

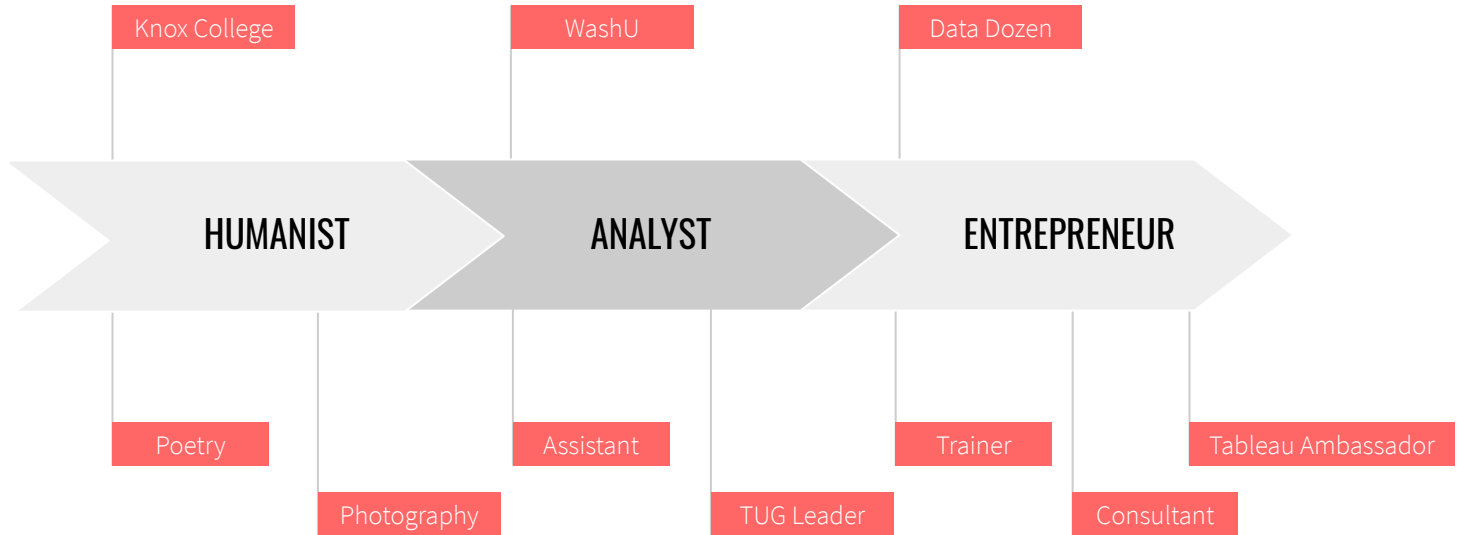
A PRESENTATION FROM
DATA DOZEN
a visual analytics studio

ERIN WALDRON FROM DATA DOZEN · OCTOBER 12, 2021 · DATA VISUALIZATION MAKEOVERS · WORKSHOP 1

FUNCTIONAL DESIGN

HOW TO BUILD VISUALIZATIONS THAT YOUR AUDIENCE TRULY UNDERSTANDS





A BIT ABOUT ME

ROADMAP

Here's what we'll cover with our three workshops.

COMPREHENSION

FUNCTIONALITY

How does the visualization work?



EMOTION

AESTHETICS

How does the visualization feel?



VALIDATION

TESTING

How does the visualization perform?



FORMAT

Each workshop will offer a collection of guided exercises.

THEORY + PRACTICE

NEW PERSPECTIVES

Examine your work from a variety of lenses.



CURRENT MAKEOVER

NEW IDEAS

Improve your current visualization.



FUTURE WORK

NEW TOOLS

Incorporate new strategies into your process.



WORKSHOP 1

Today we've got four exercises to improve our work's functional design.

PURPOSE

MADLIB

Define the intention of your work.



CHARTS

FLOWCHART

Match your data to the right chart type.



PATH

PINBALL

Curate your audience's journey.



FIT

SPECTRUMS

Calibrate your work's cognitive load.



WORKSHOP 1

Today we've got four exercises to improve our work's functional design.

PURPOSE

MADLIB

Define the intention of your work.



breakout

CHARTS

FLOWCHART

Match your data to the right chart type.



breakout

PATH

PINBALL

Curate your audience's journey.



breakout

FIT

SPECTRUMS

Calibrate your work's cognitive load.



breakout (?)

30 SECOND LIGHTNING ROUND

- ▶ What's your **name** and organization? (this is for me:)
- ▶ What's your **project** and your goal for it by the end of the module?
- ▶ Can you **share** your project? (screen sharing today? before and after in the future?)

BEFORE WE START

VISUALIZATION MAKEOVERS

Before we dive into the first exercise,
take a **before** screenshot.



WORKSHOP 1

Today we've got four exercises to improve our work's functional design.

PURPOSE

MADLIB

Define the intention of your work.



CHARTS

FLOWCHART

Match your data to the right chart type.



PATH

PINBALL

Curate your audience's journey.



FIT

SPECTRUMS

Calibrate your work's cognitive load.



FILL IN THE BLANK

____ **WHO** ____ (audience) will use this visualization about
____ **WHAT** ____ (content) to ____ **WHY** ____ (purpose).
They will interact with this visualization
on a ____ **WHERE** ____ (platform)
for ____ **WHEN** ____ (minutes) on a ____ **WHEN** ____ (frequency) basis.

FILL IN THE BLANK

_____ (audience) will use this visualization about
_____ (content) to _____ (purpose).

They will interact with this visualization
on a _____ (platform)
for _____ (minutes) on a _____ (frequency) basis.

FILL IN THE BLANK

The school board (audience) will use this visualization about
class enrollment (content) to allocate funds (purpose).
They will interact with this visualization
on a paper report (platform)
for 120 (minutes) on a annual (frequency) basis.

FILL IN THE BLANK

The cafeteria staff (audience) will use this visualization about
class enrollment (content) to prepare each day's lunch (purpose).
They will interact with this visualization
on a web portal (platform)
for 5 (minutes) on a daily (frequency) basis.

PURPOSE

Every data visualization is like a puzzle piece that connects the audience on one side to the data on the other. To be successful, your work has to respect the pre-existing contours on either side.





WHO
AUDIENCE



WHAT
CONTENT



WHEN
TIME



WHERE
PLATFORM



WHY
PURPOSE

MADLIB

- ▶ Which of these prompts feels like your project's biggest **blindspot**?
- ▶ Which of these prompts has the most underlying **friction**?
- ▶ Which of these prompts makes you most **excited** to work on your project?



WHO AUDIENCE



WHAT CONTENT



WHEN TIME



WHERE PLATFORM



WHY PURPOSE

MADLIB

► SPECIFICITY

Describe your audience. Get specific. Who are they? What are their primary concerns? What do you know about them? What don't you know about them?

► SIZE

How many folks are you trying to reach? If you could reach just one person, who would they be?

► INVESTMENT

Does your audience actively care about this data? This project? This visualization?

► DATA

What's their relationship to the data? Did they help create it? Are they in it?

► PRIMING

Are they primed to support or reject this data? (Folks are rarely completely neutral.) How could this project potentially offend, frustrate, or alienate them?

► TRUST

Describe your relationship with the audience. Is there skepticism? Camaraderie? Depth? Do they even know you? Fundamentally, do they trust you? Why or why not?



WHO
AUDIENCE



WHAT
CONTENT



WHEN
TIME



WHERE
PLATFORM



WHY
PURPOSE

MADLIB

- ▶ **VARIABLES**

Describe your data set. In particular, which variables do you want your audience to understand most? (i.e., if you could only take five variables to a desert island, which would they be?)

- ▶ **QUANTITY**

Is there enough data to actually work with in a meaningful way?

- ▶ **QUALITY**

At the end of the day, do you trust the data set? Do you understand the underlying methodology?

- ▶ **ASSUMPTIONS**

What assumptions or decisions are baked into your data set make? How can you communicate those to your audience? ("Raw data is an oxymoron" --Lisa Gitelman)

- ▶ **EXPERTISE**

How well do you know the content? Do you need a domain expert in order to make the analysis meaningful? (Remember, there's no shame in that!)



WHO
AUDIENCE



WHAT
CONTENT



WHEN
TIME



WHERE
PLATFORM



WHY
PURPOSE

MADLIB

► **DURATION**

How long will you have your audience's attention? How long do you wish you could have your audience's attention? How big is the chasm between those two answers?

► **FREQUENCY**

Describe the intended frequency with which your audience will interact with your visualization? Is this something they'll see once in a presentation or report? Or is this something that you expect them to interact with routinely? Daily? Monthly? Quarterly?

► **HABIT**

Will they see it frequently enough to remember how it works or will they need a refresher each time?



WHO
AUDIENCE



WHAT
CONTENT



WHEN
TIME



WHERE
PLATFORM



WHY
PURPOSE

MADLIB

- ▶ **PUBLICATION**

Where will the final visualization live? (ex: in a slide deck, on a blog, in a password protected online portal, etc.) What are those layout dimensions?

- ▶ **TECHNOLOGY**

Will the audience consume the visualization on their phones or on a computer?

- ▶ **INTERACTIVITY**

Is the final visualization interactive? How will the audience know what they can do?



WHO
AUDIENCE



WHAT
CONTENT



WHEN
TIME



WHERE
PLATFORM



WHY
PURPOSE

MADLIB

► JUSTIFICATION

Set aside the buzz word justification for your work. Go deeper here than just “Data-driven decision making” and “real-time insights”. Why does this specific visualization need to exist for this specific audience?

► CURRENT NEED

Does your visualization fill a current need or is just a formality, an organizational habit justified “because we’ve always done it this way”?

► MOTIVATION

Broadly speaking, what’s your motivation? What’s your organization’s motivation? What’s the audience’s motivation? How do these differ?

► PURPOSE

Describe the visualization’s purpose. Is the visualization proactive work, reactive postmortem analysis, real-time monitoring? Is the work meant to incentivize, penalize, or or simply serve as self-analysis (carrot, stick, mirror paradigms)?

► SUCCESS

What does a successful outcome look like for your project? How will you measure that?

