Sex, Gender, Transgender, Sexuality - COVID CNS

Alicia Peel

18/01/2021

Configure global options for all chunks

Clear global environment prior to initiation

```
remove(list = ls())
```

Add the add_numeric function - used to convert character variables into numeric variables. Add the sumscores function - used to generate sumscores Add the package_check function - used to install and load dependencies

```
source(file = "../functions/add_numeric_1.R")
source(file = "../functions/remove_duplicates.R")
source(file = "../functions/sumscores.R")
source(file = "../functions/package_check.R")
source(file = "../functions/imp_check.R")
```

Use package_check to install and load dependencies Load tidyverse last

```
packages <- c(
  "summarytools",
  "sjlabelled",
  "Amelia",
  "gtsummary",
  "tidyverse"
  )
package_check(packages)
```

```
Loading required package: summarytools

Warning: package 'summarytools' was built under R version 4.0.5

Registered S3 method overwritten by 'pryr':
  method      from
  print.bytes Rcpp

Loading required package: sjlabelled

Attaching package: 'sjlabelled'
```

```
The following object is masked from 'package:summarytools':

    unlabel

Loading required package: Amelia

Warning: package 'Amelia' was built under R version 4.0.5

Loading required package: Rcpp

Warning: package 'Rcpp' was built under R version 4.0.5

##
## Amelia II: Multiple Imputation
## (Version 1.8.0, built: 2021-05-26)
## Copyright (C) 2005-2022 James Honaker, Gary King and Matthew Blackwell
## Refer to http://gking.harvard.edu/amelia/ for more information
##

Loading required package: gtsummary

Warning: package 'gtsummary' was built under R version 4.0.5

Loading required package: tidyverse

Warning: package 'tidyverse' was built under R version 4.0.5

-- Attaching packages --------------------------------------- tidyverse 1.3.1 --

v ggplot2 3.3.5     v purrr   0.3.4
v tibble  3.1.6     v dplyr   1.0.7
v tidyr   1.1.4     v stringr 1.4.0
v readr   2.0.2     v forcats 0.5.1

Warning: package 'ggplot2' was built under R version 4.0.5

Warning: package 'tibble' was built under R version 4.0.5

Warning: package 'tidyr' was built under R version 4.0.5

Warning: package 'readr' was built under R version 4.0.5

Warning: package 'purrr' was built under R version 4.0.5

Warning: package 'dplyr' was built under R version 4.0.5

Warning: package 'stringr' was built under R version 4.0.5

Warning: package 'forcats' was built under R version 4.0.5
```

```
-- Conflicts ---------------------------------------- tidyverse_conflicts() --
x forcats::as_factor() masks sjlabelled::as_factor()
x dplyr::as_label()    masks ggplot2::as_label(), sjlabelled::as_label()
x dplyr::filter()      masks stats::filter()
x dplyr::lag()         masks stats::lag()
x tibble::view()       masks summarytools::view()
```

Retrieve recent date We are using the recent date to save files with paste0() as an extension to not overwrite old versions

```
date <- Sys.Date()
date
```

```
[1] "2022-02-22"
```

Read in file with path to ilovedata channel on Teams

```
source(file = "../credentials/paths.R")
```

