Project1

Sean Letzer

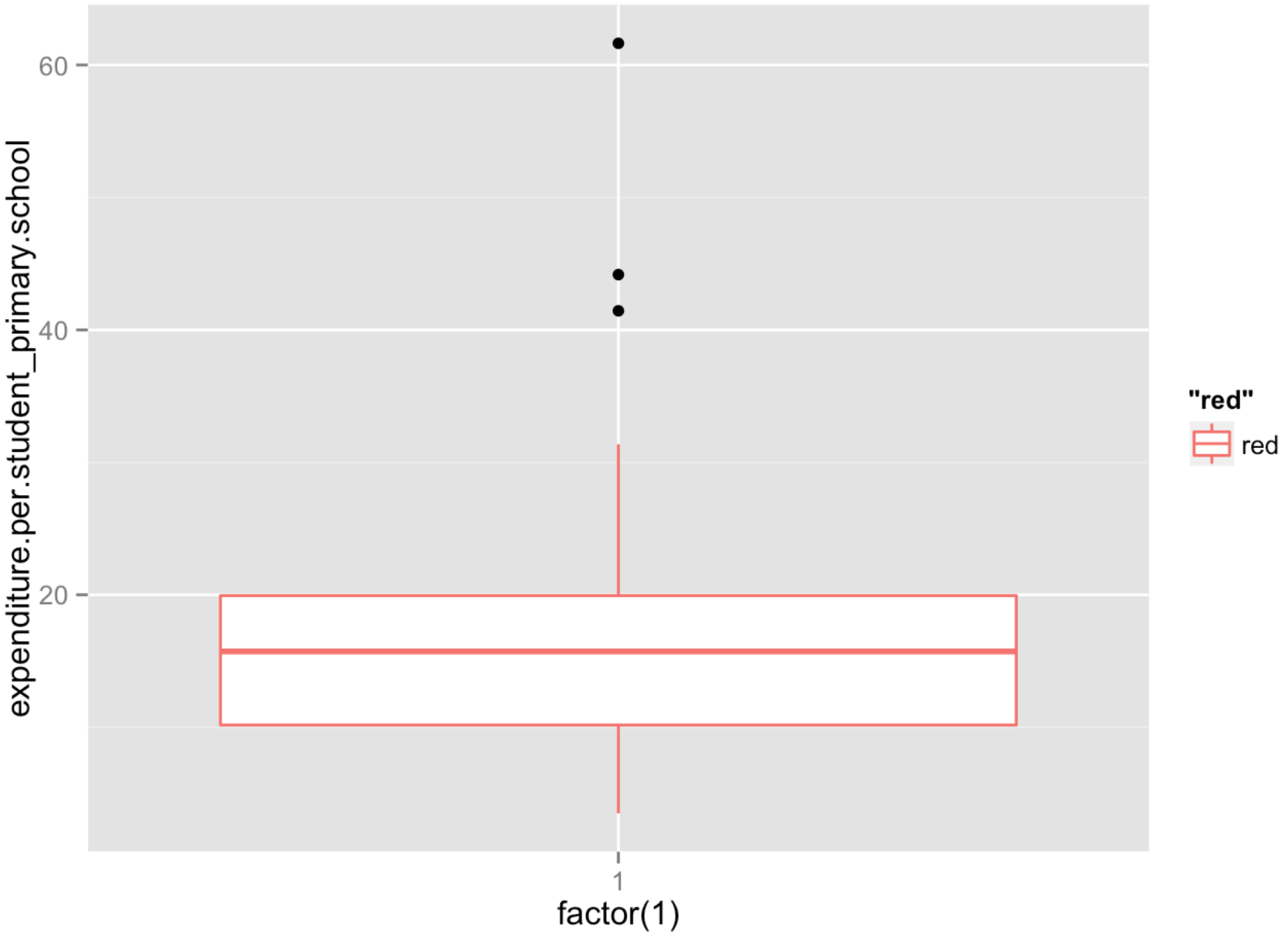
January 29, 2015

1. Calculate summary statistics and a statistical graphic appropriate for any numeric variable. Explain what the graphic and the statistics mean in the context of the variable.

