36-618 HW1

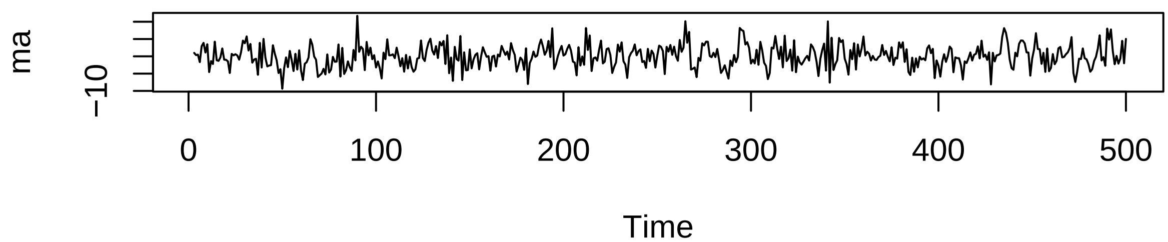
Daniel Nason

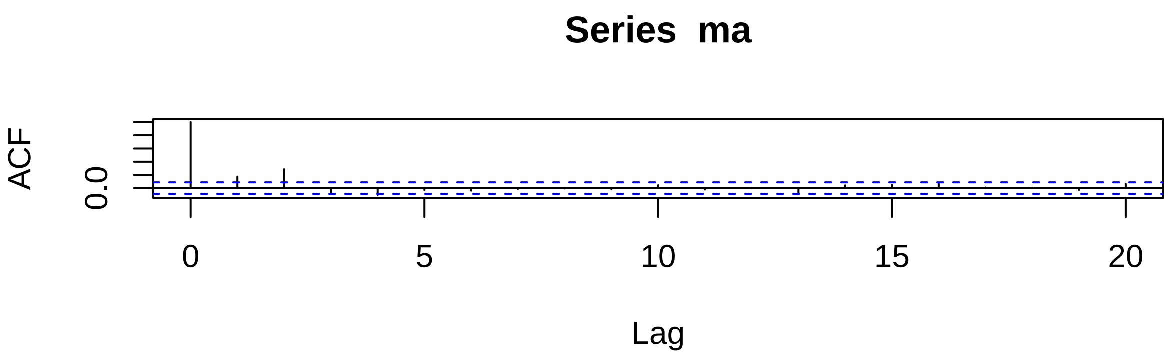
2/3/2022

Question 1

a)

```
par(mfrow = c(2,1))
ma <- stats::filter(ts(rnorm(500)), sides = 1, filter = c(1, 0.5, 3))
plot(ma)
acf(ma, na.action = na.omit, lag.max = 20)
```

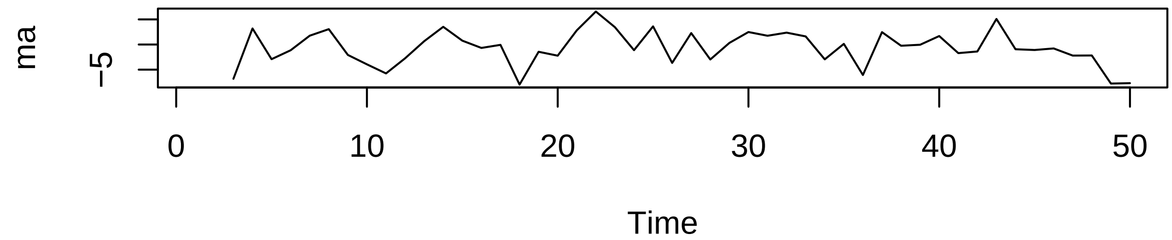


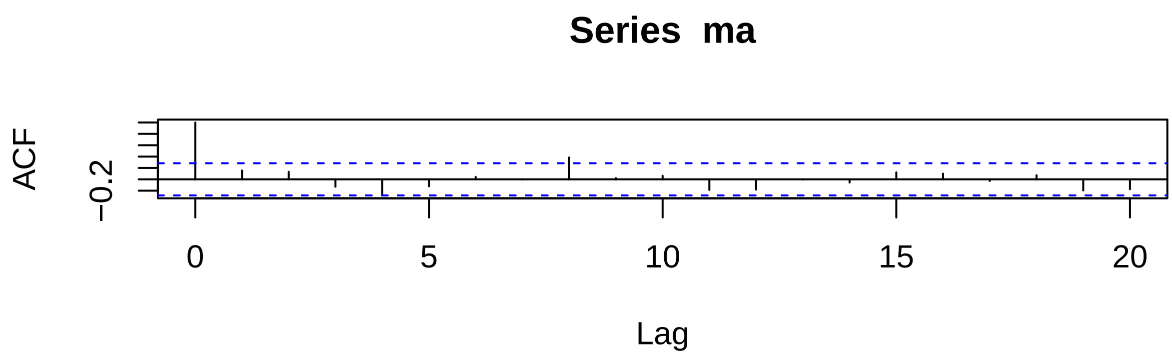



```
par(mfrow = c(1,1))
```

b)

