# Populating a Table

As an exercise we decided to populate a table from the DHS data with these concepts for rows and columns.

**HIV Cases**

|  |  |  |
| --- | --- | --- |
| Country | Male | Female |
| Angola |   |   |
| Burundi |   |   |
| Ethiopia |   |   |
| Lesotho |   |   |
| Malawi |   |   |
| Mozambique |   |   |
| Namibia |   |   |
| Rwanda |   |   |
| Zambia |   |   |
| Zimbabwe |   |   |

### Codes for Sex

One of the team knew that the Australian Bureau of Statistics had a classification scheme for sex. A google search for “abs sex classification” yielded the page:

[ttps://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1200.0.55.012Main%20Features212016](https://www.abs.gov.au/ausstats/abs%40.nsf/Latestproducts/1200.0.55.012Main%20Features212016)

Table 1 below describes the category codes, labels, and definitions of the sex classification and the gender classification.

**TABLE 1. THE SEX AND GENDER STANDARD CLASSIFICATIONS AND CODE STRUCTURES**

|  |
| --- |
|  |
| Preferred Code | Alternate Code | Label | Definition |
|  |
| **The Sex Standard Classification** |
|  |
| 1 | M | Male | Persons who have male or predominantly masculine biological characteristics, or male sex assigned at birth. |
| 2 | F | Female | Persons who have female or predominantly feminine biological characteristics, or female sex assigned at birth. |
| 3 | X | Other | Persons who have mixed or non-binary biological characteristics (if known), or a non-binary sex assigned at birth. |
|  |
| **The Gender Standard Classification** |
|  |
| 1 | M | Male  | Adults who identify themselves as men, and children who identify themselves as boys. |
| 2 | F | Female | Adults who identify themselves as women, and children who identify themselves as girls. |
| 3 | X | Other | Adults and children who identify as non-binary, gender diverse, or with descriptors other than man/boy or woman/girl. |
|  |

These were hand coded into our application.

### Codes for Countries

The DHS data has the country categories coded, but we wanted to ensure standard codes for the countries and sex categories. The original Stata file represents the country variable as a three character string like “BU7”. The metadata for the DHS files just describes this as a three character string, without any identification of the country associated with the first two characters. The numeric character is the vintage of the dataset.

We wanted to find a source for standard country codes, preferably ISO codes. We first searched Google for “iso country code”. We found this unofficial source:

<https://en.wikipedia.org/wiki/List_of_ISO_3166_country_codes>

and could have web scraped it but we wanted to find an open official source, either as a table or API. ISO codes are only available for a payment.

After considerable exploration we found there is a crosswalk for country codes available at:

<http://api.dhsprogram.com/rest/dhs/countries>

The Entry for Burundi is:

Africa","WHO\_CountryCode":"BI","FIPS\_CountryCode":"BY","ISO2\_CountryCode":"BI","ISO3\_CountryCode":"BDI","RegionOrder":13,"DHS\_CountryCode":"BU","CountryName":"Burundi","UNICEF\_CountryCode":"BRD","UNSTAT\_CountryCode":"BDI","RegionName":"Sub-Saharan Africa"},{"UNAIDS\_CountryCode":"KHM","SubregionName":"Southeast

Note that the ISO code for Burundi is “BI” but DHS uses “BU”. There is no country associated with an ISO code of “BU”. This site provides the information necessary to develop the mapping from DHS codes to ISO codes.

Gathering the ISO codes and the codes used in DHS reveals that the following transformation would be necessary to map the DHS codes into ISO codes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **DHS** |  | **ISO 3166-1** |
| Angola | AO | --> | AO |
| Burundi | BU | --> | BI |
| Ethiopia | ET | --> | ET |
| Lesotho | LS | --> | LS |
| Malawi | MW | --> | MW |
| Mozambique | MZ | --> | MZ |
| Namibia | NM | --> | NA |
| Rwanda | RW | --> | RW |
| Zambia | ZM | --> | ZM |
| Zimbabwe | ZW | --> | ZW |

Once the codes and a mapping were developed, a team member wrote R code to collect the data to populate the desired table.