1.Assume all of following R functions are relevant to a time series which you want to build an Arima model over (not using auto.arima)

R functions: ndiffs, nsdiffs, diff, tsdisplay, autoplot, monthdays, BoxCox, Arima, checkresiduals

Discuss briefly what order you should call these functions in to estimate an Arima model. Assume all functions must be called exactly once except for diff (which you can use as few or as many times). There may be more than 1 correct answer, focus on whether your sequence will work in practice.

step1 if the data is aggregated by monthly, use `monthdays()` to find out the daily average to address calendar-related issues

step2 Use `autoplot` to visualize the data to see if data has unstable variance or exhibits a clear departure from normality

step3 If so, call `BoxCox` to address these issues

step4 Use `nsdiffs` to see how many seasonal difference operations are needed to achieve seasonal stationarity

step5 Use `ndiffs` and `diff` to check how many defference opeations are need to achieve stationarity

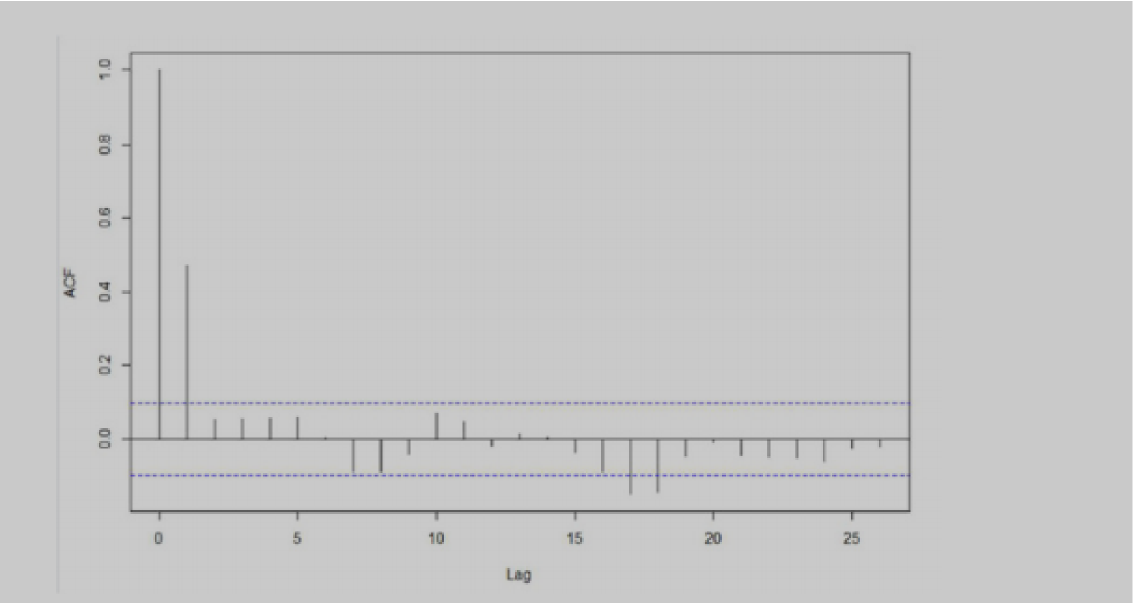
step6 Use `diff` function to difference the time series data to make it stationary

step7 Use `tsdisplay` to visualize the statinary time series, according to the ACF and PACF diagrams to suggest possible hypotheses for ARIMA model

step8 Use `Arima` and `checkresiduals` to check if the residuals of the ARIMA models meet assumption, including independence, zero mean and constant variance.

