



# **ABNORMALITY DETECTION ANALYSIS**

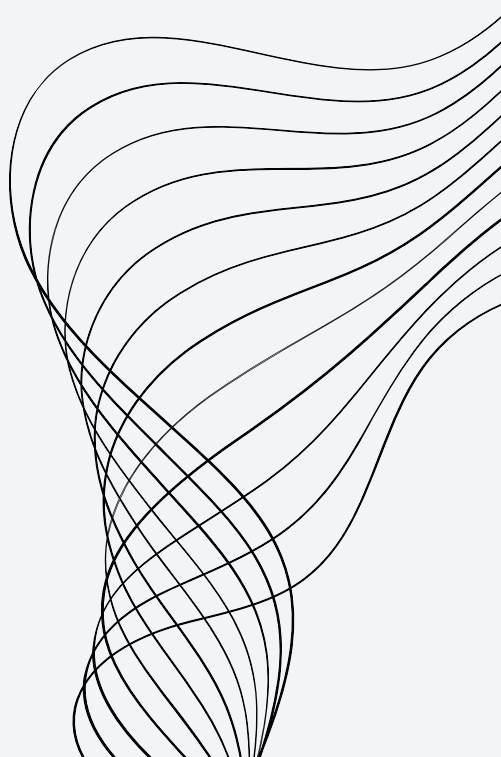
**GAURAV P LUTE**





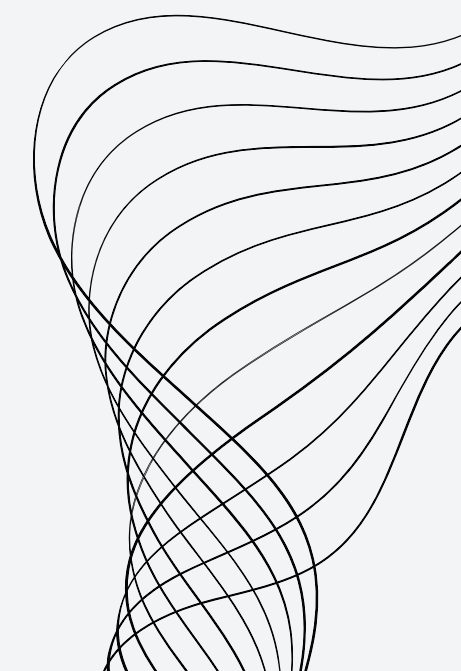
# OBJECTIVE

Using python and any algorithm of your choice, highlight time periods where this abnormality can be observed.





# DATA PREPARATION

- **Raw Data Overview:**
    - Source: Dataset containing 6 features.
  - **Preprocessing Steps:**
    - Handling datatypes
    - String convert to float
  - **Descriptive Statistics:**
    - Performed summary statistics to understand data distribution.
    - Boxplot Analysis: Identified outliers visually.
  - **QQ Plot:** Verified data normality assumptions.
  - **Feature Engineering:**
    - To create new feature like, Month, year, etc
    - Standardization: Applied StandardScaler to normalize features for consistency.
- 



# ANALYSIS STRATEGY

- **METHODOLOGY:**

- Trained Isolation Forest to identify global anomalies.
- Applied Elliptic Envelope for local and robust anomaly detection.

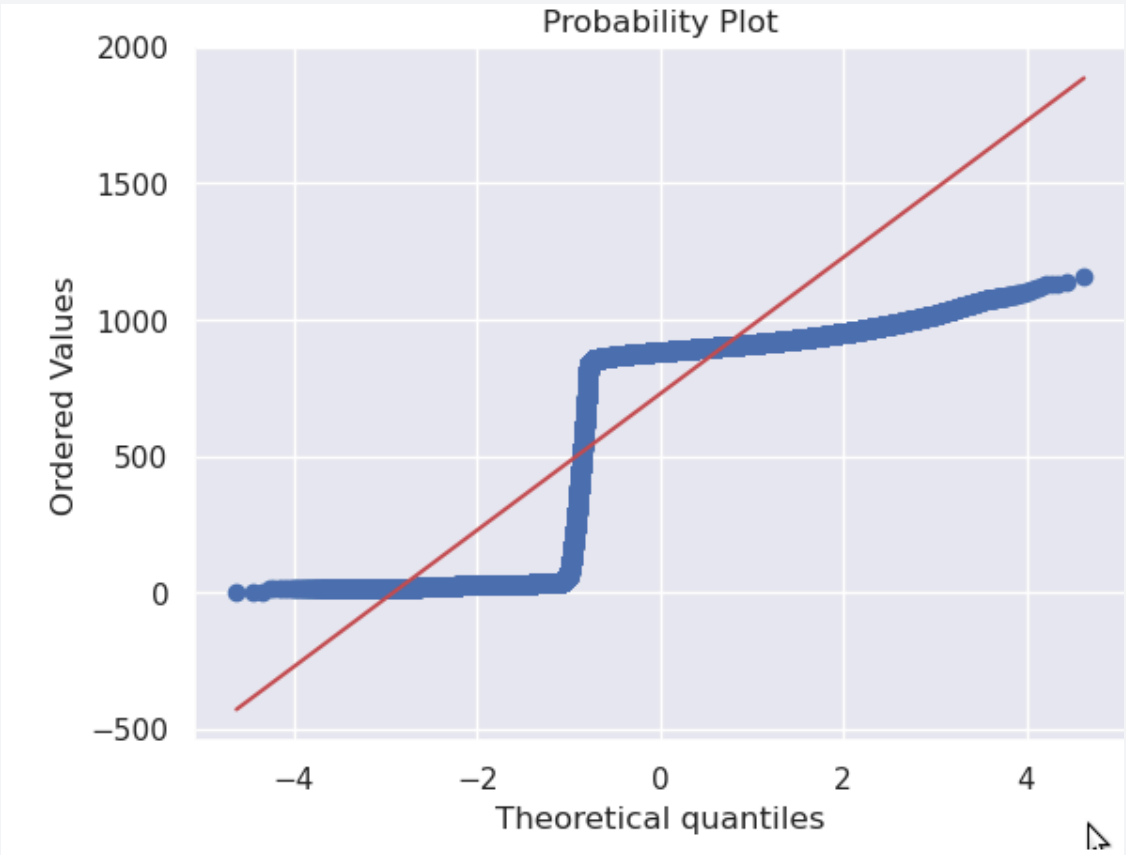
- **Visualization of Models or Highlight Abnormality:**

