

Data Visualization I: Single Variable Plots – Homework

SPEC Lab: Ben Graham and Marie Zaragoza

January 2022

Single-Variable Descriptive Figures

This homework assignment asks you to make a histogram and, for the bonus, a kernel density plot, using replication data from a recent SPEC Lab paper.

Markowitz, Jonathan, Suzie Caldwell Mulesky, Benjamin A.T. Graham, and Christopher J. Fariss. 2020. “Productive Pacifists: The Rise of Production-Oriented States and Decline of Profit-Motivated Conquest.” *International Studies Quarterly*, 64(3): 558-572.

As with most things in R, there are many different ways to complete these exercises. We post both both the uncompleted homework and the answer key. The code in the answer key is one example of how to complete this task, and represents current practice in the SPEC Lab. But the answer key code is by no means the only way to complete the exercise. If you think your code is even better than the code in our answer key, we want to hear about it! Contact benjamin.a.graham@usc.edu to share your awesomeness.

```
#Each person's setwd() function will look a little different,  
#depending on where they have saved the training data folder.  
#It may look like:  
#setwd("/Volumes/GoogleDrive/My Drive/Training Data August 2021")
```

```
library(ggplot2) #ggplot2 creates graphs  
library(tidyverse) #tidyverse is used for data management  
library(readr) #readr is needed for the read_csv() function
```

Exercise 1: Create a histogram of GDP Per Capita in the year 2000 (WorldBank_gdppc_2010_con_estimate) Make this figure in color. Make it pretty. Give it an informative title.

```
#After setting your working directory to reflect your desired folder, load in your data for the project
```

```
load("ProductivePacifists_Data.RDATA")
```

```
#First create a smaller dataset with only the relevant year.  
smalldat <- dat %>%  
  filter(year %in% c(2000))
```

```
#OR
```

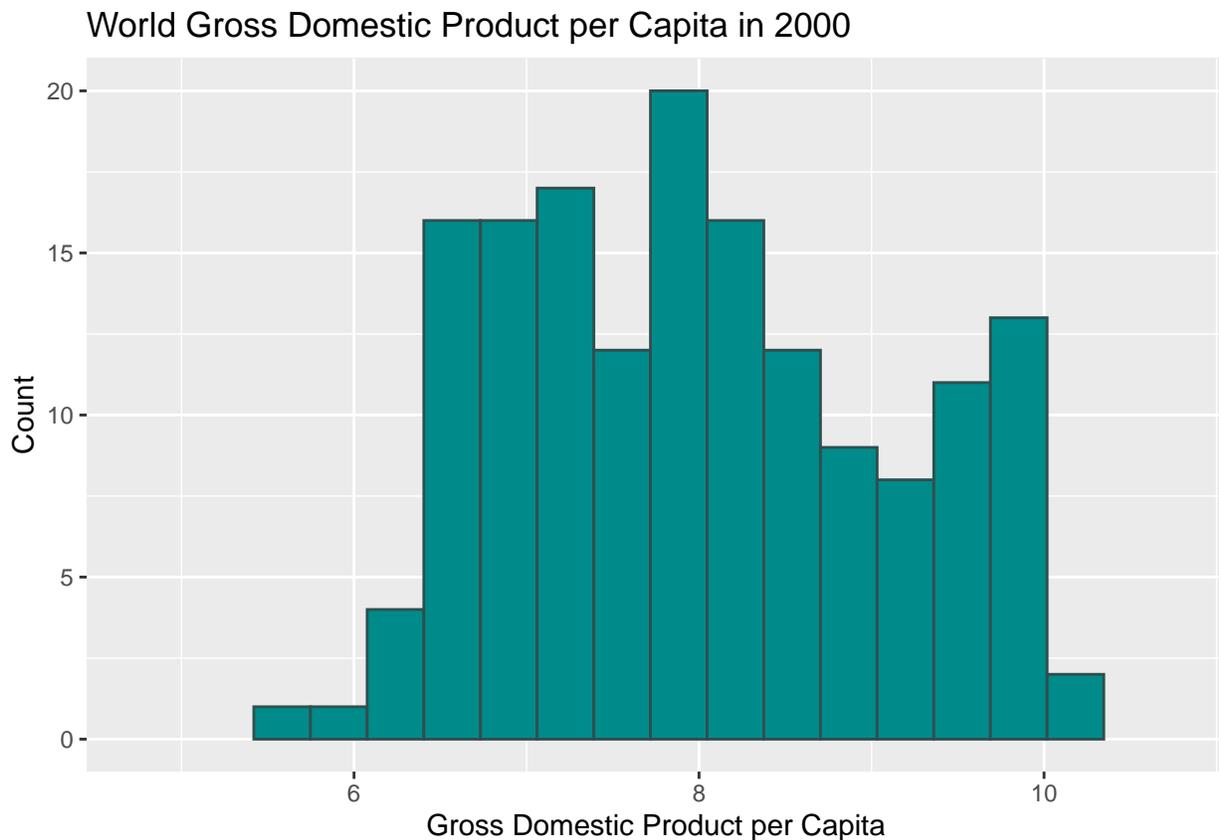
```
smalldat <- dat %>%  
  filter(year == 2000)
```

```
#To add color to the histogram, play around with the color =  
# and fill = sections within geom_histogram().
```

```
#To add labels, use labs().
```

```
#To change the coordinates of the histogram, use coord_cartesian().
```

```
gdppc_hist <- ggplot(smalldat, aes(x = WorldBank_gdppc_2010_con_estimate)) +  
  geom_histogram(bins = 15, color = "darkslategray", fill = "cyan4")+  
  labs(title = "World Gross Domestic Product per Capita in 2000",  
        x = "Gross Domestic Product per Capita",  
        y = "Count") +  
  coord_cartesian(ylim = c(0, 20), xlim = c(4.75,10.75))  
  
print(gdppc_hist)
```



Optional Bonus Use facet wrapping to create a kernel density plot of the continuous measure of land-orientation for the year 1900, 1950, and 2000 (`land_oriented_medium_continuous`). Make this figure in color. Make it pretty. Give it an informative title.

```
#Again, create different subset of the original data.
```

```
smalldat2 <- dat %>%  
  filter(year %in% c(1900,1950,2000))
```

```
#Then use that to then create the kernel density plot.
```

```
land_kernel<- ggplot(smalldat2, aes(x = land_oriented_medium_continuous))+  
  geom_density(fill = "purple", alpha = 0.4, color = NA)+  
  labs(title = "Land Orientation of National Economies",  
        subtitle = "Data source -- Markowitz et al. 2021",  
        x = "Land Use Per Year",  
        y = "Density") +
```

```
coord_cartesian(ylim = c(0, 0.7)) + #sets the range on the y axis
facet_wrap(~year)
```

```
print(land_kernel)
```

Land Orientation of National Economies

Data source -- Markowitz et al. 2021

