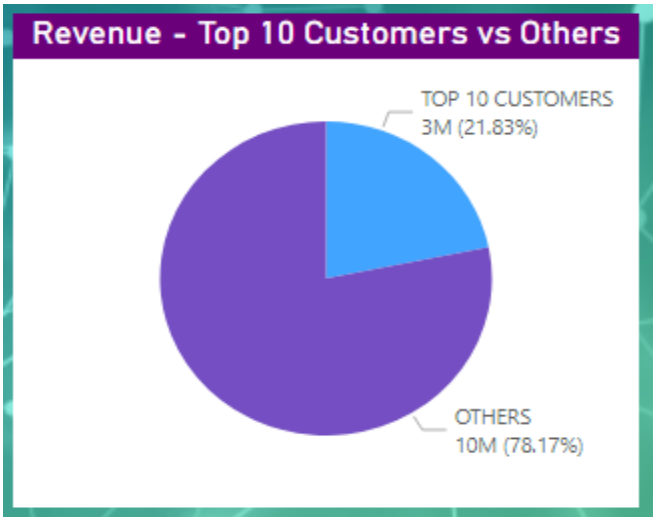
2) What is the trend in customer order volume over the past year?

- ❖ Analyzing the customer order volume, it is seen that the customers tend to order more in the month of October, November as it is the festive season.
- ❖ Orders are in peak during the festive season over the past years, it can be utilized to do more sales and increase the revenue further.



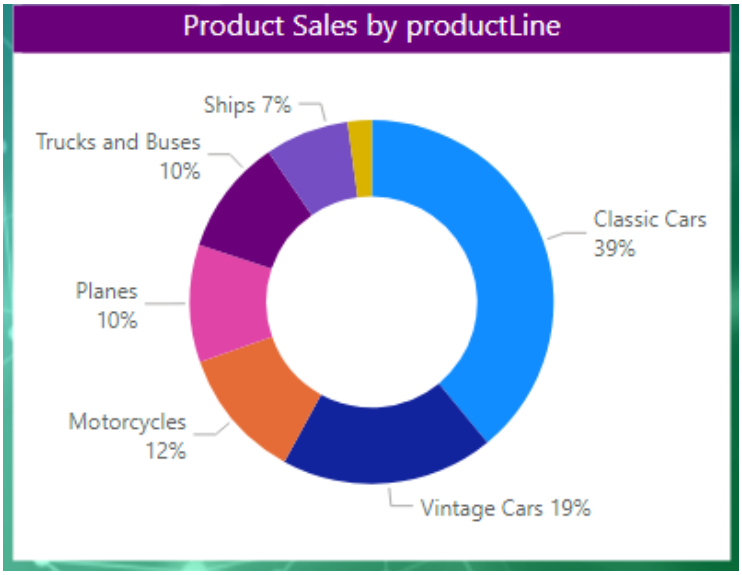
3) How does the sales performance of top customers compare to the rest?

- The analysis is done based on the revenue contribution of top customers. Here the analysis w.r.t top 10 customers.
- The contribution of top 10 customers is 20% when compared with others as shown in the visualization.



4) What is the distribution of product sales across different product lines?

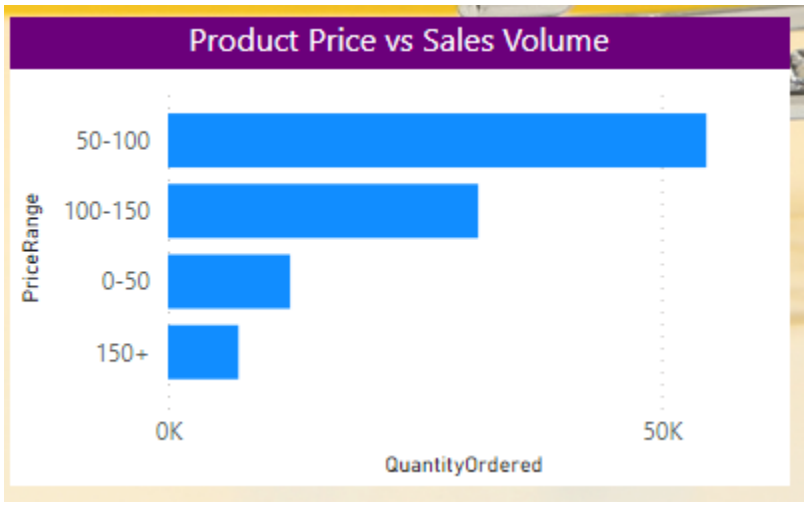
- The product sales for the product category "Classic cars and Vintage Cars" is high compared to other product categories.
- On this Analysis, we can infer that other product lines must be focused on cross selling, promoting discount, stock clearance sale in order to make revenue out of those least sold product categories.



5) How does the profitability of different products compare based on their quantity in stock?

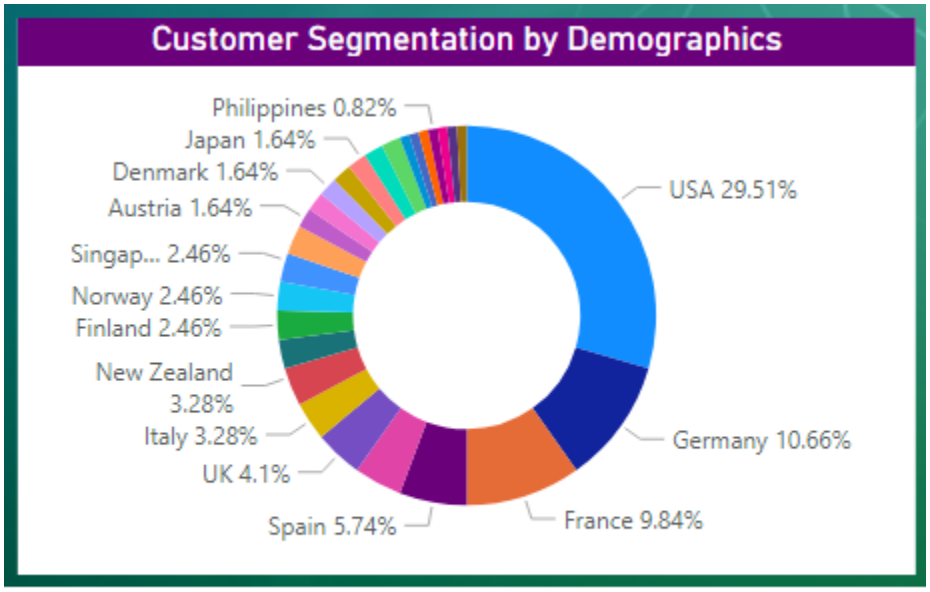
6) How does Product pricing impact sales volume?

