Assignment 3

First Chart: H1-B Petition Approval Breakdown

```
# Chart 1: (Area Chart) H1-B Petition Approval Time Series


library(tidyverse)

## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr

## Conflicts with tidy packages ----------------------------------------------

## filter(): dplyr, stats
## lag():    dplyr, stats
library(ggthemes)
library(reshape2)

##
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':
##
##     smiths
## DATA
h1b_gen <- read_csv('h1b_gen.csv')

## Parsed with column specification:
## cols(
##   Year = col_integer(),
##   Receipts = col_number(),
##   Approvals = col_number(),
##   `New Approval` = col_number()
## )
h1b_gen <- plyr::rename(h1b_gen, c('Receipts' = 'applications', 'Approvals' = 'all_approvals', 'New App:

## PLOT
h1b_gen_approved <- ggplot(h1b_gen, aes(x=Year)) +
  geom_area(aes(y=applications, fill="applications")) +
  geom_area(aes(y=all_approvals, fill="all_approvals"), alpha = 0.7) +
  geom_area(aes(y=new_approvals, fill="new_approvals")) +
  labs(title="H1-B Petition 2007-2016",
       subtitle="Breakdown of Approved Petition",
       caption="Source: USCIS & US Department of State",
       y="# of Petition") +
  theme_few() +
  scale_y_continuous(labels = scales::comma) +
```

```