Read in the data: Demographics

CNS data

```
cns_data <- readRDS(
  file = paste0(ilovedata, "/data_raw/latest_freeze/covid_cns/baseline/dem_covid_cns.rds")
  )

# Check variable names in dataframe
cns_data %>%
  colnames()
```

```
  [1] "externalDataReference"
  [2] "startDate"
  [3] "endDate"
  [4] "dem.day"
  [5] "dem.month"
  [6] "dem.year"
  [7] "dem.required_question_eligibility_criteria.txt"
  [8] "dem.what_gender_do_you_identify_with"
  [9] "dem.what_gender_do_you_identify_with.txt"
 [10] "dem.do_you_consider_yourself_to_be_transgender"
 [11] "dem.have_you_ever_been_pregnant"
 [12] "dem.what_is_your_sexual_orientation"
 [13] "dem.what_is_your_sexual_orientation.txt"
 [14] "dem.what_is_your_current_maritalrelationship_status"
 [15] "dem.what_is_your_current_maritalrelationship_status.txt"
 [16] "dem.how_would_you_describe_your_vision"
 [17] "dem.how_would_you_describe_your_hearing"
 [18] "dem.which_hand_do_you_usually_write_with"
```

```
[19] "dem.college_or_university_degree"
[20] "dem.a_levelsas_levels_or_equivalent"
[21] "dem.o_levelsgcses_or_equivalent"
[22] "dem.cses_or_equivalent"
[23] "dem.nvq_or_hnd_or_hnc_or_equivalent"
[24] "dem.other_professional_qualifications_"
[25] "dem.other_professional_qualifications_text.txt"
[26] "dem.none_of_the_above"
[27] "dem.prefer_not_to_say"
[28] "dem.british_mixed_british"
[29] "dem.irish"
[30] "dem.northern_irish"
[31] "dem.any_other_white_background"
[32] "dem.white_and_black_caribbean"
[33] "dem.white_and_black_africa"
[34] "dem.white_and_asian"
[35] "dem.any_other_mixed_background"
[36] "dem.indian_or_british_indian"
[37] "dem.pakistani_or_british_pakistani"
[38] "dem.bangladeshi_or_british_bangladeshi"
[39] "dem.any_other_asian_background"
[40] "dem.caribbean"
[41] "dem.african"
[42] "dem.any_other_black_background"
[43] "dem.chinese"
[44] "dem.any_other_ethnic_group"
[45] "dem.other"
[46] "dem.othertext.txt"
[47] "dem.english"
[48] "dem.scottish"
[49] "dem.welsh"
[50] "dem.cornish"
[51] "dem.cypriot_"
[52] "dem.greek"
[53] "dem.greek_cypriot"
[54] "dem.italian"
[55] "dem.irish_traveller"
[56] "dem.traveller"
[57] "dem.gypsyromany"
[58] "dem.polish"
[59] "dem.republics_made_ussr"
[60] "dem.kosovan"
[61] "dem.albanian"
[62] "dem.bosnian"
[63] "dem.croatian"
[64] "dem.serbian"
[65] "dem.republics_made_yugoslavia"
[66] "dem.mixed_white"
[67] "dem.other_white_european_european_unspecified_european_mix"
[68] "dem.black_and_asian"
[69] "dem.black_and_chinese"
[70] "dem.black_and_white"
[71] "dem.chinese_and_white"
[72] "dem.asian_and_chinese"
```

```
 [73] "dem.other_mixed_mixed_unspecified"
 [74] "dem.other_mixed_mixed_unspecifiedtext.txt"
 [75] "dem.mixed_asian"
 [76] "dem.punjabi"
 [77] "dem.kashmiri"
 [78] "dem.east_african_asian"
 [79] "dem.tamil"
 [80] "dem.sinhalese"
 [81] "dem.british_asian"
 [82] "dem.caribbean_asian"
 [83] "dem.other_asian_asian_unspecified"
 [84] "dem.other_asian_asian_unspecifiedtext.txt"
 [85] "dem.somali"
 [86] "dem.mixed_black"
 [87] "dem.nigerian"
 [88] "dem.black_british"
 [89] "dem.other_black_black_unspecified"
 [90] "dem.other_black_black_unspecifiedtext.txt"
 [91] "dem.is_english_your_first_language"
 [92] "dem.what_is_your_first_language"
 [93] "dem.what_is_your_first_language.txt"
 [94] "dem.please_select_your_preferred_units_of_measurement"
 [95] "dem.what_is_your_current_height"
 [96] "dem.what_is_your_current_height.1"
 [97] "dem.what_is_your_current_height.2"
 [98] "dem.pregnant_weigh_weight_provide"
 [99] "dem.pregnant_weigh_weight_provide.1"
[100] "dem.pregnant_weigh_weight_provide.2"
[101] "dem.pregnant_weighed_weight_provide"
[102] "dem.pregnant_weighed_weight_provide.1"
[103] "dem.pregnant_weighed_weight_provide.2"
[104] "dem.highest_weight"
[105] "dem.stopped_growing_adult_height"
[106] "dem.stopped_growing_adult_height.1"
[107] "dem.stopped_growing_adult_height.2"
[108] "dem.body_suffered_injury_involving"
[109] "dem.middle_wake_night_covid19"
[110] "dem.middle_wake_night_covid19.1"
[111] "dem.medical_history_birth_relevant"
[112] "dem.affects_concerned_live_memory"
[113] "dem.memory_problem_worse_year"
[114] "dem.based_confirm_living_question"
[115] "dem.diagnosed_required_question_covid19"
[116] "dem.long_ago_diagnosed_required"
[117] "dem.long_ago_diagnosed_required.1"
[118] "dem.diagnosed_covid19_experienced_similar"
[119] "dem.quality_rate_life"
[120] "dem.energy_everyday_life"
[121] "dem.opportunity_leisure_activities"
[122] "dem.money_day"
[123] "dem.middle_wake_night_trouble"
[124] "dem.affects_concerned_live_memory.1"
[125] "dem.affects_concerned_live_memory.2"
[126] "dem.has_your_memory_got_progressively_worse"
```

```
[127] "dem.vietnamese"
[128] "dem.filipino"
[129] "dem.malaysian"
[130] "dem.any_other_group"
[131] "dem.any_other_grouptext.txt"
[132] "dem.lowest_weight_adult_height"
[133] "dem.happy_general_health"
```

```
# Inspect dimensions of dataframe
cns_data %>%
  dim()
```

```
[1] 235 133
```