❖ Price and sales volume is inversely proportional to each other



7) What is the distribution of customers across different demographic segments?

❖ Maximum number of customers are from USA followed by Germany and France.

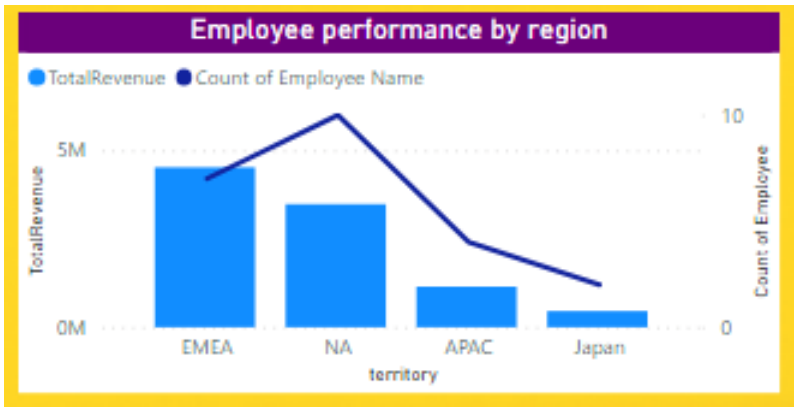


8) What are the top regions in terms of sales revenue?

❖ NORTH AMERICA and EUROPE are the top regions in terms of sales revenue inferred from the below visualization.

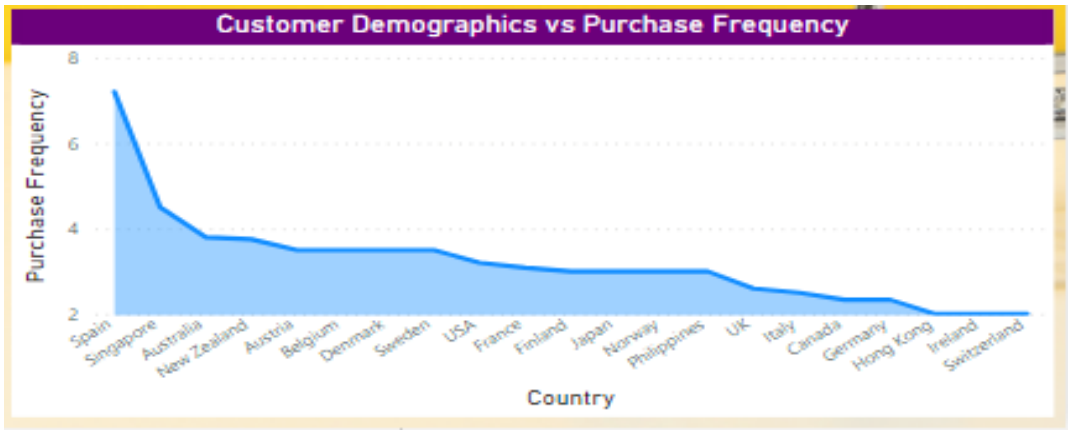
9) How does the performance of sales employees vary across different regions?

- It is observed that the employee's performance in EMEA and APAC is superior compared to other regions.



10) What is the correlation between customer demographics and purchase frequency?

- It is inferred that the customer purchase frequency is high in SPAIN and SINGAPORE followed by rest of the countries.
- In SWITZERLAND and IRELAND the purchase frequency is observed to be low.



11) How does customer lifetime value vary for different customer acquisition channel?

If Data provided to this analysis, it would help in more deep diving of the insights and recommendation.

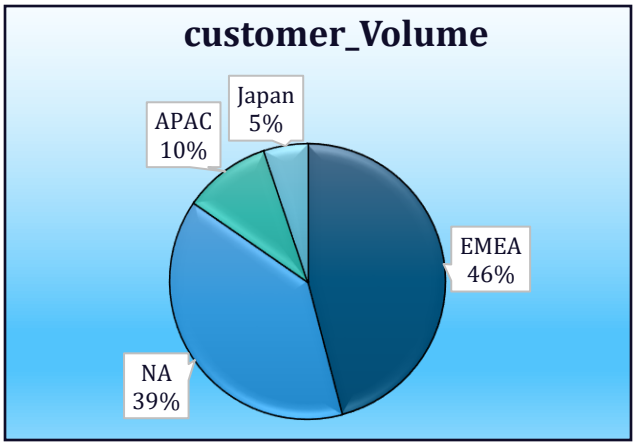
12) What is the correlation between customer age and purchase frequency?

If Data provided to this analysis, it would help in more deep diving of the insights and recommendation.

1.9 EXPLORATORY DATA ANALYSIS

1) Which factors contribute to highest sales in a particular region?

Territory	customer_Volume	order_volume	product_volume	Revenue	Employee_Volume	Office_count
EMEA	45	153	109	4520712.28	6	2
NA	38	119	109	3479191.91	6	3
APAC	10	38	109	1147176.35	2	1
Japan	5	16	81	457110.07	1	1



INSIGHTS:- Following are the factors contributing to the highest sales in regions

- Customer Volume
- Order Volume
- No of Offices & Employees in Region

2) How can customer purchasing patterns be influenced to increase average order value?

