

Problem Set #1

IRIII.2 - Quantitative Methods in the Study of International Relations

Steven V. Miller

This homework is intended to be gentle, simply introducing students to playing around in the R programming language and should not demand too much energy on behalf of the student in order to complete. Its completion assumes installation and use of the `{tidyverse}`. Students who submit their problem sets as .R scripts should make this clear, certainly if they use any `{tidyverse}` functionality in answering these questions. Failure to do this will result in a one-point deduction.

Kilowatt Hours per Capita and GDP per Capita, 2010

This homework assignment will use data included in `{stevethemes}` as the `kwh_gdp` data set. This is a very simple data set that includes GDP per capita data and kilowatt hours consumed per capita data of over 130 sovereign states in 2010.

```
?kwh_gdp
```

Here's a little preview of these data.

```
kwh_gdp
```

```
## # A tibble: 136 x 7
##   country    iso3c  year  kwhpc  gdppc  ln_kwhpc  ln_gdppc
##   <chr>      <chr> <int> <dbl> <dbl>   <dbl>   <dbl>
## 1 Albania    ALB    2010  1943.  4094.   7.57    8.32
## 2 Algeria    DZA    2010  1020.  4496.   6.93    8.41
## 3 Angola     AGO    2010   206.  3497.   5.33    8.16
## 4 Argentina  ARG    2010  2878. 10386.   7.96    9.25
## 5 Armenia    ARM    2010  1686.  3143.   7.43    8.05
## 6 Australia  AUS    2010 10727. 52134.   9.28   10.9
## 7 Austria    AUT    2010  8383. 46904.   9.03   10.8
## 8 Azerbaijan AZE    2010  1603.  5844.   7.38    8.67
## 9 Bahrain    BHR    2010 18443. 21187.   9.82    9.96
## 10 Bangladesh BGD    2010   246.   777.   5.50    6.66
## # i 126 more rows
```

Answer these questions. A successful answer of these question must include the R code you used to help you answer the question. Each question is worth a point.

1. There are multiple copies of this data frame available for manual download in Athena (check the "Data"

plan).¹ Download one of these copies, in whatever format you want, and provide for me the command that you would use to load it in R and assign it to an object called `my_data`.

2. How many columns are there in this data frame? Use a relevant command in R to tell me how many columns there are in this data frame, and what those column names are.
3. How many rows are there in this data frame? Use a relevant command in R to answer this question.
4. What is the mean GDP per capita of these more than 130 sovereign states in 2010? What is the median of this same variable? Provide the R code that produced the statistics you report.
5. What is the mean of the log-transformed GDP per capita of these more than 130 sovereign states in 2010? What is the median of this same variable? Provide the R code that produced the statistics you report.

Extra Credit (1 Point Each)

6. For the fifth question, write three-complete sentences that offer an explanation of why the differences in medians/means are so much different in the raw vs. log-transformed GDP per capita measures. This is an extra credit question, so I'm expecting some effort here. Your answer here does not require code.
7. What is the mean GDP per capita for the Nordic countries of Denmark, Finland, Iceland, Norway, and Sweden? What is the mean GDP per capita for the countries not in the Nordics? Prove it to me in code.

¹The file name is `kwh_gdp` with extensions communicating whether it's a comma- or tab-separated values file, an R serialized data frame, or a proprietary format for use in Stata or SAS. The student should take initiative to know what is what, certainly as it pertains to "answering" this "question."