

Supplement 1

Demographics

Peter Kamerman

Last knitted: 26 May 2021

Contents

1	Import and check data	1
2	Basic descriptive statistics	2
2.1	Full cohort	2
2.1.1	Descriptives	2
2.2	By study site	3
2.2.1	Sample size	3
2.2.2	Descriptives	3
3	Session information	4

1 Import and check data

```
# Import
demo <- read_rds('data-cleaned/data-demographics.rds')%>%
  # Fix site names:
  ## DD = Site 1, NM = Site 2, RESI = Site 3, RP = Site 4, STIG = Site 5
  mutate(Site = case_when(
    Site == 'DD' ~ 'Site 1',
    Site == 'NM' ~ 'Site 2',
    Site == 'RESI' ~ 'Site 3',
    Site == 'RP' ~ 'Site 4',
    Site == 'STIG' ~ 'Site 5'
  ))

# Check
## Demographics
dim(demo)

## [1] 596    8

names(demo)

## [1] "ID"           "Site"         "Sex"
## [4] "Age"         "Employment_status" "CD4_recent"
## [7] "ART_currently" "Education"
```

```
glimpse(demo)

## Rows: 596
## Columns: 8
## $ ID          <chr> "RPB73", "RPB74", "RPB75", "RPB76", "RPB77", "RPB78"~
## $ Site        <chr> "Site 4", "Site 4", "Site 4", "Site 4", "Site 4", "S~
## $ Sex         <chr> "Female", "Female", "Female", "Female", "Female", "F~
## $ Age         <dbl> 36, 27, 39, 36, 31, 32, 28, 37, 31, 25, 31, 24, 35, ~
## $ Employment_status <chr> "Other", "Unemployed", "Other", "Unemployed", "Unemp~
## $ CD4_recent   <dbl> 391, 571, 591, 207, 126, 225, 543, 410, 74, 212, 579~
## $ ART_currently <chr> "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", "Ye~
## $ Education    <chr> "Tertiary", "Secondary", "Secondary", "Primary", "Se~
```

2 Basic descriptive statistics

2.1 Full cohort

2.1.1 Descriptives

```
demo %>%
  # Remove ID column
  select(-ID, -Site) %>%
  # Convert character to factor
  mutate_if(is.character, factor) %>%
  my_skim()
```

Table 1: Data summary

Name	Piped data
Number of rows	596
Number of columns	6
Column type frequency:	
factor	4
numeric	2
Group variables	None

Variable type: factor

skim_variable	missing	complete	n_unique	top_counts
Sex	0	596	2	Fem: 481, Mal: 115
Employment_status	49	547	4	Une: 330, Ful: 131, Par: 52, Oth: 34
ART_currently	5	591	2	Yes: 460, No: 131
Education	37	559	3	Sec: 395, Pri: 99, Ter: 65

Variable type: numeric

skim_variable	missing	complete	mean	sd	p0	p25	p50	p75	p100
Age	8	588	37.28	9.06	19	31	36	42	76
CD4_recent	99	497	320.71	238.92	1	155	261	432	1232

2.2 By study site

2.2.1 Sample size

```
demo %>%  
  # Select ID  
  select(Site) %>%  
  # Group by study site  
  group_by(Site) %>%  
  summarise(Count = n()) %>%  
  kable(caption = 'Sample size by study site')
```

Table 4: Sample size by study site

Site	Count
Site 1	60
Site 2	239
Site 3	99
Site 4	148
Site 5	50

2.2.2 Descriptives

```
demo %>%  
  # Convert character to factor  
  mutate_if(is.character, factor) %>%  
  # Remove ID  
  select(-ID) %>%  
  # Group by site  
  group_by(Site) %>%  
  my_skim()
```

Table 5: Data summary

Name	Piped data
Number of rows	596
Number of columns	7
Column type frequency:	
factor	4
numeric	2

