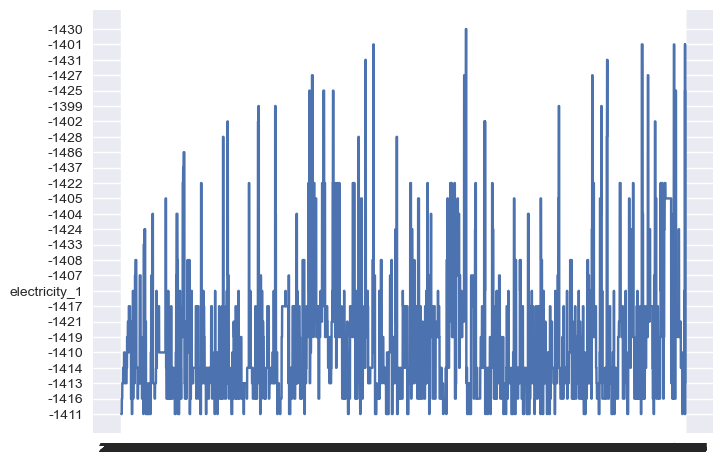
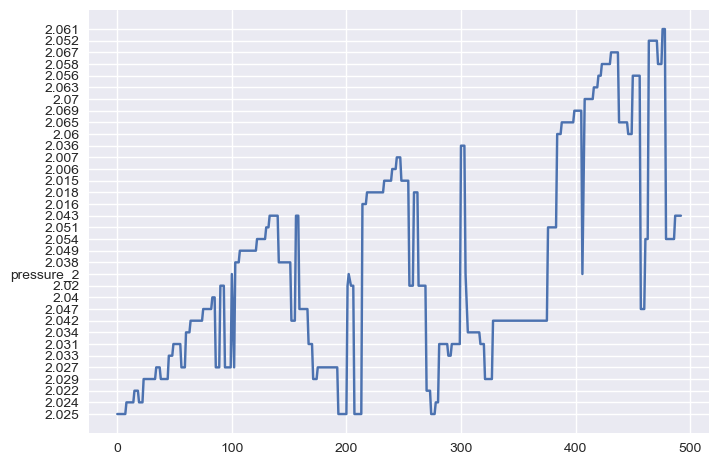
Exploratory Data Analysis (CNCITY)

