R Programming

1st Assignment

Mateus Rennó Santos

December 9th, 2016

1. **Dataset**

Title: World Mortality Database

Institution: World Health Organization

Last Updated: Sep 15th, 2016

Link: <http://www.who.int/healthinfo/statistics/mortality_rawdata/en/>

Description:

Number of deaths recorded by member countries, including the cause of death, the age at death, and the living population at each age for the calculation of population rates. Deaths are classified according to the International Classification of Diseases (ICD), a standardize catalog of causes of death used by medical practitioners globally, and which includes a specific designation for intentional homicides. Data is available by sex, between 1950 and 2014.

1. **Variables of Interest (extracted from the**
   1. Mortality Data

|  |  |  |
| --- | --- | --- |
| **Column name** | **Content** | **Type** |
| Country | Country | Text |
| Admin1 | Specified region/Category pertinent to each country | Text |
| Year | Year to which data refer | Time |
| List | List of ICD revision used | Factor |
| Cause | Cause of death | Factor/Text |
| Sex | 1 male, 2 female and 9 sex unspecified | Factor |
| Frmat | Age-group format for breakdown of deaths at 0-95+ yrs | Factor |
| Deaths1 | Deaths at all ages | Number |
| Deaths2 | Deaths at age 0 year | Number |
| Deaths3 | Deaths at age 1 year | Number |
| Deaths4 | Deaths at age 2 years | Number |
| Deaths5 | Deaths at age 3 years | Number |
| Deaths6 | Deaths at age 4 years | Number |
| Deaths7 | Deaths at age 5-9 years | Number |
| Deaths8 | Deaths at age 10-14 years | Number |
| Deaths9 | Deaths at age 15-19 years | Number |
| Deaths10 | Deaths at age 20-24 years | Number |
| Deaths11 | Deaths at age 25-29 years | Number |
| Deaths12 | Deaths at age 30-34 years | Number |
| Deaths13 | Deaths at age 35-39 years | Number |
| Deaths14 | Deaths at age 40-44 years | Number |
| Deaths15 | Deaths at age 45-49 years | Number |
| Deaths16 | Deaths at age 50-54 years | Number |
| Deaths17 | Deaths at age 55-59 years | Number |
| Deaths18 | Deaths at age 60-64 years | Number |
| Deaths19 | Deaths at age 65-69 years | Number |
| Deaths20 | Deaths at age 70-74 years | Number |
| Deaths21 | Deaths at age 75-79 years | Number |
| Deaths22 | Deaths at age 80-84 years | Number |
| Deaths23 | Deaths at age 85-89 years | Number |
| Deaths24 | Deaths at age 90-94 years | Number |
| Deaths25 | Deaths at age 95 years and above | Number |
| Deaths26 | Deaths at age unspecified | Number |

1. Population Data

|  |  |  |
| --- | --- | --- |
| **Column name** | **Content** | **Type** |
| Country | Country | Text |
| Admin1 | Specified region/Category pertinent to each country | Text |
| Year | Year to which data refer | Time |
| Sex | 1 male, 2 female | Factor |
| Frmat | Age-group format for breakdown of deaths at 0-95+ yrs | Factor/Text |
| Pop1 | Population at all ages | Factor |
| Pop2 | Population at age 0 year | Factor |
| Pop3 | Population at age1 year | Number |
| Pop4 | Population at age 2 years | Number |
| Pop5 | Population at age 3 years | Number |
| Pop6 | Population at age 4 years | Number |
| Pop7 | Population at age 5-9 years | Number |
| Pop8 | Population at age 10-14 years | Number |
| Pop9 | Population at age 15-19 years | Number |
| Pop10 | Population at age 20-24 years | Number |
| Pop11 | Population at age 25-29 years | Number |
| Pop12 | Population at age 30-34 years | Number |
| Pop13 | Population at age 35-39 years | Number |
| Pop14 | Population at age 40-44 years | Number |
| Pop15 | Population at age 45-49 years | Number |
| Pop16 | Population at age 50-54 years | Number |
| Pop17 | Population at age 55-59 years | Number |
| Pop18 | Population at age 60-64 years | Number |
| Pop19 | Population at age 65-69 years | Number |
| Pop20 | Population at age 70-74 years | Number |
| Pop21 | Population at age 75-79 years | Number |
| Pop22 | Population at age 80-84 years | Number |
| Pop23 | Population at age 85-89 years | Number |
| Pop24 | Population at age 90-94 years | Number |
| Pop25 | Population at age 95 years and over | Number |
| Pop26 | Population at age unspecified | Number |
| Lb | Live births | Number |

1. Why is the data intesting to me:

These are the main datasets that I am using for my dissertation. I am attempting to generate a clean count of homicide rate by countries since the 1950’s, for as many countries as possible. To do so, I will combine data from multiple sources into one larger file. I have already started doing this work using Stata. My goal now is to export that initial work to R, and have all my code organized in a way that I can latter publish, alongside my data. Ultimately, I intend on evaluating the impact of the age structure of a population on crime rates over time.

1. Research Questions
   1. What are the most dangerous/safe countries around the world?
   2. What is the average age distribution of homicide victims around the world?
   3. What is the association between the proportion of individuals between 15 to 25 years of age and country-level homicides?
2. Other/Broader questions:
   1. What causes crime to change within society?
   2. How stable/unstable are homicides within countries over time?