STAT 531
Midterm Redesign Project

Group #1

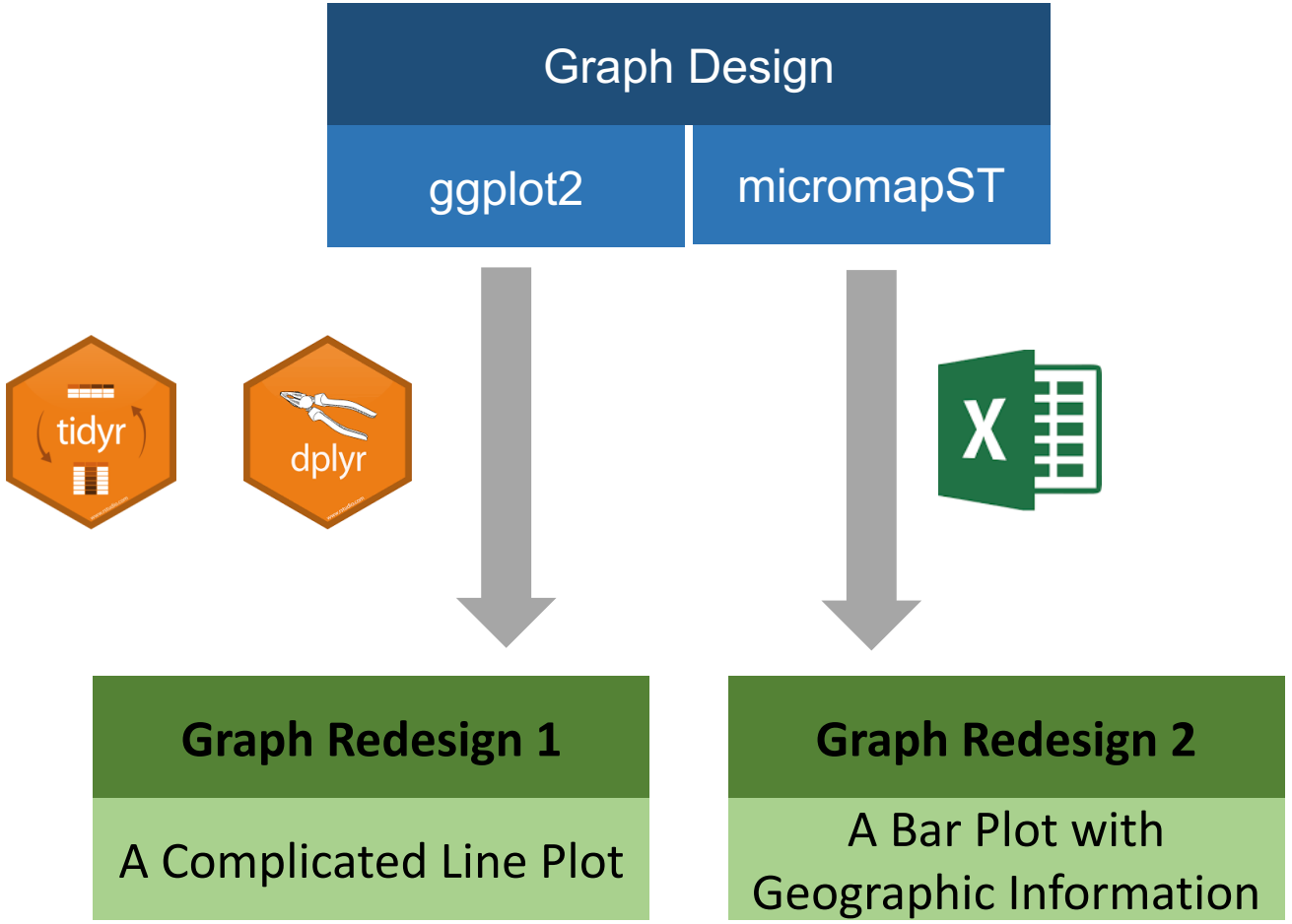
Long Zhang
Mohammad Ridwan
Baizhong Hou
03.05.2020

Content

1. Introduction
2. Redesign 1 (Redesign for a Complicated Line Plot)
3. Redesign 2 (Redesign for a Bar Plot with Geographic Information)
4. Challenges
5. Conclusion

