LETTERKENNY INSTITUTE OF TECHNOLOGY

ASSIGNMENT COVER SHEET

Lecturer's Name: **Ruth Lennon**

Assessment Title: **Time Series SAS LAB**

Work to be submitted to: **Ruth Lennon**

Date for submission of work: **3 December 2017**

Place and time for submitting work: **RM**

To be completed by the Student

Student's Name: **Prateek Parasher (L00143921)**

Class: **MSc Big Data Analytics**

Subject/Module: **Big Data Analytics**

Word Count (where applicable): **N/A**

I confirm that the work submitted has been produced solely through my own efforts.

Student's signature: **Prateek Parasher** Date: **3 / 12 /2017**

Notes

Penalties: The total marks available for an assessment is reduced by 15% for work submitted up to one week late. The total marks available are reduced by 30% for work up to two weeks late. Assessment work received more than two weeks late will receive a mark of zero. [Incidents of alleged plagiarism and cheating are dealt with in accordance with the Institute's Assessment Regulations.]

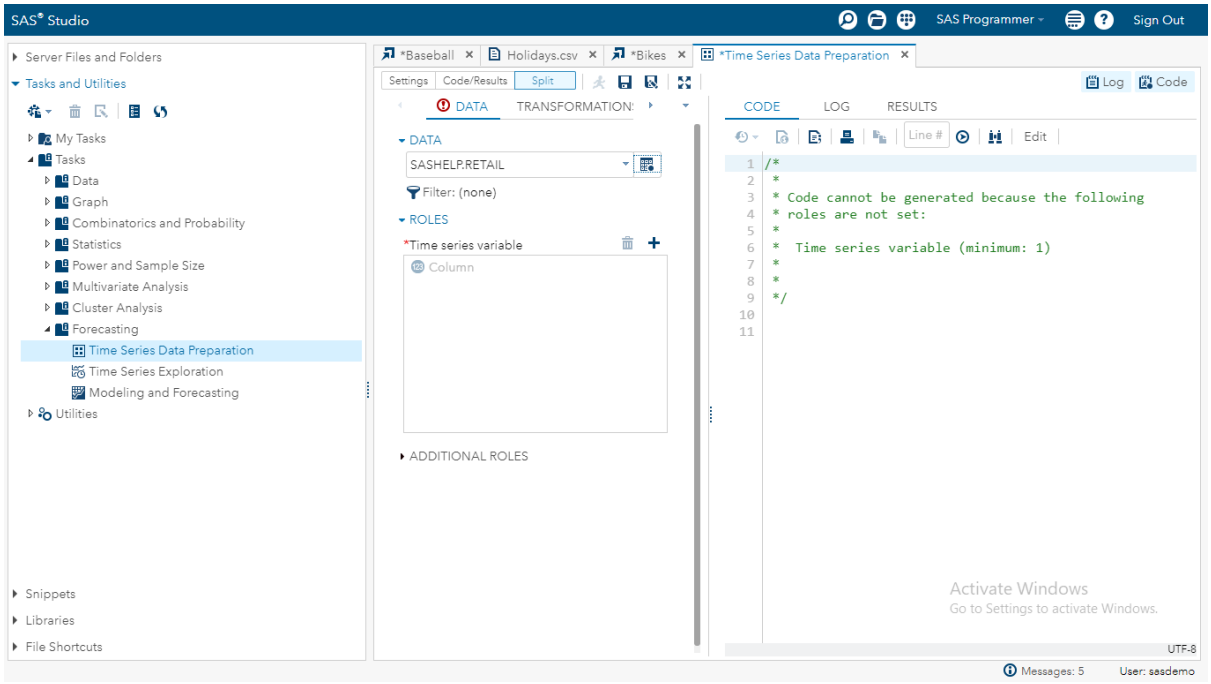
Plagiarism: Presenting the ideas etc. of someone else without proper acknowledgement (see section L1 paragraph 8).

Cheating: The use of unauthorised material in a test, exam etc., unauthorised access to test matter, unauthorised collusion, dishonest behaviour in respect of assessments, and deliberate plagiarism (see section L1 paragraph 8).

Continuous Assessment: For students repeating an examination, marks awarded for continuous assessment, shall normally be carried forward from the original examination to the repeat examination.

