Topics in Time Series (5)

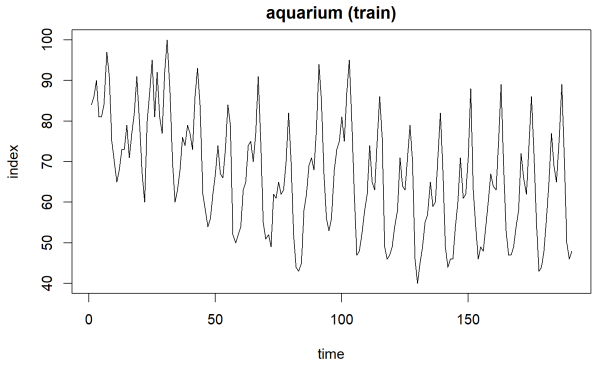
Yisen Du (Eason)

1 Question1

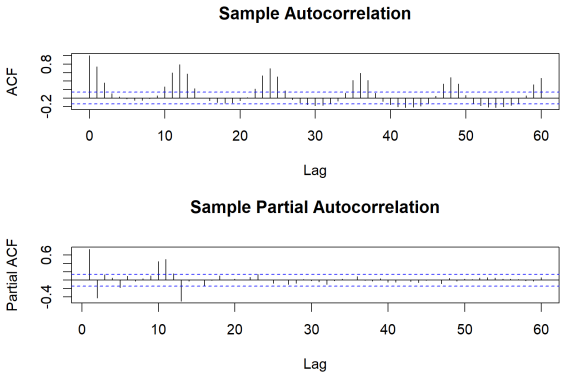
Download the Google trends datasets (2004-present and for the United States) for each of the five queries: aquarium, beer, algebra, hiking, and tutoring. For each dataset, remove the data for the last 36 months and set it aside as the test dataset. Use the Box-Jenkins modeling strategy to fit a model to the training dataset. Clearly describe each of the main steps in your analysis and include relevant plots. Use your fitted model to predict the observations in the test dataset and report the accuracy of the predictions. Include your code as an Appendix. There are twenty points for each of the five analyses.

1.1 Answer of Q1(a): aquarium

1.1.1 Preprocess the data

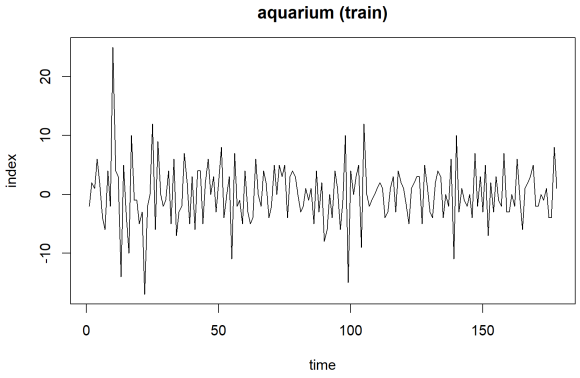


(a) Overview of data

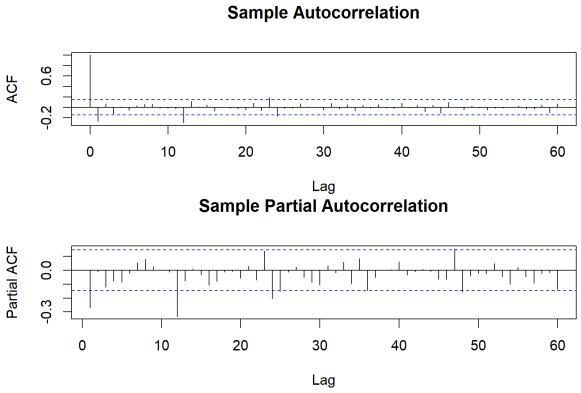


(b) ACF and PACF

Note that there's a periodicity of 12 months of data and there are some increasing and decreasing trends. Therefore, we take the seasonal differences and the first difference to preprocess the data. The below figure shows after the preprocessing, the data looks stationary.



