VisualDX capstone project 11/11/16 Documentation

What I did:

* Subset events table by year in MySQL so that queries are faster and don’t use all my disk space
* Got random sample of each subset containing 10% of events from that year (About 1.2 million observations per year)
* Made large CSV file combining all of these

Goal: Want to plot feature relationships and distributions of features to see if event and inquiry tables can be modeled similarly.