Lab 1 Assignment

Cuong Ly

2023-01-16

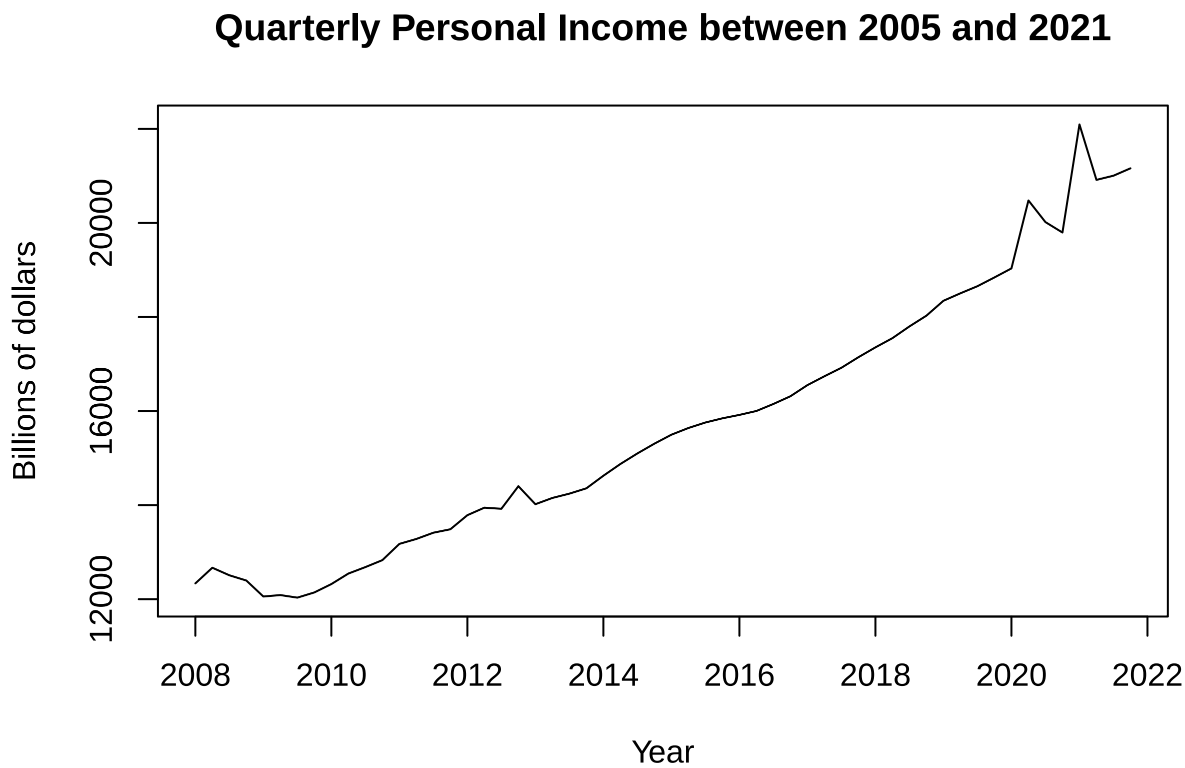
Lab Question 1:

---

1. Download a dataset from the Bureau of Economic Analysis (https://www.bea.gov/), U.S Department of Commerce.

```
data = c(12335.3,12669.4,12508.6,12396.9,12057,12087.3,12033.3,12144,12321.2,12544.7,12682,12830.2,1317
#per_inc <- read.csv('/Users/taikhanghao/Desktop/spring 23/time series/Personal-Income.csv')
ts_per_inc <- ts(data, start = 2008,frequency = 4)
```

```
plot(ts_per_inc,type = 'l', xlab='Year',ylab = 'Billions of dollars',main='Quarterly Personal Income be
```

