DataVizA Tutorial: Plotting one variable: Solutions

Department of Econometrics and Business Statistics, Monash University

Tutorial 3

Concepts

1. In the RGB color model what color is represented by #FF0000?

Since RGB stands for Red Green Blue the color FF0000 is the maximum possible amount of red (FF=255) and no green and blue. That is FF0000 is pure red.

2. In the RGB color model what would be the hex code for pure blue?

This would be #0000FF.

3. Describe the role of bandwidth in kernel density estimation.

A kernel density estimated at a value x gives highest weight to values near x. The bandwidth essentially determines how close is meant by 'near'. Large values lead to smooth (in the limit flat) density estimates, small values lead to very bumpy estimates.

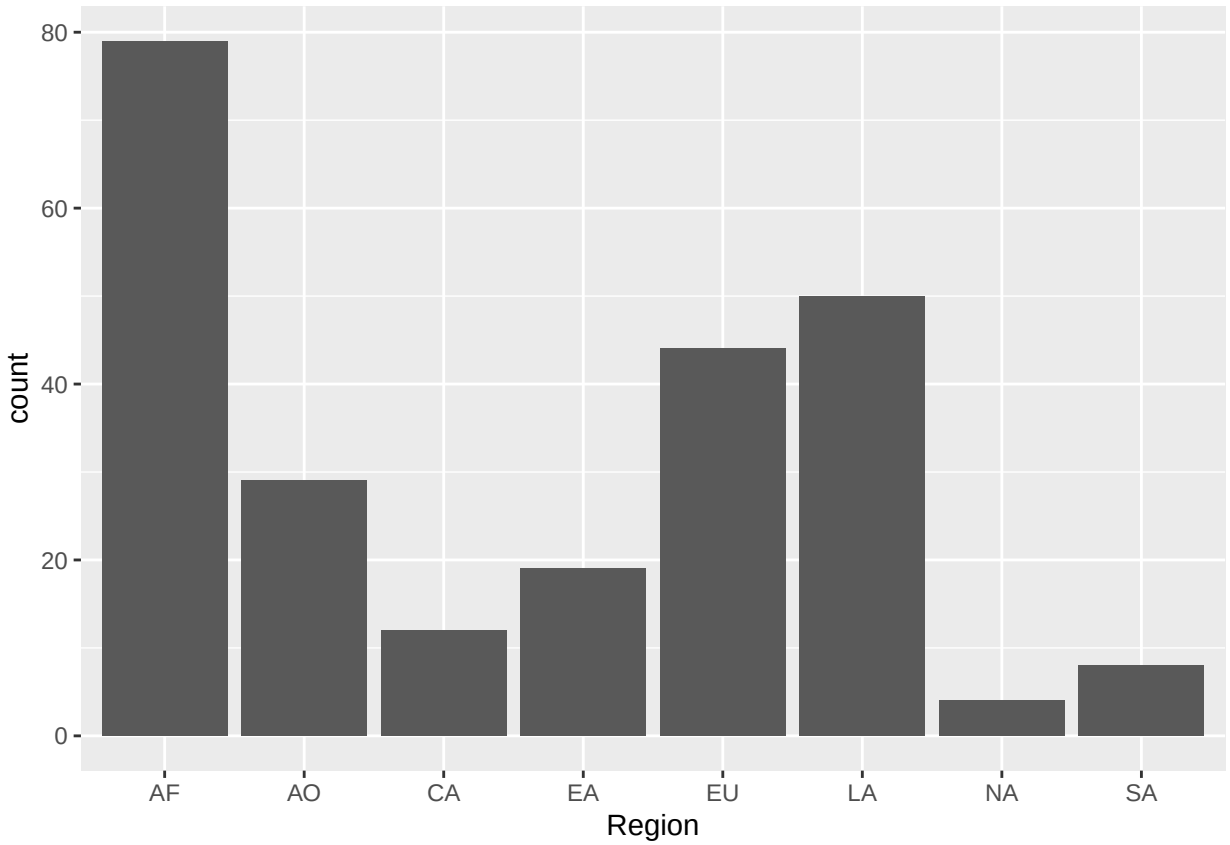
4. How are outliers defined in a boxplot?

