Question 1

1. Finding a 90% confidence interval for the average student IQ in the school

a.

```
## get the mean and standard deviation and set the same size
meany <- mean(y)
sdy <- sd(y)
n <- 25
## set the alpha (a) value as 1-0.9 for a 90% confidence interval
a <- 0.10
## get the t-value
tvalue <- qt(1-a/2, df=n-1)
## t-value is approx 1.71
## get the margin of error
margin_error <- tvalue * (sdy/sqrt(n))
## calculate confidence intervals
lowerbound <- meany - margin_error
## lower bound is approx 93.96
upperbound <- meany + margin_error
## upper bound is approx 102.92

## verified results by using the t-test function in r
confidence_interval <- t.test(y, conf.level = 0.90)
confidence_interval
## get the lower bound
confidence_interval$conf.int[1]
## get the upper bound
confidence_interval$conf.int[2]

## results from both methods are the same. the 90% confidence
## interval is approx (93.96, 102.92)
```

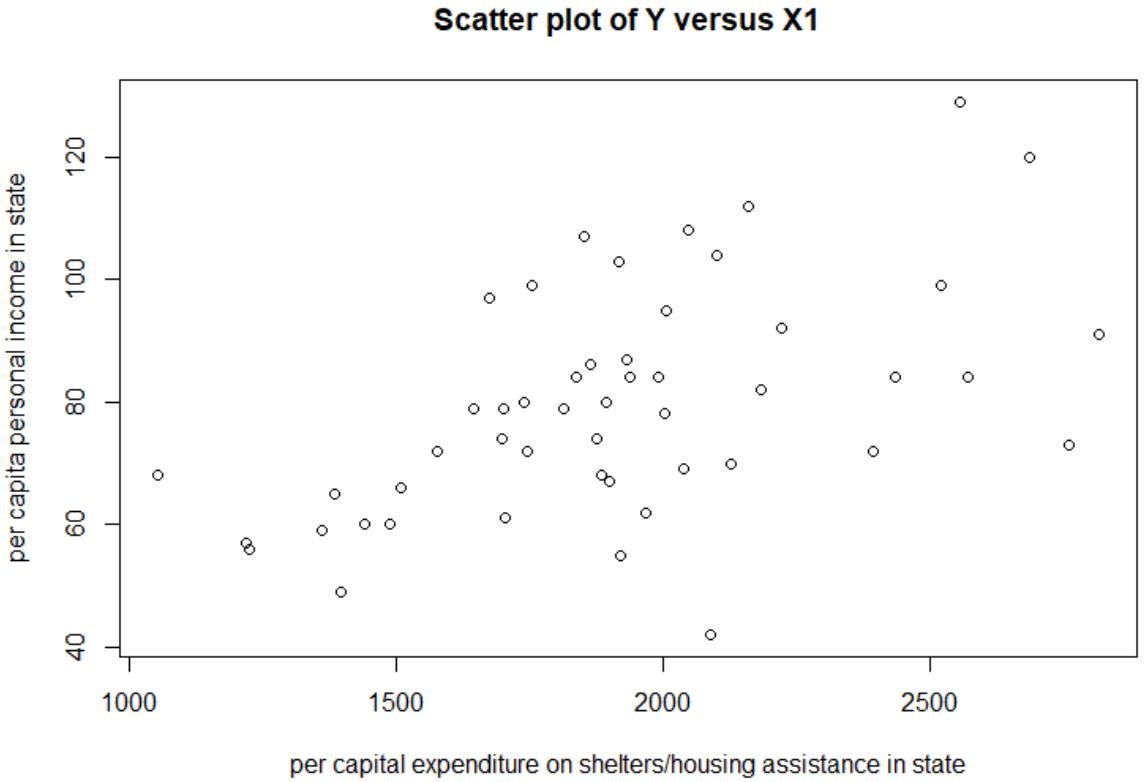
2. Conduction a hypothesis test

a.

```
## because the sample size(n) < 30, i'm going to use a one sample t-test
## in order to do this, i am assuming that the data is normally distributed
## H0 (null hypothesis) is that the mean of y is greater than 100
## setting the null hypothesis mean
null_mean <- 100
hypothesis_test_result <- t.test(y, mu= null_mean, alternative= "greater")
hypothesis_test_result
## because the p-value > 0.05, fail to reject the null hypothesis. in other words,
## there is not enough evidence to say that the class' average IQ is higher than 100 (the national average)
## can also find this out by doing calculations by hand
```

