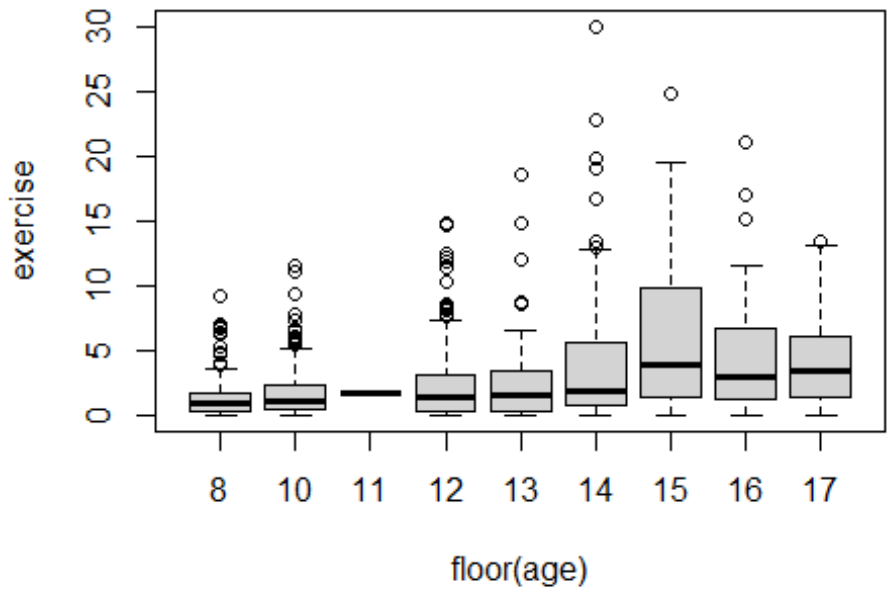
Homework 10
Syracuse University
IST 772
Summer 2021

```
# load packages
require(readr)
require(arules)
require(dplyr)
require(changepoint)
```

Question 2

```
# load the blackmore dataset
Blackmore <- read.csv("dataset-74527.csv")

# boxplot showing the exercise level at different ages
boxplot(exercise ~ floor(age),
        data = Blackmore)
```



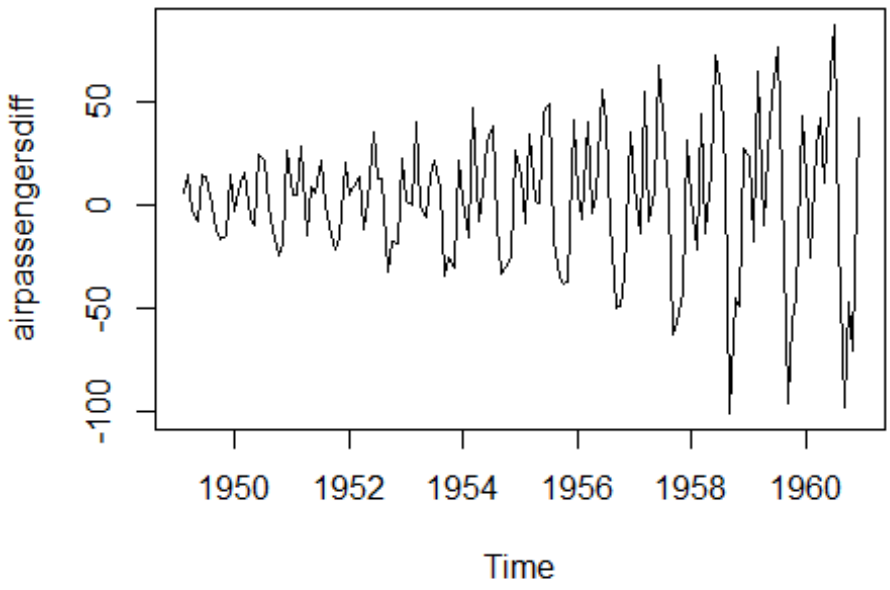

```
# check that the data is balanced
table(floor(Blackmore$age))

##
##   8  10  11  12  13  14  15  16  17
## 231 229   2 191  41 128  45  57  21
```

```r
# run repeated measures ANOVA to compare age 8, 10, 12
aovOut <- aov(exercise ~ floor(age),
              data = Blackmore[which(floor(Blackmore$age) %in% c(8, 10, 12)),
])

# summarize the ANOVA
summary(aovOut)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## floor(age)    1    122  121.97   27.01 2.72e-07 ***
## Residuals   649   2931    4.52
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 5

```r
# use diff to create a differenced data set
airpassengersdiff <- diff(AirPassengers)