(a) Transformed data



(b) ACF and PACF

1.1.2 Fit the model

The ACF and PACF of transformed data indicate $AR(1)$ or $MA(1)$ may be good models. After computing AIC and BIC of several $ARMA$ models, we select the models with the smallest AIC and BIC.
Note that there are spikes on 12, 24 of ACF and 12, 24 of PACF. We may also consider the multiplicative seasonal $ARMA$ models. The following models are considered and fitted.

Selection Criteria	Model	Accuracy (log likelihood)
the smallest AIC	$ARMA(4,5)$	-690.95
the smallest BIC	$ARMA(2,1)$	-705.66
Consider seasonal	$ARMA(4,5) \times (2,2)_{12}$	-621.28
Consider seasonal	$ARMA(4,5) \times (3,2)_{12}$	-618.85

Table 1

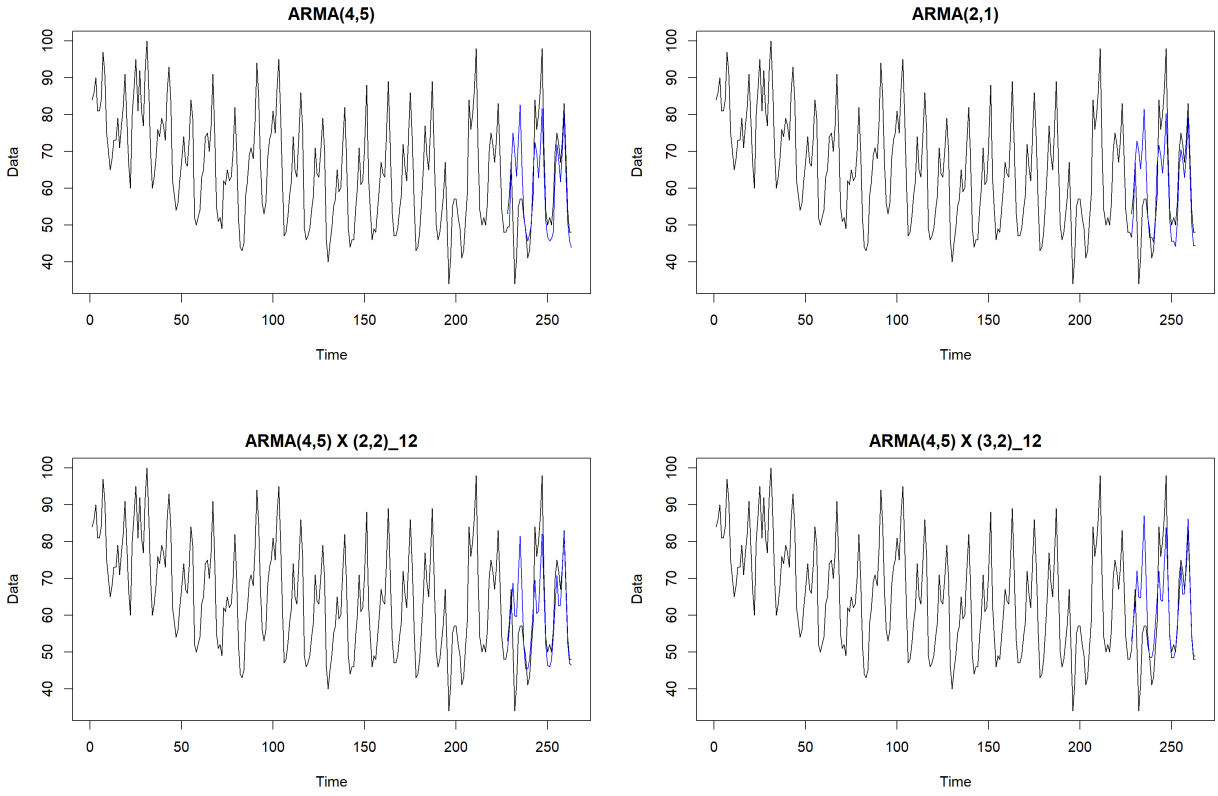


Figure 3: Fitted models

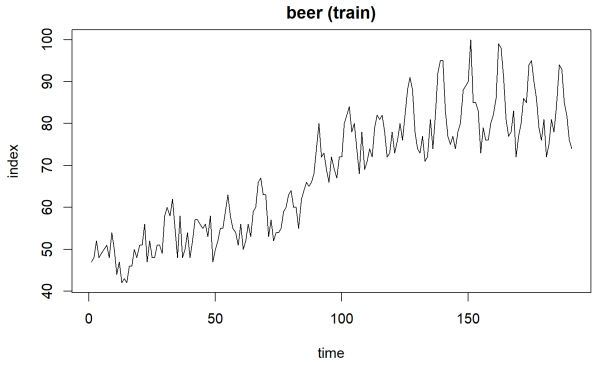
1.2 Answer of Q1(b): beer

1.2.1 Preprocess the data

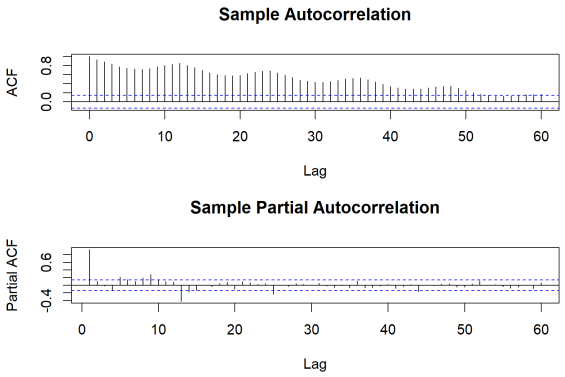
