STAT443_Assignment2_Yuting_Wen

yuting wen

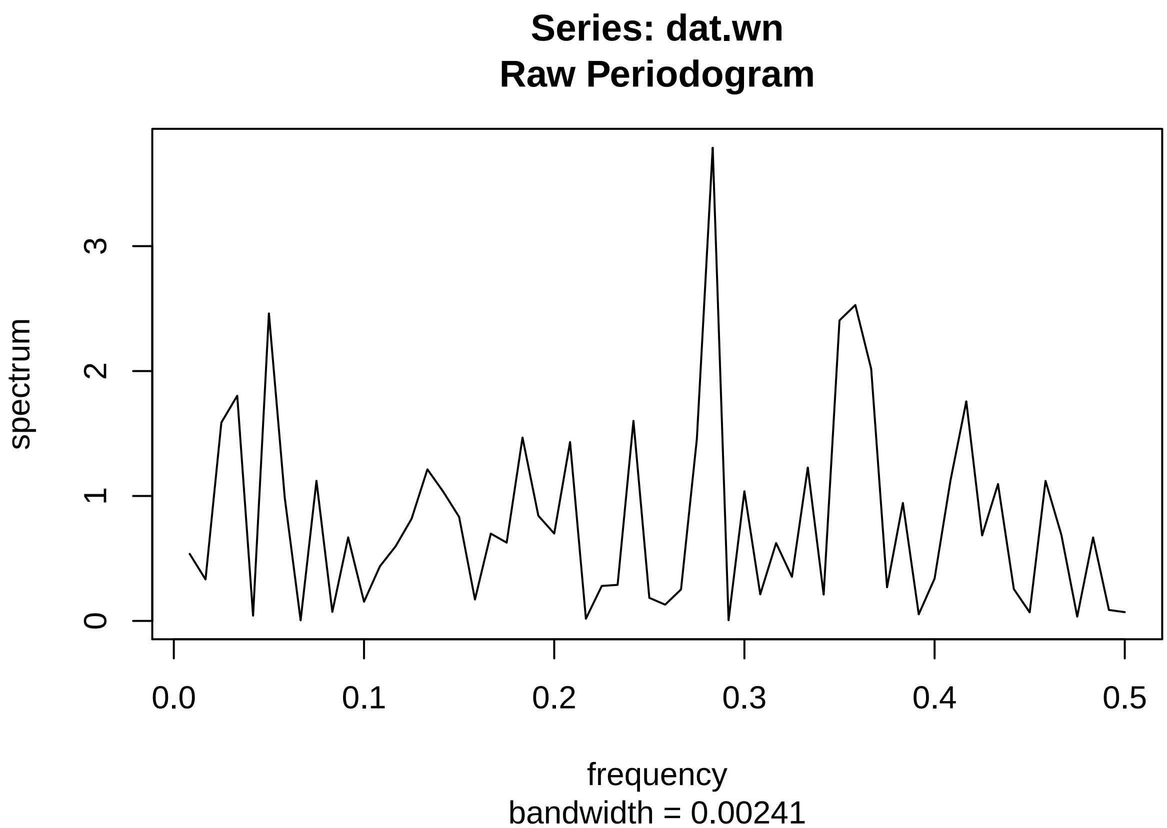
3/20/2020

Question 1

(a) raw periodogram; cumulative periodogram; perform the above test; stating clearly your test statistic and conclusion.

