

HW 4: Unicorn Genetics

Your job is to make some predictions about the next generation of unicorns.

Some unicorns have large, majestic horns. Others do not, and just appear to be horses. A third group falls in between, and has a small nub where the horn would be. This is determined by a single gene, one copy of which is inherited from each parent. Animals that are homozygous for the horned trait (**HH**) have large horns, heterozygous individuals (**HN**) have small nubs instead of big horns, and homozygous no-horn individuals (**NN**) have no horn at all.

Through much painstaking research, you have found the single SNP (single-nucleotide polymorphism) responsible. It lies on base #1933 (1-indexed) of a particular gene in the unicorn genome. A cytosine (C) at this location corresponds to the horned trait, and a thymine (T) corresponds to the no-horn trait. Since every unicorn has two copies of this gene, any unicorn may have a cytosine in both copies, a thymine in both copies, or a cytosine in one and a thymine in the other—corresponding to the observed **HH**, **NN**, and **HN** phenotypes.

You have a herd of ten unicorns, with five females and five males. You have sequenced each copy of the horn gene in each individual. This data is found in the files `female-unicorns.gene` and `male-unicorns.gene`. In each file the first two lines are the genes of the first unicorn, the second two lines are for the second, and so on.

Assume that when a foal is born, that its mother and father are both chosen randomly. Thus, each foal will have one random horn gene from the male file, and one from the female file. Ask the user for the files, and then calculate the percentage of foals that you expect to be from each of the three phenotypes.

Your solution may look like this. (Note that your numbers may be different.)

```
Please enter the male sequence file: male-unicorns.gene
Please enter the female sequence file: female-unicorns.gene
```

```
Thank you. You should expect 30.0% of foals to have horns, 50.0% to have
nubs, and 20.0% to have no horn.
```

For full credit, you must break this task down into subfunctions, which work together to create the overall program. Remember to comment well, including a comment in front of every function to explain what it does. You should call your notebook `unicorn-genetics.ipynb`.

