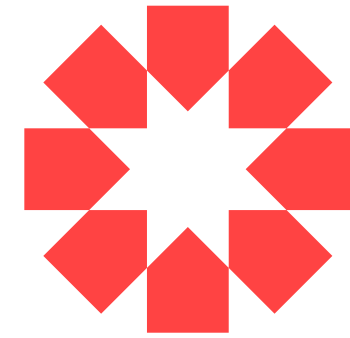


An intro to
Data Visualisation

Data Visualisation · Gyan Lakhwani · January 2025



About me



Co-founder

Public Knowledge Design Studio



Product Designer

Microsoft

IDC

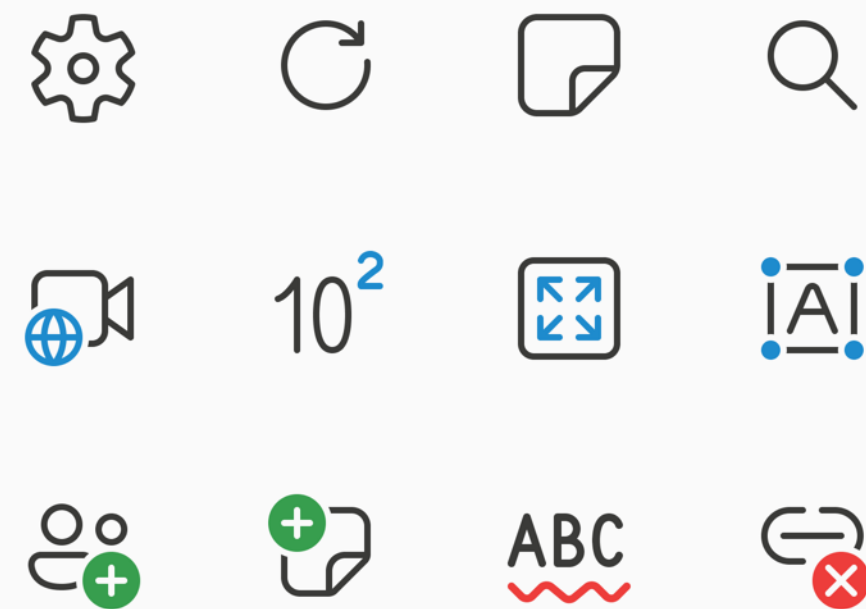
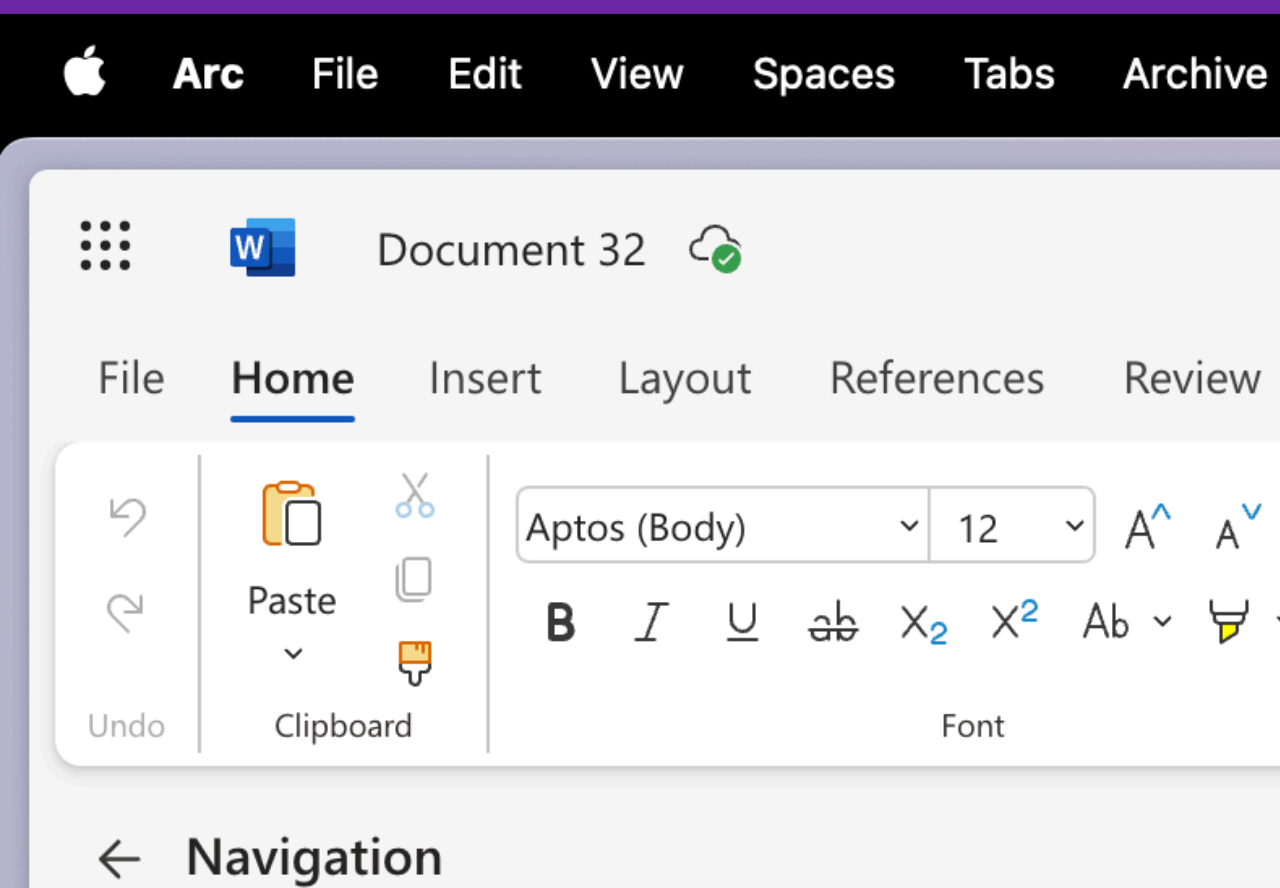
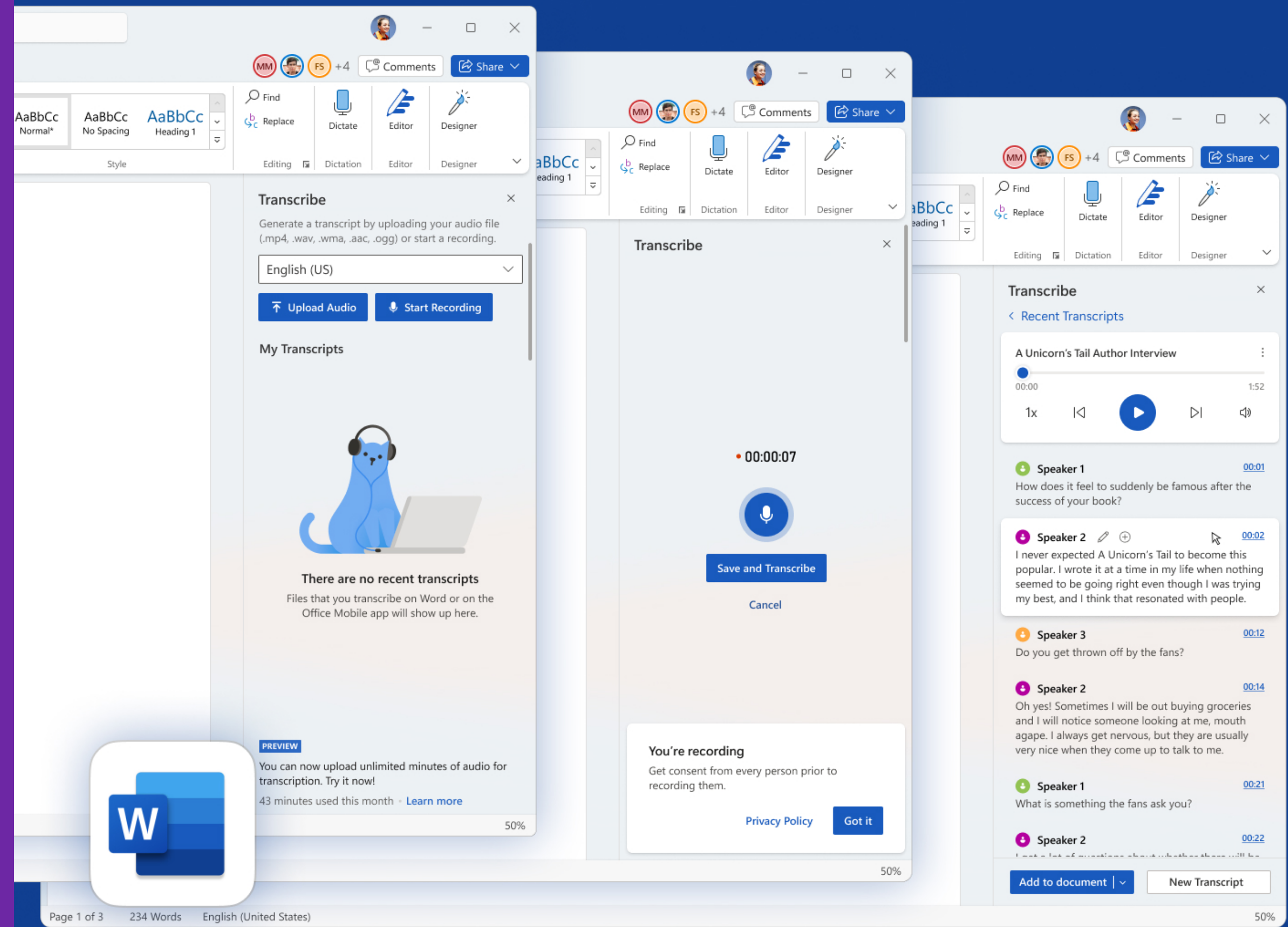
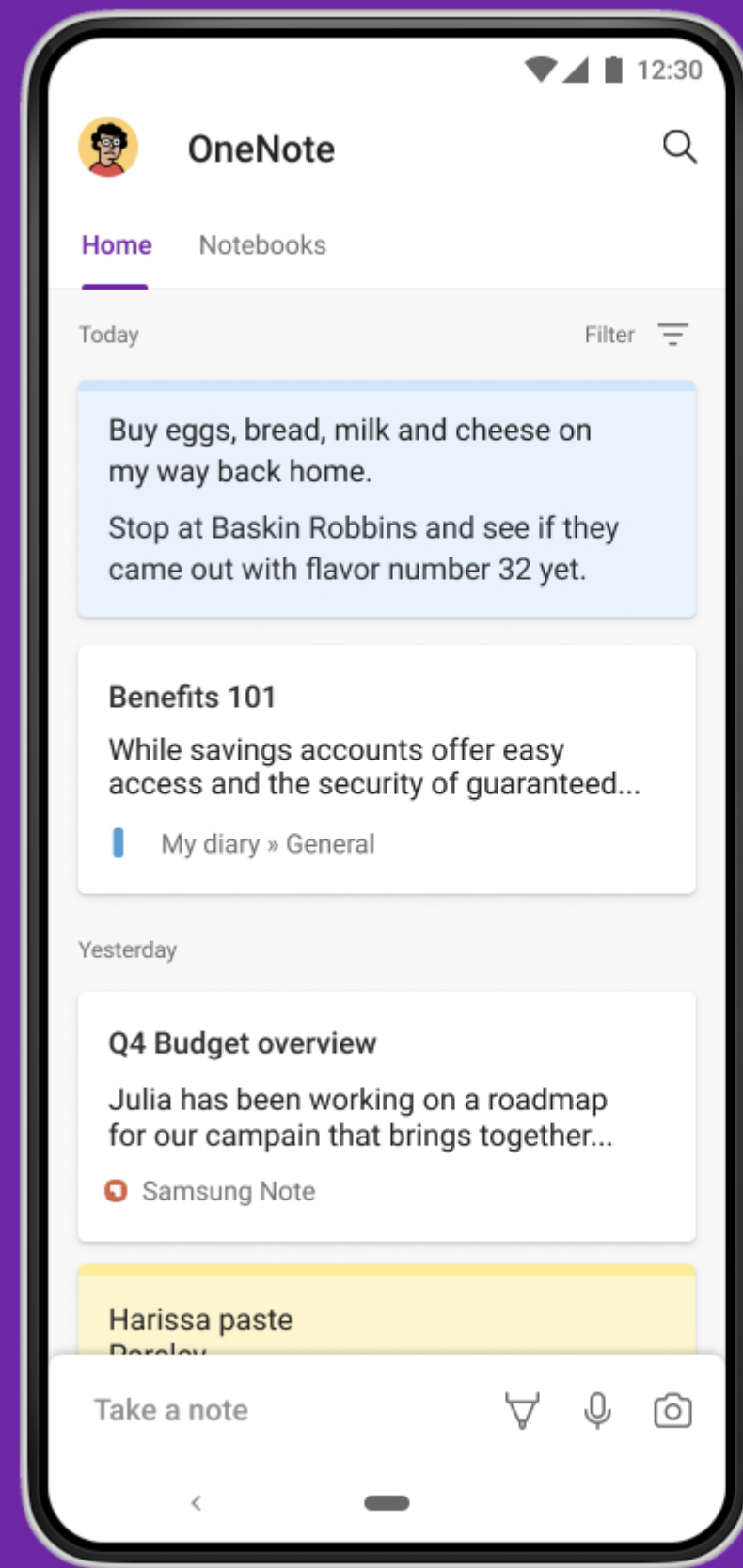
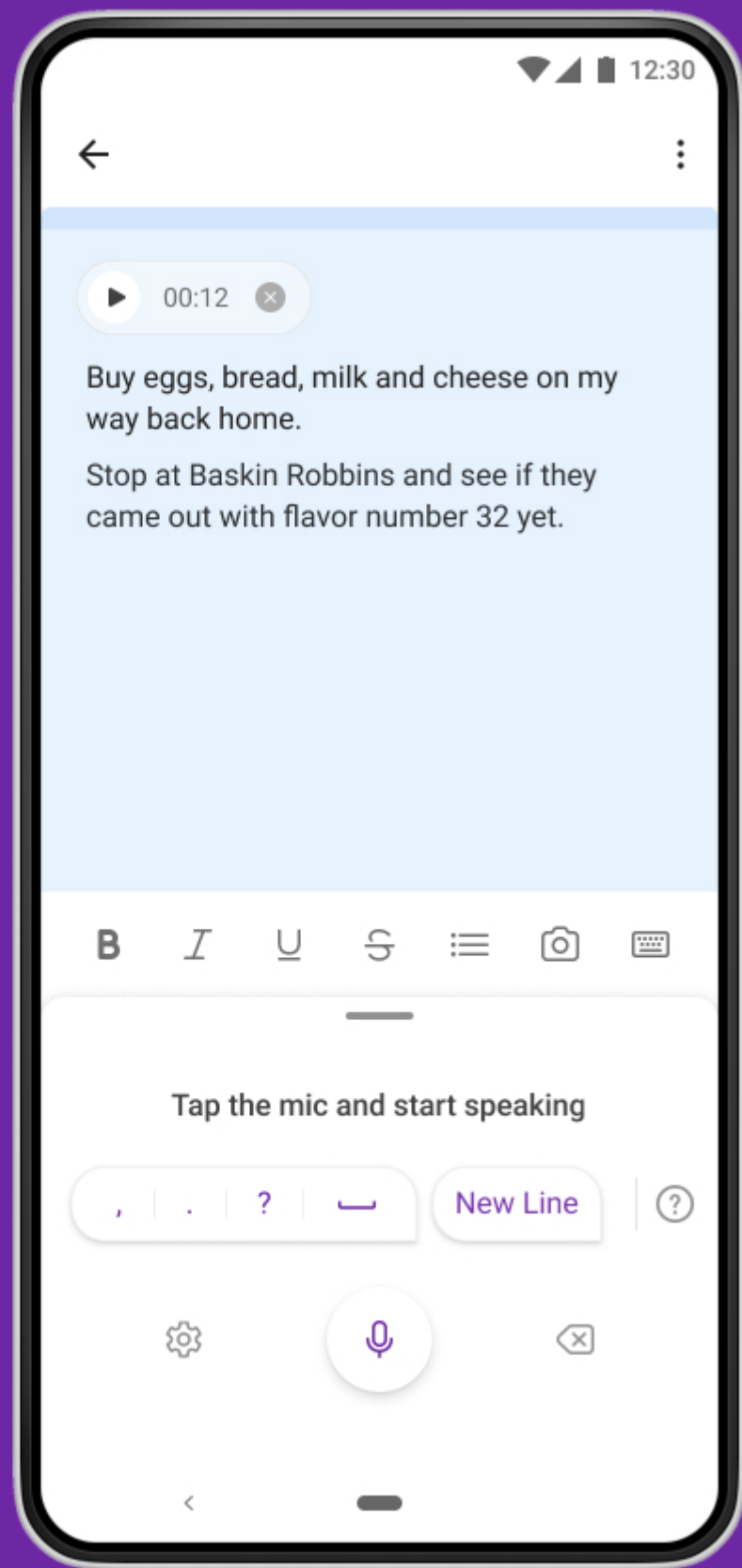
M.Des. (Interaction Design)