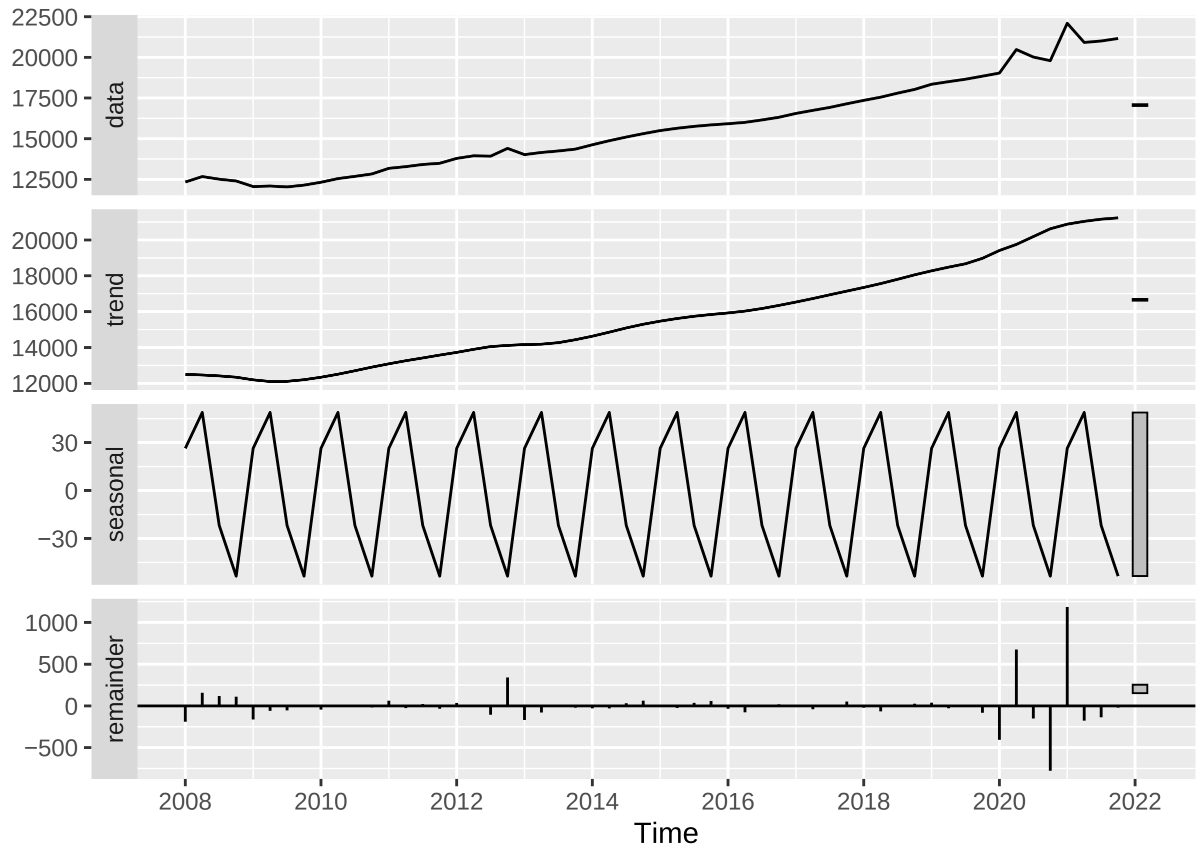


The personal income decreased significantly between 2008 and 2009. The reason for that is due to the Great Recession. People lost their jobs, did not generate income. Houses decreased in value. In 2013. That year, companies reduced wage and salary. Personal tax also gain 0.2%. For these reason, the personal income decreased. The personal income again decreased in the first half of 2020 due to the pandemic. After that, it skyrocketed. A possible explanation is the stimulus check from the government and/or the quantitative easing from the FED. From the graph, we can confidently state that the data is not stationary because there is an up trend. There is no cyclic behavior.

```
stl_inc = stl(ts_per_inc, "periodic")

seasonal_stl_inc   <- stl_inc$time.series[,1]
trend_stl_inc     <- stl_inc$time.series[,2]
random_stl_inc  <- stl_inc$time.series[,3]

# Decomposition
autoplot(stl_inc)
```

