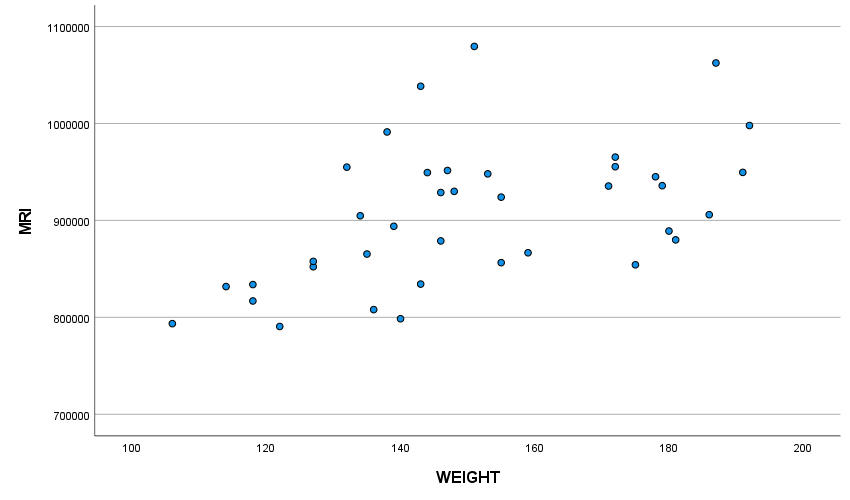
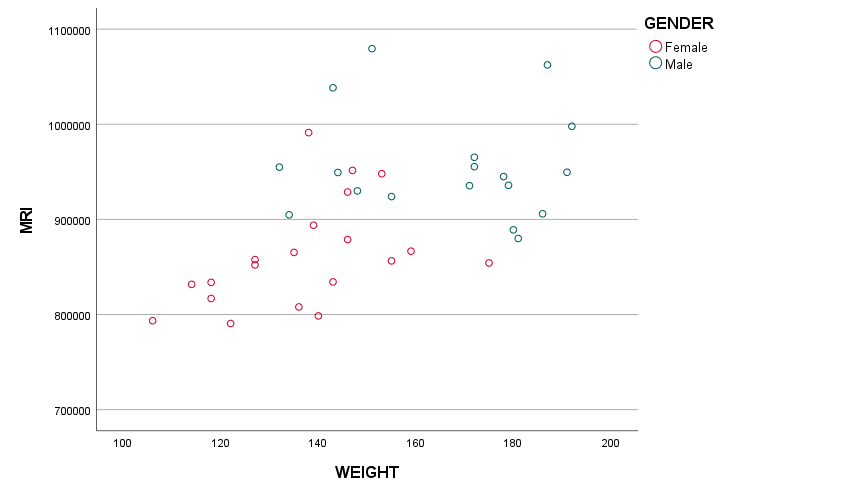
1. 1. The purpose of the study is to find out if there is a relationship between brain size and intelligence. The response variable in this study is one’s PIQ (intelligence) whereas the explanatory variable is the brain size (the variable used for this is the MRI pixel count).
   2. The subjects were randomly sampled for a southwestern USA university pool of righthanded Anglo introductory psychology students who volunteered to be a part of the study. We can make population inferences as this is an observational study and because we randomly selected students from a population. The reason we have a large range of IQ values it provide a more accurate model to be able to deal with a larger range of values.
   3. We can not make causal inferences as this is an observational study. Some other possible explanations for performance IQ is time spent studying and concentration levels. While one can not control concentration levels, we can limit the number of hours spent studying to a certain range or assign students a random range of studying(changing this into an experiment) .

