

How to Manufacture Your Personality to Gain Friends and Second Dates: A Data-Driven Field Guide for Pathetic People by Pathetic People

Group members: Beau, Richard, Saad, Kevin

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

Importing the data

(credit: Beau and Richard)

Speed Dating dataset: <https://www.kaggle.com/annavictoria/speed-dating-experiment>

Each row of the data represents one date between two people. The primary key is one of the participants in the date, so for each date there are two rows in the dataset.

We imported the data, and then turned the appropriate columns into factors. The missing data came as blank strings, which we encoded as NAs.

For now, we are filtering out rows which do not have complete data (some participants did not include their hometown or undergraduate school, which were used as proxies to quantify wealth and education level). We are probably going to add these partially-full rows back into our analysis as we drop features from our model. TODO check if true

```
# This works if you do the following:
# stick the csv file in the same directory as this Rmd
# in RStudio, go to Session -> Set Work Directory -> To Source File location
# this expects only one csv to be in the directory
project.dir <- getwd()
setwd(project.dir)
csv_file <- list.files(file.path(project.dir), pattern="*.csv", full.names=TRUE)
full_data <- fread(file.path(csv_file))

full_data[,c("race_o", "field_cd", "race", "goal", "date", "go_out", "career_c", "met", "dec", "length",
#income, tuition, and mn_sat have NA values so they aren't numeric

nrow(full_data[complete.cases(full_data[, c("income", "tuition", "mn_sat")]),])

## [1] 8378

nrow(full_data)

## [1] 8378

#we could look at subset? is there a reason that people left off this information?

#full_data <- nrow(full_data[complete.cases(full_data[, c("income")]),])
#nrow(full_data[complete.cases(full_data[, c("mn_sat", "income")]),])
#nrow(full_data[complete.cases(full_data[, c("tuition", "income")]),])
```

Save/Load current state for easy continuation (credit: Kev)

```
#setwd()
#saveRDS(full_data, file="full_speed_dating.rds")
#full_data <- readRDS("full_speed_dating.rds")
```

Predicting rate of people liking you with constant fields

(credit: Beau and Richard)

The variable 'dec_o' encodes whether the other person in the date decided they would like a second date with you (1 for yes, 0 for no). We decided to use this as the response variable instead of 'match' because 'match' means that both people mutually liked each other. In particular, for each participant, we took the mean of 'dec_o' as a proxy for desirability.

We defined a new table 'const_attris' which represents the constant attributes of the participants. For example, from one date to another your age, field of study, and gender do not change. We wanted to see whether we can predict desirability from your non-subjective attributes.

However, we could not predict anything with this lm ($R^2 = 0.14$).

```
# rate that people decide they wanted to date you after the speed date
# aggregate certain variables for each individual (avg's of attractiveness, intelligence score ... etc. to

desirability <- aggregate(dec_o ~ iid, full_data, mean)

# create subset with the fields that remain constant for an individual (regardless of the meeting)
const_attris <- full_data[, c("iid", "gender", "age", "field_cd", "mn_sat", "tuition", "imprace", "income")]

# this resulting table has lots of duplicated data
# remove duplicates so there is only one obs per person
const_attris <- const_attris[!duplicated(const_attris),]

# binding back together w desirability
const_attris <- merge(desirability, const_attris, by="iid")

# merges list of dataframes into a single df. removes iid as a predictor.
full_const_fit <- lm(dec_o ~ .-iid, data = na.omit(const_attris)) # TODO find reduced model in git
```

After we reduced the insignificant predictors in this model, it had an adjusted R^2 of just 0.14.

Predicting rate of people liking you with constant fields and partner-rankings

We augmented this linear model using the rankings from your partners: how attractive, sincere, intelligent, fun, and ambitious your partners perceived you to be during the 4 minute speed-date. In particular, for each individual we took the mean rating by their partners.

```
#additional explanatory variables. note that after this point, attr_o, sinc_o, etc are
# now aggregated means
attractive <- aggregate(attr_o ~ iid, full_data, mean)
sincere <- aggregate(sinc_o ~ iid, full_data, mean)
```

```

intelligent <- aggregate(intel_o ~ iid, full_data, mean)
fun <- aggregate(fun_o ~ iid, full_data, mean)
ambitious <- aggregate(amb_o ~ iid, full_data, mean) # this was assigned incorrectly.. do any of the a
attractive <- aggregate(attr_o ~ iid, full_data, mean)
likeability <- aggregate(like_o ~ iid, full_data, mean)

augmented_attris <- Reduce(function(x, y) merge(x, y, all=TRUE), list(attractive, sincere, intelligent

```

We removed a few columns that had pesky numerical values. We will come back and probably add these back in after fixing that.

```

# remove fields that consist of mostly NA's
# these are tuition, SAT, and predicted income
augmented_attris <- augmented_attris[, -which(names(augmented_attris) %in% c("tuition", "mn_sat", "i
# fit a full model
# NOTE: we removed NAs, remember to add these rows back later if we reduce the model
full_augmented_fit <- lm(dec_o ~ .-iid, data = na.omit(augmented_attris))
summary(full_augmented_fit)

```

```

##
## Call:
## lm(formula = dec_o ~ . - iid, data = na.omit(augmented_attris))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36520 -0.08822 -0.00524  0.08182  0.42020
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.4284378  0.1370053  -3.127   0.0019 **
## attr_o       0.1205966  0.0087454  13.790 < 2e-16 ***
## sinc_o      -0.0186533  0.0143825  -1.297   0.1954
## intel_o      0.0103971  0.0214202   0.485   0.6277
## fun_o        0.0632598  0.0102762   6.156 1.85e-09 ***
## amb_o       -0.0041783  0.0153858  -0.272   0.7861
## gender1     -0.0423535  0.0193924  -2.184   0.0296 *
## age         -0.0031902  0.0020122  -1.585   0.1137
## field_cd2    -0.0153482  0.0709494  -0.216   0.8288
## field_cd3     0.0024820  0.0615646   0.040   0.9679
## field_cd4     0.0367259  0.0818159   0.449   0.6538
## field_cd5    -0.0126334  0.0596134  -0.212   0.8323
## field_cd6    -0.0051862  0.0677090  -0.077   0.9390
## field_cd7    -0.0047066  0.0652417  -0.072   0.9425
## field_cd8     0.0374003  0.0572785   0.653   0.5142
## field_cd9    -0.0527831  0.0614286  -0.859   0.3907
## field_cd10   -0.0212944  0.0599661  -0.355   0.7227
## field_cd11   -0.0750507  0.0681438  -1.101   0.2714
## field_cd12   -0.0942384  0.1590153  -0.593   0.5538
## field_cd13   -0.0741860  0.0602051  -1.232   0.2186
## field_cd14    0.0536081  0.0827754   0.648   0.5176
## field_cd15    0.0252545  0.0764637   0.330   0.7414
## field_cd16    0.2532089  0.0923661   2.741   0.0064 **
## field_cd17   -0.0633854  0.1475603  -0.430   0.6678
## field_cd18   -0.0232298  0.1168520  -0.199   0.8425

```

```

## imprace      -0.0009147  0.0026118  -0.350  0.7264
## goal2        -0.0085901  0.0166963  -0.514  0.6072
## goal3         0.0133193  0.0264685   0.503  0.6151
## goal4        -0.0073760  0.0391689  -0.188  0.8507
## goal5         0.0029902  0.0303432   0.099  0.9215
## goal6        -0.0087543  0.0331656  -0.264  0.7920
## date2        -0.0784799  0.0725368  -1.082  0.2799
## date3        -0.0825831  0.0684856  -1.206  0.2286
## date4        -0.0583368  0.0669411  -0.871  0.3840
## date5        -0.0797517  0.0677092  -1.178  0.2396
## date6        -0.0817761  0.0673008  -1.215  0.2251
## date7        -0.0882932  0.0683041  -1.293  0.1969
## go_out2       0.0303651  0.0179792   1.689  0.0920 .
## go_out3       0.0210875  0.0197713   1.067  0.2868
## go_out4      -0.0082647  0.0343693  -0.240  0.8101
## go_out5      -0.0021154  0.0506022  -0.042  0.9667
## go_out6       0.1038222  0.0646569   1.606  0.1091
## go_out7      -0.0945670  0.1498431  -0.631  0.5283
## career_c2     0.0421632  0.0565605   0.745  0.4564
## career_c3    -0.1105502  0.0691739  -1.598  0.1108
## career_c4     0.0084988  0.0665085   0.128  0.8984
## career_c5     0.0013382  0.0711746   0.019  0.9850
## career_c6    -0.0396848  0.0622756  -0.637  0.5243
## career_c7    -0.0053737  0.0560822  -0.096  0.9237
## career_c8    -0.4795132  0.1173762  -4.085  5.34e-05 ***
## career_c9     0.0488562  0.0620896   0.787  0.4318
## career_c10    0.0155325  0.0630675   0.246  0.8056
## career_c11    0.0224088  0.0773579   0.290  0.7722
## career_c12   -0.2297543  0.1196058  -1.921  0.0555 .
## career_c13    0.1103414  0.0870643   1.267  0.2058
## career_c14    0.2365744  0.1488698   1.589  0.1128
## career_c15    0.1449811  0.1089465   1.331  0.1840
## career_c16    0.0309584  0.1218729   0.254  0.7996
## career_c17      NA         NA         NA         NA
## match_es     -0.0011390  0.0031113  -0.366  0.7145
## sports       -0.0014692  0.0036336  -0.404  0.6862
## tvsports     -0.0073857  0.0030830  -2.396  0.0171 *
## exercise      0.0028403  0.0033081   0.859  0.3911
## dining       -0.0003931  0.0045621  -0.086  0.9314
## museums       0.0059727  0.0070519   0.847  0.3975
## art           0.0003190  0.0062680   0.051  0.9594
## hiking        0.0026801  0.0031052   0.863  0.3886
## gaming        0.0015802  0.0031341   0.504  0.6144
## clubbing      0.0008773  0.0030188   0.291  0.7715
## reading      -0.0033231  0.0038386  -0.866  0.3872
## tv            0.0032676  0.0035806   0.913  0.3620
## theater      -0.0008898  0.0045579  -0.195  0.8453
## movies       -0.0043294  0.0051919  -0.834  0.4049
## concerts      0.0043669  0.0046875   0.932  0.3521
## music        -0.0034101  0.0050617  -0.674  0.5009
## shopping     -0.0019802  0.0034887  -0.568  0.5706
## yoga         -0.0063201  0.0028292  -2.234  0.0261 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
##
## Residual standard error: 0.1391 on 391 degrees of freedom
## Multiple R-squared: 0.7086, Adjusted R-squared: 0.6527
## F-statistic: 12.68 on 75 and 391 DF, p-value: < 2.2e-16
```

The full model has an adj-R² of 0.65 on 75 predictor variables. In order to reduce our model, we ran a stepwise AIC.

```
library(MASS)
step <- stepAIC(full_augmented_fit, direction="both")
```

```
## Start: AIC=-1773.21
## dec_o ~ (iid + attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + field_cd + imprace + goal + date + go_out + career_c +
##      match_es + sports + tvsports + exercise + dining + museums +
##      art + hiking + gaming + clubbing + reading + tv + theater +
##      movies + concerts + music + shopping + yoga) - iid
##
##      Df Sum of Sq    RSS    AIC
## - goal      5    0.0154  7.5826 -1782.3
## - date      6    0.0700  7.6372 -1780.9
## - go_out    6    0.1213  7.6885 -1777.8
## - art       1    0.0001  7.5673 -1775.2
## - dining    1    0.0001  7.5673 -1775.2
## - theater   1    0.0007  7.5679 -1775.2
## - amb_o     1    0.0014  7.5686 -1775.1
## - clubbing  1    0.0016  7.5688 -1775.1
## - imprace   1    0.0024  7.5696 -1775.1
## - match_es  1    0.0026  7.5698 -1775.0
## - sports    1    0.0032  7.5704 -1775.0
## - intel_o   1    0.0046  7.5718 -1774.9
## - gaming    1    0.0049  7.5721 -1774.9
## - shopping  1    0.0062  7.5734 -1774.8
## - music     1    0.0088  7.5760 -1774.7
## - movies    1    0.0135  7.5807 -1774.4
## - museums   1    0.0139  7.5811 -1774.3
## - exercise  1    0.0143  7.5815 -1774.3
## - hiking    1    0.0144  7.5816 -1774.3
## - reading   1    0.0145  7.5817 -1774.3
## - tv        1    0.0161  7.5833 -1774.2
## - concerts  1    0.0168  7.5840 -1774.2
## <none>              7.5672 -1773.2
## - sinc_o     1    0.0326  7.5998 -1773.2
## - age        1    0.0486  7.6158 -1772.2
## - field_cd   16   0.5554  8.1226 -1772.1
## - gender     1    0.0923  7.6595 -1769.5
## - yoga       1    0.0966  7.6638 -1769.3
## - tvsports   1    0.1111  7.6783 -1768.4
## - career_c   15   0.9183  8.4855 -1749.7
## - fun_o      1    0.7334  8.3006 -1732.0
## - attr_o     1    3.6802 11.2474 -1590.1
##
## Step: AIC=-1782.26
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + field_cd + imprace + date + go_out + career_c + match_es +
```

```

##      sports + tvsports + exercise + dining + museums + art + hiking +
##      gaming + clubbing + reading + tv + theater + movies + concerts +
##      music + shopping + yoga
##
##      Df Sum of Sq      RSS      AIC
## - date      6      0.0748  7.6574 -1789.7
## - go_out     6      0.1271  7.7097 -1786.5
## - dining     1      0.0000  7.5826 -1784.3
## - art        1      0.0001  7.5827 -1784.3
## - theater    1      0.0005  7.5831 -1784.2
## - imprace    1      0.0021  7.5847 -1784.1
## - amb_o      1      0.0022  7.5848 -1784.1
## - clubbing   1      0.0024  7.5850 -1784.1
## - sports     1      0.0026  7.5852 -1784.1
## - match_es   1      0.0040  7.5866 -1784.0
## - intel_o    1      0.0044  7.5870 -1784.0
## - shopping   1      0.0045  7.5871 -1784.0
## - gaming     1      0.0048  7.5874 -1784.0
## - music      1      0.0073  7.5899 -1783.8
## - exercise   1      0.0115  7.5941 -1783.5
## - museums    1      0.0120  7.5946 -1783.5
## - movies     1      0.0125  7.5951 -1783.5
## - reading    1      0.0130  7.5956 -1783.5
## - concerts   1      0.0133  7.5959 -1783.4
## - tv         1      0.0142  7.5968 -1783.4
## - hiking     1      0.0145  7.5971 -1783.4
## - sinc_o     1      0.0283  7.6109 -1782.5
## <none>              7.5826 -1782.3
## - field_cd  16      0.5518  8.1344 -1781.5
## - age        1      0.0560  7.6386 -1780.8
## - gender     1      0.0853  7.6679 -1779.0
## - yoga       1      0.0930  7.6756 -1778.6
## - tvsports   1      0.1101  7.6927 -1777.5
## + goal       5      0.0154  7.5672 -1773.2
## - career_c  15      0.9364  8.5190 -1757.9
## - fun_o      1      0.7609  8.3435 -1739.6
## - attr_o     1      3.6985 11.2811 -1598.7
##
## Step:  AIC=-1789.67
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + field_cd + imprace + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + hiking +
##      gaming + clubbing + reading + tv + theater + movies + concerts +
##      music + shopping + yoga
##
##      Df Sum of Sq      RSS      AIC
## - go_out     6      0.1261  7.7835 -1794.0
## - dining     1      0.0002  7.6576 -1791.7
## - theater    1      0.0003  7.6577 -1791.7
## - art        1      0.0006  7.6580 -1791.6
## - amb_o      1      0.0007  7.6581 -1791.6
## - sports     1      0.0007  7.6582 -1791.6
## - intel_o    1      0.0020  7.6594 -1791.5
## - shopping   1      0.0023  7.6597 -1791.5

```

```

## - imprace 1 0.0023 7.6598 -1791.5
## - clubbing 1 0.0025 7.6600 -1791.5
## - match_es 1 0.0031 7.6605 -1791.5
## - gaming 1 0.0054 7.6628 -1791.3
## - museums 1 0.0090 7.6664 -1791.1
## - music 1 0.0105 7.6679 -1791.0
## - exercise 1 0.0114 7.6688 -1791.0
## - reading 1 0.0118 7.6692 -1791.0
## - hiking 1 0.0121 7.6696 -1790.9
## - tv 1 0.0136 7.6710 -1790.8
## - movies 1 0.0144 7.6719 -1790.8
## - concerts 1 0.0145 7.6719 -1790.8
## - sinc_o 1 0.0227 7.6802 -1790.3
## - field_cd 16 0.5365 8.1939 -1790.0
## <none> 7.6574 -1789.7
## - age 1 0.0571 7.7146 -1788.2
## - gender 1 0.0777 7.7352 -1787.0
## - yoga 1 0.0908 7.7482 -1786.2
## - tvsports 1 0.1248 7.7823 -1784.1
## + date 6 0.0748 7.5826 -1782.3
## + goal 5 0.0202 7.6372 -1780.9
## - career_c 15 0.9307 8.5881 -1766.1
## - fun_o 1 0.7545 8.4119 -1747.8
## - attr_o 1 3.7487 11.4061 -1605.6
##
## Step: AIC=-1794.05
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
## age + field_cd + imprace + career_c + match_es + sports +
## tvsports + exercise + dining + museums + art + hiking + gaming +
## clubbing + reading + tv + theater + movies + concerts + music +
## shopping + yoga
##
## Df Sum of Sq RSS AIC
## - field_cd 16 0.5162 8.2997 -1796.1
## - amb_o 1 0.0001 7.7836 -1796.0
## - dining 1 0.0002 7.7837 -1796.0
## - intel_o 1 0.0003 7.7838 -1796.0
## - art 1 0.0006 7.7841 -1796.0
## - theater 1 0.0014 7.7849 -1796.0
## - sports 1 0.0015 7.7851 -1796.0
## - match_es 1 0.0024 7.7859 -1795.9
## - imprace 1 0.0032 7.7867 -1795.9
## - clubbing 1 0.0042 7.7877 -1795.8
## - shopping 1 0.0049 7.7884 -1795.8
## - gaming 1 0.0062 7.7897 -1795.7
## - exercise 1 0.0075 7.7910 -1795.6
## - music 1 0.0081 7.7916 -1795.6
## - movies 1 0.0092 7.7927 -1795.5
## - hiking 1 0.0111 7.7946 -1795.4
## - museums 1 0.0115 7.7950 -1795.4
## - concerts 1 0.0116 7.7951 -1795.3
## - reading 1 0.0154 7.7989 -1795.1
## - tv 1 0.0158 7.7993 -1795.1
## - sinc_o 1 0.0182 7.8017 -1795.0

```

```

## <none>          7.7835 -1794.0
## - age          1    0.0496 7.8331 -1793.1
## - gender       1    0.0692 7.8527 -1791.9
## - yoga         1    0.0921 7.8756 -1790.6
## + go_out       6    0.1261 7.6574 -1789.7
## - tvsports     1    0.1133 7.8968 -1789.3
## + date         6    0.0738 7.7097 -1786.5
## + goal         5    0.0233 7.7602 -1785.5
## - career_c    15    0.9516 8.7351 -1770.2
## - fun_o        1    0.7522 8.5358 -1753.0
## - attr_o       1    3.8772 11.6607 -1607.3
##
## Step:  AIC=-1796.06
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + imprace + career_c + match_es + sports + tvsports +
##      exercise + dining + museums + art + hiking + gaming + clubbing +
##      reading + tv + theater + movies + concerts + music + shopping +
##      yoga
##
##           Df Sum of Sq    RSS    AIC
## - amb_o     1     0.0000  8.2997 -1798.1
## - dining     1     0.0001  8.2998 -1798.1
## - imprace    1     0.0006  8.3003 -1798.0
## - sports     1     0.0010  8.3007 -1798.0
## - hiking     1     0.0013  8.3010 -1798.0
## - clubbing   1     0.0019  8.3016 -1798.0
## - tv         1     0.0020  8.3018 -1798.0
## - intel_o    1     0.0029  8.3026 -1797.9
## - museums    1     0.0032  8.3029 -1797.9
## - movies     1     0.0041  8.3038 -1797.8
## - reading    1     0.0057  8.3054 -1797.7
## - shopping   1     0.0073  8.3070 -1797.7
## - exercise   1     0.0081  8.3078 -1797.6
## - match_es   1     0.0084  8.3081 -1797.6
## - art        1     0.0087  8.3084 -1797.6
## - theater    1     0.0087  8.3085 -1797.6
## - gaming     1     0.0144  8.3141 -1797.2
## - music      1     0.0257  8.3255 -1796.6
## - concerts   1     0.0277  8.3274 -1796.5
## <none>          8.2997 -1796.1
## - sinc_o     1     0.0399  8.3396 -1795.8
## - age        1     0.0452  8.3449 -1795.5
## - gender     1     0.0683  8.3680 -1794.2
## + field_cd   16     0.5162  7.7835 -1794.0
## - yoga       1     0.0833  8.3830 -1793.4
## - tvsports   1     0.1073  8.4070 -1792.1
## + go_out     6     0.1058  8.1939 -1790.0
## + date       6     0.0618  8.2379 -1787.5
## + goal       5     0.0115  8.2882 -1786.7
## - career_c   16     0.8376  9.1374 -1783.2
## - fun_o      1     0.8241  9.1239 -1753.8
## - attr_o     1     4.0443 12.3440 -1612.7
##
## Step:  AIC=-1798.06

```



```

## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + imprace +
##   career_c + match_es + sports + tvsports + exercise + dining +
##   museums + art + hiking + gaming + clubbing + reading + tv +
##   theater + movies + concerts + music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC
## - dining    1    0.0001  8.2998 -1800.1
## - imprace    1    0.0006  8.3003 -1800.0
## - sports     1    0.0010  8.3007 -1800.0
## - hiking     1    0.0013  8.3010 -1800.0
## - clubbing   1    0.0019  8.3016 -1800.0
## - tv         1    0.0021  8.3018 -1799.9
## - museums    1    0.0032  8.3029 -1799.9
## - movies     1    0.0041  8.3038 -1799.8
## - intel_o    1    0.0049  8.3046 -1799.8
## - reading    1    0.0057  8.3054 -1799.7
## - shopping   1    0.0073  8.3070 -1799.7
## - exercise   1    0.0081  8.3079 -1799.6
## - match_es   1    0.0084  8.3081 -1799.6
## - art        1    0.0086  8.3084 -1799.6
## - theater    1    0.0088  8.3085 -1799.6
## - gaming     1    0.0144  8.3141 -1799.2
## - music      1    0.0257  8.3255 -1798.6
## - concerts   1    0.0277  8.3274 -1798.5
## <none>              8.2997 -1798.1
## - sinc_o     1    0.0416  8.3413 -1797.7
## - age        1    0.0454  8.3451 -1797.5
## - gender     1    0.0704  8.3701 -1796.1
## + amb_o      1    0.0000  8.2997 -1796.1
## + field_cd   16    0.5161  7.7836 -1796.0
## - yoga       1    0.0838  8.3835 -1795.4
## - tvsports   1    0.1074  8.4071 -1794.1
## + go_out     6    0.1057  8.1940 -1792.0
## + date       6    0.0617  8.2380 -1789.5
## + goal       5    0.0115  8.2882 -1788.7
## - career_c   16    0.8396  9.1393 -1785.1
## - fun_o      1    0.8687  9.1684 -1753.6
## - attr_o     1    4.0448 12.3445 -1614.7
##
## Step:  AIC=-1800.06
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + imprace +
##   career_c + match_es + sports + tvsports + exercise + museums +
##   art + hiking + gaming + clubbing + reading + tv + theater +
##   movies + concerts + music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC
## - imprace    1    0.0006  8.3003 -1802.0
## - sports     1    0.0010  8.3007 -1802.0
## - hiking     1    0.0013  8.3010 -1802.0
## - clubbing   1    0.0019  8.3016 -1802.0
## - tv         1    0.0021  8.3019 -1801.9
## - museums    1    0.0031  8.3029 -1801.9
## - movies     1    0.0041  8.3039 -1801.8
## - intel_o    1    0.0048  8.3046 -1801.8

```

```

## - reading 1 0.0057 8.3055 -1801.7
## - exercise 1 0.0081 8.3079 -1801.6
## - shopping 1 0.0085 8.3083 -1801.6
## - match_es 1 0.0085 8.3083 -1801.6
## - art 1 0.0087 8.3085 -1801.6
## - theater 1 0.0088 8.3086 -1801.6
## - gaming 1 0.0144 8.3141 -1801.2
## - music 1 0.0259 8.3257 -1800.6
## - concerts 1 0.0277 8.3275 -1800.5
## <none> 8.2998 -1800.1
## - sinc_o 1 0.0417 8.3414 -1799.7
## - age 1 0.0458 8.3456 -1799.5
## - gender 1 0.0704 8.3701 -1798.1
## + dining 1 0.0001 8.2997 -1798.1
## + amb_o 1 0.0000 8.2998 -1798.1
## + field_cd 16 0.5160 7.7837 -1798.0
## - yoga 1 0.0839 8.3837 -1797.4
## - tvsports 1 0.1073 8.4071 -1796.1
## + go_out 6 0.1057 8.1940 -1794.0
## + date 6 0.0610 8.2388 -1791.5
## + goal 5 0.0113 8.2885 -1790.7
## - career_c 16 0.8402 9.1400 -1787.0
## - fun_o 1 0.8712 9.1710 -1755.4
## - attr_o 1 4.0455 12.3453 -1616.6
##
## Step: AIC=-1802.03
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
## match_es + sports + tvsports + exercise + museums + art +
## hiking + gaming + clubbing + reading + tv + theater + movies +
## concerts + music + shopping + yoga
##
## Df Sum of Sq RSS AIC
## - sports 1 0.0009 8.3013 -1804.0
## - hiking 1 0.0012 8.3015 -1804.0
## - tv 1 0.0017 8.3021 -1803.9
## - clubbing 1 0.0021 8.3024 -1803.9
## - museums 1 0.0030 8.3034 -1803.9
## - movies 1 0.0041 8.3045 -1803.8
## - intel_o 1 0.0048 8.3051 -1803.8
## - reading 1 0.0055 8.3058 -1803.7
## - exercise 1 0.0078 8.3082 -1803.6
## - match_es 1 0.0082 8.3086 -1803.6
## - shopping 1 0.0088 8.3092 -1803.5
## - art 1 0.0089 8.3092 -1803.5
## - theater 1 0.0091 8.3094 -1803.5
## - gaming 1 0.0143 8.3147 -1803.2
## - music 1 0.0259 8.3262 -1802.6
## - concerts 1 0.0282 8.3285 -1802.4
## <none> 8.3003 -1802.0
## - sinc_o 1 0.0415 8.3418 -1801.7
## - age 1 0.0453 8.3456 -1801.5
## - gender 1 0.0698 8.3701 -1800.1
## + imprace 1 0.0006 8.2998 -1800.1
## + dining 1 0.0001 8.3003 -1800.0

```

```

## + amb_o      1      0.0000  8.3003 -1800.0
## + field_cd 16      0.5134  7.7869 -1799.8
## - yoga       1      0.0834  8.3838 -1799.4
## - tvsports   1      0.1071  8.4075 -1798.0
## + go_out      6      0.1061  8.1942 -1796.0
## + date        6      0.0609  8.2395 -1793.5
## + goal        5      0.0110  8.2893 -1792.7
## - career_c 16      0.8397  9.1400 -1789.0
## - fun_o       1      0.8718  9.1722 -1757.4
## - attr_o      1      4.0449 12.3453 -1618.6
##
## Step:  AIC=-1803.97
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
##      match_es + tvsports + exercise + museums + art + hiking +
##      gaming + clubbing + reading + tv + theater + movies + concerts +
##      music + shopping + yoga
##
##           Df Sum of Sq      RSS      AIC
## - hiking    1      0.0009  8.3021 -1805.9
## - tv         1      0.0019  8.3032 -1805.9
## - clubbing   1      0.0021  8.3034 -1805.9
## - museums    1      0.0029  8.3042 -1805.8
## - movies     1      0.0042  8.3054 -1805.7
## - intel_o    1      0.0044  8.3057 -1805.7
## - reading    1      0.0053  8.3065 -1805.7
## - exercise   1      0.0069  8.3082 -1805.6
## - theater    1      0.0086  8.3099 -1805.5
## - shopping   1      0.0088  8.3101 -1805.5
## - match_es   1      0.0089  8.3102 -1805.5
## - art        1      0.0091  8.3104 -1805.5
## - gaming     1      0.0143  8.3156 -1805.2
## - music      1      0.0256  8.3269 -1804.5
## - concerts   1      0.0278  8.3290 -1804.4
## <none>              8.3013 -1804.0
## - sinc_o     1      0.0410  8.3423 -1803.7
## - age        1      0.0451  8.3463 -1803.4
## + sports     1      0.0009  8.3003 -1802.0
## + imprace    1      0.0005  8.3007 -1802.0
## + dining     1      0.0000  8.3012 -1802.0
## + amb_o      1      0.0000  8.3013 -1802.0
## + field_cd 16      0.5130  7.7883 -1801.8
## - gender     1      0.0760  8.3773 -1801.7
## - yoga       1      0.0834  8.3847 -1801.3
## - tvsports   1      0.1422  8.4435 -1798.0
## + go_out      6      0.1067  8.1946 -1798.0
## + date        6      0.0589  8.2424 -1795.3
## + goal        5      0.0107  8.2906 -1794.6
## - career_c 16      0.8401  9.1413 -1791.0
## - fun_o       1      0.8767  9.1780 -1759.1
## - attr_o      1      4.1174 12.4187 -1617.9
##
## Step:  AIC=-1805.92
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
##      match_es + tvsports + exercise + museums + art + gaming +

```

```

##      clubbing + reading + tv + theater + movies + concerts + music +
##      shopping + yoga
##
##      Df Sum of Sq      RSS      AIC
## - tv      1      0.0017  8.3038 -1807.8
## - clubbing 1      0.0018  8.3040 -1807.8
## - museums  1      0.0032  8.3053 -1807.8
## - movies   1      0.0041  8.3062 -1807.7
## - intel_o  1      0.0044  8.3066 -1807.7
## - reading  1      0.0055  8.3076 -1807.6
## - exercise 1      0.0076  8.3098 -1807.5
## - match_es 1      0.0086  8.3108 -1807.4
## - theater  1      0.0086  8.3108 -1807.4
## - art       1      0.0092  8.3114 -1807.4
## - shopping 1      0.0100  8.3121 -1807.4
## - gaming    1      0.0176  8.3197 -1806.9
## - music     1      0.0270  8.3292 -1806.4
## - concerts  1      0.0289  8.3310 -1806.3
## <none>              8.3021 -1805.9
## - sinc_o    1      0.0403  8.3424 -1805.7
## - age       1      0.0446  8.3467 -1805.4
## + hiking    1      0.0009  8.3013 -1804.0
## + sports    1      0.0006  8.3015 -1804.0
## + imprace   1      0.0005  8.3017 -1804.0
## + dining    1      0.0000  8.3021 -1803.9
## + amb_o     1      0.0000  8.3021 -1803.9
## - gender    1      0.0794  8.3815 -1803.5
## - yoga      1      0.0831  8.3853 -1803.3
## + field_cd  16      0.5043  7.7978 -1803.2
## - tvsports  1      0.1423  8.4444 -1800.0
## + go_out    6      0.1062  8.1959 -1799.9
## + date      6      0.0584  8.2437 -1797.2
## + goal      5      0.0110  8.2911 -1796.5
## - career_c  16      0.8534  9.1555 -1792.2
## - fun_o     1      0.8759  9.1780 -1761.1
## - attr_o    1      4.1183 12.4205 -1619.8
##
## Step:  AIC=-1807.83
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
##      match_es + tvsports + exercise + museums + art + gaming +
##      clubbing + reading + theater + movies + concerts + music +
##      shopping + yoga
##
##      Df Sum of Sq      RSS      AIC
## - clubbing 1      0.0015  8.3053 -1809.8
## - museums  1      0.0029  8.3067 -1809.7
## - movies   1      0.0029  8.3067 -1809.7
## - intel_o  1      0.0042  8.3080 -1809.6
## - reading  1      0.0053  8.3091 -1809.5
## - exercise 1      0.0078  8.3116 -1809.4
## - theater  1      0.0084  8.3122 -1809.4
## - shopping 1      0.0084  8.3122 -1809.4
## - art       1      0.0092  8.3130 -1809.3
## - match_es  1      0.0093  8.3131 -1809.3

