APG 4013C

ASSIGNMENT 2- Report

Student no: MPNFLA001

Name: Flassie Mpanza

#Number 1

Output :

```
     HOVAL             INC              CRIME              OPEN        
 Min.   :17.90   Min.   : 4.477   Min.   : 0.1783   Min.   : 0.0000  
 1st Qu.:25.70   1st Qu.: 9.963   1st Qu.:20.0485   1st Qu.: 0.2598  
 Median :33.50   Median :13.380   Median :34.0008   Median : 1.0061  
 Mean   :38.44   Mean   :14.375   Mean   :35.1288   Mean   : 2.7709  
 3rd Qu.:43.30   3rd Qu.:18.324   3rd Qu.:48.5855   3rd Qu.: 3.9364  
 Max.   :96.40   Max.   :31.070   Max.   :68.8920   Max.   :24.9981  
     PLUMB            DISCBD      
 Min.   : 0.1327   Min.   :0.370  
 1st Qu.: 0.3323   1st Qu.:1.700  
 Median : 1.0239   Median :2.670  
 Mean   : 2.3639   Mean   :2.852  
 3rd Qu.: 2.5343   3rd Qu.:3.890  
 Max.   :18.8111   Max.   :5.570  
```

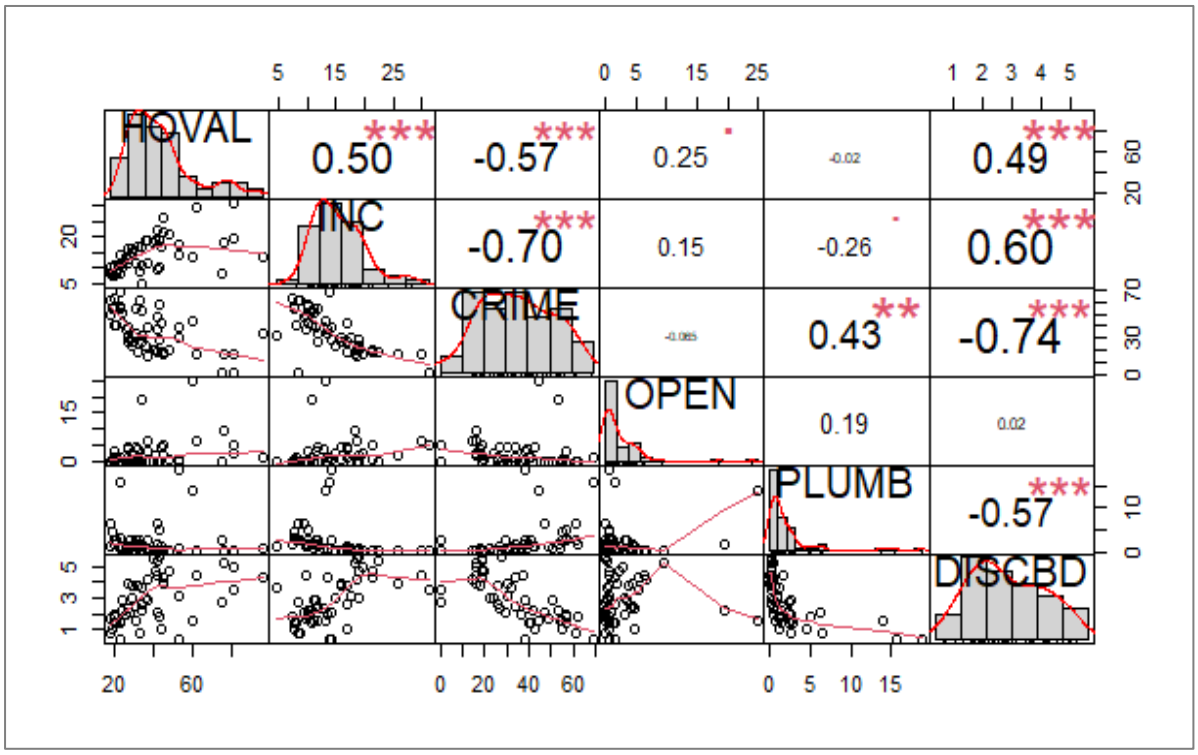
#Number 2

Output:

```
            Mean  Maximum   Minimum
HOVAL   38.436224 96.40000 17.900000
INC     14.374939 31.07000  4.477000
CRIME   35.128824 68.89204  0.178269
OPEN     2.770938 24.99807  0.000000
PLUMB    2.363944 18.81107  0.132743
DISCBD   2.852041  5.57000  0.370000
```

#Number 3

Plot:



Conclusion:

When the value of r is close to 0, it suggests a minimal linear connection. The strength of the linear connection strengthens as r deviates from 0 and approaches either -1 or 1.
The correlation coefficients between:

HOVAL and INC is 0.4999 indicating moderate positive correlation, as home values increase, income tends to increase as well.

HOVAL and CRIME is -0.5745 indicating a moderate negative correlation, areas with higher home values tend to have lower crime rates.

INC and CRIME is -0.6956 indicating moderate negative correlation, areas with higher income tend to have lower crime rates.

OPEN and PLUMB is 0.1920 indicating weak positive correlation, areas with more open space to have better plumbing.