1. Introduction

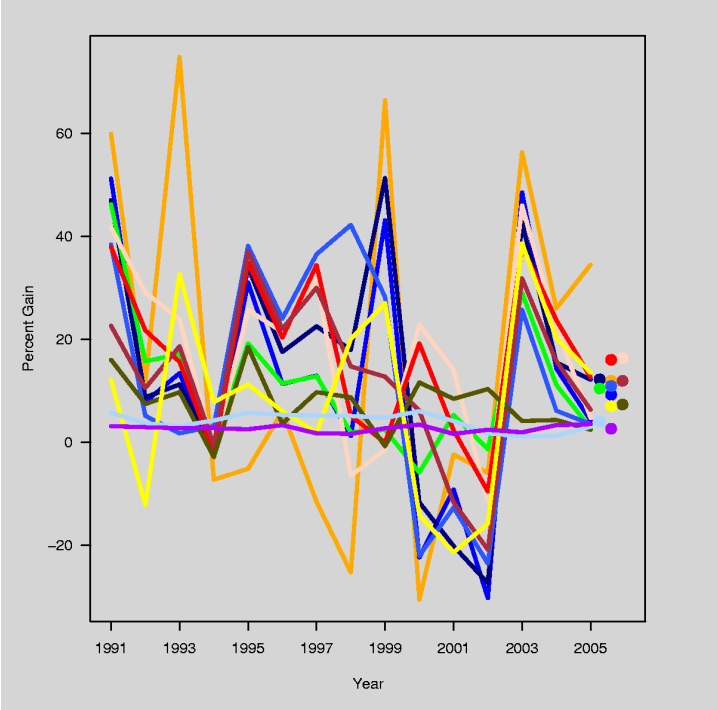
1. Introduction



2. Redesign 1

- **Redesign for a Complicated Line Plot**

2. Redesign 1: Introduction



Annual Percentage Gains for 12 funds in the Vanguard Group from 1991 to 2005 (Price, 2006)

2. Redesign 1: Redesign Goals

Goals:

Enable Accurate Comparisons

- Can easily **observe** the trends of percent gains for each fund
- Can easily **compare** the percent gains information for any two funds

