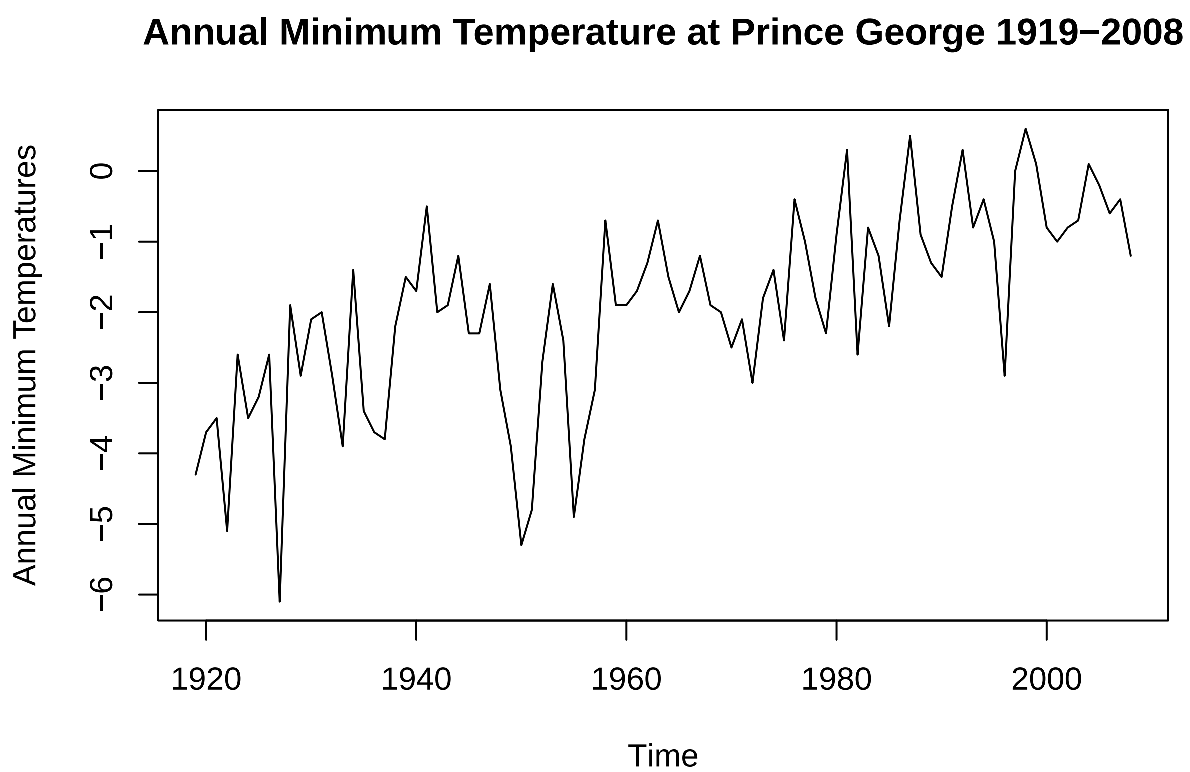
STAT 443: Lab 6

Wenxuan Zan (61336194)

