Sales Forecast Analytics

DATA SCIENCE PRODEGREE PROJECT

In this project, you will analyze and forecast sales of pesticides for a domestic pesticide manufacturing company (ABC Manufacturing) and analyze variables that impact the sales of the product. This will enable the pesticide company to manage the production to meet growth/slowdown in sales as per your sales forecast.

We Have Done Following things in this project

* Conduct state wise trend analysis of the given sales
* Calculating overall product sales of multiple chemicals by all the manufacturing companies
* Calculated most sold chemical in different Quarters
* Identify State level chemical sales in different Quarters
* Identifying which all chemical does ABC Manufacturing produces
* Identifying Which all states do ABC Manufacturing Dominant for which chemical
* identifying other reasons if they are influencing Pesticide sales
* Forecast of Pesticide sales for ABC Manufacturing

**Step 1 :** We have given a numeric number to every month to make the data ready to convert easily into quarter

**Code used:**

month\_num= {'MONTH':{'Jan': 1,'Feb': 2,'Mar': 3,'Apr': 4,'May': 5,'Jun': 6,'Jul': 7,'Aug': 8,'Sep': 9, 'Oct': 10,'Nov': 11,'Dec': 12}}

df\_sales.replace(month\_num,inplace=True)

df\_sales.head()

**Output:**

FIN\_YEAR MONTH STATE FG DISTRICT COMPANY VALUE

0 2014-2015 11 Haryana Insecticides Rohtak monous 675.00

1 2014-2015 11 Haryana Insecticides Rewari monous 73.71

2 2014-2015 11 Haryana Insecticides Gurgaon monous 405.00

3 2014-2015 11 Haryana Insecticides Mewat monous 1188.00

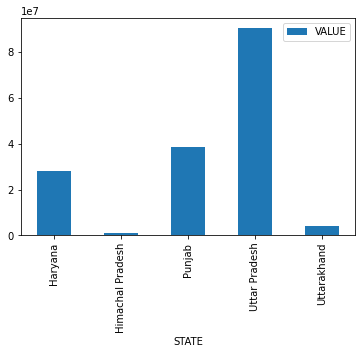
4 2014-2015 11 Haryana Insecticides Palwal monous 5188.59

**Step 2:**  We have now convert the data into quarter and grouped it on state level by using group by function to check the overall sales at state level of different pesticides

**Code Used**: df1=df\_sales.groupby('STATE').sum()['VALUE'].reset\_index()

df1.round()

**Output:** Herein we can clearly see that Uttar Pradesh is leading In terms of overall sales of pesticides with a value of 90327653 followed by Punjab with 38632144 and Haryana with 28004513

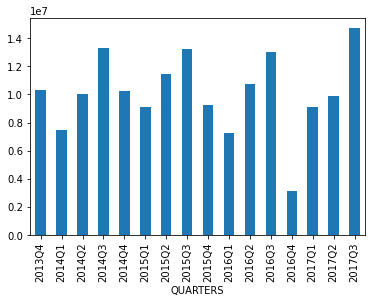


.

**Step 3:**  we have grouped all the quarters to check in how the overall quarterly sales has been every year.

**Code used:** df2=df\_sales.groupby('QUARTERS').sum()['VALUE'].round()

**Output:**  we can see now that quarter 3 of FY 2017-18 has witnessed the highest sales.

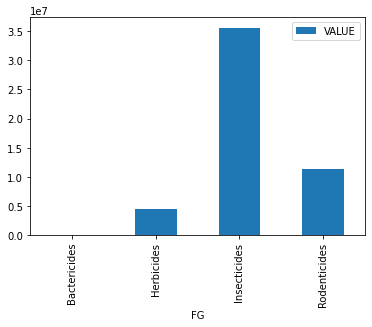
****

**Step 4:** We have now used group by to combine it on pesticides level and checked how the sales has been for every pesticide and which is highest selling pesticide.

