

The Design and Development of a Relationship Graphing Tool for Biological and Clinical Parameters

Background

The variety and amount of data collected have been increasing to better understand and conduct basic and clinical research. However, parsing through a large amount of data and finding insightful relationships between different parameters can be time-consuming. It would also be inefficient to create graphs for the sole purpose of looking for potential correlations. To optimize the process of data mining, this project aims to develop a relationship graphing tool that provides a visual of some of the basic relationships (linear, exponential, logarithmic) between the parameters selected from a curated database. This tool can inform a general direction for a research topic, generate quick and unofficial graphing, and efficiently select biological and/or clinical parameters.

Tool Description

The tool will calculate the equation of the best-fit line and generate a graph based on the input dataset. To carry out the functionality, this tool requires the following software technologies:

- SQL relational database/table: stores the information on biological and clinical parameters about different diseases in the form of tables. This information will be manually gathered and curated from publicly accessible databases, such as Alzheimer's Disease Neuroimaging Initiative.
- Python-based CGI: parses the information and plugs the input dataset into different data models (linear, exponential, and logarithmic). The Python code will calculate the equation of the best-fit line and generate statistics, such as the r value and standard deviation, to illustrate the line fitting.
- CSS/HTML/JavaScript: provides an interactive interface for the user to select input data (age, gene expression, disease stage, etc.) and generates graphs based on the input. JSON will be used to package and exchange data between JavaScript objects and the server.