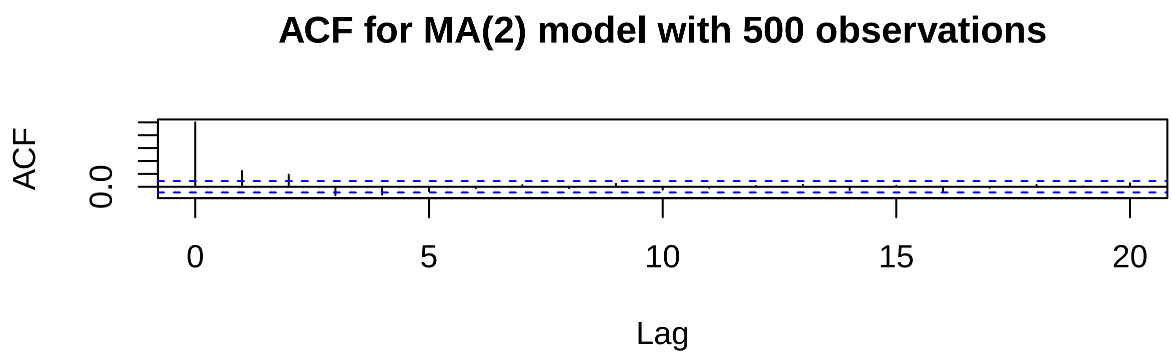
```
par(mfrow = c(2,1))
ma <- stats::filter(ts(rnorm(50)), sides = 1, filter = c(1, 0.5, 3))
plot(ma)
acf(ma, na.action = na.omit, lag.max = 20)
```

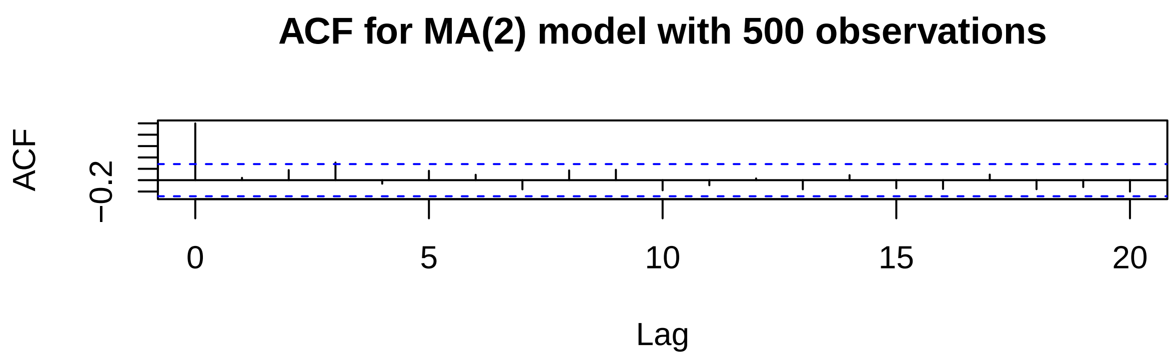





```
par(mfrow = c(1,1))
```

```
par(mfrow = c(2,1))
ma <- stats::filter(ts(rnorm(500)), sides = 1, filter = c(1, 0.5, 3))
acf(ma, na.action = na.omit, lag.max = 20,
    main = "ACF for MA(2) model with 500 observations")
ma <- stats::filter(ts(rnorm(50)), sides = 1, filter = c(1, 0.5, 3))
acf(ma, na.action = na.omit, lag.max = 20,
    main = "ACF for MA(2) model with 500 observations")
```





