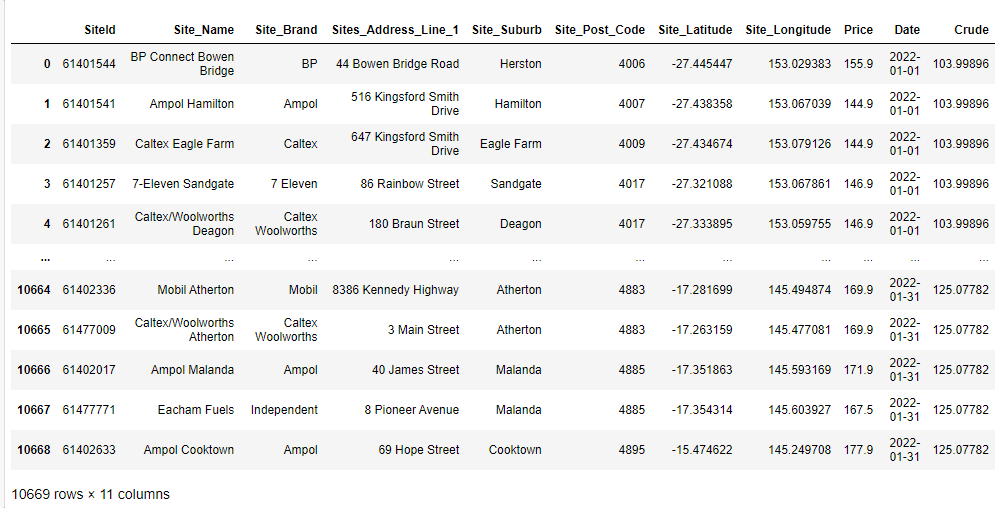
**Introduction:**

The following is a detailed report outlining the causes and predictions for the changing Oil prices. We have taken insights of Oil Retail Prices in a single month (01-01-2022 to 31-01-2022) of the Queensland dataset or you can find it on Kaggle.

The price change can be found by using regression models, particularly Linear Regression which we have applied, but before that, we have to pre-process and analyze data and its trends. The following is our dataset consisting consists of 12 columns & 10669 rows. We can perform EDA to take out further insights into the Dataset.

**Data:**

This is the given Queensland dataset.



**Data Wrangling:**

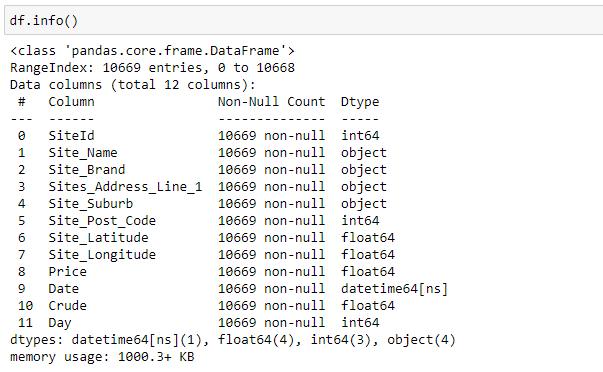
Add a new column Day extract from date to find out the results on daily basis.

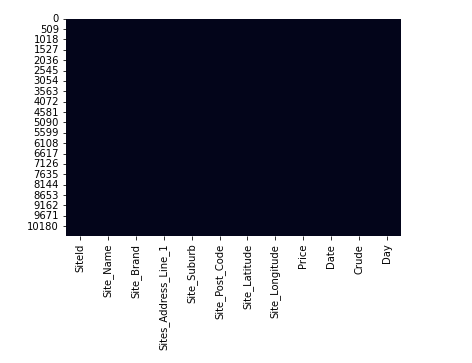


Now the Dataset adds one column Named “Day”.



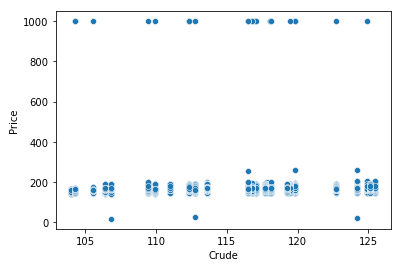
Perform the following commands and make a graph to view null values. This data contains not a single Null Value.