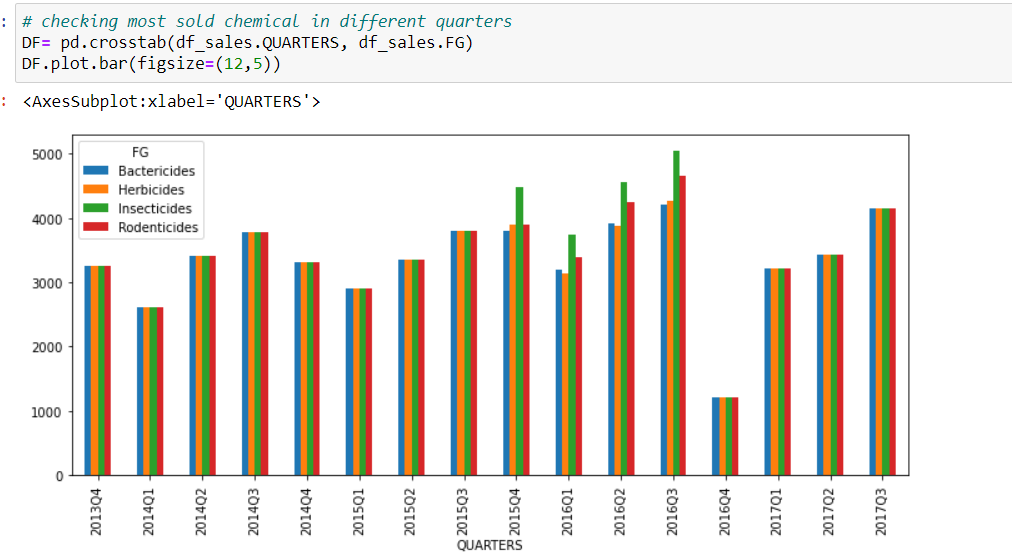
Code Used: df4=df\_sales[df3].groupby('FG').sum(['VALUE']).round()

df4.plot.bar()

Output: we can clearly see that insecticides has been the highest selling pesticides out of the given 4

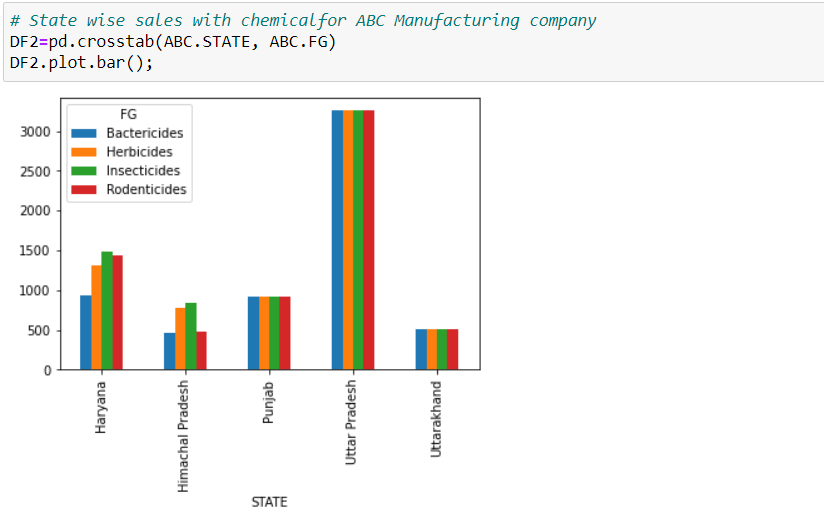


Also, we are checking which is the most sold chemical in different Quarters



Here We can clearing see that in 2016Q3 Insecticides is the most sold chemical in all over the data, after that Rodenticides is the second most sold chemical in the same Quarter.

Also, in below graph we can see that Uttar Pradesh is having higher all chemical uses for ABC Manufacturing Company



**Step 5:**  Now we have checked whether the data is stationary or not. This is done with the help of ADFULLER test

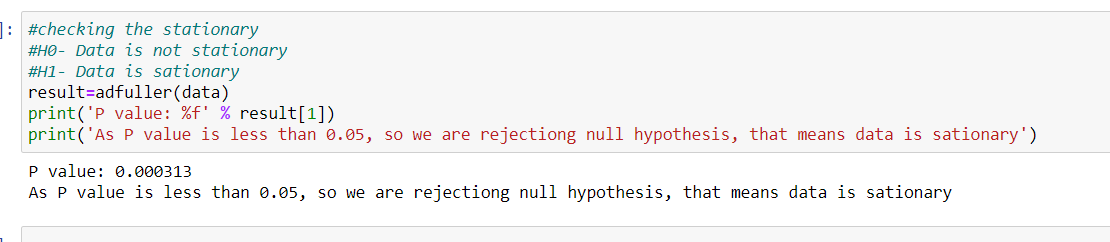
Code used: result=adfuller(data)

print ('P value: %f' % result[1])

print ('As P value is less than 0.05, so we are rejecting null hypothesis, that means data is stationary')