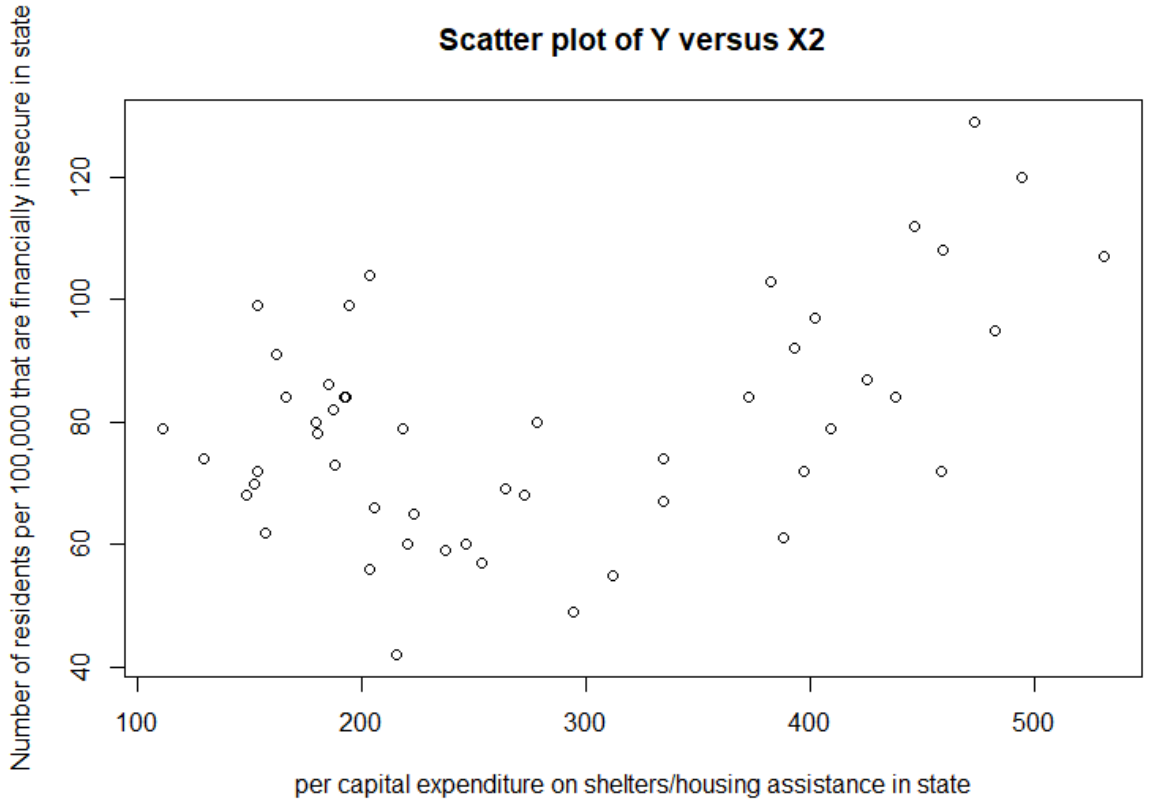
Question 2

Graph of Y and X1



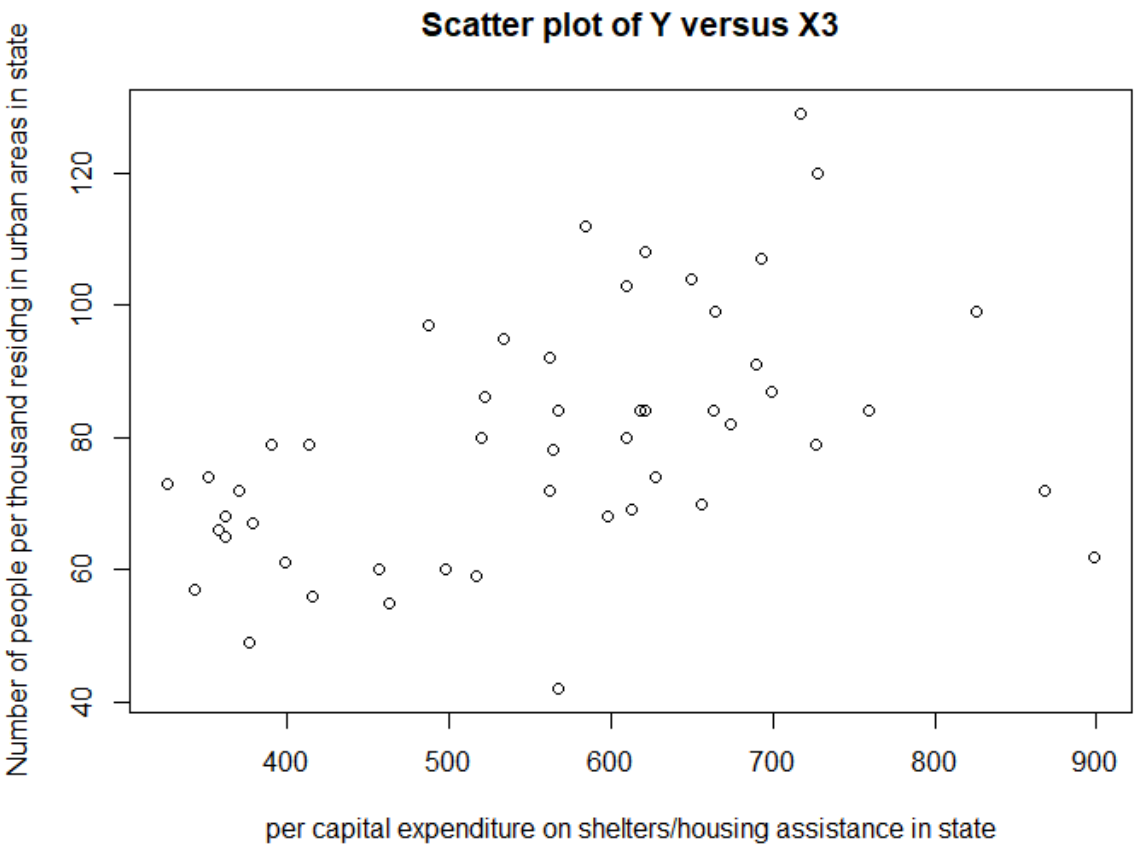
There does not seem to be a strong correlation between per capita expenditure on shelters/housing assistance in state and per capita personal income in state but it seems like there's a slight upward trend in per capita housing assistance assistance as personal income increases but it is more likely a coincidence and would have to be confirmed via additional testing.