IDC School of Design, IIT Bombay



B.Tech. (Information Technology)

MAIT, IP University



LIGHT

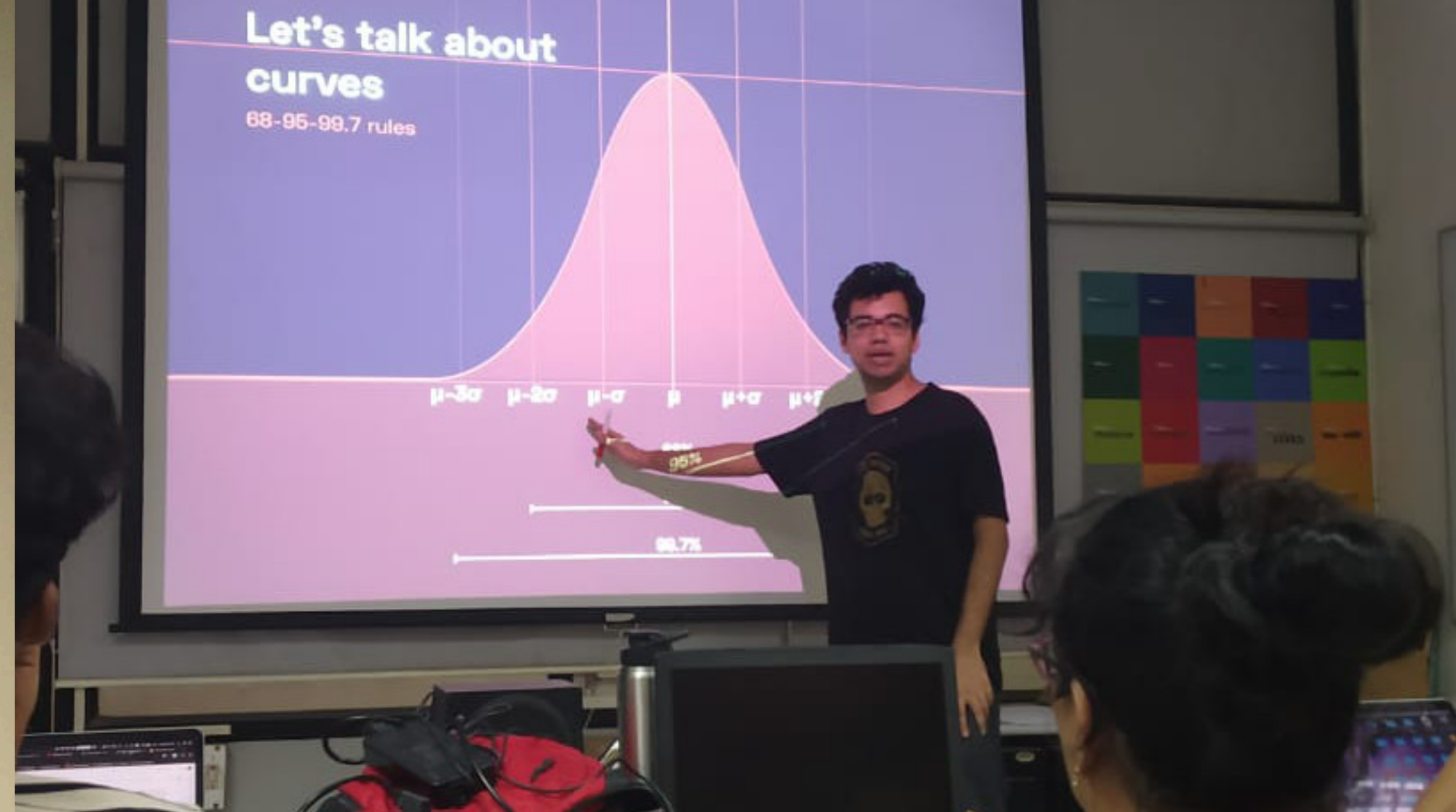


HIGH CONTRAST



DARK







Instagram browser interface showing the profile of Gyan Lakhwani (@gyanl).

Instagram

gyanl Edit profile

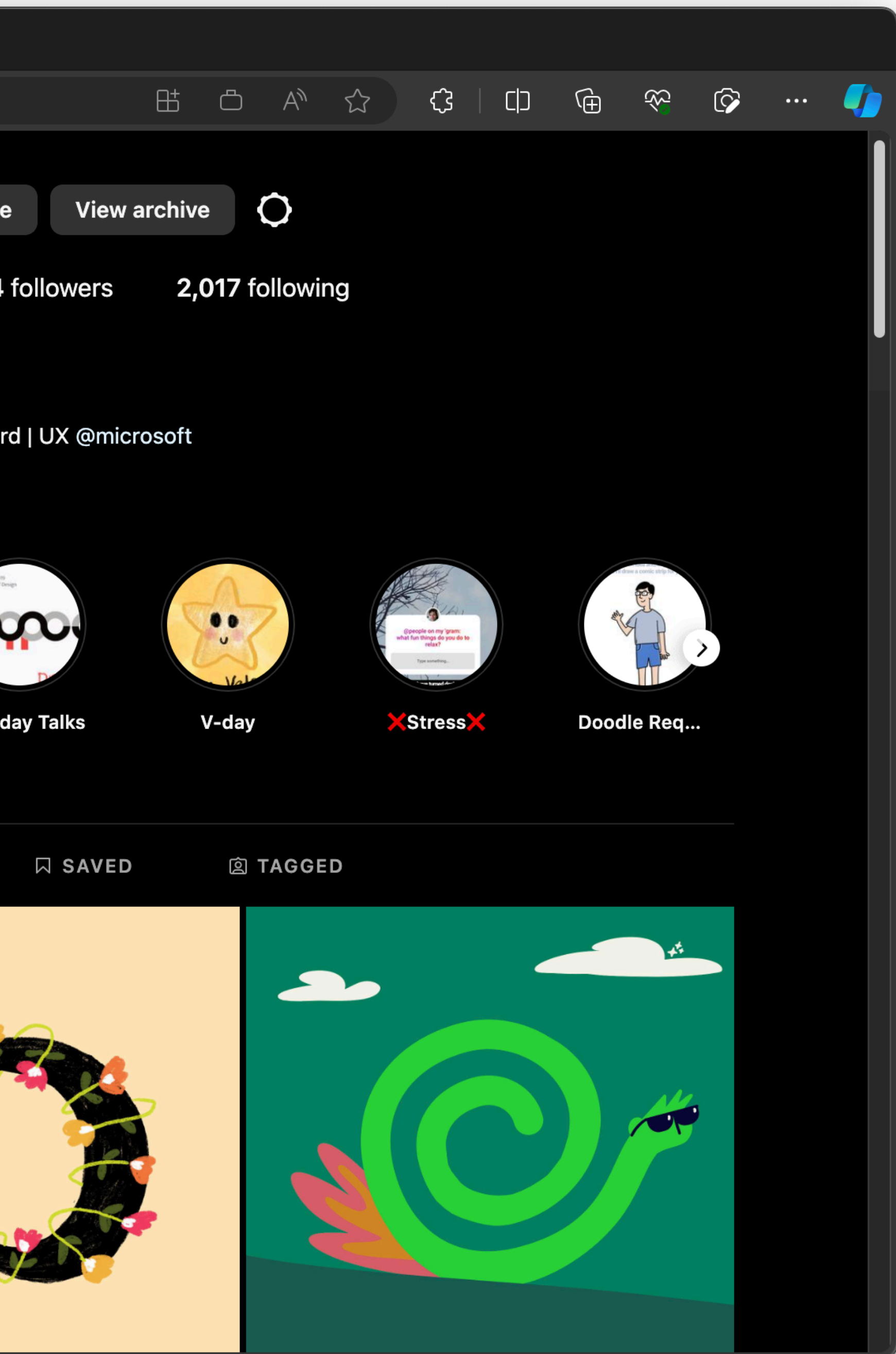
373 posts 1,174 followers

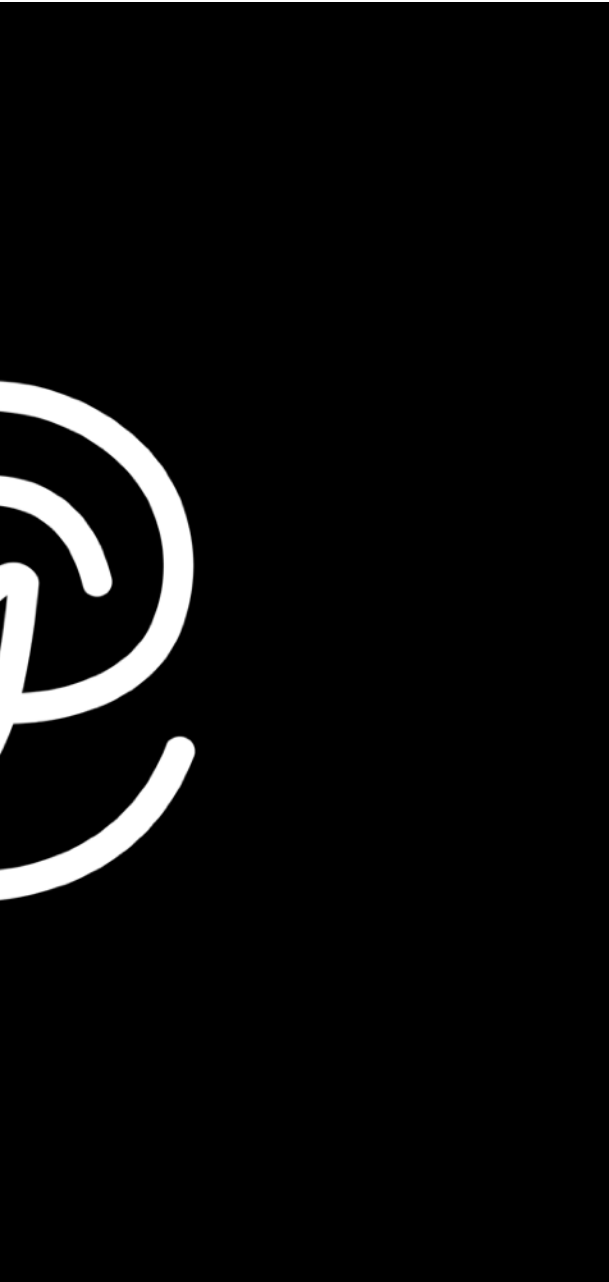
Gyan Lakhwani
Designer, illustrator, nerd | UX @
www.gyanl.com

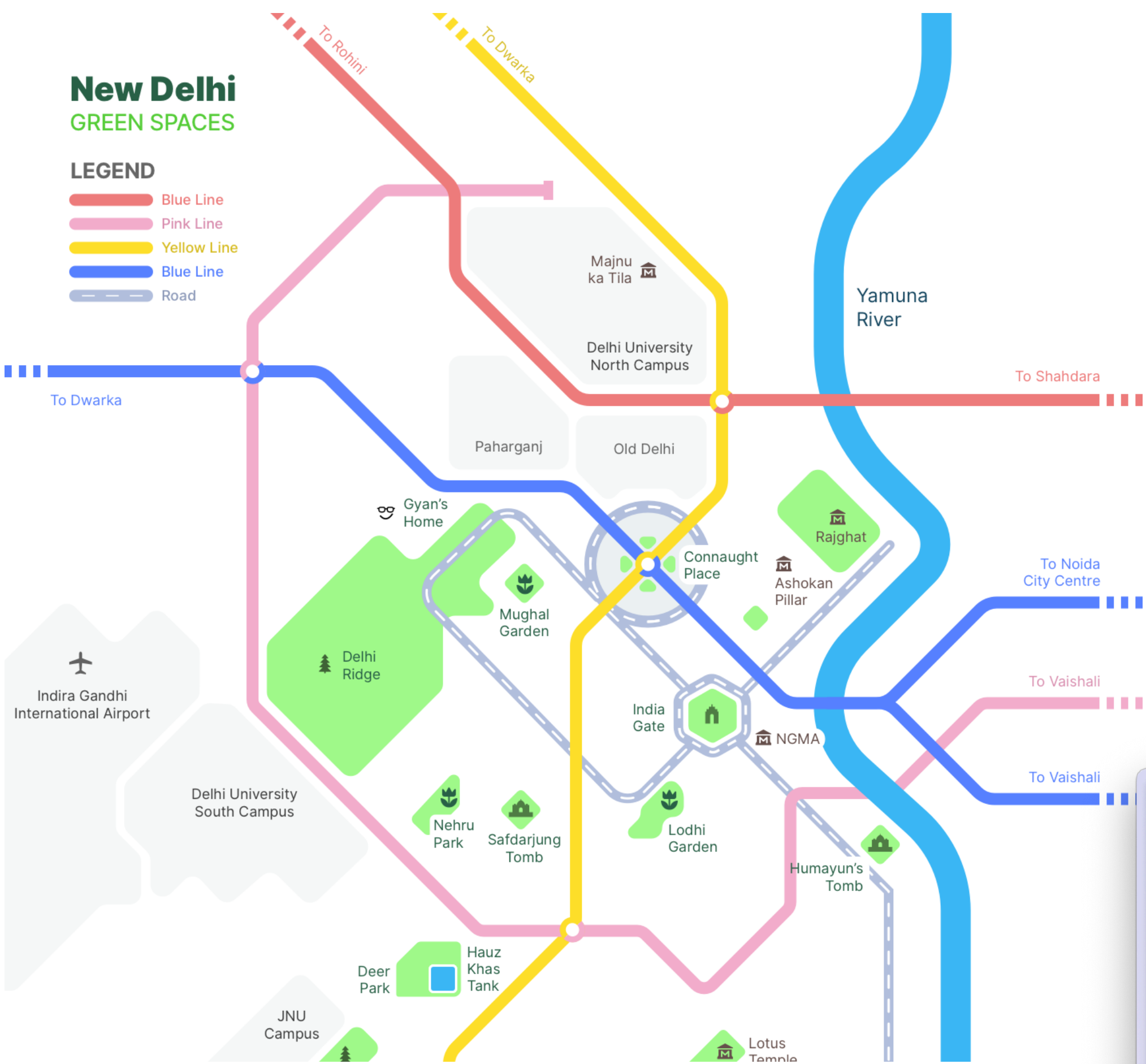
Home Search Explore Reels Messages Notifications Create Profile Threads More

Draw Reqs Summer Flo... Screenshots Typoday Talks

POSTS REELS SA







Mapping your contributions

It is possible to use Leaflet to map your contributions on an embeddable map. As an example, I mapped my classmate Gauri's contributions here:

Gauri's contributions



Exemplar Based Experience Transfer
 Paridhi Maheshwari

INDIA HCI 2019

Banners are everywhere

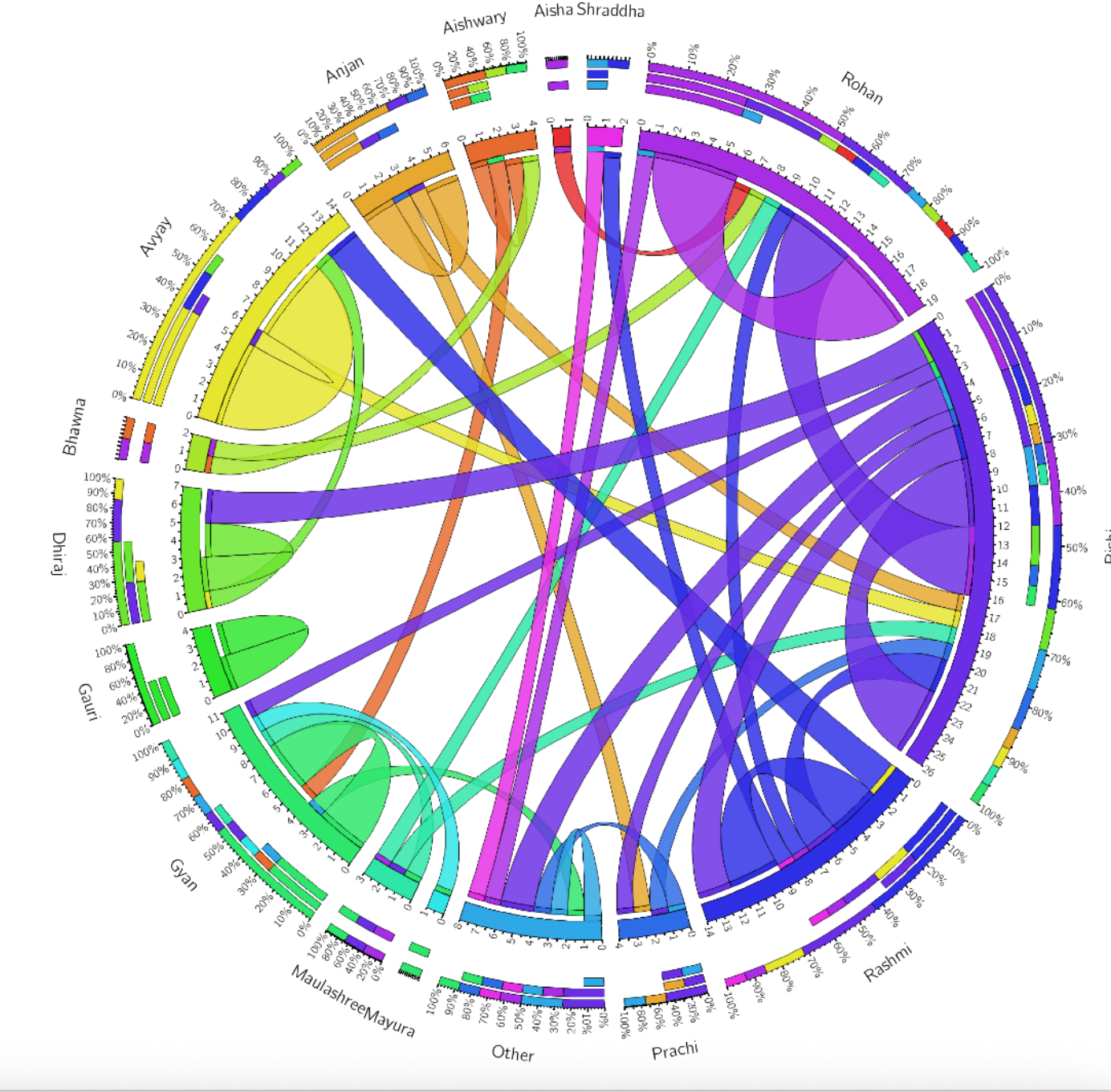
single page graphics

EXTRACT

Text
 Image
 Shape

at a pixel level.

What if the template could be extracted?



HAMMOCK

Continents Gender Awards Language

Search Feeling lucky

100+ results

Gabriela Ybarra
 Biscay, Spain

Author of
 The Dinner Guest

Thomas Hardy
 Dorset, UK

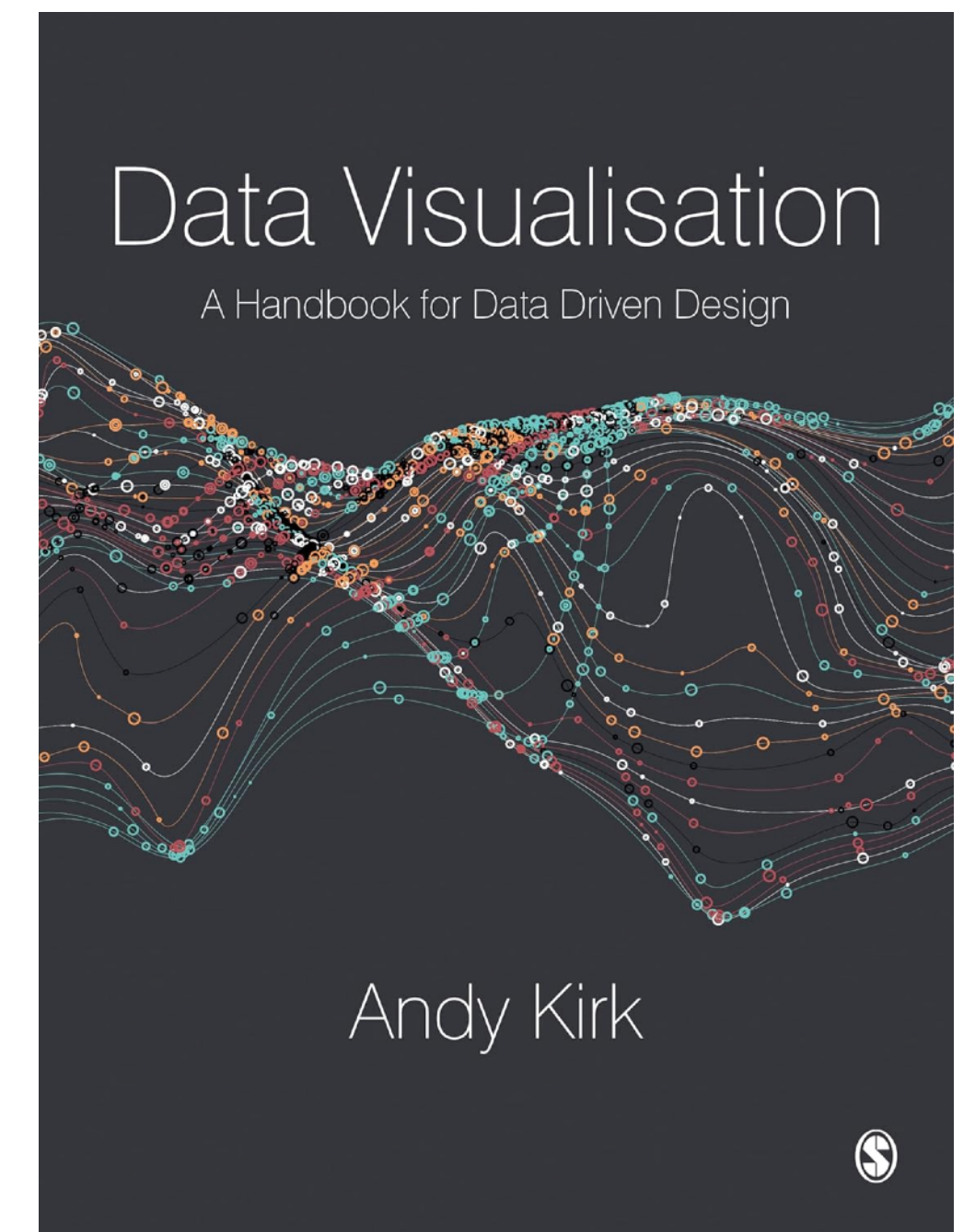
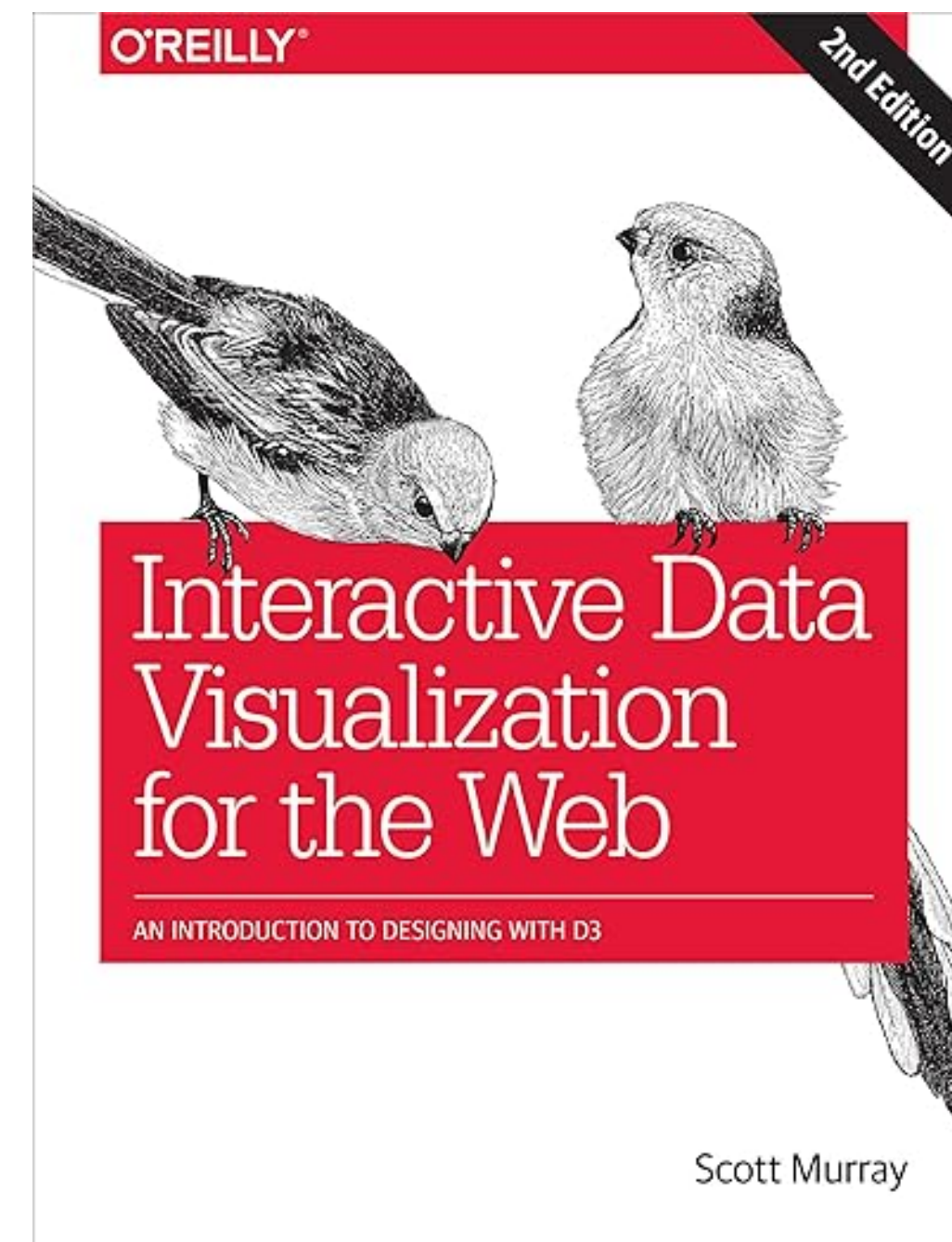
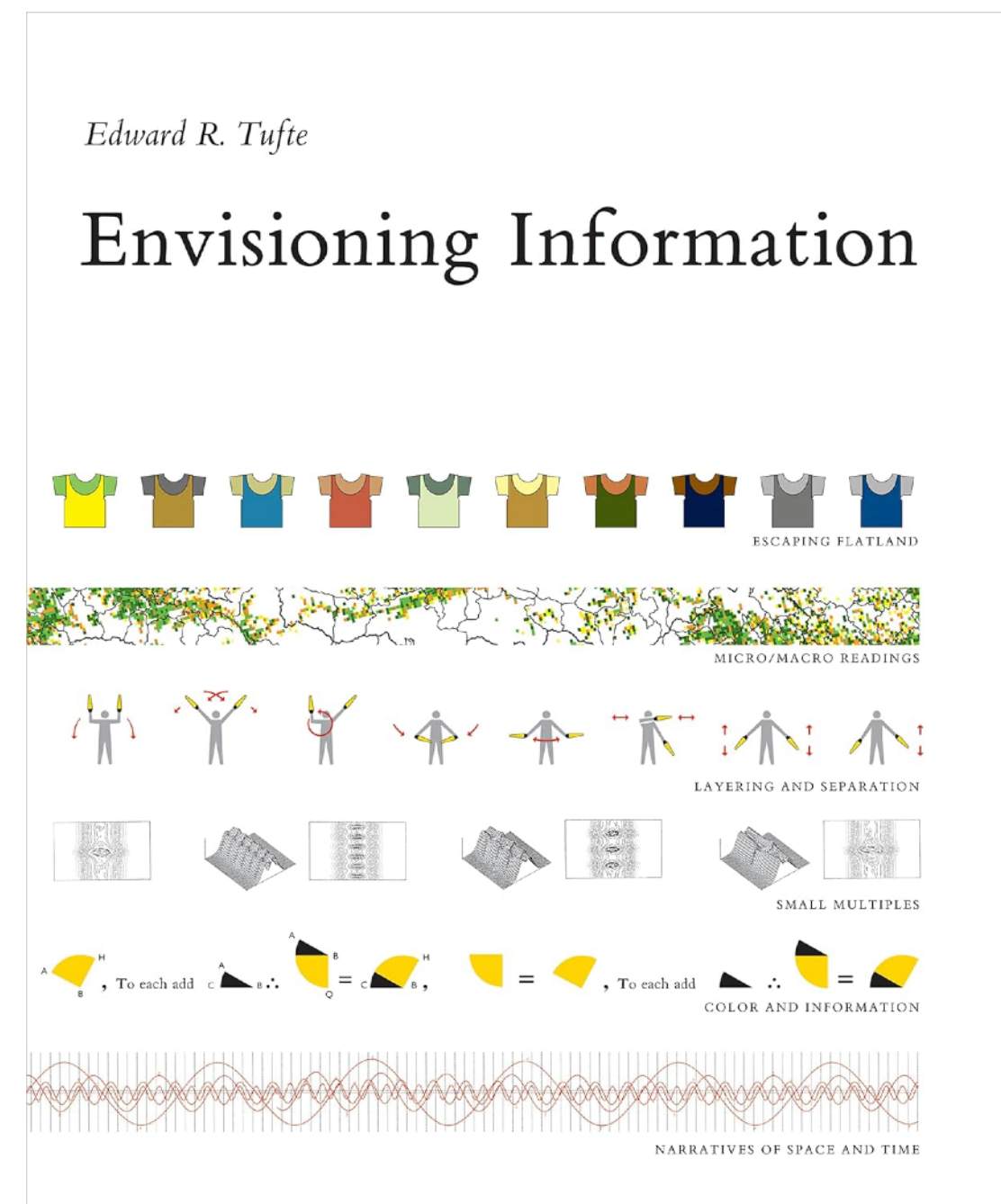
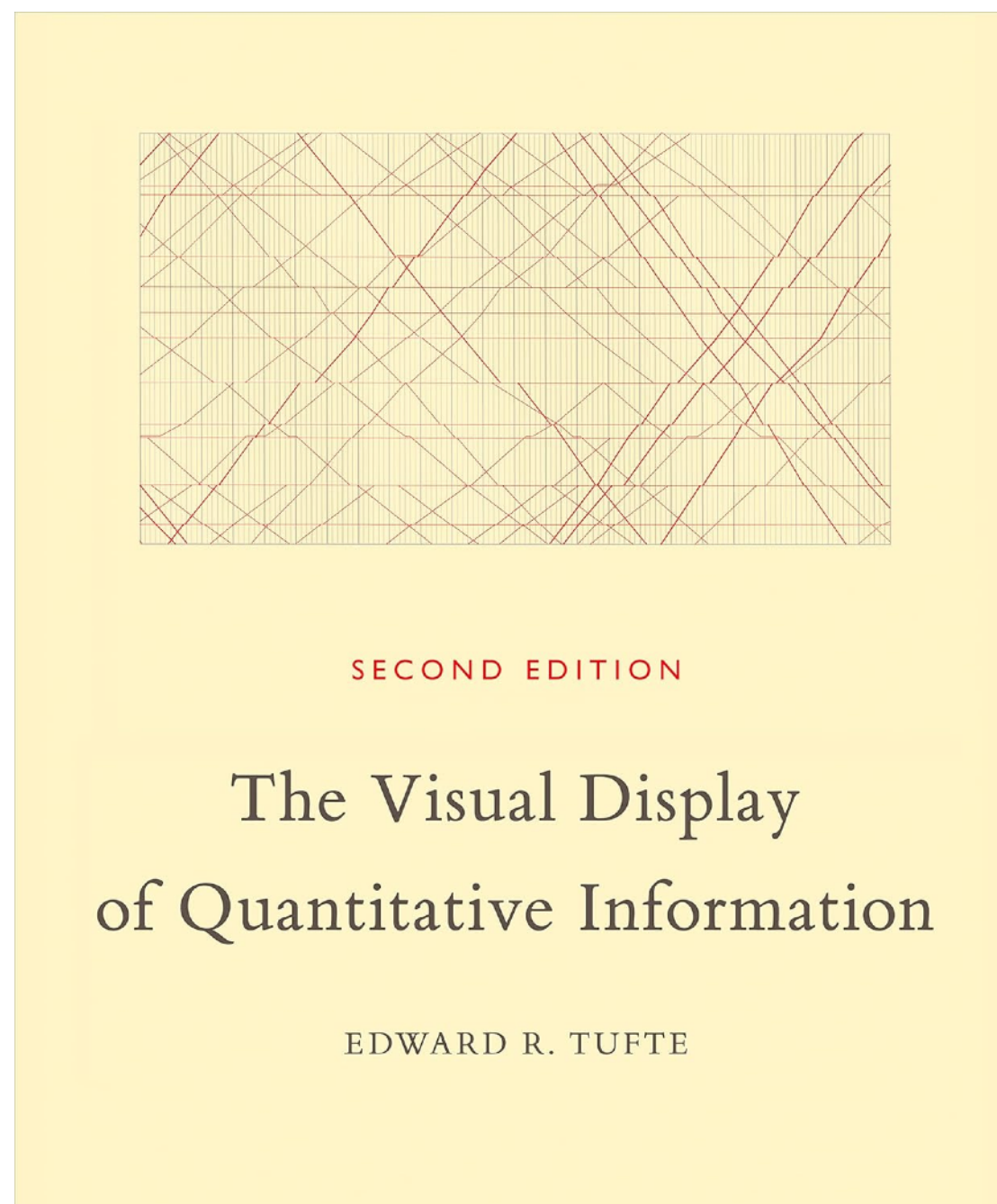
Author of
 Tess of D'Urbervilles
 Jude the Obscure

