**Report**:

* bakery sale data is between ‘01-01-2021’ and ‘09-30-2022’.
* weather dataset is between 01-01-2021and 10-31-2022.

**Data Cleaning:**

**Handling Missing Values:**

* In the weather dataset removed columns that were almost empty and with no data ('tsun' and 'snow').
* Imputed missing values using iterative imputation with a RandomForestRegressor for wind direction, speed, gust, and pressure.

**Anomaly Removal**:

* From the bakery dataset rows with non-informative entries ('.') in the Article column were removed.
* Rows with zero or negative quantities and unit prices were also filtered out.

**Feature Engineering:**

* Added features in the weather dataset such as ‘is\_holiday’,'is\_rainy’, ‘temp\_range', 'is\_hot\_day', 'is\_heavy\_wind', 'is\_high\_pressure', 'day\_of\_week', and 'day\_name'.
* Feature selection based on correlation analysis (Pearson) and feature importance from two models helped identify key features for forecasting.

**Modeling:**

* The model achieved a Mean Absolute Error (MAE) of approximately 96 units.
* Recommendations for enhancing predictions include integrating more features, optimizing model hyperparameters, and exploring advanced ensemble techniques.

**Visualization:**

All the top 10 highest sales days occurred on weekends or holidays, predominantly on Sundays. This highlights the significant impact of weekends on bakery sales. For instance, on August 14, 2022, a Sunday, the bakery recorded 1970.0 items sold despite experiencing minimal to moderate rainfall (7.6 mm) and a wind speed of 12.4 km/h. The following day being a holiday further diminished the influence of weather on sales. Notably, moderate wind speeds were typical on high sales days; for instance, on July 16, 2022, which recorded a wind speed of 17.2 km/h and a temperature of 30.0°C, resulting in €1,535 in sales.Furthermore, high sales were consistently observed on days with temperatures exceeding 30°C, indicating a positive correlation between warm weather and bakery visits.

A graph showing a number of sales

Description automatically generated with medium confidence

The following plot highlights customer preferences for various bakery products. Focusing on these top-selling items enables us to optimize production, reduce waste, and meet customer demand more effectively. TRADITIONAL BAGUETTE is the most popular item, with a total sales quantity of 118,093 units. Other items like CEREAL BAGUETTE and SPECIAL BREAD, although lower in sales, still in the top 10, indicating a significant but specialized market.

A graph showing a number of sales

Description automatically generated with medium confidence

The following plot reveals that Sundays consistently have the highest average sales, with an average of 919.71, making it the busiest day for the bakery. Weekdays (Monday to Friday) show lower average sales ranging from 480.89 units (Tuesday) to 545.55 units (Monday). However, it's noteworthy that weekdays during holidays, especially Mondays like on August 15, 2022, may not follow the usual weekday trends.

A graph of a number of blue boxes

Description automatically generated

**Insights:**

* High sales days were predominantly weekends, particularly Sundays, highlighting their significance for bakery sales.
* Holidays showed potential to boost sales, although they were less frequent among the top sales days.
* Analysis of weather conditions on high sales days revealed moderate wind speeds and high temperatures as common factors.

**Limitation:**

1. 'tsun' feature had no data available.
2. 'snow' feature had only 15 rows of data, limiting its utility.
3. Data gaps were observed in the 'date' feature, with approximately 37 missing dates within the provided period.