MADLIB

DISCUSSION QUESTIONS

- ▶ Which of these prompts feels like your project's biggest **blindspot**?
- ▶ Which of these prompts has the most underlying **friction**?
- ▶ Which of these prompts makes you most **excited** to work on your project?

_____ **WHO** (audience) will use this visualization
about _____ **WHAT** (content) to _____ **WHY** (purpose).
They will interact with this visualization on a
_____ **WHERE** (platform) for _____ **WHEN** (minutes) on a
_____ **WHEN** (frequency) basis.

WORKSHOP 1

Today we've got four exercises to improve our work's functional design.

PURPOSE

MADLIB

Define the intention of your work.



CHARTS

FLOWCHART

Match your data to the right chart type.



PATH

PINBALL

Curate your audience's journey.



FIT

SPECTRUMS

Calibrate your work's cognitive load.



CHARTS

All visualizations have the same basic building blocks.

We represent data with three types of marks:

Points



Lines



Area



And then we tell multiple marks apart by:

Position



Size



Color



Shape



DEAR DATA

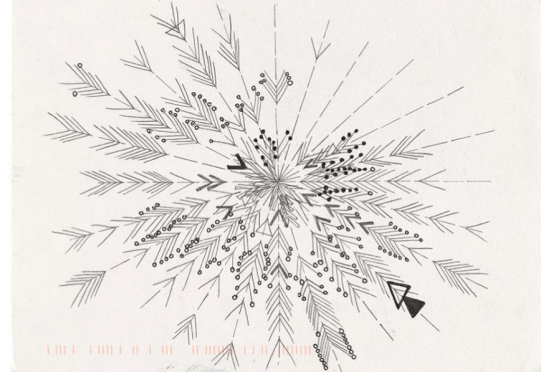
Georgia

Week One



A Week of Clocks

Stefanie



DEAR DATA
WEEK 01: WHAT'S THE TIME?

HOW TO READ IT:

- each symbol represents every moment I glanced at the clocks grouped by hours of the day
- different symbols and attributes represent why and how I checked the time

SYMBOLS:

- 00 purpose: wanted to know what the time was
- just glances: on a phone, near or else
- Because I thought of this project
- X I thought I don't look! on but I did
- 1 Because I was bored
- 10 Because I was hungry
- A Head: someone saying the time aloud

ATTRIBUTES:

- 1: fresh I'm late!
- oh, ok, I'm fine
- analog support (i.e. wrist watch)
- alarm clock rang
- glanced at the clock while talking or emailing with Stefanie

SECOND ATTEMPT
07 OCT 2014

SEND TO:

STEFANIE POSAVEC
LONDON
[UK]
ENGLAND

Drawing her first postcard, Georgia had an idea for her whole collection: from now on every time she tracks something related to Stefanie, or to Dear Data, she uses a special pen to represent it!
pink ink pen!

DEAR DATA: WEEK 01:
A WEEK OF CLOCKS

Hi Georgia! Still getting used to drawing maps, hope I get better! Lots of the clock-face clocks but I am not because I had to leave very to fly back

LEGEND

- EACH LINE = ONE HOUR OF THE DAY, MOVING CLOCKWISE
- EACH LINE = ONE LINE = ONE DAY, BEGINS IN CENTRE + MOVES OUTWARD
- OTHER INSIGHTS: I've learned I'm not really in bed every before you alarm goes off. Hence the Sun clock continuity. Sleep

- AN INSTANCE OF CLOCK-WATCHING IS INDICATED BY A SYMBOL:
- PHONE: 151
- LAPTOP: 84
- TABLET: 10
- HUSBAND'S PHONE: 3
- WATCH: 11
- CAR: 22
- MICROWAVE: 1
- FRIEND'S OVEN: 1
- CHURCH CLOCK: 1

FROM: S. POSAVEC

London, UK
15-09-2014
54007997

GEORGIA LUPI

BROOKLYN, NY
USA

BY AIR MAIL
par avion

This week Georgia and Stefanie tried gathering data in small notebooks (tedious), but soon switched to making notes on their phones (much easier). Stefanie's favourite clock to capture: a bell tolling the time in a town in Devon.

CHARTS

Typically we're looking for the most intuitive way to present the data.



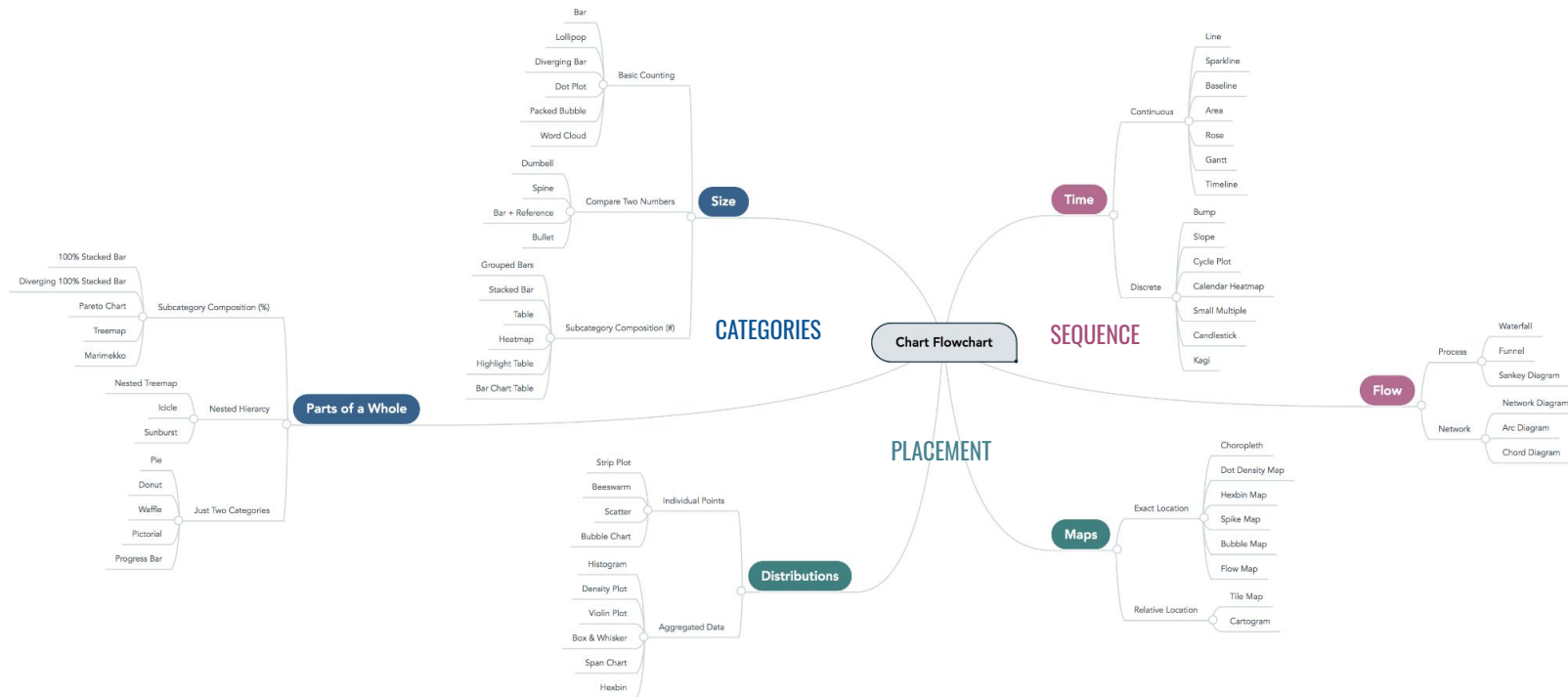
GOOD REASONS TO CHOOSE A CHART

- ▶ It's easy to read. (When in doubt, start with a bar chart.)
- ▶ It draws the eye immediately to the key point of interest.
- ▶ It works well with the project's final delivery format.

BAD REASONS TO CHOOSE A CHART

- ▶ It looks flashy, trendy, colorful, complex, or intimidating.
- ▶ It's easy to make in the software I have.
- ▶ It's the chart type I always use (or I'm asked always to use.)

FLOW CHART



FLOW CHART

- ▶ **BAR CHART**
precision
- ▶ **LOLLIPOP**
lighter, label
- ▶ **DIVERGING BAR**
positive or negative
- ▶ **DOT PLOT**
individual occurrences
- ▶ **BULLET CHART**
goal progress
- ▶ **BAR + REFERENCE**
benchmarking



- ▶ **SPINE CHART**
distribution
- ▶ **DUMBELL**
divergence
- ▶ **GROUPED BAR**
subcategory comparisons
- ▶ **STACKED BAR**
subcategory composition
- ▶ **DIVERGING STACKED**
+ / - connotation
- ▶ **PARETO CHART**
80/20 analysis

CHARTS

DISCUSSION QUESTIONS

- ▶ Which **marks** and differentiators does your visualization use?
- ▶ What good and bad chart-choosing **habits** do you have?
- ▶ Which of the charts in the **flowchart** have you never used (or heard of)?

We represent data with three types of marks:

Points



Lines

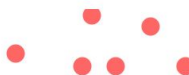


Area



And then we tell multiple marks apart by:

Position



Size



Color



Shape



WORKSHOP 1

Today we've got four exercises to improve our work's functional design.

PURPOSE

MADLIB

Define the intention of your work.



CHARTS

FLOWCHART

Match your data to the right chart type.



PATH

PINBALL

Curate your audience's journey.



FIT

SPECTRUMS

Calibrate your work's cognitive load.



PATH

How do folks **read** a visualization?



