

A PRESENTATION FROM

DATA DOZEN

a visual analytics studio

ERIN WALDRON FROM DATA DOZEN

DATA COMMUNICATION

THE 5 MOST COMMON MISTAKES WE MAKE WHEN COMMUNICATING WITH DATA



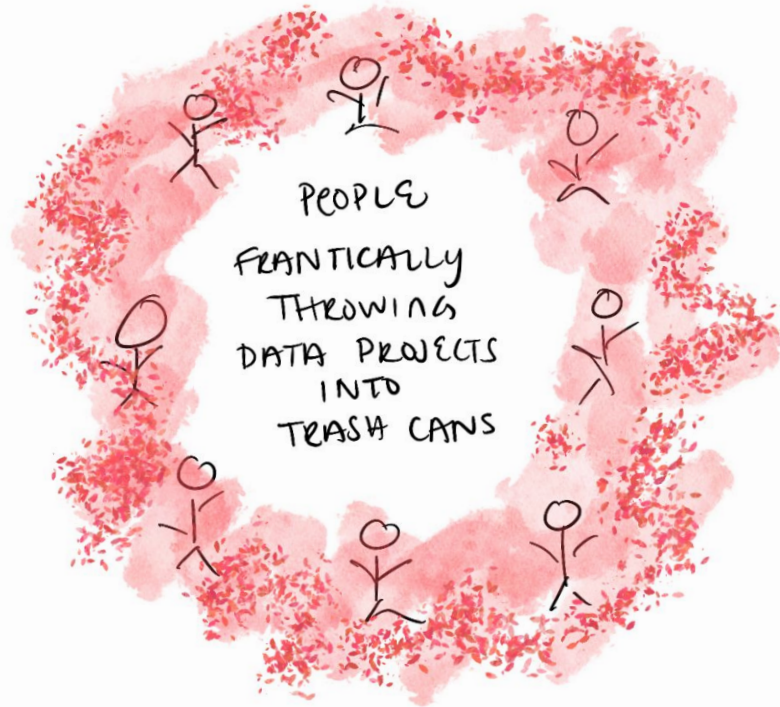
DISCLAIMER

I have **personally** made every mistake here (many times).



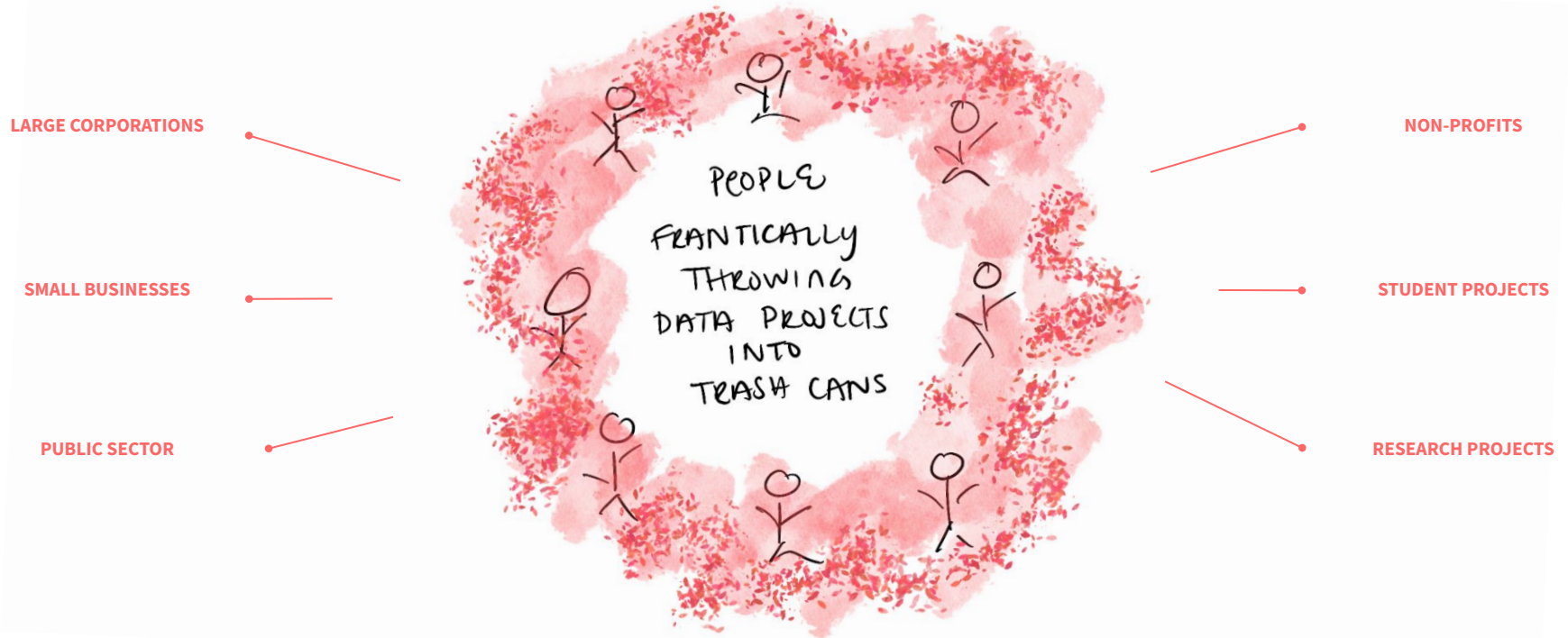
7 YEARS AS A DATA CONSULTANT

Here's what the majority of organizational data efforts look like:



THE BIG PREDICTABLE PROBLEM

Here's what the majority of organizational data efforts look like:

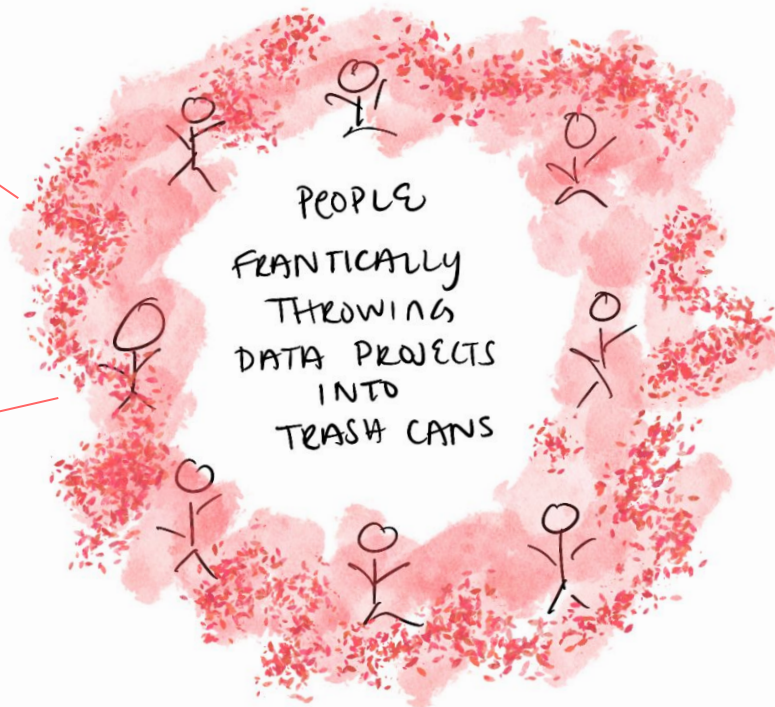


THE BIG PREDICTABLE PROBLEM

Obviously no team or manager would say that *this* is their goal.

