Mini_Project 1

Xindi Lu, Jingyu Xie

October 2, 2019

```
library(gapminder)
```

```
## Warning: package 'gapminder' was built under R version 3.5.3
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.5.1
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

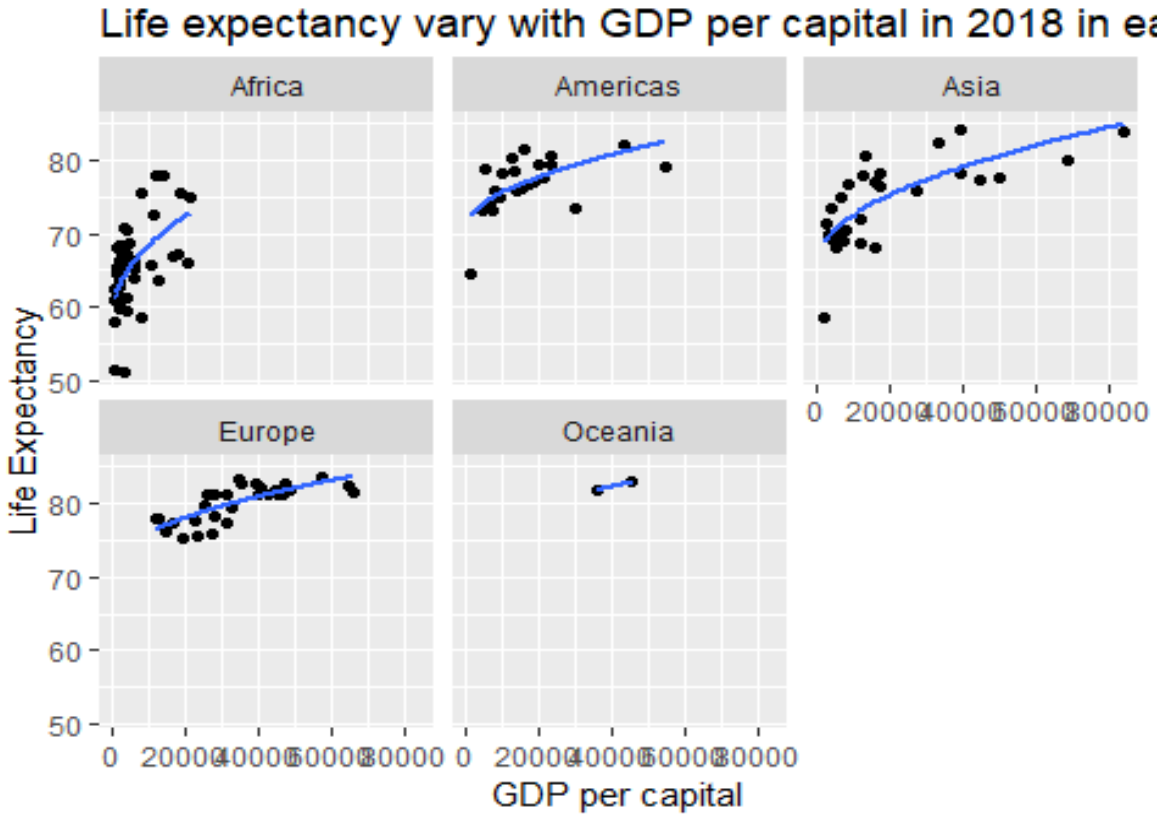
```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.5.2
```

1.GDP and life expectancy in 2018

