Visual Analysis on data of US Department of Labor's Office of Foreign Labor Certification, Employment and Training Administration

Yash Rajaram Bagayatkar

10/7/2020

Introduction

In this R file, I have aimed to create a comprehensive report with visualizations providing insights on the data provided by the US Department of Labor's Office of Foreign Labor Certification, Employment and Training Administration about the employer's request on the H1B non-immigrant workers. The information about the data's attributes is provided in the Appendix section at the bottom of this report

Initialization

In this section, I have loaded the necessary libraries and the H1B data. I have also replaced any spaces in the column names with '.'

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.3
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.1
## v tidyr   1.1.1     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## Warning: package 'ggplot2' was built under R version 3.6.3
```

```
## Warning: package 'tibble' was built under R version 3.6.3
```

```
## Warning: package 'tidyr' was built under R version 3.6.3
```

```
## Warning: package 'readr' was built under R version 3.6.3
```

```
## Warning: package 'purrr' was built under R version 3.6.3
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

```
## Warning: package 'forcats' was built under R version 3.6.3
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.6.3
```

```
##
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':
##
##     between, first, last

## The following object is masked from 'package:purrr':
##
##     transpose
```

```
library(ggplot2)
library(dplyr)
#Reading the NYC 311 Data set
h1b<-fread("h1bdata.csv")
names(h1b)<-names(h1b) %>%
  stringr::str_replace_all("\\s",".")
```

