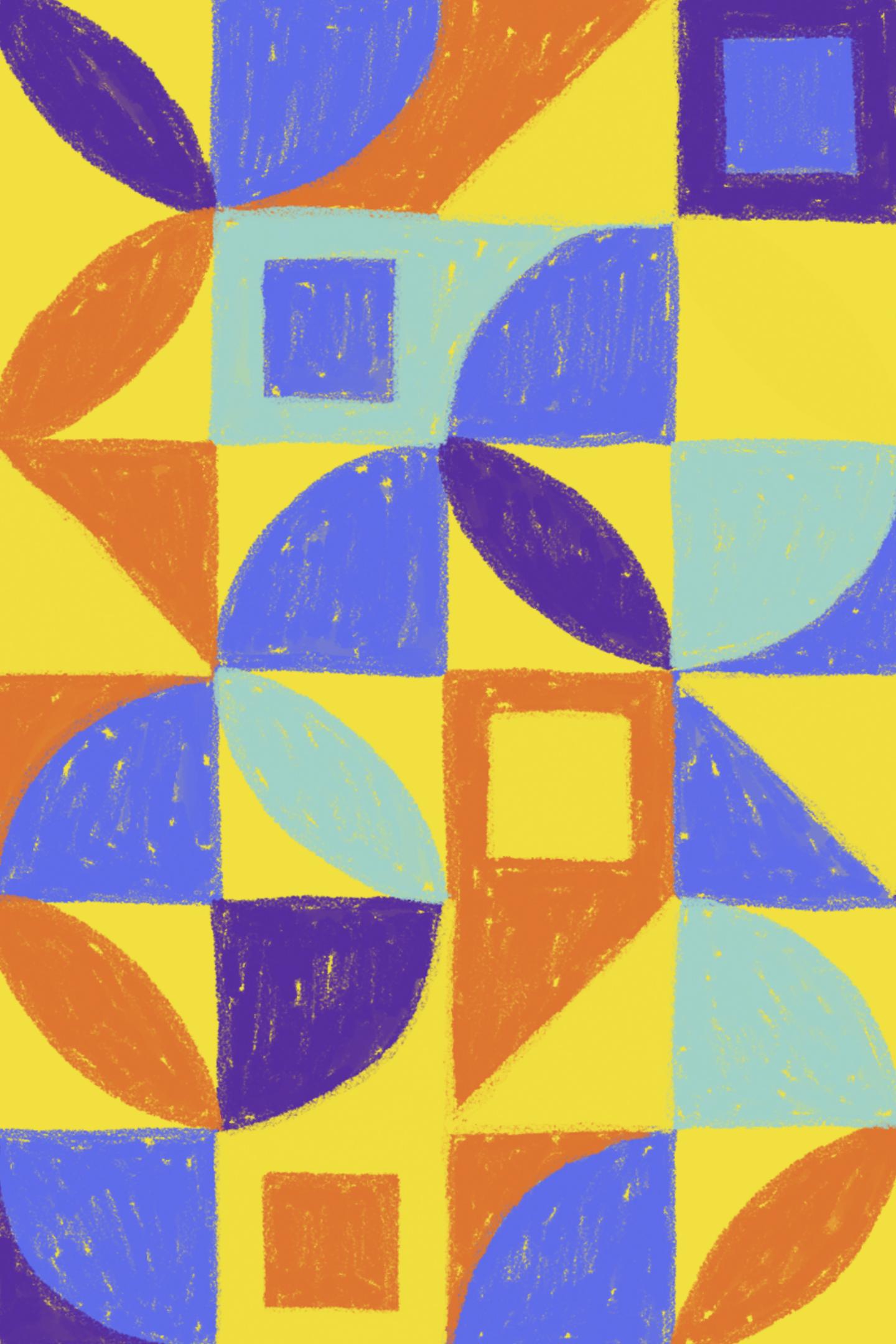


Florence Nightingale Day

10 January 2024

Nicola Rennie



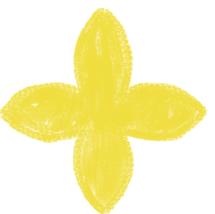


Hello!

I'm Nicola, a Lecturer in Health Data Science at Lancaster University.

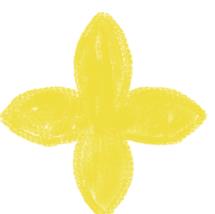


What's this talk about?



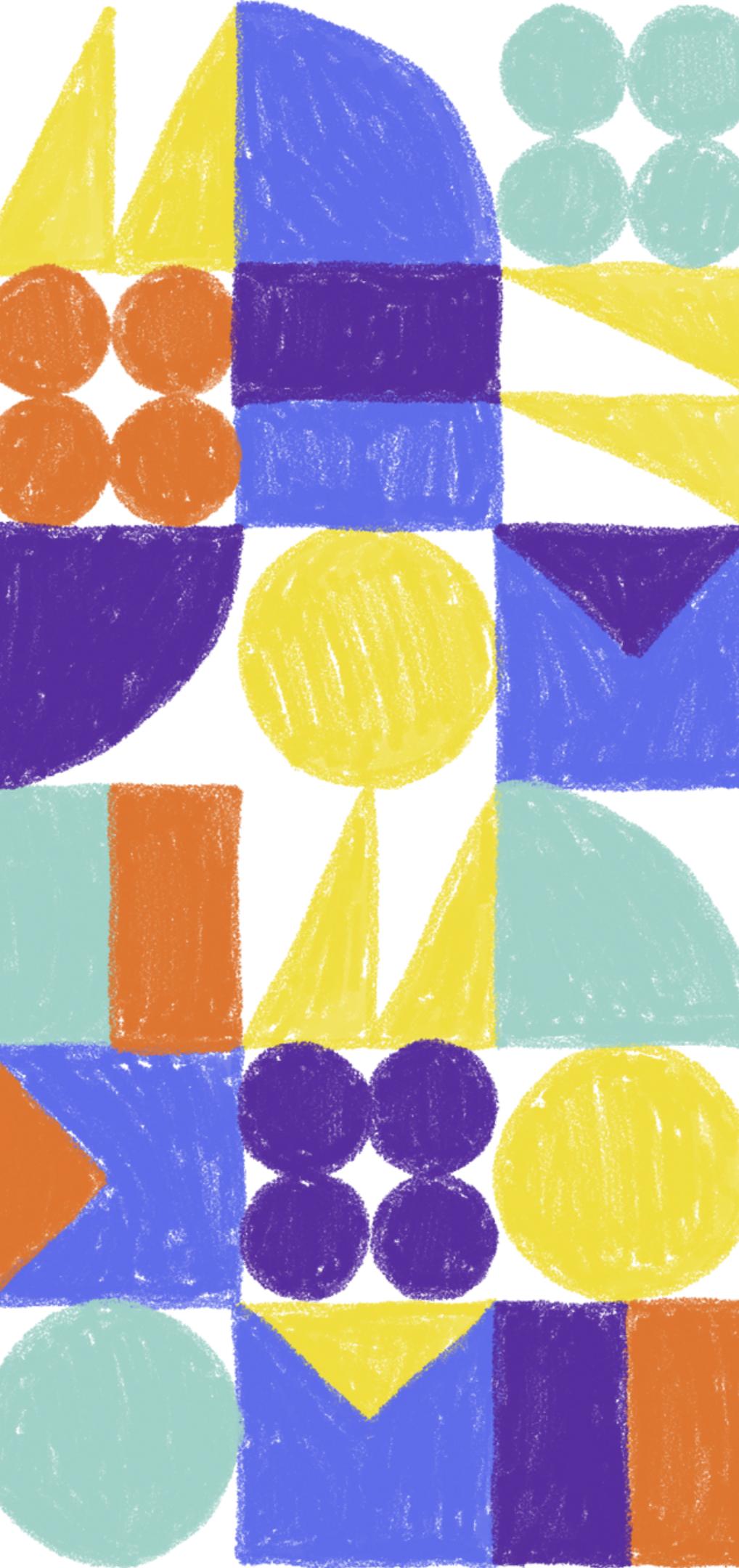
How I got here

My mathematics and statistics journey



What I do now

Using data science to improve health outcomes

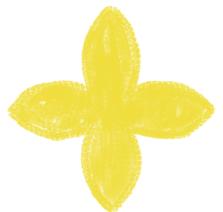




My mathematics and statistics journey

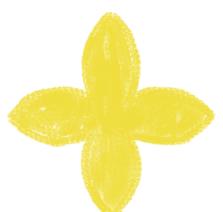


School



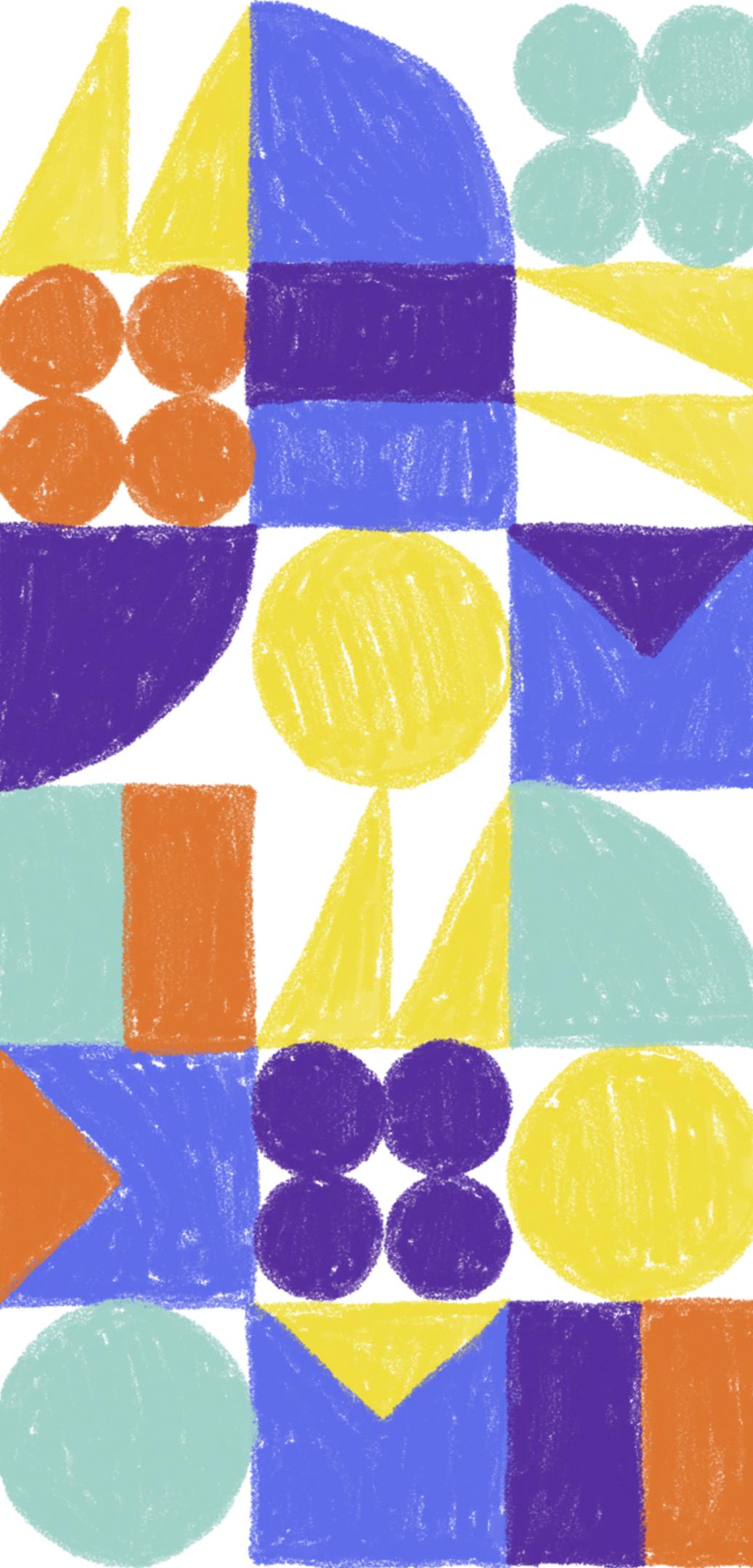
Maths was my best subject at school

Elaborate on what you want to discuss.

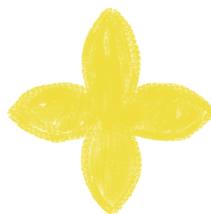


Art and music were my favourite subjects

Creativity and maths go hand-in-hand

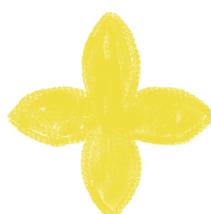


University: undergraduate



Studied Mathematics at St Andrews

Enjoyed statistics courses with applications
to real-world data

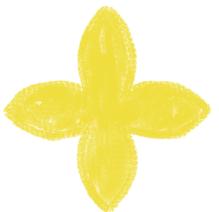


Maths opens the door to many careers

But I still didn't know what I wanted to do

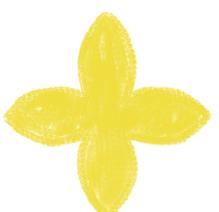


University: postgraduate



Joined STOR-i CDT in 2017

MSc and PhD in Statistics and Operational Research

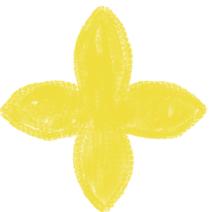


Collaboration with industry

Working with Deutsche Bahn to model railway demand

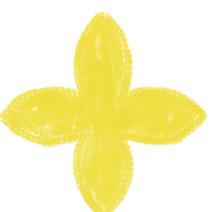


A “proper” job...



