DePaul University

Brain Size and Predictability

Is it Possible?

Jasmine Dumas

Chris Lee

Samantha Stanley

CSC 423

Professor Bill Qualls

11/20/2015

*Research Question*

In 1981, a study completed by Willerman et. al, collected a sample of 40 right-handed introductory psychology students in attendance at Southwestern University. The researchers sought to determine if an individual’s brain size was an indicator of intelligence. The “subjects were drawn from a larger pool of introductory psychology students with total Scholastic Aptitude Test scores of ≥ 1350 or ≤940” (Willerman 224). The subjects were then split into two equal groups, individuals that had full-scale IQs of less than 108 and those who had full-scale IQs over 130. The researchers took into account gender and body size to draw conclusions about the connection between brain size and intelligence. This analysis will attempt to predict brain size as result of certain independent variables.

*Data Collection and Methodology*

The volunteers then underwent MRI tests at a local testing facility to determine the actual size of the subject’s brain. According to Willerman et. al. the MRI was preformed “using the lowest margin of the cerebellum in a midsagittal view to align the first axial (horizontal) MR slice, 18 mixed-weighted images (spin-echo pulse sequence with a TR of 2000 msec and a TE of 30 msec) were obtained from a Signa MRI unit with a field strength of 1.5 Tesla” (Willerman 224). The count of the pixels from each of the 18 MRI scans were compiled to create the total MRICount which was used as a measure of brain size and is out dependent variable.

From the data collected by Willerman et. al., all of the independent variables were quantitatively measured. The only qualitative variable was gender, which was qualitatively measured as male or female. While we do not know how the original researchers coded this data, for the sake of this research project we chose dummy variables and subsequently coded female as the base. From the data set of 40 entries, there were two incomplete entries, entry #2 and entry #21, which Willerman retained for confidentiality purposes. While irrelevant to our study, it is interesting to note that both of these entries, which withheld the weight (#2) and the weight and height (#21) were both from male subjects. In an era where there is constant pressure from society to look a certain way, we would have suspected the withheld datasets to be from female volunteers. For the purposes of reference, the original dataset can be found in appendix A of this paper.

The dependent variable identified was:

* MRICount (Pixel count)

The independent variables identified were:

* *Gender*: Male or Female
* *FSIQ*: Full Scale IQ scores based on the four Wechsler (1981) subtests
* *VIQ*: Verbal IQ scores based on the four Wechsler (1981) subtests
* *PIQ*: Performance IQ scores based on the four Wechsler (1981) subtests
* *Weight*: body weight in pounds
* *Height*: height in inches

*Limitations of the Dataset*

When working with a small sample size it is important to acknowledge the limitations and issues that can arise from such a small dataset. Willerman et. al. failed to outline why they chose to use a small sample but we attribute this to the size of the subject pool. Furthermore, while we would ideally prefer a larger sample size we accept that the limitations of their research did not permit such. However, it is important to acknowledge that the possibility of a type II error increases when the sample size is small.

*Discussion of Findings:*

*Examination of the scatter plots*

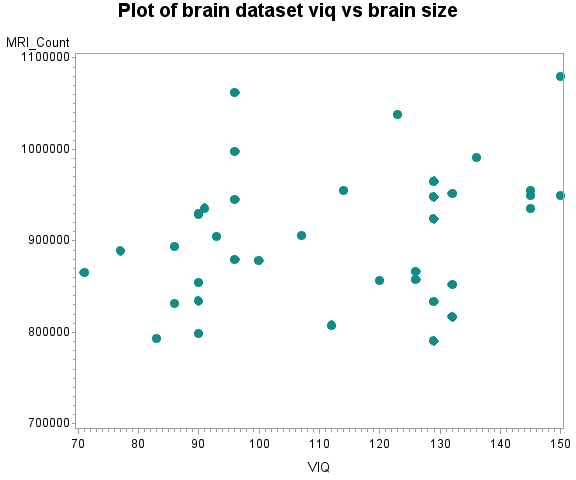


Figure 2: The plot of VIQ versus MRI\_Count does not appear to have a strong linear relationship.

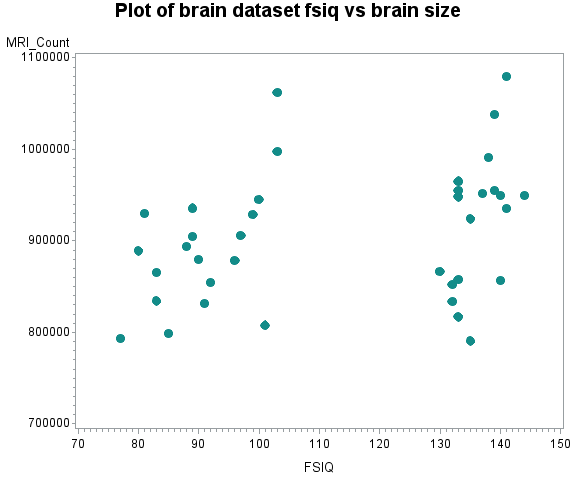


Figure 1: In the plot above we observe the FSIQ versus brain size. The data does not appear to be linear and from the plot we can observe how the data was split by Willerman et al. You can noticeably recognize the 20 subjects with FSIQs less than or equal to108 and the 20 subjects with FSIQ’s greater than or equal to 130.