An upper and lower fence are set to the the third quartile plus 1.5 times the interquartile range and the first quartile minus 1.5 times the interquartile range respectively. If the minimum value is greater than the first quartile minus 1.5 of the interquartile range, then the a line is drawn at the mimimum value rather than the lower fence. If the maximum value is less than the third quartile plus 1.5 of the interquartile range, then the upper fence is drawn at the maximum value rather than the upper fence. If there are points either above the upper fence, or below the lower fence these are considered to be outliers and drawn as a point.

Data Analysis

5. Using the Swiss Export Data, plot a bar chart of the region variable.

```
library(tidyverse)
SwissExp<-readRDS('SwissExport.rds')
ggplot(SwissExp,aes(x=Region))+geom_bar()
```

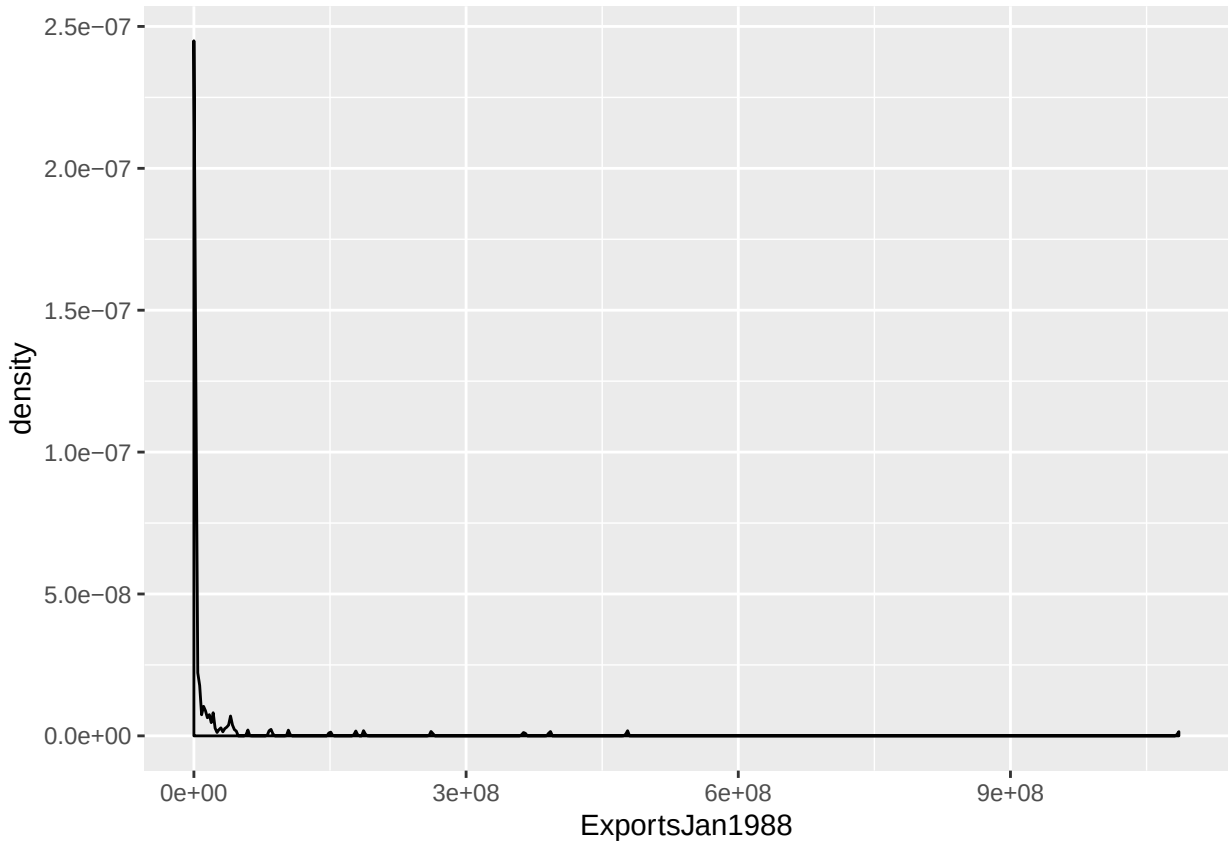


6. What does this plot tell you?

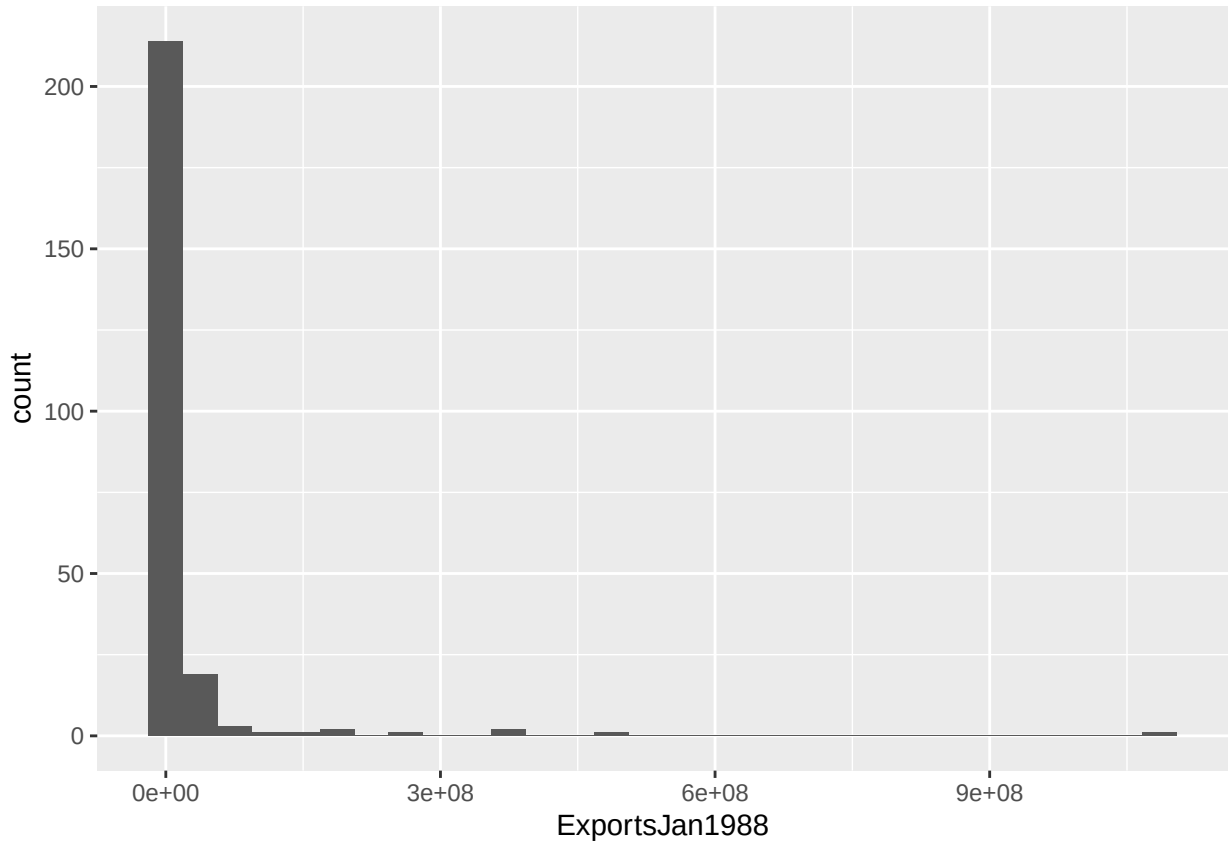
The plot merely states that most countries are in Africa and very few are in North America. This plot says nothing about Swiss Exports. In many cases this plot would not be worth including in an analysis.