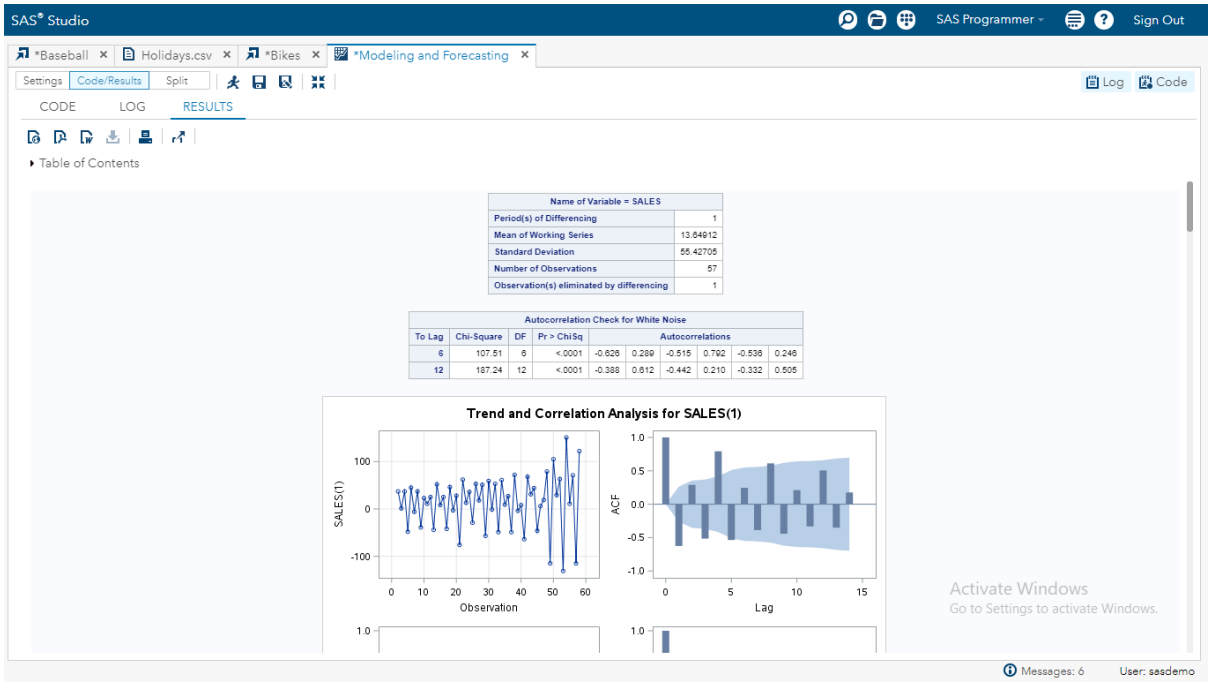
Time Series

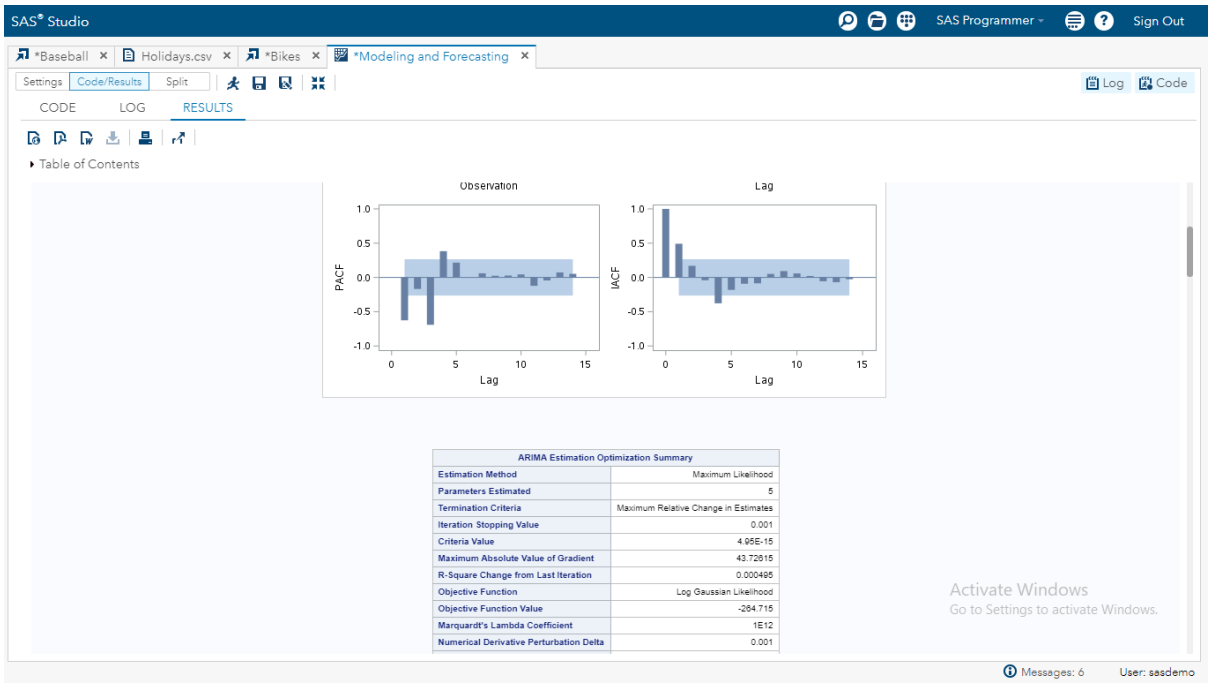
1. Time Series Analysis



Conclusion: - In this lab there are three forecasting tasks is better to go step by step although its not required but in this lab I was doing step by step with time series i can perform many tasks like prediction and analysis , initially step was prepare your data for time series and with time series exploration I did transformation data and exploration last step was modelling and forecasting using time series data there are other ways also for doing time series forecasting for those who like programming language they can do using python , R and matlab python and R both languages have many libraries which giving more power to the time series but those who don't know the programming for them sas GUI is amazing with help of SAS we can actually perform all the time series task.

2. SAS Forecasting





SAS® Studio — SAS Programmer — Sign Out

*Baseball | Holidays.csv | *Bikes | *Modeling and Forecasting

Settings | Code/Results | Split — Log | Code

CODE | LOG | RESULTS

Table of Contents

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	12.72244	1.34324	9.47	<.0001	0
MA1,1	-0.55335	0.14519	-3.81	0.0001	1
AR1,1	-1.06309	0.09419	-11.29	<.0001	1
AR1,2	-0.89391	0.11592	-7.71	<.0001	2
AR1,3	-0.82103	0.08403	-9.77	<.0001	3

Constant Estimate	48.06574
Variance Estimate	637.4319
Std Error Estimate	25.24741
AIC	539.4301
SBC	549.6453
Number of Residuals	57

Correlations of Parameter Estimates

Parameter	MU	MA1,1	AR1,1	AR1,2	AR1,3
MU	1.000	0.007	0.014	-0.002	-0.017
MA1,1	0.007	1.000	0.524	0.219	-0.071
AR1,1	0.014	0.524	1.000	0.694	-0.007
AR1,2	-0.002	0.219	0.694	1.000	0.662
AR1,3	-0.017	-0.071	-0.007	0.662	1.000

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	6.60	2	0.0369	0.115	0.250	0.096	0.140	0.032	0.035
12	7.17	8	0.5184	0.062	0.031	-0.043	-0.033	0.007	0.021

