

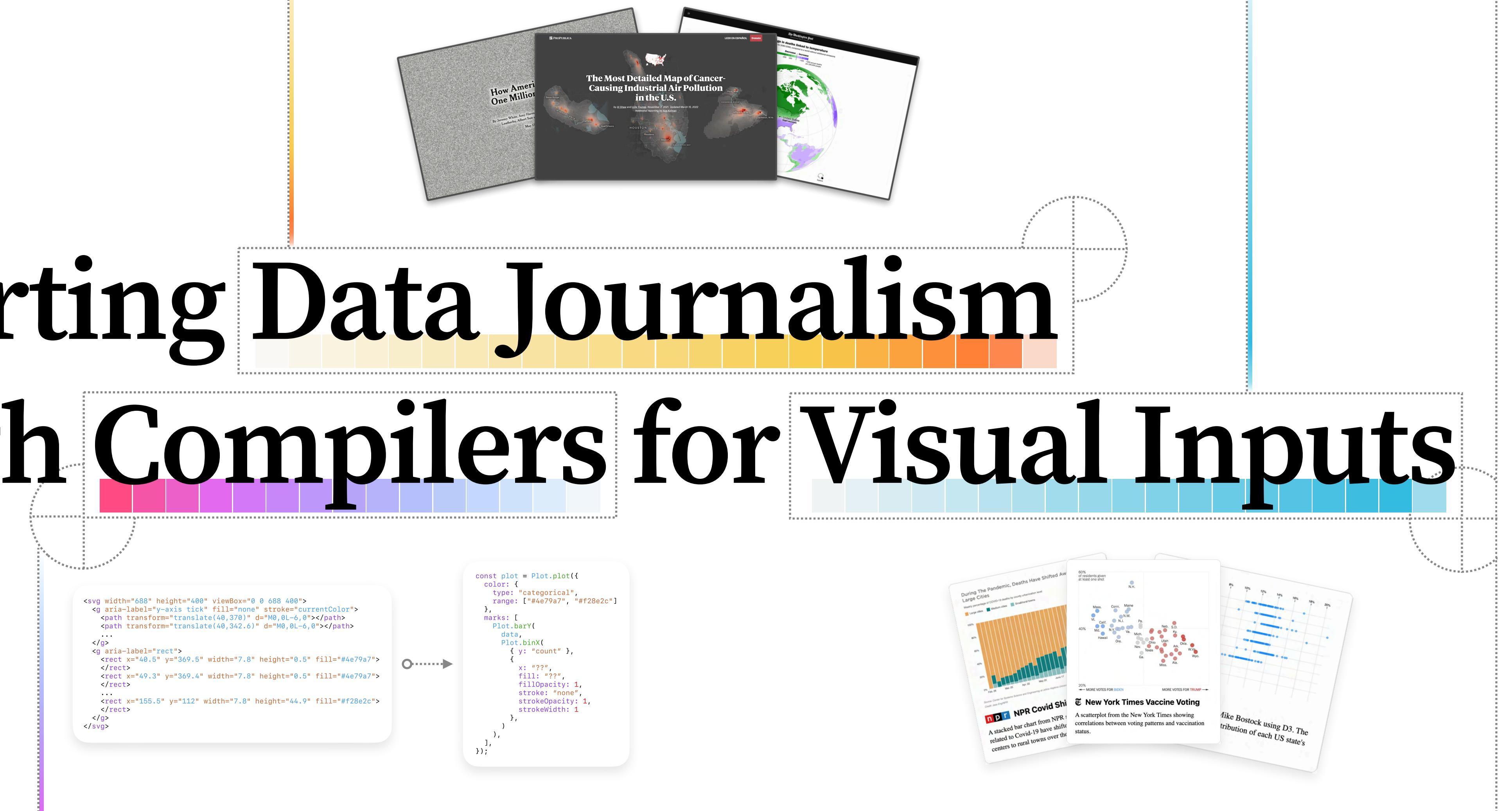
Supporting Data Journalism through Compilers for Visual Inputs

Parker Ziegler // @parker_ziegler

Ph.D. Student, UC Berkeley

∞ Strange Loop

St. Louis, MO • September 21, 2023





Data Journalism

What does
journalism have to
do with **compilation**?

Aren't compilers for transforming
programs, not **charts**?

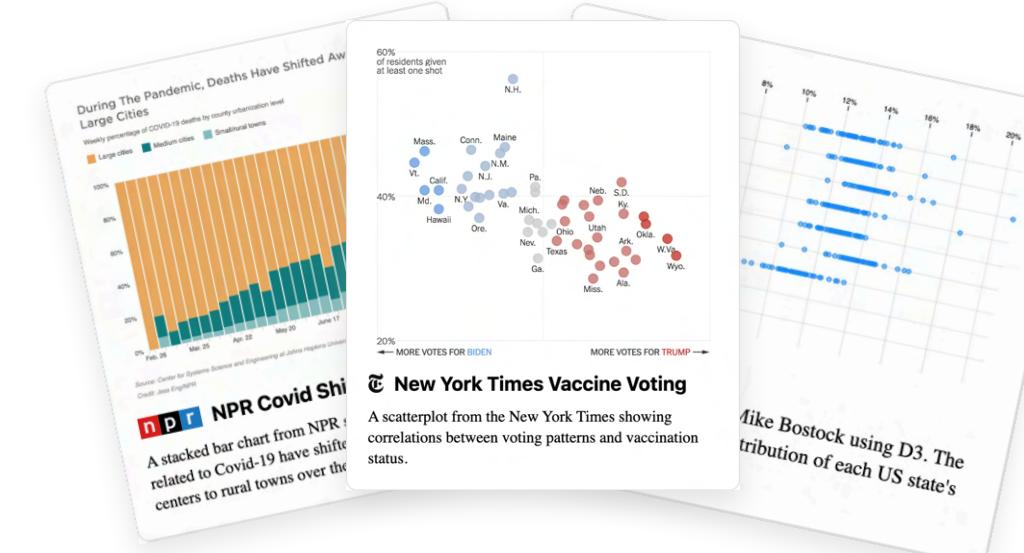
What does this
even mean?

Compilers

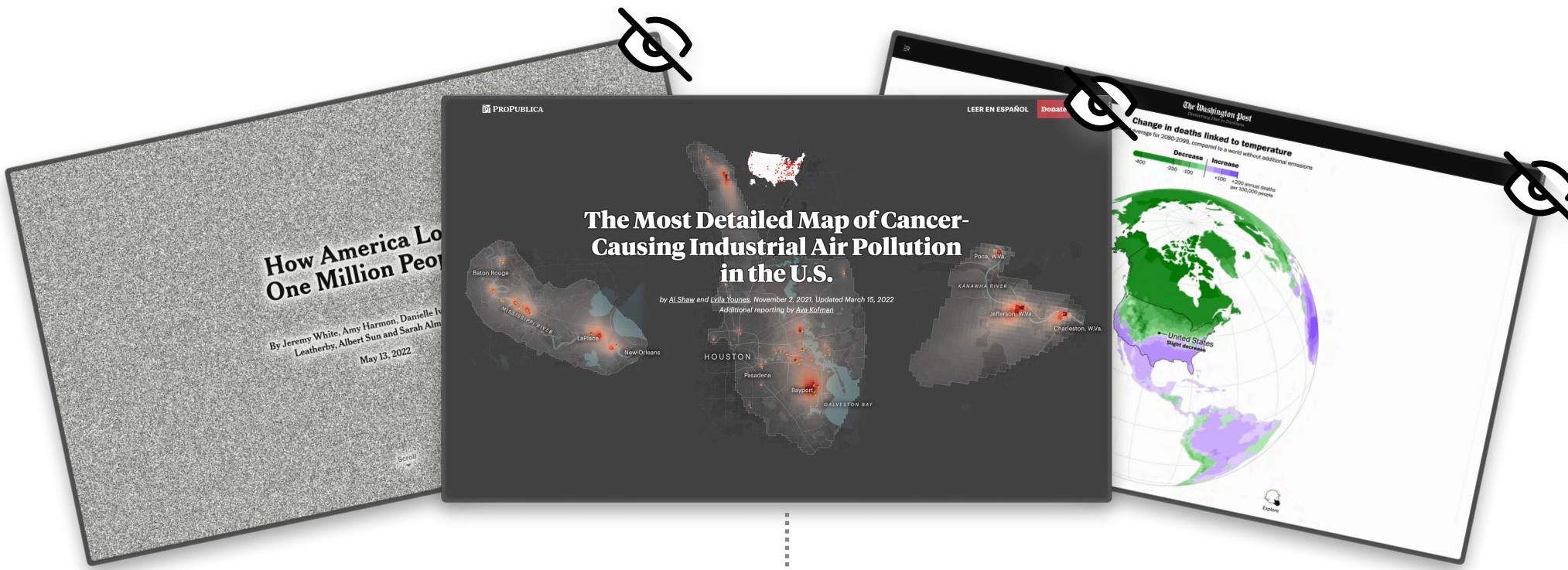
```
<svg width="688" height="400" viewBox="0 0 688 400">
  <g aria-label="y-axis tick" fill="none" stroke="currentColor">
    <path transform="translate(40,370)" d="M0,0L-6,0"></path>
    <path transform="translate(40,342.6)" d="M0,0L-6,0"></path>
  ...
</g>
<g aria-label="rect">
  <rect x="40.5" y="369.6" width="7.8" height="0.5" fill="#4e79a7">
  ...
  <rect x="49.3" y="369.4" width="7.8" height="0.5" fill="#4e79a7">
  ...
  <rect x="155.5" y="112" width="7.8" height="44.9" fill="#f28e2c">
</g>
</svg>
```

```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#4e79a7", "#f28e2c"]
  },
  marks: [
    Plot.barY(
      data,
      Plot.binX(
        { y: "count" },
        {
          x: "?",
          fill: "?",
          fillOpacity: 1,
          stroke: "none",
          strokeOpacity: 1,
          strokeWidth: 1
        }
      )
    ),
  ],
});
```

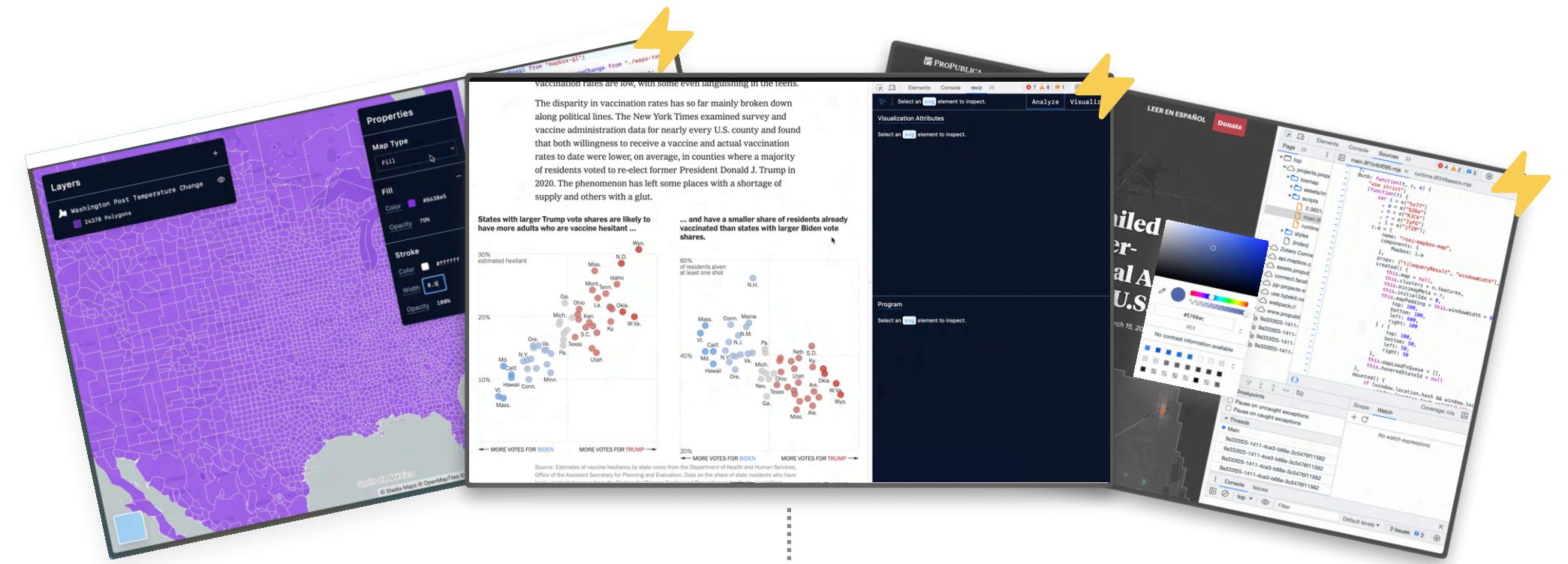
Visual Inputs



Fixed Entities



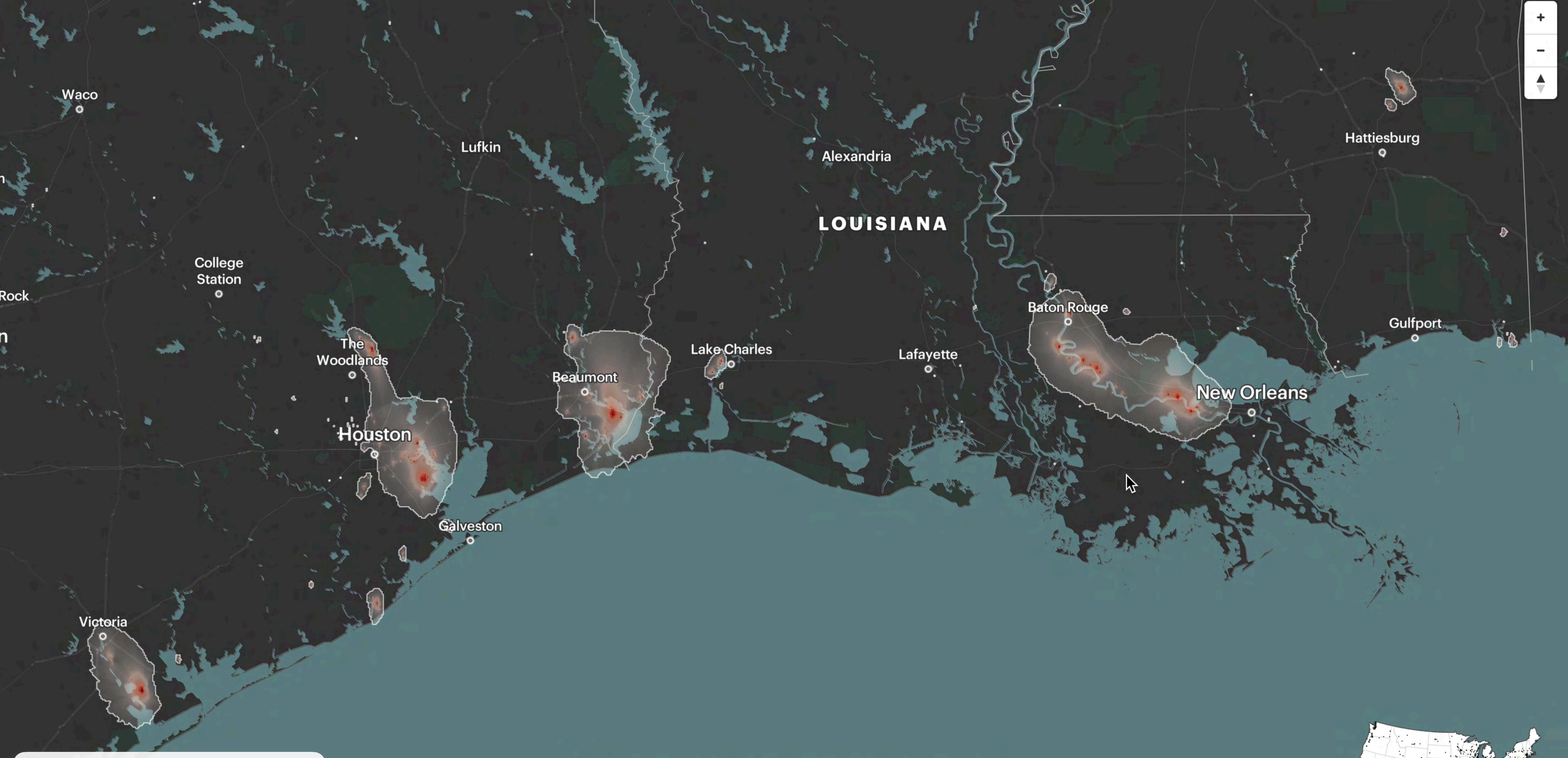
Live Objects



Compilers

Data Journalism





 PROPUBLICA

The Most Detailed Map of Cancer-Causing Industrial Air Pollution in the U.S.

Al Shaw and Lylla Younes

mapbox



Click inside a hot spot to see how risks combine there.



© Mapbox © OpenStreetMap Improve this map

The coronavirus arrived in the United States by early 2020, setting off wave after wave of infection and death in the months that followed.

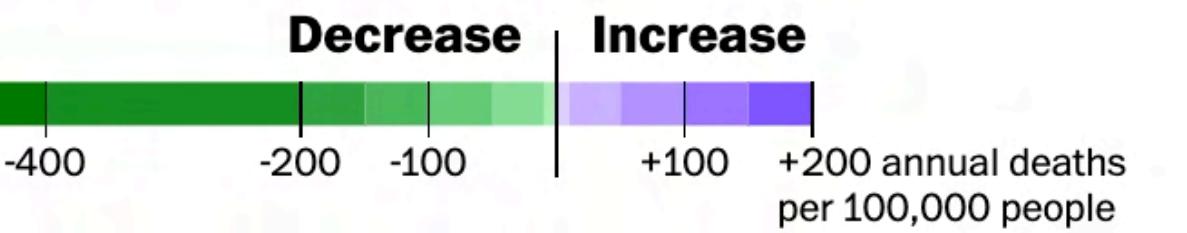
The New York Times

How America Reached One Million Covid Deaths

Jeremy White, Amy Harmon, Danielle Ivory, Lauren Leatherby, Albert Sun, Sarah Almukhtar

Change in deaths linked to temperature

Projected average for 2080-2099, compared to a world without additional emissions



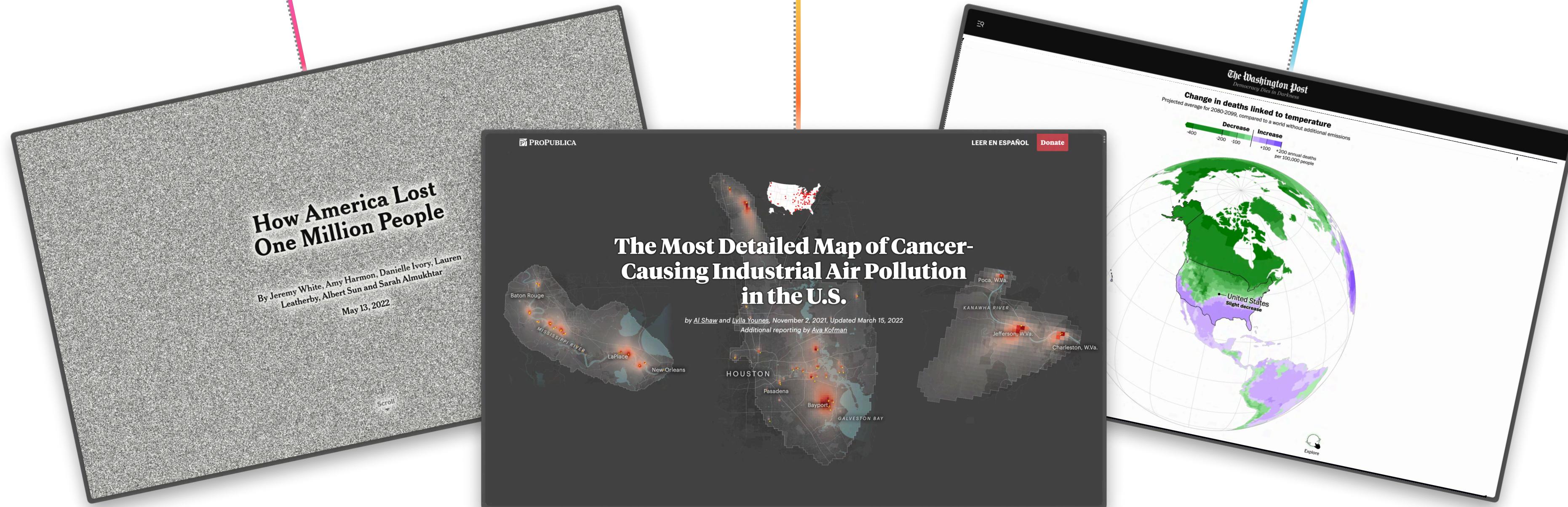
The Washington Post

Will global warming make temperature less deadly?

Harry Stevens

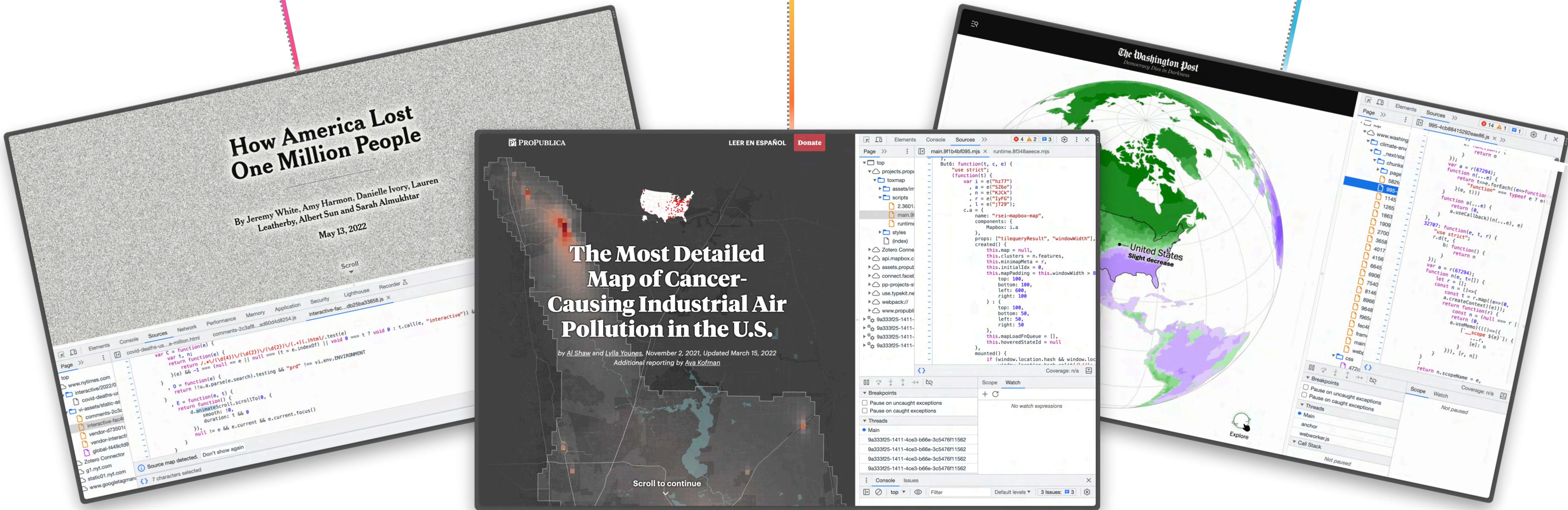


Data Journalism



“Telling **stories** with data”

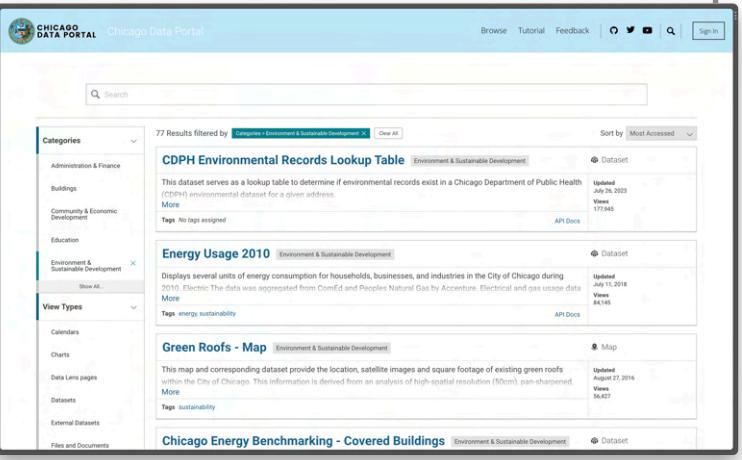
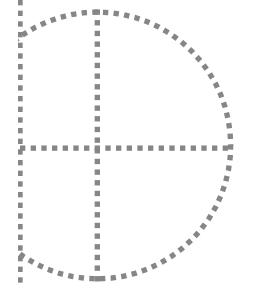
Data Journalism



(Lots and lots of)

Programming

Data Journalism

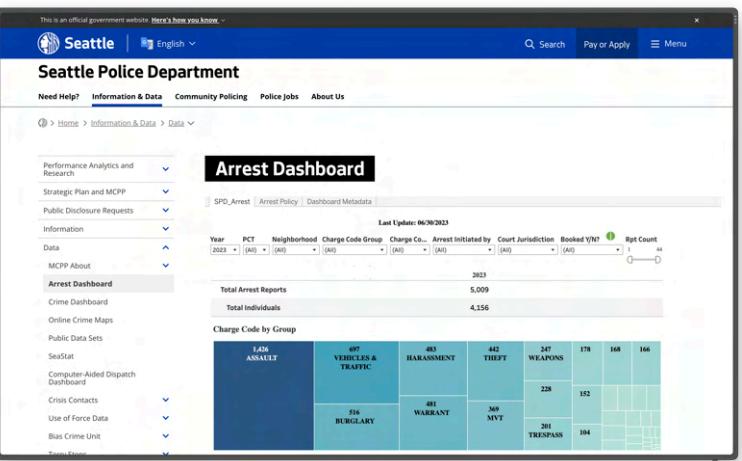


Municipal Agencies



FOIA Requests

Normalize



Police Departments



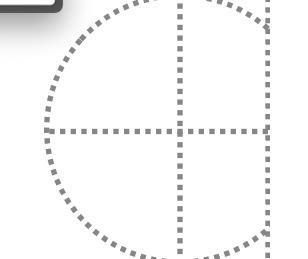
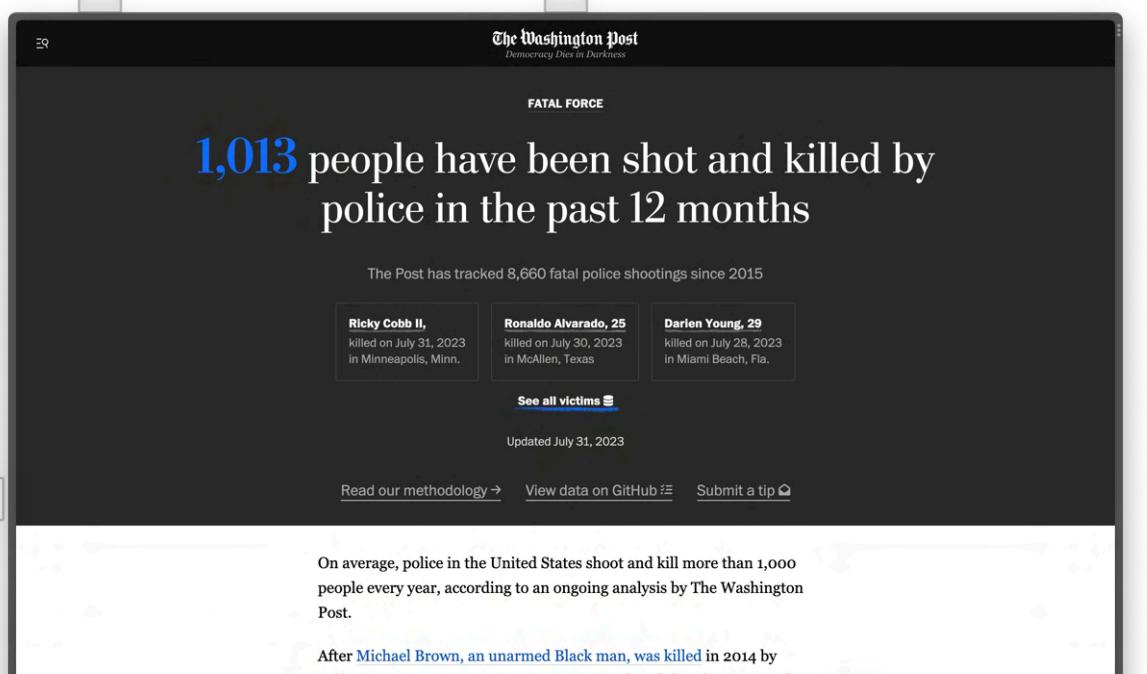
Academic Journals

Filter

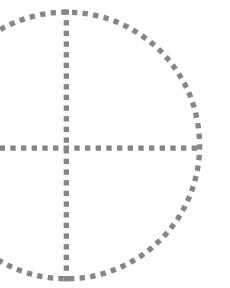
Transform

Clean

1,013 people have been shot and killed by police in the past 12 months



Data Journalism



i http://localhost:8888/lab/tree/landsat-9-surface-temp.ipynb

File Edit View Run Kernel Tabs Settings Help

Launcher landsat-9-surface-temp.ipynb Python 3 (ipykernel)

Filter files by name

Name Last Modified

- landsat-9... seconds ago
- landsat-9... seconds ago
- landsat-9... a month ago

Load Data

```
[1]: import rasterio
from rasterio.plot import show
from rasterio.warp import calculate_default_transform, reproject, Resampling
```

```
[2]: dataset = rasterio.open("landsat-9-surface-temp.tif")
```

```
[3]: show(dataset, cmap="RdBu_r")
```

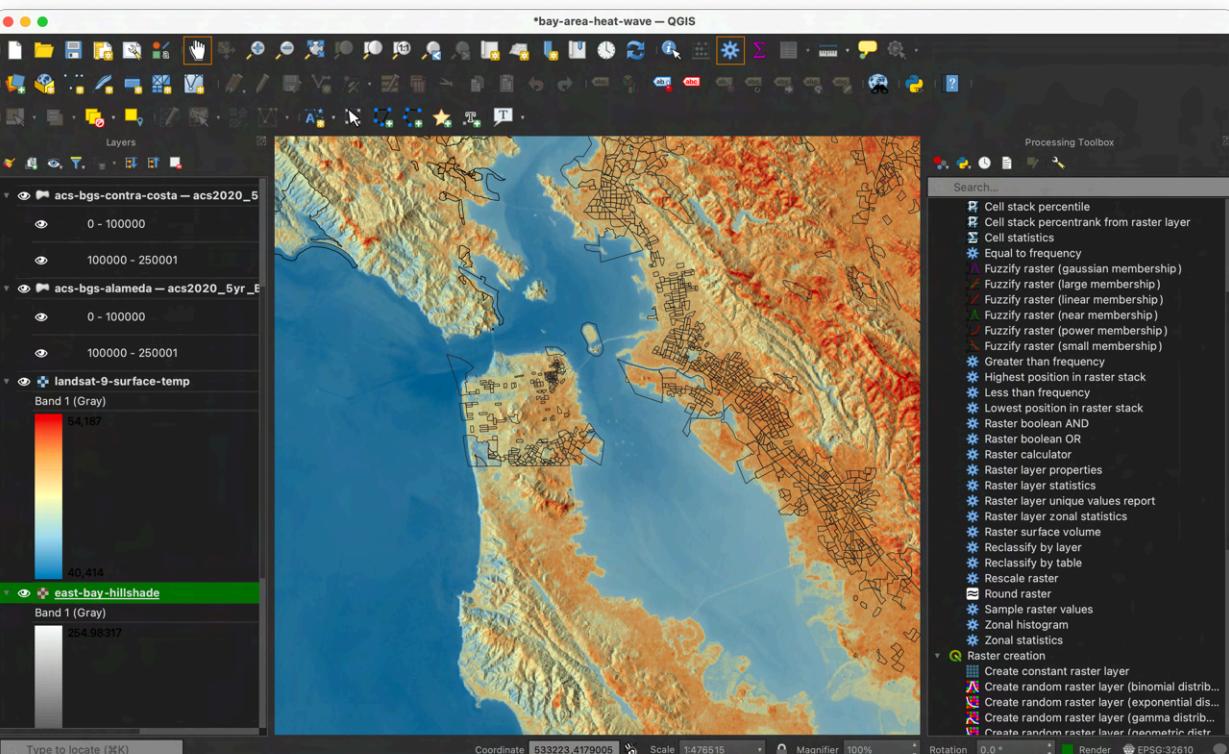
```
[3]: <AxesSubplot>
```

Reproject

Our raster is currently projected in UTM 10N. Let's reproject it into California State Plane Zone III using NAD 1983 as our control datum.

```
[4]: dst_crs = 'EPSG:26943'

with dataset as src:
    transform, width, height = calculate_default_transform(
        src.crs, dst_crs, src.width, src.height, *src.bounds)
```



```
Map.js — police-shootings

src > components > maps > Map > JS Map.js > Map > generateCircle > maxState

65  getChoroplethBreaks() {
66    const { maps } = this.props;
67
68    // for choropleth, generate a set of rects to map over our
69    // data and obtain the shootings per million of each
70    const shootingsArray = sortBy(
71      maps.geoData.objects.states.geometries.map((feature) => {
72        return (
73          feature.properties.numShootings / feature.properties.population) *
74          1000000
75        );
76      })
77    );
78
79    // use quantiles to generate our choropleth breaks
80    const quantiles = [0, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95];
81    let legendValues = uniq(
82      quantiles.map((tick) => round(d3.quantile(shootingsArray, tick)))
83    );
84
85    // if breaks don't include 0 as a value, prepend it
86    if (legendValues[0] !== 0) {
87      legendValues = concat([0], legendValues);
88    }
89
90    return legendValues;
91  }
92
93  generateCircle(geoPath, data) {
94    const { maps, mapType } = this.props;
95
96    // for circles, sort the data in descending order
97    // this ensures that smaller circles will render last - on top of larger circles
98    const statesByShootings = orderBy(
99      maps.geoData.objects.states.geometries,
100      ['properties.numShootings'],
101      ['desc']
102    );
103
104  const maxState = d3.max(statesByShootings,
105    (feature) => feature.properties.numShootings
106  );
```



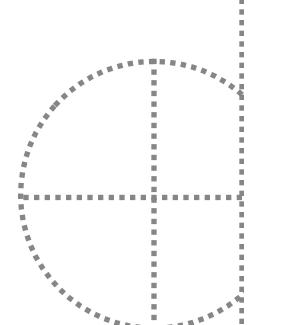
Computational Notebooks

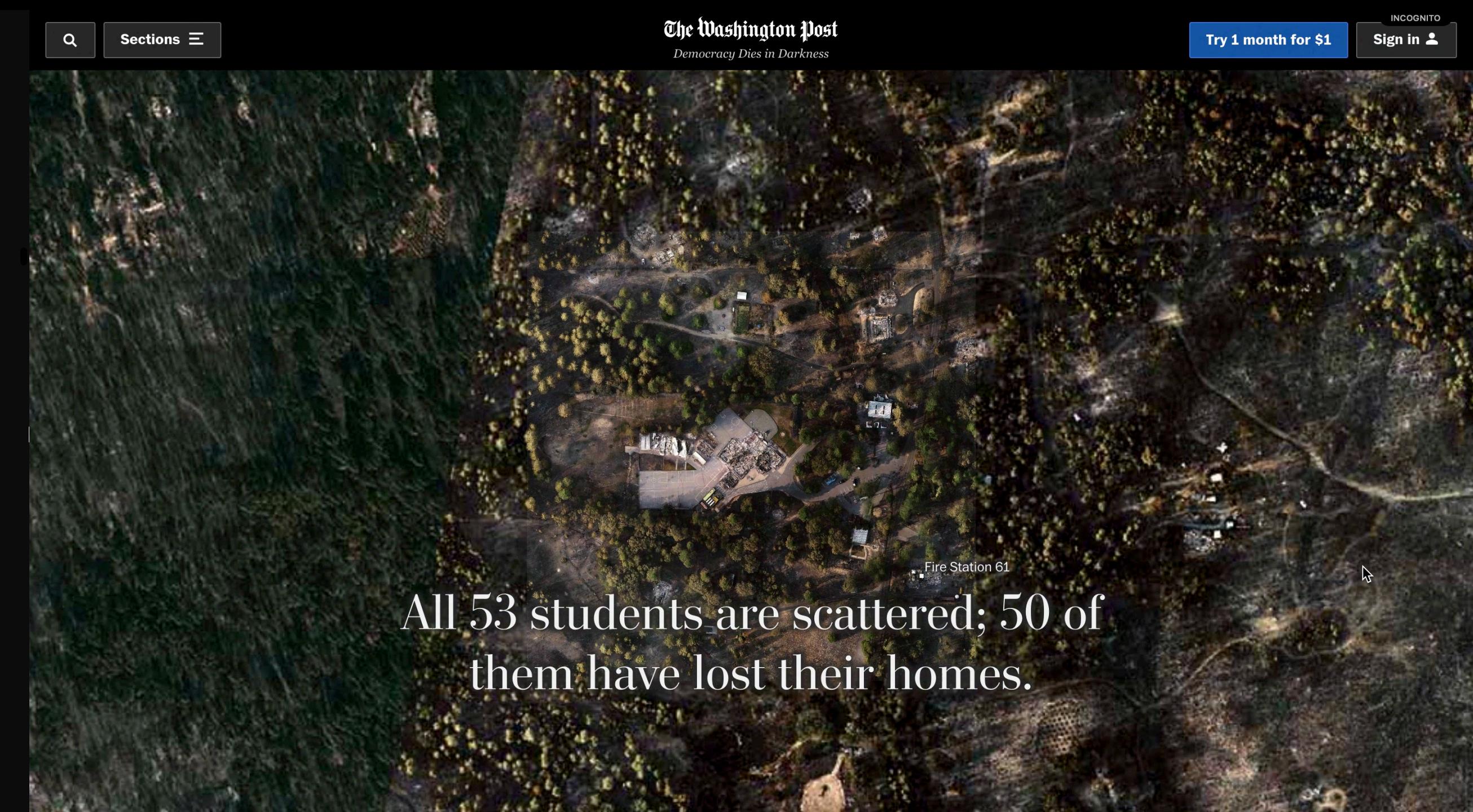
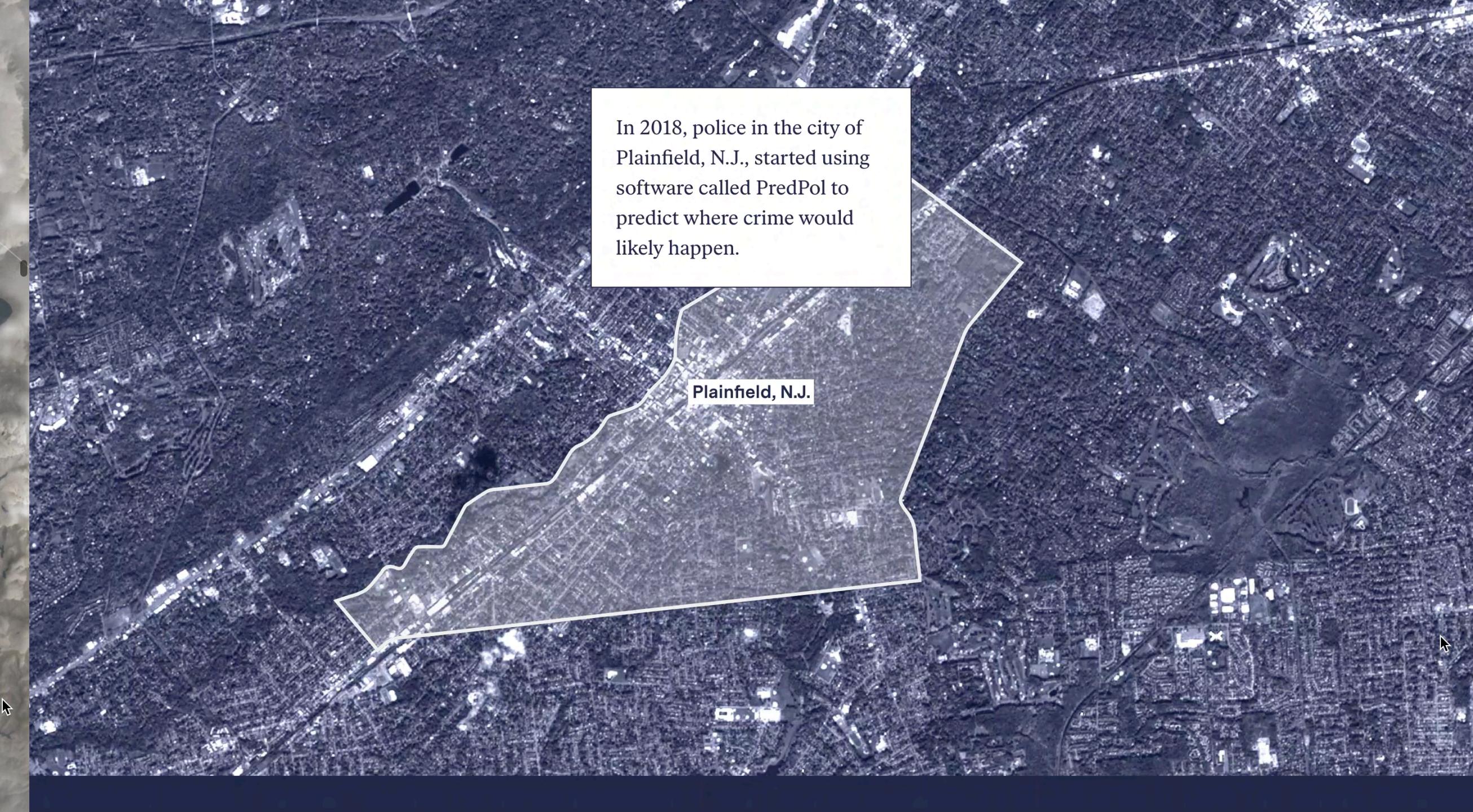
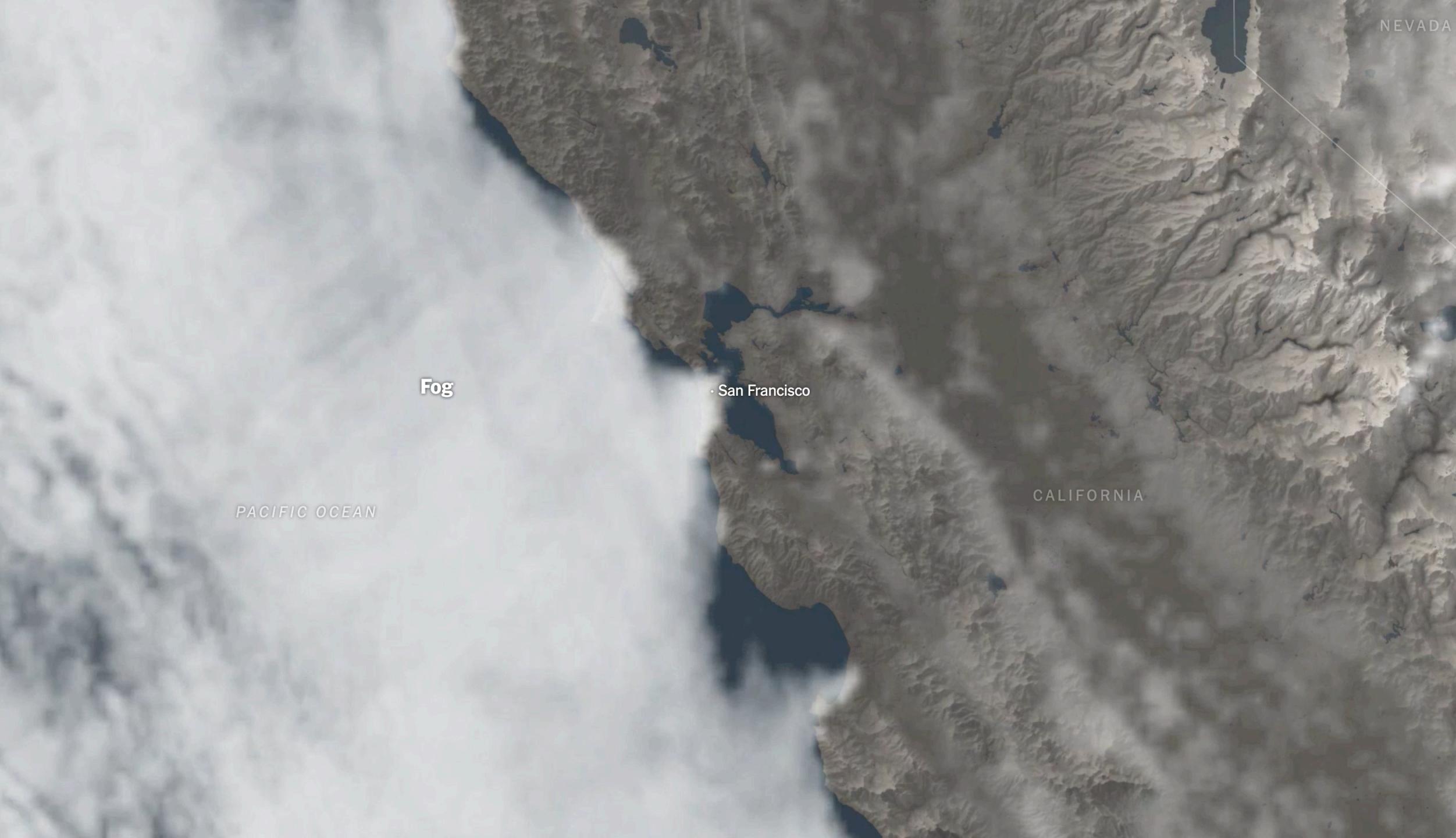


GIS Software

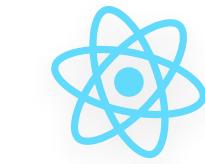
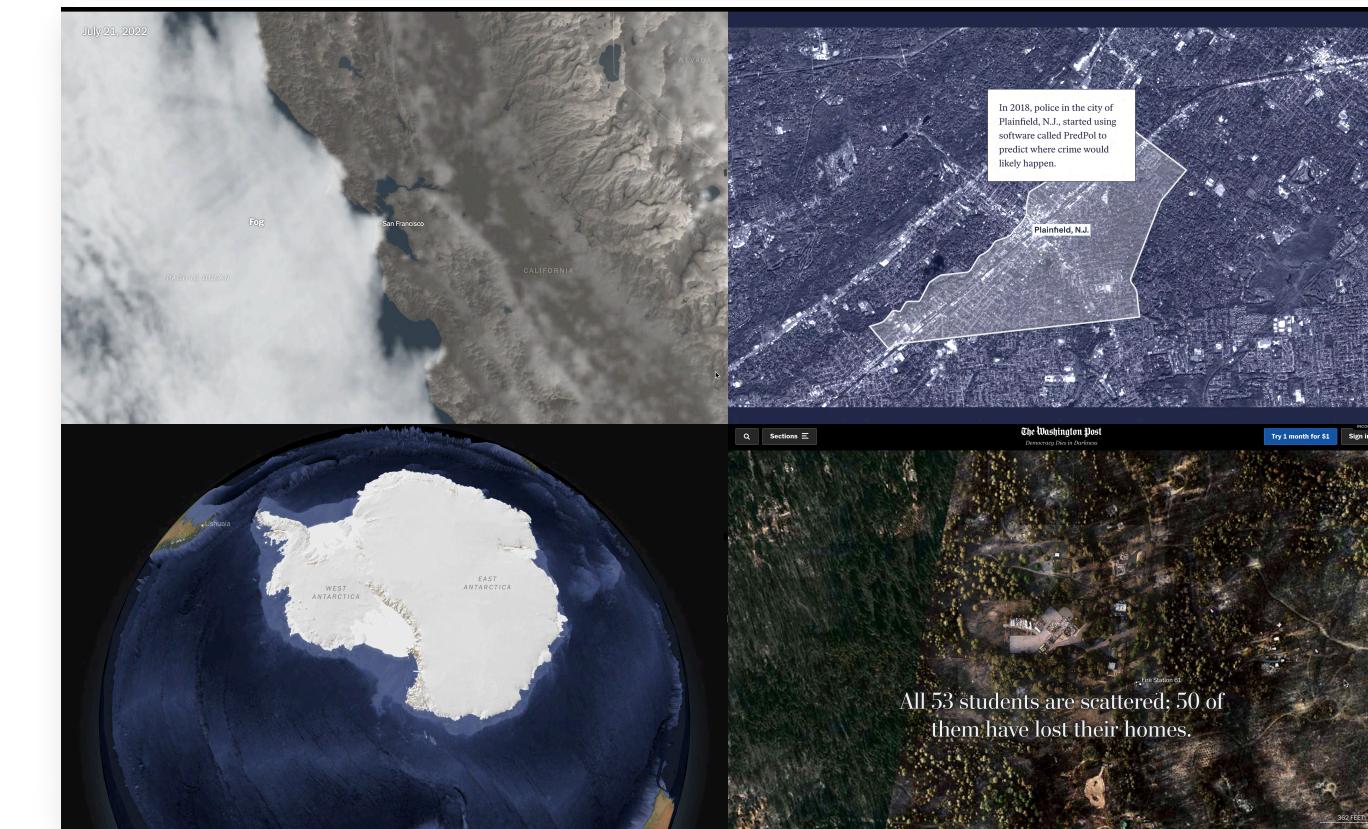
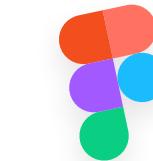
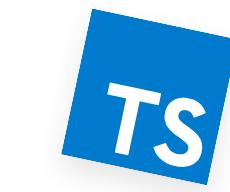