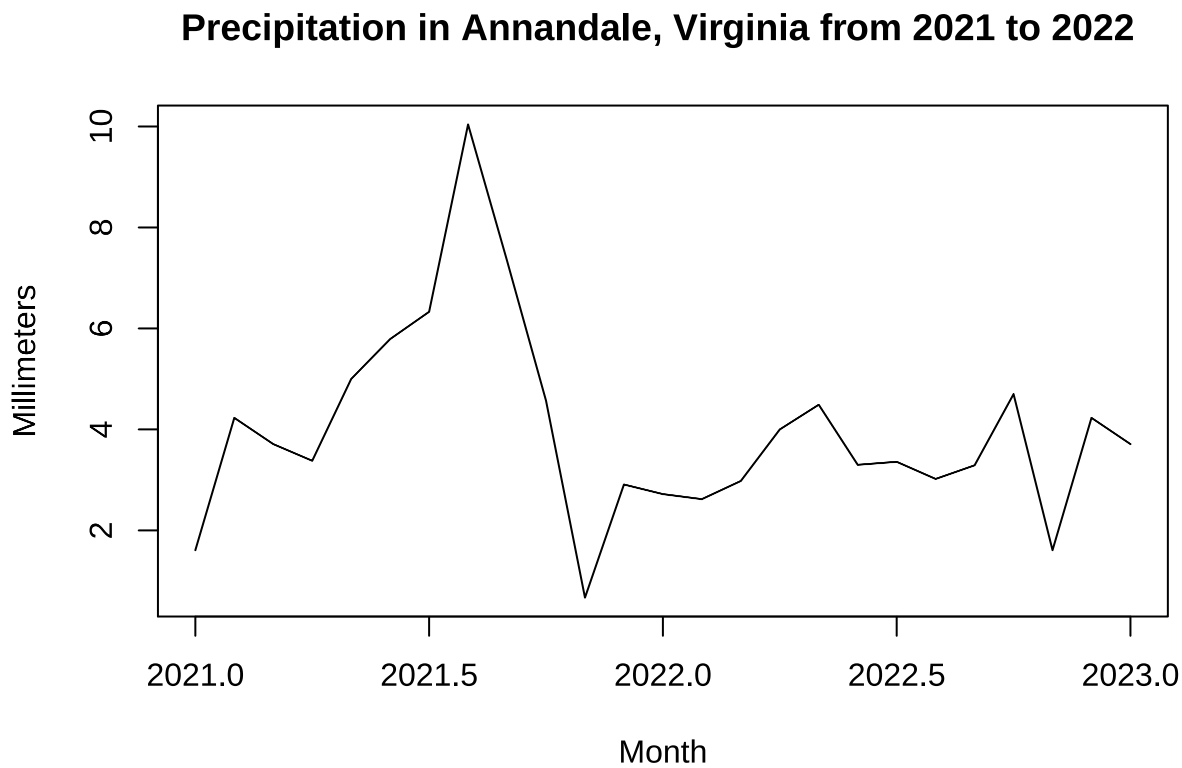


We can observe that the trend increases overall, although there are a few drops in 2009, 2013. The seasonality component stays constantly, telling us the personal income does not get affected by seasonality. There's noise between 2020 and 2021, as expected with the pandamic, people income fluctuates up and down.

2. Download a climate dataset using NOAA.

```
climate <- read.csv('/Users/taikhanghao/Desktop/spring 23/time series/climate.csv')
ts_climate <- ts(climate, start = 2021,end = 2023,frequency = 12)
```

```
plot(ts_climate,type = 'l', xlab='Month',ylab = 'Millimeters',main='Precipitation in Annandale, Virgini
```