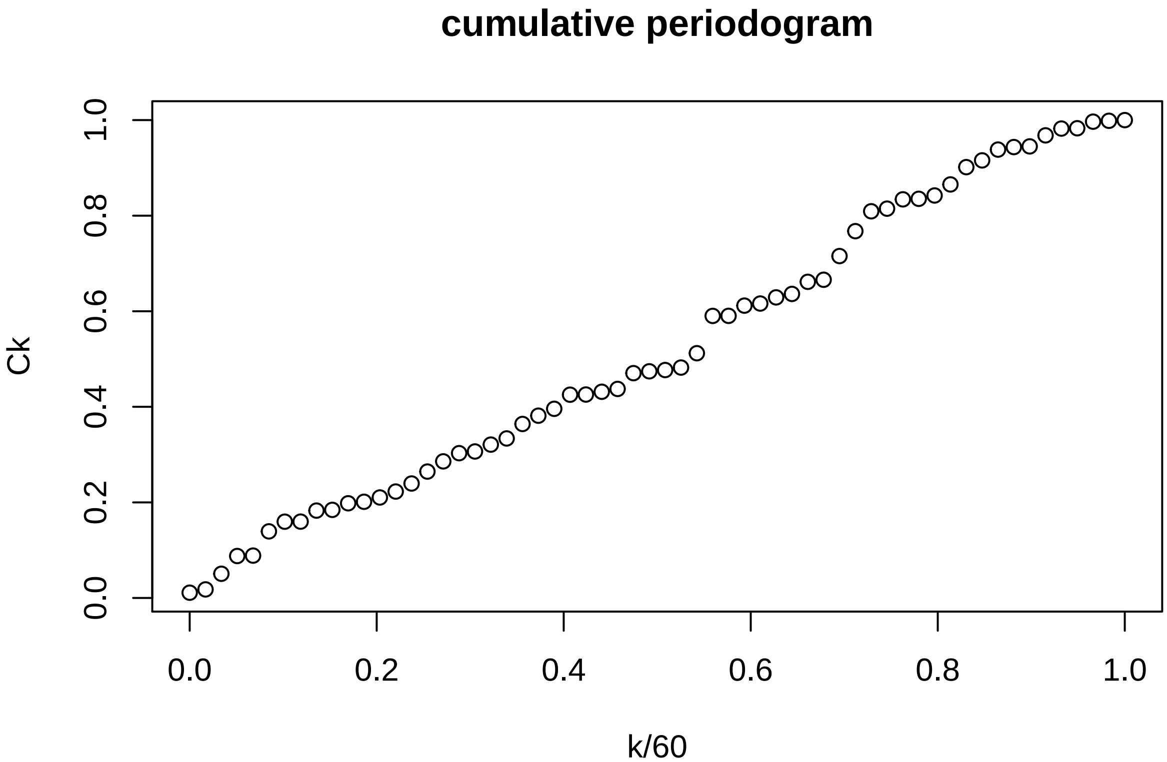
```
set.seed(123)
N <- 120
M <- 60
D.c <- 1.358/sqrt(59)

dat.wn <- rnorm(N,mean=0,sd=1)
# raw periodogram
P <- spec.pgram(dat.wn,log='no')
```




```
P.spec <- P$spec
P.k <- cumsum(P.spec)
P.m <- P.k[M]
C.k <- P.k/P.m
for (i in 1:M) {
  C.k[i] = P.k[i]/P.m
}
```

```r
# cumulative periodogram
plot(y=C.k,x=seq(0,1,length.out = M),xlab='k/60',ylab='Ck',main="cumulative periodogram")
```




```r
# test stat
D <- 0
for (i in 1:M){
  cur <- abs(C.k[i]-(i/M))
  if (cur > D){
    D <- cur
  }
}
show(D)
```

```
## [1] 0.07582877
```

