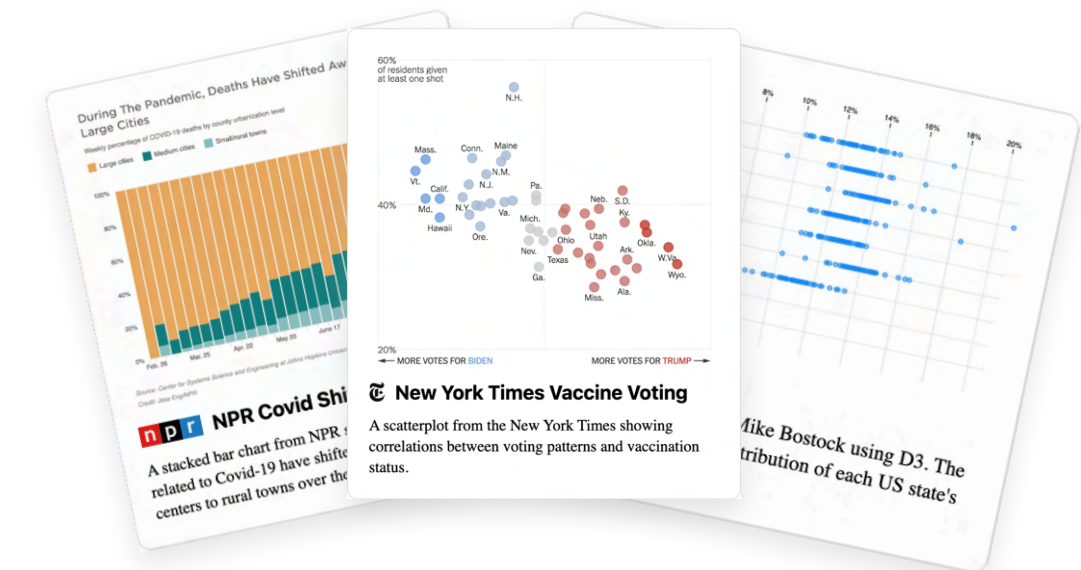


Supporting Data Journalism through Compilers for Visual Inputs

```
<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="369.5" width="7.8" height="0.5" fill="#4e79a7">
    </rect>
    <rect x="49.3" y="369.4" width="7.8" height="0.5" fill="#4e79a7">
    </rect>
    ...
    <rect x="155.5" y="112" width="7.8" height="44.9" fill="#f28e2c">
    </rect>
  </g>
</svg>
```