Web Frontends





Data Journalism



Programming

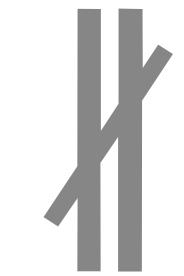
Data Journalism

Visualization

Data Science

Web Dev

Cloud Infrastructure



Software Engineering

Data Journalism

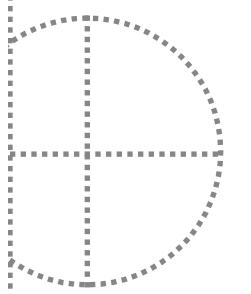
Data Science

Visualization

Web Dev

Cloud Infrastructure

Software Engineering



Reporting

Research

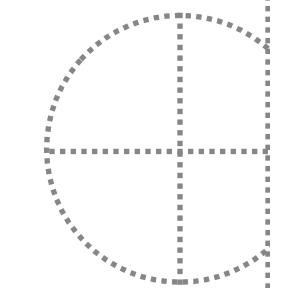
Fact Checking

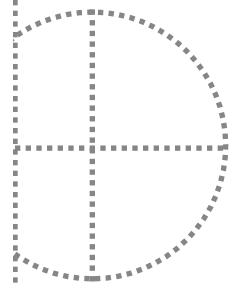
Ground Truthing

Interviews

Data Discovery

Copy Editing





NYT

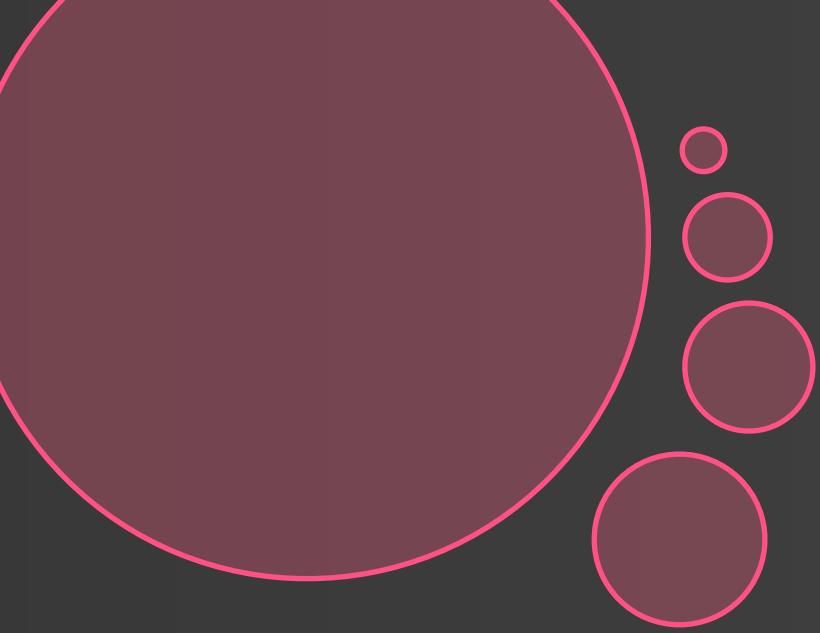
WP

Los Angeles Times

PP PROPUBLICA

FT FiveThirtyEight

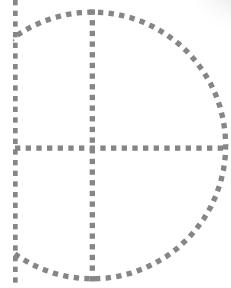
Graphics Desk



Grist



The Markup



Los Angeles Times

P PROPUBLICA

▼ FiveThirtyEight

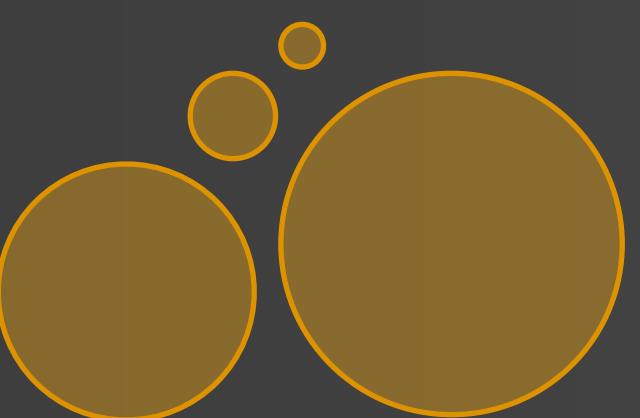
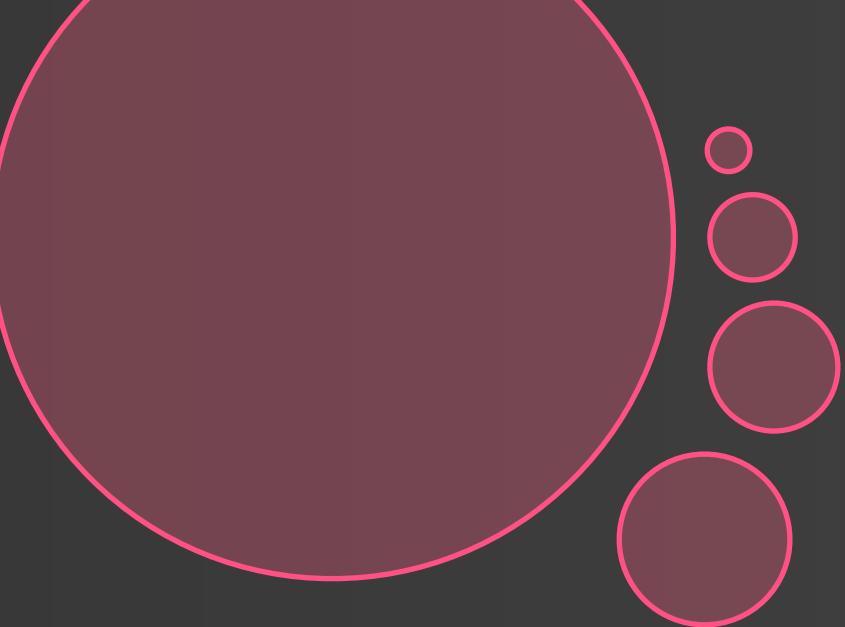
KQED

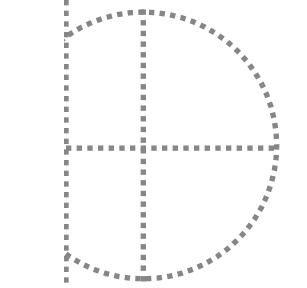
CAL MATTERS

EYE ON OHIO
Ohio Center for Journalism

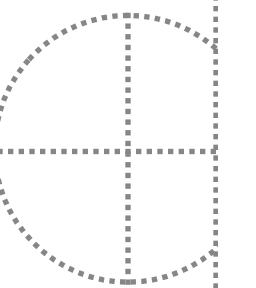
Graphics Desk

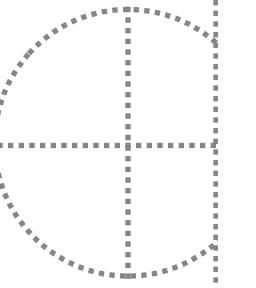
(1 or 2 data journalists)



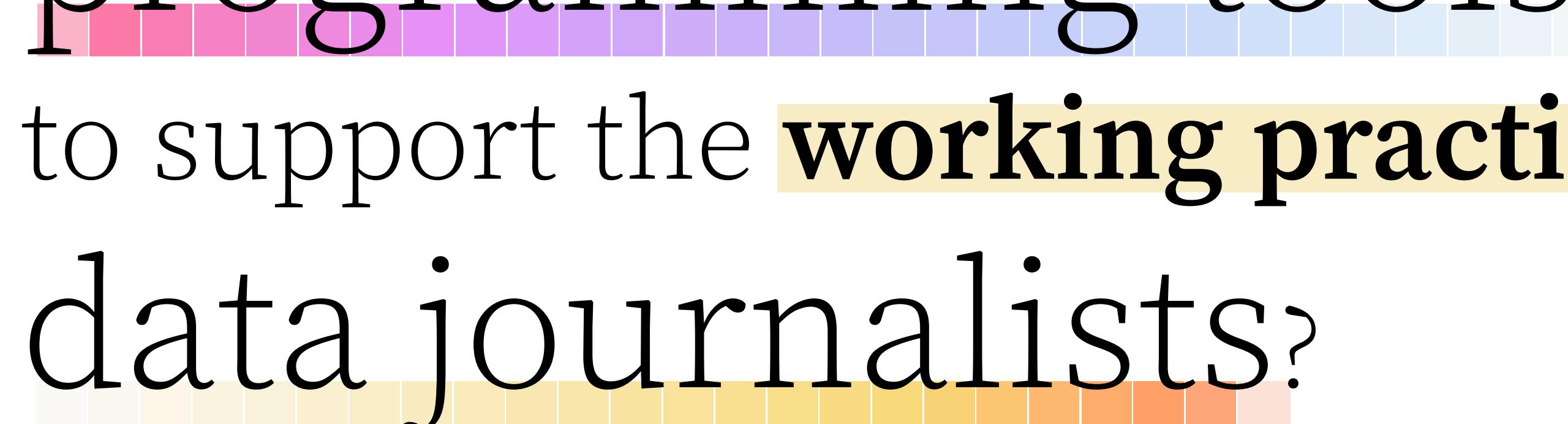


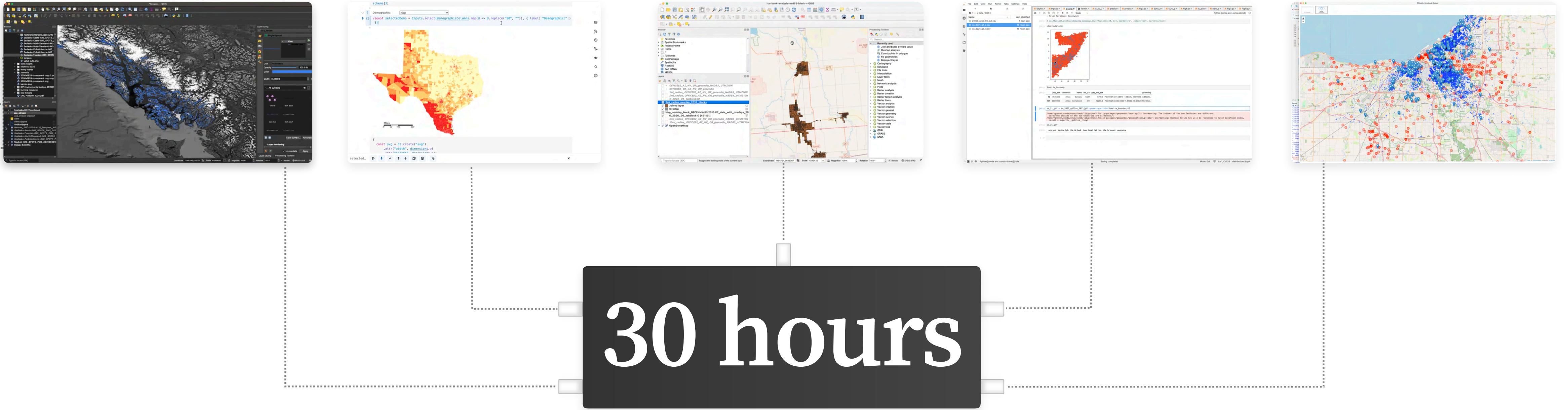
How can we design and build
programming tools
to support the working practices of
data journalists?





How can we design and build
programming tools
to support the **working practices** of
data journalists?





30 hours

of video observation with



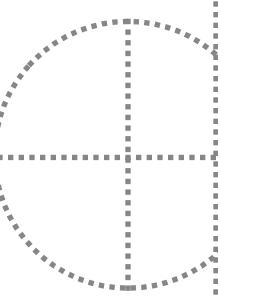
Data
Journalists



Social
Scientists



Earth and
Climate Scientists



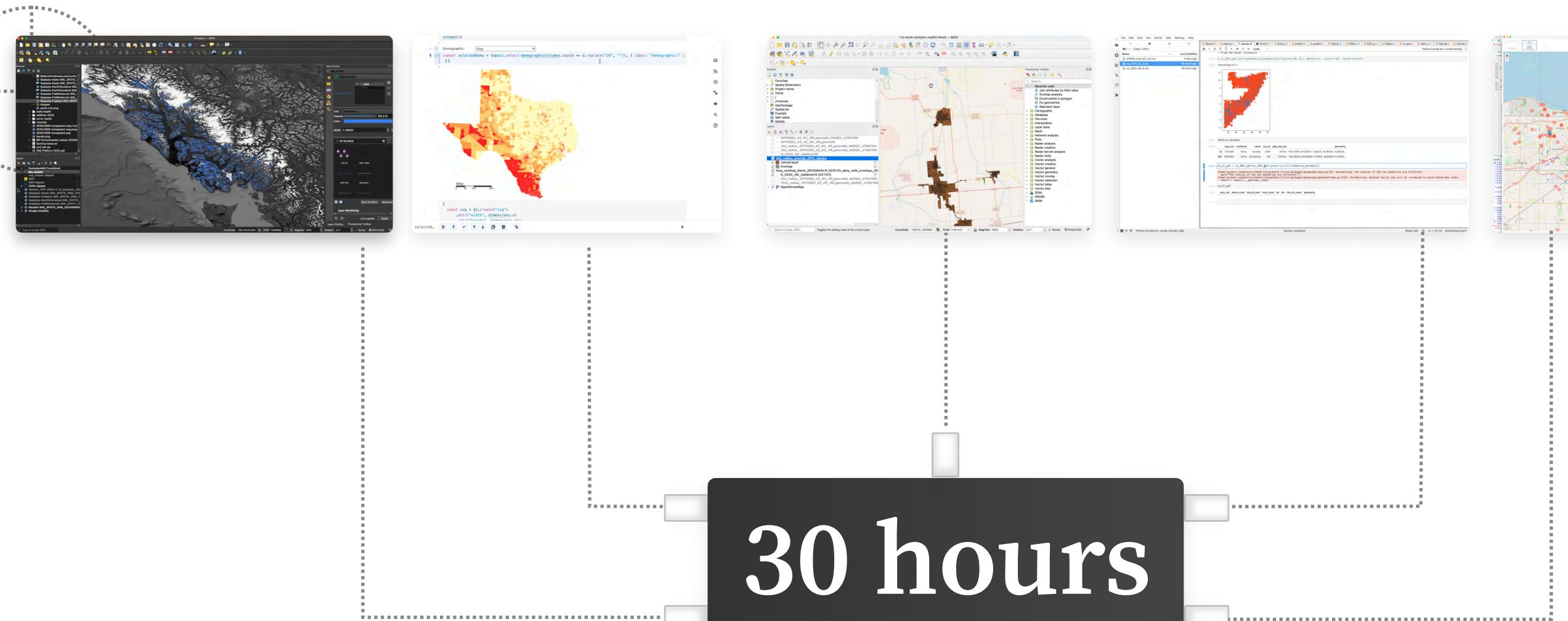
 Data
Journalists

 Social
Scientists

 Earth and
Climate Scientists

30 hours

of video observation with



A Need-Finding Study with Users of Geospatial Data

Parker Ziegler
peziegler@cs.berkeley.edu
University of California, Berkeley
Berkeley, California, USA

Sarah E. Chasins
schasins@cs.berkeley.edu
University of California, Berkeley
Berkeley, California, USA

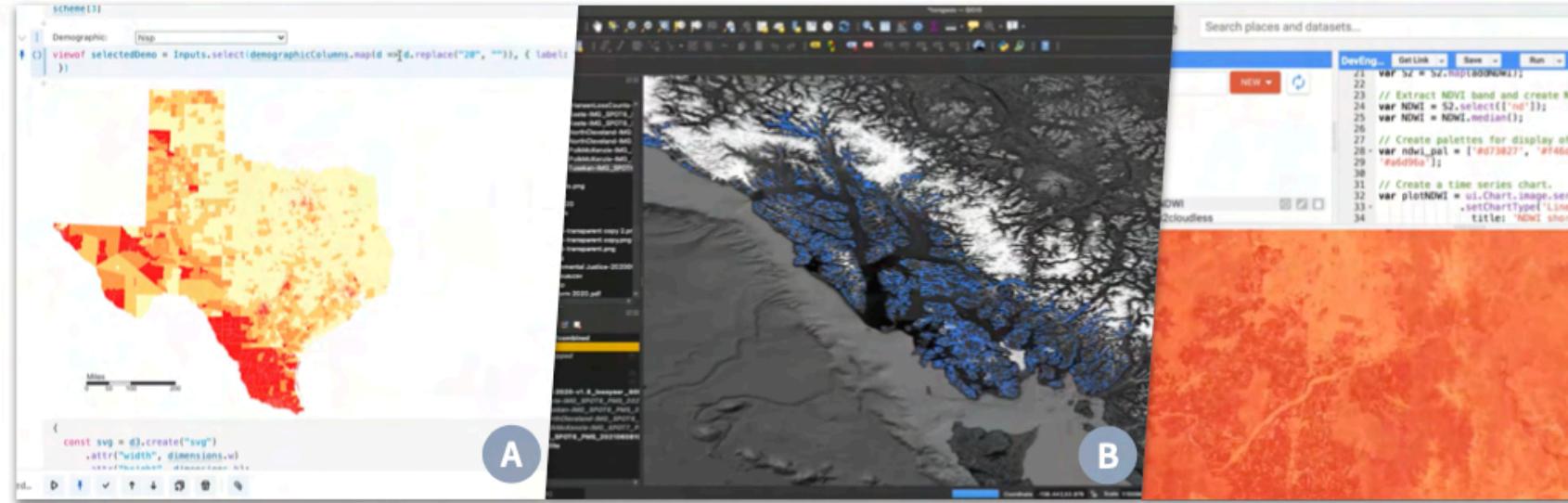


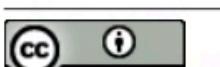
Figure 1: Example screenshots from participants' work with geospatial data. (A) PJ3 creates a choropleth map of Texas proposed electoral districts colored by majority racial demographic in Observable. (B) PJ7 combines satellite imagery data, and deforestation data in QGIS to identify illegal logging in southeast Alaska. (C) PE1 computes a Normalized Difference Water Index of their analysis region in Google Earth Engine using multispectral imagery from the Sentinel-2 satellite.

ABSTRACT

Geospatial data is playing an increasingly critical role in the work of Earth and climate scientists, social scientists, and data journalists exploring spatiotemporal change in our environment and societies. However, existing software and programming tools for geospatial analysis and visualization are challenging to learn and difficult to use. The aim of this work is to identify the unmet computing needs of the diverse and expanding community of geospatial data users. We conducted a contextual inquiry study ($n = 25$) with domain experts using geospatial data in their current work. Through a thematic analysis, we found that participants struggled to (1) find and transform geospatial data to satisfy spatiotemporal constraints, (2) understand the behavior of geospatial operators, (3) track geospatial data provenance, and (4) explore the cartographic design space. These findings suggest design opportunities for developers and designers of geospatial analysis and visualization systems.

CCS CONCEPTS

- Human-centered computing → Human computer interaction (HCI); Empirical studies in HCI; Interactive systems and tools.



This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '23, April 23–28, 2023, Hamburg, Germany
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9421-5/23/04.
<https://doi.org/10.1145/3544548.3581370>

KEYWORDS

geospatial data, GIS, geography, cartography, contextual need-finding

ACM Reference Format:

Parker Ziegler and Sarah E. Chasins. 2023. A Need-Finding Study with Users of Geospatial Data. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3544548.3581370>

1 INTRODUCTION

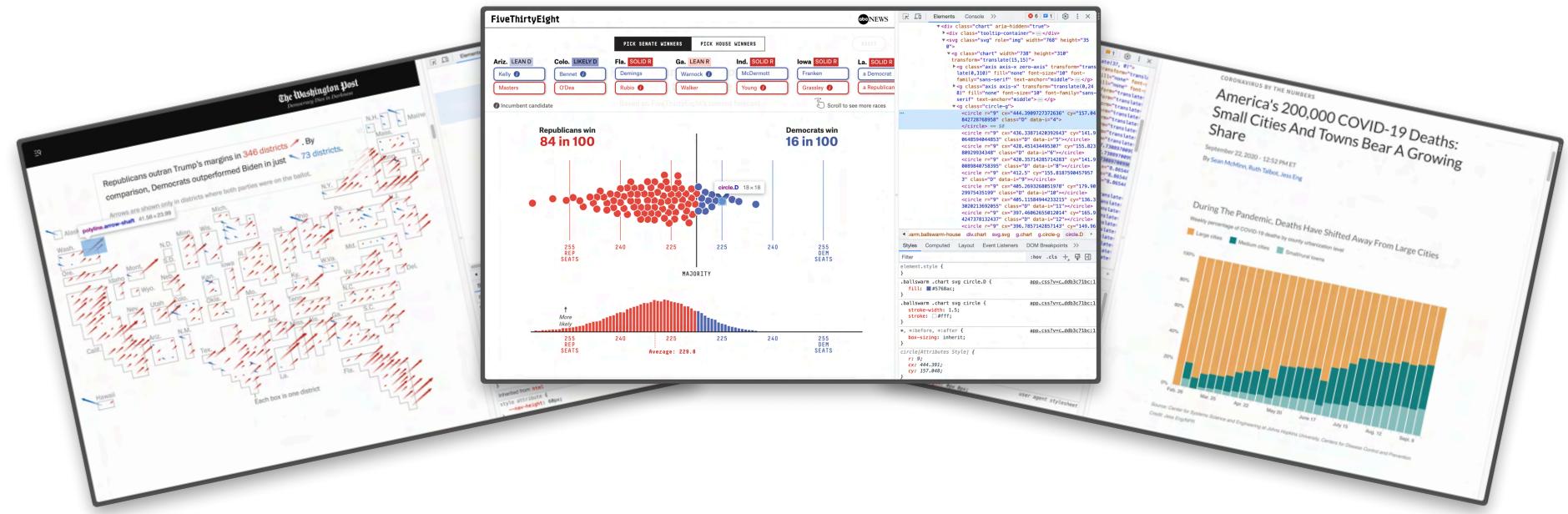
Geospatial data—data encoding the location and attributes of phenomena on the Earth's surface [59]—is growing in scale and availability at a tremendous rate [61]. Researchers estimate that Earth observation satellites generate 80TB of new imagery daily [8] due to the surface, cheap, power-efficient sensors create massive volumes of geolocated data measuring real-time environmental data [40]. Additionally, crowdsourcing efforts like OpenStreetMap have fostered an explosion in publicly available volunteered geographic information [49, 78]. Geospatial data has long played a fundamental role in the research of geographers and cartographers, and as data becomes more available, experts across a widening range of domains are turning to geospatial analysis and visualization to address challenges in climate change [17], public health [34], urban segregation [82], hazard modeling [98], and other areas.

Despite this expansion in the community of geospatial data users, research has yet to explore the specific challenges that experts face in gathering, analyzing, and visualizing geospatial information. Many domain experts are self-taught in the use of geospatial tools, and those who receive formal training often do so through informal means such as online forums or user groups. This lack of formal education can lead to a lack of understanding of key concepts and best practices in geospatial analysis and visualization, which can result in suboptimal results and increased cognitive load for users.



Data Journalists

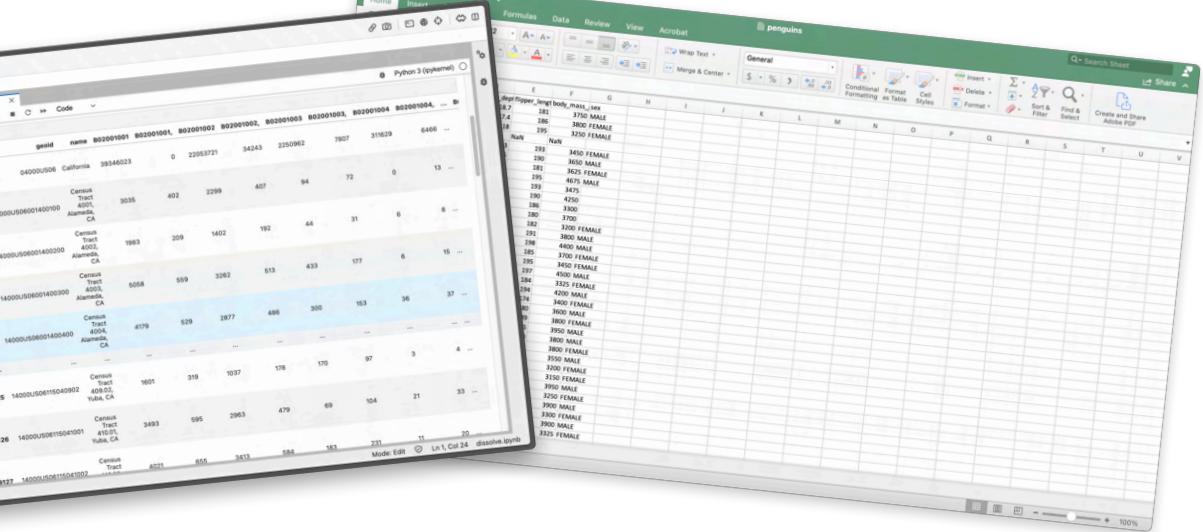
Lift **visual styles** and **graphical forms** from **examples**



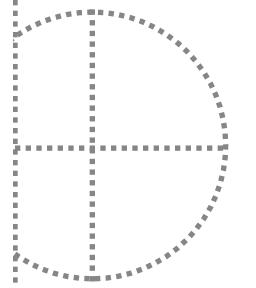
Social Scientists



Earth and Climate Scientists



Apply to new datasets



Lift visual styles and graphical forms from examples



Apply to new datasets



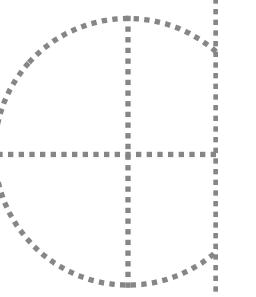
Data
Journalists

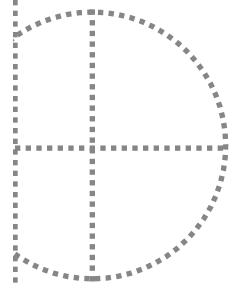


Social
Scientists

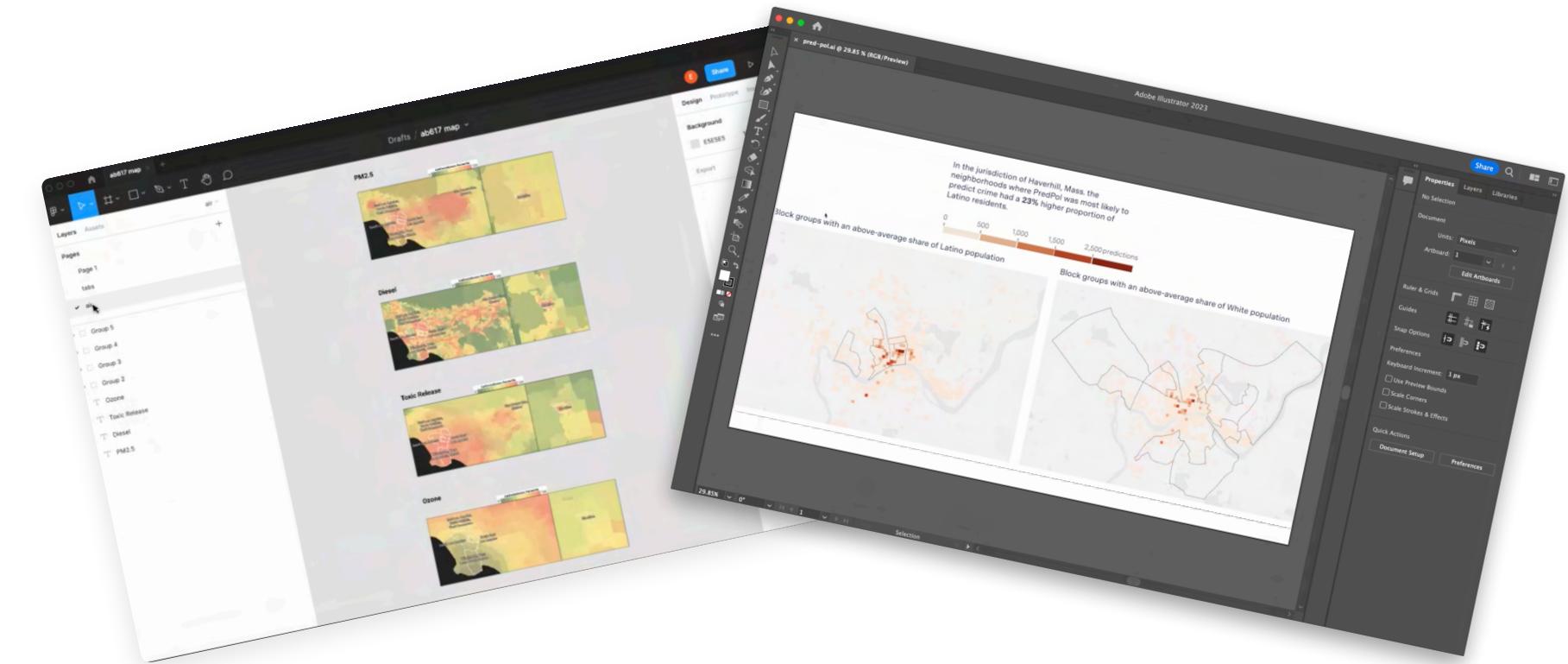


Earth and
Climate Scientists

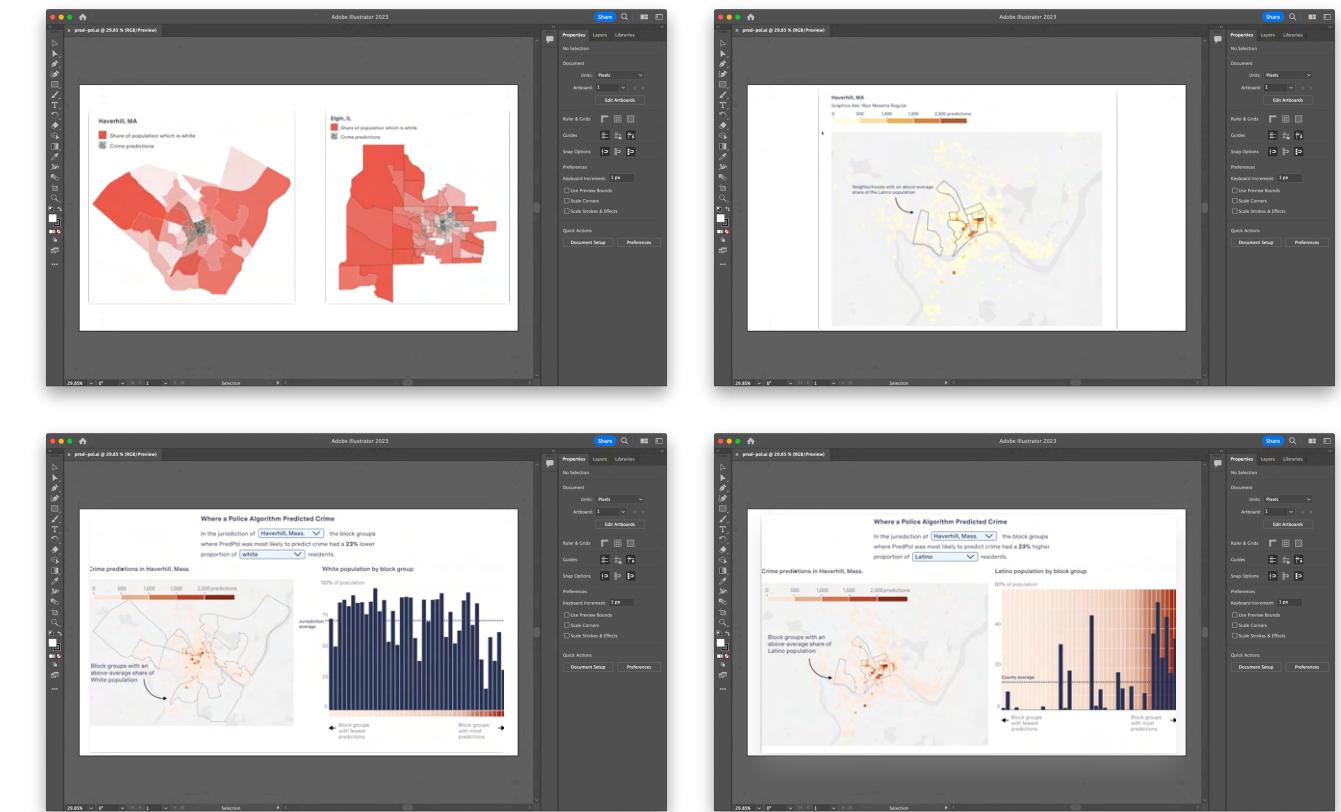




Sketch visualizations in **design software**



.....◦ Explore a wide design space **without code**



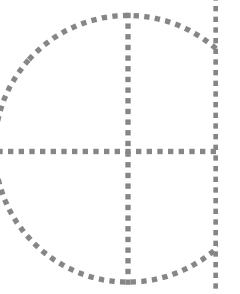
**Data
Journalists**

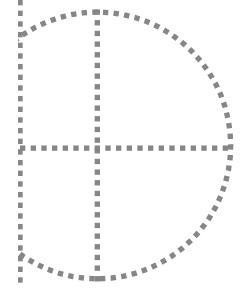


**Social
Scientists**



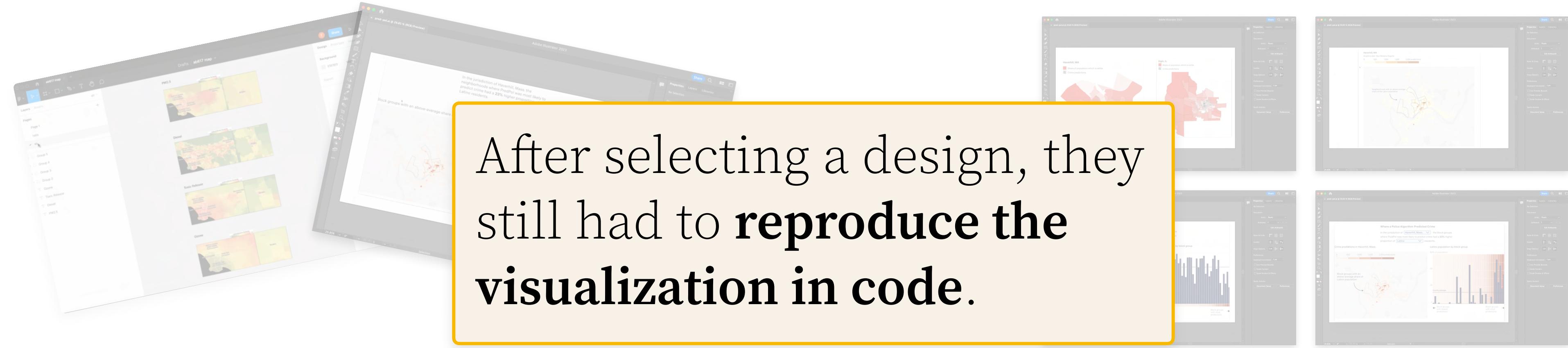
**Earth and
Climate Scientists**





Sketch visualizations in **design software**

.....◦ Explore a wide design space **without code**



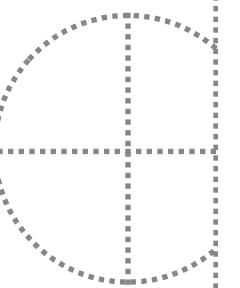
Data
Journalists

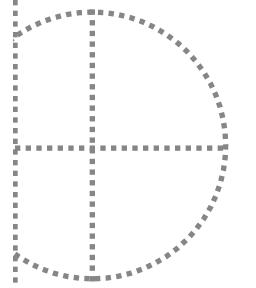


Social
Scientists

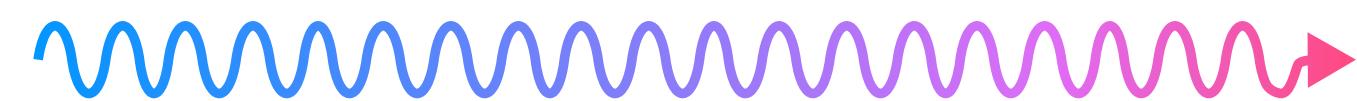


Earth and
Climate Scientists





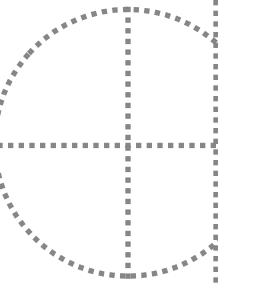
Visual Inputs



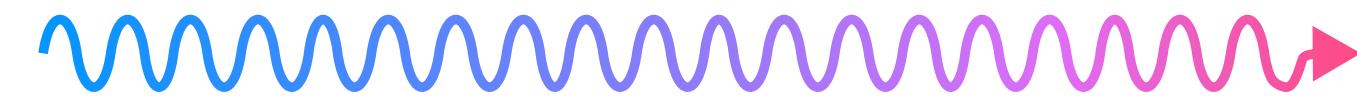
Programs



```
const x = d3.scaleLinear()  
  .domain(d3.extent(cars, d => d["weight (lb)"]))  
  .range([marginLeft, width - marginRight]);  
  
const svg = d3.create("svg")  
  .attr("width", width)  
  .attr("height", height)  
  .attr("viewBox", [0, 0, width, height])  
  .attr("style", "max-width: 100%; height: auto;");  
  
svg.append("g")  
  .attr("transform", `translate(0, ${height - marginBottom})`)  
  .call(d3.axisBottom(x).tickSizeOuter(0));  
  
svg.append("g")  
  .selectAll()  
  .data(dodge(cars, {radius: radius * 2 + padding}))  
  .join("circle")  
    .attr("cx", d => d.x)  
    .attr("cy", d => height - marginBottom - radius - padding - d.y)  
    .attr("r", radius)  
  .append("title")  
  .text(d => d.data.name);
```



Visual Inputs



Programs

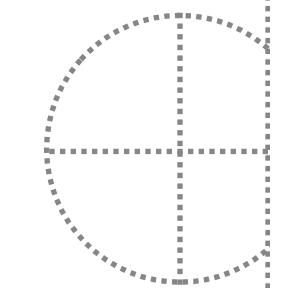


Ai

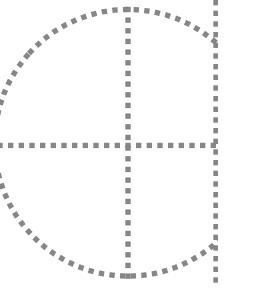
AI



```
const x = d3.scaleLinear()  
  .domain(d3.extent(cars, d => d["weight (lb)"]))  
  .range([marginLeft, width - marginRight]);  
  
const svg = d3.create("svg")  
  .attr("width", width)  
  .attr("height", height)  
  .attr("viewBox", [0, 0, width, height])  
  .attr("style", "max-width: 100%; height: auto;");  
  
svg.append("g")  
  .attr("transform", `translate(0, ${height - marginBottom})`)  
  .call(d3.axisBottom(x).tickSizeOuter(0));  
  
svg.append("g")  
  .selectAll()  
  .data(dodge(cars, {radius: radius * 2 + padding}))  
  .join("circle")  
    .attr("cx", d => d.x)  
    .attr("cy", d => height - marginBottom - radius - padding - d.y)  
    .attr("r", radius)  
  .append("title")  
  .text(d => d.data.name);
```



Could this be a **compilers** problem?



Maybe?

Could this be a **compilers** problem?



How do we compile from a **visual form** to a **textual symbolic representation**?

```
use wasm_bindgen::prelude::*;

#[wasm_bindgen]
pub fn add(a: u32, b: u32) -> {
    a + b
}
```



```
(module
  (type (;0;) (func (param i32 i32) (result i32)))
  (func (;0;) (type 0) (param i32 i32) (result i32)
        get_local 1
        get_local 0
        i32.add)
  (table (;0;) 1 1 anyfunc)
  (memory (;0;) 17)
  (global (;0;) i32 (i32.const 1049118))
  (global (;1;) i32 (i32.const 1049118))
  (export "memory" (memory 0))
  (export "__indirect_function_table" (table 0))
  (export "__heap_base" (global 0))
  (export "__data_end" (global 1))
  (export "add" (func 0))
  (data (i32.const 1049096) "invalid malloc request"))
```



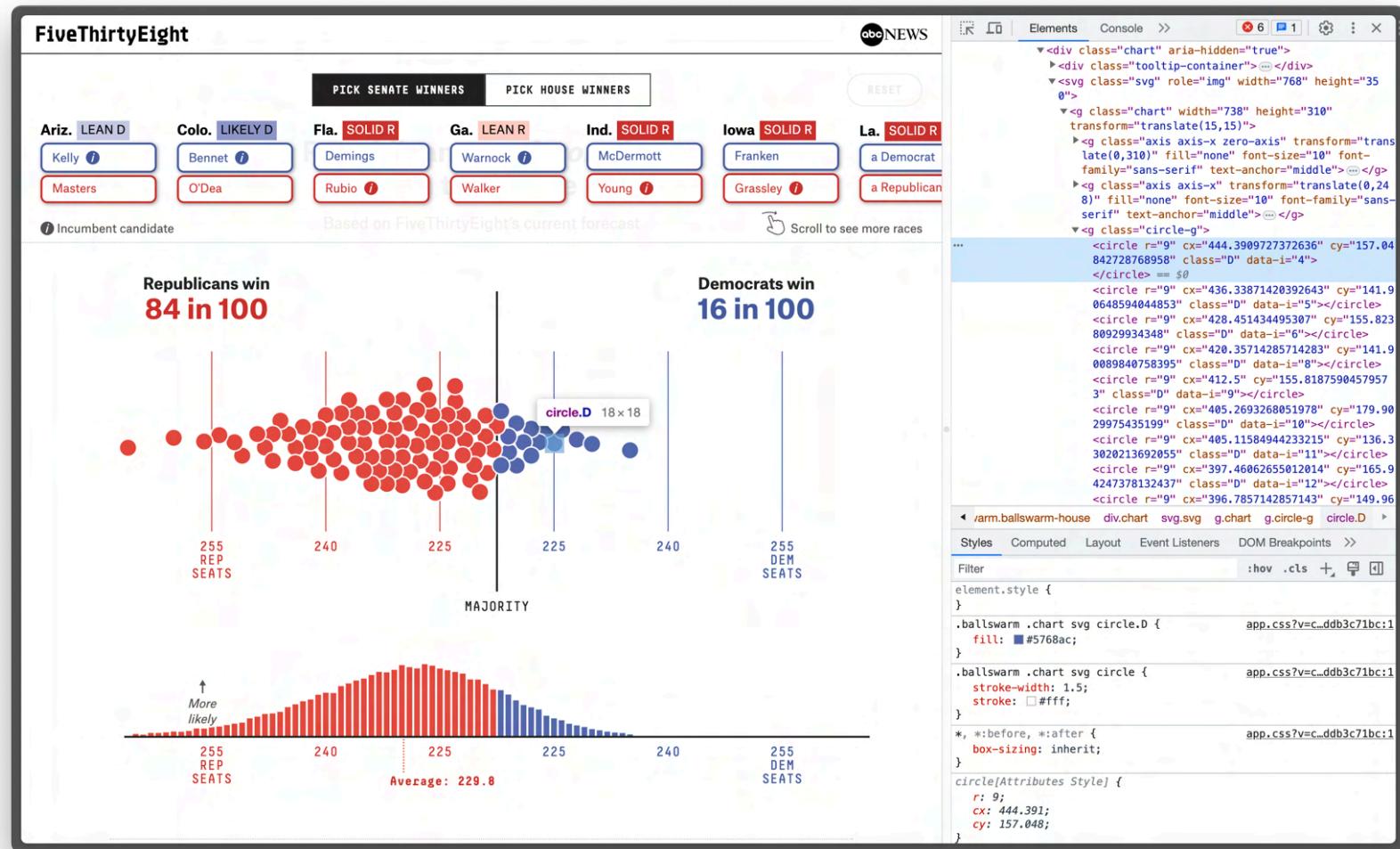
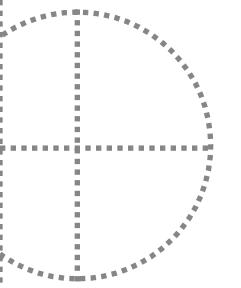
Lexer

Parser

Interm. Repr.

Code Generator

How do we compile from a visual form to a textual symbolic representation?



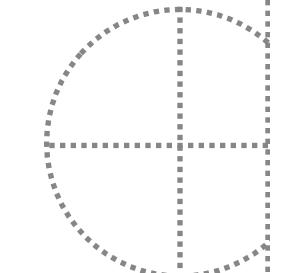
```
(module
  (type (;0;) (func (param i32 i32) (result i32)))
  (func (;0;) (type 0) (param i32 i32) (result i32)
    get_local 1
    get_local 0
    i32.add)
  (table (;0;) 1 1 anyfunc)
  (memory (;0;) 17)
  (global (;0;) i32 (i32.const 1049118))
  (global (;1;) i32 (i32.const 1049118))
  (export "memory" (memory 0))
  (export "__indirect_function_table" (table 0))
  (export "__heap_base" (global 0))
  (export "__data_end" (global 1))
  (export "add" (func 0))
  (data (i32.const 1049096) "invalid malloc request"))
```

Lexer

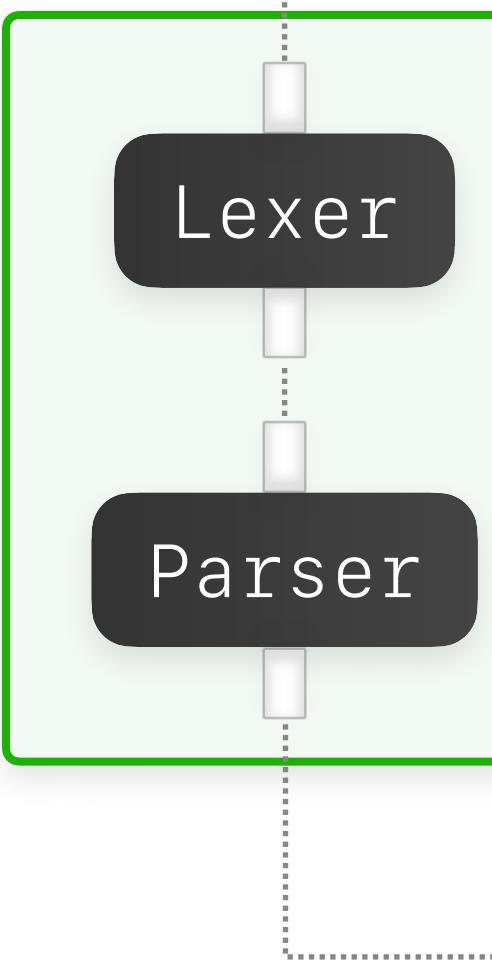
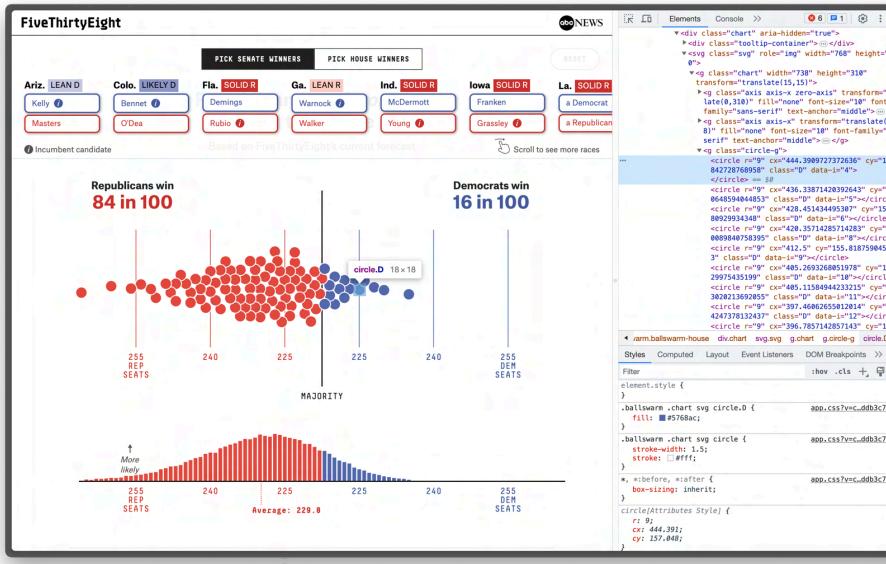
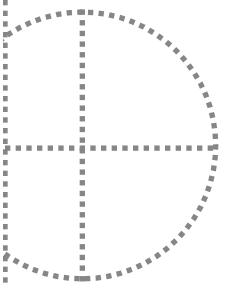
Parser

Interm. Repr.

Code Generator

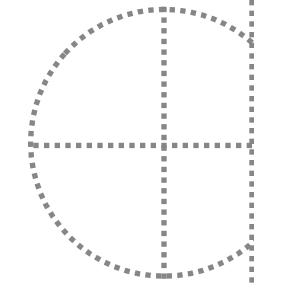


How do we compile from a **visual form** to a **textual symbolic representation**?

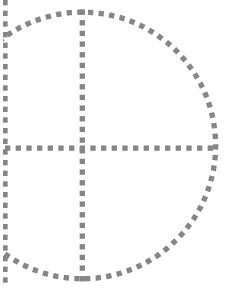


Recover (or infer) **semantic information** from the visual form that we can represent **symbolically**.