Graph of Y and X2



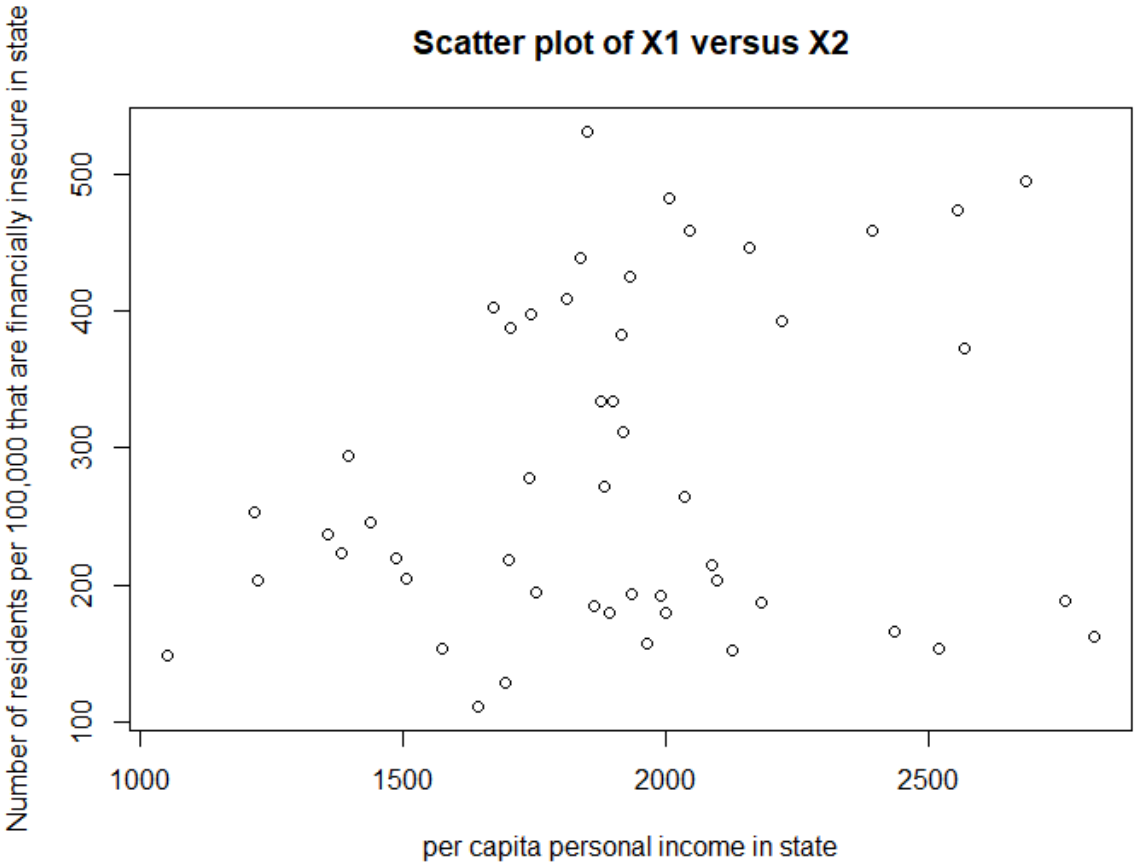
I do not think there is any strong correlation between per capita expenditure on shelters/housing assistance in state and number of residents per 100,000 that are financially insecure in state. If a line of best fit or trend line was inserted into the scatter plot, it could not be linear, which indicates no correlation.