Specify columns to be excluded from add_numeric function

```
exclude_cols_numeric <- c(
  "ID",
  "sample",
  "startDate",
  "endDate"
  )
```

Select & rename relevant columns

```
cns_data_id <- cns_data %>% #new dataset with ID
  drop_na(externalDataReference) %>% # Drop participants with no ID
  distinct(externalDataReference, .keep_all = TRUE) %>% # Changed to distinct due to NA coercion
  add_column(sample = "CNS",
             .after = "externalDataReference") %>% # Create new sample column
  select(
         ID = externalDataReference, # ID
         sample,
         startDate,
         endDate,
         dem.medical_history_birth_relevant,
         dem.what_gender_do_you_identify_with,
         dem.do_you_consider_yourself_to_be_transgender,
         dem.what_is_your_sexual_orientation,
         dem.have_you_ever_been_pregnant
         ) %>%
  add_numeric_1(exclude = exclude_cols_numeric)
```

```
New names:
* endDate -> endDate...4
* endDate -> endDate...10
```

```
# Inspect colnames
cns_data_id %>%
  colnames()
```

```
 [1] "ID"
 [2] "sample"
 [3] "startDate"
 [4] "endDate...4"
 [5] "dem.medical_history_birth_relevant"
 [6] "dem.what_gender_do_you_identify_with"
 [7] "dem.do_you_consider_yourself_to_be_transgender"
 [8] "dem.what_is_your_sexual_orientation"
 [9] "dem.have_you_ever_been_pregnant"
[10] "endDate...10"
[11] "dem.medical_history_birth_relevant_numeric"
[12] "dem.what_gender_do_you_identify_with_numeric"
[13] "dem.do_you_consider_yourself_to_be_transgender_numeric"
[14] "dem.what_is_your_sexual_orientation_numeric"
[15] "dem.have_you_ever_been_pregnant_numeric"
```

Look at number of people excluded

```
# Inspect dimensions of new data set
cns_data_id %>%
  dim()
```

```
[1] 228  15
```

```
# Inspect number of rows dropped
cns_excluded <- dim(cns_data_id)[1] - dim(cns_data)[1]
cns_excluded
```

```
[1] -7
```

