ps5

Alex Han

2023-02-18

INFO201 Problem Set: rmarkdown and plotting library(tidyverse)

1 Load and check data (5pt)

1. (1pt) For solving the problems, and answering the questions, create a new rmarkdown document with an appropriate title. See https://faculty.washington.edu/otoomet/info201-book/r-markdown.html#r-markdown-rstudio-creating.

2. (2pt) Load data. How many rows/columns do we have?

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.8     v dplyr   1.1.0
## v tidyr   1.2.1     v stringr 1.5.0
## v readr   2.1.3     v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(dplyr)


df <- read_delim("data/gapminder.csv")
```

```
## Rows: 13055 Columns: 25
## -- Column specification --------------------------------------------------------
## Delimiter: "\t"
## chr  (6): iso3, name, iso2, region, sub-region, intermediate-region
## dbl (19): time, totalPopulation, fertilityRate, lifeExpectancy, childMortali...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
x <- nrow(df)
y <- ncol(df)

cat("It has", x, "rows, and",y, "columns")
```

```
## It has 13055 rows, and 25 columns
```

3. (2pt) Print a small sample of data. Does it look OK?

```r
df %>%
  sample_n(10)
```

```
## # A tibble: 10 x 25
##    iso3  name         iso2  region sub-r~1 inter~2  time total~3 ferti~4 lifeE~5
##    <chr> <chr>        <chr> <chr>  <chr>   <chr>   <dbl>   <dbl>   <dbl>   <dbl>
##  1 ISR   Israel       IL    Asia   Wester~ <NA>     1967  2.74e6    3.79    71.5
##  2 KIR   Kiribati     KI    Ocean~ Micron~ <NA>     1982  6.09e4    5.02    56.6
##  3 UKR   Ukraine      UA    Europe Easter~ <NA>     1987  5.13e7    2.05    70.5
##  4 KHM   Cambodia     KH    Asia   South-~ <NA>     2016  1.58e7    2.56    69.0
##  5 MKD   North Maced~ MK    Europe Southe~ <NA>     2008  2.07e6    1.47    74.3
##  6 VUT   Vanuatu      VU    Ocean~ Melane~ <NA>     2016  2.78e5    3.86    70.0
##  7 LKA   Sri Lanka    LK    Asia   Southe~ <NA>     2000  1.88e7    2.24    71.3
##  8 NOR   Norway       NO    Europe Northe~ <NA>     2009  4.83e6    1.98    80.8
##  9 MDG   Madagascar   MG    Africa Sub-Sa~ Easter~  1994  1.31e7    6.03    53.8
## 10 CYP   Cyprus       CY    Asia   Wester~ <NA>     2004  1.01e6    1.54    78.5
## # ... with 15 more variables: childMortality <dbl>, youthFemaleLiteracy <dbl>,
## #   youthMaleLiteracy <dbl>, adultLiteracy <dbl>, GDP_PC <dbl>,
## #   accessElectricity <dbl>, agriculturalLand <dbl>, agricultureTractors <dbl>,
## #   cerealProduction <dbl>, fertilizerHa <dbl>, co2 <dbl>,
## #   greenhouseGases <dbl>, co2_PC <dbl>, pm2.5_35 <dbl>, battleDeaths <dbl>,
## #   and abbreviated variable names 1: `sub-region`, 2: `intermediate-region`,
## #   3: totalPopulation, 4: fertilityRate, 5: lifeExpectancy
```

The data looks OK.

2 Descriptive statistics (15pt)

1. (3pt) How many countries are there in the dataset? Analyze all three: iso3, iso2 and name.

```r
count(unique(df['iso3']))
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1   253
```

```r
count(unique(df['iso2']))
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1   249
```

```r
count(unique(df['name']))
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1   250
```

2. If you did this correctly, you saw that there are more names than iso-2 codes, and there are even more iso3 -codes. What is going on? Can you find it out?

   (a) (5pt) Find how many names are there for each iso-2 code. Are there any iso-2 codes that correspond to more than one name? What are these countries?

```
df %>%
  group_by(iso2) %>%
  summarise(number_count = n_distinct(name)) %>%
  select(iso2,number_count) %>%
  arrange(desc(number_count)) %>%
  print()
```

```
## # A tibble: 249 x 2
##    iso2  number_count
##    <chr>        <int>
##  1 <NA>             2
##  2 AD               1
##  3 AE               1
##  4 AF               1
##  5 AG               1
##  6 AI               1
##  7 AL               1
##  8 AM               1
##  9 AO               1
## 10 AQ               1
## # ... with 239 more rows
```

```
df %>%
  filter(is.na(name)) %>%
  distinct(iso2)
```

```
## # A tibble: 1 x 1
##   iso2
##   <chr>
## 1 <NA>
```

The NA iso2 code correspond to more than one name.

(b) (5pt) Now repeat the same for name and iso3-code. Are there country names that have more than one iso3-code? What are these countries?
Hint: two of these entitites are CHANISL and NLD CURACAO.

```
df %>%
  group_by(name) %>%
  summarise(name_count = n_distinct(iso3)) %>%
  select(name,name_count) %>%
  arrange(desc(name_count)) %>%
  print()
```

```
## # A tibble: 250 x 2
##    name                 name_count
##    <chr>                     <int>
##  1 <NA>                          4
##  2 Afghanistan                   1
##  3 Albania                       1
##  4 Algeria                       1
##  5 American Samoa                1
##  6 Andorra                       1
```

```
## 7 Angola 1
## 8 Anguilla 1
## 9 Antarctica 1
## 10 Antigua and Barbuda 1
## # ... with 240 more rows
```

```
df %>%
   filter(is.na(name)) %>%
   distinct(iso3)
```

```
## # A tibble: 4 x 1
##   iso3
##   <chr>
## 1 CHANISL
## 2 GBM
## 3 KOS
## 4 NLD_CURACAO
```

There are four countries with more than one iso3-code and they are: CHANISL,GBM,KOS,NLD_CURACAO

3. (2pt) What is the minimum and maximum year in these data?

