

MA615 HW#2

Xinqiao Zhang

10/2/2019

Problem 1

Load the gapminder data from the gapminder package.

How many continents are included in the data set?

There are 5 continents included in the dataset.

How many countrys are included? How many countries per continent?

There are 142 countries in the dataset, so it is 28.4 countries per continent on average.

Using the gapminder data, produce a report showing the continents in the dataset, total population per continent, and GDP per capita. Be sure that the table is properly labeled and suitable for inclusion in a printed report.

```
## # A tibble: 5 x 4
## # Groups:   year [1]
##   year continent    ttl_pop GDPperCap
##   <dbl> <fct>         <dbl>    <dbl>
## 1  2007 Africa      6187585961 1368903.
## 2  2007 Americas   7351438499 2140833.
## 3  2007 Asia       30507333901 3129252.
## 4  2007 Europe     6181115304 5209011.
## 5  2007 Oceania    212992136  446919.
```

Since the data contains data from different years, it would be reasonable to summarize the entries by years. Since there are too many years to show, we only show the data from the latest year, 2007.

Produce a well-labeled table that summarizes GDP per capita for the countries in each continent, contrasting the years 1952 and 2007.

Product a plot that summarizes the same data as the table. There should be two plots per continent. *GDP in*

Table 1: GDP per Capital in 1952

continent	GDPperCap
Africa	65133.77
Americas	101976.56
Asia	171450.97
Europe	169831.72
Oceania	20596.17

Table 2: GDP per Capital in 2007

continent	GDPperCap
Africa	160629.70
Americas	275075.79
Asia	411609.89
Europe	751634.45
Oceania	59620.38

