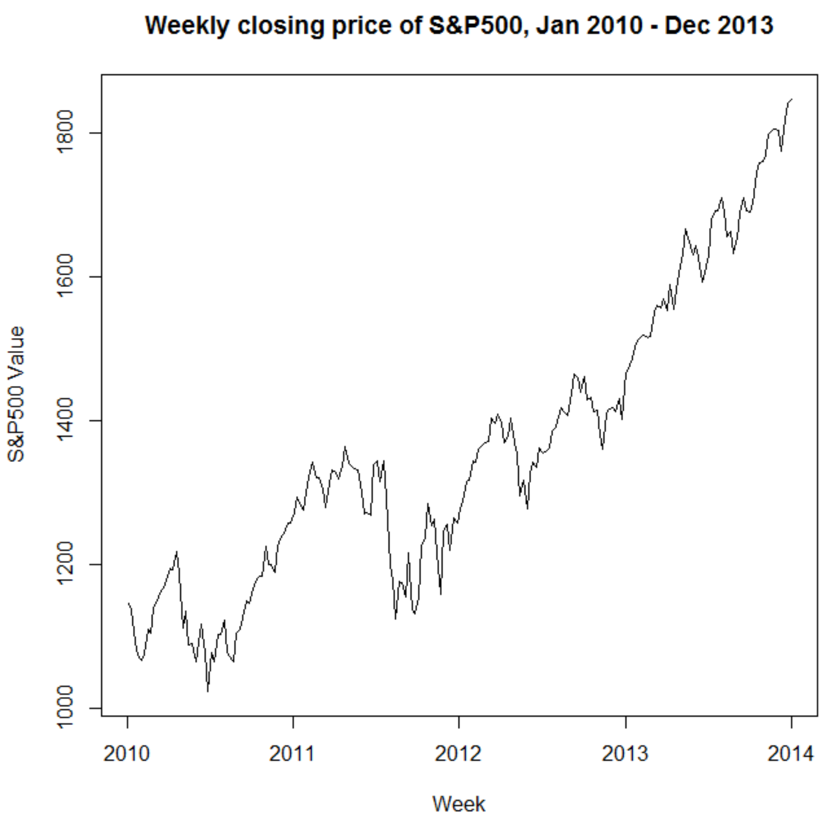
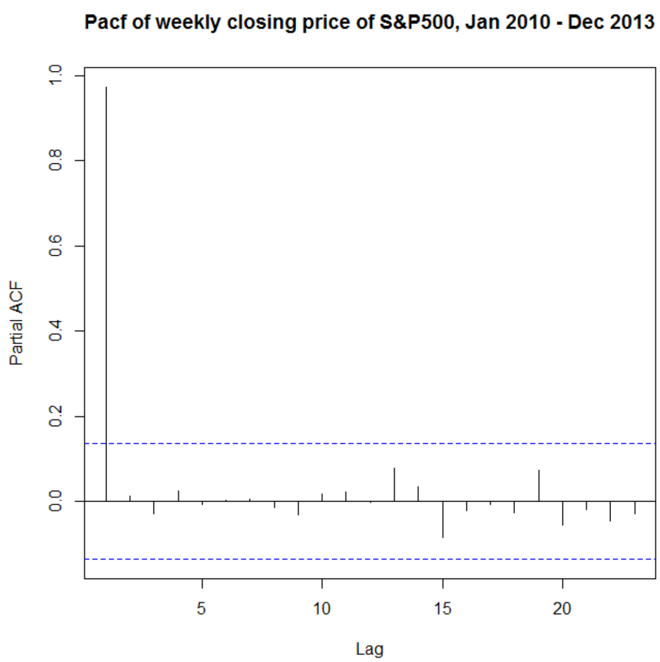
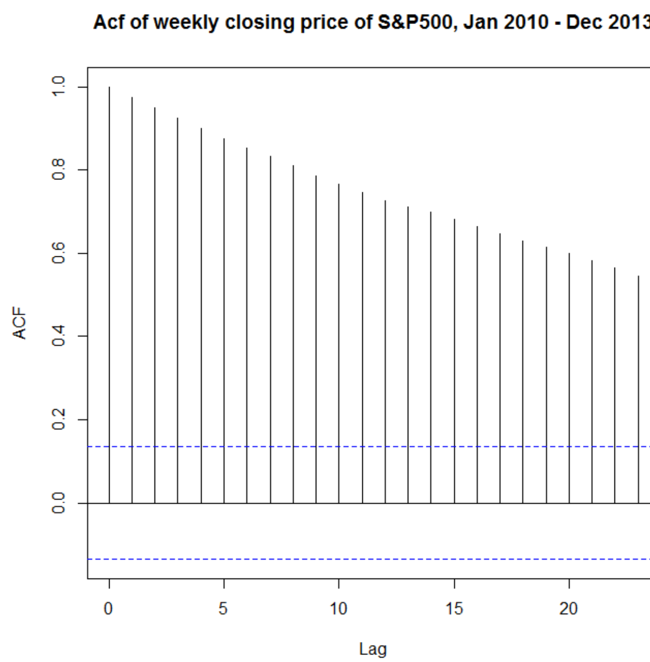
Activity Solution: Model Diagnostics