Note that there's a periodicity of 12 months of data and there is an increasing trend. Therefore, we take the seasonal differences and the first difference to preprocess the data. The below figure shows after the preprocessing, the data looks stationary.

1.2.2 Fit the model

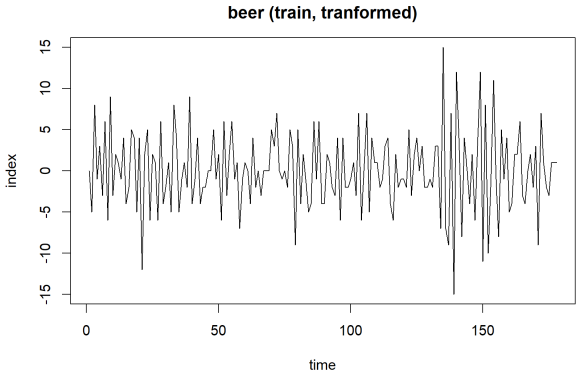
The ACF and PACF of transformed data indicate $ARMA$ models may be good models. After computing AIC and BIC of several $ARMA$ models, we select the models with the smallest AIC and BIC.



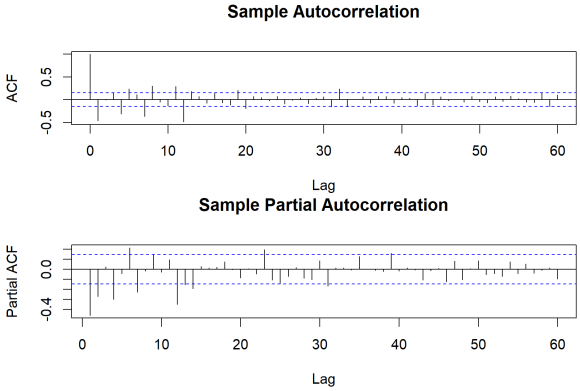
(a) Overview of data



(b) ACF and PACF



(a) Transformed data



(b) ACF and PACF

Note that there are spikes around 12 of ACF and 12, 24 of PACF. We may also consider the multiplicative seasonal $ARMA$ models. The following models are considered and fitted.

Selection Criteria	Model	Accuracy (log likelihood)
the smallest AIC	$ARMA(3,5)$	-590.68
the smallest BIC	$ARMA(3,3)$	-586.51
Consider seasonal	$ARMA(3,5) \times (1,1)_{12}$	-547.61
Consider seasonal	$ARMA(3,5) \times (2,1)_{12}$	-547.55

