Correlation Between Weather Conditions and Yield/Customer Complaints in Rendering Plants

Surena Nokham

Department of Science & Technology, Bellevue University

DSC 680-T301: Applied Data Science

Amirfarrokh Iranitalab

June 9, 2024

Topic:

Correlation Between Weather Conditions and Yield/Customer Complaints in Rendering Plants

Business Problem:

This project aims to determine if there is a significant correlation between weather conditions and the complaints regarding decreased yield/volume in rendering plants. By analyzing the correlation between weather data and production metrics, we can identify if specific weather conditions, such as precipitation or temperature fluctuations, are linked to increased customer complaints and reduced yield/volume. Establishing these correlations will help address customer concerns with evidence-based explanations and improve operational decisions by mitigating the impact of adverse weather conditions.

Datasets:

1. **LIMS Labware Data:** Contains sample data with information about moisture, fat retainment, protein values, and other relevant measurements.
    - **Columns:** SAMPLE_ID, SAMPLED_DATE, SAMPLING_TIME, PRODUCT, SAMPLE_TYPE, CUSTOMER, TESTING_FACILITY, ANALYSIS, NAME, NUMERIC_ENTRY, DISPLAY_STRING, CUSTOMER_NAME, DESCRIPTION.
2. **SAP Production Data:** Details the production pounds, including posting dates, pounds produced, and material numbers.
    - **Columns:** Plant, Posting Date, Pounds, Material Number.
3. **ICP Complaints Data:** Lists customer complaints/rejections, customer information, and complaint dates.

- **Columns:** Complaint Type, Complaint Date, Customer, Producing Plant, Material, Customer PO, Incident Type, Status, Incident Decision, Total Rejection Cost, Comments.

4. **NOAA Weather Data:** Includes weather data such as daily precipitation, monthly precipitation, snow, maximum temperature, minimum temperature, and temperature observations.
    - **Columns:** STATION, NAME, LATITUDE, LONGITUDE, ELEVATION, DATE, DAPR, MDPR, PRCP, SNOW, SNWD, TMAX, TMIN, TOBS.

Methods:

1. **Exploratory Data Analysis (EDA):** Initial examination of datasets to understand the distributions and relationships.
2. **Correlation Analysis:** Identify and quantify the potential correlations between weather variables and yield/volume or complaint occurrences.
3. **Regression Analysis:** Develop regression models to predict yield/volume changes and complaint occurrences based on weather conditions.
4. **Time Series Analysis:** Examine temporal patterns in weather data and their impact on yield/volume over time.
5. **Visualization:** Create clear and informative visualizations to effectively communicate findings and correlations.

Ethical Considerations:

1. **Data Privacy:** Ensure that any customer data used is anonymized to protect privacy.
2. **Bias and Fairness:** Avoid biased conclusions by considering all potential variables and ensuring a comprehensive analysis.

3. **Data Accuracy:** Ensure that the data from all sources, especially weather data, is accurate and reliable.

Challenges/Issues:

1. **Data Integration:** Merging data from different sources (LIMS, SAP, ICP, NOAA) may be challenging due to different formats and structures.
2. **Missing Data:** Handling missing or incomplete data entries, especially in weather records.
3. **Complex Relationships:** Identifying and modeling complex relationships between weather conditions and yield/volume changes.
4. **Temporal Alignment:** Ensuring that the timing of weather data matches the production and complaint data accurately

References

- National Oceanic and Atmospheric Administration (NOAA) for weather data.
- Internal documentation and records for LIMS, SAP, and ICP data.
- Relevant academic literature on the impact of weather on agricultural and food production.