CRIME and PLUMB is 0.4327 indicating moderate positive correlation, areas with better plumbing tend to have higher crime rates.

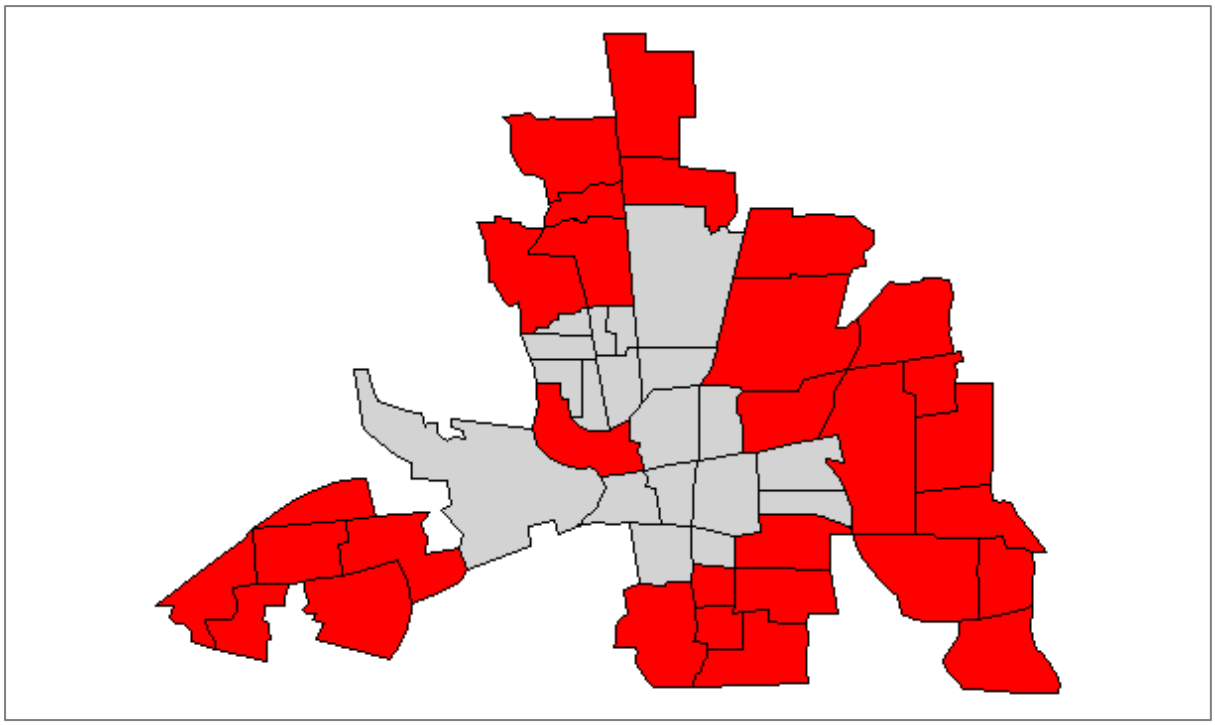
DISCBD and HOVAL is 0.4852 indicating moderate positive correlation, areas farther from the central business district tend to have higher home values.

DISCBD and INC is 0.6003 indicating moderate positive correlation, areas farther from the central business district tend to have higher incomes.

DISCBD and CRIME is -0.7408 indicating strong negative correlation, areas farther from the central business district tend to have lower crime rates.

#Number 4

Plot:



Conclusion:

High Crime Areas are marked by light grey shading, indicating elevated crime rates (above 40), primarily concentrated in central zones, while outer regions experience comparatively lower crime. These zones raise significant crime concerns. Conversely, "Low Crime Areas" shown in red signify regions with lower crime rates (below 40), potentially representing areas with reduced criminal activity.

#Number 5

Conclusion:

The data visualization reveals geographic crime rate patterns in Columbus, Ohio. Darker green shades indicate higher crime rates, while lighter shades represent lower rates. Different classification methods slightly alter class divisions and colour assignments, resulting in differences in how the maps appear.