Table 2

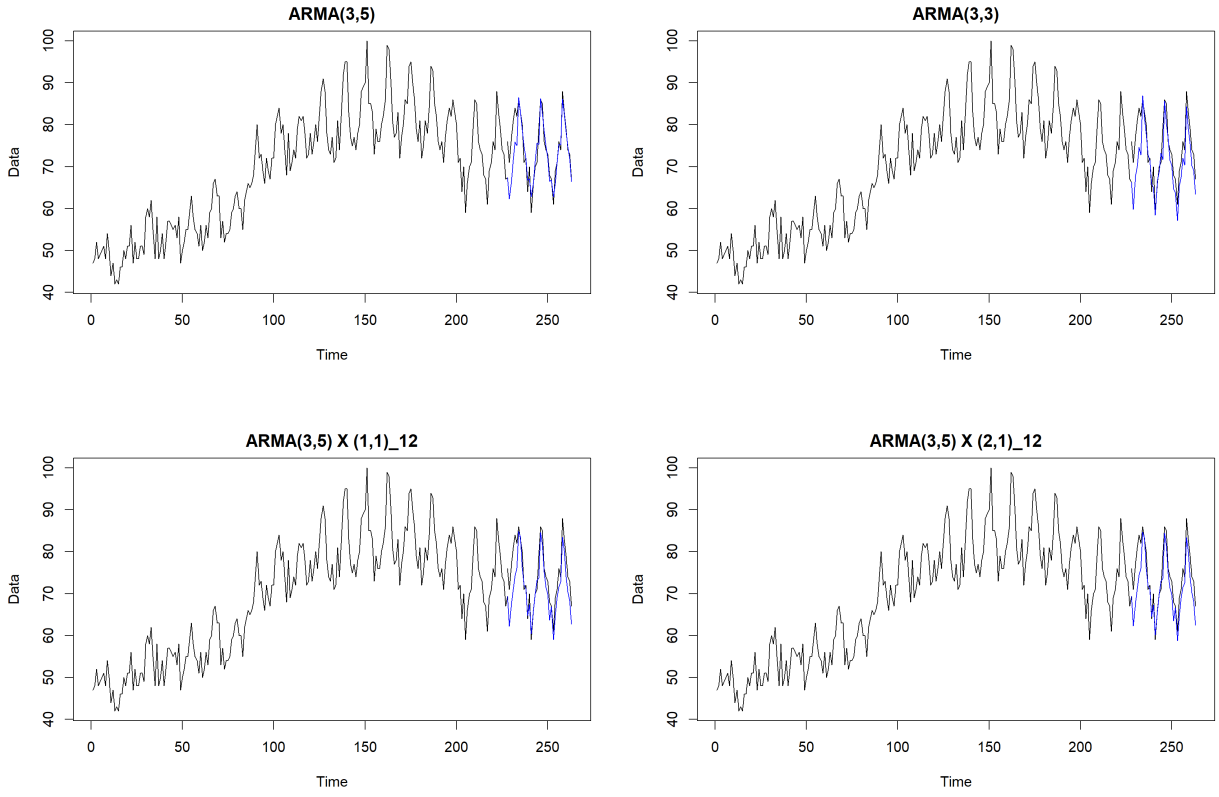
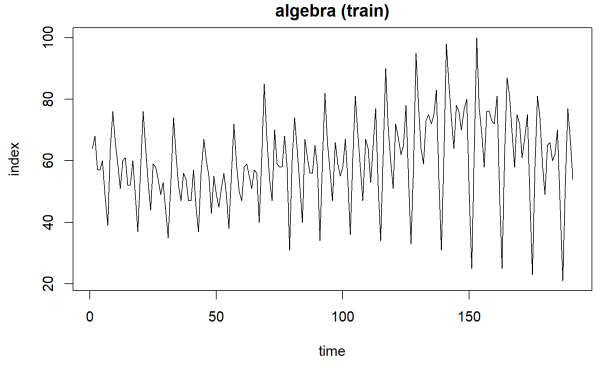


