

CL_Anthropometry_05

Group K

2025-09-19

Display Dataset content

```
## # A tibble: 5 x 29
##   ISO3      DataId Indicator Value Precision DHS_CountryCode CountryName SurveyYear
##   <chr>   <chr>   <chr>    <chr> <chr>      <chr>          <chr>      <chr>
## 1 #coun~ #meta~ #indicat~ #ind~ #indicat~ <NA>          #country+n~ #date+year
## 2 ZAF     198690 Children~ 9.8    1        ZA           South Afri~ 2016
## 3 ZAF     198687 Children~ 27.4   1        ZA           South Afri~ 2016
## 4 ZAF     198688 Mean hei~ -1.1   1        ZA           South Afri~ 2016
## 5 ZAF     597227 Children~ 0.6    1        ZA           South Afri~ 2016
## # i 21 more variables: SurveyId <chr>, IndicatorId <chr>, IndicatorOrder <dbl>,
## #   IndicatorType <chr>, CharacteristicId <dbl>, CharacteristicOrder <dbl>,
## #   CharacteristicCategory <chr>, CharacteristicLabel <chr>,
## #   ByVariableId <chr>, ByVariableLabel <chr>, IsTotal <dbl>,
## #   IsPreferred <dbl>, SDRID <chr>, RegionId <lgl>, SurveyYearLabel <dbl>,
## #   SurveyType <chr>, DenominatorWeighted <dbl>, DenominatorUnweighted <dbl>,
## #   CILow <lgl>, CIHigh <lgl>, LevelRank <lgl>
```

Remove the first row(meta data)

Percentage Missing Values

##	Column	Missing_Percentage
## 1	ISO3	0%
## 2	DataId	0%
## 3	Indicator	0%
## 4	Value	0%
## 5	Precision	0%
## 6	DHS_CountryCode	0%
## 7	CountryName	0%
## 8	SurveyYear	0%
## 9	SurveyId	0%
## 10	IndicatorId	0%
## 11	IndicatorOrder	0%
## 12	IndicatorType	0%
## 13	CharacteristicId	0%
## 14	CharacteristicOrder	0%
## 15	CharacteristicCategory	0%
## 16	CharacteristicLabel	0%
## 17	ByVariableId	0%
## 18	ByVariableLabel	100%

```
## 19          IsTotal          0%
## 20      IsPreferred          0%
## 21          SDRID            0%
## 22          RegionId        100%
## 23      SurveyYearLabel      0%
## 24          SurveyType      0%
## 25      DenominatorWeighted  10.81%
## 26      DenominatorUnweighted 10.81%
## 27          CILow           100%
## 28          CIHigh           100%
## 29          LevelRank        100%
```

```
##check for unique values
```

```
## # A tibble: 29 x 3
##   column          n_unique sample_values
##   <chr>          <int> <chr>
## 1 IS03              1 ZAF
## 2 DataId           37 198690, 198687, 198688
## 3 Indicator        33 Children severely stunted, Children stunted, Mean h~
## 4 Value            36 9.8, 27.4, -1.1
## 5 Precision         2 1, 0
## 6 DHS_CountryCode   1 ZA
## 7 CountryName       1 South Africa
## 8 SurveyYear        1 2016
## 9 SurveyId          1 ZA2016DHS
## 10 IndicatorId      37 CN_NUTS_C_HA3, CN_NUTS_C_HA2, CN_NUTS_C_HAM
## # i 19 more rows
```

```
#Drop the countries only one unique value: reason, there is no useful information - county is also always za
```

```
#Assumed pattern, the missing values can be filled with the previous non missing value in the opposite attribute
```



Boxplot of Denominator Weighted