    theme(plot.title = element_text(hjust = 0.5, face = "bold", size = 15),
          plot.subtitle = element_text(hjust = 0.5, face = "italic", size = 10),
          plot.caption = element_text(hjust = 0.5, face = "italic", size = 9)) +
    scale_fill_manual(name= NULL,
          values=c("#3c0c59", "#b32062", "#83e2c3"),
          breaks=c("applications", "all_approvals", "new_approvals"),
          labels=c("Total","All Approved\n(Continuing + New)",  "New Approved"))

h1b_gen_approved
```


H1−B Petition 2007−2016
Breakdown of Approved Petition

Source: USCIS & US Department of State
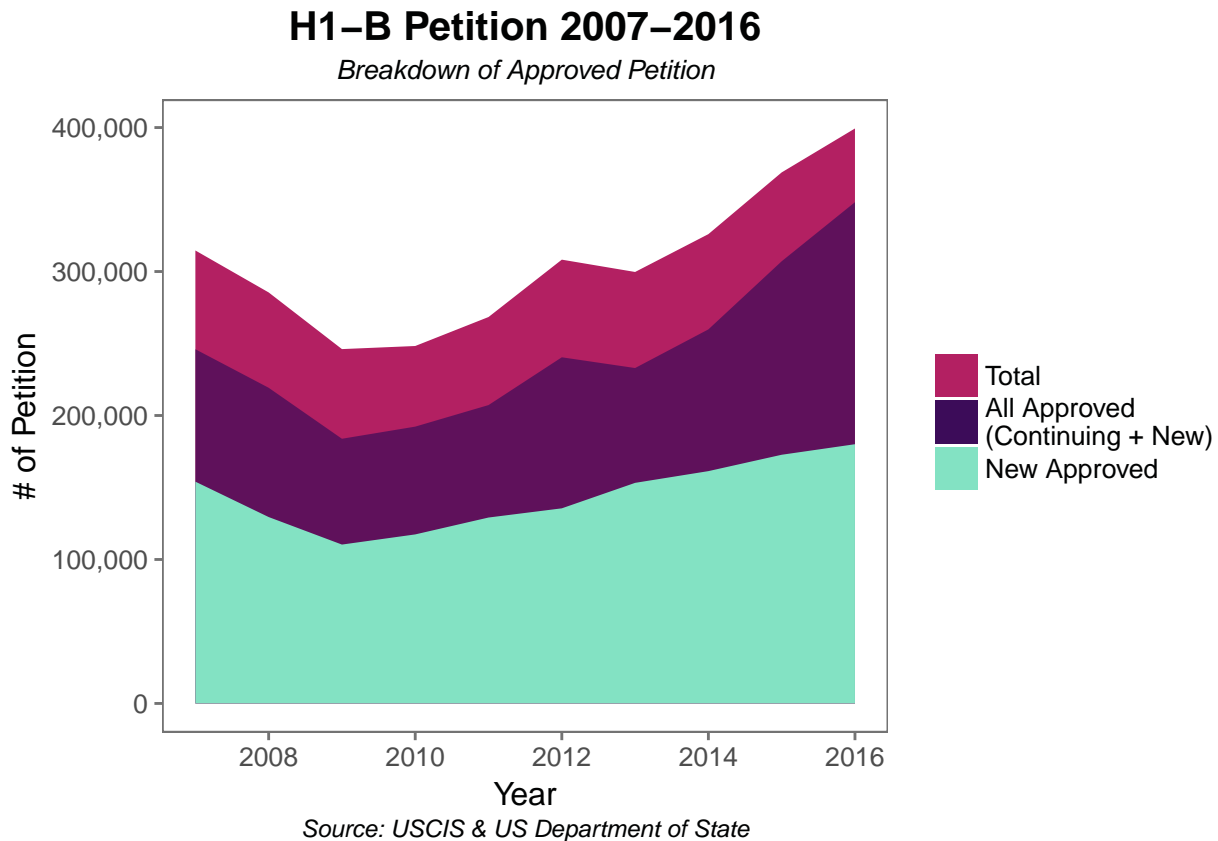
The graph is showing H1-B petition approval trends in the 10-year period from 2007 to 2016. 'Total' signifies the number of petition request that were filed to the USCIS. The 'All Approved' area signify a portion of petitions that are approved, whether for continuing workers with prior H1-B or new workers who applied for an initial, new H1-B visa. As the name suggest, the 'New Approved' area signifies only the portion of the petitions that are approved for entirely new H1-B for new workers. From this graph we can see that steep increase in petition filed are mostly due to existing H1-B workers filing for a reissue, while new H1-B have been increasing in a much slower rate.