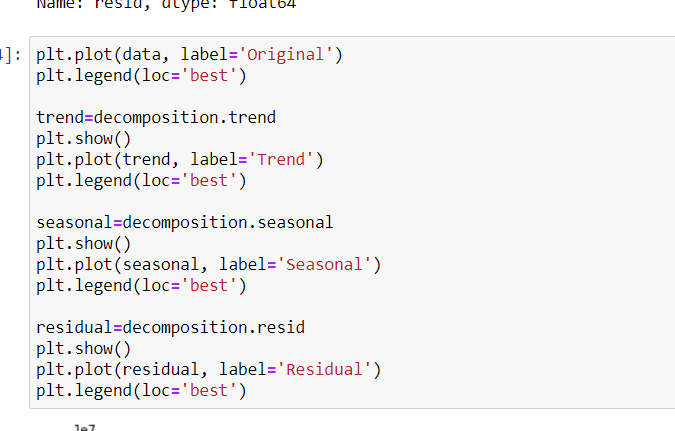
**Output:** P value: 0.000313

As P value is less than 0.05, so we are rejecting null hypothesis, that means data is stationary

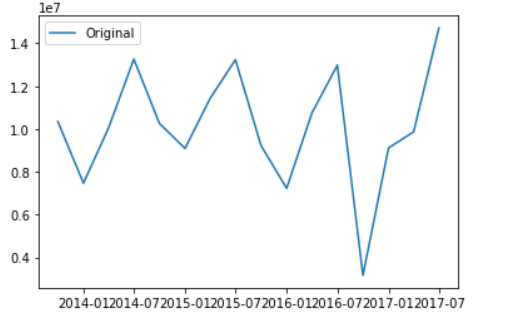


**Step 6:**  we have now decomposed the data to check about the seasonality, trend and the noise in the data.

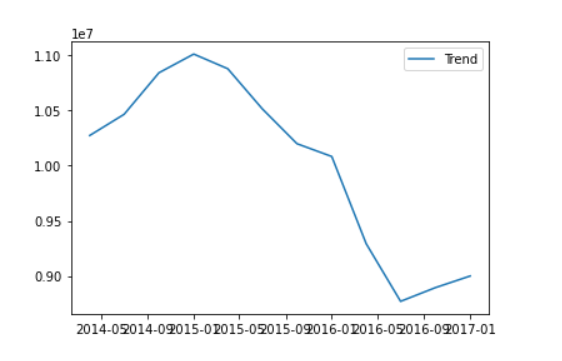
Code Used: following is the code used to check the seasonality in the data



