

1096-D5-1770

**Nicholas J Horton\*** ([nhorton@amherst.edu](mailto:nhorton@amherst.edu)), Amherst College, Department of Mathematics, Box 2239, Amherst, MA 01002, and **Benjamin S Baumer** and **Hadley Wickham**. *Big Data in the Intro Stats Class: Use of the Airline Delays Dataset to Expose Students to a Real-World, Complex Dataset*. Preliminary report.

Students in the introductory statistics course need exposure to bigger datasets and more complex questions to be able to make sense of the increasingly data-centric world that they will inhabit. In this talk, I will describe how the airline delays dataset (150 million records on all commercial flights in the US from 1987 to 2012) can be integrated into the introductory statistics course. This large dataset is introduced early in the semester through a model eliciting activity due to Garfield and colleagues that leads students to undertake informal inference when comparing the on-time performance of two airlines servicing a pair of airports. Later in the course, students are able to assess the performance of their comparison rule by repeatedly sampling from the underlying population of flights, as well as visualizing the population using straightforward commands in R to access the database using simple SQL commands. The techniques are facilitated by use of R Markdown. In addition to providing insight into flight delays, the activity helps expose students to the power of statistics to make decisions in the face of uncertainty. (Received September 16, 2013)