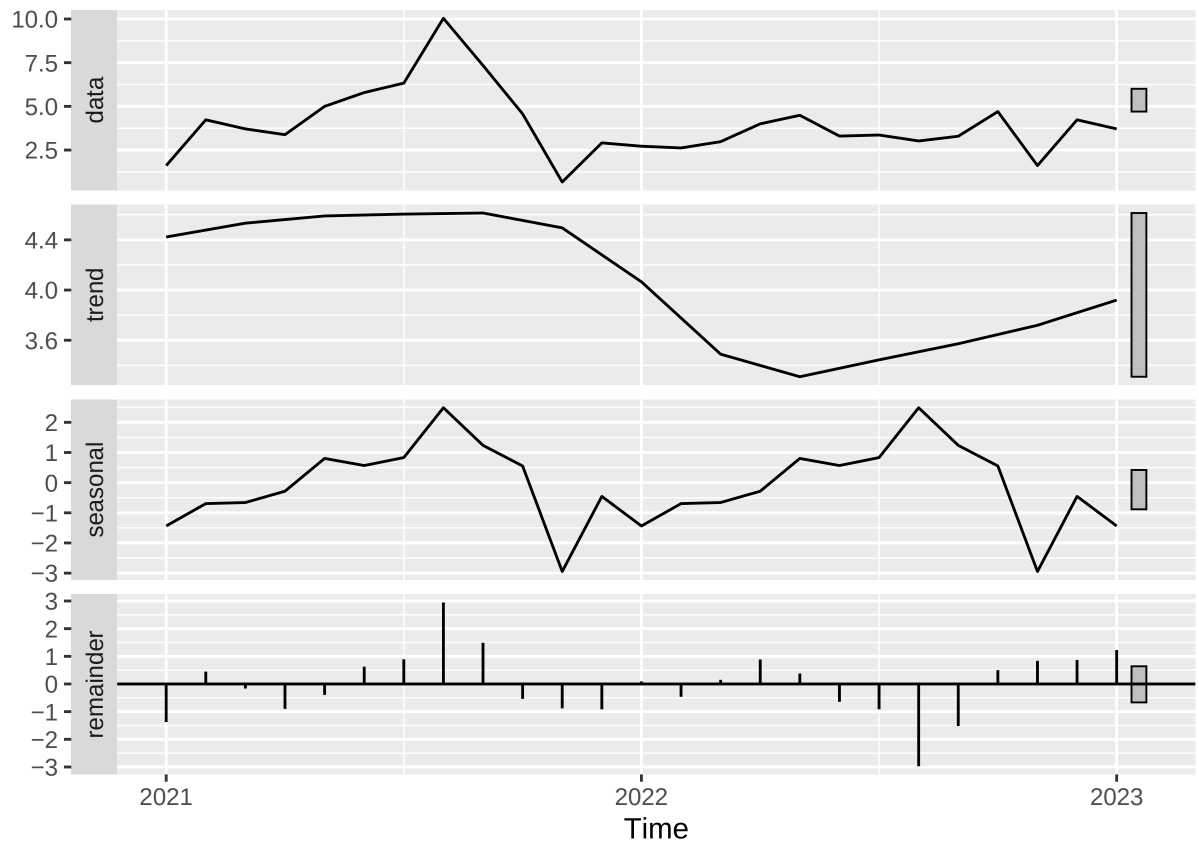


I chose Annandale because I live there. The raining season in Annandale and Virginia in general is from March to November. Months with the most rain are summer months: June, July, August. As we can see from the graph, August '21 has the highest precipitation at 10.8 mm. Compared to 2022, 2021 has more rain. Weather is unpredictable, thus, there is no clear reason why 2021 has more rain than 2022. Some claim that because of climate change, weather is more unforeseeable and intense as year passes. There data may or may not be stationary. Time series are stationary if they do not have trend or seasonal effect. When we compute the decomposition below, there is a trend but it is not clear. There is no seasonality. Thus, more tests need to be conducted in order to conclude if the data is stationary or not. There is no cyclic.

```
stl_climate = stl(ts_climate, "periodic")

seasonal_stl_climate   <- stl_climate$time.series[,1]
trend_stl_climate     <- stl_climate$time.series[,2]
random_stl_climate  <- stl_climate$time.series[,3]

# Decomposition
autoplot(stl_climate)
```