“WE’RE CREATING MORE
TRANSPARENCY &
ACCOUNTABILITY!”

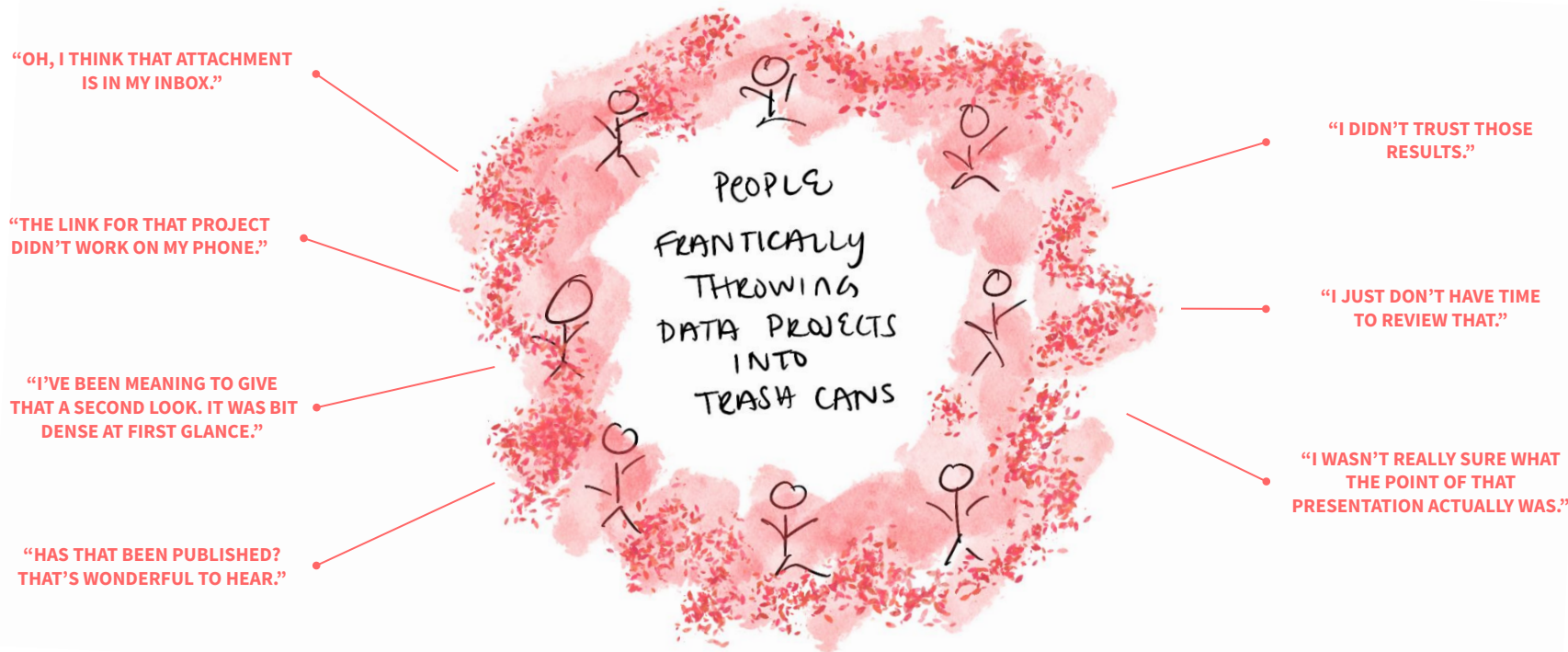
“WE’RE PROVIDING THE
ORGANIZATION WITH
ACTIONABLE INSIGHTS!”



“WE’RE EMPOWERING OUR
LEADERSHIP TO MAKE DATA
DRIVEN DECISIONS!”

THE BIG PREDICTABLE PROBLEM

Hey stakeholders: “How has the data impacted your thinking or action?”



THE BIG PREDICTABLE PROBLEM

Are you reading between the lines?

