03_Preliminary_Analyses.R

semarten

2019-12-13

Date created: December 12, 2019 Author: Sheena Martenies Contact: Sheena.Martenies@colostate.edu
Project: MAP-ERC Teacher's Study
Description: Preliminary analysis comparing ORC scores to outcomes reported on the participant survey

---

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.1
```

```
## -- Attaching packages ---------------------------------------------- tidyverse 1.2.1 --
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## Warning: package 'ggplot2' was built under R version 3.6.1
```

```
## Warning: package 'tibble' was built under R version 3.6.1
```

```
## Warning: package 'tidyr' was built under R version 3.6.1
```

```
## Warning: package 'readr' was built under R version 3.6.1
```

```
## Warning: package 'purrr' was built under R version 3.6.1
```

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

```
## Warning: package 'stringr' was built under R version 3.6.1
```

```
## Warning: package 'forcats' was built under R version 3.6.1
```

```
## -- Conflicts ------------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 3.6.1
```

Read in the clean survey and ORC data
Create a single dataset
Drop original survey variables

```
q_data <- read_csv(here::here("Data", "Cleaned_Survey_Data.csv"))
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   StartDate = col_datetime(format = ""),
##   EndDate = col_datetime(format = ""),
##   Status = col_character(),
##   IPAddress = col_character(),
##   Finished = col_logical(),
##   RecordedDate = col_datetime(format = ""),
##   ResponseId = col_character(),
##   ExternalReference = col_logical(),
##   DistributionChannel = col_character(),
##   UserLanguage = col_character(),
##   Q2 = col_character(),
##   Q3 = col_character(),
##   Q3_5_TEXT = col_character(),
##   Q4 = col_character(),
##   Q5 = col_character(),
##   Q5_7_TEXT = col_logical(),
##   Q6_1 = col_character(),
##   Q6_2 = col_character(),
##   Q6_3 = col_character(),
##   Q9 = col_character()
##   # ... with 83 more columns
## )
```

```
## See spec(...) for full column specifications.
```

```
e_data <- read_csv(here::here("Data", "Cleaned_ORC_Data.csv"))
```

```
## Parsed with column specification:
## cols(
##   school_id = col_double(),
##   school = col_character(),
##   energy_efficiency = col_double(),
##   thermal_comfort = col_double(),
##   acoustics = col_double(),
##   visual_quality = col_double(),
##   iaq = col_double(),
##   thermal_comfort_occupant = col_double(),
##   thermal_comfort_site = col_double(),
##   acoustics_occupant = col_double(),
##   acoustics_site = col_double(),
##   visual_quality_occupant = col_double(),
##   visual_quality_site = col_double(),
##   iaq_occupant = col_double(),
##   iaq_site = col_double(),
##   n_participants = col_double()
## )
```

```r
df <- left_join(q_data, e_data, by = "school") %>%
  select(RecordedDate, ResponseId, school,
         job_cat:loc_missing,
         light_dissatisfied, light_impact_work,
         sound_dissatisfied, sound_impact_work,
         temp_dissatisfied, temp_impact_work,
         air_qual_dissatisfied,
         any_headaches:freq_symptoms_count,
         environment_dissatisfied,
         energy_efficiency:iaq)
```

Summarize the study population

```r
# How many surveys did we have?
df2 <- filter(df, !is.na(school))
nrow(df2)
```

```
## [1] 134
```

```r
summary(df2)
```

```
##   RecordedDate                  ResponseId           school
##  Min.   :2019-09-03 09:37:48   Length:134         Length:134
##  1st Qu.:2019-09-03 13:21:39   Class :character   Class :character
##  Median :2019-09-04 19:03:12   Mode  :character   Mode  :character
##  Mean   :2019-09-09 03:49:19
##  3rd Qu.:2019-09-11 11:52:44
##  Max.   :2019-09-25 08:18:43
##
##    job_cat           job_teacher     job_not_teacher  job_site_aide
##  Length:134         Min.   :0.0000   Min.   :0.0000   Min.   :0.00000
##  Class :character   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000
##  Mode  :character   Median :1.0000   Median :0.0000   Median :0.00000
##                     Mean   :0.6194   Mean   :0.3806   Mean   :0.03731
##                     3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.00000
##                     Max.   :1.0000   Max.   :1.0000   Max.   :1.00000
##
##    job_coach       job_counselor      job_health        job_media
##  Min.   :0.00000   Min.   :0.0000   Min.   :0.00000   Min.   :0.00000
##  1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000
##  Median :0.00000   Median :0.0000   Median :0.00000   Median :0.00000
##  Mean   :0.01493   Mean   :0.0597   Mean   :0.02239   Mean   :0.01493
##  3rd Qu.:0.00000   3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.00000
##  Max.   :1.00000   Max.   :1.0000   Max.   :1.00000   Max.   :1.00000
##
##  job_operations      job_paraed      job_other_not_spec  job_missing
##  Min.   :0.00000   Min.   :0.00000   Min.   :0.00000    Min.   :0
##  1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000    1st Qu.:0
##  Median :0.00000   Median :0.00000   Median :0.00000    Median :0
##  Mean   :0.04478   Mean   :0.04478   Mean   :0.07463    Mean   :0
##  3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000    3rd Qu.:0
```

```
##  Max.   :1.00000   Max.   :1.00000   Max.   :1.00000    Max.   :0
##
##  primary_loc        loc_classroom    loc_cafeteria      loc_library
##  Length:134         Min.   :0.0000   Min.   :0.00000   Min.   :0.00000
##  Class :character   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.00000
##  Mode  :character   Median :1.0000   Median :0.00000   Median :0.00000
##                     Mean   :0.6716   Mean   :0.09701   Mean   :0.02985
##                     3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.00000
##                     Max.   :1.0000   Max.   :1.00000   Max.   :1.00000
##
##  loc_no_primary       loc_office       loc_other        loc_missing
##  Min.   :0.000000   Min.   :0.0000   Min.   :0.00000   Min.   :0
##  1st Qu.:0.000000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0
##  Median :0.000000   Median :0.0000   Median :0.00000   Median :0
##  Mean   :0.007463   Mean   :0.1493   Mean   :0.04478   Mean   :0
##  3rd Qu.:0.000000   3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0
##  Max.   :1.000000   Max.   :1.0000   Max.   :1.00000   Max.   :0
##
##  light_dissatisfied light_impact_work sound_dissatisfied sound_impact_work
##  Min.   :0.0000     Min.   :0.0000    Min.   :0.000      Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.:0.0000    1st Qu.:0.000      1st Qu.:0.0000
##  Median :0.0000     Median :0.0000    Median :0.000      Median :0.0000
##  Mean   :0.2687     Mean   :0.2388    Mean   :0.209      Mean   :0.1493
##  3rd Qu.:1.0000     3rd Qu.:0.0000    3rd Qu.:0.000      3rd Qu.:0.0000
##  Max.   :1.0000     Max.   :1.0000    Max.   :1.000      Max.   :1.0000
##
##  temp_dissatisfied temp_impact_work air_qual_dissatisfied any_headaches
##  Min.   :0.0000    Min.   :0.000    Min.   :0.0000        Min.   :0.0000
##  1st Qu.:0.0000    1st Qu.:0.000    1st Qu.:0.0000        1st Qu.:0.0000
##  Median :0.0000    Median :1.000    Median :0.0000        Median :1.0000
##  Mean   :0.4851    Mean   :0.597    Mean   :0.1493        Mean   :0.7317
##  3rd Qu.:1.0000    3rd Qu.:1.000    3rd Qu.:0.0000        3rd Qu.:1.0000
##  Max.   :1.0000    Max.   :1.000    Max.   :1.0000        Max.   :1.0000
##                                                           NA's   :11
##  freq_headaches     any_wheeze      freq_wheeze   any_eye_irritation
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.00   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00   1st Qu.:0.0000
##  Median :0.0000   Median :0.0000   Median :0.00   Median :1.0000
##  Mean   :0.3089   Mean   :0.1417   Mean   :0.05   Mean   :0.6148
##  3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.00   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.00   Max.   :1.0000
##  NA's   :11       NA's   :14       NA's   :14     NA's   :12
##  freq_eye_irritation any_sore_throat  freq_sore_throat any_sinus_symp
##  Min.   :0.0000      Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000      1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.0000
##  Median :0.0000      Median :1.0000   Median :0.0000   Median :1.0000
##  Mean   :0.3115      Mean   :0.6016   Mean   :0.2358   Mean   :0.7787
##  3rd Qu.:1.0000      3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:1.0000
##  Max.   :1.0000      Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
##  NA's   :12          NA's   :11       NA's   :11       NA's   :12
##  freq_sinus_symp    any_sneeze      freq_sneeze     any_eye_strain
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000   Median :1.0000   Median :0.0000   Median :1.0000
```

```
##  Mean   :0.3115   Mean   :0.6748   Mean   :0.3008   Mean   :0.7107
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
##  NA's   :12       NA's   :11       NA's   :11       NA's   :13
##  freq_eye_strain  any_tension      freq_tension   any_lack_atten
##  Min.   :0.000   Min.   :0.0000   Min.   :0.000   Min.   :0.0000
##  1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.0000
##  Median :0.000   Median :1.0000   Median :0.000   Median :0.0000
##  Mean   :0.438   Mean   :0.6116   Mean   :0.314   Mean   :0.4553
##  3rd Qu.:1.000   3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:1.0000
##  Max.   :1.000   Max.   :1.0000   Max.   :1.000   Max.   :1.0000
##  NA's   :13      NA's   :13       NA's   :13      NA's   :11
##  freq_lack_atten  any_dizziness    freq_dizziness     any_nausea
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000   Median :0.0000   Median :0.0000   Median :0.0000
##  Mean   :0.1626   Mean   :0.2683   Mean   :0.1138   Mean   :0.2439
##  3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
##  NA's   :11       NA's   :11       NA's   :11       NA's   :11
##   freq_nausea      any_depression   freq_depression   any_lethargy
##  Min.   :0.00000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.00000   Median :0.0000   Median :0.0000   Median :1.0000
##  Mean   :0.06504   Mean   :0.3008   Mean   :0.1463   Mean   :0.6585
##  3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:1.0000
##  Max.   :1.00000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
##  NA's   :11        NA's   :11       NA's   :11       NA's   :11
##  freq_lethargy   any_symptoms_count freq_symptoms_count
##  Min.   :0.000   Min.   : 0.000     Min.   : 0.000
##  1st Qu.:0.000   1st Qu.: 4.000     1st Qu.: 0.000
##  Median :0.000   Median : 7.000     Median : 2.000
##  Mean   :0.374   Mean   : 6.201     Mean   : 2.858
##  3rd Qu.:1.000   3rd Qu.: 9.000     3rd Qu.: 5.000
##  Max.   :1.000   Max.   :13.000     Max.   :10.000
##  NA's   :11
##  environment_dissatisfied energy_efficiency thermal_comfort
##  Min.   :0.0000           Min.   :48.00     Min.   :43.00
##  1st Qu.:0.0000           1st Qu.:83.00     1st Qu.:59.00
##  Median :0.0000           Median :91.00     Median :63.00
##  Mean   :0.1716           Mean   :86.04     Mean   :60.63
##  3rd Qu.:0.0000           3rd Qu.:97.00     3rd Qu.:64.00
##  Max.   :1.0000           Max.   :98.00     Max.   :68.00
##
##    acoustics     visual_quality       iaq
##  Min.   :63.00   Min.   :43.00   Min.   :45.00
##  1st Qu.:66.00   1st Qu.:47.00   1st Qu.:72.00
##  Median :67.00   Median :48.00   Median :74.00
##  Mean   :68.41   Mean   :49.16   Mean   :72.54
##  3rd Qu.:72.00   3rd Qu.:51.00   3rd Qu.:77.00
##  Max.   :72.00   Max.   :59.00   Max.   :83.00
##
```

How many schools?

```r
length(unique(df2$school))
```

```
## [1] 11
```

```r
table(df2$school, useNA = 'ifany')
```

```
##
## Arapahoe Ridge Elementary   Cotton Creek Elementary
##                         7                        18
##  Mountain View Elementary      North Mor Elementary
##                        20                         5
##     North Star Elementary    Northglenn High School
##                        10                        24
##                  STEM Lab               STEM Launch
##                         7                        17
##         Stukey Elementary       Westview Elementary
##                         9                         7
##       Woodglen Elementary
##                        10
```

Jobs represeted?

