HOMEWORK ASSIGNMENT 2

Zachary Lazerick

2/9/2023

PROBLEM 1: Power loads in the southern USA

Consider the 2006 - 2009 quarterly power loads for a utility company located in a southern part of the United States. The power load x_t is given in megawatts and contained in the file `QTRPOWER.Rdata`.

(a) Use the `load()` function to read in the data from `QTRPOWER.Rdata`. It will automatically be assigned to the data frame `QTRPOWER`. Don't forget to run `attach()` on your data frame after loading it.

```
load("QTRPOWER.Rdata")
attach(QTRPOWER)
```

(b) Look at the variables in the data frame. Use this information to convert the data frame to a time series using the `ts()` function. Then create an annual time series, where the data are aggregated over quarter.

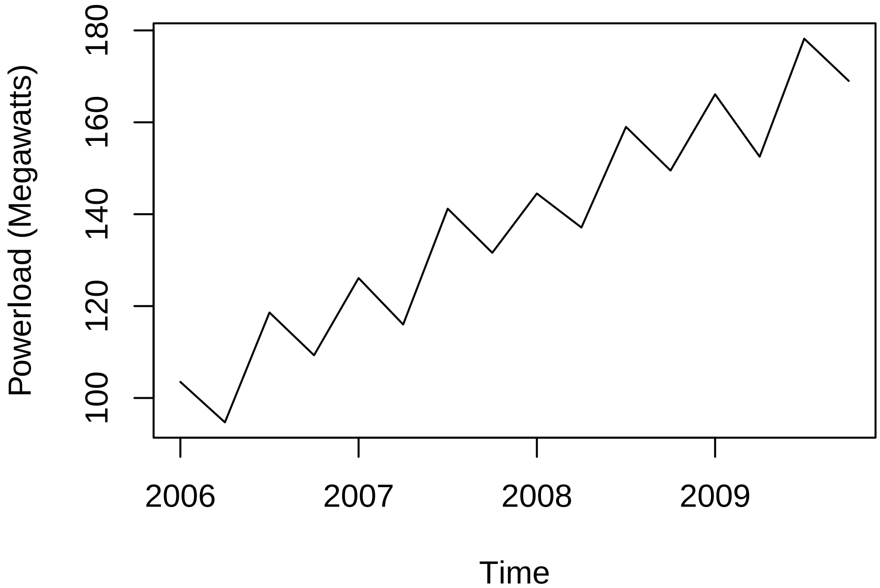
```
## Create Time Series
powload.ts <- ts(POWLOAD, start = c(2006, 1), end = c(2009, 4),
                 freq = 4)

## Aggregate by Quarter
powload.annual.ts <- aggregate(powload.ts)/4
```

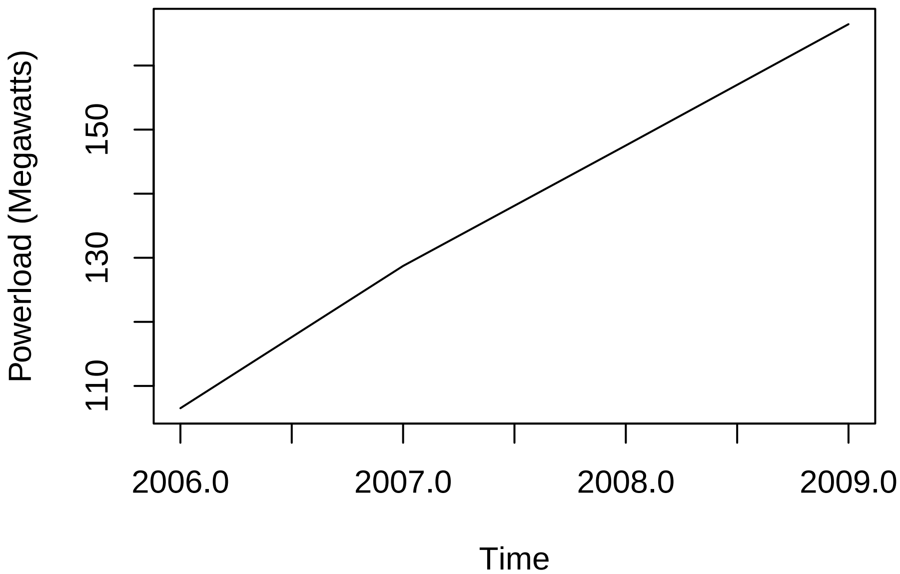
'Year' is a number that contains the years when the data was collected. 'T' is a number ranging from 1 to 16 that serves as a marker for each individual data point. 'POWLOAD' is a number that contains the power load of the utility company, measured in megawatts.

