

# Gapminder\_Exploration

*Ning Shen*

*2019/9/14*

## Simple Exploration

1. The `head` function demonstrates the first 6 rows of the dataset by default. It shows you what the dataset looks like.

```
head(gapminder)
```

```
## # A tibble: 6 x 6
##   country    continent  year lifeExp      pop gdpPercap
##   <fct>      <fct>    <int>  <dbl>    <int>    <dbl>
## 1 Afghanistan Asia      1952   28.8  8425333    779.
## 2 Afghanistan Asia      1957   30.3  9240934    821.
## 3 Afghanistan Asia      1962   32.0 10267083    853.
## 4 Afghanistan Asia      1967   34.0 11537966    836.
## 5 Afghanistan Asia      1972   36.1 13079460    740.
## 6 Afghanistan Asia      1977   38.4 14880372    786.
```

2. The `str` function outputs the categories of each column/variable in this dataset.

```
str(gapminder)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':   1704 obs. of  6 variables:
## $ country   : Factor w/ 142 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ continent : Factor w/ 5 levels "Africa","Americas",...: 3 3 3 3 3 3 3 3 3 3 ...
## $ year      : int   1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 ...
## $ lifeExp   : num   28.8 30.3 32 34 36.1 ...
## $ pop       : int   8425333 9240934 10267083 11537966 13079460 14880372 12881816 13867957 16317921 22...
## $ gdpPercap: num    779 821 853 836 740 ...
```

3. Function `nrow` gives us the number of rows of the dataset while `ncol` gives the number of columns.

```
nrow(gapminder)
```

```
## [1] 1704
```

```
ncol(gapminder)
```

```
## [1] 6
```

4. The `summary` function applying on a dataset outputs the summary information of each column, which is counts for factor variables or quantiles for numerical variables.

```
summary(gapminder)
```

```
##           country      continent      year      lifeExp
## Afghanistan: 12 Africa :624   Min.   :1952   Min.   :23.60
## Albania     : 12 Americas:300 1st Qu.:1966 1st Qu.:48.20
## Algeria     : 12 Asia    :396  Median :1980  Median :60.71
```

```

## Angola      : 12   Europe :360   Mean   :1980   Mean    :59.47
## Argentina   : 12   Oceania : 24   3rd Qu.:1993   3rd Qu.:70.85
## Australia   : 12                   Max.    :2007   Max.     :82.60
## (Other)     :1632
##      pop      gdpPercap
## Min.      :6.001e+04   Min.      : 241.2
## 1st Qu.:2.794e+06   1st Qu.: 1202.1
## Median :7.024e+06   Median : 3531.8
## Mean     :2.960e+07   Mean     : 7215.3
## 3rd Qu.:1.959e+07   3rd Qu.: 9325.5
## Max.     :1.319e+09   Max.     :113523.1
##

```