Interpretation

- Provide necessary **description**

Simplify Appearance

- Displays lines **concisely**

Engage the Reader

- **Beautiful** display

2. Redesign 1: Redesign Process

Data Preparation

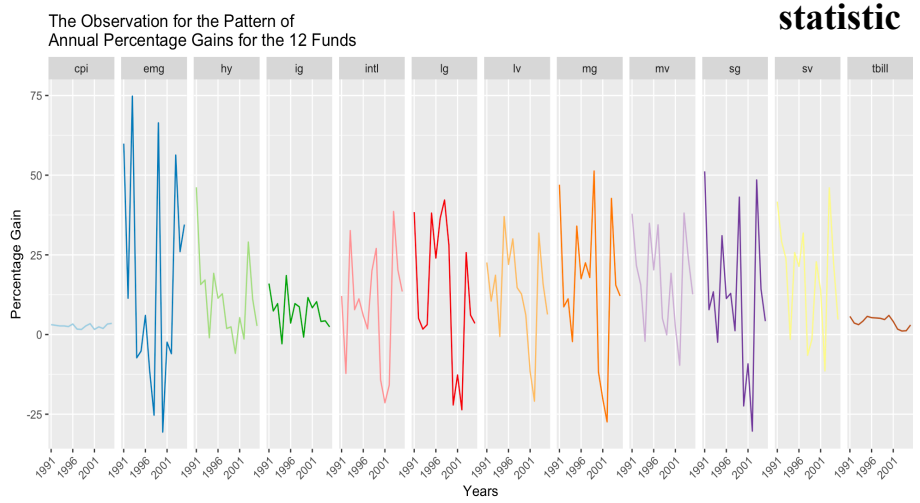
Transform the Schema

```
> str(gain_annual)
'data.frame':   15 obs. of  13 variables:
 $ class: Factor w/ 16 levels "1991","1992",..: 1 2 3 4 5 6 7 8 9 10 ...
 $ emg  : num  59.9 11.4 74.8 -7.3 -5.2 6 -11.6 -25.3 66.4 -30.6 ...
 $ sg   : num  51.2 7.8 13.4 -2.4 31 11.3 12.9 1.2 43.1 -22.4 ...
 $ mg   : num  47 8.7 11.2 -2.2 34 17.5 22.5 17.9 51.3 -11.7 ...
 $ hy   : num  46.2 15.7 17.1 -1 19.2 11.4 12.8 1.9 2.4 -5.9 ...
 $ sv   : num  41.7 29.1 23.8 -1.5 25.7 21.4 31.8 -6.5 -1.5 22.8 ...
 $ lg   : num  38.4 5.1 1.7 3.1 38.1 24 36.5 42.2 28.2 -22.1 ...
 $ mv   : num  37.9 21.7 15.6 -2.1 34.9 20.3 34.4 5.1 -0.1 19.2 ...
 $ lv   : num  22.6 10.5 18.6 -0.6 37 22 30 14.7 12.7 6.1 ...
 $ ig   : num  16 7.4 9.7 -2.9 18.5 3.6 9.7 8.7 -0.8 11.6 ...
 $ intl : num  12.1 -12.2 32.6 7.8 11.2 6 1.8 20 27 -14.2 ...
 $ tbill: num  5.7 3.6 3.1 4.2 5.7 5.3 5.2 5.1 4.7 6 ...
 $ cpi  : num  3.1 2.9 2.7 2.7 2.5 3.3 1.7 1.6 2.7 3.4 ...
> str(gain_annual2)
'data.frame':   180 obs. of  4 variables:
 $ class: Factor w/ 16 levels "1991","1992",..: 1 2 3 4 5 6 7 8 9 10 ...
 $ Funds: chr  "emg" "emg" "emg" "emg" ...
 $ gain : num  59.9 11.4 74.8 -7.3 -5.2 6 -11.6 -25.3 66.4 -30.6 ...
 $ group: chr  "Group1" "Group1" "Group1" "Group1" ...
```

```
> str(gain_annual2)
'data.frame':   180 obs.
 $ class: Factor w/ 16 l
 $ Funds: chr  "emg" "em
 $ gain : num  59.9 11.4
 $ group: chr  "Group1"
```

Group Separation

Pattern Observation



standard deviation statistic

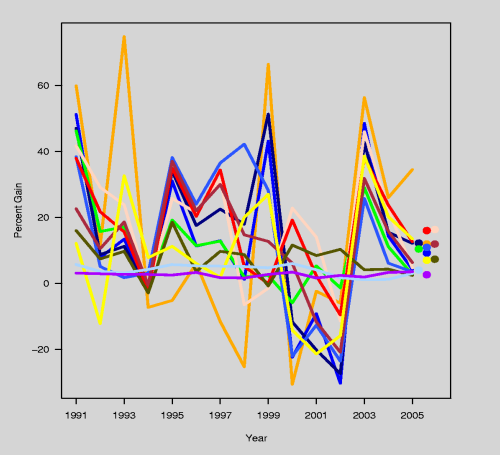
Assign Groups

```
> (group1 <- std$funds[1:3])
[1] "emg" "sg"  "mg"
> (group2 <- std$funds[4:6])
[1] "lg"   "intl" "sv"
> (group3 <- std$funds[7:9])
[1] "lv" "mv" "hy"
> (group4 <- std$funds[10:12])
[1] "ig"    "tbill" "cpi"
```

