Introduction to bootstrapping confidence intervals of proportions

Munchkin

http://puppywar.com/stats.php?id=41727

vs.

Peno

http://www.kittenwar.com/kittens/44101/

Goal: find proportion of Americans that thinks puppy is cuter

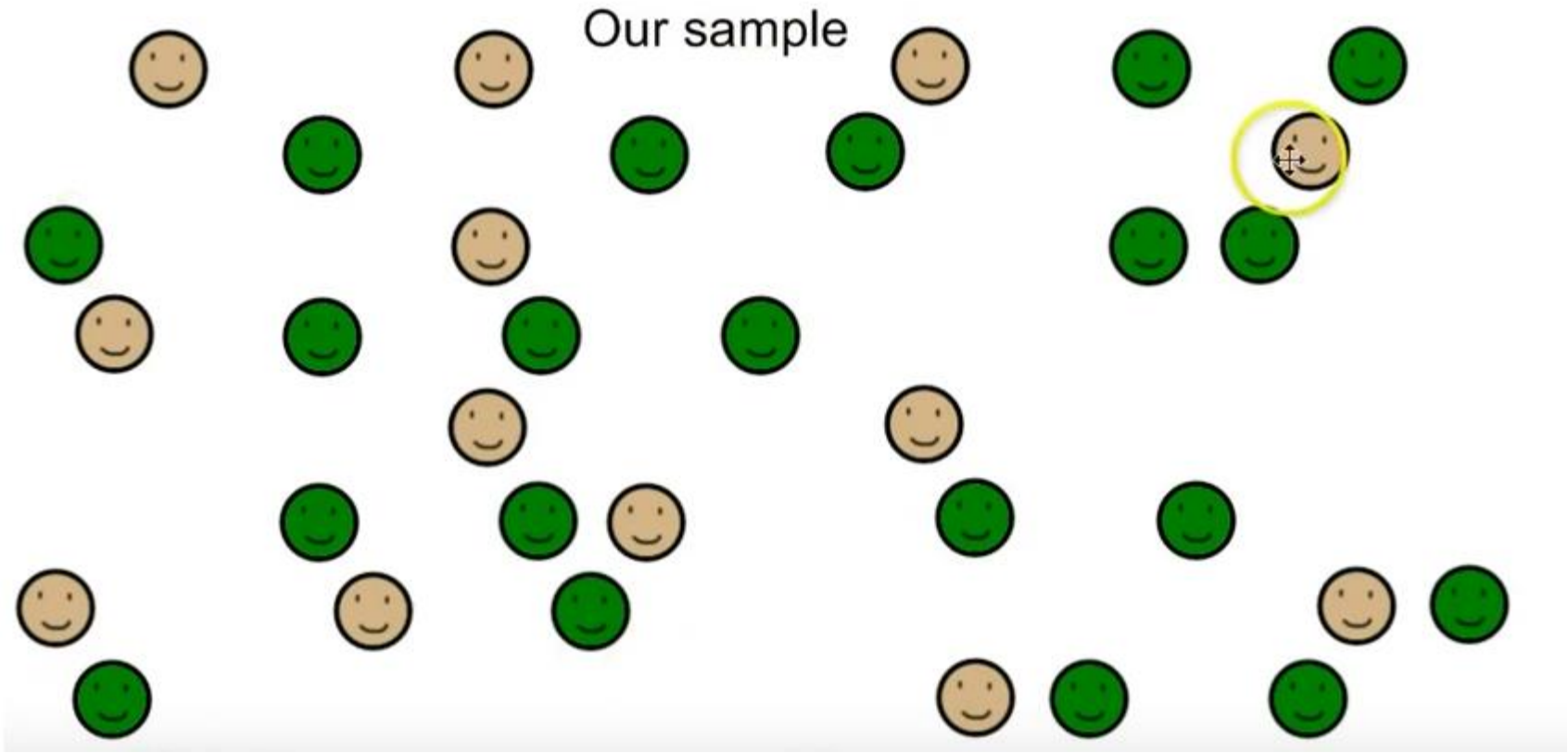
Munchkin

http://puppywar.com/stats.php?id=41727

vs.

Peno

http://www.kittenwar.com/kittens/44101/


Our sample

24

16

so…

24/40 prefer Munchkin the puppy

60% prefer Munchkin the puppy

How confident are we in this guess of 60%?

Instead of 60%, let's say...

