ma615_midterm project

xiaofan xia

2019/10/15

The dataset is from the world bank open data. This is a dataset of the health nutrition and population. It contains 259 countries and 59 years with 10 series of variables including adolescent fertility rate (births per 1,000 women ages 15-19), adults (ages 15+) and chidren (ages 0-14) living with HIV, adults and children newly infected with HIV, adults living with HIV, adults newly infected with HIV, age at first marriage (female), age at first marriage (male), age dependency ratio (% of working-age population), age dependency ratio (old),and age dependency ratio (young).

Firstly, I read in the data and clean the data.

```
## Warning: attributes are not identical across measure va
## they will be dropped
```