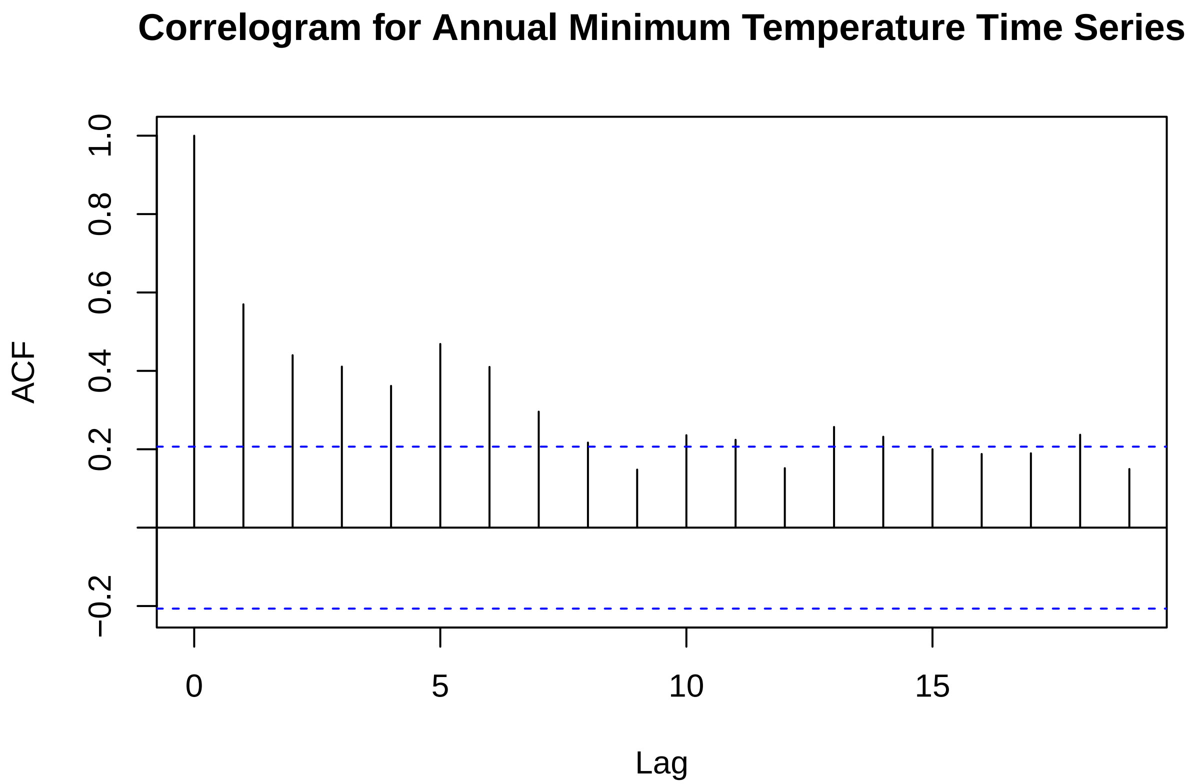
20 Feb, 2023

Question 1

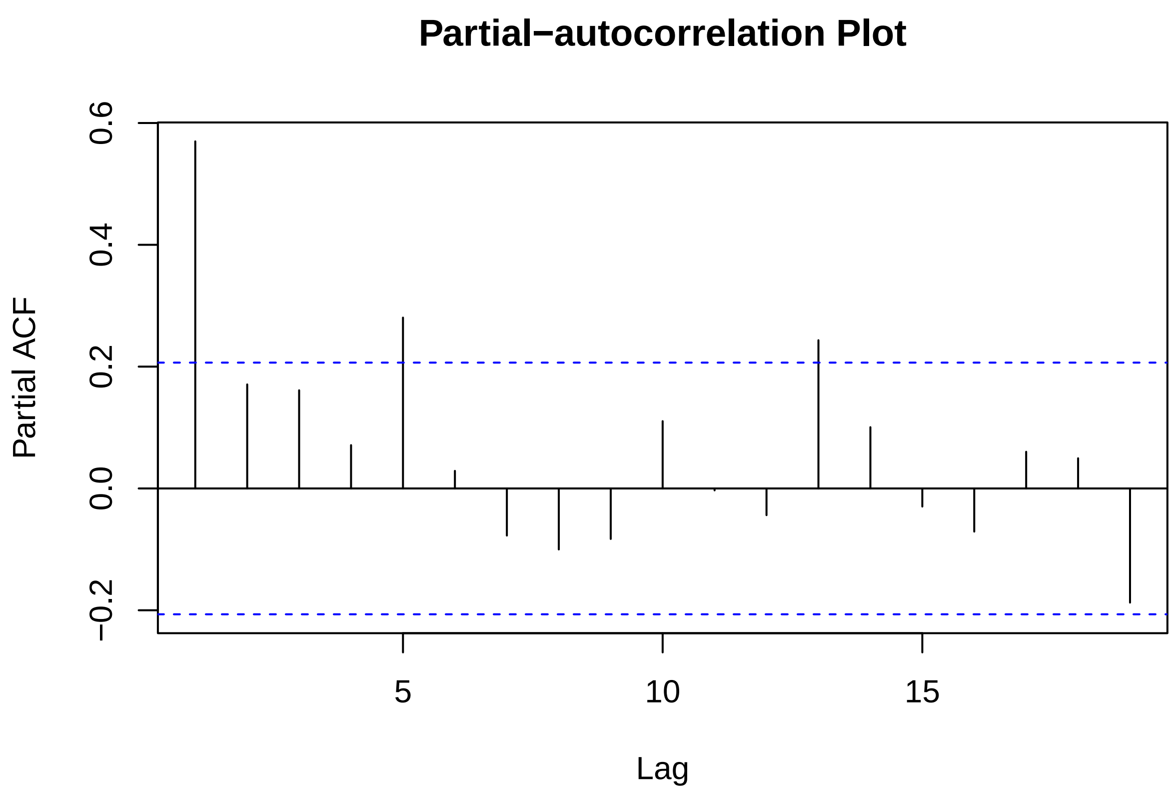
```
data <- read.csv("TempPG.csv",header = TRUE)
annual_ts <- ts(data$Annual, start = c(1919), end = c(2008))
plot(annual_ts,
     ylab = "Annual Minimum Temperatures",
     main = "Annual Minimum Temperature at Prince George 1919-2008")
```




```
acf(annual_ts,
    main = "Correlogram for Annual Minimum Temperature Time Series")
```




```r
pacf(annual_ts,
     main = "Partial-autocorrelation Plot")
```



i) Looking at the plot titled "Annual Minimum Temperature at Prince George 1919 to 2008", we can see a clear upward trend such that the annual minimum temperature at Prince George from 1919 to 2008 is increasing. Therefore this time series is likely to be non-stationary.

ii) Looking at the acf plot we observed a slow exponential decay of acf values. This pattern indicates the existence of positive temporal dependence within the data.This pattern is characteristic of a AR model.

iii) Looking at the pacf plot, a sensible cut-off is at lag 1 which suggests a AR(1) model, but the pacf value also spikes at lag 5, and 13, so perhaps a ARMA model is more suitable for the data.

iv) I would suggest a ARMA(1,0) model for the data.