7. Construct a density plot, rug plot and boxplot for Exports in January 1988.

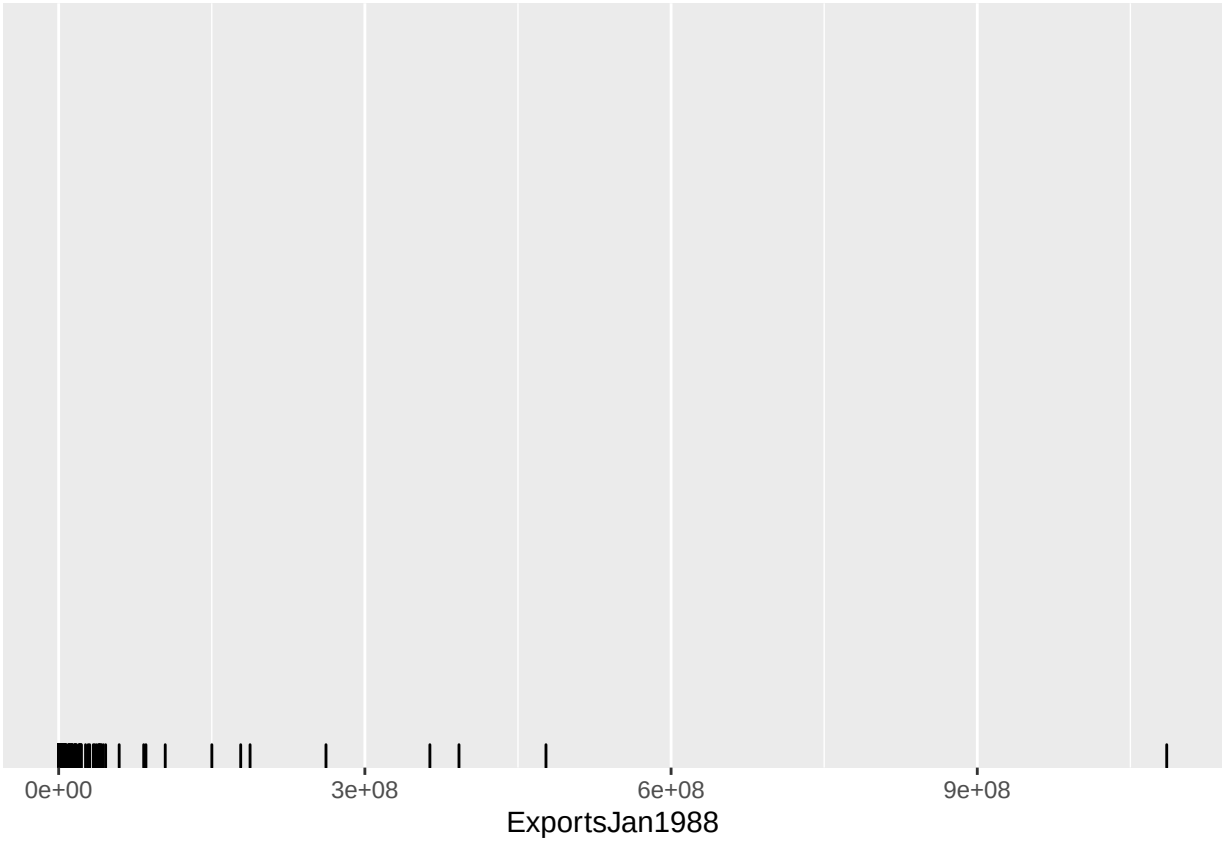
```
library(tidyverse)
SwissExp<-readRDS('SwissExport.rds')
ggplot(SwissExp,aes(x=ExportsJan1988))+geom_density()
```