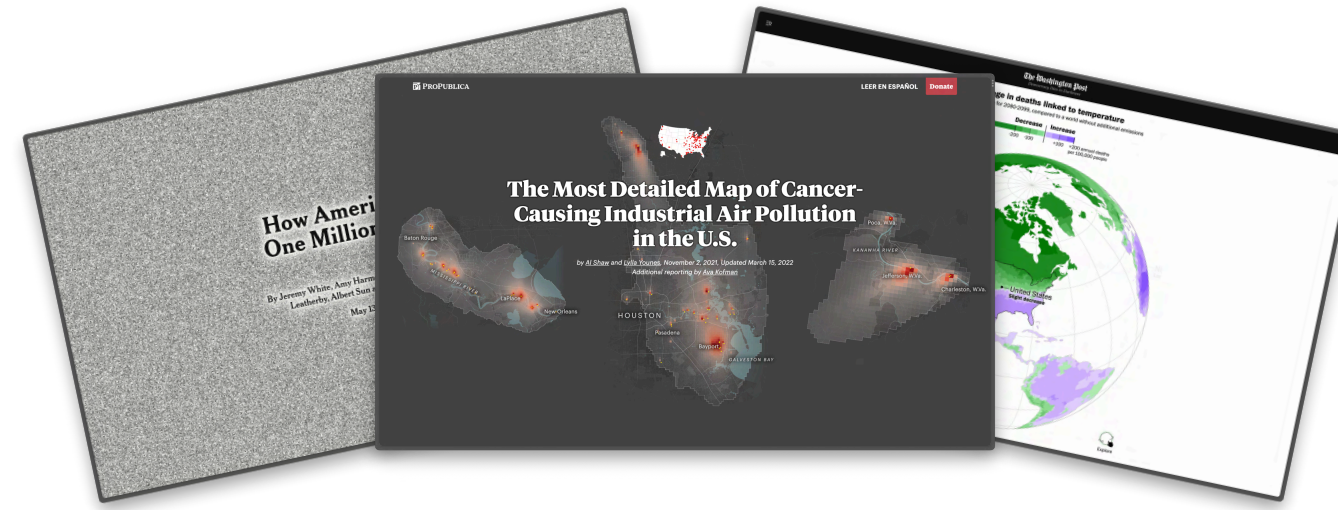


```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#4e79a7", "#f28e2c"]
  },
  marks: [
    Plot.barY(
      data,
      Plot.binX(
        { y: "count" },
        {
          x: "???",
          fill: "???",
          fillOpacity: 1,
          stroke: "none",
          strokeOpacity: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



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**?

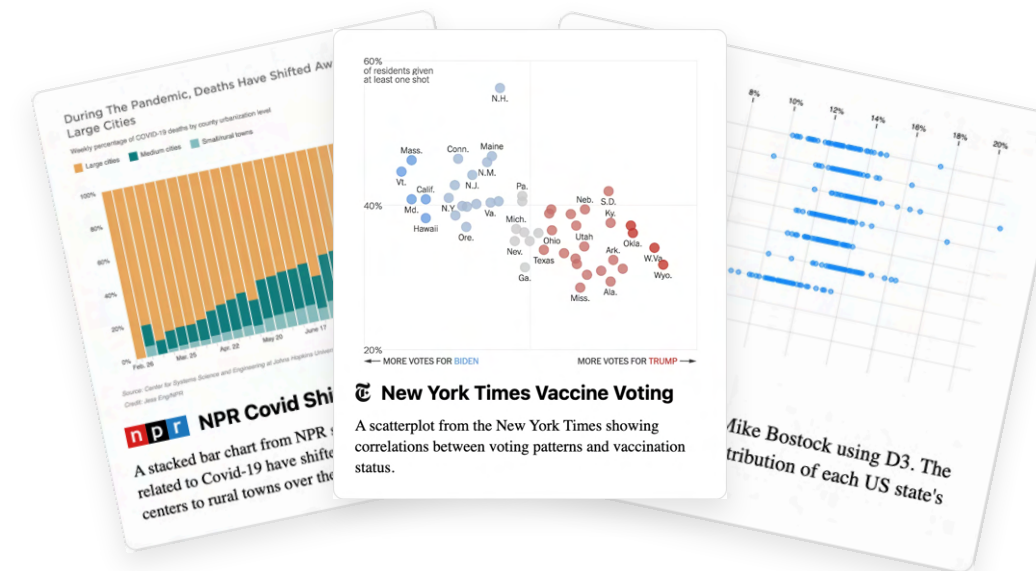
What does this **even mean**?

Compilers

Visual Inputs

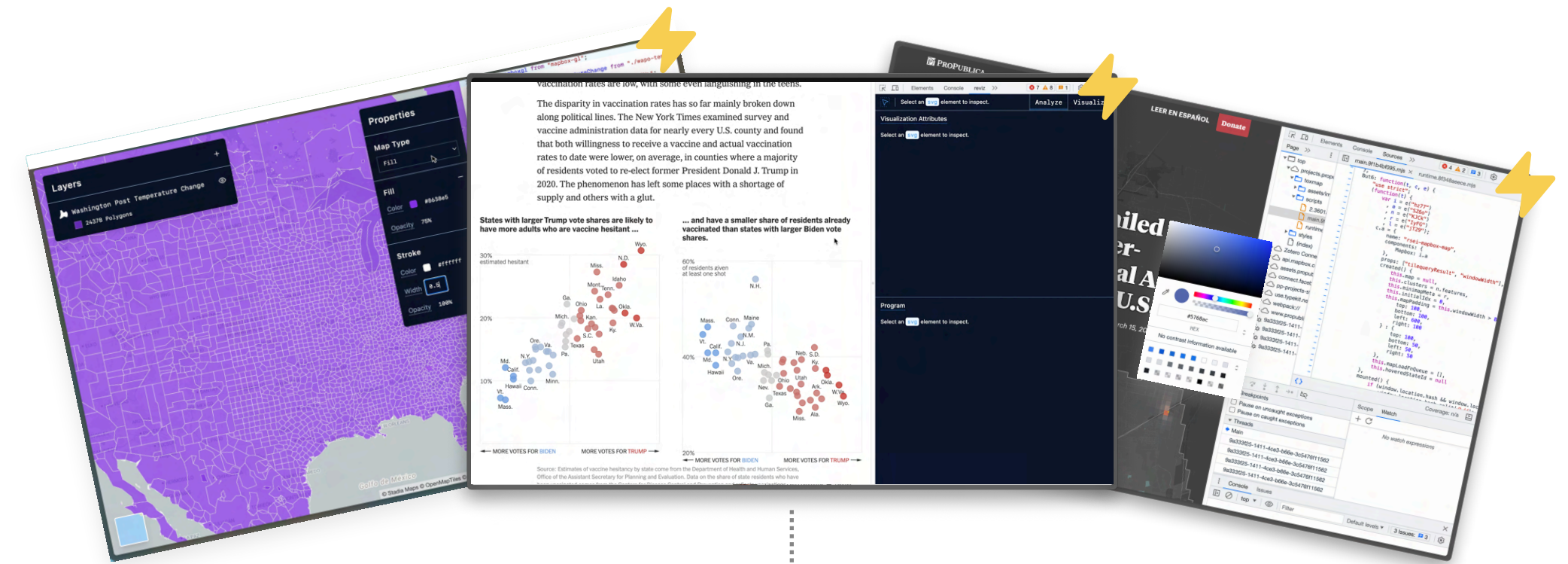
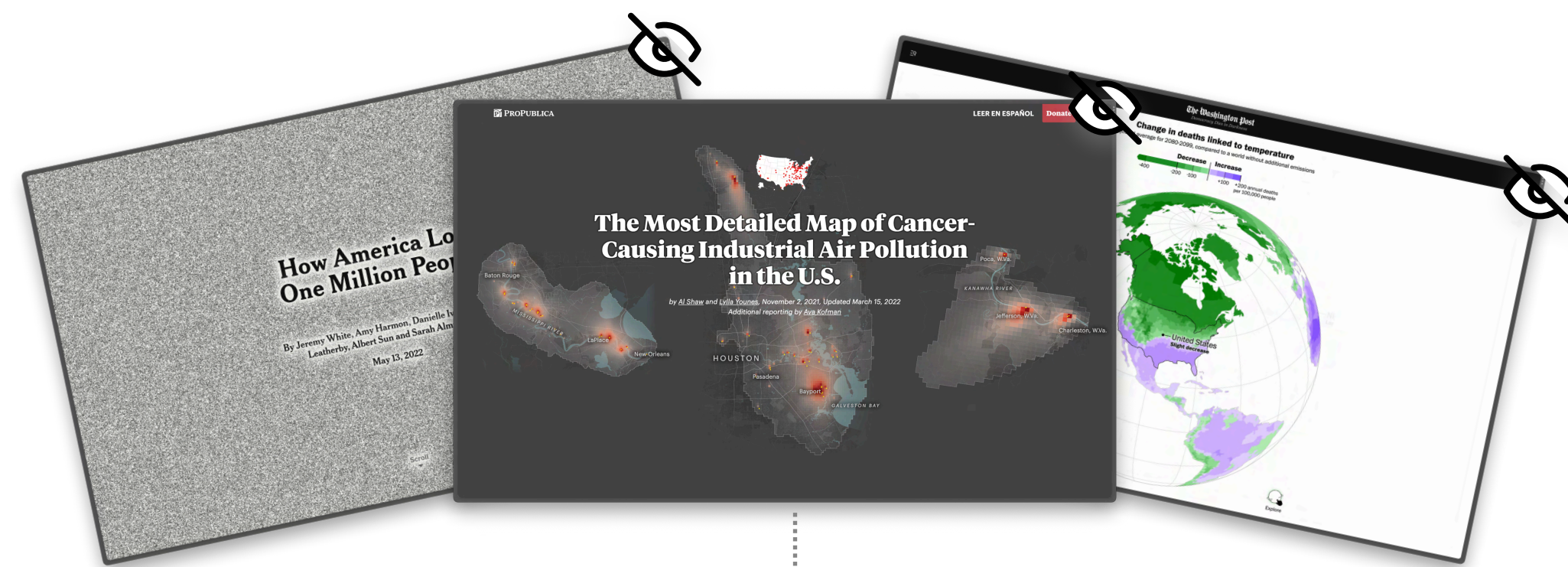
```
<svg width="688" height="400" viewBox="0 0 688 400">
  <g aria-label="y-axis tick" fill="none" stroke="currentColor">
    <path transform="translate(40,370)" d="M0,0L-6,0"></path>
    <path transform="translate(40,342.6)" d="M0,0L-6,0"></path>
  </g>
  <g aria-label="rect">
    <rect x="40.5" y="369.5" width="7.8" height="0.5" fill="#4e79a7">
    </rect>
    <rect x="49.3" y="369.4" width="7.8" height="0.5" fill="#4e79a7">
    </rect>
    <rect x="155.5" y="112" width="7.8" height="44.9" fill="#f28e2c">
    </rect>
  </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
        }
      )
    )
  ]
});
```

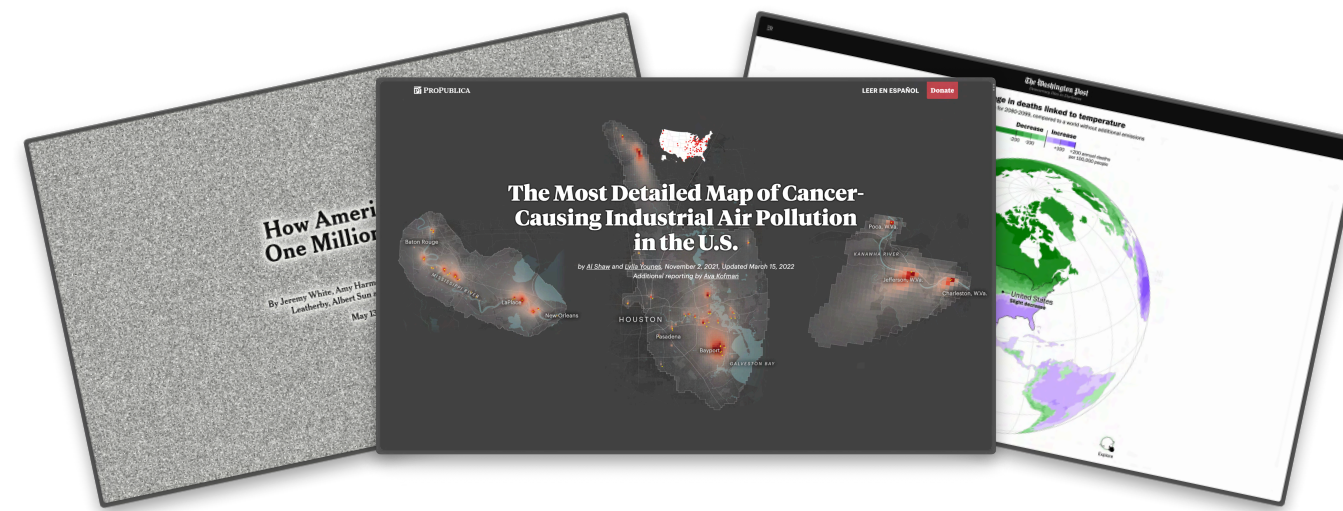


Fixed Entities

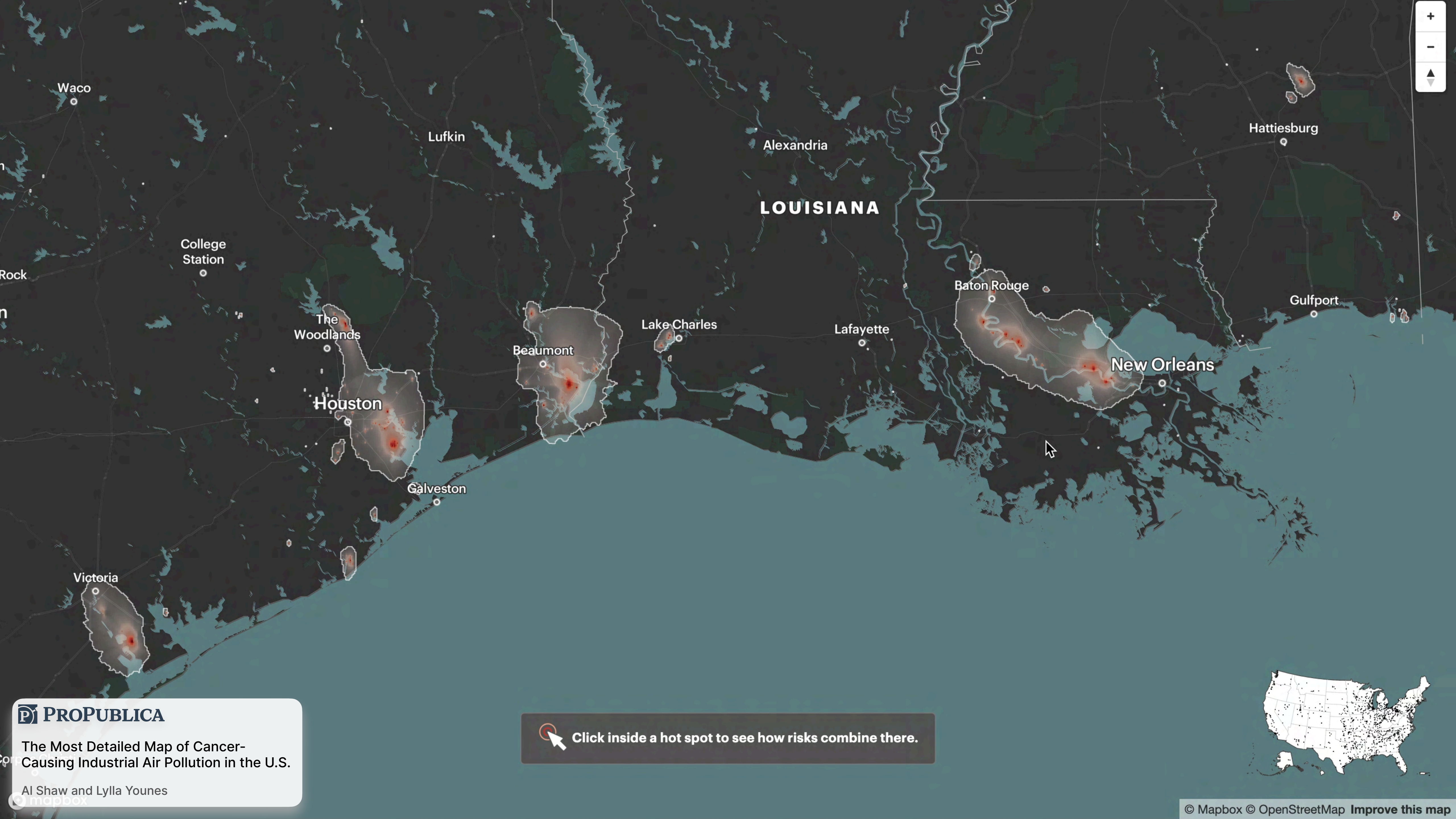
Live Objects



Compilers



Data Journalism




PROPUBLICA

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

Al Shaw and Lylla Younes



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



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

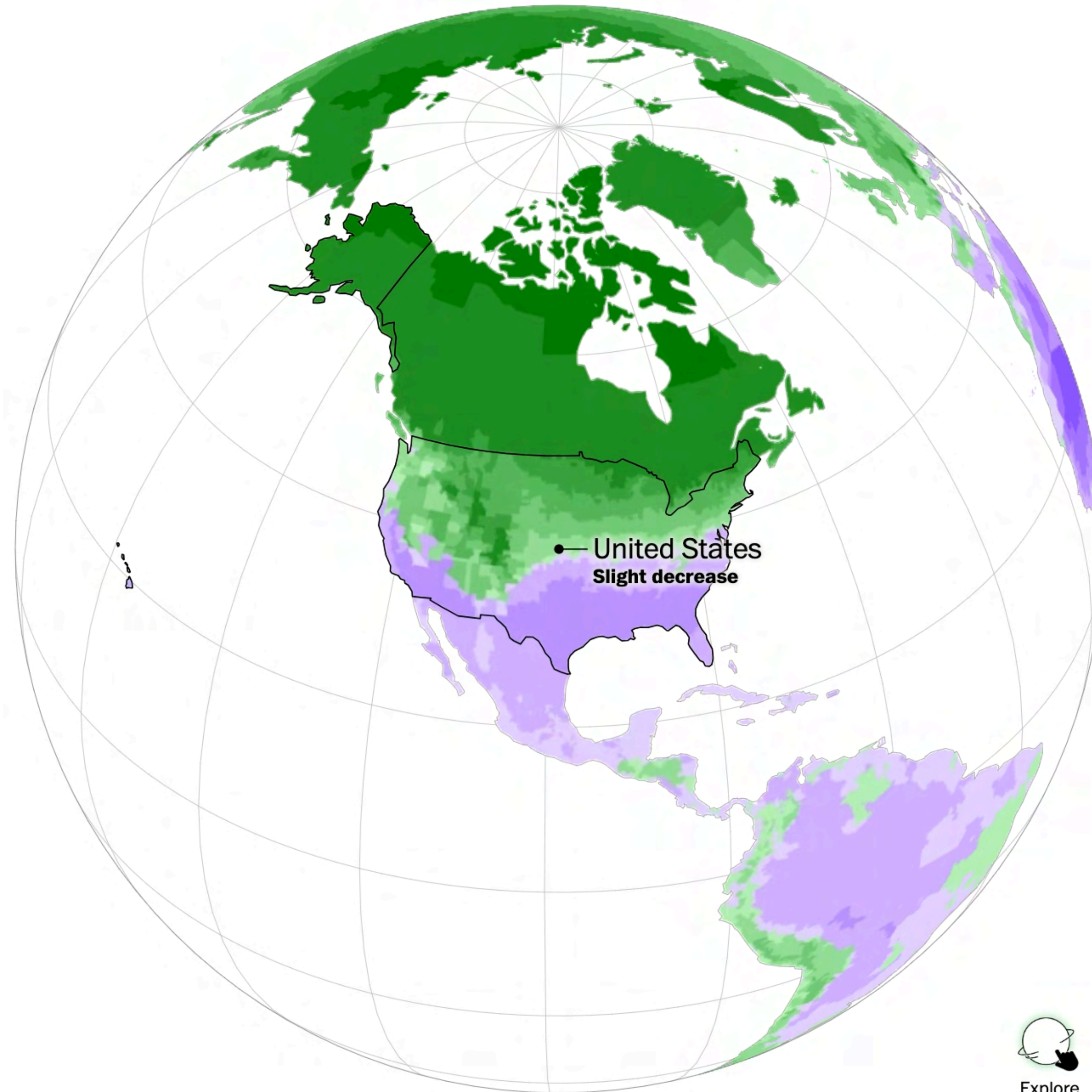
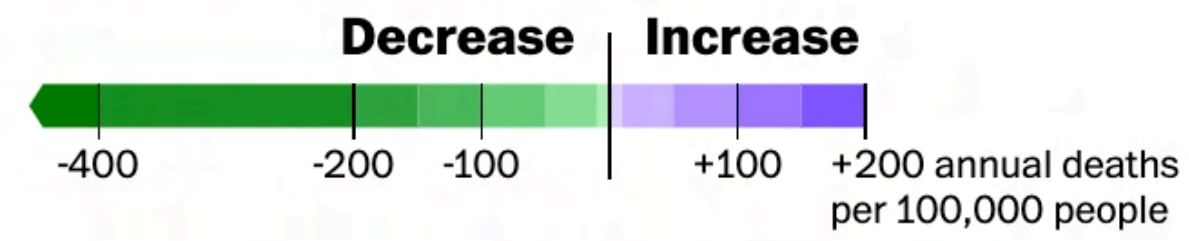
The New York Times

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

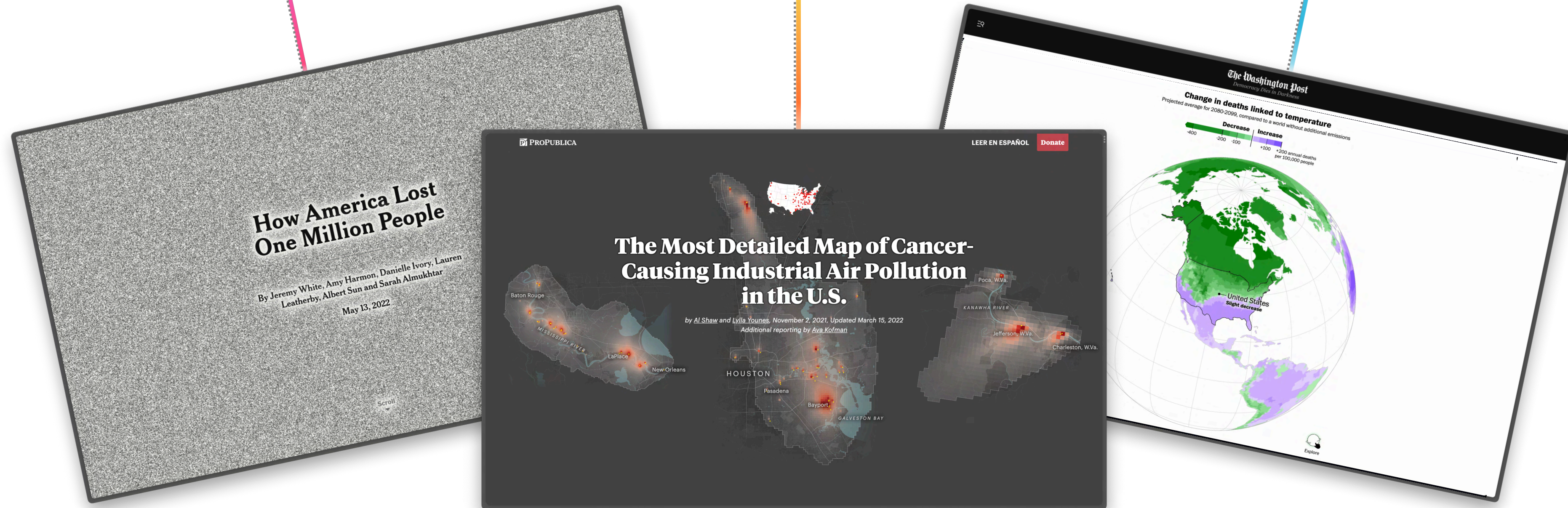
Will global warming make temperature less deadly?

Harry Stevens



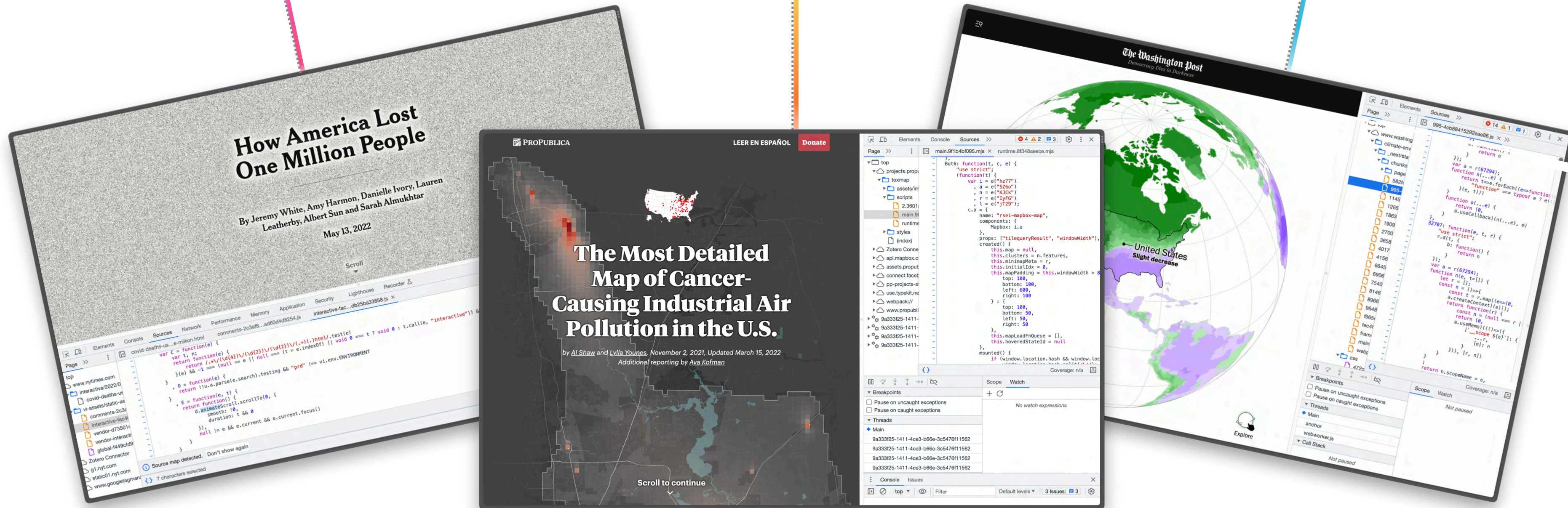
Explore

Data Journalism



“Telling **stories** with **data**”

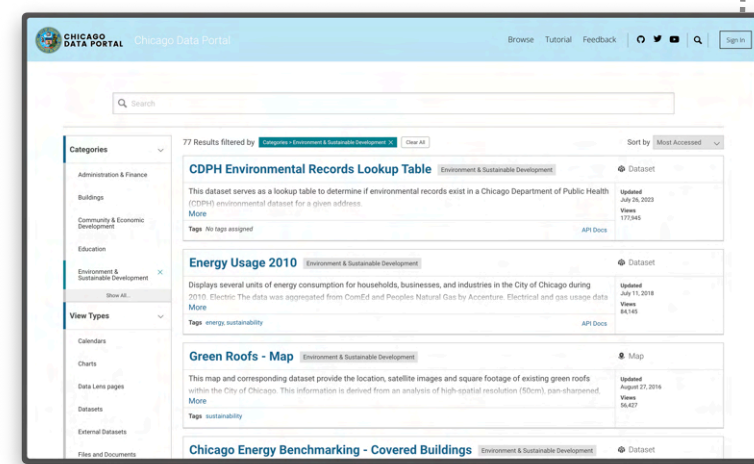
Data Journalism



(Lots and lots of)

Programming

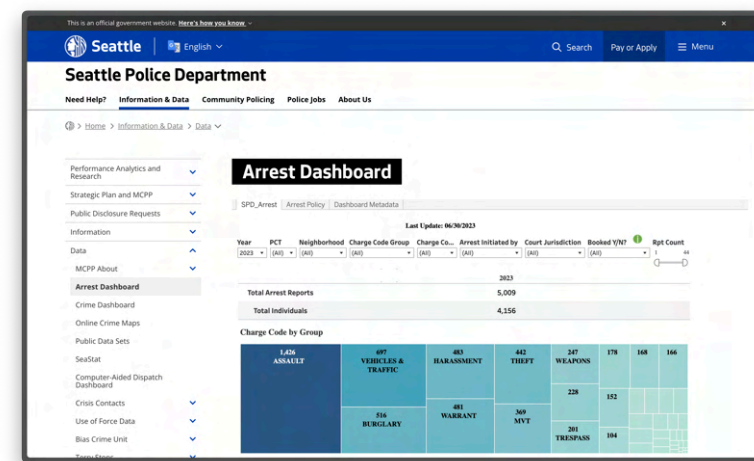
Data Journalism



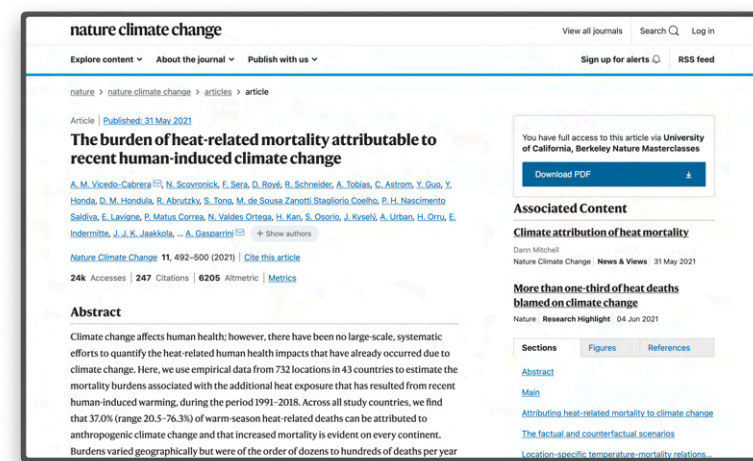
Municipal Agencies



FOIA Requests



Police Departments



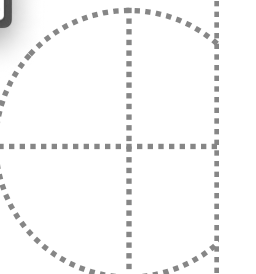
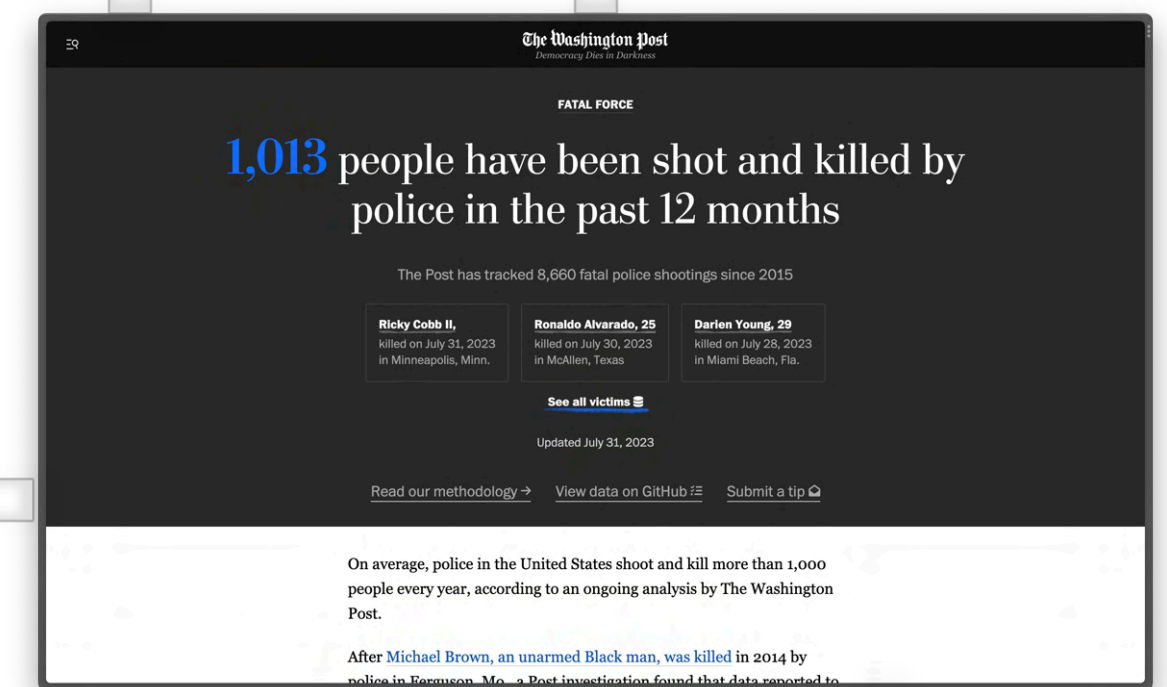
Academic Journals

Normalize

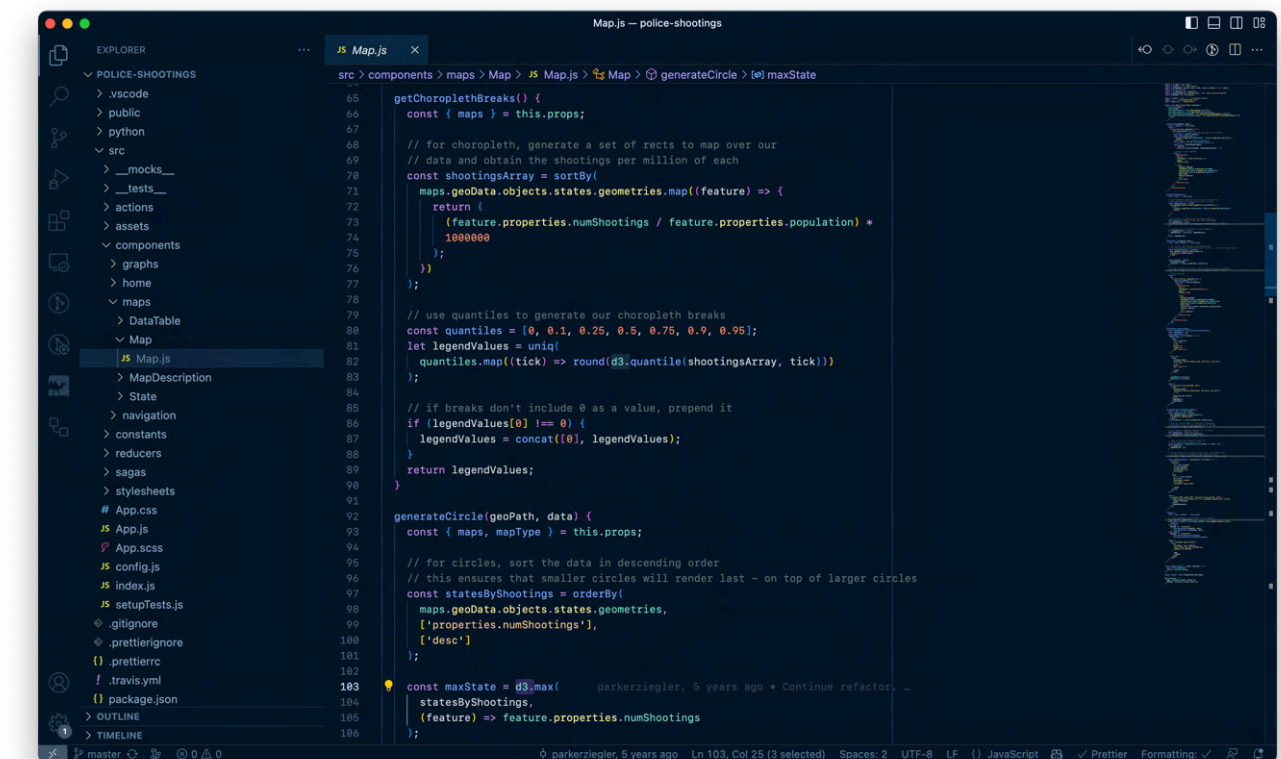
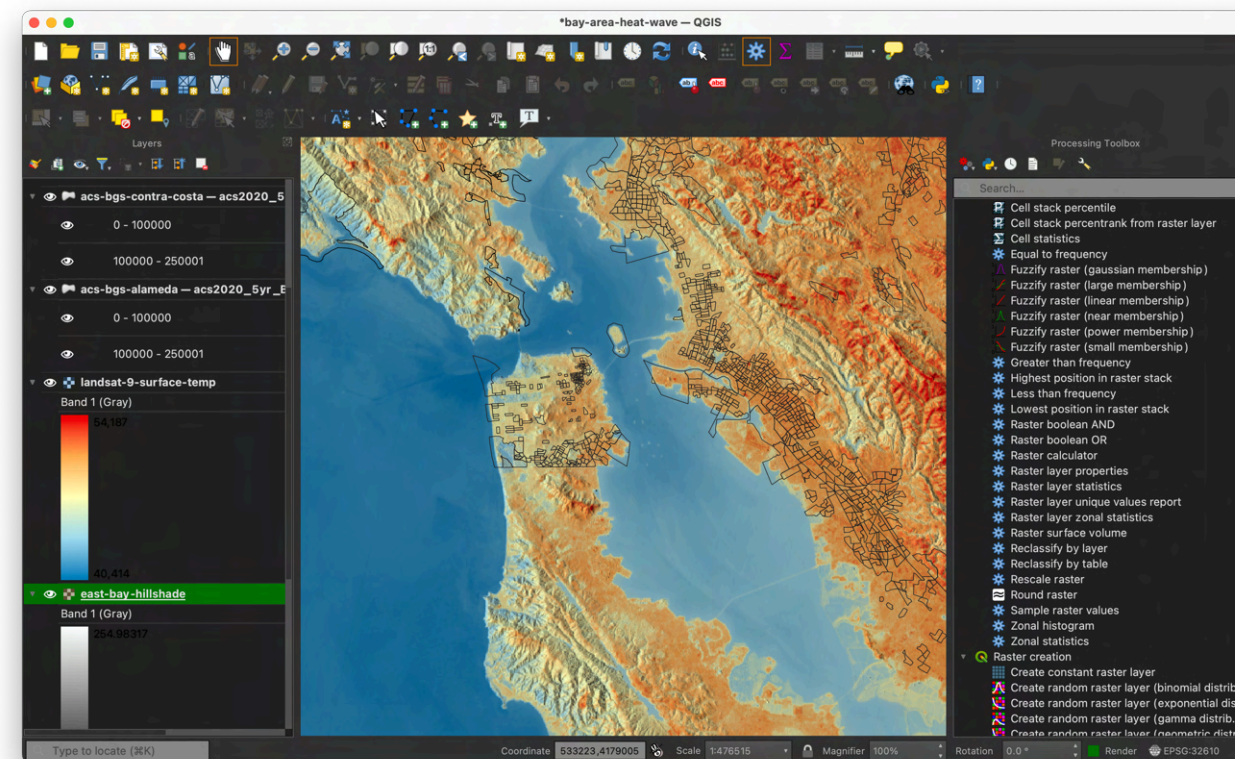
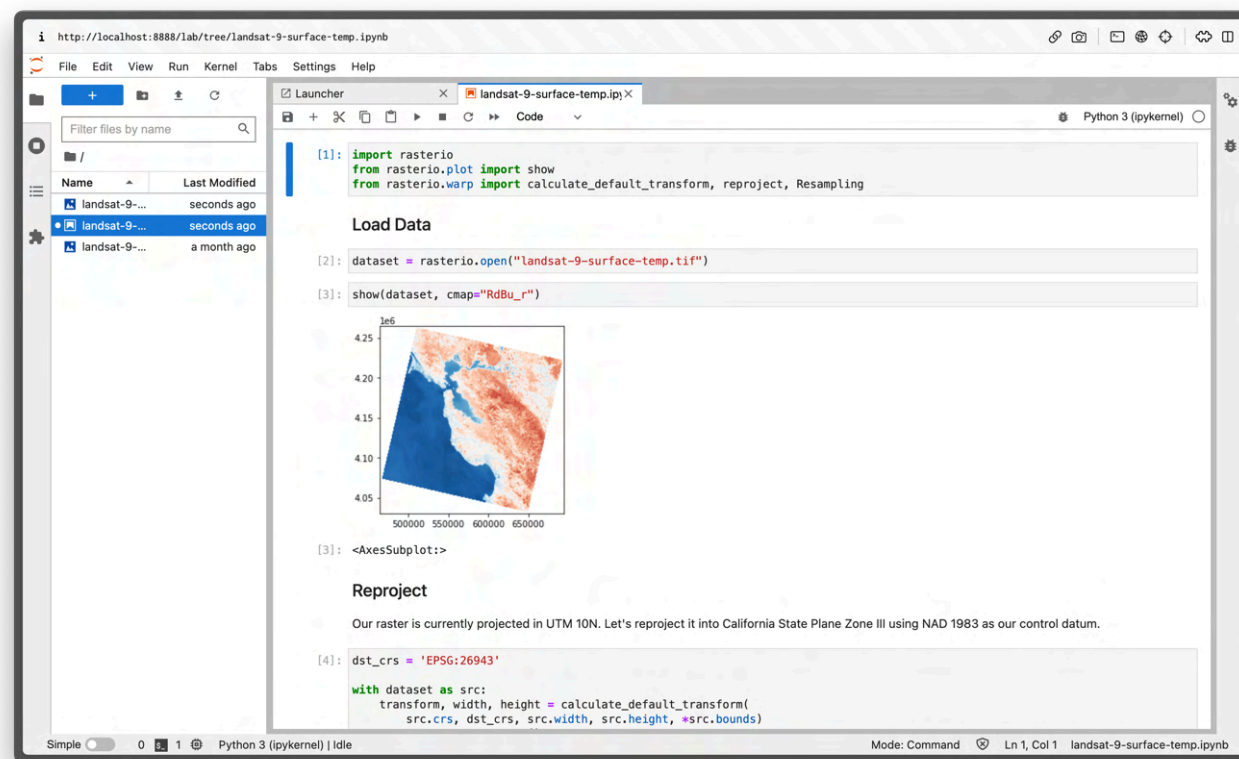
Filter

Transform

Clean



Data Journalism



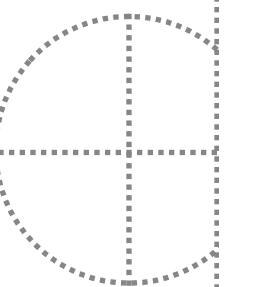
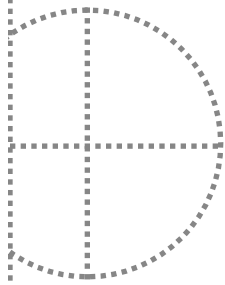
Computational Notebooks

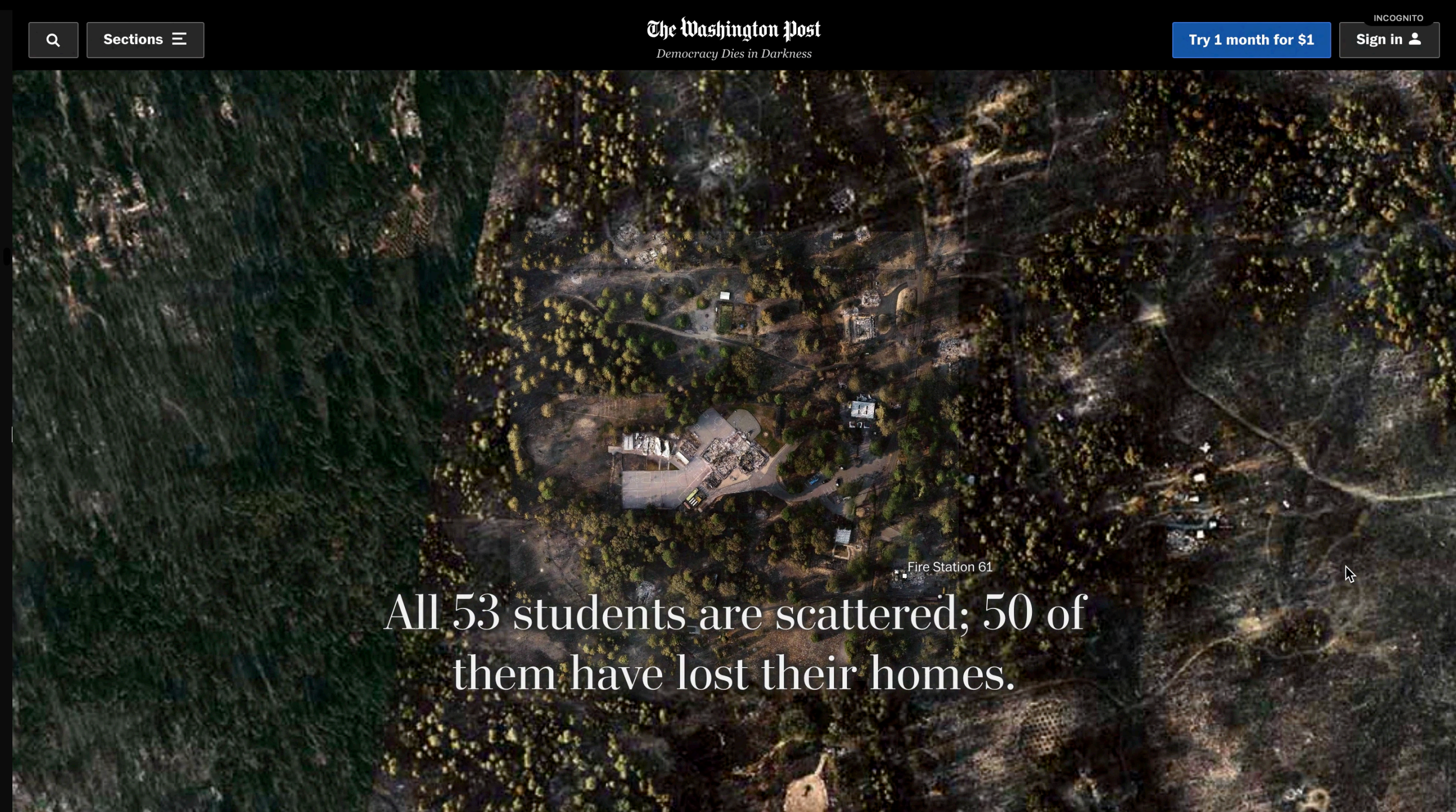
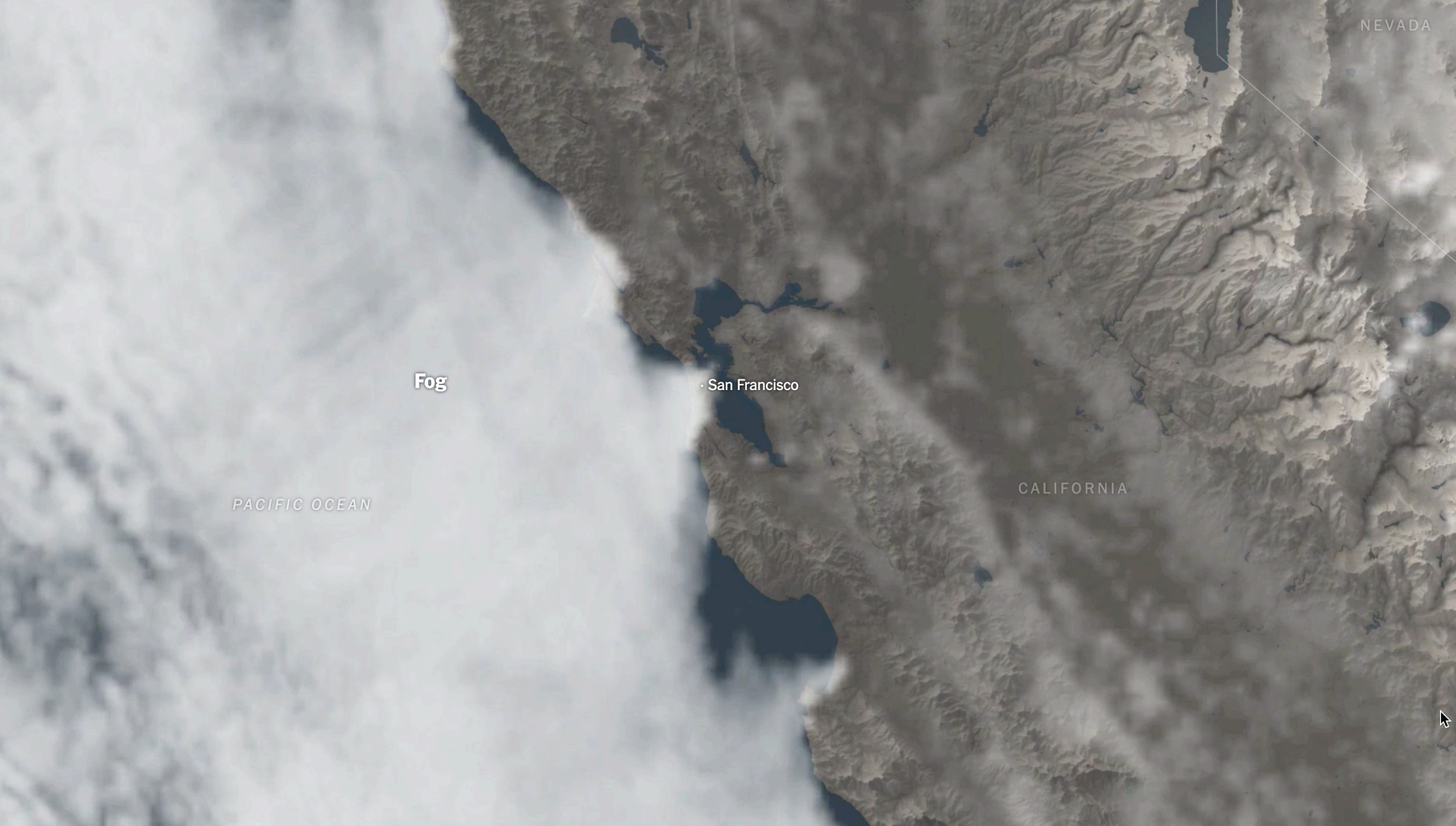


GIS Software



Web Frontends





Data Journalism



Programming

Data Journalism

Data Science

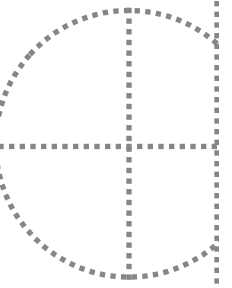
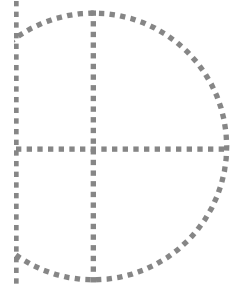
Visualization

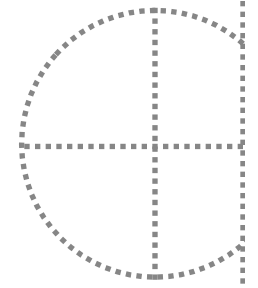
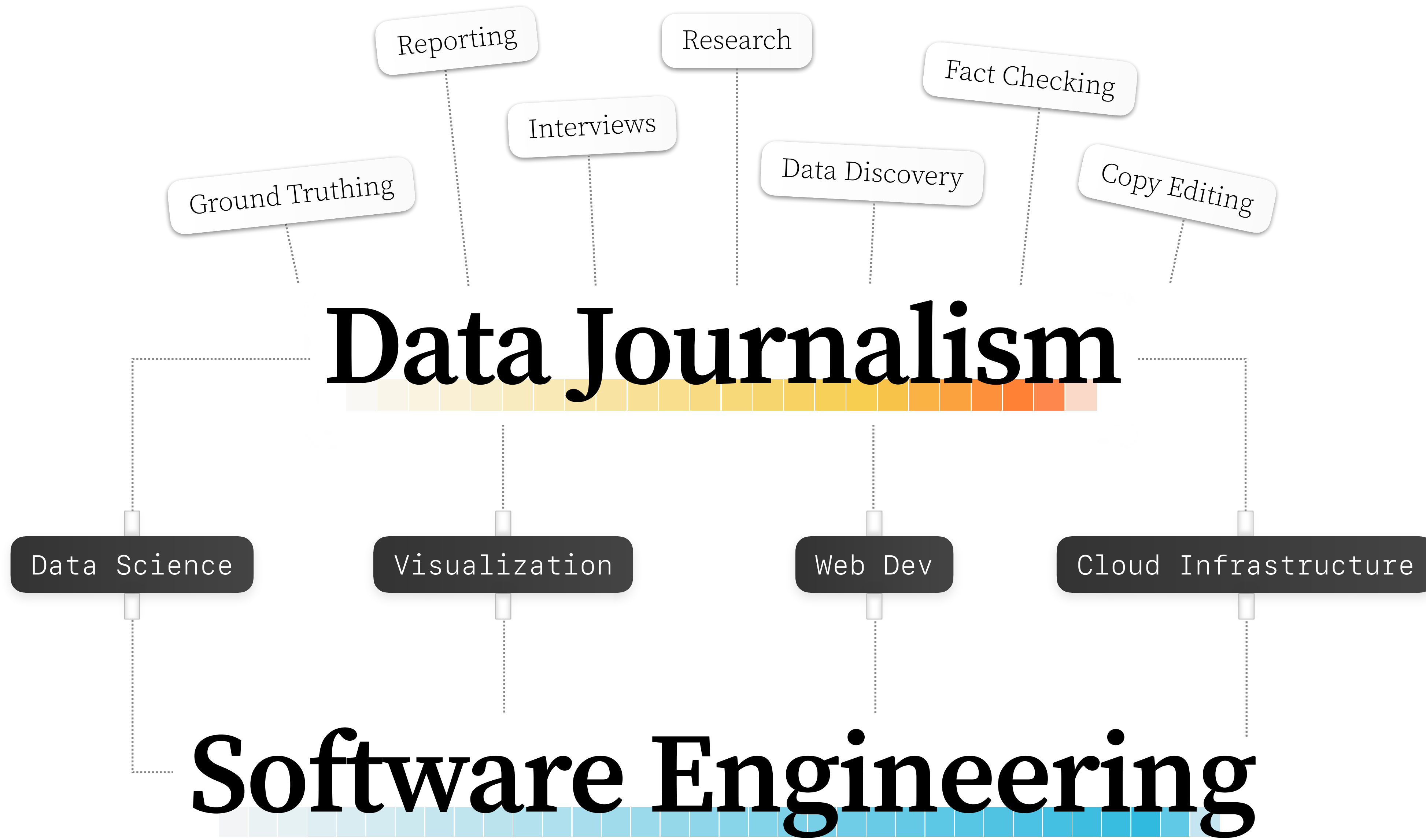
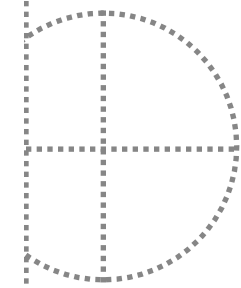


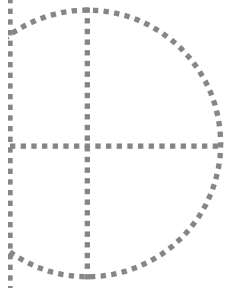
Web Dev

Cloud Infrastructure

Software Engineering







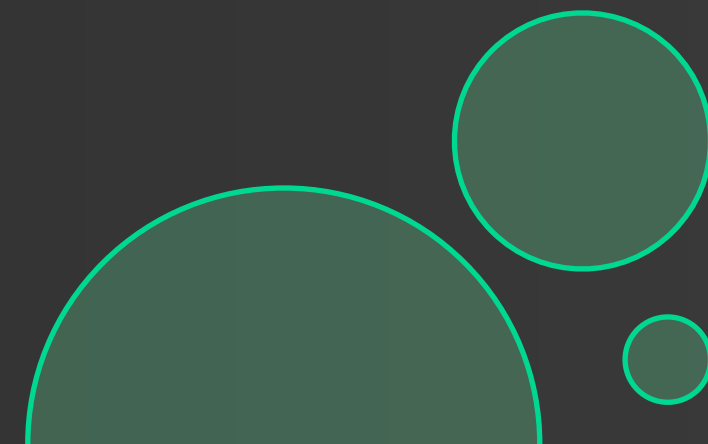
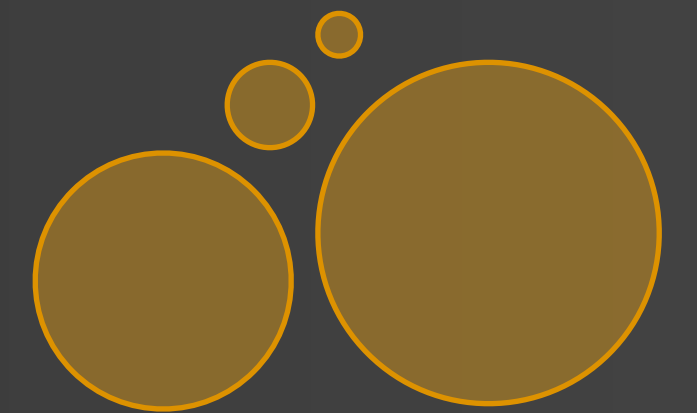
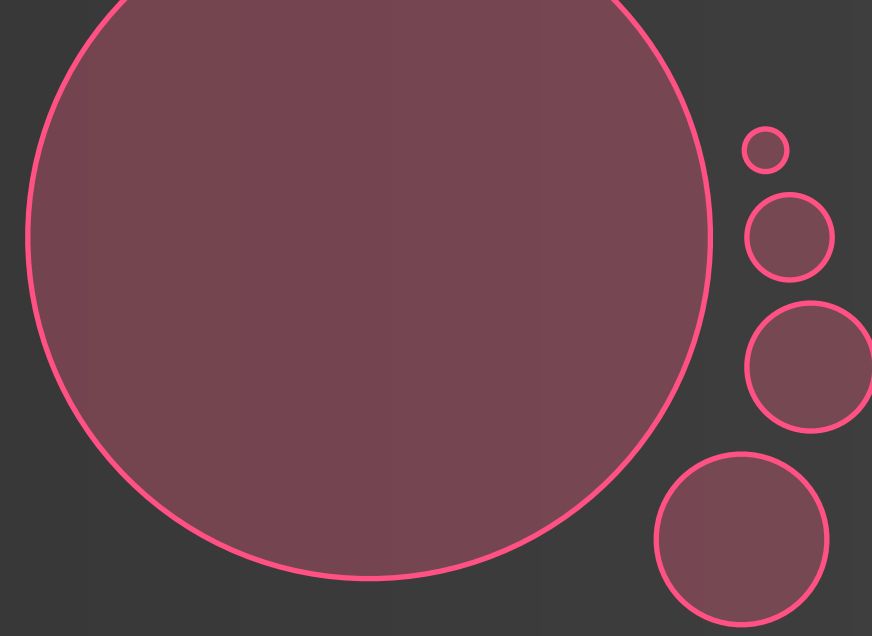
TWP

Los Angeles Times

PROPUBLICA

FiveThirtyEight

Graphics Desk



Grist

THE  CITY

The Markup



Los Angeles Times

 PROPUBLICA

 FiveThirtyEight

KQED

CAL  MATTERS

 EYE ON OHIO
Ohio Center for Journalism

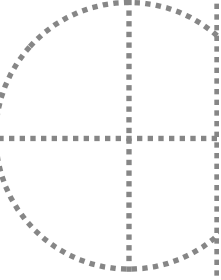
Graphics

Desk

(1 or 2 data journalists)



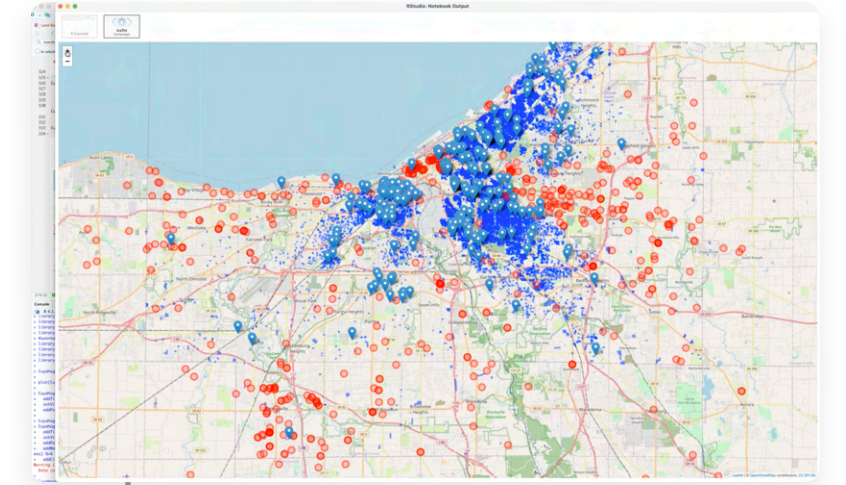
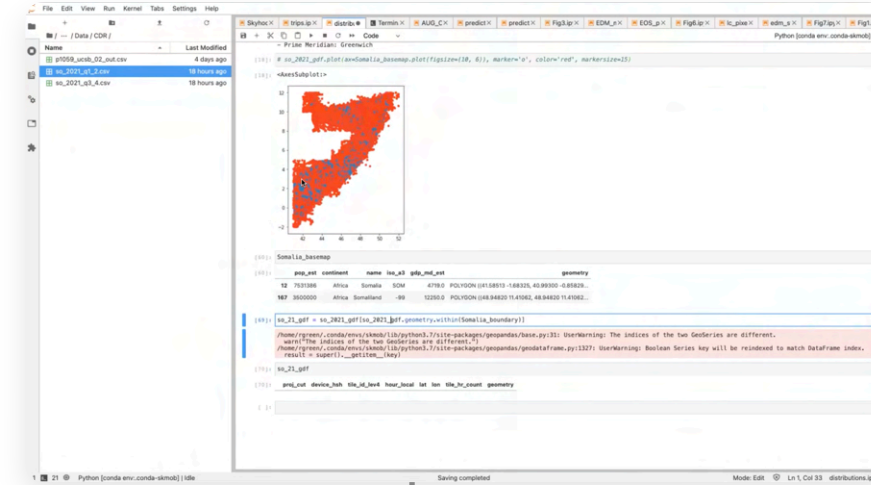
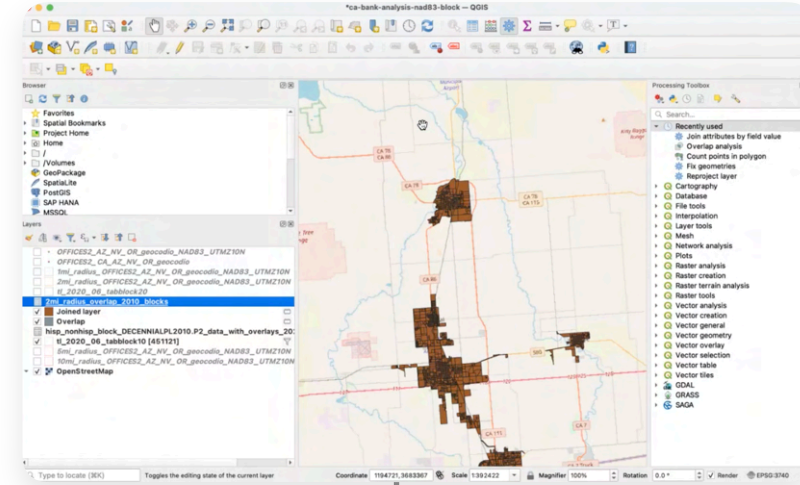
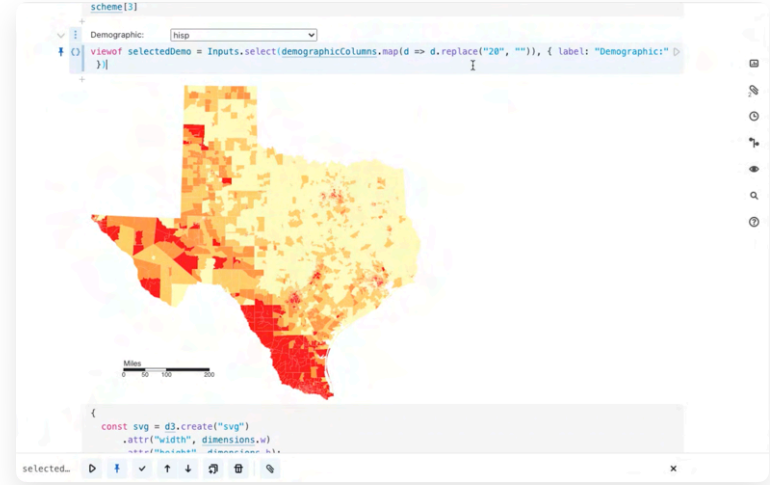
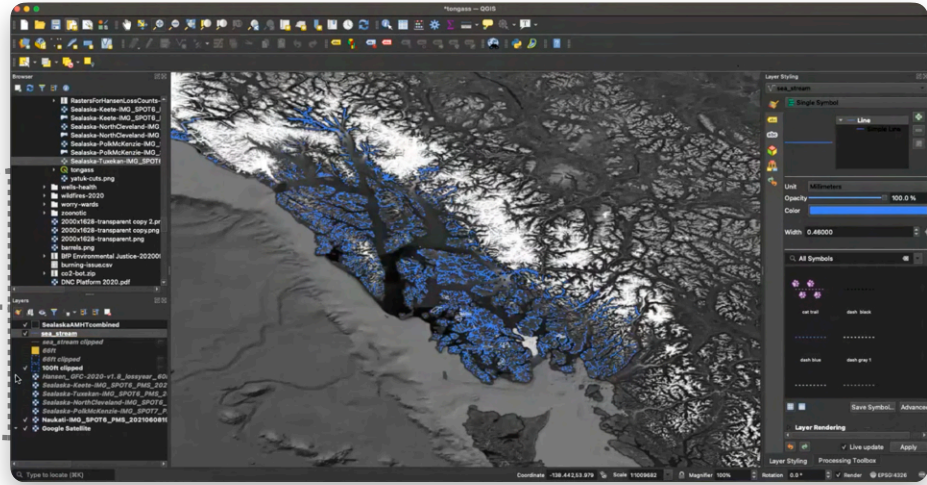
How can we design and build
programming tools
to support the working practices of
data journalists?





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





30 hours

of video observation with



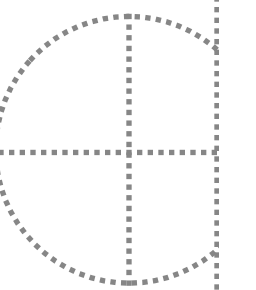
Data Journalists

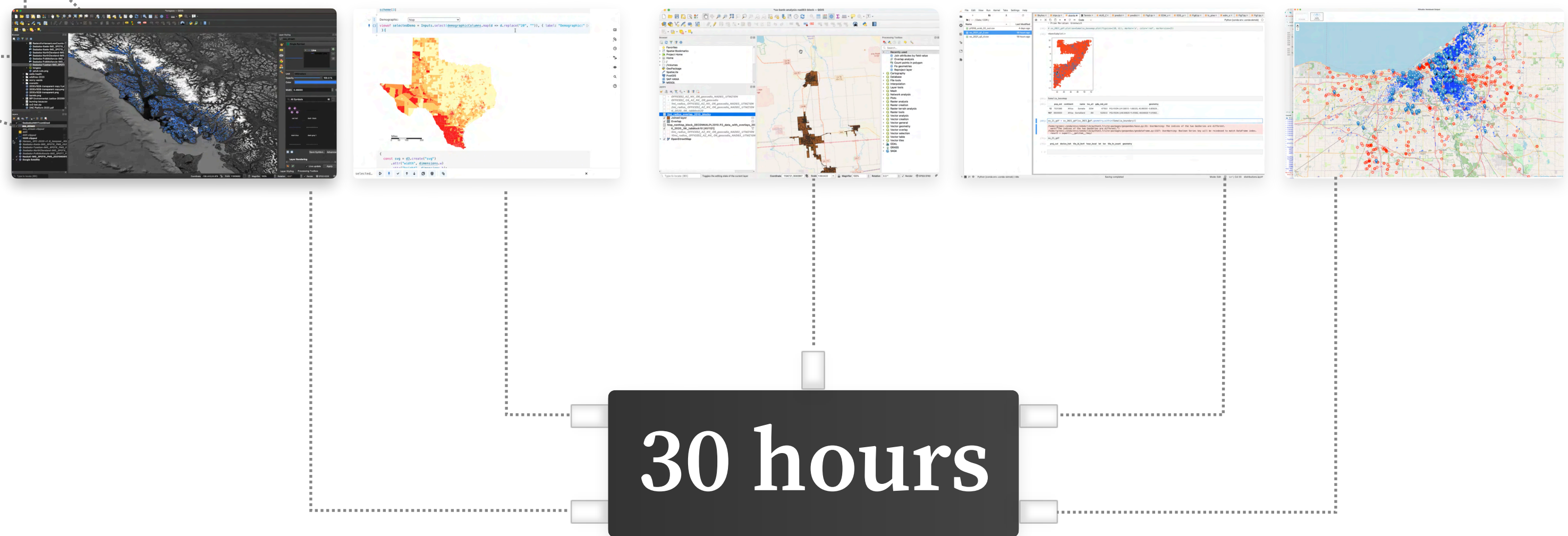


Social Scientists



Earth and Climate Scientists





30 hours

of video observation with

 **Data Journalists**

 **Social Scientists**

 **Earth and Climate Scientists**

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

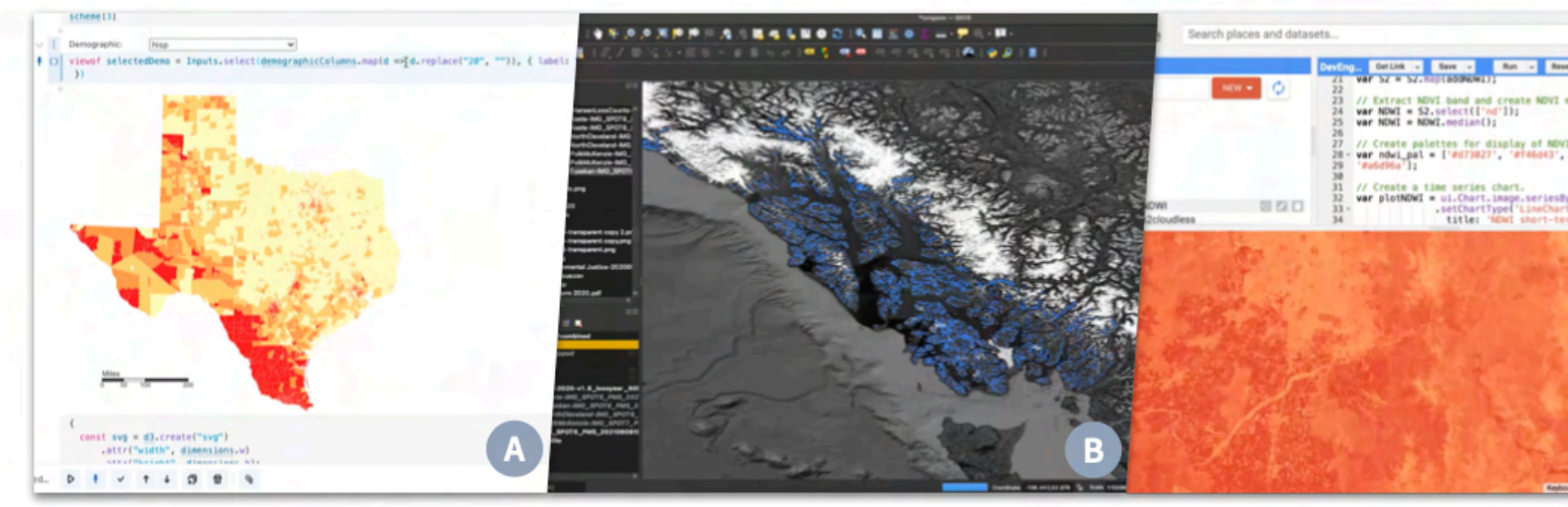


Figure 1: Example screenshots from participants' work with geospatial data. (A) PJ3 creates a choropleth map of Texas proposed electoral districts colored by majority racial demographic in Observable. (B) PJ7 combines satellite imagery data, and deforestation data in QGIS to identify illegal logging 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

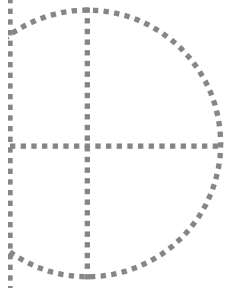
ACM Reference Format:

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>

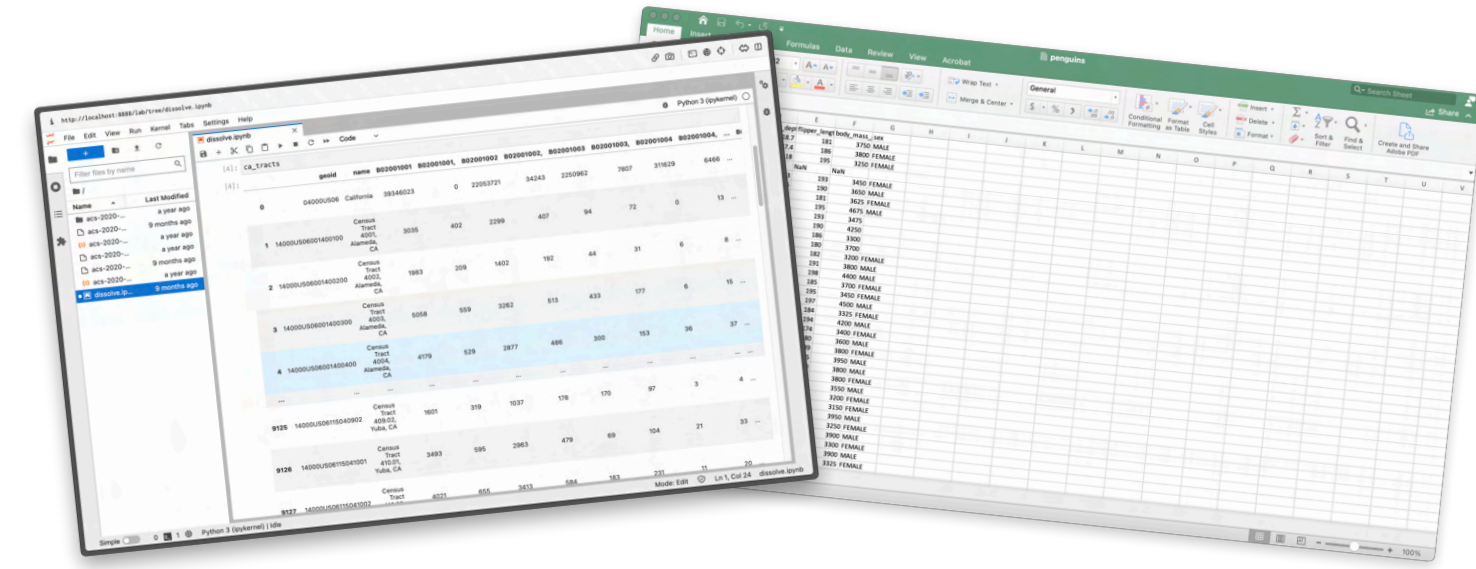
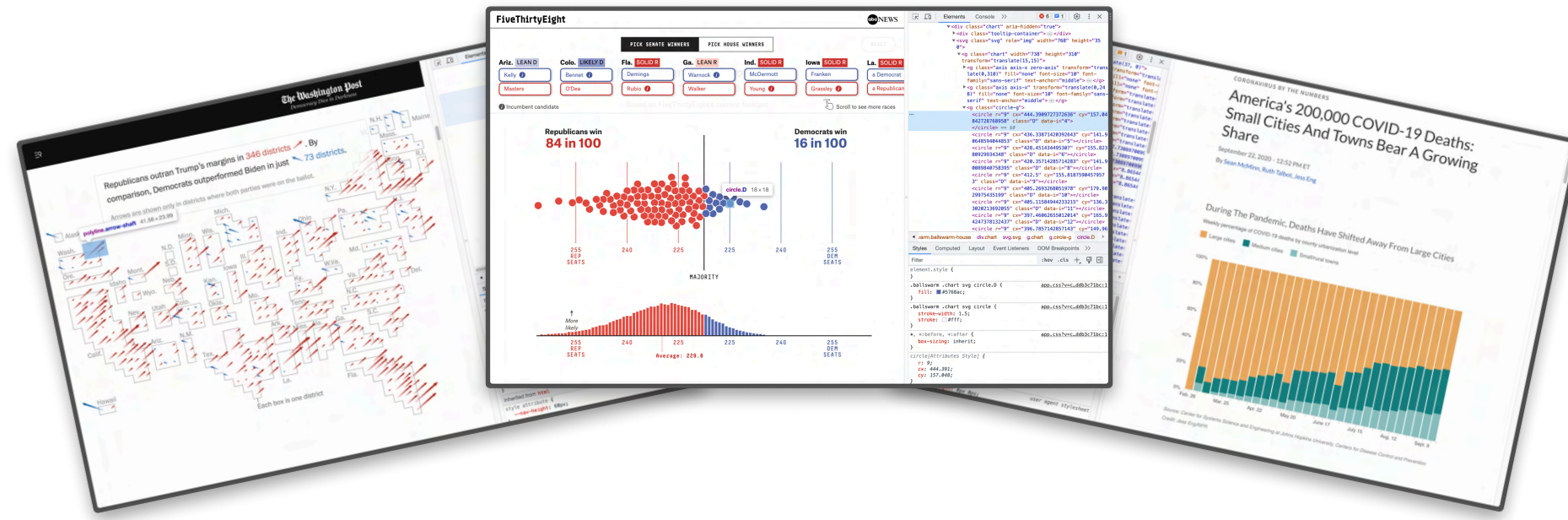
1 INTRODUCTION

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 observation satellites generate 80TB of new imagery daily [8] to the surface, cheap, power-efficient sensors create massive amounts of geolocated data measuring real-time environment [40]. Additionally, crowdsourcing efforts like OpenStreetMap 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 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], segregation [82], hazard modeling [98], and other areas.

Despite this expansion in the community of geospatial users, research has yet to explore the specific challenges domain experts face in gathering, analyzing, and visualizing geospatial information. Many domain experts are self-taught in the



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



Apply to new datasets



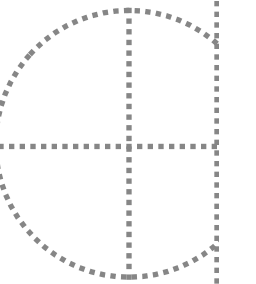
**Data
Journalists**

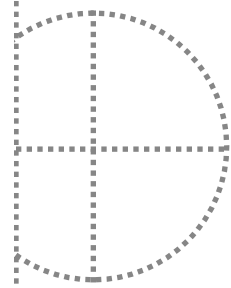


**Social
Scientists**



**Earth and
Climate Scientists**





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



Apply to new datasets



Reverse engineering was as time-consuming as developing the visualization **from scratch**.



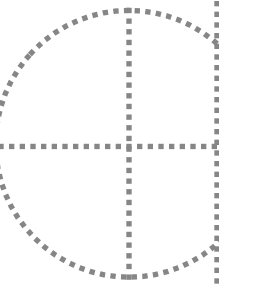
Data Journalists

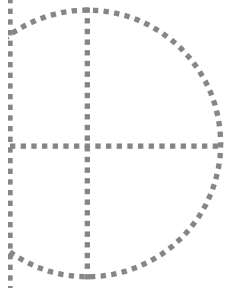


Social Scientists



Earth and Climate Scientists

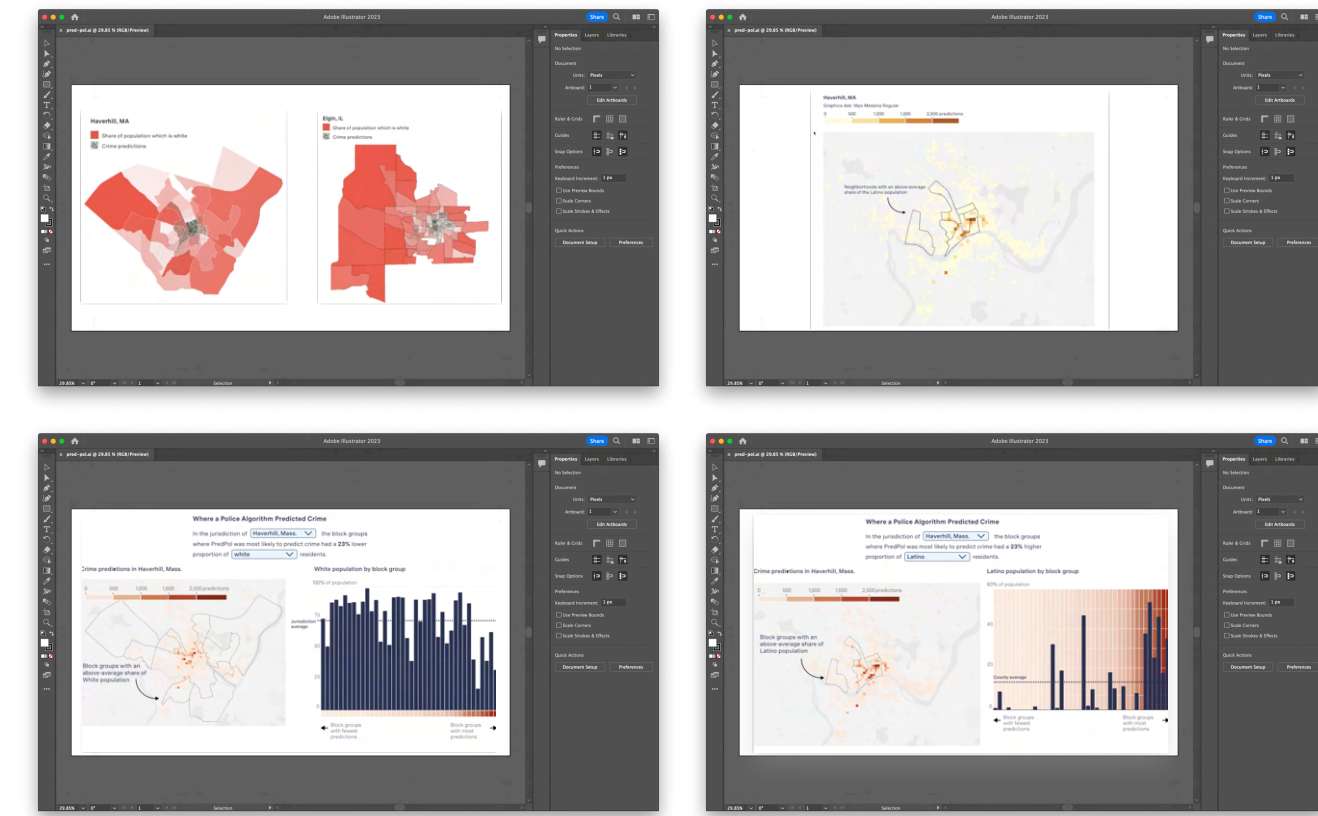




Sketch visualizations in **design software**



Explore a wide design space **without code**



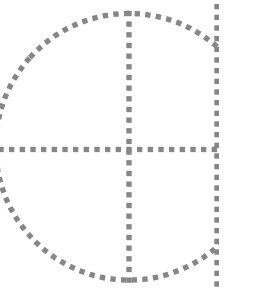
**Data
Journalists**

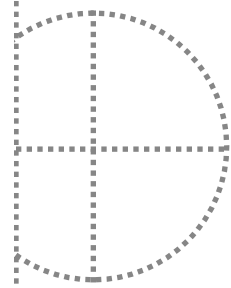


**Social
Scientists**



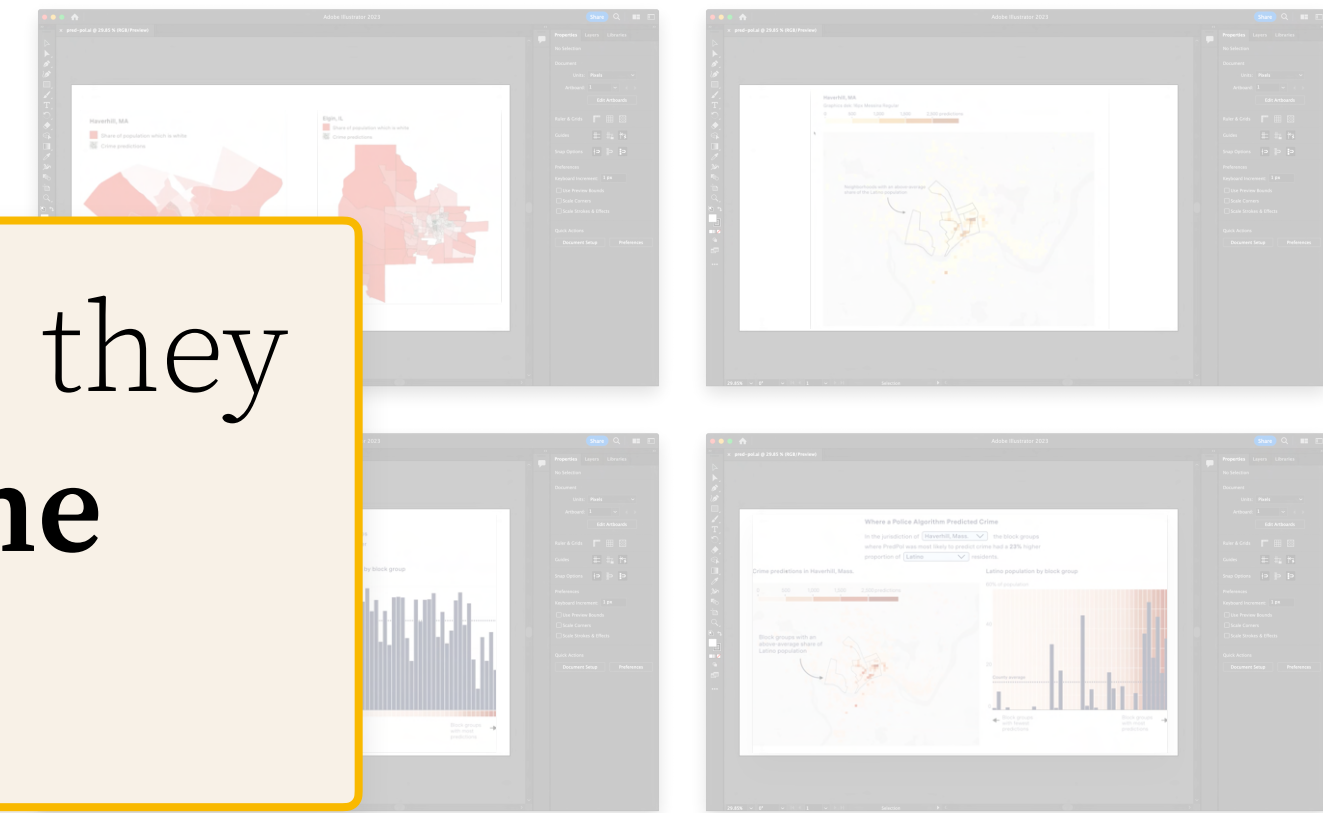
**Earth and
Climate Scientists**





Sketch visualizations in **design software**

Explore a wide design space **without code**



After selecting a design, they still had to **reproduce the visualization in code**.



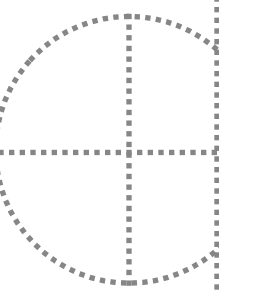
Data Journalists



Social Scientists



Earth and Climate Scientists



Visual Inputs



Programs

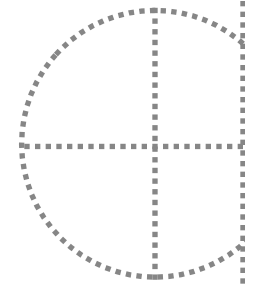
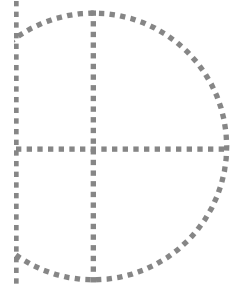


