Story - 1 : Infrastructure Investment and Jobs Act Funding Allocation

Jiaxin Zheng

2025-02-15

President Joe Biden signed the Infrastructure Investment and Jobs Act Act(IIJA) on November 12, 2021. This Act invests $1 trillion to fix our nation's infrastructure like roads, bridges, and public transit, and also helps to create over 2 million jobs over the next decade. However, several important questions remain:

-Is the allocation equitable based on the population of each of the States and Territories, or is bias apparent?
-Does the allocation favor the political interests of the Biden administration?

First: We need to prepare for the data:

1.IIJA Funding as of March 2023 2.State Population in 2023 3.Election 2022

```
# load the library
library(tidyverse)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.2
```

```
## Warning: package 'dplyr' was built under R version 4.4.2
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
# Read IIJA FUNDING AS OF MARCH 2023 CSV file
funding <- read.csv("https://raw.githubusercontent.com/Jennyjjxxzz/Data-608_Story_1/refs/heads/main/IIJ
head(funding)
```

```
##   State..Teritory.or.Tribal.Nation Total..Billions.
## 1                          ALABAMA           3.0000
```

```
## 2                         ALASKA          3.7000
## 3                 AMERICAN SAMOA          0.0686
## 4                        ARIZONA          3.5000
## 5                       ARKANSAS          2.8000
## 6                     CALIFORNIA         18.4000
```

```
# Rename the columns
colnames(funding) <- c("State", "Total_Billions")
head(funding)
```

```
##            State Total_Billions
## 1        ALABAMA         3.0000
## 2         ALASKA         3.7000
## 3 AMERICAN SAMOA         0.0686
## 4        ARIZONA         3.5000
## 5       ARKANSAS         2.8000
## 6     CALIFORNIA        18.4000
```

```
# Read Population by State 2023 CSV file
population <- read.csv("https://raw.githubusercontent.com/Jennyjjxxzz/Data-608_Story_1/refs/heads/main/
head(population)
```

```
##                                                table.with.row.headers.in.column.A.and.column.headers.
## 1 Annual Estimates of the Resident Population for the United States, Regions, States, District of Col
## 2
## 3
## 4
## 5
## 6
##                                X                                  X.1         X.2
## 1
## 2 April 1, 2020 Estimates Base Population Estimate (as of July 1)
## 3                                                              2020        2021
## 4                   331,464,948                         331,526,933 332,048,977
## 5                    57,614,141                          57,430,477  57,243,423
## 6                    68,987,296                          68,969,794  68,850,246
##           X.3         X.4
## 1
## 2
## 3        2022        2023
## 4 333,271,411 334,914,895
## 5  57,026,847  56,983,517
## 6  68,783,028  68,909,283
```

```
colnames(population) <- c("State", "April_2020_Estimates_Base", "2020", "2021", "2022", "2023")

# Keep only "State" and "2023" population columns
population_cleaned <- population %>%
  select(State, `2023`) %>%
  drop_na(State)

print(population_cleaned)
```

```
##
## 1
## 2
## 3
## 4
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17
## 18
## 19
## 20
## 21
## 22
## 23
## 24
## 25
## 26
## 27
## 28
## 29
## 30
## 31
## 32
## 33
## 34
## 35
## 36
## 37
## 38
## 39
## 40
## 41
## 42
## 43
## 44
## 45
## 46
## 47
## 48
## 49
## 50
## 51
## 52
## 53
```

