Analysis of data sets

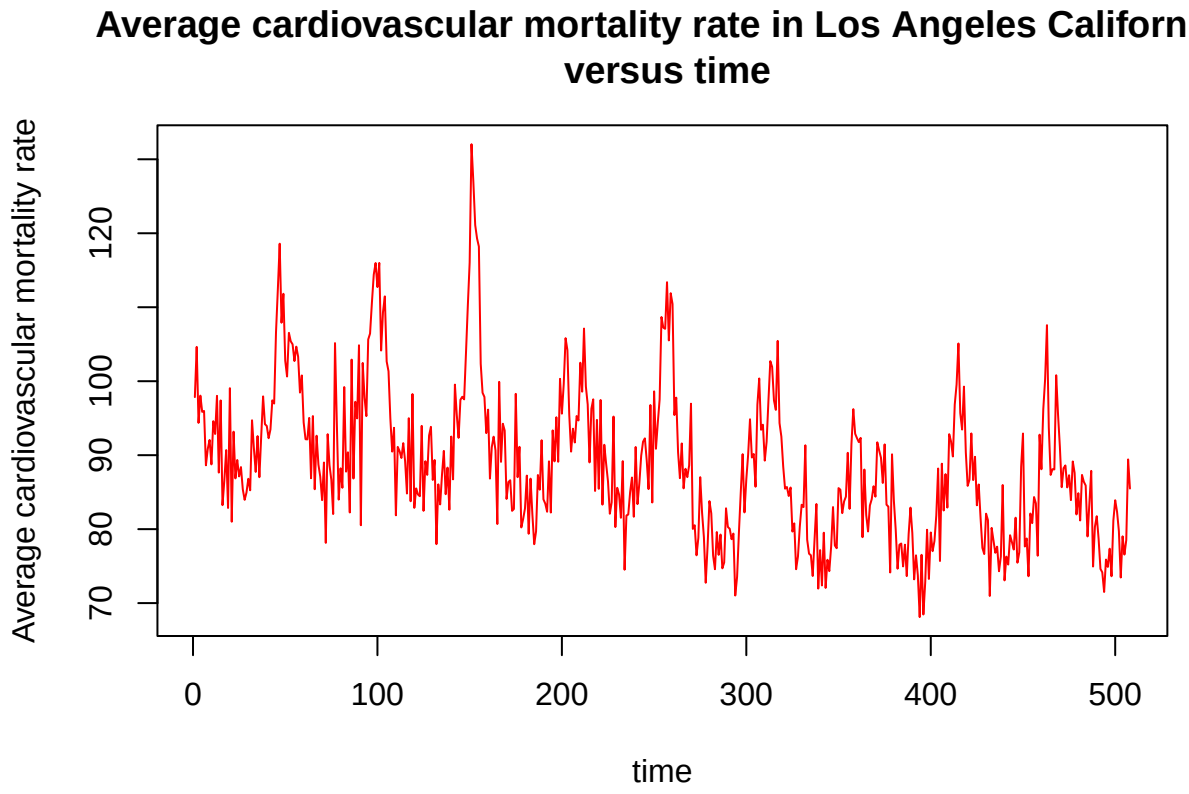
1.

(a)

There are 3 lines contain bunch of "1"s and some comments in the data set. I delete these 3 lines because i think they are meaningless.

(b)

```
cmort = read.table("cmort.dat")
plot(cmort$V1, col = 'red', type = 'l',
     xlab = 'time', ylab = 'Average cardiovascular mortality rate',
     main = 'Average cardiovascular mortality rate in Los Angeles California \n versus time')
```



(c)

i trend

There is a slightly downward trend. The average cardiovascular mortality rate goes up a lot when time is around 150.

ii periodicity and seasonality

The data seems periodic but it does not appear to be seasonality.

iii heteroskedasticity

The variability seems slightly decreasing.

iv dependence

There might be some negative dependency.

v outliers, missing data, etc.