SQL Query 1 with Results

```
1  select ordernumber,count(os.productcode) as 'no_of_products_per_order',
2  sum(TotalOrderValue) as 'Total_Revenue'
3   from Order_summary os
4  join customers c on c.customernumber = os.customernumber
5  group by ordernumber order by Total_Revenue desc limit 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

orderNumber	no_of_products_per_order	Total_Revenue
10165	18	67392.85
10287	17	61402.00
10310	17	61234.67
10212	16	59830.55
10207	16	59265.14
10127	15	58841.35
10204	17	58793.53
10126	17	57131.92
10222	18	56822.65
10142	16	56052.56

SQL Query 2 with Results

```
1  select c.customernumber,count(os.ordernumber),c.creditLimit as 'CreditLimit',
2   sum(TotalOrderValue) as 'Total_Revenue',
3  round(sum(TotalOrderValue)/c.creditLimit*100,2) as 'Credit Utilization Rate'
4   from Order_summary os join customers c on c.customernumber = os.customernumber
5   group by customernumber order by Total_Revenue desc limit 10;
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

customernumber	count(os.ordernumber)	CreditLimit	Total_Revenue	Credit Utilization Rate
141	259	227600.00	820689.54	360.58
124	180	210500.00	591827.34	281.15
114	55	117300.00	180585.07	153.95
151	48	138500.00	177913.95	128.46
119	53	118200.00	158573.12	134.16
148	43	103800.00	156251.03	150.53
323	46	88000.00	154622.08	175.71
131	49	114900.00	149085.15	129.75
187	51	136800.00	148410.09	108.49
450	40	77600.00	143536.27	184.97

INSIGHTS: - Purchasing patterns be influenced in following ways to increase average order value

- Increasing number of products per order
- Increasing the Credit Limit will increase the order volume and revenue

3) What are the key drivers of sales growth, and how can they be leveraged for future success?

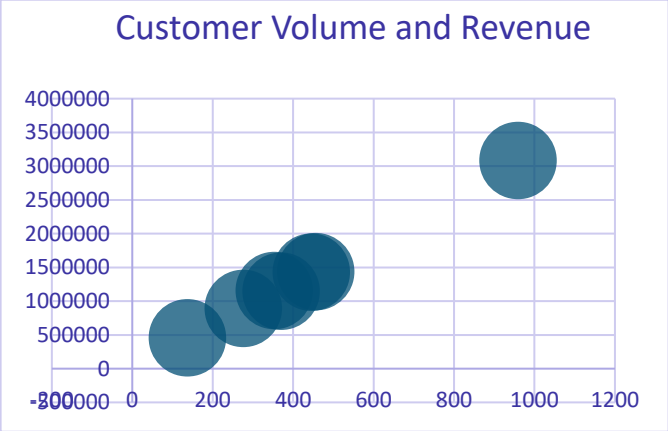
SQL Query 1 with Results

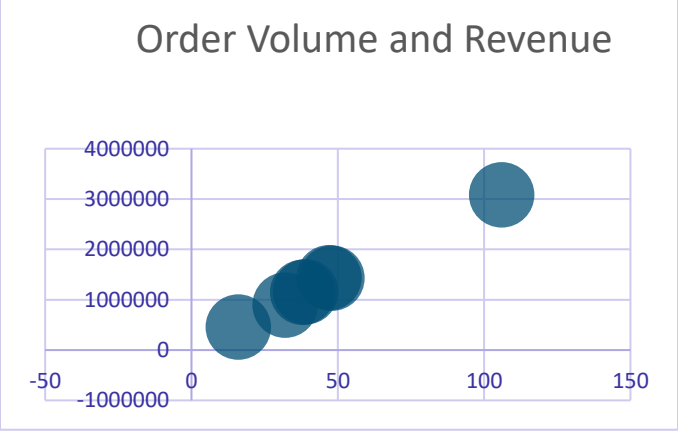
```
1  select O.territory 'Territory', O.city, count(os.customernumber) 'Customer Volume',
2  count(distinct Os.ordernumber) 'Order Volume',count(distinct os.salesRepEmployeeNumber) 'Employee Count',
3  round(sum(OS.TotalOrderValue)/count(distinct os.salesRepEmployeeNumber),2) 'Average Sales Per Emp',
4  sum(OS.TotalOrderValue) 'TotalOrderValue',count(distinct O.officeCode) 'Office Count',
5  round(avg(C.creditLimit),2) 'Average_CreditLimit',count(distinct Os.productcode) 'No of Products'
6   from Order_summary OS
7  join employees E on OS.salesRepEmployeeNumber=E.employeeNumber
8  join  offices O on O.officeCode=E.officeCode
9  right join customers C on C.customerNumber = OS.customernumber
10 group by O.territory,O.city having TotalOrderValue is not null order by TotalOrderValue desc;
```

tesult Grid | Filter Rows: | Export: | Wrap Cell Content:

Territory	city	Customer Volume	Order Volume	Employee Count	Average Sales Per Emp	TotalOrderValue	Office Count	Average_CreditLimit	No of Products
EMEA	Paris	959	106	4	770940.40	3083761.58	1	128387.70	109
EMEA	London	456	47	2	718475.35	1436950.70	1	94630.26	109
NA	San Francisco	445	48	2	714531.79	1429063.57	1	133540.90	109
NA	NYC	353	39	2	578794.86	1157589.72	1	88499.43	105
APAC	Sydney	370	38	2	573588.18	1147176.35	1	95277.84	109
NA	Boston	276	32	2	446269.31	892538.62	1	85480.43	106
Japan	Tokyo	137	16	1	457110.07	457110.07	1	88794.16	81

Territory	City	Customer Volume	Order Volume	No of Sales Rep Number	Average Revenue per Employee	Revenue	Average Credit Limit	No of Products contributed to Revenue
EMEA	Paris	959	106	4	770940.4	3083761.58	128387.7	109
EMEA	London	456	47	2	718475.35	1436950.7	94630.26	109
NA	San Francisco	445	48	2	714531.79	1429063.57	133540.9	109
NA	NYC	353	39	2	578794.86	1157589.72	88499.43	105
APAC	Sydney	370	38	2	573588.18	1147176.35	95277.84	109
NA	Boston	276	32	2	446269.31	892538.62	85480.43	106
Japan	Tokyo	137	16	1	457110.07	457110.07	88794.16	81





INSIGHTS:- Following are the Key Drivers of Sales Growth

- No of Offices or Employees
- No of Customers

4) Which product features or attributes are most appealing to customers?

