ASSIGNMENT SUBMISSION FORM

Course Name : Forecasting Analytics

Section : A

Project Report : Inventory Management through Sales Forecasting

	PGID	Name of the Member
1.	61710012	Anand Abhishek
2.	61710773	Bharath Sankaran
3.	61710053	Mayank Thapliyal
4.	61710682	Rohan Chakraborty
5.	61710703	Urvashi Surana Sunil
6.	61710741	Varun Madnani

Forecasting Analytics- Group A2

PROJECT TITLE: ***Inventory Management at Nestle through Sales Forecasting***

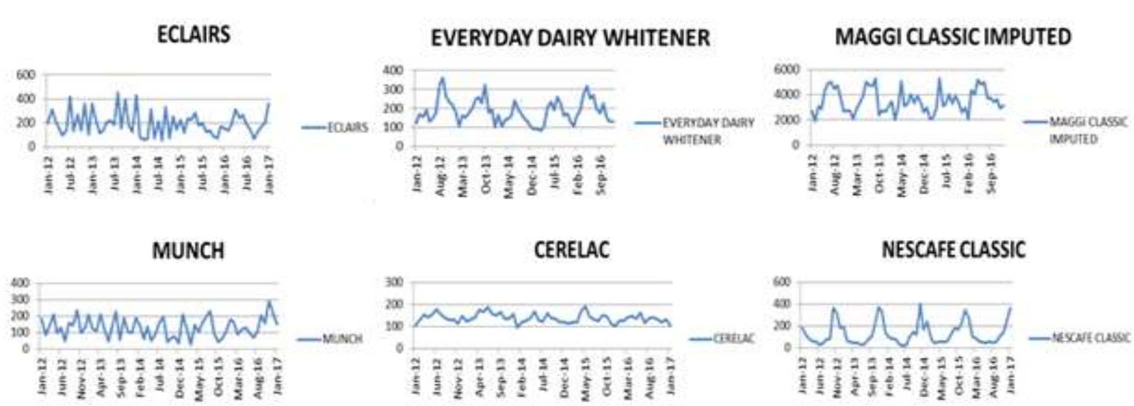
EXECUTIVE SUMMARY:

Problem Description: FMCG companies like Nestle face trouble in forecasting demand for smaller regions which comprises nearly 50% of their business and is highly critical. This is due to high volatility in demand. Due to this problem more often than not the sales force in these regions face a situation wherein they are either short of inventory and unable to meet demand or have piled up inventory at warehouses. A model that effectively forecasts sales can be tested on a small region (in this case Ferozpur sales zone) which if successful can later be deployed to other smaller regions which can be highly beneficial in management of inventory and thus production.

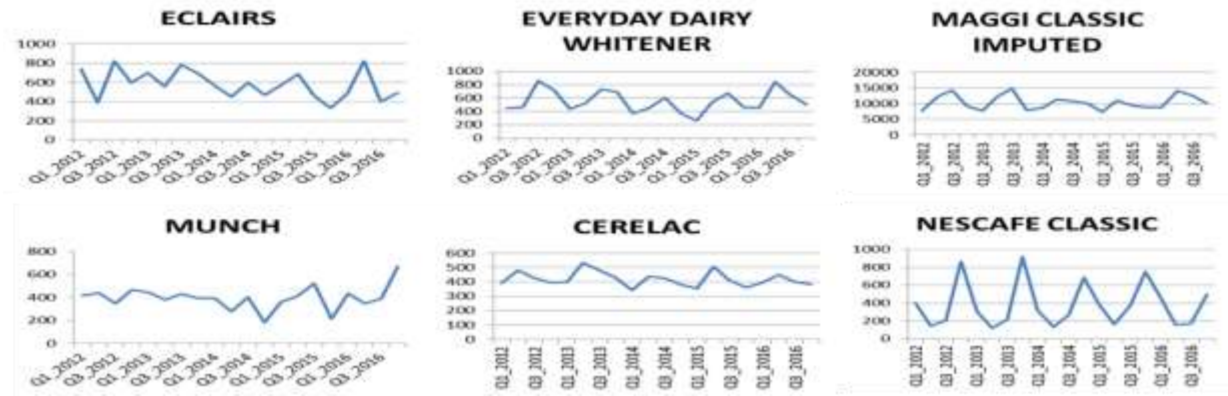
Brief Description of Data: The data has been sourced from Nestle SDS software (ERP system).The dataset obtained contains daily bill-wise product-wise sales and retailer-wise product wise sales at a distributor level for the time period Jan 2012 to Jan 2017. We have considered the top 6 products in different product categories of Nestle based on Sales volumes, i.e., Maggi, Eclairs, Nescafe Coffee Classic, Everyday daily whitener, Munch & Cerelac. Since the data is actual company sales data and Nestle has a very robust reporting mechanism, therefore the data was accurate & complete (except for 2 values which possibly got lost while extracting the data from the system. We manually confirmed those two values through the system and updated them). Data is consistent, unique and timely, i.e., it is in ascending order, there are no duplicate time stamps and they match the precision of a calendar.

