

data_w4

Celina Jang

2025-07-01

Week 4

```
chsw3 <- read.csv('chs2020_working_w3.csv')

chsw3 <- chsw3 |> mutate(
  # change binary variable values to 1 and 0
  delaypayrent = case_when(
    delaypayrent == 1 ~ 0,
    delaypayrent == 2 ~ 1,
    is.na(delaypayrent) ~ NA),
  didntgetcare0 = case_when(
    didntgetcare20 == 1 ~ 0,
    didntgetcare20 == 2 ~ 1,
    is.na(didntgetcare20) ~ NA),
  nspd = case_when(
    nspd == 1 ~ 1,
    nspd == 2 ~ 0,
    is.na(nspd) ~ NA
  ),
  rodentsstreet0 = case_when(
    rodentsstreet==1 ~ 1,
    rodentsstreet == 2 ~ 0,
    is.na(rodentsstreet) ~ NA
  ),
  # want reference group to be white, so reorder
  race_ethnicity = fct_relevel(race_ethnicity, 'White'),
  # label employment20
  employment = case_when(
    employment20 == 1 ~ 'Employed for wages or salary',
    employment20 == 2 ~ 'Self-employed',
    employment20 == 3 ~ 'Unemployed for 1 year or more',
    employment20 == 4 ~ 'Unemployed for less than 1 year',
    employment20 == 5 ~ 'A homemaker',
    employment20 == 6 ~ 'A student',
    employment20 == 7 ~ 'Retired',
    employment20 == 8 ~ 'Unable to work',
    employment20 == '.d' ~ 'Dont know',
    employment20 == '.r' ~ 'Refused',
    is.na(employment20) ~ NA
  ),
  education1 = case_when(
```

```

    education == 1 ~ 'Less than high school',
    education == 2 ~ 'High school graduate',
    education == 3 ~ 'Some college/technical school',
    education == 4 ~ 'College graduate',
    education == '.d' ~ 'Dont know',
    education == '.r' ~ 'Refused',
    is.na(education) ~ NA
  )
)

# linear regression (UNWEIGHTED)
k6.fit.lm <- lm(k6 ~ social_cohesion_rev, data=chsw3)
summary(k6.fit.lm)

##
## Call:
## lm(formula = k6 ~ social_cohesion_rev, data = chsw3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.415  -3.359  -1.124   1.994  20.463
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.00226    0.27653   21.705  <2e-16 ***
## social_cohesion_rev -0.58701    0.09206   -6.376   2e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.309 on 4328 degrees of freedom
## (6 observations deleted due to missingness)
## Multiple R-squared:  0.009307, Adjusted R-squared:  0.009078
## F-statistic: 40.66 on 1 and 4328 DF, p-value: 2.004e-10

# logistic regression (UNWEIGHTED)
nspd.fit.lg <- glm(nspd ~ social_cohesion_rev, data=chsw3)
summary(nspd.fit.lg)

##
## Call:
## glm(formula = nspd ~ social_cohesion_rev, data = chsw3)
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.878275    0.015276  122.956  < 2e-16 ***
## social_cohesion_rev 0.020976    0.005086   4.125 3.78e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.05665044)
##
##      Null deviance: 246.15  on 4329  degrees of freedom
## Residual deviance: 245.18  on 4328  degrees of freedom
## (6 observations deleted due to missingness)

```

```
## AIC: -138.8
##
## Number of Fisher Scoring iterations: 2
exp(coef(summary(nspd.fit.lg))[2, "Estimate"])

## [1] 1.021198
# social cohesion by age
sc_age_lm <- lm(social_cohesion_rev ~ age_band, data = chsw3)
anova(sc_age_lm)

## Analysis of Variance Table
##
## Response: social_cohesion_rev
##          Df Sum Sq Mean Sq F value    Pr(>F)
## age_band    2   71.5   35.751   73.001 < 2.2e-16 ***
## Residuals 4318 2114.7    0.490
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Weighted analysis

```
library(survey)

## Loading required package: grid
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
## Loading required package: survival
##
## Attaching package: 'survey'
## The following object is masked from 'package:graphics':
##
##     dotchart
# Setting the weights
chs2020_svy <- svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1, data = chsw3)

# multiple regression with continuous kessler (mental distress)
svy_lm_k6 <- svyglm(k6 ~ social_cohesion_rev + age_band + gender + race_ethnicity +
  education1 + employment + delaypayrent0 + rodentsstreet0,
  design = chs2020_svy)
summary(svy_lm_k6)

##
## Call:
## svyglm(formula = k6 ~ social_cohesion_rev + age_band + gender +
##       race_ethnicity + education1 + employment + delaypayrent0 +
```

