

# HW01\_gapminder

MINSI SUNG

12/09/2019

## Contents

Explore gapminder	1
Asia	2

## Explore gapminder

Let's load the dataset `gapminder` which is available in R.

Summary statistics for the dataset:

```
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
##
```

The dataset contains 1704 rows and 6 columns.

All the variables we have:

```
names(gapminder)
```

```
## [1] "country" "continent" "year" "lifeExp" "pop" "gdpPercap"
```

The number of unique countries in the dataset is 142.

The number of years in the dataset is 12

The mean life expectancy over all countries for all years is 59.4744394

- The highest life expectancy in all the countries:

```
gapminder[which(gapminder$lifeExp == max(gapminder$lifeExp)),]
```

```
## # A tibble: 1 x 6
##   country continent  year lifeExp      pop gdpPercap
##   <fct>   <fct>      <int>  <dbl>    <int>    <dbl>
## 1 Japan   Asia        2007   82.6 127467972  31656.
```

The mean Population over all countries for all years is  $2.9601212 \times 10^7$

- The highest Population in all the countries:

```
gapminder[which(gapminder$pop == max(gapminder$pop)),]
```

```
## # A tibble: 1 x 6
##   country continent  year lifeExp      pop gdpPercap
##   <fct>   <fct>      <int>  <dbl>    <int>    <dbl>
## 1 China   Asia        2007   73.0 1318683096  4959.
```

The mean GDP per capita over all countries for all years is 7215.3270812

- The highest GDP per capita in all the countries:

```
gapminder[which(gapminder$gdpPercap == max(gapminder$gdpPercap)),]
```

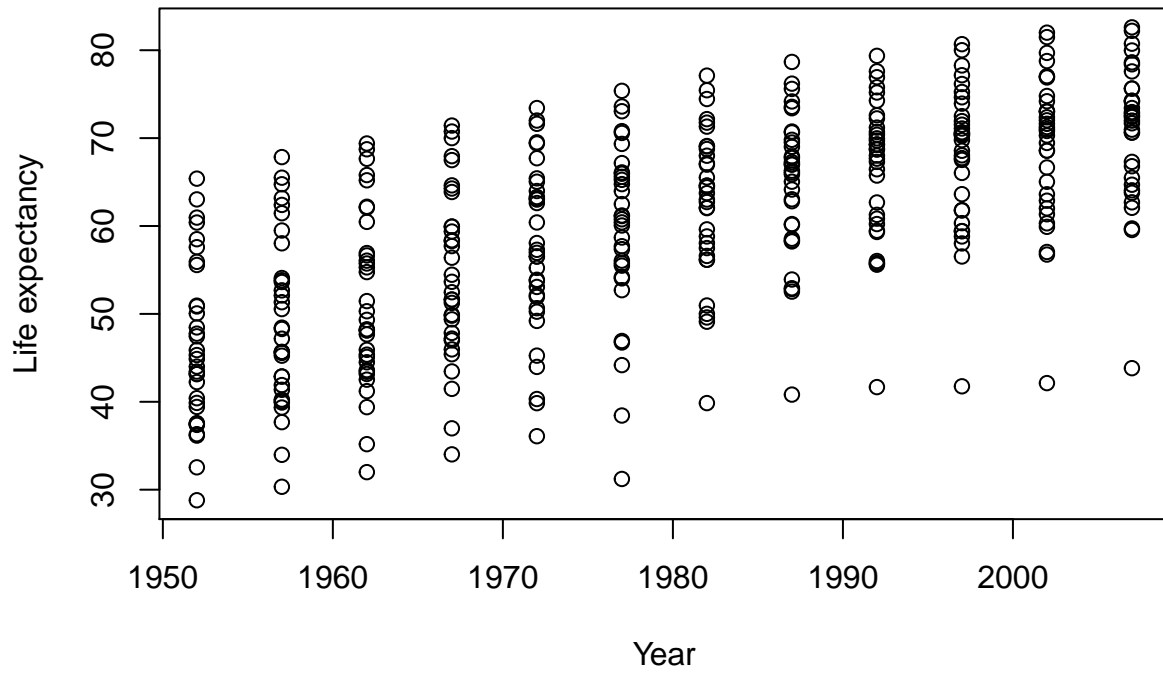
```
## # A tibble: 1 x 6
##   country continent  year lifeExp      pop gdpPercap
##   <fct>   <fct>      <int>  <dbl>    <int>    <dbl>
## 1 Kuwait   Asia        1957   58.0  212846  113523.
```

## Asia

Now let's check the gapminder data for Asia.

