Correlation between Protein Type and Fatality Rates in Zaire Ebolavirus

Jones, Ellen (School: Haines City IB East)

The student researcher tested for correlation between protein type and fatality rate in Zaire ebolavirus, hypothesizing that there is a significant difference in the fatality rates due to protein types, with the null hypothesis being no significant difference. Records were pulled from the Virus and Pathogen Resource Database for Zaire ebolavirus and sorted by protein type and by UniProtKB Accession. Fatality rates were gathered from the WHO and the CDC and then matched based on the year of collection. Protein types with the highest fatality rates were polymerase (58%), glycoprotein 1, 2 (53%), and small secreted glycoprotein (53%). An Analysis of Variance (ANOVA) was performed upon the data. The student researcher determined that the protein types and associated fatality rates supported the alternative hypothesis that presence of protein type does impact fatality rate in the instance of Zaire ebolavirus. This was concluded because the p-value, $3.05 \times 10^{4}\_71$, is less than the alpha, 0.05. The f input, 24.8, is also greater than the f critical input, 1.63. Due to the difference in the p-value and alpha, the null hypothesis fails to be accepted and the alternative hypothesis is supported beyond the realm of chance.