Final 241 Project Consolidation

author: Elizabeth Khan, Estrella Ndrianasy, Chandni Shah, Michelle Shen, Catherine Tsai date: "November 8, 2021" output: html_document —

```
# Import data and remove invalid age

d <- fread("survey_cleaned_with_states.csv")

d <- d[Age!='Other']

# Removing company test response 52.36
d <- d[ResponseId!='R_2WC5vlB2DCvHHrH']

degrees = c("Bachelors degree","Associates degree","Masters degree")

# create column to indicate whether or not subject has a college degree

# d<- d[,college_degree:= ifelse(Education %in% degrees, 1,0)]

head(d)
```

```
##           ResponseId Gender   Age             Education LocationLatitude
## 1: R_x2RQnFD8LZIZuRH Female 40-50     Associates degree         36.99040
## 2: R_O9f8Af9asil81DH Female 30-40 Some College No degree        27.99290
## 3: R_3oTqFe6v8l4jV0p Female 30-40 Some College No degree        36.83070
## 4: R_0ePUcdKMDE2jNe1 Female 30-40      Bachelors degree         39.30370
## 5: R_2owAeGSUxdN9fQR Female 40-50     Associates degree         44.14149
## 6: R_3MR7piXWpVNth0P Female 30-40 Some College No degree        38.88420
##    LocationLongitude instr_enthusiasm instr_professional instr_subject
## 1:          -85.9239                4                  4             4
## 2:          -82.4030                1                  4             3
## 3:          -76.3146                2                  2             2
## 4:          -94.9322                3                  1             1
## 5:         -103.2052                4                  4             5
## 6:          -76.9941                3                  3             3
##    instr_material instr_rating                        subject_1
## 1:              4            4                 Web Applications
## 2:              2            4                 Web Applications
## 3:              2            2                 Web Applications
