

1.

A. Q: are there any missing values?

Answer: no

Code and corresponding output: anyNA(df)

```
> anyNA(df)
[1] FALSE
```

B. Q: how many time series variable are in the data set?

Answer: there is only one time series variable in the data set.

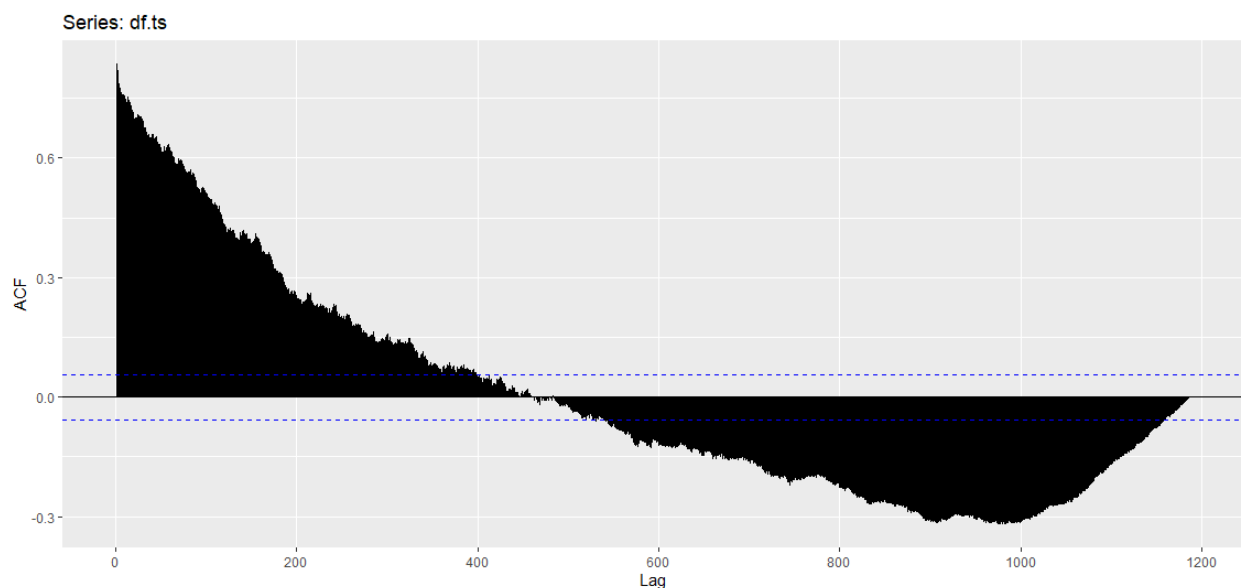
Code and output: str(df)

```
> str(df)
'data.frame': 1188 obs. of 3 variables:
 $ Year      : int  1921 1921 1921 1921 1921 1921 1921 1921 1921 1921 ...
 $ Month     : int   1  2  3  4  5  6  7  8  9 10 ...
 $ Monthly.MRSL.in.meters: num -0.092 -0.134 -0.22 -0.185 -0.119 -0.16 -0.164 -0.222 -0.186 -0.207 ...
```

C. Q: provide the correlogram of the ts object, with 1200 lags. Is there any trend in the data? Why or why not?

Answer: There is a trend in the data, as there were 3 or more consecutive instances on the correlogram where there is a positive correlation.

