Data-Viisualization-Examples

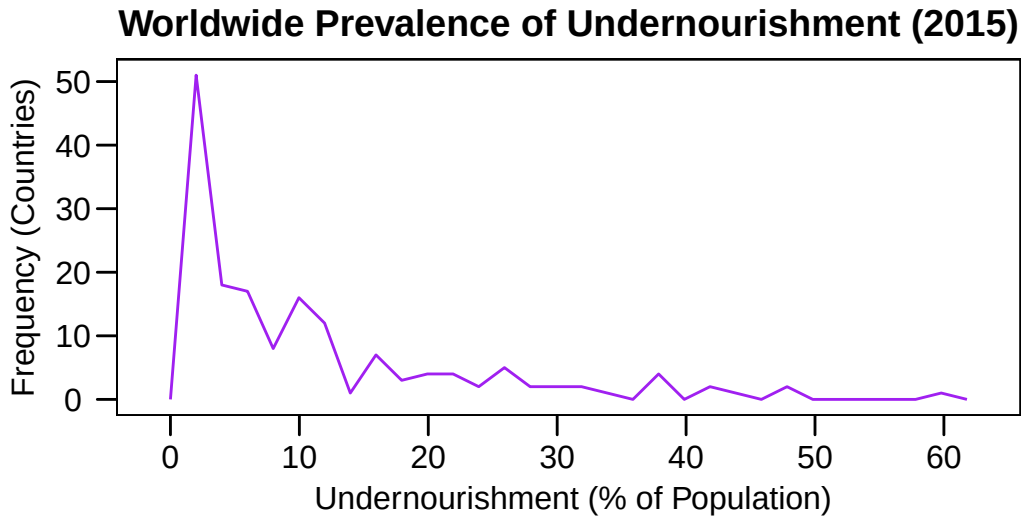
Marvin Williams

4/25/2021

Starvation

The frequency polygon shows the distribution for the rate of undernourishment worldwide according to country count. The rate of undernourishment is calculated as the percentage of those in the population whose caloric intake fails to meet the minimum energy requirements. From our data, we can see that the majority of countries have a rate of undernourishment that is less than 15%. According to the graph, 50% can be seen having a less than 5% undernourishment rate. As the rate of undernourishment according to population increases past the 15% cutoff, the frequency of countries who meet this threshold remains relatively low, at less than 5%.

```
ggplot(WDI, aes(starvation)) +
  geom_freqpoly(stat = "bin", color = 'purple', alpha = 5)+
  labs(x = " Undernourishment (% of Population)", y = "Frequency (Countries)", title = "Worldwide Preval
  theme_par()+
  scale_x_continuous(breaks = hist_breaks)
```

