STAT 443 Lab 9

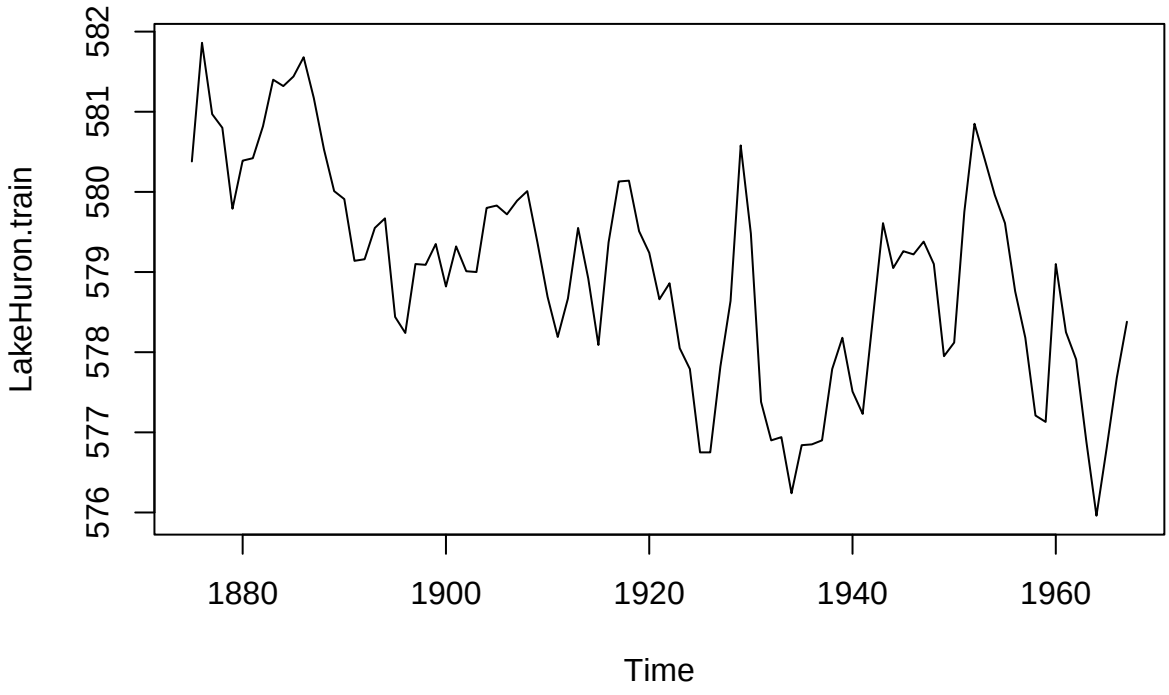
Derek Situ

2022-03-21

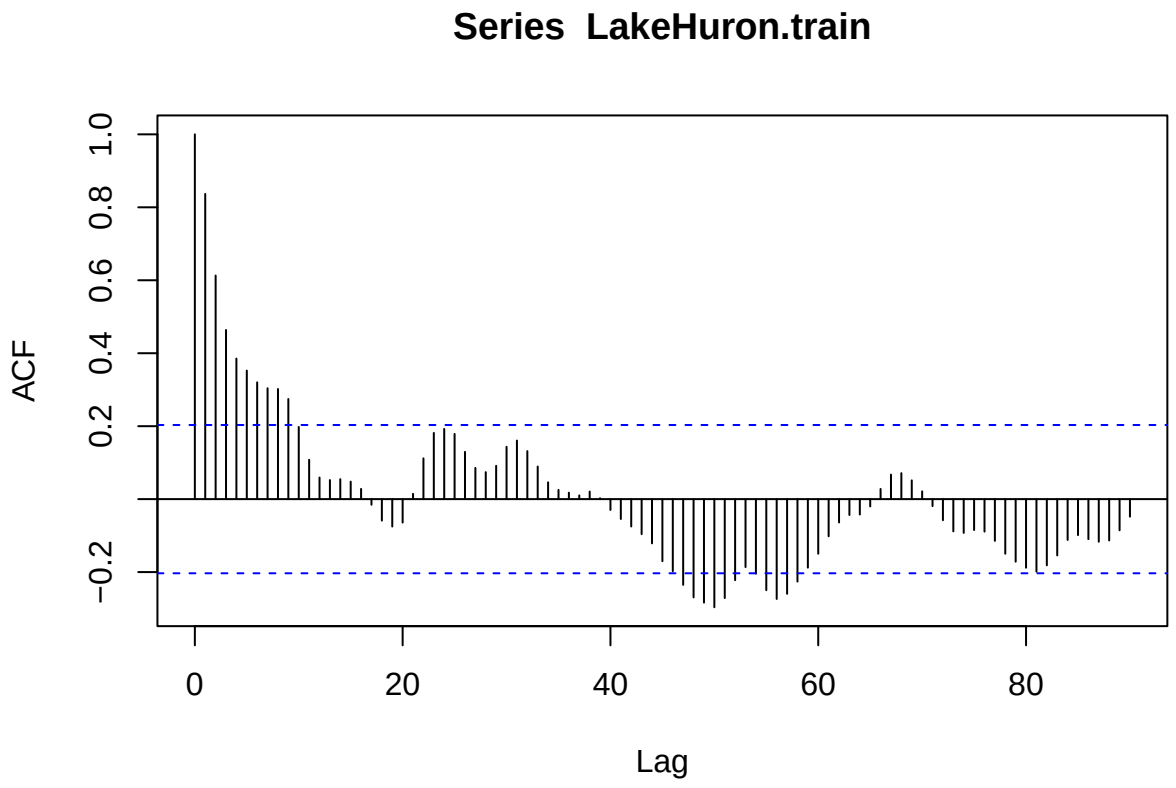
```
data(LakeHuron)
LakeHuron.train <- window(LakeHuron, start = 1875, end = 1967)
LakeHuron.test <- window(LakeHuron, start = 1968, end = 1972)
```