```

```

## - gaming      1      0.0196  8.3234 -1808.7
## - music       1      0.0270  8.3308 -1808.3
## - concerts    1      0.0276  8.3314 -1808.3
## <none>                                8.3038 -1807.8
## - sinc_o      1      0.0395  8.3433 -1807.6
## - age         1      0.0455  8.3493 -1807.3
## + tv          1      0.0017  8.3021 -1805.9
## + sports      1      0.0008  8.3030 -1805.9
## + hiking      1      0.0006  8.3032 -1805.9
## + imprace     1      0.0002  8.3036 -1805.8
## + dining      1      0.0000  8.3038 -1805.8
## + amb_o       1      0.0000  8.3038 -1805.8
## - yoga        1      0.0828  8.3866 -1805.2
## - gender      1      0.0829  8.3867 -1805.2
## + field_cd    16      0.4946  7.8092 -1804.5
## + go_out      6      0.1068  8.1970 -1801.9
## - tvsports    1      0.1444  8.4482 -1801.8
## + date        6      0.0581  8.2457 -1799.1
## + goal        5      0.0106  8.2932 -1798.4
## - career_c    16      0.8527  9.1565 -1794.2
## - fun_o       1      0.8760  9.1798 -1763.0
## - attr_o      1      4.1612 12.4650 -1620.1
##
## Step:  AIC=-1809.75
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
##      match_es + tvsports + exercise + museums + art + gaming +
##      reading + theater + movies + concerts + music + shopping +
##      yoga
##
##           Df Sum of Sq    RSS    AIC
## - museums   1      0.0029  8.3082 -1811.6
## - movies    1      0.0029  8.3082 -1811.6
## - intel_o   1      0.0040  8.3093 -1811.5
## - reading   1      0.0059  8.3112 -1811.4
## - exercise  1      0.0075  8.3128 -1811.3
## - theater   1      0.0080  8.3133 -1811.3
## - shopping  1      0.0080  8.3133 -1811.3
## - match_es  1      0.0085  8.3138 -1811.3
## - art       1      0.0096  8.3149 -1811.2
## - gaming    1      0.0220  8.3273 -1810.5
## - music     1      0.0267  8.3320 -1810.2
## - concerts  1      0.0276  8.3329 -1810.2
## <none>                                8.3053 -1809.8
## - sinc_o    1      0.0395  8.3448 -1809.5
## - age       1      0.0462  8.3515 -1809.2
## + clubbing  1      0.0015  8.3038 -1807.8
## + tv        1      0.0013  8.3040 -1807.8
## + sports    1      0.0008  8.3045 -1807.8
## + hiking    1      0.0004  8.3049 -1807.8
## + imprace   1      0.0003  8.3050 -1807.8
## + amb_o     1      0.0000  8.3053 -1807.8
## + dining    1      0.0000  8.3053 -1807.8
## - yoga      1      0.0833  8.3886 -1807.1
## - gender    1      0.0854  8.3907 -1807.0

```

```

## + field_cd 16    0.4937  7.8116 -1806.4
## + go_out      6    0.1079  8.1974 -1803.8
## - tvsports   1    0.1438  8.4491 -1803.7
## + date       6    0.0581  8.2472 -1801.0
## + goal       5    0.0109  8.2944 -1800.4
## - career_c 16    0.8574  9.1627 -1795.9
## - fun_o      1    0.8880  9.1934 -1764.3
## - attr_o     1    4.1791 12.4844 -1621.4
##
## Step:  AIC=-1811.58
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
##      match_es + tvsports + exercise + art + gaming + reading +
##      theater + movies + concerts + music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC
## - movies   1     0.0031  8.3113 -1813.4
## - intel_o   1     0.0042  8.3124 -1813.3
## - reading   1     0.0044  8.3126 -1813.3
## - theater   1     0.0067  8.3149 -1813.2
## - shopping  1     0.0072  8.3154 -1813.2
## - exercise  1     0.0074  8.3156 -1813.2
## - match_es  1     0.0094  8.3176 -1813.0
## - gaming    1     0.0211  8.3293 -1812.4
## - music     1     0.0266  8.3348 -1812.1
## - concerts  1     0.0276  8.3358 -1812.0
## <none>                        8.3082 -1811.6
## - sinc_o    1     0.0396  8.3478 -1811.4
## - age       1     0.0454  8.3536 -1811.0
## - art       1     0.0504  8.3586 -1810.8
## + museums   1     0.0029  8.3053 -1809.8
## + clubbing  1     0.0015  8.3067 -1809.7
## + tv        1     0.0011  8.3071 -1809.6
## + hiking    1     0.0006  8.3076 -1809.6
## + sports    1     0.0006  8.3076 -1809.6
## + imprace   1     0.0002  8.3080 -1809.6
## + amb_o     1     0.0001  8.3081 -1809.6
## + dining    1     0.0000  8.3082 -1809.6
## - yoga      1     0.0833  8.3915 -1808.9
## - gender    1     0.0846  8.3928 -1808.8
## + field_cd 16    0.4857  7.8225 -1807.7
## + go_out    6     0.1087  8.1995 -1805.7
## - tvsports  1     0.1410  8.4492 -1805.7
## + date      6     0.0573  8.2509 -1802.8
## + goal      5     0.0105  8.2977 -1802.2
## - career_c 16    0.8587  9.1669 -1797.7
## - fun_o     1     0.8853  9.1935 -1766.3
## - attr_o    1     4.1923 12.5005 -1622.8
##
## Step:  AIC=-1813.41
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + gender + age + career_c +
##      match_es + tvsports + exercise + art + gaming + reading +
##      theater + concerts + music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC

```

```

## - intel_o 1 0.0037 8.3150 -1815.2
## - reading 1 0.0040 8.3152 -1815.2
## - exercise 1 0.0078 8.3191 -1815.0
## - shopping 1 0.0082 8.3194 -1815.0
## - match_es 1 0.0090 8.3203 -1814.9
## - theater 1 0.0133 8.3246 -1814.7
## - gaming 1 0.0212 8.3325 -1814.2
## - concerts 1 0.0263 8.3375 -1813.9
## - music 1 0.0297 8.3410 -1813.7
## <none> 8.3113 -1813.4
## - sinc_o 1 0.0388 8.3501 -1813.2
## - age 1 0.0451 8.3563 -1812.9
## - art 1 0.0502 8.3615 -1812.6
## + movies 1 0.0031 8.3082 -1811.6
## + museums 1 0.0031 8.3082 -1811.6
## + clubbing 1 0.0015 8.3098 -1811.5
## + hiking 1 0.0007 8.3106 -1811.5
## + sports 1 0.0006 8.3107 -1811.4
## + imprace 1 0.0003 8.3110 -1811.4
## + tv 1 0.0002 8.3111 -1811.4
## + amb_o 1 0.0001 8.3112 -1811.4
## + dining 1 0.0000 8.3113 -1811.4
## - yoga 1 0.0822 8.3935 -1810.8
## - gender 1 0.0857 8.3970 -1810.6
## + field_cd 16 0.4847 7.8266 -1809.5
## - tvsports 1 0.1407 8.4520 -1807.6
## + go_out 6 0.1052 8.2061 -1807.4
## + date 6 0.0581 8.2532 -1804.7
## + goal 5 0.0104 8.3008 -1804.0
## - career_c 16 0.8724 9.1837 -1798.8
## - fun_o 1 0.8954 9.2067 -1767.6
## - attr_o 1 4.1893 12.5006 -1624.8
##
## Step: AIC=-1815.2
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
## tvsports + exercise + art + gaming + reading + theater +
## concerts + music + shopping + yoga
##
## Df Sum of Sq RSS AIC
## - reading 1 0.0037 8.3187 -1817.0
## - exercise 1 0.0075 8.3225 -1816.8
## - match_es 1 0.0085 8.3234 -1816.7
## - shopping 1 0.0095 8.3245 -1816.7
## - theater 1 0.0138 8.3287 -1816.4
## - gaming 1 0.0220 8.3370 -1816.0
## - concerts 1 0.0263 8.3413 -1815.7
## - music 1 0.0312 8.3462 -1815.5
## <none> 8.3150 -1815.2
## - sinc_o 1 0.0406 8.3556 -1814.9
## - age 1 0.0438 8.3588 -1814.8
## - art 1 0.0511 8.3660 -1814.3
## + intel_o 1 0.0037 8.3113 -1813.4
## + museums 1 0.0032 8.3118 -1813.4
## + movies 1 0.0026 8.3124 -1813.3

```

```

## + amb_o      1      0.0021  8.3129 -1813.3
## + clubbing   1      0.0013  8.3136 -1813.3
## + hiking     1      0.0007  8.3142 -1813.2
## + sports     1      0.0003  8.3146 -1813.2
## + imprace    1      0.0003  8.3147 -1813.2
## + tv         1      0.0002  8.3148 -1813.2
## + dining     1      0.0001  8.3149 -1813.2
## - yoga       1      0.0804  8.3954 -1812.7
## - gender     1      0.0821  8.3970 -1812.6
## + field_cd   16      0.4883  7.8266 -1811.5
## - tvsports   1      0.1453  8.4603 -1809.1
## + go_out     6      0.1003  8.2146 -1808.9
## + date       6      0.0577  8.2572 -1806.5
## + goal       5      0.0099  8.3050 -1805.8
## - career_c   16      0.9257  9.2407 -1797.9
## - fun_o      1      0.9327  9.2477 -1767.5
## - attr_o     1      4.2349 12.5499 -1625.0
##
## Step:  AIC=-1816.99
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      tvsports + exercise + art + gaming + theater + concerts +
##      music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC
## - exercise  1      0.0071  8.3257 -1818.6
## - match_es  1      0.0087  8.3274 -1818.5
## - shopping  1      0.0090  8.3276 -1818.5
## - theater   1      0.0164  8.3351 -1818.1
## - gaming    1      0.0215  8.3402 -1817.8
## - concerts  1      0.0269  8.3456 -1817.5
## - music     1      0.0315  8.3501 -1817.2
## <none>              8.3187 -1817.0
## - sinc_o    1      0.0391  8.3578 -1816.8
## - age       1      0.0457  8.3644 -1816.4
## - art       1      0.0484  8.3671 -1816.3
## + reading   1      0.0037  8.3150 -1815.2
## + intel_o   1      0.0034  8.3152 -1815.2
## + movies    1      0.0022  8.3164 -1815.1
## + amb_o     1      0.0020  8.3167 -1815.1
## + clubbing  1      0.0018  8.3169 -1815.1
## + museums   1      0.0017  8.3170 -1815.1
## + hiking    1      0.0008  8.3178 -1815.0
## + imprace   1      0.0002  8.3184 -1815.0
## + sports    1      0.0002  8.3185 -1815.0
## + tv        1      0.0002  8.3185 -1815.0
## + dining    1      0.0001  8.3186 -1815.0
## - yoga      1      0.0803  8.3989 -1814.5
## - gender    1      0.0804  8.3991 -1814.5
## + field_cd  16      0.4822  7.8365 -1812.9
## - tvsports   1      0.1421  8.4608 -1811.1
## + go_out    6      0.1023  8.2163 -1810.8
## + date      6      0.0568  8.2618 -1808.2
## + goal      5      0.0099  8.3088 -1807.5
## - career_c  16      0.9223  9.2410 -1799.9

```



```

## - fun_o      1    0.9332  9.2519 -1769.3
## - attr_o     1    4.2318 12.5505 -1626.9
##
## Step: AIC=-1818.6
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      tvsports + art + gaming + theater + concerts + music + shopping +
##      yoga
##
##           Df Sum of Sq    RSS    AIC
## - match_es  1    0.0086  8.3343 -1820.1
## - shopping  1    0.0086  8.3344 -1820.1
## - theater   1    0.0155  8.3413 -1819.7
## - gaming    1    0.0219  8.3476 -1819.4
## - concerts  1    0.0237  8.3494 -1819.3
## - music     1    0.0312  8.3569 -1818.8
## <none>              8.3257 -1818.6
## - sinc_o    1    0.0400  8.3658 -1818.4
## - age       1    0.0463  8.3720 -1818.0
## - art       1    0.0483  8.3741 -1817.9
## + exercise  1    0.0071  8.3187 -1817.0
## + reading   1    0.0033  8.3225 -1816.8
## + intel_o   1    0.0032  8.3226 -1816.8
## + movies    1    0.0026  8.3231 -1816.8
## + amb_o     1    0.0024  8.3234 -1816.7
## + museums   1    0.0017  8.3241 -1816.7
## + hiking    1    0.0015  8.3242 -1816.7
## + clubbing  1    0.0015  8.3243 -1816.7
## + sports    1    0.0003  8.3255 -1816.6
## + dining    1    0.0002  8.3256 -1816.6
## + tv        1    0.0002  8.3256 -1816.6
## + imprace   1    0.0001  8.3257 -1816.6
## - yoga      1    0.0752  8.4009 -1816.4
## - gender    1    0.0798  8.4056 -1816.1
## + field_cd 16    0.4827  7.8430 -1814.5
## - tvsports  1    0.1352  8.4610 -1813.1
## + go_out    6    0.0984  8.2273 -1812.2
## + date      6    0.0574  8.2684 -1809.8
## + goal      5    0.0082  8.3175 -1809.1
## - career_c 16    0.9158  9.2416 -1801.9
## - fun_o     1    0.9289  9.2546 -1771.2
## - attr_o    1    4.5005 12.8263 -1618.8
##
## Step: AIC=-1820.12
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      art + gaming + theater + concerts + music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC
## - shopping  1    0.0099  8.3442 -1821.6
## - theater   1    0.0150  8.3493 -1821.3
## - gaming    1    0.0221  8.3564 -1820.9
## - concerts  1    0.0228  8.3572 -1820.8
## - music     1    0.0313  8.3657 -1820.4
## <none>              8.3343 -1820.1
## - sinc_o    1    0.0376  8.3719 -1820.0

```

```

## - art      1    0.0449  8.3792 -1819.6
## - age      1    0.0456  8.3799 -1819.6
## + match_es 1    0.0086  8.3257 -1818.6
## + exercise 1    0.0069  8.3274 -1818.5
## + reading  1    0.0035  8.3308 -1818.3
## + intel_o  1    0.0027  8.3316 -1818.3
## + museums  1    0.0023  8.3320 -1818.2
## + movies   1    0.0023  8.3320 -1818.2
## + amb_o    1    0.0023  8.3320 -1818.2
## + hiking   1    0.0012  8.3331 -1818.2
## + clubbing 1    0.0008  8.3336 -1818.2
## + tv       1    0.0005  8.3338 -1818.2
## + dining   1    0.0001  8.3342 -1818.1
## + sports   1    0.0000  8.3343 -1818.1
## + imprace  1    0.0000  8.3343 -1818.1
## - yoga     1    0.0782  8.4125 -1817.8
## - gender   1    0.0864  8.4207 -1817.3
## + field_cd 16    0.4881  7.8462 -1816.3
## - tvsports 1    0.1420  8.4763 -1814.2
## + go_out   6    0.0976  8.2367 -1813.6
## + date     6    0.0552  8.2791 -1811.2
## + goal     5    0.0085  8.3258 -1810.6
## - career_c 16    0.9239  9.2582 -1803.0
## - fun_o    1    0.9278  9.2622 -1772.8
## - attr_o   1    4.4927 12.8270 -1620.8
##
## Step:  AIC=-1821.57
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      art + gaming + theater + concerts + music + yoga
##
##           Df Sum of Sq    RSS    AIC
## - gaming   1    0.0191  8.3632 -1822.5
## - theater   1    0.0194  8.3636 -1822.5
## - concerts  1    0.0228  8.3670 -1822.3
## - music     1    0.0349  8.3791 -1821.6
## - sinc_o    1    0.0350  8.3792 -1821.6
## <none>                      8.3442 -1821.6
## - age      1    0.0397  8.3839 -1821.3
## - art      1    0.0427  8.3869 -1821.2
## + shopping 1    0.0099  8.3343 -1820.1
## + match_es 1    0.0098  8.3344 -1820.1
## + exercise 1    0.0066  8.3376 -1819.9
## + intel_o  1    0.0039  8.3403 -1819.8
## + movies   1    0.0032  8.3410 -1819.7
## + hiking   1    0.0030  8.3412 -1819.7
## + reading  1    0.0029  8.3413 -1819.7
## + amb_o    1    0.0020  8.3421 -1819.7
## + museums  1    0.0017  8.3425 -1819.7
## + dining   1    0.0006  8.3435 -1819.6
## + clubbing 1    0.0004  8.3438 -1819.6
## + tv       1    0.0002  8.3440 -1819.6
## + imprace  1    0.0002  8.3440 -1819.6
## + sports   1    0.0001  8.3440 -1819.6
## - gender   1    0.0775  8.4217 -1819.2

```

```

## - yoga      1    0.0857  8.4299 -1818.8
## + field_cd 16    0.4954  7.8487 -1818.2
## - tvsports  1    0.1432  8.4873 -1815.6
## + go_out    6    0.0994  8.2448 -1815.2
## + date      6    0.0494  8.2947 -1812.3
## + goal      5    0.0069  8.3373 -1812.0
## - career_c 16    0.9271  9.2712 -1804.4
## - fun_o     1    0.9236  9.2678 -1774.5
## - attr_o    1    4.4933 12.8375 -1622.4
##
## Step: AIC=-1822.5
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##       art + theater + concerts + music + yoga
##
##           Df Sum of Sq    RSS    AIC
## - theater  1    0.0194  8.3826 -1823.4
## - concerts 1    0.0272  8.3904 -1823.0
## - sinc_o   1    0.0351  8.3983 -1822.5
## - music    1    0.0359  8.3991 -1822.5
## <none>                                8.3632 -1822.5
## - art      1    0.0425  8.4057 -1822.1
## - age      1    0.0502  8.4135 -1821.7
## + gaming   1    0.0191  8.3442 -1821.6
## + match_es 1    0.0097  8.3535 -1821.0
## + exercise 1    0.0070  8.3563 -1820.9
## + shopping 1    0.0069  8.3564 -1820.9
## + hiking   1    0.0068  8.3564 -1820.9
## - gender   1    0.0652  8.4284 -1820.9
## + intel_o  1    0.0044  8.3588 -1820.8
## + movies   1    0.0031  8.3602 -1820.7
## + reading  1    0.0026  8.3606 -1820.7
## + clubbing 1    0.0017  8.3615 -1820.6
## + amb_o    1    0.0016  8.3616 -1820.6
## + museums  1    0.0012  8.3620 -1820.6
## + sports   1    0.0003  8.3630 -1820.5
## + dining   1    0.0001  8.3631 -1820.5
## + tv       1    0.0001  8.3632 -1820.5
## + imprace  1    0.0001  8.3632 -1820.5
## - yoga     1    0.0802  8.4434 -1820.0
## + field_cd 16    0.5000  7.8633 -1819.3
## - tvsports 1    0.1285  8.4917 -1817.4
## + go_out    6    0.0999  8.2633 -1816.1
## + date      6    0.0478  8.3155 -1813.2
## + goal      5    0.0065  8.3567 -1812.9
## - career_c 16    0.9135  9.2767 -1806.1
## - fun_o     1    0.9244  9.2877 -1775.5
## - attr_o    1    4.4769 12.8402 -1624.3
##
## Step: AIC=-1823.42
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##       art + concerts + music + yoga
##
##           Df Sum of Sq    RSS    AIC
## - concerts 1    0.0187  8.4013 -1824.4

```

```

## - art      1      0.0272  8.4098 -1823.9
## - sinc_o   1      0.0314  8.4140 -1823.7
## - music    1      0.0359  8.4185 -1823.4
## <none>                                8.3826 -1823.4
## - age      1      0.0501  8.4327 -1822.6
## + theater  1      0.0194  8.3632 -1822.5
## + gaming   1      0.0190  8.3636 -1822.5
## - gender   1      0.0532  8.4358 -1822.5
## + movies   1      0.0118  8.3708 -1822.1
## + shopping 1      0.0106  8.3720 -1822.0
## + match_es 1      0.0094  8.3732 -1821.9
## + hiking   1      0.0079  8.3747 -1821.9
## + exercise 1      0.0059  8.3767 -1821.8
## + intel_o  1      0.0053  8.3773 -1821.7
## + reading  1      0.0050  8.3776 -1821.7
## + amb_o    1      0.0011  8.3815 -1821.5
## + sports   1      0.0011  8.3816 -1821.5
## + clubbing 1      0.0010  8.3816 -1821.5
## + dining   1      0.0008  8.3818 -1821.5
## + imprace  1      0.0006  8.3821 -1821.5
## + tv        1      0.0004  8.3822 -1821.4
## + museums  1      0.0000  8.3826 -1821.4
## + field_cd 16     0.5153  7.8673 -1821.0
## - yoga      1      0.0863  8.4689 -1820.6
## - tvsports  1      0.1234  8.5060 -1818.6
## + go_out    6      0.1031  8.2795 -1817.2
## + date      6      0.0482  8.3345 -1814.1
## + goal      5      0.0052  8.3774 -1813.7
## - career_c 16     0.9151  9.2977 -1807.0
## - fun_o     1      0.9129  9.2955 -1777.2
## - attr_o    1      4.5910 12.9736 -1621.5
##
## Step:  AIC=-1824.38
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      art + music + yoga
##
##           Df Sum of Sq    RSS    AIC
## - music      1      0.0183  8.4197 -1825.4
## - sinc_o     1      0.0325  8.4338 -1824.6
## <none>                                8.4013 -1824.4
## - art        1      0.0428  8.4441 -1824.0
## + gaming     1      0.0227  8.3787 -1823.6
## - age        1      0.0529  8.4542 -1823.5
## + concerts   1      0.0187  8.3826 -1823.4
## - gender     1      0.0560  8.4573 -1823.3
## + theater    1      0.0109  8.3904 -1823.0
## + hiking     1      0.0104  8.3910 -1823.0
## + shopping   1      0.0093  8.3920 -1822.9
## + match_es   1      0.0086  8.3927 -1822.9
## + movies     1      0.0069  8.3945 -1822.8
## + intel_o    1      0.0053  8.3961 -1822.7
## + reading    1      0.0050  8.3963 -1822.7
## + exercise   1      0.0034  8.3979 -1822.6
## + field_cd   16     0.5250  7.8764 -1822.5

```

```

## + clubbing 1 0.0015 8.3999 -1822.5
## + sports 1 0.0011 8.4003 -1822.4
## + imprace 1 0.0010 8.4004 -1822.4
## + amb_o 1 0.0008 8.4006 -1822.4
## + tv 1 0.0006 8.4007 -1822.4
## + dining 1 0.0005 8.4009 -1822.4
## + museums 1 0.0001 8.4012 -1822.4
## - yoga 1 0.0770 8.4783 -1822.1
## - tvsports 1 0.1148 8.5161 -1820.0
## + go_out 6 0.0996 8.3017 -1818.0
## + date 6 0.0490 8.3523 -1815.1
## + goal 5 0.0043 8.3970 -1814.6
## - career_c 16 0.8996 9.3009 -1808.9
## - fun_o 1 0.8982 9.2996 -1778.9
## - attr_o 1 4.5831 12.9844 -1623.1
##
## Step: AIC=-1825.36
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
## art + yoga
##
## Df Sum of Sq RSS AIC
## - art 1 0.0319 8.4516 -1825.6
## - sinc_o 1 0.0334 8.4530 -1825.5
## <none> 8.4197 -1825.4
## - age 1 0.0502 8.4699 -1824.6
## + gaming 1 0.0207 8.3989 -1824.5
## + music 1 0.0183 8.4013 -1824.4
## + theater 1 0.0156 8.4041 -1824.2
## - gender 1 0.0568 8.4765 -1824.2
## + field_cd 16 0.5368 7.8829 -1824.1
## + movies 1 0.0135 8.4061 -1824.1
## + hiking 1 0.0132 8.4065 -1824.1
## + shopping 1 0.0131 8.4065 -1824.1
## + match_es 1 0.0095 8.4102 -1823.9
## + intel_o 1 0.0072 8.4125 -1823.8
## + reading 1 0.0052 8.4144 -1823.7
## + exercise 1 0.0047 8.4149 -1823.6
## + sports 1 0.0019 8.4178 -1823.5
## + amb_o 1 0.0015 8.4182 -1823.4
## + dining 1 0.0013 8.4183 -1823.4
## + tv 1 0.0012 8.4185 -1823.4
## + concerts 1 0.0011 8.4185 -1823.4
## + clubbing 1 0.0009 8.4188 -1823.4
## + imprace 1 0.0008 8.4189 -1823.4
## + museums 1 0.0000 8.4196 -1823.4
## - yoga 1 0.0841 8.5038 -1822.7
## - tvsports 1 0.1261 8.5457 -1820.4
## + go_out 6 0.0943 8.3254 -1818.6
## + date 6 0.0525 8.3672 -1816.3
## + goal 5 0.0036 8.4161 -1815.6
## - career_c 16 0.9235 9.3432 -1808.8
## - fun_o 1 0.9001 9.3198 -1779.9
## - attr_o 1 4.5672 12.9869 -1625.0
##

```

```

## Step: AIC=-1825.59
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##     yoga
##
##           Df Sum of Sq    RSS    AIC
## - sinc_o    1    0.0332  8.4847 -1825.8
## <none>                        8.4516 -1825.6
## + art        1    0.0319  8.4197 -1825.4
## - age        1    0.0433  8.4948 -1825.2
## + museums    1    0.0234  8.4282 -1824.9
## + gaming     1    0.0226  8.4290 -1824.8
## + hiking     1    0.0178  8.4337 -1824.6
## + field_cd   16    0.5379  7.9136 -1824.3
## - yoga       1    0.0638  8.5153 -1824.1
## + concerts   1    0.0088  8.4428 -1824.1
## + music      1    0.0074  8.4441 -1824.0
## + intel_o    1    0.0070  8.4446 -1824.0
## + shopping   1    0.0065  8.4450 -1824.0
## + match_es   1    0.0055  8.4461 -1823.9
## + exercise   1    0.0040  8.4476 -1823.8
## + movies     1    0.0035  8.4480 -1823.8
## + clubbing   1    0.0032  8.4484 -1823.8
## + amb_o      1    0.0012  8.4503 -1823.7
## + sports     1    0.0011  8.4505 -1823.7
## + reading    1    0.0010  8.4505 -1823.7
## + tv         1    0.0008  8.4507 -1823.6
## + imprace    1    0.0006  8.4509 -1823.6
## + theater    1    0.0006  8.4509 -1823.6
## + dining     1    0.0003  8.4512 -1823.6
## - gender     1    0.0762  8.5277 -1823.4
## - tvsports   1    0.1272  8.5787 -1820.6
## + go_out     6    0.0999  8.3517 -1819.1
## + date       6    0.0530  8.3986 -1816.5
## + goal       5    0.0045  8.4470 -1815.8
## - career_c   16    0.9405  9.3921 -1808.3
## - fun_o      1    0.9174  9.3689 -1779.5
## - attr_o     1    4.5502 13.0017 -1626.4
##
## Step: AIC=-1825.76
## dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##     yoga
##
##           Df Sum of Sq    RSS    AIC
## <none>                        8.4847 -1825.8
## + sinc_o    1    0.0332  8.4516 -1825.6
## + art        1    0.0317  8.4530 -1825.5
## + field_cd   16    0.5512  7.9335 -1825.1
## - age        1    0.0483  8.5331 -1825.1
## + museums    1    0.0240  8.4607 -1825.1
## + gaming     1    0.0227  8.4620 -1825.0
## + hiking     1    0.0127  8.4721 -1824.5
## + concerts   1    0.0091  8.4757 -1824.3
## + music      1    0.0080  8.4767 -1824.2
## - yoga       1    0.0659  8.5506 -1824.2

```

```
## + exercise 1 0.0047 8.4800 -1824.0
## + shopping 1 0.0044 8.4804 -1824.0
## + clubbing 1 0.0039 8.4809 -1824.0
## + match_es 1 0.0036 8.4811 -1824.0
## + movies 1 0.0030 8.4817 -1823.9
## - gender 1 0.0710 8.5557 -1823.9
## + intel_o 1 0.0020 8.4827 -1823.9
## + dining 1 0.0010 8.4838 -1823.8
## + tv 1 0.0008 8.4839 -1823.8
## + sports 1 0.0007 8.4840 -1823.8
## + imprace 1 0.0005 8.4842 -1823.8
## + reading 1 0.0003 8.4844 -1823.8
## + amb_o 1 0.0002 8.4846 -1823.8
## + theater 1 0.0002 8.4846 -1823.8
## - tvsports 1 0.1338 8.6185 -1820.5
## + go_out 6 0.1016 8.3832 -1819.4
## + date 6 0.0527 8.4320 -1816.7
## + goal 5 0.0059 8.4788 -1816.1
## - career_c 16 0.9151 9.3999 -1809.9
## - fun_o 1 0.8873 9.3720 -1781.3
## - attr_o 1 4.5172 13.0019 -1628.4
```

```
step$anova # display results
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## dec_o ~ (iid + attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##   age + field_cd + imprace + goal + date + go_out + career_c +
##   match_es + sports + tvsports + exercise + dining + museums +
##   art + hiking + gaming + clubbing + reading + tv + theater +
##   movies + concerts + music + shopping + yoga) - iid
##
## Final Model:
## dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##   yoga
##
##
```

	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
## 1				391	7.567203	-1773.210
## 2	- goal	5	1.539654e-02	396	7.582600	-1782.261
## 3	- date	6	7.484334e-02	402	7.657443	-1789.674
## 4	- go_out	6	1.260687e-01	408	7.783512	-1794.048
## 5	- field_cd	16	5.162036e-01	424	8.299715	-1796.060
## 6	- amb_o	1	6.261291e-06	425	8.299722	-1798.060
## 7	- dining	1	5.409194e-05	426	8.299776	-1800.057
## 8	- imprace	1	5.688824e-04	427	8.300345	-1802.025
## 9	- sports	1	9.320928e-04	428	8.301277	-1803.973
## 10	- hiking	1	8.642941e-04	429	8.302141	-1805.924
## 11	- tv	1	1.656778e-03	430	8.303798	-1807.831
## 12	- clubbing	1	1.512689e-03	431	8.305311	-1809.746
## 13	- museums	1	2.893222e-03	432	8.308204	-1811.583
## 14	- movies	1	3.070824e-03	433	8.311275	-1813.411
## 15	- intel_o	1	3.687421e-03	434	8.314962	-1815.203

```
## 16 - reading 1 3.712978e-03 435 8.318675 -1816.995
## 17 - exercise 1 7.063253e-03 436 8.325738 -1818.599
## 18 - match_es 1 8.569040e-03 437 8.334307 -1820.118
## 19 - shopping 1 9.869137e-03 438 8.344176 -1821.565
## 20 - gaming 1 1.906035e-02 439 8.363237 -1822.500
## 21 - theater 1 1.937799e-02 440 8.382615 -1823.419
## 22 - concerts 1 1.872411e-02 441 8.401339 -1824.377
## 23 - music 1 1.833370e-02 442 8.419673 -1825.359
## 24 - art 1 3.188745e-02 443 8.451560 -1825.594
## 25 - sinc_o 1 3.317171e-02 444 8.484732 -1825.764
```

Reducing our desirableness model based on our stepAIC:

```
# the stepAIC returned this model as these as our optimal predictors
# dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports + yoga
# dec_o, attr_o, fun_o here is the mean of dec_o, attr_o, fun_o reported by partners
reduced_augmented_fit <- lm(dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports + yoga, data = )
summary(reduced_augmented_fit)
```

```
##
## Call:
## lm(formula = dec_o ~ attr_o + fun_o + gender + age + career_c +
##      tvsports + yoga, data = na.omit(augmented_attribs))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41470 -0.08912 -0.00218  0.08472  0.43041
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.577780    0.071593  -8.070 6.62e-15 ***
## attr_o       0.120673    0.007849  15.375 < 2e-16 ***
## fun_o        0.061483    0.009023   6.814 3.11e-11 ***
## gender1     -0.028455    0.014767  -1.927  0.05464 .
## age         -0.002806    0.001764  -1.591  0.11243
## career_c2    0.018852    0.026269   0.718  0.47336
## career_c3   -0.108143    0.041151  -2.628  0.00889 **
## career_c4    0.001997    0.036428   0.055  0.95630
## career_c5   -0.009262    0.044519  -0.208  0.83528
## career_c6   -0.034336    0.031669  -1.084  0.27886
## career_c7    0.004894    0.026364   0.186  0.85283
## career_c8   -0.462684    0.101430  -4.562 6.58e-06 ***
## career_c9   -0.018413    0.034970  -0.527  0.59878
## career_c10  -0.002457    0.037142  -0.066  0.94728
## career_c11  -0.039991    0.045415  -0.881  0.37903
## career_c12  -0.248414    0.101162  -2.456  0.01445 *
## career_c13   0.056597    0.066072   0.857  0.39213
## career_c14   0.208049    0.140941   1.476  0.14062
## career_c15   0.096691    0.074230   1.303  0.19339
## career_c16   0.040248    0.101192   0.398  0.69101
## career_c17  -0.071265    0.140462  -0.507  0.61215
## tvsports    -0.006275    0.002372  -2.646  0.00844 **
## yoga        -0.004584    0.002469  -1.857  0.06398 .
## ---
```



```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1382 on 444 degrees of freedom
## Multiple R-squared:  0.6732, Adjusted R-squared:  0.6571
## F-statistic: 41.58 on 22 and 444 DF,  p-value: < 2.2e-16
```

Our model stayed with approximately the same adjusted R^2 (0.657), but now we have only 24 predictors. We noticed that `career_c` did not improve our R^2 by much, so we ran an F-test

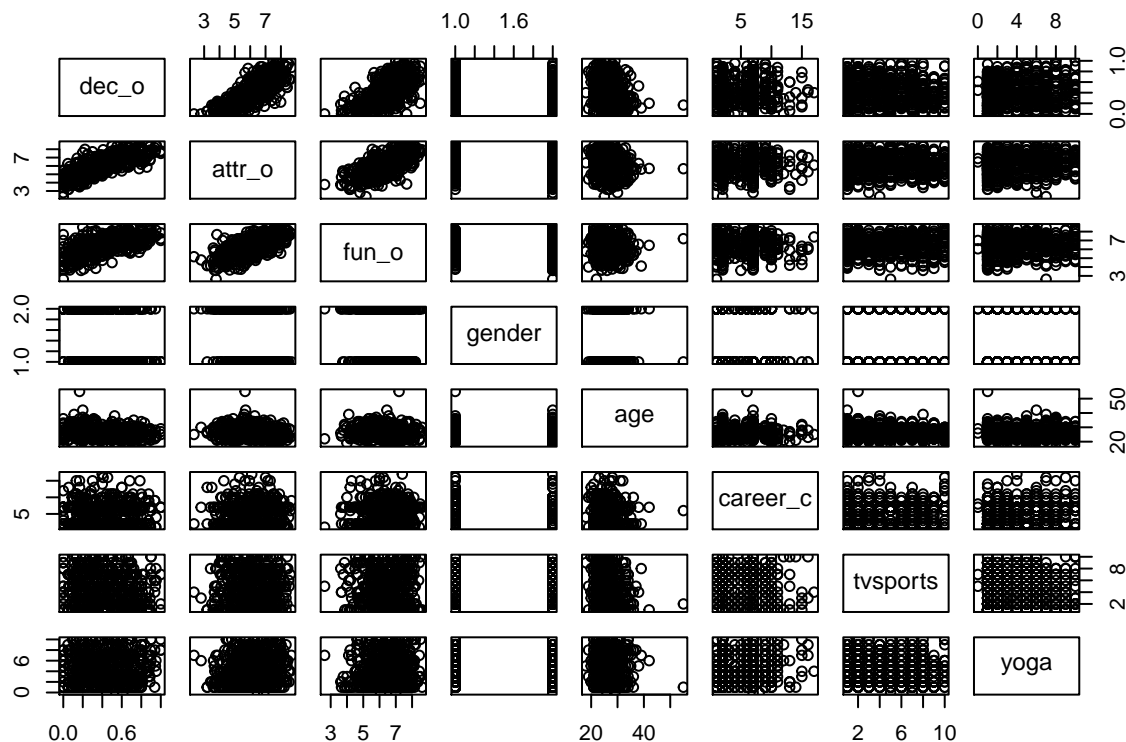
```
anova(reduced_augmented_fit, lm(dec_o ~ attr_o + fun_o + gender + age + tvsports + yoga, data = na.omit
```

```
## Analysis of Variance Table
##
## Model 1: dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##      yoga
## Model 2: dec_o ~ attr_o + fun_o + gender + age + tvsports + yoga
##   Res.Df    RSS Df Sum of Sq   F    Pr(>F)
## 1      444 8.4847
## 2      460 9.3999 -16  -0.91513 2.993 9.112e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value of the F test indicates the model with career significantly reduces the sum of squares error, so we will keep them it in model.

Checking for associations between variables

```
pairs(augmented_attribs[,c("dec_o", "attr_o", "fun_o", "gender", "age", "career_c", "tvsports", "yoga")])
```



We saw some multicollinearity with `attr_o` and `fun_o` so we decided to include an interaction term in our

model and compared it inside ANOVA.