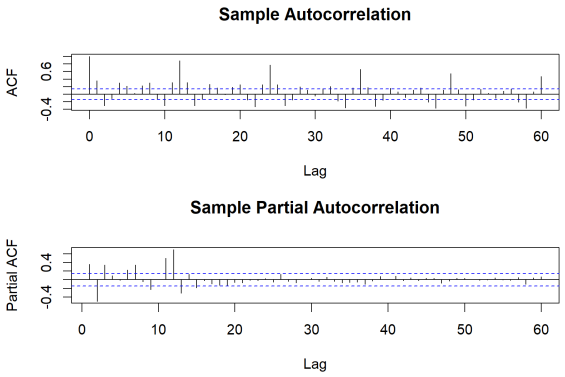
Figure 6: Fitted models

1.3 Answer of Q1(c): algebra

1.3.1 Preprocess the data

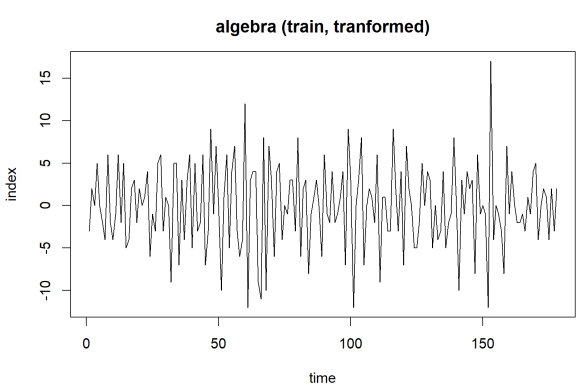


(a) Overview of data

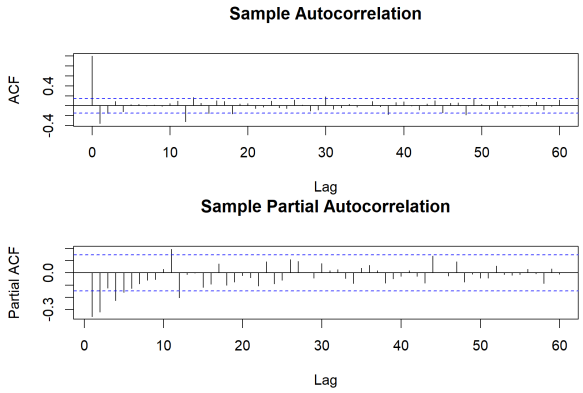


(b) ACF and PACF

Note that there's a periodicity of 12 months of data and there are some increasing and decreasing trends. Therefore, we take the seasonal differences and the first difference to preprocess the data. The below figure shows after the preprocessing, the data looks stationary.



(a) Transformed data



(b) ACF and PACF

1.3.2 Fit the model

The ACF and PACF of transformed data indicate $AR(4)$ and $MA(1)$ models may be good models. After computing AIC and BIC of several $ARMA$ models, we select the models with the smallest AIC and BIC.
Note that there are spikes around 12 of ACF and PACF. We may also consider the multiplicative seasonal $ARMA$ models. The following models are considered and fitted.

Selection Criteria	Model	Accuracy (log likelihood)
the smallest AIC	$ARMA(2,2)$	-590.68
the smallest BIC	$MA(2)$	-586.51
Consider seasonal	$ARMA(2,2) \times (1,1)_{12}$	-547.61

Table 3

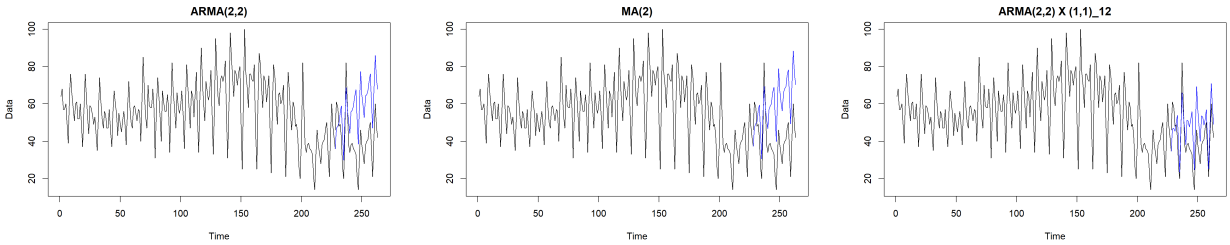
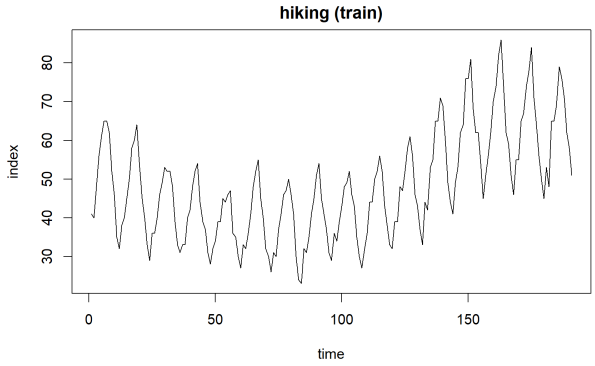