William Golding
 Cornwall, UK

Author of
 Lord of the Flies
 The Inheritors

A world map showing green circular markers with numbers, representing data points across various countries and regions.

Recommended Reading



Introductions

Tell us a little bit about yourself



- What's your name?
- What did you have for breakfast today?
- What's your date of birth?
- Where do you live currently? (Area)
- Which city/town were you born in?
- What's your favourite film?

**“You can’t improve what you don’t
measure.”**

— Peter Drucker, Austrian American management consultant

Data

Data

[23, 60, 44, 28, 11]

"25°C"

33 cm

Orange, Red, Purple

Gyan

Data

Raw, unorganized data is a collection of numbers, symbols, or text. Data can lack context when viewed individually.

Data

Raw, unorganized data is a collection of numbers, symbols, or text. Data can lack context when viewed individually.

- Unstructured
- Represents reality but without explanation
- May be quantitative or qualitative

Data

Raw, unorganized data is a collection of numbers, symbols, or text. Data can lack context when viewed individually.

Information

Information

“The current temperature in Delhi is 10°C.”

“Total sales for last month were 500 units.”

Information

- Structured and contextualized
- Provides meaning to data
- Useful for decision-making at a basic level

“The current temperature in Delhi is 10°C.”

“Total sales for last month were 500 units.”

Data

Raw, unorganized data is a collection of numbers, symbols, or text. Data can lack context when viewed individually.

Information

Processed data that has been organized, structured, or refined to provide meaning and context. Useful for decision-making, analysis, and communication.

How?

Data



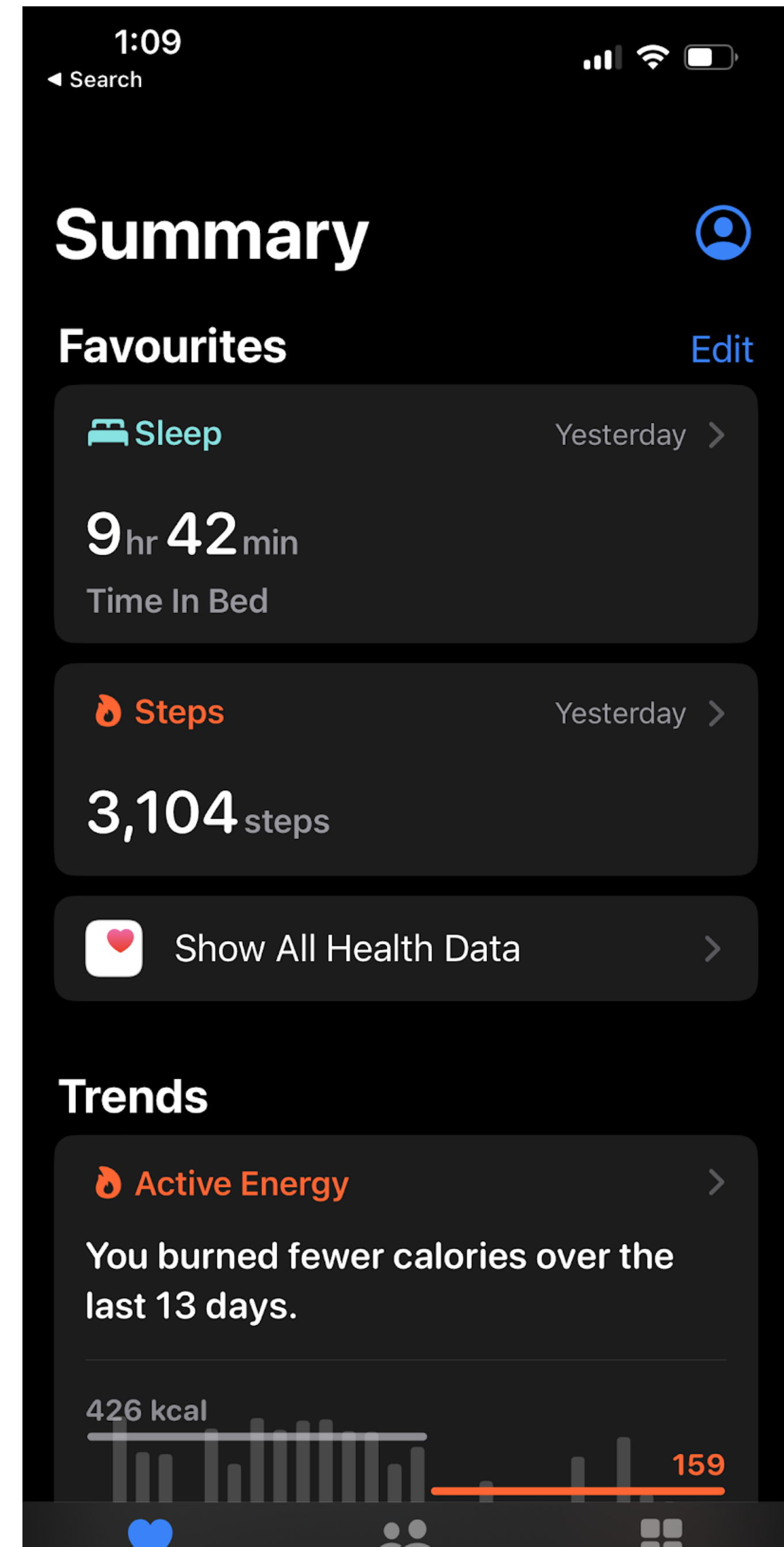
Information

Activity 1

Data selfie

We are all producing data all the time!

Phone calls
Emails/Whatsapp messages
Health/fitness data
Instagram likes
Screen time/App usage
Attendance marking
Location sharing/history
Browser history
Payment/Splitwise data on UPI



Create a pen and paper visualisation of your personal data

- Collect some personal data, around 15-20 observations should be enough.
- Write or list your data points in a simple format (e.g., date/time, category, count).
- Identify any patterns or categories (like “morning vs. evening” or “work vs. personal”). Brainstorm a rough idea of how to visually represent each data point. You can experiment with shapes, lines, colors, or symbols that connect to your theme (e.g., hearts for likes, footprints for steps).

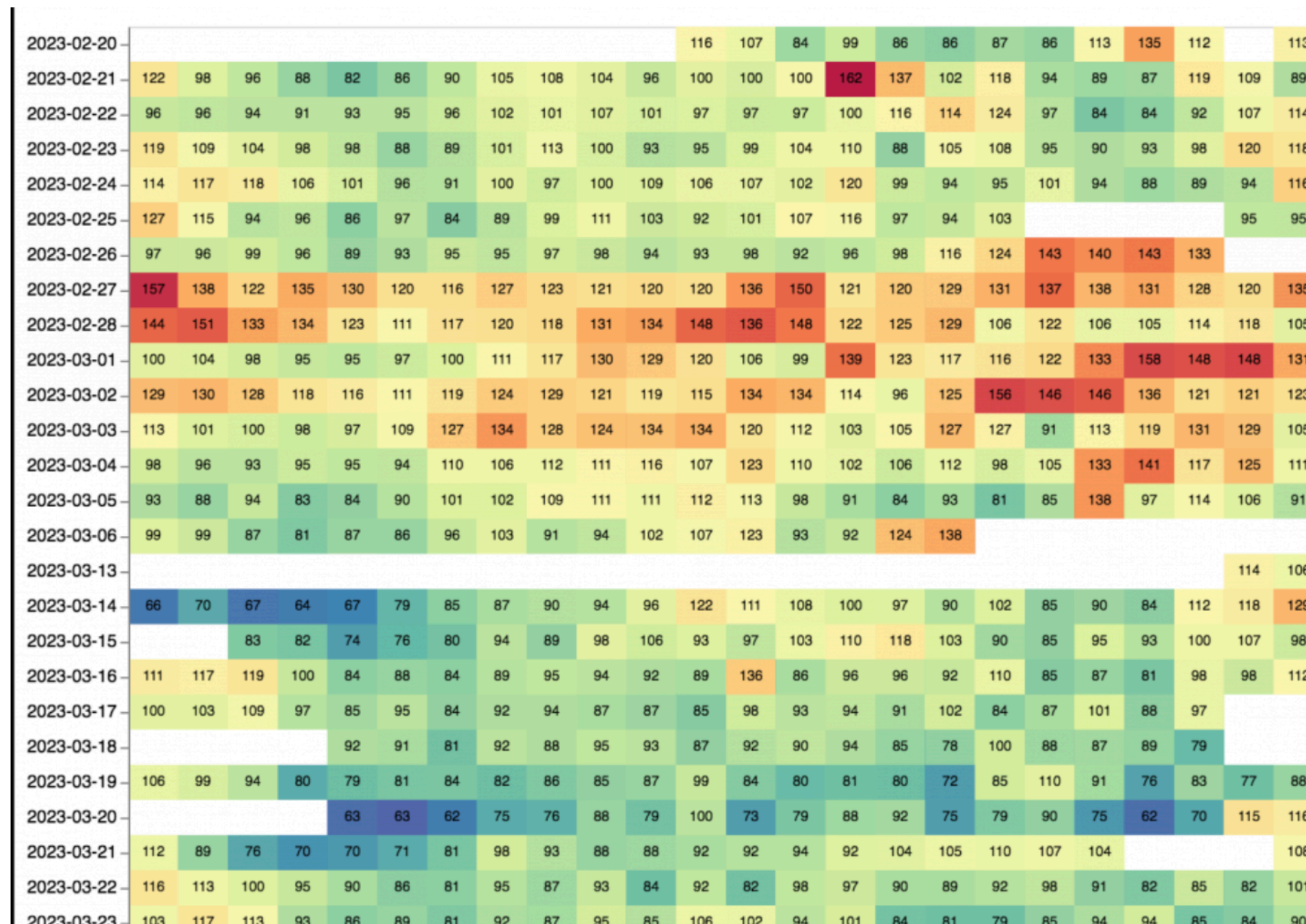
Things to keep in mind

- You will share your visualisation with the class, only use data points you are okay with sharing. Consider categorising or anonymising details you don't want to share. E.g. Categorise into hours of the day instead of exact time.
- Think about how you might visualise the information. You can encode with position, size, shape, colour, etc. Try sketching bar charts, radial diagrams, or more abstract/artistic representations.
- Use labels, legends, and a short caption explaining what each part represents.
- Try using thicker pens, colored pencils, etc. to have more options with encoding.

Quantified Self

quantifiedself.com/

QUANTIFIED SELF
SELF KNOWLEDGE THROUGH NUMBERS



THE KEATING MEMORIAL SELF RESEARCH GROUP

Would you like to get help with your self-research project from an active, experienced group of peers? You're invited to join the Keating Memorial Self Research group. We meet every Thursday at 10am Pacific time. You can find the agenda, notes & links in the full post.