The summary stats:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   3.504  10.170  15.720  16.240  19.940  61.640      77
```

Standard Deviation: 7.9908923



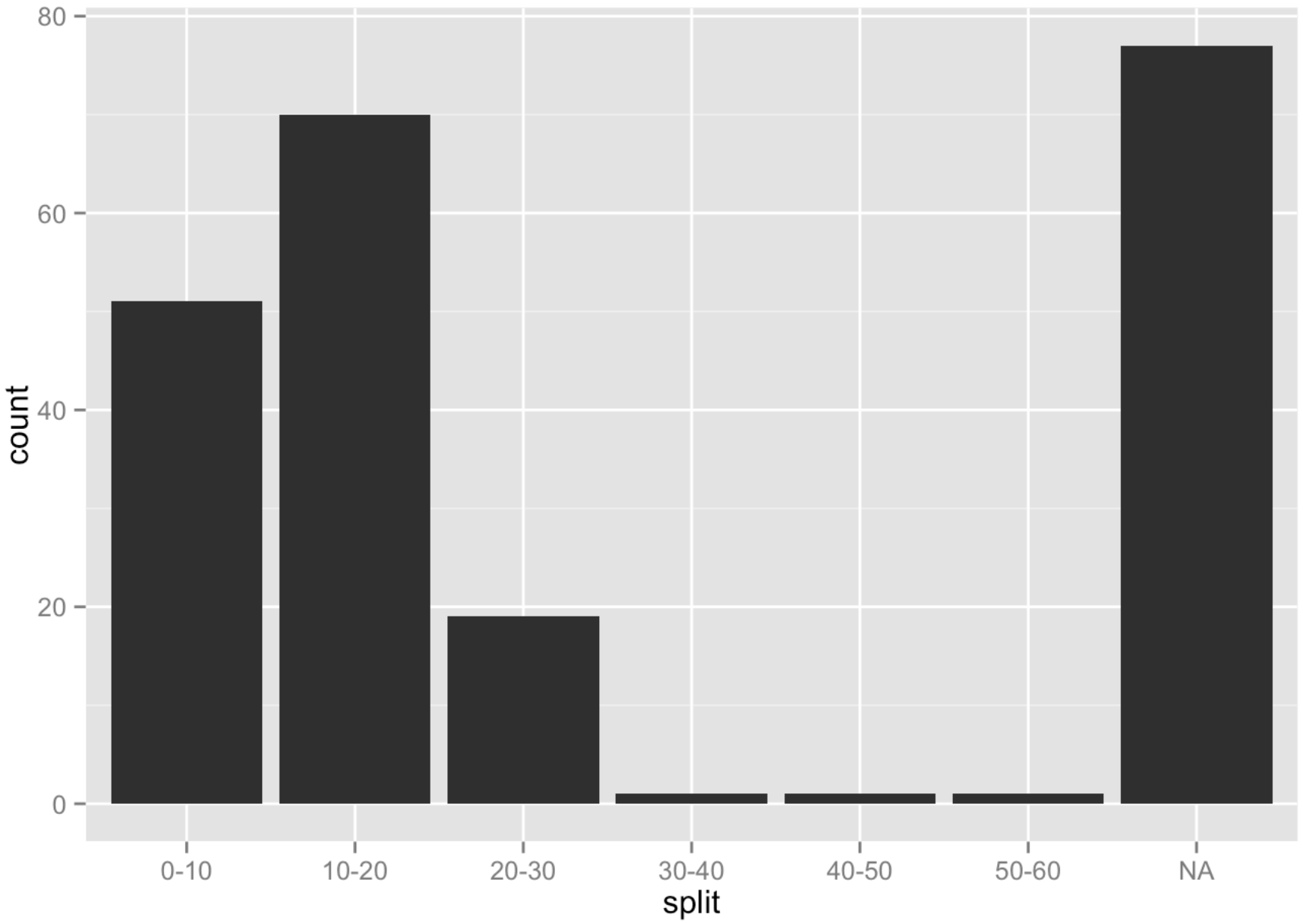
The box plot was appropriate for this numeric variable because the range of values was significantly small only about (0 to 60), and will easily surface any outliers. The data represented in the chart does a good job of depicting the summary of the data, for instance the range of the IQR is small, which implies that there is a small ammount of variance within the data which is suggested in the summary statistics by the standard deviation! This means that most kids in primary school spend roughly 16 units of currency.

