

# MRP\_analysis.R

n8than

2024-06-19

```
# multilevel regression and post-stratification analysis

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

# # load regression data
# {
#   data <-
#     haven::read_dta(
#       "C:/Users/Nathan/Documents/Newham Fellowship/data/Skills for Life Survey 2011/UKDA-7240-tab/sta
#
#   save(data, file = "data/skills_for_life_data.RData")
# }

load(here::here("data/skills_for_life_data.RData"))

# select variables

data <-
  data |>
  select(
    WORKINGSTATUS2,
    GROSS_ANNUAL_INCOME_OLDBANDS,
    BUK,
    Sex1,
    AGE1NET,
    Sesol, # is English first language
    ETHNICSIMPLE,
    HIQUAL,
    CLITSPEAK, # ENFL everyday English skills (literacy and speaking)
    IMDSCOREB4, # Index of Multiple Deprivation banded into deciles
    NSSEC7,
    # outcomes
```

```

SUMMARYCOMP,      # self-assessed computer skills (summary)
TSKILLA,          # self-assessed computer skills (summary 2)
COMBLIT,          # self-assessed reading a writing (summary)
starts_with("LiteracyScore"),      # literacy level
starts_with("LiteracyThreshold"),  # literacy threshold
starts_with("NumeracyScore"),      # literacy level
starts_with("NumeracyThreshold"),  # literacy threshold,
MultipleChoiceLevelA_1,            # ICT level
MultipleChoiceLevelA_1Thres,       # ICT threshold
# weights
rimweight2003,
rimweightLIT2003,
rimweightNUM2003,
rimweightICT2003,
rimweightNUMICT2003,
rimweightLITICT2003,
rimweightLITNUM2003
)

# these are the S4L variables used in
# Rowlands (2015) British Journal of General Practice
#
# job status: National Statistics Socioeconomic Classification 3 bands (Managerial/professional, Intermed
# employment status: employed, not employed
# gross income: >=10000, <10000
# place of birth: UK, non UK
# home ownership: Owns or part-owns home, does not own home
# sex: male, female
# age: 16-44, >=45
# first language: English, other
# ethnicity: white, black and minority ethnic
# qualification level: NQF >= level at age 16 (level 2), below level 2
# area deprivation: IMD quintiles

# matching with survey

# WORKINGSTATUS2: 0-No, 1-Yes
# GROSS_ANNUAL_INCOME_OLDBANDS: {<£5,000, £5,000 - £9,999}, {£10,000 - £14,999, £15,000 - £19,999, £20,
# BUK: 1-Yes, 2-No
# QxTenu1: 1-Own home outright or with a mortgage or loan
# Sex1: 1-Male, 2-Female
# AGE1NET: {16-24, 25-44}, 45-65
# Sesol: 1-Yes, 2-No
# ETHNICSIMPLE: 1-White, 2-BME
# HIQUAL: {1-4}, {5-Level 1 qualification or below}
# IMDSCOREB4: 1,...,9
# NSSEC7: 1 Higher managerial and professional
# 2 Lower managerial and professional
# 3 Intermediate
# 4 Small employers and own account workers
# 5 Lower supervisory and technical
# 6 Semi-routine occupations

```

```
      # 7  Routine occupations
      # 8  Never worked/ long term unemployed
      # 9  Full-time student
      # 10 Not classifiable

#####
# data cleaning

model_dat <-
  data
```