```r
table(df2$job_cat, useNA = 'ifany')
```

```
##
##          Admin   Aide_Monitor          Coach      Counselor    Health_Aide
##              9              5              2              8              3
##     Media_Tech     Operations Other_Not_Spec   Paraeducator        Teacher
##              2              6             10              6             83
```

```r
prop.table(table(df2$job_cat, useNA = 'ifany'))
```

```
##
##          Admin   Aide_Monitor          Coach      Counselor    Health_Aide
##     0.06716418     0.03731343     0.01492537     0.05970149     0.02238806
##     Media_Tech     Operations Other_Not_Spec   Paraeducator        Teacher
##     0.01492537     0.04477612     0.07462687     0.04477612     0.61940299
```

Primary work location

```r
table(df2$primary_loc, useNA = 'ifany')
```

```
##
##         Classroom Kitchen/Cafeteria           Library  No primary space
##                90                13                 4                 1
##            Office             Other
##                20                 6
```

```
prop.table(table(df2$primary_loc, useNA = 'ifany'))
```

```
##
##         Classroom Kitchen/Cafeteria           Library  No primary space
##       0.671641791       0.097014925       0.029850746       0.007462687
##            Office             Other
##       0.149253731       0.044776119
```

Summarize the outcomes– For the sake of the abstract, focusing on: dissatisfaction (somewhat or extremely) number of frequent (> 1 per week) symptoms reported

Dissatisfied with the lighting? 27%

```
r table(df2$light_dissatisfied, useNA = 'ifany')
## ##  0  1 ## 98 36
r prop.table(table(df2$light_dissatisfied, useNA = 'ifany'))
## ##         0         1 ## 0.7313433 0.2686567
```

Dissatisfied with the acoustics? 21%

```
r table(df2$sound_dissatisfied, useNA = 'ifany')
## ##   0   1 ## 106  28
r prop.table(table(df2$sound_dissatisfied, useNA = 'ifany'))
## ##         0         1 ## 0.7910448 0.2089552
```

Dissatisfied with the thermal comfort? 48%

```
r table(df2$temp_dissatisfied, useNA = 'ifany')
## ##  0  1 ## 69 65
r prop.table(table(df2$temp_dissatisfied, useNA = 'ifany'))
## ##         0         1 ## 0.5149254 0.4850746
```

Dissatisfied with the IAQ? 15%

```
r table(df2$air_qual_dissatisfied, useNA = 'ifany')
## ##   0   1 ## 114  20
r prop.table(table(df2$air_qual_dissatisfied, useNA = 'ifany'))
## ##         0         1 ## 0.8507463 0.1492537
```

Overall environment? 17%

```
r table(df2$environment_dissatisfied, useNA = 'ifany')
## ##   0   1 ## 111  23
r prop.table(table(df2$environment_dissatisfied, useNA = 'ifany'))
## ##         0         1 ## 0.8283582 0.1716418
```

Summary of frequent health symptoms reported

```
r summary(df2$freq_symptoms_count)
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. ##   0.000   0.000   2.000   2.858
5.000  10.000
```

Summarize the ORC scores Calculate an "average orc" for the 5 variables assessed

```
df2 <- df2 %>%
  mutate(mean_orc = (energy_efficiency + thermal_comfort +
                       acoustics + visual_quality + iaq)/5)
View(head(select(df2, energy_efficiency:mean_orc)))

summary(df2$energy_efficiency)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
##   48.00   83.00   91.00   86.04   97.00   98.00
```

```
summary(df2$thermal_comfort)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   43.00   59.00   63.00   60.63   64.00   68.00
```

```
summary(df2$acoustics)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   63.00   66.00   67.00   68.41   72.00   72.00
```

```
summary(df2$visual_quality)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   43.00   47.00   48.00   49.16   51.00   59.00
```

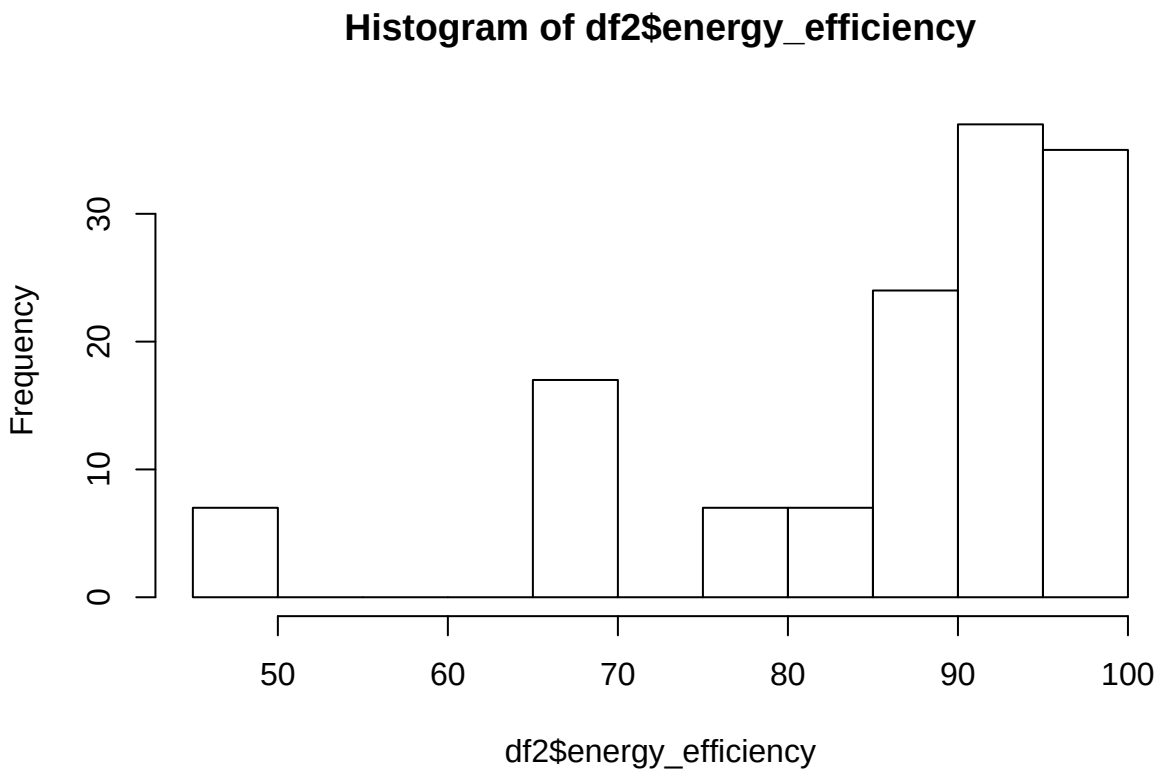
```
summary(df2$iaq)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   45.00   72.00   74.00   72.54   77.00   83.00
```

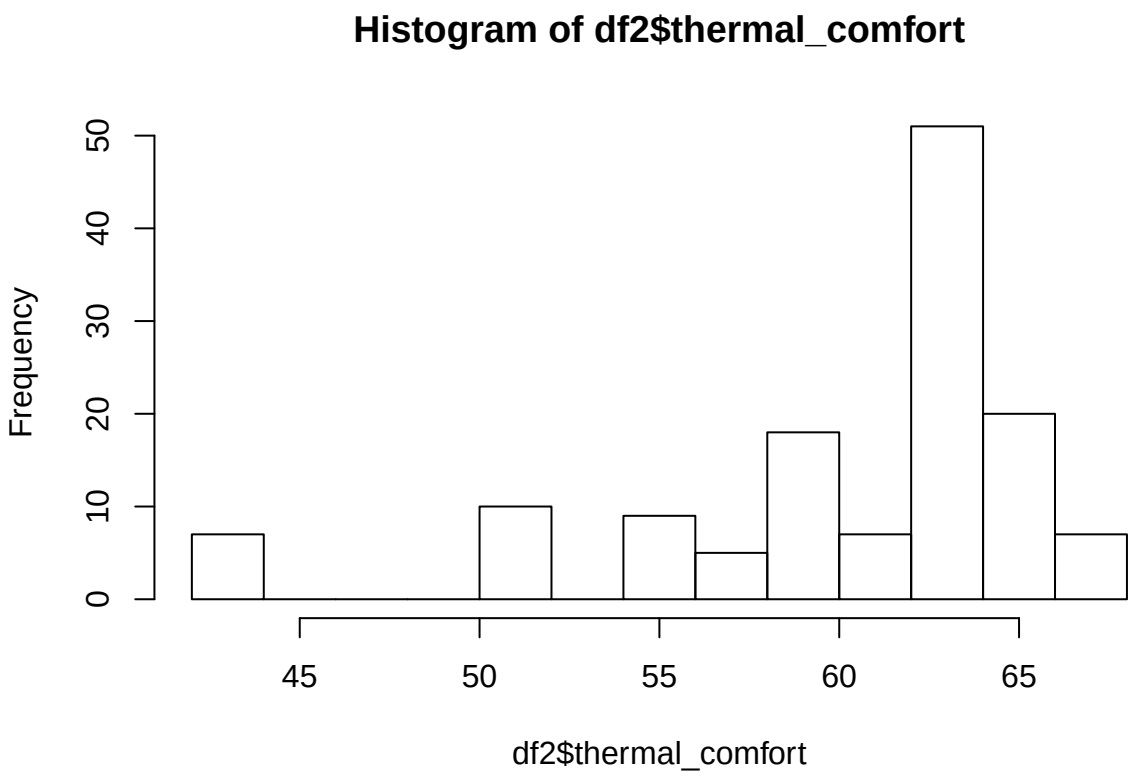
```
summary(df2$mean_orc)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   54.40   64.10   67.80   67.36   70.00   73.80
```

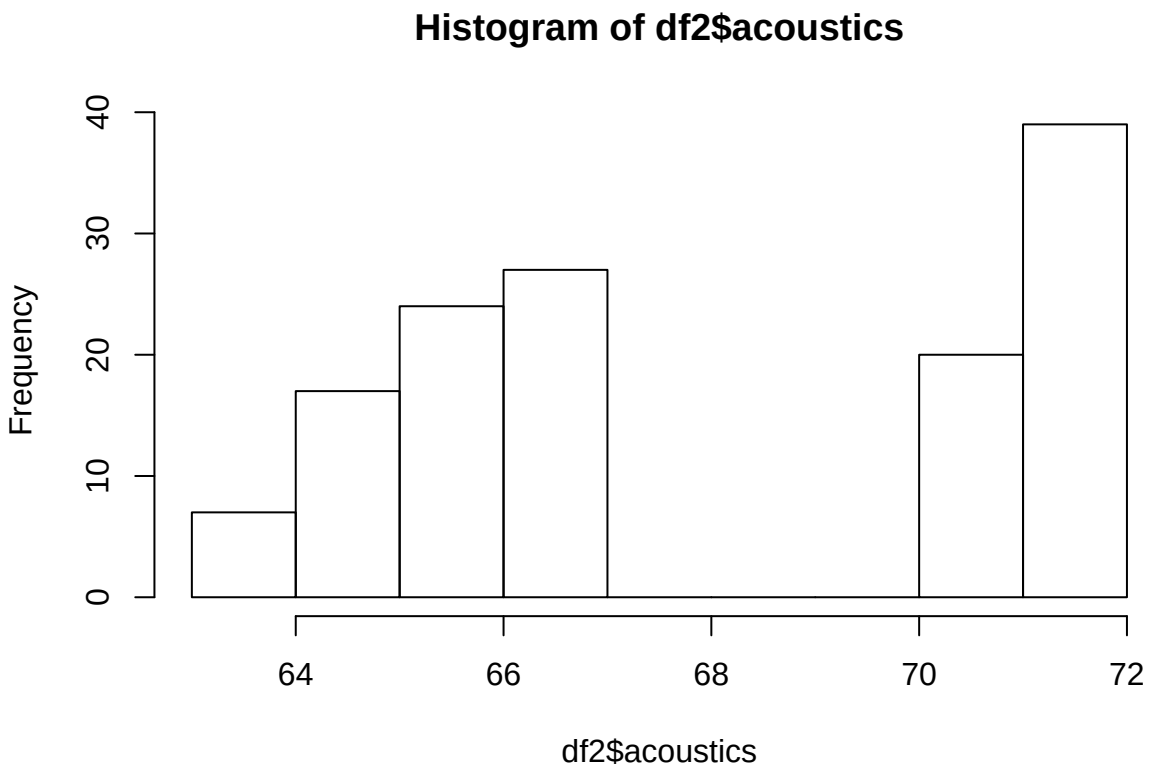
```
hist(df2$energy_efficiency)
```



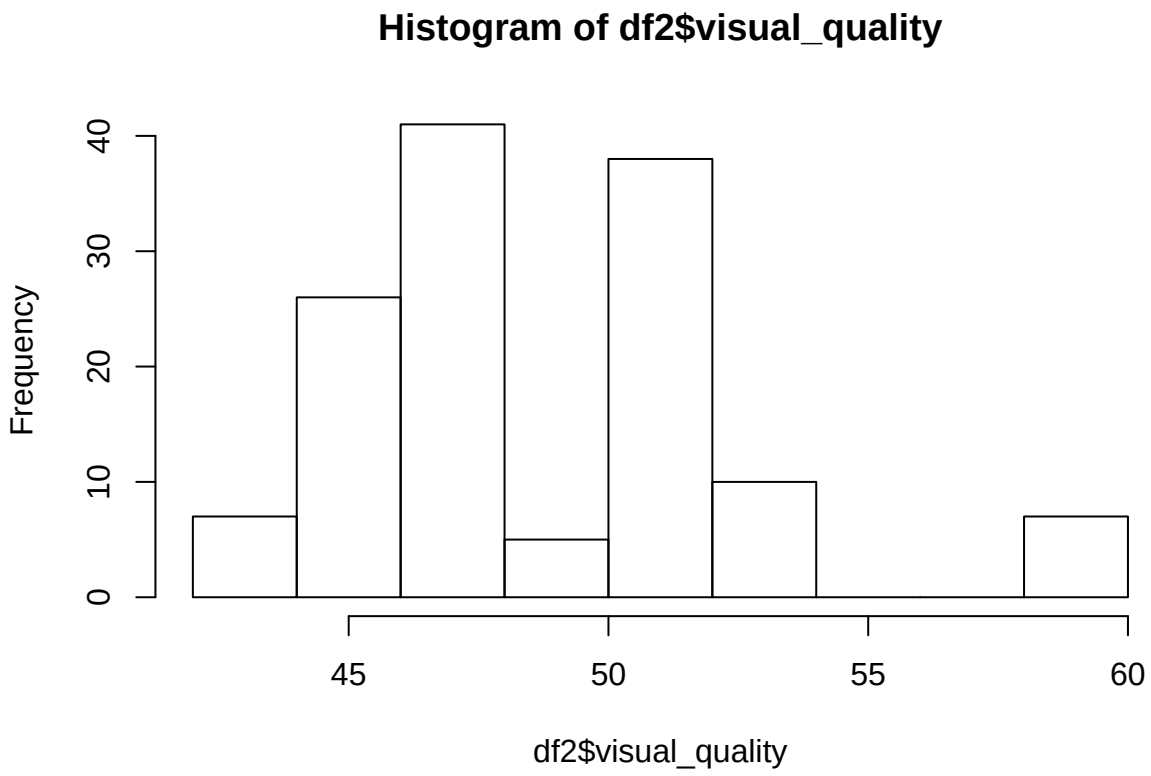

