

Structures of Data:

Structures of Data:

UNSTRUCTURED

Data which has no form, and likely needs to be processed in some way to create organization/structure.

Examples:

1. Word Docs
2. Emails
3. PDFs
4. Images
5. Etc.

Structures of Data:

UNSTRUCTURED

Data which has no form, and likely needs to be processed in some way to create organization/structure.

SEMI-STRUCTURED

Data which has some level of organization and structure, but itself doesn't represent truly relational data. Semi-Structured data often exists in a form which can be easily translated to structured data.

Examples:

1. Word Docs
2. Emails
3. PDFs
4. Images
5. Etc.

Examples:

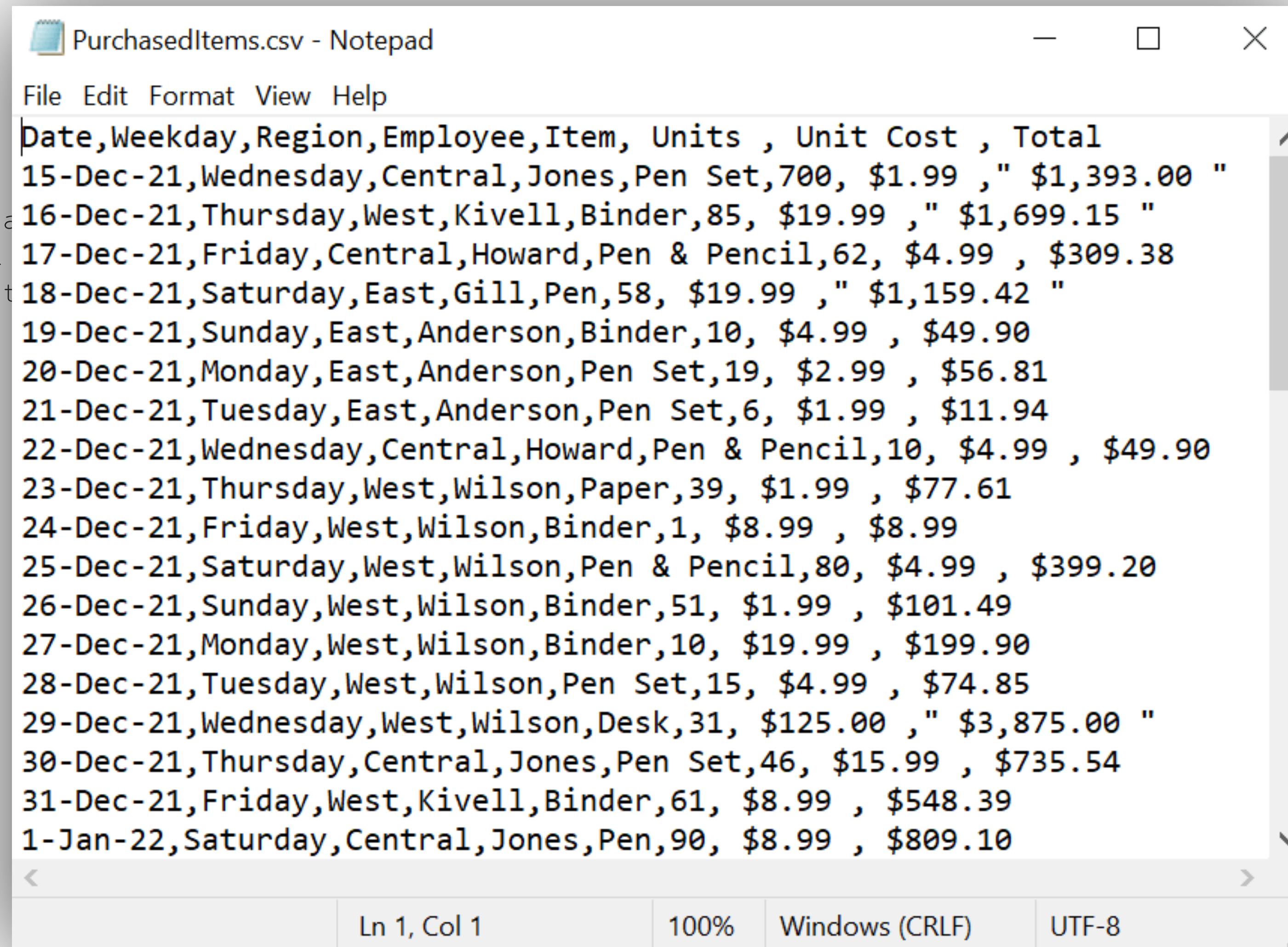
1. XML
2. JSON
3. CSV

UNSTRUCTURED

Data which has no form, a needs to be processed in to create organization/st

Examples:

1. Word Docs
2. Emails
3. PDFs
4. Images
5. Etc.



PurchasedItems.csv - Notepad

File Edit Format View Help

Date	Weekday	Region	Employee	Item	Units	Unit Cost	Total
15-Dec-21	Wednesday	Central	Jones	Pen Set	700	\$1.99	\$1,393.00
16-Dec-21	Thursday	West	Kivell	Binder	85	\$19.99	\$1,699.15
17-Dec-21	Friday	Central	Howard	Pen & Pencil	62	\$4.99	\$309.38
18-Dec-21	Saturday	East	Gill	Pen	58	\$19.99	\$1,159.42
19-Dec-21	Sunday	East	Anderson	Binder	10	\$4.99	\$49.90
20-Dec-21	Monday	East	Anderson	Pen Set	19	\$2.99	\$56.81
21-Dec-21	Tuesday	East	Anderson	Pen Set	6	\$1.99	\$11.94
22-Dec-21	Wednesday	Central	Howard	Pen & Pencil	10	\$4.99	\$49.90
23-Dec-21	Thursday	West	Wilson	Paper	39	\$1.99	\$77.61
24-Dec-21	Friday	West	Wilson	Binder	1	\$8.99	\$8.99
25-Dec-21	Saturday	West	Wilson	Pen & Pencil	80	\$4.99	\$399.20
26-Dec-21	Sunday	West	Wilson	Binder	51	\$1.99	\$101.49
27-Dec-21	Monday	West	Wilson	Binder	10	\$19.99	\$199.90
28-Dec-21	Tuesday	West	Wilson	Pen Set	15	\$4.99	\$74.85
29-Dec-21	Wednesday	West	Wilson	Desk	31	\$125.00	\$3,875.00
30-Dec-21	Thursday	Central	Jones	Pen Set	46	\$15.99	\$735.54
31-Dec-21	Friday	West	Kivell	Binder	61	\$8.99	\$548.39
1-Jan-22	Saturday	Central	Jones	Pen	90	\$8.99	\$809.10

Ln 1, Col 1 100% Windows (CRLF) UTF-8

Structures of Data:

UNSTRUCTURED

Data which has no form, and likely needs to be processed in some way to create organization/structure.

SEMI-STRUCTURED

Data which has some level of organization and structure, but itself doesn't represent truly relational data. Semi-Structured data often exists in a form which can be easily translated to structured data.

STRUCTURED

Data which exists in a relational structured form. Most often this means in the form of a table with rows and columns.

Examples:

1. Word Docs
2. Emails
3. PDFs
4. Images
5. Etc.

Examples:

1. XML
2. JSON
3. CSV

Examples:

1. Spreadsheets
2. Databases

**Note!

Simply by loading a CSV, etc. into sheets for example, changes it from semi-structured to structured.