Inspect numeric variables

```
cns_data_id %>%
  select(all_of(ends_with("numeric"))) %>%
  tbl_summary(missing_text = "Missing")
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**N = 228**
What was your assigned sex at birth?* This question is required.Note: This question refers to assigned sex at birth, not gender. Responses to this question are used to select questionnaire items that may be relevant to your medical history.	93 (41%)
Missing	1
What gender do you identify with?	
-999	2 (0.9%)

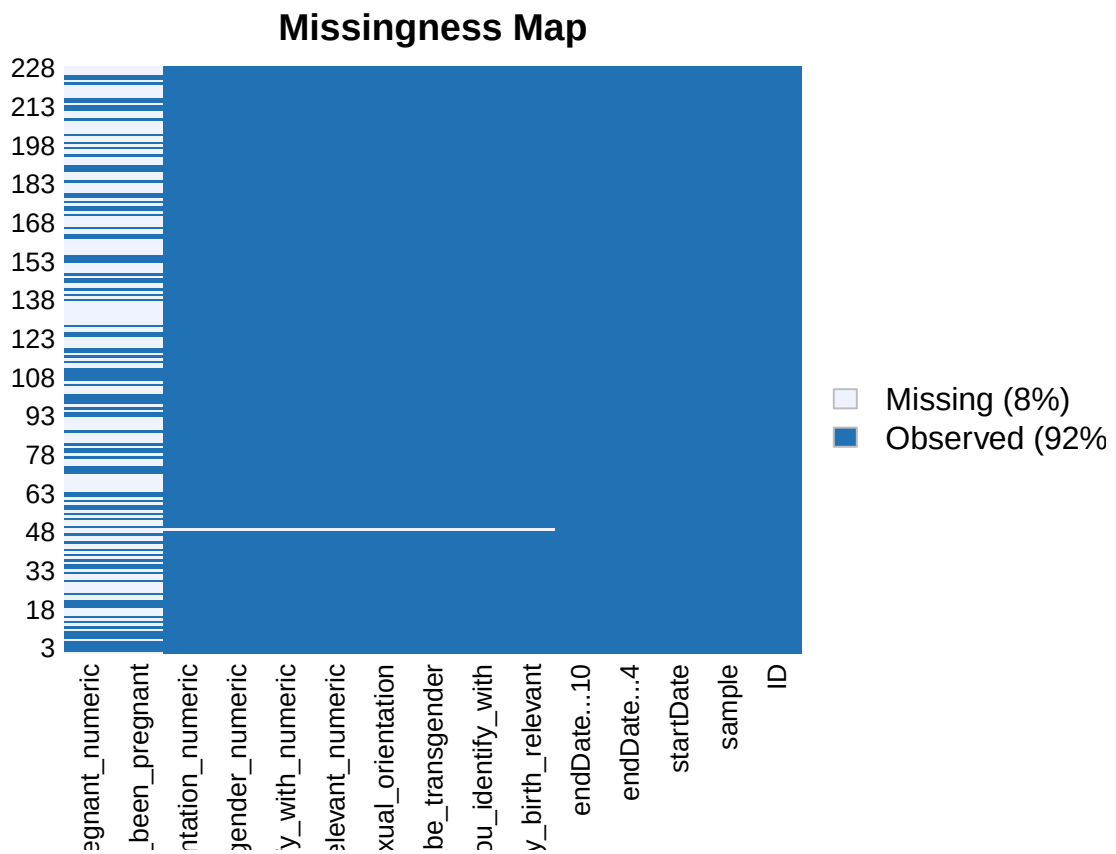
Characteristic	N = 228
0	133 (59%)
1	92 (41%)
Missing	1
Do you consider yourself to be transgender?	1 (0.4%)
Missing	1
What is your sexual orientation?	
-999	4 (1.8%)
0	212 (93%)
1	5 (2.2%)
2	5 (2.2%)
4	1 (0.4%)
Missing	1
Have you ever been pregnant?	70 (75%)
Missing	135

Check missingness by missmap

```
cns_miss_map <- cns_data_id %>%
  missmap()
```

```
Warning: Unknown or uninitialised column: ‘arguments‘.
Unknown or uninitialised column: ‘arguments‘.