Second Chart: H1-B Petition Beneficiary's Education Level

```
# Chart 2: (LINE CHART) H1−B Petition Beneficiary's Education Level

## DATA
h1b_edu <- read_csv('h1b_edu.csv')

## Parsed with column specification:
```

```
## cols(
##   Year = col_integer(),
##   `No Diploma` = col_integer(),
##   `High School Graduate` = col_integer(),
##   `Some College Credit (Less than 1 year)` = col_integer(),
##   `One or More Years of College (No Degree)` = col_integer(),
##   `Associates Degree` = col_integer(),
##   `Bachelor's Degree` = col_number(),
##   `Master's Degree` = col_number(),
##   `Professional Degree` = col_number(),
##   `Doctorate Degree` = col_number()
## )
h1b_edu <- plyr::rename(h1b_edu, c('Doctorate Degree' = 'doctorate',
                                   'Professional Degree' = 'professional',
                                   "Master's Degree" = 'masters',
                                   "Bachelor's Degree" = 'bachelors',
                                   'Associates Degree' = 'associates',
                                   'One or More Years of College (No Degree)' = 'some_college2',
                                   'Some College Credit (Less than 1 year)' = 'some_college1',
                                   'High School Graduate' = 'highschool',
                                   'No Diploma' = 'no_diploma'))
h1b_edu <- mutate(h1b_edu,
                  some_college = associates +
                    no_diploma +
                    highschool +
                    some_college1 +
                    some_college2)
h1b_edu_flat = melt(h1b_edu, id=c("Year"))


## PLOT
h1b_edu_petition <- ggplot(h1b_edu_flat, aes(x=Year, y = value, color = variable)) +
  geom_line() + geom_point() +
  labs(title="H1-B Petition 2007-2016",
       subtitle="Breakdown of Petition by Beneficiary Education",
       caption="Source: USCIS",
       y="# of Petition") +
  theme_few() +
  scale_y_continuous(labels = scales::comma) +
  theme(plot.title = element_text(hjust = 0.5, face = "bold", size = 15),
        plot.subtitle = element_text(hjust = 0.5, face = "italic", size = 10),
        plot.caption = element_text(hjust = 0.5, face = "italic", size = 9)) +
  scale_color_discrete(name= "Degree Level",
                       breaks=c("doctorate","professional","masters","bachelors", "some_college"),
                       labels=c("Doctorate","Professional","Master's","Bachelor's","Less than a Bachelor'

h1b_edu_petition
```
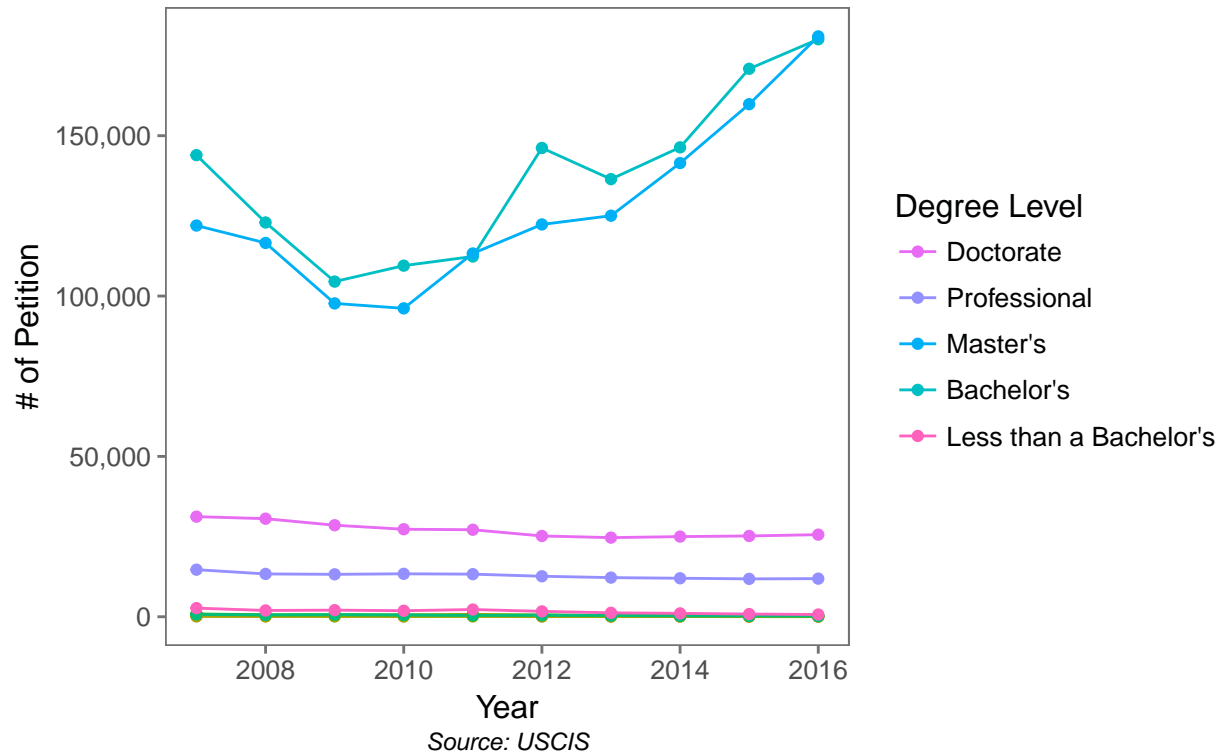
H1−B Petition 2007−2016