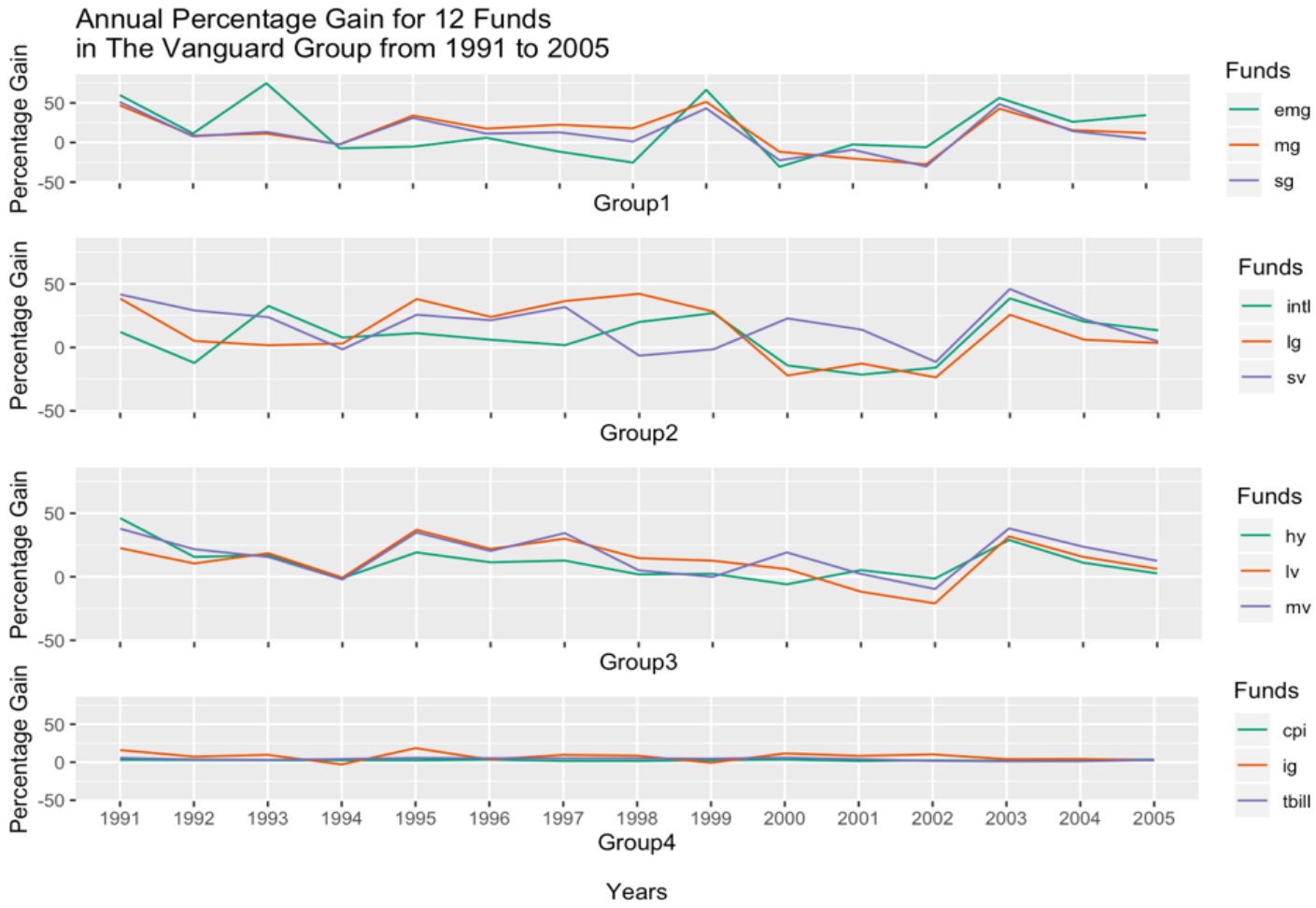
Filter the Data

Design the Plot

2. Redesign 1: Results and Analysis



Original

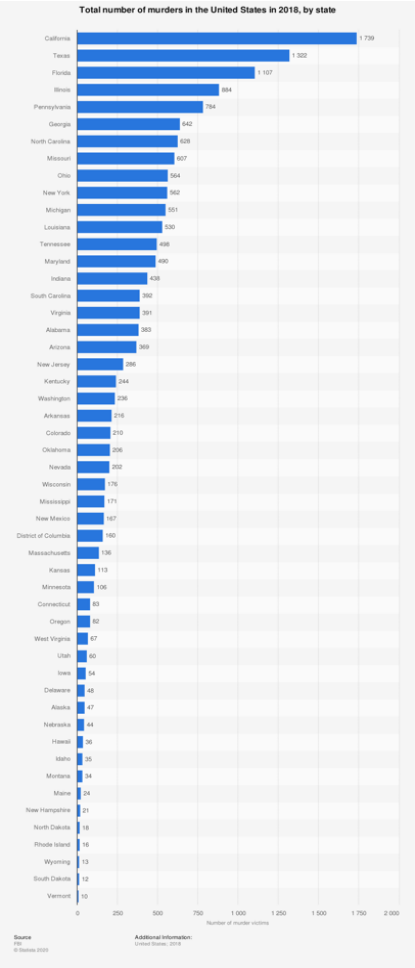


Redesigned Graph

3. Redesign 2

- **Redesign for a Bar Plot with Geographic Information**

3. Redesign 2: Introduction



Total number of murders in the United States in 2018 (Statista, 2019)