## 4:              3            2 Artificial intelligence projects
## 5:              4            4                 Web Applications
## 6:              3            3                 All of the above
##                                             subject_2
## 1:                                        Both A & B
## 2: C. A low level implementation programming language
## 3:                 B. A flexible programming language
```

```
## 4:                  B. A flexible programming language
## 5:                  B. A flexible programming language
## 6:                                    All of the above
##
## 1: Python is a compiled programming language with a faster and more efficient execution time than in
## 2:                                                                                 Python is an objec
## 3:                                                                                    Python was crea
## 4:                                                                                    Python was crea
## 5:                                                                                 Python is an objec
## 6:                                                                                     Python is an i
##                subject_4       video_watched assigned assignment ques_1
## 1: Great Advanced Language   Video3-2(Control)        0    Control      0
## 2:           Easy to Learn   Video2-2(Control)        0    Control      0
## 3:     Wonderful Community Video2-1(Treatment)        1  Treatment      0
## 4:     Wonderful Community   Video1-2(Control)        0    Control      0
## 5:     Wonderful Community   Video2-2(Control)        0    Control      0
## 6:  Great Starter Language   Video3-2(Control)        0    Control      1
##    ques_2 ques_3 ques_4 quiz_avg avg_rating                state state_code
## 1:      1      1      1     