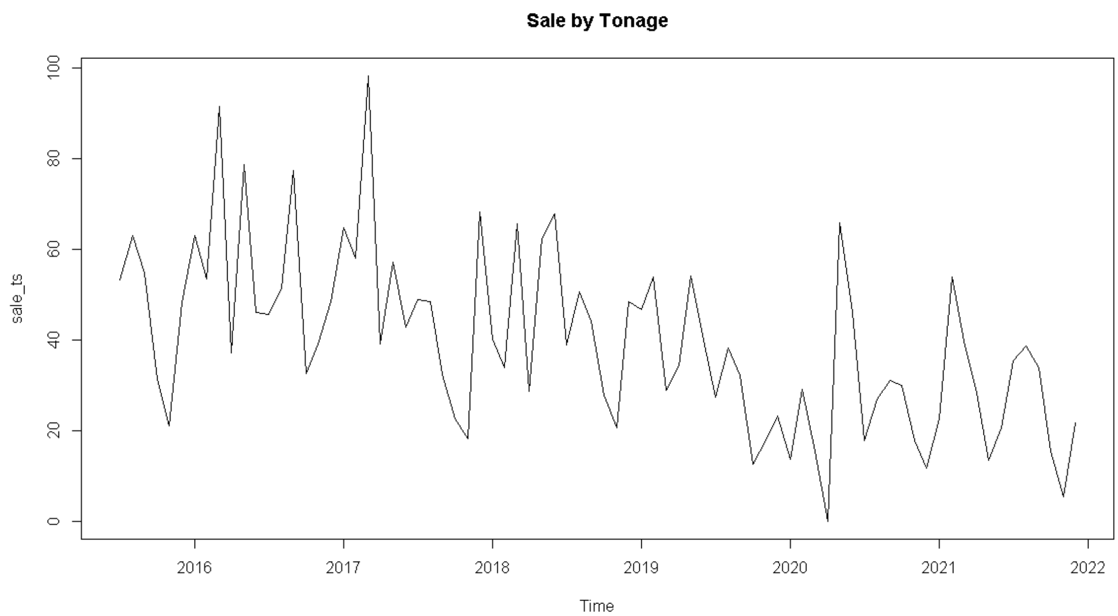
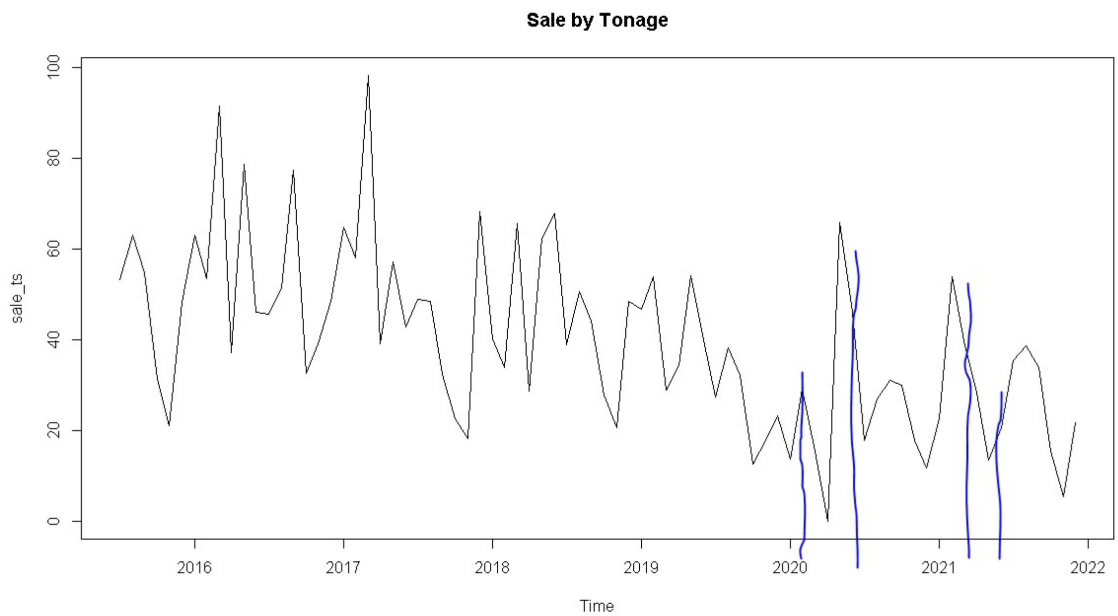
In March 2020, due to Covid-19 the Government of India was forced to enforce a complete lockdown which immensely affected the Indian Market and all the businesses. The first lockdown was relaxed in June 2020. In the following year, the second Covid-19 lockdown was implemented in April 2021 and was lifted in June 2021. Incorporated in 1987, X is one of the India's largest and highly regarded PVC pipe manufacturing and multi-polymer processors. They've been awarded the Brand of the Year – Pipes Award at INEX Reality+ Awards 2021. Y is a Pipes and Fittings Company which had a stronghold in South India was taken over by X in the Year 2017. In 2019 the company became a public limited company and was open for investors to invest in the company. Below is the graph showing us the overall trend of sales from 2015 to 2022 of Y. The considerable ups and down in the sales caused by the lockdowns made it difficult for the company to plan their production which led to the company facing a really high number of MNA (Material Not Available). As a solution to overcome this issue, the company decided to use the Forecasting method (Machine Learning). After a little R&D, it was discovered that the Holt-Winters model was the best fit for the company's data. In this article we are about to see how the company used the Holt-Winters model to comprehensively plan the production of more than 7,167 products and extending this even to the distributor level which consists of 766 distributors across India.

