CAPSTONE PROJECT

Retail Stores

DHEERAJ PATEL
INTELLIPAAT

Tables of Content

Problem Statement

A retail company with multiple outlet stores is having bad revenue returns from the stores with most of them facing bankruptcy. This project undertakes to review the sales of records from the stores with a view to provide useful insights to the company and also to forecast sales outlook for the next 12-weeks

Project Objective

The Retail Company with multiple outlets across the country are facing issues with inventory management. The task is to come up with useful insights using the provided data and make prediction models to forecast the sales of next twelve weeks

Data Description

The available dataset contains total entries of 6,435*8 having 6,435 rows and eight columns as shown in the Table below:

Feature Name	Description
Store	Store Numbers
Date	Week of Sales
Weekly_Sales	Sales for the given store in that week
Holiday_Flag	If it is a holiday week
Temperature	Temperature on the day of the sale
Fuel_Price	Cost of the fuel in the region
CPI	Consumer Price Index
Unemployment	Unemployment Rate

Data Description, various insights from the data.

From the given dataset of the retail company, it is observed that the data consists of six thousand four hundred and thirty-five (6,435) records with seven features (recorded weekly) as follows:

1. Stores: There are 45 stores and each store has 143 entries of:
2. Date of record (weekly),
3. Total sales record for the week,
4. Holiday flag for the week (1 or 0),
5. Temperature: average temperature recorded during the week
6. Fuel Price: average fuel price for the week,
7. CPI: average Consumer Price Index for the week
8. Unemployment: rate of the unemployment for the week of record

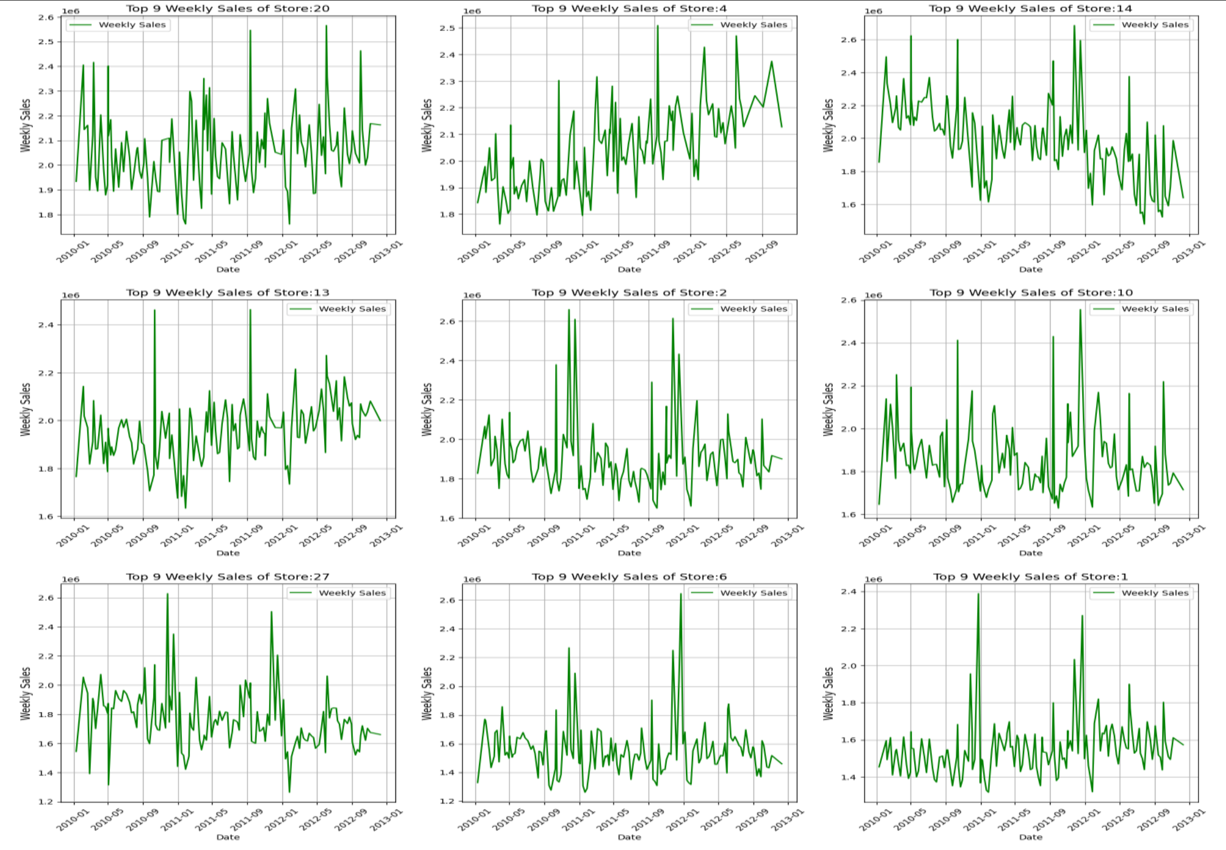
Data Pre-processing Steps and Inspiration

The pre-processing of the data included the following steps:

1. Step 1: Load Data
2. Step 2: Perform Exploratory Data Analysis
    a. Confirm number of records in the data and how they are distributed
    b. Check data types
    c. Check for missing data, invalid entries, duplicates
    d. Examine the correlation of the independent features with the target (Weekly_Sales) variables.
    e. Check for outliers that are known to distort predictions and forecasts.
3. Step 3: See relations between independent and dependent variables and make inferences.
4. Step 4: Model Predictions, two approaches:
    a. Linear Regression Models.
    b. Time Series Model (ARIMA, SARIMAX).
5. Step 5: Forecast
6. Step 6: Compare result from different models

Model Evaluation and Technique

Plot of Weekly Sales for Selected Stores



Model Selection:

Examination of the plot of the target feature, Weekly Sales as shown above shows a time varying data.

A Linear Regression models (Linear Regression, Gradient Boosting, Random Forest) for predictions. Attempt will be made on Time Series model (ARIMA, SARIMAX) will be employed for the predictions and forecast and compare results with Linear Regression Predictions.

1. **Regression Models**

a. **Linear Regression** is a basic predictive analytics technique that uses historical data to predict an output variable and is used to estimate the relationship between variables. It is a popular algorithm employed to predict continuous (dependent) variables such as price, based on their correlation with the other independent variables. https://www.mathworks.com/help/stats/what-is-linear-regression.html It is based on the following assumptions:

   i. **Linear Relationship:** The relationship between the independent and dependent variables should be linear.

   ii. **Multivariate Normal:** All the variables together should be multivariate normal, which means that each variable separately has to be univariate normal means, a bell shaped curve.

   iii. **No Multicollinearity:** There is little or no multicollinearity in the data which means that the independent variables should have minimal correlation with each other.

   iv. **Homoscedasticity:** There should be homoscedasticity or 'same variance" across regression lines. In other words, residuals are equal across regression line.

b. **Gradient Boosting:** Gradient boosting stands out for its prediction speed and accuracy, particularly with large and complex datasets. The algorithm has produced the best results from Kaggle competitions and machine learning solutions for businesses. In machine learning algorithm, two types of errors, otherwise called loss functions, are encountered, **bias error** and **variance error**. Gradient boosting algorithm is based on minimizing the bias error or the loss function of the model. The gradient boosting algorithm is based on building models sequentially where the subsequent models try to reduce the errors of the previous model. The subsequent models are built on the errors or residuals of the previous model. The process is repeated until there is no more significant change on the error.

c. **Random Forest:** Random Forest is a commonly-used machine learning algorithm trademarked by **Leo Bremian** and **Adele Cutler**, which combines the output of multiple decision trees to reach a single result. It ease of use and flexibility have fuelled its adoption, as it handles both classification and regression problems.

2. **The ARIMA Models:**

Autoregressive Integrated Moving Average (ARIMA) is defined as a statistical analysis model that uses time series data to better understand the data set or to predict future trends.

https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp#:~:text=Autoregressive%20integrated%20moving%20verage%20(ARIMA)%20models%20predict%20future%20values%20based,to%20forecast%20future%20security%20prices.