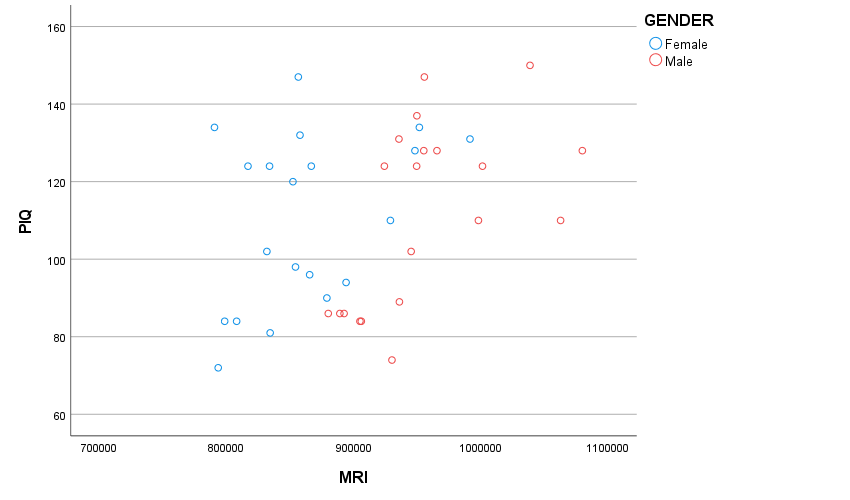
The shape of the above scatter plot shows a somewhat weak positive correlation which indicates a weak postive relationship between MRI and Weight. It does seem to show that it is not necessary to need a large weight to have a a better MRI score as the largest MRI value is acheived by an average body weight.

Comparing the data for males against the data of females, we can see that females have a more linear data(not fully linear but show a weak positive relationship) in comparison to males (Their data does not show a linear relationship at all). but males have higher MRI and weight values than those of females on average. Comparing to part a), females data matched the conclusion we found for a. in the sense that there is a weak positive correlation whereas male data does not match the conclusion for a) at all as they have no correlation at all.

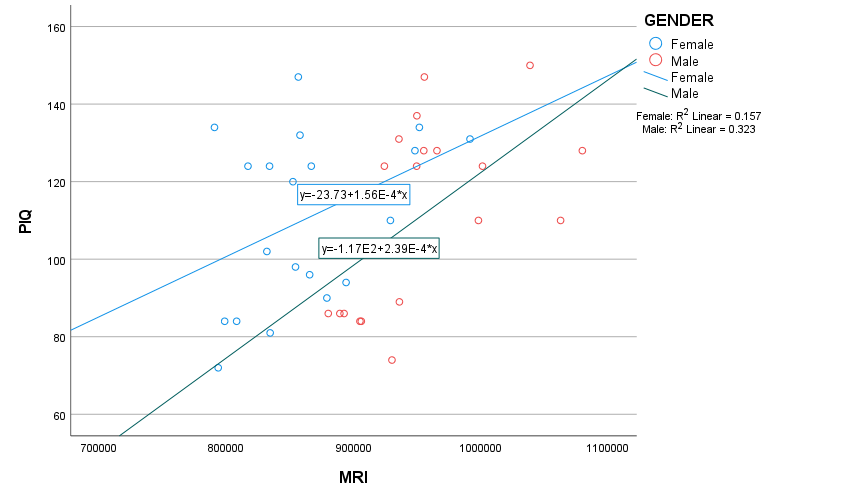
|  |  |  |  |
| --- | --- | --- | --- |
| **Correlations** | | | |
|  | | MRI | WEIGHT |
| MRI | Pearson Correlation | 1 | .513\*\* |
| Sig. (2-tailed) |  | <.001 |
| N | 40 | 38 |
| WEIGHT | Pearson Correlation | .513\*\* | 1 |
| Sig. (2-tailed) | <.001 |  |
| N | 38 | 38 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Correlations** | | | | |
| GENDER | | | MRI | WEIGHT |
| Female | MRI | Pearson Correlation | 1 | .446\* |
| Sig. (2-tailed) |  | .049 |
| N | 20 | 20 |
| WEIGHT | Pearson Correlation | .446\* | 1 |
| Sig. (2-tailed) | .049 |  |
| N | 20 | 20 |
| Male | MRI | Pearson Correlation | 1 | -.077 |
| Sig. (2-tailed) |  | .762 |
| N | 20 | 18 |
| WEIGHT | Pearson Correlation | -.077 | 1 |
| Sig. (2-tailed) | .762 |  |
| N | 18 | 18 |
| \*. Correlation is significant at the 0.05 level (2-tailed). | | | | |

Given the correlations without splitting data between male and female, we see that there is a weak positive correlation (.513) which seems to match what I said for part a). When splitting the data between male and female, we see that there is no relationship between MRI and Weights for males (-0.077) and a somewhat weak positive correlation between MRI and weight for female data(.446).



