STAT 443: Lab 6

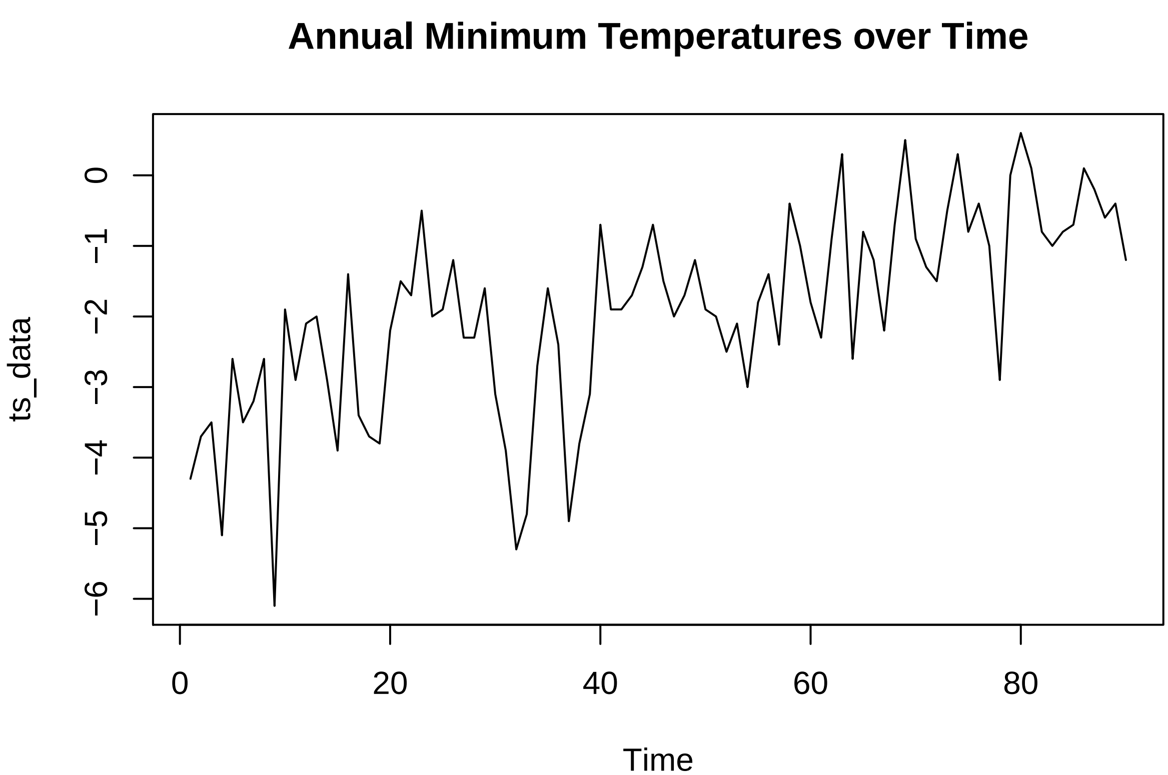
Flora Zhang 52135365

28 Feb, 2022

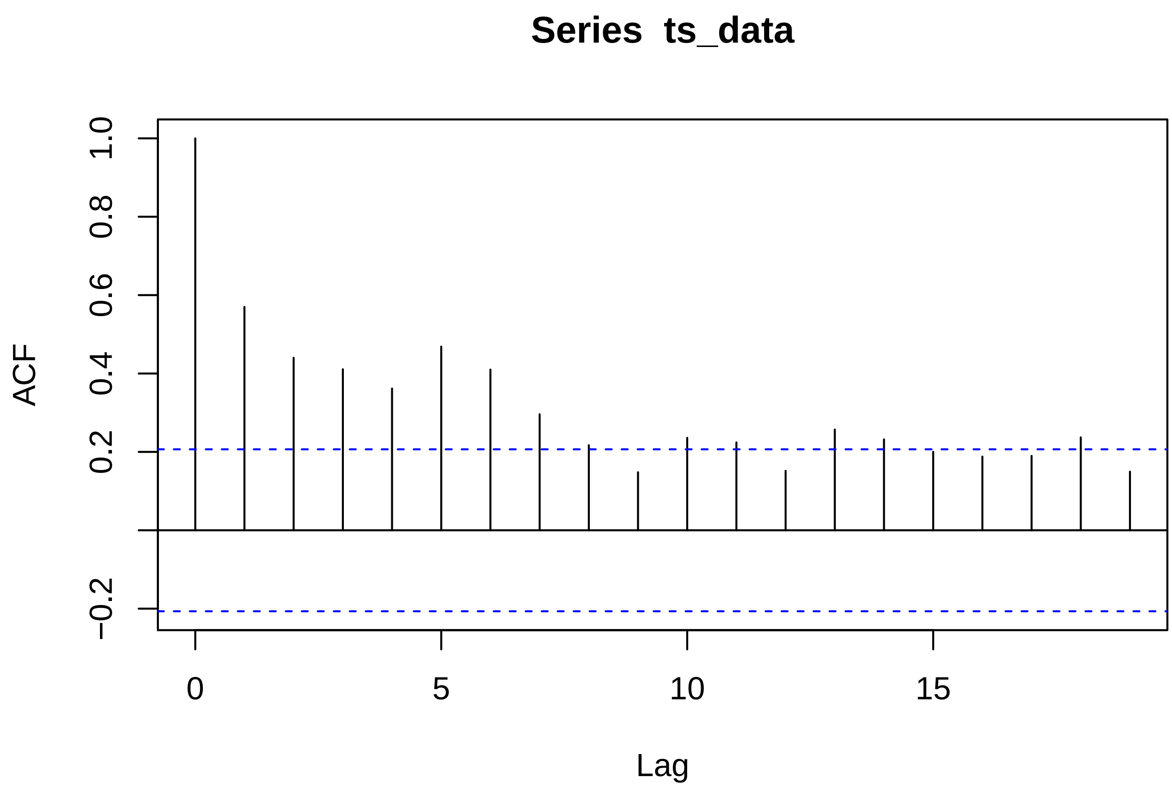
```
data = read.csv("TempPG.csv")
```

1. The column labelled "Annual" includes the annual minimum temperatures. Extract those data, and coerce them into a time series object. Plot the time series, its acf and pacf. Comment on what you observe. If you were to fit an ARMA model to the above data, which would you select?

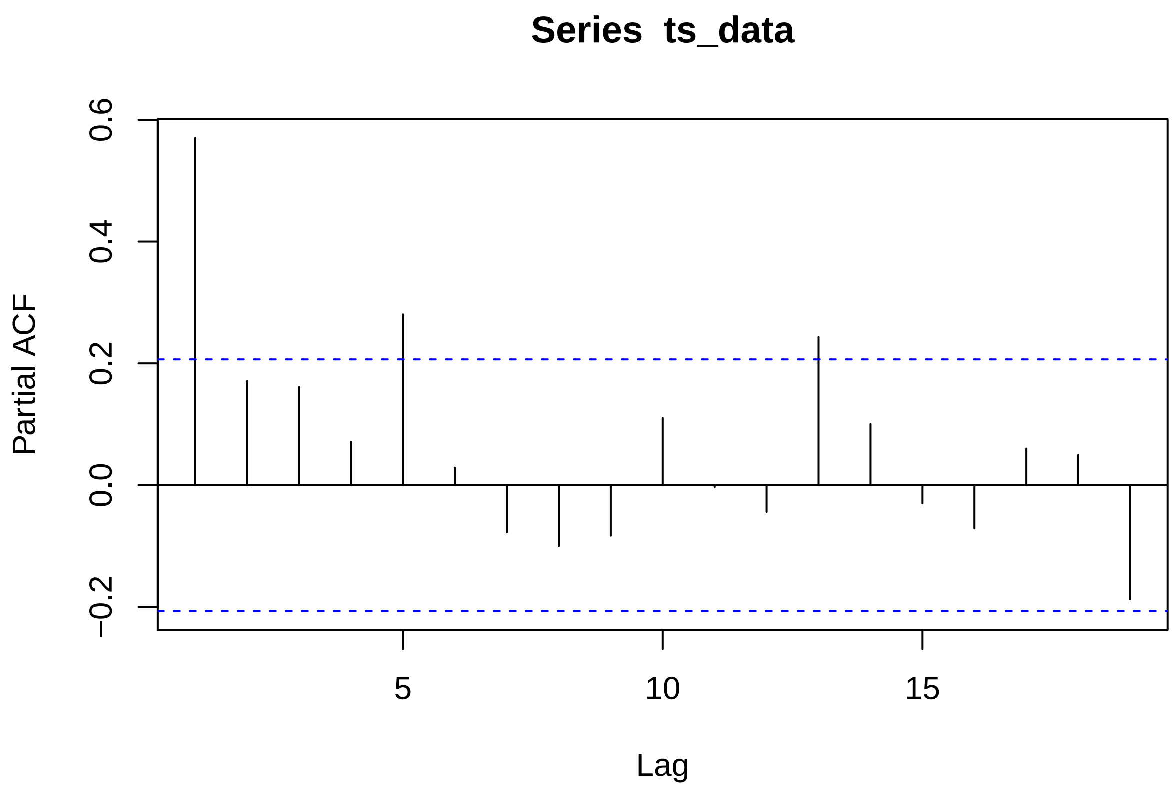
```
ts_data = ts(data=data$Annual)
#plot time series
plot(ts_data, main="Annual Minimum Temperatures over Time")
```




```
acf(ts_data)
```




```
pacf(ts_data)
```



From the time series plot, we see an upwards long-run trend over time. The acf plot appears to be tapering, with seasonalty as it tapers slowly at multiples of 5. The pacf also demonstrates spikes at lags 1, 5, and 13. Judging from these observations, this appears to be an AR model. Our candidate pool for model selection, from most to least preferred, will be AR(1), AR(5), or AR(13). In this scenario, we will fit AR(1).

