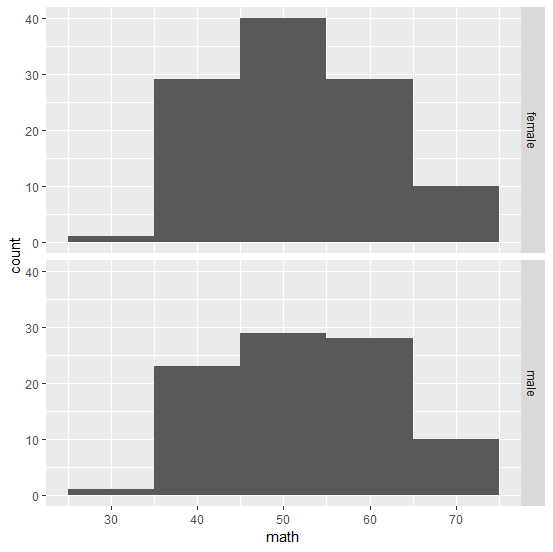
Question 1

Part 1: Describe in graphs and numbers the distribution of math scores between male and female

Table

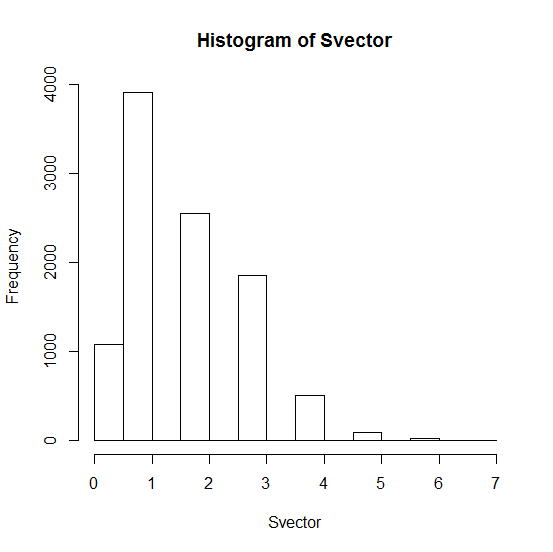
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 20-39 | 40-59 | 60-69 | Above 70 |
| Female | 5 | 78 | 22 | 4 |
| Male | 6 | 62 | 16 | 7 |

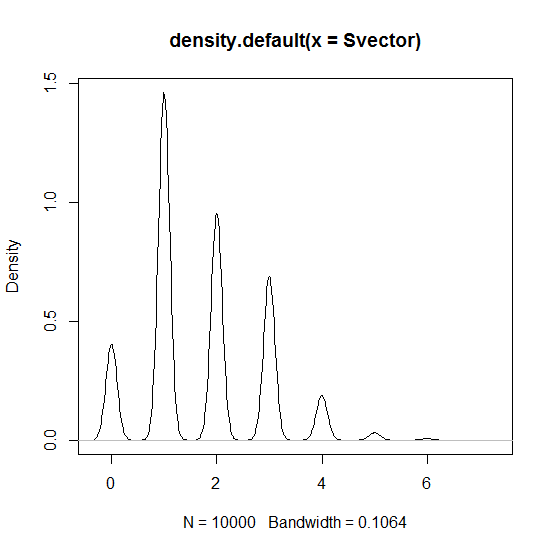


From initial look at the table and histogram, the median score looks to be in the range of lower 50s. The data was recoded into four groups to analyze it better. There are 109 girls and 91 boys in the sample. On the 20-39 range, more boys seem to score lower than girls. Boys also seem to have more high scores (greater than 70).

Part 2: Is there a significant difference in the median score between these two groups? Use a permutation test to find out. Remember to set seed so that the grader can reproduce your result.

* H0: There is no significant difference between median score of boys and girls.
* HA: There is a significant difference between median score of boys and girls.
* S = Difference in median math scores between female and male





s = |median(female) -median(male)|

s = 53 -52 = 1

P value obtained- P(S >= 1) = 0.8876

Table

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1080 | 3911 | 2553 | 1851 | 504 | 83 | 17 | 1 |

Over repeating the experiment 10000 times the p-value obtained was 0.8876, which is very high. So, our H0 is true i.e There is no significant difference in median scores between boys and girls.

Part 3: Is there a significant difference between male and female in the proportion of those who math score is 65 or more? Use a test of your choice.