```
## 54
## 55
## 56
## 57
## 58
## 59
## 60
## 61
## 62
## 63 Note: The estimates are developed from a base that integrates the 2020 Census, Vintage 2020 estima
## 64
## 65
## 66
## 67
##           2023
## 1
## 2
## 3         2023
## 4  334,914,895
## 5   56,983,517
## 6   68,909,283
## 7  130,125,290
## 8   78,896,805
## 9    5,108,468
## 10     733,406
## 11   7,431,344
## 12   3,067,732
## 13  38,965,193
## 14   5,877,610
## 15   3,617,176
## 16   1,031,890
## 17     678,972
## 18  22,610,726
## 19  11,029,227
## 20   1,435,138
## 21   1,964,726
## 22  12,549,689
## 23   6,862,199
## 24   3,207,004
## 25   2,940,546
## 26   4,526,154
## 27   4,573,749
## 28   1,395,722
## 29   6,180,253
## 30   7,001,399
## 31  10,037,261
## 32   5,737,915
## 33   2,939,690
## 34   6,196,156
## 35   1,132,812
## 36   1,978,379
## 37   3,194,176
## 38   1,402,054
## 39   9,290,841
```

```
## 40   2,114,371
## 41  19,571,216
## 42  10,835,491
## 43     783,926
## 44  11,785,935
## 45   4,053,824
## 46   4,233,358
## 47  12,961,683
## 48   1,095,962
## 49   5,373,555
## 50     919,318
## 51   7,126,489
## 52  30,503,301
## 53   3,417,734
## 54     647,464
## 55   8,715,698
## 56   7,812,880
## 57   1,770,071
## 58   5,910,955
## 59     584,057
## 60
## 61   3,205,691
## 62
## 63
## 64
## 65
## 66
## 67
```

```r
# Remove the not necessary rows
population_cleaned <- population_cleaned [-c(1,2,3,4,5,6,7,8,60,62,63,64,65,66,67),]
population_cleaned$State <- gsub("^\\.", "", population_cleaned$State)
```

```r
print(population_cleaned)
```

```
##                   State       2023
## 9               Alabama  5,108,468
## 10               Alaska    733,406
## 11              Arizona  7,431,344
## 12             Arkansas  3,067,732
## 13           California 38,965,193
## 14             Colorado  5,877,610
## 15          Connecticut  3,617,176
## 16             Delaware  1,031,890
## 17 District of Columbia    678,972
## 18              Florida 22,610,726
## 19              Georgia 11,029,227
## 20               Hawaii  1,435,138
## 21                Idaho  1,964,726
## 22             Illinois 12,549,689
## 23              Indiana  6,862,199
## 24                 Iowa  3,207,004
## 25               Kansas  2,940,546
```

```
## 26             Kentucky  4,526,154
## 27            Louisiana  4,573,749
## 28                Maine  1,395,722
## 29             Maryland  6,180,253
## 30        Massachusetts  7,001,399
## 31             Michigan 10,037,261
## 32            Minnesota  5,737,915
## 33          Mississippi  2,939,690
## 34             Missouri  6,196,156
## 35              Montana  1,132,812
## 36             Nebraska  1,978,379
## 37               Nevada  3,194,176
## 38        New Hampshire  1,402,054
## 39           New Jersey  9,290,841
## 40           New Mexico  2,114,371
## 41             New York 19,571,216
## 42       North Carolina 10,835,491
## 43         North Dakota    783,926
## 44                 Ohio 11,785,935
## 45             Oklahoma  4,053,824
## 46               Oregon  4,233,358
## 47         Pennsylvania 12,961,683
## 48         Rhode Island  1,095,962
## 49       South Carolina  5,373,555
## 50         South Dakota    919,318
## 51            Tennessee  7,126,489
## 52                Texas 30,503,301
## 53                 Utah  3,417,734
## 54              Vermont    647,464
## 55             Virginia  8,715,698
## 56           Washington  7,812,880
## 57        West Virginia  1,770,071
## 58            Wisconsin  5,910,955
## 59              Wyoming    584,057
## 61          Puerto Rico  3,205,691
```

