****

**SQL CAPSTONE PROJECT:**

**ANALYSIS OF CUSTOMER RELATIONSHIP MANAGEMENT (CRM) DATA**

**BY**

**MUHAMMAD ABDULMATEEN ADEIZA**

**DATE: Thursday, 27th of June, 2024**

**INTRODUCTION**

**BACKGROUND**

Customer Relationship Management (CRM) lies at the heart of modern businesses as well as their strategies, focusing on maintaining and optimizing relationships with customers. CRM systems are important in gathering, soring, and utilizing vast amount of customer data to enhance business operations, drive sales, and improve customer satisfaction.

CRM systems consolidate customer interactions and data across various channels, providing organizations with valuable insights into customer behavior, preferences, and trends. By analyzing CRM data, businesses can tailor their marketing strategies, personalize customer experiences, and identify opportunities for growth and retention.

SQL plays a very important role in extracting actionable insights from CRM data. It enables data analysts and business professionals to retrieve, manipulate, and analyze data stored in relational databases efficiently. Its wide querying capabilities allow for complex data aggregation, filtering, and transformation, empowering organizations to derive meaningful conclusions from large datasets.

In the field of business analytics, SQL serves as a foundational tool for querying and analyzing CRM data to uncover patterns, correlations, and trends. This enables businesses to make data driven decisions, optimize marketing campaigns, forecast sales trends, and enhance overall operational efficiency.

**PROJECT OVERVIEW**

This project makes use of SQL to analyze CRM data, demonstrating the practical application of SQL in extracting valuable insights that drive business decisions and strategies. By harnessing the CRM data provided by Techx through SQL, this analysis aims to illustrate how businesses can make use of data analytics to gain a competitive advantage in the dynamic market landscape.

**PROBLEM STATEMENT**

Techx, a medium scale firm that deals with the sales of tech products is facing challenges in accurately keeping track of their growth, as well as challenges in its transactional processes. These loses indicate inefficiencies within the processes potentially leading to revenue loss and diminished customer satisfaction. Identifying the root causes of the transactional losses and implementing strategies to mitigate these issues are critical to improving operational efficiency and enhancing overall business performance.

**OBJECTIVES**

1. Clean and transform CRM data
2. Create Dashboard in excel to track sales.
3. Identify factors contributing to low deal success rate.
4. Propose recommendations to improve the deal success rate.

**DATA TRANSFORMATION**

**DATA ACQUISITION AND DATABASE INITIALIZATION**

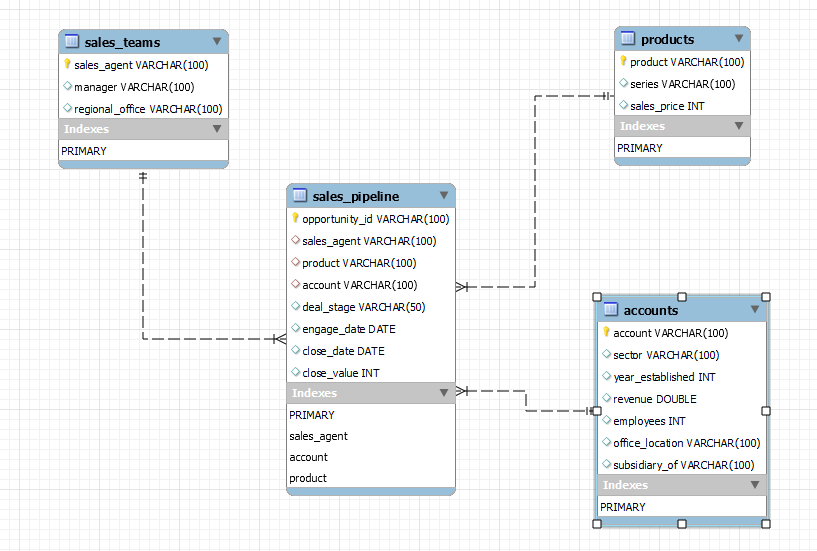
Initially, the CSV files containing the CRM data were gathered and subsequently loaded into a newly establish SQL database. This database served as the foundation platform for efficient data management and also facilitated seamless data querying and retrieval.

**DATA CLEANING AND STANDARDIZATION**

Upon importing the CSV files into the SQL database, meticulous data cleaning procedures were implemented. Each dataset underwent through scrutiny to rectify discrepancies, such as correcting typos, standardizing date formats for SQL compatibility, and ensuring overall data consistency across multiple datasets.

