# Part 1

**Question 1:** Descriptive Statistics

The following table shows the main statistics of the statedata.csv dataset.

Table 1: Statistics of the statedata.csv dataset

| Statistic | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|
| Population | 4,246.420 | 4,464.491 | 365 | 21,198 |
| Income | 4,435.800 | 614.470 | 3,098 | 6,315 |
| Illiteracy | 1.170 | 0.610 | 0.500 | 2.800 |
| Life.Exp | 70.879 | 1.342 | 67.960 | 73.600 |
| Murder | 7.378 | 3.692 | 1.400 | 15.100 |
| HS.Grad | 53.108 | 8.077 | 37.800 | 67.300 |
| Frost | 104.460 | 51.981 | 0 | 188 |
| Area | 70,735.880 | 85,327.300 | 1,049 | 566,432 |
| state.area | 72,367.980 | 88,278.010 | 1,214 | 589,757 |

**Question 2:** Life Expectancy Analysis

The life expectancy ranges from 67.96 to 73.60 across all the states. The following table shows the first five states with the highest life expectancy:

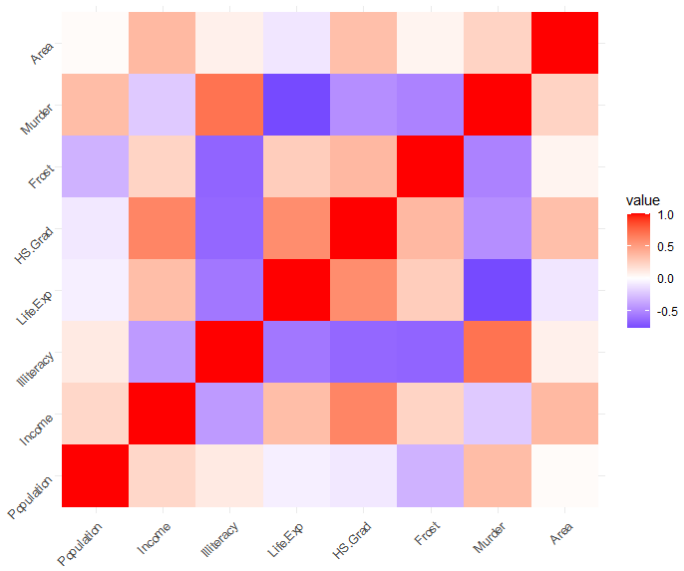
Table 2: The highest-five life expectancy

| | state.abb <chr> | Life.Exp <dbl> |
|---|---|---|
| 1 | HI | 73.60 |
| 2 | MN | 72.96 |
| 3 | UT | 72.90 |
| 4 | ND | 72.78 |
| 5 | NE | 72.60 |

Among all the states, Hawaii has the highest life expectancy, and South Atlantic has the lowest.

**Question 3:** Correlation Heatmap

The following heatmap shows the correlation map of key variables in the dataset.




We can observe income and HS. Grad has the strongest positive correlation with Life.Exp and Murder have the strongest negative correlation. These results seem reasonable.

# Part 2

**Question 1:** Winners

The following table shows the total votes for the party with the highest votes each year.

Table 3: Highest-voted party in each year

| | year | candidatevotes | party_detailed |
|---|---|---|---|
| | <int> | <int> | <chr> |
| 1 | 1976 | 3882244 | REPUBLICAN |
| 2 | 1980 | 4522994 | REPUBLICAN |
| 3 | 1984 | 5467009 | REPUBLICAN |
| 4 | 1988 | 5054917 | REPUBLICAN |
| 5 | 1992 | 5121325 | DEMOCRAT |
| 6 | 1996 | 5119835 | DEMOCRAT |
| 7 | 2000 | 5861203 | DEMOCRAT |
| 8 | 2004 | 6745485 | DEMOCRAT |
| 9 | 2008 | 8274473 | DEMOCRAT |
| 10 | 2012 | 7854285 | DEMOCRAT |
| 11 | 2016 | 8753788 | DEMOCRAT |
| 12 | 2020 | 11110250 | DEMOCRAT |

**Question 2:** Election Map

The following graph shows the winning party in each state in 2020.