A statistical model is autoregressive if it predicts future values based on past values.

ARIMA model is based on a number of assumptions including:

a. Data does not contain anomalies,
b. Model parameters and error term is constant,
c. Historic time points dictate behaviour of present time points,
d. Time series is stationary.

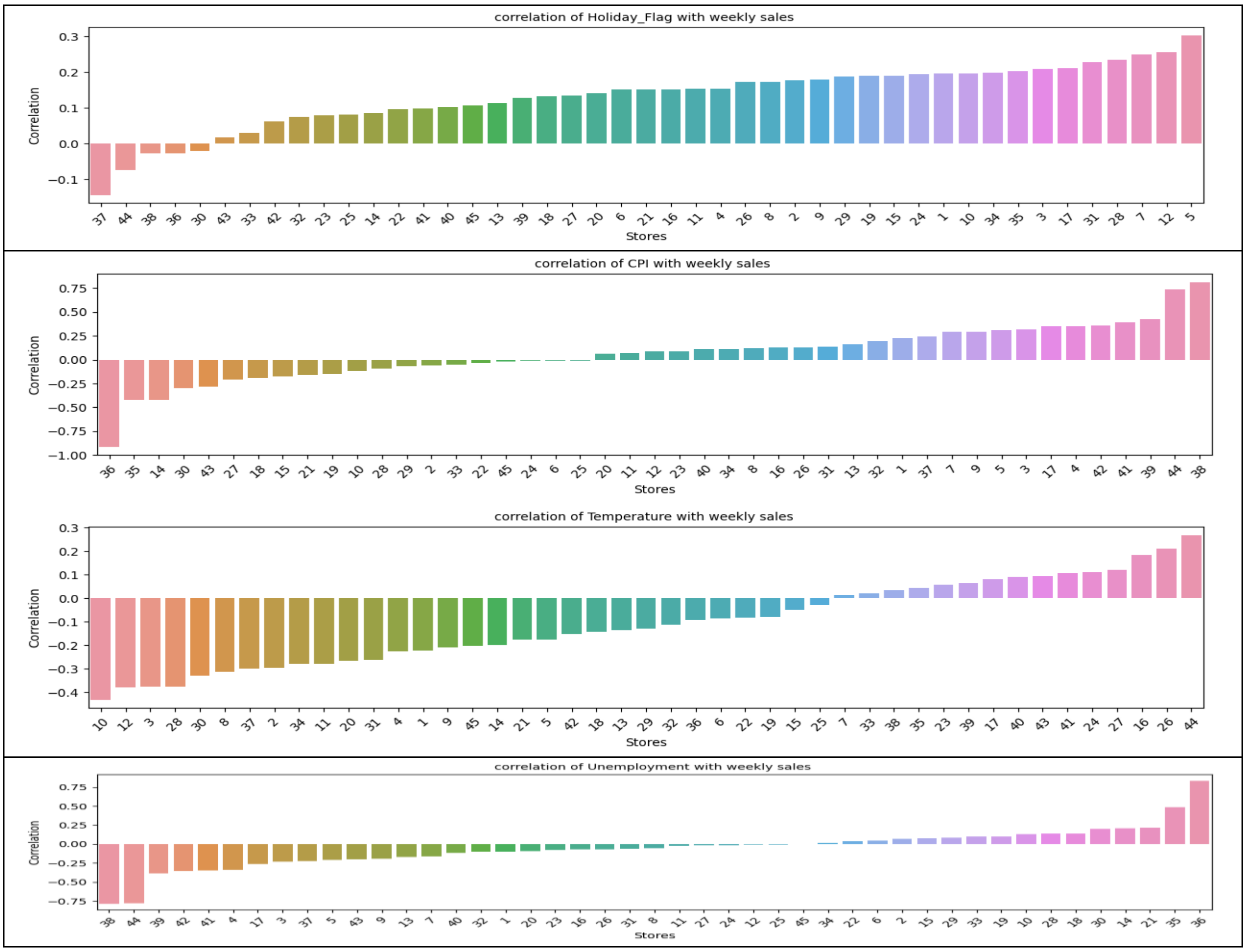
Model Evaluation:

The following techniques and steps were involved in the evaluation of the model

1. Load necessary libraries
2. Load the dataset
3. Perform Exploratory Data Analysis (EDA) on the dataset
   a. Find the shape or size of the data
   b. Check the invalid and null entries
   c. Explore data description
   d. Examine the correlations of the independent variables to the target variable (Weekly_Sales).
   e. Line plot of the effect of the independent variables on the target variable.
   f. Box plot of the features to identify outliers
4. Model Prediction
5. Forecast

Model Evaluation and Technique

Model Design:



It was observed from the EDA that the effects of the independent variables (Unemployment, Temperature, Holiday_Flag, and CPI) on the target variable, weekly sales differ greatly by the store. For example, as shown above the effects of unemployment vary by the stores whereas it appears to have positive effects on some and negative effects on others. The same is also true for Temperature, CPI, and Holiday Flag to some extent.

Model Evaluation and Technique

Premised on the findings, the decision was taken to handle the model, predictions by the stores as a single prediction for all the stores may not be reasonable given the peculiar conditions prevalent in each region of the stores.

For simplicity and ease of presentation, I have also decide to limit my predictions for top eight stores with highest weekly revenue, however the model provided cane be used for predictions for other stores or all stores.

Model Approach:

1. **Regression Models:**
    - Gradient Boosting, Linear Regression and Random forest models were also used for the prediction. The best of the three predictions will then be compared to the predictions by ARIMA or SARIMAX model predictions.
2. **Time Series Model, ARIMA**
    - The first step for this model is to check stationarity of the dataset (p-value less than 0.05).
    - Next is to find stationarity of the data of the given stores.
    - Using the best ARIMA order, make predictions for the selected stores.
    - Forecast using SARIMAX
        - Detrend the dataset if necessary,
        - Using SARIMAX estimate 12 weeks forecast

Inferences from the Project

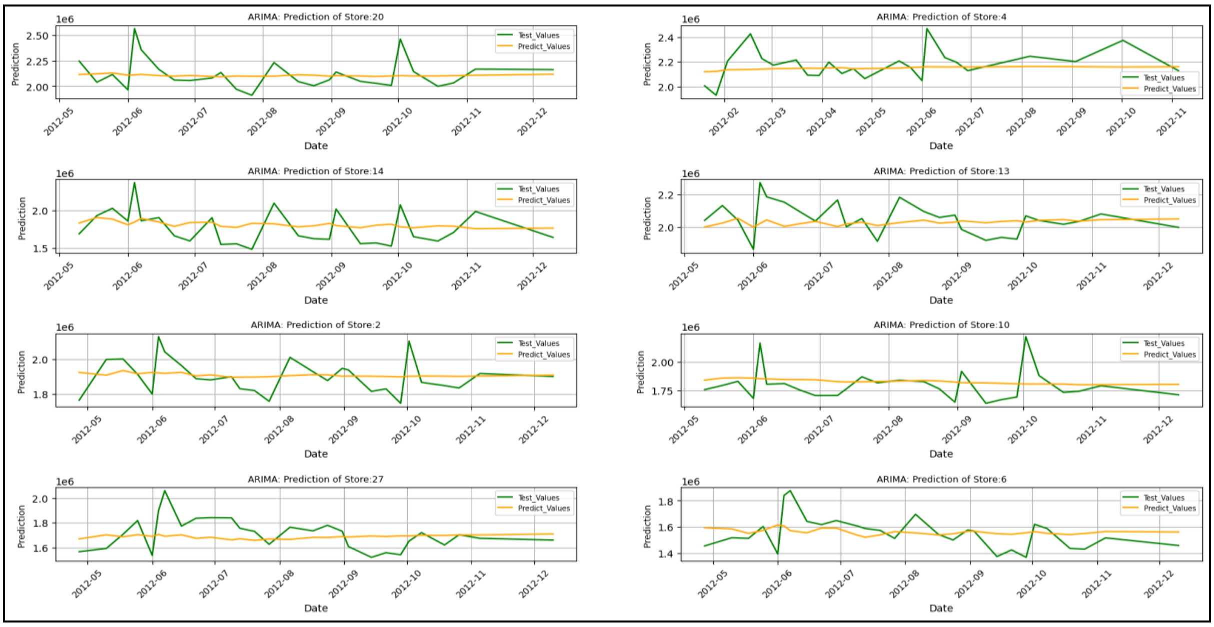
Model Results:

