Exploratory Data Science

Course 4

Brian Pollack

11/05/19

Univariate Data

Univariate: single variable

Typical Situations:
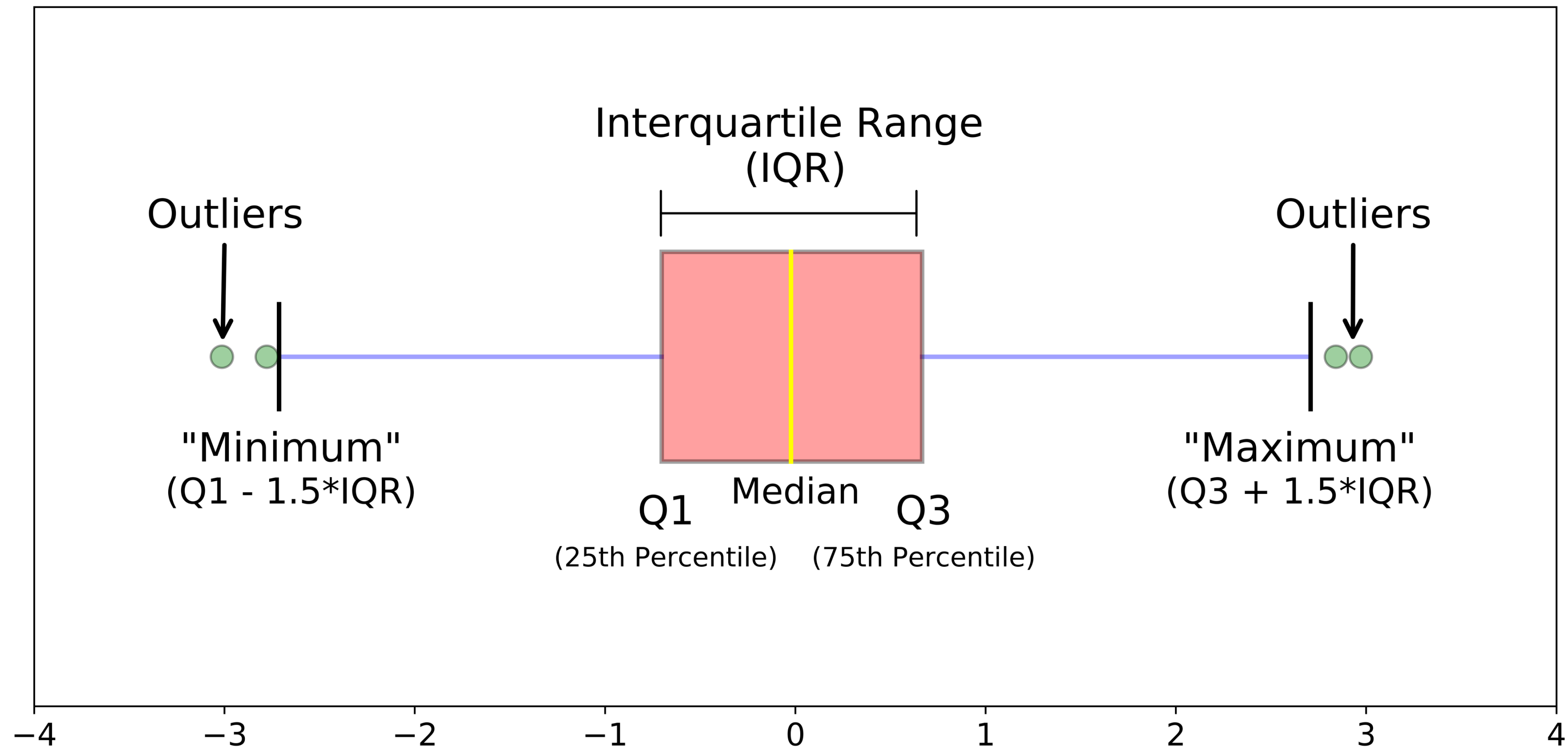
-Time

-Samples from underlying PDF

```
Index   Value