2. Fit the ARMA model you proposed above using the arima() command. Write down your fitted model. Note that in the output of the arima command, 'intercept' refers to the mean of the process, which we denote by µ in class.

```
#fit the model
x = arima(ts_data, order=c(1,0,0) ,include.mean=T)
```

The parameter is: ar1=0.865, intercept = -1.9591 The model is: y_t = -1.9591 + 0.5843y_{t-1} + e_t

3. Use the confint() command to find 95% confidence intervals for relevant parameters

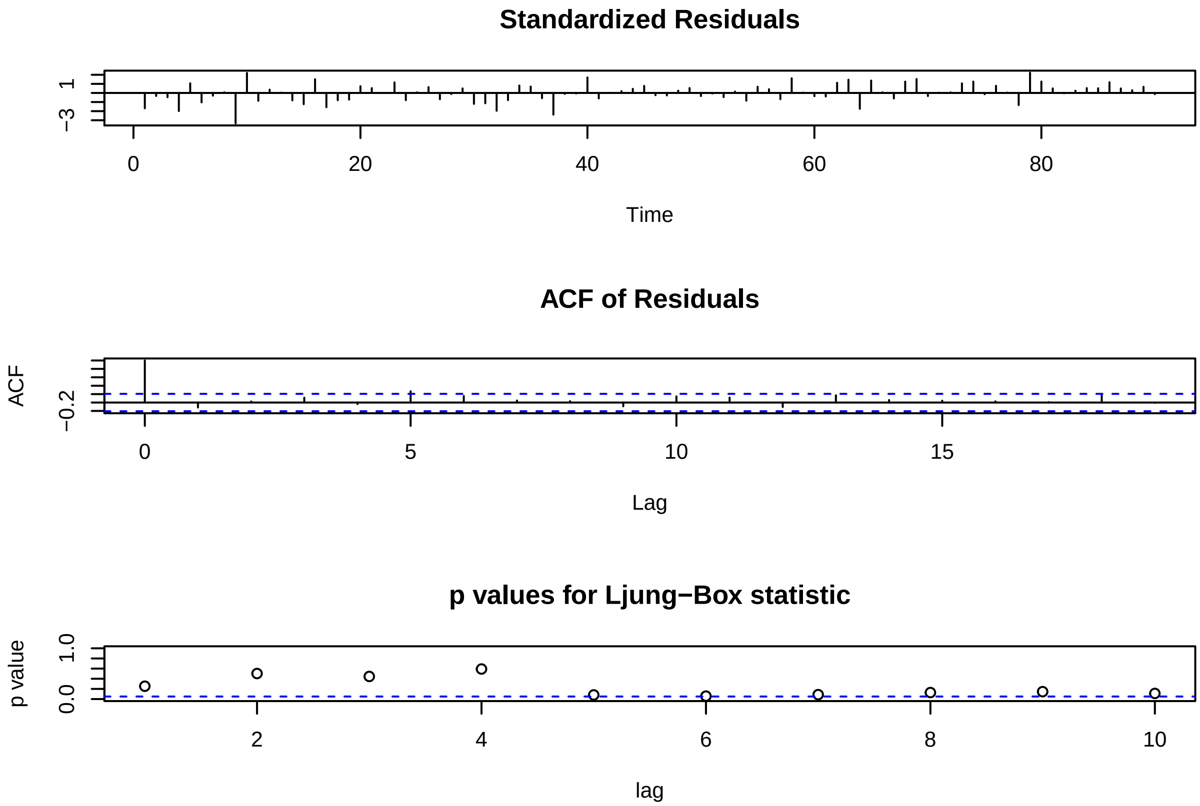
```
confint(x, level=0.95)
```

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

The 95% CI for ar1 is: [0.4150038, 0.753554] And for the intercept: [-2.5098255, -1.408472]

4. Use the tsdiag() function to see diagnostic plots for the model you have fitted. Comment on each plot. How well does the model you proposed appear to fit?

```
tsdiag(x)
```



Lookng at the standardized residuals plot, we can see that the residuals appear to be scattered around a zero horizontal level, and there seems to be no trend. There also seems to be no outliers as all points are contained within the critical values +/- 3.15.

Looking at the acf of residuals, there is no obvious serial correlation.

From the Ljung-Box statistic, the p-values appear to be large, meaning we cannot reject the null hypothesis, and therefore can assume that the residuals are independent and the model does not show lack of fit.