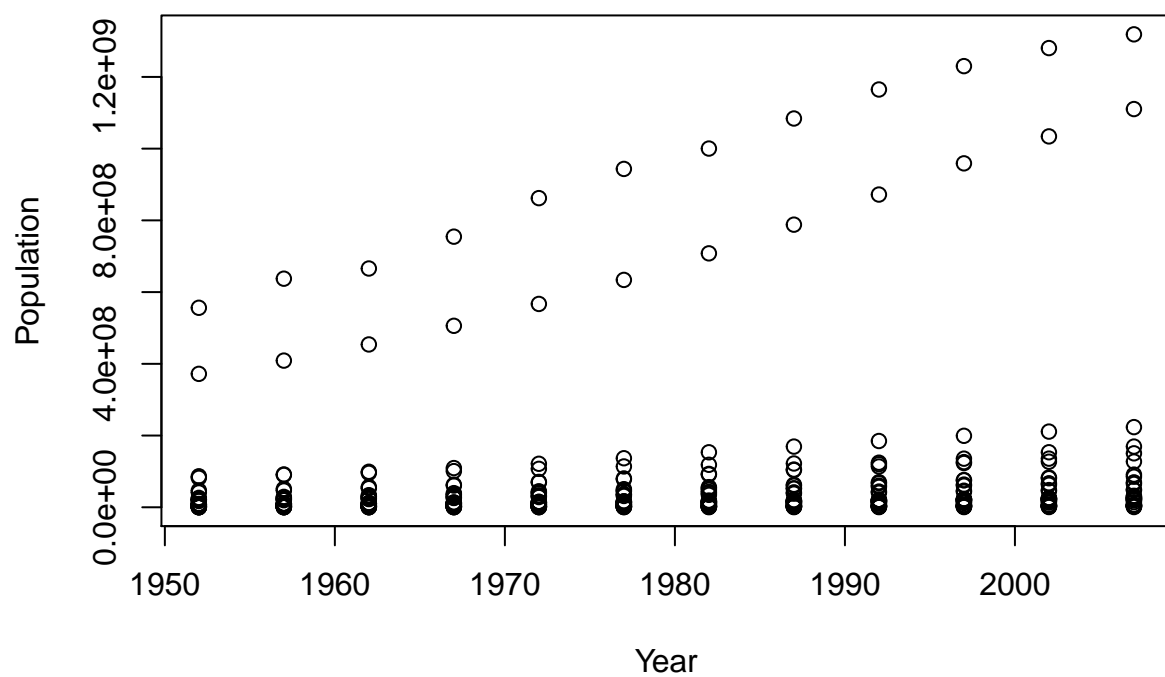
```
plot(gapminder$year[gapminder$continent=="Asia"], gapminder$lifeExp[gapminder$continent=="Asia"], ylab = "Life Expectancy")
```

