```r
# ////////////////////////////////////////////////////////////////
#
#       PROJECT ENVIRONMENT MANAGEMENT (renv)
#
# ////////////////////////////////////////////////////////////////

# This project uses the 'renv' package to manage dependencies and
# ensure reproducibility. All collaborators must use the same package
# versions defined in the 'renv.lock' file.

# 1. INITIAL SETUP FOR NEW USERS:
#    If this is the first time working on the project, the environment
#    should activate automatically when you open the 'Project.Rproj' file.
#    If prompted, confirm that you wish to restore the library.

# 2. MANUAL RESTORE (If activation fails):
#    If the environment doesn't load, run this command in the R Console
#    to install/restore the required packages and versions. This ensures
#    your setup matches the committed environment.
#
# renv::restore()

# 3. UPDATING PACKAGES:
#    If you install or update packages necessary for the project, run:
#    renv::snapshot()
#    Then, COMMIT the updated 'renv.lock' file to GitHub.
#
# ////////////////////////////////////////////////////////////////


# The file has been loaded
load("Group_3.RData")

# Changing the name of my variable
View(dat) # As table
raw_data <- dat
head(raw_data)
```

```
##       isced11_20 bkfamstd labgro20 expft20 expue20 exppt20 lfs20 siops08_20
## 12117          6        1       -2   10.00   3.833   3.417     1      32.00
## 13263          3        1     2900   21.25   0.500   0.000    11      25.00
## 4806           3        1     2300   10.50   0.000  24.250    11      49.00
## 11014          5        1     7100   26.25   0.000   0.000    11      42.17
## 8465           2        3      800    1.25   0.000   3.000    11      23.00
## 21853          1        3       -2    0.00   0.000   0.667     1      16.00
##           cid bkpbirthy sex bkp_05_02 bkp_01_06 bkp_01_03 bkp_02_01 bkp_02_02
## 12117 2155799      1982   2         6         0        -2         2         2
## 13263 2167347      1980   1         3         7         8         3         1
## 4806   314668      1960   2         4         5         5         3         2
## 11014 2117153      1973   1         2         8         8         3         2
## 8465  2013722      2000   1         6         0         2         4         2
## 21853 3178543      2000   2         6         5        -2         2         2
##       bkp_02_03 bkp_02_04 bkp_03 bkp_07_01 bkp_06_11 bkp_168_02 bkp_204 bkp_123
```

```
## 12117         5         2     10         0         7         1       7       3
## 13263         4         2      7         5         5         2       9       2
## 4806          3         3      5         3         6         3       5       3
## 11014         4         2      7         2         5         3       9       2
## 8465          2         3      7         3         2         3       1       2
## 21853         4         2      2        10         6         3       8       2
##       bkp_169 bkp_168_09 migback sumkids
## 12117       3          2       2       0
## 13263       2          2       1       2
## 4806        3          3       1       3
## 11014       2          2       1       2
## 8465        2          2       1       0
## 21853       3          2       2       0
```

```r
# ============================================================ #
# ============================================================ #
#                         DATA CLEANING
# ============================================================ #
# ============================================================ #

# Load the necessary library for data manipulation
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

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

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

```r
# Rename the columns in the 'new_data' data frame from 'raw_data'
new_data <- raw_data %>%
  rename(
    # New_Name = Actual_Name

    # Key Dependent Variable
    personal_achievement_deserved = bkp_05_02,

    # Demographics and Background #Ale
    education_level             = isced11_20,
    marital_status              = bkfamstd,
    birth_year                  = bkpbirthy,
    gender                      = sex,
    migration_background        = migback,
    total_children              = sumkids,
    household_id                = cid,

    # Financial and Employment  # Ale Meyer
    gross_labor_income          = labgro20,
```

```
    # to test for fulltime experience with 10 year steps
    # instead of single years
    exp_fulltime_years          = expft20 / 10,

    exp_unemployment_years      = expue20,
    exp_parttime_years          = exppt20,
    labor_force_status          = lfs20,

    # to test for jumps in prestige by 10 points
    # instead of just 1.0
    job_prestige_siops          = siops08_20 / 10,

    # Satisfaction and Attitudes #Deniz and Filippo
    satisfaction_income         = bkp_01_06,
    satisfaction_job            = bkp_01_03,
    feeling_angry               = bkp_02_01,
    feeling_worried             = bkp_02_02,
    feeling_happy               = bkp_02_03,
    feeling_sad                 = bkp_02_04,
    life_value_usefulness       = bkp_03,
    positive_attitude           = bkp_06_11,
    concern_economic_situation  = bkp_168_02,
    life_satisfaction_general   = bkp_204,
    health_status               = bkp_123,
    political_interest          = bkp_169,
    concern_social_cohesion     = bkp_168_09,
    number_close_friends        = bkp_07_01
  )

head(new_data)
```

```
##       education_level marital_status gross_labor_income exp_fulltime_years
## 12117               6              1                -2              10.00
## 13263               3              1              2900              21.25
## 4806                3              1              2300              10.50
## 11014               5              1              7100              26.25
## 8465                2              3               800               1.25
## 21853               1              3                -2               0.00
##       exp_unemployment_years exp_parttime_years labor_force_status
## 12117                  3.833             3.417                  1
## 13263                  0.500             0.000                 11
## 4806                   0.000            24.250                 11
## 11014                  0.000             0.000                 11
## 8465                   0.000             3.000                 11
## 21853                  0.000             0.667                  1
##       job_prestige_siops household_id birth_year gender
## 12117              32.00      2155799       1982      2
## 13263              25.00      2167347       1980      1
## 4806               49.00       314668       1960      2
## 11014              42.17      2117153       1973      1
## 8465               23.00      2013722       2000      1
## 21853              16.00      3178543       2000      2
##       personal_achievement_deserved satisfaction_income satisfaction_job
```

```
## 12117                            6                   0               -2
## 13263                            3                   7                8
## 4806                             4                   5                5
## 11014                            2                   8                8
## 8465                             6                   0                2
## 21853                            6                   5               -2
##       feeling_angry feeling_worried feeling_happy feeling_sad
## 12117             2               2             5           2
## 13263             3               1             4           2
## 4806              3               2             3           3
## 11014             3               2             4           2
## 8465              4               2             2           3
## 21853             2               2             4           2
##       life_value_usefulness number_close_friends positive_attitude
## 12117                    10                    0                 7
## 13263                     7                    5                 5
## 4806                      5                    3                 6
## 11014                     7                    2                 5
## 8465                      7                    3                 2
## 21853                     2                   10                 6
##       concern_economic_situation life_satisfaction_general health_status
## 12117                          1                         7             3
## 13263                          2                         9             2
## 4806                           3                         5             3
## 11014                          3                         9             2
## 8465                           3                         1             2
## 21853                          3                         8             2
##       political_interest concern_social_cohesion migration_background
## 12117                  3                       2                    2
## 13263                  2                       2                    1
## 4806                   3                       3                    1
## 11014                  2                       2                    1
## 8465                   2                       2                    1
## 21853                  3                       2                    2
##       total_children
## 12117              0
## 13263              2
## 4806               3
## 11014              2
## 8465               0
## 21853              0
```

```
# ============================================================ #
#                        END DATA CLEANING
# ============================================================ #


# ============================================================ #
# ============================================================ #
#                        DUMMY VARIABLES
# ============================================================ #
# ============================================================ #

# We add all dummy variables to the dataframe
```

```r
new_data <- new_data |>
  mutate(

    # ========================================-
    # =====-- Demographics and Background ====--
    # ========================================-

    # -- Education_level --

    # We make 3 education classes for low, average (reference),
    # and high education as professor suggested
    education_low = case_when(
      education_level %in% c(1, 2) ~ 1,
      education_level == 0 ~ NA_real_,
      education_level < 0 ~ NA_real_,
      TRUE ~ 0,
    ),

    education_high = case_when(
      education_level %in% c(6, 7, 8) ~ 1,
      education_level == 0 ~ NA_real_,
      education_level < 0 ~ NA_real_,
      TRUE ~ 0
    ),

    # -- Marital status --
    # with married living together, married living separate
    # and single/unmarried (reference)
    # Reference: single/unmarried
    married_together = case_when(
      marital_status %in% c(1, 7) ~ 1,
      marital_status < 0 ~ NA_real_,
      TRUE ~ 0
    ),
    married_separate = case_when(
      marital_status %in% c(2, 6, 8) ~ 1,
      marital_status < 0 ~ NA_real_,
      TRUE ~ 0
    ),

    # -- Birth_year --

    # We use conditional logic to handle missing birth years (-1, -2, etc.)
    # We use 2020 as the reference year to match the survey wave (labgro20)
    age = case_when(
      birth_year < 0 ~ NA_real_, # If negative (missing code), set Age to NA
      # age calculation, we look at the age variable in 10 year steps (decades)
      TRUE ~ (2020 - birth_year)/10
    ),

    # -- Gender --

    # Reference: Male(0)
```

```
  gender_female = case_when(
    gender == 2 ~ 1,       # Female (code 2) becomes 1
    gender == 1 ~ 0,       # Male (code 1) becomes 0 (reference)
    TRUE ~ NA_real_        # Invalid/Non-response codes become NA
  ),

  # -- Migration_background --

  # Reference: No background
  migration_direct = case_when(
    migration_background == 2 ~ 1,
    migration_background < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  migration_indirect = case_when(
    migration_background == 3 ~ 1,
    migration_background < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- Total_children --

  # Reference: having no kids
  one_or_two_children = case_when(
    total_children %in% c(1,2) ~ 1,
    total_children < 0 ~ NA_real_, #Error!
    TRUE ~ 0
  ),

  more_than_two_children = case_when(
    total_children > 2 ~ 1,
    total_children < 0 ~ NA_real_, #Error fix
    TRUE ~ 0
  ),

  # -- Household_id --


  # ========================================-
  # =====-- Financial and Employment ====--
  # ========================================-

  # -- gross_labor_income --

  # transforms income into 1,000 unit steps
  # (could also use 500 unit steps or income classes instead)
  #income_per_1000 = (gross_labor_income/1000),
  #Delete this


  # -- exp_fulltime_years --
```

```r
# -- exp_unemployment_years --


# -- exp_parttime_years --


# -- labor_force_status --

# dummy variables for labor force status, with "working" as reference
# unemployed or only secondary job
unemployed_or_minimal = case_when(
  labor_force_status %in% c(1, 6, 8, 9, 10, 13) ~ 1,
  labor_force_status < 0 ~ NA_real_,
  TRUE ~ 0
),
# non working due to reasons: education, pension, military, parental leave
non_working = case_when(
  labor_force_status %in% c(2, 3, 4, 5) ~ 1,
  labor_force_status < 0 ~ NA_real_,
  TRUE ~ 0
),

# -- job_prestige_siops --


# ========================================-
# =====-- Satisfaction and Attitudes ====--
# ========================================-

# -- satisfaction_income --

# -- satisfaction_job --

# -- feeling_angry --

# Reference: very rarely or rarely angry
# Because want to test the effect of having the negative feelings:
# angriness, worriedness and sadness
angry_often = case_when(
  feeling_angry %in% c(3, 4, 5) ~ 1, # Very often, Often, somewhat = 1
  feeling_angry < 0 ~ NA_real_,
  TRUE ~ 0
),

# -- feeling_worried --

# Reference: very rarely or rarely worried
worried_often = case_when(
  feeling_worried %in% c(3, 4, 5) ~ 1, # Very often, Often, somewhat = 1
  feeling_worried < 0 ~ NA_real_,
  TRUE ~ 0
),
```