3. Redesign 2: Redesign Goals

Goals:

Enable Accurate Comparisons

- Ease the readers to focus on some areas

Complete the Information

- Ease the readers to see the increment or decrement of the level

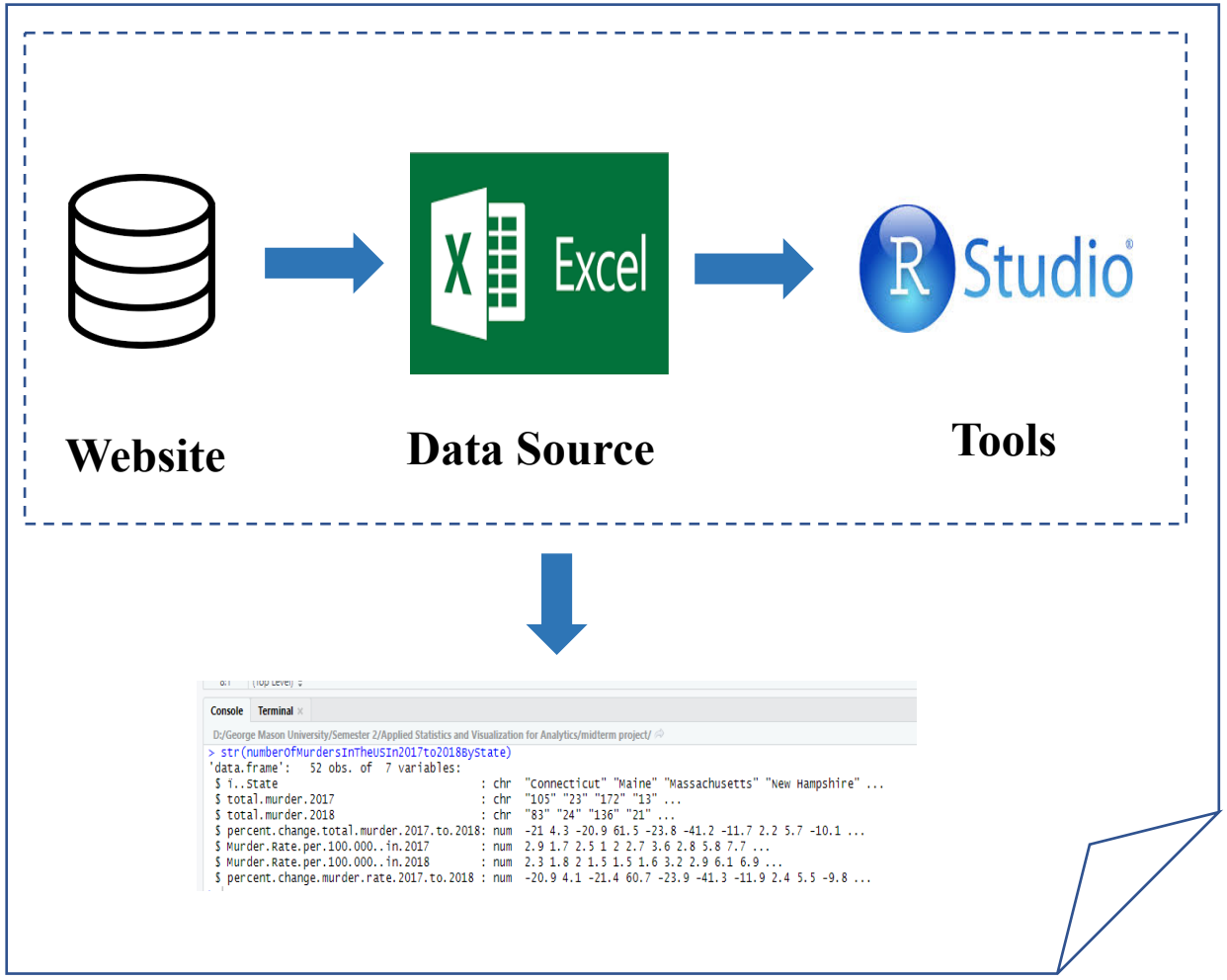
Interpretation

- Straighten the conclusion

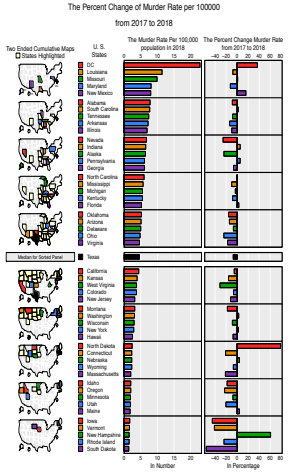
Attract the Reader

- Maps and Colors

3. Redesign 2: Redesign Process

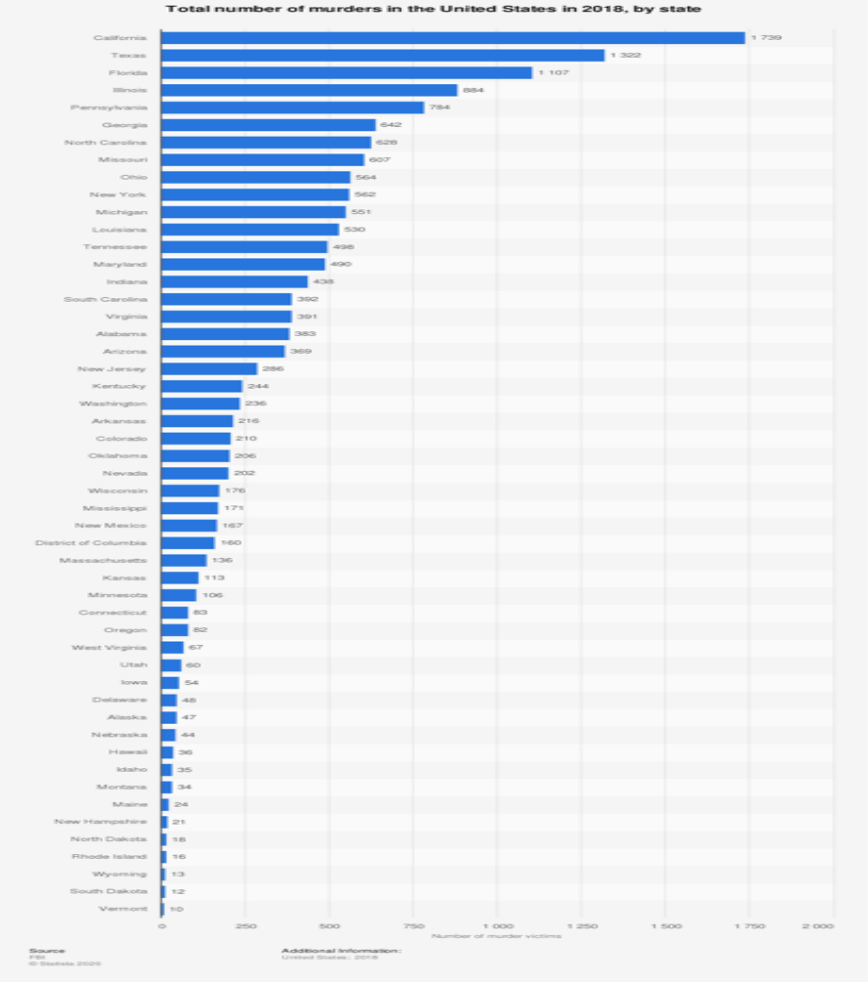


Data Preparation

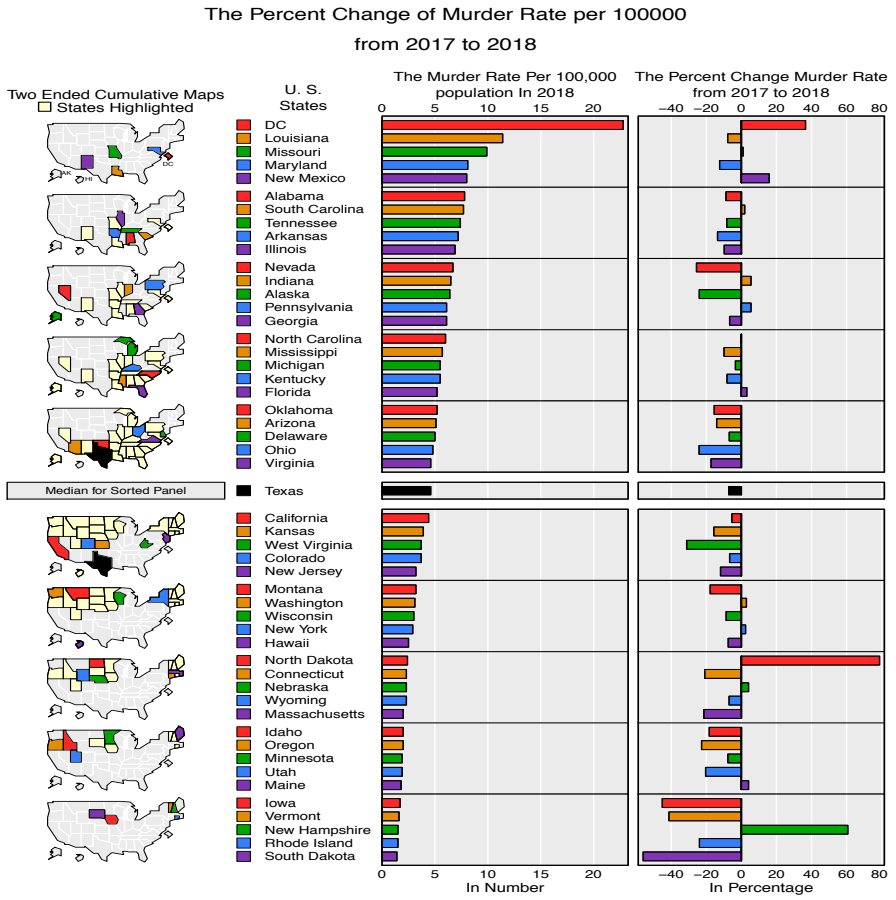


