

MindX Technology School
X-Data | PNL-BI41
Khang Nguyen

Business Context:

This is a transnational data set that contains all the transactions occurring between 01/12/2020 and 09/12/2021 for a UK-based and registered non-store online retail. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

Information is a main asset of businesses nowadays. The success of a business in a competitive environment depends on its ability to acquire, store, and utilize information. Data is one of the main sources of information. Therefore, data analysis is an important activity for acquiring new and useful information. Analyze this dataset and try to answer the following questions. The following table is the description and data type of each column.

Column	Data Type	Description
InvoiceNo		A six-digit unique number that defines each transaction.
StockCode	Categorical	A five or six-digit unique character used to identify a specific product.
Description	Categorical	Product or item name.
Quantity	Numeric	The quantity of each product per transaction. Negative values related to canceled transactions.
InvoiceDate		The date when each transaction was generated.
UnitPrice		The price of each product per unit is in pounds sterling.
CustomerID ^(*)	Categorical	A five-digit unique number that defines each customer.
CusName		Customer's name.
Country		Customer's country.
Gender		Customer's gender.

Requirements:

A. Connect the data into Power BI (1 point)

B. Build a simple analysis dashboard (3 points)

- Highlight Cards: Total Sales, Total Quantity, Total Number of Customers, Average Sales of Orders, Average Quantity of Orders.
- Visualizations: Trending Sales by Month, Sales by Country, Sales by Customer, Average Sales of Orders by Month, Average Quantity of Orders by Month.
- Slicers: Year & Country.

C. Advanced analytics techniques (4 points)

- RFM Analysis: Each segment contributes how many percent in Total Sales?
- Cohort Analysis: Retention rate by month in 2021?

D. Insights—Make recommendations for increasing sales (2 points)

(The answers to this test are provided on the following pages.)

A. Connect the data into Power BI.

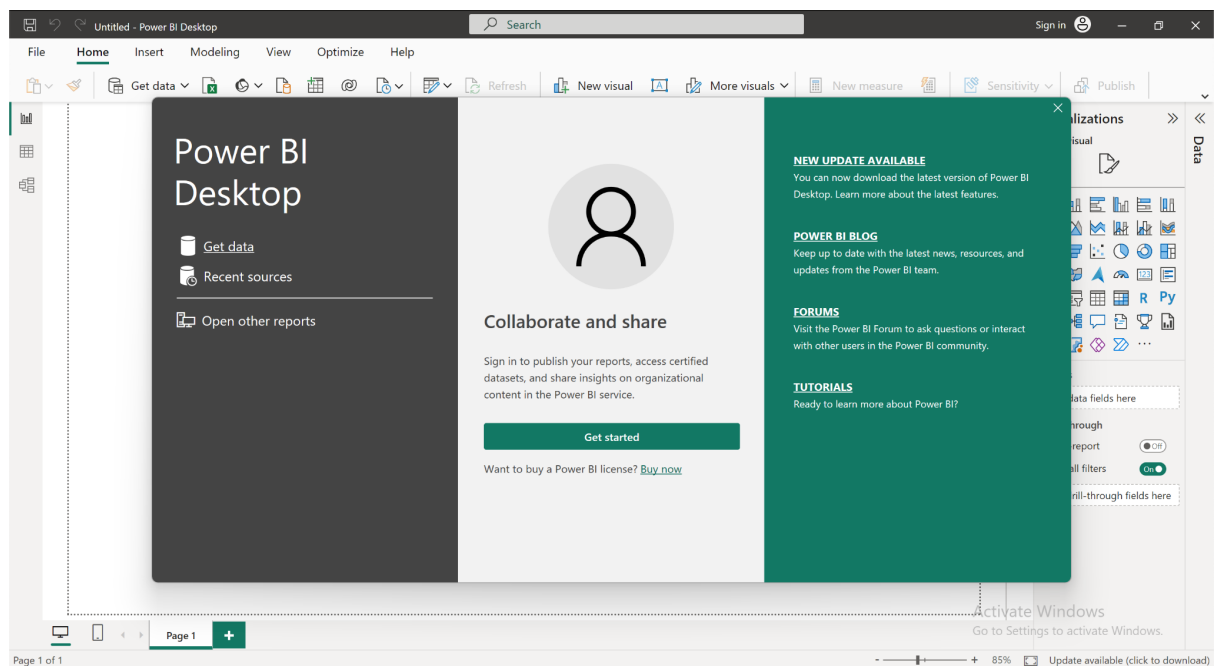


Figure 01: Open Power BI and click 'Get data' on the left of the application.

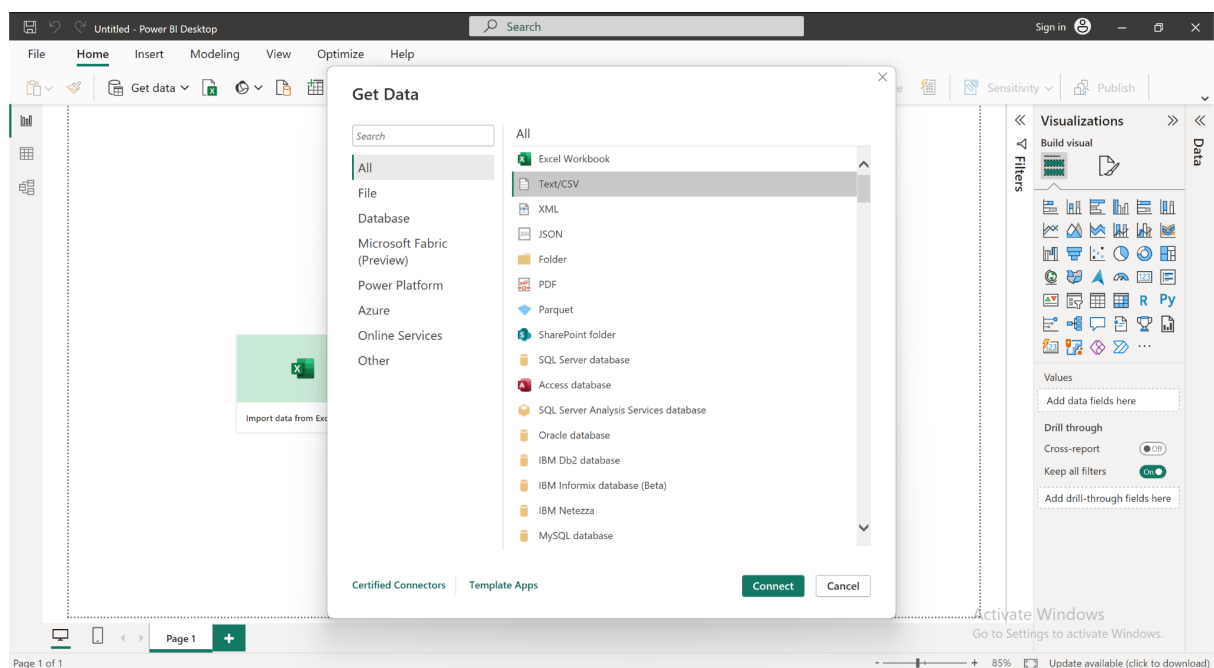


Figure 02: Select the data source as 'Text/CSV' and press 'Connect'.