The Standard and Poor 500 index (usually abbreviated S&P500) is a stock market index based on the valuations of 500 companies listed on either the New York Stock Exchange or NASDAQ, the largest US stock exchange. In simple terms the S&P500 is a weighted sum of the unit stock values of the 500 companies selected. Many mutual funds attempt to track or replicate the S&P500, hence many investors worldwide are impacted by its growth. Here we attempt to model the S&P500. A plot of the weekly closing prices between January 2010 and December 2013 (inclusive) is given below, along with the corresponding acf and pacf plots. There are 209 values in the time series here.







1. Comment on the key features of series, informed by the above plots.
   The series appears to fluctuate considerably, and may exhibit an upward trend over the time period given. The acf decays slowly, which would be consistent with a non-stationary process, such as one with a trend. The mean is clearly not zero. The pacf appears to cut off after lag one.

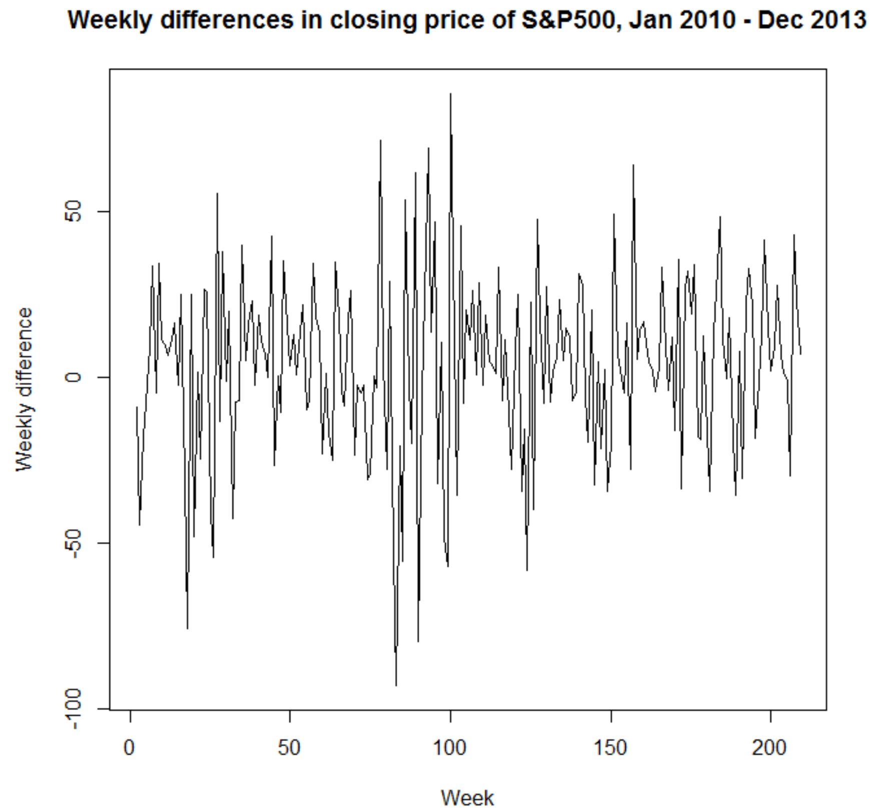
2. *Clickers question:* If you were to fit an ARMA model to the above series, which model would you try first?
   Although the acf decays very slowly, an $AR(1)$ with non-zero mean is just tenable given the behaviour of the pacf. The parameter α would be close to +1 though to exhibit a similar acf to that observed. However, fitting this model in R is not possible, since
   `AR1quoteSPw <- arima(quoteSPw, order = c(1, 0, 0), seasonal = list(order = c(0, 0, 0), method = c("CCS-ML", "ML"), period = NA), xreg = NULL, include.mean = TRUE)`
   returns an error of a non-stationary AR whichever method of model fitting is chosen.

