Economic Clustering Summary Report: South Africa 2016

Andy Kim

24 June, 2023

# Summary of Data

**Country Code-year:** ZA16

**Number of observations:** 11083

**Number of clusters:** 5

**Number of variables used:** 23

**Distance used:** Hamming

**Variables used in the algorithm:** hv201a, hv206, hv207, hv208, hv209, hv212, hv225, hv237, hv243b, hv243e, hv246, hv252, sh121g, sh121h, sh121i, sh121j, sh124b, water, toilet, floor, roof, cookfuel, wall

# Top Variable Groupings (Sorted by ASW)

| Group | ASW | Var. 1 | Var. 2 | Var. 3 | Var. 4 |
| --- | --- | --- | --- | --- | --- |
| 1 | 0.9411 | has electricity | has refrigerator | vacuum cleaner or floor polisher | washing machine |
| 2 | 0.9408 | has electricity | has refrigerator | vacuum cleaner or floor polisher | microwave oven |
| 3 | 0.9376 | has electricity | has a computer | vacuum cleaner or floor polisher | roof |
| 4 | 0.9375 | has electricity | vacuum cleaner or floor polisher | microwave oven | electric/gas stove |
| 5 | 0.9369 | has electricity | vacuum cleaner or floor polisher | electric/gas stove | roof |
| 6 | 0.9364 | has electricity | has a computer | washing machine | roof |
| 7 | 0.9358 | has refrigerator | vacuum cleaner or floor polisher | electric/gas stove | roof |
| 8 | 0.9355 | has electricity | has a computer | microwave oven | roof |
| 9 | 0.9355 | has electricity | microwave oven | washing machine | roof |
| 10 | 0.935 | has electricity | has refrigerator | vacuum cleaner or floor polisher | roof |

# Marginal Distributions

| Variable | Description | % Time in Top Clusters | Distribution |
| --- | --- | --- | --- |
| hv206 | has electricity | 90% | Binary, 90% 1s (or yes) |
| hv209 | has refrigerator | 40% | Binary, 74.5% 1s (or yes) |
| hv243e | has a computer | 30% | Binary, 19.6% 1s (or yes) |
| sh121g | vacuum cleaner or floor polisher | 70% | Binary, 12.2% 1s (or yes) |
| sh121h | microwave oven | 40% | Binary, 49% 1s (or yes) |
| sh121i | electric/gas stove | 30% | Binary, 83.2% 1s (or yes) |
| sh121j | washing machine | 30% | Binary, 34.8% 1s (or yes) |
| roof | roof | 70% | Categorical, 0 (finished) = 95.6%,   1 (natural) = 2.2%,   2 (other) = 0.1%,   3 (rudimentary) = 2% |

# Cluster #1 Configuration

| Cluster Group | has electricity | has refrigerator | vacuum cleaner or floor polisher | washing machine | Proportion (%) |
| --- | --- | --- | --- | --- | --- |
| 1 | 1 | 0 | 0 | 0 | 15.37 |
| 1 | 1 | 0 | 1 | 0 |  |
| 2 | 0 | 0 | 0 | 1 | 23.85 |
| 2 | 0 | 1 | 0 | 1 |  |
| 2 | 1 | 0 | 0 | 1 |  |
| 2 | 1 | 1 | 0 | 1 |  |
| 3 | 0 | 0 | 0 | 0 | 9.16 |
| 4 | 0 | 1 | 1 | 1 | 10.91 |
| 4 | 1 | 0 | 1 | 1 |  |
| 4 | 1 | 1 | 1 | 1 |  |
| 5 | 0 | 1 | 0 | 0 | 40.71 |
| 5 | 1 | 1 | 0 | 0 |  |
| 5 | 1 | 1 | 1 | 0 |  |

# Validation Tables

## a.1) Using Children Deceased (Sorted by proportion of 0%)

| Cluster ID/Children Deceased | 0% | 1-33% | 34-66% | 67+% |
| --- | --- | --- | --- | --- |
| 1\* | 630 (95.9%) | 14 (2.1%) | 12 (1.8%) | 1 (0.2%) |
| 2 | 2,231 (93.6%) | 82 (3.4%) | 54 (2.3%) | 17 (0.7%) |
| 5 | 3,399 (91.6%) | 169 (4.6%) | 103 (2.8%) | 38 (1.0%) |
| 4 | 907 (90.1%) | 48 (4.8%) | 36 (3.6%) | 16 (1.6%) |
| 3 | 665 (87.8%) | 60 (7.9%) | 24 (3.2%) | 8 (1.1%) |
| Total | 7,832 (92.0%) | 373 (4.4%) | 229 (2.7%) | 80 (0.9%) |
| \*The chi-squared p-value is 0 | | | | |

## a.2) Aggregating proportions greater than 0%

| Cluster ID/Children Deceased | 0% | >0% |
| --- | --- | --- |
| 1\* | 630 (95.9%) | 27 (4.1%) |
| 2 | 2,231 (93.6%) | 153 (6.4%) |
| 5 | 3,399 (91.6%) | 310 (8.4%) |
| 4 | 907 (90.1%) | 100 (9.9%) |
| 3 | 665 (87.8%) | 92 (12.2%) |
| Total | 7,832 (92.0%) | 682 (8.0%) |
| \*The chi-squared p-value is 0 | | |

## b) Using Individual Education Level Attained (Sorted by weighted average by row)

| Cluster ID/Education | 0 | 1 | 2 | 3 | 4 | 5 | W. Avg. |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1\*a | 2 (0.3%) | 5 (0.8%) | 6 (0.9%) | 192 (29.2%) | 235 (35.8%) | 217 (33.0%) | 3.98 |
| 2 | 34 (1.4%) | 83 (3.5%) | 53 (2.2%) | 1,145 (48.0%) | 724 (30.4%) | 345 (14.5%) | 3.46 |
| 5 | 69 (1.9%) | 227 (6.1%) | 145 (3.9%) | 2,042 (55.1%) | 937 (25.3%) | 289 (7.8%) | 3.19 |
| 4 | 35 (3.5%) | 108 (10.7%) | 68 (6.8%) | 586 (58.2%) | 189 (18.8%) | 21 (2.1%) | 2.84 |
| 3 | 50 (6.6%) | 101 (13.3%) | 66 (8.7%) | 444 (58.7%) | 87 (11.5%) | 9 (1.2%) | 2.59 |
| Total | 190 (2.2%) | 524 (6.2%) | 338 (4.0%) | 4,409 (51.8%) | 2,172 (25.5%) | 881 (10.3%) | 3.23 |
| \*The chi-squared p-value is 0 | | | | | | | |
| a0=none,1=incomplete primary, 2=primary, 3=incomplete secondary, 4=secondary, 5=higher | | | | | | | |

## c) Using Primary Healthcare Source (Sorted by % enrolled in public healthcare [ascending order])

| Cluster ID/Primary Healthcare Source | 0 | 1 | 2 |
| --- | --- | --- | --- |
| 1\*a | 149 (50.9%) | 120 (41.0%) | 24 (8.2%) |
| 2 | 885 (79.9%) | 127 (11.5%) | 96 (8.7%) |
| 5 | 1,500 (85.6%) | 131 (7.5%) | 122 (7.0%) |
| 4 | 437 (89.5%) | 17 (3.5%) | 34 (7.0%) |
| 3 | 336 (94.4%) | 4 (1.1%) | 16 (4.5%) |
| Total | 3,307 (82.7%) | 399 (10.0%) | 292 (7.3%) |
| \*The chi-squared p-value is 0 | | | |
| a0=public/government, 1=private, 2=other | | | |