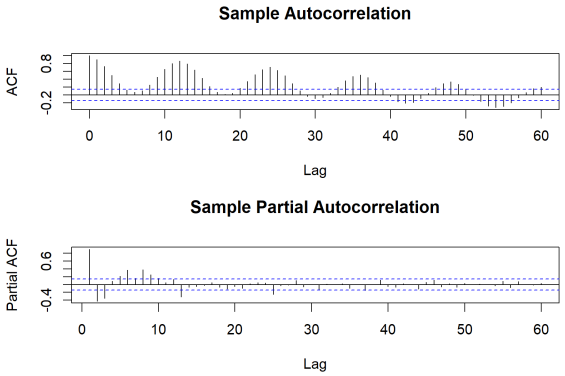
Figure 9: Fitted models

1.4 Answer of Q1(d): hiking

1.4.1 Preprocess the data

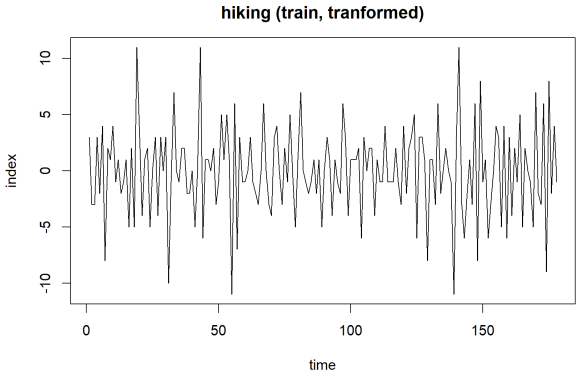


(a) Overview of data

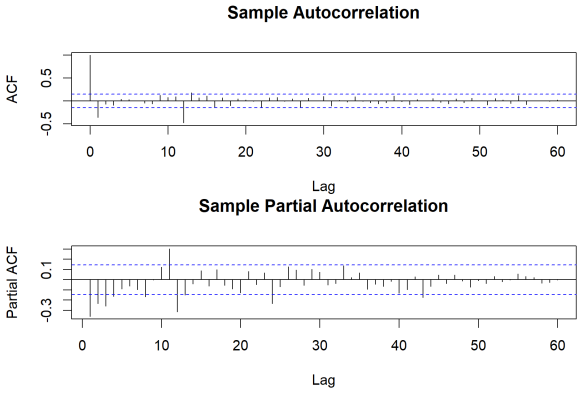


(b) ACF and PACF

Note that there's a periodicity of 12 months of data and there are some increasing and decreasing trends. Therefore, we take the seasonal differences and the first difference to preprocess the data. The below figure shows after the preprocessing, the data looks stationary.



(a) Transformed data



(b) ACF and PACF

1.4.2 Fit the model

The ACF and PACF of transformed data indicate $AR(4)$ and $MA(1)$ models may be good models. After computing AIC and BIC of several $ARMA$ models, we select the models with the smallest AIC and BIC.
Note that there are spikes around 12 of ACF and 12, 24 of PACF. We may also consider the multiplicative seasonal $ARMA$ models. The following models are considered and fitted.