```
df %>%
  filter(!is.na(time),!is.na(time)) %>%
  summarise(maximum_year = max(time), minimum_year = min(time))
```

```
## # A tibble: 1 x 2
##   maximum_year minimum_year
##          <dbl>        <dbl>
## 1         2019         1960
```

The maximum year in this dataset is 2019 and the minimum year is 1960.

3 CO2 emissions (30pt)

Next, let's analyze CO2 emissions.

1. (2pt) How many missing co2 emissions are there for each year? Analyze both missing CO2 and co2_PC. Which years have most missing data?

```
df %>%
    filter(is.na(co2),is.na(co2_PC)) %>%
    group_by(time) %>%
    summarise(years = n()) %>%
    arrange(-years) %>%
    head(3)
```

```
## # A tibble: 3 x 2
##    time years
##   <dbl> <int>
## 1  2017   217
## 2  2018   217
## 3  2019   217
```

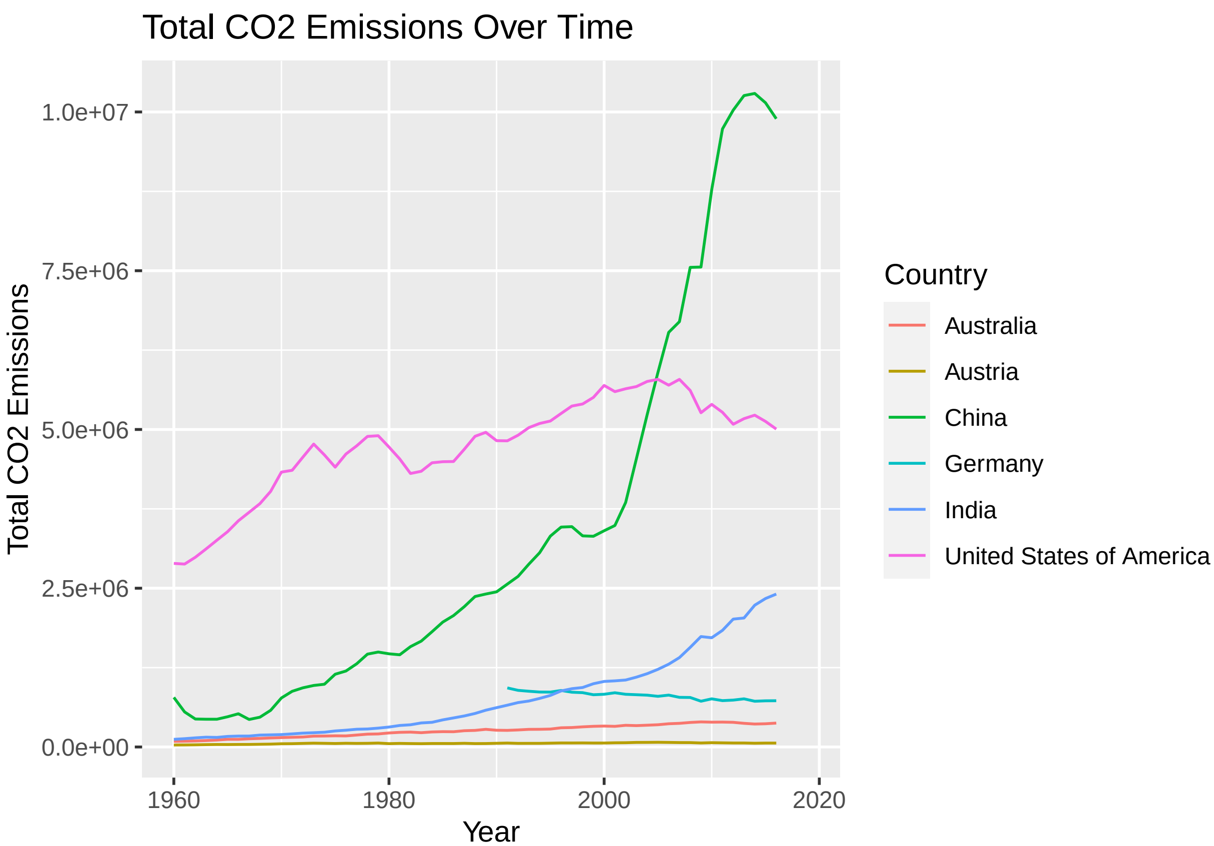
The three years with the most missing data are: 2019, 2018, 2017

2. (5pt) Make a plot of total CO2 emissions over time for the U.S, China, and India. Add a few more countries of your choice. Explain what do you see.

```r
df %>%
    filter(name == "United States of America" | name == "China" | name == "India"
           | name == "Austria" | name == "Germany" | name == "Australia") %>%
    group_by(name, time) %>%
    summarise(t_co2 = sum(co2)) %>%
    ggplot(aes(x = time, y = t_co2, color = name)) +
    geom_line() +
    labs(title = "Total CO2 Emissions Over Time", x = "Year", y = "Total CO2 Emissions", color = "Count
```

```
## `summarise()` has grouped output by 'name'. You can override using the
## `.groups` argument.

