

*“Why are things as they are and not otherwise?”*

Johannes Kepler (1571-1630) German astronomer

Large complex data sets can be challenging to analyze. Enormous amounts of data are being collected, stored, and analyzed in an ever increasing array of fields. In transportation, this evolution is no different. As vehicles increasingly become probes — connecting subsecond vehicle pulses of data containing position, speed, throttle, brake, and other monitoring information — operational data will only expand. On the planning side, the use of smartphones and other devices constantly reporting position and travel patterns will be invaluable sources of data to better plan, operate, and design our cities and rural environments. Even seemingly ubiquitous data streams such as information from traffic signals, freeway sensors, and bus fleet information produce very large datasets in a short amount of time. These large datasets quickly exceed the capacity of traditional analysis tools that most engineering students are very familiar with, such as Excel.

The purpose of this chapter is to introduce the concepts of data, databases, and structured query language (SQL). The chapter begins by using a very familiar data, in Excel, to demonstrate some very basic concepts of data. From this strawman, students are introduced to an open source database: PostgreSQL. Students are introduced to data types, tables, schemas and then create their first database. Finally, students are challenged to attempt the use of SQL language in order to generate simple queries that will assist with future activities.

## ACTIVITY LIST

Activity Type	Number and Title	Assessment Type
Out of class	Activity 5: Setting Up Your Accounts	Participation
In class	Activity 6: Creating a Simple Database in Excel	Short response
Out of class	Activity 7: Overview of SQL	Short Response
In class	Activity 8: Creating a Simple Database in PostgreSQL	Short Response
In class	Activity 9: Writing and Testing Some Simple SQL	Participation