- Scatter plots showing data points with anomalies highlighted.
- 

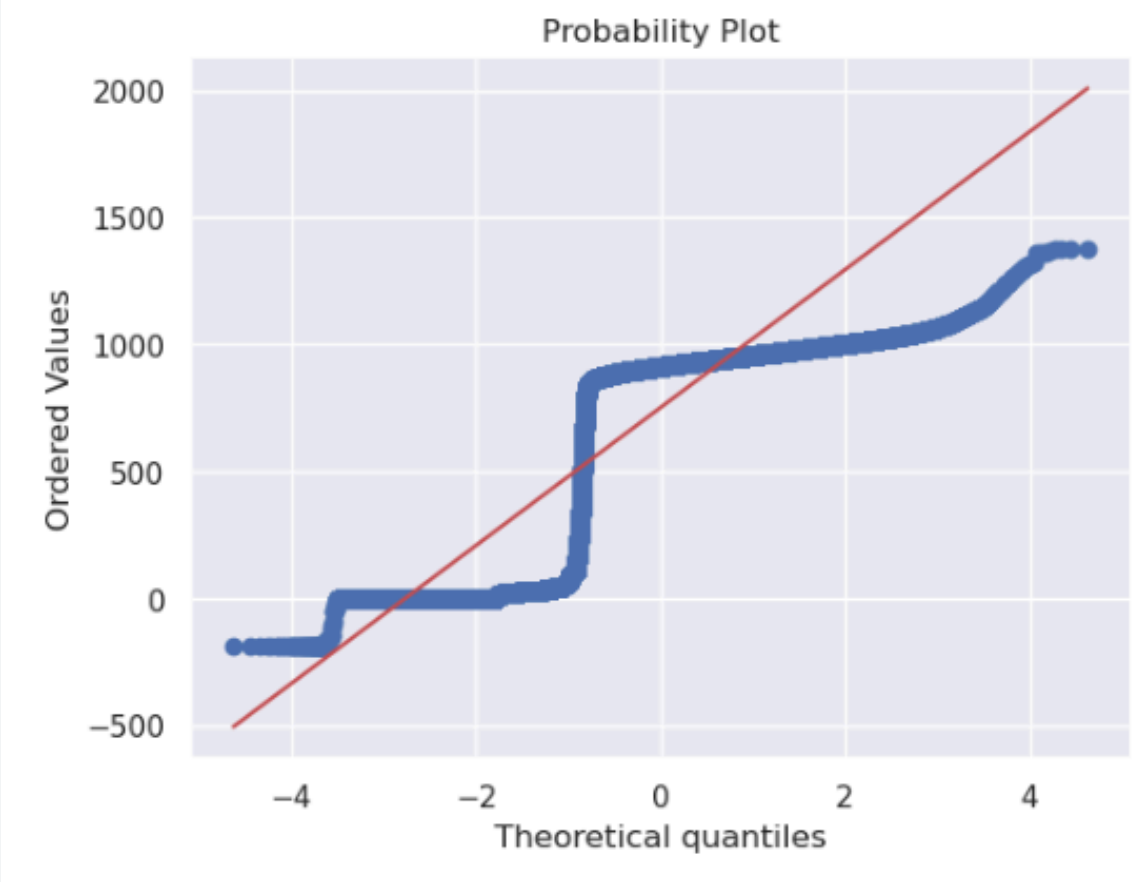


# VISUAL REPRESENTATION QQ-PLOT

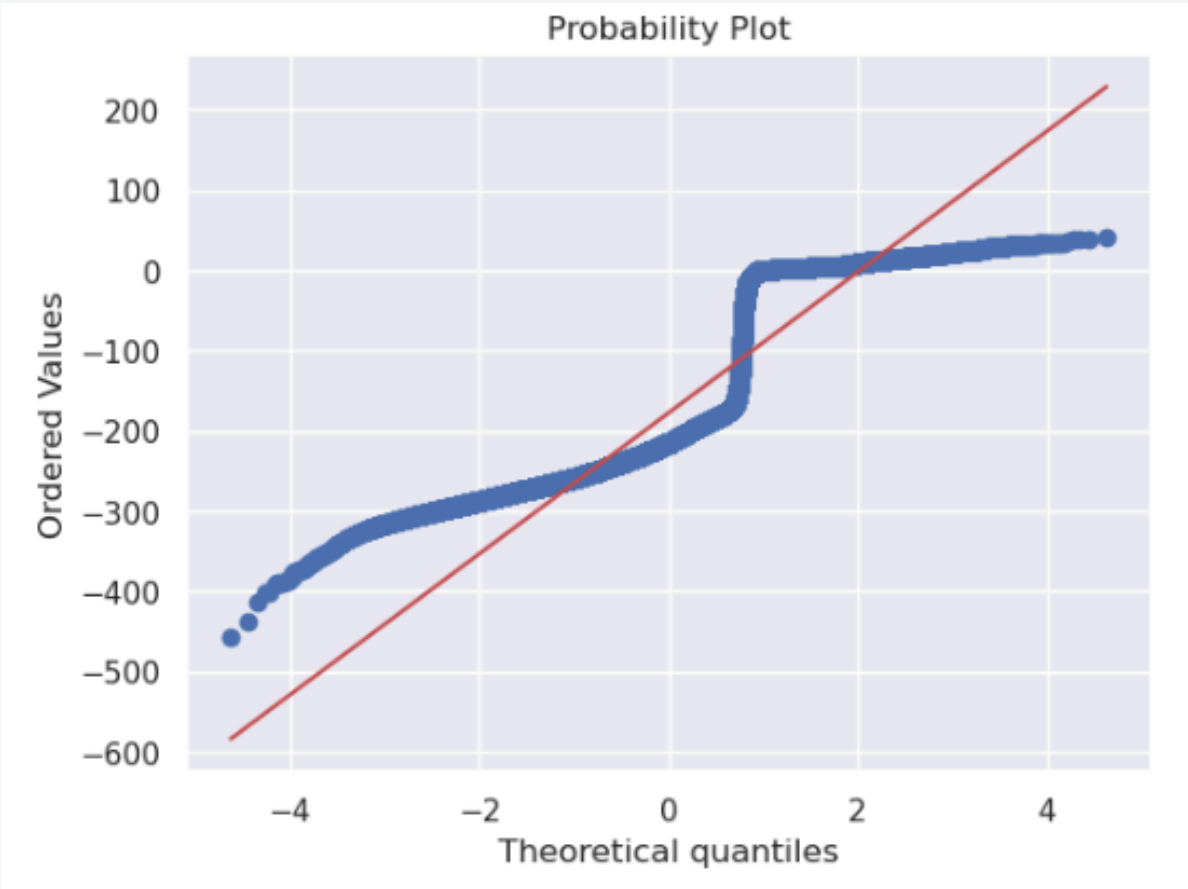
