

EDA_Anthropometry_National_02

Group K

2025-09-09

1. Loading the dataset

Dislay the first 5 rows

Table 1: First 5 rows of the dataset

ISO3	DataId	Indicator	Value	CountryName	SurveyYear
ZAF	198690	Children severely stunted	9.8	South Africa	2016
ZAF	198687	Children stunted	27.4	South Africa	2016
ZAF	198688	Mean height for age of children	-1.1	South Africa	2016
ZAF	597227	Children severely wasted	0.6	South Africa	2016
ZAF	597228	Children wasted	2.5	South Africa	2016

2. Data Overview

Summary statistics

Table 2: Summary Statistics for Key Numeric Variables in the Dataset

Variable	Statistic	Value
CIHigh	mean	NaN
CIHigh	sd	NA
CIHigh	min	Inf
CIHigh	max	-Inf
CILow	mean	NaN
CILow	sd	NA
CILow	min	Inf
CILow	max	-Inf
DenominatorUnweighted	mean	2432.9394
DenominatorUnweighted	sd	765.4351
DenominatorUnweighted	min	1449.0000
DenominatorUnweighted	max	3405.0000
DenominatorWeighted	mean	2336.7576
DenominatorWeighted	sd	747.1613
DenominatorWeighted	min	1384.0000
DenominatorWeighted	max	3272.0000

Variable	Statistic	Value
Value	mean	663.7676
Value	sd	1265.5961
Value	min	-1.1000
Value	max	3405.0000

3. Data Quality

Find duplicated values

```
## [1] 0
```

Inspect columns for missing/empty data as values and pecentages respectively

```
##          ISO3          DataId          Indicator
##          0              0              0
##          Value          Precision          DHS_CountryCode
##          0              0              1
##          CountryName      SurveyYear          SurveyId
##          0              0              0
##          IndicatorId      IndicatorOrder          IndicatorType
##          0              1              1
##          CharacteristicId  CharacteristicOrder  CharacteristicCategory
##          1              1              1
##          CharacteristicLabel  ByVariableId      ByVariableLabel
##          1              0              37
##          IsTotal          IsPreferred          SDRID
##          1              1              1
##          RegionId        SurveyYearLabel          SurveyType
##          38              1              1
##          DenominatorWeighted  DenominatorUnweighted          CILow
##          5              5              38
##          CIHigh          LevelRank
##          38              38
```

```
##          ISO3          DataId          Indicator
##          0.00          0.00          0.00
##          Value          Precision          DHS_CountryCode
##          0.00          0.00          2.63
##          CountryName      SurveyYear          SurveyId
##          0.00          0.00          0.00
##          IndicatorId      IndicatorOrder          IndicatorType
##          0.00          2.63          2.63
##          CharacteristicId  CharacteristicOrder  CharacteristicCategory
##          2.63          2.63          2.63
##          CharacteristicLabel  ByVariableId      ByVariableLabel
##          2.63          0.00          97.37
##          IsTotal          IsPreferred          SDRID
##          2.63          2.63          2.63
##          RegionId        SurveyYearLabel          SurveyType
```

##	100.00	2.63	2.63
##	DenominatorWeighted	DenominatorUnweighted	CILow
##	13.16	13.16	100.00
##	CIHigh	LevelRank	
##	100.00	100.00	

Check for unique values

##	ISO3	DataId	Indicator
##	2	38	34
##	Value	Precision	DHS_CountryCode
##	37	3	2
##	CountryName	SurveyYear	SurveyId
##	2	2	2
##	IndicatorId	IndicatorOrder	IndicatorType
##	38	38	4
##	CharacteristicId	CharacteristicOrder	CharacteristicCategory
##	3	3	3
##	CharacteristicLabel	ByVariableId	ByVariableLabel
##	3	2	2
##	IsTotal	IsPreferred	SDRID
##	2	2	38
##	RegionId	SurveyYearLabel	SurveyType
##	1	2	2
##	DenominatorWeighted	DenominatorUnweighted	CILow
##	8	8	1
##	CIHigh	LevelRank	
##	1	1	

4. Visualizations