Warning: Unknown or uninitialised column: ‘imputations‘.
```




```
cns_miss_map
```

```
NULL
```

Rename data for cleaning

```
dat <- cns_data_id
#Check
dat %>% glimpse()
```

```
Rows: 228
Columns: 15
$ ID                                                  <chr> "\nCNS01004", "~
$ sample                                              <chr> "CNS", "CNS", "~
$ startDate                                           <dttm> 2021-03-25 12:~
$ endDate...4                                         <dttm> 2021-03-25 14:~
$ dem.medical_history_birth_relevant                  <fct> Male, Male, Mal~
$ dem.what_gender_do_you_identify_with                <fct> Male, Male, Mal~
$ dem.do_you_consider_yourself_to_be_transgender      <fct> No, No, No, No,~
$ dem.what_is_your_sexual_orientation                 <fct> Heterosexual, H~
$ dem.have_you_ever_been_pregnant                     <fct> NA, NA, NA, No,~
$ endDate...10                                        <dttm> 2021-03-25 14:~
$ dem.medical_history_birth_relevant_numeric          <dbl> 0, 0, 0, 1, 1, ~
$ dem.what_gender_do_you_identify_with_numeric        <dbl> 0, 0, 0, 1, 1, ~
$ dem.do_you_consider_yourself_to_be_transgender_numeric <dbl> 0, 0, 0, 0, 0, ~
```

```
$ dem.what_is_your_sexual_orientation_numeric            <dbl> 0, 0, 0, 2, 0, ~
$ dem.have_you_ever_been_pregnant_numeric                <dbl> NA, NA, NA, 0, ~
```

Recode Non-answer values to 3 digits -555 'Not applicable' response from participant -777 Seen but not answered -888 Don't know -999 Prefer not to answer/Prefer not to say `NA` Were not shown the question (genuinely missing value) When we code someone as being 'not applicable' by deduction, we use `NA_real_`

```
dat <- dat %>%
  mutate(across(ends_with("numeric"),
                ~case_when(
                  . == -55 ~ -555,
                  . == -77 ~ -777,
                  . == -88 ~ -888,
                  . == -99 ~ -999,
                  TRUE ~ .)))
```

Numeric 0-1 variables

Cleaning numeric variables

```
values_numeric_0 <- c(
  0,
  1,
  -777,
  NA
  )
values_numeric_0
```

```
[1]    0    1 -777   NA
```

Create vector of variable names for numeric variables

```
variables_numeric_0 <- c(
    "dem.medical_history_birth_relevant_numeric",
    "dem.do_you_consider_yourself_to_be_transgender_numeric",
    "dem.have_you_ever_been_pregnant_numeric"
  )
variables_numeric_0
```

```
[1] "dem.medical_history_birth_relevant_numeric"
[2] "dem.do_you_consider_yourself_to_be_transgender_numeric"
[3] "dem.have_you_ever_been_pregnant_numeric"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_numeric_0,
          values = values_numeric_0)
```

```
[1] "There are no implausible values in the dataset. Can leave these variables as they are."
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**N = 228**
dem.medical_history_birth_relevant_numeric	93 (41%)
Missing	1
dem.do_you_consider_yourself_to_be_transgender_numeric	1 (0.4%)
Missing	1
dem.have_you_ever_been_pregnant_numeric	70 (75%)
Missing	135

Numeric 0-3 variables

Cleaning numeric variables

```
values_numeric_0_3 <- c(
  0,
  1,
  2,
  3,
  -777,
  -888,
  -999,
  NA
  )
values_numeric_0_3
```

```
[1]    0    1    2    3 -777 -888 -999   NA
```

Vector of numeric variables

```
variables_numeric_0_3 <- c(
    "dem.what_gender_do_you_identify_with_numeric"
  )
variables_numeric_0_3
```

```
[1] "dem.what_gender_do_you_identify_with_numeric"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_numeric_0_3,
          values = values_numeric_0_3)
```

```
[1] "There are no implausible values in the dataset. Can leave these variables as they are."
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	N = 228
dem.what_gender_do_you_identify_with_numeric	
-999	2 (0.9%)
0	133 (59%)
1	92 (41%)
Missing	1

Numeric 1-5 variables

Create vector of numeric values

```
values_numeric_5 <- c(
  0,
  1,
  2,
  3,
  4,
  -777,
  -999,
  NA
  )
values_numeric_5
```

```
[1]    0    1    2    3    4 -777 -999   NA
```

Vector of numeric variables

```
variables_numeric_5 <- c(
    "dem.what_is_your_sexual_orientation_numeric"
  )
variables_numeric_5
```

```
[1] "dem.what_is_your_sexual_orientation_numeric"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_numeric_5,
          values = values_numeric_5)
```