* 1. Comparing both Male and Female Scatter Points, we see that males seem to indicate a weak positive correlation between MRI and PIQ whereas female data indicate that there is no relation between the two data. Both data seem to go in the positive direction(as x increases y increases).



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Correlations** | | | | |
| Control Variables | | | PIQ | MRI |
| WEIGHT | PIQ | Correlation | 1.000 | .439 |
| Significance (2-tailed) | . | .007 |
| df | 0 | 35 |
| MRI | Correlation | .439 | 1.000 |
| Significance (2-tailed) | .007 | . |
| df | 35 | 0 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Correlations** | | | | | |
| GENDER | Control Variables | | | PIQ | MRI |
| Female | WEIGHT | PIQ | Correlation | 1.000 | .388 |
| Significance (2-tailed) | . | .100 |
| df | 0 | 17 |
| MRI | Correlation | .388 | 1.000 |
| Significance (2-tailed) | .100 | . |
| df | 17 | 0 |
| Male | WEIGHT | PIQ | Correlation | 1.000 | .524 |
| Significance (2-tailed) | . | .031 |
| df | 0 | 15 |
| MRI | Correlation | .524 | 1.000 |
| Significance (2-tailed) | .031 | . |
| df | 15 | 0 |

We can see that the significance when we get the correlation for controlling weight for both genders pooled is 0.007 whereas when we separate the two genders we have a significance of 0.1 for females and 0.031 for males.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Descriptives** | | | | | |
|  | GENDER | | | Statistic | Std. Error |
| PIQ | Female | Mean | | 110.45 | 4.907 |
| 95% Confidence Interval for Mean | Lower Bound | 100.18 |  |
| Upper Bound | 120.72 |  |
| 5% Trimmed Mean | | 110.56 |  |
| Median | | 115.00 |  |
| Variance | | 481.629 |  |
| Std. Deviation | | 21.946 |  |
| Minimum | | 72 |  |
| Maximum | | 147 |  |
| Range | | 75 |  |
| Interquartile Range | | 39 |  |
| Skewness | | -.169 | .512 |
| Kurtosis | | -1.286 | .992 |
| Male | Mean | | 111.60 | 5.264 |
| 95% Confidence Interval for Mean | Lower Bound | 100.58 |  |
| Upper Bound | 122.62 |  |
| 5% Trimmed Mean | | 111.56 |  |
| Median | | 117.00 |  |
| Variance | | 554.147 |  |
| Std. Deviation | | 23.540 |  |
| Minimum | | 74 |  |
| Maximum | | 150 |  |
| Range | | 76 |  |
| Interquartile Range | | 42 |  |
| Skewness | | -.064 | .512 |
| Kurtosis | | -1.350 | .992 |

We can see that males have a higher mean PIQ(111.6) than females(111.45) by approximately 0.15. Males also have a higher standard deviation(23.54) than females (21.946). That means there is a wider range of PIQ values for men than women, However, given that the ratio of the larger Standard Deviation over the lower one is less than 2, we can say the two are similar to each other.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Independent Samples Test** | | | | | | | | | | | |
|  | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | |
| F | Sig. | t | df | Significance | | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| One-Sided p | Two-Sided p | Lower | Upper |
| PIQ | Equal variances assumed | .133 | .718 | -.160 | 38 | .437 | .874 | -1.150 | 7.196 | -15.718 | 13.418 |
| Equal variances not assumed |  |  | -.160 | 37.815 | .437 | .874 | -1.150 | 7.196 | -15.721 | 13.421 |

H0:µMales= µFemales Ha:µMales≠ µFemales

t-statistic=-.160 This means the null distribution of the test statistic is the under the t-distribution with a df of 38.

