STAT 443 Lab 6

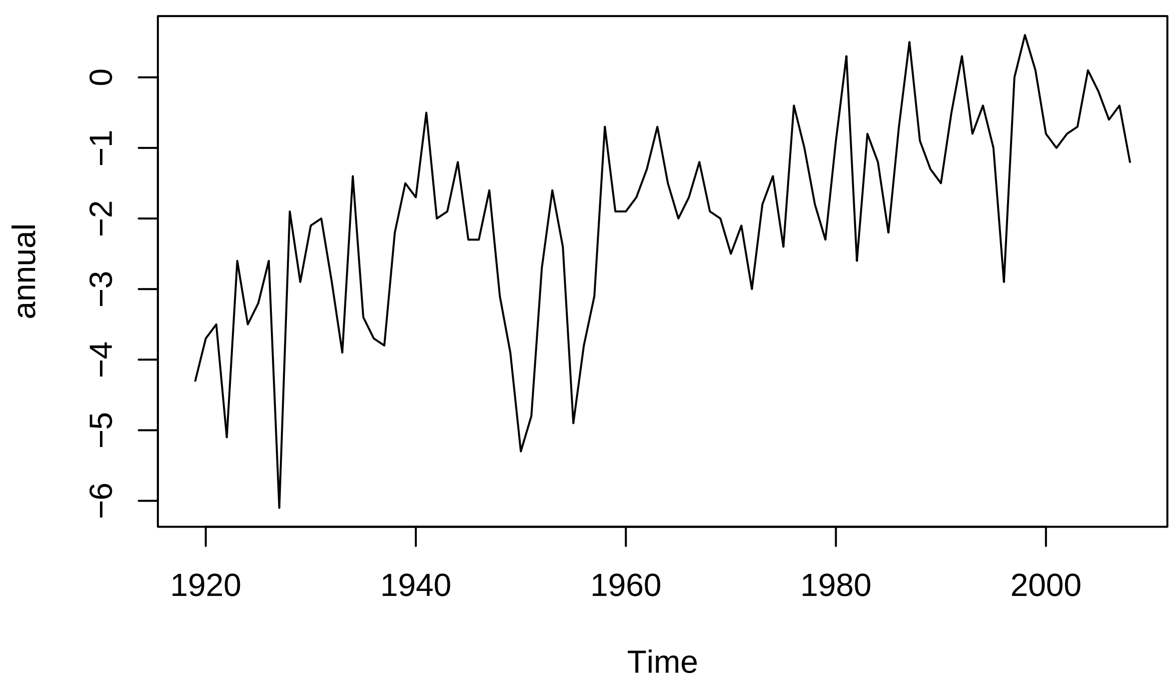
Derek Situ (62222807)

2/26/2022

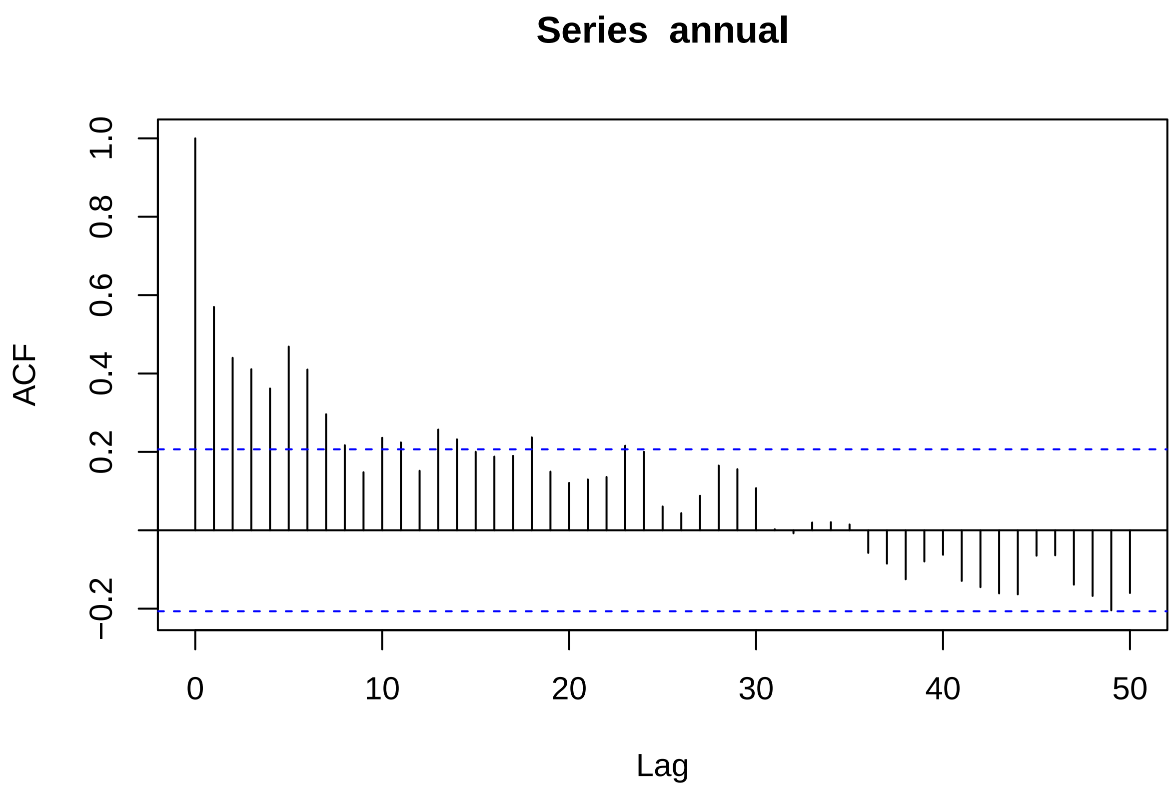
```r
temp <- read.csv("TempPG.csv", header = TRUE)
```

