Forecasting and Time Series Methods

Non-Stationary Time Series Case

**Ashwita Saxena**

# Problem Statement

The series analyzed here is the monthly volume of commercial bank real-estate loans, in billions of dollars, from January 1973 to October 1978, a total of 70 observations. The data are derived from reports to the Federal Reserve System from large commercial banks.

For such a data set, check for stationarity, build an ARIMA model, perform data transformation/difference (if needed), model identification, model selection, diagnostic checking, parameter estimation, and forecast the next two years. Write a short report summarizing your results, but **no longer than 15 pages**.

# Data

**DATA** CASE;

**DATA** CASE;

DO YEAR=**73** TO **78**;

DO MONTH=**1** TO **12**;

DATE=MDY(MONTH,**1**,YEAR);

one=**1**;

INPUT LOAN @@;

Output;

END;

END;

KEEP DATE LOAN one;

FORMAT DATE MONYY5.;

Title 'COMMERCIAL BANK REAL-ESTATE LOANS';

CARDS;

46.5 47 47.5 48.3 49.1 50.1 51.1 52 53.2 53.9 54.5 55.2 55.6 55.7 56.1

56.8 57.5 58.3 58.9 59.4 59.8 60 60 60.3 60.1 59.7 59.5 59.4 59.3 59.2

59.1 59 59.3 59.5 59.5 59.5 59.7 59.7 60.5 60.7 61.3 61.4 61.8

62.4 62.4 62.9 63.2 63.4 63.9 64.5 65 65.4 66.3 67.7 69 70 71.4

72.5 73.4 74.6 75.2 75.9 76.8 77.9 79.2 80.5 82.6 84.4 85.9 87.6

;

**PROC PRINT**;**RUN**;



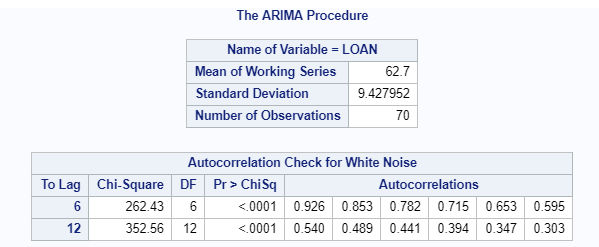
This dataset has 70 observtions

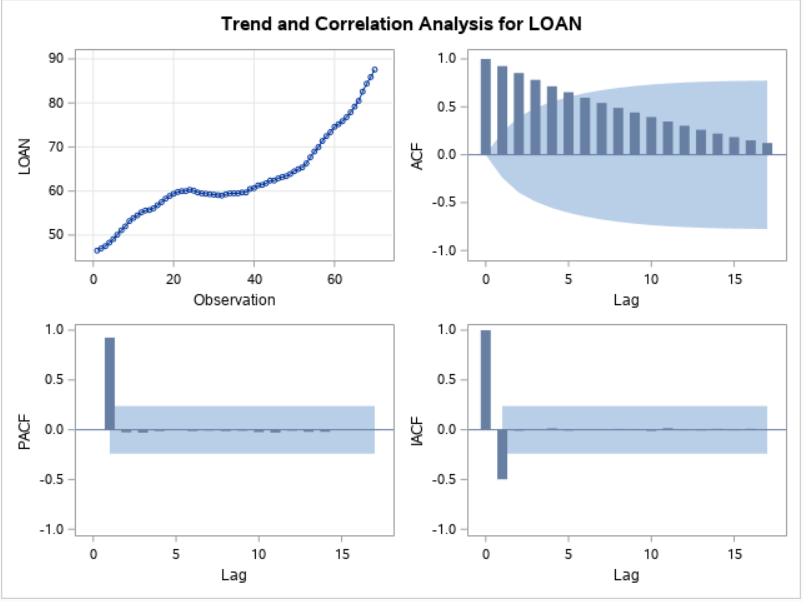
# Identification of the Model

**PROC** **ARIMA**;

IDENTIFY VAR=LOAN;