```r
  # -- feeling_sad --

  # Reference: very rarely or rarely sad
  sad_often = case_when(
    feeling_sad %in% c(3, 4, 5) ~ 1,      # Very often, Often, somewhat = 1
    feeling_sad < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- feeling_happy --

  # Reference: often or very often happy, because we assume that
  # being happy is the default.
  # > we want to test the effect of NOT being happy
  not_happy = case_when(
    feeling_happy %in% c(1, 2) ~ 1,    # very rarely (1) or rarely (2) = 1
    feeling_happy < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- life_value_usefulness --
  # likert scale is from 0-10 > large enough to assume metric data
  # and take the raw values

  # -- positive_attitude --
  # likert scale is from 1-7 > large enough to assume metric data
  # and take the raw values

  # -- concern_economic_situation --

  # dummy for having no concerns about own economic situation
  # with having some worries (concern_economic_situation = 2) as reference
  economy_not_worried = case_when(
    concern_economic_situation == 3 ~ 1,
    concern_economic_situation < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # dummy for having a lot of concerns about own economic situation
  economy_worried = case_when(
    concern_economic_situation == 1 ~ 1,
    concern_economic_situation < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- life_satisfaction_general --

  # -- health_status --

  # Reference: good health
  # dummy for the health status with very good=1 and good=2 as reference
  # (this is the reference because the largest group is health being good)
  # dummy variable = 1 means their health status is only satisfactory or below
```

```
  health_not_good = case_when(
    health_status %in% c(3,4,5) ~ 1,
    health_status < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- political_interest --

  # Reference: political interest being not so strong (largest group)
  # dummy for having strong political interest
  strong_political_interest = case_when(
    political_interest %in% c(1,2) ~ 1,
    political_interest < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # dummy for having strong political interest at all
  no_political_interest = case_when(
    political_interest == 4 ~ 1,
    political_interest < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- concern_social_cohesion --

  # Reference: Having some worries (concern_social_cohesion = 2)
  # dummy for having a lot of concerns about social cohesion
  social_great_concern = case_when(
    concern_social_cohesion == 1 ~ 1,
    concern_social_cohesion < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # dummy for having no concerns about social cohesion
  social_no_concern = case_when(
    concern_social_cohesion == 3 ~ 1,
    concern_social_cohesion < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  # -- number_close_friends --

  # Reference is average number of close friends (2 - 4)
  low_friends = case_when(
    number_close_friends %in% c(0, 1) ~ 1,
    labor_force_status < 0 ~ NA_real_,
    TRUE ~ 0
  ),

  high_friends = case_when(
    number_close_friends %in% c(5,6,7,8) ~ 1,
    labor_force_status < 0 ~ NA_real_,
    TRUE ~ 0
```

```
    ),

    very_high_friends = case_when(
      number_close_friends > 8 ~ 1,
      labor_force_status < 0 ~ NA_real_,
      TRUE ~ 0
    ),

    # --- Further dummy variables for grouping multiple variables

    # dummy for grouping bad feelings in general
    # Reference: not experiencing bad feelings across all 3 categories
    # bad_feeling_overall = 1 when person experiences angriness, sadness
    # and lack of happiness all at the same time
    # worriedness is not included because it has been identified as a
    # different type of bad feeling and doesn't correlate well with others
    bad_feeling_overall = case_when(
      feeling_angry %in% c(3,4,5) &
        feeling_sad %in% c(3,4,5) &
        feeling_happy %in% c(1,2) ~ 1,
      TRUE ~ 0
    ),

    # Dummy for checking extremely high life satisfaction
    # Reference: values from 1-8 on the likert scale
    high_life_satisfaction = case_when(
      life_satisfaction_general %in% c(9,10) ~ 1,
      life_satisfaction_general < 0 ~ NA_real_,
      TRUE ~ 0
    ),

    # Dummy for experiencing high satisfaction and psychological values
    # across 3 categories
    # Reference: not experiencing high satisfaction (> 7) in all 3 categories
    high_satisfaction_values = case_when(
      satisfaction_income > 7 &
        satisfaction_job > 7 &
        life_value_usefulness > 7 ~ 1,
      TRUE ~ 0
    ),

    # lOGTEST
    # Dummy for optional log tests, please do NOT delete
    achievement_deserved = case_when(
      personal_achievement_deserved %in% c(1,2,3) ~ 1,
      personal_achievement_deserved < 0 ~ NA_real_,
      TRUE ~ 0
    )

  )

# replace all other negative values we haven't replaced yet also with NA
new_data[new_data < 0] <- NA
```

```r
sum(new_data$gross_labor_income == 0, na.rm = TRUE)
```

```
## [1] 50
```

```r
# Because we use log() in the modelling phase for gross labor income
# We can not have income values that equal 0
# (there are 50 gross labor income values in the data set with the value 0)
new_data$gross_labor_income[new_data$gross_labor_income == 0] <- NA

# ============================================================== #
#                         END DUMMY VARIABLES
# ============================================================== #


# Check the result
head(new_data)
```

```
##       education_level marital_status gross_labor_income exp_fulltime_years
## 12117               6              1                 NA              10.00
## 13263               3              1               2900              21.25
## 4806                3              1               2300              10.50
## 11014               5              1               7100              26.25
## 8465                2              3                800               1.25
## 21853               1              3                 NA               0.00
##       exp_unemployment_years exp_parttime_years labor_force_status
## 12117                  3.833             3.417                  1
## 13263                  0.500             0.000                 11
## 4806                   0.000            24.250                 11
## 11014                  0.000             0.000                 11
## 8465                   0.000             3.000                 11
## 21853                  0.000             0.667                  1
##       job_prestige_siops household_id birth_year gender
## 12117              32.00      2155799       1982      2
## 13263              25.00      2167347       1980      1
## 4806               49.00       314668       1960      2
## 11014              42.17      2117153       1973      1
## 8465               23.00      2013722       2000      1
## 21853              16.00      3178543       2000      2
##       personal_achievement_deserved satisfaction_income satisfaction_job
## 12117                             6                  0              NA
## 13263                             3                  7               8
## 4806                              4                  5               5
## 11014                             2                  8               8
## 8465                              6                  0               2
## 21853                             6                  5              NA
##       feeling_angry feeling_worried feeling_happy feeling_sad
## 12117             2               2             5           2
## 13263             3               1             4           2
## 4806              3               2             3           3
## 11014             3               2             4           2
## 8465              4               2             2           3
## 21853             2               2             4           2
```

```
##       life_value_usefulness number_close_friends positive_attitude
## 12117                    10                    0                 7
## 13263                     7                    5                 5
## 4806                      5                    3                 6
## 11014                     7                    2                 5
## 8465                      7                    3                 2
## 21853                     2                   10                 6
##       concern_economic_situation life_satisfaction_general health_status
## 12117                          1                         7             3
## 13263                          2                         9             2
## 4806                           3                         5             3
## 11014                          3                         9             2
## 8465                           3                         1             2
## 21853                          3                         8             2
##       political_interest concern_social_cohesion migration_background
## 12117                  3                       2                    2
## 13263                  2                       2                    1
## 4806                   3                       3                    1
## 11014                  2                       2                    1
## 8465                   2                       2                    1
## 21853                  3                       2                    2
##       total_children education_low education_high married_together
## 12117              0             0              1                 1
## 13263              2             0              0                 1
## 4806               3             0              0                 1
## 11014              2             0              0                 1
## 8465               0             1              0                 0
## 21853              0             1              0                 0
##       married_separate age gender_female migration_direct migration_indirect
## 12117                0 3.8             1                1                  0
## 13263                0 4.0             0                0                  0
## 4806                 0 6.0             1                0                  0
## 11014                0 4.7             0                0                  0
## 8465                 0 2.0             0                0                  0
## 21853                0 2.0             1                1                  0
##       one_or_two_children more_than_two_children unemployed_or_minimal
## 12117                   0                      0                     1
## 13263                   1                      0                     0
## 4806                    0                      1                     0
## 11014                   1                      0                     0
## 8465                    0                      0                     0
## 21853                   0                      0                     1
##       non_working angry_often worried_often sad_often not_happy
## 12117           0           0             0         0         0
## 13263           0           1             0         0         0
## 4806            0           1             0         1         0
## 11014           0           1             0         0         0
## 8465            0           1             0         1         1
## 21853           0           0             0         0         0
##       economy_not_worried economy_worried health_not_good
## 12117                   0               1               1
## 13263                   0               0               0
## 4806                    1               0               1
## 11014                   1               0               0
```

```
## 8465                    1              0              0
## 21853                   1              0              0
##       strong_political_interest no_political_interest social_great_concern
## 12117                         0                     0                    0
## 13263                         1                     0                    0
## 4806                          0                     0                    0
## 11014                         1                     0                    0
## 8465                          1                     0                    0
## 21853                         0                     0                    0
##       social_no_concern low_friends high_friends very_high_friends
## 12117                 0           1            0                 0
## 13263                 0           0            1                 0
## 4806                  1           0            0                 0
## 11014                 0           0            0                 0
## 8465                  0           0            0                 0
## 21853                 0           0            0                 1
##       bad_feeling_overall high_life_satisfaction high_satisfaction_values
## 12117                   0                     0                        0
## 13263                   0                     1                        0
## 4806                    0                     0                        0
## 11014                   0                     1                        0
## 8465                    1                     0                        0
## 21853                   0                     0                        0
##       achievement_deserved
## 12117                    0
## 13263                    1
## 4806                     0
## 11014                    1
## 8465                     0
## 21853                    0
```

```r
summary(new_data$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.800   3.400   4.900   4.887   6.200   9.800       1
```

```r
# ================================================================ #
# ================================================================ #
#                EXPERIMENTATION WITH MODELLING lm()
# ================================================================ #
# ================================================================ #


mod1 <- lm(personal_achievement_deserved ~ gender_female + age,
           data = new_data)
summary(mod1)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age, data = new_data)
##
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
## -2.166 -1.129 -0.112  1.021  4.032
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.03956    0.05821  52.218   <2e-16 ***
## gender_female -0.09516    0.03711  -2.564   0.0104 *
## age            0.01293    0.01052   1.229   0.2190
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.739 on 8790 degrees of freedom
##   (207 observations deleted due to missingness)
## Multiple R-squared:  0.000931,   Adjusted R-squared:  0.0007036
## F-statistic: 4.095 on 2 and 8790 DF,  p-value: 0.01668
```

```
# we see no correlation between age and main variable, there is an effect
# due to gender p value is < 0.05, but the effect is quite small


mod2 <- lm(personal_achievement_deserved ~ gender_female + age + education_low
           + education_high,
           data = new_data)
summary(mod2)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high, data = new_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -2.873 -1.450 -0.297  1.325  4.544
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     2.95317    0.06277  47.044  < 2e-16 ***
## gender_female  -0.12549    0.03673  -3.417 0.000636 ***
## age             0.04647    0.01071   4.338 1.46e-05 ***
## education_low   0.53845    0.05457   9.867  < 2e-16 ***
## education_high -0.49708    0.04217 -11.786  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.707 on 8674 degrees of freedom
##   (321 observations deleted due to missingness)
## Multiple R-squared:  0.03604,    Adjusted R-squared:  0.0356
## F-statistic: 81.08 on 4 and 8674 DF,  p-value: < 2.2e-16
```

```
# p values are all significant, the effects from gender and age are negligible,
# we can see a very significant effect due to high or low education though
```