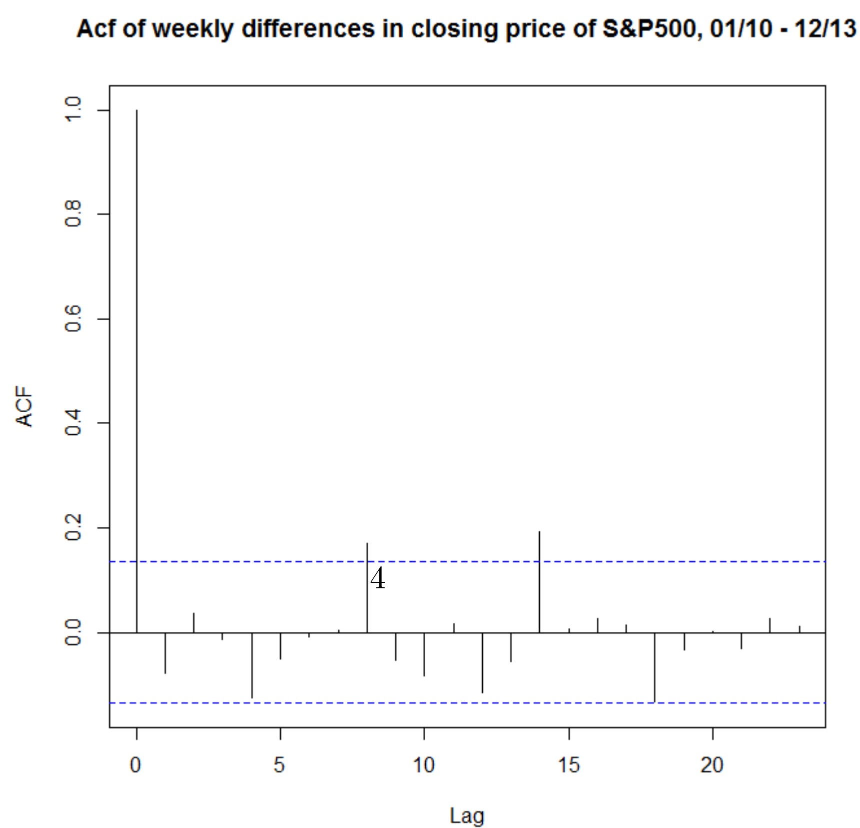
3. What objections could there be to the model you selected?

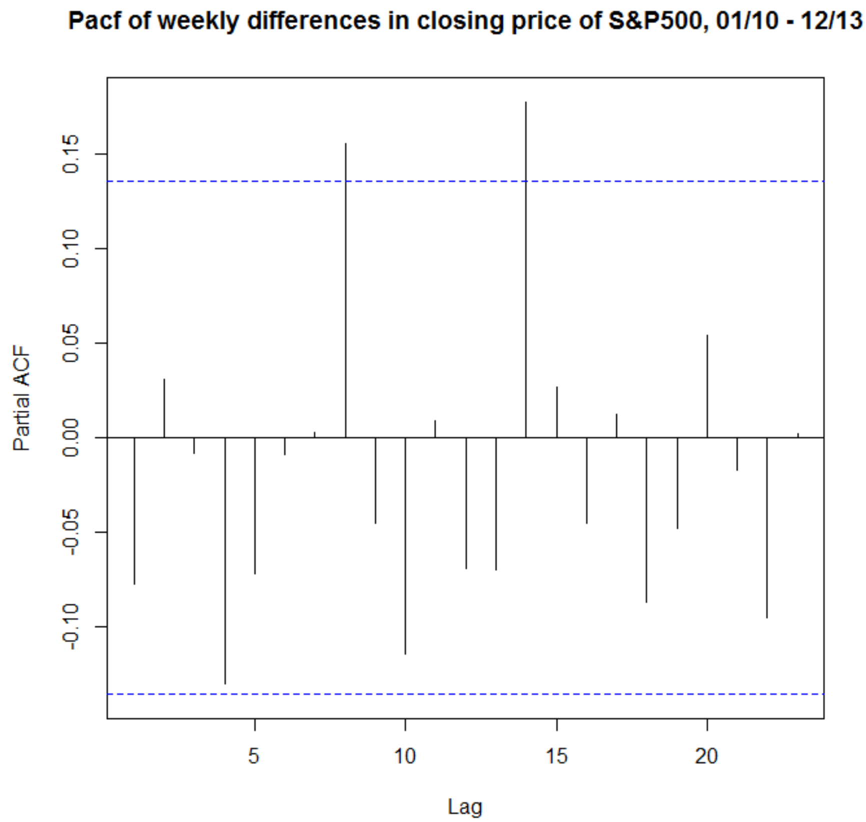
Aside from a stationary $AR(1)$ model not being possible to fit in R, one would ask what the model suggests about the behaviour of the S&P500. The model implies the index fluctuates about a fixed mean, which does not appear realistic in the context of a financial time series process.

