Sydney Louit
Ziqi Li
Zewei Ren
Luna Shen

BIST/STAT 5225 Group Project: CDC Crime Rate Analytics App

1 TECHNICAL PART

Packages

The following R packages are necessary for everything to work:
xlsx - Package to read .xlsx files
sf - Package to read shapefiles
leaflet - Package to plot interactive maps
leaflet.minicharts - Package to update and modify maps
dplyr - Package for data manipulation and transformation
rgdal - Package for geographical data manipulation
htmltools - Package to display text in HTML format
shiny - Package to display RShiny App
shinydashboard - Package to section the RShiny into a customizable three-part dashboard

Packages can be installed using the install.packages() function, and can then be accessed using the library() function, (i.e. library(leaflet))

Dataset

The dataset was formed after combining the datasets from 1999-2019 into one dataset containing the crime rates for each state, crime, and year. The dataset only contains statewide crime rate data (No total crime or MSA-level statistics). Puerto Rico does not have data for 2010, so a new (blank) row was added. In total, the dataset contains 1092 observations of 11 variables.

Shapefile

The shapefile originates from the U.S. Census Bureau, and includes the District of Columbia and Puerto Rico. The shapefile was condensed using mapshaper.org, and then was loaded as a data frame using the st_read() function Then, territories with no data, such as Guam, were eliminated, and the resulting file was sorted alphabetically by jurisdiction.

Functions

Crime rate map

The rate data is subsetted for the selected year and selected crime. Once these values are selected, a color scale is generated based on the range of the data, then fitting the result in a way that the lowest crime rate will appear in white and the highest crime rate will appear in red. Consequently, the map can appear to be in shades from white to red, depending on the crime rate value of a specific year and state.

Using the built-in ColorRampPalette() function with a large number of breaks, it is possible to construct a color scale once the data has been transformed into a scale between 0 and 1.

Time series plot

To conduct a time series plot for a specific crime type in a user-selected state, we do it in two steps.

Firstly, we generate a subset of crime data for the user-selected state, and then plot the scatter plot for the specific crime type against years. To show both point and line in this plot, we add option ‘type=”b”’.

Change of crime rate map

The rate data is subsetted for the two selected years. Once that step is completed, the data from the later year is divided by the data from the earlier year. Once these values are computed, a color scale is generated by taking the log of the data, then fitting the result in a way that increases in crime rate will appear in red and decreases will appear in shades of green. Consequently, the map can appear to be entirely in shades of green or entirely in shades of red, depending on whether crime rates have increased or decreased in the time span. The start and end years can

Visualization-R Shiny

There are three basic components in R shiny: user-interaction (UI), server and running function.

In the UI part, we use the ‘titlePanel’ function to create a title for the results page. Then, we generate four user-input variables, ‘state’, ‘crime type’, ‘year’ and ‘range of year’, through the ‘selectInput’ function, which allows the user to select one value of the variables, and the ‘sliderInput’ function, which allows the user to use the slider to choose the specific year. After that, through the ‘tabPanel’ function, we assign three results panels for the results we will generate later.

In the server part, we assign the three functions we’ve already conducted to the results panels we created before.

Finally, we use the ‘shinyAPP’ function to run the code.

GitHub

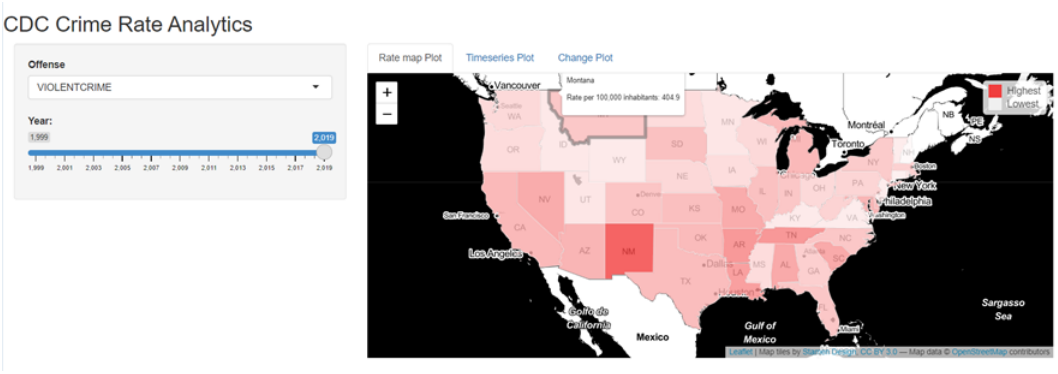
The repository is at sydneylouit539/uscrimerate . From April 5th onwards, GitHub was the primary mode of collaboration on the project.

2 CLIENTS-ORIENT PRODUCT MANUAL

Instructions

Step1: open shiny.R and run the app.
Step2: click on the html file.

Crime Rate Map



Benefits:

1. Observe specific crime types all over the country easily and compare these crime rates between states intuitively.
2. Clients could get the accurate crime rate through mouse interaction.

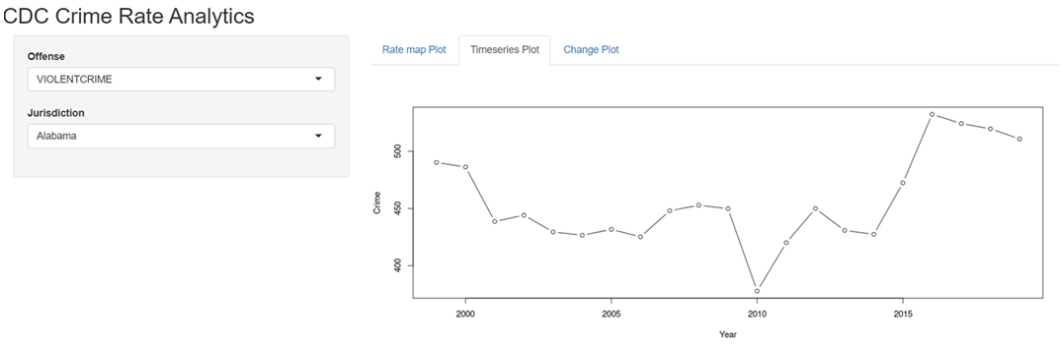
Drawbacks:

1. Cannot notice the year-to-year trend for a specific crime in states.
2. Cannot find the exact difference for a specific crime over years.

To rectify these particular drawbacks, we also created a time series plot and a change of crime rate plot.

Time Series Plot

In this plot, we can notice a trend for a specific crime in each state clearly. The U.S. Government's definition of rape changed in 2017, so most states will show a large spike around that time when viewing rape data.



Change of Crime Rate Map

In this map, we can find the exact difference for a specific crime over a given time period. Additionally, the colors make it very easy to see where the crime rate is increasing or decreasing. However, as with the time series plot, it is important to note that the U.S. Government's definition of rape changed in 2017, so the change maps may be misleading for that particular offense.