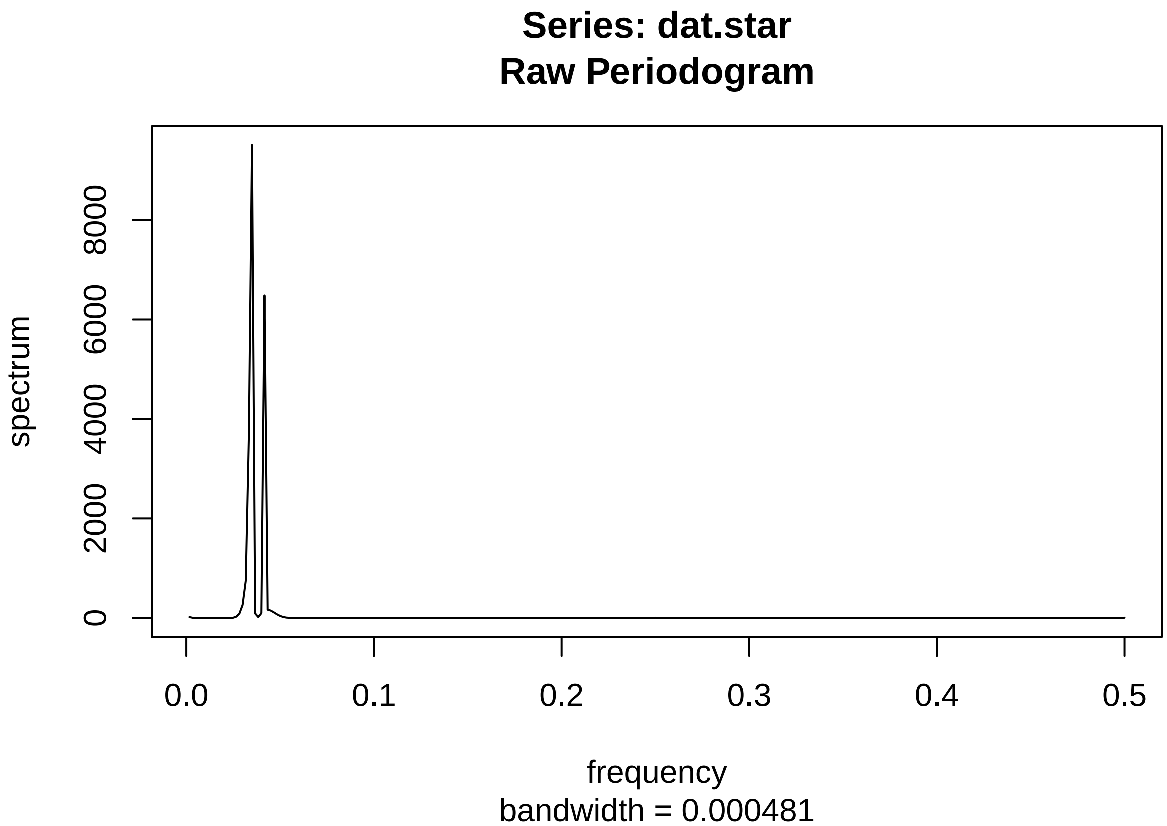
The test stat I got is 0.0758.

Comparing with Dc (0.1768), Clearly the test statistic is non—significant at the 5% level, and there is no reason to reject the (as we know that is true) null hypothesis that the data arise from a white noise process.