Question 1

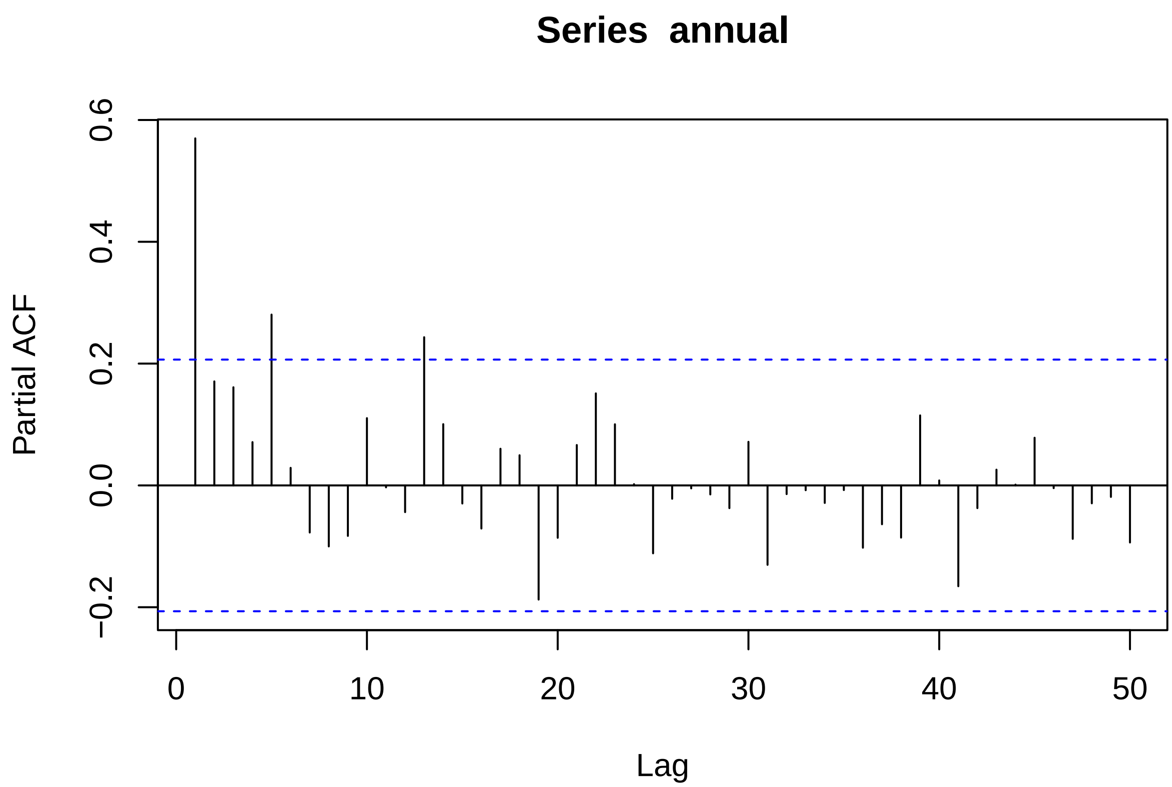
```r
annual <- ts(temp$Annual, start = 1919, end = 2008)
plot(annual)
```




```r
acf(annual, lag.max = 50)
```




```
pacf(annual, lag.max = 50)
```



The acf stays positive for many lags before it tails off, probably because of the trend in the data, and the pacf tails off slowly with no pattern. If I had to fit an ARMA model to this data (rather than SARIMA to address the non-stationarity), I would try an ARMA(3,1) model since it is relatively simple as it has a low order and yet may still capture the correlations we want.