```
const x = d3.scaleLinear()
  .domain(d3.extent(cars, d => d["weight (lb)"]))
  .range([marginLeft, width - marginRight]);

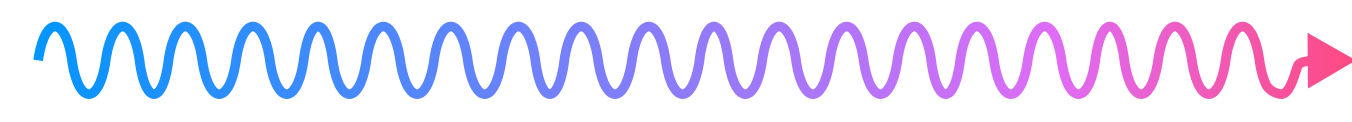
const svg = d3.create("svg")
  .attr("width", width)
  .attr("height", height)
  .attr("viewBox", [0, 0, width, height])
  .attr("style", "max-width: 100%; height: auto;");

svg.append("g")
  .attr("transform", `translate(0,${height - marginBottom})`)
  .call(d3.axisBottom(x).tickSizeOuter(0));

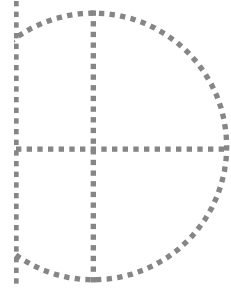
svg.append("g")
  .selectAll()
  .data(dodge(cars, {radius: radius * 2 + padding}))
  .join("circle")
  .attr("cx", d => d.x)
  .attr("cy", d => height - marginBottom - radius - padding - d.y)
  .attr("r", radius)
  .append("title")
  .text(d => d.data.name);
```



Visual Inputs



Programs



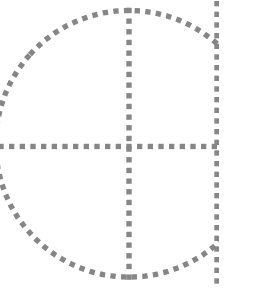
```
const x = d3.scaleLinear()
  .domain(d3.extent(cars, d => d["weight (lb)"]))
  .range([marginLeft, width - marginRight]);

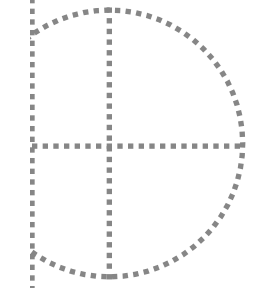
const svg = d3.create("svg")
  .attr("width", width)
  .attr("height", height)
  .attr("viewBox", [0, 0, width, height])
  .attr("style", "max-width: 100%; height: auto;");

svg.append("g")
  .attr("transform", `translate(0,${height - marginBottom})`)
  .call(d3.axisBottom(x).tickSizeOuter(0));

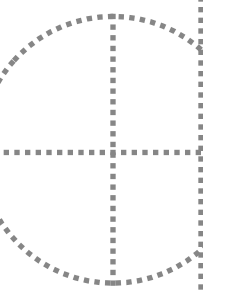
svg.append("g")
  .selectAll()
  .data(dodge(cars, {radius: radius * 2 + padding}))
  .join("circle")
  .attr("cx", d => d.x)
  .attr("cy", d => height - marginBottom - radius - padding - d.y)
  .attr("r", radius)
  .append("title")
  .text(d => d.data.name);
```

Could this be a **compilers** problem?






Maybe?



Could this be a **compilers** problem?



PL RESEARCHERS

**LITERALLY
ANYTHING**

IS THIS 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"))
```

WA

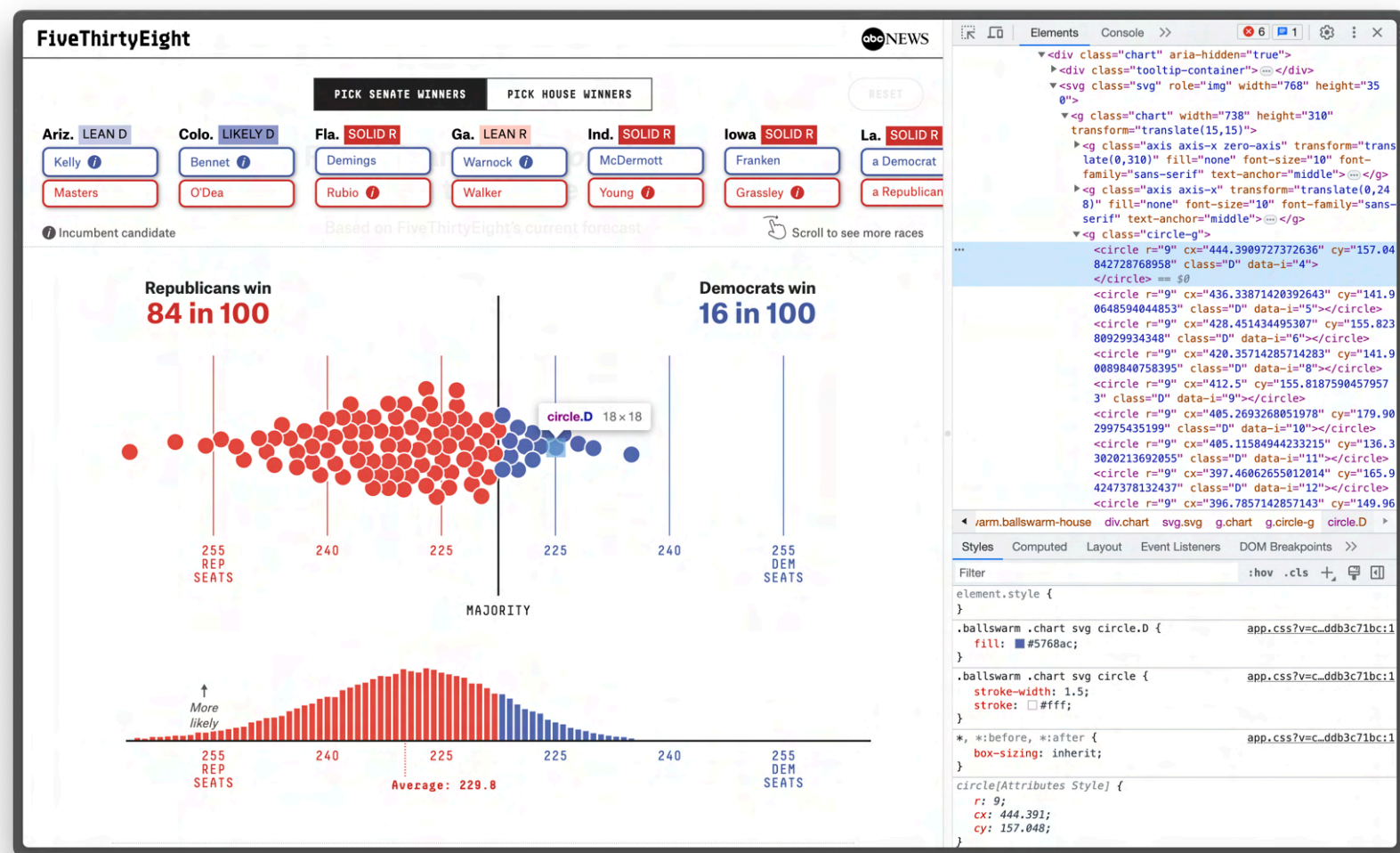
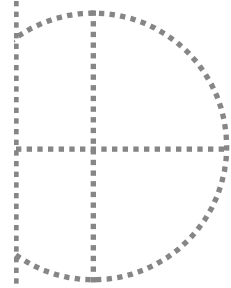
Lexer

Parser

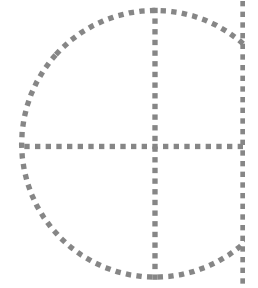
Intern. Repr.

Code Generator

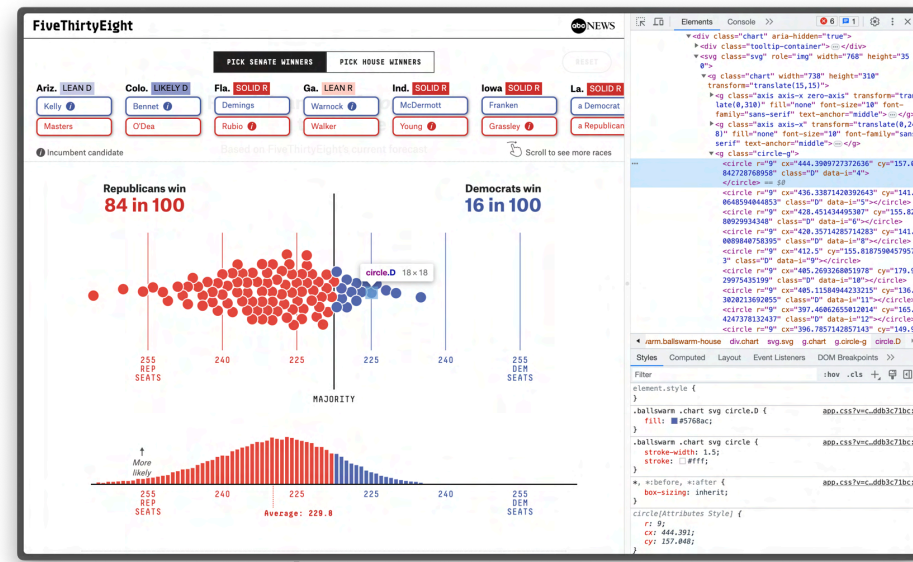
How do we compile from a **visual form** to a **textual symbolic representation**?



```
(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"))
```

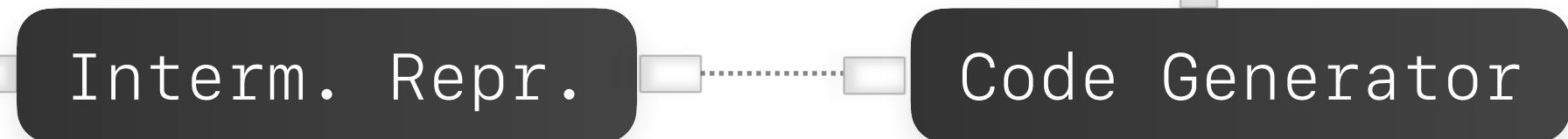
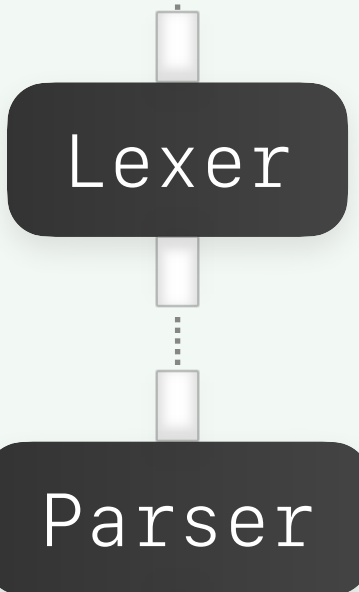


How do we compile from a **visual form** to a **textual symbolic representation**?



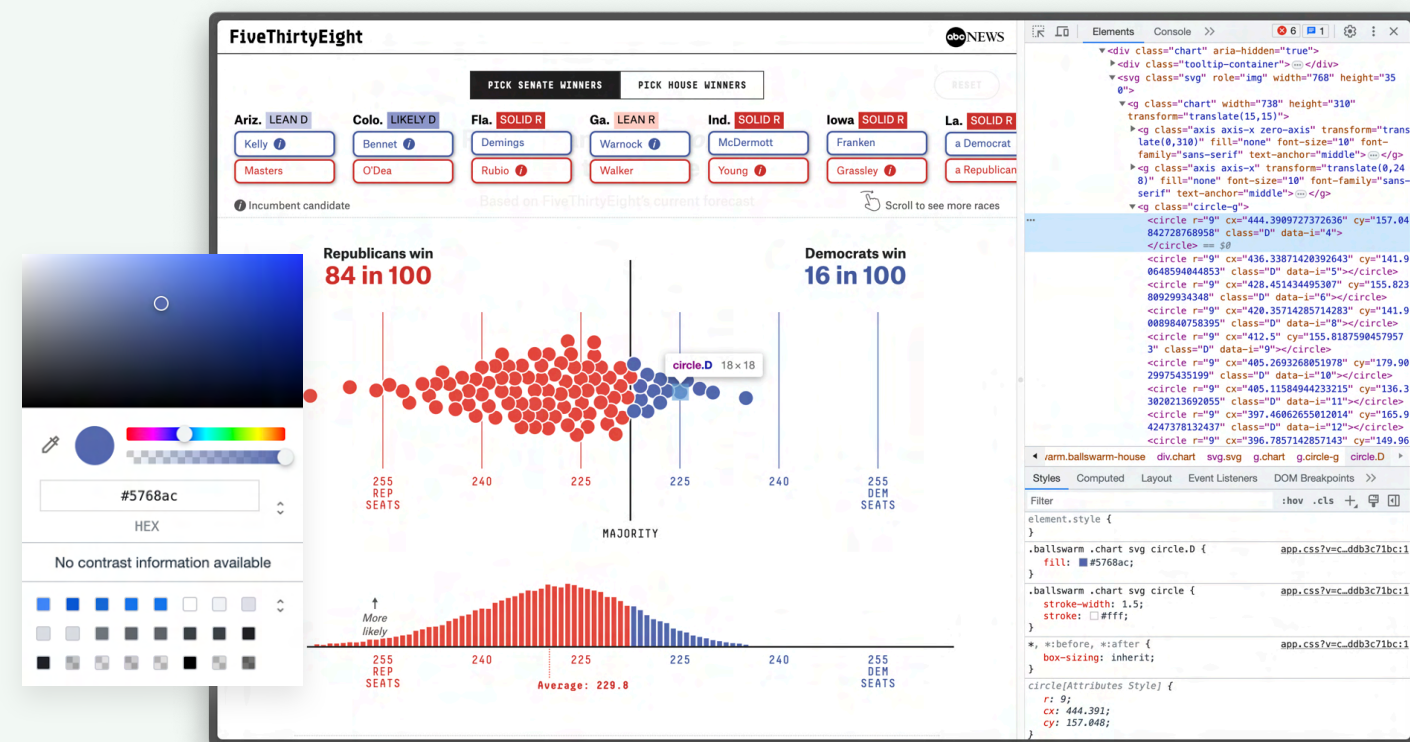
```
(module
  (type (;0;) (func (param i32 i32) (result i32)))
  (func (;0;) (type 0) (param i32 i32) (result i32)
    get_local 1
    get_local 0
    i32.add)
  (table (;0;) 1 1 anyfunc)
  (memory (;0;) 17)
  global (;0;) i32 (i32.const 1049118)
  global (;1;) i32 (i32.const 1049118)
  export "memory" (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"))
```

