Martina A. Pillay

GA Data Science

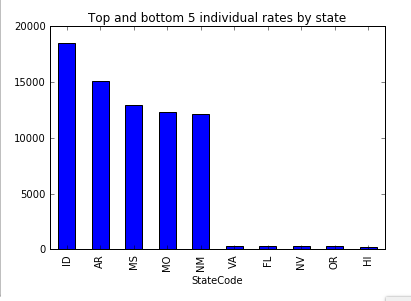
May 19, 2016

**Problem Statement:**

There is little transparency into health care costs; co payments and premiums can vary across health plans and states and depend on the demographic of the patient. I would like to examine the Health Insurance Marketplace data released by the Center of Medicare and Medicaid Services (CMS) which contains health and dental plan information across the US. Specifically, I would like to better understand the relationship between the insurance rate and the state, insurer, and age.

**What I’ve done so far:**

* Data retrieval: available in csv format on the CMS website <https://www.cms.gov/cciio/resources/data-resources/marketplace-puf.html>
* Some data exploration: on my github page (the data is not on the github because the files are too large) <https://github.com/martinaap/health-insurance-marketplace>
* Example of mean premium rate differences across state (will most likely turn this into a map:



**To be completed:**

* Finish data exploration of all files
* Define features to use and begin analysis using different statistical models