0       33.6
1       26.6
2       23.3
3       28.1
4       43.1
5       25.6
6       31.0
7       35.3
8       30.5
9        0.0
10      37.6
11      38.0
12      27.1
13      30.1
14      25.8
15      30.0
16      45.8
17      29.6
18      43.3
19      34.6
Name: BMI, dtype: float64
```
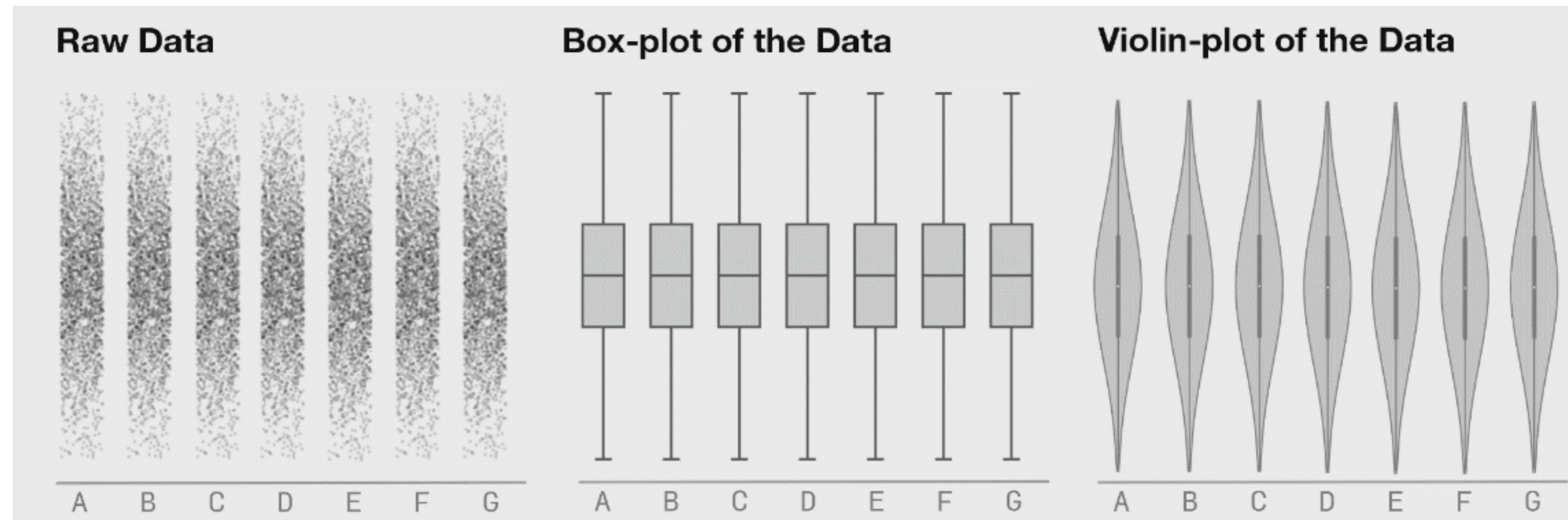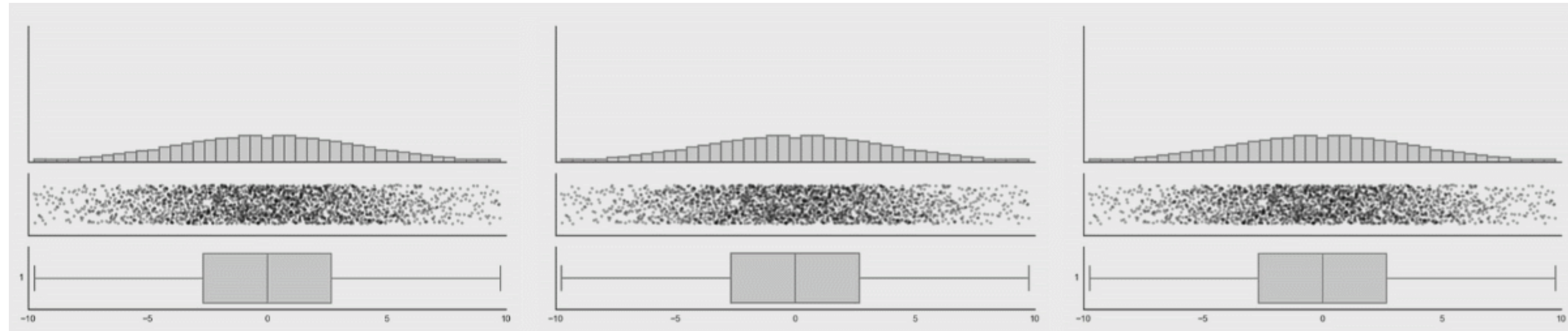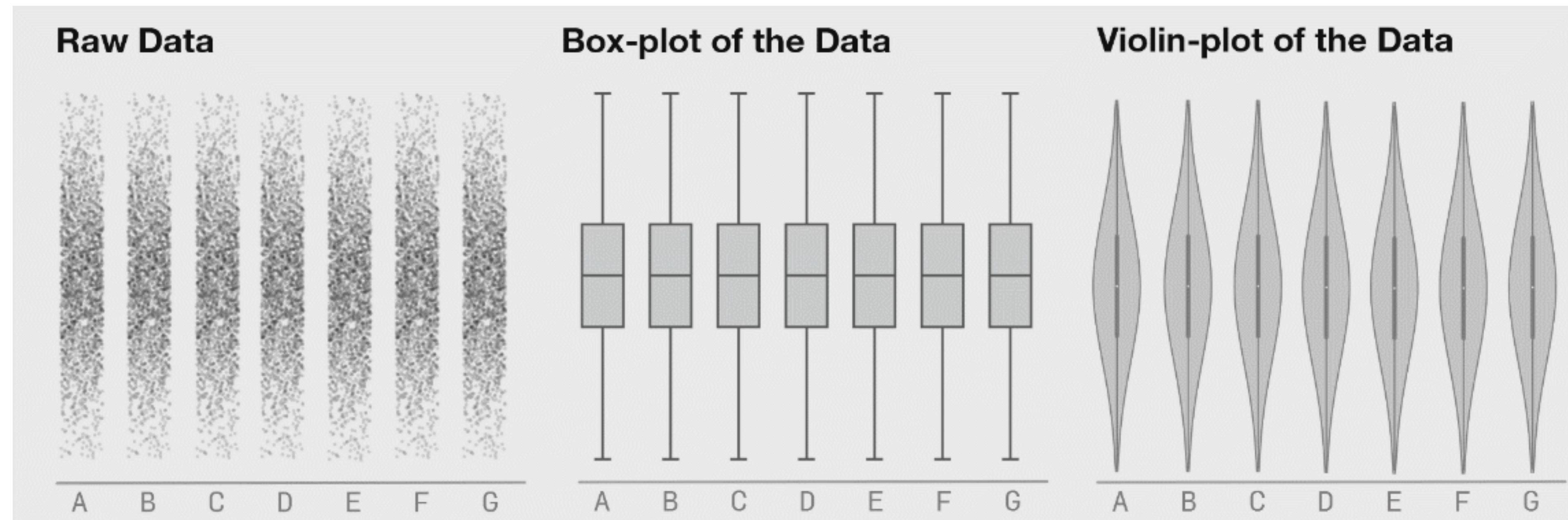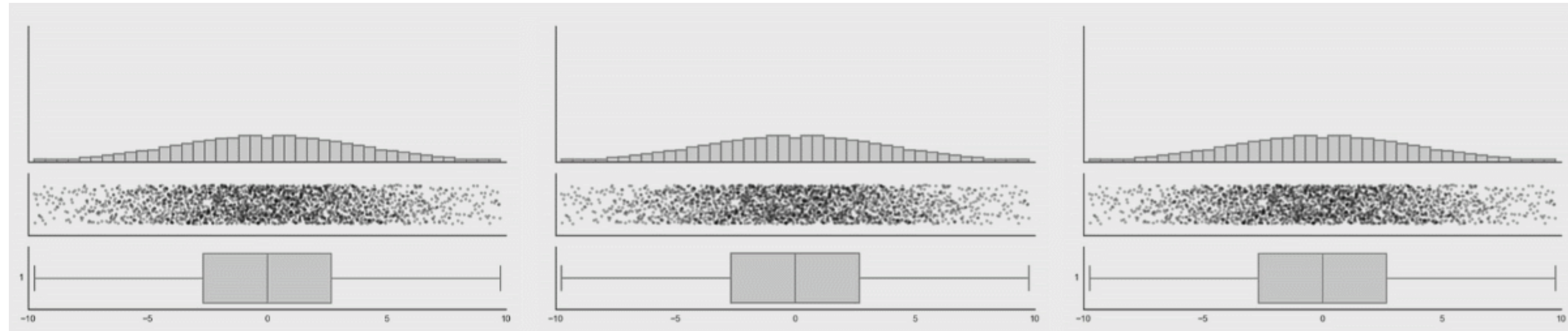
Boring Boxplot

