Things That keep Me up at night

(as a data scientist)

I am not a

- Mathematician
- Statistician

I am going to talk about:

- Confirmation Bias
- Bayesian Traps
- Simpson’s Paradox

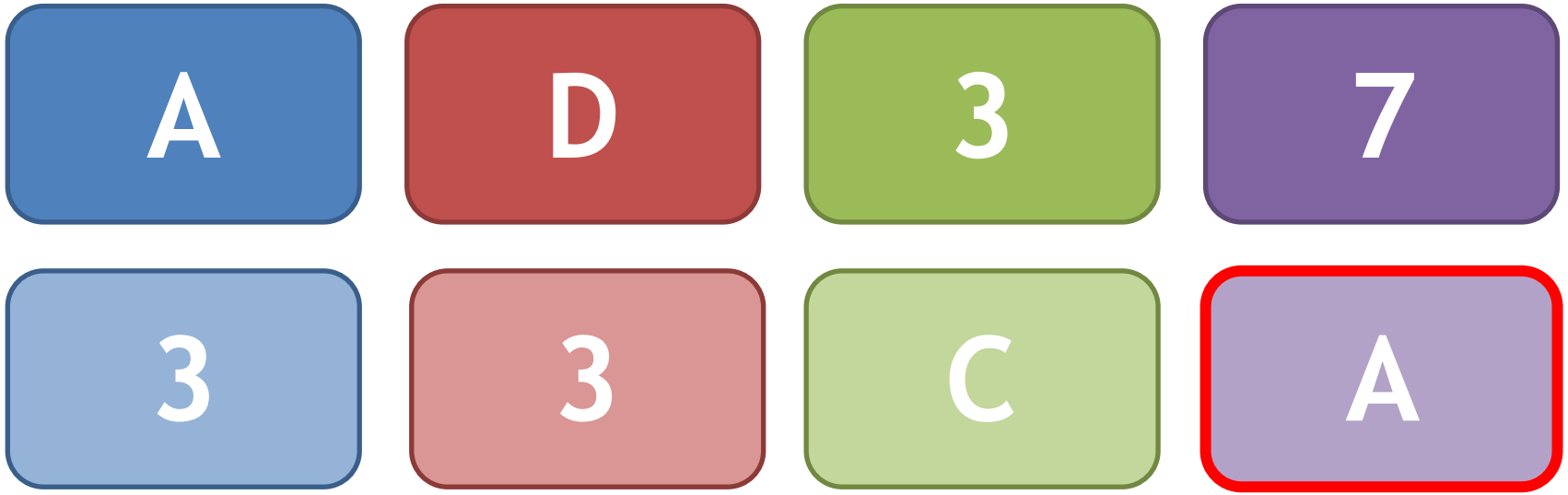
Confirmation Bias

What do you know

Each card has a letter on one side and a number on the other

What you want to prove

If there is an **A** on one side of the card, there is a **3** on the other side



Bayesian Traps 2/3

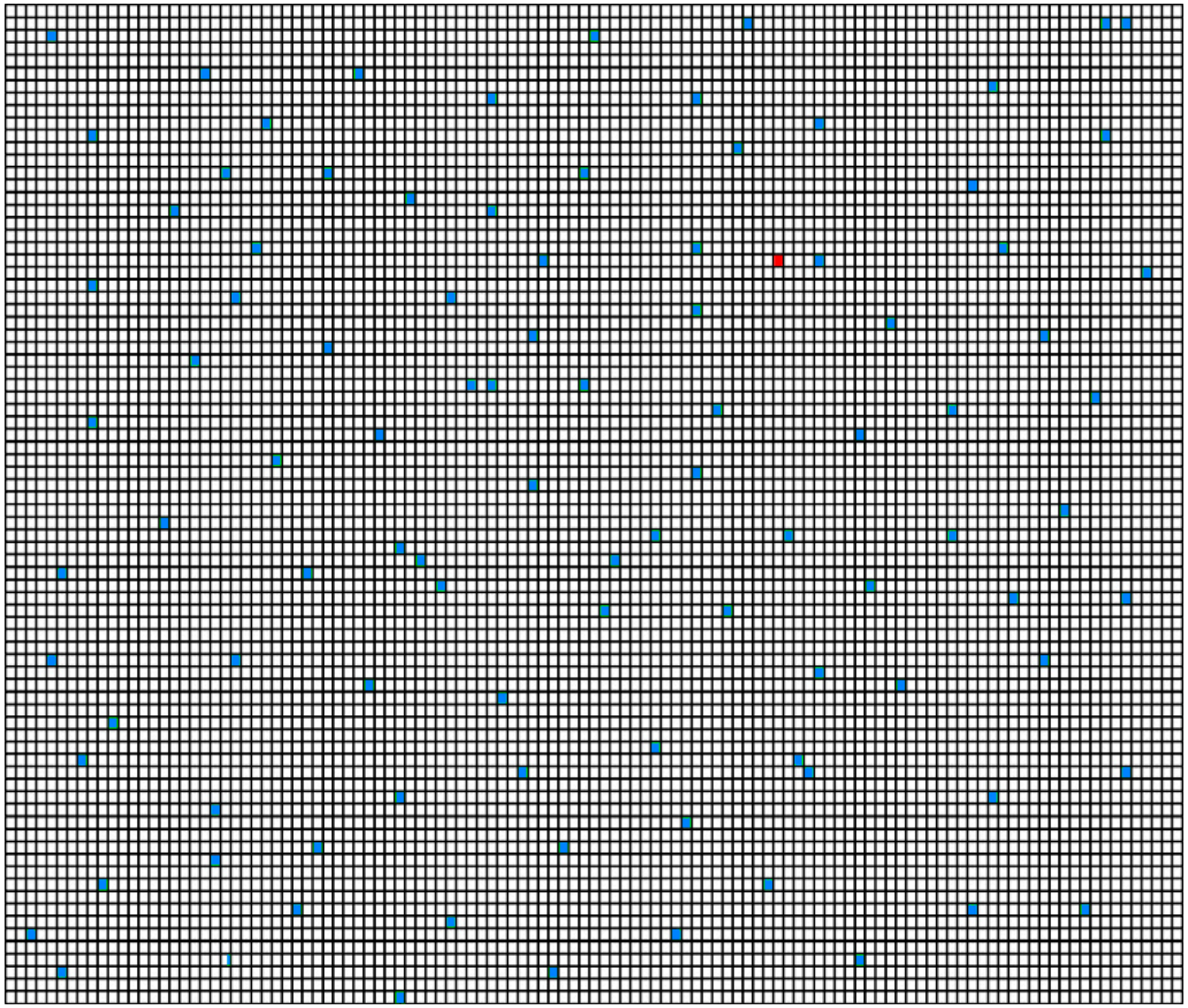
You hear about this terrible disease going around that affects around **1:10000** people and is lethal.

The doctor tells you there is a simple blood test that is **99% accurate** and you take it.