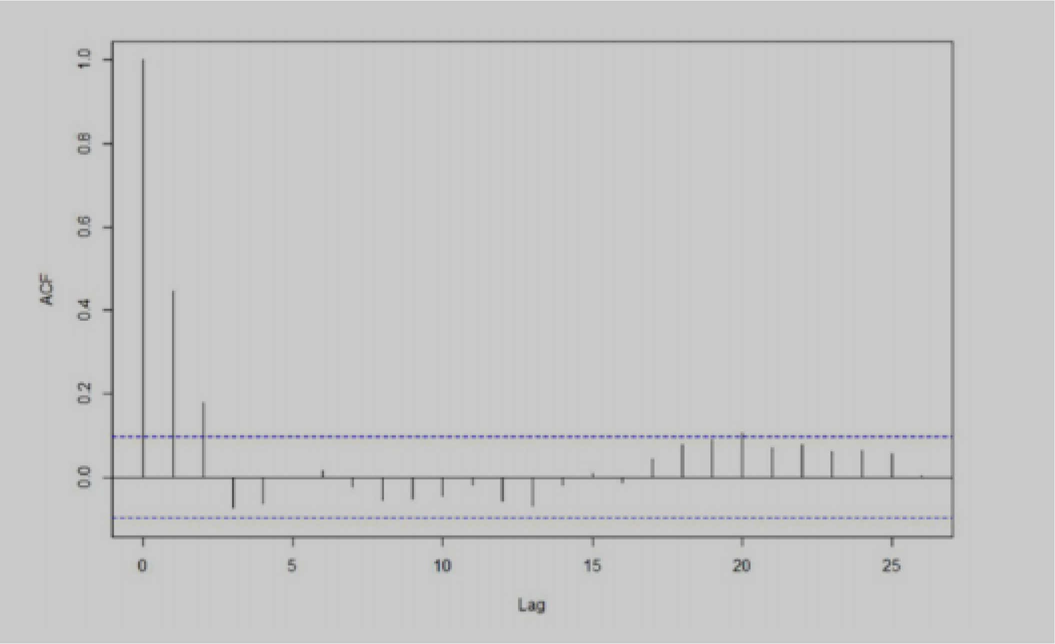
2.The following 3 ACFs have been estimated over artificially simulated time series.

Propose possible ARIMA model(s). There may be more than 1 correct answer, although combining clearly ‘impossible’ suggestions with feasible answers may result in net 0 points awarded

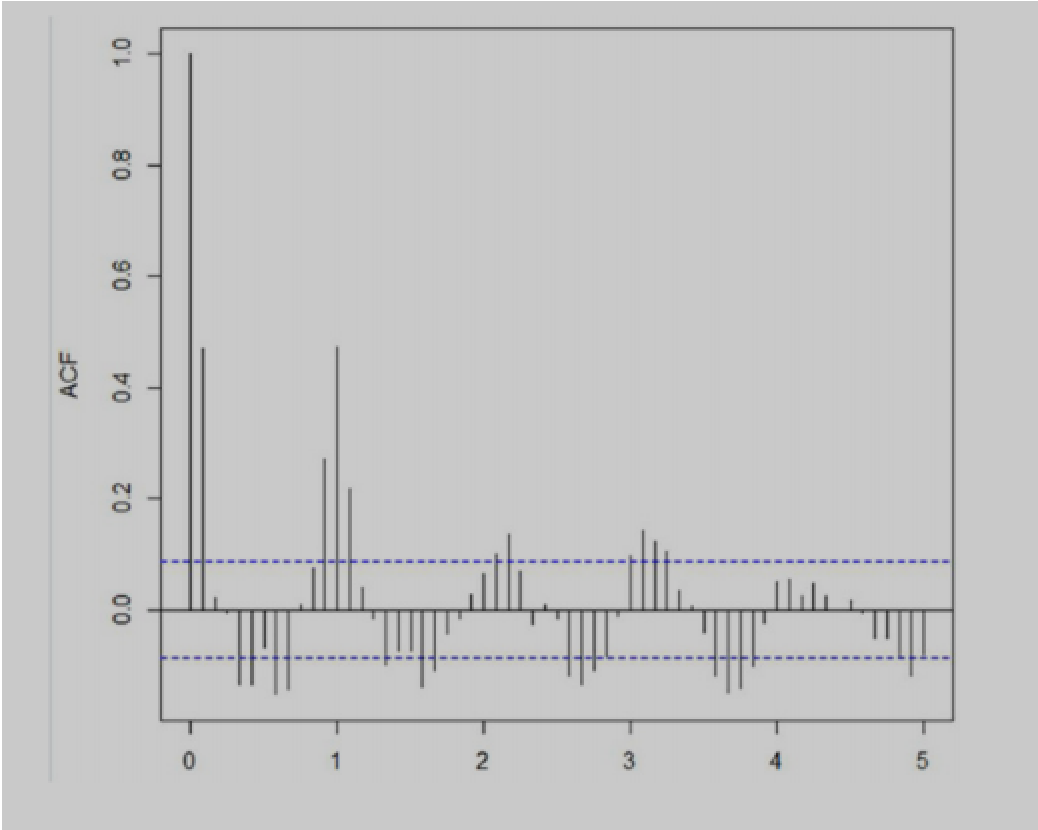


Fast decaying pattern in ACF probably eliminates possibility of AR model, although we are not 100% sure.

