**Customer Segmentation and Sales Performance Analysis (Power BI & K-Means)**

**Project Overview**

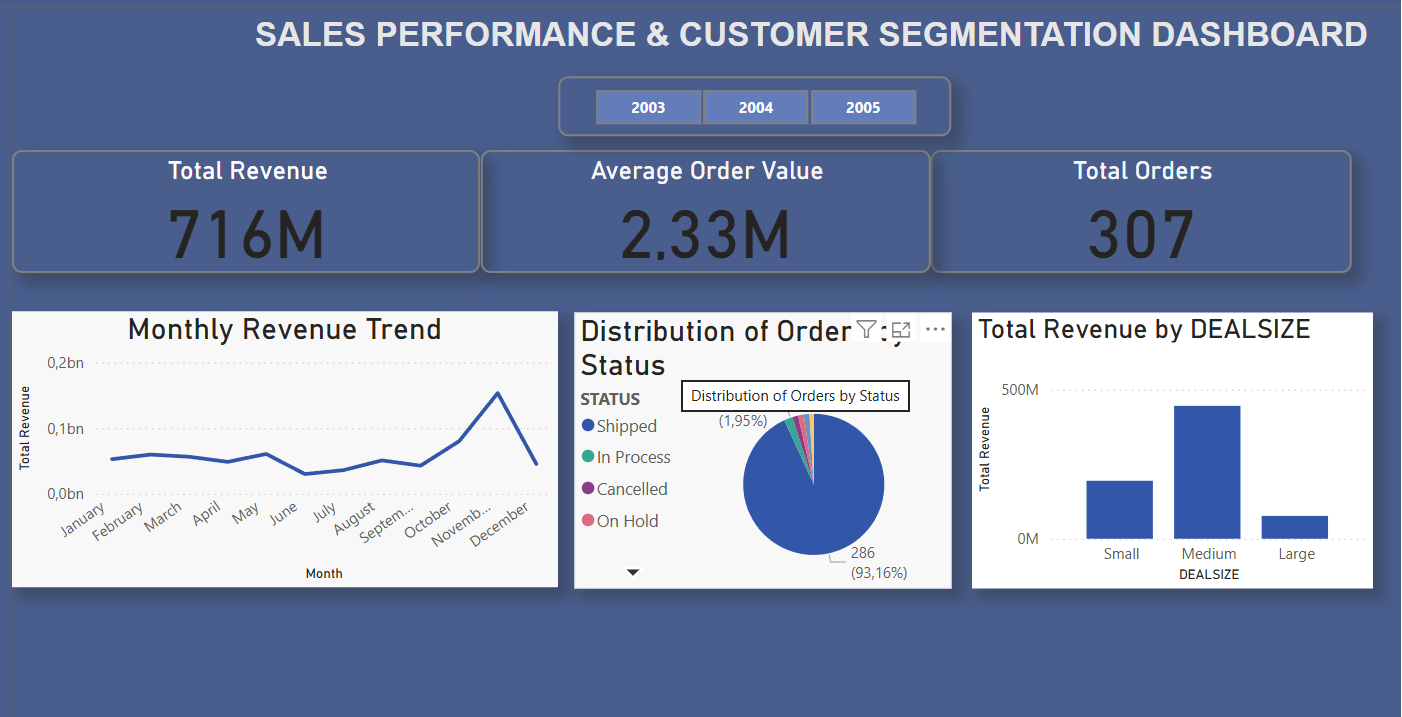
This project aimed to create a comprehensive Business Intelligence (BI) report using sales data from a retail company spanning the years 2003-2005. The core focus was to utilize the Machine Learning (K-Means) model for customer segmentation and develop concrete, **actionable marketing strategies** aimed at increasing company revenue.

**Technologies Used**

* **Data Processing & Modeling:** Python (Pandas, Scikit-learn, Matplotlib)
* **Data Visualization:** Power BI Desktop
* **Analytical Methodology:** RFM (Recency, Frequency, Monetary) Analysis
* **Machine Learning:** K-Means Clustering Algorithm

**📊 Power BI Dashboard Outputs**

1. **Executive Summary**

****

**Key Findings**

* **Financial Status:** Total Revenue was determined to be $716M with an Average Order Value (AOV) of $2.33M.
* **Revenue Dependency:** The company derives the majority of its total revenue from **Medium-sized** deals ($0.4Bn). This necessitates a focus on **Large Deals** to mitigate potential risk and diversify the revenue stream.
* **Anomaly and Trend:** A **sharp peak in revenue** was observed in December in the annual revenue trend. Analyzing the discount/sales volume balance during this peak is crucial for maximizing the profitability of next year's campaigns.

1. **Customer Segmentation (ML Segmentation)**

**metin, yazılım, bilgisayar simgesi, multimedya yazılımı içeren bir resim

Yapay zeka tarafından oluşturulmuş içerik yanlış olabilir.**

**Machine Learning Insights (ML Insights)**

Customers were divided into 4 main segments based on their RFM scores: Champions, Loyal, At Risk, and Lost.

|  |  |  |  |
| --- | --- | --- | --- |
| **Segment** | **Key Characteristic (RFM)** | **CRITICAL FINANCIAL IMPACT** | **Strategic Action Plan** |
| **At Risk** | High M, High R | **HIGHEST AVERAGE MONETARY VALUE.** Losing these customers represents the biggest financial risk. | **PRIORITY:** Highly personalized win-back offers and dedicated customer representative outreach to prevent churn. |
| **Champions** | Lowest R, Highest F | Most loyal and profitable customer base. | **RETENTION:** Reward programs, early access to new product launches, and exclusive perks to secure continued loyalty. |
| **Loyal** | Medium R,  High F | Provides a reliable, consistent revenue stream. | **GROWTH:** Use upsell and cross-sell offers to increase their Average Order Value (AOV). |
| **Lost** | Highest R, Lowest F, M | Low recovery probability, low potential value. | **EFFICIENCY:** Focus on automated, low-cost re-engagement efforts (e.g., surveys) instead of high-cost campaigns. |

**🔗 Project Links**

* **Power BI Report (PBIX File):** <https://github.com/canakikol/customer-segmentation-rfm-powerbi>
* **Python Notebook (RFM & K-Means) :**  <https://www.kaggle.com/code/canakikol/sales-data-analysis-comprehensive-eda-rfm>