## Warning: Removed 49 row(s) containing missing values (geom_path).
```



It appears to me that China had a exponential growth in CO2 emissions in the early 2000's, this might due to its rapid economic reforms in the late 90's. The US is also showing a slow decline in CO2 emissions in recent years due to the mass replacement of coal power with solar, electric and other forms of green energy. India is also steadily climbing in total CO2 emissions as it is a fast developing nation. The other countries I've chosen: Austria, Australia and Germany are all pioneers in green energy and it's absolutely no surprise that they produce substantially less CO2 emissions compared to China, the US, and India.

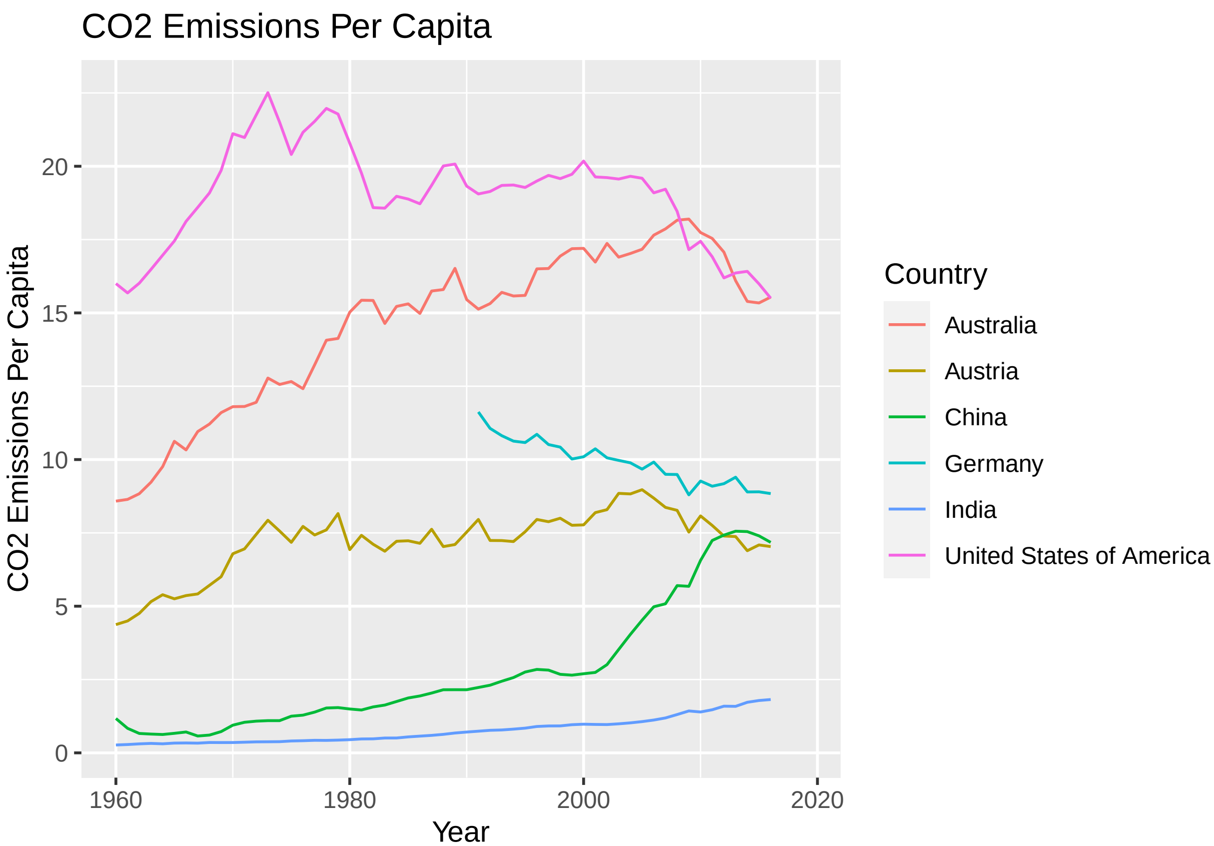
3. (5pt) Now let's analyze the CO2 emissions per capita (co2_PC). Make a similar plot of the same countries. What does this figure suggest?

```r
df %>%
    filter(name == "United States of America" | name == "China" | name == "India"
           | name == "Austria" | name == "Germany" | name == "Australia") %>%
    group_by(name, time) %>%
```

```
    summarise(t_co2pc = co2_PC) %>%
    ggplot(aes(x = time, y = t_co2pc, color = name)) +
    geom_line() +
    labs(title = "CO2 Emissions Per Capita", x = "Year", y = "CO2 Emissions Per Capita", color = "Count
```

```
## `summarise()` has grouped output by 'name'. You can override using the
## `.groups` argument.

## Warning: Removed 49 row(s) containing missing values (geom_path).
```



It is no surprise that the US would still remain at the top of the chart for having the highest CO2 emissions per capita. Australia, being one of the most developed and industrialized nations of the world also situates itself on the top of the chart. Germany and Austria while also developed find themselves sitting in between Australia and China. The two still developing countries of China and India have the lowest CO2 emission per capita. This suggests that rapidly developing countries such as China and India have much lower CO2 emissions per capita compared to the developed nations. However, their total CO2 emissions are propelled higher than the developed nations to due their enormous population.

4. (6pt) Compute average CO2 emissions per capita across the continents (assume region is the same as continent). Comment what do you see.
Note: just compute averages over countries and ignore the fact that countries are of different size.
Hint: Americas 2016 should be 4.80.