```
# Read Elections2020 CSV file
elections <- read.csv("https://raw.githubusercontent.com/Jennyjjxxzz/Data-608_Story_1/refs/heads/main/E
head(elections)
```

```
##        state state_abr trump_pct biden_pct trump_vote biden_vote trump_win
## 1     Alaska        AK      53.1      43.0     189543     153502         1
## 2     Hawaii        HI      34.3      63.7     196864     366130         0
## 3 Washington        WA      39.0      58.4    1584651    2369612         0
## 4     Oregon        OR      40.7      56.9     958448    1340383         0
## 5 California        CA      34.3      63.5    5982194   11082293         0
## 6      Idaho        ID      63.9      33.1     554128     287031         1
##   biden_win
## 1         0
## 2         1
## 3         1
## 4         1
## 5         1
## 6         0
```

```r
# Select the necessary columns
elections_result <- elections %>% select(state, trump_win, biden_win)
colnames(elections_result) <- c("State", "Trump_Win","Biden_Win")
print(elections_result)
```

```
##                  State Trump_Win Biden_Win
## 1               Alaska         1         0
## 2               Hawaii         0         1
## 3           Washington         0         1
## 4               Oregon         0         1
## 5           California         0         1
## 6                Idaho         1         0
## 7              Montana         1         0
## 8               Nevada         0         1
## 9              Wyoming         1         0
## 10                Utah         1         0
## 11             Arizona         0         1
## 12            Colorado         0         1
## 13          New Mexico         0         1
## 14        North Dakota         1         0
## 15        South Dakota         1         0
## 16            Nebraska         1         0
## 17              Kansas         1         0
## 18            Oklahoma         1         0
## 19               Texas         1         0
## 20           Minnesota         0         1
## 21                Iowa         1         0
## 22            Missouri         1         0
## 23            Arkansas         1         0
## 24           Louisiana         1         0
## 25           Wisconsin         0         1
## 26            Illinois         0         1
## 27         Mississippi         1         0
## 28            Michigan         0         1
## 29             Indiana         1         0
## 30            Kentucky         1         0
## 31           Tennessee         1         0
## 32             Alabama         1         0
## 33                Ohio         1         0
## 34       West Virginia         1         0
## 35             Georgia         0         1
## 36             Florida         1         0
## 37               Maine         0         1
## 38       New Hampshire         0         1
## 39             Vermont         0         1
## 40       Massachusetts         0         1
## 41        Rhode Island         0         1
## 42         Connecticut         0         1
## 43            New York         0         1
## 44        Pennsylvania         0         1
## 45            Virginia         0         1
## 46          New Jersey         0         1
## 47            Delaware         0         1
```