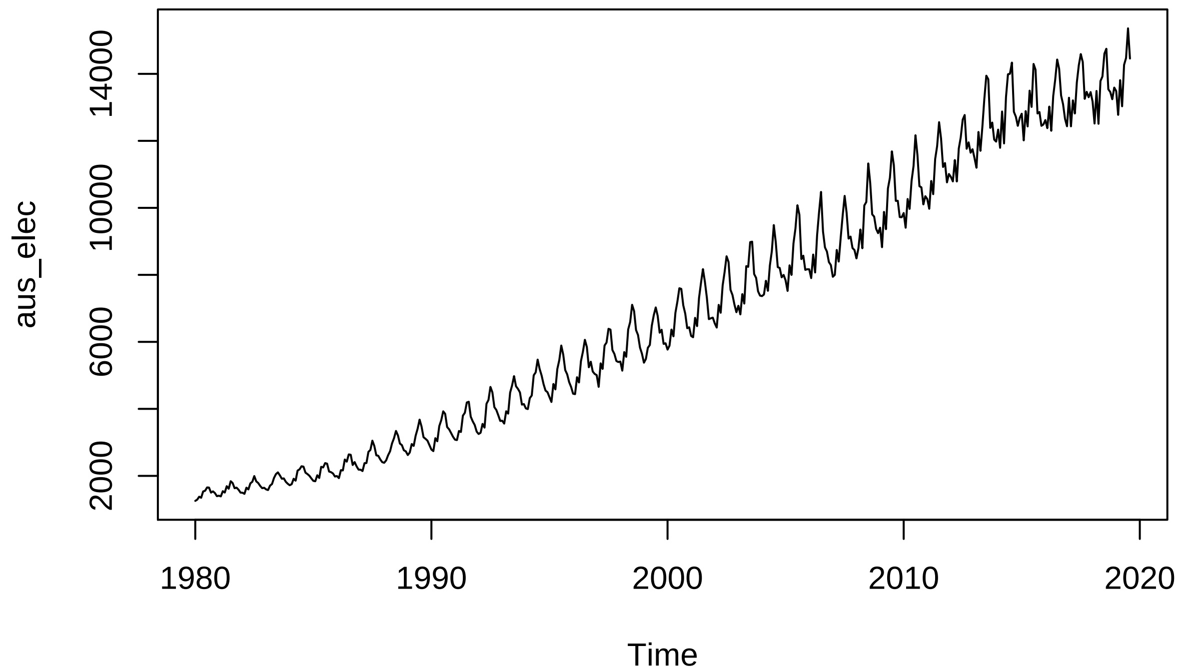
We can see that trend decreases, which saying that 2021 has more rain than 2022. Seasonality graphs confirm that summer months have the most rain.

Lab Question 2:

(a). Plot The monthly Australian electricity demand and comment on any patterns you see. What are the time series components you can visualize from the time series plot? Is it multiplicative or additive? What else can you conclude from it?

Dataset: `"elec" {fma}` ; Choose the starting year as 1980 and make sure to convert the data to a time series object.

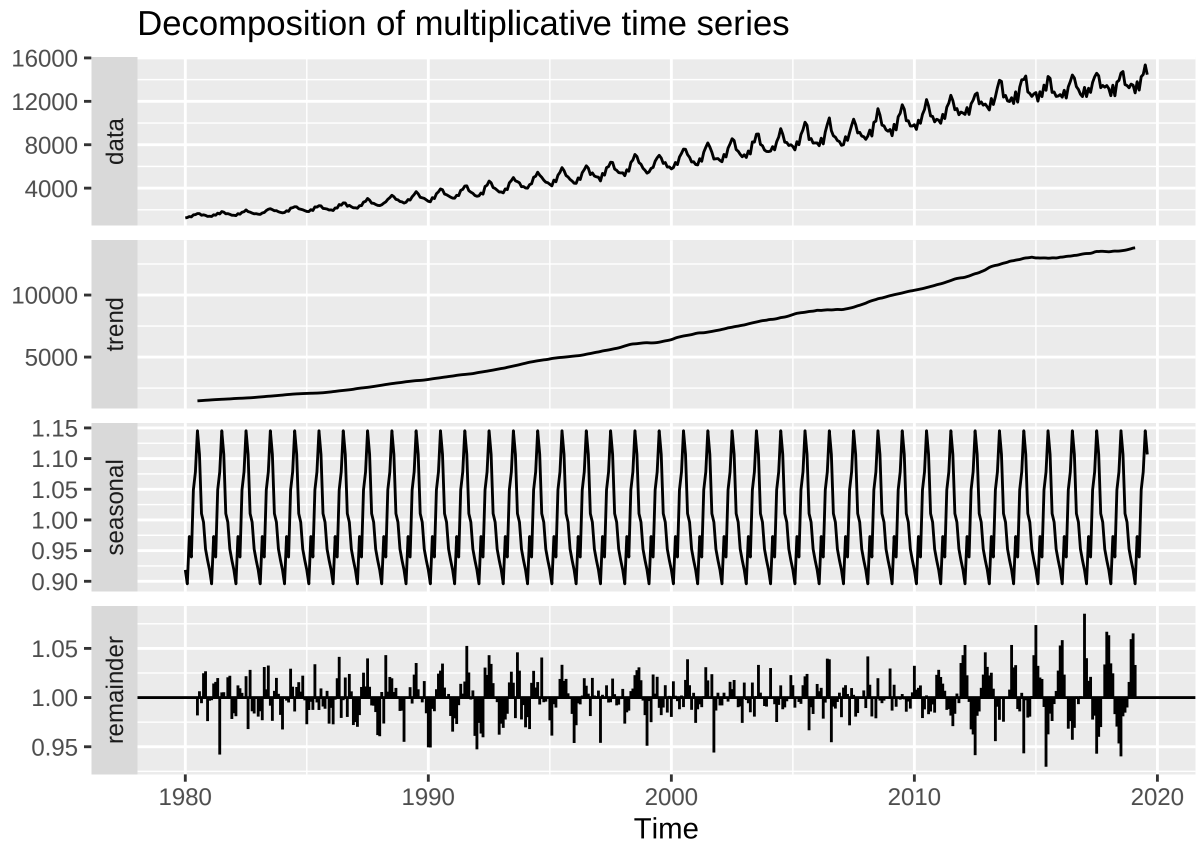
```
aus_elec = elec
aus_elec <- ts(aus_elec, start = 1980, frequency = 12)
plot(aus_elec)
```