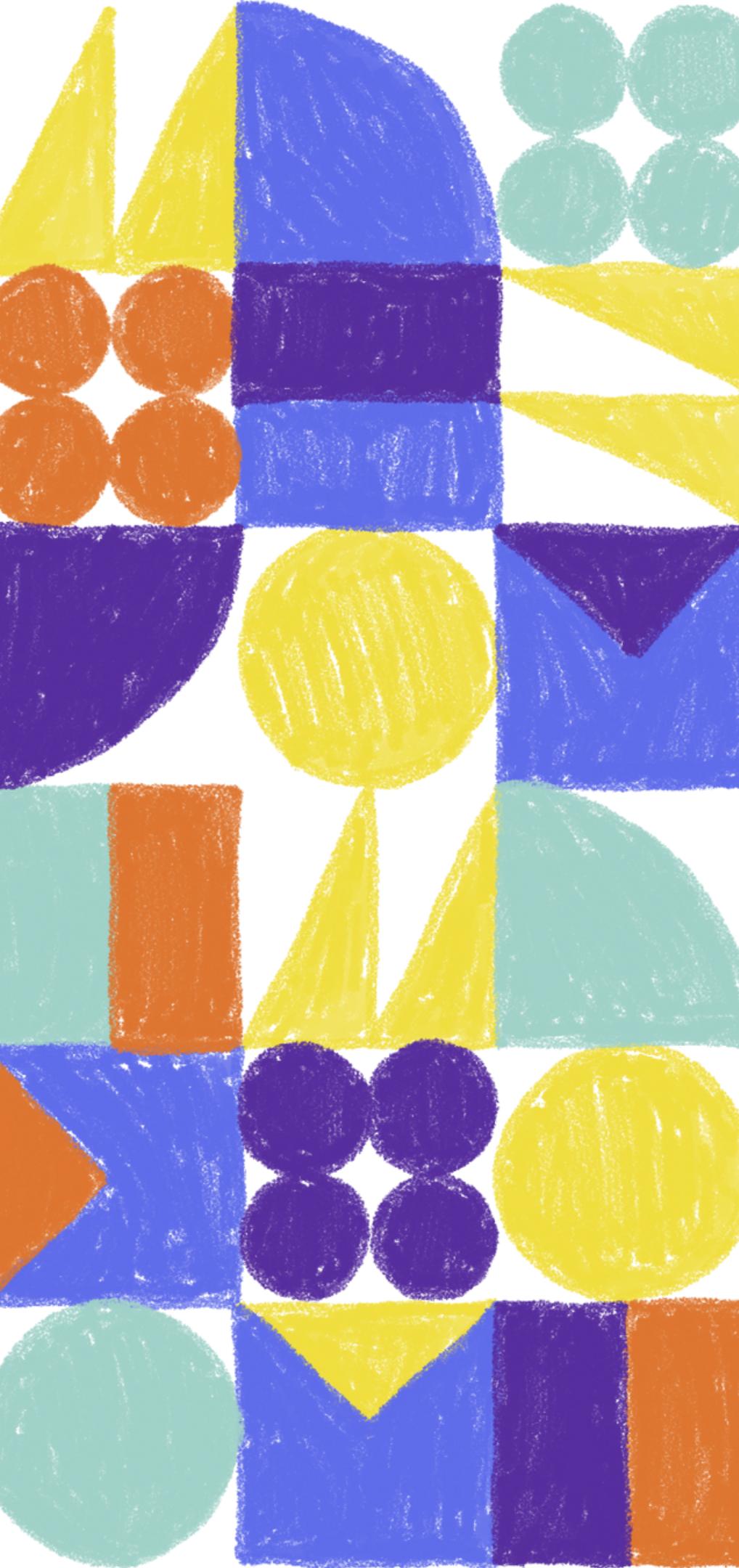
Data science consultancy

Solving problems with a wide range of companies

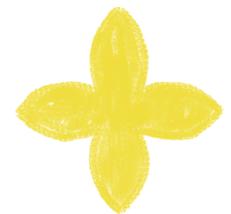


Running training courses

Teaching people about programming and statistics

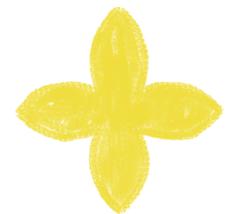


Lecturer



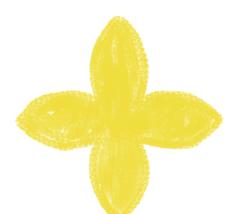
Lecturer in Health Data Science

Teaching and supervising epidemiology, data science, and medical students



Collaborating with the NHS

Working on data with local NHS trusts and developing software

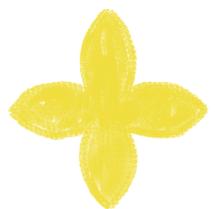


Building communities

Organising R-Ladies meetups and delivering workshops



Why I like my job



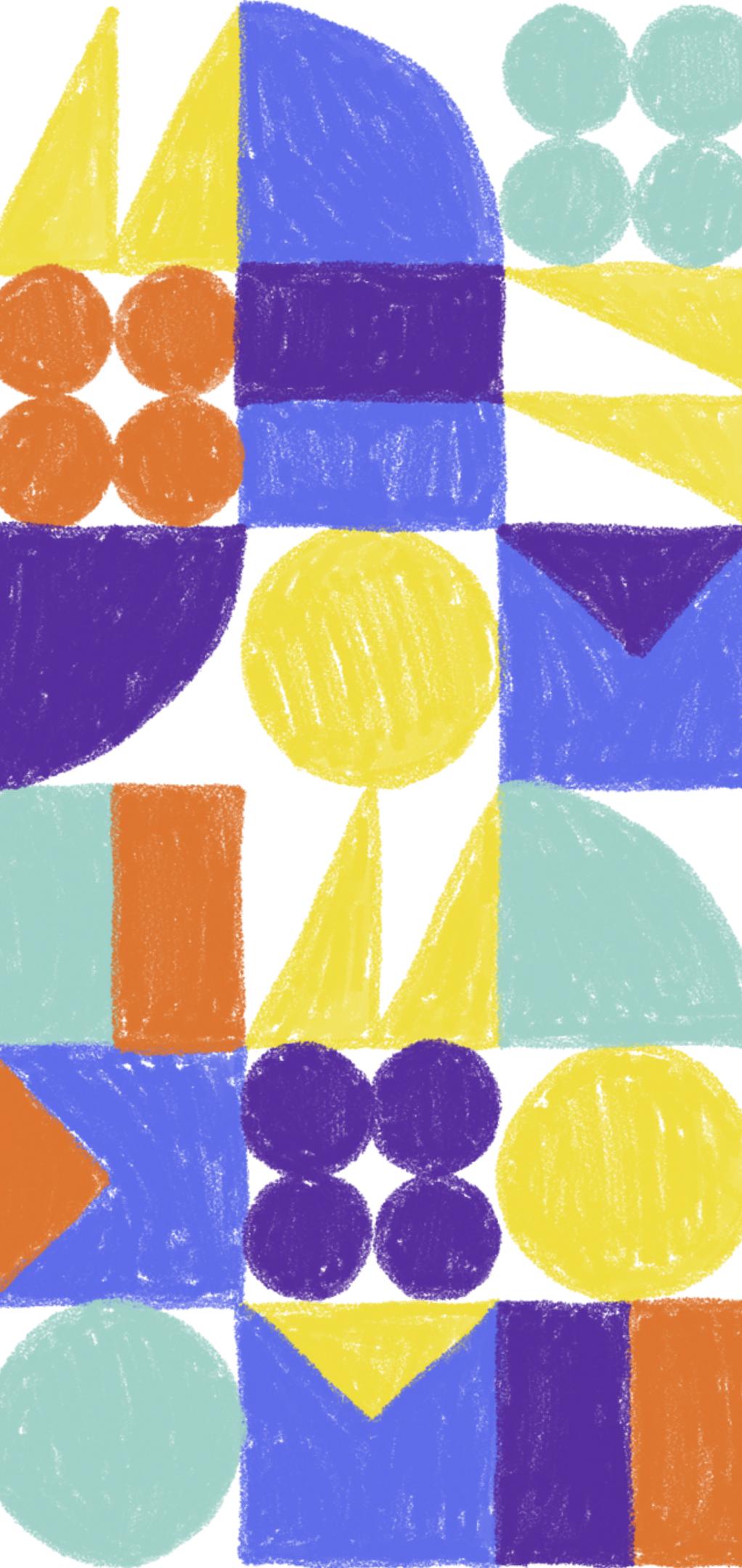
I feel like I'm doing something useful



I work with an amazing group of people



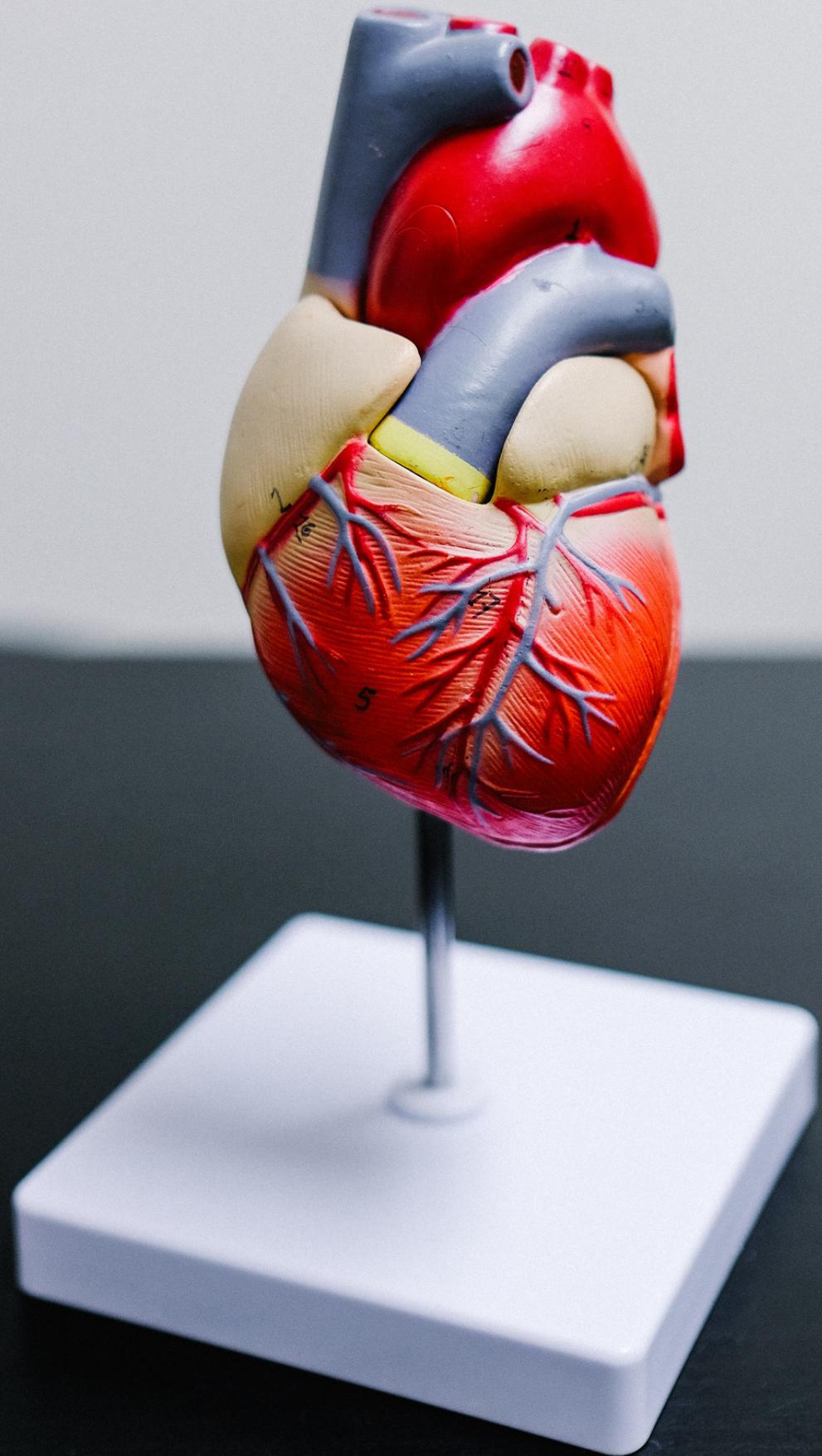
I get to do lots of different, interesting things





Using data science to improve health outcomes

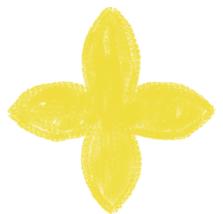




Can you
detect a
heart
murmur?

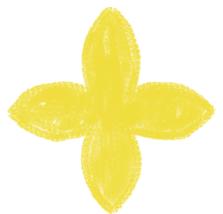
Using statistics to aid diagnosis

Heart murmurs



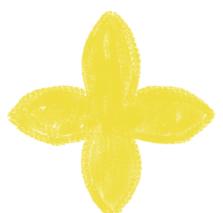
What are heart murmurs?

A heart murmur is an extra, unusual sound in your heartbeat.



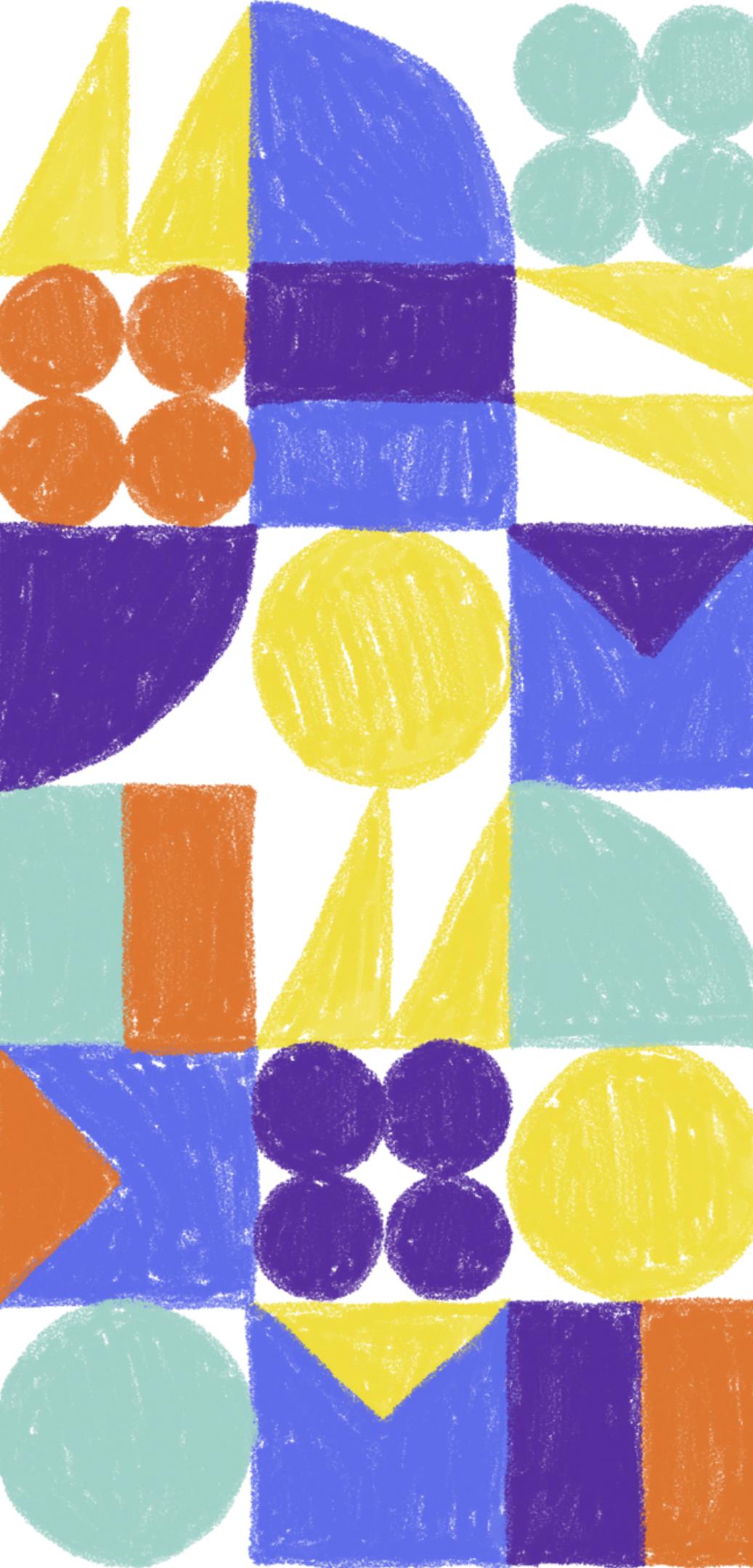
Are they dangerous?

