DSC640_Exercise_4-2_JakeMeyer

Jake Meyer

2023-01-28

Assignment Instructions

Submit 1 scatterplot, 1 bubble chart, and 1 density plot with R.

Show Working directory

```
getwd()
```

```
## [1] "C:/Users/jkmey/Documents/Github/DSC640_Course_Assignments/DSC640_Repository/Weeks7&8"
```

```
dir()
```

```
##  [1] "~$DSC640_Exercise_4-2_JakeMeyer.pptx"
##  [2] "birth-rate.csv"
##  [3] "birth-rates-yearly.csv"
##  [4] "birthdensity.txt"
##  [5] "crimerates-by-state-2005.csv"
##  [6] "Data+Visualizations+-+DataCamp.pdf"
##  [7] "DSC640_Exercise_4-2_JakeMeyer.ipynb"
##  [8] "DSC640_Exercise_4-2_JakeMeyer.pbix"
##  [9] "DSC640_Exercise_4-2_JakeMeyer.pptx"
## [10] "DSC640_Exercise_4-2_JakeMeyer_Power_BI_Images.pdf"
## [11] "DSC640_Exercise_4-2_JakeMeyerRCode.Rmd"
## [12] "DSC640_Exercise_4-2_PythonCode.pdf"
## [13] "life-expectancy.csv"
## [14] "tv_sizes.txt"
```

```
 # setwd("C:/Users/jkmey/Documents/Github/DSC640_Course_Assignments/DSC640_Repository")
```

Import the necessary libraries

```
library(readxl)
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## v purrr   0.3.5
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(dplyr)
library(scales)
```

```
##
## Attaching package: 'scales'
##
## The following object is masked from 'package:purrr':
##
##      discard
##
## The following object is masked from 'package:readr':
##
##      col_factor
```

```
library(plotly)
```

```
##
## Attaching package: 'plotly'
##
## The following object is masked from 'package:ggplot2':
##
##      last_plot
##
## The following object is masked from 'package:stats':
##
##      filter
##
## The following object is masked from 'package:graphics':
##
##      layout
```

```
theme_set(theme_minimal())
```

Import the data

```
df1 <- read.csv("crimerates-by-state-2005.csv", header = TRUE, sep = ",")
```

Scatterplot using crime rate data by state from 2005.

```
df2 <- df1[df1$state != 'District of Columbia',]
df2 <- df2[df2$state != 'United States',]
```

```
scatter.smooth(df2$murder, df2$burglary, xlim = c(0,10),
                ylim = c(0, 1400), xlab = 'Murder Rate (per 100,000 population)',
                ylab = 'Burglary Rate (per 100,000 population',
                main = 'Scatterplot of Burglary Rate vs. Murder Rate in U.S.')
```

Scatterplot of Burglary Rate vs. Murder Rate in U.S.
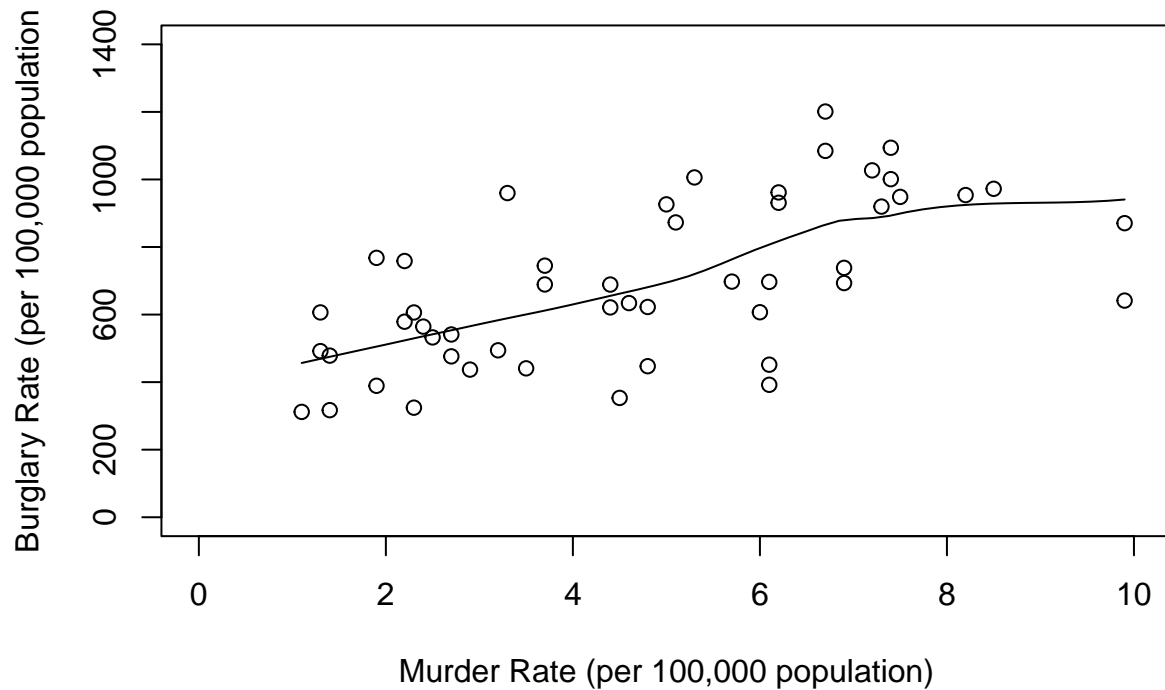
Bubble Chart using crime rate data by state from 2005.

```
radius <- sqrt(df2$population / pi)
symbols(df2$murder, df2$burglary, circles = radius,inches = 0.35,
        fg="white", bg="red", xlab = "Murder Rate (per 100,000 population)",
        ylab = "Burglary Rate (per 100,000 population)",
        main = 'Burglary Rate vs. Murder Rate in U.S. by State Population (2005)')
text(df2$murder, df2$burglary, df2$state, cex = 0.5)
```

