Jacqueline Amaya, Hannah Cook, Becky Davis, Rebecca Mills

Math 301

**First-Year GPA Prediction Report Sheet**

In order to create a prediction for your first-year GPA, we gathered 219 samples from first-year students at a midwestern college. The boxplot in Figure 1 displays the data we gathered. This data collection has been used to build a model to predict a first-year college GPA from factors such as SAT score, humanities credits, social science credits, and high school GPA. We strongly believe that this data has provided us with a greater understanding of how well a student will do in their first year of college based on various predictors. Below we have outlined possible outcomes in GPA and credits that could alter the overall first-year undergraduate GPA.

In our analysis of the data collected, we selected a four-predictor model for first-year GPA. Our model predictors are high school GPA (HSGPA), Verbal SAT score (SATV), number of humanities credits (HU), and an indicator for whether or not a student is white (White). With our model, we can predict that a white student applying to this college has a high school GPA of 3.20, a 600 score on his verbal SAT, and completion of 10 humanities credits,  his first year GPA on a 4.0 scale would be 2.97, with a prediction interval with a lower bound of 2.21 to an upper bound of 3.73. Our prediction of a first-year GPA for this student would project a successful first year of college for a white student with the completion of 10 humanities courses, 3.2 high school GPA and a verbal SAT score of 600.

With our data and models run, we have been able to conduct intervals containing a predicted GPA, with a 95% level of confidence. With a high school GPA of 3.20, we have predicted that your first-year college GPA will be between 2.21 and 3.73. This 95% confidence interval only increases as your high school GPA increases. If you were to raise your GPA to a 3.40, your new first-year college GPA would have a lower bound of 2.31 and an upper bound of 3.82. Continuing, a 3.60 high school GPA would result in a college GPA of 2.40- 3.92, with 95% confidence. The upper bound for this GPA is nearly a 4.00, a full GPA in college. In our data analyses, the upper bounds calculated came out to be above 4.00, but please recognize that the GPA of a college student only extends to 4.00. For a high school GPA of 3.8, the GPA prediction interval was 2.50- 4.01. Finally, with a high school GPA of 4.00, your first-year college GPA would 95% confidently be between 2.60 and 4.11. Again, realistically a college GPA is limited to 4.00, but with rising high school GPAs, it is evident that you will receive a rising GPA in your first year of college.

Adding variables to a statistical model can often alter the overall results. We collected data regarding the relationship between a student’s high school GPA and the number of social science credits they have in order to make a better estimate of their first-year undergraduate GPA. Given the information you have already provided above, we will create intervals for the same GPAs above and other variables, while also adding that you have taken 10 social science credits and see how the prediction intervals change. With a 3.2 GPA and the addition of 10 social science credits on top of the previous variables, the interval changed to (2.229526, 3.740715). This was about a .015 increase on average for both the upper and lower bound. This increase continuously gets smaller as the given high school GPA increases. For example, if your high school GPA is 4.0, the 95% prediction interval for undergraduate first-year GPA would be (2.512773, 4.026266). This actually causes a decrease in both upper and lower bounds by about 0.08. To put this in layman’s terms, taking 10 social science classes in addition to your previous statistical standings does not give your GPA much of a boost, and sometimes it can even cause a decrease. This concludes that you might not want to put too much time and energy into social science credits. Your high school GPA will be something more worth it to work on, as the intervals in the previous paragraph show a promising increase in the first-year GPA given a higher high school GPA. 

Appendix

 >gpa.data <- read.csv("http://www.stat2.org/datasets/FirstYearGPA.csv")

> head(gpa.data)

> lm.out.gpa <- lm(GPA ~ HSGPA + SATV + HU + White, data=gpa.data)

> predict(lm.out.gpa,newdata=new.data, interval="predict")

> new.data <- data.frame(HSGPA=3.40, SATV=600, HU=10, White=1)

> predict(lm.out.gpa,newdata=new.data, interval="predict")

> new.data <- data.frame(HSGPA=3.60, SATV=600, HU=10, White=1)

> predict(lm.out.gpa,newdata=new.data, interval="predict")

> new.data <- data.frame(HSGPA=3.80, SATV=600, HU=10, White=1)

> predict(lm.out.gpa,newdata=new.data, interval="predict")

> new.data <- data.frame(HSGPA=4.00, SATV=600, HU=10, White=1)

> predict(lm.out.gpa,newdata=new.data, interval="predict")

> boxplot(gpa.data)