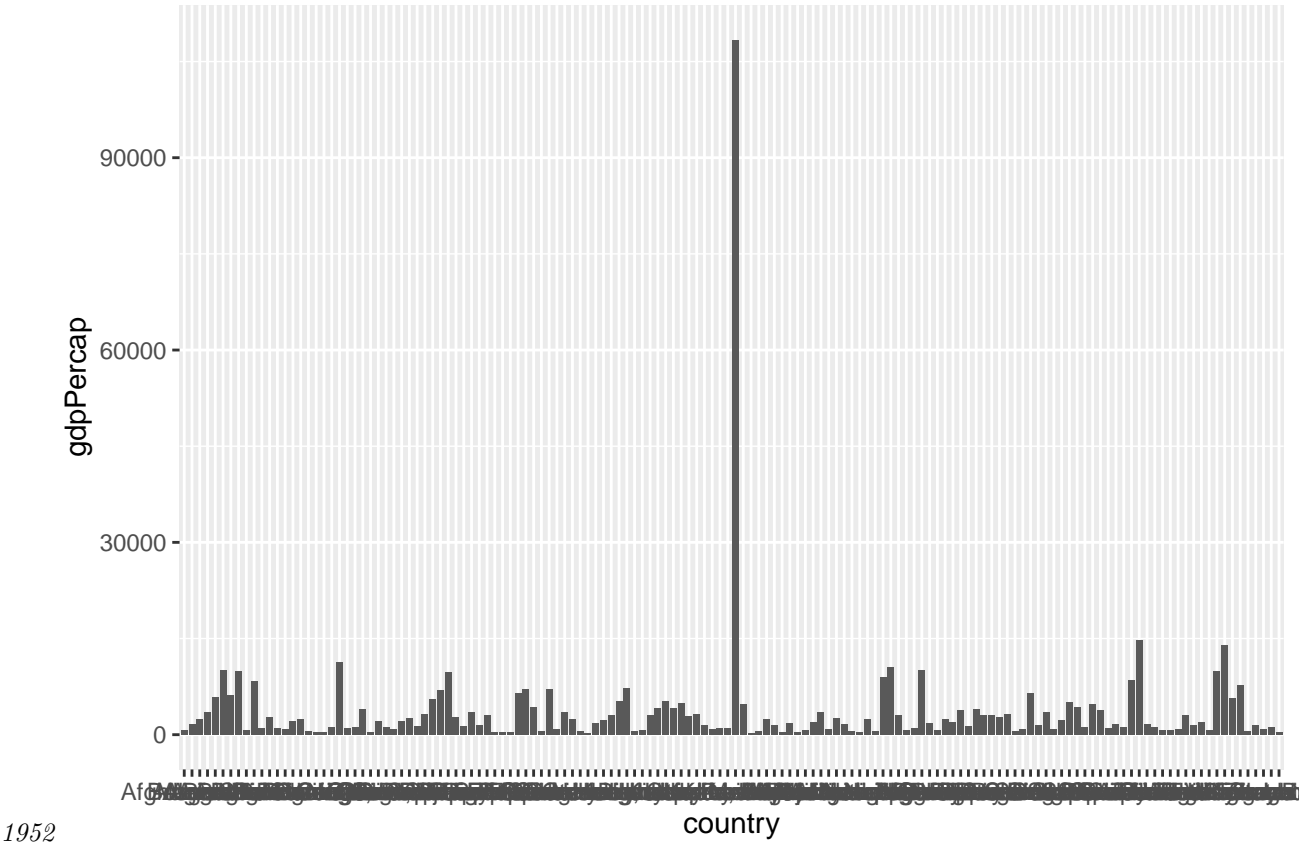
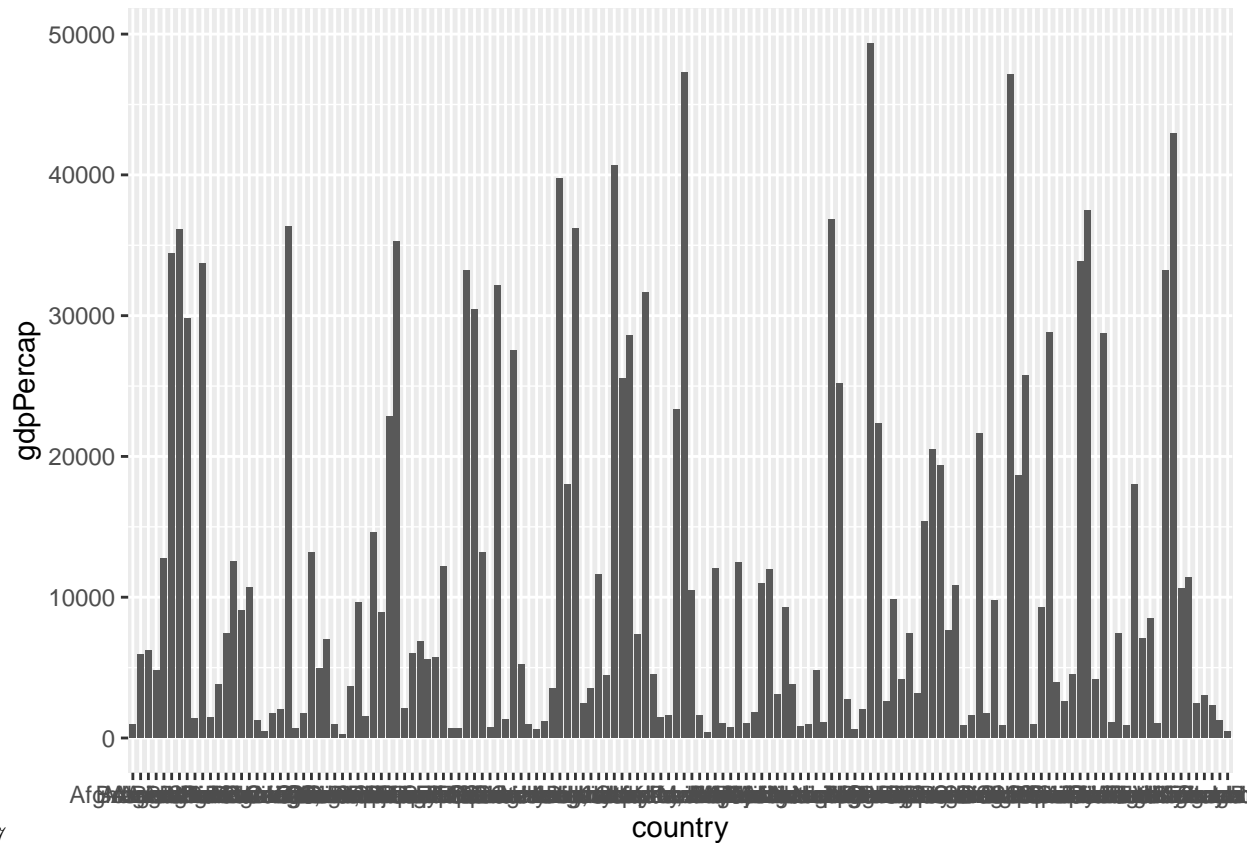