**SETTING UP DATABASE RELATIONSHIPS**

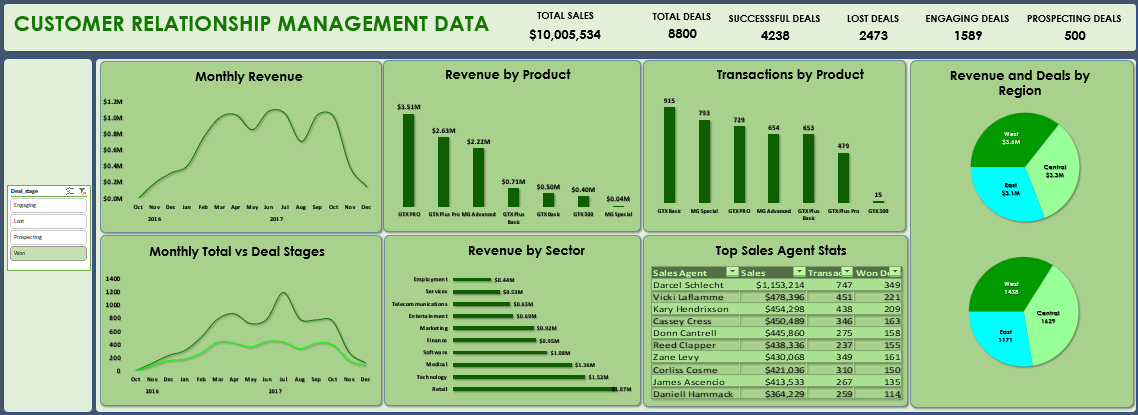
Following data cleaning and relational mapping, the cleaned datasets were merged into a centralized dataset within the SQL database. This merging process involved merging relevant data elements from separate tables into one, facilitating a comprehensive view of CRM data and enabling holistic analysis of business operations. The image below shows the model of the relationship in the CRM data. It follows a star schema arrangement with the sale pipeline acting as the facts table while the remaining are the dimension tables.



**CREATION OF DATA VISUALIZATION DASHBOARD**

To facilitate meaningful insights and decision-making, the centralized dataset was exported to Excel for the development of a dynamic dashboard. This dashboard served as a visual representation of key performance indicators (KPIs), trends, and actionable insights derived from the CRM data. It enabled stakeholders to interactively explore data patterns and trends, supporting informed decision-making and strategic planning initiatives.

The important KPIs tracked in the dashboard are the revenues, total transactions, and successful deals.

****

**EXPLORATORY DATA ANALYSIS**

**TRANSACTIONS AND REVENUE**

From October 2016 to December 2017, there have been a total of 8800 transactions yielding $10,005,534. It is observed that there was a steady increase in number of transactions and revenue from October 2016 to March 2017, and a rapid decline from October 2017 to December 2017. However, while the number of transactions peaked in July 2017 with over 1000 transactions, there are several peak periods of revenue with depressions in-between.

Overall, 48% of the transactions have been successful total just over 4200 deals with another 2473 lost deals. The rest of the deals are either still engaged or just in the prospecting stage,

**PRODUCTS**

There a total of 7 different products sold which are GTX PRO, GTX Plus Pro, GTX Basic, GTX Plus Basic, GTK 500, MG Advanced, MG Special.

Among them, the GTX Basic is the highest selling with 915 units sold while the GTK 500 is the lowest selling with only 15 units.

In terms of revenue generation, the GTX PRO has generated the most with over $3.5m while the MG Special has generated the lowest revenue od only about $40,000.

**REGION**

The office locations are spread across 3 regions: East, West, and Central each performing well in terms of revenue generation. Among them, the West region is the best performing with the highest revenue generating about $3.6m while the Central comes in second place with $3.3m. The East comes in last place with about $3.1m. However, when it comes to the number of transactions, there is a stark difference in their performances with the Central successfully completing 1629 deals and the West completing 1438 deals. The East still retains its last position with only 1171 successful deals.

**SECTOR**

Techx operates in different sectors ranging from retail, services and employment to technology, medical, and entertainment. This enables them to reach as many customers as possible. Among these sectors the ones in which the most revenues are generated are the retail, technology, and medical sectors, while the sectors with the least revenue are the employment, service, and telecommunication sectors.

