Problem Set 1

Applied Stats/Quant Methods 1

Charis Rogers (19333698)

Question 1: Education

A school counselor was curious about the average of IQ of the students in her school and took a random sample of 25 students' IQ scores. The following is the data set:

```
IQ <- c(105, 69, 86, 100, 82, 111, 104, 110, 87, 108, 87, 90, 94, 113, 112,
    98,
```

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

```
t_score <- qt(0.95, df=length(IQ)-1)
t_score
t_lower_99<- mean (IQ) -(t_score)*(sd(IQ)/sqrt(n))
t_upper_99<- mean (IQ) +(t_score)*(sd(IQ)/sqrt(n))
```

   For a 90% confidence interval [93.96, 102.92]

2. Next, the school counselor was curious whether the average student IQ in her school is higher than the average IQ score (100) among all the schools in the country. Using the same sample, conduct the appropriate hypothesis test with $\alpha = 0.05$.

   H0 $\mu \leq 100$. Null hypothesis: the average IQ of students in this specific school is less than or equal to the average IQ among all the schools in the country.

   HA $\mu > 100$. Alternative hypothesis: the average IQ of students in this specific school is greater than the average IQ among all the schools in the country.

```
t_stat <- (mean(IQ)-100)/(sd/sqrt(n))
t_stat #-0.5957439
p_value <- (pt(abs(t_stat), 24))
p_value # 0.7215383
```

   $0.7215 > 0.05$ which means we cannot reject the null hypothesis. Therefore there is no evidence to support the idea that the average IQ of this specific class is greater than the average IQ among all the schools in the country.

