RecallResultsAnalysis

```
#setwd("MIDS/DATASCI_W241/Assignments/Project/")

load(file = "CourseProjectCleanedData.rda")

## Load relevant libraries
library(pryr)
library(memisc)
```

```
## Loading required package: lattice
## Loading required package: MASS
##
## Attaching package: 'memisc'
##
## The following objects are masked from 'package:stats':
##
##     contr.sum, contr.treatment, contrasts
##
## The following objects are masked from 'package:base':
##
##     as.array, trimws
```

```
library(stringr)
library(stargazer)
```

```
##
## Please cite as:
##
##  Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##  R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:memisc':
##
##     collect, query, rename
##
## The following object is masked from 'package:MASS':
##
##     select
##
## The following object is masked from 'package:pryr':
##
##     %.%
##
## The following objects are masked from 'package:stats':
```

```
##
##     filter, lag
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(data.table)
```

```
##
## Attaching package: 'data.table'
##
## The following objects are masked from 'package:dplyr':
##
##     between, last
##
## The following object is masked from 'package:pryr':
##
##     address
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
##
## The following objects are masked from 'package:data.table':
##
##     hour, mday, month, quarter, wday, week, yday, year
##
## The following object is masked from 'package:memisc':
##
##     is.interval
```

```
library(RDSTK)
```

```
## Loading required package: plyr
## ------------------------------------------------------------------------------
## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)
## ------------------------------------------------------------------------------
##
## Attaching package: 'plyr'
##
## The following object is masked from 'package:lubridate':
##
##     here
##
## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
```

```
##     summarize
##
## The following object is masked from 'package:memisc':
##
##     rename
##
## Loading required package: rjson
## Loading required package: RCurl
## Loading required package: bitops
```

```
d<- pd.cleaned %>%
  filter(recall_Invalid == 0)


## preliminary analysis
table(d$recall_PercentageCorrect)
```

```
##
##   0 0.5   1
##  37 131 229
```

```
## review recall means of percentage correct by exp1_Treatment group
d %>%
  dplyr::group_by(exp1_Treatment) %>%
  dplyr::summarise(outcome = mean(recall_PercentageCorrect))
```

```
## Source: local data frame [6 x 2]
##
##          exp1_Treatment   outcome
##                   (fctr)     (dbl)
## 1      Nonordered table 0.6893939
## 2 Nonordered chart junk 0.7134146
## 3      Nonordered tufte 0.8333333
## 4    Ordered chart junk 0.7745098
## 5         Ordered Table 0.7622951
## 6         Ordered tufte 0.6911765
```

```
## code dummy variable to look @ performance between "tufte optimized" vs. non as well as ordered
d <- d %>%
  dplyr::mutate(exp1_TreatmentRecallRollupTufte =
                cases("Not Optimized"=exp1_Treatment %in% c("Nonordered chart junk"
                                                            ,"Ordered chart junk"
                                                            ,"Ordered Table"
                                                            ,"Nonordered table"),
                    "Tufte Optimized"=exp1_Treatment %in% c("Nonordered tufte", "Ordered tufte"))
                ,exp1_TreatmentRecallRollupOrdered =
                  cases("Not Ordered"=exp1_Treatment %in% c("Nonordered chart junk"
                                                            ,"Nonordered tufte"
                                                            ,"Nonordered table"),
                        "Ordered"=exp1_Treatment %in% c("Ordered chart junk", "Ordered tufte","Ordered T
                      )
```

```
## review recall means of percentage correct by exp1_Treatment group dummy (when tufte treatments are c
d %>%
  dplyr::group_by(exp1_TreatmentRecallRollupTufte) %>%
  dplyr::summarise(outcome = mean(recall_PercentageCorrect))
```

```
## Source: local data frame [2 x 2]
##
##   exp1_TreatmentRecallRollupTufte   outcome
##                             (fctr)     (dbl)
## 1                    Not Optimized 0.7307692
## 2                  Tufte Optimized 0.7627737
```

```
## review recall means of percentage correct by exp1_Treatment group dummy (when non ordered treatments
d %>%
  dplyr::group_by(exp1_TreatmentRecallRollupOrdered) %>%
  dplyr::summarise(outcome = mean(recall_PercentageCorrect))
```

```
## Source: local data frame [2 x 2]
##
##   exp1_TreatmentRecallRollupOrdered   outcome
##                               (fctr)     (dbl)
## 1                        Not Ordered 0.7442396
## 2                            Ordered 0.7388889
```

```
#Regress recall_PercentageCorrect on various treatment variables

recall.treat <- lm(recall_PercentageCorrect ~ exp1_Treatment, d)

recall.ordertreat <- lm(recall_PercentageCorrect ~ exp1_TreatmentRecallRollupOrdered, d)

recall.tuftetreat <- lm(recall_PercentageCorrect ~ exp1_TreatmentRecallRollupTufte, d)