IMAGINED PATH

IMAGINED PATH

Vaccination rates by county social vulnerability

IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

<20% 25% 30% 35% 40% 45% 50% 55% 60% 65% 70% 75% >80%

IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

<20% 25% 30% 35% 40% 45% 50% 55% 60% 65% 70% 75% >80%

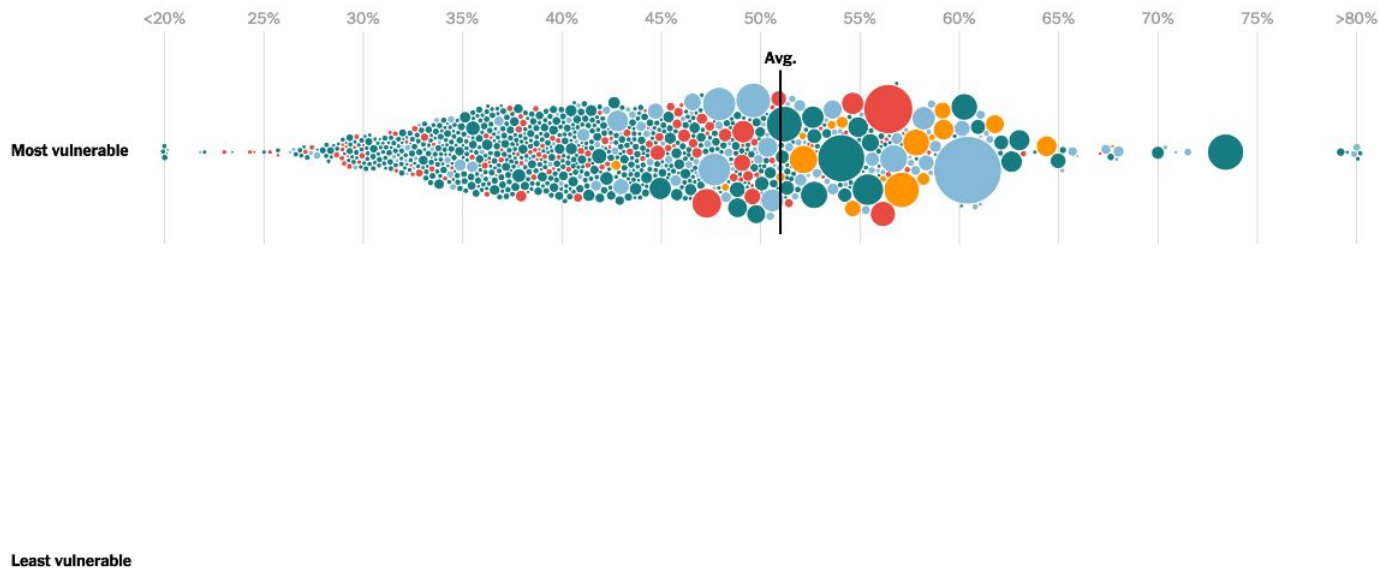
Most vulnerable

Least vulnerable

IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

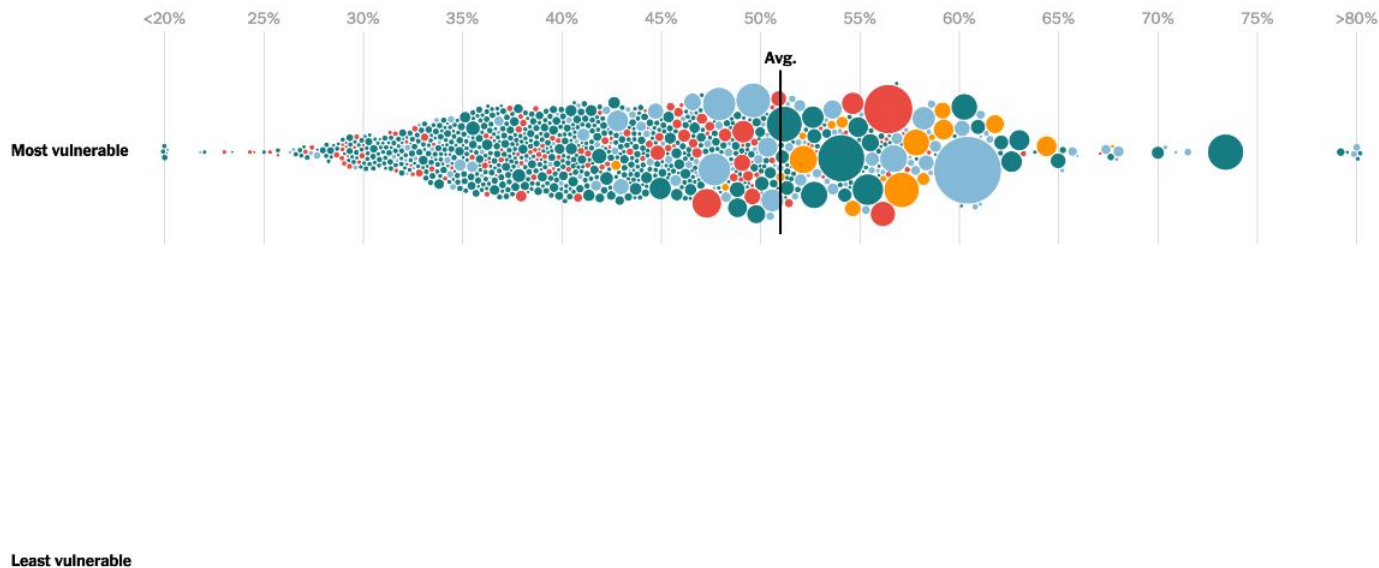


IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

Region ● Midwest ● Northeast ● South ● West

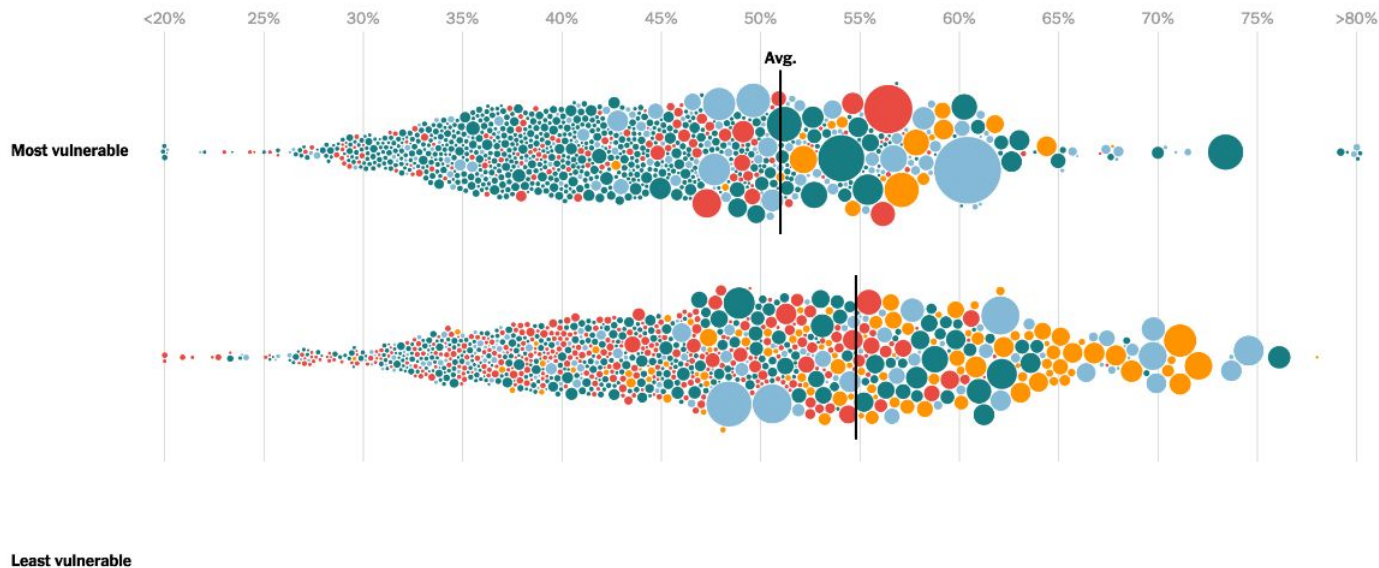


IMAGINED PATH

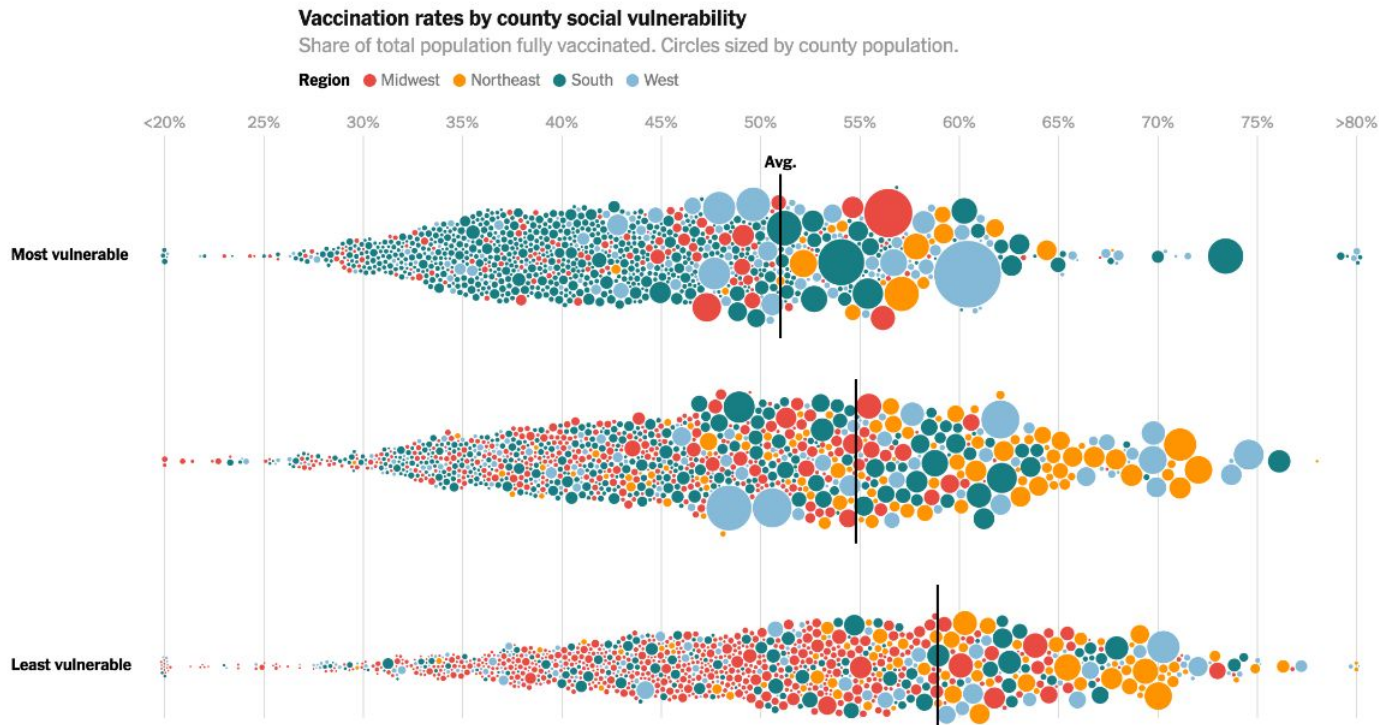
Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

