

# Sampling and Inference

1) Evaluate the `big-animals-table` in the Interactions Area. This is the *complete* population of animals from the shelter! Below is a true statement about that population:

The population is 47.7% fixed and 52.3% unfixed.

2) How close to these percentages do we get with random samples?

Type each of the following lines into the Interactions Area and hit "Enter".

```
random-rows(big-animals-table, 10)
random-rows(big-animals-table, 40)
```

3) What do you get?

---

Two tables, with 10 and 40 rows. Rows seem randomly chosen.

4) What is the contract for `random-rows` ? # random-rows :: (t :: Table, size :: Number) -> Table

5) What does the `random-rows` function do?

---

Given a table and a sample size, it randomly selects that many rows from the table.

---

6) In the Definitions Area, define `tiny-sample` and `small-sample` to be these two random samples.

7) Make a `pie-chart` for the animals in each sample, showing percentages of fixed and unfixed.

- The percentage of fixed animals in the entire population is 47.7%.
- The percentage of fixed animals in `tiny-sample` is (answers will vary).
- The percentage of fixed animals in `small-sample` is (answers will vary less).

8) Make a `pie-chart` for the animals in each sample, showing percentages for each species.

- The percentage of tarantulas in the entire population is roughly 5%.
- The percentage of tarantulas in `tiny-sample` is (answers will vary).
- The percentage of tarantulas in `small-sample` is (answers will vary less).

9) Click "Run" to direct the computer to generate a different set of random samples of these sizes. Make a new `pie-chart` for each sample, showing percentages for each species.

- The percentage of tarantulas in the entire population is roughly 5%.
- The percentage of tarantulas in `tiny-sample` is (answers will vary).
- The percentage of tarantulas in `small-sample` is (answers will vary less).

10) Which repeated sample gave us a more accurate inference about the whole population? Why?

---

The small-sample was more accurate, because larger samples are more likely to represent the population they are drawn from.

---