#Compare results

stargazer(recall.treat, recall.ordertreat, recall.tuftetreat, type = "text")
```

```
##
## ==========================================================================================================
##                                                                  Dependent variable:
##                                                ----------------------------------------------------------
##                                                                recall_PercentageCorrect
##                                                       (1)                   (2)                  (3)
## ----------------------------------------------------------------------------------------------------------
## exp1_TreatmentNonordered chart junk                  0.024
##                                                     (0.054)
##
## exp1_TreatmentNonordered tufte                      0.144**
##                                                     (0.057)
##
## exp1_TreatmentOrdered chart junk                     0.085
##                                                     (0.061)
```

```
##
## exp1_TreatmentOrdered Table                                0.073
##                                                           (0.058)
##
## exp1_TreatmentOrdered tufte                                0.002
##                                                           (0.057)
##
## exp1_TreatmentRecallRollupOrderedOrdered                                  -0.005
##                                                                           (0.033)
##
## exp1_TreatmentRecallRollupTufteTufte Optimized                                                0.032
##                                                                                              (0.035)
##
## Constant                                                  0.689***          0.744***          0.731***
##                                                           (0.040)           (0.022)           (0.021)
##
## ------------------------------------------------------------------------------------------------------
## Observations                                               397               397               397
## R2                                                        0.025             0.0001            0.002
## Adjusted R2                                               0.012             -0.002            -0.0004
## Residual Std. Error                                 0.329 (df = 391)  0.331 (df = 395)  0.331 (df =
## F Statistic                                        2.001* (df = 5; 391) 0.026 (df = 1; 395) 0.840 (df =
## ======================================================================================================
## Note:                                                                        *p<0.1; **p<0.05; ***
```

```
#Regress recall_PercentageCorrect on various treatment variables controlling for various factors

summary(lm(recall_PercentageCorrect ~ exp1_Treatment + recall_PageFiveTimeMinutes, d))
```

```
##
## Call:
## lm(formula = recall_PercentageCorrect ~ exp1_Treatment + recall_PageFiveTimeMinutes,
##     data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.7873 -0.2211  0.1658  0.2798  0.3828
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        0.704745   0.043440  16.223   <2e-16
## exp1_TreatmentNonordered chart junk  0.027666   0.054466   0.508   0.6118
## exp1_TreatmentNonordered tufte       0.144826   0.056580   2.560   0.0109
## exp1_TreatmentOrdered chart junk     0.086140   0.061268   1.406   0.1605
## exp1_TreatmentOrdered Table          0.078739   0.058668   1.342   0.1803
## exp1_TreatmentOrdered tufte          0.004014   0.056822   0.071   0.9437
## recall_PageFiveTimeMinutes          -0.045355   0.046835  -0.968   0.3335
##
## (Intercept)                        ***
## exp1_TreatmentNonordered chart junk
## exp1_TreatmentNonordered tufte      *
## exp1_TreatmentOrdered chart junk
## exp1_TreatmentOrdered Table
## exp1_TreatmentOrdered tufte
```

```
## recall_PageFiveTimeMinutes
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3286 on 390 degrees of freedom
## Multiple R-squared:  0.02729,    Adjusted R-squared:  0.01232
## F-statistic: 1.823 on 6 and 390 DF,  p-value: 0.09329
```

```
summary(lm(recall_PercentageCorrect ~ exp1_Treatment + recall_PageFiveTimeMinutes + gen_Gender + gen_Ag
```

```
