

EDA_Literacy_07

Group K

2025-09-09

1. Loading the dataset

Display the first 5 rows

Table 1: First 5 rows of the Access to Health Care Dataset (excluding first row)

ISO3	DataId	Indicator	Value	CountryName	SurveyYear
ZAF	563770	Women with secondary or higher education	11.8	South Africa	2016
ZAF	563771	Women who can read a whole sentence	76.2	South Africa	2016
ZAF	563772	Women who can read part of a sentence	8.2	South Africa	2016
ZAF	563773	Women who cannot read at all	3.5	South Africa	2016
ZAF	563769	Women for whom no card with required language was available	0.1	South Africa	2016

2. Data Overview

Summary statistics

```
## Warning: There were 4 warnings in 'summarise()'.  
## The first warning was:  
## i In argument: 'across(...)'.  
## Caused by warning in 'min()':  
## ! no non-missing arguments to min; returning Inf  
## i Run 'dplyr::last_dplyr_warnings()' to see the 3 remaining warnings.
```

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

Variable	Statistic	Value
CIHigh	mean	NaN
CIHigh	sd	NA

Variable	Statistic	Value
CIHigh	min	Inf
CIHigh	max	-Inf
CILow	mean	NaN
CILow	sd	NA
CILow	min	Inf
CILow	max	-Inf
DenominatorUnweighted	mean	7492.000
DenominatorUnweighted	sd	4438.040
DenominatorUnweighted	min	3179.000
DenominatorUnweighted	max	11805.000
DenominatorWeighted	mean	5858.000
DenominatorWeighted	sd	2733.001
DenominatorWeighted	min	3202.000
DenominatorWeighted	max	8514.000
Value	mean	1364.555
Value	sd	3203.736
Value	min	0.000
Value	max	11805.000

3. Data Quality

Missing values per column

##	IS03	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	20
##	IsTotal	IsPreferred	SDRID
##	1	1	1
##	RegionId	SurveyYearLabel	SurveyType
##	21	1	1
##	DenominatorWeighted	DenominatorUnweighted	CILow
##	3	3	21
##	CIHigh	LevelRank	
##	21	21	

Dupilcates

[1] 0

4. Visualizations

Literacy Status Distribution by Gender (2016)

