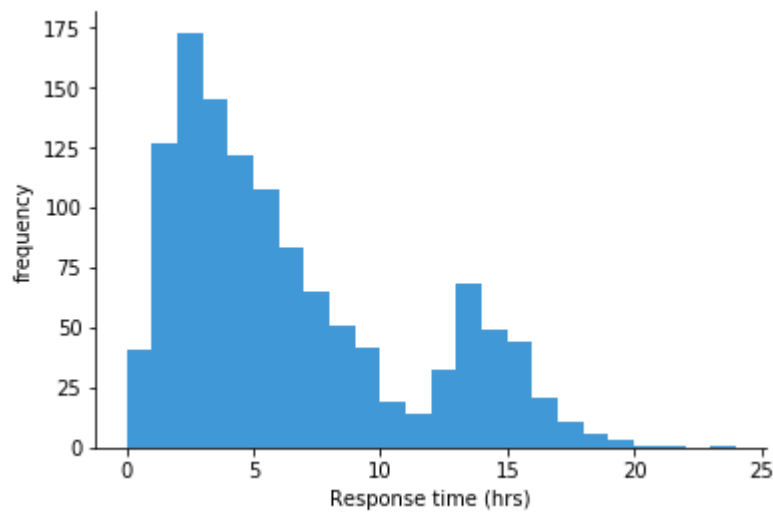# Historgram

- Plots values as a series of bars
- Each Bar covers an interval of values --> Bins, Class
- Bar height indicate the frequency of the data that falls in that Bin

- The mean and standard deviation, often makes us miss the fact there were these two peaks that can effect the conclusions and statistics inference we get

Best practices for histograms plots

- Using Zero valued baseline
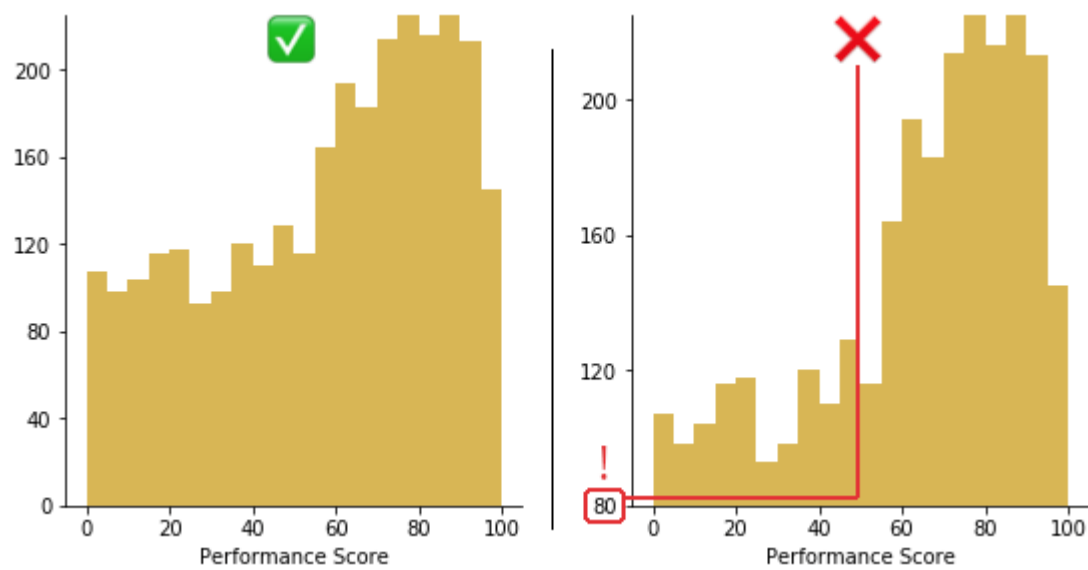
- The right number of bins, It also have a relation with the size of the bins
  - too many bins will make the distribution look rough plus a lot of noise
  - few bins the histogram will lack details and makes us miss patterns and peaks
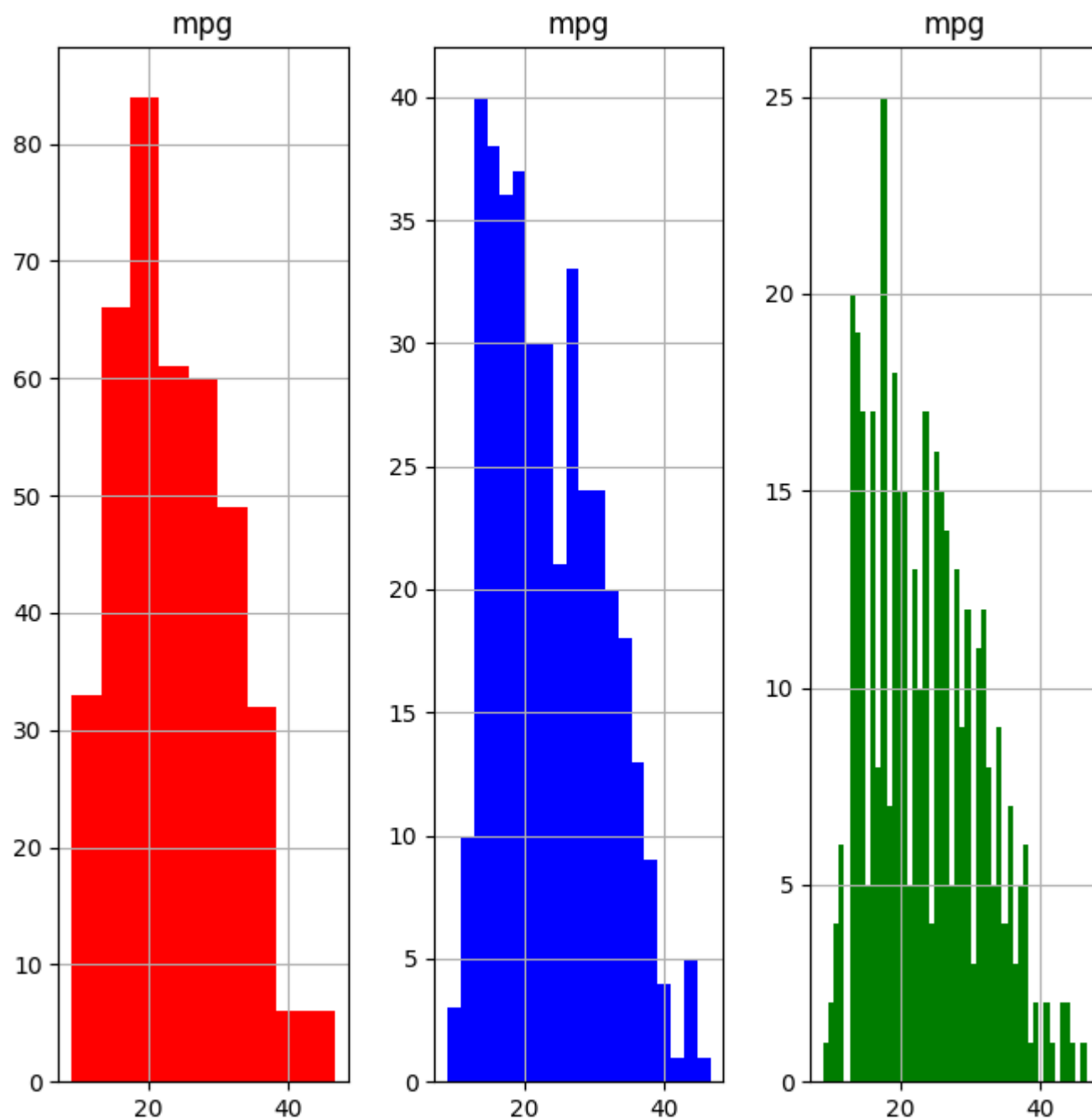
- The right bin boundaries --> 0 5 10 15 20 is much more interpretable
  than 0 3.7 7.4 11.1
  caution : If you got a fractional interval `[0,2,5,5,7.5]` and your observations are integers 0->2.5 can fit 0-1-2 but 2.5->5 will
  only fit 3-4, This can make your histogram looks bumpy