Question 2

```
model <- arima(annual_ts, order = c(1,0,0), include.mean = TRUE)
model
```

```
##
## Call:
## arima(x = annual_ts, order = c(1, 0, 0), include.mean = TRUE)
##
## Coefficients:
##          ar1  intercept
##       0.5843    -1.9591
## s.e.  0.0864     0.2810
##
## sigma^2 estimated as 1.265:  log likelihood = -138.49,  aic = 282.99
```

The fitted model is:

$$X_t + 1.9591 = 0.5843(X_{t-1} + 1.9591) + Z_t; \qquad Z_t \sim WN(0, 1.265)$$

Question 3

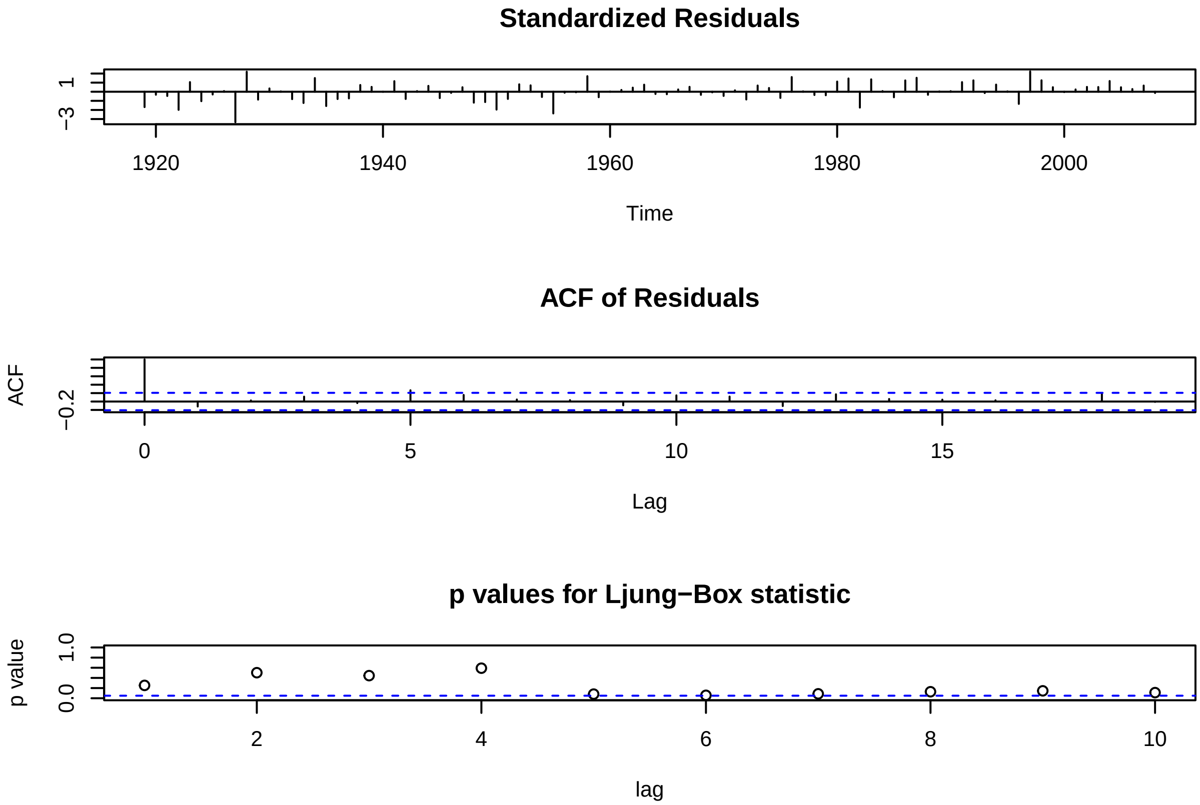
```
confint(model)
```

```
##                2.5 %    97.5 %
## ar1        0.4150038  0.753554
## intercept -2.5098255 -1.408472
```

The 95% CI for alpha is $[0.415, 0.754]$ and the 95% CI for μ is $[-2.510, -1.408]$

Question 4

```
tsdiag(model)
```



i) Most of the residuals are within the ± 3 bound.

ii) Most acf of residuals are within the $\pm 2/\sqrt{n}$ bound.

iii) p-values are non-significant for lag smaller than 5, and appears to be close to significant level for lag greater than 5.

iv) Overall, the model seems to fit the data relatively well.