

Assignment 3 & 4

UCID : 12149052

Name : Regina Widjaya

Assignment 3: Markdown

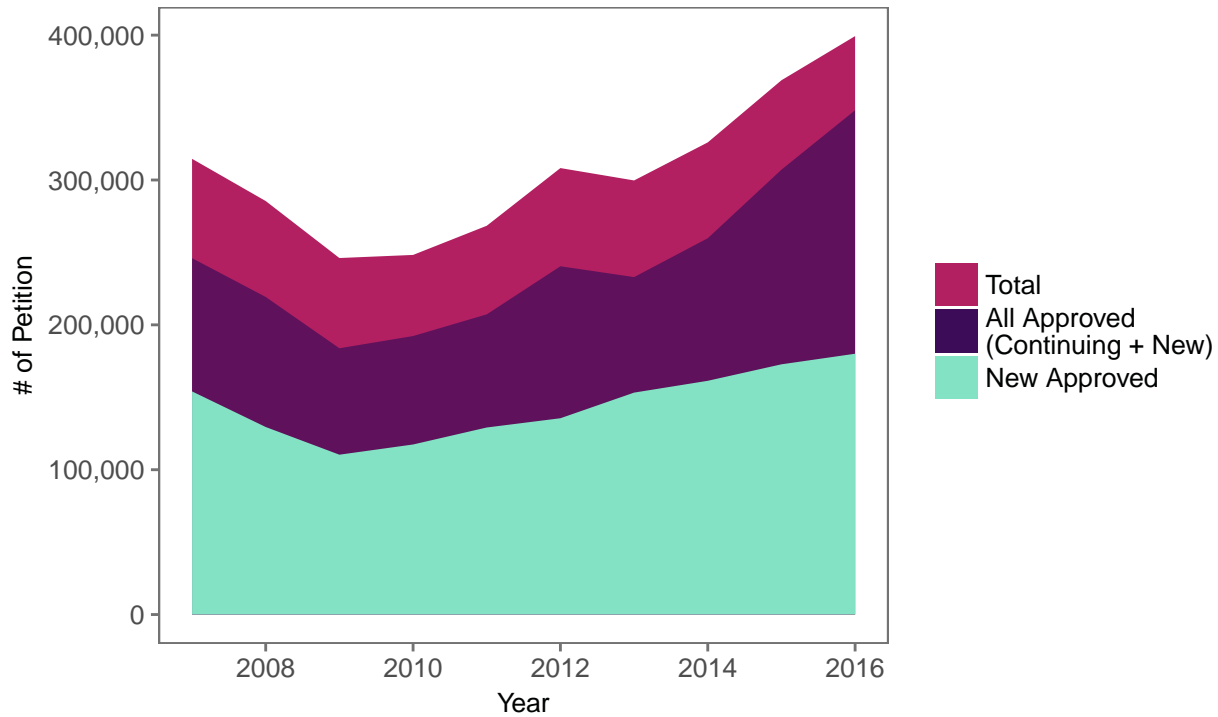
First Chart: H1-B Petition Approval Breakdown

```
## 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 Petitions",
       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),
        axis.title.y = element_text(size = 10),
        axis.title.x = element_text(size = 10)) +
  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 Petitions



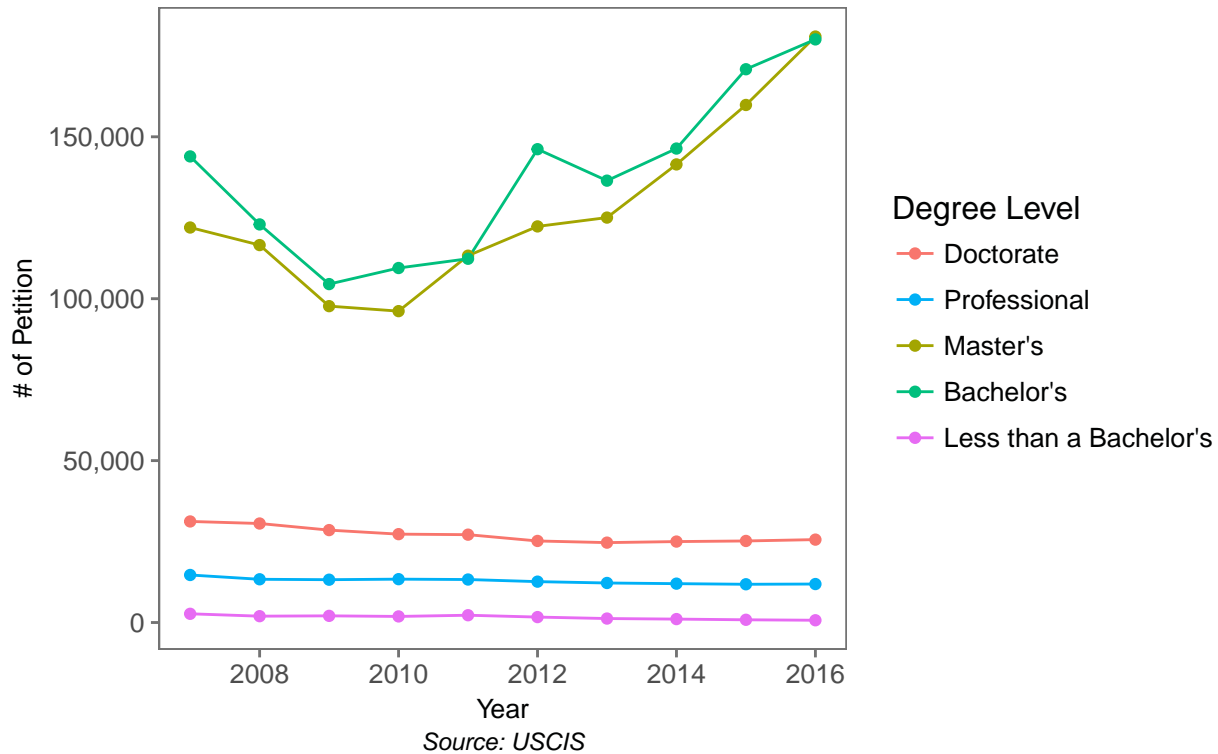
The graph is showing H1-B petition approval trends in the 10-year period from 2007 to 2016. ‘Total’ signifies the number of petition request that were filed to the USCIS. The ‘All Approved’ area signify a portion of petitions that are approved, whether for continuing workers with prior H1-B or new workers 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