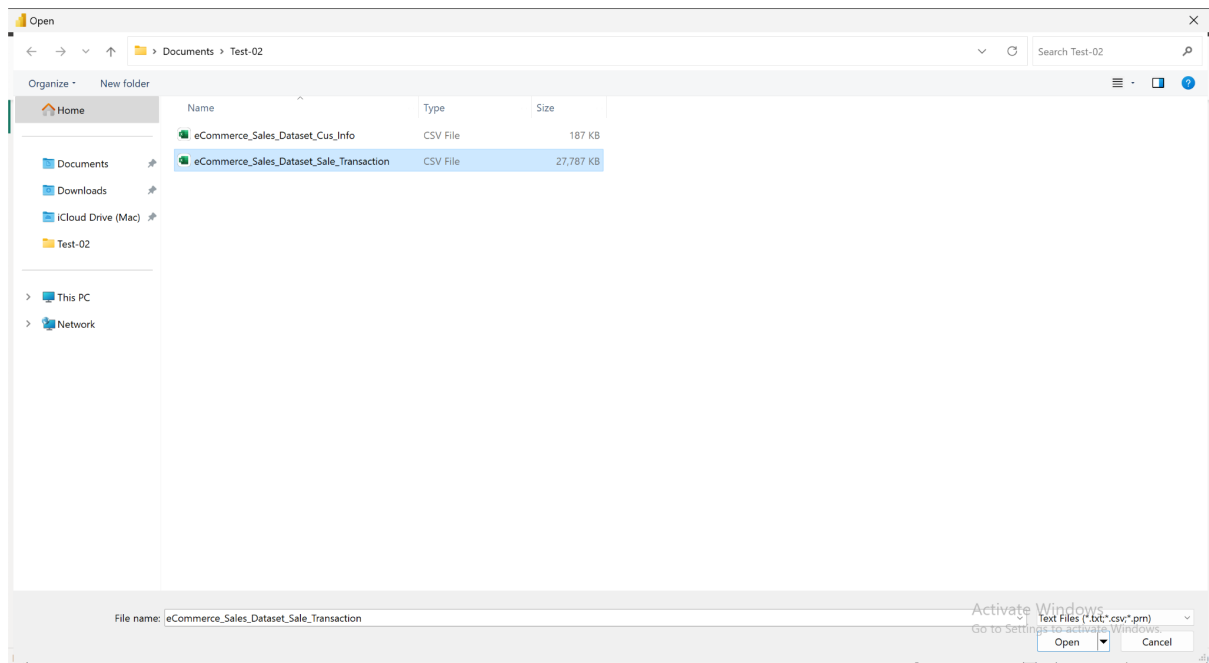


Figure 03: Select the file and open it to connect, now we need-
[eCommerce_Sales_Dataset_Sale_Transaction.csv].

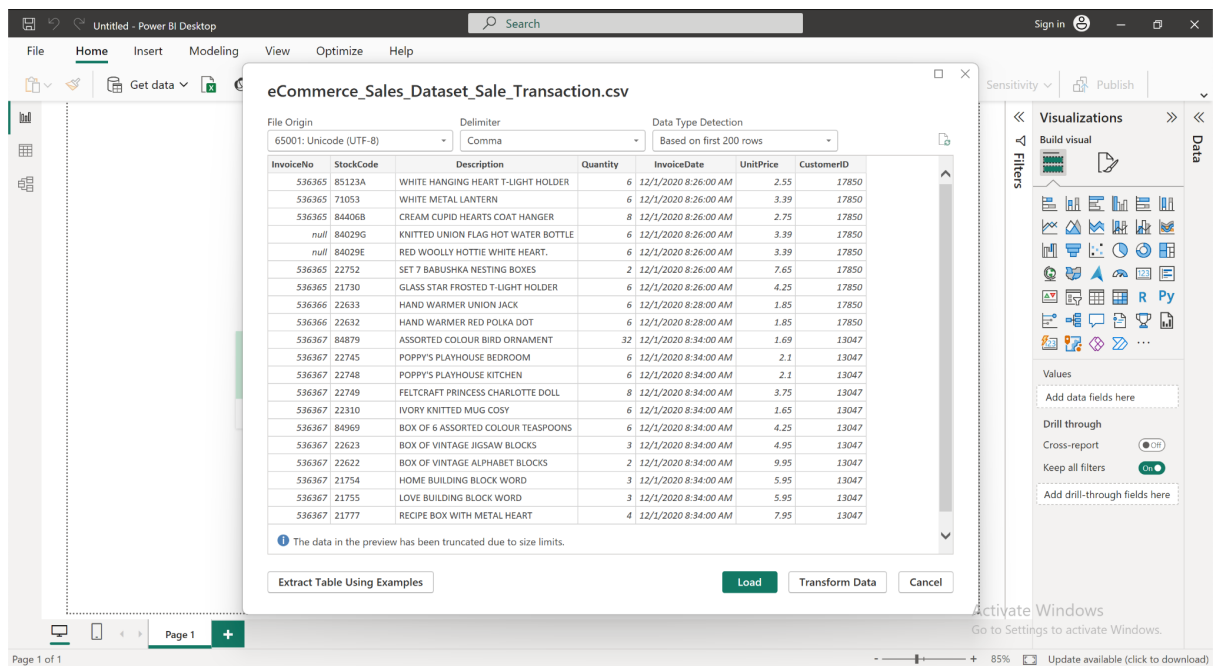


Figure 04: Check the values and click 'Load' below the preview pop-up session.