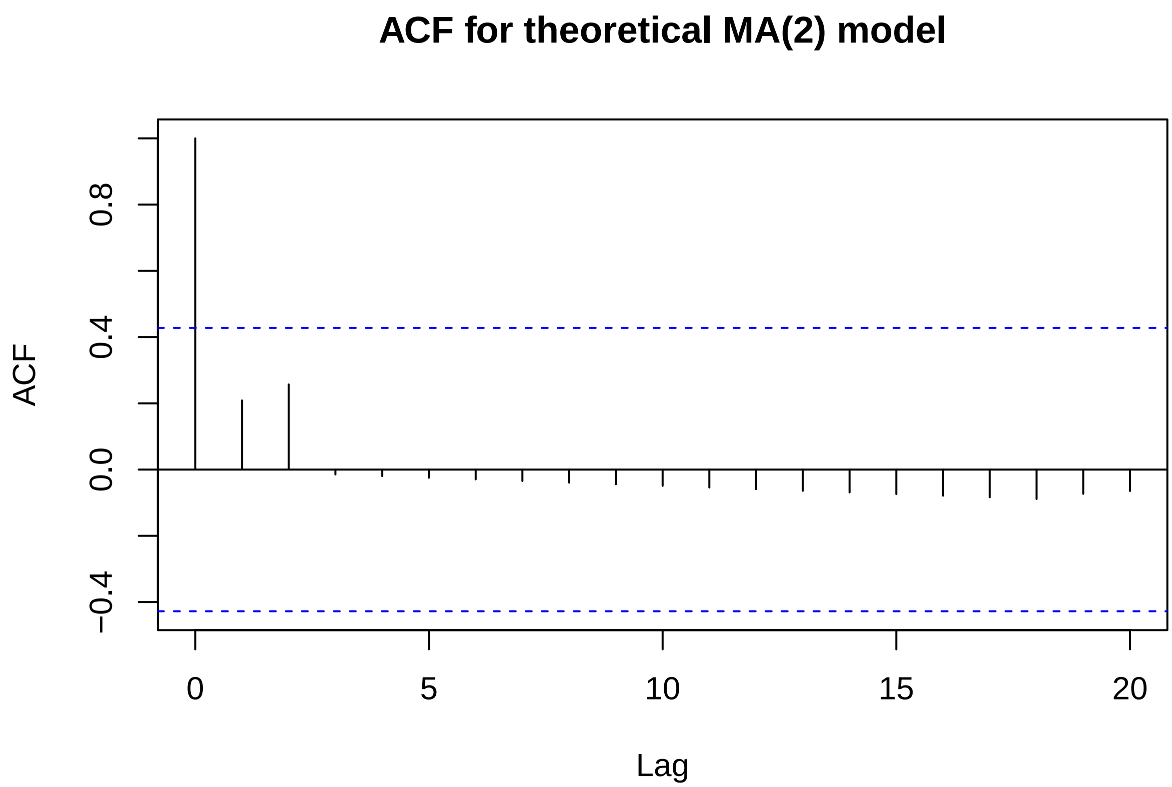

```r
par(mfrow = c(1,1))
```

We see from the ACF plot that the 95% error bars are further from 0 as we drop n from 500 to 50 since the square root of n is the denominator of error bars. The ACF values also have a smaller magnitude for larger values of n.

```r
ARMAacf(ma=c(0.5,3))
```

```
##         0         1         2         3
## 1.0000000 0.1951220 0.2926829 0.0000000
```

```r
acf(ARMAacf(ma=c(0.5,3), lag.max = 20),
    na.action = na.omit,
    lag.max = 20,
    main = "ACF for theoretical MA(2) model")
```



Comparing the theoretical ACF for the MA(2) model to the plots from a) and b), we see that the first two lags of the theoretical ACF positive and non-zero and the third value is 0. The ACFs of all the plots are relatively similar in that each plot has the first two ACFs with values that are outside of the 95% error bars (non-zero) and the remainder that lie within the bars closer to 0. We see that as we get more observations, the empirical ACF more closely resembles the theoretical ACF as plot a) more closely resembles the theoretical distribution than plot b).

Question 2

a)

Several indices monitor different regions of the tropical Pacific, with the Nino 3.4 index and Oceanic Nino Index (ONI) most commonly used to define El Nino and La Nina events. Nino 3.4 and ONI cover the same geographic region of the Pacific (5N-5S, 170W-120W), which is considered to represent the average equatorial SSTs (Sea Surface Temperature) across the Pacific from the dateline to the South American coast. Nino 3.4 uses a 5-month running mean while ONI uses a 3-month running mean; the NOAA classifies a full-fledged event (Nino/Nina) as having anomalies that exceed +/- 0.5 Celsius for at least 5 straight months.

b)

```
oni_raw <- read_delim("https://origin.cpc.ncep.noaa.gov/data/indices/oni.ascii.txt") %>%
  dplyr::select(3,4,6,8,9) %>%
  rename('SEAS' = ...3, 'TOTAL' = ...6, 'ANOM1' = ...8, 'ANOM2' = ...9) %>%
  mutate(ANOM = coalesce(ANOM1, ANOM2)) %>%
```

```
  select(-c(ANOM1, ANOM2))
# test_df <- data.frame(a = oni %>% select(ANOM1), b = oni %>% select(ANOM2))
# test_df[is.na(test_df)] <- 0
# test_df$comb <- test_df$ANOM1 + test_df$ANOM2
# identical(oni$ANOM, test_df$comb)
# head(oni)
# colSums(is.na(oni))
```