I



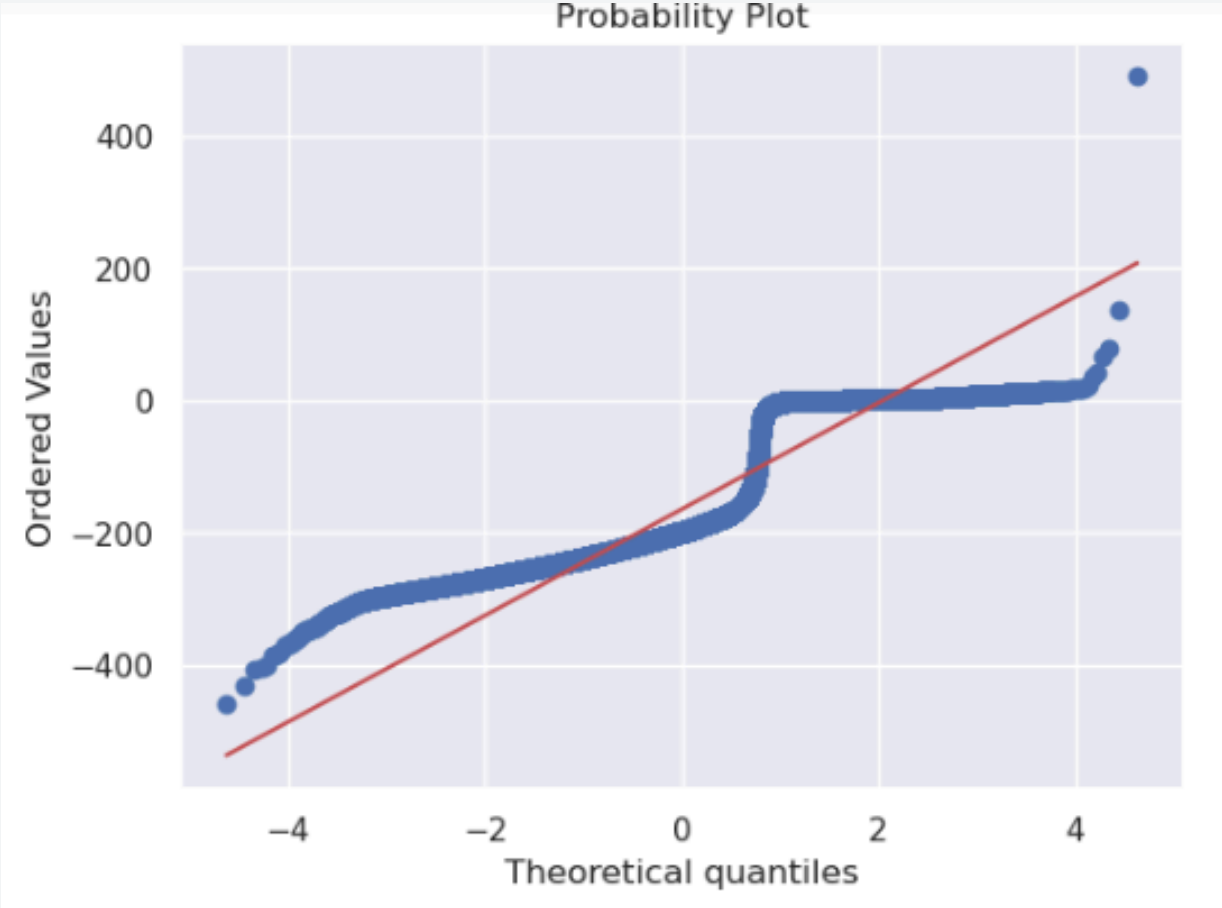
II



III

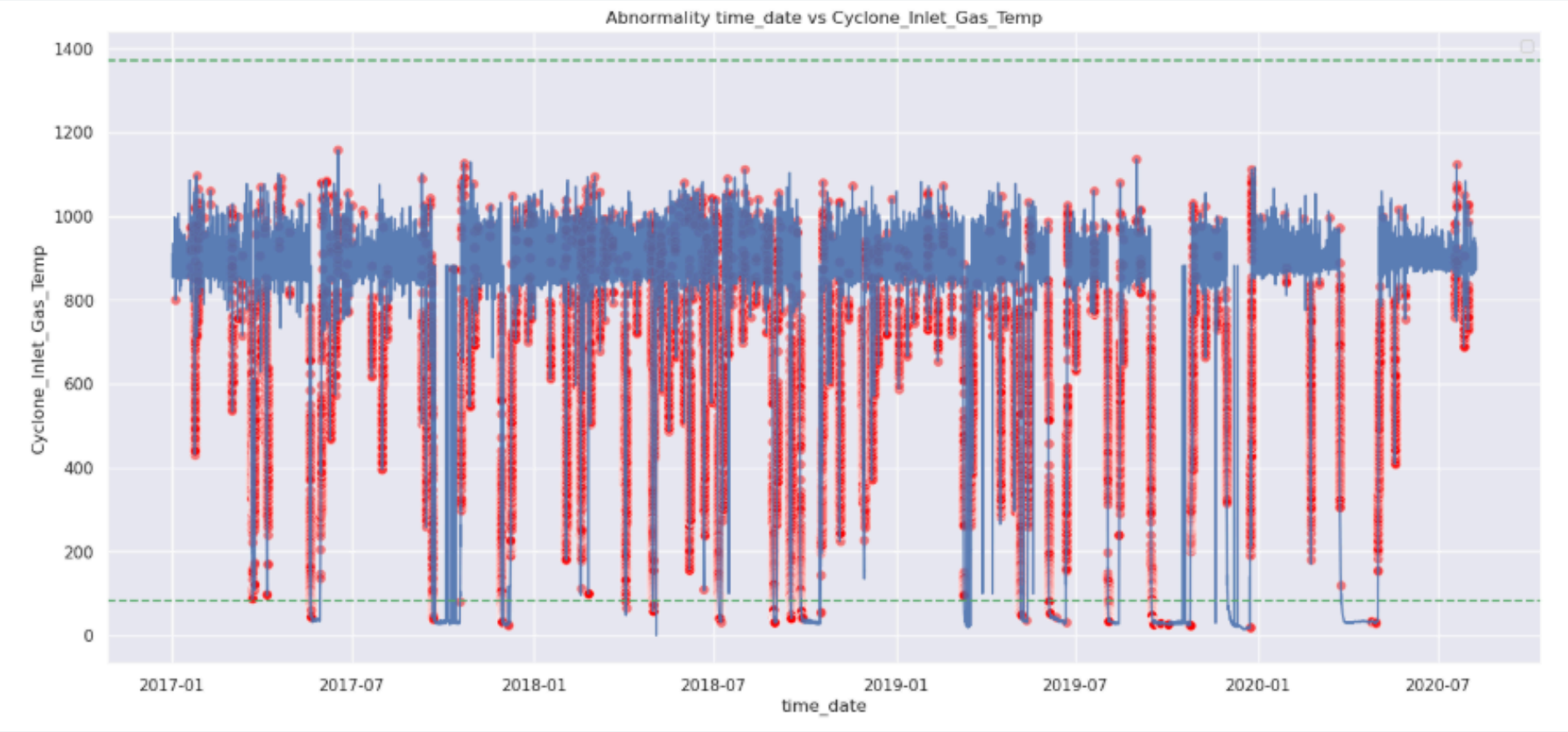
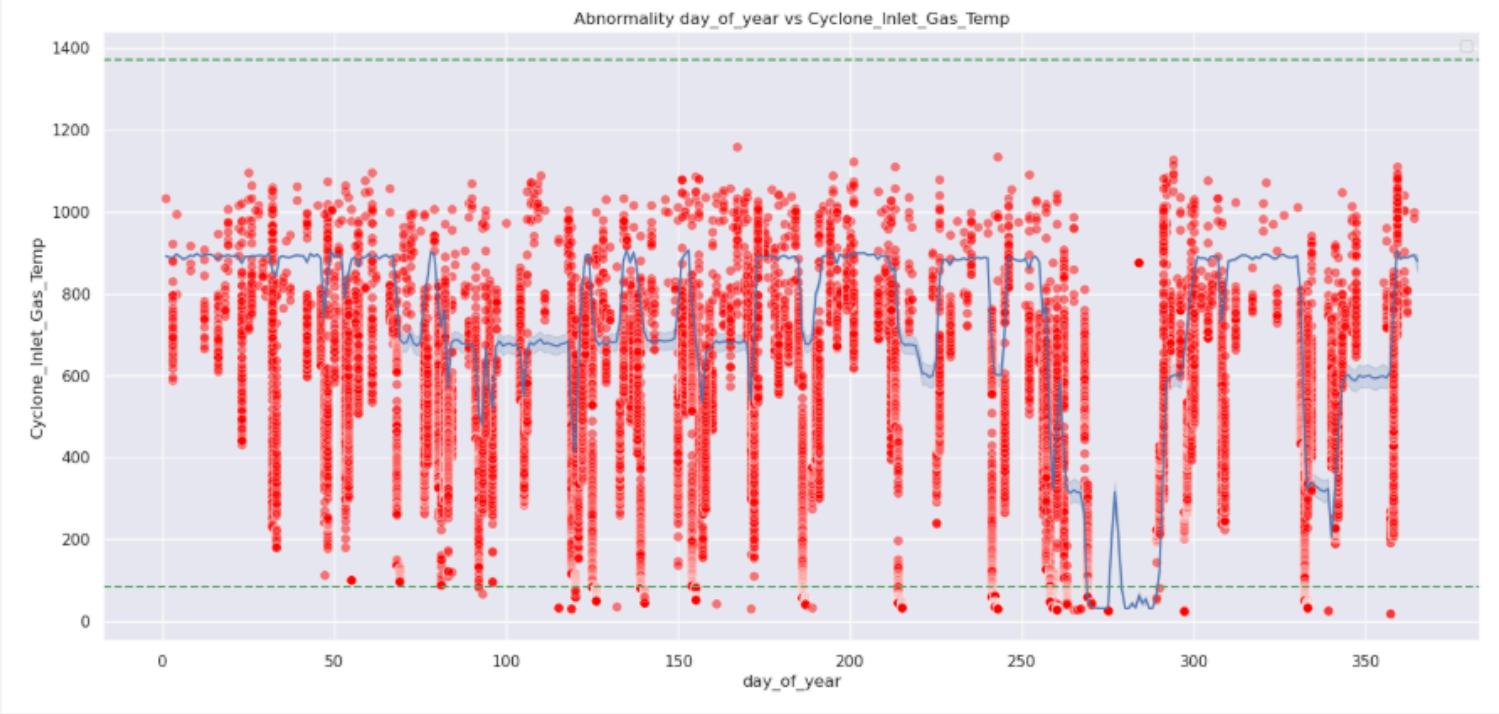


