Homework 1

YOU

add your date here

**Remember: show your code!**

### this is a comment inside of a code chunk. You can use command + option + i or ctrl + alt + i to generate a code chunk. Click 'knit' above (or cmd + shift + k) to see the html output file.  
  
2 + 2

## [1] 4

a<-4  
b<-3  
  
a^b

## [1] 64

1. Read Chapter 2 of Kieran Healy’s Data Visualization (available for free here: <https://socviz.co/gettingstarted.html#gettingstarted>)
2. Confirm you have the needed packages installed by loading the tidyverse package. Install (if needed) and load the gapminder package

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5 ✓ purrr 0.3.4  
## ✓ tibble 3.1.3 ✓ dplyr 1.0.7  
## ✓ tidyr 1.1.3 ✓ stringr 1.4.0  
## ✓ readr 2.0.0 ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(gapminder)

1. Create a new variable in the tibble gapminder called pop1000 that is equal to the population of a country in 1,000s of people (pop/1000).

gapminder$pop1000 <- gapminder$pop / 1000

1. Compute the mean life expectancy for all observations in the dataset

mean(gapminder$lifeExp)

## [1] 59.47444

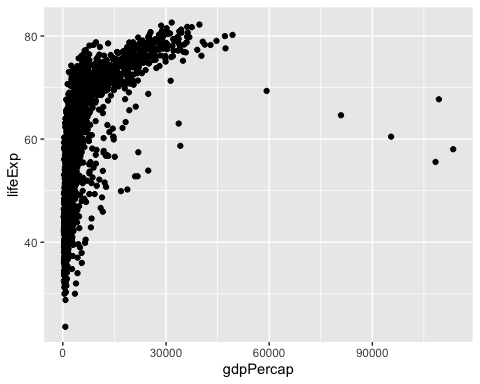
1. Use the unique() function to list all of the countries with observations in the gapminder tibble.

unique(gapminder$country)

## [1] Afghanistan Albania Algeria   
## [4] Angola Argentina Australia   
## [7] Austria Bahrain Bangladesh   
## [10] Belgium Benin Bolivia   
## [13] Bosnia and Herzegovina Botswana Brazil   
## [16] Bulgaria Burkina Faso Burundi   
## [19] Cambodia Cameroon Canada   
## [22] Central African Republic Chad Chile   
## [25] China Colombia Comoros   
## [28] Congo, Dem. Rep. Congo, Rep. Costa Rica   
## [31] Cote d'Ivoire Croatia Cuba   
## [34] Czech Republic Denmark Djibouti   
## [37] Dominican Republic Ecuador Egypt   
## [40] El Salvador Equatorial Guinea Eritrea   
## [43] Ethiopia Finland France   
## [46] Gabon Gambia Germany   
## [49] Ghana Greece Guatemala   
## [52] Guinea Guinea-Bissau Haiti   
## [55] Honduras Hong Kong, China Hungary   
## [58] Iceland India Indonesia   
## [61] Iran Iraq Ireland   
## [64] Israel Italy Jamaica   
## [67] Japan Jordan Kenya   
## [70] Korea, Dem. Rep. Korea, Rep. Kuwait   
## [73] Lebanon Lesotho Liberia   
## [76] Libya Madagascar Malawi   
## [79] Malaysia Mali Mauritania   
## [82] Mauritius Mexico Mongolia   
## [85] Montenegro Morocco Mozambique   
## [88] Myanmar Namibia Nepal   
## [91] Netherlands New Zealand Nicaragua   
## [94] Niger Nigeria Norway   
## [97] Oman Pakistan Panama   
## [100] Paraguay Peru Philippines   
## [103] Poland Portugal Puerto Rico   
## [106] Reunion Romania Rwanda   
## [109] Sao Tome and Principe Saudi Arabia Senegal   
## [112] Serbia Sierra Leone Singapore   
## [115] Slovak Republic Slovenia Somalia   
## [118] South Africa Spain Sri Lanka   
## [121] Sudan Swaziland Sweden   
## [124] Switzerland Syria Taiwan   
## [127] Tanzania Thailand Togo   
## [130] Trinidad and Tobago Tunisia Turkey   
## [133] Uganda United Kingdom United States   
## [136] Uruguay Venezuela Vietnam   
## [139] West Bank and Gaza Yemen, Rep. Zambia   
## [142] Zimbabwe   
## 142 Levels: Afghanistan Albania Algeria Angola Argentina Australia ... Zimbabwe

1. Create a scatterplot (using ggplot) showing lifeExp on the y axis and gdpPercap on the x axis. Provide a brief interpretation of your plot.

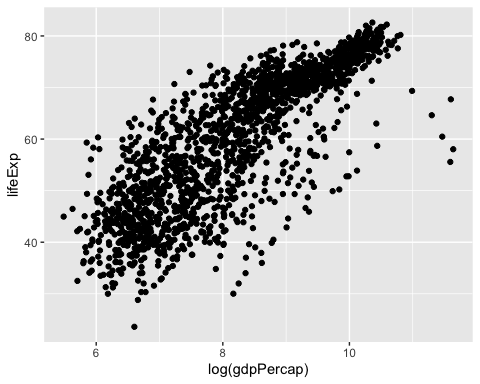
ggplot(gapminder,   
 aes(x = gdpPercap,  
 y = lifeExp)) +   
 geom\_point()



This plot shows that countries with higher levels of economic output, as measured by GDP per capita, tend to have much higher life expectancies than countries with lower levels of economic output.

1. Reproduce the plot from 6, but instead, plot log(gdpPercap) on the x axis. Provide a brief interpretation of your plot.

ggplot(gapminder,   
 aes(x = log(gdpPercap),  
 y = lifeExp)) +   
 geom\_point()



We see here that the relationship between GDP per capita and life expectancy is approximately linear after a log transformation of GDP per capita (log-linear for short). This suggests that small increases in GDP per capita when GDP per capita is low are associated with large increases in life expectancy. When GDP per capita is high, we see relatively smaller increases in life expectancy as GDP per capita increases. In other words, the relationship between economic output and life expectancy is positive, but non-linear.