1. ARIMA model:

   a. **Predictions:**

   Predictions were performed for eight stores (stores: 20, 4, 14, 13, 2, 10, 27 and 6) in order of decreasing weekly sales revenue. The predictions results are summarized in the Table and graphs below:

	Store 20	**Store 4**	**Store 14**	**Store 13**	**Store 2**	**Store 10**	**Store 27**	**Store 6**
Median Error %	3.21	13.19	15.75	6.79	0.43	1.65	9.67	0.89
Mean Error %	2.67	12.14	16.75	5.86	0.90	3.26	6.69	0.53

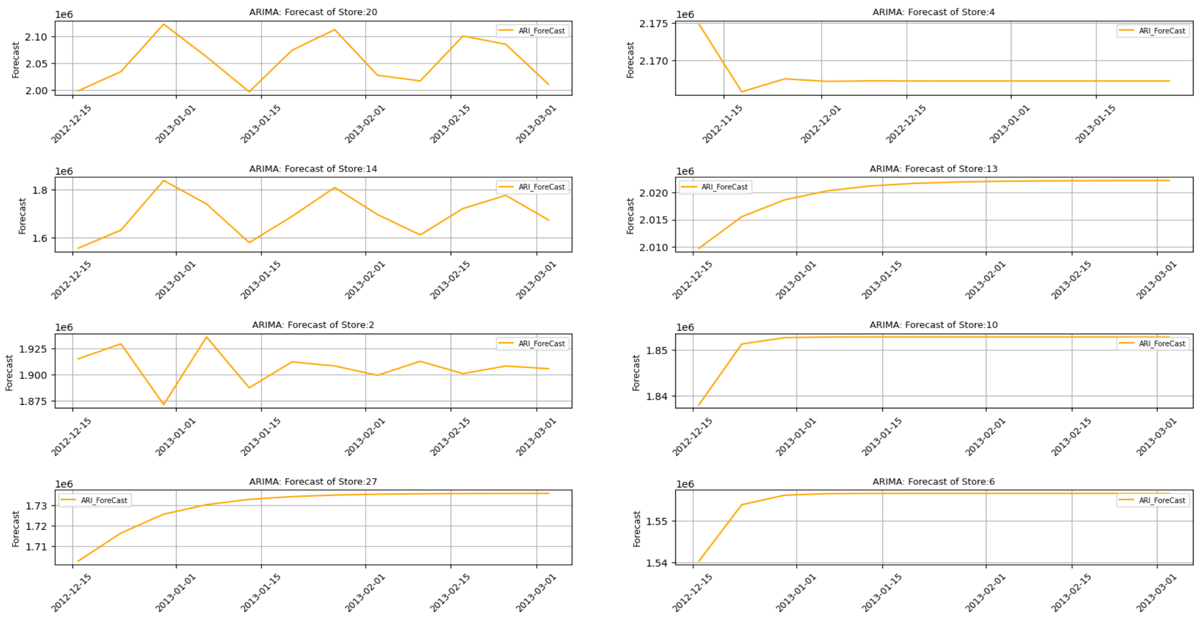


Inferences from the Project

b. **Forecast:**

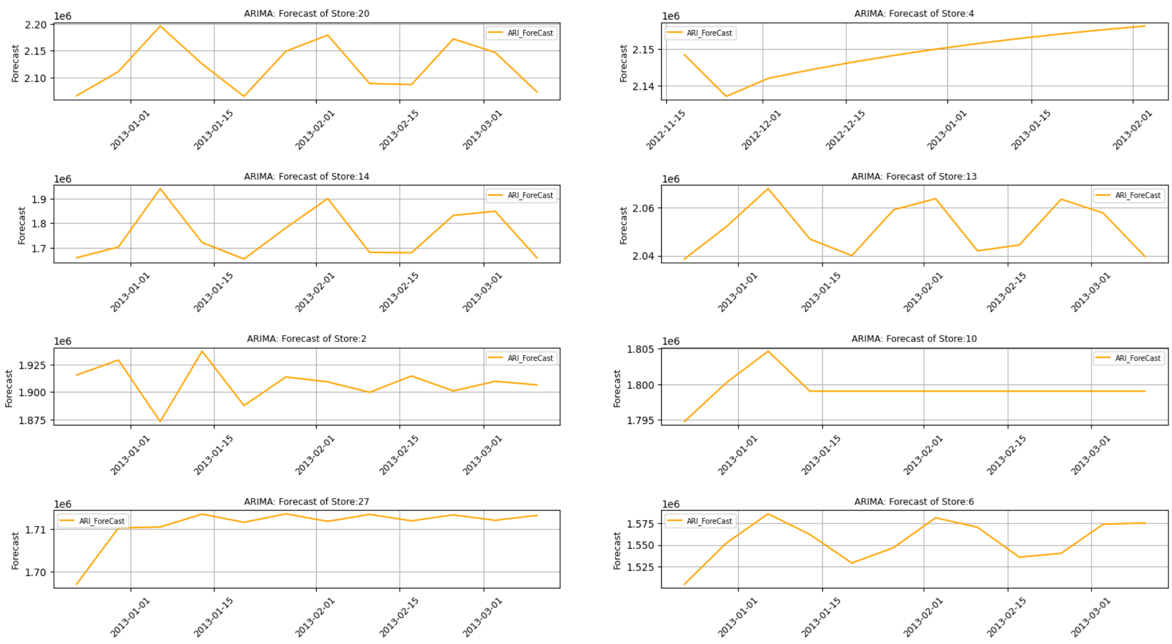
The initial results of the forecast shown above are not very good showing evidence of noise which maybe as a result of trends, and the observed outliers in the dataset which are distorting the forecasts. As a result, the dataset was detrended and the forecast repeated.

