

CEO Diversity - Richie Zweigenhaft

Ryan Larson - TSP

07/12/2019

```
setwd("C:/Users/DELL/Documents/UMN/TSP/ceo diversity")
#set to local path with data to reproduce

library(readxl)
library(tidyr)
library(dplyr)

## Warning: package 'dplyr' was built under R version 3.6.1

library(stringr)
library(forcats)
library(ggplot2)

#Was not sent compensation data
#could do a plot of compensation amounts
#colored by gender/race to highlight points made in 4/5 paragraphs
```

2018 Distributions by Race, Gender, and List

```
triad <- read_xlsx("Triad's top compensation, 2018, 5.27.19.xlsx") %>%
  rename(id = ...1, name = ...2, race = ...5, gender = Gender) %>%
  mutate(race_cat = case_when(
    race=="Af Am" ~ "African American",
    race=="Canadian (Indian background)" ~ "Other",
    race=="Latino" ~ "Latino",
    race=="white" ~ "White"
  ),
  gender = str_to_title(gender),
  list = "Triad") %>%
  select(name, race_cat, gender, list)

nyt <- read_xlsx("Top CEO compensation, Fortune 500, top 200, 5.27.19.xlsx",
  col_names = F) %>%
  rename(name = ...1, race = ...3, gender = ...4) %>%
  mutate(race_cat = case_when(
    race=="AFAM" ~ "African American",
    race=="Indian" ~ "Indian/Ind. Am.",
    race=="Indian American" ~ "Indian/Ind. Am.",
    race=="Iranian-born" ~ "Other",
    race=="Latino" ~ "Latino",
    race=="Latino (Argentina)" ~ "Latino",
    race=="Latino (Brazil)" ~ "Latino",
    race=="white" ~ "White",
    race=="white (Australian)" ~ "White",
```

```

race=="white (French)" ~ "White",
race=="white (Italian)" ~ "White",
race=="white (Swedish)" ~ "White"
),
gender = str_replace_all(gender,
                          pattern = "mle",
                          replacement = "male"),
gender = str_to_title(gender),
list = "NYT") %>%
select(name, race_cat, gender, list)

```

```
listcomb <- bind_rows(nyt, triad)
```

Race Distribution (by List)

