

Name: Tofik Mussa

Submitted to: Dr. Paul Bamberg

Final Project

My application is a tool for performing an exploratory data analysis on a dataset which may be useful for statisticians to derive key insights out of. The dataset is composed of most popular Google search words in US states/regions and personality profiles. I am trying to predict personality types as a function of what people spend the most time on the internet searching

US regions were classified as one of

- 1) Friendly and Conventional
- 2) Relaxed and Creative
- 3) Temperamental

Extraversion, agreeableness, conscientiousness, neuroticism and openness played into these classifications.

Luckily the dataset was scaled to relative proportions which made the analysis straight-forward. I conducted the following analysis on the data

- 1) Contingency table
- 2) t test and pairwise t test
- 3) Analysis of variance
- 4) Linear regression
- 5) Finding outliers with boxplots
- 6) Density plots, bar plots, scatter plots
- 7) Correlation coefficient analysis
- 8) Hierarchical clustering for classification
- 9) Quantile regression to test without outliers
- 10) Chi-Square test
- 11) Correlation graph

I am a candidate for ALM in Software Engineering at the Extension school and has done course on front end web development and statistics in the past. I have tried putting both skills into a test with this application. I have applied most of what I learnt in the weeks where we covered data analysis and general R/R Shiny programming techniques into the project. Here are some of the topics I applied into the project from this course

- 1) Chi-Square test
- 2) ggplot and plotting in general
- 3) Contingency tables
- 4) Reading data from file
- 5) Dynamic rendering
- 6) functions
- 7) Global/local scopes
- 8) Dataframes
- 9) Vectors
- 10) How to wrangle data
- 11) Grouping/filtering operations
- 12) Correlation analysis
- 13) Linear regression
- 14) t Tests
- 15) Finding outliers

URL - <https://tofikmussa.shinyapps.io/exploaratory-data-analysis/>