STAT 443: Lab 1

Saksham Sudershan (Student #31339427)

17 January, 2022

Question 1

(a)

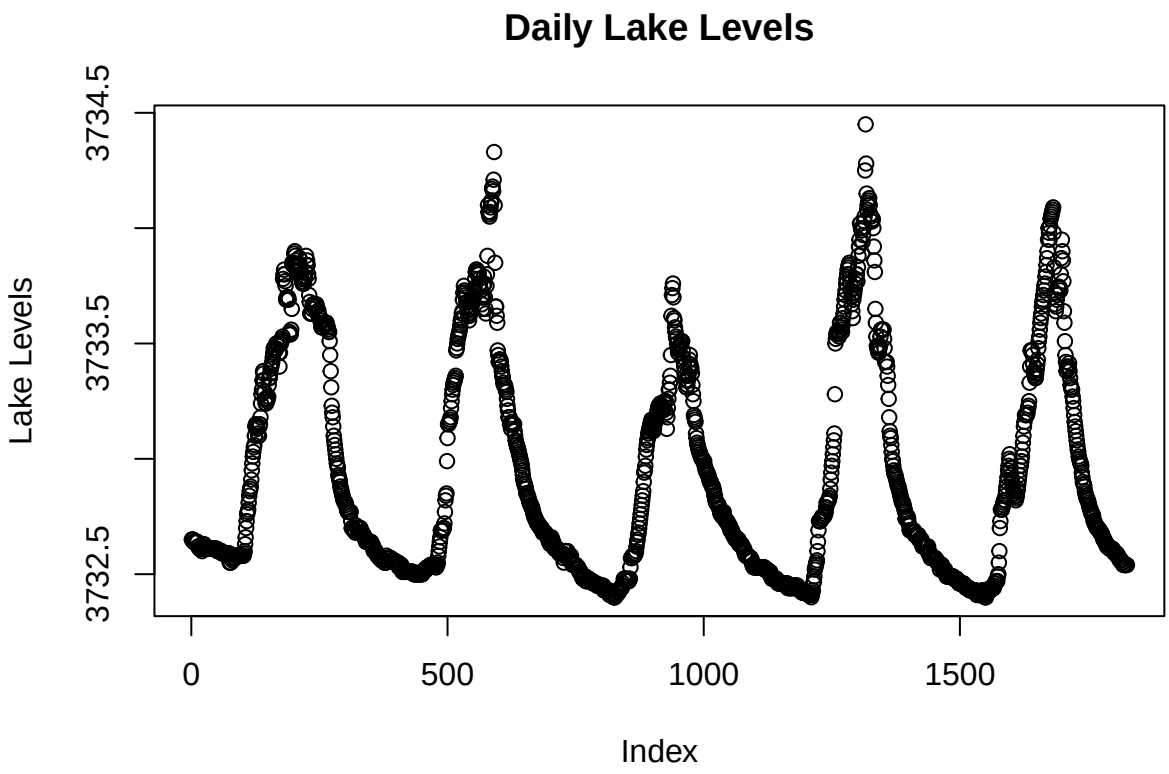
```
# Reading data
data <- read.csv("LakeLevels.csv")

# Looking at data
head(data,10)
```

```
##         Date LakeLevel
## 1   1/1/2007   3732.65
## 2   1/2/2007   3732.65
## 3   1/3/2007   3732.65
## 4   1/4/2007   3732.64
## 5   1/5/2007   3732.64
## 6   1/6/2007   3732.64
## 7   1/7/2007   3732.64
## 8   1/8/2007   3732.64
## 9   1/9/2007   3732.64
## 10 1/10/2007   3732.64
```

```
# List of vectors in dataset
names(data)
```

```
## [1] "Date"      "LakeLevel"
```

```
# Creating plot
plot(data$LakeLevel, ylab = "Lake Levels", main = "Daily Lake Levels")
```



The plot is disjoint, and the x-axis is indexed for the time variable. For time series data, time should be on the x-axis and the points should be joined.

(b)