```

##      rodentsstreet0, design = chs2020_svy)
##
## Survey design:
## svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1,
##      data = chsw3)
##
## Coefficients:
##
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.471012   0.689371   9.387 < 2e-16
## social_cohesion_rev -0.350410   0.140451  -2.495 0.012639
## age_band45-64      -0.921701   0.215514  -4.277 1.94e-05
## age_band65+       -1.679147   0.354379  -4.738 2.23e-06
## gendermale        -0.787687   0.192956  -4.082 4.55e-05
## race_ethnicityAsian/Pacific Islander -0.716561   0.320823  -2.234 0.025569
## race_ethnicityBlack -0.942921   0.274082  -3.440 0.000587
## race_ethnicityHispanic -0.421392   0.268242  -1.571 0.116274
## race_ethnicityNorth African/Mid Eastern  1.335607   0.917007   1.456 0.145336
## race_ethnicityOther  0.974199   0.761504   1.279 0.200862
## education1High school graduate  0.006541   0.241291   0.027 0.978375
## education1Less than high school  0.022603   0.319113   0.071 0.943536
## education1Some college/technical school -0.100518   0.274388  -0.366 0.714135
## employmentA student  1.791193   0.631136   2.838 0.004562
## employmentEmployed for wages or salary  0.977580   0.402940   2.426 0.015304
## employmentRetired    2.127199   0.509292   4.177 3.02e-05
## employmentSelf-employed  1.063422   0.495941   2.144 0.032072
## employmentUnable to work  3.860980   0.619358   6.234 5.02e-10
## employmentUnemployed for 1 year or more  1.486199   0.571774   2.599 0.009376
## employmentUnemployed for less than 1 year  2.060827   0.500317   4.119 3.88e-05
## delaypayrent0      -1.929357   0.334318  -5.771 8.47e-09
## rodentsstreet0      0.928649   0.215214   4.315 1.63e-05
##
## (Intercept)      ***
## social_cohesion_rev      *
## age_band45-64      ***
## age_band65+      ***
## gendermale      ***
## race_ethnicityAsian/Pacific Islander      *
## race_ethnicityBlack      ***
## race_ethnicityHispanic
## race_ethnicityNorth African/Mid Eastern
## race_ethnicityOther
## education1High school graduate
## education1Less than high school
## education1Some college/technical school
## employmentA student      **
## employmentEmployed for wages or salary      *
## employmentRetired      ***
## employmentSelf-employed      *
## employmentUnable to work      ***
## employmentUnemployed for 1 year or more      **
## employmentUnemployed for less than 1 year ***
## delaypayrent0      ***
## rodentsstreet0      ***
## ---

```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 16.41754)
##
## Number of Fisher Scoring iterations: 2
# multiple regression with binary psychological distress
svy_glm_nspd <- svyglm(nspd0 ~ social_cohesion_rev + age_band + gender + race_ethnicity +
  education1 + employment + delaypayrent0 + rodentsstreet0,
  design = chs2020_svy, family = quasibinomial())
summary(svy_glm_nspd)

##
## Call:
## svyglm(formula = nspd0 ~ social_cohesion_rev + age_band + gender +
##       race_ethnicity + education1 + employment + delaypayrent0 +
##       rodentsstreet0, design = chs2020_svy, family = quasibinomial())
##
## Survey design:
## svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1,
##       data = chsw3)
##
## Coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -2.587447   0.716080  -3.613 0.000306
## social_cohesion_rev -0.124995   0.122962  -1.017 0.309439
## age_band45-64      -0.239272   0.236335  -1.012 0.311394
## age_band65+        -0.620763   0.380791  -1.630 0.103138
## gendermale         -0.481078   0.195672  -2.459 0.013989
## race_ethnicityAsian/Pacific Islander  0.067596   0.377982   0.179 0.858077
## race_ethnicityBlack -0.349864   0.324695  -1.078 0.281314
## race_ethnicityHispanic -0.004964   0.278507  -0.018 0.985781
## race_ethnicityNorth African/Mid Eastern  0.992585   0.723843   1.371 0.170366
## race_ethnicityOther  0.609464   0.513335   1.187 0.235193
## education1High school graduate  0.050312   0.271944   0.185 0.853232
## education1Less than high school -0.167753   0.324481  -0.517 0.605193
## education1Some college/technical school  0.362423   0.284499   1.274 0.202773
## employmentA student  1.362664   0.694819   1.961 0.049926
## employmentEmployed for wages or salary  0.618147   0.581578   1.063 0.287900
## employmentRetired    1.553928   0.665058   2.337 0.019512
## employmentSelf-employed  0.921639   0.673757   1.368 0.171416
## employmentUnable to work  2.483018   0.597693   4.154 3.33e-05
## employmentUnemployed for 1 year or more  0.742959   0.724051   1.026 0.304899
## employmentUnemployed for less than 1 year 1.515084   0.595268   2.545 0.010958
## delaypayrent0       -1.004009   0.240040  -4.183 2.94e-05
## rodentsstreet0       0.486171   0.212599   2.287 0.022259
##
## (Intercept)      ***
## social_cohesion_rev
## age_band45-64
## age_band65+
## gendermale      *
## race_ethnicityAsian/Pacific Islander
## race_ethnicityBlack
## race_ethnicityHispanic
```

```

## race_ethnicityNorth African/Mid Eastern
## race_ethnicityOther
## education1High school graduate
## education1Less than high school
## education1Some college/technical school
## employmentA student *
## employmentEmployed for wages or salary
## employmentRetired *
## employmentSelf-employed
## employmentUnable to work ***
## employmentUnemployed for 1 year or more
## employmentUnemployed for less than 1 year *
## delaypayrent0 ***
## rodentsstreet0 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 0.9394312)
##
## Number of Fisher Scoring iterations: 6
# interaction term (cohesion x age)
svy_lm_k6_int <- svyglm(k6 ~ social_cohesion_rev * age_band + gender + race_ethnicity +
                        education1 + employment + delaypayrent0 + rodentsstreet0,
                        design = chs2020_svy)
summary(svy_lm_k6_int)

##
## Call:
## svyglm(formula = k6 ~ social_cohesion_rev * age_band + gender +
##       race_ethnicity + education1 + employment + delaypayrent0 +
##       rodentsstreet0, design = chs2020_svy)
##
## Survey design:
## svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1,
##       data = chsw3)
##
## Coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.277245    0.781653   8.031 1.26e-15
## social_cohesion_rev -0.280025    0.199217  -1.406 0.15991
## age_band45-64     -0.956931    0.987983  -0.969 0.33282
## age_band65+      -0.131098    1.233000  -0.106 0.91533
## gendermale       -0.783195    0.193057  -4.057 5.07e-05
## race_ethnicityAsian/Pacific Islander -0.720325    0.321799  -2.238 0.02525
## race_ethnicityBlack -0.921722    0.275811  -3.342 0.00084
## race_ethnicityHispanic -0.408059    0.266977  -1.528 0.12648
## race_ethnicityNorth African/Mid Eastern  1.349429    0.921071   1.465 0.14298
## race_ethnicityOther  0.983402    0.760645   1.293 0.19614
## education1High school graduate  0.002682    0.241161   0.011 0.99113
## education1Less than high school -0.002304    0.318458  -0.007 0.99423
## education1Some college/technical school -0.102928    0.274090  -0.376 0.70729
## employmentA student  1.788636    0.633860   2.822 0.00480
## employmentEmployed for wages or salary  0.970859    0.402582   2.412 0.01593
## employmentRetired    2.114945    0.510148   4.146 3.46e-05