(c) Plot the quarterly and annual time series. Which plot allows you to clearly see the secular trend? How would you describe this trend?

```r
## Plot Quarterly Time Series
plot(powload.ts, ylab = "Powerload (Megawatts)")
```




```r
##Plot Annual Time Series
plot(powload.annual.ts, ylab = "Powerload (Megawatts)")
```



The secular trend can be seen from both plots, however, the secular trend is most clearly seen in the annual time series. It appears as though all noise has been smoothed from performing this aggregation. The secular trend is linear, strictly increasing, suggesting that the power load for this utility company continues to rise as time goes on.

(d) Use the `window()` function to sample the original time series for each quarter separately. Then calculate the ratio of power loads for each quarter, relative to the mean for the original time series. Describe the power loads for each quarter relative to the mean. By what percentage are they higher or lower?

```r
## Quarter 1
powload.1 <- window(powload.ts, start = c(2006, 1), freq = TRUE)
Q1.ratio <- mean(powload.1) / mean(powload.ts)
Q1.ratio
```

```
## [1] 0.9835678
```

The power load for Q1 is about 1.65% below the mean for the whole series.

```r
## Quarter 2
powload.2 <- window(powload.ts, start = c(2006, 2), freq = TRUE)
Q2.ratio <- mean(powload.2) / mean(powload.ts)
Q2.ratio
```

```
## [1] 0.9109199
```

The power load for Q2 is about 8.9% below the mean for the whole series, the lowest on average for any quarter.

```r
## Quarter 3
powload.3 <- window(powload.ts, start = c(2006, 3), freq = TRUE)
Q3.ratio <- mean(powload.3) / mean(powload.ts)
Q3.ratio
```

```
## [1] 1.086986
```

The power load for Q3 is about 8.69% above the mean for the whole series, the highest on average for any quarter.

```r
## Quarter 4
powload.4 <- window(powload.ts, start = c(2006, 4), freq = TRUE)
Q4.ratio <- mean(powload.4) / mean(powload.ts)
Q4.ratio
```

```
## [1] 1.018526
```

The power load for Q4 is about 1.8% above the mean for the whole series.

PROBLEM 2: Quarterly earnings of Johnson & Johnson

Consider the time series data of quarterly earnings of Johnson & Johnson between 1960 and 1980. There are 84 quarters over 21 years.

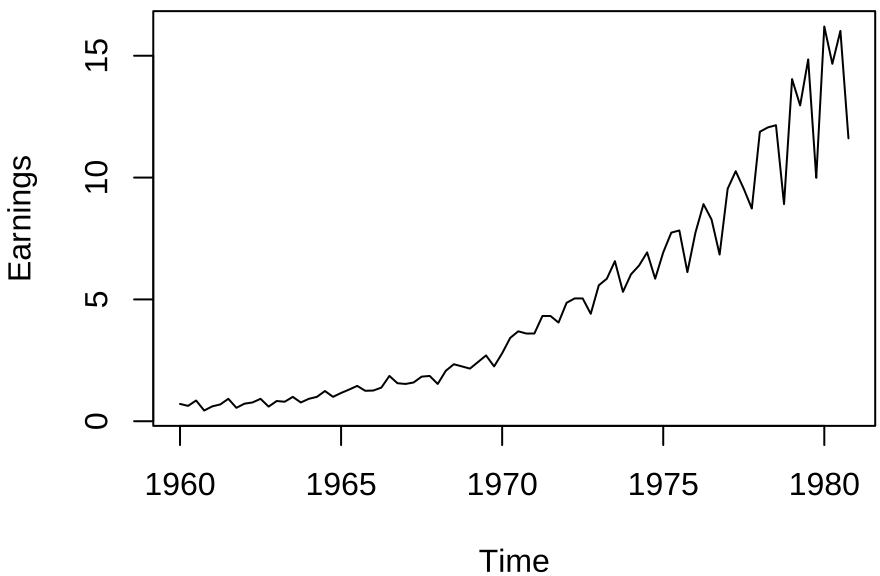
(a) Use the `load()` function to read in the data from `tsa3.rda`. You will see several data frames loaded, as well as the time series of interest to us, `jj`.

```
load("tsa3.rda")
```

(b) Create two time series variables, `JJ.ts` (the quarterly time series) and `JJ.annual` (the annual time series). Use the `aggregate()` function to remove any seasonal effects within each year and produce an annual series of quarterly earnings. Then plot both series.