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



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"))
```

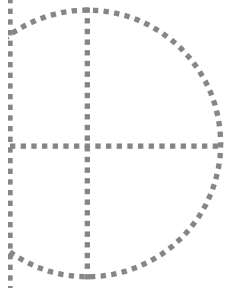
Lexer

Parser

Intern. Repr.

Code Generator

WA



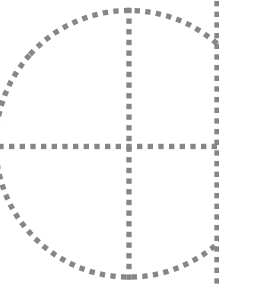
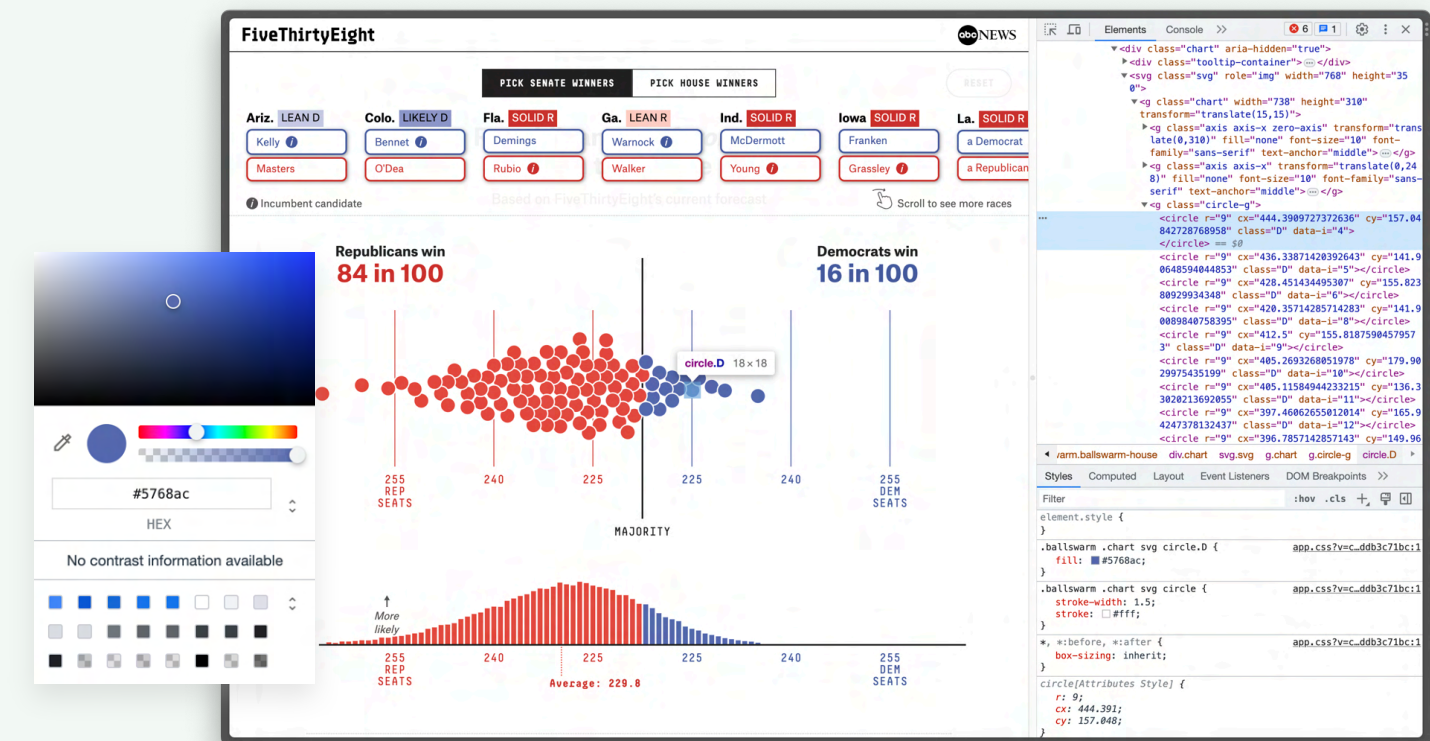
Lexer

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

Parser



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



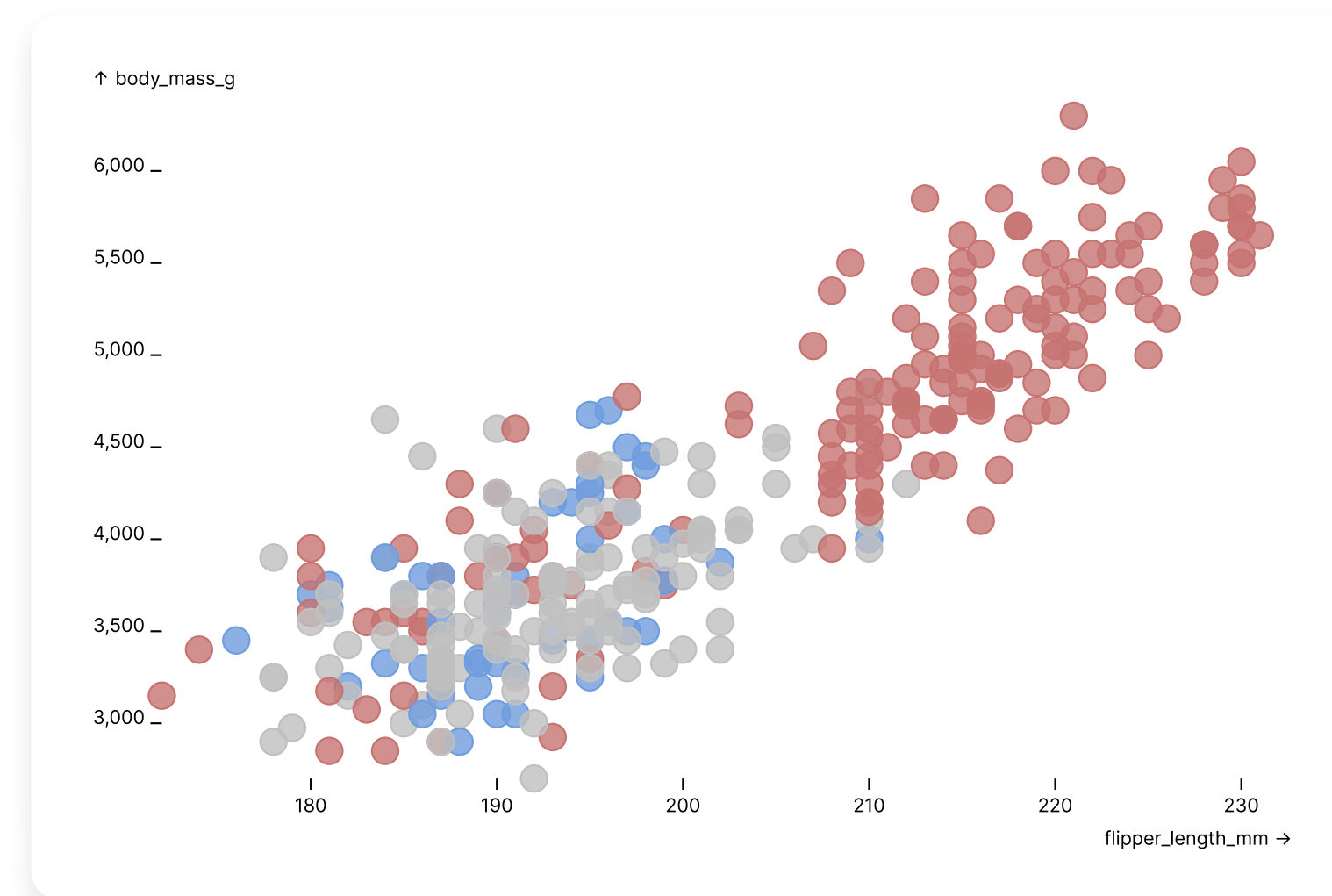
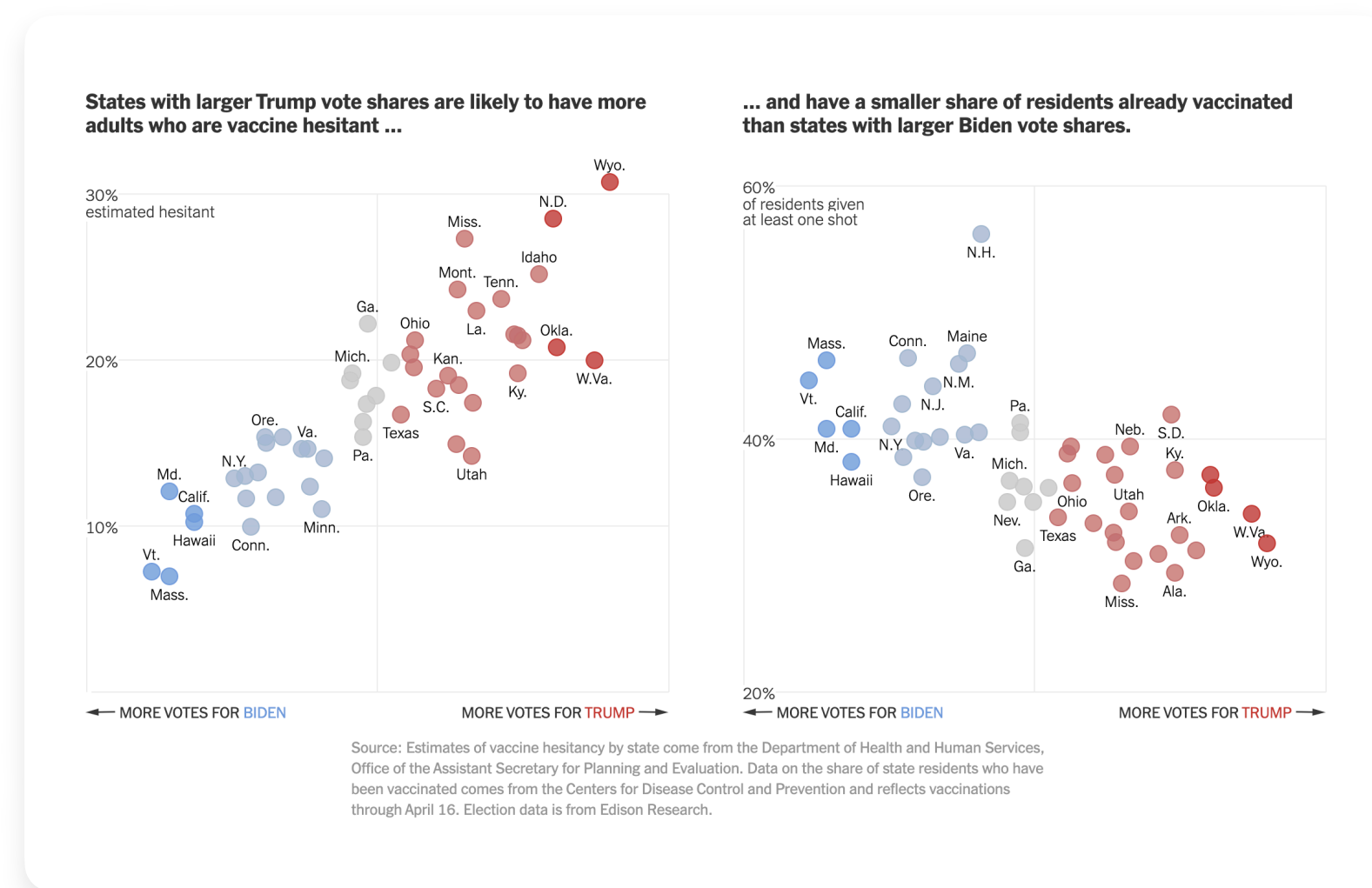
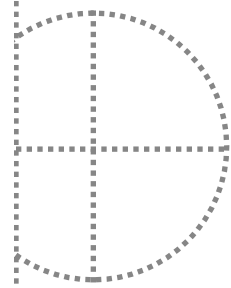


A compiler from **SVG subtrees** to **data visualizations**



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

Demo



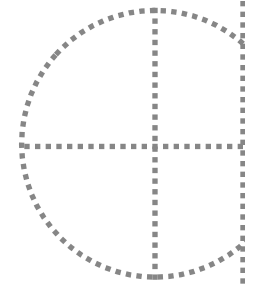
The New York Times

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

Danielle Ivory, Lauren Leatherby and Robert Gebeloff

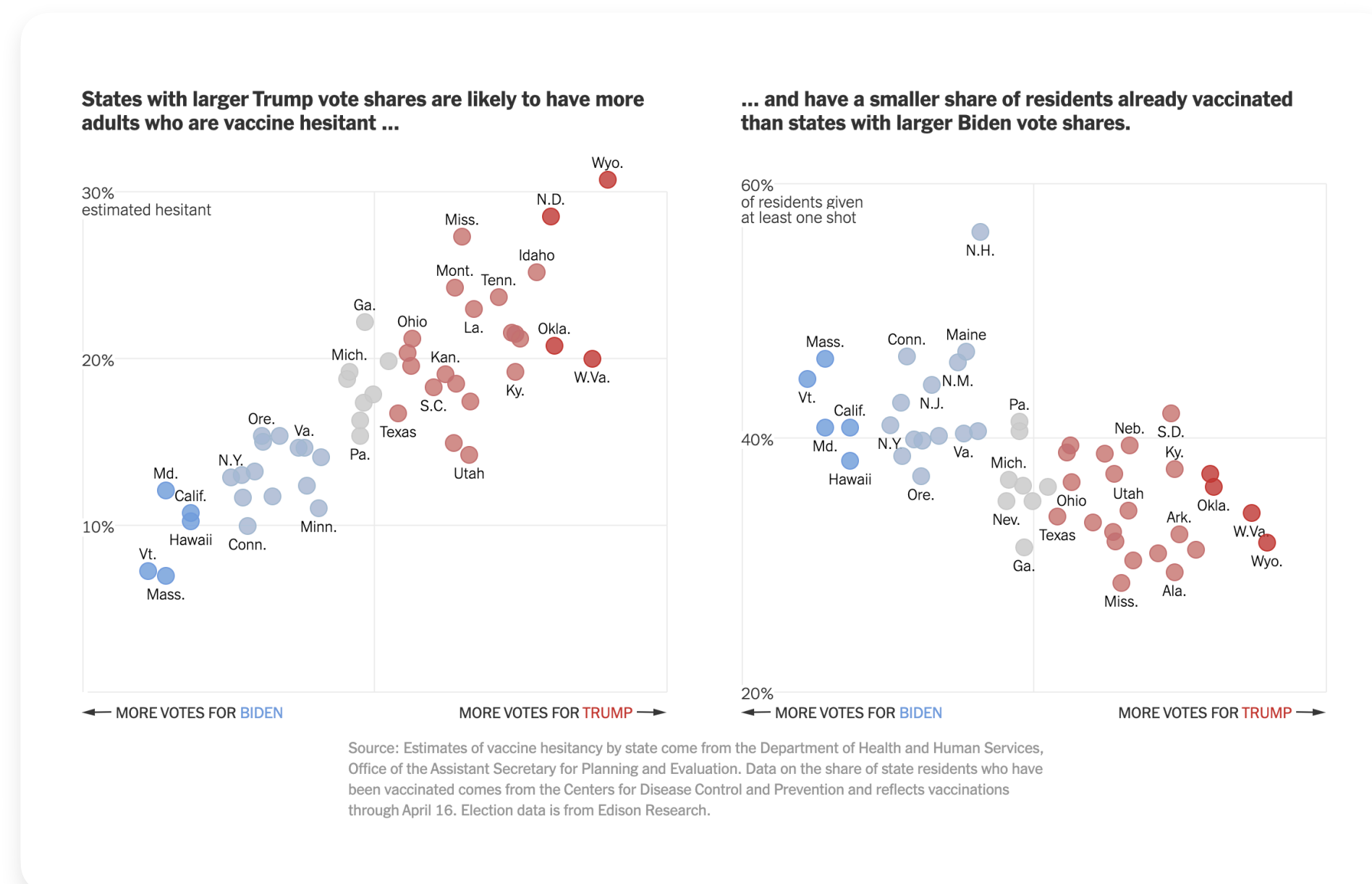


Comparing Flipper Length and Body Mass of the Palmer Penguins



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

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



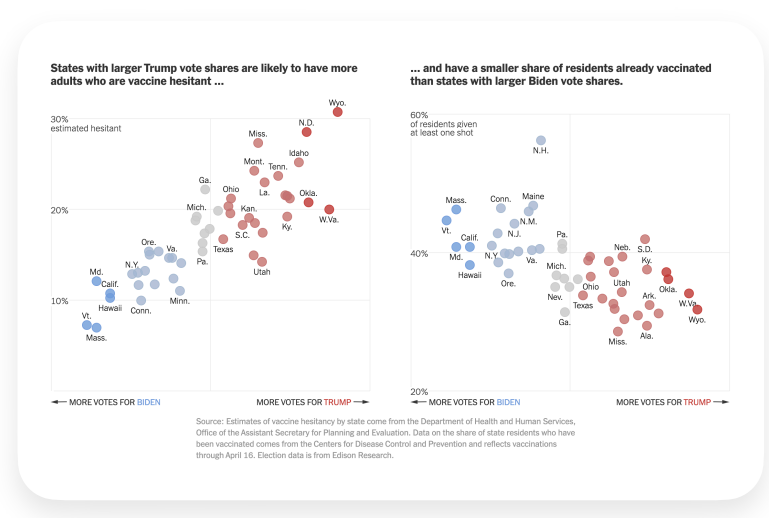
The New York Times

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

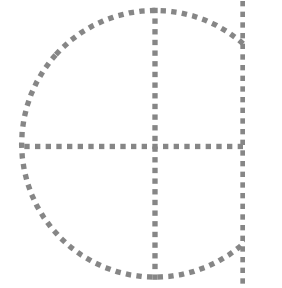
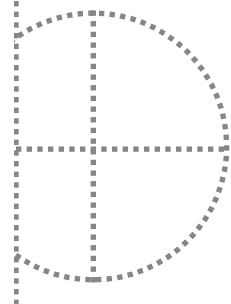
Danielle Ivory, Lauren Leatherby and Robert Gebeloff

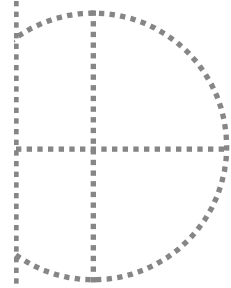


They pass its `svg` subtree to

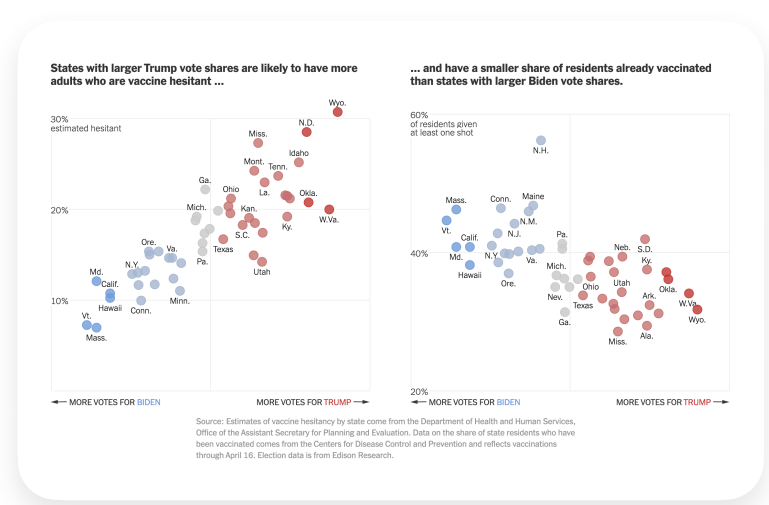


1.

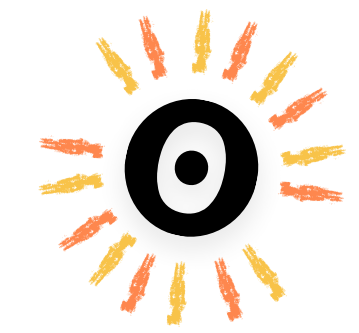




1.



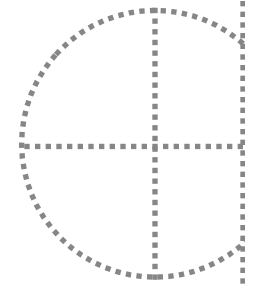
reviz compiles a **partial JavaScript program** using Observable's Plot library.

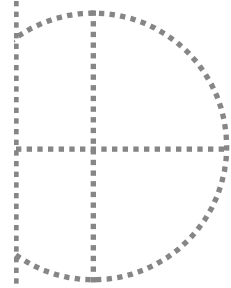


2.

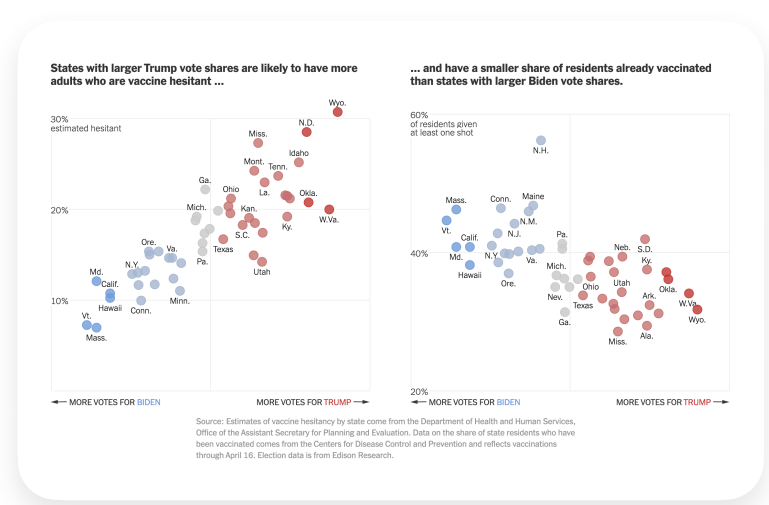


```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE", "#A7B9D3", "#C23734"]
  },
  marks: [
    Plot.dot(data, {
      x: "??",
      y: "??",
      fill: "??",
      fillOpacity: 1,
      stroke: "??",
      strokeOpacity: 1,
      strokeWidth: 1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





1.



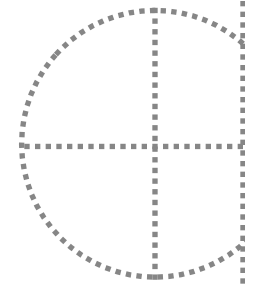
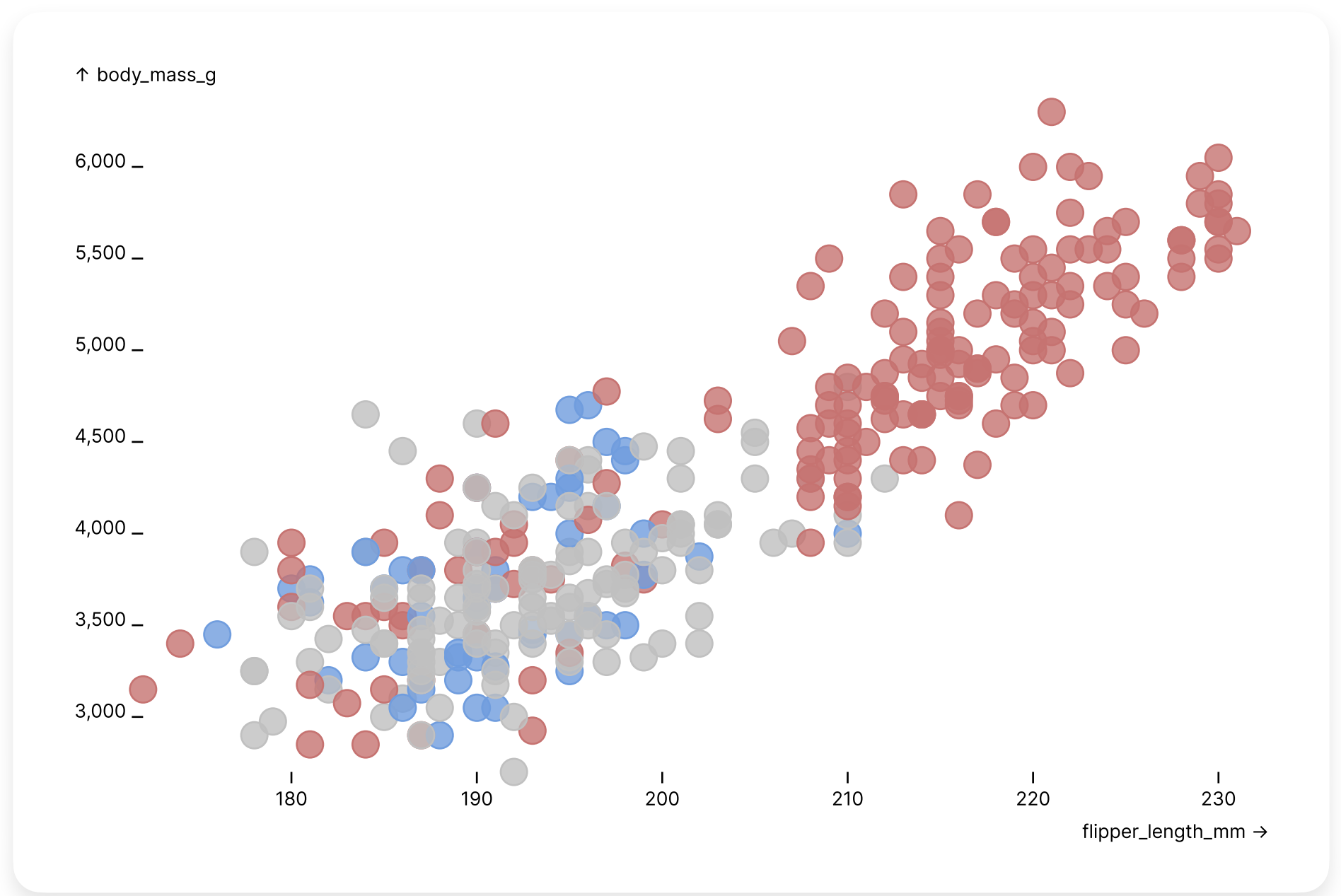
A user **fills in the holes** ("??") in the partial program to produce a new visualization.

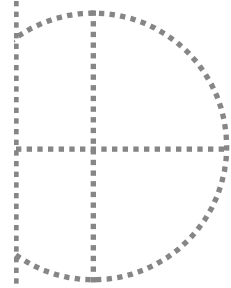
2.



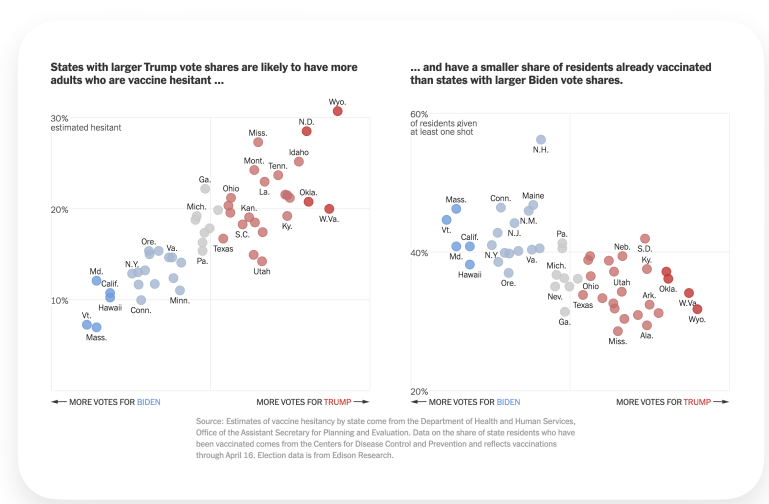
3.

```
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
    })
  ],
});
```





1.



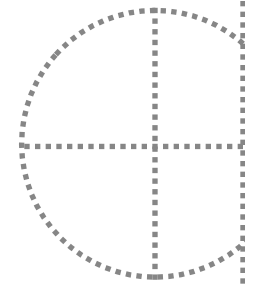
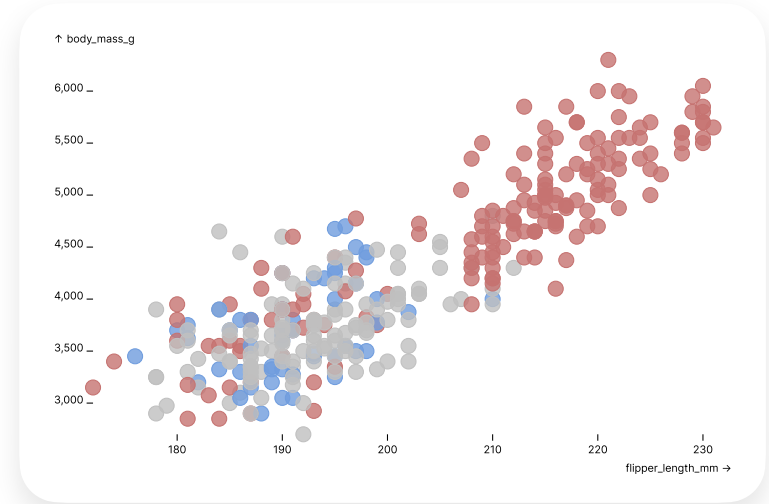
2.

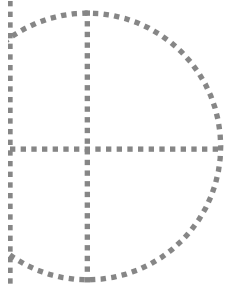


3.

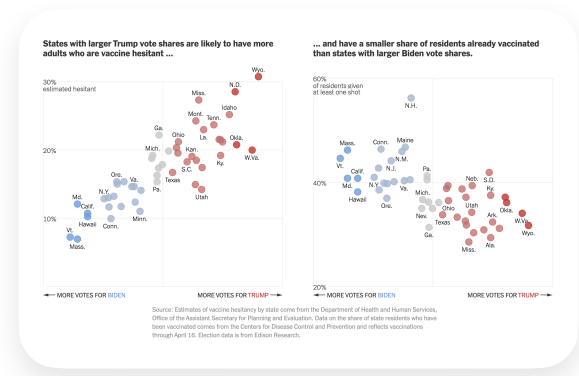
```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE", "#A7B9D3", "#C23734"]
  },
  marks: [
    Plot.dot(data, {
      x: "??",
      y: "??",
      fill: "??",
      fillOpacity: 1,
      stroke: "??",
      strokeOpacity: 1,
      strokeWidth: 1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

4.





1.



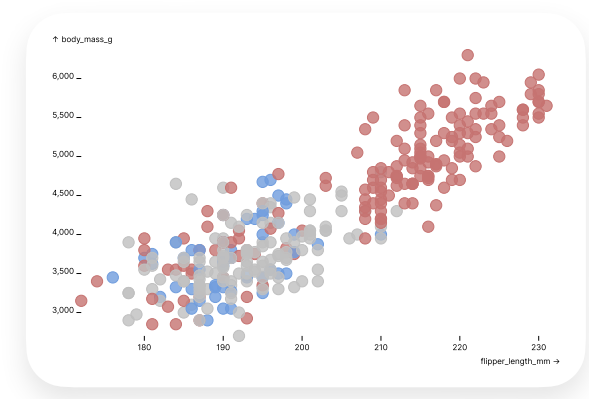
2.

reviz

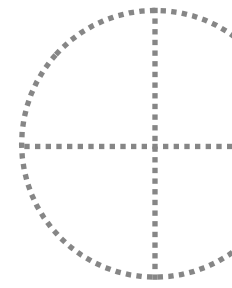
3.

```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#667371", "#ccc", "#7890DE", "#A799D3", "#C23734"]
  },
  marks: [
    Plot.dot(data, {
      x: "???",
      y: "???",
      fill: "???",
      fillOpacity: 1,
      stroke: "???",
      strokeOpacity: 1,
      strokeWidth: 1
    })
  ]
});
```

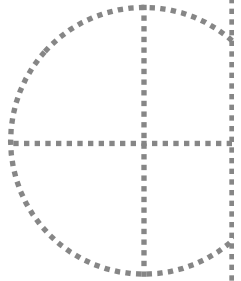
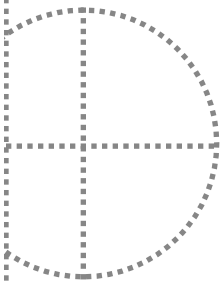
4.



What **actually happens** in this step?



It all starts with a **walk**.



Visit each node in the `svg` subtree.

```
<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" />
```

...

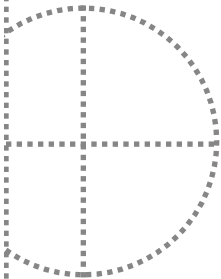
```
<g aria-label="x-axis tick label" fill="none" stroke="currentColor">
```

```
<text y="0.71em" transform="translate(56, 420)" stroke="currentColor">10</text>
```

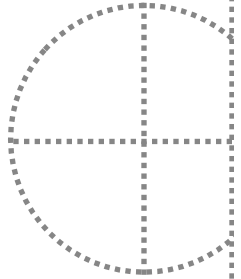
```
<text y="0.71em" transform="translate(56, 420)" stroke="currentColor">20</text>
```

...

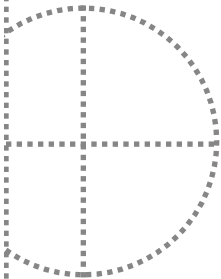
It all starts with a **walk**.



Visit each node
in the **svg** subtree.



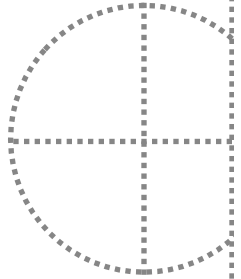
It all starts with a **walk**.



```
<svg viewBox="0 0 493.5 450" width="493.5" height="450">  
  <g aria-label="dot">  
    <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>  
    ...  
</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

cy → 349

r → 7

fill → "#C67371"

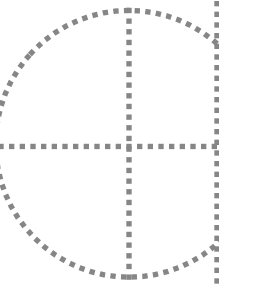
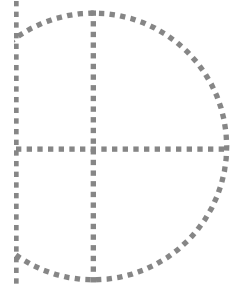
fill-opacity → 0.8

stroke → "#C67371"

stroke-opacity → 1

stroke-width → 2px

Styles	Computed
▶ 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 , ... }`

`cy: { cy → 349 , ... }`

`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 }`

Bar Chart

`hasMarkType("rect")`

`hasConsistentGeomAttr("width")`

`hasXScaleType("discrete")`

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

Bar Chart (0/3 predicates)

`hasMarkType("rect")` ::::

`hasConsistentGeomAttr("width")` ::::

`hasXScaleType("width")` ::::

Scatterplot (2/2 predicates)

`hasMarkType("circle")` ::::

`hasConsistentGeomAttr("r")` ::::

Strip Plot (2/3 predicates)

`hasMarkType("circle")` ::::

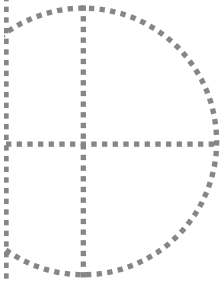
`hasConsistentGeomAttr("r")` ::::

`hasSiblingsWithConsistentCyAttr` ::::

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", ... }

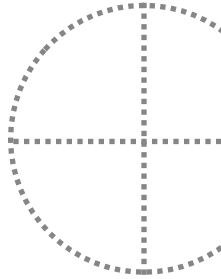
fill-opacity: { fill-opacity → 0.8 }

stroke: { stroke → "#C67371", ... }

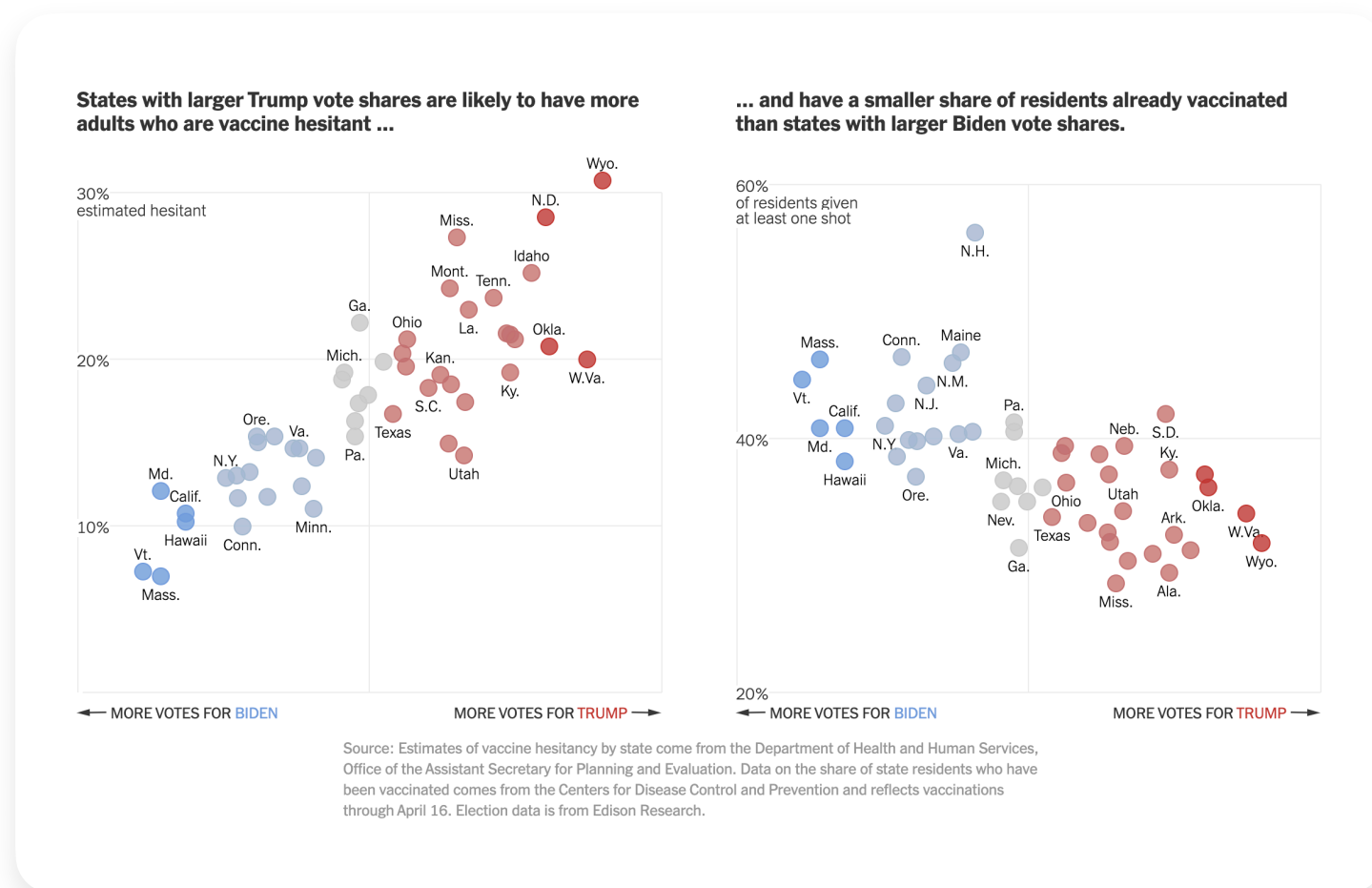
stroke-opacity: { stroke-opacity → 1 }

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



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

Input Visualization

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"]
}
```

Intermediate Representation



```
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
    })
  ],
});
```

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

```
`${EVAL_HOLE}`
```

Insertion point for
generated code

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"]  
}
```

Intermediate Representation

evalType

```
`${EVAL_HOLE}`
```

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

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"]  
}
```

Intermediate Representation

evalGeomAttr

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

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

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"]  
}
```

Intermediate Representation

evalPresAttr

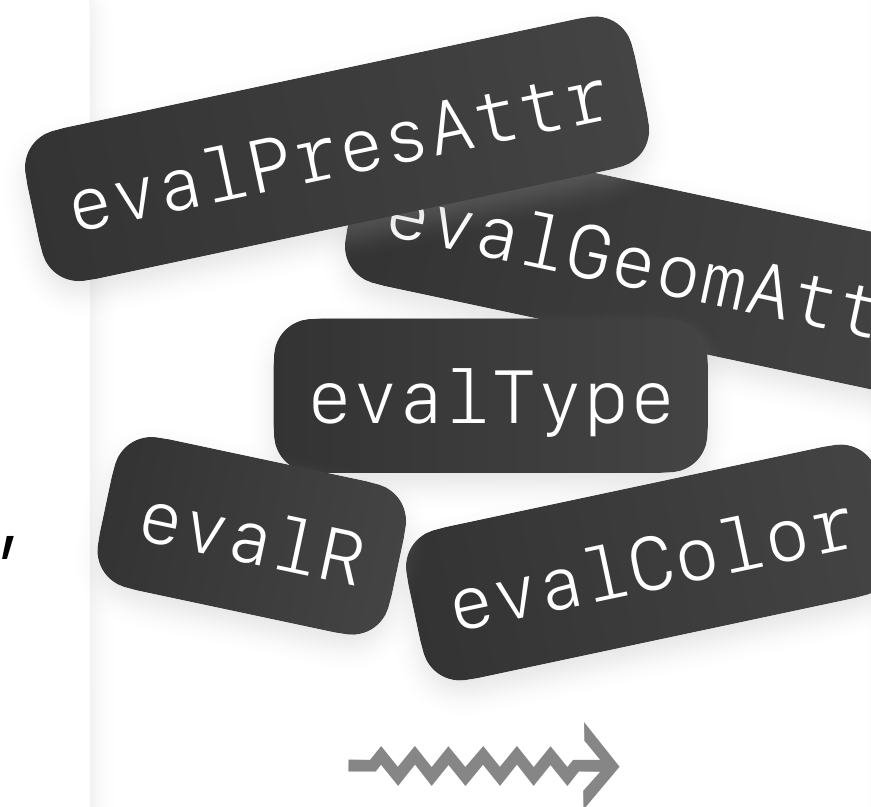
```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  r: 7,  
  ${EVAL_HOLE}  
})`
```

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

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"]  
}
```

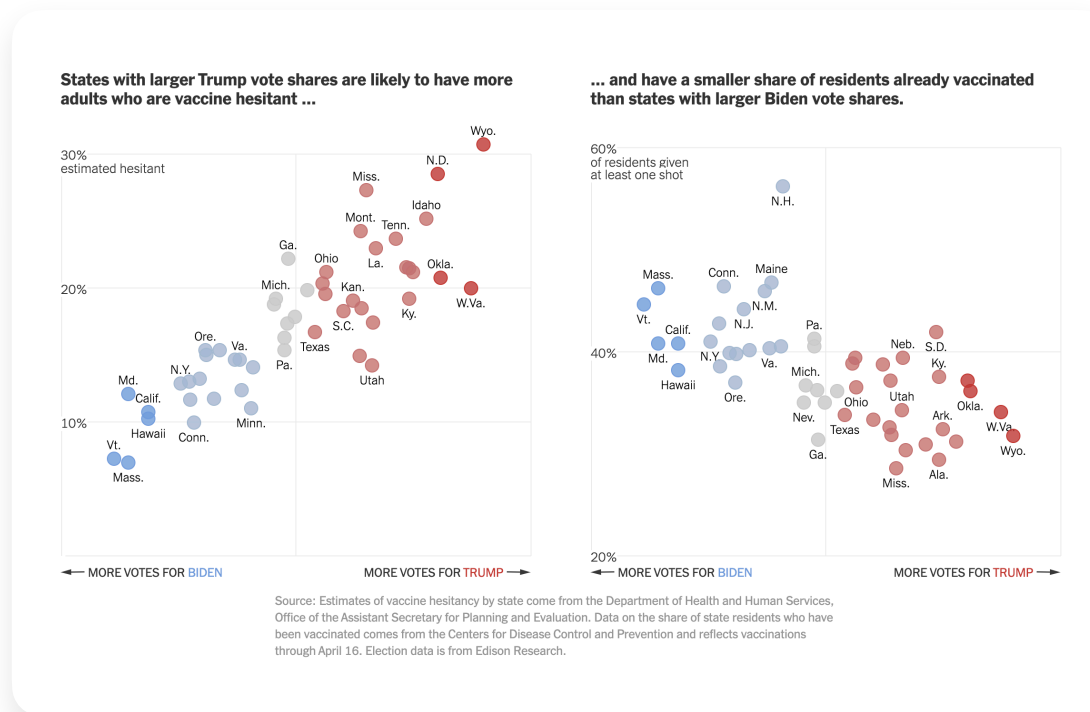