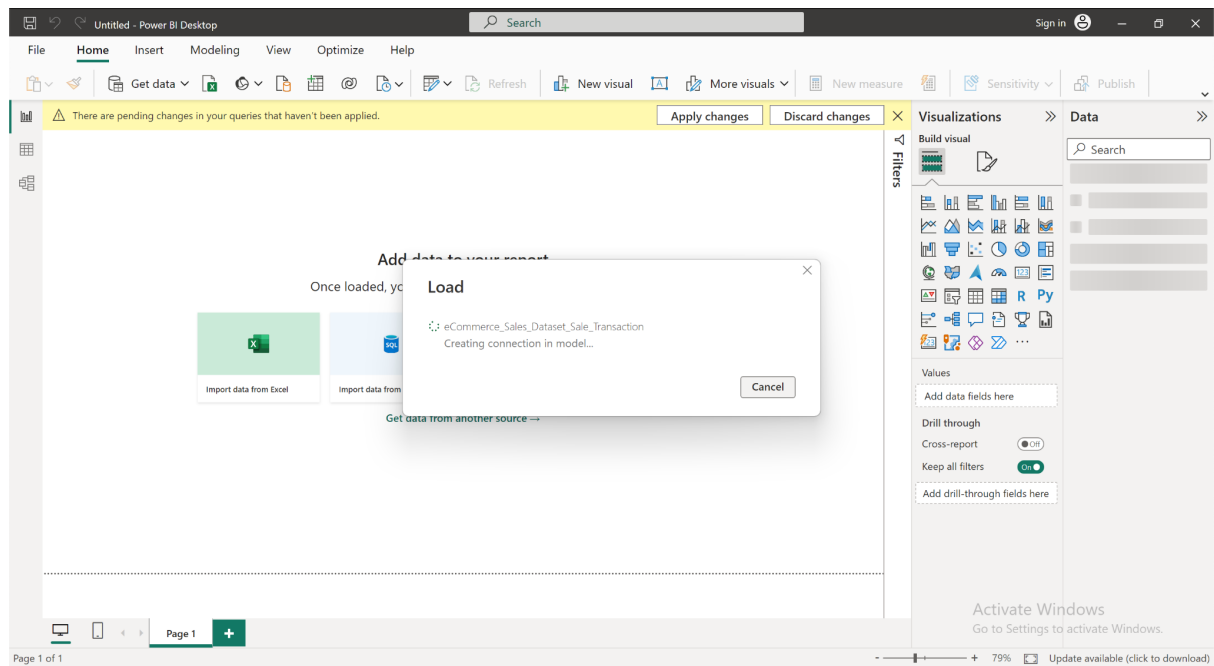


Figure 05: Data is being loaded.

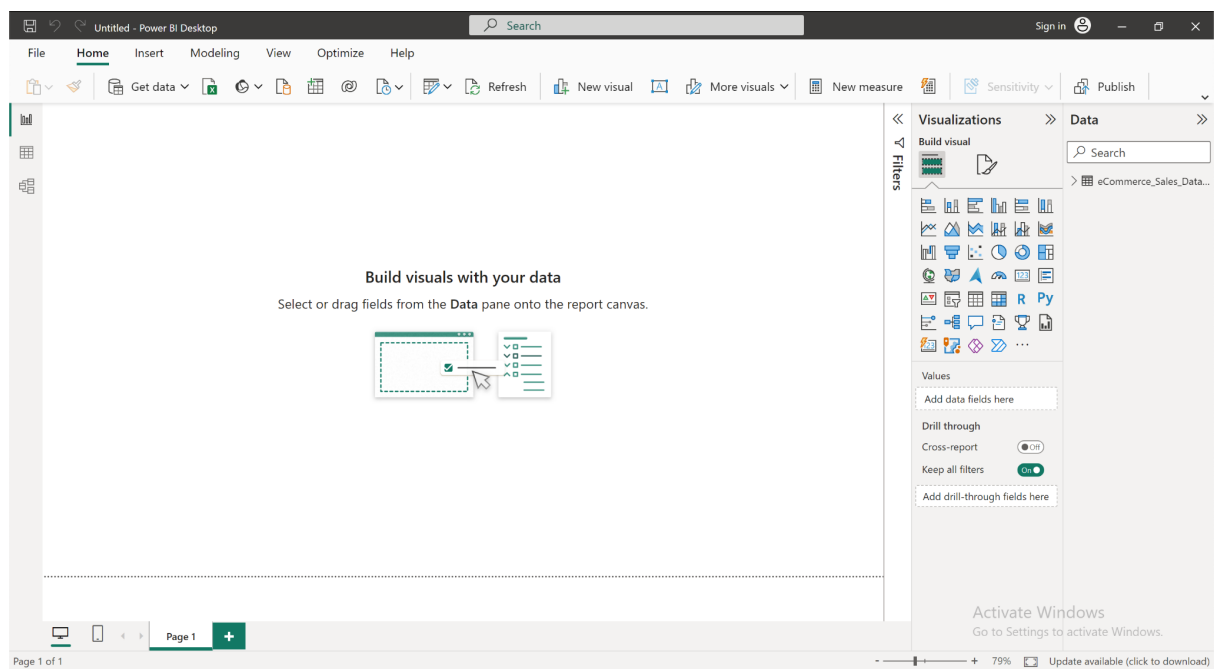


Figure 06: We have successfully connected to the table [eCommerce_Sales_Dataset_Sale_Transaction.csv].