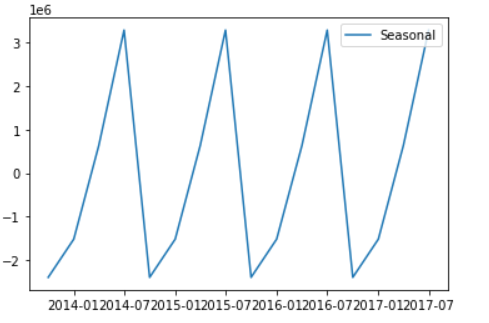
**Output:**



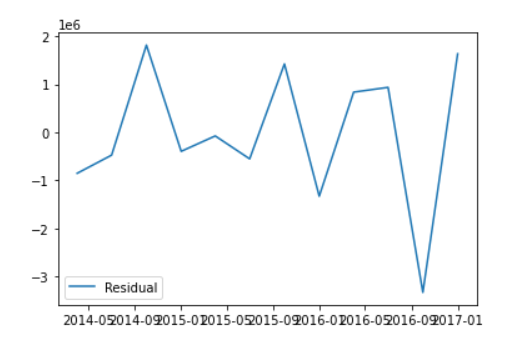
**Trend:**



**Seasonality:**

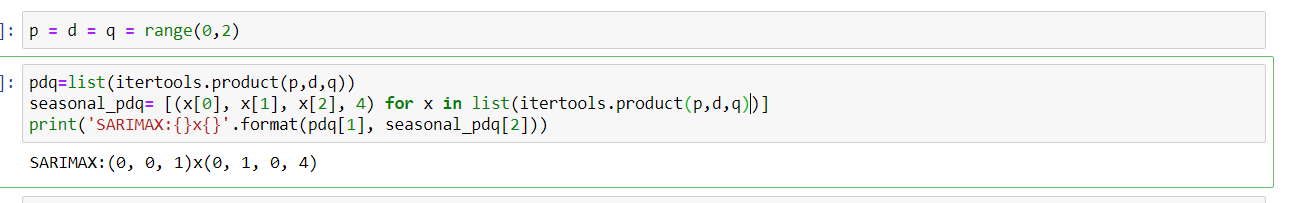


**Residual (Noise):**

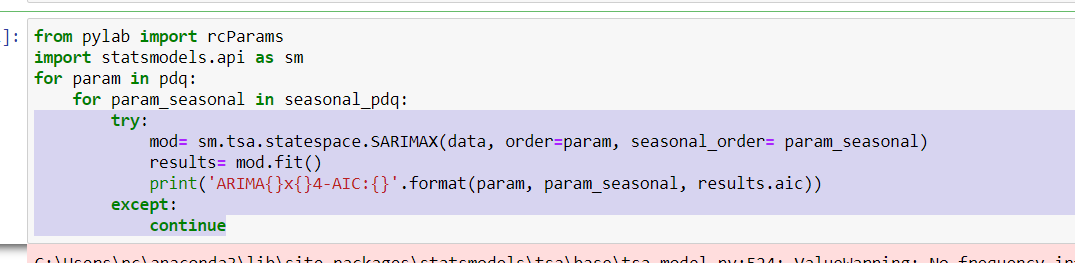


**Step 7:** now trying to find the best value of p, d, q where p stands for number of Autoregressive term, d stands for nonseasonal differences and q stands for no. lagged forecast error.

Code used: p=d=q =range(0,2)

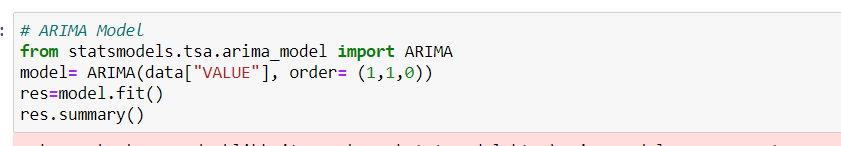


and entered it in the SARIMAX model,

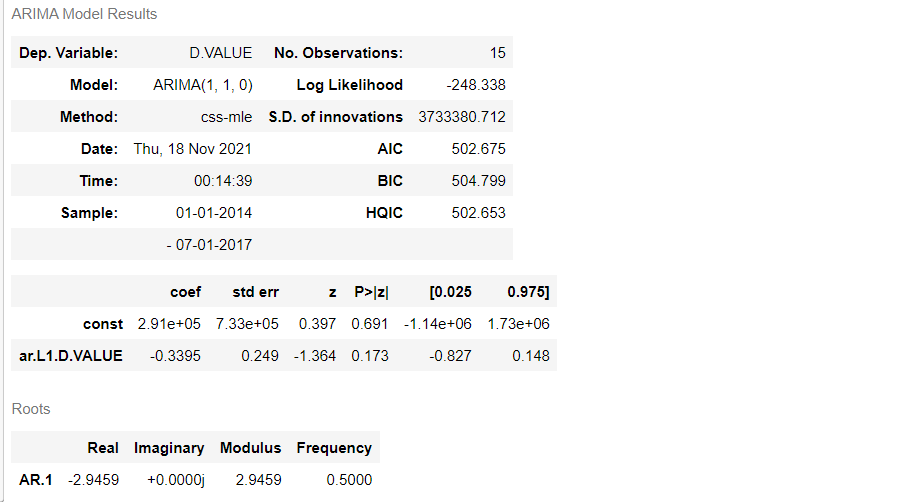


For (1, 1, 0)x(0, 1, 0, 4) we got lowed AIC value AIC: 361.9580657646517

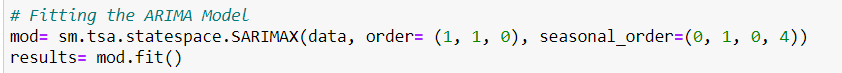
**Step 8:** As we have p, d, q value, so we entered it in ARIMA Model.

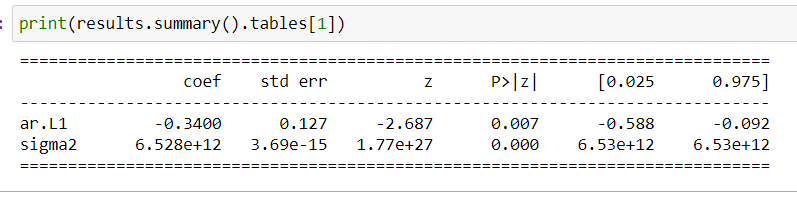


And got below matrices but as per the seasonality graph data have some seasonality, in this case Seasonal ARIMA model will be best choice for prediction.