SQL Query 1 with Results

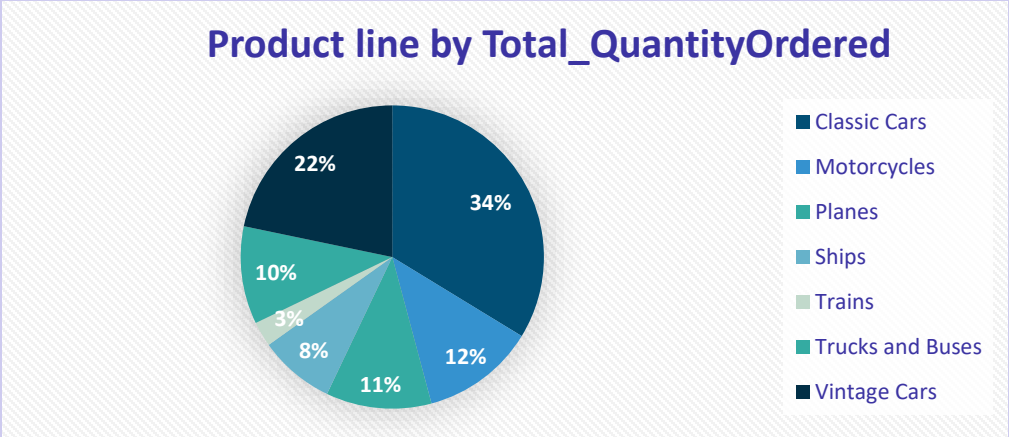
```
1  select OD.productCode 'Product Code',P.productName 'Product Name',P.productLine 'Product Line',
2  P.productVendor 'Product Vendor',P.productScale 'Product Scale',count(OD.orderNumber) as 'OrderVolume',
3  sum(OD.quantityOrdered) as 'Total_QuantityOrdered', P.MSRP 'MSRP', P.buyPrice 'Buy Price',
4  round((P.MSRP-P.buyPrice)/P.MSRP*100,2) as 'Discount%'
5  from orderdetails OD
6  inner join products P
7  on P.productCode = OD.productCode
8  group by od.productCode
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Product Code	Product Name	Product Line	Product Vendor	Product Scale	OrderVolume	Total_QuantityOrdered	MSRP	Buy Price	Discount%
S18_3232	1992 Ferrari 360 Spider red	Classic Cars	Unimax Art Galleries	1:18	53	1808	169.34	77.90	54.00
S18_1342	1937 Lincoln Berline	Vintage Cars	Motor City Art Classics	1:18	28	1111	102.74	60.62	41.00
S700_4002	American Airlines: MD-11S	Planes	Second Gear Diecast	1:700	28	1085	74.03	36.27	51.01
S18_3856	1941 Chevrolet Special Deluxe Cabriolet	Vintage Cars	Exoto Designs	1:18	28	1076	105.87	64.58	39.00
S50_1341	1930 Buick Marquette Phaeton	Vintage Cars	Studio M Art Models	1:50	28	1074	43.64	27.06	37.99
S18_4600	1940s Ford truck	Trucks and Buses	Motor City Art Classics	1:18	28	1061	121.08	84.76	30.00
S10_1678	1969 Harley Davidson Ultimate Chopper	Motorcycles	Min Lin Diecast	1:10	28	1057	95.70	48.81	49.00
S12_4473	1957 Chevy Pickup	Trucks and Buses	Exoto Designs	1:12	28	1056	118.50	55.70	53.00
S18_2319	1964 Mercedes Tour Bus	Trucks and Buses	Unimax Art Galleries	1:18	28	1053	122.73	74.86	39.00
S24_3856	1956 Porsche 356A Coupe	Classic Cars	Classic Metal Creations	1:18	27	1052	140.43	98.30	30.00
S24_3949	Corsair F4U (Bird Cage)	Planes	Second Gear Diecast	1:24	28	1051	68.24	29.34	57.00
S700_3167	F/A 18 Hornet 1/72	Planes	Motor City Art Classics	1:72	28	1047	80.00	54.40	32.00
S18_1662	1980s Black Hawk Helicopter	Planes	Red Start Diecast	1:18	28	1040	157.69	77.27	51.00

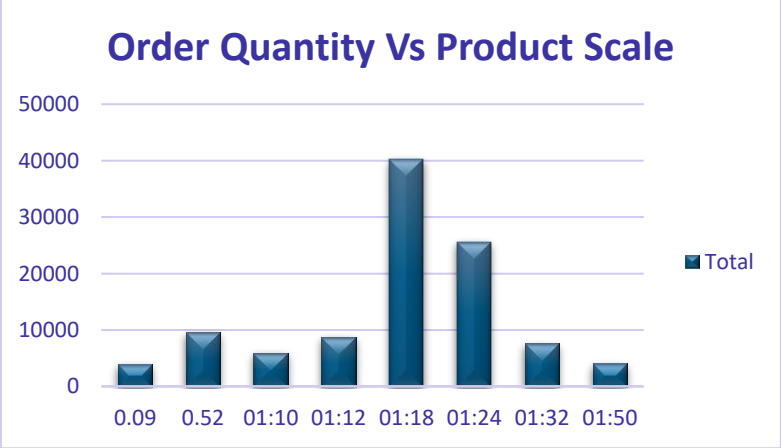
PIVOT 1

Row Labels	Sum of Total_ QuantityOrdered	Count of OrderVolume	Average of buyPrice	Average of MSRP	Average of Discount_perc
Classic Cars	33.72%	33.94%	64.65	118.30	45.22
Motorcycles	12.11%	11.93%	50.69	97.18	47.08
Planes	11.25%	11.01%	49.63	89.52	44.25
Ships	8.09%	8.26%	47.01	86.56	46.11
Trains	2.67%	2.75%	43.92	73.85	42.00
Trucks and Buses	10.43%	10.09%	56.33	103.18	46.45
Vintage Cars	21.73%	22.02%	46.07	87.10	46.62



PIVOT 2

Row Labels	Total_Quantity_Ordered
0.09	3882
0.52	9533
01:10	5964
01:12	8790
01:18	40243
01:24	25515
01:32	7578
01:50	4011



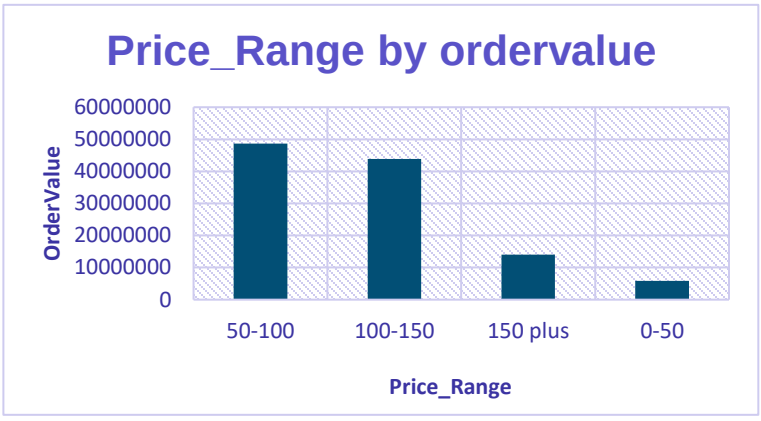
SQL Query 2 with Results