2. Turn your variable into a categorical variable using the cut function (you may also want the quantile function). You may name your categories with words (eg. high, medium, low) or as ranges. Choose what you feel is an appropriate number of categories. Calculate summary statistics and a statistical graphic appropriate for the categorical variable. Explain what the graphic and the statistics mean in the context of the variable.

The summary stats:

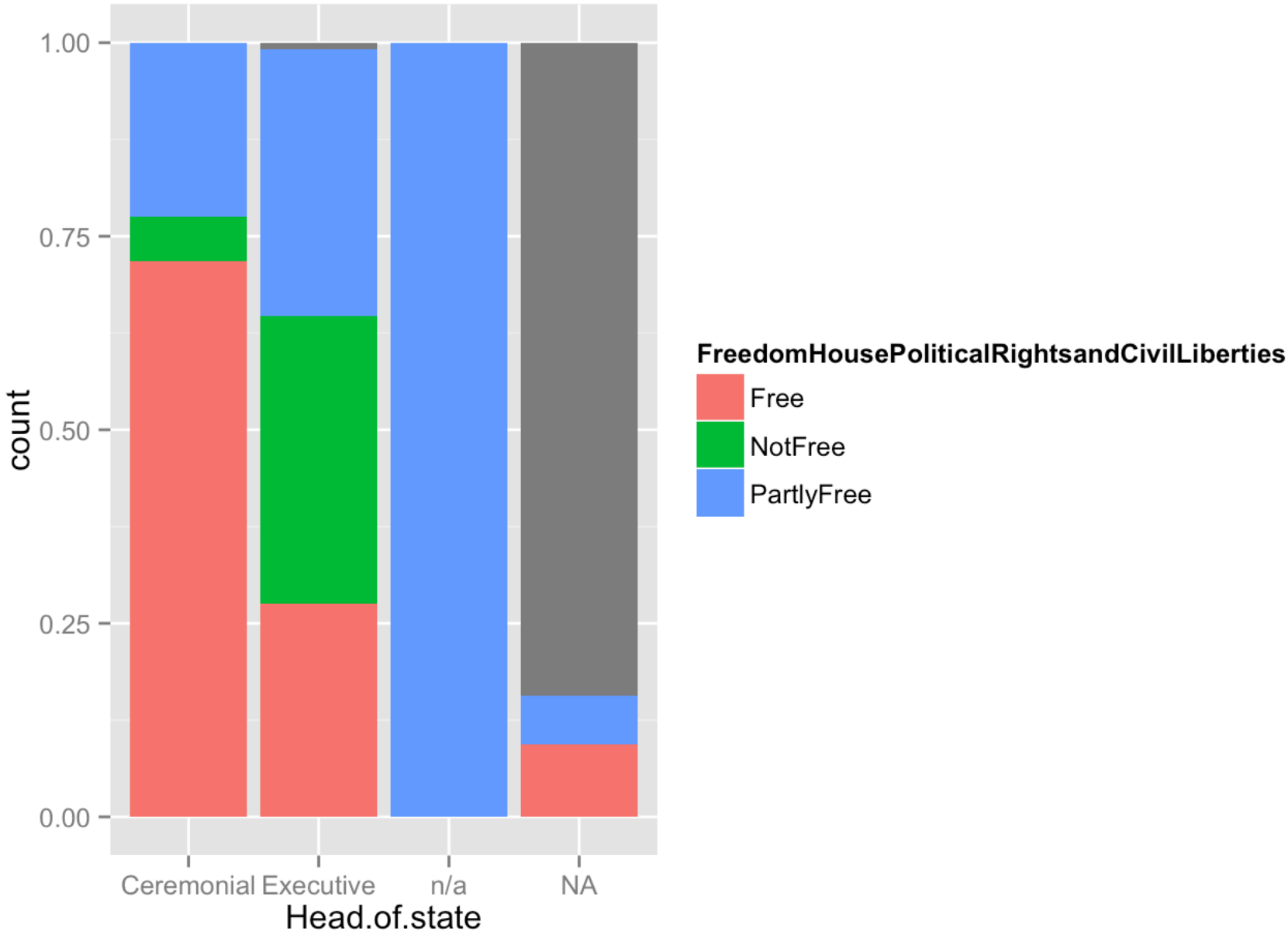
```
##  0-10 10-20 20-30 30-40 40-50 50-60  NA's
##    51    70    19     1     1     1    77
```