Tidying the Dataset

After initialization, I have performed cleaning of the data to deal with missing values and also made changes with respect to some columns depending on the need for analysis. Instead of adding mean, or default values for numeric or categorical missing values, I have removed those rows altogether as the missing rows are very less as compared to the length of the whole dataset. I have removed empty string values from EMPLOYEE_STATE, PW_UNIT_OF_PAY and WAGE_UNIT_OF_PAY. I have also converted all the PREVAILING_WAGE, WAGE_RATE_OF_PAY_FROM, WAGE_RATE_OF_PAY_TO values to Hourly Wages out of the possible categories of Yearly, Monthly, Weekly, Bi-Weekly and Hourly. I have done this so that there is consistency in the wage related analysis that I have performed in the later section.

```
sum(is.na(h1b))
```

```
## [1] 33
```

```
nrow(h1b)
```

```
## [1] 528134
```

```
h1b<-na.omit(h1b)
sapply(h1b, class)
```

```
##    CASE_SUBMITTED_DAY  CASE_SUBMITTED_MONTH   CASE_SUBMITTED_YEAR
##             "integer"             "integer"            "integer"
##          DECISION_DAY        DECISION_MONTH         DECISION_YEAR
##             "integer"             "integer"            "integer"
##            VISA_CLASS         EMPLOYER_NAME        EMPLOYER_STATE
##           "character"           "character"           "character"
##      EMPLOYER_COUNTRY              SOC_NAME            NAICS_CODE
##           "character"           "character"            "integer"
##         TOTAL_WORKERS    FULL_TIME_POSITION       PREVAILING_WAGE
##             "integer"           "character"            "numeric"
##        PW_UNIT_OF_PAY             PW_SOURCE        PW_SOURCE_YEAR
##           "character"           "character"            "integer"
##       PW_SOURCE_OTHER WAGE_RATE_OF_PAY_FROM   WAGE_RATE_OF_PAY_TO
##           "character"             "numeric"            "numeric"
##      WAGE_UNIT_OF_PAY        H-1B_DEPENDENT      WILLFUL_VIOLATOR
##           "character"           "character"           "character"
##        WORKSITE_STATE  WORKSITE_POSTAL_CODE           CASE_STATUS
```

```
##           "character"           "character"           "character"
```

```
head(h1b)
```

```
##    CASE_SUBMITTED_DAY CASE_SUBMITTED_MONTH CASE_SUBMITTED_YEAR DECISION_DAY
## 1:                 24                    2                2016           1
## 2:                  4                    3                2016           1
## 3:                 10                    3                2016           1
## 4:                 28                    9                2016           1
## 5:                 22                    2                2015           2
## 6:                 12                    3                2015           2
##    DECISION_MONTH DECISION_YEAR VISA_CLASS             EMPLOYER_NAME
## 1:             10          2016        H1B     DISCOVER PRODUCTS INC
## 2:             10          2016        H1B          DFS SERVICES LLC
## 3:             10          2016        H1B EASTBANC TECHNOLOGIES LLC
## 4:             10          2016        H1B         INFO SERVICES LLC
## 5:             10          2016        H1B        BBandT CORPORATION
## 6:             10          2016        H1B        SUNTRUST BANKS INC
##    EMPLOYER_STATE         EMPLOYER_COUNTRY            SOC_NAME NAICS_CODE
## 1:             IL UNITED STATES OF AMERICA            ANALYSTS     522210
## 2:             IL UNITED STATES OF AMERICA            ANALYSTS     522210
## 3:             DC UNITED STATES OF AMERICA            ANALYSTS     541511
## 4:             MI UNITED STATES OF AMERICA COMPUTER OCCUPATION     541511
## 5:             NC UNITED STATES OF AMERICA            ANALYSTS     522110
## 6:             GA UNITED STATES OF AMERICA            ANALYSTS     522110
##    TOTAL_WORKERS FULL_TIME_POSITION PREVAILING_WAGE PW_UNIT_OF_PAY PW_SOURCE
## 1:             1                  Y           59197          Year       OES
## 2:             1                  Y           49800          Year     Other
## 3:             2                  Y           76502          Year       OES
## 4:             1                  Y           90376          Year       OES
## 5:             1                  Y          116605          Year       OES
## 6:             1                  Y           59405          Year       OES
##    PW_SOURCE_YEAR             PW_SOURCE_OTHER WAGE_RATE_OF_PAY_FROM
## 1:           2015     OFLC ONLINE DATA CENTER                 65811
## 2:           2015 WILLIS TOWERS WATSON SURVEY                 53000
## 3:           2015     OFLC ONLINE DATA CENTER                 77000
## 4:           2016     OFLC ONLINE DATA CENTER                102000
## 5:           2015     OFLC ONLINE DATA CENTER                132500
## 6:           2015     OFLC ONLINE DATA CENTER                 71750
##    WAGE_RATE_OF_PAY_TO WAGE_UNIT_OF_PAY H-1B_DEPENDENT WILLFUL_VIOLATOR
## 1:               67320            Year              N               N
## 2:               57200            Year              N               N
## 3:                   0            Year              Y               N
## 4:                   0            Year              Y               N
## 5:                   0            Year              N               N
## 6:                   0            Year              N               N
##    WORKSITE_STATE WORKSITE_POSTAL_CODE        CASE_STATUS
## 1:             IL                60015 CERTIFIEDWITHDRAWN
## 2:             IL                60015 CERTIFIEDWITHDRAWN
## 3:             DC                20007 CERTIFIEDWITHDRAWN
## 4:             NJ                 7302          WITHDRAWN
## 5:             NY                10036 CERTIFIEDWITHDRAWN
## 6:             GA                30303 CERTIFIEDWITHDRAWN
```

```
h1b<-dplyr::filter(h1b,h1b$EMPLOYER_STATE!="")
h1b<-dplyr::filter(h1b,h1b$PW_UNIT_OF_PAY!="")
h1b<-dplyr::filter(h1b,h1b$WAGE_UNIT_OF_PAY!="")
h1b<-within(h1b,{
  temp<-PW_UNIT_OF_PAY=="Year"
  PREVAILING_WAGE[temp]<-PREVAILING_WAGE/(52*40)
  PW_UNIT_OF_PAY<-"Hour"
})
```

```
## Warning in PREVAILING_WAGE[temp] <- PREVAILING_WAGE/(52 * 40): number of items
## to replace is not a multiple of replacement length
```

```
h1b<-within(h1b,{
  temp<-PW_UNIT_OF_PAY=="Month"
  PREVAILING_WAGE[temp]<-PREVAILING_WAGE*(12/(52*40))
  PW_UNIT_OF_PAY<-"Hour"
})
h1b<-within(h1b,{
  temp<-PW_UNIT_OF_PAY=="Week"
  PREVAILING_WAGE[temp]<-PREVAILING_WAGE/40
  PW_UNIT_OF_PAY<-"Hour"
})
h1b<-within(h1b,{
  temp<-PW_UNIT_OF_PAY=="Bi-Weekly"
  PREVAILING_WAGE[temp]<-PREVAILING_WAGE/80
  PW_UNIT_OF_PAY<-"Hour"
})

h1b<-within(h1b,{
  temp<-WAGE_UNIT_OF_PAY=="Year"
  WAGE_RATE_OF_PAY_FROM[temp]<-WAGE_RATE_OF_PAY_FROM/(52*40)
  WAGE_RATE_OF_PAY_TO[temp]<-WAGE_RATE_OF_PAY_TO/(52*40)
  WAGE_UNIT_OF_PAY<-"Hour"
})
```

```
## Warning in WAGE_RATE_OF_PAY_FROM[temp] <- WAGE_RATE_OF_PAY_FROM/(52 * 40):
## number of items to replace is not a multiple of replacement length
```

```
## Warning in WAGE_RATE_OF_PAY_TO[temp] <- WAGE_RATE_OF_PAY_TO/(52 * 40): number of
## items to replace is not a multiple of replacement length
```

```
h1b<-within(h1b,{
  temp<-WAGE_UNIT_OF_PAY=="Month"
  WAGE_RATE_OF_PAY_FROM[temp]<-WAGE_RATE_OF_PAY_FROM*(12/(52*40))
  WAGE_RATE_OF_PAY_TO[temp]<-WAGE_RATE_OF_PAY_TO*(12/(52*40))
  WAGE_UNIT_OF_PAY<-"Hour"
})
h1b<-within(h1b,{
  temp<-WAGE_UNIT_OF_PAY=="Week"
  WAGE_RATE_OF_PAY_FROM[temp]<-WAGE_RATE_OF_PAY_FROM/40
  WAGE_RATE_OF_PAY_TO[temp]<-WAGE_RATE_OF_PAY_TO/40
  WAGE_UNIT_OF_PAY<-"Hour"
})
h1b<-within(h1b,{
  temp<-WAGE_UNIT_OF_PAY=="Bi-Weekly"
```

```
  WAGE_RATE_OF_PAY_FROM[temp]<-WAGE_RATE_OF_PAY_FROM/80
  WAGE_RATE_OF_PAY_TO[temp]<-WAGE_RATE_OF_PAY_TO/80
  WAGE_UNIT_OF_PAY<-"Hour"
})
head(h1b)
```

```
##    CASE_SUBMITTED_DAY CASE_SUBMITTED_MONTH CASE_SUBMITTED_YEAR DECISION_DAY
## 1:                 24                    2                2016            1
## 2:                  4                    3                2016            1
## 3:                 10                    3                2016            1
## 4:                 28                    9                2016            1
## 5:                 22                    2                2015            2
## 6:                 12                    3                2015            2
##    DECISION_MONTH DECISION_YEAR VISA_CLASS             EMPLOYER_NAME
## 1:             10          2016        H1B     DISCOVER PRODUCTS INC
## 2:             10          2016        H1B           DFS SERVICES LLC
## 3:             10          2016        H1B EASTBANC TECHNOLOGIES LLC
## 4:             10          2016        H1B          INFO SERVICES LLC
## 5:             10          2016        H1B         BBandT CORPORATION
## 6:             10          2016        H1B         SUNTRUST BANKS INC
##    EMPLOYER_STATE         EMPLOYER_COUNTRY             SOC_NAME NAICS_CODE
## 1:             IL UNITED STATES OF AMERICA             ANALYSTS     522210
## 2:             IL UNITED STATES OF AMERICA             ANALYSTS     522210
## 3:             DC UNITED STATES OF AMERICA             ANALYSTS     541511
## 4:             MI UNITED STATES OF AMERICA COMPUTER OCCUPATION     541511
## 5:             NC UNITED STATES OF AMERICA             ANALYSTS     522110
## 6:             GA UNITED STATES OF AMERICA             ANALYSTS     522110
##    TOTAL_WORKERS FULL_TIME_POSITION PREVAILING_WAGE PW_UNIT_OF_PAY PW_SOURCE
## 1:             1                  Y        28.46010           Hour       OES
## 2:             1                  Y        23.94231           Hour     Other
## 3:             2                  Y        36.77981           Hour       OES
## 4:             1                  Y        43.45000           Hour       OES
## 5:             1                  Y        56.06010           Hour       OES
## 6:             1                  Y        28.56010           Hour       OES
##    PW_SOURCE_YEAR            PW_SOURCE_OTHER WAGE_RATE_OF_PAY_FROM
## 1:           2015     OFLC ONLINE DATA CENTER              31.63990
## 2:           2015 WILLIS TOWERS WATSON SURVEY              25.48077
## 3:           2015     OFLC ONLINE DATA CENTER              37.01923
## 4:           2016     OFLC ONLINE DATA CENTER              49.03846
## 5:           2015     OFLC ONLINE DATA CENTER              63.70192
## 6:           2015     OFLC ONLINE DATA CENTER              34.49519
##    WAGE_RATE_OF_PAY_TO WAGE_UNIT_OF_PAY H-1B_DEPENDENT WILLFUL_VIOLATOR
## 1:            32.36538             Hour              N                N
## 2:            27.50000             Hour              N                N
## 3:             0.00000             Hour              Y                N
## 4:             0.00000             Hour              Y                N
## 5:             0.00000             Hour              N                N
## 6:             0.00000             Hour              N                N
##    WORKSITE_STATE WORKSITE_POSTAL_CODE        CASE_STATUS  temp
## 1:             IL                60015 CERTIFIEDWITHDRAWN FALSE
## 2:             IL                60015 CERTIFIEDWITHDRAWN FALSE
## 3:             DC                20007 CERTIFIEDWITHDRAWN FALSE
## 4:             NJ                 7302          WITHDRAWN FALSE
## 5:             NY                10036 CERTIFIEDWITHDRAWN FALSE
```

```
## 6:             GA                30303 CERTIFIEDWITHDRAWN FALSE
```

Exploring the Dataset

The exploring section deals with the core of this report highlighting observations made using the visualizations. I have performed series of visualizations for seasonal analysis, OOccupational analysis and geographical analysis in succession.