Below is the graph with respect to the Tonnage from July 2015 to Dec 2021.



For better understanding. let us look at the covid affected regions.

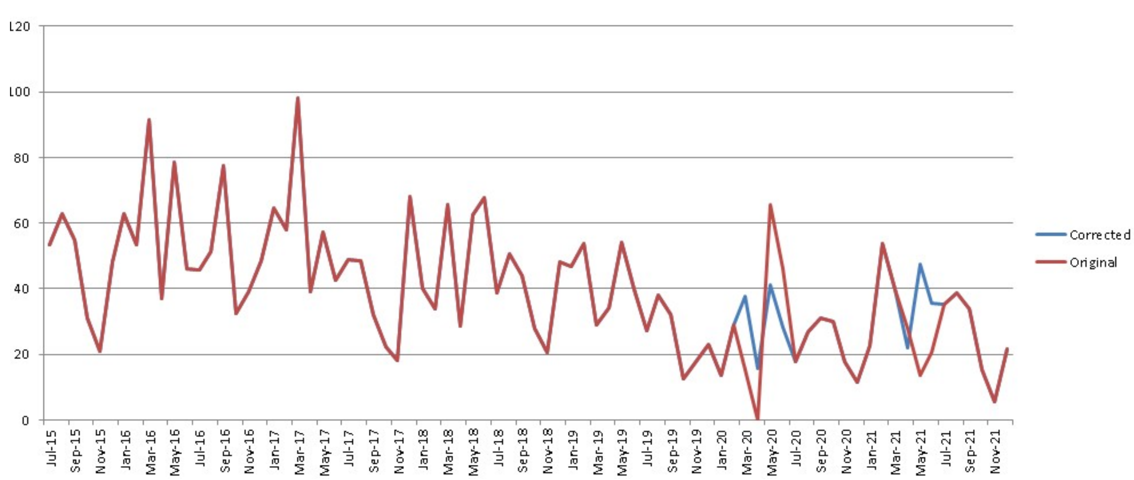


The region between the first 2 blue lines from the left is the lockdown affected region during the first lockdown (March 2020 to June 2020) and to the right is the lockdown affected region during the second lockdown (April 2021 to June 2021).

In order to do a successful forecast, we need to minimise the effect of lockdown as much as possible because the effect cannot be nullified completely. So we decided to do a forecast of the lockdown affected region and replace the original values by the forecasted value to minimise the effect of lockdowns.

Data Preparation:

The diagram shown below is the Data Processing goal we had to achieve.



Replacing the lockdown affected values with the possible values had the lockdown not occurred. The red trend denotes Original values and the blue trend denotes the Corrected values.

We used Machine Learning in R to achieve the above.