c)

```
oni <- oni_raw %>% select(ANOM) %>% ts(frequency = 12)
oni
```

```
##      Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
## 1  -1.53 -1.34 -1.16 -1.18 -1.07 -0.85 -0.54 -0.42 -0.39 -0.44 -0.60 -0.80
## 2  -0.82 -0.54 -0.17  0.18  0.36  0.58  0.70  0.89  0.99  1.15  1.04  0.81
## 3   0.53  0.37  0.34  0.29  0.20  0.00 -0.08  0.00  0.15  0.10  0.04  0.15
## 4   0.40  0.60  0.63  0.66  0.75  0.77  0.75  0.73  0.78  0.84  0.84  0.81
## 5   0.76  0.47 -0.05 -0.41 -0.54 -0.50 -0.64 -0.84 -0.90 -0.77 -0.73 -0.66
## 6  -0.68 -0.62 -0.69 -0.80 -0.79 -0.72 -0.68 -0.75 -1.09 -1.42 -1.67 -1.47
## 7  -1.11 -0.76 -0.63 -0.54 -0.52 -0.51 -0.57 -0.55 -0.46 -0.42 -0.43 -0.43
## 8  -0.25  0.06  0.41  0.72  0.92  1.11  1.25  1.32  1.33  1.39  1.53  1.74
## 9   1.81  1.66  1.27  0.93  0.74  0.64  0.57  0.43  0.39  0.44  0.50  0.61
## 10  0.61  0.62  0.52  0.33  0.20 -0.07 -0.18 -0.28 -0.09 -0.03  0.05 -0.04
## 11 -0.10 -0.10 -0.07  0.03  0.02  0.03  0.13  0.24  0.27  0.20  0.12  0.05
## 12  0.04  0.03  0.04  0.09  0.23  0.27  0.14 -0.13 -0.30 -0.26 -0.19 -0.16
## 13 -0.24 -0.22 -0.20 -0.26 -0.28 -0.20 -0.04 -0.07 -0.11 -0.22 -0.31 -0.43
## 14 -0.40 -0.15  0.15  0.27  0.31  0.52  0.86  1.14  1.22  1.29  1.37  1.31
## 15  1.07  0.62  0.12 -0.33 -0.58 -0.58 -0.60 -0.66 -0.76 -0.80 -0.82 -0.78
## 16 -0.59 -0.28 -0.07  0.18  0.46  0.83  1.22  1.54  1.85  1.98  1.97  1.72
## 17  1.37  1.17  0.98  0.66  0.35  0.24  0.24  0.12 -0.05 -0.10 -0.18 -0.30
## 18 -0.41 -0.48 -0.53 -0.45 -0.24  0.00  0.05 -0.16 -0.30 -0.38 -0.34 -0.44
## 19 -0.64 -0.74 -0.62 -0.44 -0.04  0.28  0.58  0.53  0.45  0.55  0.73  0.98
## 20  1.13  1.09  0.95  0.77  0.61  0.43  0.36  0.51  0.79  0.86  0.81  0.63
## 21  0.51  0.34  0.29  0.19  0.04 -0.30 -0.63 -0.76 -0.77 -0.74 -0.86 -1.15
## 22 -1.36 -1.38 -1.12 -0.85 -0.73 -0.74 -0.80 -0.77 -0.82 -0.85 -0.96 -0.90
## 23 -0.71 -0.35  0.06  0.41  0.67  0.92  1.13  1.37  1.58  1.84  2.09  2.12
## 24  1.84  1.25  0.54 -0.10 -0.54 -0.87 -1.11 -1.28 -1.45 -1.71 -1.95 -2.03
## 25 -1.84 -1.55 -1.23 -1.03 -0.91 -0.77 -0.53 -0.37 -0.41 -0.61 -0.75 -0.64
## 26 -0.54 -0.57 -0.65 -0.73 -0.83 -0.98 -1.13 -1.20 -1.37 -1.43 -1.55 -1.65
## 27 -1.56 -1.17 -0.73 -0.47 -0.28 -0.05  0.18  0.35  0.62  0.81  0.86  0.85
## 28  0.71  0.64  0.34  0.23  0.21  0.34  0.35  0.42  0.57  0.73  0.81  0.79
## 29  0.69  0.42  0.06 -0.18 -0.31 -0.29 -0.36 -0.42 -0.42 -0.29 -0.08  0.00
## 30  0.03  0.07  0.20  0.28  0.23  0.05  0.04  0.17  0.33  0.45  0.52  0.64
## 31  0.59  0.46  0.34  0.38  0.48  0.46  0.25  0.03 -0.07  0.02  0.11 -0.01
## 32 -0.26 -0.50 -0.47 -0.37 -0.26 -0.29 -0.30 -0.25 -0.16 -0.13 -0.15 -0.08
## 33 -0.05  0.07  0.19  0.47  0.66  0.72  0.79  1.07  1.58  1.97  2.18  2.23
## 34  2.18  1.92  1.54  1.29  1.06  0.72  0.31 -0.08 -0.46 -0.81 -1.00 -0.91
## 35 -0.60 -0.42 -0.34 -0.43 -0.51 -0.45 -0.30 -0.16 -0.24 -0.56 -0.92 -1.14
## 36 -1.04 -0.85 -0.77 -0.78 -0.78 -0.63 -0.49 -0.46 -0.40 -0.35 -0.27 -0.36
## 37 -0.49 -0.47 -0.31 -0.20 -0.12 -0.04  0.22  0.44  0.71  0.94  1.14  1.22
```