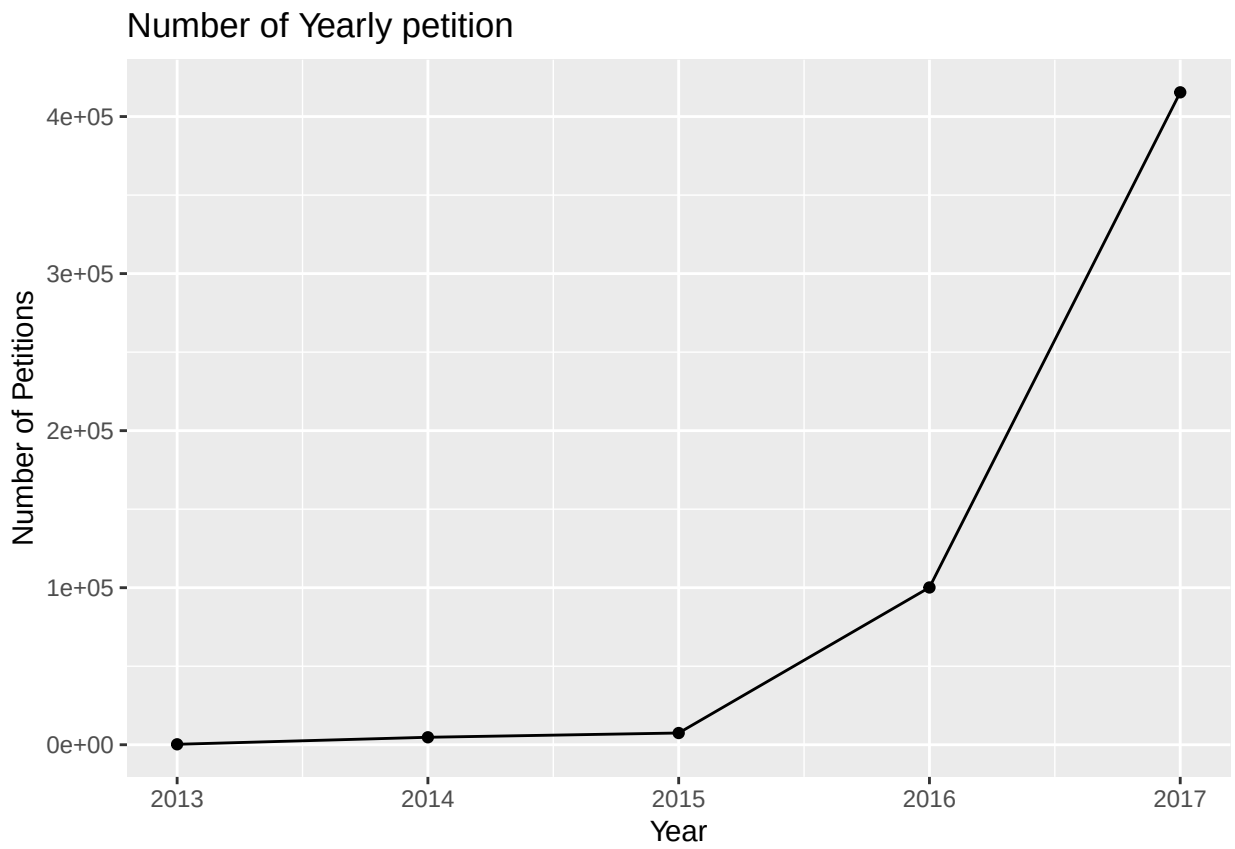
```
# Frequency per year
h1b %>%
  group_by(CASE_SUBMITTED_YEAR) %>%
  summarise(count=n())
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 7 x 2
##   CASE_SUBMITTED_YEAR  count
##                 <int>  <int>
## 1                2011      2
## 2                2012     11
## 3                2013    279
## 4                2014   4750
## 5                2015   7466
## 6                2016 100147
## 7                2017 415426
```