4. Consider the differences between weekly closing prices, displayed below

along with their acf and pacf. Comment on the key features.







The differences exhibit no obvious trend or pattern, though do appear to become more volatile (i.e., higher variance) during the middle of the time period. There are only two significant acf values, at lags 8 and 14. It is difficult to explain these correlations in terms of the original series, so possibly they are just large values by chance and the underlying correlations are zero. A similar argument may apply to the pacf. Remember that even for pure noise, about one in twenty sample acf and pacf values will appear "significant". Seeing two such values out of about 25 is not highly unexpected, and given the lags at which they occur it is hard to explain them as being anything other than due to chance.

5. *Clickers question:* Which model would you consider reasonable to fit to the first differences of the original series? Write down your model

using suitable notation.
The model that treats first differences as white noise is the ARIMA$(0,1,0)$ *, which can be written*

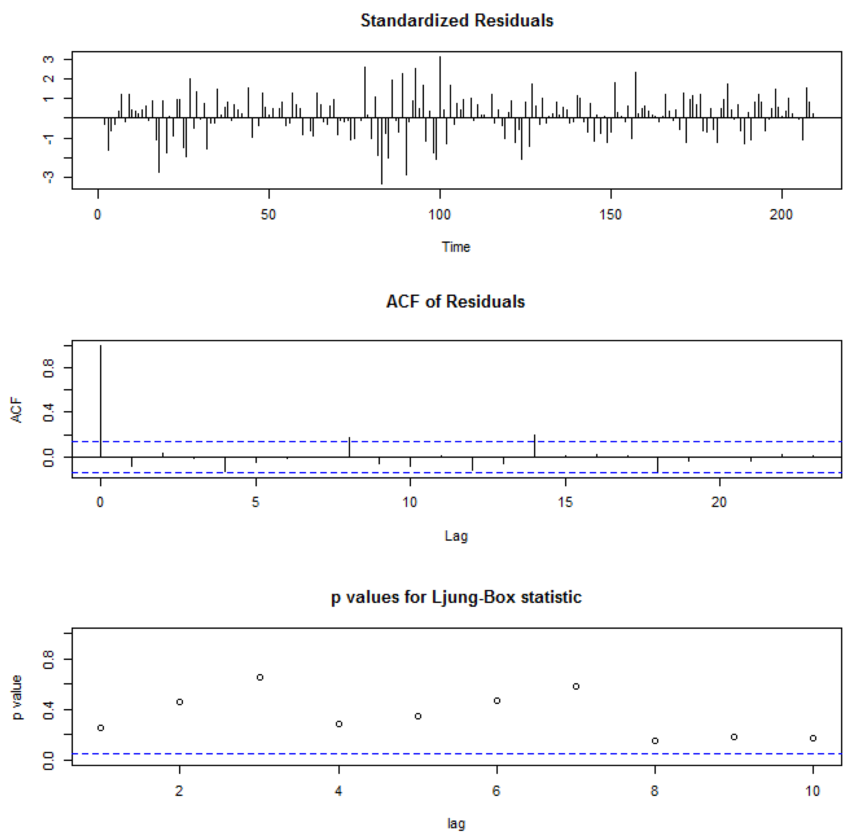
$$X(t) = X(t-1) + Z(t),$$

where $Z(t)$ *is white noise with variance* σ^2*, for some* σ*.*

6. Model diagnostics often involves consideration of the **residuals** of the model, i.e., the differences between the observed values and the corresponding fitted values. Let