Question 1

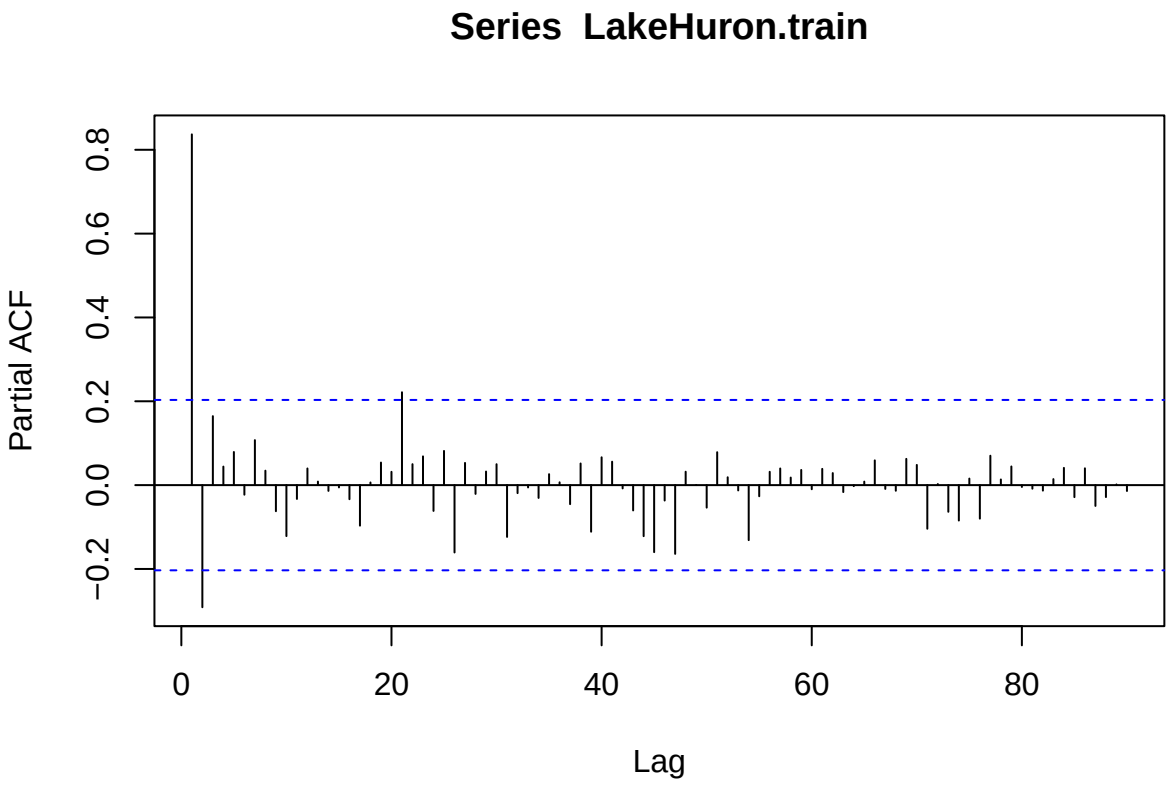
```
plot(LakeHuron.train)
```




```
acf(LakeHuron.train, lag.max = 90)
```




```
pacf(LakeHuron.train, lag.max = 90)
```



The ARMA model I would pick is ARMA(2, 0) or AR(2) because the acf decays slowly and the pacf cuts off at lag 2.