New Show&Tell Event: Tracking Blood Glucose

Please join us for an hour of short "QS Show&Tell" talks about diet and metabolic discovery series using personal science. This session will focus on



Dear Data

[the book](#) [the project](#) [press](#) [the authors](#) [get in touch](#) [news!](#)

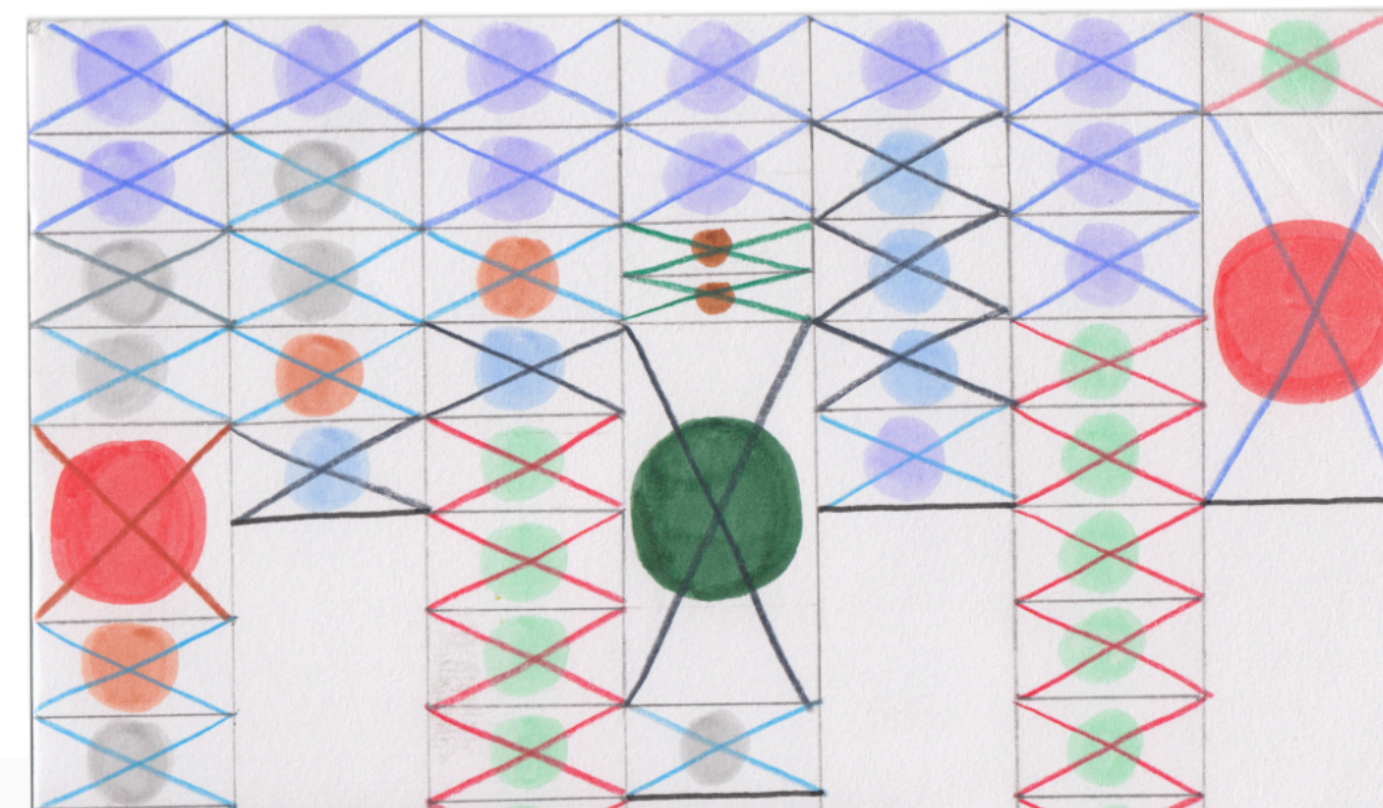
Dear Data is a year-long, analog data drawing project
by [Giorgia Lupi](#) and [Stefanie Posavec](#).

We did it! We reached **52 weeks of drawing our data!**

Week 52

A week of goodbyes

Giorgia Stefanie



Activity 1

Data selfie

Lets analyse

Why did you choose this data point?

How did you decide on the visual style?

Did you discover discover something interesting about yourself?



How?

Data → **Information** → **Knowledge** → **Wisdom**

Knowledge

A deeper understanding derived from combining multiple pieces of information and experience.

“When temperatures rise above 30°C, demand for cold beverages increases.”

“Sales typically increase by 10% in December due to holiday promotions.”

Wisdom

The ability to make sound judgments and decisions based on knowledge, values, and long-term vision.

“Given rising temperatures, we should diversify our cold beverage offerings to maintain competitive advantage.”

“Invest in sustainable supply chains to prepare for long-term growth and environmental impacts.”

Data as design material

What can we do with data?

Exhibit

Show raw data

List, Table, Infographic

Explain

Answer Questions

Data journalism, report

Explore

Finding what to ask

Dashboards, simulations

Experience

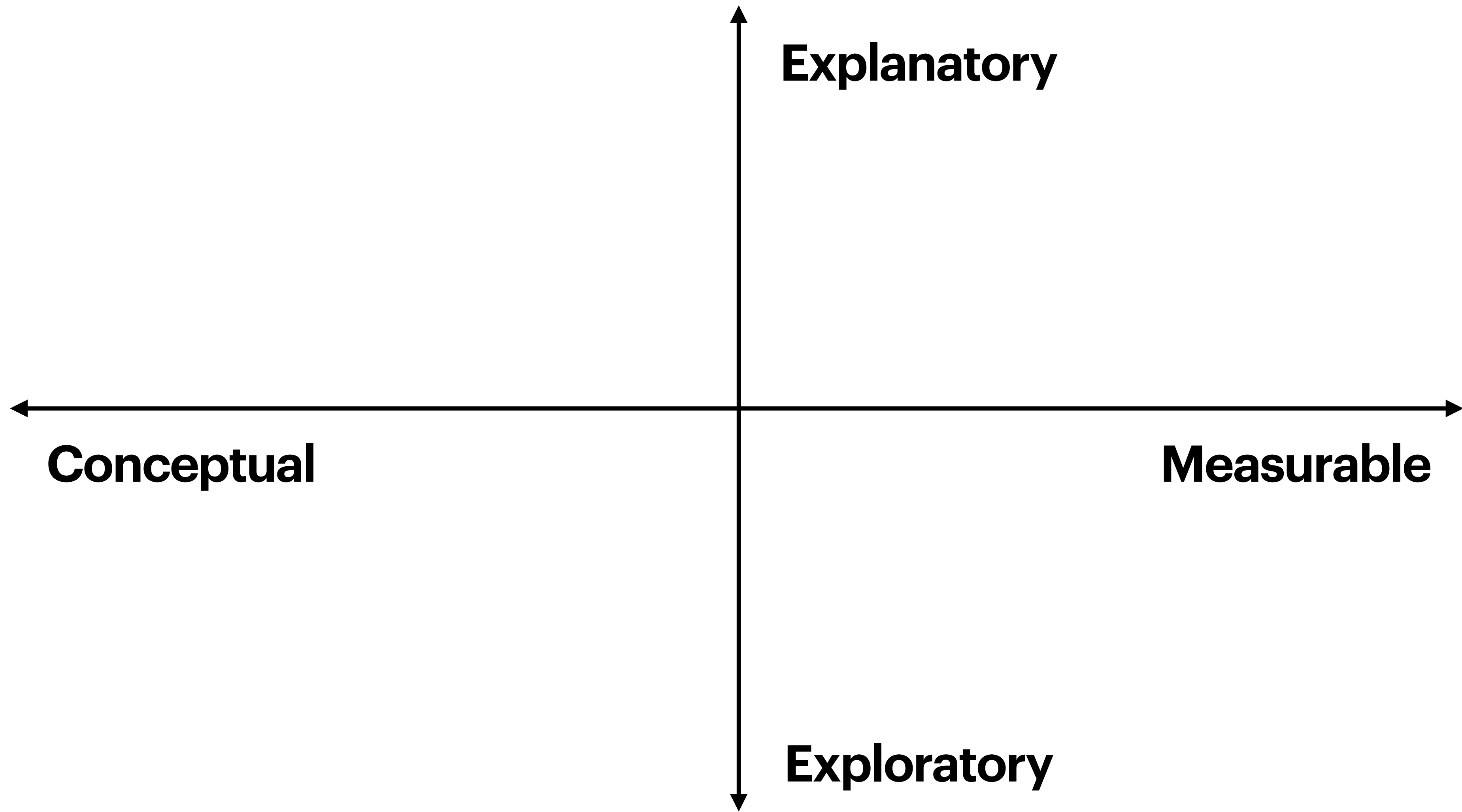
Finding meaning in data

Data art piece, New media
installation

Enable

Building tools to visualise
specific use cases


Software for data viz




Datasets

data.gov.in

A Digital India Initiative

Choose your theme 

 data.gov.in
Open Government Data (OGD) Platform India


HOME CATALOG APIs SECTOR CHIEF DATA OFFICER METRICS | LOGIN | REGISTER

Kisan Call Centre (KCC)


Transcripts Of Farmers Queries & Answers Data

Search Catalog/Resources/API Search

Trending Data Crime Rainfall Pincode Population Market





 ANALYTICS	505,060 RESOURCES	12,462 CATALOG	10.56 M TIMES DOWNLOADED	583 CHIEF DATA OFFICERS
	3,321 VISUALIZATIONS	181 SOURCED WEBSERVICES/APIs	36.89 M TIMES VIEWED	262,052 APIs

Suggest Dataset

NEED HELP? ASK NICCI 

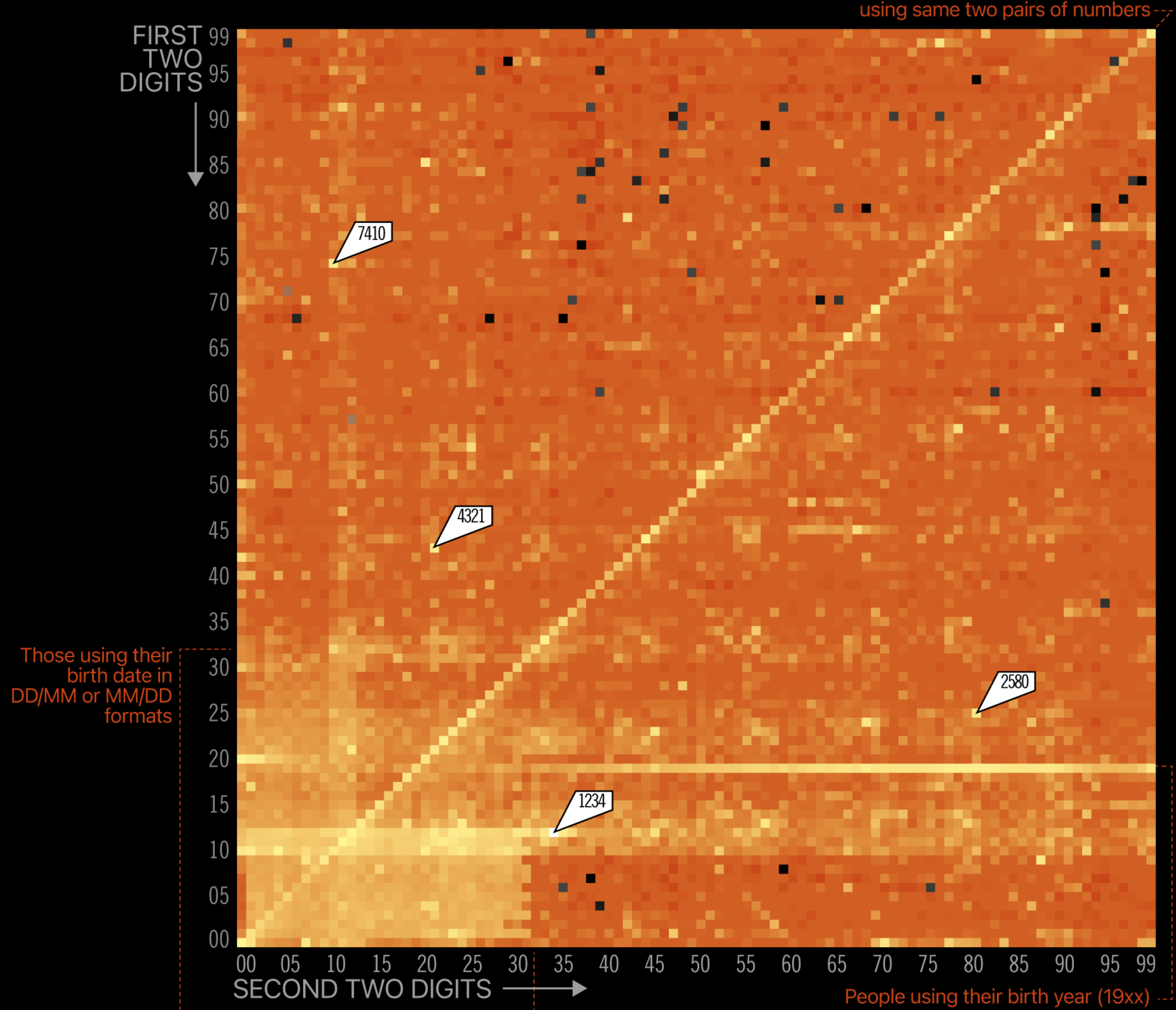
Discover Datasets By

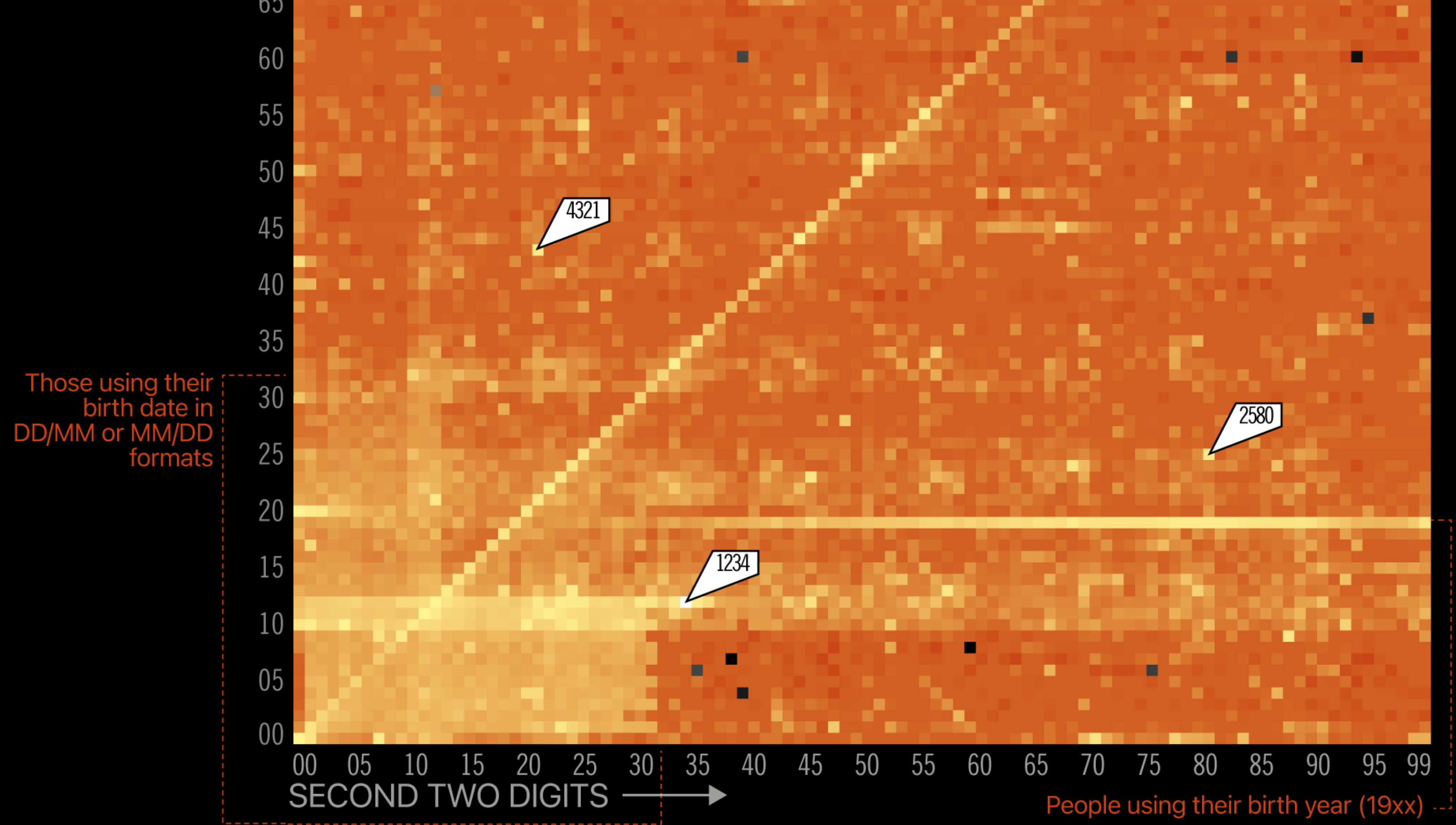
Sector >

Pin Point

The most common
4-digit PIN numbers





most common

1234 0000 7777 2000 2222 9999 5555 1122 8888 2001
 1111 1212 1004 4444 6969 3333 6666 1313 4321 1010

27% of all PIN numbers

least common

8557 8438 9539 7063 6827 0859 6793 0738 6835 8093
 9047 0439 8196 6093 7394 9480 8398 7637 9629 8068