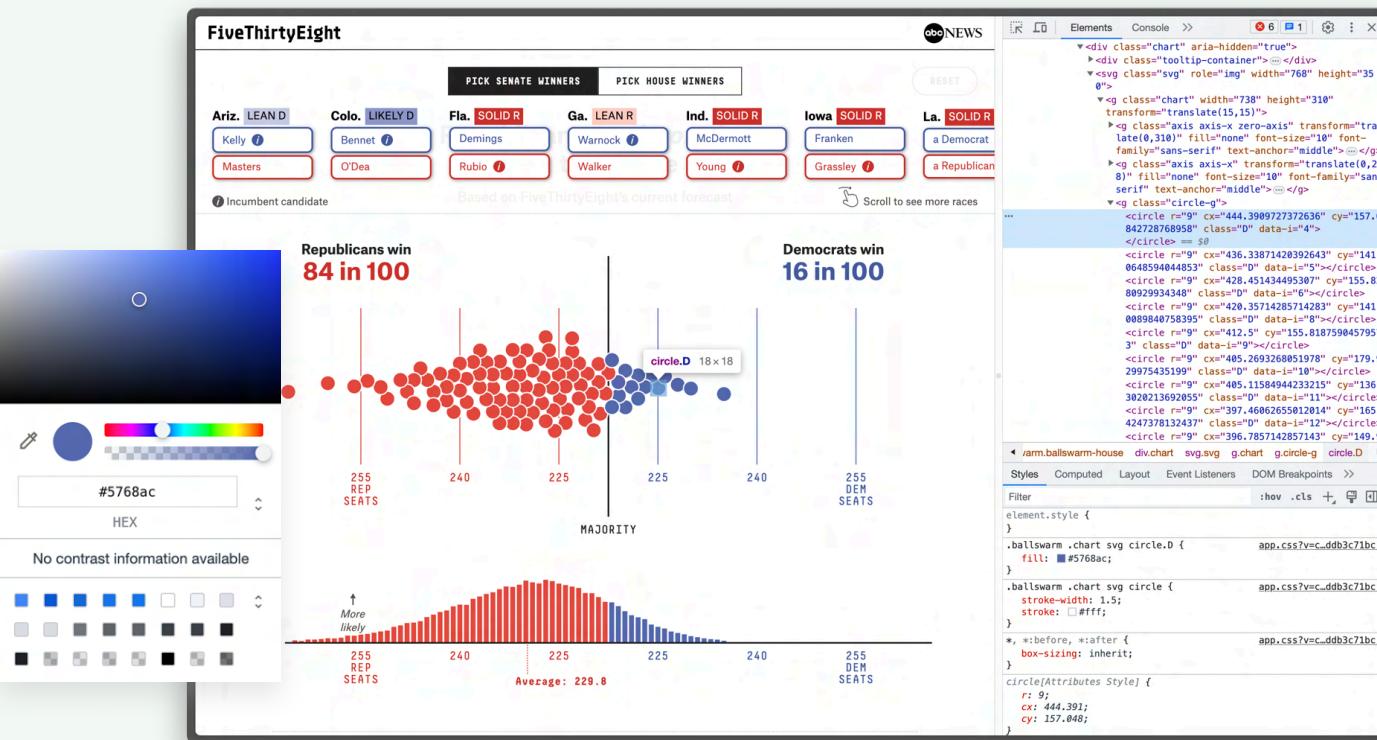
```
(module
  (type (;0;) (func (param i32 i32) (result i32)))
  (func (;0;) (type 0) (param i32 i32) (result i32)
    get_local 1
    get_local 0
    i32.add)
  (table (;0;) 1 1 anyfunc)
  (memory (;0;) 17)
  global (;0;) i32 (i32.const 1049118))
  global (;1;) i32 (i32.const 1049118))
  export "memory" (memory 0))
  export "__indirect_function_table" (table 0))
  export "__heap_base" (global 0))
  export "__data_end" (global 1))
  export "add" (func 0))
  data (i32.const 1049096) "invalid malloc request"))
```



How do we compile from a **visual form** to a **textual symbolic representation**?

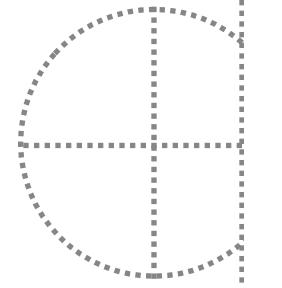


Map **interactions** with the visual form to **program edits**.



```
(module
  (type (;0;) (func (param i32 i32) (result i32)))
  (func (;0;) (type 0) (param i32 i32) (result i32)
    get_local 1
    get_local 0
    i32.add)
  (table (;0;) 1 1 anyfunc)
  (memory (;0;) 17)
  (global (;0;) i32 (i32.const 1049118))
  (global (;1;) i32 (i32.const 1049118))
  (export "memory" (memory 0))
  (export "__indirect_function_table" (table 0))
  (export "__heap_base" (global 0))
  (export "__data_end" (global 1))
  (export "add" (func 0))
  (data (i32.const 1049096) "invalid malloc request"))
```

WA

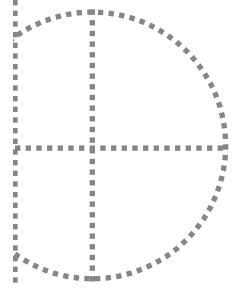


Lexer

Parser

Interm. Repr.

Code Generator

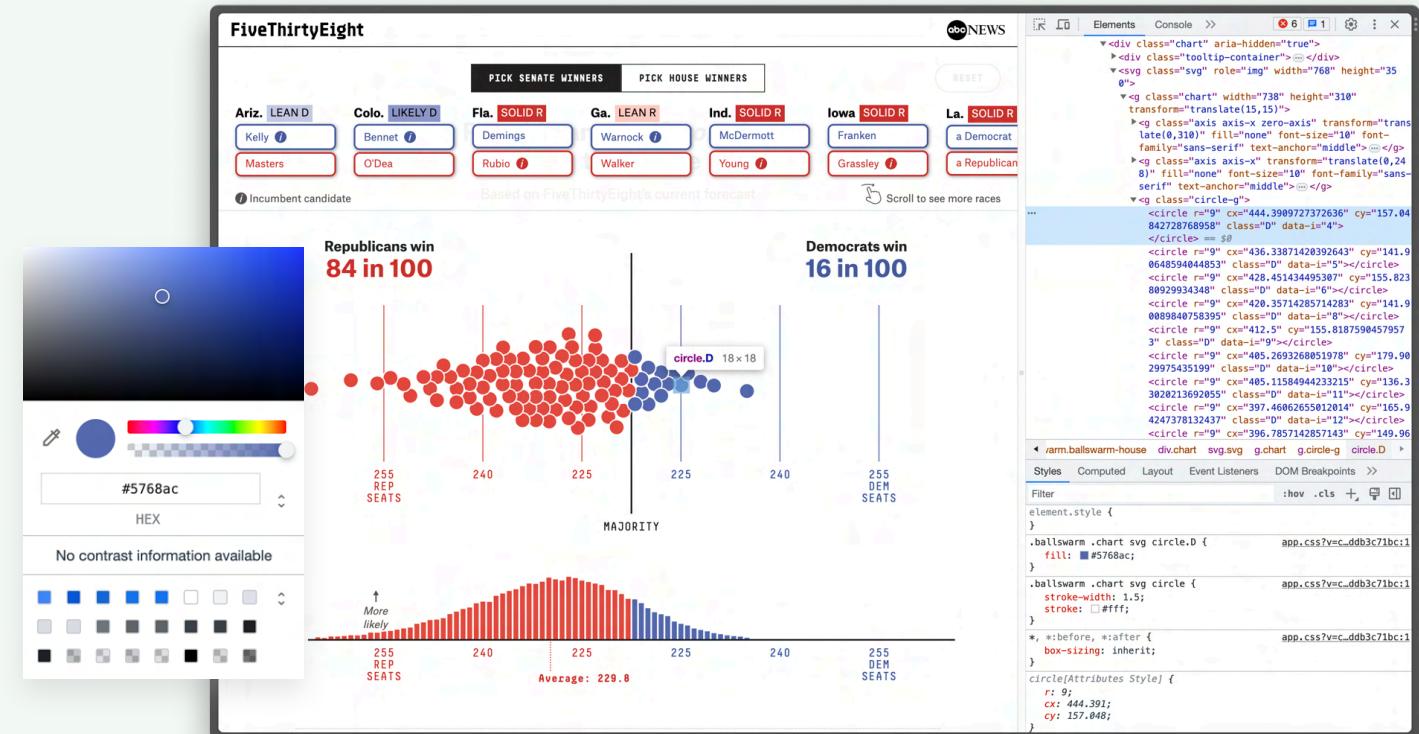


Lexer

Parser

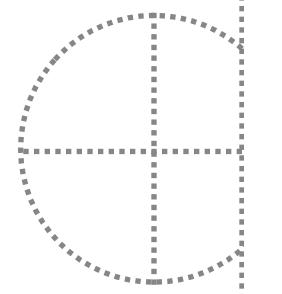
Recover (or infer) **semantic information** from the visual form that we can represent **symbolically**.

Map **interactions** with the visual form to **program edits**.



reviz
A tool for recovering semantic information from visual forms.

cartokit
A library for interacting with and editing visual forms.



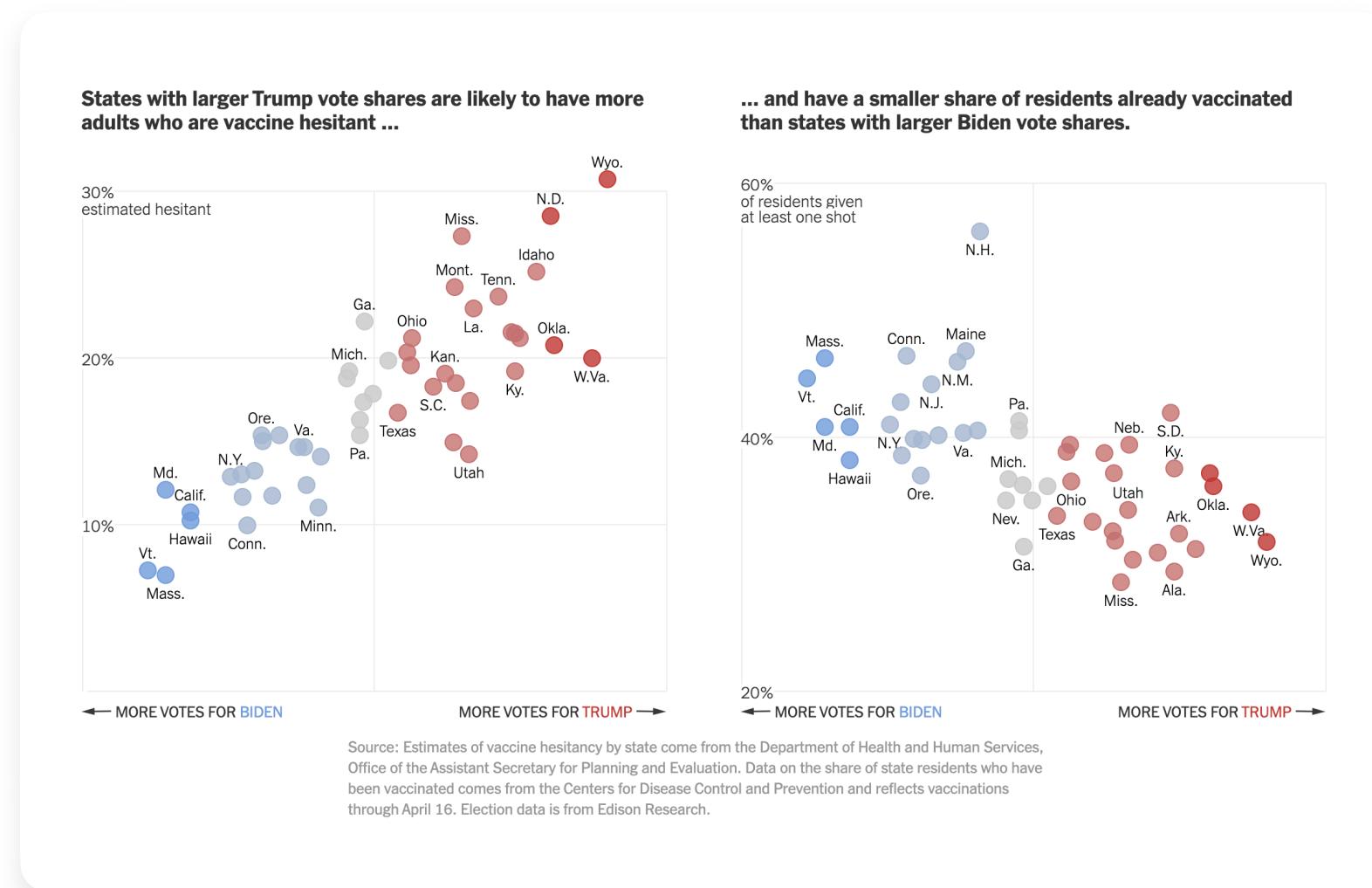


A compiler from **SVG subtrees** to **data visualizations**



A compiler from **SVG subtrees** to **data visualizations**
partial programs

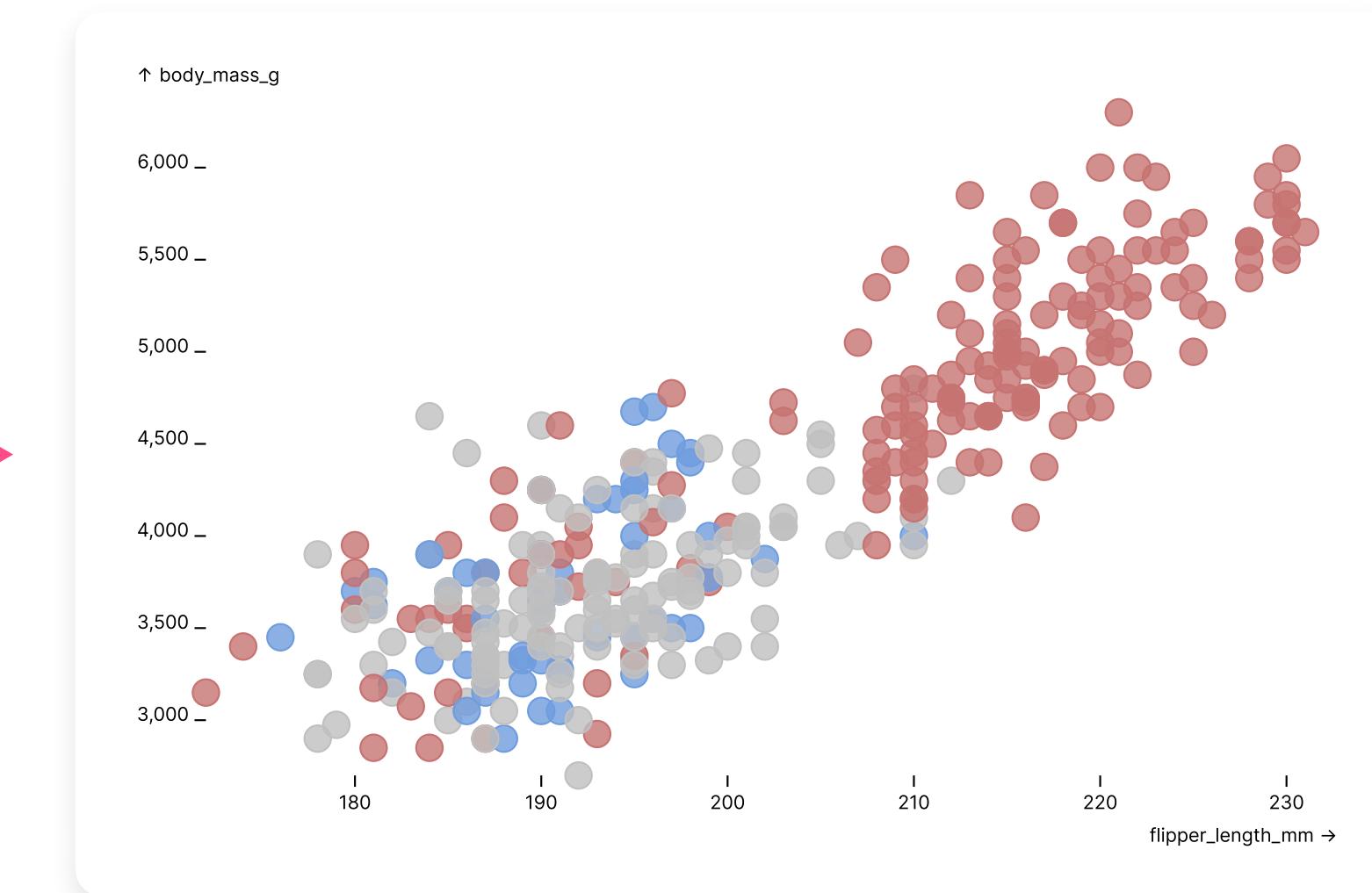
Demo



The New York Times

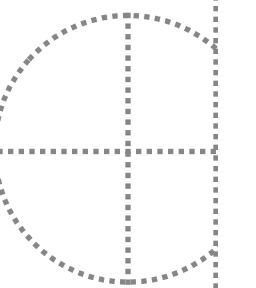
Least Vaccinated U.S. Counties Have Something in Common: Trump Voters

Danielle Ivory, Lauren Leatherby and Robert Gebeloff

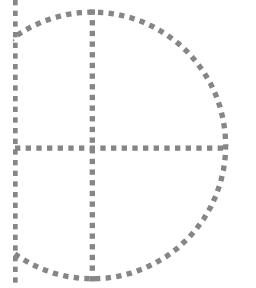


reviz

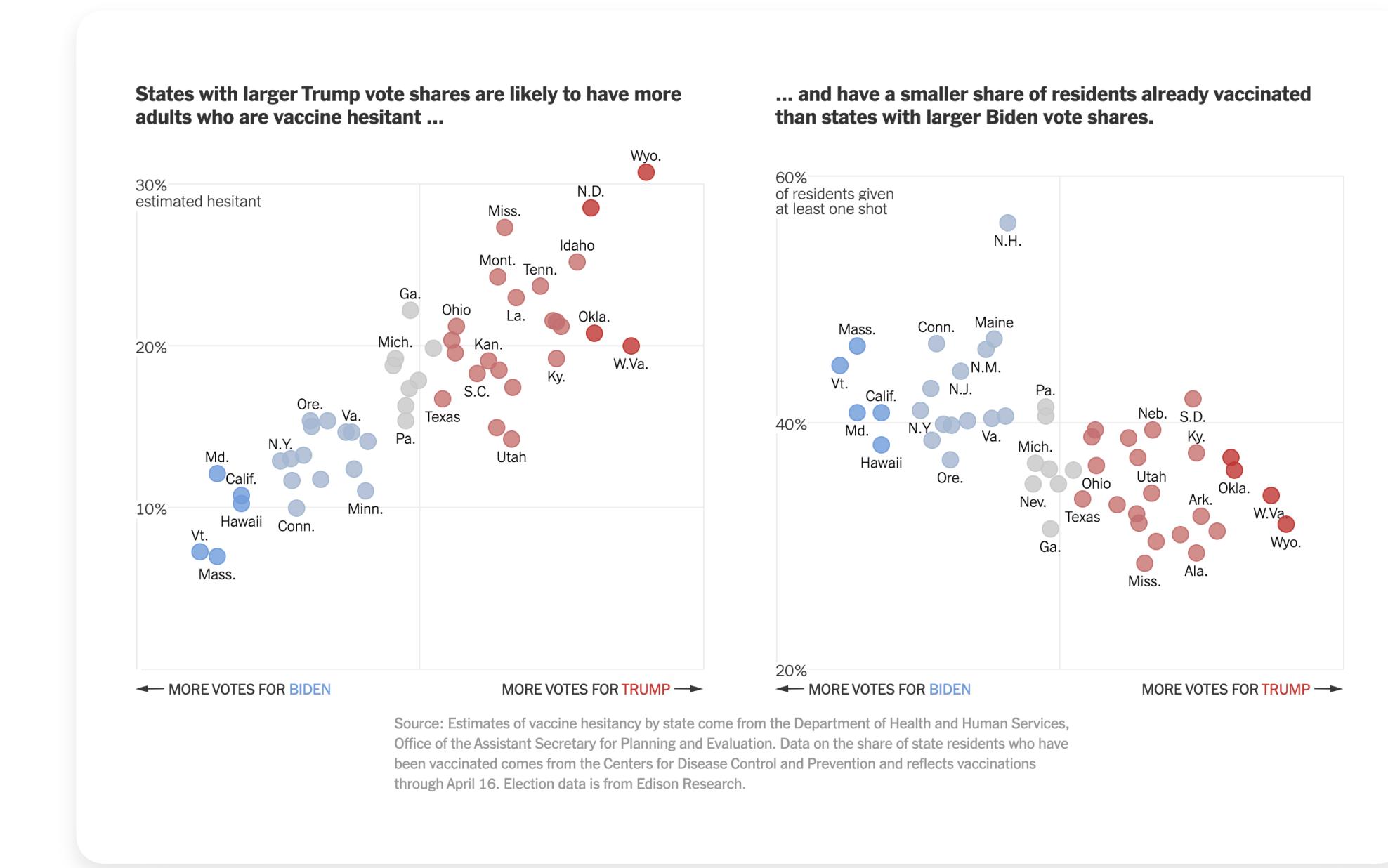
Comparing Flipper Length and Body Mass of the Palmer Penguins



Can we automatically retarget **existing visualizations** at **new datasets**?



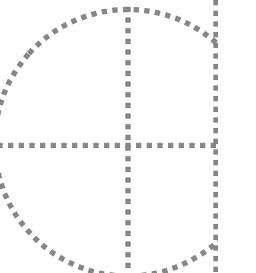
A user identifies a visualization they want to use as a visual input.



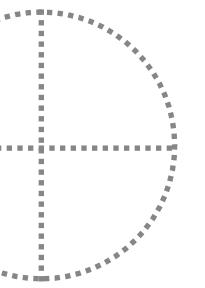
The New York Times

Least Vaccinated U.S. Counties Have Something in Common: Trump Voters

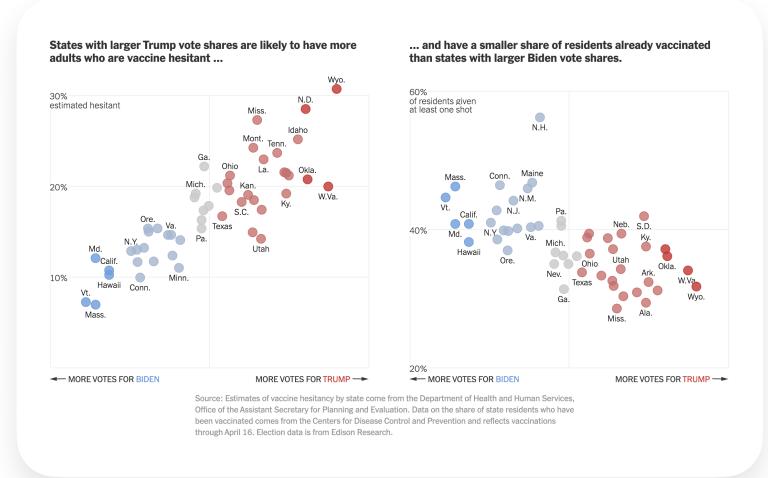
Danielle Ivory, Lauren Leatherby and Robert Gebeloff



reviz

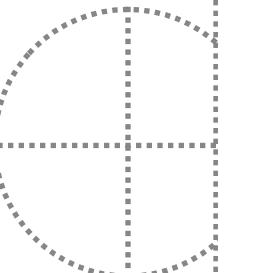


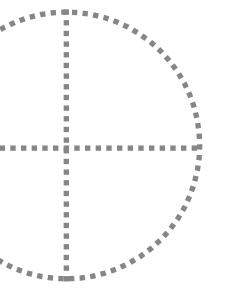
1.



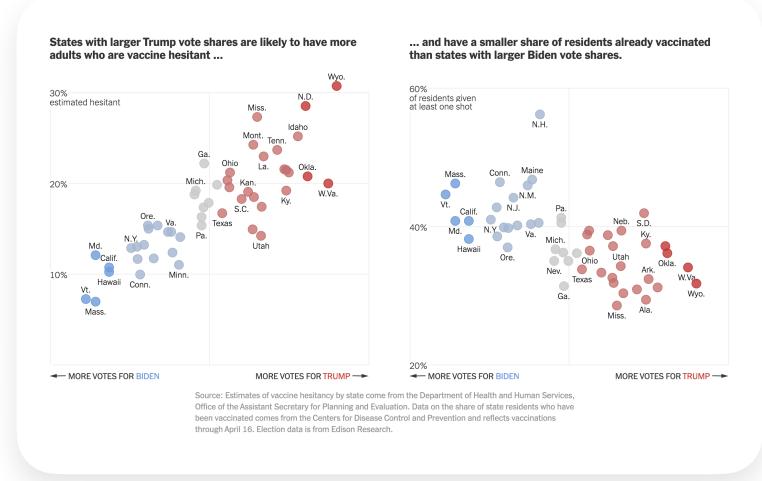
They pass its **svg**
subtree to

reviz





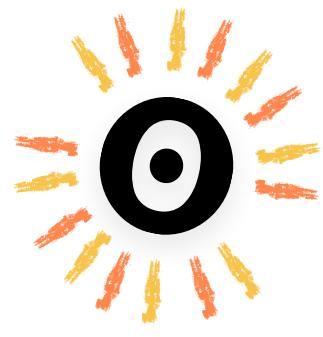
1.



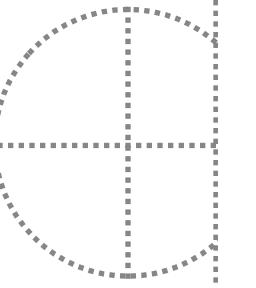
2.

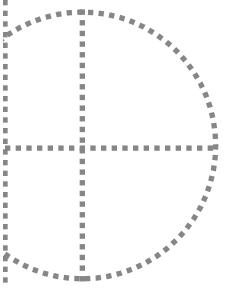


reviz compiles a **partial**
JavaScript program using
Observable's Plot library.

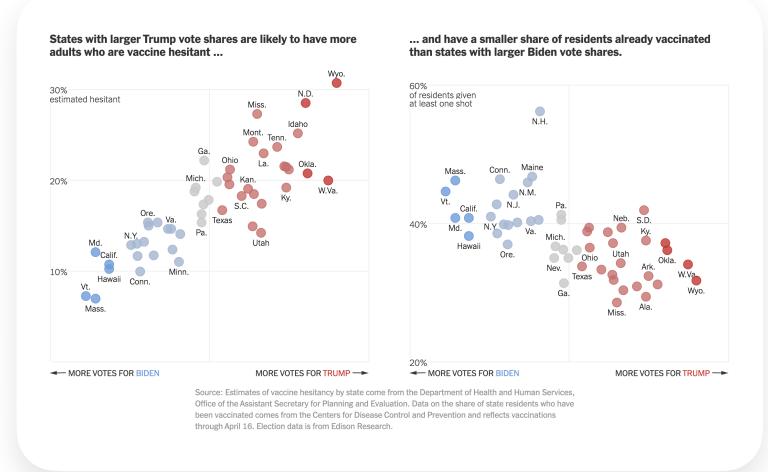


```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE", "#A7B9D3", "#C23734"]
  },
  marks: [
    Plot.dot(data, {
      x: "?",
      y: "?",
      fill: "?",
      fillOpacity: 1,
      stroke: "?",
      strokeOpacity: 1,
      strokeWidth: 1
    }),
  ],
});
```





1.



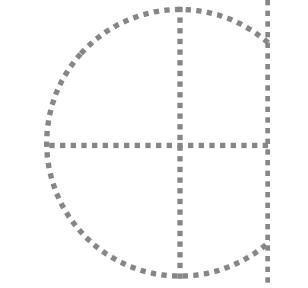
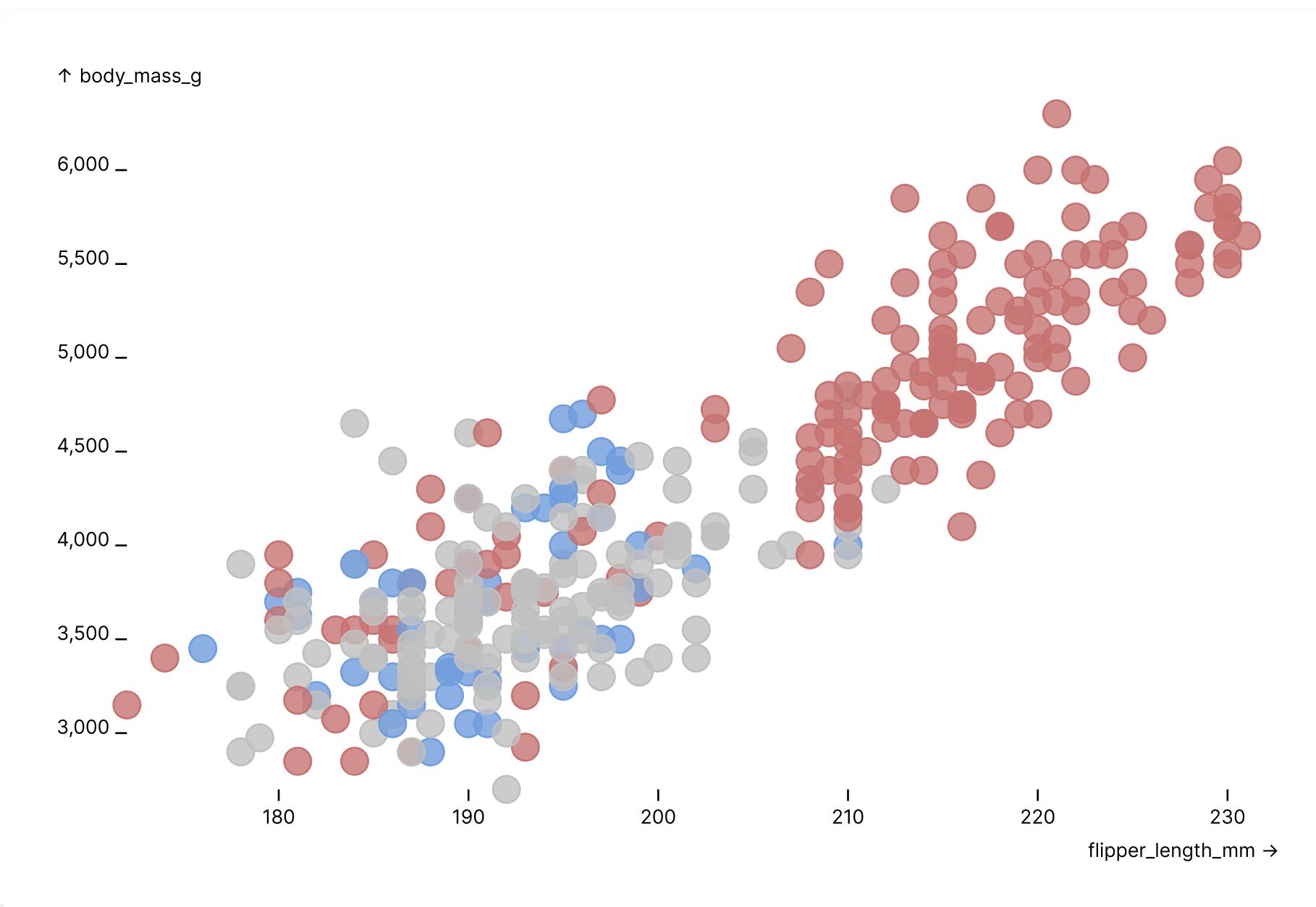
2.

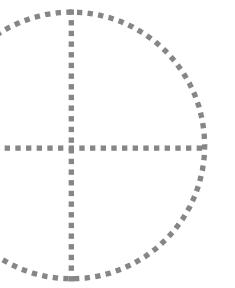


3.

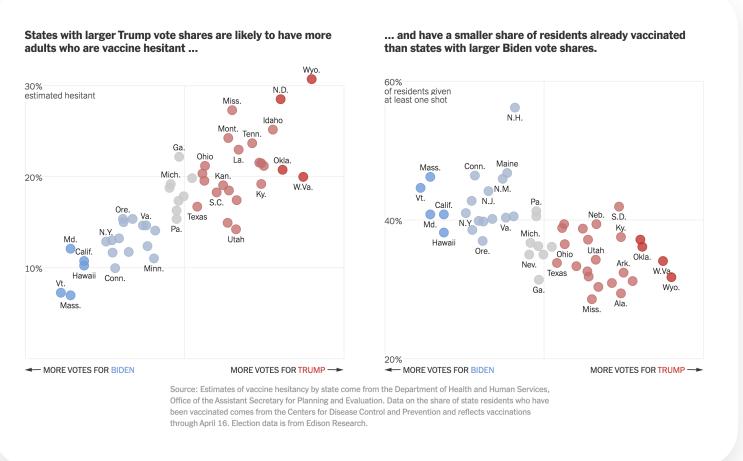
```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE", "#A7B9D3", "#C23734"]
  },
  marks: [
    Plot.dot(data, {
      x: "?",
      y: "?",
      fill: "?",
      fillOpacity: 1,
      stroke: "?",
      strokeOpacity: 1,
      strokeWidth: 1
    })
  ]
});
```

A user **fills in the holes ("?"")**
in the partial program to
produce a new visualization.





1.



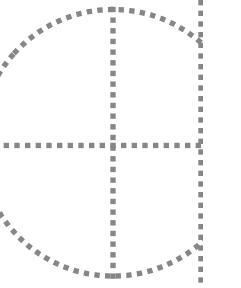
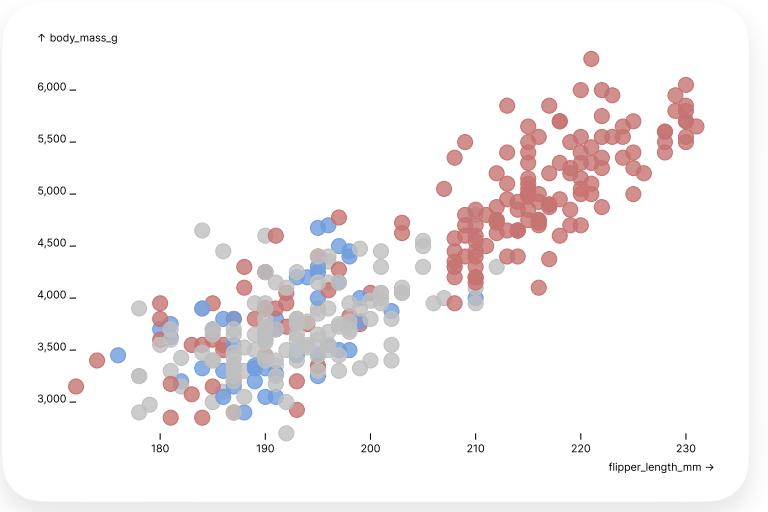
2.

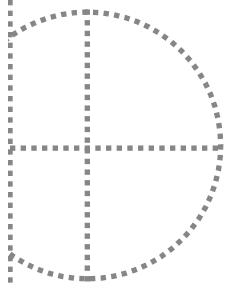


3.

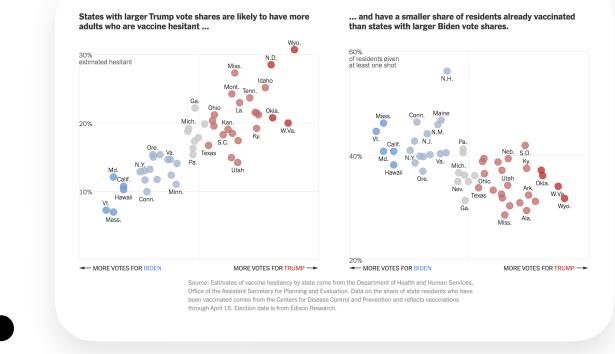
```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE", "#A7B9D3", "#C23734"]
  },
  marks: [
    Plot.dot(data, {
      x: "?",
      y: "?",
      fill: "?",
      fillOpacity: 1,
      stroke: "?",
      strokeOpacity: 1,
      strokeWidth: 1
    })
  ],
});
```

4.





1.



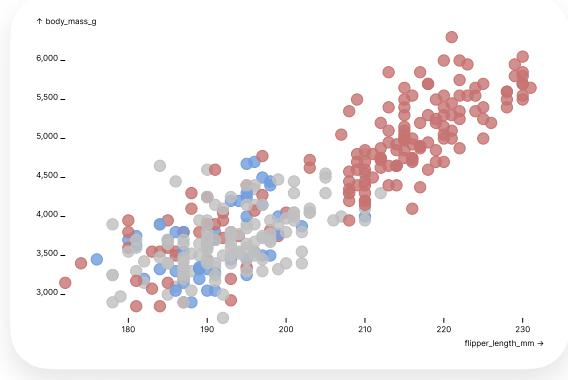
2.



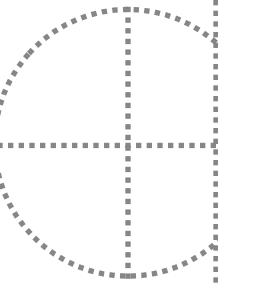
3.



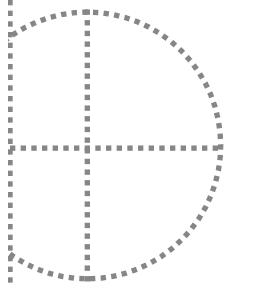
4.



What actually happens in this step?



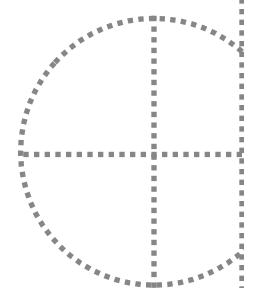
It all starts with a walk.



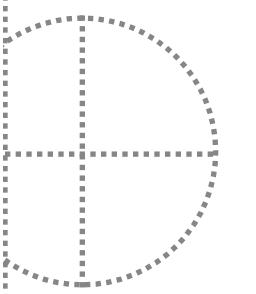
```
<svg viewBox="0 0 493.5 450" width="493.5" height="450">  
  <g aria-label="dot">  
    <circle cx="366" cy="349" r="7" fill="#C67371" fill-opacity=".8" stroke="#C67371" />  
    <circle cx="140" cy="167" r="7" fill="#A7B9D3" fill-opacity=".8" stroke="#A7B9D3" />  
    <circle cx="121" cy="119" r="7" fill="#709DDE" fill-opacity=".8" stroke="#709DDE" />  
    ...  
  </g>  
  <g aria-label="x-axis tick label" fill="none" stroke="currentColor">  
    <text y="0.71em" transform="translate(56, 420)" stroke="currentColor">10</text>  
    <text y="0.71em" transform="translate(56, 420)" stroke="currentColor">20</text>  
    ...  
</g>
```



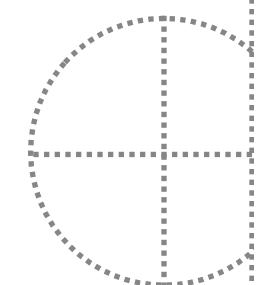
reviz Visit each node
in the `svg` subtree.



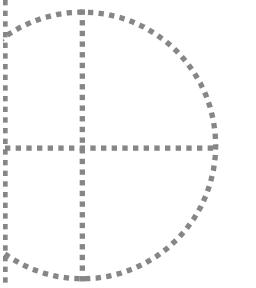
It all starts with a **walk**.



```
<svg viewBox="0 0 493.5 450" width="493.5" height="450">  
  <g aria-label="dot">  
    reviz Visit each node  
    in the svg subtree.  
    <circle cx="366" cy="349" r="7" fill="#C67371" fill-opacity=".8" stroke="#C67371" />  
    <circle cx="140" cy="167" r="7" fill="#A7B9D3" fill-opacity=".8" stroke="#A7B9D3" />  
    <circle cx="121" cy="119" r="7" fill="#709DDE" fill-opacity=".8" stroke="#709DDE" />  
    ...  
  <g aria-label="x-axis tick label" fill="none" stroke="currentColor">  
    <text y="0.71em" transform="translate(56, 420)" stroke="currentColor">10</text>  
    <text y="0.71em" transform="translate(56, 420)" stroke="currentColor">20</text>  
    ...  
</g></g>
```



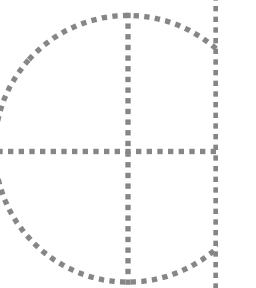
It all starts with a **walk**.



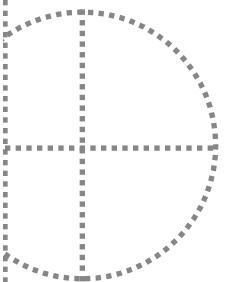
```
<svg viewBox="0 0 493.5 450" width="493.5" height="450">  
  <g aria-label="dot">  
    <circle cx="366" cy="349" r="7" fill="#C67371" fill-opacity=".8" stroke="#C67371" />  
    <circle cx="140" cy="167" r="7" fill="#A7B9D3" fill-opacity=".8" stroke="#A7B9D3" />  
    <circle cx="121" cy="119" r="7" fill="#709DDE" fill-opacity=".8" stroke="#709DDE" />  
    ...  
  </g>  
  <g aria-label="x-axis tick label" fill="none" stroke="currentColor">  
    <text y="0.71em" transform="translate(56, 420)" stroke="currentColor">10</text>  
    <text y="0.71em" transform="translate(56, 420)" stroke="currentColor">20</text>  
    ...  
  </g>  
</svg>
```



Visit each node
in the `svg` subtree.



Read **geometric** and **presentational** attributes off of each DOM node and its computed styles.



```
<circle cx="366" cy="349" r="7" fill="#C67371" fill-opacity=".8" stroke="#C67371" />
```

cx → 366

fill → "#C67371"

cy → 349

fill-opacity → 0.8

r → 7

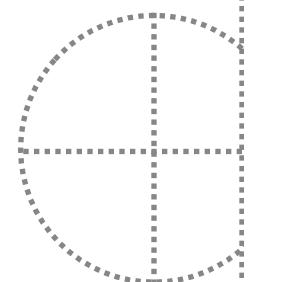
stroke → "#C67371"

Styles Computed

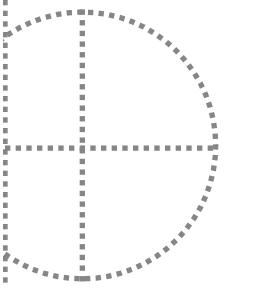
- stroke-opacity 1
- stroke-width 2px

stroke-opacity → 1

stroke-width → 2px



Store collected attributes in **attribute sets**.



cx: { cx → 366 , cx → 398 , cx → 422 , cx → 543 , ... }

cy: { cy → 349 , cy → 218 , cy → 265 , cy → 121 , ... }

r: { r → 7 }

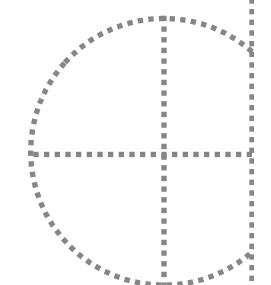
fill: { fill → "#C67371" , fill → "#A7B9D3" , fill → "#709DDE" , ... }

fill-opacity: { fill-opacity → 0.8 }

stroke: { stroke → "#C67371" , stroke → "#A7B9D3" , stroke → "#709DDE" , ... }

stroke-opacity: { stroke-opacity → 1 }

stroke-width: { stroke-width → 2px }





Apply a set of **predicate functions** associated with a **visualization type**.

`cx: { cx → 366 , ... }`

Bar Chart

`cy: { cy → 349 , ... }`

`hasMarkType("rect")`

`r: { r → 7 }`

`hasConsistentGeomAttr("width")`

`fill: { fill → "#C67371" , ... }`

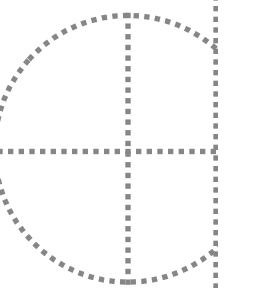
`hasXScaleType("discrete")`

`fill-opacity: { fill-opacity → 0.8 }`

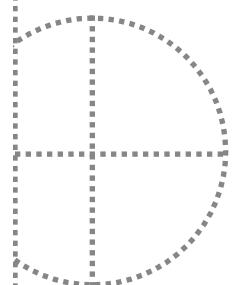
`stroke: { stroke → "#C67371" , ... }`

`stroke-opacity: { stroke-opacity → 1 }`

`stroke-width: { stroke-width → 2px }`



Compute the **ratio of predicates** returning true for each **visualization type**.



Bar Chart (0/3 predicates)

hasMarkType("rect")

⋮ ⋮ ⋮

hasConsistentGeomAttr("width")

⋮ ⋮ ⋮

hasXScaleType("width")

⋮ ⋮ ⋮

Strip Plot (2/3 predicates)

hasMarkType("circle")

⋮ ⋮ ⋮

hasConsistentGeomAttr("r")

⋮ ⋮ ⋮

hasSiblingsWithConsistentCyAttr

⋮ ⋮ ⋮

Scatterplot (2/2 predicates)

hasMarkType("circle")

⋮ ⋮ ⋮

hasConsistentGeomAttr("r")

⋮ ⋮ ⋮

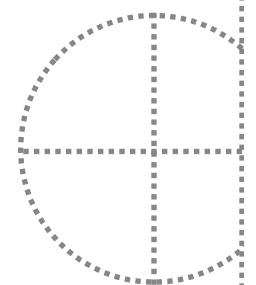
Bubble Chart (1/2 predicates)

hasMarkType("circle")

⋮ ⋮ ⋮

hasDivergentGeomAttr("r")

⋮ ⋮ ⋮





Merge the inferred **visualization type** with the **attribute sets** to produce the **intermediate representation (IR)**.

Scatterplot (2/2 predicates)

```
r: { r → 7 }

fill: { fill → "#C67371" ,... }

fill-opacity: { fill-opacity → 0.8 }

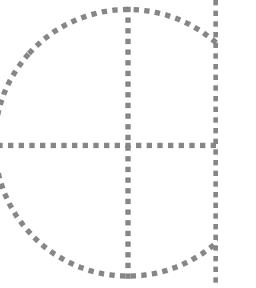
stroke: { stroke → "#C67371" ,... }

stroke-opacity: { stroke-opacity → 1 }

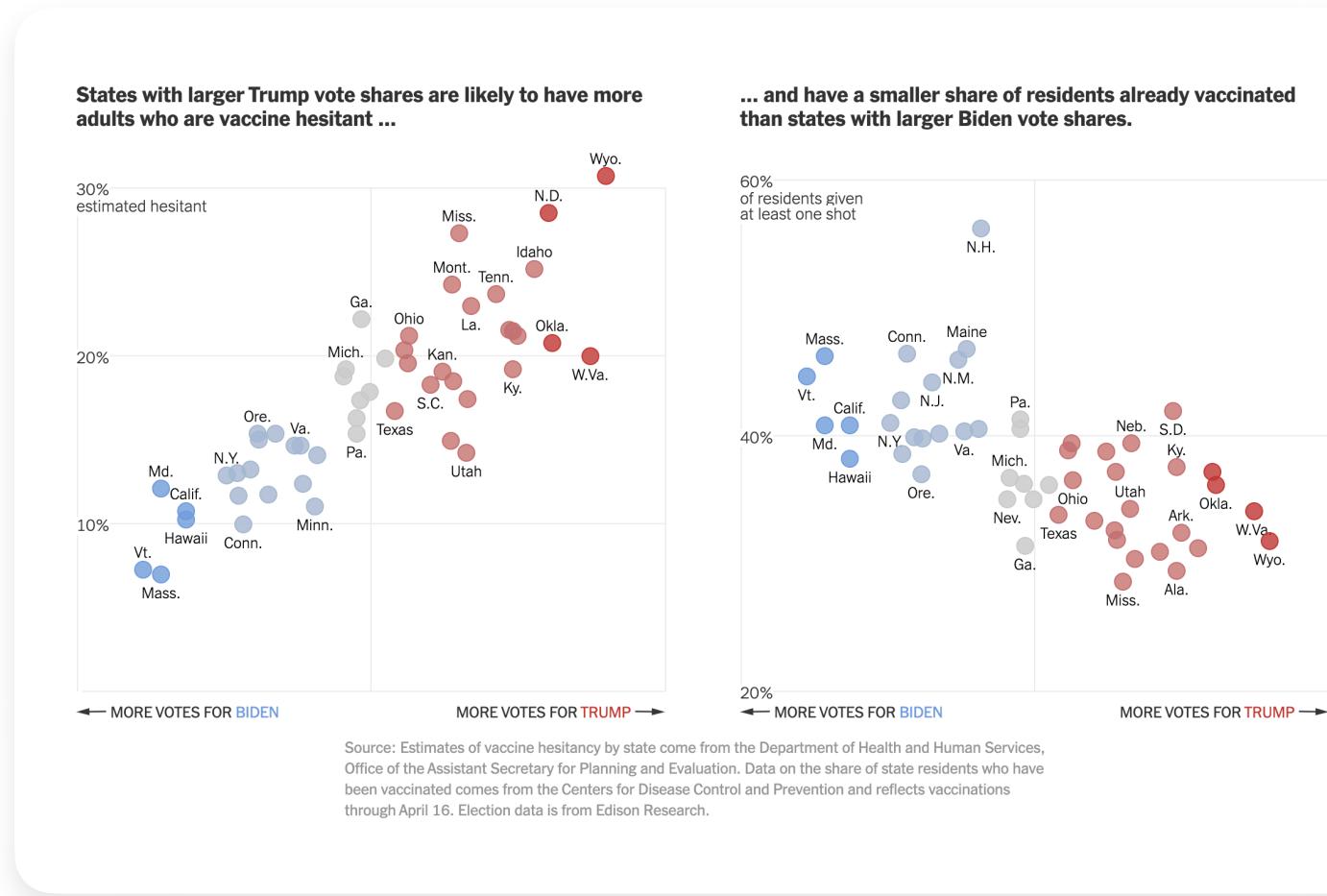
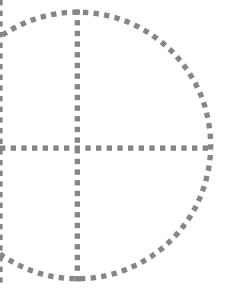
stroke-width: { stroke-width → 2px }
```



```
{
  type: "Scatterplot",
  r: [7],
  fill: ["#C67371", "#ccc", "#709DDE"],
  fill-opacity: [0.8],
  stroke: ["#C67371", "#ccc", "#709DDE"],
  stroke-opacity: [1],
  stroke-width: ["2px"]
}
```

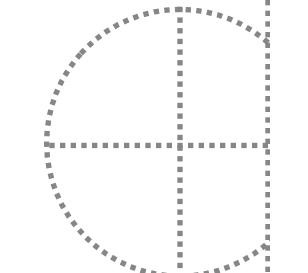


The *frontend* of the compiler



Input Visualization

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```



Intermediate Representation

The *backend* of the compiler

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```



```
const plot = Plot.plot({  
  color: {  
    type: "categorical",  
    range: ["#C67371", "#ccc", "#709DDE"]  
  },  
  marks: [  
    Plot.dot(data, {  
      x: "??",  
      y: "??",  
      fill: "??",  
      fillOpacity: 1,  
      stroke: "??",  
      strokeOpacity: 1,  
      strokeWidth: 1  
    }),  
  ],  
});
```

Intermediate Representation

Partial Program

Start with the IR and an empty program,
signified by an **evaluation hole**.