A closer look at JSON:

A closer look at JSON:

JAVASCRIPT OBJECT NOTATION

A closer look at JSON:

1. Semi-Structured Data Format
2. Stores & Transmits Data
3. While it suggests closest alignment with JS, it is possible to work with JSON in an array of languages.
4. It is the industry standard for data storage and transmission.

```
{"name":"Anna", "age":30, "isStudent":false}
```

Keys

Must be strings written
with double quotes.

```
{"name":"Anna", "age":30, "isStudent":false}
```

Values

Values can only be:

1. String
2. Number
3. Object
4. Array
5. Boolean
6. Null

```
const data = `{
    "name": "Anna",
    "age": 30,
    "isStudent": false,
    "address": {
        "street": "123 Main St",
        "city": "New York",
        "postalCode": "10001"
    },
    "hobbies": ["travelling", "photography", "cooking"],
    "hasPets": null
}`;

const dataObject = JSON.parse(data);

console.log(dataObject);
```

javascript.js:19

```
▼ {name: 'Anna', age: 30, isStudent: false, address: {...},  
  hobbies: Array(3), ...} ⓘ  
  ► address: {street: '123 Main St', city: 'New York', pos  
    age: 30  
    hasPets: null  
  ► hobbies: (3) ['travelling', 'photography', 'cooking']  
    isStudent: false  
    name: "Anna"  
  ► [[Prototype]]: Object
```

APIs:

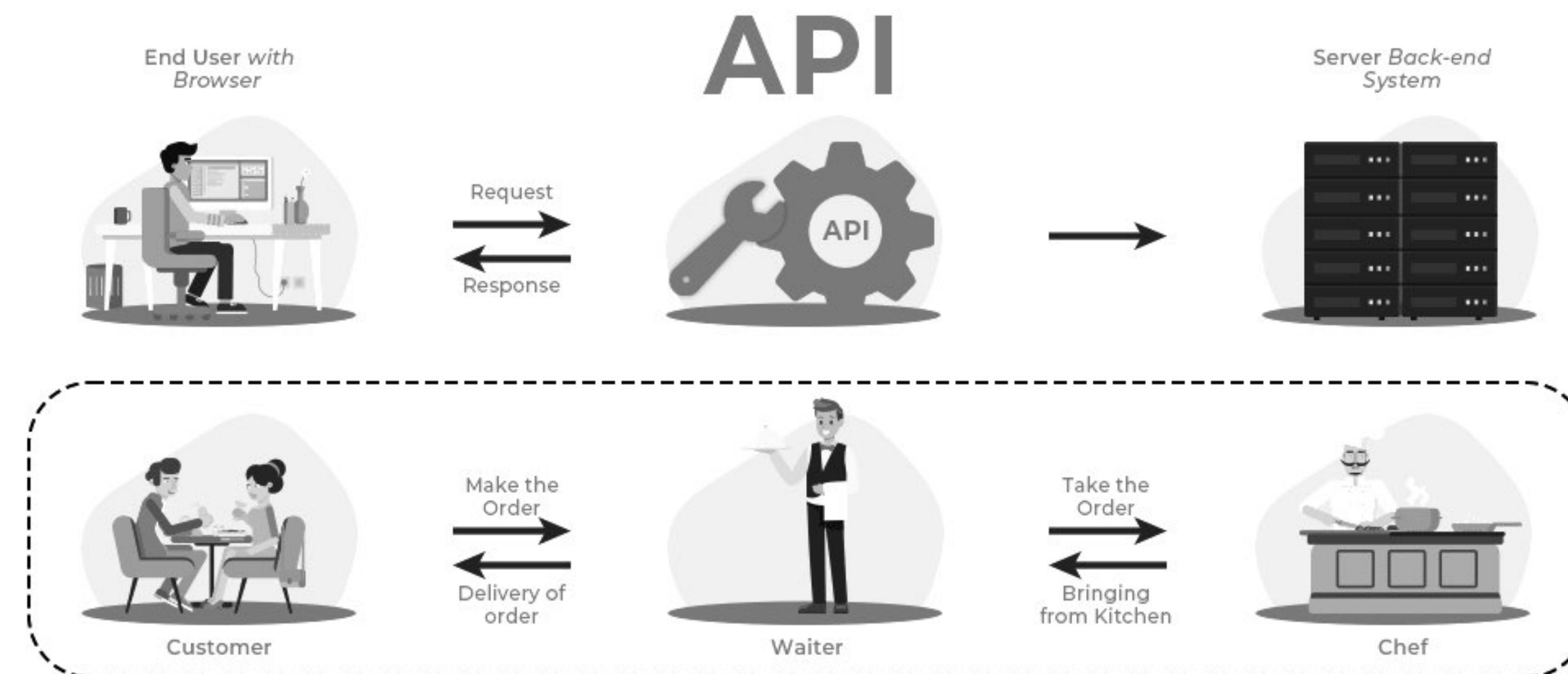
APIs:

APPLICATION PROGRAMMING INTERFACE

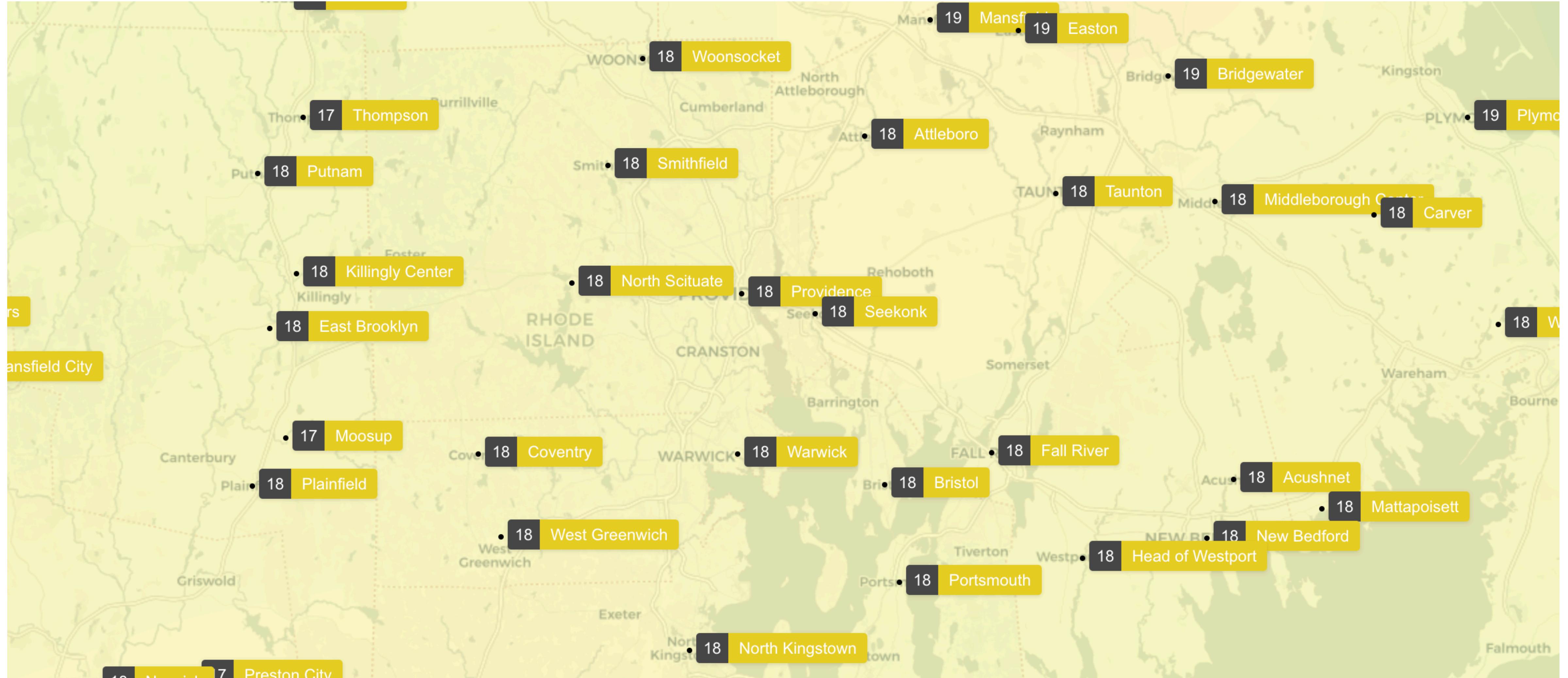
APIs:

Allows two systems to exchange data and functionality

APIs:



Live Data



Payment Processing



Image Processing

Form JSON Node.js Python HTTP Cog Docker

image



Preview JSON



Generated in

3.9 seconds

⤒ Tweak it

⤓ Share

⤔ Download

ⓘ Report

Web Scraping:

Web Scraping:

Allows the Extraction of Data from a Website

Web Scraping:

1. Product Names & Prices
2. Images
3. Profiles Information, Comments, and Likes
4. Sports Statistics
5. Financial Data
6. And More!