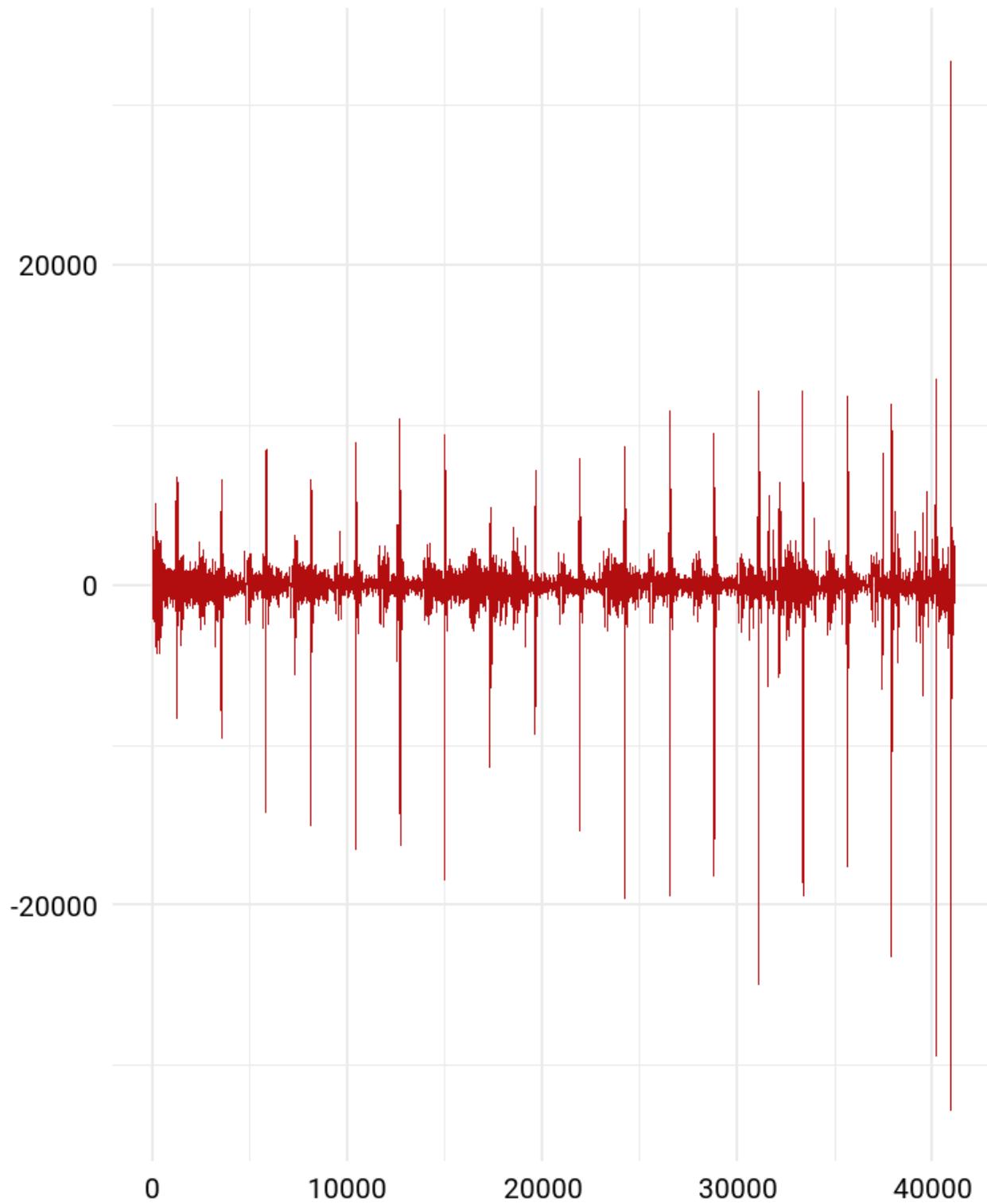
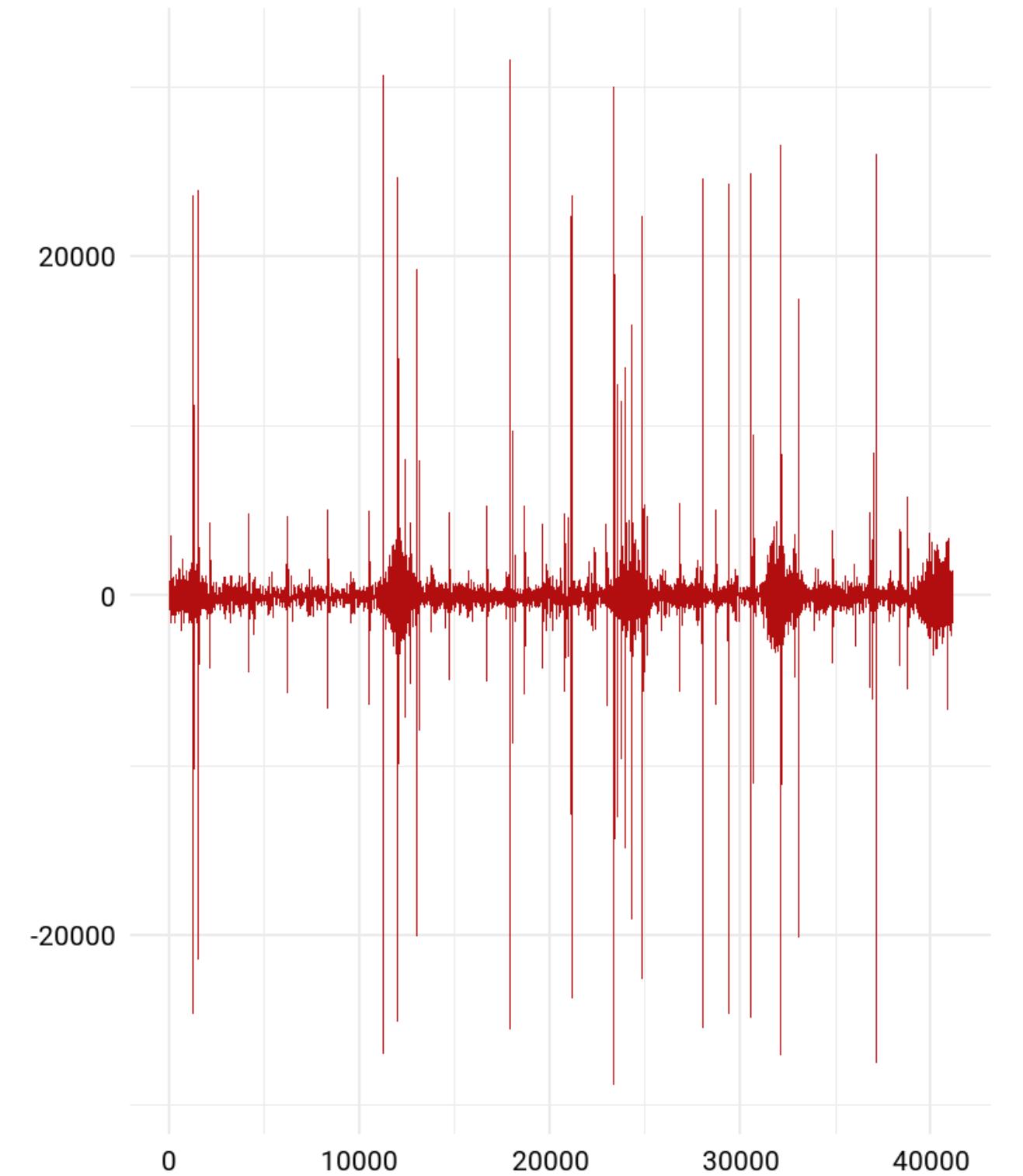
There are different types of heart murmurs. Some are more dangerous than others.



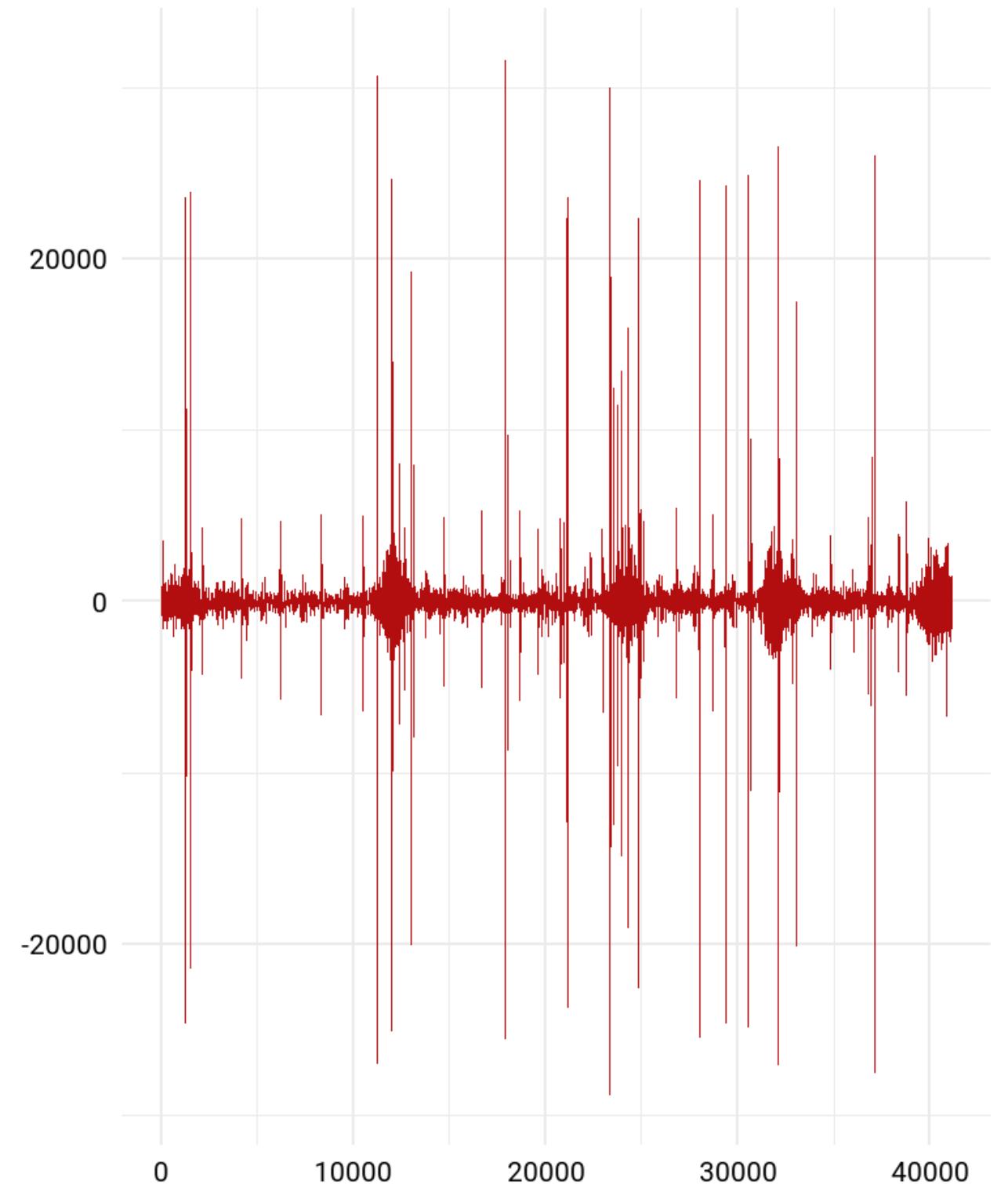
How are they diagnosed?

Heart murmurs can be diagnosed with a PCG (phonocardiogram).

