Hi. This tutorial covers paradoxes. So let's start with a definition of what a paradox is. And a paradox is a situation where phenomenon can be looked at in two seemingly contradictory ways. Paradoxes and statistics are common. Data interpretation often utilizes probability. And often, our intuition about probability is often misleading.

Probability gets pretty tricky. And a lot of times, what we think is true isn't actually true. So there are several paradoxes that come up in stats, and we're going to look at one in particular called Benford's Law.

So what Benford's Law is is a property that states that in many real world data sets, the first digit in whatever value you're measuring is not uniformly distributed. So if we're thinking about a data set that involves the values of houses, so if we have a big list of data, chances are that leading digit is not going to be uniformly distributed. Or if we think of the number of pages in a book, the first digit and all of those number of pages might not be uniformly distributed.

Our intuition might say, well, yeah, all of that stuff is fairly random. It would seem to be uniformly distributed. But Benford's Law says otherwise. So this is basically what he came up with. So the percent chances of each digit as a leading digit of a real world data value is predicted by Benford's Law.

So what that's saying is that the digit 1 has about a 30% of being the leading digit, OK? And if we think about that, if we go back to that housing value example, there are going to be a lot of houses that will be about $100,000. So because a lot of those values have a leading digit of 1, that probability becomes so much higher.

And actually, it ends up becoming a positively skewed distribution. So as you move to the right, the probabilities get lower and lower. So the chance of a 9 being a leading digit in a real world data set is pretty low, less than 5%. And I'll give you an example of this.

So this data from 2010. So there were 237 countries that were part of this survey. That was all the countries in the world. And we looked at the first digit of their population. So basically, what this graph does is you look at the leading digit, and the value on the left here is the percent of countries with that number as the leading digit. So 67 out of 237 had a leading digit of 1 in their population. So it might have been 17 million something. That country would be in this bar. So that bar goes up to about 27. Leading digit of 2, OK, that looks to be about 18%.

Now, what these black dots are is these black dots show what Benford's law predicted. So Benford's Law predicted this would actually be up a little bit more. But notice as you get lower and lower, those predictions become more accurate.

All right, so although our intuition might tell us that the leading digit of real world data would have a uniform distribution, Benford's Law tells us otherwise. Since our intuition in this case is wrong, Benford's Law is a type of statistical paradox. And there are others, but I just want to make sure that you had a good example of one specific statistical paradox. So that has been your tutorial on paradoxes. Thanks for watching.