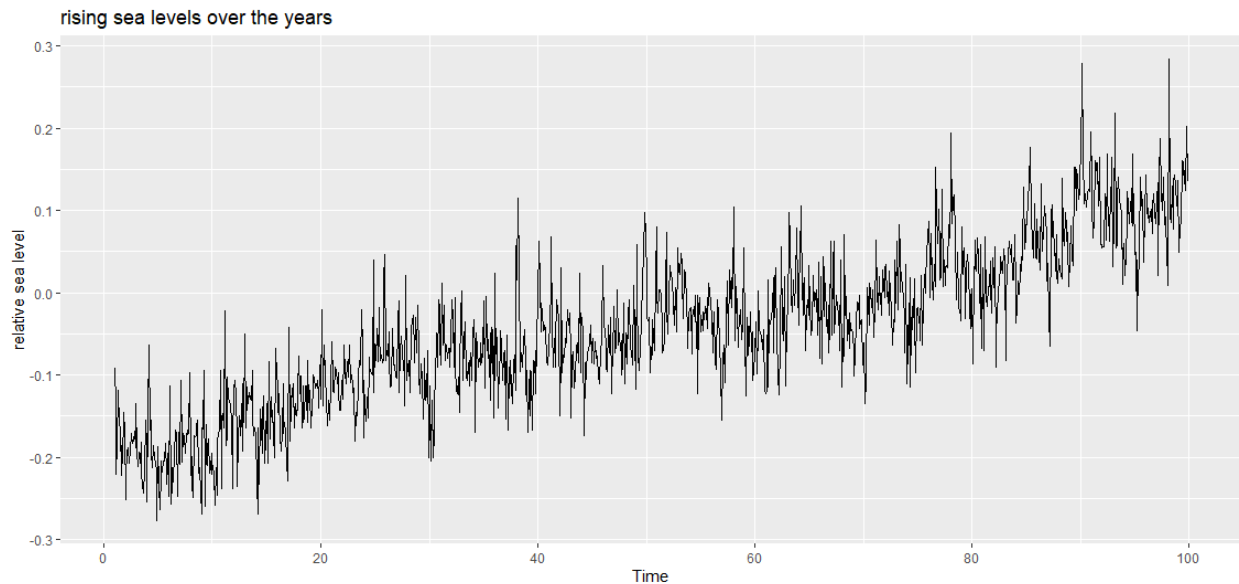
Code and output: ggAcf(df.ts,lag=1200)



D: Q: Define a ts object for the time series data, with frequency of 12. Provide the time series plot of the data. Is there any trend? Is it an upward or downward trend?

A: there is an upward trend in the data as there is an increase in the values in the same seasonal position. So month 13th has a larger value than month 1, and so on and so forth.

Code and Output: autoplot(df.ts)+ylab('relative sea level')+ggtitle('rising sea levels over the years')



E: Q: paste the fully functional, complete R code for this problem.

A: # homework 2

#pre work

rm(list=ls()) #clear workspace

gc()#garbage collections/ clear occupied memeory

cat("\f") #clear console

#import the data and library

library('fpp2')

setwd('C:/Users/xtorres1/Desktop/R/Datasets/Time Series/session 7/')

df=read.csv('Sealevel.csv')

are there any missing values:

anyNA(df)

how many time serieres variables are in the data set: there is only one variable in hte data set

View(df)

```

str(df)

df$Year=NULL

df$Month=NULL

df.ts=ts(df,start = c(1,1),frequency = 12)

ggAcf(df.ts,lag=1200)

autoplot(df.ts)+ylab('relative sea level')+ggtitle('rising sea levels over the years')

```

Code or output: code posted above, output not available.

2.

A: Q:is there any missing values?

A:there are no missing values

Code and output: anyNA(dj)

```

> anyNA(dj)
[1] FALSE

```

B: Q: how many time series variables are there?

A: since the time series object of jd came predefined there is only one time series element which is the trading day end values.

