Lab Assignment 1

The following are the emergency room charges (dollars) made to a sample of 25 patients at two city hospitals.

Hospital A

229.10	202.50	222.20	114.40	205.95
214.30	95.10	213.30	225.50	191.40
201.20	239.80	245.70	213.00	238.80
171.10	222.00	212.50	201.70	184.90
248.30	209.70	233.90	229.80	217.90

Hospital B

199.50	184.00	173.20	186.00	214.10
125.50	133.50	190.40	152.00	165.70
154.70	135.30	154.60	190.30	135.40
167.70	203.40	186.70	155.30	195.90
168.90	166.70	178.60	150.20	232.40

Data Preparation

#Enter Data

```
ER <- data.frame(charges=c(229.10, 202.50, 222.20, 114.40, 205.95,
214.30,95.10, 213.30, 225.50, 191.40, 201.20, 239.80, 245.70, 213.00, 238.80,
171.10, 222.00, 212.50, 201.70, 184.90, 248.30, 209.70, 233.90, 229.80,
217.90, 199.50, 184.00,  173.20, 186.00, 214.10, 125.50, 133.50, 190.40,
152.00, 165.70, 154.70, 135.30, 154.60, 190.30, 135.40, 167.70, 203.40,
186.70, 155.30, 195.90, 168.90, 166.70, 178.60, 150.20, 232.40
                          ),
                 Hospital=c("A", "A", "A", "A", "A","A", "A", "A", "A", "A",
                            "A", "A", "A", "A", "A","A", "A", "A", "A", "A",
                            "A", "A", "A", "A", "A","B", "B", "B", "B", "B",
                            "B", "B", "B", "B", "B","B", "B", "B", "B", "B",
                            "B", "B", "B", "B", "B","B", "B", "B", "B", "B"))
```

The data are coded as follows

```
str(ER)
```

```
## 'data.frame':    50 obs. of  2 variables:
##  $ charges : num  229 202 222 114 206 ...
##  $ Hospital: chr  "A" "A" "A" "A" ...
```

Assignment

From now on, the output from R is not shown below. You are required to run the code indicated and answer the questions. First, compute the mean charge for the two Hospitals together.

```
mean(ER$charges)
```

Question 1: What is the mean charge? Give your answer to 2 decimal places.

Now compute the median charge for the two Hospitals together.

```
median(ER$charges)
```

Question 2: What is the median charge? Give your answer to 2 decimal places.

Next, look at the charges for different hospitals

```
aggregate(charges ~ Hospital, data = ER, FUN = mean)
```

Question 3: What is the mean charge for Hospital A? Give your answer to the nearest dollar.

Question 4: What is the mean charge for Hospital B? Give your answer to the nearest dollar.

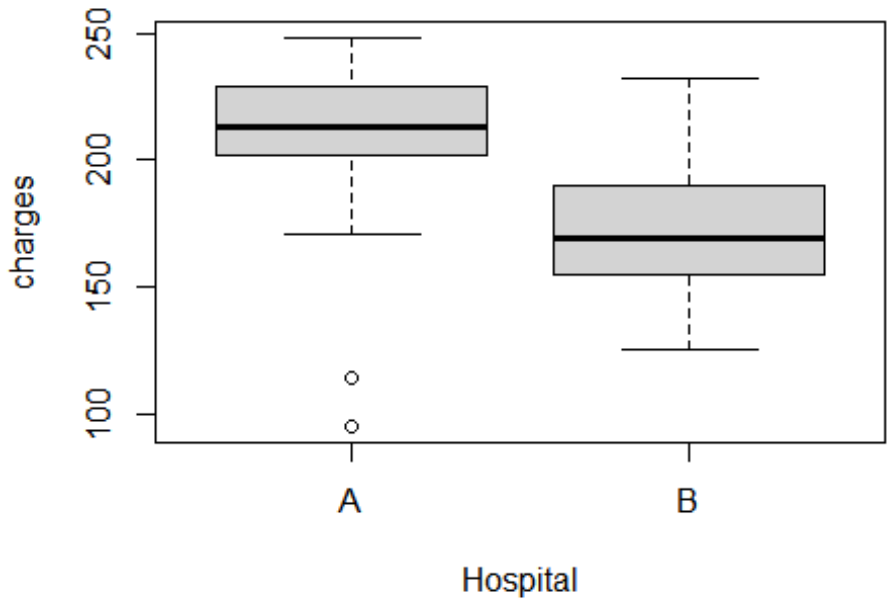
Draw a histogram of the charges data for the two Hospitals together with the density plot superimposed.

```
dens <- density(ER$charges)
hist(ER$charges, main = "Histogram of Charges ", freq = FALSE)
lines(dens)
```

Question 5: Does the distribution look skewed or roughly symmetric?

Produce a boxplot comparing charges of Hospital A and Hospital B

```
boxplot(charges ~ Hospital, data = ER)
```



Question 6: What does the thick black line in the box represent?

Question 7: From the boxplots, are the charges lower or higher on average for Hospital A than Hospital B?

Question 8: From the boxplots, do charges look more or less variable for Hospital B than Hospital A?

Compute the standard deviation of the charges for Hospital A and Hospital B.

```
aggregate(charges ~ Hospital, data = ER, FUN = sd)
```

Question 9: What is the standard deviation of charges for Hospital A? Give your answer to 2 decimal places.

Question 10: What is the standard deviation of charges for Hospital B? Give your answer to 2 decimal places.

Question 11: What units are the standard deviations of charges in?

Next, compute several summary statistics from charges

```
summary(ER$charges)
```

Answer the following questions about the results produced by R:

Question 12: Calculate the range of the charges of two Hospitals combined.

Question 13: What is the value of the lower quartile of the charges in Hospital A and Hospital B combined?

Question 14: What is the value of the upper quartile of the charges in Hospital A and Hospital B combined?