

## Do Problets Optimally Test Students?

Problets help assist in the learning environment by providing small-scale problem solving concepts to students. Broad introductory programming courses are extremely daunting to new programmers so problets aim to supplement and reinforce the material taught in these courses through repeated exercises.

During the research conducted by Professor Amruth Kumar, responses of these problet topics, including scores, race, and education level, were recorded to judge the effectiveness of each question. My task was to analyze the pretest data and calculate the conditional probability that a student who solves question  $i$  correctly will solve question  $i+1$  correctly. Essentially, the goal of calculating these probabilities was to determine, in a data-driven manner, whether the problem sequence for pretest questions is in the most optimal order. To analyze the data, I created a script using Python, that took the excel data, and split it. Each method corresponded to a new sheet that was built upon the previous data. First, I took the raw excel data, in the first sheet, and searched for the rows that contained pretest data. The script took this data and wrote it to a new sheet, so that the previous data would not be overwritten. This method helped minimize errors by tracking each step of any changes to the data.

After extracting the pretest data, I focused on the scores that every student got, depending on how many questions the problet contained. I extracted all the grades, including those that were skipped and those that were incorrect. I wrote this data to a new sheet as well. After extracting all the grades, I focused on the scores that were answered, even if it was incorrect. Therefore, I excluded all empty cells and only copied the cells that contained scores. I wrote this data to a new sheet, which contained all the grades that were answered. To get into the real analyzation of the data, I had to extract the correct scores from the data, so I searched for scores with the threshold of 1 out of 1, and copied this data to a new sheet in the excel file. Using the correct scores, we wanted to find any details about the data that we could, so I extracted the conditional scores. To find this, I located the cells in which  $i$  was correct *and*  $i+1$  was correct. After this, I extracted cumulative scores which had a similar method as the conditional scores. I found values that were correct in the sequential columns and the previous columns, and extracted only those scores. I wrote this to a new sheet as well. Finally, the main calculations took place in a new sheet, in which I used certain values to calculate the regular, conditional, cumulative, and other types of probabilities.

There are some certain conclusions that can be drawn from the data and there are other uncertainties with the data. For certain topics, especially the three topics that were analyzed thoroughly. For example, one conclusion that we are certain about is the necessity of specific questions. Looking at the results, we can see that for *AirthmeticExpln C++* questions 11 and 12 both had cumulative probabilities of 1 which means that question 11 is not necessary to ask students. Other similar perfect answers proved that certain questions were not necessary in the question pool, so they could be removed.

There are certain things that could be improved with the process to analyze the data. I believe checking the data manually with excel to see if the calculations were correct was inefficient because the script

calculated the probabilities much faster than manually. I understand the idea of checking for mistakes but I did not see a clear use in repeatedly checking the same topics manually.

The conclusions could be stronger if we had more data for some topics. When all the data was extracted, polished, and cleaned, for some topics, there were very few relevant responses. For certain topics, there were a maximum of 30 respondents who correctly answered the questions, without skipping any problems. This is an extremely small pool of students to use analytics that could be considered highly accurate.