[redacted]

```
knitr::opts_chunk$set(echo = TRUE)
packageload <- c("tidyverse", "dplyr", "here", "Hmisc", "purrr", "stats","boot")
library(readxl) # read excel files
sapply(packageload, library, character.only = T, quietly = T)
```

input raw data

```
rawdata<- readxl::read_excel(paste(here::here(), "[redacted].xlsx", sep = "/"),
                              sheet = "Data", col_names = c("ID","Auto", "SE1", "SE2","SE3", "SS1", "S
# load demo data
rawdemo <- readxl::read_excel(paste(here::here(), "[redacted].xlsx", sep = "/"),
                              sheet = "Demographics",
                              col_names = TRUE,
                              na = c("99"))
#create mean scores for the items in data sheet and demographics
cleandata <- rawdata %>% mutate(
    SE = rowMeans(dplyr::select(., starts_with("SE")), na.rm = T), # self-effeicacy
    SS = rowMeans(dplyr::select(., starts_with("SS")), na.rm = T), # supervisor support
    TS = rowMeans(dplyr::select(., starts_with("TS")), na.rm = T), # task signifiance
    JS = rowMeans(dplyr::select(., starts_with("JS")), na.rm = T), # job satisfaction
    FP = rowMeans(dplyr::select(., FP1:FP2), na.rm = T), # fair procedures
    OA = rowMeans(dplyr::select(., OA1:OA2), na.rm = T), # organizational authenticity
    JE = rowMeans(dplyr::select(., JE1:JE2), na.rm = T), # job embeddedness
    CR = rowMeans(dplyr::select(., CR1:CR3), na.rm = T)  # coworker relationships
) %>% dplyr::select(ID, Auto, SE:CR, everything())
# combine cleaned data with demographics data
cleandata_cb <- cleandata %>% left_join(rawdemo, by = c("ID" = "ID"))
head(cleandata_cb)
```

```
## # A tibble: 6 x 40
##      ID  Auto    SE    SS    TS    JS    FP    OA    JE    CR   SE1   SE2
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  2582     3  3.33  3     1     2.25  5     1     2     4.33     1     3
## 2  5225     7  7     6.33  6.33  5.25  6     6     6.5   5.67     7     7
## 3  4689     5  4.33  4     5     3.5   4.5   4.5   6     3.67     4     3
## 4  3851     7  6     6.67  6.33  6.25  7     6.5   6.5   6.33     6     6
## 5  9637     7  6.33  6.67  6.33  6.5   5.5   7     6     5.67     5     7
## 6  7127     4  3.67  5.67  5     5.5   5.5   5.5   3.5   5.33     3     3
## # ... with 28 more variables: SE3 <dbl>, SS1 <dbl>, SS2 <dbl>, SS3 <dbl>,
## #   TS1 <dbl>, TS2 <dbl>, TS3 <dbl>, JS1 <dbl>, JS2 <dbl>, JS3 <dbl>,
## #   JS4 <dbl>, FP1 <dbl>, FP2 <dbl>, OA1 <dbl>, OA2 <dbl>, JE1 <dbl>,
## #   JE2 <dbl>, CR1 <dbl>, CR2 <dbl>, CR3 <dbl>, gender <dbl>,
## #   manager <chr>, tenure <dbl>, hourly_salaried <dbl>, income <dbl>,
## #   region <chr>, highest_ed <dbl>, department <chr>
```

```
#reorganize data (code demographics as factors)
cleandata_cb_2 <- cleandata_cb %>% select(ID:CR) %>% bind_cols(cleandata_cb %>%
  select(gender:department) %>% mutate_if(is.numeric, ~as.factor(.)))
cleandata_cb_2_demo <- cleandata_cb_2 %>% select(ID,JS,c(gender:department))
```

Exploratory Analysis

Descriptives and Correlation Matrix among Variables

This step is to examine correlations among key variables: all predictors are moderately to highly correlated with each other, and all predictors are moderately to highly correlated with Job Satisfaciton. OLS regressions would not be a good choice in this situation. Relative weight analysis would be a good option in this case to examine the variance contributed by predictors to the outcome (job satisfaction).