**Gas Pressure Regulator 1:**

**Electricity 1: (5000 lines)**

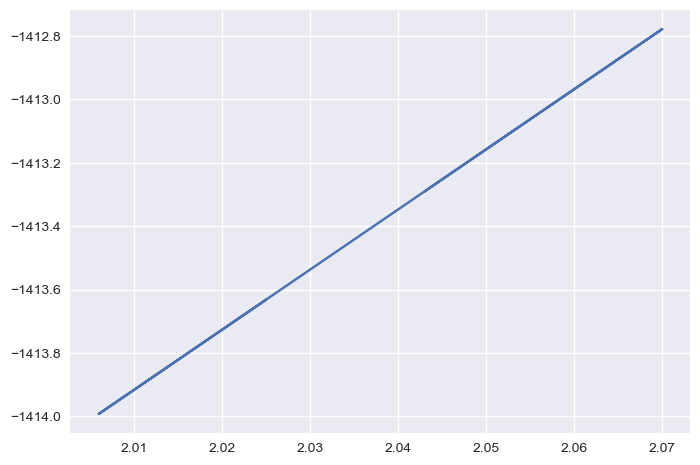
****

**Linear Regression: (pressure\_2, is\_maintenance) - epic failure**

****

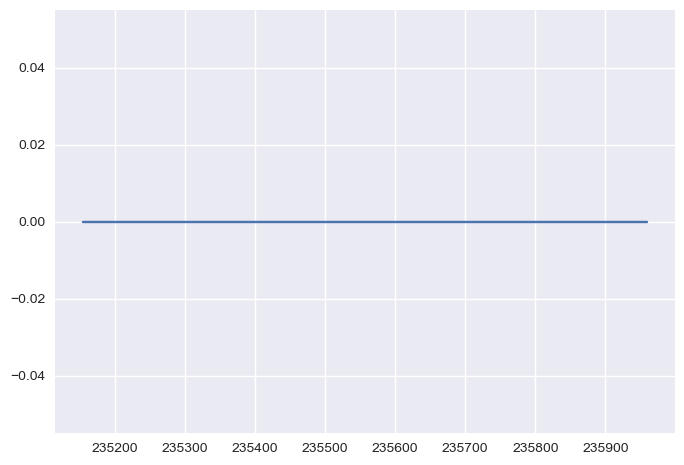
**Key takeaways from this graph:**

* **Since linear regression gives us a float between 0 and 1, logistic regression seems to be a better model when predicting is\_maintenance. (is\_maintenance is a binomial variable)**
* **It turns out that the ‘is\_maintenance’ column is just 0’s**

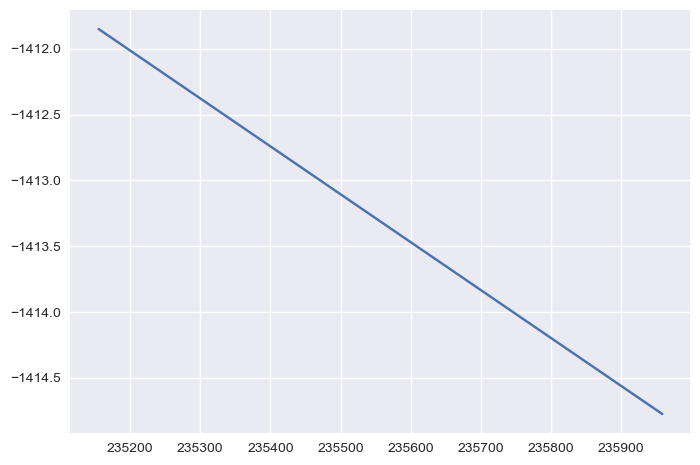
****

**Ran linear regression with ‘electricity\_1’ and ‘pressure\_2’**

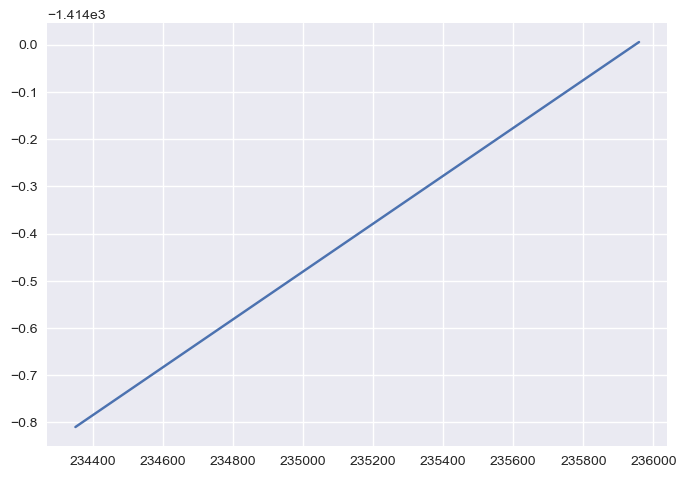
* **‘Electricity\_1’ and ‘pressure\_2’ are positively/linearly correlated**

****

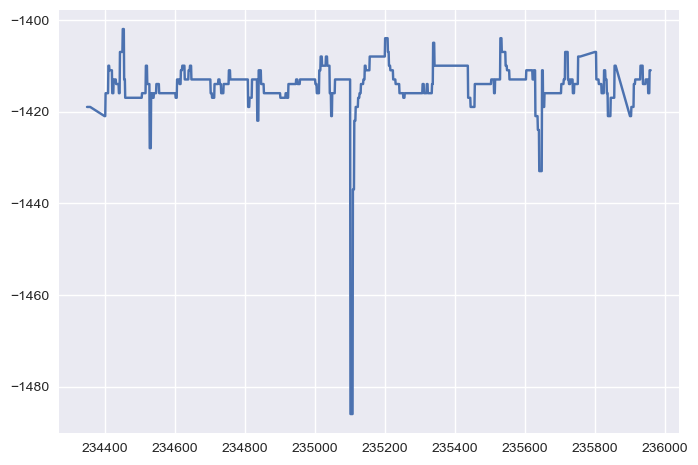
**X label: time, y label: electricity 2**