Messages: 6 — User: sasdemo

SAS® Studio — SAS Programmer — Sign Out

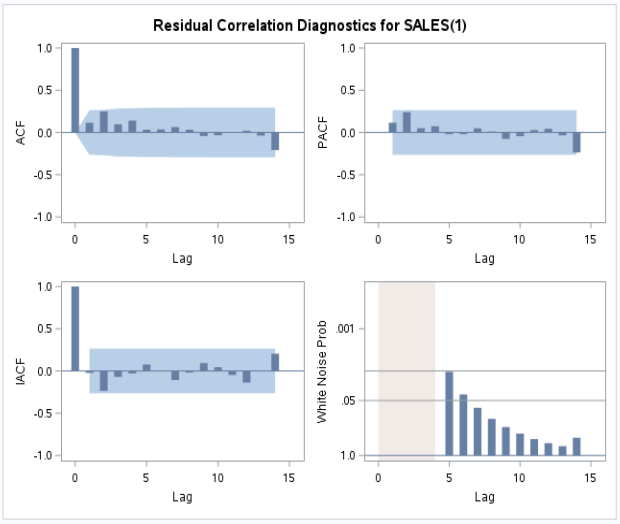
*Baseball | Holidays.csv | *Bikes | *Modeling and Forecasting

Settings | Code/Results | Split — Log | Code

CODE | LOG | RESULTS

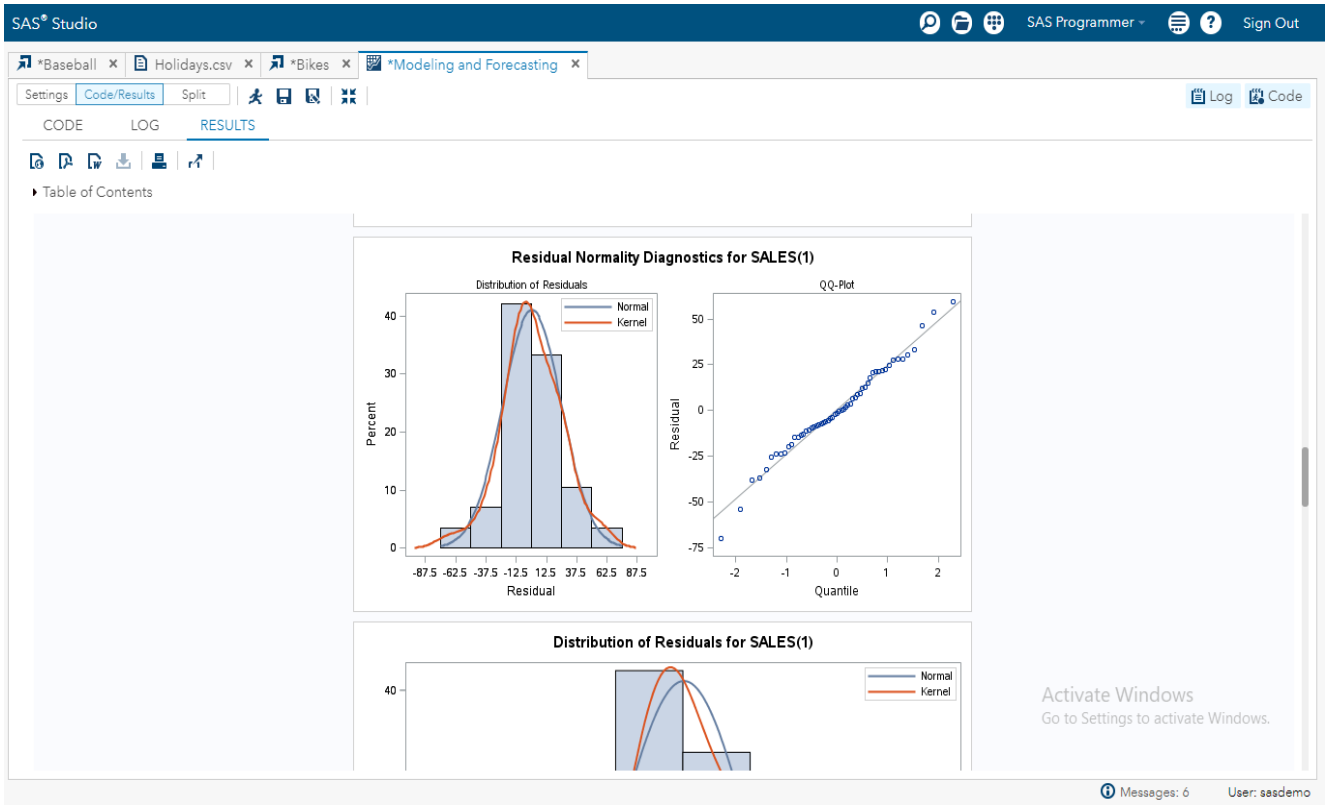
Table of Contents

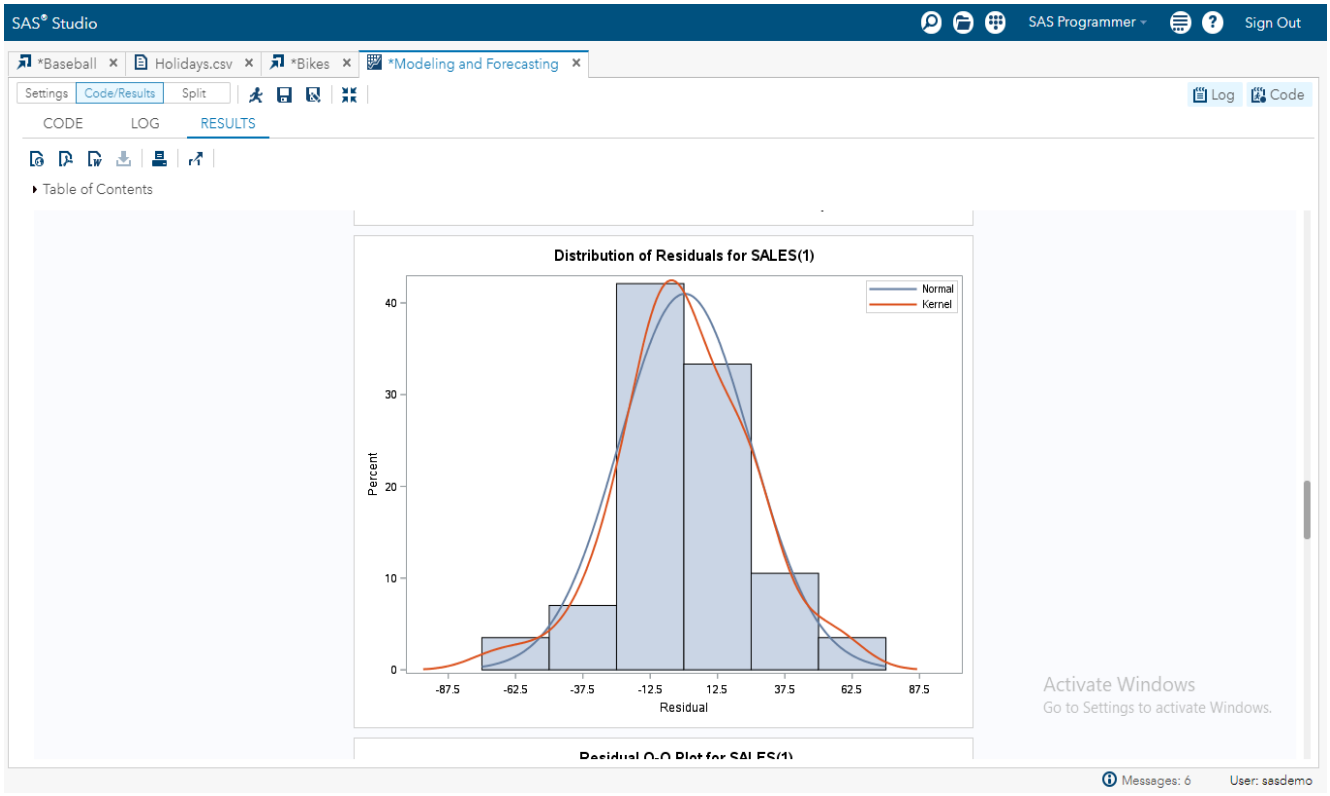
24	13.96	20	0.8323	-0.018	0.100	-0.037	-0.027	0.032	0.055



Activate Windows
Go to Settings to activate Windows.

Messages: 6 — User: sasdemo


SAS® Studio
SAS Programmer
Sign Out
*Baseball
Holidays.csv
*Bikes
*Modeling and Forecasting
Settings
Code/Results
Split
Log
Code
CODE
LOG
RESULTS
Table of Contents
Residual Normality Diagnostics for SALES(1)
Distribution of Residuals
QQ-Plot
Normal
Kernel
Percent
Residual
Quantile
Distribution of Residuals for SALES(1)
Activate Windows
Go to Settings to activate Windows.
Messages: 6
User: sasdemo


SAS® Studio
SAS Programmer
Sign Out
*Baseball
Holidays.csv
*Bikes