```
anova(reduced_augmented_fit, lm(dec_o ~ attr_o + fun_o + career_c + gender + age + tvsports + yoga + attr_o * fun_o))

## Analysis of Variance Table
##
## Model 1: dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##      yoga
## Model 2: dec_o ~ attr_o + fun_o + career_c + gender + age + tvsports +
##      yoga + attr_o * fun_o
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      444 8.4847
## 2      443 8.2983   1    0.18648 9.955 0.001713 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# it appears that fun_o and attractiveness_o have a linear relationship with our response AND (maybe) e
```

Here, we find that the model with the interaction term reduces the sums of squares better than the reduced model without it.

In the following chunk, we rerun the Step AIC and it added the concerts and music variable. We then did another ANOVA model comparison and found that adding those variables reduced the sums of squares better than without it. We saw from the summary that age was not significant, so we did another ANOVA comparison without age, and found that this new reduced model reduced the sums of squares better than our full model.

```
reduced_augmented_fit<-lm(dec_o ~ attr_o + fun_o + career_c + gender + age + tvsports + yoga + attr_o*fun_o)

summary(reduced_augmented_fit)

##
## Call:
## lm(formula = dec_o ~ attr_o + fun_o + career_c + gender + age +
##      tvsports + yoga + attr_o * fun_o, data = na.omit(augmented_attribs))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.43265 -0.08633 -0.00321  0.08104  0.44211
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.010399   0.193292  -0.054   0.95712
## attr_o        0.019333   0.033046   0.585   0.55882
## fun_o       -0.030650   0.030537  -1.004   0.31606
## career_c2     0.025153   0.026085   0.964   0.33544
## career_c3    -0.104333   0.040760  -2.560   0.01081 *
## career_c4     0.008133   0.036118   0.225   0.82195
## career_c5    -0.008849   0.044077  -0.201   0.84098
## career_c6    -0.024113   0.031522  -0.765   0.44470
## career_c7     0.008893   0.026133   0.340   0.73379
## career_c8    -0.492488   0.100866  -4.883 1.46e-06 ***
## career_c9    -0.012222   0.034678  -0.352   0.72468
## career_c10    0.003635   0.036824   0.099   0.92142
## career_c11   -0.026998   0.045152  -0.598   0.55019
## career_c12   -0.235326   0.100243  -2.348   0.01934 *
## career_c13    0.039950   0.065628   0.609   0.54301
```

```
## career_c14    0.211939    0.139547    1.519    0.12953
## career_c15    0.106547    0.073559    1.448    0.14820
## career_c16    0.057115    0.100329    0.569    0.56946
## career_c17   -0.055345    0.139158   -0.398    0.69103
## gender1      -0.031970    0.014663   -2.180    0.02976 *
## age          -0.002569    0.001748   -1.470    0.14232
## tvsports     -0.006058    0.002349   -2.579    0.01024 *
## yoga         -0.004803    0.002445   -1.964    0.05012 .
## attr_o:fun_o  0.015823    0.005015    3.155    0.00171 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1369 on 443 degrees of freedom
## Multiple R-squared:  0.6804, Adjusted R-squared:  0.6638
## F-statistic: 41.01 on 23 and 443 DF,  p-value: < 2.2e-16

summary(lm(dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
  concerts + music + yoga + attr_o:fun_o, data = na.omit(augmented_attris)))

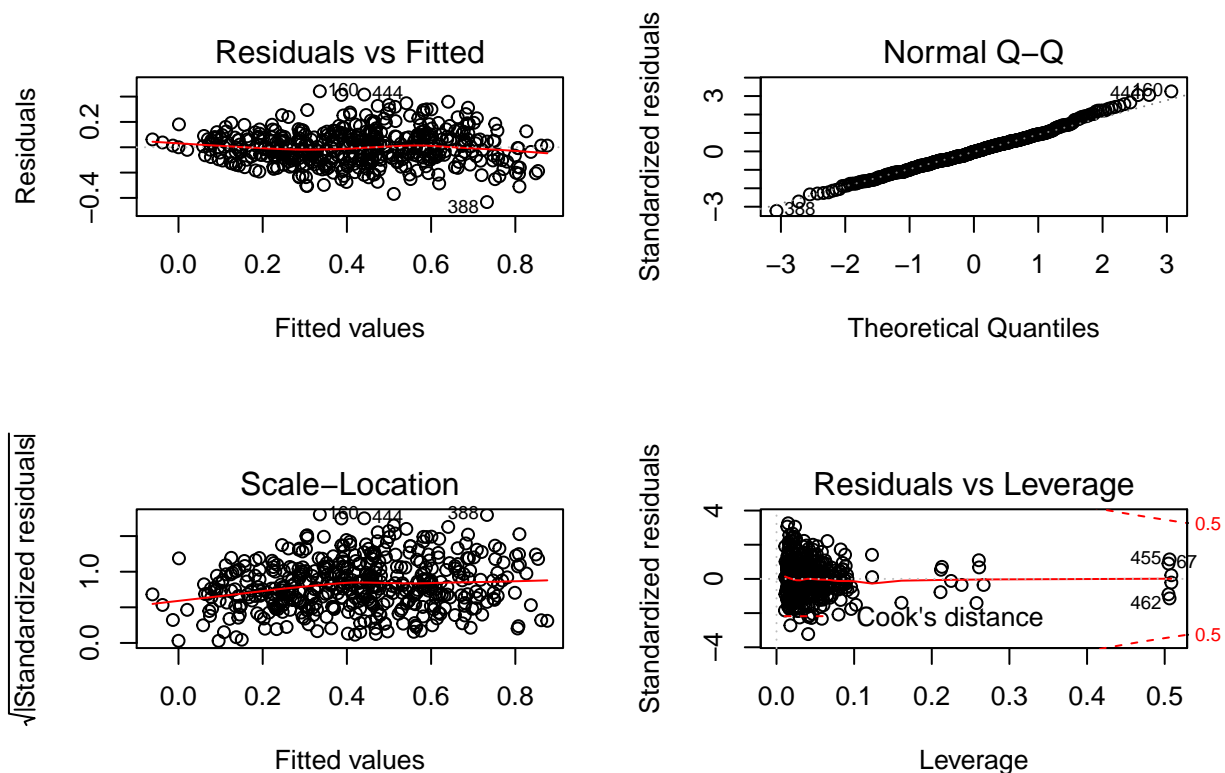
##
## Call:
## lm(formula = dec_o ~ attr_o + fun_o + gender + age + career_c +
##     tvsports + concerts + music + yoga + attr_o:fun_o, data = na.omit(augmented_attris))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41960 -0.08858 -0.00191  0.08020  0.43427
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.004220   0.194249   0.022  0.98268
## attr_o         0.017285   0.033075   0.523  0.60151
## fun_o        -0.031902   0.030538  -1.045  0.29675
## gender1       -0.030351   0.014709  -2.063  0.03965 *
## age          -0.002535   0.001747  -1.451  0.14750
## career_c2      0.025690   0.026073   0.985  0.32502
## career_c3     -0.104504   0.040767  -2.563  0.01069 *
## career_c4      0.008371   0.036086   0.232  0.81666
## career_c5     -0.012965   0.044136  -0.294  0.76908
## career_c6     -0.023914   0.031542  -0.758  0.44877
## career_c7      0.010462   0.026168   0.400  0.68950
## career_c8     -0.490645   0.101018  -4.857 1.66e-06 ***
## career_c9     -0.011240   0.034662  -0.324  0.74588
## career_c10     0.007169   0.036860   0.194  0.84588
## career_c11    -0.028874   0.045132  -0.640  0.52265
## career_c12    -0.238328   0.100169  -2.379  0.01777 *
## career_c13     0.043373   0.065639   0.661  0.50910
## career_c14     0.218969   0.139484   1.570  0.11717
## career_c15     0.111603   0.073647   1.515  0.13039
## career_c16     0.071038   0.100622   0.706  0.48056
## career_c17    -0.062551   0.139487  -0.448  0.65406
## tvsports      -0.006258   0.002369  -2.642  0.00854 **
## concerts       0.006323   0.004024   1.571  0.11683
## music         -0.006605   0.004656  -1.418  0.15676
## yoga          -0.005361   0.002520  -2.127  0.03394 *
```

```
## attr_o:fun_o 0.016216 0.005021 3.230 0.00133 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1367 on 441 degrees of freedom
## Multiple R-squared: 0.6824, Adjusted R-squared: 0.6644
## F-statistic: 37.91 on 25 and 441 DF, p-value: < 2.2e-16
```

```
par(mfrow = c(2,2))
plot(reduced_augmented_fit)
```

```
## Warning: not plotting observations with leverage one:
## 212, 444
```

```
## Warning: not plotting observations with leverage one:
## 212, 444
```



QQ plot indicates normally distributed errors. The scale-location plot show that there are possibly some outliers.

#next step is to maybe introduce interaction terms?

```
full_augmented_fit <- lm(dec_o ~ . + attr_o*fun_o -iid, data = na.omit(augmented_attribs))
step <- stepAIC(full_augmented_fit, direction="both")
```

```
## Start: AIC=-1781.15
## dec_o ~ iid + attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
## age + field_cd + imprace + goal + date + go_out + career_c +
## matches + sports + tvsports + exercise + dining + museums +
## art + hiking + gaming + clubbing + reading + tv + theater +
## movies + concerts + music + shopping + yoga + attr_o * fun_o -
## iid
```

```

##
##           Df Sum of Sq   RSS   AIC
## - goal      5    0.01619 7.4240 -1790.1
## - date      6    0.07633 7.4841 -1788.4
## - go_out    6    0.11283 7.5206 -1786.1
## - theater   1    0.00010 7.4079 -1783.2
## - sports    1    0.00024 7.4080 -1783.1
## - art       1    0.00043 7.4082 -1783.1
## - imprace   1    0.00058 7.4084 -1783.1
## - clubbing  1    0.00090 7.4087 -1783.1
## - dining    1    0.00125 7.4090 -1783.1
## - intel_o   1    0.00168 7.4095 -1783.0
## - match_es  1    0.00316 7.4109 -1783.0
## - amb_o     1    0.00470 7.4125 -1782.9
## - gaming    1    0.00475 7.4125 -1782.8
## - shopping  1    0.00553 7.4133 -1782.8
## - exercise  1    0.00615 7.4139 -1782.8
## - reading   1    0.00962 7.4174 -1782.5
## - music     1    0.01003 7.4178 -1782.5
## - hiking    1    0.01505 7.4228 -1782.2
## - movies    1    0.01616 7.4240 -1782.1
## - museums   1    0.02057 7.4284 -1781.9
## - concerts  1    0.02319 7.4310 -1781.7
## - sinc_o    1    0.02426 7.4320 -1781.6
## - tv        1    0.02561 7.4334 -1781.5
## - field_cd  16    0.52263 7.9304 -1781.3
## <none>              7.4078 -1781.2
## - age       1    0.03451 7.4423 -1781.0
## - gender    1    0.09930 7.5071 -1776.9
## - yoga      1    0.11002 7.5178 -1776.3
## - tvsports  1    0.13193 7.5397 -1774.9
## - attr_o:fun_o 1    0.15942 7.5672 -1773.2
## - career_c  15    0.97300 8.3808 -1753.5
##
## Step:  AIC=-1790.13
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##       age + field_cd + imprace + date + go_out + career_c + match_es +
##       sports + tvsports + exercise + dining + museums + art + hiking +
##       gaming + clubbing + reading + tv + theater + movies + concerts +
##       music + shopping + yoga + attr_o:fun_o
##
##           Df Sum of Sq   RSS   AIC
## - date      6    0.07828 7.5023 -1797.2
## - go_out    6    0.11957 7.5435 -1794.7
## - theater   1    0.00003 7.4240 -1792.1
## - art       1    0.00031 7.4243 -1792.1
## - sports    1    0.00035 7.4243 -1792.1
## - imprace   1    0.00036 7.4243 -1792.1
## - dining    1    0.00056 7.4245 -1792.1
## - clubbing  1    0.00122 7.4252 -1792.1
## - intel_o   1    0.00158 7.4256 -1792.0
## - shopping  1    0.00381 7.4278 -1791.9
## - match_es  1    0.00439 7.4284 -1791.9
## - gaming    1    0.00474 7.4287 -1791.8

```

```

## - exercise      1  0.00493 7.4289 -1791.8
## - amb_o         1  0.00546 7.4294 -1791.8
## - reading       1  0.00878 7.4328 -1791.6
## - music         1  0.00896 7.4329 -1791.6
## - movies        1  0.01461 7.4386 -1791.2
## - hiking        1  0.01464 7.4386 -1791.2
## - museums       1  0.01833 7.4423 -1791.0
## - concerts      1  0.02009 7.4441 -1790.9
## - sinc_o        1  0.02022 7.4442 -1790.9
## - tv            1  0.02203 7.4460 -1790.8
## - field_cd      16  0.51884 7.9428 -1790.6
## <none>          7.4240 -1790.1
## - age           1  0.03760 7.4616 -1789.8
## - gender        1  0.09073 7.5147 -1786.5
## - yoga          1  0.10570 7.5297 -1785.5
## - tvsports      1  0.13133 7.5553 -1784.0
## - attr_o:fun_o  1  0.15863 7.5826 -1782.3
## + goal          5  0.01619 7.4078 -1781.2
## - career_c      15  0.98795 8.4119 -1761.8
##
## Step: AIC=-1797.24
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + field_cd + imprace + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + hiking +
##      gaming + clubbing + reading + tv + theater + movies + concerts +
##      music + shopping + yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS      AIC
## - go_out      6  0.12393 7.6262 -1801.6
## - theater     1  0.00001 7.5023 -1799.2
## - art         1  0.00003 7.5023 -1799.2
## - dining      1  0.00006 7.5023 -1799.2
## - intel_o     1  0.00037 7.5026 -1799.2
## - imprace     1  0.00037 7.5026 -1799.2
## - clubbing    1  0.00140 7.5037 -1799.2
## - sports      1  0.00157 7.5038 -1799.1
## - shopping    1  0.00192 7.5042 -1799.1
## - amb_o       1  0.00305 7.5053 -1799.0
## - field_cd    16  0.50136 8.0036 -1799.0
## - match_es    1  0.00396 7.5062 -1799.0
## - exercise    1  0.00471 7.5070 -1798.9
## - gaming      1  0.00500 7.5073 -1798.9
## - reading     1  0.00731 7.5096 -1798.8
## - hiking      1  0.01204 7.5143 -1798.5
## - music       1  0.01303 7.5153 -1798.4
## - museums     1  0.01506 7.5173 -1798.3
## - sinc_o      1  0.01622 7.5185 -1798.2
## - movies      1  0.01694 7.5192 -1798.2
## - tv          1  0.02038 7.5226 -1798.0
## - concerts    1  0.02189 7.5241 -1797.9
## <none>        7.5023 -1797.2
## - age         1  0.03955 7.5418 -1796.8
## - gender      1  0.08599 7.5882 -1793.9
## - yoga        1  0.10152 7.6038 -1793.0

```

```

## - tvsports      1    0.14570 7.6480 -1790.2
## + date          6    0.07828 7.4240 -1790.1
## - attr_o:fun_o  1    0.15519 7.6574 -1789.7
## + goal          5    0.01814 7.4841 -1788.4
## - career_c     15    0.97443 8.4767 -1770.2
##
## Step:  AIC=-1801.58
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + field_cd + imprace + career_c + match_es + sports +
##      tvsports + exercise + dining + museums + art + hiking + gaming +
##      clubbing + reading + tv + theater + movies + concerts + music +
##      shopping + yoga + attr_o:fun_o
##
##              Df Sum of Sq   RSS    AIC
## - field_cd    16   0.48099 8.1072 -1805.0
## - intel_o      1   0.00000 7.6262 -1803.6
## - dining       1   0.00002 7.6262 -1803.6
## - art          1   0.00005 7.6262 -1803.6
## - theater      1   0.00035 7.6265 -1803.6
## - sports       1   0.00063 7.6268 -1803.5
## - imprace      1   0.00073 7.6269 -1803.5
## - amb_o        1   0.00190 7.6281 -1803.5
## - exercise     1   0.00257 7.6288 -1803.4
## - clubbing     1   0.00274 7.6289 -1803.4
## - match_es     1   0.00371 7.6299 -1803.4
## - shopping     1   0.00458 7.6308 -1803.3
## - gaming       1   0.00537 7.6315 -1803.3
## - music        1   0.01001 7.6362 -1803.0
## - reading      1   0.01059 7.6368 -1802.9
## - movies       1   0.01133 7.6375 -1802.9
## - hiking       1   0.01174 7.6379 -1802.9
## - sinc_o       1   0.01335 7.6395 -1802.8
## - concerts     1   0.01743 7.6436 -1802.5
## - museums      1   0.01784 7.6440 -1802.5
## - tv           1   0.02201 7.6482 -1802.2
## <none>                7.6262 -1801.6
## - age          1   0.03584 7.6620 -1801.4
## - gender       1   0.07707 7.7033 -1798.9
## - yoga         1   0.10147 7.7277 -1797.4
## + go_out       6   0.12393 7.5023 -1797.2
## - tvsports     1   0.13093 7.7571 -1795.6
## + date         6   0.08264 7.5435 -1794.7
## - attr_o:fun_o 1   0.15733 7.7835 -1794.0
## + goal         5   0.02244 7.6037 -1793.0
## - career_c     15   0.98930 8.6155 -1774.6
##
## Step:  AIC=-1805.02
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + imprace + career_c + match_es + sports + tvsports +
##      exercise + dining + museums + art + hiking + gaming + clubbing +
##      reading + tv + theater + movies + concerts + music + shopping +
##      yoga + attr_o:fun_o
##
##              Df Sum of Sq   RSS    AIC

```

```

## - imprace      1  0.00005 8.1072 -1807.0
## - dining       1  0.00083 8.1080 -1807.0
## - clubbing     1  0.00120 8.1084 -1807.0
## - sports       1  0.00134 8.1085 -1806.9
## - intel_o      1  0.00141 8.1086 -1806.9
## - amb_o        1  0.00222 8.1094 -1806.9
## - hiking       1  0.00233 8.1095 -1806.9
## - reading      1  0.00356 8.1107 -1806.8
## - art          1  0.00366 8.1108 -1806.8
## - exercise     1  0.00379 8.1110 -1806.8
## - movies       1  0.00544 8.1126 -1806.7
## - tv           1  0.00557 8.1127 -1806.7
## - theater      1  0.00645 8.1136 -1806.7
## - museums      1  0.00700 8.1142 -1806.6
## - shopping     1  0.00725 8.1144 -1806.6
## - match_es     1  0.01045 8.1176 -1806.4
## - gaming       1  0.01344 8.1206 -1806.2
## - music        1  0.02700 8.1342 -1805.5
## - sinc_o       1  0.02998 8.1372 -1805.3
## <none>         8.1072 -1805.0
## - age          1  0.03556 8.1427 -1805.0
## - concerts     1  0.03594 8.1431 -1805.0
## - gender       1  0.07585 8.1830 -1802.7
## - yoga         1  0.09034 8.1975 -1801.8
## + field_cd     16  0.48099 7.6262 -1801.6
## - tvsports     1  0.12954 8.2367 -1799.6
## + go_out       6  0.10356 8.0036 -1799.0
## + date         6  0.06665 8.0405 -1796.9
## - attr_o:fun_o 1  0.19254 8.2997 -1796.1
## + goal         5  0.01233 8.0948 -1795.7
## - career_c     16  0.88598 8.9932 -1788.6
##
## Step:  AIC=-1807.02
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##       age + career_c + match_es + sports + tvsports + exercise +
##       dining + museums + art + hiking + gaming + clubbing + reading +
##       tv + theater + movies + concerts + music + shopping + yoga +
##       attr_o:fun_o
##
##              Df Sum of Sq  RSS    AIC
## - dining      1  0.00082 8.1080 -1809.0
## - clubbing    1  0.00117 8.1084 -1809.0
## - sports      1  0.00132 8.1085 -1808.9
## - intel_o     1  0.00141 8.1086 -1808.9
## - amb_o       1  0.00222 8.1094 -1808.9
## - hiking      1  0.00236 8.1096 -1808.9
## - reading     1  0.00363 8.1108 -1808.8
## - art         1  0.00364 8.1109 -1808.8
## - exercise    1  0.00387 8.1111 -1808.8
## - movies      1  0.00545 8.1127 -1808.7
## - tv          1  0.00602 8.1132 -1808.7
## - theater     1  0.00642 8.1136 -1808.7
## - museums     1  0.00708 8.1143 -1808.6
## - shopping    1  0.00721 8.1144 -1808.6

```



```

## - match_es      1    0.01058 8.1178 -1808.4
## - gaming        1    0.01345 8.1207 -1808.2
## - music         1    0.02700 8.1342 -1807.5
## - sinc_o        1    0.03004 8.1373 -1807.3
## <none>           8.1072 -1807.0
## - concerts      1    0.03590 8.1431 -1807.0
## - age           1    0.03607 8.1433 -1807.0
## + imprace       1    0.00005 8.1072 -1805.0
## - gender        1    0.07677 8.1840 -1804.6
## - yoga          1    0.09166 8.1989 -1803.8
## + field_cd      16    0.48031 7.6269 -1803.5
## - tvsports      1    0.12959 8.2368 -1801.6
## + go_out        6    0.10334 8.0039 -1801.0
## + date          6    0.06669 8.0405 -1798.9
## - attr_o:fun_o  1    0.19306 8.3003 -1798.0
## + goal          5    0.01238 8.0948 -1797.7
## - career_c      16    0.88629 8.9935 -1790.6
##
## Step:  AIC=-1808.97
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + career_c + match_es + sports + tvsports + exercise +
##      museums + art + hiking + gaming + clubbing + reading + tv +
##      theater + movies + concerts + music + shopping + yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS      AIC
## - clubbing    1    0.00102 8.1091 -1810.9
## - intel_o     1    0.00133 8.1094 -1810.9
## - sports      1    0.00138 8.1094 -1810.9
## - hiking      1    0.00218 8.1102 -1810.8
## - amb_o       1    0.00235 8.1104 -1810.8
## - reading     1    0.00360 8.1116 -1810.8
## - exercise    1    0.00367 8.1117 -1810.8
## - art         1    0.00375 8.1118 -1810.8
## - movies      1    0.00567 8.1137 -1810.6
## - tv          1    0.00609 8.1141 -1810.6
## - museums     1    0.00650 8.1145 -1810.6
## - theater     1    0.00656 8.1146 -1810.6
## - shopping    1    0.00966 8.1177 -1810.4
## - match_es    1    0.01097 8.1190 -1810.3
## - gaming      1    0.01312 8.1212 -1810.2
## - music       1    0.02746 8.1355 -1809.4
## - sinc_o      1    0.02947 8.1375 -1809.3
## <none>         8.1080 -1809.0
## - concerts    1    0.03583 8.1439 -1808.9
## - age         1    0.03698 8.1450 -1808.8
## + dining      1    0.00082 8.1072 -1807.0
## + imprace     1    0.00004 8.1080 -1807.0
## - gender      1    0.07604 8.1841 -1806.6
## - yoga        1    0.09086 8.1989 -1805.8
## + field_cd    16    0.48111 7.6269 -1805.5
## - tvsports    1    0.12902 8.2371 -1803.6
## + go_out      6    0.10332 8.0047 -1803.0
## + date        6    0.06512 8.0429 -1800.7
## - attr_o:fun_o 1    0.19230 8.3003 -1800.0

```

```

## + goal          5    0.01183 8.0962 -1799.7
## - career_c      16    0.88559 8.9936 -1792.6
##
## Step: AIC=-1810.91
## dec_o ~ attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##      age + career_c + match_es + sports + tvsports + exercise +
##      museums + art + hiking + gaming + reading + tv + theater +
##      movies + concerts + music + shopping + yoga + attr_o:fun_o
##
##              Df Sum of Sq    RSS    AIC
## - intel_o      1    0.00122 8.1103 -1812.8
## - sports        1    0.00137 8.1104 -1812.8
## - hiking         1    0.00191 8.1110 -1812.8
## - amb_o          1    0.00227 8.1113 -1812.8
## - exercise       1    0.00358 8.1126 -1812.7
## - art            1    0.00395 8.1130 -1812.7
## - reading        1    0.00399 8.1131 -1812.7
## - movies         1    0.00549 8.1146 -1812.6
## - tv            1    0.00562 8.1147 -1812.6
## - theater        1    0.00626 8.1153 -1812.5
## - museums        1    0.00652 8.1156 -1812.5
## - shopping       1    0.00925 8.1183 -1812.4
## - match_es       1    0.01035 8.1194 -1812.3
## - gaming         1    0.01519 8.1243 -1812.0
## - music          1    0.02738 8.1364 -1811.3
## - sinc_o         1    0.02921 8.1383 -1811.2
## <none>              8.1091 -1810.9
## - concerts       1    0.03586 8.1449 -1810.8
## - age            1    0.03743 8.1465 -1810.8
## + clubbing       1    0.00102 8.1080 -1809.0
## + dining         1    0.00067 8.1084 -1809.0
## + imprace        1    0.00001 8.1090 -1808.9
## - gender         1    0.07882 8.1879 -1808.4
## - yoga           1    0.09096 8.2000 -1807.7
## + field_cd       16    0.47916 7.6299 -1807.4
## - tvsports       1    0.12825 8.2373 -1805.6
## + go_out         6    0.10412 8.0049 -1805.0
## + date           6    0.06518 8.0439 -1802.7
## - attr_o:fun_o   1    0.19334 8.3024 -1801.9
## + goal           5    0.01186 8.0972 -1801.6
## - career_c       16    0.89150 9.0006 -1794.2
##
## Step: AIC=-1812.84
## dec_o ~ attr_o + sinc_o + fun_o + amb_o + gender + age + career_c +
##      match_es + sports + tvsports + exercise + museums + art +
##      hiking + gaming + reading + tv + theater + movies + concerts +
##      music + shopping + yoga + attr_o:fun_o
##
##              Df Sum of Sq    RSS    AIC
## - amb_o          1    0.00112 8.1114 -1814.8
## - sports          1    0.00165 8.1119 -1814.8
## - hiking          1    0.00189 8.1122 -1814.7
## - exercise        1    0.00322 8.1135 -1814.7
## - reading          1    0.00379 8.1141 -1814.6

```

```

## - art          1  0.00409 8.1144 -1814.6
## - movies       1  0.00502 8.1153 -1814.5
## - tv           1  0.00532 8.1156 -1814.5
## - museums      1  0.00645 8.1167 -1814.5
## - theater      1  0.00678 8.1171 -1814.5
## - match_es     1  0.01010 8.1204 -1814.3
## - shopping     1  0.01023 8.1205 -1814.2
## - gaming       1  0.01607 8.1264 -1813.9
## - music        1  0.02809 8.1384 -1813.2
## - sinc_o       1  0.03468 8.1450 -1812.8
## <none>         8.1103 -1812.8
## - concerts     1  0.03608 8.1464 -1812.8
## - age          1  0.03661 8.1469 -1812.7
## + intel_o      1  0.00122 8.1091 -1810.9
## + clubbing     1  0.00091 8.1094 -1810.9
## + dining       1  0.00060 8.1097 -1810.9
## + imprace      1  0.00002 8.1103 -1810.8
## - gender       1  0.07852 8.1888 -1810.3
## - yoga         1  0.09104 8.2013 -1809.6
## + field_cd     16  0.48037 7.6299 -1809.4
## - tvsports     1  0.13096 8.2412 -1807.4
## + go_out       6  0.10181 8.0085 -1806.7
## + date         6  0.06445 8.0458 -1804.6
## - attr_o:fun_o 1  0.19477 8.3051 -1803.8
## + goal         5  0.01174 8.0985 -1803.5
## - career_c     16  0.93384 9.0441 -1794.0
##
## Step: AIC=-1814.78
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      sports + tvsports + exercise + museums + art + hiking + gaming +
##      reading + tv + theater + movies + concerts + music + shopping +
##      yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS    AIC
## - sports      1  0.00145 8.1128 -1816.7
## - hiking      1  0.00191 8.1133 -1816.7
## - exercise    1  0.00317 8.1146 -1816.6
## - reading     1  0.00382 8.1152 -1816.6
## - art         1  0.00428 8.1157 -1816.5
## - movies      1  0.00496 8.1164 -1816.5
## - tv          1  0.00507 8.1165 -1816.5
## - museums     1  0.00622 8.1176 -1816.4
## - theater     1  0.00698 8.1184 -1816.4
## - match_es    1  0.01014 8.1215 -1816.2
## - shopping    1  0.01026 8.1217 -1816.2
## - gaming      1  0.01649 8.1279 -1815.8
## - music       1  0.02767 8.1391 -1815.2
## <none>         8.1114 -1814.8
## - concerts    1  0.03653 8.1479 -1814.7
## - age         1  0.03658 8.1480 -1814.7
## - sinc_o      1  0.04096 8.1524 -1814.4
## + amb_o       1  0.00112 8.1103 -1812.8
## + clubbing    1  0.00092 8.1105 -1812.8
## + dining      1  0.00076 8.1106 -1812.8

```