```
## 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 Petitions 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),
        axis.title.y = element_text(size = 10),
        axis.title.x = element_text(size = 10)) +
  scale_color_discrete(name= "Degree Level",
                       breaks=c("doctorate","professional","masters","bachelors", "some_college"), labels=
```

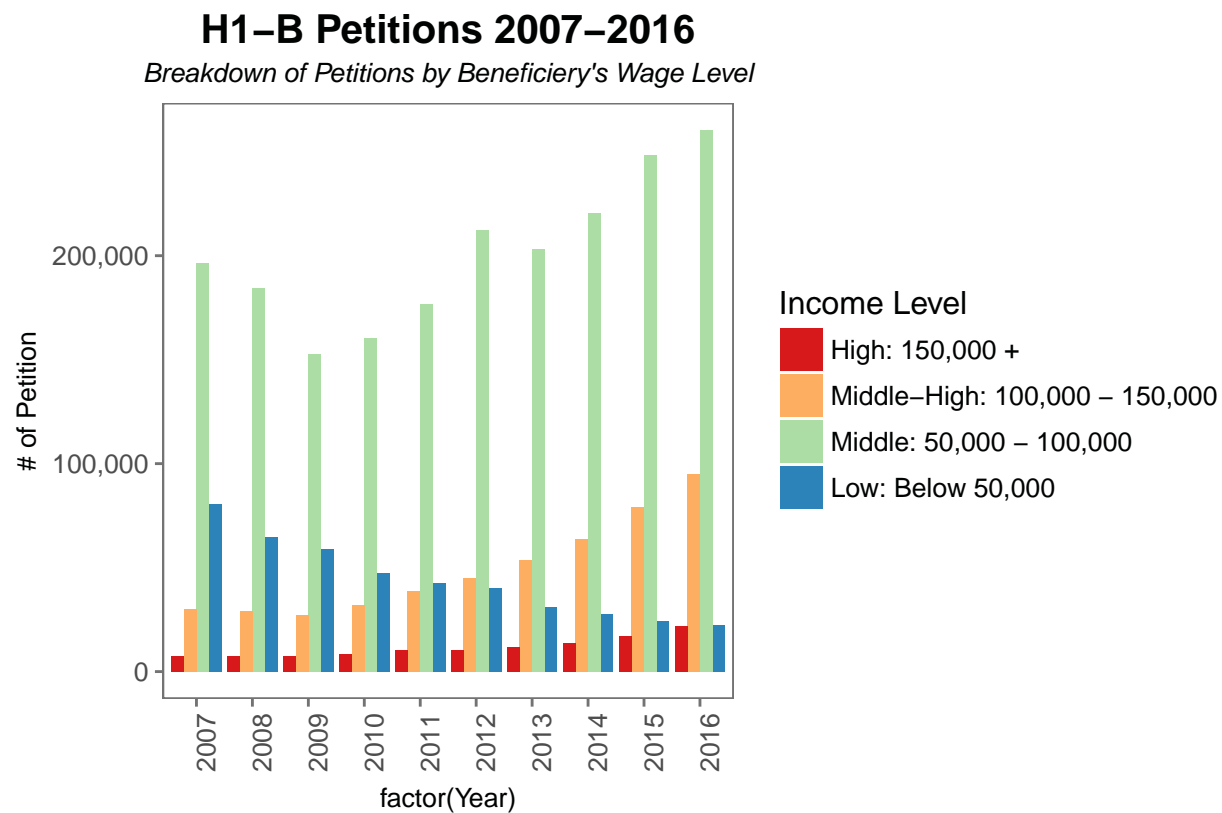
H1-B Petition 2007–2016