Graph of Y and X3



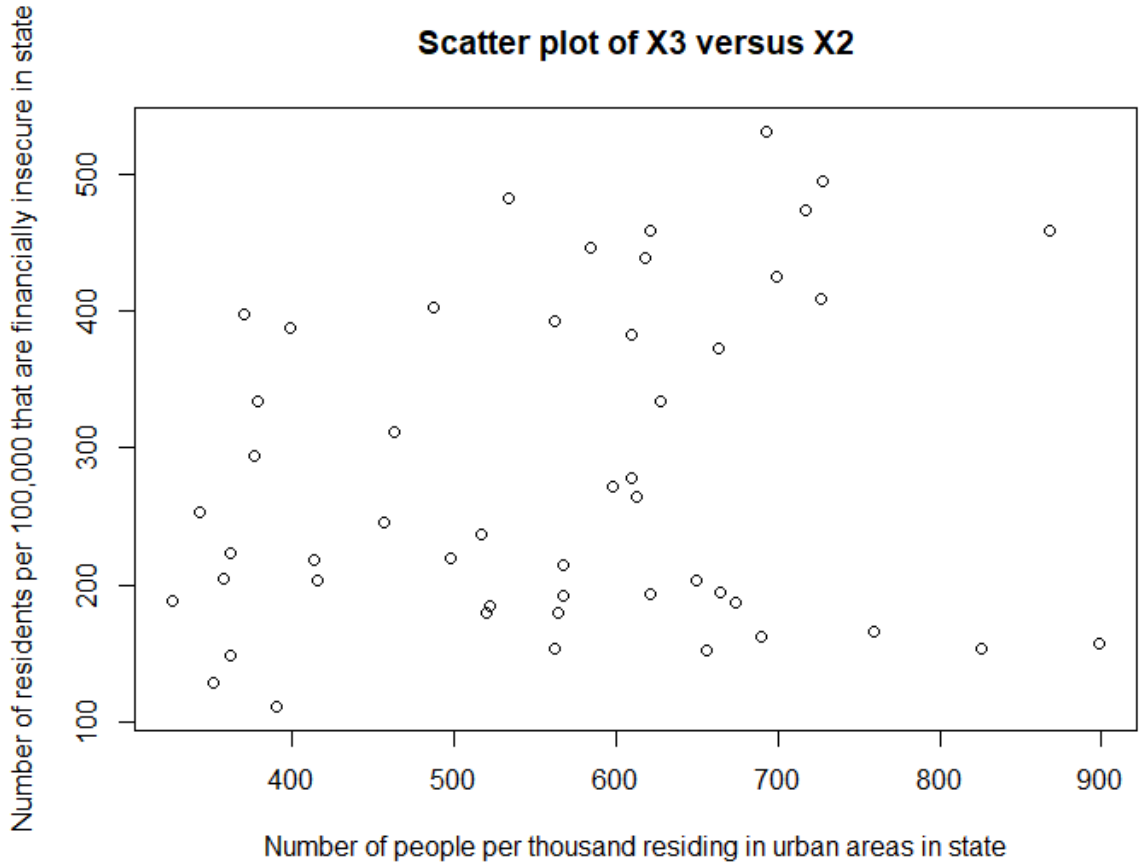
It also appears unlikely that there is much correlation between expenditure on shelters/housing assistance and the number of people per thousand residing in urban areas in state. The data appears to be fairly randomly distributed in the scatter plot but its possible that with a larger sample size that an upward, positive trend could become apparent. There seems to be a small indication that as the number of people residing in urban areas increases, so does the expenditure on shelters/housing assistance in state. However, this is difficult to prove without additional testing and evidence.

