At this point, you have obtained the data set for your Capstone project, cleaned and wrangled it into a form that's ready for analysis. It's now time to apply the inferential statistics techniques you have learned to explore the data. For example, are there variables that are particularly significant in terms of explaining the answer to your project question? Are there strong correlations between pairs of independent variables, or between an independent and a dependent variable?

**Submission:** Write a short report (1-2 pages) on the inferential statistics steps you performed and your findings. Check this report into your github and submit a link to it. Eventually, this report can be incorporated into your Milestone report.

Auction P = 4.7 e-9

Time P = 0.01

Url P = 0.43

bids /auction P = 1.7e-8

Countries chi2 p = 0

Once we have analyzed that data and created visualizations of it to search quickly for possibly relevant categories we apply statistical methods to determine which categories and correlations are significant. The main statistical tests we will focus on in this analysis are t-tests and chi-squared tests. The quantities up for analysis are the unique number of bids per auction, unique urls and ips used, total number of bids, and the average number of bids per auction from a user. By comparing the differences in the bot and human populations between these categories we will be able to identify the most important qualities for distinguishing bots and humans.

T-tests

First I will discuss the results from the t-test analysis. T-tests were used to analyze the majority of the categories. By splitting the data into bots and humans and then grouping by bidder id populations of each category are created.

(insert table head)

The populations of interest are auctions, time (representing total number of unique bids), url and a self-created feature bids per auction.

Chi-squared tests

Conclusions

After having created features in out data such as a bids per