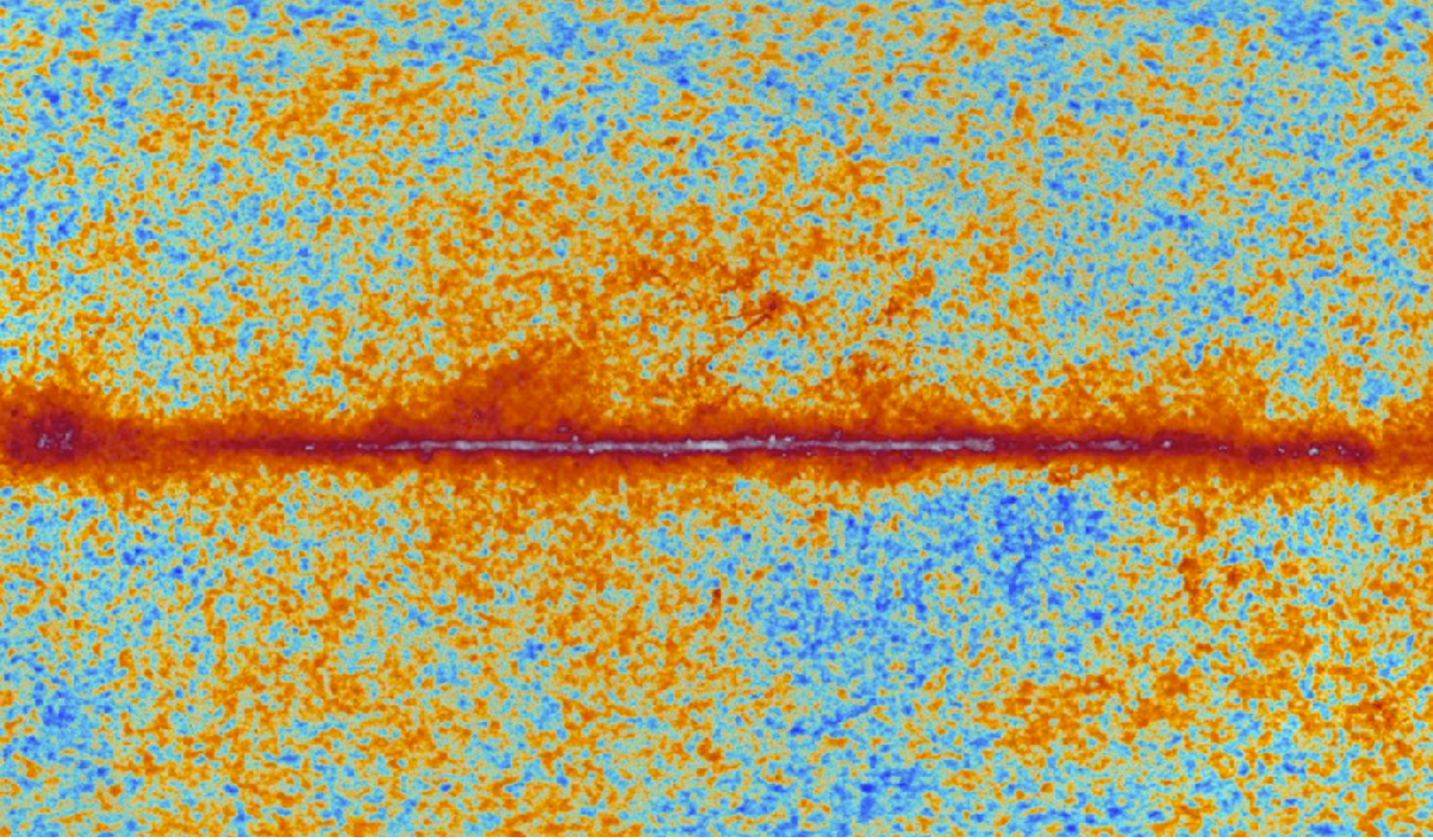


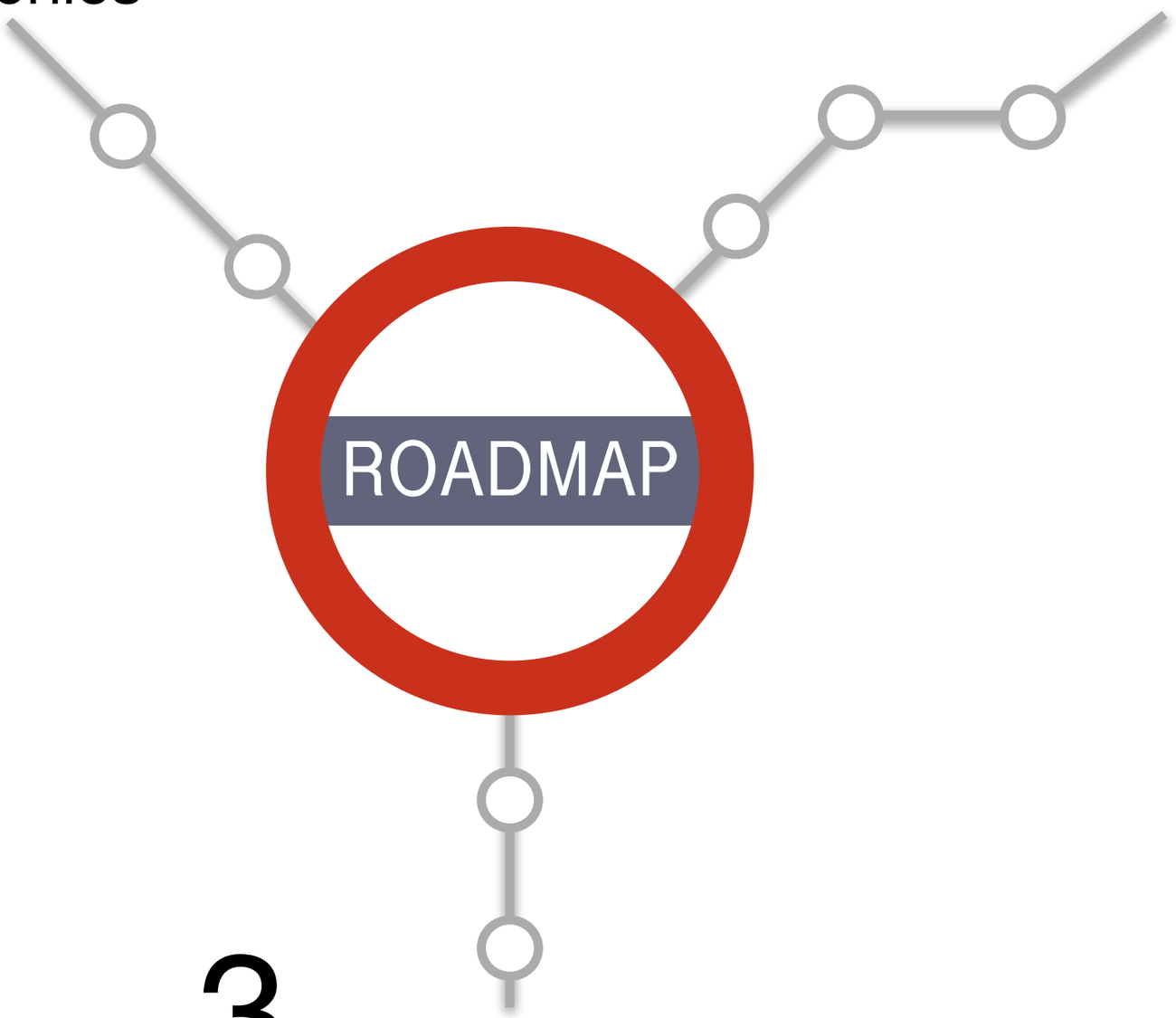
wandering in graphicsland

Raquel H Ribeiro

Senior Data Scientist at Feedzai

1. advice from the grammar of graphics

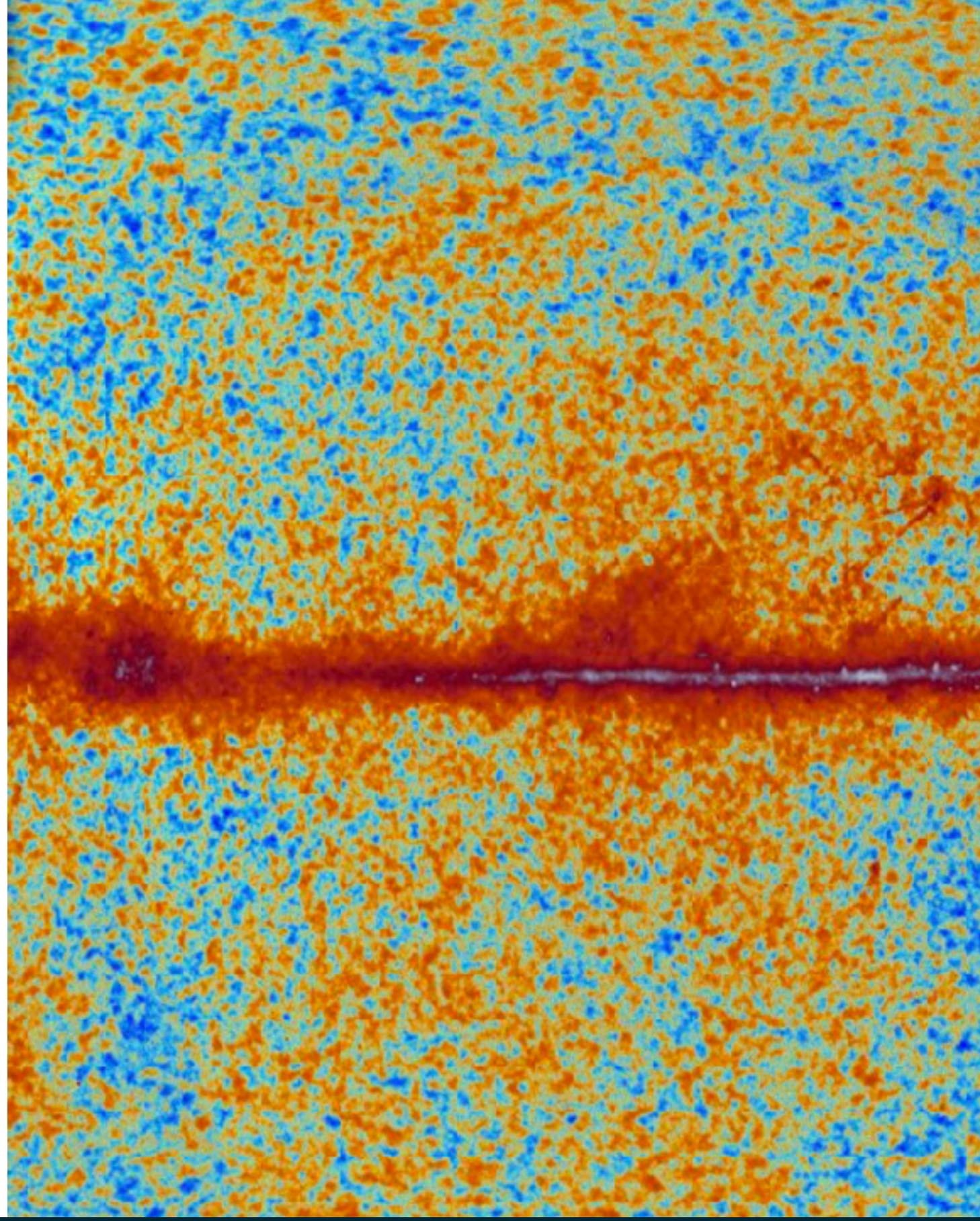
2. layering operations



3. practice session rendering maps

1

advice from the grammar of graphics



graphics are everywhere

But good graphics are hard to find. If you need more than 10s to understand what the basic message the plot is conveying, it is not a good plot.

Usually good graphics make transparent a clear interpretation and organisation of data insights. Good graphics should be an intrinsic part of the data science pipeline.