Table 3: Countries with negative growth in GDP

continent	country	growth
Africa	Central African Republic	-365.29418
Africa	Comoros	-116.84306
Africa	Congo, Dem. Rep.	-502.99047
Africa	Djibouti	-587.04791
Africa	Liberia	-161.06565
Africa	Madagascar	-398.24159
Africa	Niger	-142.20248
Africa	Sierra Leone	-17.24698
Africa	Somalia	-209.60877
Americas	Haiti	-638.72978
Americas	Nicaragua	-363.04298
Asia	Kuwait	-61075.36312



GDP in 2007

Which countries in the dataset have had periods of negative population growth? Illustrate your answer with a table or plot.

The countries listed above had negative growth in the period.

Which countries in the dataset have had the highest rate of growth in per capita GDP? Illustrate your answer with a table or plot.

According to the table, Singapore, Asia has the highest growth in GDP from 1952 to 2007.

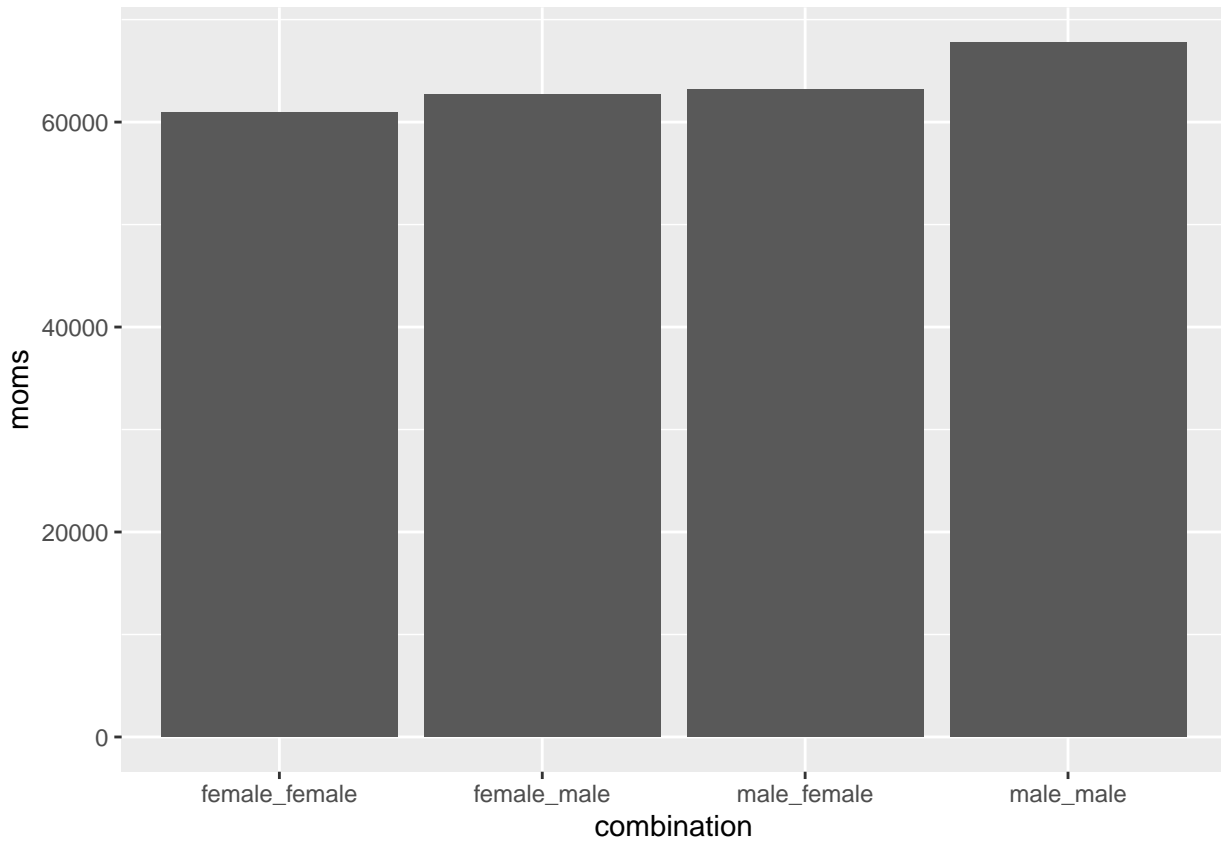
Problem 2 The data for Problem 2 is the Fertility data in the AER package. This data is from the 1980 US Census and is comprised of data on married women aged 21-35 with two or more children. The data report the gender of each woman's first and second child, the woman's race, age, number of weeks worked in 1979, and whether the woman had more than two children.