Question 2

```
model_fit <- arima(LakeHuron.train, order = c(2, 0 , 0))
c(model_fit$coef, sigma2 = model_fit$sigma2)
```

```
##         ar1         ar2   intercept      sigma2
##   1.0626979  -0.2691039 578.9887676   0.4814825
```

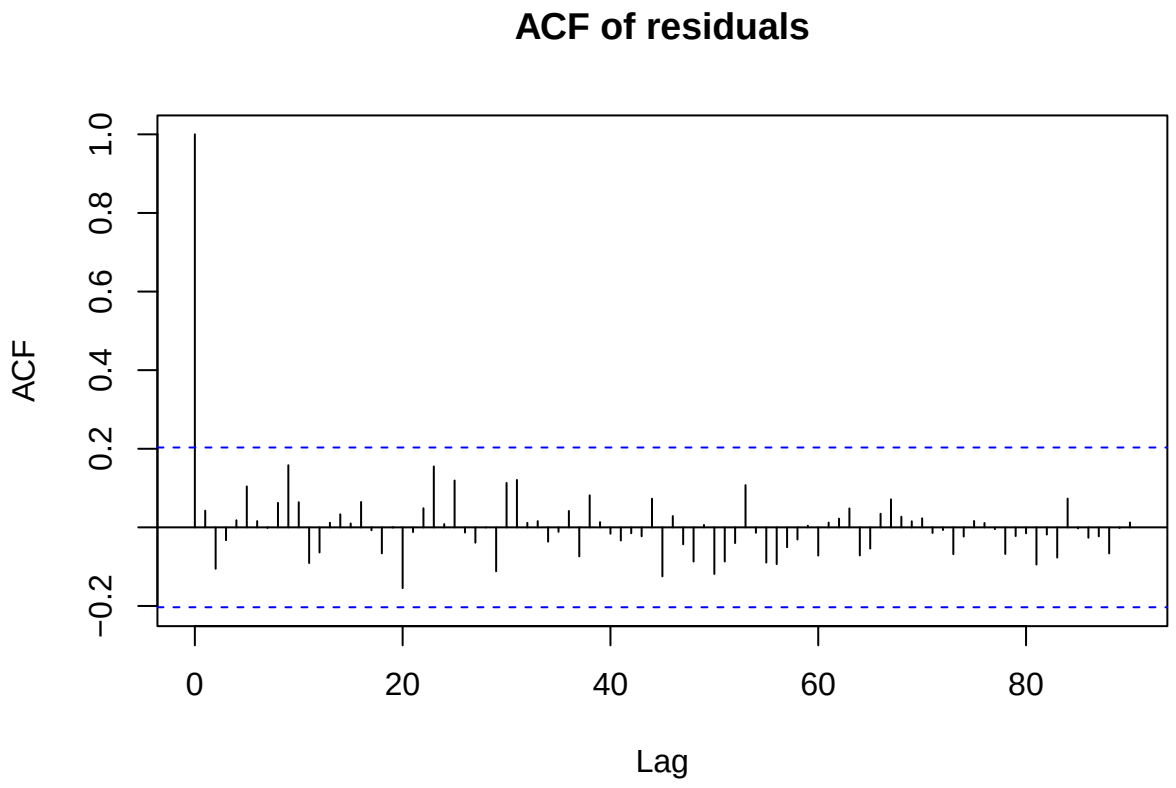
The fitted model is

$$X_t - 578.989 = 1.063(X_{t-1} - 578.989) - 0.269(X_{t-2} - 578.989) + Z_t$$

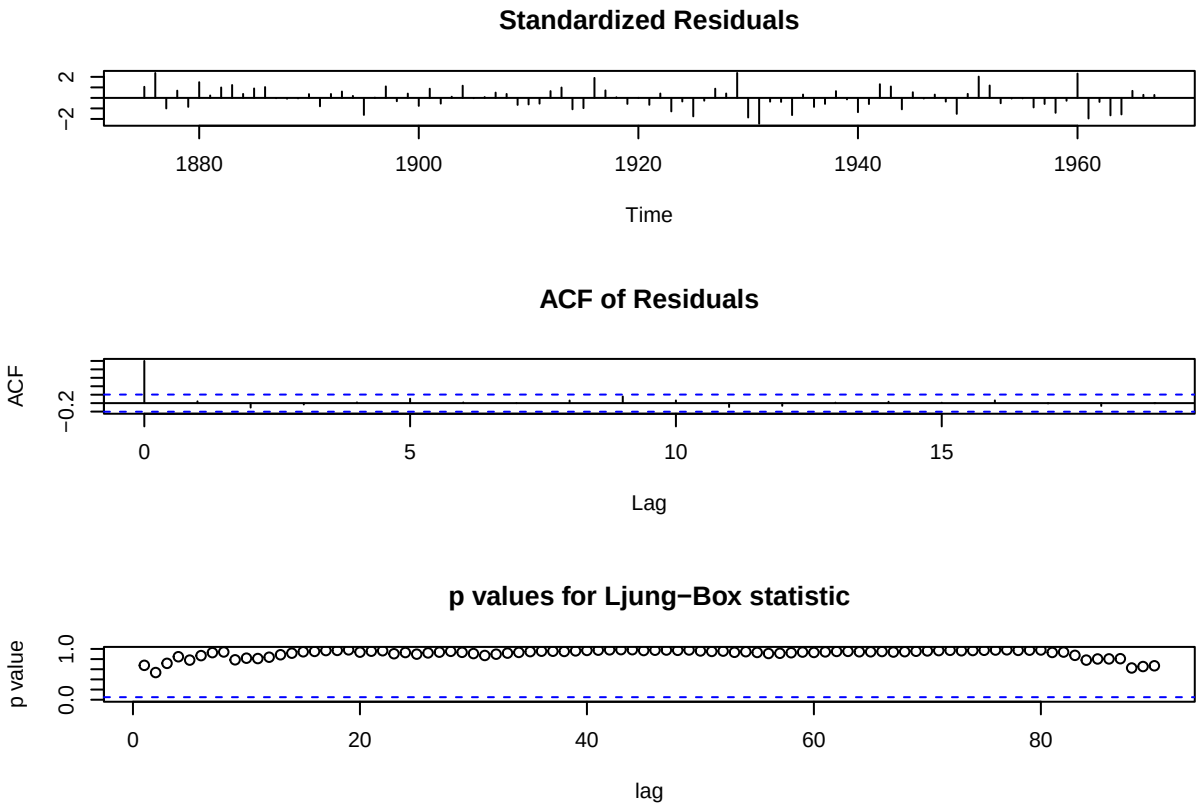
with $\sigma^2 = 0.481$.

Question 3

```
acf(model_fit$residuals, lag.max = 90, main = "ACF of residuals")
```




```
tsdiag(model_fit, gof.lag = 90)
```



From the plot of the acf of the residuals, we see that there are no significant autocorrelations after lag 0. The plot of the standardized residuals shows no pattern and the p-values for the Ljung-Box test are very high. These are all signs that the model fits well.

Question 4

```
prediction <- predict(model_fit, n.ahead = 3)
knitr::kable(data.frame(Year = c(1968, 1969, 1970),
                        Estimate = prediction$pred,
                        CI_LowerBound = prediction$pred - 1.96 * prediction$se,
                        CI_UpperBound = prediction$pred + 1.96 * prediction$se),
             booktabs = TRUE)
```

Year	Estimate	CI_LowerBound	CI_UpperBound
1968	578.6940	577.3340	580.0540
1969	578.8394	576.8548	580.8239
1970	578.9093	576.6056	581.2131

Question 5

```
knitr::kable(data.frame(Year = c(1968, 1969, 1970),
                        Observed = LakeHuron.test[1:3],
                        Estimate = prediction$pred,
                        Residual = LakeHuron.test[1:3] - prediction$pred))
```

Year	Observed	Estimate	Residual
1968	578.52	578.6940	-0.1740260
1969	579.74	578.8394	0.9006319
1970	579.31	578.9093	0.4006829

We see that the observed values are within the 95% confidence intervals found in question 4, and the residuals are under 1.