```
## 48             Maryland         0         1
## 49       North Carolina         1         0
## 50       South Carolina         1         0
## 51 District of Columbia         0         1
```

```
# Fix known mismatches, and fix the misspelled
population_cleaned$State[population_cleaned$State == "DISTRICT OF COLUMBIA"] <- "WASHINGTON D.C."
funding$State[funding$State == "DELEWARE"] <- "DELAWARE"
```

```
# Standardize states in all data
funding_cleaned <- funding %>%
  mutate(State = toupper(trimws(State)))

population_cleaned <- population_cleaned %>%
  mutate(State = toupper(trimws(State)))

elections_cleaned <- elections_result %>%
  mutate(State = toupper(trimws(State)))
```

```
# Merge all three data in to one data
merged_data <- funding_cleaned %>%
  left_join(population_cleaned, by = "State") %>%
  left_join(elections_cleaned, by = "State")

head(merged_data)
```

```
##            State Total_Billions       2023 Trump_Win Biden_Win
## 1        ALABAMA         3.0000  5,108,468         1         0
## 2         ALASKA         3.7000    733,406         1         0
## 3 AMERICAN SAMOA         0.0686       <NA>        NA        NA
## 4        ARIZONA         3.5000  7,431,344         0         1
## 5       ARKANSAS         2.8000  3,067,732         1         0
## 6     CALIFORNIA        18.4000 38,965,193         0         1
```

```
# Remove the NA
final_df <- merged_data %>%
  drop_na()

print(final_df)
```

```
##                   State Total_Billions       2023 Trump_Win Biden_Win
## 1               ALABAMA         3.0000  5,108,468         1         0
## 2                ALASKA         3.7000    733,406         1         0
## 3               ARIZONA         3.5000  7,431,344         0         1
## 4              ARKANSAS         2.8000  3,067,732         1         0
## 5            CALIFORNIA        18.4000 38,965,193         0         1
## 6              COLORADO         3.2000  5,877,610         0         1
## 7           CONNECTICUT         2.5000  3,617,176         0         1
## 8              DELAWARE         0.7920  1,031,890         0         1
## 9  DISTRICT OF COLUMBIA         1.1000    678,972         0         1
## 10              FLORIDA         8.2000 22,610,726         1         0
## 11              GEORGIA         5.0000 11,029,227         0         1
```

```
## 12              HAWAII         1.0000  1,435,138         0         1
## 13               IDAHO         1.2000  1,964,726         1         0
## 14            ILLINOIS         8.4000 12,549,689         0         1
## 15             INDIANA         3.4000  6,862,199         1         0
## 16                IOWA         2.4000  3,207,004         1         0
## 17              KANSAS         1.5000  2,940,546         1         0
## 18            KENTUCKY         3.9000  4,526,154         1         0
## 19           LOUISIANA         4.3000  4,573,749         1         0
## 20               MAINE         1.1000  1,395,722         0         1
## 21            MARYLAND         2.7000  6,180,253         0         1
## 22       MASSACHUSETTS         3.6000  7,001,399         0         1
## 23            MICHIGAN         5.2000 10,037,261         0         1
## 24           MINNESOTA         2.7000  5,737,915         0         1
## 25         MISSISSIPPI         2.3000  2,939,690         