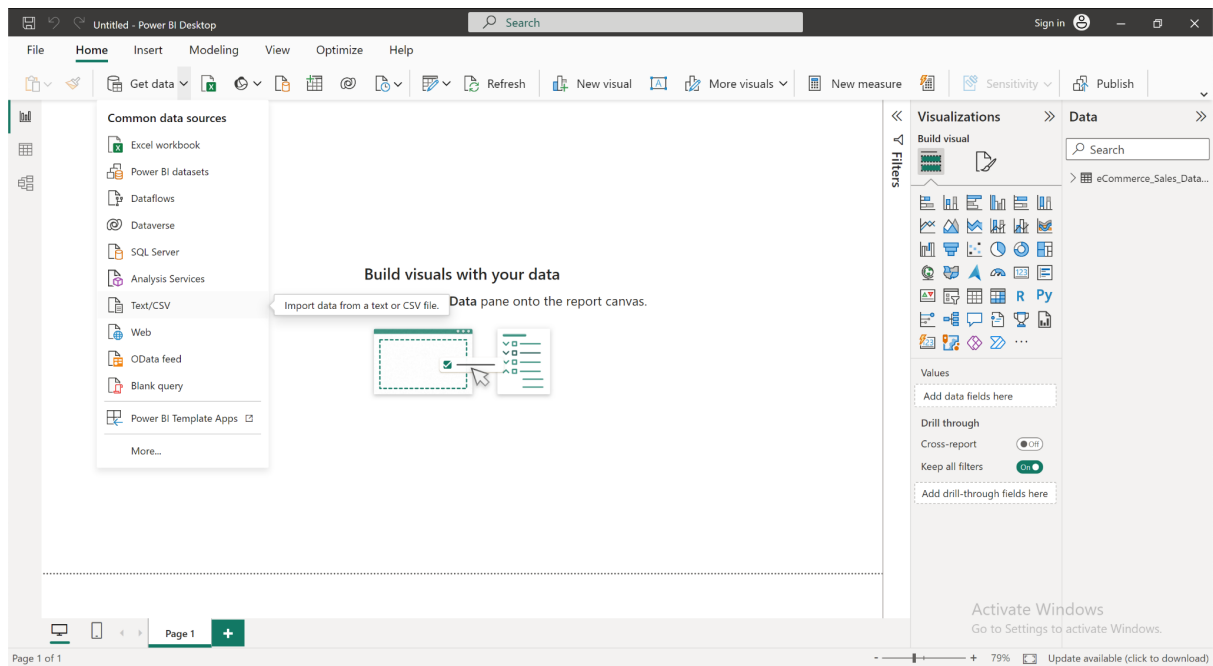


Figure 07: Click on the data source next to 'Get data' on the left of the application, then select 'Text/CSV'.

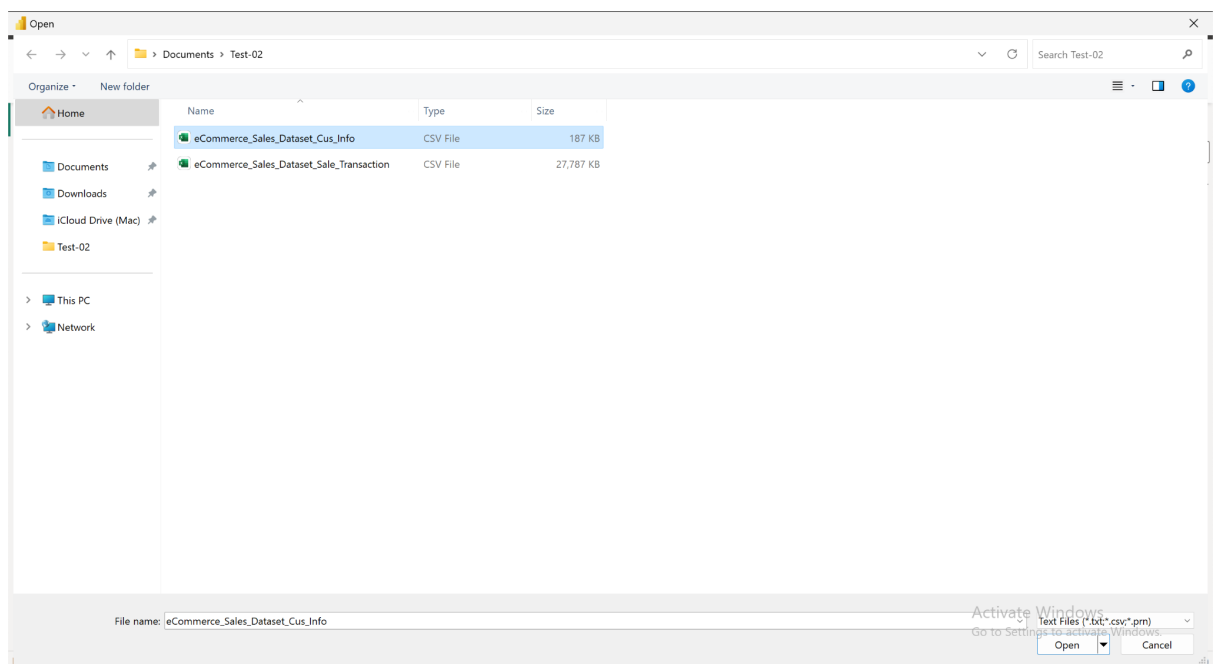


Figure 08: Select the file and open it to connect, now we need [eCommerce_Sales_Dataset_Cus_Info.csv].

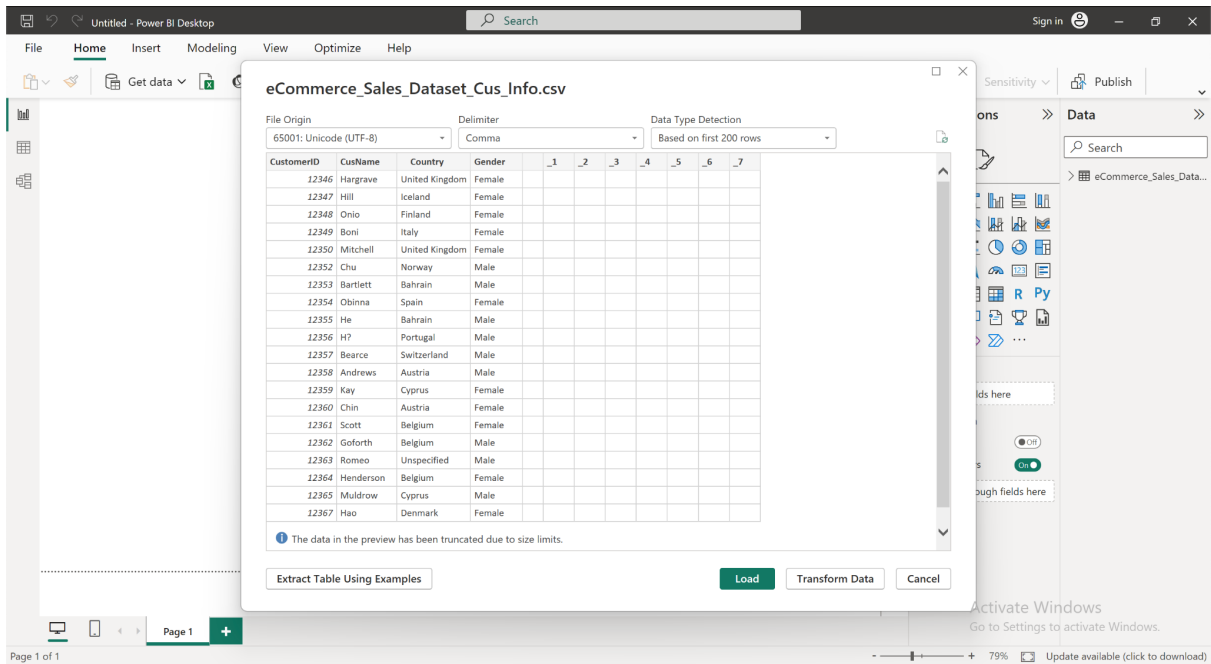


Figure 09: Check the values and click 'Load' below the preview pop-up session.