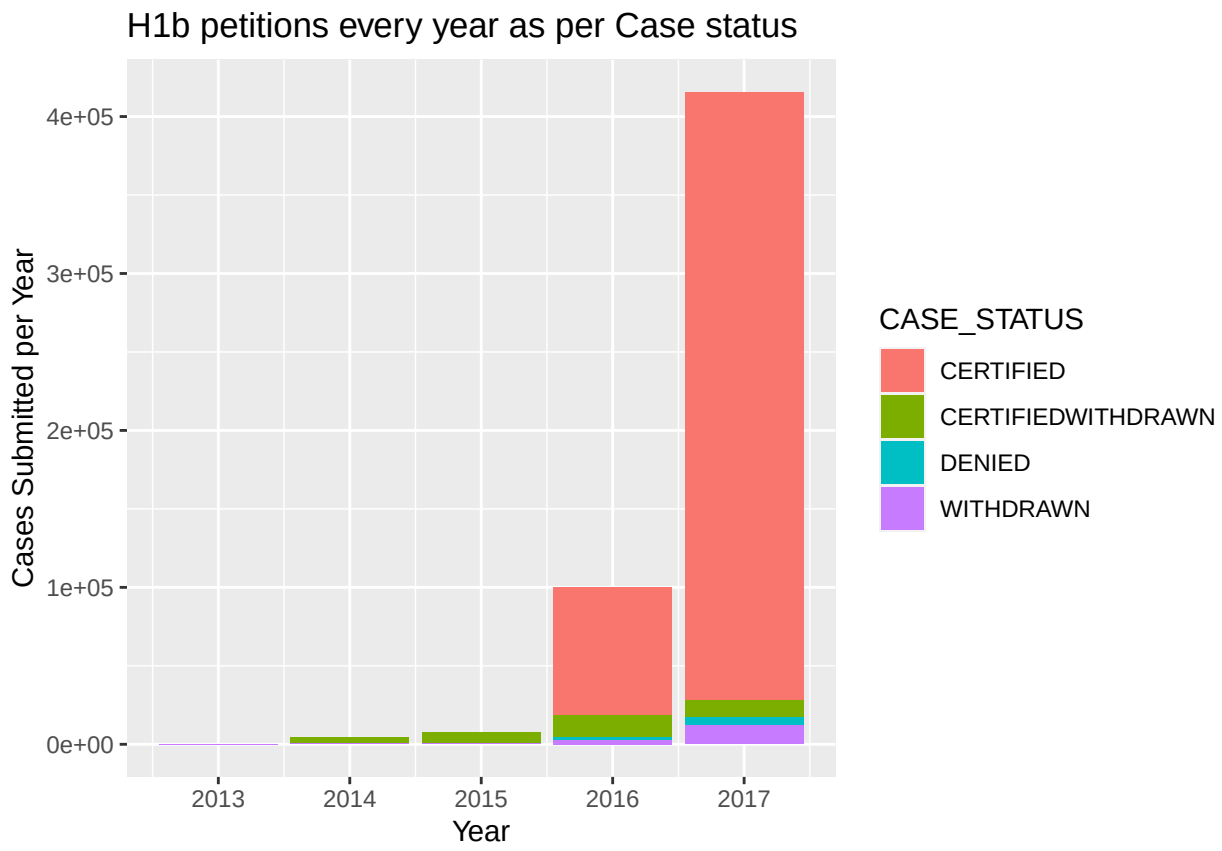
```
h1b<-dplyr::filter(h1b,CASE_SUBMITTED_YEAR!=2011 & CASE_SUBMITTED_YEAR!=2012)
#Number of Yearly petition
h1b %>%
  group_by(CASE_SUBMITTED_YEAR) %>%
  summarise(count=n()) %>%
  ggplot(aes(x=CASE_SUBMITTED_YEAR,y=count,group=1)) +
  geom_line() +
  geom_point() +
  xlab("Year") +
  ylab("Number of Petitions") +
  ggtitle("Number of Yearly petition")
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```



I have removed the years 2011 and 2012 years from CASE_SUBMITTED_YEAR column as they amount to just a handful observations which do not contribute to a data with 530000 observations. In the above visualization, the drastic increase can be seen at the year 2016 and then a further increase 2017.

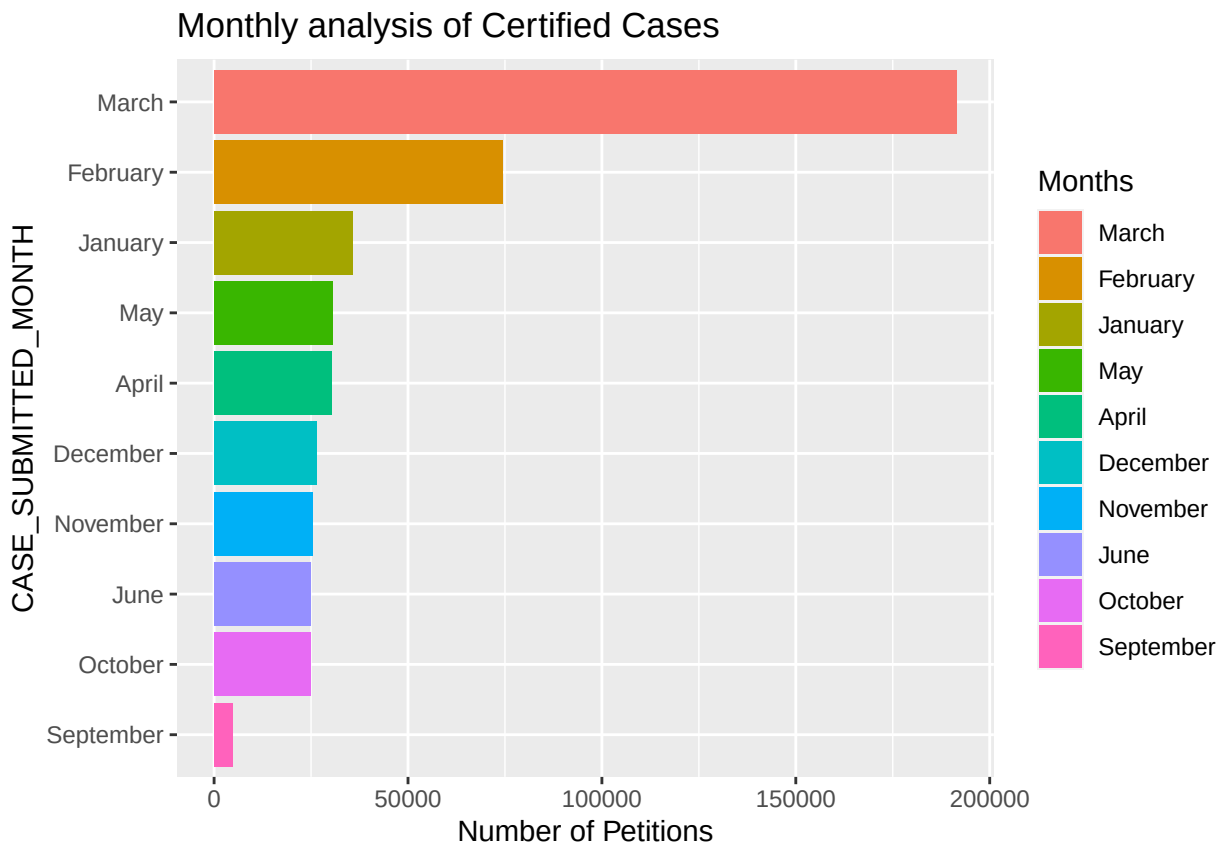
```
#H1b petitions every year as per Case_status
ggplot(h1b) +
  geom_bar(aes(x=CASE_SUBMITTED_YEAR,fill=CASE_STATUS)) +
  xlab("Year") +
  ylab("Cases Submitted per Year") +
  ggtitle("H1b petitions every year as per Case status")
```



The above visualization is performed to understand how the case statuses out of CERTIFIED, CERTIFIED-WITHDRAWN, DENIED AND WITHDRAWN are distributed along the years. We can see that primarily, the status for the cases in the dataset is certified.

