UNAIDS_2024_Clean_Estimates

Autogenerated data summary from dataReporter

2024-08-03 16:46:44.50579

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	242993
Number of variables	18

Checks performed

The following variable checks were performed, depending on the data type of each variable:

	character	factor	labelled	haven labelled	numeric	integer	logical	Date
Identify miscoded missing values	×	×	×	×	×	×		×
Identify prefixed and suffixed whitespace	×	×	×	×				
Identify levels with < 6 obs.	×	×	×	×				
Identify case issues	×	×	×	×				
Identify misclassified numeric or integer variables	×	×	×	×				
Identify outliers					×	×		×

Please note that all numerical values in the following have been rounded to 2 decimals.

Codebook summary table

Label	Variable	Class	# unique values	Missing	Description
Year of estimate	**year**	numeric	34	0.00 %	
Three digit ISO country code	**iso**	character	183	0.00 %	
Country name	**country**	character	183	0.00 %	
UNAIDS regional groupings	**region**	character	10	1.37 %	
Epidemic indicator	**indicator**	character	19	0.00 %	
Age groups	**age**	character	5	0.00 %	
Sex groups	**sex**	character	3	0.00 %	
Modeled estimate value	**estimate**	numeric	35592	29.64 %	

| Label	Variable	Class	# unique values	Missing	Description
Lower bound estimate from model	**lower_bound**	numeric	21761	40.73 %	
Upper bound estimate from model	**upper_bound**	numeric	21876	40.73 %	
demarcates estimates that have $<$ | $>$ signs associated with them	**estimate_flag**	logical	3	31.25 %	
Sheet from which data are derived	**sheet**	character	17	0.00 %	
nNature of the estimate, percent or integer	**indic_type**	character	2	0.00 %	
Logical flagging PEPFAR operating units	**pepfar**	logical	3	6.88 %	
Flag indicating that a country has achieved all 3 95 UNAIDS goals using the PLHIV base	**achv_95_plhiv**	logical	3	20.81 %	
Flag indicating that a country has achieved all 3 95 UNAIDS goals using the relative (sliding) base	**achv_95_relative**	logical	3	20.62 %	
Ratio of new HIV Infections to total deaths to HIV population in 2023	**epi_ratio_2023**	numeric	130	30.85 %	
Flag indicating a country has reached epi countrol – where total number of new HIV infections falls below the total number of deaths to HIV population	**epi_control**	logical	3	27.69 %	

Variable list

year

Year of estimate

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	34
Median	2013
1st and 3rd quartiles	2006; 2018
Min. and max.	1990; 2023
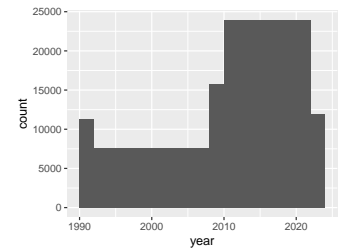
iso

Three digit ISO country code

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	183
Mode	"03M49WLD"
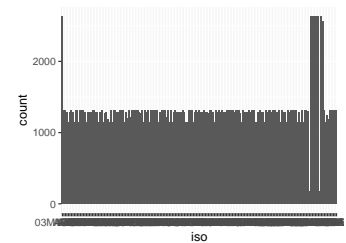
country

Country name

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	183
Mode	"Asia and the Pacific"
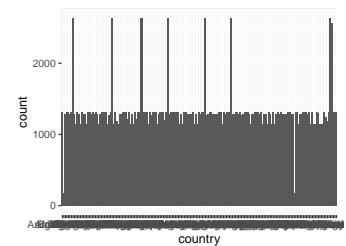
region

