HW03: dplyr/ggplot2 Part II

Carleena Ortega

29/09/2019

Instructions

Pick three of the six tasks below, and produce:

- a tibble, using dplyr as your data manipulation tool
- an accompanying plot of data from the tibble, using ggplot2 as your visualization tool
- some dialogue about what your tables/figures show (doesn't have to be much).

Or, make up your own task(s)! If you want to do something comparable but different, i.e. swap one quantitative variable for another, be my guest! If you are feeling inspired and curious, then we're doing this right. Go for it. Just write down the task that you are now doing. Between the dplyr coverage in class and the list below, I think you get the idea.

Tasks

Task Option 1

Report the absolute and/or relative abundance of countries with low life expectancy over time by continent: Compute some measure of worldwide life expectancy – you decide – a mean or median or some other quantile or perhaps your current age. Then determine how many countries on each continent have a life expectancy less than this benchmark, for each year.

```r
t1<-gapminder %>%
  select(continent,year,lifeExp) %>%
  group_by(year,continent) %>%
  filter(lifeExp<60) %>%
  group_by(year,continent) %>%
  summarize(low_LE=length(continent))
knitr::kable(t1) %>%
  kable_styling("basic")
```

year	continent	low_LE
1952	Africa	52
1952	Americas	19
1952	Asia	29
1952	Europe	7
1957	Africa	52
1957	Americas	15
1957	Asia	27
1957	Europe	3
1962	Africa	51
1962	Americas	13
1962	Asia	25
1962	Europe	1
1967	Africa	50
1967	Americas	11
1967	Asia	25
1967	Europe	1
1972	Africa	50
1972	Americas	10
1972	Asia	19
1972	Europe	1
1977	Africa	50
1977	Americas	7
1977	Asia	14
1977	Europe	1
1982	Africa	44
1982	Americas	5
1982	Asia	12
1987	Africa	40
1987	Americas	2
1987	Asia	8
1992	Africa	39
1992	Americas	2
1992	Asia	7
1997	Africa	39
1997	Americas	1
1997	Asia	6
2002	Africa	41
2002	Americas	1
2002	Asia	4
2007	Africa	40
2007	Asia	3

```
ggplot(t1)+
  geom_point(aes(year,low_LE,group=continent,color=continent))+
  labs(x="Year",y="Number of Countries",title="The number of Countries in Each Continent with Life Expe
    Less than 60 Years")
```