```
[1] "There are no implausible values in the dataset. Can leave these variables as they are."
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	N = 228
dem.what_is_your_sexual_orientation_numeric	
-999	4 (1.8%)
0	212 (93%)
1	5 (2.2%)
2	5 (2.2%)
4	1 (0.4%)
Missing	1

Categorical 0-1 variables (sex)

Create vector of categorical values for variables

```
values_categorical_sex <- c(
  "Male",
  "Female",
  "Seen but not answered",
  NA
  )
values_categorical_sex
```

```
[1] "Male"                  "Female"                "Seen but not answered"
[4] NA
```

Create vector of variable names

```
variables_categorical_sex <- c(
    "dem.medical_history_birth_relevant"
  )
variables_categorical_sex
```

```
[1] "dem.medical_history_birth_relevant"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_categorical_sex,
          values = values_categorical_sex)
```

```
[1] "There are no implausible values in the dataset. Can leave these variables as they are."

Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	N = 228
What was your assigned sex at birth?* This question is required.Note: This question refers to assigned sex at birth, not gender. Responses to this question are used to select questionnaire items that may be relevant to your medical history.	
Seen but not answered	0 (0%)
Male	134 (59%)
Female	93 (41%)
Missing	1

Categorical 0-1 variables (yes/no)

Create vector of categorical values for variables

```
values_categorical_0 <- c(
  "No",
  "Yes",
  "Seen but not answered",
  NA
  )
values_categorical_0
```

```
[1] "No"                   "Yes"                  "Seen but not answered"
[4] NA
```

Create vector of variable names

```
variables_categorical_0 <- c(
    "dem.do_you_consider_yourself_to_be_transgender",
    "dem.have_you_ever_been_pregnant"
  )
variables_categorical_0
```

```
[1] "dem.do_you_consider_yourself_to_be_transgender"
[2] "dem.have_you_ever_been_pregnant"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_categorical_0,
          values = values_categorical_0)
```

```
[1] "There are no implausible values in the dataset. Can leave these variables as they are."
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	N = 228
Do you consider yourself to be transgender?	
Seen but not answered	0 (0%)
No	226 (100%)
Yes	1 (0.4%)
Prefer not to say	0 (0%)
Missing	1
Have you ever been pregnant?	
Seen but not answered	0 (0%)
No	23 (25%)
Yes	70 (75%)
Prefer not to say	0 (0%)
Missing	135

Categorical 0-3 variables

Create vector of categorical values for variables

```
values_categorical_0_3 <- c(
  "Male",
  "Female",
  "Non-binary",
  "Self-define",
  "Seen but not answered",
  "Don't know",
  "Prefer not to answer",
  NA
  )
values_categorical_0_3
```

```
[1] "Male"                  "Female"                "Non-binary"
[4] "Self-define"           "Seen but not answered" "Don't know"
[7] "Prefer not to answer"  NA
```

Create vector of variable names

```
variables_categorical_0_3 <- c(
    "dem.what_gender_do_you_identify_with"
  )
variables_categorical_0_3
```

```
[1] "dem.what_gender_do_you_identify_with"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_categorical_0_3,
          values = values_categorical_0_3)
```

```
[1] "The number of implausible values in the dataset is 2. Please investigate."
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**N = 228**
What gender do you identify with?	
Prefer not to say	2 (0.9%)
Don't know	0 (0%)
Seen but not answered	0 (0%)
Male	133 (59%)
Female	92 (41%)
Non-binary	0 (0%)
Prefer to self-define (please tell us more):	0 (0%)
Missing	1

Remove once cleaned on Qualtrics

```
dat <- dat %>%
  mutate(
    across(all_of(variables_categorical_0_3),
           .fns = ~dplyr::recode(.,
             "Prefer to self-define (please tell us more):" = "Self-define"
            )
          )
  )
```

```
dat %>%
  select(all_of(variables_categorical_0_3)) %>%
  freq()
```

```
Frequencies
dat$dem.what_gender_do_you_identify_with
Label: What gender do you identify with?
Type: Factor

