358_RDS_Project

Jodeci Wheaden

2023-01-18

Libraries

```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Set Working Directory

```
# set working directory
setwd('/Volumes/GoogleDrive/My Drive/Reproducible Data Science/Reproducible_Data_Science_Final_Project/c
```

Global Variables

```
# Initializing city vector
city = c()
# Initializing No. of Years Vector
no.ofYears = c()
# Initializing Total Vector
Total = c()
# Initializing Yearly Avg vector
YearlyAvg = c()
```

Atlanta, GA

```
atlanta_data = read.csv(file = './atlanta_edited.csv')
```

```
# Found on Read.me in the List of cities FOIA'd and what happened with each table
number_years_GA = 2020-2015
total_settlement_GA = sum(atlanta_data$amount_awarded)
yearly_average_GA = total_settlement_GA/number_years_GA

# Append elements to vector
city = append(city, "Atlanta, GA")
no.ofYears = append(no.ofYears, number_years_GA)
Total = append(Total, total_settlement_GA)
YearlyAvg = append(YearlyAvg, yearly_average_GA)
```

Baltimore, MD

```
baltimore_data = read.csv(file = './baltimore_edited.csv')
```

```
number_years_MD = 2020-2015
total_settlement_MD = sum(baltimore_data$amount_awarded)
yearly_average_MD = total_settlement_MD/number_years_MD

# Append elements to vector
city = append(city, "Baltimore, MD")
no.ofYears = append(no.ofYears, number_years_MD)
Total = append(Total, total_settlement_MD)
YearlyAvg = append(YearlyAvg, yearly_average_MD)
```

Baton Rouge, LA

```
baton_rouge_data = read.csv(file = './baton_rouge_edited.csv')
```

```
number_years_LA = 2019-2010 + 1
total_settlement_LA = sum(baton_rouge_data$amount_awarded)
yearly_average_LA = total_settlement_LA/number_years_LA

# Append elements to vector
city = append(city, "Baton Rouge, LA")
no.ofYears = append(no.ofYears, number_years_LA)
Total = append(Total, total_settlement_LA)
YearlyAvg = append(YearlyAvg, yearly_average_LA)
```

Boston, MA

```
boston_data = read.csv(file = './boston_edited.csv')
```

```
number_years_MA = 2019-2010 + 1
total_settlement_MA = sum(boston_data$amount_awarded)
yearly_average_MA = total_settlement_MA/number_years_MA

# Append elements to vector
city = append(city, "Boston, MA")
no.ofYears = append(no.ofYears, number_years_MA)
Total = append(Total, total_settlement_MA)
YearlyAvg = append(YearlyAvg, yearly_average_MA)
```

Cambridge, MA

```
cambridge_data = read.csv(file = './cambridge_edited.csv')
```

```
number_years_MA2 = 2019-2010 + 1
total_settlement_MA2 = sum(cambridge_data$amount_awarded)
yearly_average_MA2 = total_settlement_MA2/number_years_MA2

# Append elements to vector
city = append(city, "Cambridge, MA")
no.ofYears = append(no.ofYears, number_years_MA2)
Total = append(Total, total_settlement_MA2)
YearlyAvg = append(YearlyAvg, yearly_average_MA2)
```

Charleston, SC

```
charleston_data = read.csv(file = './charleston_edited.csv')
```

```
number_years_SC = 2019-2010 + 1
total_settlement_SC = sum(charleston_data$amount_awarded)
yearly_average_SC = total_settlement_SC/number_years_SC

# Append elements to vector
city = append(city, "Charleston, SC")
no.ofYears = append(no.ofYears, number_years_SC)
Total = append(Total, total_settlement_SC)
YearlyAvg = append(YearlyAvg, yearly_average_SC)
```

Chicago, IL

```
chicago_data = read.csv(file = './chicago_edited.csv')
```

```
number_years_IL = 2019-2010 + 1
total_settlement_IL = sum(chicago_data$amount_awarded)
yearly_average_IL = total_settlement_IL/number_years_IL
```