```
# Seasonal analysis of Certified case status Petitions
Months<-c("January","February","March","April","May","June","July","August","September","October","Nover
h1b$Case_Submitted_Month_Name<-Months[h1b$CASE_SUBMITTED_MONTH]
h1b.Certified<-dplyr::filter(h1b,h1b$CASE_STATUS=="CERTIFIED")
h1b.Certified%>%
  group_by(Case_Submitted_Month_Name)%>%
  summarise(count=n())%>%
  ggplot(aes(x=reorder(Case_Submitted_Month_Name,count),y=count,fill=reorder(Case_Submitted_Month_Name,
  scale_fill_discrete(name="Months") +
  geom_bar(stat="identity") +
  xlab("CASE_SUBMITTED_MONTH") +
  ylab("Number of Petitions") +
  ggtitle("Monthly analysis of Certified Cases") +
  coord_flip()
```

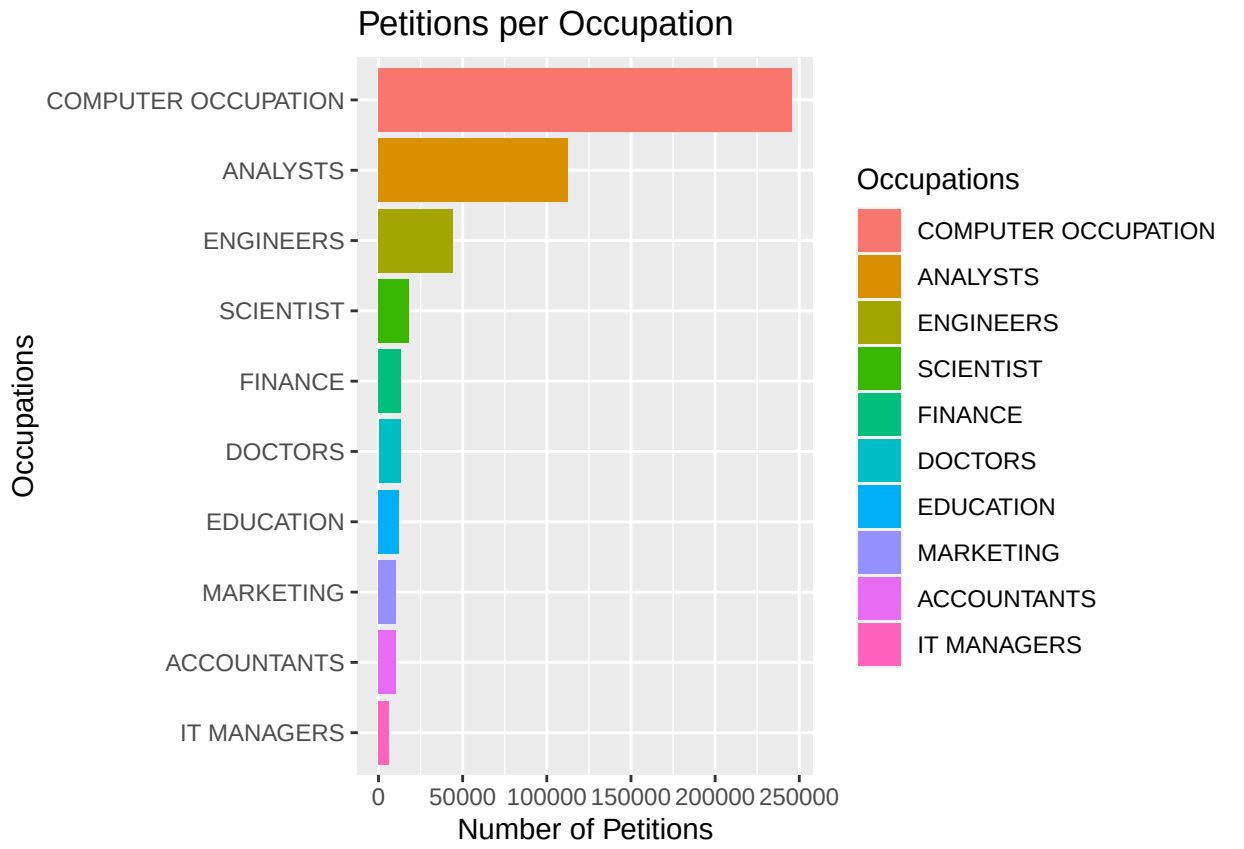
```
## `summarise()` ungrouping output (override with `.groups` argument)
```



For the next visualization, I have created a list of Month names as the CASE_SUBMITTED_MONTH column is numerical. The visualization shows a seasonal analysis throughout the years for every month. We can see from the above chart that H1B cases are primarily submitted by employers in the month of March followed by a steep decline on the month of February and then January. This states that the beginning of the year is when the maximum cases are submitted for H1B.

```
# Petitions per Occupation
h1b %>%
  group_by(SOC_NAME)%>%
  summarise(count=n())%>%
  arrange(desc(count))%>%
  head(10)%>%
  ggplot(aes(x=reorder(SOC_NAME,count),y=count,fill=reorder(SOC_NAME,-count)))+
  scale_fill_discrete(name="Occupations") +
  geom_bar(stat="identity") +
  xlab("Occupations") +
  ylab("Number of Petitions") +
  ggtitle("Petitions per Occupation") +
  coord_flip()
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```



I have then moved towards an occupational analysis of the dataset. In the above dataset, I have shown the number of H1B petitions per occupation highlighting the top 10 occupations. The highest number in this case is of the field of computer related occupations. This overwhelming majority is followed by analysts and then engineers which are the in demand jobs for which employers apply for H1B status.

```
library(scales)
```

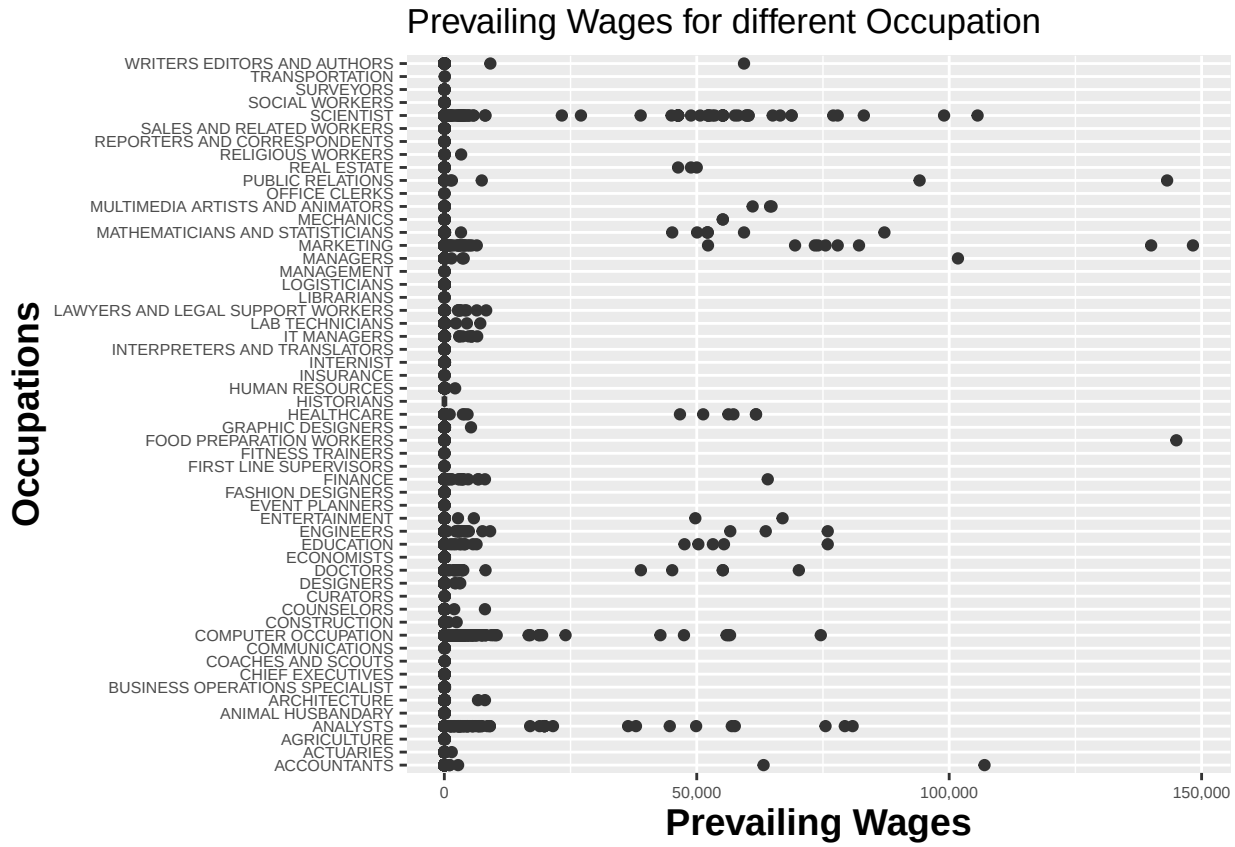
```
## Warning: package 'scales' was built under R version 3.6.3
```

```
##
## Attaching package: 'scales'
```

```
## The following object is masked from 'package:purrr':
##
##     discard
```