*Modeling and Forecasting
Settings
Code/Results
Split
Log
Code
CODE
LOG
RESULTS
Table of Contents
Distribution of Residuals for SALES(1)
Normal
Kernel
Percent
Residual
Activate Windows
Go to Settings to activate Windows.
Messages: 6
User: sasdemo

SAS® Studio

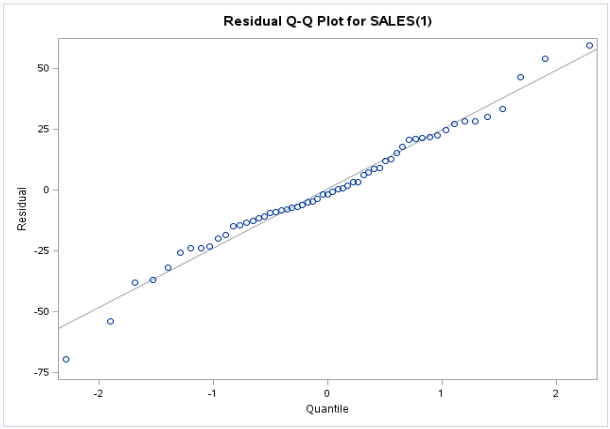
SAS Programmer | Sign Out

*Baseball | Holidays.csv | *Bikes | *Modeling and Forecasting

Settings | Code/Results | Split | Log | Code

CODE | LOG | RESULTS

Table of Contents



Model for variable SALES	
Estimated Mean	12.72244

Messages: 6 User: sasdemo

SAS® Studio

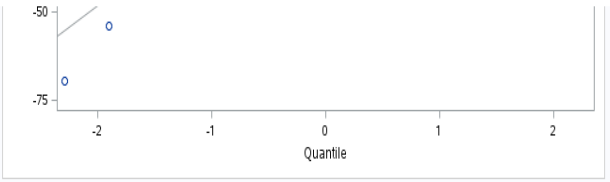
SAS Programmer | Sign Out

*Baseball | Holidays.csv | *Bikes | *Modeling and Forecasting

Settings | Code/Results | Split | Log | Code

CODE | LOG | RESULTS

Table of Contents



Model for variable SALES	
Estimated Mean	12.72244
Period(s) of Differencing	1

Autoregressive Factors	
Factor 1:	1 + 1.06309 B**(1) + 0.89391 B**(2) + 0.82103 B**(3)