- 1. Product Names & Prices
- 2. Images
- 3. Profiles Information, Comments, and Likes
- 4. Sports Statistics
- 5. Financial Data
- 6. And More!



Web Scraping:

- 1. Market Research
- 2. Price Monitoring & Competitive Analysis
- 3. Lead Generation for Sales
- 4. Sentiment Analysis
- 5. Journalism

Untitled Project

Enter a website you'd like to extract data from

?cflt=restaurants&find_loc=Toronto%2C+ON

Start project on this URL

Connecting... https://www.parsehub.com/landing

Welcome!

Beginner Tutorials

1. Start Here: Create your first project
2. Extract text from a web page
3. Extract data from many pages (pagination)
4. Run project & download Excel & JSON data
5. Use the REST API

Advanced Tutorials

1. Collect data on a schedule
2. Enter text into a search box
3. Get data from behind a log-in
4. Infinite scrolling pages
5. Enter URLs for ParseHub to crawl

API Docs

Use our REST API and webhooks.

[View API >](#)

More Tutorials

Browse 80 + tutorials, videos & examples.

[View Tutorials >](#)

Contact Us

Need help with your project? Reach out.

[Contact >](#)

Product Demo

Get a custom product walk-through.

[Schedule >](#)

Waiting for www.yelp.ca...

Live preview is disabled when working on interactive templates. Use [Test Run](#) instead.

API Tutorials Contact

Show me data | View my results | Refresh



THE BEST 10 Restaurants in ... +

https://www.yelp.ca/search?cflt=restaurants&find_loc=Toronto%2C+ON

yelp Find tacos, cheap dinner, Max's Near Toronto, ON Select Mode Log In Sign Up

Restaurants Home Services Auto Services More Write a Review For Businesses

The Best 10 Restaurants in Toronto, ON Showing 1-30 of 8808

Toronto > Restaurants

All Filters \$ \$\$ \$\$\$ \$\$\$ Open Now Reservations Good for Groups Good for Dessert

Wed, 19 Jun 7:00 pm 2 people Find a Table

Sponsored Results ⓘ

 **Trattoria Leonardo** (647) 749-6150
★★★★★ 56 reviews 4740 Dundas Street W
\$\$ - Italian Etobicoke
"I live right next door to the place so I had to see what everyone"

Mo' Map Redo search when map is moved

CSV/Excel JSON CSV/Excel Wide (beta)

A preview of your data will appear here

THE BEST 10 Restaurants in Toronto, ON

Find tacos, cheap dinner, Max's Near Toronto, ON

Restaurants Home Services Auto Services More Write a Review For Businesses

The Best 10 Restaurants in Toronto, ON Showing 1-30 of 8808

Toronto > Restaurants

All Filters \$ \$\$ \$\$\$ Open Now Reservations Good for Groups Good for Dessert

Wed, 19 Jun 7:00 pm 2 people Find a Table

Sponsored Results

Trattoria Leonardo (647) 749-6150 4740 Dundas Street W Etobicoke

4 reviews \$\$ - Italian

Mo' Map Redo search when map is moved

CSV/Excel JSON CSV/Excel Wide (beta)

A preview of your data will appear here

Share your data View my saved filters

THE BEST 10 Restaurants in Toronto, ON

Find tacos, cheap dinner, Max's Near Toronto, ON

Restaurants Home Services Auto Services More Write a Review For Businesses

The Best 10 Restaurants in Toronto, ON Showing 1-30 of 8808

Toronto > Restaurants

All Filters \$ \$\$ \$\$\$ \$\$\$ Open Now Reservations Good for Groups Good for Dessert

Wed, 19 Jun 7:00 pm 2 people Find a Table

Sponsored Results

Trattoria Leonardo (647) 749-6150 4740 Dundas Street W Etobicoke

★★★★ 3 reviews \$\$ - Italian

Mo' Map Redo search when map is moved

CSV/Excel JSON CSV/Excel Wide (beta)

selection1 selection1_url

Trattoria Leonardo https://www.yelp.ca/adredir?ad_business_id=yRUZ_sIxV6FyNygZD6MNTg&campaign_id=BjSKAJv...40&signature=342176d4ba9b2b559e1e28ea116af81c14c6c063848b6a9c58c01284180a5e12&slot=0

This is a live preview. When you are ready to run your project, click Get Data.

THE BEST 10 Restaurants in  Toronto

https://www.yelp.ca/search?cflt=restaurants&find_loc=Toronto%2C+ON

main template

Select page 

Select restaurant 

Extract name

Extract url

Empty Relative selection1 (0) 

Get Data

Click an element on the page to select it.

Offers reservations 

Find a Table

7. **Kōjin**  (647) 253-8000

★ ★ ★ ★ 80 reviews

Canadian (New)

190 University Ave Core

"The Lucky Peach delivers the whimsical happiness that is their culinary space. Playfully flavorful with textures and visuals that make your meal memorable..." [read more](#)

8. **Seven Lives Tacos Y Mariscos**  (416) 393-4636

★ ★ ★ ★ 1238 reviews

\$ - Mexican

69 Kensington Avenue
Kensington Market

"I am @ Niagara Falls for my b-day, but on this day I decided I am too close to Toronto not to rent a car and go check out this great city for myself, eh? Over..." [read more](#)

CSV/Excel JSON CSV/Excel Wide (beta)

restaurant_name	restaurant_url
Trattoria Leonardo	https://www.yelp.ca/adredir?ad_business_id=yRUZ_sIxV6FyNygZD6MNTg&campaign_id=BjSKAJvqRPwThVjrXG8Fbw&click_origin...
Nord Bistro	https://www.yelp.ca/adredir?ad_business_id=4P_HNoH4cjoCdPHPb2LTXA&campaign_id=gNU-QtEOrjstiYSpma14rQ&click_origin...

This is a live preview. When you are ready to run your project, click Get Data.

