Exploratory Data Analysis (EDA) on Food Service Data

1. Introduction

This project performs an Exploratory Data Analysis (EDA) on a food service dataset to uncover patterns related to food waste, staffing levels, and environmental factors. The objective is to extract insights that can help optimize kitchen operations and reduce unnecessary food waste.

The dataset includes columns such as meals served, kitchen staff, temperature, humidity, food waste levels, special event indicators, and staff experience.

2. Data Cleaning

The following steps were taken during data cleaning:

- - Missing values in categorical columns ('staff_experience' and 'waste_category') were filled using the mode.

- - The 'date' column was converted to datetime format.

- - Duplicate rows were identified and removed.

- Data types were verified for consistency.

3. Exploratory Data Analysis (EDA)

- - Summary statistics were calculated for numerical columns such as meals_served, temperature_C, humidity_percent, and food_wasted.

- - Countplots and bar charts were used to visualize categorical variables like staff_experience and waste_category.

- - Line plots were used to show trends in food served and food wasted over time.

- - Boxplots and histograms were used to assess distributions and detect outliers.

4. Correlation Analysis

A correlation matrix was created for numeric variables using a heatmap. It helped identify relationships such as:
- Positive correlation between number of meals served and food waste.
- Weak or negligible correlation between humidity and waste.
- Some correlation between number of staff and waste levels.

5. Hypothesis Testing

Two key hypothesis tests were conducted:

- - H0: There is no significant difference in food waste between staff with '<1 year' and '>5 years' experience.

- - H1: Staff experience significantly affects food waste.

A t-test was performed. The p-value was below 0.05, indicating that staff experience significantly influences food waste.

6. Key Insights and Recommendations

- - Staff with more experience tend to waste less food.

- - Food waste peaks during certain days, possibly linked to special events or environmental factors.

- - Dairy and meat categories show the highest waste rates.

Recommendations include:

- - Provide additional training for new staff.

- - Improve waste forecasting for event days.

- - Focus waste reduction strategies on high-waste food categories.

7. Conclusion

The EDA provided useful insights into operational inefficiencies and food waste patterns. Staff experience and special events were found to have a measurable impact on waste levels. Future work could involve predictive modeling to forecast food waste more accurately and evaluating the effect of implemented changes over time.

8. Appendix

Please refer to the google colab Notebook (`ProjectEDA`) for all charts, code, and intermediate steps.