```
`Plot.dot(data, {  
  x: "${PROGRAM_HOLE}",  
  y: "${PROGRAM_HOLE}",  
  r: 7,  
  fill: "${PROGRAM_HOLE}",  
  fillOpacity: 0.8,  
  stroke: "${PROGRAM_HOLE}",  
  strokeOpacity: 1,  
  strokeWidth: "2px"  
})`
```

Intermediate Representation

Partial Program

Input Visualization



Attribute Sets

cx → 366

r → 7

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

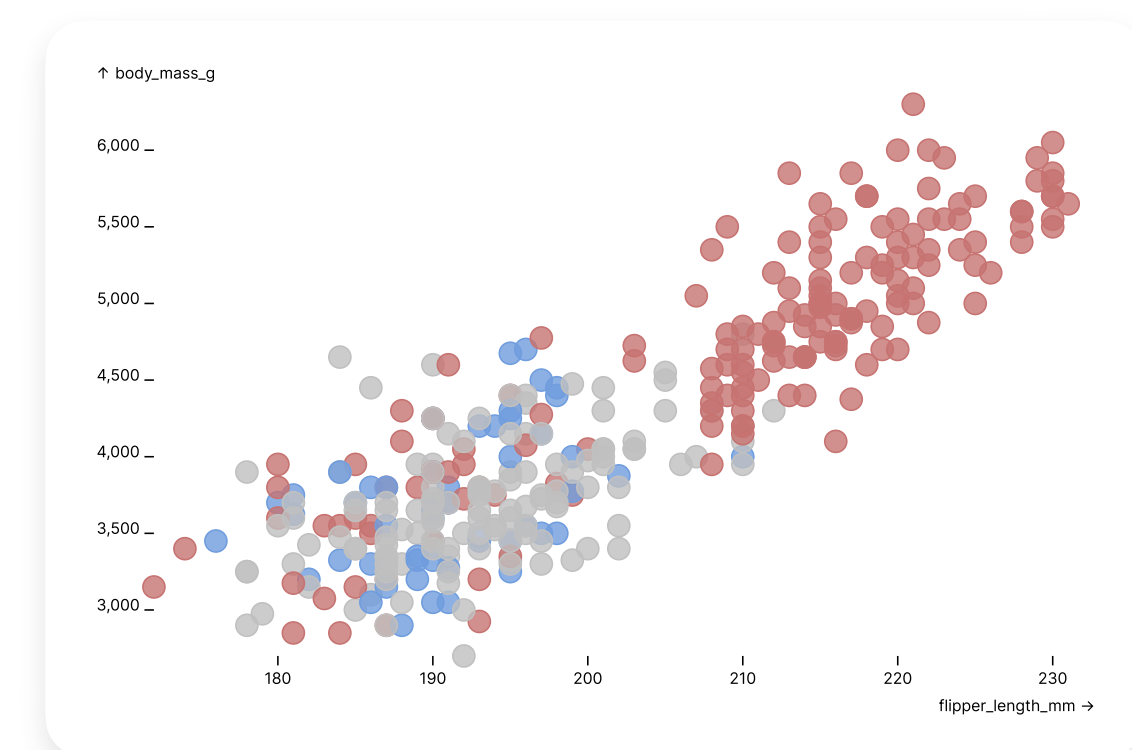
evalR

evalColor

Partial Program

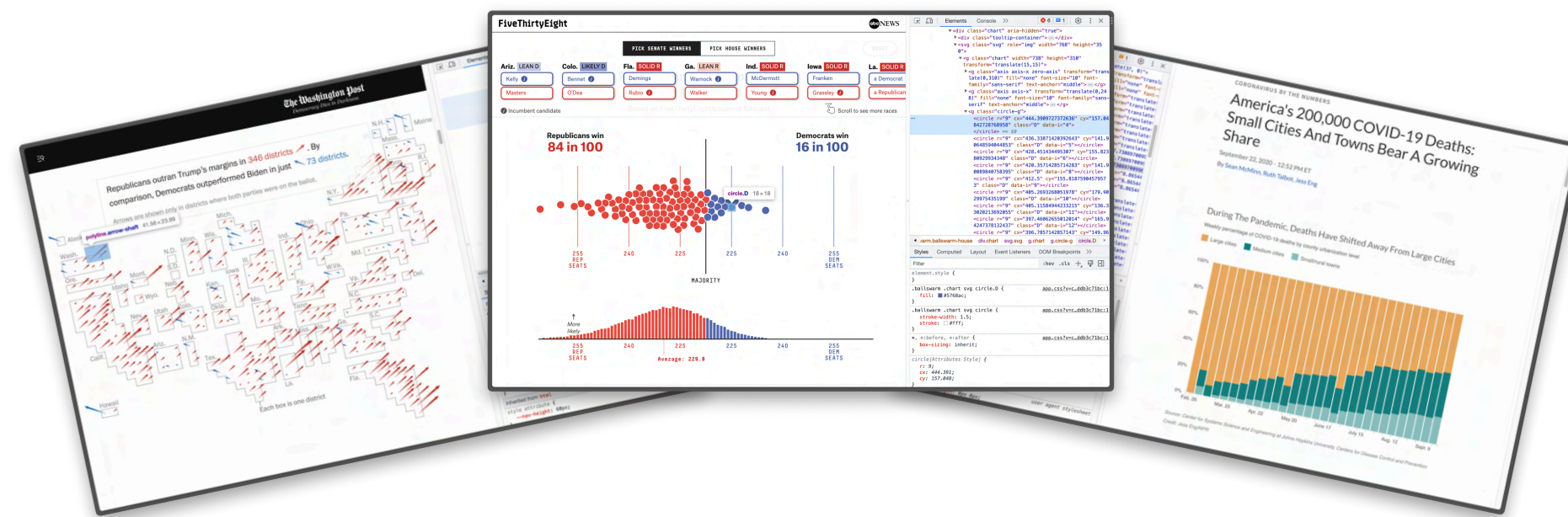
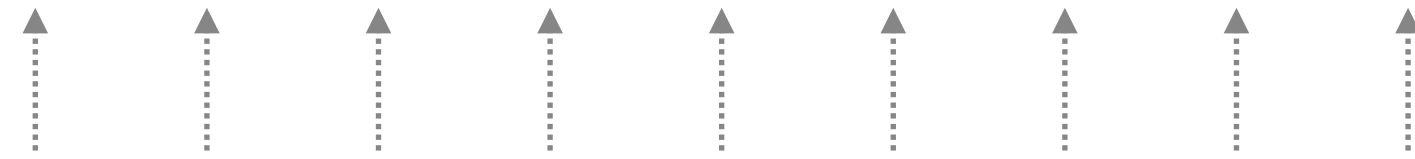
```
const plot = Plot.plot({
  color: {
    type: "categorical",
    range: ["#C67371", "#ccc", "#709DDE"]
  },
  marks: [
    Plot.dot(data, {
      x: "??",
      y: "??",
      fill: "??",
      fillOpacity: 1,
      stroke: "??",
      strokeOpacity: 1,
      strokeWidth: 1
    })
  ],
});
```

Output Visualization

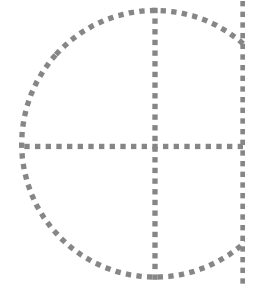
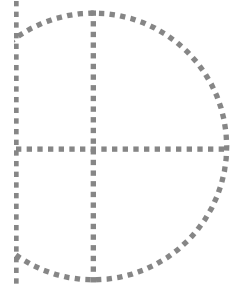
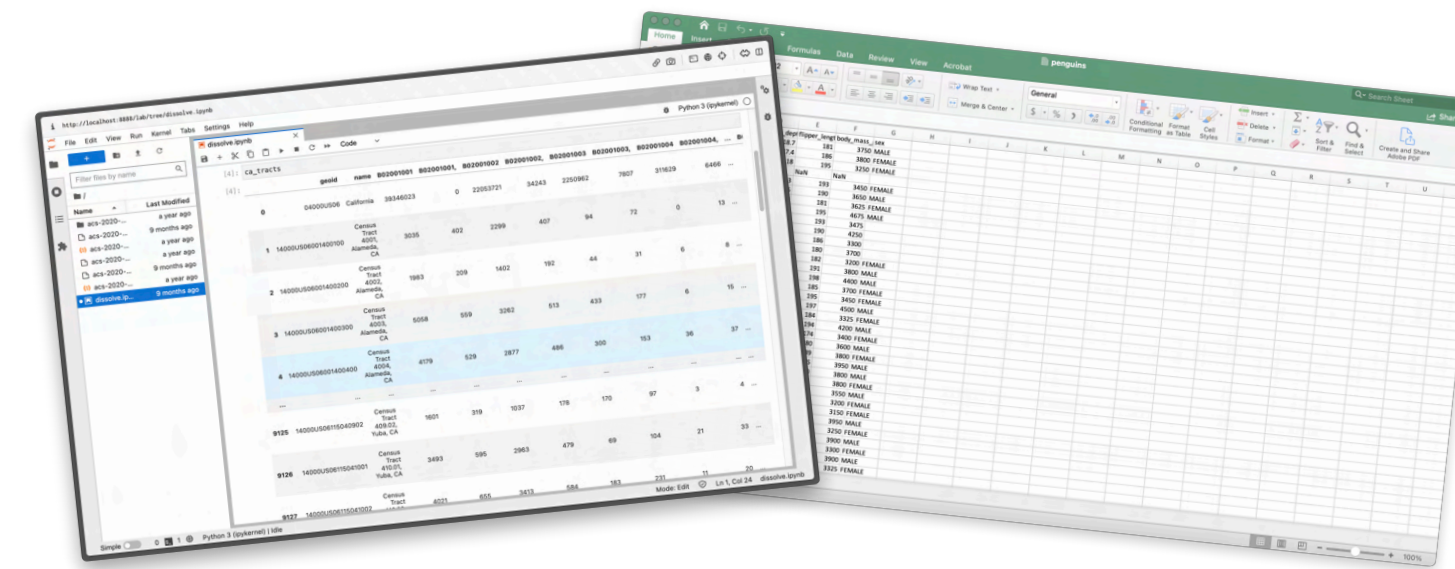


Working Practice

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



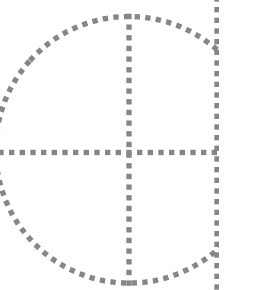
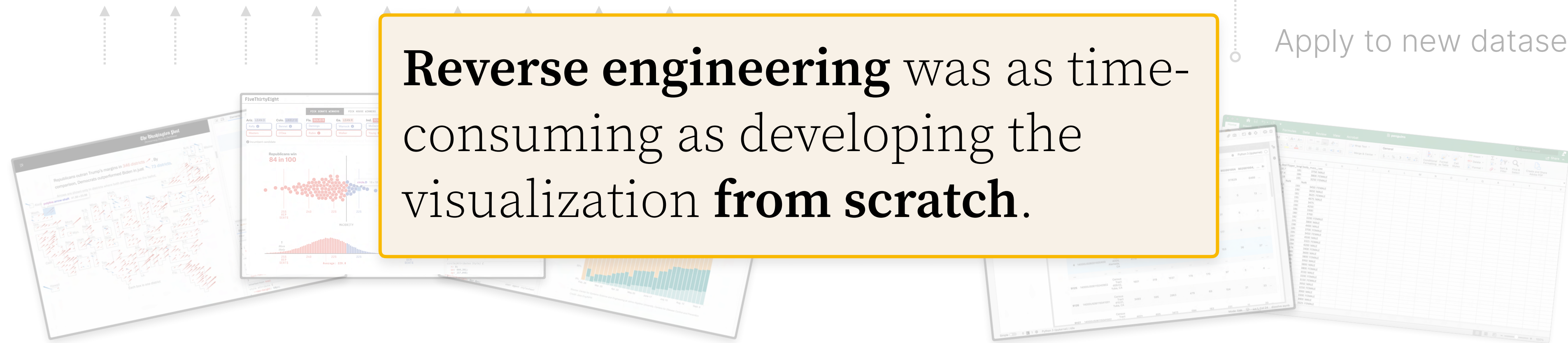
Apply to new datasets





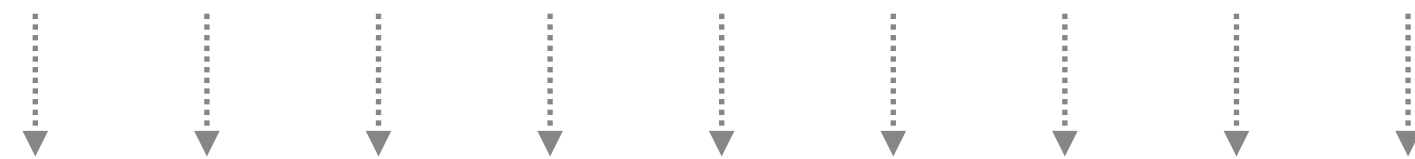
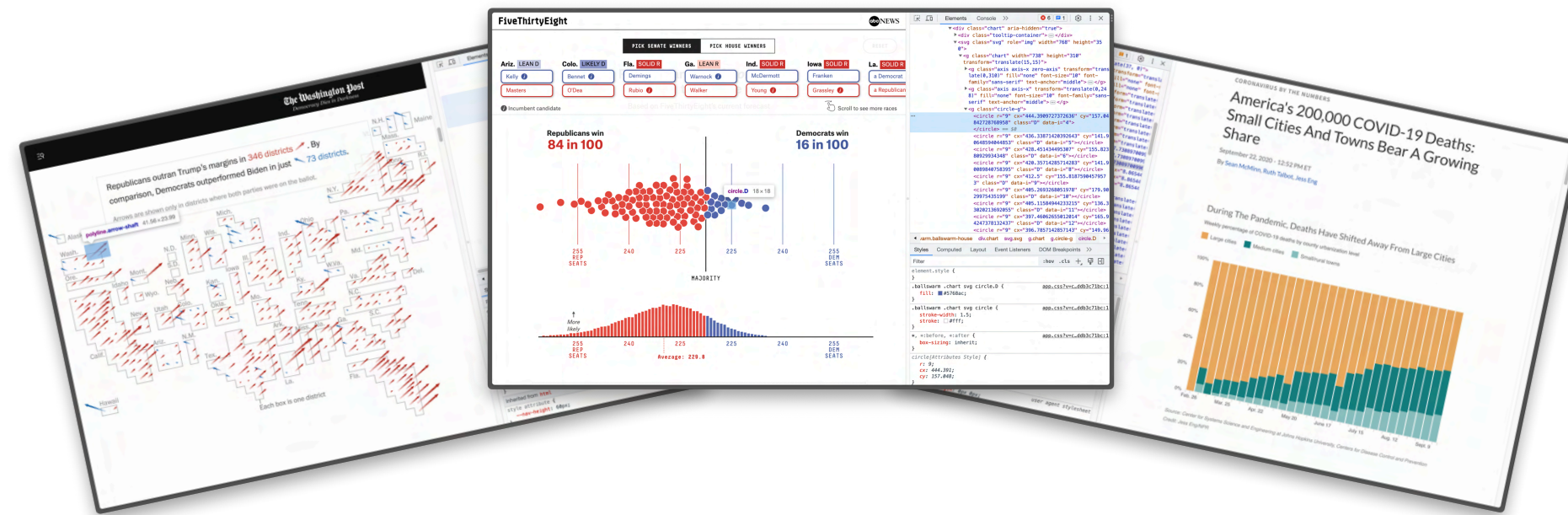
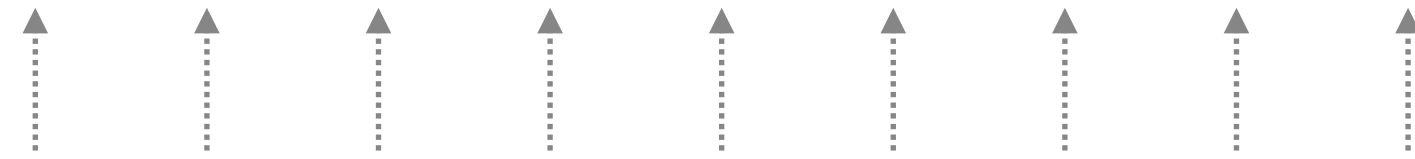
Working Practice

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



Working Practice

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

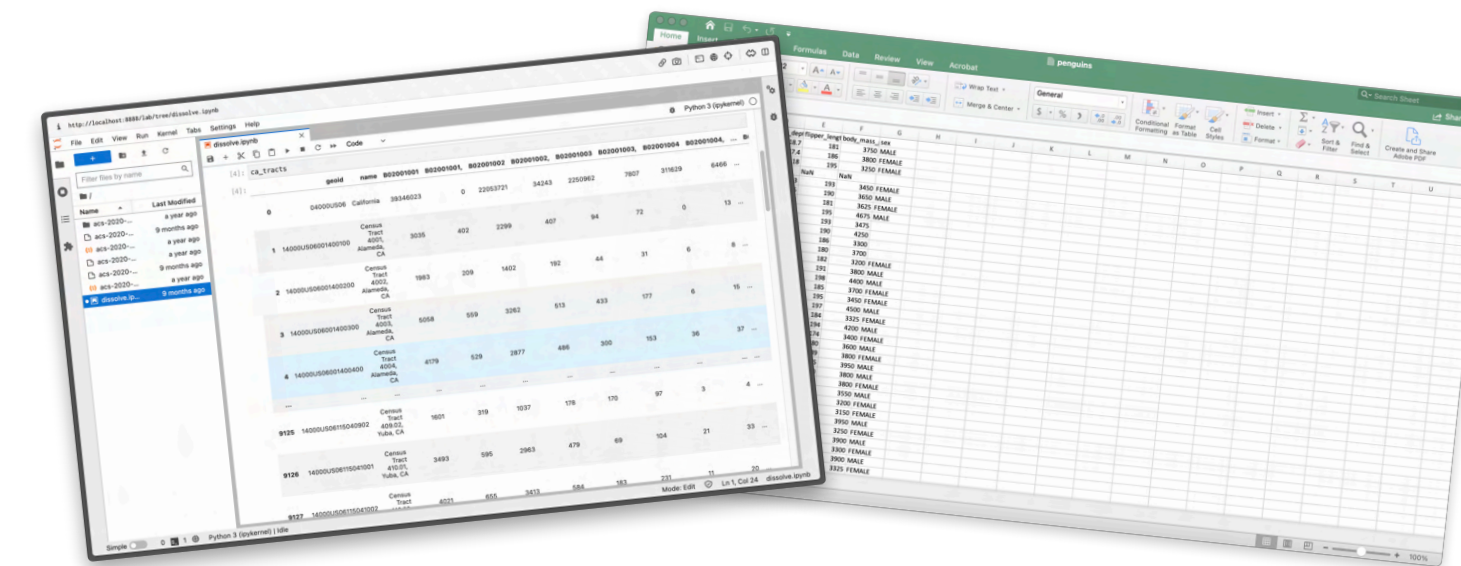


Automatically compile **partial programs** from **examples**

Milliseconds

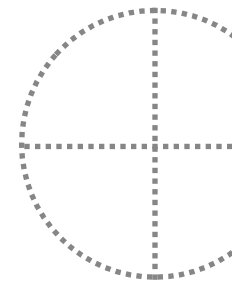
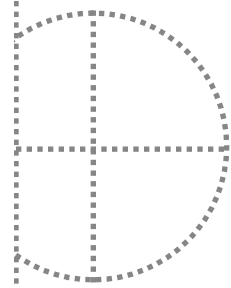
Hours

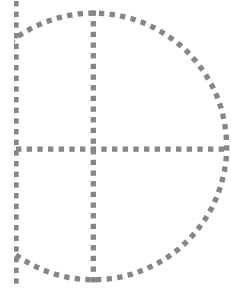
Apply to new datasets



Feed in any dataset