Mo' Map Redo search when map is moved

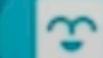
Select Mode

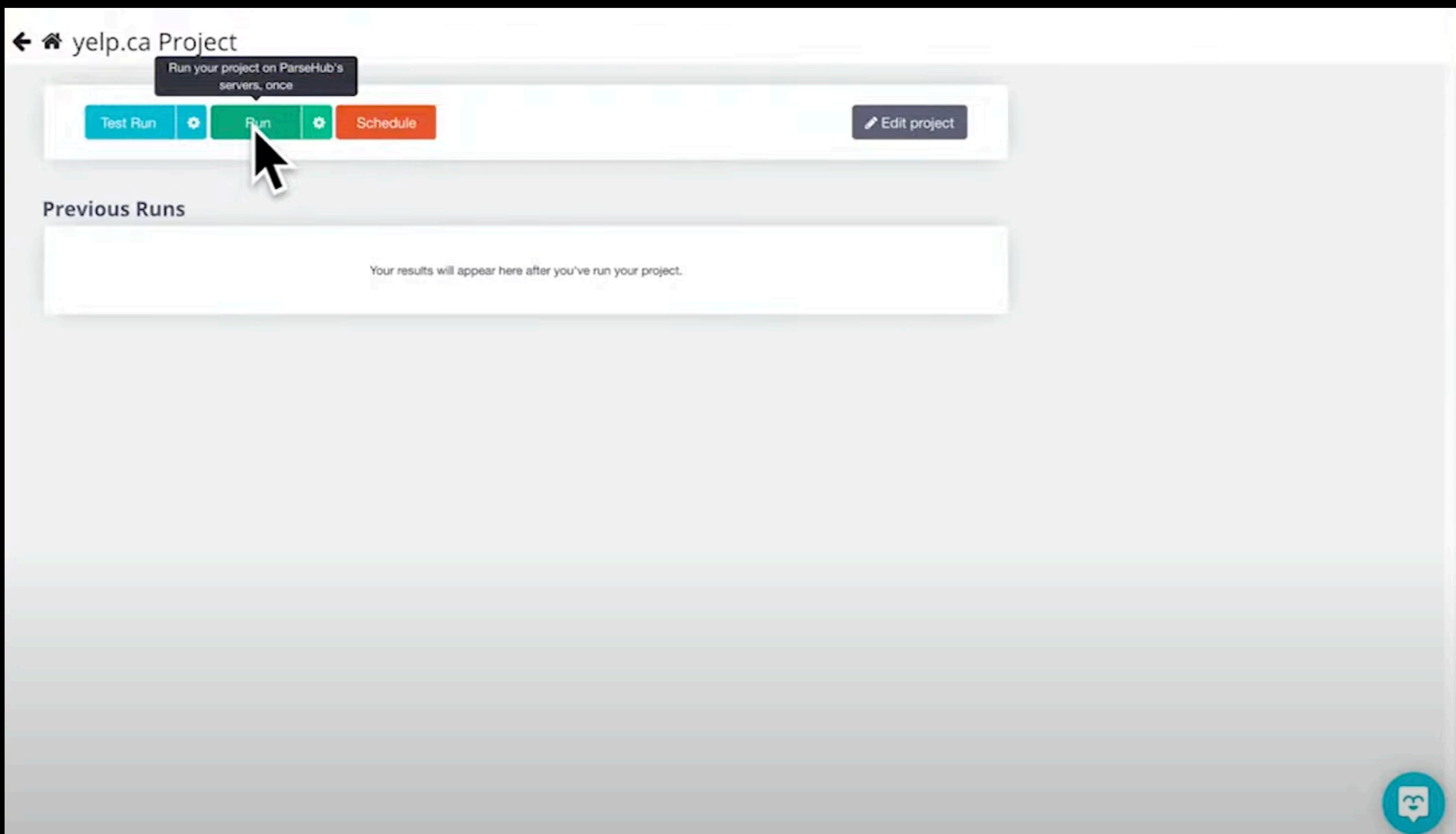
Map data ©2019 Google [Terms of Use](#)

 firehousesubs.ca

Firehouse Subs - Rooted In Authenticity - Founded By Firemen

(Ad) Connect With Your Loved Ones Over Delicious Foods And Top Notch Service. Firehouse Subs Is A Place Everyone Can Enjoy. Find A Location Near You. Founded by Firemen. Styles: Smoked Turkey, Beef Brisket, Pastrami, Venison, Grilled Chicken





yelp.ca Project

 Edit project

Data is being collected. Please wait.

Starting up...

Refreshing status in 1 second(s). [Refresh now](#)

Download Data

CSV/Excel  JSON API


All dates and times are in UTC +0000.

Empty file with no results? [Click here](#) to fix.

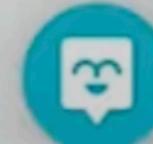
CSV file too big? Save the JSON file and [click here](#) to convert to CSV.

Run Details

Status	initialized
Pages	0 collected
Initialized	2019-06-19T13:42:01
Start Time	Not yet started
Finished	Not yet finished
API Key	tjfY-a0QTaU
Project Token	tsB5tTM10GdG

Settings

URL	<code>https://www.yelp.ca/search?cflt=restaurants&find_loc=Toronto%2C+ON</code>
Starting Template	main_template



D3.js:

D3.js:

JavaScript Library for bespoke data visualization

D3.js:

D3 gallery

Looking for a good D3 example? Here's a few (okay, 173...) to peruse.

Animation

D3's [data join](#), [interpolators](#), and [easings](#) enable flexible [animated transitions](#) between views while preserving [object constancy](#).



Case Studies:

Mimi Onuoha

US, AGGREGATED



Us, Aggregated 3.0 (2019)



The Library of Missing Datasets (2016)

Case Studies:

Jer Thorp



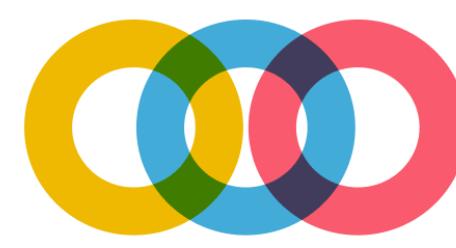
Every Bird



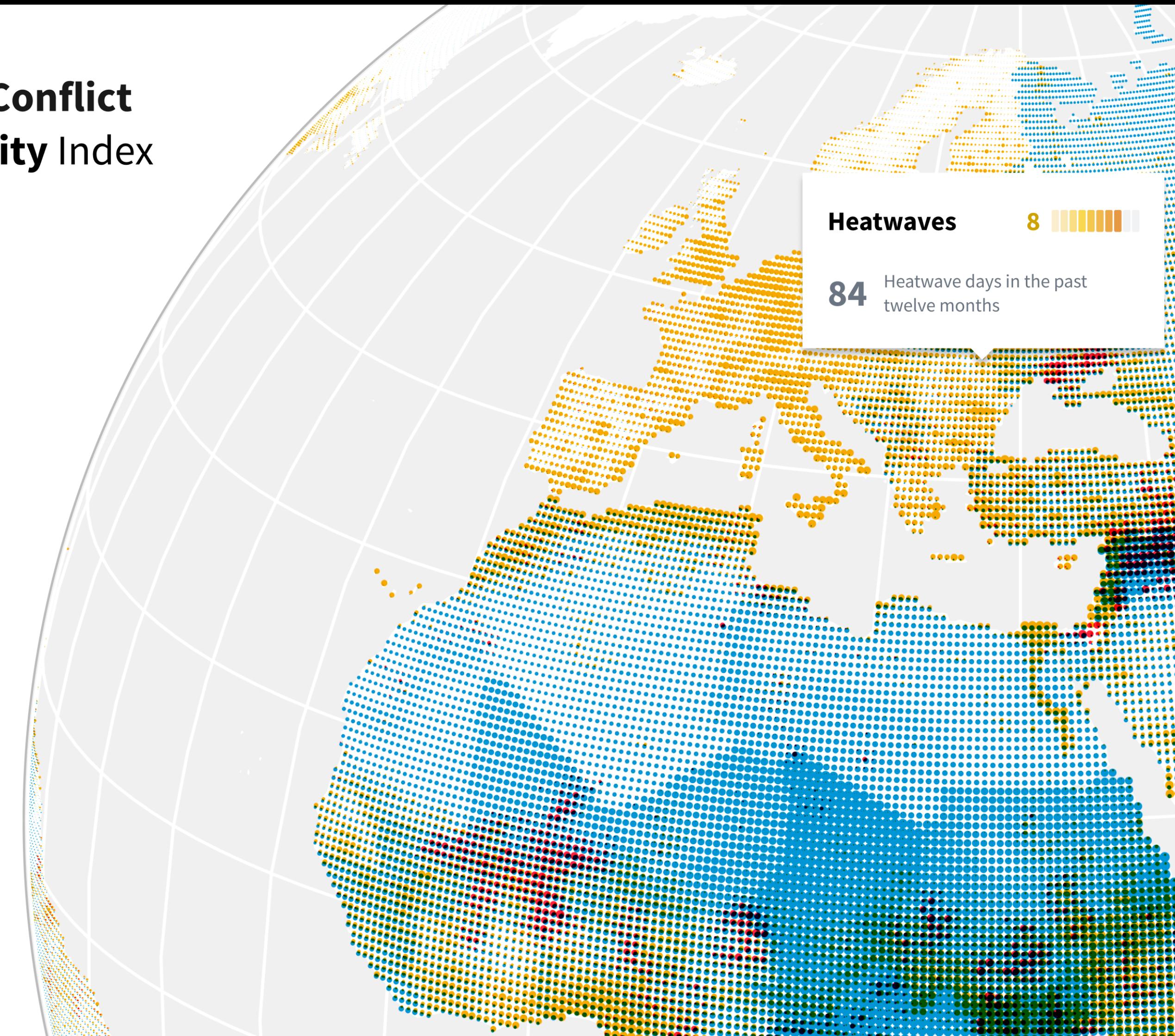
Herald Harbinger (2018)