Good graphics don't start with

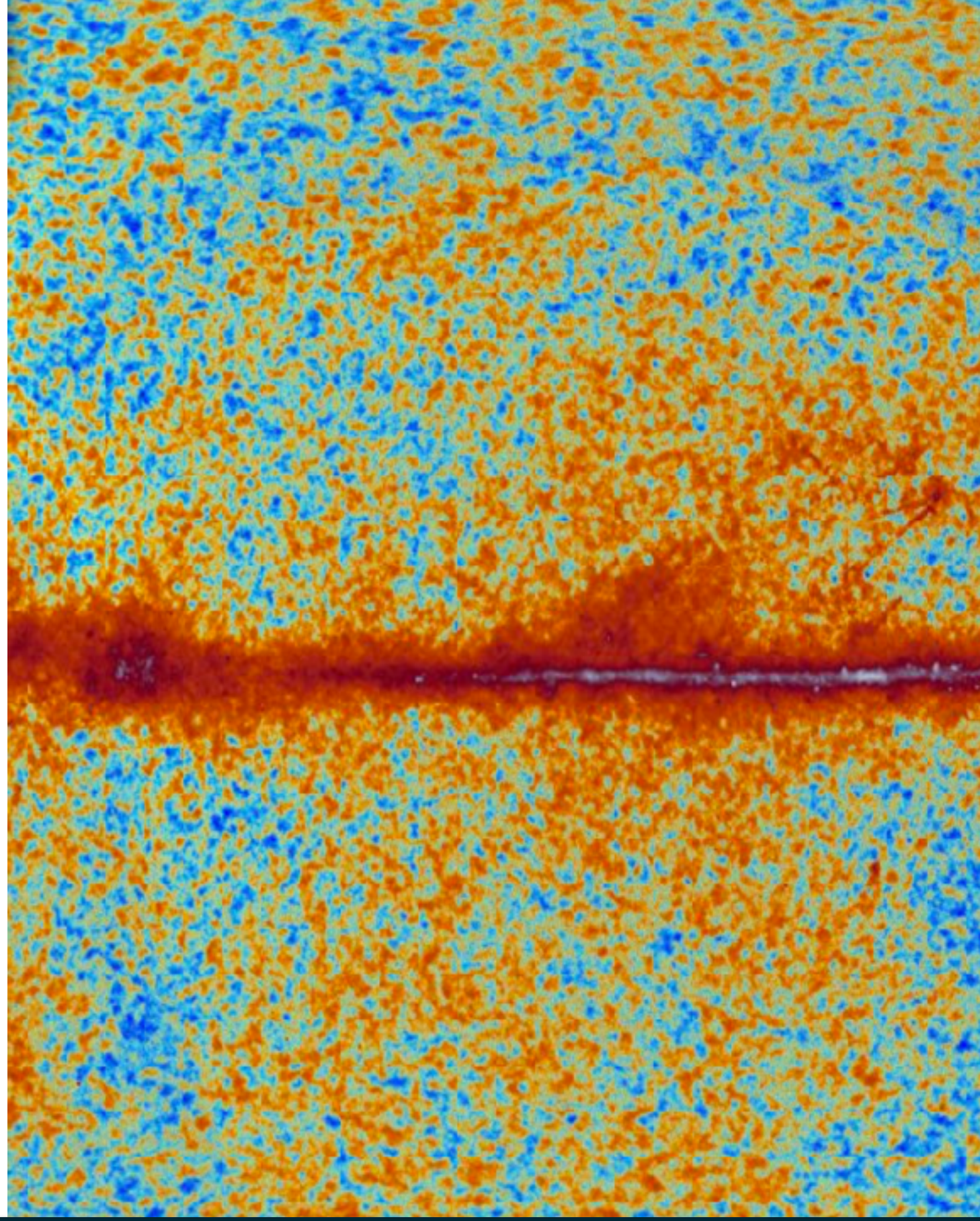
```
sudo make me a plot
```

They start with

```
let's focus on the data
```

2

layering



meet grammar of graphics

“An abstraction which makes thinking, reasoning and communicating graphics easier”.

Hadley Wickham

tidyverse

R packages for data science

The tidyverse is an opinionated **collection of R packages** designed for data science. All packages share an underlying design philosophy, grammar, and data structures.