UNAIDS regional groupings
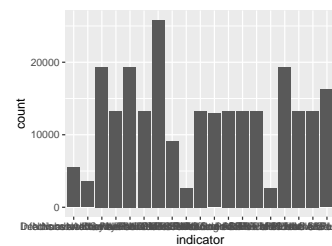
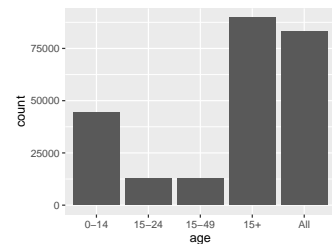| Feature	Result
Variable type	character
Number of missing obs.	3332 (1.37 %)
Number of unique values	9
Mode	"Western and central Europe and North America"
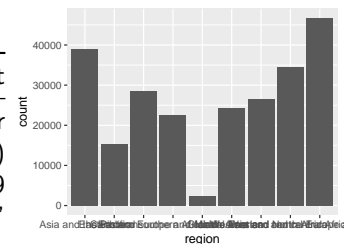
indicator

Epidemic indicator

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	19
Mode	"Number PLHIV"

age

Age groups

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	"15+"

sex

Sex groups

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"All"
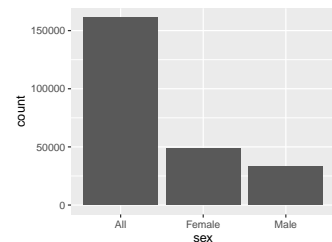
estimate

Modeled estimate value

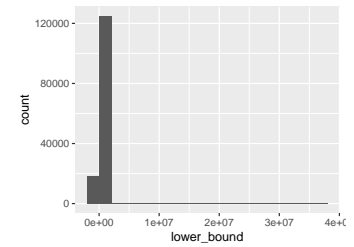| Feature	Result
Variable type	numeric
Number of missing obs.	72022 (29.64 %)
Number of unique values	35591
Median	100
1st and 3rd quartiles	35; 4700
Min. and max.	-526.09; 39900000
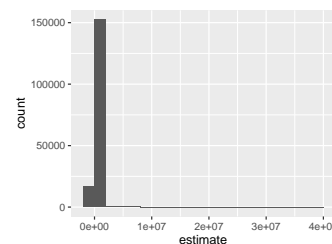
- Note that the following possible outlier values were detected: "-526.09", "-291.65", "-246.92", "-218.38", "-214.17", ..., "38600000", "38700000", "39200000", "39500000", "39900000" (3376 values omitted).

lower_bound

Lower bound estimate from model

| Feature	Result
Variable type	numeric
Number of missing obs.	98964 (40.73 %)
Number of unique values	21760
Median	82
1st and 3rd quartiles	15; 1500
Min. and max.	-461.04; 36100000

- Note that the following possible outlier values were detected: "-461.04", "-228.46", "-194.77", "-193.42", "-191.36", ..., "34900000", "3.5e+07", "35400000", "35800000", "36100000" (1771 values omitted).

upper_bound

Upper bound estimate from model

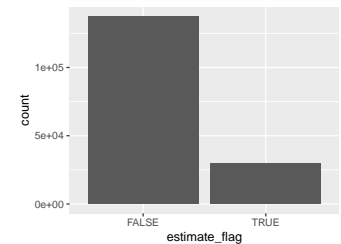| Feature	Result
Variable type	numeric
Number of missing obs.	98964 (40.73 %)
Number of unique values	21875
Median	98
1st and 3rd quartiles	24; 2500
Min. and max.	-613.32; 44600000
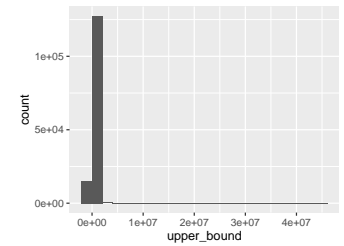
- Note that the following possible outlier values were detected: "-613.32", "-379.37", "-321.19", "-257.61", "-244.68", ..., "43100000", "43200000", "43800000", "44200000", "44600000" (1616 values omitted).

estimate_flag

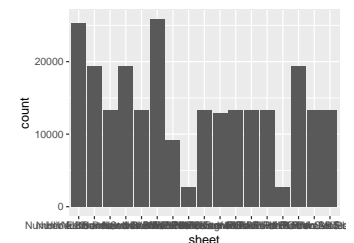*demarcates estimates that have $<|>$ signs associated with them*