Breakdown of Petitions 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 getting 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



Source: USCIS

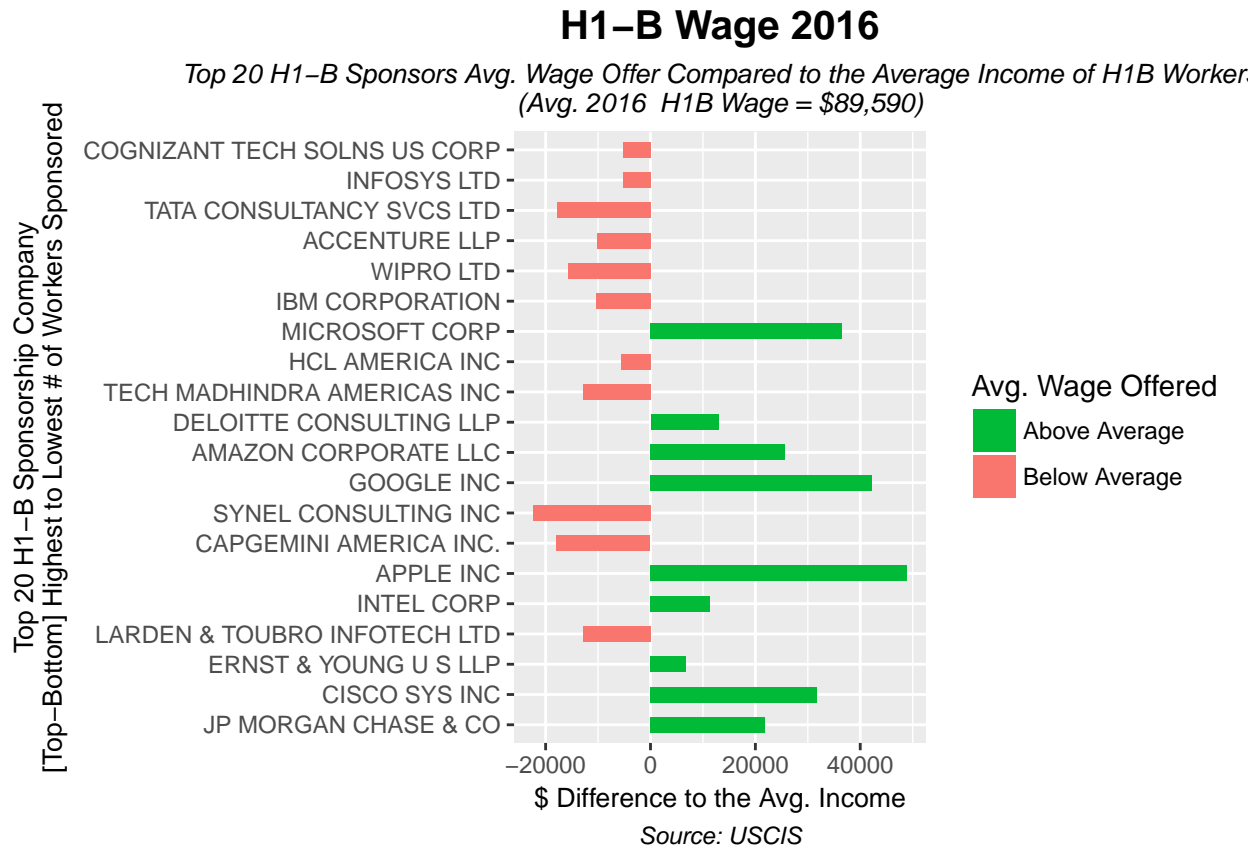
In this chart, we can explicitly see that H1-B petitioners over the years are getting paid higher than their predecessors. This trend might be caused by the prevalence of high-paying technology jobs that attracts top international talent and therefore requires H1-B certification. It might also be caused 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