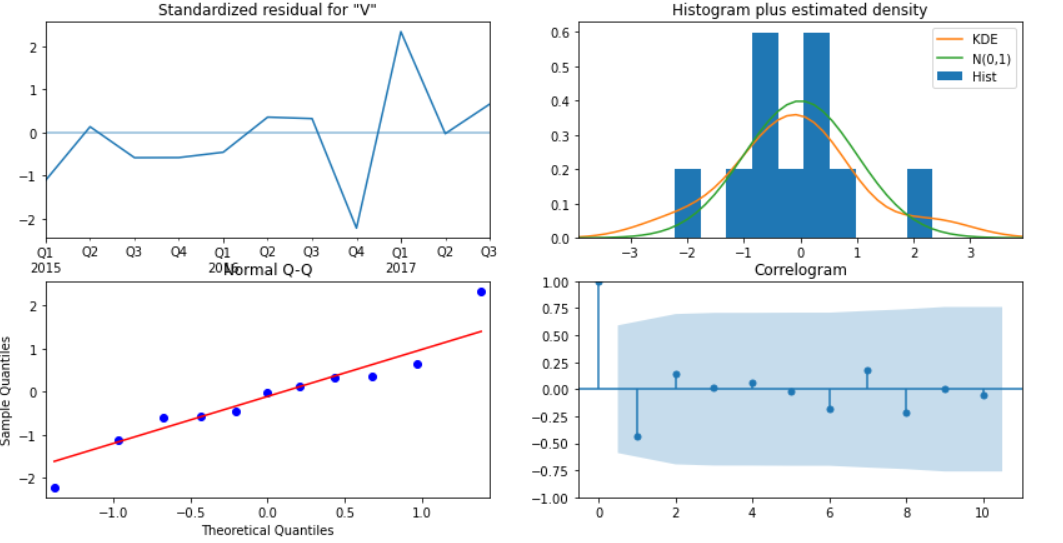


**Step 9:** Now we have p, d, q & seasonal P, D, Q value in the interval of quarters, so we are entering the final values in SARIMAX model and got output

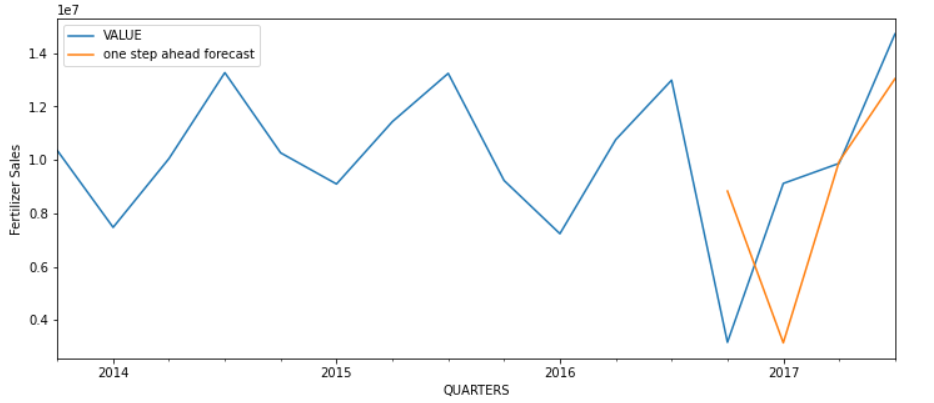




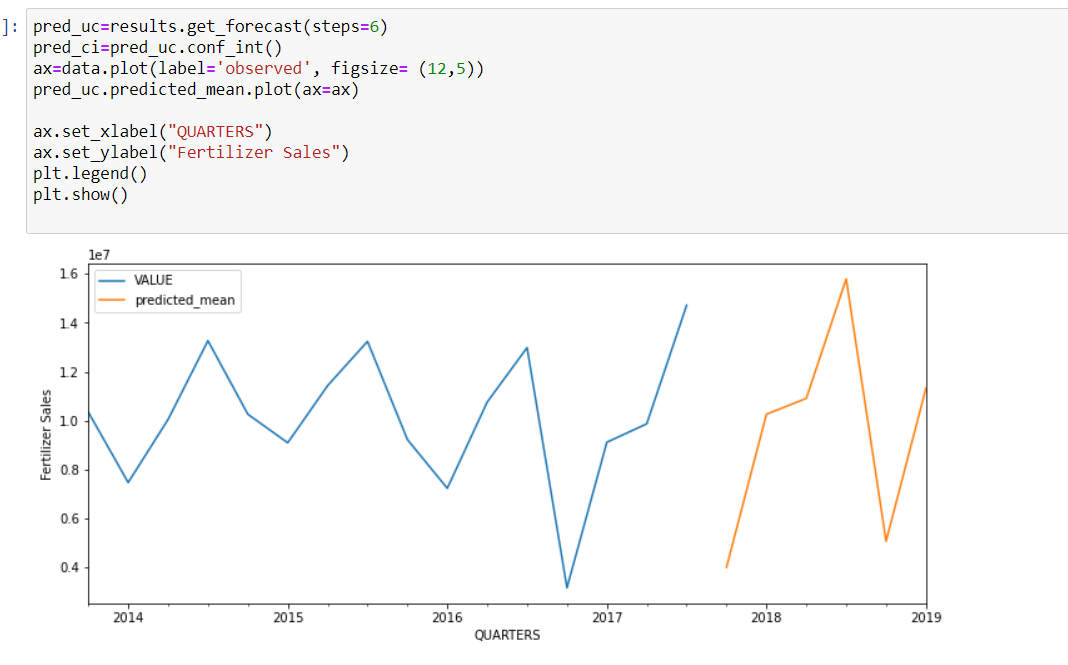
**Step 10**: now from the model got below diagnostics, by reviewing the Q-Q plot we can clearly say that the data is normal distributed which have few outliers.