```r
mod3 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + married_together
  + married_separate,
  data = new_data
)
summary(mod3)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + married_together +
##     married_separate, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0780 -1.4426 -0.3009  1.3101  4.5427
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        2.95309    0.06377  46.307  < 2e-16 ***
## gender_female     -0.13395    0.03704  -3.617   0.0003 ***
## age                0.04899    0.01126   4.349 1.38e-05 ***
## education_low      0.54593    0.05544   9.847  < 2e-16 ***
## education_high    -0.49413    0.04249 -11.629  < 2e-16 ***
## married_together  -0.03366    0.04013  -0.839   0.4016
## married_separate   0.24587    0.10834   2.269   0.0233 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.705 on 8534 degrees of freedom
##   (459 observations deleted due to missingness)
## Multiple R-squared:  0.0372, Adjusted R-squared:  0.03653
## F-statistic: 54.96 on 6 and 8534 DF,  p-value: < 2.2e-16
```

```r
# we add marriage and see bad p values for it, meaning no effect,
# this could be due to the data not capturing regular relationships
# but just marriages


mod4 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + unemployed_or_minimal
  + non_working,
  data = new_data
)
summary(mod4)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + unemployed_or_minimal +
##     non_working, data = new_data)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1381 -1.4134 -0.3847  1.2863  4.5981
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            2.86539    0.06555  43.710  < 2e-16 ***
## gender_female         -0.14509    0.03667  -3.957 7.65e-05 ***
## age                    0.05745    0.01229   4.673 3.01e-06 ***
## education_low          0.48561    0.05556   8.740  < 2e-16 ***
## education_high        -0.47350    0.04211 -11.244  < 2e-16 ***
## unemployed_or_minimal  0.48261    0.05660   8.527  < 2e-16 ***
## non_working           -0.08602    0.05268  -1.633    0.103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.699 on 8672 degrees of freedom
##   (321 observations deleted due to missingness)
## Multiple R-squared:  0.04558,    Adjusted R-squared:  0.04492
## F-statistic: 69.03 on 6 and 8672 DF,  p-value: < 2.2e-16
```

```r
# being unemployed or only having a minimal job > great effect of 0.54


mod5 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2),
  data = new_data
)
summary(mod5)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2), data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9505 -1.3125 -0.3768  1.1898  5.3105
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    0.02572    0.71309   0.036 0.971225
## gender_female                 -0.29315    0.04716  -6.216 5.46e-10 ***
## age                            0.06891    0.01773   3.886 0.000103 ***
## education_low                  0.45612    0.07471   6.105 1.09e-09 ***
## education_high                -0.25287    0.05220  -4.844 1.30e-06 ***
## log(gross_labor_income)        1.08932    0.19167   5.683 1.39e-08 ***
## I(log(gross_labor_income)^2) -0.09267    0.01322  -7.008 2.71e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 1.667 on 5650 degrees of freedom
##   (3343 observations deleted due to missingness)
## Multiple R-squared:  0.05317,    Adjusted R-squared:  0.05216
## F-statistic: 52.88 on 6 and 5650 DF,  p-value: < 2.2e-16
```

```
# incomes effect seems to be very small too, until now education prevails
# especially the effect of low education, additionally the gender_female
# variable seems to get traction


mod6 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general,
  data = new_data
)
summary(mod6)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9600 -1.2225 -0.3671  1.1161  5.5390
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.18570    0.72332   3.022  0.00252 **
## gender_female                 -0.28700    0.04635  -6.192 6.36e-10 ***
## age                            0.05657    0.01771   3.194  0.00141 **
## education_low                  0.48407    0.07316   6.617 4.02e-11 ***
## education_high                -0.23096    0.05104  -4.525 6.16e-06 ***
## log(gross_labor_income)        1.07011    0.19089   5.606 2.17e-08 ***
## I(log(gross_labor_income)^2) -0.08832    0.01311  -6.737 1.78e-11 ***
## satisfaction_job             -0.10119    0.01183  -8.550  < 2e-16 ***
## life_satisfaction_general    -0.19597    0.01557 -12.588  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.612 on 5513 degrees of freedom
##   (3478 observations deleted due to missingness)
## Multiple R-squared:  0.1128, Adjusted R-squared:  0.1115
## F-statistic: 87.58 on 8 and 5513 DF,  p-value: < 2.2e-16
```

```
# for life and job satisfaction we can see good effects with very good p values
# for every 5 points in life satisfaction (scale is 1-10), the main variable
# decreases by a whole point (they think they achieved what they deserved)
```

```r
mod7 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job
  + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends,
  data = new_data
)
summary(mod7)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7763 -1.2026 -0.3535  1.0815  5.4565
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.43749    0.72405   3.366 0.000767 ***
## gender_female                 -0.28286    0.04655  -6.076 1.31e-09 ***
## age                            0.05881    0.01783   3.299 0.000977 ***
## education_low                  0.45374    0.07348   6.175 7.08e-10 ***
## education_high                -0.23569    0.05114  -4.608 4.15e-06 ***
## log(gross_labor_income)        1.07970    0.19056   5.666 1.54e-08 ***
## I(log(gross_labor_income)^2) -0.08862    0.01309  -6.771 1.41e-11 ***
## satisfaction_job              -0.08640    0.01219  -7.090 1.51e-12 ***
## life_satisfaction_general     -0.15837    0.01706  -9.283  < 2e-16 ***
## life_value_usefulness         -0.06300    0.01493  -4.221 2.47e-05 ***
## positive_attitude             -0.04140    0.01955  -2.118 0.034224 *
## low_friends                    0.08948    0.07279   1.229 0.219017
## high_friends                  -0.05048    0.05000  -1.010 0.312771
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.607 on 5486 degrees of freedom
##   (3501 observations deleted due to missingness)
## Multiple R-squared:  0.118,  Adjusted R-squared:  0.1161
## F-statistic: 61.16 on 12 and 5486 DF,  p-value: < 2.2e-16
```

```r
# number of close friends is completely insignificant, life_value_usefulness
# has a moderate effect


mod8 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
```

```
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2) + satisfaction_job
  + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends
  + migration_direct
  + migration_indirect,
  data = new_data
)
summary(mod8)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0595 -1.1845 -0.3362  1.0506  5.4652
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.74607    0.71855   3.822 0.000134 ***
## gender_female                 -0.26375    0.04616  -5.714 1.16e-08 ***
## age                            0.08896    0.01808   4.921 8.88e-07 ***
## education_low                  0.34908    0.07351   4.748 2.10e-06 ***
## education_high                -0.27549    0.05086  -5.416 6.35e-08 ***
## log(gross_labor_income)        0.94707    0.18925   5.004 5.78e-07 ***
## I(log(gross_labor_income)^2) -0.07825    0.01301  -6.016 1.90e-09 ***
## satisfaction_job              -0.08213    0.01208  -6.798 1.18e-11 ***
## life_satisfaction_general     -0.16169    0.01691  -9.564  < 2e-16 ***
## life_value_usefulness         -0.06956    0.01480  -4.699 2.67e-06 ***
## positive_attitude             -0.06462    0.01950  -3.314 0.000926 ***
## low_friends                    0.04903    0.07222   0.679 0.497264
## high_friends                  -0.03771    0.04955  -0.761 0.446643
## migration_direct               0.54909    0.05342  10.278  < 2e-16 ***
## migration_indirect             0.18472    0.08910   2.073 0.038206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.592 on 5484 degrees of freedom
##   (3501 observations deleted due to missingness)
## Multiple R-squared:  0.1347, Adjusted R-squared:  0.1325
## F-statistic: 60.96 on 14 and 5484 DF,  p-value: < 2.2e-16
```

```
# next we can see that the feeling of being happy is a good predictor with a
# very good p value, same with being a direct migrant
# conclusion for now: good predictors: education, life satisfaction in general,
# feeling of happiness, direct migration
# no correlation: number of close friends, age, gender (only very minimal)
# > But we still have to keep age and gender in tests as controlling variables,
```

```r
# same with income


# new tests added start here
# Starting from here, we use the previous model as a base
# and gradually tested some other variables one by one using this base

# testing of all the feelings variables
mod9 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + not_happy
  + migration_direct + migration_indirect,
  data = new_data
)
summary(mod9)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + not_happy + migration_direct + migration_indirect,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9646 -1.1895 -0.3289  1.0569  5.4184
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.56433    0.71960   3.564 0.000369 ***
## gender_female                 -0.26052    0.04614  -5.646 1.72e-08 ***
## age                            0.08315    0.01813   4.587 4.61e-06 ***
## education_low                  0.34455    0.07347   4.690 2.80e-06 ***
## education_high                -0.27328    0.05085  -5.374 8.01e-08 ***
## log(gross_labor_income)        0.95634    0.18912   5.057 4.40e-07 ***
## I(log(gross_labor_income)^2) -0.07886    0.01300  -6.067 1.39e-09 ***
## satisfaction_job              -0.08269    0.01208  -6.847 8.35e-12 ***
## life_satisfaction_general     -0.14751    0.01735  -8.503  < 2e-16 ***
## life_value_usefulness         -0.06703    0.01481  -4.527 6.11e-06 ***
## positive_attitude             -0.05947    0.01953  -3.045 0.002339 **
## low_friends                    0.03602    0.07226   0.498 0.618193
## high_friends                  -0.03945    0.04954  -0.796 0.425871
## not_happy                      0.32167    0.08833   3.642 0.000273 ***
## migration_direct               0.53747    0.05349  10.047  < 2e-16 ***
## migration_indirect             0.18031    0.08903   2.025 0.042902 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.591 on 5477 degrees of freedom
##   (3507 observations deleted due to missingness)
## Multiple R-squared:  0.1367, Adjusted R-squared:  0.1343
## F-statistic: 57.81 on 15 and 5477 DF,  p-value: < 2.2e-16
```

```r
# added not_happy


mod10 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + not_happy
  + angry_often + migration_direct + migration_indirect,
  data = new_data
)
summary(mod10)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + not_happy + angry_often + migration_direct +
##     migration_indirect, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7778 -1.1759 -0.3247  1.0306  5.2905
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.26403    0.72024   3.143 0.001679 **
## gender_female                 -0.27447    0.04612  -5.952 2.82e-09 ***
## age                            0.08571    0.01809   4.738 2.21e-06 ***
## education_low                  0.34369    0.07331   4.688 2.82e-06 ***
## education_high                -0.26699    0.05073  -5.263 1.47e-07 ***
## log(gross_labor_income)        0.94522    0.18860   5.012 5.56e-07 ***
## I(log(gross_labor_income)^2) -0.07818    0.01296  -6.031 1.73e-09 ***
## satisfaction_job              -0.07422    0.01216  -6.103 1.11e-09 ***
## life_satisfaction_general     -0.13796    0.01741  -7.925 2.74e-15 ***
## life_value_usefulness         -0.06776    0.01478  -4.585 4.63e-06 ***
## positive_attitude             -0.05052    0.01956  -2.583 0.009825 **
## low_friends                    0.04061    0.07207   0.564 0.573102
## high_friends                  -0.03713    0.04941  -0.751 0.452405
## not_happy                      0.30668    0.08813   3.480 0.000506 ***
## angry_often                    0.24146    0.04685   5.154 2.64e-07 ***
## migration_direct               0.56213    0.05364  10.480  < 2e-16 ***
## migration_indirect             0.18130    0.08879   2.042 0.041203 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.586 on 5473 degrees of freedom
##   (3510 observations deleted due to missingness)
## Multiple R-squared:  0.1408, Adjusted R-squared:  0.1383
## F-statistic: 56.07 on 16 and 5473 DF,  p-value: < 2.2e-16
```