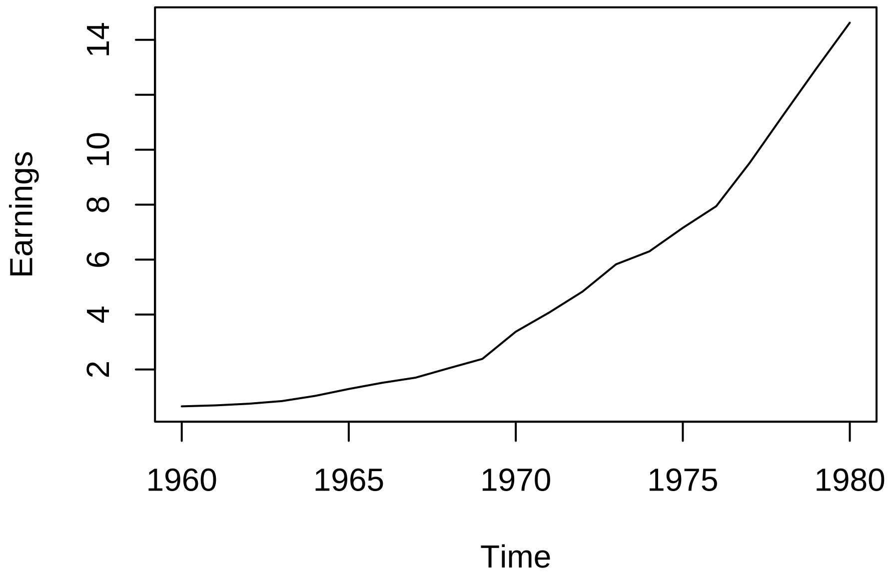
```
## Quarterly Time Series
JJ.ts <- jj

plot(JJ.ts, ylab = "Earnings")
```



```r
## Annual Time Series
JJ.annual <- aggregate(JJ.ts)/4

plot(JJ.annual, ylab = "Earnings")
```



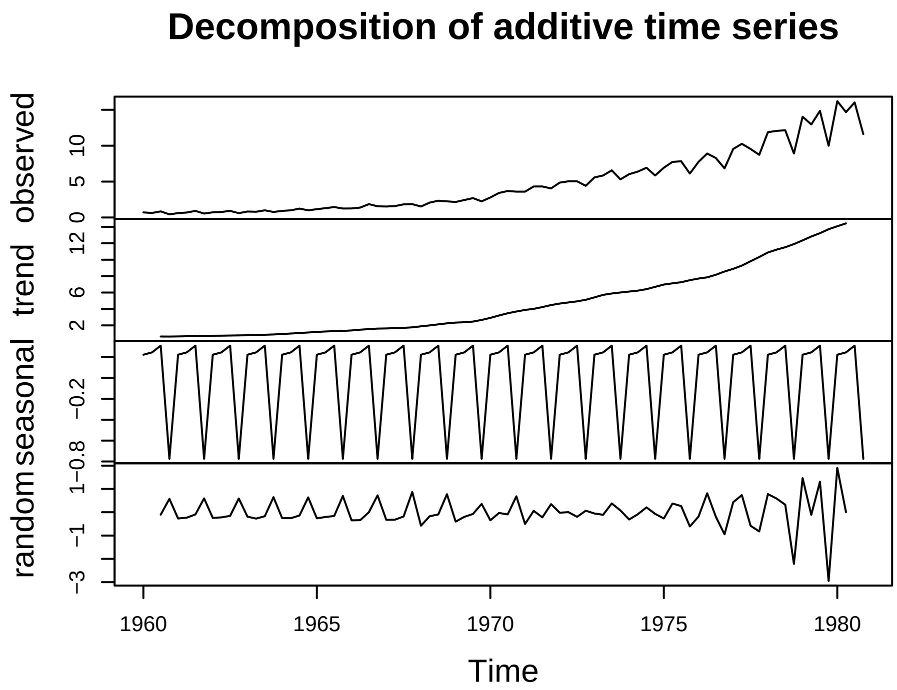
(c) How would you describe the secular trend: linear, quadratic, or exponential?