I would say it's a `ARIMA(0,0,1)`



Same reasoning here, I would say it's a `ARIMA(0,0,2)`



Fast decaying nonseasonal and seasonal pattern in ACF probably eliminates possibility of AR model.

I would say it's a `ARIMA(0,0,1)(0,0,1)_12` , `ARIMA(0,0,1)(0,0,2)_12` or `ARIMA(0,0,1)(0,0,3)_12`

3.We wish to forecast two variables into the future, and both depend on each other.We consider whether to build a VECM. As a feasibility test, we run the Johansen procedure.

How many cointegrating relationships are there in the data, based on output below?

```
########################
# Johansen-Procedure #
########################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.073463811 0.005621199

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  2.71  6.50  8.18 11.65
r = 0  | 36.63 12.91 14.90 19.19

Eigenvectors, normalised to first column:
(These are the cointegration relations)

             short.run.12 long.run.12
short.run.12     1.000000   1.0000000
long.run.12     -1.022065  -0.2409262

Weights W:
(This is the loading matrix)

            short.run.12  long.run.12
short.run.d  -0.01161738 -0.009806832
long.run.d    0.08880969 -0.014155748
```

There are only 2 variables.

Hence, we either have 0 or 1 cointegrating relationship.

We always choose smallest number of cointegrating relationships where we fail to reject the null.

In this case, `r=1`

4.Regardless of your answer to B3, outline an appropriate course of action if there are no cointegrating relationships.

i.e. what model should we then build, should we render the data stationary, outline how we go about building such a model

If we can reject null for all values, all data series appear to be stationary, so we can jsut `bulid a VAR on the raw data`

If r = 0, there are no cointegrating relationships, and we `build a VAR on the differenced variables`

5.Regardless of your answer to B3, outline an appropriate course of action if:

5.1. There is exactly 1 cointegrated relationship

We bulid a VECM model

step1:Number of VAR lag will be order of VECM model. Use `VARselect` and `SC(n)` to determine number of VAR lag.

step2:We build VECM model with lag = 2 and 1 cointegrating relationship.

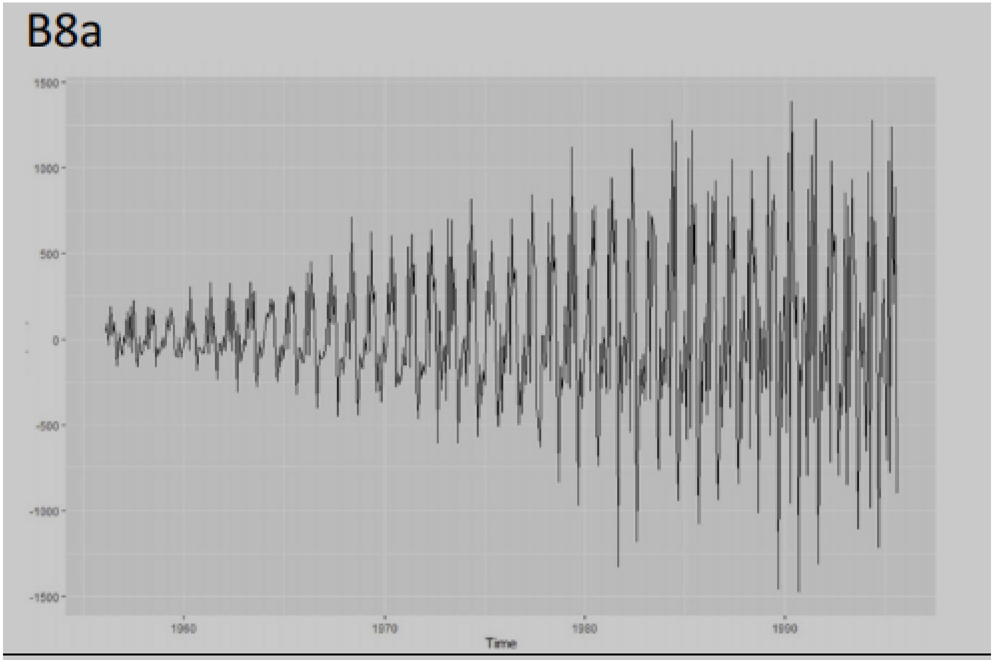
5.2. We are unable to reject hypothesis that there is more than 1 cointegrating relationship

We build a VAR model on the raw data.