```
## 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 H1-B Sponsors Avg. Wage Offer Compared to the Average Income of H1B Workers",
       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),
```

```
axis.title.y = element_text(size = 10),
axis.title.x = element_text(size = 10)) +
coord_flip()

plot(h1b_wage_diff)
```



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

```
## 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) +
  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)
```

```

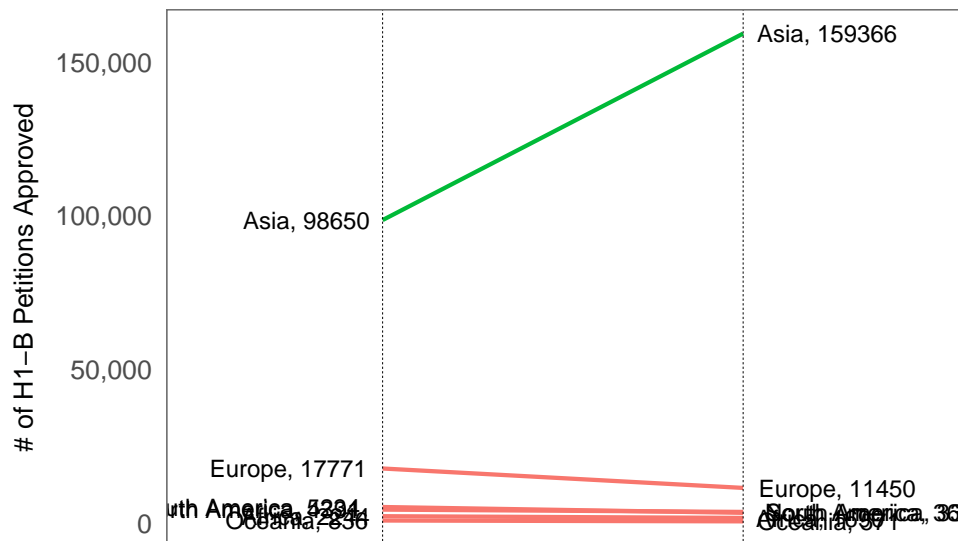
geom_text(label=rgt_label, y=h1b_continent$Year2016, x=rep(2, NROW(h1b_continent)), hjust=-0.1, size=12)
geom_text(label="Time 1", x=1, y=0.8*(max(h1b_continent$Year2011, h1b_continent$Year2016)), hjust=1.2)
geom_text(label="Time 2", x=2, y=0.8*(max(h1b_continent$Year2011, h1b_continent$Year2016)), hjust=-0.1)