$$\hat{z}(t) := x(t) - \hat{x}(t)$$

(for $t = 1, 2, \ldots, 208$) define the residuals of a model, where $\hat{x}(t)$ is the value fitted by the model at time t. Fitting the above model using R, the diagnostic plot is given below. Are the standardized residuals consistent with a good model fit? What about the acf of the residuals?



For a good model fit, the residuals should appear like pure noise, likely approximately independent standard Normal values. Only about one in a hundred should fall outside ± 3, *which looks a reasonable here on eye-balling the residual plot. Perhaps rather more lie outside* ± 2 *than might be expected. A potential problem is that outlier values are tending to cluster during weeks 80 to 100 , and so do not appear distributed at random. As with the series of differences, the acf exhibits significantly non-zero values at lags 8 and 14, though these may reasonably be ascribed to chance.*

7. *Clickers question:* Tests for goodness–of–fit exist based on the residuals. Let the acf at lag k of the residuals be denoted $r_k(\hat{z})$. One such test is the **Ljung–Box test**, which suggests test statistic

$$N(N+2)\sum_{k=1}^{M}\frac{r_k(\hat{z})^2}{(N-k)}$$

where N is the number of terms in the series and M is an integer lag less than N. The above should be consistent with χ^2_{M-p-q} if the ARMA(p,q) model we have fitted to the data is reasonable. If the model fitted is poor, the value of the statistic becomes inflated, and too large to be consistent with the χ^2_{M-p-q} distribution. The output above gives P-values of the test for certain choices of M. Given that the first three values of the acf of the residuals here are –0.077, 0.037, and –0.013, find the value of the test statistic when $M = 3$. If you have access to software, confirm the corresponding P-value.
The test statistic at $M = 3$ *would be*

$$208 \times 210\left(\frac{(-0.077)^2}{(208-1)}+\frac{(0.037)^2}{(208-2)}+\frac{(-0.013)^2}{(208-3)}\right) = 1.577\,.$$

Under the null hypothesis that the fitted model is adequate, the above value should be consistent with the χ^2_3. *The P-value is about 0.665. From the above plot of the P-values from various choices of* M, *there is nothing to suggest the model is inadequate based on residual correlations.*

8. What concerns would you have about the model fitted? In particular, how would you use the model for forecasting future weekly closing

values?
The model asserts that the weekly closing value of the S&P500 index is a nonstationary process, and in fact behaves like a random walk. This model implies that weekly price differences are purely random and therefore cannot be predicted in that they do not have dependence on past values. Hence our best forecast for next week's value would be today's value, which hardly seems an appealing approach to forecasting the underlying index.

9. Re-cap what you have done during this activity. Why do you think you were asked to do this activity? What have you learned?
We have examined a well–known time series, the S&P500, over a four–year period, looking at just the end–of–week closing prices. We observed that the series is likely to be nonstationary, and although an $AR(1)$ looked tenable, a stationary form of such a model could not be fitted in R. Taking first differences, the resulting series appeared similar to white noise, suggesting and $ARIMA(0,1,0)$ model, that is, a random walk model for the original data. We explored the fit diagnostics given by R, in particular a goodness–of–fit test based on the acf of the residuals. The test appeared to support the general fit of the model.
The main point to take from this activity is that model fitting for time series is rarely a straightforward procedure, and in many cases either more than one model looks viable or no simple model appears a good fit. A more sophisticated approach here would attempt to model the non-constant variance apparent in the weekly differences. Candidate models include so–called ARCH and GARCH models, which are not considered in this course.
Learning outcomes encountered in this activity include:

    (a) *Describe the issues related to fitting AR, MA and ARMA models, in particular the choice of the order and general approaches to parameter estimation.*

    (b) *Interpret results of model diagnostic tests based on the residuals of a fitted time series model.*