Region ● Midwest ● Northeast ● South ● West



IMAGINED PATH

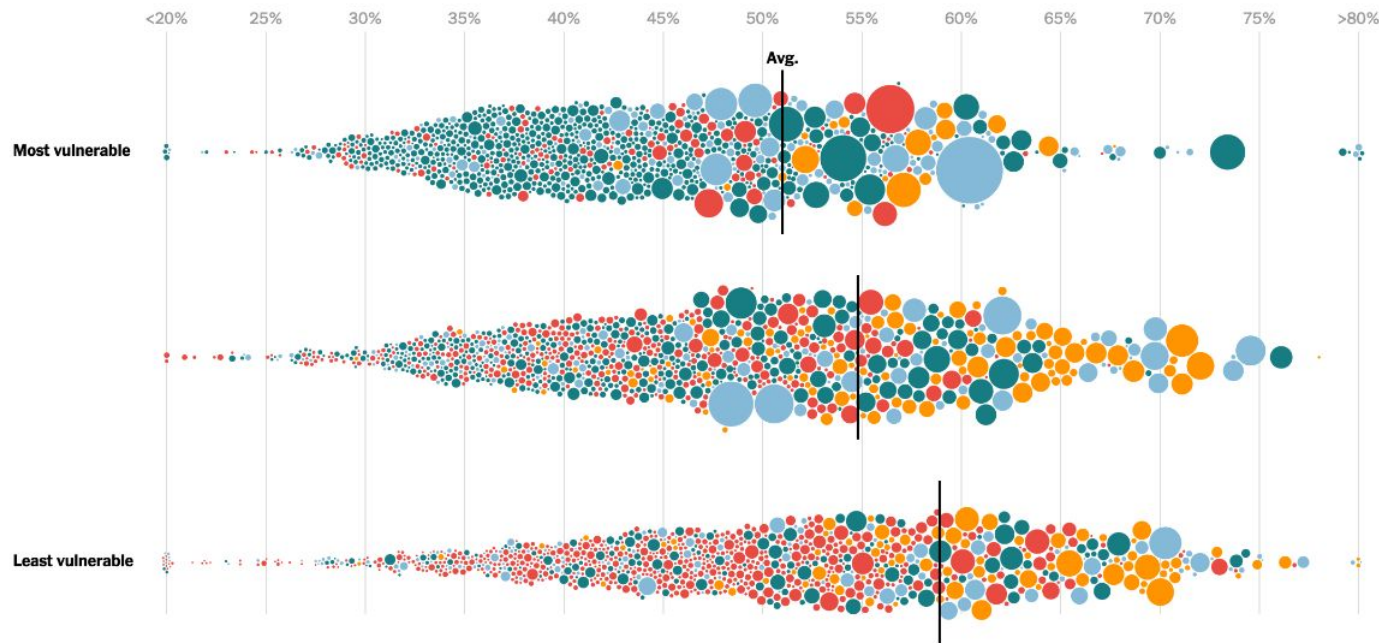


IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

Region ● Midwest ● Northeast ● South ● West



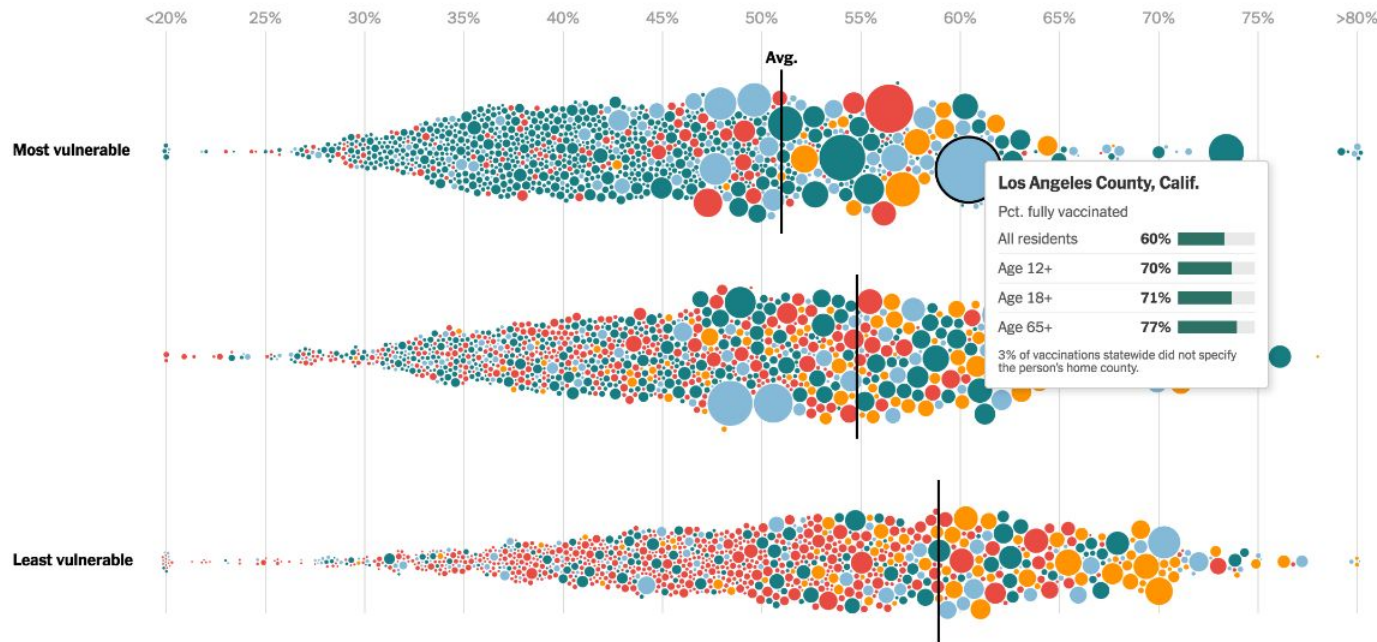
Sources: [Centers for Disease Control and Prevention](#); [Texas Department of State Health Services](#); [Colorado Department of Public Health & Environment](#); [Massachusetts Department of Public Health](#); U.S. Census Bureau | Note: No C.D.C. data available for Hawaii, Texas and some counties. Georgia, Vermont and West Virginia were excluded because more than a quarter of data is missing.

IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

Region ● Midwest ● Northeast ● South ● West



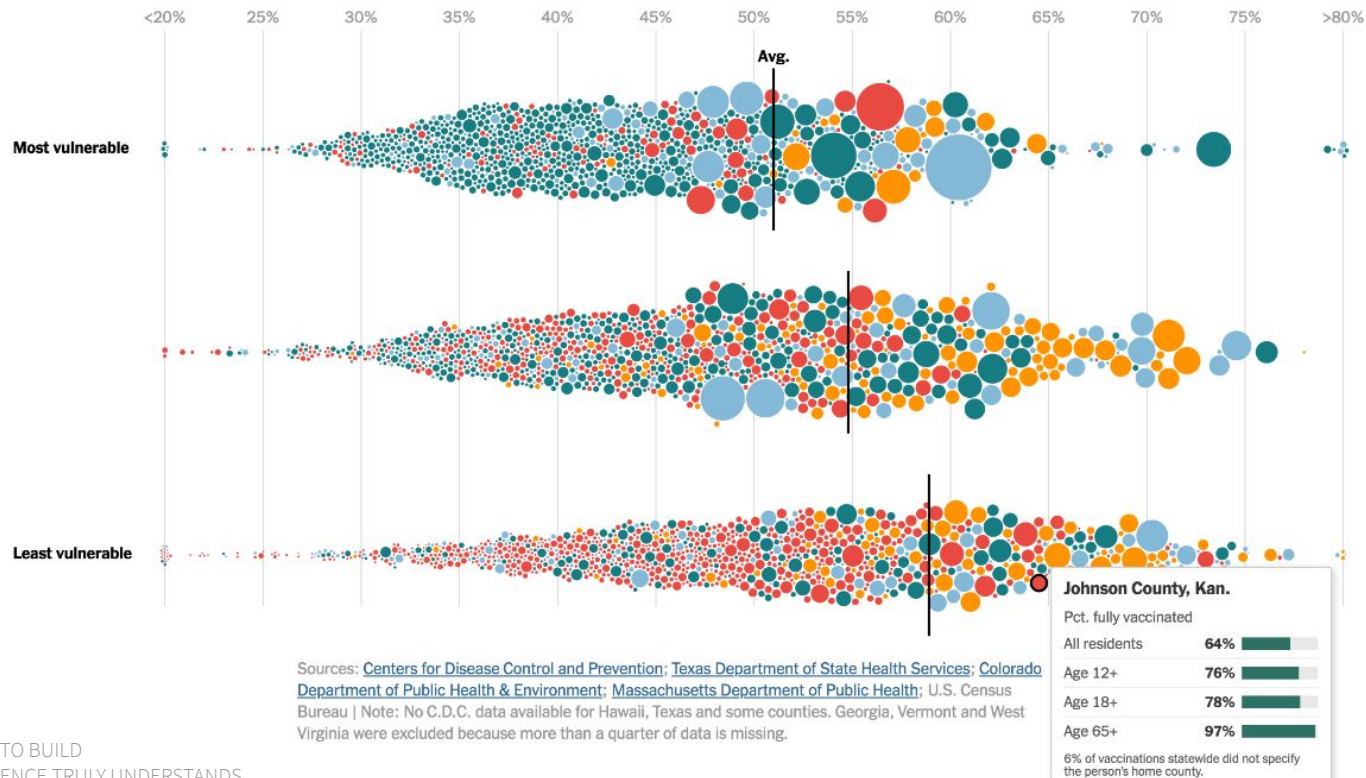
Sources: [Centers for Disease Control and Prevention](#); [Texas Department of State Health Services](#); [Colorado Department of Public Health & Environment](#); [Massachusetts Department of Public Health](#); U.S. Census Bureau | Note: No C.D.C. data available for Hawaii, Texas and some counties. Georgia, Vermont and West Virginia were excluded because more than a quarter of data is missing.