TRASH CAN

TRASH CAN

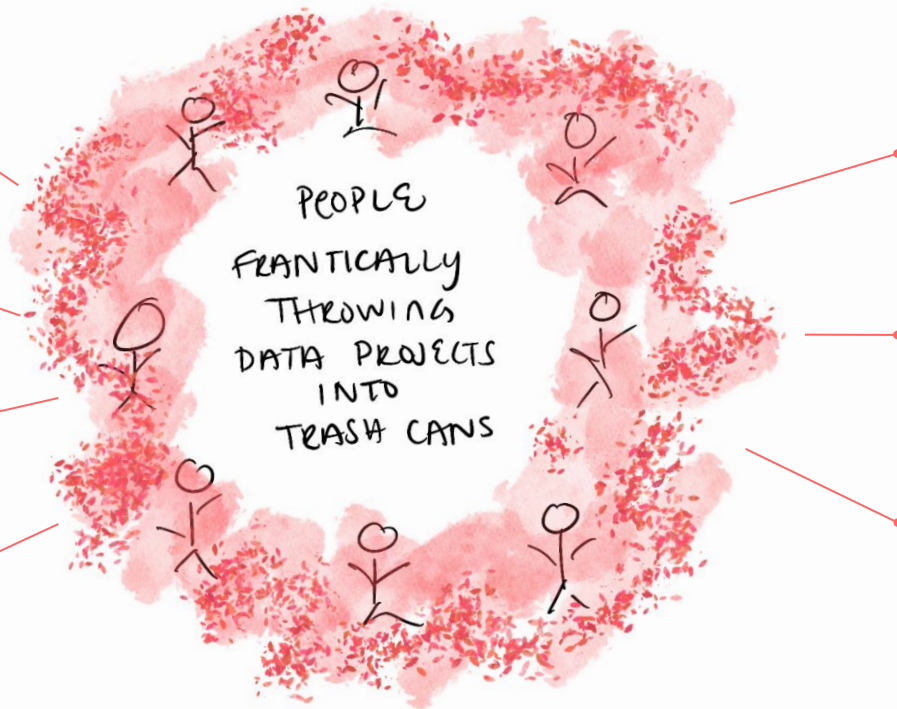
TRASH CAN

TRASH CAN

TRASH CAN

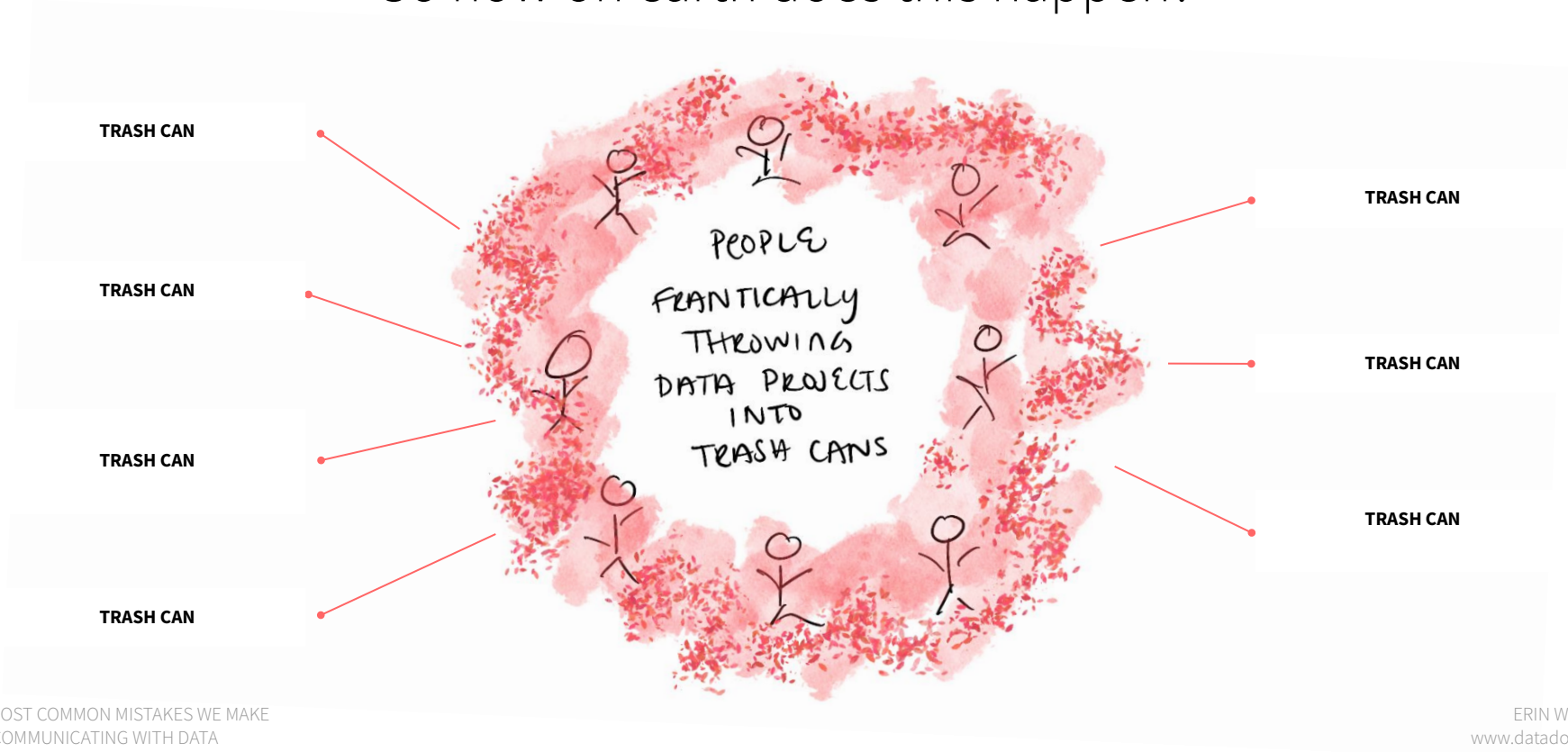
TRASH CAN

TRASH CAN



THE BIG PREDICTABLE PROBLEM

So how on earth does this happen?



a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE
PURPOSE

Here's what this step is about.



PROBLEM SOLVING
QUESTIONS

We're often asking things like:

a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE
PURPOSE

Here's what this step is about.

- ▷ Gathering is about the creation or identification of the dataset.
- ▷ Let's demystify what I actually mean when I say "data."



PROBLEM SOLVING
QUESTIONS

We're often asking things like:

TIDY DATA

In his article “Tidy Data,” Hadley Wickham explains how we organize data so computers can read it.

ONE COLUMN PER VARIABLE

DIMENSIONS, MEASURES, FIELDS



ONE ROW PER OBSERVATION

RECORDS, TRANSACTIONS, EVENTS



ONE TABLE PER OBSERVATION TYPE

CONSISTENT GRANULARITY



TIDY DATA

Imagine a table in Excel.

TIDY DATA

One column per variable.

TIDY DATA

One column per variable.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE

TIDY DATA

One row per observation.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE

TIDY DATA

One row per observation.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE
11/1/23	9:52 AM	Manager	Me	Great report!

TIDY DATA

One row per observation.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE
11/1/23	9:52 AM	Manager	Me	Great report!
11/1/23	2:14 PM	Colleague A	Me	Meet tomorrow

TIDY DATA

One row per observation.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE
11/1/23	9:52 AM	Manager	Me	Great report!
11/1/23	2:14 PM	Colleague A	Me	Meet tomorrow
11/1/23	2:46 PM	Me	Colleague A	Re: Meet tomorrow

TIDY DATA

One row per observation.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE
11/1/23	9:52 AM	Manager	Me	Great report!
11/1/23	2:14 PM	Colleague A	Me	Meet tomorrow
11/1/23	2:46 PM	Me	Colleague A	Re: Meet tomorrow
11/1/23	2:59 PM	Me	Manager	Re: Great report!

TIDY DATA

One row per observation.

DATE	TIME	SENDER	RECIPIENT	SUBJECT LINE
11/1/23	9:52 AM	Manager	Me	Great report!
11/1/23	2:14 PM	Colleague A	Me	Meet tomorrow
11/1/23	2:46 PM	Me	Colleague A	Re: Meet tomorrow
11/1/23	2:59 PM	Me	Manager	Re: Great report!
11/1/23	4:59 PM	Colleague B	Me	New Deadline

TIDY DATA

One table per observational unit.

TABLE A

DATE	TIME	SENDER	RECIPIENT	SUBJECT
11/1/23	9:52 AM	Manager	Me	Great report!
11/1/23	2:14 PM	Colleague A	Me	Meet tomorrow
11/1/23	2:46 PM	Me	Colleague A	Re: Meet tomorrow
11/1/23	2:59 PM	Me	Manager	Re: Great report!
11/1/23	4:59 PM	Colleague B	Me	New Deadline

TABLE B

YEAR	QUARTER	DEPARTMENT	DAILY EMAIL AVG
2023	Q3	Marketing	35
2023	Q4	Marketing	47
2023	Q1	Analytics	23

IS THAT IT? THAT'S DATA?

Yes, that's all **data** fundamentally is.



a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE PURPOSE

Here's what this step is about.

- Gathering is about the creation or identification of the dataset.
- Data is simply individual events captured in columns and rows in a table.



PROBLEM SOLVING QUESTIONS

We're often asking things like:

- What data do I need? What variables at what granularity?
- Does this data already exist? Is it trustworthy?
- If not, how will I collect the data? Measure what matters.
- Where will I store the data?

a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE PURPOSE

Here's what this step is about.

- Access is about getting the dataset to your computer.
- If you made the dataset, this step is practically nonexistent.
- But if a large organization has the data, this step can get tricky.



PROBLEM SOLVING QUESTIONS

We're often asking things like:

- Data storage: Where is the data?
- Data governance: Am I allowed to access it? At what granularity?
- Data security: How do I securely access the data?
- Data querying: Can I access just the data I need for this project?

a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE PURPOSE

Here's what this step is about.

- Cleaning is about preparing the data for computer analysis.
- Lisa Gitelman quote: "Raw data is an oxymoron."
- Data is just as fallible as any form of human knowledge.



PROBLEM SOLVING QUESTIONS

We're often asking things like:

- Content cleaning: How do I standardize and correct what's inside the cells for analysis?
- Structural cleaning: How do I organize the cells to make the file readable for computers?
- Automation: Can I be done yet?

a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE PURPOSE

Here's what this step is about.

- Finally! It's time to analyze the data!
- Analyzing is all about exploring the data and see what it can teach you.