The standard Deviation: 0.8192263



Tansforming the numerical variables into categorical variables, surfaces that the data is skewed to the right, and is multi-modal. This corresponds to the summary that gives the number of values in each category, you can even see the multi-modal relationship just by looking at the first four categories. What is interesting that this model demonstrates, is that given this variable, it must have been hard to collect data, because the largest number of entries were “Not Available” or NA.

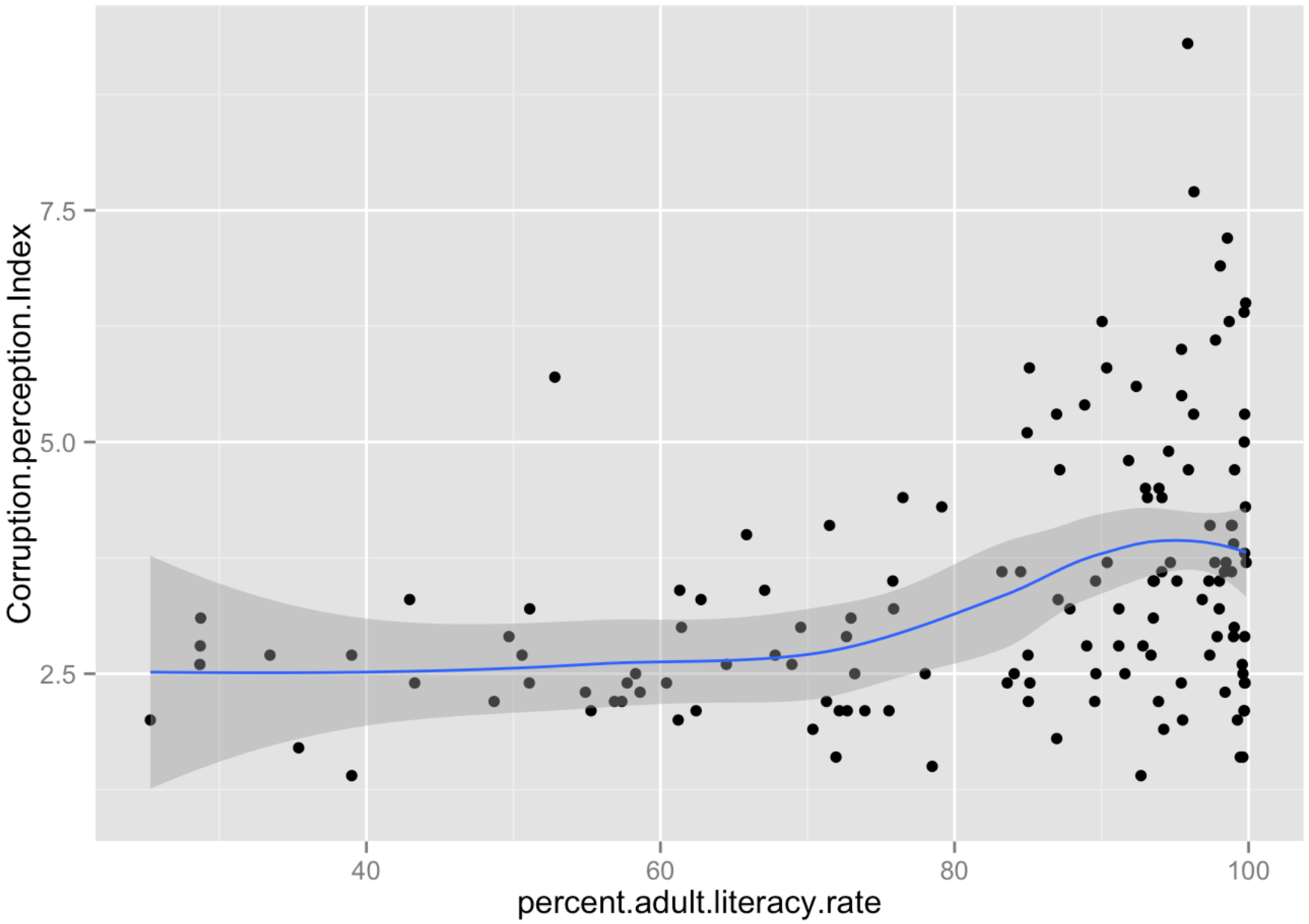