IV

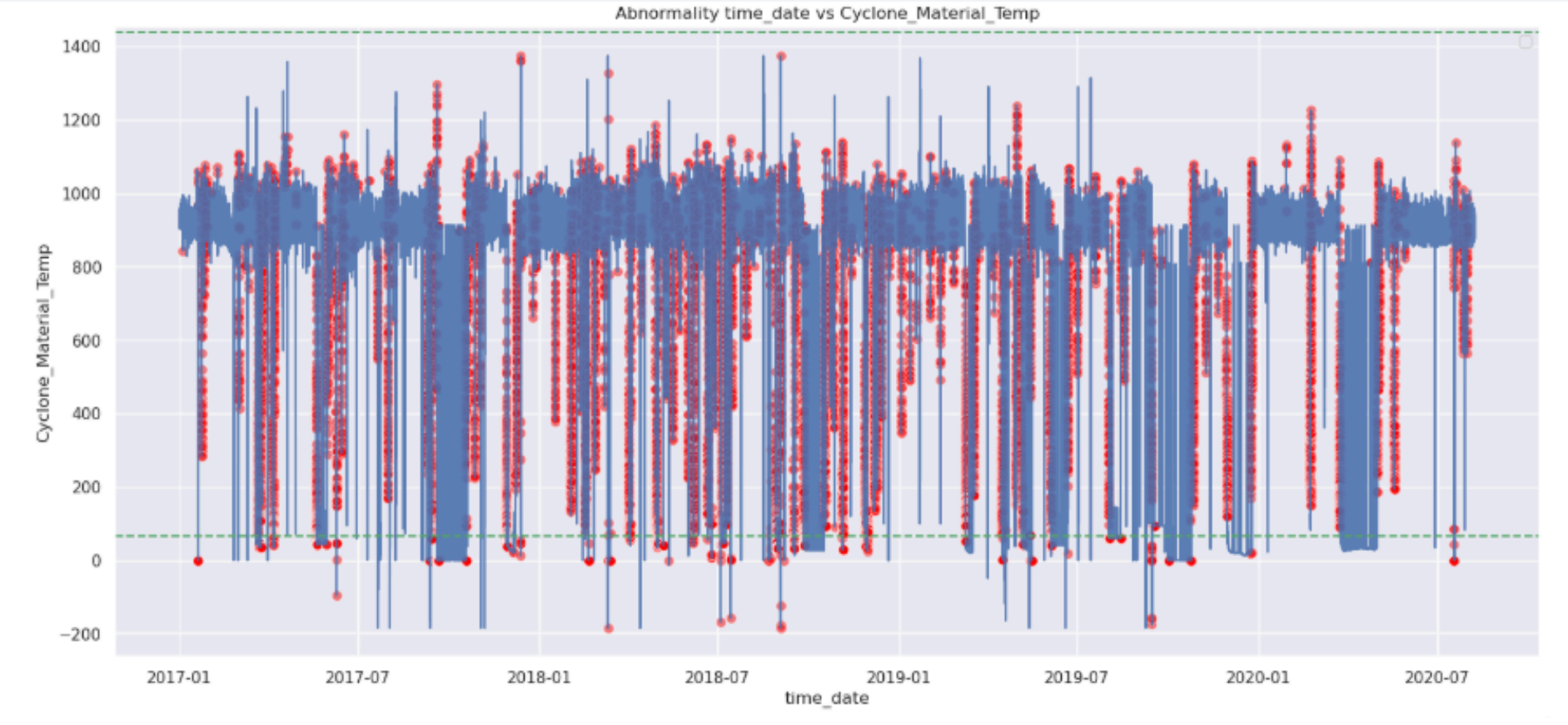
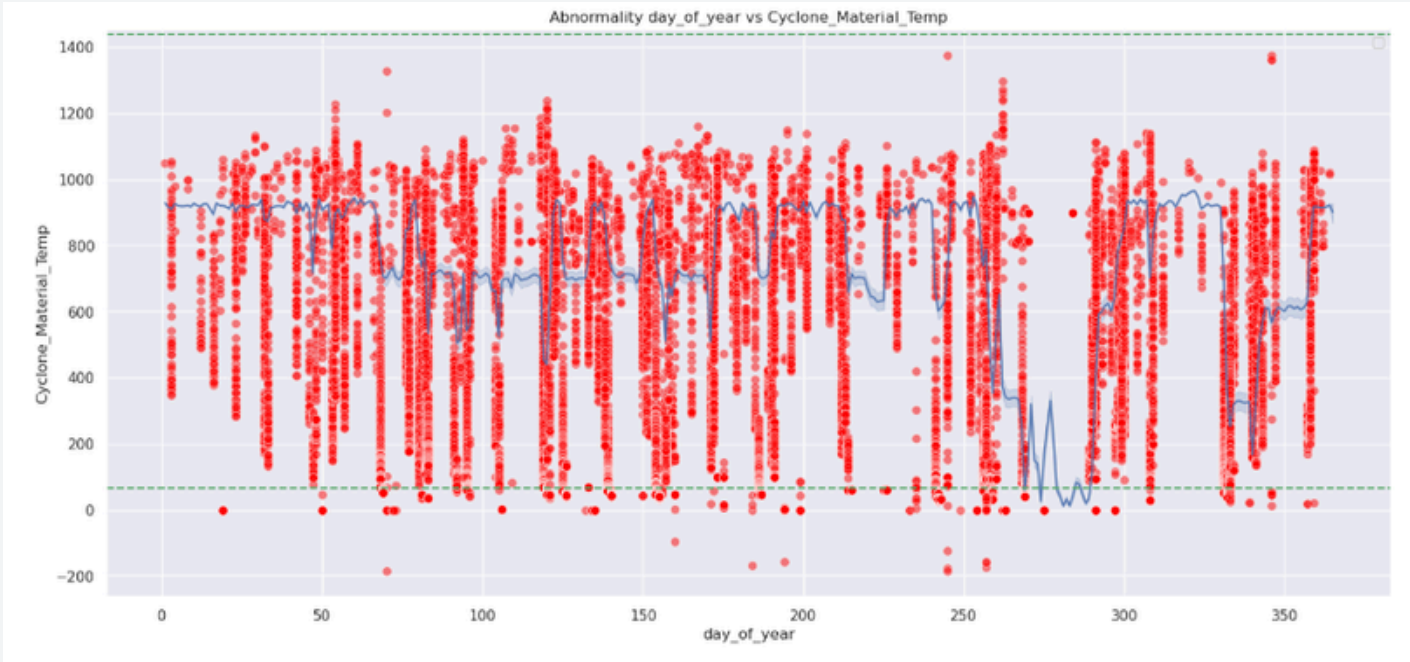


# VISUAL REPRESENTATION

'CYCLONE\_INLET\_GAS\_TEMP'