-
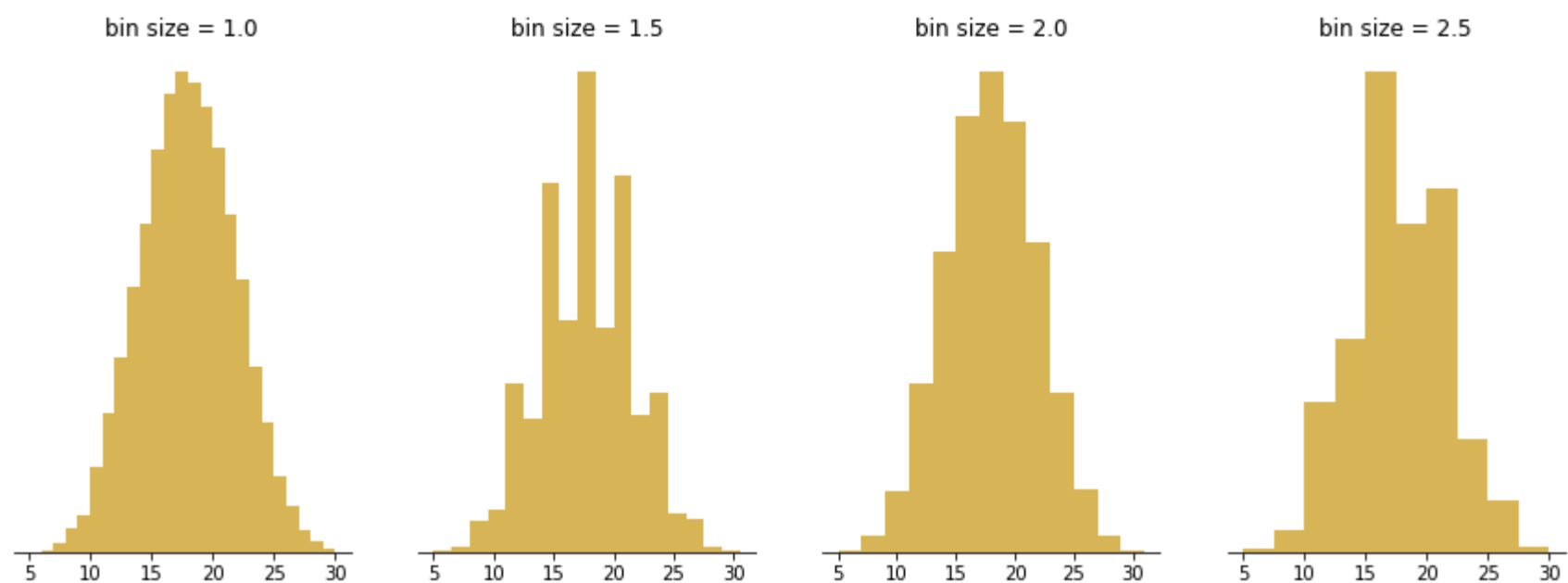
Ref : [histogram guide Article](#)
tags: #stats   #plots   #machine-learning