

Home Work-6(Balaji Kothandaraman Kalanidhi)

kk

3/25/2020

```
library(tidyverse)
```

```
## -- Attaching packages -----  
----- tidyvers  
e 1.3.0 --
```

```
## v ggplot2 3.3.0      v purrr  0.3.3  
## v tibble  2.1.3      v dplyr  0.8.5  
## v tidyr   1.0.2      v stringr 1.4.0  
## v readr   1.3.1      v forcats 0.5.0
```

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

```
library(ggplot2)  
library(plotly)
```

```
##  
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':  
##  
## last_plot
```

```
## The following object is masked from 'package:stats':  
##  
## filter
```

```
## The following object is masked from 'package:graphics':  
##  
## layout
```

```
library(usmap)  
library(maps)
```

```
##
## Attaching package: 'maps'
```

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

```
h1b<- read_csv('C:\\Users\\kkbal\\OneDrive\\Desktop\\Neu\\computation and visualization\\Assignments\\HW-6\\h1b_datahubexport-2019.csv')
```

```
## Parsed with column specification:
## cols(
##   `Fiscal Year` = col_double(),
##   Employer = col_character(),
##   `Initial Approvals` = col_double(),
##   `Initial Denials` = col_double(),
##   `Continuing Approvals` = col_double(),
##   `Continuing Denials` = col_double(),
##   NAICS = col_double(),
##   `Tax ID` = col_character(),
##   State = col_character(),
##   City = col_character(),
##   ZIP = col_character()
## )
```

```
## Warning: 38 parsing failures.
## row          col          expected actual
file
## 2650 Continuing Approvals no trailing characters    ,293 'C:\Users\kkbal\OneDrive\Desktop\Neu
\computation and visualization\Assignments\HW-6\h1b_datahubexport-2019.csv'
## 2650 Continuing Denials   no trailing characters    ,304 'C:\Users\kkbal\OneDrive\Desktop\Neu
\computation and visualization\Assignments\HW-6\h1b_datahubexport-2019.csv'
## 3691 Initial Approvals    no trailing characters    ,164 'C:\Users\kkbal\OneDrive\Desktop\Neu
\computation and visualization\Assignments\HW-6\h1b_datahubexport-2019.csv'
## 4022 Continuing Approvals no trailing characters    ,254 'C:\Users\kkbal\OneDrive\Desktop\Neu
\computation and visualization\Assignments\HW-6\h1b_datahubexport-2019.csv'
## 5060 Initial Approvals    no trailing characters    ,580 'C:\Users\kkbal\OneDrive\Desktop\Neu
\computation and visualization\Assignments\HW-6\h1b_datahubexport-2019.csv'
## ....
.....

## See problems(...) for more details.
```

```
zip<-read.csv('C:\\Users\\kkbal\\OneDrive\\Desktop\\Neu\\computation and visualization\\Assignments\\HW-6\\us-zip-code-latitude-and-longitude.csv',sep=';')
```