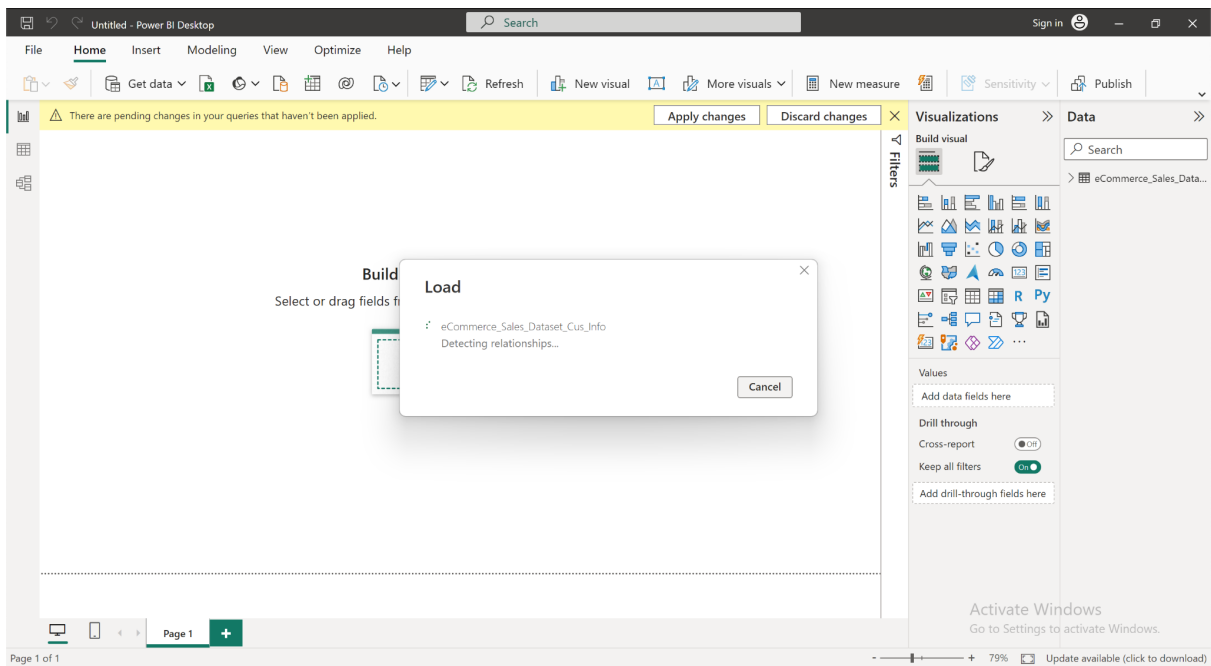


Figure 10: Data is being loaded.

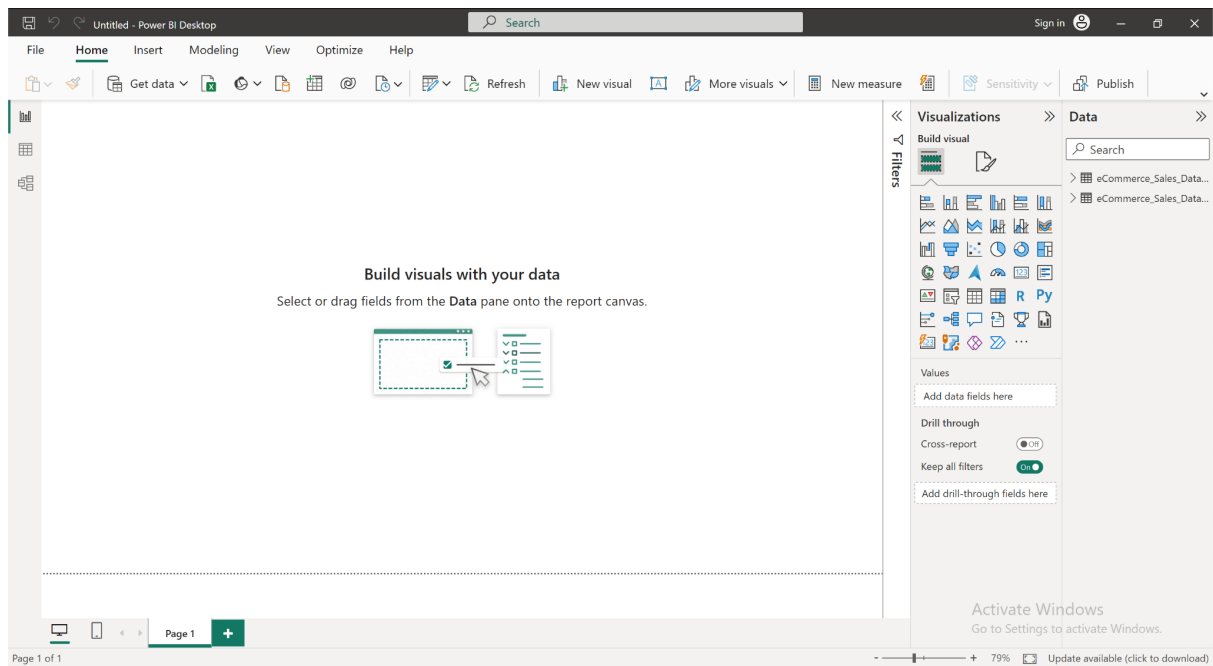


Figure 11: We have successfully connected to the table [eCommerce_Sales_Dataset_Cus_Info.csv].