First, we need to Import a few Libraries.

```
library(forecast)
library(TTR)
library(tseries)
library(dplyr)
```

Later we Import the Data

df = read.csv("C:\\Users\\rdp\\Documents\\Excel\\"File Name" - Sales Data.csv")

We use Select to Shortlist the Important Columns Required for our time series forecasting and then we Summarise the Tonnage sold during the same Period

```
impds = df %>% select(Period, ItemTn)

impds = aggregate(.~Period,data=impds,FUN=sum)
```

We check if the Data is Prepared to be fitted inside the modal

```
head(impds)   # V
tail(impds)   # V



impds   ## Cutting
```

We Split the Data into 5 different parts

Part 1: Before the First Lockdown

Part 2: During the First Lockdown (March 2020 to June 2020)

Part 3: Between the First and the Second Lockdown

Part 4: During the Second Lockdown (April 2021 to June 2021)

Part 5: After the Second Lockdown

```
##_________Splitting_The_Data_into_5_parts___________________________________

impds1 = impds[1:56, ]  ##-Before Covid 1st Lockdown from July 15 to Feb 20

head(impds1)  ##View Top 6 ROWS
tail(impds1)  ##View Bottom 6 ROWS

##Convert Into Time Series Data
sale_ts = ts(impds1$ItemTn,start = c(2015,7),end = c(2020,2),frequency = 12)
sale_ts

plot(sale_ts) ##Plot Time Series Data

decompose(sale_ts) ##Decompose to Observe Trend and Seasonality
plot(decompose(sale_ts)) ##Plot the Decomposition

adf.test(sale_ts) ## Augmented Dicky Fuller Test


HW_temp_forecast = HoltWinters(sale_ts) ## Fitting the data into Holts Winter
HW_temp_forecast
plot(HW_temp_forecast) ##Plot Holts Winter Fit


HW_pred_model = forecast:::forecast(HW_temp_forecast,h=4) ##Forecast of next 4 Months
HW_pred_model
plot(HW_pred_model) ##Plot Forecast
```

We Split the data and use Part 1 depending on the Decomposition of part 1 we learn that Trend and Seasonality Exists hence we decide to go ahead with the Case 3 of Holts Winter where Alpha = TRUE , Beta = TRUE and Gamma = TRUE (Level, Trend and Seasonality Exists)

We Get

```
> plot(HW_temp_forecast) ##Plot Holts Winter Fit
> HW_pred_model = forecast:::forecast(HW_temp_forecast,h=4) ##Forecast
> HW_pred_model
         Point Forecast      Lo 80    Hi 80      Lo 95    Hi 95
Mar 2020       37.53941 18.572231 56.50658   8.531615 66.54720
Apr 2020       15.85833 -3.272845 34.98951 -13.400280 45.11695
May 2020       41.09012 21.770472 60.40977  11.543267 70.63698
Jun 2020       28.20891  8.675137 47.74268  -1.665418 58.08324
```

We have to replace these Forecast Values with the Original Values to do so we first need to convert the Forecast Values to a Data frame

```

impds12 = as.data.frame(HW_pred_model) ##Convert the Forecast values to Dataframe


colnames(impds12) <- c('ItemTn','a','b','c','d') ##Change the names of the columns


impds12 = subset(impds12, select = -c(a,b,c,d) ) ##Delete unwanted Columns

impds12$Period = seq(57,60) ##Add Separate column for Period

```

We Convert the forecast values to a data frame and then we change the column names, we remove the unnecessary columns like High and Low Confidence Interval and add a separate period column

Now we have to replace the forecast value with the original value

```
##______________Minimising_the_affect_of_first_lockdown______________

clds1 = impds[57:60, ]   ##-Covid 1st Lockdown Affected Dataset fro
head(clds1) ##View top 6 Rows of the 1st Lockdown Period

ids = merge(x = clds1, y = impds12, by = "Period", all.x = TRUE) ##
ids ## Call the Merged Table

ids = subset(ids, select = -c(ItemTn.x) ) ## Delete the Unwanted Co

colnames(ids) <- c('Period','ItemTn') ##Replace the Original Value

impds_new = rbind(impds1,ids) ##Append the new Table

head(impds_new) ##View Top 6 Rows of New Data
tail(impds_new) ##View Bottom 6 Rows of New Data

impds_new ##Call the New Data
plot(impds_new) ##Plot the New Data
```

We callout the Part 2 data that is of the first Lockdown and we merge the 2 data frames creating a separate data frame after which we delete the column from the Original data frame and Replace the Original Values with the forecast Values and then Append the part 2 df to part 1 df (Add the part 2 Data frame below part 1 Data frame) and create an entirely new Data Frame