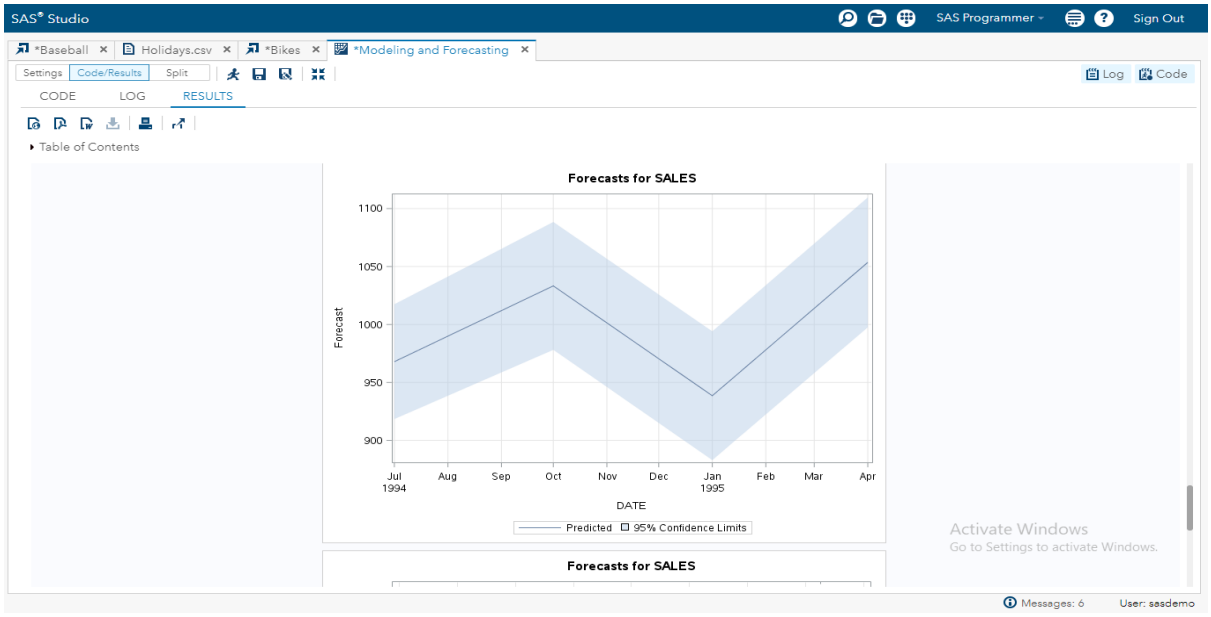
Moving Average Factors	
Factor 1:	1 + 0.55335 B**(1)

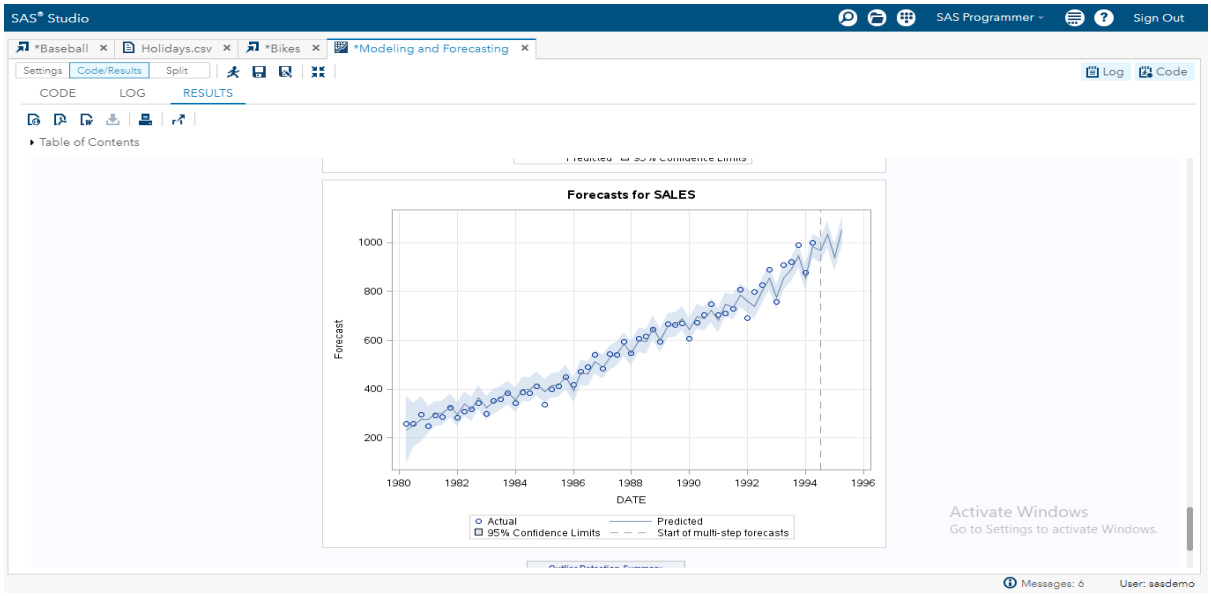
Forecasts for variable SALES

Obs	Forecast	Std Error	95% Confidence Limits	
59	967.9879	25.2474	918.4839	1017.4520
60	1033.3220	28.1184	978.2109	1088.4332
61	938.5908	28.3342	883.0567	994.1249
62	1053.6008	28.5766	997.5915	1109.6097