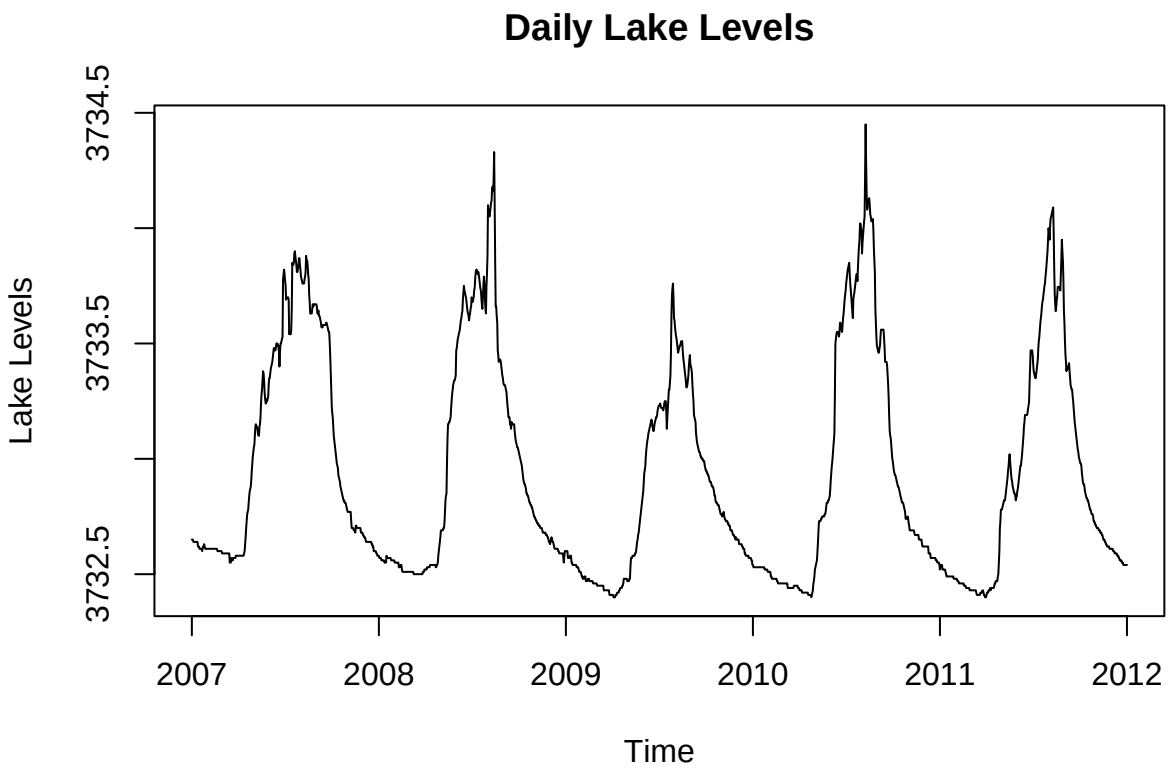
```
# Checking whether dataframe is a time series object
is.ts(data)
```

```
## [1] FALSE
```

```
# Fitting data into time series object
x <- ts(data = data$LakeLevel, start = c(2007,1), frequency = 365)
```

(c)

```
# Plotting time series
plot(x, xlab="Time", ylab= "Lake Levels", main = "Daily Lake Levels")
```



The points are joint and the x-axis shows time in years. It has no discernible trend but does have a seasonality effect; the lake levels seem to increase every summer and go back down towards the winter of each year.