**RUN**;





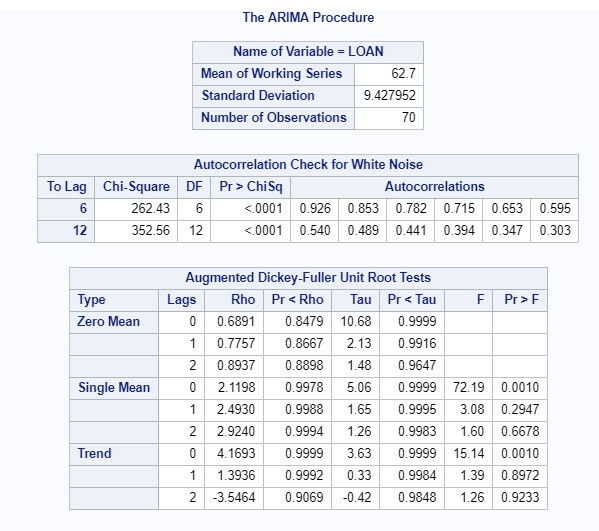
* Mean of this series is 62.7 and standard deviation is 9.428
* Looking at the autocorrelation check, all p-values are significant. Hence we reject the null hypothesis that the data is white noise.
* ACF decays slowly but not exponentially. PACF cuts of lag 1. This looks like non stationary data.

# Stationarity check using ADF test

**PROC** **ARIMA**;

IDENTIFY VAR=LOAN STATIONARITY=(ADF);

**RUN**;

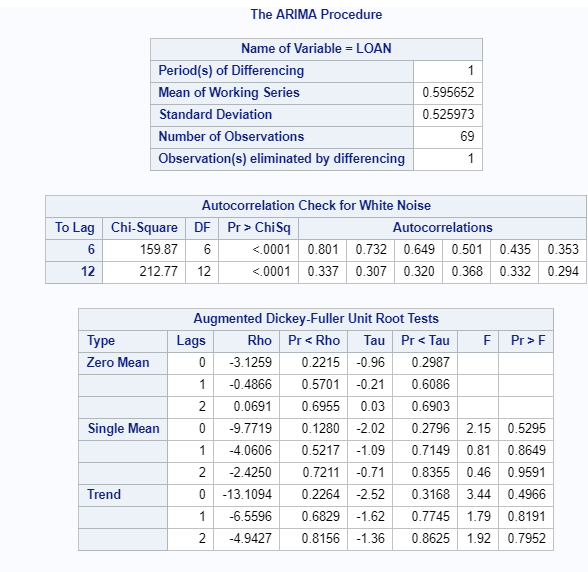


* Null Hypothesis: Data is homogenous non-stationary, Based on these p-values we cannot reject the null.
* We should try first order differentiation

**PROC** **ARIMA**;

IDENTIFY VAR=LOAN (**1**) STATIONARITY=(ADF);

**RUN**;

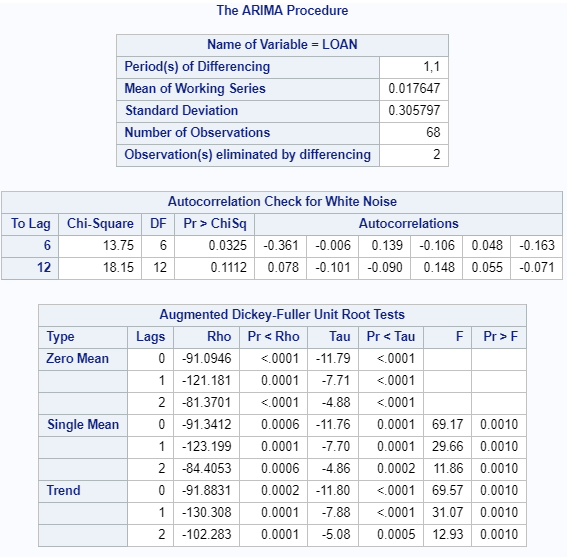


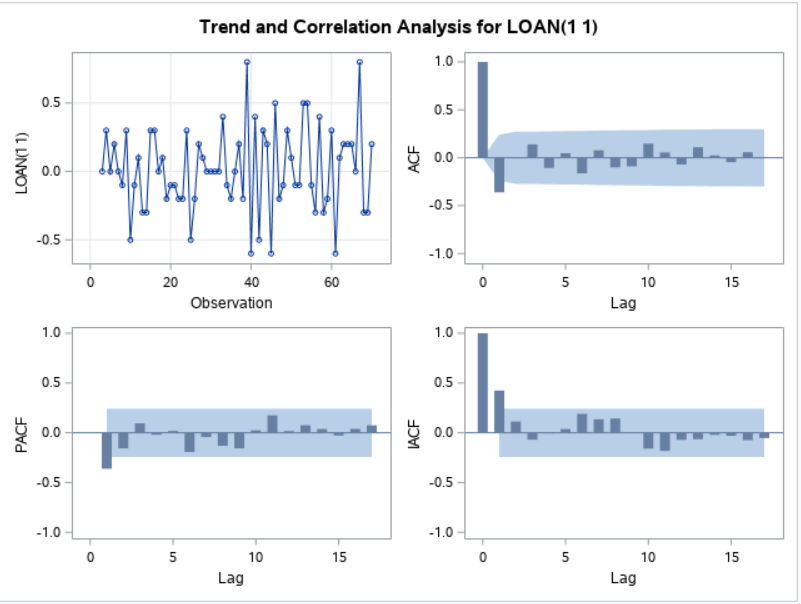
* We still cannot reject the null based on these P-values.
* We need to try second order differentiation