Feature	Result
Variable type	logical
Number of missing obs.	75925 (31.25 %)
Number of unique values	2
Mode	"FALSE"

sheet

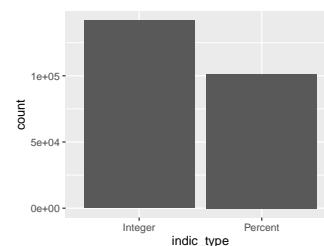*Sheet from which data are derived*

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	17
Mode	"Number PLHIV"

indic_type

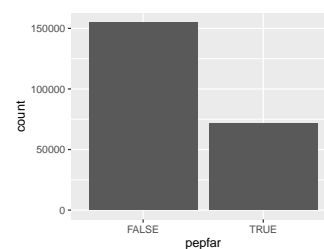*nNature of the estimate, percent or integer*

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"Integer"

pepfar

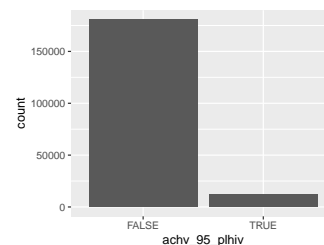*Logical flagging PEPFAR operating units*

Feature	Result
Variable type	logical
Number of missing obs.	16729 (6.88 %)
Number of unique values	2
Mode	"FALSE"

achv_95_plhiv

*Flag indicating that a country has achieved all 3 95 UNAIDS goals using the PLHIV base*

Feature	Result
Variable type	logical
Number of missing obs.	50560 (20.81 %)
Number of unique values	2
Mode	"FALSE"

achv_95_relative

*Flag indicating that a country has achieved all 3 95 UNAIDS goals using the relative (sliding) base*

Feature	Result
Variable type	logical
Number of missing obs.	50094 (20.62 %)
Number of unique values	2
Mode	"FALSE"
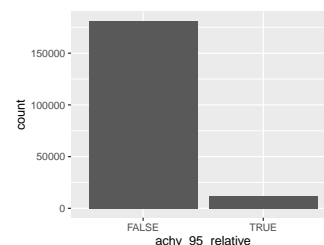
epi_ratio_2023

Ratio of new HIV Infections to total deaths to HIV population in 2023

Feature	Result
Variable type	numeric
Number of missing obs.	74970 (30.85 %)
Number of unique values	129
Median	2.06
1st and 3rd quartiles	1.06; 4.08
Min. and max.	0.52; 44.44
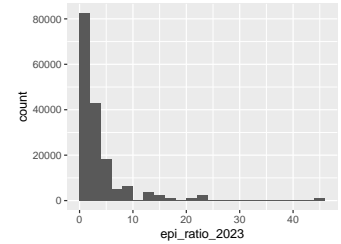
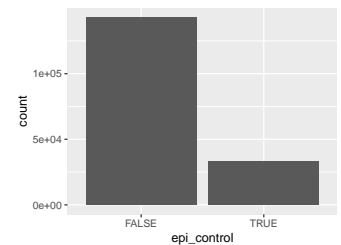- Note that the following possible outlier values were detected: "20.26", "22.7", "23.25", "44.44".

epi_control

Flag indicating a country has reached epi countrol – where total number of new HIV infections falls below the total number of deaths to HIV population

Feature	Result
Variable type	logical
Number of missing obs.	67278 (27.69 %)
Number of unique values	2
Mode	"FALSE"

Report generation information:

- Created by: tessam30 (username: `tessam`).

- Report creation time: Sat Aug 03 2024 16:47:02

- Report was run from directory: `C:/Users/tessam/Documents/Github/mindthegap`

- dataReporter v1.0.2 [Pkg: 2021-11-11 from CRAN (R 4.4.1)]

- R version 4.4.0 (2024-04-24 ucrt).

- Platform: x86_64-w64-mingw32/x64(America/New_York).

- Function call: `makeDataReport(data = UNAIDS_2024_Clean_Estimates, file = "data-raw/UNAIDS_2024_Clean_Estin` `replace = TRUE, codebook = TRUE)`