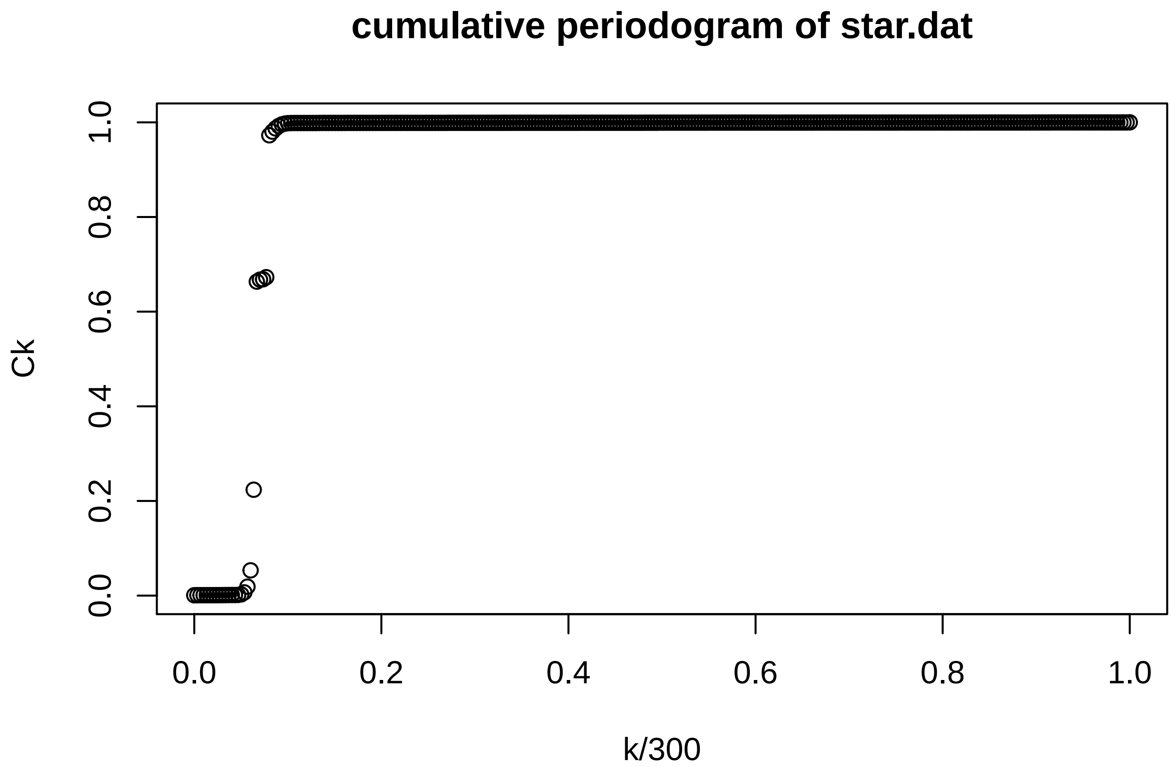
(b) raw periodogram; cumulative periodogram;

```r
dat.star <- read.table("~/Desktop/STAT443/assignments/star.dat")

# raw periodogram
M <- 300
P <- spec.pgram(dat.star,log='no')
```




```r
# cumulative periodogram
P.spec <- P$spec
P.k <- cumsum(P.spec)
P.m <- P.k[M]
C.k <- P.k/P.m
for (i in 1:M) {
  C.k[i] = P.k[i]/P.m
}
plot(y=C.k,x=seq(0,1,length.out = M),xlab='k/300',ylab='Ck',main="cumulative periodogram of star.dat")
```


cumulative periodogram of star.dat
Ck
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
k/300

(c) Perform the test; stating clearly your test statistic and conclusion

```
D <- 0
for (i in 1:M){
  cur <- abs(C.k[i]-(i/M))
  if (cur > D){
    D <- cur
  }
}
show(D)
```

```
## [1] 0.8989481
```

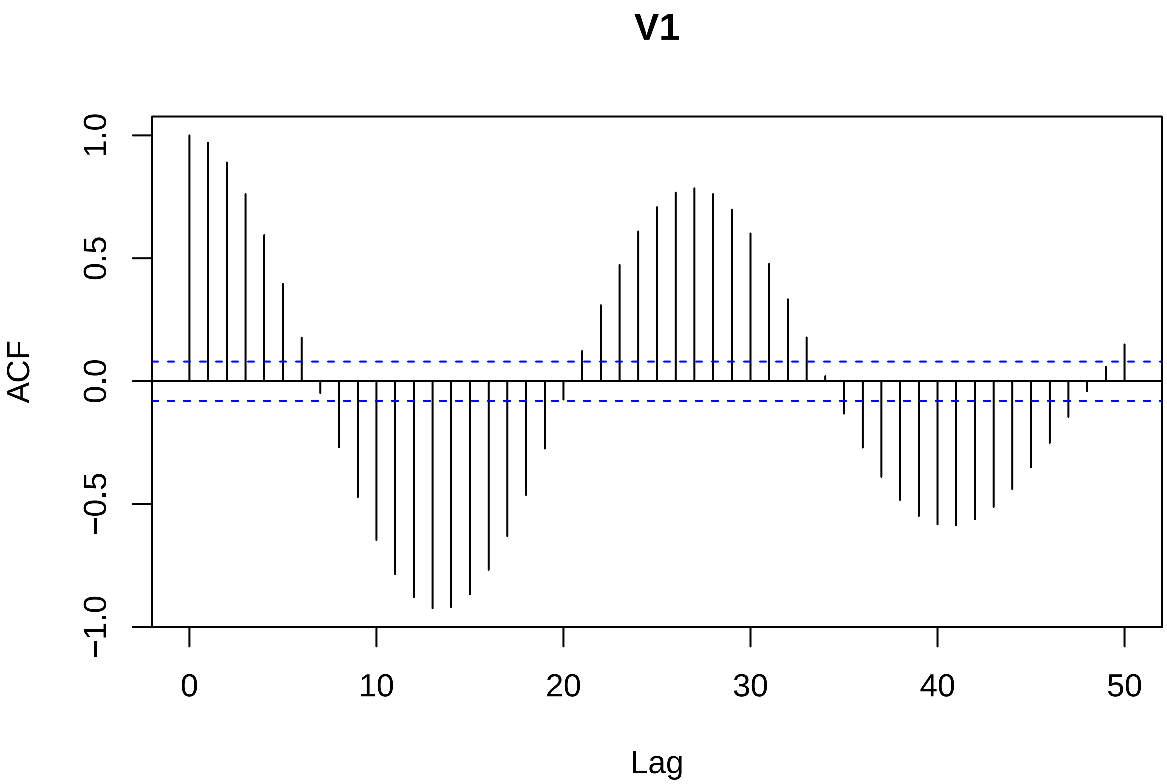
The test stat I got is 0.8989.