Graph of X1 and X2



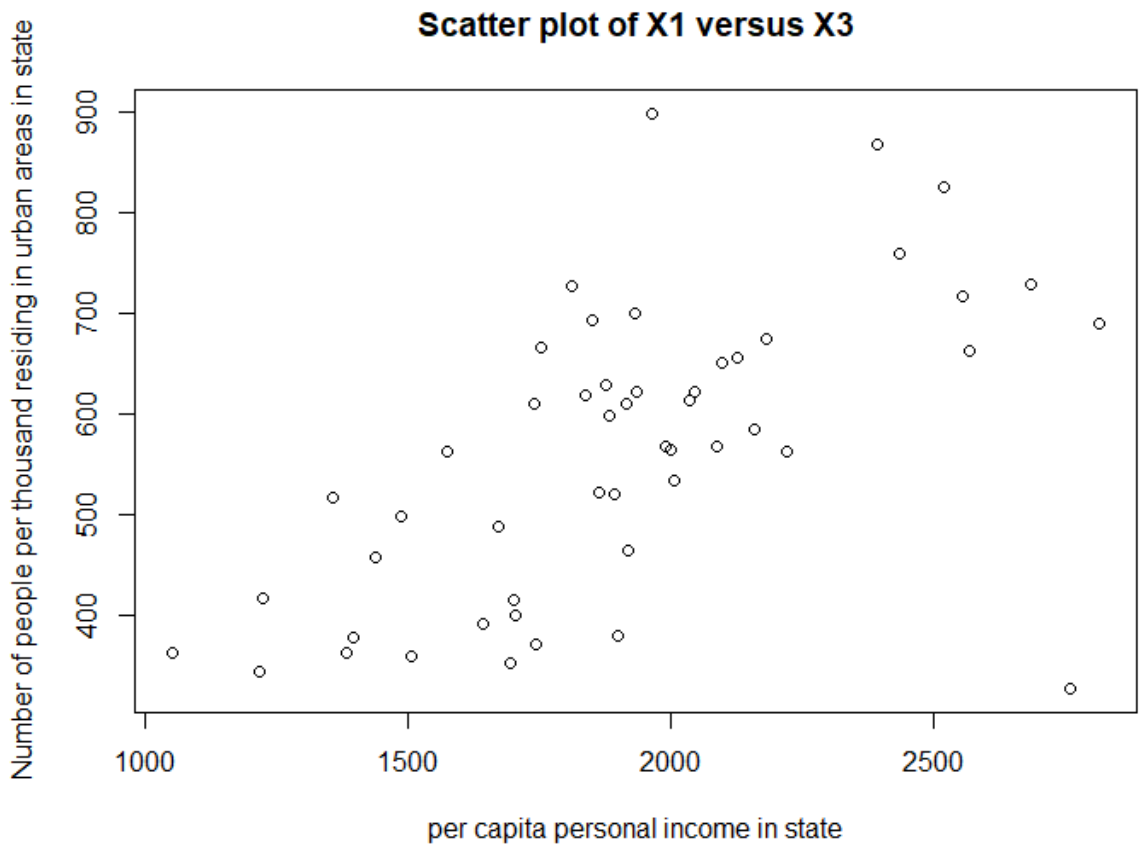
The data points the plot of per capita personal income in state and the number of people per 100,000 that are financially insecure in state seem to be very scattered and with no apparent correlation. It would be impossible to identify a trend or insert a line of best fit for this graph. It also makes sense that there is no correlation between these two variables when considering what they represent.

Graph of X2 and X3



There does not seem to be any correlation between the number of people per thousand residing in urban areas and the number of residents per 100,000 that are financially insecure in state. The data points don't form any patterns or clusters and are fairly randomly placed throughout the scatter plot.