0.75       4.00             kentucky         KY
## 2:      0      0      0     0.00       2.75              florida         FL
## 3:      0      0      0     0.00       2.00             virginia         VA
## 4:      0      0      0     0.00       2.25               kansas         KS
## 5:      0      0      0     0.00       4.00         south dakota         SD
## 6:      0      0      0     0.25       3.00 district of columbia         DC
##    dem_percent rep_percent
## 1:       0.362       0.621
## 2:       0.479       0.512
## 3:       0.541       0.440
## 4:       0.416       0.562
## 5:       0.356       0.618
## 6:       0.921       0.054
```

Figures and Tables

```
summary_response_table <- d %>%
  count(Gender, Age, assignment)

summary_response_table
```

```
##     Gender   Age assignment  n
##  1: Female 20-30    Control 14
##  2: Female 20-30  Treatment 11
##  3: Female 30-40    Control 27
##  4: Female 30-40  Treatment 29
##  5: Female 40-50    Control 15
##  6: Female 40-50  Treatment 17
##  7: Female   50+    Control 12
##  8: Female   50+  Treatment 17
##  9:   Male 20-30    Control  9
## 10:   Male 20-30  Treatment 10
## 11:   Male 30-40    Control 12
## 12:   Male 30-40  Treatment 14
```

```
## 13:   Male 40-50    Control 11
## 14:   Male 40-50  Treatment  5
## 15:   Male   50+    Control  8
## 16:   Male   50+  Treatment 10
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

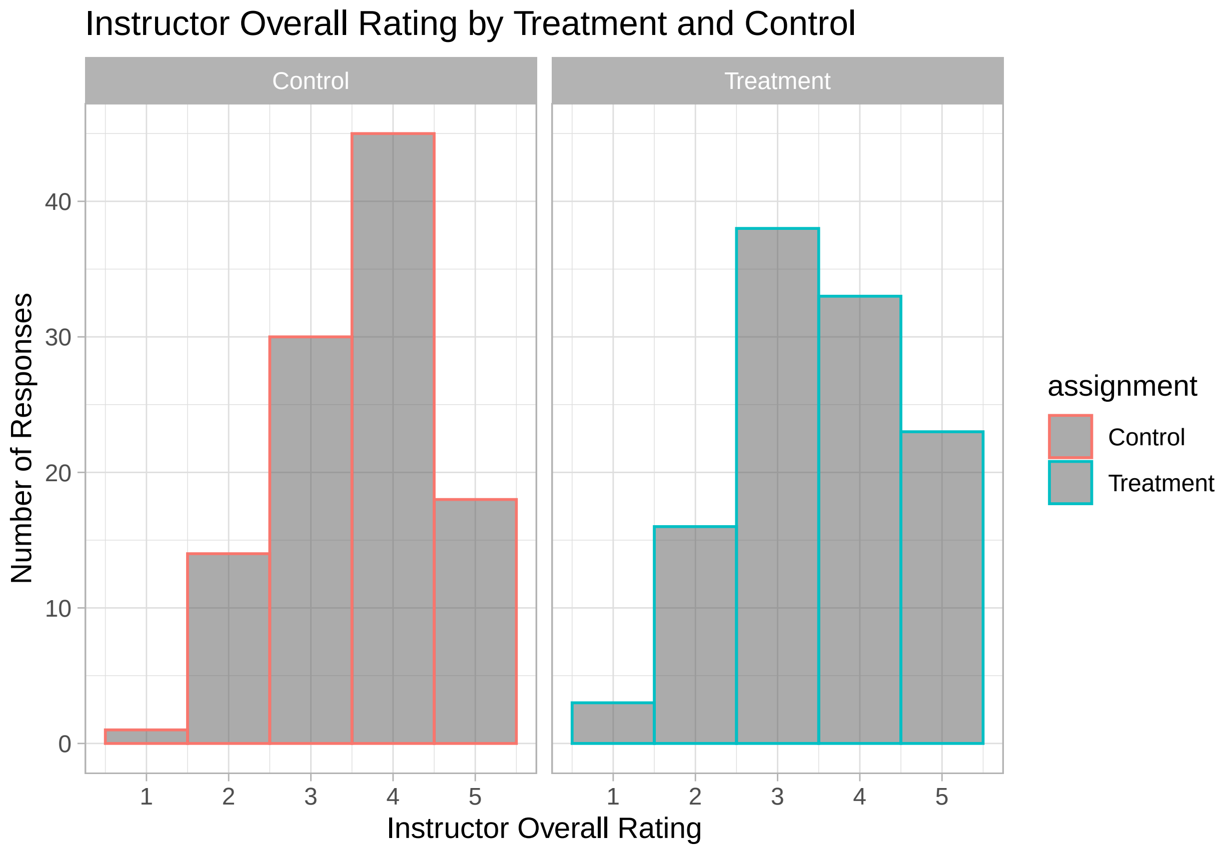
Characteristic	**Overall**, N = 221	**Female**, N = 142	**Male**, N = 79
assignment			
Control	108 (49%)	68 (48%)	40 (51%)
Treatment	113 (51%)	74 (52%)	39 (49%)
Age			
20-30	44 (20%)	25 (18%)	19 (24%)
30-40	82 (37%)	56 (39%)	26 (33%)
40-50	48 (22%)	32 (23%)	16 (20%)
50+	47 (21%)	29 (20%)	18 (23%)

```
d %>%
  select(Gender, Education, assignment) %>%
  tbl_summary(by = assignment) %>%
  add_overall()
```

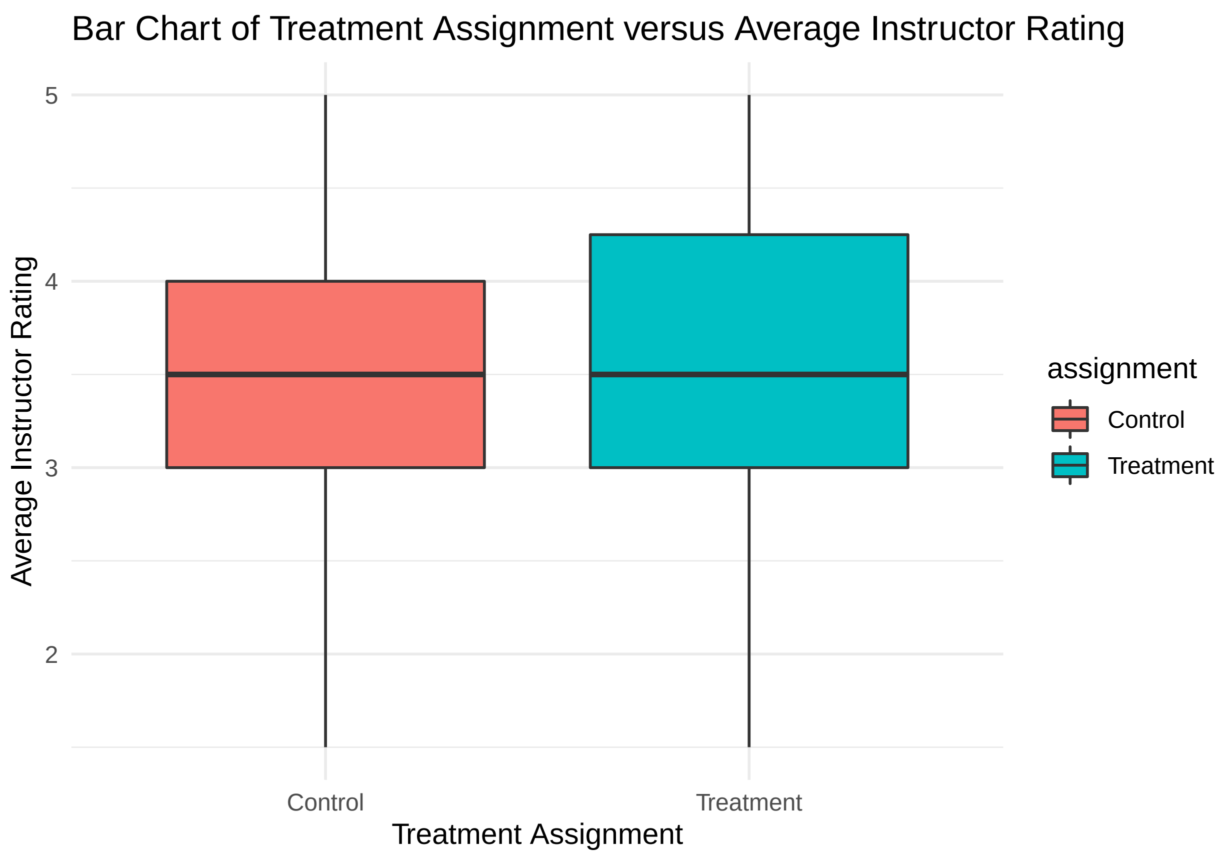