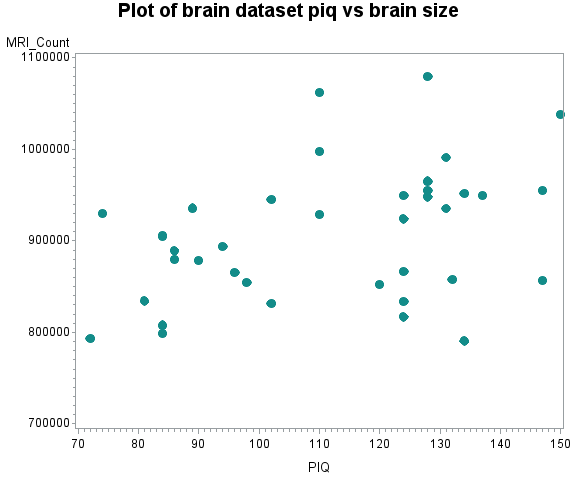


Figure 3: The plot of PIQ versus MRI\_Count does not appear to have a strong linear relationship.

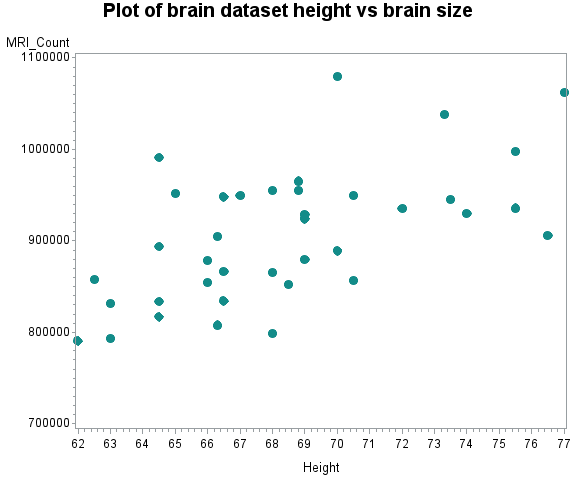
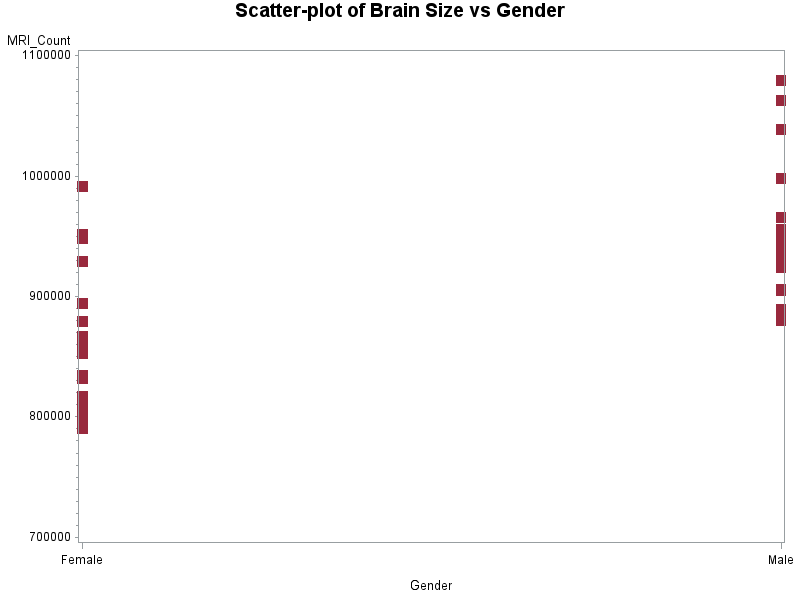


Figure 4: The plot of height versus MRI\_Count does appear to have a linear relationship.



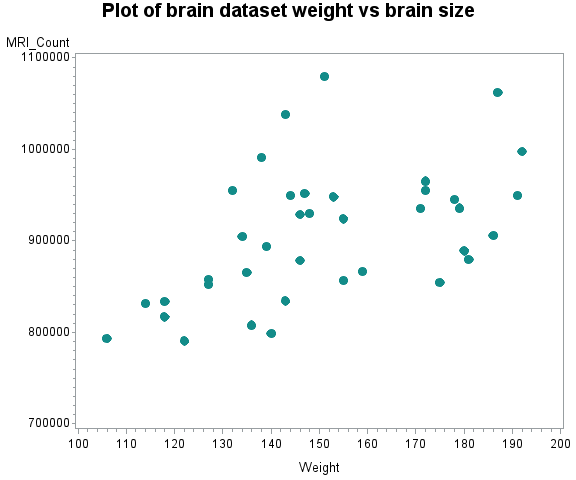


Figure 5: The plot of weight versus MRI\_Count does appear to have a linear relationship.