**PROC** **ARIMA**;

IDENTIFY VAR=LOAN (**1**,**1**) STATIONARITY=(ADF);

**RUN**;





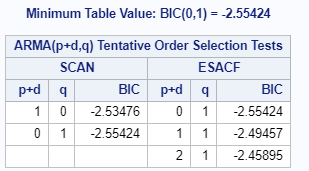
* P-values here are significant and hence we can reject the null hypothesis and conform second order differentiation.
* The ACF and PACF graph do not tell us any specific values for p and q.
* We will perform a BIC test

# Estimation of Model Parameters

**PROC** **ARIMA**;

IDENTIFY VAR=LOAN(**1**,**1**) MINIC ESACF SCAN;

**RUN**;



* According to BIC, the model can be an MA(1) model.
* According to scan and esacf, it might be an AR model, however, we see that this test also recommends MA(1). Also, here, d is 2 if p is 0
* Final model is ARIMA (0,2,1), IMA (2,1)

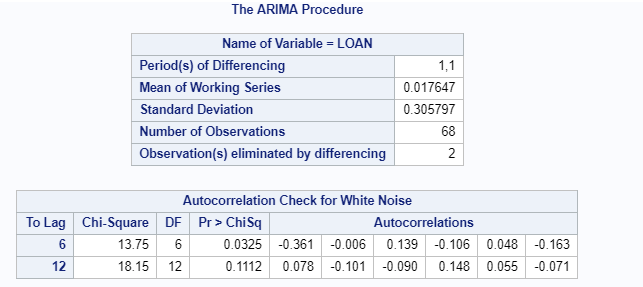
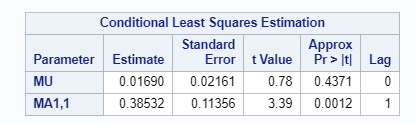
# Residual Diagnostics

**PROC** **ARIMA**;

IDENTIFY VAR=LOAN(**1**,**1**);

estimate q=**1**;

**RUN**;

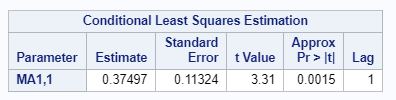
* Here mu is insignificant, hence we remove it from our model using noconstant command

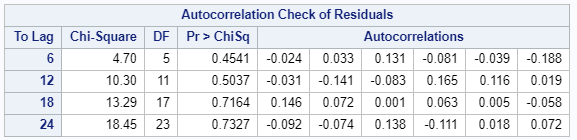
**PROC** **ARIMA**;

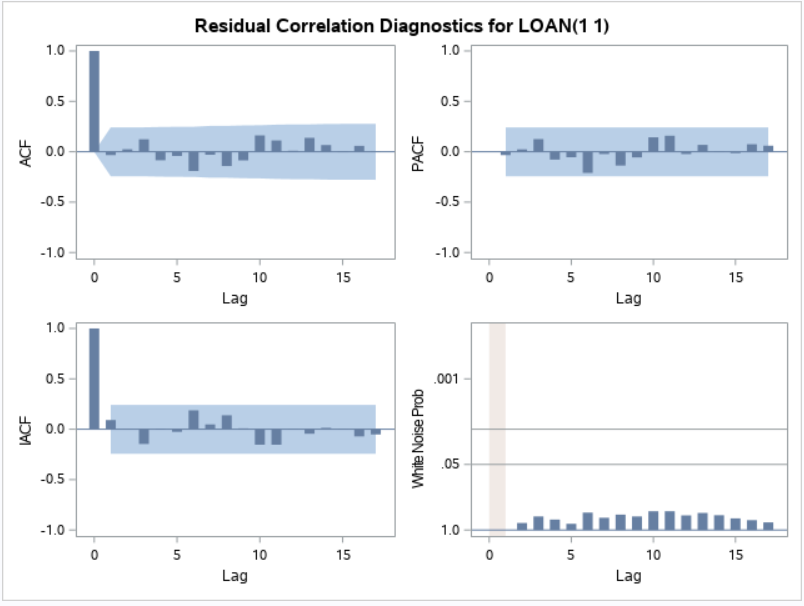
IDENTIFY VAR=LOAN(**1**,**1**);