```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**Overall**, N = 221	**Control**, N = 108	**Treatment**, N = 113
Gender			
Female	142 (64%)	68 (63%)	74 (65%)
Male	79 (36%)	40 (37%)	39 (35%)
Education			
Associates degree	27 (12%)	16 (15%)	11 (9.7%)
Bachelors degree	43 (19%)	21 (19%)	22 (19%)
High school diploma	66 (30%)	33 (31%)	33 (29%)
Less than High school	7 (3.2%)	1 (0.9%)	6 (5.3%)
Masters degree	11 (5.0%)	5 (4.6%)	6 (5.3%)
Some College No degree	67 (30%)	32 (30%)	35 (31%)

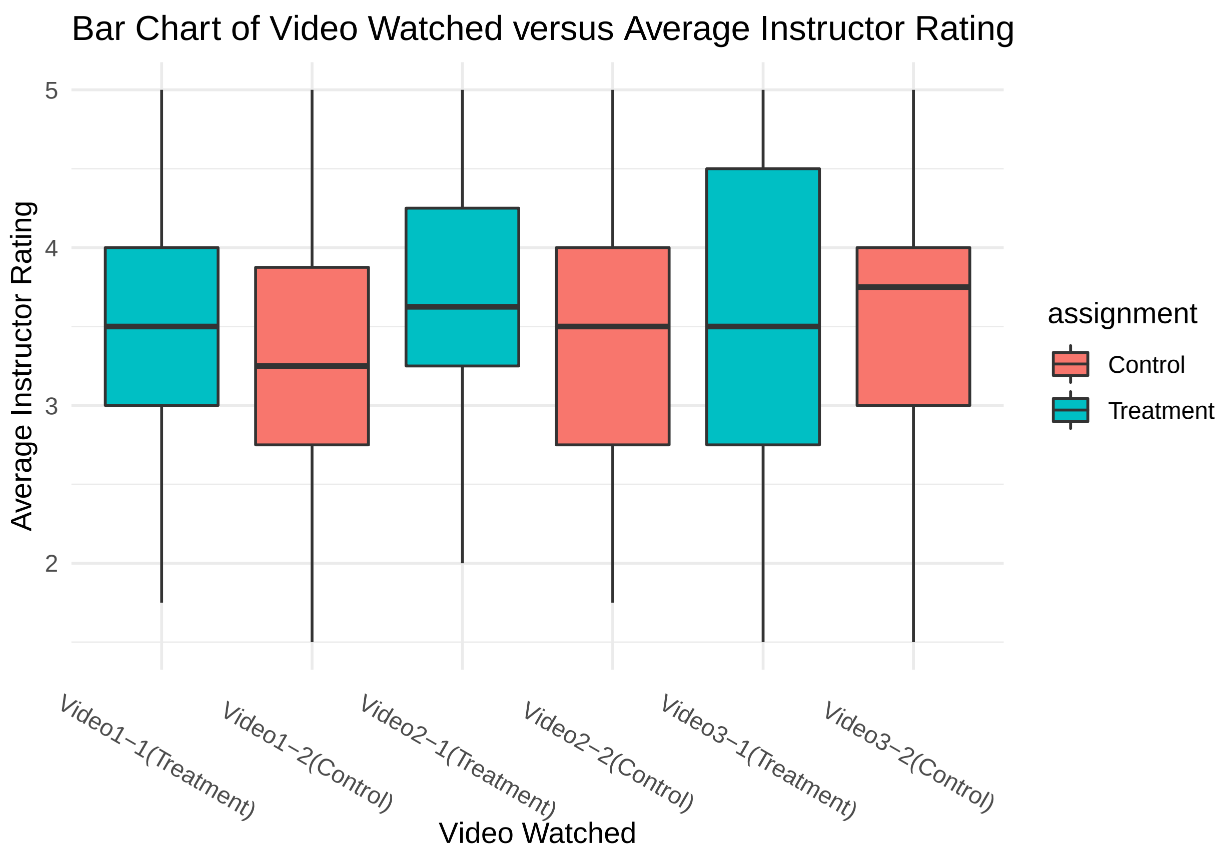
```
response_hist <- d %>%
  ggplot() +
  aes(x = instr_rating, color=assignment) +
  geom_histogram(alpha = 0.5, bins=5) +
  labs(
    title  = 'Instructor Overall Rating by Treatment and Control',
    x      = 'Instructor Overall Rating',
    y      = 'Number of Responses'
  ) +
  theme_light() +
  facet_wrap(~ assignment)

response_hist
```



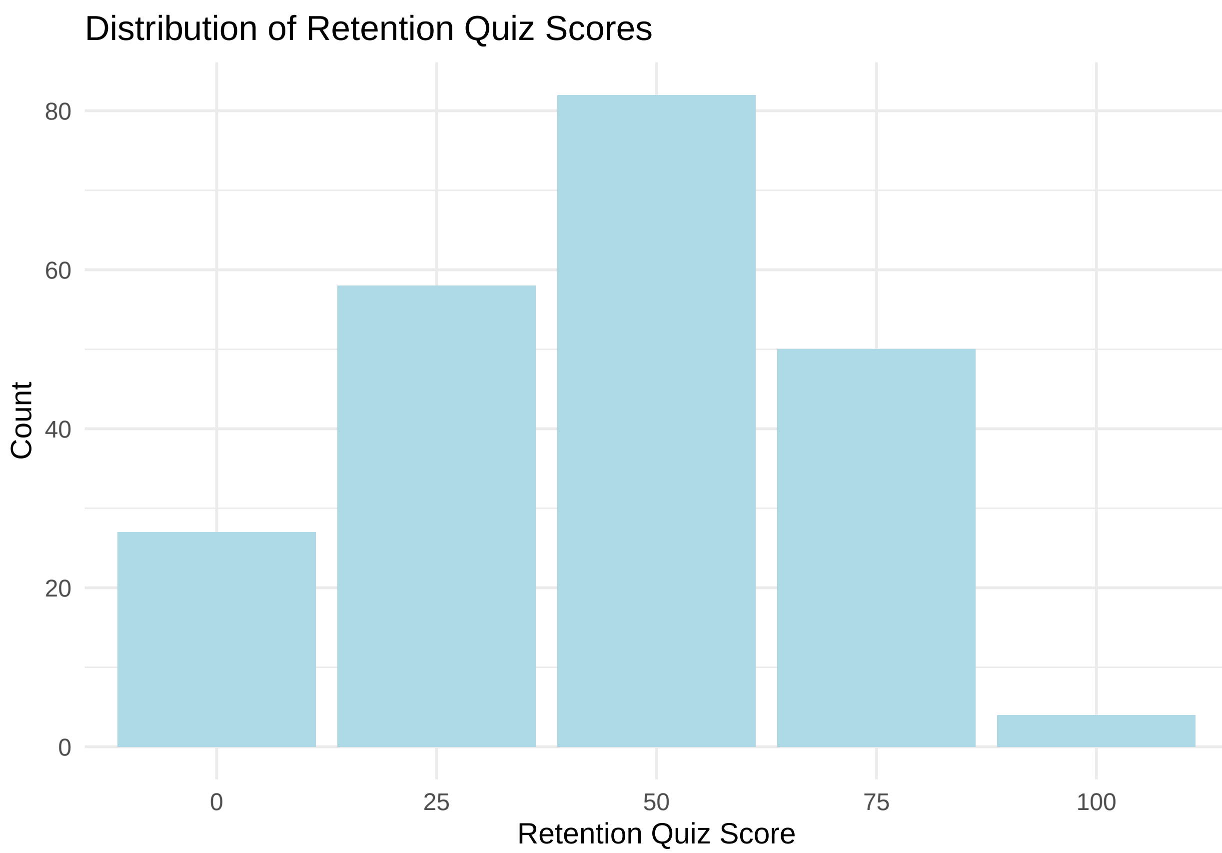

```
ggplot(d, aes(x=assignment, y=avg_rating, fill=assignment))+
  geom_boxplot()+
  labs(title="Bar Chart of Treatment Assignment versus Average Instructor Rating",
       x ="Treatment Assignment", y = "Average Instructor Rating")
```




```
ggplot(d, aes(x=video_watched, y=avg_rating, fill=assignment))+
  geom_boxplot()+
  labs(title="Bar Chart of Video Watched versus Average Instructor Rating",
        x ="Video Watched", y = "Average Instructor Rating")+theme(axis.text.x = element_text(angle = -3
```




```
d$quiz_grade <- as.factor(d$quiz_avg*100)
ggplot(d, aes(x=quiz_grade))+
geom_bar(aes(y = (..count..)),fill="lightblue")+
labs(title="Distribution of Retention Quiz Scores",
        x ="Retention Quiz Score", y = "Count")
```



Randomness Check

In this section, we verify if the block randomization was successful; there should be no increase in predicting whether or not someone is in the treatment or control group based on their blocks.