Group variables	Site
-----------------	------

Variable type: factor

skim_variable	Site	missing	complete	n_unique	top_counts
Sex	Site 1	0	60	2	Fem: 39, Mal: 21
Sex	Site 2	0	239	2	Fem: 184, Mal: 55
Sex	Site 3	0	99	2	Fem: 66, Mal: 33
Sex	Site 4	0	148	1	Fem: 148, Mal: 0
Sex	Site 5	0	50	2	Fem: 44, Mal: 6
Employment_status	Site 1	0	60	4	Une: 36, Ful: 17, Par: 6, Oth: 1
Employment_status	Site 2	49	190	3	Une: 126, Ful: 55, Par: 9, Oth: 0
Employment_status	Site 3	0	99	4	Une: 47, Ful: 35, Par: 16, Oth: 1
Employment_status	Site 4	0	148	4	Une: 99, Oth: 31, Par: 14, Ful: 4
Employment_status	Site 5	0	50	4	Une: 22, Ful: 20, Par: 7, Oth: 1
ART_currently	Site 1	0	60	2	Yes: 58, No: 2
ART_currently	Site 2	3	236	2	Yes: 130, No: 106
ART_currently	Site 3	1	98	2	Yes: 95, No: 3
ART_currently	Site 4	0	148	2	Yes: 128, No: 20
ART_currently	Site 5	1	49	1	Yes: 49, No: 0
Education	Site 1	2	58	3	Sec: 40, Pri: 9, Ter: 9
Education	Site 2	27	212	3	Sec: 160, Pri: 44, Ter: 8
Education	Site 3	4	95	3	Sec: 61, Pri: 19, Ter: 15
Education	Site 4	2	146	3	Sec: 108, Ter: 20, Pri: 18
Education	Site 5	2	48	3	Sec: 26, Ter: 13, Pri: 9

Variable type: numeric

skim_variable	Site	missing	complete	mean	sd	p0	p25	p50	p75	p100
Age	Site 1	0	60	38.63	5.87	27	34.00	38.5	42.00	58
Age	Site 2	6	233	36.73	8.54	20	31.00	36.0	41.00	76
Age	Site 3	0	99	43.40	9.32	19	35.00	42.0	49.00	66
Age	Site 4	2	146	30.82	4.72	20	27.00	31.0	34.00	40
Age	Site 5	0	50	44.90	9.67	25	37.25	44.5	52.50	65
CD4_recent	Site 1	18	42	428.60	255.93	16	256.00	374.5	577.25	1095
CD4_recent	Site 2	37	202	222.10	180.15	1	103.75	185.5	308.75	963
CD4_recent	Site 3	16	83	452.76	278.07	11	256.50	405.0	651.00	1232
CD4_recent	Site 4	4	144	322.19	205.43	26	174.00	270.0	415.00	1040
CD4_recent	Site 5	24	26	482.88	292.90	50	291.00	434.5	582.00	1138

3 Session information

```
sessionInfo()
```

```
## R version 4.0.4 (2021-02-15)
```

```

## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Catalina 10.15.7
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] knitr_1.31      skimr_2.1.3     forcats_0.5.1  stringr_1.4.0
## [5] dplyr_1.0.5     purrr_0.3.4     readr_1.4.0    tidyr_1.1.3
## [9] tibble_3.1.0    ggplot2_3.3.3   tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
## [1] tidyselect_1.1.0 xfun_0.22       repr_1.1.3      haven_2.3.1
## [5] colorspace_2.0-0 vctrs_0.3.6     generics_0.1.0  htmltools_0.5.1.1
## [9] base64enc_0.1-3  yaml_2.2.1      utf8_1.2.1      rlang_0.4.10
## [13] pillar_1.5.1     glue_1.4.2      withr_2.4.1     DBI_1.1.1
## [17] dbplyr_2.1.0     modelr_0.1.8    readxl_1.3.1    lifecycle_1.0.0
## [21] munsell_0.5.0    gtable_0.3.0    cellranger_1.1.0 rvest_1.0.0
## [25] evaluate_0.14    fansi_0.4.2     highr_0.8       broom_0.7.5
## [29] Rcpp_1.0.6       scales_1.1.1    backports_1.2.1 jsonlite_1.7.2
## [33] fs_1.5.0         hms_1.0.0       digest_0.6.27   stringi_1.5.3
## [37] grid_4.0.4       cli_2.3.1       tools_4.0.4     magrittr_2.0.1
## [41] crayon_1.4.1     pkgconfig_2.0.3 ellipsis_0.3.1  xml2_1.3.2
## [45] reprex_1.0.0     lubridate_1.7.10 assertthat_0.2.1 rmarkdown_2.7
## [49] httr_1.4.2       rstudioapi_0.13 R6_2.5.0        compiler_4.0.4

```