Breakdown of Petition by Beneficiary Education

As the title suggests, this chart aims to display the academic make up of yearly H1-B petitioners. Since H1-B is a visa aimed for admitting highly-skilled individual, it is not surprising to see people who have 'less than a bachelor degree' would be less likely to apply. At the other end of the spectrum, Master's Degree petitioners are slowly taking over their Bachelors Degree counterpart in terms of volume. This might mean that H1-B petitioners are gettting smarter, and the additional 20,000 master's quota might be just as overcrowded as the regular visa cap.

Third Chart: H1-B Petition Income Make Up

```
# Chart 3: (CLUSTERED BAR) H1-B Petition Income Make Up

## DATA
h1b_income <- read_csv('h1b_income.csv')

## Parsed with column specification:
## cols(
##   Year = col_integer(),
##   `Low: Below 50,000` = col_integer(),
##   `Middle: 50,000 - 100,000` = col_integer(),
##   `Middle-High: 100,000 - 150,000` = col_integer(),
##   `High: 150,000 +` = col_integer()
## )

h1b_income <- h1b_income %>%
  filter(Year < 2017) %>%
  melt(id=c("Year"))
```

```
## PLOT
h1b_income_bar <- ggplot(data=h1b_income, aes(factor(Year), y = value, fill = forcats::fct_rev(variable)
  geom_bar(stat = 'identity', position="dodge") +
  labs(title="H1-B Petition 2007-2016",
       subtitle="Breakdown of Petition by Income Level",
       caption="Source: USCIS",
       y="# of Petition") +
  theme_few() +
  scale_y_continuous(labels = scales::comma) +
  theme(plot.title = element_text(hjust = 0.5, face = "bold", size = 15),
        plot.subtitle = element_text(hjust = 0.5, face = "italic", size = 10),
        plot.caption = element_text(hjust = 0.5, face = "italic", size = 9)) +
  scale_fill_brewer(name= "Income Level", palette = "Spectral")

h1b_income_bar
```
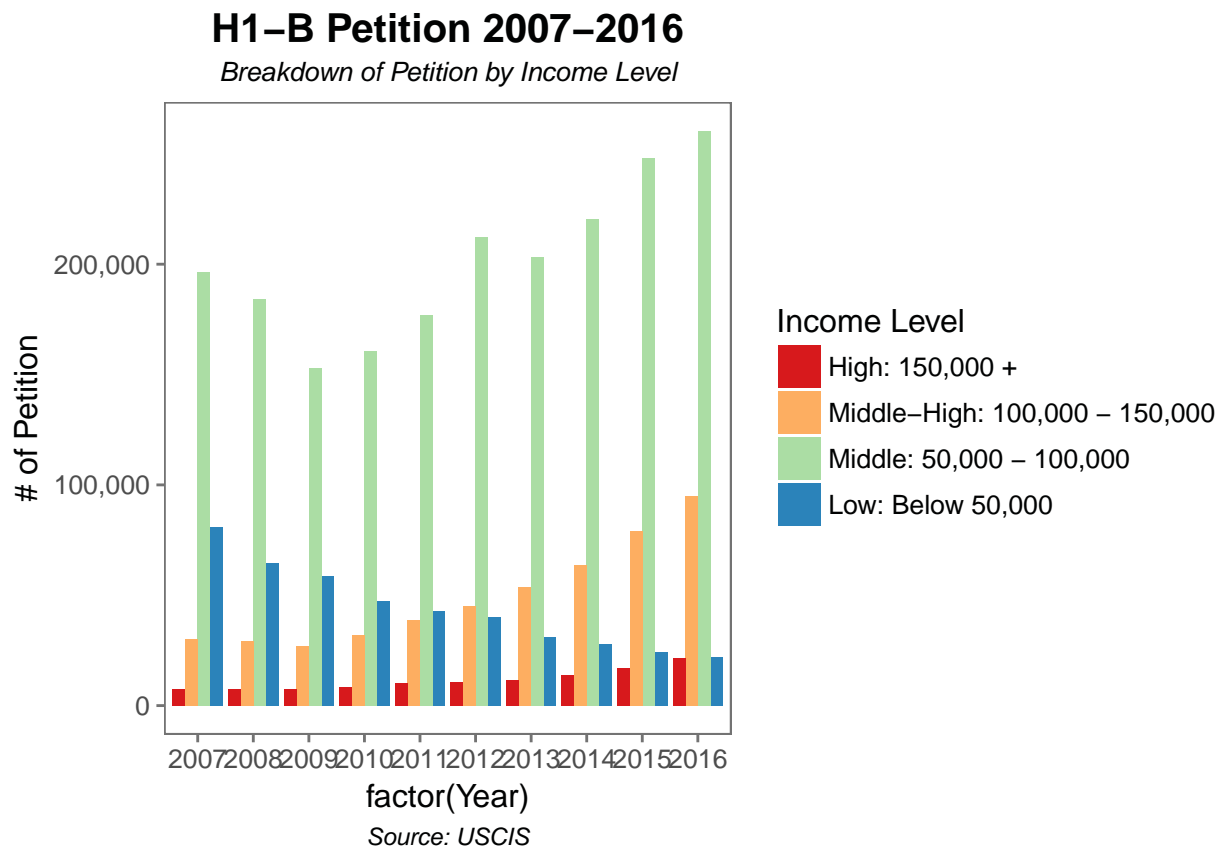