```

```

## employmentSelf-employed          1.051154    0.496438    2.117    0.03429
## employmentUnable to work          3.857169    0.622348    6.198    6.30e-10
## employmentUnemployed for 1 year or more  1.471283    0.574307    2.562    0.01045
## employmentUnemployed for less than 1 year 2.056278    0.500590    4.108    4.07e-05
## delaypayrent0                      -1.935061    0.333691   -5.799    7.18e-09
## rodentsstreet0                     0.934522    0.215108    4.344    1.43e-05
## social_cohesion_rev:age_band45-64      0.006561    0.323438    0.020    0.98382
## social_cohesion_rev:age_band65+       -0.498521    0.380646   -1.310    0.19038
##
## (Intercept)                        ***
## social_cohesion_rev
## age_band45-64
## age_band65+
## gendermale                         ***
## race_ethnicityAsian/Pacific Islander    *
## race_ethnicityBlack                   ***
## race_ethnicityHispanic
## race_ethnicityNorth African/Mid Eastern
## race_ethnicityOther
## education1High school graduate
## education1Less than high school
## education1Some college/technical school
## employmentA student                  **
## employmentEmployed for wages or salary  *
## employmentRetired                    ***
## employmentSelf-employed              *
## employmentUnable to work             ***
## employmentUnemployed for 1 year or more *
## employmentUnemployed for less than 1 year ***
## delaypayrent0                       ***
## rodentsstreet0                      ***
## social_cohesion_rev:age_band45-64
## social_cohesion_rev:age_band65+
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 16.40213)
##
## Number of Fisher Scoring iterations: 2
# nspd and social cohesion log reg model
summary(svyglm(nspd0 ~ social_cohesion_rev ,design=chs2020_svy, family=binomial))

## Warning in eval(family$initialize): non-integer #successes in a binomial glm!
##
## Call:
## svyglm(formula = nspd0 ~ social_cohesion_rev, design = chs2020_svy,
##         family = binomial)
##
## Survey design:
## svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1,
##          data = chsw3)
##
## Coefficients:

```

```
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -1.8235      0.3655  -4.989 6.33e-07 ***
## social_cohesion_rev -0.2863      0.1284  -2.229  0.0259 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 0.9988666)
##
## Number of Fisher Scoring iterations: 5
exp(-0.2863)

## [1] 0.7510373
# k6 and social cohesion linear reg model
summary(svyglm(k6 ~ social_cohesion_rev ,design=chs2020_svy))

##
## Call:
## svyglm(formula = k6 ~ social_cohesion_rev, design = chs2020_svy)
##
## Survey design:
## svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1,
##      data = chsw3)
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)         6.0627      0.4386   13.82 < 2e-16 ***
## social_cohesion_rev -0.5925      0.1445   -4.10 4.22e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 18.58422)
##
## Number of Fisher Scoring iterations: 2
```

Visualizing regression results

```
# tidy the model
tidy_model <- broom::tidy(svy_lm_k6)

# filter for significant variables (p < 0.05)
signif_vars <- tidy_model |>
  filter(p.value < 0.05) |>
  mutate(across(where(is.numeric), ~ round(., 3))) |>
  rename(
    Variable = term,
    Estimate = estimate,
    "Std. Error" = std.error,
    "p-value" = p.value,
    "t value" = statistic
```



```

) |>
select(Variable, Estimate, 'Std. Error', 't value', "p-value")

# create the table
signif_vars |>
  kable(caption = "Significant Predictors of K6 (Survey-Weighted Linear Regression)", escape=T)