Charts (Various products plotted against time):

Monthly data



Quarterly data



Final Forecasting Model:

The final model is a tool that is able to forecast the sales of top 6 products by choosing the method most suited for forecasting sales for that product. The manager will need to input the product name in the model and the model will use historical data to come up with the forecast. The most suited model is determined based on the predictive accuracy (the one with least expected error in forecasting) of the method on a product time series. The overall performance of the model as compared to the benchmark is 12-15% better in terms of % error.

Conclusion:

Based on our analysis, we found that:

1. Quarterly forecasting performs much better on predictive accuracy as compared to monthly forecasting (**Refer Exhibit 1**).
2. Use of different forecasting models to estimate sales for different products gives maximum predictive accuracy.
3. Predictive power of some models might be impacted by lack of sales promotion and marketing campaign information.
4. The forecasts for the month of Feb ’17, Mar ’17 and Apr ’17 have been provided in **Exhibit 2**

Recommendations:

1. We should go ahead with quarterly forecasting as it provides a better estimate.
2. Forecasting horizon should not be more than one quarter as it is a learning based model.

3. Predictive power of the model should be improved by taking into account the effect of sales and marketing initiatives along with other external factors (macroeconomic, competitor strategy).

TECHNICAL SUMMARY:

Data Preparation: Initially we received the bill-wise data at daily level. We then aggregated the data as per product category at monthly and quarterly level using the tool “Spotfire”. We took the top products across six different product categories as per sales volume in Firozpur sales zone in Punjab. These top six products across categories are Maggi (snacks), Nescafe classic (beverages), Everyday Daily Whitener (dairy), Cerelac (infant nutrition), Munch (chocolate), and Eclairs (confectionaries).

However, one of these products (Maggi) was hit by a lawsuit controversy in May 2015 and hence was withdrawn by Nestle in the June, 2015. After thorough quality and process check the product was re-launched in November, 2015. However, as Firozpur sales region is small so the product was relaunched in this area by the end of March. In order to compensate for this unforeseen event, we imputed the affected data of this duration in our dataset by comparing with the sales of previous year in the same period and thus finally multiplying with a correction factor (depending on the growth/de-growth rate observed in sales data till April, 2015 over the previous year).

There were also four outliers in the entire dataset. After investigation it was determined that these values were either due to unavailability (shortage) of the product or due to flooding of the product in the market in the subsequent period. These values were corrected for by taking average of the shortage + flooding sales values.