HOME / DATA

Cyber fraud in banking transactions surges in FY24: Data

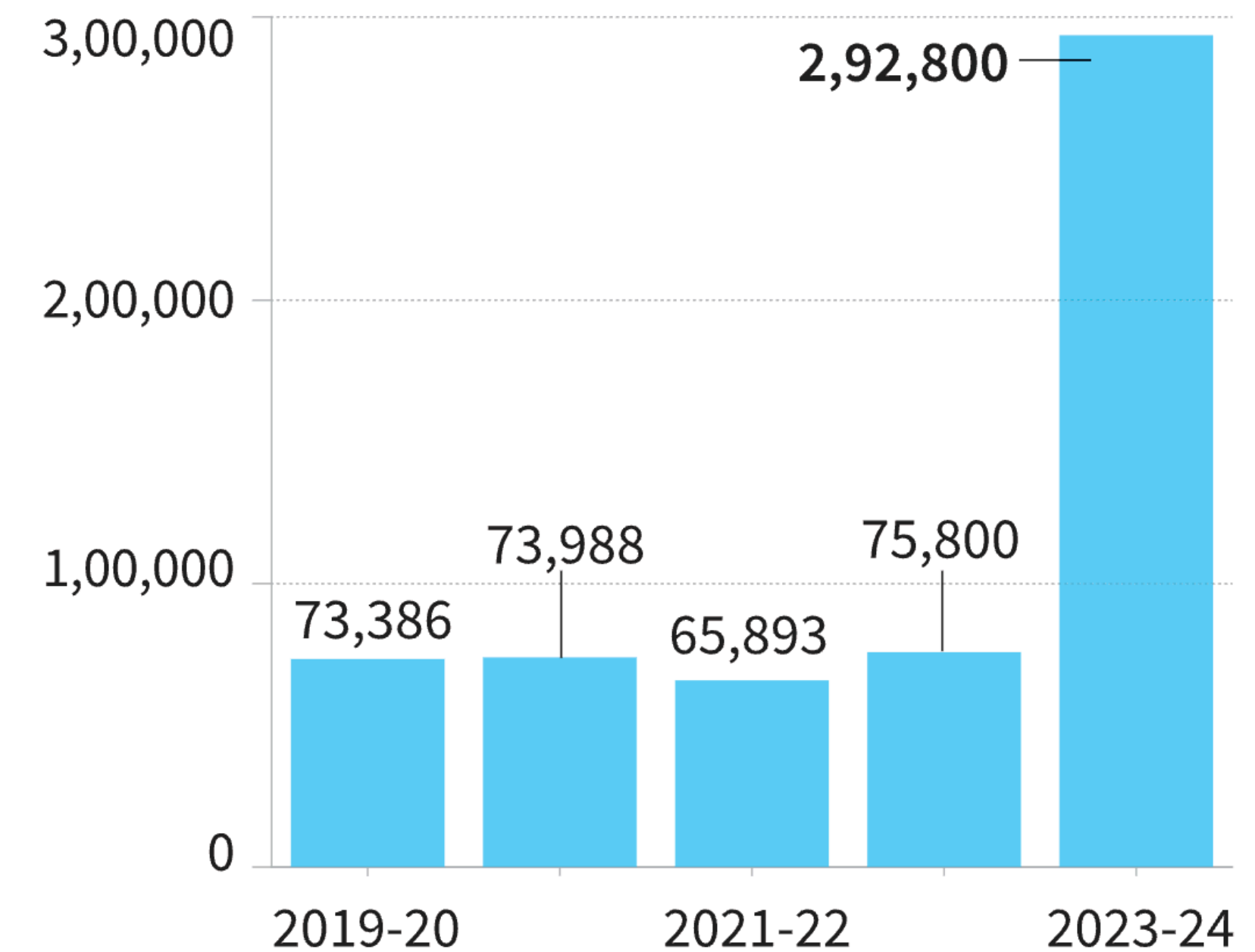
Maharashtra accounts for more than one-fourth of the amount lost due to cyber fraud in India

Updated - November 13, 2024 04:43 pm IST

MD ZAKARIA SIDDIQUI, SABIR AHAMED



Chart 1(a) | *The chart shows the number cyberfrauds across the years in banking transactions*



Why are tomato prices rising in October: Data

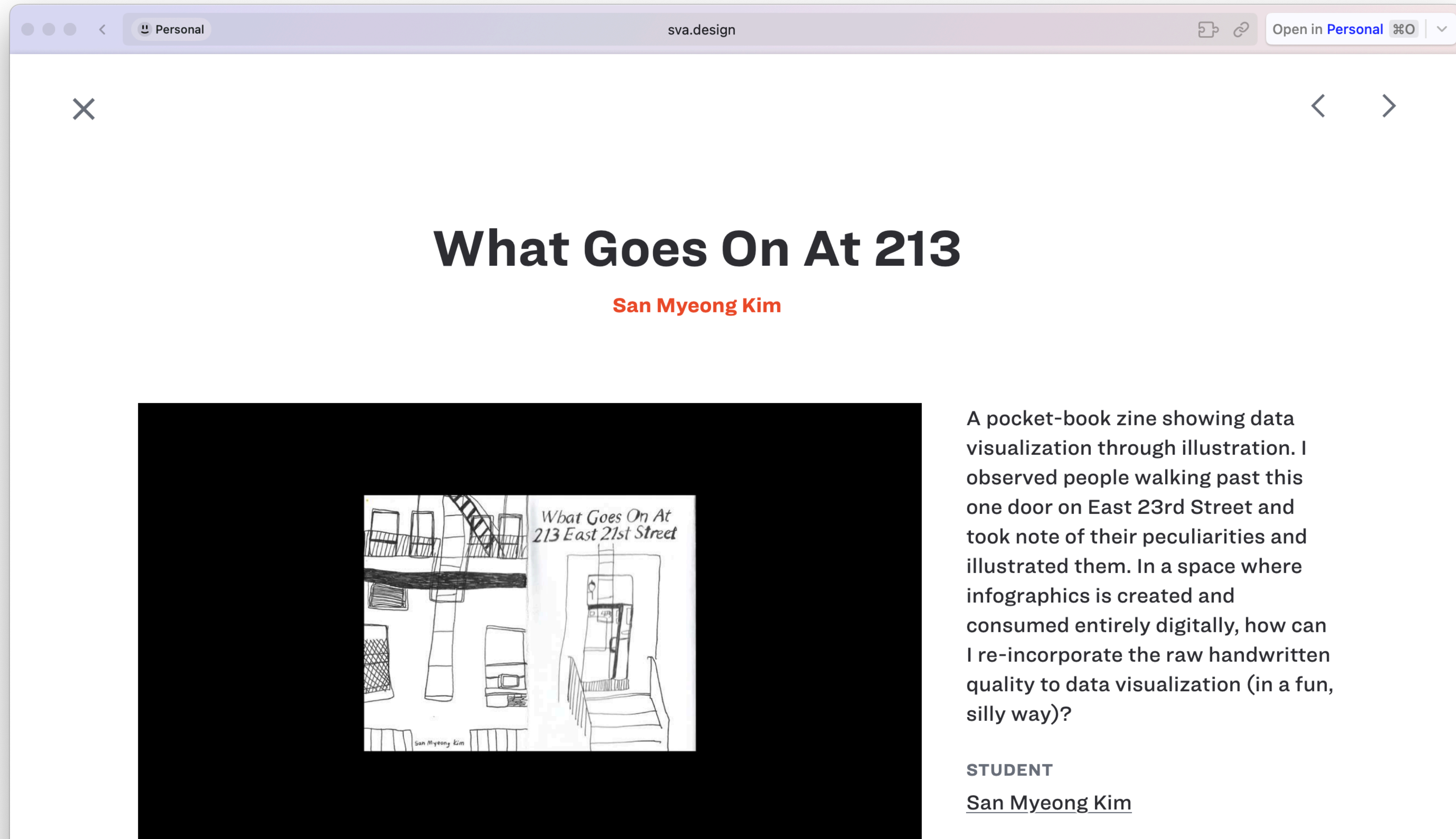
The average retail prices shot up to ₹65-70 per kilo in the north and the eastern regions in October

Updated - October 19, 2024 11:12 am IST

Divisions in tomato belt	June 1-5				July 3-9				August 7-13				September 4-10					
Rayalaseema	LE	LE	LE	D	D	N	N	N	LD	E	E	LE	LD	LE	D	LE	N	D
S.I. Karnataka*	E	LE	D	D	N	N	LE	LE	LE	N	N	LE	D	N	D	LD	D	D
West M.P.	N	D	D	E	N	E	N	N	E	E	N	D	LE	N	LE	LE	D	LE
Odisha	LE	LD	LD	D	E	D	D	LE	N	E	D	D	N	N	E	N	N	LE
Gujarat Region	LD	D	LD	N	E	D	N	E	E	N	D	LD	LE	LE	LE	D	D	LE
Madhya Maharashtra	D	LE	D	N	N	E	E	LE	LE	E	D	E	LE	N	N	LD	N	LE

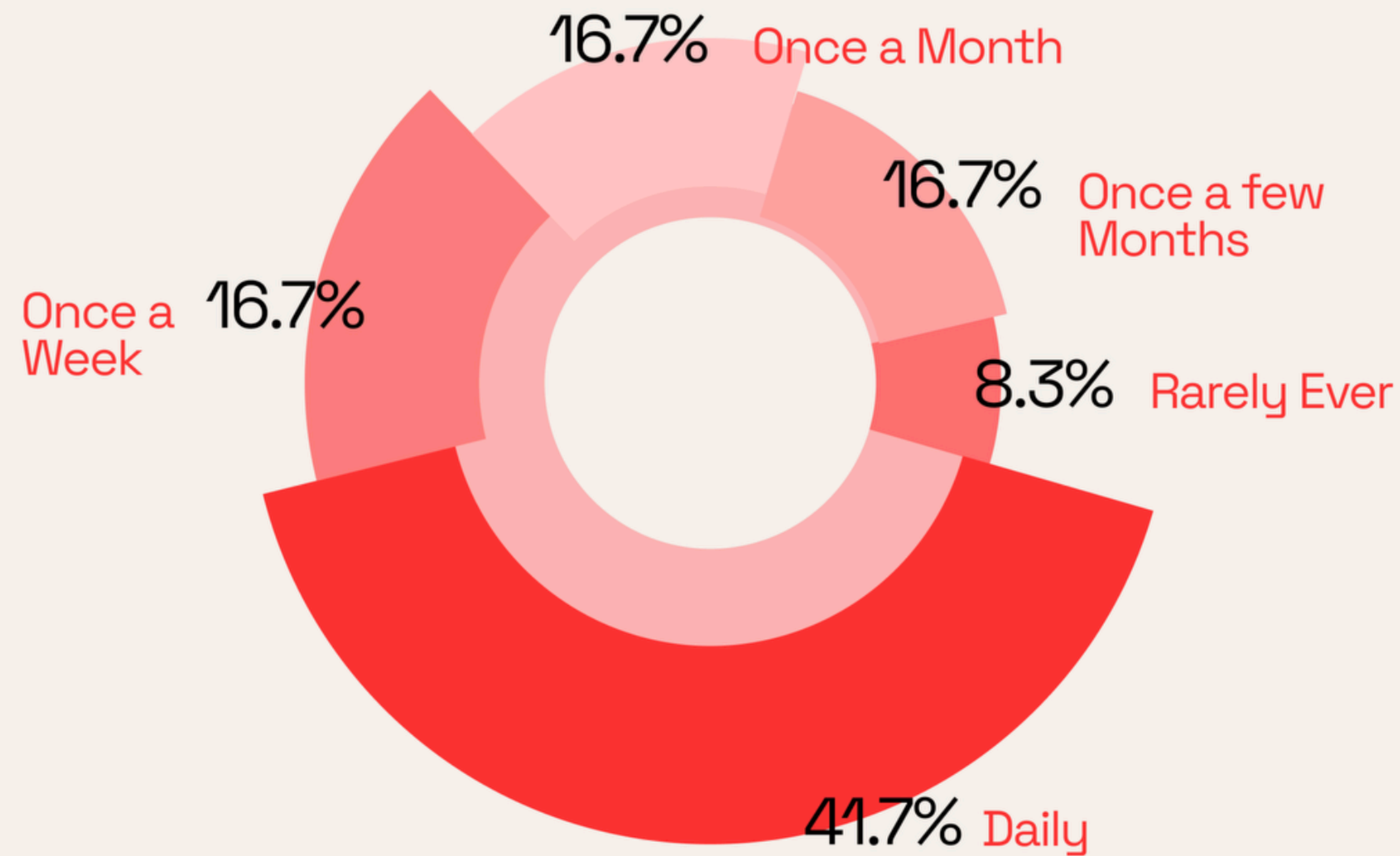
In the graph, **LE**: Large excess rainfall; **E**: Excess, **N**: Normal, **D**: Deficient, **LD**: Large deficient. In the last week of September, large excess rainfall in four of the six major tomato production belts destroyed tomatoes ready for sale





Infographics

Harshwardhan Somwanshey



13 people per 100 person own a vehicle in India



2.2 million injuries in India that warranted hospital admission, and 18 million injuries warranted an emergency room visit

(Bhalla et al. 2014).

The number of cars registered in 2020 was 43.73 M

(Official RTI Statistics).