3. Choose any two categorical variables and create a statistical graphic showing their relationship. Explain what the graphic means in the context of the variables.



This histogram is based on proportionalities of the categorical variables, and it demonstrates that for countries whose head of state is ceremonial, they generally are free countries. If the head of state is executive they have limited to no freedom.

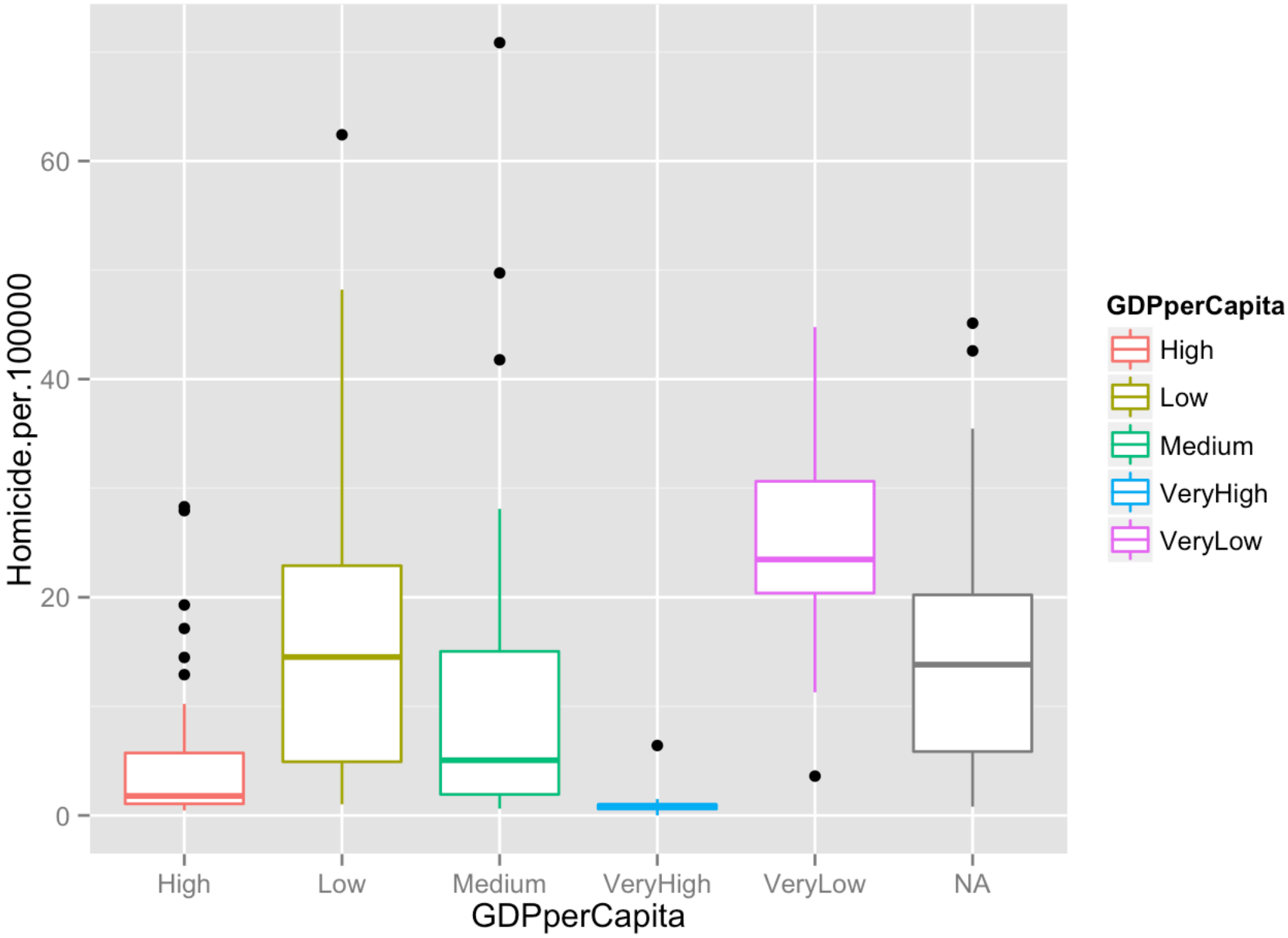
4. Choose any two numeric variables and create a statistical graphic showing their relationship. Explain what the graphic means in the context of the variables.