```r
hist(df2$thermal_comfort)
```



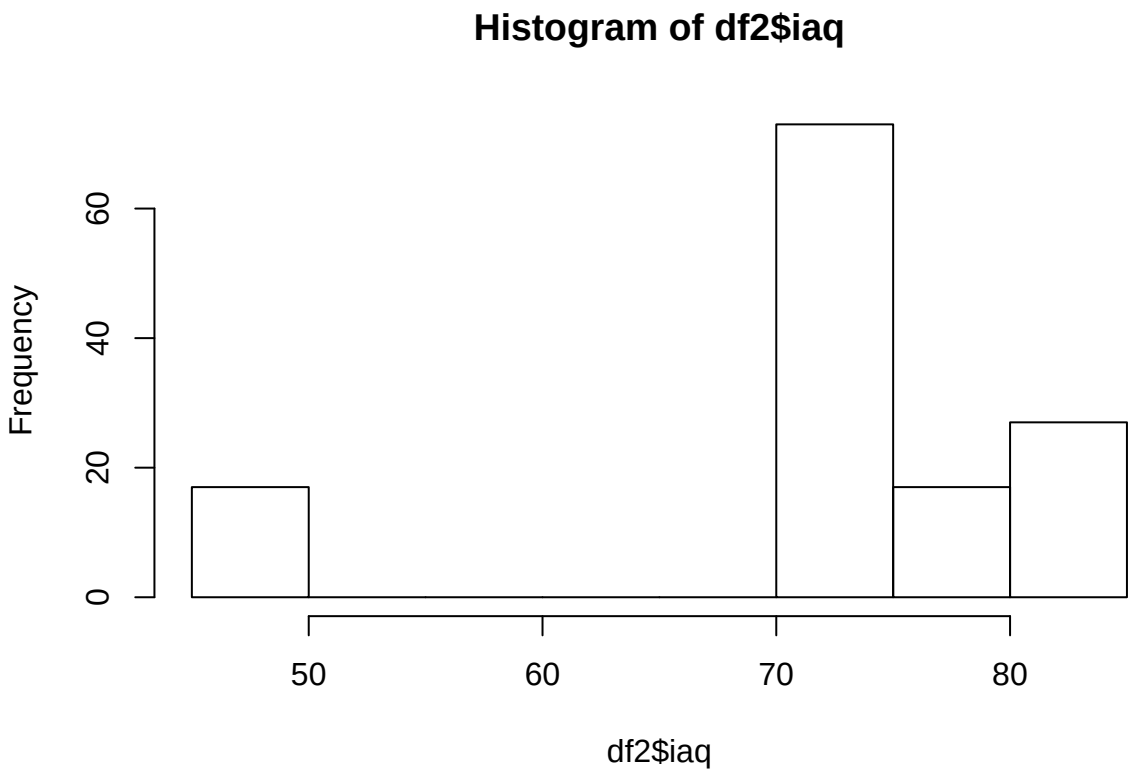

```r
hist(df2$acoustics)
```




```
hist(df2$visual_quality)
```

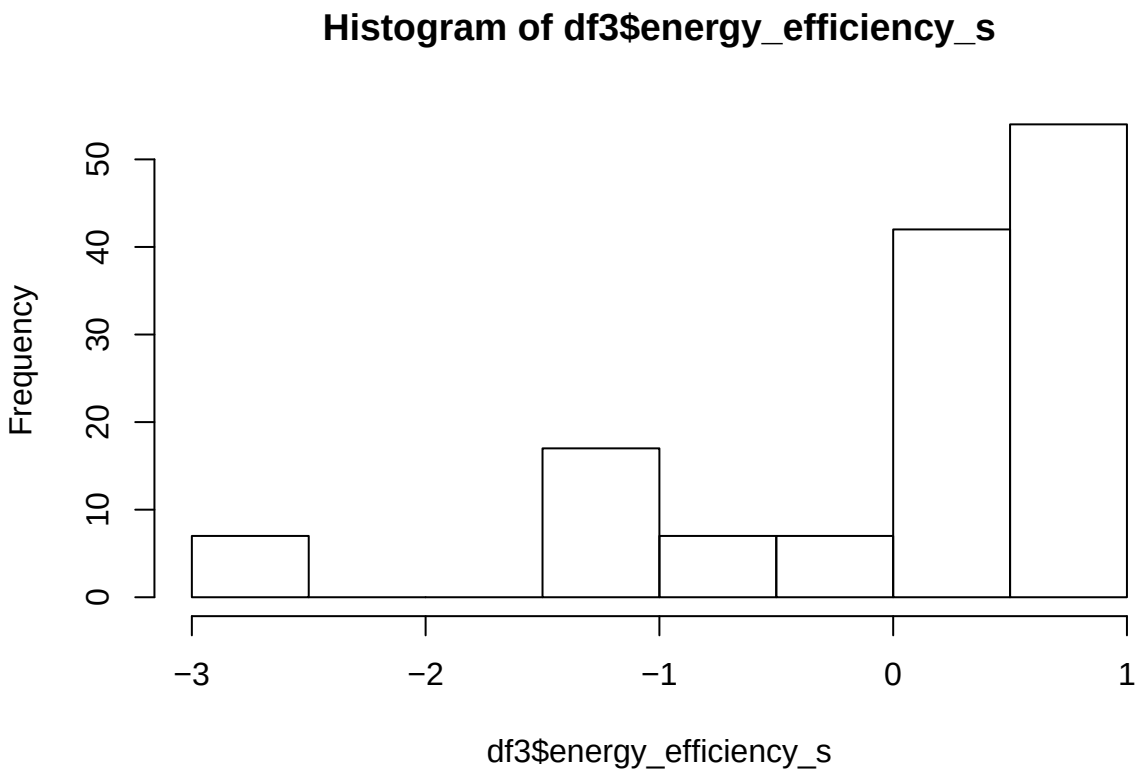

Histogram of df2$visual_quality
Frequency
0
10
20
30
40
45
50
55
60
df2$visual_quality

```
hist(df2$iaq)
```

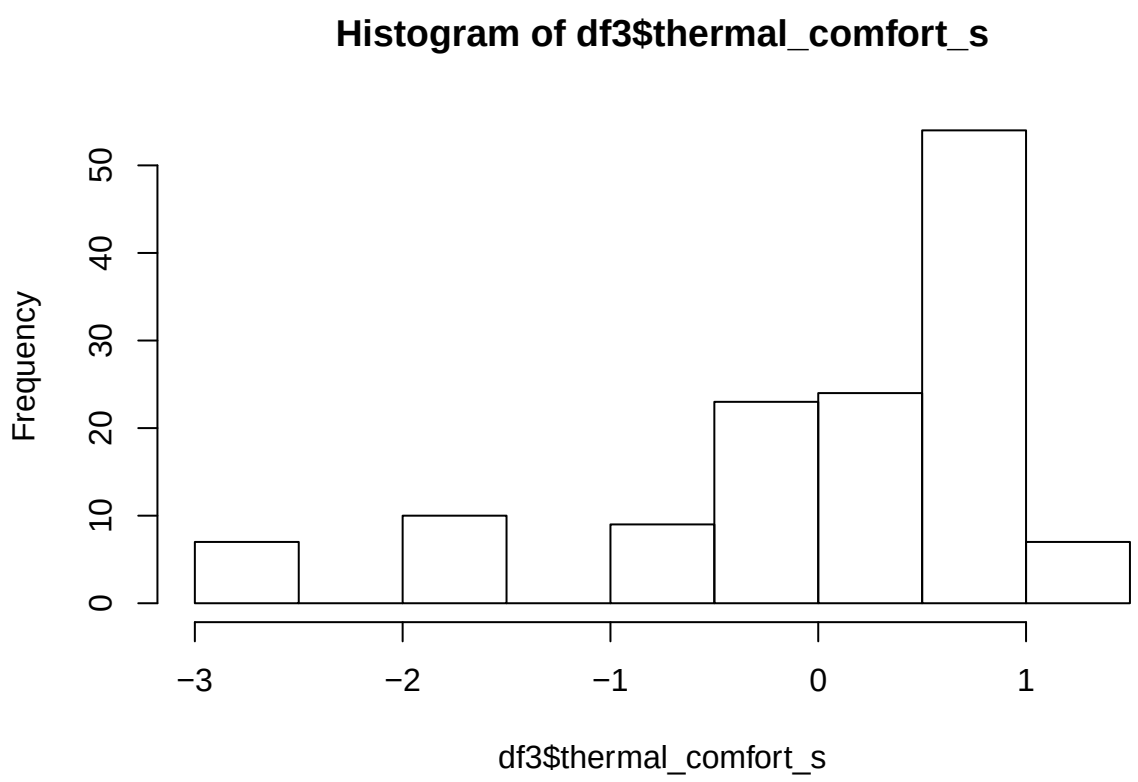


scale the ORC scores

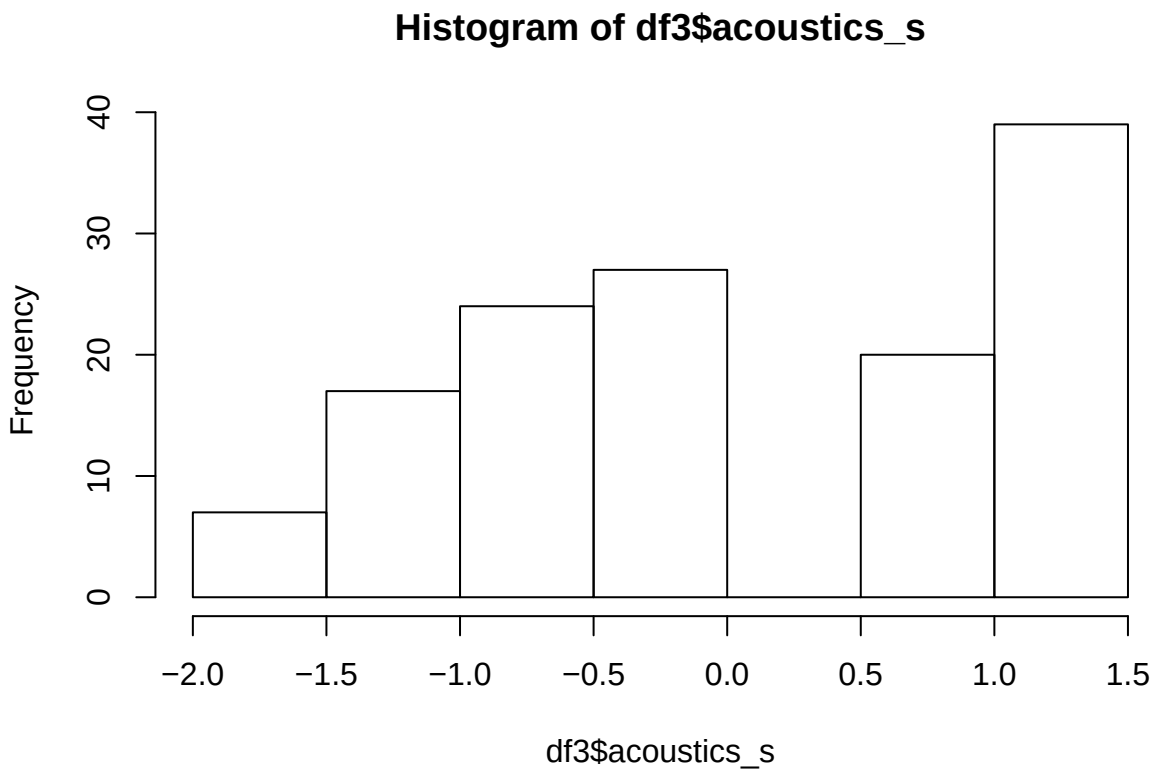
```
df3 <- df2 %>%
  mutate(energy_efficiency_s = scale(energy_efficiency),
         thermal_comfort_s = scale(thermal_comfort),
         acoustics_s = scale(acoustics),
         visual_quality_s = scale(visual_quality),
         iaq_s = scale(iaq))