```
df %>%
   filter(!is.na(co2_PC)) %>%
   group_by(region) %>%
   summarise(avg_co2_PC = mean(co2_PC)) %>%
```

```
  print()
```

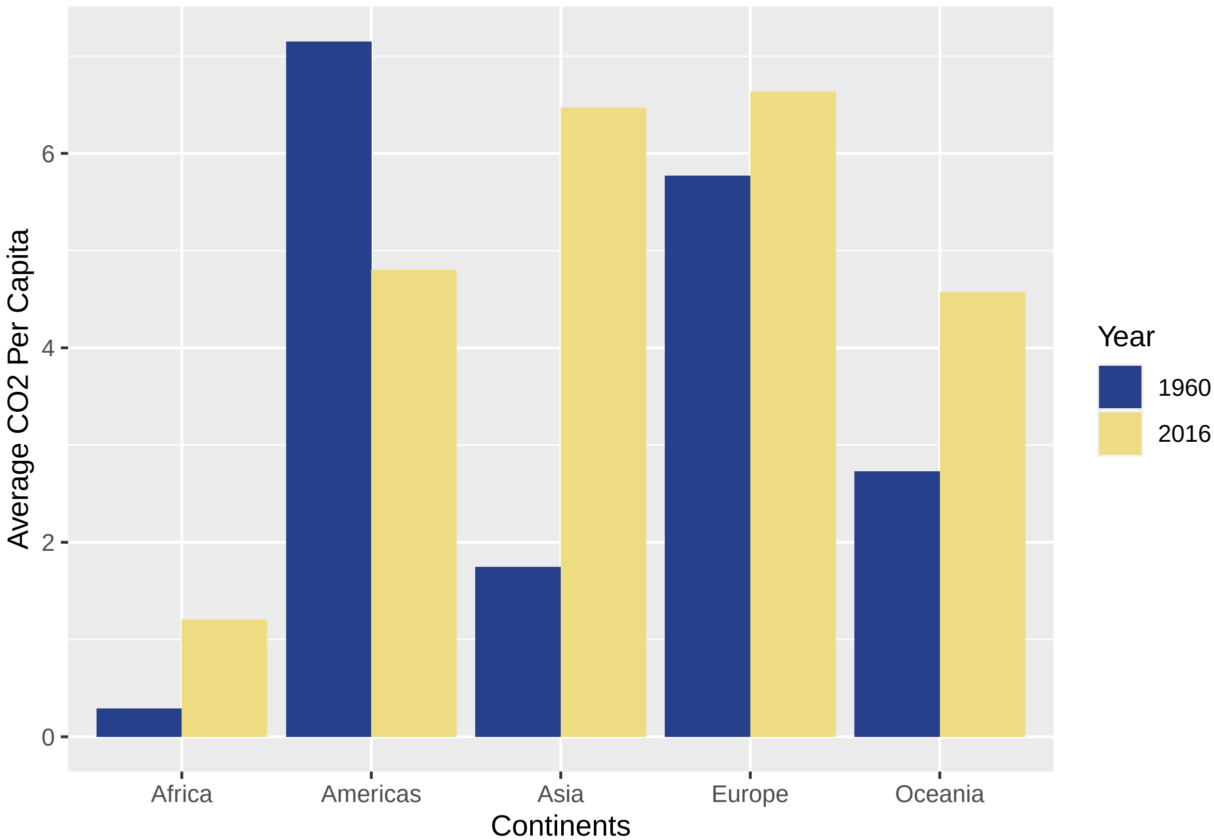
```
## # A tibble: 6 x 2
##   region   avg_co2_PC
##   <chr>         <dbl>
## 1 Africa        0.930
## 2 Americas      6.46
## 3 Asia          6.21
## 4 Europe        7.95
## 5 Oceania       4.39
## 6 <NA>         16.2
```

This shows that continents with more developing nations would produce less CO2 per capita even if their population is much higher than that of a continent with more developed nations.

5. (7pt) Make a barplot where you show the previous results–average CO2 emissions per capita across continents in 1960 and 2016.
Hint: it should look something along these lines:

```
ycolors <- c("royalblue4", "lightgoldenrod2")
df %>%
  filter(time == "1960" | time == "2016") %>%
  filter(!is.na(co2_PC)) %>%
  filter(!is.na(region)) %>%
  group_by(region, time) %>%
  summarise(avg_co2pc = mean(co2_PC)) %>%
  ggplot(aes(x= region, y = avg_co2pc, fill = as.factor(time))) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.9))+
  scale_fill_manual(values = ycolors) +
  labs(x= "Continents", y = "Average CO2 Per Capita",
       fill = "Year")
```

```
## `summarise()` has grouped output by 'region'. You can override using the
## `.groups` argument.
```



This shows that the average CO2 per capita has actually dropped for the Americas. Personally, I believe this to be the effect of the US switching over to green energy in the span of the period from 1960 to 2016. However, other continents have also increased in average CO2 emission per capita as the two developing powerhouses of Asia— China and India are fully dedicated to the path of becoming world economical powers. Increasing Asia's average CO2 per capita by a massive amount from 1960 to 2016.

6. Which countries are the three largest, and three smallest CO2 emitters (in terms of CO2 per capita) in 2019 for each continent? (Assume region is continent).

