Google Data Analysis Proposal

Insight Innovators team

Project Description

The Brazilian E-Commerce Dataset is a comprehensive dataset that captures transactional data from a Brazilian e-commerce platform. It provides valuable insights into customer behavior, product performance, and operational efficiency.

- It reflects real-world e-commerce operations in a growing market (Brazil).
- It covers multiple aspects of the e-commerce lifecycle, from orders to reviews.
- It is rich in features, allowing for diverse analyses and insights.

Team Leader

- **Youssef Mohamed Sayed**

Group Members & Roles

- **Data Cleaning and Preprocessing :**
  - Youssef Mohamed
  - Abdelrahman Ashraf
- **Data Visualization :**
  - Esraa Salama
  - Ganna Hassan
- **Data Exploration & preparing Documents & Presentation :**
  - Fatema Yasser
  - Aya Saeed

Objectives

- Analyze customer behavior to identify purchasing patterns and customer lifetime value (CLV).
- Evaluate product performance to identify the best-selling products and underperforming categories.
- Assess operational efficiency, including order fulfillment times and delivery performance.
- Perform sentiment analysis on customer reviews to gauge satisfaction levels.
- Examine regional sales trends to understand market demand variations across Brazil.
- Identify the impact of discounts and promotions on sales performance.
- Assess shipping and logistics efficiency to improve delivery times.
- Understand the relationship between product categories and customer demographics.
- **Sales Performance Analysis:**
  - Calculate total revenue from successfully delivered orders (price + freight value).
  - Determine expected revenue from all approved orders.
  - Analyze the number of canceled orders.
  - Measure late deliveries where actual delivery exceeds estimated date.

- **Payment Accuracy Analysis:**
  - Compare actual payments received with expected revenue to check accuracy.
- **Monthly Financial Dashboard Requirements:**
  - **Financial Overview:**
    - Total revenue from delivered orders.
    - Total payments received.
    - Expected revenue from approved orders.
    - Percentage of revenue reconciliation (actual vs. expected).
  - **Order Breakdown:**
    - Total orders by month.
    - Percentage of canceled, pending, and delivered orders.
    - Average order value.
  - **Delivery Insights:**
    - Count of late deliveries by month.
    - Average delay in days.
  - **Interactive Features:**
    - Filters by order status, payment type, and date range.
    - Drill-down to order-level details.
- the relationship between product categories and customer demographics.

Tools & Technologies

- **Programming Languages:** Python.
- **Data Visualization:** Tableau, Power BI.
- **Collaboration Tools:** GitHub.
- **Analysis Tools:** Python, Excel, SQL.

Milestones & Deadlines

1. **Week 1 (March 14 - March 20, 2025)**
   - Data Collection & Understanding
   - Explore dataset structure and identify missing values
   - Define key metrics for analysis
2. **Week 2 (March 21 - March 27, 2025)**
   - Data Cleaning & Preprocessing
   - Handle missing and inconsistent data
   - Ensure data is well-structured for analysis
3. **Week 3 (March 28 - April 3, 2025)**
   - Exploratory Data Analysis (EDA) & Visualization
   - Generate statistical summaries and visual insights
   - Identify trends in customer behavior and product performance
4. **Week 4 (April 4 - April 10, 2025)**
   - Advanced Analysis & Modeling
   - Perform sentiment analysis on customer reviews
   - Analyze operational efficiency and forecast trends
5. **Final Week (April 11 - April 14, 2025)**
   - Dashboard & Report Preparation
   - Create final visualizations and insights using Tableau/Power BI
   - Prepare project presentation and documentation

References:

- project on GitHub
- Dataset On Kaggle

THE PROJECT IS BEING PROCESSED.....

By: Insight Innovators team