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```

Intermediate Representation

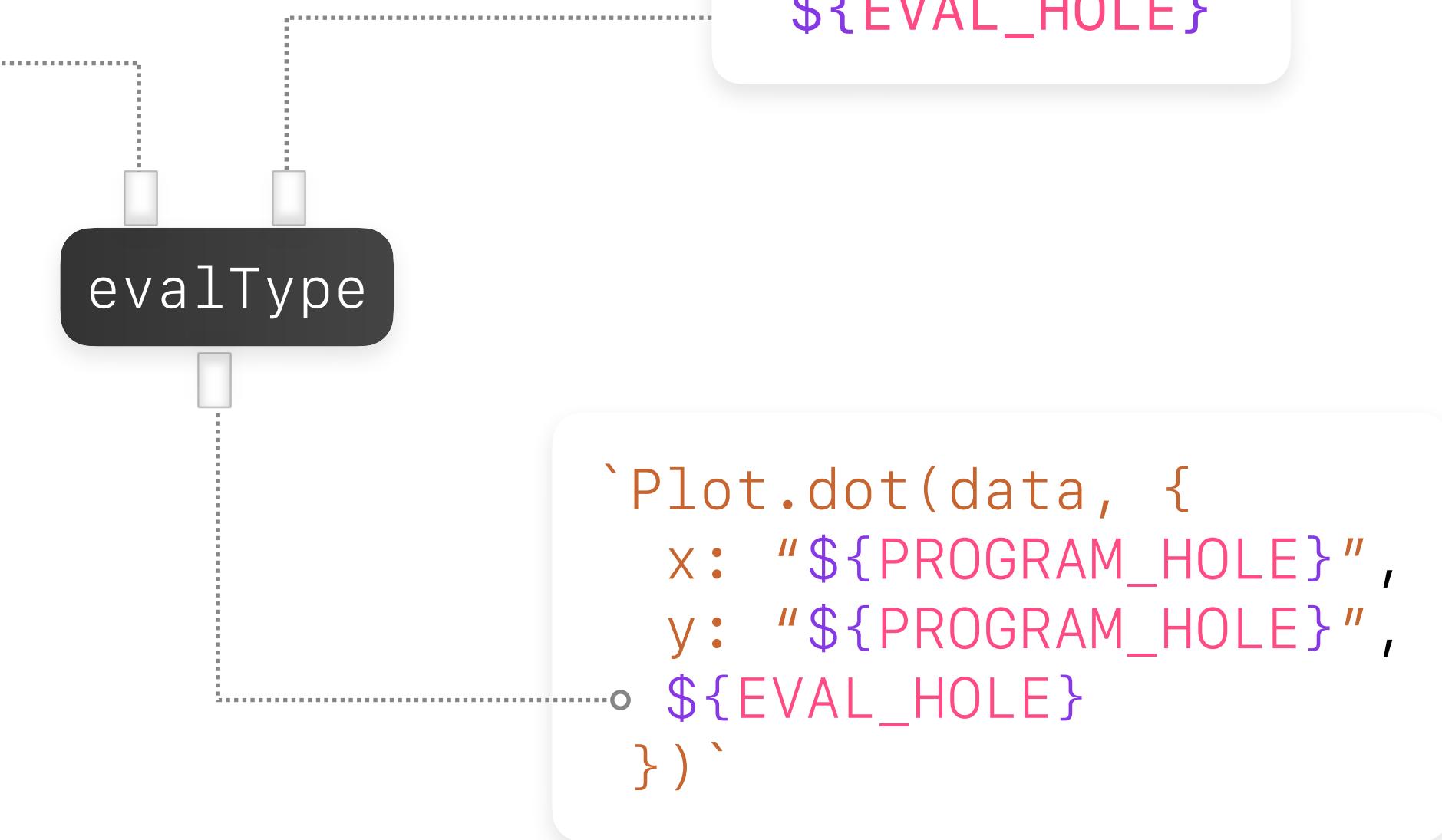
Partial Program

`\${EVAL_HOLE}`

Insertion point for
generated code

Apply a set of **rewrite rules** to transform **key-value pairs** in the IR into **program fragments**.

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```

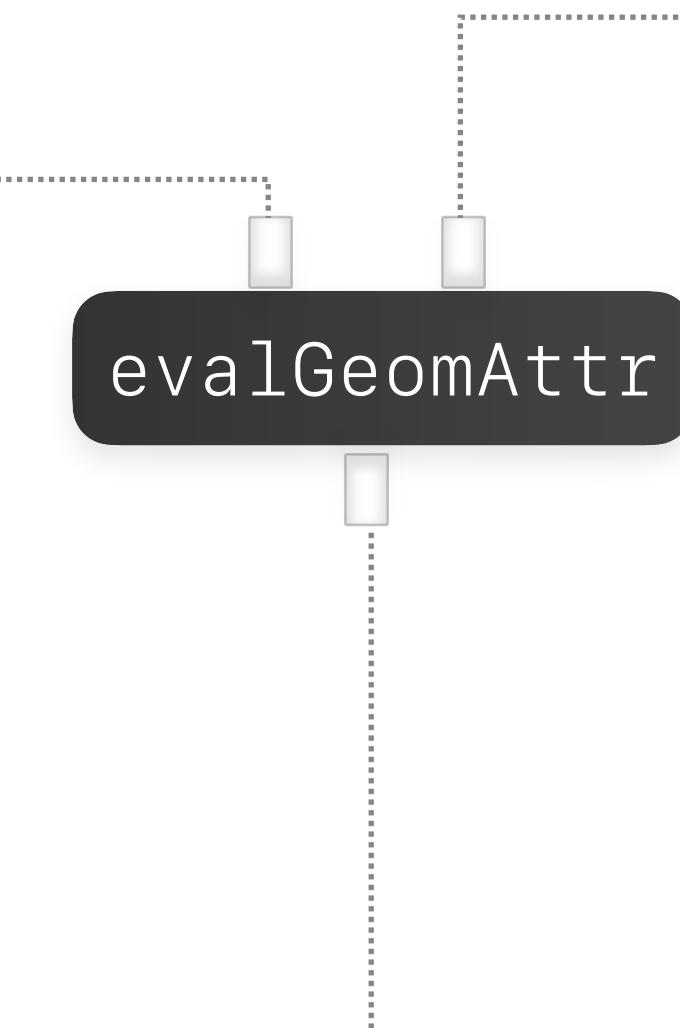


Intermediate Representation

Partial Program

Apply a set of **rewrite rules** to transform **key-value pairs** in the IR into **program fragments**.

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```



```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  ${EVAL_HOLE}  
} )`
```

```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  r: 7,  
  ${EVAL_HOLE}  
} )`
```

Intermediate Representation

Partial Program

Continue applying rewrites until we've read all key-value pairs in the IR.

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```

evalPresAttr

```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  r: 7,  
  ${EVAL_HOLE}  
})`
```

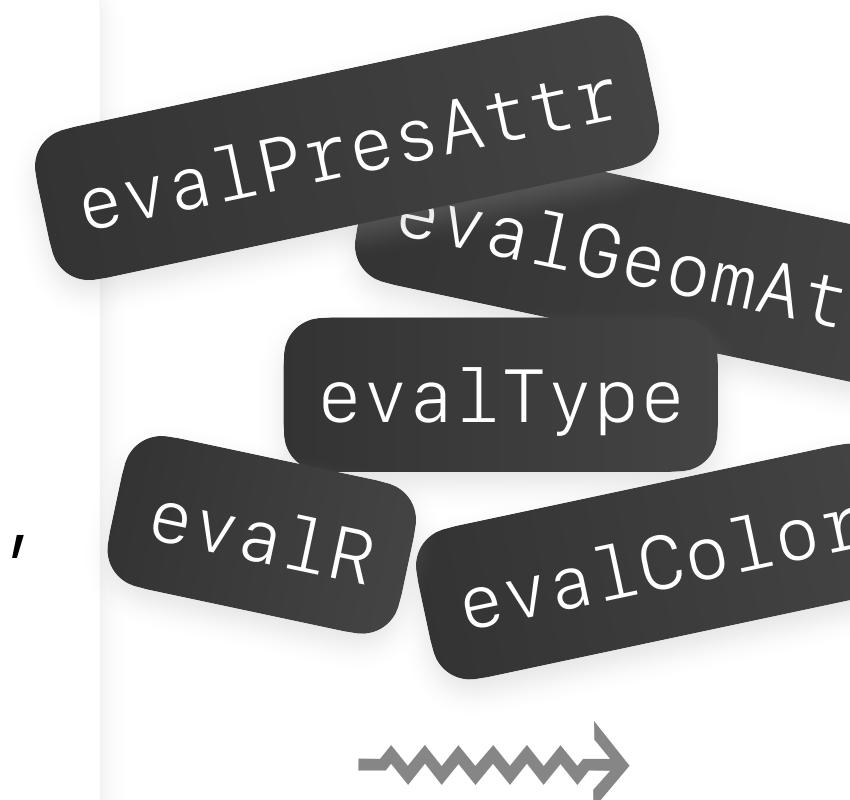
```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  r: 7,  
  fill: "${PROGRAM_HOLE}",  
  ${EVAL_HOLE}  
})`
```

Intermediate Representation

Partial Program

Continue applying rewrites until we've read all key-value pairs in the IR.

```
{  
  type: "Scatterplot",  
  r: [7],  
  fill: ["#C67371", "#ccc", "#709DDE"],  
  fill-opacity: [0.8],  
  stroke: ["#C67371", "#ccc", "#709DDE"],  
  stroke-opacity: [1],  
  stroke-width: ["2px"]  
}
```

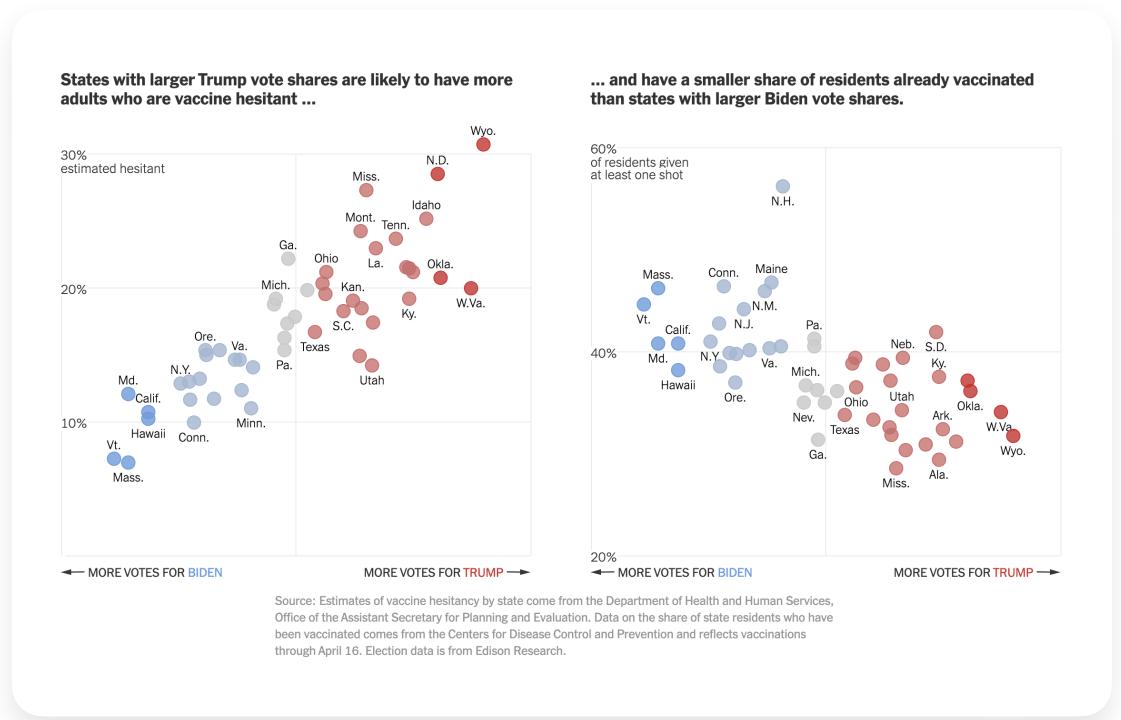


```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  r: 7,  
  fill: "${PROGRAM_HOLE}",  
  fillOpacity: 0.8,  
  stroke: "${PROGRAM_HOLE}",  
  strokeOpacity: 1,  
  strokeWidth: "2px"  
})`
```

Intermediate Representation

Partial Program

Input Visualization



Attribute Sets

```
r → 7
cx → 366
fill → "#C67371"
fill-opacity → 0.8
```

Intermediate Representation

```
{
  type: "Scatterplot",
  r: [7],
  fill: ["#C67371", "#ccc", "#709DDE"],
  fill-opacity: [0.8],
  stroke: ["#C67371", "#ccc", "#709DDE"],
  stroke-opacity: [1],
  stroke-width: ["2px"]
}
```

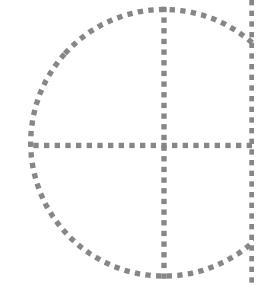
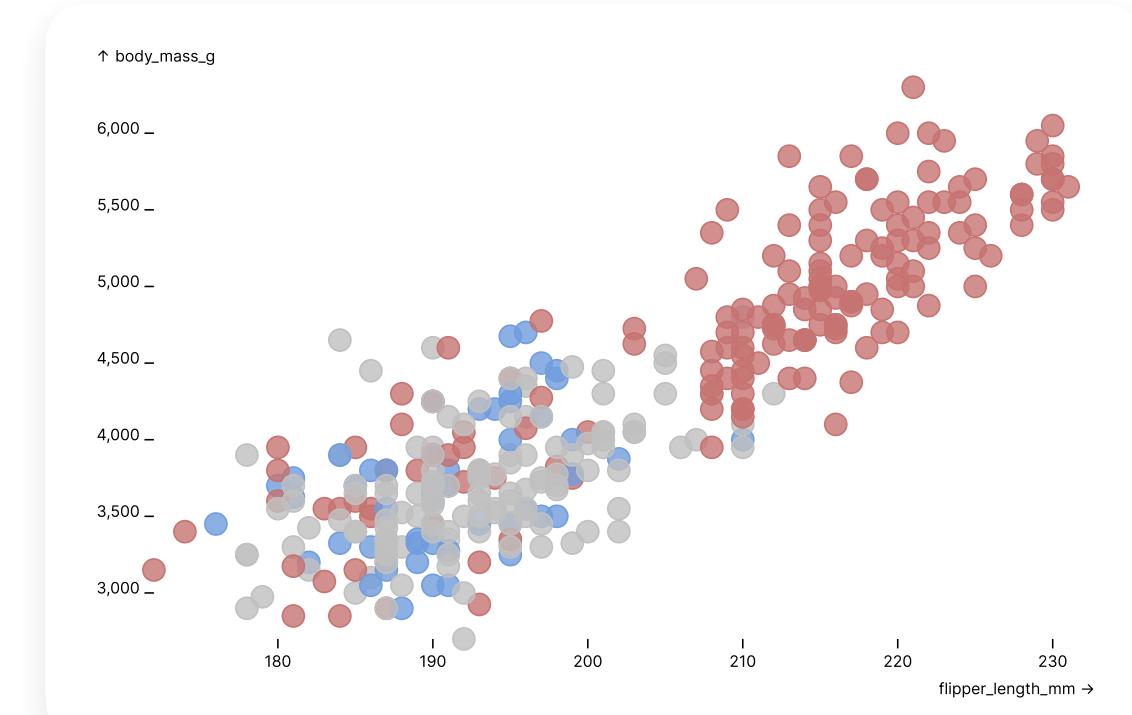
Rewrite Rules and Codegen

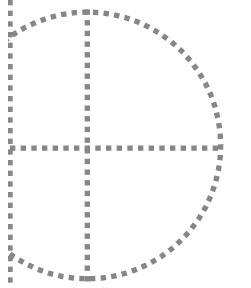
evalPresAttr
evalGeomAttr
evalType
evalR
evalColor

Partial Program

```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE"]
  },
  marks: [
    Plot.dot(data, {
      x: "?",
      y: "?",
      fill: "?",
      fillOpacity: 1,
      stroke: "?",
      strokeOpacity: 1,
      strokeWidth: 1
    })
  ],
});
```

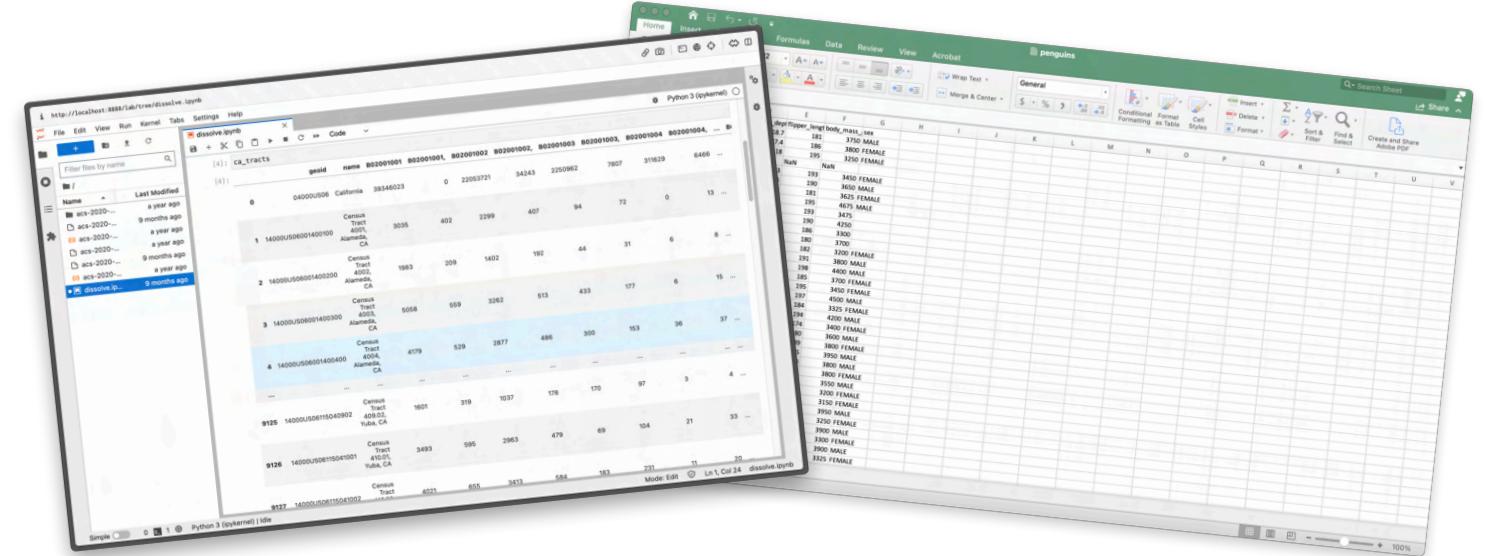
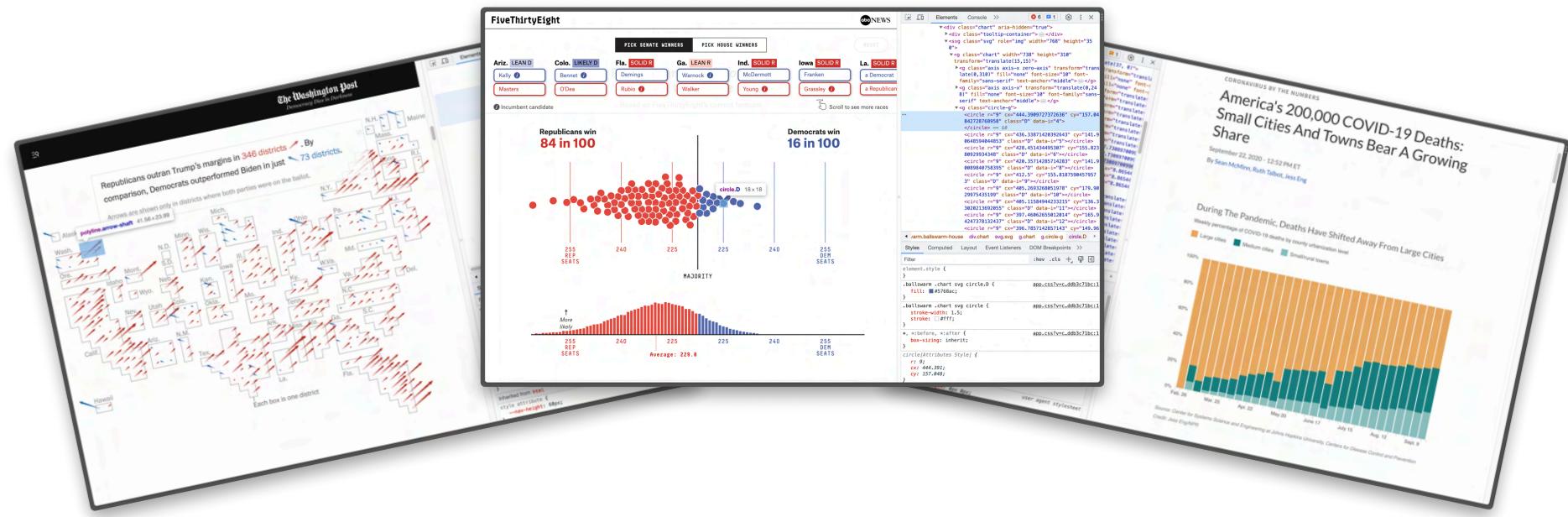
Output Visualization



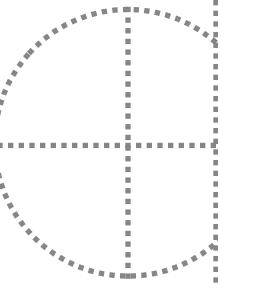


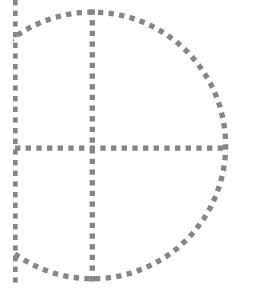
Working Practice

Lift **visual styles** and **graphical forms** from **examples**



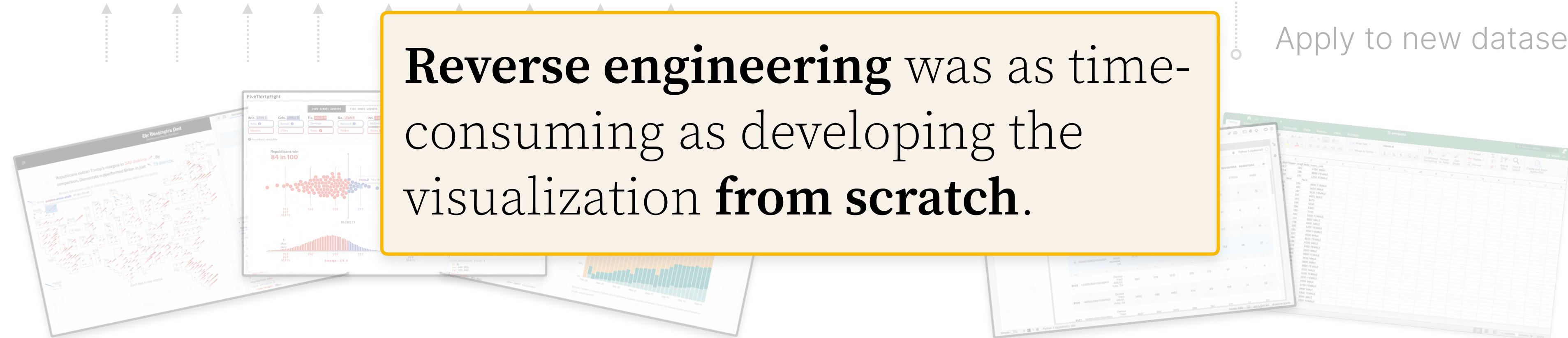
Apply to new datasets



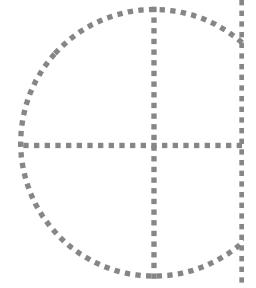


Working Practice

Lift visual styles and graphical forms from examples



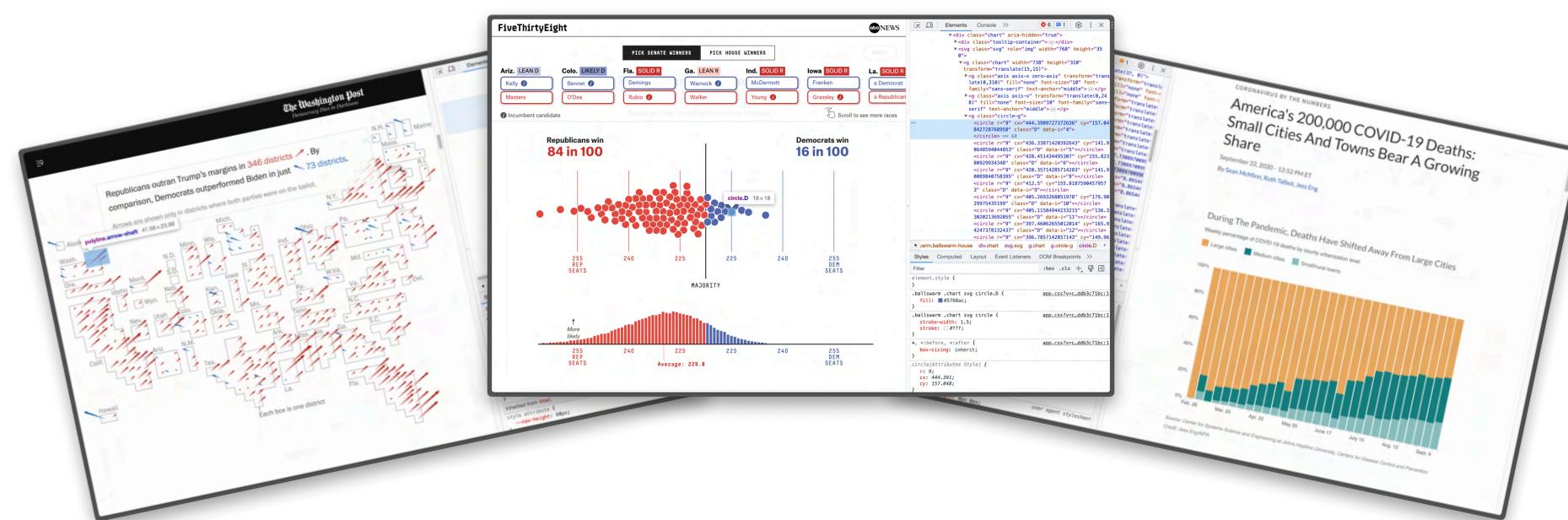
Apply to new datasets





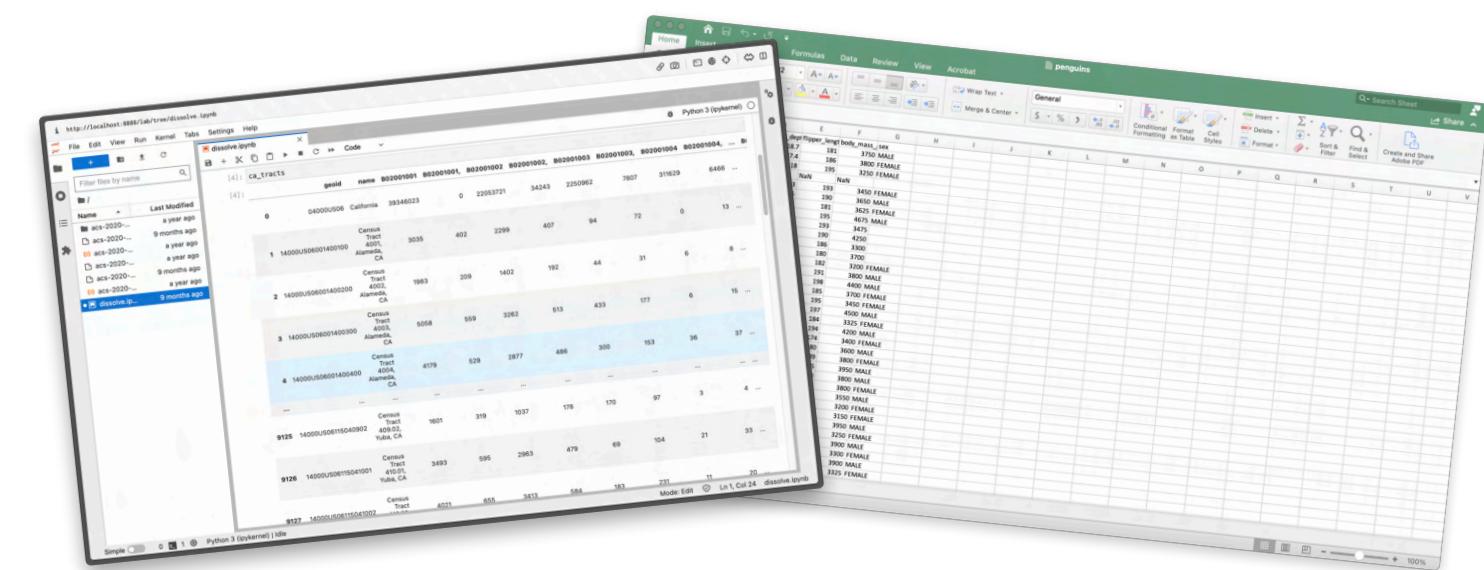
Working Practice

Lift visual styles and graphical forms from examples



Hours

Apply to new datasets



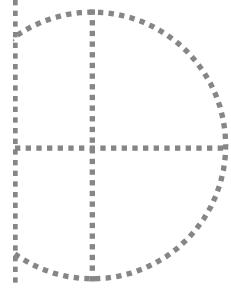
Automatically compile partial programs from examples



Milliseconds

```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE"]
  },
  marks: [
    Plot.dot(data, {
      x: "?",
      y: "?",
      fill: "?",
      fillOpacity: 1,
      stroke: "?",
      strokeOpacity: 1,
      strokeWidth: 1
    })
  ],
});
```

Feed in any dataset

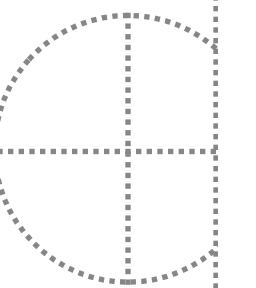


“

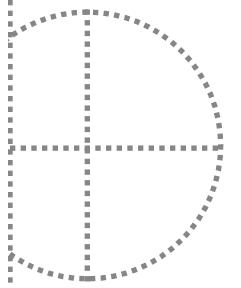
Don't commit to a specific visual form before seeing your data in it. A given visual form — say the pie chart or treemap — isn't “good” or “bad” in an absolute sense, but it may or may not be appropriate to your data and the specific question you want answered.

The only way to know whether a form is effective is if it communicates: you must put your data in it and see.

”



Mike Bostock • “10 Years of Open-Source Visualization”

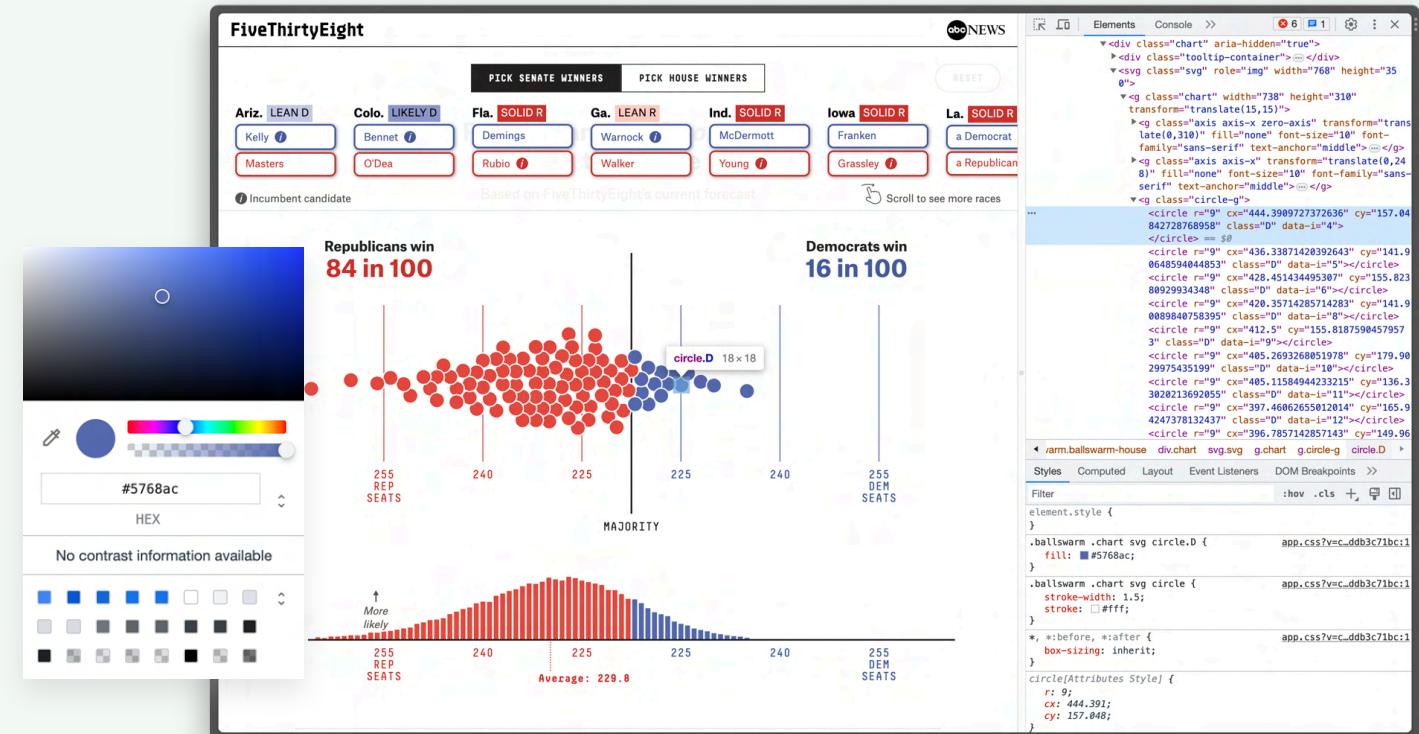


Lexer

Parser

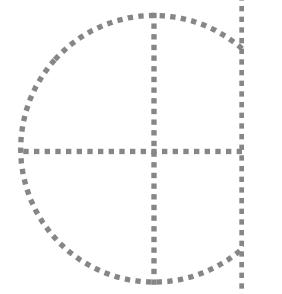
Recover (or infer) **semantic information** from the visual form that we can represent **symbolically**.

Map **interactions** with the visual form to **program edits**.



reviz
└─┐ ┌─┘ └─┐ ┌─┘
└─┐ ┌─┘ └─┐ ┌─┘
└─┐ ┌─┘ └─┐ ┌─┘

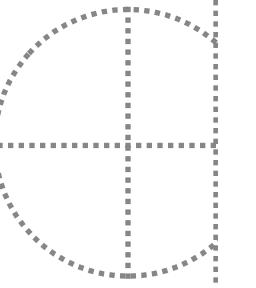
cartokit





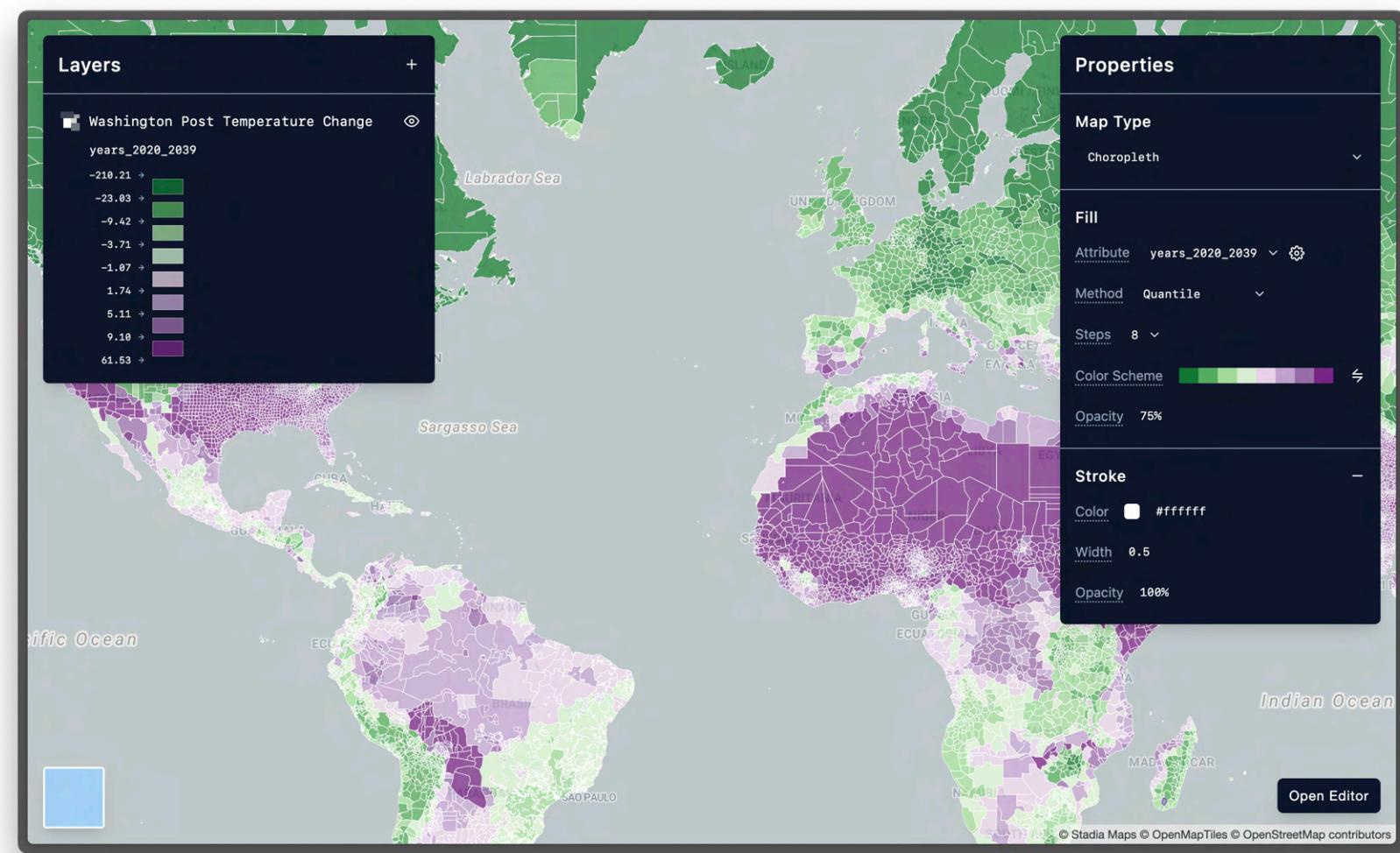
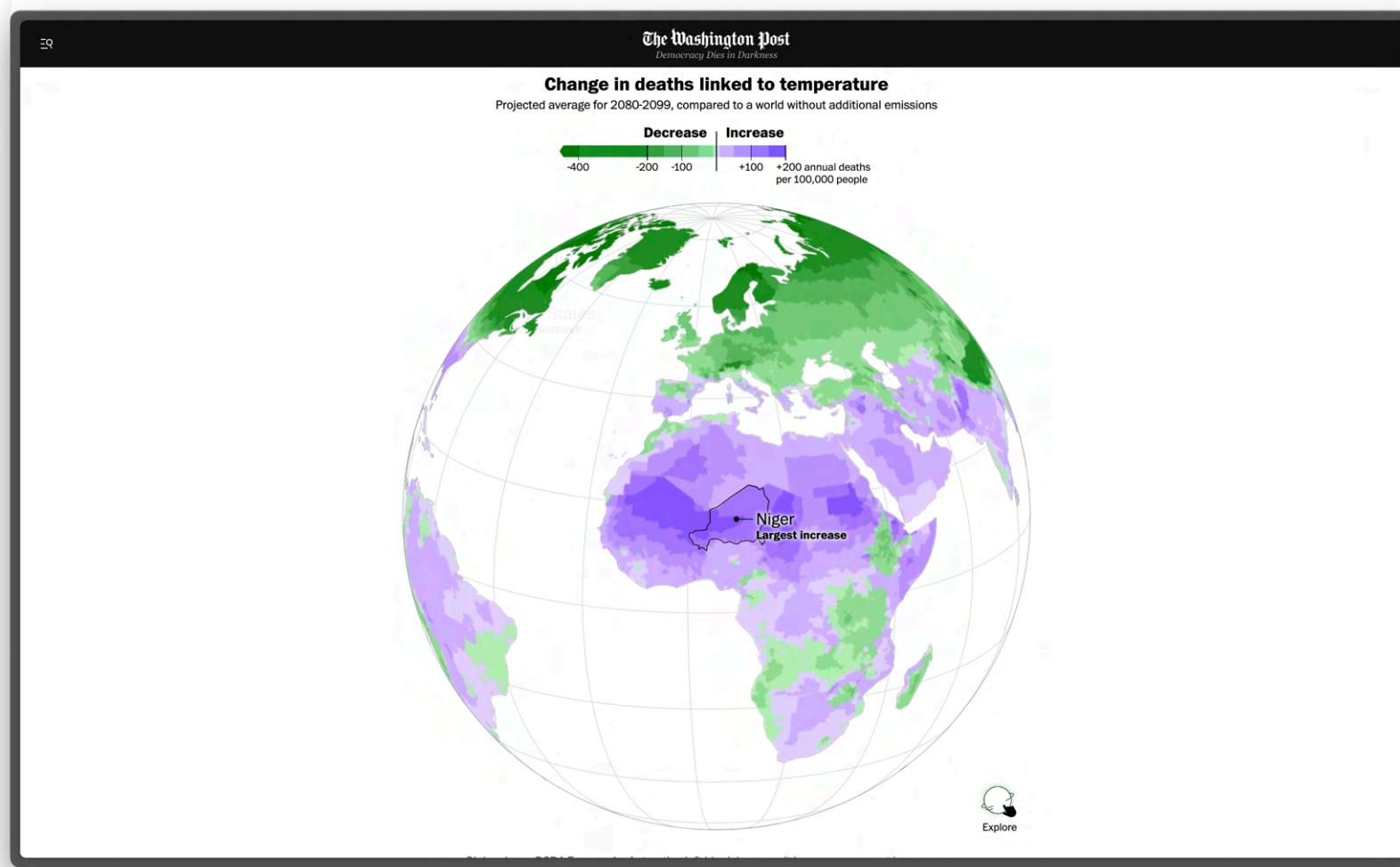
cartokit

A **direct manipulation** programming
environment for **interactive cartography**



cartokit

Demo



The Washington Post

Will global warming make temperature less deadly?

Harry Stevens

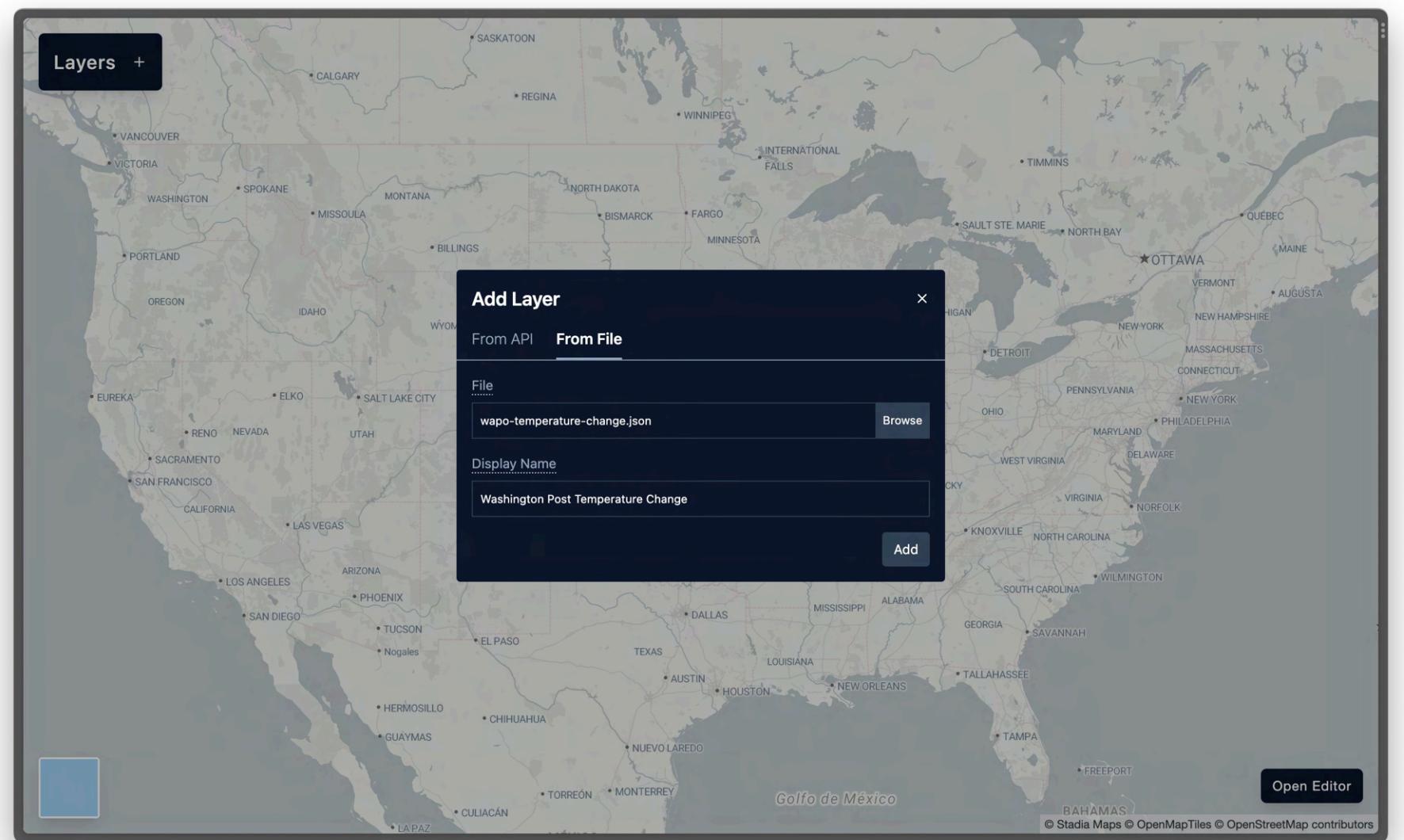
cartokit

A reproduction of the map in cartokit.

Can we author programs through **direct manipulation** of their outputs?

cartokit

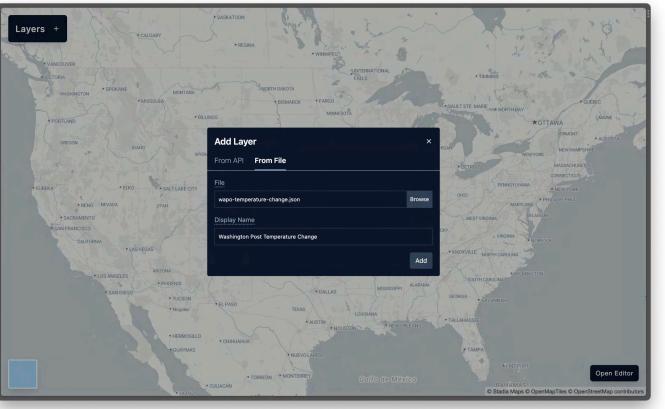
A user uploads their geospatial data to **cartokit**.



cartokit

cartokit renders and displays the data while simultaneously generating a program.

1.

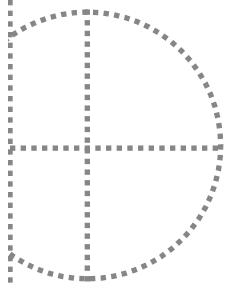


The screenshot shows a more detailed view of the cartokit interface. On the left is a map of North America with state and country boundaries. On the right, there are two panels: a "Properties" panel and a "Code" panel. The "Properties" panel is open for the "Washington Post Temperature Change" layer, showing settings for "Map Type" (Fill), "Fill" (Color: #8638e5, Opacity: 75%), and "Stroke" (Color: #8638e5, Width: 1, Opacity: 100%). The "Code" panel contains a snippet of Mapbox GL JS code that generates the map. Below the code is a table titled "wapo-temperature-change.json" with columns for Region_ID, years_2020_2039, years_2040_2059, and year. The table lists several data points. At the bottom of the interface, there is a copyright notice: "© Stadia Maps © OpenMapTiles © OpenStreetMap contributors".