```
# added angriness


mod11 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + worried_often
  + sad_often + migration_direct + migration_indirect,
  data = new_data
)
summary(mod11)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + worried_often + sad_often + migration_direct +
##     migration_indirect, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9546 -1.1813 -0.3323  1.0390  5.4920
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.32255    0.72031   3.224  0.00127 **
## gender_female                 -0.31273    0.04690  -6.668 2.84e-11 ***
## age                            0.08550    0.01805   4.738 2.21e-06 ***
## education_low                  0.36009    0.07335   4.910 9.40e-07 ***
## education_high                -0.27643    0.05075  -5.447 5.34e-08 ***
## log(gross_labor_income)        0.95264    0.18868   5.049 4.58e-07 ***
## I(log(gross_labor_income)^2) -0.07814    0.01297  -6.026 1.79e-09 ***
## satisfaction_job              -0.07622    0.01209  -6.305 3.11e-10 ***
## life_satisfaction_general     -0.14176    0.01721  -8.239  < 2e-16 ***
## life_value_usefulness         -0.06871    0.01476  -4.654 3.33e-06 ***
## positive_attitude             -0.04675    0.01971  -2.372  0.01774 *
## low_friends                    0.05307    0.07200   0.737  0.46110
## high_friends                  -0.04197    0.04943  -0.849  0.39583
## worried_often                  0.13687    0.05226   2.619  0.00885 **
## sad_often                      0.20739    0.04934   4.204 2.67e-05 ***
## migration_direct               0.51306    0.05365   9.563  < 2e-16 ***
## migration_indirect             0.15547    0.08901   1.747  0.08075 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.587 on 5476 degrees of freedom
##   (3507 observations deleted due to missingness)
## Multiple R-squared:  0.1399, Adjusted R-squared:  0.1374
## F-statistic: 55.66 on 16 and 5476 DF,  p-value: < 2.2e-16
```

```r
# tested worriedness and sadness instead of happy and angry


mod12 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + health_not_good,
  data = new_data
)
summary(mod12)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + health_not_good,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9706 -1.1793 -0.3375  1.0468  5.4984
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.67101    0.71956   3.712 0.000208 ***
## gender_female                 -0.26870    0.04622  -5.813 6.48e-09 ***
## age                            0.08012    0.01856   4.318 1.61e-05 ***
## education_low                  0.34275    0.07358   4.658 3.26e-06 ***
## education_high                -0.26915    0.05097  -5.281 1.34e-07 ***
## log(gross_labor_income)        0.93497    0.18934   4.938 8.12e-07 ***
## I(log(gross_labor_income)^2) -0.07734    0.01301  -5.942 2.98e-09 ***
## satisfaction_job              -0.08021    0.01212  -6.615 4.06e-11 ***
## life_satisfaction_general     -0.15358    0.01727  -8.894  < 2e-16 ***
## life_value_usefulness         -0.06832    0.01481  -4.612 4.09e-06 ***
## positive_attitude             -0.05935    0.01964  -3.022 0.002523 **
## low_friends                    0.04485    0.07227   0.621 0.534917
## high_friends                  -0.03584    0.04958  -0.723 0.469720
## migration_direct               0.55064    0.05344  10.304  < 2e-16 ***
## migration_indirect             0.18383    0.08911   2.063 0.039164 *
## health_not_good                0.10325    0.04912   2.102 0.035583 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.592 on 5477 degrees of freedom
```

```
##   (3507 observations deleted due to missingness)
## Multiple R-squared:  0.1344, Adjusted R-squared:  0.132
## F-statistic: 56.69 on 15 and 5477 DF,  p-value: < 2.2e-16
```

```r
# testing of the health status instead of the feelings variables


mod13 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + one_or_two_children + more_than_two_children,
  data = new_data
)
summary(mod13)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + one_or_two_children +
##     more_than_two_children, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9973 -1.1721 -0.3422  1.0658  5.3687
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.84157    0.71805   3.957 7.68e-05 ***
## gender_female                 -0.28217    0.04627  -6.098 1.15e-09 ***
## age                            0.05748    0.02078   2.766 0.005688 **
## education_low                  0.36160    0.07349   4.920 8.89e-07 ***
## education_high                -0.26348    0.05085  -5.182 2.28e-07 ***
## log(gross_labor_income)        0.93785    0.18895   4.964 7.13e-07 ***
## I(log(gross_labor_income)^2) -0.07800    0.01299  -6.007 2.01e-09 ***
## satisfaction_job              -0.08108    0.01207  -6.720 2.00e-11 ***
## life_satisfaction_general     -0.16314    0.01688  -9.664  < 2e-16 ***
## life_value_usefulness         -0.07168    0.01481  -4.840 1.33e-06 ***
## positive_attitude             -0.06434    0.01947  -3.305 0.000956 ***
## low_friends                    0.04275    0.07217   0.592 0.553629
## high_friends                  -0.03378    0.04953  -0.682 0.495299
## migration_direct               0.54099    0.05337  10.137  < 2e-16 ***
## migration_indirect             0.18199    0.08898   2.045 0.040867 *
## one_or_two_children            0.23044    0.05581   4.129 3.70e-05 ***
## more_than_two_children         0.05540    0.07081   0.782 0.433992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.589 on 5482 degrees of freedom
```

```
##   (3501 observations deleted due to missingness)
## Multiple R-squared:  0.1379, Adjusted R-squared:  0.1354
## F-statistic: 54.83 on 16 and 5482 DF,  p-value: < 2.2e-16
```

```r
# testing of number of children instead of health status


mod14 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + strong_political_interest + no_political_interest,
  data = new_data
)
summary(mod14)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + strong_political_interest +
##     no_political_interest, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1603 -1.1645 -0.3237  1.0315  5.4884
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.96537    0.72429   4.094 4.30e-05 ***
## gender_female                 -0.29330    0.04677  -6.271 3.87e-10 ***
## age                            0.10148    0.01825   5.560 2.83e-08 ***
## education_low                  0.30834    0.07400   4.167 3.13e-05 ***
## education_high                -0.23405    0.05176  -4.522 6.25e-06 ***
## log(gross_labor_income)        0.87743    0.19059   4.604 4.24e-06 ***
## I(log(gross_labor_income)^2) -0.07335    0.01310  -5.599 2.26e-08 ***
## satisfaction_job              -0.08181    0.01207  -6.777 1.35e-11 ***
## life_satisfaction_general     -0.16091    0.01690  -9.520  < 2e-16 ***
## life_value_usefulness         -0.06882    0.01480  -4.651 3.38e-06 ***
## positive_attitude             -0.06696    0.01948  -3.437 0.000593 ***
## low_friends                    0.03060    0.07230   0.423 0.672078
## high_friends                  -0.02389    0.04955  -0.482 0.629719
## migration_direct               0.50311    0.05421   9.280  < 2e-16 ***
## migration_indirect             0.17142    0.08898   1.926 0.054104 .
## strong_political_interest     -0.12272    0.04923  -2.493 0.012700 *
## no_political_interest          0.21325    0.06533   3.264 0.001104 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.589 on 5475 degrees of freedom
```

```
##   (3508 observations deleted due to missingness)
## Multiple R-squared:  0.1381, Adjusted R-squared:  0.1356
## F-statistic: 54.85 on 16 and 5475 DF,  p-value: < 2.2e-16
```

```r
# testing of political interest instead of number of children


mod15 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + economy_worried + economy_not_worried
  + social_great_concern + social_no_concern,
  data = new_data
)
summary(mod15)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + economy_worried +
##     economy_not_worried + social_great_concern + social_no_concern,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1403 -1.1695 -0.3182  1.0249  5.5140
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.40806    0.71388   3.373 0.000748 ***
## gender_female                 -0.26051    0.04587  -5.679 1.43e-08 ***
## age                            0.09773    0.01805   5.414 6.42e-08 ***
## education_low                  0.33837    0.07378   4.586 4.62e-06 ***
## education_high                -0.21956    0.05081  -4.322 1.58e-05 ***
## log(gross_labor_income)        0.92699    0.18750   4.944 7.88e-07 ***
## I(log(gross_labor_income)^2) -0.07531    0.01289  -5.844 5.40e-09 ***
## satisfaction_job              -0.07176    0.01210  -5.930 3.21e-09 ***
## life_satisfaction_general     -0.13002    0.01714  -7.586 3.87e-14 ***
## life_value_usefulness         -0.06267    0.01473  -4.256 2.12e-05 ***
## positive_attitude             -0.06836    0.01938  -3.527 0.000424 ***
## low_friends                    0.03534    0.07185   0.492 0.622825
## high_friends                  -0.03148    0.04930  -0.639 0.523118
## migration_direct               0.44940    0.05450   8.246  < 2e-16 ***
## migration_indirect             0.16880    0.08847   1.908 0.056433 .
## economy_worried                0.29410    0.07098   4.143 3.47e-05 ***
## economy_not_worried           -0.36311    0.04780  -7.596 3.57e-14 ***
## social_great_concern           0.06748    0.05120   1.318 0.187575
## social_no_concern              0.13353    0.06057   2.205 0.027518 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.576 on 5440 degrees of freedom
##   (3541 observations deleted due to missingness)
## Multiple R-squared:  0.1494, Adjusted R-squared:  0.1466
## F-statistic: 53.09 on 18 and 5440 DF,  p-value: < 2.2e-16
```

```r
# testing of concern for own economic situation and social cohesion concern
# instead of political interest. Concern for economic situation seems important
# but for social cohesion not.


# Test for remaining independent variables relating to Job and Income
mod16 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + job_prestige_siops,
  data = new_data
)
summary(mod16)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + job_prestige_siops,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0461 -1.1345 -0.3416  1.0005  5.5250
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   4.025203   0.744545   5.406 6.73e-08 ***
## gender_female                -0.263376   0.047558  -5.538 3.22e-08 ***
## age                           0.076570   0.018707   4.093 4.32e-05 ***
## education_low                 0.291447   0.080250   3.632 0.000284 ***
## education_high               -0.134451   0.055467  -2.424 0.015386 *
## log(gross_labor_income)       0.720613   0.193920   3.716 0.000205 ***
## I(log(gross_labor_income)^2) -0.059584   0.013401  -4.446 8.93e-06 ***
## satisfaction_job             -0.086470   0.012500  -6.917 5.17e-12 ***
## life_satisfaction_general    -0.157914   0.017362  -9.095  < 2e-16 ***
## life_value_usefulness        -0.067970   0.015347  -4.429 9.68e-06 ***
## positive_attitude            -0.068361   0.020078  -3.405 0.000667 ***
## low_friends                   0.032348   0.074093   0.437 0.662426
## high_friends                 -0.055563   0.050716  -1.096 0.273326
## migration_direct              0.472077   0.058129   8.121 5.75e-16 ***
```

```
## migration_indirect            0.195027   0.090652   2.151 0.031493 *
## job_prestige_siops           -0.013543   0.002044  -6.626 3.81e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.56 on 5033 degrees of freedom
##   (3951 observations deleted due to missingness)
## Multiple R-squared:  0.1485, Adjusted R-squared:  0.146
## F-statistic: 58.52 on 15 and 5033 DF,  p-value: < 2.2e-16
```