Forecast Before Detrending: ARIMA



Inferences from the Project

Forecast After Detrending: ARIMA

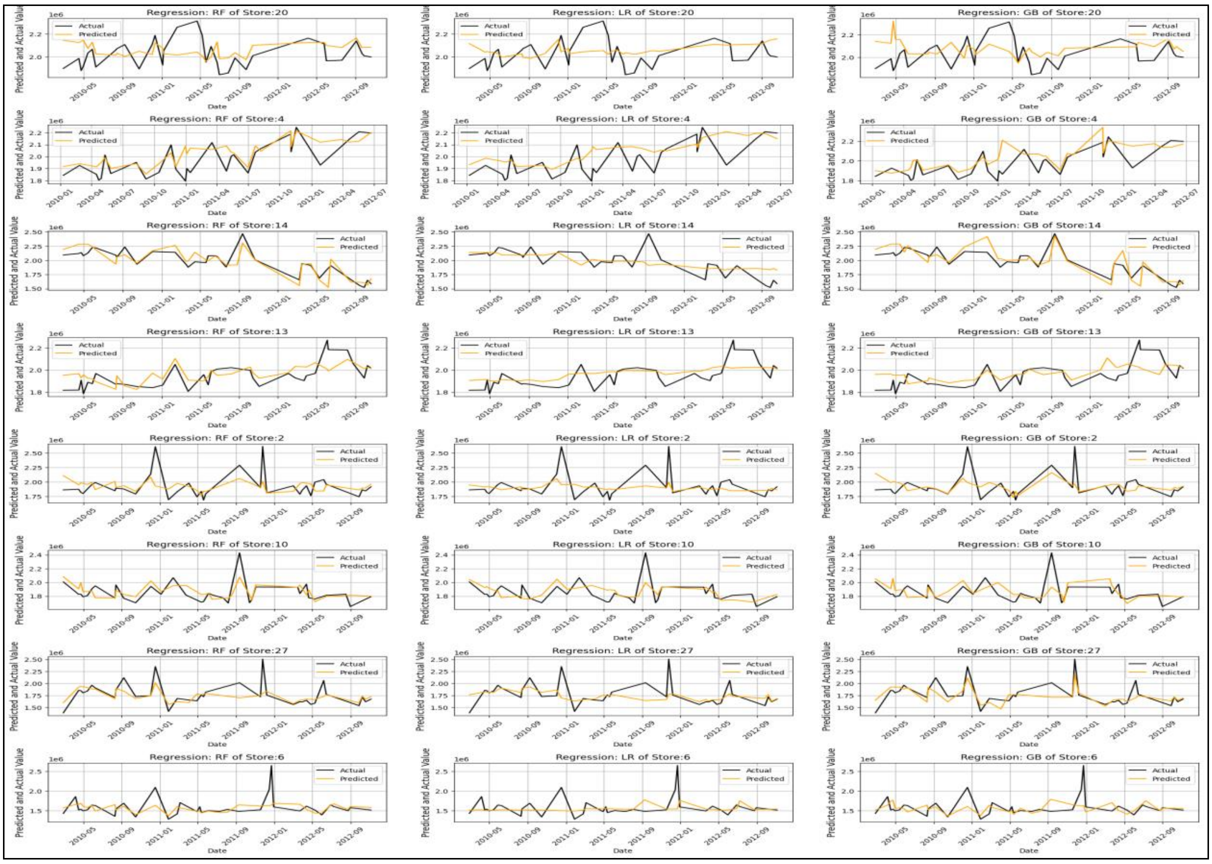


The forecast after detrending sales shows the anticipated variabilities. However the overall projected sales outlook for the next 12 weeks is down for all the stores studied.

2. **Regression Models:**

The predictions of the three models Gradient Boosting, Linear Regression and Random Forest are shown below:

Inferences from the Project



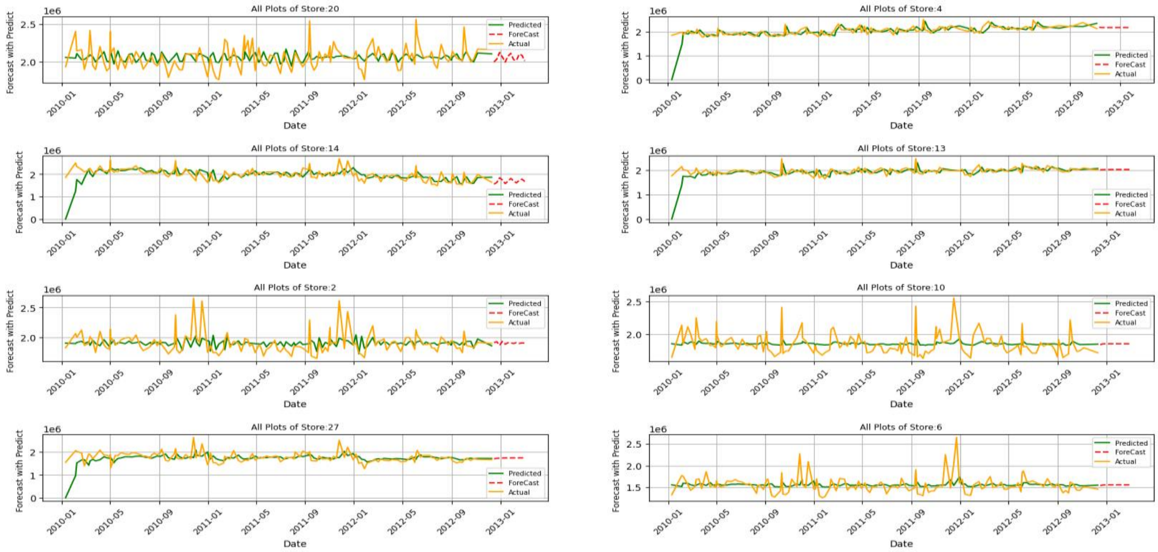
Model Evaluation:

1. **ARIMA/SARIMAX Models:**

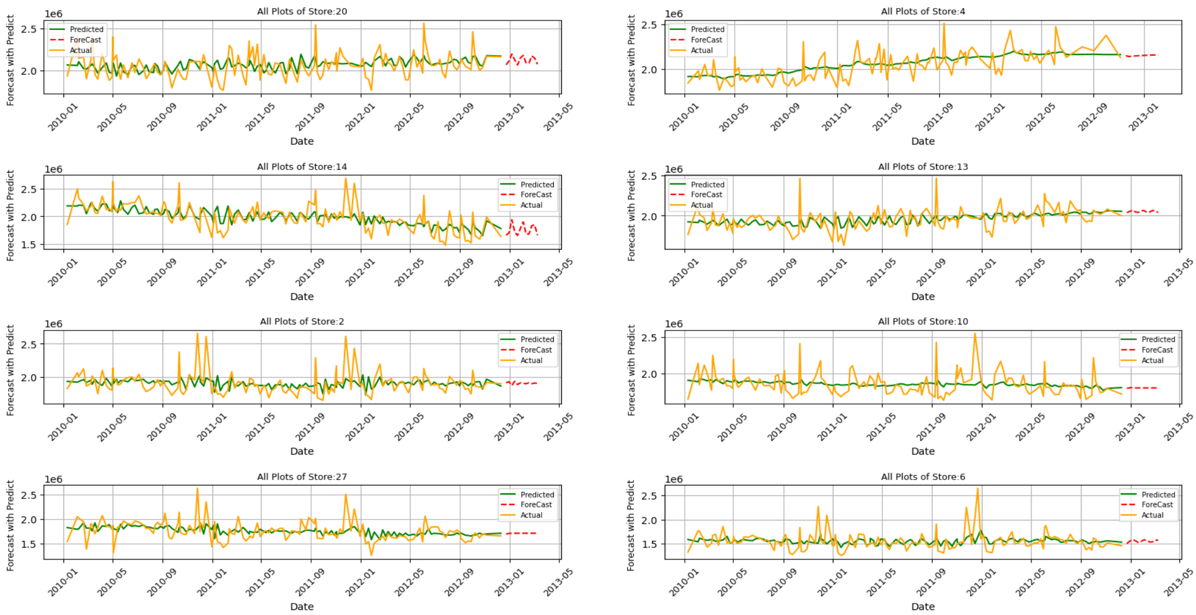
   The model predictions for the selected stores were okay. The forecast after detrending was also okay showing variabilities of the weekly sales in line with sales history as shown below:

Inferences from the Project

Forecast with Prediction before Detrending



Forecast with Prediction after Detrending



Inferences from the Project

ARIMA/SARIMAX Prediction Report for store 20:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	4.56	-20.79	2.53	1.08
After Detrending	11.75	-12.91	2.67	3.21

ARIMA/SARIMAX Prediction Report for store 4:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	1.00	-100.00	16.32	12.63
After Detrending	21.80	0.64	12.14	13.19

ARIMA/SARIMAX Prediction Report for store 14:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	49.34	-100.00	15.29	30.08
After Detrending	-1.25	-30.55	16.75	-15.75

ARIMA/SARIMAX Prediction Report for store 13:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	0.12	-100.00	13.92	8.34
After Detrending	13.36	-5.46	5.86	6.79

ARIMA/SARIMAX Prediction Report for store 2:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	11.46	-9.38	0.60	0.97
After Detrending	8.77	-11.12	0.90	-0.43

ARIMA/SARIMAX Prediction Report for store 10:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	12.93	-16.56	3.78	5.33
After Detrending	11.87	-17.96	3.26	1.65

ARIMA/SARIMAX Prediction Report for store 27:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	17.54	-100.00	5.70	1.07
After Detrending	26.78	-18.81	6.69	9.67

ARIMA/SARIMAX Prediction Report for store 6:

	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Before Detrending	15.16	-17.45	1.01	0.03
After Detrending	20.02	-14.31	0.53	0.89