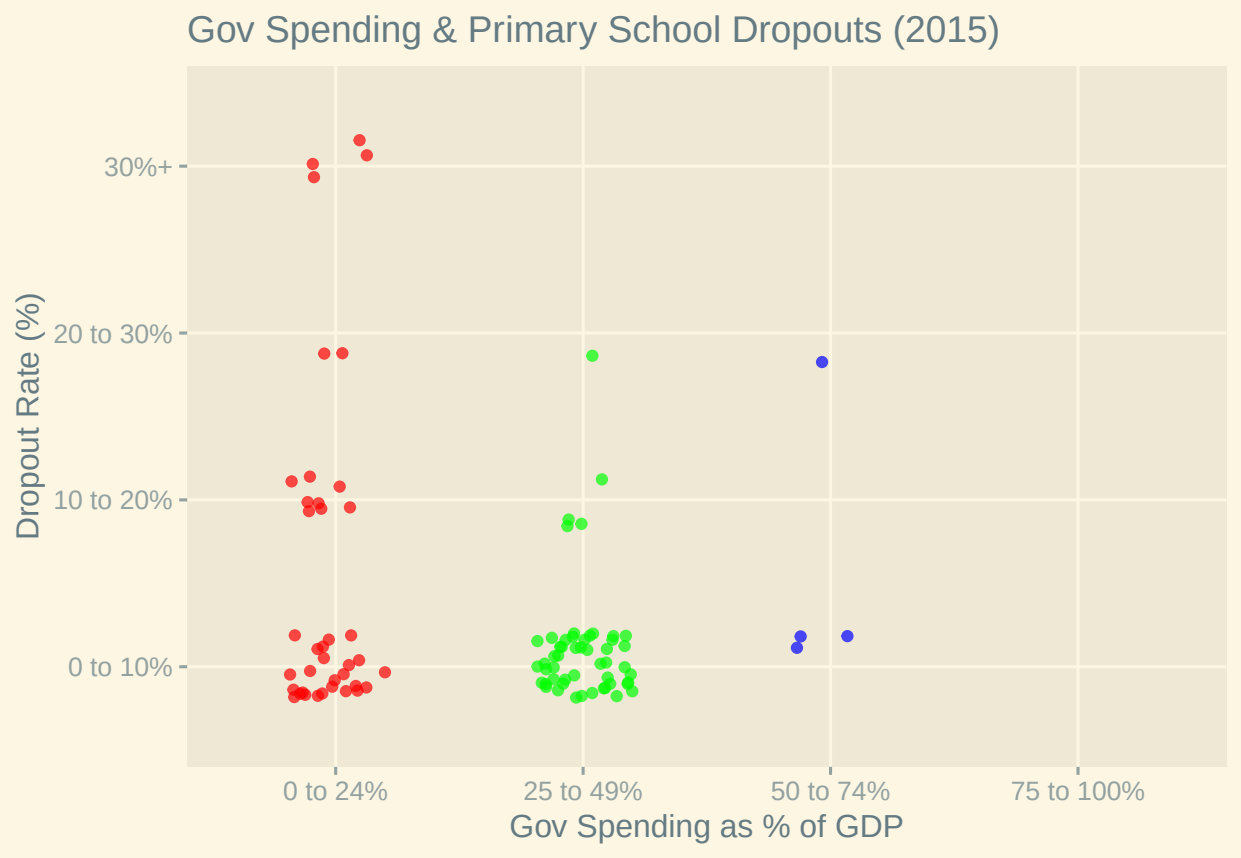


Government Spending & GDP Rates

The jitter plot shows us a bivariate relationship between government spending as a percentage of GDP and primary school dropout rates. Gov spending and primary school dropout rates have been recorded into 4 different categories. From our plot, we see that there is a negative correlation between government spending as % of GDP and primary school dropout rates. Per our plot, we see that the highest rate of primary school dropout occurs in countries with the lowest GDP spending, at 0 to 25%. As for spending increases, we see dropout rates decreasing, with lower rates for 25 to 50%, and even lower for those in the 50 to 75% threshold, and practically non-existent for countries spending 75 to 100%. From this association, we can infer that an

increase in government spending is associated with a lower primary school dropout rate.

```
ggplot(WDI, aes(x = WDI_2, y = WDI_3))+
  geom_jitter(position = position_jitter(width = 0.2, height = 0.2), alpha = .7, aes(color = WDI_2))+
  scale_x_discrete(na.translate= FALSE)+
  scale_y_discrete(na.translate= FALSE)+
  labs(title = 'Gov Spending & Primary School Dropouts (2015)', x = "Gov Spending as % of GDP" , y  = "
  theme_solarized_2()+
  scale_color_manual(name="Dropout Rate",
                     labels=c("0 to 24%","25 to 49%","50 to 74%", '75 to 100%', "    -"),
                     values=c("red","green","blue", 'purple', " "), guide = FALSE)
```