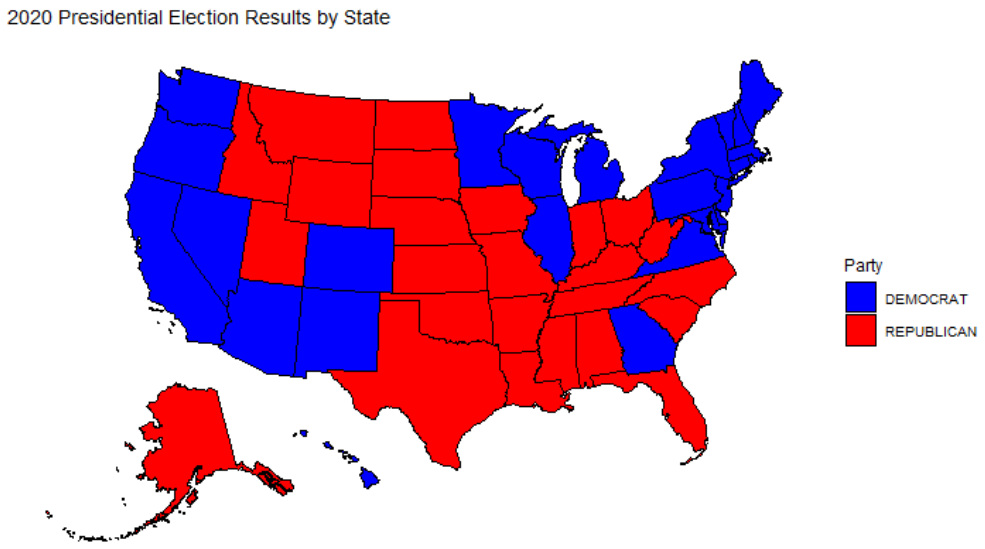



Figure 1: 2020 presidential election results by state

**Question 3:** Facet plot maps

The following facet graph shows the winning party in each state from 2004 to 2016.

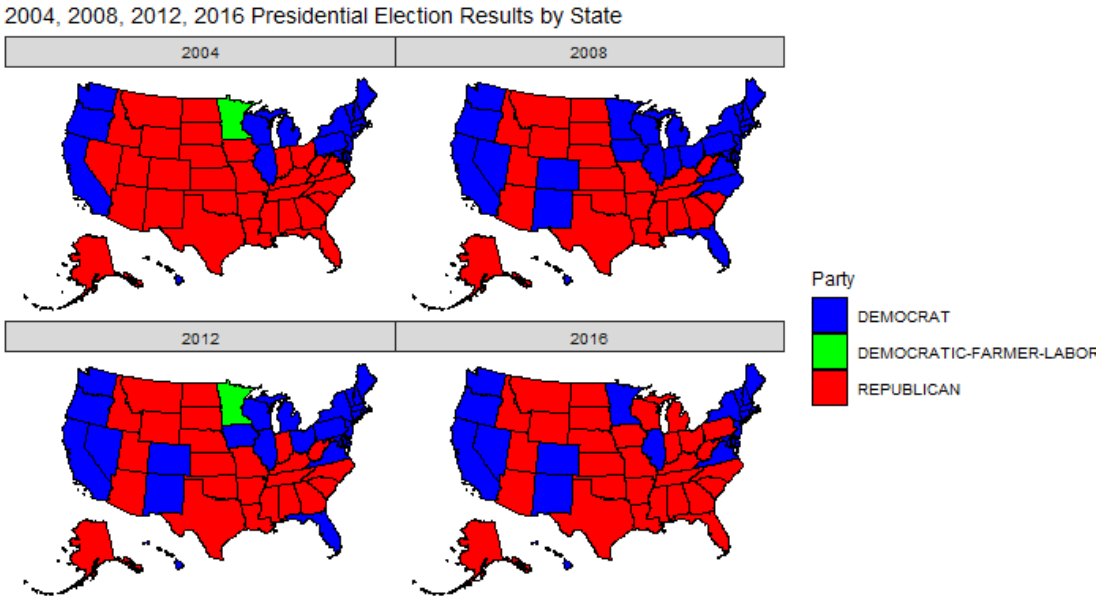



Figure 2: Presidential election results by state

We can observe that states New Mexico and Colorado swayed from Republican to Democrat in 2008 and remained Democrat in the subsequent years. Meanwhile, more and more states in the Rust area swayed from Democrat to Republican.

**Bonus Question 1:** see the file "assignment_2_beamer.pdf".

**Bonus Question 2:** see website: https://yezheng.shinyapps.io/assignment_2/

# Part 3

The given code snippet has the following bugs:

1. Line 5: “choices datasets” should be “choice=dataset”.
2. Line 13: “summmry” is a typo, it should be “summary”.
3. Line 17: “plot(dataset)” should be “plot(dataset())”.