```
null_model <- lm(assigned~1,data=d)
full_model <- lm(assigned ~ 1 + as.factor(Age)+as.factor(Gender), data=d)
f_test <- anova(full_model, null_model, test ='F')
f_test
```

```
## Analysis of Variance Table
##
## Model 1: assigned ~ 1 + as.factor(Age) + as.factor(Gender)
## Model 2: assigned ~ 1
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    216 54.797
## 2    220 55.222 -4  -0.42483 0.4186 0.7951
```

```
covs <- c('Age','Gender','Education', 'assignment','avg_rating')
subset <- d[,..covs]

head(subset)
```

```
##       Age Gender              Education assignment avg_rating
```

```
## 1: 40-50 Female      Associates degree    Control       4.00
## 2: 30-40 Female Some College No degree    Control       2.75
## 3: 30-40 Female Some College No degree  Treatment       2.00
## 4: 30-40 Female       Bachelors degree    Control       2.25
## 5: 40-50 Female      Associates degree    Control       4.00
## 6: 30-40 Female Some College No degree    Control       3.00
```

```
options(qwraps2_markup = 'markdown')
covariate_summary <-
  list("Gender" =
       list("Male"      = ~ qwraps2::n_perc(Gender == 'Male'),
            "Female"    = ~ qwraps2::n_perc(Gender == 'Female')),
      "Age" =
       list("20-30" = ~ qwraps2::n_perc(Age == '20-30'),
            "30-40"  = ~ qwraps2::n_perc(Age == '30-40'),
            "40-50"  = ~ qwraps2::n_perc(Age == '40-50'),
            "50+" = ~ qwraps2::n_perc(Age == '50+')),
       "Education" =
       list("Less than High school" = ~ qwraps2::n_perc(Education == 'Less than High school'),
            "High school diploma"  = ~ qwraps2::n_perc(Education == 'High school diploma'),
            "Some College No degree"  = ~ qwraps2::n_perc(Education == 'Some College No degree'),
            "Associates degree"  = ~ qwraps2::n_perc(Education == 'Associates degree'),
            "Bachelors degree"  = ~ qwraps2::n_perc(Education == 'Bachelors degree'),
            "Masters degree"  = ~ qwraps2::n_perc(Education == 'Masters degree'))
       )

table <- summary_table(subset, covariate_summary, by = c("assignment"))


table
```

```
##
##
## |                                   |Control (N = 108) |Treatment (N = 113) |
## |:----------------------------------|:-----------------|:-------------------|
## |**Gender**                         |        |          |
## |   Male                  |40 (37.04%)       |39 (34.51%)         |
## |   Female                |68 (62.96%)       |74 (65.49%)         |
## |**Age**                            |        |          |
## |   20-30                 |23 (21.30%)       |21 (18.58%)         |
## |   30-40                 |39 (36.11%)       |43 (38.05%)         |
## |   40-50                 |26 (24.07%)       |22 (19.47%)         |
## |   50+                   |20 (18.52%)       |27 (23.89%)         |
## |**Education**                      |        |          |
## |   Less than High school |1 (0.93%)         |6 (5.31%)           |
## |   High school diploma   |33 (30.56%)       |33 (29.20%)         |
## |   Some College No degree |32 (29.63%)       |35 (30.97%)         |
## |   Associates degree     |16 (14.81%)       |11 (9.73%)          |
## |   Bachelors degree      |21 (19.44%)       |22 (19.47%)         |
## |   Masters degree        |5 (4.63%)         |6 (5.31%)           |
```

```
unique(d$Education)
```

```
## [1] "Associates degree"      "Some College No degree" "Bachelors degree"
```

```
## [4] "Less than High school"  "High school diploma"    "Masters degree"
```

```
ages <- unique(d[,Age])
p_values_ages <- data.table(Age = character(), mean_rating_control = numeric(), mean_rating_treatment =



for(age in ages)

{ print(age)

treatment_avg = round(mean(d[(Age==age)&(assignment=='Treatment'),instr_rating]),2)
control_avg = round(mean(d[(Age==age)&(assignment=='Control'),instr_rating]),2)
p_val =round(t.test(d[(Age==age)&(assignment=='Treatment'),instr_rating], d[(Age==age)&(assignment=='Co
table1 <-data.table(Age=age,mean_rating_control = control_avg, mean_rating_treatment = treatment_avg, p_
p_values_ages<-rbind(p_values_ages, table1)
}
```

```
## [1] "40-50"
## [1] "30-40"
## [1] "50+"
## [1] "20-30"
```

```
p_values_ages <- unique(p_values_ages)
```

```
t.test(d[(Age==age)&(assignment=='Treatment'),instr_rating], d[(Age==age)&(assignment=='Control'),instr_
```

```
##
##  Welch Two Sample t-test
##
## data:  d[(Age == age) & (assignment == "Treatment"), instr_rating] and d[(Age == age) & (assignment =
## t = 0.16838, df = 40.218, p-value = 0.8671
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.5921768  0.6998373
## sample estimates:
## mean of x mean of y
##  3.619048  3.565217
```

```
print(age)
```

```
## [1] "20-30"
```

```
genders <- unique(d[,Gender])
p_values_genders<- data.table(Gender = character(), mean_rating_control = numeric(), mean_rating_treatme



for(gender in genders)

{ print(gender)
```

```r
treatment_avg =round(mean(d[(Gender==gender)&(assignment=='Treatment'),avg_rating]),2)
control_avg =round(mean(d[(Gender==gender)&(assignment=='Control'),avg_rating]),2)
p_val =round(t.test(d[(Gender==gender)&(assignment=='Treatment'),avg_rating], d[(Gender==gender)&(assig
table1 <-data.table(Gender=gender,mean_rating_control = control_avg, mean_rating_treatment = treatment_a
p_values_genders<-rbind(p_values_genders, table1)
}
```

```
## [1] "Female"
## [1] "Male"
```

```r
educations <- unique(d[,Education])
p_values_educations<- data.table(Education = character(), mean_rating_control = numeric(), mean_rating_t




for(education in educations)

{ print(education)
treatment_avg =round(mean(d[(Education==education)&(assignment=='Treatment'),avg_rating]),2)
control_avg =round(mean(d[(Education==education)&(assignment=='Control'),avg_rating]),2)
if (education!= 'Less than High school') {
p_val =round(t.test(d[(Education==education)&(assignment=='Treatment'),avg_rating], d[(Education==educat
else {p_val = NA}
table1 <-data.table(Education=education,mean_rating_control = control_avg, mean_rating_treatment = treat
p_values_educations<-rbind(p_values_educations, table1)
}
```

```
## [1] "Associates degree"
## [1] "Some College No degree"
## [1] "Bachelors degree"
## [1] "Less than High school"
## [1] "High school diploma"
## [1] "Masters degree"
```

```r
p_values_all <- left_join(subset, p_values_ages, by="Age")
p_values_all <- left_join(p_values_all, p_values_educations, by="Education")