```

Table 1: Significant Predictors of K6 (Survey-Weighted Linear Regression)

Variable	Estimate	Std. Error	t value	p-value
(Intercept)	6.471	0.689	9.387	0.000
social_cohesion_rev	-0.350	0.140	-2.495	0.013
age_band45-64	-0.922	0.216	-4.277	0.000
age_band65+	-1.679	0.354	-4.738	0.000
gendermale	-0.788	0.193	-4.082	0.000
race_ethnicityAsian/Pacific Islander	-0.717	0.321	-2.234	0.026
race_ethnicityBlack	-0.943	0.274	-3.440	0.001
employmentA student	1.791	0.631	2.838	0.005
employmentEmployed for wages or salary	0.978	0.403	2.426	0.015
employmentRetired	2.127	0.509	4.177	0.000
employmentSelf-employed	1.063	0.496	2.144	0.032
employmentUnable to work	3.861	0.619	6.234	0.000
employmentUnemployed for 1 year or more	1.486	0.572	2.599	0.009
employmentUnemployed for less than 1 year	2.061	0.500	4.119	0.000
delaypayrent0	-1.929	0.334	-5.771	0.000
rodentsstreet0	0.929	0.215	4.315	0.000

```

# Step 1: Tidy the model
log_table <- tidy(svy_glm_nspd)

# Step 2: Filter significant variables (before modifying p-value format)
significant_terms <- log_table %>%
  filter(p.value < 0.05 & term != "(Intercept)")

# Step 3: Add formatting (after filtering)
significant_terms <- significant_terms %>%
  mutate(
    "p-value" = ifelse(p.value < 1e-4, "<0.0001", round(p.value, 4)),
    "Odds Ratio" = round(exp(estimate), 3),
    Estimate = round(estimate, 3),
    "Std. Error" = round(std.error, 3)
  ) %>%
  select(Term = term, Estimate, "Std. Error", "Odds Ratio", "p-value")

```

```
# Display full table or significant terms only
significant_terms %>%
  kable(caption = "Significant Predictors of Psychological Distress (nspd)",
        escape = T, align = "lcccc")
```

Table 2: Significant Predictors of Psychological Distress (nspd)

Term	Estimate	Std. Error	Odds Ratio	p-value
gendermale	-0.481	0.196	0.618	0.014
employmentA student	1.363	0.695	3.907	0.0499
employmentRetired	1.554	0.665	4.730	0.0195
employmentUnable to work	2.483	0.598	11.977	<0.0001
employmentUnemployed for less than 1 year	1.515	0.595	4.550	0.011
delaypayrent0	-1.004	0.240	0.366	<0.0001
rodentsstreet0	0.486	0.213	1.626	0.0223

```
# tidy table
reg_table <- tidy(svy_lm_k6_int)

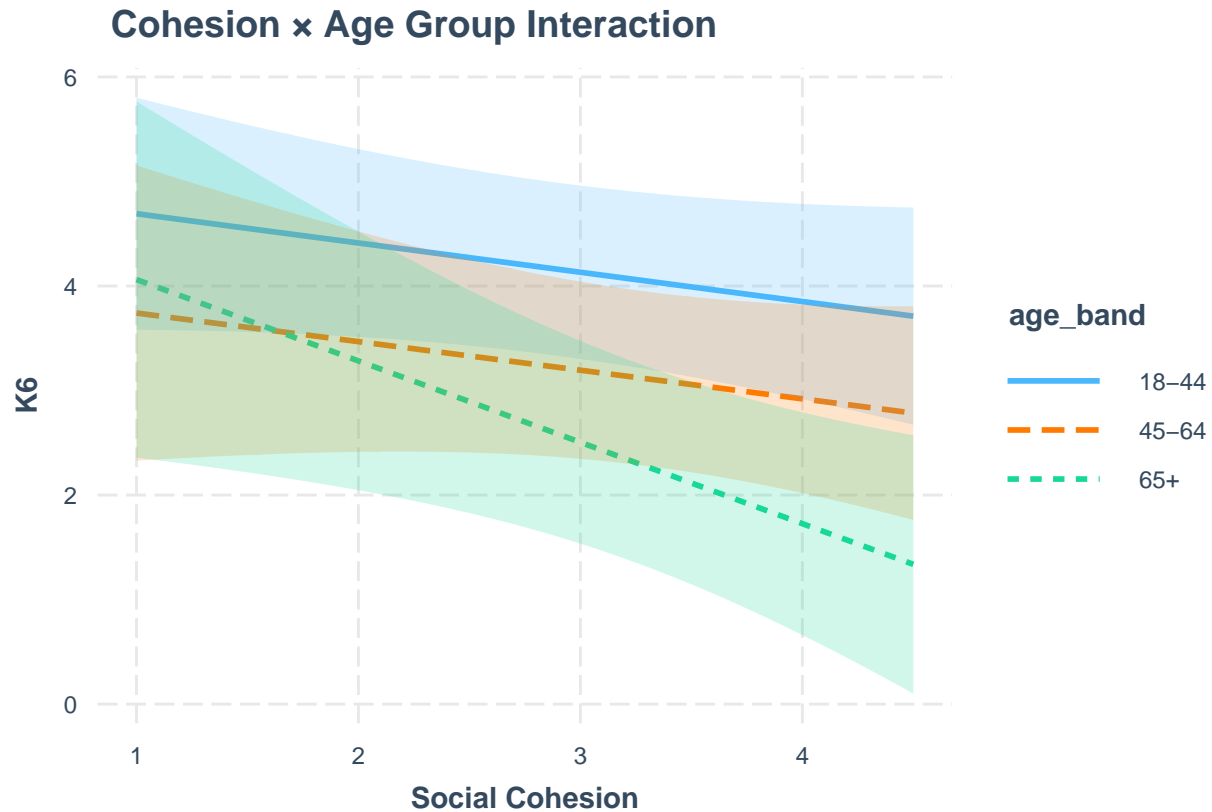
# select significant variables
reg_table_signif <- reg_table |>
  filter(p.value < 0.05) |>
  select(
    Term = term,
    Estimate = estimate,
    "Std. Error" = std.error,
    "t value" = statistic,
    "p-value" = p.value
  ) |>
  mutate(across(where(is.numeric), ~ round(., 3)))

# put into table
reg_table_signif |>
  kable(caption = "Significant Predictors in Cohesion X Age", align = "lcccc")
```