Forecasts for SALES

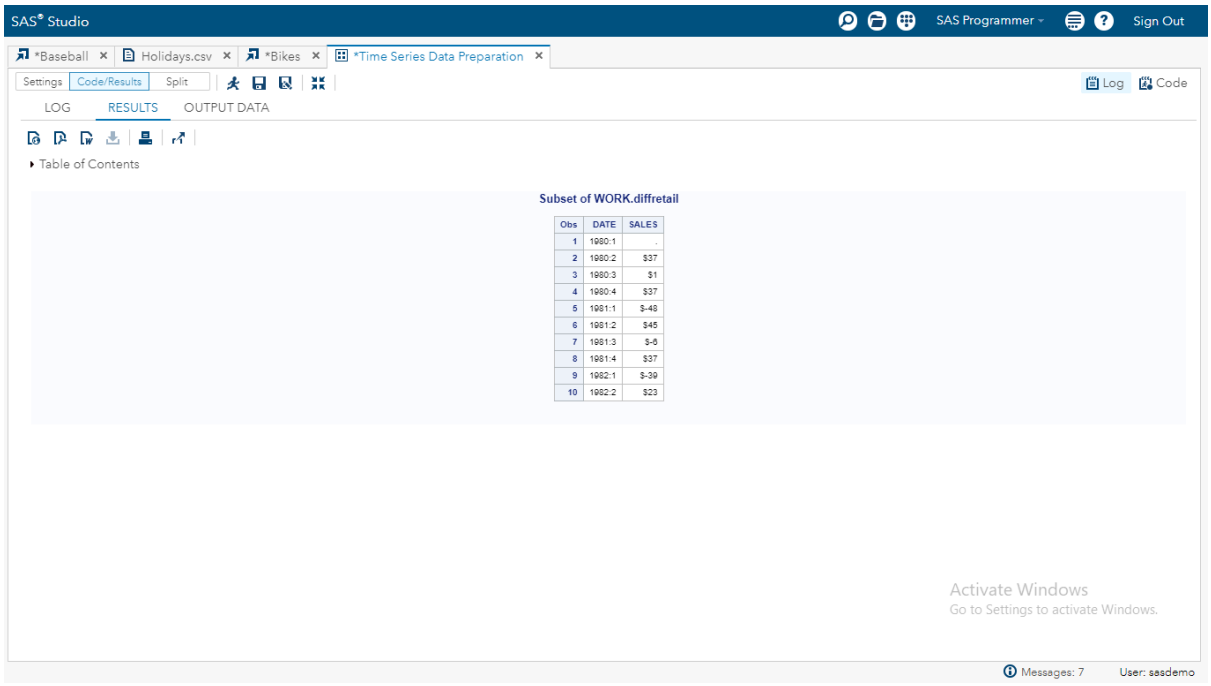
Messages: 6 User: sasdemo





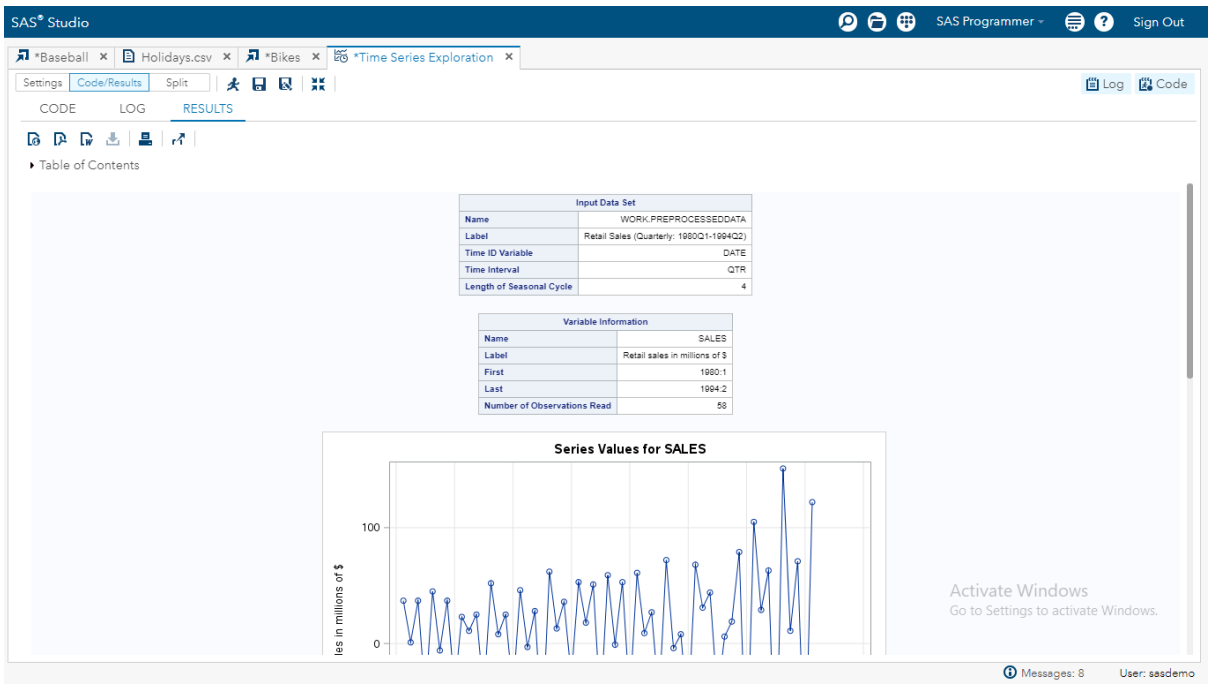
Conclusion:- In this lab I did last step of time series data forecasting and modelling before that we performed data pre-processing and data exploration. I used same retail data from SAS studio we could picked any forecasting model I used in this lab ARIMA model for forecasting. 1st pic shows tables main standard deviation and auto correlation and maximum likelyhood and many other sales forecasting related details this how easy and powerful SAS studio is in term of playing with data or predicting the data I used tableau and power BI by Microsoft but those software are more focused on visualization but on other hand SAS studio help In really core part.