1         0
## 26            MISSOURI         3.8000  6,196,156         1         0
## 27             MONTANA         3.3000  1,132,812         1         0
## 28            NEBRASKA         1.3000  1,978,379         1         0
## 29              NEVADA         1.7000  3,194,176         0         1
## 30       NEW HAMPSHIRE         0.7518  1,402,054         0         1
## 31          NEW JERSEY         5.1000  9,290,841         0         1
## 32          NEW MEXICO         2.6000  2,114,371         0         1
## 33            NEW YORK        10.1000 19,571,216         0         1
## 34      NORTH CAROLINA         4.5000 10,835,491         1         0
## 35        NORTH DAKOTA         1.8000    783,926         1         0
## 36                OHIO         6.6000 11,785,935         1         0
## 37            OKLAHOMA         2.9000  4,053,824         1         0
## 38              OREGON         2.3000  4,233,358         0         1
## 39        PENNSYLVANIA         8.1000 12,961,683         0         1
## 40        RHODE ISLAND         1.1000  1,095,962         0         1
## 41      SOUTH CAROLINA         2.3000  5,373,555         1         0
## 42        SOUTH DAKOTA         1.3000    919,318         1         0
## 43           TENNESSEE         3.7000  7,126,489         1         0
## 44               TEXAS        14.2000 30,503,301         1         0
## 45                UTAH         1.8000  3,417,734         1         0
## 46             VERMONT         0.8521    647,464         0         1
## 47            VIRGINIA         4.5000  8,715,698         0         1
## 48          WASHINGTON         4.0000  7,812,880         0         1
## 49       WEST VIRGINIA         2.0000  1,770,071         1         0
## 50           WISCONSIN         2.8000  5,910,955         0         1
## 51             WYOMING         2.3000    584,057         1         0
```

```r
# Rename the columns
final_df <- final_df %>%
  rename(Population = `2023`)

print(final_df)
```

```
##                  State Total_Billions Population Trump_Win Biden_Win
## 1              ALABAMA         3.0000  5,108,468         1         0
## 2               ALASKA         3.7000    733,406         1         0
## 3              ARIZONA         3.5000  7,431,344         0         1
## 4             ARKANSAS         2.8000  3,067,732         1         0
## 5           CALIFORNIA        18.4000 38,965,193         0         1
## 6             COLORADO         3.2000  5,877,610         0         1
```

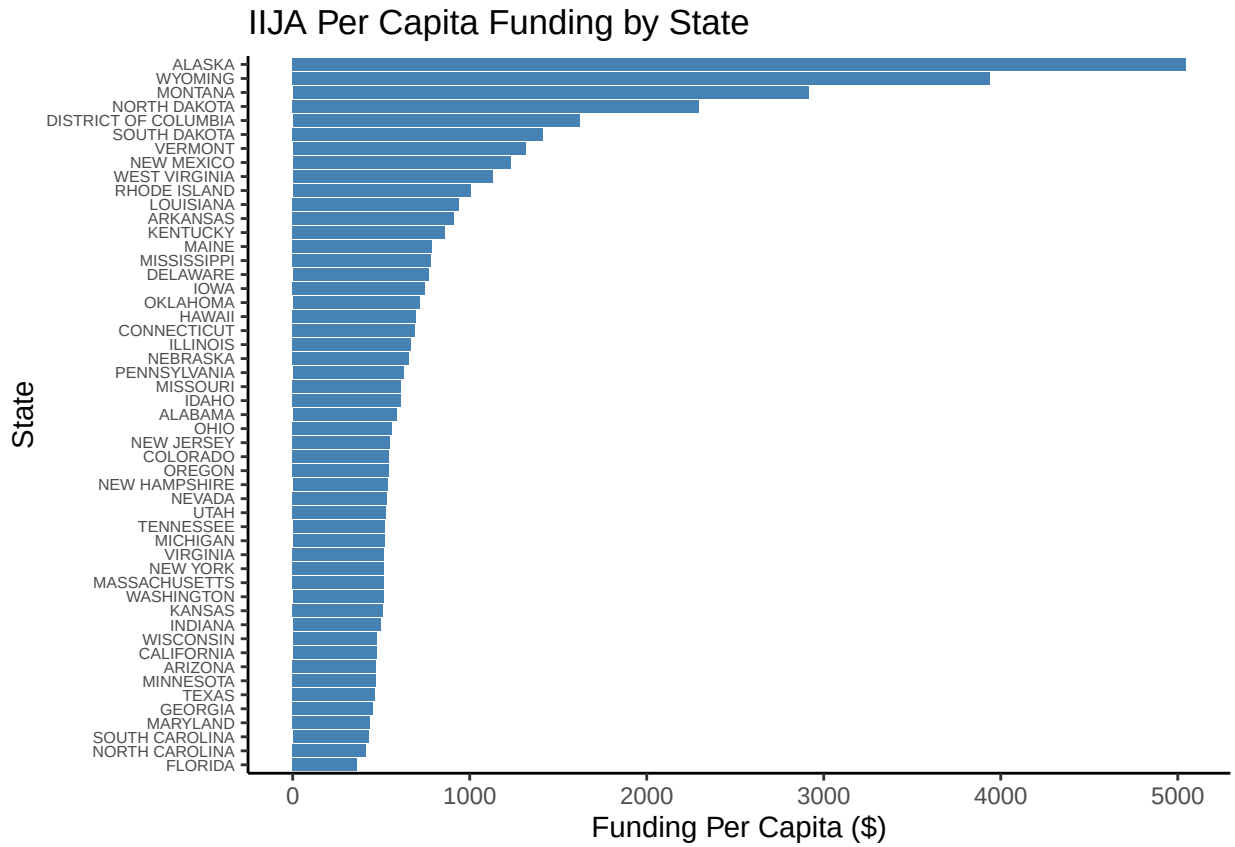
```
## 7           CONNECTICUT         2.5000  3,617,176         0         1
## 8              DELAWARE         0.7920  1,031,890         0         1
## 9  DISTRICT OF COLUMBIA         1.1000    678,972         0         1
## 10              FLORIDA         8.2000 22,610,726         1         0
## 11              GEORGIA         5.0000 11,029,227         0         1
## 12               HAWAII         1.0000  1,435,138         0         1
## 13                IDAHO         1.2000  1,964,726         1         0
## 14             ILLINOIS         8.4000 12,549,689         0         1
## 15              INDIANA         3.4000  6,862,199         1         0
## 16                 IOWA         2.4000  3,207,004         1         0
## 17               KANSAS         1.5000  2,940,546         1         0
## 18             KENTUCKY         3.9000  4,526,154         