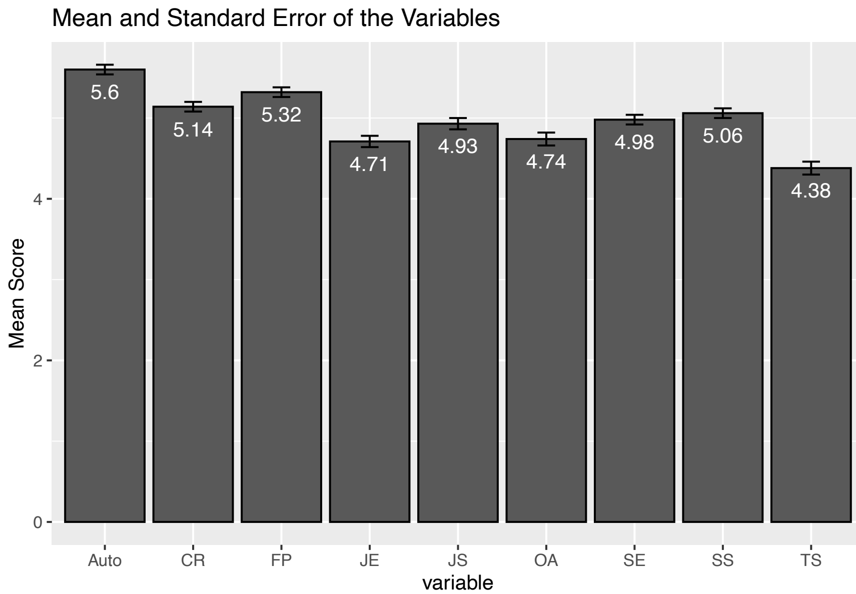
```
se <- function(x,na.rm = TRUE) sqrt(var(x,na.rm = TRUE)/length(x))
cleandata_cb %>% select(ID:CR) %>%
  gather( "variable", "value", -ID, na.rm = FALSE) %>%
  group_by(variable) %>%
  summarise(Mean = round(mean(value,na.rm = TRUE),2),
            SD = round(sd(value, na.rm = TRUE),2),
            StdEr = round(se(value, na.rm = TRUE),2)) -> var_descriptive
## import function I wrote before to get correlation matrix
# This function requires the following package
# library(Hmisc)
source("Correlation_matrix_function.R")
correlations <- cormatrix(dplyr::select(cleandata_cb, Auto:CR))
correlations$variable <-  names(correlations)
correlations <- correlations %>% left_join(var_descriptive, by = c("variable")) %>%
  select(variable, Mean, SD, everything())
# rownames(correlations) = names(correlations)
correlations2 <- correlations %>% select(-StdEr)
print(correlations2)
```

```
##   variable Mean   SD    Auto      SE      SS      TS      JS      FP
## 1     Auto 5.60 1.28       1 0.72*** 0.47*** 0.45*** 0.56***  0.6***
## 2       SE 4.98 1.35 0.72***       1  0.7*** 0.72*** 0.78*** 0.72***
## 3       SS 5.06 1.30 0.47***  0.7***       1 0.76*** 0.77*** 0.59***
## 4       TS 4.38 1.67 0.45*** 0.72*** 0.76***       1 0.82*** 0.67***
## 5       JS 4.93 1.49 0.56*** 0.78*** 0.77*** 0.82***       1 0.74***
## 6       FP 5.32 1.22  0.6*** 0.72*** 0.59*** 0.67*** 0.74***       1
## 7       OA 4.74 1.63 0.47*** 0.69*** 0.69*** 0.78*** 0.82*** 0.65***
## 8       JE 4.71 1.58 0.44*** 0.74*** 0.74*** 0.82*** 0.77*** 0.65***
## 9       CR 5.14 1.27 0.56*** 0.78*** 0.73*** 0.76*** 0.85*** 0.76***
##        OA      JE      CR
## 1 0.47*** 0.44*** 0.56***
## 2 0.69*** 0.74*** 0.78***
## 3 0.69*** 0.74*** 0.73***
## 4 0.78*** 0.82*** 0.76***
## 5 0.82*** 0.77*** 0.85***
## 6 0.65*** 0.65*** 0.76***
## 7       1 0.72*** 0.76***
## 8 0.72***       1 0.76***
## 9 0.76*** 0.76***       1
```

```
# write.csv(correlations2,"correlations2.csv")
```

