Visualizing Uncertainty
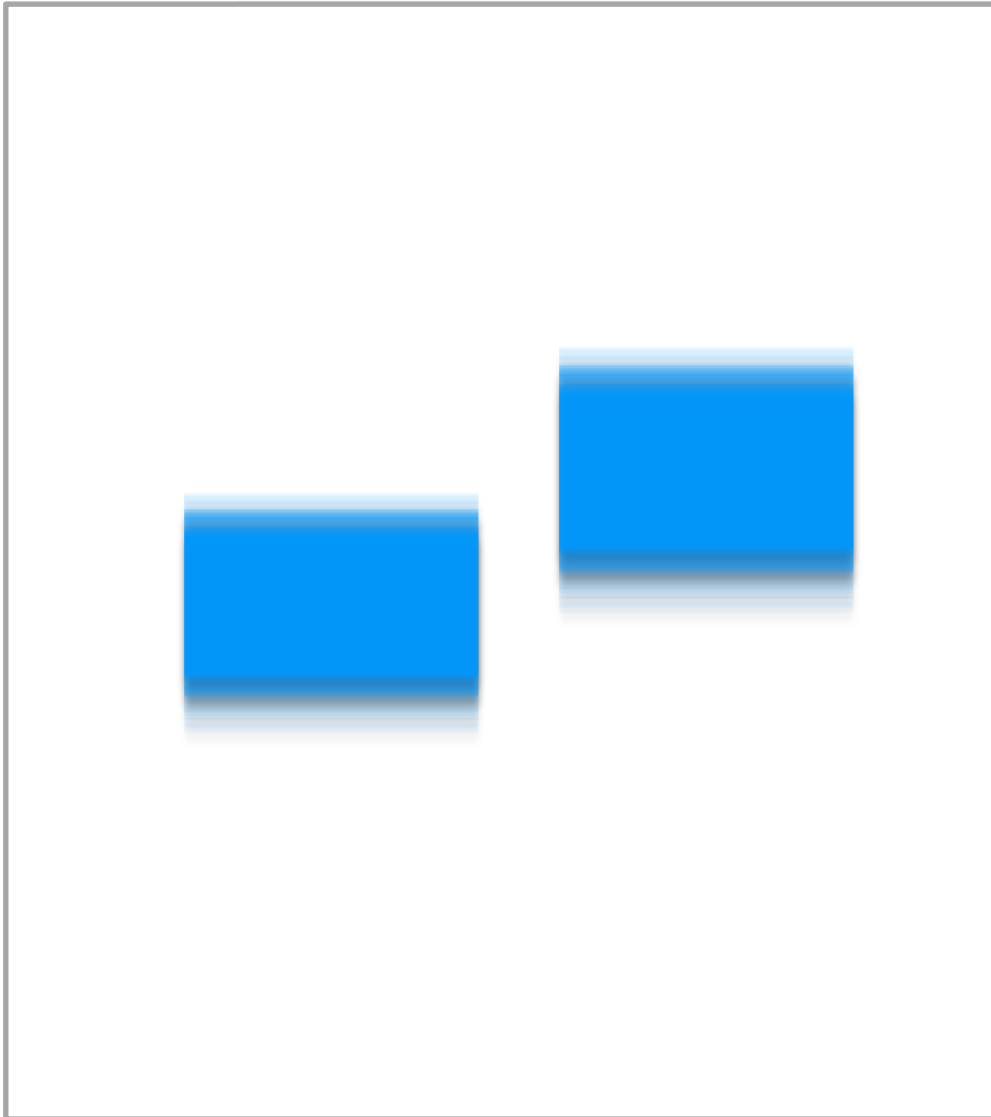
Paul Resnick
Michael D. Cohen Collegiate Professor