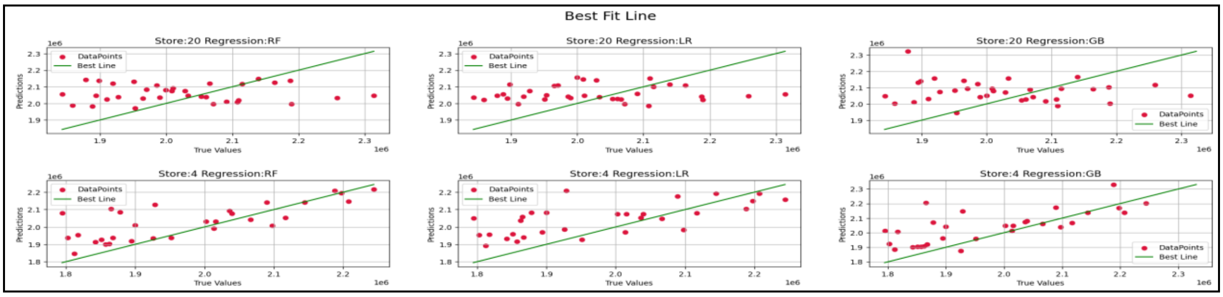
2. **Regression Models:** The regression model results is summarize below:

Inferences from the Project

Regression Prediction Report for Store: 20				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	23.61	-11.39	2.94	2.62
Linear Regression	11.46	-11.19	2.10	2.46
Random Forest	13.19	-11.71	1.96	3.18
Regression Prediction Report for Store: 4				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	19.02	-4.08	3.39	2.05
Linear Regression	14.50	-5.41	3.60	3.35
Random Forest	15.82	-3.81	3.06	1.95
Regression Prediction Report for Store: 14				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	13.59	-17.06	0.98	1.67
Linear Regression	20.63	-22.61	1.24	0.40
Random Forest	9.87	-16.82	0.25	1.50
Regression Prediction Report for Store: 13				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	9.40	-10.81	1.58	2.32
Linear Regression	8.91	-11.25	1.28	1.61
Random Forest	8.95	-10.58	1.33	0.87
Regression Prediction Report for Store: 2				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	15.17	-25.31	1.94	2.62
Linear Regression	15.77	-25.13	-0.36	1.17
Random Forest	11.05	-25.72	1.04	2.15
Regression Prediction Report for Store: 10				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	12.83	-20.53	1.16	0.68
Linear Regression	10.13	-17.84	1.01	1.91
Random Forest	9.44	-13.22	1.28	1.60
Regression Prediction Report for Store: 27				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	15.00	-17.44	-0.37	1.77
Linear Regression	26.91	-29.04	-0.32	0.71
Random Forest	18.90	-28.21	0.26	2.07
Regression Prediction Report for Store: 6				
	Max Err (%)	Min Err (%)	Mean Err (%)	Median Err (%)
Gradient Boosting	21.43	-37.97	0.02	-0.32
Linear Regression	19.63	-41.56	-1.09	0.54
Random Forest	17.28	-38.42	-0.48	1.01

As seen, the mean percentage error is between 3.6% to 9.1% for all the models which is within acceptable range. As seen in the prediction report table, the results from the three models are very comparable.

Inferences from the Project



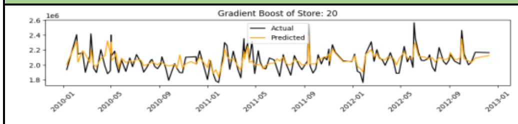
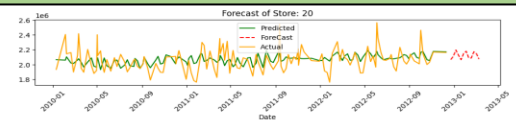
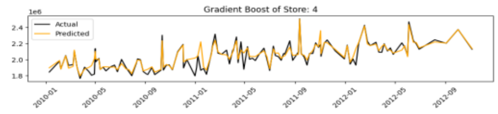
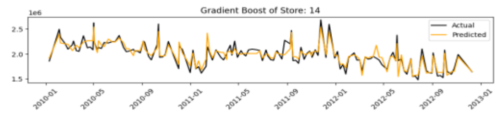
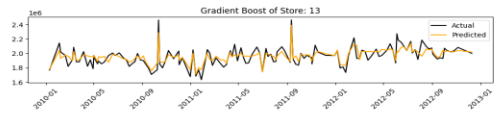
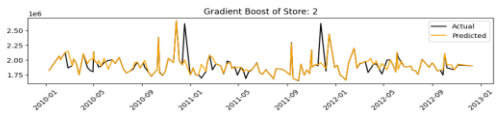
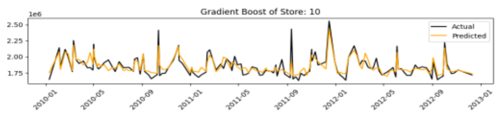
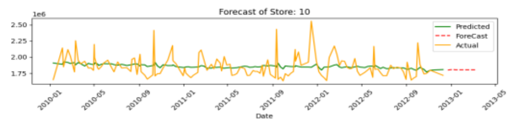
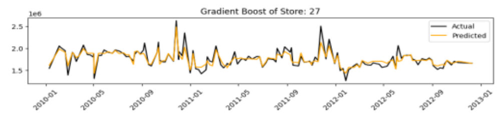
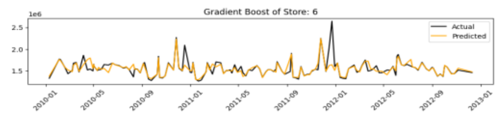
The summary of the evaluation report (r2, ad_r2, root_mean_squared_error and mean_absolute_percentage_error) is presented in the table below. It’s noted below:-

Store No 20: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	-0.68843	-0.75287	148683.49	82693.78	0.0595	0.0318
Linear Regression	-0.20987	-0.25605	125861.15	155365.74	0.0534	0.0577
Random Forest	-0.28476	-0.33380	129697.56	62162.03	0.0525	0.0212
Store No 4: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	0.25816	0.22591	115637.41	29988.15	0.0427	0.0110
Linear Regression	0.21087	0.17656	119265.74	111178.87	0.0506	0.0384
Random Forest	0.33360	0.30463	109599.36	42034.99	0.0408	0.0147
Store No 14: Evaluation Report:-						
	r2	ad_r2	Rmse_test	Rmse_train	Mape_test	mape_train
Gradient Boosting	0.66372	0.65117	127832.59	76155.42	0.0506	0.0308
Linear Regression	0.41740	0.39566	168258.58	216623.88	0.0644	0.0828
Random Forest	0.66831	0.65593	126956.79	69877.84	0.0483	0.0234
Store No 13: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	0.16777	0.13625	99896.27	55843.68	0.0400	0.0215
Linear Regression	0.35388	0.32941	88020.39	126360.40	0.0347	0.0464
Random Forest	0.32806	0.30261	89761.86	45466.72	0.0360	0.0164
Store No 2: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	0.08839	0.05487	194000.84	2750.17	0.0649	0.0011
Linear Regression	0.15042	0.11919	187283.63	146702.82	0.0582	0.0546
Random Forest	0.11914	0.08676	190700.34	59207.11	0.0636	0.0190
Store No 10: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	0.11270	0.07934	131204.63	62064.24	0.0494	0.0265
Linear Regression	0.18779	0.15726	125530.32	148023.28	0.0517	0.0571
Random Forest	0.41620	0.39425	106425.72	54324.94	0.0449	0.0203
Store No 27: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	0.59491	0.58002	142537.08	60824.05	0.0553	0.0274
Linear Regression	0.16705	0.13643	204391.30	164971.22	0.0625	0.0631
Random Forest	0.37631	0.35338	176862.52	65491.62	0.0614	0.0224
Store No 6: Evaluation Report:-						
	r2	ad_r2	rmse_test	rmse_train	mape_test	mape_train
Gradient Boosting	0.01580	-0.02012	238353.50	21623.35	0.0891	0.0105
Linear Regression	-0.10162	-0.14183	252170.83	157085.94	0.0818	0.0773
Random Forest	0.13954	0.10814	222866.69	56010.53	0.0738	0.0226