```
# Append elements to vector
city = append(city, "Chicago, IL")
no.ofYears = append(no.ofYears, number_years_IL)
Total = append(Total, total_settlement_IL)
YearlyAvg = append(YearlyAvg, yearly_average_IL)
```

Cincinnati, OH

```
cincinnati_data = read.csv(file = './cincinnati_edited.csv')
```

```
number_years_OH = 2020-2010 + 1
total_settlement_OH = sum(cincinnati_data$amount_awarded)
yearly_average_OH = total_settlement_OH/number_years_OH

# Append elements to vector
city = append(city, "Cincinnati, OH")
no.ofYears = append(no.ofYears, number_years_OH)
Total = append(Total, total_settlement_OH)
YearlyAvg = append(YearlyAvg, yearly_average_OH)
```

Cleveland, OH

```
cleveland_data = read.csv(file = './cleveland_edited.csv')
```

```
number_years_OH2 = 2020-2010 + 1
total_settlement_OH2 = sum(cleveland_data$amount_awarded)
yearly_average_OH2 = total_settlement_OH2/number_years_OH2

# Append elements to vector
city = append(city, "Cleveland, OH")
no.ofYears = append(no.ofYears, number_years_OH2)
Total = append(Total, total_settlement_OH2)
YearlyAvg = append(YearlyAvg, yearly_average_OH2)
```

Columbia, SC

```
columbia_data = read.csv(file = './columbia_edited.csv')
```

```
number_years_SC2 = 2019-2010 + 1
total_settlement_SC2 = sum(columbia_data$amount_awarded)
yearly_average_SC2 = total_settlement_SC2/number_years_SC2

# Append elements to vector
city = append(city, "Columbia, SC")
no.ofYears = append(no.ofYears, number_years_SC2)
Total = append(Total, total_settlement_SC2)
YearlyAvg = append(YearlyAvg, yearly_average_SC2)
```

Detroit, MI

```
detroit_data = read.csv(file = './detroit_edited.csv')
```

```
number_years_MI = 2019-2010
total_settlement_MI = sum(detroit_data$amount_awarded)
yearly_average_MI = total_settlement_MI/number_years_MI

# Append elements to vector
city = append(city, "Detroit, MI")
no.ofYears = append(no.ofYears, number_years_MI)
Total = append(Total, total_settlement_MI)
YearlyAvg = append(YearlyAvg, yearly_average_MI)
```

Fort Lauderdale, FL

```
fort_lauderdale_data = read.csv(file = './fort_lauderdale_edited.csv')
```

```
number_years_FL = 2019-2011 + 1
total_settlement_FL = sum(fort_lauderdale_data$amount_awarded)
yearly_average_FL = total_settlement_FL/number_years_FL

# Append elements to vector
city = append(city, "Fort Lauderdale, FL")
no.ofYears = append(no.ofYears, number_years_FL)
Total = append(Total, total_settlement_FL)
YearlyAvg = append(YearlyAvg, yearly_average_FL)
```

Indianapolis, IN

```
indianapolis_data = read.csv(file = './indianapolis_edited.csv')
```

```
number_years_IN = 2019-2010 + 1
total_settlement_IN = sum(indianapolis_data$amount_awarded)
yearly_average_IN = total_settlement_IN/number_years_IN

# Append elements to vector
city = append(city, "Indianapolis, IN")
no.ofYears = append(no.ofYears, number_years_IN)
Total = append(Total, total_settlement_IN)
YearlyAvg = append(YearlyAvg, yearly_average_IN)
```

Little Rock, AR

```
little_rock_data = read.csv(file = './little_rock_edited.csv')
```

```
number_years_AR = 2019-2010 + 1
total_settlement_AR = sum(little_rock_data$amount_awarded)
yearly_average_AR = total_settlement_AR/number_years_AR

# Append elements to vector
city = append(city, "Little Rock, AR")
no.ofYears = append(no.ofYears, number_years_AR)
Total = append(Total, total_settlement_AR)
YearlyAvg = append(YearlyAvg, yearly_average_AR)
```

Los Angeles, CA