```
## 38  1.23  1.19  1.06  0.95  0.97  1.22  1.51  1.70  1.65  1.48  1.25  1.11
## 39  0.81  0.54  0.14 -0.31 -0.88 -1.30 -1.30 -1.11 -1.19 -1.48 -1.80 -1.85
## 40 -1.69 -1.43 -1.08 -0.83 -0.58 -0.40 -0.31 -0.27 -0.24 -0.22 -0.16 -0.05
## 41  0.14  0.21  0.28  0.29  0.29  0.31  0.33  0.38  0.39  0.35  0.40  0.41
## 42  0.41  0.26  0.22  0.26  0.45  0.64  0.73  0.64  0.62  0.79  1.21  1.53
## 43  1.71  1.63  1.48  1.29  1.06  0.73  0.37  0.09 -0.13 -0.25 -0.28 -0.13
## 44  0.09  0.30  0.50  0.67  0.70  0.57  0.32  0.25  0.15  0.10  0.04  0.06
## 45  0.06  0.07  0.17  0.31  0.42  0.41  0.44  0.43  0.55  0.74  1.01  1.09
## 46  0.96  0.72  0.53  0.30  0.14 -0.03 -0.24 -0.54 -0.81 -0.97 -1.00 -0.98
## 47 -0.90 -0.75 -0.59 -0.39 -0.31 -0.30 -0.27 -0.32 -0.35 -0.40 -0.45 -0.49
## 48 -0.50 -0.36 -0.10  0.28  0.75  1.22  1.60  1.90  2.14  2.33  2.40  2.39
## 49  2.24  1.93  1.44  0.99  0.45 -0.13 -0.78 -1.12 -1.31 -1.35 -1.48 -1.57
## 50 -1.55 -1.30 -1.07 -0.98 -1.02 -1.04 -1.10 -1.11 -1.16 -1.26 -1.46 -1.65
## 51 -1.66 -1.41 -1.07 -0.81 -0.71 -0.64 -0.55 -0.51 -0.55 -0.63 -0.75 -0.74
## 52 -0.68 -0.52 -0.44 -0.34 -0.25 -0.12 -0.08 -0.13 -0.19 -0.29 -0.35 -0.31
## 53 -0.15  0.03  0.09  0.20  0.43  0.65  0.79  0.86  1.01  1.21  1.31  1.14
## 54  0.92  0.63  0.38 -0.04 -0.26 -0.16  0.08  0.21  0.26  0.29  0.35  0.35
## 55  0.37  0.31  0.23  0.17  0.17  0.28  0.47  0.64  0.70  0.67  0.66  0.69
## 56  0.64  0.58  0.45  0.43  0.29  0.11 -0.06 -0.14 -0.11 -0.29 -0.57 -0.84
## 57 -0.85 -0.77 -0.57 -0.37 -0.14 -0.03  0.10  0.30  0.54  0.77  0.94  0.94
## 58  0.66  0.22 -0.12 -0.32 -0.38 -0.47 -0.56 -0.81 -1.07 -1.34 -1.50 -1.60
## 59 -1.64 -1.52 -1.29 -1.01 -0.84 -0.61 -0.37 -0.23 -0.24 -0.35 -0.55 -0.73
## 60 -0.85 -0.79 -0.61 -0.33  0.01  0.28  0.45  0.58  0.71  1.01  1.36  1.56
## 61  1.50  1.22  0.84  0.35 -0.17 -0.66 -1.05 -1.35 -1.56 -1.64 -1.64 -1.59
## 62 -1.42 -1.19 -0.93 -0.73 -0.55 -0.44 -0.48 -0.62 -0.83 -1.01 -1.09 -1.04
## 63 -0.86 -0.72 -0.59 -0.47 -0.26 -0.01  0.25  0.37  0.37  0.27  0.05 -0.21
## 64 -0.43 -0.43 -0.34 -0.30 -0.36 -0.41 -0.40 -0.32 -0.26 -0.18 -0.17 -0.27
## 65 -0.42 -0.46 -0.27  0.04  0.21  0.16  0.05  0.07  0.23  0.49  0.64  0.66
## 66  0.55  0.47  0.53  0.70  0.93  1.18  1.52  1.86  2.16  2.42  2.57  2.64
## 67  2.48  2.14  1.58  0.94  0.39 -0.07 -0.36 -0.54 -0.63 -0.69 -0.67 -0.56
## 68 -0.34 -0.16  0.05  0.20  0.30  0.31  0.14 -0.11 -0.38 -0.65 -0.84 -0.97
## 69 -0.92 -0.85 -0.70 -0.50 -0.22 -0.01  0.09  0.23  0.49  0.76  0.90  0.81
## 70  0.75  0.72  0.71  0.66  0.54  0.45  0.28  0.14  0.19  0.35  0.51  0.55
## 71  0.50  0.48  0.40  0.19 -0.08 -0.30 -0.41 -0.57 -0.89 -1.17 -1.27 -1.19
## 72 -1.05 -0.93 -0.84 -0.66 -0.48 -0.38 -0.40 -0.49 -0.68 -0.81 -0.98 -0.99
```