```r
# Testing the values for job prestige


mod17 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + satisfaction_income,
  data = new_data
)
summary(mod17)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + satisfaction_income,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0732 -1.1583 -0.3487  1.0293  5.4846
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    3.06938    0.71551   4.290 1.82e-05 ***
## gender_female                 -0.24731    0.04597  -5.380 7.74e-08 ***
## age                            0.08626    0.01800   4.793 1.69e-06 ***
## education_low                  0.34785    0.07322   4.751 2.08e-06 ***
## education_high                -0.26057    0.05062  -5.147 2.73e-07 ***
## log(gross_labor_income)        0.82990    0.18870   4.398 1.11e-05 ***
## I(log(gross_labor_income)^2) -0.06485    0.01304  -4.974 6.74e-07 ***
## satisfaction_job              -0.04553    0.01276  -3.569 0.000362 ***
## life_satisfaction_general     -0.13730    0.01708  -8.039 1.10e-15 ***
## life_value_usefulness         -0.06390    0.01473  -4.338 1.46e-05 ***
## positive_attitude             -0.06062    0.01941  -3.124 0.001793 **
## low_friends                    0.04606    0.07185   0.641 0.521541
## high_friends                  -0.03135    0.04928  -0.636 0.524639
## migration_direct               0.51679    0.05334   9.688  < 2e-16 ***
## migration_indirect             0.16589    0.08859   1.872 0.061193 .
```

```
## satisfaction_income          -0.10610    0.01268  -8.369  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.582 on 5473 degrees of freedom
##   (3511 observations deleted due to missingness)
## Multiple R-squared:  0.1452, Adjusted R-squared:  0.1429
## F-statistic: 61.99 on 15 and 5473 DF,  p-value: < 2.2e-16
```

```r
# Testing the variables for income satisfaction


mod18 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + exp_fulltime_years + exp_parttime_years
  + exp_unemployment_years,
  data = new_data
)
summary(mod18)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + exp_fulltime_years +
##     exp_parttime_years + exp_unemployment_years, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0787 -1.1802 -0.3324  1.0383  5.4432
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   2.741995   0.724081   3.787 0.000154 ***
## gender_female                -0.185842   0.050525  -3.678 0.000237 ***
## age                           0.031980   0.043252   0.739 0.459698
## education_low                 0.316259   0.073961   4.276 1.94e-05 ***
## education_high               -0.244975   0.052658  -4.652 3.36e-06 ***
## log(gross_labor_income)       0.960035   0.189232   5.073 4.04e-07 ***
## I(log(gross_labor_income)^2) -0.079180   0.013062  -6.062 1.44e-09 ***
## satisfaction_job             -0.083704   0.012077  -6.931 4.66e-12 ***
## life_satisfaction_general    -0.153274   0.016926  -9.056  < 2e-16 ***
## life_value_usefulness        -0.069470   0.014766  -4.705 2.61e-06 ***
## positive_attitude            -0.067393   0.019498  -3.456 0.000552 ***
## low_friends                   0.034274   0.072175   0.475 0.634891
## high_friends                 -0.033996   0.049526  -0.686 0.492477
## migration_direct              0.536250   0.053683   9.989  < 2e-16 ***
## migration_indirect            0.183894   0.088818   2.070 0.038456 *
```

```
## exp_fulltime_years            0.007371   0.004382   1.682 0.092594 .
## exp_parttime_years           -0.004386   0.005533  -0.793 0.428035
## exp_unemployment_years        0.051153   0.009752   5.246 1.62e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.587 on 5457 degrees of freedom
##   (3525 observations deleted due to missingness)
## Multiple R-squared:  0.1405, Adjusted R-squared:  0.1378
## F-statistic: 52.45 on 17 and 5457 DF,  p-value: < 2.2e-16
```

```r
# Testing the experience of years spent in full time, part time or unemployment


mod19 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + satisfaction_job + life_satisfaction_general + life_value_usefulness
  + positive_attitude + low_friends + high_friends + migration_direct
  + migration_indirect + bad_feeling_overall,
  data = new_data
)
summary(mod19)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + satisfaction_job + life_satisfaction_general +
##     life_value_usefulness + positive_attitude + low_friends +
##     high_friends + migration_direct + migration_indirect + bad_feeling_overall,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9669 -1.1896 -0.3299  1.0593  5.4166
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.51062    0.72017   3.486 0.000494 ***
## gender_female                 -0.26915    0.04612  -5.836 5.65e-09 ***
## age                            0.08497    0.01809   4.698 2.69e-06 ***
## education_low                  0.35132    0.07342   4.785 1.76e-06 ***
## education_high                -0.26878    0.05083  -5.288 1.28e-07 ***
## log(gross_labor_income)        0.96982    0.18910   5.129 3.02e-07 ***
## I(log(gross_labor_income)^2) -0.07972    0.01300  -6.135 9.13e-10 ***
## satisfaction_job              -0.08252    0.01207  -6.838 8.88e-12 ***
## life_satisfaction_general     -0.14722    0.01729  -8.515  < 2e-16 ***
## life_value_usefulness         -0.06824    0.01479  -4.615 4.02e-06 ***
## positive_attitude             -0.05863    0.01953  -3.001 0.002699 **
## low_friends                    0.03881    0.07217   0.538 0.590743
## high_friends                  -0.04044    0.04949  -0.817 0.413920
## migration_direct               0.54355    0.05337  10.184  < 2e-16 ***
```

```
## migration_indirect            0.18022    0.08899   2.025 0.042899 *
## bad_feeling_overall           0.44005    0.11294   3.896 9.88e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.59 on 5483 degrees of freedom
##   (3501 observations deleted due to missingness)
## Multiple R-squared:  0.137,  Adjusted R-squared:  0.1347
## F-statistic: 58.05 on 15 and 5483 DF,  p-value: < 2.2e-16
```

```r
# testing for a new dummy variable called "bad_feeling_overall"
# this tests the effect of experiencing bad feelings overall
# (being sad, angry and not happy at the same time)


mod20 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + high_life_satisfaction + migration_direct + migration_indirect
  + bad_feeling_overall + economy_worried
  + economy_not_worried + one_or_two_children + more_than_two_children,
  data = new_data
)
summary(mod20)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + high_life_satisfaction + migration_direct +
##     migration_indirect + bad_feeling_overall + economy_worried +
##     economy_not_worried + one_or_two_children + more_than_two_children,
##     data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4857 -1.2066 -0.3029  1.0975  5.3413
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   0.482290   0.689140   0.700 0.484054
## gender_female                -0.301386   0.045497  -6.624 3.82e-11 ***
## age                           0.070403   0.020160   3.492 0.000483 ***
## education_low                 0.390563   0.072889   5.358 8.74e-08 ***
## education_high               -0.191075   0.050542  -3.780 0.000158 ***
## log(gross_labor_income)       0.905186   0.184992   4.893 1.02e-06 ***
## I(log(gross_labor_income)^2) -0.076211   0.012762  -5.972 2.49e-09 ***
## high_life_satisfaction       -0.564779   0.048621 -11.616  < 2e-16 ***
## migration_direct              0.393254   0.053104   7.405 1.50e-13 ***
## migration_indirect            0.159574   0.088606   1.801 0.071765 .
## bad_feeling_overall           0.674959   0.107998   6.250 4.41e-10 ***
## economy_worried               0.416342   0.069578   5.984 2.31e-09 ***
## economy_not_worried          -0.398463   0.047191  -8.444  < 2e-16 ***
```

```
## one_or_two_children           0.179857   0.055231   3.256 0.001135 **
## more_than_two_children       -0.009365   0.069784  -0.134 0.893246
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.595 on 5616 degrees of freedom
##   (3369 observations deleted due to missingness)
## Multiple R-squared:  0.1331, Adjusted R-squared:  0.1309
## F-statistic: 61.57 on 14 and 5616 DF,  p-value: < 2.2e-16
```

```r
# using the new dummy variable for extremely high life satisfaction
# instead raw values (high_life_satisfaction)


mod21 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2)
  + high_life_satisfaction + migration_direct + migration_indirect
  + bad_feeling_overall + economy_worried
  + economy_not_worried + one_or_two_children + more_than_two_children
  + satisfaction_income + satisfaction_job + life_value_usefulness
  + log(gross_labor_income),
  data = new_data
)
summary(mod21)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + high_life_satisfaction + migration_direct +
##     migration_indirect + bad_feeling_overall + economy_worried +
##     economy_not_worried + one_or_two_children + more_than_two_children +
##     satisfaction_income + satisfaction_job + life_value_usefulness +
##     log(gross_labor_income), data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9545 -1.1398 -0.3143  1.0162  5.2957
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    1.93950    0.71019   2.731 0.006335 **
## gender_female                 -0.26650    0.04573  -5.828 5.93e-09 ***
## age                            0.05924    0.02053   2.885 0.003927 **
## education_low                  0.38989    0.07304   5.338 9.76e-08 ***
## education_high                -0.19725    0.05040  -3.914 9.19e-05 ***
## log(gross_labor_income)        0.82607    0.18774   4.400 1.10e-05 ***
## I(log(gross_labor_income)^2) -0.06550    0.01296  -5.052 4.52e-07 ***
## high_life_satisfaction       -0.34755    0.05161  -6.734 1.82e-11 ***
## migration_direct              0.41657    0.05320   7.830 5.83e-15 ***
## migration_indirect            0.15445    0.08818   1.752 0.079913 .
## bad_feeling_overall           0.52110    0.10919   4.772 1.87e-06 ***
```

```
## economy_worried               0.28366    0.07022   4.039 5.43e-05 ***
## economy_not_worried          -0.28561    0.04803  -5.946 2.92e-09 ***
## one_or_two_children           0.19234    0.05527   3.480 0.000506 ***
## more_than_two_children        0.02340    0.06993   0.335 0.737872
## satisfaction_income          -0.08290    0.01289  -6.432 1.36e-10 ***
## satisfaction_job             -0.05038    0.01254  -4.018 5.96e-05 ***
## life_value_usefulness        -0.08270    0.01392  -5.941 3.00e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.572 on 5469 degrees of freedom
##   (3513 observations deleted due to missingness)
## Multiple R-squared:  0.1562, Adjusted R-squared:  0.1536
## F-statistic: 59.56 on 17 and 5469 DF,  p-value: < 2.2e-16
```

```
# using the new dummy variable for extremely high life satisfaction
# instead raw values (high_life_satisfaction)
# testing for extremely high life_satisfaction and comparing the effect
# of other satisfaction variables

# checking for multicollinearity
# -> by this test we can see that we cannot test for income and labor
# force status at the same time because they are multicollinear
sum(new_data$income_per_1000[new_data$non_working == 1], na.rm = TRUE)
```

```
## [1] 0
```

```
mod22 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + log(gross_labor_income)
  + I(log(gross_labor_income)^2) + high_life_satisfaction
  + migration_direct + migration_indirect + bad_feeling_overall
  + economy_worried + economy_not_worried + one_or_two_children
  + more_than_two_children + log(gross_labor_income)
  + satisfaction_income + satisfaction_job + job_prestige_siops
  + positive_attitude + life_value_usefulness + exp_fulltime_years
  + exp_unemployment_years + married_together + married_separate
  + worried_often + no_political_interest + strong_political_interest,
  data = new_data
)
summary(mod22)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + log(gross_labor_income) +
##     I(log(gross_labor_income)^2) + high_life_satisfaction + migration_direct +
##     migration_indirect + bad_feeling_overall + economy_worried +
##     economy_not_worried + one_or_two_children + more_than_two_children +
##     log(gross_labor_income) + satisfaction_income + satisfaction_job +
##     job_prestige_siops + positive_attitude + life_value_usefulness +
##     exp_fulltime_years + exp_unemployment_years + married_together +
##     married_separate + worried_often + no_political_interest +
```