Bar Plot

```

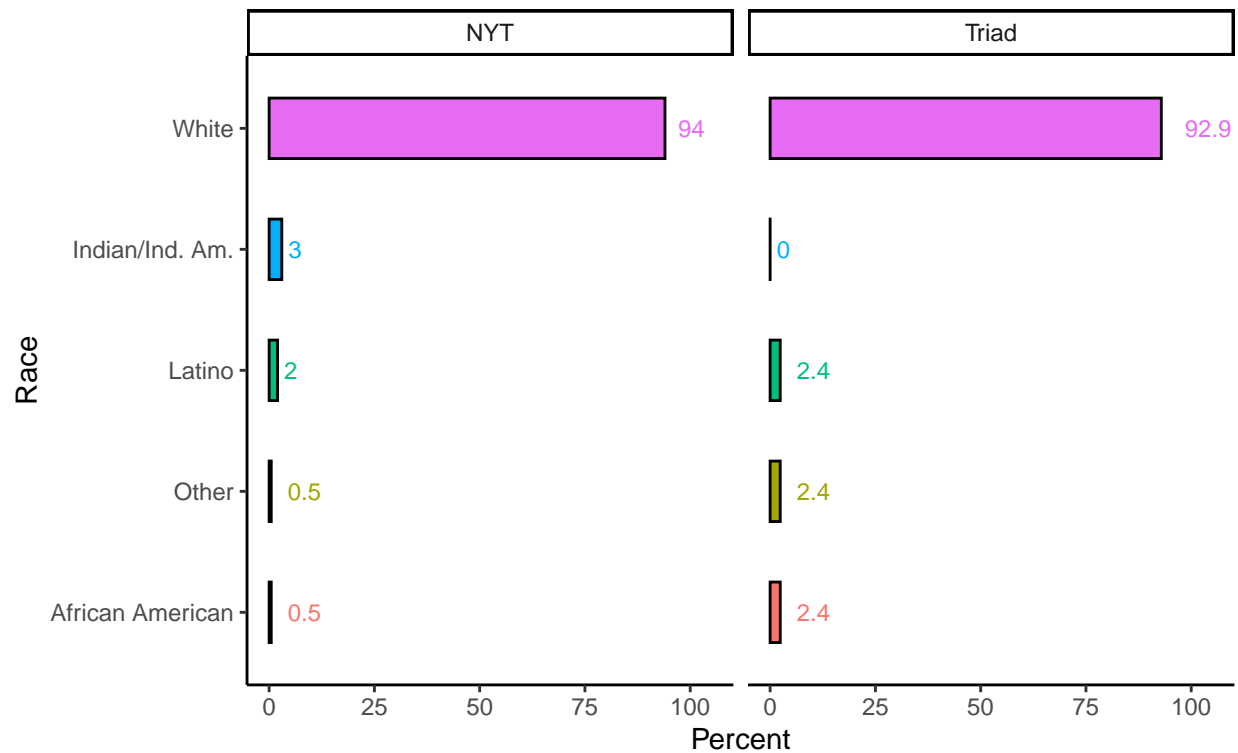
race <- listcomb %>% count(list, race_cat) %>%
  complete(list, race_cat, fill = list(n=0)) %>%
  group_by(list) %>%
  mutate(perc = round(n/sum(n, na.rm = T), digits = 3)*100) %>%
  ungroup() %>%
  mutate(race_cat = fct_reorder(race_cat, n, .desc = F))

#bar plot - race
ggplot(race, aes(x = race_cat, y = perc, fill = race_cat))+
  geom_bar(stat="identity", width = .5, color="black") +
  geom_text(aes(label = perc, color=race_cat), size = 3, hjust=-.5)+
  scale_y_continuous(limits = c(0,105))+
  facet_grid(~list)+
  labs(title = "Race Distribution of CEOs by List, 2018",
       subtitle = "Source: The New York Times and News & Record (Greensboro, NC)" ,
       y = "Percent",
       x = "Race")+
  theme_classic()+
  coord_flip() +
  theme(legend.position = "none")

```

Race Distribution of CEOs by List, 2018

Source: The New York Times and News & Record (Greensboro, NC)