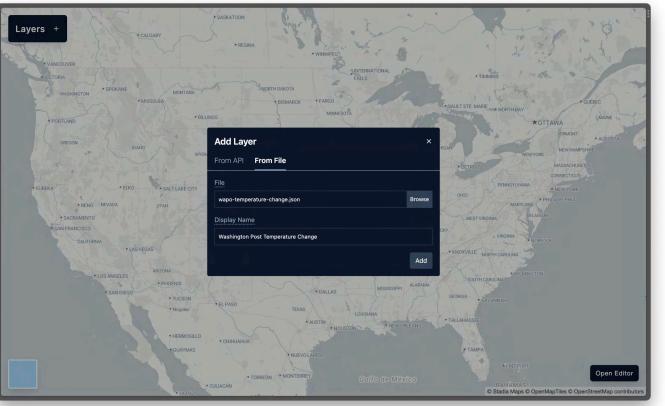
Region_ID	years_2020_2039	years_2040_2059	year
CAN.1.2.28	-65.3281598200321	-163.334143536894	-471
CAN.1.17.403	-59.1677151807567	-146.68549469866	-415
CAN.2.34.951	-52.9761497092573	-128.46697888707	-342
CAN.11.259.4274	-87.1755278245942	-275.897268533766	-856
CAN.11.269.4448	-78.1681517253813	-241.2227686884024	-749
CAN.2.34.941	-41.7140861264925	-96.348160306398	-231
CAN.8.117.2357	-130.286632328369	-486.757162116164	-127
CAN.6.98.2246	-152.228643773868	-433.465834690033	-127
CAN.8.118.2365	-131.1424850282	-365.139013292446	-107
CAN.8.117.2356	-108.036617353292	-332.204738759988	-112
CAN.8.117.2347	-133.329028398639	-406.148634475286	-127
CAN.4.100.2214	95.12017064550	230.47540702000	-120

cartokit

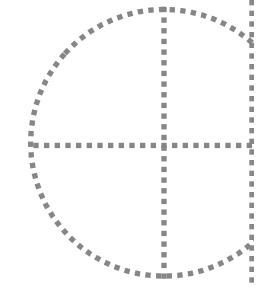
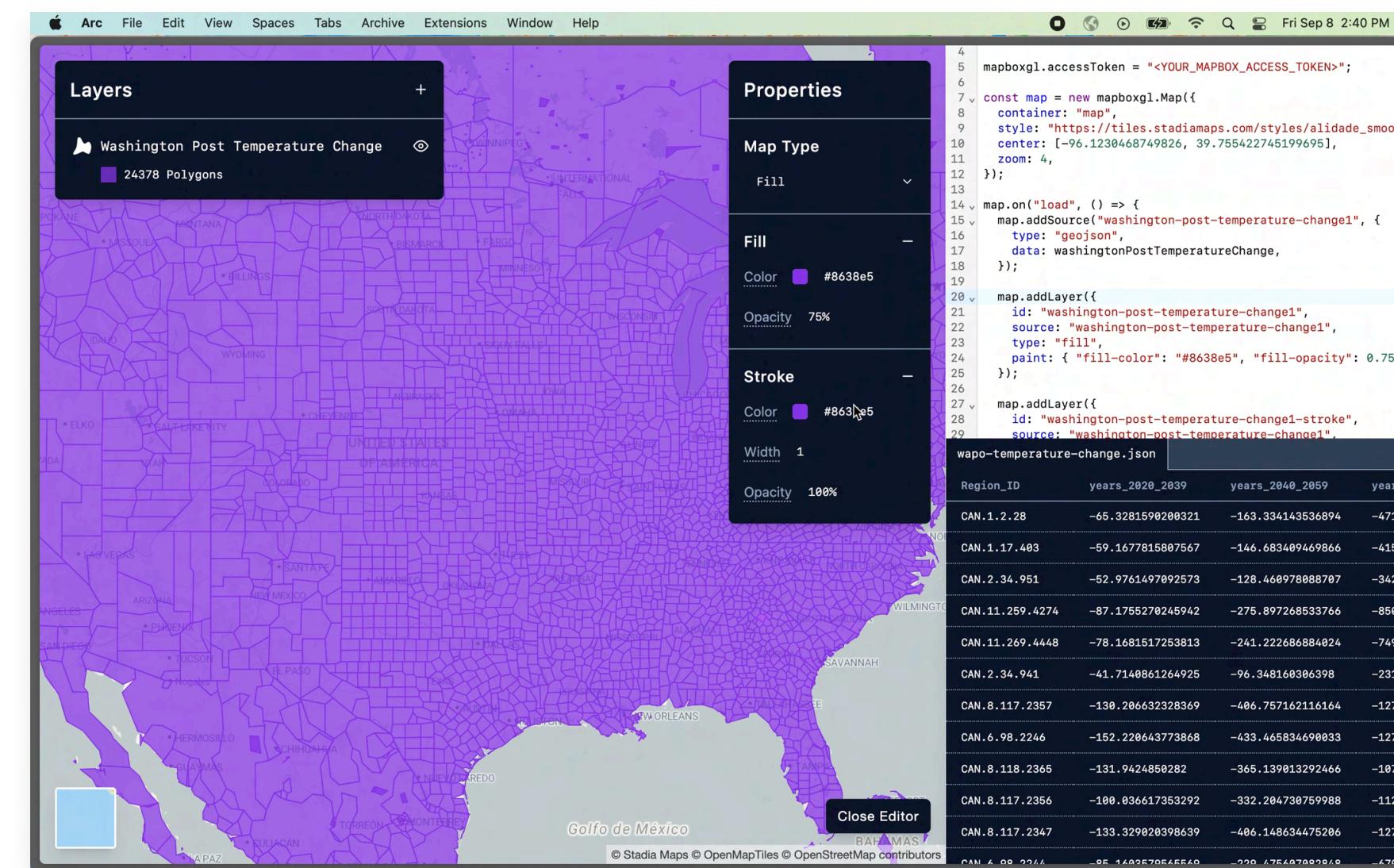
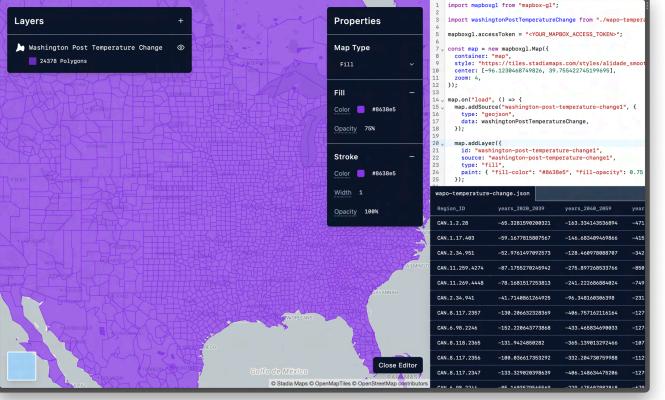
A user styles the map via **direct manipulation** while **cartokit** recompiles a matching program.



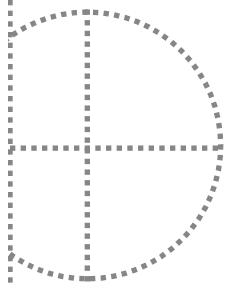
1.



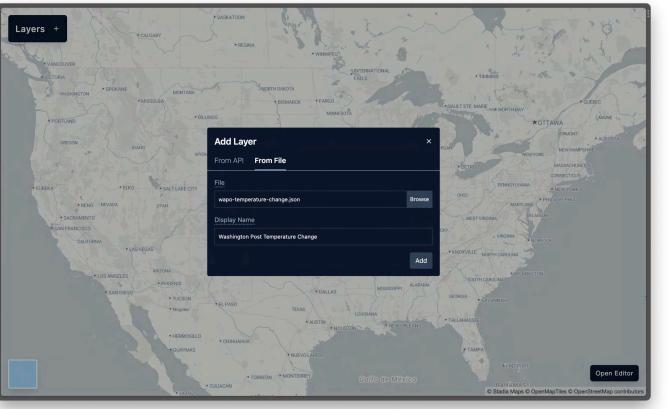
2.



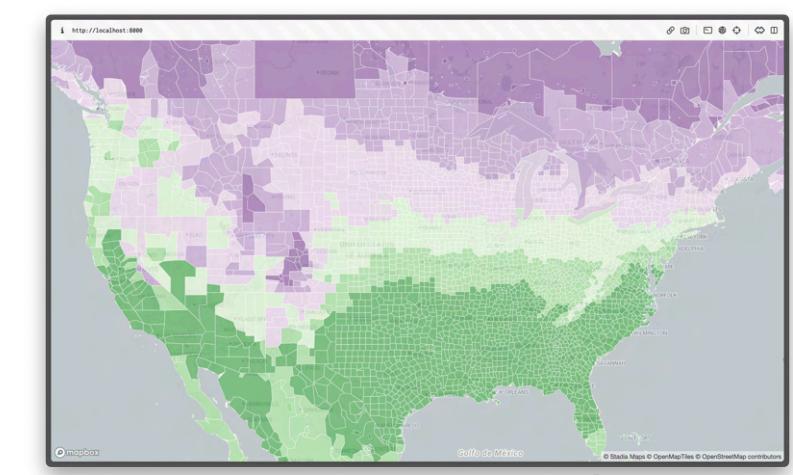
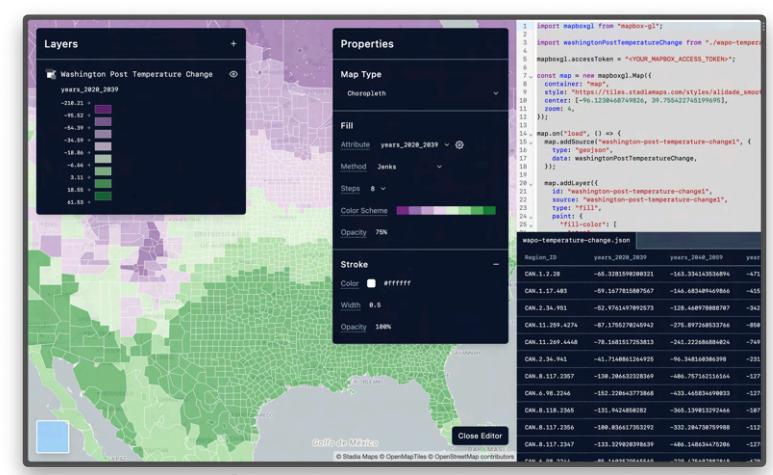
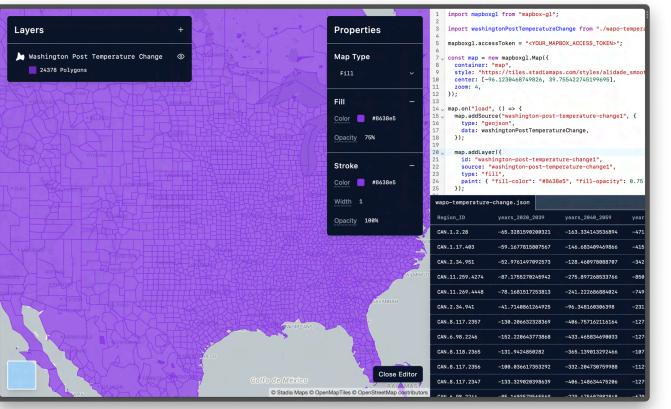
A user can **copy**, **modify**, and **deploy** the compiled program as desired.



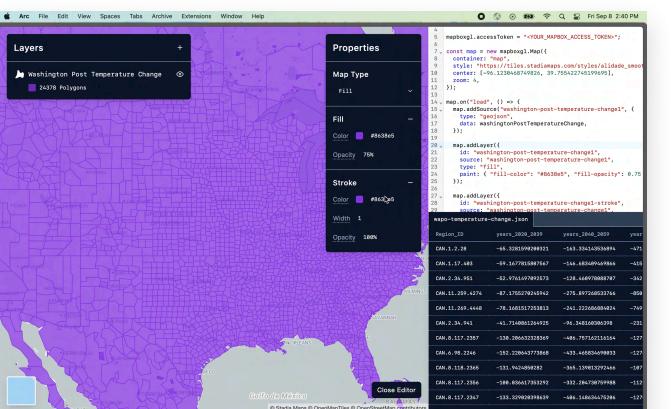
1.



2.



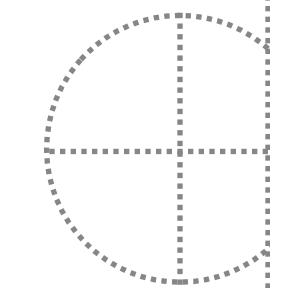
3.



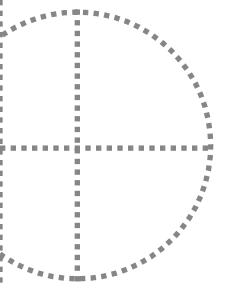
Copy

Deploy

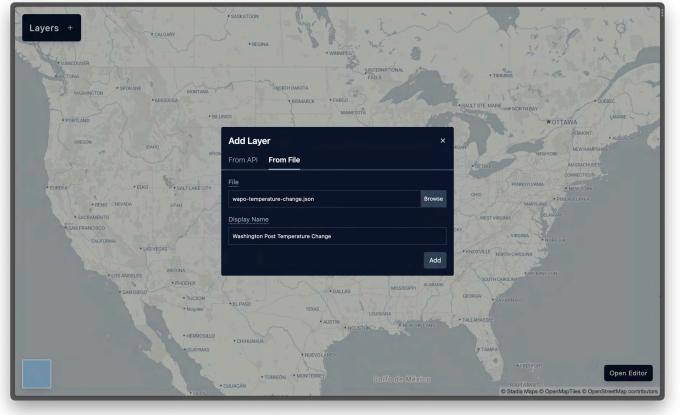
Modify



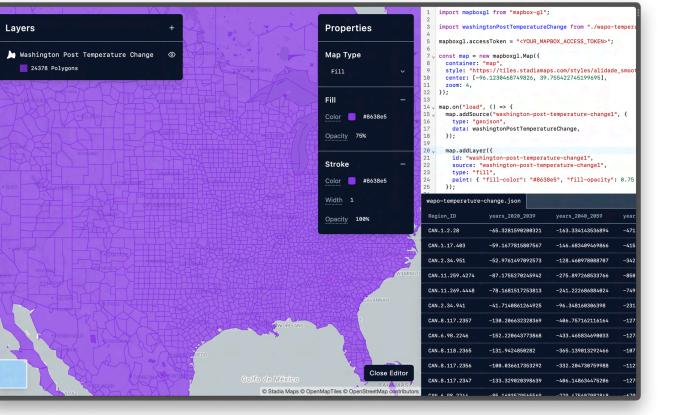
cartokit



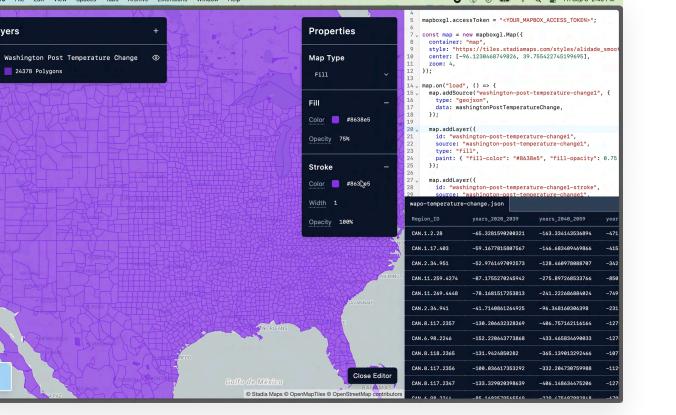
1.



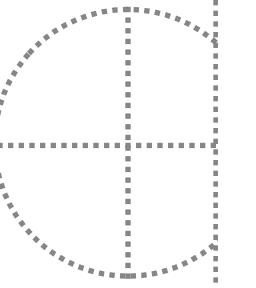
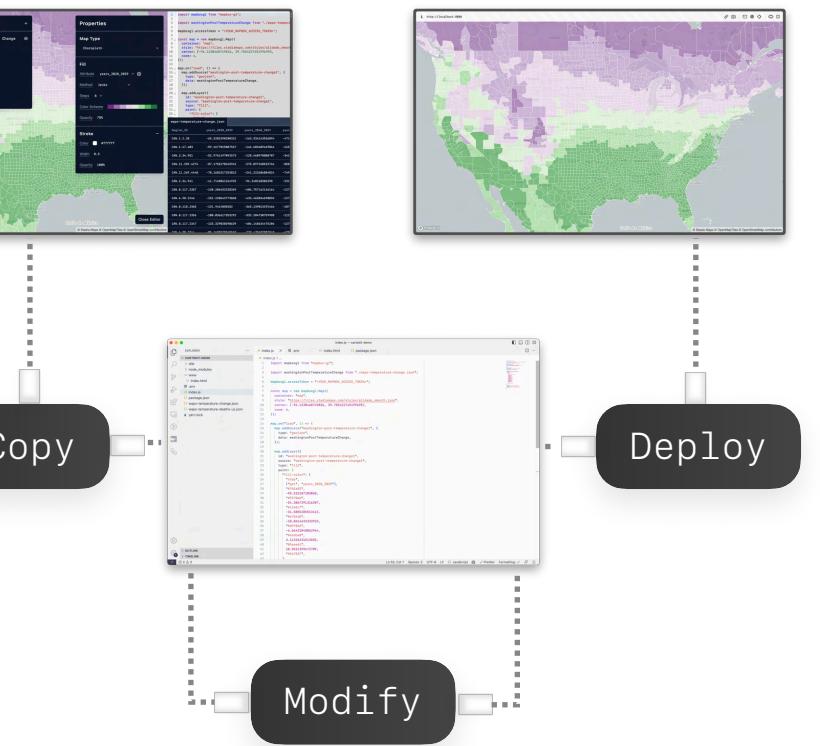
2.

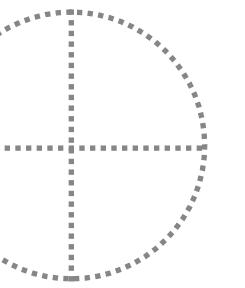


3.



4.

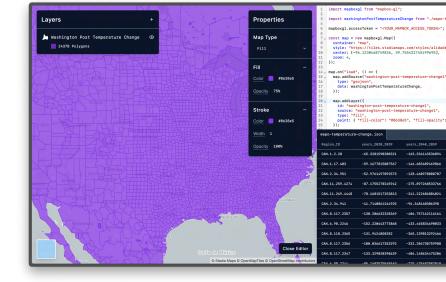




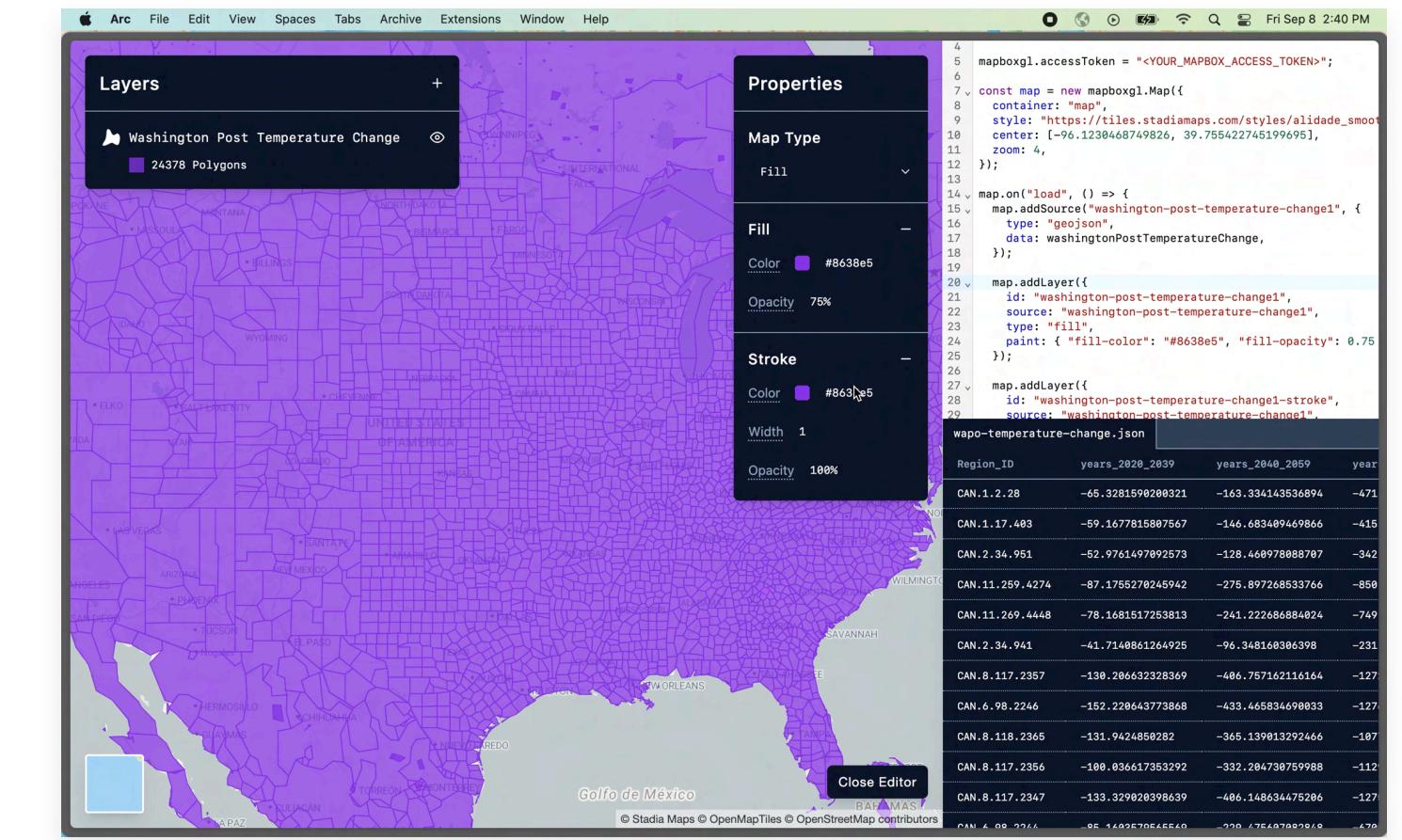
1.



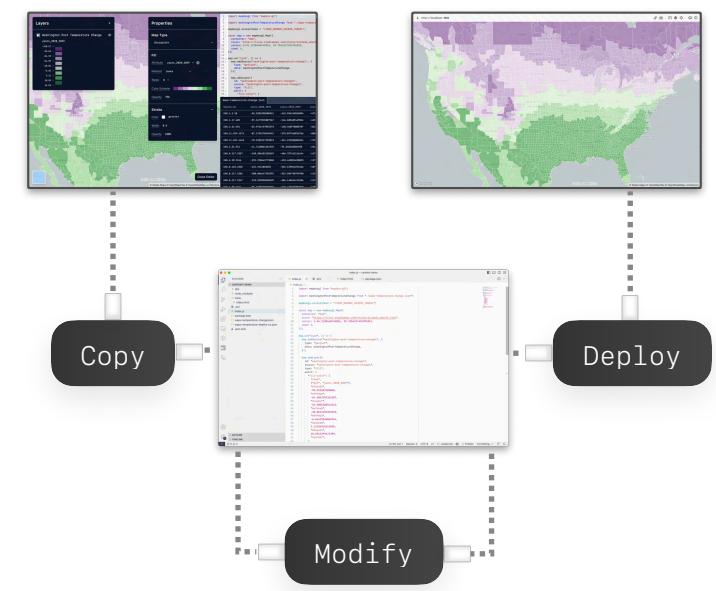
2.



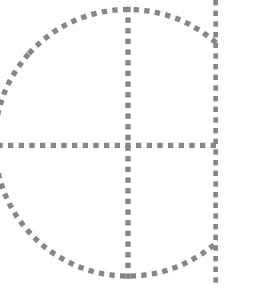
3.



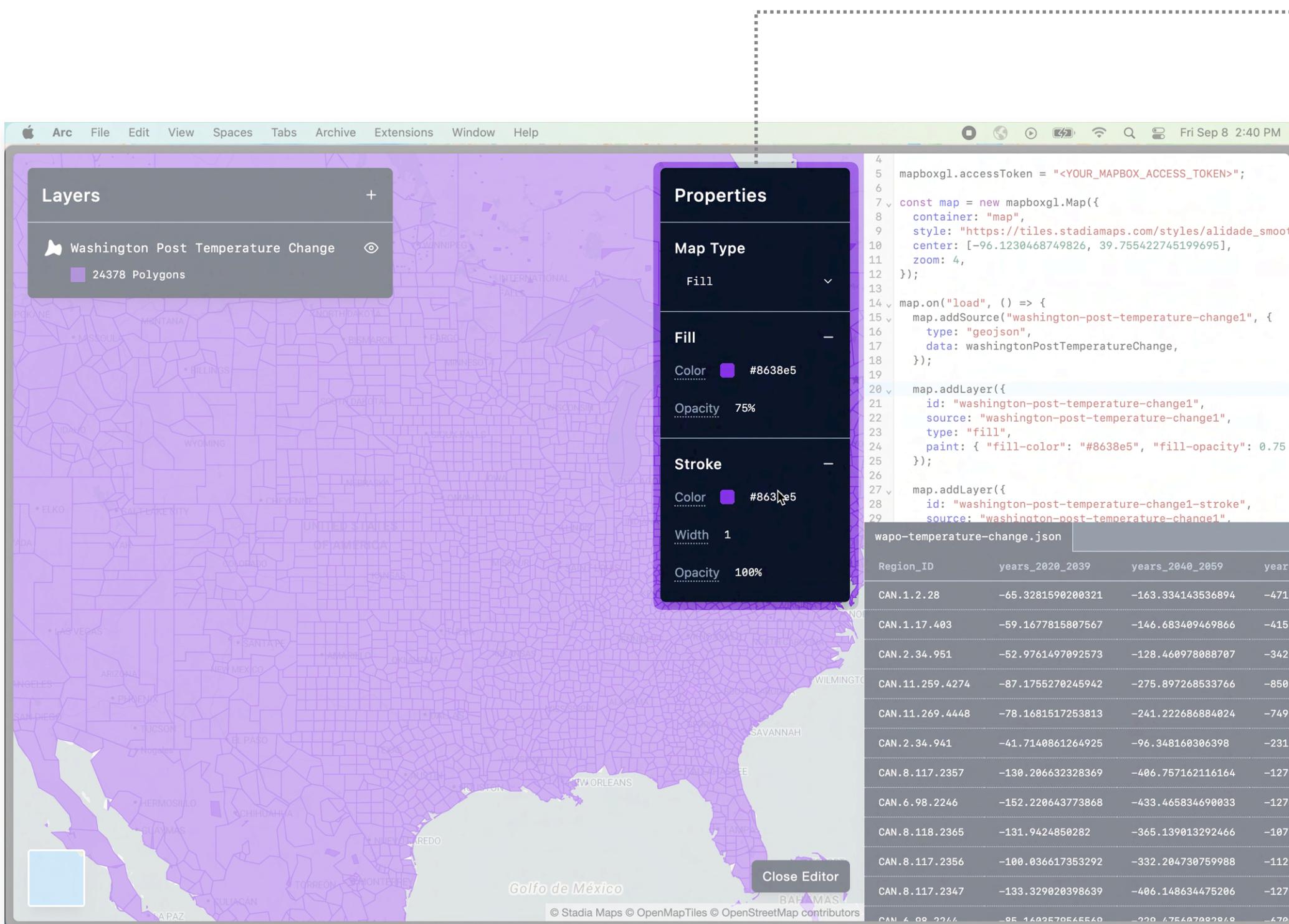
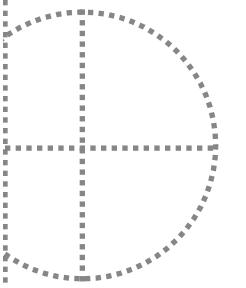
4.



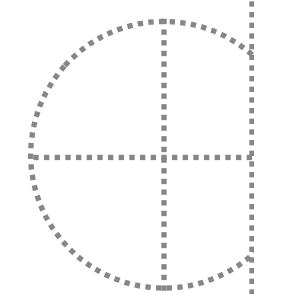
How do we get from **interactions** to **program edits**?



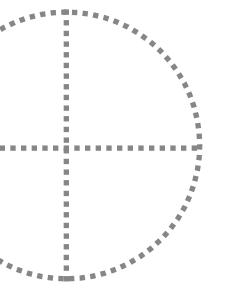
Map property controls in the UI to an intermediate representation.



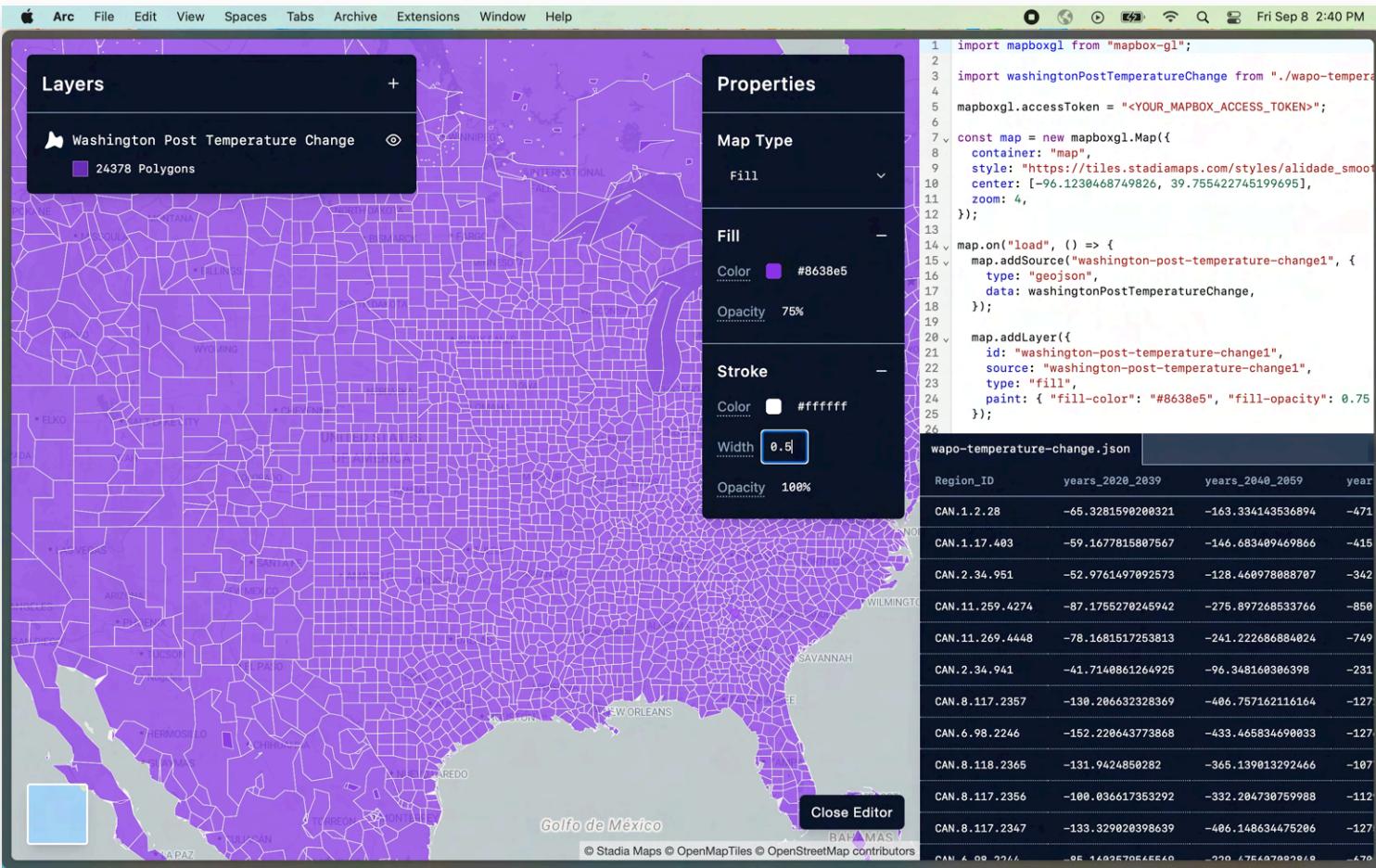
```
const ir = {
  center: [-106.1086, 37.7531],
  zoom: 4,
  basemap: {
    url: "https://tiles.stadiamaps.com/styl...",
    provider: "Stadia Maps",
  },
  layers: [
    {
      type: "Fill",
      style: {
        fill: {
          color: "#8638e5",
          opacity: 0.75,
        },
        stroke: {
          color: "#ffffff",
          width: 1,
          opacity: 0.5,
        },
      },
    },
  ],
};
```



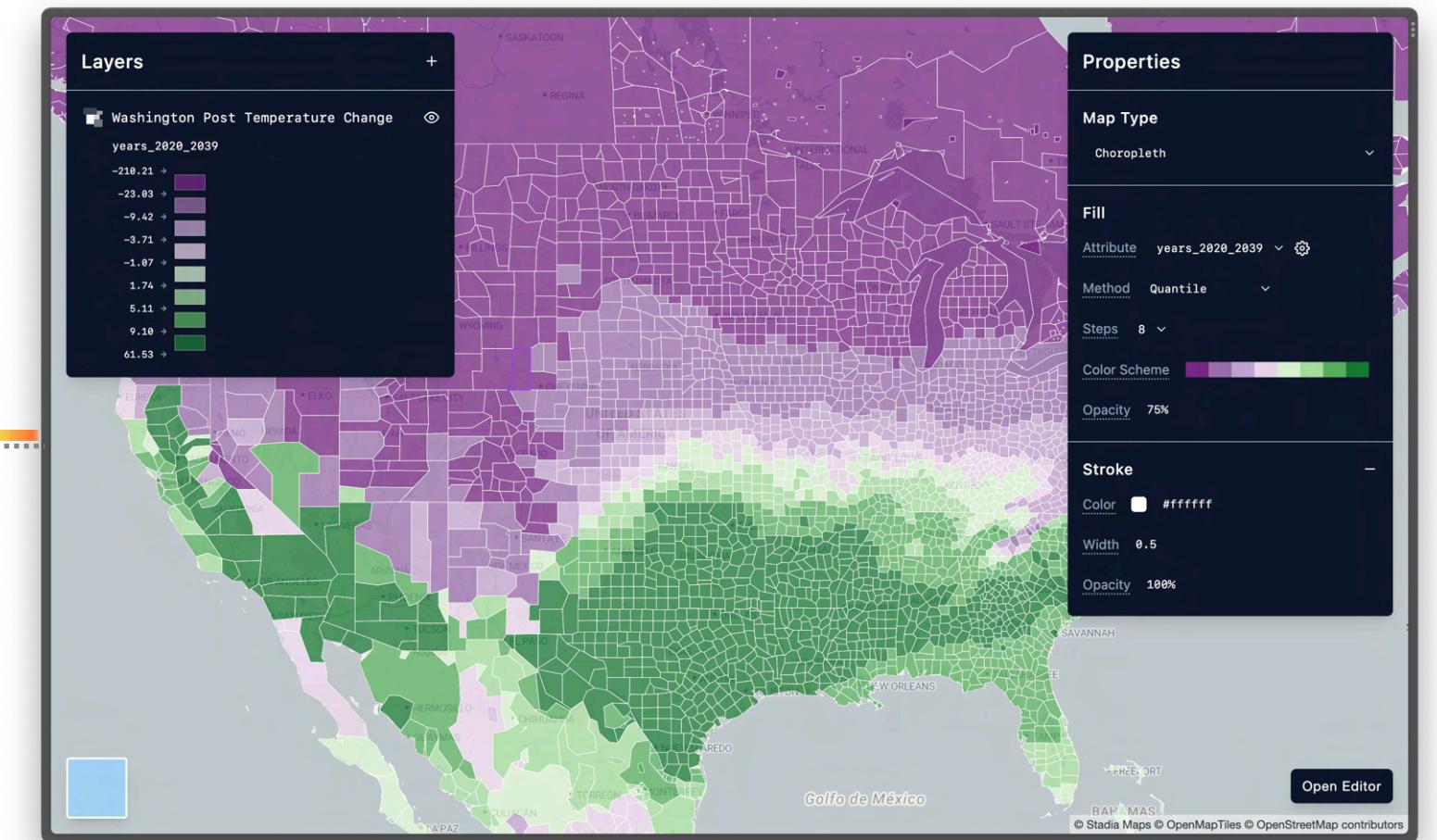
When a user alters a property in the UI, **dispatch an update** to the IR.



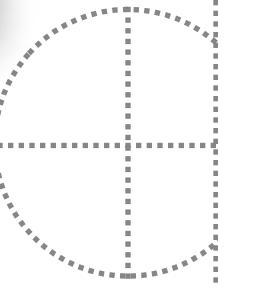
Fill Layer



Choropleth Layer



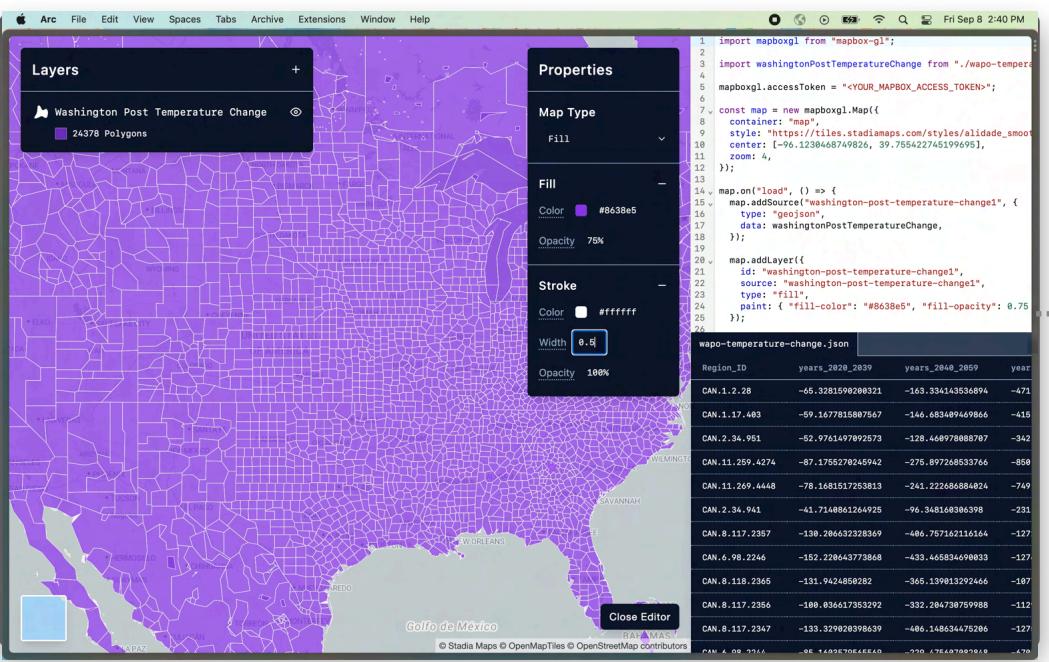
dispatchLayerUpdate



When a user alters a property in the UI, **dispatch an update** to the IR.

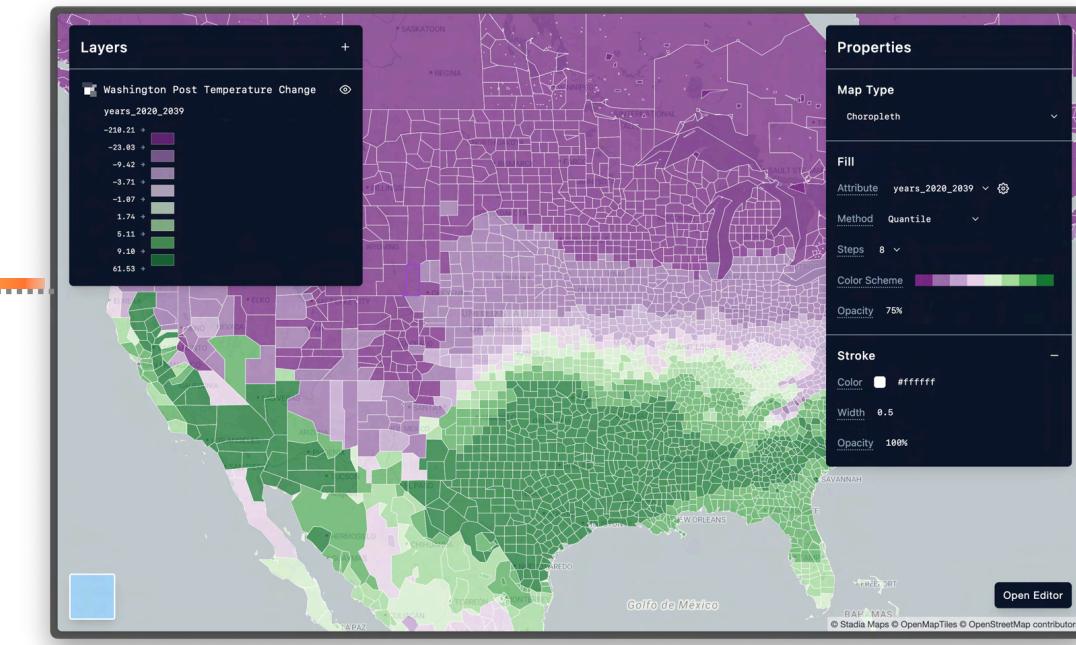


Fill Layer



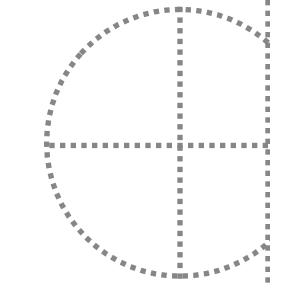
dispatchLayerUpdate

Choropleth Layer



```
const ir = {
  center: [-106.1086, 37.7531],
  zoom: 4,
  basemap: {
    url: "https://tiles.stadiamaps.com/styl...", provider: "Stadia Maps",
  },
  layers: [
    {
      type: "Fill",
      style: {
        fill: {
          color: "#8638e5",
          opacity: 0.75,
        },
        stroke: {
          color: "#ffffff",
          width: 1,
          opacity: 0.5,
        },
      },
    },
  ],
};
```

```
const ir = {
  center: [-106.1086, 37.7531],
  zoom: 4,
  basemap: {
    url: "https://tiles.stadiamaps.com/styl...", provider: "Stadia Maps",
  },
  layers: [
    {
      type: "Choropleth",
      style: {
        fill: {
          attribute: "years_2020_2039",
          scale: "Quantile",
          count: 8,
          scheme: d3.schemeBuPu,
          thresholds: [-210.21, -23.03, ...],
          opacity: 0.75,
        },
        stroke: {
          color: "#ffffff",
          width: 1,
          opacity: 0.5,
        },
      },
    },
  ],
};
```



Begin **code generation**, starting with dependencies.

```
const ir = {  
  center: [-106.1086, 37.7531],  
  zoom: 4,  
  basemap: {  
    url: "https://tiles.stadiamaps.com/styl...",  
    provider: "Stadia Maps",  
  },  
  layers: [  
    {  
      type: "Choropleth",  
      style: {  
        fill: {  
          attribute: "years_2020_2039",  
          scale: "Quantile",  
          count: 8,  
          scheme: d3.schemeBuPu,  
          thresholds: [-210.21, -23.03, ...],  
          opacity: 0.75,  
        },  
        stroke: {  
          color: "#ffffff",  
          width: 1,  
          opacity: 0.5,  
        },  
      },  
    },  
  ],  
};
```

codegenImports

```
`import mapboxgl from "mapbox-gl";  
${transformTable.size > 0  
  ? 'import * as turf from "@turf/turf"';  
  : ""}  
${codegenFileImports(ir)}  
${codegenFns(ir, transformTable)}  
${codegenMap({ map, ir, uploadTable, transformTable })}`
```

Pass execution to **codegen functions** to produce **program fragments** for distinct parts of the IR.

```
const ir = {  
  center: [-106.1086, 37.7531],  
  zoom: 4,  
  basemap: {  
    url: "https://tiles.stadiamaps.com/styl...",  
    provider: "Stadia Maps",  
  },  
  layers: [  
    {  
      type: "Choropleth",  
      style: {  
        fill: {  
          attribute: "years_2020_2039",  
          scale: "Quantile",  
          count: 8,  
          scheme: d3.schemeBuPu,  
          thresholds: [-210.21, -23.03, ...],  
          opacity: 0.75,  
        },  
        stroke: {  
          color: "#ffffff",  
          width: 1,  
          opacity: 0.5,  
        },  
      },  
    },  
  ],  
};
```

codegenImports

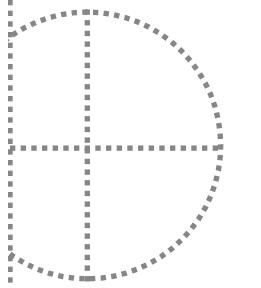
```
`import mapboxgl from "mapbox-gl";  
${transformTable.size > 0  
? 'import * as turf from "@turf/turf"';  
: ""}  
${codegenFileImports(ir)}  
${codegenFns(ir, transformTable)}  
${codegenMap({ map, ir, uploadTable, transformTable })}`
```

codegenFileImports

codegenFns

codegenMap

Pass execution to **codegen functions** to produce **program fragments** for distinct parts of the IR.



```
const ir = {  
  center: [-106.1086, 37.7531],  
  zoom: 4,  
  basemap: {  
    url: "https://tiles.stadiamaps.com/styl...",  
    provider: "Stadia Maps",  
  },  
  layers: [  
    {  
      type: "Choropleth",  
      style: {  
        fill: {  
          attribute: "years_2020_2039",  
          scale: "Quantile",  
          count: 8,  
          scheme: d3.schemeBuPu,  
          thresholds: [-210.21, -23.03, ...],  
          opacity: 0.75,  
        },  
        stroke: {  
          color: "#ffffff",  
          width: 1,  
          opacity: 0.5,  
        },  
      },  
    },  
  ],  
};
```

codegenImports

```
`import mapboxgl from "mapbox-gl";  
${transformTable.size > 0  
? 'import * as turf from "@turf/turf"';  
: ""}  
${codegenFileImports(ir)}  
${codegenFns(ir, transformTable)}  
${codegenMap({ map, ir, uploadTable, transformTable })}`
```

codegenFileImports

codegenFns

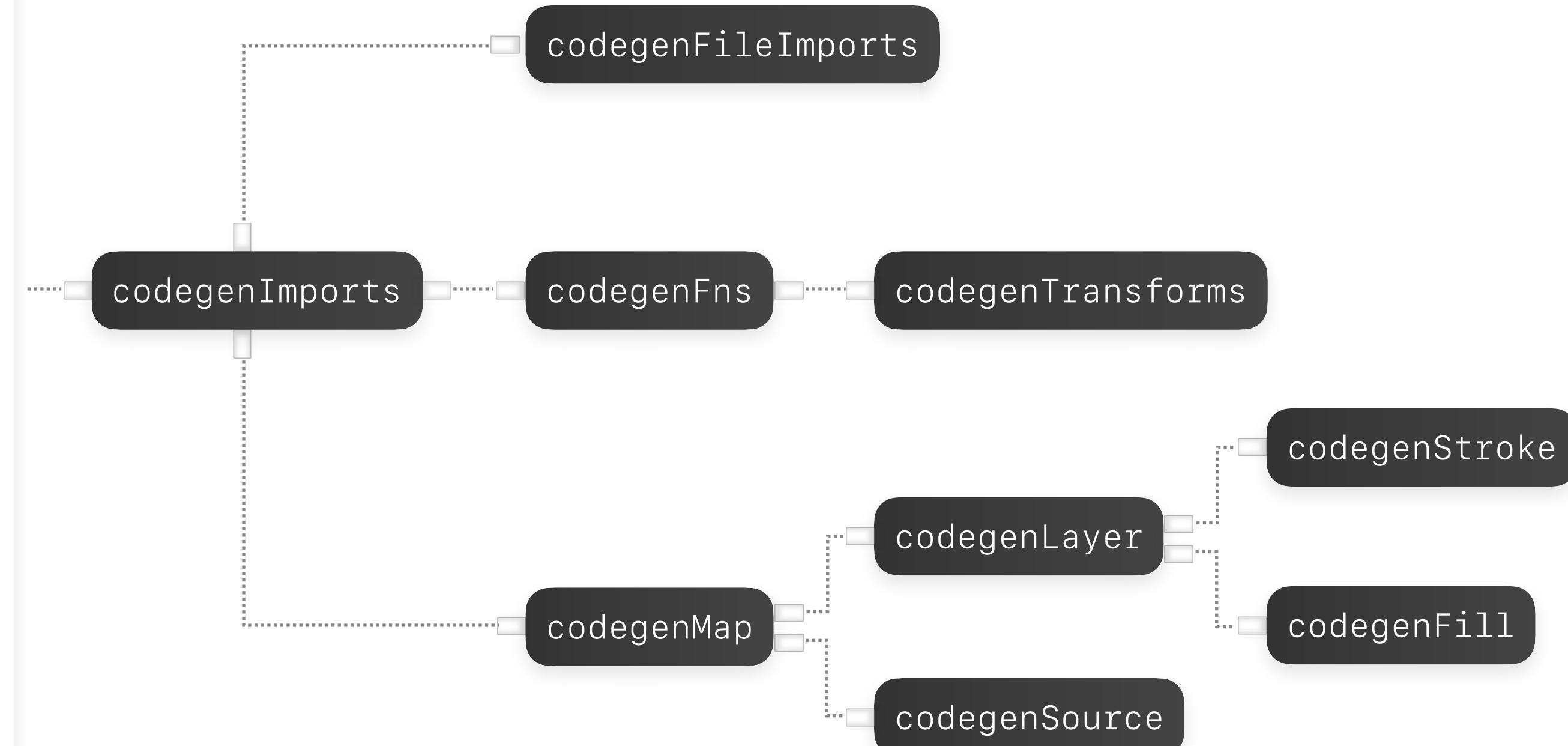
codegenMap

```
`const map = new mapboxgl.Map({  
  container: "map",  
  style: "${ir.basemap.url}",  
  center: [${ir.center[0]}, ${ir.center[1]}],  
  zoom: ${map.getZoom()}  
});  
  
map.on('load', ${isLoadSync ? 'async' : ''}() => {  
  ${Object.values(ir.layers).reduce((p, layer) => {  
    return p.concat("\n\n" + codegenSource(layer, uploadTable));  
  }, "")}  
  ...`
```

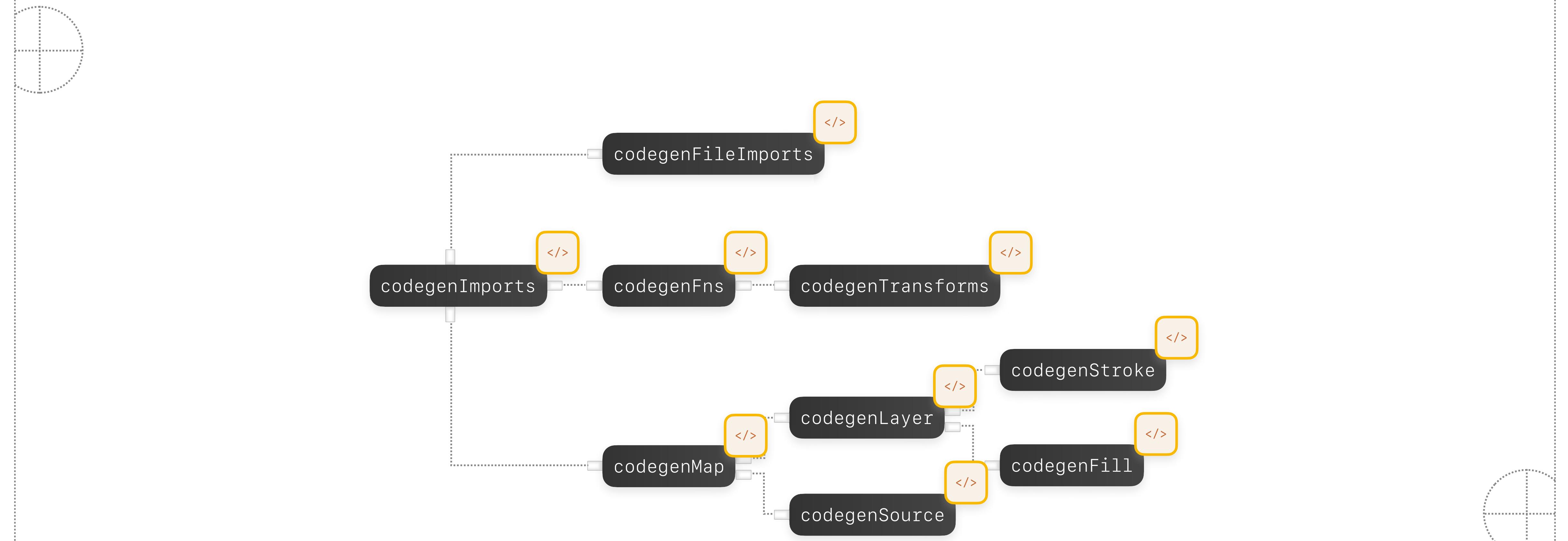


Pass execution to **codegen functions** to produce **program fragments** for distinct parts of the IR.