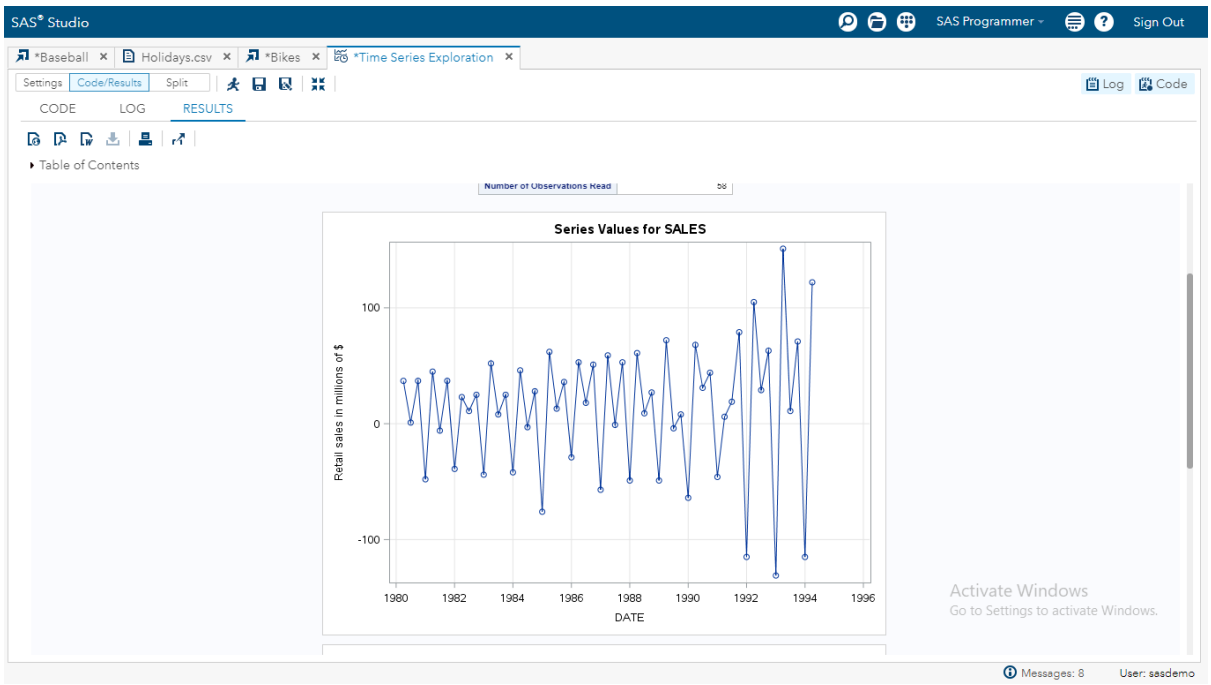
3. Time Series Data Prep

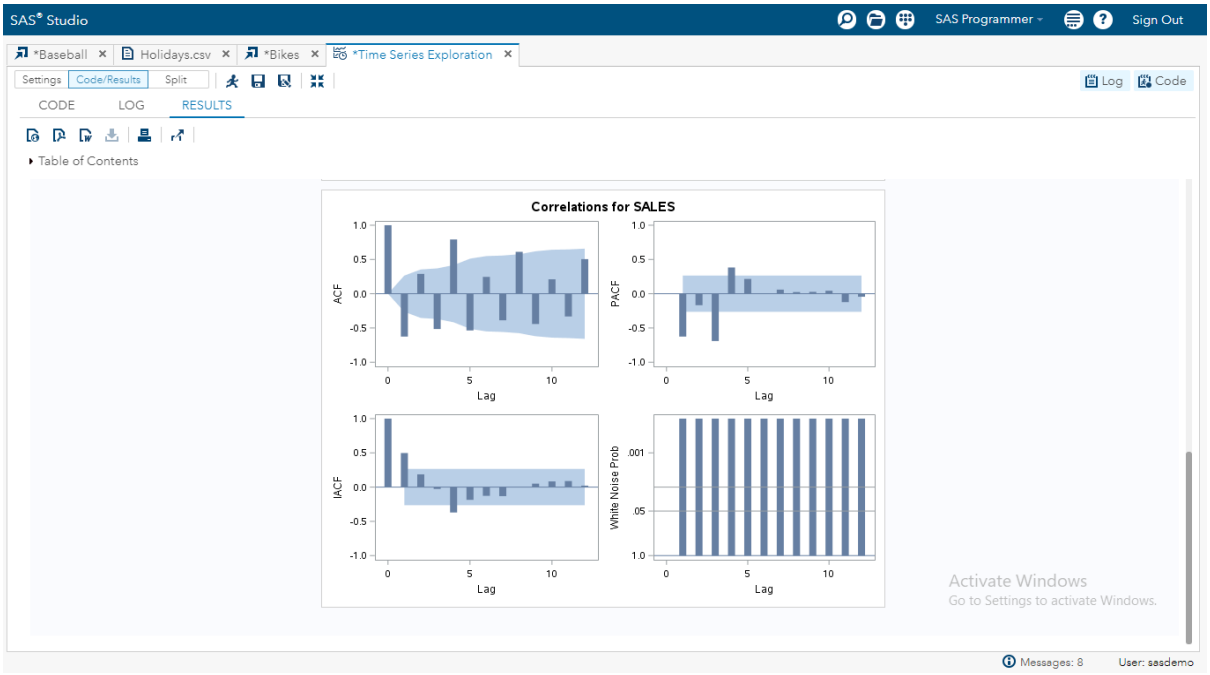


Conclusion: - In this lab I did first step of our time series task which data preparation in lab I was using retail data from sas studio after uploading the data we changed the few settings and I generated first 10 observation and this data set I used later for data exploration and after that will use that for forecasting.

