

# HW week 2 Data 608

Fredrick Jones

2024-02-04

First load the data into R

```
library(readxl)
```

Let's see a few rows of the given data

```
df<- readxl::read_xlsx("IIJA FUNDING AS OF MARCH 2023.xlsx")
head(df, 5)
```

```
## # A tibble: 5 x 2
##   'State, Territory or Tribal Nation' 'Total (Billions)'
##   <chr>                                <dbl>
## 1 ALABAMA                             3
## 2 ALASKA                              3.7
## 3 AMERICAN SAMOA                     0.0686
## 4 ARIZONA                             3.5
## 5 ARKANSAS                             2.8
```

Renaming the feature and target so that we can easily handle them in R

```
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
```

```
df <- df |>
  rename(State_TTN = `State, Territory or Tribal Nation`,
         Total_billions=`Total (Billions)`)
head(df,3)
```

```
## # A tibble: 3 x 2
##   State_TTN      Total_billions
##   <chr>          <dbl>
## 1 ALABAMA        3
## 2 ALASKA         3.7
## 3 AMERICAN SAMOA 0.0686
```

```
library(ggplot2)
```

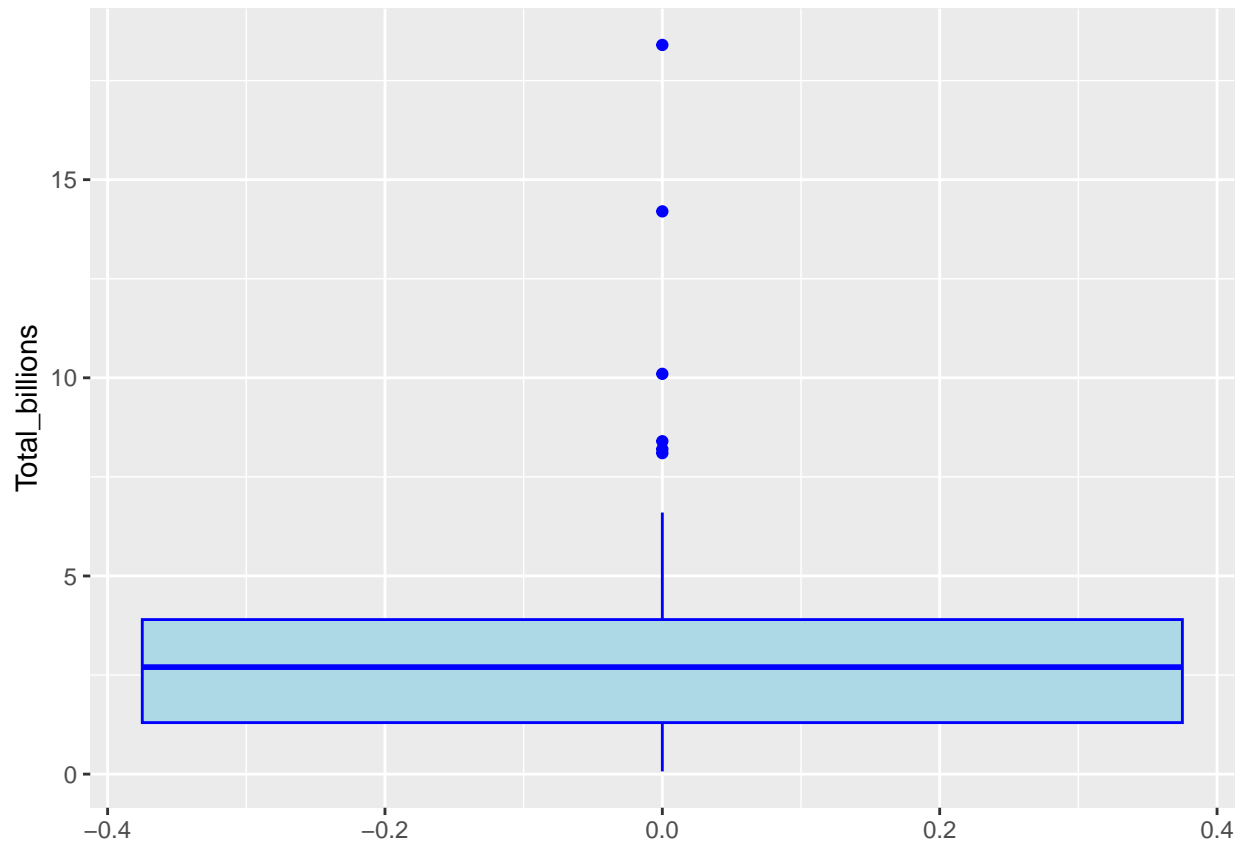
```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

```
ggplot(data = df, mapping= aes(x=State_TTN, y=Total_billions))+
  geom_point(col='red')+
  theme(axis.text.x = element_text(angle=90, vjust=0, hjust=1))
```



It can be seen that some State, Territory or Tribal Nation received more fund than others. But allocation of fund is also based on the population size and other factors such as unemployment rate. To check if the allocation of fund is affected by political interest of Biden administration, it will need to find fund of allocation per capita. So, the fund should be divided by the population. Based on the data given we can not claim that Biden administration allocated fund according to their political interest.

```
ggplot(data = df, mapping= aes(Total_billions))+
  geom_boxplot(color = 'blue', fill='lightblue')+
  coord_flip()
```



The box plot shows that there are some outliers in the allocation of funds. This might be due to the fact that Biden administration might have allocated fund according to their political interest. But to state it firmly, we will have to study other parameters of the states that received much more fund than other states.