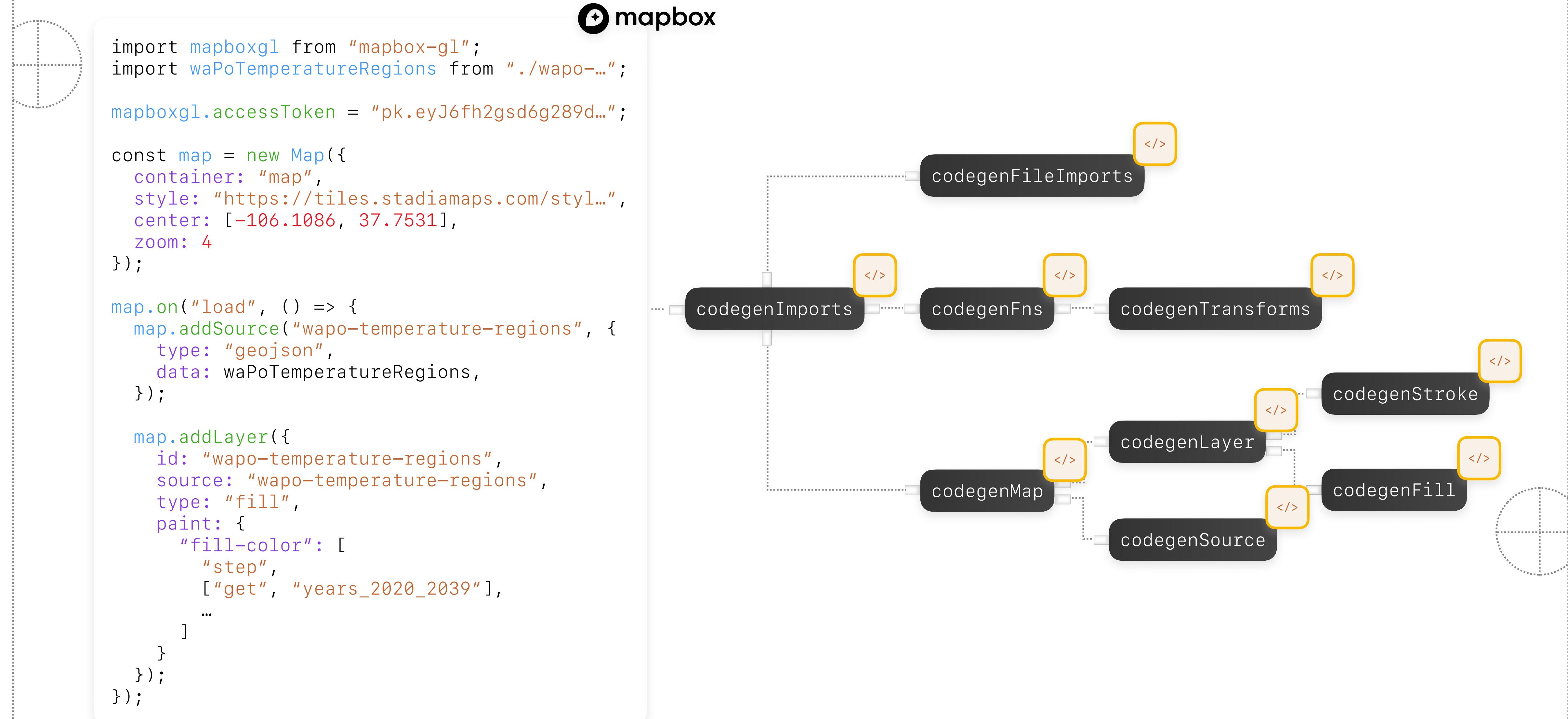
```
const ir = {  
  center: [-106.1086, 37.7531],  
  zoom: 4,  
  basemap: {  
    url: "https://tiles.stadiamaps.com/styl...",  
    provider: "Stadia Maps",  
  },  
  layers: [  
    {  
      type: "Choropleth",  
      style: {  
        fill: {  
          attribute: "years_2020_2039",  
          scale: "Quantile",  
          count: 8,  
          scheme: d3.schemeBuPu,  
          thresholds: [-210.21, -23.03, ...],  
          opacity: 0.75,  
        },  
        stroke: {  
          color: "#ffffff",  
          width: 1,  
          opacity: 0.5,  
        },  
      },  
    },  
  ],  
};
```



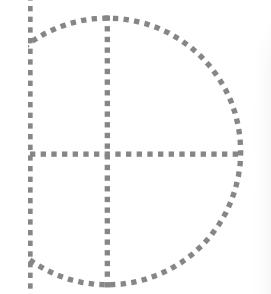
Each codegen function produces its own **program fragment**, with callers deciding **insertion points**.



The **program fragments** are assembled into an output JavaScript program.



The **program fragments** are assembled into an output JavaScript program.



```
import mapboxgl from "mapbox-gl";
import waPoTemperatureRegions from "./wapo-...";

mapboxgl.accessToken = "pk.eyJ1IjoiZmF5b2JhM2JhIiwiYSI6ImNqY2cxaGZz...";

const map = new Map({
  container: "map",
  style: "https://tiles.stadiamaps.com/styl...",
  center: [-106.1086, 37.7531],
  zoom: 4
});

map.on("load", () => {
  map.addSource("wapo-temperature-regions", {
    type: "geojson",
    data: waPoTemperatureRegions,
  });

  map.addLayer({
    id: "wapo-temperature-regions",
    source: "wapo-temperature-regions",
    type: "fill",
    paint: {
      "fill-color": [
        "step",
        ["get", "years_2020_2039"],
        ...
      ]
    }
  });
});
```



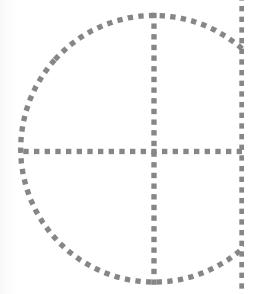
```
import L from "leaflet";
import waPoTemperatureRegions from "./wapo-...";

const map = L.map("map")
  .setView([-106.1086, 37.7531], 4);

L.tileLayer(
  "https://tiles.stadiamaps.com/tiles/alidavinci-0.0.1/20/{z}/{x}/{y}.png",
  {
    maxZoom: 20,
    attribution:
      '&copy; <a href="https://stadiamaps.com/> Stadia Maps',
  }
).addTo(map);

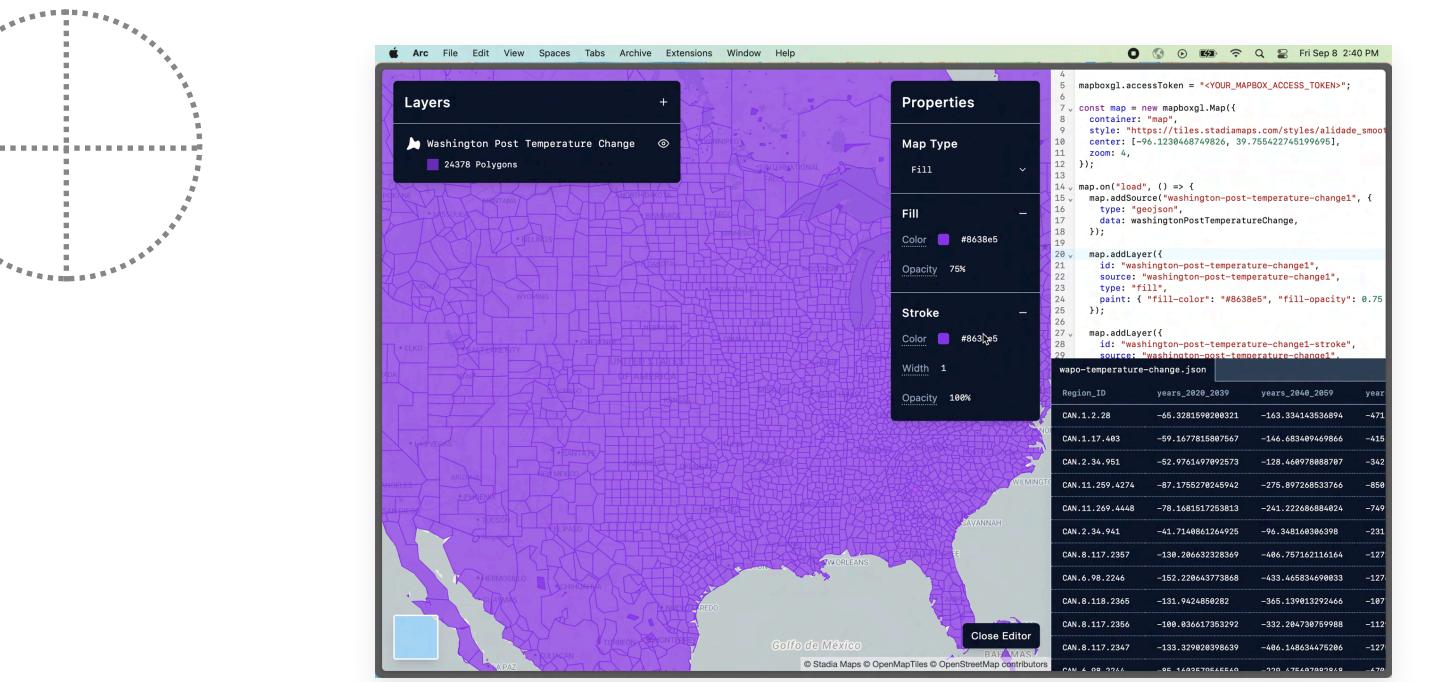
L.geoJSON(waPoTemperatureRegions, {
  style: (feature) => {
    const attr =
      feature.properties["years_2020_2039"];

    if (attr < -23.03) {
      return {
        fillColor: "#762a83",
        fillOpacity: 0.75,
        color: "#FFFFFF",
        weight: 0.5,
      };
    } else if (attr < 9.42) {
      ...
    },
  }
}).addTo(map);
```

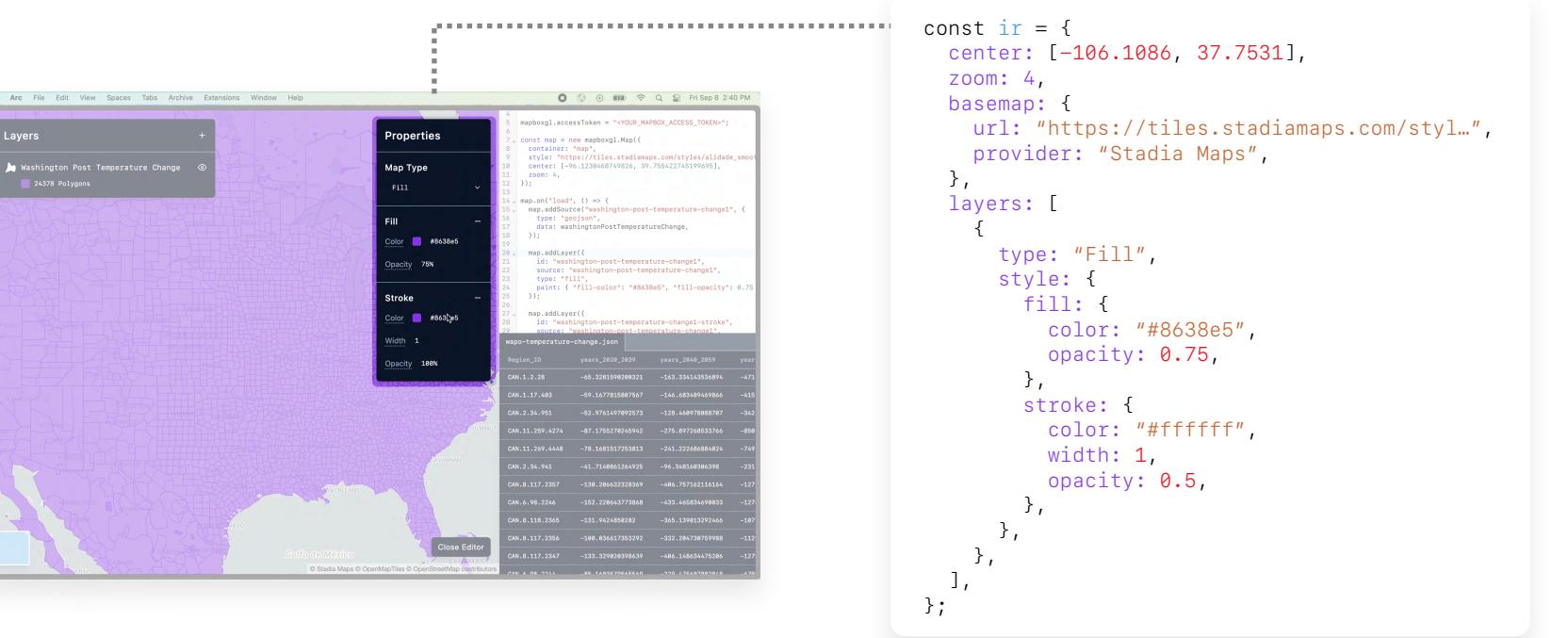


cartokit

Direct Manipulation



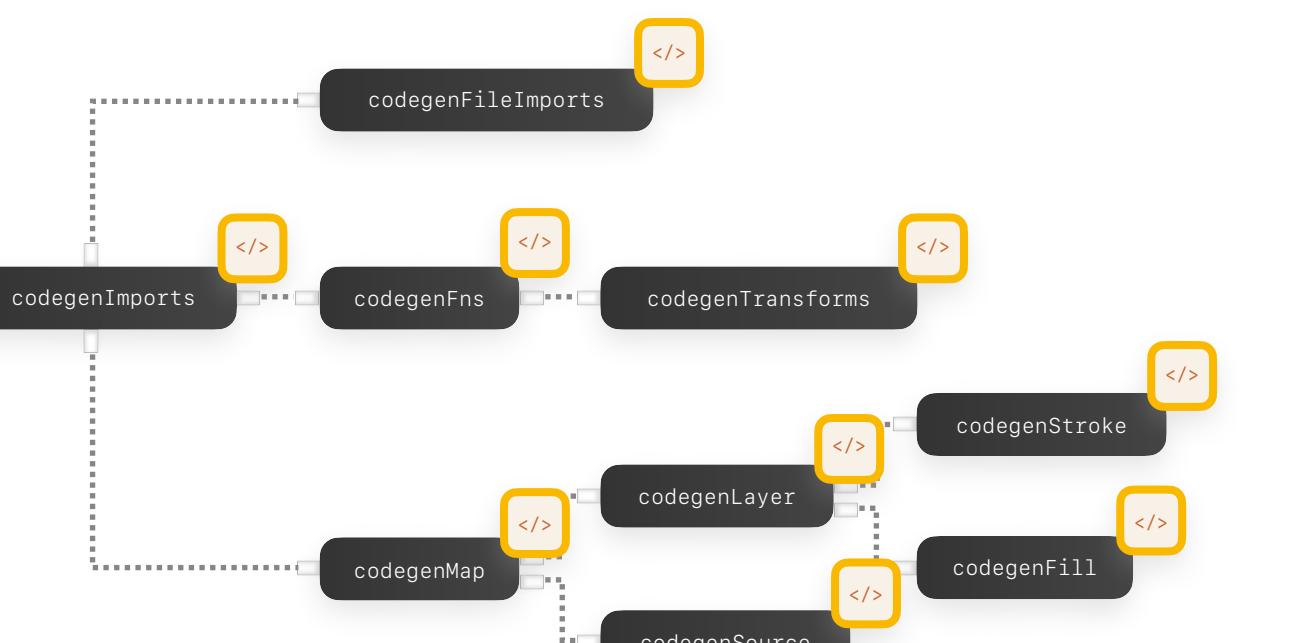
UI controls → Intermediate Representation



Update IR



Code Generation



JavaScript Program

```

import mapboxgl from "mapbox-gl";
import waPoTemperatureRegions from "./wapo-...";

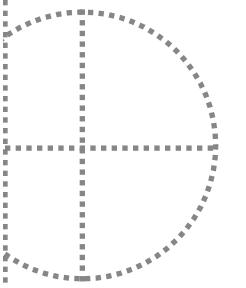
mapboxgl.accessToken = "pk.eyJ6fh2gsd6g289d...";

const map = new Map({
  container: "map",
  style: "https://tiles.stadiamaps.com/style...", center: [-106.1086, 37.7531], zoom: 4});

map.on("load", () => {
  map.addSource("wapo-temperature-regions", {
    type: "geojson",
    data: waPoTemperatureRegions,
  });

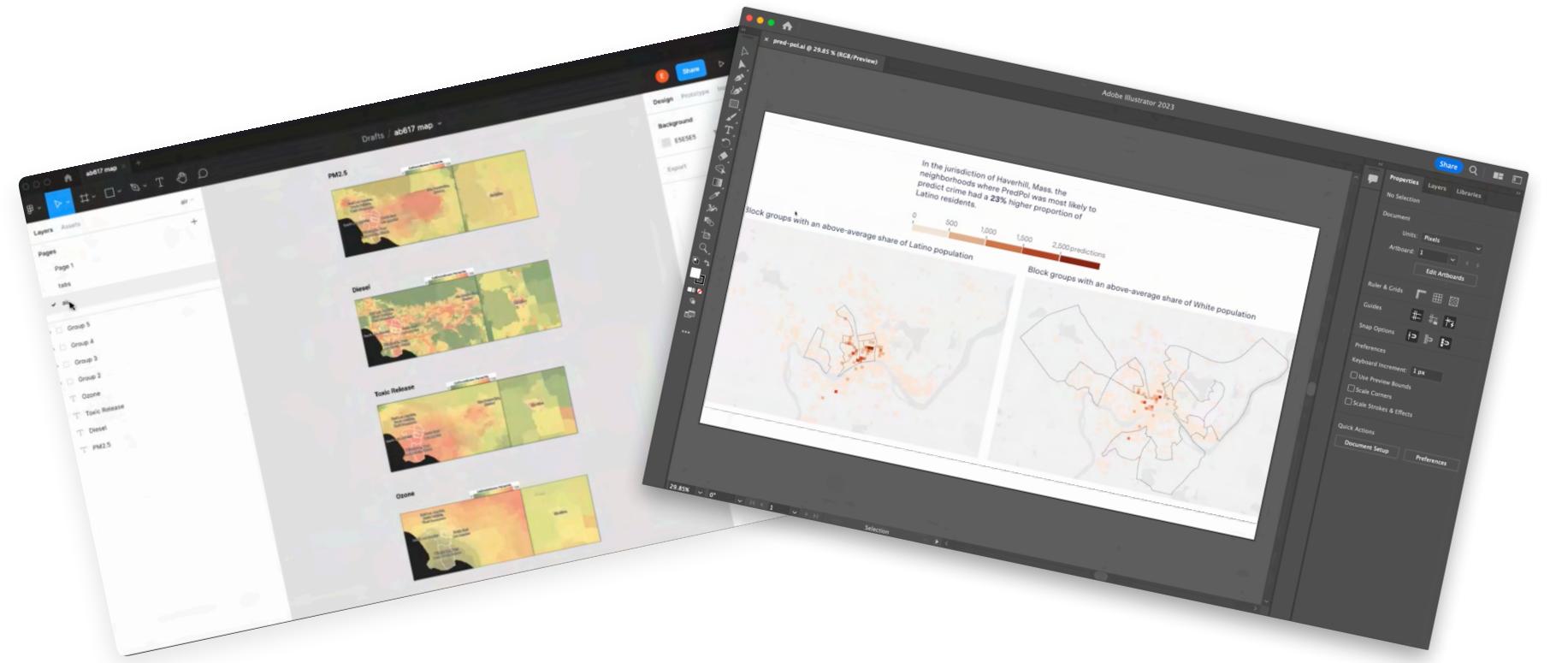
  map.addLayer({
    id: "wapo-temperature-regions",
    source: "wapo-temperature-regions",
    type: "fill",
    paint: {
      "fill-color": [
        "step",
        ["get", "years_2020_2039"],
        ...
      ];
    };
  });
}
  
```



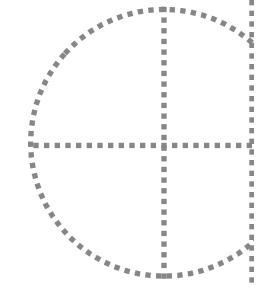
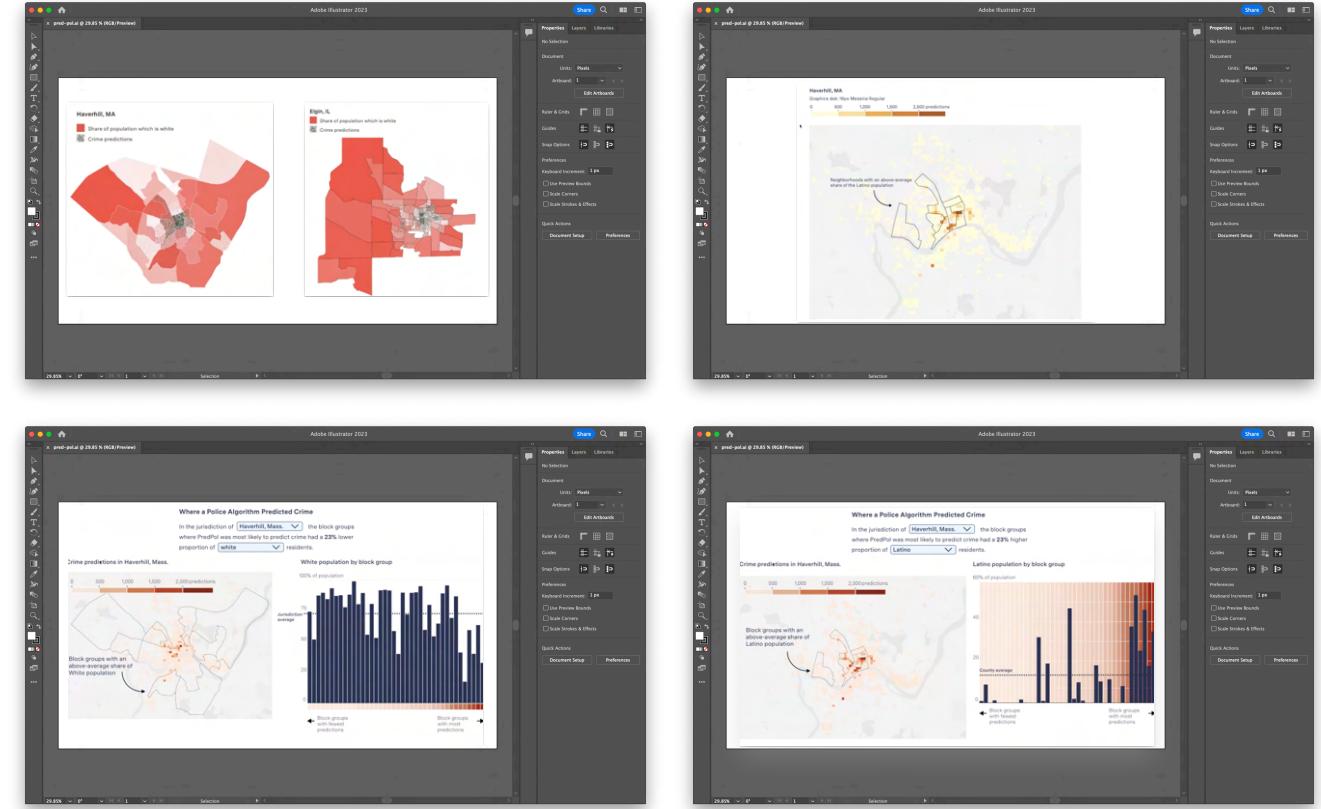


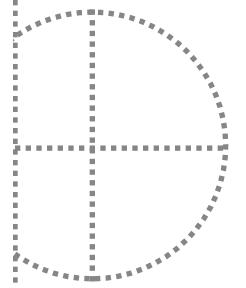
Working Practice

Sketch visualizations in **design software**



..... • Explore a wide design space **without code**

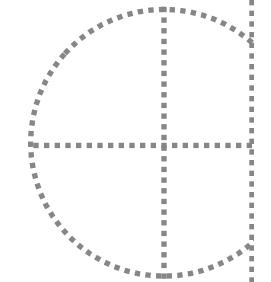
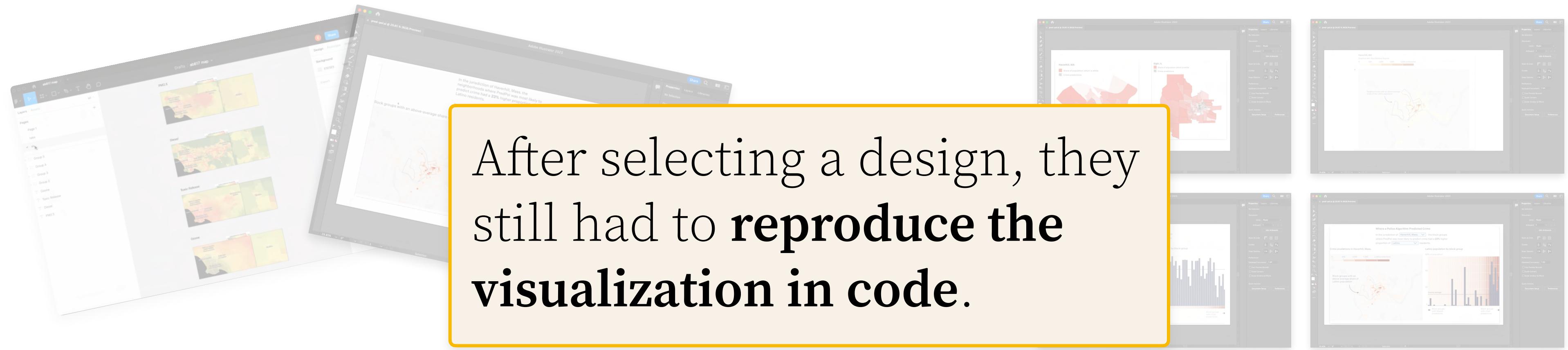




Working Practice

Sketch visualizations in **design software**

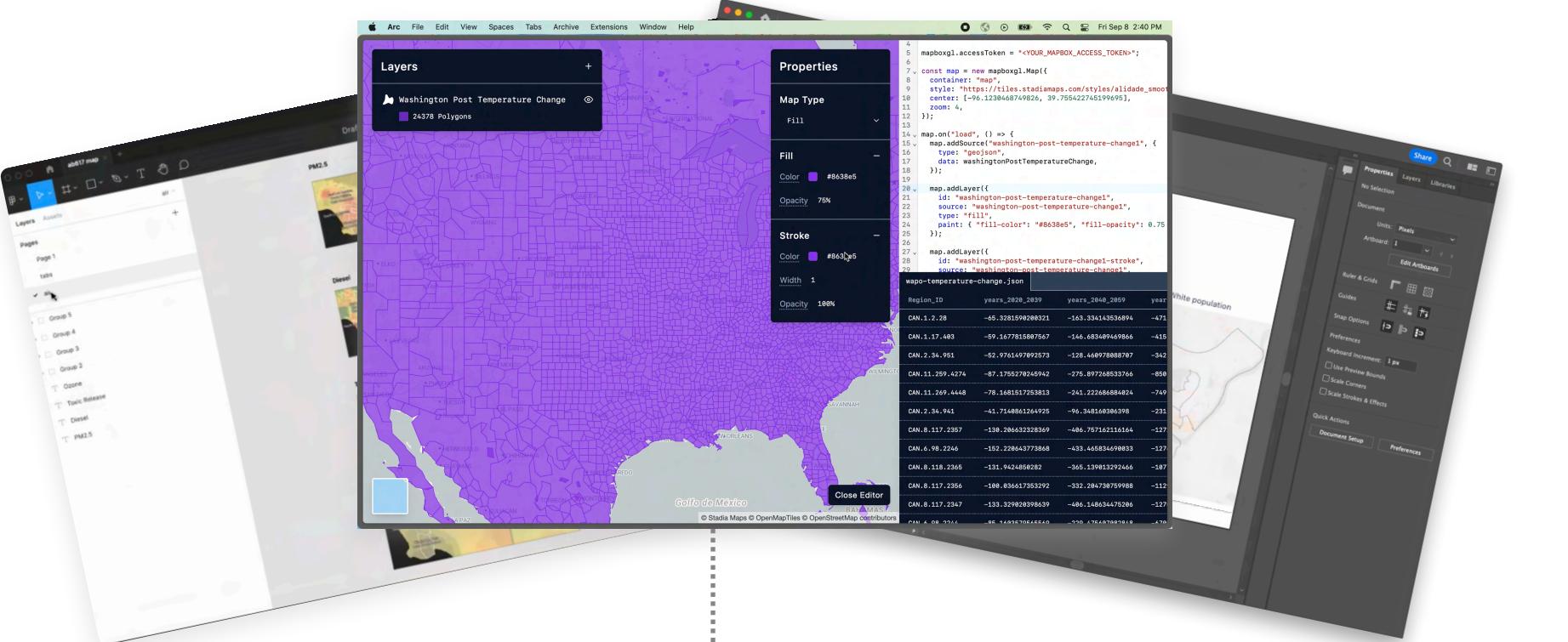
.....◦ Explore a wide design space **without code**



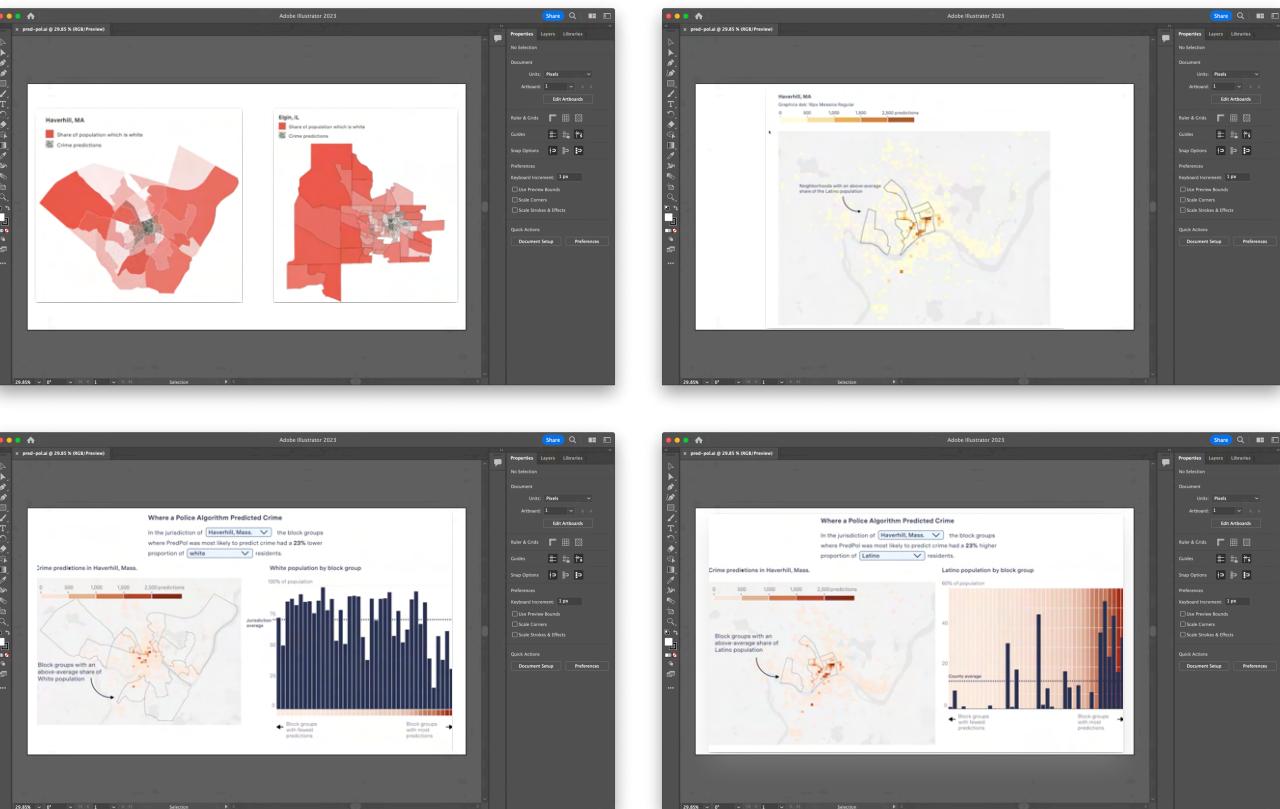
Working Practice



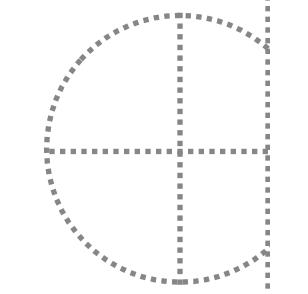
Sketch visualizations in **design software**

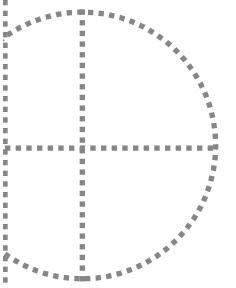


..... • Explore a wide design space **without code**



Milliseconds



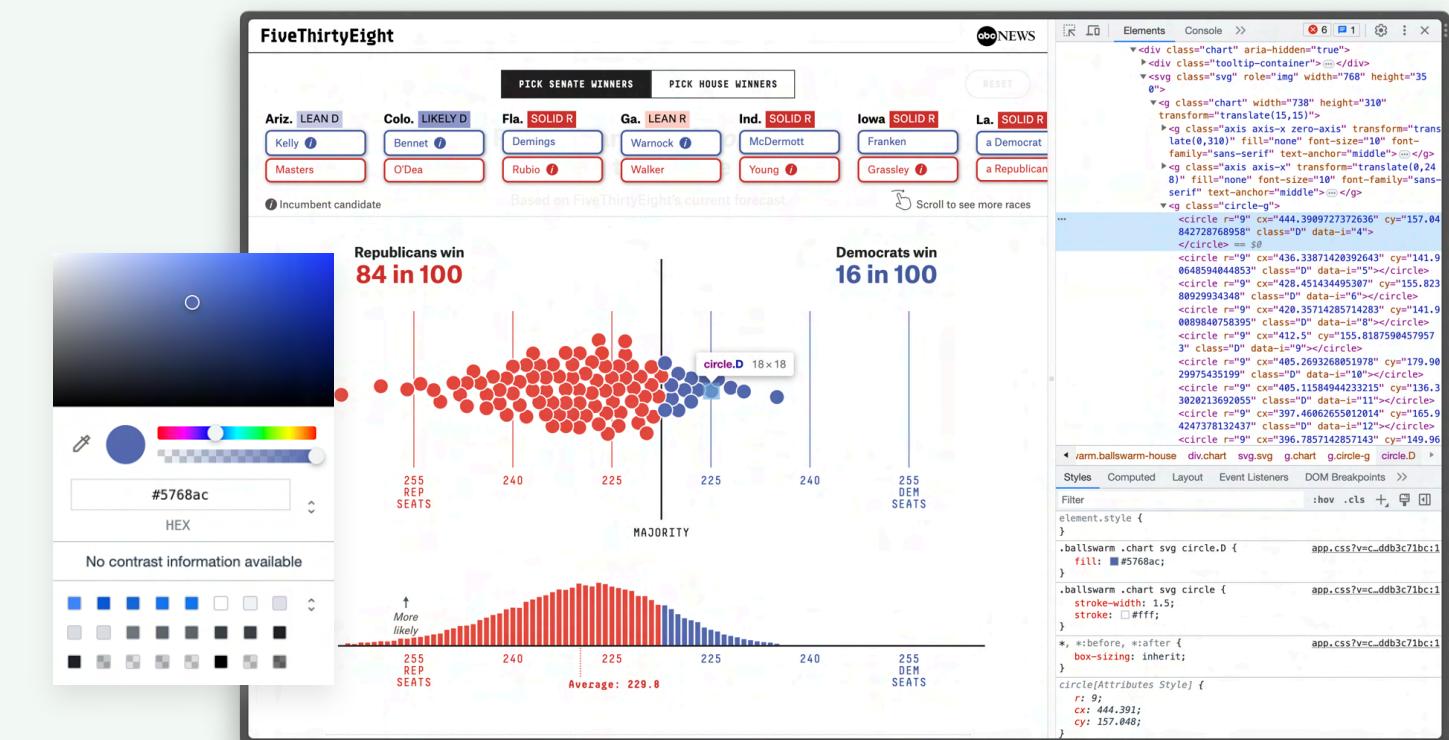


Lexer

Parser

Recover (or infer) **semantic information** from the visual form that we can represent **symbolically**.

Map **interactions** with the visual form to **program edits**.



cartokit

Compilers for Visual Inputs

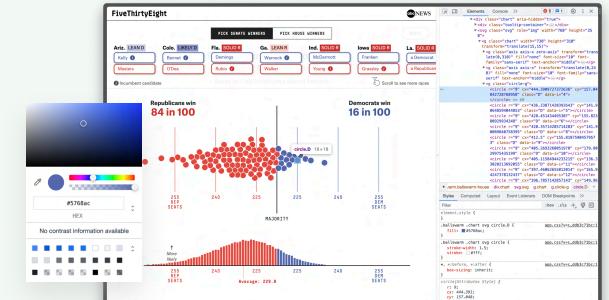


Lexer

Parser

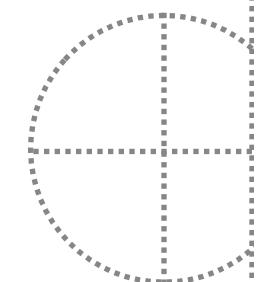
Recover (or infer) **semantic information** from the visual form that we can represent **symbolically**.

Map **interactions** with the visual form to **program edits**.