Result

3. Redesign 2: Results and Analysis



Original

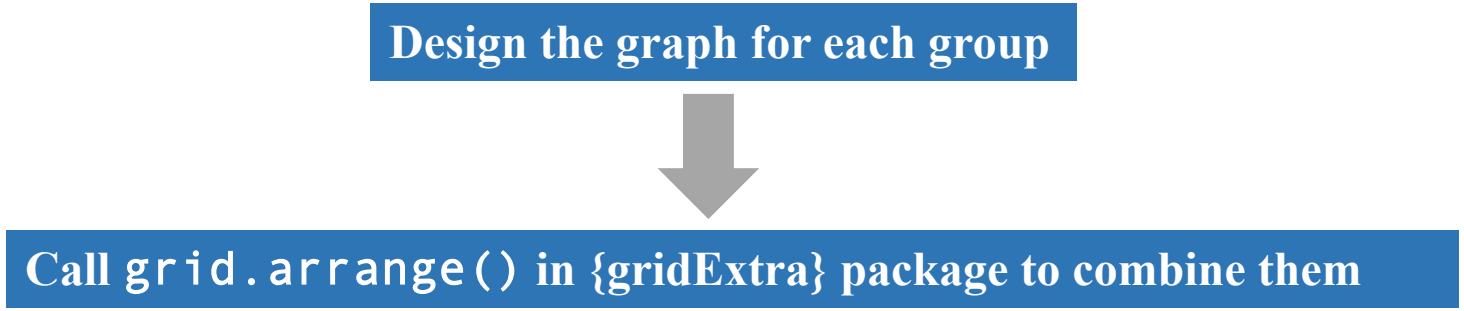


Redesigned Graph

4. Challenges

4. Challenges

- The selection of cases
- Hard to assign individual legend for each subgraph when we use facet_grid() in {ggplot2} package.



5. Conclusion

5. Conclusion

- Redesign two different types of bad graphs
- Plot graphs via {ggplot2} and {micromapST} packages
- Learn the principles for how to make nice graphs
- Understand the power of {tidyr} and {dplyr} packages for data preparation

THANK YOU

Price, P. (2006, May 23). *A bad graph but not clear how to make it better*. Retrieved from Statistical Modeling, Causal Inference, and Social Science: https://statmodeling.stat.columbia.edu/2006/05/23/post_8/

Statista. (2019, September). *Total number of murders in the United States in 2018, by state*. Retrieved from Statista: https://www.statista.com/statistics/195331/number-of-murders-in-the-us-by-state/