



**CALIFORNIA STATE SCIENCE FAIR
2007 PROJECT SUMMARY**

Name(s) Craig T. McHugh	Project Number J1310
Project Title Benford's Law: Fact or Fiction?	
<p style="text-align: center;">Abstract</p> <p>Objectives/Goals The goal of this experiment is to determine whether Benford's Law is valid or not. Benford's Law states that lower digits appear as the first digits of a number of a numerical data set more often than the higher digits.</p> <p>Methods/Materials Experimental Method: Collect data sources, Purchase data spreadsheets, Select data sources to test, Record occurrences of first number in a data spreadsheet, Create graphs/tables from data, Compare results to Benford's Law.</p> <p>Materials: Data Pad Spreadsheets, Binder, The World Almanac 2007, The World Almanac 2007 for Kids, A reliable helper, Calculator, Access to Microsoft Excel.</p> <p>Results Of the 34 data sets tested to determine the actual occurrences of the numbers one through nine being the first significant digit or leading digit, the majority of the data sets yielded results consistent with the probability determined in Benford's Law. When the results for each data set used in this experiment were plotted and compared to the slope of the data sets used by Benford, 19 of the 34 data sets, or approximately 56%, were similar to the curve for Benford's Law. Of the remaining 15 data sets, 14 of them had some relationship to Benford's curve, but varied in that some of the numbers were either significantly higher or lower than Benford's curve. Only one data set had results that yielded a curve that did not resemble Benford's curve. For the numbers five through nine the tested results were extremely close to the probabilities in Benford's Law. The combined probability of numbers five through nine occurring as the first significant digit under Benford's Law would be 30.3%. Of the 5,900 numbers tested the numbers five through nine appeared as the first digit 1,782 times or 30.2%, just .1% less than Benford's Law. The curve for the combined totals for all 34 data sets was extremely close to Benford's curve. The testing on the data sets that contained the most numbers had results that were the closest to Benford's Law, and in evaluating the mean results for the total of all data sets, it was observed that as the number of sets tested increased, the results converged with Benford's Law.</p> <p>Conclusions/Discussion My hypothesis was correct and Benford's Law stands valid. Individually, some data sets did not exactly</p>	
Summary Statement I was trying to test if Benford's Law was valid or not	
Help Received # I would first like to thank my dad and mom because they were references to me if I didn't understand something while I was doing research. They also taught me cool tricks on Microsoft Excel that quickened the process of making all my graphs, and they purchased my data sources for me.	