```
1  select
2  case
3  when priceEach between 0 and 50 then '0-50'
4  when priceEach between 50 and 100 then '50-100'
5  when priceEach between 100 and 150 then '100-150'
6  else '150 plus'
7  end as 'Price_Range',
8  sum(OD.quantityOrdered*OD.priceEach) as 'ordervalue'
9  from  order_summary OS
10 JOIN orderdetails OD
11 on OS.orderNumber = OD.OrderNumber
12 group by Price_Range
13 order by ordervalue desc
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Price_Range	ordervalue
50-100	48640594.08
100-150	43905799.62
150 plus	14005006.62
0-50	5903769.61

Price_Range	Ordervalue
50-100	48640594.08
100-150	43905799.62
150 plus	14005006.62
0-50	5903769.61



INSIGHTS:- Following are the probable product feature or attributes attractive to the customers

- Product Category Cars (Classic Cars and Vintage Cars) yield high revenue and order volume
- Product Scale of 1:18 and 1:24 yield high revenue and order volume compare to others
- Product with price ranging 50 to 100 is yield high revenue followed by 100-150 which clearly states affordable pricing attracts customer the most

5) How can the product mix be optimized to cater to changing market demands?

SQL Query 1 with Results

```
1  select P.productCode 'Product Code',P.productName 'Product Name',P.productLine 'Product Line',
2  P.quantityInStock 'Quantity in Stock',sum(os.quantityOrdered) as 'Total_Quantity'
3  from order_summary os
4  join products P
5  on P.productCode = os.productCode
6  group by 1
7  order by Total_Quantity desc
8  limit 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Product Code	Product Name	Product Line	Quantity in Stock	Total_Quantity
S18_3232	1992 Ferrari 360 Spider red	Classic Cars	8347	1808
S18_1342	1937 Lincoln Berline	Vintage Cars	8693	1111
S700_4002	American Airlines: MD-11S	Planes	8820	1085
S18_3856	1941 Chevrolet Special Deluxe Cabriolet	Vintage Cars	2378	1076
S50_1341	1930 Buick Marquette Phaeton	Vintage Cars	7062	1074

SQL Query 2 with Results

```
1  select P.productCode 'Product Code',P.productName 'Product Name',P.productLine 'Product Line',
2  P.quantityInStock 'Quantity in Stock',sum(os.quantityOrdered) as 'Total_Quantity'
3  from order_summary os
4  join products P
5  on P.productCode = os.productCode
6  group by 1
7  order by Total_Quantity
8  limit 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Product Code	Product Name	Product Line	Quantity in Stock	Total_Quantity
S18_4933	1957 Ford Thunderbird	Classic Cars	3209	767
S24_1046	1970 Chevy Chevelle SS 454	Classic Cars	1005	803
S24_3969	1936 Mercedes Benz 500k Roadster	Vintage Cars	2081	824
S18_2248	1911 Ford Town Car	Vintage Cars	540	832
S18_2870	1999 Indy 500 Monte Carlo SS	Classic Cars	8164	855

RECOMMENDATION: - Product Mix can be Optimized by following ways

- Bundling Top selling Products and providing a discount
- Bundling Top selling Quantity with least sold Quantity
- Increasing the Inventory by procuring "1941 Chevrolet Special Deluxe Cabriolet" since the stock in hand is less when compared to other top selling products

6) Are there any specific market segments where a particular product is underperforming, and how can it be improved?