```
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
    })
  ]
});
```



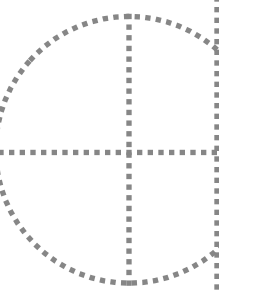


“

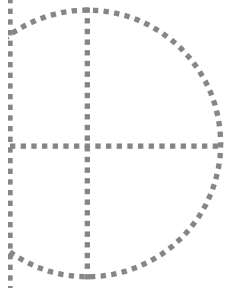
Don't commit to a specific visual form before seeing your data in it. A given visual form — say the pie chart or treemap — isn't “good” or “bad” in an absolute sense, but it may or may not be appropriate to your data and the specific question you want answered.

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”*



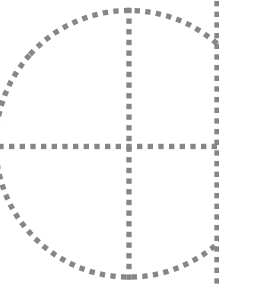
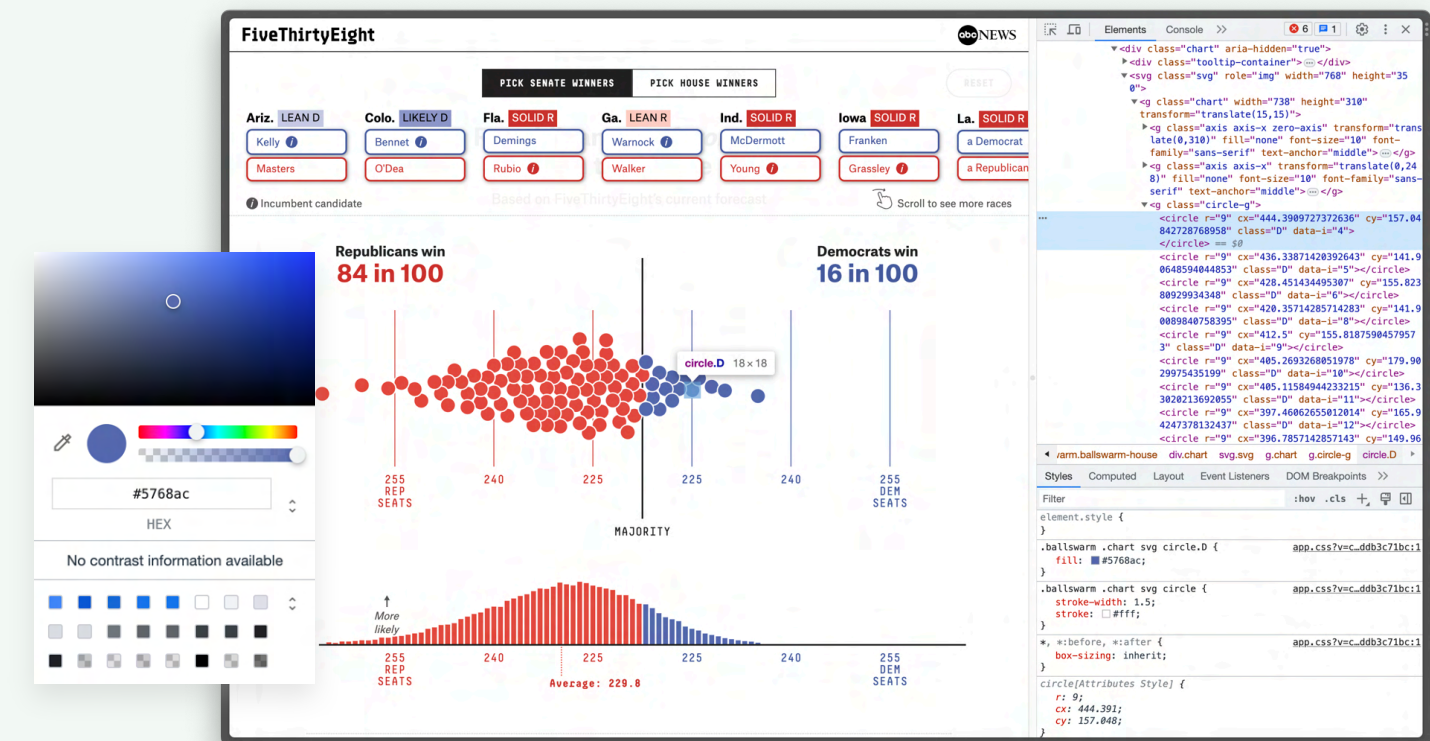
Lexer

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

Parser



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



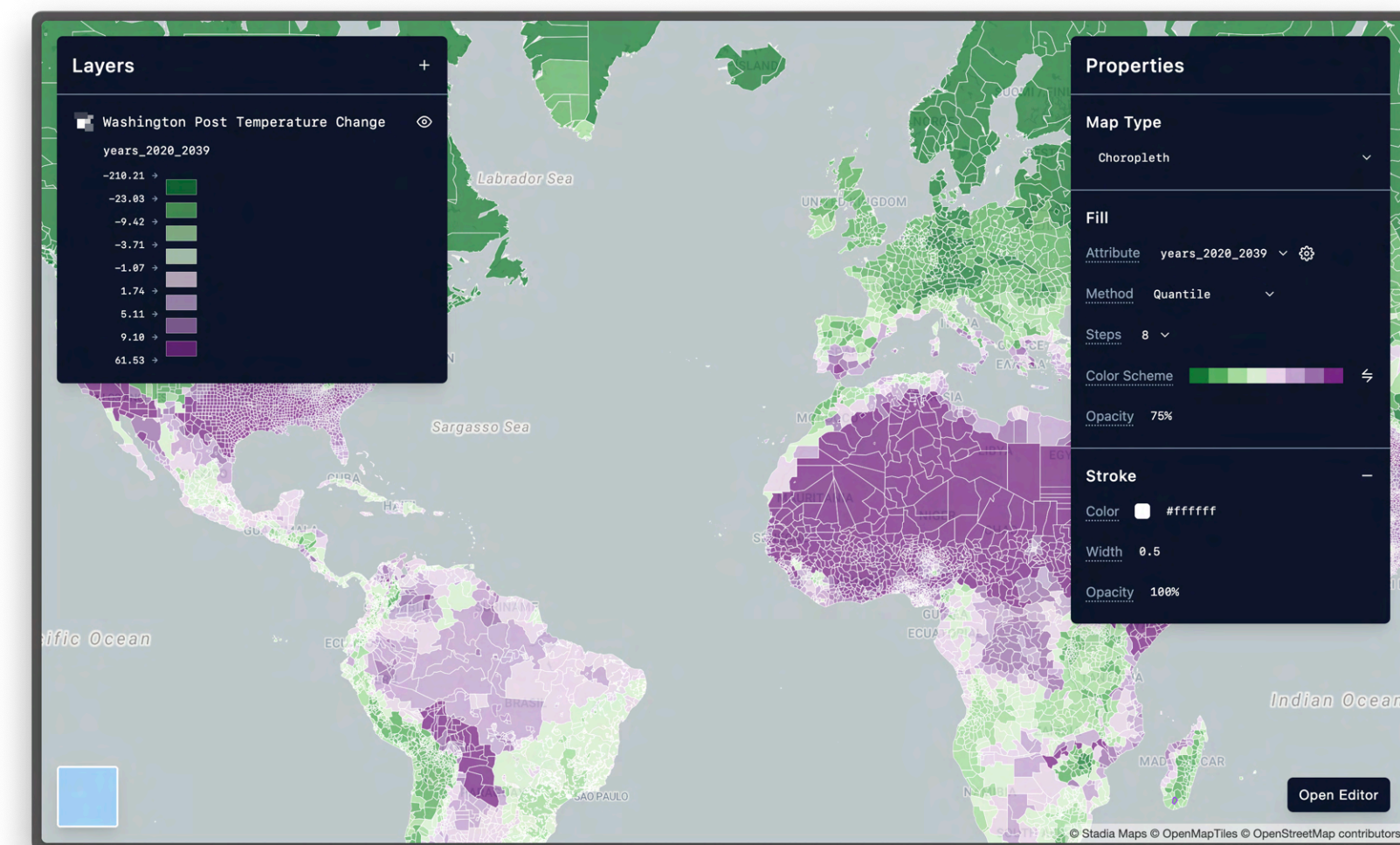
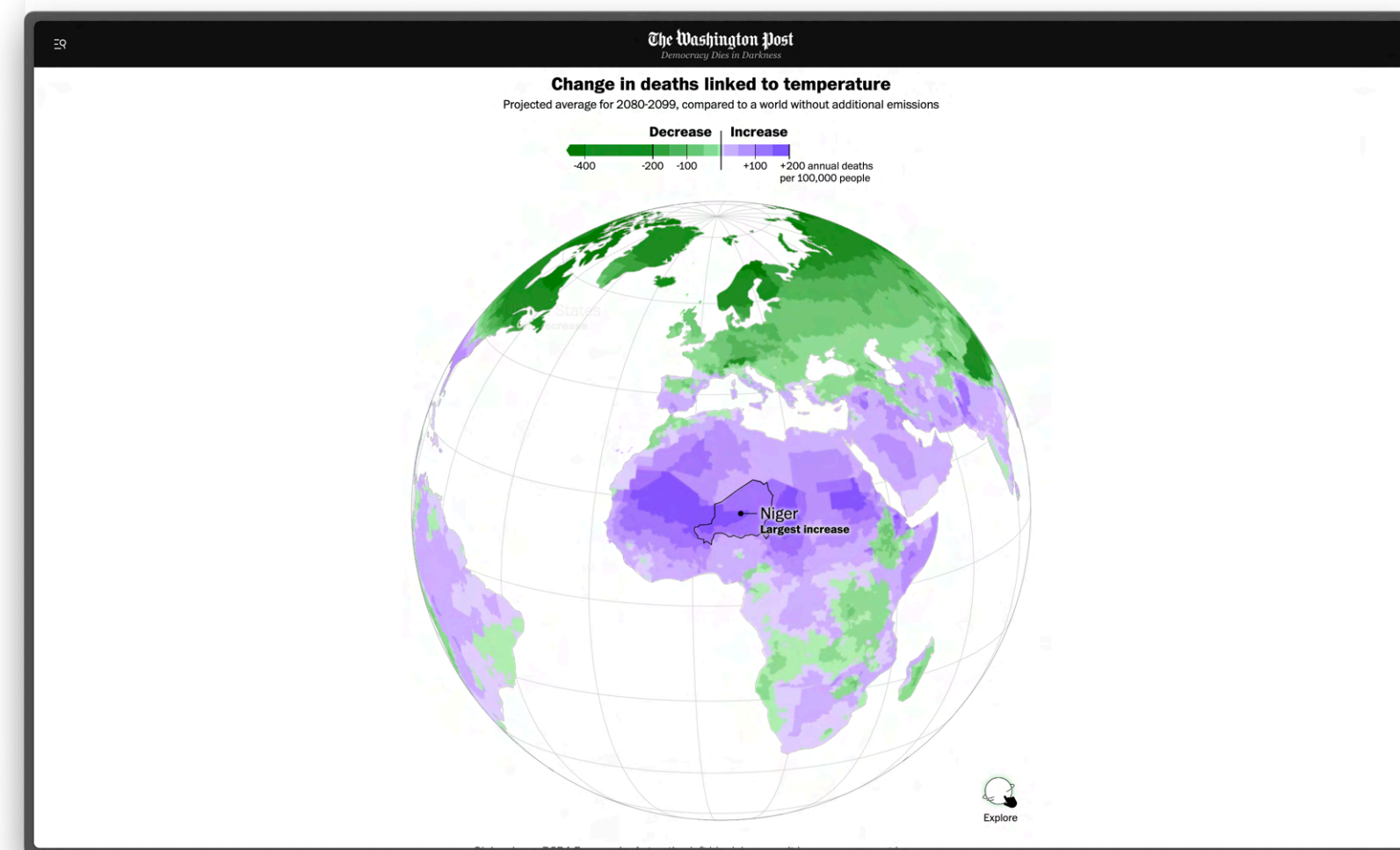


cartokit

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



Demo



The Washington Post

Will global warming make temperature less deadly?

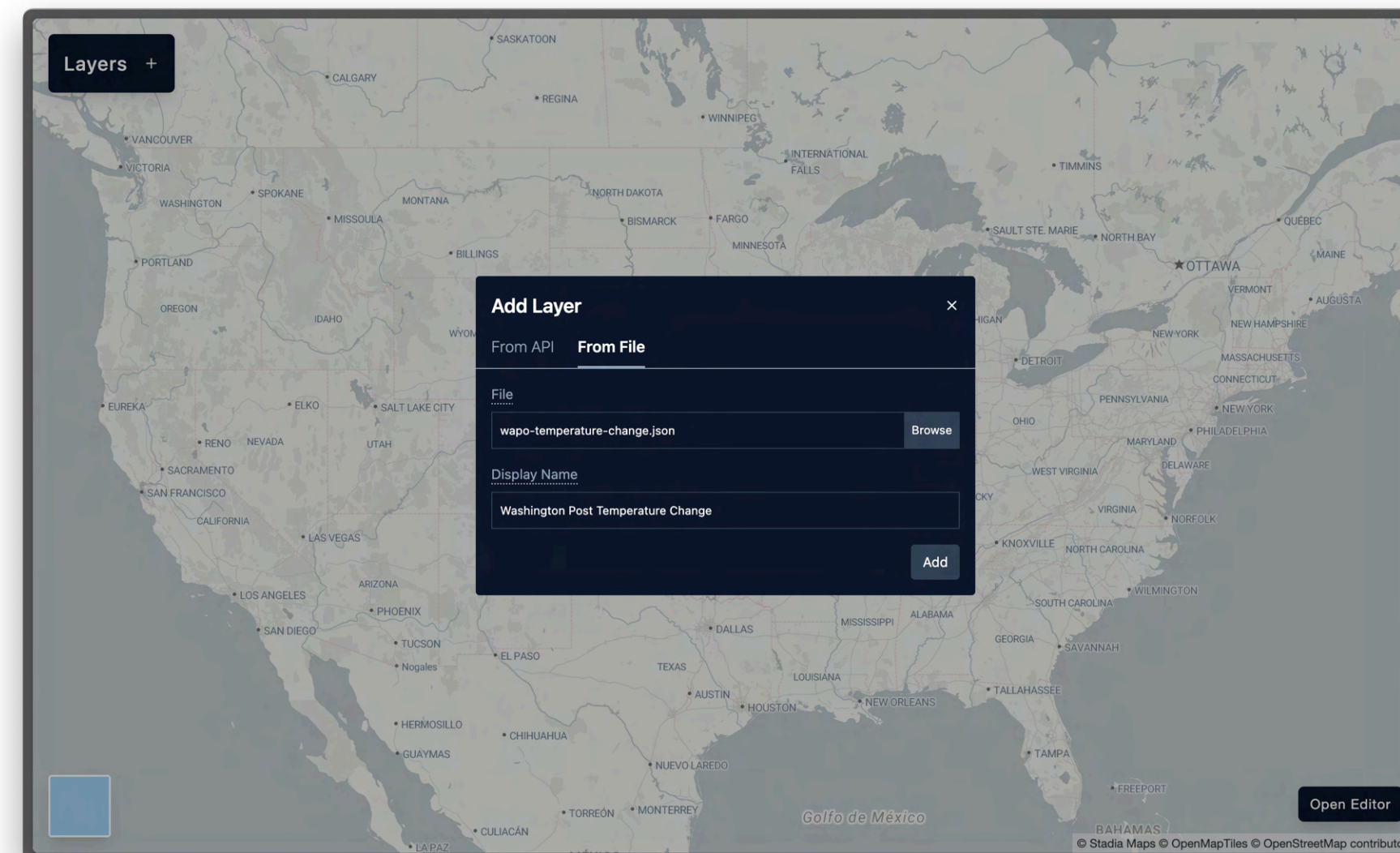
Harry Stevens

cartokit

A reproduction of the map in cartokit.

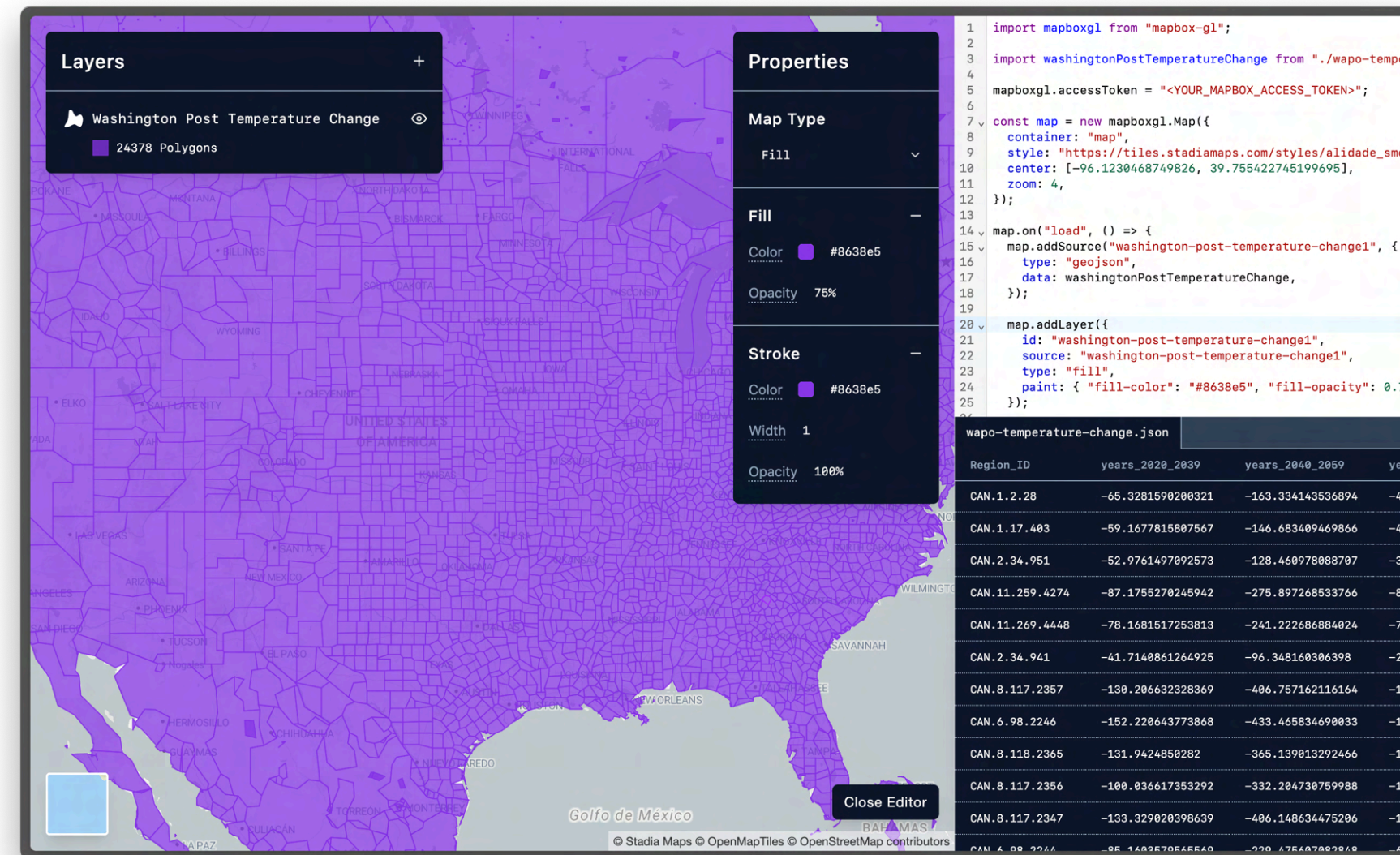
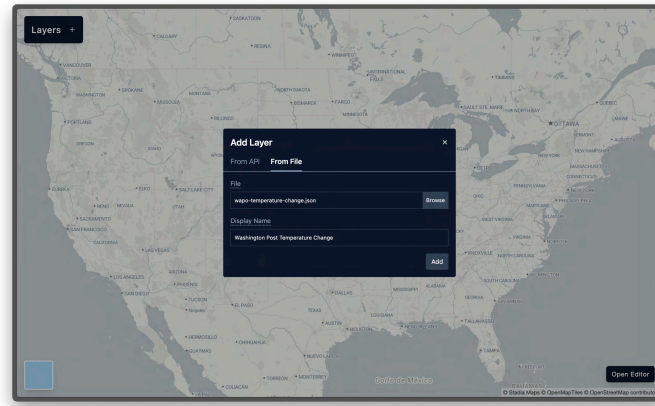
Can we author programs through **direct manipulation** of their outputs?

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



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

1.



The screenshot shows a map editor interface with a map of the United States. The map is filled with a purple color, representing a polygon layer. The interface includes a Layers panel on the left, a Properties panel on the right, and a code editor on the right side. The Layers panel shows a layer named "Washington Post Temperature Change" with 24378 Polygons. The Properties panel shows the Map Type set to "Fill", the Fill Color set to "#8638e5", the Fill Opacity set to 75%, the Stroke Color set to "#8638e5", the Stroke Width set to 1, and the Stroke Opacity set to 100%. The code editor shows the following code:

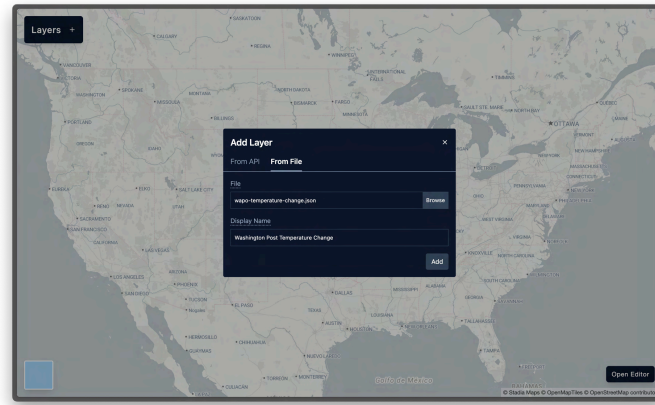
```
1 import mapboxgl from "mapbox-gl";
2 import washingtonPostTemperatureChange from "./wapo-temper
3 mapboxgl.accessToken = "<YOUR_MAPBOX_ACCESS_TOKEN>";
4
5 const map = new mapboxgl.Map({
6   container: "map",
7   style: "https://tiles.stadiamaps.com/styles/alidade_smooth
8   center: [-96.1238468749826, 39.755422745199695],
9   zoom: 4,
10 });
11
12 map.on("load", () => {
13   map.addSource("washington-post-temperature-change1", {
14     type: "geojson",
15     data: washingtonPostTemperatureChange,
16   });
17   map.addLayer({
18     id: "washington-post-temperature-change1",
19     source: "washington-post-temperature-change1",
20     type: "fill",
21     paint: { "fill-color": "#8638e5", "fill-opacity": 0.75
22   });
23 });
```

The code editor also shows a JSON file named "wapo-temperature-change.json" with the following data:

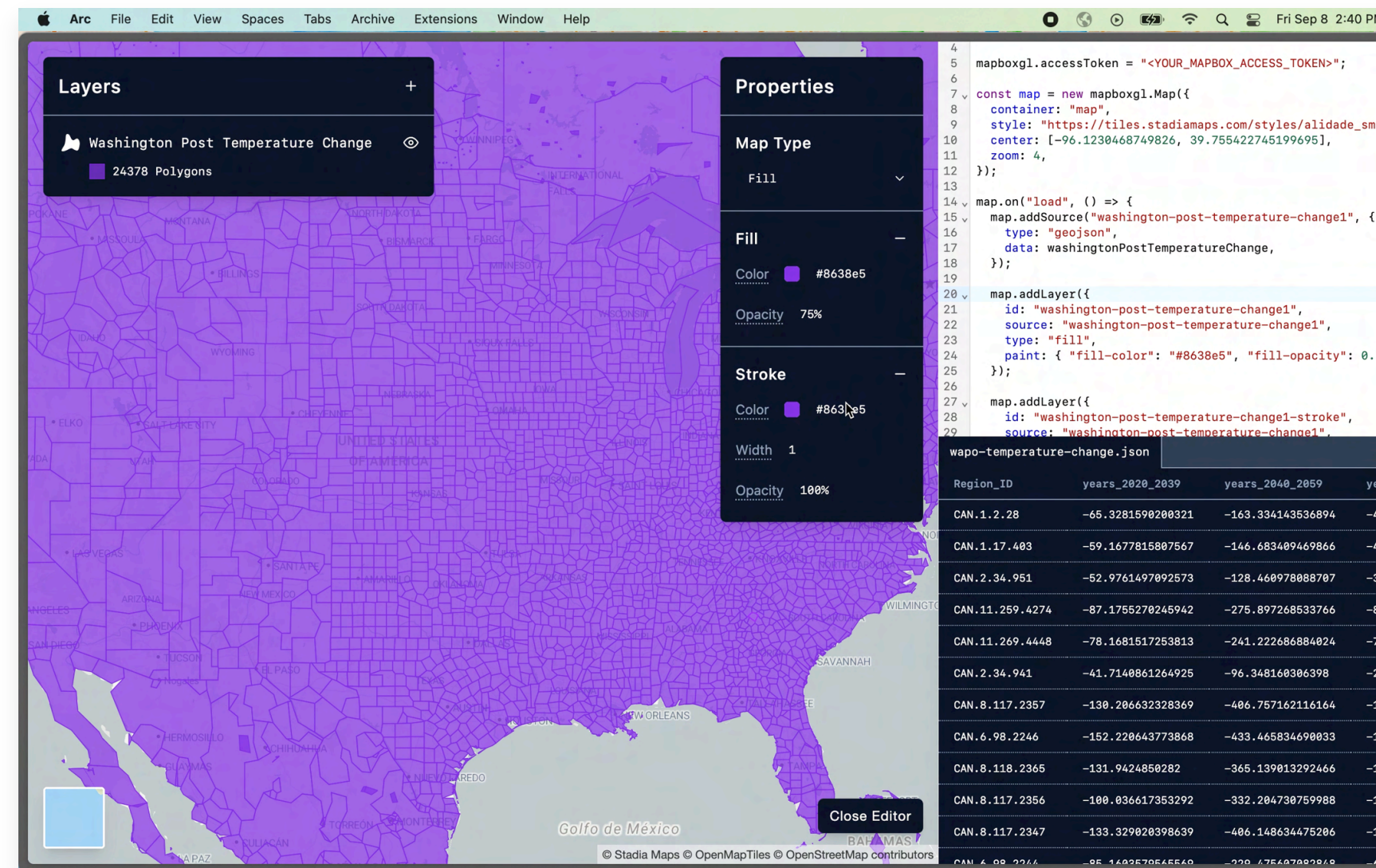
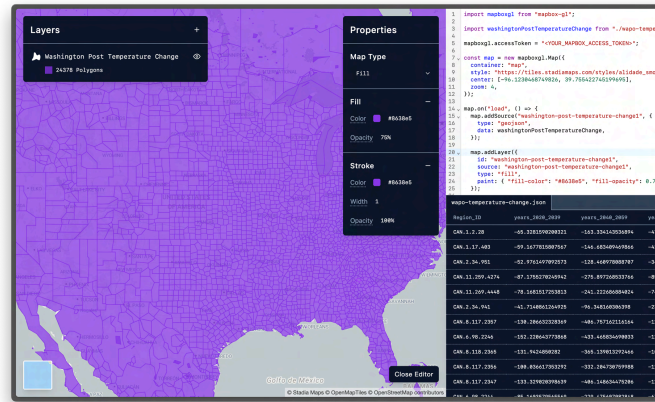
Region_ID	years_2020_2039	years_2040_2059	year
CAN.1.2.28	-65.3281599280321	-163.334143536894	-471
CAN.1.17.483	-59.1677815887567	-146.683489469866	-415
CAN.2.34.951	-52.9761497892573	-128.46897888787	-342
CAN.11.259.4274	-87.1755278245942	-275.897268533766	-858
CAN.11.269.4448	-78.1681517253813	-241.222686884824	-749
CAN.2.34.941	-41.7148861264925	-96.348168386398	-231
CAN.8.117.2357	-138.286632328369	-486.757162116164	-127
CAN.6.98.2246	-152.228643773868	-433.465834698833	-127
CAN.8.118.2365	-131.9424858282	-365.139813292465	-187
CAN.8.117.2356	-188.836617353292	-332.284738759988	-112
CAN.8.117.2347	-133.329828398639	-486.148634475286	-127
CAN.4.08.2366	86.1487670845540	310.176487881818	476

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

1.

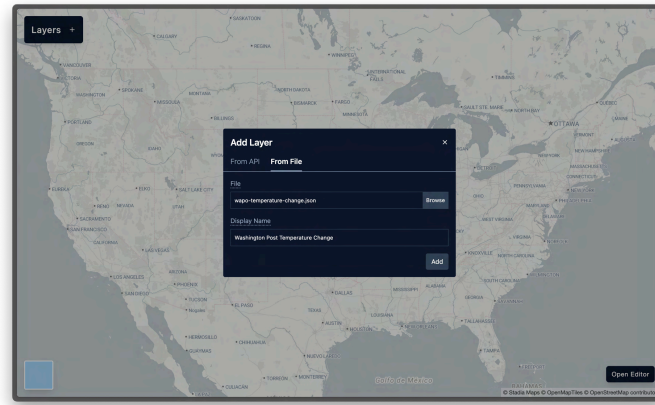


2.

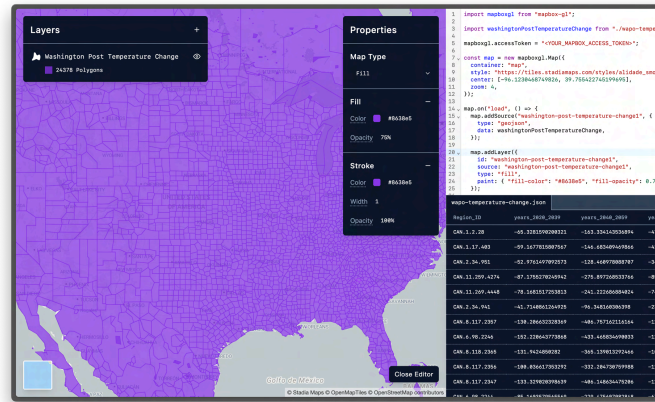


A user can **copy**, **modify**, and **deploy** the compiled program as desired.

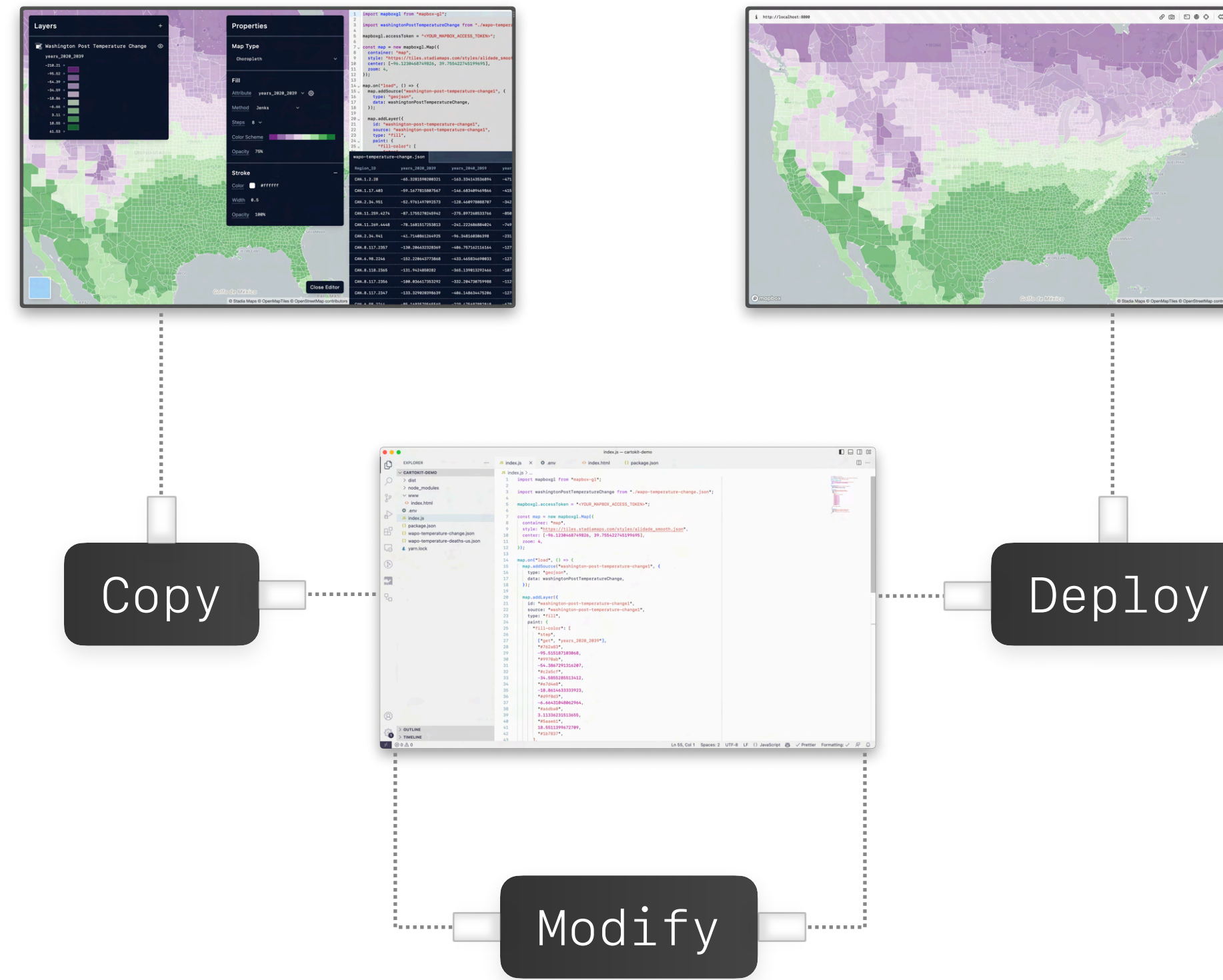
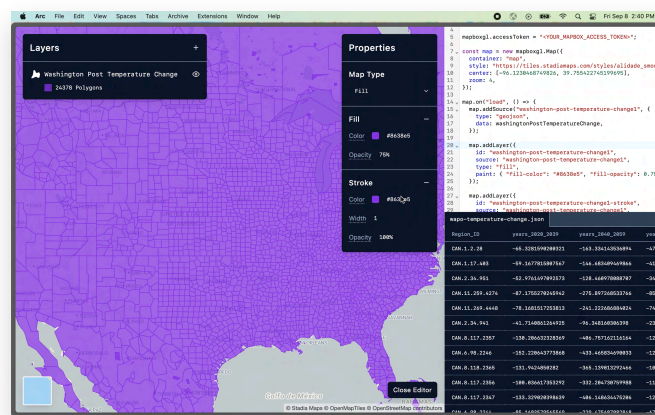
1.



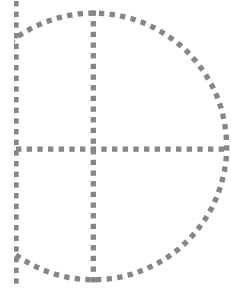
2.



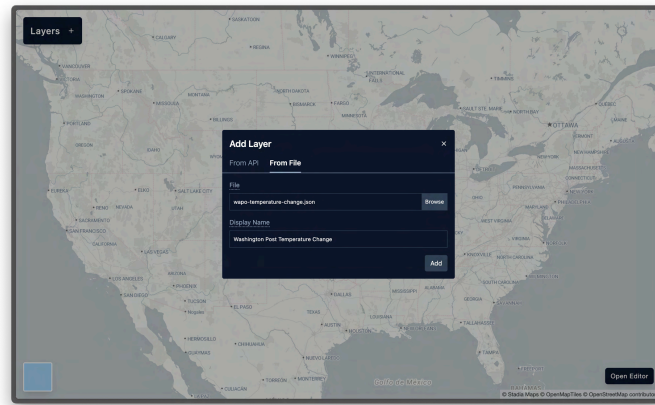
3.



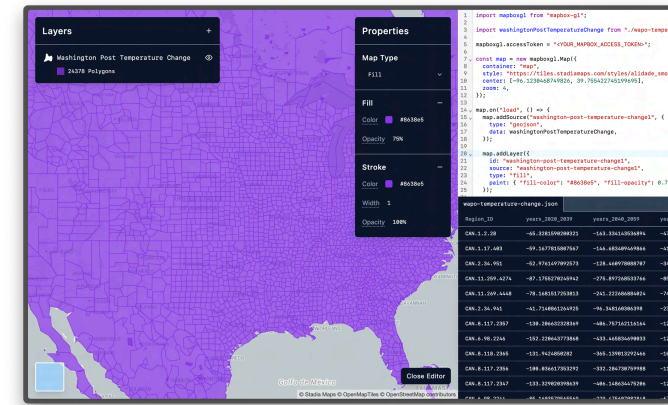
cartokit



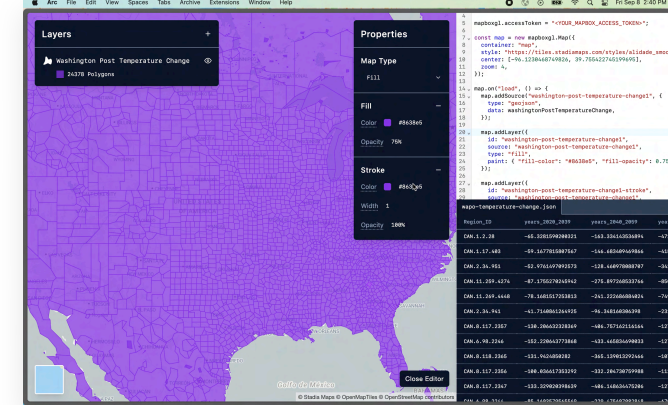
1.



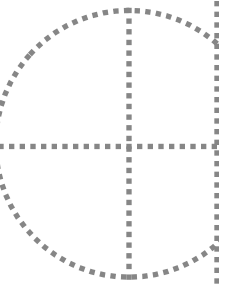
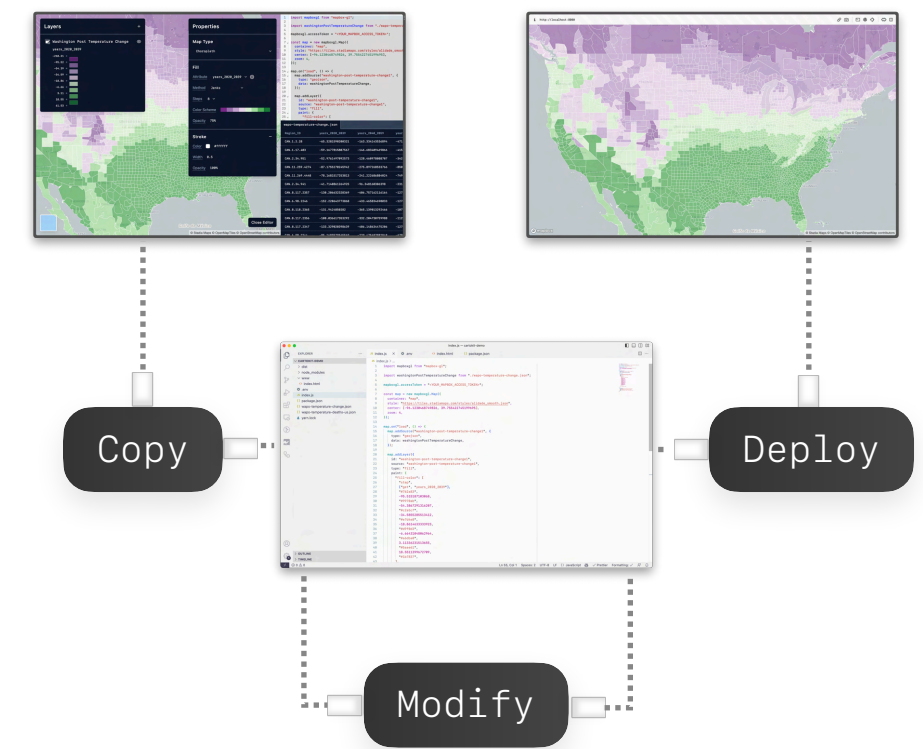
2.

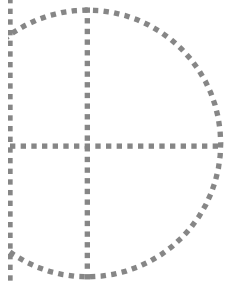


3.

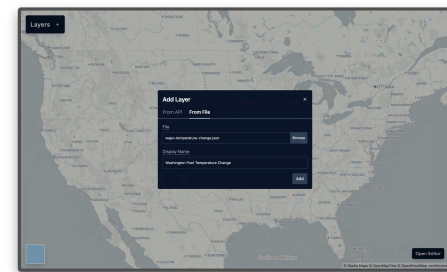


4.

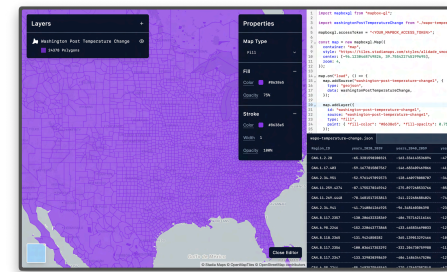




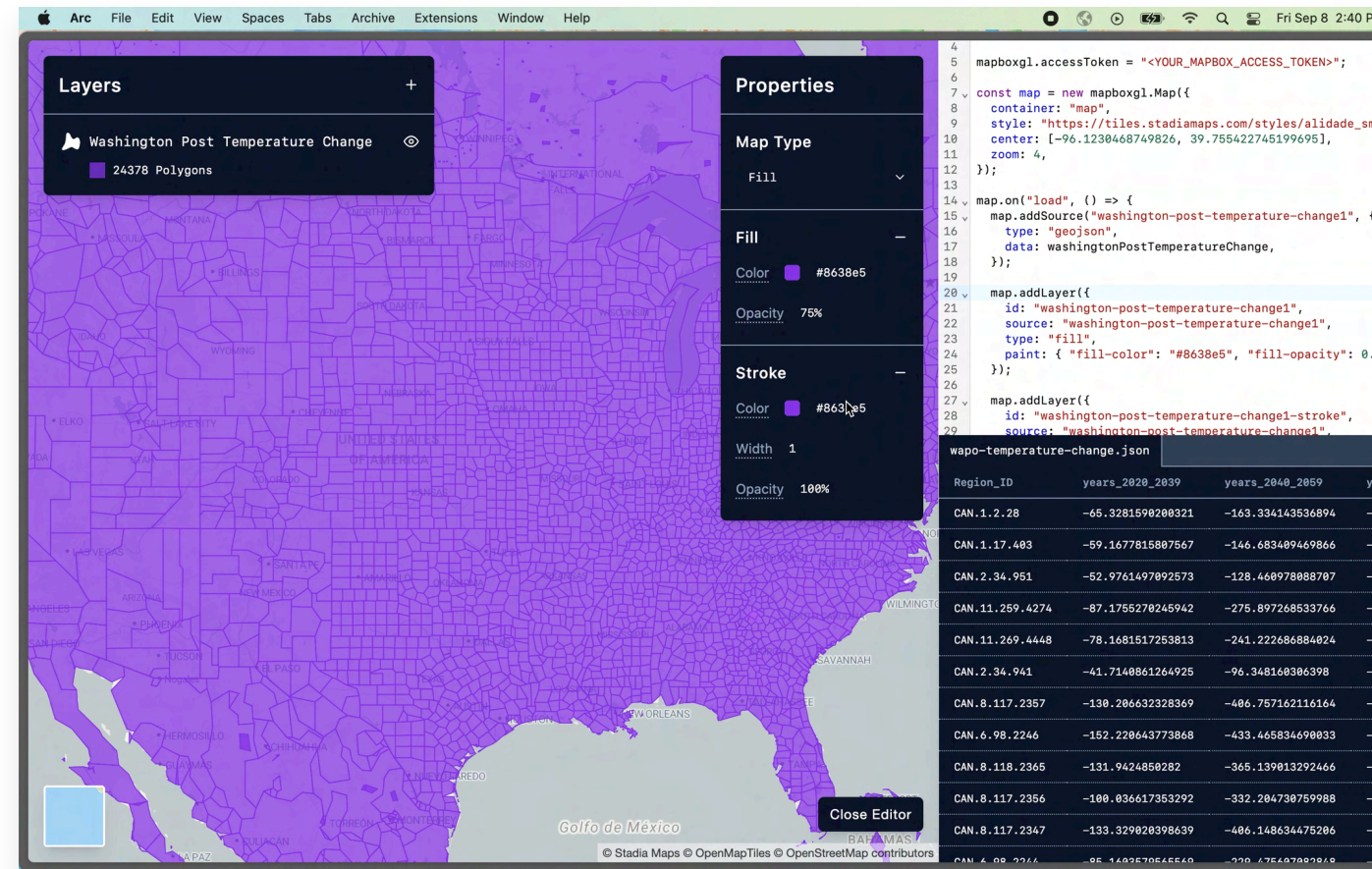
1.



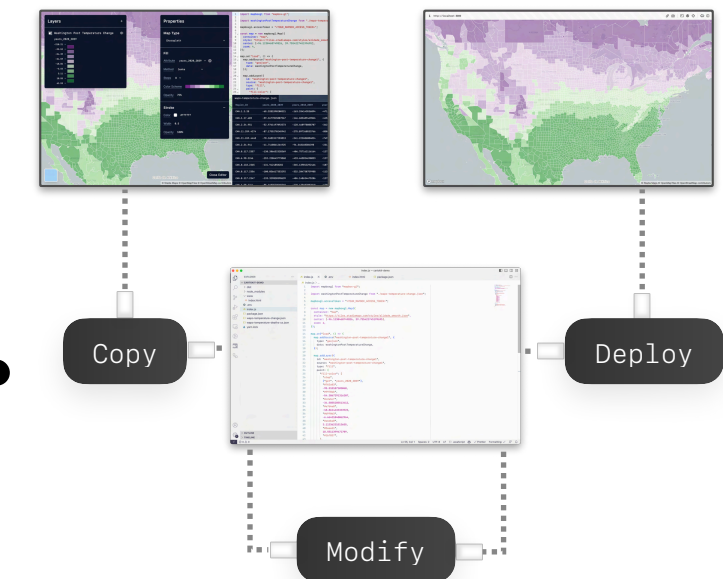
2.



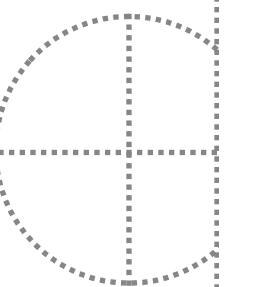
3.



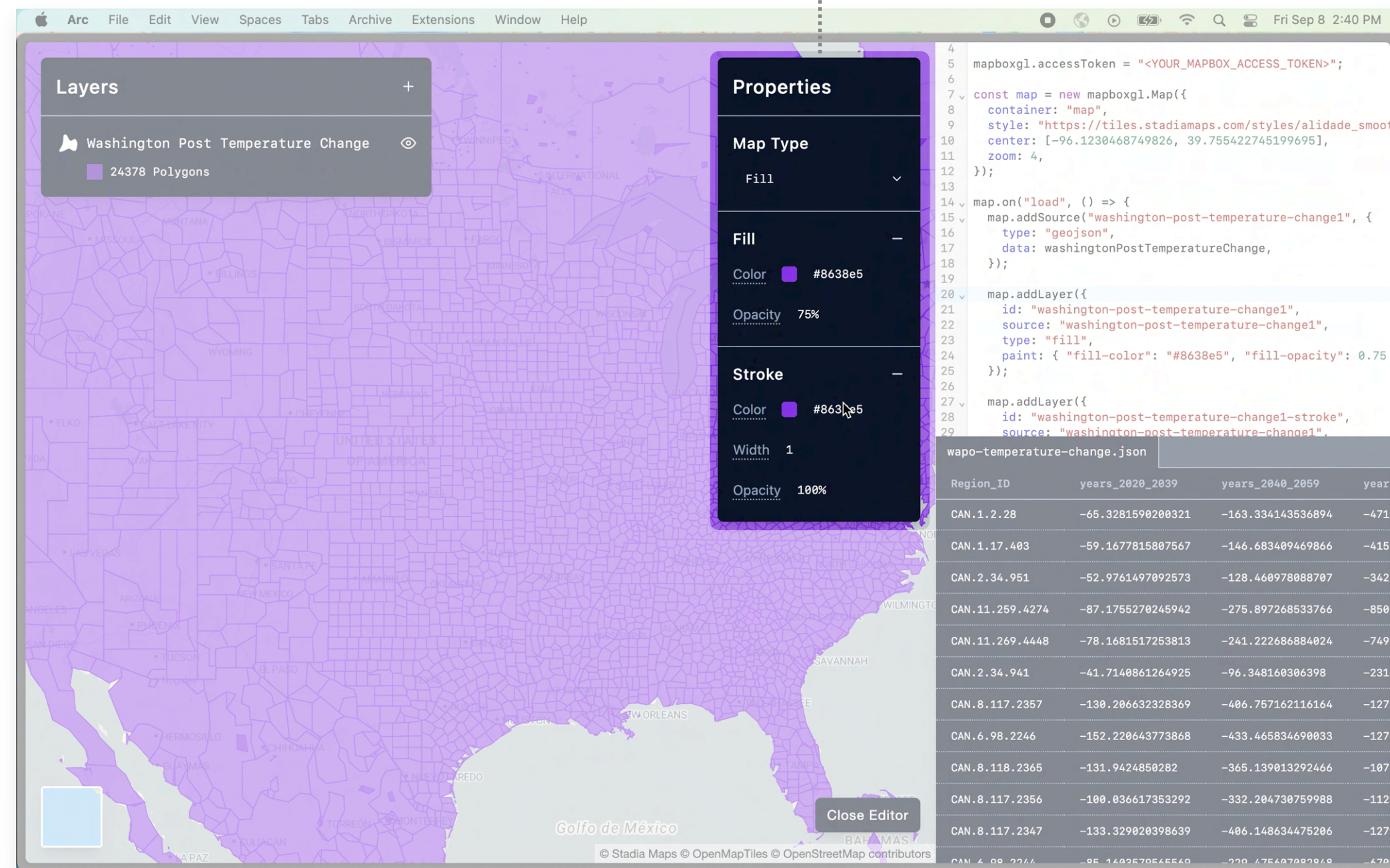
4.



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

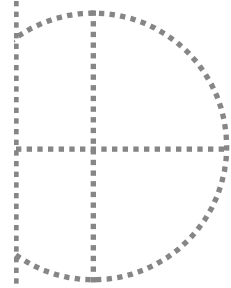


Map **property controls** in the UI to an **intermediate representation**.



```
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

```
1 import mapboxgl from "mapbox-gl";
2 import washingtonPostTemperatureChange from "./wapo-temper
3
4 mapboxgl.accessToken = "<YOUR_MAPBOX_ACCESS_TOKEN>";
5
6
7 const map = new mapboxgl.Map({
8   container: "map",
9   style: "https://tiles.stadiamaps.com/styles/aldide_smooth
10   center: [-96.1238468749826, 39.75842274519695],
11   zoom: 4,
12 });
13
14 map.on("load", () => {
15   map.addSource("washington-post-temperature-change1", {
16     type: "geojson",
17     data: washingtonPostTemperatureChange,
18   });
19   map.addLayer({
20     id: "washington-post-temperature-change1",
21     source: "washington-post-temperature-change1",
22     type: "fill",
23     paint: { "fill-color": "#8638e5", "fill-opacity": 0.75
24   });
25 });
26
```

Region_ID	years_2020_2039	years_2040_2059	year
CAN.1.2.28	-65.32815982989221	-163.334143536894	-471
CAN.1.17.483	-59.1677815887567	-146.683499469866	-415
CAN.2.34.951	-52.9761497892573	-128.468978888787	-342
CAN.11.289.4274	-87.1768278246942	-278.897268533766	-858
CAN.11.269.4448	-78.1681617253813	-241.22268884824	-749
CAN.2.34.941	-41.7148861264925	-96.348168386398	-231
CAN.8.117.2367	-138.286632328369	-486.757162116164	-127
CAN.6.98.2246	-152.228643773868	-433.465834698833	-127
CAN.8.118.2365	-131.9424858282	-365.139813292466	-187
CAN.8.117.2356	-188.836617353292	-332.284738789988	-112
CAN.8.117.2347	-133.329828396459	-486.148634475286	-127

dispatchLayerUpdate

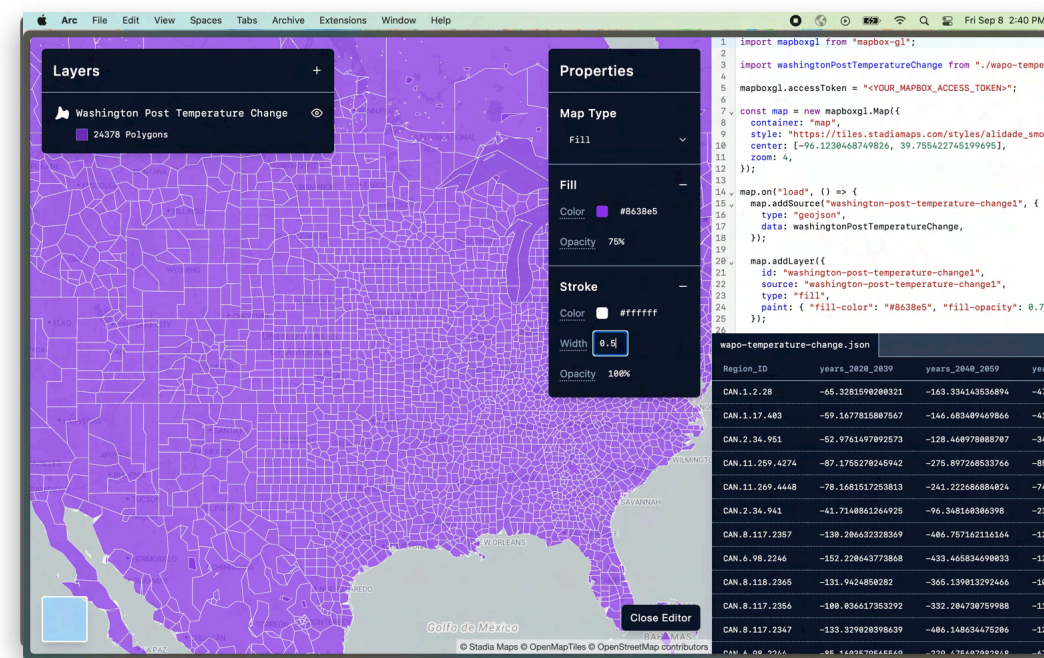
Choropleth Layer

years_2020_2039
-218.21
-23.83
-9.42
-3.71
-1.87
1.74
5.11
9.18
61.53



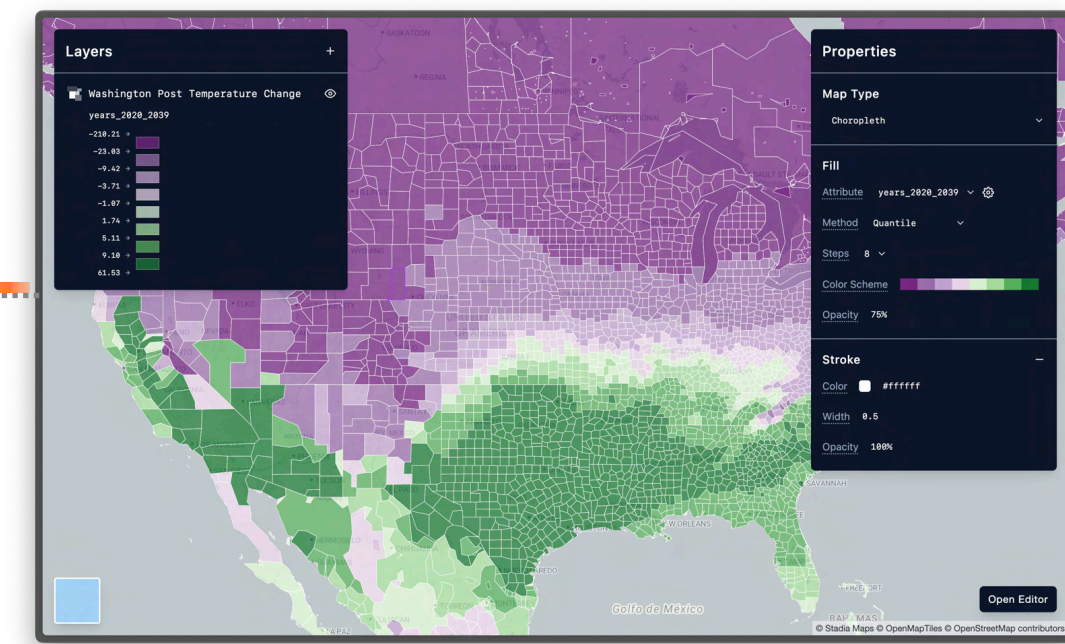
When a user alters a property in the UI, **dispatch an update** to the IR.