There is no obvious outliers, and missing data.

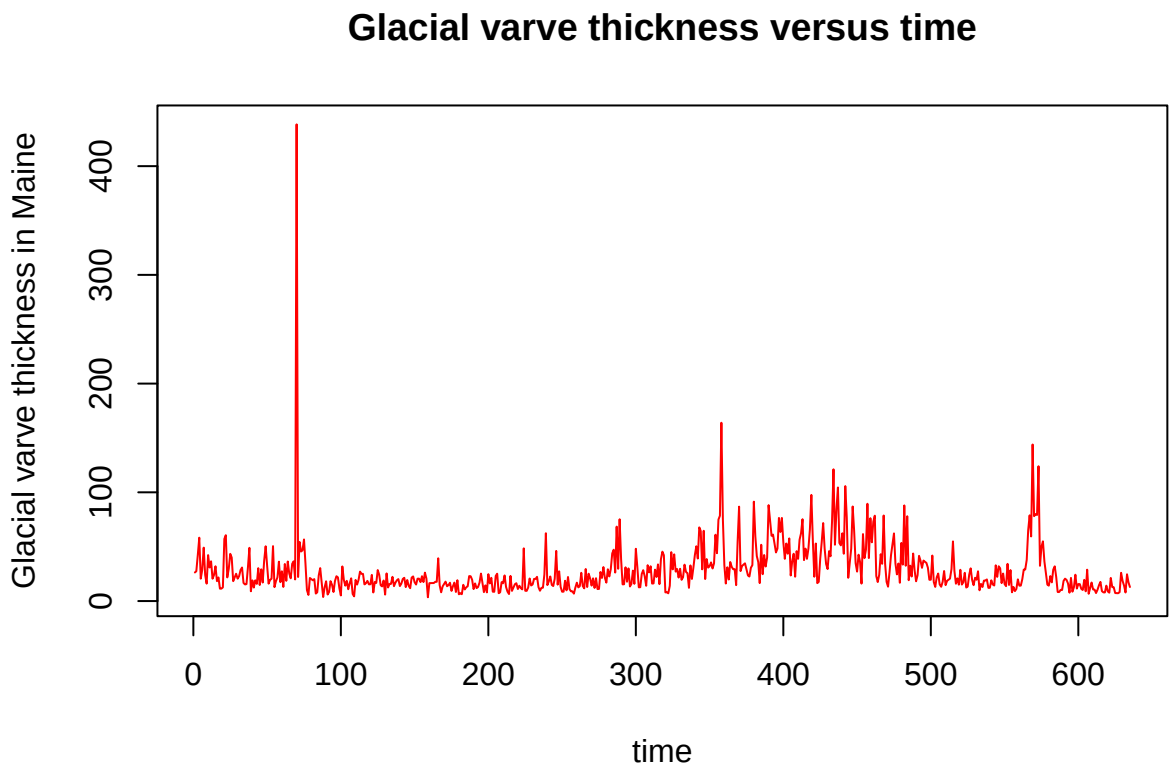
2.

(a)

There is nothing strange in this data file, so i did not make any adjustment.

(b)

```
varve = read.table('varve.dat')
plot(varve$V1, col = 'red', type = 'l',
     xlab = 'time', ylab = 'Glacial varve thickness in Maine',
     main = 'Glacial varve thickness versus time')
```



(c)

i trend

There is no obvious trend. The Glacial varve thickness goes up extremely when time is around 80. This might be a outlier.

ii periodicity and seasonality

There is no obvious periodicity and seasonality.

iii heteroskedasticity

The variablity seems increasing after time is aournd 350.

iv dependence

There might be some negative dependency.

v outliers, missing data, etc.

data point around time is 80 might be a outlier which need further investigation.