##
## Call:
## lm(formula = recall_PercentageCorrect ~ exp1_Treatment + recall_PageFiveTimeMinutes +
##     gen_Gender + gen_Age + gen_Education + iq_PercentageCorrect,
##     data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8297 -0.2487  0.1381  0.2468  0.5138
##
## Coefficients:
##                                          Estimate Std. Error t value
## (Intercept)                              0.221260   0.233330   0.948
## exp1_TreatmentNonordered chart junk      0.035508   0.054193   0.655
## exp1_TreatmentNonordered tufte           0.128261   0.056322   2.277
## exp1_TreatmentOrdered chart junk         0.082656   0.060899   1.357
## exp1_TreatmentOrdered Table              0.094211   0.058406   1.613
## exp1_TreatmentOrdered tufte              0.001934   0.056061   0.034
## recall_PageFiveTimeMinutes              -0.049922   0.046285  -1.079
## gen_GenderFemale                         0.003214   0.033388   0.096
## gen_GenderPrefer not to respond         -0.189265   0.325960  -0.581
## gen_Age                                  0.001817   0.001360   1.336
## gen_EducationHigh School / GED           0.158537   0.239939   0.661
## gen_EducationSome College                0.194755   0.237181   0.821
## gen_Education2-year College Degree       0.300675   0.239714   1.254
## gen_Education4-year College Degree       0.243421   0.236908   1.027
## gen_EducationMasters Degree              0.251497   0.241287   1.042
## gen_EducationDoctoral Degree             0.312257   0.263498   1.185
## gen_EducationProfessional Degree (JD, MD)  0.184946   0.287888   0.642
## iq_PercentageCorrect                     0.245294   0.066497   3.689
##                                         Pr(>|t|)
## (Intercept)                             0.343597
## exp1_TreatmentNonordered chart junk     0.512730
## exp1_TreatmentNonordered tufte          0.023325 *
## exp1_TreatmentOrdered chart junk        0.175506
## exp1_TreatmentOrdered Table             0.107566
## exp1_TreatmentOrdered tufte             0.972504
## recall_PageFiveTimeMinutes              0.281464
## gen_GenderFemale                        0.923355
## gen_GenderPrefer not to respond         0.561829
## gen_Age                                 0.182307
## gen_EducationHigh School / GED          0.509182
## gen_EducationSome College               0.412092
## gen_Education2-year College Degree      0.210502
```

```
## gen_Education4-year College Degree        0.304844
## gen_EducationMasters Degree               0.297931
## gen_EducationDoctoral Degree              0.236742
## gen_EducationProfessional Degree (JD, MD) 0.520987
## iq_PercentageCorrect                      0.000258 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3212 on 379 degrees of freedom
## Multiple R-squared:  0.09641,    Adjusted R-squared:  0.05588
## F-statistic: 2.379 on 17 and 379 DF,  p-value: 0.001691
```

```
summary(lm(recall_PercentageCorrect ~ exp1_TreatmentRecallRollupOrdered + recall_PageFiveTimeMinutes, d
```

```
##
## Call:
## lm(formula = recall_PercentageCorrect ~ exp1_TreatmentRecallRollupOrdered +
##     recall_PageFiveTimeMinutes, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.7542 -0.2454  0.2483  0.2555  0.4083
##
## Coefficients:
##                                          Estimate Std. Error t value
## (Intercept)                              0.761172   0.028537  26.673
## exp1_TreatmentRecallRollupOrderedOrdered -0.003906   0.033409  -0.117
## recall_PageFiveTimeMinutes               -0.045148   0.046895  -0.963
##                                         Pr(>|t|)
## (Intercept)                               <2e-16 ***
## exp1_TreatmentRecallRollupOrderedOrdered   0.907
## recall_PageFiveTimeMinutes                 0.336
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3311 on 394 degrees of freedom
## Multiple R-squared:  0.002412,   Adjusted R-squared:  -0.002652
## F-statistic: 0.4763 on 2 and 394 DF,  p-value: 0.6214
```

```
summary(lm(recall_PercentageCorrect ~ exp1_TreatmentRecallRollupOrdered + recall_PageFiveTimeMinutes + g
```

```
##
## Call:
## lm(formula = recall_PercentageCorrect ~ exp1_TreatmentRecallRollupOrdered +
##     recall_PageFiveTimeMinutes + gen_Gender + gen_Age + gen_Education +
##     iq_PercentageCorrect, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8319 -0.2442  0.1681  0.2505  0.4640
##
## Coefficients:
##                                           Estimate Std. Error t value
```

```
## (Intercept)                                 0.216988   0.233685   0.929
## exp1_TreatmentRecallRollupOrderedOrdered    0.001554   0.032814   0.047
## recall_PageFiveTimeMinutes                 -0.047122   0.046223  -1.019
## gen_GenderFemale                            0.007417   0.033405   0.222
## gen_GenderPrefer not to respond            -0.240982   0.326474  -0.738
## gen_Age                                     0.001913   0.001363   1.403
## gen_EducationHigh School / GED              0.208724   0.240026   0.870
## gen_EducationSome College                   0.244883   0.237031   1.033
## gen_Education2-year College Degree          0.342958   0.239898   1.430
## gen_Education4-year College Degree          0.298714   0.236839   1.261
## gen_EducationMasters Degree                 0.288916   0.241770   1.195
## gen_EducationDoctoral Degree                0.364798   0.263284   1.386
## gen_EducationProfessional Degree (JD, MD)   0.256508   0.286405   0.896
## iq_PercentageCorrect                        0.247788   0.066438   3.730
##                                            Pr(>|t|)
## (Intercept)                                0.353708
## exp1_TreatmentRecallRollupOrderedOrdered   0.962264
## recall_PageFiveTimeMinutes                 0.308629
## gen_GenderFemale                           0.824398
## gen_GenderPrefer not to respond            0.460885
## gen_Age                                    0.161497
## gen_EducationHigh School / GED             0.385071
## gen_EducationSome College                  0.302196
## gen_Education2-year College Degree         0.153647
## gen_Education4-year College Degree         0.207985
## gen_EducationMasters Degree                0.232825
## gen_EducationDoctoral Degree               0.166686
## gen_EducationProfessional Degree (JD, MD) 0.371021
## iq_PercentageCorrect                       0.000221 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3232 on 383 degrees of freedom
## Multiple R-squared:  0.07564,    Adjusted R-squared:  0.04426
## F-statistic: 2.411 on 13 and 383 DF,  p-value: 0.003919
```

