# Time Series and Forecasting - Case 02

Krishna Teja Jagarlapudi

## Input data

Given data is pertaining to monthly volume of commercial bank real-estate loans, in billions of dollars, from January 1973 to October 1978, a total of 70 observations. A time series model is to be built on this data and forecasted for the next two years or 24 months. The data is input into SAS.

## White noise check

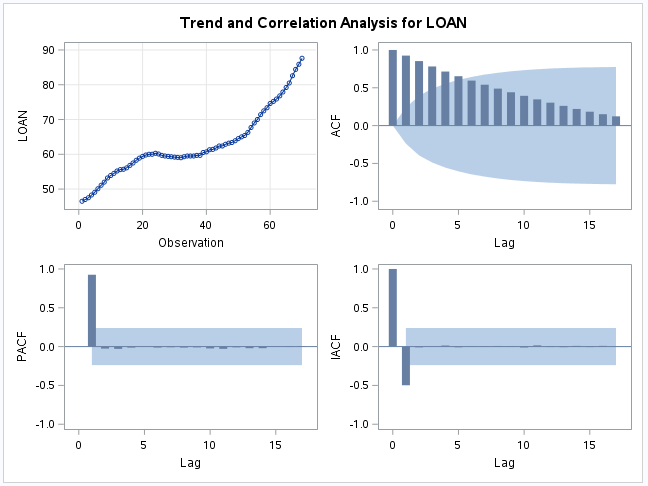
Below code is used to plot the time series data, ACF and PACF plots; a check on auto-correlation in white noise is performed.

proc arima data=case;

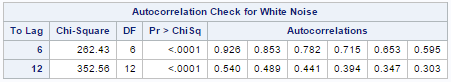
identify var=LOAN;

run;

The plot of the data points and associated ACF and PACF plots are shown below.



Below is the auto-correlation check for white noise. The p-values indicate that the null hypothesis can be rejected – the given data is not white noise.



## Stationarity test

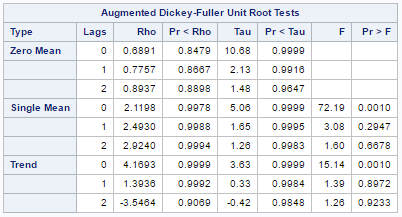
The below code is used to perform the ADF unit root test.

proc arima data=case;

identify var=LOAN stationarity=(adf);

run;

P-values for tau in the ADF test are greater than 5% indicating that the null hypothesis cannot be rejected. This series has a unit root; hence, the series is homogenous non-stationary.



We need to take the first order difference to transform the model and check for stationarity again.

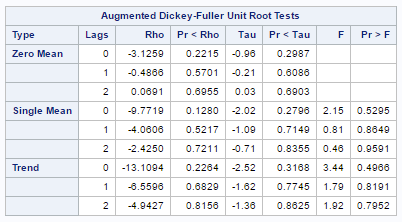
## Transform the data

Below code is used to take the first order difference and check for stationarity.

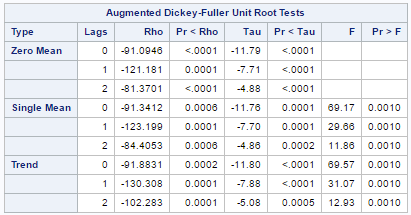
proc arima data=case;

identify var=LOAN(1) stationarity=(adf);

run;



The p-values against Tau show that the data is still not stationary. The next step is to take the second order difference and check the model again for stationarity.



The p-values are for tau are all below the 5% significance level. This proves that the second order difference model is stationary.

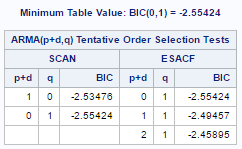
## Model fitting

The ESACF and SCAN methods are used to identify which model parameters for ARIMA(p,d,q) best fits the given data. The code used is shown below.

proc arima data=case;

identify var=LOAN(1,1) esacf scan;

run;



The above table shows possible values for (p+d) and q, based on SCAN and ESACF methods. Considering the estimated model with least BIC ((p+d)=0, q=1); the parameters chosen to test first are p=0 and q=1, i.e., ARIMA(0,2,1); we are already taking difference d=2. The following code is used to perform this check.

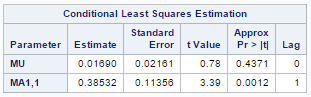
proc arima data=case;

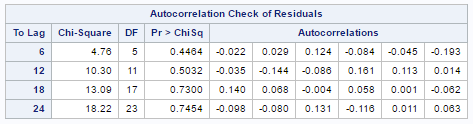
identify var=LOAN(1,1);

estimate q=1;

run;

The mean is 0, as indicated by the p-value against ‘MU’ in below table. The value of is significant at 0.385.





The p-values above indicate that there is no autocorrelation among the residuals and is white noise.

## Forecast

The below code was used for forecasting.

proc arima data=case;

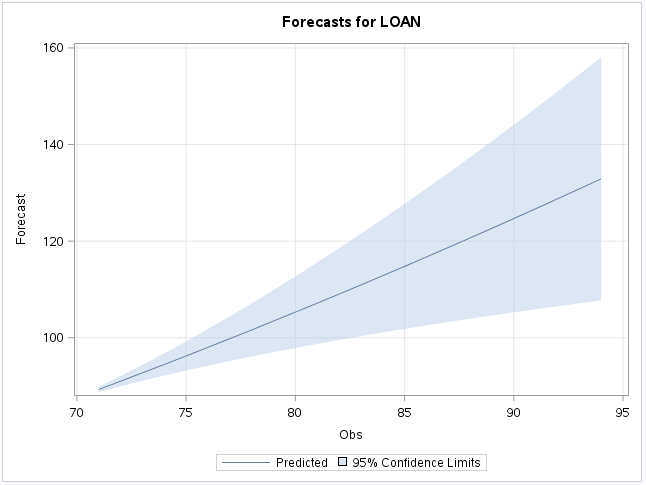
identify var=LOAN(1,1);

estimate q=1;

forecast lead=24;

run;

The plot below shows the forecast values with confidence intervals. As seen, the confidence band is very wide as we go further in the future. Thus, it is not recommended to build a forecast for longer than 12 months. The confidence limits are (99.53,118.54) for the forecast for the 12th month. This limit becomes very wide (107.72, 159.09) for the 24th month forecast. This is a window of 52 billion USD.



## Summary

The given data is found to fit with an ARIMA(0,2,1). Forecast has been made using this predicted model parameters.

It was initially found that the data was non-stationary and first order difference is needed to transform the data into stationary time series. SCAN and ESACF methods are used to estimate values of p and q to identify the best fitting ARIMA model. The diagnostic plots are used to confirm that the data indeed is ARIMA(0,2,1)