Preliminary Analysis of the Dataset:

The plots were visualized using “Tableau” and the analysis are summarized as below:

COMPONENTS	TREND	SEASONALITY	LEVEL
MAGGI	NO	NO	3349.214
CERELAC	Slightly downward linear trend	NO	139.27
EVERYDAY WHITENER	Downward Linear trend	NO	183.60
ECLAIRS	Linear	NO	195.98
NESCAFE COFFEE	NO	YES(Quarterly & Monthly)	129.98
MUNCH	Polynomial of degree two.	NO	132.47

<u>Methods Employed:</u> Other than naïve prediction, which we used as a benchmark to compare other method's performance, we deployed different methods based on individual time-series components observed. For all methods the forecasting horizon used was three months for monthly data and 1 quarter for quarterly data. The methods used have been summarized below:

1). <u>Moving Averages:</u> Wherever we did not observe any seasonality or trend in the dataset we used this method. The method uses trailing moving averages method to forecast. Values of intervals tried were 2, 4, 8 for quarterly data and 2, 4, 8, 12 for monthly data.

2). <u>Exponential Smoothing:</u> Again whenever no trend or seasonality was observed we applied this method. The default Alpha (0.2) and optimized Alpha values were tried. The optimized model performed worse for all different products.

3). <u>Double Exponential Smoothing/Holt's Method:</u> If a time series had trend but no seasonality, this method was applied. The default Alpha (0.2) & Beta (0.15) and optimized Alpha & Beta values were tried. In one case the optimized model performed better but on checking the RMSE values there was a hint of overfitting, so it was ignored.

4). <u>Holt-Winters Method:</u> If a time series had both seasonality and trend we employed this method. The default Alpha (0.2), Beta (0.15), Gamma (0.05) and optimized values of these were also tried. The optimized model performed worse.

5). <u>Multi Linear Regression (MLR):</u> This method was applied to all product datasets. If the trend observed was linear the "t" variable was used to capture trend. If the trend was polynomial another variable "t^2" was used for capturing trend. To capture seasonality, if present, we used a dummy variable based on seasonality. For example, if there was a monthly seasonality, 12 dummy variables were used.

6). <u>MLR + AR:</u> This method was used whenever, the residuals for training set in MLR output had autocorrelation. Depending on the dataset and autocorrelation observed AR (1) and AR (2) was used in pairing with MLR.

7). <u>Neural Networks:</u> This method involves using a input layer, a few hidden layers and an output layer connected to each other through different functions and weights. For our estimations we used 1 hidden layer and tried different node values like 4, 6, 12, 25. The best output amongst these were obtained when node values in hidden layer were 25.

Performance Measure: We have used MAPE (Mean Absolute Percentage Error) as our performance measure as it can be compared over different models even if the two models have different actual values. RMSE for the same MAPE is higher for data with higher actual values. MAPE is free of scale effects of actual value. Thus, it provides us the benefit of being able to compare quarterly data forecasts against monthly data forecasts.

Benchmark: For all the six products we used naïve prediction as our benchmark to compare different models and hence the best fit. The predictive performance on training, validation set and forecast for Naïve as compared to a few of the models is shown in the appendix in **Exhibit 2 charts.**

APPENDIX:

Exhibit 1: PREDICTIVE ACCURACY - QUARTERLY V/S MONTHLY:

	Monthly		Quarterly	
Product	Method	MAPE	Method	MAPE
Munch	MLR	29.86%	Moving Average	16.96%
Cerelac	MLR	8.87%	Holts Winter	4.78%
Nescafe Classic	MLR	47.17%	MLR	38.28%
Maggi Classic	Moving Average	23.12%	Double Exponential	19.85%
Eclairs	MLR	35.59%	Double Exponential	17.52%
Everyday Dairy	Holts	24.22%	Exponential	18.29%