Case Studies:

Moritz Stefaner



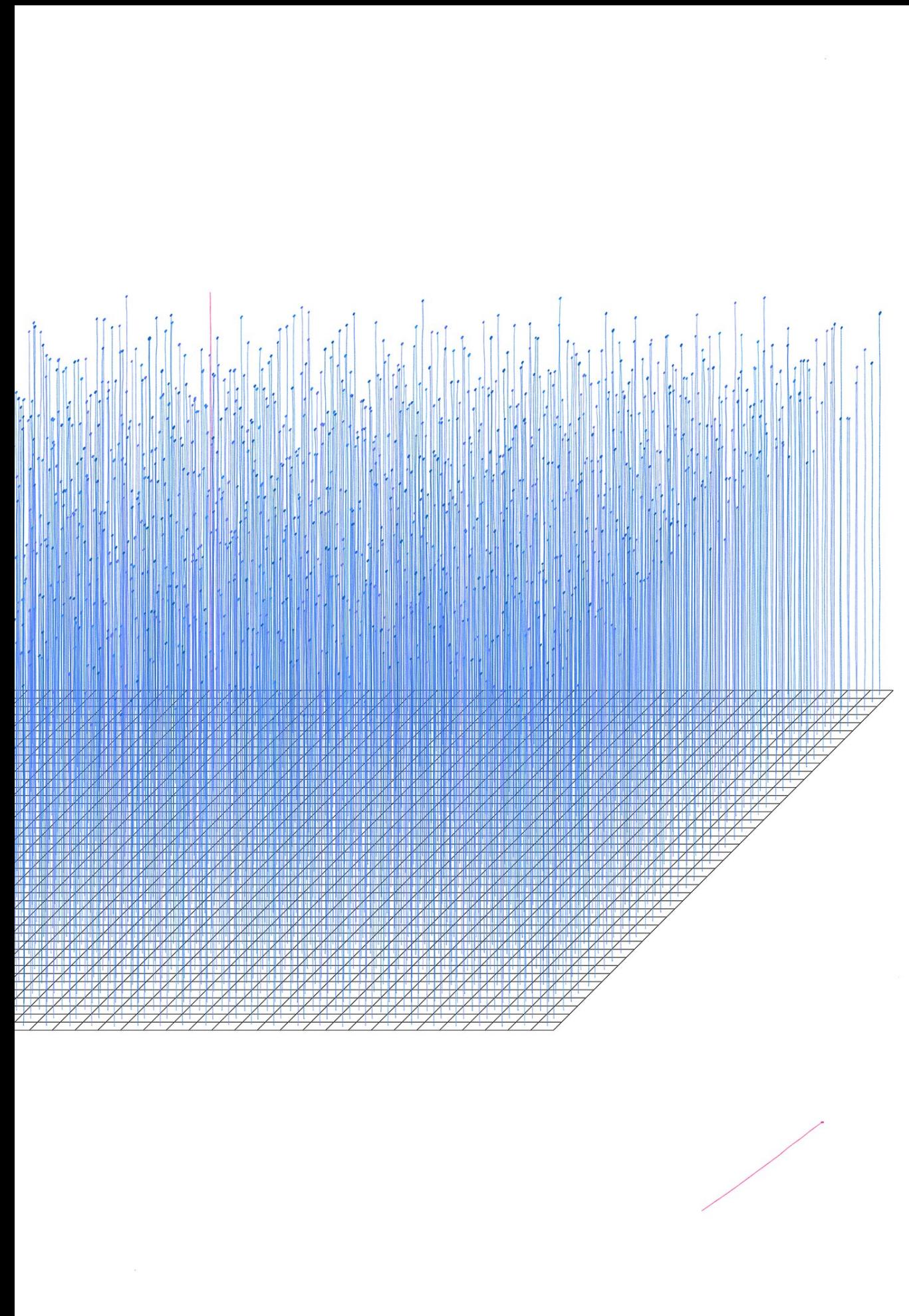
Climate–Conflict Vulnerability Index



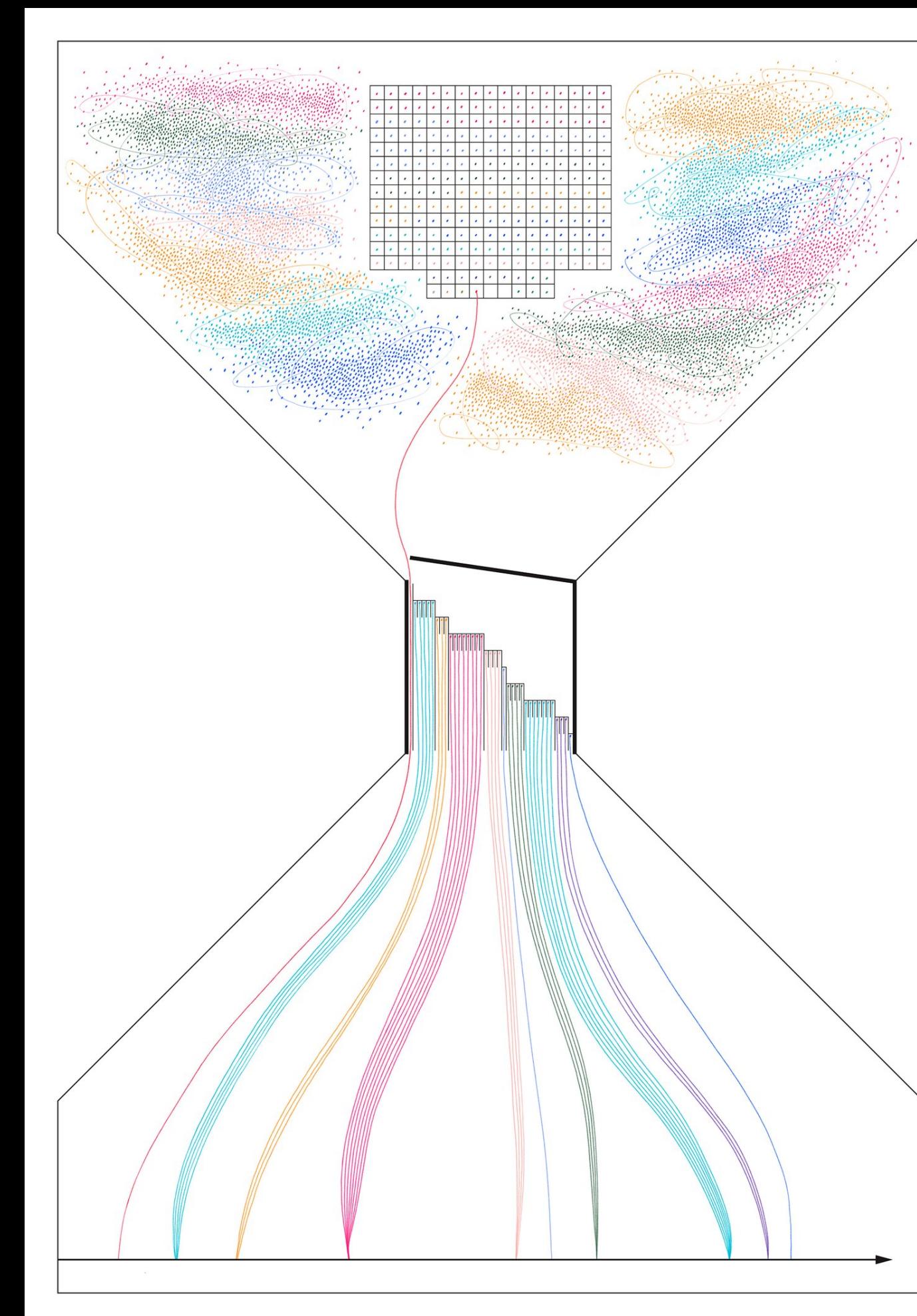
Climate–Conflict Vulnerability Index

Case Studies:

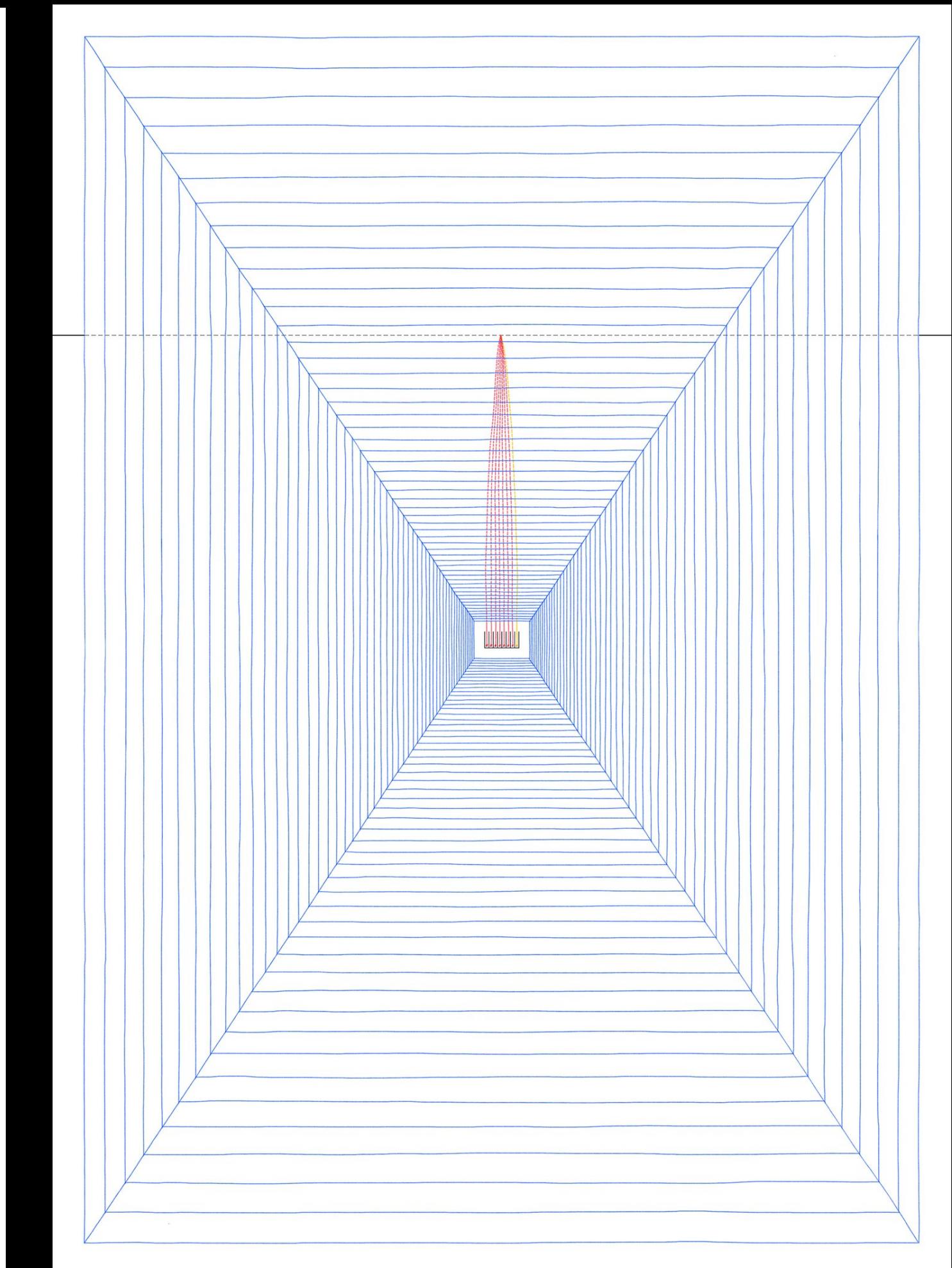
Stefanie Posavec



The Researcher



The Database Manager



The Lab Techs



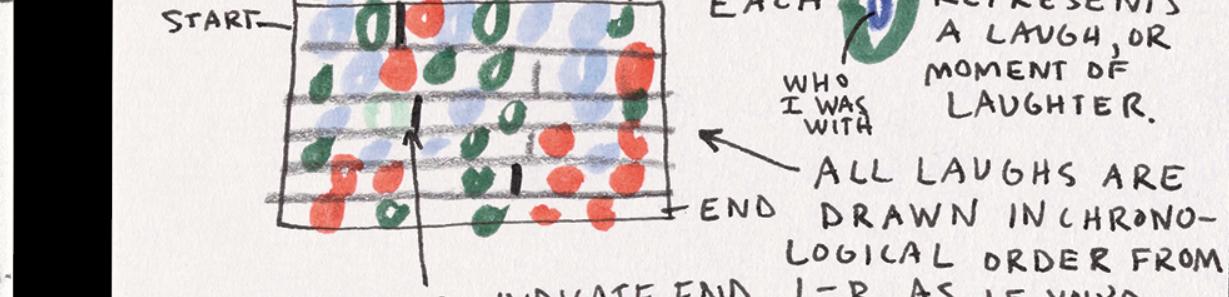
DEAR DATA - WEEK 42

A WEEK OF LAUGHTER FOR A CARD ABOUT LAUGHTER I AM SAD ABOUT HOW THIS CARD TURNED OUT

ABOUT THE DATA: I TRIED TO CAPTURE MY LAUGHS WHICH WAS REALLY HARD + GOT IN THE WAY OF ENJOYING LIFE, HENCE THE DATA VOIDS :)

HOW TO READ IT:

IN CHRONO ORDER



MARKERS INDICATE END OF ONE DAY + BEGINNING OF OTHER. IF MARKER IS IN PEN, IT MEANS I HAD A DATA VOID DUE TO THE FOLLOWING: DRINKING W FRIENDS, MY BDAY!! , MY B-DAY DINNER :)

SIZE OF LAUGH:

000

TO MYSELF, LARGE + REGULAR HEARTFELT!

00 A DOT MEANS YOU KNOW LAUGHING WITH LAUGHING THEM AT PEOPLE (GENERALLY IN FUN!!)

I WAS LAUGHING WITH:

- MYSELF
- BEING IN A GOOD MOOD, GENERAL TV SHOW
- FRIENDS
- STRANGERS BEHAVIOUR
- MYSELF, BEING TICKLED!
- DEAR DATA
- STUDIO MATES
- GROUP OF FRIENDS
- PROFESSIONAL CONTACT
- YOU!
- PARENT
- SCHADENFRUENDE
- PHONE AUTOCORRECT
- YOU!!