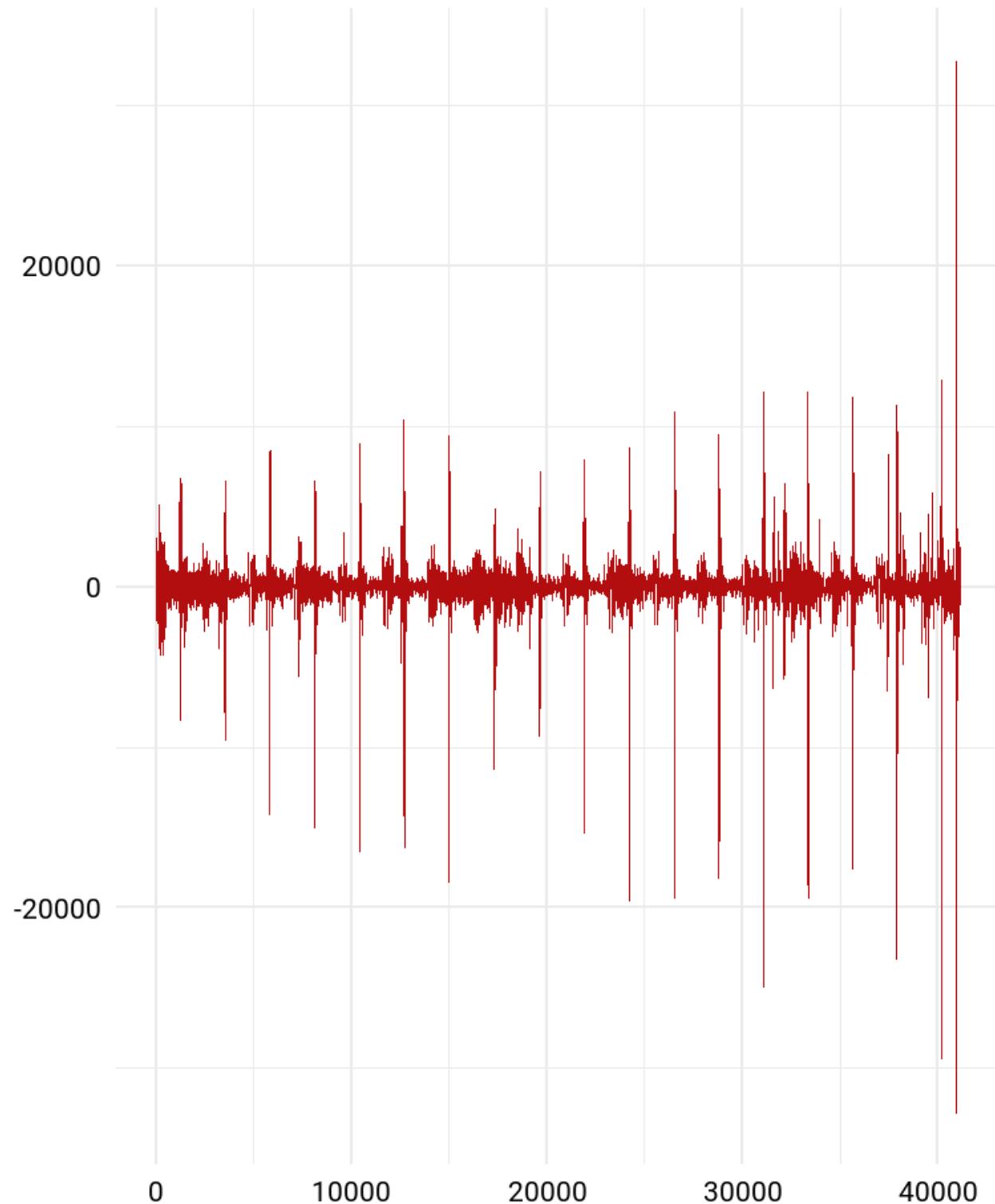




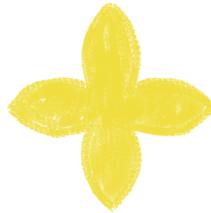
Subject: 23625 (Absent)



Subject: 13918 (Present)



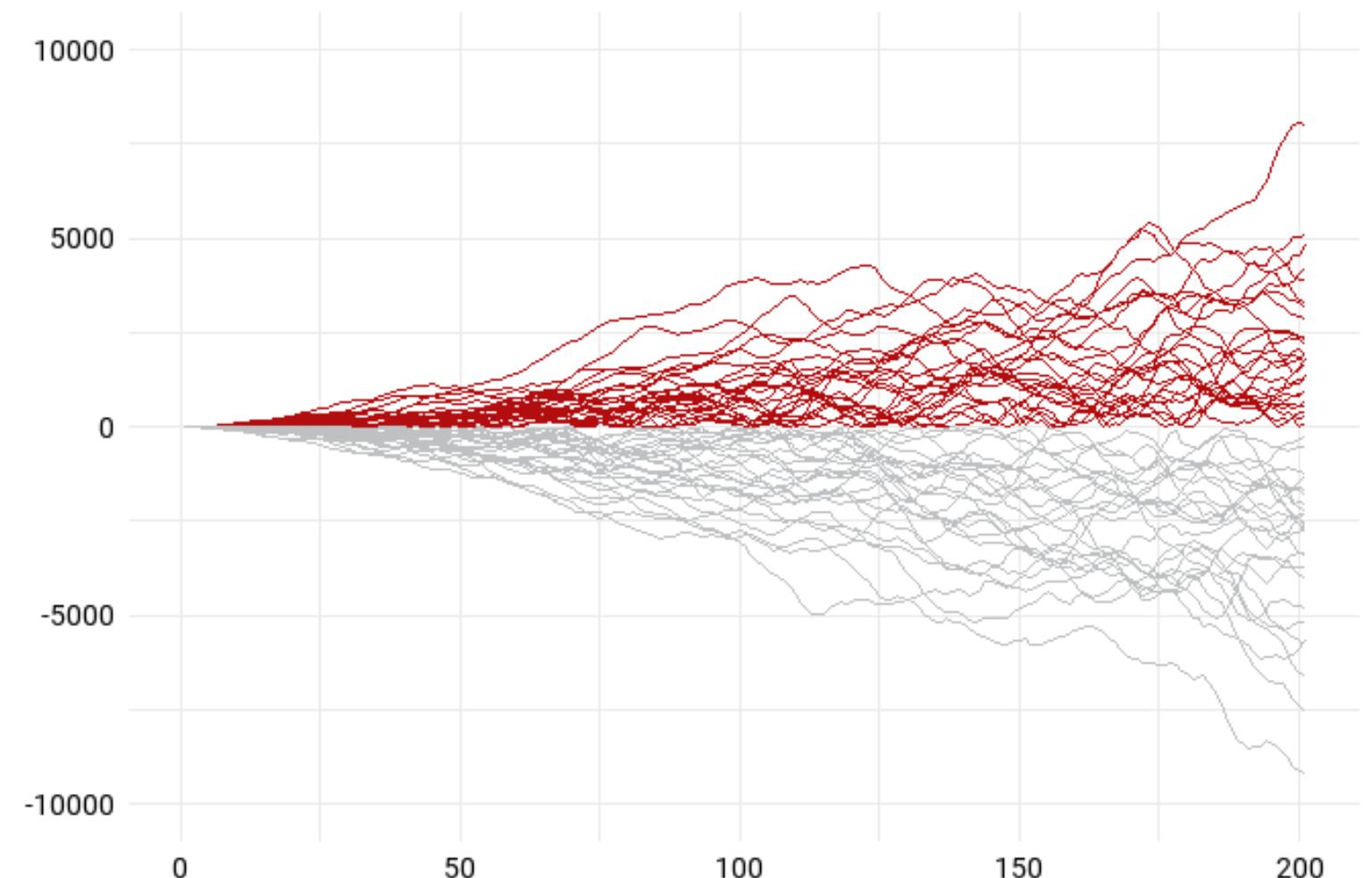
Aim



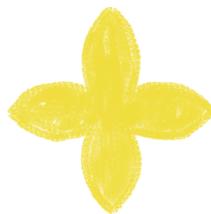
Group phonocardiogram signals

Two groups: heart murmur and no heart murmur

If only classifying time series was this easy...



Challenge

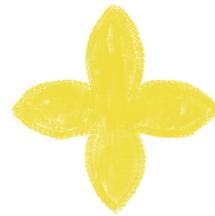


Data is very noisy

There are around 40,000 observations every 10 seconds.

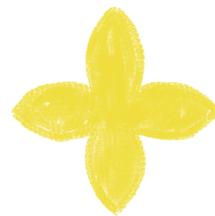


Method



Calculate *features* of the PCG signals

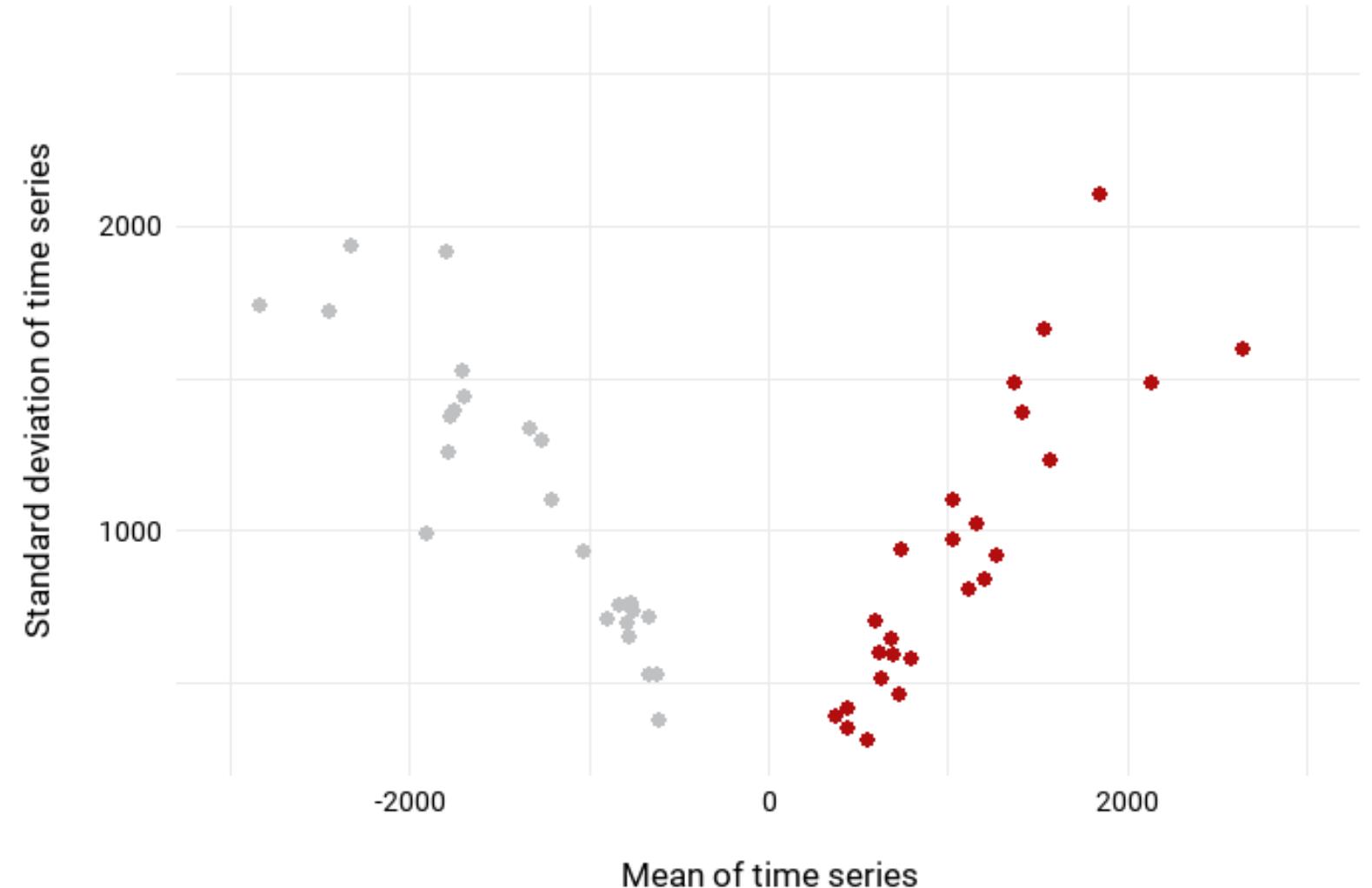
For example, measures of averages or variability.



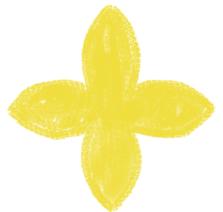
Group new features instead

Two groups: heart murmur and no heart murmur

This looks a little more familiar...

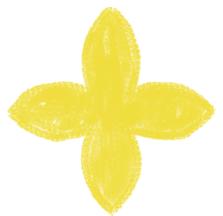


What's next?



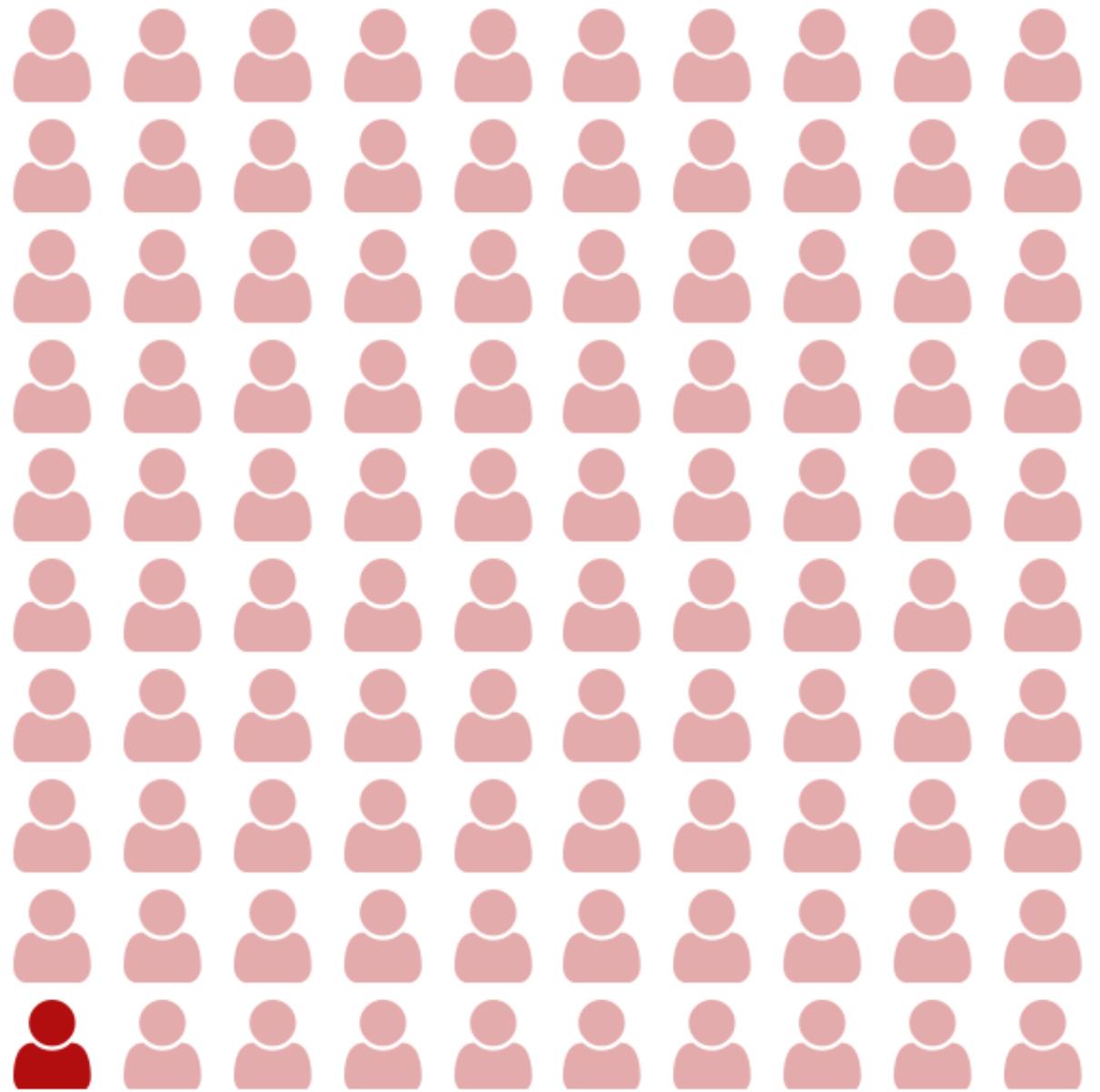
Improving predictions

Current methodology sometimes misses people who have heart murmurs.