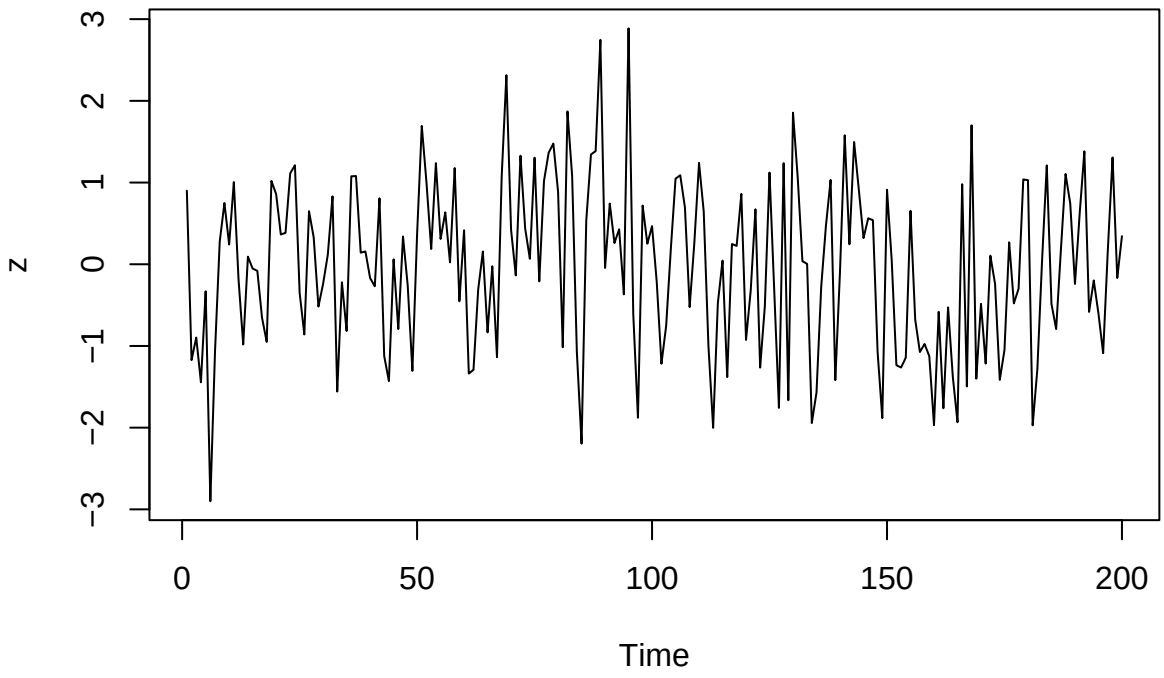
Question 2

(a)

```
# Generating independent observations from standard normal distribution
set.seed(2022)
y <- rnorm(200)

# Fitting into class ts
z <- ts(y)
```

(b)

```
# Plotting
plot(z)
```



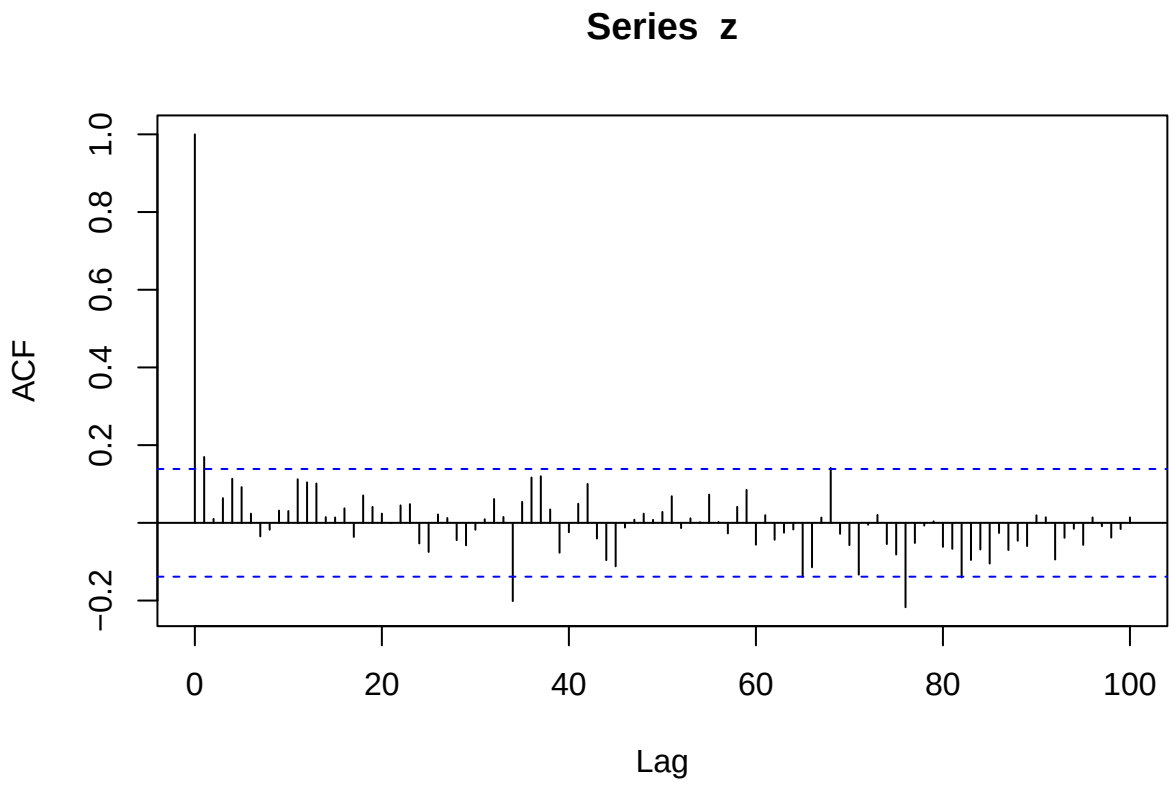
Since the standard deviation of our normal distribution is 1, values outside the range of 2 or -2 would be 2 standard deviations away from the mean which is 0. Using the empirical rule, we can say that we would expect only 5% of values to be outside, which is 10 values.

```
# Counting no. of observations outside the range +-2
sum(abs(y)>2)
```

```
## [1] 6
```

(c)

```
# Creating sample autocorrelation function
acf(z, lag.max = 100, plot = TRUE)
```



The autocorrelation function shows that there is little correlation between past and present values. The autocorrelation for h=0 is 1, and almost all of the following autocorrelations fall within the 95% confidence limits and tends to stay within $\frac{\pm 2}{\sqrt{n}}$. We can say that the data is random.