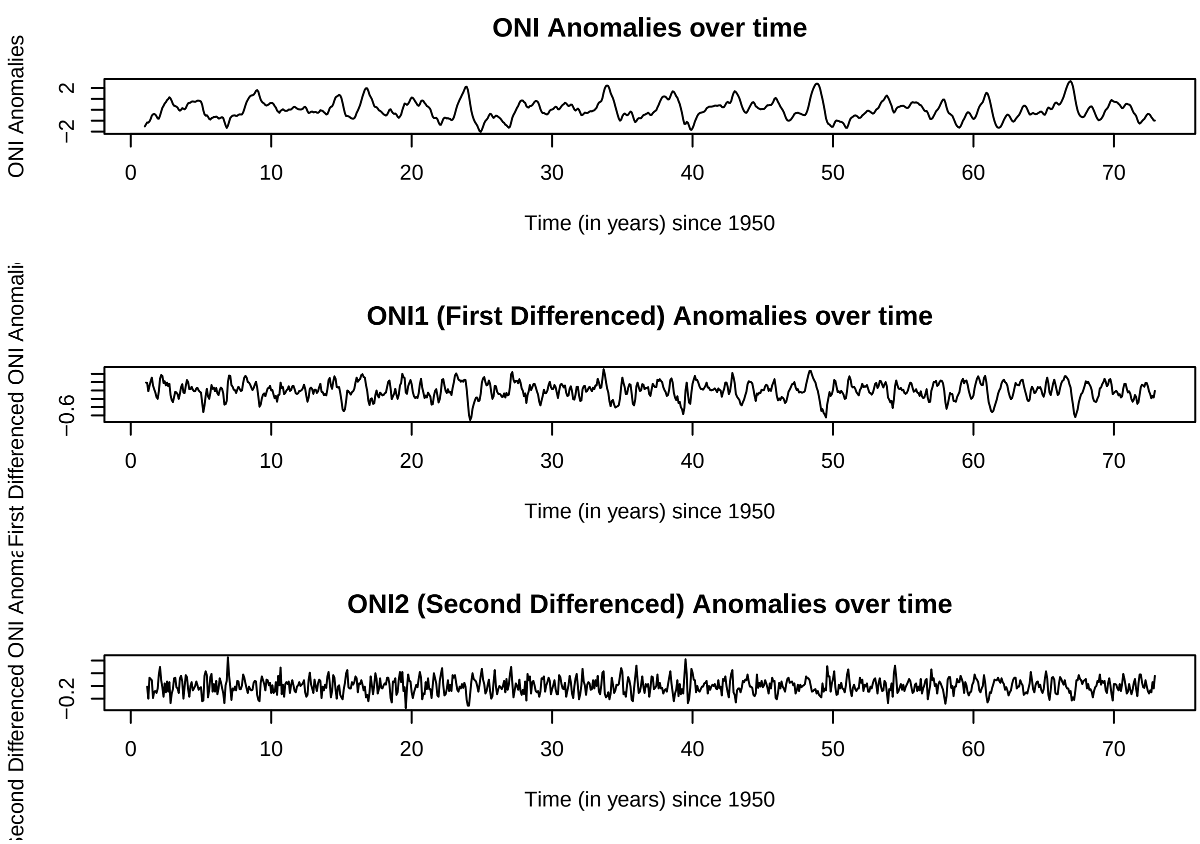
d)

```
oni1 <- diff(oni)
oni2 <- diff(oni1)
oni_combined <- cbind(oni, oni1, oni2)
oni_combined %>% head()
```

```
##        oni  oni1  oni2
## [1,] -1.53    NA    NA
## [2,] -1.34  0.19    NA
## [3,] -1.16  0.18 -0.01
## [4,] -1.18 -0.02 -0.20
## [5,] -1.07  0.11  0.13
## [6,] -0.85  0.22  0.11
```