Detect different types of heart murmur

Some heart murmurs are easier to detect than others. Some are more important to detect than others.





Evaluating NHS services

Collaborating with local NHS Trusts
in Lancashire

Collaboration with NHS Trusts



**Lancashire &
South Cumbria
NHS Foundation Trust**



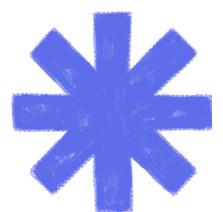
**University Hospitals of
Morecambe Bay
NHS Foundation Trust**



**Lancashire Teaching
Hospitals
NHS Foundation Trust**



Research questions



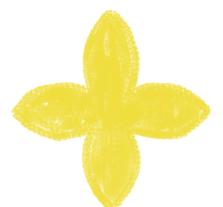
What affects how long someone stays in hospital?



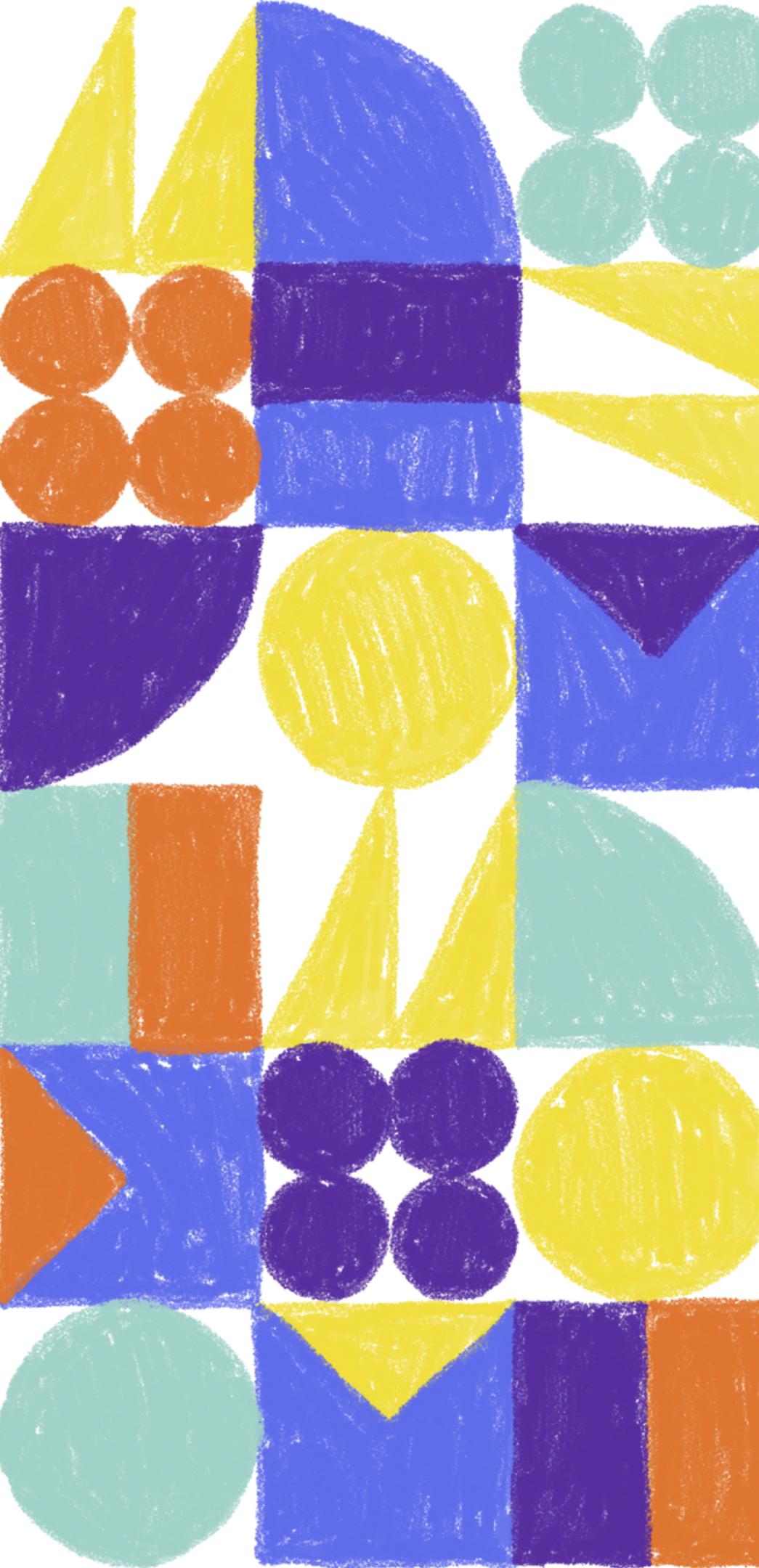
How many beds are required?



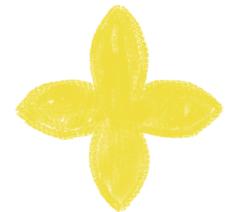
How do patients move through different NHS services?



How can we link different data sources?



NHS Data



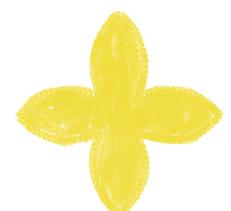
Data from many sources

The NHS collect a lot of data from different sources including primary care (e.g. GP), secondary care (e.g. hospital), and operational data.



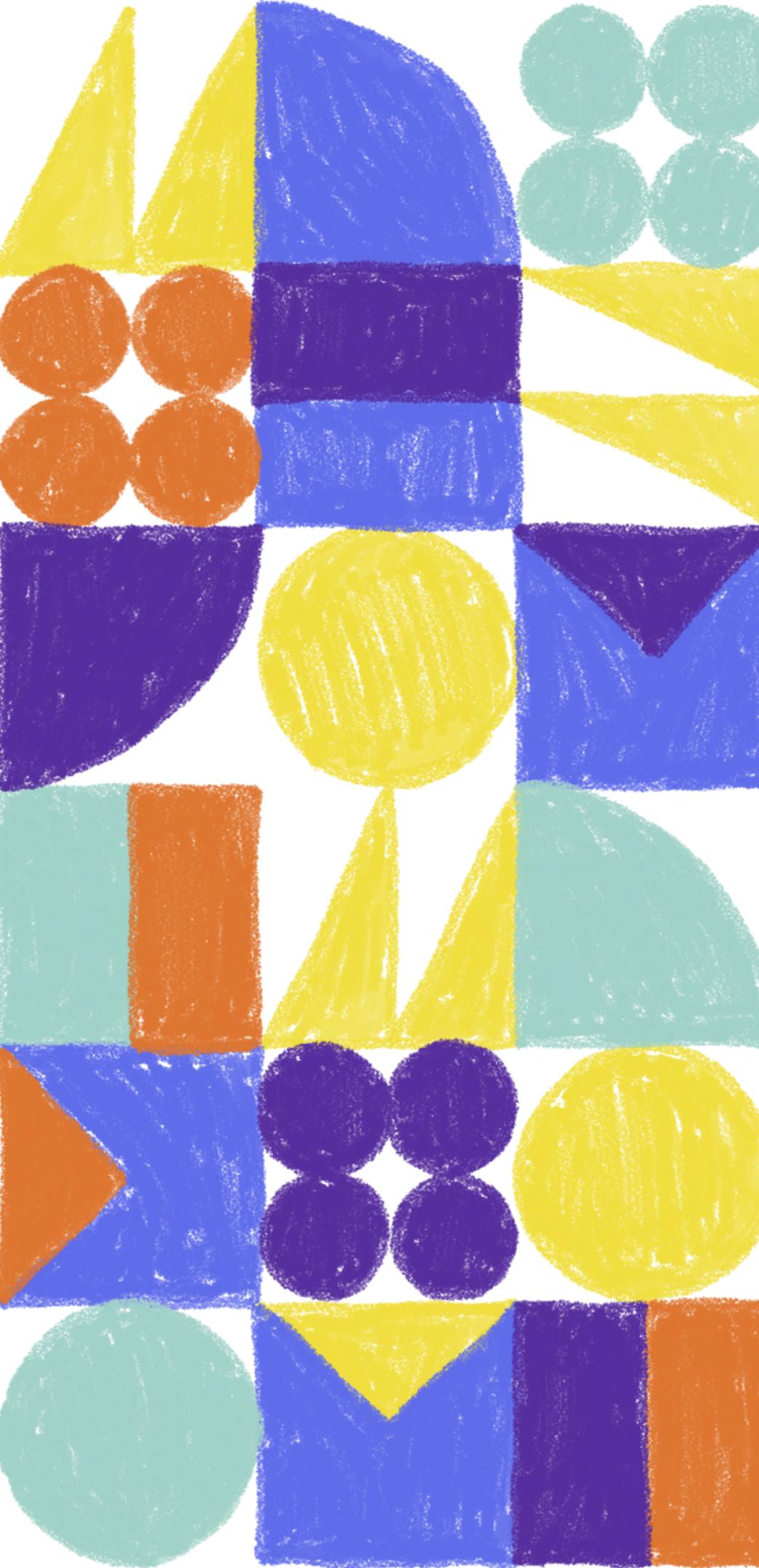
Data in many formats

Data can be in the form of spreadsheets, letters from doctors to patients, scan images, and may other things!

