Asian Institute of Technology

School of Engineering and Technology
Department of ICT
Computer Science and Information Management
Data Science and AI Program

Business Intelligence System for Global Sales Analysis and Customer Segmentation

Final Report

[AT 82 - DSAI] Business Intelligence and Analytics (BI&A)

Group7

Nathas Sungworawongpana (st124323)

Wut Yee Aung   (st124377)

Kyi Thin Nu   (st124087)

Business Intelligence System for Global Sales Analysis and Customer Segmentation

This proposal outlines the implementation of a comprehensive Business Intelligence System focusing on global sales analysis and customer segmentation. The project aims to address current challenges in decision-making by providing real-time insights through an interactive BI dashboard and facilitating strategic planning through a Decision Support System (DSS) for customer segmentation. The scope includes analyzing sales data across multiple countries, designing an intuitive BI dashboard, and implementing advanced algorithms for effective customer segmentation. By leveraging a robust technology stack and adhering to a structured methodology, the proposed BI system is expected to enhance overall business intelligence, leading to improved decision-making, increased operational efficiency, and heightened customer satisfaction. The project timeline, budget, and expected outcomes are detailed to provide a comprehensive overview of the initiative's scope and impact.

1. Project Background

In today's dynamic business landscape, organizations face the challenge of managing and extracting meaningful insights from vast and diverse datasets. As we recognize the critical role that data plays in informed decision-making and strategic planning. For the businesses operating across multiple countries, the need for a robust Business Intelligence (BI) System has become increasingly evident to analyze global sales trends and enhance customer segmentation for more effective decision support.

Traditional methods of analysis lack the agility and depth required to keep pace with the rapidly evolving market dynamics. Additionally, the absence of a dedicated Decision Support System (DSS) for customer segmentation hinders the business ability to tailor strategies and offerings to diverse customer segments effectively.

In light of these challenges, we envision the implementation of a sophisticated BI system that not only consolidates and analyzes sales data from various

countries but also incorporates a robust DSS for customer segmentation. This initiative aligns with our commitment to fostering data-driven decision-making, enhancing operational efficiency, and ultimately elevating the overall competitiveness of the multi-countries platform organization.

The proposed BI system will not only address current shortcomings but also position the organization at the forefront of data-driven innovation. By leveraging advanced analytics and visualization tools, we aim to empower the teams with actionable insights, enabling them to make informed decisions that drive business growth. This initiative reflects our dedication to staying ahead in the market, meeting customer expectations, and ensuring sustained success in a rapidly evolving global business landscape.

2. Objectives

Our objectives may include:

- Analyzing sales data across multiple countries.
- Creating an interactive BI dashboard for real-time insights.
- Implementing a Decision Support System (DSS) for customer segmentation.
- Enhancing overall business intelligence for informed decision-making.

3. Description

3.1 Data Sources

For the data sources, we will use the dataset that is provided from the paper, "Online Retail Dataset II".

The dataset contains 2 years of sales data (Year 2009-2010 and Year 2010-2011) including total 1,067,371 samples and 8 features.

Features:

- Invoice
- StockCode

- Description
- Quantity
- InvoiceDate
- Price
- CustomerID
- Country

3.2. Methodology

3.2.1. BI dashboard

- Analyze historical data and define key metrics to answer 'what' and 'how' of past business operations and performance.
- Extract and prepare data for the key metrics from datasets.
- Develop the periodical reports and Key Performance dashboard with Tableau.

3.2.2 Customer Segmentation

According to Chongkolnee R. et. al. (2023)[†], the authors quantify customers' purchasing amount, purchasing frequency, and time since last purchase into quantiles and segment customers based on that. The issue is the heavily skewed nature of the data, which distorts the edge quantiles. Specifically, in heavily right-skewed distributions, the edge quantiles cover an excessively wide range, rendering their qualitative information practically meaningless. To fix this, we propose to make a scoring based on the standard deviation instead.

Our Customer Segmentation reference:

Segment ID	Segment Name	RFM Score (125)
1	Champion	555, 554, 544, 545, 454, 455, 445.
2	Loyal Customer	543, 444, 435, 355, 354, 345, 344, 335, 553, 551, 552, 541, 542, 533, 532, 531, 452, 451, 442, 441, 431, 453, 433, 432, 423, 353, 352, 351, 342, 341, 333, 323.
3	Need Attention	535, 534, 443, 434, 343, 334, 325, 324, 155, 154, 144, 214,215,115, 114, 113, 525, 524, 523, 522, 521, 515, 514, 513, 425,424, 413,414,415, 315, 314, 313, 331, 321, 312, 221, 213.
4	Losing	111, 112, 121, 131, 141, 151, 332, 322, 231, 241, 251, 233, 232, 223, 222, 132, 123, 122, 212, 211, 255, 254, 245, 244, 253, 252, 243, 242, 235, 234, 225, 224, 153, 152, 145, 143, 142, 135, 134, 133, 125, 124.
5	New Customer	512, 511, 422, 421 412, 411, 311.

https://connectif.ai/en/blog/what-are-rfm-scores-and-how-to-calculate-them/

3.2.3 End Users of the system

Target Audience:

Our end users encompass various stakeholders crucial for business development, including:

- Top Management Level:
    - Executives and senior leaders who oversee the overall strategic direction of the company.
    - They require insights from the dashboards to make informed decisions regarding resource allocation, goal setting, and strategic planning.
- Strategy Team:
    - Professionals responsible for developing and refining the company's overarching strategies.
    - They rely on the dashboards to identify trends, opportunities, and areas for improvement to align business strategies with market dynamics and customer needs.
- Sales Management Team:
    - Sales leaders and managers tasked with driving revenue growth and managing sales operations.
    - They utilize the dashboards to track sales performance, monitor customer interactions, and optimize sales strategies to meet revenue targets and enhance customer satisfaction.

3.2.4 Software for implementation

We will implement the following dashboards with the Tableau.

Dashboards:

- Sales Dashboard:
    - Provides insights into sales performance, revenue trends, product performance, and sales team effectiveness.
    - Enables users to track key sales metrics, monitor progress towards sales goals, and identify opportunities for revenue growth.
- Customer Dashboard:

- - Offers a comprehensive view of customer behavior and customer retention.
    - Helps users understand customer demographics, purchase patterns, and engagement metrics to tailor marketing strategies and improve customer retention.
  - Customer Segmentation Dashboard:
    - Facilitates segmentation of customers based on RFM (Recency, Frequency, Monetary) analysis.
    - Empowers users to identify high-value customers, prioritize marketing efforts, and personalize customer experiences based on segmentation insights.