PROBLEM SOLVING QUESTIONS

We're often asking things like:

- What tool will I use? (Excel, Tableau, Power BI, R, SPSS, etc.)
- Descriptive statistics: sum, avg, distributions, categories, etc.
- Inferential statistics: What can I infer about the broader population? What's causal?

a typical data
project process

GATHER

ACCESS

CLEAN

ANALYZE

COMMUNICATE



BIG PICTURE
PURPOSE

Here's what this step is about.

- ▷ Communication is all about sharing what you've learned with your intended audience.
- ▷ This is usually where the data project breaks down...



PROBLEM SOLVING
QUESTIONS

We're often asking things like:

- ▷
- ▷
- ▷
- ▷ ... because we feel like we don't have any time left to ask good questions.

THE TYPICAL DATA PROCESS

Here's the catch. These five steps are not created equal.

THE TYPICAL DATA PROCESS

Here's the catch. These five steps are not created equal.

100% —

—

—

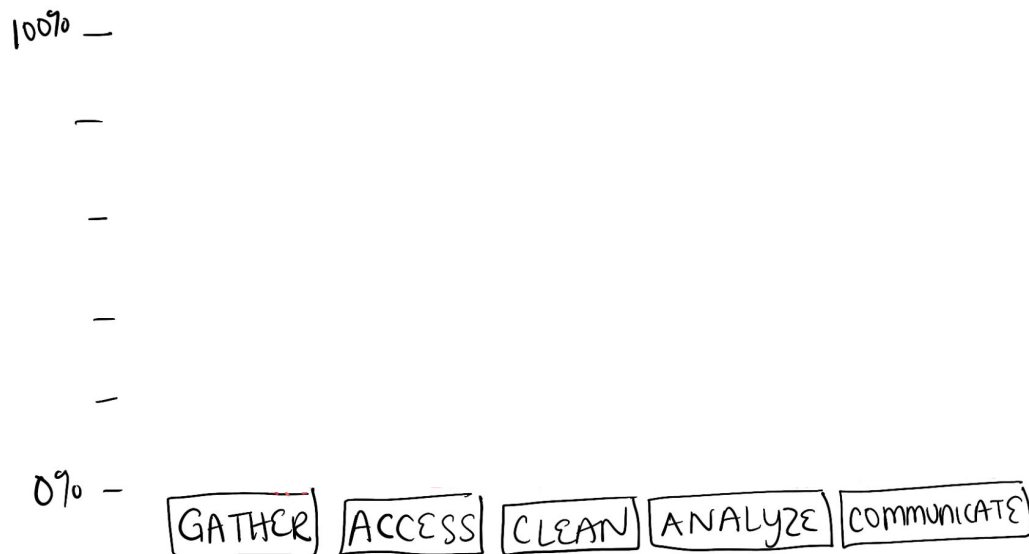
—

—

0% —

THE TYPICAL DATA PROCESS

Here's the catch. These five steps are not created equal.



THE TYPICAL DATA PROCESS

Here's the catch. These five steps are not created equal.



THE TYPICAL DATA PROCESS

Different circumstances will have different time distributions.



THE TYPICAL DATA PROCESS

The one consistent similarity is minimal time spent on communication.



THE TYPICAL DATA PROCESS

Why? Because the thinking usually ends like this.



THE TYPICAL DATA PROCESS

Why? Because the thinking usually ends like this.



THE TYPICAL DATA PROCESS

Why? Because the thinking usually ends like this.



THE TYPICAL DATA PROCESS

Why? Because the thinking usually ends like this.



THE TYPICAL DATA PROCESS

Why? Because the thinking usually ends like this.



THE TYPICAL DATA PROCESS

Why? Because the thinking usually ends like this.



THE BIG PREDICTABLE PROBLEM

The communication step is usually the **last step** in a data project, when we're most stressed for time.



THE BIG PREDICTABLE PROBLEM

So we either skip the communication step altogether,
or we **rush** through it to call it done.



THE BIG PREDICTABLE PROBLEM

And when we rush through the communication step,
we make very predictable **mistakes**.
And our communication fails.



EVEN IF THE FIRST FOUR
STEPS WERE DONE WELL

EVEN IF THE WORK IS
GROUNDBREAKING

EVEN IF THERE'S AN AUDIENCE
WAITING FOR THE DATA

THIS IS THE WHOLE POINT

If the communication step fails,
the **entire** data project fails.

DATA PROJECTS EXIST TO
CONNECT THE DATA & AUDIENCE

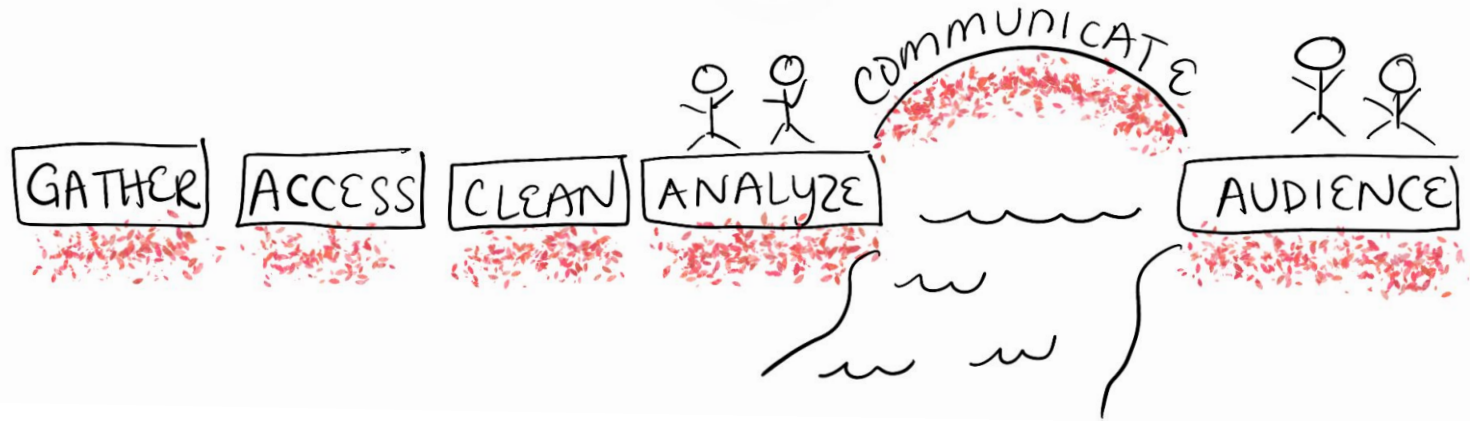
SO A DATA PROJECT'S ROI
HINGES ON COMMUNICATION