```
## Warning: package 'tidyverse' was built under R version 3.5.3
```

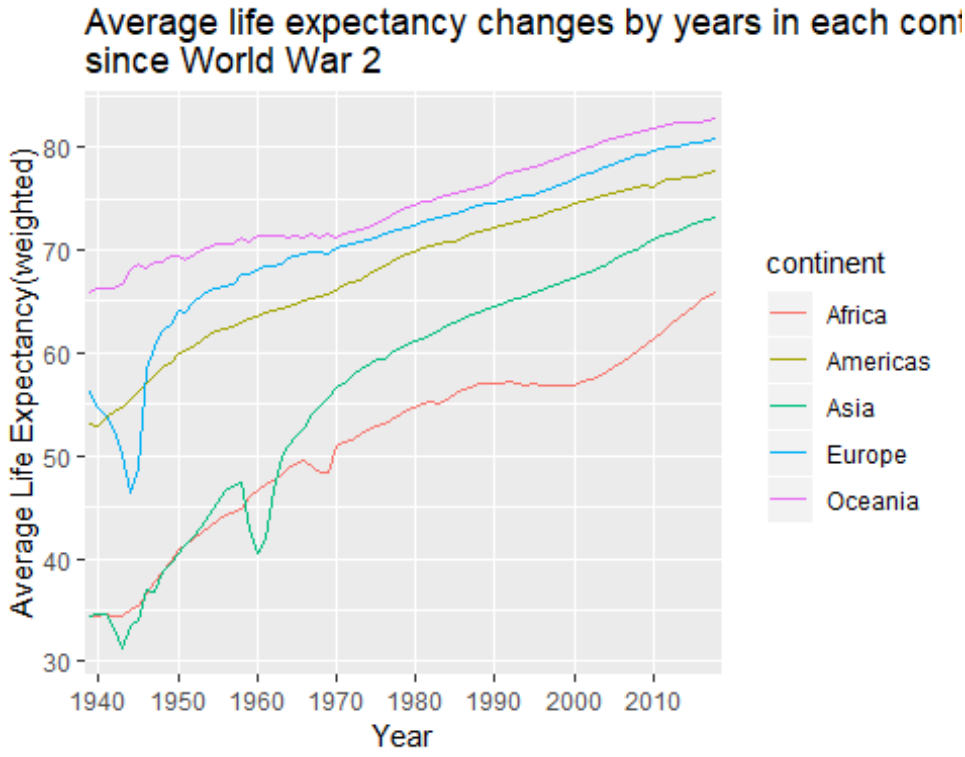
```
## Warning: package 'readr' was built under R version 3.5.1

## Warning: package 'forcats' was built under R version 3.5.1
```



From the graph, we can see that the trends between life expectancy and GDP per capital could not be described as a simple liner model, instead, it maybe some complicated model. It's obvious that the pattern for every continent is different. Continents such as Africa and Oceania are quite different than others.

For continents of Americas, Asia, and Europe, they tend to have a positive realtionship between life expectancy and GDP per capital, starting from around age 70. For the Africa, life expecatancy starting from 60, has a very sharp slope with GDP per capital. However, because Africa has much lower GDP values, then the curve of Africa is also shorter than other continents. For Oceania, we only get two observations, therefore, the curve may not be accurate enough to describe the relationship between life expectancy and GDP per capital in Oceania.

2.Life expectancy over time by continent



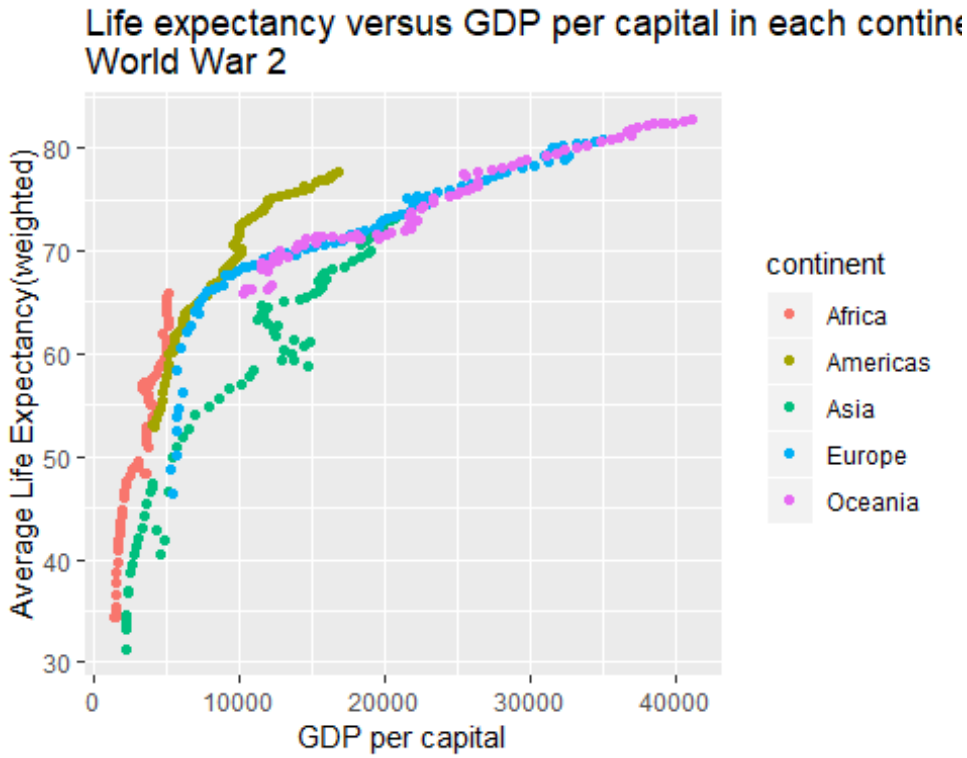
The average life expectancy changed since WorldWar2 in each continent is a lot. For the Africa continent, its life expectancy starts at 35 in the year 1939 which is the one of the lowest. And the tendency of life expectancy changed since WW2 is increased until the 1990. Then there is something happened to Africa, so the life expectancy drops a little bit. But it starts raising at 2000 again, to the highest point of itself at 61. For the Americans, its life expectancy starts at 54 at year 1930 which is the middle position within the five continents. Then it keeps an increasing tendency all the way until the 2018 to the highest point of itself on 77. For the Asia, its life expectancy starts at around 34 in year 1939 which also the one of the lowest. Then it drops a little in the year 1940 and then starts raising up at 1943 until the year 1958, it starts drop again until 1960 around age 41. But after the years 1960, Asia starts a continuing increasing tendency that all the way to the 2018 to the highest point of itself around age 74. For the Europe its life expectancy starts at age 56, then it drops utile 1943 same period with Asia, after 1943 it keeps increasing until 2018 reach the age 81. For the Oceania, it starts at age 76 which is the highest within five continents. And it appears a continuing increasing tendency until 2018 reaches the highest age with five continents around age 83.

Yes, there's several continents caught up to others, Asia is a typical example, and Europe after 1943 caught up to others as well. Since we weighted average life expectancy, so it is more general for continents that caught up to others. The linear change of the continents is only Oceania, Americas and Africa(kinda). Oceania is kind of flat, Americas faster than Oceania, so does Africa. Asia and Europe show an increase as general, but not kind of linear. And Asia seems a little faster than Europe. The period between 1939 to 1950 have a

dramatically change for both Asia and Europe. And this period is the WW2 period. The main battleground of the WW2 is on the Asia, and the Europe as a main invaders so these may be the main reasons that has faster change.

3.Changes in the relationship between GDP and life expectancy over time

