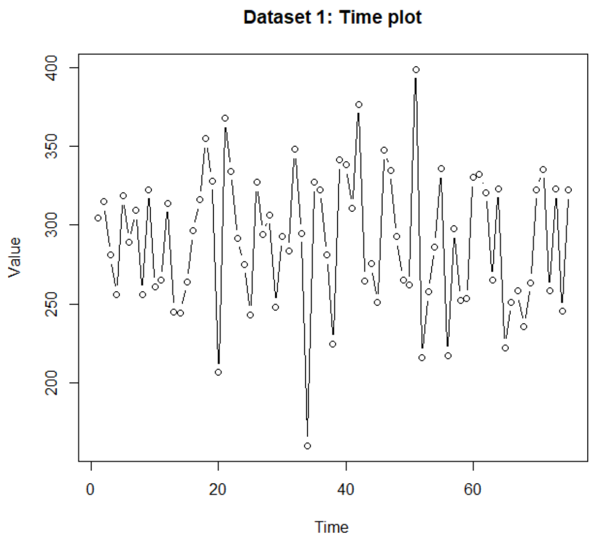
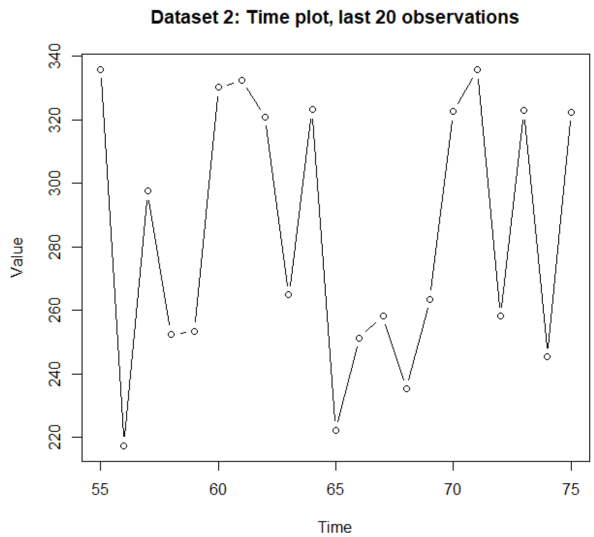
Technical Appendix – data set 2

Summary of the time series, 75 observations,

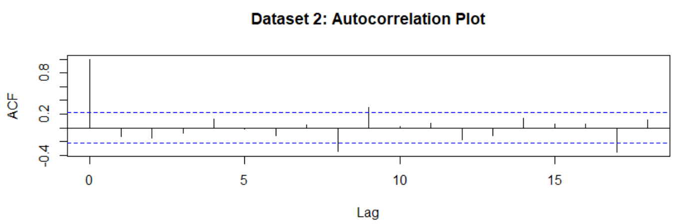
Min.	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*Max.*
159.80	257.90	292.60	290.08	322.85	399

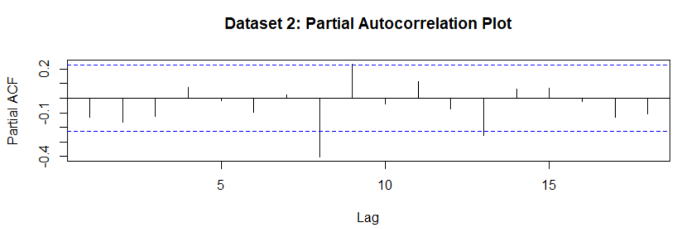
Time series plots





The timeseries plot data set indicates that data is stationary. There is no obvious trend in data. The Augmented Dickey-Fuller test confirmed that a good starting point for modeling this time series is an ARMA model.