FROM:

S POSAVEC

on ~~Get tips at~~ ~~mail~~ ~~UK~~ ~~dogawareness~~



GIORGIA LUPI

BROOKLYN, NY

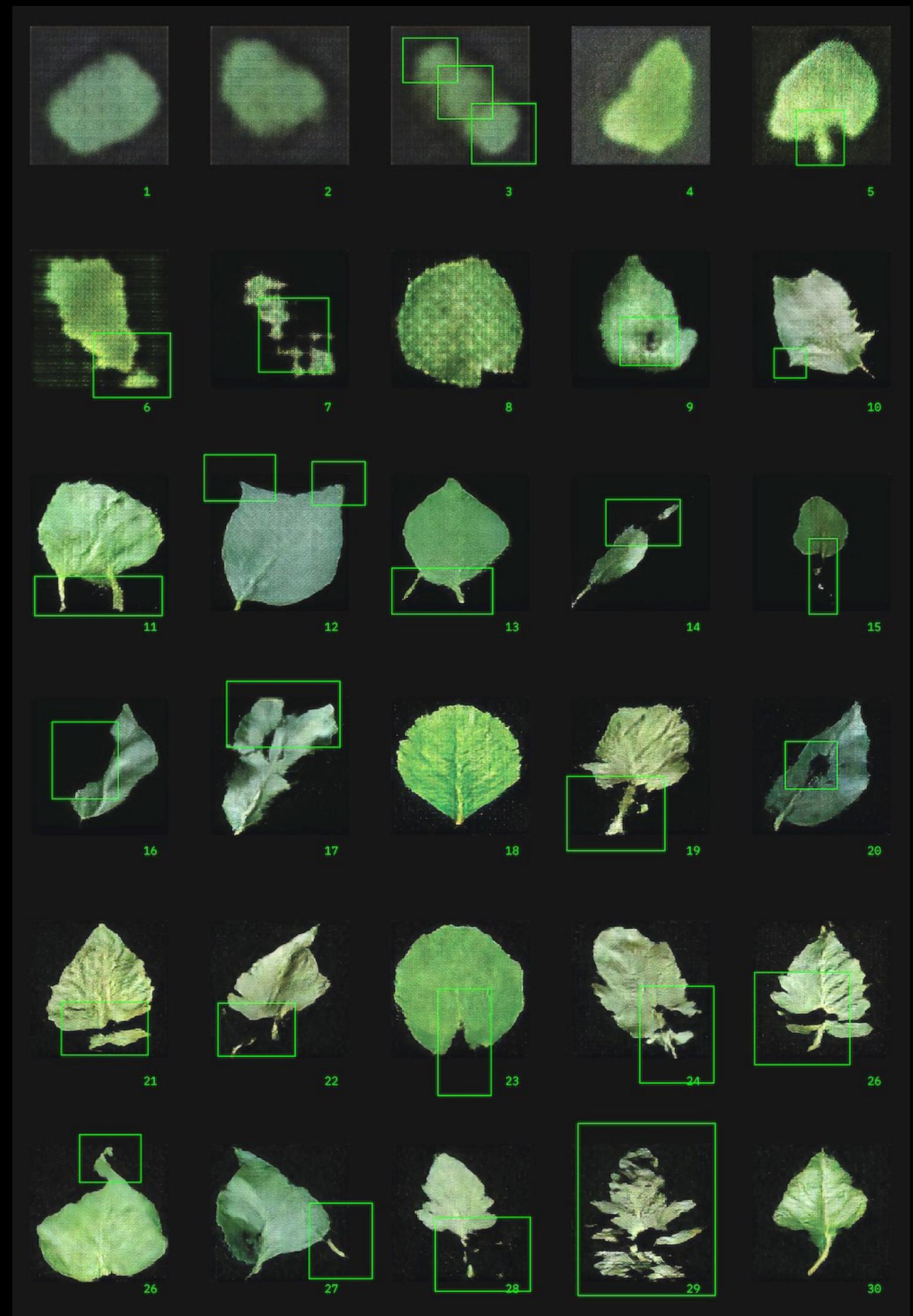
USA

AIRMAIL!