Table 3: Significant Predictors in Cohesion X Age

Term	Estimate	Std. Error	t value	p-value
(Intercept)	6.277	0.782	8.031	0.000
gendermale	-0.783	0.193	-4.057	0.000
race_ethnicityAsian/Pacific Islander	-0.720	0.322	-2.238	0.025
race_ethnicityBlack	-0.922	0.276	-3.342	0.001
employmentA student	1.789	0.634	2.822	0.005
employmentEmployed for wages or salary	0.971	0.403	2.412	0.016
employmentRetired	2.115	0.510	4.146	0.000
employmentSelf-employed	1.051	0.496	2.117	0.034
employmentUnable to work	3.857	0.622	6.198	0.000
employmentUnemployed for 1 year or more	1.471	0.574	2.562	0.010
employmentUnemployed for less than 1 year	2.056	0.501	4.108	0.000
delaypayrent0	-1.935	0.334	-5.799	0.000
rodentsstreet0	0.935	0.215	4.344	0.000

```
library(interactions)
interact_plot(svy_lm_k6_int, pred = social_cohesion_rev, modx = age_band,
  plot.points = F, interval = TRUE, main.title = "Cohesion × Age Group Interaction",
  x.label='Social Cohesion', y.label = 'K6')
```



Checking Multicollinearity

```
library(car)

## Loading required package: carData
##
## Attaching package: 'car'
## The following object is masked from 'package:dplyr':
##
##   recode
## The following object is masked from 'package:purrr':
##
##   some

lm_k6 <- lm(k6 ~ social_cohesion_rev + age_band + gender + race_ethnicity +
  education1 + employment + delaypayrent0 + rodentsstreet0, data = chsw3)
vif(lm_k6)

##
##          GVIF Df GVIF^(1/(2*Df))
```

```
## social_cohesion_rev 1.065670 1 1.032313
## age_band 2.279699 2 1.228767
## gender 1.063521 1 1.031271
## race_ethnicity 1.459850 5 1.038558
## education1 1.422960 3 1.060552
## employment 2.653128 7 1.072182
## delaypayrent0 1.103273 1 1.050368
## rodentsstreet0 1.045885 1 1.022685
```

