http://topdrawer.aamt.edu.au/Statistics/Good-teaching/Making-informal-inferences/Single-measurement-variables/Paperclips

Paperclips Investigation video transcript

I have a container of paperclips here, quite colourful ones, and originally there were a thousand in there, but there are some missing now.

(Opening the container of paperclips)

Now, if I wanted to know exactly how many paperclips were in that container, I could take them all out and count them or I could use a method that scientists use when they want to determine how big a population is when they can't possibly count it and that's called tag, release and recapture. Let's try it on the paperclips.

(Explaining method to count paperclips)

Now, my tagged paperclips are going to be these silver ones because I'll be able to distinguish them quite easily.

(Taking out silver paperclips from a box)

So I am going to add 100 tagged paperclips to my population.

(Adding silver tagged paperclips to untagged coloured paperclips)

Now, I have to make sure that this is really well distributed throughout my population so I make sure I give it a really good shake.

(Shaking and mixing all paperclips)

When I think that I have it quite well distributed, I'm going to draw out a sample of 50 paperclips. I think they will be representative of the population. So the proportion of tagged paperclips in my sample should tell me something about my population.

(Explaining process of calculating proportion from a sample)

So there were 5 tagged paperclips in a sample of 50. So our 100 tagged clips should be

the same fraction of our unknown population. 5 out of 50 would be equal to 100 out of our unknown population.

(Calculating fraction in a given population)

So using equivalent fractions, if we multiply the numerator by 20 and we multiply the denominator by 20 then we see that the population that we didn't know previously is 50 times 20, which would be 1000.

(Calculating population from sample)

But remember, I added 100 paperclips to the container before I took the sample so the original number that was in there was not 1000. Because I added 100, the original number was 900.

(Calculating original population)

Now, that prediction was based on just one sample. I can do a much more accurate estimate of the population by taking a number of samples and looking at the average.

(Explaining average calculation from different samples)