4. Time Series Exploration

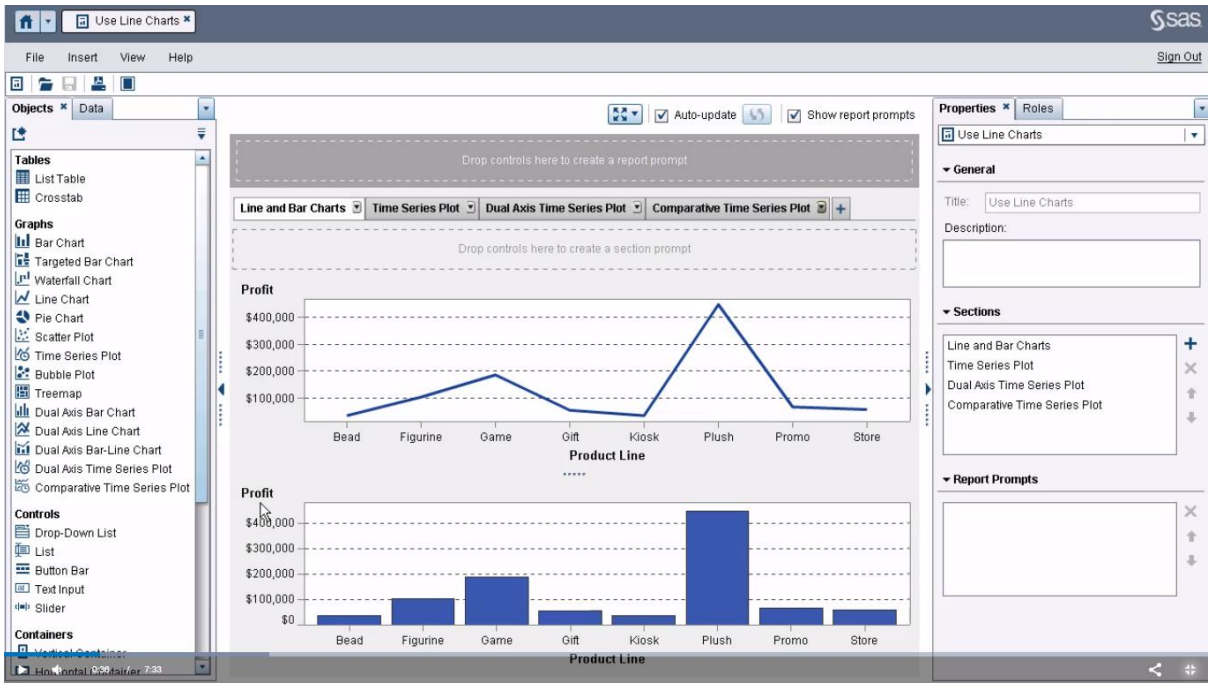


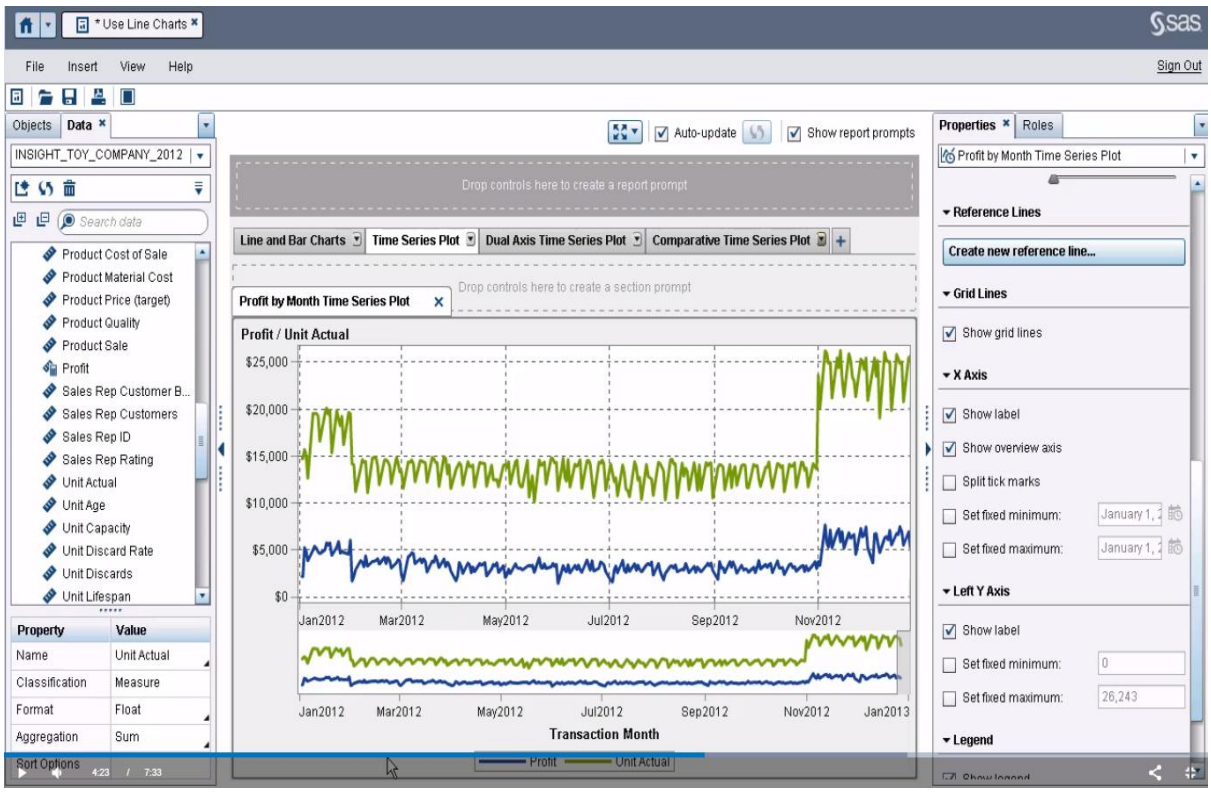




Conclusion: - in this lab I explored our retail data and in retail we explored sales data specifically i changed the SAS studio setting like time , first difference , date , year , interval after that I choose our different visualization charts and graph first graph shows 1st difference sales data quarterly SAS studio is very easy to use all of these complicated task we can use different model and tasks till now I haven't across any other software for data analysis which gives that much power to users

5. Time Series Plotting





Conclusion: - In this lab I learned how to use line chart , time series plots & dual axis charts initially in 1st pic there two graph bar chart and line chart using sas studio I was comparing them with each other then I switched to time series and dual time series axis user interface of sas is very easy after small introduction and practice GUI is easy to learn and it's hard to get all of that feature in other data analysis software