```

## + intel_o      1  0.00006 8.1113 -1812.8
## + imprace      1  0.00001 8.1114 -1812.8
## - gender       1  0.08738 8.1988 -1811.8
## - yoga         1  0.09389 8.2053 -1811.4
## + field_cd     16  0.47862 7.6328 -1811.2
## - tvsports     1  0.12985 8.2412 -1809.4
## + go_out       6  0.10252 8.0089 -1808.7
## + date         6  0.06384 8.0476 -1806.5
## - attr_o:fun_o  1  0.19558 8.3070 -1805.7
## + goal         5  0.01188 8.0995 -1805.5
## - career_c     16  0.93435 9.0457 -1795.9
##
## Step: AIC=-1816.7
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      tvsports + exercise + museums + art + hiking + gaming + reading +
##      tv + theater + movies + concerts + music + shopping + yoga +
##      attr_o:fun_o
##
##           Df Sum of Sq  RSS    AIC
## - hiking      1  0.00267 8.1155 -1818.5
## - reading      1  0.00411 8.1170 -1818.5
## - art          1  0.00412 8.1170 -1818.5
## - tv          1  0.00467 8.1175 -1818.4
## - movies      1  0.00482 8.1177 -1818.4
## - exercise    1  0.00550 8.1184 -1818.4
## - museums     1  0.00643 8.1193 -1818.3
## - theater     1  0.00782 8.1207 -1818.2
## - match_es    1  0.00943 8.1223 -1818.2
## - shopping    1  0.01039 8.1232 -1818.1
## - gaming      1  0.01657 8.1294 -1817.7
## - music       1  0.02827 8.1411 -1817.1
## <none>                8.1128 -1816.7
## - age         1  0.03683 8.1497 -1816.6
## - concerts    1  0.03715 8.1500 -1816.6
## - sinc_o      1  0.04062 8.1535 -1816.4
## + sports      1  0.00145 8.1114 -1814.8
## + amb_o       1  0.00092 8.1119 -1814.8
## + clubbing    1  0.00090 8.1119 -1814.8
## + dining      1  0.00079 8.1121 -1814.7
## + intel_o     1  0.00015 8.1127 -1814.7
## + imprace     1  0.00000 8.1128 -1814.7
## - gender      1  0.08687 8.1997 -1813.7
## - yoga        1  0.09360 8.2064 -1813.3
## + field_cd    16  0.47954 7.6333 -1813.2
## + go_out      6  0.10088 8.0120 -1810.5
## - tvsports    1  0.14482 8.2577 -1810.4
## + date        6  0.06492 8.0479 -1808.5
## - attr_o:fun_o 1  0.19470 8.3075 -1807.6
## + goal        5  0.01204 8.1008 -1807.4
## - career_c    16  0.93477 9.0476 -1797.8
##
## Step: AIC=-1818.54
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      tvsports + exercise + museums + art + gaming + reading +

```

```

##      tv + theater + movies + concerts + music + shopping + yoga +
##      attr_o:fun_o
##
##      Df Sum of Sq   RSS   AIC
## - tv      1    0.00397 8.1195 -1820.3
## - art      1    0.00422 8.1197 -1820.3
## - reading  1    0.00440 8.1199 -1820.3
## - movies   1    0.00470 8.1202 -1820.3
## - exercise 1    0.00660 8.1221 -1820.2
## - museums  1    0.00717 8.1227 -1820.1
## - theater  1    0.00800 8.1235 -1820.1
## - match_es 1    0.00898 8.1245 -1820.0
## - shopping 1    0.01271 8.1282 -1819.8
## - gaming   1    0.02160 8.1371 -1819.3
## - music    1    0.03078 8.1463 -1818.8
## <none>           8.1155 -1818.5
## - age      1    0.03596 8.1515 -1818.5
## - sinc_o   1    0.03882 8.1543 -1818.3
## - concerts 1    0.03919 8.1547 -1818.3
## + hiking   1    0.00267 8.1128 -1816.7
## + sports   1    0.00221 8.1133 -1816.7
## + amb_o    1    0.00090 8.1146 -1816.6
## + dining   1    0.00063 8.1149 -1816.6
## + clubbing 1    0.00060 8.1149 -1816.6
## + intel_o  1    0.00017 8.1153 -1816.5
## + imprace  1    0.00002 8.1155 -1816.5
## - yoga     1    0.09097 8.2065 -1815.3
## - gender   1    0.09161 8.2071 -1815.3
## + field_cd 16    0.47080 7.6447 -1814.5
## + go_out   6    0.10027 8.0152 -1812.3
## - tvsports 1    0.14516 8.2607 -1812.3
## + date     6    0.06393 8.0516 -1810.2
## - attr_o:fun_o 1 0.19267 8.3082 -1809.6
## + goal     5    0.01259 8.1029 -1809.3
## - career_c 16    0.95244 9.0680 -1798.7
##
## Step: AIC=-1820.31
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      tvsports + exercise + museums + art + gaming + reading +
##      theater + movies + concerts + music + shopping + yoga + attr_o:fun_o
##
##      Df Sum of Sq   RSS   AIC
## - movies      1    0.00269 8.1222 -1822.2
## - reading      1    0.00405 8.1235 -1822.1
## - art          1    0.00423 8.1237 -1822.1
## - museums      1    0.00647 8.1260 -1821.9
## - exercise     1    0.00685 8.1263 -1821.9
## - theater      1    0.00765 8.1271 -1821.9
## - shopping     1    0.00934 8.1288 -1821.8
## - match_es     1    0.01018 8.1297 -1821.7
## - gaming       1    0.02446 8.1439 -1820.9
## - music        1    0.03060 8.1501 -1820.6
## <none>           8.1195 -1820.3
## - concerts     1    0.03664 8.1561 -1820.2

```

```

## - age          1    0.03720 8.1567 -1820.2
## - sinc_o       1    0.03790 8.1574 -1820.1
## + tv           1    0.00397 8.1155 -1818.5
## + hiking       1    0.00197 8.1175 -1818.4
## + sports       1    0.00163 8.1179 -1818.4
## + amb_o        1    0.00072 8.1188 -1818.3
## + dining       1    0.00072 8.1188 -1818.3
## + clubbing     1    0.00033 8.1192 -1818.3
## + imprace      1    0.00029 8.1192 -1818.3
## + intel_o      1    0.00013 8.1194 -1818.3
## - yoga         1    0.09025 8.2097 -1817.2
## - gender       1    0.09678 8.2163 -1816.8
## + field_cd     16    0.45896 7.6605 -1815.5
## - tvsports     1    0.14272 8.2622 -1814.2
## + go_out       6    0.10067 8.0188 -1814.1
## + date         6    0.06310 8.0564 -1812.0
## - attr_o:fun_o 1    0.18986 8.3093 -1811.5
## + goal         5    0.01152 8.1080 -1811.0
## - career_c     16    0.94979 9.0693 -1800.7
##
## Step: AIC=-1822.16
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
##      tvsports + exercise + museums + art + gaming + reading +
##      theater + concerts + music + shopping + yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS    AIC
## - reading      1    0.00373 8.1259 -1823.9
## - art           1    0.00405 8.1262 -1823.9
## - museums       1    0.00670 8.1289 -1823.8
## - exercise      1    0.00724 8.1294 -1823.7
## - match_es      1    0.00980 8.1320 -1823.6
## - shopping      1    0.01036 8.1325 -1823.6
## - theater       1    0.01423 8.1364 -1823.3
## - gaming        1    0.02453 8.1467 -1822.8
## - music         1    0.03366 8.1558 -1822.2
## <none>                8.1222 -1822.2
## - concerts     1    0.03529 8.1575 -1822.1
## - age          1    0.03701 8.1592 -1822.0
## - sinc_o       1    0.03792 8.1601 -1822.0
## + movies       1    0.00269 8.1195 -1820.3
## + hiking       1    0.00205 8.1201 -1820.3
## + tv           1    0.00196 8.1202 -1820.3
## + sports       1    0.00163 8.1205 -1820.2
## + dining       1    0.00087 8.1213 -1820.2
## + amb_o        1    0.00073 8.1214 -1820.2
## + clubbing     1    0.00032 8.1219 -1820.2
## + imprace      1    0.00021 8.1220 -1820.2
## + intel_o      1    0.00006 8.1221 -1820.2
## - yoga         1    0.08933 8.2115 -1819.0
## - gender       1    0.09853 8.2207 -1818.5
## + field_cd     16    0.45733 7.6648 -1817.2
## - tvsports     1    0.14227 8.2645 -1816.0
## + go_out       6    0.09774 8.0244 -1815.8
## + date         6    0.06395 8.0582 -1813.8

```

```

## - attr_o:fun_o 1 0.18962 8.3118 -1813.4
## + goal 5 0.01131 8.1109 -1812.8
## - career_c 16 0.96005 9.0822 -1802.0
##
## Step: AIC=-1823.94
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
## tvsports + exercise + museums + art + gaming + theater +
## concerts + music + shopping + yoga + attr_o:fun_o
##
## Df Sum of Sq RSS AIC
## - art 1 0.00487 8.1308 -1825.7
## - museums 1 0.00489 8.1308 -1825.7
## - exercise 1 0.00679 8.1327 -1825.5
## - shopping 1 0.00957 8.1355 -1825.4
## - match_es 1 0.01027 8.1362 -1825.3
## - theater 1 0.01611 8.1420 -1825.0
## - gaming 1 0.02382 8.1497 -1824.6
## - music 1 0.03395 8.1599 -1824.0
## <none> 8.1259 -1823.9
## - concerts 1 0.03597 8.1619 -1823.9
## - sinc_o 1 0.03651 8.1624 -1823.8
## - age 1 0.03841 8.1643 -1823.7
## + reading 1 0.00373 8.1222 -1822.2
## + movies 1 0.00238 8.1235 -1822.1
## + hiking 1 0.00231 8.1236 -1822.1
## + sports 1 0.00198 8.1239 -1822.1
## + tv 1 0.00182 8.1241 -1822.0
## + dining 1 0.00080 8.1251 -1822.0
## + amb_o 1 0.00075 8.1252 -1822.0
## + clubbing 1 0.00056 8.1254 -1822.0
## + imprace 1 0.00030 8.1256 -1822.0
## + intel_o 1 0.00003 8.1259 -1822.0
## - yoga 1 0.08921 8.2151 -1820.8
## - gender 1 0.09668 8.2226 -1820.4
## + field_cd 16 0.45089 7.6750 -1818.6
## - tvsports 1 0.13874 8.2646 -1818.0
## + go_out 6 0.09992 8.0260 -1817.7
## + date 6 0.06259 8.0633 -1815.5
## - attr_o:fun_o 1 0.19108 8.3170 -1815.1
## + goal 5 0.01141 8.1145 -1814.6
## - career_c 16 0.95666 9.0826 -1804.0
##
## Step: AIC=-1825.66
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
## tvsports + exercise + museums + gaming + theater + concerts +
## music + shopping + yoga + attr_o:fun_o
##
## Df Sum of Sq RSS AIC
## - exercise 1 0.00677 8.1375 -1827.3
## - match_es 1 0.00884 8.1396 -1827.2
## - shopping 1 0.00994 8.1407 -1827.1
## - theater 1 0.01472 8.1455 -1826.8
## - gaming 1 0.02462 8.1554 -1826.2
## - music 1 0.03344 8.1642 -1825.8

```

```

## <none> 8.1308 -1825.7
## - sinc_o 1 0.03553 8.1663 -1825.6
## - age 1 0.03756 8.1683 -1825.5
## - concerts 1 0.03900 8.1698 -1825.4
## - museums 1 0.04035 8.1711 -1825.3
## + art 1 0.00487 8.1259 -1823.9
## + reading 1 0.00455 8.1262 -1823.9
## + hiking 1 0.00242 8.1284 -1823.8
## + movies 1 0.00216 8.1286 -1823.8
## + tv 1 0.00186 8.1289 -1823.8
## + sports 1 0.00184 8.1289 -1823.8
## + amb_o 1 0.00093 8.1298 -1823.7
## + dining 1 0.00090 8.1299 -1823.7
## + clubbing 1 0.00075 8.1300 -1823.7
## + imprace 1 0.00024 8.1305 -1823.7
## + intel_o 1 0.00003 8.1308 -1823.7
## - yoga 1 0.08667 8.2175 -1822.7
## - gender 1 0.09922 8.2300 -1822.0
## + field_cd 16 0.45564 7.6751 -1820.6
## + go_out 6 0.10057 8.0302 -1819.5
## - tvsports 1 0.14433 8.2751 -1819.5
## + date 6 0.06293 8.0679 -1817.3
## - attr_o:fun_o 1 0.19717 8.3280 -1816.5
## + goal 5 0.01181 8.1190 -1816.3
## - career_c 16 0.96225 9.0930 -1805.4
##
## Step: AIC=-1827.28
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + match_es +
## tvsports + museums + gaming + theater + concerts + music +
## shopping + yoga + attr_o:fun_o
##
## Df Sum of Sq RSS AIC
## - match_es 1 0.00871 8.1463 -1828.8
## - shopping 1 0.00960 8.1472 -1828.7
## - theater 1 0.01388 8.1514 -1828.5
## - gaming 1 0.02503 8.1626 -1827.8
## - music 1 0.03317 8.1707 -1827.4
## <none> 8.1375 -1827.3
## - concerts 1 0.03531 8.1729 -1827.2
## - sinc_o 1 0.03637 8.1739 -1827.2
## - age 1 0.03805 8.1756 -1827.1
## - museums 1 0.04035 8.1779 -1827.0
## + exercise 1 0.00677 8.1308 -1825.7
## + art 1 0.00485 8.1327 -1825.5
## + sports 1 0.00483 8.1327 -1825.5
## + reading 1 0.00405 8.1335 -1825.5
## + hiking 1 0.00345 8.1341 -1825.5
## + movies 1 0.00252 8.1350 -1825.4
## + tv 1 0.00194 8.1356 -1825.4
## + amb_o 1 0.00071 8.1368 -1825.3
## + dining 1 0.00064 8.1369 -1825.3
## + clubbing 1 0.00056 8.1370 -1825.3
## + imprace 1 0.00044 8.1371 -1825.3
## + intel_o 1 0.00001 8.1375 -1825.3

```



```

## - yoga          1    0.08156 8.2191 -1824.6
## - gender        1    0.09856 8.2361 -1823.7
## + field_cd      16    0.45790 7.6797 -1822.3
## - tvsports      1    0.13786 8.2754 -1821.4
## + go_out        6    0.09690 8.0406 -1820.9
## + date          6    0.06319 8.0744 -1818.9
## - attr_o:fun_o  1    0.19744 8.3350 -1818.1
## + goal          5    0.01072 8.1268 -1817.9
## - career_c      16    0.95580 9.0933 -1807.4
##
## Step:  AIC=-1828.78
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      museums + gaming + theater + concerts + music + shopping +
##      yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS      AIC
## - shopping    1    0.01096 8.1572 -1830.2
## - theater      1    0.01391 8.1602 -1830.0
## - gaming       1    0.02516 8.1714 -1829.3
## - music        1    0.03340 8.1797 -1828.9
## - concerts     1    0.03390 8.1802 -1828.8
## - sinc_o       1    0.03411 8.1804 -1828.8
## <none>                8.1463 -1828.8
## - age          1    0.03774 8.1840 -1828.6
## - museums      1    0.03923 8.1855 -1828.5
## + match_es     1    0.00871 8.1375 -1827.3
## + exercise     1    0.00664 8.1396 -1827.2
## + reading      1    0.00438 8.1419 -1827.0
## + art          1    0.00343 8.1428 -1827.0
## + sports       1    0.00339 8.1429 -1827.0
## + hiking       1    0.00283 8.1434 -1826.9
## + tv           1    0.00279 8.1435 -1826.9
## + movies       1    0.00218 8.1441 -1826.9
## + dining       1    0.00101 8.1453 -1826.8
## + imprace      1    0.00093 8.1453 -1826.8
## + amb_o        1    0.00074 8.1455 -1826.8
## + clubbing     1    0.00016 8.1461 -1826.8
## + intel_o      1    0.00000 8.1463 -1826.8
## - yoga         1    0.08530 8.2316 -1825.9
## - gender       1    0.10553 8.2518 -1824.8
## + field_cd     16    0.46272 7.6835 -1824.1
## - tvsports     1    0.14438 8.2906 -1822.6
## + go_out       6    0.09666 8.0496 -1822.3
## + date         6    0.06212 8.0841 -1820.3
## - attr_o:fun_o 1    0.19476 8.3410 -1819.7
## + goal         5    0.01064 8.1356 -1819.4
## - career_c     16    0.96308 9.1093 -1808.6
##
## Step:  AIC=-1830.15
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      museums + gaming + theater + concerts + music + yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS      AIC
## - theater      1    0.01781 8.1750 -1831.1

```

```

## - gaming          1    0.02173 8.1790 -1830.9
## - sinc_o          1    0.03153 8.1888 -1830.3
## - age             1    0.03189 8.1891 -1830.3
## - concerts        1    0.03387 8.1911 -1830.2
## <none>              8.1572 -1830.2
## - museums         1    0.03597 8.1932 -1830.1
## - music           1    0.03727 8.1945 -1830.0
## + shopping        1    0.01096 8.1463 -1828.8
## + match_es        1    0.01007 8.1472 -1828.7
## + exercise        1    0.00628 8.1509 -1828.5
## + hiking          1    0.00559 8.1516 -1828.5
## + sports          1    0.00421 8.1530 -1828.4
## + dining          1    0.00408 8.1531 -1828.4
## + art             1    0.00366 8.1536 -1828.4
## + reading         1    0.00353 8.1537 -1828.3
## + movies          1    0.00313 8.1541 -1828.3
## + amb_o           1    0.00087 8.1564 -1828.2
## + imprace         1    0.00023 8.1570 -1828.2
## + tv              1    0.00014 8.1571 -1828.2
## + intel_o         1    0.00010 8.1571 -1828.2
## + clubbing        1    0.00001 8.1572 -1828.2
## - yoga            1    0.09341 8.2506 -1826.8
## - gender          1    0.09541 8.2526 -1826.7
## + field_cd        16    0.47077 7.6865 -1825.9
## - tvsports        1    0.14532 8.3025 -1823.9
## + go_out          6    0.09951 8.0577 -1823.9
## + date            6    0.05662 8.1006 -1821.4
## - attr_o:fun_o    1    0.19464 8.3519 -1821.1
## + goal            5    0.00875 8.1485 -1820.7
## - career_c        16    0.96888 9.1261 -1809.7
##
## Step:  AIC=-1831.13
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      museums + gaming + concerts + music + yoga + attr_o:fun_o
##
##              Df Sum of Sq  RSS    AIC
## - gaming          1    0.02126 8.1963 -1831.9
## - museums         1    0.02158 8.1966 -1831.9
## - concerts        1    0.02472 8.1998 -1831.7
## - sinc_o          1    0.02841 8.2034 -1831.5
## - age             1    0.03150 8.2065 -1831.3
## <none>              8.1750 -1831.1
## - music           1    0.03734 8.2124 -1831.0
## + theater         1    0.01781 8.1572 -1830.2
## + shopping        1    0.01486 8.1602 -1830.0
## + movies          1    0.01142 8.1636 -1829.8
## + match_es        1    0.01036 8.1647 -1829.7
## + hiking          1    0.00683 8.1682 -1829.5
## + sports          1    0.00654 8.1685 -1829.5
## + dining          1    0.00594 8.1691 -1829.5
## + exercise        1    0.00531 8.1697 -1829.4
## + reading         1    0.00520 8.1698 -1829.4
## + art             1    0.00239 8.1726 -1829.3
## + amb_o           1    0.00117 8.1739 -1829.2

```

```

## + intel_o      1  0.00027 8.1748 -1829.1
## + tv           1  0.00027 8.1748 -1829.1
## + clubbing     1  0.00005 8.1750 -1829.1
## + imprace      1  0.00000 8.1750 -1829.1
## - gender       1  0.08261 8.2576 -1828.4
## + field_cd     16  0.48400 7.6910 -1827.6
## - yoga         1  0.10022 8.2752 -1827.4
## - tvsports     1  0.13849 8.3135 -1825.3
## + go_out       6  0.10116 8.0739 -1824.9
## + date         6  0.05776 8.1173 -1822.4
## - attr_o:fun_o 1  0.19560 8.3706 -1822.1
## + goal         5  0.00769 8.1673 -1821.6
## - career_c     16  0.97252 9.1476 -1810.6
##
## Step: AIC=-1831.92
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      museums + concerts + music + yoga + attr_o:fun_o
##
##      Df Sum of Sq  RSS    AIC
## - museums      1  0.02045 8.2167 -1832.8
## - sinc_o       1  0.02848 8.2248 -1832.3
## - concerts     1  0.02996 8.2263 -1832.2
## <none>                8.1963 -1831.9
## - music        1  0.03826 8.2346 -1831.7
## - age          1  0.04115 8.2374 -1831.6
## + gaming       1  0.02126 8.1750 -1831.1
## + theater      1  0.01734 8.1790 -1830.9
## + hiking       1  0.01255 8.1837 -1830.6
## + movies       1  0.01113 8.1852 -1830.5
## + shopping     1  0.01081 8.1855 -1830.5
## + match_es     1  0.01025 8.1860 -1830.5
## + sports       1  0.00734 8.1890 -1830.3
## + exercise     1  0.00571 8.1906 -1830.2
## + reading      1  0.00458 8.1917 -1830.2
## + dining       1  0.00346 8.1928 -1830.1
## + art          1  0.00290 8.1934 -1830.1
## + amb_o        1  0.00155 8.1947 -1830.0
## - gender       1  0.06909 8.2654 -1830.0
## + intel_o      1  0.00045 8.1958 -1829.9
## + clubbing     1  0.00031 8.1960 -1829.9
## + imprace      1  0.00007 8.1962 -1829.9
## + tv           1  0.00004 8.1963 -1829.9
## + field_cd     16  0.49032 7.7060 -1828.7
## - yoga         1  0.09357 8.2899 -1828.6
## - tvsports     1  0.12288 8.3192 -1827.0
## + go_out       6  0.10101 8.0953 -1825.7
## + date         6  0.05494 8.1414 -1823.1
## - attr_o:fun_o 1  0.19428 8.3906 -1823.0
## + goal         5  0.00682 8.1895 -1822.3
## - career_c     16  0.95779 9.1541 -1812.3
##
## Step: AIC=-1832.75
## dec_o ~ attr_o + sinc_o + fun_o + gender + age + career_c + tvsports +
##      concerts + music + yoga + attr_o:fun_o

```

```

##
##          Df Sum of Sq    RSS    AIC
## - sinc_o      1  0.02908 8.2458 -1833.1
## - age          1  0.03512 8.2519 -1832.8
## <none>                8.2167 -1832.8
## - music        1  0.03576 8.2525 -1832.7
## - concerts     1  0.04445 8.2612 -1832.2
## + art          1  0.02191 8.1948 -1832.0
## + museums      1  0.02045 8.1963 -1831.9
## + gaming       1  0.02013 8.1966 -1831.9
## + hiking       1  0.01605 8.2007 -1831.7
## + match_es     1  0.00895 8.2078 -1831.3
## + sports       1  0.00670 8.2100 -1831.1
## + shopping     1  0.00652 8.2102 -1831.1
## + exercise     1  0.00623 8.2105 -1831.1
## + movies       1  0.00609 8.2107 -1831.1
## + theater      1  0.00343 8.2133 -1831.0
## + amb_o        1  0.00112 8.2156 -1830.8
## + clubbing     1  0.00087 8.2159 -1830.8
## + intel_o      1  0.00054 8.2162 -1830.8
## + reading      1  0.00047 8.2163 -1830.8
## + imprace      1  0.00033 8.2164 -1830.8
## + dining       1  0.00014 8.2166 -1830.8
## + tv           1  0.00009 8.2167 -1830.8
## - yoga         1  0.08221 8.2989 -1830.1
## - gender       1  0.08475 8.3015 -1830.0
## + field_cd     16  0.48009 7.7366 -1828.9
## - tvsports     1  0.12444 8.3412 -1827.7
## + go_out       6  0.10258 8.1142 -1826.6
## - attr_o:fun_o 1  0.19304 8.4098 -1823.9
## + date         6  0.05462 8.1621 -1823.9
## + goal         5  0.00780 8.2089 -1823.2
## - career_c     16  0.99043 9.2072 -1811.6
##
## Step:  AIC=-1833.1
## dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##      concerts + music + yoga + attr_o:fun_o
##
##          Df Sum of Sq    RSS    AIC
## <none>                8.2458 -1833.1
## - music        1  0.03762 8.2834 -1833.0
## - age          1  0.03936 8.2852 -1832.9
## + sinc_o       1  0.02908 8.2167 -1832.8
## - concerts     1  0.04617 8.2920 -1832.5
## + art          1  0.02165 8.2242 -1832.3
## + museums      1  0.02104 8.2248 -1832.3
## + gaming       1  0.02019 8.2256 -1832.2
## + hiking       1  0.01129 8.2345 -1831.7
## + intel_o      1  0.00748 8.2383 -1831.5
## + exercise     1  0.00712 8.2387 -1831.5
## + match_es     1  0.00663 8.2392 -1831.5
## + amb_o        1  0.00601 8.2398 -1831.4
## + sports       1  0.00572 8.2401 -1831.4
## + movies       1  0.00534 8.2405 -1831.4

```

```
## + shopping      1  0.00442 8.2414 -1831.3
## + theater       1  0.00217 8.2437 -1831.2
## + clubbing      1  0.00123 8.2446 -1831.2
## + imprace       1  0.00044 8.2454 -1831.1
## + tv            1  0.00009 8.2457 -1831.1
## + reading       1  0.00008 8.2457 -1831.1
## + dining        1  0.00000 8.2458 -1831.1
## - gender        1  0.07961 8.3254 -1830.6
## - yoga          1  0.08462 8.3304 -1830.3
## + field_cd      16  0.48990 7.7559 -1829.7
## - tvsports      1  0.13051 8.3763 -1827.8
## + go_out        6  0.10270 8.1431 -1827.0
## - attr_o:fun_o  1  0.19506 8.4409 -1824.2
## + date          6  0.05402 8.1918 -1824.2
## + goal          5  0.00757 8.2383 -1823.5
## - career_c      16  0.96729 9.2131 -1813.3
```

```
step$anova # display results
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
```

```
## Initial Model:
```

```
## dec_o ~ iid + attr_o + sinc_o + intel_o + fun_o + amb_o + gender +
##   age + field_cd + imprace + goal + date + go_out + career_c +
##   match_es + sports + tvsports + exercise + dining + museums +
##   art + hiking + gaming + clubbing + reading + tv + theater +
##   movies + concerts + music + shopping + yoga + attr_o * fun_o -
##   iid
```

```
##
```

```
## Final Model:
```

```
## dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##   concerts + music + yoga + attr_o:fun_o
```

```
##
```

```
##
```

	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
## 1				390	7.407788	-1781.153
## 2	- goal	5	1.618567e-02	395	7.423973	-1790.134
## 3	- date	6	7.827952e-02	401	7.502253	-1797.236
## 4	- go_out	6	1.239291e-01	407	7.626182	-1801.584
## 5	- field_cd	16	4.809918e-01	423	8.107174	-1805.022
## 6	- imprace	1	4.608595e-05	424	8.107220	-1807.019
## 7	- dining	1	8.236220e-04	425	8.108043	-1808.972
## 8	- clubbing	1	1.018732e-03	426	8.109062	-1810.913
## 9	- intel_o	1	1.217020e-03	427	8.110279	-1812.843
## 10	- amb_o	1	1.116978e-03	428	8.111396	-1814.779
## 11	- sports	1	1.452237e-03	429	8.112848	-1816.695
## 12	- hiking	1	2.667011e-03	430	8.115515	-1818.542
## 13	- tv	1	3.968803e-03	431	8.119484	-1820.313
## 14	- movies	1	2.694675e-03	432	8.122179	-1822.158
## 15	- reading	1	3.733004e-03	433	8.125912	-1823.944
## 16	- art	1	4.871151e-03	434	8.130783	-1825.664
## 17	- exercise	1	6.766252e-03	435	8.137549	-1827.275
## 18	- match_es	1	8.713508e-03	436	8.146263	-1828.776
## 19	- shopping	1	1.096179e-02	437	8.157225	-1830.148

```
## 20 - theater 1 1.780664e-02      438  8.175031 -1831.129
## 21 - gaming 1 2.126151e-02      439  8.196293 -1831.916
## 22 - museums 1 2.044857e-02     440  8.216741 -1832.753
## 23 - sinc_o 1 2.908015e-02     441  8.245822 -1833.103
```

```
anova(lm(dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
  concerts + music + yoga + attr_o:fun_o, data = na.omit(augmented_attris)), reduced_augmented_fit)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Model 1: dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##      concerts + music + yoga + attr_o:fun_o
```

```
## Model 2: dec_o ~ attr_o + fun_o + career_c + gender + age + tvsports +
##      yoga + attr_o * fun_o
```

```
##      Res.Df    RSS Df Sum of Sq      F Pr(>F)
```

```
## 1      441 8.2458
```

```
## 2      443 8.2983 -2 -0.052433 1.4021 0.2472
```

```
anova(lm(dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports + yoga + attr_o:fun_o, data = na.omit(augmented_attris)), reduced_augmented_fit)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Model 1: dec_o ~ attr_o + fun_o + gender + age + career_c + tvsports +
##      yoga + attr_o:fun_o
```

```
## Model 2: dec_o ~ attr_o + fun_o + gender + career_c + tvsports + yoga +
##      attr_o:fun_o
```

```
##      Res.Df    RSS Df Sum of Sq      F Pr(>F)
```

```
## 1      443 8.2983
```

```
## 2      444 8.3387 -1 -0.040468 2.1604 0.1423
```

```
summary(lm(dec_o ~ attr_o + fun_o + gender + career_c + tvsports + yoga + attr_o:fun_o, data = na.omit(augmented_attris)))
```

```
##
```

```
## Call:
```

```
## lm(formula = dec_o ~ attr_o + fun_o + gender + career_c + tvsports +
##      yoga + attr_o:fun_o, data = na.omit(augmented_attris))
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -0.43002 -0.08564 -0.00169  0.07653  0.44276
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -0.0674879  0.1895965  -0.356  0.72204
```

```
## attr_o      0.0186050  0.0330850   0.562  0.57417
```

```
## fun_o      -0.0330816  0.0305317  -1.084  0.27917
```

```
## gender1    -0.0317748  0.0146816  -2.164  0.03098 *
```

```
## career_c2   0.0191187  0.0257934   0.741  0.45895
```

```
## career_c3  -0.1046515  0.0408127  -2.564  0.01067 *
```

```
## career_c4   0.0025634  0.0359659   0.071  0.94321
```

```
## career_c5  -0.0100657  0.0441264  -0.228  0.81966
```

```
## career_c6  -0.0302373  0.0312858  -0.966  0.33433
```

```
## career_c7   0.0008058  0.0255804   0.031  0.97489
```

```
## career_c8  -0.5019695  0.1007908  -4.980 9.11e-07 ***
```

```
## career_c9  -0.0162589  0.0346141  -0.470  0.63879
```

```
## career_c10  0.0027604  0.0368671   0.075  0.94035
```

```
## career_c11    -0.0267629  0.0452110  -0.592  0.55418
## career_c12    -0.2345931  0.1003723  -2.337  0.01987 *
## career_c13     0.0394800  0.0657128   0.601  0.54828
## career_c14     0.2187081  0.1396527   1.566  0.11804
## career_c15     0.0987607  0.0734641   1.344  0.17952
## career_c16     0.0541056  0.1004389   0.539  0.59037
## career_c17    -0.0558462  0.1393390  -0.401  0.68877
## tvsports      -0.0056223  0.0023336  -2.409  0.01639 *
## yoga          -0.0049050  0.0024472  -2.004  0.04564 *
## attr_o:fun_o   0.0161394  0.0050170   3.217  0.00139 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.137 on 444 degrees of freedom
## Multiple R-squared:  0.6789, Adjusted R-squared:  0.663
## F-statistic: 42.66 on 22 and 444 DF,  p-value: < 2.2e-16
final_reduced<-lm(dec_o ~ attr_o + fun_o + gender + career_c + tvsports + yoga + attr_o:fun_o, data = na.omit(augmented_data))
```

The residual plot appears to indicate some heteroskedasticity. Let's check it out.

```
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
```

```
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      as.Date, as.Date.numeric
```

```
bptest(dec_o ~ attr_o + fun_o + gender + career_c + tvsports + yoga + attr_o:fun_o, data = na.omit(augmented_data))
```

```
##
```

```
## studentized Breusch-Pagan test
```

```
##
```

```
## data:  dec_o ~ attr_o + fun_o + gender + career_c + tvsports + yoga + attr_o:fun_o
```

```
## BP = 33.372, df = 22, p-value = 0.05682
```

The p-value of the BP test is above 0.05 so we are kosher.

How can we get dates?

(credit: Kev) Based on the model identified above, we want to help you get a date. What attributes predict fun or attractiveness? Are we any good at knowing how fun/attractive we are?

Predicting likeability

This question is “can we predict how fun someone is perceived based on self-reported attributes” not other’s subjective opinions (i.e. can we predict if you will be perceived as fun based on your favorite activities, not on if someone else thinks you are attractive).

So what makes us fun?