https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51

Problem Plots

Raw Data | Box-plot of the Data | Violin-plot of the Data

https://www.autodeskresearch.com/publications/samestats

Problem Plots

Raw Data — Box-plot of the Data — Violin-plot of the Data

https://www.autodeskresearch.com/publications/samestats

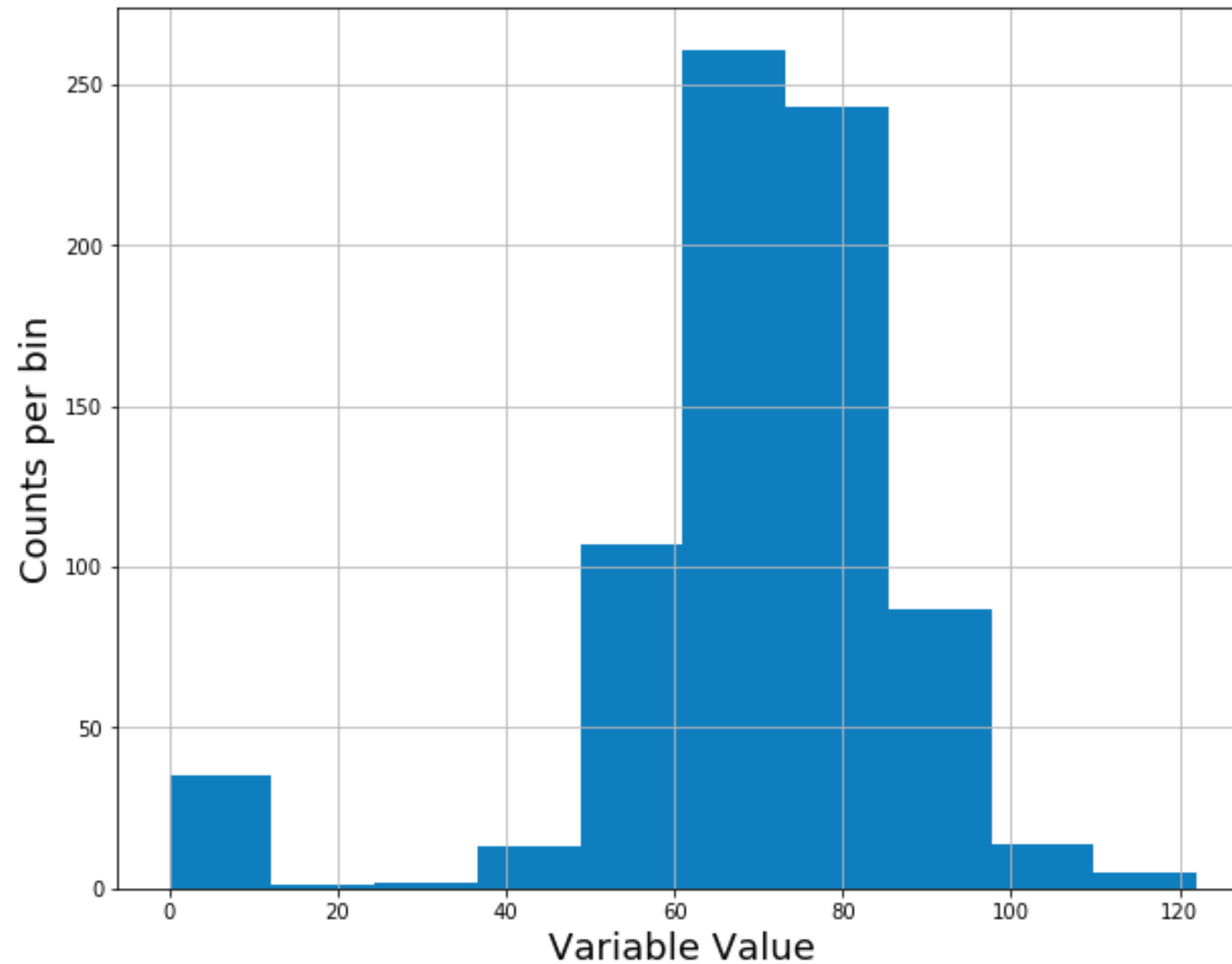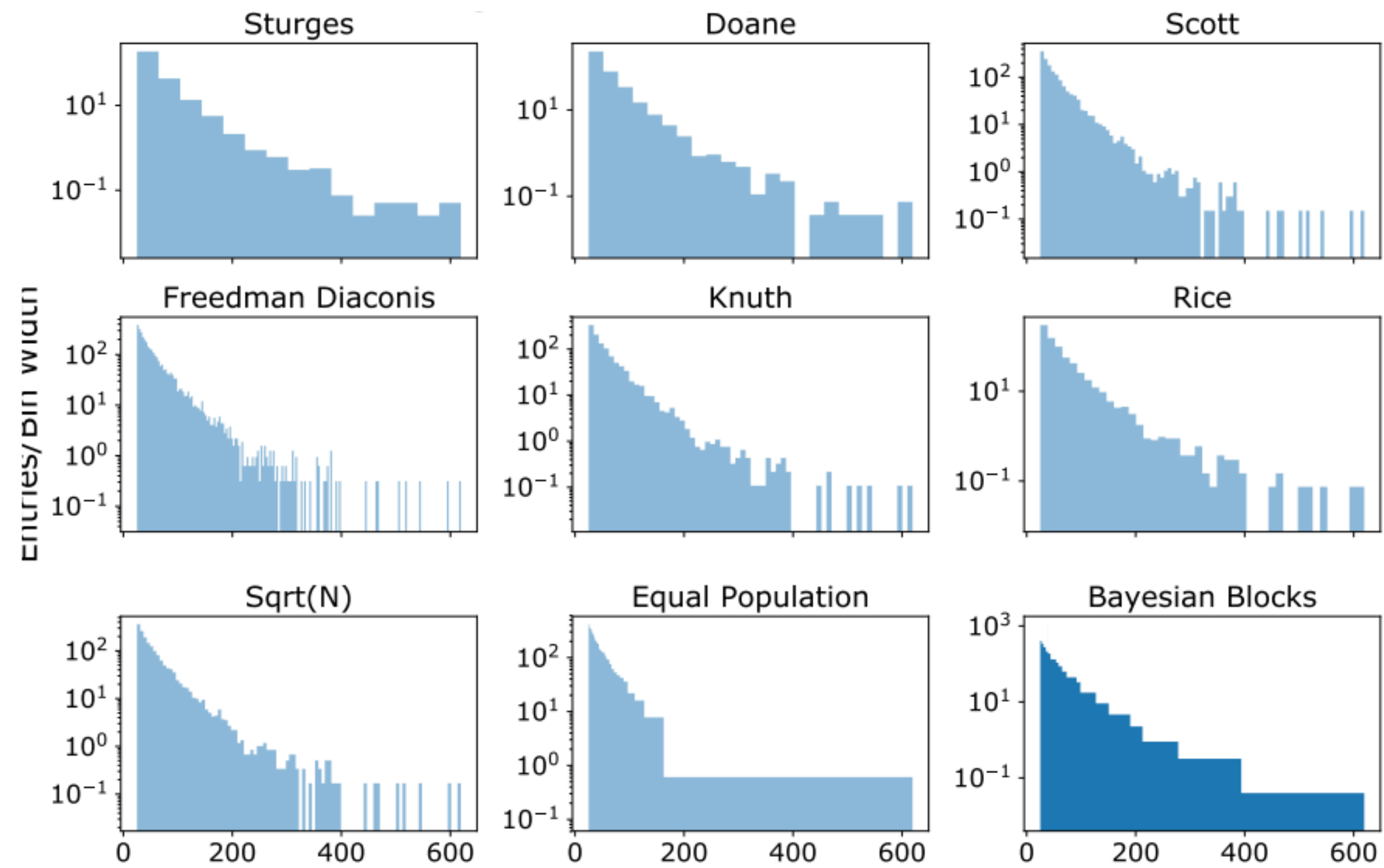# Problem Plots