reviz

cartokit





Cartography Studio



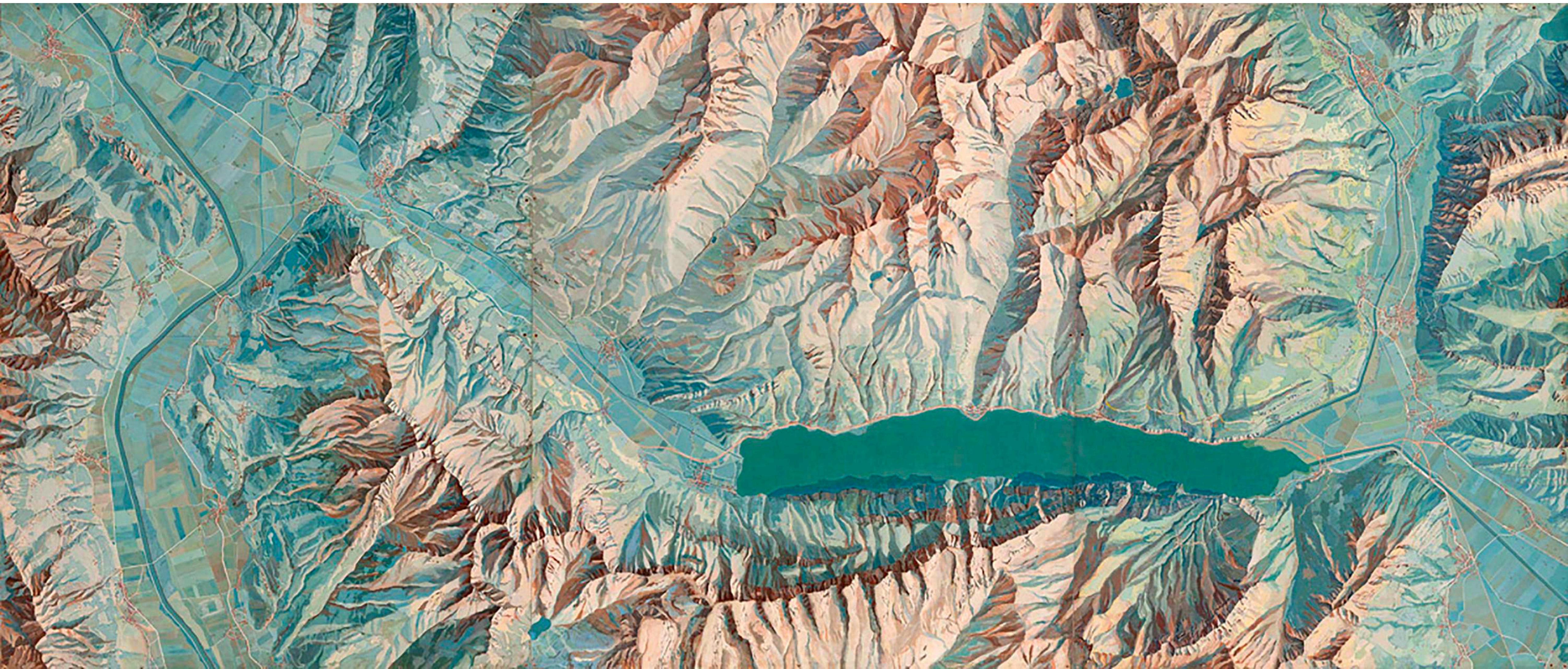
Linework

“London Underground Tube Map” · Harry Beck



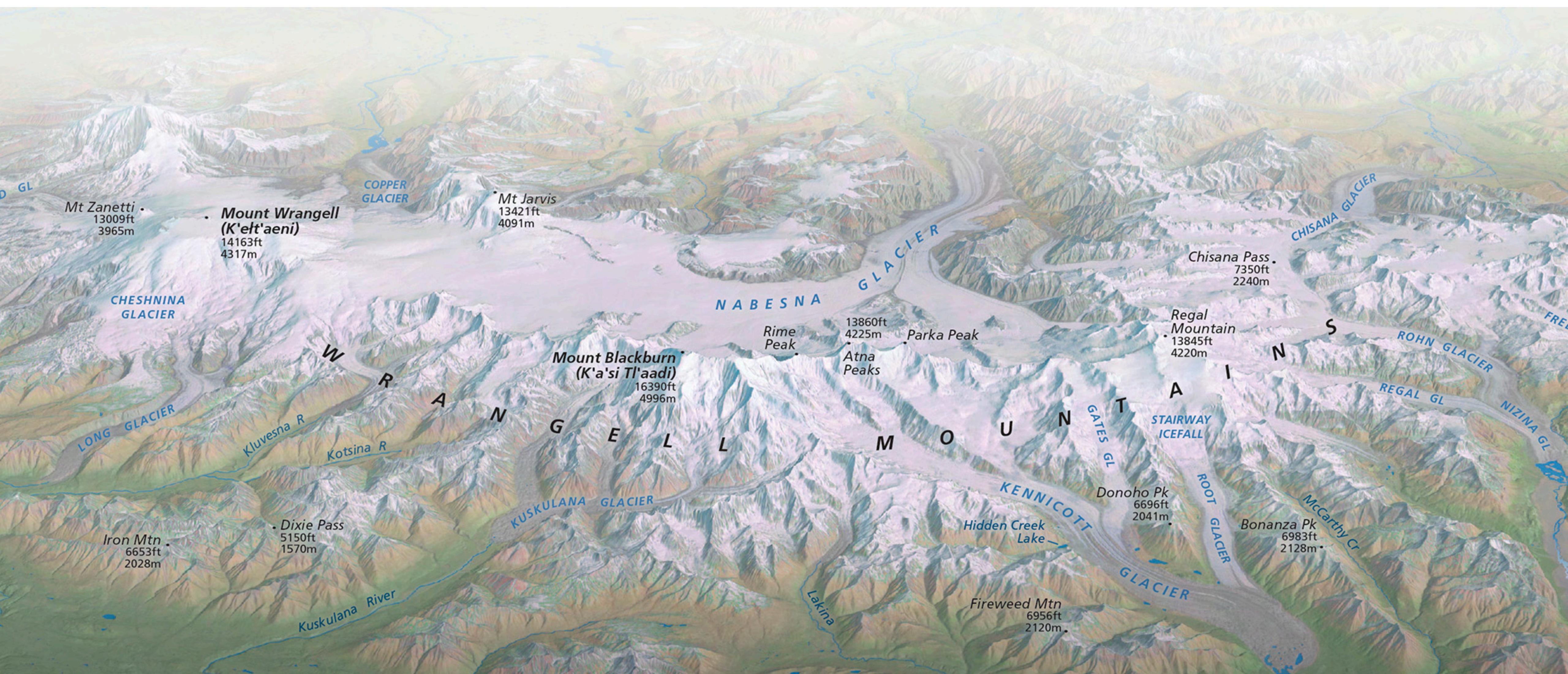
Shaded Relief

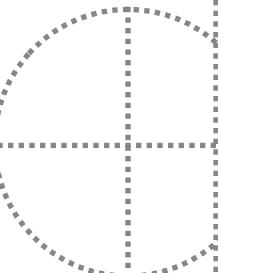
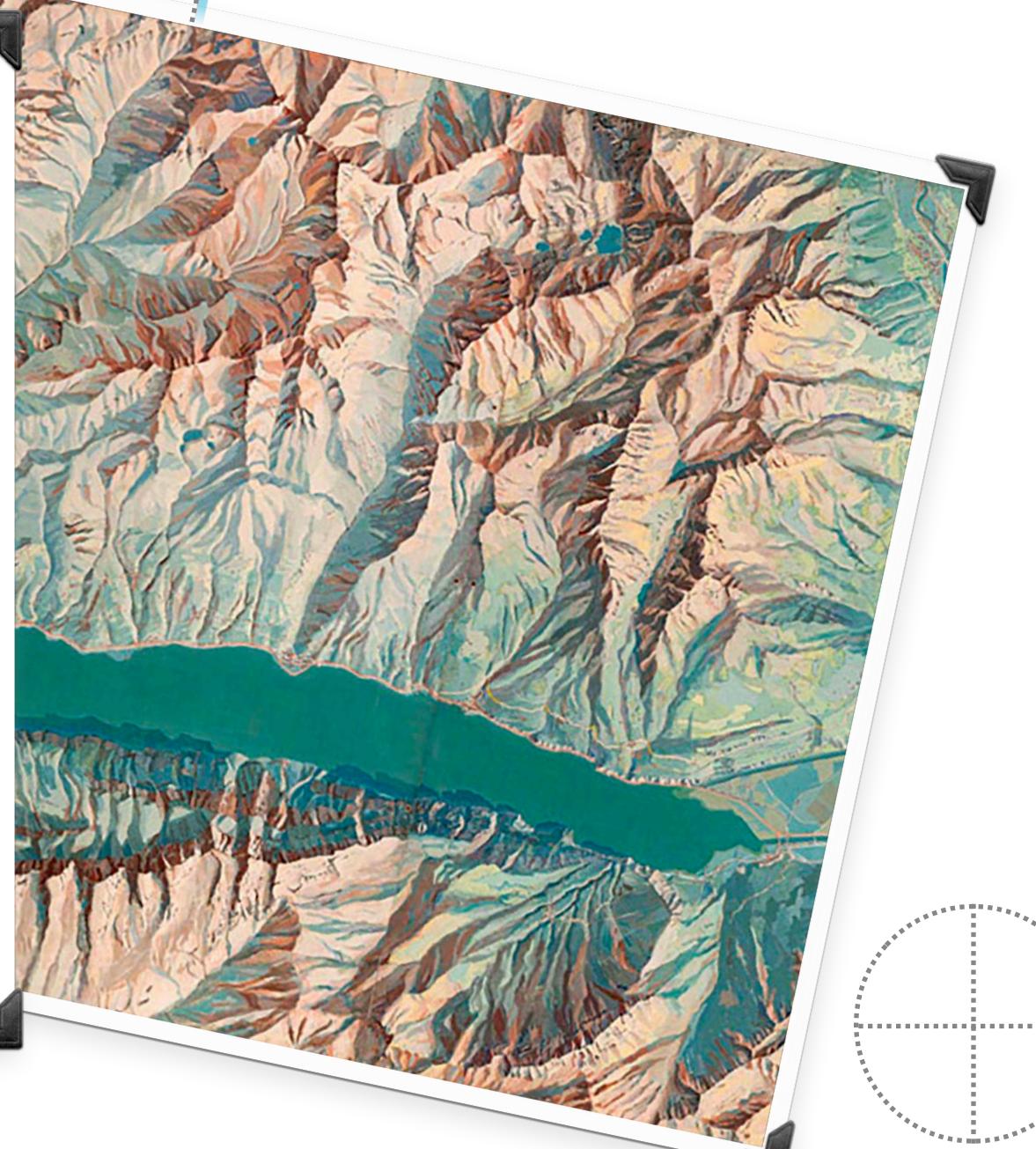
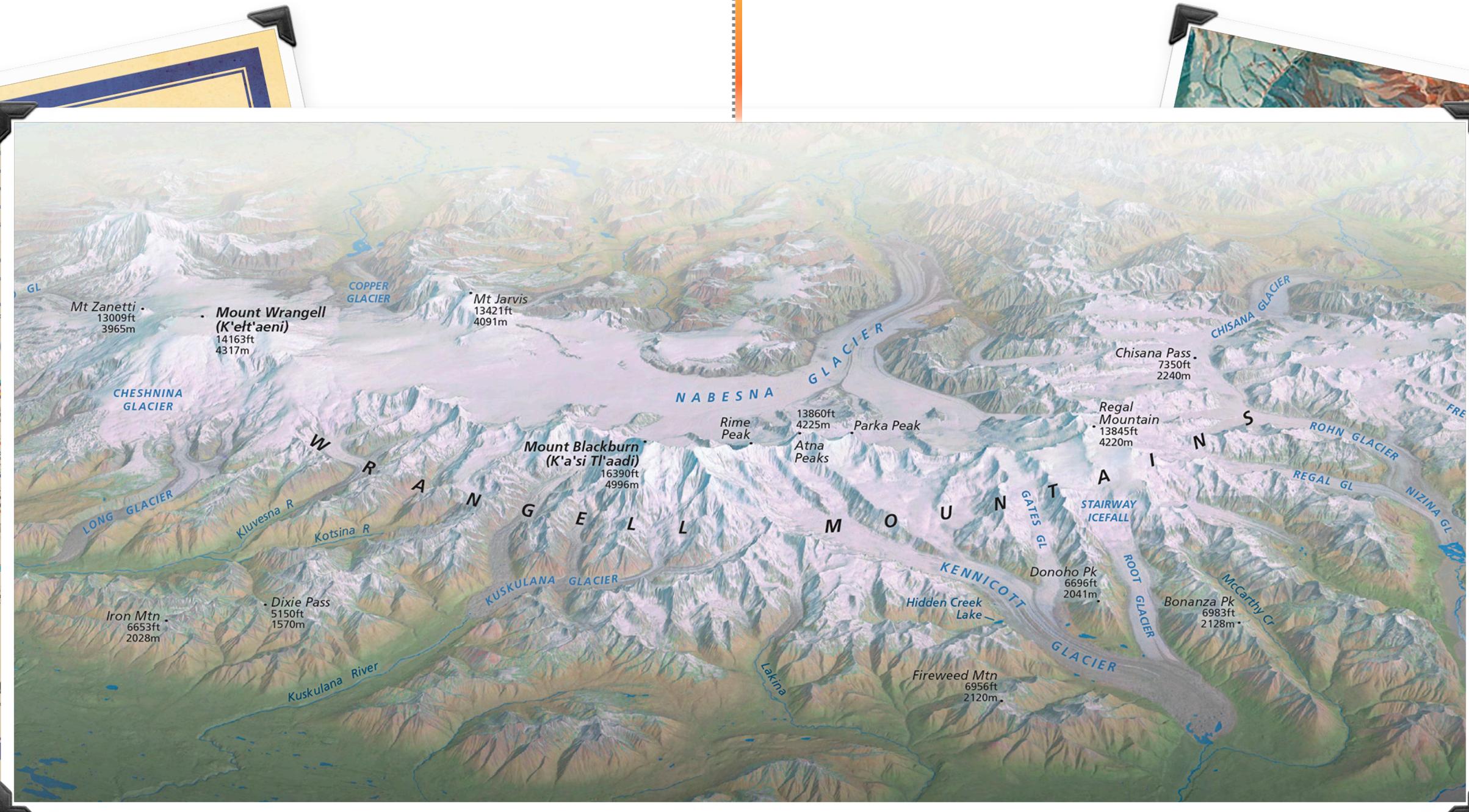
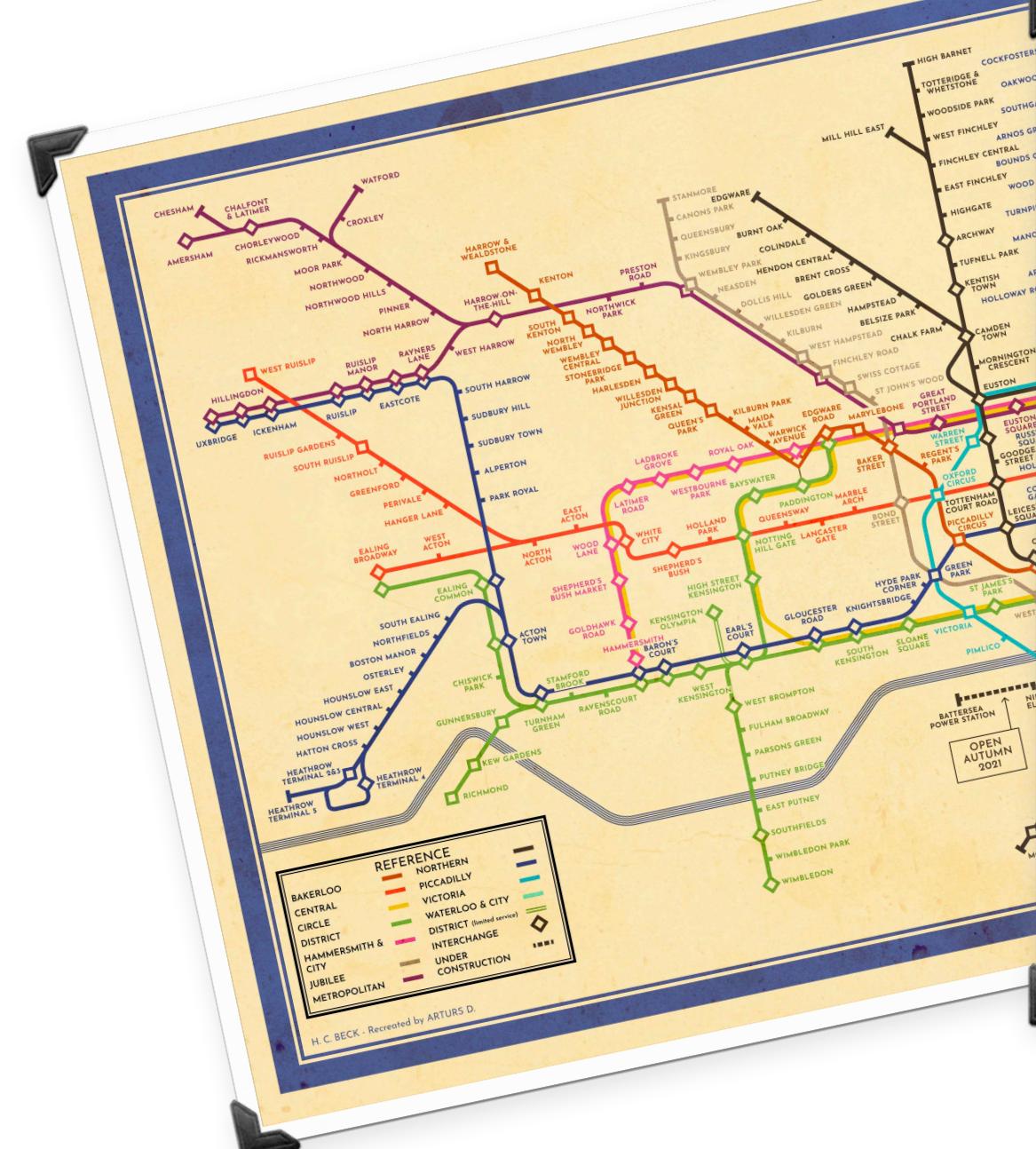
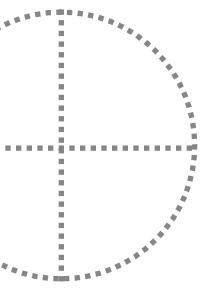
“Walensee und Seetal” • Eduard Imhof



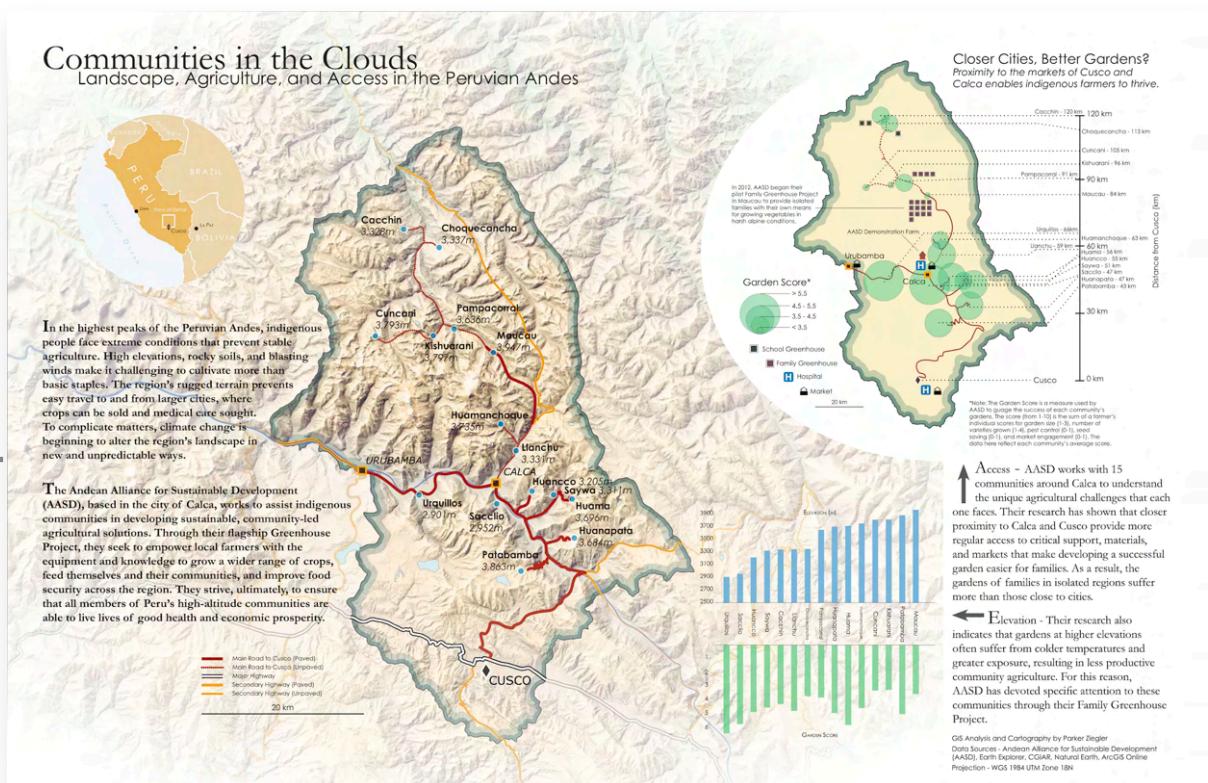
Perspective

“Wrangell Mountains Shaded Relief Map” • Tom Patterson



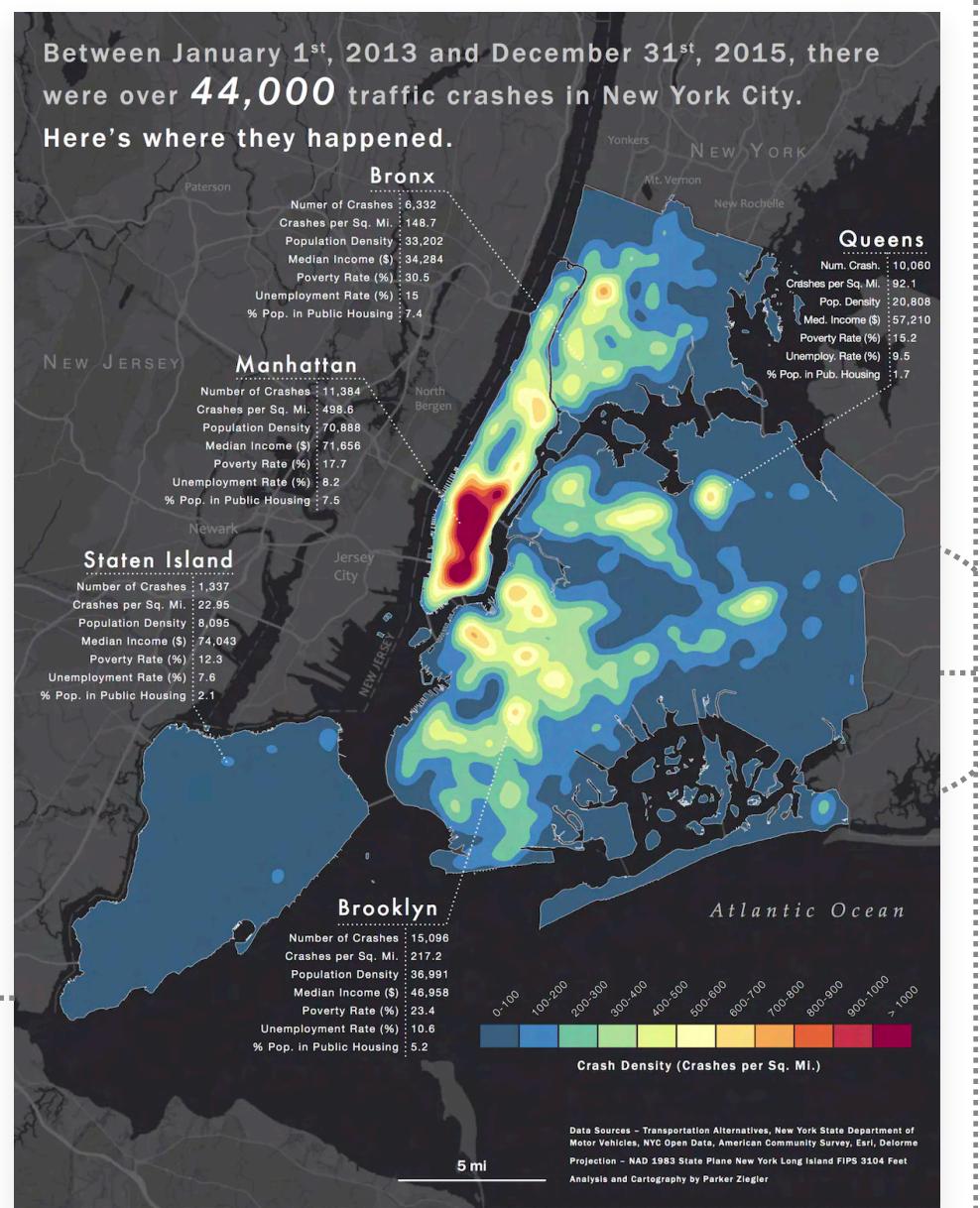
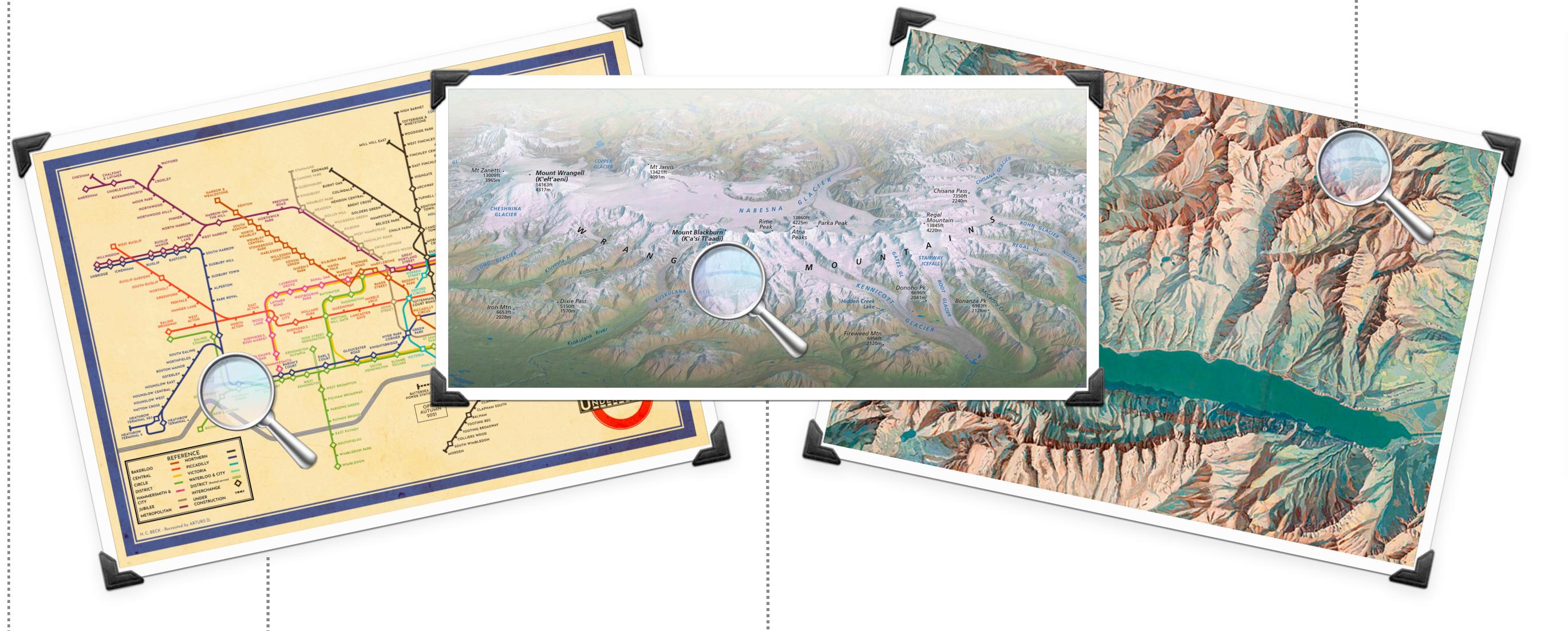


Color Palettes

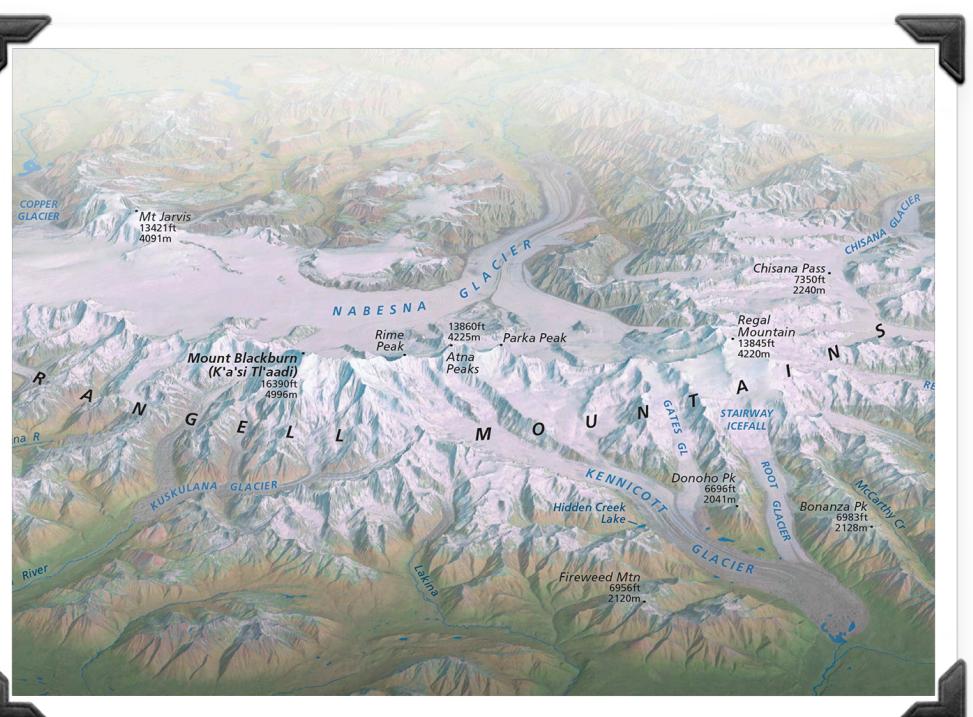


Typography

Visual Hierarchy

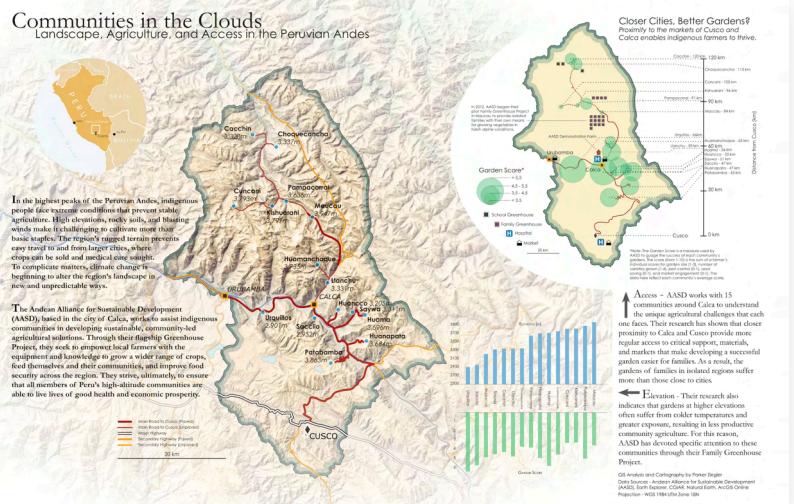


Examples



Many, Many (Many) Hours of Painstaking Work

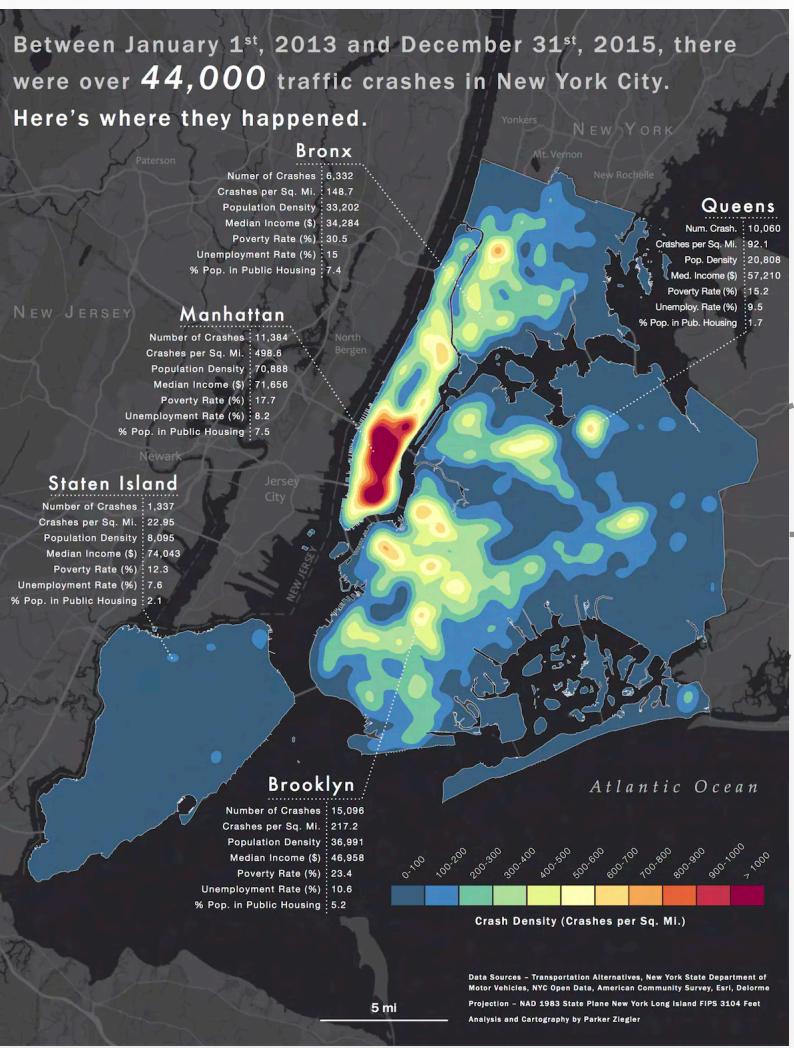
Outputs



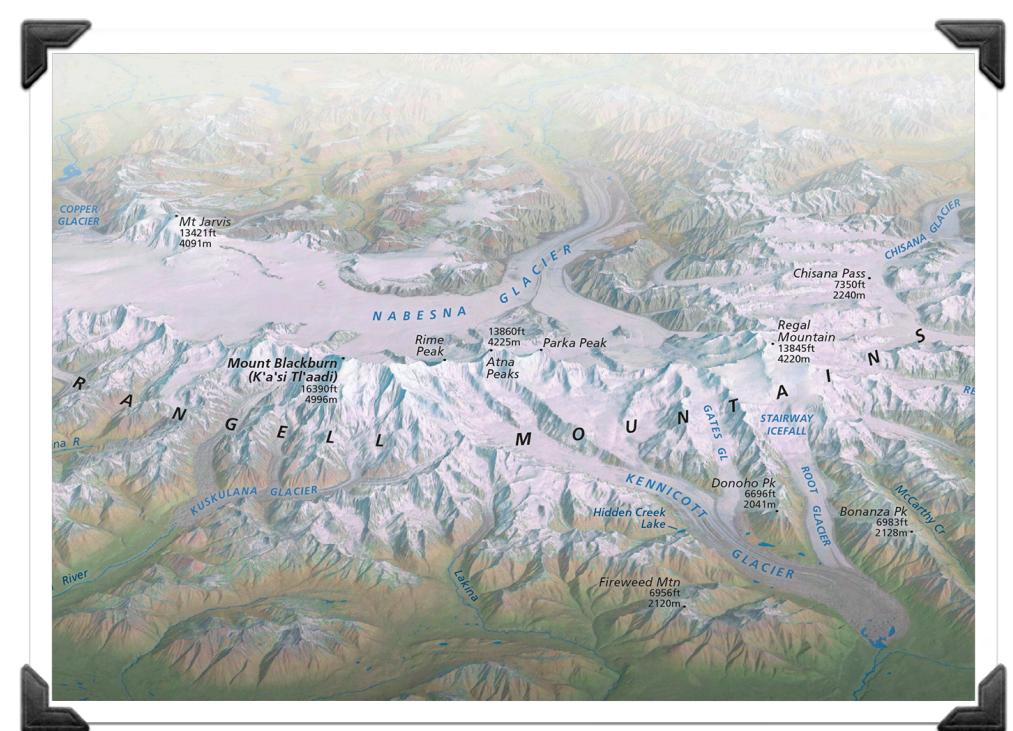
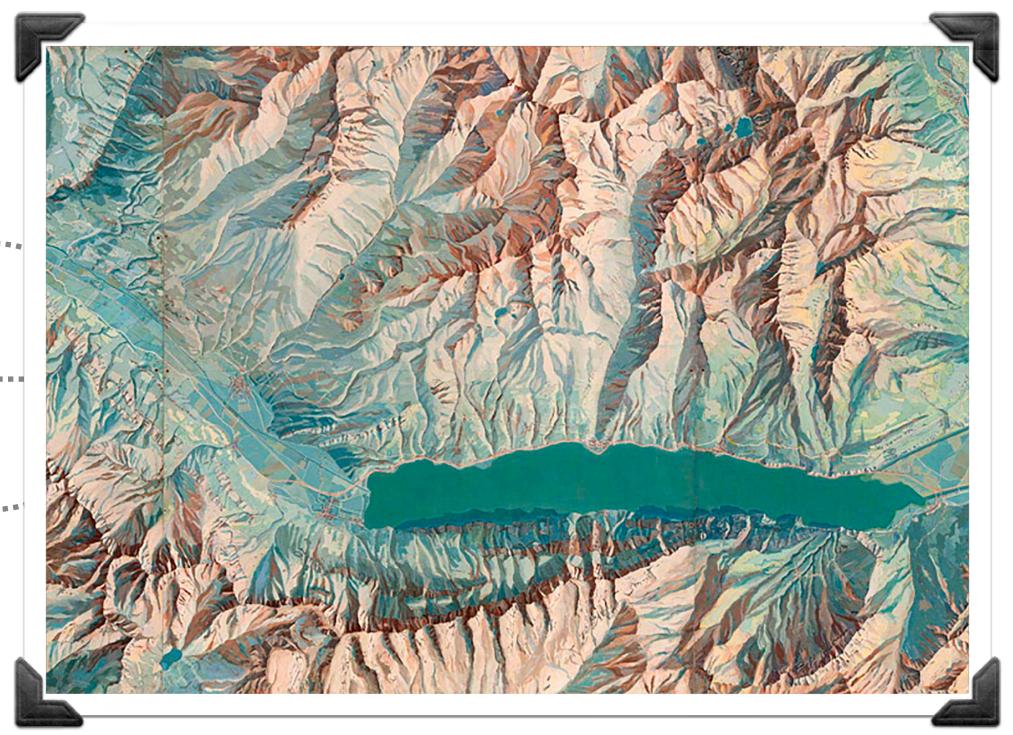
Unseen, Unheard, and Forgotten
Child Migration from Central America to the United States



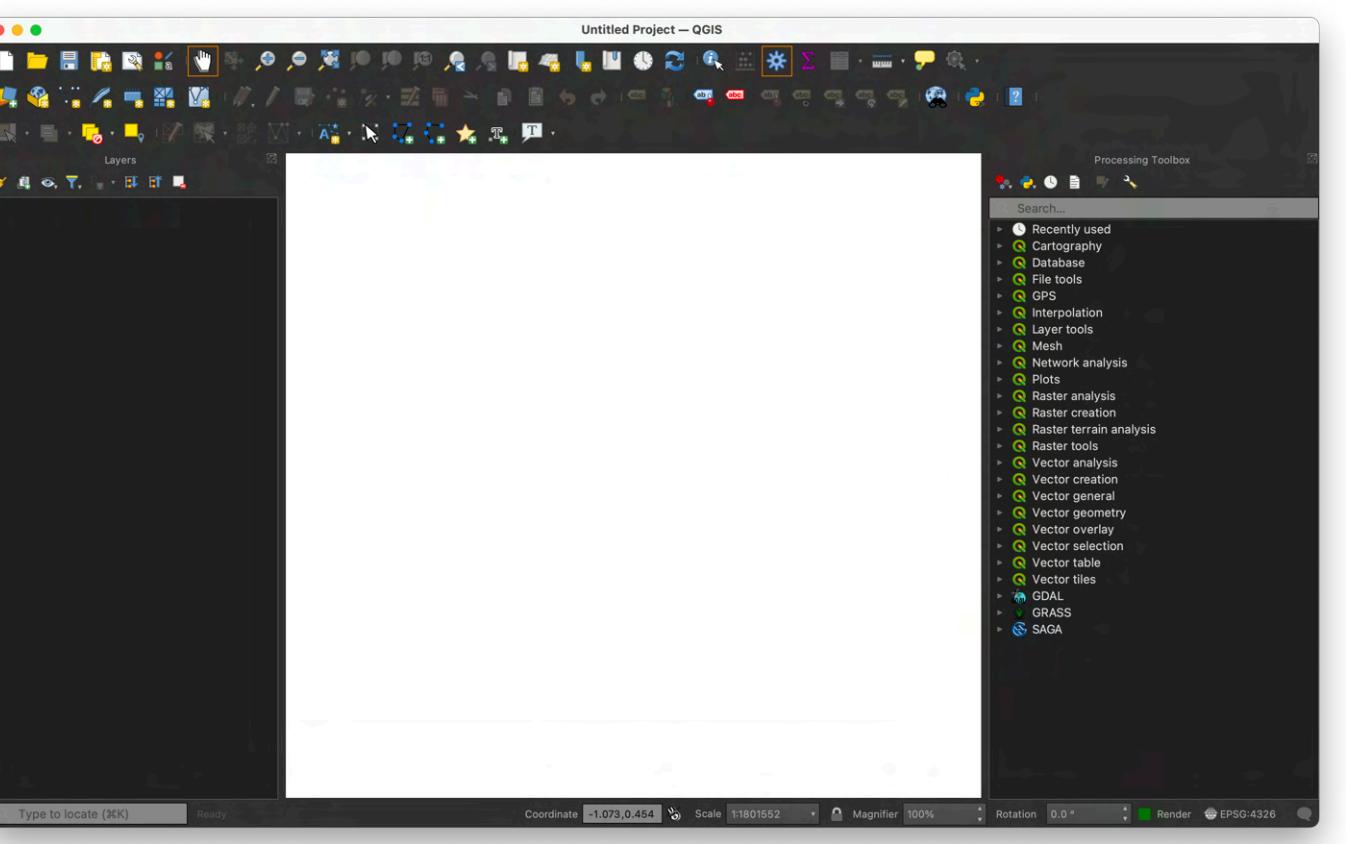
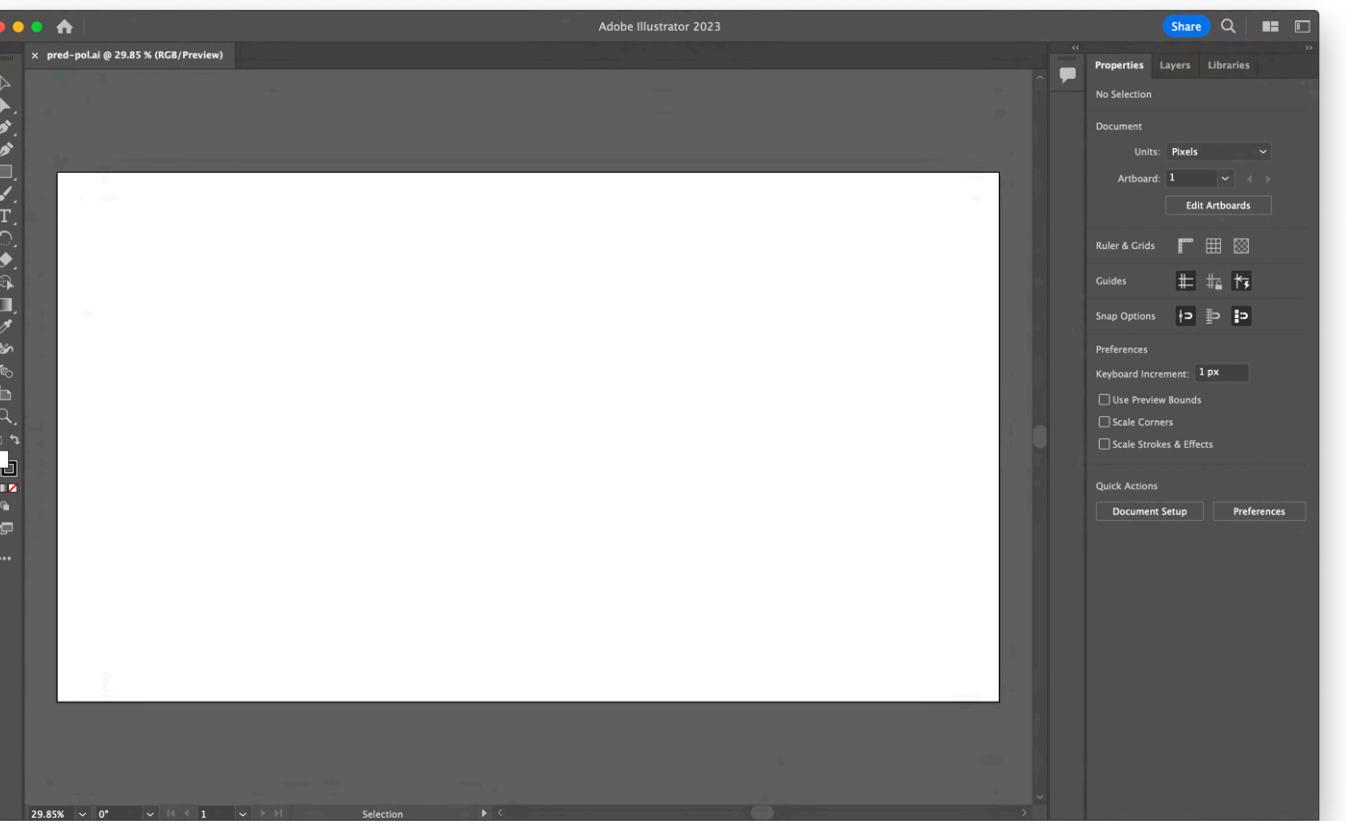
Between January 1st, 2013 and December 31st, 2015, there were over 44,000 traffic crashes in New York City.
Here's where they happened.



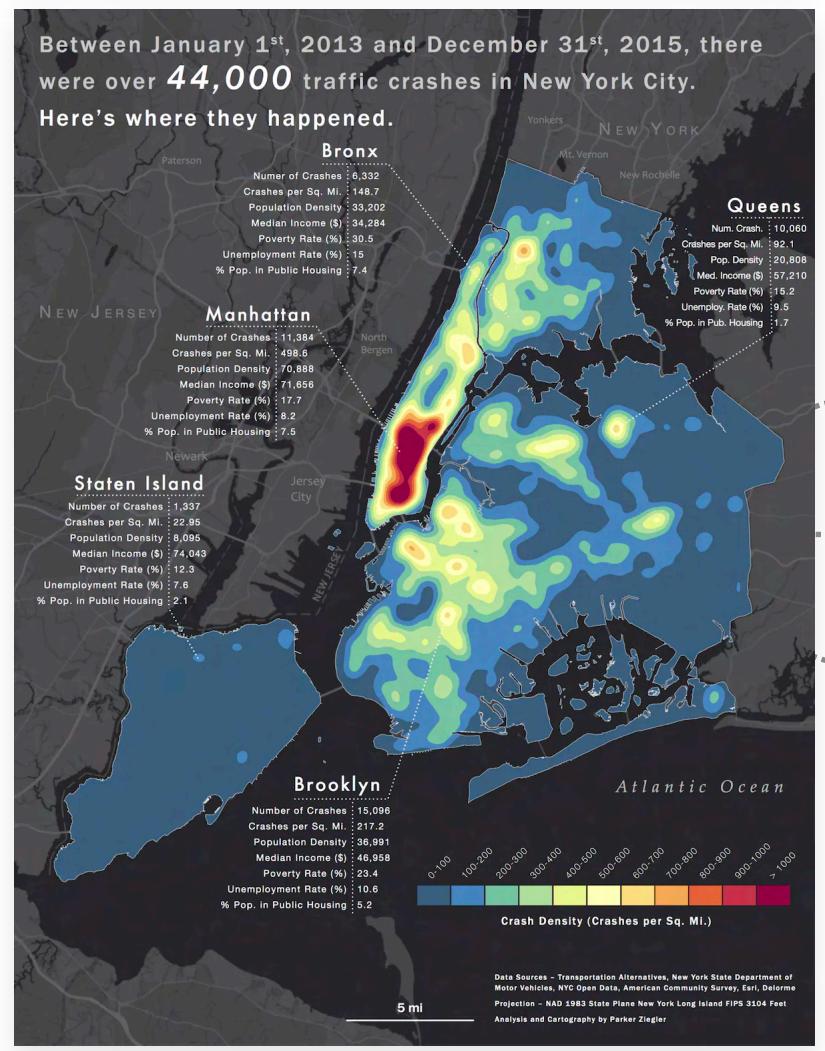
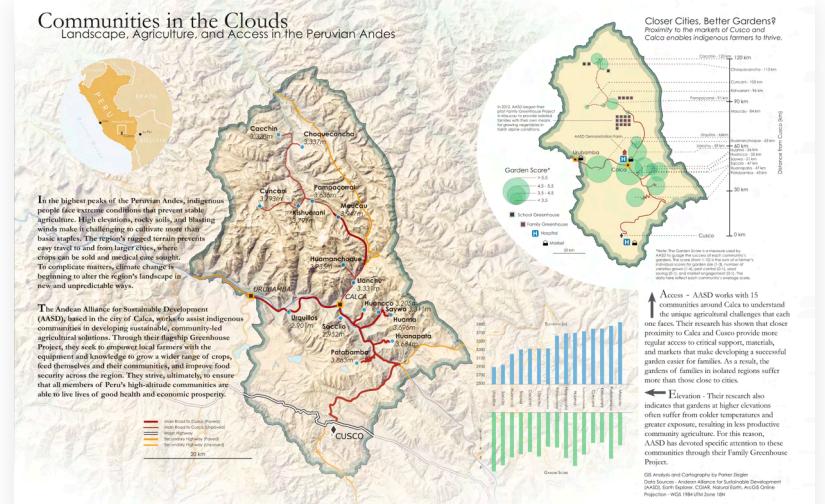
Examples



Many, Many (Many) Hours of Painstaking Work

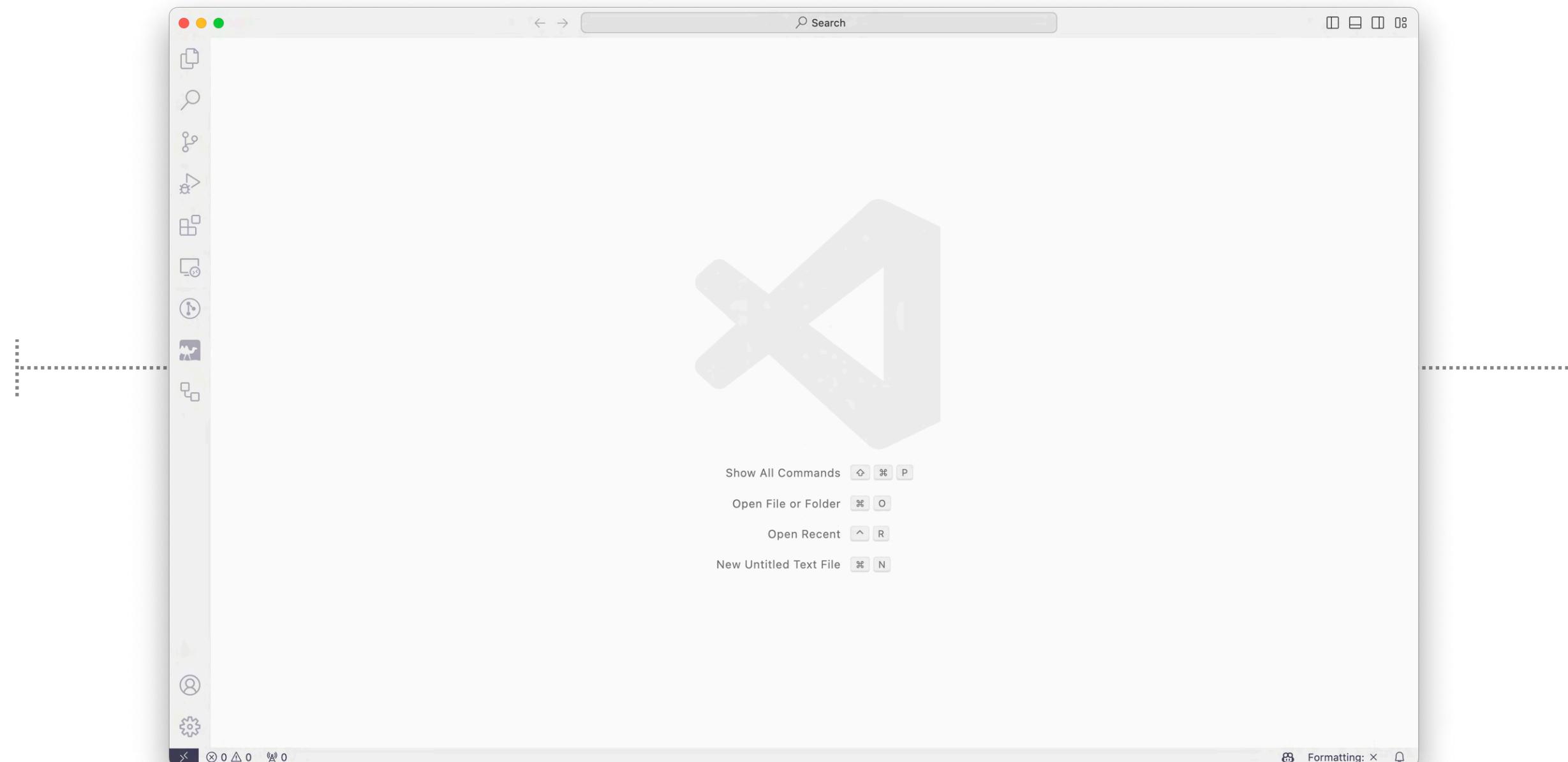
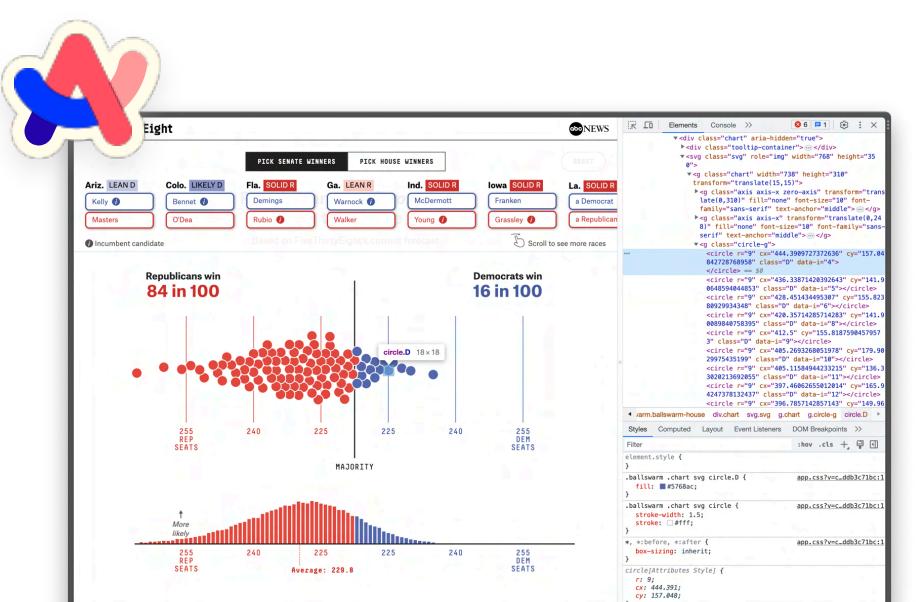
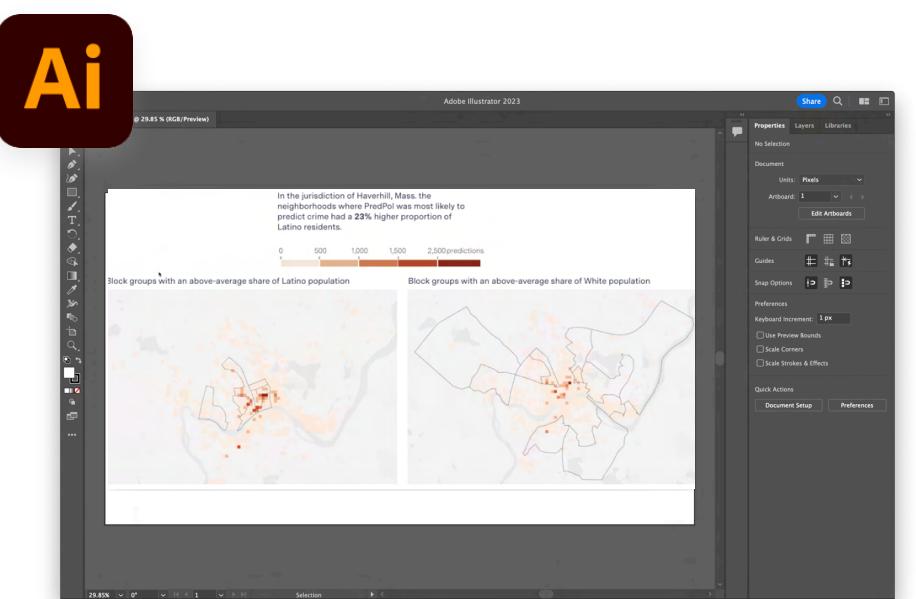
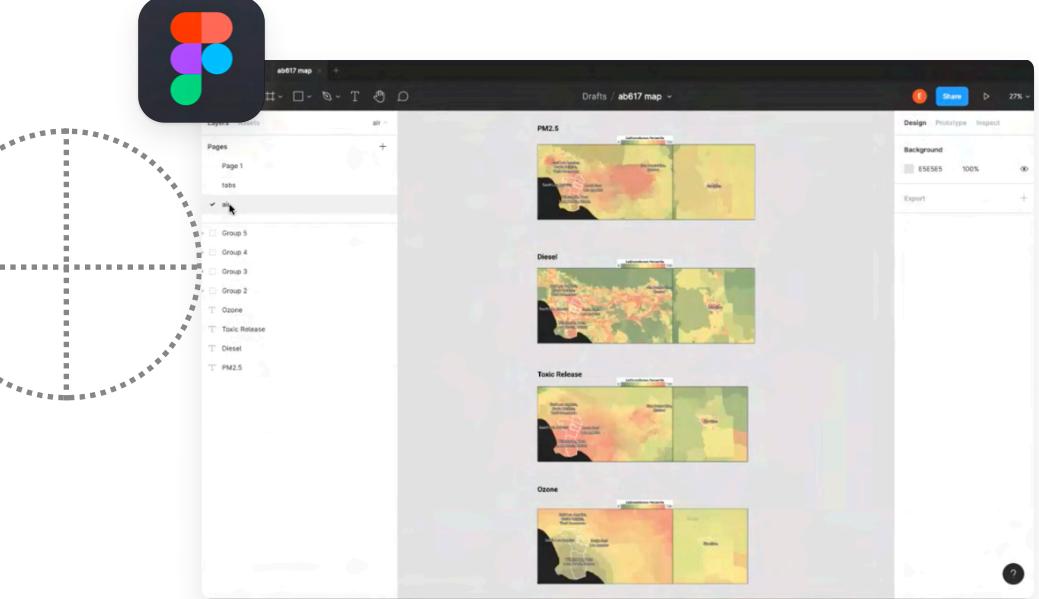