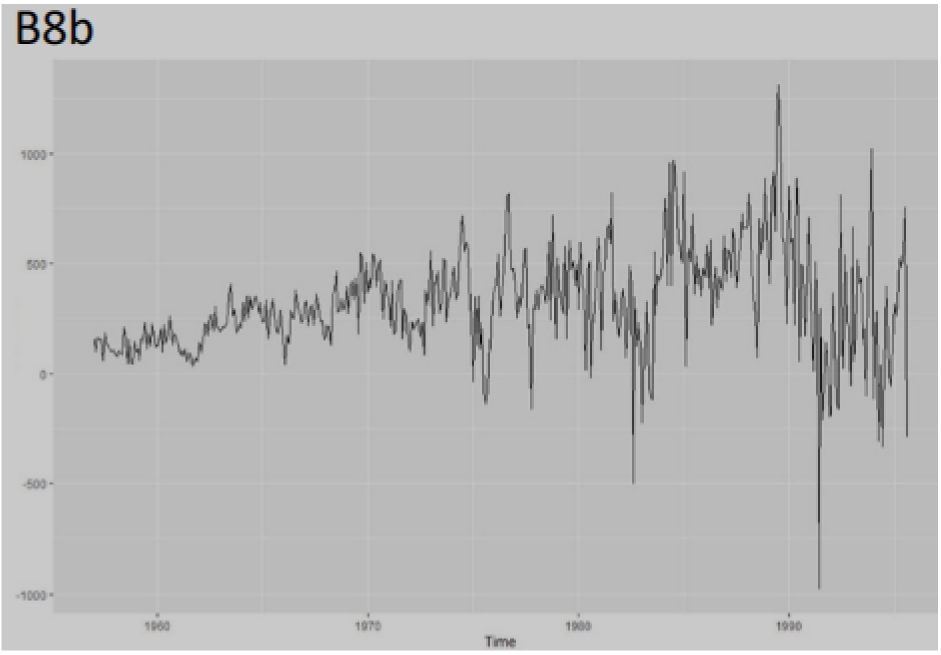
step1:Use `VARselect` and `SC(n)` to determine number of VAR lag.

step2:We will iterate on the selection of VAR(number of lag determined in step 1) until there is no longer any residual serial correlation.

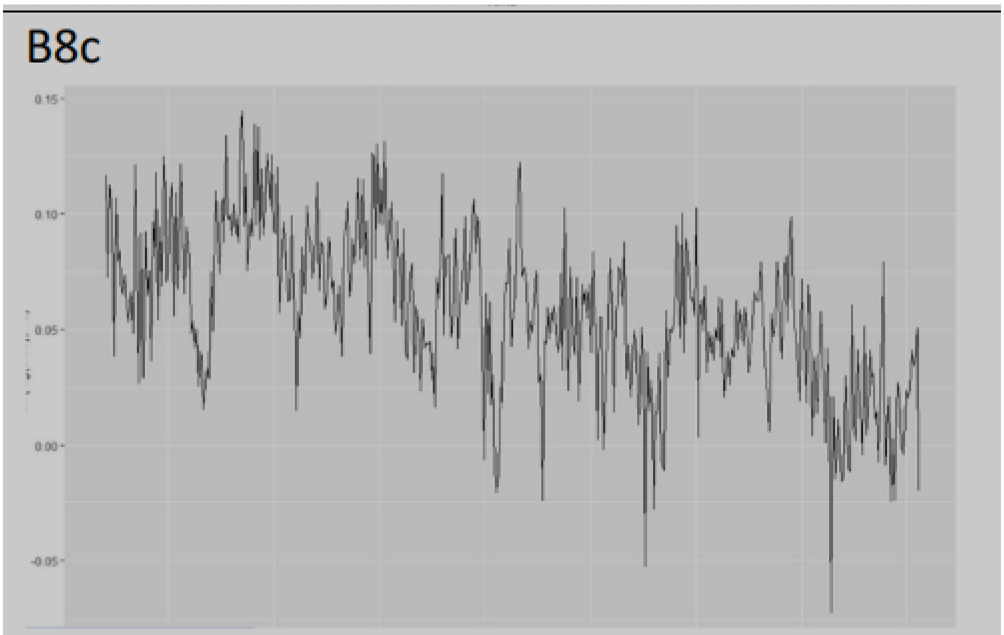
6.Discuss whether following time series are covariance stationary