These dashboards are designed to provide actionable insights to our target audience, enabling them to make data-driven decisions and drive business growth effectively.

4. Expected Outcomes:

Our BI system will offer periodic insights into customer behavior and sales performance featuring key metrics. Additionally, it will facilitate comparative analysis, allowing users to assess sales performance across various products, regions, or time periods. This dashboard aims to enhance operational efficiency by automating data analysis and facilitating quick access to critical information.

Concerning customer segmentation, our system will generate segmented groups applicable to targeted marketing campaigns, high-value customer focus, customer retention, winning back lost customers, and engaging average spenders. This strategic approach is expected to result in improved customer satisfaction, loyalty, and revenue growth.

5. Actual Result

Data Cleaning Process:

- We cleaned the dataset to ensure its applicability for our task. To achieve this, we applied the following criteria:
  - Each sale entry must have a quantity of 1 or more.

- - Customer IDs and invoice IDs must not be blank.
  - Stock codes must not be blank.
  - Prices must be greater than 0.
- After applying these criteria, the dataset was reduced from 1,067,371 to 805,602 entries, with no null values.

RFM Model Implementation:

- The RFM (Recency, Frequency, Monetary) model was utilized for analysis.
- Recency score: Calculated as the time difference, in days, between each purchase date and the end date of the quarter. This approach enables quarterly analysis.
- Revenue calculation: Determined by multiplying the quantity and price for each entry.
- Data grouping: Entries were grouped into quarterly time frames spanning from 2009Q4 to 2011Q4.
- For each quarter:
  - Recency score: Selected from the most recent order, identified by the lowest date difference.
  - Revenue score: Computed as the sum of revenue for each customer and quarter.
  - Frequency score: Derived from the count of distinct invoices received by each customer during the quarter.

	Customer ID	Quarter	Revenue	Invoice	Last Seen
0	12346	2009Q4	113.50	5	14
1	12346	2010Q1	117.05	9	30
2	12346	2010Q2	142.31	19	3
3	12346	2011Q1	77183.60	1	73
4	12347	2010Q4	2035.11	102	25
...	...	...	...	...	...
18021	18287	2010Q2	1071.61	54	45
18022	18287	2010Q3	892.60	21	10
18023	18287	2010Q4	381.50	10	40
18024	18287	2011Q2	765.28	29	40
18025	18287	2011Q4	1072.00	41	65

18026 rows × 5 columns

**RFM Scoring and Segmentation**:

- Scoring Methodology:
  - We assign discrete scores ranging from 1 to 5 for recency, frequency, and monetary value, based on even quantiles.
  - These scores are combined into a three-digit score, with each digit representing R, F, and M scores, respectively.
- Segmentation:
  - Among the possible 125 distinct scores, we assign them into 5 segments according to predefined groups:
    - Champion: High scores across all three dimensions (e.g., 555, 554).
    - Loyal Customer: High scores in some dimensions, indicating consistent engagement (e.g., 543, 444).
    - Need Attention: Moderate scores in certain dimensions, suggesting potential for improvement (e.g., 525, 434).
    - Losing: Low scores across multiple dimensions, indicating declining engagement (e.g., 331, 221).
    - New Customer: Initial engagement, reflected in specific score combinations (e.g., 512, 411).

- ○ Additionally, the 'lost' segment is included in the customer segmentation shift chart, representing customers who haven't made any purchases in the given quarter, potentially indicating churn between quarters.

5.1 Sales Performance Dashboard

The Sales Performance Dashboard provides comprehensive insights into the sales performance of the company over a selected period. This report covers various key performance indicators (KPIs) such as total sales amount, sales trend, sales by region, top 10 countries by sales amount, and top 10 products by sales amount. These metrics enable stakeholders to assess the company's sales performance effectively and make informed decisions.
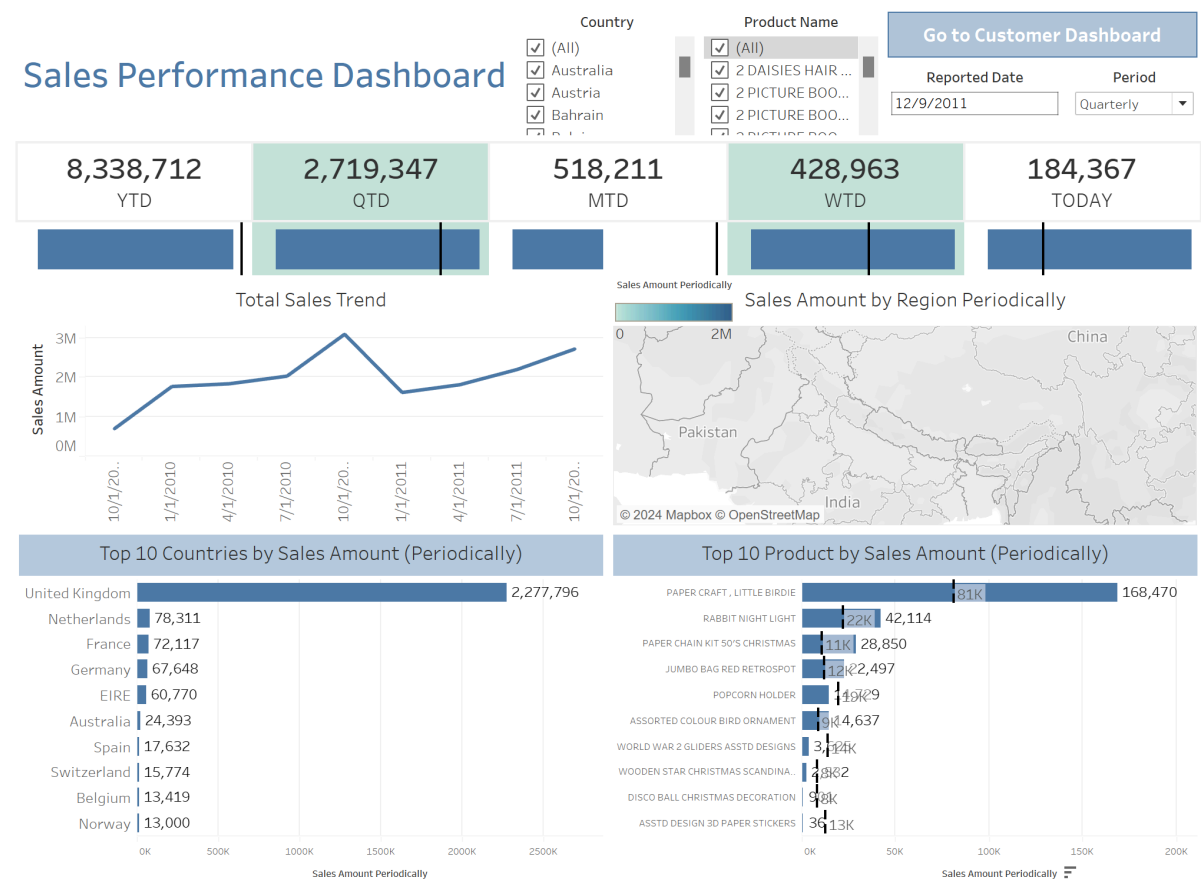
5.1.1 Total Sales Amount