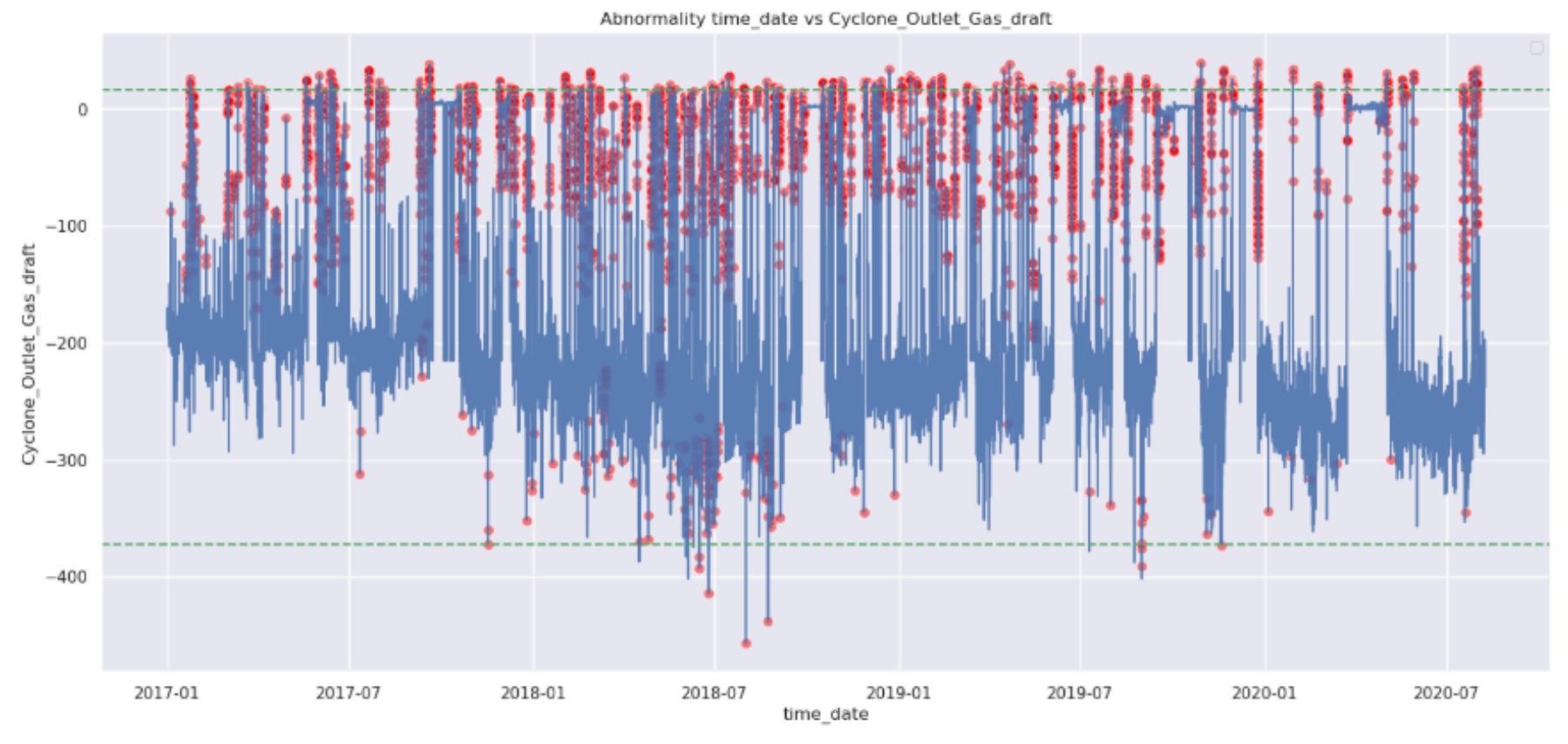
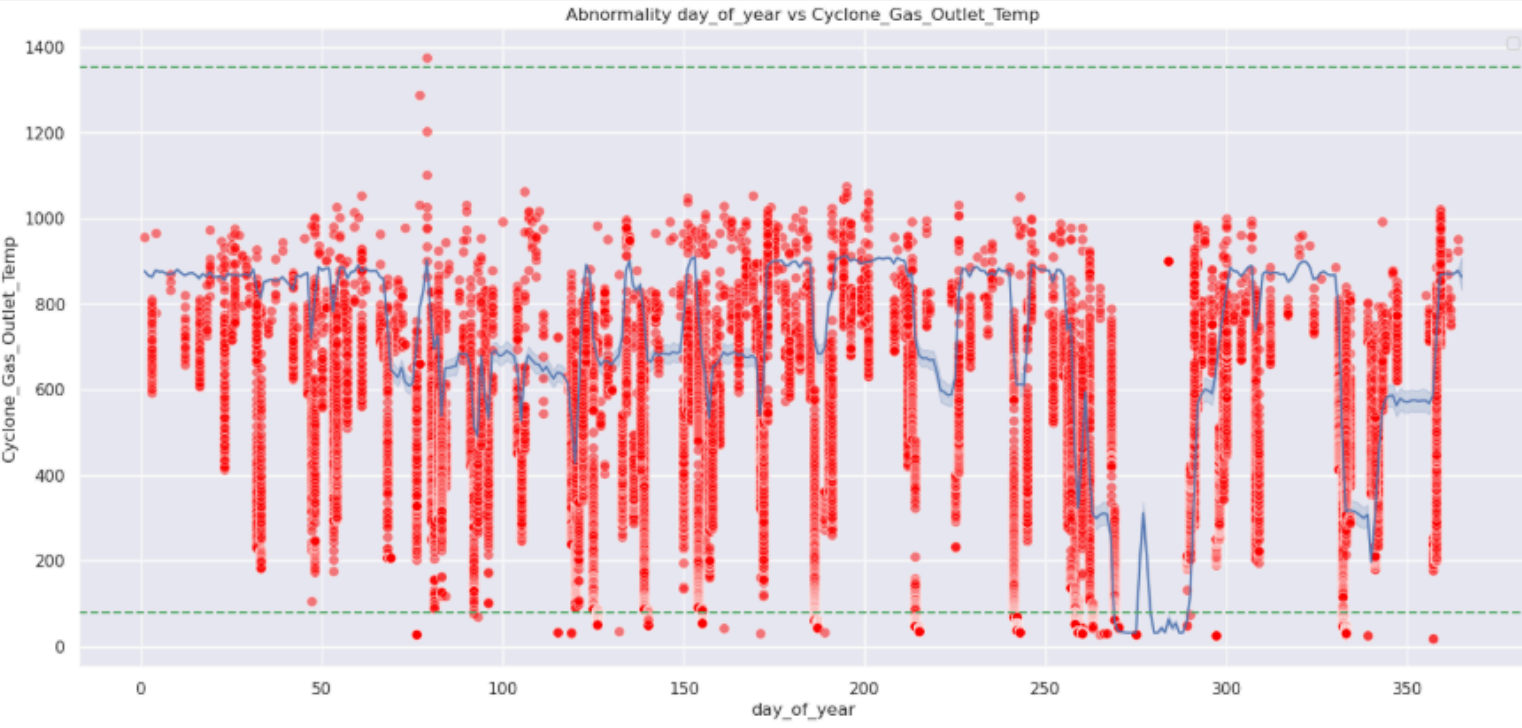
'CYCLONE\_MATERIAL\_TEMP'



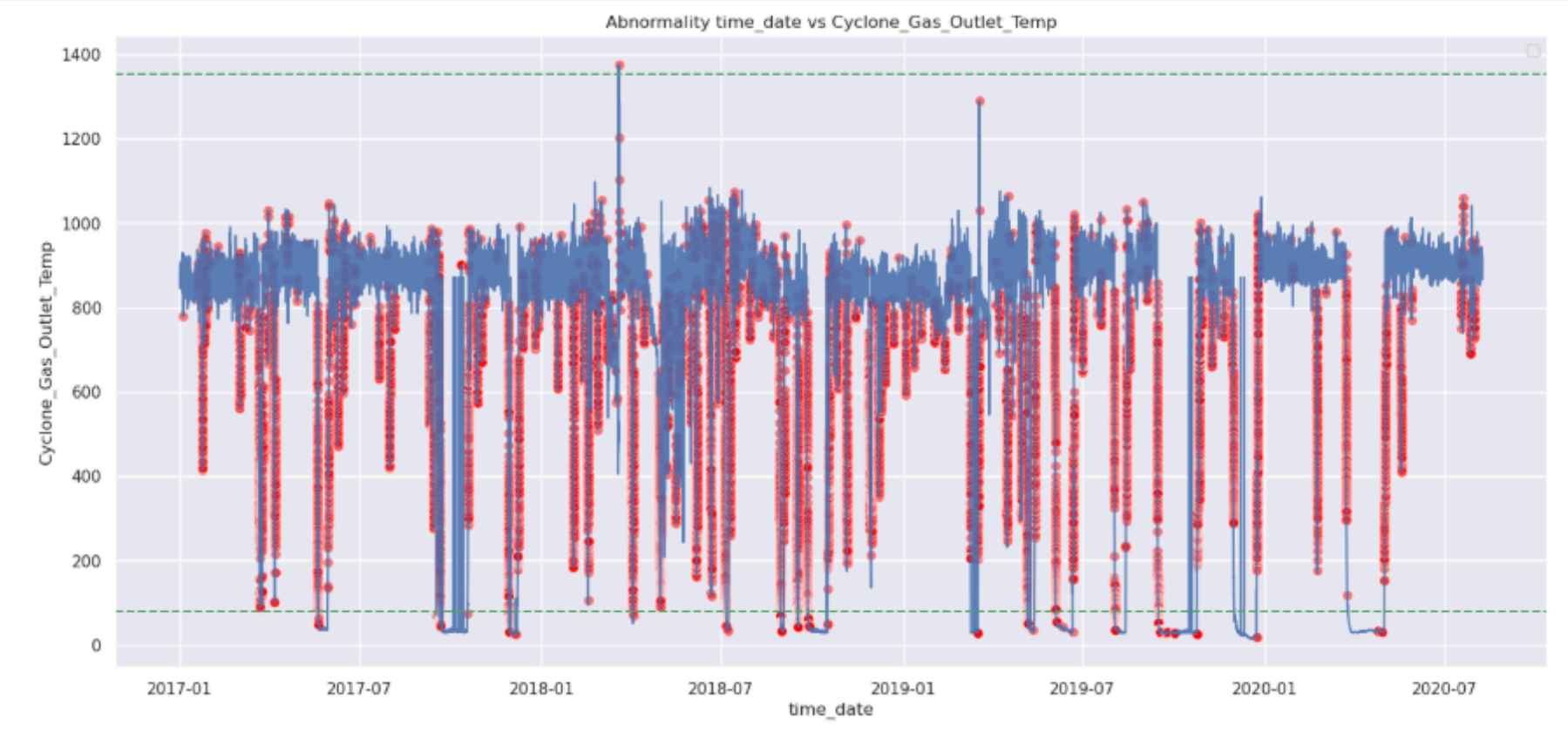
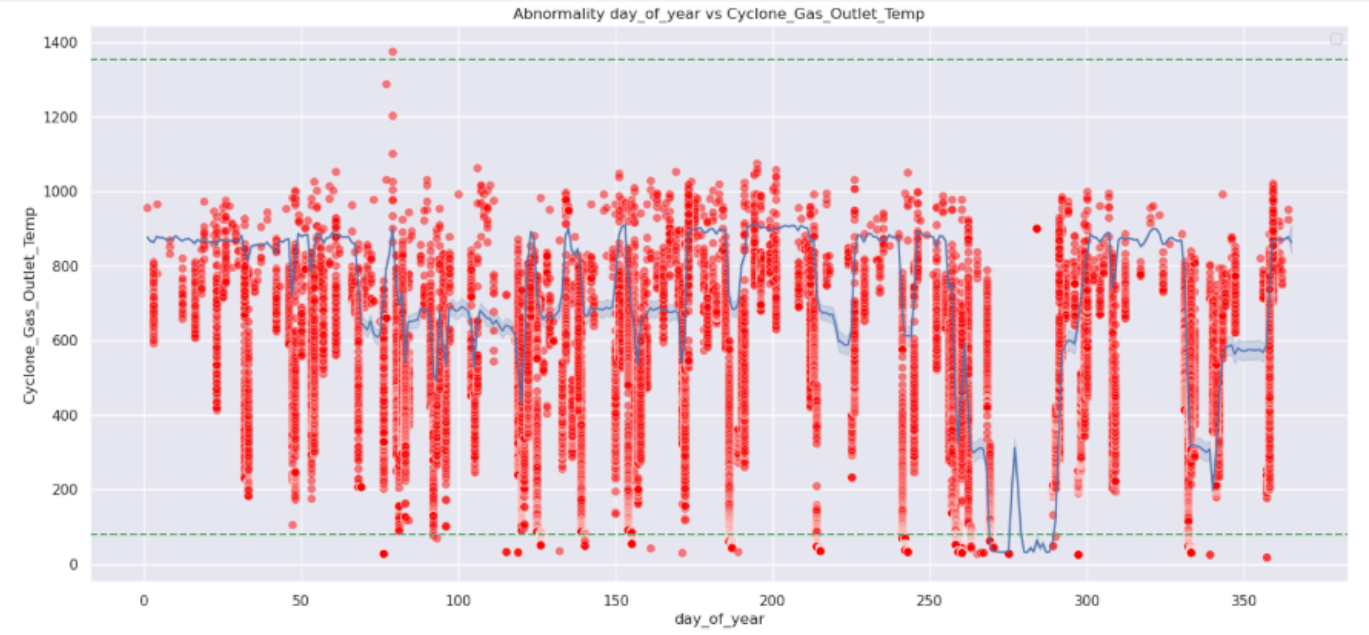