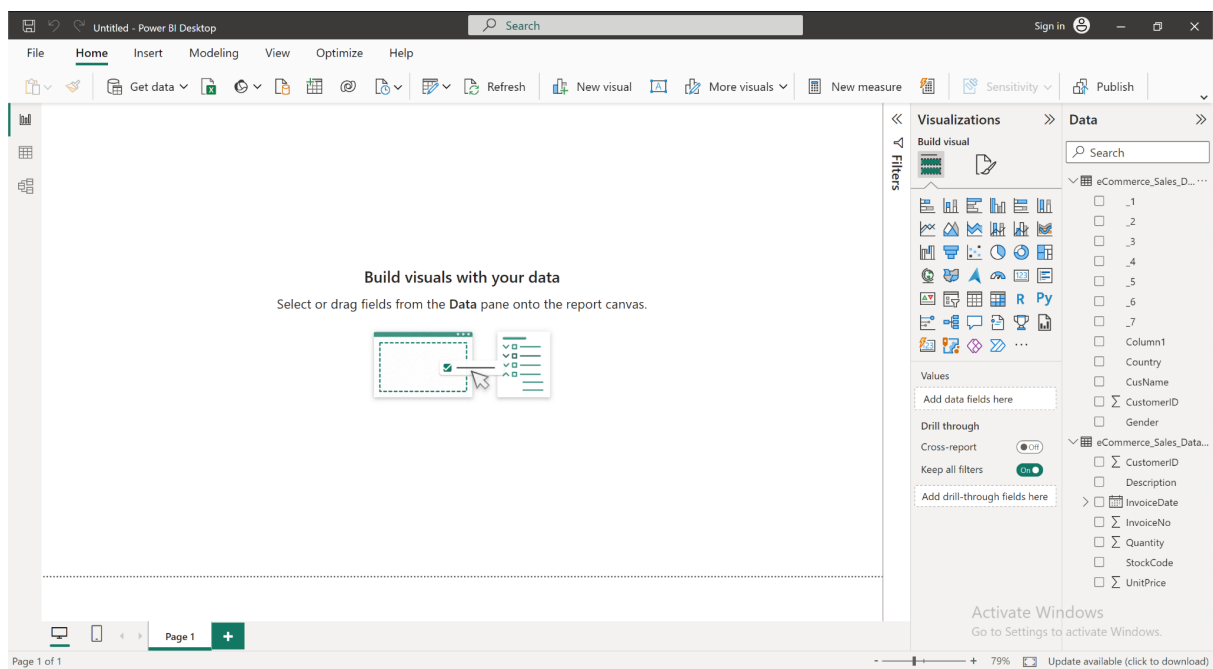


Figure 12: Check the 2 connected tables, we have successfully connected the data to Power BI.

B. Build a simple analysis dashboard.

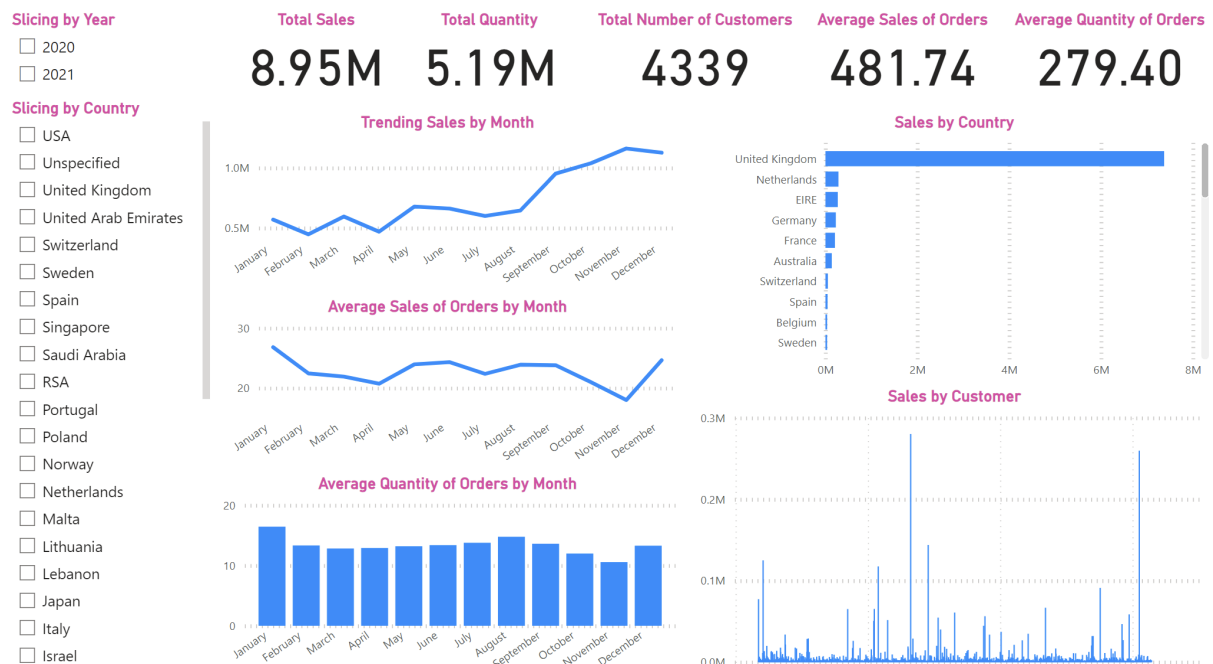


Figure 13: The dashboard above includes the following visuals: Highlight Cards, Visualizations & Slicers.

C. Advanced analytics techniques.

RFM Analysis: Each segment contributes how many percent in Total Sales?

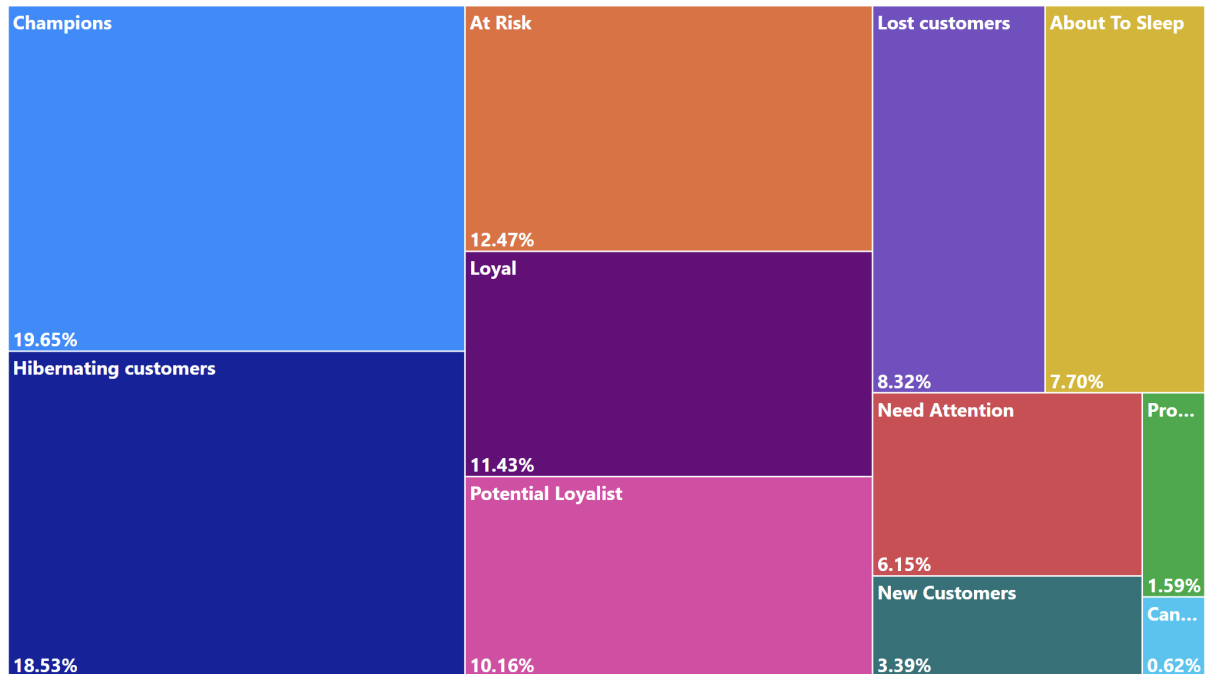


Figure 14: RFM Analysis and Customer Segmentation using Treemap.