Selection Criteria	Model	Accuracy (log likelihood)
the smallest AIC	$ARMA(5,5)$	-640.1
the smallest BIC	$MA(1)$	-660.69
Consider seasonal	$ARMA(5,5) \times (1,1)_{12}$	-593.65
Consider seasonal	$ARMA(5,5) \times (2,1)_{12}$	-594.34

Table 4

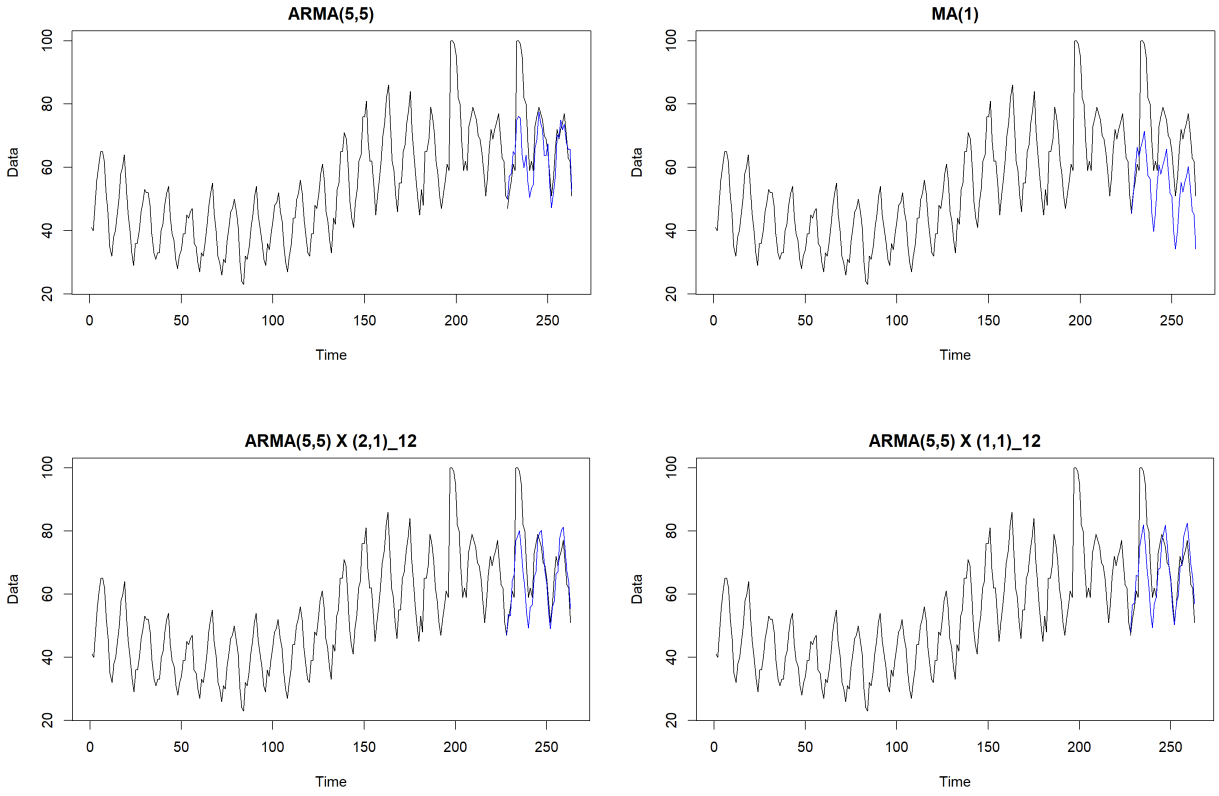


Figure 12: Fitted models

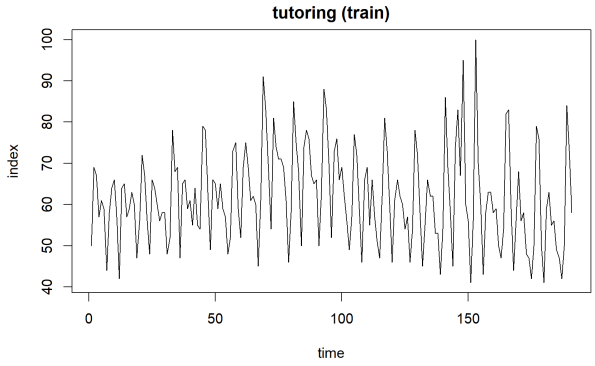
1.5 Answer of Q1(e): tutoring

1.5.1 Preprocess the data

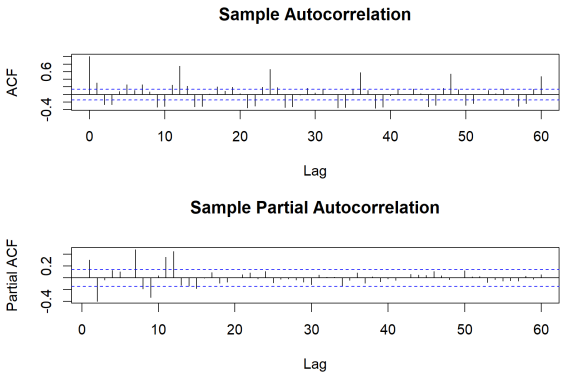
Note that there's a periodicity of 12 months of data and there are some increasing and decreasing trends. Therefore, we take the seasonal differences and the first difference to preprocess the data. The below figure shows after the preprocessing, the data looks stationary.

1.5.2 Fit the model

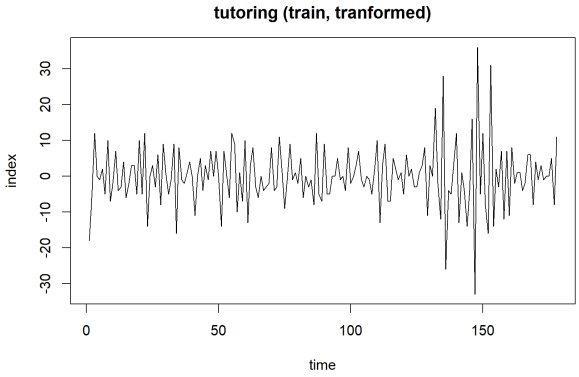