# VISUAL REPRESENTATION

'CYCLONE\_OUTLET\_GAS\_DRAFT'

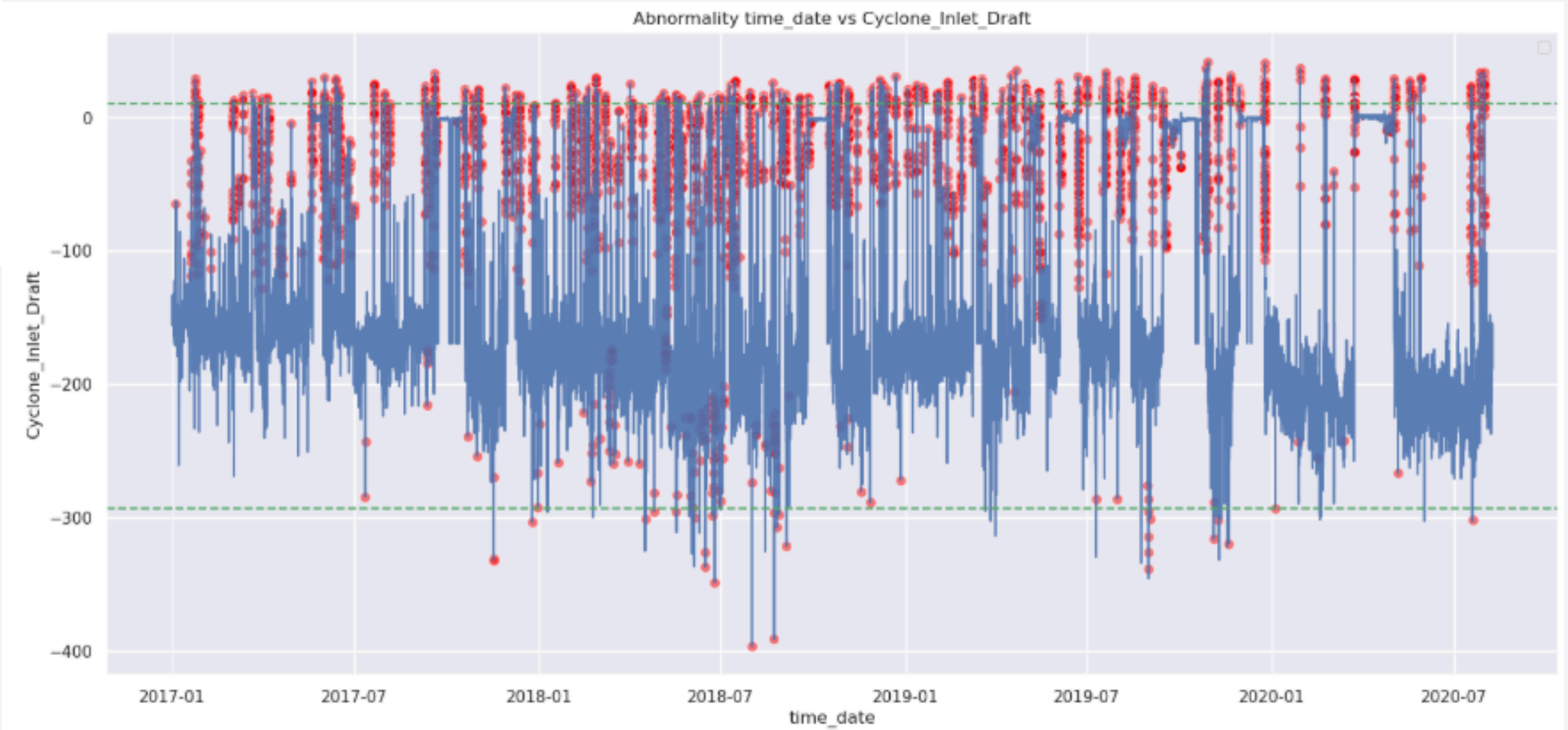
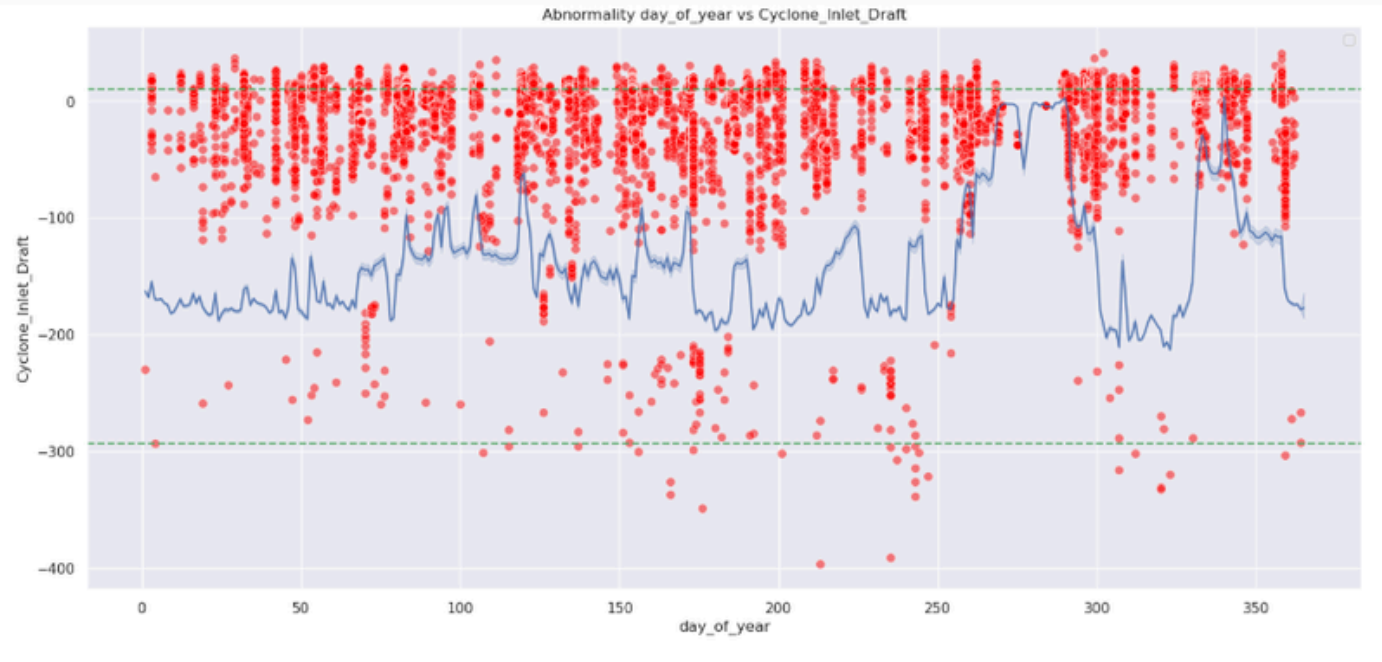