Calculate Each Segment's Sales and Percent Contribution:

By visualizing Total Sales by Each Segment, we can see how much each segment contributes in both sales and percentage to Total Sales. Here I use a Pie Chart to show that.

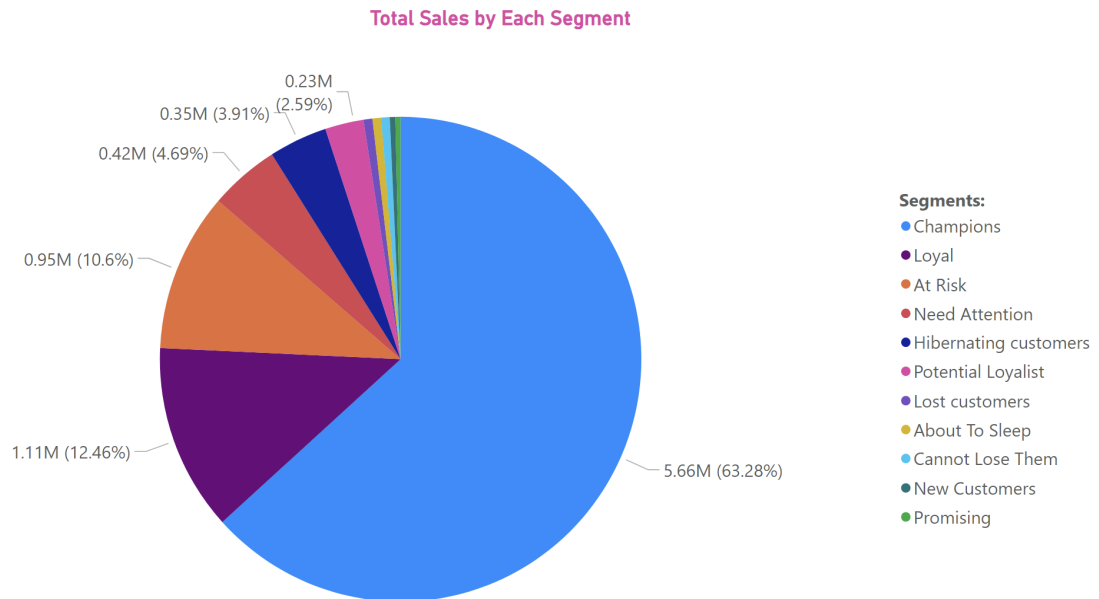


Figure 15: Total Sales Contribution by Each Segment by using Pie Chart.

The table below lists the percent contribution of each segment to Total Sales in descending order.

Segment Type	Segment Sales	Percent Contribution
Total Sales	8,947,369.22	100.00%
Champions	5,661,544.88	63.28%
Loyal	1,114,695.13	12.46%
At Risk	948,313.85	10.60%
Need Attention	419,607.09	4.69%
Hibernating Customers	350,265.52	3.91%
Potential Loyalist	231,831.94	2.59%
Lost Customers	53,694.68	0.60%
About To Sleep	52,144.03	0.58%
Cannot Lose Them	51,075.55	0.57%
New Customers	33,185.43	0.37%
Promising	31,011.12	0.35%

Figure 16: Total Sales Contribution by Each Segment by using Table.

Cohort Analysis: Retention rate by month in 2021?

Count of CustomerID	Column Labels													
Row Labels	0	1	2	3	4	5	6	7	8	9	10	11	12	Grand Total
12/31/2020	28405	10191	8777	11268	9728	12055	11191	11428	11417	14232	13545	22086	7254	171577
1/31/2021	11041	2220	3039	2314	3627	3243	2793	2880	3793	5056	6398	1371		47775
2/28/2021	8931	1412	1935	2505	2278	2027	2248	2740	2981	2946	451			30454
3/31/2021	11458	1429	2609	2384	2445	2115	2962	3579	4297	996				34274
4/30/2021	7238	1405	1287	1301	1364	2038	1904	2186	523					19246
5/31/2021	6121	1090	917	917	1509	2357	1980	791						15682
6/30/2021	5712	924	714	1528	1762	2603	627							13870
7/31/2021	4992	501	1333	1368	1814	526								10534
8/31/2021	4857	1600	2849	2804	899									13009
9/30/2021	8295	2383	2646	867										14191
10/31/2021	11773	3583	885											16241
11/30/2021	11202	1130												12332
12/31/2021	984													984
Grand Total	121009	27868	26991	27256	25426	26964	23705	23604	23011	23230	20394	23457	7254	400169
Row Labels	0	1	2	3	4	5	6	7	8	9	10	11	12	Grand Total
12/31/2020	100%	36%	31%	40%	34%	42%	39%	40%	40%	50%	48%	78%	26%	171577
1/31/2021	100%	20%	28%	21%	33%	29%	25%	26%	34%	46%	58%	12%	0%	47775
2/28/2021	100%	16%	22%	28%	26%	23%	25%	31%	33%	33%	5%	0%	0%	30454
3/31/2021	100%	12%	23%	21%	21%	18%	26%	31%	38%	9%	0%	0%	0%	34274
4/30/2021	100%	19%	18%	18%	19%	28%	26%	30%	7%	0%	0%	0%	0%	19246
5/31/2021	100%	18%	15%	15%	25%	39%	32%	13%	0%	0%	0%	0%	0%	15682
6/30/2021	100%	16%	13%	27%	31%	46%	11%	0%	0%	0%	0%	0%	0%	13870
7/31/2021	100%	10%	27%	27%	36%	11%	0%	0%	0%	0%	0%	0%	0%	10534
8/31/2021	100%	33%	59%	58%	19%	0%	0%	0%	0%	0%	0%	0%	0%	13009
9/30/2021	100%	29%	32%	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	14191
10/31/2021	100%	30%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	16241
11/30/2021	100%	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	12332
12/31/2021	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	984
Grand Total	121009	27868	26991	27256	25426	26964	23705	23604	23011	23230	20394	23457	7254	400169

Figure 17: Cohort Analysis in Excel.

D. Insights—Make recommendations for increasing sales.

Below are my recommendations for the company based on Months, Countries, and Customers.