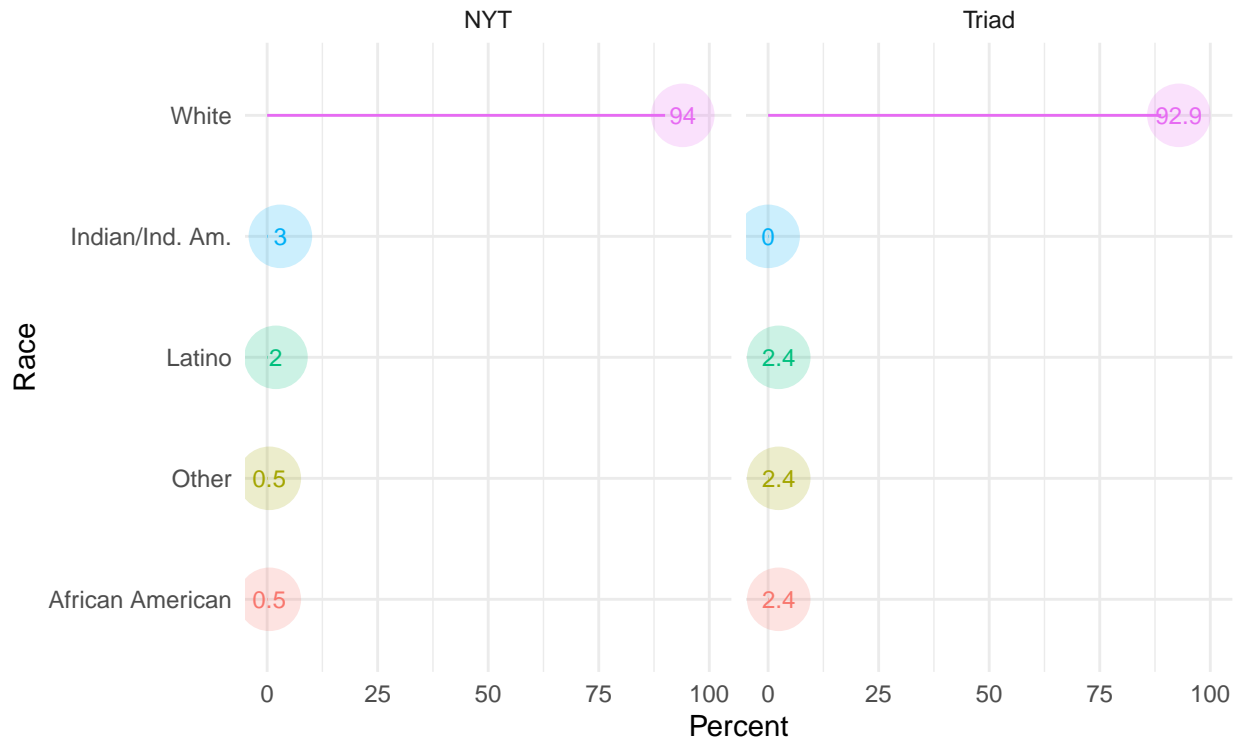


Lollipop Chart

```
#lollipop chart - race
ggplot(race, aes(x = race_cat, y = perc, color=race_cat))+
  geom_linerange(aes(ymin=0, max=ifelse(perc-4 < 0, 0.1, perc-4))) +
  geom_point(size=10.5, alpha=0.2)+
  geom_text(aes(label = perc),size = 3)+
  scale_y_continuous(limits = c(0,100))+
  facet_grid(~list)+
  labs(title = "Race Distribution of CEOs by List, 2018",
       subtitle = "Source: The New York Times and News & Record (Greensboro, NC)" ,
       y = "Percent",
       x = "Race")+
  theme_minimal()+
  coord_flip()+
  theme(legend.position = "none")
```

Race Distribution of CEOs by List, 2018

Source: The New York Times and News & Record (Greensboro, NC)



Gender Distribution (by List)