Fill Layer



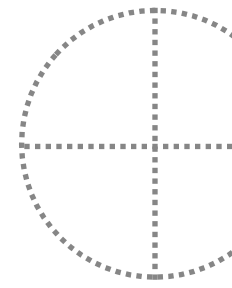
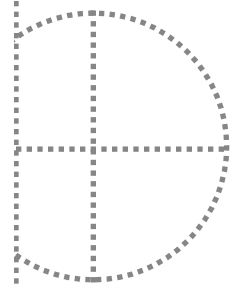
dispatchLayerUpdate

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

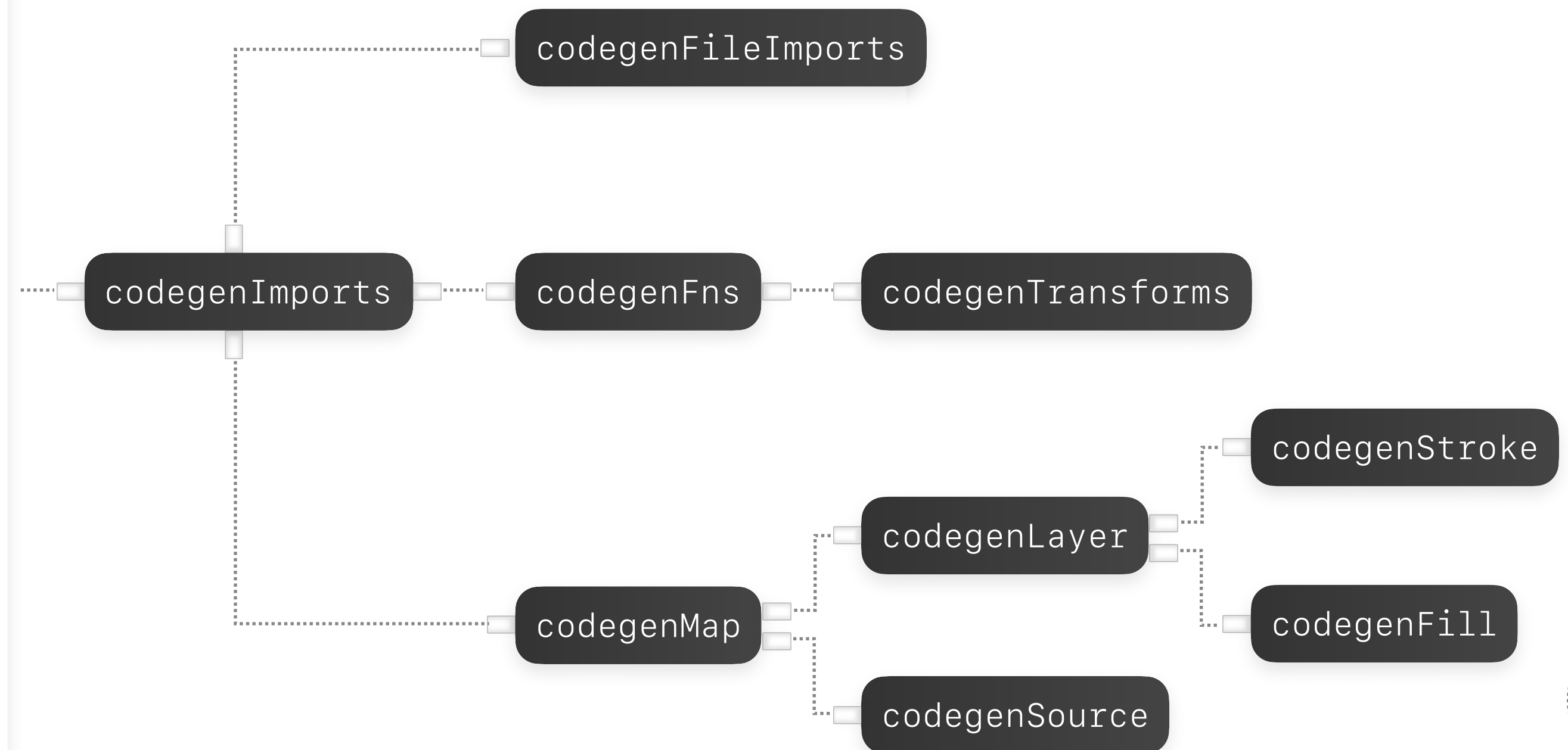
codegenMap

```
`const map = new mapboxgl.Map({
  container: "map",
  style: "${ir.basemap.url}",
  center: [${ir.center[0]}, ${ir.center[1]}],
  zoom: ${map.getZoom()}
});

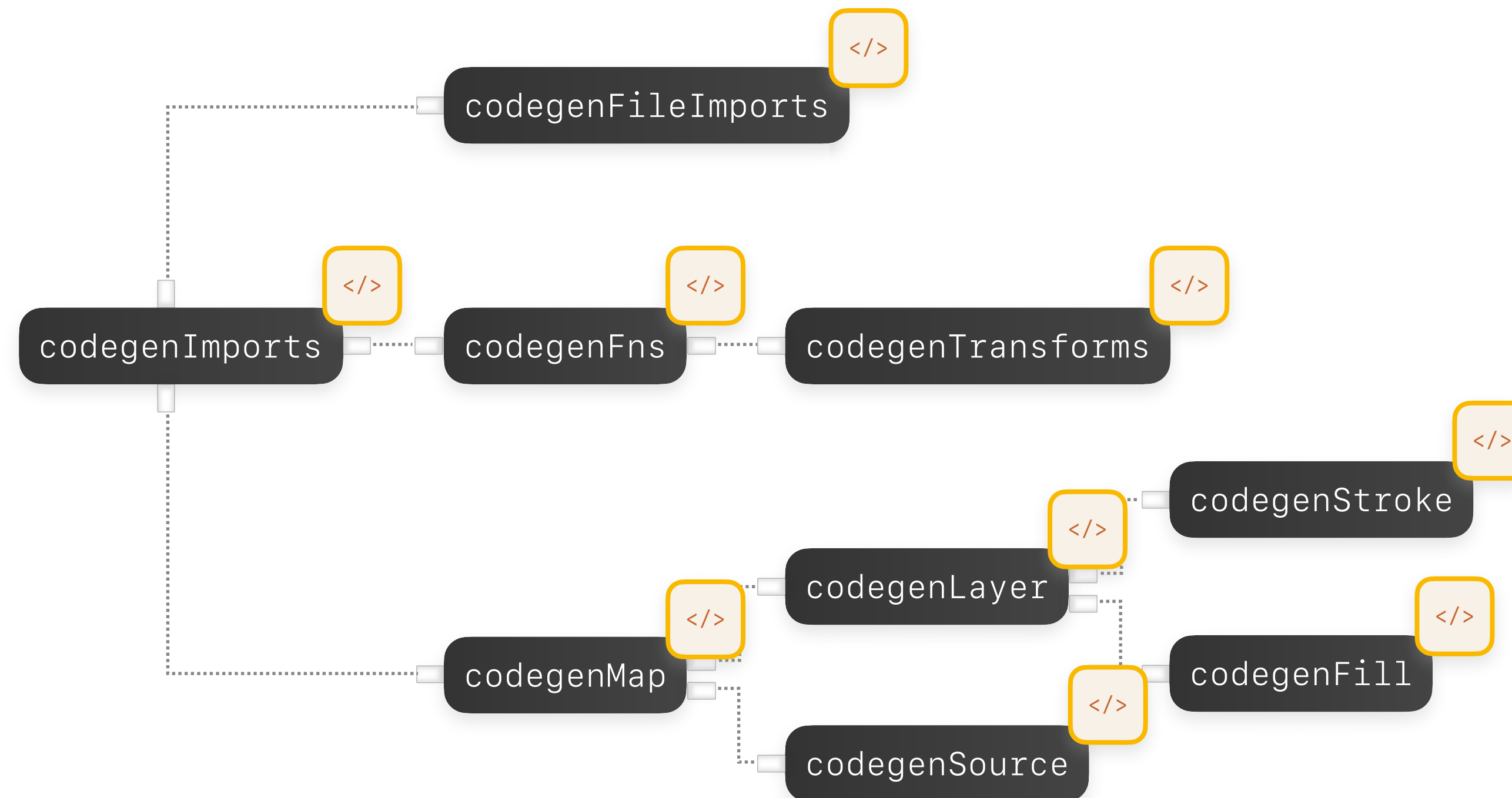
map.on('load', ${isLoadSync ? 'async ' : ''}() => {
  ${Object.values(ir.layers).reduce((p, layer) => {
    return p.concat("\n\n" + codegenSource(layer, uploadTable));
  }, "")}
  ...
});
```

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.



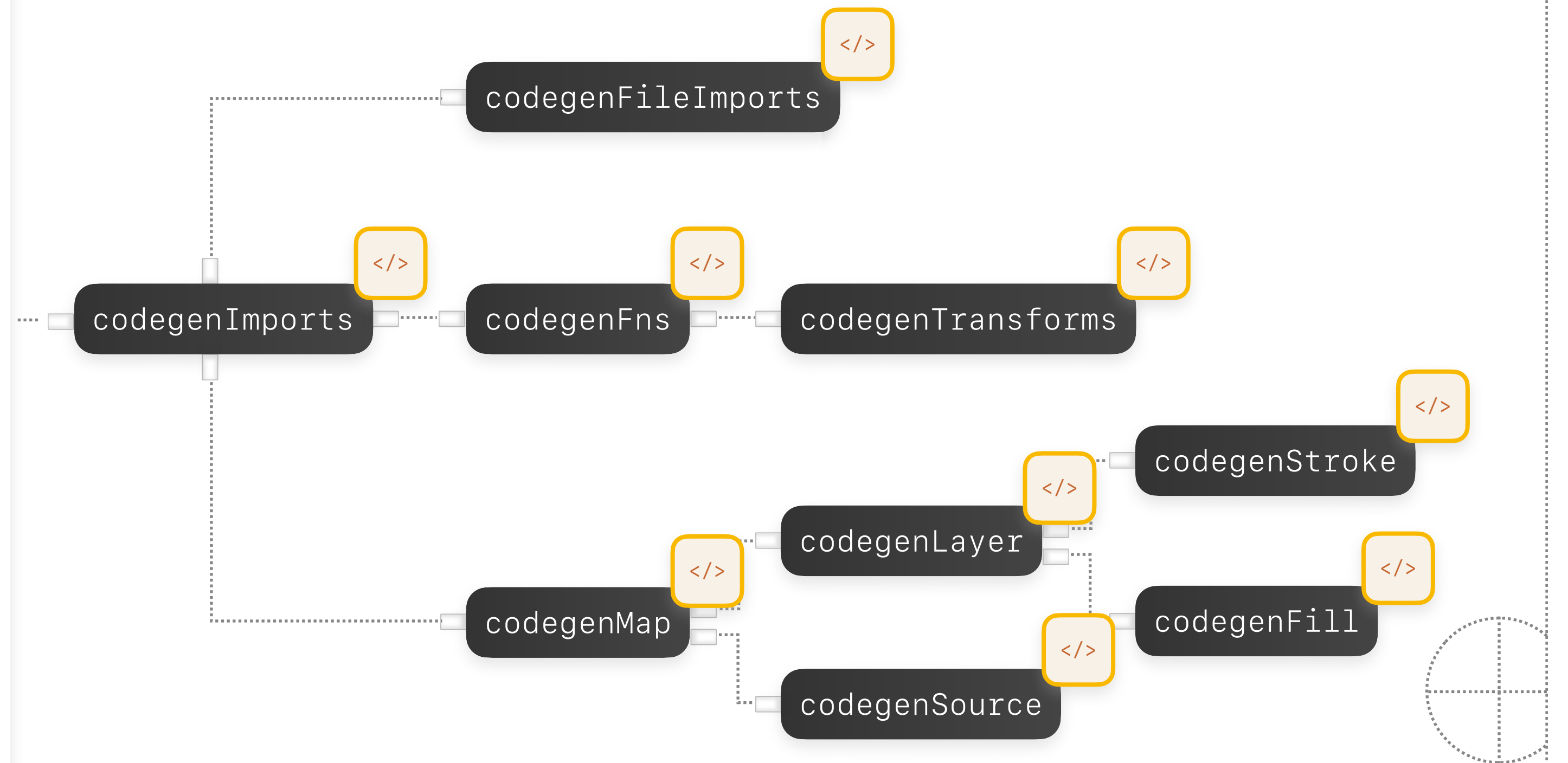
```
import mapboxgl from "mapbox-gl";
import waPoTemperatureRegions from "./wapo-...";

mapboxgl.accessToken = "pk.eyJ6f6fh2gsd6g289d...";

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"],
        ...
      ]
    }
  });
});
```



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



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

mapboxgl.accessToken = "pk.eyJ66fh2gsd6g289d...";

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"],
        ...
      ]
    }
  });
});
```

cartokit IR



MapLibre



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

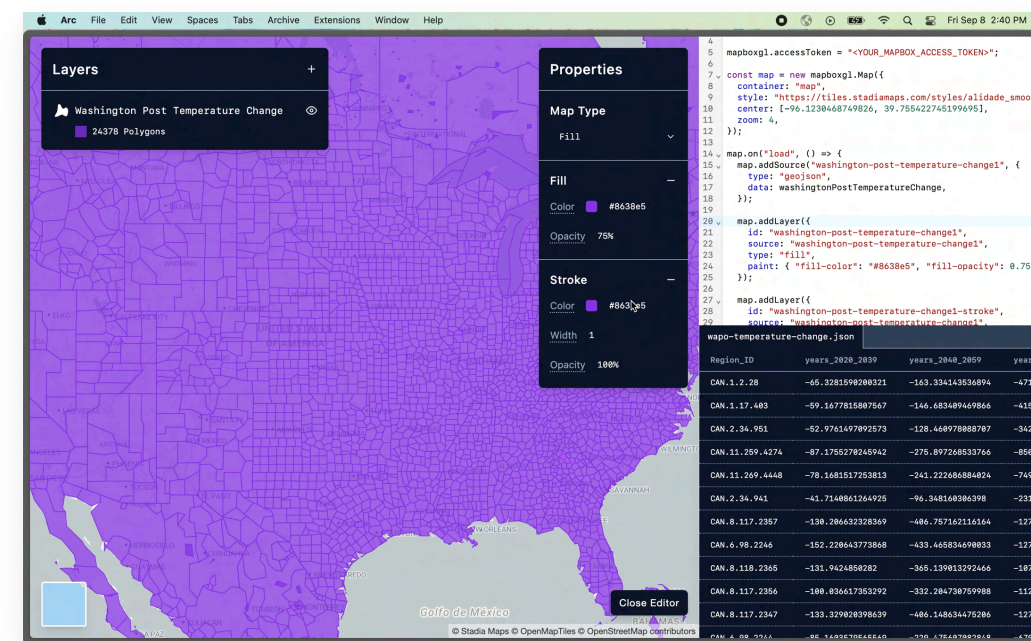
const map = L.map("map")
  .setView([-106.1086, 37.7531], 4);

L.tileLayer(
  "https://tiles.stadiamaps.com/tiles/alida...",
  {
    maxZoom: 20,
    attribution:
      '&copy; <a href="https://stadiamaps..."',
  }
).addTo(map);

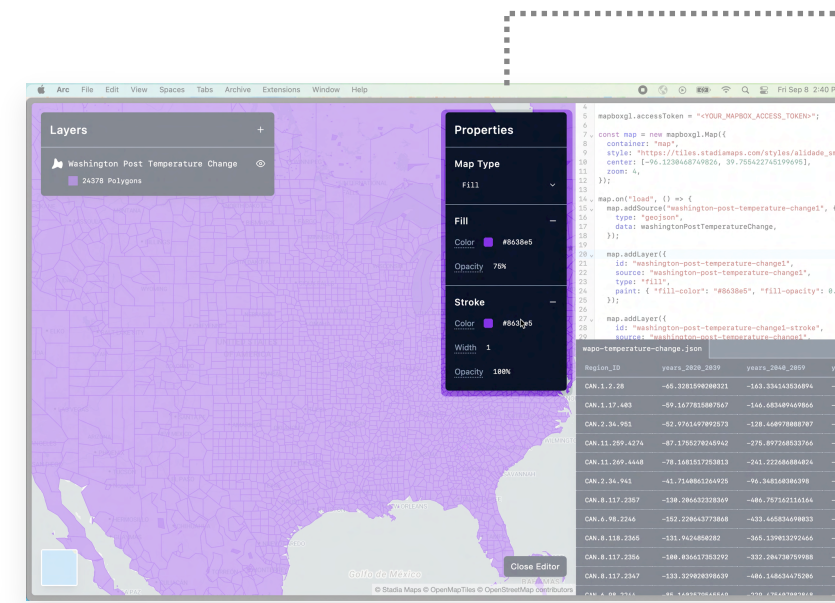
L.geoJSON(waPoTemperatureRegions, {
  style: (feature) => {
    const attr =
      feature.properties["years_2020_2039"];

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

Direct Manipulation



UI controls → Intermediate Representation

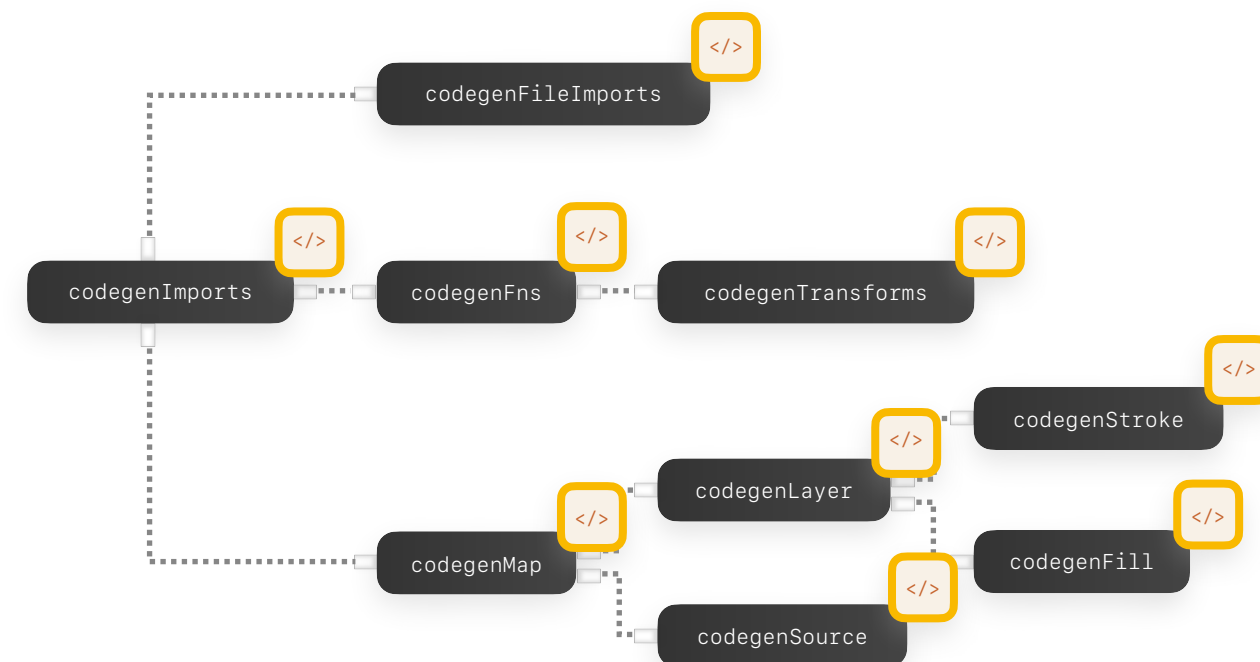


```
const ir = {
  center: [-106.1086, 37.7531],
  zoom: 4,
  basemap: {
    url: "https://tiles.stadiamaps.com/styl.",
    provider: "Stadia Maps",
  },
  layers: [
    {
      type: "Fill",
      style: {
        fill: {
          color: "#8638e5",
          opacity: 0.75,
        },
        stroke: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

Update 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: [-219.21, -23.83, ...],
          opacity: 0.75,
        },
        stroke: {
          color: "#ffffff",
          width: 1,
          opacity: 0.5,
        },
      },
    },
  ],
};
```

Code Generation



JavaScript Program

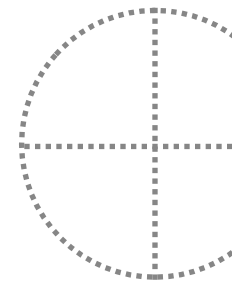
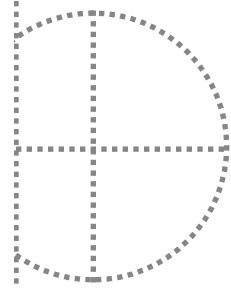
```
import mapboxgl from "mapbox-gl";
import waPoTemperatureRegions from "./wapo-...";

mapboxgl.accessToken = "pk.eyJ6J6fh2gsd6g289d...";

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,
  });

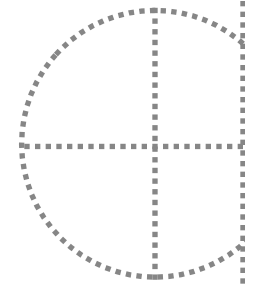
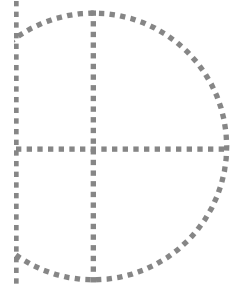
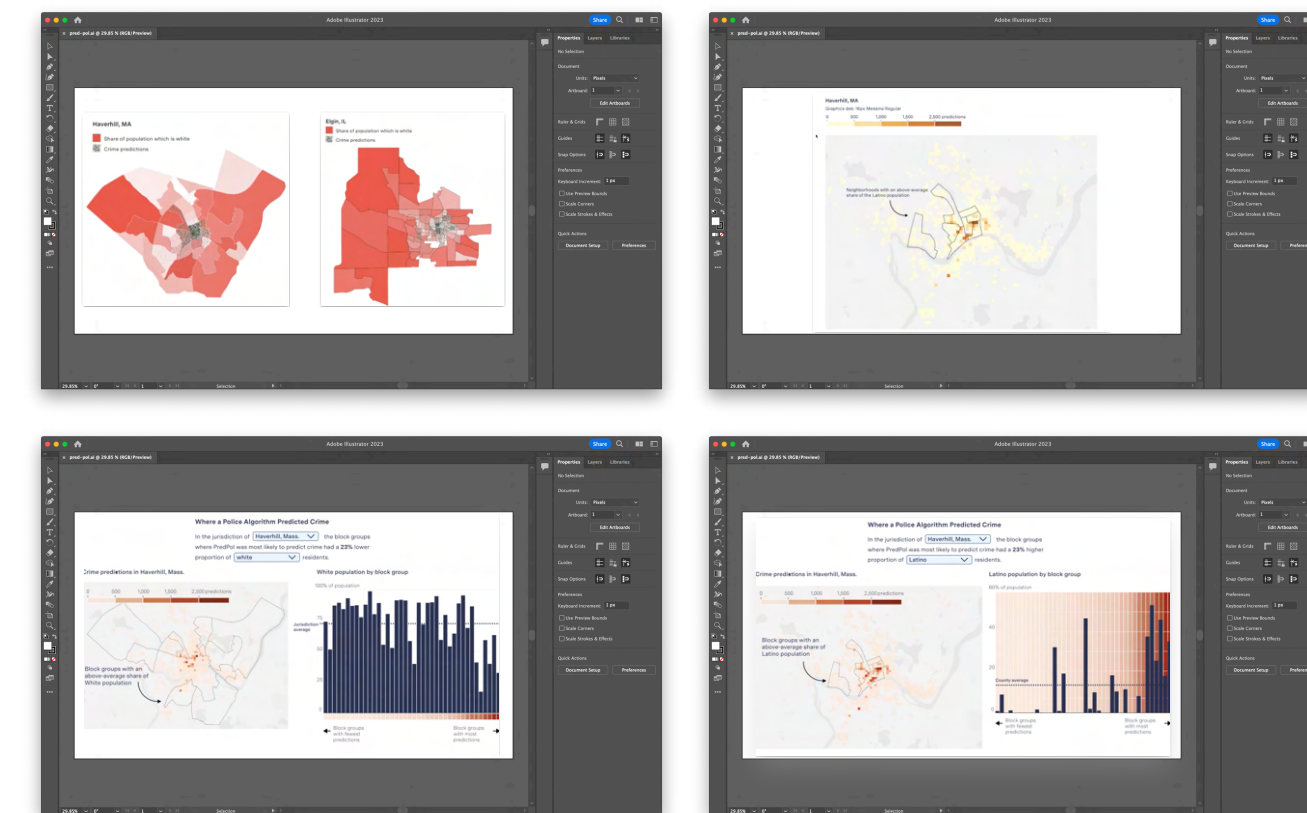
  map.addLayer({
    id: "wapo-temperature-regions",
    source: "wapo-temperature-regions",
    type: "fill",
    paint: {
      "fill-color": [
        "step",
        ["get", "years_2020_2039"],
        ...
      ]
    }
  });
});
```

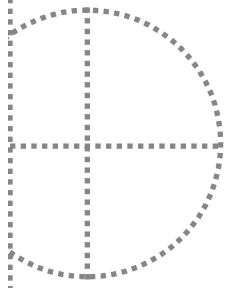


Working Practice

Sketch visualizations in **design software**

Explore a wide design space **without code**





Working Practice

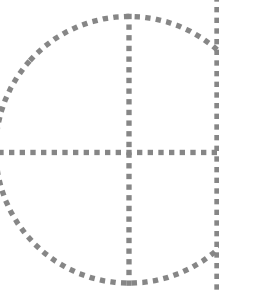
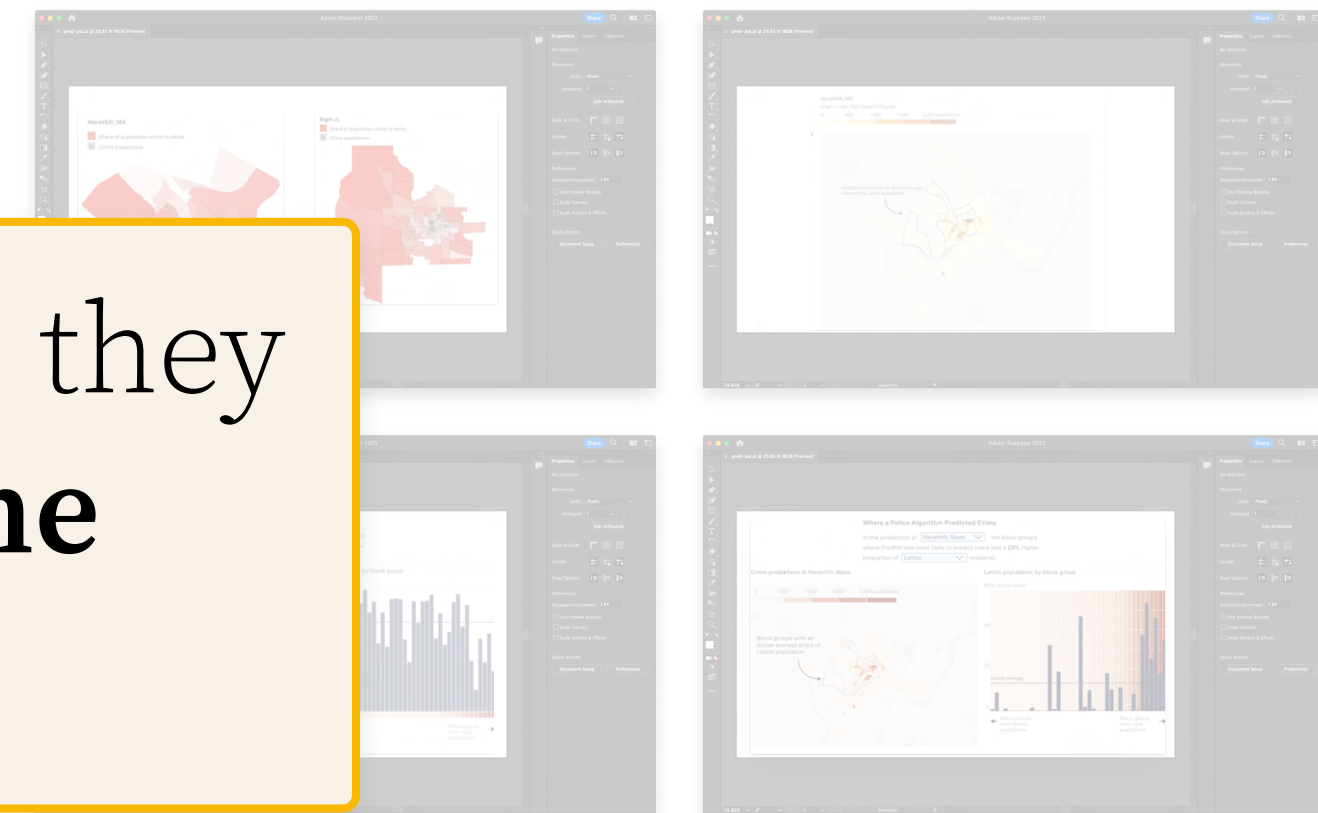
Sketch visualizations in **design software**



Explore a wide design space **without code**



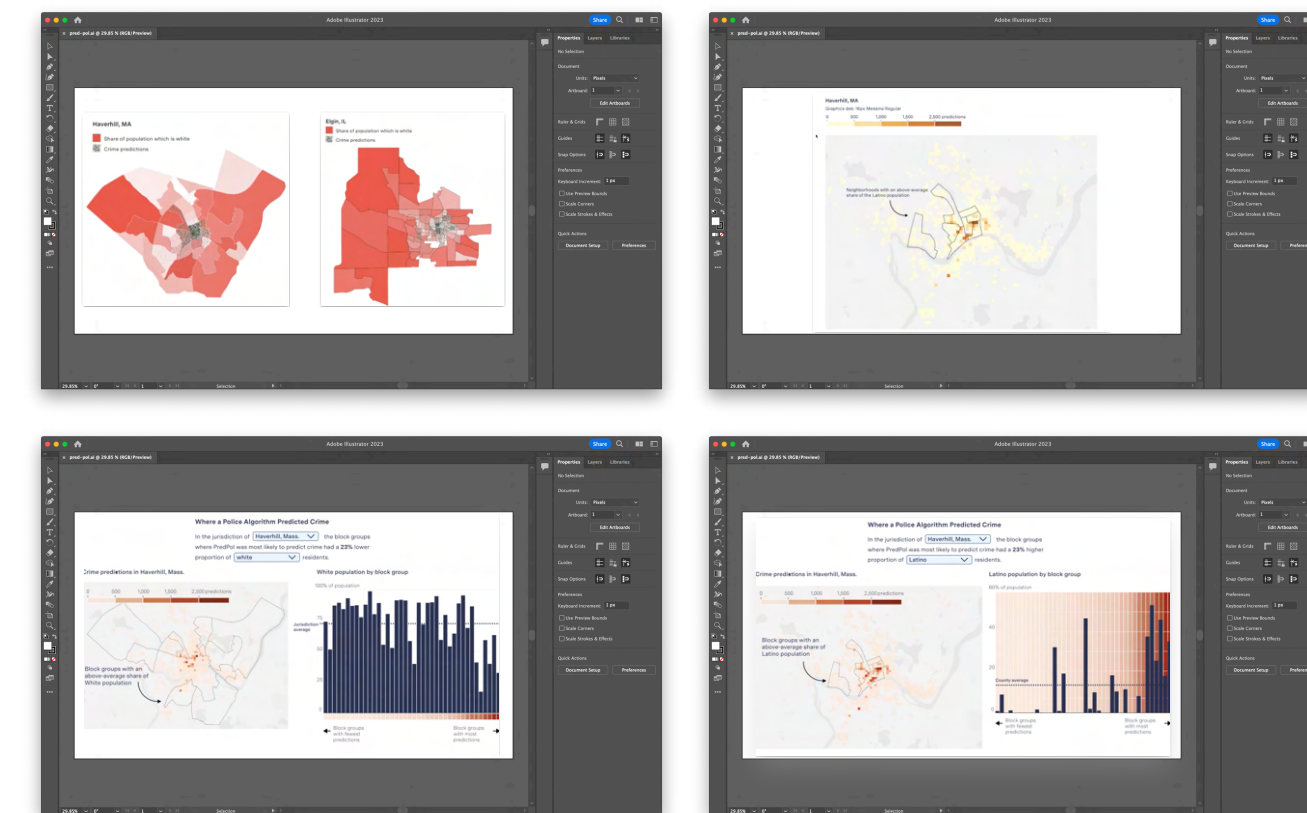
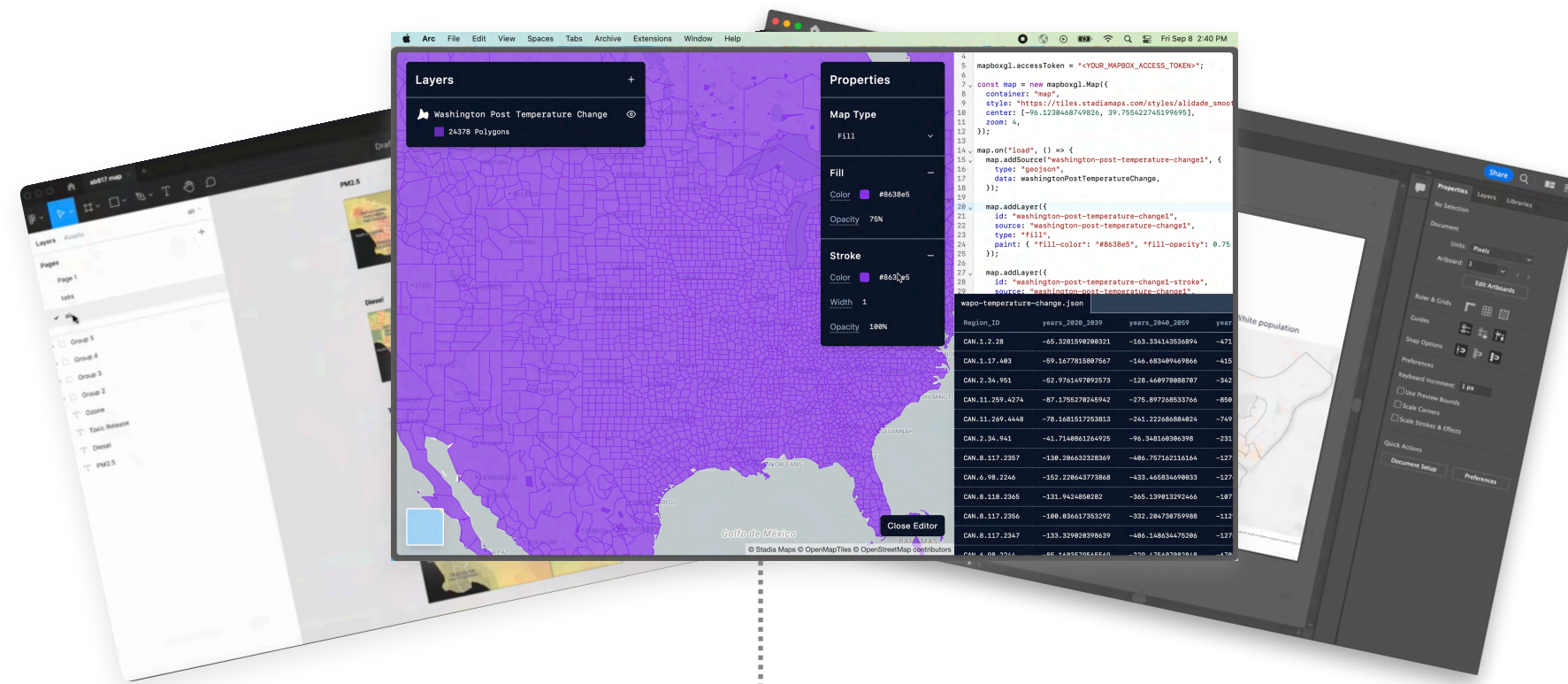
After selecting a design, they still had to **reproduce the visualization in code.**



Working Practice

Sketch visualizations in **design software**

Explore a wide design space **without code**



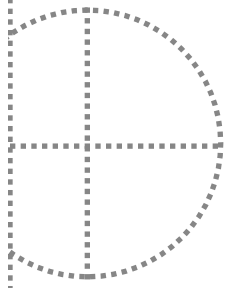
Milliseconds

```
import mapboxgl from "mapbox-gl";
import cancerRegions from "cancer-regions";

mapboxgl.accessToken = "pk.eyJ66fh2...";

const map = new Map({
  container: "map",
  style: "https://tiles.stadiamaps.com/tiles/hybrid/1/v1/styles/v1.json",
  center: [-106.1086, 37.7531],
  zoom: 4,
});

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



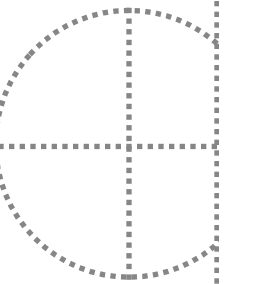
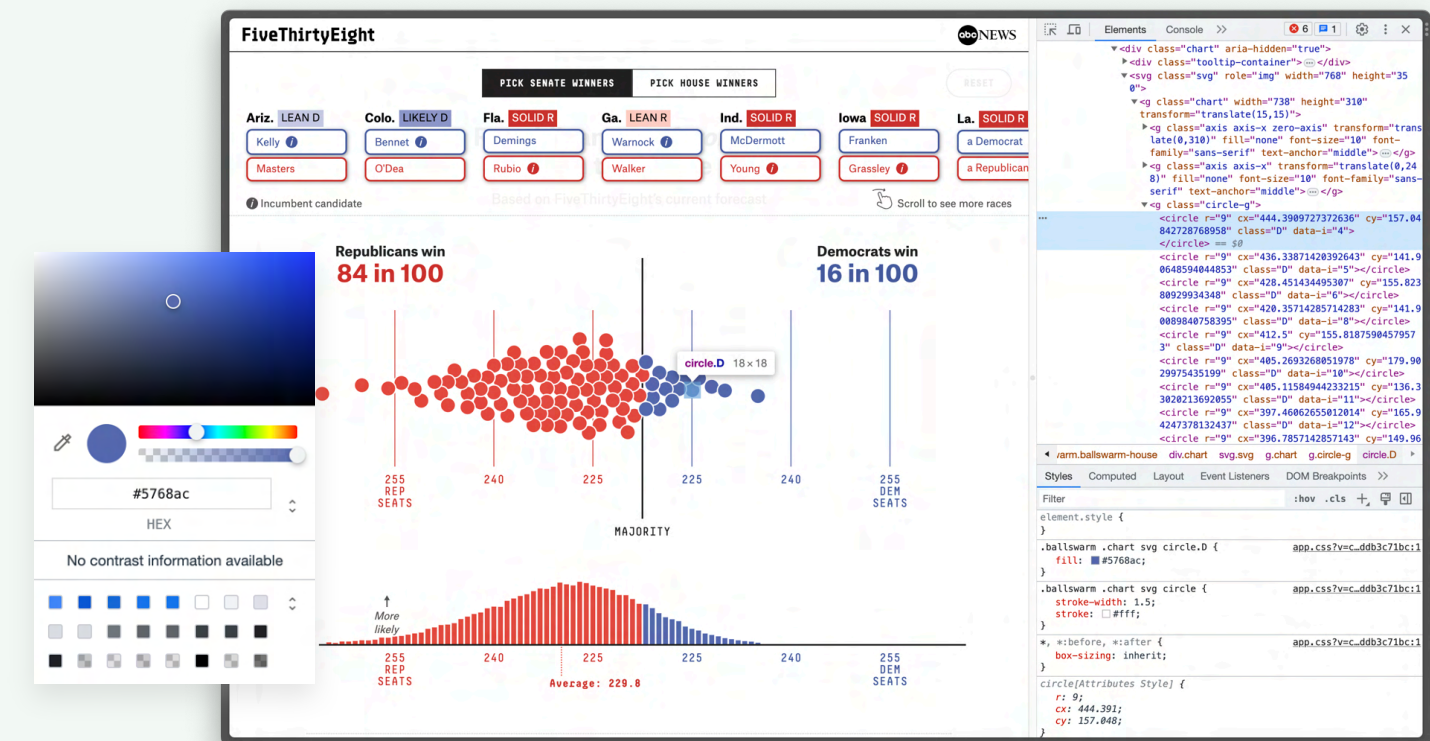
Lexer