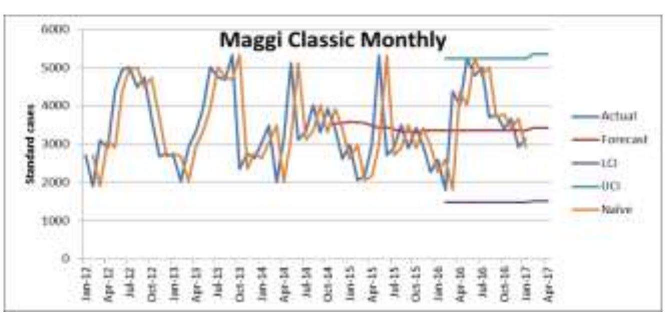
Exhibit 2: FORECASTED DATA ALONG WITH CHARTS:

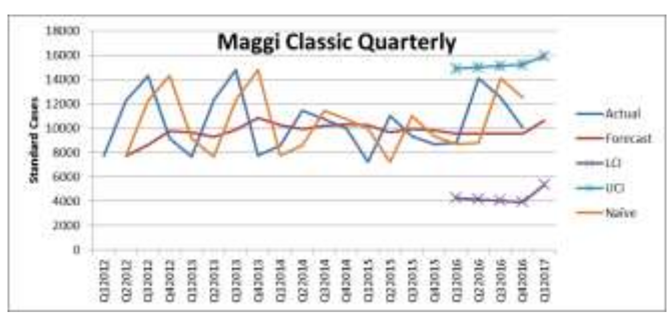
Month on month Forecasts

Product	Time	Forecast	LCI	UCI
Munch	Jan-17	132.6446	98.99364	264.568
	Feb-17	133.8324	100.1814	265.7558
	Mar-17	135.0692	101.4182	266.9926
Cerelac	Jan-17	130.5595	110.8912	153.2976
	Feb-17	130.29	110.6217	153.0287
	Mar-17	130.0205	110.3522	152.7587
Nescafe Classic	Jan-17	94.18824	N/A	N/A
	Feb-17	67.78824	N/A	N/A
	Mar-17	55.18824	289.576	399.9528
Everyday Dairy Whitener	Jan-17	189.0603	141.1371	313.9849
	Feb-17	185.8776	137.9544	310.8022
	Mar-17	182.6499	134.7717	307.6196
Eclairs	Jan-17	172.1737	-35.189	379.5368
	Feb-17	171.405	-36.2807	379.0921
	Mar-17	170.637	-37.3819	378.232
Maggi Classic Imputed	Jan-17	3433.676	1509.329	5358.024
	Feb-17	3433.676	1509.329	5358.024
	Mar-17	3433.676	1509.329	5358.024

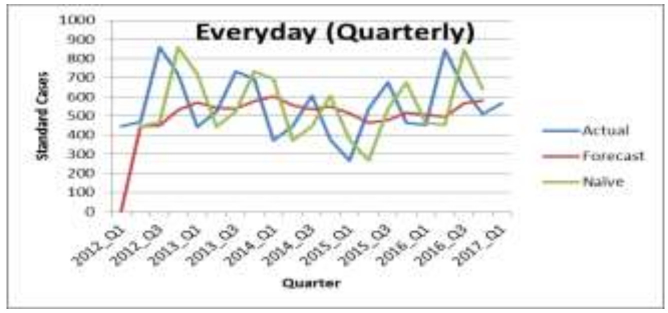
Quarterly Forecast

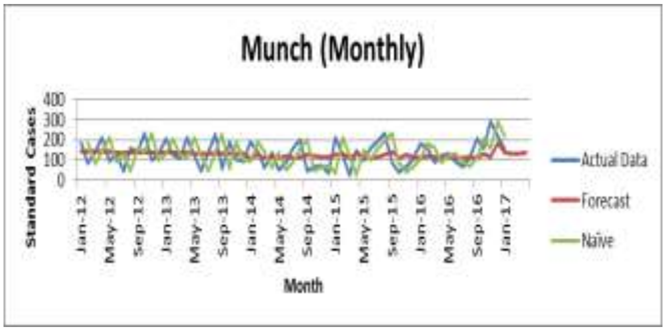
Product	Time	Forecast	LCI	UCI
Munch	Q1_2017	460	244.4423	675.5577
Cerelac	Q1_2017	344.6342	286.4324	402.836
Nescafe Classic	Q1_2017	354.5	75.8875	449.1625
Maggi Classic Imputed	Q1_2017	10614.83	5336.514	15893.14
Eclairs	Q1_2017	482.6824	188.073	777.2918
Everyday Dairy Whitener	Q1_2017	566.7037	223.2655	910.142