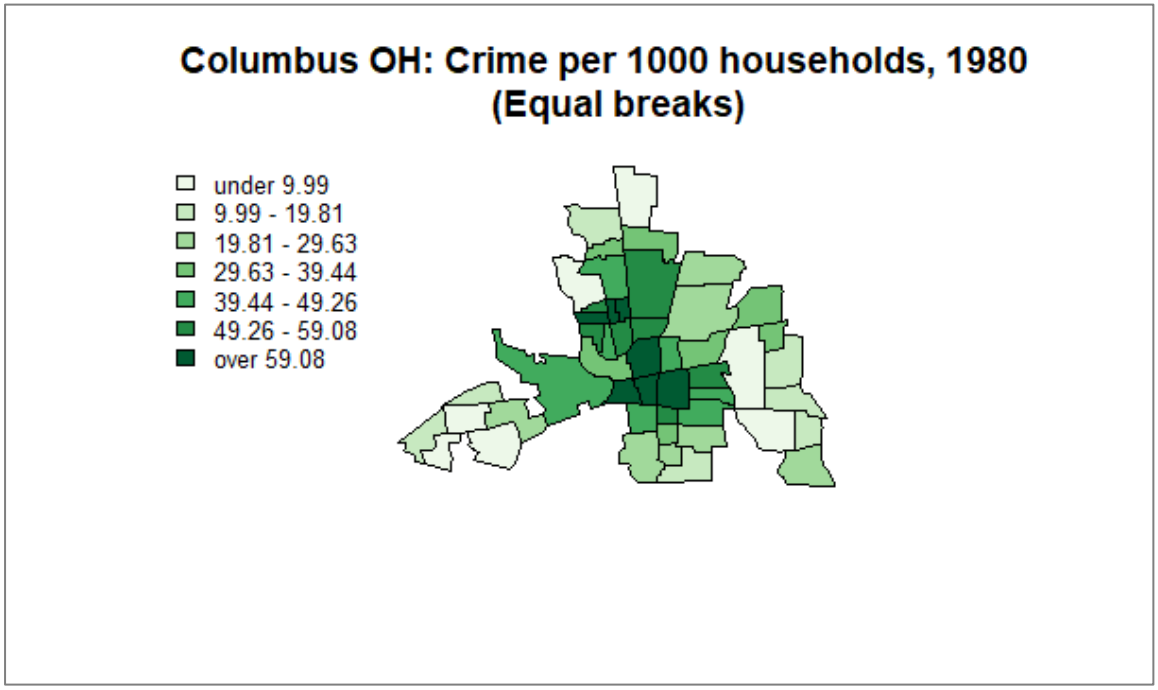
Answer:

The key differences between the plots are in how they classify and display the data:

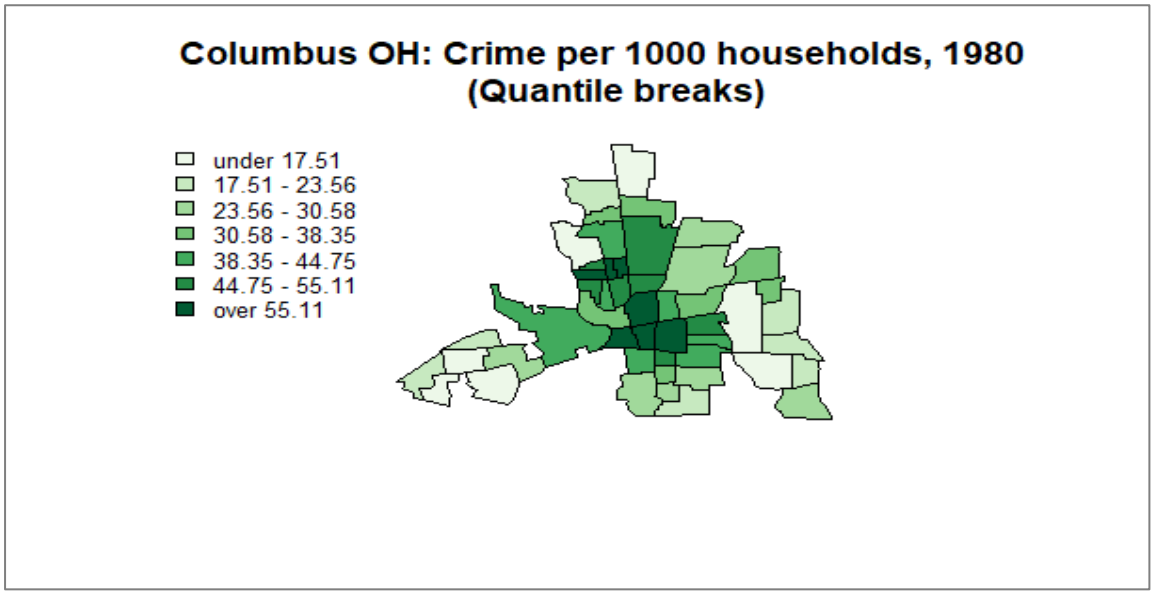
- **Equal Breaks** emphasize uniformity in class intervals but might not highlight the variations in crime rates effectively.
- **Quantile Breaks** focus on an equal distribution of data points in each interval, making it helpful for identifying areas with significantly different crime rates.
- **Jenks Breaks** aim to reveal natural groupings based on data variance, which can be useful for highlighting areas with abrupt changes in crime rates.

Plots:

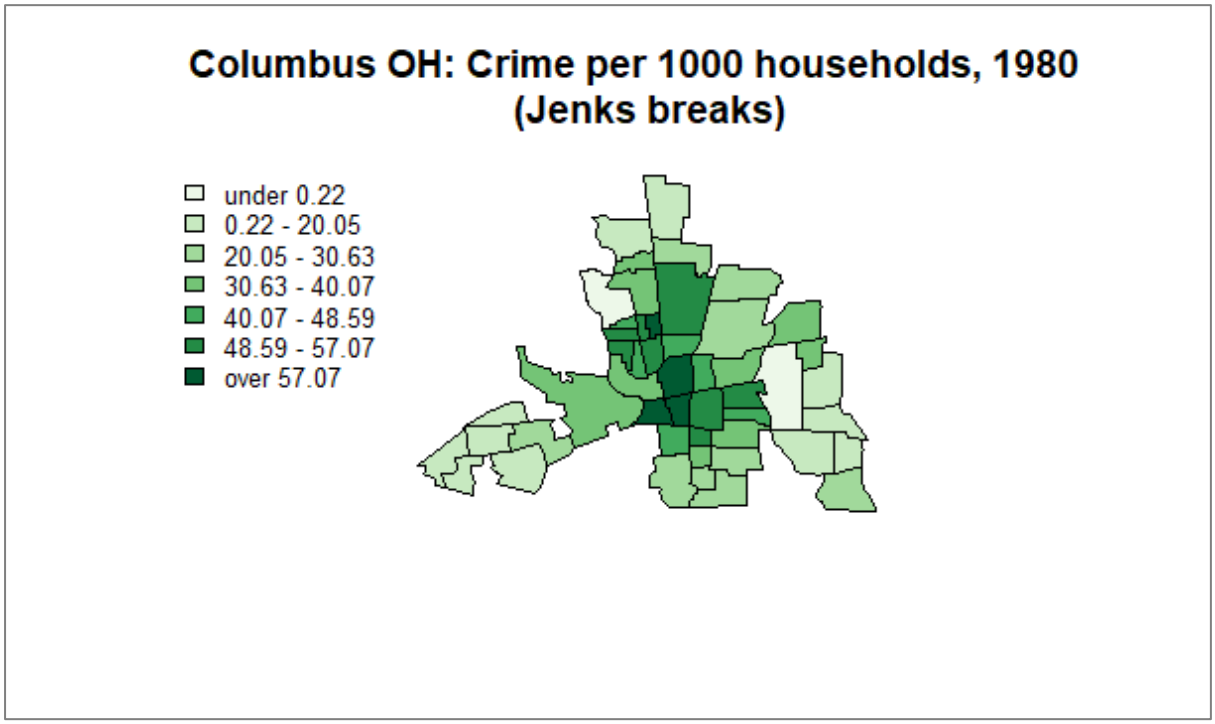
Equal Breaks



Quantile Breaks



Jenks Breaks



#Number 6

weight matrix summary output:

```
Characteristics of weights list object:
Neighbour list object:
Number of regions: 49
Number of nonzero links: 236
Percentage nonzero weights: 9.829238
Average number of links: 4.816327
Link number distribution:

 2  3  4  5  6  7  8  9 10
 5  9 12  5  9  3  4  1  1
5 least connected regions:
0 5 41 45 46 with 2 links
1 most connected region:
19 with 10 links

Weights style: W
Weights constants summary:
   n   nn S0       S1       S2
W 49 2401 49 22.75119 203.7091
```

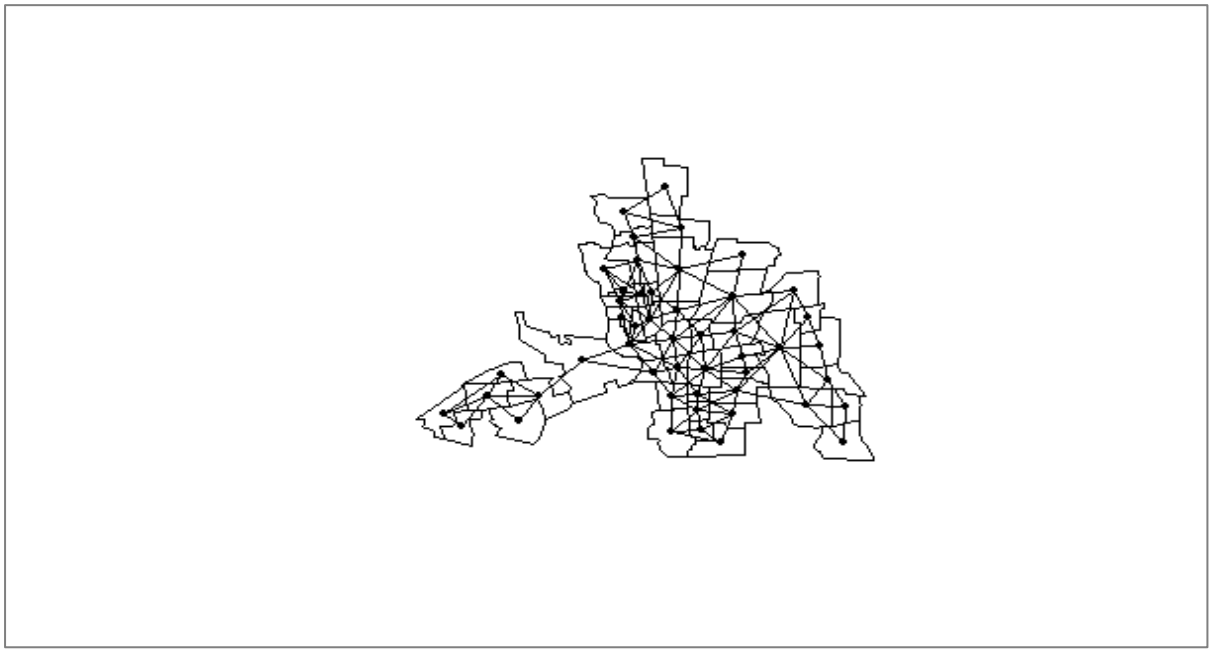
Answers:

Average number of links is **4.816327.**

The least connected regions are **0**, **5**, **41**, **45** and **46** with 2 links.