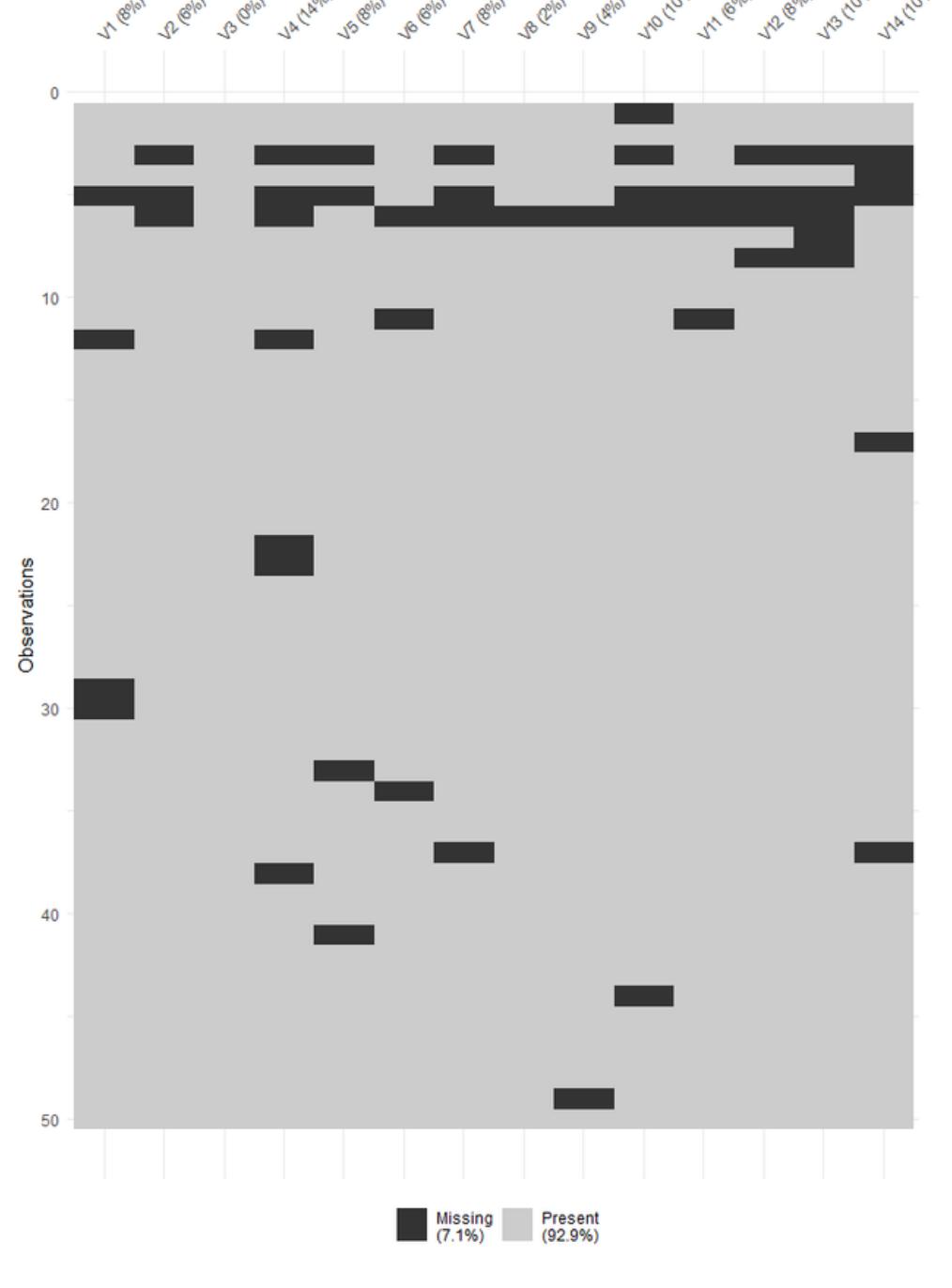
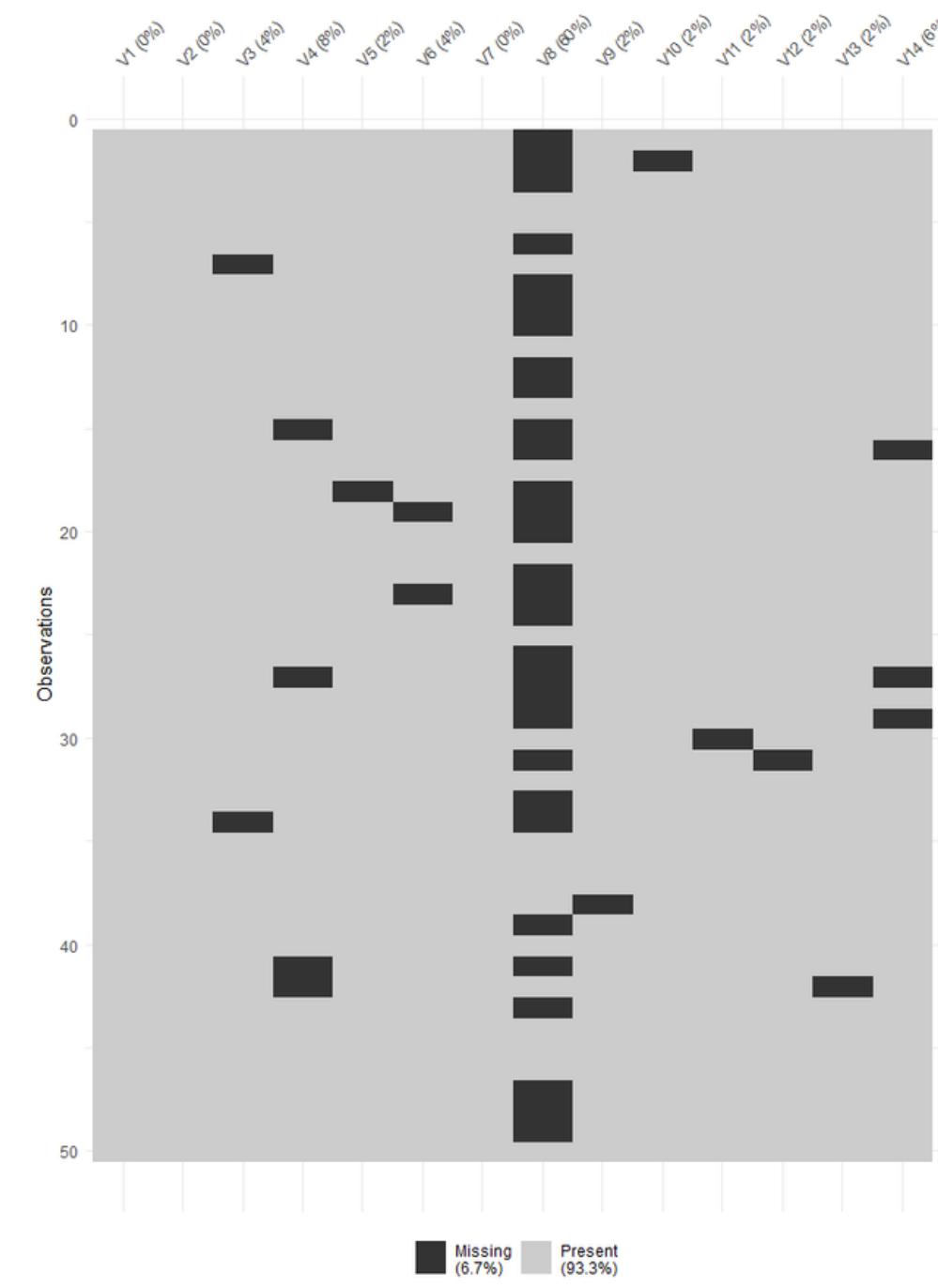
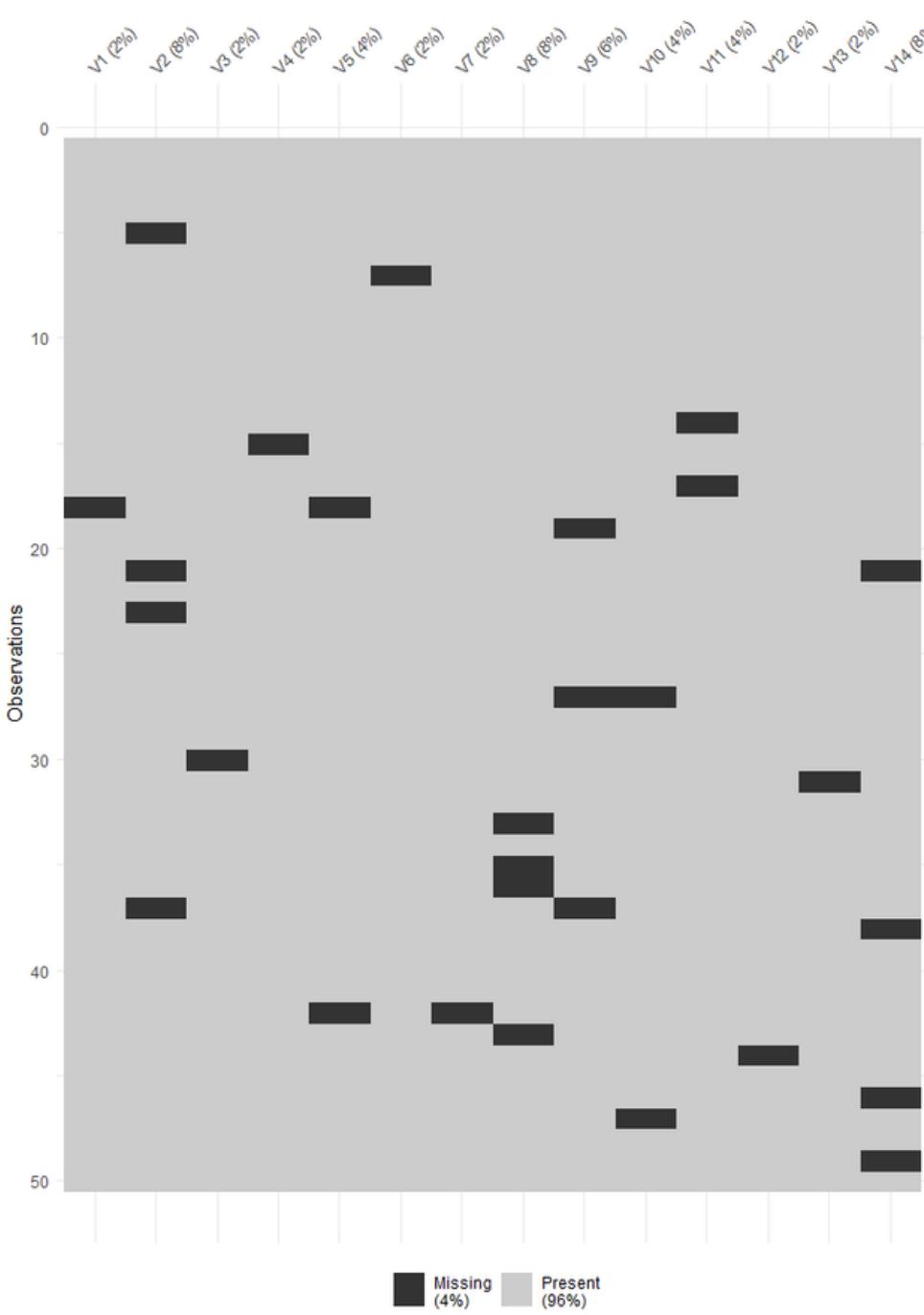


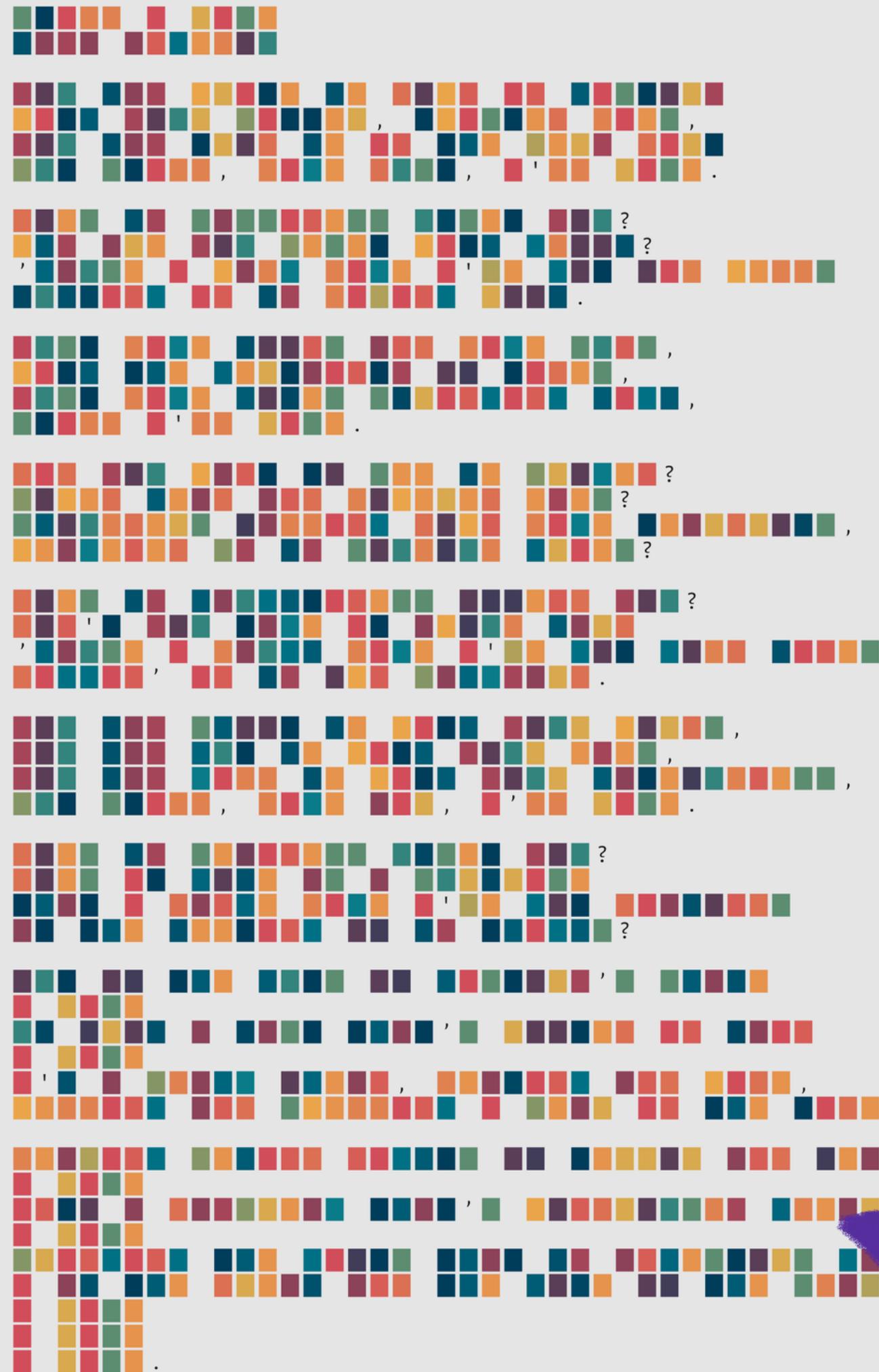
Data quality issues

Data can sometimes be incomplete, inaccurate, and duplicated.



Dealing with missing data

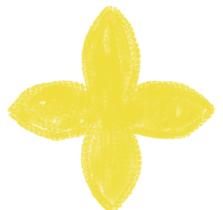




Data visualisation

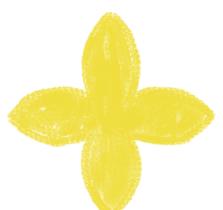
How do we effectively
communicate information?

Why visualise data?