```
los_angeles_data = read.csv(file = './los_angeles_edited.csv')
```

```
number_years_CA = 2019-2010 + 1
total_settlement_CA = sum(los_angeles_data$amount_awarded)
yearly_average_CA = total_settlement_CA/number_years_CA

# Append elements to vector
city = append(city, "Los Angeles, CA")
no.ofYears = append(no.ofYears, number_years_CA)
Total = append(Total, total_settlement_CA)
YearlyAvg = append(YearlyAvg, yearly_average_CA)
```

Memphis, TN

```
memphis_data = read.csv(file = './memphis_edited.csv')
```

```
number_years_TN = 2019-2013 + 1
total_settlement_TN = sum(memphis_data$amount_awarded)
yearly_average_TN = total_settlement_TN/number_years_TN

# Append elements to vector
city = append(city, "Memphis, TN")
no.ofYears = append(no.ofYears, number_years_TN)
Total = append(Total, total_settlement_TN)
YearlyAvg = append(YearlyAvg, yearly_average_TN)
```

Miami, FL

```
miami_data = read.csv(file = './miami_edited.csv')
```

```
number_years_FL = 2020-2010 + 1
total_settlement_FL = sum(miami_data$amount_awarded)
yearly_average_FL = total_settlement_FL/number_years_FL

# Append elements to vector
city = append(city, "Miami, FL")
no.ofYears = append(no.ofYears, number_years_FL)
Total = append(Total, total_settlement_FL)
YearlyAvg = append(YearlyAvg, yearly_average_FL)
```

Milwaukee, WI

```
milwaukee_data = read.csv(file = './milwaukee_edited.csv')
```

```
number_years_WI = 2019-2010 + 1
total_settlement_WI = sum(milwaukee_data$amount_awarded)
yearly_average_WI = total_settlement_WI/number_years_WI

# Append elements to vector
city = append(city, "Milwaukee, WI")
no.ofYears = append(no.ofYears, number_years_WI)
Total = append(Total, total_settlement_WI)
YearlyAvg = append(YearlyAvg, yearly_average_WI)
```

New Orleans, LA

```
new_orleans_data = read.csv(file = './new_orleans_edited.csv')
```

```
number_years_LA = 2019-2010 + 1
total_settlement_LA = sum(new_orleans_data$amount_awarded)
yearly_average_LA = total_settlement_LA/number_years_LA

# Append elements to vector
city = append(city, "New Orleans, LA")
no.ofYears = append(no.ofYears, number_years_LA)
Total = append(Total, total_settlement_LA)
YearlyAvg = append(YearlyAvg, yearly_average_LA)
```

New York, NY

```
new_york_data = read.csv(file = './new_york_edited.csv')
```

```
number_years_NY = 2019-2010 + 1
total_settlement_NY = sum(new_york_data$amount_awarded)
yearly_average_NY = total_settlement_NY/number_years_NY
```

```
# Append elements to vector
city = append(city, "New York, NY")
no.ofYears = append(no.ofYears, number_years_NY)
Total = append(Total, total_settlement_NY)
YearlyAvg = append(YearlyAvg, yearly_average_NY)
```

North Charleston, SC

```
north_charleston_data = read.csv(file = './north_charleston_edited.csv')
```

```
number_years_SC3 = 2019-2010 + 1
total_settlement_SC3 = sum(north_charleston_data$amount_awarded)
yearly_average_SC3 = total_settlement_SC3/number_years_SC3

# Append elements to vector
city = append(city, "North Charleston, SC")
no.ofYears = append(no.ofYears, number_years_SC3)
Total = append(Total, total_settlement_SC3)
YearlyAvg = append(YearlyAvg, yearly_average_SC3)
```

Orlando, FL

```
orlando_data = read.csv(file = './orlando_edited.csv')
```

```
number_years_FL2 = 2018-2010 + 1
total_settlement_FL2 = sum(orlando_data$amount_awarded)
yearly_average_FL2 = total_settlement_FL2/number_years_FL2

