Internship Assignment Report

Introduction

This report summarizes the completion of two analytical tasks assigned by Null Class. These tasks involved using data analytics techniques to investigate sales Analysis and customer Order behaviour.

Background

1. **Sales Analysis** – To investigate the sales performance of the "Others" category of products between 2021 and 2022.

2. **Customer Order Behaviour** –To investigate how long it takes for customers to place their first order after registration.

Learning Objectives

- Analysis and Compare datasets.
- Learn how to calculate metrics such as percentage change and average customer delay.
- Enhance data visualization skills.
- Improve data cleaning, aggregation and merging.
- Understand customer behaviour.

Activities and Tasks

For Task 1- Sales Decrease Analysis:

1. Filter sales data by category=" Others" and is_valid=1 for year 2021 and 2022.
2. Group by sku_name and calculate total qty_ordered per year.
3. Merge 2021 and 2022 databases by sku_name.
4. Calculate:
    - difference = qty_ordered – qty_ordered
    - Percentage= ((qty_ordered – qty_ordered)*100 / qty_ordered)
5. Classify the result:
    - "Down" if < 0
    - "Up" if > 0
    - "Fair" Otherwise
6. Sort by largest negative difference and display top 20 products.
7. Visualize results using a horizontal bar chart.

For Task 2 – First Order Timing Analysis:

1. For each customer, identify the earliest order_date after their registered_date.

2. Use the Calculated field:

    - Date_diff(order_date, registered_date, Day)

3. Create a scorecard showing the average number of days between registration and first order.

4. Generate a histogram to visualize the distribution of days taken across customers

Skills and Competencies

- Data Cleaning and Filtering.

- Aggregation and Grouping.

- Merging and Joining Datasets.

- Time-based calculations.

- Visualization techniques (bar chart, scorecard, histogram).

- Analytical Thinking.

Feedback and Evidence

- Merged dataset showing year changes.

- List of top 20 decreasing products.

- Horizontal Bar Chart displaying 20 decreasing products.

- Calculated field days_to_first_order.

- Average via Scorecard.

- Histogram showing spread of time intervals.

Challenges and Solutions

- Work independently without mentor support.

- Data accuracy and merging discrepancies.

Outcomes and Impact

- Improve self-research and troubleshooting skills.

- Develop strong visualization techniques.

- Gained hands-on-experience in analytics tools.

Conclusion

First, we discovered the sales dropped in 2022 compared to 2021.We also identified the top 20 products with the decline sales and showed them clearly in a bar chart. Second, we analysed how long customers take to purchase their first order after registering. Then create a scorecard to show average number of days for customers to place their first order and visualize it through histogram.