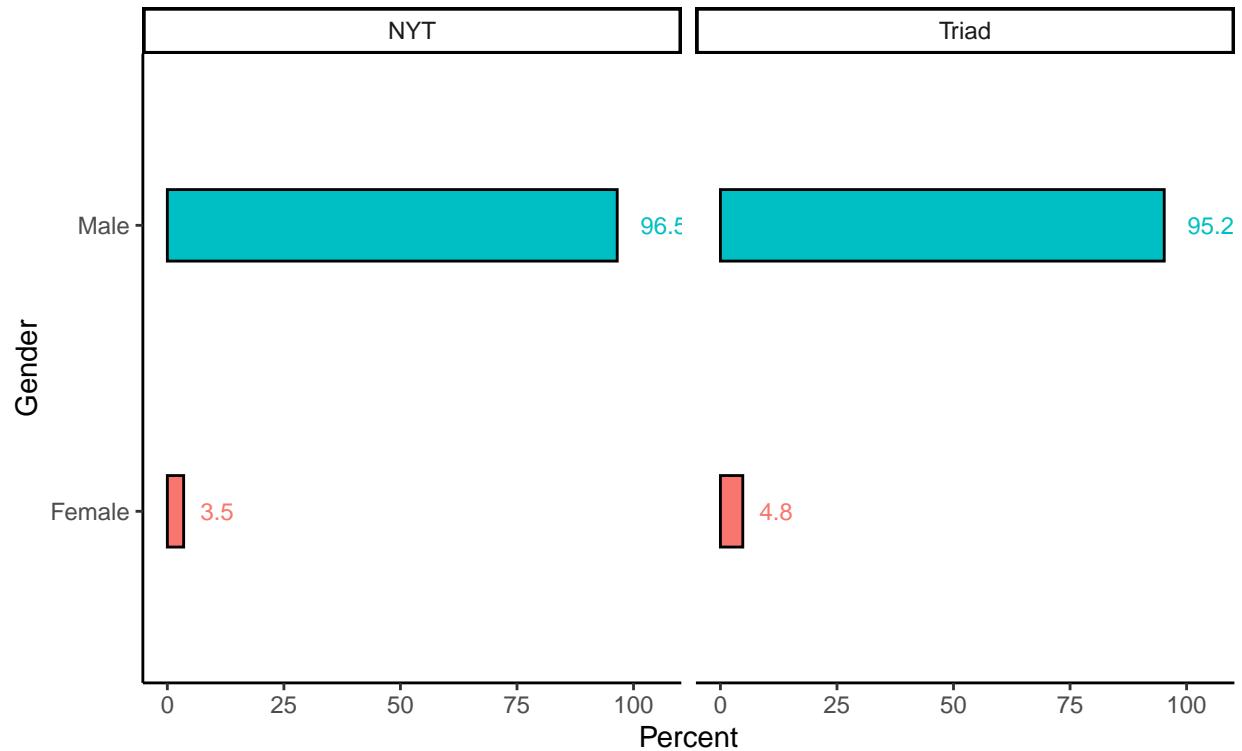
Bar Plot

```
gender <- listcomb %>% count(list, gender) %>%
  group_by(list) %>%
  mutate(perc = round(n/sum(n, na.rm = T), digits = 3)*100) %>%
  ungroup() %>%
  mutate(gender = fct_reorder(gender, n, .desc = F))

#bar plot - gender
ggplot(gender, aes(x = gender, y = perc, fill = gender))+
  geom_bar(stat="identity", width = .25, color="black") +
  geom_text(aes(label = perc, color=gender), size = 3, hjust = -.5)+
  scale_y_continuous(limits = c(0,105))+
  facet_grid(~list)+
  labs(title = "Gender Distribution of CEOs by List, 2018",
       subtitle = "Source: The New York Times and News & Record (Greensboro, NC)",
       y = "Percent",
       x = "Gender")+
  theme_classic()+
  coord_flip()+
  theme(legend.position = "none")
```