# Append elements to vector
city = append(city, "Orlando, FL")
no.ofYears = append(no.ofYears, number_years_FL2)
Total = append(Total, total_settlement_FL2)
YearlyAvg = append(YearlyAvg, yearly_average_FL2)
```

Paterson, NJ

```
paterson_data = read.csv(file = './paterson_edited.csv')
```

```
number_years_NJ = 2019-2010 + 1
total_settlement_NJ = sum(paterson_data$amount_awarded)
yearly_average_NJ = total_settlement_NJ/number_years_NJ

# Append elements to vector
city = append(city, "Paterson, NJ")
no.ofYears = append(no.ofYears, number_years_NJ)
Total = append(Total, total_settlement_NJ)
YearlyAvg = append(YearlyAvg, yearly_average_NJ)
```

Philadelphia, PA

```
philly_data = read.csv(file = './philly_edited.csv')
```

```
number_years_PA = 2019-2009 + 1
total_settlement_PA = sum(philly_data$amount_awarded)
yearly_average_PA = total_settlement_PA/number_years_PA

# Append elements to vector
city = append(city, "Philadelphia, PA")
no.ofYears = append(no.ofYears, number_years_PA)
Total = append(Total, total_settlement_PA)
YearlyAvg = append(YearlyAvg, yearly_average_PA)
```

Richmond, VA

```
richmond_data = read.csv(file = './richmond_edited.csv')
```

```
number_years_VA = 2019-2010 + 1
total_settlement_VA = sum(richmond_data$amount_awarded)
yearly_average_VA = total_settlement_VA/number_years_VA

# Append elements to vector
city = append(city, "Richmond, VA")
no.ofYears = append(no.ofYears, number_years_VA)
Total = append(Total, total_settlement_VA)
YearlyAvg = append(YearlyAvg, yearly_average_VA)
```

Roanoke, VA

```
roanoke_data = read.csv(file = './roanoke_edited.csv')
```

```
number_years_VA2 = 2019-2010 + 1
total_settlement_VA2 = sum(roanoke_data$amount_awarded)
yearly_average_VA2 = total_settlement_VA2/number_years_VA2

# Append elements to vector
city = append(city, "Roanoke, VA")
no.ofYears = append(no.ofYears, number_years_VA2)
Total = append(Total, total_settlement_VA2)
YearlyAvg = append(YearlyAvg, yearly_average_VA2)
```

San Francisco, CA

```
san_francisco_data = read.csv(file = './san_francisco_edited.csv')
```

```
number_years_CA2 = 2019-2010 + 1
total_settlement_CA2 = sum(san_francisco_data$amount_awarded)
yearly_average_CA2 = total_settlement_CA2/number_years_CA2

# Append elements to vector
city = append(city, "San Francisco, CA")
no.ofYears = append(no.ofYears, number_years_CA2)
Total = append(Total, total_settlement_CA2)
YearlyAvg = append(YearlyAvg, yearly_average_CA2)
```

Springfield, MA

```
springfield_data = read.csv(file = './springfield_edited.csv')
```

```
number_years_MA3 = 2020-2006 + 1
total_settlement_MA3 = sum(springfield_data$amount_awarded)
yearly_average_MA3 = total_settlement_MA3/number_years_MA3

# Append elements to vector
city = append(city, "Springfield, MA")
no.ofYears = append(no.ofYears, number_years_MA3)
Total = append(Total, total_settlement_MA3)
YearlyAvg = append(YearlyAvg, yearly_average_MA3)
```

St Louis, MO

```
stlouis_data = read.csv(file = './stlouis_edited.csv')
```

```
number_years_MO = 2019-2015 + 1
total_settlement_MO = sum(stlouis_data$amount_awarded)
yearly_average_MO = total_settlement_MO/number_years_MO

# Append elements to vector
city = append(city, "St Louis, MO")
no.ofYears = append(no.ofYears, number_years_MO)
Total = append(Total, total_settlement_MO)
YearlyAvg = append(YearlyAvg, yearly_average_MO)
```

Washington, DC