```
## geom_smooth: method="auto" and size of largest group is <1000, so using loess. Use
 'method = x' to change the smoothing method.
```



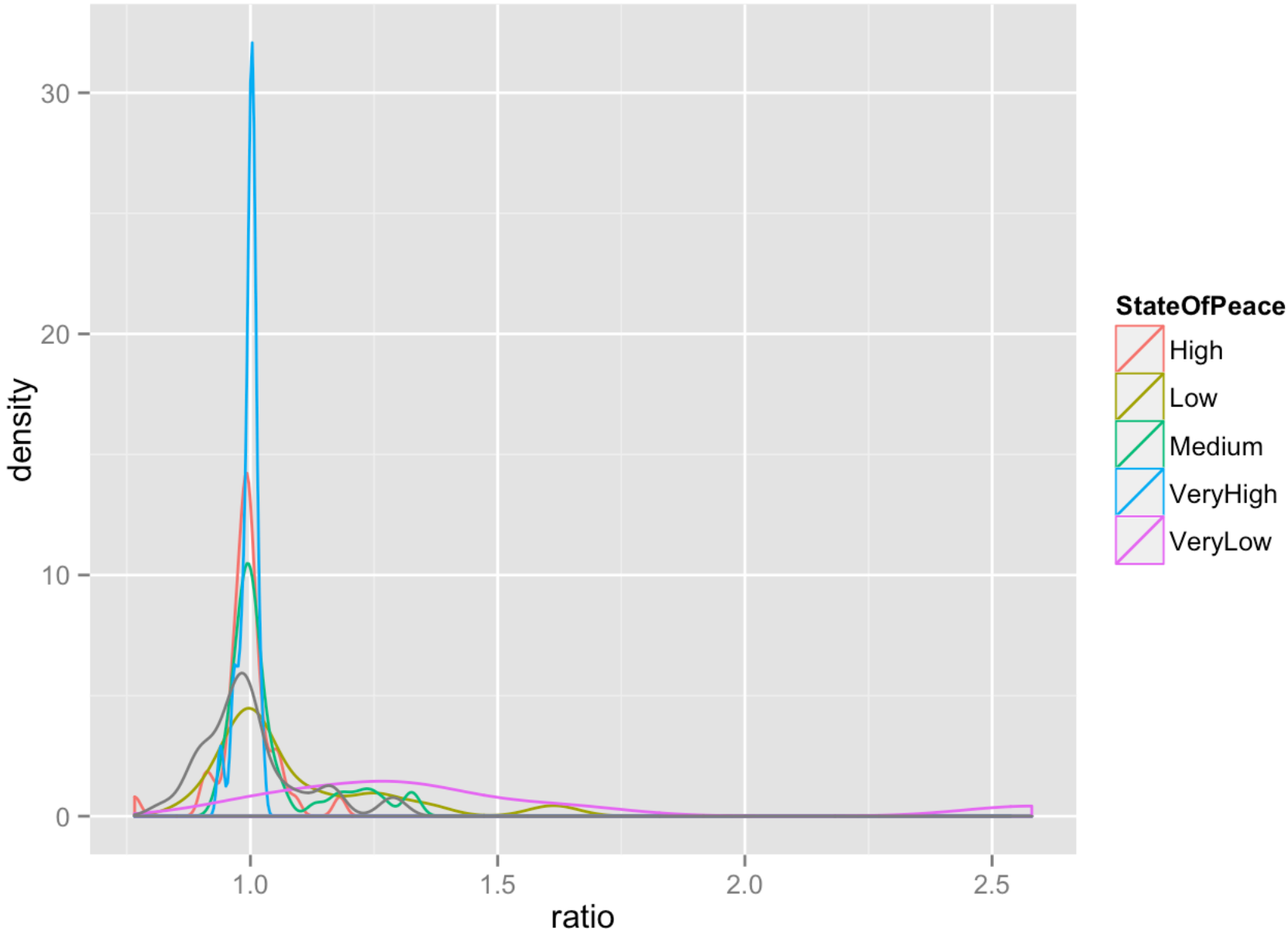
Given the slightly downward slope of the “best-fit” line curve, It is evident that the less literate a given person is, the less they perceive corruption in their country. Speculations and conspiracies can be heavily deduced from this.

5. Choose any categorical variable and any numeric variable. Create a statistical graphic showing their relationship. Explain what the graphic means in the context of the variables.



Given this box plot graph, it is evident that countries that are in bad economic conditions, and not making any GDP, will have more homicides. Interestingly, the low category of GDP has a pretty wide IQR, and thus depicts a lot of variance within the data. However, the “very low” category has little variance and has the highest median, which almost guarentees homicide is rampid in countries that are destitute. As expected, very successfull countries have virtually no homicides!

