HW 3

Frank Edwards

2/8/2019

Required reading:

Healy, Kieran. "Data Visualization: A Practical Introduction" chapters 1 and 3.

Available for free through a pre-print of the book:

http://socviz.co/lookatdata.html#lookatdata

and

http://socviz.co/makeplot.html#makeplot

For this homework, you will use data on national life expectancies from the gapminder package.

You can install the package with the following command

```
install.packages("gapminder")
```

And load the data with the following command

```
library(gapminder)
```

You must use ggplot for all graphics on this assignment

Question 1

1. Provide a histogram of life expectancy (lifeExp). Interpret this histogram.
2. Provide a histogram of GDP per capita (gdpPercap). Interpret this histogram.
3. Provide a scatterplot of life expectancy by GDP per capita
4. Transform GDP per capita using a log scale either manually or using scale_x_log10()
5. Interpret the relationship between life expectancy and gdp per capita

Question 2

1. Using your scatterplot from 1.5., add a color aesthetic to your plot for continent.
2. Now, replace your color aesthetic with a shape aesthetic.
3. Which plot is easier to interpret, 2.1 or 2.2?

Question 3

1. Estimate a regression equation for the relationship between life expectancy, GDP per capita, and continent.
2. Provide a display of the estimated model using broom::tidy()
3. Provide a visual that summarizes your estimated relationship

Question 4

If you have already obtained your data for the final project, provide three interesting and/or useful data visuals of your dataset. If you have not obtained the data yet, you may complete this question next week.