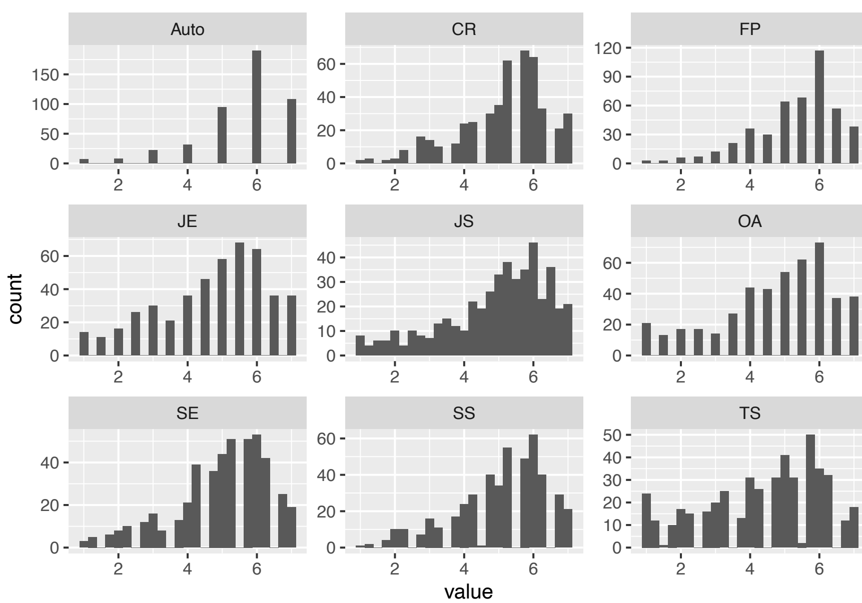
As shown in the graphic below, mean scores of the variables are all high. Task Significance, Organizational Authenticity, and Job Embeddedness are relatively lower than others. Job satisfaction is also high.

```
ggplot(correlations, aes(x=variable, y=Mean)) +
  geom_bar(stat="identity", color="black",
           position=position_dodge()) +
  geom_errorbar(aes(ymin=Mean-StdEr, ymax=Mean+StdEr), width=.2,
                position=position_dodge(.9)) +
  labs(title = "Mean and Standard Error of the Variables", y = "Mean Score") +
  geom_text(aes(label=Mean), vjust=2, color="white", size=4)
```



Graphs below show the score distributions of all variables. Scores are slightly skewed; no variables are highly skewed. For the time being, no transformations will be conducted. If time permits, transformations of the distrubtions may be conducted and sensitivity analysis needs to be conducted to see if the transformations have a large effect on the results.

```
cleandata_cb %>% dplyr::select(Auto:CR) %>% gather() %>% ggplot(aes(value)) +
  facet_wrap(~ key, scales = "free") +
  geom_histogram(binwidth = 0.25)
```