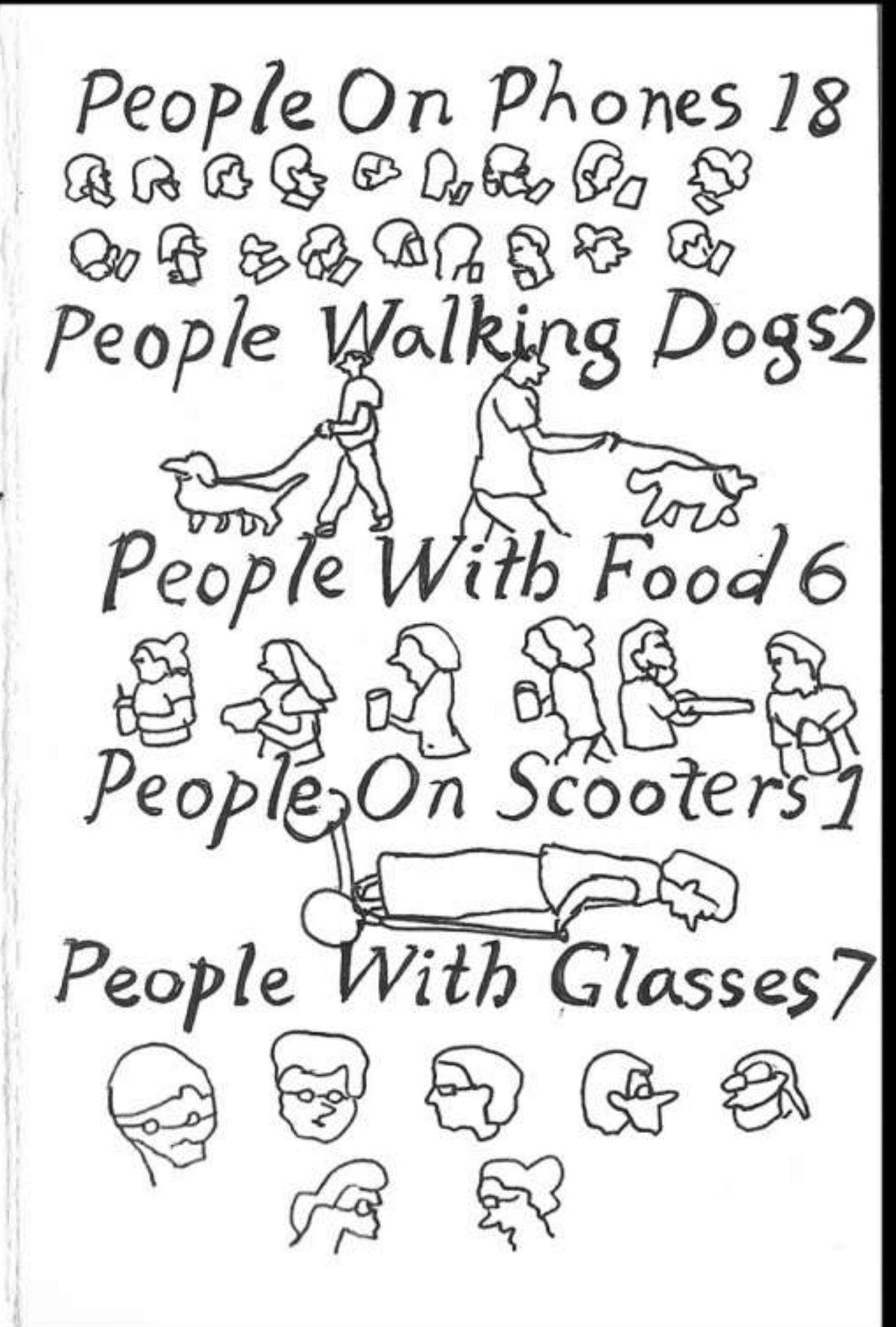
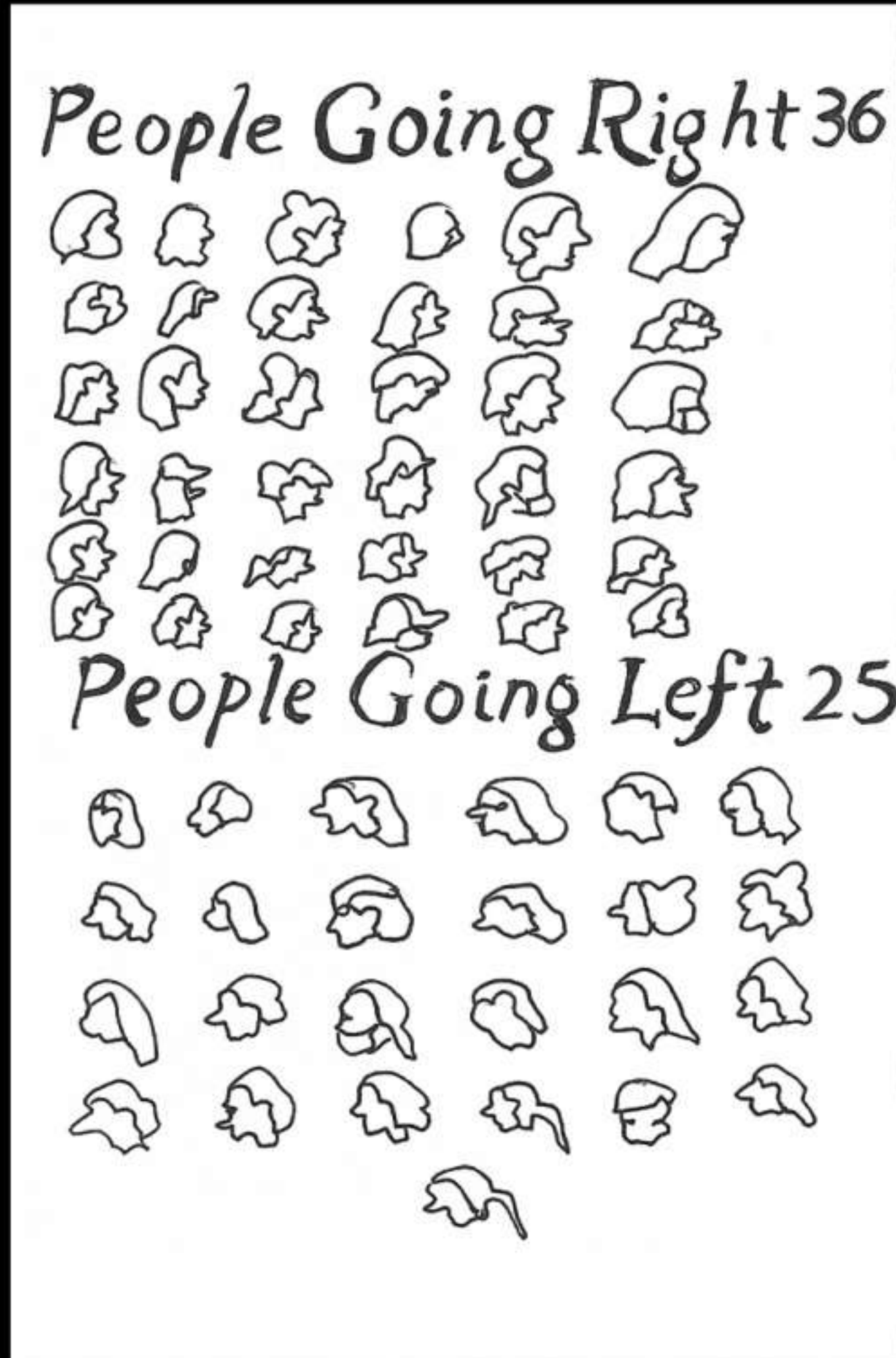
155,622 persons were killed in road traffic crashes in India in 2021

(Official RTI Statistics).



Data Art/Storytelling

sva.design

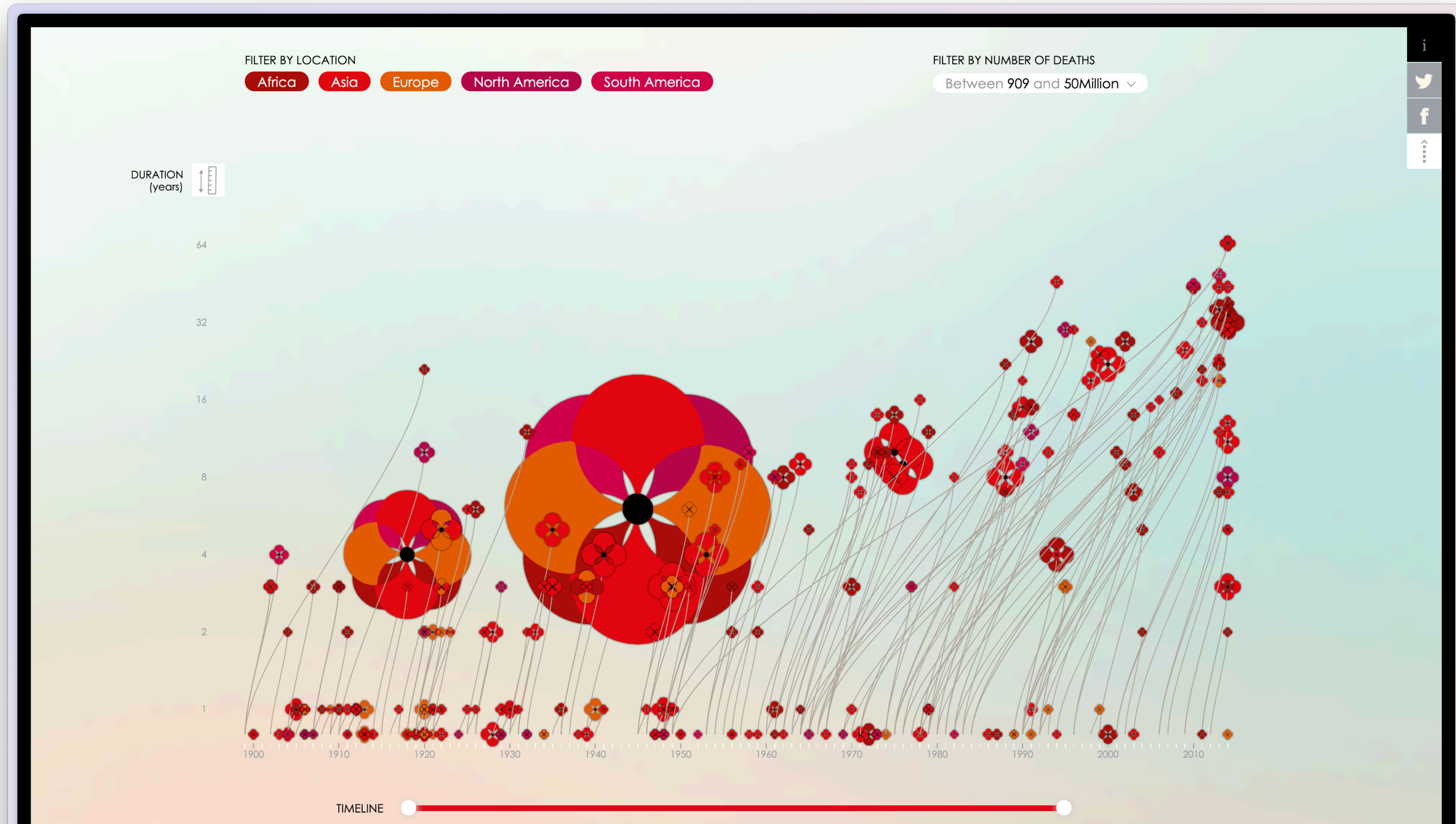




What's for dinner? · Gyan Lakhwani

Data Art

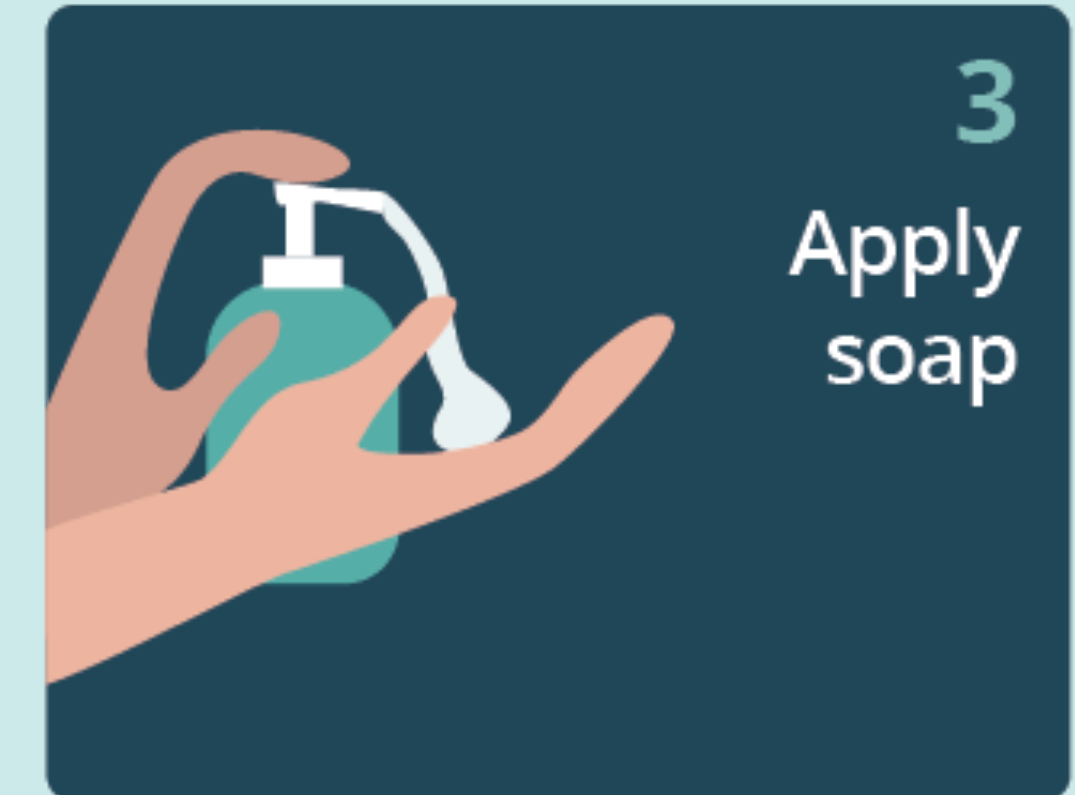
poppyfield.org



Infographic

healthdirect.gov.au

Hand washing tips



Data Storytelling

dear-data.com

Dear Data

[the book](#) [the project](#) [press](#) [the authors](#) [get in touch](#) [news!](#)

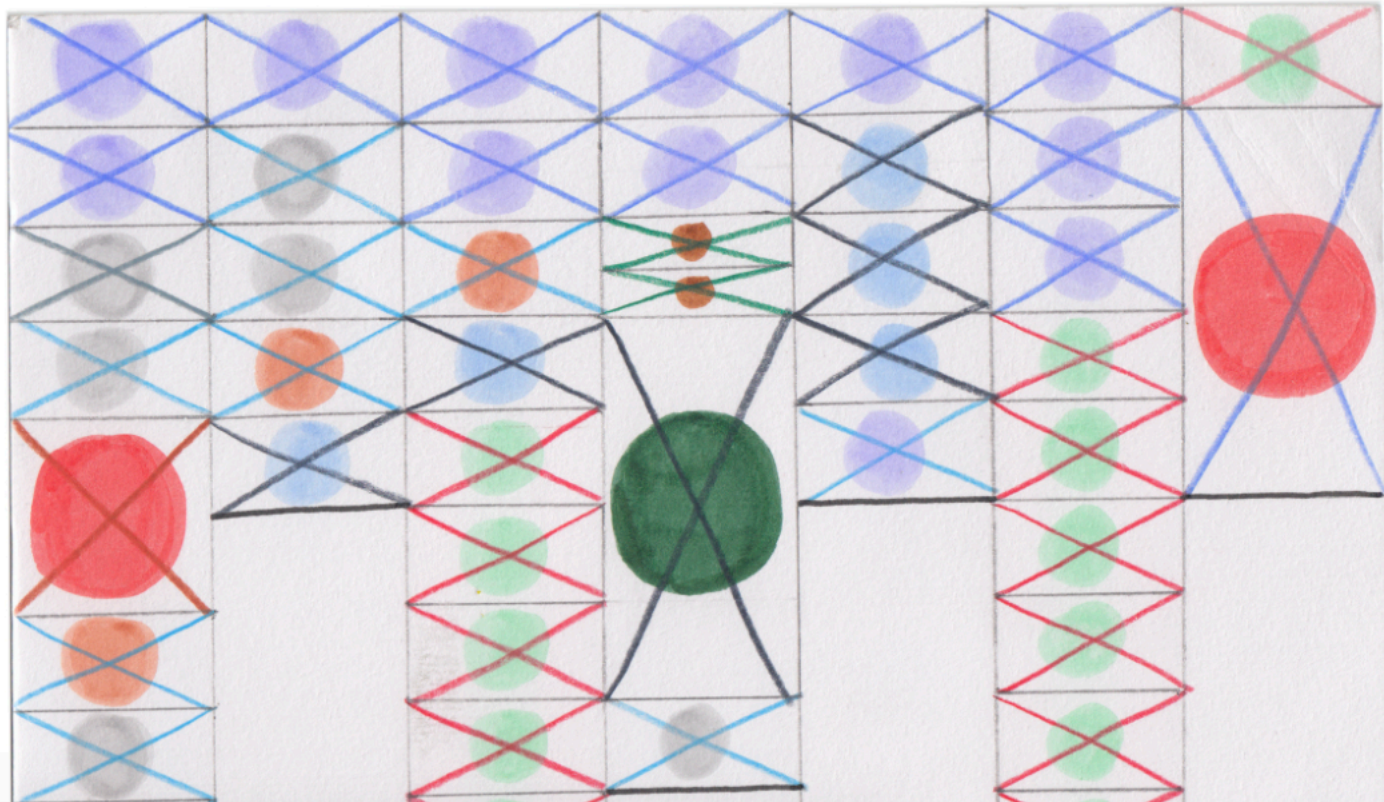
Dear Data is a year-long, analog data drawing project
by [Giorgia Lupi](#) and [Stefanie Posavec](#).