```
fun_augmented_fit <- lm(fun_o ~ .-iid -attr_o -sinc_o -intel_o -dec_o -amb_o, data = na.omit(augmented_
summary(fun_augmented_fit)
```

```
##
## Call:
## lm(formula = fun_o ~ . - iid - attr_o - sinc_o - intel_o - dec_o -
##     amb_o, data = na.omit(augmented_attris))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.6511 -0.5688  0.0840  0.6334  2.0586
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.210572   0.669942   9.270  <2e-16 ***
## gender1      -0.229676   0.123252  -1.863   0.0631 .
## age         -0.008037   0.013337  -0.603   0.5471
## field_cd2    -0.897353   0.469088  -1.913   0.0565 .
## field_cd3    -0.530350   0.410281  -1.293   0.1969
## field_cd4     0.402071   0.541581   0.742   0.4583
## field_cd5    -0.866024   0.394683  -2.194   0.0288 *
## field_cd6    -0.316010   0.451710  -0.700   0.4846
## field_cd7    -0.376142   0.435122  -0.864   0.3879
## field_cd8    -0.260141   0.380137  -0.684   0.4942
## field_cd9    -0.637274   0.406784  -1.567   0.1180
## field_cd10   -0.553891   0.399249  -1.387   0.1661
## field_cd11   -0.548934   0.453778  -1.210   0.2271
## field_cd12   -0.037692   1.054175  -0.036   0.9715
## field_cd13   -0.099640   0.400671  -0.249   0.8037
## field_cd14   -0.139998   0.551015  -0.254   0.7996
## field_cd15   -0.094374   0.509408  -0.185   0.8531
## field_cd16   -0.443167   0.614224  -0.722   0.4710
## field_cd17    1.064434   0.979778   1.086   0.2780
## field_cd18    0.006051   0.778049   0.008   0.9938
## imprace     -0.005097   0.017456  -0.292   0.7705
## goal2        -0.205157   0.110543  -1.856   0.0642 .
## goal3        -0.169922   0.175742  -0.967   0.3342
## goal4        -0.379071   0.258070  -1.469   0.1427
## goal5         0.146505   0.202010   0.725   0.4687
## goal6        -0.108374   0.221033  -0.490   0.6242
## date2        -0.086379   0.483600  -0.179   0.8583
## date3         0.409029   0.455771   0.897   0.3700
## date4         0.126883   0.446141   0.284   0.7763
## date5         0.062396   0.450605   0.138   0.8899
## date6         0.094690   0.448104   0.211   0.8328
## date7         0.064855   0.454852   0.143   0.8867
## go_out2      -0.017249   0.119545  -0.144   0.8853
## go_out3      -0.195459   0.131641  -1.485   0.1384
## go_out4      -0.268010   0.225247  -1.190   0.2348
## go_out5      -0.195132   0.335274  -0.582   0.5609
## go_out6      -0.756045   0.424445  -1.781   0.0756 .
```



```

## go_out7      -1.301412    0.995211   -1.308    0.1917
## career_c2     0.089209    0.372323    0.240    0.8108
## career_c3     0.763740    0.457264    1.670    0.0957 .
## career_c4     0.171764    0.441788    0.389    0.6976
## career_c5     0.547319    0.468742    1.168    0.2437
## career_c6     0.324049    0.410779    0.789    0.4307
## career_c7     0.253097    0.372929    0.679    0.4977
## career_c8     1.837990    0.775311    2.371    0.0182 *
## career_c9     0.416178    0.412072    1.010    0.3131
## career_c10    0.424030    0.414201    1.024    0.3066
## career_c11    0.391842    0.512022    0.765    0.4446
## career_c12    0.059283    0.785894    0.075    0.9399
## career_c13   -0.482718    0.576560   -0.837    0.4030
## career_c14   -1.314336    0.987929   -1.330    0.1842
## career_c15   -1.525895    0.716411   -2.130    0.0338 *
## career_c16   -0.459790    0.810198   -0.568    0.5707
## career_c17      NA         NA         NA         NA
## match_es     -0.014087    0.020678   -0.681    0.4961
## sports        0.037949    0.023802    1.594    0.1117
## tvsports     -0.013686    0.020423   -0.670    0.5032
## exercise      0.046365    0.021560    2.151    0.0321 *
## dining        0.040898    0.030209    1.354    0.1766
## museums       -0.042559    0.047033   -0.905    0.3661
## art           0.057768    0.041676    1.386    0.1665
## hiking        0.008638    0.020641    0.419    0.6758
## gaming        -0.026038    0.020658   -1.260    0.2082
## clubbing      0.045650    0.020003    2.282    0.0230 *
## reading       0.001733    0.025570    0.068    0.9460
## tv            -0.035526    0.023617   -1.504    0.1333
## theater       -0.010336    0.030219   -0.342    0.7325
## movies        -0.023670    0.034329   -0.690    0.4909
## concerts     -0.052725    0.031086   -1.696    0.0907 .
## music         0.060682    0.033553    1.809    0.0713 .
## shopping      0.002255    0.023150    0.097    0.9224
## yoga          0.009521    0.018740    0.508    0.6117
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9299 on 396 degrees of freedom
## Multiple R-squared:  0.2894, Adjusted R-squared:  0.1638
## F-statistic: 2.304 on 70 and 396 DF,  p-value: 2.471e-07
# reduce model with stepAIC
step <- stepAIC(fun_augmented_fit, direction="both")

## Start:  AIC=-2.85
## fun_o ~ (iid + dec_o + attr_o + sinc_o + intel_o + amb_o + gender +
##   age + field_cd + imprace + goal + date + go_out + career_c +
##   match_es + sports + tvsports + exercise + dining + museums +
##   art + hiking + gaming + clubbing + reading + tv + theater +
##   movies + concerts + music + shopping + yoga) - iid - attr_o -
##   sinc_o - intel_o - dec_o - amb_o
##
##           Df Sum of Sq  RSS    AIC
## - date      6   4.4816 346.94 -8.7800

```

```

## - go_out      6      6.6130 349.07 -5.9199
## - field_cd 16     22.2163 364.68 -5.4984
## - goal        5      5.6584 348.12 -5.1987
## - reading     1      0.0040 342.46 -4.8464
## - shopping    1      0.0082 342.47 -4.8406
## - imprace     1      0.0737 342.53 -4.7513
## - theater     1      0.1012 342.56 -4.7139
## - hiking      1      0.1515 342.61 -4.6453
## - yoga        1      0.2232 342.68 -4.5475
## - age         1      0.3140 342.77 -4.4238
## - tvsports    1      0.3883 342.85 -4.3225
## - match_es    1      0.4013 342.86 -4.3048
## - movies      1      0.4111 342.87 -4.2915
## - museums     1      0.7081 343.17 -3.8872
## - gaming      1      1.3740 343.83 -2.9819
## <none>                342.46 -2.8518
## - dining      1      1.5850 344.04 -2.6953
## - art         1      1.6616 344.12 -2.5915
## - tv          1      1.9568 344.42 -2.1910
## - sports      1      2.1982 344.66 -1.8638
## - concerts    1      2.4878 344.95 -1.4715
## - music       1      2.8286 345.29 -1.0104
## - gender      1      3.0030 345.46 -0.7745
## - career_c 15     25.0310 367.49  0.0922
## - exercise    1      3.9996 346.46  0.5708
## - clubbing    1      4.5039 346.96  1.2499
##
## Step:  AIC=-8.78
## fun_o ~ gender + age + field_cd + imprace + goal + go_out + career_c +
##      match_es + sports + tvsports + exercise + dining + museums +
##      art + hiking + gaming + clubbing + reading + tv + theater +
##      movies + concerts + music + shopping + yoga
##
##           Df Sum of Sq   RSS   AIC
## - field_cd 16     22.1580 369.10 -11.8680
## - shopping  1      0.0003 346.94 -10.7797
## - reading   1      0.0009 346.94 -10.7788
## - hiking    1      0.0700 347.01 -10.6858
## - theater   1      0.1095 347.05 -10.6326
## - imprace   1      0.1658 347.11 -10.5570
## - yoga      1      0.2211 347.16 -10.4825
## - goal      5      6.2816 353.22 -10.4003
## - match_es  1      0.3065 347.25 -10.3677
## - age       1      0.3181 347.26 -10.3520
## - movies    1      0.4543 347.40 -10.1689
## - tvsports  1      0.5046 347.45 -10.1014
## - go_out    6      8.2064 355.15  -9.8625
## - museums   1      0.7198 347.66  -9.8121
## - gaming    1      1.3161 348.26  -9.0118
## <none>                346.94  -8.7800
## - dining    1      1.5479 348.49  -8.7011
## - tv        1      1.5895 348.53  -8.6453
## - art       1      1.8835 348.82  -8.2516
## - sports    1      2.4265 349.37  -7.5252

```

```

## - gender      1      2.5123 349.45 -7.4105
## - concerts   1      2.6322 349.57 -7.2503
## - music       1      2.8453 349.79 -6.9657
## - career_c   15     25.1908 372.13 -6.0465
## - clubbing    1      4.3696 351.31 -4.9351
## - exercise    1      4.3776 351.32 -4.9244
## + date        6      4.4816 342.46 -2.8518
##
## Step: AIC=-11.87
## fun_o ~ gender + age + imprace + goal + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + hiking +
##      gaming + clubbing + reading + tv + theater + movies + concerts +
##      music + shopping + yoga
##
##           Df Sum of Sq    RSS    AIC
## - shopping  1      0.000 369.10 -13.8676
## - hiking    1      0.024 369.12 -13.8374
## - age       1      0.040 369.14 -13.8168
## - movies    1      0.087 369.19 -13.7580
## - yoga      1      0.120 369.22 -13.7166
## - imprace   1      0.148 369.25 -13.6807
## - reading   1      0.252 369.35 -13.5489
## - match_es  1      0.270 369.37 -13.5268
## - theater   1      0.275 369.37 -13.5196
## - goal      5      6.832 375.93 -13.3024
## - tvsports  1      0.460 369.56 -13.2862
## - museums   1      0.614 369.71 -13.0917
## <none>              369.10 -11.8680
## - tv        1      1.743 370.84 -11.6683
## - art        1      1.750 370.85 -11.6593
## - dining     1      1.966 371.07 -11.3873
## - gaming     1      2.233 371.33 -11.0512
## - sports     1      2.314 371.41 -10.9489
## - concerts  1      2.569 371.67 -10.6294
## - music      1      2.722 371.82 -10.4368
## - exercise   1      2.843 371.94 -10.2847
## - clubbing   1      3.467 372.57 -9.5022
## - gender     1      3.743 372.84 -9.1556
## - go_out     6     12.082 381.18 -8.8257
## + field_cd  16     22.158 346.94 -8.7800
## + date       6      4.423 364.68 -5.4984
## - career_c   16     37.226 406.33  1.0049
##
## Step: AIC=-13.87
## fun_o ~ gender + age + imprace + goal + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + hiking +
##      gaming + clubbing + reading + tv + theater + movies + concerts +
##      music + yoga
##
##           Df Sum of Sq    RSS    AIC
## - hiking    1      0.024 369.12 -15.8372
## - age       1      0.040 369.14 -15.8167
## - movies    1      0.087 369.19 -15.7580
## - yoga      1      0.120 369.22 -15.7155

```

```

## - imprace 1 0.149 369.25 -15.6792
## - reading 1 0.254 369.35 -15.5460
## - match_es 1 0.273 369.37 -15.5226
## - theater 1 0.276 369.38 -15.5180
## - tvsports 1 0.461 369.56 -15.2850
## - goal 5 6.857 375.96 -15.2715
## - museums 1 0.619 369.72 -15.0851
## <none> 369.10 -13.8676
## - art 1 1.752 370.85 -13.6558
## - tv 1 1.939 371.04 -13.4204
## - dining 1 2.134 371.23 -13.1758
## - gaming 1 2.241 371.34 -13.0406
## - sports 1 2.315 371.41 -12.9482
## - concerts 1 2.586 371.69 -12.6077
## - music 1 2.723 371.82 -12.4354
## - exercise 1 2.844 371.94 -12.2835
## + shopping 1 0.000 369.10 -11.8680
## - clubbing 1 3.471 372.57 -11.4964
## - gender 1 3.856 372.96 -11.0137
## - go_out 6 12.090 381.19 -10.8163
## + field_cd 16 22.158 346.94 -10.7797
## + date 6 4.424 364.68 -7.4983
## - career_c 16 37.289 406.39 -0.9219
##
## Step: AIC=-15.84
## fun_o ~ gender + age + imprace + goal + go_out + career_c + match_es +
## sports + tvsports + exercise + dining + museums + art + gaming +
## clubbing + reading + tv + theater + movies + concerts + music +
## yoga
##
## Df Sum of Sq RSS AIC
## - age 1 0.047 369.17 -17.7782
## - movies 1 0.087 369.21 -17.7268
## - yoga 1 0.106 369.23 -17.7027
## - imprace 1 0.152 369.28 -17.6444
## - reading 1 0.258 369.38 -17.5110
## - theater 1 0.273 369.40 -17.4917
## - match_es 1 0.275 369.40 -17.4900
## - tvsports 1 0.448 369.57 -17.2706
## - goal 5 6.835 375.96 -17.2692
## - museums 1 0.633 369.76 -17.0373
## <none> 369.12 -15.8372
## - art 1 1.740 370.86 -15.6410
## - tv 1 1.920 371.04 -15.4140
## - dining 1 2.129 371.25 -15.1520
## - sports 1 2.318 371.44 -14.9143
## - gaming 1 2.531 371.65 -14.6462
## - concerts 1 2.648 371.77 -14.4989
## - music 1 2.818 371.94 -14.2855
## - exercise 1 2.831 371.95 -14.2690
## + hiking 1 0.024 369.10 -13.8676
## + shopping 1 0.000 369.12 -13.8374
## - clubbing 1 3.577 372.70 -13.3338
## - gender 1 3.844 372.97 -12.9996

```

```

## - go_out      6      12.120 381.24 -12.7503
## + field_cd 16      22.108 347.02 -12.6799
## + date        6       4.445 364.68  -9.4953
## - career_c 16      37.536 406.66  -2.6112
##
## Step:  AIC=-17.78
## fun_o ~ gender + imprace + goal + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + gaming +
##      clubbing + reading + tv + theater + movies + concerts + music +
##      yoga
##
##           Df Sum of Sq    RSS      AIC
## - movies    1      0.086 369.26 -19.6700
## - yoga       1      0.103 369.27 -19.6478
## - imprace    1      0.139 369.31 -19.6022
## - reading    1      0.248 369.42 -19.4651
## - match_es   1      0.265 369.44 -19.4431
## - theater    1      0.285 369.46 -19.4172
## - tvsports   1      0.442 369.61 -19.2188
## - museums    1      0.634 369.80 -18.9766
## - goal       5      7.426 376.60 -18.4771
## <none>                369.17 -17.7782
## - art         1      1.719 370.89 -17.6082
## - tv           1      1.890 371.06 -17.3932
## - dining      1      2.099 371.27 -17.1303
## - sports      1      2.309 371.48 -16.8658
## - gaming      1      2.485 371.66 -16.6454
## - concerts    1      2.614 371.78 -16.4828
## - music       1      2.815 371.99 -16.2304
## - exercise    1      2.907 372.08 -16.1147
## + age         1      0.047 369.12 -15.8372
## + hiking      1      0.030 369.14 -15.8167
## + shopping    1      0.002 369.17 -15.7806
## - clubbing    1      3.647 372.82 -15.1870
## - gender      1      3.859 373.03 -14.9220
## - go_out      6     12.254 381.42 -14.5287
## + field_cd 16     21.864 347.31 -14.2888
## + date        6      4.439 364.73 -11.4277
## - career_c 16     37.521 406.69  -4.5744
##
## Step:  AIC=-19.67
## fun_o ~ gender + imprace + goal + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + gaming +
##      clubbing + reading + tv + theater + concerts + music + yoga
##
##           Df Sum of Sq    RSS      AIC
## - yoga       1      0.113 369.37 -21.5277
## - imprace    1      0.135 369.39 -21.4992
## - match_es   1      0.256 369.51 -21.3457
## - reading    1      0.260 369.52 -21.3412
## - tvsports   1      0.407 369.66 -21.1559
## - theater    1      0.472 369.73 -21.0734
## - museums    1      0.630 369.89 -20.8739
## - goal       5      7.504 376.76 -20.2749

```

```

## <none>          369.26 -19.6700
## - art           1      1.711 370.97 -19.5109
## - dining        1      2.062 371.32 -19.0697
## - sports         1      2.267 371.52 -18.8111
## - tv             1      2.391 371.65 -18.6563
## - gaming         1      2.440 371.70 -18.5941
## - music          1      2.738 371.99 -18.2202
## - concerts       1      2.777 372.03 -18.1716
## - exercise       1      2.981 372.24 -17.9145
## + movies         1      0.086 369.17 -17.7782
## + age            1      0.045 369.21 -17.7268
## + hiking         1      0.031 369.22 -17.7092
## + shopping       1      0.003 369.25 -17.6739
## - clubbing       1      3.613 372.87 -17.1224
## - gender         1      3.946 373.20 -16.7056
## - go_out         6     12.264 381.52 -16.4112
## + field_cd      16     21.506 347.75 -15.6926
## + date           6      4.464 364.79 -13.3501
## - career_c      16     37.472 406.73  -6.5325
##
## Step: AIC=-21.53
## fun_o ~ gender + imprace + goal + go_out + career_c + match_es +
##      sports + tvsports + exercise + dining + museums + art + gaming +
##      clubbing + reading + tv + theater + concerts + music
##
##           Df Sum of Sq   RSS   AIC
## - imprace   1      0.162 369.53 -23.323
## - match_es  1      0.238 369.61 -23.227
## - reading   1      0.257 369.63 -23.203
## - theater   1      0.452 369.82 -22.956
## - tvsports  1      0.461 369.83 -22.946
## - museums   1      0.617 369.99 -22.749
## - goal      5      7.502 376.87 -22.137
## <none>          369.37 -21.528
## - art        1      1.787 371.16 -21.274
## - dining     1      2.026 371.39 -20.973
## - sports     1      2.317 371.69 -20.608
## - tv         1      2.323 371.69 -20.600
## - gaming     1      2.383 371.75 -20.525
## - concerts   1      2.678 372.05 -20.154
## - music      1      2.721 372.09 -20.101
## + yoga       1      0.113 369.26 -19.670
## + movies     1      0.095 369.27 -19.648
## + age        1      0.041 369.33 -19.580
## + hiking     1      0.014 369.35 -19.546
## + shopping   1      0.008 369.36 -19.538
## - exercise   1      3.180 372.55 -19.525
## - clubbing   1      3.587 372.96 -19.014
## - gender     1      4.264 373.63 -18.167
## - go_out     6     12.537 381.91 -17.940
## + field_cd   16     21.328 348.04 -17.302
## + date       6      4.465 364.90 -15.208
## - career_c   16     37.366 406.73  -8.525
##

```

```

## Step: AIC=-23.32
## fun_o ~ gender + goal + go_out + career_c + match_es + sports +
##      tvsports + exercise + dining + museums + art + gaming + clubbing +
##      reading + tv + theater + concerts + music
##
##      Df Sum of Sq    RSS    AIC
## - match_es  1      0.211 369.74 -25.057
## - reading   1      0.283 369.81 -24.966
## - tvsports  1      0.462 369.99 -24.740
## - theater   1      0.474 370.00 -24.724
## - museums   1      0.659 370.19 -24.492
## - goal      5      7.452 376.98 -23.999
## <none>                369.53 -23.323
## - art       1      1.852 371.38 -22.989
## - dining    1      2.008 371.54 -22.793
## - sports    1      2.358 371.89 -22.353
## - gaming    1      2.404 371.93 -22.296
## - concerts  1      2.615 372.14 -22.030
## - music     1      2.719 372.25 -21.900
## - tv        1      2.810 372.34 -21.785
## + imprace   1      0.162 369.37 -21.528
## + yoga      1      0.139 369.39 -21.499
## + movies    1      0.091 369.44 -21.439
## - exercise  1      3.103 372.63 -21.418
## + age       1      0.028 369.50 -21.358
## + hiking    1      0.015 369.52 -21.342
## + shopping  1      0.005 369.52 -21.330
## - clubbing  1      3.744 373.27 -20.616
## - gender    1      4.141 373.67 -20.119
## - go_out    6     12.459 381.99 -19.838
## + field_cd 16     21.338 348.19 -19.100
## + date      6      4.557 364.97 -17.118
## - career_c 16     37.464 406.99 -10.227
##
## Step: AIC=-25.06
## fun_o ~ gender + goal + go_out + career_c + sports + tvsports +
##      exercise + dining + museums + art + gaming + clubbing + reading +
##      tv + theater + concerts + music
##
##      Df Sum of Sq    RSS    AIC
## - reading   1      0.254 369.99 -26.737
## - theater    1      0.481 370.22 -26.450
## - tvsports   1      0.493 370.23 -26.435
## - museums    1      0.595 370.34 -26.306
## - goal       5      7.532 377.27 -25.639
## <none>                369.74 -25.057
## - art       1      1.741 371.48 -24.863
## - dining    1      1.910 371.65 -24.651
## - sports    1      2.227 371.97 -24.253
## - gaming    1      2.370 372.11 -24.073
## - concerts  1      2.642 372.38 -23.731
## - music     1      2.690 372.43 -23.672
## - tv        1      2.738 372.48 -23.612
## + match_es  1      0.211 369.53 -23.323

```

```

## + imprace 1 0.135 369.61 -23.227
## + yoga 1 0.117 369.62 -23.205
## + movies 1 0.083 369.66 -23.162
## - exercise 1 3.155 372.90 -23.090
## + age 1 0.022 369.72 -23.085
## + hiking 1 0.016 369.72 -23.078
## + shopping 1 0.001 369.74 -23.058
## - clubbing 1 3.588 373.33 -22.547
## - gender 1 4.329 374.07 -21.621
## - go_out 6 12.643 382.38 -21.355
## + field_cd 16 21.321 348.42 -20.794
## + date 6 4.368 365.37 -18.607
## - career_c 16 37.254 407.00 -12.226
##
## Step: AIC=-26.74
## fun_o ~ gender + goal + go_out + career_c + sports + tvsports +
## exercise + dining + museums + art + gaming + clubbing + tv +
## theater + concerts + music
##
## Df Sum of Sq RSS AIC
## - theater 1 0.411 370.41 -28.218
## - museums 1 0.456 370.45 -28.162
## - tvsports 1 0.560 370.55 -28.030
## - goal 5 7.642 377.64 -27.190
## <none> 369.99 -26.737
## - art 1 1.643 371.64 -26.668
## - dining 1 1.894 371.89 -26.352
## - sports 1 2.143 372.14 -26.040
## - gaming 1 2.310 372.30 -25.830
## - concerts 1 2.679 372.67 -25.368
## - music 1 2.696 372.69 -25.346
## - tv 1 2.787 372.78 -25.233
## + reading 1 0.254 369.74 -25.057
## + match_es 1 0.181 369.81 -24.966
## + imprace 1 0.159 369.84 -24.937
## + yoga 1 0.118 369.88 -24.886
## + movies 1 0.095 369.90 -24.857
## + hiking 1 0.019 369.98 -24.761
## + age 1 0.015 369.98 -24.756
## + shopping 1 0.000 369.99 -24.737
## - exercise 1 3.310 373.30 -24.578
## - clubbing 1 3.438 373.43 -24.417
## - gender 1 4.413 374.41 -23.199
## - go_out 6 12.519 382.51 -23.197
## + field_cd 16 21.573 348.42 -22.792
## + date 6 4.368 365.63 -20.283
## - career_c 16 37.155 407.15 -14.049
##
## Step: AIC=-28.22
## fun_o ~ gender + goal + go_out + career_c + sports + tvsports +
## exercise + dining + museums + art + gaming + clubbing + tv +
## concerts + music
##
## Df Sum of Sq RSS AIC

```



```

## - tvsports 1      0.510 370.92 -29.575
## - museums  1      0.670 371.08 -29.374
## - goal     5      7.507 377.91 -28.848
## - art      1      1.526 371.93 -28.298
## <none>          370.41 -28.218
## - dining   1      1.809 372.21 -27.943
## - gaming   1      2.249 372.65 -27.392
## - sports   1      2.426 372.83 -27.169
## + theater  1      0.411 369.99 -26.737
## - music    1      2.821 373.23 -26.675
## + movies   1      0.268 370.14 -26.556
## + match_es 1      0.191 370.22 -26.458
## + reading  1      0.184 370.22 -26.450
## + imprace  1      0.175 370.23 -26.439
## + yoga     1      0.100 370.31 -26.344
## - exercise 1      3.149 373.55 -26.265
## + age      1      0.024 370.38 -26.249
## + hiking   1      0.017 370.39 -26.240
## + shopping 1      0.000 370.41 -26.218
## - clubbing 1      3.293 373.70 -26.085
## - tv       1      3.389 373.79 -25.965
## - concerts 1      3.556 373.96 -25.756
## - gender   1      4.223 374.63 -24.923
## - go_out   6      12.575 382.98 -24.627
## + field_cd 16     21.573 348.83 -24.241
## + date     6      4.313 366.09 -21.687
## - career_c 16     36.967 407.37 -15.793
##
## Step: AIC=-29.58
## fun_o ~ gender + goal + go_out + career_c + sports + exercise +
##       dining + museums + art + gaming + clubbing + tv + concerts +
##       music
##
##           Df Sum of Sq    RSS    AIC
## - museums  1      0.740 371.66 -30.645
## - goal     5      7.433 378.35 -30.310
## <none>          370.92 -29.575
## - art      1      1.683 372.60 -29.462
## - sports   1      1.919 372.83 -29.166
## - dining   1      1.931 372.85 -29.151
## - gaming   1      2.518 373.43 -28.416
## + tvsports 1      0.510 370.41 -28.218
## - music    1      2.750 373.67 -28.126
## + theater  1      0.361 370.55 -28.030
## + reading  1      0.244 370.67 -27.882
## + match_es 1      0.217 370.70 -27.848
## + movies   1      0.188 370.73 -27.813
## + imprace  1      0.176 370.74 -27.797
## + yoga     1      0.153 370.76 -27.769
## - exercise 1      3.079 373.99 -27.715
## - clubbing 1      3.142 374.06 -27.636
## + age      1      0.018 370.90 -27.598
## + hiking   1      0.005 370.91 -27.581
## + shopping 1      0.001 370.92 -27.576

```

```

## - concerts 1      3.797 374.71 -26.819
## - gender 1       4.311 375.23 -26.178
## - tv 1          4.452 375.37 -26.003
## - go_out 6       12.805 383.72 -25.725
## + field_cd 16     21.574 349.34 -25.560
## + date 6         4.418 366.50 -23.172
## - career_c 16     36.740 407.66 -17.468
##
## Step: AIC=-30.64
## fun_o ~ gender + goal + go_out + career_c + sports + exercise +
##      dining + art + gaming + clubbing + tv + concerts + music
##
##           Df Sum of Sq    RSS    AIC
## - art      1      1.041 372.70 -31.338
## - goal     5      7.529 379.19 -31.278
## - dining   1      1.553 373.21 -30.697
## <none>                371.66 -30.645
## - sports   1      1.829 373.49 -30.351
## - gaming   1      2.383 374.04 -29.660
## + museums  1      0.740 370.92 -29.575
## + tvsports 1      0.580 371.08 -29.374
## + theater  1      0.573 371.08 -29.365
## - music    1      2.744 374.40 -29.209
## + movies   1      0.226 371.43 -28.928
## + imprace  1      0.224 371.43 -28.926
## + match_es 1      0.157 371.50 -28.841
## + yoga     1      0.142 371.51 -28.823
## - exercise 1      3.101 374.76 -28.764
## + reading  1      0.073 371.58 -28.736
## + age      1      0.024 371.63 -28.674
## + hiking   1      0.011 371.64 -28.659
## + shopping 1      0.000 371.66 -28.645
## - clubbing 1      3.239 374.89 -28.592
## - concerts 1      3.889 375.54 -27.783
## - gender   1      4.225 375.88 -27.366
## - tv       1      4.450 376.11 -27.086
## - go_out   6     13.017 384.67 -26.568
## + field_cd 16     21.298 350.36 -26.204
## + date     6      4.448 367.21 -24.267
## - career_c 16     37.259 408.91 -18.028
##
## Step: AIC=-31.34
## fun_o ~ gender + goal + go_out + career_c + sports + exercise +
##      dining + gaming + clubbing + tv + concerts + music
##
##           Df Sum of Sq    RSS    AIC
## - goal     5      7.144 379.84 -32.472
## <none>                372.70 -31.338
## - sports   1      1.702 374.40 -31.210
## + art      1      1.041 371.66 -30.644
## - dining   1      2.377 375.07 -30.370
## + tvsports 1      0.674 372.02 -30.184
## - gaming   1      2.542 375.24 -30.164
## + yoga     1      0.315 372.38 -29.733

```

```

## - music      1      2.906 375.60 -29.711
## + reading    1      0.233 372.46 -29.631
## + imprace    1      0.231 372.47 -29.628
## + match_es   1      0.111 372.59 -29.478
## + movies     1      0.107 372.59 -29.472
## + theater    1      0.103 372.59 -29.467
## + museums    1      0.098 372.60 -29.461
## - concerts   1      3.184 375.88 -29.365
## - exercise   1      3.188 375.89 -29.360
## + hiking     1      0.004 372.69 -29.343
## + age        1      0.002 372.70 -29.340
## + shopping   1      0.001 372.70 -29.339
## - clubbing   1      3.639 376.34 -28.800
## - go_out     6     12.685 385.38 -27.708
## - tv         1      4.548 377.25 -27.674
## - gender     1      4.839 377.54 -27.314
## + field_cd   16     21.321 351.38 -26.849
## + date       6      4.700 368.00 -25.265
## - career_c   16     36.643 409.34 -19.543
##
## Step: AIC=-32.47
## fun_o ~ gender + go_out + career_c + sports + exercise + dining +
##      gaming + clubbing + tv + concerts + music
##
##      Df Sum of Sq    RSS    AIC
## <none>                379.84 -32.472
## - exercise   1      2.206 382.05 -31.767
## - gaming      1      2.235 382.08 -31.732
## - sports      1      2.496 382.34 -31.413
## + goal        5      7.144 372.70 -31.338
## - dining      1      2.596 382.44 -31.291
## + art         1      0.655 379.19 -31.278
## + tvsports    1      0.581 379.26 -31.186
## + age         1      0.387 379.45 -30.948
## + reading     1      0.270 379.57 -30.804
## + yoga        1      0.259 379.58 -30.791
## + movies      1      0.177 379.66 -30.689
## + imprace     1      0.172 379.67 -30.684
## + match_es    1      0.168 379.67 -30.678
## + theater     1      0.084 379.76 -30.575
## + shopping    1      0.059 379.78 -30.544
## + museums     1      0.014 379.83 -30.489
## + hiking      1      0.006 379.83 -30.479
## - tv          1      3.598 383.44 -30.069
## - music       1      3.832 383.67 -29.783
## - concerts    1      4.839 384.68 -28.560
## - clubbing    1      5.184 385.02 -28.141
## - gender      1      5.671 385.51 -27.550
## + field_cd    16     21.239 358.60 -27.343
## + date        6      5.274 374.57 -27.001
## - go_out      6     15.387 395.23 -25.927
## - career_c    16     35.494 415.33 -22.753

```

```
step$anova # display results
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## fun_o ~ (iid + dec_o + attr_o + sinc_o + intel_o + amb_o + gender +
##   age + field_cd + imprace + goal + date + go_out + career_c +
##   match_es + sports + tvsports + exercise + dining + museums +
##   art + hiking + gaming + clubbing + reading + tv + theater +
##   movies + concerts + music + shopping + yoga) - iid - attr_o -
##   sinc_o - intel_o - dec_o - amb_o
##
## Final Model:
## fun_o ~ gender + go_out + career_c + sports + exercise + dining +
##   gaming + clubbing + tv + concerts + music
##
##
##           Step Df      Deviance Resid. Df Resid. Dev      AIC
## 1
## 2      - date  6 4.481620e+00      402   346.9414  -8.780028
## 3 - field_cd 16 2.215804e+01      418   369.0994 -11.867960
## 4 - shopping  1 2.824428e-04      419   369.0997 -13.867603
## 5   - hiking  1 2.400366e-02      420   369.1237 -15.837233
## 6     - age  1 4.665603e-02      421   369.1703 -17.778210
## 7   - movies  1 8.556078e-02      422   369.2559 -19.669988
## 8     - yoga  1 1.125064e-01      423   369.3684 -21.527722
## 9   - imprace 1 1.616828e-01      424   369.5301 -23.323348
## 10 - match_es 1 2.107973e-01      425   369.7409 -25.057026
## 11 - reading  1 2.535269e-01      426   369.9944 -26.736919
## 12 - theater  1 4.113948e-01      427   370.4058 -28.217953
## 13 - tvsports 1 5.100970e-01      428   370.9159 -29.575275
## 14 - museums  1 7.400379e-01      429   371.6559 -30.644462
## 15    - art   1 1.040968e+00      430   372.6969 -31.338274
## 16    - goal  5 7.143779e+00      435   379.8407 -32.471621
```

Reducing our model based on our anova results

```
# fun_o ~ gender + go_out + career_c + sports + exercise + dining + gaming + clubbing + tv + concerts +
fun_reduced_augmented_fit <- lm(fun_o ~ gender + go_out + career_c + sports + exercise + dining +
  gaming + clubbing + tv + concerts + music, data = na.omit(augmented_attris))
summary(fun_reduced_augmented_fit)
```

```
##
## Call:
## lm(formula = fun_o ~ gender + go_out + career_c + sports + exercise +
##   dining + gaming + clubbing + tv + concerts + music, data = na.omit(augmented_attris))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.9682 -0.5683  0.0323  0.6391  2.2438
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.03751    0.35955  16.792  < 2e-16 ***
```

```
## gender1      -0.27384    0.10745   -2.549    0.01116 *
## go_out2      -0.08254    0.10894   -0.758    0.44907
## go_out3      -0.31131    0.12094   -2.574    0.01038 *
## go_out4      -0.42787    0.21174   -2.021    0.04392 *
## go_out5      -0.11523    0.31432   -0.367    0.71410
## go_out6      -1.14682    0.39907   -2.874    0.00425 **
## go_out7      -1.49557    0.96791   -1.545    0.12304
## career_c2    -0.43093    0.17679   -2.437    0.01519 *
## career_c3      0.22806    0.27954    0.816    0.41503
## career_c4     -0.06810    0.24789   -0.275    0.78368
## career_c5     -0.26568    0.30406   -0.874    0.38272
## career_c6      0.09284    0.21343    0.435    0.66380
## career_c7     -0.08688    0.17644   -0.492    0.62269
## career_c8      1.49079    0.68507    2.176    0.03008 *
## career_c9      0.12300    0.24013    0.512    0.60875
## career_c10    -0.09314    0.25737   -0.362    0.71762
## career_c11    -0.17864    0.30928   -0.578    0.56382
## career_c12    -0.25396    0.68805   -0.369    0.71223
## career_c13    -0.89627    0.45025   -1.991    0.04715 *
## career_c14    -1.12556    0.95665   -1.177    0.24001
## career_c15    -1.43809    0.50100   -2.870    0.00430 **
## career_c16    -0.72572    0.68458   -1.060    0.28969
## career_c17     1.30057    0.96143    1.353    0.17684
## sports        0.03378    0.01998    1.691    0.09162 .
## exercise      0.03278    0.02062    1.590    0.11266
## dining        0.04592    0.02663    1.724    0.08536 .
## gaming        -0.03048    0.01905   -1.600    0.11037
## clubbing      0.04530    0.01859    2.437    0.01523 *
## tv            -0.03758    0.01852   -2.030    0.04298 *
## concerts     -0.06468    0.02748   -2.354    0.01901 *
## music         0.06710    0.03203    2.095    0.03675 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9345 on 435 degrees of freedom
## Multiple R-squared:  0.2118, Adjusted R-squared:  0.1557
## F-statistic: 3.772 on 31 and 435 DF,  p-value: 2.805e-10
```

The R^2 is only 0.1557.

Does the model come out any better if we go forwards?