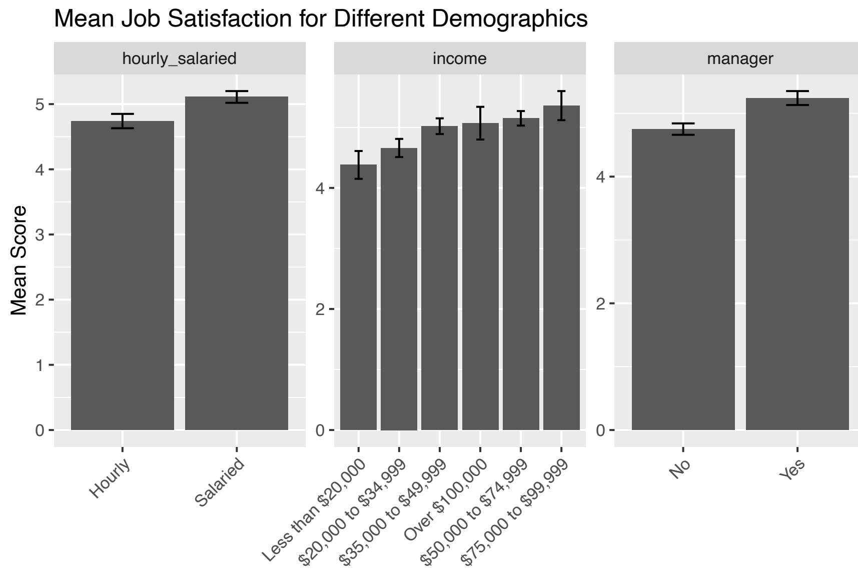
Demographic Differences in Job Satisfaction

This step is to examine if there are any demographic differences in job satisfaction using ANOVA. There are significant differences in hourly_salaried, income, manager, but not others. Results of relative weight analysis will be compared for hourly_salaried(F(1,460) = 6.906, p < .05), income(F(5,456) = 3.446, p < .01), manager(F(1,460) = 11.74, p < .001). ANOVA output is shown at the end of the document (supplemental material section)

```
anova.hourly_salaried2 <- aov(JS ~ factor(hourly_salaried), data = cleandata_cb_2_demo)
summary(anova.hourly_salaried2)
Tukey.means.hourly_salaried <- TukeyHSD(x=anova.hourly_salaried2, 'factor(hourly_salaried)', conf.lev
Tukey.means.hourly_salaried
anova.income2 <- aov(JS ~ factor(income), data = cleandata_cb_2_demo)
summary(anova.income2)
Tukey.means.income <- TukeyHSD(x=anova.income2, 'factor(income)', conf.level=0.95)
Tukey.means.income
anova.manager2 <- aov(JS ~ factor(manager), data = cleandata_cb_2_demo)
summary(anova.manager2)
Tukey.means.manager <- TukeyHSD(x=anova.manager2, 'factor(manager)', conf.level=0.95)
Tukey.means.manager
anova.gender<- lm(JS ~ factor(gender), data = cleandata_cb_2_demo); anova(anova.gender)
anova.tenure <- lm(JS ~ factor(tenure), data = cleandata_cb_2_demo); anova(anova.tenure)
anova.highest_ed <- lm(JS ~ factor(highest_ed), data = cleandata_cb_2_demo); anova(anova.highest_ed)
anova.region <- lm(JS ~ factor(region), data = cleandata_cb_2_demo); anova(anova.region)
anova.department <- lm(JS ~ factor(department), data = cleandata_cb_2_demo); anova(anova.department)
```