LET'S FOCUS ON HOW TO DO
THIS FINAL STEP WELL



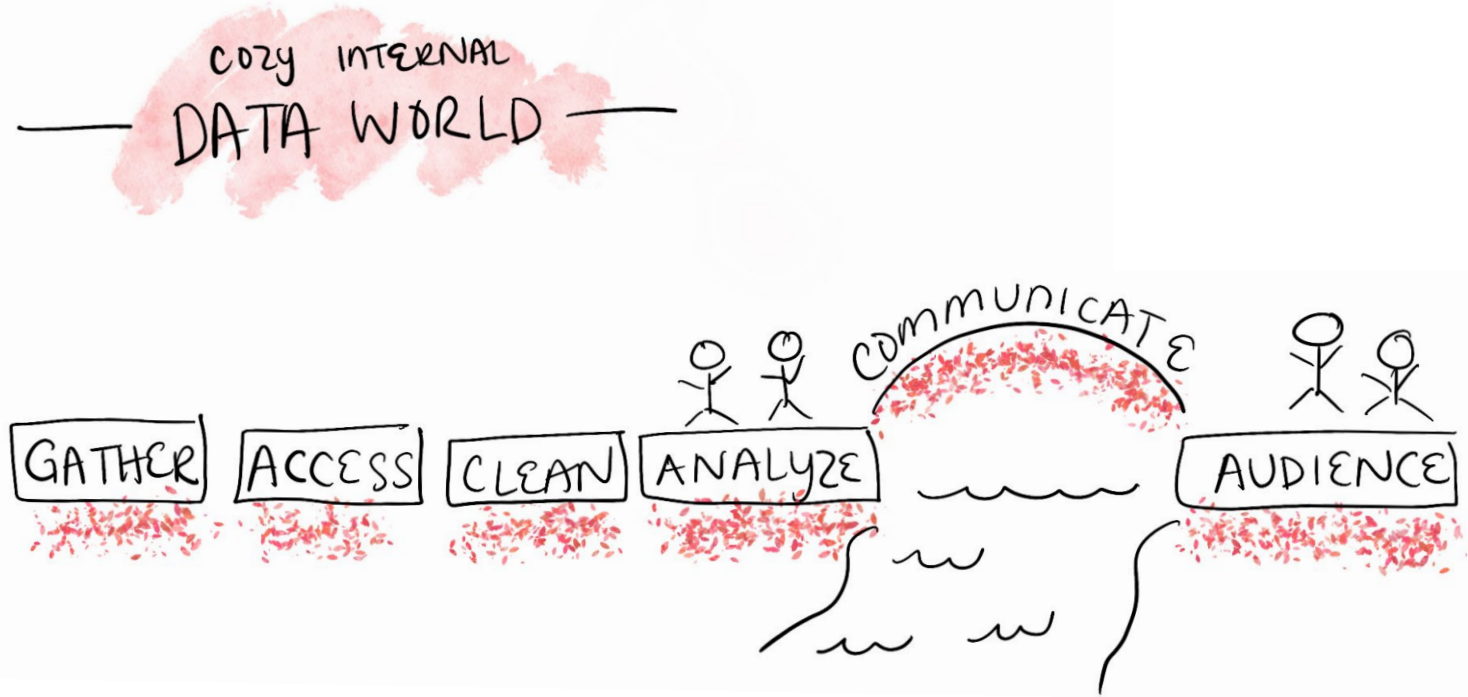
EFFECTIVE DATA COMMUNICATION

It's the bridge that connects the data and the audience.



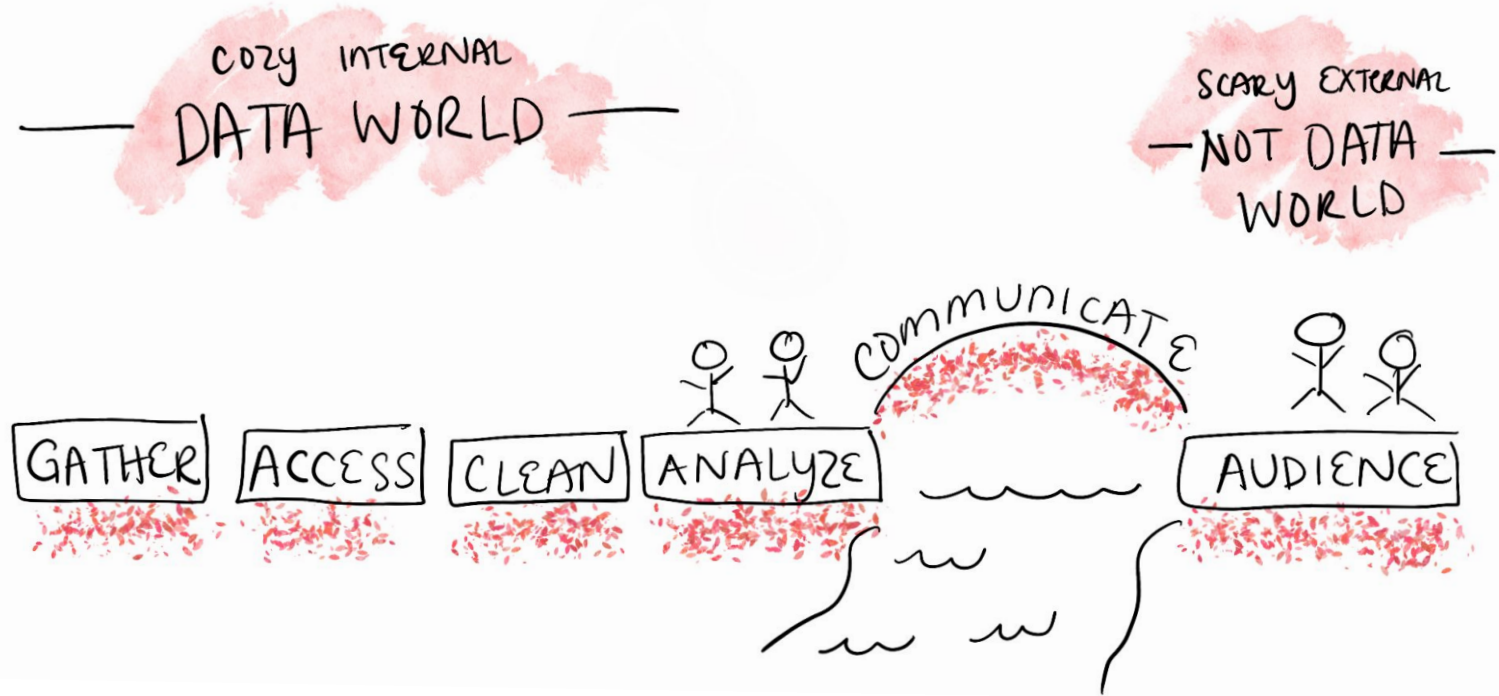
EFFECTIVE DATA COMMUNICATION

It's the bridge that connects the data and the audience.