'CYCLONE\_OUTLET\_GAS\_DRAFT'



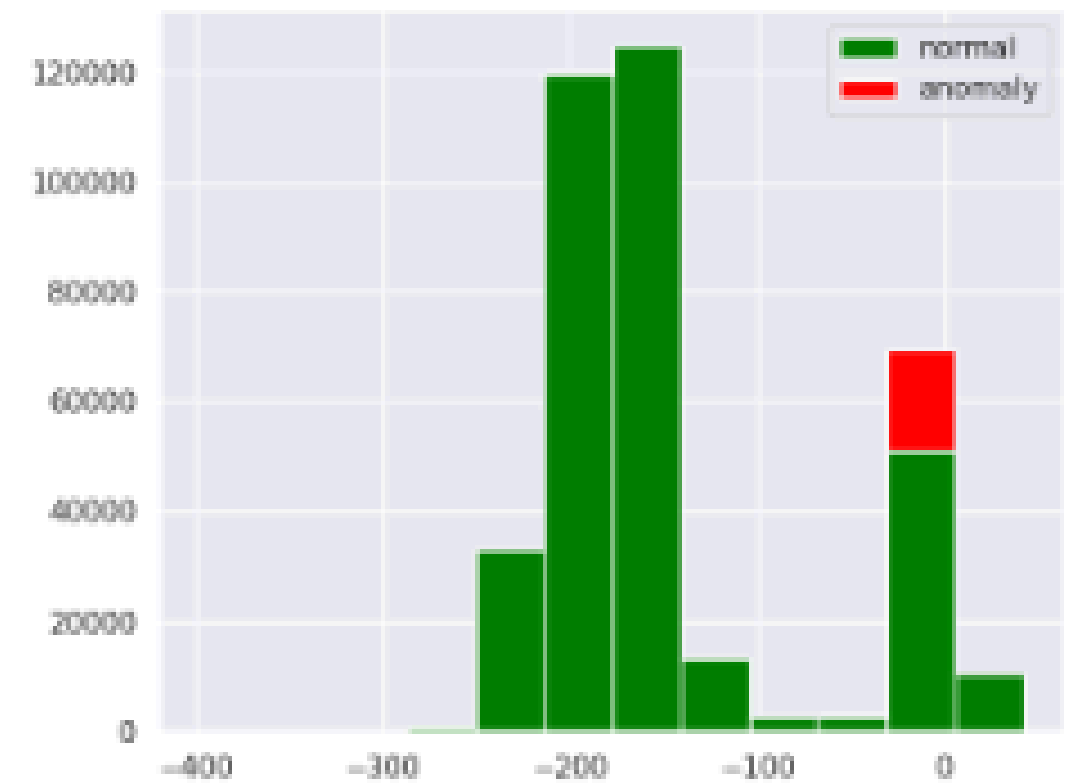
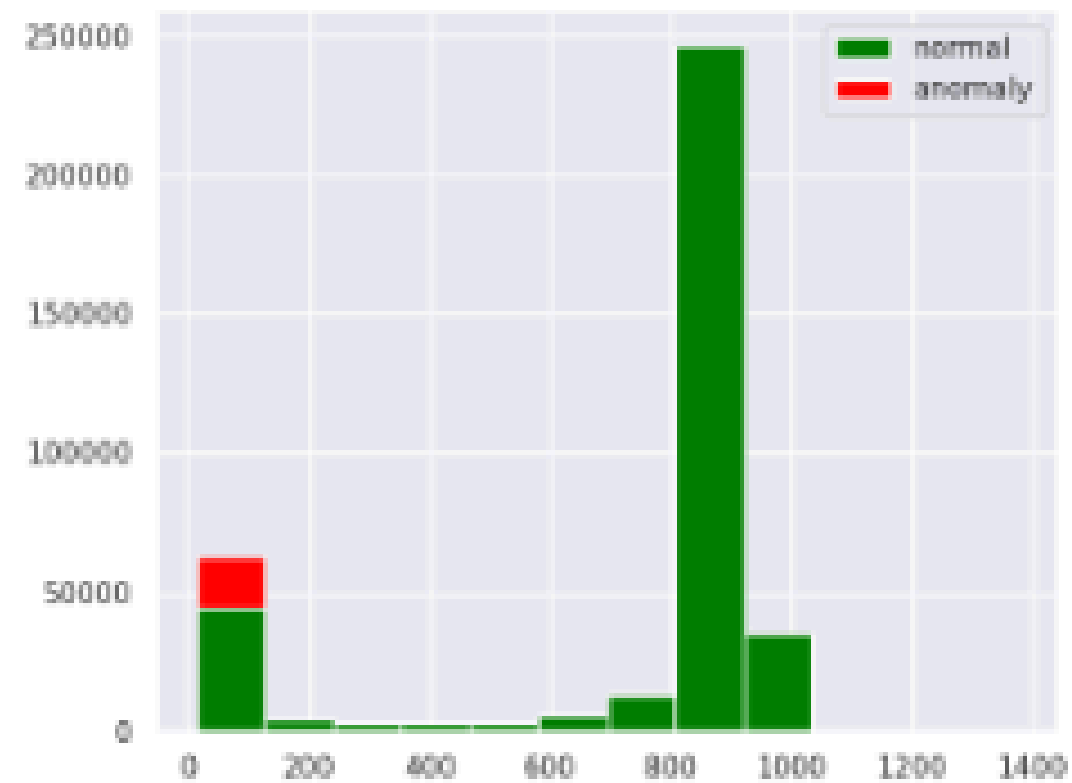
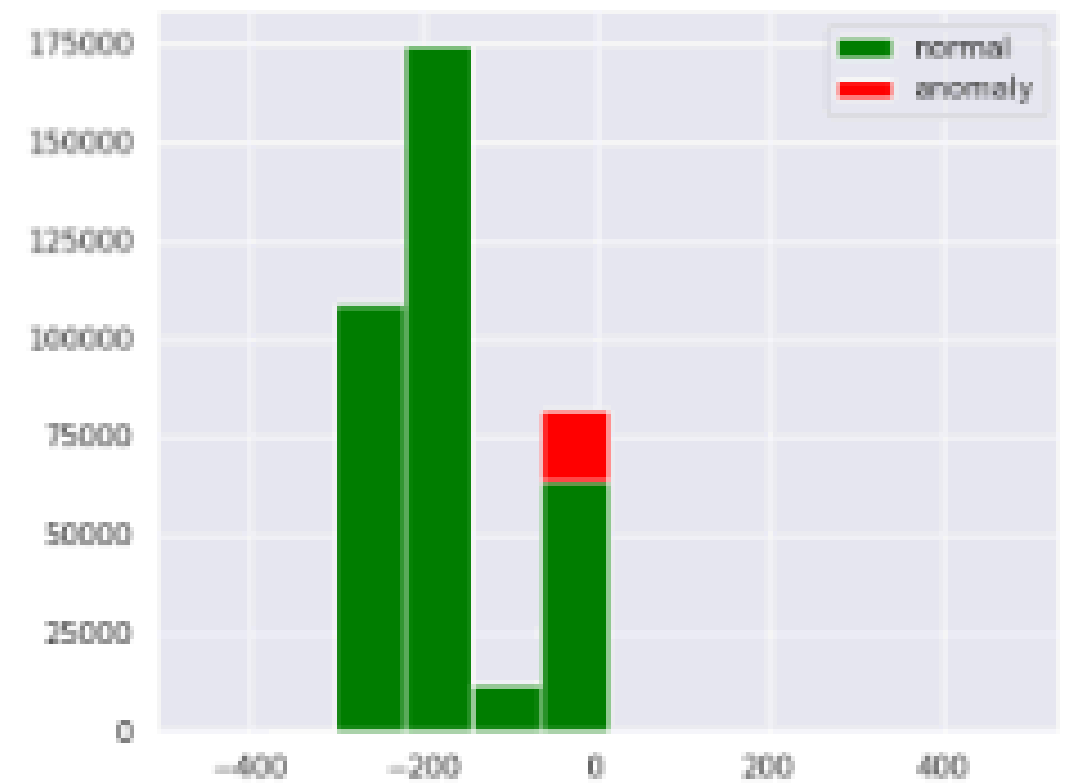
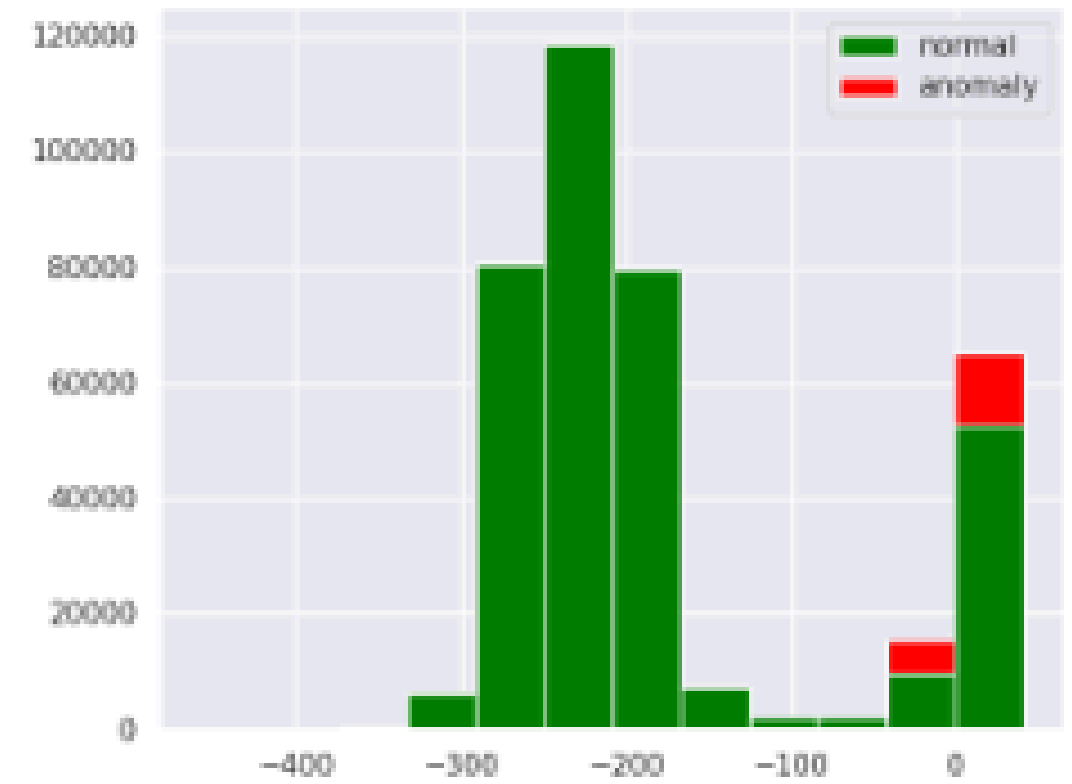
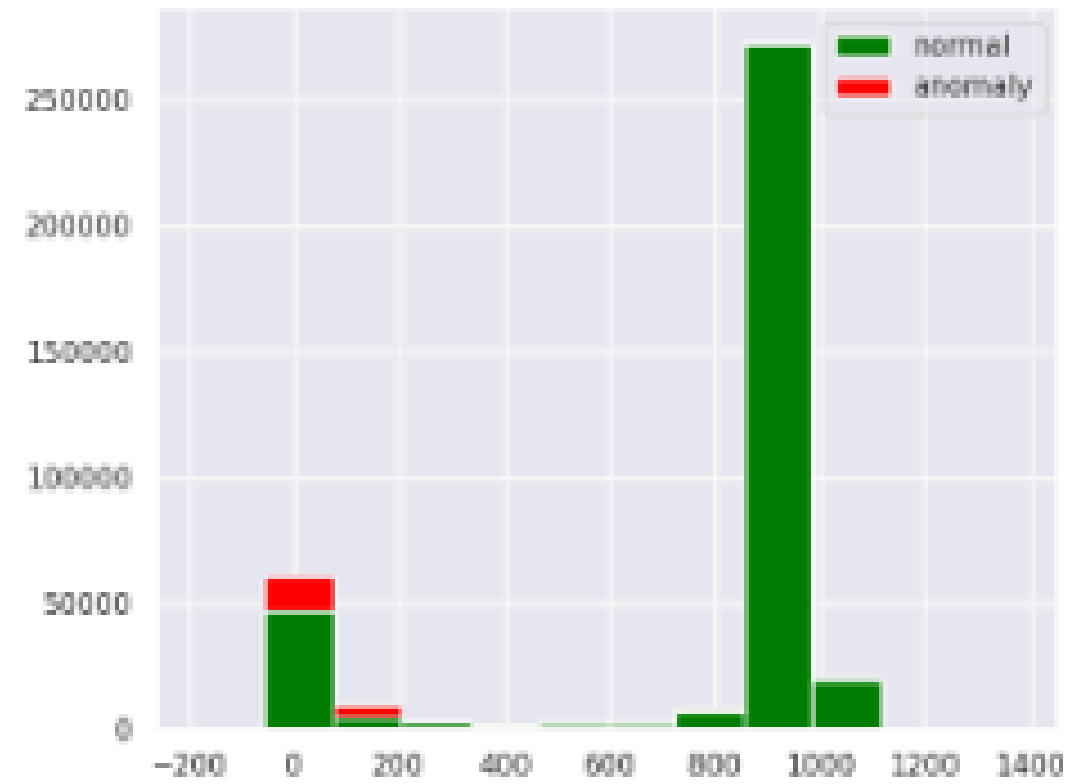