Graphs below show job satisfaction scores for different levels of hourly_salaried, income, manager, since ANOVA revealed that there are differences in job satisfaction for these demographic variables.

```
cleandata_cb_2_demo %>% select(ID, JS, manager, hourly_salaried, income) %>%
  mutate( hourly_salaried = case_when( is.na(hourly_salaried) ~ "NA",
              hourly_salaried == 1 ~ "Hourly", hourly_salaried == 2 ~ "Salaried"),
          income = case_when(is.na(income) ~ "NA", income == 1 ~ "Less than $20,000",
              income == 2 ~ "$20,000 to $34,999", income == 3 ~ "$35,000 to $49,999",
              income == 4 ~ "$50,000 to $74,999", income == 5 ~ "$75,000 to $99,999",
              income == 6 ~ "Over $100,000")) %>%
  gather( "demo", "value", -ID, -JS, na.rm = FALSE) %>%
  dplyr::select(-ID) %>% group_by(demo, value) %>%
  summarise(demo_mean = round(mean(JS,na.rm = TRUE),2),demo_StdEr =round(se(JS,na.rm = TRUE),2)) %>%
  ggplot(aes(x = reorder(factor(value), demo_mean), y = demo_mean)) +
  geom_bar(stat="identity") +
  geom_errorbar(aes(ymin=demo_mean-demo_StdEr, ymax=demo_mean+demo_StdEr), width=.2,
                position=position_dodge(.9)) +
  labs(title = "Mean Job Satisfaction for Different Demographics", x = "", y = "Mean Score") +
  facet_wrap(~ demo, scales = "free") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Explore Missing Data

This is to examine missing data for raw scores and scale scores: since there are only two missing data points. those observations will be excluded to simplify analysis (e.g., relative weight analysis, etc.); if time permits, sensitivity analysis needs to be conducted to examine if excluding those cases would create differences in the results. Multiple imputation could be used to impute the missing data.

```
# rawdata %>% dplyr::filter_all(any_vars(is.na(.)))
cleandata_cb %>% dplyr::select(ID:CR) %>% dplyr::filter_all(any_vars(is.na(.)))
```

```
## # A tibble: 2 x 10
##      ID  Auto    SE    SS    TS    JS    FP    OA    JE    CR
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1    98     7  6.33  6.33  6.33  7        7   NaN     7     7
## 2    70     6  4.33  4.5   1.5   4.25     5   NaN     6     4
```

Relative Weight Analysis

To take into account that all variables are moderately to highly correlated with each other and with the outcome job satisfaction, relative weights analysis is conducted to examine the relative cocntributions of each predictor to the outcome Job Satisfaciton.

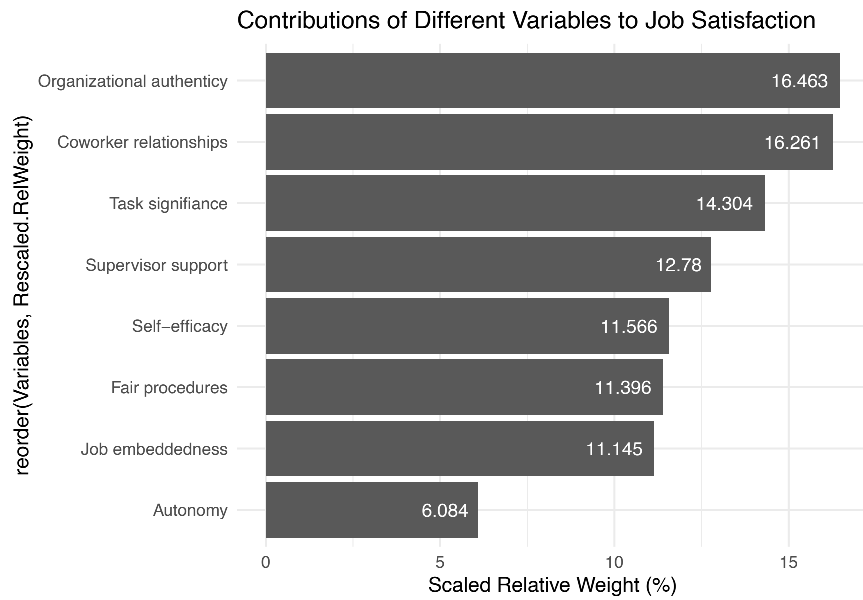
```
## import function I used before to conduct relative weights analysis with bootstrapped resamples
source("Relative_weights_analysis_function_submitted.R")
# remove missing data so the script can computer the CIs for relative weights
# with bootstrapped resamples
thedata <- cleandata_cb %>% dplyr::select(Auto:CR) %>%
  dplyr::filter_all(all_vars(!is.na(.))) %>% dplyr::select(JS, everything()) %>% as.data.frame()
# rwa analysis
rwa_results <- rwa(thedata,iter = 5000)
RSQ.Results<-rwa_results[[1]]
RW.Results<-rwa_results[[2]]
CI.Results<-rwa_results[[3]]
CI.Significance<-rwa_results[[4]]
rwa.output <- RW.Results %>% left_join(CI.Results, by = c("Variables")) %>%
```