```
##     strong_political_interest, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8558 -1.0901 -0.3202  0.9650  5.4234
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    3.402153   0.744539   4.569 5.01e-06 ***
## gender_female                 -0.271210   0.053092  -5.108 3.37e-07 ***
## age                           -0.006115   0.033128  -0.185 0.853565
## education_low                  0.267698   0.080757   3.315 0.000924 ***
## education_high                -0.049287   0.056562  -0.871 0.383589
## log(gross_labor_income)        0.589746   0.193328   3.050 0.002297 **
## I(log(gross_labor_income)^2) -0.046847   0.013479  -3.476 0.000514 ***
## high_life_satisfaction        -0.333947   0.053670  -6.222 5.31e-10 ***
## migration_direct               0.365038   0.059831   6.101 1.13e-09 ***
## migration_indirect             0.154698   0.089475   1.729 0.083880 .
## bad_feeling_overall            0.427623   0.113214   3.777 0.000161 ***
## economy_worried                0.218016   0.073632   2.961 0.003082 **
## economy_not_worried           -0.232732   0.049621  -4.690 2.80e-06 ***
## one_or_two_children            0.235396   0.061820   3.808 0.000142 ***
## more_than_two_children         0.096090   0.077357   1.242 0.214238
## satisfaction_income           -0.090866   0.013404  -6.779 1.35e-11 ***
## satisfaction_job              -0.048649   0.013043  -3.730 0.000194 ***
## job_prestige_siops            -0.010037   0.002057  -4.880 1.09e-06 ***
## positive_attitude             -0.058456   0.019870  -2.942 0.003277 **
## life_value_usefulness         -0.064492   0.015045  -4.287 1.85e-05 ***
## exp_fulltime_years             0.007501   0.003176   2.362 0.018210 *
## exp_unemployment_years         0.039018   0.009660   4.039 5.44e-05 ***
## married_together              -0.057338   0.054157  -1.059 0.289769
## married_separate               0.012801   0.123877   0.103 0.917703
## worried_often                  0.135502   0.052438   2.584 0.009794 **
## no_political_interest          0.163622   0.068453   2.390 0.016872 *
## strong_political_interest     -0.069453   0.049593  -1.400 0.161439
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.53 on 4944 degrees of freedom
##   (4029 observations deleted due to missingness)
## Multiple R-squared:  0.1814, Adjusted R-squared:  0.1771
## F-statistic: 42.13 on 26 and 4944 DF,  p-value: < 2.2e-16
```

```r
# For this one I just wanted test as many variables as possible
# to see if there are any effects on the model if I try to include anything


mod23 <- lm(
  personal_achievement_deserved ~ gender_female + age + education_low
  + education_high + high_life_satisfaction + migration_direct
  + migration_indirect + bad_feeling_overall
  + economy_worried + economy_not_worried + one_or_two_children
  + log(gross_labor_income)
  + I(log(gross_labor_income)^2) + satisfaction_income + satisfaction_job
```

```r
  + positive_attitude + life_value_usefulness + exp_unemployment_years
  + worried_often + no_political_interest + strong_political_interest,
  data = new_data
)
summary(mod23)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + high_life_satisfaction +
##     migration_direct + migration_indirect + bad_feeling_overall +
##     economy_worried + economy_not_worried + one_or_two_children +
##     log(gross_labor_income) + I(log(gross_labor_income)^2) +
##     satisfaction_income + satisfaction_job + positive_attitude +
##     life_value_usefulness + exp_unemployment_years + worried_often +
##     no_political_interest + strong_political_interest, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6839 -1.1320 -0.3219  1.0016  5.4225
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   2.228711   0.716222   3.112 0.001869 **
## gender_female                -0.309838   0.046945  -6.600 4.50e-11 ***
## age                           0.065324   0.019031   3.432 0.000603 ***
## education_low                 0.319969   0.073901   4.330 1.52e-05 ***
## education_high               -0.181522   0.051291  -3.539 0.000405 ***
## high_life_satisfaction       -0.304601   0.052134  -5.843 5.44e-09 ***
## migration_direct              0.387489   0.054582   7.099 1.42e-12 ***
## migration_indirect            0.127611   0.087849   1.453 0.146388
## bad_feeling_overall           0.423971   0.110973   3.820 0.000135 ***
## economy_worried               0.259947   0.070681   3.678 0.000238 ***
## economy_not_worried          -0.260004   0.048174  -5.397 7.06e-08 ***
## one_or_two_children           0.178175   0.045102   3.951 7.90e-05 ***
## log(gross_labor_income)       0.755143   0.188359   4.009 6.18e-05 ***
## I(log(gross_labor_income)^2) -0.058831   0.013020  -4.519 6.36e-06 ***
## satisfaction_income          -0.081597   0.012877  -6.337 2.54e-10 ***
## satisfaction_job             -0.049387   0.012554  -3.934 8.46e-05 ***
## positive_attitude            -0.056358   0.019277  -2.924 0.003475 **
## life_value_usefulness        -0.068803   0.014379  -4.785 1.76e-06 ***
## exp_unemployment_years        0.038540   0.009204   4.187 2.87e-05 ***
## worried_often                 0.135655   0.050717   2.675 0.007501 **
## no_political_interest         0.190320   0.064767   2.939 0.003312 **
## strong_political_interest    -0.104211   0.048603  -2.144 0.032068 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.562 on 5420 degrees of freedom
##   (3558 observations deleted due to missingness)
## Multiple R-squared:  0.166,  Adjusted R-squared:  0.1628
## F-statistic: 51.38 on 21 and 5420 DF,  p-value: < 2.2e-16
```

```r
# This is a test where I tried to include every variable that seemed
# SOMEWHAT important so far


robustnesstest <- glm(
  achievement_deserved ~ gender_female + age + education_low
  + education_high + high_life_satisfaction + migration_direct
  + migration_indirect + bad_feeling_overall
  + economy_worried + economy_not_worried + one_or_two_children
  + more_than_two_children + log(gross_labor_income)
  + log(gross_labor_income) + I(log(gross_labor_income)^2)
  + satisfaction_income + satisfaction_job + job_prestige_siops
  + positive_attitude + life_value_usefulness + exp_fulltime_years
  + exp_unemployment_years + married_together + married_separate
  + worried_often + no_political_interest + strong_political_interest,
  family = "binomial",
  data = new_data
)
summary(robustnesstest)
```

```
##
## Call:
## glm(formula = achievement_deserved ~ gender_female + age + education_low +
##     education_high + high_life_satisfaction + migration_direct +
##     migration_indirect + bad_feeling_overall + economy_worried +
##     economy_not_worried + one_or_two_children + more_than_two_children +
##     log(gross_labor_income) + log(gross_labor_income) + I(log(gross_labor_income)^2) +
##     satisfaction_income + satisfaction_job + job_prestige_siops +
##     positive_attitude + life_value_usefulness + exp_fulltime_years +
##     exp_unemployment_years + married_together + married_separate +
##     worried_often + no_political_interest + strong_political_interest,
##     family = "binomial", data = new_data)
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   3.498067   1.696019   2.063 0.039159 *
## gender_female                 0.352007   0.079687   4.417 9.99e-06 ***
## age                          -0.026898   0.049744  -0.541 0.588701
## education_low                -0.354159   0.114437  -3.095 0.001969 **
## education_high                0.118624   0.086163   1.377 0.168593
## high_life_satisfaction        0.388659   0.084182   4.617 3.90e-06 ***
## migration_direct             -0.522648   0.085803  -6.091 1.12e-09 ***
## migration_indirect           -0.238208   0.130368  -1.827 0.067671 .
## bad_feeling_overall          -0.349405   0.160366  -2.179 0.029347 *
## economy_worried              -0.269677   0.102766  -2.624 0.008686 **
## economy_not_worried           0.317506   0.074581   4.257 2.07e-05 ***
## one_or_two_children          -0.234612   0.092329  -2.541 0.011052 *
## more_than_two_children       -0.084867   0.116541  -0.728 0.466483
## log(gross_labor_income)      -1.657582   0.454082  -3.650 0.000262 ***
## I(log(gross_labor_income)^2)  0.123897   0.031343   3.953 7.72e-05 ***
## satisfaction_income           0.097450   0.019372   5.031 4.89e-07 ***
## satisfaction_job              0.048563   0.018860   2.575 0.010027 *
## job_prestige_siops            0.012630   0.003061   4.126 3.70e-05 ***
```

```
## positive_attitude             0.053975   0.028811   1.873 0.061011 .
## life_value_usefulness         0.095788   0.021830   4.388 1.14e-05 ***
## exp_fulltime_years           -0.008011   0.004760  -1.683 0.092398 .
## exp_unemployment_years       -0.037520   0.014155  -2.651 0.008036 **
## married_together              0.183560   0.080034   2.294 0.021817 *
## married_separate              0.017873   0.179009   0.100 0.920469
## worried_often                -0.138247   0.076397  -1.810 0.070360 .
## no_political_interest        -0.242750   0.096849  -2.506 0.012194 *
## strong_political_interest     0.097110   0.074377   1.306 0.191670
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 6408.2  on 4970  degrees of freedom
## Residual deviance: 5641.1  on 4944  degrees of freedom
##   (4029 observations deleted due to missingness)
## AIC: 5695.1
##
## Number of Fisher Scoring iterations: 5
```

```
exp(coef(robustnesstest))
```

```
##                   (Intercept)                 gender_female
##                    33.0515026                     1.4219181
##                           age                 education_low
##                     0.9734609                     0.7017637
##                education_high       high_life_satisfaction
##                     1.1259468                     1.4750008
##             migration_direct           migration_indirect
##                     0.5929483                     0.7880388
##          bad_feeling_overall              economy_worried
##                     0.7051076                     0.7636263
##          economy_not_worried          one_or_two_children
##                     1.3736971                     0.7908777
##       more_than_two_children      log(gross_labor_income)
##                     0.9186348                     0.1905992
## I(log(gross_labor_income)^2)          satisfaction_income
##                     1.1318994                     1.1023566
##              satisfaction_job           job_prestige_siops
##                     1.0497615                     1.0127098
##            positive_attitude         life_value_usefulness
##                     1.0554584                     1.1005257
##           exp_fulltime_years       exp_unemployment_years
##                     0.9920209                     0.9631753
##             married_together              married_separate
##                     1.2014866                     1.0180334
##                worried_often         no_political_interest
##                     0.8708837                     0.7844677
##    strong_political_interest
##                     1.1019821
```

```r
# This is a robustness test with logistic regression
# It tests if the variables would also be relevant if we used logistic
# regression instead.
# It seems like the chosen variables stay relevant / significant (robust)



# ============================================================== #
#                   END EXPERIMENTATION MODELLING
# ============================================================== #


# ============================================================== #
# ============================================================== #
#                   FINAL MODEL (MOD 20) - STRUCTURED
# ============================================================== #
# ============================================================== #

mod29 <- lm(
  formula = personal_achievement_deserved ~

    # --- Demographics and Background ---
    gender_female
  + age
  + education_low
  + education_high
  + migration_direct
  + migration_indirect
  + one_or_two_children
  + more_than_two_children  #New

  # --- Financial and Employment ---
  + log(gross_labor_income)
  + I(log(gross_labor_income)^2) #New
  + exp_unemployment_years


  # --- Satisfaction and Attitudes ---
  + satisfaction_income
  + satisfaction_job
  + high_life_satisfaction
  + life_value_usefulness
  + positive_attitude
  + bad_feeling_overall
  + worried_often