Comparing with Dc (0.1768), Clearly the test statistic is significant at the 5% level, and there is evidence to reject the null hypothesis that the data arise from a white noise process.

In conclusion, we say the data given is not from a white noise process.

(d) Plot the acf up to lag 50; approximate what appears to be the wavelength of the main periodic component in the series; estimate what may be the important contributing frequency to the spectrum of the series.

```
acf(dat.star,lag.max = 50)
```



The wavelength would be 28 approximately based on acf.

This also means that the important contributing frequency would be 2*pi/28 = 0.224.

(e) Compare your estimate in (d) with the largest component from the periodogram.

```
ans <- 2*pi*P$freq[which.max(P$spec)]
show(ans)
```

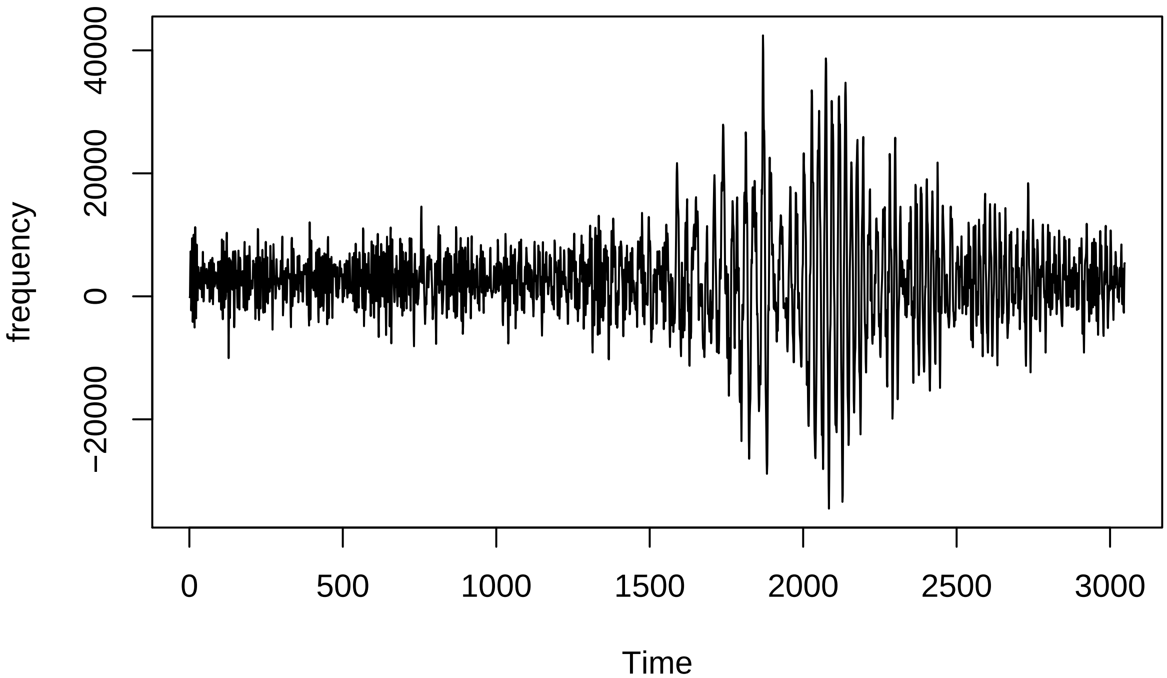
```
## [1] 0.2199115
```

the largest component from the periodogram is 0.220, which is super similar to our estimate in (d).