```
DC_data = read.csv(file = './DC_edited.csv')
```

```
number_years_DC = 2019-2010
total_settlement_DC = sum(DC_data$amount_awarded)
yearly_average_DC = total_settlement_DC/number_years_DC

# Append elements to vector
city = append(city, "Washington, DC")
no.ofYears = append(no.ofYears, number_years_DC)
Total = append(Total, total_settlement_DC)
YearlyAvg = append(YearlyAvg, yearly_average_DC)
```

Waterbury, CT

```
waterbury_data = read.csv(file = './waterbury_edited.csv')
```

```
number_years_CT = 2019-2011 + 1
total_settlement_CT = sum(waterbury_data$amount_awarded)
yearly_average_CT = total_settlement_CT/number_years_CT

# Append elements to vector
city = append(city, "Waterbury, CT")
no.ofYears = append(no.ofYears, number_years_CT)
Total = append(Total, total_settlement_CT)
YearlyAvg = append(YearlyAvg, yearly_average_CT)
```

Figure 1

```
df <- data.frame(city,no.ofYears,Total,YearlyAvg)
# Used formatC function to add commas to numbers
df <- df %>% mutate(Total = formatC(round(Total), format = "f", big.mark = ",", drop0trailing = TRUE), 
df
```

```
##                    city no.ofYears         Total   YearlyAvg
## 1           Atlanta, GA          5     4,761,182     952,236
## 2         Baltimore, MD          5    18,432,748   3,686,550
## 3       Baton Rouge, LA         10     2,879,795     287,979
## 4            Boston, MA         10    11,905,482   1,190,548
## 5         Cambridge, MA         10       114,000      11,400
## 6        Charleston, SC         10     1,520,250     152,025
## 7           Chicago, IL         10   467,586,464  46,758,646
## 8        Cincinnati, OH         11     2,472,787     224,799
## 9         Cleveland, OH         11    28,573,475   2,597,589
## 10         Columbia, SC         10     1,352,435     135,244
## 11           Detroit, MI          9    57,702,989   6,411,443
## 12  Fort Lauderdale, FL          9     2,471,384     274,598
## 13     Indianapolis, IN         10    13,149,775   1,314,977
## 14      Little Rock, AR         10       943,950      94,395
## 15      Los Angeles, CA         10   329,925,620  32,992,562
## 16          Memphis, TN          7     8,772,884   1,253,269
## 17            Miami, FL         11     7,284,684     662,244
```