```
#Americas
df %>%
   filter(region == "Americas") %>%
   filter(time == "2019") %>%
   group_by(region) %>%
   mutate(rank = rank(co2_PC)) %>%
   filter(rank(rank) <= 3 | rank >= n() -2) %>%
   select(name, region, rank) %>%
   arrange(rank)
```

```
## # A tibble: 6 x 3
## # Groups:   region [1]
##   name                               region    rank
##   <chr>                              <chr>    <dbl>
## 1 Aruba                              Americas     1
## 2 Argentina                          Americas     2
## 3 Antigua and Barbuda                Americas     3
## 4 Venezuela (Bolivarian Republic of) Americas    43
## 5 Virgin Islands (British)           Americas    44
```

```
## 6 Virgin Islands (U.S.)              Americas    45
```

```r
#Asia
df %>%
   filter(region == "Asia") %>%
   filter(time == "2019") %>%
   group_by(region) %>%
   mutate(rank = rank(co2_PC)) %>%
   filter(rank(rank) <= 3 | rank >= n() -2) %>%
   select(name, region, rank) %>%
   arrange(rank)
```

```
## # A tibble: 6 x 3
## # Groups:   region [1]
##   name                 region  rank
##   <chr>                <chr>  <dbl>
## 1 Afghanistan          Asia       1
## 2 United Arab Emirates Asia       2
## 3 Armenia              Asia       3
## 4 Uzbekistan           Asia      48
## 5 Viet Nam             Asia      49
## 6 Yemen                Asia      50
```

```r
#Europe
df %>%
   filter(region == "Europe") %>%
   filter(time == "2019") %>%
   group_by(region) %>%
   mutate(rank = rank(co2_PC)) %>%
   filter(rank(rank) <= 3 | rank >= n() -2) %>%
   select(name, region, rank) %>%
   arrange(rank)
```

```
## # A tibble: 6 x 3
## # Groups:   region [1]
##   name     region  rank
##   <chr>    <chr>  <dbl>
## 1 Albania  Europe     1
## 2 Andorra  Europe     2
## 3 Austria  Europe     3
## 4 Slovenia Europe    43
## 5 Sweden   Europe    44
## 6 Ukraine  Europe    45
```

```r
#Africa
df %>%
   filter(region == "Africa") %>%
   filter(time == "2019") %>%
   group_by(region) %>%
   mutate(rank = rank(co2_PC)) %>%
   filter(rank(rank) <= 3 | rank >= n() -2) %>%
   select(name, region, rank) %>%
   arrange(rank)
```

```
## # A tibble: 6 x 3
## # Groups:   region [1]
```

```
##   name         region  rank
##   <chr>        <chr>  <dbl>
## 1 Angola       Africa     1
## 2 Burundi      Africa     2
## 3 Benin        Africa     3
## 4 South Africa Africa    52
## 5 Zambia       Africa    53
## 6 Zimbabwe     Africa    54
```

```
#Oceania
df %>%
   filter(region == "Oceania") %>%
   filter(time == "2019") %>%
   group_by(region) %>%
   mutate(rank = rank(co2_PC)) %>%
   filter(rank(rank) <= 3 | rank >= n() -2) %>%
   select(name, region, rank) %>%
   arrange(rank)
```

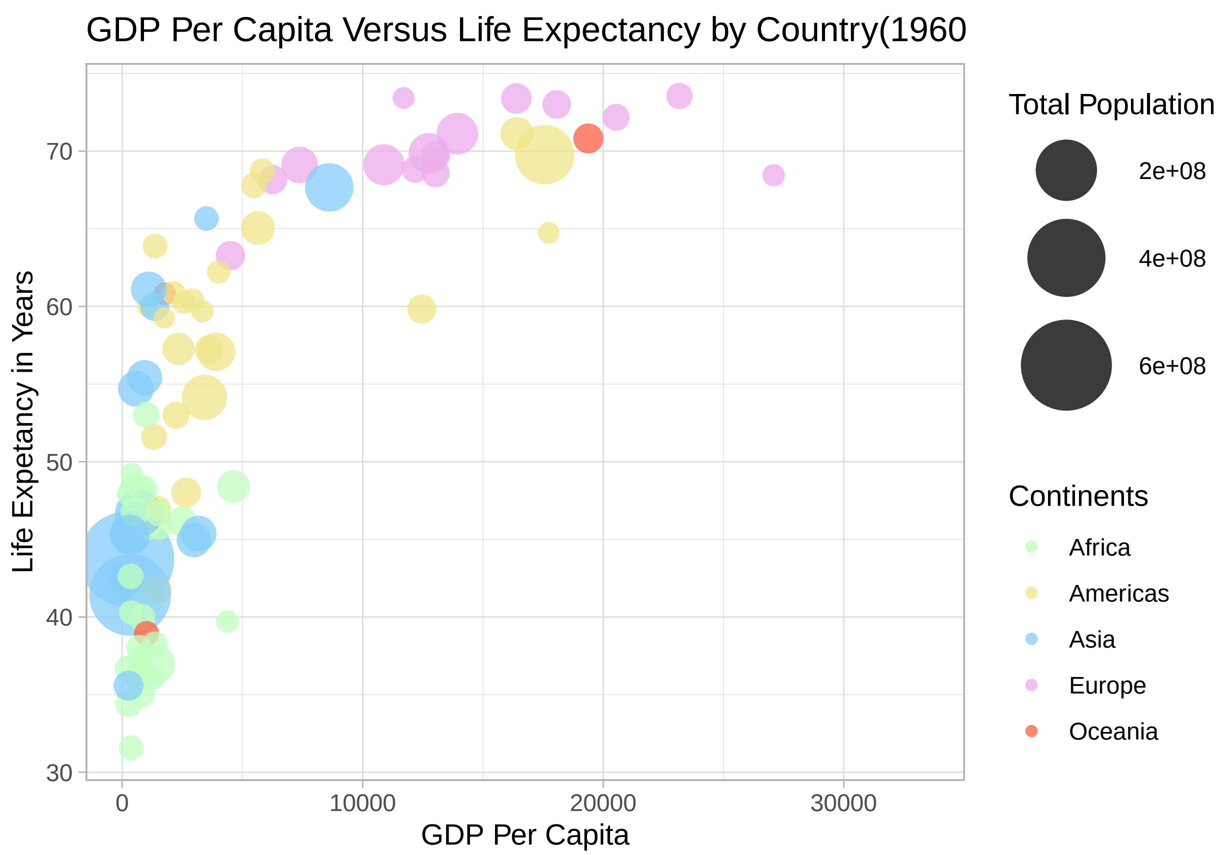
```
## # A tibble: 6 x 3
## # Groups:   region [1]
##   name           region   rank
##   <chr>          <chr>   <dbl>
## 1 American Samoa Oceania     1
## 2 Australia      Oceania     2
## 3 Fiji           Oceania     3
## 4 Tuvalu         Oceania    17
## 5 Vanuatu        Oceania    18
## 6 Samoa          Oceania    19
```

4 GDP per capita (50pt)

Let's look at GDP per capita (GDP_PC).
1. (8pt) Make a scatterplot of GDP per capita versus life expectancy by country, using data for 1960. Make the point size dependent on the country size, and color those according to the continent. Feel free to adjust the plot in other ways to make it better.
Comment what do you see there.