The demand grew exponentially from 1980 to 2000. After that, it increased linearly until 2014 and stay relatively constant until 2020. The trend is increasing. There is seasonality. There is no cyclic behavior. It is multiplicative because the magnitude gets bigger as time passes. One interesting point is that there is a kink around 2013. The demade drops slightly before coming back up in 2014.

(b). Use the `decompose()` or `stl()` to decompose the series and comment on each ts component as we did during the Lab and in the lecture.Compare your answers with your answers to part a.

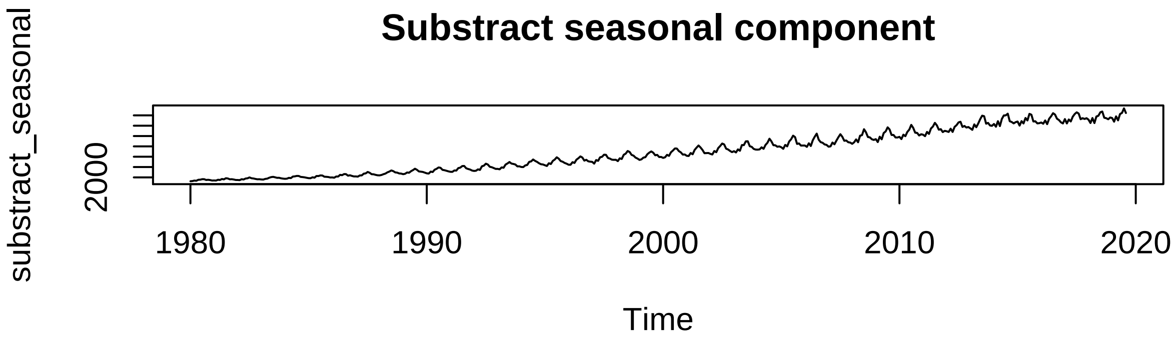
```
decompose_demand = decompose(aus_elec, "multiplicative")
autoplot(decompose_demand)
```

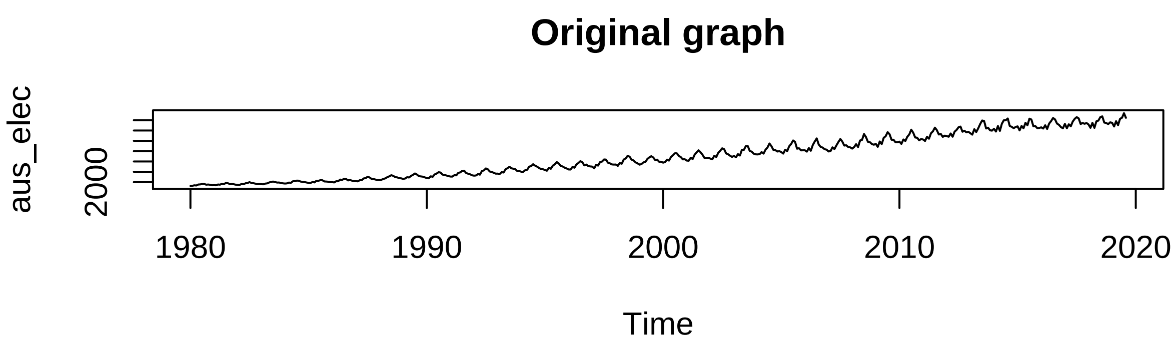


As shown from the trend component in the graph above, the demand grew linearly from 1980 to 2013. However, after 2013, the demand still grew but at a slower pace than before. I did a little research. The reason for a slower demand is because the energy price in 2013 in Australia skyrockted. Energy prices in major cities increased between 30% - 107%. There is a seasonality, but it stays constant yearly. Regarding the noise, we can see there was lots of noises between 2010 and 2020. As explained earlier, energy price contributed to this volatility. These findings corroborated my answer in part a.

