Data Analytics Internship Project Report

Author: Aditya Soni
Date: 13 December 2024

1. Introduction

This project involves analyzing an Excel file containing data across three sheets: **UserDetails**, **CookingSessions**, and **OrderDetails**. The aim of the project is to identify key insights into customer behavior, such as popular dishes, the relationship between cooking sessions and orders, demographic preferences, and trends over time. The analysis helps to uncover patterns that can drive business strategies for improving sales, marketing, and inventory management.

2. Data Cleaning and Preprocessing

2.1 Handling Missing Data

The initial datasets contained missing values. Here's how they were handled:

- For **UserDetails**, missing values were forward-filled.
- In **CookingSessions**, rows with missing `dish_name` values were dropped.
- The **OrderDetails** dataset was processed, with `order_date` being converted to a `datetime` format.

2.2 Duplicate Removal

Duplicates were identified and removed from all datasets, ensuring the integrity of the analysis.

2.3 Data Transformation

- `order_date` in **OrderDetails** was converted to the appropriate `datetime` type to facilitate time-based analysis.

3. Data Analysis

3.1 Popular Dishes

By analyzing the **dish_name** column from the merged dataset, the top 10 most popular dishes were identified based on the number of orders. This insight can help businesses focus on promoting high-demand dishes and streamline their menu offerings.

3.2 Cooking Sessions and Orders

A correlation analysis was performed to understand the relationship between **cooking_time** and **order_quantity**. This helped identify whether longer cooking times are associated with larger or smaller order quantities, which can be useful for operational decisions.

3.3 Demographic Analysis

The **UserDetails** dataset was segmented by **age_group**, and the number of orders per dish was aggregated. This analysis revealed:

- Certain age groups prefer specific dishes, offering opportunities for targeted marketing and personalized promotions.

3.4 Time-Based Trends

By aggregating the data based on the **order_date**, monthly trends in the number of orders were examined. This analysis identified:

- High-demand periods that could be used for planning marketing campaigns or inventory adjustments.

4. Key Findings

- **Top Popular Dishes**: The top 10 dishes were identified, which provide actionable insights for inventory and marketing efforts.
- **Demographic Preferences**: Age groups show distinct dish preferences, enabling personalized promotional strategies.
- **Monthly Trends**: The business can optimize promotional activities and inventory management by focusing on peak order months.

5. Business Recommendations

- **Dish Promotions**: Focus on promoting popular dishes to maximize sales. Special offers on high-demand dishes can also improve customer engagement.
- **Targeted Marketing**: Tailor marketing campaigns to specific age groups based on dish preferences to enhance conversion rates.
- **Seasonal and Monthly Campaigns**: Based on the time-based trends, create seasonal marketing campaigns and adjust inventory to align with peak demand months.

6. Conclusion

This analysis provided valuable insights into dish preferences, customer demographics, and time-based order trends. By focusing on popular dishes, leveraging demographic preferences, and adjusting to seasonal trends, businesses can improve their sales strategies and operational efficiency. The findings and recommendations outlined in this report are intended to help businesses refine their marketing strategies and enhance customer satisfaction.