  # --- Worries and Interests ---
  + economy_worried
  + economy_not_worried
  + no_political_interest
  + strong_political_interest,

  data = new_data
```

```
)

summary(mod29)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ gender_female +
##     age + education_low + education_high + migration_direct +
##     migration_indirect + one_or_two_children + more_than_two_children +
##     log(gross_labor_income) + I(log(gross_labor_income)^2) +
##     exp_unemployment_years + satisfaction_income + satisfaction_job +
##     high_life_satisfaction + life_value_usefulness + positive_attitude +
##     bad_feeling_overall + worried_often + economy_worried + economy_not_worried +
##     no_political_interest + strong_political_interest, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6810 -1.1300 -0.3202  1.0046  5.4242
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    2.235829   0.716902   3.119 0.001826 **
## gender_female                 -0.310227   0.046977  -6.604 4.39e-11 ***
## age                            0.063243   0.020925   3.022 0.002520 **
## education_low                  0.319444   0.073939   4.320 1.59e-05 ***
## education_high                -0.181548   0.051296  -3.539 0.000405 ***
## migration_direct               0.387246   0.054597   7.093 1.48e-12 ***
## migration_indirect             0.128144   0.087885   1.458 0.144876
## one_or_two_children            0.185808   0.055238   3.364 0.000774 ***
## more_than_two_children         0.016718   0.069845   0.239 0.810832
## log(gross_labor_income)        0.754485   0.188396   4.005 6.29e-05 ***
## I(log(gross_labor_income)^2) -0.058797   0.013022  -4.515 6.46e-06 ***
## exp_unemployment_years         0.038485   0.009208   4.180 2.97e-05 ***
## satisfaction_income           -0.081538   0.012881  -6.330 2.64e-10 ***
## satisfaction_job              -0.049370   0.012556  -3.932 8.53e-05 ***
## high_life_satisfaction        -0.305001   0.052165  -5.847 5.30e-09 ***
## life_value_usefulness         -0.069016   0.014408  -4.790 1.71e-06 ***
## positive_attitude             -0.056399   0.019280  -2.925 0.003455 **
## bad_feeling_overall            0.423500   0.111000   3.815 0.000138 ***
## worried_often                  0.135702   0.050722   2.675 0.007486 **
## economy_worried                0.259627   0.070700   3.672 0.000243 ***
## economy_not_worried           -0.259729   0.048192  -5.389 7.37e-08 ***
## no_political_interest          0.190110   0.064779   2.935 0.003352 **
## strong_political_interest     -0.104060   0.048612  -2.141 0.032347 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.562 on 5419 degrees of freedom
##   (3558 observations deleted due to missingness)
## Multiple R-squared:  0.166,  Adjusted R-squared:  0.1626
## F-statistic: 49.04 on 22 and 5419 DF,  p-value: < 2.2e-16
```

```r
mod30 <- lm(
  formula = personal_achievement_deserved ~

    # --- Demographics and Background ---
    #gender_female
    #+ age
    + education_low
  #+ education_high
  #+ migration_direct
  #+ migration_indirect
  + one_or_two_children

  # --- Financial and Employment ---
  #+ log(gross_labor_income)
  + exp_unemployment_years

  # --- Satisfaction and Attitudes ---
  + satisfaction_income
  + satisfaction_job
  #+ high_life_satisfaction
  + life_value_usefulness
  #+ positive_attitude
  #+ bad_feeling_overall
  #+ worried_often

  # --- Worries and Interests ---
  #+ economy_worried
  + economy_not_worried
  + no_political_interest
  + strong_political_interest,

  data = new_data
)

summary(mod30)
```

```
##
## Call:
## lm(formula = personal_achievement_deserved ~ +education_low +
##     one_or_two_children + exp_unemployment_years + satisfaction_income +
##     satisfaction_job + life_value_usefulness + economy_not_worried +
##     no_political_interest + strong_political_interest, data = new_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7292 -1.1924 -0.3337  1.0597  5.2306
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                4.974483   0.117814  42.223  < 2e-16 ***
## education_low              0.461643   0.070371   6.560 5.86e-11 ***
## one_or_two_children        0.171495   0.043147   3.975 7.14e-05 ***
## exp_unemployment_years     0.053617   0.008991   5.964 2.62e-09 ***
```

```
## satisfaction_income       -0.116940   0.011859  -9.861  < 2e-16 ***
## satisfaction_job          -0.060246   0.012405  -4.857 1.23e-06 ***
## life_value_usefulness     -0.100852   0.013487  -7.478 8.74e-14 ***
## economy_not_worried       -0.424710   0.046336  -9.166  < 2e-16 ***
## no_political_interest      0.236382   0.064321   3.675  0.00024 ***
## strong_political_interest -0.142471   0.046876  -3.039  0.00238 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.593 on 5551 degrees of freedom
##   (3439 observations deleted due to missingness)
## Multiple R-squared:  0.131,  Adjusted R-squared:  0.1296
## F-statistic: 92.95 on 9 and 5551 DF,  p-value: < 2.2e-16
```

```
# testing of multicollinearity of models
# test any model by replacing mod_x
# might need to install packages when using these functions
library(car)
```

```
## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##     recode
```

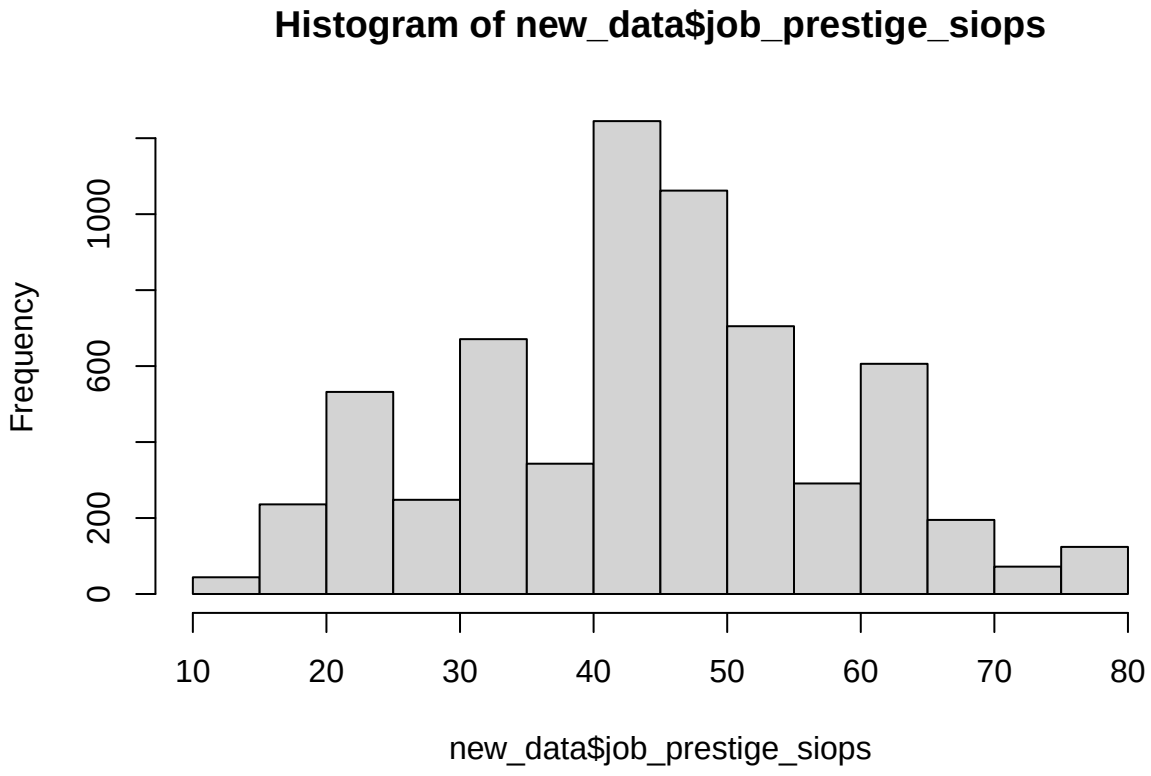
```
# We removed vif(mod20) to focus on the final model validation

vif(mod29) #Final Model
```

```
##                gender_female                         age
##                     1.229829                    1.608630
##                education_low               education_high
##                     1.213028                    1.273931
##             migration_direct           migration_indirect
##                     1.203103                    1.071952
##          one_or_two_children       more_than_two_children
##                     1.700325                    1.594236
##      log(gross_labor_income) I(log(gross_labor_income)^2)
##                    60.822645                   63.606753
##       exp_unemployment_years          satisfaction_income
##                     1.091963                    1.594818
##              satisfaction_job       high_life_satisfaction
##                     1.386169                    1.220180
##        life_value_usefulness            positive_attitude
##                     1.349908                    1.287931
##          bad_feeling_overall                worried_often
##                     1.102584                    1.172824
##              economy_worried          economy_not_worried
##                     1.172010                    1.243985
##        no_political_interest    strong_political_interest
##                     1.230597                    1.277906
```

```r
# ============================================================ #
#                   END FINAL MODEL STRUCTURED
# ============================================================ #


# ============================================================ #
# ============================================================ #
#                       VISUALIZATIONS AND GRAPHS
# ============================================================ #
# ============================================================ #

# BELOW HERE ARE SOME EXPERIMENTATIONS TO VISUALIZE AND GET MORE
# FAMILIARIZED WITH THE DATA

mean(new_data$personal_achievement_deserved)
```

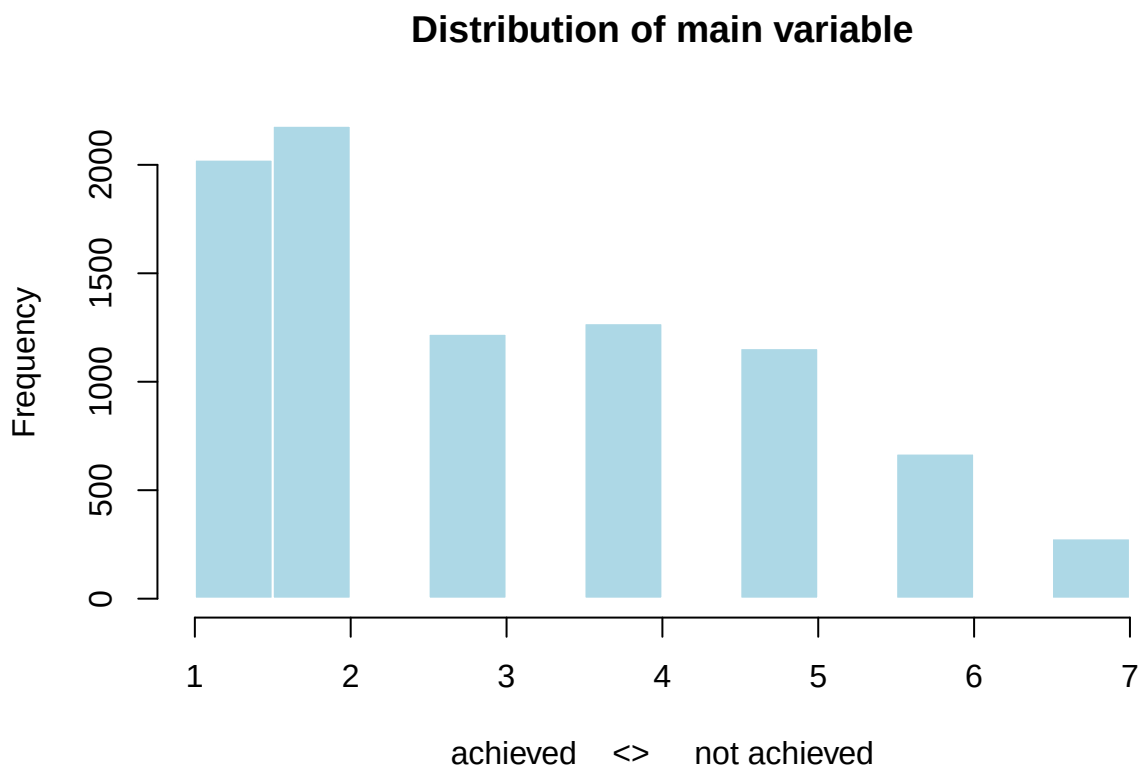
```
## [1] NA
```