The number of Countries in Each Continent with Life Expectancies Less than 60 Years
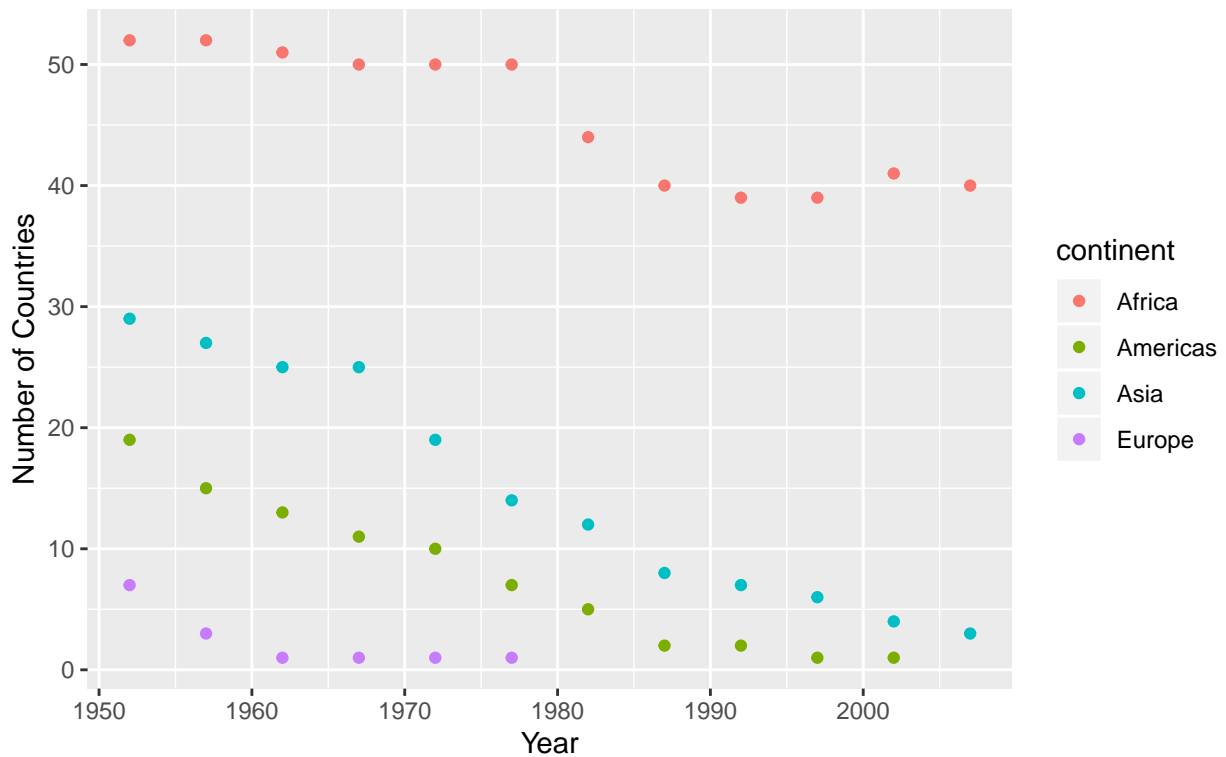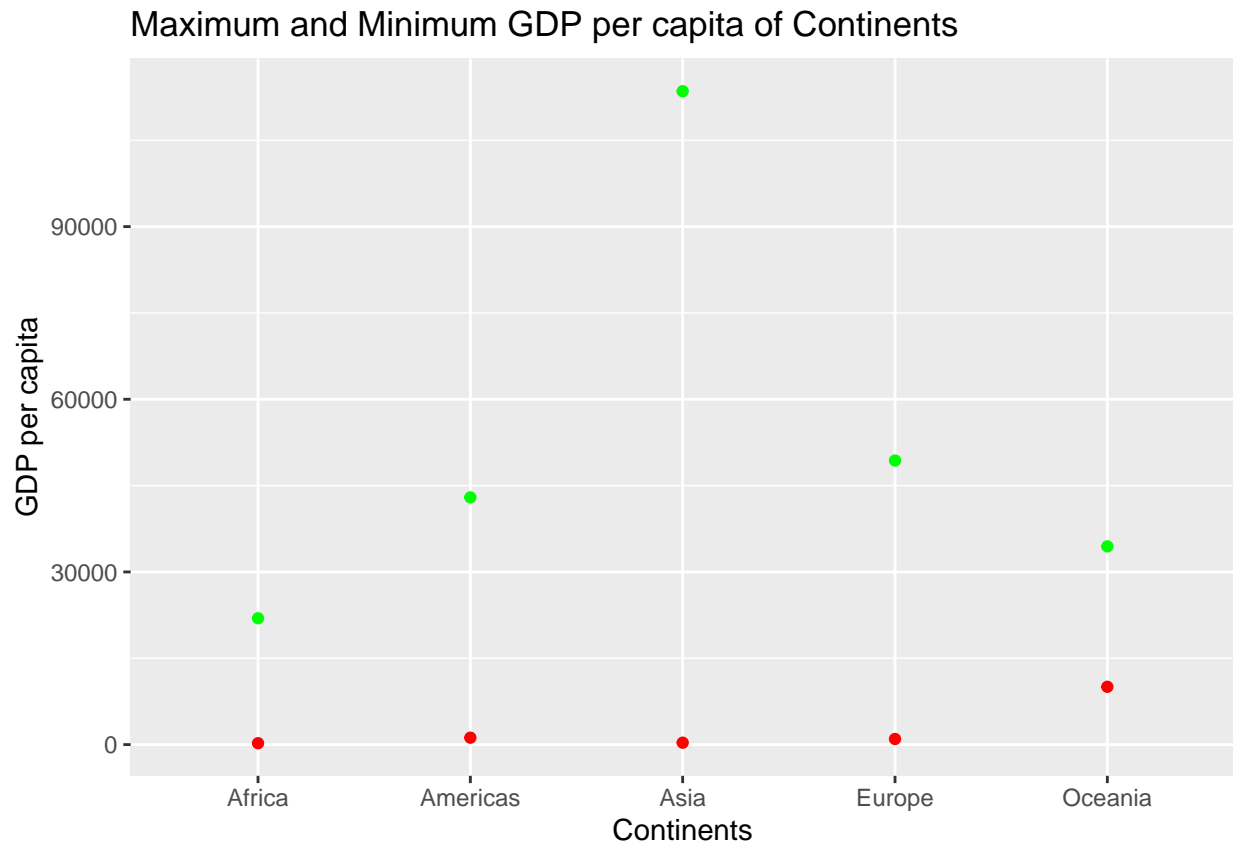

Task Option 2

Get the maximum and minimum of GDP per capita for all continents.

```r
t2<-gapminder %>%
  select(continent,gdpPercap) %>%
  group_by(continent) %>%
  arrange(continent)%>%
#mutate(min=min(gdpPercap),max=max(gdpPercap)) #if this is chosen instead of summarize, can plot but ta
  summarize(min_GDP=min(gdpPercap),max_GDP=max(gdpPercap)) #if this is chosen instead of mutate, cannot
knitr::kable(t2) %>%
  kable_styling("striped",position="center")
```

continent	min_GDP	max_GDP
Africa	241.1659	21951.21
Americas	1201.6372	42951.65
Asia	331.0000	113523.13
Europe	973.5332	49357.19
Oceania	10039.5956	34435.37

```r
#?is it possible to add the tibble and ggplot in one string?
```

```
ggplot(t2)+
  geom_point(aes(continent, min_GDP), color = 'red') + geom_point(aes(continent,max_GDP), color = 'green
  labs(x="Continents",y="GDP per capita",title="Maximum and Minimum GDP per capita of Continents")+
  theme_get()
```