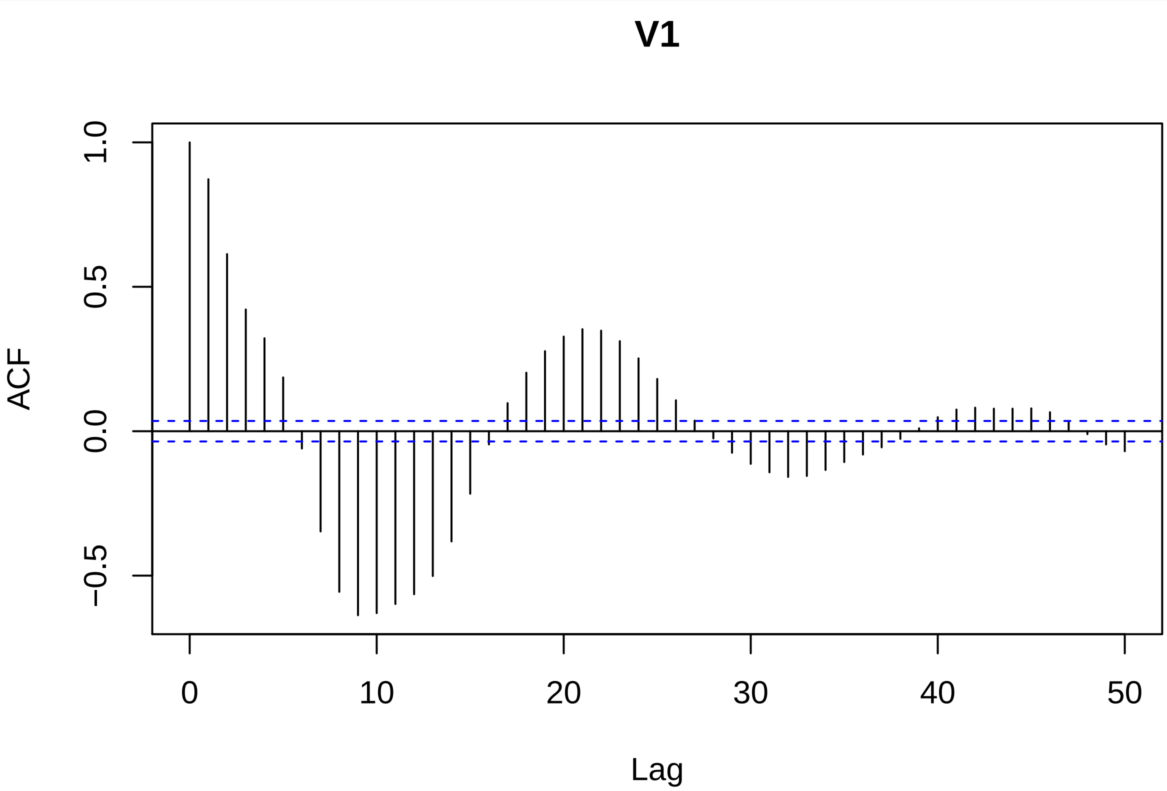
Question 2

(a) Plot the data. Plot the acf of the series up to lag 50, and comment on what you observe.

```
dat.kobe <- read.table("~/Desktop/STAT443/assignments/Kobe.csv")
dat.kobe <- as.ts(dat.kobe)
plot(dat.kobe, ylab='frequency')
```




```
acf(dat.kobe,lag.max=50)
```