```
con_colors <- c("darkseagreen1","khaki", "lightskyblue", "plum2", "tomato1")
df %>%
  filter(time == "1960") %>%
  filter(!is.na(region)) %>%
  filter(!is.na(GDP_PC)) %>%
  ggplot(aes(x = GDP_PC, y = lifeExpectancy)) +
  geom_point(aes(size = totalPopulation, color = region), alpha = 0.77) +
  scale_size(range = c(3, 16)) +
  scale_color_manual(values = con_colors) +
  labs(title = "GDP Per Capita Versus Life Expectancy by Country(1960",
       x = "GDP Per Capita", y = "Life Expetancy in Years",
       size = "Total Population", color = "Continents") +
  theme_light()
```

```
## Warning: Removed 3 rows containing missing values (geom_point).
```

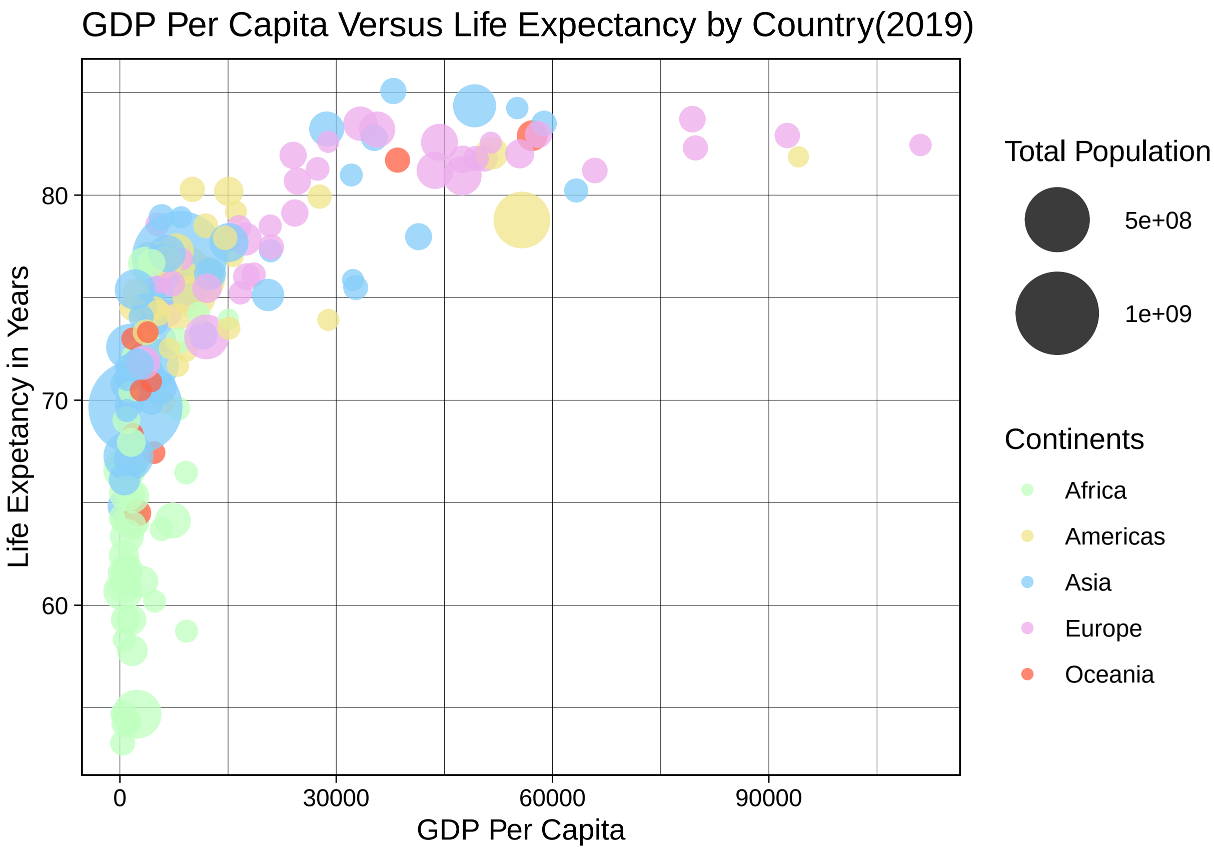


From the plot I'm observing that populations from countries with **higher GDP per capita** demonstrate **higher longevity**. European countries are often much higher on the chart in terms of GDP per capita and life expectancy. Another observation I've made is that *population size does not seem to have a direct effect on life expectancy* albeit China and India are ranked relatively low in terms of life expectancy on the chart. I believe that dots plotted toward the lower left portion of the chart are countries yet to develop and countries plotted on the top right portion of the graph are developed nations.

2. (4pt) Make a similar plot, but this time use 2019 data only.