Code and output: >dj

```

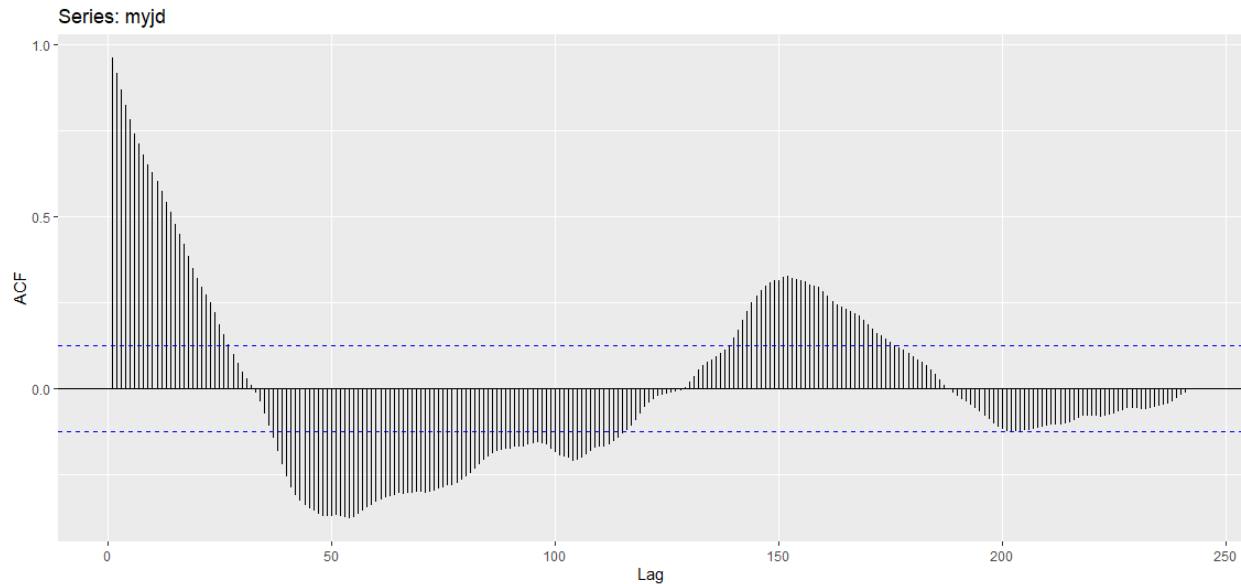
Time Series:
Start = 1
End = 292
Frequency = 1
 [1] 3651.0 3645.0 3626.0 3634.0 3620.5 3607.0 3589.0 3590.0 3622.0 3634.0 3616.0
 [12] 3634.0 3631.0 3613.0 3576.0 3537.0 3547.0 3540.0 3543.0 3568.0 3566.0 3566.0
 [23] 3555.0 3581.0 3578.0 3587.0 3599.0 3584.0 3585.0 3593.0 3593.0 3603.0 3622.0
 [34] 3630.0 3642.0 3635.0 3645.0 3636.0 3649.0 3674.0 3672.0 3665.0 3688.0 3681.0
 [45] 3693.0 3698.0 3662.0 3625.0 3643.0 3648.0 3640.0 3664.0 3662.0 3684.0 3678.0
 [56] 3711.0 3704.0 3685.0 3694.0 3670.0 3674.0 3688.0 3686.0 3684.0 3678.0 3684.0
 [67] 3697.0 3702.0 3704.0 3710.0 3719.0 3734.0 3730.0 3741.0 3764.0 3743.0 3717.0
 [78] 3726.0 3752.0 3755.0 3745.0 3762.0 3758.0 3776.0 3794.0 3776.0 3757.0 3784.0
 [89] 3799.0 3804.0 3821.0 3866.0 3850.0 3849.0 3842.0 3867.0 3870.0 3870.0 3884.0
[100] 3892.0 3914.0 3913.0 3895.0 3926.0 3945.0 3978.0 3964.0 3976.0 3968.0 3871.0
[111] 3906.0 3906.0 3932.0 3895.0 3895.0 3904.0 3928.0 3937.0 3923.0 3888.0 3900.0
[122] 3912.0 3892.0 3840.0 3839.0 3832.0 3809.0 3832.0 3824.0 3832.0 3856.0 3852.0
[133] 3853.0 3831.0 3863.0 3863.0 3850.0 3848.0 3865.0 3896.0 3865.0 3863.0 3869.0
[144] 3821.0 3775.0 3762.0 3699.0 3627.0 3636.0 3675.0 3680.0 3693.0 3674.0 3689.0
[155] 3682.0 3662.0 3663.0 3662.0 3620.0 3620.0 3599.0 3653.0 3649.0 3700.0 3684.0

```

C: Q: extract a portion of the ts object by: mydj <-window(dj,start=(50)). Provide the correlogram of the portion, mydj. Identify the time series elements in mydj. Is there any seasonality? Why or why not? If yes, also identify frequency. Also provide evidence on how the ts elements and frequency value are determined.

A: There is no seasonality, as none of the points that are not white noise are not massively more correlated than their neighbors. The TS element of trend is found by finding statistically relevant points that are consecutively positive. Frequency cannot be found as the data is not seasonal.