IMAGINED PATH

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

Region ● Midwest ● Northeast ● South ● West

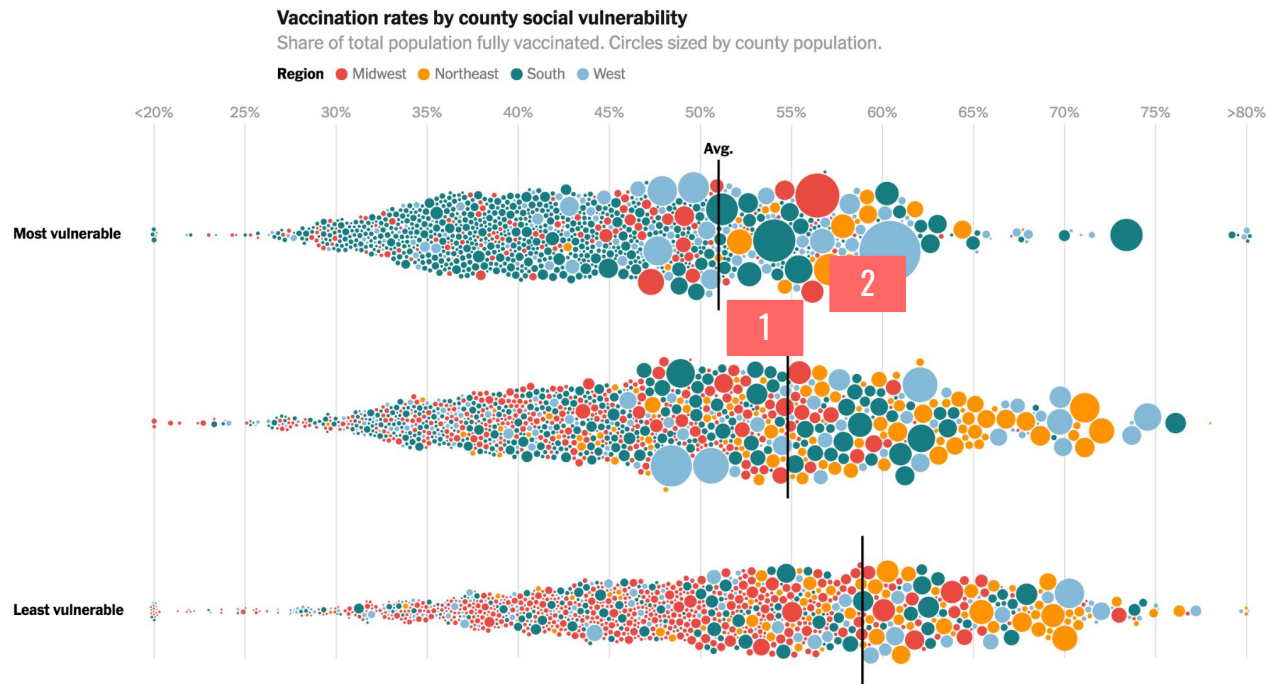


PATH

Some folks read all of the instructions from Ikea
before they start assembling their furniture.
But a lot of people just grab their allen wrench and dive in.
For these folks, we need **just in time information**.

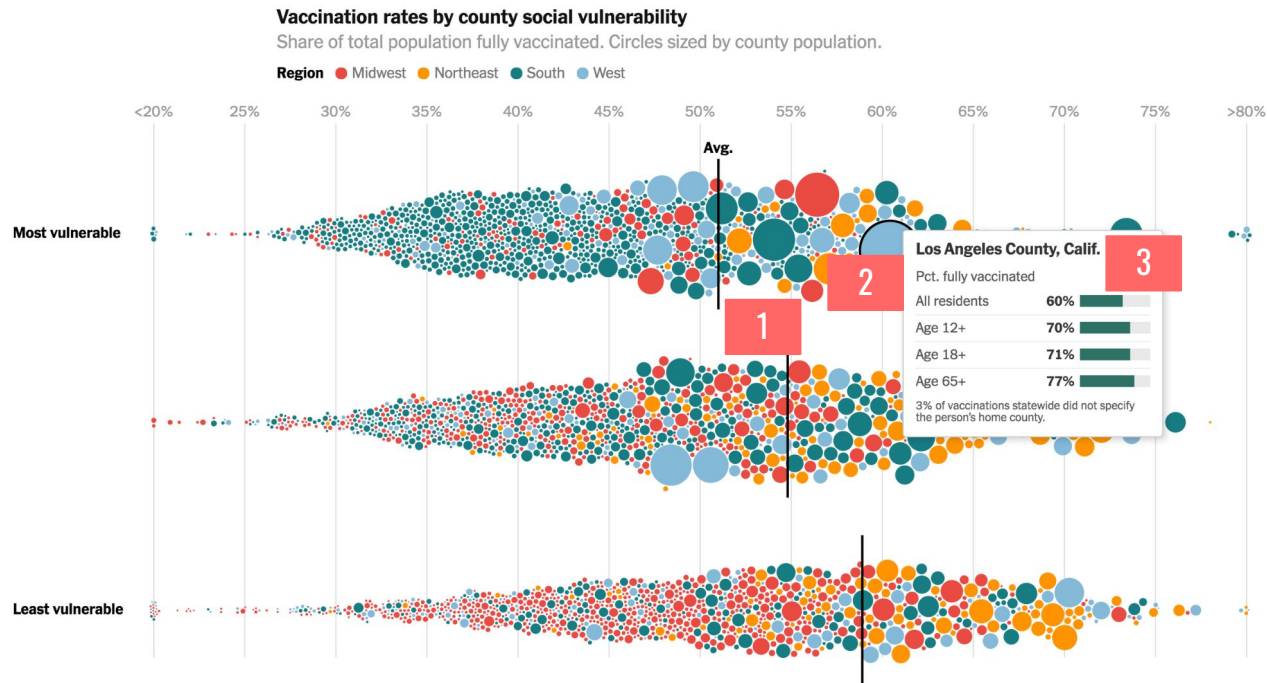


ACTUAL PATH



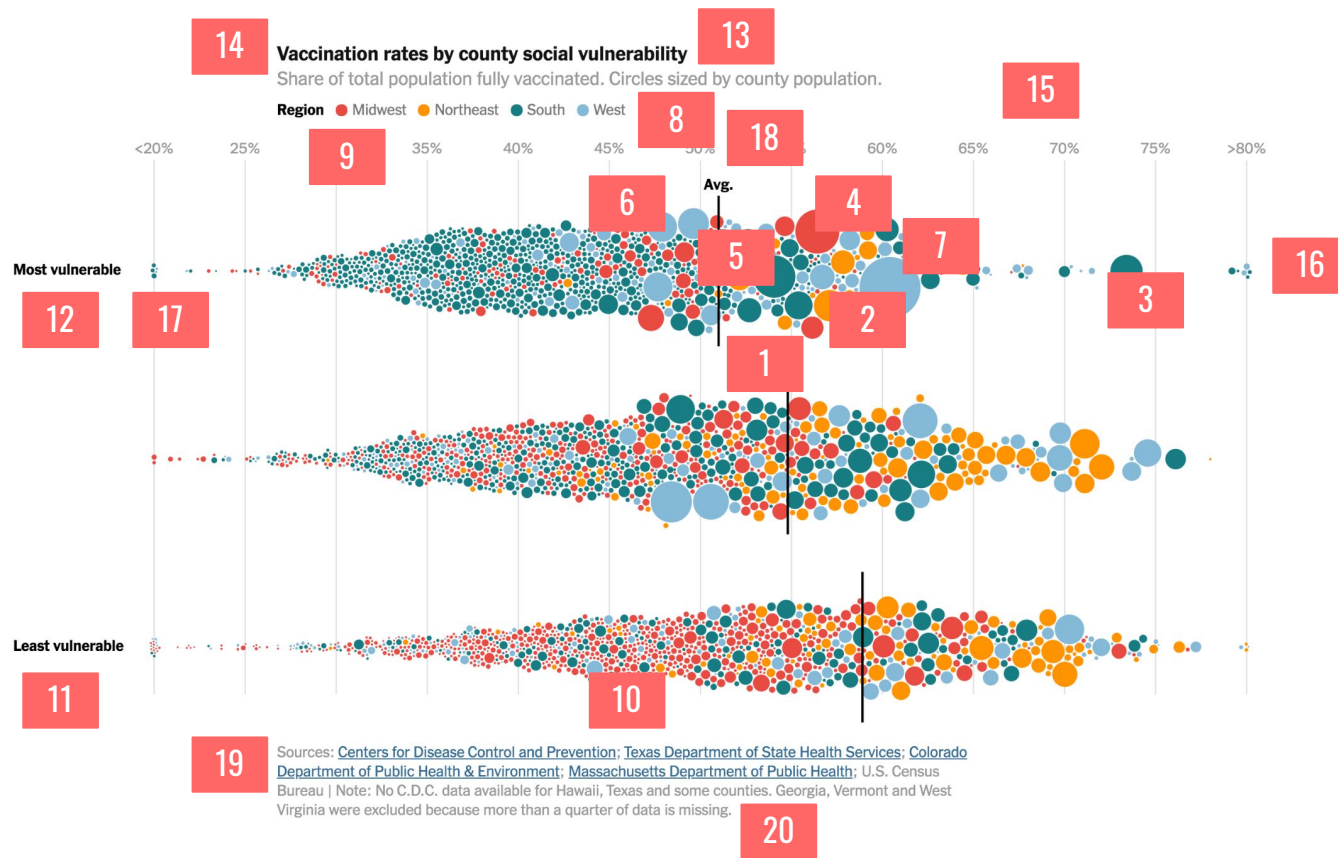
Sources: [Centers for Disease Control and Prevention](#); [Texas Department of State Health Services](#); [Colorado Department of Public Health & Environment](#); [Massachusetts Department of Public Health](#); U.S. Census Bureau | Note: No C.D.C. data available for Hawaii, Texas and some counties. Georgia, Vermont and West Virginia were excluded because more than a quarter of data is missing.