p_values_all <- left_join(p_values_all, p_values_genders, by="Gender")

head(p_values_all)
```

```
##      Age Gender              Education assignment avg_rating
## 1: 40-50 Female      Associates degree    Control       4.00
## 2: 30-40 Female Some College No degree    Control       2.75
## 3: 30-40 Female Some College No degree  Treatment       2.00
## 4: 30-40 Female       Bachelors degree    Control       2.25
## 5: 40-50 Female      Associates degree    Control       4.00
## 6: 30-40 Female Some College No degree    Control       3.00
##    mean_rating_control.x mean_rating_treatment.x p_value.x
## 1:                  3.73                    3.59     0.638
## 2:                  3.46                    3.44     0.933
```

```
## 3:                 3.46                   3.44     0.933
## 4:                 3.46                   3.44     0.933
## 5:                 3.73                   3.59     0.638
## 6:                 3.46                   3.44     0.933
##    mean_rating_control.y mean_rating_treatment.y p_value.y mean_rating_control
## 1:                 3.52                   3.36     0.720                3.38
## 2:                 3.23                   3.43     0.285                3.38
## 3:                 3.23                   3.43     0.285                3.38
## 4:                 3.50                   3.82     0.259                3.38
## 5:                 3.52                   3.36     0.720                3.38
## 6:                 3.23                   3.43     0.285                3.38
##    mean_rating_treatment p_value
## 1:                 3.47   0.535
## 2:                 3.47   0.535
## 3:                 3.47   0.535
## 4:                 3.47   0.535
## 5:                 3.47   0.535
## 6:                 3.47   0.535
```

```
options(qwraps2_frmtp_digits= 2)

covariate_summary2<-
  list("Gender" =
       list("Male"      = ~ max(as.numeric(ifelse(Gender == 'Male',mean_rating_control,0))),
            "Female"    = ~ max(as.numeric(ifelse(Gender == 'Female',mean_rating_control,0)))),
       "Age" =
       list("20-30" = ~ max(as.numeric(ifelse(Age == '20-30',mean_rating_control.x,0))),
            "30-40"  = ~ max(as.numeric(ifelse(Age == '30-40',mean_rating_control.x,0))),
            "40-50"  = ~ max(as.numeric(ifelse(Age == '40-50',mean_rating_control.x,0))),
            "50+" = ~ max(as.numeric(ifelse(Age == '50+',mean_rating_control.x,0)))),
       "Education" =
       list("Less than High school" = ~ max(as.numeric(ifelse(Education == 'Less than High school',mean_
            "High school diploma"  = ~ max(as.numeric(ifelse(Education == 'High school diploma',mean_rat
            "Some College No degree"  = ~ max(as.numeric(ifelse(Education == 'Some College No degree',me
            "Associates degree"  = ~ max(as.numeric(ifelse(Education == 'Associates degree',mean_rating_
            "Bachelors degree"  = ~ max(as.numeric(ifelse(Education == 'Bachelors degree',mean_rating_co
            "Masters degree"  = ~ max(as.numeric(ifelse(Education == 'Masters degree',mean_rating_contro
       )

table2<- summary_table(p_values_all , covariate_summary2)


covariate_summary3<-
  list("Gender" =
       list("Male"      = ~ max(as.numeric(ifelse(Gender == 'Male',mean_rating_treatment,0))),
            "Female"    = ~ max(as.numeric(ifelse(Gender == 'Female',mean_rating_treatment,0)))),
       "Age" =
       list("20-30" = ~ max(as.numeric(ifelse(Age == '20-30',mean_rating_treatment.x,0))),
            "30-40"  = ~ max(as.numeric(ifelse(Age == '30-40',mean_rating_treatment.x,0))),
            "40-50"  = ~ max(as.numeric(ifelse(Age == '40-50',mean_rating_treatment.x,0))),
            "50+" = ~ max(as.numeric(ifelse(Age == '50+',mean_rating_treatment.x,0)))),
       "Education" =
       list("Less than High school" = ~ max(as.numeric(ifelse(Education == 'Less than High school',mean_
            "High school diploma"  = ~ max(as.numeric(ifelse(Education == 'High school diploma',mean_rat
```

```r
            "Some College No degree"  = ~ max(as.numeric(ifelse(Education == 'Some College No degree',me
            "Associates degree"  = ~ max(as.numeric(ifelse(Education == 'Associates degree',mean_rating_
            "Bachelors degree"  = ~ max(as.numeric(ifelse(Education == 'Bachelors degree',mean_rating_tr
            "Masters degree"  = ~ max(as.numeric(ifelse(Education == 'Masters degree',mean_rating_treatm
       )

table3<- summary_table(p_values_all , covariate_summary3)

covariate_summary4<-
  list("Gender" =
       list("Male"      = ~ max(as.numeric(ifelse(Gender == 'Male',p_value,0))),
            "Female"    = ~ max(as.numeric(ifelse(Gender == 'Female',p_value,0)))),
       "Age" =
       list("20-30" = ~ max(as.numeric(ifelse(Age == '20-30',p_value.x,0))),
            "30-40"  = ~ max(as.numeric(ifelse(Age == '30-40',p_value.x,0))),
            "40-50"  = ~ max(as.numeric(ifelse(Age == '40-50',p_value.x,0))),
            "50+" = ~ max(as.numeric(ifelse(Age == '50+',p_value.x,0)))),
       "Education" =
       list("Less than High school" = ~ max(as.numeric(ifelse(Education == 'Less than High school',p_val
            "High school diploma"  = ~ max(as.numeric(ifelse(Education == 'High school diploma',p_value
            "Some College No degree"  = ~ max(as.numeric(ifelse(Education == 'Some College No degree',p_
            "Associates degree"  = ~ max(as.numeric(ifelse(Education == 'Associates degree',p_value.y,0)
            "Bachelors degree"  = ~ max(as.numeric(ifelse(Education == 'Bachelors degree',p_value.y,0)))
            "Masters degree"  = ~ max(as.numeric(ifelse(Education == 'Masters degree',p_value.y,0))))
       )

options(digits=3)
table4<- summary_table(p_values_all , covariate_summary4)
final <-cbind(table, table2,table3,table4)
print(final, cname = c("Control (N = 108)","Treatment (N = 113)", "Mean Rating - Control", "Mean Rating
```

```