Inferences from the Project

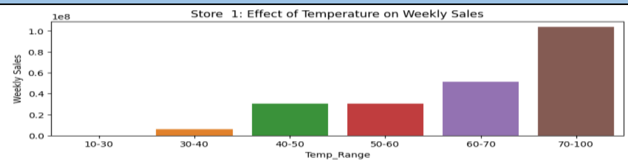
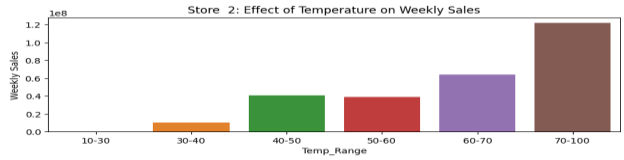
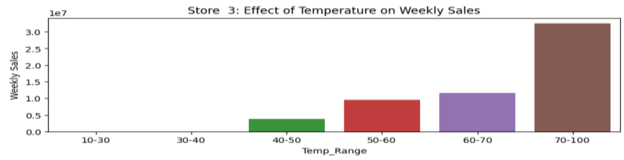
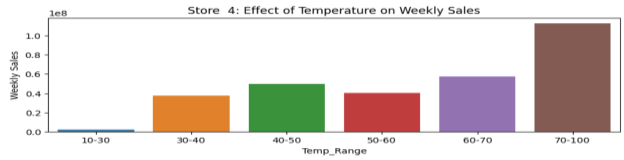
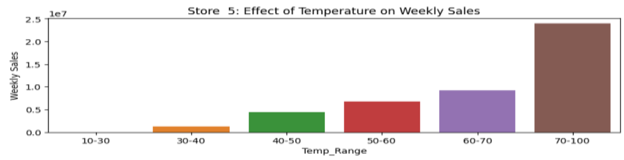
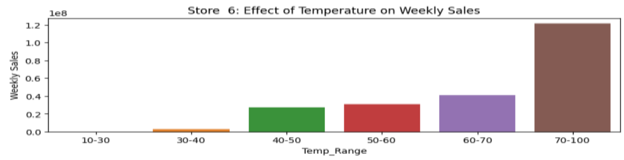
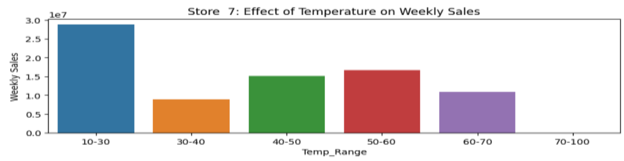
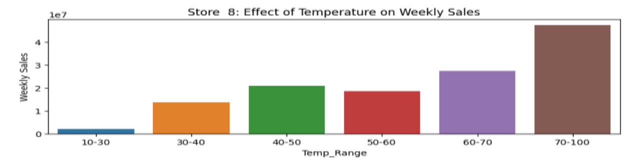
Mean_squared_error, r2, which is a measure of the goodness of fit of the model to the data is negative for all the models. However, the mean absolute percentage error (for both test and train) which is a measure of the accuracy of the model is very good for all the models.

3. **Comparing Models:**

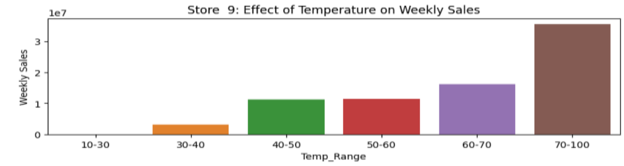
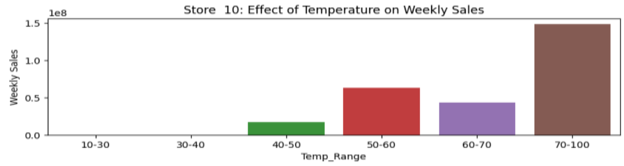
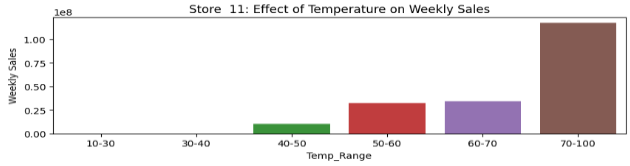
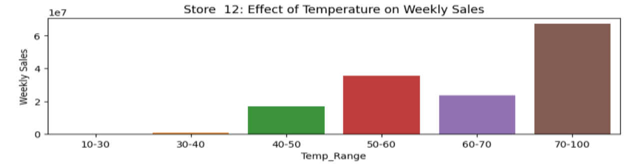
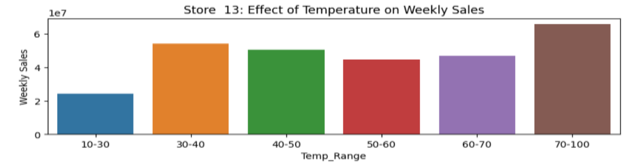
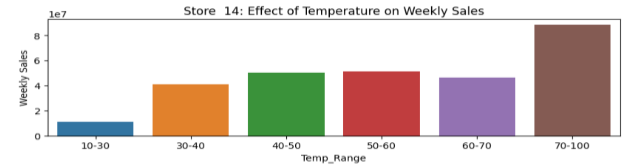
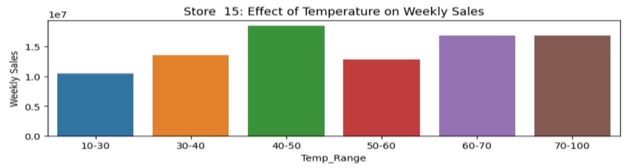
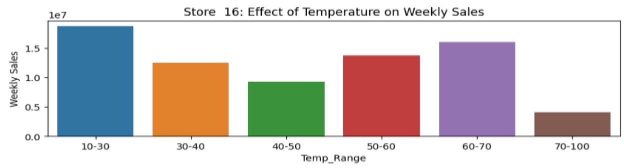
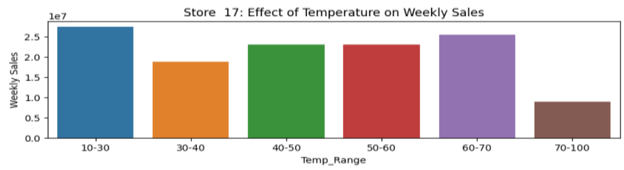
Regression Model Predictions	ARIMA Model Predictions
	
	
	
	
	
	
	
	