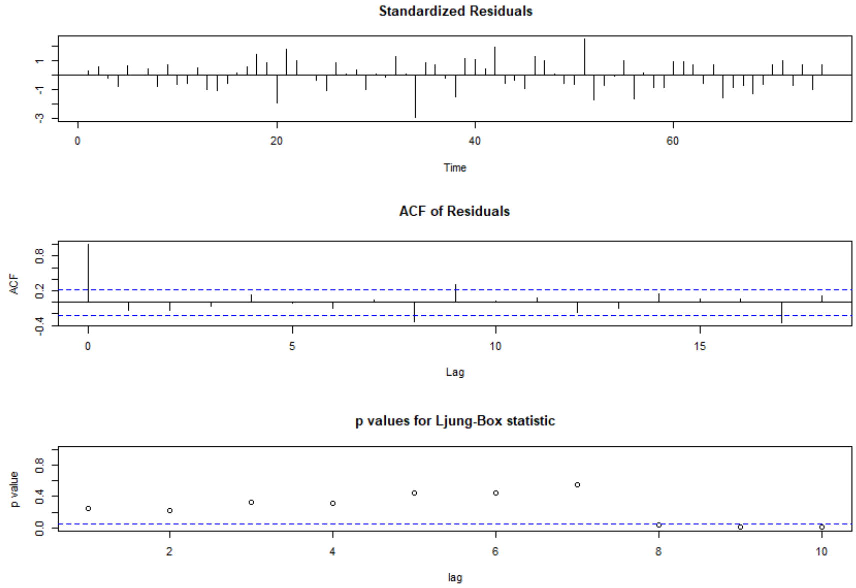




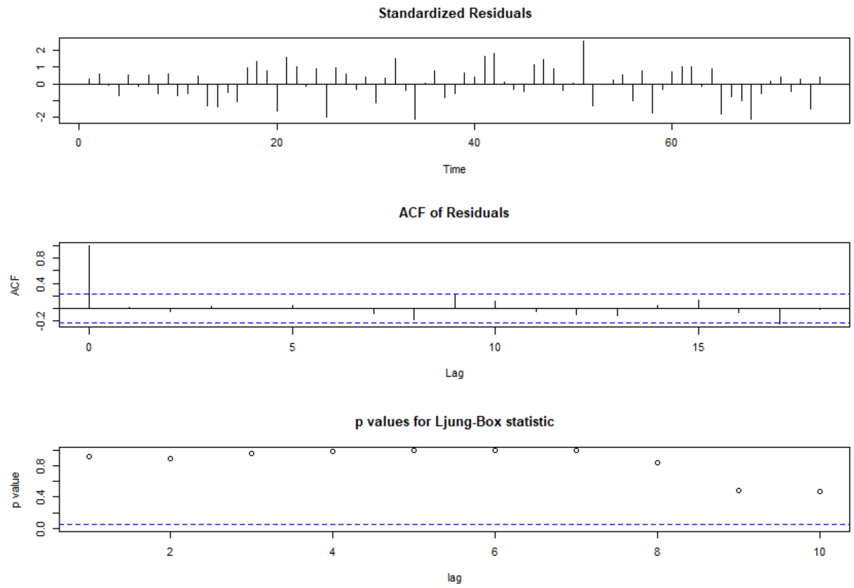
There is a first significant value at lag 9 at **ACF** plot, and then at lag 10 and 17 which might be significant by chance. We can expect around 5% of false significant values. Similarly, the **PACF** plot shows the first significant value at lag 8.
Therefore, we can assume that the process that is behind this time series is in fact a white noise process.
The auto.arima() R function also recommends the ARMA (0,0) model, that in fact is a purely random, white noise process.

AIC 10x10 matrix gave us that the smallest AIC value has model ARMA(6,3), but that can be due a large number of predictors (models with large number of predictors have smaller AIC values).

Judging by the significance of the last coefficient model ARMA(6,3) has insignificant coefficient AR(6) , therefore, not relevant.



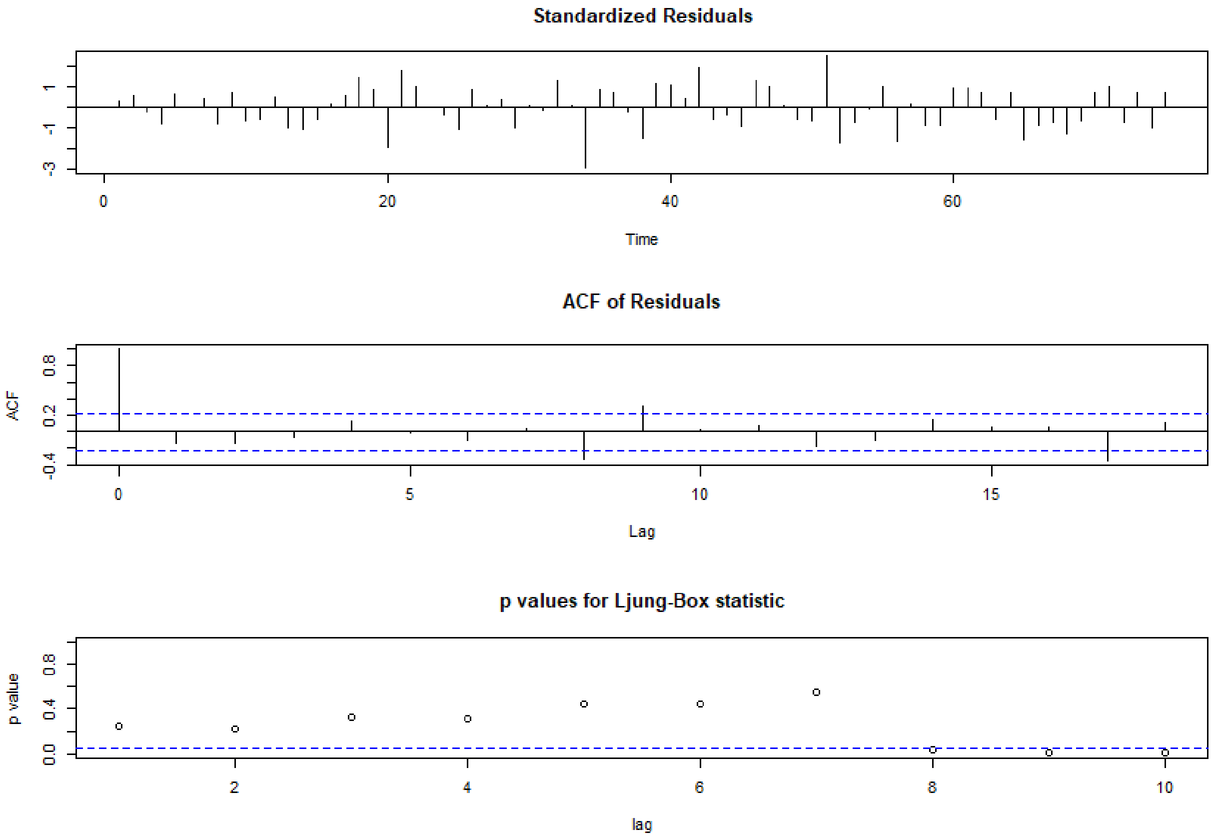
1 - Residual Plot of the model ARMA(0,0)



2 - Residual Plot of the model ARMA(6,3)

The closer examination of the residuals indicated that the smallest model that had random residuals with nonsignificant ACF values is ARMA(6,3). The non-randomness of the residuals of a white noise process ARMA(0,0) illustrates how some amount of ACF values at lags may be falsefully look significant.

3 - Model ARMA(6,3)



Executive report – data set 2

By the thrall analysis of the data from DataSet 2 we concluded that our dataset is best modeled by a white noise, therefore we are 95% confident that all future values will be between

202.26 and 377.9