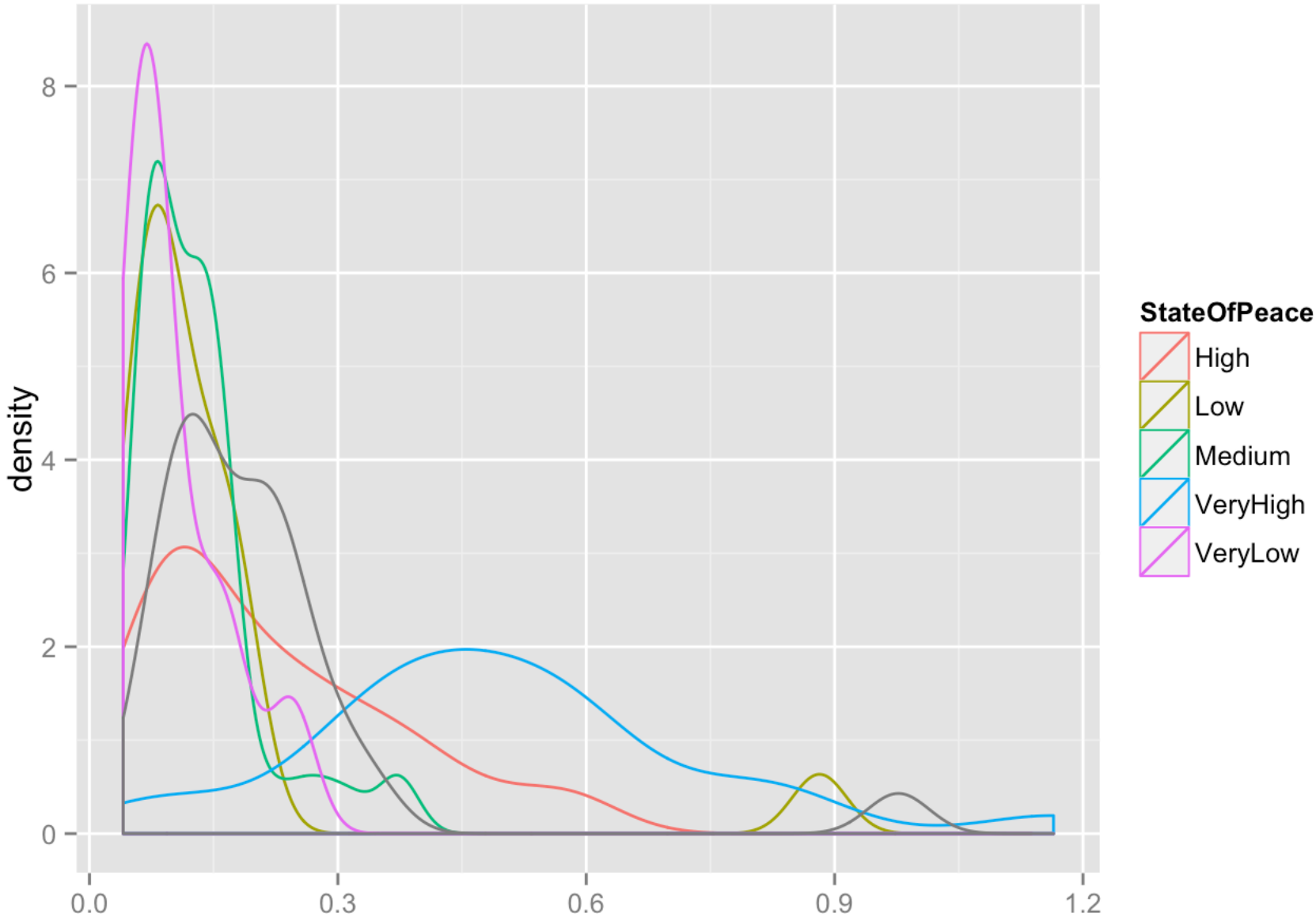
6. Describe the relationship between primary school completion ratio and State of Peace. Can you say that having more girls complete school will create a more peaceful society?



When the ratio is 1:1, and the country is peaceful, then most people complete school. The complete rate decreases proportionally to the state of peace (i.e as the peace goes from high to low so does completion rate)

To the left of the 1.0 ratio mark, is the area that indicates that more girls completed than boys. So looking to the left of 1.0, its evident that having more girls completing school will increase the peace, but not by orders of magnitude greater. Only slightly.

7. Create a single graphic which shows something about this data set. Describe what the graphic tells us. You may use as many or as few variables as you'd like, they can be categorical or numeric, you may use outside data to supplement this data, and you may combine variables.



Taking the ratio of average number of children born per woman and infant mortality rates per 1000 births, and plotting that against the categorical variable state of peace, we can see from the chart (which is speaking a generalization about the entire data set), given the political wellness of a country, increases the overall chances at a well functioning and developing society (evolutionarily speaking).