##
##
## |                                    |Control (N = 108) |Treatment (N = 113) |Mean Rating - Control
## |:-----------------------------------|:-----------------|:-------------------|:--------------------
## |**Gender**                          |        |          |  
## |   Male                   |40 (37.04%)       |39 (34.51%)         |3.56
## |   Female                 |68 (62.96%)       |74 (65.49%)         |3.38
## |**Age**                             |        |          |  
## |   20-30                  |23 (21.30%)       |21 (18.58%)         |3.57
## |   30-40                  |39 (36.11%)       |43 (38.05%)         |3.46
## |   40-50                  |26 (24.07%)       |22 (19.47%)         |3.73
## |   50+                    |20 (18.52%)       |27 (23.89%)         |3.75
## |**Education**                       |        |          |  
## |   Less than High school  |1 (0.93%)         |6 (5.31%)           |3.25
## |   High school diploma    |33 (30.56%)       |33 (29.20%)         |3.58
## |   Some College No degree |32 (29.63%)       |35 (30.97%)         |3.23
## |   Associates degree      |16 (14.81%)       |11 (9.73%)          |3.52
## |   Bachelors degree       |21 (19.44%)       |22 (19.47%)         |3.5
## |   Masters degree         |5 (4.63%)         |6 (5.31%)           |3.55
```

```r
# Table 1: Baseline models Outcome ~ Treatment
model_1_1 <- lm(avg_rating ~ assignment, data=d)
model_1_2 <- lm(instr_rating ~ assignment, data=d)
model_1_3 <- lm(instr_subject ~ assignment, data=d)
model_1_4 <- lm(instr_material ~ assignment, data=d)
model_1_5 <- lm(instr_enthusiasm  ~ assignment, data=d)

# Table 1
stargazer(model_1_1,model_1_2, model_1_3, model_1_4, model_1_5, type='text')
```

```
##
## ==================================================================================================
##                                                   Dependent variable:
##                                   ----------------------------------------------------------------
##                                   avg_rating instr_rating instr_subject instr_material instr_enthusiasm
##                                      (1)         (2)           (3)           (4)             (5)
## --------------------------------------------------------------------------------------------------
## assignmentTreatment                0.122       -0.097         0.068         0.176          0.364**
##                                   (0.114)     (0.135)       (0.122)       (0.137)        (0.152)
##
## Constant                          3.450***    3.600***      3.860***      3.630***       2.780***
##                                   (0.082)     (0.096)       (0.087)       (0.098)        (0.109)
##
## --------------------------------------------------------------------------------------------------
## Observations                        221         221           221           221            221
## R2                                 0.005       0.002         0.001         0.007          0.025
## Adjusted R2                        0.001       -0.002        -0.003        0.003          0.021
## Residual Std. Error (df = 219)     0.849       1.000         0.907         1.020          1.130
## F Statistic (df = 1; 219)          1.140       0.521         0.311         1.640          5.710**
## ==================================================================================================
## Note:                                                              *p<0.1; **p<0.05; ***p<0.01
```

```r
# Table 2: Outcome with Pre-treatment blocks (Age and Gender)
model_2_1 <- lm(avg_rating ~ as.factor(Age) + as.factor(Gender)+assignment, data=d)
model_2_2 <- lm(instr_rating ~ as.factor(Age) + as.factor(Gender)+ assignment, data=d)
model_2_3 <- lm(instr_subject ~ as.factor(Age) + as.factor(Gender)+assignment, data=d)
model_2_4 <- lm(instr_material ~ as.factor(Age) + as.factor(Gender)+assignment, data=d)
model_2_5 <- lm(instr_enthusiasm ~ as.factor(Age) + as.factor(Gender)+assignment , data=d)

# Table 2
stargazer(model_2_1,model_2_2, model_2_3, model_2_4, model_2_5, type='text')
```

```
##
## ==================================================================================================
##                                                   Dependent variable:
##                                   ----------------------------------------------------------------
##                                   avg_rating instr_rating instr_subject instr_material instr_enthusiasm
##                                      (1)         (2)           (3)           (4)             (5)
## --------------------------------------------------------------------------------------------------
## as.factor(Age)30-40                -0.118      -0.109        -0.212        -0.098          0.043
##                                   (0.159)     (0.188)       (0.170)       (0.192)        (0.208)
##
```

```
## as.factor(Age)40-50              -0.094     0.097      -0.049      0.002        -0.306
##                                  (0.177)    (0.210)     (0.190)     (0.213)      (0.232)
##
## as.factor(Age)50+                 -0.131     0.003       0.064      -0.035       -0.144
##                                  (0.178)    (0.211)     (0.191)     (0.215)      (0.233)
##
## as.factor(Gender)Male             0.229*     0.237*      0.108       0.220        0.479***
##                                  (0.120)    (0.141)     (0.128)     (0.144)      (0.156)
##
## assignmentTreatment               0.133     -0.085       0.069       0.185        0.369**
##                                  (0.115)    (0.135)     (0.123)     (0.138)      (0.150)
##
## Constant                         3.450***   3.530***    3.900***    3.590***     2.680***
##                                  (0.149)    (0.176)     (0.159)     (0.179)      (0.195)
##
## ----------------------------------------------------------------------------------------------------
## Observations                       221        221         221         221          221
## R2                                0.026      0.022       0.021       0.021        0.081
## Adjusted R2                       0.003     -0.001      -0.002      -0.002        0.059
## Residual Std. Error (df = 215)    0.848      1.000       0.907       1.020        1.110
## F Statistic (df = 5; 215)         1.150      0.962       0.913       0.902        3.770***
## ====================================================================================================
## Note:                                                            *p<0.1; **p<0.05; ***p<0.01
```

```r
# Table 3: Outcome with Pre-treatment blocks (Age and Gender) and interaction term (gender*treatment)
model_3_1 <- lm(avg_rating ~ as.factor(Age) + as.factor(Gender)+assignment +as.factor(Gender)*assignmen