                           Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
-------------------------- ------ --------- -------------- --------- --------------
          Prefer not to say      2      0.88           0.88      0.88           0.88
                Don't know      0      0.00           0.88      0.00           0.88
      Seen but not answered      0      0.00           0.88      0.00           0.88
                       Male    133     58.59          59.47     58.33          59.21
                     Female     92     40.53         100.00     40.35          99.56
                 Non-binary      0      0.00         100.00      0.00          99.56
                Self-define      0      0.00         100.00      0.00          99.56
                       <NA>      1                                0.44         100.00
                      Total    228    100.00         100.00    100.00         100.00
```

Categorical 1-5 variables

Create vector of categorical values for variables

```
values_categorical_5 <- c(
  "Heterosexual",
  "Homosexual",
  "Bisexual",
  "Asexual",
  "Other",
  "Seen but not answered",
  "Don't know",
  "Prefer not to answer",
  NA
  )
values_categorical_5
```

```
[1] "Heterosexual"          "Homosexual"            "Bisexual"
[4] "Asexual"               "Other"                 "Seen but not answered"
[7] "Don't know"            "Prefer not to answer"  NA
```

Create vector of variable names

```
variables_categorical_5 <- c(
    "dem.what_is_your_sexual_orientation"
  )
variables_categorical_5
```

```
[1] "dem.what_is_your_sexual_orientation"
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
           variables = variables_categorical_5,
          values = values_categorical_5)
```

```
[1] "The number of implausible values in the dataset is 4. Please investigate."
```

```
Table printed with `knitr::kable()`, not {gt}. Learn why at
https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**N = 228**
What is your sexual orientation?	
Prefer not to say	4 (1.8%)
Seen but not answered	0 (0%)
Heterosexual	212 (93%)
Homosexual	5 (2.2%)
Bisexual	5 (2.2%)
Asexual	0 (0%)
Other	1 (0.4%)
Missing	1

Remove once cleaned on Qualtrics

```
dat <- dat %>%
  mutate(
    across(all_of(variables_categorical_5),
           .fns = ~dplyr::recode(.,
             "Prefer not to say" = "Prefer not to answer"
            )
          )
  )
```

```
dat %>%
  select(all_of(variables_categorical_5)) %>%
  freq()
```

```
Frequencies
dat$dem.what_is_your_sexual_orientation
Label: What is your sexual orientation?
Type: Factor
```

	Freq	% Valid	% Valid Cum.	% Total	% Total Cum.
Prefer not to answer	4	1.76	1.76	1.75	1.75
Seen but not answered	0	0.00	1.76	0.00	1.75
Heterosexual	212	93.39	95.15	92.98	94.74
Homosexual	5	2.20	97.36	2.19	96.93
Bisexual	5	2.20	99.56	2.19	99.12
Asexual	0	0.00	99.56	0.00	99.12
Other	1	0.44	100.00	0.44	99.56
<NA>	1			0.44	100.00
Total	228	100.00	100.00	100.00	100.00

Save cleaned data

Check colnames before exporting final dataset

```
colnames(dat)
```

```
 [1] "ID"
 [2] "sample"
 [3] "startDate"
 [4] "endDate...4"
 [5] "dem.medical_history_birth_relevant"
 [6] "dem.what_gender_do_you_identify_with"
 [7] "dem.do_you_consider_yourself_to_be_transgender"
 [8] "dem.what_is_your_sexual_orientation"
 [9] "dem.have_you_ever_been_pregnant"
[10] "endDate...10"
[11] "dem.medical_history_birth_relevant_numeric"
[12] "dem.what_gender_do_you_identify_with_numeric"
[13] "dem.do_you_consider_yourself_to_be_transgender_numeric"
[14] "dem.what_is_your_sexual_orientation_numeric"
[15] "dem.have_you_ever_been_pregnant_numeric"
```

Combined object for excluded participants

```
cns_excluded <- as.data.frame(cns_excluded)
colnames(cns_excluded) <- c("Number of Participants Excluded")
rownames(cns_excluded) <- c("CNS")
cns_excluded
```

```
    Number of Participants Excluded
CNS                              -7
```

CNS

```
dat %>%
  filter(sample == "CNS") %>%
  saveRDS(
    file = paste0(ilovedata, "/data/latest_freeze/covidcns/sex_gender_sexuality_covidcns_clean.rds")
    )
```