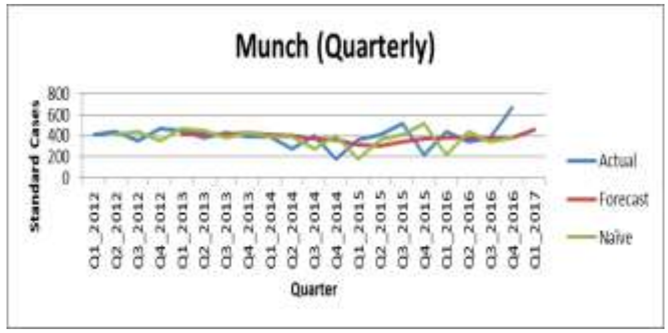


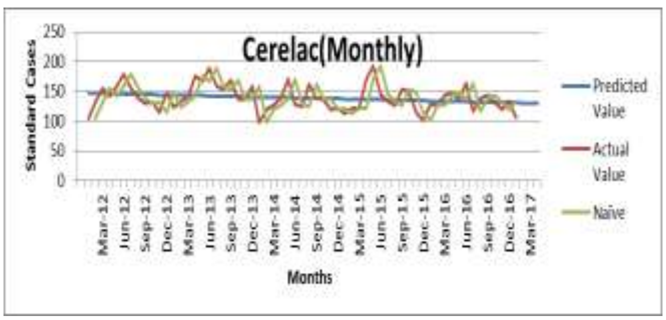


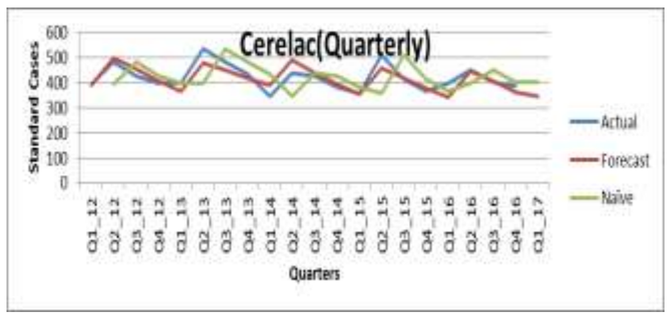

Everyday Sales (Monthly)
Standard Cases
Months
Actual
Predicted
Naive Prediction

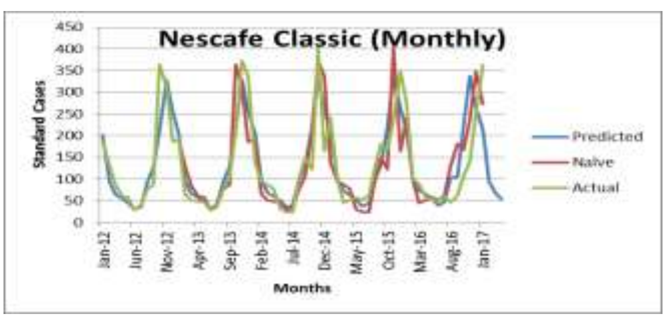

Everyday (Quarterly)
Standard Cases
Quarter
Actual
Forecast
Naive

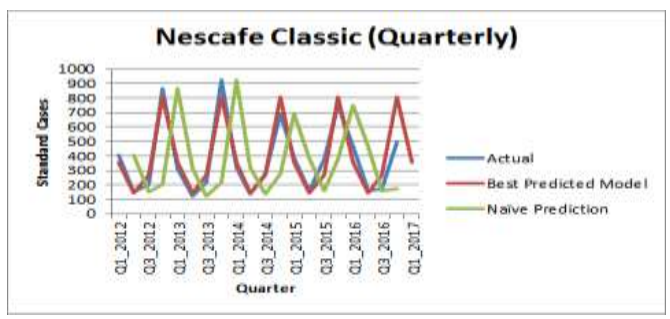

Munch (Monthly)
Standard Cases
Month
Actual Data
Forecast
Naive

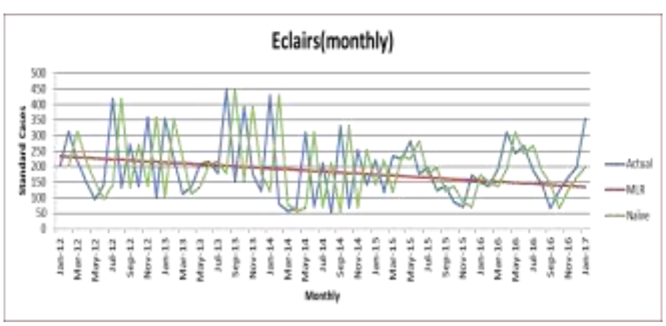

Munch (Quarterly)
Standard Cases
Quarter
Actual