Separate models by age group

```
# 65+
svy_65plus <- subset(chs2020_svy, age_band == "65+")
model_65plus <- svyglm(k6 ~ social_cohesion_rev + gender + race_ethnicity +
  education1 + employment + delaypayrent0 + rodentsstreet0,
  design = svy_65plus)
summary(model_65plus)
```

```
##
## Call:
## svyglm(formula = k6 ~ social_cohesion_rev + gender + race_ethnicity +
##   education1 + employment + delaypayrent0 + rodentsstreet0,
##   design = svy_65plus)
##
## Survey design:
## subset(chs2020_svy, age_band == "65+")
##
## Coefficients:
##
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.33787 2.13737 3.433 0.000627
## social_cohesion_rev -0.74943 0.32870 -2.280 0.022873
## gendermale -1.44641 0.41048 -3.524 0.000450
## race_ethnicityAsian/Pacific Islander -1.44659 0.57469 -2.517 0.012026
## race_ethnicityBlack -0.78000 0.68576 -1.137 0.255709
## race_ethnicityHispanic -0.15437 0.66396 -0.233 0.816208
## race_ethnicityNorth African/Mid Eastern 4.31012 0.80649 5.344 1.19e-07
## race_ethnicityOther 1.55157 1.08245 1.433 0.152138
## education1High school graduate 0.06466 0.47871 0.135 0.892590
## education1Less than high school 0.97798 0.72205 1.354 0.175978
## education1Some college/technical school 0.60826 0.65159 0.933 0.350847
## employmentEmployed for wages or salary -0.45802 0.98378 -0.466 0.641653
## employmentRetired 0.97256 0.90687 1.072 0.283849
## employmentSelf-employed -0.29295 1.00129 -0.293 0.769926
## employmentUnable to work 2.69813 1.29174 2.089 0.037046
## employmentUnemployed for 1 year or more -0.15572 1.13005 -0.138 0.890434
## employmentUnemployed for less than 1 year 1.66380 1.63702 1.016 0.309766
## delaypayrent0 -2.31129 1.42770 -1.619 0.105867
## rodentsstreet0 1.16380 0.52327 2.224 0.026422
##
## (Intercept) ***
## social_cohesion_rev *
## gendermale ***
## race_ethnicityAsian/Pacific Islander *
```

```

## race_ethnicityBlack
## race_ethnicityHispanic
## race_ethnicityNorth African/Mid Eastern ***
## race_ethnicityOther
## education1High school graduate
## education1Less than high school
## education1Some college/technical school
## employmentEmployed for wages or salary
## employmentRetired
## employmentSelf-employed
## employmentUnable to work *
## employmentUnemployed for 1 year or more
## employmentUnemployed for less than 1 year
## delaypayrent0
## rodentsstreet0 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 14.7298)
##
## Number of Fisher Scoring iterations: 2
# 45-64
svy_45 <- subset(chs2020_svy, age_band == '45-64')
model_45 <- svyglm(k6 ~ social_cohesion_rev + gender + race_ethnicity +
                  education1 + employment + delaypayrent0 + rodentsstreet0,
                  design = svy_45)
summary(model_45)

##
## Call:
## svyglm(formula = k6 ~ social_cohesion_rev + gender + race_ethnicity +
##       education1 + employment + delaypayrent0 + rodentsstreet0,
##       design = svy_45)
##
## Survey design:
## subset(chs2020_svy, age_band == "45-64")
##
## Coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.1239    1.1412   5.366 9.53e-08
## social_cohesion_rev -0.1848    0.2440  -0.757  0.44904
## gendermale       -0.7224    0.2906  -2.486  0.01306
## race_ethnicityAsian/Pacific Islander -0.8099    0.4438  -1.825  0.06823
## race_ethnicityBlack -1.2111    0.3919  -3.090  0.00204
## race_ethnicityHispanic  0.1345    0.4259   0.316  0.75215
## race_ethnicityNorth African/Mid Eastern -0.8055    0.6262  -1.286  0.19854
## race_ethnicityOther   0.5500    0.9060   0.607  0.54393
## education1High school graduate -0.2600    0.3724  -0.698  0.48514
## education1Less than high school -0.2666    0.5005  -0.533  0.59432
## education1Some college/technical school -0.3896    0.4214  -0.925  0.35527
## employmentA student   4.9447    3.4218   1.445  0.14869
## employmentEmployed for wages or salary  0.3694    0.5560   0.664  0.50652
## employmentRetired     1.6683    0.7825   2.132  0.03320
## employmentSelf-employed  0.5271    0.6596   0.799  0.42434

```

```
## employmentUnable to work          4.4033      0.8447      5.213 2.16e-07
## employmentUnemployed for 1 year or more 0.7908      0.8149      0.970 0.33202
## employmentUnemployed for less than 1 year 1.7295      0.8031      2.154 0.03146
## delaypayrent0                      -2.3880      0.4997     -4.779 1.96e-06
## rodentsstreet0                     0.4112      0.3156      1.303 0.19275
##
## (Intercept)                        ***
## social_cohesion_rev
## gendermale                          *
## race_ethnicityAsian/Pacific Islander .
## race_ethnicityBlack                 **
## race_ethnicityHispanic
## race_ethnicityNorth African/Mid Eastern
## race_ethnicityOther
## education1High school graduate
## education1Less than high school
## education1Some college/technical school
## employmentA student
## employmentEmployed for wages or salary
## employmentRetired                   *
## employmentSelf-employed
## employmentUnable to work            ***
## employmentUnemployed for 1 year or more
## employmentUnemployed for less than 1 year *
## delaypayrent0                       ***
## rodentsstreet0
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for gaussian family taken to be 15.18208)
```

```
##
```

```
## Number of Fisher Scoring iterations: 2
```

```
# 18-44
```

```
svy_18 <- subset(chs2020_svy, age_band == '18-44')
model_18 <- svyglm(k6 ~ social_cohesion_rev + gender + race_ethnicity +
                    education1 + employment + delaypayrent0 + rodentsstreet0,
                    design = svy_18)
summary(model_18)
```

```
##
```

```
## Call:
```

```
## svyglm(formula = k6 ~ social_cohesion_rev + gender + race_ethnicity +
##       education1 + employment + delaypayrent0 + rodentsstreet0,
##       design = svy_18)
```

```
##
```

```
## Survey design:
```

```
## subset(chs2020_svy, age_band == "18-44")
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.87659    0.95122      6.178 8.09e-10
## social_cohesion_rev -0.28217    0.20091     -1.404 0.16037
## gendermale       -0.53882    0.29943     -1.800 0.07211
## race_ethnicityAsian/Pacific Islander -0.63037    0.47250     -1.334 0.18235
```

```

## race_ethnicityBlack -0.85949 0.41570 -2.068 0.03883
## race_ethnicityHispanic -0.65271 0.38370 -1.701 0.08911
## race_ethnicityNorth African/Mid Eastern 2.05529 1.25611 1.636 0.10197
## race_ethnicityOther 0.78499 1.36839 0.574 0.56627
## education1High school graduate 0.13888 0.36426 0.381 0.70306
## education1Less than high school -0.27972 0.47281 -0.592 0.55419
## education1Some college/technical school -0.03644 0.39404 -0.092 0.92634
## employmentA student 1.51389 0.73816 2.051 0.04043
## employmentEmployed for wages or salary 1.04165 0.57505 1.811 0.07025
## employmentRetired 4.15099 2.34877 1.767 0.07736
## employmentSelf-employed 1.09707 0.76903 1.427 0.15389
## employmentUnable to work 1.92808 1.02607 1.879 0.06040
## employmentUnemployed for 1 year or more 1.82963 0.88200 2.074 0.03819
## employmentUnemployed for less than 1 year 1.95515 0.66434 2.943 0.00329
## delaypayrent0 -1.58564 0.38824 -4.084 4.63e-05
## rodentsstreet0 1.12029 0.30258 3.702 0.00022
##
## (Intercept) ***
## social_cohesion_rev
## gendermale .
## race_ethnicityAsian/Pacific Islander
## race_ethnicityBlack *
## race_ethnicityHispanic .
## race_ethnicityNorth African/Mid Eastern
## race_ethnicityOther
## education1High school graduate
## education1Less than high school
## education1Some college/technical school
## employmentA student *
## employmentEmployed for wages or salary .
## employmentRetired .
## employmentSelf-employed
## employmentUnable to work .
## employmentUnemployed for 1 year or more *
## employmentUnemployed for less than 1 year **
## delaypayrent0 ***
## rodentsstreet0 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 16.97771)
##
## Number of Fisher Scoring iterations: 2