We're POSITIVE its between 0%-100%

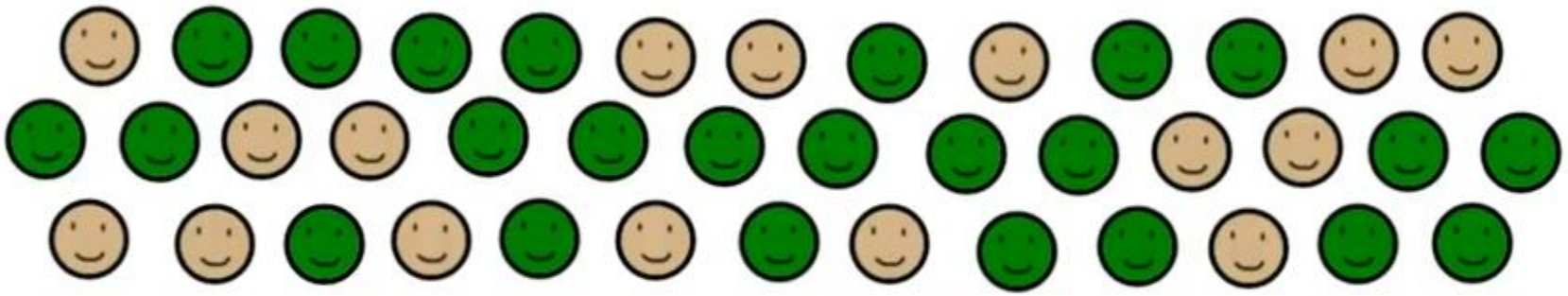
Instead of 60%, let's say...

We're POSITIVE its between 0%-100%

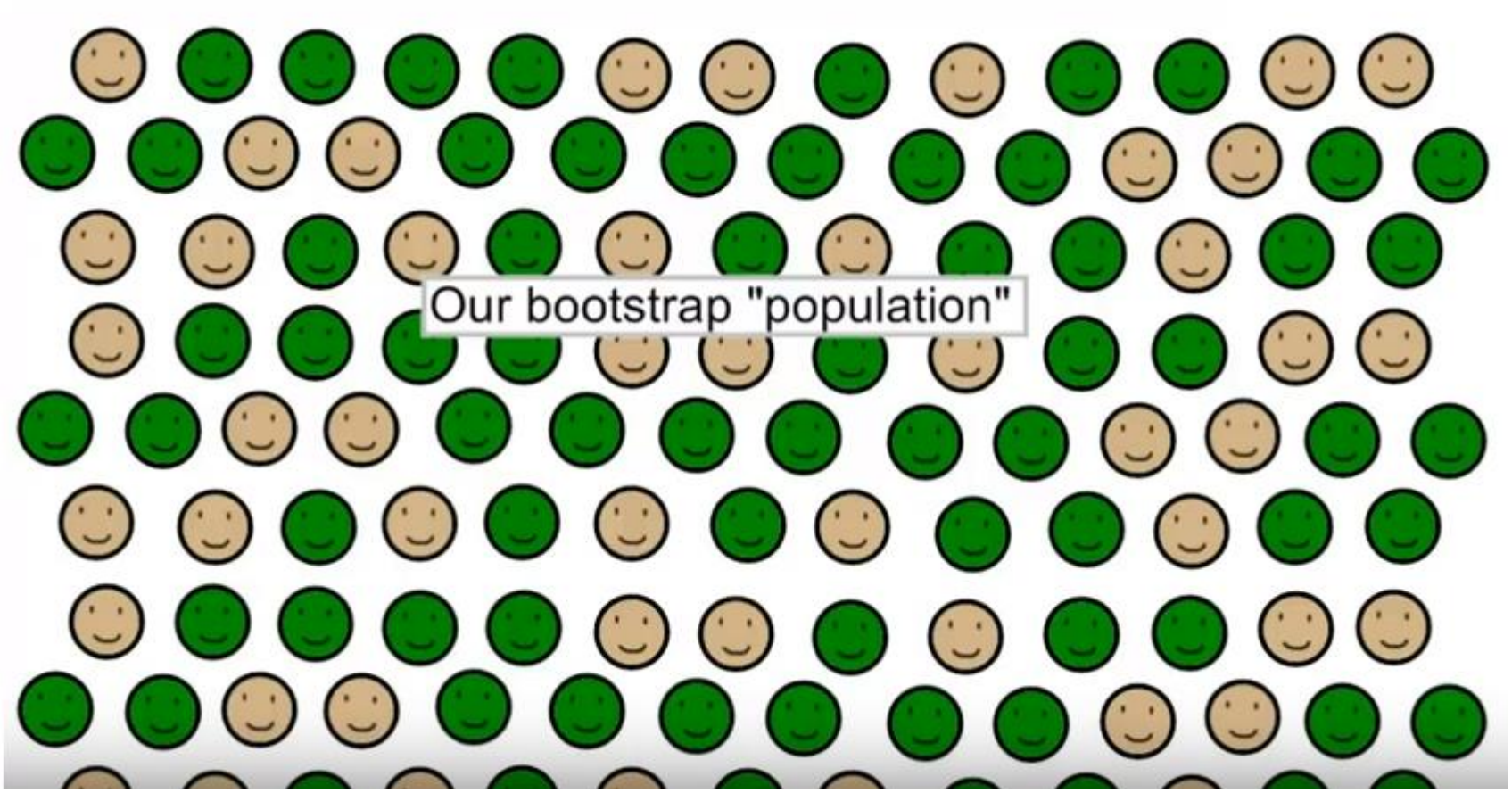
60

We're very confident its between 40%-80%

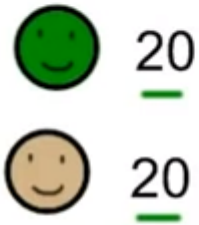
We're 95% confident its between ___ - ___