Associate Dean for Research and Faculty Affairs
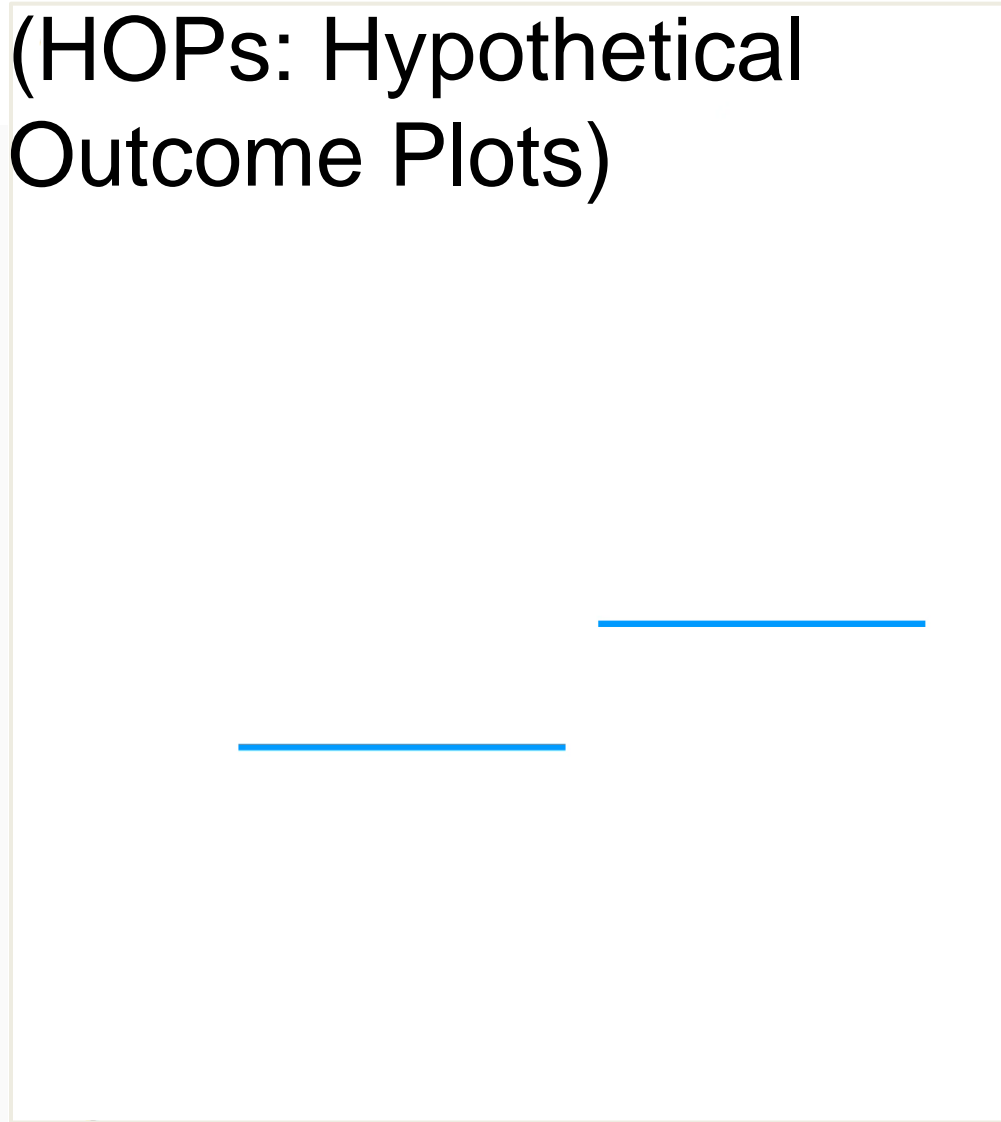October 5, 2017