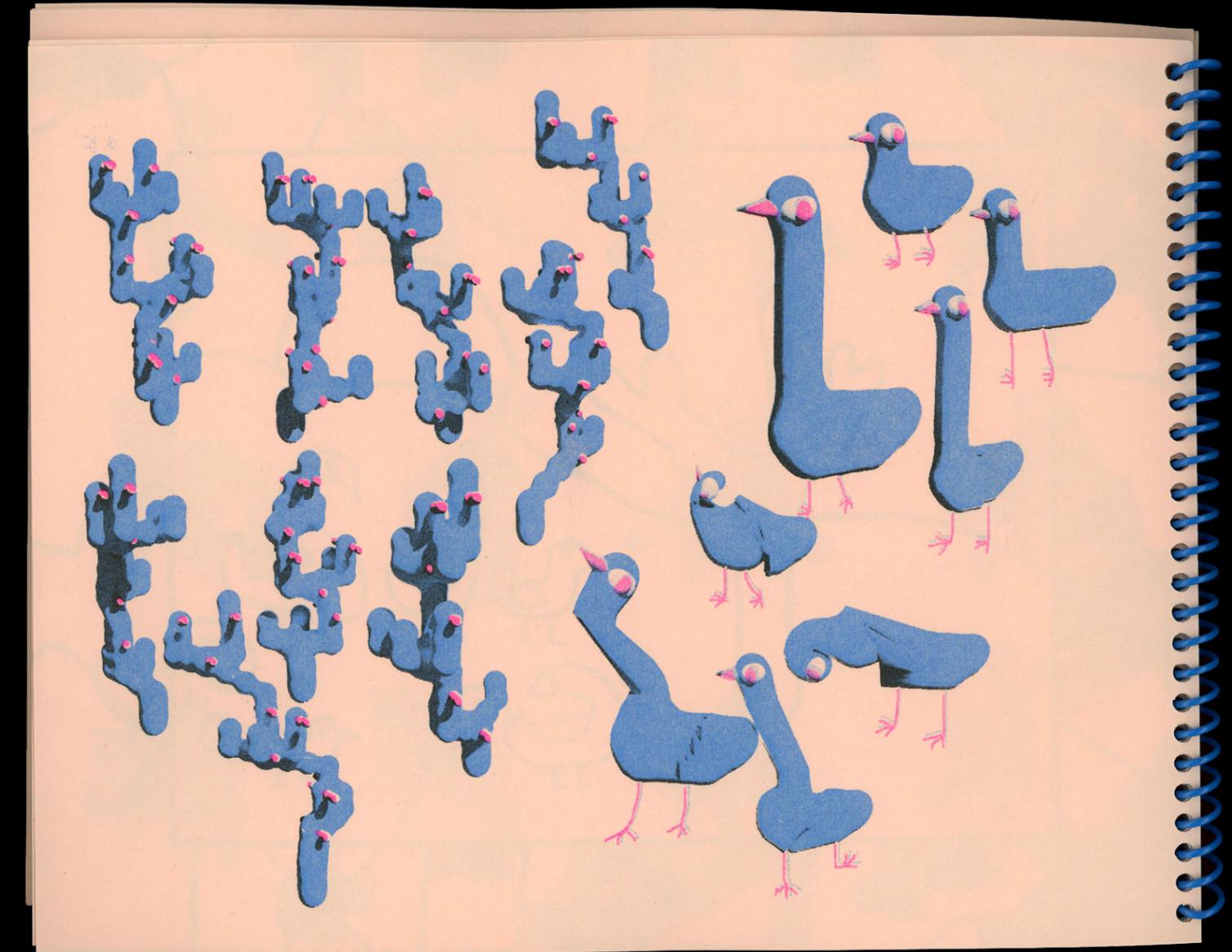
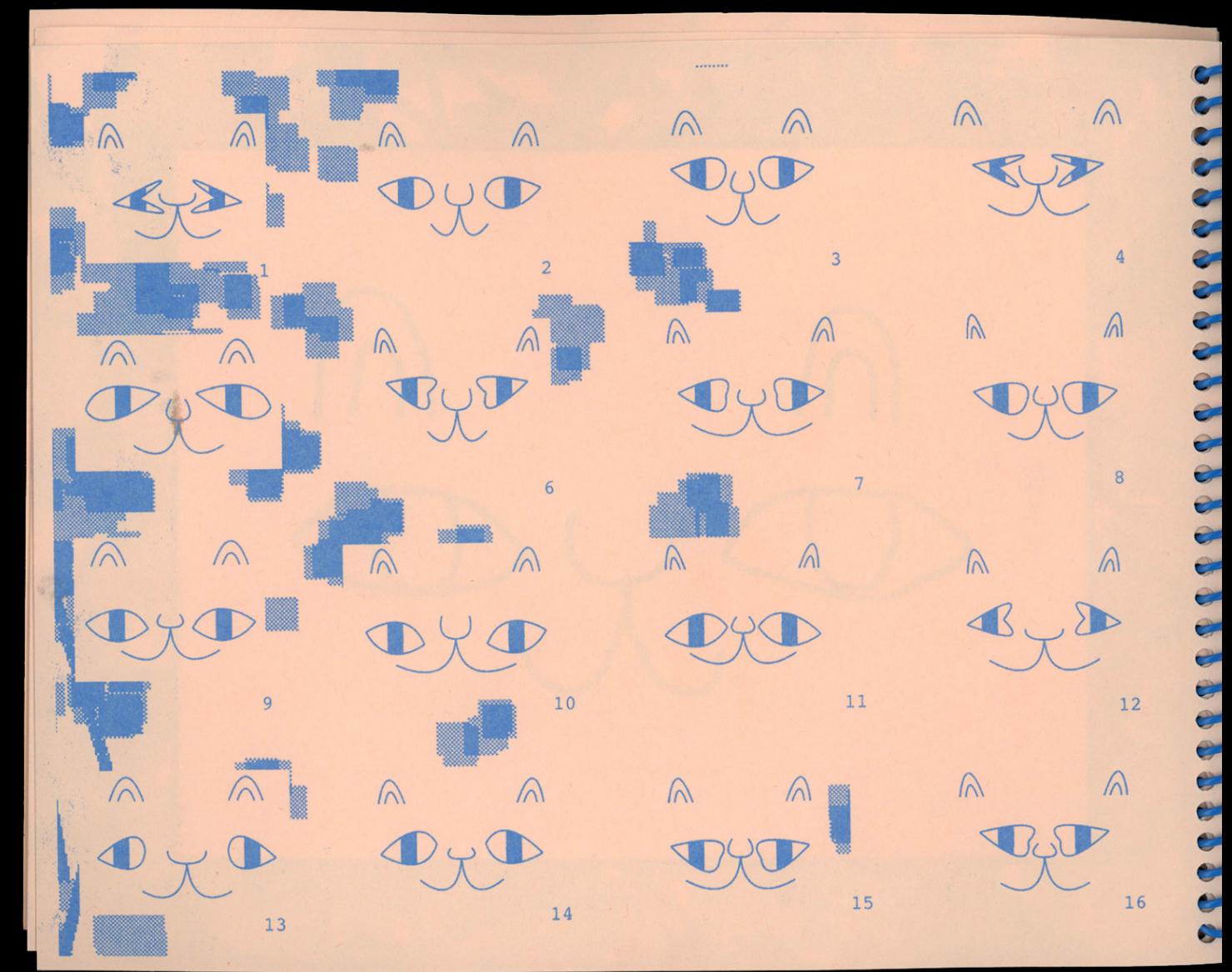
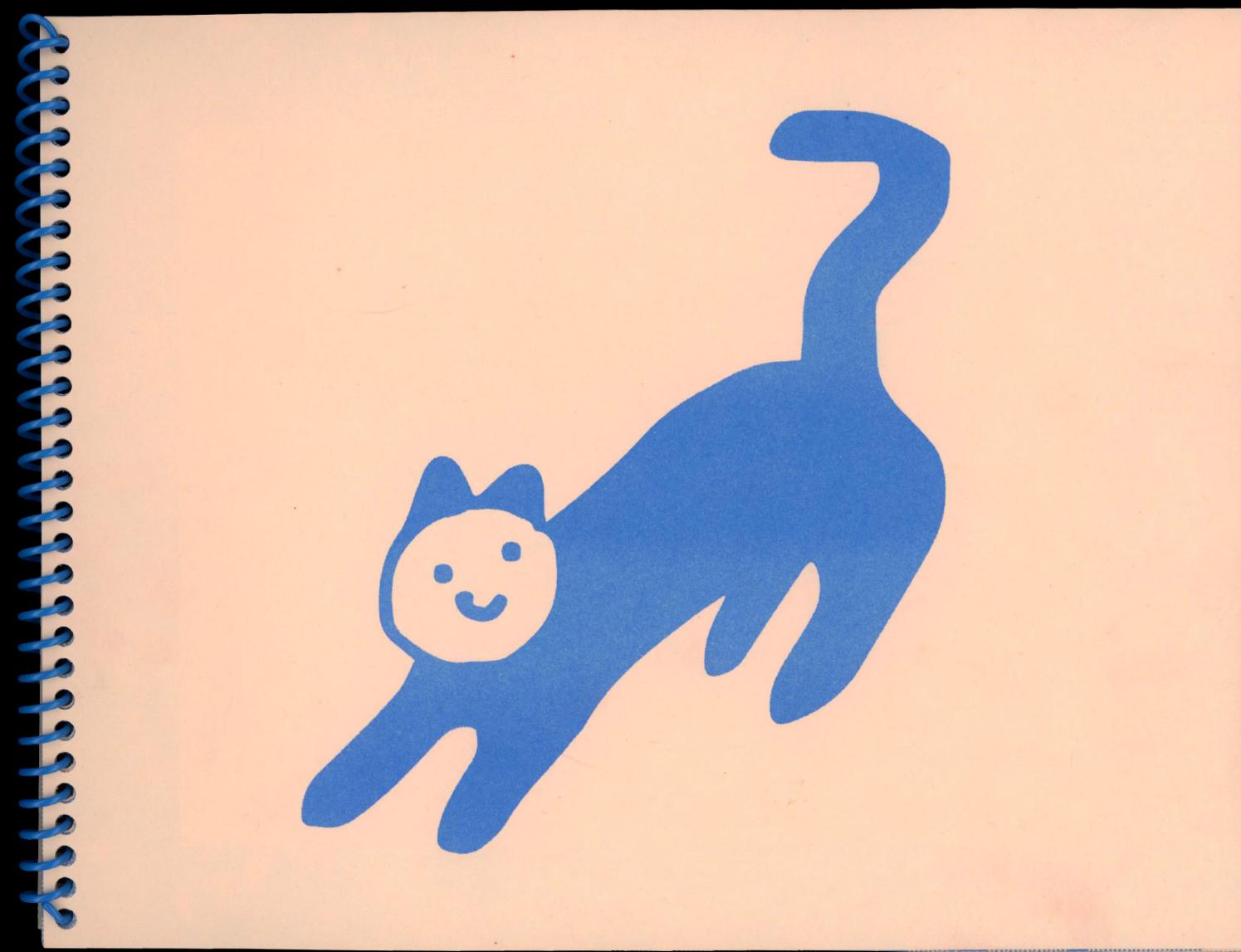
Dear Data

Case Studies:

Aarati Akkapeddi



Faux Foliage (2020)



Permutations Book

Case Studies:

NYT Graphics & The Pudding

A digital publication that...
is the best internet rabbit hole

The Pudding

[ABOUT](#) [SUBSCRIBE](#) [MORE](#)

Find a story...

OUR FAVES POPULAR UPDATING YOUR INPUT VIDEO AUDIO

#203 APR 2025 #202 APR 2025 #201 MAR 2025

wine animals

We try to predict the price and quality of a bottle of wine based on the animal on the label.

music dna legacies

We trace musical DNA from the classical "Hall of the Mountain King" to 2pac's "Hit Em Up."

rhoslc apologies

What we can learn from watching reality stars apologize.

The New York Times

U.S.

ARTICLES FREE TO READ FOR A LIMITED TIME

This year, our visual stories covered a range of subjects: the U.S. presidential race and election, the Olympics, artificial intelligence, extreme weather events, and continued war in Ukraine and Gaza. **2024: The Year in Visual Stories and Graphics.**

By The New York Times Dec. 20, 2024

January February March April May June July August September October November December

January

Michigan, Arizona, Pennsylvania, Georgia, Justice Dept., Mike Pence

Opinion

Relief granted, In ongoing proceedings or legal limbo, Deportation order