Gender Distribution of CEOs by List, 2018

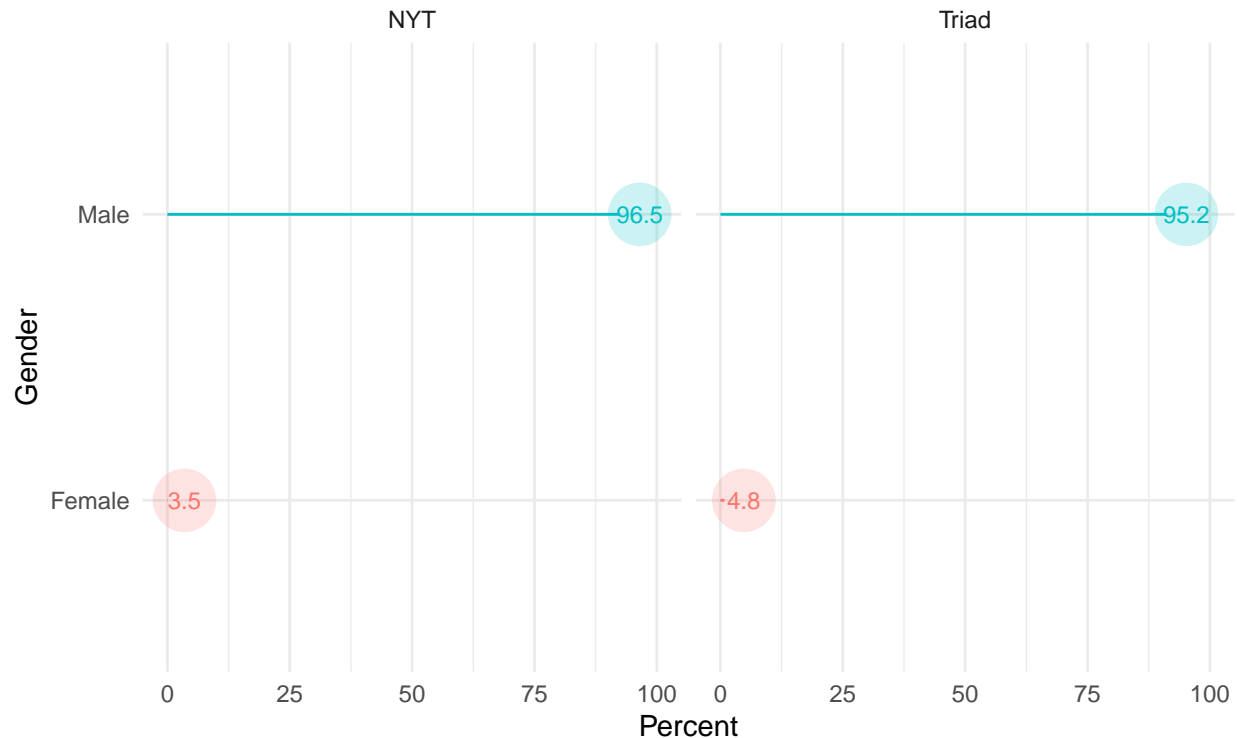
Source: The New York Times and News & Record (Greensboro, NC)



```
#lollipop chart - gender
ggplot(gender, aes(x = gender, y = perc, color = gender)) +
  geom_linerange(aes(ymin=0, max=ifelse(perc-4 < 0, 0.1, perc-4))) +
  geom_point(size=10.5, alpha=0.2) +
  geom_text(aes(label = perc), size = 3) +
  scale_y_continuous(limits = c(0,100)) +
  facet_grid(~list) +
  labs(title = "Gender Distribution of CEOs by List, 2018",
        subtitle = "Source: The New York Times and News & Record (Greensboro, NC)" ,
        y = "Percent",
        x="Gender") +
  theme_minimal() +
  coord_flip() +
  theme(legend.position = "none")
```

Gender Distribution of CEOs by List, 2018

Source: The New York Times and News & Record (Greensboro, NC)