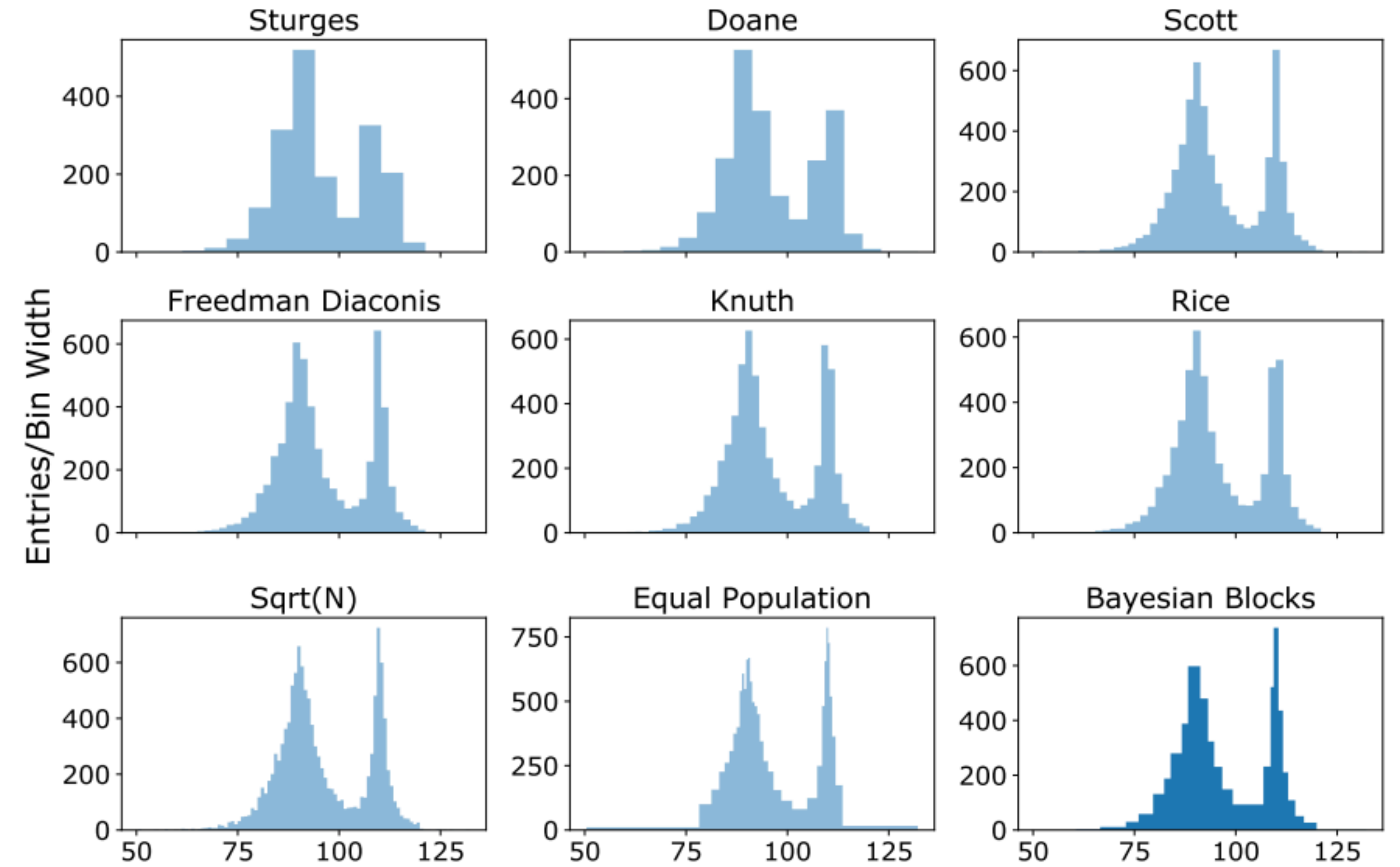https://www.autodeskresearch.com/publications/samestats

Our Hero, the Histogram

Binning

Falling Spectrum

Double Peaks

https://github.com/scikit-hep/scikit-stats