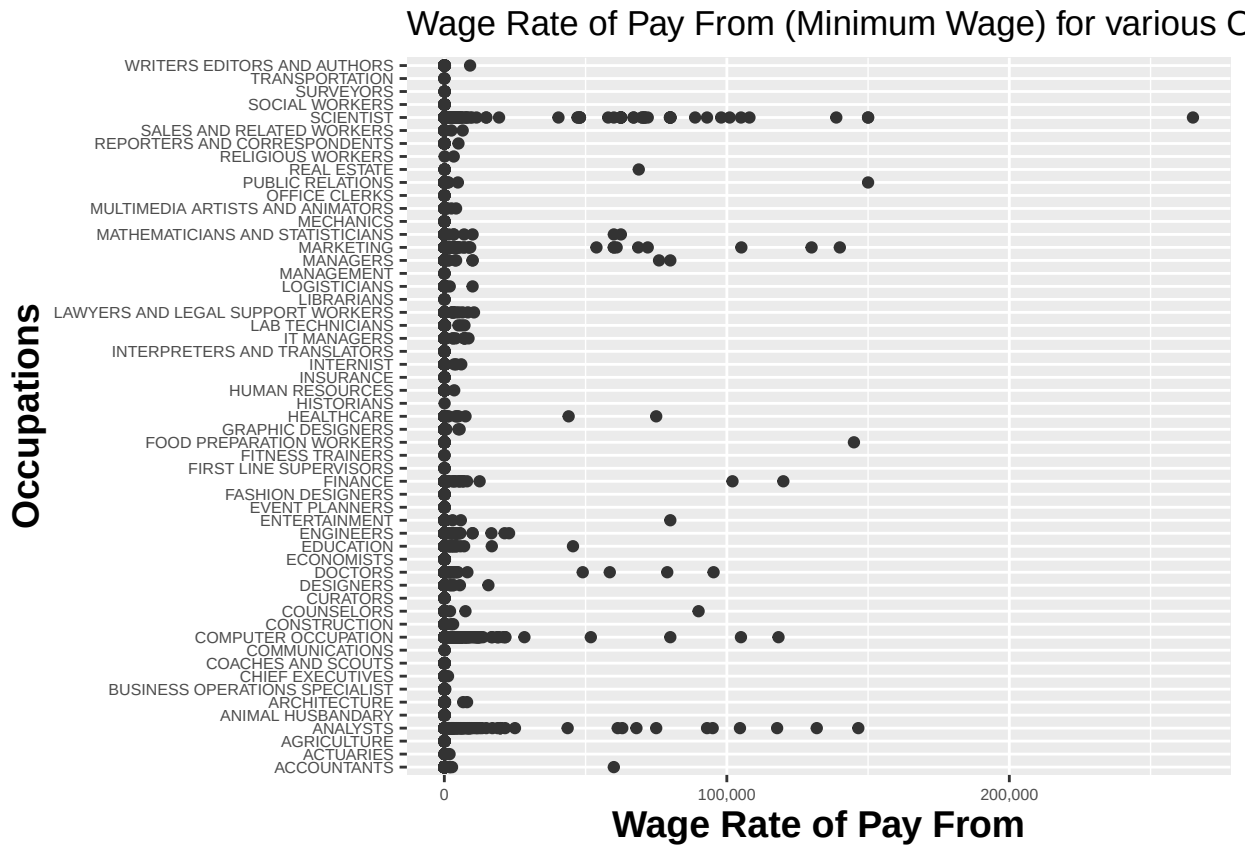
```
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
# Range Prevailing Wage for different Occupation
ggplot(h1b) +
  geom_boxplot(aes(PREVAILING_WAGE,SOC_NAME))+
  xlab("Prevailing Wages") +
  ylab("Occupations") +
  ggtitle("Prevailing Wages for different Occupation") +
  theme(axis.text=element_text(size=5.5),
        axis.title=element_text(size=14,face="bold")) +
  scale_x_continuous(labels=comma)
```



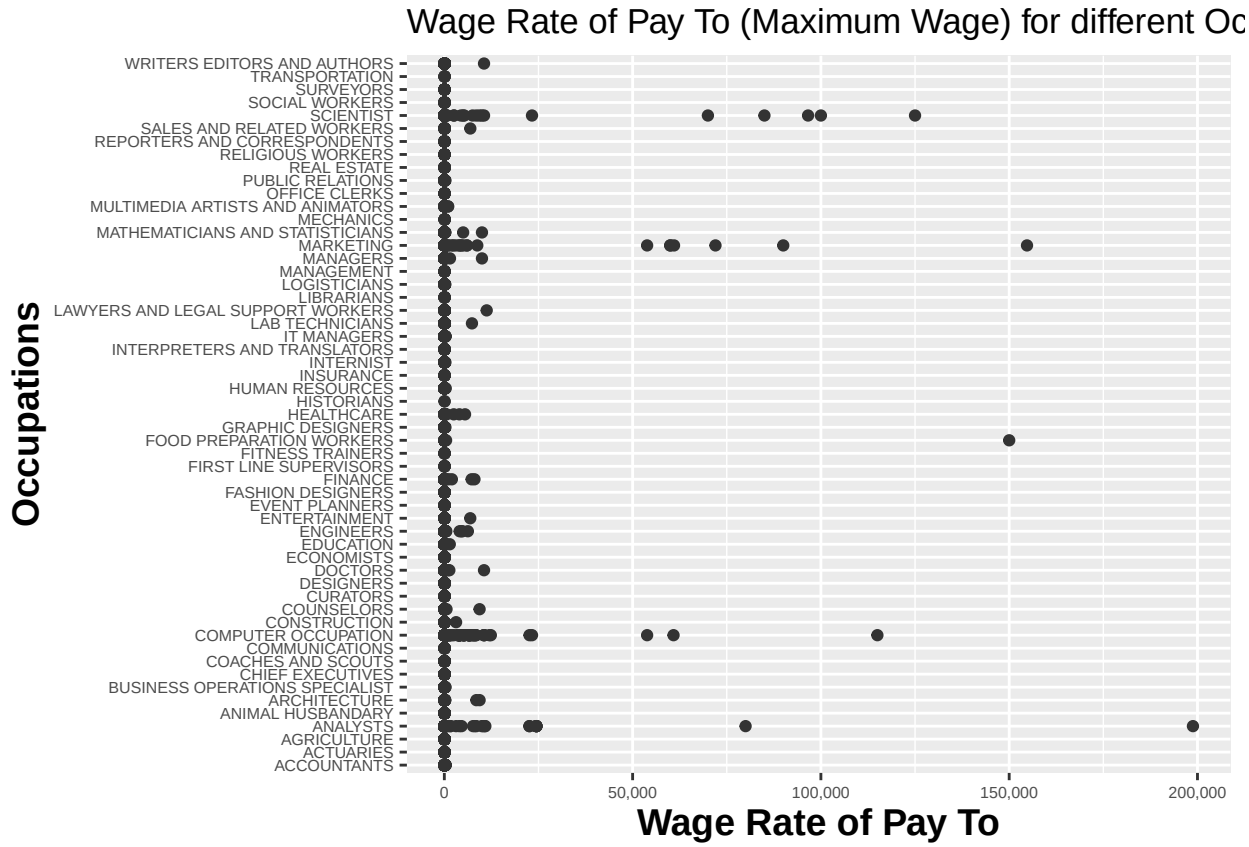
After seeing the number of petitions for each occupation, we can now move on to the wages where in the above visualization, we can see the various jobs and their boxplots showing the range for hourly prevailing wages. Although the previous chart leaders of Computer occupation, Analysts show variation, Engineers shown minimal variation in prevailing wages. On the contrary, considerable variation is shown for occupations in Marketing, and that of Scientists.

```
library(scales)
# Range of Wage Rate per Pay From (Minimum) for different Occupation
ggplot(h1b) +
  geom_boxplot(aes(WAGE_RATE_OF_PAY_FROM,SOC_NAME))+
  xlab("Wage Rate of Pay From") +
  ylab("Occupations") +
  ggtitle("Wage Rate of Pay From (Minimum Wage) for various Occupation") +
  theme(axis.text=element_text(size=5.5),
        axis.title=element_text(size=14,face="bold")) +
  scale_x_continuous(labels=comma)
```



