

Program Development Exercise: Engineering

Background:

You are the library liaison for the Engineering departments of High Tech University. You are walking through the Civil Engineering department one day and decide to stop by Prof. Ortega's office for a brief chat. Prof. Ortega studies real-time traffic signal performance measures in which he observes the movement of traffic, specifically the number of vehicles passing through an intersection and the amount of time they spend at an intersection over a 24 hour period. He uses this information to generate "profiles" of traffic movement at intersections. These profiles also take into consideration contextual elements including weather conditions, local zoning or the surrounding area (heavy commercial, light residential, etc.) and the condition of the roads. His objective is to improve traffic conditions through urban areas by studying and accounting for factors that either improve or impede traffic flow. He works closely with the state department of transportation (his major source of funding), and his data is used to make decisions on public works spending and as a means of informing public policy.

You find him squinting at a spreadsheet on his computer and muttering to himself. He explains that one of his graduate students just sent him some data in preparation for a meeting that they will be having later that day and he cannot make heads or tails of the numbers shown in the spreadsheet. You start up a conversation with Prof. Ortega about how his graduate students are generating and managing data.

Potential Areas of Need:

Prof. Ortega tells you that his graduate students are all over the map in terms of their skills and abilities in handling data. A few of them do a nice job of tracking the work that they do and explaining how and why they do things, others do not. Most of his students make some annotations in the data to explain perceived anomalies, but their annotations vary in their descriptive quality. He does try to guide his students during their one on one meetings, but sheepishly admits that he does not always have the time to review the students' work with the data with as much depth as he would like. Another area where his students could improve is on their use of visualizations. Although they are generally knowledgeable in generating tables and graphs (or at least they seem to be able to figure it out), they don't always use these visualizations very effectively to convey their findings.

He delivers a copy of his data to the state department of transportation, but they do not ask him to follow a particular standard or format. The state will occasionally ask some questions about the data, which he will answer or refer to his students. Prof. Ortega believes that his data could be very interesting and potentially very useful to the public. He would definitely support sharing with others but is not sure how the state would feel about that.

Software and file types:

Prof. Ortega places sensors at each lane of traffic at an intersection. These sensors produce a raw data stream of data measuring traffic flow in a format that is proprietary to the sensor. The raw data is exported into a .csv files which are then cleaned up and added into a MySQL database. Prof. Ortega's



Jake Carlson – PI (Purdue), Camille Andrews (Cornell), Marianne Bracke (Purdue), Michael Fosmire (Purdue), Jon Jeffryes (Univ. of MN), Lisa Johnston (Univ. of MN), Megan Sapp Nelson (Purdue), Dean Walton, (Univ. of OR), Brian Westra (Univ. of OR), Sara Wright (Cornell)

Program Development Exercise: Engineering

lab supports multiple mySQL databases and each one holds between 3-4 months' worth of data. Data are extracted from the mySQL database into MS Excel for creating graphs and charts. Other components of the data set include videos taken of traffic flows at the intersection, a database of road conditions maintained by the state transportation department and data about weather conditions that may be needed to explain deviations in data points from what was expected.



Jake Carlson – PI (Purdue), Camille Andrews (Cornell), Marianne Bracke (Purdue), Michael Fosmire (Purdue), Jon Jeffryes (Univ. of MN), Lisa Johnston (Univ. of MN), Megan Sapp Nelson (Purdue), Dean Walton, (Univ. of OR), Brian Westra (Univ. of OR), Sara Wright (Cornell)