hist(df3$energy_efficiency_s)
```



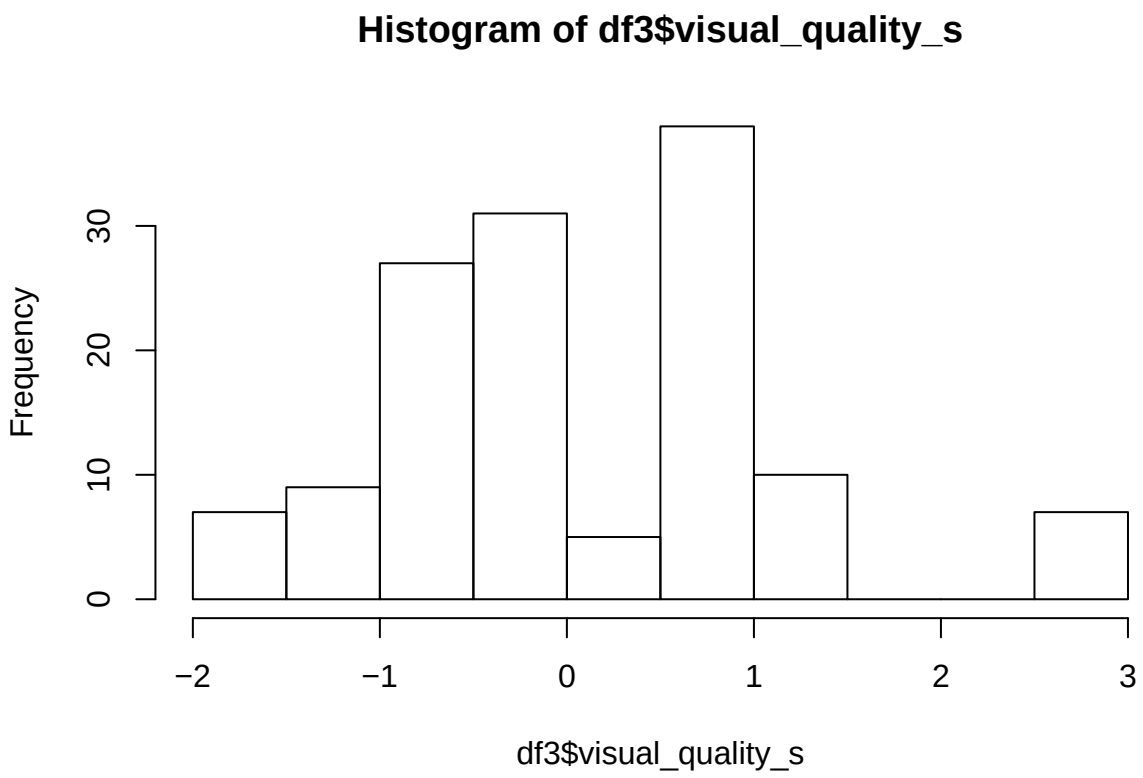

```r
hist(df3$thermal_comfort_s)
```



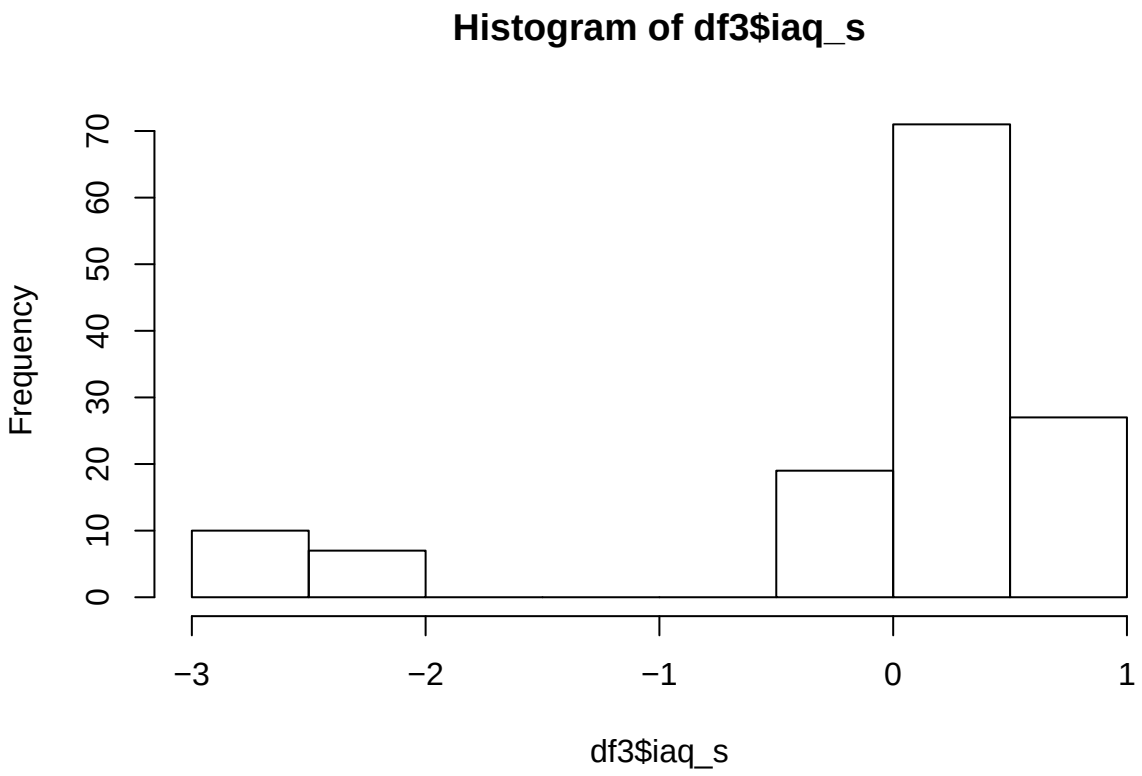

```r
hist(df3$acoustics_s)
```




```
hist(df3$visual_quality_s)
```

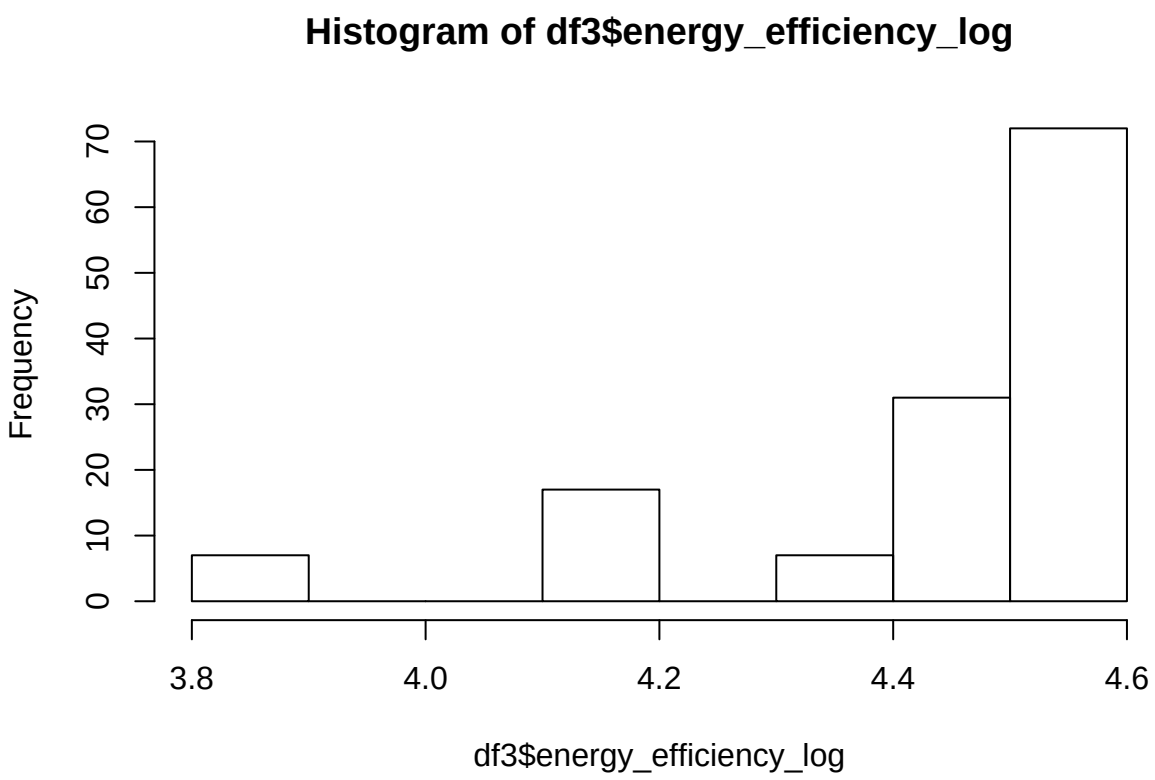



```
hist(df3$iaq_s)
```

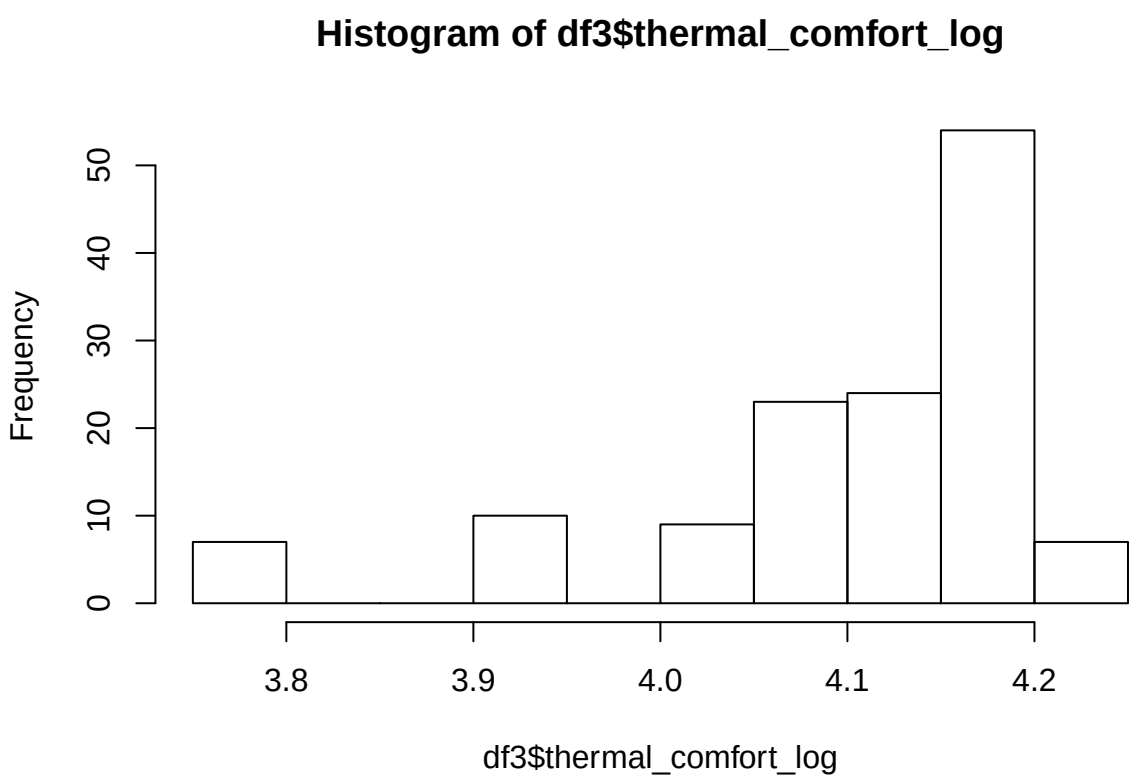


log-transform

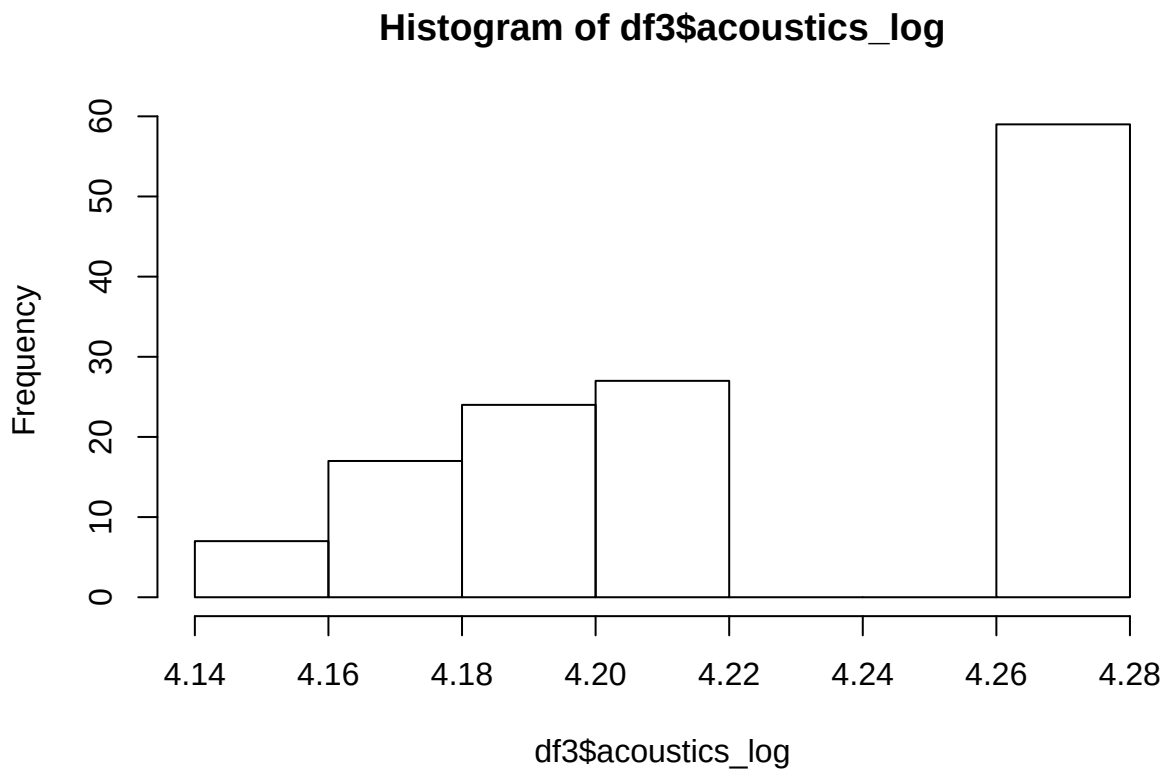
```
df3 <- df3 %>%
  mutate(energy_efficiency_log = log(energy_efficiency),
         thermal_comfort_log = log(thermal_comfort),
         acoustics_log = log(acoustics),
         visual_quality_log = log(visual_quality),
         iaq_log = log(iaq))