8,338,712	2,719,347	1,161,817	248,713	41,481
YTD	QTD	MTD	WTD	TODAY
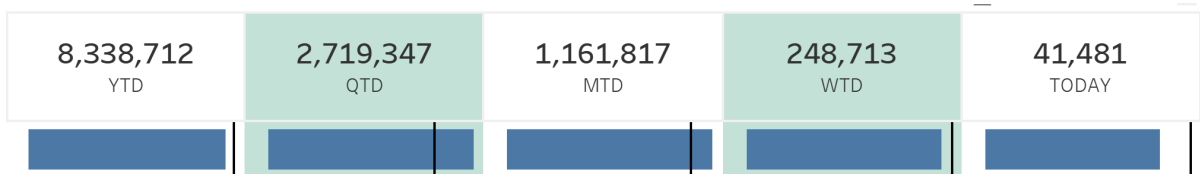
- This figure shows the Total Sales Amount for the selected period, country, and product according to the reported date.

| Country	Product Name	Go to Customer Dashboard
☑ (All)	☑ (All)	
☑ Australia	☑ 2 DAISIES HAIR ...	Reported Date / Period
☑ Austria	☑ 2 PICTURE BOO...	12/9/2011 / Quarterly ▼
☑ Bahrain	☑ 2 PICTURE BOO...	

- The total sales amount reflects the overall revenue generated by the company within the specified timeframe. It serves as a fundamental indicator of the company's financial health and growth trajectory.

5.1.2 Total Sales Trend

Total Sales Trend

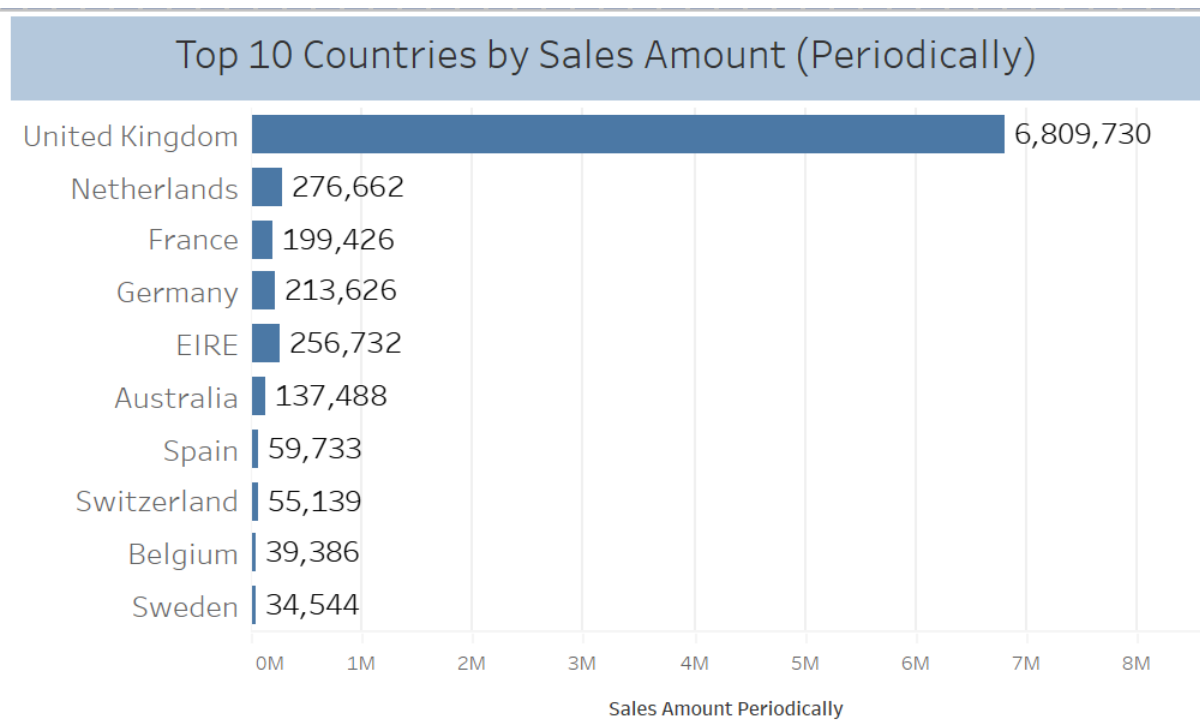Top 10 Countries by Sales Amount (Periodically)