```
# I don't want the forward approach to grab the perceived data
smaller_augmented_attribs <- augmented_attribs[, -which(names(augmented_attribs) %in% c("attr_o", "sinc
# since we have a bunch of predictor variables I'm going to run a stepwise AIC
fun_augmented_fit_2 <- lm(fun_o ~ 1, data = na.omit(smaller_augmented_attribs))
summary(fun_augmented_fit_2)
```

```
##
## Call:
## lm(formula = fun_o ~ 1, data = na.omit(smaller_augmented_attribs))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.7713 -0.6325  0.1037  0.7465  2.2190
##
```

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.39635    0.04706   135.9  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.017 on 466 degrees of freedom
step <- stepAIC(fun_augmented_fit_2, direction="forward", scope = list(upper = fun_augmented_fit, lower

## Start:  AIC=16.7
## fun_o ~ 1
##
##           Df Sum of Sq    RSS    AIC
## + go_out    6    30.650 451.29 -1.9830
## + exercise   1    13.010 468.93  5.9234
## + career_c  16    40.360 441.58  7.8585
## + clubbing   1    10.551 471.39  8.3653
## + dining     1     9.259 472.68  9.6439
## + field_cd  17    40.223 441.71 10.0034
## + gender     1     7.622 474.32 11.2588
## + goal       5    12.222 469.72 14.7077
## + date       6    14.191 467.75 14.7458
## + sports     1     4.022 477.92 14.7899
## + shopping   1     2.662 479.28 16.1166
## <none>                481.94 16.7032
## + music      1     1.971 479.97 16.7897
## + tv         1     1.809 480.13 16.9474
## + gaming     1     1.644 480.29 17.1071
## + yoga       1     1.625 480.31 17.1256
## + art        1     1.123 480.81 17.6141
## + age        1     1.086 480.85 17.6496
## + concerts   1     0.449 481.49 18.2676
## + museums    1     0.263 481.67 18.4486
## + match_es   1     0.129 481.81 18.5785
## + movies     1     0.126 481.81 18.5815
## + imprace    1     0.085 481.85 18.6210
## + tvsports   1     0.063 481.87 18.6421
## + reading    1     0.053 481.88 18.6521
## + theater    1     0.016 481.92 18.6881
## + hiking     1     0.000 481.94 18.7029
##
## Step:  AIC=-1.98
## fun_o ~ go_out
##
##           Df Sum of Sq    RSS    AIC
## + exercise   1     9.883 441.40 -10.3238
## + clubbing    1     8.755 442.53  -9.1318
## + gender      1     8.216 443.07  -8.5634
## + career_c   16    33.825 417.46  -6.3666
## + dining     1     5.489 445.80  -5.6980
## + shopping    1     2.526 448.76  -2.6043
## + sports     1     2.217 449.07  -2.2830
## + music      1     2.000 449.29  -2.0576
## <none>                451.29  -1.9830
```

```

## + art      1      1.916 449.37 -1.9700
## + gaming   1      1.328 449.96 -1.3592
## + field_cd 17     31.085 420.20 -1.3124
## + yoga     1      1.118 450.17 -1.1417
## + museums  1      0.695 450.59 -0.7031
## + tv       1      0.686 450.60 -0.6932
## + age      1      0.377 450.91 -0.3734
## + concerts 1      0.327 450.96 -0.3212
## + reading  1      0.239 451.05 -0.2306
## + goal     5      7.885 443.40 -0.2150
## + match_es 1      0.190 451.10 -0.1800
## + imprace  1      0.170 451.12 -0.1588
## + theater  1      0.115 451.17 -0.1015
## + tvsports 1      0.022 451.27 -0.0059
## + movies   1      0.003 451.28  0.0134
## + hiking   1      0.003 451.28  0.0141
## + date     6      6.417 444.87  3.3287
##
## Step:  AIC=-10.32
## fun_o ~ go_out + exercise
##
##           Df Sum of Sq   RSS     AIC
## + clubbing 1      8.157 433.25 -17.0347
## + gender   1      7.780 433.62 -16.6288
## + dining   1      4.458 436.95 -13.0644
## + career_c 16     31.333 410.07 -12.7096
## + field_cd 17     32.717 408.69 -12.2874
## + art      1      2.087 439.32 -10.5369
## + music    1      2.056 439.35 -10.5043
## + shopping 1      1.947 439.46 -10.3885
## <none>          441.40 -10.3238
## + gaming   1      1.757 439.65 -10.1862
## + goal     5      8.907 432.50 -9.8438
## + tv       1      1.064 440.34 -9.4509
## + tvsports 1      0.823 440.58 -9.1954
## + museums  1      0.724 440.68 -9.0907
## + yoga     1      0.473 440.93 -8.8250
## + imprace  1      0.424 440.98 -8.7725
## + age      1      0.219 441.19 -8.5554
## + reading  1      0.204 441.20 -8.5398
## + theater  1      0.137 441.27 -8.4684
## + concerts 1      0.116 441.29 -8.4461
## + match_es 1      0.104 441.30 -8.4334
## + hiking   1      0.100 441.30 -8.4293
## + sports   1      0.018 441.39 -8.3432
## + movies   1      0.017 441.39 -8.3419
## + date     6      5.599 435.81 -4.2849
##
## Step:  AIC=-17.03
## fun_o ~ go_out + exercise + clubbing
##
##           Df Sum of Sq   RSS     AIC
## + gender   1      6.394 426.85 -21.978
## + field_cd 17     32.843 400.40 -19.851

```

```

## + career_c 16      30.300 402.95 -18.893
## + gaming      1       3.354 429.89 -18.664
## + dining      1       2.401 430.85 -17.630
## <none>                433.25 -17.035
## + music      1       1.330 431.92 -16.470
## + tvsports   1       1.283 431.96 -16.420
## + tv         1       1.225 432.02 -16.357
## + art        1       1.150 432.10 -16.275
## + shopping   1       0.725 432.52 -15.817
## + concerts   1       0.475 432.77 -15.547
## + reading    1       0.421 432.83 -15.489
## + museums    1       0.300 432.95 -15.358
## + yoga       1       0.259 432.99 -15.314
## + imprace    1       0.250 433.00 -15.305
## + age        1       0.055 433.19 -15.094
## + hiking     1       0.039 433.21 -15.076
## + movies     1       0.034 433.21 -15.072
## + sports     1       0.013 433.23 -15.048
## + theater    1       0.004 433.24 -15.039
## + match_es   1       0.003 433.24 -15.039
## + goal       5       6.863 426.38 -14.492
## + date       6       5.552 427.69 -11.058
##
## Step:  AIC=-21.98
## fun_o ~ go_out + exercise + clubbing + gender
##
##           Df Sum of Sq   RSS   AIC
## + field_cd 17      32.828 394.03 -25.350
## + career_c 16      30.268 396.59 -24.326
## + tv       1       2.164 424.69 -22.352
## <none>                426.85 -21.978
## + gaming   1       1.493 425.36 -21.614
## + concerts 1       1.240 425.61 -21.336
## + dining   1       1.092 425.76 -21.174
## + theater  1       0.931 425.92 -20.998
## + sports   1       0.930 425.92 -20.996
## + music    1       0.851 426.00 -20.910
## + imprace  1       0.848 426.01 -20.906
## + tvsports 1       0.540 426.31 -20.570
## + movies   1       0.351 426.50 -20.362
## + art      1       0.187 426.67 -20.183
## + hiking   1       0.165 426.69 -20.159
## + reading  1       0.059 426.79 -20.043
## + age      1       0.024 426.83 -20.004
## + match_es 1       0.022 426.83 -20.003
## + museums  1       0.013 426.84 -19.992
## + yoga     1       0.009 426.84 -19.988
## + shopping 1       0.005 426.85 -19.983
## + goal     5       6.941 419.91 -19.634
## + date     6       6.819 420.03 -17.499
##
## Step:  AIC=-25.35
## fun_o ~ go_out + exercise + clubbing + gender + field_cd
##

```



```

##          Df Sum of Sq    RSS    AIC
## + tv      1     3.7006 390.32 -27.756
## <none>                394.03 -25.350
## + sports   1     1.5450 392.48 -25.184
## + concerts 1     1.2483 392.78 -24.831
## + gaming   1     1.0945 392.93 -24.649
## + movies   1     1.0591 392.97 -24.607
## + dining   1     0.9887 393.04 -24.523
## + imprace  1     0.9645 393.06 -24.494
## + theater  1     0.8761 393.15 -24.389
## + music    1     0.7567 393.27 -24.247
## + tvsports 1     0.5591 393.47 -24.013
## + age      1     0.3676 393.66 -23.786
## + art       1     0.3659 393.66 -23.784
## + goal      5     6.9210 387.10 -23.625
## + shopping 1     0.1296 393.90 -23.503
## + match_es 1     0.0776 393.95 -23.442
## + hiking    1     0.0679 393.96 -23.430
## + reading   1     0.0064 394.02 -23.357
## + museums   1     0.0037 394.02 -23.354
## + yoga      1     0.0033 394.02 -23.354
## + date      6     6.0122 388.01 -20.530
## + career_c 15    19.8101 374.22 -19.439
##
## Step:  AIC=-27.76
## fun_o ~ go_out + exercise + clubbing + gender + field_cd + tv
##
##          Df Sum of Sq    RSS    AIC
## <none>                390.32 -27.756
## + dining   1     1.4977 388.83 -27.552
## + sports   1     1.3986 388.93 -27.433
## + goal      5     7.8956 382.43 -27.300
## + music    1     1.1131 389.21 -27.090
## + concerts 1     0.9462 389.38 -26.890
## + age      1     0.7940 389.53 -26.707
## + art       1     0.4093 389.92 -26.246
## + theater  1     0.3468 389.98 -26.172
## + gaming    1     0.3416 389.98 -26.165
## + imprace  1     0.1835 390.14 -25.976
## + shopping 1     0.1833 390.14 -25.976
## + movies    1     0.1462 390.18 -25.931
## + tvsports 1     0.0563 390.27 -25.824
## + match_es 1     0.0487 390.28 -25.815
## + yoga      1     0.0210 390.30 -25.781
## + reading   1     0.0105 390.31 -25.769
## + museums   1     0.0091 390.32 -25.767
## + hiking    1     0.0004 390.32 -25.757
## + career_c 15    21.0488 369.28 -23.644
## + date      6     6.1968 384.13 -23.230

```

```
step$anova # display results
```

```

## Stepwise Model Path
## Analysis of Deviance Table
##

```

```
## Initial Model:
## fun_o ~ 1
##
## Final Model:
## fun_o ~ go_out + exercise + clubbing + gender + field_cd + tv
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1              6  30.649718      466   481.9371  16.703198
## 2    + go_out   6  30.649718      460   451.2874  -1.983026
## 3  + exercise   1   9.883010      459   441.4044 -10.323783
## 4  + clubbing   1   8.157185      458   433.2472 -17.034716
## 5    + gender   1   6.393789      457   426.8534 -21.977984
## 6  + field_cd  17  32.827927      440   394.0255 -25.349691
## 7      + tv     1   3.700564      439   390.3249 -27.756334
```

Reducing our model based on our anova results

```
# Forward AIC = fun_o ~ go_out + exercise + clubbing + gender + field_cd + tv
fun_reduced_augmented_fit_2 <- lm(fun_o ~ go_out + exercise + clubbing + gender + field_cd + tv, data =
summary(fun_reduced_augmented_fit_2))
```

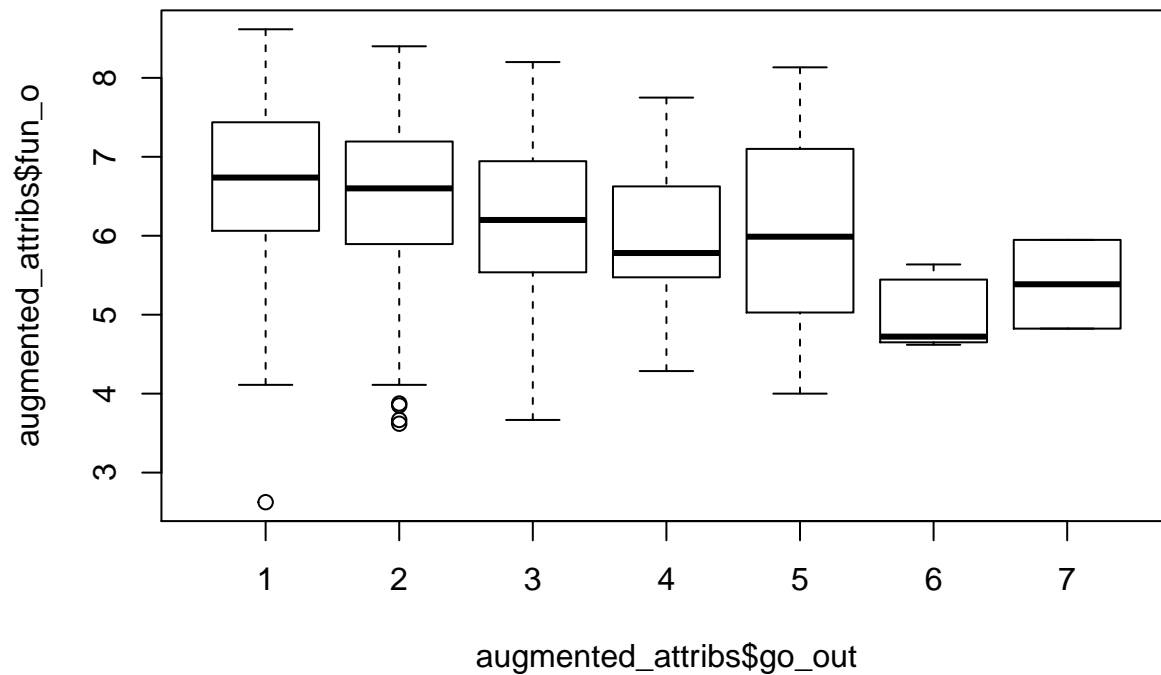
```
##
## Call:
## lm(formula = fun_o ~ go_out + exercise + clubbing + gender +
##      field_cd + tv, data = na.omit(smaller_augmented_attris))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.1794 -0.5175  0.0286  0.6445  2.3316
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   6.436726   0.264105   24.372 < 2e-16 ***
## go_out2       -0.092142   0.109721   -0.840 0.401488
## go_out3       -0.310933   0.121885   -2.551 0.011079 *
## go_out4       -0.471626   0.210271   -2.243 0.025398 *
## go_out5       -0.408490   0.317929   -1.285 0.199522
## go_out6       -1.156091   0.403127   -2.868 0.004332 **
## go_out7       -1.380294   0.961100   -1.436 0.151669
## exercise       0.065153   0.018255    3.569 0.000398 ***
## clubbing       0.047721   0.018263    2.613 0.009284 **
## gender1       -0.307308   0.106095   -2.897 0.003962 **
## field_cd2      -0.670767   0.326160   -2.057 0.040319 *
## field_cd3      -0.218843   0.216248   -1.012 0.312095
## field_cd4       0.563170   0.356395    1.580 0.114784
## field_cd5      -0.563166   0.207212   -2.718 0.006831 **
## field_cd6      -0.173453   0.279286   -0.621 0.534882
## field_cd7      -0.280855   0.308084   -0.912 0.362471
## field_cd8      -0.009421   0.182912   -0.052 0.958947
## field_cd9      -0.553539   0.227453   -2.434 0.015346 *
## field_cd10     -0.452911   0.208655   -2.171 0.030495 *
## field_cd11     -0.351553   0.243975   -1.441 0.150315
## field_cd12      0.198140   0.959094    0.207 0.836425
## field_cd13      0.098945   0.222517    0.445 0.656782
```

```
## field_cd14  0.078275  0.391888  0.200 0.841778
## field_cd15  0.047481  0.353030  0.134 0.893072
## field_cd16 -0.289022  0.569382 -0.508 0.611984
## field_cd17  1.200144  0.962667  1.247 0.213177
## field_cd18 -0.741276  0.573506 -1.293 0.196852
## tv          -0.036438  0.017861 -2.040 0.041938 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9429 on 439 degrees of freedom
## Multiple R-squared:  0.1901, Adjusted R-squared:  0.1403
## F-statistic: 3.816 on 27 and 439 DF,  p-value: 1.599e-09
```

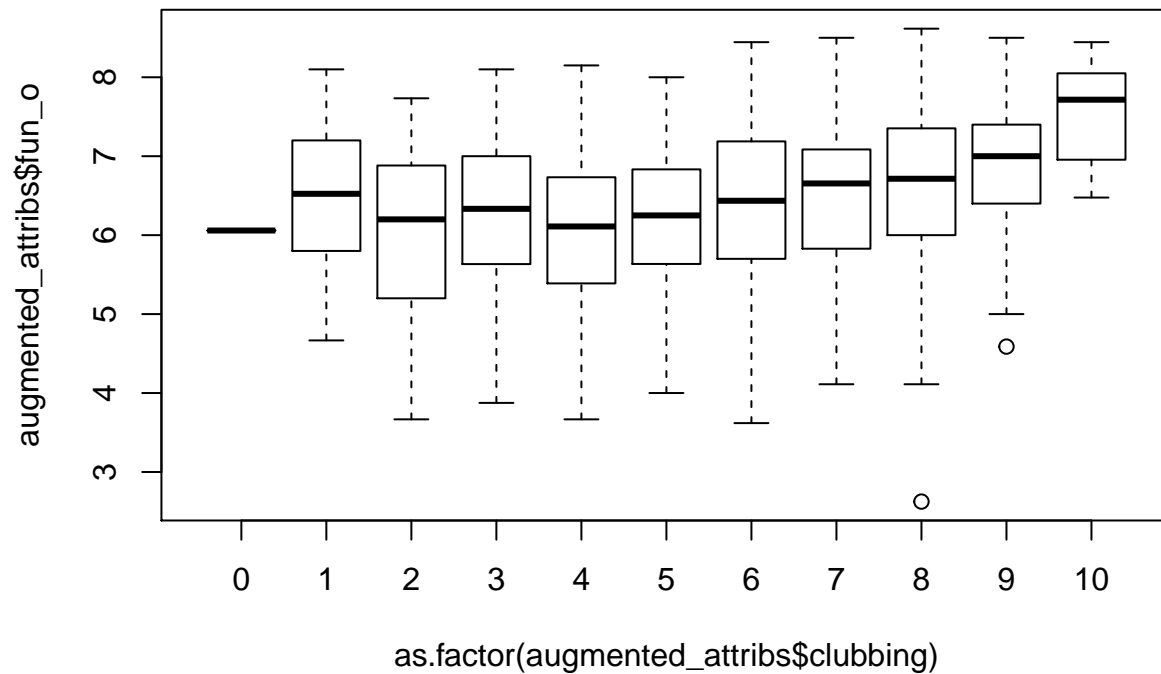
The adjusted R^2 is quite low, at 0.1403 with 27 variables.

Although these appear statistically significant, the Adj. R. Square for this model is very low and there is not much linear relationship from the predicted values

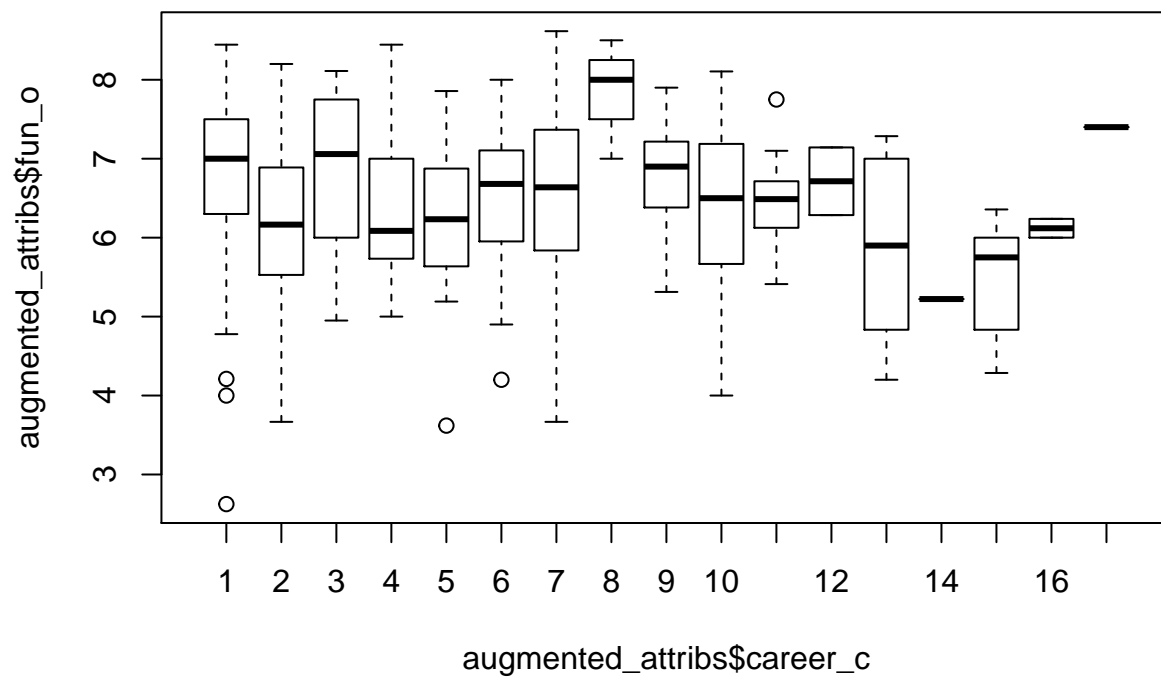
```
plot(augmented_attribs$fun_o ~ augmented_attribs$go_out)
```



```
plot(augmented_attribs$fun_o ~ as.factor(augmented_attribs$clubbing))
```



```
plot(augmented_attrib$fun_o ~ augmented_attrib$career_c)
```



Where should we go from here with this idea? - real estate agents are more fun? (but they showed up as statistically significant with a negative coefficient in the decision model above!) - visually, it is possible that the less you go out (7) the less fun you are - it could be possible that the more important clubbing is to you, the more fun you are

How accurate is our self-perception?

We see that personality traits like fun and attractiveness are correlated with if your speed date decides to say “yes” to you. What does this mean for us as far as what we think about ourselves? Do we even know how fun/attractive we are to others?

Alternatively, are others able to pick up on our less obvious traits, like sincerity and ambition, in 4 minutes?

plot self ratings vs. avg of trait

```
# rate that people decide they like you after the speed date
desirableness <- aggregate(dec_o ~ iid, full_data, mean)

#additional explanatory variables
attractive <- aggregate(attr_o ~ iid, full_data, mean)
sincere <- aggregate(sinc_o ~ iid, full_data, mean)
intelligent <- aggregate(intel_o ~ iid, full_data, mean)
fun <- aggregate(fun_o ~ iid, full_data, mean)
ambitious <- aggregate(amb_o ~ iid, full_data, mean)

personality_attribs <- full_data[, c("iid", "attr3_1", "sinc3_1", "fun3_1", "intel3_1", "amb3_1")]

personality_attribs <- personality_attribs[!duplicated(personality_attribs),]

# binding back together
personality_attribs <- merge(desirableness, personality_attribs, by="iid")

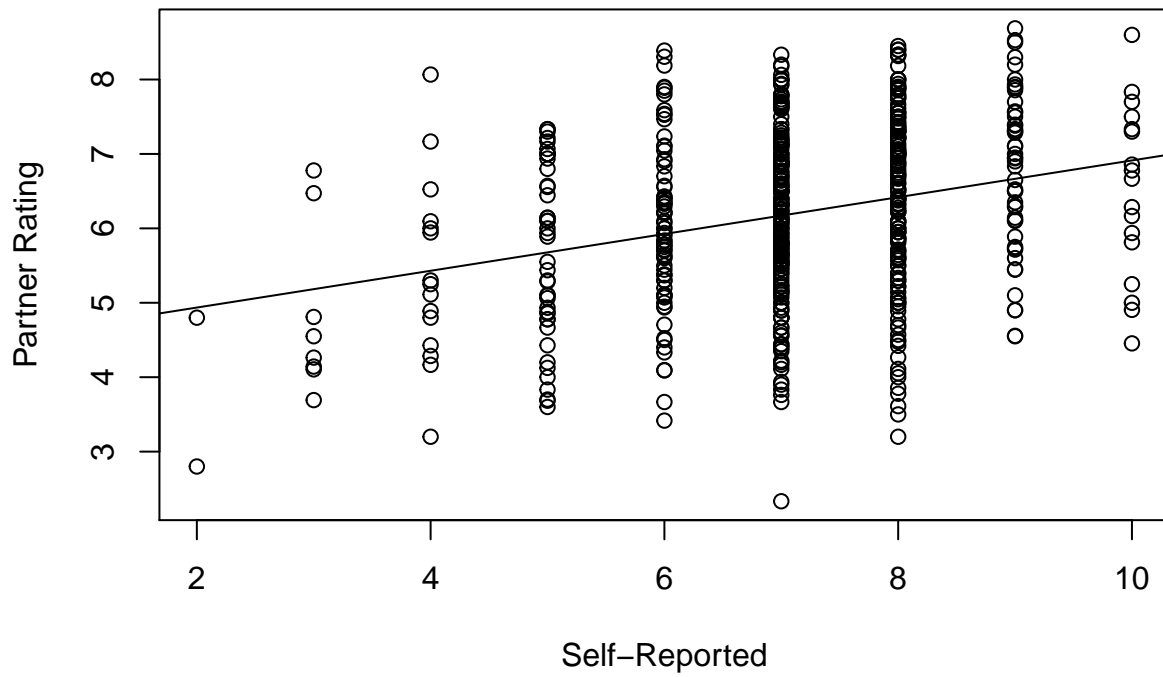
# we end up with three columns of dec_o here
personality_attribs <- Reduce(function(x, y) merge(x, y, all=TRUE), list(attractive, sincere, intelligent))

attr_lm <- lm(personality_attribs$attr_o ~ personality_attribs$attr3_1)
summary(attr_lm)

##
## Call:
## lm(formula = personality_attribs$attr_o ~ personality_attribs$attr3_1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.8381 -0.7651  0.0491  0.8329  2.6366
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.44154    0.25427  17.468 < 2e-16 ***
## personality_attribs$attr3_1  0.24712    0.03518   7.024 6.52e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.138 on 540 degrees of freedom
## (9 observations deleted due to missingness)
## Multiple R-squared:  0.08371, Adjusted R-squared:  0.08202
## F-statistic: 49.33 on 1 and 540 DF, p-value: 6.521e-12

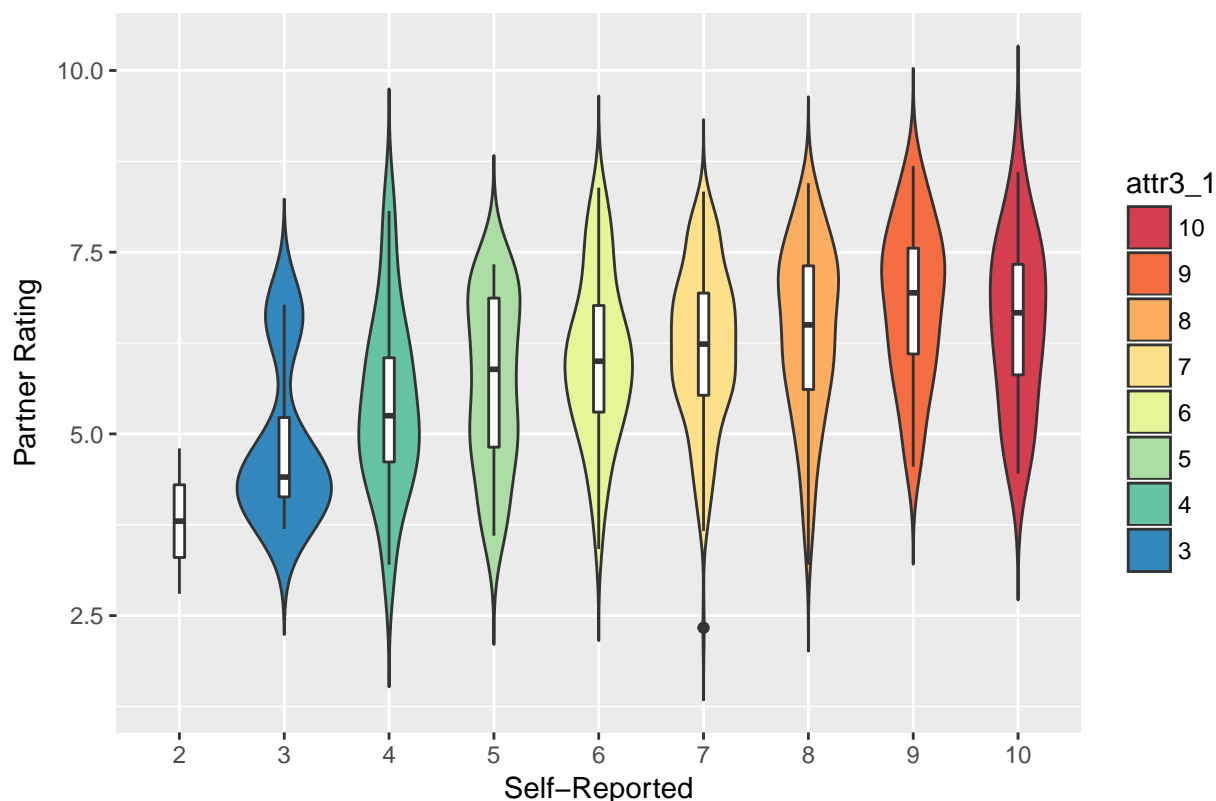
plot(personality_attribs$attr_o ~ personality_attribs$attr3_1, xlab="Self-Reported", ylab = "Partner Rating",
abline(attr_lm))
```