SQL Query 1 with Results

```
1  select T.productCode,T.productName,t.quantityinstock,count(distinct territory) as 'Marketcount' from
2  (select O.territory,P.ProductCode,p.productName,p.quantityinstock,
3  sum(OS.TotalOrderValue) 'Total_Sales',round((P.msrp - P.buyPrice)/p.msrp*100,2) as 'Discount%'
4  from order_summary OS
5  join employees E on  E.employeeNumber = OS.salesRepEmployeeNumber
6  join offices O on O.officeCode = E.officeCode
7  right join products P on P.productCode = Os.ProductCode
8  group by O.territory, P.ProductCode,p.productName
9  order by Total_Sales) T group by T.productCode,T.productName
10 having Marketcount<4 order by Marketcount
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

productCode	productName	quantityinstock	Marketcount
S18_3233	1985 Toyota Supra	7733	0
S10_4757	1972 Alfa Romeo GTA	3252	3
S18_3029	1999 Yamaha Speed Boat	4259	3
S12_2823	2002 Suzuki XREO	9997	3
S700_3505	The Titanic	1956	3
S18_1367	1936 Mercedes-Benz 500K Special Roadster	8635	3
S24_2766	1949 Jaguar XK 120	2350	3
S18_4409	1932 Alfa Romeo 8C2300 Spider Sport	6553	3
S24_4258	1936 Chrysler Airflow	4710	3
S18_1749	1917 Grand Touring Sedan	2724	3

INSIGHTS:-

- Product '1985 Toyota Supra' is not being sold anywhere
- Above products are underperforming that these products are not even being sold in a particular territory

RECOMMENDATION: -

- Product '1985 Toyota Supra' should be discontinued for further procurement
- Existing stock of above products should be forced for sales as a complement or Grand Sale to clear the stock
- Above Products can be bundled with other products for discounted rate - for top customers, for promotion

7) What are the main factors that influence customer loyalty and repeat purchases?

SQL Query 1 with Results

```
1  select tt.territory,tt.city,tt.Status,count(*) from
2  (select O.territory,O.city, case
3  when ord.status = 'Disputed' then 'Failure'
4  when ord.status = 'Cancelled' then 'Failure'
5  when ord.status = 'On Hold' then 'Failure'
6  else 'Success' end as 'Status',
7  Os.ordernumber from Order_summary OS
8  join employees E on OS.salesRepEmployeeNumber=E.employeeNumber
9  join  offices O on O.officeCode=E.officeCode
```

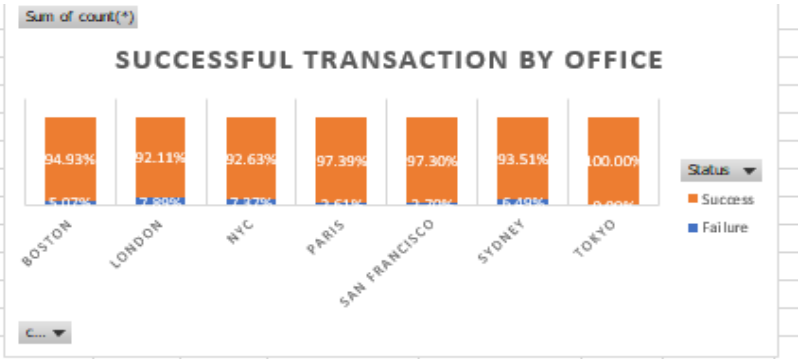
esult Grid | Filter Rows: | Export: | Wrap Cell Content:

territory	city	Status	count(*)
EMEA	London	Failure	36
EMEA	London	Success	420
EMEA	Paris	Failure	25
EMEA	Paris	Success	934
Japan	Tokyo	Success	137
NA	Boston	Failure	14
NA	Boston	Success	262
NA	NYC	Failure	26
NA	NYC	Success	327
NA	San Fr...	Failure	12
NA	San Fr...	Success	433

Territory	City	Status	count(*)
APAC	Sydney	Failure	24
APAC	Sydney	Success	346
EMEA	London	Failure	36
EMEA	London	Success	420
EMEA	Paris	Failure	25
EMEA	Paris	Success	934
Japan	Tokyo	Success	137
NA	Boston	Failure	14
NA	Boston	Success	262
NA	NYC	Failure	26
NA	NYC	Success	327
NA	San Francisco	Failure	12
NA	San Francisco	Success	433

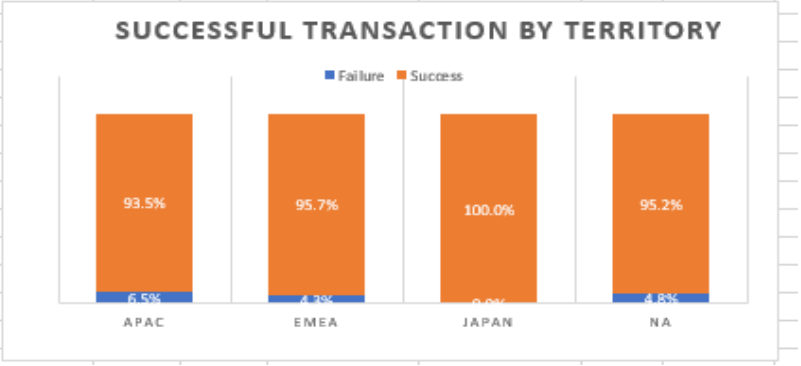
PIVOT 1

Sum of count(	Column Labels	
Row Labels	Failure	Success
Boston	5.07%	94.93%
London	7.89%	92.11%
NYC	7.37%	92.63%
Paris	2.61%	97.39%
San Francisco	2.70%	97.30%
Sydney	6.49%	93.51%
Tokyo	0.00%	100.00%
Grand Total	**4.57%**	**95.43%**



PIVOT 2

Sum of count(	Column Labels	
Row Labels	Failure	Success
APAC	6.5%	93.5%
EMEA	4.3%	95.7%
Japan	0.0%	100.0%
NA	4.8%	95.2%
Grand Total	**4.6%**	**95.4%**



INSIGHTS: - High Success ratio in terms of shipping status is the main factor for customer loyalty and repeat purchases

8) How do customer preferences differ based on geographic location, and how can marketing campaigns be customized accordingly?

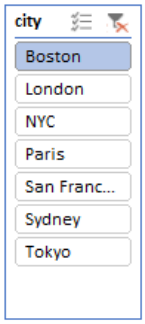
SQL Query 1 with Results