**Exploratory Data Analysis:**

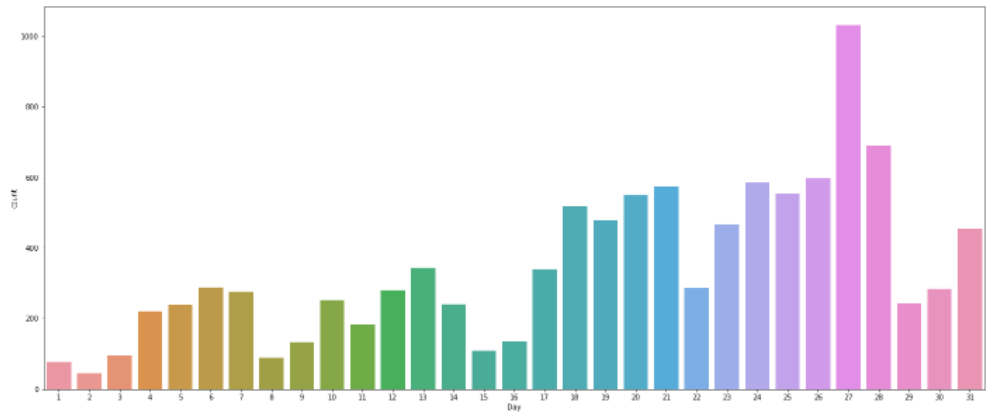
Look at data trends that can show the outliers and not provide accurate trends which can relate to other features.



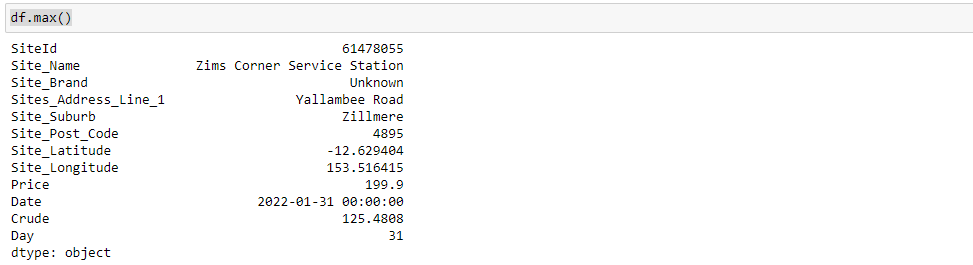
Next, we remove outliers present in a dataset that can affect our data trends after removing outliers it can provide clear insights.



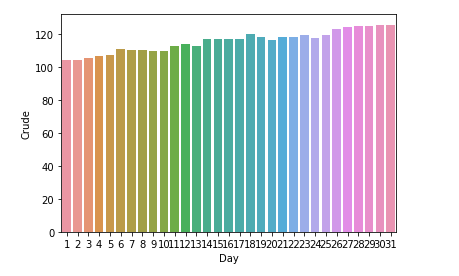
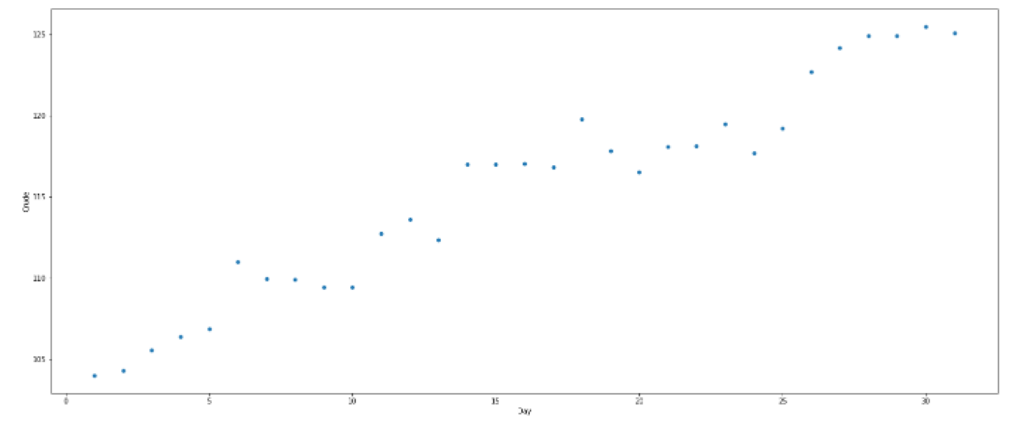
The graph provides information on how crude prices change daily. 27 of January shows the high price of Crude.