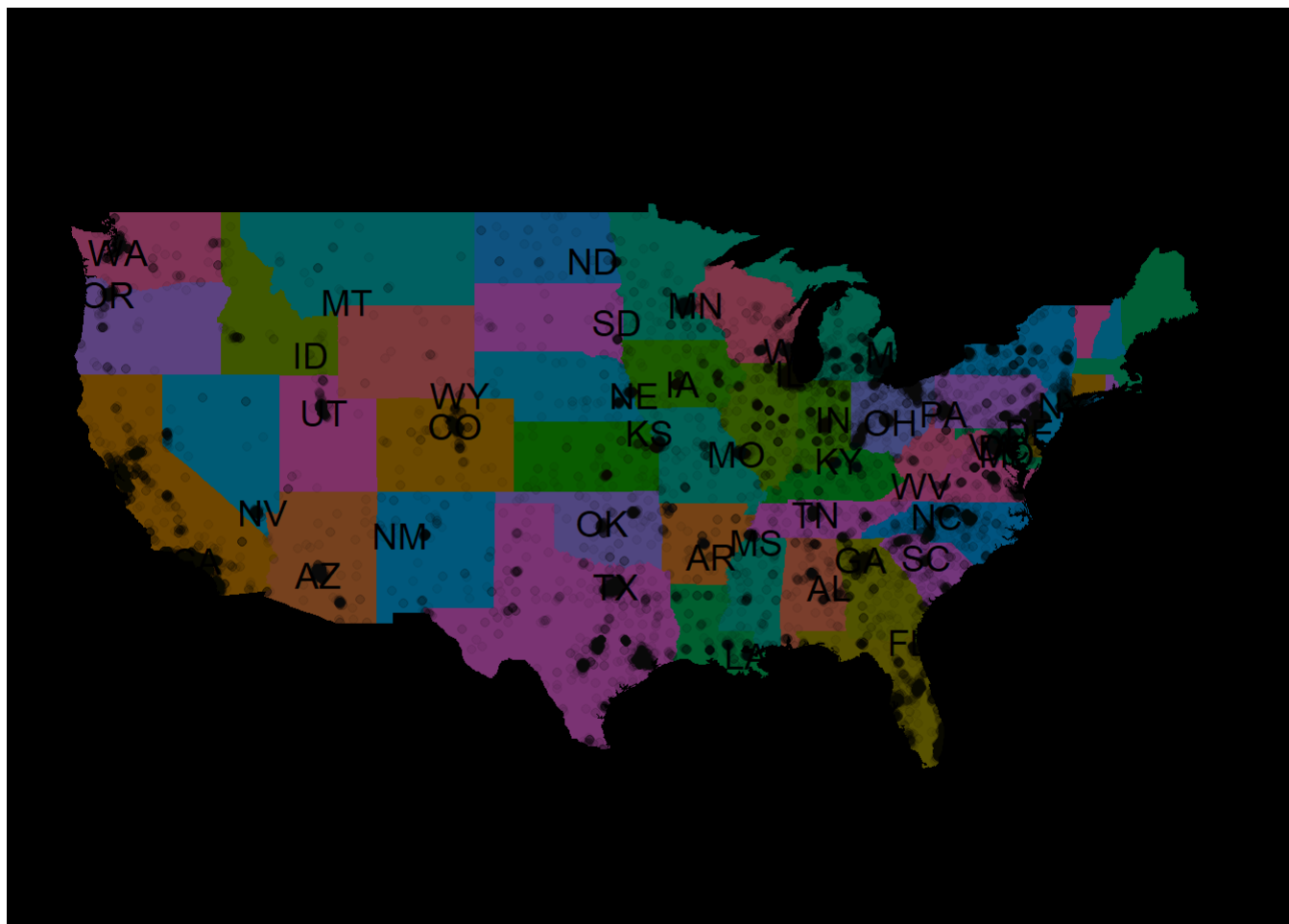
Data Wrangling

```
h1b<-h1b%>%select(-c(`Tax ID`))
h1b2019<-h1b%>%merge(zip,by.x=c('ZIP'),by.y=c('Zip'))%>%select(-c('State.y','City.y'))%>%rename
(State='State.x',City='City.x')
head(h1b2019)
```

```
##      ZIP Fiscal Year      Employer Initial Approvals
## 1 10001      2019  AFFINITY SOLUTIONS INC              0
## 2 10001      2019  CATCHPOINT SYSTEMS INC              0
## 3 10001      2019      RNK JEWELRY INC                 0
## 4 10001      2019  HBO SERVICE CORPORATION            0
## 5 10001      2019    BEACON PRODUCTS INC              1
## 6 10001      2019  ARCHSTONE BUILDERS LLC             5
##      Initial Denials Continuing Approvals Continuing Denials NAICS State
## 1              0              1              0      99      NY
## 2              0              1              0      99      NY
## 3              0              1              0      33      NY
## 4              0             12              0      51      NY
## 5              0              0              0      99      NY
## 6              2              3              0      23      NY
##      City Latitude Longitude Timezone Daylight.savings.time.flag
## 1 NEW YORK 40.75074 -73.99653      -5              1
## 2 NEW YORK 40.75074 -73.99653      -5              1
## 3 NEW YORK 40.75074 -73.99653      -5              1
## 4 NEW YORK 40.75074 -73.99653      -5              1
## 5 NEW YORK 40.75074 -73.99653      -5              1
## 6 NEW YORK 40.75074 -73.99653      -5              1
##      geoint
## 1 40.750742,-73.99653
## 2 40.750742,-73.99653
## 3 40.750742,-73.99653
## 4 40.750742,-73.99653
## 5 40.750742,-73.99653
## 6 40.750742,-73.99653
```