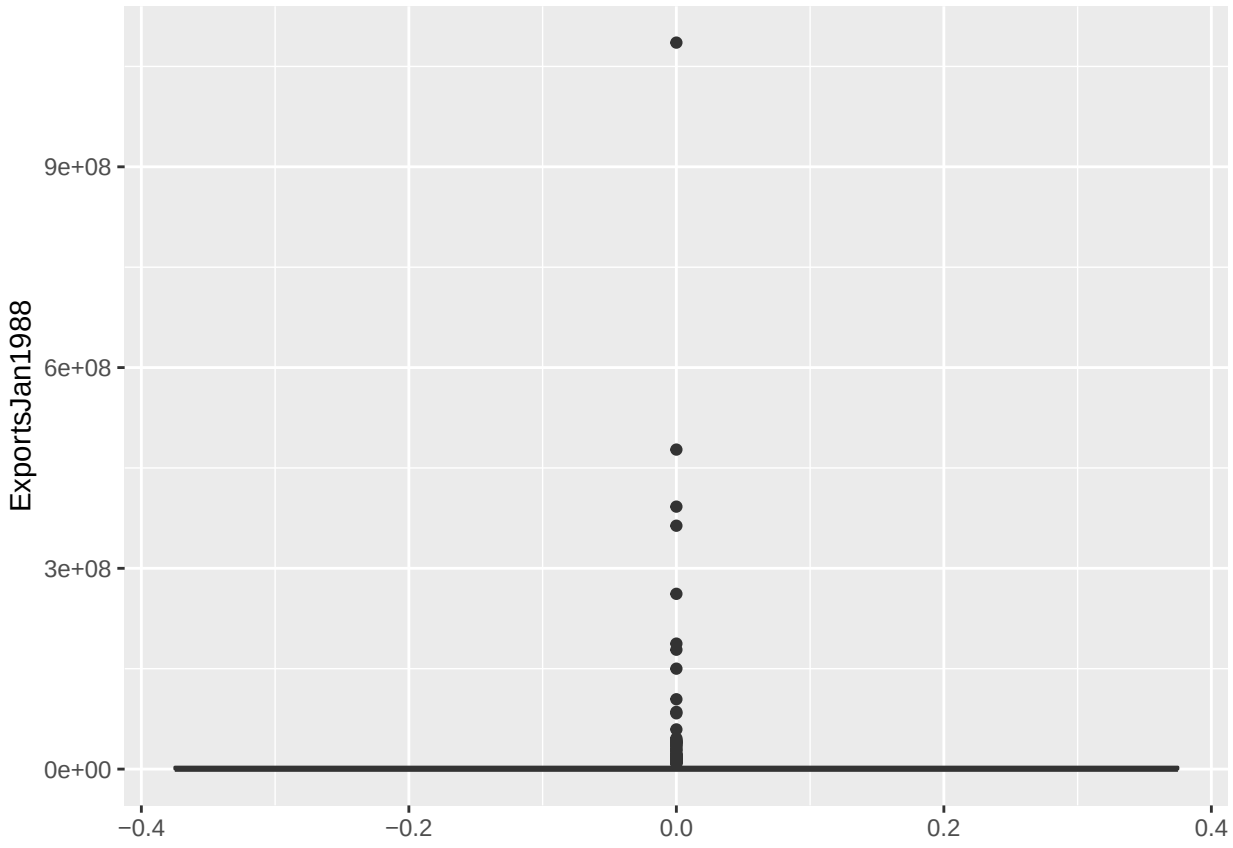

```
ggplot(SwissExp,aes(x=ExportsJan1988))+geom_histogram()
```




```
ggplot(SwissExp,aes(x=ExportsJan1988))+geom_rug()
```




```
ggplot(SwissExp,aes(y=ExportsJan1988))+geom_boxplot()
```



8. What do we see in these plots?

All plots show that the Switzerland had a very small level of exports to most countries. In the box plots, the quartiles and medians are all close to 0. There is one very large export destination which is Germany.