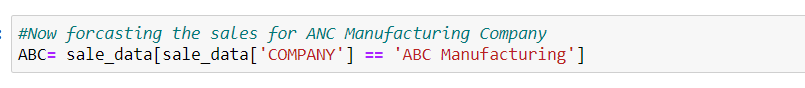
**Steps 11: now predicting the value form 2016-10-01 in the original data and comparing the actual data**



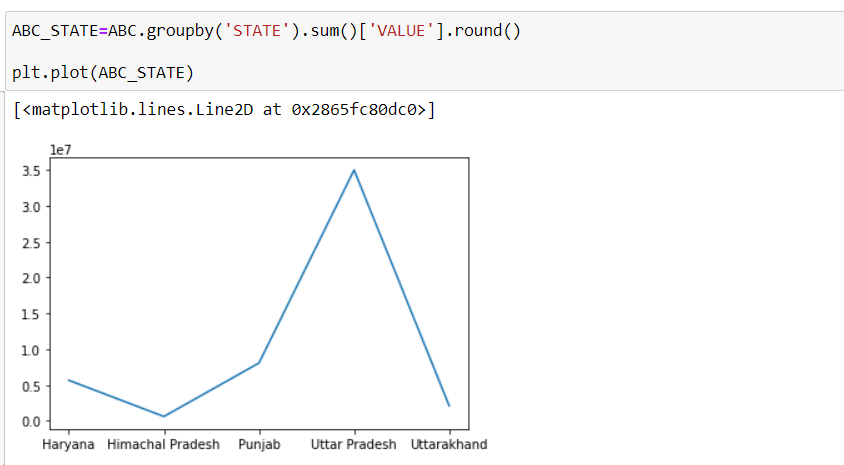
**Step 12: Forecasting the value for next 6 quarters.**

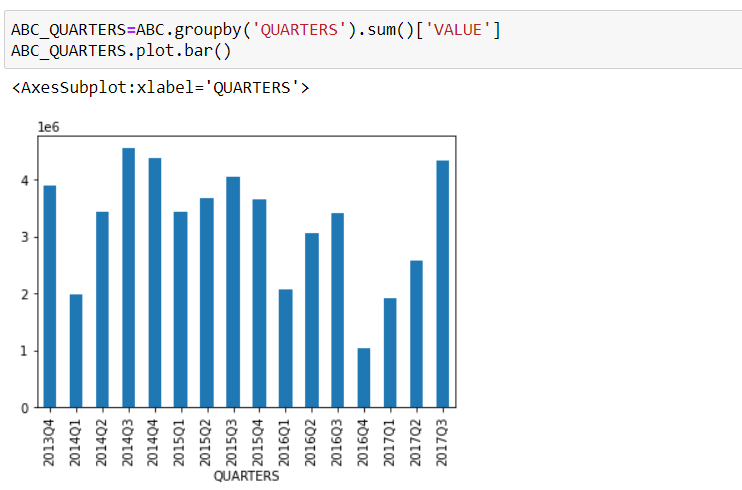


**Step13: Now forecasting the sales for ABC Manufacturing company, for that first filtered the data for ABC Manufacturing company.**