Speaking of boxplots and wages, the next chart is also an occupational boxplot showing variation among the Beginning spectrum of the wage rate. Again, Computer occupation, Analysts show variation along with Scientists and Marketing occupations.

```
library(scales)
# Range of Wage Rate per Pay To (Maximum) for different Occupation
ggplot(h1b) +
  geom_boxplot(aes(WAGE_RATE_OF_PAY_TO,SOC_NAME))+
  xlab("Wage Rate of Pay To") +
  ylab("Occupations") +
  ggtitle("Wage Rate of Pay To (Maximum Wage) for different Occupation") +
  theme(axis.text=element_text(size=5.5),
        axis.title=element_text(size=14,face="bold")) +
  scale_x_continuous(labels=comma)
```

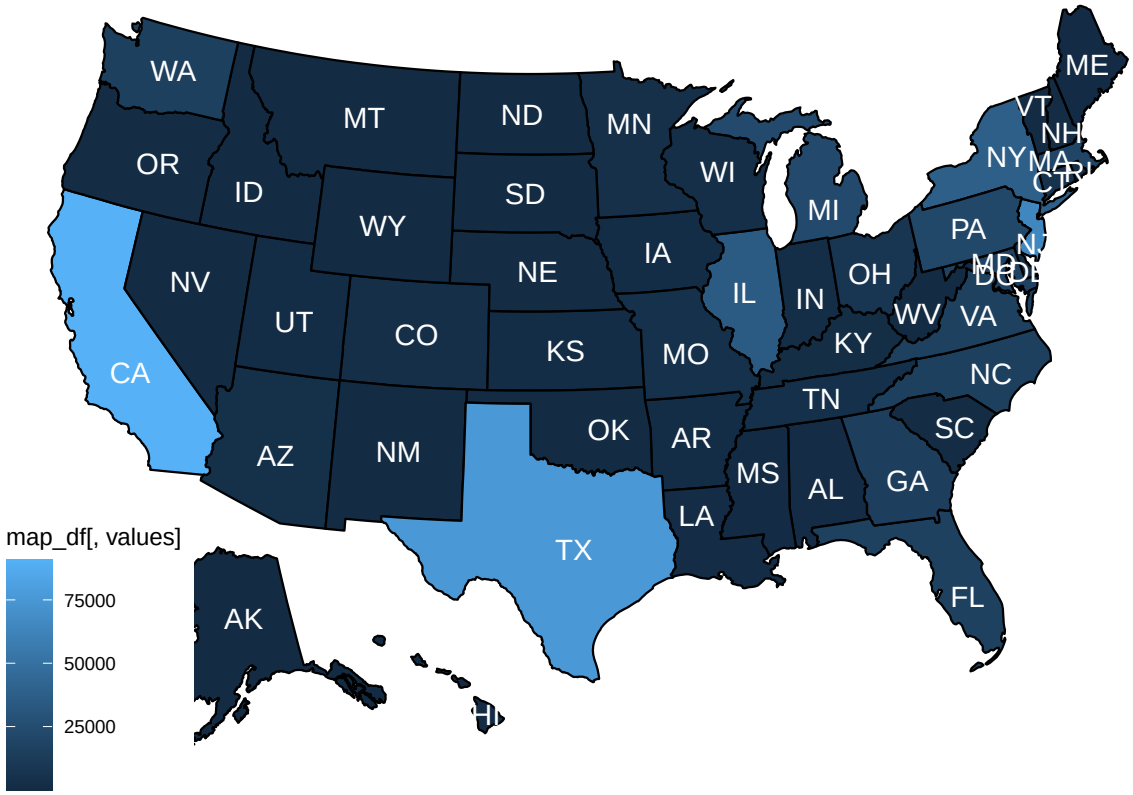


The last graph in the boxplot section is for the Wage Rate of Pay To column which is the higher side of spectrum for wage rates. The Computer occupation and Analysts do not show variation here hinting that there is a relatively constant wage rate for the higher end of the wage rate range. The scientists and marketing occupation still shows some variation in wage, but not drastic.

```
# USA Map for number of Petitions
library(usmap)
df1<-h1b%>%
  group_by(EMPLOYER_STATE)%>%
  summarise(count=n())
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
colnames(df1)[colnames(df1)=="EMPLOYER_STATE"]<-"state"
plot_usmap(regions=c("states"), data=df1,values="count",labels=TRUE,label_color="white")
```