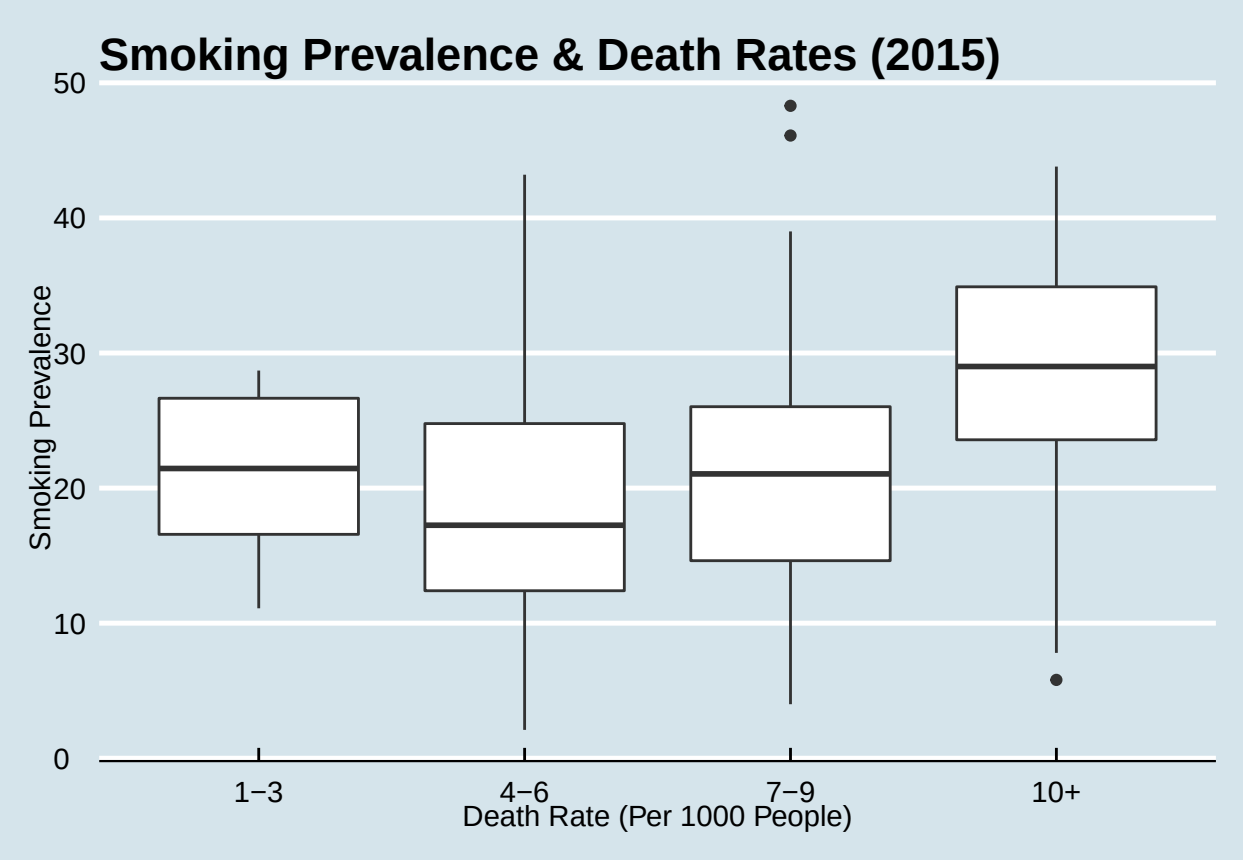


Smoking & Death Rates

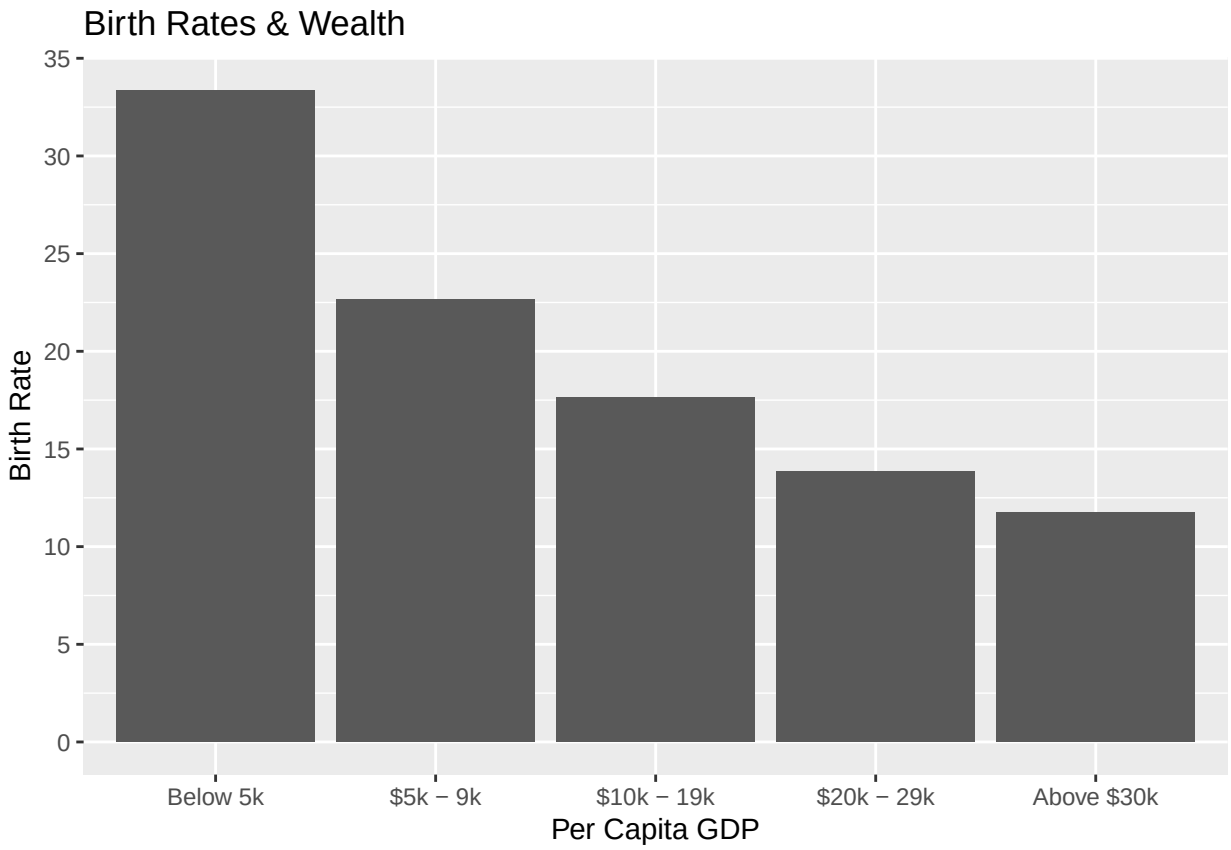
The boxplot shows the association between a country's smoking population and crude death rate per 1000 people. The lower whisker for those countries whose death rate is 1-4 is higher than other groups at about 11, as well as having a higher median of about 22 compared to about 16 for countries with a 4-7 rate, 21 for countries with a 7-10 death rate, and about 28 for countries with a 10+ death rate. Overall, the IQR for countries with a death rate of about 1-4 is higher than all groups, aside from 10+. Although it may be tempting to conclude that there is a link between countries' smoking prevalence there are many other spurious variables that need to be taken into account.

```
ggplot(WDI, aes(x = WDI_death, y = WDI$smoking))+
  geom_boxplot(rnscale=  'area')+
  scale_x_discrete(na.translate = FALSE)+
  theme_economist()+
```

```
  labs(title = "Smoking Prevalence & Death Rates (2015)", x = "Death Rate (Per 1000 People)", y = "Smoki
```




```
ggplot(BR, aes(x= WDI_Income, y = BR$`WDI$birth.rate`))+
  geom_bar(stat = 'identity')+
            labs(title = "Birth Rates & Wealth", x = 'Per Capita GDP', y = "Birth Rate") +
  scale_y_continuous(breaks= y_breaks)
```


Birth Rates & Wealth
Birth Rate
35
30
25
20
15
10
5
0
Below 5k
$5k − 9k
$10k − 19k
$20k − 29k
Above $30k
Per Capita GDP