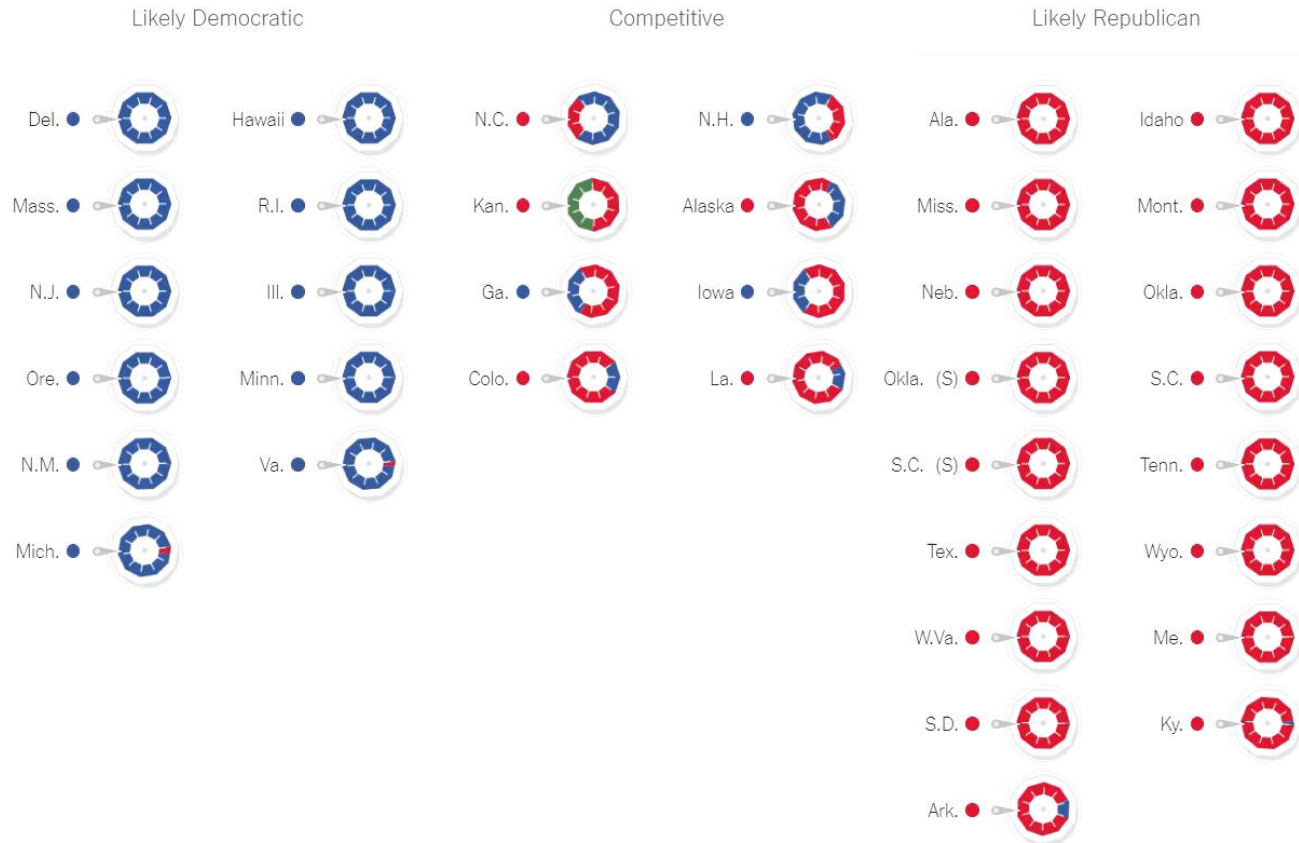# Two approaches

Static Aggregate Distribution plots (SAD)

Sample-based (experience-able) plots (HOPs: Hypothetical Outcome Plots)

