Data Profile - World Health Organization API

Kurt Wydrinski

11/22/2019

Source

During the formation of the United Nations in 1945, one of the key objectives for its ongoing mission was to foster cooperation between countries to address many social well-being issues, including health.[1] This objective was taken on as part of the World Health Organization (WHO) which was established on April 7, 1948.[2] The WHO continues this mission today with 194 countries cooperating at various levels to promote global health concerns. In doing so, the WHO regularly collects data from its members that are related to health issues. The data contains various time series regarding diseases, illnesses, economics, social demographics, etc. This data is collectively captured in the WHO's online database *Global Health Observatory (GHO)*.[3]

API Overview

The R package `WHO`, provides a simple API to access the GHO. It only provides two functions: `get_codes()` and `get_data()`. Inside the GHO, each time series that exists is identifed by a label. Each label is a code that uniquely identifies the series. These labels are then used a a parameter to `get_data()` to retrieve the time series observations.

```r
# install.packages("WHO")
library(WHO)
```

The code below uses the `extra` parameter to download all metadata available for the GHO codes.

```r
who_codes <- get_codes(extra = TRUE)
glimpse(who_codes)
```

```
## Observations: 3,287
## Variables: 9
## $ label          <chr> "MDG_0000000001", "MDG_0000000003", "MDG_0000000005"...
## $ display        <chr> "Infant mortality rate (probability of dying between...
## $ url            <chr> "https://www.who.int/data/gho/indicator-metadata-reg...
## $ display_fr     <chr> "Taux de mortalité des nourrissons (probabilité de d...
## $ display_es     <chr> "Tasa de mortalidad de menores de 1 año (probabilida...
## $ definition_xml <chr> "http://apps.who.int/gho/indicatorregistryservice/pu...
## $ category       <chr> "Mortality and global health estimates", "Sustainabl...
## $ imr_id         <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, "16", NA...
## $ renderer_id    <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
```

There are 3,287 indicators or time series in this dataset. The `category` variable is a WHO grouping of the indicators into 51 sets of indicators as follows:

```r
who_codes %>%
  group_by(category) %>%
  summarise(indicator_count = n()) %>%
  knitr::kable()
```

---

[1] See https://www.history.com/this-day-in-history/the-united-nations-is-born.

[2] See https://www.who.int/about/who-we-are/history.

[3] See https://www.who.int/gho/about/en/.

category	indicator__count
AMR GLASS Coordination	4
AMR GLASS Quality assurance	5
AMR GLASS Surveillance	3
Child health	37
Demographic and socioeconomic statistics	19
Essential health technologies	10
FINANCIAL PROTECTION	28
Global Observatory for eHealth (GOe)	92
Health Equity Monitor	64
Health financing	13
Health systems	381
Health workforce	24
HIV/AIDS and other STIs	37
Infectious diseases	31
Infrastructure	6
Injuries and violence	59
Insecticide resistance	9
International Health Regulations (2005) monitoring framework	13
Malaria	126
Medical equipment	8
Millennium Development Goals (MDGs)	15
Mortality and global health estimates	58
Negelected tropical diseases	9
Neglected tropical diseases	17
Neglected Tropical Diseases	17
Noncommunicable diseases	50
Noncommunicable diseases and mental health	2
Noncommunicable diseases CCS	82
Nutrition	27
Oral health	1
Public health and environment	177
RSUD: GOVERNANCE, POLICY AND FINANCING : PREVENTION	3
RSUD: GOVERNANCE, POLICY AND FINANCING: FINANCING	7
RSUD: GOVERNANCE, POLICY AND FINANCING: TREATMENT	12
RSUD: HUMAN RESOURCES	8
RSUD: INFORMATION SYSTEMS	6
RSUD: SERVICE ORGANIZATION AND DELIVERY: PHARMACOLOGICAL TREATMENT	11
RSUD: SERVICE ORGANIZATION AND DELIVERY: PREVENTION PROGRAMS AND PROVIDERS	5
RSUD: SERVICE ORGANIZATION AND DELIVERY: SCREENING AND BRIEF INTERVENTIONS	4
RSUD: SERVICE ORGANIZATION AND DELIVERY: SPECIAL PROGRAMMES AND SERVICES	12
RSUD: SERVICE ORGANIZATION AND DELIVERY: TREATMENT CAPACITY AND TREATMENT COVERAGE	2
RSUD: SERVICE ORGANIZATION AND DELIVERY: TREATMENT SECTORS AND PROVIDERS	8
RSUD: YOUTH	3
Substance use and mental health	1121
Sustainable development goals	28
Tobacco	3
Tuberculosis	68
Universal Health Coverage	16
Urban health	32
World Health Statistics	67
NA	447

Note the following about the above categories.

1. There are 447 indicators that do not have a category assigned by the WHO. These should be considered as a category called `Uncategorized` or `No category`.
2. The naming convention is not consistent across categories. For example, some use single words, phrases, all capital letters, irregular case, etc. Category names should be somewhat standardized while retaining the original meaning in the dataset.
3. A number of categories appear to be fragmented. For example, there are three premutations of *Neglected Tropical Diseases* that need to be combined into a single category.

```
who_codes %>%
  filter(str_detect(display, "influenza"))
```

```
## # A tibble: 1 x 9
##   label display url   display_fr display_es definition_xml category imr_id
##   <chr> <chr>   <chr> <chr>      <chr>      <chr>          <chr>    <chr>
## 1 WHS3~ H5N1 i~ ""    <NA>       <NA>       http://apps.w~ Infecti~ 53
## # ... with 1 more variable: renderer_id <chr>
```