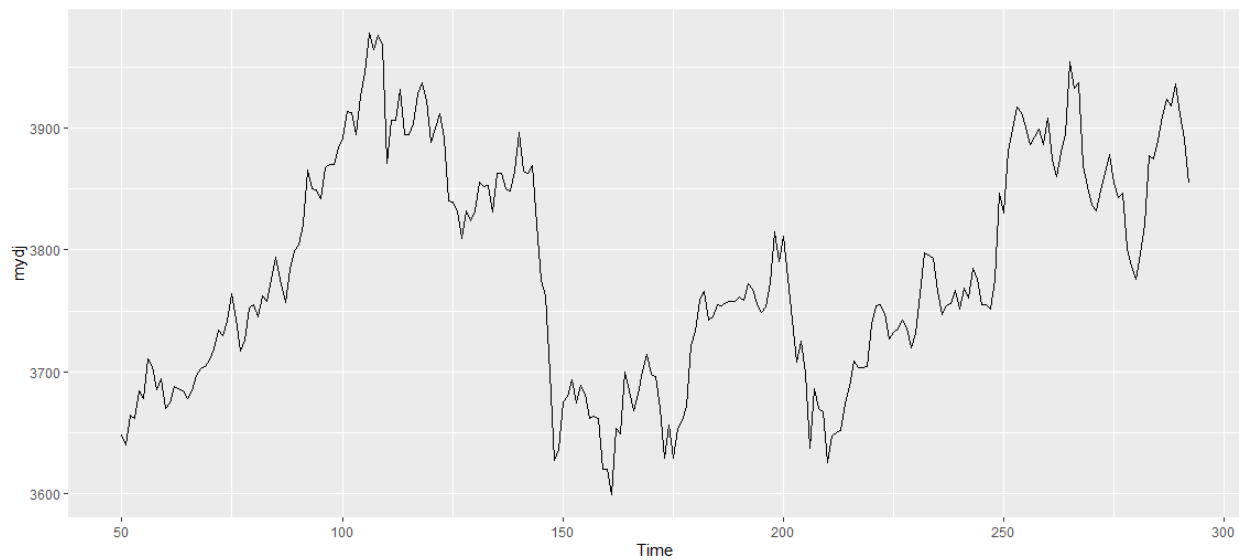
Code and output: `ggAcf(myjd,lag=300)`



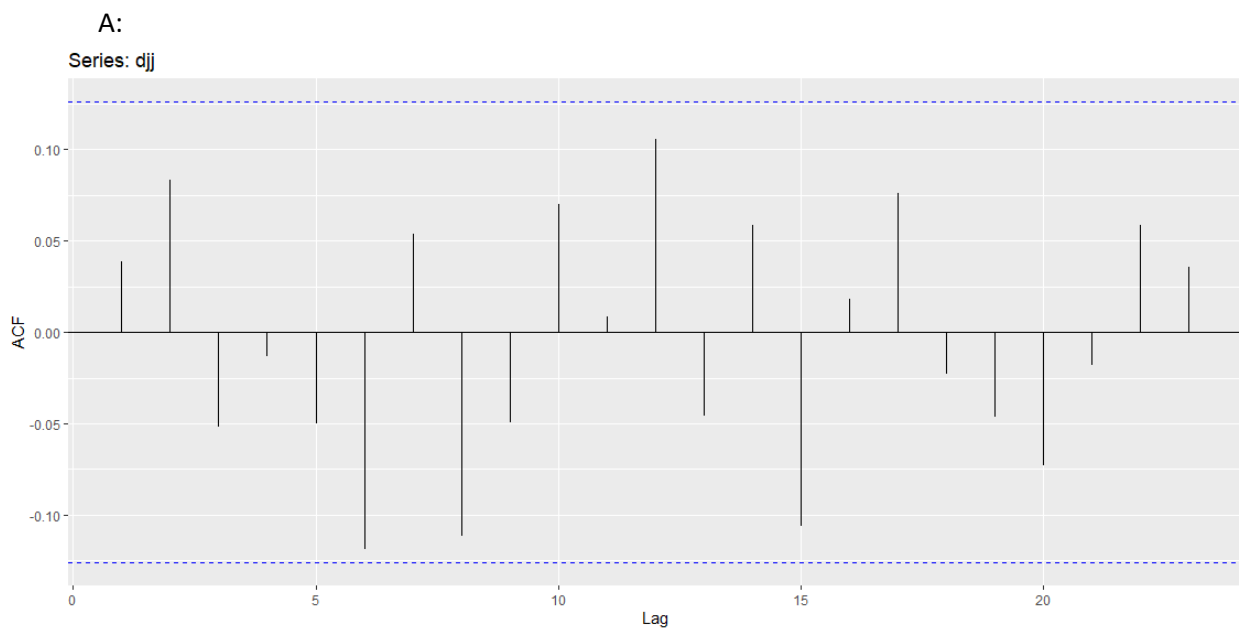
D: Q: Provide the time series plot of the portion, mydj. Identify the time series elements there. Are your answers consistent with the question above?