Table 4: 5 Countries with the highest growth in GDP

continent	country	growth
Europe	Austria	29989.42
Europe	Ireland	35465.72
Asia	Hong Kong, China	36670.56
Europe	Norway	39261.77
Asia	Singapore	44828.04

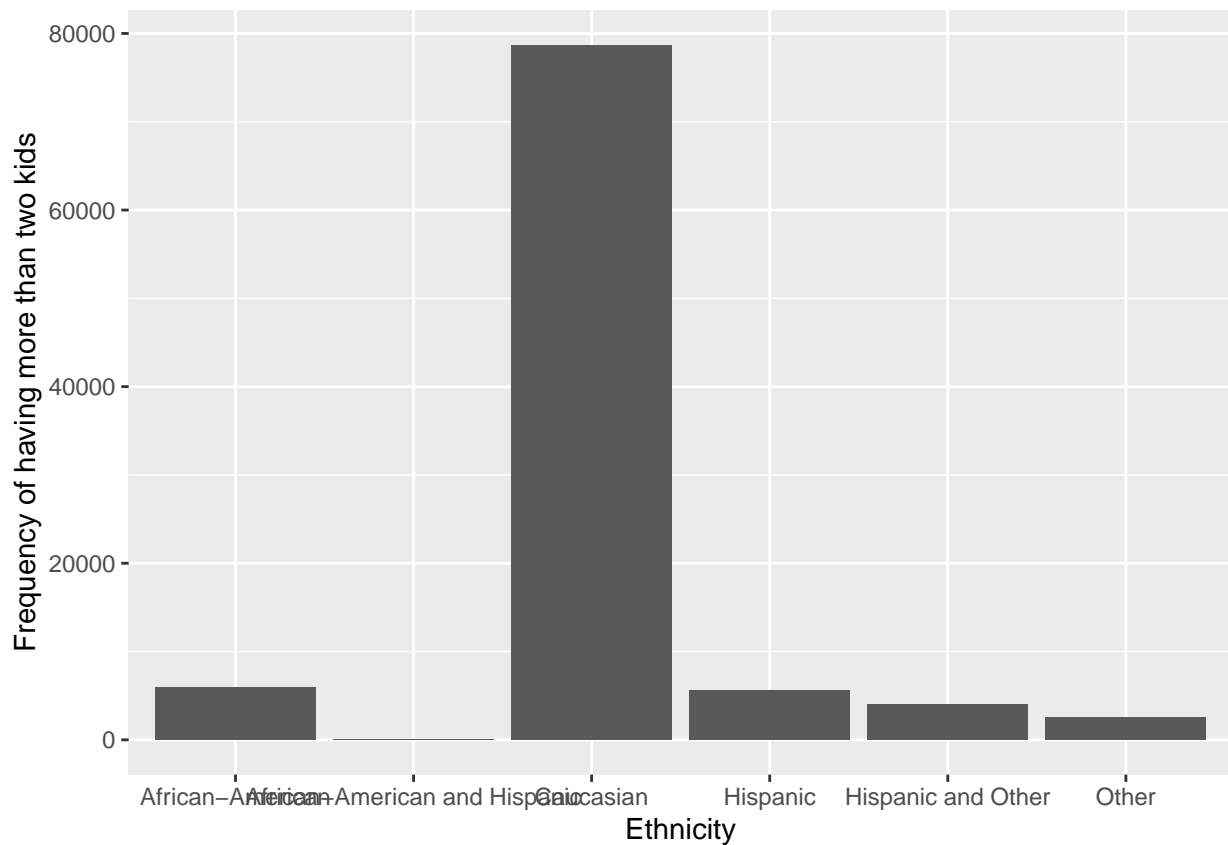
There are four possible gender combinations for the first two Children. Product a plot the contracts the frequency of these four combinations.

combination	moms
female_female	60946
female_male	62724
male_female	63185
male_male	67799



Are the frequencies different for women in their 20s and women who are older than 29?