Bivariate Distribution

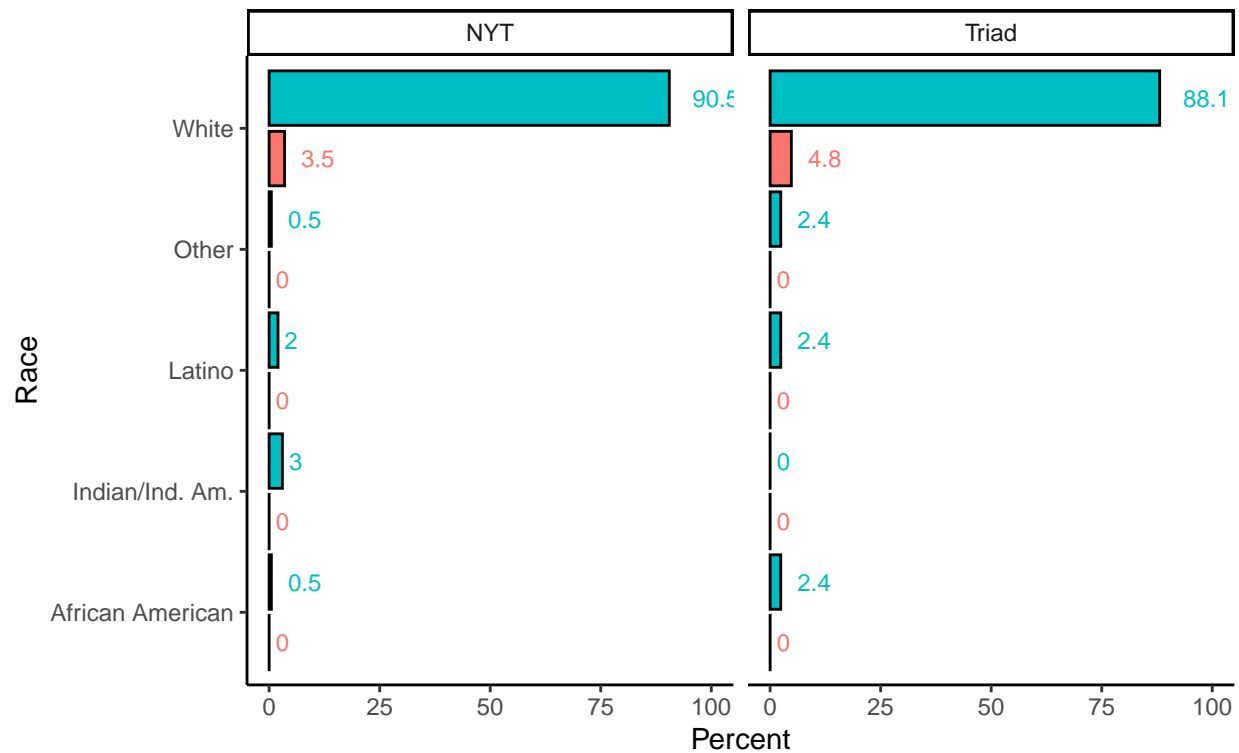
Bar Plot

```
biv <- listcomb %>% count(list, race_cat, gender) %>%
  complete(list, race_cat, gender, fill = list(n=0)) %>%
  group_by(list) %>%
  mutate(perc = round(n/sum(n, na.rm = T), digits = 3)*100)

#bar plot - bivariate
ggplot(biv, aes(x = race_cat, y = perc, fill = gender, color=gender))+
  geom_bar(position = position_dodge(width=1), stat="identity", color = "black") +
  geom_text(aes(label = perc), position = position_dodge(width=1),
    size = 3, hjust=-.5)+
  scale_y_continuous(limits = c(0,100))+
  facet_grid(~list)+
  labs(title = "Demographic Distribution of CEOs by List, 2018",
    subtitle = "Source: The New York Times and News & Record (Greensboro, NC)",
    y = "Percent",
    x = "Race")+
  theme_classic()+
  coord_flip()+
  theme(legend.position = "none")
```

Demographic Distribution of CEOs by List, 2018

Source: The New York Times and News & Record (Greensboro, NC)



Lollipop Chart

```
#lollipop chart - bivariate
ggplot(biv, aes(x = race_cat, y = perc, color = gender)) +
  geom_linerange(aes(ymin=0, max=ifelse(perc-4 < 0, 0.1, perc-4)),
    position = position_dodge(width=1)) +
  geom_point(size=10.5, alpha=0.2, position = position_dodge(width=1)) +
  geom_text(aes(label = perc), size = 3,
    position = position_dodge(width=1), show.legend = F) +
  scale_y_continuous(limits = c(0,100)) +
  facet_grid(~list) +
  labs(title = "Demographic Distribution of CEOs by List, 2018",
    subtitle = "Source: The New York Times and News & Record (Greensboro, NC)",
    y = "Percent",
    x = "Race",
    color="Gender") +
  theme_classic() +
  coord_flip()
```

Demographic Distribution of CEOs by List, 2018

Source: The New York Times and News & Record (Greensboro, NC)