Perceived Attractiveness



```
library(ggplot2)
personality_attribs$attr3_1 <- as.factor(personality_attribs$attr3_1)
p <- ggplot(na.omit(personality_attribs), aes(x=attr3_1, y=attr_o, fill=attr3_1)) +
  geom_violin(trim=FALSE)
# plot with median and quartile
p + geom_boxplot(width=0.1, fill="white") + labs(title="Perceived Attractiveness", x="Self-Reported", y="Partner Rating")
```

Perceived Attractiveness



is it possible to add the lm to this plot?

```
sinc_lm <- lm(personality_attribs$sinc_o ~ as.numeric(personality_attribs$sinc3_1))
intel_lm <- lm(personality_attribs$intel_o ~ as.numeric(personality_attribs$intel3_1))
fun_lm <- lm(personality_attribs$fun_o ~ as.numeric(personality_attribs$fun3_1))
amb_lm <- lm(personality_attribs$amb_o ~ as.numeric(personality_attribs$amb3_1))
```

```
summary(sinc_lm)
```

```
##
## Call:
## lm(formula = personality_attribs$sinc_o ~ as.numeric(personality_attribs$sinc3_1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.05351 -0.41264  0.07983  0.49891  1.77983
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)      7.265203   0.177344  40.967
## as.numeric(personality_attribs$sinc3_1) -0.005629   0.021099  -0.267
##              Pr(>|t|)
## (Intercept)      <2e-16 ***
## as.numeric(personality_attribs$sinc3_1)      0.79
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.693 on 540 degrees of freedom
## (9 observations deleted due to missingness)
## Multiple R-squared: 0.0001318, Adjusted R-squared: -0.00172
## F-statistic: 0.07117 on 1 and 540 DF, p-value: 0.7897
```

```
summary(intel_lm)
```

```
##
## Call:
## lm(formula = personality_attri$intel_o ~ as.numeric(personality_attri$intel3_1))
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-2.57062	-0.44042	0.04232	0.45906	1.70438

```
##
## Coefficients:
```

	Estimate	Std. Error	t value
(Intercept)	7.14878	0.20778	34.406
as.numeric(personality_attri\$intel3_1)	0.02968	0.02457	1.208

```
## Pr(>|t|)
## (Intercept) <2e-16 ***
## as.numeric(personality_attri$intel3_1) 0.228
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6228 on 540 degrees of freedom
## (9 observations deleted due to missingness)
## Multiple R-squared: 0.002695, Adjusted R-squared: 0.0008486
## F-statistic: 1.459 on 1 and 540 DF, p-value: 0.2275
```

```
summary(fun_lm)
```

```
##
## Call:
## lm(formula = personality_attri$fun_o ~ as.numeric(personality_attri$fun3_1))
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-4.2271	-0.6367	0.0246	0.7546	2.1291

```
##
## Coefficients:
```

	Estimate	Std. Error	t value
(Intercept)	5.02305	0.21200	23.693
as.numeric(personality_attri\$fun3_1)	0.18290	0.02699	6.777

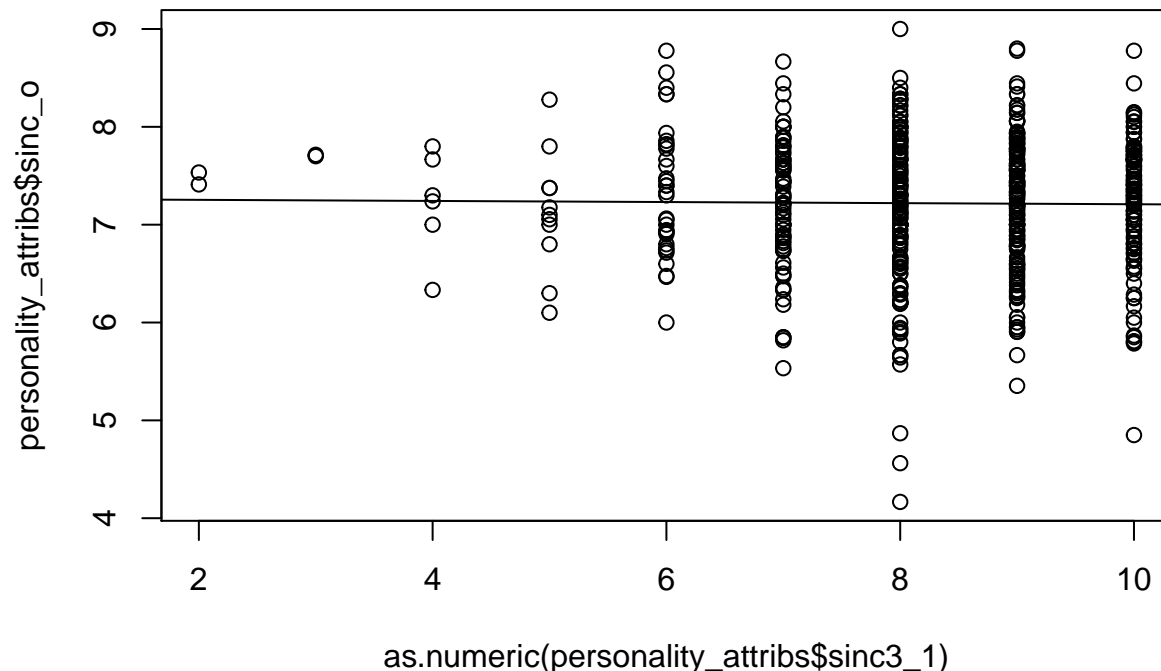
```
## Pr(>|t|)
## (Intercept) < 2e-16 ***
## as.numeric(personality_attri$fun3_1) 3.23e-11 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9715 on 540 degrees of freedom
## (9 observations deleted due to missingness)
## Multiple R-squared: 0.07837, Adjusted R-squared: 0.07667
## F-statistic: 45.92 on 1 and 540 DF, p-value: 3.23e-11
```



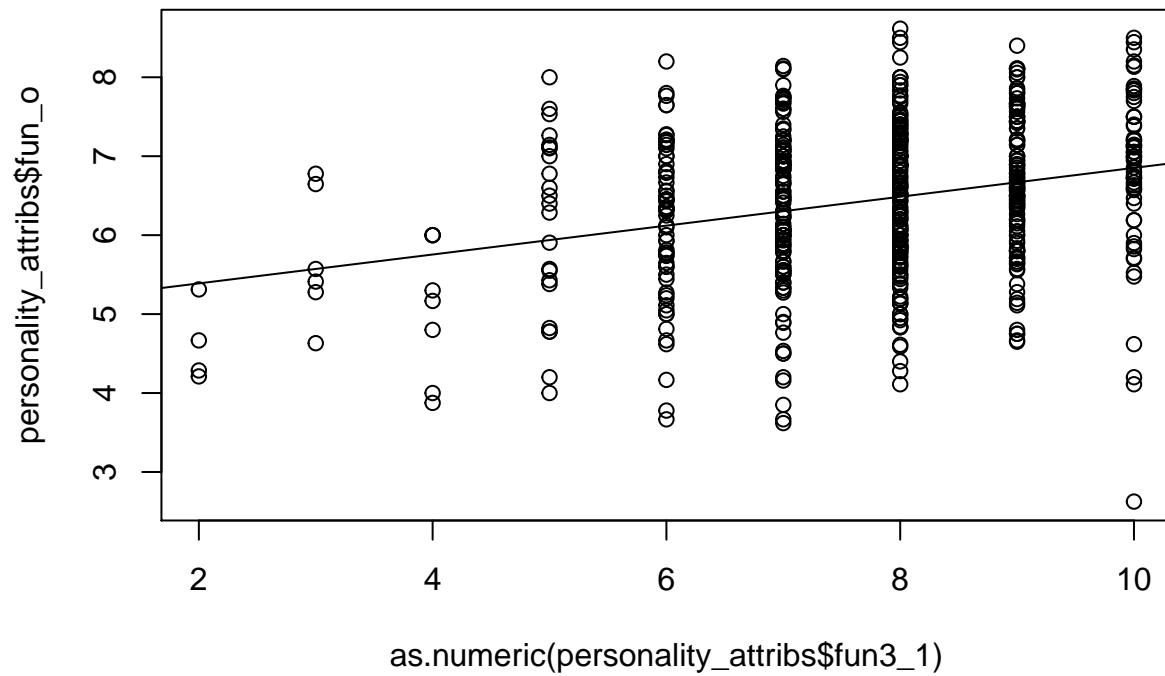
```
summary(amb_lm)
```

```
##
## Call:
## lm(formula = personality_attribs$amb_o ~ as.numeric(personality_attribs$amb3_1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.84567 -0.48655 -0.01689  0.50715  1.90091
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)      6.30664    0.14184  44.464
## as.numeric(personality_attribs$amb3_1)  0.06781    0.01822   3.722
##              Pr(>|t|)
## (Intercept)      < 2e-16 ***
## as.numeric(personality_attribs$amb3_1) 0.000219 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.757 on 540 degrees of freedom
## (9 observations deleted due to missingness)
## Multiple R-squared:  0.02501,    Adjusted R-squared:  0.0232
## F-statistic: 13.85 on 1 and 540 DF,  p-value: 0.0002187
```

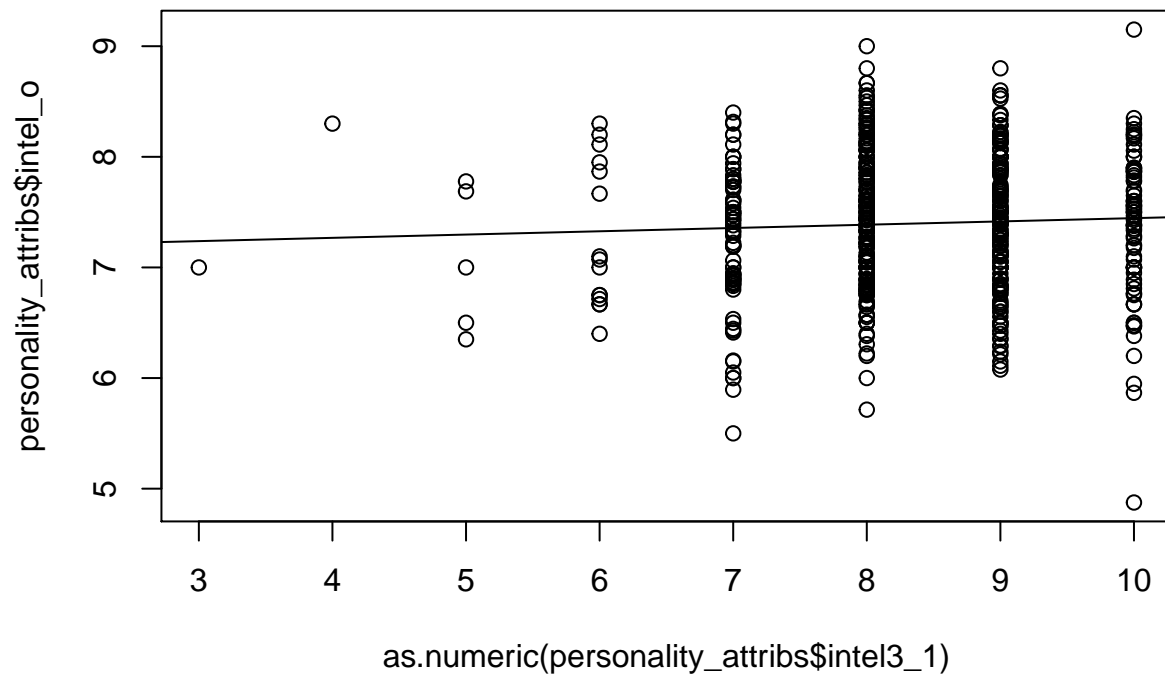
```
plot(personality_attribs$sinc_o ~ as.numeric(personality_attribs$sinc3_1))
abline(sinc_lm)
```



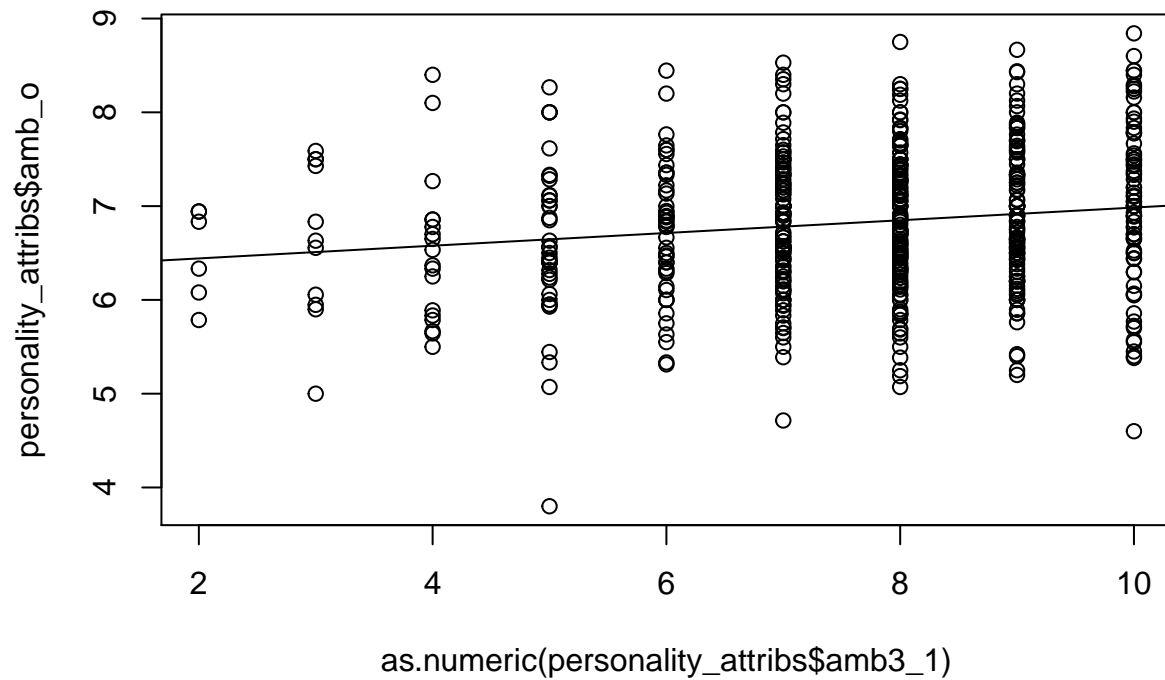
```
plot(personality_attribs$fun_o ~ as.numeric(personality_attribs$fun3_1))
abline(fun_lm)
```



```
plot(personality_attribs$intel_o ~ as.numeric(personality_attribs$intel3_1))
abline(intel_lm)
```



```
plot(personality_attribs$amb_o ~ as.numeric(personality_attribs$amb3_1))
abline(amb_lm)
```



```

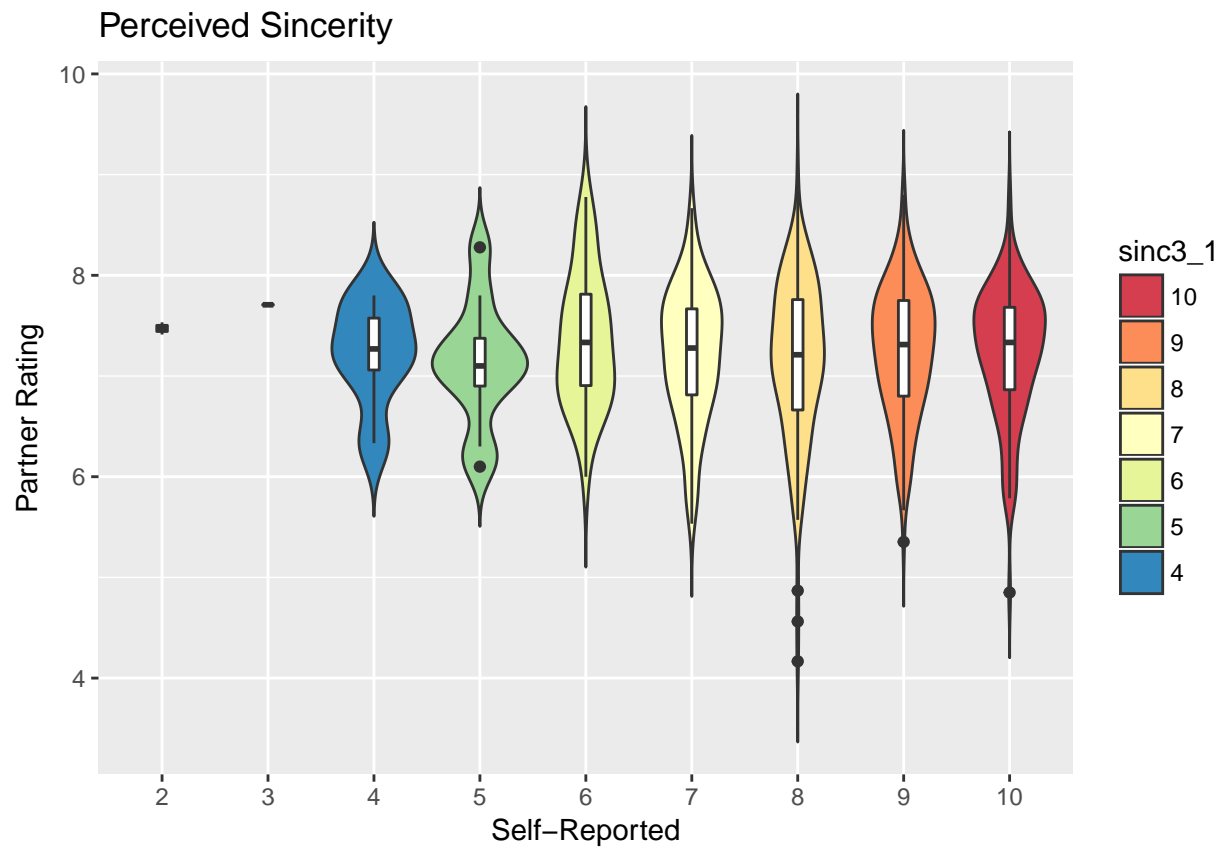
personality_attribs$sinc3_1 <- as.factor(personality_attribs$sinc3_1)
personality_attribs$fun3_1 <- as.factor(personality_attribs$fun3_1)
personality_attribs$intel3_1 <- as.factor(personality_attribs$intel3_1)
personality_attribs$amb3_1 <- as.factor(personality_attribs$amb3_1)

```

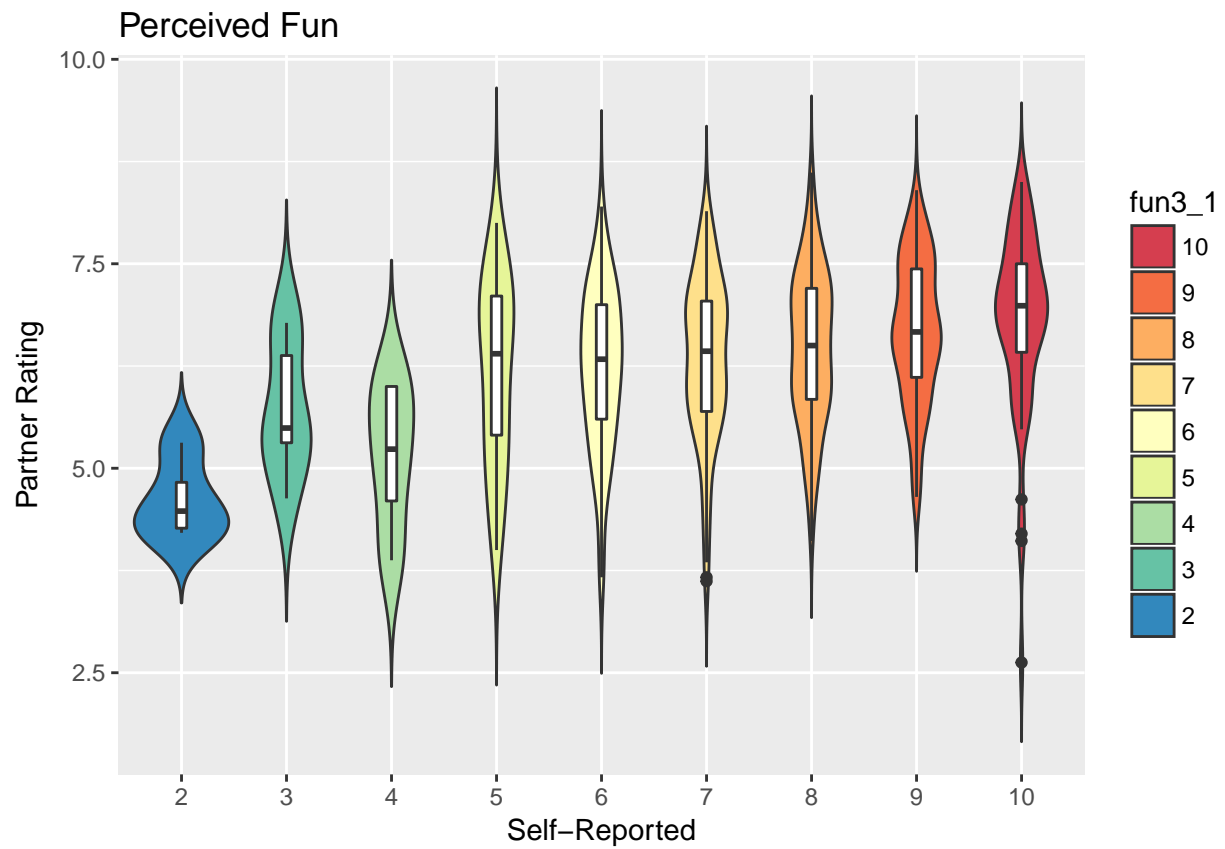
```

ggplot(na.omit(personality_attribs), aes(x=sinc3_1, y=sinc_o, fill=sinc3_1)) +
  geom_violin(trim=FALSE) + geom_boxplot(width=0.1, fill="white") + labs(title="Perceived Sincerity", x=

```



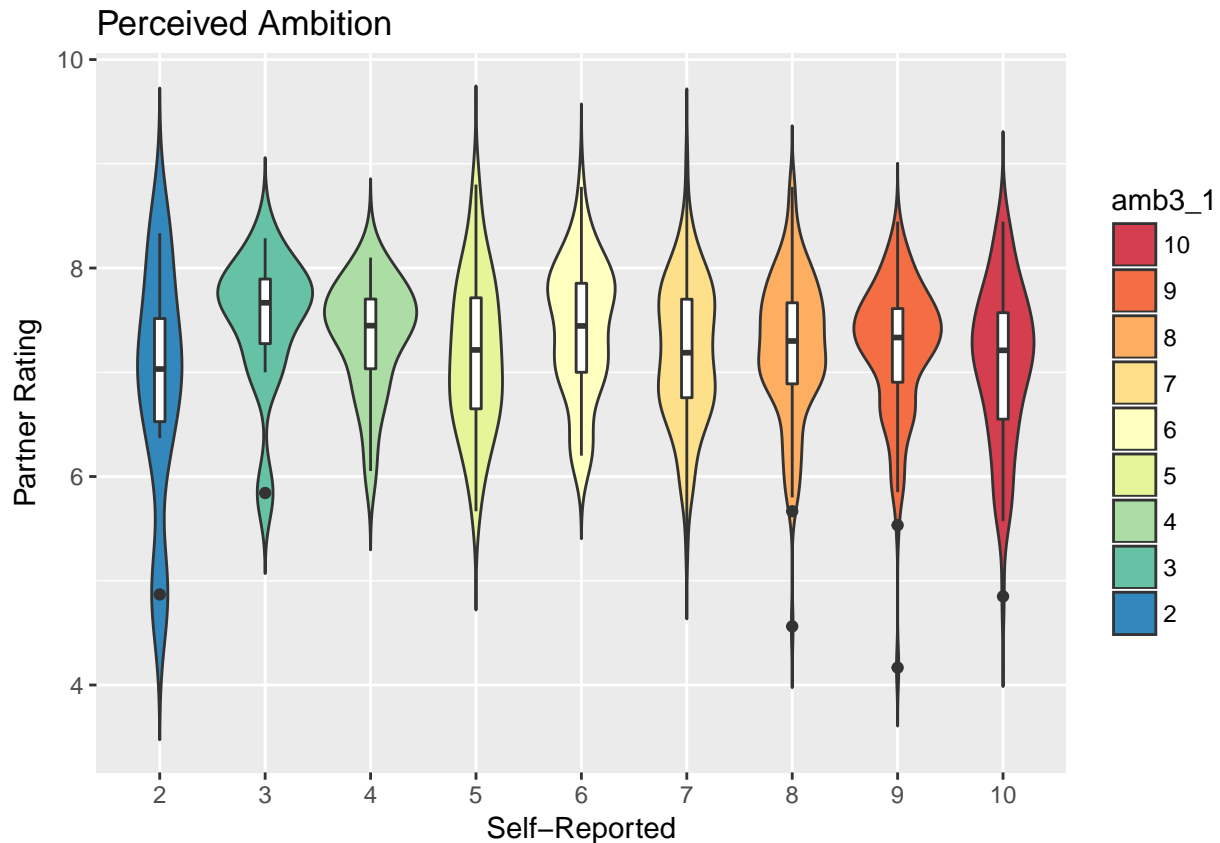
```
ggplot(na.omit(personality_attris), aes(x=fun3_1, y=fun_o, fill=fun3_1)) +  
  geom_violin(trim=FALSE) + geom_boxplot(width=0.1, fill="white") + labs(title="Perceived Fun", x="Self
```



```
ggplot(na.omit(personality_attris), aes(x=intel3_1, y=intel_o, fill=intel3_1)) +
  geom_violin(trim=FALSE) + geom_boxplot(width=0.1, fill="white") + labs(title="Perceived Intelligence")
```



```
ggplot(na.omit(personality_attris), aes(x=amb3_1, y=sinc_o, fill=amb3_1)) +
  geom_violin(trim=FALSE) + geom_boxplot(width=0.1, fill="white") + labs(title="Perceived Ambition", x=
```



```
#library(tidyverse)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:MASS':
##
##   select

## The following objects are masked from 'package:data.table':
##
##   between, first, last

## The following objects are masked from 'package:stats':
##
##   filter, lag

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

sincerity <- full_data[,c("sinc_o", "sinc3_1")]
sincerity <- sincerity %>% count(sinc3_1, sinc_o)
# round very few decimal ratings to integers
sincerity$sinc_o <- as.integer(sincerity$sinc_o)
sincerity$sinc3_1 <- as.factor(sincerity$sinc3_1)
sincerity$sinc_o <- as.factor(sincerity$sinc_o)

fun <- full_data[,c("fun_o", "fun3_1")]
```

```

fun <- fun %>% count(fun3_1, fun_o)
fun$fun3_1 <- as.factor(fun$fun3_1)
fun$fun_o <- as.factor(as.integer(fun$fun_o))

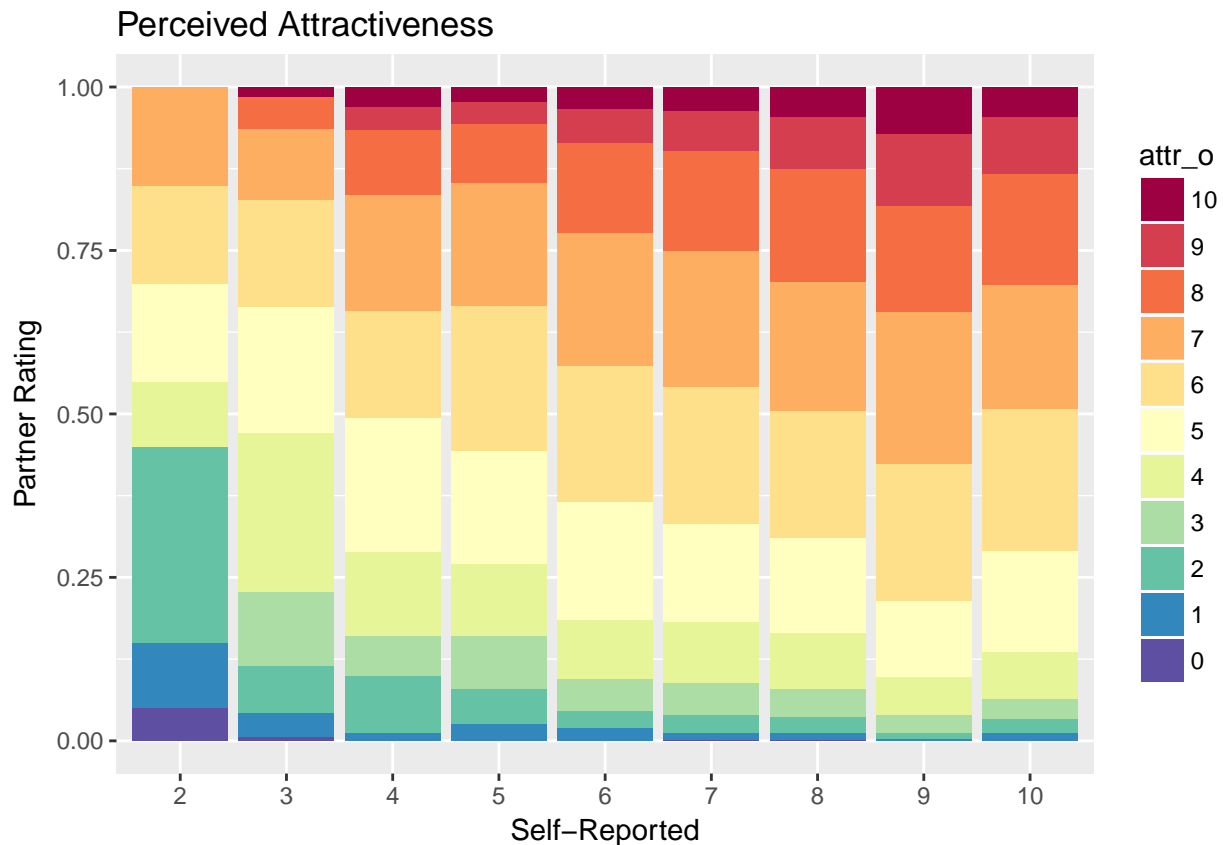
ambition <- full_data[,c("amb_o", "amb3_1")]
ambition <- ambition %>% count(amb3_1, amb_o)
ambition$amb3_1 <- as.factor(ambition$amb3_1)
ambition$amb_o <- as.factor(as.integer(ambition$amb_o))

intelligence <- full_data[,c("intel_o", "intel3_1")]
intelligence <- intelligence %>% count(intel3_1, intel_o)
intelligence$intel3_1 <- as.factor(intelligence$intel3_1)
intelligence$intel_o <- as.factor(as.integer(intelligence$intel_o))

hotness <- full_data[,c("attr_o", "attr3_1")]
hotness <- hotness %>% count(attr3_1, attr_o)
hotness$attr3_1 <- as.factor(hotness$attr3_1)
hotness$attr_o <- as.factor(as.integer(hotness$attr_o))

ggplot(data = na.omit(hotness), aes(x=attr3_1, y=n, fill=attr_o)) +
  geom_bar(position=position_fill(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")

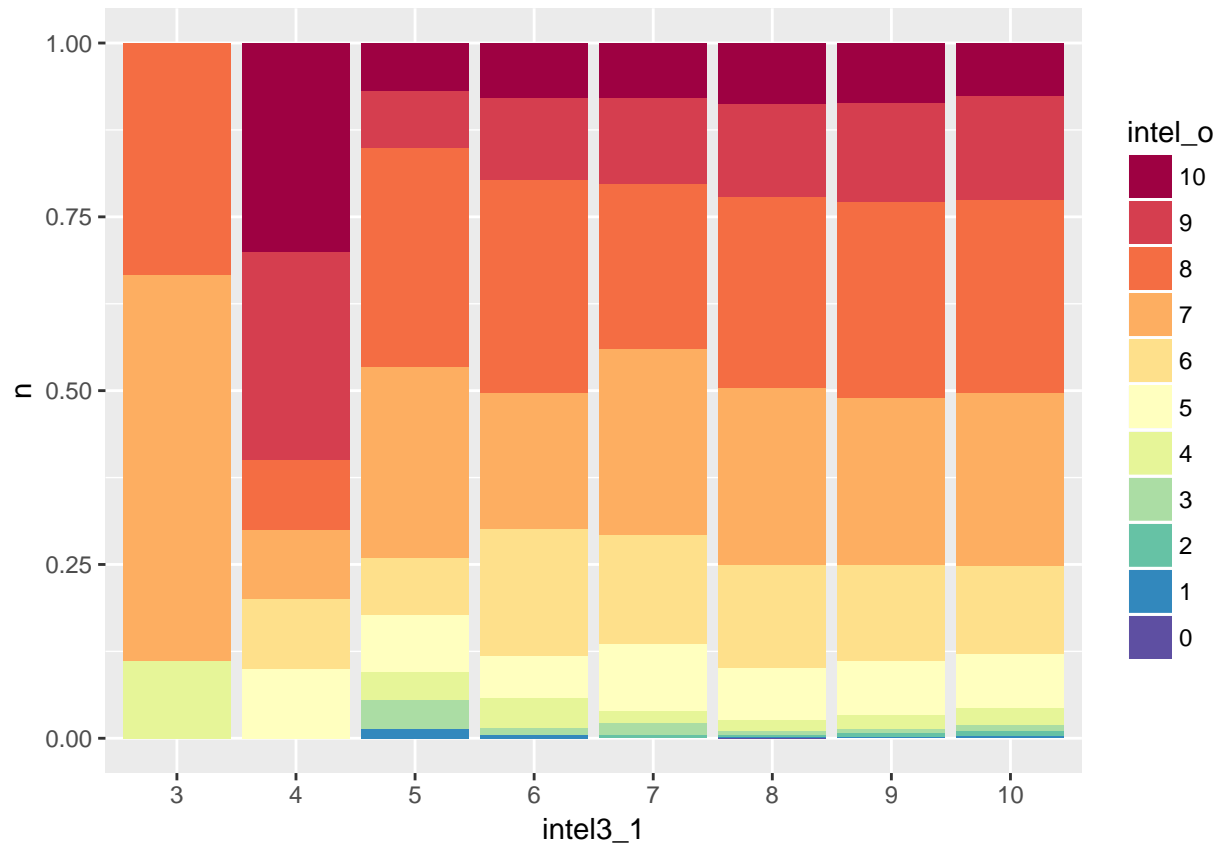
```