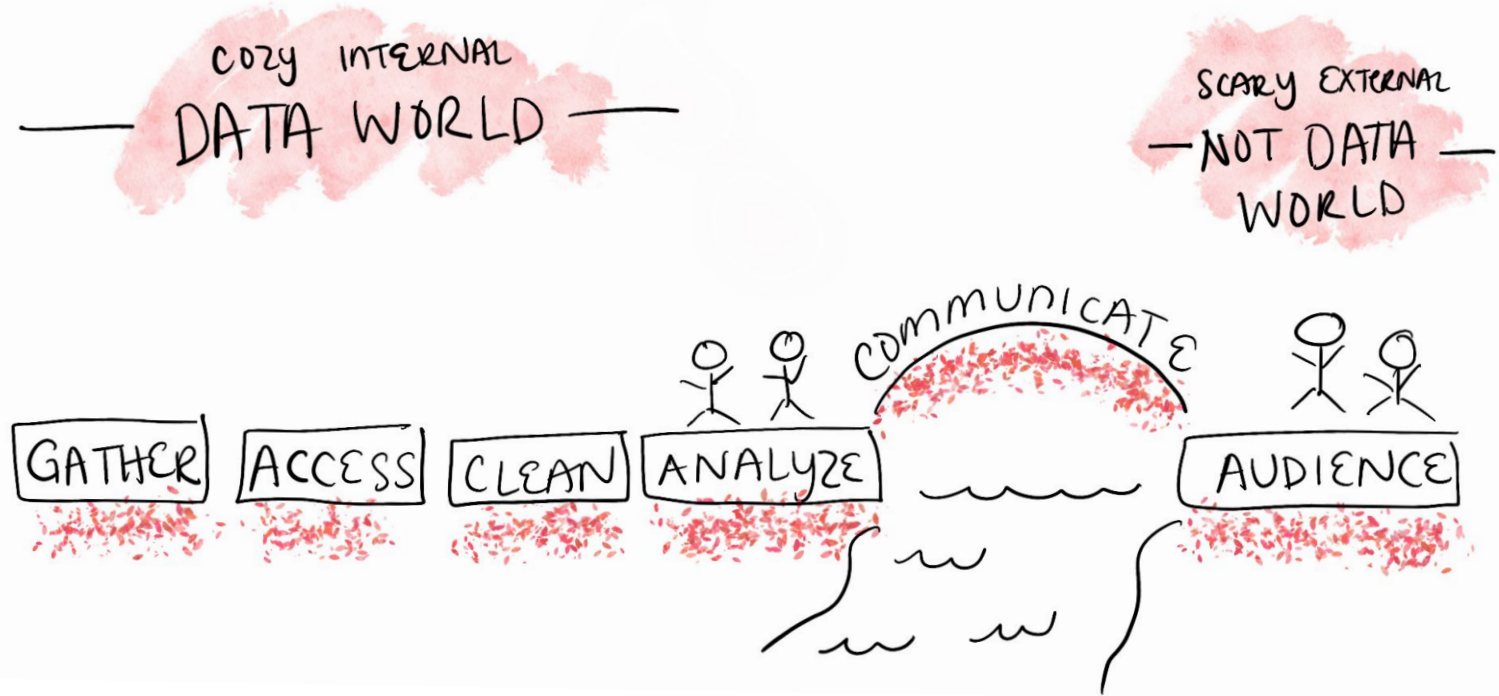
EFFECTIVE DATA COMMUNICATION

It's the bridge that connects the data and the audience.



EFFECTIVE DATA COMMUNICATION

A sturdy communication bridge is both essential and rare.



FINALLY!

Here are the 5 most common data communication **mistakes**,
and how to fix them.



MISTAKE #1

THE HALF-BAKED BRIDGE



THE MISTAKE

WE DON'T UNDERSTAND OUR AUDIENCE

- The Half-Baked Bridge is only anchored on the data side of the river (the side we know).
- We get stressed for time, and leave it just floating in the audience's general direction.
- The kicker: Then we get irritated nobody is using our bridge to reach our data.



THE SOLUTION

START EACH PROJECT WITH YOUR AUDIENCE

- Who exactly are your audience members?
- What are their questions for the data? Why?
- What is their relationship with this dataset?
- How does your audience learn best?
- How will you measure the bridge's strength?

EUROPE

SHARE

The Scale of the Migrant Crisis, From 160 to Millions

By GREGOR AISCH, SARAH ALMUKHTAR, JOSH KELLER and WILSON ANDREWS UPDATED SEPT. 22, 2015



4

ARTICLES REMAINING

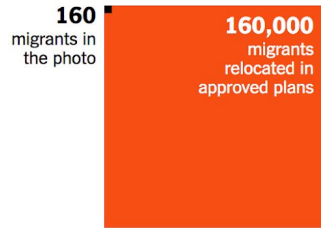
AUDIENCE CLARITY

THE SCALE OF THE MIGRANT CRISIS

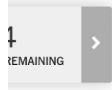
NEW YORK TIMES

2015

This photo shows about **160** migrants in southern Hungary waiting to board a bus to a registration center on Sept. 9. They will be excluded from the only two relocation plans approved by European leaders so far. The plans will relocate a total of **160,000** people in Greece and Italy to other countries in Europe.



The plans add up to the largest agreement ever to distribute asylum seekers among member states. But they address only a small fraction of the expected total.



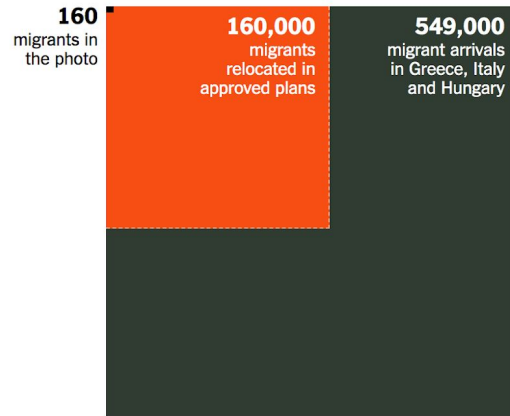
AUDIENCE CLARITY

THE SCALE OF THE MIGRANT CRISIS

NEW YORK TIMES

2015

Almost **549,000** migrants have arrived in Greece, Italy and Hungary alone this year, often by boat. Many want to move on to Germany or Sweden, which have stronger job markets and more benefits for refugees.