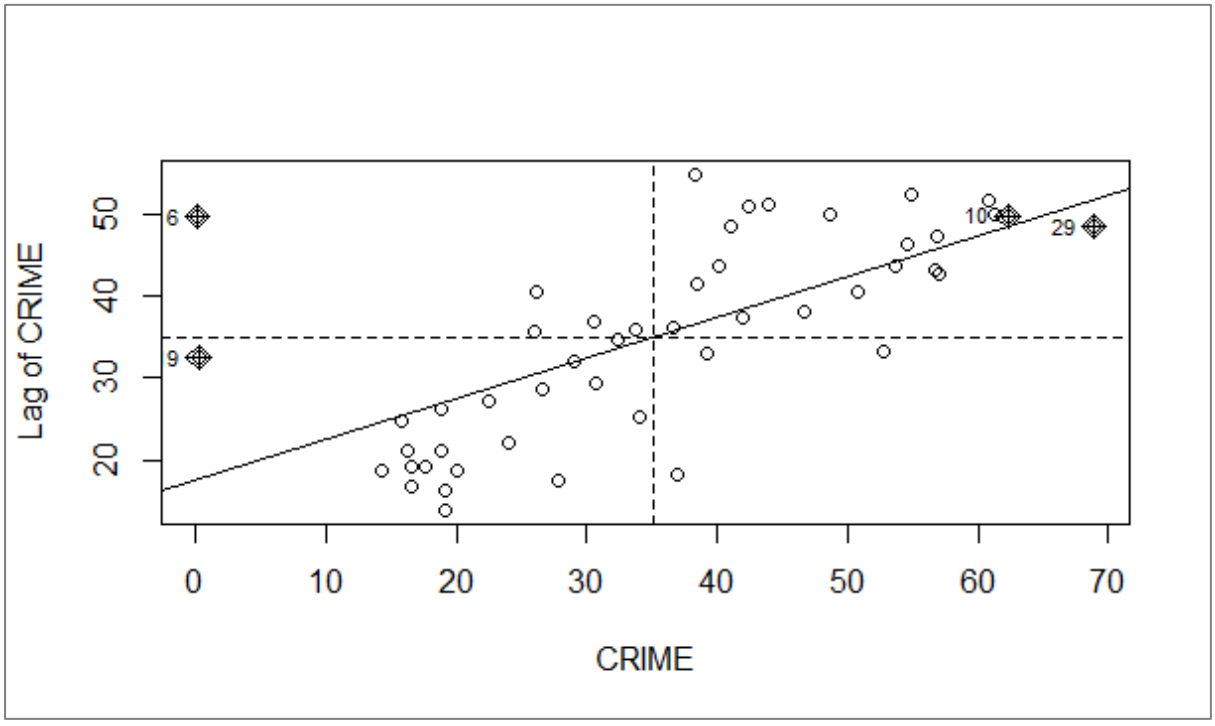
The most connected region is **19** with 10 links.

Plot:



#Number 7

Plot:



Conclusion:

The p-value of 1.139e-08, which is a very small value (close to zero), indicates strong evidence against the null hypothesis. In the context of the Moran's I test, this p-value suggests that the

observed spatial autocorrelation in the "CRIME" variable among the neighboring regions is highly unlikely to occur purely by random chance.

#Number 8

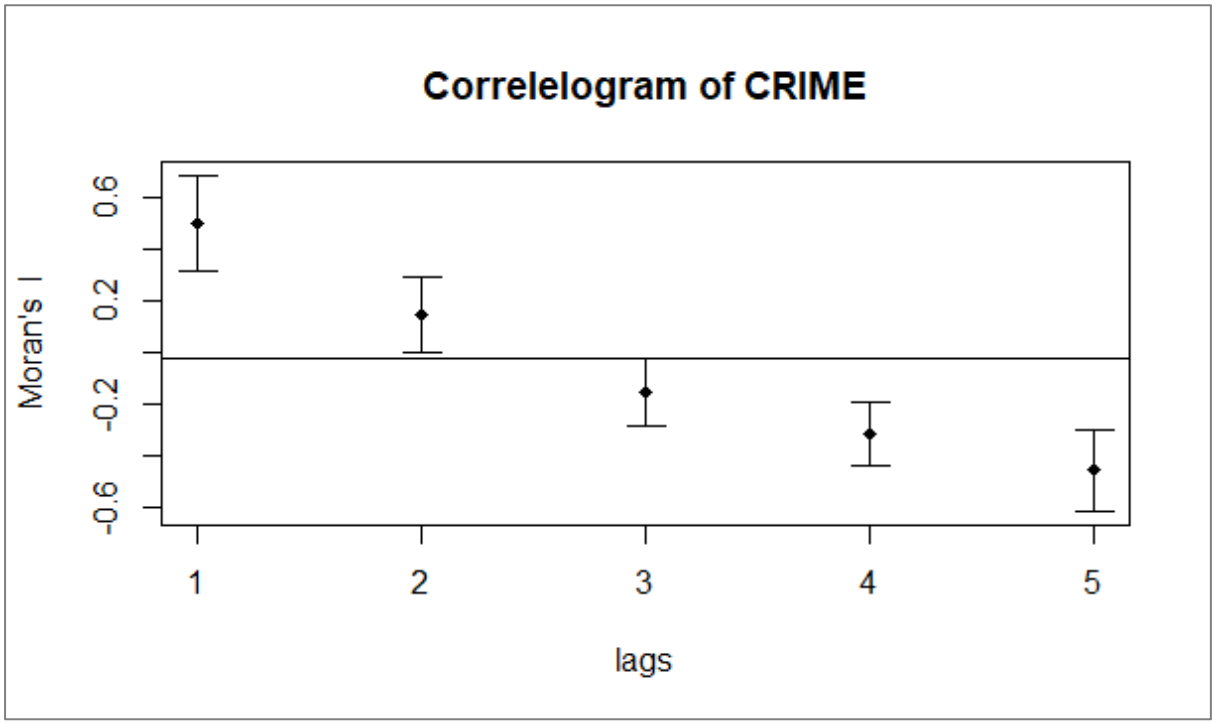
Conclusion:

Positive Spatial Autocorrelation: If the correlogram values are positive and decrease as the distance lag increases, it indicates that similar crime rates tend to cluster together in nearby regions. This suggests the presence of positive spatial autocorrelation, where regions with high crime rates are surrounded by regions with similarly high crime rates, and vice versa.

columbus$CRIME summary output:

```
Length Class  Mode
res     15     -none- numeric
method   1     -none- character
cardnos  5     -none- list
var      1     -none- character
```

Plot:



#Number 9

columbus_locm summary output:

```
       Ii                 E.Ii                   Var.Ii              Z.Ii
 Min.   :-1.86059   Min.   :-9.279e-02   Min.   :0.001108   Min.   :-1.799
2
```

```
 1st Qu.: 0.06795   1st Qu.:-2.857e-02   1st Qu.:0.029554   1st Qu.: 0.398
2
 Median : 0.46768   Median :-1.728e-02   Median :0.149293   Median : 1.285
7
 Mean   : 0.50019   Mean   :-2.083e-02   Mean   :0.206937   Mean   : 1.216
8
 3rd Qu.: 0.96930   3rd Qu.:-2.827e-03   3rd Qu.:0.289330   3rd Qu.: 2.061
1
 Max.   : 1.65517   Max.   :-9.665e-05   Max.   :0.965424   Max.   : 3.360
7
 Pr(z != E(Ii))
 Min.   :0.0007776
 1st Qu.:0.0392923
 Median :0.1619765
 Mean   :0.2930883
 3rd Qu.:0.4518551
 Max.   :0.9522213
```

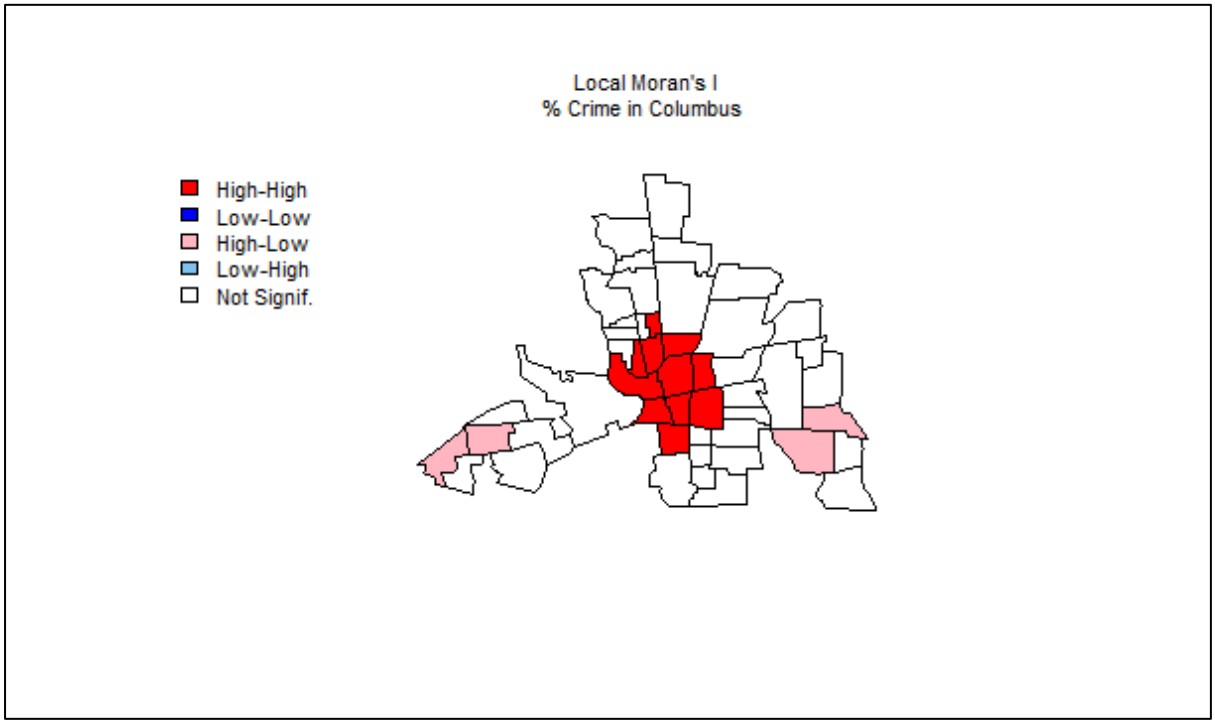
CRIME summary output:

```
V1
 Min.   :-2.08883
 1st Qu.:-0.90128
 Median :-0.06741
 Mean   : 0.00000
 3rd Qu.: 0.80424
 Max.   : 2.01787
```

columbus$lag_sCRIME summary output:

```
V1
 Min.   :-1.271457
 1st Qu.:-0.622430
 Median : 0.046334
 Mean   :-0.004665
 3rd Qu.: 0.676146
 Max.   : 1.181964
```