Maximum and Minimum GDP per capita of Continents


```
#add legend for green and red
```

Task Option 3

Look at the spread of GDP per capita within the continents.

```
gapminder %>%
  select(continent,gdpPercap) %>%
  group_by(continent) %>%
  arrange(continent)
```

```
## # A tibble: 1,704 x 2
## # Groups:   continent [5]
##    continent gdpPercap
##    <fct>         <dbl>
## 1 Africa        2449.
## 2 Africa        3014.
## 3 Africa        2551.
## 4 Africa        3247.
```

```
## 5 Africa         4183.
## 6 Africa         4910.
## 7 Africa         5745.
## 8 Africa         5681.
## 9 Africa         5023.
## 10 Africa        4797.
## # ... with 1,694 more rows
```

```
gapminder %>%
  select(continent,gdpPercap) %>%
  group_by(continent) %>%
  arrange(continent)%>%
  ggplot()+
  geom_boxplot(aes(continent,gdpPercap)) +
  scale_y_log10()+
  labs(x="Continents",y="GDP per capita",title="GDP per capita of Continents")+
  theme_get()
```

GDP per capita of Continents
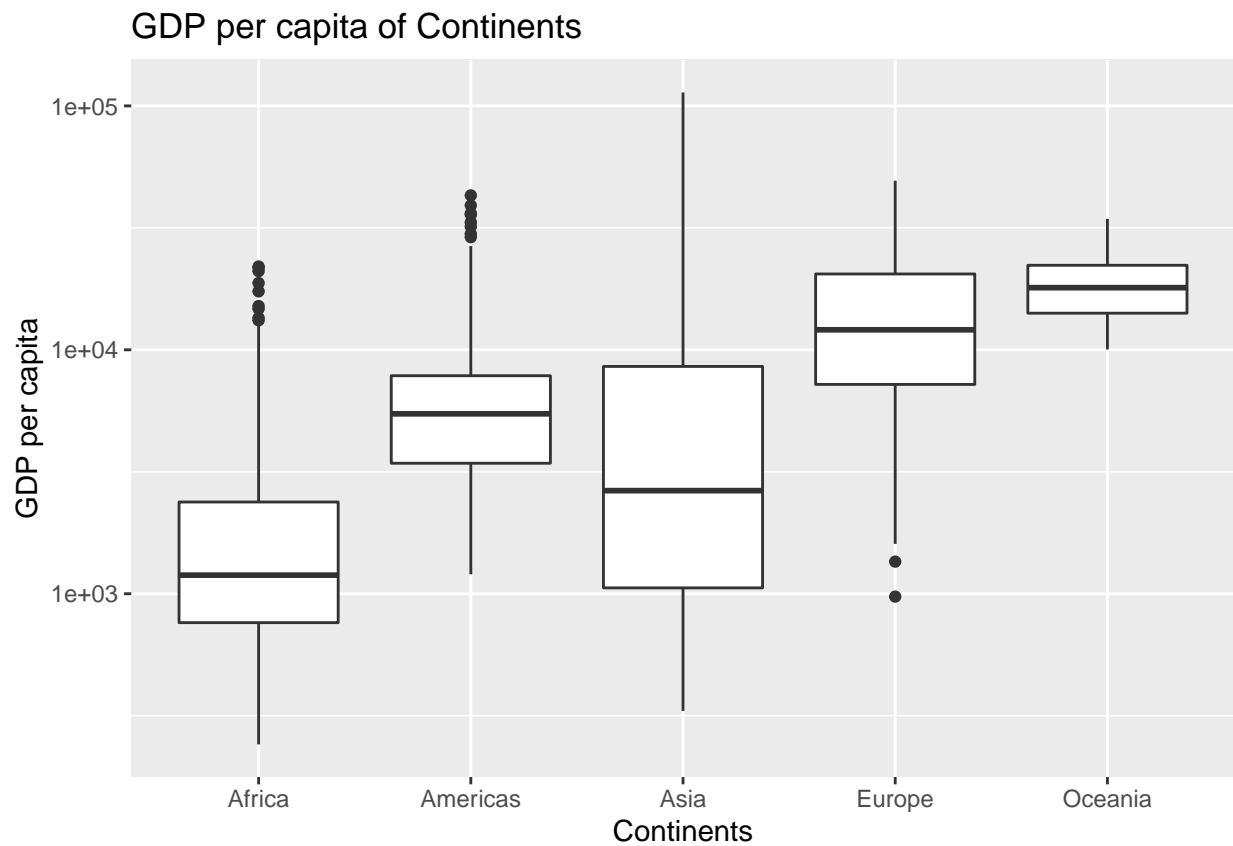

```
#add how to put the y range in better format
```

Task Option 4

Compute a trimmed mean of life expectancy for different years. Or a weighted mean, weighting by population.
Just try something other than the plain vanilla mean.

Task Option 5

How is life expectancy changing over time on different continents?

Task Option 6

Find countries with interesting stories. Open-ended and, therefore, hard. Promising but unsuccessful attempts are encouraged. This will generate interesting questions to follow up on in class.