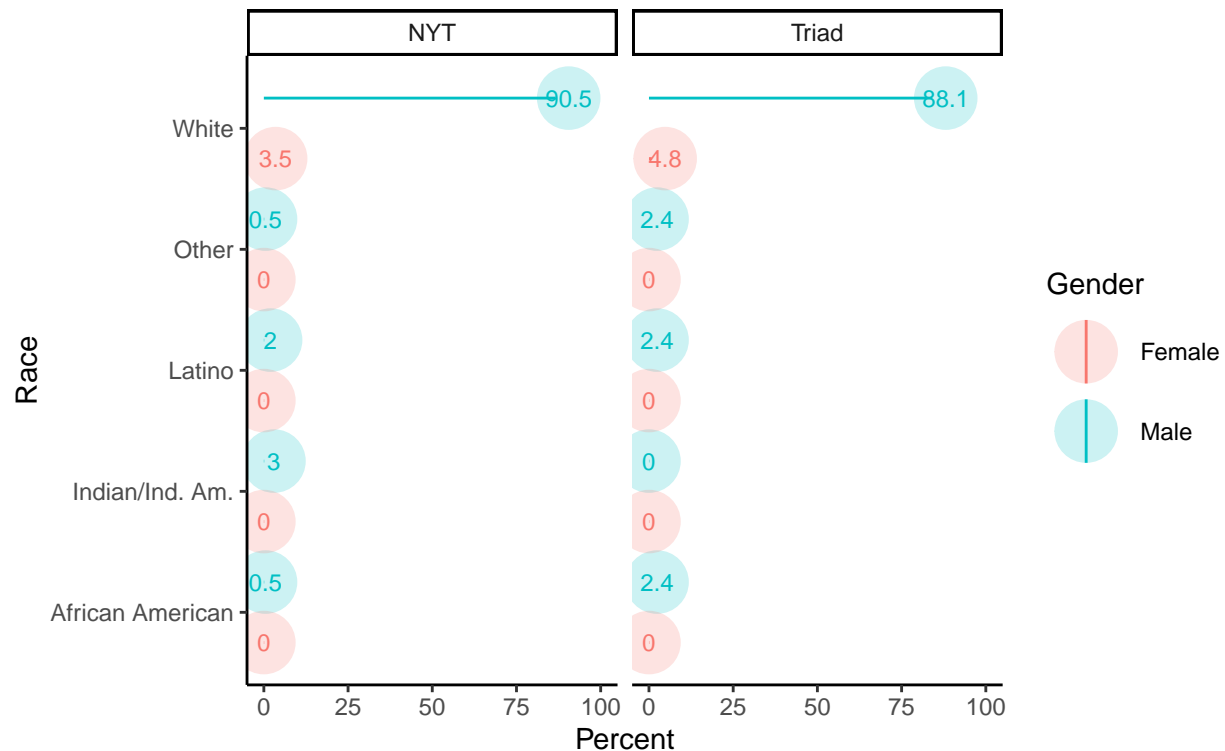


Table 2 Visualization: Fortune 500 2005-2017

Bar Plot

#using table 2 numbers (from word doc) to construct 2005v2017 plot

```
x1 <- c("2005", "2017")
x2 <- c("White", "African American", "Asian American", "Latino")
x3 <- c("Male", "Female")
x4 <- c(467,10,6,0,7,1,9,0,447,28,4,0,8,2,10,1)
```

```
fortune <- expand.grid(year = x1, race = x2, gender = x3) %>% arrange(year, race) %>% cbind(n = x4) %>%
  group_by(year) %>%
  mutate(perc = round(n/sum(n, na.rm = T), digits = 3)*100) %>%
  mutate(race = fct_reorder(race, n, .desc = F),
         gender = fct_reorder(gender, n, .desc = F))
```