P-Value=.874

That means as our P-value is greater than our α, we fail to reject the null hypothesis. This means that there is insufficient evidence to say there is a difference between male and female mean PIQ’s

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summary** | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .026a | .001 | -.026 | 22.757 |
| a. Predictors: (Constant), GENDER | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 13.225 | 1 | 13.225 | .026 | .874b |
| Residual | 19679.750 | 38 | 517.888 |  |  |
| Total | 19692.975 | 39 |  |  |  |
| a. Dependent Variable: PIQ | | | | | | |
| b. Predictors: (Constant), GENDER | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 110.450 | 5.089 |  | 21.705 | <.001 |
| GENDER | 1.150 | 7.196 | .026 | .160 | .874 |
| a. Dependent Variable: PIQ | | | | | | |

We have an estimated intercept of 110.450 and our estimated slope of 1.15. Comparing the estimated intercept to the sample mean for the female PID, we can see it is the same and when comparing the sum estimated slope and intercept against mean male PID of 111.6 we can see the sum is 10x that of the mean male PID. The reason for the above two happening is because…

* 1. Our test statistic in this case is .160 with a P-value at .874. This means we fail to reject the null hypothesis which means that there is not sufficient evidence that there the slope for gender is different than 0. The p-values for part b and d are the exact same at .874. The relationship between independent-Samples T Test and the test we performed for regression is that the they both are the exact same thing but with regression we have added variables.

1. 1. PIQ=β0 + β1 \* MRI + ϵ . We have several assumption, our 1st one is that the relationship between the two variables are linear. The two variables must have two equal standard variances. Both must be independent and normally distributed with no outliers within them.
   2. 0.00012\*MRI +1.744=PIQ. That means on average as x increases by 1, y would increase by 0.00012 and when x = 0, we would expect y would be 1.744.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summary** | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .387a | .150 | .127 | 20.993 |
| a. Predictors: (Constant), MRI | | | | |

From this, we can see that our R2 value is equal to 0.15. This means that 15% of the variation of PIQ is explained by MRI.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 2946.613 | 1 | 2946.613 | 6.686 | .014b |
| Residual | 16746.362 | 38 | 440.694 |  |  |
| Total | 19692.975 | 39 |  |  |  |
| a. Dependent Variable: PIQ | | | | | | |
| b. Predictors: (Constant), MRI | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 1.744 | 42.392 |  | .041 | .967 |
| MRI | .000 | .000 | .387 | 2.586 | .014 |
| 1. Dependent Variable: PIQ | | | | | | |

H0: The slope for MRI =0 Ha: The slope for MRI ≠0

Test statistic: 2.586. This means the null distribution of the test statistic is the under the t-distribution with df of 38.

P-Value=.014

Since P value<0.05, we reject the null hypothesis which means that our slope for MRI at the 5% significance level is different from 0

1. 1. For females: 0.000156\*MRI -23.733=PIQ. That means on average as x increases by 1, y would increase by 0.000156 and when x = 0, we would expect y would be -23.733. For males: 0.000239\*MRI -116.844=PIQ. That means on average as x increases by 1, y would increase by 0.000239 and when x = 0, we would expect y would be -116.844

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model Summary** | | | | | |
| GENDER | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| Female | 1 | .396a | .157 | .110 | 20.703 |
| Male | 1 | .568a | .323 | .285 | 19.901 |
| a. Predictors: (Constant), MRI | | | | | |

For females: 15.7% of variance for PIQ is explained by MRI. For Males: 32.3% of variance for PIQ is explained by MRI

