



CALIFORNIA SCIENCE & ENGINEERING FAIR 2019 PROJECT SUMMARY

Name(s) Rinoa Oliver	Project Number J0415
Project Title Factors Affecting the California Science and Engineering Fair Results	
<p style="text-align: center;">Abstract</p> <p>Objectives The purpose of this project is to investigate how income, population, grade and school history impact the performance of a project at the California State Science and Engineering Fair. The hypothesis was that winning projects are more likely to come from cities with greater median income and populations as well as be of a higher grade level than other projects. Winning projects are also more likely to come from a school that consistently performs well in the science fair.</p> <p>Methods 1990 to 2018 data was retrieved from the CSEF website using BeautifulSoup for Python. The median household income, population, longitude and latitude were found for each city. Three one-sided Welch t-tests were performed comparing the median city household income, city population and grade of winning projects to those of other projects in a statistical package. Graphs were made comparing income, grade, population for the projects as well as see if there is a correlation with geographic location. Finally, a statistical package was used to create logistic regression models to predict the top four projects from 8 senior continuous categories. The data from 1990 was used to train the models and the data from 2018 was used to test the models.</p> <p>Results Since the p-value for all three t-tests was less than 0.001, the alternate hypotheses that the median city household income, city population and grade level is higher for winners than other projects were accepted. The average city median income of all projects is higher today than in 1990 and there is more of a correlation with these factors and winning than ever before. Schools from an area with higher median income sent on average more projects than schools from an area with a lower median income. Many of the projects were from cities on the coast of California. Some categories were more spread out than others; computer science projects were almost entirely from the coast unlike botany. Some categories such as zoology and botany are losing popularity, while other categories such as biochemistry and computer science are gaining popularity. The logistic regression models predicted 6 winners in 5 categories using only the history of a school in that category.</p> <p>Conclusions The data shows that there seems to be some inequality in the science fair. Winning projects were more likely to be from cities with higher median income and populations and be by students in a higher grade level than other projects. The logistic regression models showed that school history also helps determine a winner,</p>	
Summary Statement Higher city median income, city population, and grade level correlate to projects winning in the CSEF.	
Help Received Dr. Jane Zones of UCSF for advising, Professor Christopher Gould of CSEF for Santa Cruz County Data, Monica Hernandez from Kirby for guidance.	