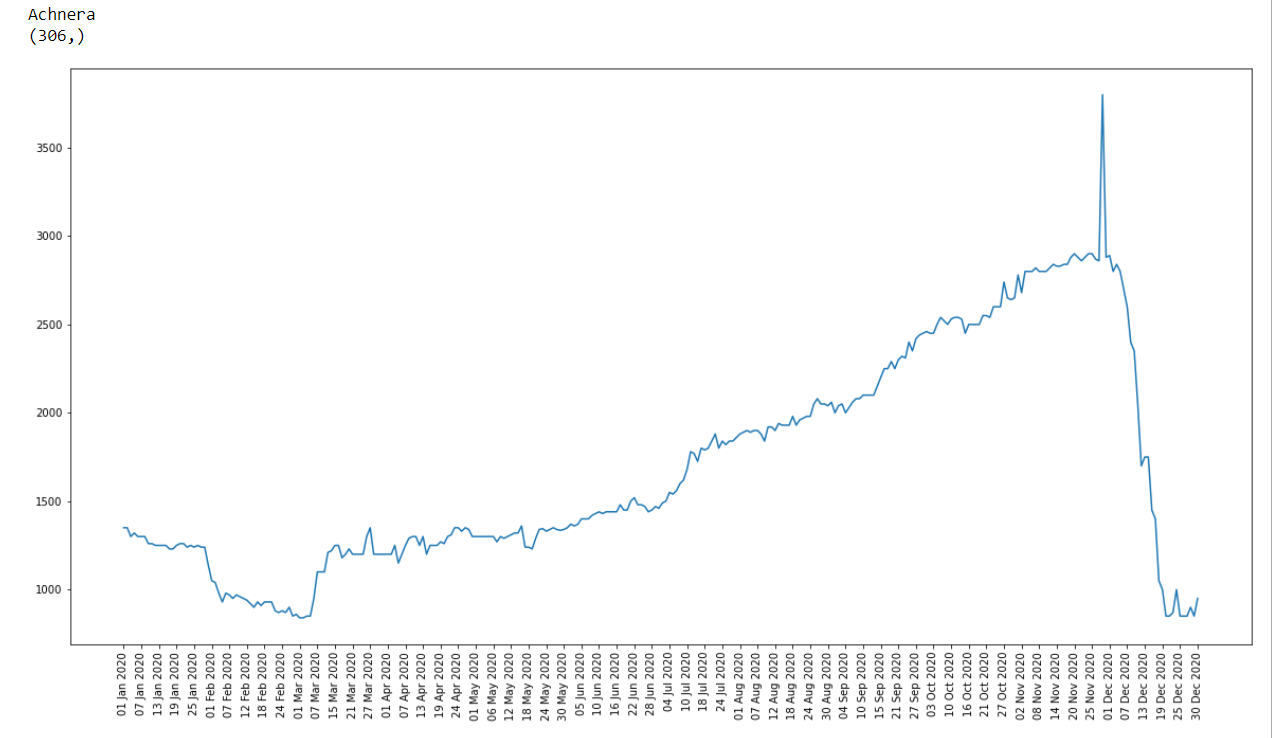
**Vishwas upadhyaya(vishwasu2001@gmail.com) – agrilinks assignment**