Forecast
Naive


Cerelac(Monthly)
Standard Cases
Months
Predicted Value
Actual Value
Naive


Cerelac(Quarterly)
Standard Cases
Quarters
Actual
Forecast
Naive


Nescafe Classic (Monthly)
Standard Cases
Months
Predicted
Naive
Actual


Nescafe Classic (Quarterly)
Standard Cases
Quarter
Actual
Best Predicted Model
Naive Prediction


Eclairs(monthly)
Standard Cases
Monthly
Actual
MLR
Naive

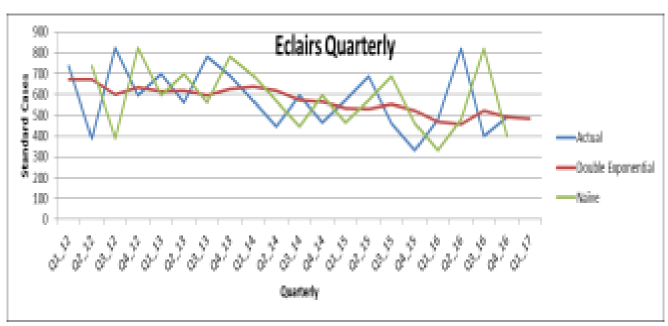
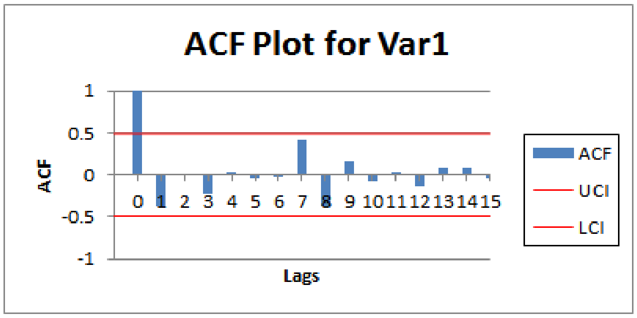
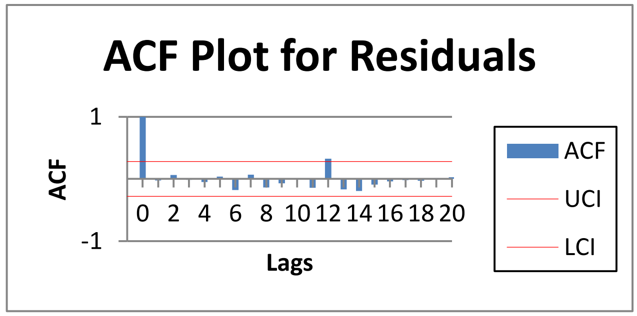