hist(df3$energy_efficiency_log)
```



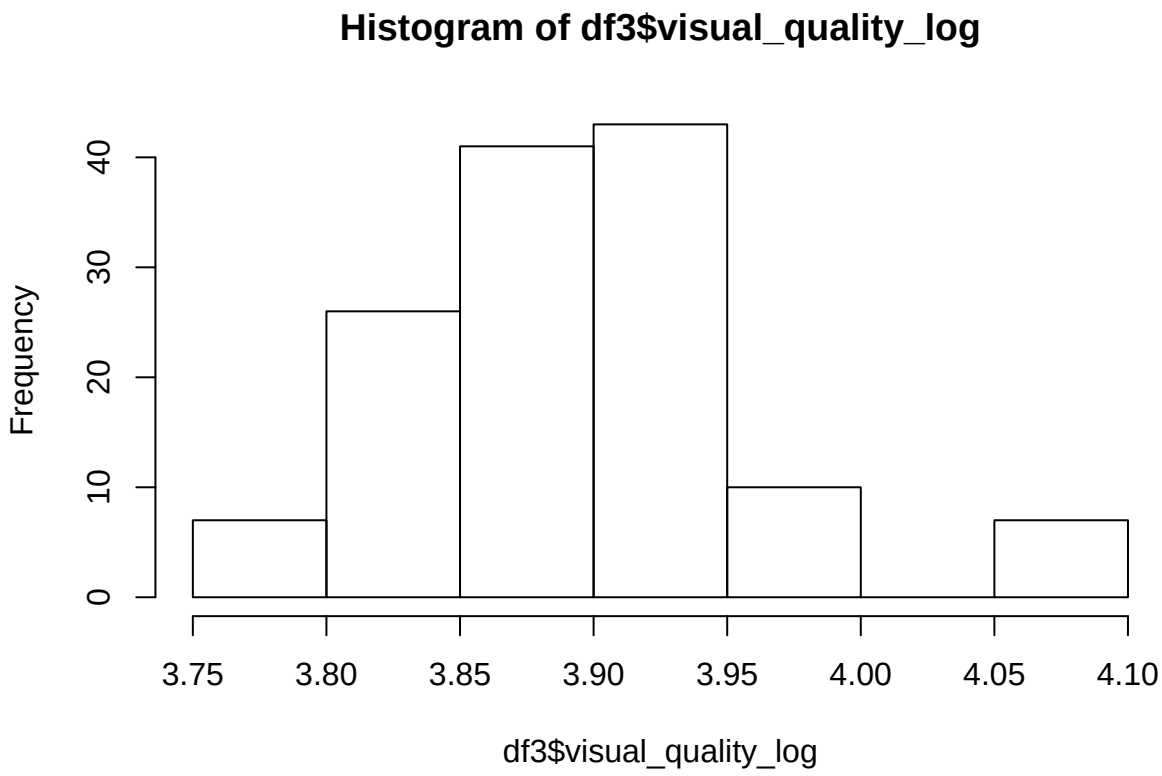

```
hist(df3$thermal_comfort_log)
```



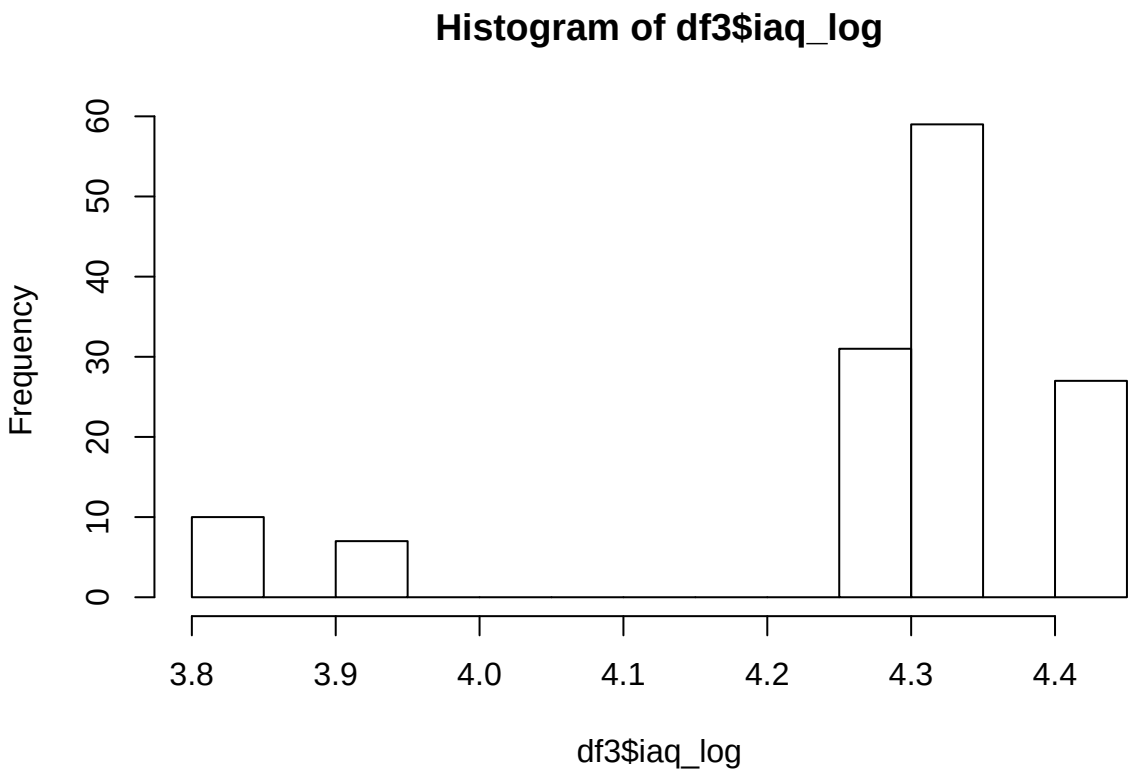

```
hist(df3$acoustics_log)
```




```
hist(df3$visual_quality_log)
```




```r
hist(df3$iaq_log)
```



Correlations between outcomes and exposures

```
library(ggcorrplot)
```

```
## Warning: package 'ggcorrplot' was built under R version 3.6.1
```

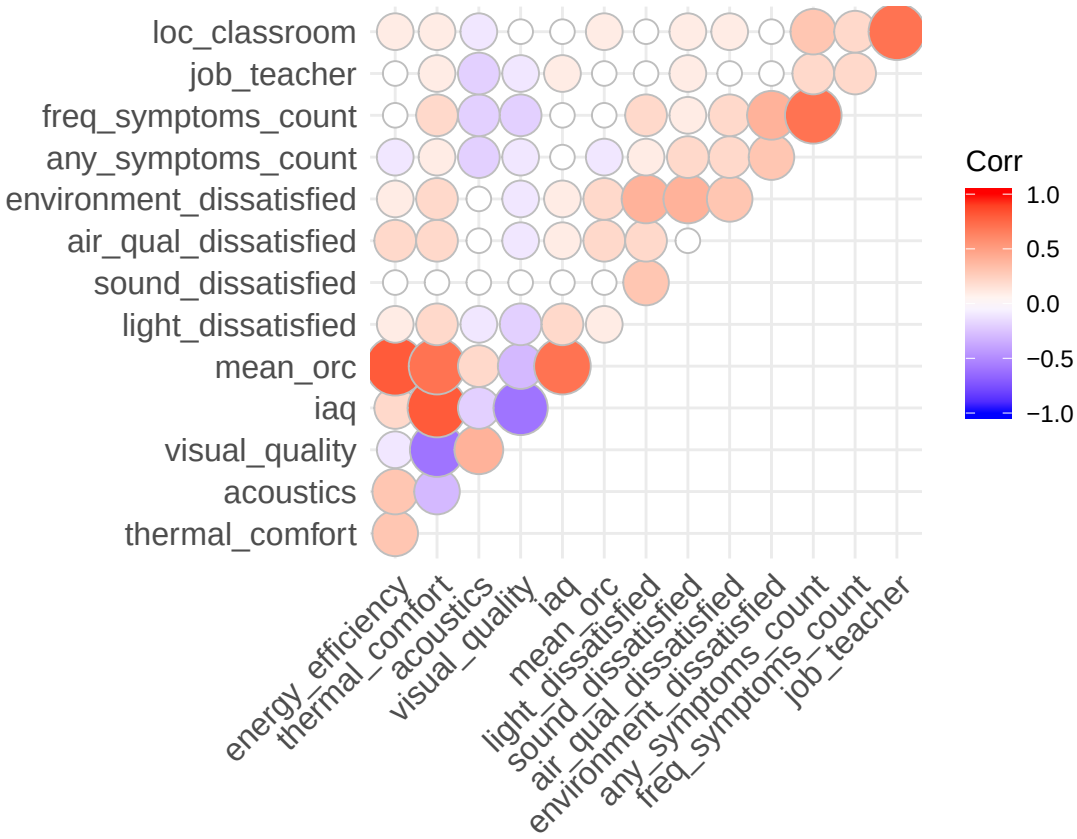
```
df_corr <- df3 %>%
  select(energy_efficiency:mean_orc,
         light_dissatisfied, sound_dissatisfied,
         air_qual_dissatisfied, environment_dissatisfied,
         any_symptoms_count, freq_symptoms_count,
         job_teacher, loc_classroom)
cor(df_corr)
```

```
##                          energy_efficiency thermal_comfort   acoustics
## energy_efficiency              1.00000000      0.33950889  0.31714886
## thermal_comfort                0.33950889      1.00000000 -0.34118566
## acoustics                      0.31714886     -0.34118566  1.00000000
## visual_quality                -0.12711041     -0.64534359  0.41809631
## iaq                            0.23799078      0.81389770 -0.16208000
## mean_orc                       0.82887856      0.70285290  0.22910895
## light_dissatisfied             0.08232075      0.16440212 -0.07092621
## sound_dissatisfied             0.03945706      0.03884819 -0.02113884
## air_qual_dissatisfied          0.15990150      0.19748992 -0.01525428
```