```
con_colors <- c("darkseagreen1","khaki", "lightskyblue", "plum2", "tomato1")
df %>%
  filter(time == "2019") %>%
  filter(!is.na(region)) %>%
  filter(!is.na(GDP_PC)) %>%
  ggplot(aes(x = GDP_PC, y = lifeExpectancy)) +
  geom_point(aes(size = totalPopulation, color = region), alpha = 0.77) +
  scale_size(range = c(3, 16)) +
  scale_color_manual(values = con_colors) +
  labs(title = "GDP Per Capita Versus Life Expectancy by Country(2019)",
       x = "GDP Per Capita", y = "Life Expetancy in Years",
       size = "Total Population", color = "Continents") +
  theme_linedraw()
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



From the graph, we can observe that countries across the world has made **dramatic improvements** in terms of life expectancy and GDP per capita. This is understandable as there were many **social** and **technological advancements** made over the course of 60 years.

3. (6pt) Compare these two plots and comment what do you see. How has world developed through the last 60 years?

While comparing the two charts, we can observe that the world has *developed drastically* in the span of **60 years** as more dots are moving toward the upper right corner with countries from each continent making progress toward development and modernization. Overall, it is clearly that countries yet to develop in the 60's are showing great signs of *economical* and *social* **advancements**.

4. (6pt) Compute the average life expectancy for each continent in 1960 and 2019. Do the results fit with what do you see on the figures?
Note: here as average I mean just average over countries, ignore the fact that countries are of different size.

```
#1960
df %>%
   filter(!is.na(lifeExpectancy)) %>%
   filter(grepl('1960', time)) %>%
   group_by(region) %>%
   summarise(avg_le60 = mean(lifeExpectancy))
```

```
## # A tibble: 6 x 2
##   region   avg_le60
##   <chr>       <dbl>
```

```
## 1 Africa       41.5
## 2 Americas     58.6
## 3 Asia         51.6
## 4 Europe       68.3
## 5 Oceania      56.4
## 6 <NA>         70.7
```

```r
#2019
df %>%
   filter(!is.na(lifeExpectancy)) %>%
   filter(grepl('2019', time)) %>%
   group_by(region) %>%
   summarise(avg_le19 = mean(lifeExpectancy))
```

```
## # A tibble: 6 x 2
##   region   avg_le19
##   <chr>       <dbl>
## 1 Africa       64.1
## 2 Americas     75.8
## 3 Asia         74.6
## 4 Europe       79.4
## 5 Oceania      73.5
## 6 <NA>         77.8
```

With no surprise, life expectancy have *dramatically increased* for across **all continents** and especially for Asia and Africa. This data **supports** my graphs by showing an *uptrend in life expectancy* for world populations.

5. (8pt) Compute the average LE growth from 1960-2019 across the continents. Show the results in the order of growth. Explain what do you see.
Hint: these data (data in long form) is not the simplest to compute growth. But you may want to check out the lag() function. And do not forget to group data by continent when using lag(), otherwise your results will be messed up! See https://faculty.washington.edu/otoomet/info201-book/dplyr.html#dplyr-helpers-compute.

```r
df %>%
   filter(!is.na(lifeExpectancy)) %>%
   filter(time == "1960"| time == "2019" ) %>%
   filter(!is.na(region)) %>%
   group_by(region, time) %>%
   summarise(avgle = mean(lifeExpectancy)) %>%
   mutate(prev = lag(avgle), growth = avgle - prev) %>%
   filter(!is.na(growth)) %>%
   arrange(-growth)
```

```
## `summarise()` has grouped output by 'region'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 5 x 5
## # Groups:   region [5]
##   region    time avgle  prev growth
##   <chr>    <dbl> <dbl> <dbl>  <dbl>
## 1 Asia      2019  74.6  51.6   23.0
## 2 Africa    2019  64.1  41.5   22.6
## 3 Americas  2019  75.8  58.6   17.2
## 4 Oceania   2019  73.5  56.4   17.1
## 5 Europe    2019  79.4  68.3   11.1
```

From the output, one can observe that all continents have shown improvements in life expectancy with the Asia and Africa demonstrating the most substantial

