

ACT Predictor for Utah College Students
Daniel Perkins and Savannah Bjorkman
Level of Interest: 10/10

Description

We want to predict the ACT scores of current Utah college students based on the lives they had in high school. Every student has a different background; they have different family situations, different interests, and come from different places. If we can use data about their lives to predict ACT scores, we may be able to determine how to help future high school students improve their test scores.

Data

We hypothesize that the information to predict ACT scores can be found by looking at different attributes of Utah college students from when they attended high school. This includes factors that directly reflect academic success and others that illustrate other priorities and life circumstances. This could include factors regarding family life, extracurricular sports and hobbies, and media consumption. For example, the data could take the following form:

Student	GPA	Amount of AP classes	2 parents in the home	Part time job	Years on Varsity sports	Musical instruments	ACT score
A	3.70	5	Yes	Yes	0	Yes	27
B	3.97	9	No	No	2	No	33

There is not enough space in the table to reflect how many factors we would want to use in our prediction. The nominal factors could include level of mother and father's education, if they received standardized testing tutoring, if they were a member of NHS, the political party of home state, if they spoke another language in the home, if they are a racial minority group, and if they took medication regularly for the majority of their high school years. Real factors could include GPA, amount of siblings, club membership, number of parents in the home, class size, average video game weekly hours, how many times moved out of state, height, and how many committed relationships they had.

Gathering the Data

To gather all of this data, we will create a Google Forms survey and send it out to students at BYU. We will talk to our roommates and have them send out the survey to their friends and family who live in Utah. Also, we will talk to BYU professors and encourage them to share the survey with their students. In order to maximize the amount of students who take the survey, we may include an incentive in the form of a raffle. We also may be able to find a data set through a third-party source, such as a BYU faculty member or an online database.

To convert the data into a usable form for the machine learning regression algorithms, we will put it in a spreadsheet, convert the nominal data into 0's and 1's, and normalize the other inputs.