```
1  select o.city,count(os.ordernumber),sum(os.TotalOrderValue),month(os.orderdate) 'Mon' from  employees E
2  join offices O
3  on O.officeCode = E.officeCode
4  join order_summary OS
5  on OS.salesRepEmployeeNumber = E.employeeNumber group by o.city, month(os.orderdate) order by city,Mon
```

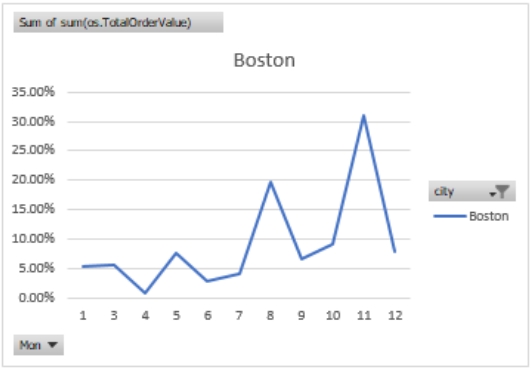
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

city	count(os.ordernumber)	sum(os.TotalOrderValue)	Mon
Boston	15	46853.94	1
Boston	14	50649.01	3
Boston	2	6276.60	4
Boston	21	68499.74	5
Boston	9	24510.67	6
Boston	11	36246.21	7
Boston	57	174945.08	8
Boston	19	58023.03	9
Boston	26	80722.24	10
Boston	82	276888.22	11
Boston	20	68923.88	12
London	55	160634.24	1
London	15	56141.64	2

PIVOT



Sum of sum(os.TotalOrderValue)	Column Labels
Row Labels	Boston
1	5.25%
3	5.67%
4	0.70%
5	7.67%
6	2.75%
7	4.06%
8	19.60%
9	6.50%
10	9.04%
11	31.02%
12	7.72%



INSIGHTS:

- November month is the peak revenue month for almost all cities due to the Christmas Season
- January is the second peak revenue month for Japan whereas that's not the case with Australia which depicts April
- This Infers that the Sale is good during Christmas or New year in most of the city.
- Highest sales is also attributed to the summer holidays which varies per region

RECOMMENDATION:-

- 2 months prior to the festival and summer holidays, the inventory check for in stock vs demand to be analyzed and corrected
- Off seasons when the market is having less demand, requires more promotion and discounts to improve sales

9) What are the characteristics of high-value customers, and how can similar customers be targeted for acquisition?

SQL Query 1 with Results

```
1  select C.customerNumber,count(distinct os.orderNumber) as 'Total_orders',
2  count(distinct os.productCode) as 'Products',
3  count(distinct E.officeCode) as 'Office_Count',
4  sum(os.quantityOrdered) as 'Total_Quantity',
5  sum(os.TotalOrderValue) as 'Total_sales',
6  round(avg(os.TotalOrderValue),2) as 'Average_ordervalue',
7  round(sum(os.TotalOrderValue)/C.creditLimit*100,2) as 'Credit_Utilisation'
8  from order_summary os join customers C on C.customerNumber = os.customerNumber
9  join employees E on E.employeeNumber = os.salesRepEmployeeNumber
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

customerNumber	Total_orders	Products	Office_Count	Total_Quantity	Total_sales	Average_ordervalue	Credit_Utilisation
141	26	106	1	9327	820689.54	3168.69	360.58
124	17	77	1	6366	591827.34	3287.93	281.15
114	5	40	1	1926	180585.07	3283.36	153.95
151	4	32	1	1775	177913.95	3706.54	128.46
119	4	43	1	1832	158573.12	2991.95	134.16
148	5	37	1	1524	156251.03	3633.74	150.53
323	5	42	1	1691	154622.08	3361.35	175.71
131	4	41	1	1631	149085.15	3042.55	129.75
187	3	37	1	1778	148410.09	2910.00	108.49
450	4	39	1	1656	143536.27	3588.41	184.97

customerNumber	Total_orders	Products	Office_Count	Total_Quantity	Total_sales	Average_ordervalue	Credit_Utilisation
141	26	106	1	9327	820689.54	3168.69	360.58
124	17	77	1	6366	591827.34	3287.93	281.15
114	5	40	1	1926	180585.07	3283.36	153.95
151	4	32	1	1775	177913.95	3706.54	128.46
119	4	43	1	1832	158573.12	2991.95	134.16
148	5	37	1	1524	156251.03	3633.74	150.53
323	5	42	1	1691	154622.08	3361.35	175.71
131	4	41	1	1631	149085.15	3042.55	129.75
187	3	37	1	1778	148410.09	2910	108.49
450	4	39	1	1656	143536.27	3588.41	184.97

INSIGHTS:

- High Order Volume in a given period
- High Average Sales per Order
- High Quantity of products (Bulk Order)
- Who utilizes and purchases over and above credit limit

RECOMMENDATION: - To acquire higher valued customer

- Sales team should focus on evaluating existing customer for untapped potential
- Marketing team to target big corporates for acquisition, who can purchase in bulk and in regular frequency.

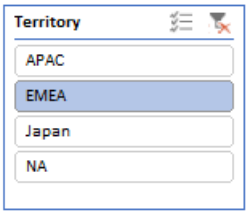
10) How can marketing strategies be tailored to target specific demographic segments in different regions?

SQL Query 1 with Results

```
1  select o.territory,sum(os.totalordervalue),month(os.orderdate) 'Monthly' from  employees E
2  join offices O
3  on O.officeCode = E.officeCode
4  join order_summary OS
5  on OS.salesRepEmployeeNumber = E.employeeNumber group by o.territory,
6  month(os.orderdate) order by o.territory,Monthly
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

territory	sum(os.totalordervalue)	Monthly
APAC	27083.78	1
APAC	88801.22	2
APAC	95112.21	3
APAC	156313.11	4
APAC	84389.29	5
APAC	122091.66	6
APAC	136347.39	7
APAC	90161.53	9
APAC	203214.17	11
APAC	143661.99	12
EMEA	397827.94	1
EMEA	428589.58	2
EMEA	350080.59	3
EMEA	286834.05	4
EMEA	470908.91	5
EMEA	167532.11	6
EMEA	202887.95	7
EMEA	194776.54	8
EMEA	262012.52	9



Row Labels	Sum of Revenue
1	397827.94
2	428589.58
3	350080.59
4	286834.05
5	470908.91
6	167532.11
7	202887.95
8	194776.54
9	262012.52
10	551617.49
11	965730.45
12	241914.15
Grand Total	**4520712.28**

11) What are the potential untapped markets based on demographic indicators, and how can market penetration be increased?

SQL Query with Results

