Country: **Namibia**

Year: **2013**

Number of Observations: **9849**

Number of Clusters: **5**

Number of variables used: **30**

Distance used: **Hamming**

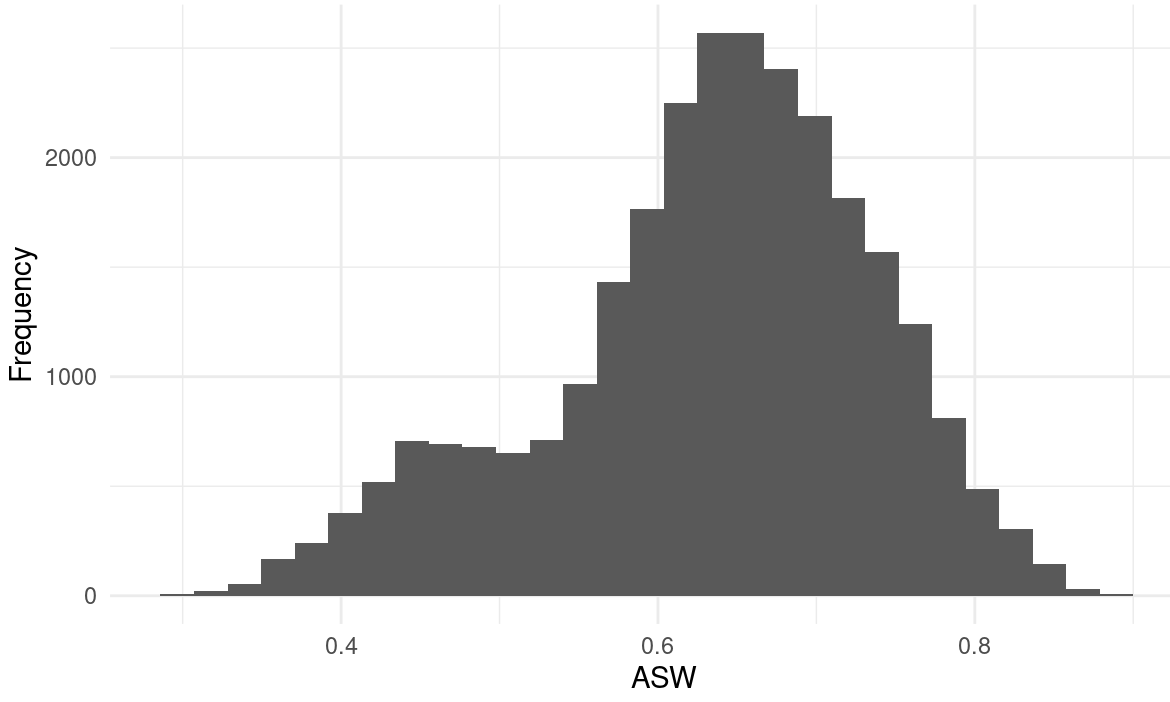
**Variables used in the algorithm:** hv206,hv207,hv208,hv209,hv210,hv212,hv227,hv237,hv243a,hv243b,hv244,hv246,hv247,hv252,hv253,sh110g,h,i,k,l,n,sh141,sh147,water,toilet,floor,roof,cookfuel,wall

Binary variables with less than 10% class imbalance excluded, other variables excluded for collinearity: hv253a through hv253z, sh132ba through sh123ex

**Summary of Top Cluster Configurations (ranked by ASW)**

| **Cluster #** | **ASW** | **Variable 1** | **Variable 2** | **Variable 3** | **Variable 4** |
| --- | --- | --- | --- | --- | --- |
| **1** | 0.902276259093039 | hv206 | hv209 | hv243a | sh110i |
| **2** | 0.891360977492431 | hv209 | hv237 | hv243a | sh110i |
| **3** | 0.88826516191456 | hv206 | hv209 | hv243a | sh110g |
| **4** | 0.887205360077326 | hv237 | hv243a | sh110h | sh110i |
| **5** | 0.883733283297226 | hv209 | hv243a | sh110g | sh110i |
| **6** | 0.880840691692939 | hv209 | hv243a | hv247 | sh110i |
| **7** | 0.88078812376457 | hv206 | hv237 | hv243a | sh110i |
| **8** | 0.88066562678344 | hv243a | sh110g | sh110h | sh110i |
| **9** | 0.880343698200144 | hv209 | hv243a | sh110h | sh110i |
| **10** | 0.876198772395493 | hv243a | hv247 | sh110h | sh110i |

**Distribution of ASW values in all clusters in Namibia**



**Marginal Distributions**

| Variable | Description | % time in top clusters | Distribution |
| --- | --- | --- | --- |
| hv206 | Has electricity | 30% | Binary,48.7% 1s (or yes) |
| hv209 | Has refrigerator | 60% | Binary,41.7% 1s (or yes) |
| hv237 | Anything done to water to make safe to drink | 30% | Binary,8.6% 1s (or yes) |
| hv243a | Has mobile telephone | 100% | Binary,87.2% 1s (or yes) |
| hv247 | Has bank account | 20% | Binary,73.7% 1s (or yes) |
| sh110g | computer/laptop | 30% | Binary,18.8% 1s (or yes) |
| sh110h | stove | 40% | Binary,46.8% 1s (or yes) |
| sh110i | microwave | 90% | Binary,23.8% 1s (or yes) |

**Summary of variable distributions in top clusters**

Currently our method is choosing five distinct clusters of individuals within each cluster variable configuration. Here are the medioids for each of these five clusters:

**Cluster #1 Configuration**

| **Config#** | **Node** | **hv206** | **hv209** | **hv243a** | **sh110i** | **Proportion** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | 1 | 0 | 0 | 0 | 0 | 11.90% |
| **2** | 1 | 0 | 0 | 0 | 1 |
| **3** | 1 | 0 | 1 | 0 | 0 |
| **4** | 1 | 1 | 0 | 0 | 0 |
| **5** | 2 | 1 | 0 | 1 | 0 | 8.02% |
| **6** | 3 | 0 | 0 | 1 | *0* | 40.36 |
| **7** | 3 | 0 | 0 | 1 | 1 |
| **8** | 3 | 0 | 1 | 1 | 0 |
| **9** | 4 | 1 | *1* | 0 | 0 | 16.04% |
| **10** | 4 | 1 | *1* | 1 | 0 |
| **11** | 5 | 0 | 1 | 0 | 1 | 23.69% |
| **12** | 5 | 0 | 1 | 1 | 1 |
| **13** | 5 | 1 | 0 | 0 | 1 |
| **14** | 5 | 1 | 0 | 1 | 1 |
| **15** | 5 | 1 | 1 | 0 | 1 |
| **16** | 5 | 1 | 1 | 1 | 1 |

**Validation of Top Cluster with Chi-Sq Distribution**

Note that in the tables, the rows are the validation variables, and the columns are the node number from the cluster distributions in the table above.