ACTUAL PATH

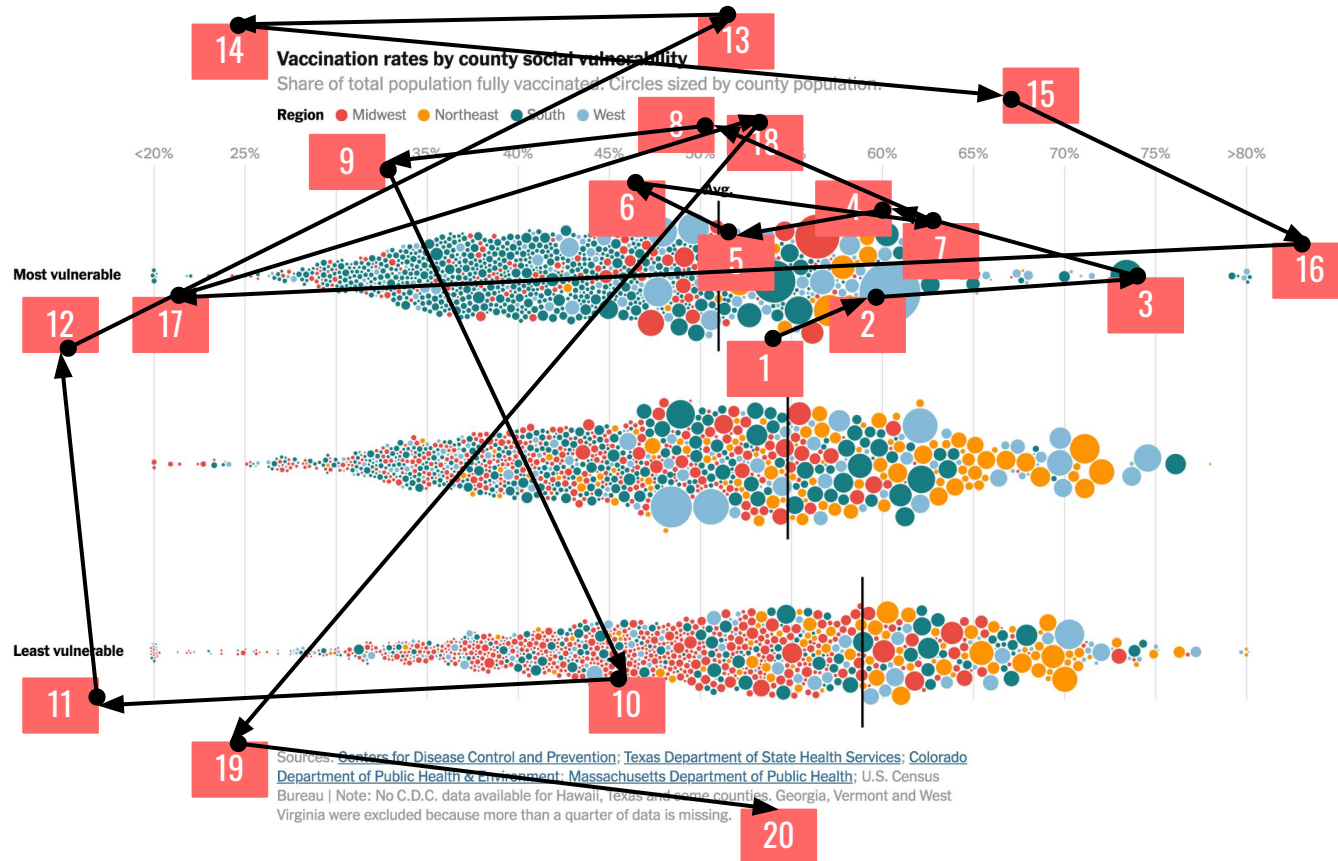


Sources: [Centers for Disease Control and Prevention](#); [Texas Department of State Health Services](#); [Colorado Department of Public Health & Environment](#); [Massachusetts Department of Public Health](#); U.S. Census Bureau | Note: No C.D.C. data available for Hawaii, Texas and some counties. Georgia, Vermont and West Virginia were excluded because more than a quarter of data is missing.

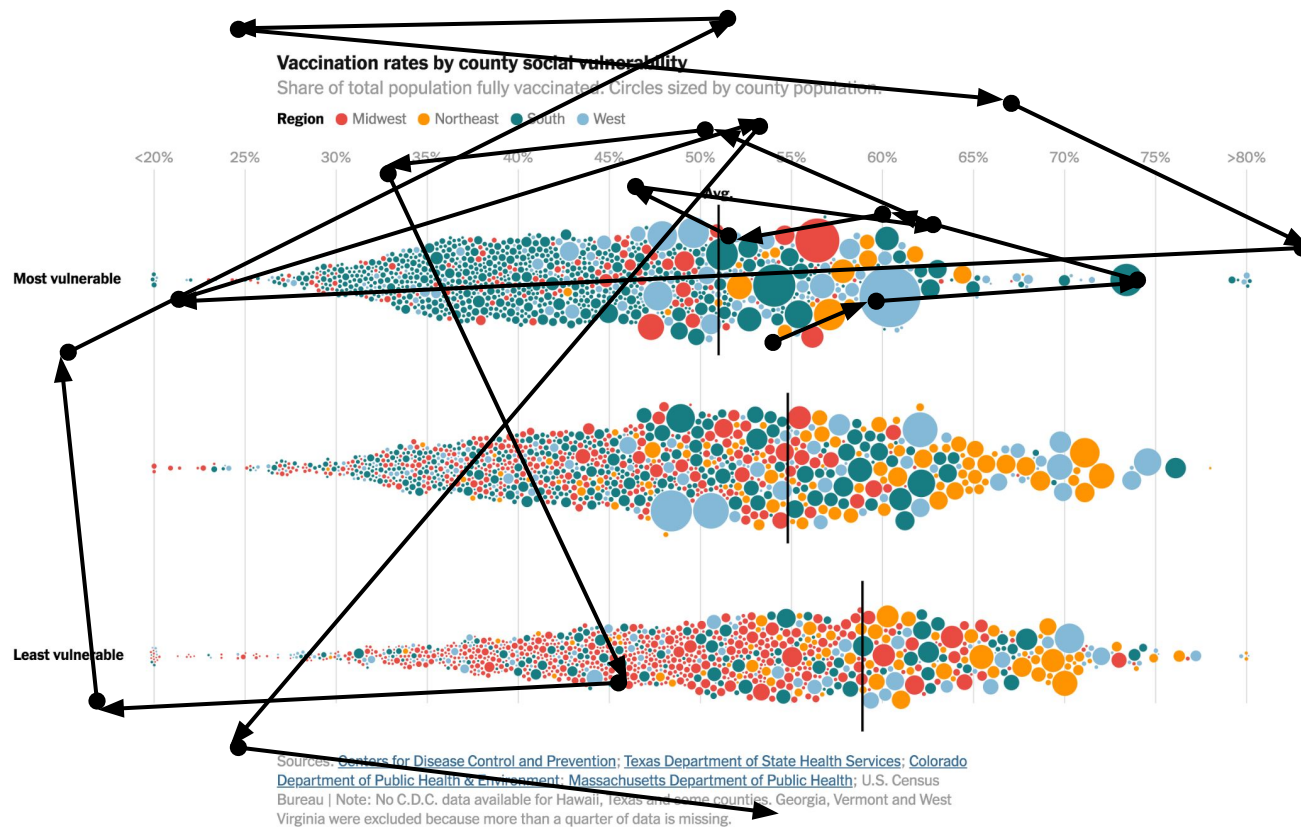
ACTUAL PATH



ACTUAL PATH



ACTUAL PATH



PINBALL MACHINES

2 Components

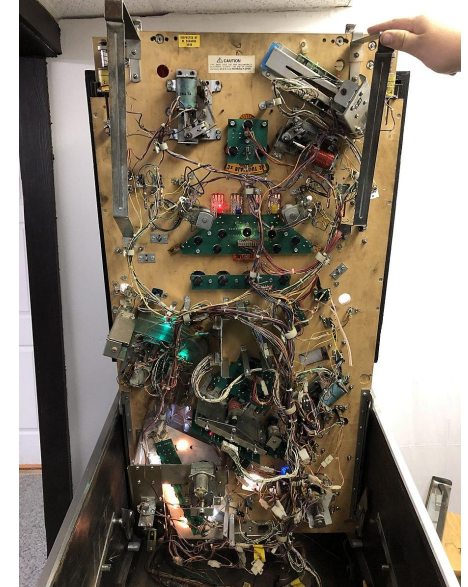
- 2.1 Cabinet
- 2.2 Backbox/Head
- 2.3 Lower Cabinet
- 2.4 Playfield
- 2.5 Plunger
- 2.6 Flippers
- 2.7 Bumpers
- 2.8 Kickers and slingshots
- 2.9 Targets
- 2.10 Holes and saucers
- 2.11 Spinners and rollovers
- 2.12 Switches, gates, and stoppers
- 2.13 Ramps
- 2.14 Toys, magnets and captive balls
- 2.15 Backglass