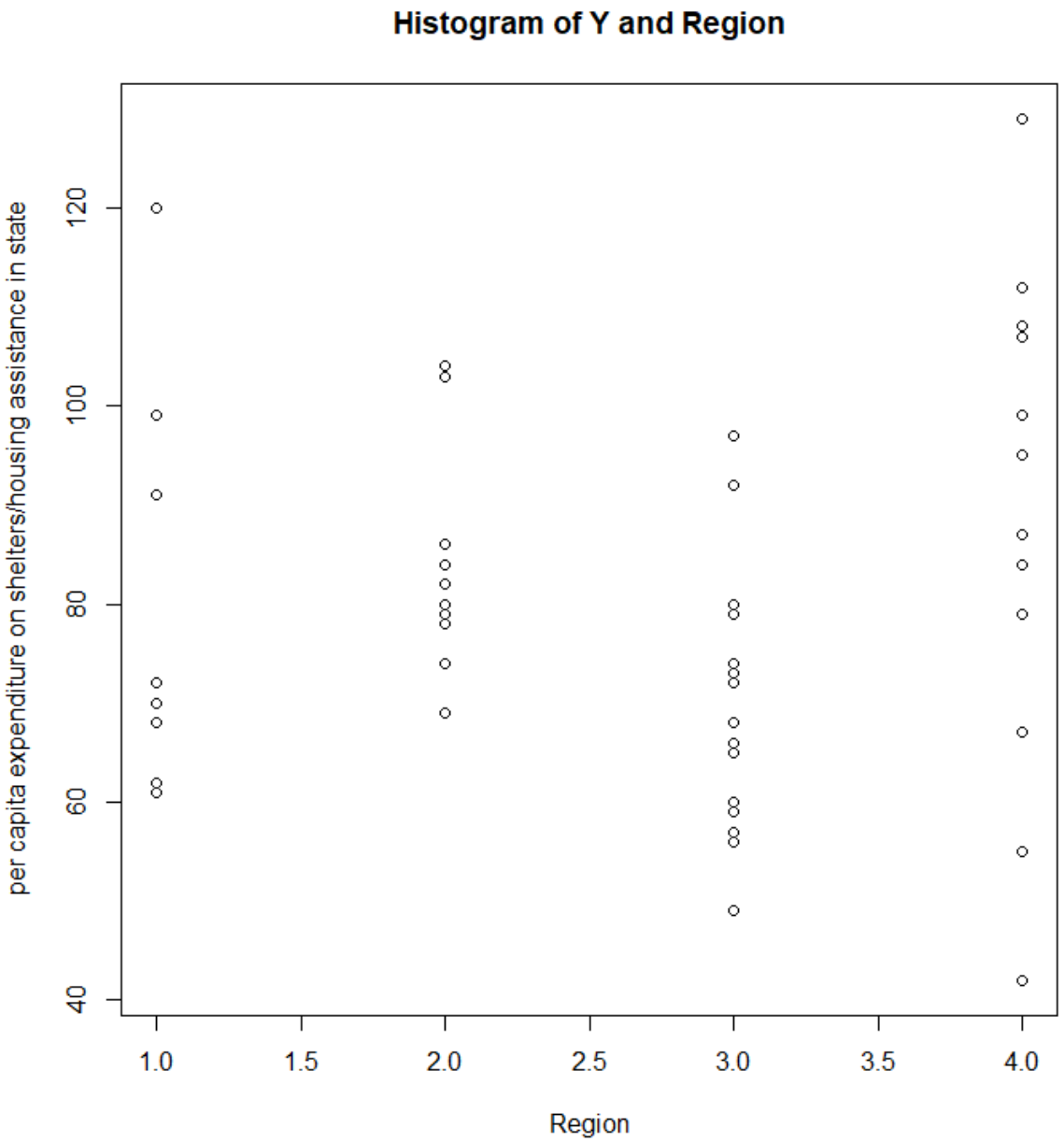
Graph of X1 and X3



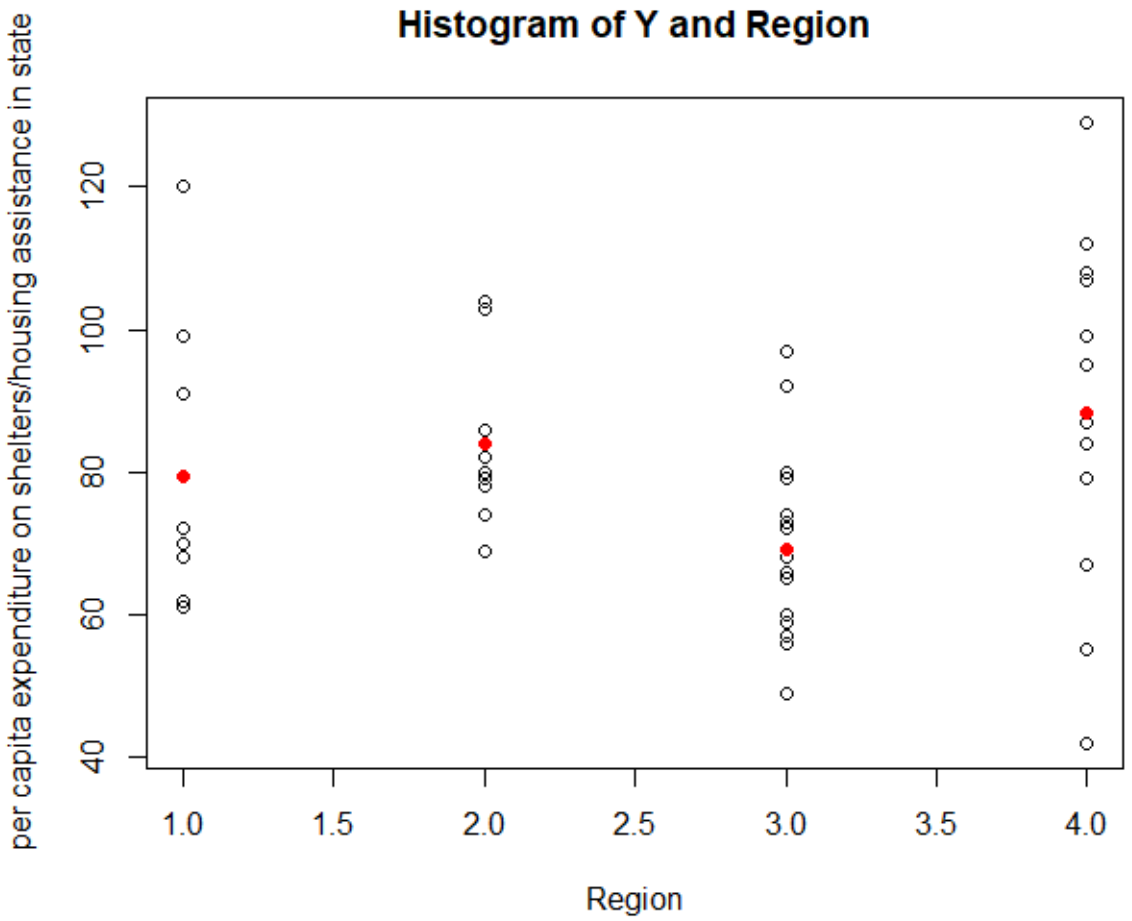
There seems to be a correlation between per capita personal income in state and the number of people per thousand residing in urban areas in state. As personal income increases, so does the number of people residing in urban areas. On the plot, this is shown through the data points trending upwards along the x-axis representing income. If a line of best fit was inserted into the graph, it is safe to assume that it would be linear and have a positive slope.