```
  left_join(CI.Significance, by = c("Variables" = "Labels")) %>%
  mutate_if(vars(is.numeric(.)), ~round(.,digits = 3)) %>%
  mutate(#CI.Raw = paste0("[",CI.Lower.Bound.x,", ",CI.Upper.Bound.x,"]"),
         CI.Sig = paste0("[",CI.Lower.Bound.y,", ",CI.Upper.Bound.y,"]")) %>%
  select(Variables:Rescaled.RelWeight, CI.Sig) %>%
  mutate(Variables = case_when(
    Variables %in% c("Auto") ~ "Autonomy", Variables %in% c("SE") ~ "Self-efficacy",
    Variables %in% c("SS") ~ "Supervisor support",Variables %in% c("TS") ~ "Task signifiance",
    Variables %in% c("FP") ~ "Fair procedures",Variables %in% c("OA") ~ "Organizational authenticy",
    Variables %in% c("JE") ~ "Job embeddedness",Variables %in% c("CR") ~ "Coworker relationships"
  ) ) %>% arrange(desc(Rescaled.RelWeight))
rwa.output2 <- rwa.output  %>%
  rename('Raw Weight'= Raw.RelWeight,'Scaled Weight (%)' = Rescaled.RelWeight,'CI Ho testing' = CI.Si
print(rwa.output2) #; write.csv(rwa.output2, "rwa.output2.csv")
```

```
##                    Variables Raw Weight Scaled Weight (%)  CI Ho testing
## 1 Organizational authenticy      0.137            16.463  [0.119, 0.16]
## 2    Coworker relationships      0.136            16.261 [0.119, 0.155]
## 3          Task signifiance      0.119            14.304 [0.105, 0.135]
## 4        Supervisor support      0.107            12.780 [0.089, 0.124]
## 5             Self-efficacy      0.097            11.566 [0.081, 0.112]
## 6           Fair procedures      0.095            11.396 [0.076, 0.116]
## 7          Job embeddedness      0.093            11.145 [0.079, 0.107]
## 8                  Autonomy      0.051             6.084 [0.033, 0.069]
```

Below is a graphic display of relative contributions of the predictors to Job Satisfaction (how much variances in the outcome explained by the predictors). As shown in the figure, **Organizational Authenticy** and **Coworker Relationships** contributed to Job Satisfaction the most (explained 16.46% and 16.26% variance respectively). **Task Significance** and **Supervisor Support** also explained a good amount of variance, 14.30% and 12.78% respectively. These areas are the ones to focus on to improve to improve Job Satisfaction. Considering Organizational Authenticy and Task Significance have relatively lower mean scores, these two areas should especially be focused on for improvement.

```
rwa.output %>% ggplot(aes(x = reorder(Variables, Rescaled.RelWeight), y = Rescaled.RelWeight)) +
  geom_bar(stat = "identity") +
  labs(title = "Contributions of Different Variables to Job Satisfaction",
       y = "Scaled Relative Weight (%)") +
  coord_flip() +
  geom_text(aes(label=Rescaled.RelWeight), hjust=1.2, color="white", size=3.5) +
  theme_minimal()
```



Relative Weight Analysis for Different Demographic Groups

This step is to compare if there are demographic differences in the results of relative weight analysis. To simplifiy analysis, levels of income are divided to low vs. high income groups $50,000 is the cutoff); each group roughly have the same number of people: 265 for low income group and 197 for high income group.

As shown below, because bootstrapped confidence intervals cover zero, there's no signifiant difference for low vs. high income groups, whether they are managers or not, or whether they are hourly or not. Thus, factors influencing job satisfaction are the same across these three different demographic groups. Strategies proposed to improve the identified factors that have the most impact on job satisfaction are applicable across different groups within the organization.