Total Sales Trend
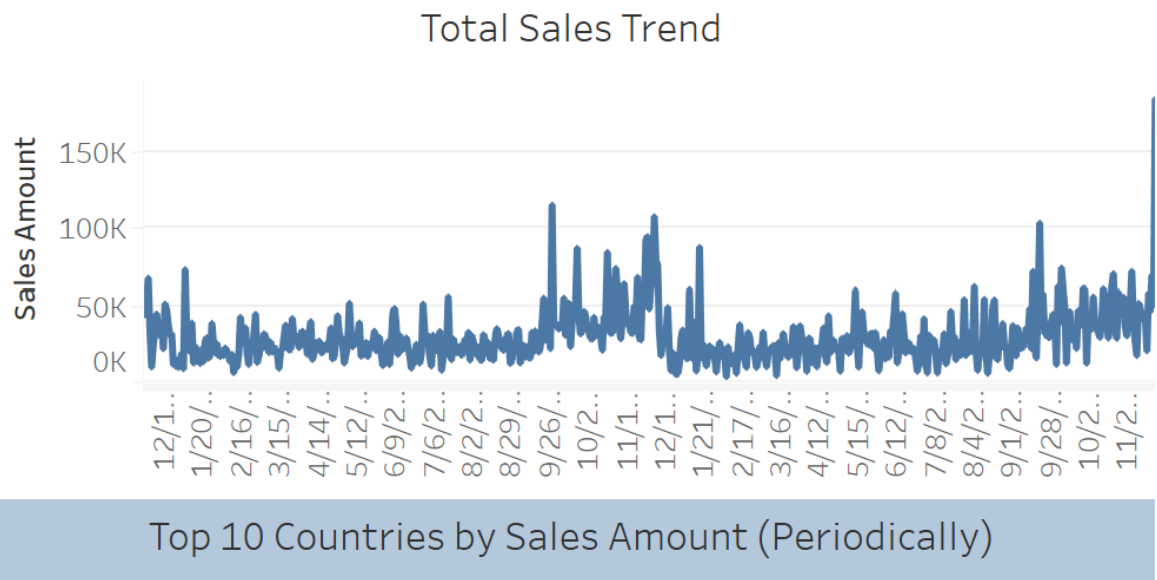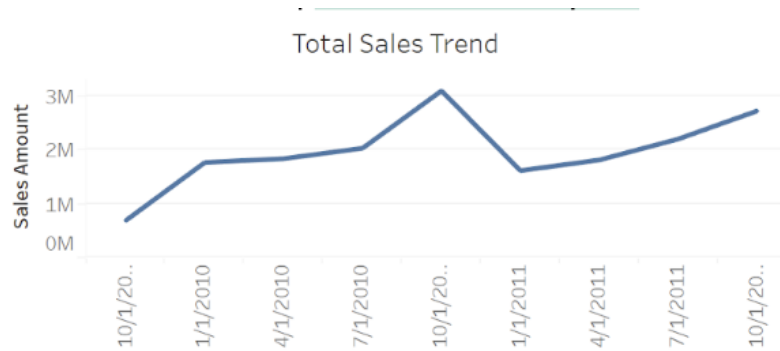
- The sales trend analysis provides insights into the trajectory of sales over the selected period. It helps identify patterns, fluctuations, and potential growth opportunities.
- The trend analysis indicates [increasing/stable/decreasing] sales over the selected period, suggesting [positive/negative] momentum in sales performance.

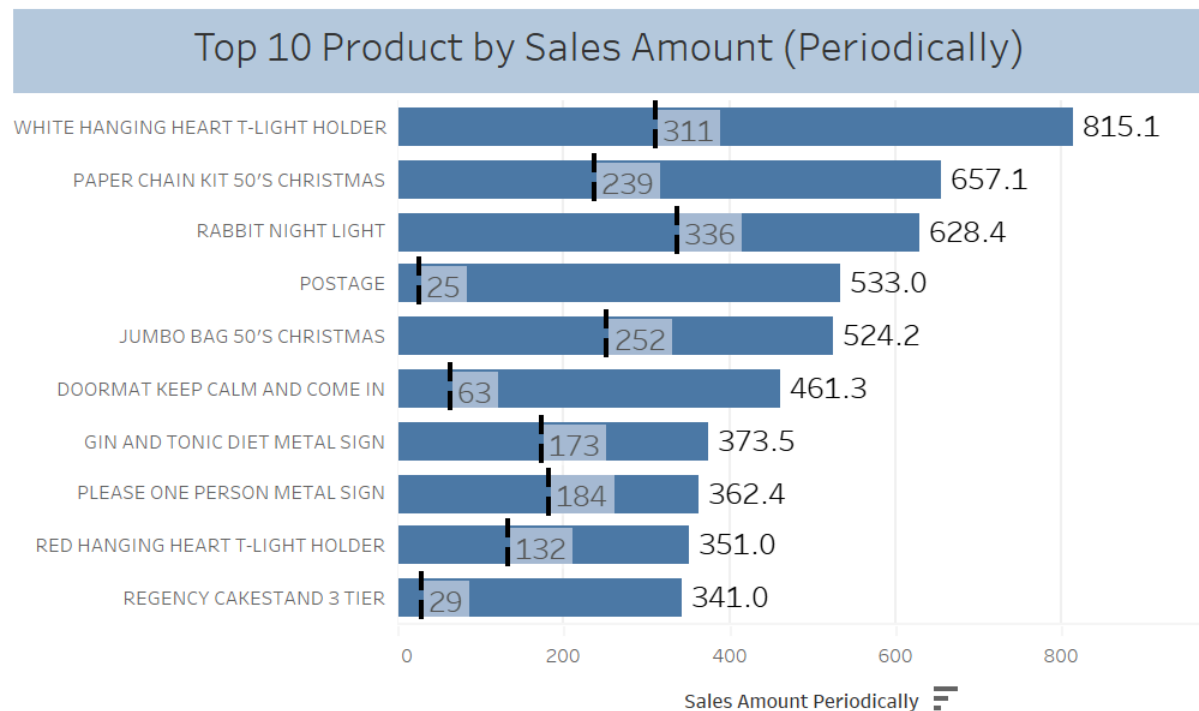### 5.1.3 Sales amount by Region Periodically