```
summary(lm(recall_PercentageCorrect ~ exp1_TreatmentRecallRollupTufte + recall_PageFiveTimeMinutes, d))
```

```
##
## Call:
## lm(formula = recall_PercentageCorrect ~ exp1_TreatmentRecallRollupTufte +
##     recall_PageFiveTimeMinutes, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.7753 -0.2392  0.2319  0.2628  0.4129
##
## Coefficients:
##                                              Estimate Std. Error t value
## (Intercept)                                   0.74829    0.02773  26.987
## exp1_TreatmentRecallRollupTufteTufte Optimized  0.03087    0.03494   0.884
## recall_PageFiveTimeMinutes                   -0.04397    0.04683  -0.939
##                                              Pr(>|t|)
## (Intercept)                                    <2e-16 ***
```

```
## exp1_TreatmentRecallRollupTufteTufte Optimized    0.377
## recall_PageFiveTimeMinutes                        0.348
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3307 on 394 degrees of freedom
## Multiple R-squared:  0.004351,   Adjusted R-squared:  -0.0007033
## F-statistic: 0.8608 on 2 and 394 DF,  p-value: 0.4236
```

```
summary(lm(recall_PercentageCorrect ~ exp1_TreatmentRecallRollupTufte + recall_PageFiveTimeMinutes + ge
```

```
##
## Call:
## lm(formula = recall_PercentageCorrect ~ exp1_TreatmentRecallRollupTufte +
##     recall_PageFiveTimeMinutes + gen_Gender + gen_Age + gen_Education +
##     iq_PercentageCorrect, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8403 -0.2484  0.1662  0.2510  0.4617
##
## Coefficients:
##                                                Estimate Std. Error
## (Intercept)                                    0.209591   0.233747
## exp1_TreatmentRecallRollupTufteTufte Optimized 0.016319   0.034697
## recall_PageFiveTimeMinutes                    -0.046134   0.046197
## gen_GenderFemale                               0.007116   0.033401
## gen_GenderPrefer not to respond               -0.237063   0.326256
## gen_Age                                        0.001895   0.001363
## gen_EducationHigh School / GED                 0.212493   0.240085
## gen_EducationSome College                      0.251260   0.237333
## gen_Education2-year College Degree             0.347763   0.240033
## gen_Education4-year College Degree             0.302868   0.236911
## gen_EducationMasters Degree                    0.294241   0.241913
## gen_EducationDoctoral Degree                   0.371921   0.263576
## gen_EducationProfessional Degree (JD, MD)      0.267911   0.286965
## iq_PercentageCorrect                           0.244981   0.066489
##                                               t value Pr(>|t|)
## (Intercept)                                     0.897 0.370467
## exp1_TreatmentRecallRollupTufteTufte Optimized  0.470 0.638384
## recall_PageFiveTimeMinutes                     -0.999 0.318601
## gen_GenderFemale                                0.213 0.831412
## gen_GenderPrefer not to respond                -0.727 0.467904
## gen_Age                                         1.390 0.165371
## gen_EducationHigh School / GED                  0.885 0.376671
## gen_EducationSome College                       1.059 0.290413
## gen_Education2-year College Degree              1.449 0.148208
## gen_Education4-year College Degree              1.278 0.201882
## gen_EducationMasters Degree                     1.216 0.224617
## gen_EducationDoctoral Degree                    1.411 0.159040
## gen_EducationProfessional Degree (JD, MD)       0.934 0.351098
## iq_PercentageCorrect                            3.685 0.000262 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.3231 on 383 degrees of freedom
## Multiple R-squared:  0.07617,    Adjusted R-squared:  0.04481
## F-statistic: 2.429 on 13 and 383 DF,  p-value: 0.003638
```

```
#########################################################################
#### Swap Variables, look @ time to complete instead of correct score
#########################################################################

## review recall means of time to complete  by exp1_Treatment group
d %>%
  dplyr::group_by(exp1_Treatment) %>%
  dplyr::summarise(outcome = mean(recall_PageFiveTimeMinutes))
```

```
## Source: local data frame [6 x 2]
##
##          exp1_Treatment   outcome
##                  (fctr)     (dbl)
## 1      Nonordered table 0.3384581
## 2 Nonordered chart junk 0.4188370
## 3      Nonordered tufte 0.3580092
## 4    Ordered chart junk 0.3610408
## 5         Ordered Table 0.4671689
## 6         Ordered tufte 0.3876510
```