The Perfect Fit

STAT 385 SP2019 - Team 24

Jacob Hart (jbhart2)
Landon Hess (lahess)
Jami Tugas (jtugas2)
Joseph Wang (josephw5)

May 9, 2019

Abstract

This project examines possible variables that might affect the desirability of a living in a state, including variables such as climate, population density, and median home value. This allows a user to find the most affordable area to live in while simultaneously satisfying their personal preferences of what type of environment they value.

Contents

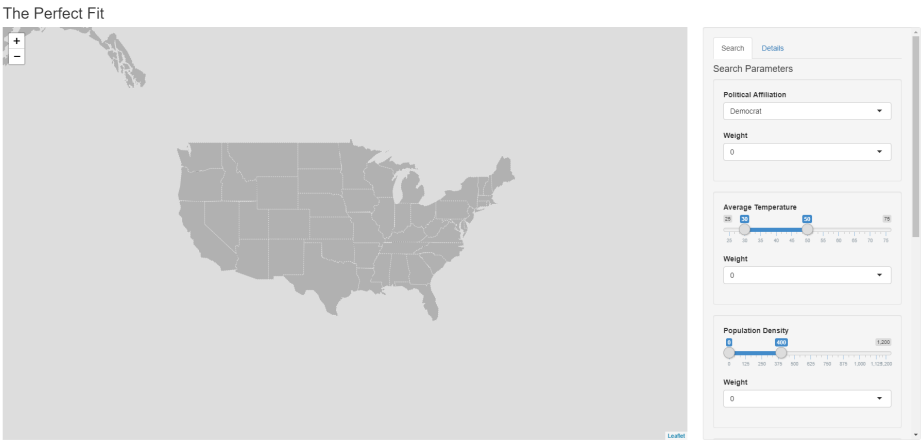

Figure 1: UI Default

Introduction

For our Project, we decided to make a program that makes it easy for the user to find a state in the United States that fits their desires. When people are choosing where to live, there are some standard factors that most people focus on in order to choose the best county to live in (*Best Counties to Live in America* 2019).

With our program you are able to customize the parameters of multiple factors, and immediately see a map of the results that match your selections. This program uses `R` and `shiny` in order to make an easy to use and highly beneficial tool for people that are looking for states with certain attributes. The data being used is taken from multiple sources that have different information about each state in the United States. These factors include Political Affiliation, Average Temperature, Density, Cost of Living, Median Home Income, and School Rank. The user is able to select ranges and weights for each one in order to find the best state for them.

Methods

When the program starts, data is imported by web scraping. Web scraping for each variable begins with `rvest` (Wickham 2016). The HTML of each page is parsed into data frames. Variables of interest are selected from the data sets and combined into a master data frame.

On the UI side, the main section is an interactive map of the United States, created with `leaflet` (Cheng, Karambelkar, and Xie 2018). There is a sidebar with two tabs. The first tab is a search panel. The user is able to select bounds for variables of interest (Avg. temperature, Crime rate, etc.), apply weights for how much they care about the variable, and search for states that match these criteria. A model is formed to light the states green based on how well they match the search parameters. States are completely gray do not match the selected parameters at all, while completely green states are a perfect fit. This communicates to the user which states are best for them. At any time, the user is able to select a state, which automatically switches to the second tab, Details.

In the details tab, there is a stacked bar graph to show political affiliation, and gauges created with `flexdashboard` (Iannone, Allaire, and Borges 2018) that show every variable of interest for the selected state.

Best Counties to Live in America. 2019. Niche.com Inc. https://www.niche.com/places-to-live/search/best-counties/.

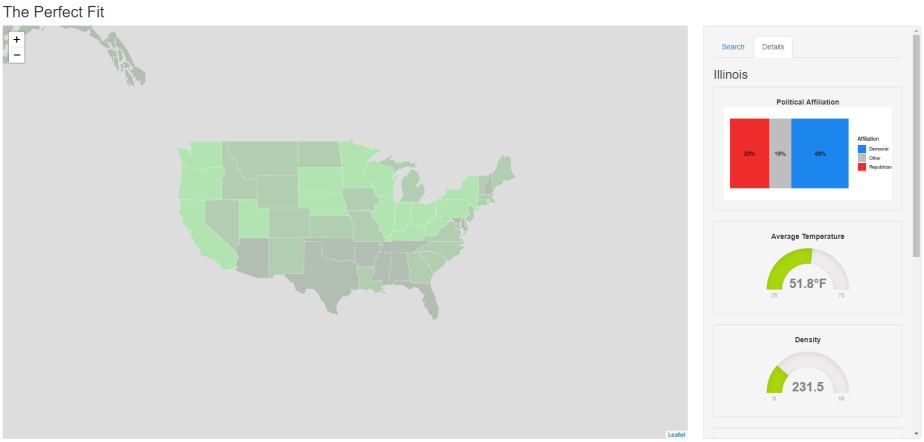

Figure 2: UI Details

Cheng, Joe, Bhaskar Karambelkar, and Yihui Xie. 2018. *Leaflet: Create Interactive Web Maps with the Javascript 'Leaflet' Library*. http://rstudio.github.io/leaflet/.

Iannone, Richard, JJ Allaire, and Barbara Borges. 2018. *Flexdashboard: R Markdown Format for Flexible Dashboards*. http://rmarkdown.rstudio.com/flexdashboard.

Wickham, Hadley. 2016. *Rvest: Easily Harvest (Scrape) Web Pages*. https://CRAN.R-project.org/package=rvest.