USA map

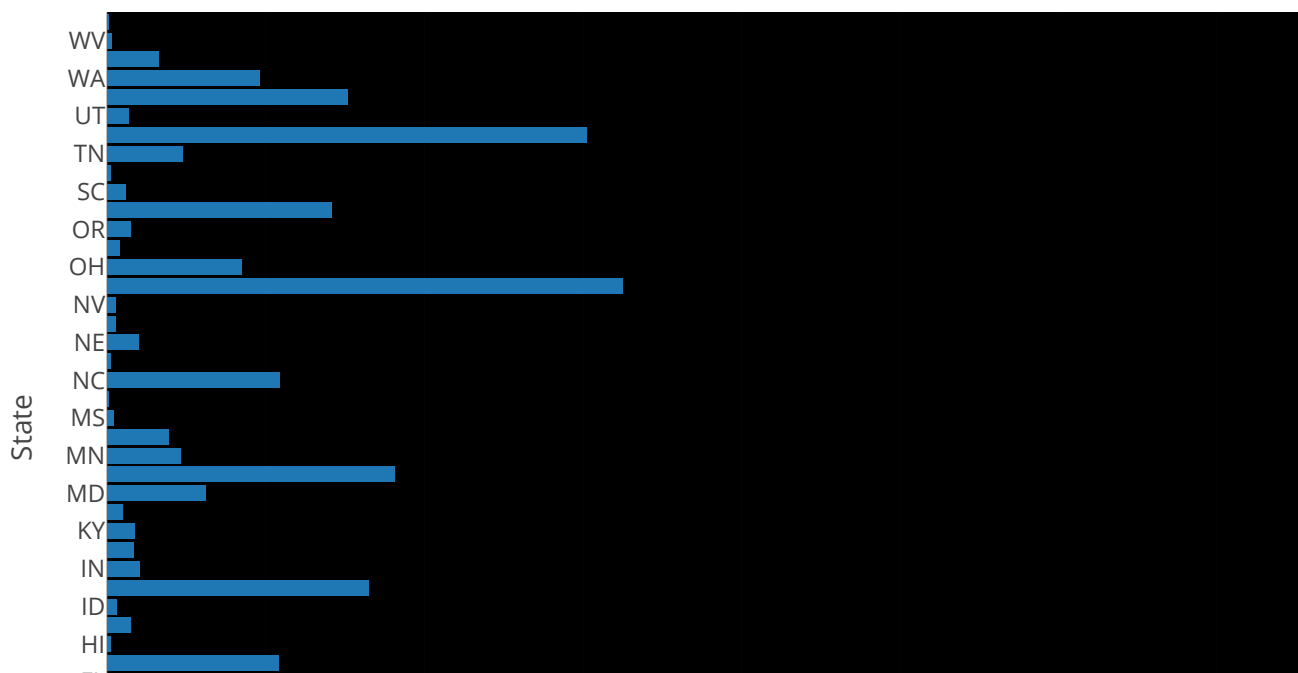
```
USA<-map_data('state')
usa_map<-ggplot()+geom_polygon(data=USA,aes(x=long,y=lat,group=`group`,fill=region),alpha=0.5)+t
heme_void()+theme(legend.position='none')+coord_cartesian(xlim=c(-125,-65),ylim=c(20,56))
us_map<-usa_map+geom_point(data=h1b2019,aes(y=Latitude,x=Longitude,fill=State),alpha=0.05)
labels<-distinct_if(h1b2019[,c('State','Latitude','Longitude')],is.character,.keep_all = TRUE)
us_map+geom_text(data=labels,aes(x=Longitude,y=Latitude,label=State),size=5,color='black')+ them
e(panel.background = element_rect(fill = 'black', colour = 'black'))
```

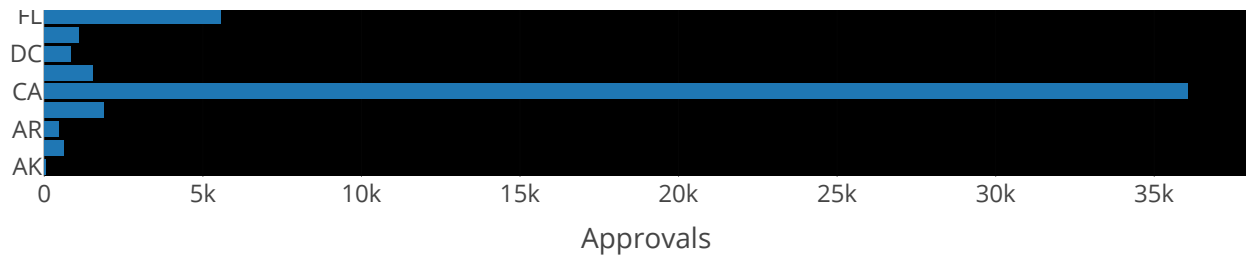


Most H1B Approvals State Wise

```
h1bstate<-aggregate(~State,h1b2019[,c('State','Continuing Approvals')],sum)
fig<-plot_ly(h1bstate,y=~State,x=~`Continuing Approvals`,type='bar',orientation='h')
fig<-fig %>% layout(xaxis=list(title='Approvals'),yaxis=list(title='State'))
fig%>% layout(title='H1B Approvals of 2019')%>%layout(plot_bgcolor='black')
```

H1B Approvals of 2019





Top H1B Sponsors in California

```
ca<-h1b2019%>%select(Employer,`Continuing Approvals`,State)%>%group_by(Employer)%>%filter(State=
='CA')
ca<-aggregate(.~Employer,ca[,c('Continuing Approvals','Employer')],sum)
ca<-ca[order(ca$`Continuing Approvals`,decreasing = T),]%>%head(5)
ca<-ca%>%select(-Employer)
Employer<-c('QUALCOMM TECHNOLOGIES','UBER','Linkedin','VMWARE','PAYPAL')
ca<-cbind(ca,Employer)
ca<-ggplot(ca,aes(x=Employer,y=`Continuing Approvals`,fill=`Employer`))+geom_bar(stat='identity'
,width=1,color='white')+coord_polar('x')+ylab('Approvals')+ggtitle('Top H1B Approvals in Califor
nia')+theme(plot.title = element_text(hjust=0.5))+ylab('')+geom_text(aes(y=`Continuing Approvals
`,label=`Continuing Approvals`),position =position_dodge(width=1),hjust=-0.2)+ theme(panel.backg
round = element_rect(fill = 'grey', colour = 'black'))+theme_void()
```

ca

Top H1B Approvals in California