```
## environment_dissatisfied        0.13759204      0.18030413 -0.01592537
## any_symptoms_count              -0.05409301      0.09463274 -0.24513860
## freq_symptoms_count              0.04659974      0.18139335 -0.22041211
## job_teacher                      0.01067005      0.10445352 -0.15238437
## loc_classroom                    0.13305458      0.08803245 -0.08876578
##                          visual_quality          iaq    mean_orc
## energy_efficiency           -0.12711041  0.237990780  0.82887856
## thermal_comfort             -0.64534359  0.813897697  0.70285290
## acoustics                    0.41809631 -0.162079998  0.22910895
## visual_quality               1.00000000 -0.621965271 -0.32204337
## iaq                         -0.62196527  1.000000000  0.70815857
## mean_orc                    -0.32204337  0.708158574  1.00000000
## light_dissatisfied          -0.16261346  0.175246247  0.13989897
## sound_dissatisfied           0.04919257  0.005196547  0.04165675
## air_qual_dissatisfied       -0.13508586  0.112860079  0.17820195
## environment_dissatisfied    -0.05851730  0.088431349  0.16096341
## any_symptoms_count          -0.14961182 -0.031614833 -0.08021712
## freq_symptoms_count         -0.18478816  0.036043784  0.03318639
## job_teacher                 -0.08145551  0.096300795  0.04550732
## loc_classroom               -0.01818573 -0.020335604  0.07902310
##                          light_dissatisfied sound_dissatisfied
## energy_efficiency             0.082320754        0.039457059
## thermal_comfort               0.164402123        0.038848186
## acoustics                    -0.070926211       -0.021138842
## visual_quality               -0.162613456        0.049192574
## iaq                           0.175246247        0.005196547
## mean_orc                      0.139898972        0.041656754
## light_dissatisfied            1.000000000        0.309650286
## sound_dissatisfied            0.309650286        1.000000000
## air_qual_dissatisfied         0.218604943        0.042285679
## environment_dissatisfied      0.393847217        0.350202852
## any_symptoms_count            0.113373454        0.196593575
## freq_symptoms_count           0.159537677        0.134572426
## job_teacher                  -0.010350744        0.062632088
## loc_classroom                -0.006420951        0.085756646
##                          air_qual_dissatisfied environment_dissatisfied
## energy_efficiency                  0.15990150               0.13759204
## thermal_comfort                    0.19748992               0.18030413
## acoustics                         -0.01525428              -0.01592537
## visual_quality                    -0.13508586              -0.05851730
## iaq                                0.11286008               0.08843135
## mean_orc                           0.17820195               0.16096341
## light_dissatisfied                 0.21860494               0.39384722
## sound_dissatisfied                 0.04228568               0.35020285
## air_qual_dissatisfied              1.00000000               0.25366386
## environment_dissatisfied           0.25366386               1.00000000
## any_symptoms_count                 0.17079983               0.27125338
## freq_symptoms_count                0.15987796               0.38928713
## job_teacher                        0.02639504              -0.01003840
## loc_classroom                      0.11448364               0.02327334
##                          any_symptoms_count freq_symptoms_count
## energy_efficiency               -0.05409301          0.04659974
## thermal_comfort                  0.09463274          0.18139335
## acoustics                       -0.24513860         -0.22041211
```

```
## visual_quality                  -0.14961182         -0.18478816
## iaq                             -0.03161483          0.03604378
## mean_orc                        -0.08021712          0.03318639
## light_dissatisfied               0.11337345          0.15953768
## sound_dissatisfied               0.19659357          0.13457243
## air_qual_dissatisfied            0.17079983          0.15987796
## environment_dissatisfied         0.27125338          0.38928713
## any_symptoms_count               1.00000000          0.73613000
## freq_symptoms_count              0.73613000          1.00000000
## job_teacher                      0.23119549          0.15996376
## loc_classroom                    0.25887863          0.17645572
##                          job_teacher loc_classroom
## energy_efficiency         0.01067005   0.133054584
## thermal_comfort           0.10445352   0.088032450
## acoustics                -0.15238437  -0.088765775
## visual_quality           -0.08145551  -0.018185730
## iaq                       0.09630080  -0.020335604
## mean_orc                  0.04550732   0.079023100
## light_dissatisfied       -0.01035074  -0.006420951
## sound_dissatisfied        0.06263209   0.085756646
## air_qual_dissatisfied     0.02639504   0.114483639
## environment_dissatisfied -0.01003840   0.023273340
## any_symptoms_count        0.23119549   0.258878633
## freq_symptoms_count       0.15996376   0.176455719
## job_teacher               1.00000000   0.695613910
## loc_classroom             0.69561391   1.000000000
```

```
ggcorrplot(round(cor(df_corr), 1), method = "circle", type = "upper")
```



Associations between environment dissatisfaction and ORC scores
To help with model convergence, use scaled predictors

```
library(lme4)
```

```
## Warning: package 'lme4' was built under R version 3.6.1
```

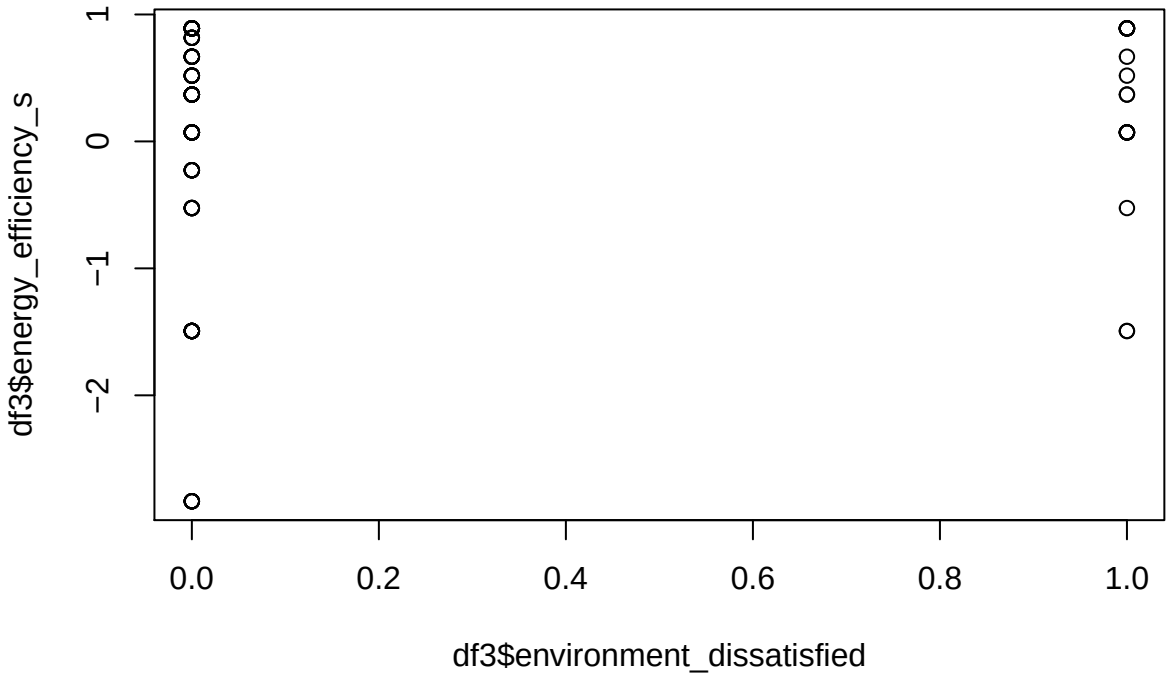
```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

Overall environment dissatisfaction

```
plot(df3$environment_dissatisfied, df3$energy_efficiency_s)
```




```
env_effic <- glmer(environment_dissatisfied ~ energy_efficiency_s + job_teacher +
                     loc_classroom + (1|school),
                   family = binomial, data = df3)
```

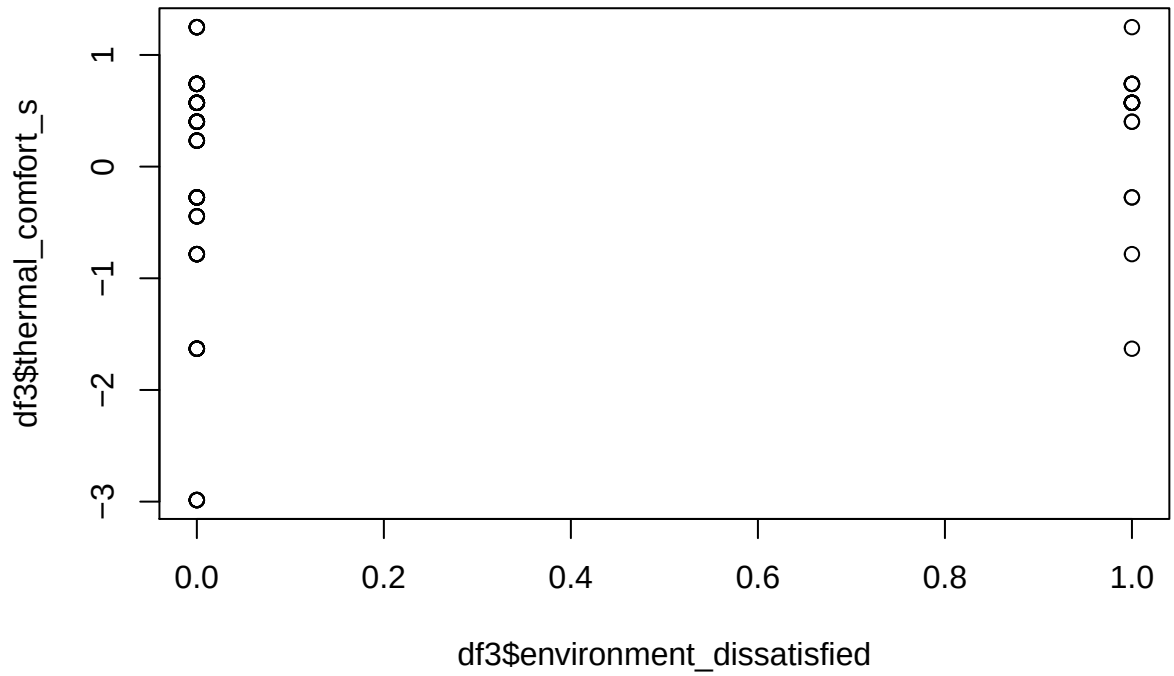
```
## boundary (singular) fit: see ?isSingular
```

```
summary(env_effic)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: environment_dissatisfied ~ energy_efficiency_s + job_teacher +
##     loc_classroom + (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    129.9    144.4    -59.9    119.9      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.5804 -0.5052 -0.4471 -0.3094  3.4757
##
## Random effects:
##  Groups Name        Variance Std.Dev.
```

```
##  school (Intercept) 0        0
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -1.6597     0.4297  -3.862 0.000112 ***
## energy_efficiency_s   0.4598     0.3055   1.505 0.132315
## job_teacher          -0.1453     0.6721  -0.216 0.828819
## loc_classroom         0.1622     0.7096   0.229 0.819209
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) enrg__ jb_tch
## enrgy_ffcn_ -0.173
## job_teacher -0.172  0.120
## loc_classrm -0.447 -0.113 -0.706
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
plot(df3$environment_dissatisfied, df3$thermal_comfort_s)
```



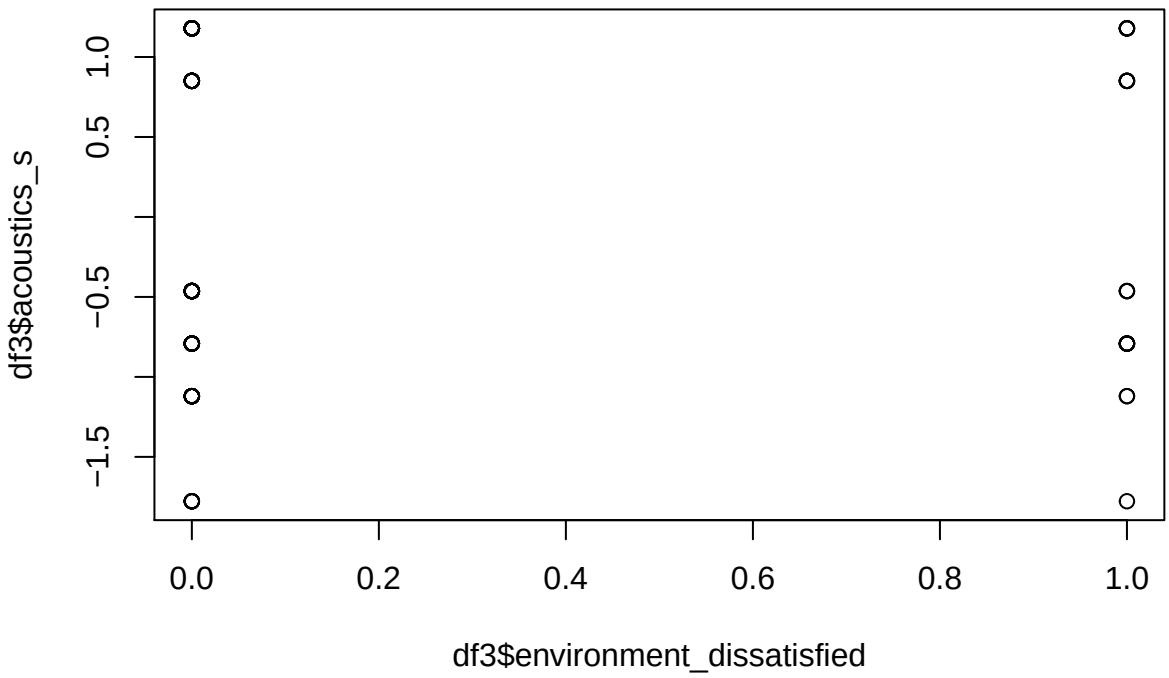

```r
env_therm <- glmer(environment_dissatisfied ~ thermal_comfort_s + job_teacher +
                     loc_classroom + (1|school),
                   family = binomial, data = df3)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(env_therm)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: environment_dissatisfied ~ thermal_comfort_s + job_teacher +
##     loc_classroom + (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    127.1    141.5    -58.5    117.1      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7682 -0.5115 -0.3866 -0.2331  4.3159
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0        0
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)        -1.7026     0.4410  -3.861 0.000113 ***
## thermal_comfort_s   0.7192     0.3593   2.002 0.045324 *
## job_teacher        -0.3261     0.6463  -0.505 0.613841
## loc_classroom       0.2769     0.6810   0.407 0.684269
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) thrm__ jb_tch
## thrml_cmfr_ -0.237
## job_teacher -0.186 -0.008
## loc_classrm -0.453 -0.039 -0.668
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
plot(df3$environment_dissatisfied, df3$acoustics_s)
```