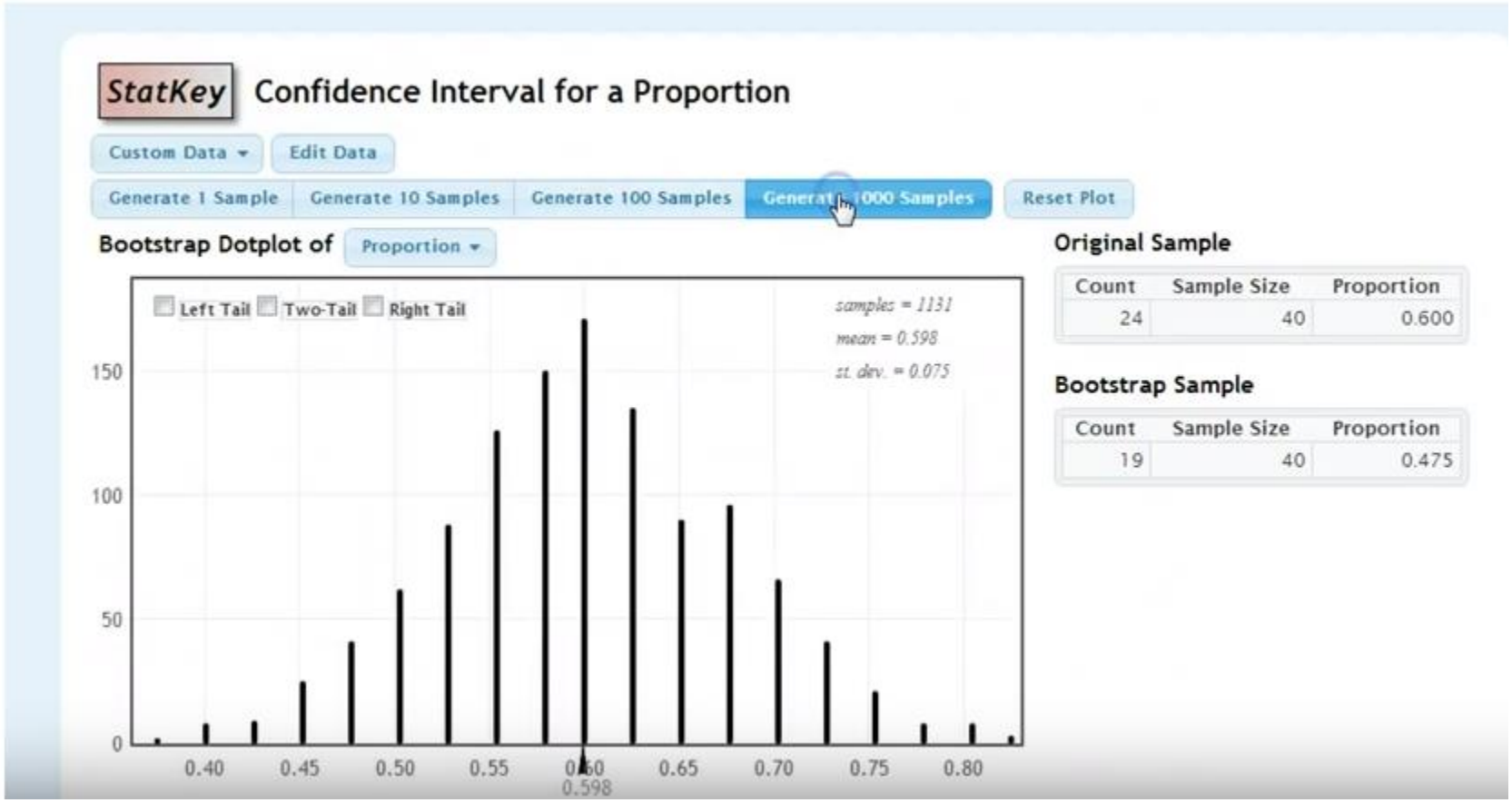
Our sample


Our bootstrap "population"

BOOTSTRAPPING



Repeat this process thousands of times...


StatKey
Confidence Interval for a Proportion
Custom Data
Edit Data
Generate 1 Sample
Generate 10 Samples
Generate 100 Samples
Generate 1000 Samples
Reset Plot
Bootstrap Dotplot of
Proportion
Left Tail
Two-Tail
Right Tail
samples = 1131
mean = 0.598
st. dev. = 0.075
150
100
50
0
0.40
0.45
0.50
0.55
0.60
0.65
0.70
0.75
0.80
0.598
Original Sample
Count
Sample Size
Proportion
24
40
0.600
Bootstrap Sample
Count
Sample Size
Proportion
19
40
0.475

Instead of 60%, let's say...

We're POSITIVE its between 0%-100%

We're very confident its between 40%-80%

We're 95% confident its between 45% - 75%

In context, we would say:

We're 95% confident that the proportion of Americans that think Munchkins is cuter than Peno is between 45% and 75%.