```
source("Relative_weights_analysis_group_comparison_function_submitted.R")
cleandata_cb_3 <- cleandata_cb %>%
  select(ID:CR, gender:department) %>%
  mutate(income2 = as.numeric(case_when(
    is.na(income) ~ "NA", income <= 3 ~ "1", income >3 ~ "2")),
  manager2 = as.numeric(case_when(
    is.na(manager) ~ "NA", manager == "Yes" ~ "1", manager == "No" ~ "2")))
thedata <- cleandata_cb_3 %>% select(JS, Auto, SE, SS, TS, FP, OA, JE, CR) %>% as.data.frame()
### income
grpdata<-cleandata_cb_3 %>%
  select(JS, Auto, SE, SS, TS, FP, OA, JE, CR, income2) %>%
  rename(grouping = income2) %>% as.data.frame()
rwa_compare_results_income<-rwa_compare(thedata, grpdata,iter = 5000)
### manager
grpdata<-cleandata_cb_3 %>%
  select(JS, Auto, SE, SS, TS, FP, OA, JE, CR, manager2) %>%
  rename(grouping = manager2) %>% as.data.frame()
rwa_compare_results_manager<-rwa_compare(thedata, grpdata,iter = 5000)
### hourly salary
grpdata<-cleandata_cb_3 %>%
  select(JS, Auto, SE, SS, TS, FP, OA, JE, CR, hourly_salaried) %>%
  rename(grouping = hourly_salaried) %>% as.data.frame()
rwa_compare_results_hourly<-rwa_compare(thedata, grpdata,iter = 5000)
# comparing predictors across 2 groups
```

```r
# If Zero is not included, Weights are Significantly different between the groups
rwa_compare_results_income[[3]] %>%
  mutate_if(vars(is.numeric(.)), ~round(.,digits = 3)) %>%
  mutate('Income:High vs. Low' = paste0("[",CI.Lower.Bound,", ",CI.Upper.Bound,"]")) %>%
    select(-contains("CI")) %>%
  left_join(rwa_compare_results_manager[[3]] %>%
  mutate_if(vars(is.numeric(.)), ~round(.,digits = 3)) %>%
  mutate('Manager:Yes vs. No' = paste0("[",CI.Lower.Bound,", ",CI.Upper.Bound,"]")) %>%
    select(-contains("CI")), by = c("Labels")
  ) %>% left_join(rwa_compare_results_hourly[[3]] %>%
  mutate_if(vars(is.numeric(.)), ~round(.,digits = 3)) %>%
  mutate('Hourly vs. Salary' = paste0("[",CI.Lower.Bound,", ",CI.Upper.Bound,"]")) %>%
    select(-contains("CI")),by = c("Labels")) %>%
  rename(Variable = Labels) -> rwa_compare_results_all
print(rwa_compare_results_all)
```

```
##   Variable Income:High vs. Low Manager:Yes vs. No Hourly vs. Salary
## 1     Auto     [-0.027, 0.042]    [-0.044, 0.026]   [-0.026, 0.046]
## 2       SE      [-0.02, 0.038]    [-0.022, 0.038]   [-0.025, 0.032]
## 3       SS      [-0.056, 0.01]    [-0.024, 0.047]   [-0.054, 0.017]
## 4       TS      [-0.03, 0.029]    [-0.031, 0.034]   [-0.021, 0.037]
## 5       FP      [-0.02, 0.056]    [-0.067, 0.011]    [-0.03, 0.047]
## 6       OA     [-0.043, 0.036]    [-0.008, 0.078]   [-0.047, 0.035]
## 7       JE      [-0.04, 0.016]    [-0.039, 0.018]   [-0.042, 0.015]
## 8       CR     [-0.053, 0.015]    [-0.045, 0.025]   [-0.023, 0.046]
```

Supplemental Analysis

ANOVA to examine if there are demographic differences in job satisfaction

