Problem Set #3

EH6105 - Quantitative Methods

Steven V. Miller

This homework makes use of data available in `{peacesciencer}` and necessitates the use of `{tidyverse}` to at least generate the data. `{tidyverse}` is not necessary to answer these questions though it will assuredly make the process easier. Load these two libraries to get started answering these questions.

```
library(tidyverse)
library(peacesciencer)
```

Democracy and Economic Development in 2015

This homework will deal with what is, perhaps, the biggest "chicken or egg" problem in the literature on international relations and international development. What is the relationship between democracy and economic wealth/development? There have been a lot of trees sacrificed for paper for this topic, and it's debatable what exactly causes what.[1] We will have to sidestep this debate completely and just roll with an assumption democracy is an independent variable of interest to us and economic development is a dependent variable so that we might learn about bivariate OLS. We are going to limit our analysis to just all sovereign states in the Correlates of War state system data in 2015.

You will have to run the following commands in R to create the data.

```
# Create all state-years possible from 1816 to present, but return just 2015, and...
create_stateyears(subset_years = 2015) %>%
  # add assorted democracy data, and...
  add_democracy() %>%
  # add some economic data from Anders et al. (2020), and
  #  then assign to object called `Data`.
  add_sdp_gdp() -> Data
```

Here's a preview of these data. Note: you won't use all these data you see here.

```
Data
```

```
## # A tibble: 195 x 10
##    ccode statenme    year v2x_polyarchy polity2 xm_qudsest wbgdp2011est wbpopest
```

[1] From one perspective: economic development is only likely to emerge when property rights are stable and governments make credible commitments to protect private property for investment (i.e. democracy precedes economic growth). From another perspective: nascent economic development creates the kind of capital necessary to challenge rulers and demand these protections in the first place (i.e. the conditions for economic development precede demands for democracy).

```
##    <dbl> <chr>      <dbl>         <dbl>   <dbl>      <dbl>         <dbl>    <dbl>
##  1     2 United St~  2015         0.89       10       1.83          30.6     19.6
##  2    20 Canada      2015         0.839      10       2.32          28.1     17.3
##  3    31 Bahamas     2015        NA          NA       1.59          22.9     12.8
##  4    40 Cuba        2015         0.182      -7      -0.420         25.9     16.2
##  5    41 Haiti       2015         0.425       0       0.254         23.6     16.1
##  6    42 Dominican~  2015         0.555       8       1.03          25.7     16.1
##  7    51 Jamaica     2015         0.833       9       1.24          23.9     14.8
##  8    52 Trinidad ~  2015         0.733      10       1.35          24.5     14.1
##  9    53 Barbados    2015         0.78       NA       1.57          22.1     12.5
## 10    54 Dominica    2015        NA          NA       1.59          20.5     11.2
## # ... with 185 more rows, and 2 more variables: sdpest <dbl>,
## #   wbgdppc2011est <dbl>
```

The student must read the codebooks {peacesciencer} for these functions in order to understand the questions, which you can do here:

```
# for the functions
# library(peacesciencer)
?add_democracy()
?add_sdp_gdp()

# for the data
?ccode_democracy # where the data for add_democracy() come from
?cow_sdp_gdp # where add_sdp_gdp() data come from
```

Answer these questions. A successful answer of these question must include the R code you used to help you answer the question.

1. It is not a formal assumption of OLS that the dependent variable is normally distributed, but the normality assumption for the conditional distribution of errors wants to imply the marginal distribution of the dependent variable is also normal. Let's assume our dependent variable is `wbgdppc2011est`. Show me how you might look at the distribution of this variable for these observations. Describe it to me.
2. Let's assume our primary independent variable for democracy is Xavier Marquez' "Quick UDS" extentions (`xm_qudsest`). Show me how you might look at the distribution of this variable for these observations. Describe it to me.
3. (2 POINTS) Let's run a naive bivariate OLS regression that regresses `wbgdppc2011est` on `xm_qudsest`. Describe the results to me.
4. You can do a proto-bivariate OLS with {ggplot2} and a combination of `geom_scatter()` and `geom_smooth(method = "lm")`. Create a scatter plot with linear trend for our independent variable and dependent variable.
5. Change the smoother on this scatter plot to a LOESS smoother (i.e. `geom_smooth(method = "loess")`). What do you see and what do you think this implies for the effect of democracy on economic development?
6. It's not uncommon you'll be asked to do "robustness tests" for an independent variable of interest. We have a few other democracy indicators in the data frame as well. Regress `wbgdppc2011est` on `polity2` and describe the results to me. Make sure you understand the distribution of the independent variable.

7. Regress `wbgdppc2011est` on `v2x_polyarchy` and describe the results to me. Make sure you understand the distribution of the independent variable.