Inferences from the Project

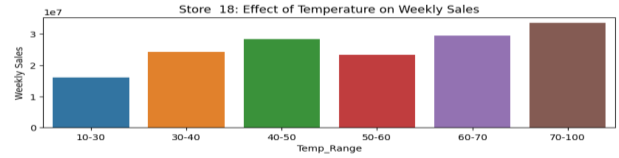
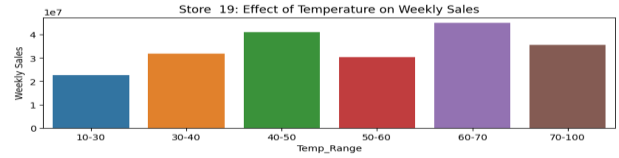
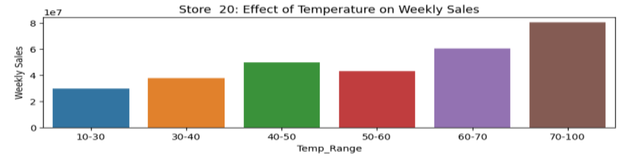
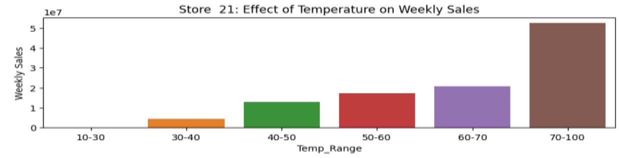
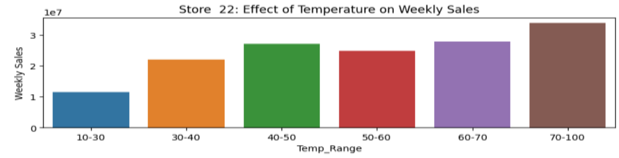
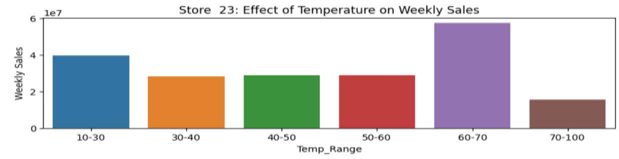
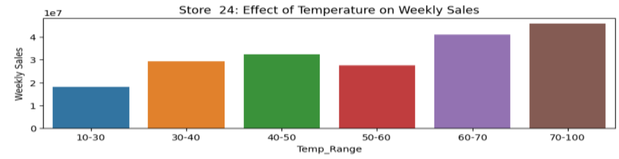
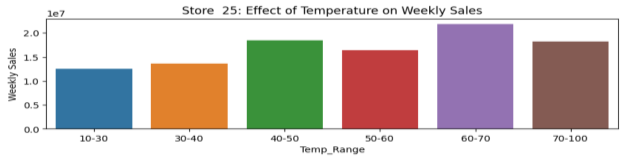
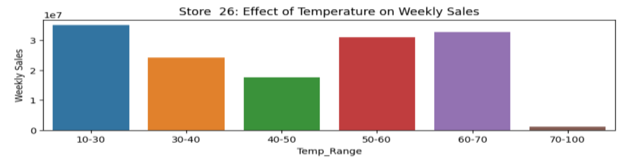
Temperature Effects on weekly sales: evaluation of how changes in temperature effects the weekly sales revenue is presented below:

Store	Outlook – Recommendation (s)	
1		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • ***– must fill up inventory for hot***
2		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • ***– must fill up inventory for hot***
3		• *Significant sales happened in hot weather months* • *Mild sales in summer weather* • ***– must fill up inventory for hot***
4		• *Significant sales happened in hot weather months* • *Mild sales in summer, warm and cold weather* • ***– must fill up inventory for hot***
5		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • ***– must fill up inventory for hot***
6		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • ***– must fill up inventory for hot***
7		• *Significant sales happened in cold weather months* • *Moderate sales in warm weather* • *Mild sales in summer weather* • ***– must fill up inventory for Winter***
8		• *Significant sales happened in hot weather months* • *Mild sales in summer, warm and cold weather* • ***– must fill up inventory for hot***

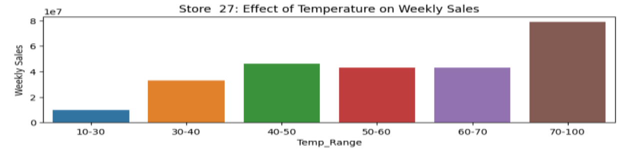
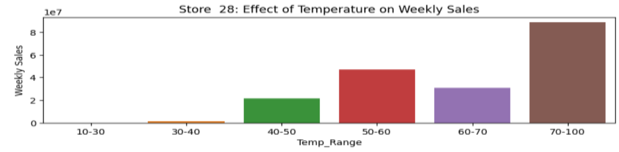
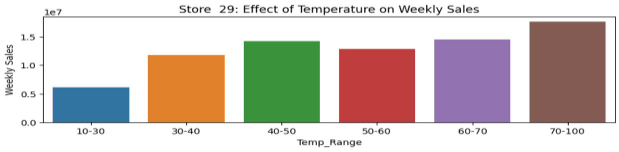
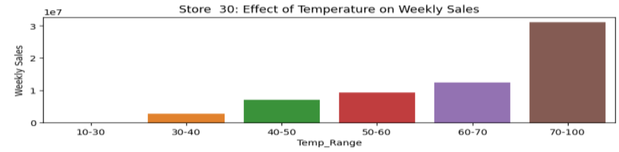
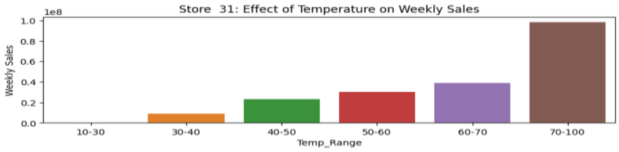
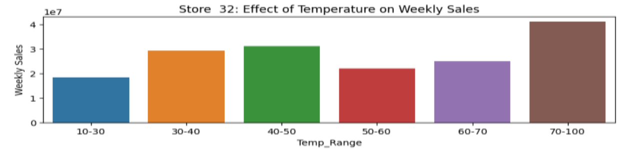
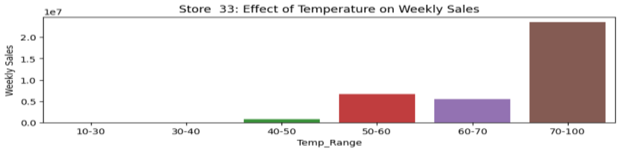
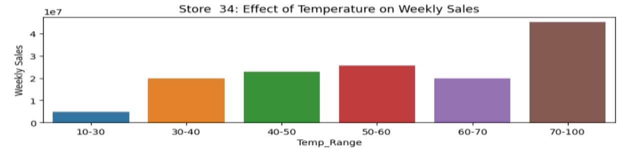
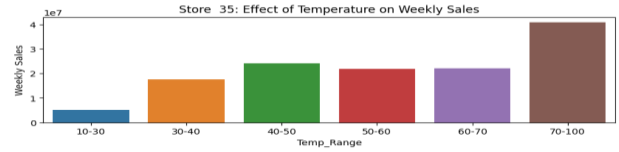
Inferences from the Project

Store	Outlook – Recommendation (s)	
9		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
10		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
11		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
12		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
13		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot and warm weather***
14		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
15		• *Significant sales happened in warm weather months* • *Moderate sales in hot, warm and cold weather* • ***must fill up inventory for warm weather***
16		• *Significant sales happened in cold weather months* • *Moderate sales in warm weather* • ***must fill up inventory for winter***
17		• *Significant sales happened in cold weather months* • *Moderate sales in warm weather* • ***must fill up inventory for winter***

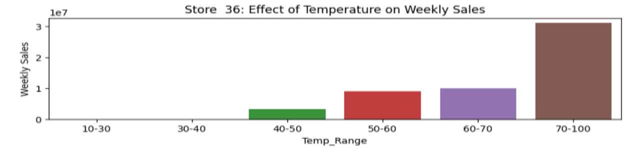
Inferences from the Project

Store	Outlook – Recommendation (s)	
18		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
19		• *Significant sales happened in warm weather months* • *Moderate sales in warm weather* • *Mild sales in cold weather* • ***must fill up inventory for summer***
20		• *Significant sales happened in hot weather months* • *Moderate sales in warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
21		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
22		• *Significant sales happened in hot weather months* • *Moderate sales in warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
23		• *Significant sales happened in summer weather months* • *Moderate sales in cold and warm weather* • *Mild sales in hot weather* • ***– must fill up inventory for summer***
24		• *Significant sales happened in hot weather months* • *Moderate sales in summer and fall* • ***– must fill up inventory for hot***
25		• *Significant sales happened in summer weather months* • *Moderate sales in summer, fall and cold weather* • ***– must fill up inventory for summer***
26		• *Significant sales happened in cold weather months* • *Moderate sales in warm weather* • *Mild sales in hot weather* • ***must fill up inventory for winter***

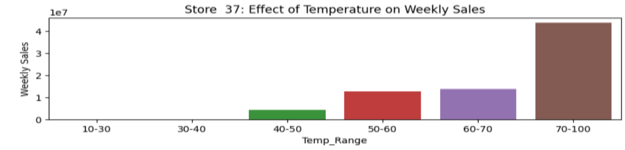
Inferences from the Project

Store	Outlook – Recommendation (s)	
27		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
28		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
29		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
30		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
31		• *Significant sales happened in hot weather months* • *Mild sales in summer and warm weather* • *No sales in cold weather* • ***– must fill up inventory for hot***
32		• *Significant sales happened in hot weather months* • *Moderate sales in warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
33		• *Significant sales happened in hot weather months* • *Mild sales in summer* • *No sales in cold and fall weather* • ***– must fill up inventory for hot***
34		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***
35		• *Significant sales happened in hot weather months* • *Moderate sales in summer and warm weather* • *Mild sales in cold weather* • ***must fill up inventory for hot***