Question 2

```
arma_fit <- arima(annual, order = c(3,0,1), include.mean = TRUE, method = "CSS")
c(arma_fit$coef, sigma2 = arma_fit$sigma2)
```

```
##         ar1         ar2         ar3         ma1   intercept      sigma2
##  1.25509562 -0.28034046  0.01296494 -1.06783960 -0.62085109  0.93300387
```

Let $\{Z_t\}_{t\in\mathbb{N}}$ be a white noise process with mean zero and variance 0.93. The fitted model is

$$X_t + 0.62 = 1.26(X_{t-1} + 0.62) - 0.28(X_{t-2} + 0.62) + 0.013(X_{t-3} + 0.62) + Z_t - 1.07Z_{t-1}.$$

Question 3

```
confint(arma_fit, level = 0.95) |>
  knitr::kable(booktabs = TRUE,
               col.names = c("Lower bound", "Upper bound"),
```

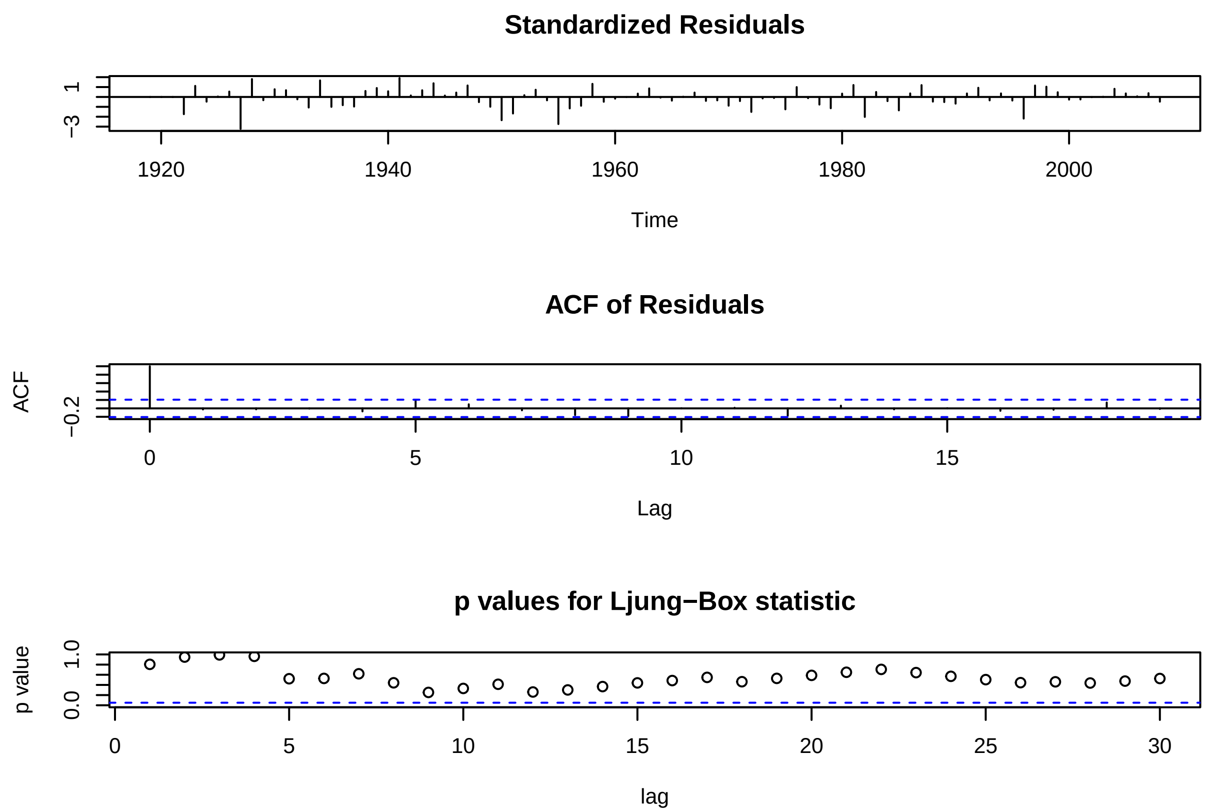
```r
                caption =
                  "95% confidence intervals for the parameters of the fitted model")
```

Table 1: 95% confidence intervals for the parameters of the fitted model

	Lower bound	Upper bound
ar1	1.2361574	1.2740338
ar2	-0.3101291	-0.2505518
ar3	0.0018592	0.0240707
ma1	-1.0814570	-1.0542222
intercept	-0.7124481	-0.5292540

Question 4

```r
tsdiag(arma_fit, gof.lag = 30)
```



Plot of standardized residuals

The residuals appear to be random, which is a good sign that the model fits well.

Plot of acf of residuals

The acf of the residuals is very small for all lags, which is another sign that the residuals are random, which is good in terms of how well the model fits.

Plot of p-values for Ljung-Box test

The p-values for the Ljung-Box test are quite high, which means that the Portmanteau test statistic is consistent with the chi-squared distribution. This is a sign that the model fits well.

Overall fit

Overall, the model appears to fit well, as each of the plots suggest that the residuals are not correlated. I compared the diagnostic plots generated from fitting this model with several other models and found that this model was one of the better-fitting ones while still being relatively simple. For example, fitting an ARMA(1,1) model would result in less parameters, but the Portmanteau p-values would be quite low, indicating correlation between the residuals.