Putting together a probabilistic model-- that is, a model of a random phenomenon or a random experiment-- involves two steps.

First step, we describe the possible outcomes of the phenomenon or experiment of interest.

Second step, we describe our beliefs about the likelihood of the different possible outcomes by specifying a probability law.

Here, we start by just talking about the first step, namely, the description of the possible outcomes of the experiment.

So we carry out an experiment.

For example, we flip a coin.

Or maybe we flip five coins simultaneously.

Or maybe we roll a die.

Whatever that experiment is, it has a number of possible outcomes, and we start by making a list of the possible outcomes-- or, a better word, instead of the word "list", is to use the word "set", which has a more formal mathematical meaning.

So we create a set that we usually denote by capital omega.

That set is called the sample space and is the set of all possible outcomes of our experiment.

The elements of that set should have certain properties.

Namely, the elements should be mutually exclusive and collectively exhaustive.

What does that mean?

Mutually exclusive means that, if at the end of the experiment, I tell you that this outcome happened, then it should not be possible that this outcome also happened.

At the end of the experiment, there can only be one of the outcomes that has happened.

Being collectively exhaustive means something else-- that, together, all of these elements of the set exhaust all the possibilities.

So no matter what, at the end, you will be able to point to one of the outcomes and say, that's the one that

occurred.

To summarize-- this set should be such that, at the end of the experiment, you should be always able to point to one, and exactly one, of the possible outcomes and say that this is the outcome that occurred.

Physically different outcomes should be distinguished in the sample space and correspond to distinct points.

But when we say physically different outcomes, what do we mean?

We really mean different in all relevant aspects but perhaps not different in irrelevant aspects.

Let's make more precise what I mean by that by looking at a very simple, and maybe silly, example, which is the following.

Suppose that you flip a coin and you see whether it resulted in heads or tails.

So you have a perfectly legitimate sample space for this experiment which consists of just two points-- heads and tails.

Together these two outcomes exhaust all possibilities.

And the two outcomes are mutually exclusive.

So this is a very legitimate sample space for this experiment.

Now suppose that while you were flipping the coin, you also looked outside the window to check the weather.

And then you could say that my sample space is really, heads, and it's raining.

Another possible outcome is heads and no rain.

Another possible outcome is tails, and it's raining, and, finally, another possible outcome is tails and no rain.

This set, consisting of four elements, is also a perfectly legitimate sample space for the experiment of flipping a coin.

The elements of this sample space are mutually exclusive and collectively exhaustive.

Exactly one of these outcomes is going to be true, or will have materialized, at the end of the experiment.

So which sample space is the correct one?

This sample space, the second one, involves some irrelevant details.

So the preferred sample space for describing the flipping of a coin, the preferred sample space is the simpler one, the first one, which is sort of at the right granularity, given what we're interested in.

But ultimately, the question of which one is the right sample space depends on what kind of questions you want to answer.

For example, if you have a theory that the weather affects the behavior of coins, then, in order to play with that theory, or maybe check it out, and so on, then, in such a case, you might want to work with the second sample space.

This is a common feature in all of science.

Whenever you put together a model, you need to decide how detailed you want your model to be.

And the right level of detail is the one that captures those aspects that are relevant and of interest to you.