```
##                            Df Sum Sq Mean Sq F value  Pr(>F)
## factor(hourly_salaried)   1   15.2  15.176   6.906 0.00888 **
## Residuals               460 1010.8   2.197
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = JS ~ factor(hourly_salaried), data = cleandata_cb_2_demo)
##
## $`factor(hourly_salaried)`
##          diff        lwr      upr     p adj
## 2-1 0.3625694 0.09144678 0.633692 0.0088772

##                  Df Sum Sq Mean Sq F value  Pr(>F)
## factor(income)   5   37.4   7.472   3.446 0.00456 **
## Residuals      456  988.7   2.168
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = JS ~ factor(income), data = cleandata_cb_2_demo)
```

```
##
## $`factor(income)`
##            diff         lwr       upr     p adj
## 2-1  0.27604259 -0.44285549 0.9949407 0.8817282
## 3-1  0.63714892 -0.09648477 1.3707826 0.1305247
## 4-1  0.76811030  0.06245123 1.4737694 0.0238330
## 5-1  0.97448980  0.04188007 1.9070995 0.0346503
## 6-1  0.68992758 -0.27714366 1.6569988 0.3203769
## 3-2  0.36110633 -0.21355190 0.9357646 0.4678040
## 4-2  0.49206771 -0.04641945 1.0305549 0.0956526
## 5-2  0.69844720 -0.11504810 1.5119425 0.1393080
## 6-2  0.41388499 -0.43889897 1.2666690 0.7337271
## 4-3  0.13096138 -0.42704610 0.6889689 0.9849582
## 5-3  0.33734088 -0.48920527 1.1638870 0.8518626
## 6-3  0.05277866 -0.81246373 0.9180211 0.9999775
## 5-4  0.20637950 -0.59544023 1.0081992 0.9772895
## 6-4 -0.07818271 -0.91983629 0.7634709 0.9998198
## 6-5 -0.28456221 -1.32388263 0.7547582 0.9702329
```

```
##                  Df Sum Sq Mean Sq F value   Pr(>F)
## factor(manager)   1   25.5  25.542   11.74 0.000665 ***
## Residuals       460 1000.5   2.175
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = JS ~ factor(manager), data = cleandata_cb_2_demo)
##
## $`factor(manager)`
##             diff       lwr       upr     p adj
## Yes-No 0.4881657 0.2082285 0.7681029 0.0006654
```

```
## Analysis of Variance Table
##
## Response: JS
##                 Df  Sum Sq Mean Sq F value Pr(>F)
## factor(gender)   1    0.29  0.2914  0.1307 0.7179
## Residuals      460 1025.73  2.2298
```

```
## Analysis of Variance Table
##
## Response: JS
##                 Df Sum Sq Mean Sq F value  Pr(>F)
## factor(tenure)   6  27.26  4.5428  2.0695 0.05551 .
## Residuals      455 998.76  2.1951
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Analysis of Variance Table
##
## Response: JS
##                     Df  Sum Sq Mean Sq F value Pr(>F)
## factor(highest_ed)   6   17.76  2.9608  1.3361 0.2394
```

```
## Residuals          455 1008.25  2.2159
```

```
## Analysis of Variance Table
##
## Response: JS
##                 Df  Sum Sq Mean Sq F value Pr(>F)
## factor(region)   3    5.63  1.8767  0.8424 0.4712
## Residuals      458 1020.39  2.2279
```

```
## Analysis of Variance Table
##
## Response: JS
##                     Df Sum Sq Mean Sq F value Pr(>F)
## factor(department)  15  32.39  2.1591  0.9691 0.4871
## Residuals          446 993.63  2.2279
```