Sales Amount by Region Periodically
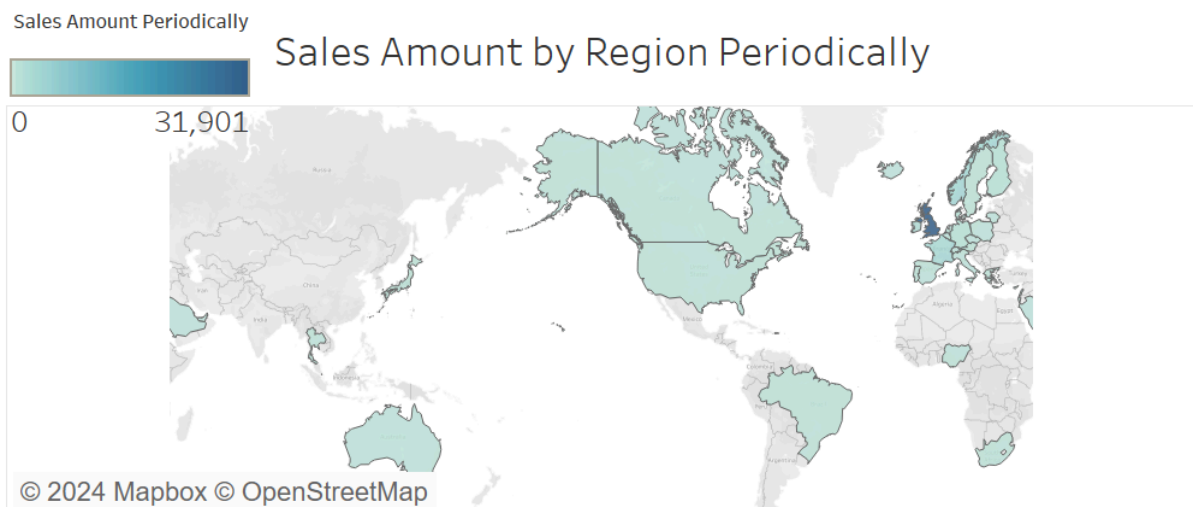
- Sales distribution across different regions provides valuable insights into geographic sales performance.

5.1.4 Top 10 Countries by Sales Amount (Periodically)

Top 10 Countries by Sales Amount (Periodically)

Country	Sales Amount Periodically
United Kingdom	6,809,730
Netherlands	276,662
France	199,426
Germany	213,626
EIRE	256,732
Australia	137,488
Spain	59,733
Switzerland	55,139
Belgium	39,386
Sweden	34,544

- Identifying top-performing countries in terms of sales amount helps prioritize markets and allocate resources efficiently.
- The interesting point is the United Kingdom is the highest one over all the periods.

5.1.5 Top 10 Products by Sales Amount:

Top 10 Product by Sales Amount (Periodically)

Product		
WHITE HANGING HEART T-LIGHT HOLDER	311	815.1
PAPER CHAIN KIT 50'S CHRISTMAS	239	657.1
RABBIT NIGHT LIGHT	336	628.4
POSTAGE	25	533.0
JUMBO BAG 50'S CHRISTMAS	252	524.2
DOORMAT KEEP CALM AND COME IN	63	461.3
GIN AND TONIC DIET METAL SIGN	173	373.5
PLEASE ONE PERSON METAL SIGN	184	362.4
RED HANGING HEART T-LIGHT HOLDER	132	351.0
REGENCY CAKESTAND 3 TIER	29	341.0

Sales Amount Periodically

- This figure offers insights into the products that contribute the most to sales, facilitating strategic product management and marketing efforts. By identifying top-performing products, we can prioritize resource allocation and refine marketing strategies to maximize revenue and profitability.

Conclusion:

The Sales Performance Dashboard provides a comprehensive overview of the company's sales performance, encompassing various KPIs such as total sales amount, sales trend, sales by region, top countries, and top products by sales amount. These insights empower stakeholders to make data-driven decisions, optimize sales strategies, and drive business growth. Continuous monitoring and analysis of these metrics will be crucial for maintaining a competitive edge in the market.

5.2 Customer Dashboard

The Customer KPI Dashboard offers a comprehensive overview of customer-related metrics, enabling stakeholders to assess customer engagement, retention, and regional distribution. This report covers key

performance indicators (KPIs) such as active user count, quarterly active vs. churn analysis, customer distribution by region, customer retention rate, and customer retention trend. These metrics provide valuable insights into the company's customer base dynamics and loyalty.
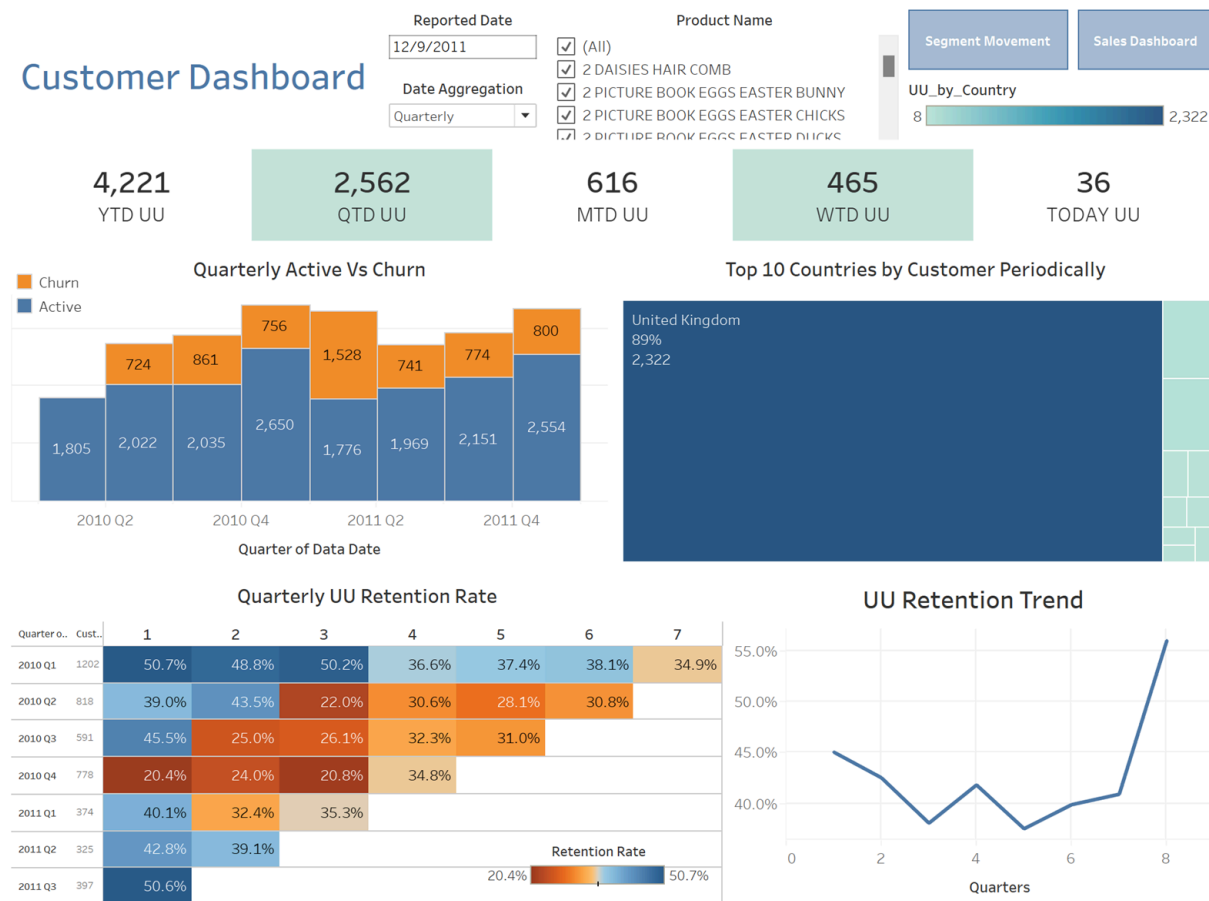
5.2.1 Active User Count:

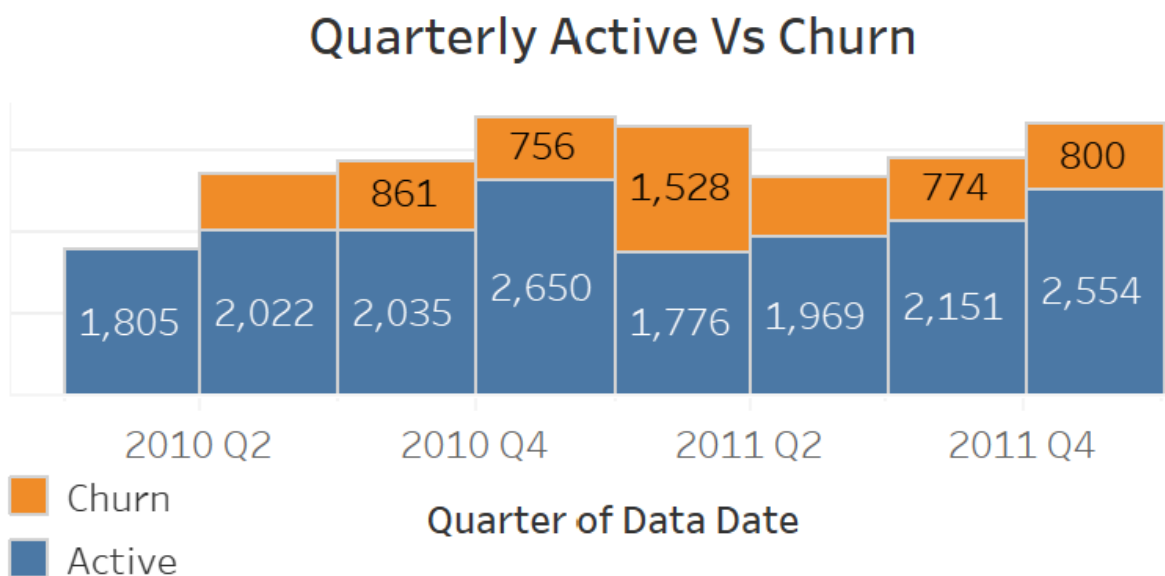- This figure shows the Total Active Users for the selected period according to the report date and Date Aggregation.

- Active User Count reflects the total number of users engaged with the company's products or services within the specified timeframe. It serves as a fundamental indicator of customer engagement and overall business performance.

5.2.2 Quarterly Active vs. Churn:

Quarterly Active Vs Churn

| | 756 | | 1,528 | | 774 | 800 |
| 1,805 | 2,022 | 2,035 | 2,650 | 1,776 | 1,969 | 2,151 | 2,554 |
861

2010 Q2   2010 Q4   2011 Q2   2011 Q4

■ Churn
■ Active

Quarter of Data Date

- Definition for Churn:
  - Our definition of churn is the customer who is active in a reported quarter and not active in the next coming quarter.
- The figure shows the Quarterly analysis of active users versus churned users provides insights into customer retention efforts.
- The comparison between active and churned users over the quarters helps identify trends and assess the effectiveness of retention strategies.
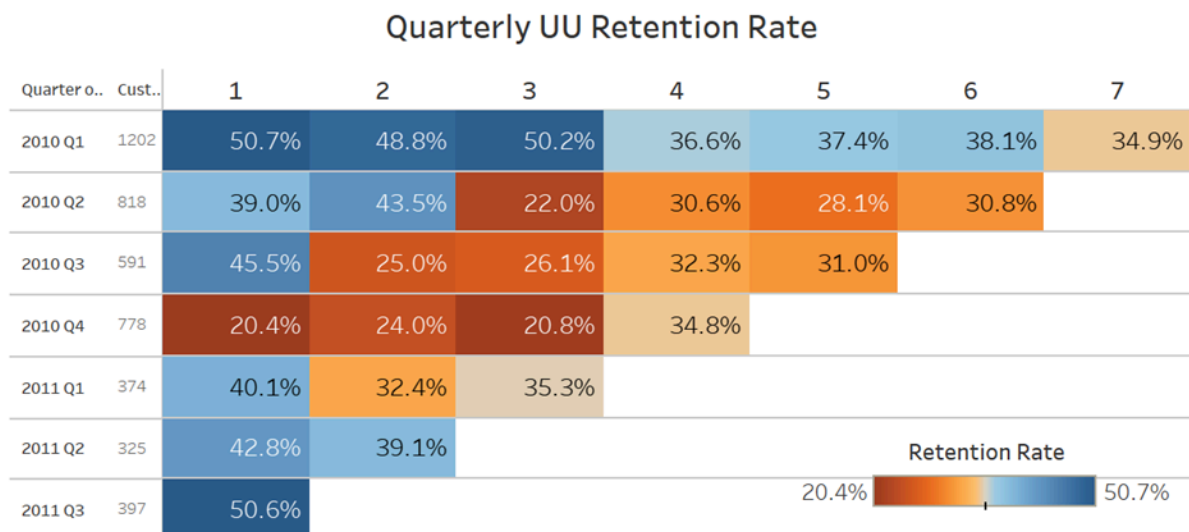
5.2.3 Top 10 Countries by Customer Periodically:

Top 10 Countries by Customer Periodically

- Understanding customer distribution across different regions facilitates targeted marketing and sales efforts. UK hold 89% of total customer.