Burglary Rate vs. Murder Rate in U.S. by State Population (2005)
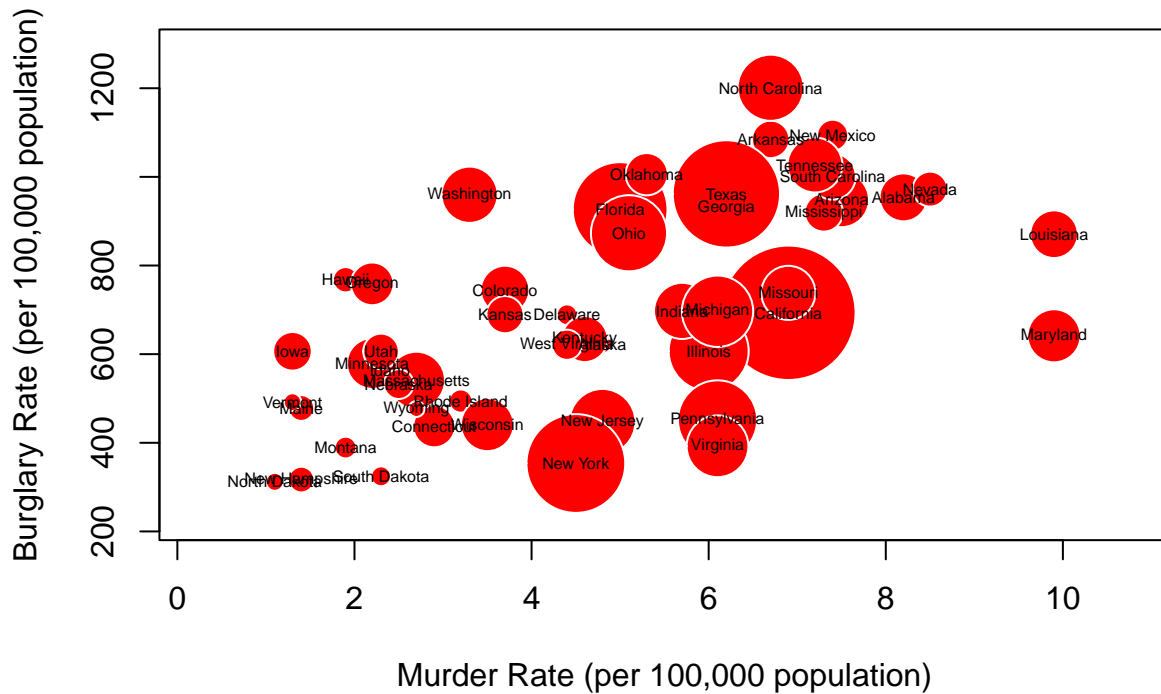
Density Plot using birth rate data from 2008.

```
df3 <- read.csv("birth-rate.csv", header = TRUE, sep = ",")
```

Eliminate the missing values.

```
birth2008 <- df3$X2008[!is.na(df3$X2008)]
```

Create the density data for the birth rates from 2008.

```
d2008 <- density(birth2008)
```

Store the x and y density data in a dataframe. Write to a .txt file.

```
d2008frame <- data.frame(d2008$x, d2008$y)
write.table(d2008frame, 'birthdensity.txt', sep = ',' , row.names = FALSE)
```

Generate the Density Plot.

```
plot(d2008, type = 'n', xlab = "Birth Rate", ylab = "Density",
     main = "Density Plot of Birth Rates (2008)")
polygon(d2008, col = '#821122', border = "#cccccc")
```

Density Plot of Birth Rates (2008)