h1b_continent_line +
  labs(title="H1-B Approved Petition 2011 - 2016",
        subtitle="Change in the # 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),
        axis.title.y = element_text(size = 10),
        axis.title.x = element_text(size = 10),
        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"))

```

H1-B Approved Petition 2011 – 2016

Change in the # of Approved H1-B Visa in the Last 5 Years



Source: US Department of State

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 shown to have an upward trend when it comes to the work visa application. This might have been caused 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.

Assignment 4: Theming & Design of Static Visualization

```

## Barbel Chart
bb <- ggplot(h1b_state, aes(x=h1b_state$MED_WAGE, xend=median_wage, y=state, group = state)) +

```

```

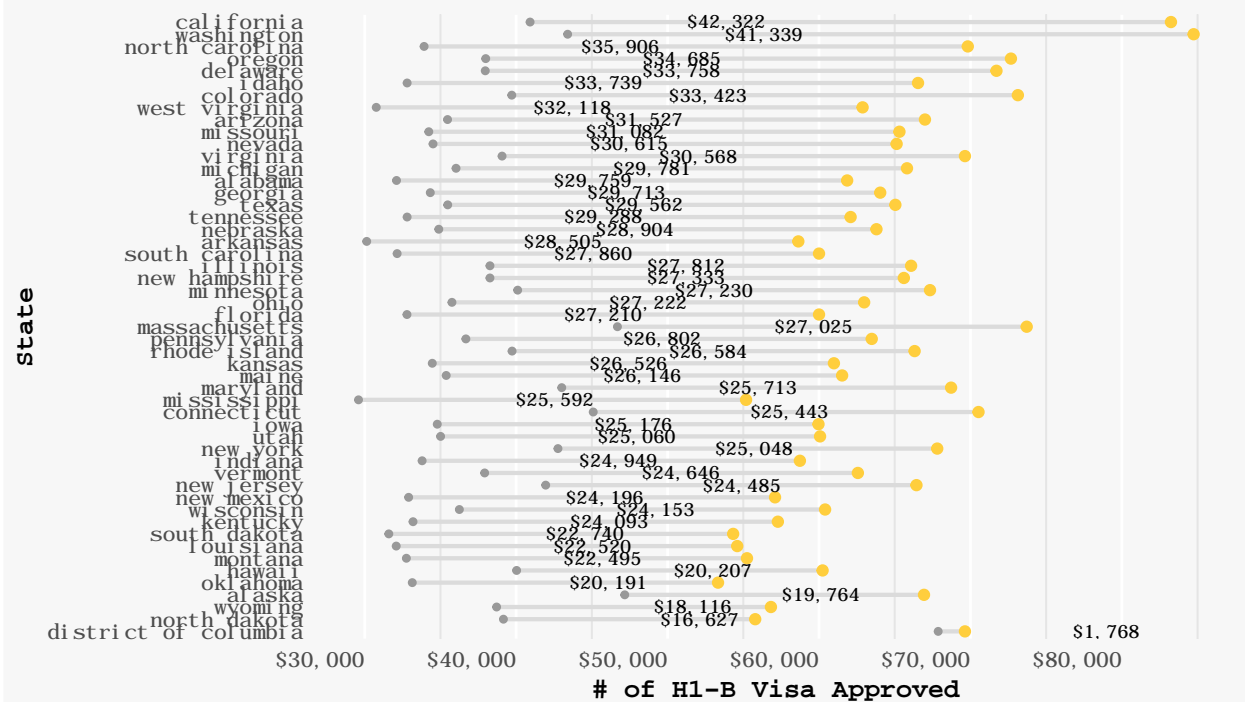
geom_dumbbell(colour = "#dbdbdb",
              size=0.75,
              colour_x = "#999999",
              colour_xend = "#ffce3d",
              size_xend = 1.5)+
geom_text(aes(label = scales::dollar(round(wage_diff)), family = "Andale Mono"),
          hjust = -2, size = 2.5)+
scale_x_continuous(label= dollar) +
labs(x="# of H1-B Visa Approved",
     y="State",
     title="How much more are H1-B workers get paid in every state?",
     subtitle="The H1-B workers generally receive much better payment than state median.\nTechnology",
     caption="Source: U.S. Department of Labor Statistics & Department of State 2016") +
theme(plot.title = element_text(family = 'Courier New', hjust = 0, face = "bold", size = 15),
      plot.subtitle = element_text(family = 'Courier New', face = "italic", size = 10),
      plot.caption = element_text(family = 'Courier New', face = "italic", size = 9),
      plot.background=element_rect(fill="#f7f7f7"),
      panel.background=element_rect(fill="#f7f7f7"),
      panel.grid.minor=element_blank(),
      panel.grid.major.y=element_blank(),
      panel.grid.major.x=element_line(),
      panel.grid.minor.x=element_line(color="gray90"),
      panel.border=element_blank(),
      axis.ticks=element_blank(),
      axis.title.x = element_text(family = 'Courier New', face = 'bold', size = 10),
      axis.text.x = element_text(angle = 0, hjust = 1, family = 'Andale Mono', size = 8),
      axis.title.y = element_text(family = 'Courier New', face = 'bold', size = 10),
      axis.text.y = element_text(angle = 0, hjust = 1, family = 'Andale Mono', size = 8))