The ACF and PACF of transformed data indicate $MA(1)$ and $AR(5)$ models may be good models. After computing AIC and BIC of several $ARMA$ models, we select the models with the smallest AIC and BIC.



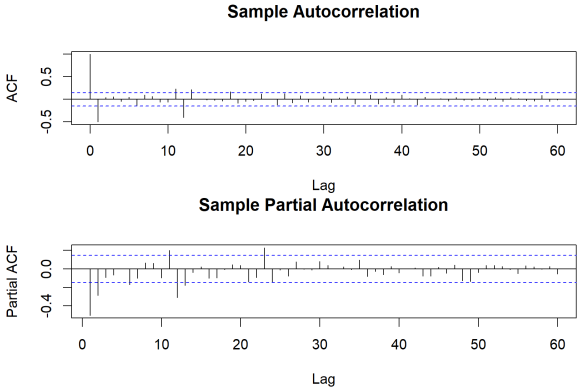
(a) Overview of data



(b) ACF and PACF



(a) Transformed data



(b) ACF and PACF

Note that there are spikes around 12 of ACF and 12, 24 of PACF. We may also consider the multiplicative seasonal $ARMA$ models. The following models are considered and fitted.

Selection Criteria	Model	Accuracy (log likelihood)
the smallest AIC	$ARMA(5,3)$	-720.97
the smallest BIC	$MA(1)$	-733.61
Consider seasonal	$ARMA(5,3) \times (2,1)_{12}$	-684.36

Table 5

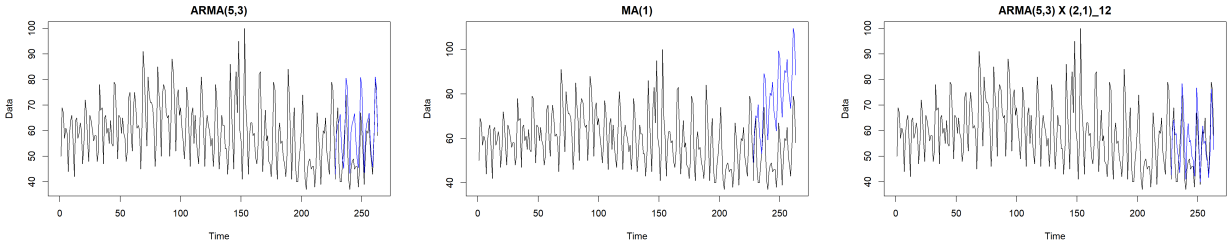


Figure 15: Fitted models