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | | |
| GENDER | Model | | Sum of Squares | df | Mean Square | F | Sig. |
| Female | 1 | Regression | 1436.155 | 1 | 1436.155 | 3.351 | .084b |
| Residual | 7714.795 | 18 | 428.600 |  |  |
| Total | 9150.950 | 19 |  |  |  |
| Male | 1 | Regression | 3399.679 | 1 | 3399.679 | 8.584 | .009b |
| Residual | 7129.121 | 18 | 396.062 |  |  |
| Total | 10528.800 | 19 |  |  |  |
| a. Dependent Variable: PIQ | | | | | | | |
| b. Predictors: (Constant), MRI | | | | | | | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | |
| GENDER | Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| Female | 1 | (Constant) | -23.733 | 73.449 |  | -.323 | .750 |
| MRI | .000 | .000 | .396 | 1.831 | .084 |
| Male | 1 | (Constant) | -116.844 | 78.100 |  | -1.496 | .152 |
| MRI | .000 | .000 | .568 | 2.930 | .009 |
| a. Dependent Variable: PIQ | | | | | | | |

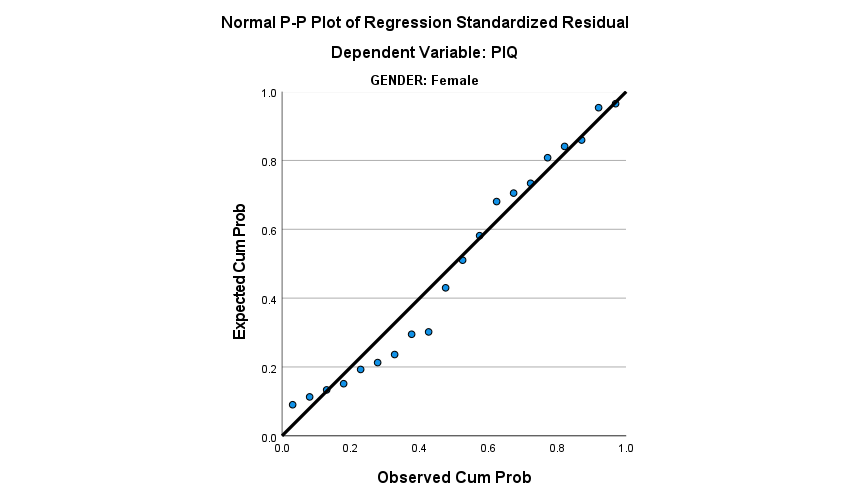
H0: The slope for MRI =0 Ha: The slope for MRI ≠0. This is for both males and females.

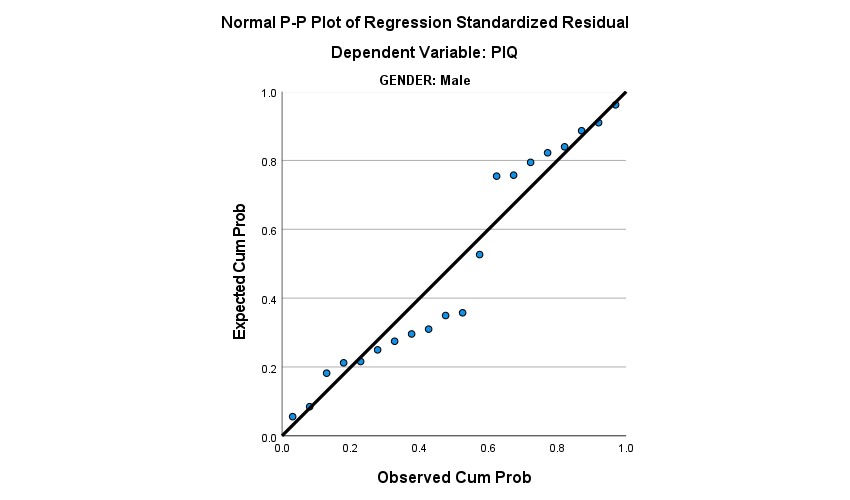
Test statistic: for males is 2.93 while the male test statistic is 1.831. This means the null distribution of the test statistic is the under the t-distribution with df of 18.