From the bar chart we can tell that the frequencies of the four combinations of moms under 29 are obviously less than the frequencies of those above 29. However, the rough relationship of the four combinations does not vary too much. The frequencies of having two male children are leading in both groups. Produce a plot that contrasts the frequency of having more than two children by race and ethnicity.



Problem 3

Use the mtcars and mpg datasets.

How many times does the letter “e” occur in mtcars rownames?

There are 25 times that letter “e” occurs in mtcars rownames.

How many cars in mtcars have the brand Merc?

There are 7 cars from Mercedes.

How many cars in mpg have the brand(“manufacturer” in mpg) Merc?

There are 4 cars in mpg have the brand Mercedes.

Contrast the mileage data for Merc cars as reported in mtcars and mpg. Use tables, plots, and a short explanation. Is it referring to mpg by mileage?

Problem 4

Install the babynames package.

Draw a sample of 500,000 rows from the babynames data

Produce a tibble that displays the five most popular boy names and girl names in the years 1880,1920, 1960, 2000.

year	sex	name	n	prop
1880	M	William	9532	0.0805068
1880	M	Henry	2444	0.0206419
1880	M	Robert	2415	0.0203970
1880	M	Edward	2364	0.0199662
1880	M	Walter	1755	0.0148226

year	sex	name	n	prop
1920	M	June	73	6.63e-05
1920	M	Gentle	7	6.40e-06
1920	M	Arron	9	8.20e-06
1920	M	Perfecto	8	7.30e-06
1920	M	Franz	17	1.54e-05

year	sex	name	n	prop
1960	M	Jeffrey	28831	0.0133126
1960	M	Kevin	28388	0.0131080
1960	M	Gary	21684	0.0100125
1960	M	Anthony	19498	0.0090031
1960	M	Stephen	16259	0.0075075

year	sex	name	n	prop
2000	M	Nicholas	24652	0.0118100
2000	M	Daniel	22312	0.0106889
2000	M	David	19767	0.0094697
2000	M	James	17981	0.0086141
2000	M	Christian	16056	0.0076919

year	sex	name	n	prop
1880	F	Mary	7065	0.0723836
1880	F	Elizabeth	1939	0.0198658
1880	F	Annie	1258	0.0128887
1880	F	Laura	1012	0.0103683
1880	F	Nellie	995	0.0101942

year	sex	name	n	prop
1920	F	Betty	14017	0.0112673
1920	F	Marie	12743	0.0102432
1920	F	Alice	11596	0.0093212
1920	F	Irene	10073	0.0080970
1920	F	Catherine	8878	0.0071364

year	sex	name	n	prop
1960	F	Mary	51474	0.0247490
1960	F	Susan	39200	0.0188476
1960	F	Lisa	33702	0.0162041
1960	F	Pamela	22093	0.0106224
1960	F	Sharon	20424	0.0098200

year	sex	name	n	prop
2000	F	Elizabeth	15094	0.0075664
2000	F	Kayla	13312	0.0066731
2000	F	Victoria	10923	0.0054755
2000	F	Sydney	10242	0.0051341
2000	F	Destiny	9850	0.0049376

What names overlap boys and girls?

name	overlaps
Aaden	2
Aadyn	2
Aalijah	2
Aaliyah	2
Aaliyan	2
Aamari	2
Aaren	2
Aarian	2
Aarin	2
Aarion	2

Here are 10 illustrations.

What names were used in the 19th century but have not been used in the 21st century? 20th century?

```
## [1] "Rosanna" "Harriet" "Hollis" "Nona" "Jim" "Adda"
## [7] "Basil" "Cleve" "Johnie" "Jenny" "Al" "Honora"
## [13] "William" "Cynthia" "Roderick" "Isabelle" "Melvina" "Landon"
## [19] "Tina" "Irena"
```

Here are 20 illustrations.

Produce a chart that shows the relative frequency of the names “Donald”, “Hilary”, “Hillary”, “Joe”, “Barrack”, over the years 1880 through 2017.