[By Months]

Overall, monthly revenue exhibited steady growth throughout the year, with a notable increase in the number of orders, up to 50%. In the early months, while several festivals occur, total sales, average sales, and the average number of items sold remain low, accounting for less than half of peak months. Leveraging major holidays like Valentine's Day and Easter during this period presents an opportunity for businesses to drive sales by offering thematic products and running awareness-enhancing advertising and incentive programs.

Mid-year, although total sales are moderate, the stability of average sales and product sales averages indicates potential for selling holiday-themed items during this period. Alternatively, businesses can utilize this time to clear inventory from earlier months through discount programs, "buy one get one" offers, and shock deals, while also promoting pre-orders for products with high year-end demand.

September marked a significant sales spike, suggesting the influence of a specific marketing campaign or promotional activity. The last quarter of the year experienced a sharp increase in sales, with November being the peak month. With numerous festivals focusing on family and friends during this period, demand surpasses that of the earlier months (primarily oriented towards individuals or couples). Promoting festival-themed products during this period is recommended.

Regarding average order quantity, August records the highest average count at 14.77, while November shows the lowest count at 10.56. This underscores the seasonal trend of heightened orders during December and the summer months of August.

In summary, the data reflects consistent patterns from January to December, punctuated by fluctuations. This stability in customer ordering behavior underscores the importance of maintaining a year-round marketing and engagement strategy. By recognizing seasonal trends, businesses can effectively prepare for peak demand during the holiday season and summer months while tailoring their marketing strategies accordingly. Furthermore, the observed spike in September and drop in November call for targeted marketing efforts during those months to boost sales.

[By Countries]

Major sales primarily originate from neighboring European countries, with England being the highest revenue-generating country. These European markets share similar cultures and benefit from convenient goods transportation due to proximity and the use of a common currency.

Among the top 10 countries by sales volume, Australia, located in the Southern Hemisphere, holds the 6th position, signifying its potential as a market. The geographical distinction of the southern hemisphere, with its opposing seasons to northern hemisphere countries, provides an avenue for diversifying products and stimulating innovation.

Japan and Singapore, ranked 11th and 15th, respectively, exhibit substantial sales, indicating promising areas for market expansion. These culturally rich countries have also served as inspiration for product diversification and tailored customer engagement strategies.

Regional Tailoring: Develop specific marketing and product strategies for each region or country, taking into account cultural nuances, consumer preferences, and market dynamics.

UK Market: Acknowledge the UK's pivotal role, with a substantial revenue contribution of \$7,375,315.07. Consider diversification and risk mitigation strategies to reduce dependence on a single market while maintaining focus on its importance.

Global Presence: Recognize the diverse international footprint, with notable contributions from countries like the Netherlands, EIRE, Germany, and France. Explore growth opportunities in countries with lower sales, implementing targeted marketing and strategic initiatives.

In summary, the sales data highlights a robust international revenue distribution, with the UK as the primary market. The presence of numerous countries with varying revenue levels underscores global reach and presents opportunities for expansion and diversification. Lower sales regions, such as Saudi Arabia, Hong Kong, and Bahrain, should be considered for market expansion or optimization efforts. This data underlines the potential for growth and strategic development within the company's global operations.

[By Customers via RFM Analysis]

The data exhibits the Pareto principle, where a small set of high-value customers significantly impacts total sales. Prioritizing retention and engagement strategies for these top spenders is pivotal for consistent revenue growth.

Simultaneously, a notable segment comprises customers with minimal or no purchases. Reconnecting with these individuals through targeted marketing or incentives can potentially convert them into active, valuable customers.

In essence, this data underscores the necessity of a customer-focused approach in sales and marketing. By segmenting and understanding the customer base, businesses can optimize strategies to retain high-value customers and reignite the engagement of less active customers. This approach fosters revenue growth and enhances overall customer satisfaction.

High-Value Customer Focus: Implement personalized strategies to retain and enhance relationships with top-spending customers.

Re-engagement Initiatives: Design targeted marketing or incentive programs to reactivate less engaged customers, potentially converting them into active patrons.

Customer-Centric Approach: Ensure marketing and sales strategies prioritize understanding and meeting the diverse needs of customers for sustained growth and satisfaction.

Segmented Marketing Campaigns: Invest in targeted marketing campaigns aimed at specific customer segments to optimize results. Focus on the At Risk segment by offering discount codes to re-engage them.

Product Recommendations: Recommend new products based on customers' purchase history, tailoring suggestions to their preferences.

Champions and Loyal Customer Groups: These high-value segments represent significant revenue. Implement special incentives such as vouchers, early access to discounts, or product samples to encourage repeat business. Ensure they are informed about upcoming promotions and consider loyalty programs. Encourage them to refer new customers through an affiliate program.

Potential Loyalists: Care for and retain this group, providing product updates and opportunities to transition to the Champions segment. Encourage them to introduce new buyers and offer commissions for successful referrals.

Hibernating and At Risk Groups: Implement re-engagement campaigns and offer suitable products to entice them back. Consider incentive programs, lucky draws, and discount codes to incentivize future purchases.

Need Attention, New Customers, Promises, and Cannot Lose Them Groups: Though smaller, actively engage with these customer segments. Gather insights to create tailored programs, such as developing products aligned with specific interests, like Japanese culture for customers interested in.

It's noteworthy that the "Champions" segment, representing 19.65%, consists of highly engaged and valuable customers, while the "Cannot Lose Them" segment, at 0.62%, may include high-value customers in a smaller quantity. These insights help prioritize customer relationship management efforts effectively.

[By Customers via Cohort Analysis]

Customer retention and revenue analysis is a critical aspect of business strategy. In this examination, we delve into customer return rates, revenue patterns, and engagement factors throughout the year. We also emphasize the importance of personalized engagement and customer lifecycle analysis in enhancing retention and fostering long-term growth.

Mid-Year Revenue Review: Evaluate mid-year activities and product offerings to align with customer needs. Adjust pricing as necessary to boost revenue during this period.

Year-End Sales Strategy: Capitalize on peak activity in September, October, and November by implementing holiday promotions and early-bird discounts. Address delivery concerns to stimulate year-end sales.

Retention Analysis: Identify retention declines over time and target personalized marketing for cohorts with stronger retention. Cultivate customer loyalty through tailored engagement.

Customer Lifecycle Examination: Examine the customer journey to pinpoint retention drop-off points. Implement precise retention strategies for customers during these stages.

In summary, the cohort analysis reveals insights into retention and engagement trends. While a common decline in retention is observed, tailored engagement and lifecycle analysis offer opportunities to improve customer loyalty and drive long-term success. Continuously adapt strategies based on these insights to enhance customer retention and business growth.