The secular trend looks exponential.

(d) Decompose the quarterly time series using the `decompose` function with an additive model. Then plot the decomposed time series. You should see four stacked plots in the plotting area.

```r
## Decompose Series
JJ.decom.add <- decompose(JJ.ts, type = "additive")

## Plot Decomposed Series
plot(JJ.decom.add)
```



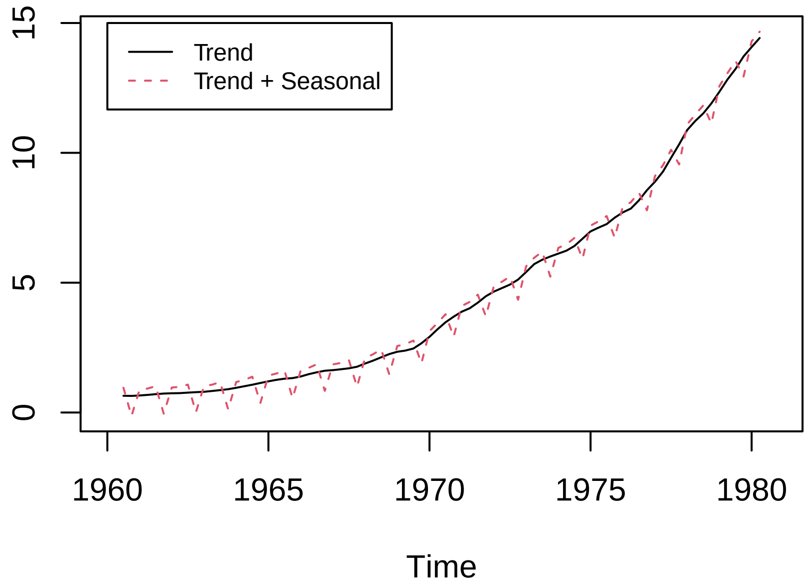
(e) Based on the random component, do you think the additive model is appropriate for this data? Briefly explain.

Based on the plot, I think that the additive model is appropriate for the data from the years 1960 to about 1975. During this time frame, the size of the random error term stays largely consistent. However, after the year 1975, the error term grows almost triples in size, suggesting that the additive model is not appropriate for this portion of the data.

(f) Separate out the `trend` and `seasonal` components into variables. Then use the `ts.plot()` and `cbind()` functions to plot the trend overlayed by the trend + seasonal components. Does the resulting plot agree with the original plot of the quarterly time series?

```
## Isolate Trend and Seasonal Components
JJ.add.trend <- JJ.decom.add$trend
JJ.add.seasonal <- JJ.decom.add$seasonal

## Plot Decomposed Components with a Legend
ts.plot(cbind(JJ.add.trend, JJ.add.trend + JJ.add.seasonal),
        col = 1:2, lty = 1:2)
legend(x = 1960, y = 15, legend = c("Trend", "Trend + Seasonal"),
       col = 1:2, lty = 1:2, cex = .75)
```

