Judges' Comments on the 4-6 Project Competition

Some strong projects were entered in this year's competition. The following suggestions are based on this year's judging.

Selecting a question

Selecting a good question for a statistics project is important. Students should clearly define the problem/question they wish to address. They should think about why that problem/question is important to research and how to best collect data to answer the posed problem/question. Questions need not be complicated or complex, but should serve a purpose other than just to collect data. Sometimes the simple questions are the best.

Not only should the question be interesting, it should give rise to data that lend themselves to statistical treatment. For example, if the question leads to a categorical response (i.e., "What is your favorite color?"), one may be left with nothing more than a few counts (one for each category). This limits both the graphical and statistical analyses that can be used. Be sure the question can be answered with the data collected.

Questions need to be stated clearly. If more than one question is posed, each should be answered. Finally, upon completion of the project, it should be reviewed to be certain the question being posed was actually answered.

Design of Study

Students should strive to create or design an experiment or instrument that answers the question. Think about how you would like to collect your data. Will you be able to collect enough data to answer your specific question? If it is a survey, how will you try to get a "random sample" so that you data can be generalized from the sample to the group you sampled? What problems can you think of prior to collecting your data that might make it hard to answer your question? Can you think of ways to eliminate those problems? If you cannot eliminate them, you need to be able to address them in your reflections and conclusion.

Replication is important in any study. For example, the purpose of a study may be to compare the growth of a corn plant with and without fertilizer. Suppose two pots are used and two corn seeds are planted in each pot. Then it is randomly determined which pot gets which treatment (fertilizer or no fertilizer). Even though there are two plants under each treatment, there is no replication. The reason for this is that treatments were assigned randomly to pots (not plants). More than one pot would have to be used for each treatment for there to be true replication.

Collecting data:

Collecting data properly is challenging. Students who find data that have already been compiled often do not realize the pitfalls and potential errors of data collection. As a consequence, they miss an opportunity to understand this vital phase of any project. For this reason, the scoring rubric emphasizes data collection by the students, making projects in which students collect data "from scratch" more highly viewed than those in which students utilize existing data. It is rare for a project based on data that the students did not collect to receive recognition in this competition.

The data collection process should be described clearly, and the student's role in the data collection should be clear. The variables in the study should be defined in terms of what is to

be measured and how. If a random sample is taken, the randomization process should be given. Haphazard or other unplanned sampling is not random sampling and can lead to biased results.

If a survey is conducted, a copy of the survey should be included in an appendix. For all projects, raw data should be included as an appendix.

Graphs:

Graphs and charts should be clear in the data they are displaying and what they are trying to tell the reader about the collected data. All graphs and charts should be properly labeled. Although you might make many graphs or charts to decide how to best display your data, only include those that are clear in their intent.

Many projects fail to take advantage of this important statistical tool. In projects using at least one graphical display, the graphs often are only the most rudimentary pie and bar charts. Stem-and-leaf, dot plots, box plots, and scatter plots are some of the methods that might provide more insight into the data. Displaying sample means with error bars also may be helpful. Care should be taken to use appropriate graphs. For example, line plots and scatter plots are used sometimes when bar charts would be better. Replication permits variability to be captured by the data; appropriate graphs make it visible.

Analysis of Data:

Summarize the data appropriately. Estimates of the center (mean, median, mode) and spread (range) of the distribution are important. The mean, median, and mode may be used to identify a "typical" value. The range provides a measure of variability and helps convey the uncertainty associated with the measurements. Think about how sure you are that the conclusions you made are correct. If you repeated your data collection at a different time or with different participants, do you think you would get the same type of results? If you have time, try that and discuss it in your reflections and conclusions. Would there be a way to extend your work or collect more data so that you could make sure your conclusions are headed in the right direction.

Students should understand fully the methods they use. Sometimes the conclusions did not follow from the analysis presented. It is better to use informal (but appropriate) methods correctly than to apply more sophisticated procedures improperly.

Conclusions and Reflection:

Conclusions should be directed toward answering the question. Clearly explain how the data and the analysis of that data answer the question.

Students should reflect on their project and be honest in assessing the shortcomings and flaws. It is okay if, after the data are collected, you think of a way that could have improved on your project. You need to state those thoughts in your reflections. It is expected that you will think of ways to make your project better if you were to do it again.

Presentation:

A good project and a good write-up take time. If a student hurries to complete a project and write it up at the last moment, it is very difficult for that haste not to be evident in the final project.

Font size should be at least 12 pt., and complete sentences and standard grammar should be used. The emphasis in writing should be on the statistical aspects of the study. Background information should lead to a precise statement of the question to be considered. Some projects benefit from a more detailed description of the data collection phase of the study. Details of the statistical analysis should be presented. The statistical methods should be clearly outlined and discussed. The analysis should serve as the foundation for any conclusions drawn.