(c). Subtract the seasonal component from the original process and plot it and the original graph in the same window (You can use `par(mfrow=c(2,1))`). Comment.

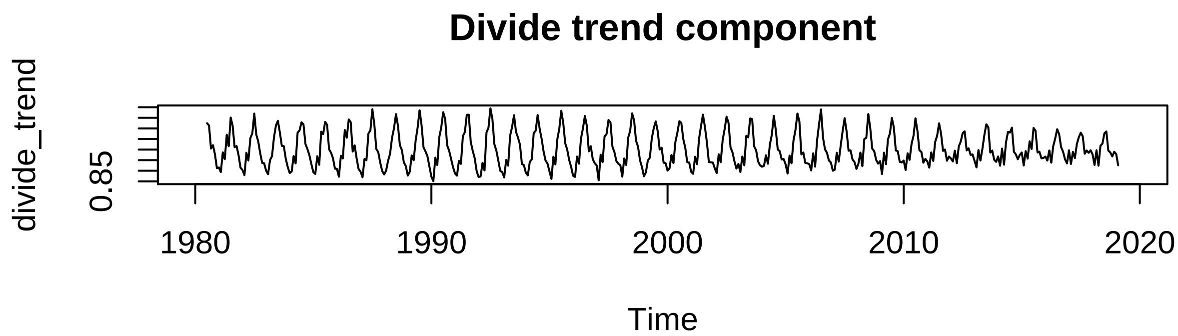
```
substract_seasonal = aus_elec - decompose_demand$seasonal
par(mfrow=c(2,1))
plot(substract_seasonal, main = 'Substract seasonal component')
plot(aus_elec, main = 'Original graph')
```

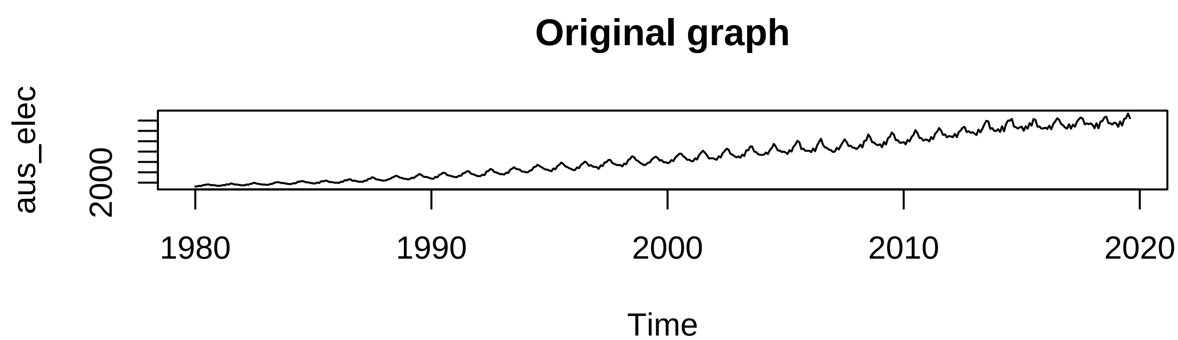




(d). Divide the trend component from the original process and plot it and the original graph in the same window.(You can use `par(mfrow=c(2,1))`). Comment.

```
divide_trend = aus_elec / decompose_demand$trend
par(mfrow=c(2,1))
plot(divide_trend, main = 'Divide trend component')
plot(aus_elec, main = 'Original graph')
```





(e). Is one of these methods (b) or (c) not entirely correct? If so, which one? (Hint: Substracting or deviding depends on whether the series is "Additive" or "Multiplicative")

Since the change in magnitude increases as year pass, we use multiplicative method when compose the deposition. So the method is part c is not entirely corret. We substract when the series is "Addictive", which it is not in this case.