Table:

|  |  |
| --- | --- |
|  | Greater than 65 score |
| Female | 13 |
| Male | 10 |

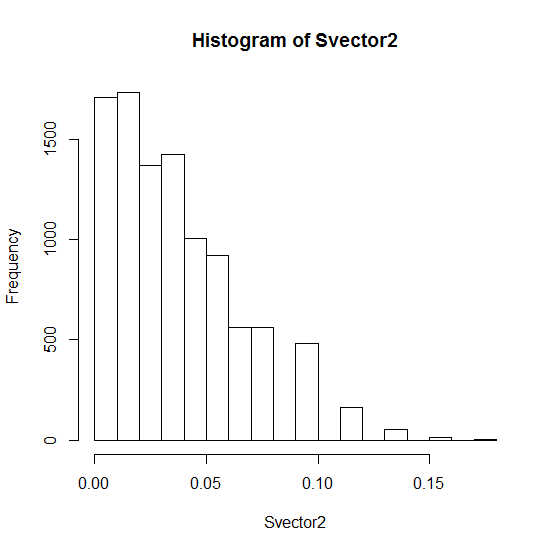
* H0: There is no significant difference between male and female in the proportion of those who math score is 65 or more
* HA: There is no significant difference between male and female in the proportion of those who math score is 65 or more
* S = Difference in proportion of girls and boys who scored 65 or more

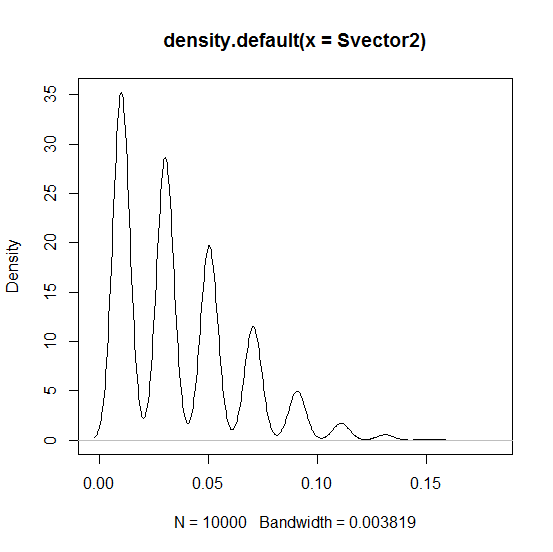
|S| = |(female)-(male)|

|S| = |(13 /109)-(10/91)|

= 0.0094

* Type of test used: permutation





P-value

P (S >= 0.0094) = 0.8294

S calculated = 0.0094

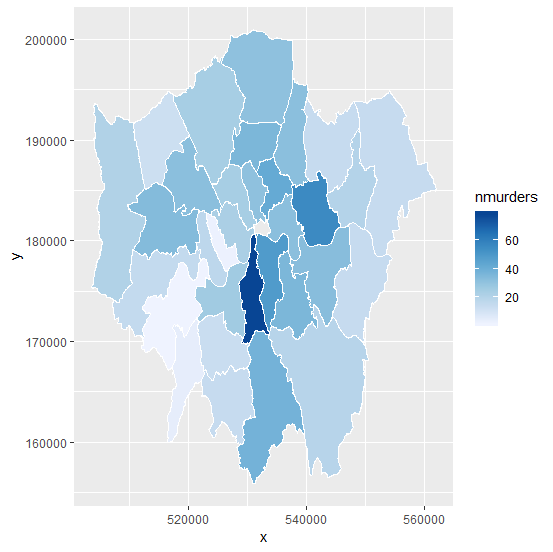
The p value and graph suggest that our H0 is correct and Ha is false i.e there is no significant difference in the proportion of those who math score is 65 or more

Part 4: Does your analysis disprove or support the claim that top math students tend to be male”?

My analysis shows that top math students can be both male or female.

**Question2**

Part 1: Produce a map to visualize the murders by borough



Part 2. Produce a table that counts the number of murders by borough.

|  |  |
| --- | --- |
| Borough Name | Number of Murders |
| Barking & Dagenham | 20 |
| Barnet | 24 |
| Bexley | 14 |
| Brent | 32 |
| Bromley | 19 |
| Camden | 24 |
| Croydon | 38 |
| Ealing | 34 |
| Enfield | 31 |
| Greenwich | 33 |
| Hackney | 42 |
| Hammersmith & Fulham | 17 |
| Haringey | 36 |
| Harrow | 12 |
| Havering | 14 |
| Hillingdon | 21 |
| Hounslow | 15 |
| Islington | 31 |
| Kensington & Chelsea | 2 |
| Kingston | 4 |
| Lambeth | 79 |
| Lewisham | 36 |
| Merton | 13 |
| Newham | 56 |
| Redbridge | 14 |
| Richmond | 1 |
| Southwark | 49 |
| Sutton | 14 |
| Tower Hamlets | 32 |
| Waltham Forest | 32 |
| Wandsworth | 26 |
| Westminster | 23 |

Part 3: Is the count itself meaningful? What other statistics are we missing to compute the murder rate? Go online to find them and compute the murder rate by borough.