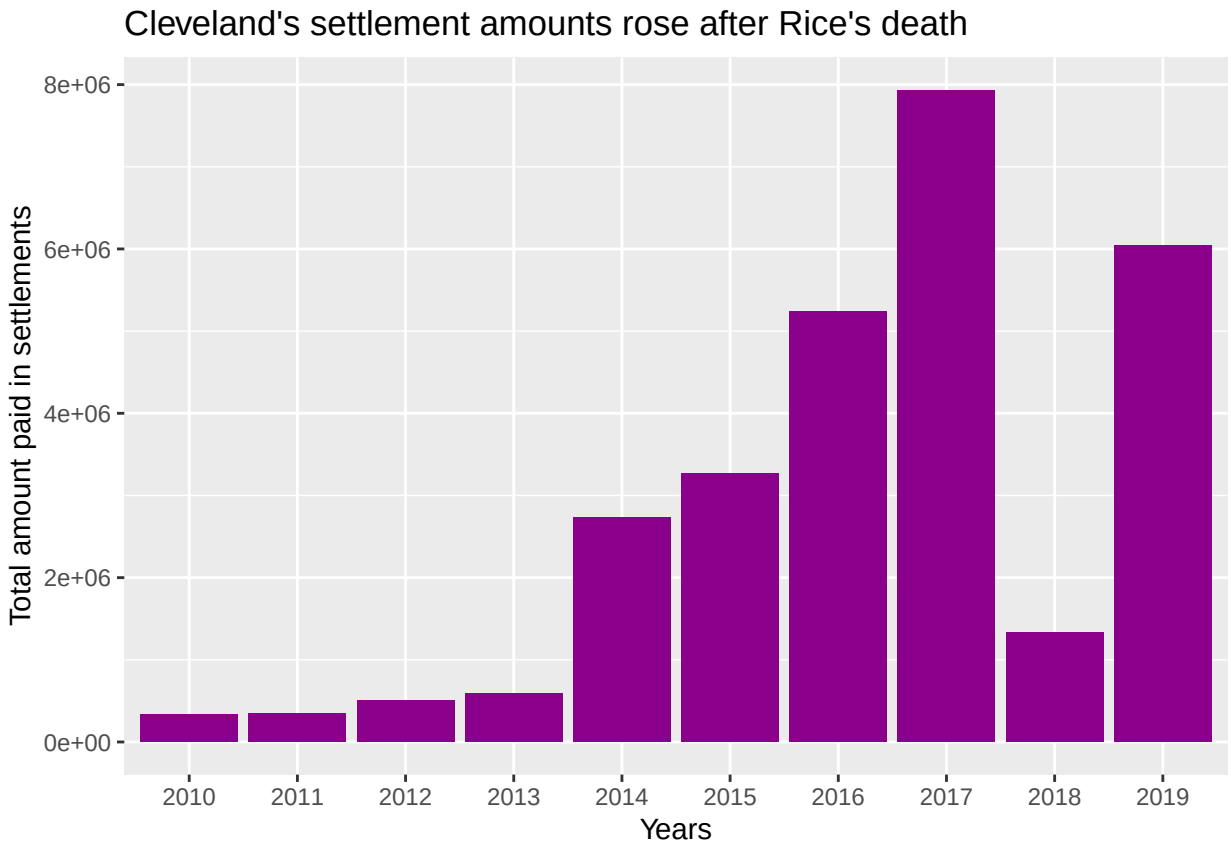
```
## 18        Milwaukee, WI         10    40,017,822   4,001,782
## 19      New Orleans, LA         10     3,510,642     351,064
## 20         New York, NY         10 1,704,120,487 170,412,049
## 21 North Charleston, SC         10     3,333,750     333,375
## 22          Orlando, FL          9     3,611,879     401,320
## 23         Paterson, NJ         10     7,742,498     774,250
## 24     Philadelphia, PA         11   116,881,088  10,625,553
## 25         Richmond, VA         10       748,500      74,850
## 26          Roanoke, VA         10       132,500      13,250
## 27    San Francisco, CA         10    27,873,298   2,787,330
## 28      Springfield, MA         15    32,846,089   2,189,739
## 29         St Louis, MO          5     3,117,847     623,569
## 30       Washington, DC          9   114,841,449  12,760,161
## 31        Waterbury, CT          9     2,227,250     247,472
```

Figure 2

```
# Dataframe for settlement amounts per year
df1 <- cleveland_data %>%
  group_by(calendar_year) %>%
  summarise(sum_settlement = sum(amount_awarded)) %>%
  filter(calendar_year != 2020)
df1
```

```
## # A tibble: 10 x 2
##    calendar_year sum_settlement
##            <int>          <dbl>
##  1          2010        329500
##  2          2011        350454.
##  3          2012        503830.
##  4          2013        589405.
##  5          2014       2734400
##  6          2015       3269500
##  7          2016       5240500
##  8          2017       7934438.
##  9          2018       1330400
## 10          2019       6046547.
```

```
# Change calendar_year variable from discrete variable to continuous variable by using factor function
ggplot(df1, aes(x=factor(calendar_year), y = sum_settlement)) + geom_bar(stat = "identity", fill="darkma
  xlab("Years") +
  ylab("Total amount paid in settlements") +
  ggtitle("Cleveland's settlement amounts rose after Rice's death")
```


Cleveland's settlement amounts rose after Rice's death
Total amount paid in settlements
8e+06
6e+06
4e+06
2e+06
0e+00
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Years