e)

```r
par(mfrow=c(3,1))
plot(oni_combined[,1],
     main = "ONI Anomalies over time",
     ylab = "ONI Anomalies",
     xlab = "Time (in years) since 1950")
plot(oni_combined[,2],
     main = "ONI1 (First Differenced) Anomalies over time",
     ylab = "First Differenced ONI Anomalies",
     xlab = "Time (in years) since 1950")
plot(oni_combined[,3],
     main = "ONI2 (Second Differenced) Anomalies over time",
     ylab = "Second Differenced ONI Anomalies",
     xlab = "Time (in years) since 1950")
```




```r
par(mfrow=c(1,1))
```

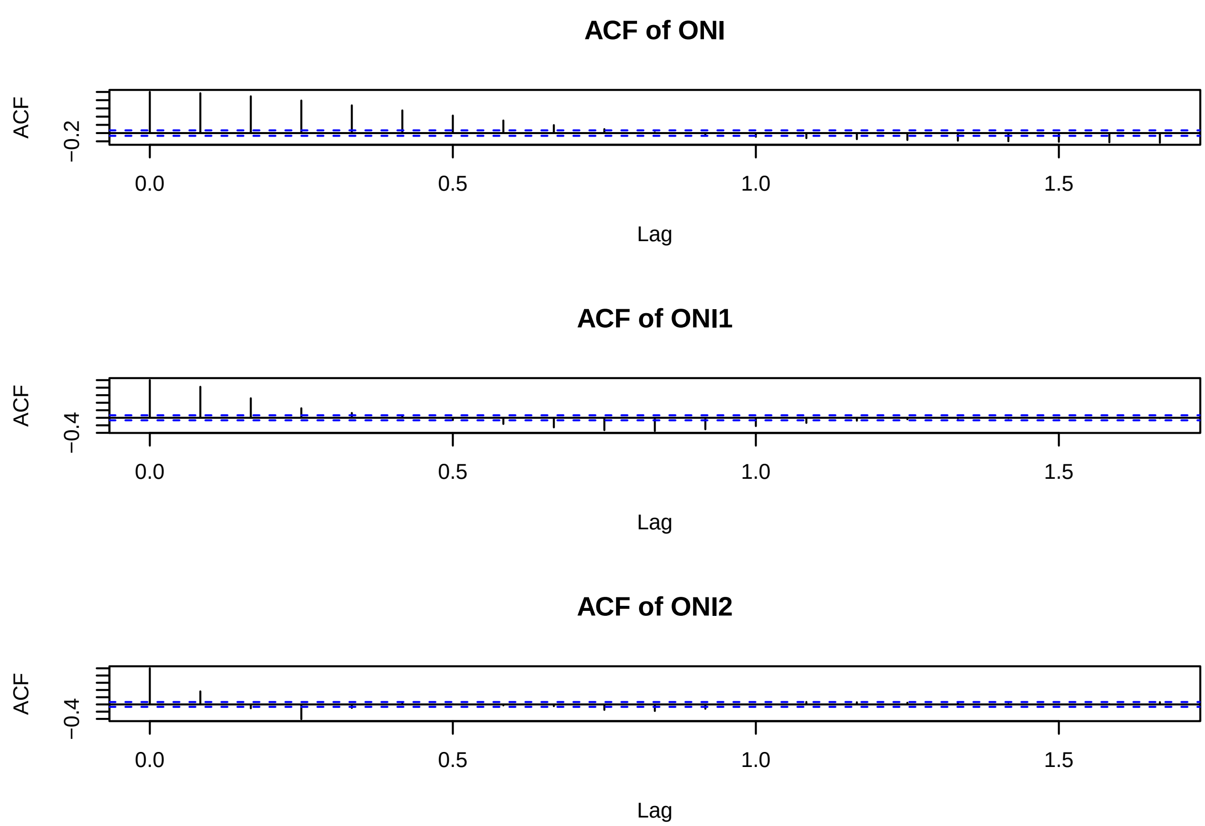
The plot for ONI suggests that there is evidence of seasonality since the ONI anomalies seem to fluctuate in regular intervals depending on the month of the year. There is also some evidence of heteroscedasticity in the data that aligns periodic fluctuations of the year. However, there is no trend in the data, as the mean seems to be roughly centered around 0. The periodicity and heteroscedasticity suggest that the ONI data are not stationary.

The plot for ONI1 shows no evidence of trend and periodicity. However, there does not appear to be constant variance (heteroscedasticity), which suggests that the ONI1 data are not stationary.

The plot for ONI2 shows no evidence of trend, relatively constant variance, and no periodicity. This implies that the ONI2 data are roughly stationary.

f)

```
par(mfrow=c(3,1))
acf(oni_combined[,1],
    na.action = na.omit,
    lag.max = 20,
    main = 'ACF of ONI')
acf(oni_combined[,2],
    na.action = na.omit,
    lag.max = 20,
    main = 'ACF of ONI1')
acf(oni_combined[,3],
    na.action = na.omit,
    lag.max = 20,
    main = 'ACF of ONI2')
```




```
par(mfrow=c(1,1))
```

The ACF plot for ONI illustrates that there is periodicity present in the data because of the periodic trends in the ACFs depending on the lag values. This also illustrates that autocorrelation is present in the data since the ACF values are outside the 95% error bar lines for the majority of the first 20 lags.