Moving on to part 3

```
##_____________Append_the_next_Unaffected_timeline___________________________________

impds2 = impds[61:69, ]  ##-After 1st Lockdown and before 2nd Lockdown from Jul 20 to
head(impds2)
tail(impds2)

impds_new = rbind(impds_new,impds2) ##Append
plot(impds_new) ##Plot Updated
impds_new ##Call Updated


## Converting in Timeseries Data
sale_ts2 = ts(impds_new$ItemTn,start = c(2015,7),end = c(2021,3),frequency = 12)
sale_ts2
plot(sale_ts2)


##Decompose to Observe Trend and Seasonality
decompose(sale_ts2)
plot(decompose(sale_ts2))
```

Part 3 is the unaffected timeline which took place between the first lockdown and the second lockdown we append (add below) the part 3 timeline to new data frame and then convert it to time series data after which we decompose the data and learn that Level, Trend and Seasonality Exists and hence we again use Case 3 of Holts Winters Model and Repeat the same process for forecasting the value of Part 4 which is the affected timeline of 2nd lockdown replacing the values of the part 4 re append part5 and now our data is ready for forecasting

We get

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2022	23.553619	6.295784	40.81145	-2.8399622	49.94720
Feb 2022	43.229575	25.674628	60.78452	16.3815999	70.07755
Mar 2022	38.302900	20.442481	56.16332	10.9877461	65.61805
Apr 2022	20.035247	1.861123	38.20937	-7.7596771	47.83017
May 2022	45.370046	26.874116	63.86598	17.0829622	73.65713
Jun 2022	33.498205	14.672500	52.32391	4.7067733	62.28964
Jul 2022	27.241100	8.077783	46.40442	-2.0666648	56.54887
Aug 2022	32.551632	13.042999	52.06026	2.7157531	62.38751
Sep 2022	30.062581	10.201064	49.92410	-0.3129883	60.43815
Oct 2022	15.415768	-4.806072	35.63761	-15.5108674	46.34240
Nov 2022	7.557599	-13.031869	28.14707	-23.9312745	39.04647
Dec 2022	19.009761	-1.954511	39.97403	-13.0523262	51.07185

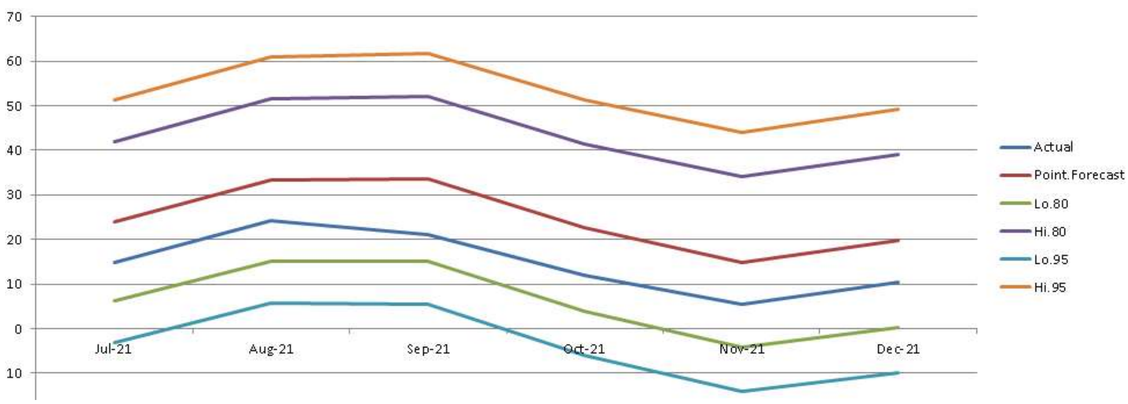
```
> plot(HW_pred_model)
```

Let us check the Accuracy of the Model By splitting the Data into Train and Test

Train: July 2015 to June 2021

Test: July 2021 to December 2021

Over here we fit the Holts winter Forecast Model in the Train Data for the above mentioned time period and calculate the forecast for next 6 months and compare it with the Test Data for the above mentioned time period



As we can see that The Red Line is the Forecast trend and Blue is the Actual Value as we can observe from the Graph Our Predicted Value lies under 80% Confidence Interval and RMSE of 10.11375

Using this forecast Model the Company would plan their production for the coming Future