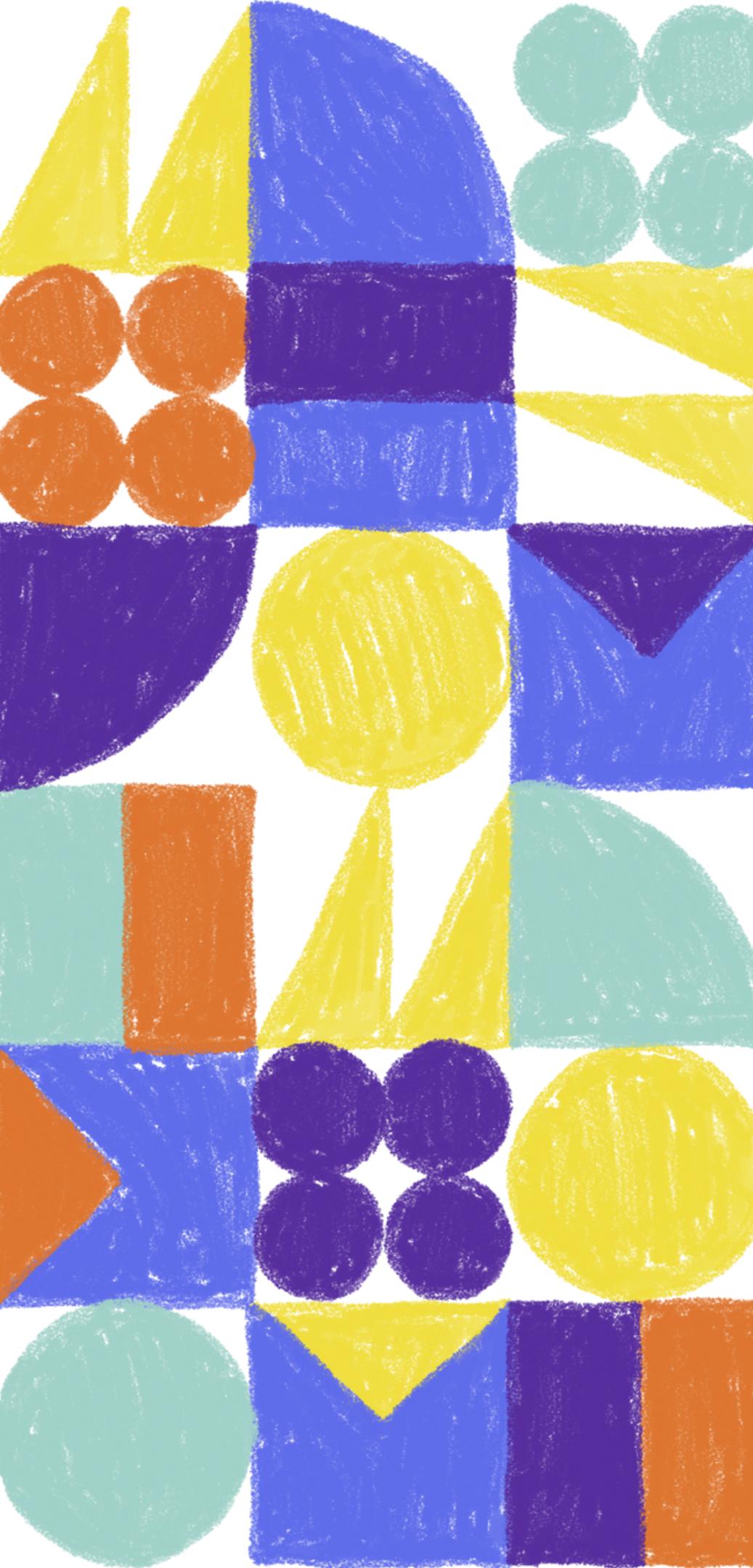
Explorating data

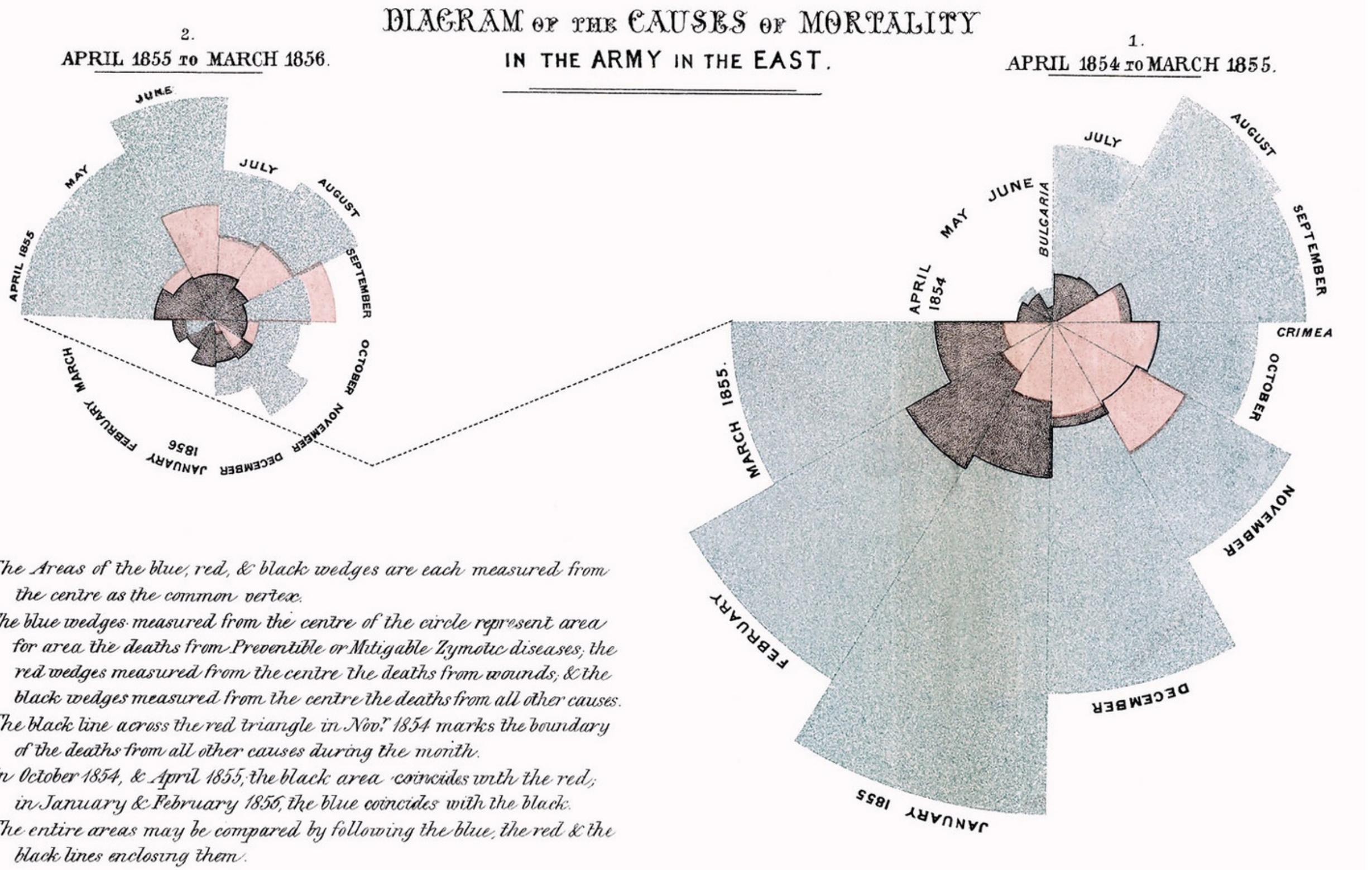
Identifying data issues and modelling approaches.



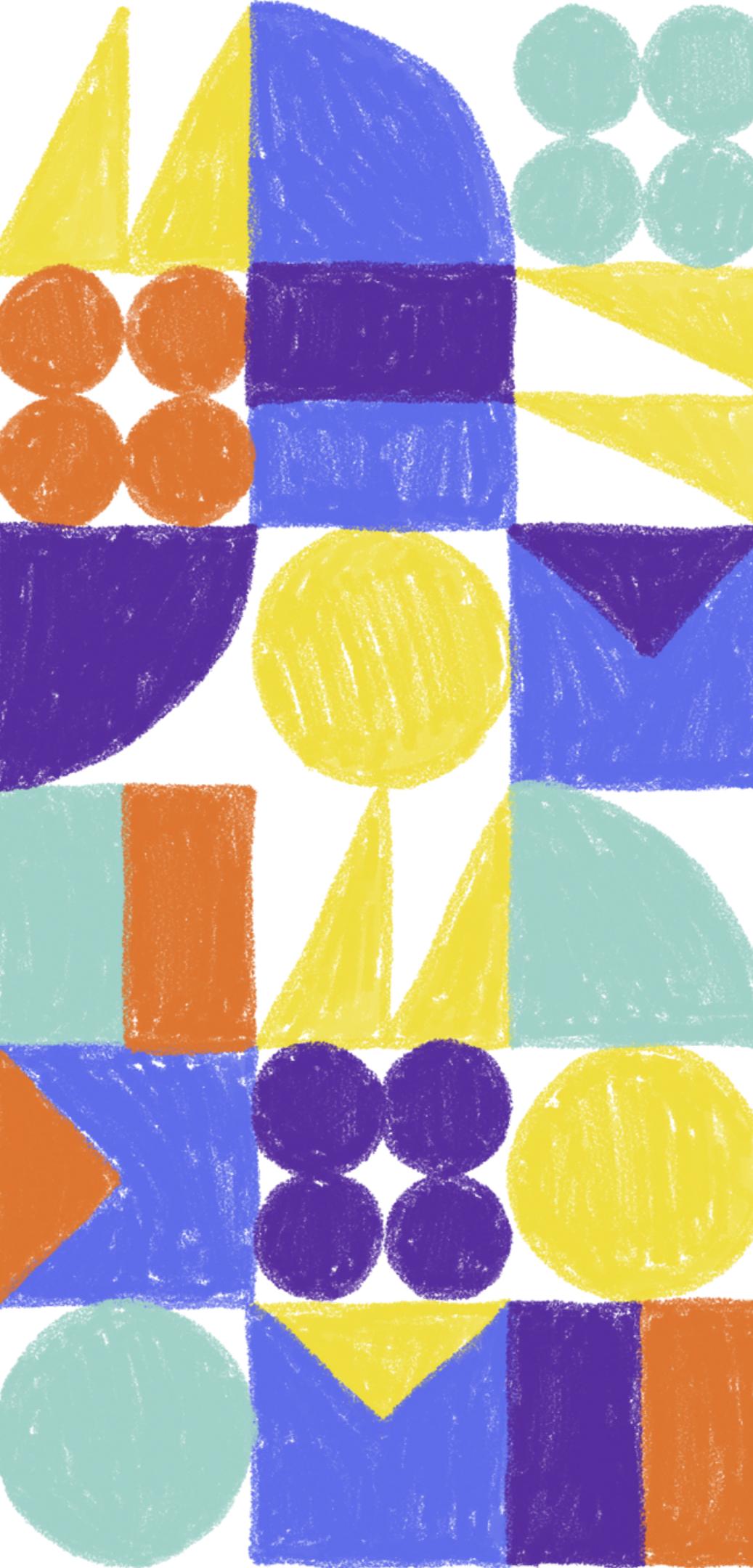
Communicating insights

Telling stories with data.





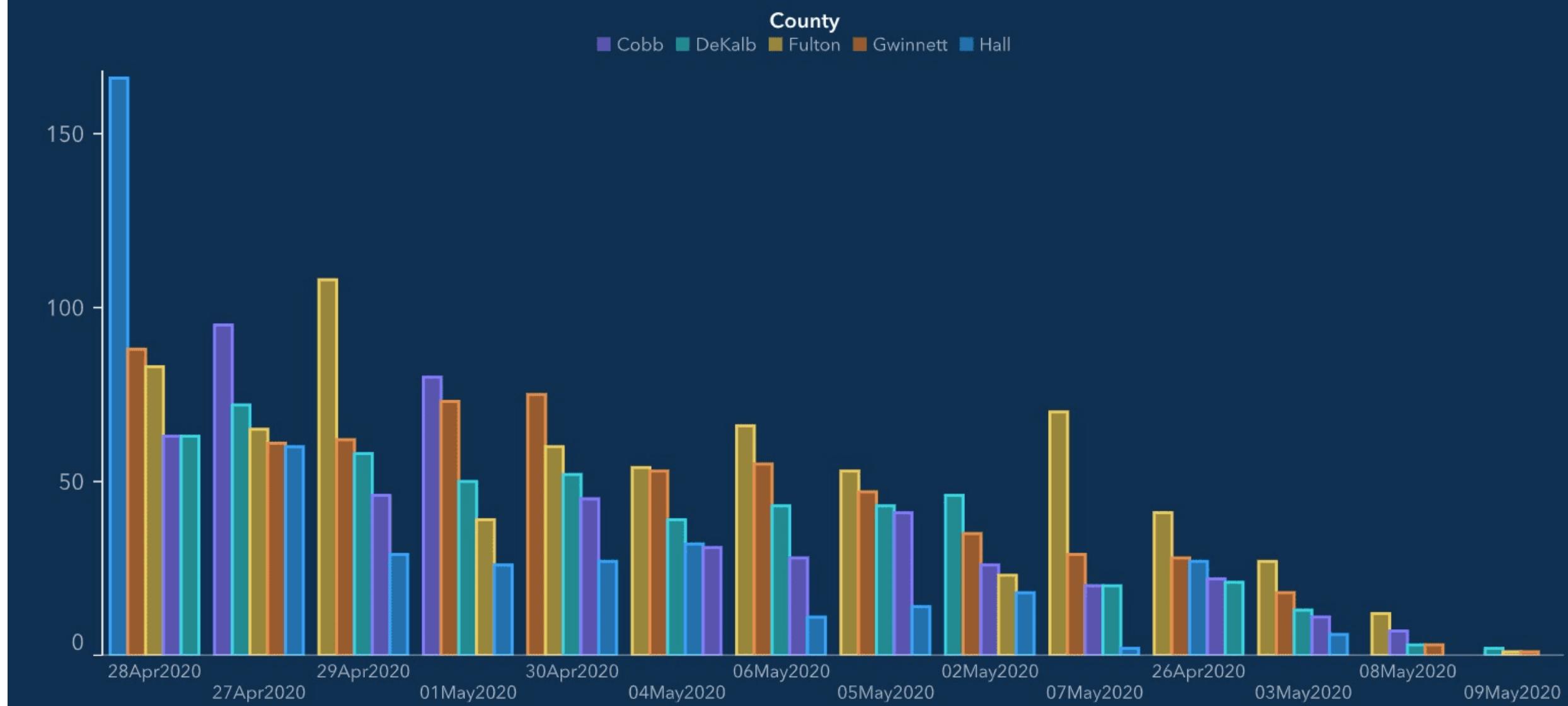
Florence Nightingale's Rose Diagram. 1859.



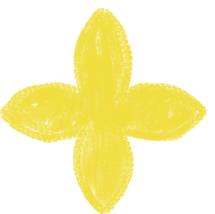
Not all data visualisations are *good* data visualisations...

Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.

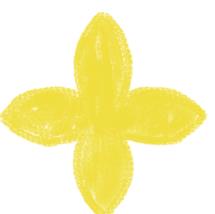


How do we help people make better charts?