1         0
## 19            LOUISIANA         4.3000  4,573,749         1         0
## 20                MAINE         1.1000  1,395,722         0         1
## 21             MARYLAND         2.7000  6,180,253         0         1
## 22        MASSACHUSETTS         3.6000  7,001,399         0         1
## 23             MICHIGAN         5.2000 10,037,261         0         1
## 24            MINNESOTA         2.7000  5,737,915         0         1
## 25          MISSISSIPPI         2.3000  2,939,690         1         0
## 26             MISSOURI         3.8000  6,196,156         1         0
## 27              MONTANA         3.3000  1,132,812         1         0
## 28             NEBRASKA         1.3000  1,978,379         1         0
## 29               NEVADA         1.7000  3,194,176         0         1
## 30        NEW HAMPSHIRE         0.7518  1,402,054         0         1
## 31           NEW JERSEY         5.1000  9,290,841         0         1
## 32           NEW MEXICO         2.6000  2,114,371         0         1
## 33             NEW YORK        10.1000 19,571,216         0         1
## 34       NORTH CAROLINA         4.5000 10,835,491         1         0
## 35         NORTH DAKOTA         1.8000    783,926         1         0
## 36                 OHIO         6.6000 11,785,935         1         0
## 37             OKLAHOMA         2.9000  4,053,824         1         0
## 38               OREGON         2.3000  4,233,358         0         1
## 39         PENNSYLVANIA         8.1000 12,961,683         0         1
## 40         RHODE ISLAND         1.1000  1,095,962         0         1
## 41       SOUTH CAROLINA         2.3000  5,373,555         1         0
## 42         SOUTH DAKOTA         1.3000    919,318         1         0
## 43            TENNESSEE         3.7000  7,126,489         1         0
## 44                TEXAS        14.2000 30,503,301         1         0
## 45                 UTAH         1.8000  3,417,734         1         0
## 46              VERMONT         0.8521    647,464         0         1
## 47             VIRGINIA         4.5000  8,715,698         0         1
## 48           WASHINGTON         4.0000  7,812,880         0         1
## 49        WEST VIRGINIA         2.0000  1,770,071         1         0
## 50            WISCONSIN         2.8000  5,910,955         0         1
## 51              WYOMING         2.3000    584,057         1         0
```

```r
# Convert Population column to numeric
final_df$Population <- as.numeric(gsub(",", "", final_df$Population))

# Calculate per capita funding
final_df <- final_df %>%
  mutate(Per_Capita_Funding = (Total_Billions * 1e9) / Population)

head(final_df)
```

```
##        State Total_Billions Population Trump_Win Biden_Win Per_Capita_Funding
## 1    ALABAMA             3.0    5108468         1         0           587.2602
## 2     ALASKA             3.7     733406         1         0          5044.9546
## 3    ARIZONA             3.5    7431344         0         1           470.9781
## 4   ARKANSAS             2.8    3067732         1         0           912.7264
## 5 CALIFORNIA            18.4   38965193         0         1           472.2163
## 6   COLORADO             3.2    5877610         0         1           544.4390
```

First, I calculate the per capita funding for each state. Alaska, Wyoming, and Montana are the top three states that have the most funding.

