Data Science Math Final Project

In the SQL ans R portions of the final project, we create databases that demonstrate the relationships between users, libraries (and library groupings), and the materials the users have access to. We chose to create these databases based on inspiration provided by Hawaii Library data. We use data collected on the libraries to showcase what we learned using R to create visual displays from the data.

But why is this data being collected? What is the driving force behind the the data visualization? Libraries collect and maintain all kinds of data and compile statistics for a number of reasons. One reason is to demonstrate the value that libraries provide. They also use the assessments to demonstrate effective management of resources. Another important use of library statistics is to show policy makers and funders, for purposes of strategic planning and decision making on value added services.

Measuring the funding put into library systems displays the commitment and engagement of stakeholders. Measuring output, for instance, circulation data, shows that the services are appropriate for the intended population. It can also be used to look at trends in use, to see if, for instance, users are increasingly borrowing digital as opposed to print media.

Data science math can be used to analyze if there is an unusual upswing in activity related to things like time of year. For instance, plotting the number of movies borrowed by month may show an increase in the colder months, indicating that is a good time to add new material to the collection. It can also show that perhaps more homework help resources are needed around finals time. This can help with budget planning, and staffing issues. Data science math can also be used for predictive analytics, for things like whether more computers will be needed to meet future demand, whether more materials should be purchased on certain topics, depending on the shifting needs of the population served.

The math behind data science allows libraries to attempt to meet community needs in real time, assessing what people are asking for, and are interested in over time. It helps them to utilize their physical space and monetary resources to the fullest advantage, by culling collections of little used material (like outdated media formats), and staffing more at busier times, and also in sharing resources across communities to save resources. There are innumerable ways that data can be harnessed to make improvements in library services, and the math behind the visualizations makes it all possible.