**Weekly Report - 3**

**CSE 523 Machine Learning**

Group Name: **Mavericks**

**Store Sales – Time series Forecasting**

| **Group Members** | **Roll No.** |
| --- | --- |
| Dhruv Prajapati | AU1940192 |
| Aaryan Mori | AU1940194 |
| Harshil Doshi | AU1940279 |
| Dhaval Chaudhary | AU1940180 |

1. **Task Performed in the Third week:**

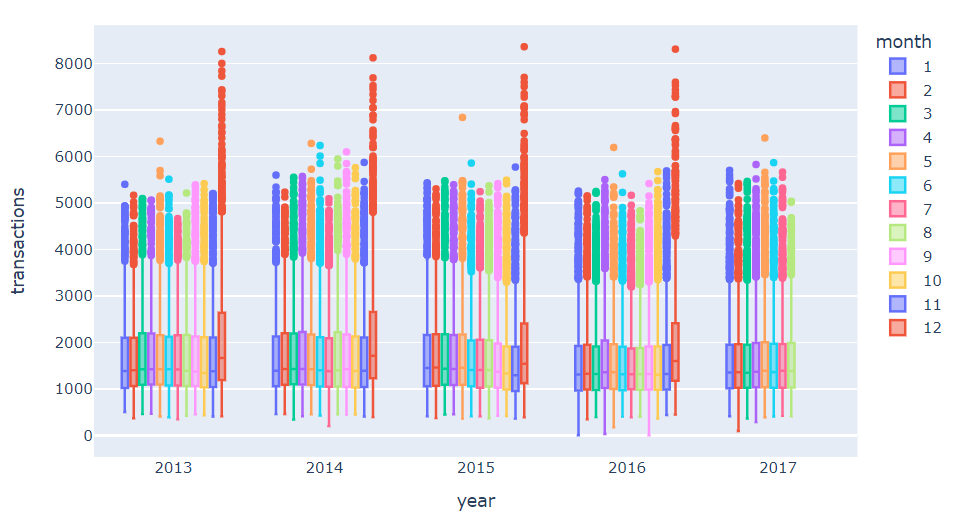
* Visualized the data for a better understanding.
* Imported the Transactions data and found the stable pattern that increases sales.
* Found the relationship between Transactions and Sales data.

1. **Outcome of the task in the Third week:**

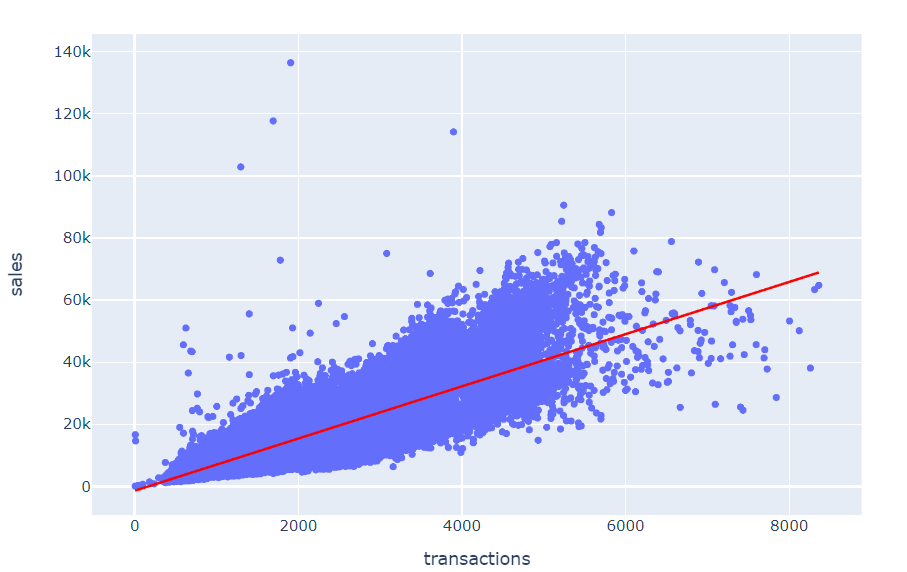
* By visualizing the data, we got to know the following things:

1. We observed that to find the relationship between sales and transactions, we need to sum up the sales feature.
2. Sales gives the total sales for a product family at a particular store at a given date. Fractional values are possible since products can be sold in fractional units (1.5 kg of cheese, for instance, as opposed to 1 bag of chips).
3. So, transactions will be one of the relevant features in the model.

* Results after importing the Transaction data are as follows:



1. We found that there is a stable pattern in Transaction. All the months are similar except December from 2013 to 2017 by boxplot. In addition, we observed that the Store sales had always increased at the end of the year.
2. We've just visualized a pattern that increases sales. It was the end of the year. We saw that the transactions increase in spring and decrease after spring.

* We found the relationship between sales and transactions by plotting a graph. 

We came to know that by looking at their relationship, we can see that there is a highly correlation between total sales and transactions also. Both sales and transaction features are highly correlated to each other.

1. **Task to be performed in the next week:**

* Researching Ecuador’s economy to be able to understand the change in oil prices.
* Finding the missing data points in the daily oil data by using various imputation methods.
* Exploring the best fit algorithm for the model.