5.2.4 Customer Retention Rate by Quarterly:

Quarterly UU Retention Rate

| Quarter o..	Cust..	1	2	3	4	5	6	7
2010 Q1	1202	50.7%	48.8%	50.2%	36.6%	37.4%	38.1%	34.9%
2010 Q2	818	39.0%	43.5%	22.0%	30.6%	28.1%	30.8%	
2010 Q3	591	45.5%	25.0%	26.1%	32.3%	31.0%		
2010 Q4	778	20.4%	24.0%	20.8%	34.8%			
2011 Q1	374	40.1%	32.4%	35.3%				
2011 Q2	325	42.8%	39.1%					
2011 Q3	397	50.6%						

Retention Rate

20.4% 50.7%

- Our retention rate is calculated based on the recency of the specific customers.
- The figure shows the Customer retention rate which indicates the percentage of customers retained over next quarter.

- A higher retention rate signifies better customer satisfaction and loyalty, contributing to long-term business success.
- The color of this graph indicates the level of safety to risks, ranging from blue to red

5.3.5 Customer Retention Trend:
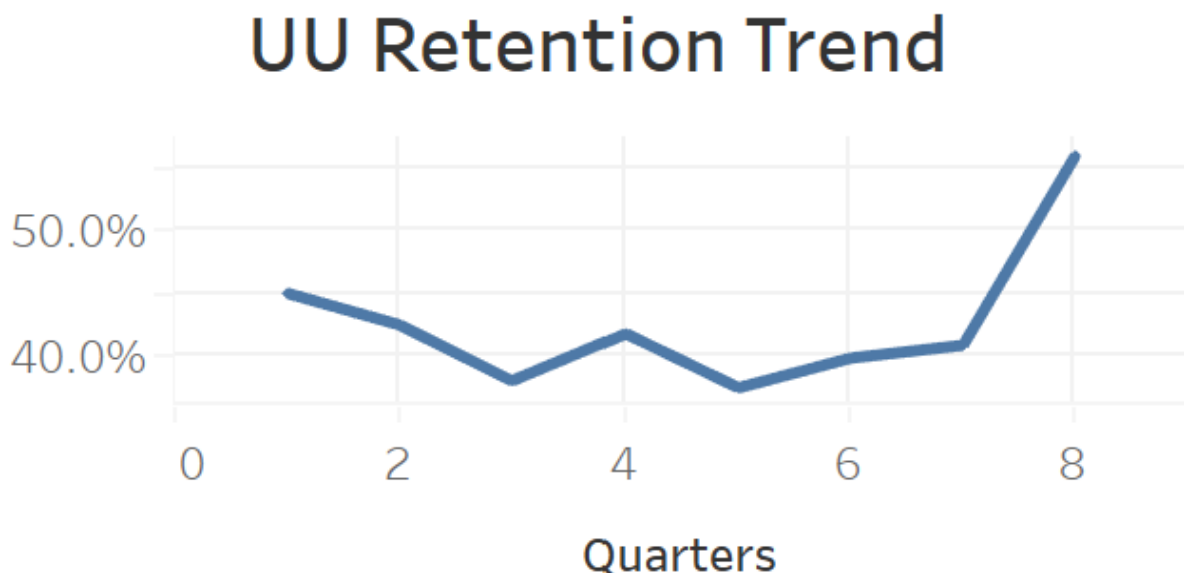

UU Retention Trend

- Analyzing the trend of customer retention over time helps identify patterns and evaluate the effectiveness of retention strategies.
- The retention trend indicates [increasing/stable/decreasing] retention rates over the selected period, suggesting [positive/negative] customer loyalty trends.

Conclusion:

The Customer KPI Dashboard provides a comprehensive view of customer engagement, retention, and regional distribution. Key metrics such as active user count, quarterly active vs. churn analysis, customer distribution by region, retention rate, and retention trend offer valuable insights for optimizing customer relationships and driving business growth. Continuous monitoring and analysis of these metrics are essential for maintaining customer satisfaction and fostering long-term customer loyalty.

5.3 Customer Segmentation Dashboard
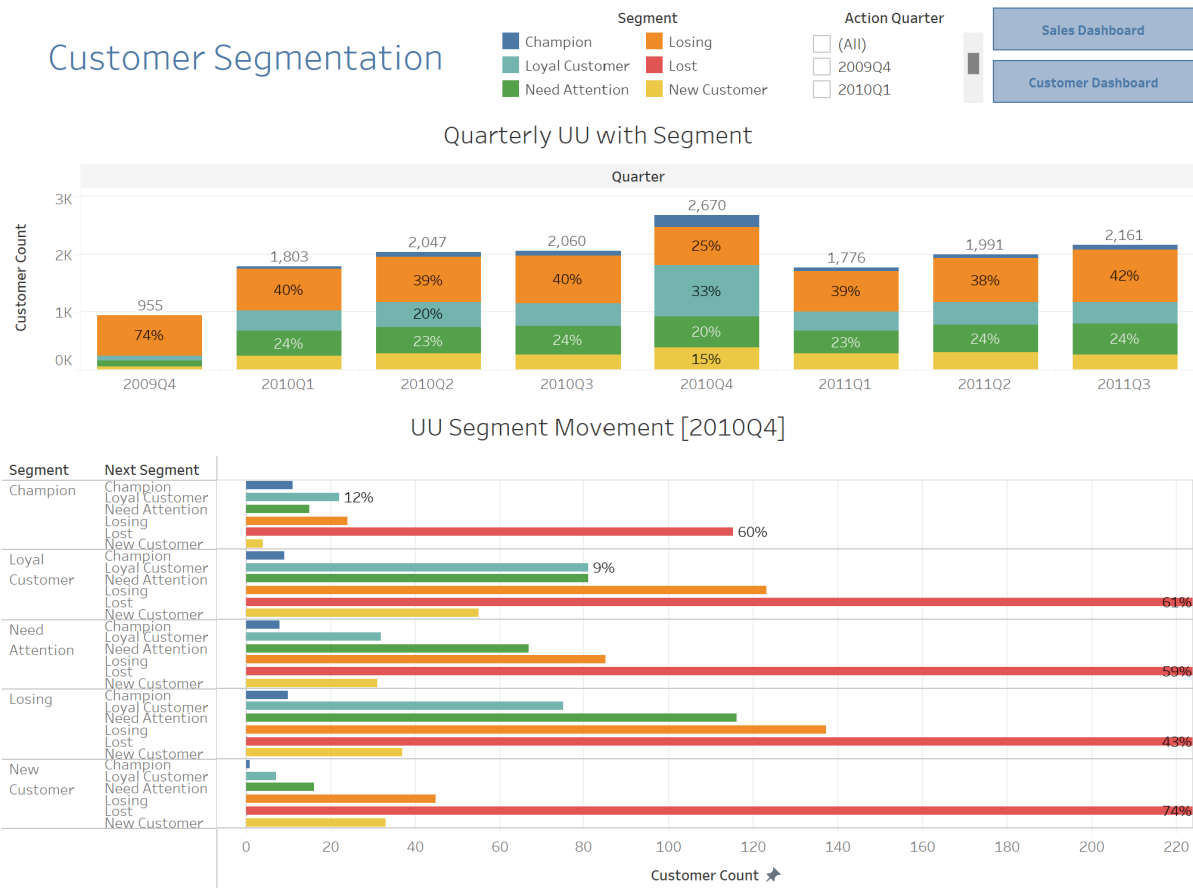
In our segmentation strategy, we categorize our customers into five distinct segments for each quarter, each represented by a unique color:

- Champion
- Loyal Customer
- Losing
- Need Attention
- New Customer

These segments enable us to tailor our marketing efforts effectively, ensuring that we address the specific needs and behaviors of each group.

This dashboard provides insights into how customer segments vary across different quarters. By analyzing the distribution of customer segments over time, we can track changes in customer behavior and identify trends that may impact our marketing strategies and business decisions.
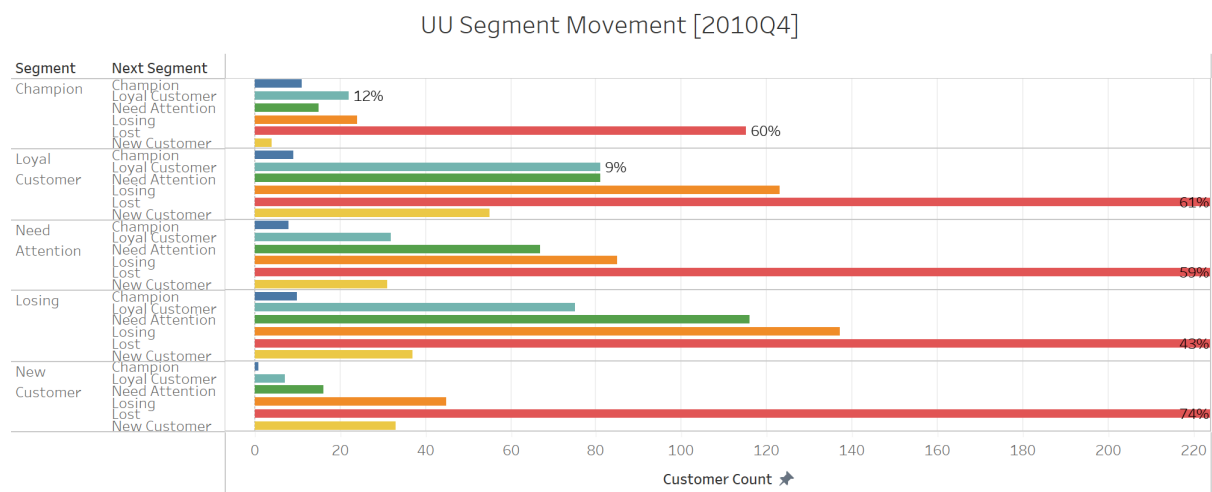
Segment

Champion   Losing

Loyal Customer   Lost

Need Attention   New Customer

5.3.1 Quarterly UU with Segment

Quarterly UU with Segment

UU Segment Movement [2009Q4]
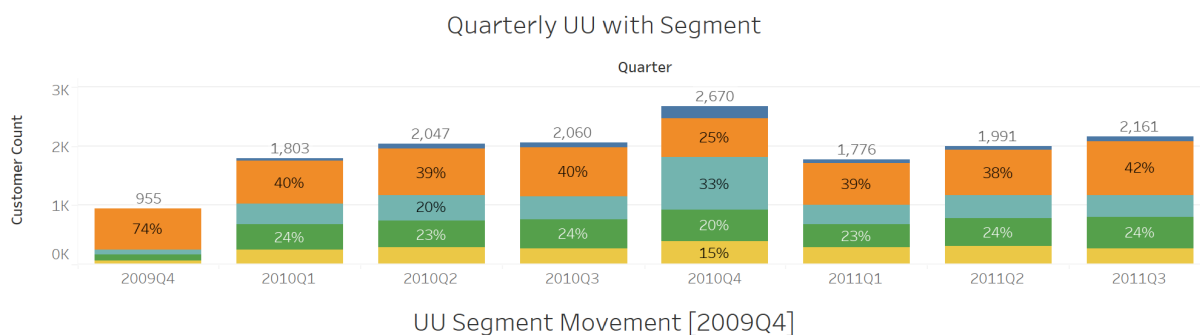
This figure illustrates the percentage distribution of each customer segment alongside the total number of users in each quarter. By presenting both the relative proportion of each segment and the absolute user count, we gain a comprehensive understanding of the composition of our customer base over time.

5.3.2 UU Segment Movement

UU Segment Movement [2010Q4]

This figure depicts the anticipated variations in each user segment for the next quarter. By forecasting the changes in segment distribution, we can proactively adapt our marketing strategies and customer engagement initiatives to address evolving trends and preferences.

6. Conclusion:

The proposed Business Intelligence System, with its focus on global sales analysis, an interactive BI dashboard, and a Decision Support System (DSS) for customer segmentation, is poised to revolutionize our approach to data-driven decision-making. By unraveling intricate patterns in sales data and providing real-time insights, the system promises to enhance agility and strategic responsiveness. The integration of a DSS for customer segmentation reflects our commitment to personalized strategies and precision in decision-making. This initiative signifies not only a technological advancement but a cultural shift towards a more data-centric organization. Anticipating transformative outcomes, we are poised to embrace a future where business intelligence becomes a strategic asset, empowering us to navigate the complexities of the global market with heightened foresight and efficiency.

Appendices:

- RFM model customer segmentation based on hierarchical approach using FCA [Chongkolnee Rungruang [a], Pakwan Riyapan [b], Arthit Intarasit [b], Khanchit Chuarkham [c], Jirapond Muangprathub]

- https://connectif.ai/en/blog/what-are-rfm-scores-and-how-to-calculate-them/