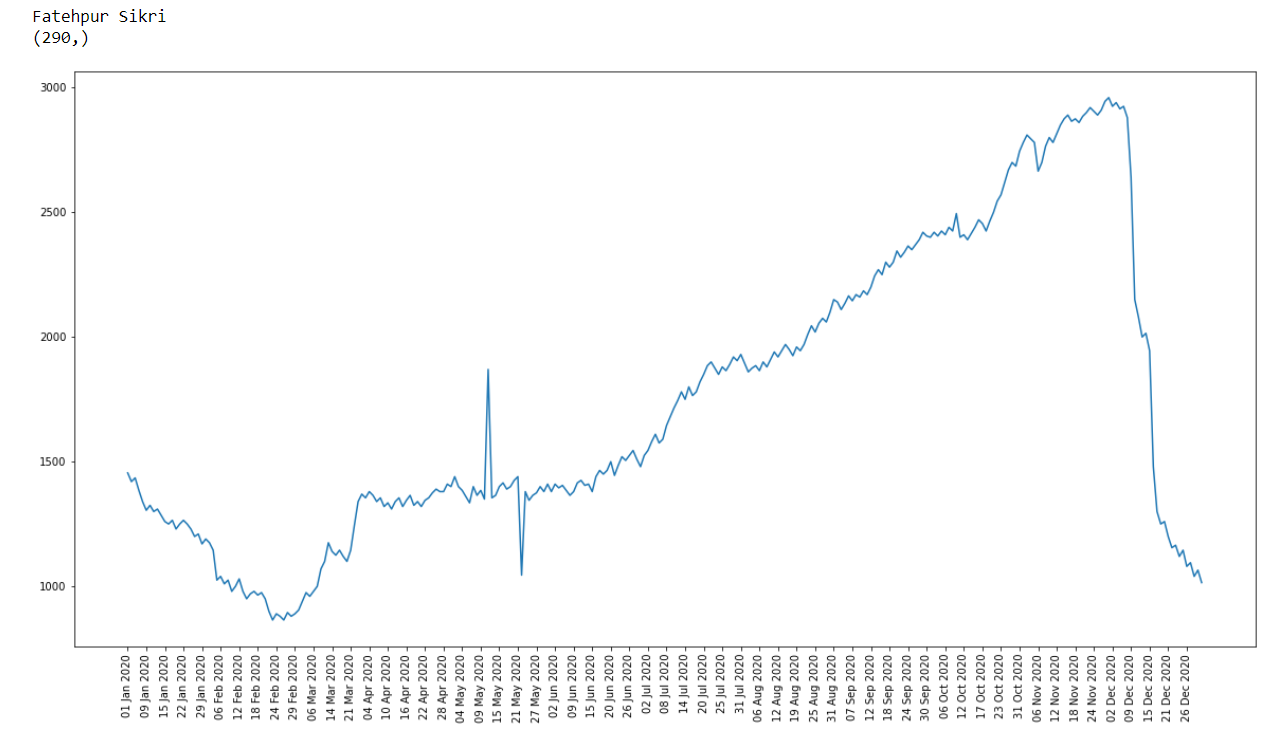
**Point b outputs---**

1. Top 4 major markets are – 1. Achnera 2. Fatehpur Sikri 3. Agra 4. Fatehabad
2. Price patterns for each of them is --

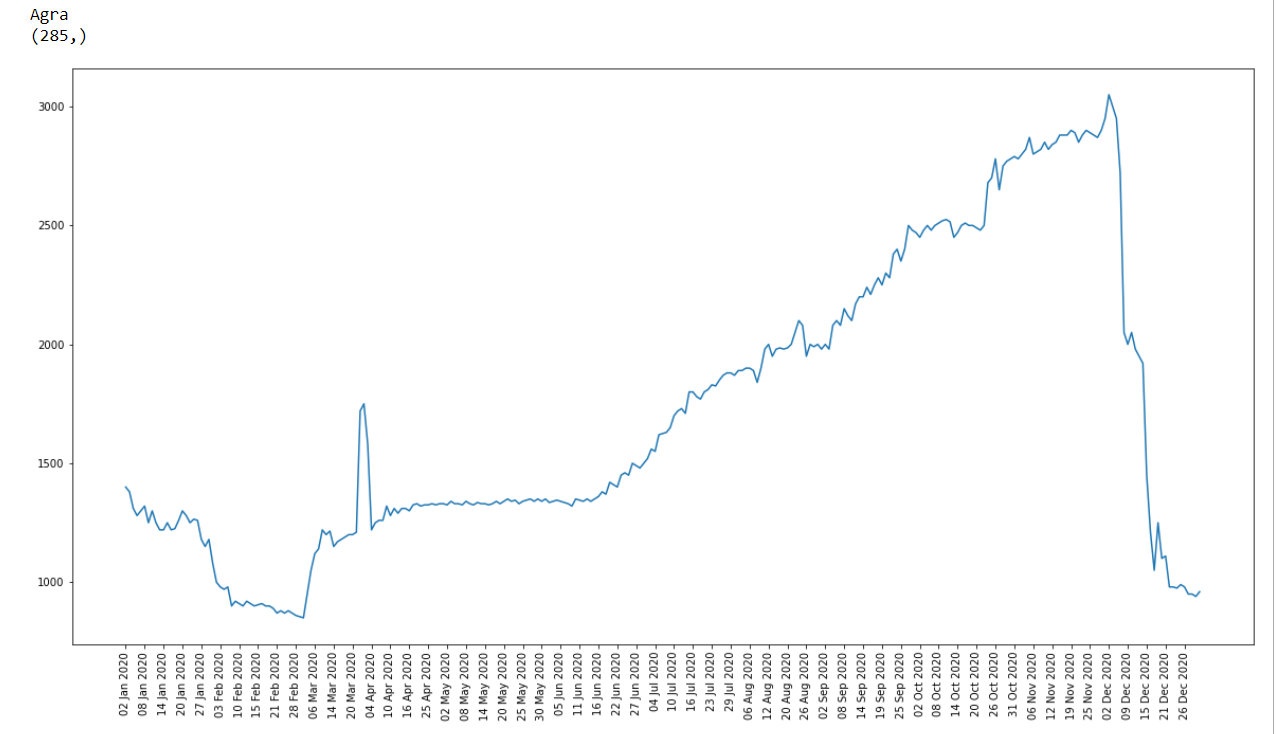
Achnera -

- 

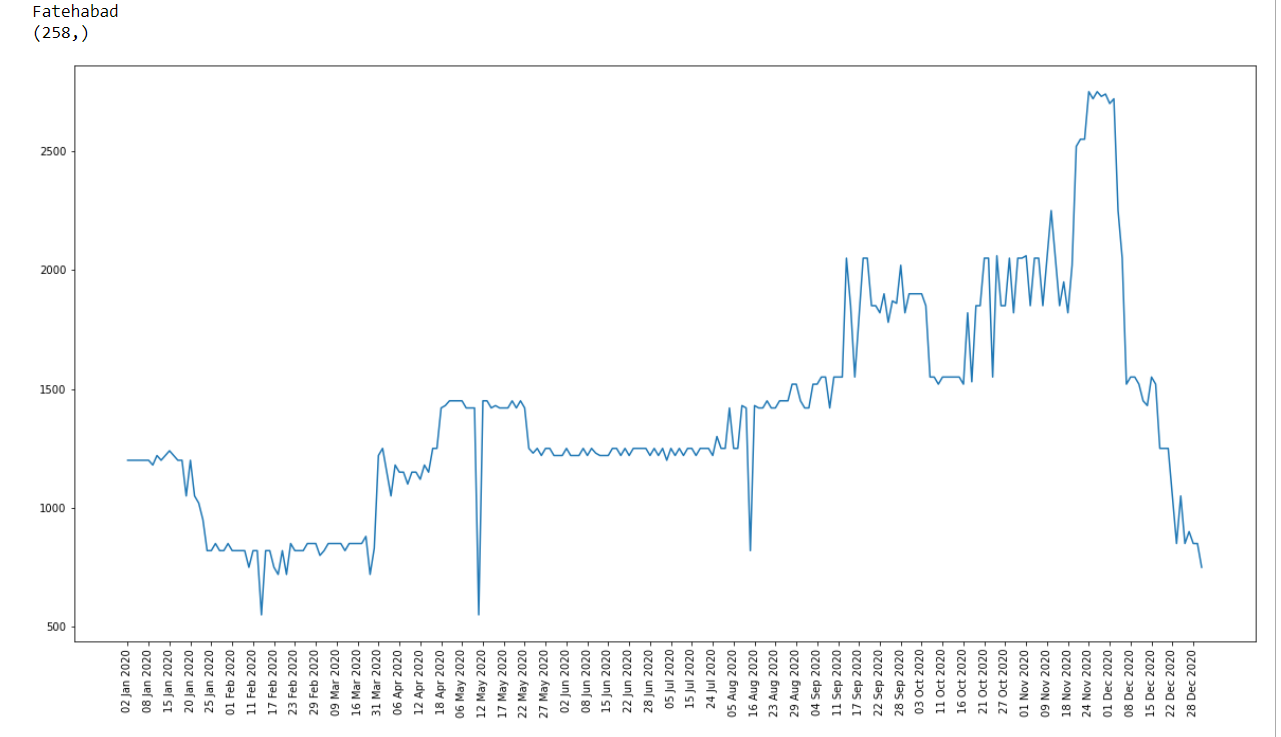
Fatehpur Sikri --



Agra --



Fatehabad --



In all these price trends plots in the starting months of the year(Jan) the price is little bit high but in feb and march the price is dropping and after march till november price is increasing and in december the price is falling enormously.

**Point c outputs --**

1. **What are the data pre-processing / cleaning techniques you would apply?**   
   **Ans –** here in this problem where we have to predict price, I think there is nothing to clean. But we can consider to remove some data points which have unpredictable spikes in between the graph. For preprocessing we would convert data column to two columns where one is day column and another is month column.
2. **What are the features you would use to create the model?**  
   **Ans –** features I will use is markets name and date.
3. **How would you frame this problem as a machine learning problem? What would be the target variable?**  
   **Ans –** here we have to predict price so this is the regression problem means given some features we have to predict real values(price). Here the target variable would be price.
4. **Which algorithm would you use for price prediction?**  
   **Ans –** First I will use simple KNN algorithm and then after seeing error, I can see error on multiple regression problem then after seeing all the results and do hyperparameter tuning I can select one algorithm.
5. **What would be the loss function you would use?**  
   **Ans –** loss function I will use is mean squared error.