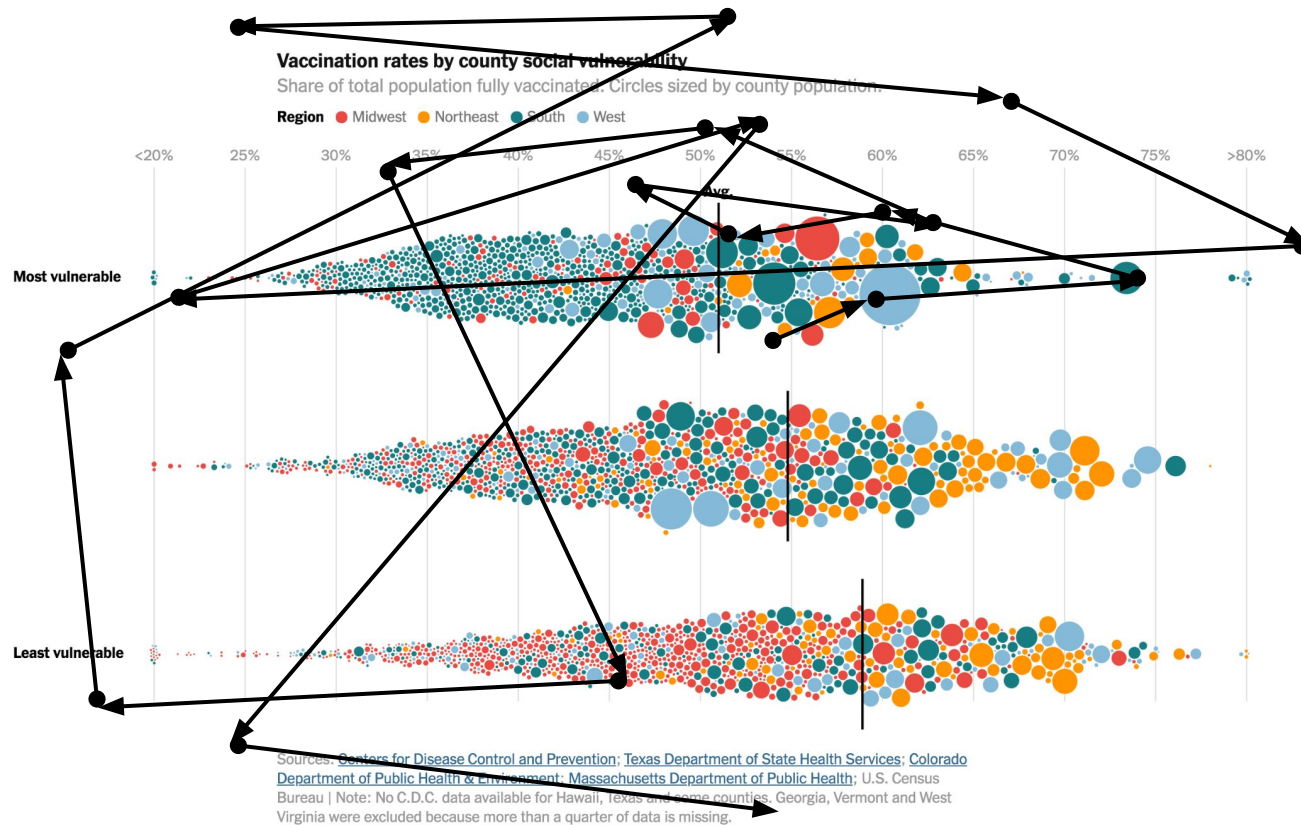
PINBALL MACHINES

2 Components

- 2.1 Cabinet
- 2.2 Backbox/Head
- 2.3 Lower Cabinet
- 2.4 Playfield
- 2.5 Plunger
- 2.6 Flippers
- 2.7 Bumpers
- 2.8 Kickers and slingshots
- 2.9 Targets
- 2.10 Holes and saucers
- 2.11 Spinners and rollovers
- 2.12 Switches, gates, and stoppers
- 2.13 Ramps
- 2.14 Toys, magnets and captive balls
- 2.15 Backglass



ACTUAL PATH

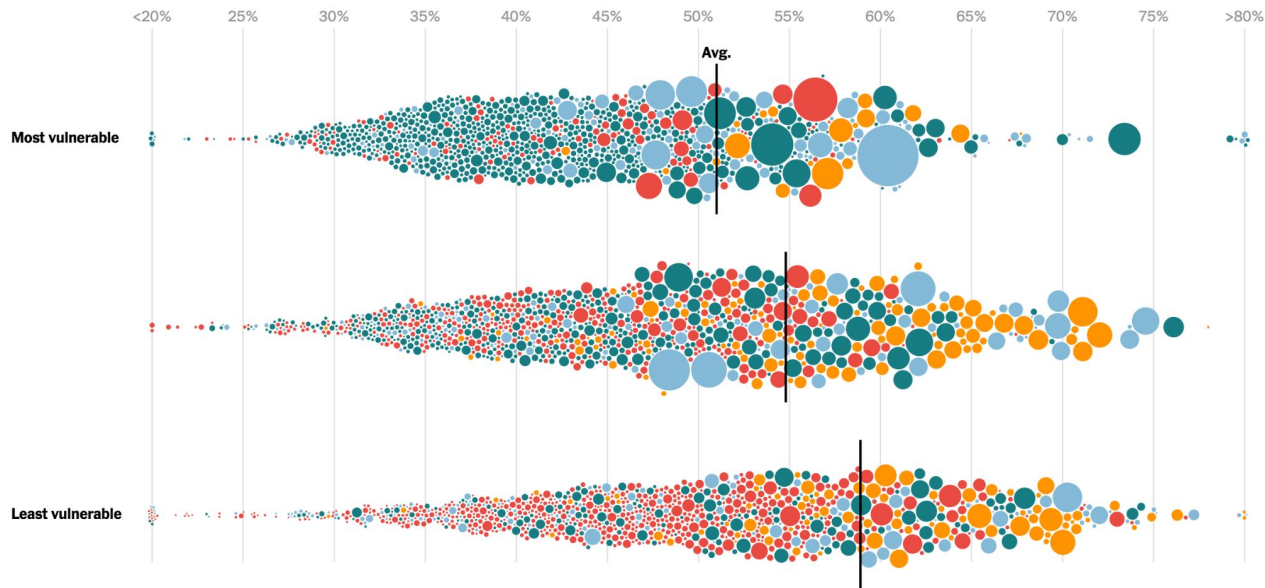


SCAFFOLDING

Vaccination rates by county social vulnerability

Share of total population fully vaccinated. Circles sized by county population.

Region ● Midwest ● Northeast ● South ● West



Sources: [Centers for Disease Control and Prevention](#); [Texas Department of State Health Services](#); [Colorado Department of Public Health & Environment](#); [Massachusetts Department of Public Health](#); U.S. Census Bureau | Note: No C.D.C. data available for Hawaii, Texas and some counties. Georgia, Vermont and West Virginia were excluded because more than a quarter of data is missing.

WHERE WE TEND TO GO FIRST

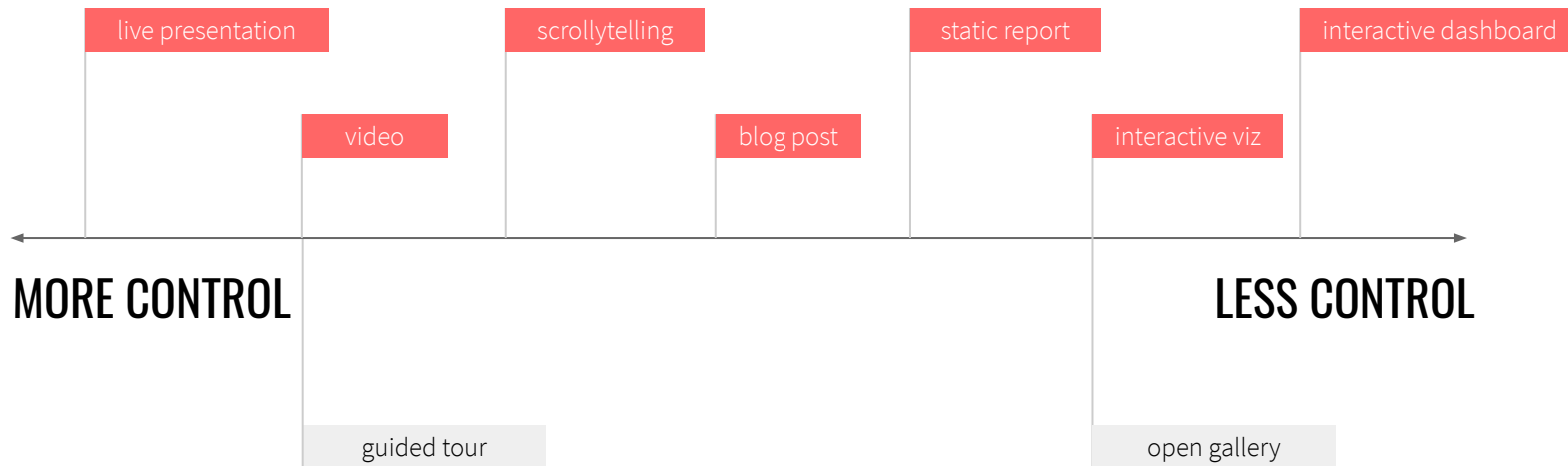
- ▶ Outliers
- ▶ Patterns
- ▶ Me
- ▶ Color / Weight
- ▶ Embedded Writing

COMPONENTS TO CONSIDER

- ▶ Title
- ▶ Subtitle
- ▶ Axes
- ▶ Legends
- ▶ Keys
- ▶ Reference Lines
- ▶ Annotations
- ▶ Popup Boxes
- ▶ Footnotes
- ▶ Instructions