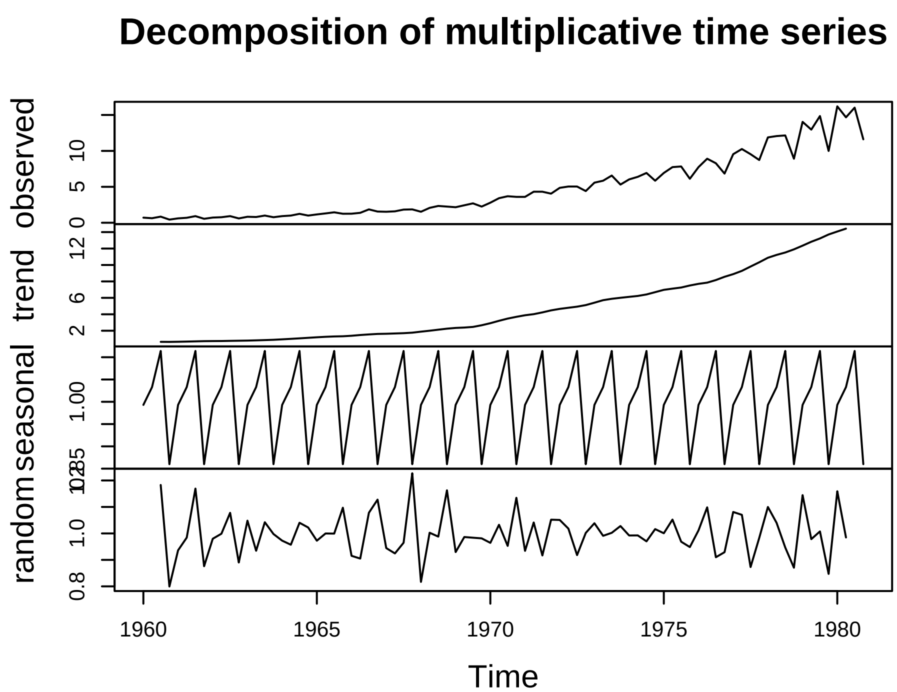


The resulting plot does mostly agree with the original quarterly time series.

(g) Decompose the quarterly time series using the `decompose` function with a multiplicative model. Then plot the decomposed time series. You should see four stacked plots in the plotting area.

```
## Decompose Series
JJ.decom.mult <- decompose(JJ.ts, type = "multiplicative")

## Plot Decomposed Series
plot(JJ.decom.mult)
```



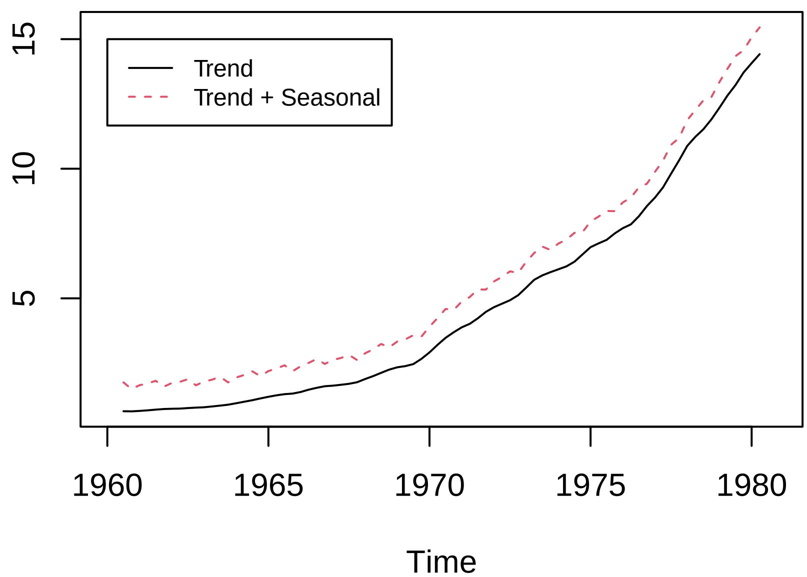
(h) Based on the random component, do you think the multiplicative model is appropriate for this data? Briefly explain.

Based on the plot, I think that the multiplicative model is not appropriate for the data overall. The size of the random error term varies greatly from datapoint to datapoint, suggesting that the multiplicative model does not fit the data well.

(i) Separate out the `trend` and `seasonal` components into variables. Then use the `ts.plot()` and `cbind()` functions to plot the trend overlayed by the trend + seasonal components. Does the resulting plot agree with the original plot of the quarterly time series?

```
## Isolate Trend and Seasonal Components
JJ.mult.trend <- JJ.decom.mult$trend
JJ.mult.seasonal <- JJ.decom.mult$seasonal

## Plot Decomposed Components with a Legend
ts.plot(cbind(JJ.mult.trend, JJ.mult.trend + JJ.mult.seasonal),
        col = 1:2, lty = 1:2)
legend(x = 1960, y = 15, legend = c("Trend", "Trend + Seasonal"),
       col = 1:2, lty = 1:2, cex = .75)
```



The resulting plot does mostly agree with the original quarterly time series. However, it does not model the original time series as well as the additive model.