In this chart, we can explicitly see that H1-B petitioners over the years are getting paid higher than their predecessors. This trend might be caused by the prevalence of high-paying technology jobs that attracts top international talent and therefore requires H1-B certification. It might also be caused by the perpetually rising administration cost for h1-b visa that can be cost-prohibitive for interested applicants in lower income brackets. There are numerous claims by resercher that H1-B workers are generally paid less than their American counterpart. The project will try to dig deeper on that claim.

Fourth Chart: (DIVERGING BARS) Top 20 H1-B Sponsors of Approved Petition Wage Comparison

```r
# Chart 4: (BAR PANEL) Top 20 H1-B Sponsors of Approved Petitions Avg. Wage Comparison

## DATA
h1b_comp <- read_csv('h1b_company.csv')

## Parsed with column specification:
## cols(
##   `Employer Name` = col_character(),
##   `Total Number of Approved Petitions` = col_integer(),
##   `Average Salary ($)` = col_number(),
##   `Wage Difference` = col_integer()
## )

h1b_comp <- plyr::rename(h1b_comp, c('Employer Name' = 'employer',
                                     'Total Number of Approved Petitions' = 'num_approved',
                                     "Average Salary ($)" = 'avg_salary',
                                     "Wage Difference" = 'wage_diff'))
h1b_comp$wage_type <- ifelse(h1b_comp$wage_diff < 0, "below", "above")
h1b_comp <- h1b_comp %>% arrange(num_approved)
h1b_comp$employer <- factor(h1b_comp$employer, levels = h1b_comp$employer)

## PLOT
h1b_wage_diff <- ggplot(h1b_comp, aes(x=employer, y=wage_diff, label=wage_diff)) +
  geom_bar(stat='identity', aes(fill=wage_type), width=.5)  +
  scale_fill_manual(name= "Avg. Wage Offered",
                    labels = c("Above Average", "Below Average"),
                    values = c("above"="#00ba38", "below"="#f8766d")) +
  labs(title="H1-B Wage 2016",
       subtitle="Top 20 H1B Sponsor Companies Avg. Wage Offer Compared to the Average Income of H1
       caption="Source: USCIS",
       y="$ Difference to the Avg. Income",
       x= "Top 20 H1-B Sponsorship Company\n[Top-Bottom] Highest to Lowest # of Workers Sponsored")
  theme(plot.title = element_text(hjust = 0.5, face = "bold", size = 15),
        plot.subtitle = element_text(hjust = 0.5, face = "italic", size = 10),
        plot.caption = element_text(hjust = 0.5, face = "italic", size = 9)) +
  coord_flip()

h1b_wage_diff
```
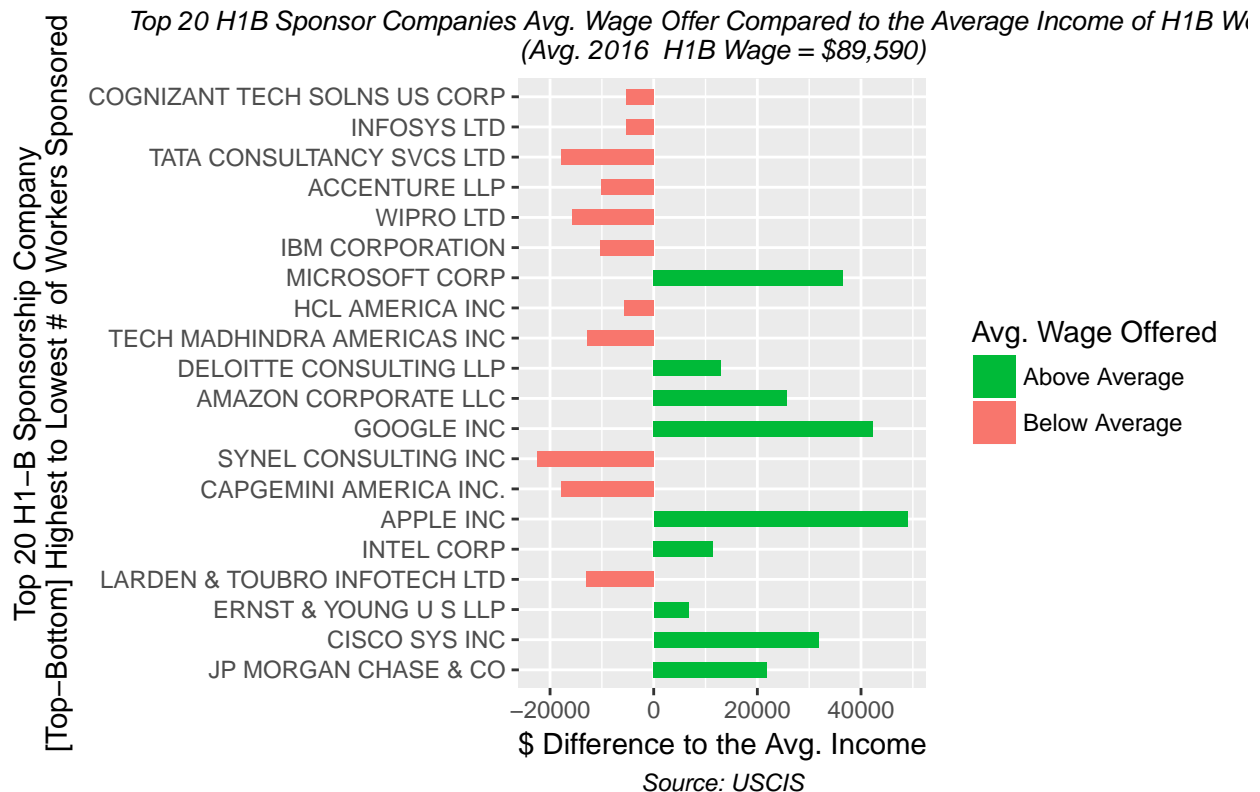
H1-B Wage 2016

