Summary of Recodes

Simon J. Kiss, Matthew J. Arp and Scott Lu

09/03/2021

Introduction

This document tracks and summarizes basic data wrangling and recodes conducted at LISPOP for the 2021 Canadian Public Health Workforce and General Population Survey. The purpose of this document is to facilitate multi-institutional collaboration. LISPOP staff will take on the task of basic variable recodes and export data files for analysis by project researchers so that they can dedicate their attention to substantive analyses.

Scientific Literacy Questions

We have recoded variables Q14_1, Q14_2, Q14_3 and Q14_4 into variables `know1`, `know2`, 'know3' and `know4` such that respondents were coded 1 if they were correct and 0 if they were incorrect.

If Q14_1 = 1 or 2 (Definitely or Probably false), it is correct "Ordinary tomatoes do not contain genes, while genetically modified tomatoes do."

If Q14_2 = 1 or 2 (Definitely or Probably false), it is correct "Antibiotics kill viruses as well as bacteria."

```
know1=case_when(
Q14_1<3 ~ 1,
Q14_1>2 ~ 0,
TRUE ~ 0

know1=case_when(
Q14_2<3 ~ 1,
Q14_2>2 ~ 0,
TRUE ~ 0
```

If Q14_3 = 3 or 4 (Definitely or Probably true), it is correct "The cloning of living things produces genetically identical copies."

If Q14_4 = 3 or 4 (Definitely or Probably true), it is correct "All plants and animals have DNA."

```
know3=case_when(
    Q14_3>2 ~ 1,
    Q14_3<3 ~ 0,
    TRUE ~ 0

know3=case_when(
    Q14_4>2 ~ 1,
    Q14_4<3 ~ 0,
    TRUE ~ 0
```

```
full %>%
  select(Q14_1:Q14_4, starts_with('know')) %>%
  as_factor()
```

```
## # A tibble: 2,078 x 8
##    Q14_1        Q14_2        Q14_3        Q14_4        know1 know2 know3 know4
##    <fct>        <fct>        <fct>        <fct>        <dbl> <dbl> <dbl> <dbl>
##  1 Definitely F~ Probably True Probably Tr~ Definitely ~     1     0     1     1
##  2 Probably Fal~ Probably True Probably Fa~ Probably Fa~     1     0     0     0
##  3 Probably True Probably Fal~ Probably Tr~ Definitely ~     0     1     1     1
##  4 Definitely F~ Probably True Probably Tr~ Definitely ~     1     0     1     1
##  5 Definitely F~ Definitely F~ Definitely ~ Definitely ~     1     1     1     1
##  6 Probably True Definitely T~ Probably Tr~ Definitely ~     0     0     1     1
##  7 Definitely F~ Probably True Probably Tr~ Definitely ~     1     0     1     1
##  8 Probably Fal~ Probably True Probably Fa~ Definitely ~     1     0     0     1
##  9 Probably Fal~ Definitely F~ Probably Tr~ Probably Tr~     1     1     1     1
## 10 Probably Fal~ Probably Fal~ Probably Tr~ Probably Tr~     1     1     1     1
## # ... with 2,068 more rows
```

Cognitive Reflection test Questions

We have recoded variables Q18_1, Q19_1, Q20_1 and Q21_1 into variables `crt1`, `crt2` etc.such that respondents were coded 1 if they got the correct response and 0 if they were incorrect.

CRT Question Q18_1 coded 1 when answer (X) is a variation of "Second Place" "If you're running a race and you pass the person in second place, what place are you in?"

CRT Question Q19_1 coded 1 when answer (X) is a variation of "Eight" "A farmer had 15 sheep and all but 8 died. How many are left?"

CRT Question Q20_1 coded 1 when answer (X) is a variation of "Emily" "Emily's father has three daughters. The first two are named April and May. What is the third daughter's name?"

CRT Question Q21_1 coded 1 when answer (X) is a variation of "None" "How many cubic feet of dirt are there in a hole that is 3' deep x 3' wide x 3' long?"

```
full %>%
  select(Q18_1:Q21_1, starts_with('crt')) %>%
  as_factor()
```

```
## # A tibble: 2,078 x 8
##    Q18_1  Q19_1 Q20_1 Q21_1  crt1  crt2  crt3  crt4
##    <chr>  <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl>
##  1 second 8     june  27        1     1     0     0
##  2 First  7     June  3'        0     0     0     0
##  3 Second 7     Emily Idk       0     0     1     0
##  4 1      8     Emily 16        0     1     1     0
##  5 first  8     Emily 27        0     1     1     0
##  6 first  7     June  3'        0     0     0     0
##  7 first  7     June  27        0     0     0     0
##  8 First  8     Emily 9         0     1     1     0
##  9 second 15    Emily 3         1     0     1     0
## 10 1st    8     emily 27        0     1     0     0
## # ... with 2,068 more rows
```

Most Important Problem in Public Health

We have provided more meaningful variable names (e.g. `Obesity`, `Vaccine Hesitancy` , etc.) and *appended* these to the data frame. The original variables are `Q1_1` through `Q1_9`. In addition, we have gone through the open ended responses that respondents provided and added new variables that appeared frequently in the data set. The index currently reads:

1 = COVID-19

2 = Public health messaging/gov't handling of COVID

3 = Healthcare (access to care, short supply)

4 = Long term care

5 = Chronic disease (cancer, heart disease)

6 = Mental health

7 = Access to housing and food

8 = Drug abuse

9 = Inequality

10 = Economy

11 = Corrupt gov't

12 = Climate change/environmental

13 = Abortion

14 = Reliance on meat

15 = Domestic abuse

16 = Misinformation

17 = Internet addiction

18 = Don't know