```
ggplot(final_df, aes(y = reorder(State, Per_Capita_Funding),
                     x = Per_Capita_Funding)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  labs(title = "IIJA Per Capita Funding by State",
       x = "Funding Per Capita ($)",
       y = "State") +
  theme_classic() +
  theme(axis.text.y = element_text(size = 6))  # Reduce font size of state names
```

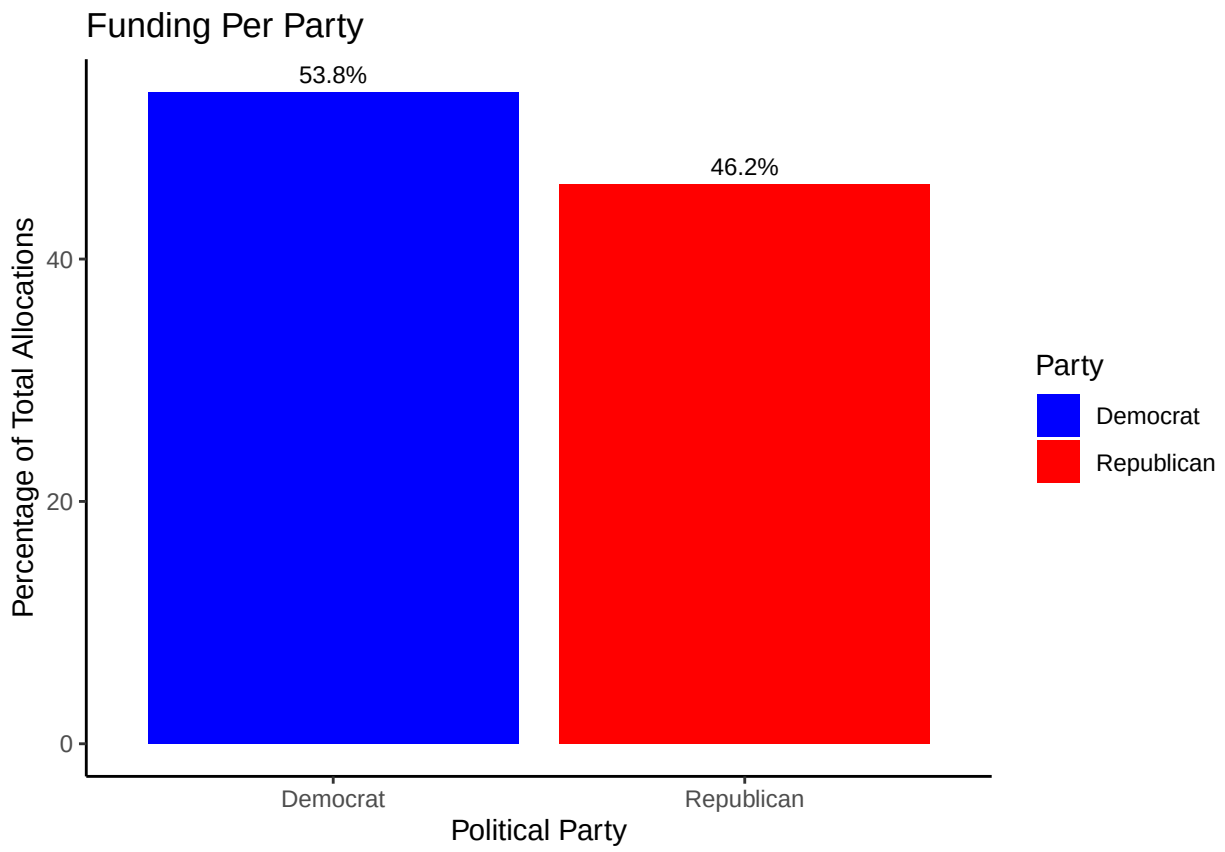


Next, I calculated the funding by party, there are 53.8% funding goes to Democrat, and 46.2% funding goes to Republican. It seems biased that Democrat got more funding than Republican.

```
final_df$Party <- ifelse(final_df$Biden_Win == 1, "Democrat",
                         ifelse(final_df$Trump_Win == 1, "Republican", "NA"))
```

```
party_funding <- final_df %>%
  group_by(Party) %>%
  summarise(Total_Funding = sum(Total_Billions)) %>%
  mutate(Percentage = round((Total_Funding / sum(Total_Funding)) * 100, 1))  # Convert to percentage
```

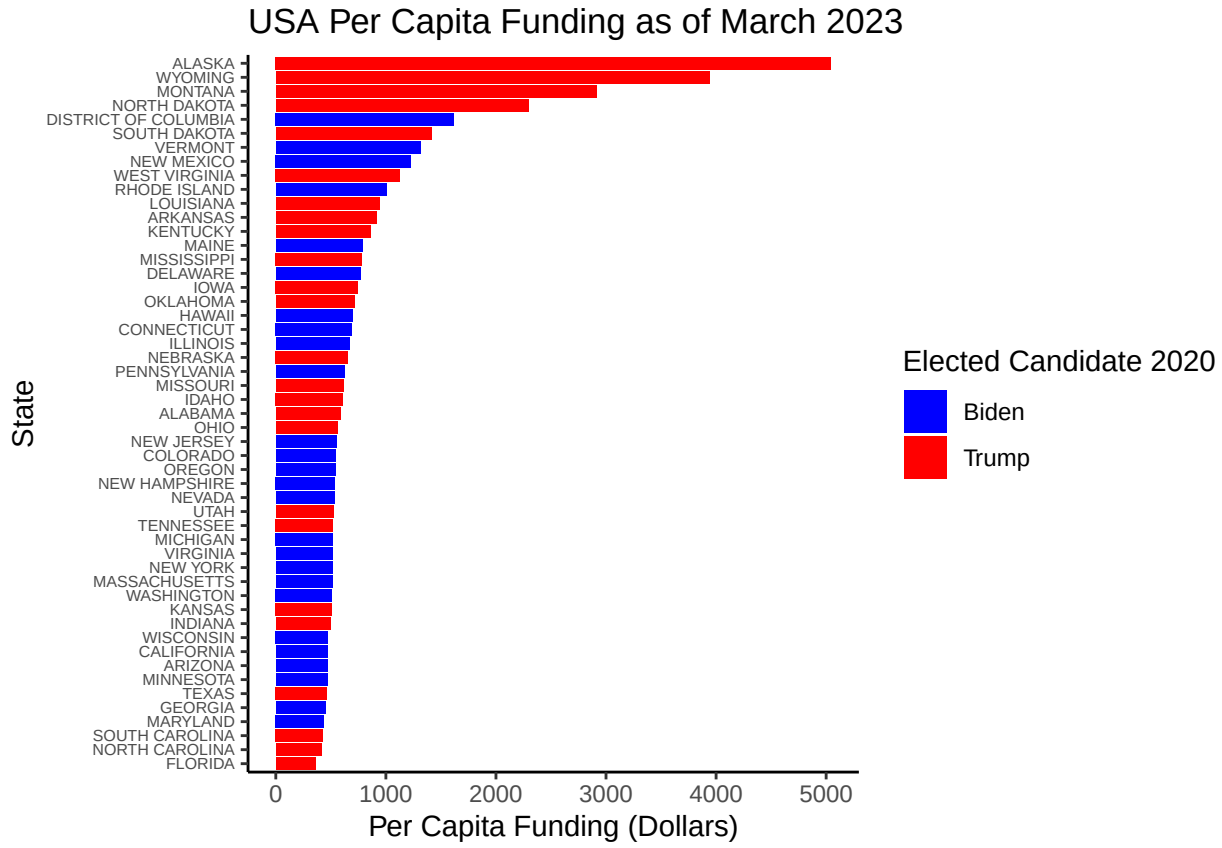
```
ggplot(party_funding, aes(x = Party, y = Percentage, fill = Party)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = paste0(Percentage, "%")), vjust = -0.5, size = 3) +  # Add percentage labels
  scale_fill_manual(values = c("Democrat" = "blue", "Republican" = "red", "NA" = "gray")) +
  labs(title = "Funding Per Party",
       x = "Political Party",
       y = "Percentage of Total Allocations",
       fill = "Party") +
  theme_classic()
```



Last, I plot the bar chart showing what party each state is below and how much funding they got. The top three states are all Republican.

```
final_df$Election_Result <- ifelse(final_df$Biden_Win == 1, "Biden", "Trump")
```

```
ggplot(final_df, aes(y = reorder(State, Per_Capita_Funding),
                     x = Per_Capita_Funding,
                     fill = Election_Result)) +
  geom_bar(stat = "identity") +
  labs(title = "USA Per Capita Funding as of March 2023",
       x = "Per Capita Funding (Dollars)",
       y = "State",
       fill = "Elected Candidate 2020") +
  scale_fill_manual(values = c("Biden" = "blue", "Trump" = "red")) +
  theme_classic()+
  theme(axis.text.y = element_text(size = 6))
```



Conclusions: Based on the graphs, it does not seem like there is an equitable distribution of funds for each state. While there are disparities exist, but our analysis suggests that the IIJA funding is driven more by state size, small and less populated states received higher funding. There is no clear partisan bias.