Plot:



#Number 10

Conclusion:

The linear regression model (reg1) utilizes INC and HOVAL as predictors, showing a moderately strong fit. The Multiple R-squared value indicates around 56% of CRIME variability explained. The Adjusted R-squared, considering model complexity, remains near 54%, balancing explanatory power and simplicity. Income and housing value offer meaningful insights into Columbus' crime rate variation. Moran's I test on model residuals suggests spatial autocorrelation, implying unaccounted spatial factors. Exploring these patterns could enhance model accuracy or reveal more variables behind spatial crime rate trends.

reg1 summary output:

```
Call:
lm(formula = columbusdata$CRIME ~ columbusdata$INC + columbusdata$HOVAL)

Residuals:
    Min      1Q  Median      3Q     Max
-34.418  -6.388  -1.580   9.052  28.649

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       68.6190     4.7355  14.490  < 2e-16 ***
columbusdata$INC  -1.5973     0.3341  -4.780 1.83e-05 ***
columbusdata$HOVAL -0.2739     0.1032  -2.654   0.0109 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.43 on 46 degrees of freedom
```

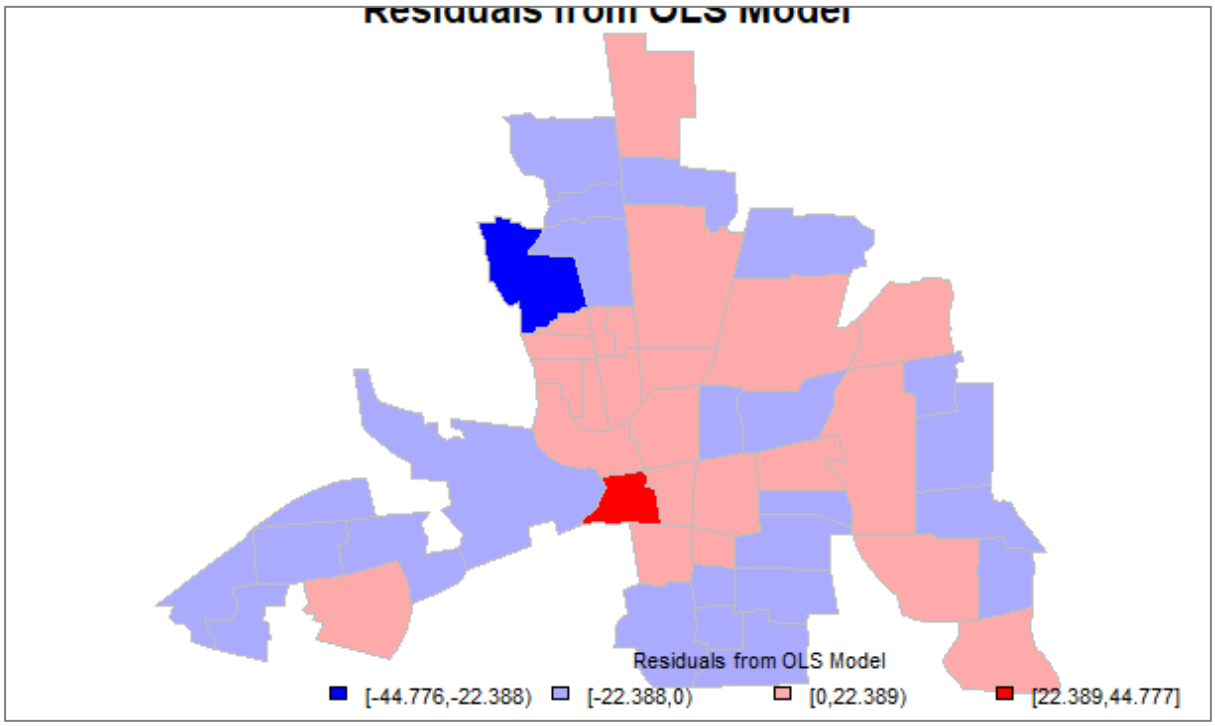
```
Multiple R-squared:  0.5524,   Adjusted R-squared:  0.5329
F-statistic: 28.39 on 2 and 46 DF,  p-value: 9.341e-09
```

#Number 11

Answer:

Yes, there is autocorrelation in the residuals.