```

Weighted ANOVA

```

svy_aov <- svyglm(social_cohesion_rev ~ age_band, design = chs2020_svy)
summary(svy_aov)

##
## Call:
## svyglm(formula = social_cohesion_rev ~ age_band, design = chs2020_svy)
##

```

```
## Survey design:
## svydesign(ids = ~1, strata = ~strata_q1, weights = ~wt21_dual_q1,
##   data = chsw3)
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.78647    0.02423 115.022 < 2e-16 ***
## age_band45-64 0.26713    0.03740   7.143 1.07e-12 ***
## age_band65+   0.33974    0.04412   7.700 1.69e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.5131945)
##
## Number of Fisher Scoring iterations: 2
# Post hoc comparisons (Bonferroni adjustment)
library(emmeans)

## Welcome to emmeans.
## Caution: You lose important information if you filter this package's results.
## See '? untidy'

emmeans(svy_aov, pairwise ~ age_band, adjust = "bonferroni")

## Warning: Model has 4321 prior weights, but we recovered 4327 rows of data.
## So prior weights were ignored.

## $emmeans
##   age_band emmean      SE    df lower.CL upper.CL
## 18-44      2.79 0.0242 4196     2.74     2.83
## 45-64      3.05 0.0285 4196     3.00     3.11
## 65+        3.13 0.0369 4196     3.05     3.20
##
## Confidence level used: 0.95
##
## $contrasts
##   contrast      estimate      SE    df t.ratio p.value
## (18-44) - (45-64) -0.2671 0.0374 4196  -7.143 <.0001
## (18-44) - (65+)   -0.3397 0.0441 4196  -7.700 <.0001
## (45-64) - (65+)   -0.0726 0.0467 4196  -1.556 0.3591
##
## P value adjustment: bonferroni method for 3 tests
```