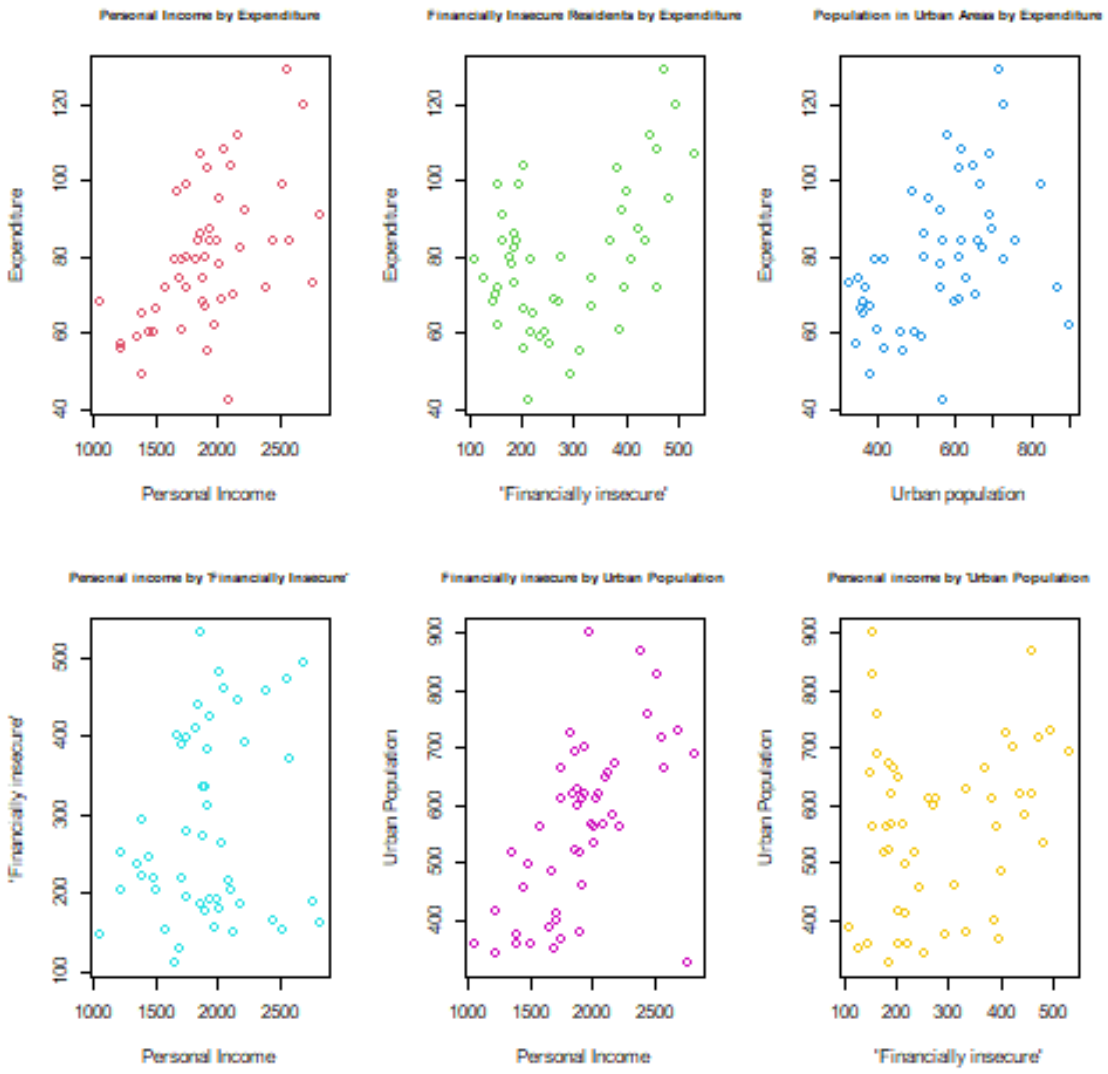
Question 2: Political Economy

Researchers are curious about what affects the amount of money communities spend on addressing homelessness. The following variables constitute our data set about social welfare expenditures in the USA.

`State`	*50 states in US*
`Y`	*per capita expenditure on shelters/housing assistance in state*
`X1`	*per capita personal income in state*
`X2`	*Number of residents per 100,000 that are "financially insecure" in state*
`X3`	*Number of people per thousand residing in urban areas in state*
`Region`	*1=Northeast, 2= North Central, 3= South, 4=West*

Explore the `expenditure` data set and import data into `R`.

- Please plot the relationships among *Y*, *X1*, *X2*, and *X3*? What are the correlations among them (you just need to describe the graph and the relationships among them)?



There seems to be a positive linear relationship between Y and X1. This means that states with a higher personal income per capita spend more on shelters/housing assistance in state.

There seems to be a non-linear between Y and X2 which means there is no linear association between the number of residents in a state that are 'financially insecure' and expenditure on shelters/housing assistance.

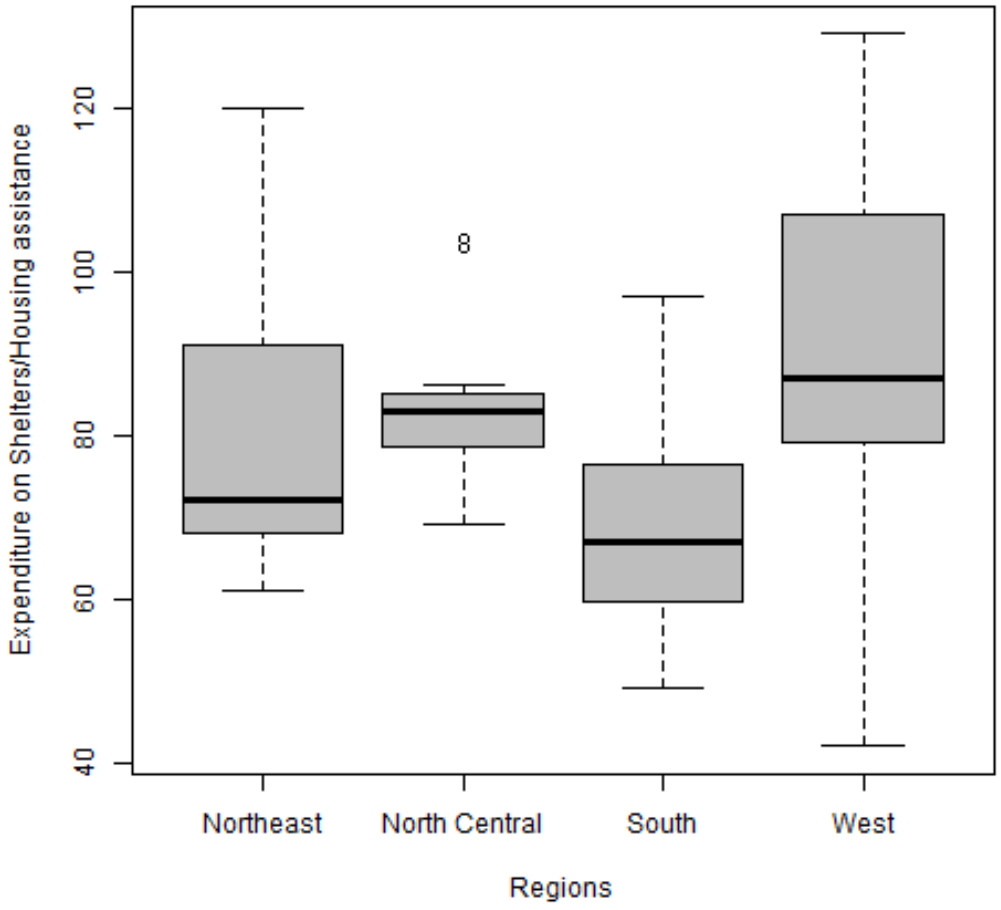
There seems to be a weak positive linear relationship between Y and X3. This means that states with a greater urban population spend more on shelters/housing assistance in state compared to states with less people living in urban areas.

There seems to be no association between X1 and X2. That means that there is no relationship between the average personal income in a state and the number of residents in a state that are 'financially secure'.