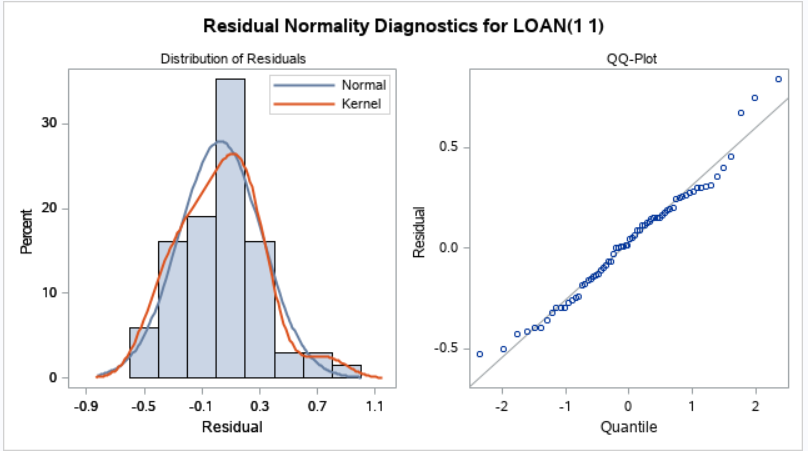
estimate q=**1** noconstant;

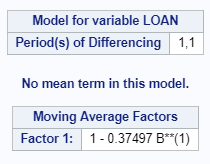
**RUN**;











* MA parameters are significant
* Both ACF and PACF residuals are within the confidence band and hence ARIMA(0,2,1) is our final model.

(1-B)2 \* ZT = (1 - 0.375\*B) (at)

* Residuals follow normal distribution

# Forecasting

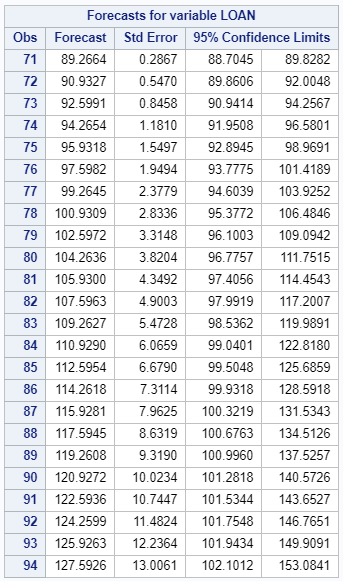
**PROC** **ARIMA**;

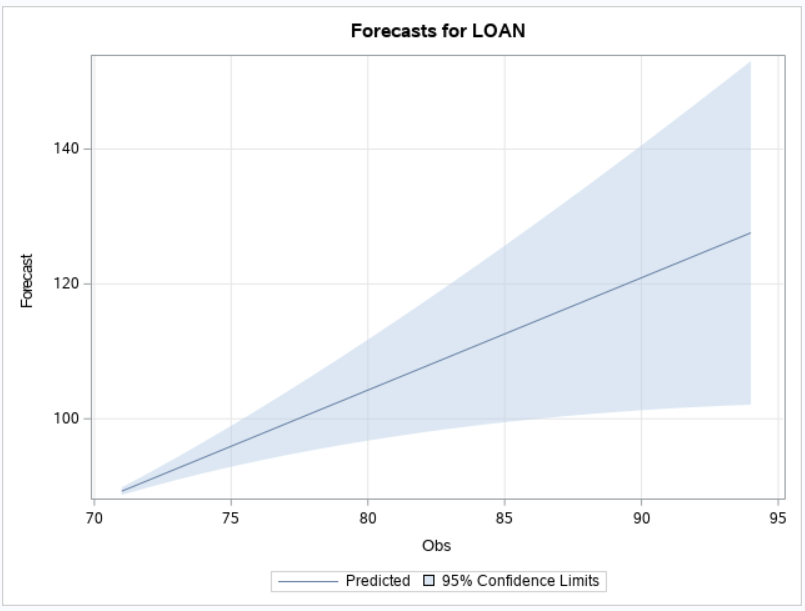
IDENTIFY VAR=loan(**1**,**1**) ;

estimate q=**1** noconstant;

FORECAST LEAD=**24** OUT= results;

**RUN**;





* In the next 2 years loans increase from $89 Billion to $153 Billion according to our forecast