

# Mean Comparison

*Syed Badruddoza*

*August 27, 2019*

## Income disparity graph

An example of concisely visualizing mean differences.

```
rm(list=ls())
library(ggplot2)
#library(directlabels)
#my data set:
oc <- read.csv("http://www.columbia.edu/~mad2200/oc.csv")
oc$percent_women <- oc$W_employment/(oc$M_employment+oc$W_employment)
oc$size <- oc$W_employment+oc$M_employment
ggplot(oc, aes(M_w_earnings, W_w_earnings,
               label = as.character(Occupational.Group)))+
  geom_point(aes(size = size, colour=percent_women))+
  #scale_size_continuous(range=c(0,30)) + #scale_area()+
  #geom_point(aes(colour = oc$percent_women)) +
  coord_equal() +
  scale_colour_gradient(high = "red")+
  ylim(700, 1700) +
  xlim(700, 1700) +
  geom_abline(slope=1) +
  labs(title = "Income Disparity by Occupation and Gender") +
  ylab("Women's Weekly Earnings in $") +
  xlab("Men's Weekly Earnings in $")+
  geom_text(aes(label=Occupational.Group), size=3)
```