# plot the differenced data set
plot(airpassengersdiff)
```

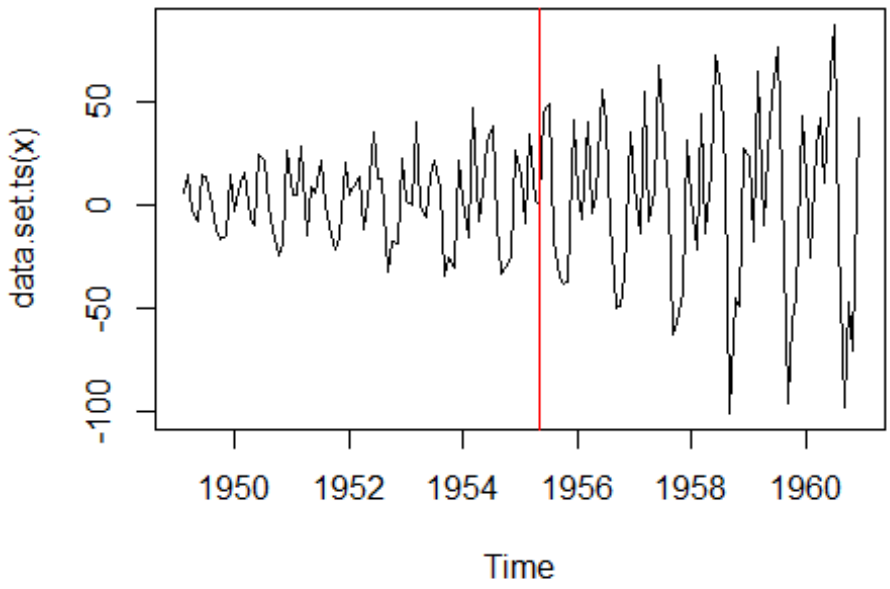



```r
# use cpt.var() to find the change point in the variability
cpt.var(airpassengersdiff)
```

```
## Class 'cpt' : Changepoint Object
##         ~~   : S4 class containing 12 slots with names
##                cpttype date version data.set method test.stat pen.type
```

```
pen.value minseglen cpts ncpts.max param.est
##
## Created on  : Fri Sep 03 15:14:26 2021
##
## summary(.)  :
## ----------
## Created Using changepoint version 2.2.2
## Changepoint type      : Change in variance
## Method of analysis    : AMOC
## Test Statistic  : Normal
## Type of penalty       : MBIC with value, 14.88853
## Minimum Segment Length : 2
## Maximum no. of cpts   : 1
## Changepoint Locations : 76
```

```
# plot the result
plot(cpt.var(airpassengersdiff))
```



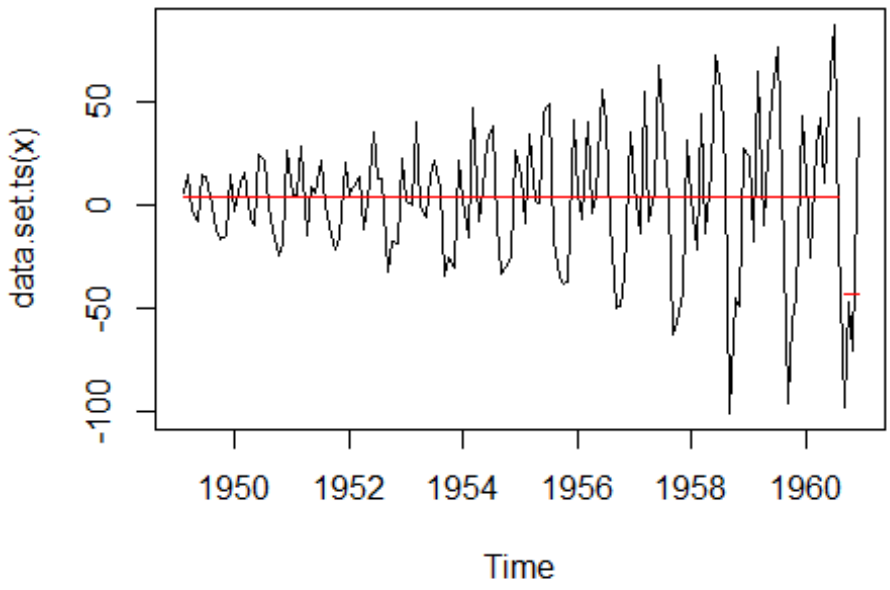

```
# the change point identifies the point in time where there was a change
# in the variance of the y value.
```

Question 6

```
# use cpt.mean() on the airpassengers time series
cpt.mean(airpassengersdiff)
```

```
## Class 'cpt' : Changepoint Object
##         ~~   : S4 class containing 12 slots with names
```

```
##                cpttype date version data.set method test.stat pen.type
pen.value minseglen cpts ncpts.max param.est
##
## Created on  : Fri Sep 03 15:14:26 2021
##
## summary(.)  :
## ----------
## Created Using changepoint version 2.2.2
## Changepoint type      : Change in mean
## Method of analysis    : AMOC
## Test Statistic  : Normal
## Type of penalty       : MBIC with value, 14.88853
## Minimum Segment Length : 1
## Maximum no. of cpts   : 1
## Changepoint Locations : 139

# compare the change point of the mean
plot(cpt.mean(airpassengersdiff))
```




```
# according to the statistical model, the mean of the time series changed
later than
# the variation of the time series.
```

Question 7

```
# The air passengers data set contains monthly airline passenger numbers
# between 1949-1960. We can easily see from looking at the plot that
# the number of passengers is increasing with time. From analyzing the
```

```
# statistics we can see that the change point in variation is at about
# 1955 while the change point in the mean is at about 1961. The reason for
# this is because the seasonal dips start to become much larger, but it takes
# a little bit longer for this to affect the mean than it does to affect
# the variation.
```

Question 8