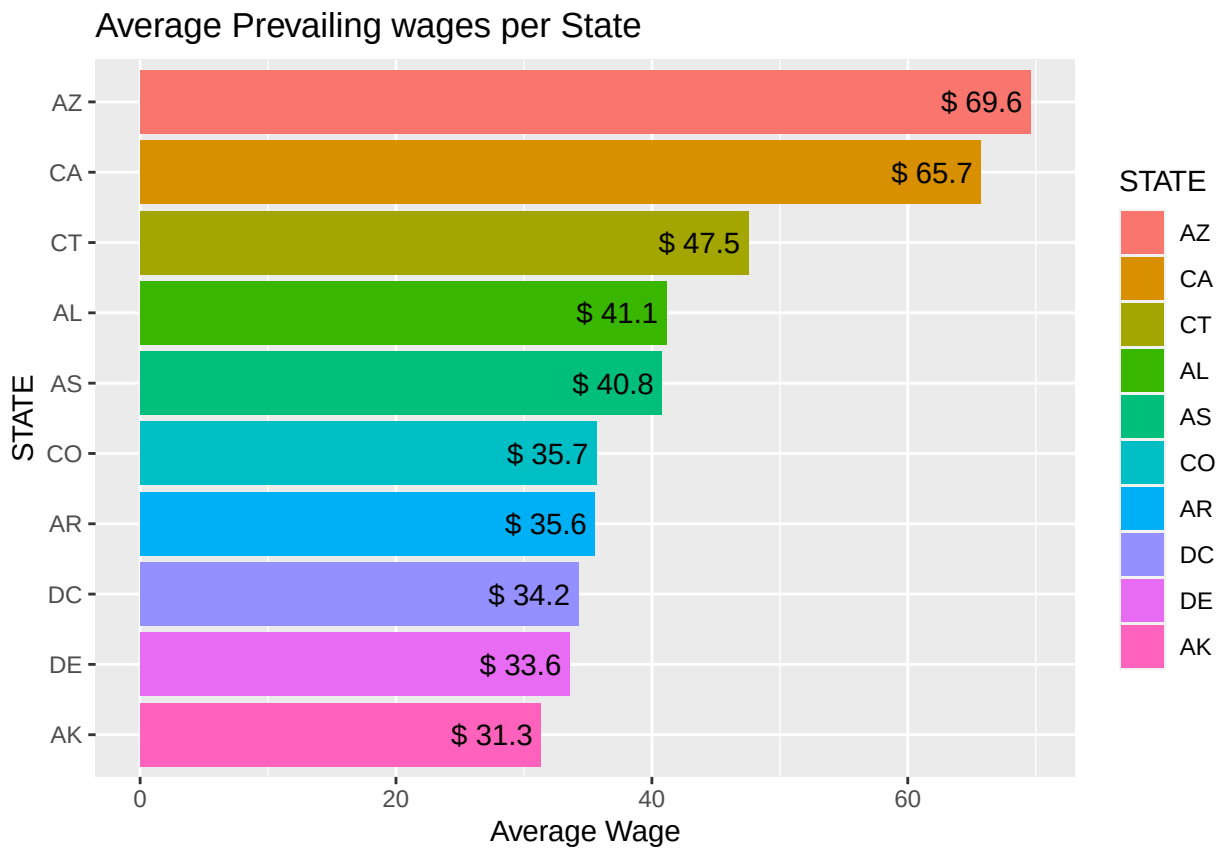


Above is the map that I have used to show the geographical distribution of the number of H1B petitions where largely many states are seemingly almost evenly distributed distributed towards the lower side of the spectrum. However two light blue states show drastic increase in the number of petitions as compared to the other states and they are California and Texas followed by New Jeysey in the North East part of the United States of America.

```
#Average Prevailing wages per State

h1b%>%
  group_by(EMPLOYER_STATE)%>%
  summarise(Mean=mean(PREVAILING_WAGE))%>%
  head(10)%>%
  ggplot(aes(x=reorder(EMPLOYER_STATE,Mean),y=Mean,fill=reorder(EMPLOYER_STATE,-Mean)))+
  scale_fill_discrete(name="STATE") +
  geom_bar(stat="identity") +
  geom_text(aes(label=paste("$",round(Mean,1))),hjust=1.1) +
  xlab("STATE") +
  ylab("Average Wage") +
  ggtitle("Average Prevailing wages per State") +
  coord_flip()
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```



Next, I have checked for the average prevailing wages per hour for the States with top 10 prevailing wages. It can be found that Arizona is leading this chart with an average prevailing wage of $69.6 . The next highest prevailing wage is for the state of california with average of $65.7 . After this, Conneticut and Alabama average a preaviling wage of $47.5 and $41.1 per hour

Conclusion

Using this report, I have attempted to solve some of intriguing questions with the help of powerful and insightful visualizations. It can be seen that US provides very good opportunity to many good employers to expand their work force by bringing in employees from all over the world using the H1B visa status. Even though states like California and Arizona have higher prevailing wage rates, states like Texas, New York and New Jersey also provide great opportunity for workers of different occupation, as shown by the number of petitions on the US Map visualization. I was also able to gain insights from the boxplots of the various occupation with varying range of prevailing wage, Wage rate to pay From and Wage rate to pay To. For the Scientists and Marketing employees even though varied greatly, along with more variance, there are more number of petitions for Computer related occupation, Analysts and Engineers. I have also performed seasonal analysis for the Certified status cases and found that most cases are sibmitted by the employer in the beginning of the year, with an overwhelming increase in the month of March followed by February and January.

Appendix A: Data Dictionary for H1B Dataset

The dataset consists of the following columns:

1. CASE_STATUS: status associated with the last significant event of decision. Valid values include “Certified”, “Certified-Withdrawn”,“Denied” and “Withdrawn”

2. CASE_SUMBITTED_DAY: The day the application was submitted
3. CASE_SUMBITTED_MONTH: The month the application was submitted
4. CASE_SUMBITTED_YEAR: The year the application was submitted
5. DECISION_DAY: The day on which the last significant event of decision was recorded
6. DECISION_MONTH: The month on which the last significant event of decision was recorded
7. DECISION_YEAR: The year on which the last significant event of decision was recorded
8. VISA_CLASS: Indicates the type of temporary application submitted for processing
9. EMPLOYER_NAME: Name of the employer submitting labor condition application
10. EMPLOYER_STATE: State of the employer submitting labor condition application
11. EMPLOYER_COUNTRY: Country of the employer submitting labor condition application
12. SOC_NAME: Occupation name associated with the job being requested for temporary labor condition
13. NAICS_CODE: Industry code associated with the employer requesting permanent labor condition
14. TOTAL_WORKERS: Total number of foreign workers requested by the employer
15. FULL_TIME_POSITION: Y=Full time position, N= Part time position
16. PREVAILING_WAGE: Prevailing wage for the job being requested for temporary labor
17. PW_UNIT_OF_PAY: Unit of pay. Valid values are “Year”, “Month”, “Week”, “Bi-Weekly” and “Hour”
18. PW_SOURCE: Variables include: “OES”, “CBA”, “DBA”, “SCA” or “Other”
19. PW_SOURCE_YEAR: Year the prevailing wage source was issued
20. PW_SOURCE_OTHER: If “Other” wage source, provide source of wage
21. WAGE_RATE_OF_PAY_FROM: Employer’s proposed wage rate
22. WAGE_RATE_OF_PAY_TO: Maximum wage rate
23. WAGE_UNIT_OF_PAY: Unit of pay. Valid values are “Year”, “Month”, “Week”, “Bi-Weekly” and “Hour”
24. H-1b_DEPENDENT: Y= Employer is H-1B dependent, N= Employer is not H-1B dependent
25. WILLFUL_VIOLATOR: Y= Employer has been previously found to be willful violator, N= Employer has not been a willful violator
26. WORKSITE_STATE: State information of the foreign worker’s intended area of employment
27. WORKSITE_POSTAL_CODE: Zip code information of the foreign worker’s intended area of employment