http://www.nytimes.com/newsgraphics/2014/senate-model/
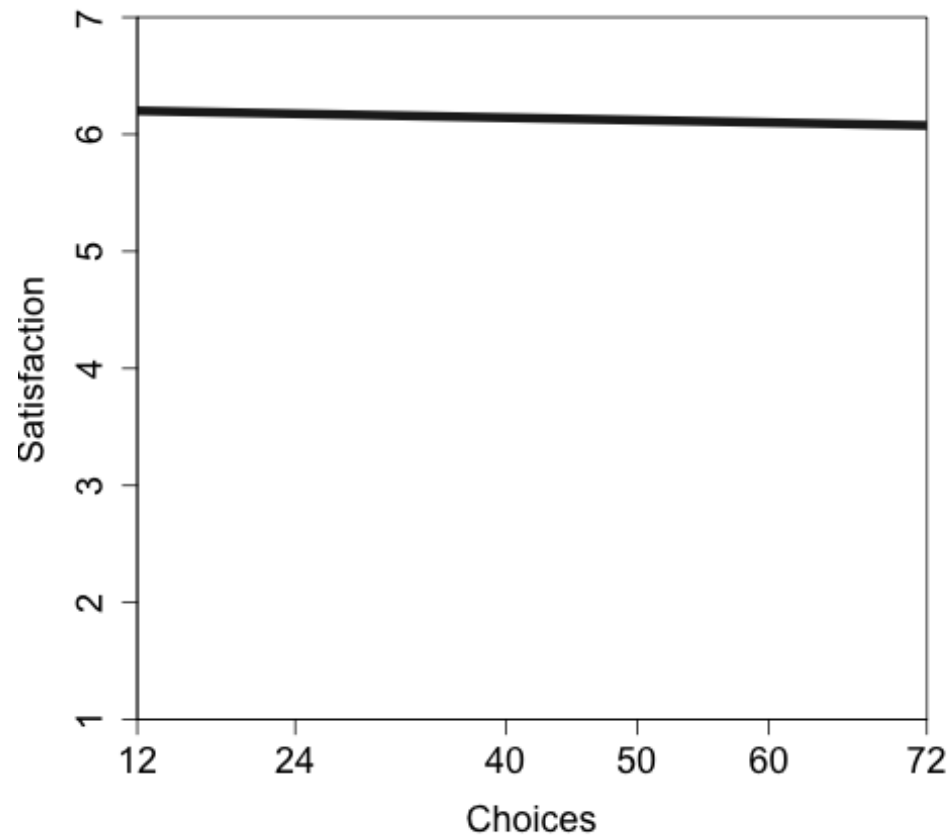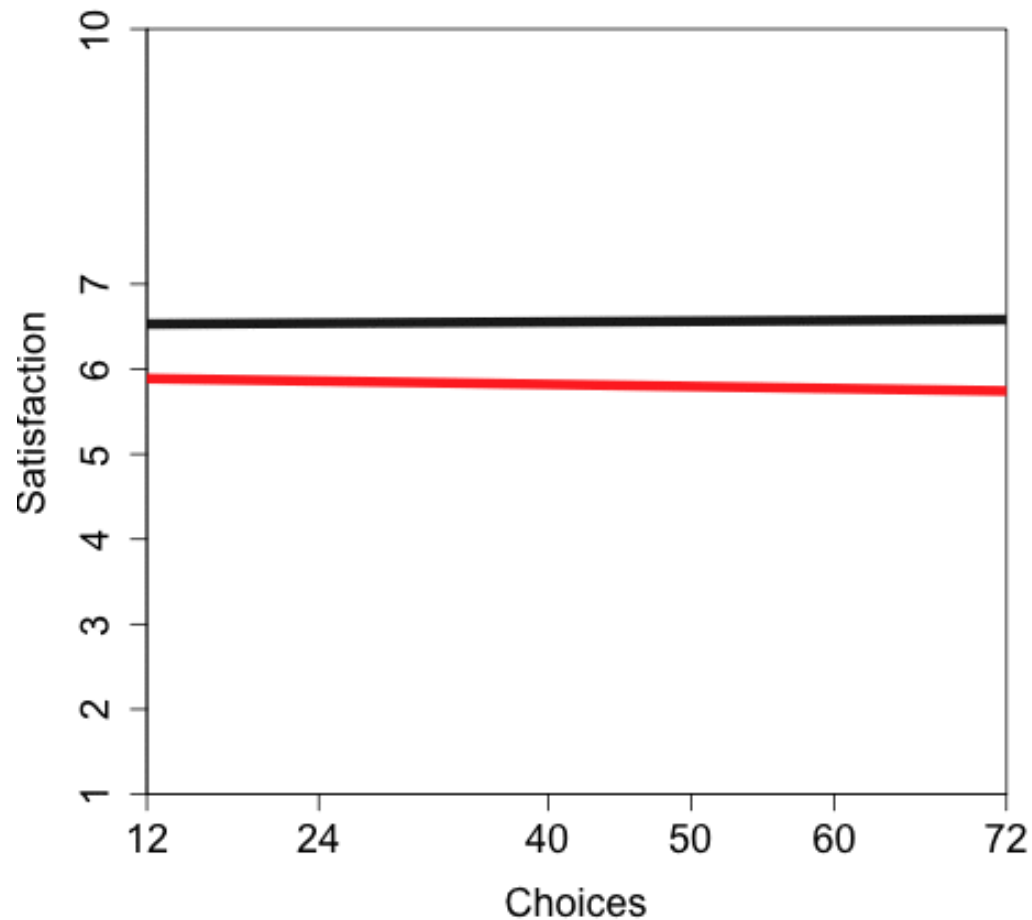
Visualizing Bayesian Posteriors

```
##              Mean    StdDev lower 0.95 upper 0.95
## a        6.228085 0.092183   6.059158   6.418798
## b_lin  -0.002147 0.001902  -0.005847   0.001623
```