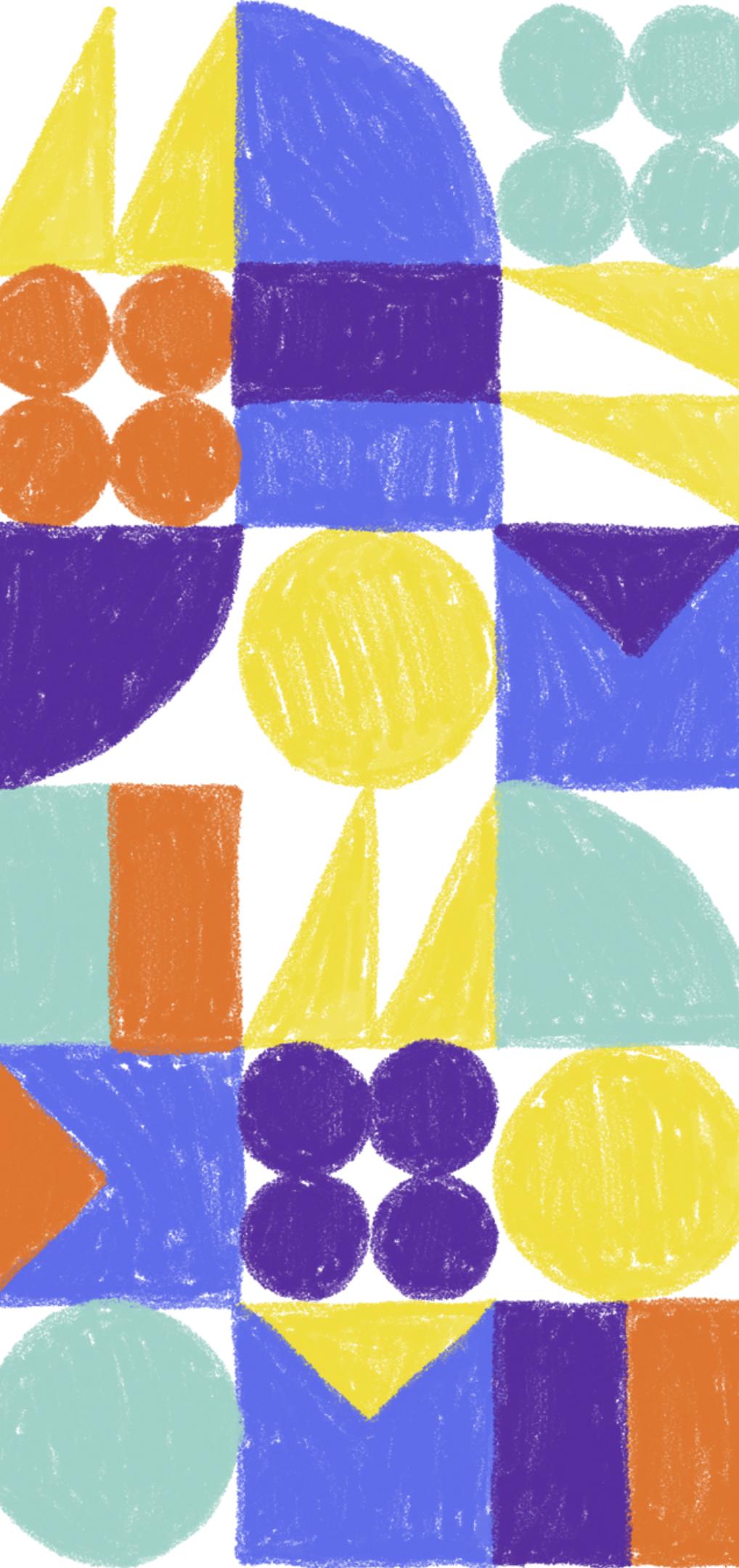
Develop data visualisation guidance

Project with the Royal Statistical Society to develop and publish a website



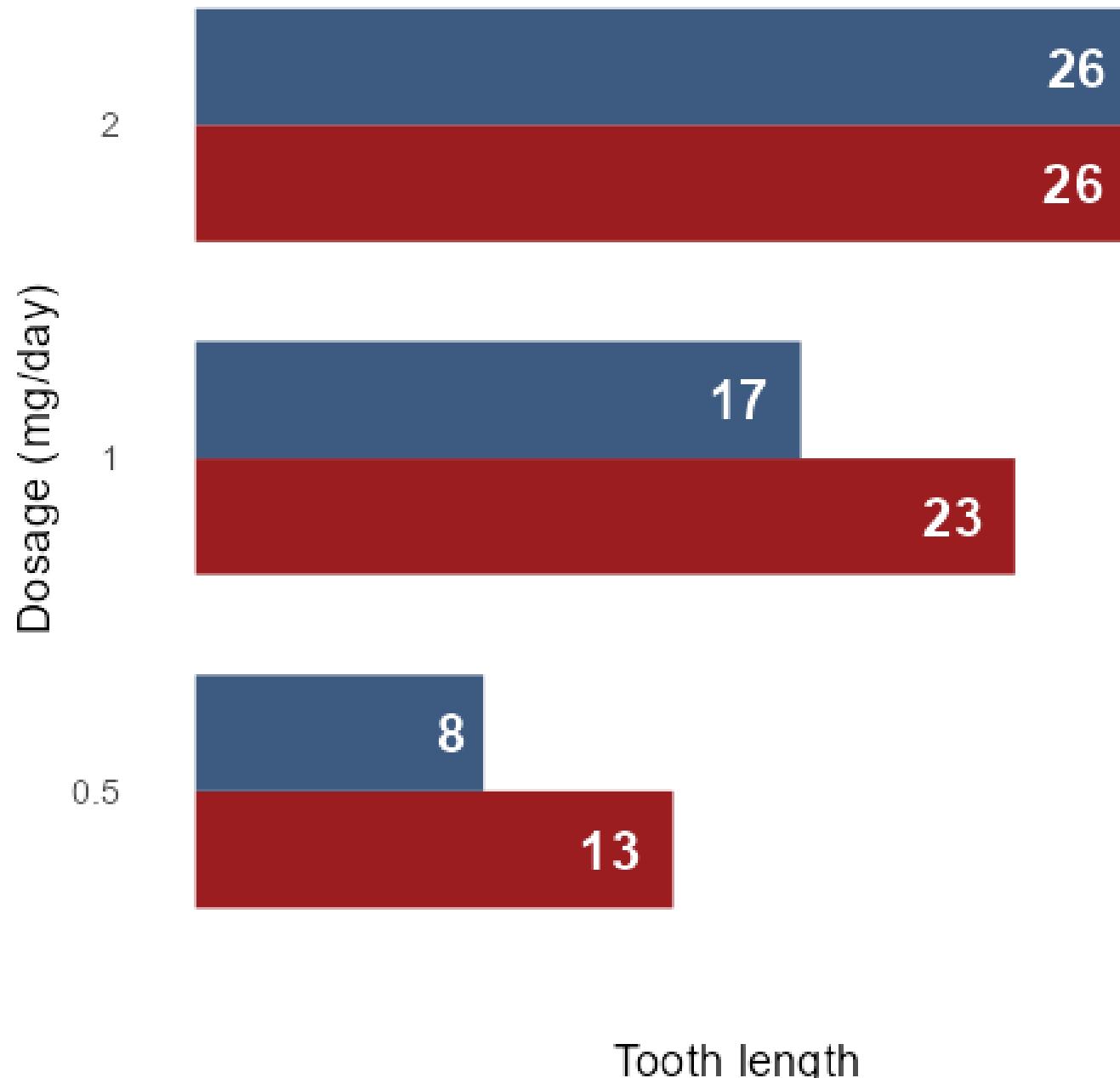
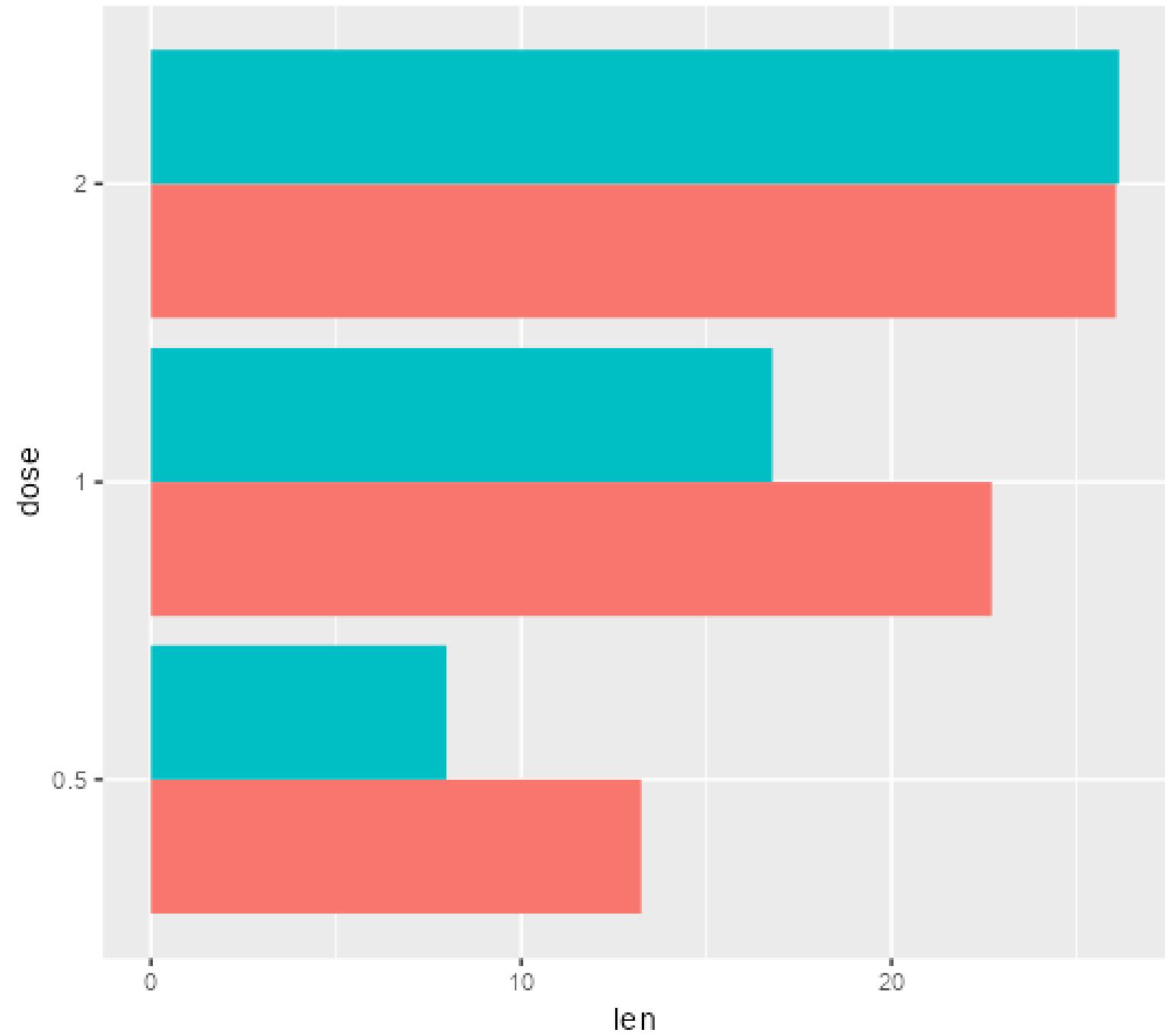
Build software to do it for them

Create better default settings for charts



Tooth Growth

Each of 60 guinea pigs received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods: **orange juice** or **ascorbic acid**.



[Introduction](#)[How to use this guide](#)[The guide](#) ▾[Why we visualise data](#)[Principles and elements of visualisations](#)[Choosing a visualisation type](#)[Styling for accessibility](#)[Styling for RSS publications](#)[References and resources](#)[About the authors](#)[Join us at RSS](#)[Conference](#)[Terms and conditions](#)

Best Practices for Data Visualisation

Insights, advice, and examples (with code) to make data outputs more readable, accessible, and impactful



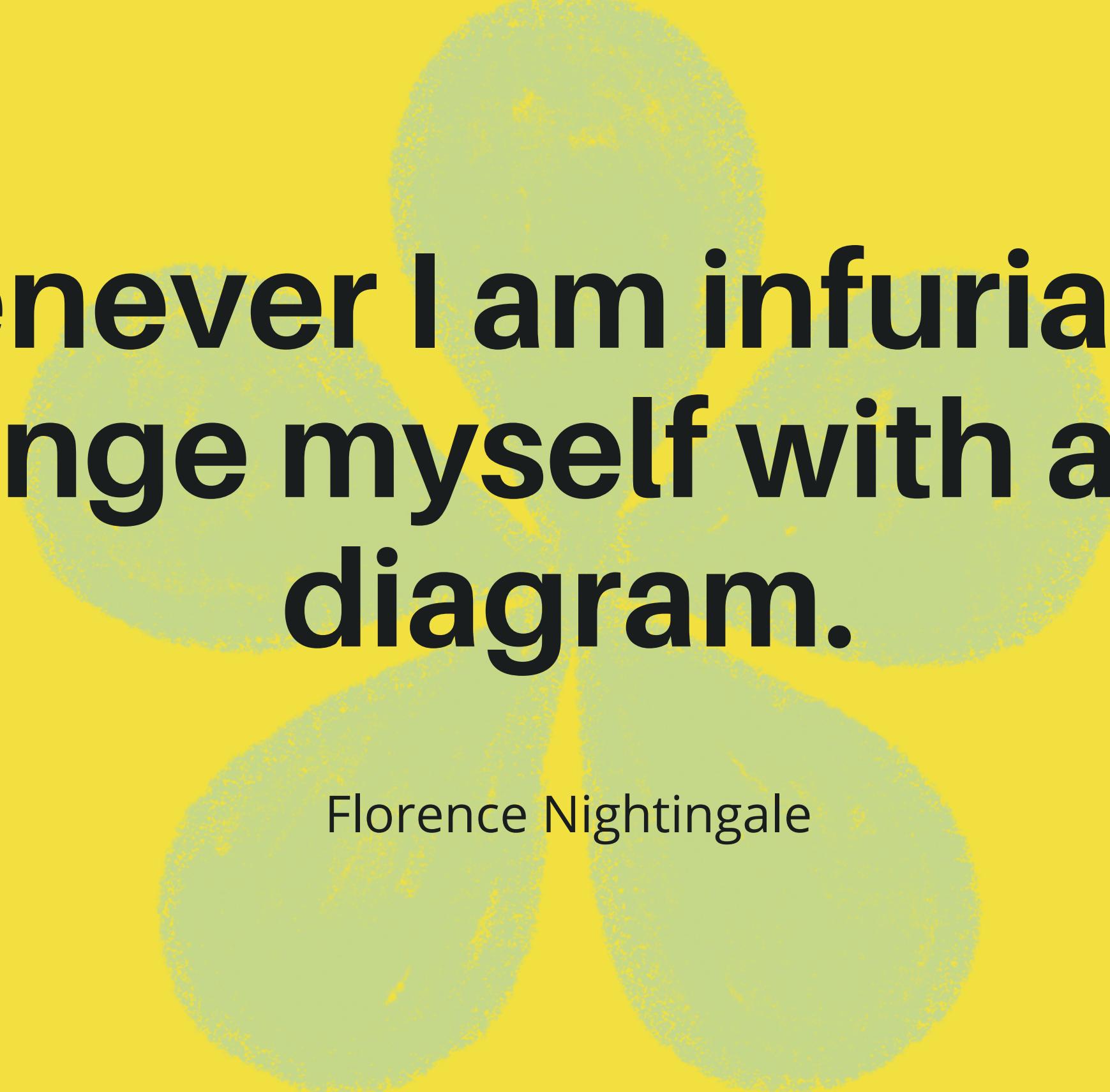
Coming to the RSS Conference this September? Find out more about [our conference session](#) and how to become a guide contributor!

Statistics is “the science of collecting, analyzing, presenting, and interpreting data” ([Williams, Anderson, and Sweeney 2023](#)). Presentation of data is a key means to support and guide interpretation and subsequent decision making. Techniques exist for effective display. This is what this guide is all about.

Good data visualisation requires appreciation and careful consideration of the technical aspects of data presentation. But it also involves a creative element. Authorial choices are made about the “story” we want to tell, and design decisions are driven by the need to convey that story most effectively to our audience. Software systems use default settings for most graphical elements. However, each visualisation has its own story to tell, and so we must actively consider and choose settings for the visualisation under construction.

This guide covers both aspects of data visualisation: the art and the science. It is written primarily for contributors to [Royal Statistical Society publications](#) – chiefly, [Significance magazine](#), the [Journal of the Royal Statistical Society Series A](#), and [Real World Data Science](#) – but we trust you will find the information and advice within to be of broad relevance and use to any data visualisation task.





**Whenever I am infuriated,
revenge myself with a new
diagram.**

Florence Nightingale



n.rennie@lancaster.ac.uk