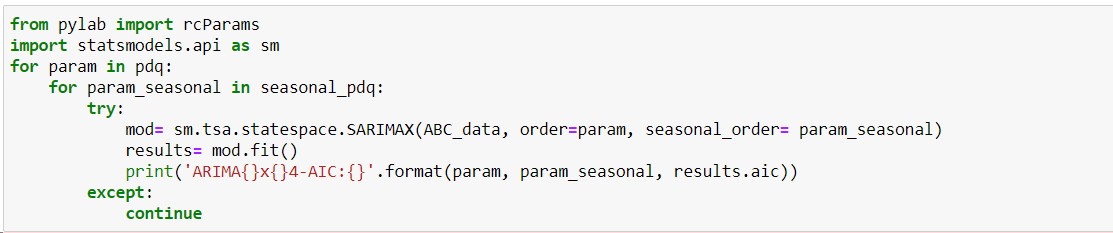


**Step14: checking the State wise & Quarter wise company sales.**

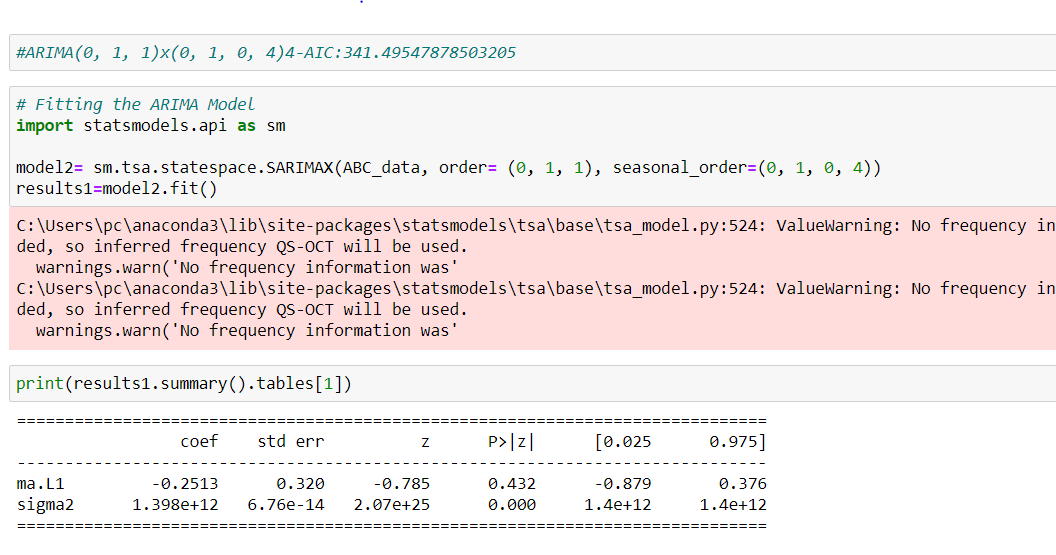




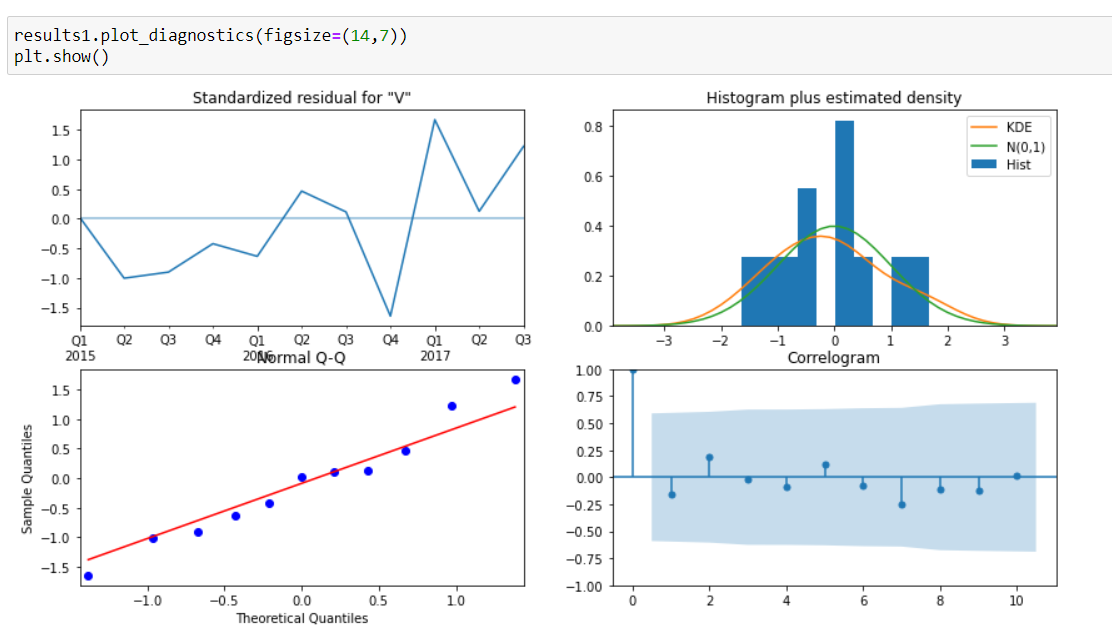
**Step 15: setting the Quarters as index and again calculating p, d, q values for ABC company**