```
1  select O.territory as 'Territory',count(distinct OS.customerNumber) as 'Customer_Volume',
2  count(distinct OS.orderNumber) as 'Order_volume',
3  count(distinct OS.productcode) as 'Product_volume',
4  sum(OS.TotalOrderValue) as 'Revenue',
5  count(distinct E.employeeNumber) as 'Employee_Volume',
6  count(distinct O.officeCode) as 'Office_count'
7  from employees E
8  join offices O
9  on O.officeCode = E.officeCode
10 join order_summary OS
11 on OS.salesRepEmployeeNumber = E.employeeNumber
12 group by O.territory
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Territory	Customer_Volume	Order_volume	Product_volume	Revenue	Employee_Volume	Office_count
EMEA	45	153	109	4520712.28	6	2
NA	38	119	109	3479191.91	6	3
APAC	10	38	109	1147176.35	2	1
Japan	5	16	81	457110.07	1	1

Territory	Average Order Volume Per Customer	Average Revenue Per Customer	Average Revenue per Employee	Average Revenue per Office	Rank - Average Order Volume Per Customer	Rank - Average Revenue Per Customer	Rank - Average Revenue per Employee	Rank - Average Revenue per Office
EMEA	3.4	100460.273	753452.0467	2260356	2	2	1	1
NA	3.1	91557.6818	579865.3183	1159731	4	3	2	2
APAC	3.8	114717.635	573588.175	1147176	1	1	3	3
Japan	3.2	91422.014	457110.07	457110	3	4	4	4

INSIGHTS:-

- APAC is the market untapped which has lot of potential
- APAC has highest Order per customer and Highest Revenue per customer which is operating with just 1 office and 2 employees
- Japan has been exposed to just 81 products for orders out of 109 product variances.

RECOMMENDATION:-

- More offices can be operated in APAC region to increase customer base and revenue
- Marketing strategies should be devised for Japan to expose all product variance unless if there is any country specific restriction on any products

12) How do customer preferences and behavior differ based on demographic factors, and how can they be leveraged for personalized marketing campaigns?

SQL Query 1 with Results

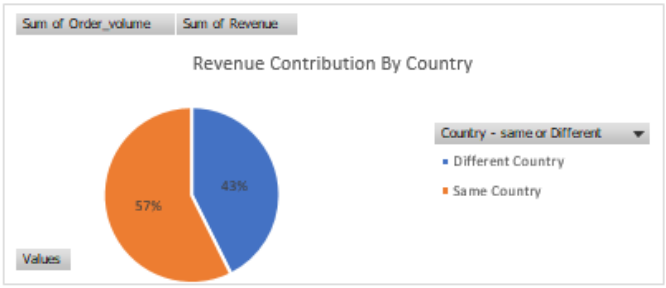
```
select c.city 'Customer City',o.city 'Office City',c.country 'Customer Country', o.country 'Office Country',
count(distinct OS.customerNumber) as 'customer_Volume',count(distinct OS.orderNumber) as 'Order_volume',
count(distinct OS.productcode) as 'Product_volume',
sum(OS.TotalOrderValue) as 'Revenue' from employees E
join offices O on O.officeCode = E.officeCode
join order_summary OS on OS.salesRepEmployeeNumber = E.employeeNumber
join customers c on c.customerNumber=os.customernumber
group by c.city,o.city,c.country, o.country order by Revenue desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Customer City	Office City	Customer Country	Office Country	customer_Volume	Order_volume	Product_volume	Revenue
Madrid	Paris	Spain	France	3	31	108	979880.77
San Rafael	San Francisco	USA	USA	1	17	77	591827.34
NYC	NYC	USA	USA	5	16	81	497941.50
Auckland	Sydney	New Zealand	Australia	3	12	85	386514.63
Paris	Paris	France	France	3	9	50	240649.68
San Francisco	San Francisco	USA	USA	2	7	45	199051.34

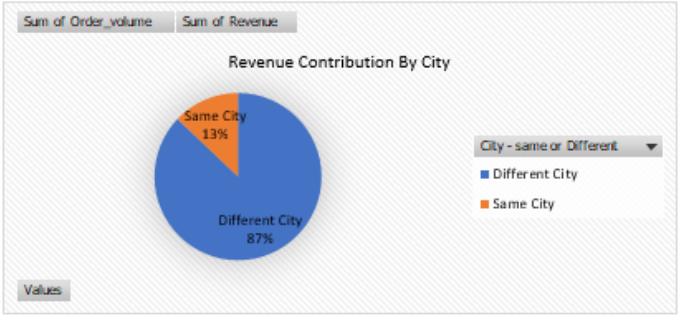
PIVOT 1

Row Labels	Sum of Order_volume	Sum of Revenue
Different Country	139	4156096.56
Same Country	187	5448094.05
Grand Total	**326**	**9604190.61**



PIVOT 2

Row Labels	Sum of Order_volume	Sum of Revenue
Different City	284	8406523.15
Same City	42	1197667.46
Grand Total	**326**	**9604190.61**



INSIGHTS:-

- Most Customers prefer buying it from same country.
- City does not play major role in orders.
- Madrid is the City where the customers order the most for highest revenue from office in Paris

RECOMMENDATION:-

- ABC Retail should open new offices in country for much more smoother order experience in terms of low shipping cost, to tailor products or sales region wise.
- ABC Retail should aim at centralizing the offices within a country, as customers don't really order based on the city of the office.

1.10 FUTURE ANALYSIS

Following Insights can be brought up if we get following info

Data	Insights
Customer Age & Gender	Categorization of customers based on the Age as Young and Old and Gender, to promote marketing and products for target audience.
Relationship of Order and Payment	To establish the number of days between Order Date and Payment Date to identify the customers based on the delinquency of payment and credit limit utilization for loyalty-based promotions.
Customer Onboarding Date	To determine the lifetime of the customer and company relationship for loyalty-based promotions.