****

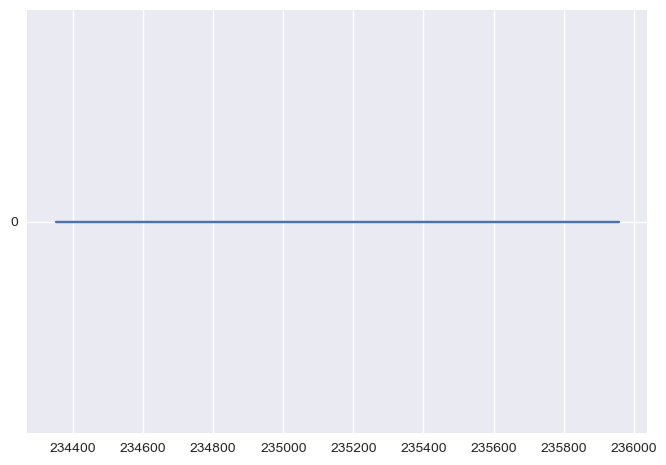
**Linear Regression: time, electricity\_1**

****

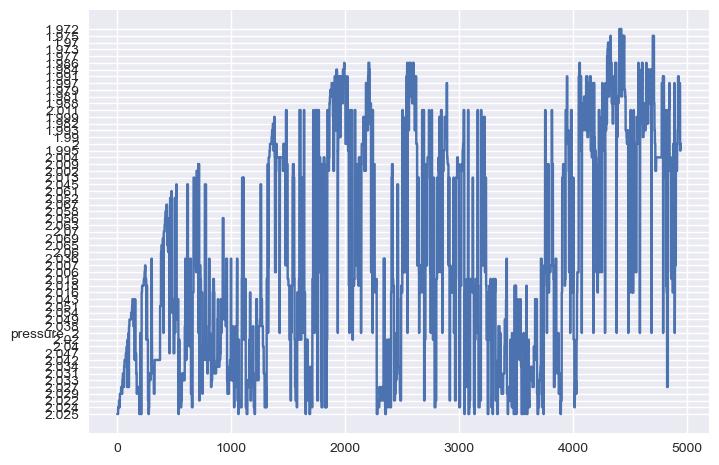
**The graph below is far more informative and intuitive.**

****

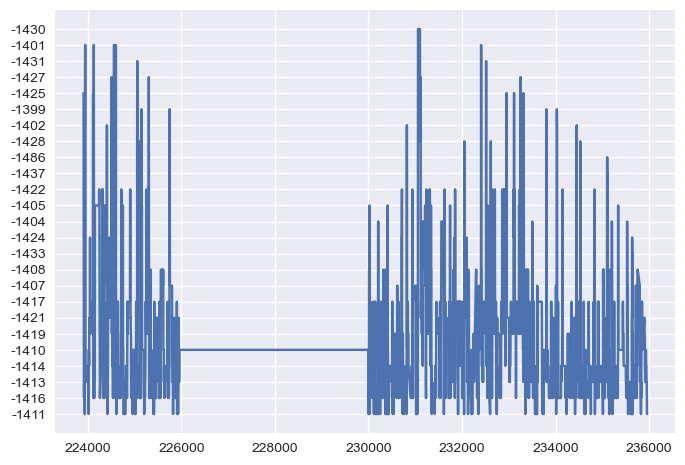
**Pressure\_1\_status and time**

****

**Pressure\_2:**

****

**X axis: time, Y axis: status\_electricity\_1**

****

**X axis: time, Y axis: electricity\_1**

[**https://www.w3schools.com/python/python\_ml\_knn.asp**](https://www.w3schools.com/python/python_ml_knn.asp)

**6/11/2024:**

* **Query + connection manager**
* **Table\_df (5000+ lines)**