## Life expectancy in Asia by year



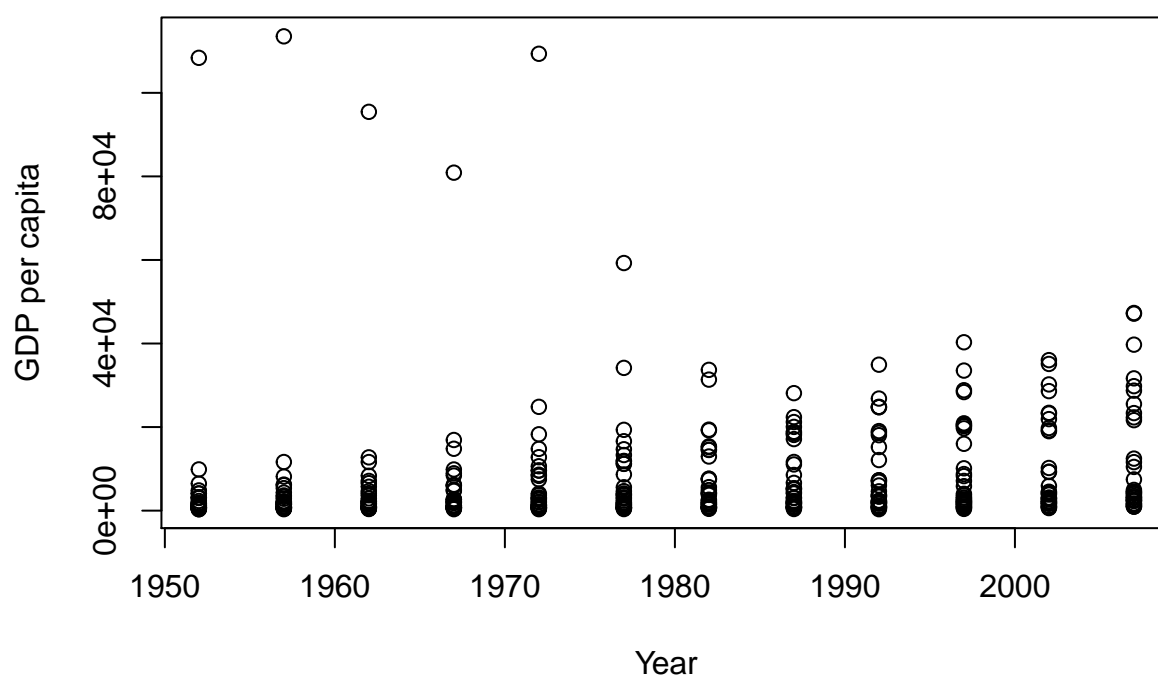
```
plot(gapminder$year[gapminder$continent=="Asia"], gapminder$pop[gapminder$continent=="Asia"], ylab = "P
```

## Population in Asia by year



```
plot(gapminder$year[gapminder$continent=="Asia"], gapminder$gdpPerCap[gapminder$continent=="Asia"], ylab="Population", xlab="Year", main="Population in Asia by year")
```

## GDP per capita in Asia by year



- The highest life expectancy in Asia:

```
gapminder[which(gapminder$lifeExp == max(gapminder$lifeExp[gapminder$continent=="Asia"])),]
```

```
## # A tibble: 1 x 6
##   country continent  year lifeExp      pop gdpPercap
##   <fct>   <fct>     <int>  <dbl>    <int>    <dbl>
## 1 Japan   Asia       2007   82.6 127467972  31656.
```

- The highest population in Asia:

```
gapminder[which(gapminder$pop == max(gapminder$pop[gapminder$continent=="Asia"])),]
```

```
## # A tibble: 1 x 6
##   country continent  year lifeExp      pop gdpPercap
##   <fct>   <fct>     <int>  <dbl>    <int>    <dbl>
## 1 China   Asia       2007   73.0 1318683096  4959.
```

- The highest GDP per capita in Asia:

```
gapminder[which(gapminder$gdpPercap == max(gapminder$gdpPercap[gapminder$continent=="Asia"])),]
```

```
## # A tibble: 1 x 6
##   country continent  year lifeExp      pop gdpPercap
##   <fct>   <fct>     <int>  <dbl>    <int>    <dbl>
## 1 Kuwait  Asia       1957   58.0  212846 113523.
```

### Summary statistics for Asia:

- Population:

```
summary(gapminder$pop[gapminder$continent=="Asia"])
```

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## 1.204e+05 3.844e+06 1.453e+07 7.704e+07 4.630e+07 1.319e+09
```

- Life expectancy:

```
summary(gapminder$lifeExp[gapminder$continent=="Asia"])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   28.80   51.43   61.79   60.06   69.51   82.60
```

- GDP per capita:

```
summary(gapminder$gdpPercap[gapminder$continent=="Asia"])
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     331   1057   2647   7902   8549  113523
```