Part 2

Graph of Y and Region



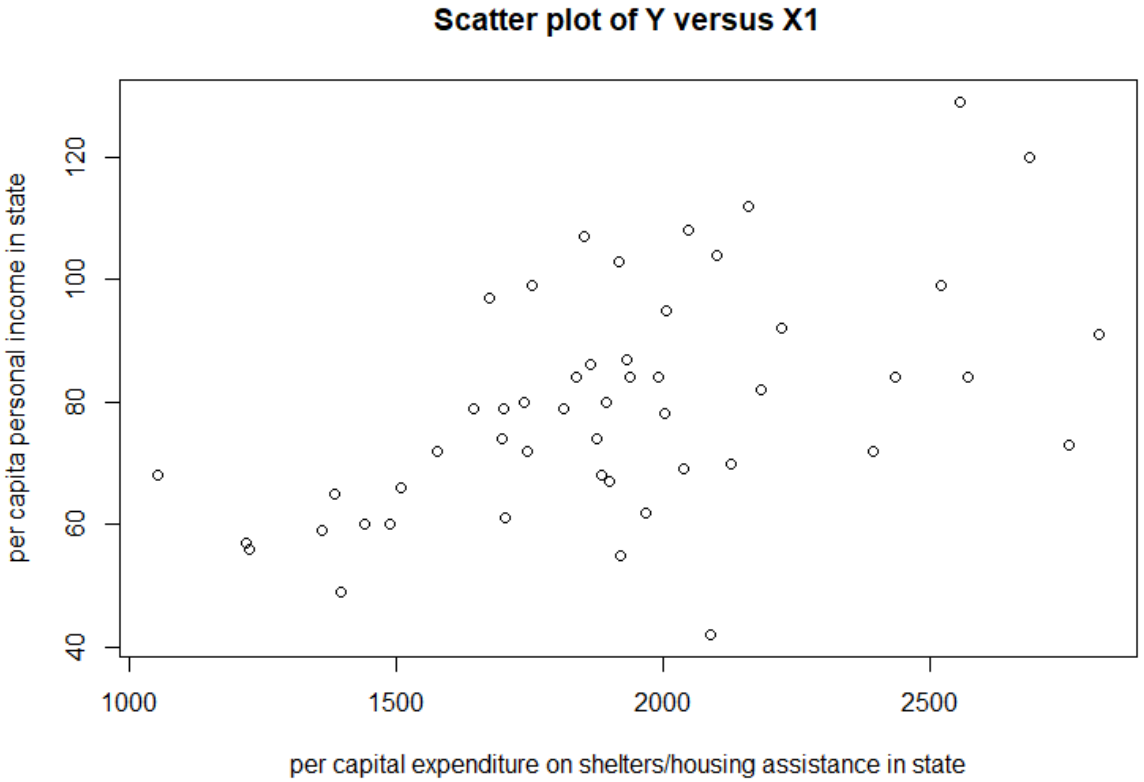
Along the x-axis representing region, each whole number value represents a different area. 1 is Northeast, 2 is North Central, 3 is South, and 4 is West. Looking at the plot of per capita expenditure on shelters/housing assistance in state versus region, it seems that spending is highest in the West, followed closely by the Northeast. This could be explored further by calculating the average expenditure for each region and plotting that.



This confirms that the West has the highest spending and that the North Central region actually has the second highest expenditure on shelters/housing assistance.