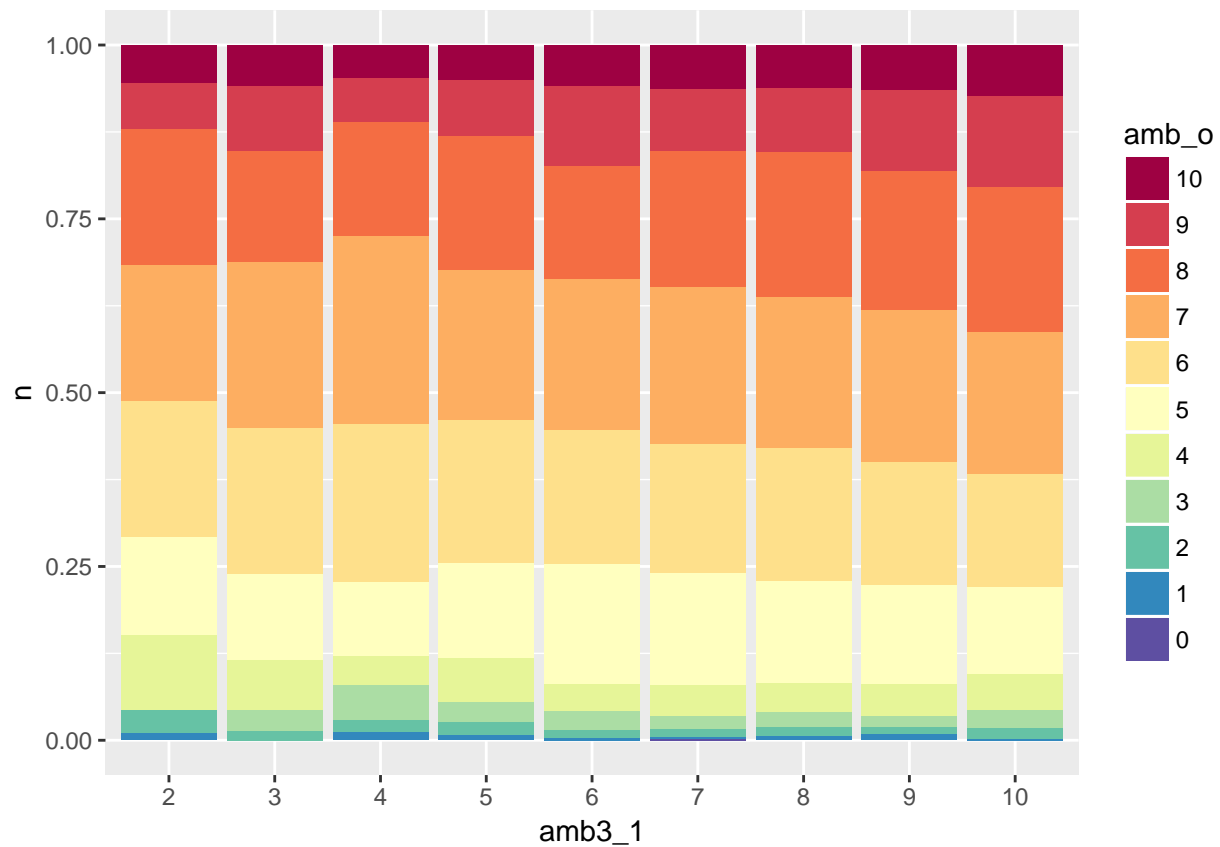
```

ggplot(data = na.omit(intelligence), aes(x=intel3_1, y=n, fill=intel_o)) +
  geom_bar(position=position_fill(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")

```

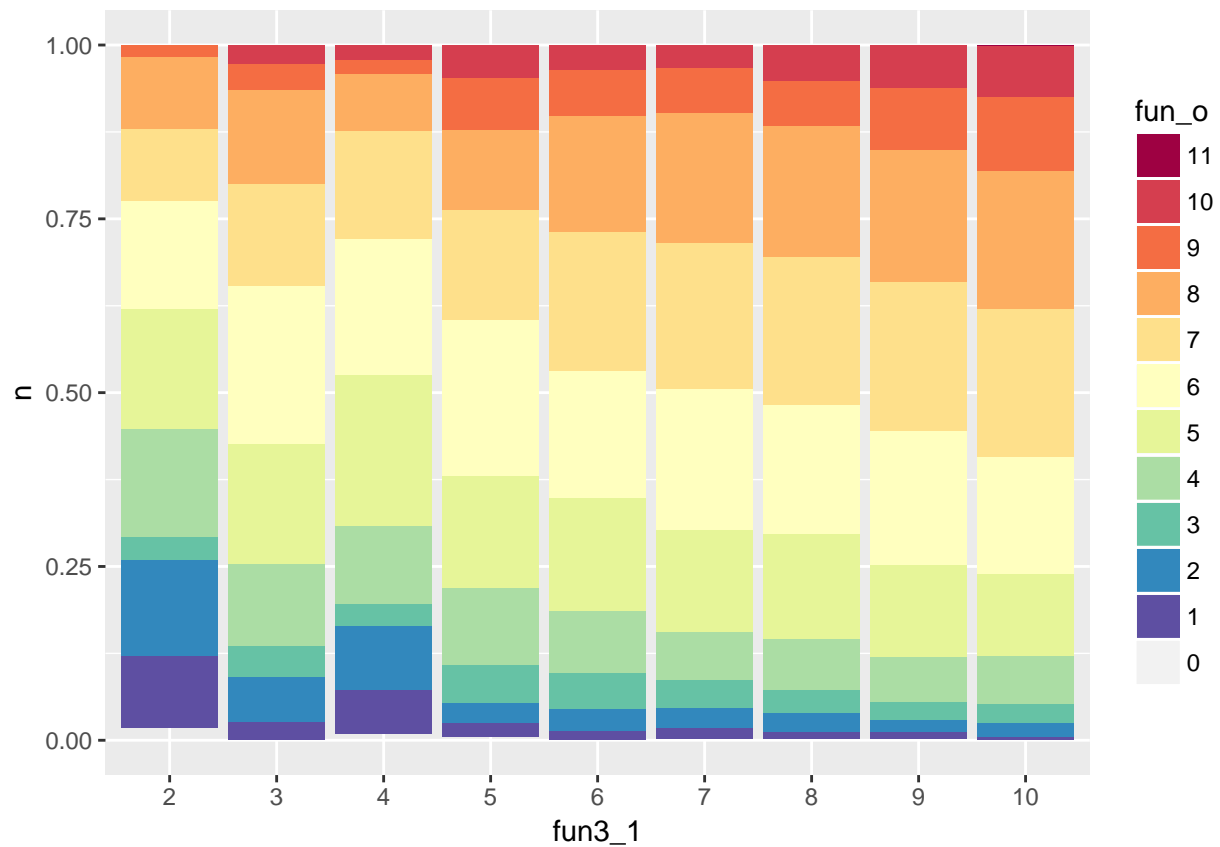



```
ggplot(data = na.omit(ambition), aes(x=amb3_1, y=n, fill=amb_o)) +
  geom_bar(position=position_fill(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")
```

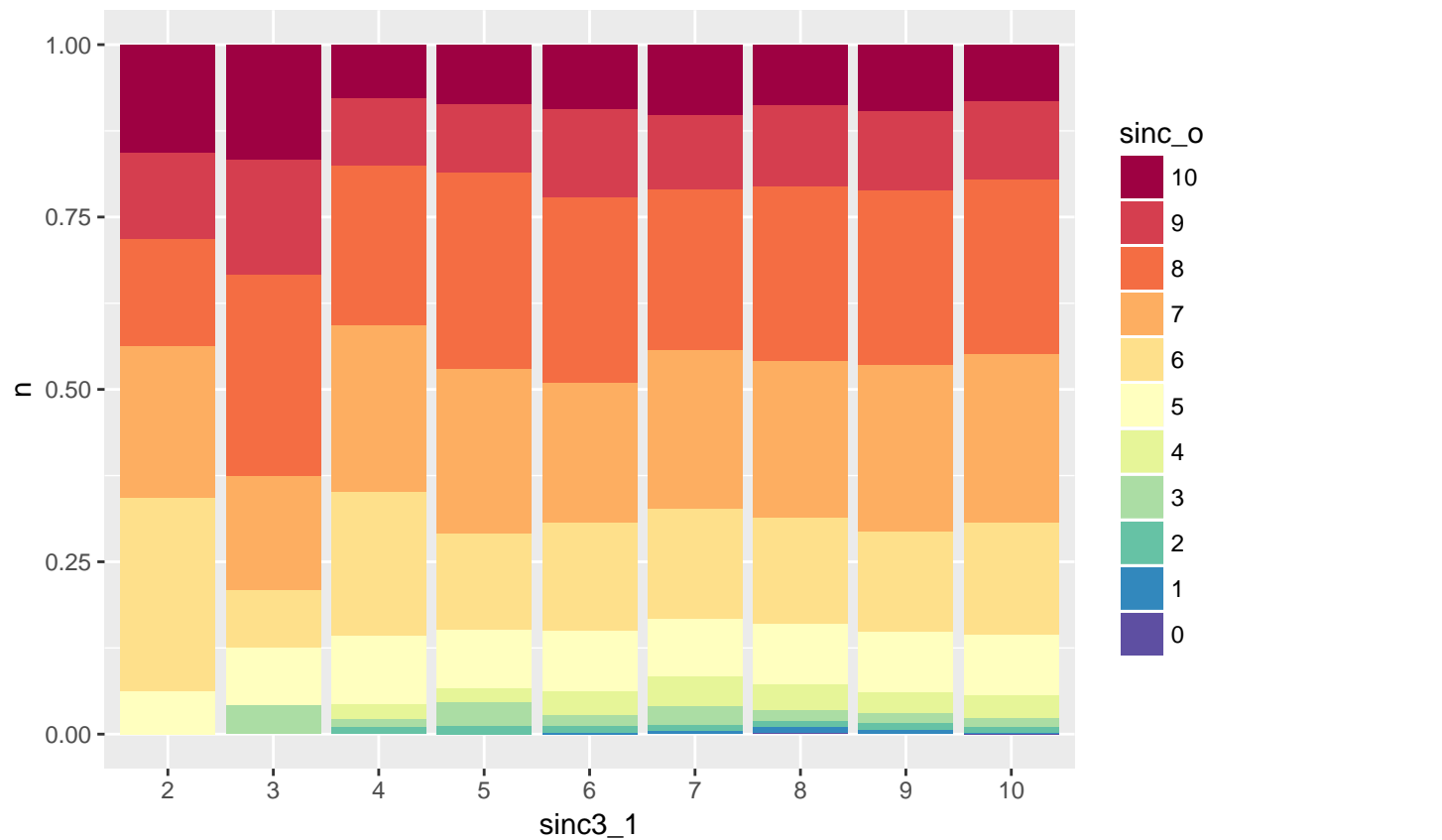


```
ggplot(data = na.omit(fun), aes(x=fun3_1, y=n, fill=fun_o)) +
geom_bar(position=position_fill(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")

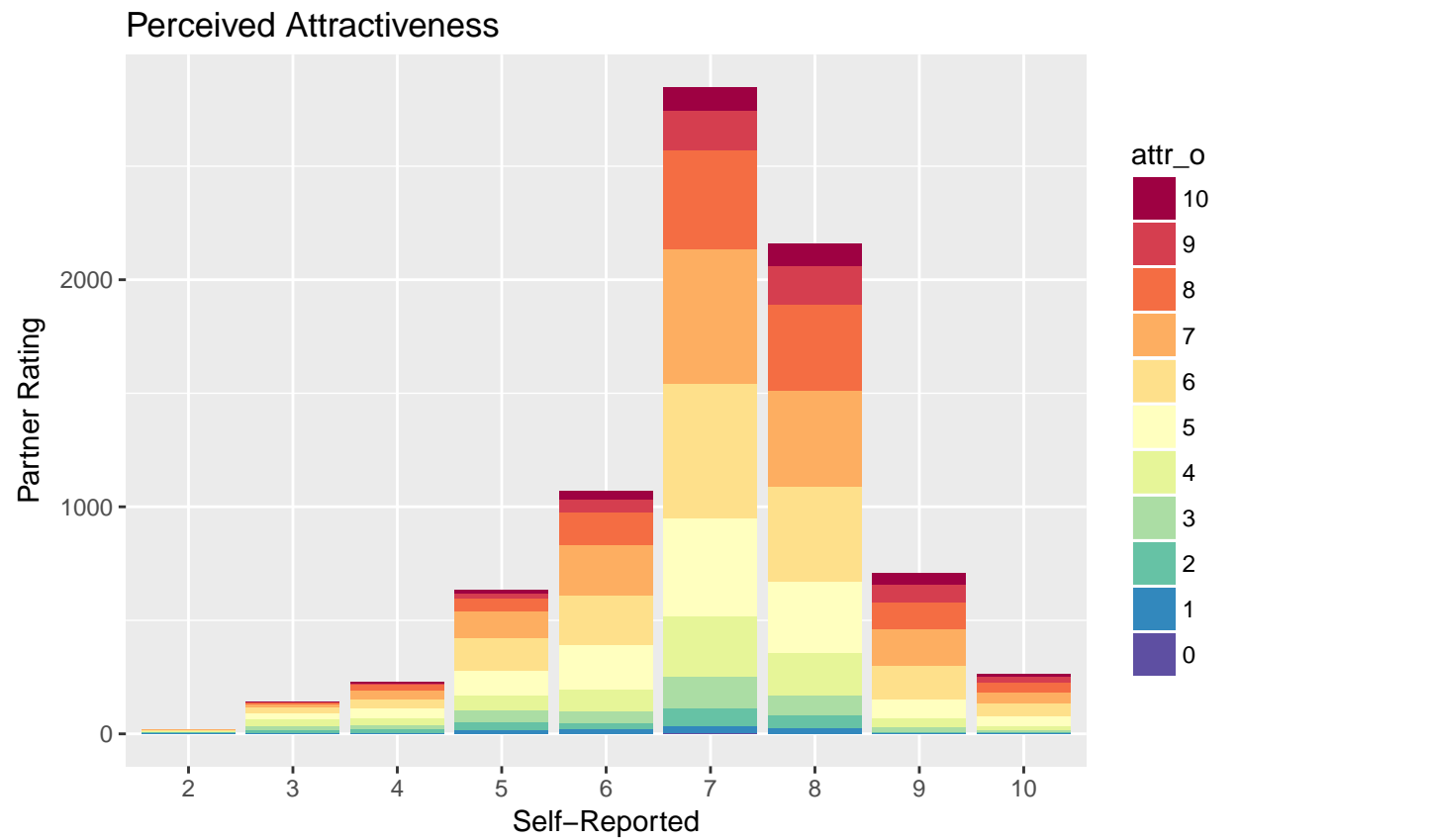
## Warning in RColorBrewer::brewer.pal(n, pal): n too large, allowed maximum for palette Spectral is 11
## Returning the palette you asked for with that many colors
```



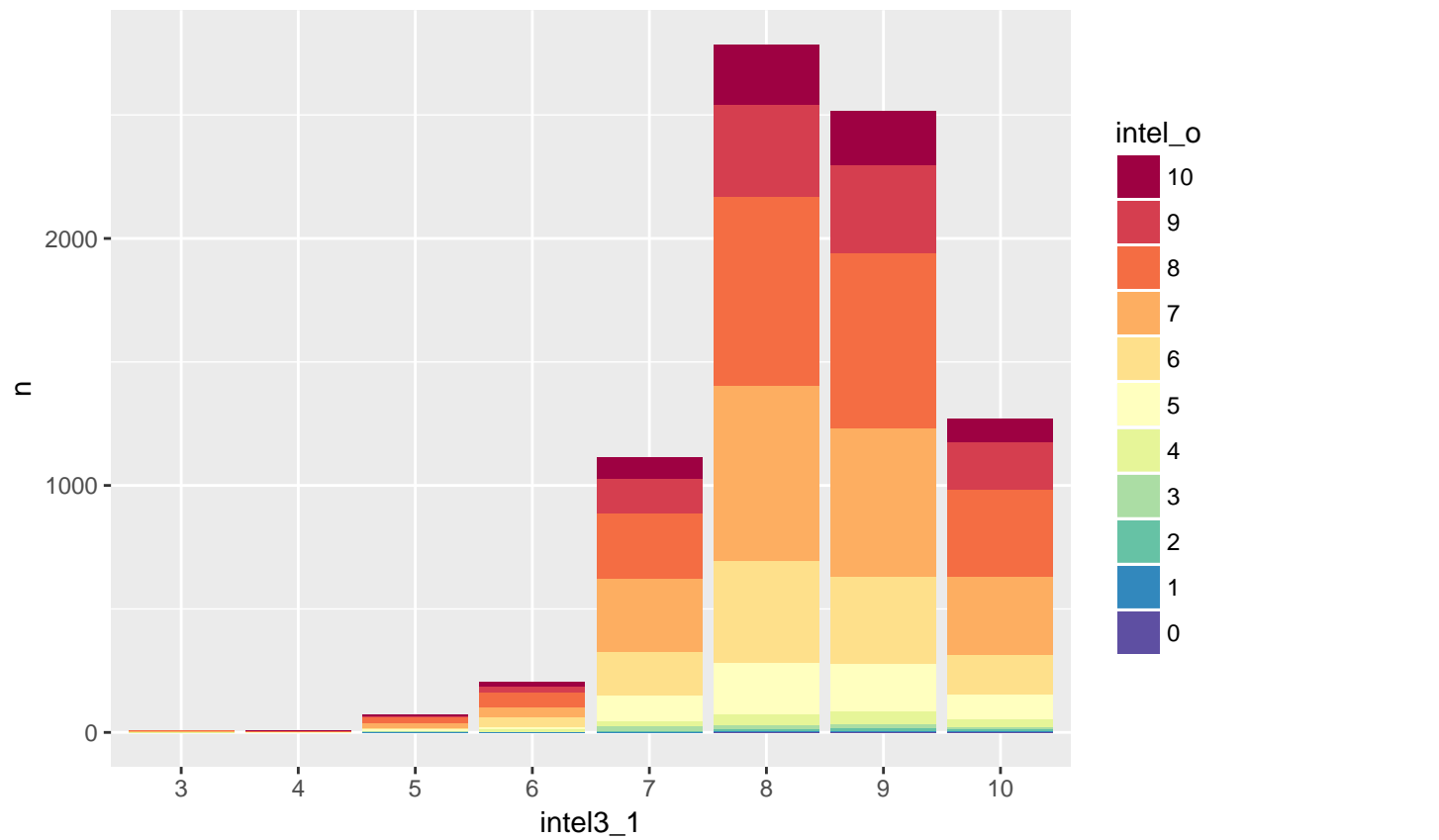
```
ggplot(data = na.omit(sincerity), aes(x=sinc3_1, y=n, fill=sinc_o)) +
  geom_bar(position=position_fill(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")
```



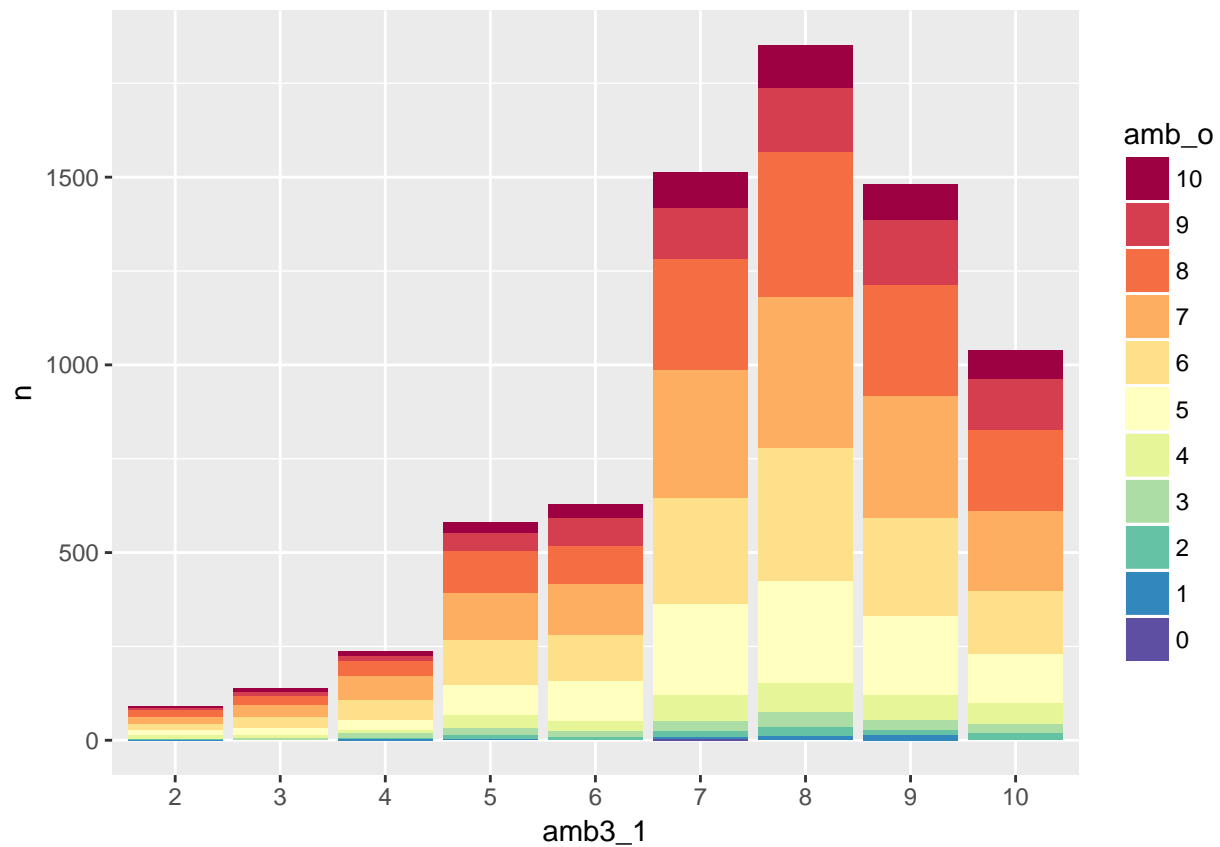
```
# show with frequency
ggplot(data = na.omit(hotness), aes(x=attr3_1, y=n, fill=attr_o)) +
  geom_bar(position=position_stack(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")
```



```
ggplot(data = na.omit(intelligence), aes(x=intel3_1, y=n, fill=intel_o)) +
  geom_bar(position=position_stack(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")
```

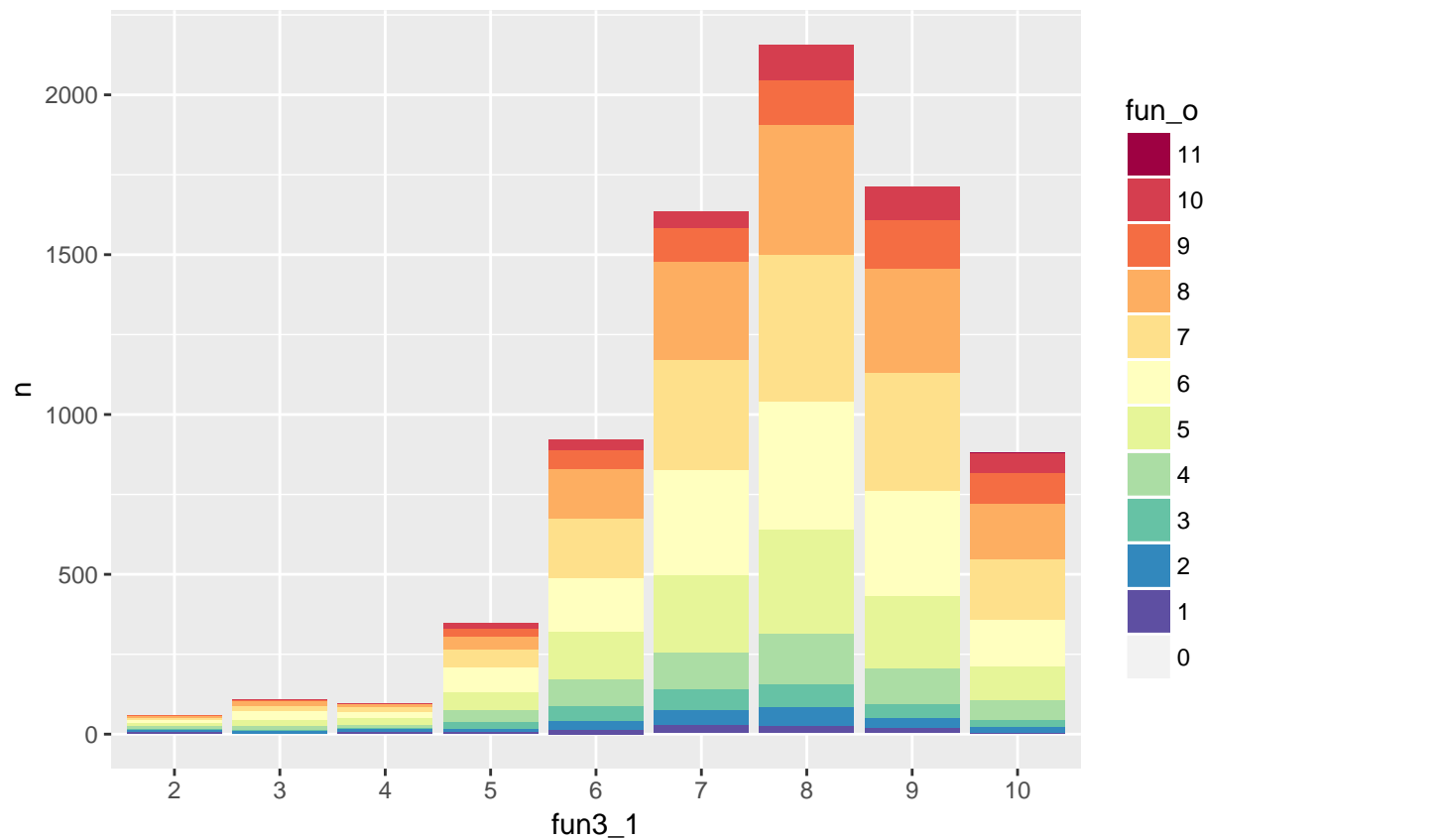


```
ggplot(data = na.omit(ambition), aes(x=amb3_1, y=n, fill=amb_o)) +  
geom_bar(position=position_stack(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")
```

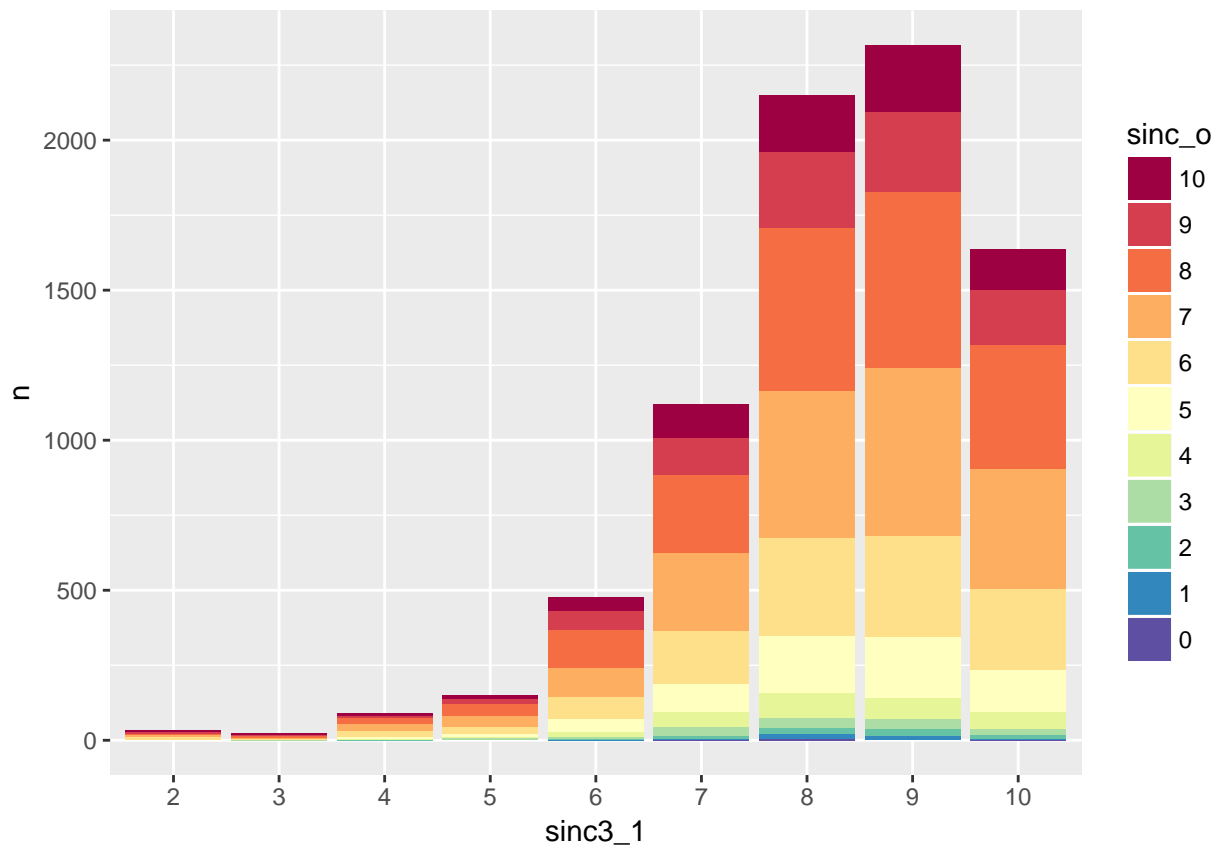


```
ggplot(data = na.omit(fun), aes(x=fun3_1, y=n, fill=fun_o)) +
geom_bar(position=position_stack(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")

## Warning in RColorBrewer::brewer.pal(n, pal): n too large, allowed maximum for palette Spectral is 11
## Returning the palette you asked for with that many colors
```



```
ggplot(data = na.omit(sincerity), aes(x=sinc3_1, y=n, fill=sinc_o)) +
  geom_bar(position=position_stack(reverse = TRUE), stat="identity") + scale_fill_brewer(palette="Spectral")
```

General Takeaway:

Attractiveness - the linear model does not explain much of the variation in peer rating. There may be some trend in increased perceived attractiveness with increased self-rating, however sample sizes of self-ratings ≤ 4 are small. Most people rated themselves a 7 or 8 in attractiveness. No one's.

Fun - very similar. Most people rated themselves a 7-9 in fun. A handful of one's.

Perhaps it is hard to determine traits like ambition, sincerity and intelligence in a 4 minute speed date. But it appears that people default to giving a peer review of 7 in all cases.

This is interesting because most people give themselves around a 7 as well. I've seen this trend in other 0-10 rating scales in other datasets. Is there discussion in the stats field about this phenomenon?

Finally: Turns out.. we kind of suck at predicting how attractive we are to others. What we think about ourselves is not correlated to what others see in us. Maybe there is no point to worrying about how strangers see us, because either they don't or we don't.