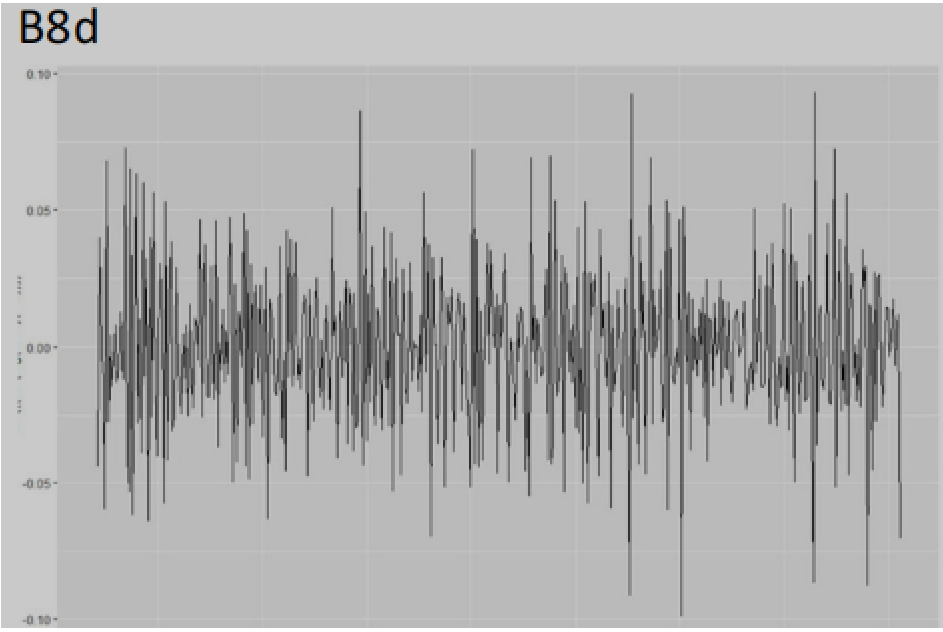
`non covariance stationary` :increasing variance



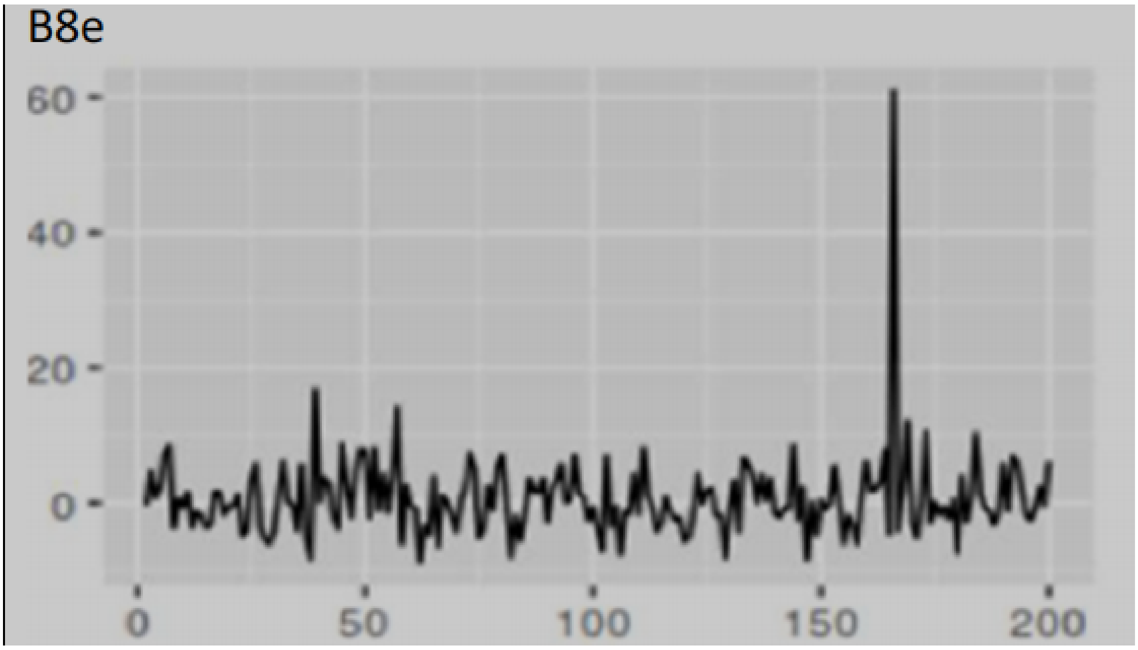
`non covariance stationary` :increasing variance



non covariance stationary :trend



covariance stationary




```
covariance stationary
```