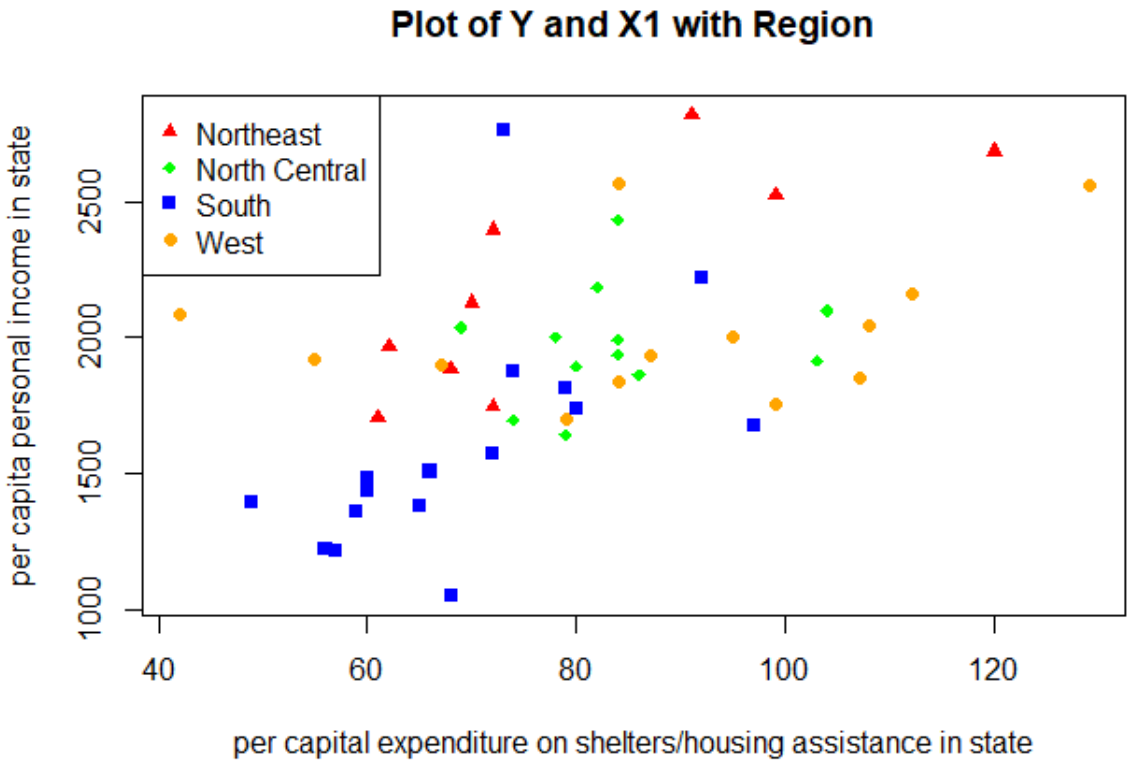
Part 3

Plot of Y and X1



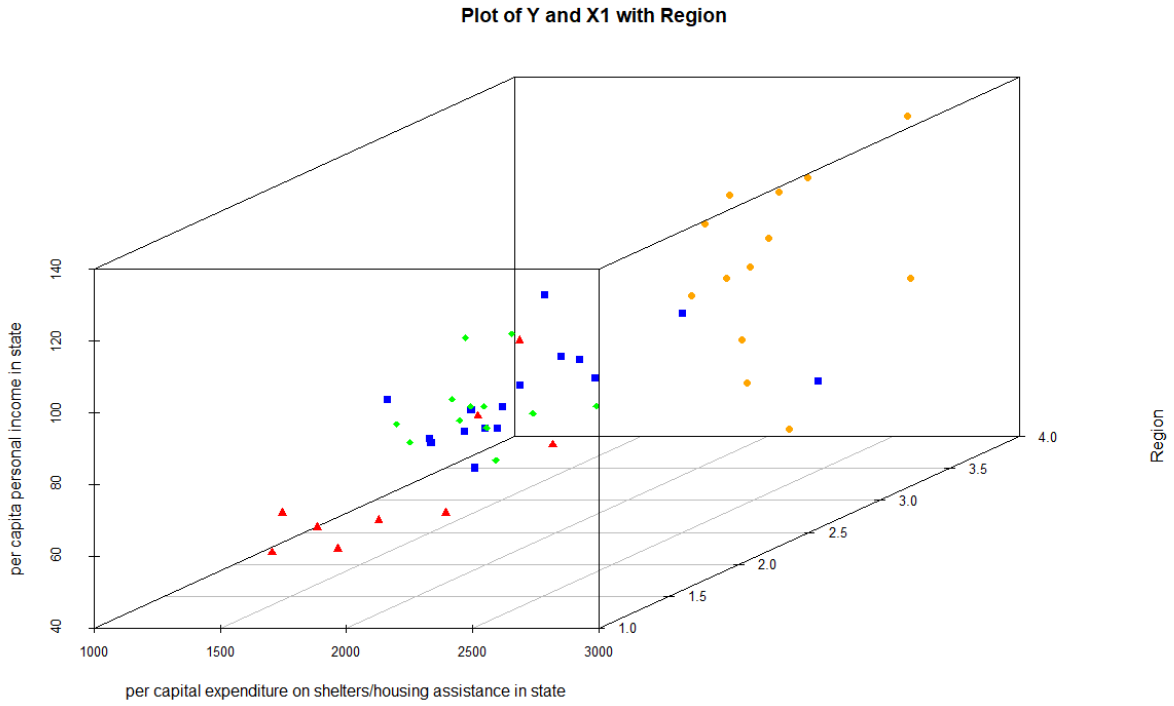
Again, there seems to be a slight positive correlation between per capita expenditure on shelters/housing assistance in state and per capita personal income in state. That is, as personal income increases, so does expenditure on housing assistance.

Plot of Y and X1 with Region



According to this plot, it seems that blue squares, representing the South, show that that region has the lowest per capita expenditure on shelters/housing assistance in state as well as the lowest per capita personal income. The highest area is the West, represented by the orange circles.

3D plot of Y and X1 with Region


Plot of Y and X1 with Region
per capita personal income in state
40
60
80
100
120
140
1000
1500
2000
2500
3000
per capital expenditure on shelters/housing assistance in state
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Region