The ACF plot for ONI1 also illustrates that periodicity and autocorrelation are present in the data, but to a lesser extent than the ONI data. The ACF values are outside the bars for approximately half the lags and they periodically decay faster to 0 in the first 20 lags.

The ACF plot for ONI2 illustrates that there is much less periodicity than the first two plots, however autocorrelation is still noticeably present in the first and third lags since the ACF values lie outside the error bars.

None of these time series should be regarded as white noise, since there is clearly autocorrelation present, although the autocorrelation is reduced as the data is differenced. In each of the plots, there are at least 2 ACF lag values that are noticeably larger than the 95% error bars, and there is a pattern between sequential lag values. In a white noise ACF, we would expect approximately 1-2 ACF values to lie only slightly outside of the error bars and there would be no periodic trend in the direction of the ACF values.

36-618 HW1 Q3

a) $\{X_t\}$ is a stationary series with $\mu = 0$ & acf $\rho(\ell)$ for $t+\ell$, $\ell > 0$.

Show: $MSE(a) = E((X_{t+\ell} - aX_t)^2)$ st $a^* = \rho(\ell)$

Note: $Var(X) = E(X^2) - (E(X))^2$, $Cov(X,Y) = E(XY) - E(X)E(Y)$,

$\gamma(\ell+t, t) = \gamma(\ell) = Cov(X_{t+\ell}, X_t)$, $Var(X_t) = Cov(X_t, X_t) = \gamma(t,t) = \gamma(0)$

$\rho(\ell) = \gamma(\ell)/\gamma(0)$

$\Rightarrow MSE(a) = E((X_{t+\ell} - aX_t)^2) = E(X_{t+\ell}^2 - 2a(X_{t+\ell}X_t) + a^2 X_t^2)$

$= E(X_{t+\ell}^2) - 2a\,E(X_{t+\ell}X_t) + a^2 E(X_t^2)$

$= (Var(X_{t+\ell}) + (E(X_{t+\ell}))^2) - 2a(Cov(X_{t+\ell}, X_t) + E(X_{t+\ell})(X_t))$

$+ a^2(Var(X_t) + (E(X_t))^2$

$= (Var X_{t+\ell} + 0^2) - 2a(Cov(X_{t+\ell}, X_t) + 0 \cdot 0) + a^2(Var(X_t) + 0^2)$

$= Var(X_{t+\ell}) - 2a(Cov(X_{t+\ell}, X_t) + a^2(Var(X_t))) \Rightarrow \frac{\partial}{\partial a} = 0$

$\frac{\partial}{\partial a} \Rightarrow 0 - 2\,Cov(X_{t+\ell}, X_t) + 2a\,Var(X_t) = 0$

$\Rightarrow \cancel{2}a\,Var(X_t) = \cancel{2}Cov(X_{t+\ell}, X_t) \Rightarrow a = \dfrac{Cov(X_{t+\ell}, X_t)}{Var(X_t)}$

$\Rightarrow a = \dfrac{\gamma(t+\ell, t)}{\gamma(t,t)} = \dfrac{\gamma(\ell)}{\gamma(0)} = \rho(\ell) \Rightarrow \boxed{a^* = \rho(\ell)}$ minimize MSE

b) Show: $MSE(a^*) = \sigma^2(1 - \rho(\ell)^2)$, $Var(X_t) = \sigma^2$

$\Rightarrow E((X_{t+\ell} - a^* X_t)^2) = E(X_{t+\ell}^2) - 2a^* E(X_{t+\ell}X_t) + a^{*2} E(X_t^2)$

$= Var(X_{t+\ell}) - E(X_{t+\ell})^2 - 2a^*(Cov(X_{t+\ell}, X_t) + E(X_{t+\ell})E(X_t))$

$+ a^{*2}(Var(X_t) + (E(X_t))^2)$

$= Var(X_{t+\ell}) - 2a^*\,Cov(X_{t+\ell}, X_t) + a^{*2}\,Var(X_t)$

$= \sigma^2 - 2a^*\gamma(\ell) + a^*\sigma^2 = \sigma^2(1 - 2a^*\gamma(\ell)/\sigma^2 + a^*)$

$= \sigma^2\left(1 - 2a^*\dfrac{\gamma(\ell)}{\gamma(0)} + a^{*2}\right) = \sigma^2(1 - 2a^* a^* + a^{*2})$

$= \sigma^2(1 - a^{*2}) = \boxed{\sigma^2(1 - \rho(\ell)^2)}$, $MSE(a^*)$