There seems to be a weak linear positive relationship between X1 and X3. This means that states with a higher personal income are also states which have a higher number of people living in urban areas.

There does not seem to be any association between X2 and X3. This means there is no relationship between the number of residents in a state who are 'financially secure' and the number of people living in urban areas in a state.

- Please plot the relationship between Y and *Region*? On average, which region has the highest per capita expenditure on housing assistance?

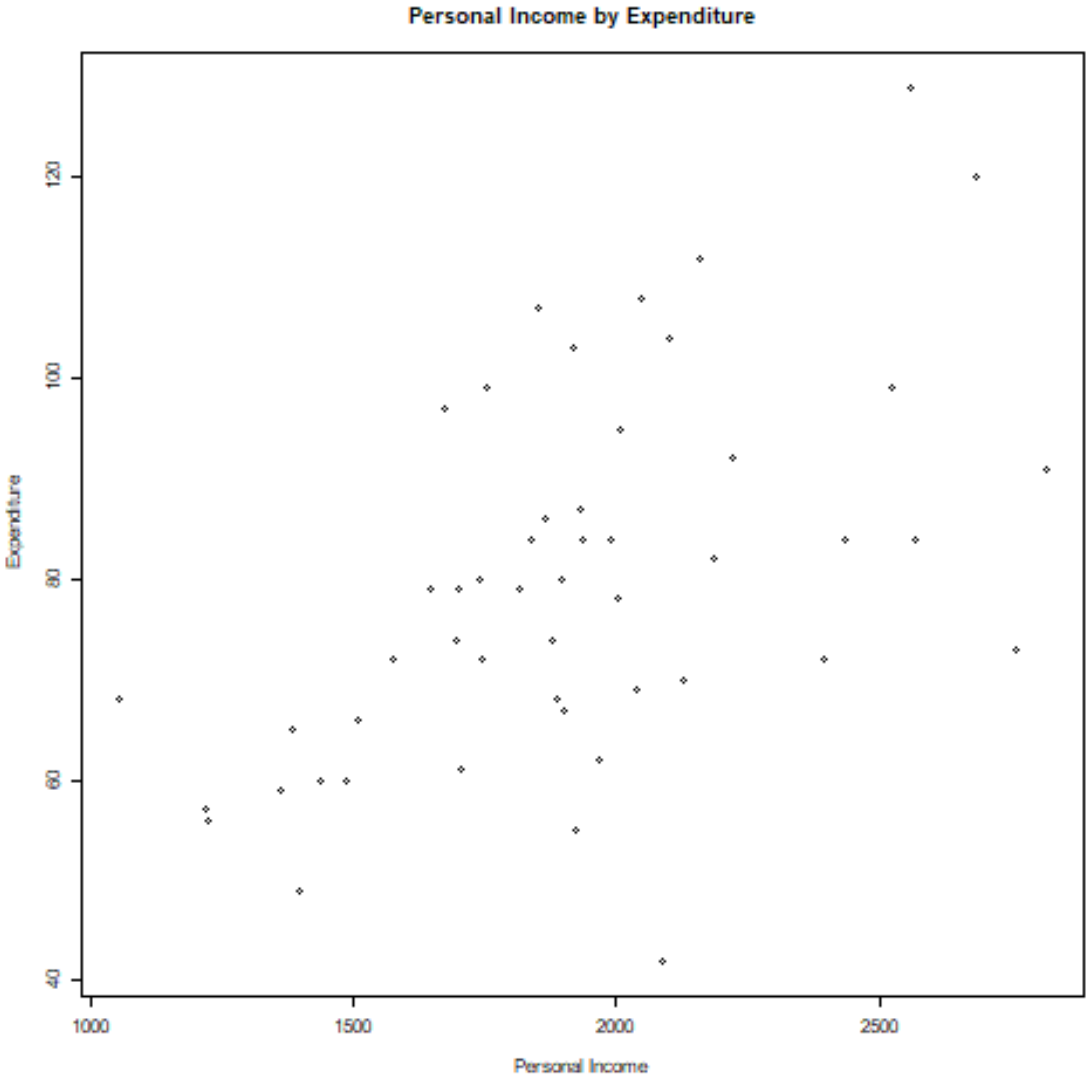



```
mean(expenditure$Y[expenditure$Region == "1"])
mean(expenditure$Y[expenditure$Region == "2"])
mean(expenditure$Y[expenditure$Region == "3"])
mean(expenditure$Y[expenditure$Region == "4"])
```

Northeast	*79.44444*
North Central	*83.91667*
South	*69.1875*
West	*88.30769*

As seen in the table above, the West has the highest per capita expenditure on housing assistance

- Please plot the relationship between Y and $X1$? Describe this graph and the relationship. Reproduce the above graph including one more variable *Region* and display different regions with different types of symbols and colors.



As seen in the graph above, the relationship between Y and X1 seems to be a positive linear relationship. This means that in states with an average higher personal income, the spending on housing assistance will, on average, be higher.