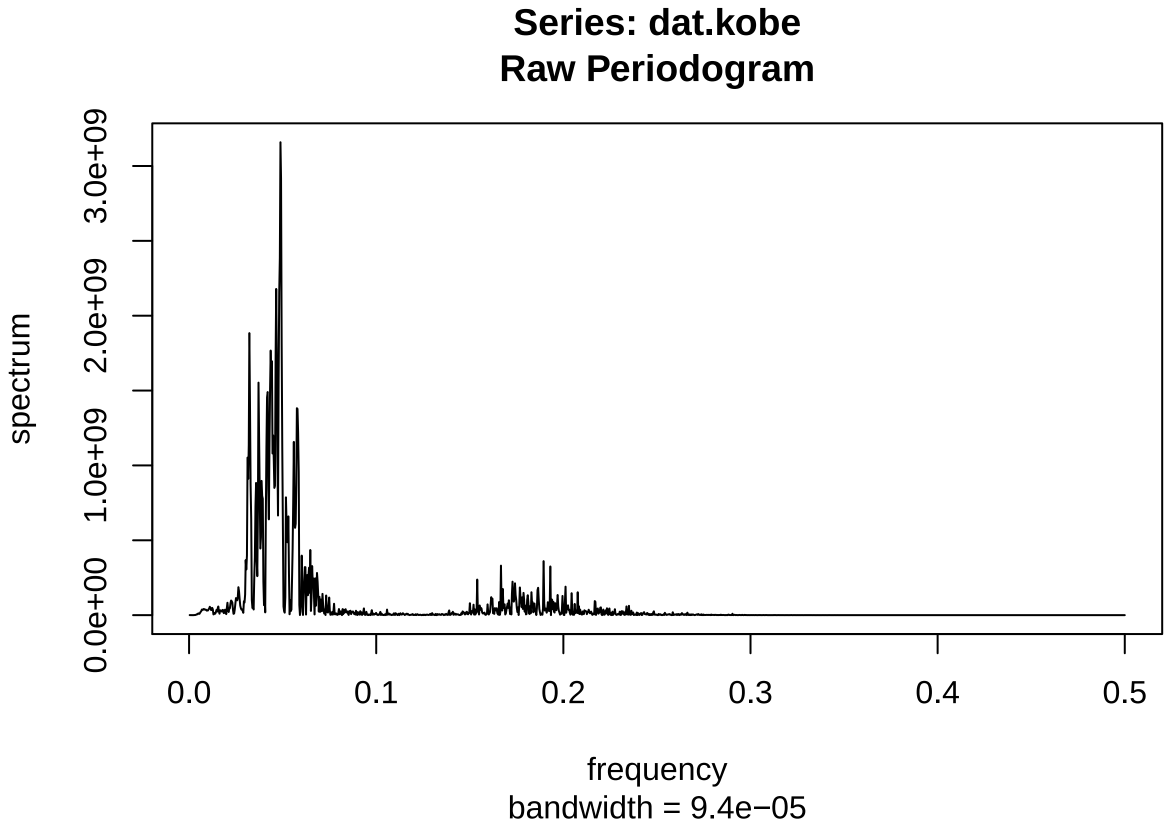
Observation:

plot: it frequently oscillates around 0; amplitude of the frequency starts to increase at around 1600, peak at around 1800 and 2100, then decreases again;

acf: it shows a damped cosine wave, with wavelength approximately 21.

(b) raw periodogram; Comment on what you observe.

```
P.kobe <- spec.pgram(dat.kobe,log='no')
```



Observation:

raw periodogram: it shows a strong peak at a low frequency between 0 and 0.1, and cuts off to 0 qucikly; then it shows another peak at frequency around 0.2 which is much lower than first peak, then cuts off to 0 again.

(c) Use the raw periodogram to estimate the wavelength of the most important cyclical component in the time series. Comment on your findings in relation to the acf plot of the series.