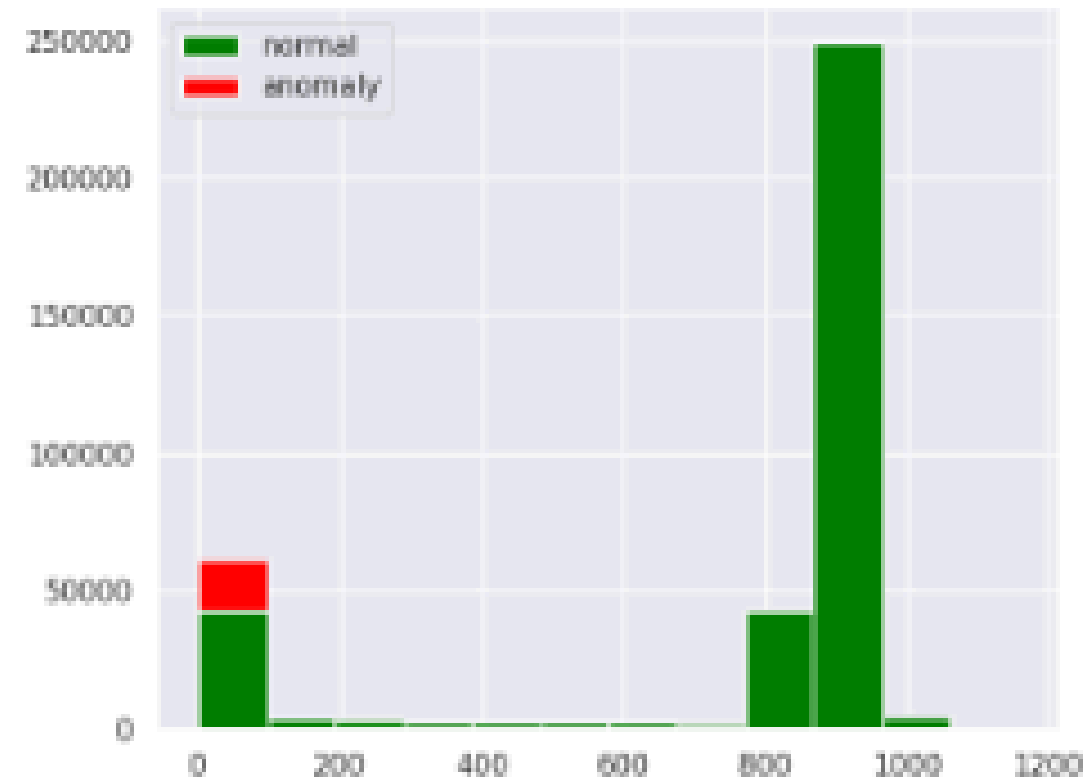
# VISUAL REPRESENTATION

' 'CYCLONE\_INLET\_DRAFT''





# VISUAL REPRESENTATION





# CONCLUSION

- Isolation Forest is effective for detecting broad anomalies.
  - Elliptic Envelope excels in detecting nuanced, localized anomalies
- 