Install the complete tidyverse with:

```
install.packages("tidyverse")
```

https://www.tidyverse.org/

you can contribute to tidyverse

Don't contribute code

The easiest way to contribute to an open source package is not to contribute code at all. Find a typo in the documentation, add a reproducible example to an open issue without one, post a solution to a question in an issue, on twitter or stackoverflow. These types of contributions are among the easiest things for maintainers to review and accept, so it is a great place to start getting used to the contribution process.

https://www.tidyverse.org/

let’s switch to RStudio

What is the story that data is telling you?

merge

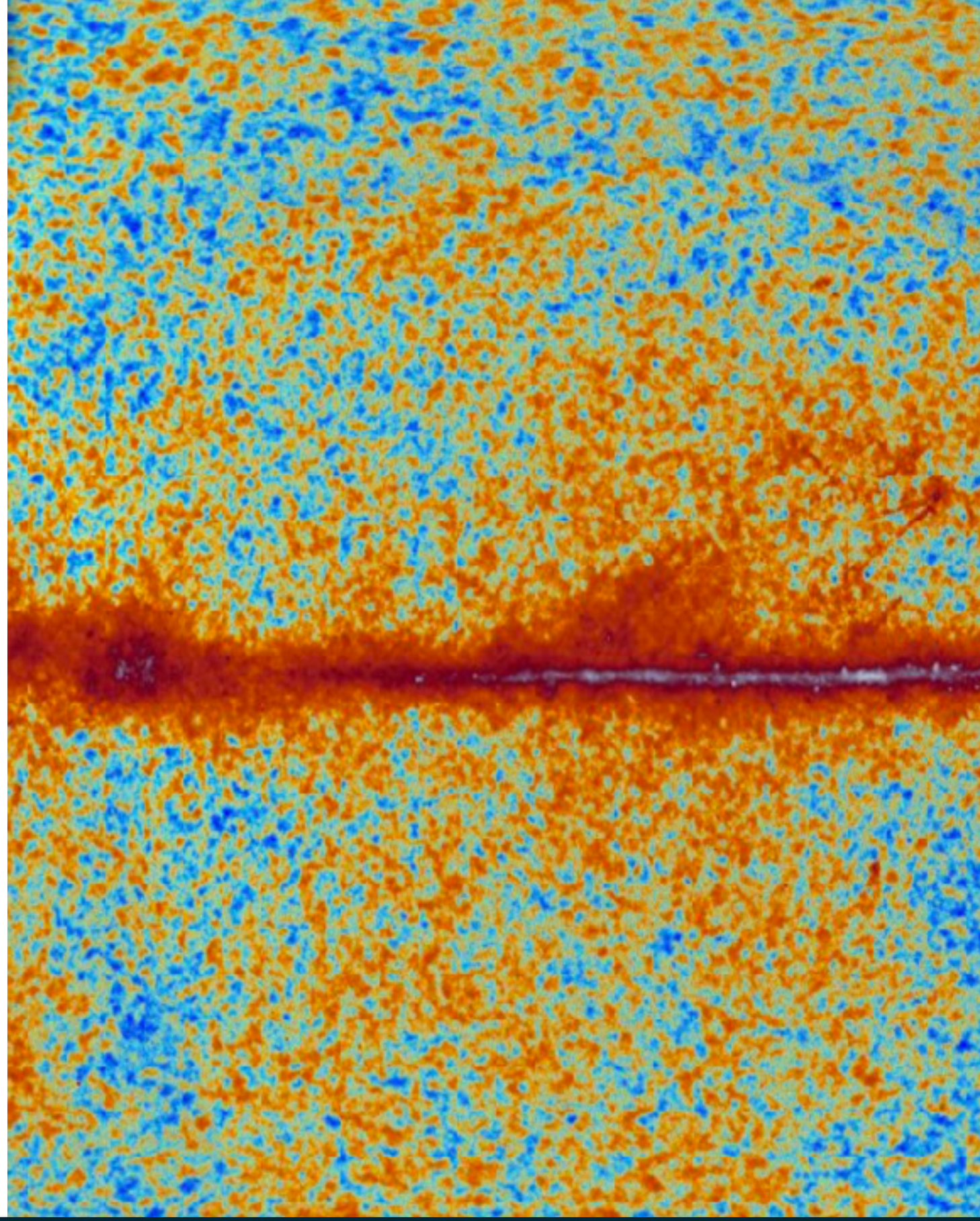
melt

filter

group by

3

practice session
rendering maps



Raquel H Ribeiro

need ideas?

Population data by US state and council: https://simplemaps.com/data/us-cities

Medical store geospatial challenge: http://databits.io/challenges/medical-store-geospatial-challenge

Data.gov: https://tinyurl.com/ya8xeor4