For More Complex Model
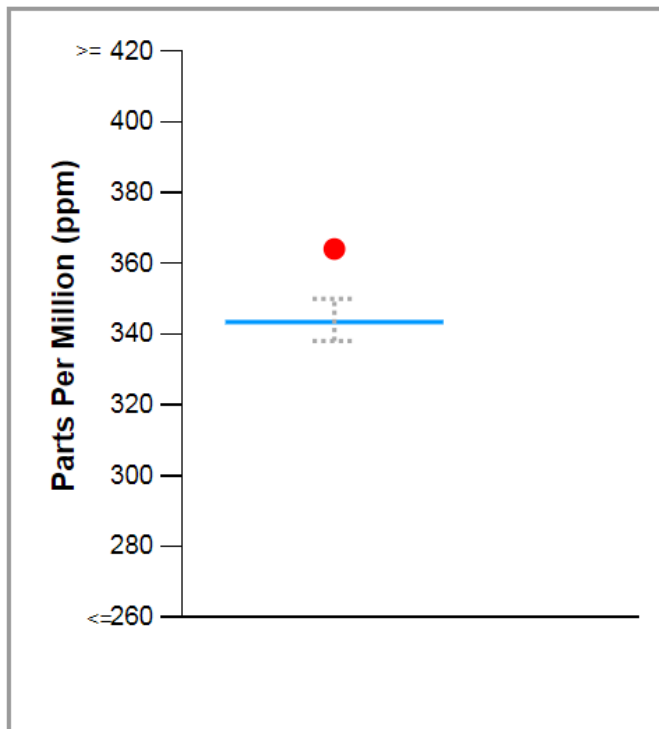
```
##                       Mean    StdDev lower 0.95 upper 0.95
## a               6.613922 0.361735   5.876673   7.278685
## b_lin           0.001519 0.006887  -0.011496   0.014941
## b_max          -0.007651 0.006335  -0.020324   0.004377
## b interaction  -0.000044 0.000122  -0.000289   0.000189
```

How well can you guess?

"Interpret Data Charts" - Number 2 out of 9

- The chart summarizes data that scientists have gathered by measuring the concentration of a chemical solute (measured in parts per million) in many vials of sea water.

Q1: What is the value of solute in parts per million (ppm) at the red dot?

Q2: What is the average measurement of solute in parts per million (ppm)?

Q3: How often are the measurements above the value of the red dot? Answer in terms of the number of times out of 100.

times out of 100
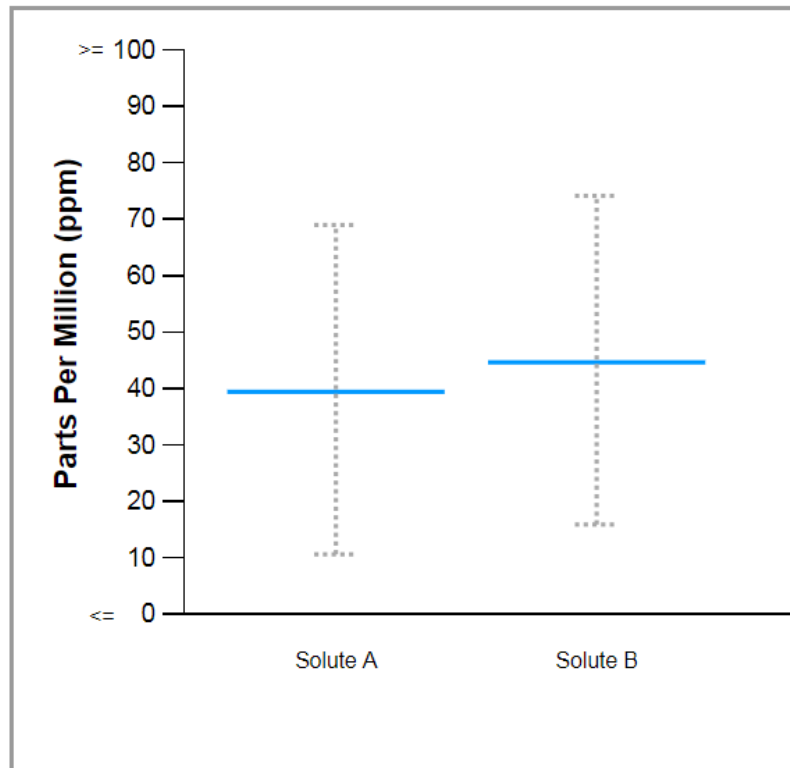
Q4: How often will the measurements lie **between 340 and 370**? Answer in terms of the number of times out of 100.

times out of 100

The blue line shows the average amount of solute in all the seawater vials. The dashed lines show an error bar, a range above and below the average. 95% of the collected vials fall in the range defined by the dashed lines.