model_3_2 <- lm(instr_rating ~ as.factor(Age) + as.factor(Gender)+ assignment +as.factor(Gender)*assign
model_3_3 <- lm(instr_subject ~ as.factor(Age) + as.factor(Gender)+assignment+as.factor(Gender)*assignm
model_3_4 <- lm(instr_material ~ as.factor(Age) + as.factor(Gender)+assignment+as.factor(Gender)*assign
model_3_5 <- lm(instr_enthusiasm ~ as.factor(Age) + as.factor(Gender)+assignment+as.factor(Gender)*assi


# Table 3
stargazer(model_3_1,model_3_2, model_3_3, model_3_4, model_3_5, type='text')
```

```
##
## ====================================================================================================
##                                                        Dependent variable:
##                                         ------------------------------------------------------------
##                                         avg_rating instr_rating instr_subject instr_material instr_
##                                            (1)          (2)          (3)           (4)
## ----------------------------------------------------------------------------------------------------
## as.factor(Age)30-40                       -0.116       -0.108       -0.210        -0.092
##                                          (0.160)      (0.189)      (0.171)       (0.191)        (
##
## as.factor(Age)40-50                       -0.087        0.100       -0.040         0.029
##                                          (0.179)      (0.211)      (0.191)       (0.214)        (
##
## as.factor(Age)50+                         -0.129        0.004        0.067        -0.026
##                                          (0.179)      (0.211)      (0.191)       (0.214)        (
##
## as.factor(Gender)Male                      0.180        0.219        0.042         0.023         0
##                                          (0.170)      (0.201)      (0.181)       (0.203)        (
##
```

```
## assignmentTreatment                          0.098       -0.097       0.022        0.045        0
##                                             (0.143)     (0.169)      (0.153)      (0.172)      (
##
## as.factor(Gender)Male:assignmentTreatment    0.097       0.035        0.132        0.393
##                                             (0.240)     (0.283)      (0.256)      (0.287)      (
##
## Constant                                    3.470***    3.540***     3.920***     3.650***     2
##                                             (0.154)     (0.182)      (0.165)      (0.185)      (
##
## ------------------------------------------------------------------------------------------------
## Observations                                  221         221          221          221
## R2                                           0.027       0.022        0.022        0.029        0
## Adjusted R2                                 -0.001      -0.005       -0.005        0.002        0
## Residual Std. Error (df = 214)               0.850       1.000        0.909        1.020        1
## F Statistic (df = 6; 214)                    0.981       0.800        0.802        1.070        3
## ================================================================================================
## Note:                                                                          *p<0.1; **p<0.05
```

```r
# Table 4:

# base model
model_4_1 <- lm(quiz_avg~ assignment, data=d)
model_4_2 <- lm(quiz_avg ~ as.factor(Age) + as.factor(Gender)+assignment, data=d)
model_4_3 <- lm(quiz_avg ~ as.factor(Age) + as.factor(Gender)+assignment +as.factor(Gender)*assignment,


#models with blocking across quiz averages
stargazer(model_4_1, model_4_2,model_4_3, type = 'text')
```

```
##
## ================================================================================================
##                                                            Dependent variable:
##                                              ---------------------------------------------------
##                                                                 quiz_avg
##                                                     (1)              (2)                 (3)
## ------------------------------------------------------------------------------------------------
## as.factor(Age)30-40                                                -0.032              -0.032
##                                                                    (0.046)             (0.046)
##
## as.factor(Age)40-50                                                 0.032               0.031
##                                                                    (0.051)             (0.052)
##
## as.factor(Age)50+                                                   0.095*              0.095*
##                                                                    (0.052)             (0.052)
##
## as.factor(Gender)Male                                              -0.052              -0.045
##                                                                    (0.035)             (0.049)
##
## assignmentTreatment                                -0.020          -0.024              -0.019
##                                                    (0.034)         (0.033)             (0.042)
##
## as.factor(Gender)Male:assignmentTreatment                                              -0.013
##                                                                                        (0.070)
##
```

```
## Constant                                  0.449***            0.454***            0.452***
##                                           (0.024)             (0.043)             (0.045)
##
## --------------------------------------------------------------------------------------------
## Observations                                221                 221                 221
## R2                                         0.002               0.047               0.048
## Adjusted R2                               -0.003               0.025               0.021
## Residual Std. Error                  0.250 (df = 219)    0.246 (df = 215)    0.247 (df = 214)
## F Statistic                     0.350 (df = 1; 219) 2.140* (df = 5; 215) 1.780 (df = 6; 21
## ============================================================================================
## Note:                                                          *p<0.1; **p<0.05; ***p<0.
```