The retail sector generated the most revenue of about $1.87m followed by the technology sector with about $1.52m. The least revenue comes from the employment sector with just over $400k.

**SALES AGENT**

Techx has over 30 sales agents spread across the 3 different regions with some performing very well and others grossly underperforming.

The best performing sales agent is Darcel Schlecht from the Central region with over 300 successful deals resulting in over $1min revenue for the company. Below is a table showing a list of the top performing sales agents:

****

**ANALYTICAL SOLUTION**

**PRODUCT SALES**

It is evident that some of the products have high supply but generate really low revenue while others have relatively low supply but generate the most revenue. This can be attributed to the price of the products. For instance, the GTX PRO generates the most revenue of $3,510,578 with its relatively high price and successful transactions while the MG Special is the second highest selling product but generates the lowest total revenue of $43,768 due to its very low price. The table below shows the products and their average price.

|  |  |  |  |
| --- | --- | --- | --- |
| **Product** | **Total Revenue** | **Successful Deals** | **Average Price** |
| **GTX PRO** | **$3,510,578** | **729** | **$4,815.61** |
| **GTX Plus Pro** | **$2,629,651** | **479** | **$5,489.88** |
| **MG Advanced** | **$2,216,387** | **654** | **$3,388.97** |
| **GTX Plus Basic** | **$705,275** | **653** | **$1,080.05** |
| **GTX Basic** | **$499,263** | **915** | **$545.64** |
| **GTK 500** | **$400,612** | **15** | **$26,707.47** |
| **MG Special** | **$43,768** | **793** | **$55.19** |

This illustrates how the prices have impacts on the distribution of the products.

**REGIONAL SALES**

The Central performed best with the most number of total and successful transactions. The East ranked in the last place with the lowest number of total and successful transactions. There is a 35% and 28% difference in the total and successful transactions respectively between the top Central region and the bottom East region. This indicates an underperformance in the East region compared to the other two regions. It is however important to note that when combined with the revenue generation data, the East region does not perform too poorly as expected indicating the possibility of large sales of high price products.

**SALES AGENTS**

It is observed that there is a wide disparity in the transaction volumes of the best and worst performing sales agents. The chart below indicates that the top sales agents are consistently engaged in multiple transactions monthly while the bottom sales agent show significantly fewer transactions over the same period. This also indicate a severe underperformance of some of the sales agent.

Also, there are 5 sales agents who have not engaged in any transactions at all.

**MANAGERS**

The data also reveals some differences in the performance of the managers particular of those from the East region. They have less transactions as well as overall revenue generated.

The table below ranked the managers according to their performances in both areas of transactions as well as revenue. The final rank is based on the weighted average of the rankings in the revenue and transactions putting more weight on revenue than number of transactions using the formula (0.6\*R) + (0.4\*T) where R is revenue and T is total transactions.

|  |  |  |  |
| --- | --- | --- | --- |
| **Manager** | **Revenue** | **Total Transactions** | **Weighted Average** |
| **Cara Losch** | **5** | **6** | **5.4** |
| **Celia Rouche** | **4** | **5** | **4.4** |
| **Dustin Brinkmann** | **6** | **3** | **4.8** |
| **Melvin Marxen** | **1** | **1** | **1** |
| **Rocco Neubert** | **3** | **4** | **3.4** |
| **Summer Sewald** | **2** | **2** | **2** |

It is observed that the best performing manager is Melvin Marxen while the worst performing manager is Cara Losch.

**CONCLUSIONS AND RECOMMENDATIONS**

**CONCLUSIONS**

1. The prices of the products are a huge factor in their demand and consequently revenue generation.
2. The Central region is performs very well compared to the other regions.
3. There are significant discrepancies in the performance of the sales agents.
4. There is a massive underperformance in the Employment and Services sectors.

**RECOMMENDATIONS**

1. The price of MG Special should be increased to capitalize on its high demand.
2. The price of the GTK 500 should be lowered to encourage customers to purchase them.
3. The least performing sales agents should either be incentivized to improve or replaced.
4. There should be an investigation into why some of the sales agents have not engaged in any deal.
5. There should be increased efforts in the East region to boost their performance.
6. There should be increased efforts in the Employment and Services sectors to boost their performance.