A: originally there seemed like there was a trend in the data, here it is difficult to tell, there could be a slight trend as the troughs of the later dates are less than in the beginning but it's hard to tell. There are similar patterns here with large peaks and low valleys so it looks like its seasonal. These answers are not consistent with the question above.

Code and output: `autoplot(mydj)`



E: Q: Use `ddj <-diff(mydj)` to compute the daily changes in the portion, `mydj`. Provide the correlogram of `ddj`, and identify the time series elements in `ddj`. Also provide evidence on how ts elements are determined.

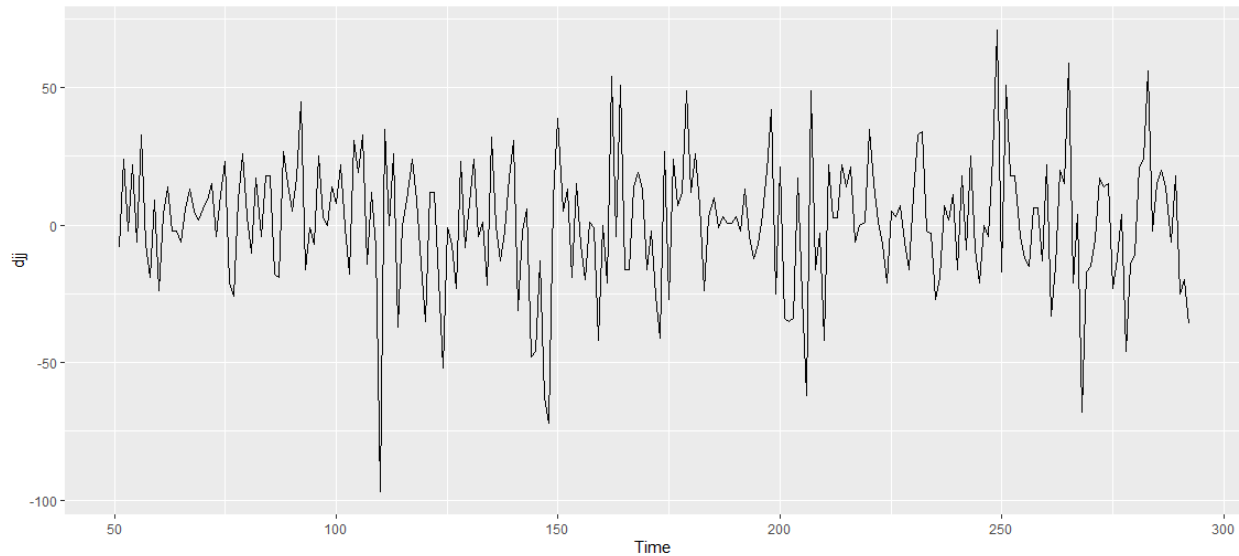


There are no time series elements in `ddj` as it is white noise. None of the points go past our barriers and as such they are not statistically relevant.

Code and output: `ggAcf(ddj)` and output is above in answer section.

F: Q: Provide the time series plot of the daily changes in the portion, `ddj`. Identify the time series elements. Are your answers consistent with those for part 5) above?

A:



there are no time series elements to identify in the autoplot as it is random. These would be consistent with the answer above as it is white noise and as such should not have any patterns.

Code and output: autoplot(djj) and plot above.

G: Q: Paste the fully functional, complete R code for this problem.

A: # homework 2.2

#pre work

rm(list=ls()) #clear workspace

gc()#garbage collections/ clear occupied memeory

cat("\f") #clear console

#import the data and library

library('fpp2')

no need to load the data or set WD as the data comes preloaded on fpp2

dj=dj

anyNA(dj)

how many time series variables are there? there is only one variable in the time series since it is preset therefore

str(dj)

```
View(dj)
```

```
mydj=window(dj,start=50)
```

```
ggAcf(mydj,lag=300)
```

```
autoplot(mydj)
```

```
djj=diff(mydj)
```

```
ggAcf(djj)
```

```
autoplot(djj)
```

Code and output: code above and output not available.