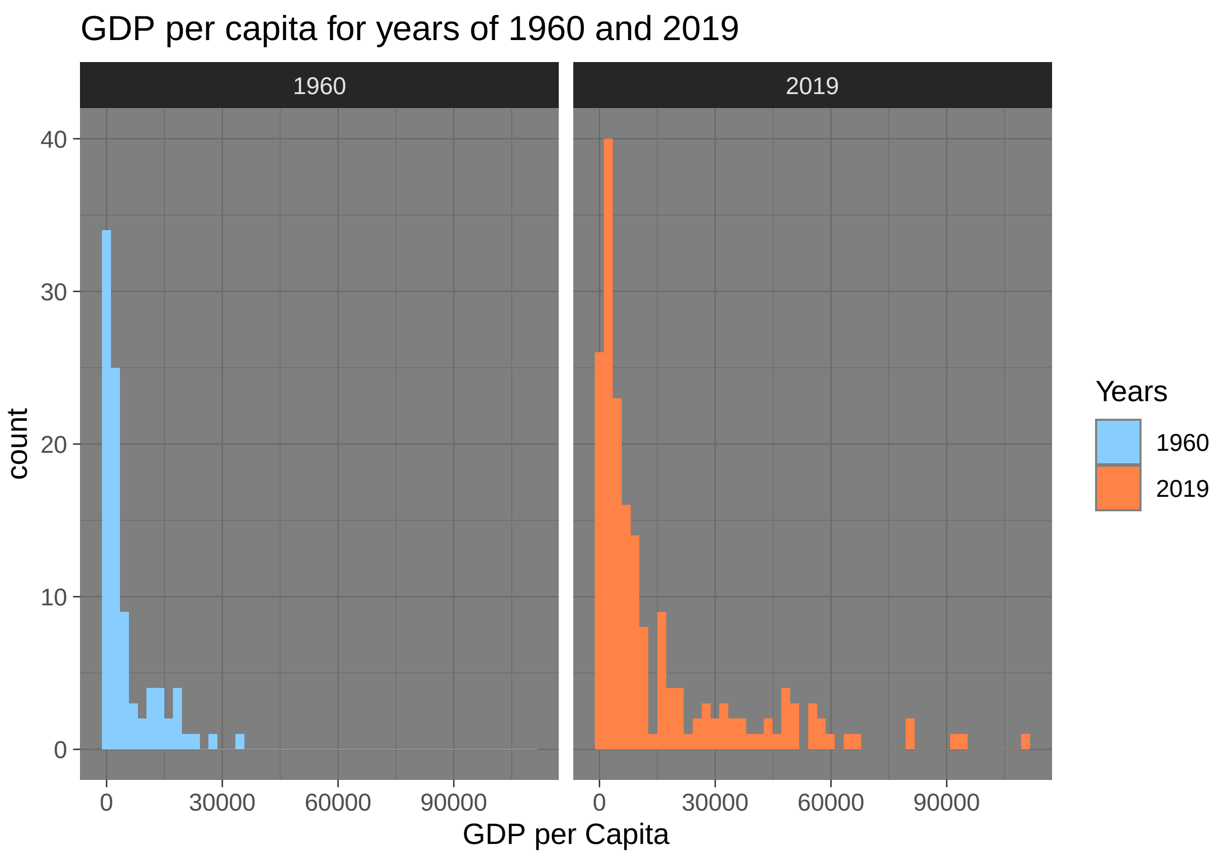
improvements.

6. (6pt) Show the histogram of GDP per capita for years of 1960 and 2019. Try to put both histograms on the same graph, see how well you can do it!

```
hcolors <- c("skyblue1", "sienna1")

df %>%
  filter(time == "1960" | time == "2019") %>%
  ggplot(aes(x = GDP_PC, fill = as.factor(time)))+
  geom_histogram(binwidth = 2300, position = position_dodge(width = 0.8)) +
  facet_wrap(~ time) +
  scale_fill_manual(values = hcolors) +
  labs(title = "GDP per capita for years of 1960 and 2019", x = "GDP per Capita", fill = "Years")+
  theme_dark()
```

```
## Warning: Removed 157 rows containing non-finite values (stat_bin).
```



From the histograms, we can observe that more countries are joining the higher end of GDP per capita in in 2019. Overall, *almost all countries* have *shifted* over bin width to the *right* in the span of 60 years.

7. (6pt) What was the ranking of US in terms of life expectancy in 1960 and in 2019? (When counting from top.)
Hint: check out the function rank()!
Hint2: 17 for 1960.

```r
df %>%
    filter(time == "1960") %>%
    filter(!is.na(lifeExpectancy)) %>%
    filter(!is.na(region)) %>%
    mutate(le_rank = rank(desc(lifeExpectancy))) %>%
    select(name, le_rank, time,region) %>%
    filter(name == "United States of America") %>%
    print()
```

```
## # A tibble: 1 x 4
##   name                     le_rank  time region
##   <chr>                      <dbl> <dbl> <chr>
## 1 United States of America      17  1960 Americas
```

```r
df %>%
    filter(time == "2019") %>%
    filter(!is.na(lifeExpectancy)) %>%
    filter(!is.na(region)) %>%
    mutate(le_rank = rank(desc(lifeExpectancy))) %>%
    select(name, le_rank, time, region) %>%
    filter(name == "United States of America") %>%
    print()
```

```
## # A tibble: 1 x 4
##   name                     le_rank  time region
##   <chr>                      <dbl> <dbl> <chr>
## 1 United States of America      46  2019 Americas
```

The US ranked 17th in life expectancy in 1960 and 46th in 2019.

8. (6pt) If you did this correctly, then you noticed that US ranking has been falling quite a bit. But we also have more countries in 2019–what about the relative rank divided by the corresponding number of countries that have LE data in the corresponding year?
Hint: 0.0904 for 1960.

```r
df %>%
    filter(time == "1960") %>%
    filter(!is.na(lifeExpectancy)) %>%
    filter(!is.na(region)) %>%
    mutate(le_rank = rank(desc(lifeExpectancy)), nc = n(), r_rank = le_rank/nc) %>%
    select(name, le_rank, time, nc, r_rank) %>%
    filter(name == "United States of America") %>%
    print()
```

```
## # A tibble: 1 x 5
##   name                     le_rank  time    nc r_rank
##   <chr>                      <dbl> <dbl> <int>  <dbl>
## 1 United States of America      17  1960   188 0.0904
```

```r
df %>%
    filter(time == "2019") %>%
    filter(!is.na(lifeExpectancy)) %>%
    filter(!is.na(region)) %>%
    mutate(le_rank = rank(desc(lifeExpectancy)), nc = n(), r_rank = le_rank/nc) %>%
    select(name, le_rank, time, nc, r_rank) %>%
    filter(name == "United States of America") %>%
```

```
    print()
```

```
## # A tibble: 1 x 5
##   name                     le_rank  time    nc r_rank
##   <chr>                      <dbl> <dbl> <int>  <dbl>
## 1 United States of America      46  2019   196  0.235
```

Using the relative rank, we can still observe that the US has dropped significantly in terms of life expectancy over the course of 60 years. In 1960, the US was in the 90th percentile in terms of life expectancy. However in 2019, the US is now in the 77th percentile.

Finally tell us how many hours did you spend on this PS. Around 7 hours