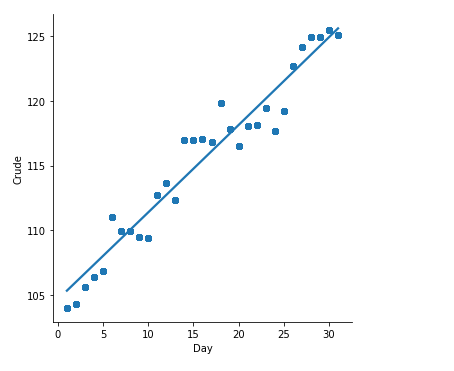
After removing the outliers the values vary in the range. It can show a maximum range in a dataset.



Now, this scatterplot can show a trend of crude prices that can change and continuously increase daily.

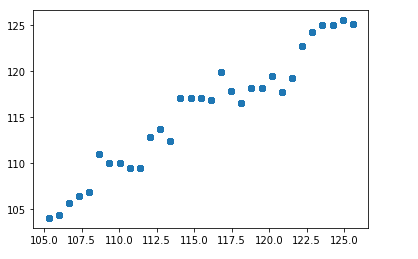


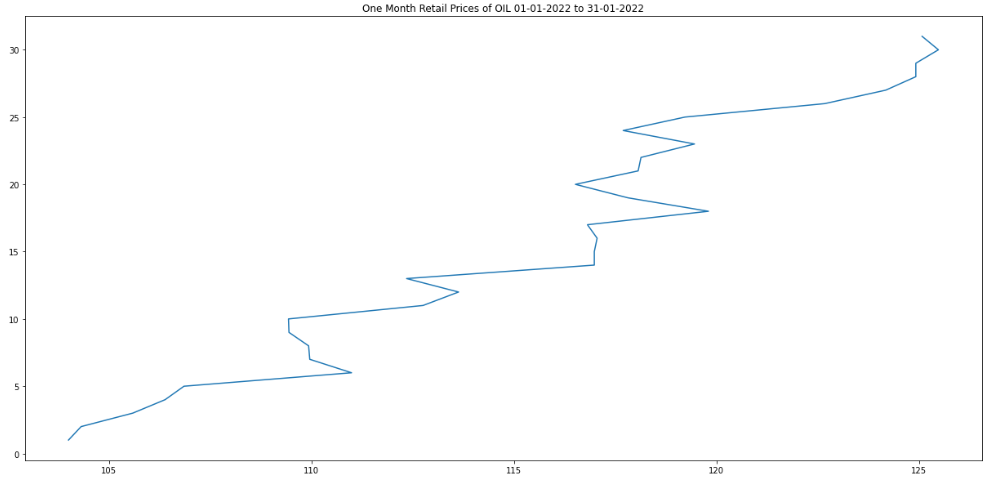
Plot a linear line in the given data set to verify the trends and whether the linear regression is acceptable or not in this acquired dataset.



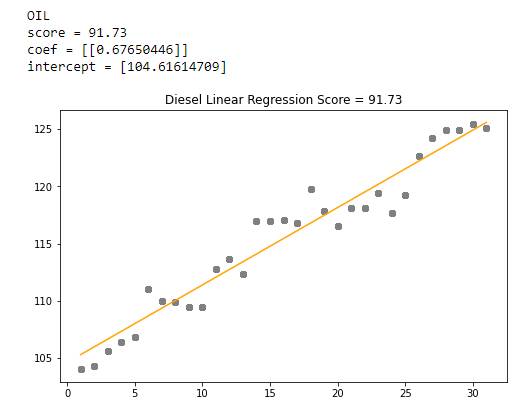
**Deploy Model:**

Apply Linear Regression on a training dataset and it can show the given results on a testing dataset which can be quite similar to the training data set.





The below graph can show a result 91.73% score with a coefficient of 0.67 and indicates a periodic rise in price according to the coefficient. Plot a training data set with scatterplot and prediction linear line with the testing dataset. It can produce more than 90% accuracy on a given dataset.



**Conclusion:**

From the above diagram, we have to find the trends of oil price which seems to have quite fluctuated. However, it has experienced a bottom to an uptrend which is what our model predicted too. Our model performed quite well on the test data for 2022 with a regression score of 91.73% and a coefficient of 0. 67 which is acceptable.