```
GDPpercapita_cc_upto2018 <- subset(GDPpercapita_cc, select = country:`2018`)
dataset5 <- gather(GDPpercapita_cc_upto2018, key = "year", value = "GDPpercapita", -c(country, continent))
average_GDPpercapita_continent_year <- dataset5 %>% filter(year >= 1939) %>%
group_by(continent, year) %>%
                                                    summarise(averageGDPpercapita = mean(GDPpercapita))
dataset6 <- cbind(weighted_average_lifeExpectancy_continent_year, average_GDPpercapita_continent_year)
ggplot(dataset6, aes(x = averageGDPpercapita, y = weighted_average_LE_pop, group = continent, color = continent)) + geom_point() + ggtitle("Life expectancy versus GDP per capital in each continent since \nWorld War 2") + labs(x = "GDP per capital", y = "Average Life Expectancy(weighted)")
```



They all have a direct ratio that while the increasing of the GDP, the life expectancy raised as well. Unlike Oceania and Europe and Americas start at comparatively high GDP per capita and life expectancy, Asia and Africa, their starting GDP per capita almost 0, and the life expectancy is relatively much lower than other three continents. But if we look at them as a general, GDP per capita has a direct ratio with life expectancy. Changes in life expectancy cannot be entirely explained by changes in GDP per capita, because the line of the Asia, that when the life expectancy changed form 65 drops to 59, the GDP per capita still increased. Time not only effect on life expectancy, that while the increase of the years, the increase of the life expectancy, but also the GDP per capita which the increase of the years, GDP also increased. Yes, it has kind of "convergence" in the sense that perhaps GDP and/or continent don't matter as much as they used to. GDP per capita and life expectancy has convergence somehow at the top ending. Americas has a little exceptions, but still follow the most of other continents.