

Infuriatingly, the answer to this final problem is one of the simplest in the entire book. It is this simplistic genius in the face of a seemingly impossible predicament that makes it such a great problem.

Firstly, let's remind ourselves briefly of the key facts of the problem:

1. We need to survey a population to determine the spread of the Van Basten Virus.
2. Everyone with the virus wants the true infection rate to be established accurately to ensure the government funding to fight the virus is secured.
3. No-one will do anything that will leave them on record as being infected with the virus, for fear of being sent to Wales.
4. Furthermore, the government insists that the actual answer given by everyone surveyed is recorded alongside their name to avoid any fraudulent answers.

In short, we need to be able to ask someone a question, be able to use their answer to discover the spread of the virus, but without the answer actually telling us whether they have the virus or not. Very tricky indeed.

The first thing that becomes quite obvious is that there is no single question that you could ask that would satisfy all of the above criteria, so you immediately realise that it must be a two-step process.

The next thing you conclude is that you cannot avoid asking people the direct 'Do you have the virus?' question.

Then, you appear stumped, since the answers must be published.

A lot of people make a mistake here, in thinking that it is an individual's ANSWER that needs to be kept secret. Not at all. In fact, the answers cannot be kept secret due to the government's stipulations.

The real genius of this problem is that it is the QUESTION that needs to be kept secret.

Let us explain.

There will be three steps to this process.

STEP ONE

The first step will introduce a random element. Let's say in this case that the person being surveyed tosses a coin.

STEP TWO

Answer one of two questions, depending on the result of the coin toss in step 1:

If it was heads, answer the question 'Do you have the virus?'

If it was tails, toss the coin again and answer the question 'Did you get heads on your second toss?'

STEP THREE

Put 'Yes' or 'No' to whichever question you answered on your ballot paper. All that is recorded on the ballot paper is your name and either a 'Yes' or a 'No'.

For example:

'Dr Hans Crikey Mosey – Yes'

So the government, if suspicious, can very easily ring up Dr Hans to check that he did indeed answer YES on the survey, but without knowing which question he was answering.

Now, using maths, we know that the probability of getting a head or a tail is 50%, so that when enough people have been surveyed, we know that 50% of them will have actually answered the virus question in step two.

Imagine 1000 people did the survey and we ended up with 350 YES's and 650 NO's in the ballot box at the end.

50% of these would have not answered the virus question.

So, 500 people would have answered the coin toss question in step two. As the probability is 50% of throwing a heads or tails, we know that roughly half of these 500 would have answered 'yes' and the rest would have answered 'no'. So we can assume that we would get close to 250 yeses and 250 nos, give or take a few due to random variation.

We need to subtract these coin toss answers from the total answers to find the answer to the virus question.

This would leave 100 ($350 - 250$) yeses for the Virus question and 400 ($650 - 250$) nos for the virus question.

Thus, 100 people would actually have the virus and 400 would not.

However, only 500 people (50%) would have answered the virus question, so the overall infection percentage of the population would be 20% of the population (100 out of 500).

So this would mean that of the 1000 people surveyed, 200 would be infected and 800 would be clear.

So we've got an accurate indication of the spread of the virus, a record of everyone's answers, but nothing in these answers that gives the government any reason to suspect they have the virus. Genius.

The larger the sample, the more accurate the survey results.