plot(bb)

```

How much more are H1-B workers get paid

The H1-B workers generally receive much better payment than Technology 'hot-spot's likely to display larger gap.



Source: U.S. Department of Labor Statistics & Department of State 2016

In this chart, we can clearly see that H1-B workers are paid much higher salary compared to the general median wage in every state. Technology 'hot-spot' states such as California and Washington displayed widest gap between the two subset of median wage. This is unsurprising because H1-B caters disproportionately to technology graduates that are more likely to be paid higher than other occupations.

```
ff <- ggplot(h1b_facet_long, aes(y=value, x= state)) +
  facet_wrap(~variable, scales="free") +
  geom_bar(stat="identity", aes(fill = (state %in% tech_states))) +
  geom_text(aes(label = scales::comma(round(value,1)), family = "Andale Mono"),
    hjust = 0, size = 2.5,
    position = position_dodge(width = 1)) +
  scale_fill_manual(values = c('#999999', '#ffce3d'))+
  facet_grid(~variable, scales = "free_x") +
  coord_flip() +
  theme_few() +
  labs(title="Which state hosts the most H1-B workers?",
    subtitle="California hosts the majority of H1-B workers, \nbut New Jersey is more likely to empl",
    caption="Source: U.S. Department of Labor Statistics & Department of State 2016",
    y = "State")+
  guides(fill=FALSE) +
  theme(axis.title.x=element_blank(),
    axis.text.x = element_text(angle = 0, hjust = 1, family = 'Andale Mono', size = 8),
    axis.title.y = element_text(family = 'Courier New', face = 'bold', size = 10),
    axis.text.y = element_text(angle = 0, hjust = 1, family = 'Andale Mono', size = 8),
    panel.background=element_rect(fill="#f7f7f7"),
    plot.background=element_rect(fill="#f7f7f7"),
    strip.background = element_rect(fill="#f7f7f7"),
```