P-Value(male)=.009. P Value(female)= 0.084

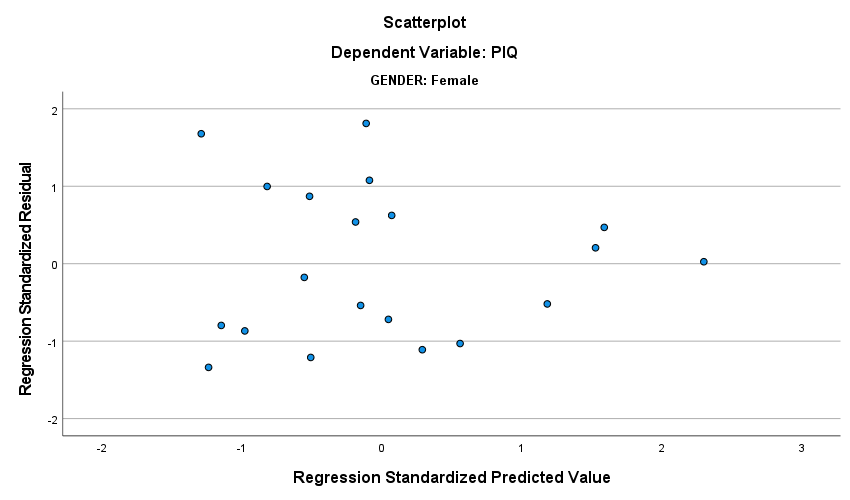
Since P value<0.05, we reject the null hypothesis for both males and females which means that our slope for MRI at the 5% significance level is different from 0 for both male and females.

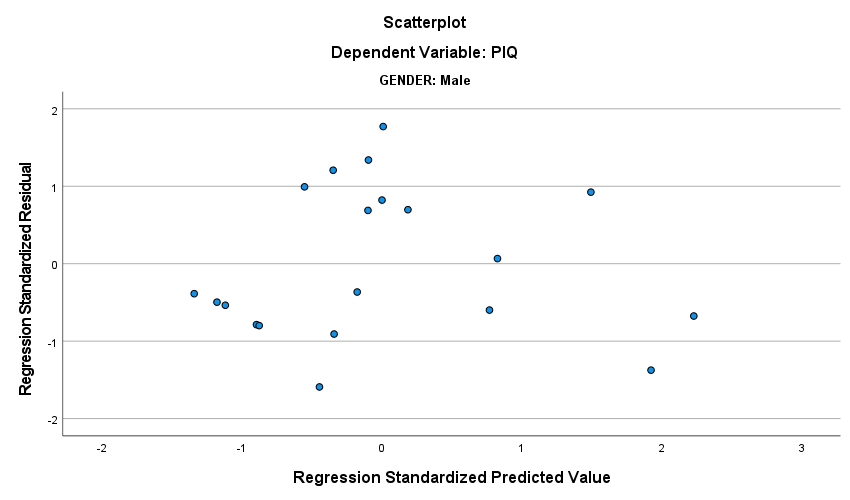






For female data, we can see the points lie a lot closer to the line than those of males with a few points lying a bit far away from the line. The male data suggests a violation in normality due to a lot of points lying far away from the line whereas for females, I would say the data slightly violates normality. Despite, there being more points on/near the line, the points which aren’t are quite far.





For female data, it appears to have a cone-shaped pattern being followed which means the assumption of homoscedasticity is violated for female data. In comparison, the male data seems randomly scattered about a horizontal line around zero so the assumption of homoscedasticity is not violated.

* 1. Estimation= :0.000239\*1001121-116.844 = 122.424.

95% Confidence Interval (110.405, 134.933)

95% Prediction Interval(79.096, 166.241)

From this, we can see that the Prediction Interval is wider than that of the Confidence Interval. Seeing as normality is being violated for males, it is safer to say that the Prediction Interval is more trustworthy.