```r
hist(new_data$job_prestige_siops)
```

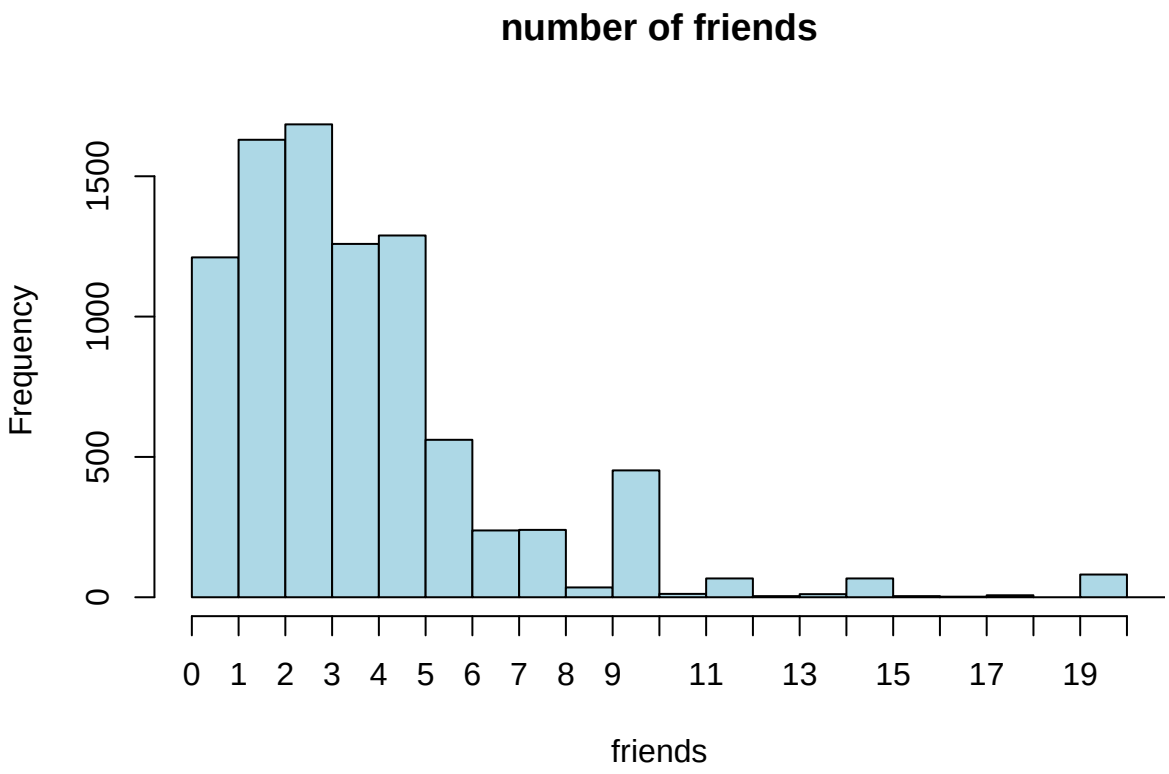



```r
# visualization of main variable
hist(new_data$personal_achievement_deserved,
     main = "Distribution of main variable",
     xlab = "achieved    <>     not achieved",
```

```
     col = "lightblue",
     border = "white")
```




```
hist(new_data$number_close_friends, main = "number of friends",
     xlab = "friends"
     , col = "lightblue", breaks = seq(0 , 100, by = 1)
     , xlim = c(0,20), xaxt = "n"
)
axis(1, at = seq(0, 20, by = 1))
```

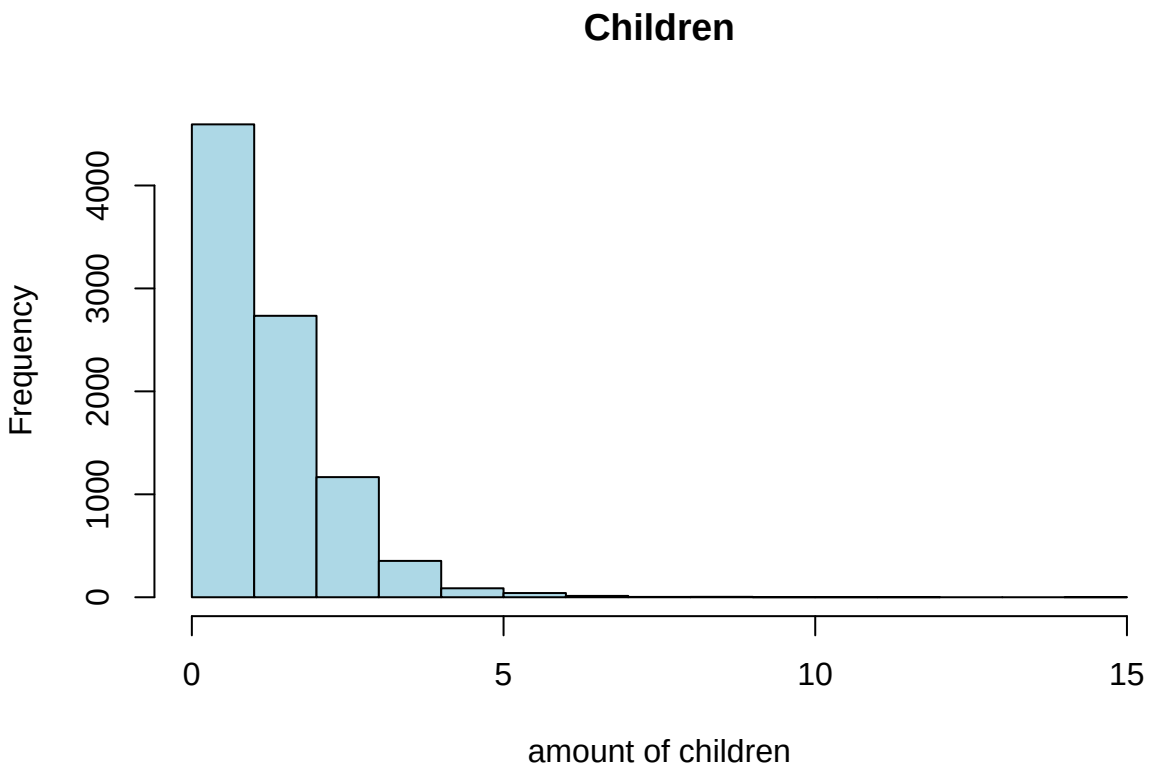



```
median(new_data$number_close_friends, na.rm = TRUE)
```

```
## [1] 3
```

```
hist(new_data$feeling_happy, main = "happy", xlab = "happiness",
     col = "lightblue", breaks = seq(0.5, 5.5, by = 1))
axis(1, at = seq(1,6, by = 1))
```


```
hist(new_data$total_children, main = "Children", xlab = "amount of children",
     col = "lightblue", breaks = seq(0, max(new_data$total_children), by = 1))
```




```
# This checks how the gross labor income behaves (in this case from
# 2,000 to 6,000 units, an increase of 200%)
old <- log(2000)
new <- log(6000)
effect <- 0.754485 * (new - old) + -0.058797 * (new^2 - old^2)
effect
```

```
## [1] -0.2240407
```

```
# ============================================================ #
#                      END VISUALIZATIONS AND GRAPHS
# ============================================================ #


# ============================================================ #
# ============================================================ #
#                          COMPARING MODELS
# ============================================================ #
# ============================================================ #

install.packages("stargazer")
```

```
## The following package(s) will be installed:
## - stargazer [5.2.3]
## These packages will be installed into "C:/Users/alexm/OneDrive/Escritorio/Master DBM/Semestre 1/Data
```

```
##
## # Installing packages --------------------------------------------------------
## - Installing stargazer ...                      OK [linked from cache]
## Successfully installed 1 package in 14 milliseconds.
```

```
library(stargazer)
```

```
##
## Please cite as:

##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
stargazer(mod23, mod29, mod30,
          type = "latex",
          header = FALSE,
          font.size = "footnotesize",
          title = "Evolution of Models",
          dep.var.labels = "Achievement Gap (High = Worse)",
          omit = c("Constant") # We remove the renaming list to avoid errors
)
```

```
#================-

# 1. Frequency Table (How many Men vs Women)
gender_distribution <- new_data %>%
  count(gender_code = gender) %>%
  mutate(
    label = case_when(
      gender_code == 1 ~ "Male",
      gender_code == 2 ~ "Female",
      TRUE ~ "Unknown/Missing"
    )
  )

# View the result in a table
View(gender_distribution)

# 2. Base R Table (Quick check)
table(new_data$gender, useNA = "ifany")
```

1 2 3 4352 4645 3

```
# ================================================================ #
#                          TEST TO VALIDATE DUMMIES
# ================================================================ #


table(Original = new_data$education_level,
      Dummy = new_data$education_low, useNA = "ifany")
```

Table 1: Evolution of Models

	Dependent variable:		
	Achievement Gap (High = Worse)		
	(1)	(2)	(3)
gender_female	−0.310***	−0.310***	
	(0.047)	(0.047)	
age	0.065***	0.063***	
	(0.019)	(0.021)	
education_low	0.320***	0.319***	0.462***
	(0.074)	(0.074)	(0.070)
education_high	−0.182***	−0.182***	
	(0.051)	(0.051)	
high_life_satisfaction	−0.305***	−0.305***	
	(0.052)	(0.052)	
migration_direct	0.387***	0.387***	
	(0.055)	(0.055)	
migration_indirect	0.128	0.128	
	(0.088)	(0.088)	
bad_feeling_overall	0.424***	0.424***	
	(0.111)	(0.111)	
economy_worried	0.260***	0.260***	
	(0.071)	(0.071)	
economy_not_worried	−0.260***	−0.260***	−0.425***
	(0.048)	(0.048)	(0.046)
one_or_two_children	0.178***	0.186***	0.171***
	(0.045)	(0.055)	(0.043)
more_than_two_children		0.017	
		(0.070)	
log(gross_labor_income)	0.755***	0.754***	
	(0.188)	(0.188)	
I(log(gross_labor_income)^2)	−0.059***	−0.059***	
	(0.013)	(0.013)	
satisfaction_income	−0.082***	−0.082***	−0.117***
	(0.013)	(0.013)	(0.012)
satisfaction_job	−0.049***	−0.049***	−0.060***
	(0.013)	(0.013)	(0.012)
positive_attitude	−0.056***	−0.056***	
	(0.019)	(0.019)	
life_value_usefulness	−0.069***	−0.069***	−0.101***
	(0.014)	(0.014)	(0.013)
exp_unemployment_years	0.039***	0.038***	0.054***
	(0.009)	(0.009)	(0.009)
worried_often	0.136***	0.136***	
	(0.051)	(0.051)	
no_political_interest	0.190***	0.190***	0.236***
	(0.065)	(0.065)	(0.064)
strong_political_interest	−0.104**	−0.104**	−0.142***
	(0.049)	(0.049)	(0.047)
Observations	5,442	5,442	5,561
R^2	0.166	0.166	0.131
Adjusted R^2	0.163	0.163	0.130

Dummy

Original 0 1 0 0 0 11 1 0 385 0 2 0 959 0 3 3786 0 0 4 829 0 0 5 399 0 0 6 1702 0 0 7 711 0 0 8 110 0 0 0 0 0 108

```
# New table
validation_education <- new_data %>%
  count(original = education_level,
        nuevo_dummy = education_low)

View(validation_education)

# ================================================================ #
#                          END PROJECT
# ================================================================ #
```