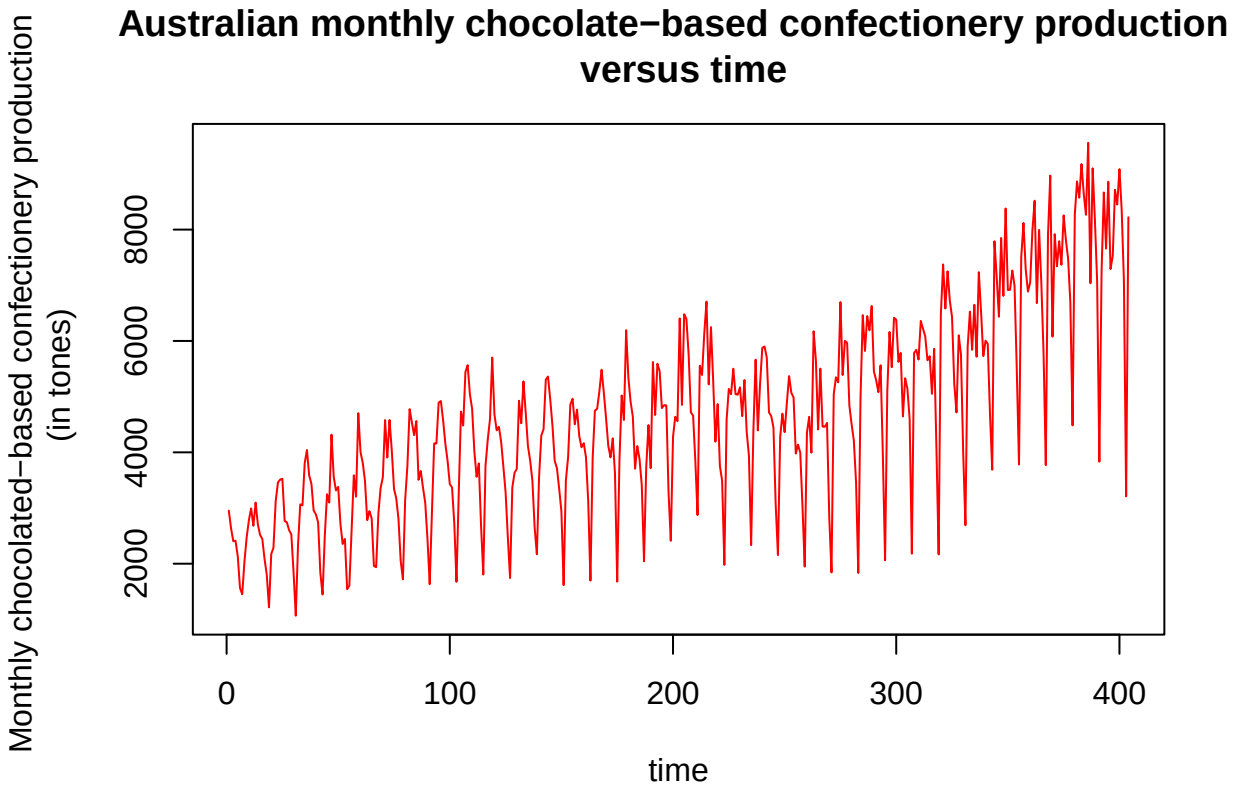
3.

(a)

There are 4 dates in the end of data set, and keep them will affect how R read our data, so i delete these 4 dates.

(b)

```
choc = read.table('choc.dat')
par(mar = c(5.1,5.1,4.1,2.1))
plot(choc$V1, col = 'red', type = 'l',
     xlab = 'time',
     ylab = 'Monthly chocolated-based confectionery production \n (in tones)',
     main = 'Australian monthly chocolate-based confectionery production \n versus time')
```



(c)

i trend

There is a slightly upward trend with structural break aournd time is 150 and 250.

ii periodicity and seasonality

There is a seasonality on a 12 month interval.

iii heteroskedasticity

The variability seems increasing.

iv dependence

There is an oscillatory pattern. There might be a negative dependency.

v outliers, missing data, etc.

There is no obvious outliers, and missing data.