Top 20 H1B Sponsor Companies Avg. Wage Offer Compared to the Average Income of H1B Wo
(Avg. 2016 H1B Wage = $89,590)

Source: USCIS

The top 5 h1-b sponsors are actually paying their workers less than the average wage of h1-b workers in 2016. This might be an interesting starting point in finding proof to the h1-b underpayment problem. Further in-depth investigation towards industry and job specific income standard are in order.

Fifth Chart: (SLOPE Chart) 5-Year Change in H1-B Approved Petition

```
# Chart 5: (SLOPE Chart) Progression of Approved H1-B per Continent in the last 5 Years (2011-2016)

## DATA
h1b_continent <- read_csv('h1b_continent.csv')

## Parsed with column specification:
## cols(
##   continent = col_character(),
##   Year2016 = col_number(),
##   Year2011 = col_number()
## )
## PLOT
lft_label <- paste(h1b_continent$continent, round(h1b_continent$`Year2011`),sep=", ")
rgt_label <- paste(h1b_continent$continent, round(h1b_continent$`Year2016`),sep=", ")
h1b_continent$class <- ifelse((h1b_continent$`Year2016` - h1b_continent$`Year2011`) < 0, "red", "green")

h1b_continent_line <- ggplot(h1b_continent) + geom_segment(aes(x=1, xend=2, y=Year2011, yend=Year2016, 
  geom_vline(xintercept=1, linetype="dashed", size=.1) +
  geom_vline(xintercept=2, linetype="dashed", size=.1) +
```

```r
  scale_color_manual(labels = c("Up", "Down"),
                     values = c("green"="#00ba38", "red"="#f8766d")) +
  labs(x="", y="# of H1-B Petitions Approved") +
  xlim(.5, 2.5) + ylim(0,(0.8*(max(h1b_continent$Year2011, h1b_continent$Year2016)))) +
  geom_text(label=lft_label, y=h1b_continent$Year2011, x=rep(1, NROW(h1b_continent)), hjust=1.1, size=3
  geom_text(label=rgt_label, y=h1b_continent$Year2016, x=rep(2, NROW(h1b_continent)), hjust=-0.1, size=
  geom_text(label="Time 1", x=1, y=0.8*(max(h1b_continent$Year2011, h1b_continent$Year2016)), hjust=1.2
  geom_text(label="Time 2", x=2, y=0.8*(max(h1b_continent$Year2011, h1b_continent$Year2016)), hjust=-0.

h1b_continent_line +
  labs(title="H1-B Approved Petition 2011 - 2016",
       subtitle="Progression of Approved H1-B Visa in the Last 5 Years",
       caption="Source: US Department of State") +
  theme_few() +
  scale_y_continuous(labels = scales::comma) +
  theme(plot.title = element_text(hjust = 0.5, face = "bold", size = 15),
        plot.subtitle = element_text(hjust = 0.5, face = "italic", size = 10),
        plot.caption = element_text(hjust = 0.5, face = "italic", size = 9)) +
  theme(panel.background = element_blank(),
        panel.grid = element_blank(),
        axis.ticks = element_blank(),
        axis.text.x = element_blank(),
        plot.margin = unit(c(1,2,1,2), "cm"))
```

```
## Scale for 'y' is already present. Adding another scale for 'y', which
## will replace the existing scale.

## Warning: Removed 1 rows containing missing values (geom_segment).

## Warning: Removed 1 rows containing missing values (geom_text).


## Warning: Removed 1 rows containing missing values (geom_text).

## Warning: Removed 7 rows containing missing values (geom_text).


## Warning: Removed 7 rows containing missing values (geom_text).
```
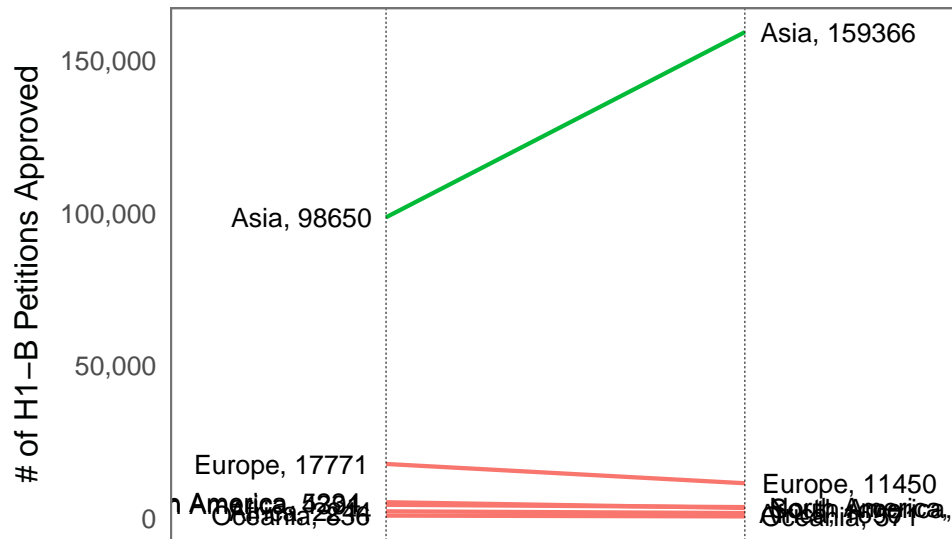
H1–B Approved Petition 2011 – 2016

Progression of Approved H1–B Visa in the Last 5 Years

Asia, 159366

Asia, 98650

150,000

# of H1–B Petitions Approved

100,000

50,000

Europe, 17771

Europe, 11450

America, 5294

North America,

0

Oceania, 836

Oceania, 971

In this chart, international workers from Asia not only dominates the general make up of approved h1-b visa, but it also the only one that is showm to have an upward trend when it comes to the work visa application. This might have been cause by the historically high traffic of Indian worker in the h1-b system. Further investigation must be made in order to ensure fairness between applicants.