The count itself would not be meaningful. It would be useful if we had population by bourough, so we could calculate the murder rate.

From Greater London Authority’s website population average from 2011

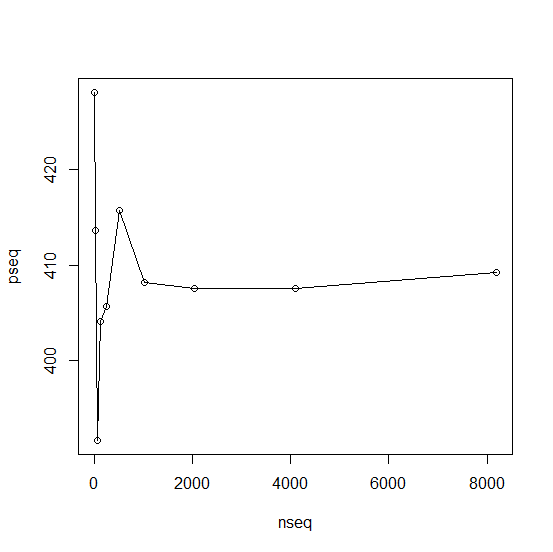
|  |  |  |  |
| --- | --- | --- | --- |
| Borough Name | Number of Murders | Population | Murder Rate per 100k people |
| Barking & Dagenham | 20 | 187,029 | 10.69 |
| Barnet | 24 | 357,538 | 6.71 |
| Bexley | 14 | 232,774 | 6.01 |
| Brent | 32 | 312,245 | 10.25 |
| Bromley | 19 | 310,554 | 6.12 |
| Camden | 24 | 220,087 | 10.90 |
| Croydon | 38 | 364,815 | 10.42 |
| Ealing | 34 | 339,314 | 10.02 |
| Enfield | 31 | 313,935 | 9.87 |
| Greenwich | 33 | 255,483 | 12.92 |
| Hackney | 42 | 247,182 | 16.99 |
| Hammersmith & Fulham | 17 | 182,445 | 9.32 |
| Haringey | 36 | 255,540 | 14.09 |
| Harrow | 12 | 240,499 | 4.99 |
| Havering | 14 | 237,927 | 5.88 |
| Hillingdon | 21 | 275,499 | 7.62 |
| Hounslow | 15 | 254,927 | 5.88 |
| Islington | 31 | 206,285 | 15.03 |
| Kensington & Chelsea | 2 | 158,251 | 1.26 |
| Kingston | 4 | 160,436 | 2.49 |
| Lambeth | 79 | 304,481 | 25.95 |
| Lewisham | 36 | 276,938 | 13.00 |
| Merton | 13 | 200,543 | 6.48 |
| Newham | 56 | 310,460 | 18.04 |
| Redbridge | 14 | 281,395 | 4.98 |
| Richmond | 1 | 187,527 | 0.53 |
| Southwark | 49 | 288,717 | 16.97 |
| Sutton | 14 | 191,123 | 7.33 |
| Tower Hamlets | 32 | 256,012 | 12.50 |
| Waltham Forest | 32 | 259,742 | 12.32 |
| Wandsworth | 26 | 307,710 | 8.45 |
| Westminster | 23 | 219,582 | 10.47 |

Question 3

Part 1: Is there a significant difference in the rate of murders between boroughs? Answer this question by performing a permutation test.

* H0: There is no difference between murder rates between boroughs. Let p1…p32 be the true murder rates of 32 boroughs respectively.
* HA: There is a significant difference in murder rates between boroughs
* S = Difference in murder rates between London and a borough

S = 315.3778



Computed p value P(S >= 315.3778) = 0.04078424

The computed p values is very small, so we can conclude that our hypothesis is false. That is, there is a significant difference in murder rates between boroughs. Some boroughs are safer than others.