```

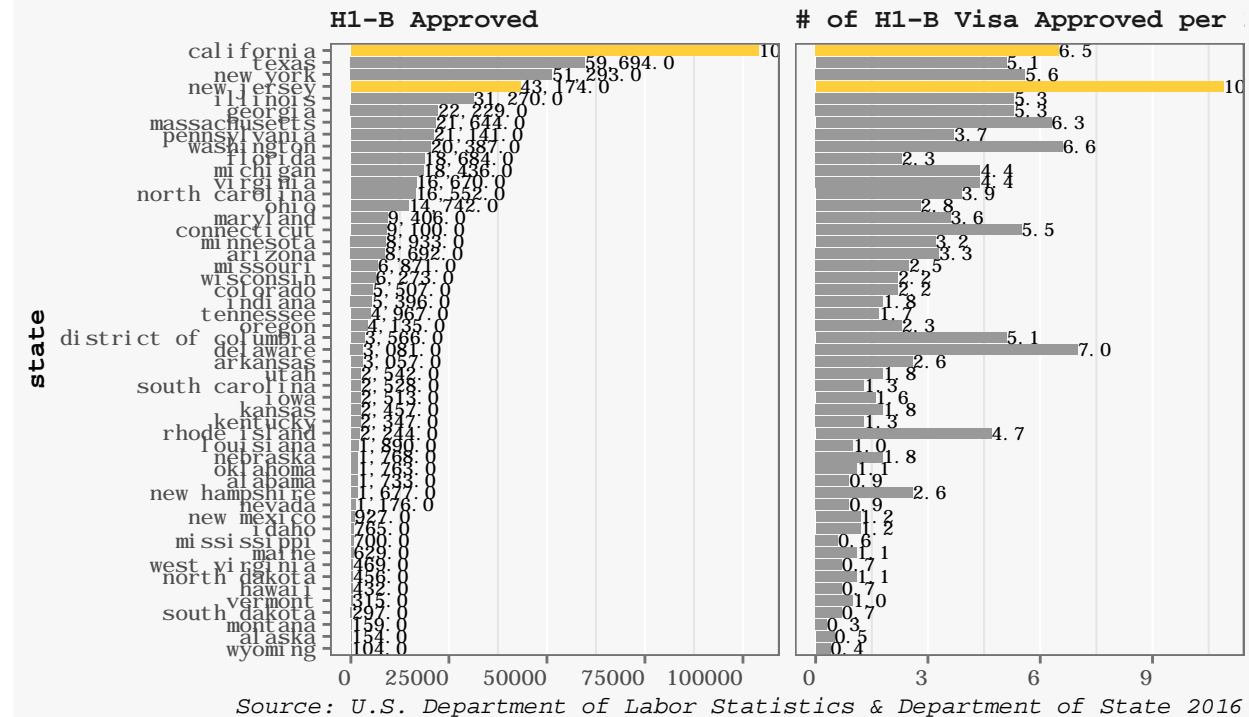
plot.title = element_text(family = 'Courier New', hjust = 0, face = "bold", size = 15),
plot.subtitle = element_text(family = 'Courier New', face = "italic", size = 10),
plot.caption = element_text(family = 'Courier New', face = "italic", size = 9),
strip.text.x = element_text(family = 'Courier New', hjust = 0, face = "bold", size = 10),
panel.grid.major.x=element_line(color="gray100"),
panel.grid.minor.x=element_line(color="gray90")

```

plot(ff)

Which state hosts the most H1-B workers

*California hosts the majority of H1-B workers,
but New Jersey is more likely to employ H1-B workers.*



This chart showed the distribution of population and density (per 1000 workers) of H1-B approved workers in 2016. California, where Silicon Valley is based, unsurprisingly tops the chart in terms of H1-B intake. New Jersey, in the other hand, surprisingly appears to have the highest concentration of H1-B workers in their workforce.