Plot:



Moran's I output:

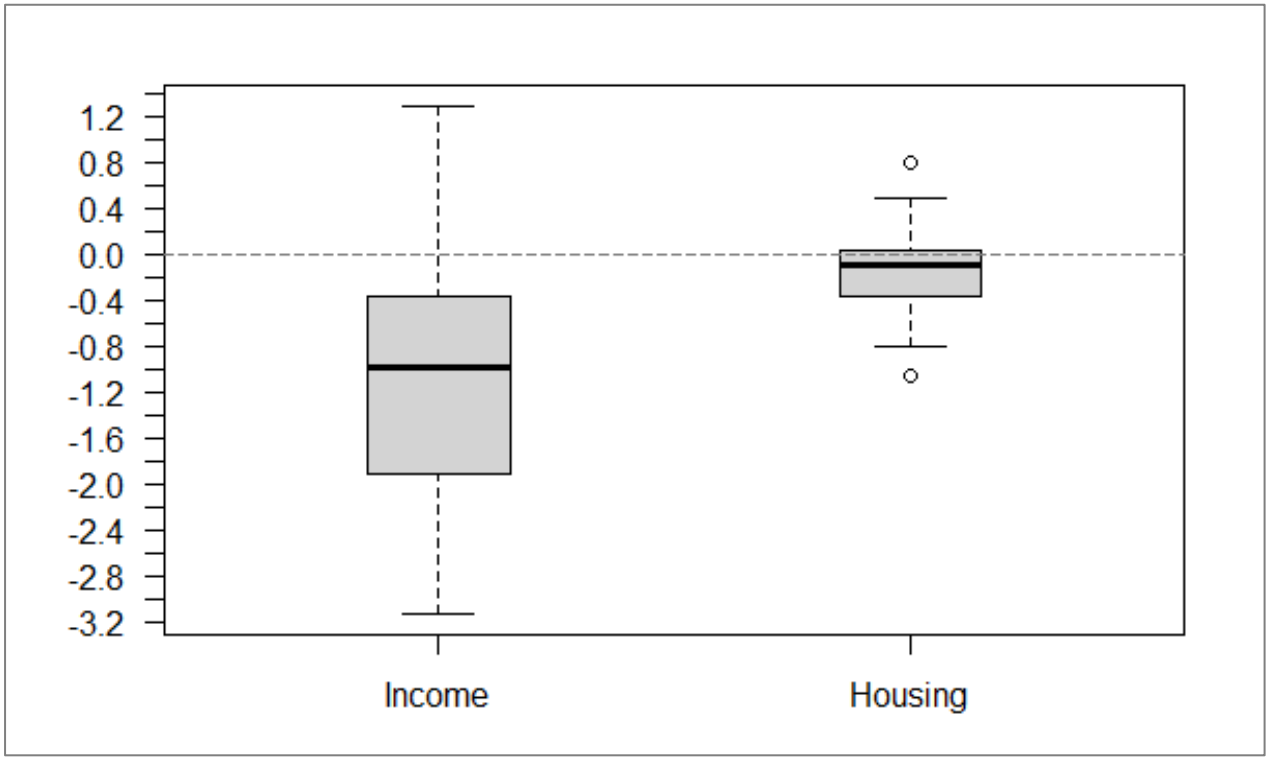
```
Moran I test under randomisation

data:  reg1$residuals
weights: columbus.wts

Moran I statistic standard deviate = 2.6521, p-value = 0.007999
alternative hypothesis: two.sided
sample estimates:
Moran I statistic       Expectation          Variance
      0.222109407      -0.020833333       0.008391173
```

#Number 12

Plot:



Code:

```
library(spgwr)
data(columbus) # Load the "columbus" dataset
crime.bw <- gwr.sel(CRIME ~ INC + HOVAL,
                    data=columbus,
                    coords=cbind(columbus$X, columbus$Y))
crime.gauss <- gwr(CRIME ~ INC + HOVAL,
                   data=columbus,
                   coords=cbind(columbus$X, columbus$Y),
                   bandwidth=crime.bw)
#Plotting the distribution of beta coefficients
d <- cbind(crime.gauss$SDF$INC, crime.gauss$SDF$HOVAL)
#Setting plot margins and create the boxplot
par(mar=c(3,4,2,2))
boxplot(d, xaxt="n", yaxt="n", pars=list(boxwex=0.3))
axis(1, at=1:2, label=c("Income", "Housing"))
axis(2, at=seq(-4, 2, 0.2), las=1)
```

```
abline(h=0, lty="4343", col="#7E7E7E")
mtext("Beta i", 2, line=3)
```

Conclusion:

The Geographic Weighted Regression (GWR) analysis of the Ohio dataset examines the relationship between CRIME and independent variables INC and HOVAL. Spatial patterns are revealed, with a boxplot illustrating varying impact across Columbus regions. Higher income generally associates with lower crime, emphasizing its deterrent role. However, HOVAL's effect on crime varies spatially, demanding customized strategies for crime intervention. Acknowledging these spatial nuances aids policymakers and urban planners in informed, localized decisions for crime prevention and community development.

#Number 13

Conclusion:

The boxplot highlights the spatial variability in the impact of income and housing value on crime rates across different regions in Columbus. While higher income appears to generally correlate with lower crime rates, the relationship with housing value is more complex and region-specific. This underscores the importance of considering local factors and tailoring interventions for effective crime prevention and community development strategies in different parts of the city.