Standardizing

```
# Standardize selected predictors
chsw3$z_cohesion <- scale(chsw3$social_cohesion_rev)
chsw3$z_delaypayrent <- scale(chsw3$delaypayrent0)
chsw3$z_rodents <- scale(chsw3$rodentsstreet0)

# Now update the survey object with the new variables
chs2020_svy_z <- update(chs2020_svy, z_cohesion = chsw3$z_cohesion,
                        z_delaypayrent = chsw3$z_delaypayrent,
                        z_rodents = chsw3$z_rodents)
```



```

# Run the weighted model
svy_lm_k6_z <- svyglm(k6 ~ z_cohesion + age_band + gender + race_ethnicity +
                     education1 + employment + z_delaypayrent + z_rodents,
                     design = chs2020_svy_z)

summary(svy_lm_k6_z)

##
## Call:
## svyglm(formula = k6 ~ z_cohesion + age_band + gender + race_ethnicity +
##       education1 + employment + z_delaypayrent + z_rodents, design = chs2020_svy_z)
##
## Survey design:
## update(chs2020_svy, z_cohesion = chsw3$z_cohesion, z_delaypayrent = chsw3$z_delaypayrent,
##       z_rodents = chsw3$z_rodents)
##
## Coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.151479   0.421090   9.859 < 2e-16
## z_cohesion       -0.249256   0.099906  -2.495 0.012639
## age_band45-64    -0.921701   0.215514  -4.277 1.94e-05
## age_band65+      -1.679147   0.354379  -4.738 2.23e-06
## gendermale       -0.787687   0.192956  -4.082 4.55e-05
## race_ethnicityAsian/Pacific Islander -0.716561   0.320823  -2.234 0.025569
## race_ethnicityBlack -0.942921   0.274082  -3.440 0.000587
## race_ethnicityHispanic -0.421392   0.268242  -1.571 0.116274
## race_ethnicityNorth African/Mid Eastern 1.335607   0.917007   1.456 0.145336
## race_ethnicityOther 0.974199   0.761504   1.279 0.200862
## education1High school graduate 0.006541   0.241291   0.027 0.978375
## education1Less than high school 0.022603   0.319113   0.071 0.943536
## education1Some college/technical school -0.100518   0.274388  -0.366 0.714135
## employmentA student 1.791193   0.631136   2.838 0.004562
## employmentEmployed for wages or salary 0.977580   0.402940   2.426 0.015304
## employmentRetired 2.127199   0.509292   4.177 3.02e-05
## employmentSelf-employed 1.063422   0.495941   2.144 0.032072
## employmentUnable to work 3.860980   0.619358   6.234 5.02e-10
## employmentUnemployed for 1 year or more 1.486199   0.571774   2.599 0.009376
## employmentUnemployed for less than 1 year 2.060827   0.500317   4.119 3.88e-05
## z_delaypayrent   -0.699658   0.121236  -5.771 8.47e-09
## z_rodents         0.445121   0.103157   4.315 1.63e-05
##
## (Intercept)      ***
## z_cohesion        *
## age_band45-64     ***
## age_band65+       ***
## gendermale        ***
## race_ethnicityAsian/Pacific Islander  *
## race_ethnicityBlack ***
## race_ethnicityHispanic
## race_ethnicityNorth African/Mid Eastern
## race_ethnicityOther
## education1High school graduate
## education1Less than high school

```

```

## education1Some college/technical school
## employmentA student **
## employmentEmployed for wages or salary *
## employmentRetired ***
## employmentSelf-employed *
## employmentUnable to work ***
## employmentUnemployed for 1 year or more **
## employmentUnemployed for less than 1 year ***
## z_delaypayrent ***
## z_rodents ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 16.41754)
##
## Number of Fisher Scoring iterations: 2
# Step 1: Tidy the model
std_table <- tidy(svy_lm_k6_z)

# filter for significant variables (p < 0.05)
signif_vars <- std_table |>
  filter(p.value < 0.05) |>
  mutate(across(where(is.numeric), ~ round(., 3))) |>
  rename(
    Variable = term,
    Estimate = estimate,
    "Std. Error" = std.error,
    "p-value" = p.value,
    "t value" = statistic
  ) |>
  select(Variable, Estimate, 'Std. Error', 't value', "p-value")

# create the table
signif_vars |>
  kable(caption = "Significant Predictors of K6 (Standardized Linear Regression)", escape=T)

```

Table 4: Significant Predictors of K6 (Standardized Linear Regression)

Variable	Estimate	Std. Error	t value	p-value
(Intercept)	4.151	0.421	9.859	0.000
z_cohesion	-0.249	0.100	-2.495	0.013
age_band45-64	-0.922	0.216	-4.277	0.000
age_band65+	-1.679	0.354	-4.738	0.000
gendermale	-0.788	0.193	-4.082	0.000
race_ethnicityAsian/Pacific Islander	-0.717	0.321	-2.234	0.026
race_ethnicityBlack	-0.943	0.274	-3.440	0.001
employmentA student	1.791	0.631	2.838	0.005
employmentEmployed for wages or salary	0.978	0.403	2.426	0.015
employmentRetired	2.127	0.509	4.177	0.000
employmentSelf-employed	1.063	0.496	2.144	0.032

Variable	Estimate	Std. Error	t value	p-value
employmentUnable to work	3.861	0.619	6.234	0.000
employmentUnemployed for 1 year or more	1.486	0.572	2.599	0.009
employmentUnemployed for less than 1 year	2.061	0.500	4.119	0.000
z_delaypayment	-0.700	0.121	-5.771	0.000
z_rodents	0.445	0.103	4.315	0.000

Visualizing effect of social cohesion across age

```
# First, extract confidence intervals
ci_18 <- confint(model_18)["social_cohesion_rev", ]
ci_45 <- confint(model_45)["social_cohesion_rev", ]
ci_65 <- confint(model_65plus)["social_cohesion_rev", ]

# Build a summary data frame
coef_data <- data.frame(
  age_group = c("18-44", "45-64", "65+"),
  coef = c(coef(model_18)["social_cohesion_rev"],
            coef(model_45)["social_cohesion_rev"],
            coef(model_65plus)["social_cohesion_rev"]),
  lower = c(ci_18[1], ci_45[1], ci_65[1]),
  upper = c(ci_18[2], ci_45[2], ci_65[2])
)

# plot
library(ggplot2)

ggplot(coef_data, aes(x = age_group, y = coef)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  labs(title = "Effect of Social Cohesion on Psychological Distress by Age Group",
       x = "Age Group", y = "Coefficient from Weighted Regression") +
  theme_minimal() +
  geom_hline(yintercept = 0, color = "black", linewidth = 0.5)
```