How well can you guess?

The chart summarizes two data sets that scientists have gathered by measuring the concentration of two chemical solutes (A and B, each measured in parts per million) in many vials of sea water.

Q: How often is the measurement of solute B larger than the measurement of solute A? Answer in terms of the number of times out of 100.

The blue lines show the average amounts of solute A and B in all the seawater samples. The dashed lines show error bars, ranges above and below the averages. 95% of the collected samples fall in the ranges defined by the dashed lines.

How well can you guess?

The chart summarizes two data sets that scientists have gathered by measuring the concentration of two chemical solutes (A and B, each measured in parts per million) in many vials of sea water.
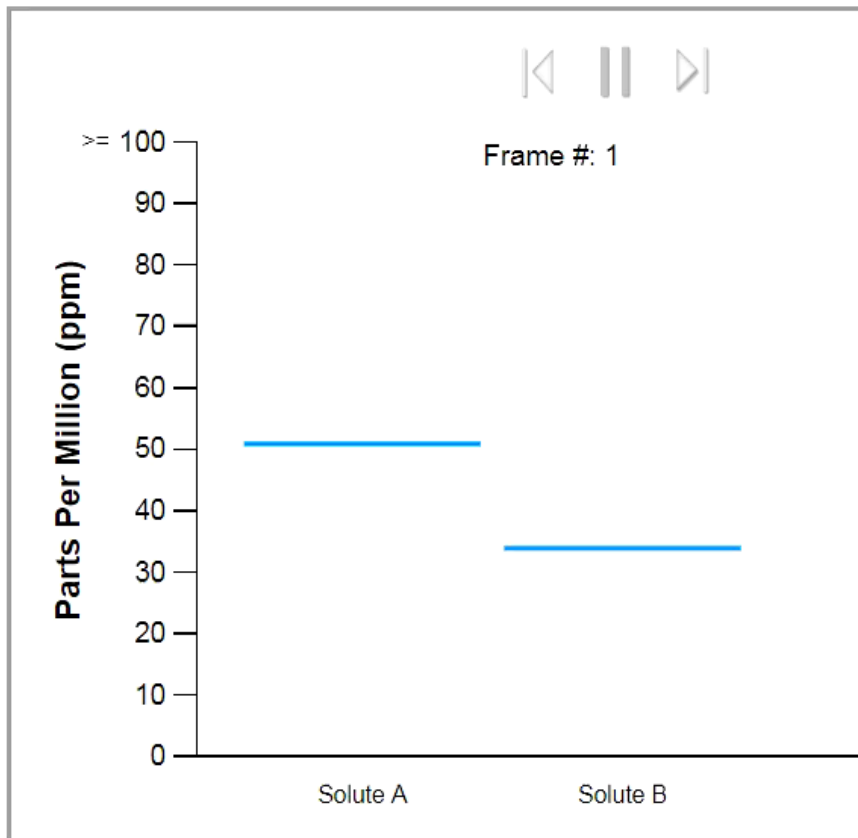

Q: How often is the measurement of solute B larger than the measurement of solute A? Answer in terms of the number of times out of 100.

Each plot shows the quantity of solute in one vial of seawater for each of A and B. **Use the buttons at the top of the plot to pause, restart, or step forward and back through the plots** if you want to see them at your own pace.

How well can you guess?

"Interpret Data Charts" - Number 8 out of 9

This task was included in the experiment, but omitted from reported results since A and B are correlated and this information is not available in the violin plots or error bars.

The chart summarizes two data sets that scientists have gathered by measuring the concentration of two chemical solutes (A and B, each measured in parts per million) in many vials of sea water.

Q: How often is the measurement of solute B larger than the measurement of solute A? Answer in terms of the number of times out of 100.