Eclairs Quarterly
Standard Cases
Quarterly
Actual
Double Exponential
Naive

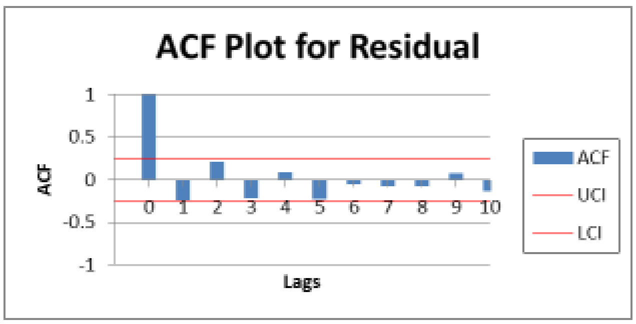
Exhibit 3: ACF plots (Wherever applicable)

NESCAFE ACF

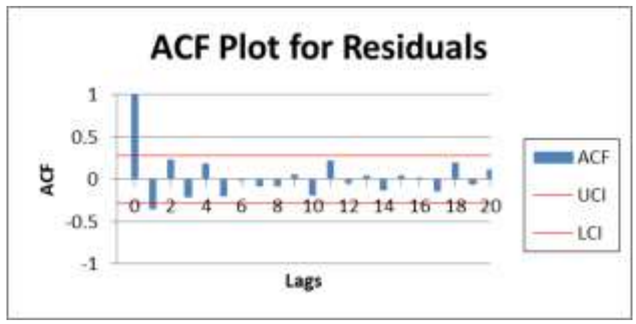




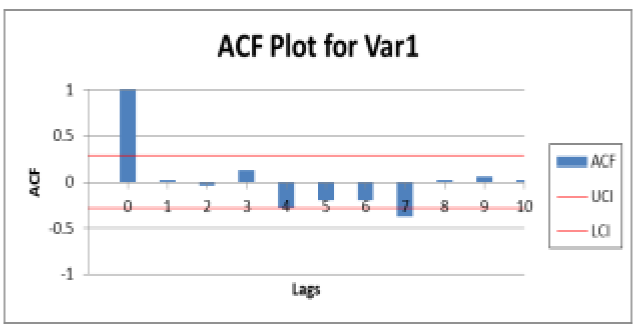
ECLAIRS ACF

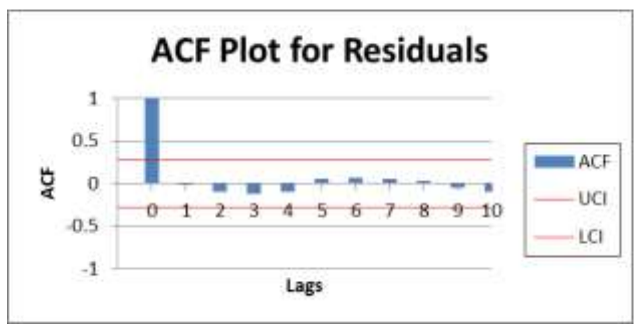


CERELAC ACF

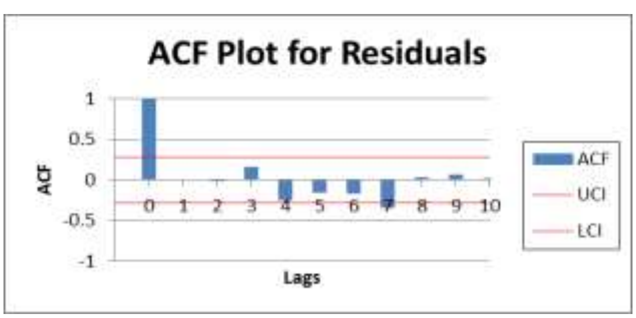


MAGGI ACF





EVERYDAY ACF



MUNCH ACF