PATH CURATION

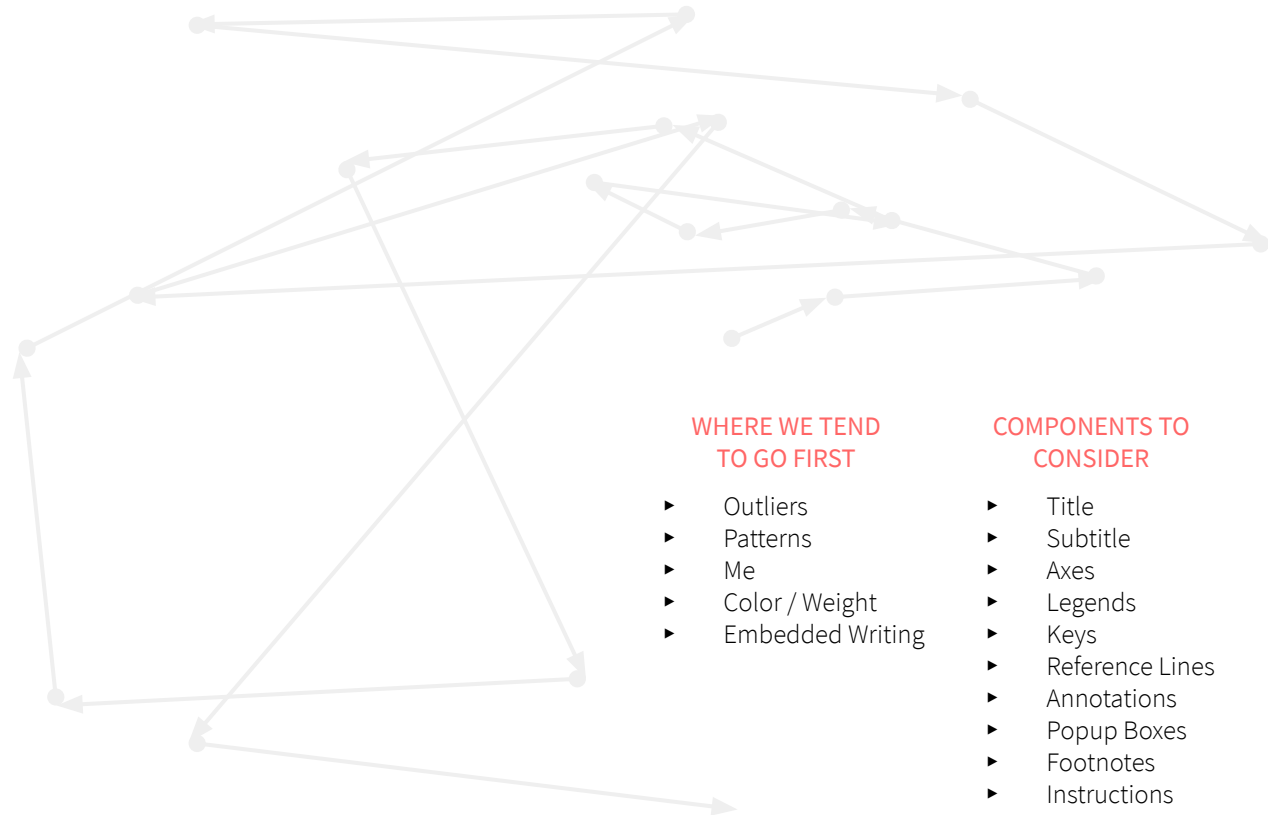
It's a balance between the designer's direction and the audience's agency.



DISCUSSION QUESTIONS

- ▶ Where are likely **starting points** in your work?
- ▶ Where do you want to **direct** your audience?
- ▶ Which components are essential to provide **just in time** information for your audience?
- ▶ How much **control** do you have in directing your audience's path?
- ▶ Do you feel any **hesitation** in actively curating the path for your audience?

PINBALL



WORKSHOP 1

Today we've got four exercises to improve our work's functional design.

PURPOSE

MADLIB

Define the intention of your work.



CHARTS

FLOWCHART

Match your data to the right chart type.



PATH

PINBALL

Curate your audience's journey.



FIT

SPECTRUMS

Calibrate your work's cognitive load.



REVISIONS

When we ask ourselves, “How can we make our data visualizations better?”, the answer is usually some derivative of, “Well, **it depends** on the visualization.”



COGNITIVE LOAD

How easy or hard is it to make sense of something new?

DATA

INTRINSIC LOAD

the new information's inherent complexity



AUDIENCE

GERMANE LOAD

the audience's familiarity (or lack thereof)



VISUALIZATION

EXTRANEIOUS LOAD

how the new information is presented



COGNITIVE LOAD

simpler information + knowledgeable audience + pared-down delivery = lighter cognitive load

complex information + novice audience + detailed delivery = heavier cognitive load



COGNITIVE LOAD

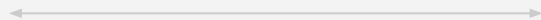
No visualization should have a
heavy cognitive load because it's poorly made.
But plenty of well-made visualizations carry
a heavier cognitive load.



DATA

How simple or complex is your data?

MEASUREMENT



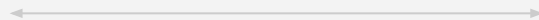
QUANTITATIVE

quarterly earnings
patient's heart rate
country's coastline

QUALITATIVE

company culture
patient's pain level
country's democratic health

KNOWABILITY



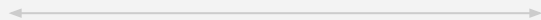
CERTAIN

NBA win-loss records
flu vaccines administered
number of Target stores

UNCERTAIN

NBA score predictions
cases of the flu
people inside a Target now

SPECIFICITY



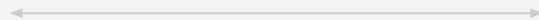
PRECISE

zip codes
blood type
standardized test scores

AMBIGUOUS

neighborhoods
gender
letter grades

RELATABILITY



CONCRETE

cost of lunch
plastic in a yogurt container
your age

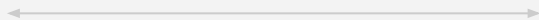
ABSTRACT

gross domestic product
all plastic ever produced
the age of the earth

AUDIENCE

How much bandwidth does your audience have?

CONNECTION



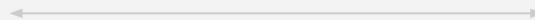
INTENTIONAL

subscriber
I have a ticket
this is my jam!

COINCIDENTAL

social media feed
conference walk-ins
hmm...this could be cool?

PACE



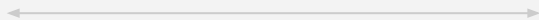
SLOW

what can I discover?
note-taking
desktop-oriented

FAST

what's the takeaway?
skimming
mobile-oriented

KNOWLEDGE



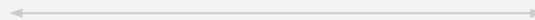
EXPERT

this is similar to that
have you considered...?
let's see the details

NOVICE

this is brand new
where do I start?
what's the big picture?

CONFIDENCE



CONFIDENT

I know the routine
this button should do this
I'm excited to see the data

ANXIOUS

what will happen next?
will I break this?
I'm frustrated/embarrassed

VISUALIZATION

What visualization best connects your audience to your data?

CHART TYPE

COMMON

bar charts, line charts, maps
familiar axes
one or two variables

RARE

treemaps, distributions, etc.
unusual axes (or none at all)
multiple variables

INTERPRETATION

ACCURATE

length, position, shape
absolute values
2-4 colors

APPROXIMATE

angles, area, volume
relational values
5+ colors

COMPOSITION

CONCISE

minimal data marks
key takeaways
context outside the graph

DETAILED

densely-encoded marks
disaggregated details
context on the graph

DELIVERY

EXPLANATORY

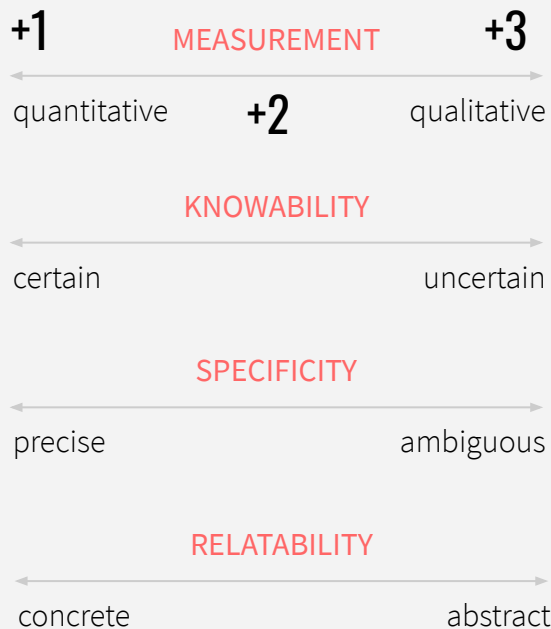
deliberate sequence
prepackaged thesis
created by domain expert

EXPLORATORY

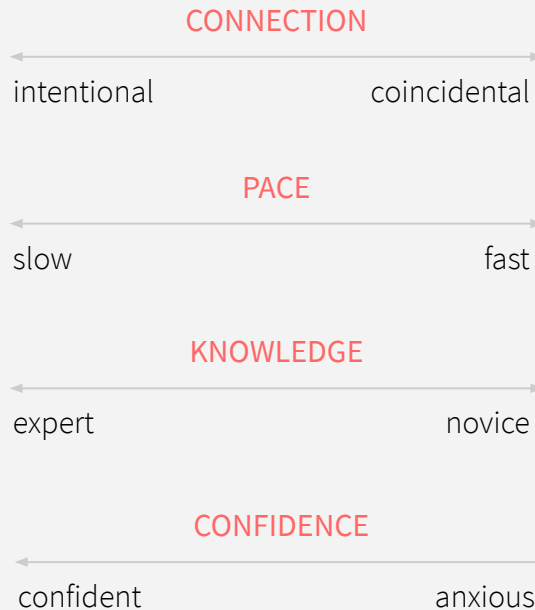
built for discovery
audience adds knowledge
created by viz expert



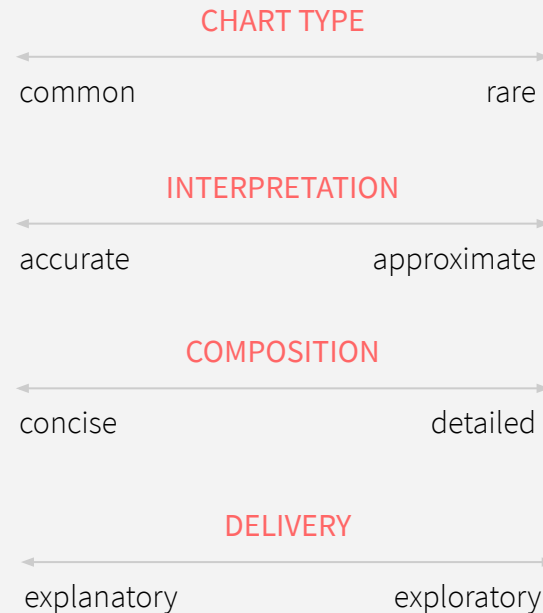
DATA



AUDIENCE



VISUALIZATION



ERIN WALDRON FROM DATA DOZEN · CONTACT@DATADOZEN.COM · 913-303-1125

THANK YOU!

WWW.DATADOZEN.COM/HARVARD-SDP-2021