#bar plot - bivariate

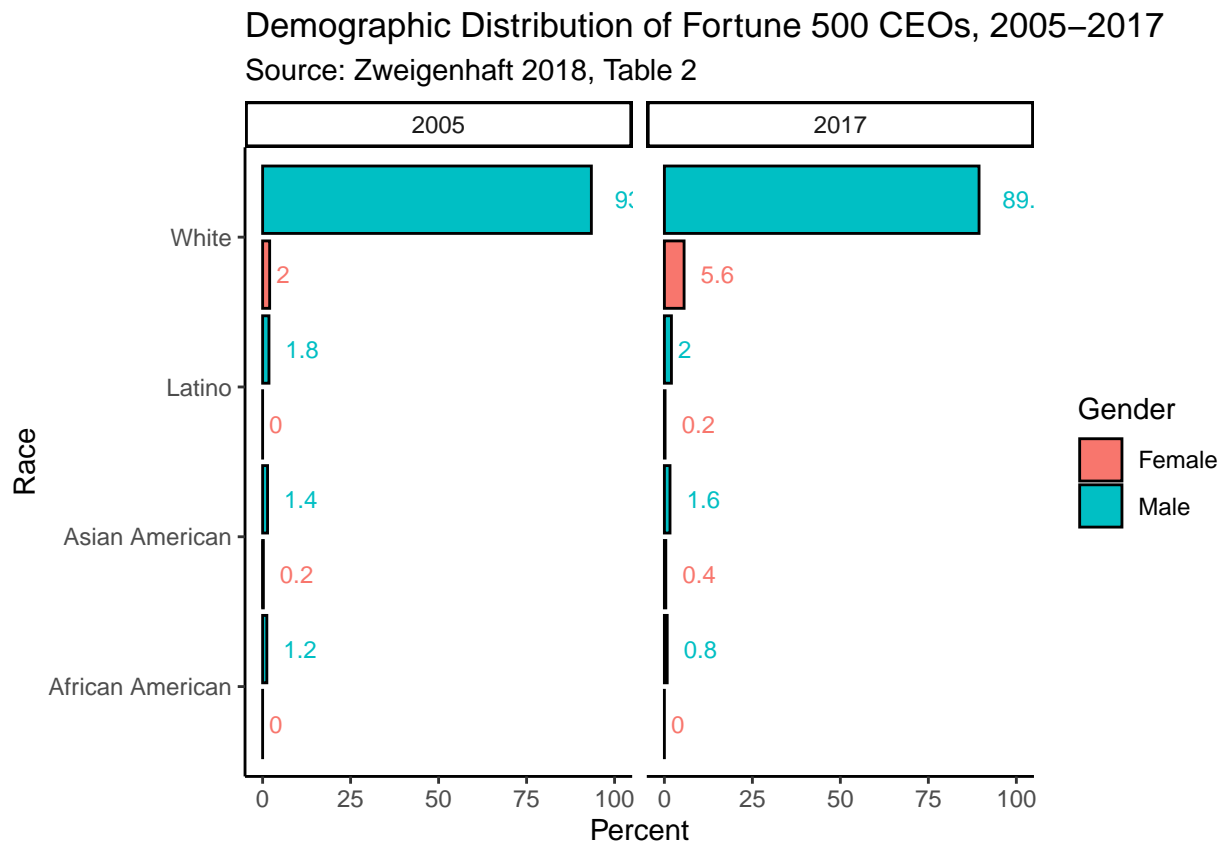
```
ggplot(fortune, aes(x = race, y = perc, fill = gender, color=gender))+
  geom_bar(position = position_dodge(width=1), stat="identity", color = "black") +
  geom_text(aes(label = perc), position = position_dodge(width=1),
            size = 3, hjust=-.5, show.legend = F)+
  scale_y_continuous(limits = c(0,100))+
```



```

facet_grid(~year)+
labs(title = "Demographic Distribution of Fortune 500 CEOs, 2005-2017",
      subtitle = "Source: Zweigenhaft 2018, Table 2",
      y = "Percent",
      x = "Race",
      fill="Gender")+
theme_classic()+
coord_flip()

```



Lollipop Chart

```

#lollipop chart - bivariate
ggplot(fortune, aes(x = race, y = perc, color = gender))+
  geom_linerange(aes(ymin=0, max=ifelse(perc-4 < 0, 0.1, perc-4)),
                position = position_dodge(width=1)) +
  geom_point(size=10.5, alpha=0.2, position = position_dodge(width=1))+
  geom_text(aes(label = perc),size = 3,
            position = position_dodge(width=1), show.legend = F)+
  scale_y_continuous(limits = c(0,100))+
  facet_grid(~year)+
  labs(title = "Demographic Distribution of Fortune 500 CEOs, 2005-2017",
        subtitle = "Source: Zweigenhaft 2018, Table 2",
        y = "Percent",
        x = "Race",

```

```

color="Gender")+
theme_classic()+
coord_flip()

```

Demographic Distribution of Fortune 500 CEOs, 2005–2017

Source: Zweigenhaft 2018, Table 2