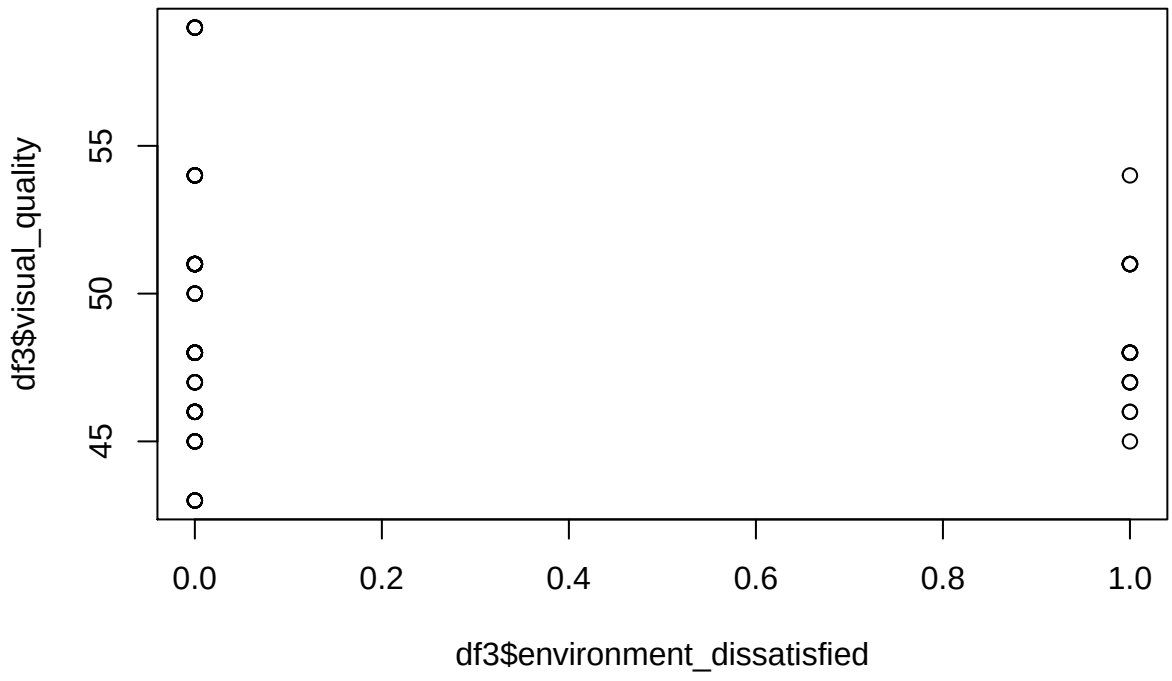



```
env_sound <- glmer(environment_dissatisfied ~ acoustics_s + job_teacher +
                     loc_classroom + (1|school),
                   family = binomial, data = df3)
summary(env_sound)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## environment_dissatisfied ~ acoustics_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    132.5    147.0    -61.2    122.5      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.5670 -0.4536 -0.4308 -0.4079  2.6204
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.08292  0.288
## Number of obs: 134, groups:  school, 11
##
```

```
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.66051    0.45265  -3.668 0.000244 ***
## acoustics_s   -0.05887    0.25516  -0.231 0.817545
## job_teacher   -0.27396    0.65306  -0.420 0.674849
## loc_classroom  0.29294    0.69345   0.422 0.672701
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) acstc_ jb_tch
## acoustics_s -0.043
## job_teacher -0.190  0.093
## loc_classrm -0.394  0.021 -0.677
```

```
plot(df3$environment_dissatisfied, df3$visual_quality)
```

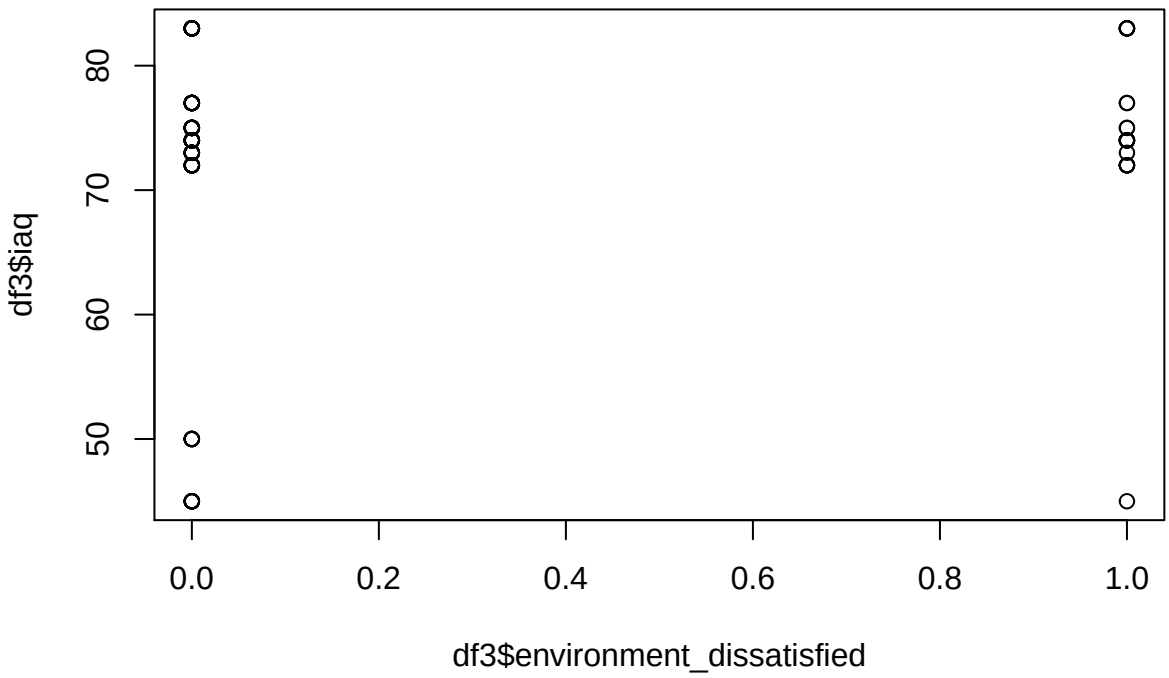



```
env_visual <- glmer(environment_dissatisfied ~ visual_quality_s + job_teacher +
                      loc_classroom + (1|school),
                    family = binomial, data = df3)
summary(env_visual)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
```

```
## Formula:
## environment_dissatisfied ~ visual_quality_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    132.0    146.5    -61.0    122.0      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.5727 -0.4790 -0.4411 -0.3853  2.5954
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.06298  0.251
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -1.6624     0.4500  -3.695  0.00022 ***
## visual_quality_s  -0.1808     0.2616  -0.691  0.48952
## job_teacher       -0.2965     0.6494  -0.457  0.64791
## loc_classroom      0.3199     0.6925   0.462  0.64414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) vsl_q_ jb_tch
## visl_qlty_s  0.072
## job_teacher -0.191  0.066
## loc_classrm -0.393 -0.028 -0.679
```

```
plot(df3$environment_dissatisfied, df3$iaq)
```



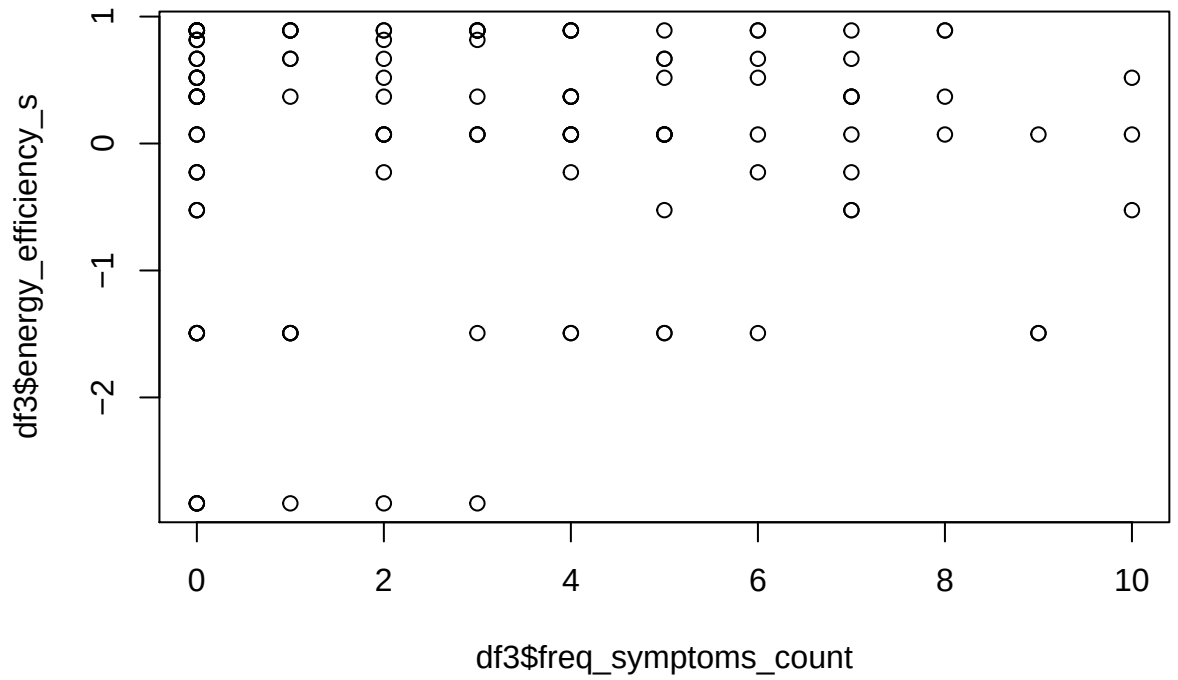

```
env_iaq <- glmer(environment_dissatisfied ~ iaq_s + job_teacher +
                   loc_classroom + (1|school),
                   family = binomial, data = df3)
summary(env_iaq)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: environment_dissatisfied ~ iaq_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    131.3    145.8    -60.6    121.3      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.5718 -0.4718 -0.4494 -0.3203  3.3419
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.01105  0.1051
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
```

```
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.6580     0.4366  -3.798 0.000146 ***
## iaq_s           0.3020     0.2813   1.074 0.282929
## job_teacher    -0.3843     0.6625  -0.580 0.561801
## loc_classroom   0.4221     0.7092   0.595 0.551790
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) iaq_s  jb_tch
## iaq_s       -0.113
## job_teacher -0.178 -0.132
## loc_classrm -0.408  0.110 -0.696
```

Number of frequent symptoms reported

```
plot(df3$freq_symptoms_count, df3$energy_efficiency_s)
```

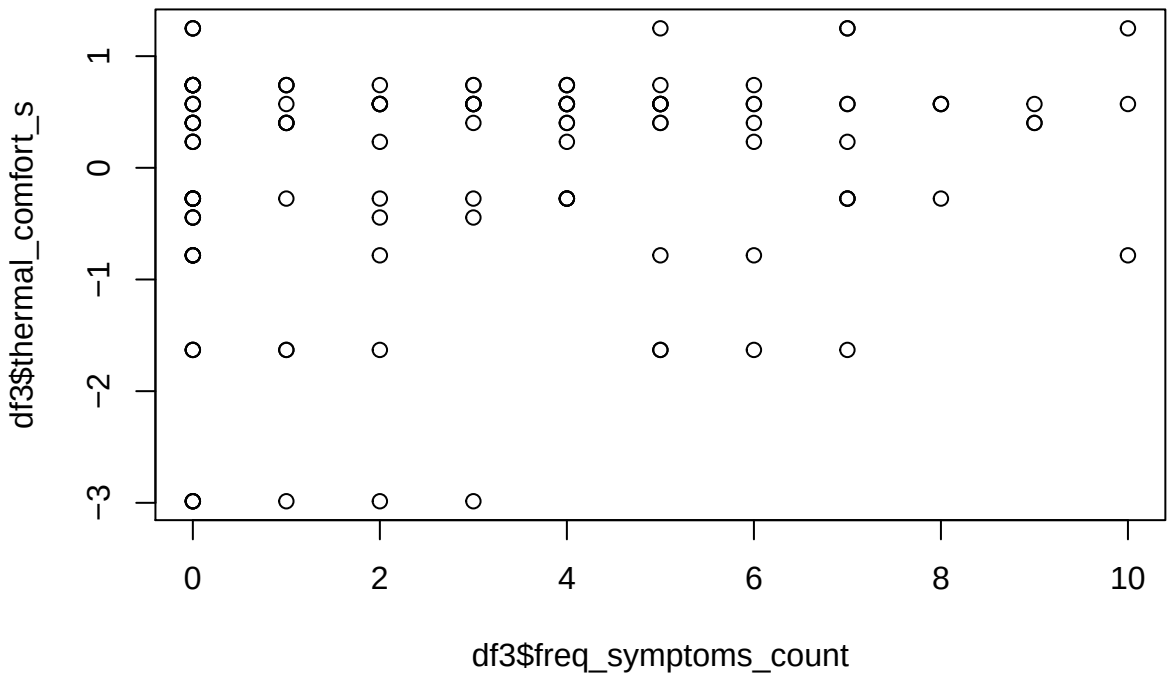



```
sym_effic <- glmer(freq_symptoms_count ~ energy_efficiency_s + job_teacher +
                     loc_classroom + (1|school),
                   family = poisson, data = df3)
summary(sym_effic)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
```

```
##  Family: poisson  ( log )
## Formula:
## freq_symptoms_count ~ energy_efficiency_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    704.4    718.9   -347.2    694.4      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0481 -1.4405 -0.3272  1.1636  5.4555
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.1107   0.3327
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)           0.7219     0.1495   4.829 1.37e-06 ***
## energy_efficiency_s   0.1052     0.1160   0.907    0.364
## job_teacher           0.1993     0.1516   1.315    0.189
## loc_classroom         0.1535     0.1645   0.933    0.351
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) enrg__ jb_tch
## enrgy_ffcn_  0.023
## job_teacher -0.163  0.054
## loc_classrm -0.330 -0.088 -0.658
```

```r
plot(df3$freq_symptoms_count, df3$thermal_comfort_s)
```