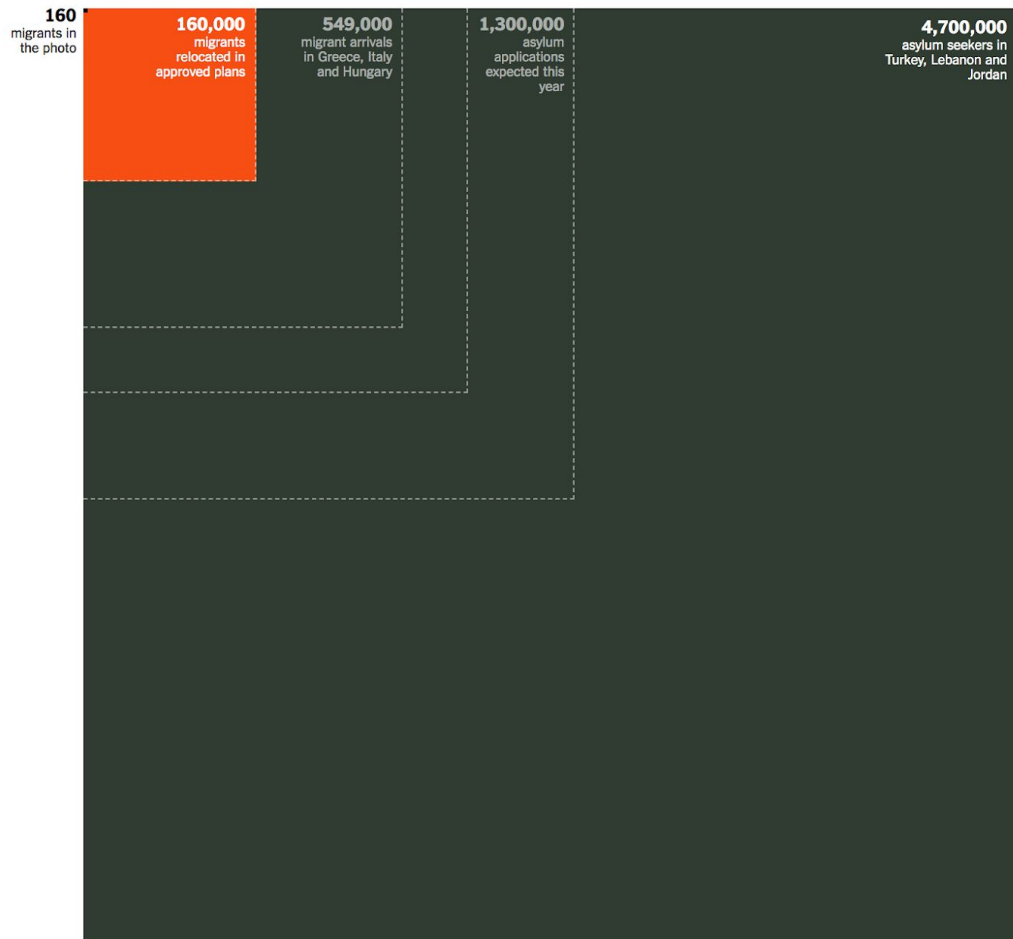


AUDIENCE CLARITY

THE SCALE OF THE MIGRANT CRISIS

NEW YORK TIMES

2015



AUDIENCE CLARITY

THE SCALE OF THE MIGRANT CRISIS

NEW YORK TIMES
2015

MISTAKE #2

THE JUNGLE BRIDGE



THE MISTAKE

WE OVERWHELM OUR AUDIENCE

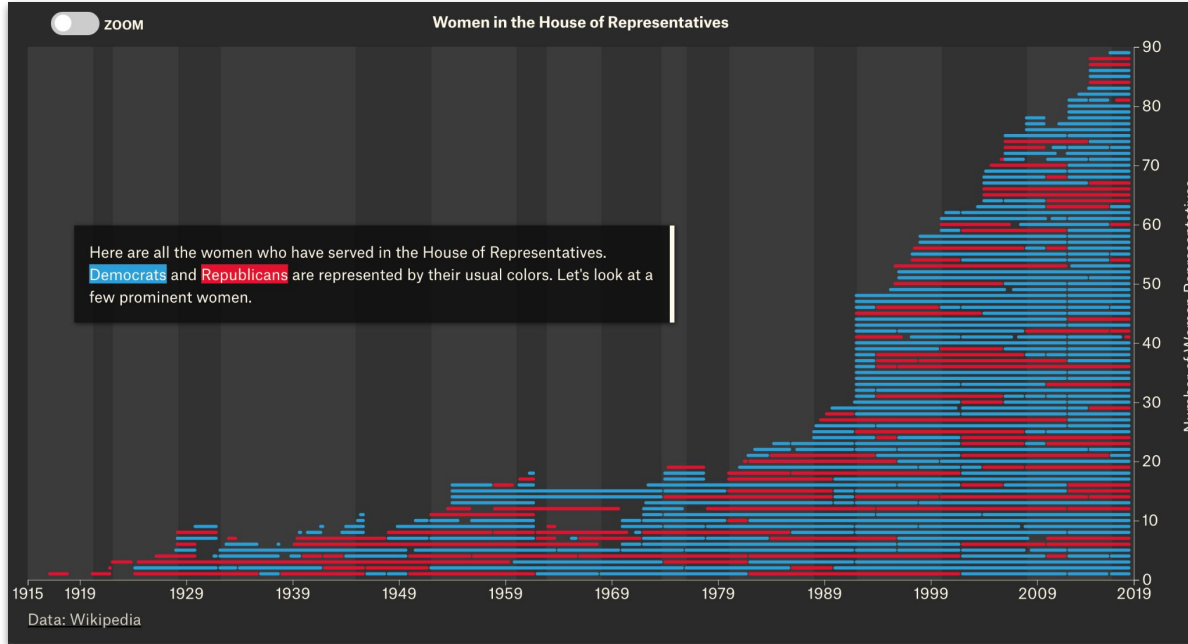
- The Jungle Bridge bombards our audience with every possible detail from the project.
- It's dense, scary, and requires folks to hack their way through searching for the point.
- 250-page reports, 150 chart slides, infinite excel attachments, crowded dashboards.



THE SOLUTION

CURATE A PRODUCTIVE PATH THROUGH THE DATA

- Knaflic: Don't the oysters, show the pearls.
- Curate a path based on your audience:
- Eliminate: Good curation means cutting.
- Order: Create an arc (story, argument, etc.)
- Drip: Direct attention to one item at a time.



CURATED PATH

WOMEN IN THE HOUSE OF REPRESENTATIVES

DURAND D'SOUZA

2018

MISTAKE #3

THE CIRCUS BRIDGE



THE MISTAKE

WE OVERSIMPLIFY THE DATA

- The Circus Bridge flattens all of the data's complexities into a caricature of itself.
- We tend to make this mistake from a place of desperation or arrogance.
- It leads to black and white thinking and emotionally charged interpretations.



THE SOLUTION

SHARE THE PROJECT'S NUANCES WITH YOUR AUDIENCE

- Your audience is just as smart as you.
- What limitations does the data itself have?
- What fundamental assumptions were made?
- What are likely incorrect interpretations?
- What internal jargon could trip up folks?
- Example: Raise your hand if you understand:
- “You have to clean expo before you can be cut.”

STEP 1

Collect annual household income data from all families



STEP 2

Sort all families in order of lowest household income to highest



SHARE NUANCE

HOW RIGID IS THE MIDDLE CLASS IN THE US, REALLY?

ALICE FENG & AMBER THOMAS

2020

MISTAKE #4

THE BLIND SPOT BRIDGE



THE MISTAKE

WE WORK IN ISOLATION (AND IT SHOWS)

- The Blind Spot Bridge happens when we do the entire data project by ourselves.
- You've seen this anytime the creator's weaknesses are the project's weaknesses.
- We forget that data viz and communication sit at an intersection of many fields.

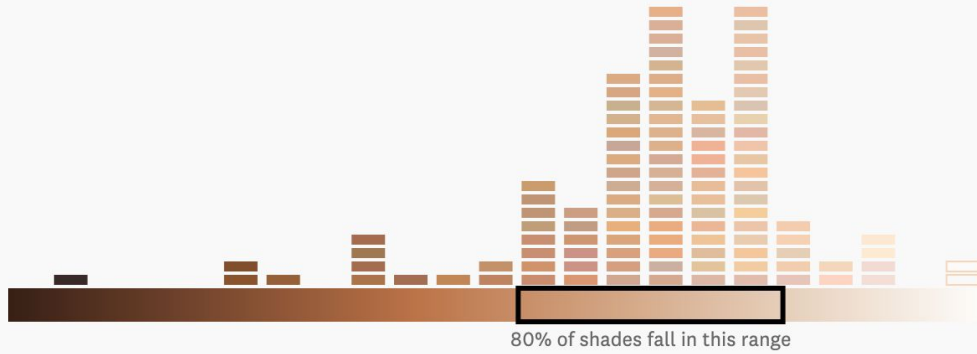


THE SOLUTION

COLLABORATE WITH COLLEAGUES WITH COMPLEMENTARY SKILLS

- Design: branding, layout, user experience
- Accessibility: screen readers, color blindness
- DEI: inclusivity, complexity, policy, culture
- Automation: ongoing maintenance, accuracy
- Marketing: connection, measurement
- Where are you and your project weakest?