We did it! We reached **52 weeks of drawing our data!**

Week 52

A week of goodbyes

Giorgia Stefanie

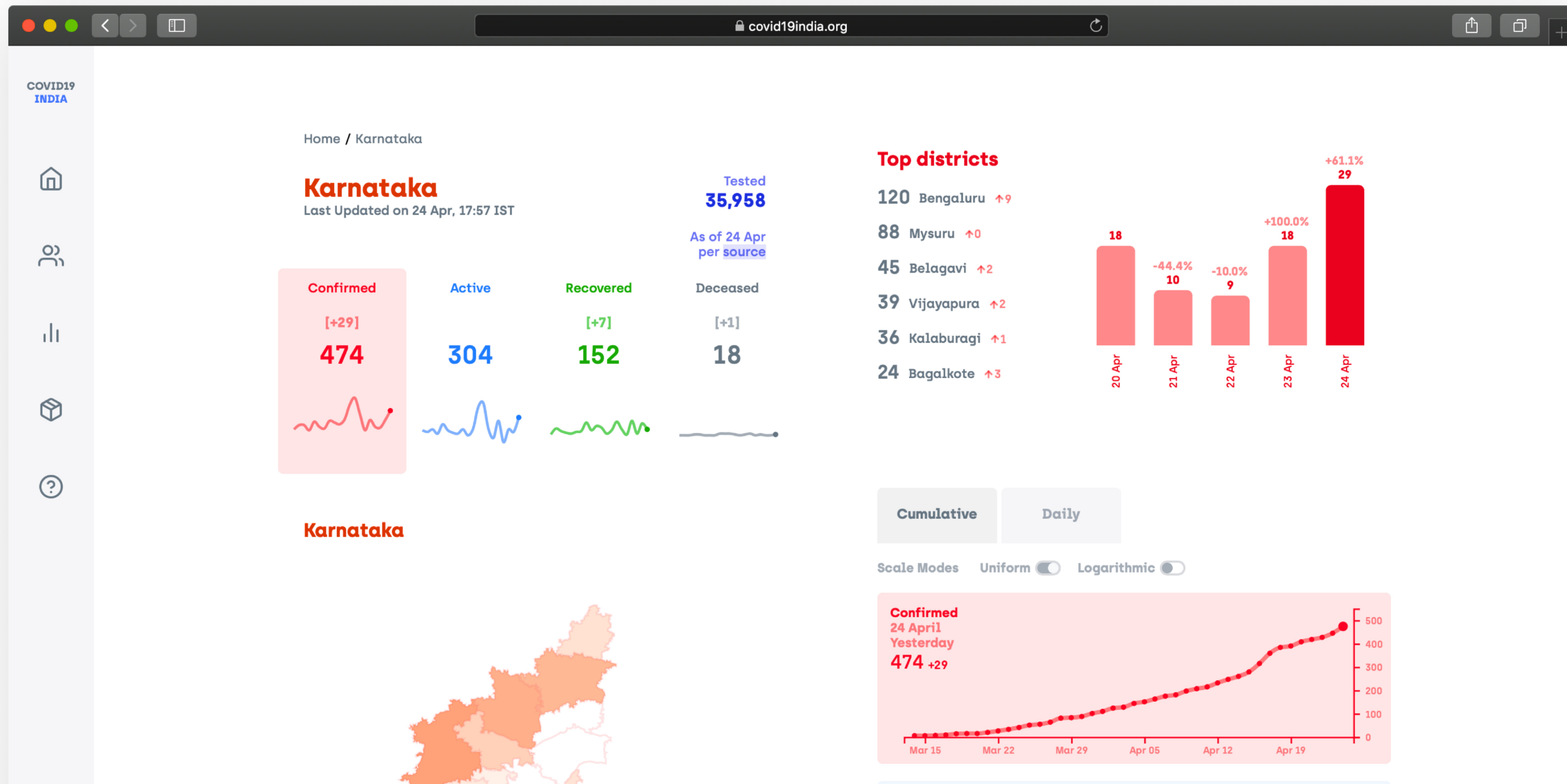


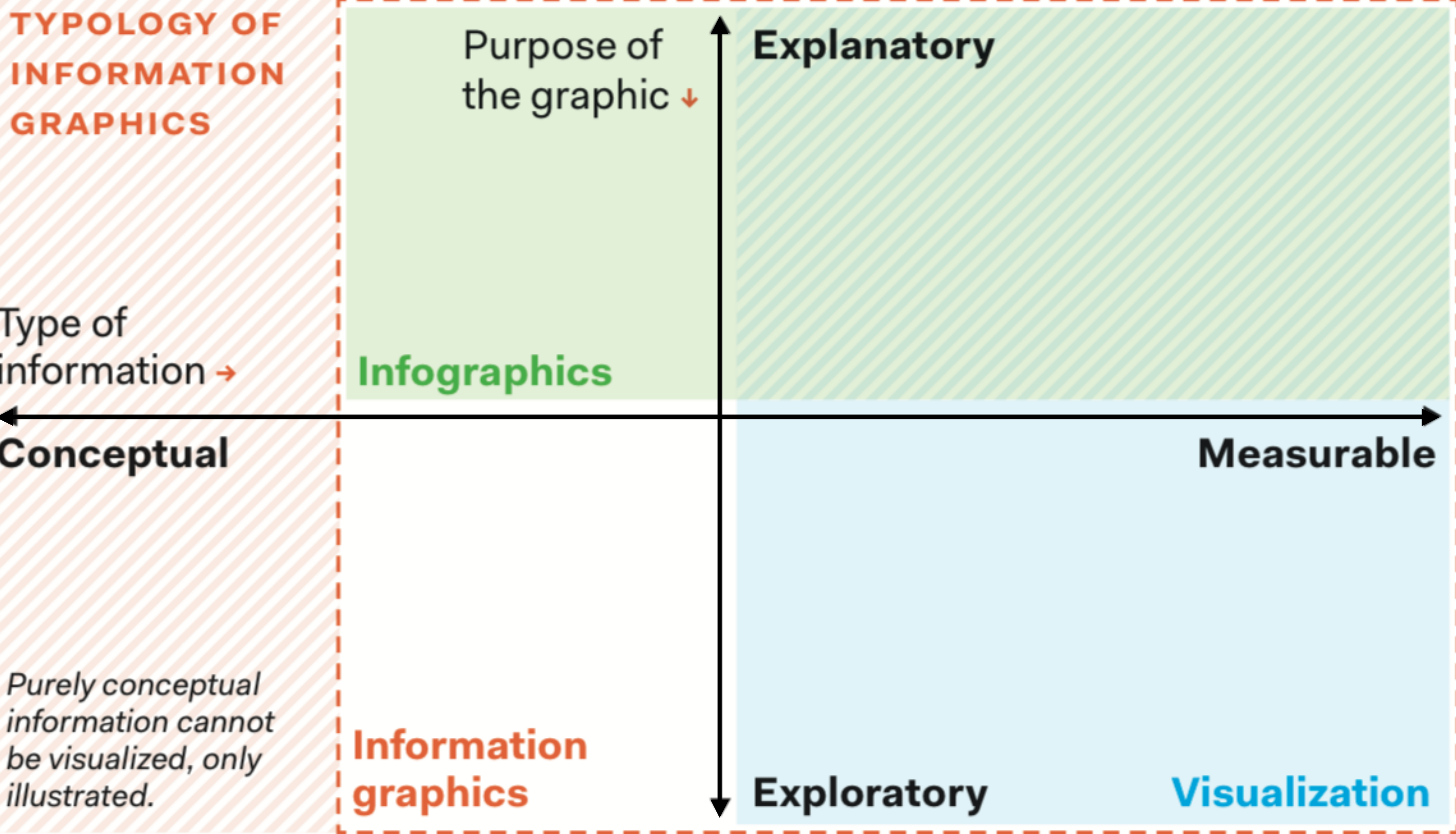
Dashboards

covid19india



COVID19INDIA





Levels of measurement

- **Nominal:** the data can only be categorized
- **Ordinal:** the data can be categorized and ranked
- **Interval:** the data can be categorized, ranked, and evenly spaced
- **Ratio:** the data can be categorized, ranked, evenly spaced, and has a natural zero

Ratio

Categories, Rankable, Equally Spaced, True Zero

- Height (If I have 0cm of cloth, how much cloth do I have?)
- Age
- Weight (10kg is half of 20kg)
- Temperature in Kelvin

Best Encodings

Length / Height

Bar charts are highly effective because starting at zero is meaningful.

Position on a Common Axis

Plots (scatter plots, line graphs) that begin at zero can show proportional differences accurately.

Area or Volume

For special cases (e.g., bubble charts), ratio data can be encoded in areas or volumes, but clarity and scaling must be carefully managed.

Tips

Including a zero baseline is crucial to reflect accurate proportions. If the axis doesn't start at zero, it can misrepresent the magnitude of differences.

Interval

Categories, Rankable, Equally Spaced

- Temperature in Fahrenheit or Celsius
- Calendar Years (e.g. 5 years ago it was 2020, 10 years ago it was 2015)
- Time of the day (1pm to 2pm is the same interval as 3pm to 4pm. However, 4pm is not half the time as 2pm. 00:00 hrs is not the absence of time, it's just the start point for the daily cycle)

Best Encodings

Position on a Common Scale

For continuous variables, a line graph or horizontal axis that shows the numeric scale works well.

Color Gradients

Continuous gradients can indicate intervals, but be mindful of how you handle zero or the “origin.”

Length / Height

Bar charts can work, but labeling is key to remind viewers that zero isn't absolute in interval scales.

Tips

Because interval scales can go below zero (e.g., -5°C), color ramps or dual-axis encodings (positive vs. negative) can help clarify where zero is relative to the data.

Ordinal

Categories, Rankable

- Language ability
(e.g., beginner < intermediate < fluent)
- Spiciness Levels on a Menu
(e.g. Mild < Medium < Hot < Extra Hot)
- Medal placements in competitions
(e.g. Gold > Silver > Bronze)

Best Encodings

1. Ordered Color Gradients (Saturation or Lightness)

- Use a consistent light-to-dark gradient or a single hue varying in saturation to reflect an inherent ranking (e.g., from less to more).

2. Position Along a Single Axis

- Placing items on a vertical or horizontal axis in order, from smallest rank to largest rank.

3. Size (with caution)

- Increasing size (such as the length of bars) can show a progression, but remember ordinal data doesn't ensure precise numerical intervals.

4. Shape Variation that Suggests Progression

- For instance, incremental shapes (thin bar → thicker bar → thickest bar) but only if differences are clearly incremental.

Tips

- Avoid using too many color hues for ordinal data, because changing hue can undermine the sense of order. A single hue in varying saturation is more intuitive.

Nominal Categories

- City of birth
(Delhi, Ranchi, Kolkata, Mysore)
- Gender
(Male, Female)
- Car brands
(Hyundai, Toyota, Mercedes)
- Marital status
(Unmarried, married, divorced)

Best Encodings

Color Hue

Different hues (e.g., red, green, blue) help distinguish categories without implying any hierarchy.

Shape

Vary the shape (circle, triangle, square) to denote distinct categories.

Spatial Grouping

Group items by category in different regions of a chart, map, or diagram.

Tip

Avoid using size, saturation, or brightness for nominal variables, since they can accidentally imply magnitude or importance that isn't intended.

Pop Quiz!

A researcher collects demographic data from her participants.

She asks participants for their city of birth.

Which level of measurement is this?

A researcher collects demographic data from her participants.
She asks participants for their city of birth.
Which level of measurement is this?

[A] Nominal

[B] Ordinal

[C] Interval

[D] Ratio

A Likert scale is a rating system that measures attitudes, opinions, or behaviors by asking respondents to choose from a range of answer options.

What type of measurement is a Likert scale?

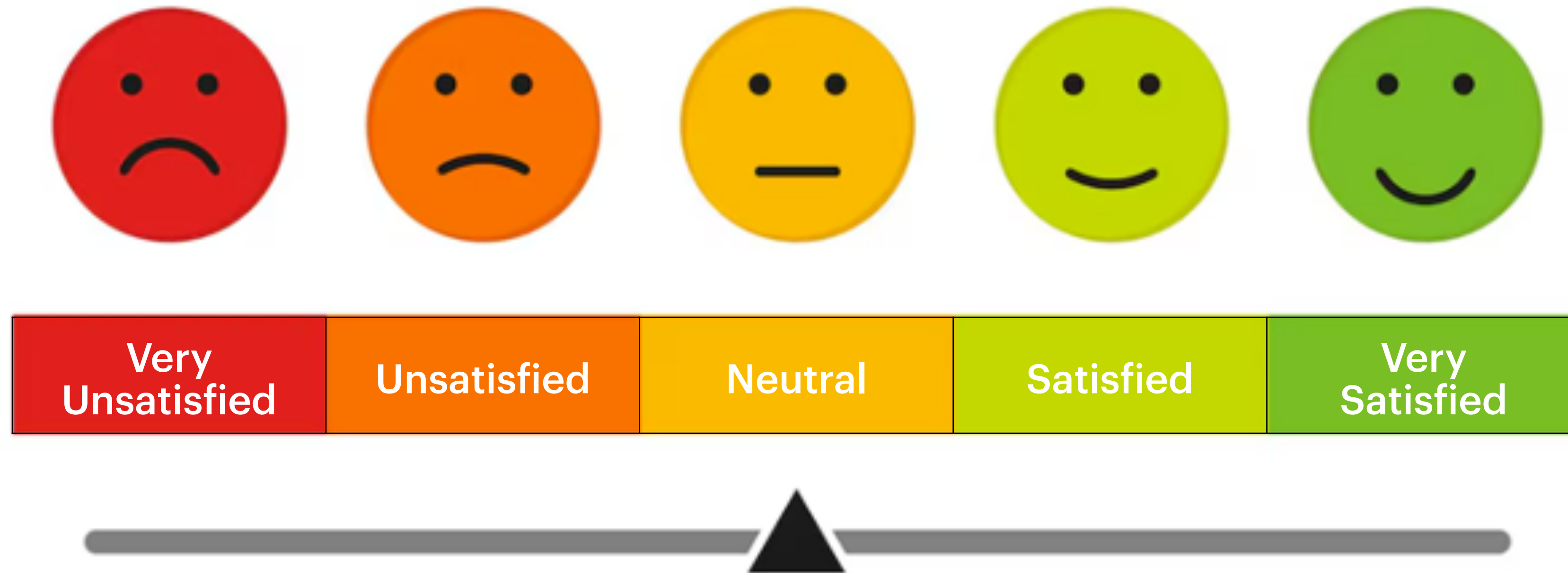
[A] Nominal

[B] Ordinal

[C] Interval

[D] Ratio

How satisfied are you with our service?



Level of measurement	Description	Examples
Nominal	Data can only be categorized	Gender, eye color, hair color
Ordinal	Data can be categorized and ranked, but intervals between categories are not necessarily equal	Education level, Likert scale responses, income
Interval	Data can be categorized, ranked, and evenly spaced	Temperature measurements in Celsius or Fahrenheit
Ratio	Data can be categorized, ranked, evenly spaced, and has a natural zero	Height, weight

Data Encoding

Mapping

data variables → visual channels

Nominal

Ordinal

Interval

Ratio

Lines

Bars

Shapes

Colors

Positions

Certain types of variables map better to certain ways of visualising information

Data visualisation tools

Or, homework #1!

Group Activity: Look into these Data visualisation tools. Create a presentation on how they work and present them to the class in the next session (14 jan?)

- Flourish
- Tableau
- Infogram
- ChartBlocks
- Datawrapper
- Google Charts
- D3.js
- Carto
- Mapbox
- Power BI
- Charticulator
- RAWGraphs
- Visme
- Venngage
- Piktochart
- RAW Graphs
- Orange
- P5.js
- chart.js
- Leaflet.js
- dygraphs
- Timeline.js
- Gephi
- Candela
- Rshiny
- Google Data Studio
- Datawrapper
- Palladio
- OpenRefine
- Circos
- Plotly
- Observable
- Tangle