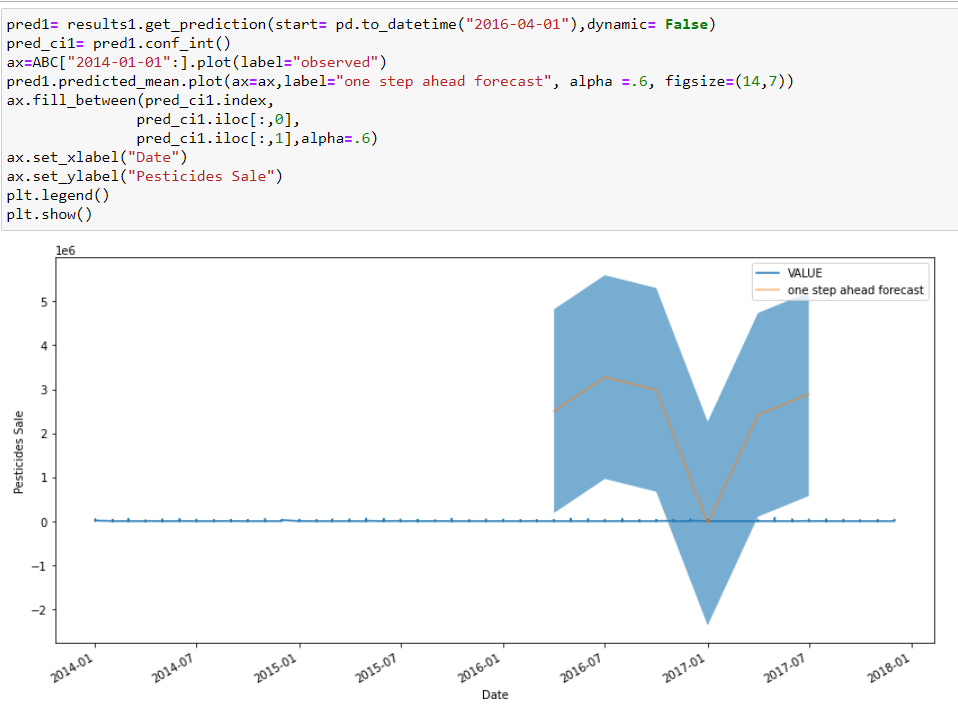
**Got p, d, q and P, D, Q values than by using SARIMAX model, completing the modeling part**



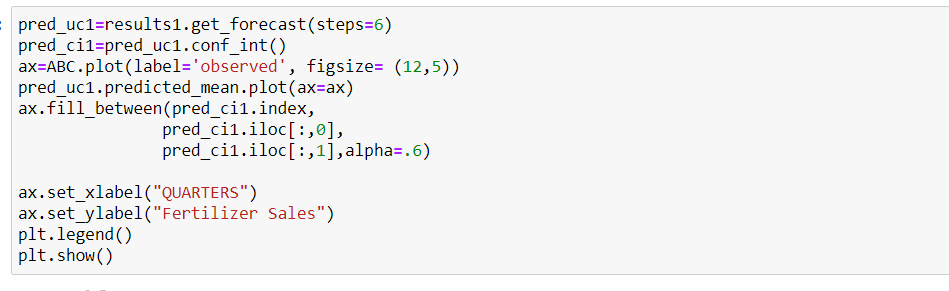
**Finding Diagnostics for ABC Company,**

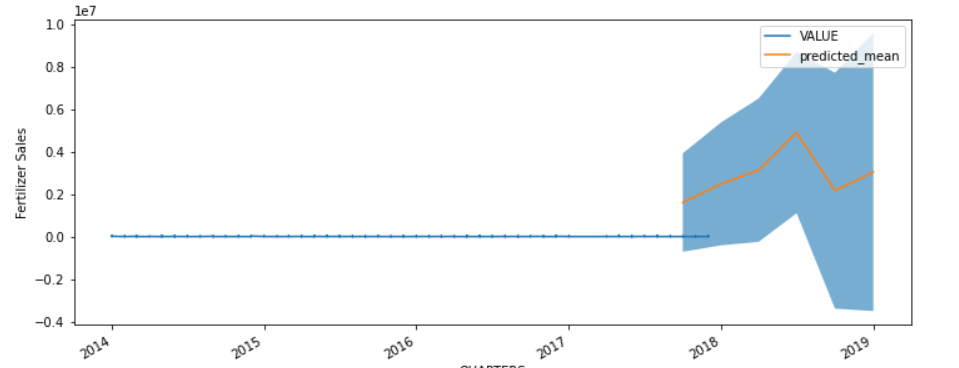


**Step17: Predicting sales from 2016 to 2017**



**Step 18: forecasting sales for next6 quarters,**





**Conculsion:**

**So as we can see the actual chemical sales is up in every third quarter that is between October to December, which is because of multiple factors like from oct-Dec is the harvesting season in India which affecting the chemical sales. Also, in every first quarter from April- June sales have down due to summer season as there will be no rain or no monsoon,**

**If we want to see state wise chemical uses so we can see in above graphs that in Utter Pradesh the Consumption of chemical is much high comparing to other states and they have equally uses all four chemical in Utter Pradesh and we want to see the chemical wise sales the uses of Insecticides is much high in comparison for other three chemical.**