110 shades with "nude" in the name



SHARE NUANCE

THE NAKED TRUTH: BIAS IN BEAUTY

OFFUNE AMAKA & AMBER THOMAS
2021

MISTAKE #5

THE ONE-WAY BRIDGE



THE MISTAKE

WE DON'T HELP OUR AUDIENCE ACT UPON THE DATA

- The One-Way Bridge acts like a museum folks visit and then leave empty handed.
- You've seen this anytime you felt excited in the moment, but never returned to the data.
- When we don't enable folks to take the data with them, the data remains on the island.



THE SOLUTION

CREATE TANGIBLE TAKEAWAYS FOLKS CAN CARRY FORWARD

- Web: digital home audiences revisit and share
- Visualization: a picture is worth 1,000 words
- Data: downloadable CSV or Excel file
- Video: verbal annotation of screen recording
- Next Steps: specific questions or actions
- One Pager: this is my personal favorite



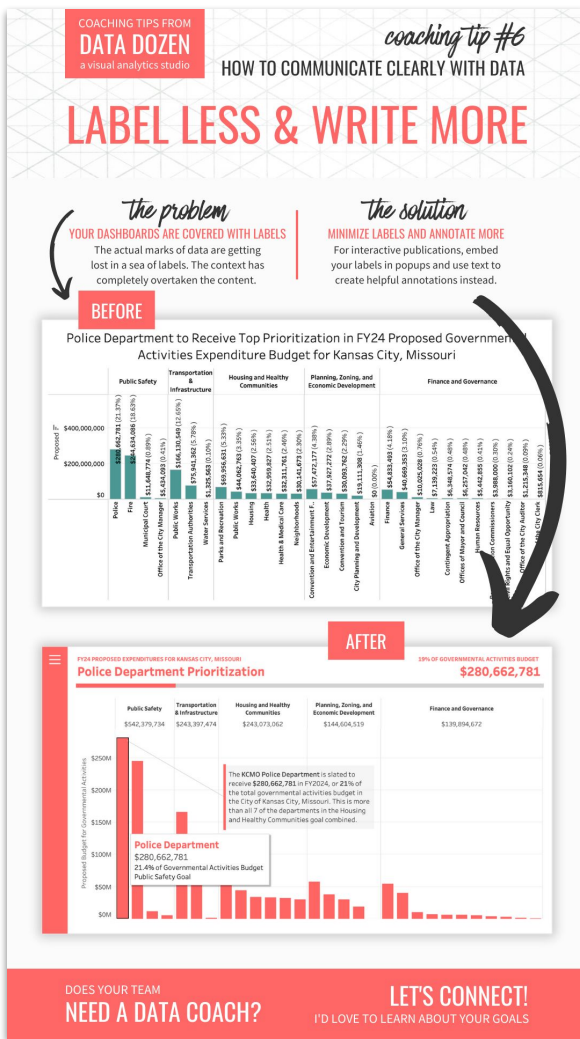
Erin Waldron • You
Data Communication Consultant • Speaker • Tableau Expert
1w • 📢

THE PROBLEM: Your data dashboards are covered with labels. The actual marks of data—the bars or lines or points—are getting lost in a sea of tiny text. You're eager to provide your audience context, but the labels have completely overtaken the content itself. The sheer number of labels scare off potential viewers, and the brave souls who try to understand them discover that the vertical orientation makes them nearly impossible to actually read.

THE SOLUTION: Minimize labels by using them strategically and instead use your text to create helpful annotations. So many of our data visualizations are built to live on screens these days. The interactivity of so many data tools means that the information that would have to go on a label if the dashboards were printed can now live in an interactive popup. The labels are there when they're helpful and disappear when they're not needed. (And on the off chance you do need to print it, move the label collection into a helpful side table instead of drowning out your chart with text.) What's more, now you can turn your writing attention to actually explaining the visualization. Use your annotations to call out the most interesting pieces of the visualization, and write down what's happening using the same language you would use to explain the visualization to a colleague over coffee. (Hint, hint: that means by writing full, complete sentences :)

NEED A DATA COACH? Does your team need a data coach to help you streamline your data communication strategy? If so, I'd love to connect! Send me a message and we'll set up a time to chat about your team's goals and how I might be able to help. Learn more about me and my work here: <https://datadozen.com/>

ABOUT THIS PROJECT: In this series, How to Communicate Clearly with Data, I've gathered up my top 10 go-to coaching strategies that I leverage on a daily basis with my clients as we work to improve their data communication and reporting. These simple solutions make such a substantial difference when it comes to getting your audience to actually engage with your data. Whether you're a fellow analyst or a business leader looking to get more out of your data, I hope you'll find these strategies useful. A special thanks to the **City of Kansas City, MO** for making their annual budget data available to the public on OpenDataKC and to **Michael Jenkins, MPA** for explaining the ins and outs of this dataset to me. As a local citizen, this dataset has been a blast to explore during this project. Thank you!



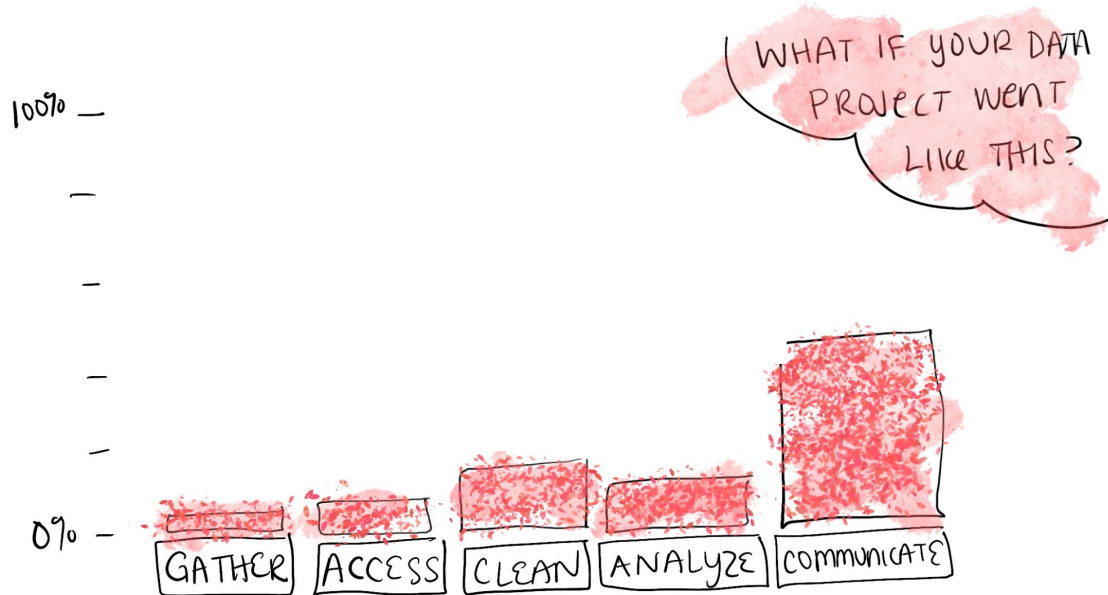
TANGIBLE TAKEAWAYS

BEFORE & AFTER SERIES

ERIN WALDRON 2023

ONE FINAL THOUGHT

What would happen if you spent more time
sharing your data project than creating it?



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

THANK YOU!

WWW.DATADOZEN.COM/COMMUNICATE