Figure 6: While this scatter plot of gender versus brain size isn’t necessarily related to our final model, it is interesting to see that males seem to have larger brains than females.

After examining the scatter plots for each of the independent variables it looks as though none of the data needs to be transformed. Furthermore, at this point in the analysis, we did not identify any suspicious data points that seem out of place.

*Building the Model*

Following our examination of our scatter plots, we examined the data for any correlations among the independent variables. We observed strong linear correlations between FSIQ and VIQ, between FSIQ and PIQ, and between PIQ and VIQ. While these correlations, .945, .934, and .776 respectively, might seem high we were not surprise by these results. Given that FSIQ, PIQ, and VIQ are all measures of intelligence, we would anticipate that an individual who scored high on the FSIQ would also score high on the PIQ and VIQ tests.

Additional correlations were observed between x1 (males) and weight as well as between weight and height. Again, we expected these correlations to be exist given that males tend to weight more than females and tend to be taller.

*Full Model*

After examining the correlations, we preformed linear regression on the data in attempts to build a useful model for the prediction of brain size (Appendix C). Running the regression procedure resulted in the creation of our first model with an adjusted R-square of .5847:

However, WEIGHT had a p(value) = .8589 which indicated that it should ultimately be removed from the model. The initial model had a p(value) = .0941 for x1 and a p(value) = .0601 for VIQ thus indicating that x1 and VIQ should be removed from the model. Therefore, our resulting model would be:

The initial model supported that brain size can be calculated by FSIQ, PIQ, and by the height of an individual. However, as previous stated, FSIQ and PIQ have a high correlation and we concluded that additional variable screen methods would be required. As a result, we preformed backward, forward, stepwise, and best fit regression on the full model in attempts to obtain a better model for prediction purposes (Appendix D). These processes resulted in slightly different modes however, we settled the model which was produced by both forward regression and stepwise. That model contained x1, PIQ, and HEIGHT:

After preforming the regression procedure on the modified model the data returned the above model and an adjusted R-square of .5701. Roughly, 57.01% of all variation in the data can be described by this model. The modified model has contains 2 less independent variables. Furthermore, this model managed to address the correlation among the testing scores as we would have ultimately have had to consider removing one of the score variables.

*Interaction and Higher Order Terms*

After the initial model building process, interaction terms were addressed. Ultimately, interaction terms were not identified for either the full model or the model produced through stepwise regression. Additionally, adding interaction terms lowered the adjusted R-square of the model and were deemed unnecessary as none of the terms were statistically significant. Furthermore, we did not observe the need for a higher order model using both the full model and the modified model. The full model contained only 1 parameter, weight that indicated we might want to use a high order term. However, this parameter was shown to be statistically insignificant in the linear model and we did not perform and further analysis.

*Further Examination of the Model*

In the plots of the residuals, we identified some signs of heteroscedasticity in the plots of PIQ versus residuals and predicted y-values versus the residuals. We attempted to modify the PIQ plot using a 1/PIQ inverse transformation and a ln(PIQ) log transformation. However, neither transformation altered the plot in any substantial manner. Unfortunately, we believe that due to the high correlation among the original 5 independent variables and the manner in which the sample was originally collected that there is nothing that we can do to obtain a normal distribution of the residuals. A similar issue occurred when we plotted the predicted y-values against the residuals. We observed a linear trend among the residuals and were unsure how to combat that issue. We also observed less than 5% of the residuals outside of 2s and none outside of 3s which is ideal.

*Outliers and Autocorrelation*

We observed no apparent outliers in the original dataset. However, during the examination of the residuals for our data, there were 3 influential data points identified: #6, #11, and number #13. Each of these points had higher studenstized residual values, however, none of these values exceeded |3|. Observation #13 provided a value of .22 which was over the .21 that we should have observed. However, given that the hat-value and the DFFITS value were within the range that we would expect to see, we concluded that we would not remove observation #13 from the dataset. Additionally, observations #6 and #11 were not removed from the dataset because their influential point statistics fell well within the values that we would expect to observe. (Appendix G)

For the chosen model:

We calculated a Durbin Watson Test Statistic of 1.879 (Appendix H). At k=3 and n=38 this value fell outside of the range that would indicate a correlation [1.32 to 1.66] Therefore, we concluded that there was insufficient evidence of autocorrelation among the residuals for our data.

*Conclusion*

While the reduced model gives an adjusted R-square value of just over 57% we saw signs of extreme heteroscedasticity within our data. We believe that this was due to high correlations among FSIQ, PIQ, and VIQ and additionally among height and weight. Furthermore, we believe that the manner in which the data was collected and the small sample size influenced the heteroscedacity to occur within our model. Perhaps the box-cox model would have proved useful in our manipulation of the data; however, that is beyond the scope of this course. This dataset provided multiple opportunities to utilize all of the regression analysis techniques as we learned in the course, such as:

* removing incomplete data
* removing and detecting outliers and influential points
* recognizing and classifying linear relationships from exploratory scatterplots
* interpreting criterion values to rate the potential of our iterative models
* automated covariate selection methods
* statistical transformations to improve independent-dependent relationships.