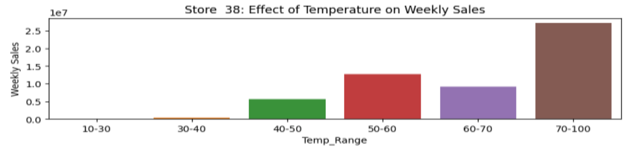
Inferences from the Project

<table>
<tr><th>Store</th><th colspan="2">Outlook – Recommendation (s)</th></tr>
<tr><td>36</td><td>
</td><td>

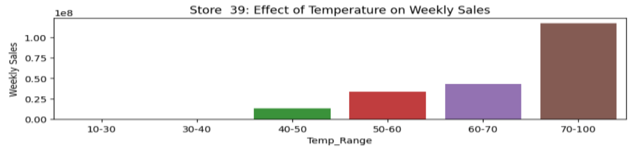
- *Significant sales happened in hot weather months*
- *Mild sales in summer and warm weather*
- *No sales in cold weather*
- ***– must fill up inventory for hot***

</td></tr>
<tr><td>37</td><td>
</td><td>

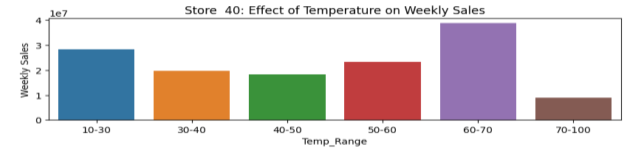
- *Significant sales happened in hot weather months*
- *Mild sales in summer and warm weather*
- *No sales in cold weather*
- ***– must fill up inventory for hot***

</td></tr>
<tr><td>38</td><td>
</td><td>

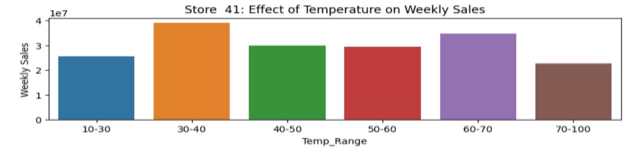
- *Significant sales happened in hot weather months*
- *Mild sales in summer and warm weather*
- *No sales in cold weather*
- ***– must fill up inventory for hot***

</td></tr>
<tr><td>39</td><td>
</td><td>

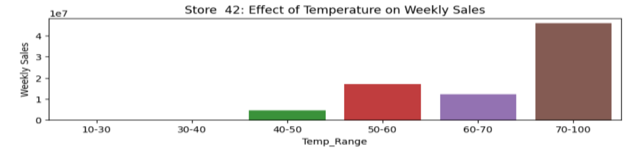
- *Significant sales happened in hot weather months*
- *Mild sales in summer and warm weather*
- *No sales in cold weather*
- ***– must fill up inventory for hot***

</td></tr>
<tr><td>40</td><td>
</td><td>

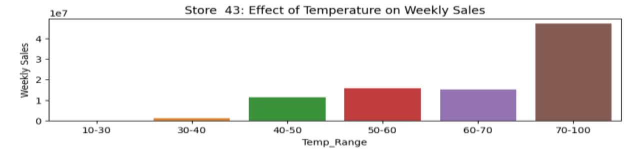
- *Significant sales happened in summer weather months*
- *Moderate sales in summer, fall and cold weather*
- ***– must fill up inventory for summer***

</td></tr>
<tr><td>41</td><td>
</td><td>

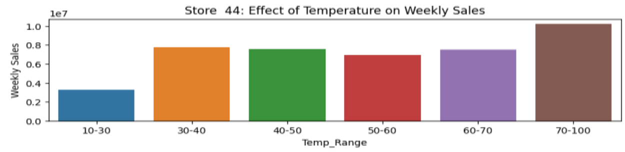
- *Significant sales happened in summer weather months*
- *Moderate sales in summer, fall and cold weather*
- ***– must fill up inventory for summer***

</td></tr>
<tr><td>42</td><td>
</td><td>

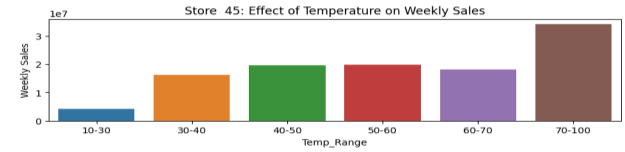
- *Significant sales happened in hot weather months*
- *Mild sales in summer and warm weather*
- *No sales in cold weather*
- ***– must fill up inventory for hot***

</td></tr>
<tr><td>43</td><td>
</td><td>

- *Significant sales happened in hot weather months*
- *Mild sales in summer and warm weather*
- *No sales in cold weather*
- ***– must fill up inventory for hot***

</td></tr>
<tr><td>44</td><td>
</td><td>

- *Significant sales happened in hot weather months*
- *Moderate sales in summer and warm weather*
- *Mild sales in cold weather*
- ***must fill up inventory for hot***

</td></tr>
<tr><td>45</td><td>
</td><td>

- *Significant sales happened in hot weather months*
- *Moderate sales in summer and warm weather*
- *Mild sales in cold weather*
- ***must fill up inventory for hot***

</td></tr>
</table>

Future Possibilities

Future of Machine learning is revolutionary and exciting. At present it is used in healthcare with personalized treatments, empower autonomous systems for safer transportation and industry, and drive ethical AI innovations ensuring fairness and transparency in decision-making across sectors.

By helping enterprises to better understand both customer's and business functioning behaviour, Machine learning enables companies to offer better/targeted customer service leading to more loyal customers and ultimately improved sales revenue.

All over the world, almost all big corporate organizations (Facebook, Apple, Amazon and Google) and many more companies employ machine learning in their daily operations.

Conclusion

The project undertook a study of a retail company with 45 outlets stores. Some of the key findings from the survey include the following:

1. Sales projection of next 12 weeks for the most of the stores is down.
2. Mild no of stores have no sales in some period of year mostly in cold season.
3. To improve sales revenue, the following steps are recommended:
    a. Concerted efforts by the company to find out though local market surveys and past sales records what products are in high demand by the local population at any given period of the year and make effort to replenish those stocks.
    b. Create increased local awareness of the products on offer at each store through commercial outreach: social media, television commercials and trade shows to name a few, could help improve sales.
    c. Have detailed records of inventory of the items on offer at each store indicating amount and dates if sold as it is needed for effective inventory tracking which could help in maintaining stores offer.
    d. Explore other service options that have worked well for similar companies, such as same day or next day delivery or services provided for the product provided.

It may result to wounding some stores as sales revenue does not improve.

References

Authors	Books
Robert H. Shumway and David S. Stoffer	Time Series Analysis and Its Applications
Rob J Hyndman and George Athanasopoulos	Forecasting: Principles and Practice
Wes McKinney	Python for Data Analysis

1. **Time Series Analysis and Its Applications** by Robert H. Shumway and David S. Stoffer: This textbook covers various aspects of time series analysis, including stationarity, decomposition, and forecasting techniques.
2. **Forecasting: Principles and Practice** by Rob J Hyndman and George Athanasopoulos: This online textbook provides a comprehensive overview of forecasting methods, including ARIMA, SARIMA, and SARIMAX models, along with practical examples.
3. **Python for Data Analysis** by Wes McKinney: This book covers data analysis techniques using Python, including time series analysis with pandas, statsmodels, and other libraries.
4. **Statsmodels Documentation**: The official documentation for statsmodels library offers detailed information on time series analysis and forecasting methods available in Python
5. **Blogs and Articles**:

   Data science blogs like Towards Data Science, Analytics Vidhya, and Medium often publish articles on time series analysis and forecasting.