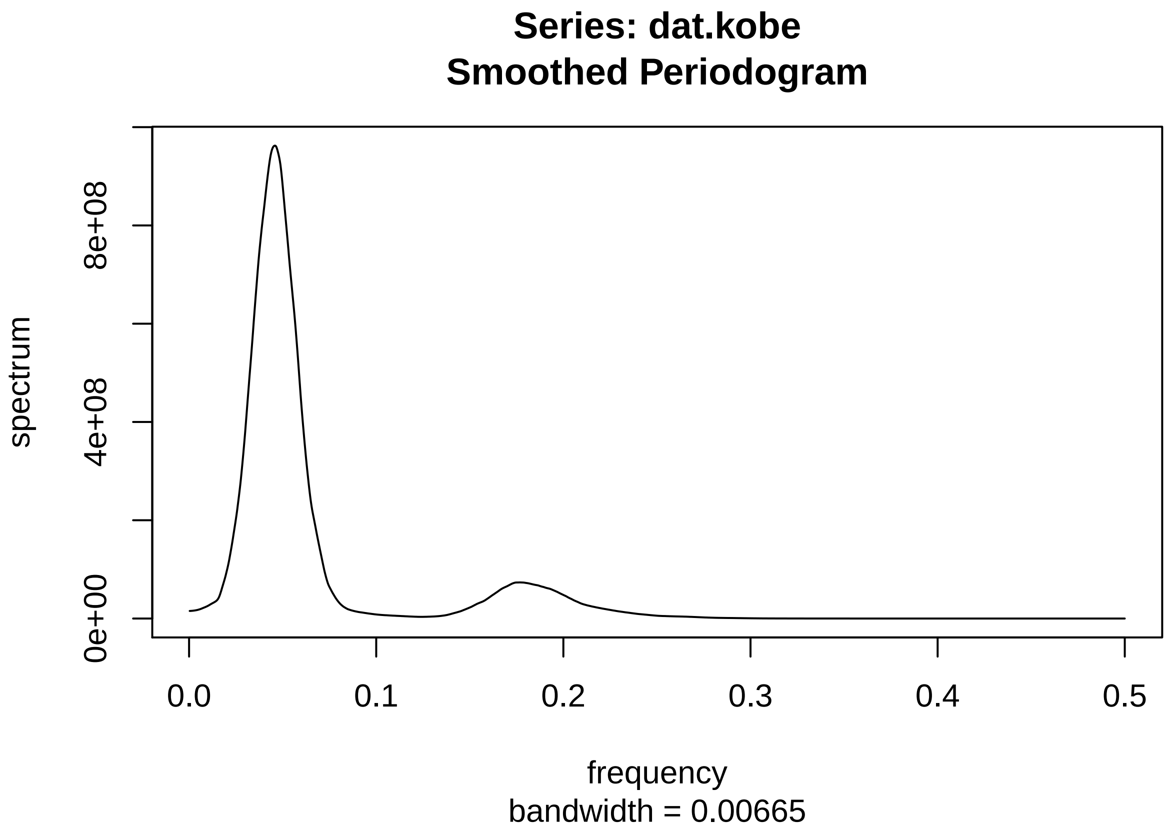
```
wl.est <- 1/P.kobe$freq[which.max(P.kobe$spec)]
show(wl.est)
```

```
## [1] 20.48
```

The wavelength of the most important cyclical component in the time series we estimated based on the raw periodogram is 20.48, which is similar to what we found based on acf (21).

(d) smooth the raw periodogram; estimate the amplitude of the most important frequency contributing to the spectrum of the data.

```
P.kobe.s <- spec.pgram(dat.kobe, log="no", span=c(50,50))
```




```
amp.max <- max(P.kobe.s$spec)/pi
amp.max <- sqrt(amp.max*4*pi/length(dat.kobe))
show(amp.max)
```

```
## [1] 1123.821
```

the amplitude of the most important frequency contributing to the spectrum of the data I estimated is 1123.821.

(e) State clearly any reservations you may have regarding your estimate in (d).

As the span tried increases, the estimated amplitude decreases. Also as the span increases, the low frequency part would have more effect and be higher compared to high frequency part, when I tried span(500,500), the graph just showed a clear decreasing trend. Also the series looks non-stationary, while we assume stationarity to apply to the spectrum and periodogram, which is also limitation.