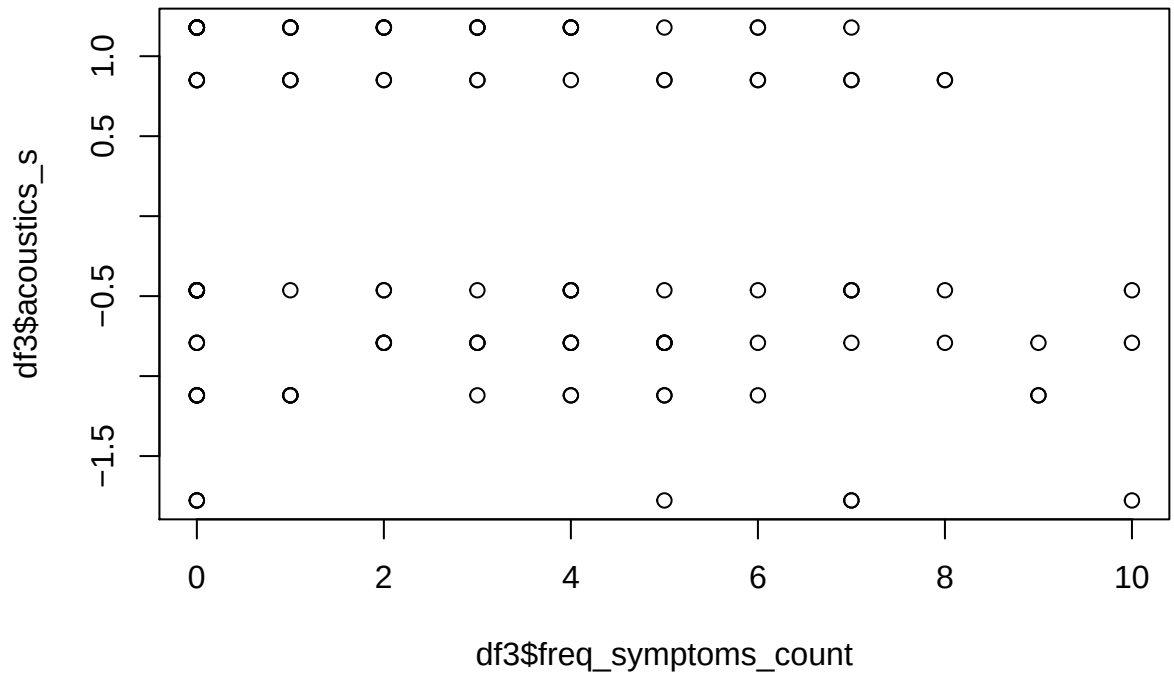



```
sym_therm <- glmer(freq_symptoms_count ~ thermal_comfort_s + job_teacher +
                     loc_classroom + (1|school),
                   family = poisson, data = df3)
summary(sym_therm)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## freq_symptoms_count ~ thermal_comfort_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    699.7    714.2   -344.8    689.7      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0703 -1.3928 -0.3667  1.0899  5.7869
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.06305  0.2511
## Number of obs: 134, groups:  school, 11
##
```

```
## Fixed effects:
##                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)         0.75322    0.13357   5.639 1.71e-08 ***
## thermal_comfort_s   0.23412    0.09543   2.453   0.0142 *
## job_teacher         0.17359    0.15096   1.150   0.2502
## loc_classroom       0.17659    0.16375   1.078   0.2808
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) thrm__ jb_tch
## thrml_cmfr_  0.046
## job_teacher -0.185 -0.007
## loc_classrm -0.377 -0.049 -0.656
```

```
plot(df3$freq_symptoms_count, df3$acoustics_s)
```

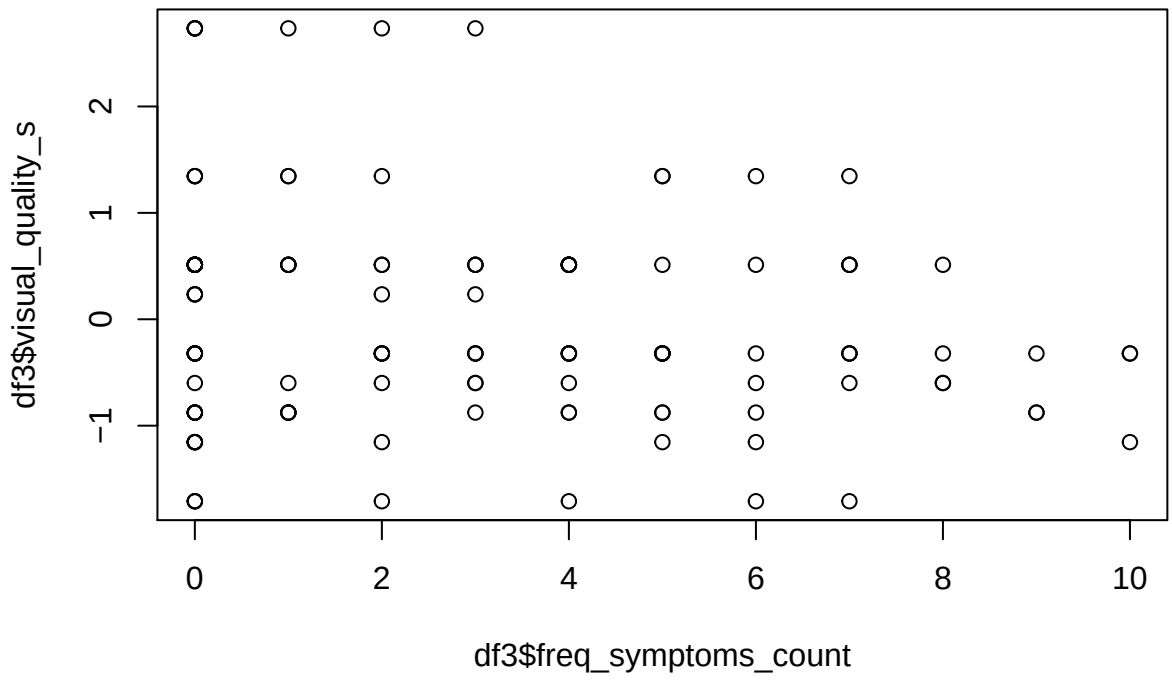



```
sym_sound <- glmer(freq_symptoms_count ~ acoustics_s + job_teacher +
                   loc_classroom + (1|school),
                   family = poisson, data = df3)
summary(sym_sound)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
```

```
## Formula:
## freq_symptoms_count ~ acoustics_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    700.3    714.8   -345.2    690.3      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.1043 -1.3790 -0.3399  1.1876  5.3007
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.06045  0.2459
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.74780    0.13261   5.639 1.71e-08 ***
## acoustics_s   -0.21975    0.09193  -2.390   0.0168 *
## job_teacher    0.16118    0.15198   1.061   0.2889
## loc_classroom  0.18998    0.16523   1.150   0.2502
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) acstc_ jb_tch
## acoustics_s -0.037
## job_teacher -0.191  0.030
## loc_classrm -0.371  0.034 -0.661
```

```
plot(df3$freq_symptoms_count, df3$visual_quality_s)
```

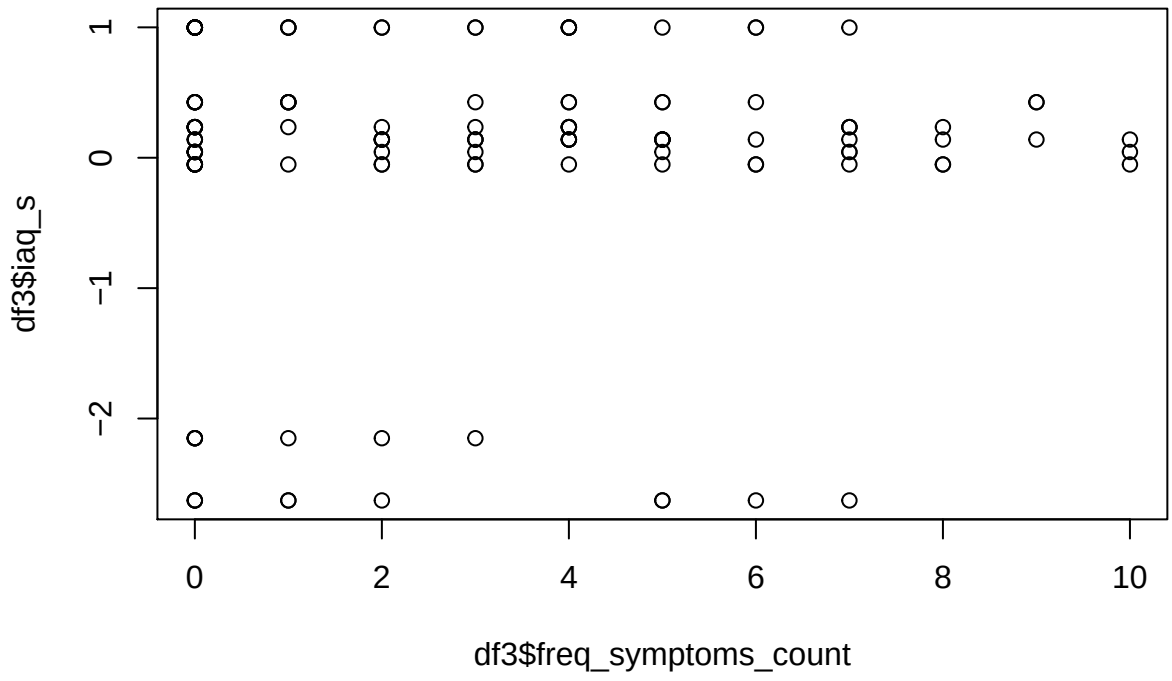



```
sym_visual <- glmer(freq_symptoms_count ~ visual_quality_s + job_teacher +
                      loc_classroom + (1|school),
                    family = poisson, data = df3)
summary(sym_visual)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## freq_symptoms_count ~ visual_quality_s + job_teacher + loc_classroom +
##     (1 | school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    700.4    714.9   -345.2    690.4      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0361 -1.4353 -0.3371  1.1530  5.1393
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.06404  0.2531
## Number of obs: 134, groups:  school, 11
##
```

```
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)        0.71720    0.13406   5.350  8.8e-08 ***
## visual_quality_s -0.21557    0.09267  -2.326    0.020 *
## job_teacher        0.16604    0.15183   1.094    0.274
## loc_classroom      0.19546    0.16587   1.178    0.239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) vsl_q_ jb_tch
## visl_qlty_s  0.063
## job_teacher -0.185  0.031
## loc_classrm -0.361  0.002 -0.661
```

```
plot(df3$freq_symptoms_count, df3$iaq_s)
```




```
sym_iaq <- glmer(freq_symptoms_count ~ iaq_s + job_teacher +
                   loc_classroom + (1|school),
                 family = poisson, data = df3)
summary(sym_iaq)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
```

```
## Formula: freq_symptoms_count ~ iaq_s + job_teacher + loc_classroom + (1 |
##     school)
##    Data: df3
##
##      AIC      BIC   logLik deviance df.resid
##    704.7    719.2   -347.3    694.7      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0492 -1.4314 -0.3396  1.1974  5.5224
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  school (Intercept) 0.1119   0.3345
## Number of obs: 134, groups:  school, 11
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.72690    0.15026   4.838 1.31e-06 ***
## iaq_s          0.08565    0.11171   0.767    0.443
## job_teacher    0.18428    0.15201   1.212    0.225
## loc_classroom  0.17038    0.16500   1.033    0.302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) iaq_s  jb_tch
## iaq_s        0.083
## job_teacher -0.171 -0.062
## loc_classrm -0.321  0.030 -0.660
```