Outputs



Visual Inputs

Many, Many (Many) Hours of Painstaking Work



Programs

```
import * as d3 from "d3";
const context = canvas.getContext("2d");
regionsGeo.features.forEach(feature) => {
  context.beginPath();
  path(feature);
}
const c = color(
  props.years_2080_2099
);
```

```
import mapboxgl from "mapbox-gl";
import cancerRegions from "./canc-...";
mapboxgl.accessToken = "pk.eyJ1...";

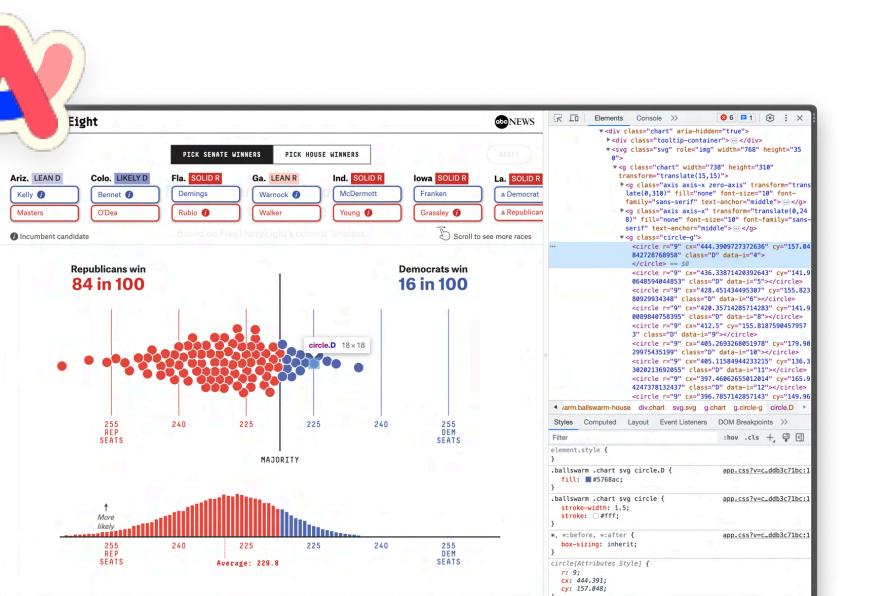
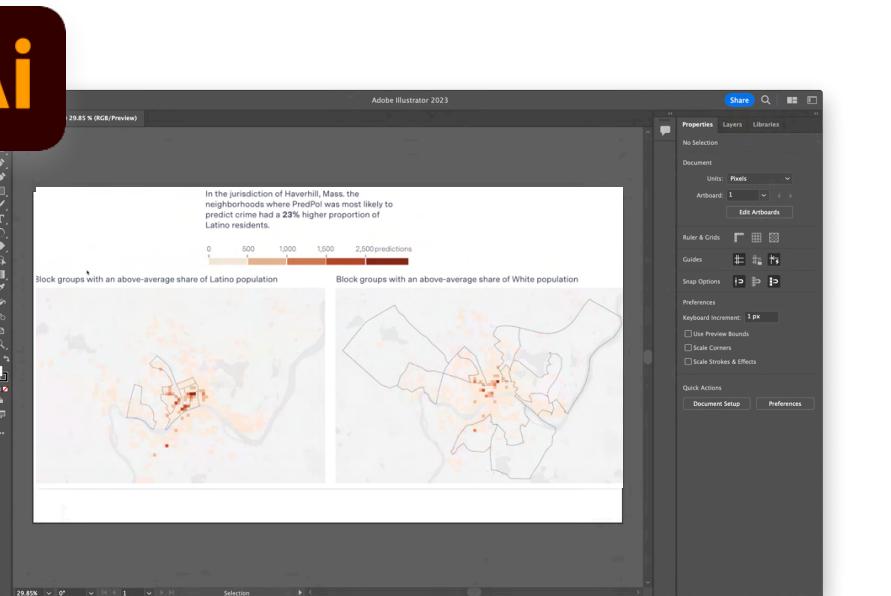
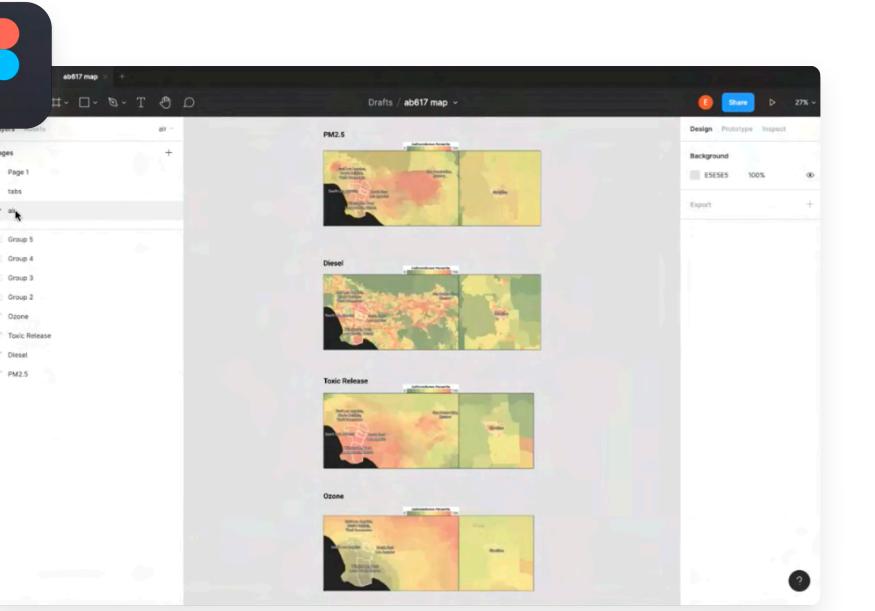
const map = new Map({
  container: "map",
  style: "https://tiles.stadiamaps....",
  center: [-106.1086, 37.7531],
  zoom: 4,
});

map.on("load", () => {
  map.addSource("cancer-regions", {
    type: "geojson",
    data: cancerRegions,
  });
});
```

```
import * as Plot from "@obs...";

const plot = Plot.plot({
  y: { grid: true },
  marks: [
    Plot.line(covidDeaths,
      Plot.binX(
        { y: "sum" },
        { x: "date" },
      )
    ),
    Plot.ruleY([0])
  ],
});
```

Visual Inputs



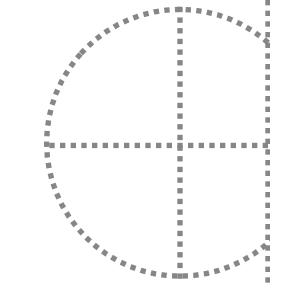
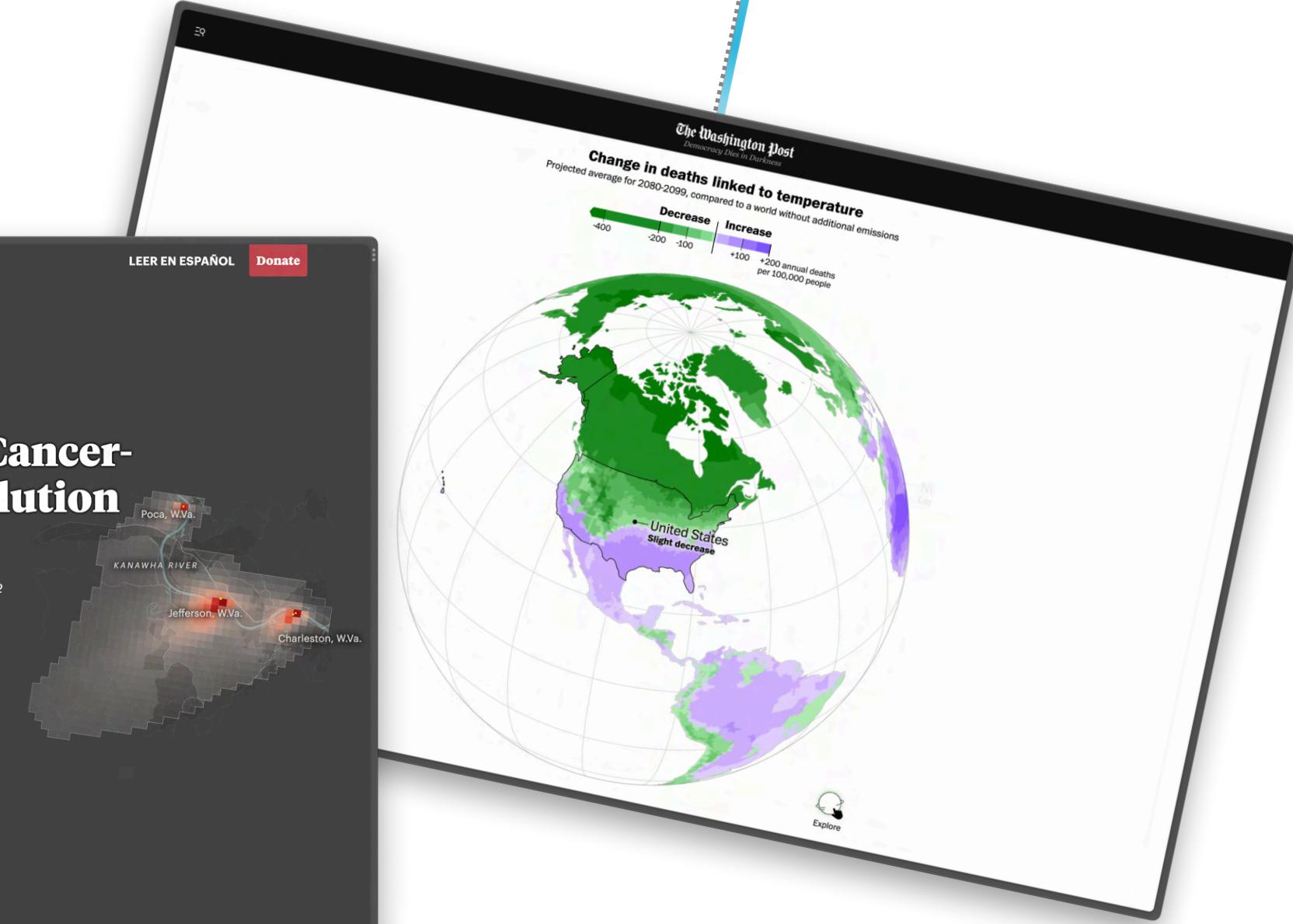
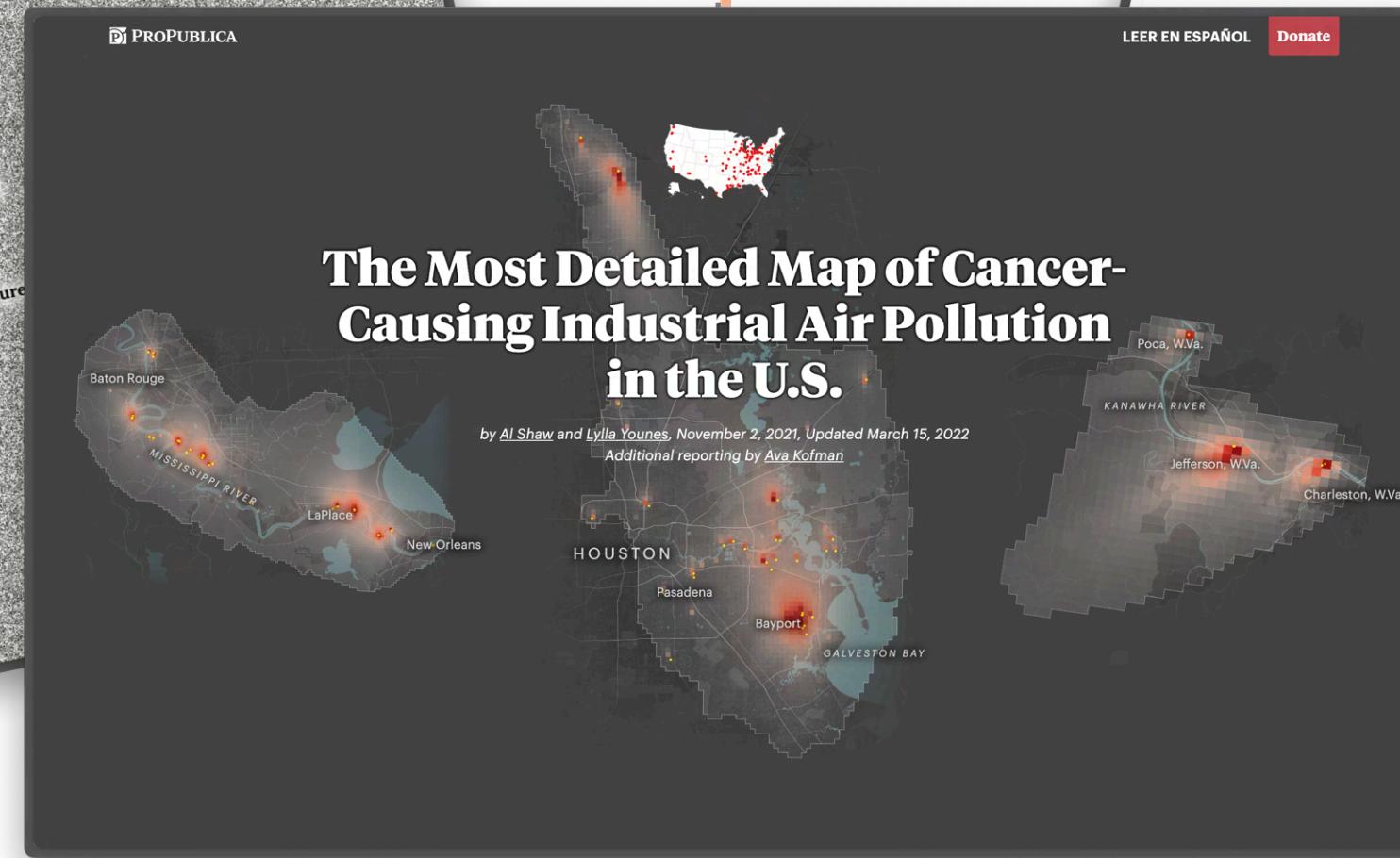
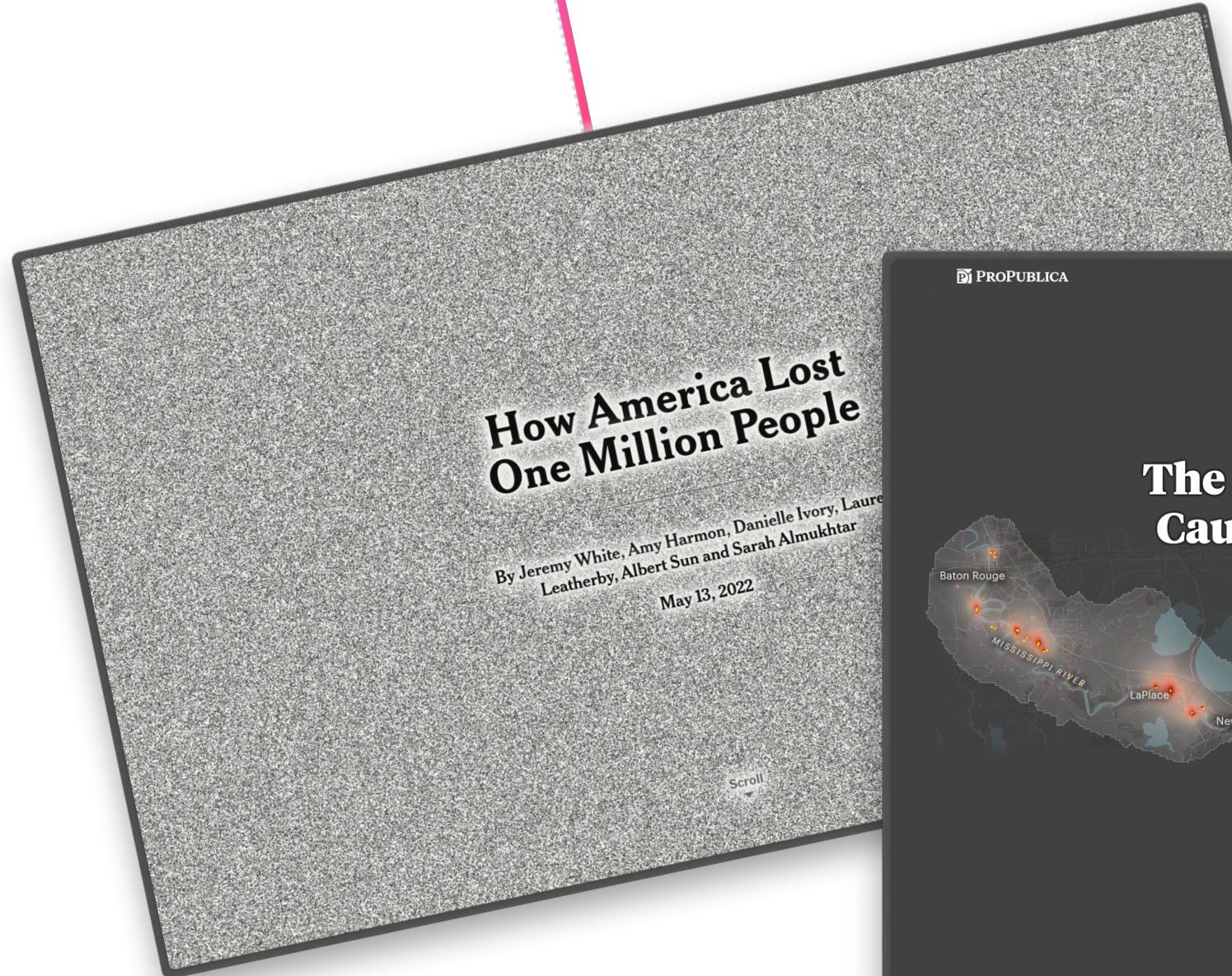
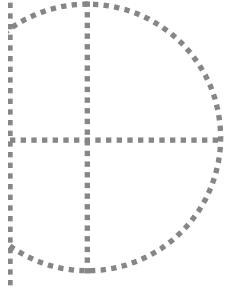
Programs

```
import * as d3 from "d3";
const context = canvas.getContext("2d");
regionsGeo.features.forEach(feature) => {
  context.beginPath();
  path(feature);
  const c = color(
    props.years_2080_2099
  );
};
```

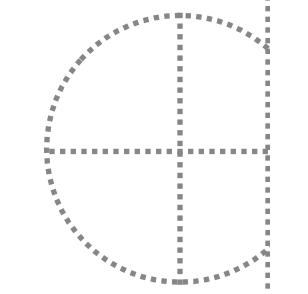
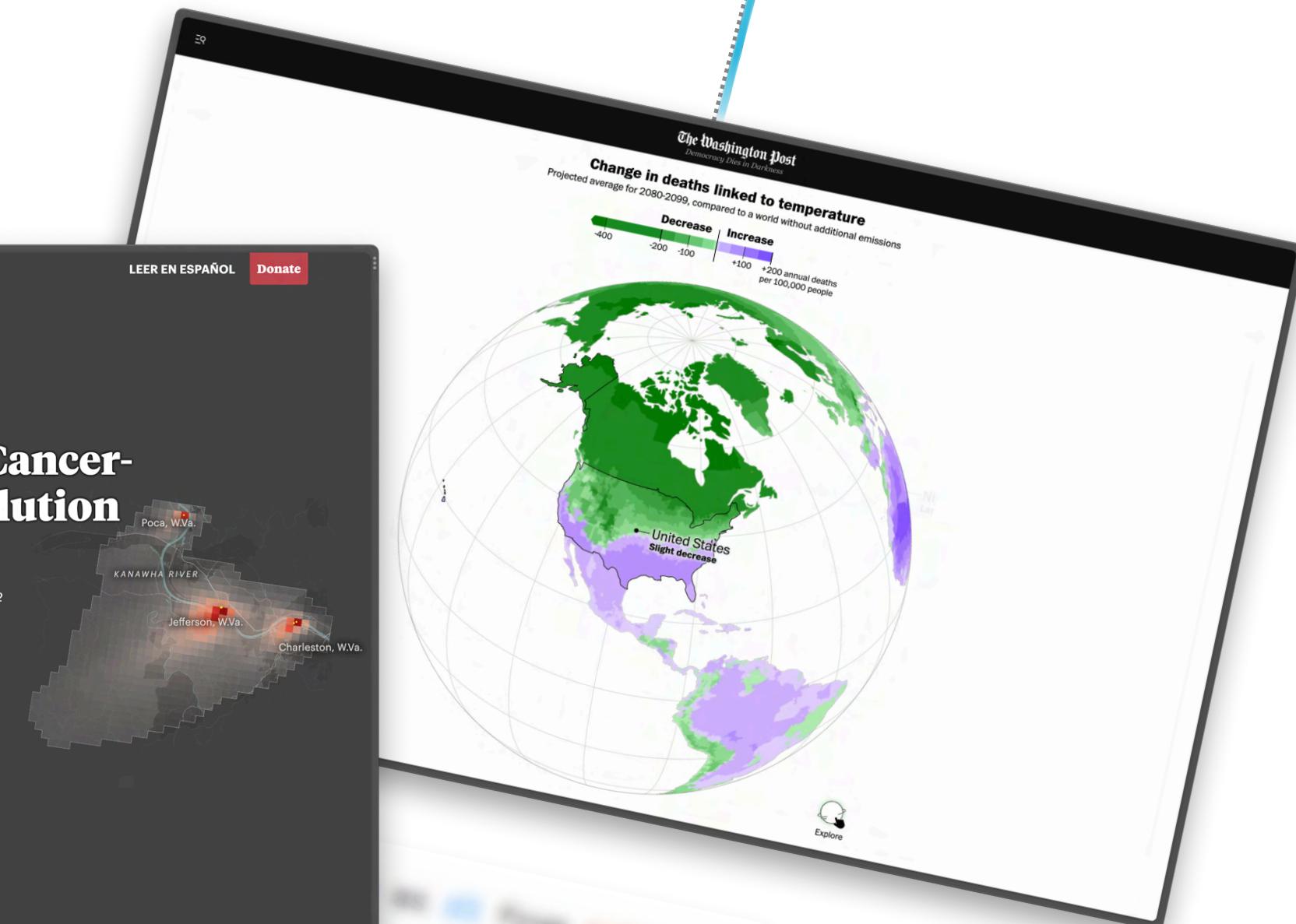
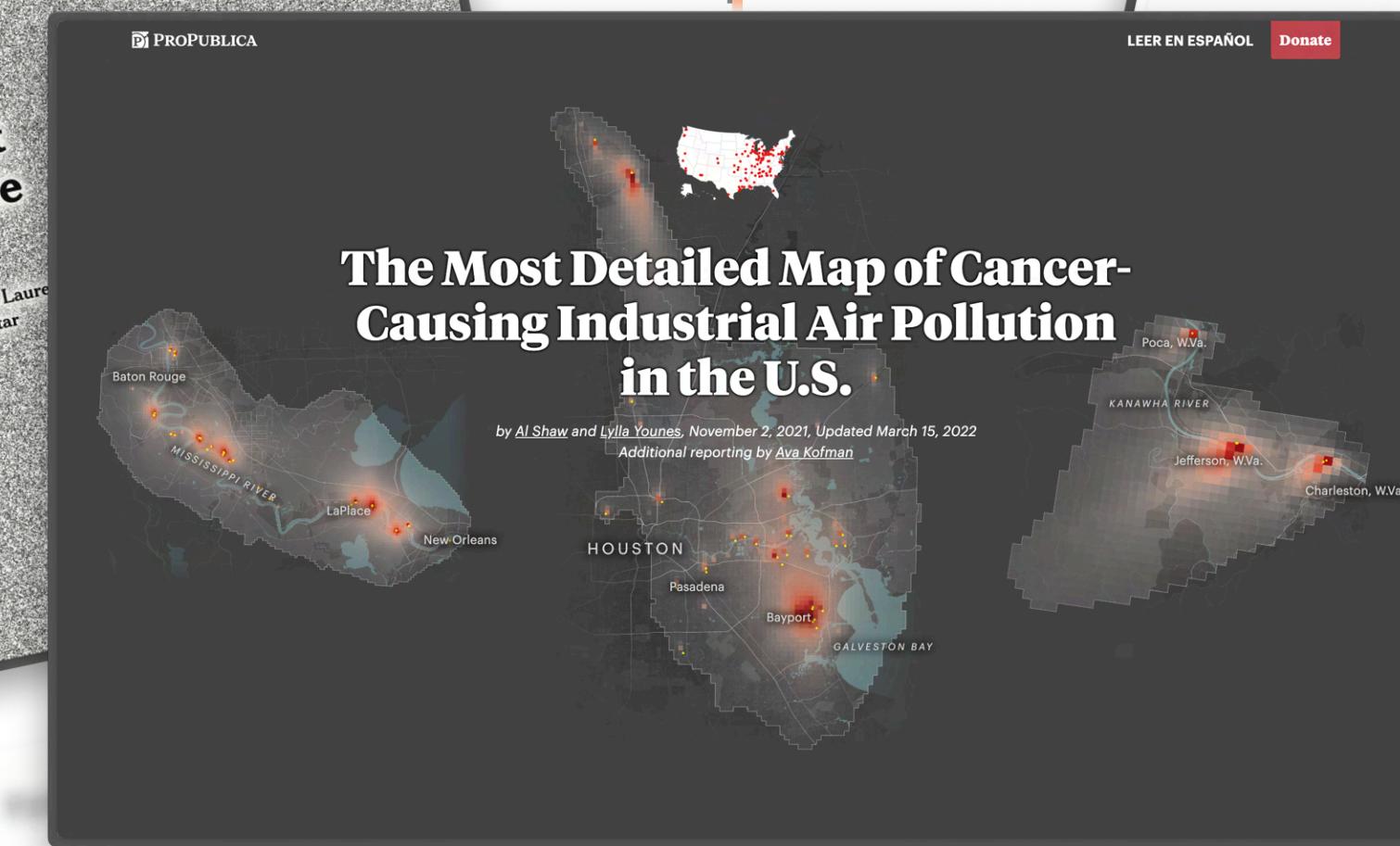
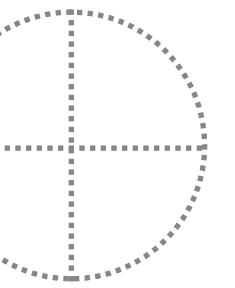
```
import mapboxgl from "mapbox-gl";
import cancerRegions from "./cancer-regions";
mapboxgl.accessToken = "pk.eyJ1Ij...";
const map = new Map({
  container: "map",
  style: "https://tiles.stadiamaps.com/tiles/carto-dark-10m.v1.json",
  center: [-106.1086, 37.7531],
  zoom: 4,
});
map.on("load", () => {
  map.addSource("cancer-regions", {
    type: "geojson",
    data: cancerRegions,
  });
});
```

```
import * as Plot from "@observablehq/plot";
const plot = Plot.plot({
  y: { grid: true },
  marks: [
    Plot.line(covidDeaths,
      Plot.binX(
        { y: "sum" },
        { x: "date" },
      )
    ),
    Plot.ruleY([0])
  ],
});
```

Fixed Entities



Fixed Entities



Live Objects

The collage illustrates various live data visualization and analysis tools:

- Top Left:** A screenshot of a web browser showing a map titled "The Most Detailed Map of Cancer-Causing Industrial Pollution in the U.S." by ProPublica. The map uses a color-coded legend to represent pollution levels across the United States.
- Top Center:** A screenshot of a browser developer tools' Elements tab, showing the DOM structure of a page with some code snippets visible.
- Bottom Left:** A screenshot of a data visualization tool (possibly Plotly or similar) showing a scatter plot of US states based on vaccination rates and political leanings. The plot includes a color legend and a text overlay about vaccine hesitancy.
- Bottom Center:** A screenshot of a data visualization tool (possibly Tableau or similar) showing a scatter plot of US states. The plot includes a color legend and a text overlay about vaccine hesitancy.
- Right:** A screenshot of a map application (possibly Mapbox) showing temperature change polygons across the Washington Post area. It includes a sidebar with properties for the selected layer and a large amount of JSON code for the map's configuration.

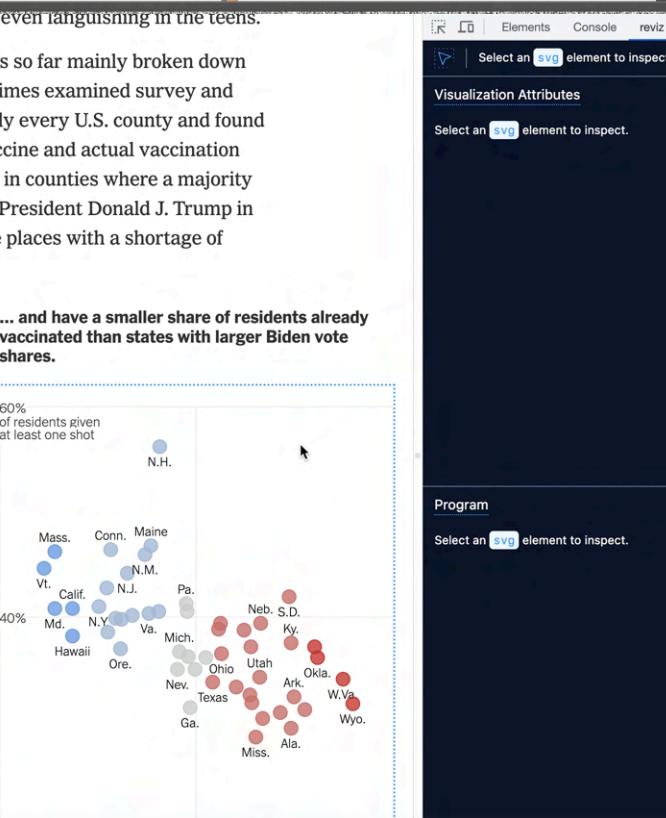
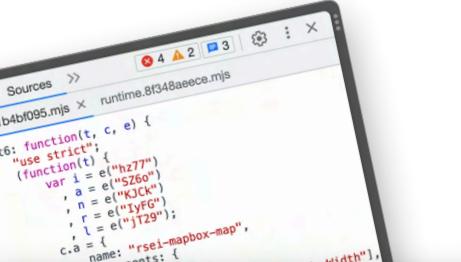
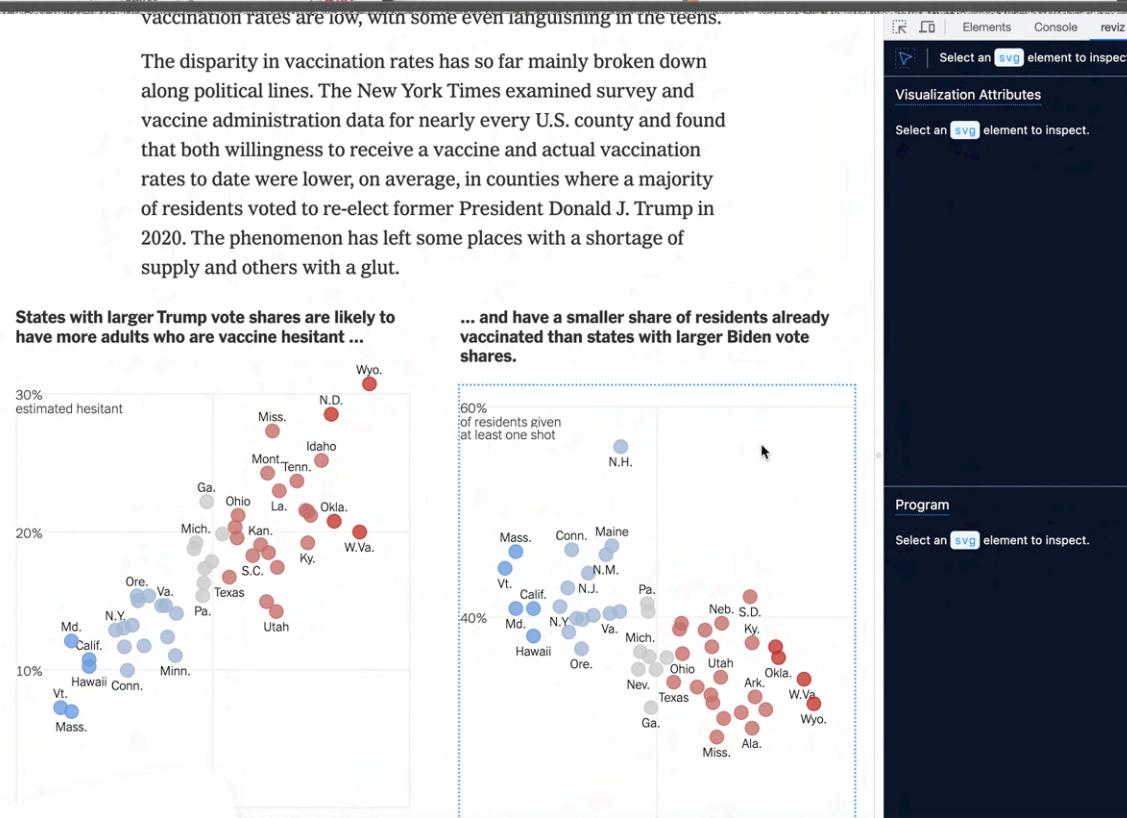
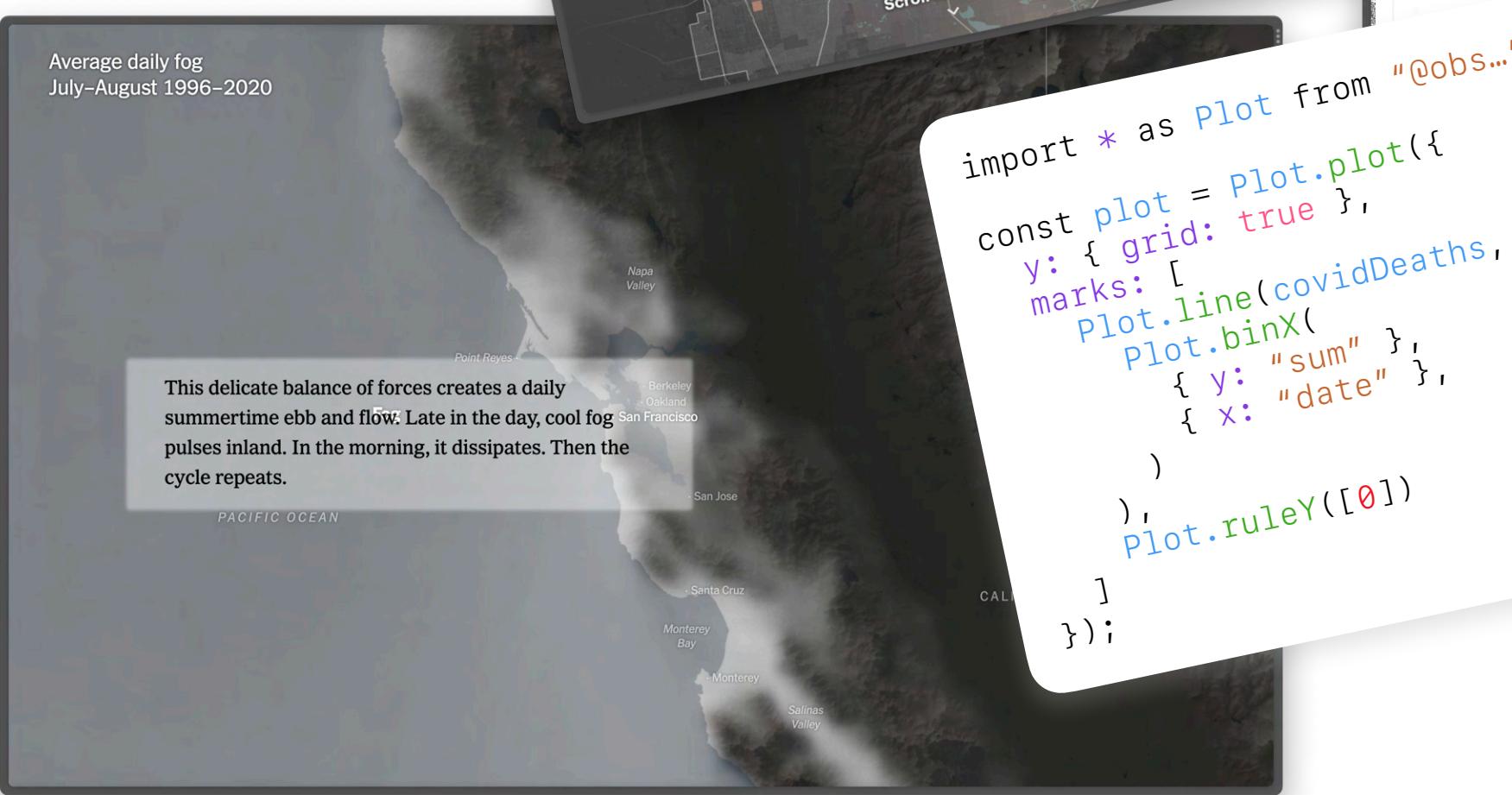
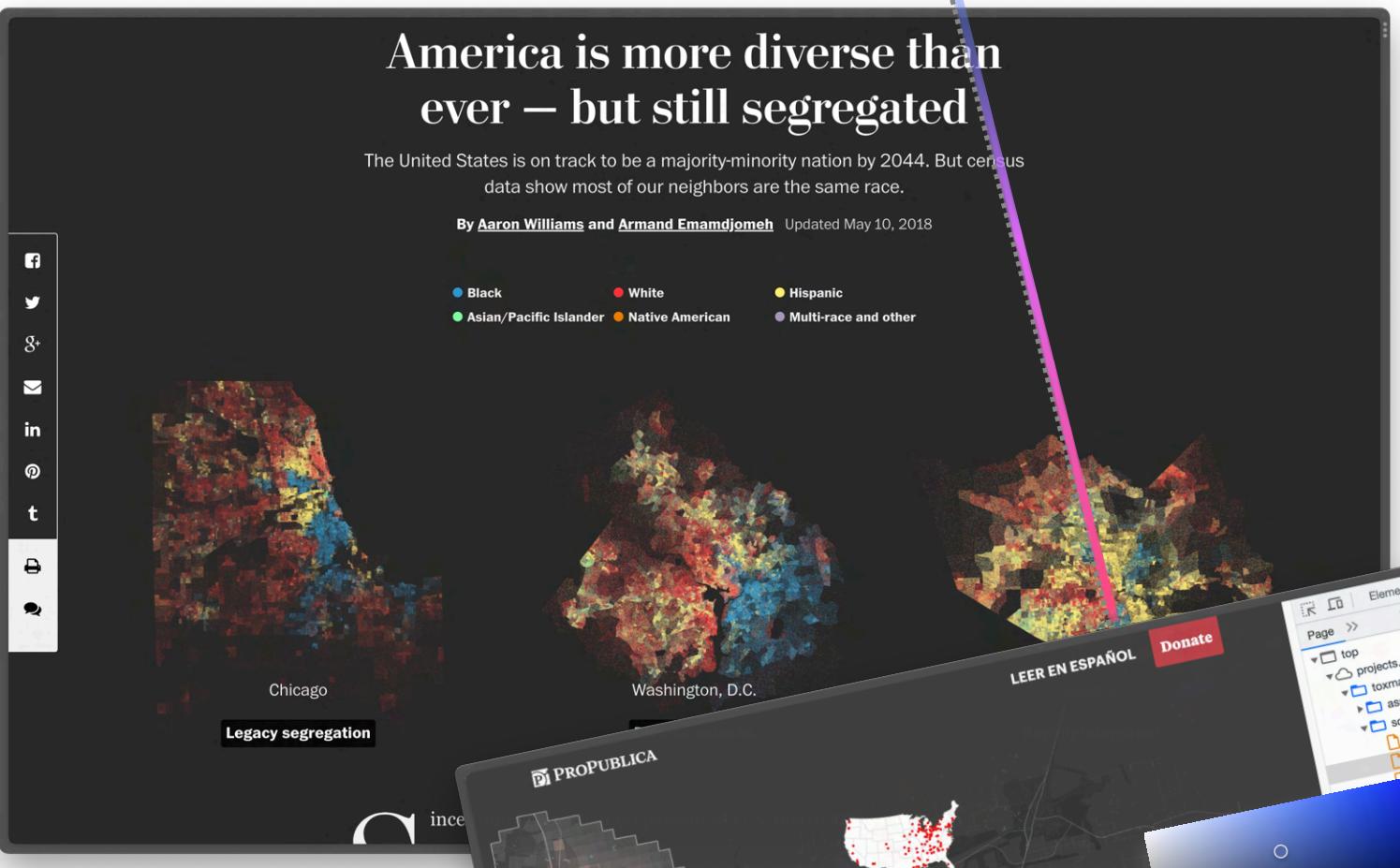
Code Examples:

```
import * as Plot from "@obs...";  
const plot = Plot.plot({  
  y: { grid: true },  
  marks: [  
    Plot.line(covidDeaths,  
    Plot.binX(  
      { y: "sum" },  
      { x: "date" }  
    ),  
    Plot.ruleY([0])  
  ]  
});
```

```
import mapboxgl from "mapbox-gl";  
import cancerRegions from "./canc...";  
  
mapboxgl.accessToken = "pk.eyJ1IjoiZmFyZWJhZG9tIiwiYSI6ImN...";  
  
const map = new Map({  
  container: "map",  
  style: "https://tiles.stadiamaps.com/styles/alidade_smooth.json",  
  center: [-106.1086, 37.7531],  
  zoom: 4,  
});  
  
map.on("load", () => {  
  map.addSource("cancer-regions", {  
    type: "geojson",  
    data: cancerRegions,  
  });  
});
```

```
import * as d3 from "d3";  
const context = canvas.getContext("2d");  
regionsGeo.features.forEach(feature) => {  
  context.beginPath();  
  path(feature);  
}  
const c = color(  
  props.years_2080_2099  
);
```

Live Objects

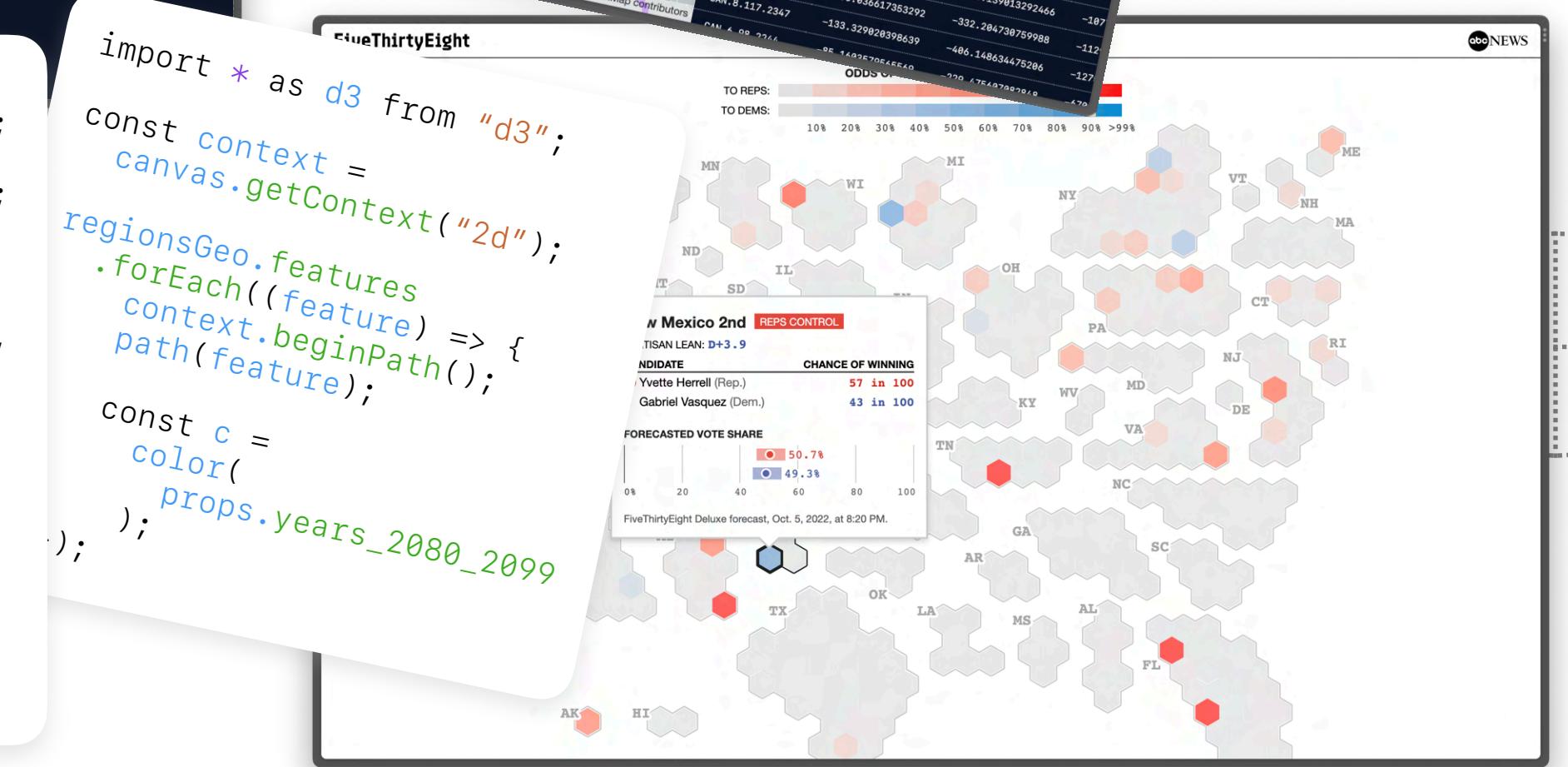
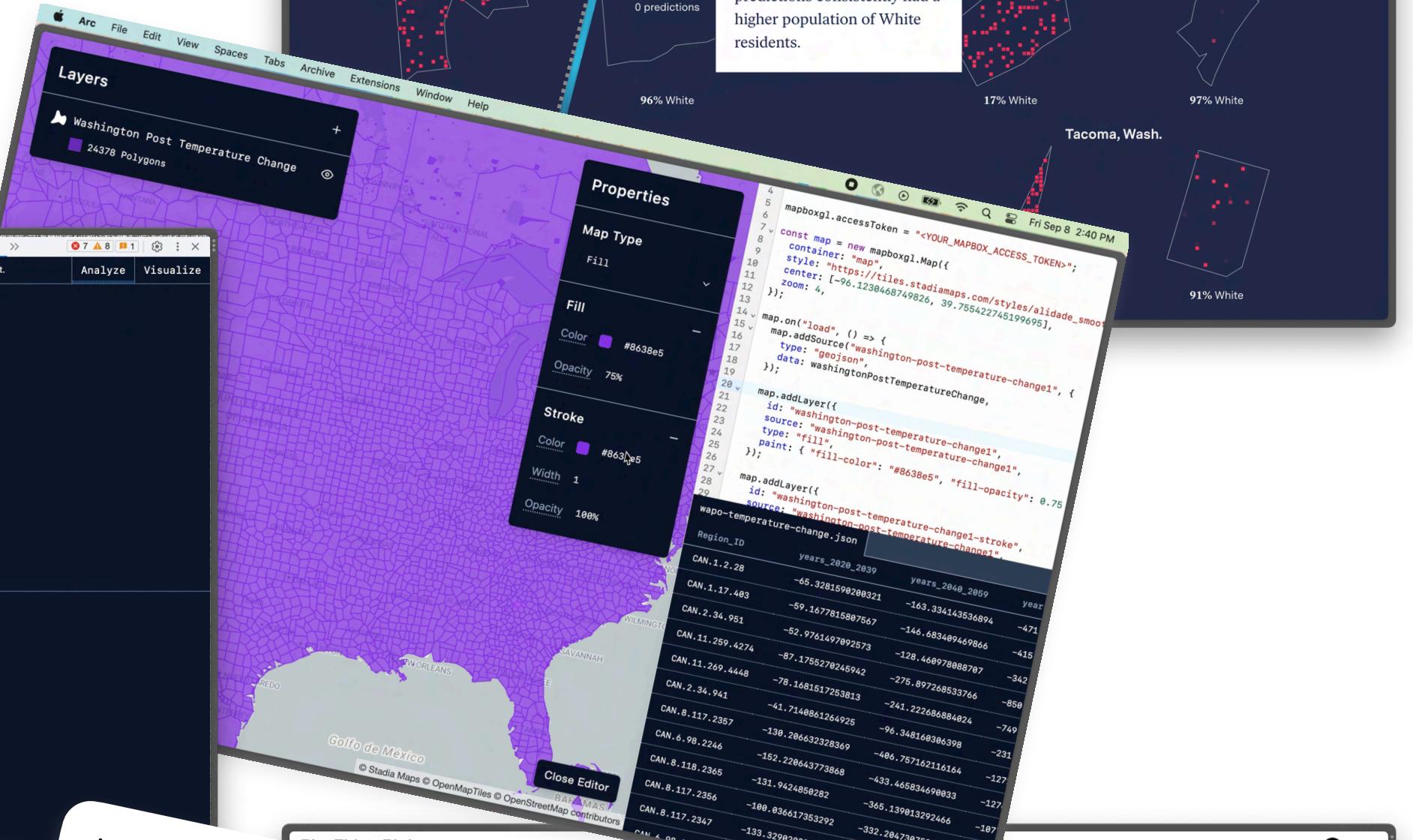


```

import mapboxgl from "mapbox-gl";
import cancerRegions from "./canc...";
mapboxgl.accessToken = "pk.eyJ1IjoiZmFyZWJhZGJhIiwiZWZpZWQiOiJkMzVlNzEwNzUwMzI4In0.eyJ6fh2...";

const map = new Map({
  container: "map",
  style: "https://tiles.stadiamaps...",
  center: [-106.1086, 37.7531],
  zoom: 4,
});

map.on("load", () => {
  map.addSource("cancer-regions", {
    type: "geojson",
    data: cancerRegions,
  });
});
  
```



Supporting Data Journalism through Compilers for Visual Inputs



github.com/parkerziegler/reviz

observablehq.com/@parkerziegler/hello-reviz



github.com/parkerziegler/cartokit

alpha.cartokit.dev

Parker Ziegler // [@parker_ziegler](https://twitter.com/parker_ziegler)

Ph.D. Student, UC Berkeley

Get in touch!

peziegler@cs.berkeley.edu

parkie-doo.sh

Strange Loop
St. Louis, MO • September 21, 2023