A week later results come in and the test indicates that you have the disease.

Should you start writing your will?

Bayesian Traps

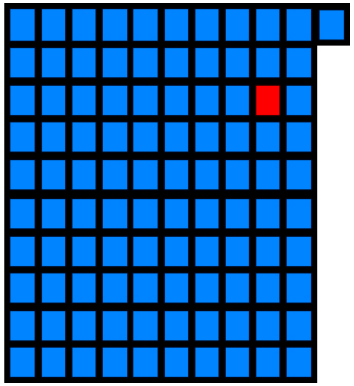


Affects **1:10000**

99% accurate

1% False positives

100 F.P.

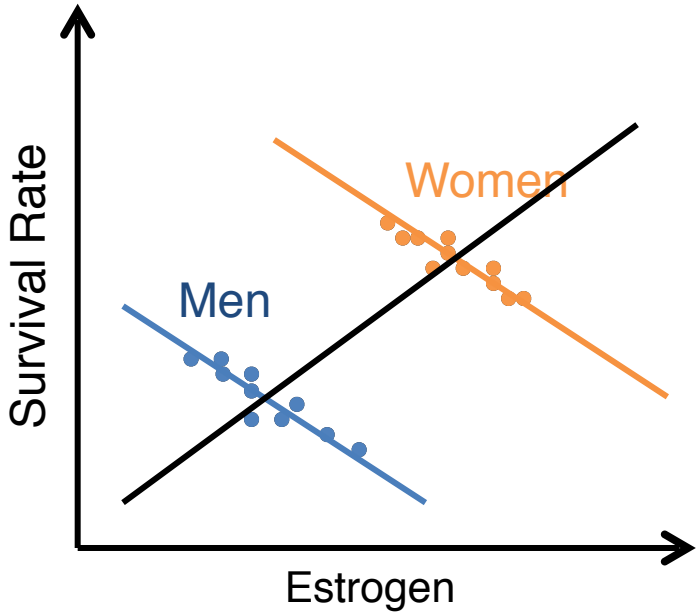


1/101 (<1%) chance you actually have the disease

Simpson's paradox

Say you are running a drug trial and for 2 weeks you go around, you grab people that are sick and you test 2 different vaccines:

	Vaccine A	Vaccine B
Week 1	45/65 69%	25/35 **71%**
Week 2	15/35 43%	30/65 **46%**
Total	60/100 **60%**	55/100 55%



Thank you

tiagovieira.pt

email@tiagovieira.pt

@TVieira

TMMV

tiagommvieira