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

Parser



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



Compilers for Visual Inputs

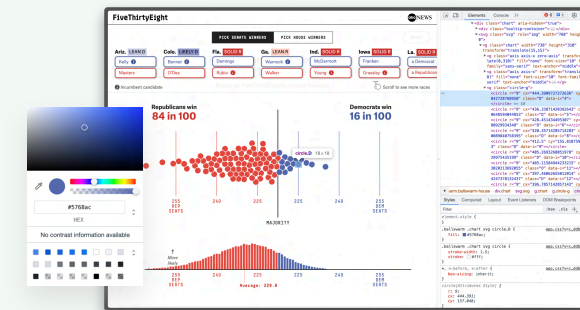
Lexer

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

Parser



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



cartokit

Cartography Studio



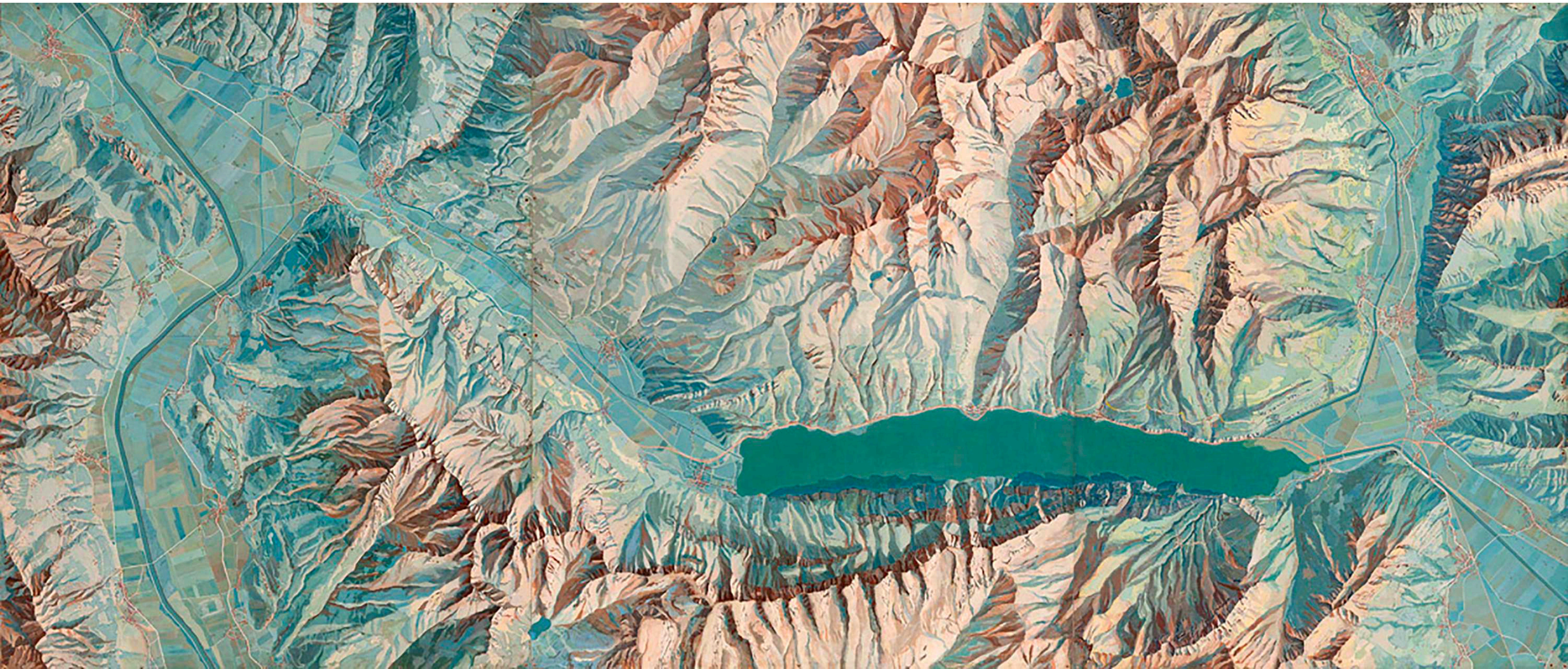
Linework

“London Underground Tube Map” • *Harry Beck*



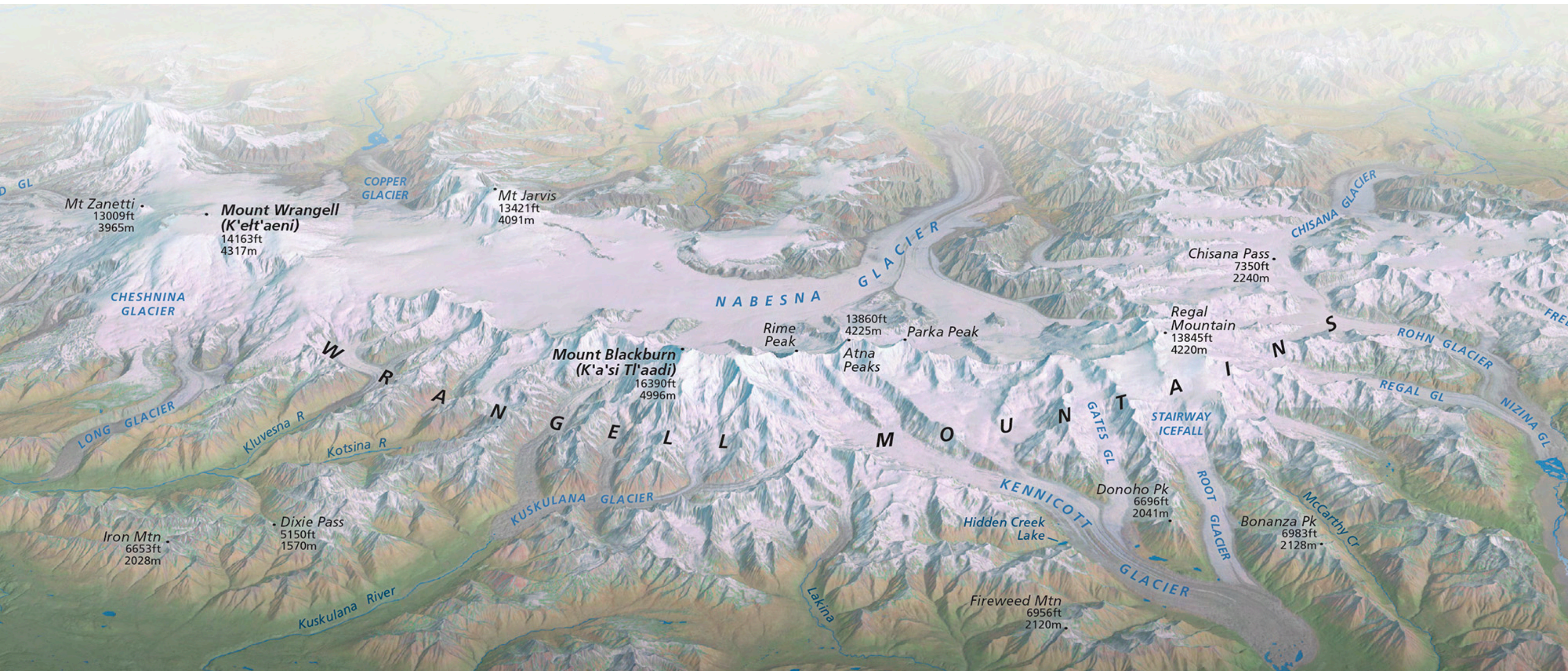
Shaded Relief

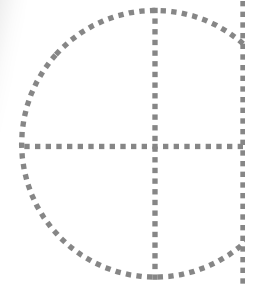
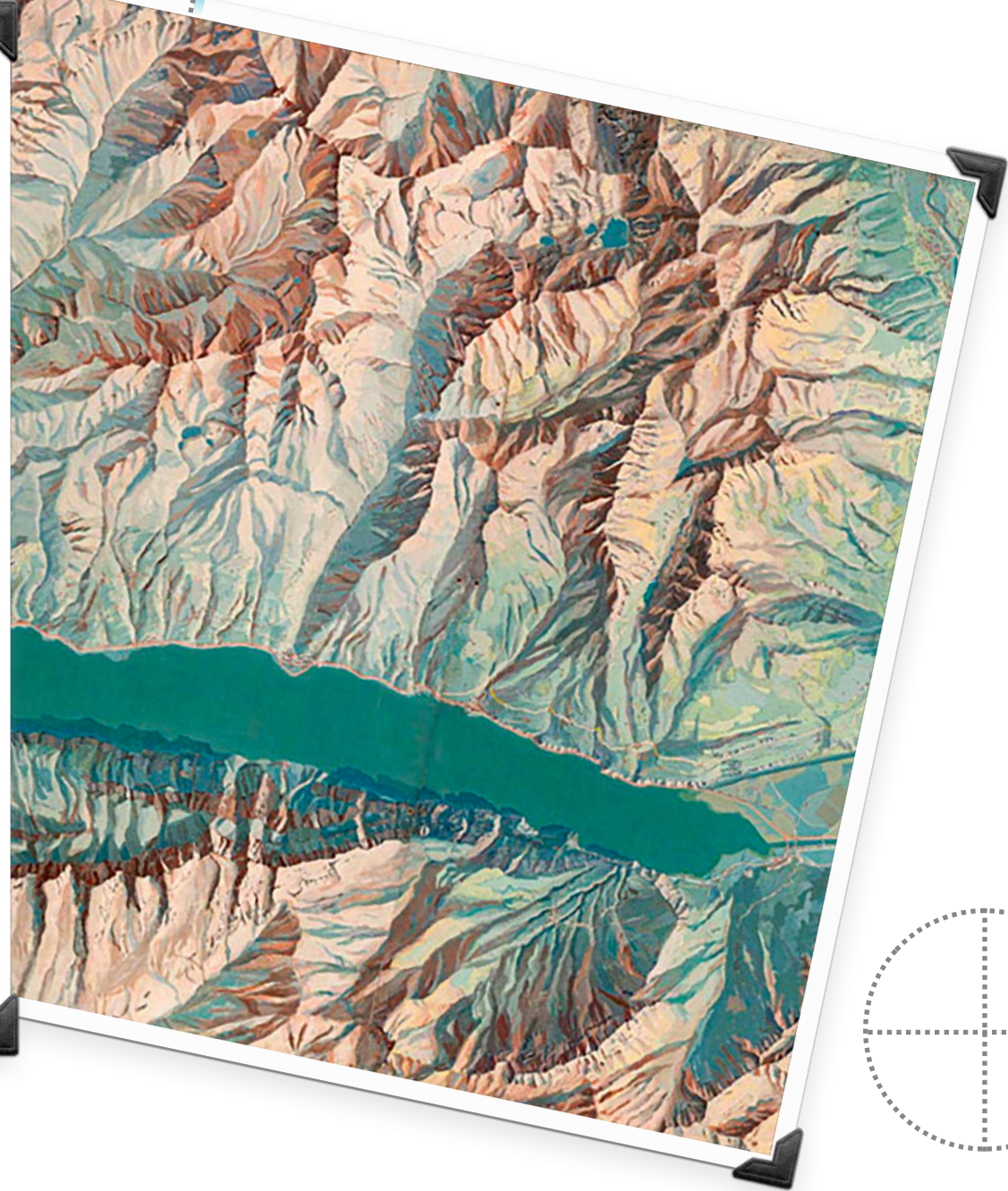
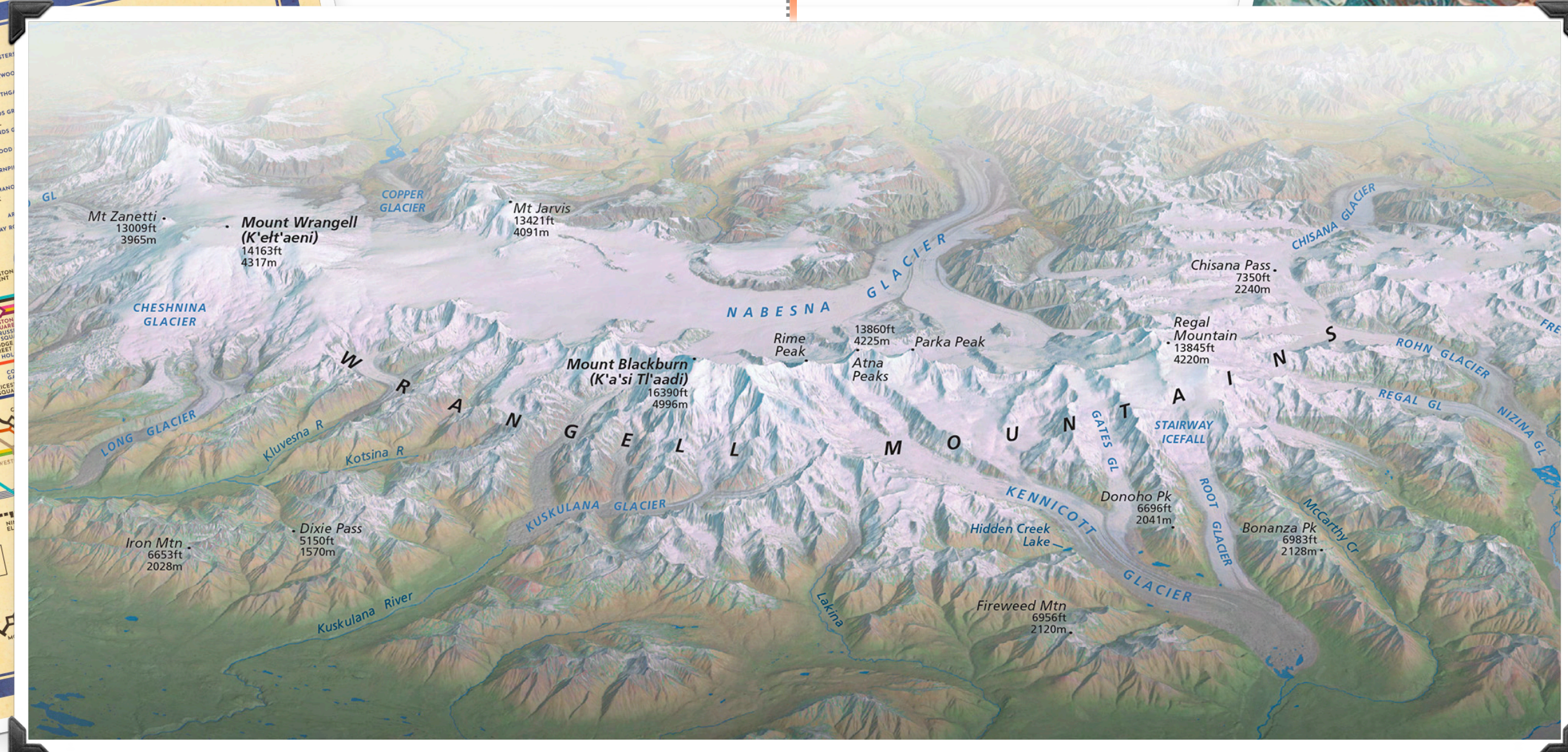
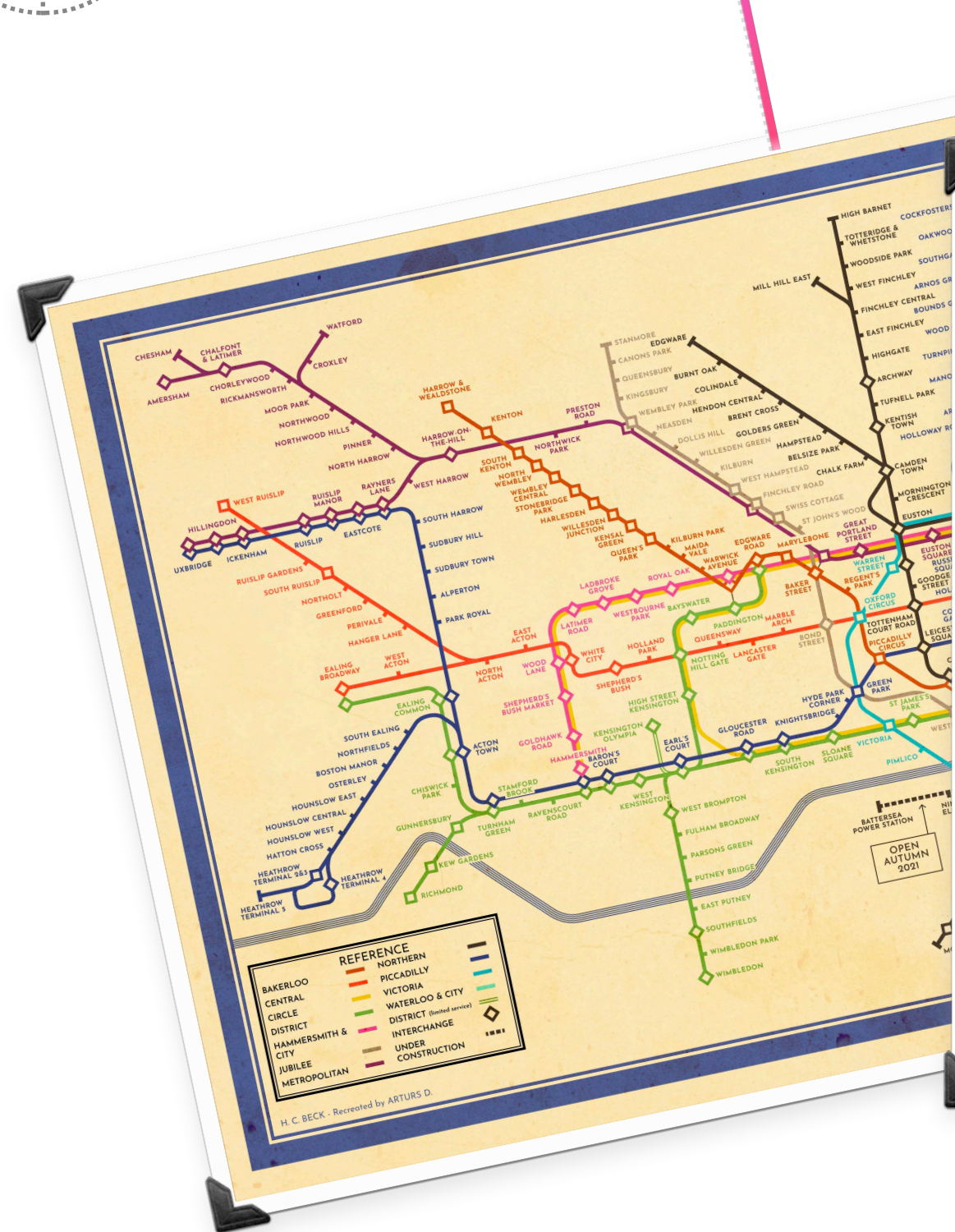
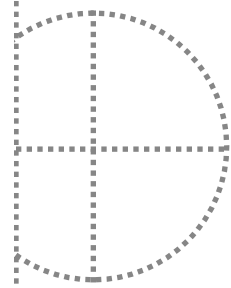
“Walensee und Seeztal” • *Eduard Imhof*



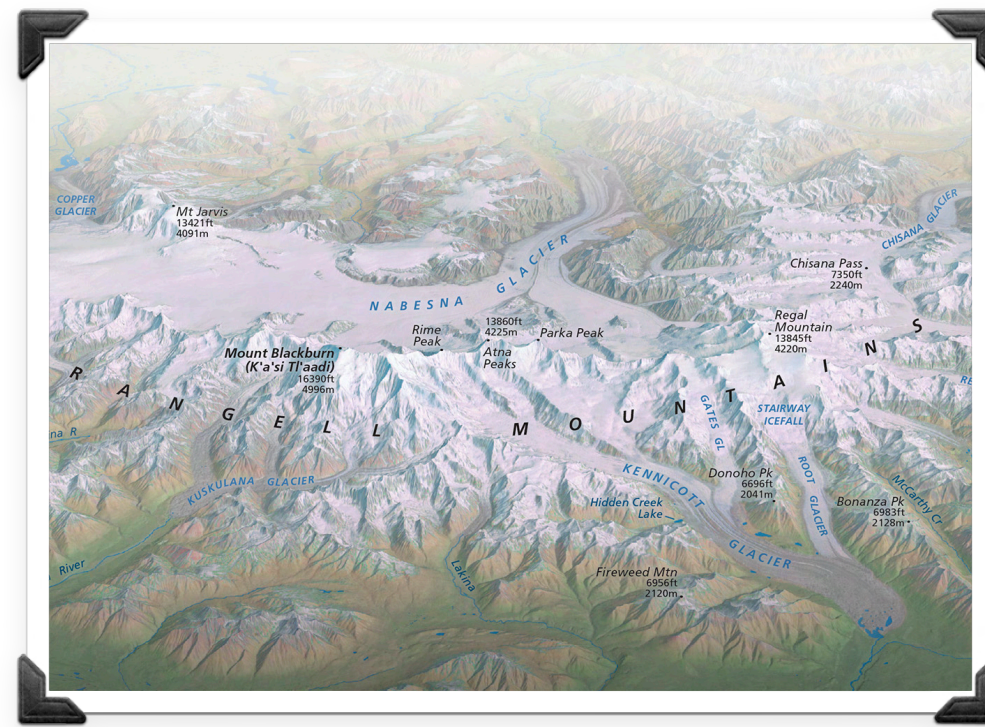
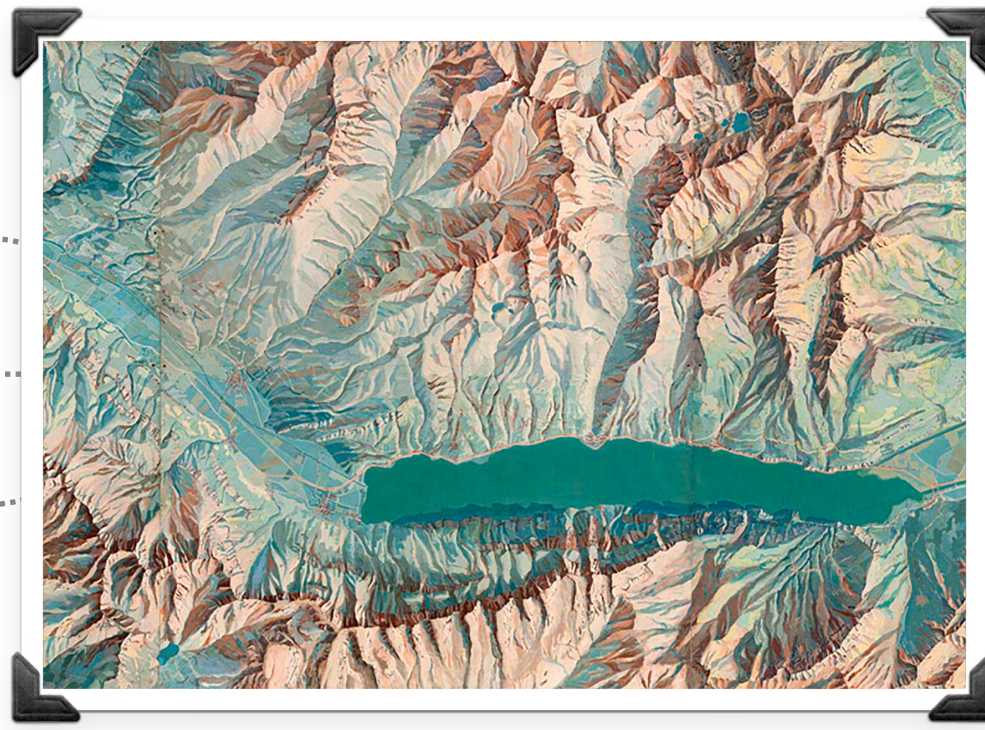
Perspective

“Wrangell Mountains Shaded Relief Map” • *Tom Patterson*

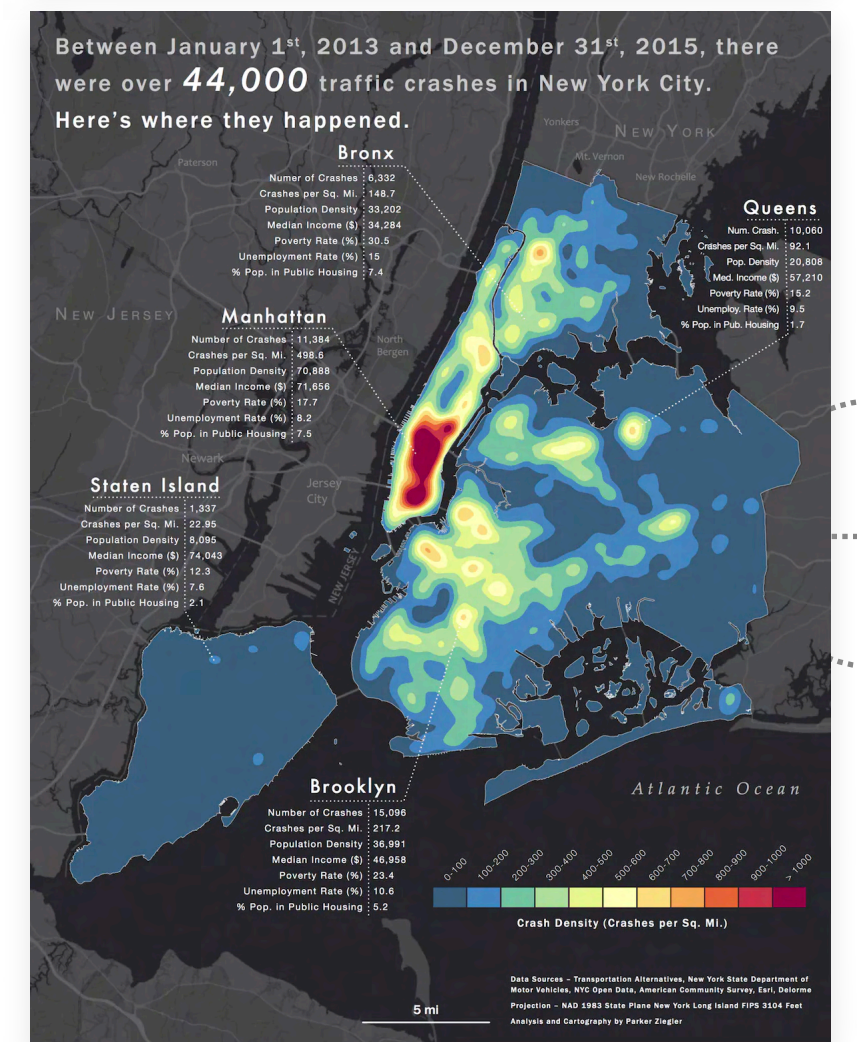
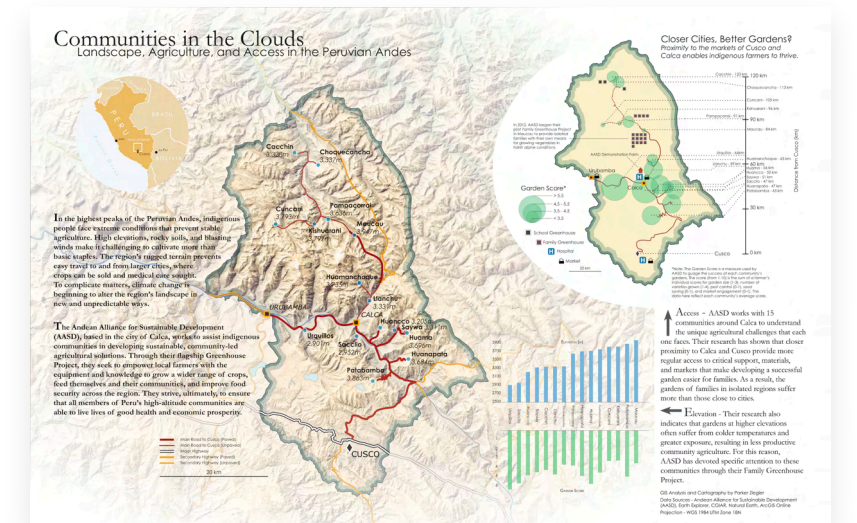




Examples

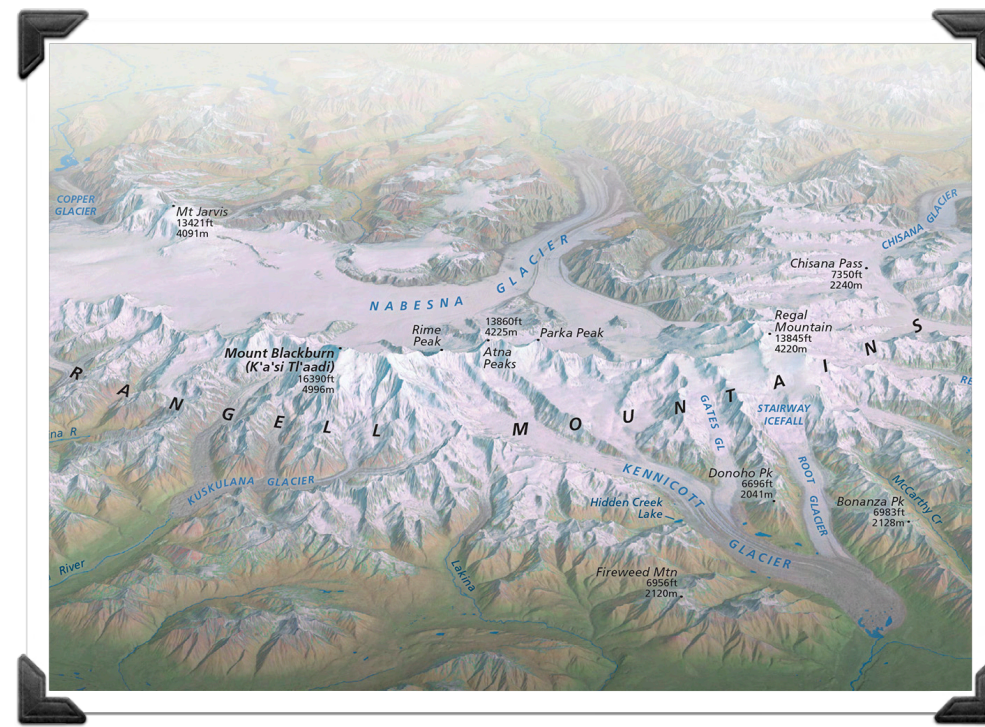
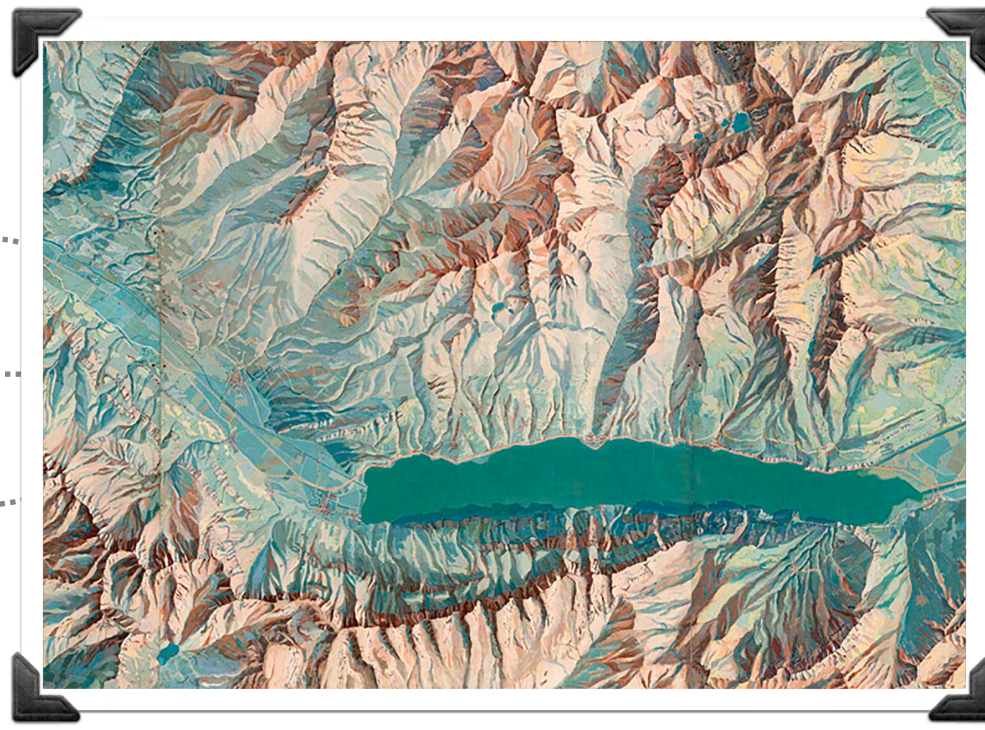


Outputs

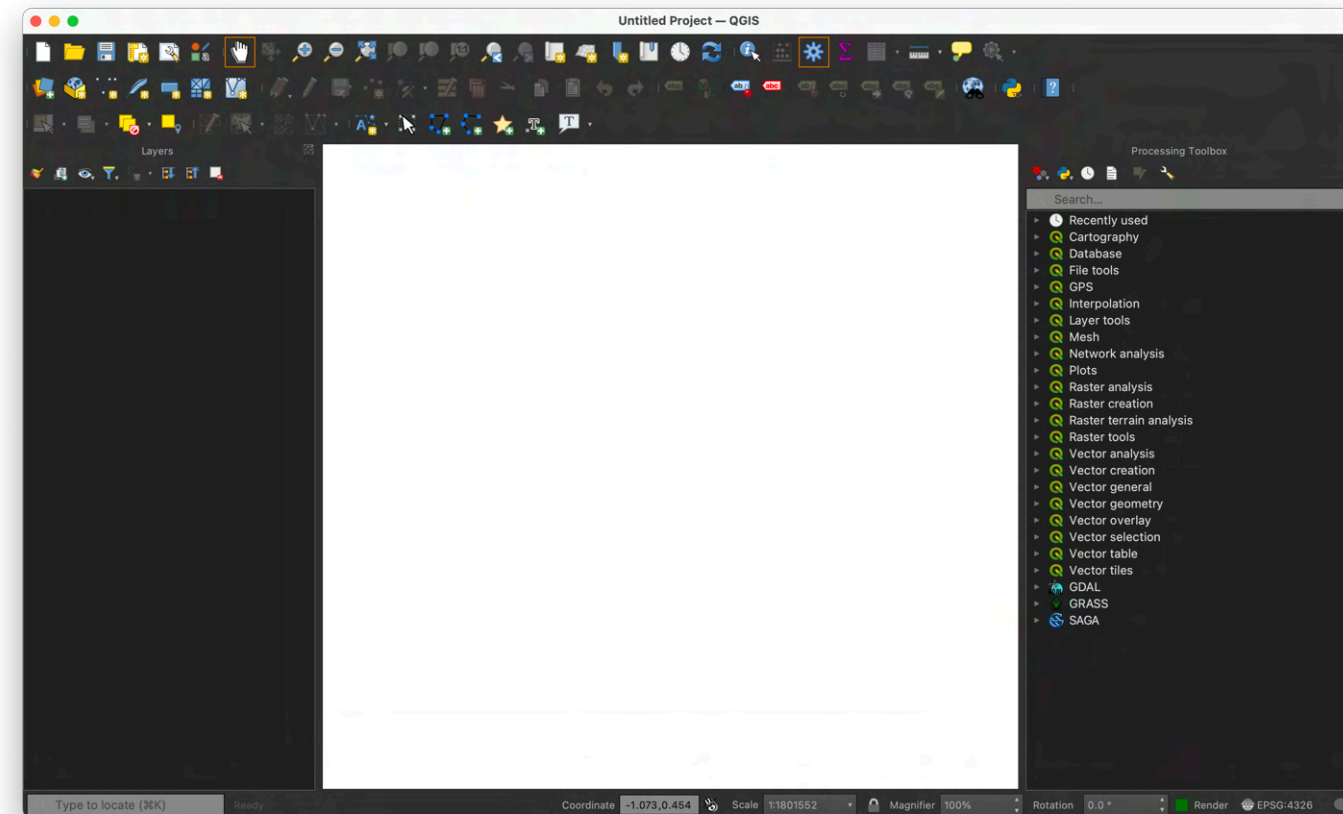
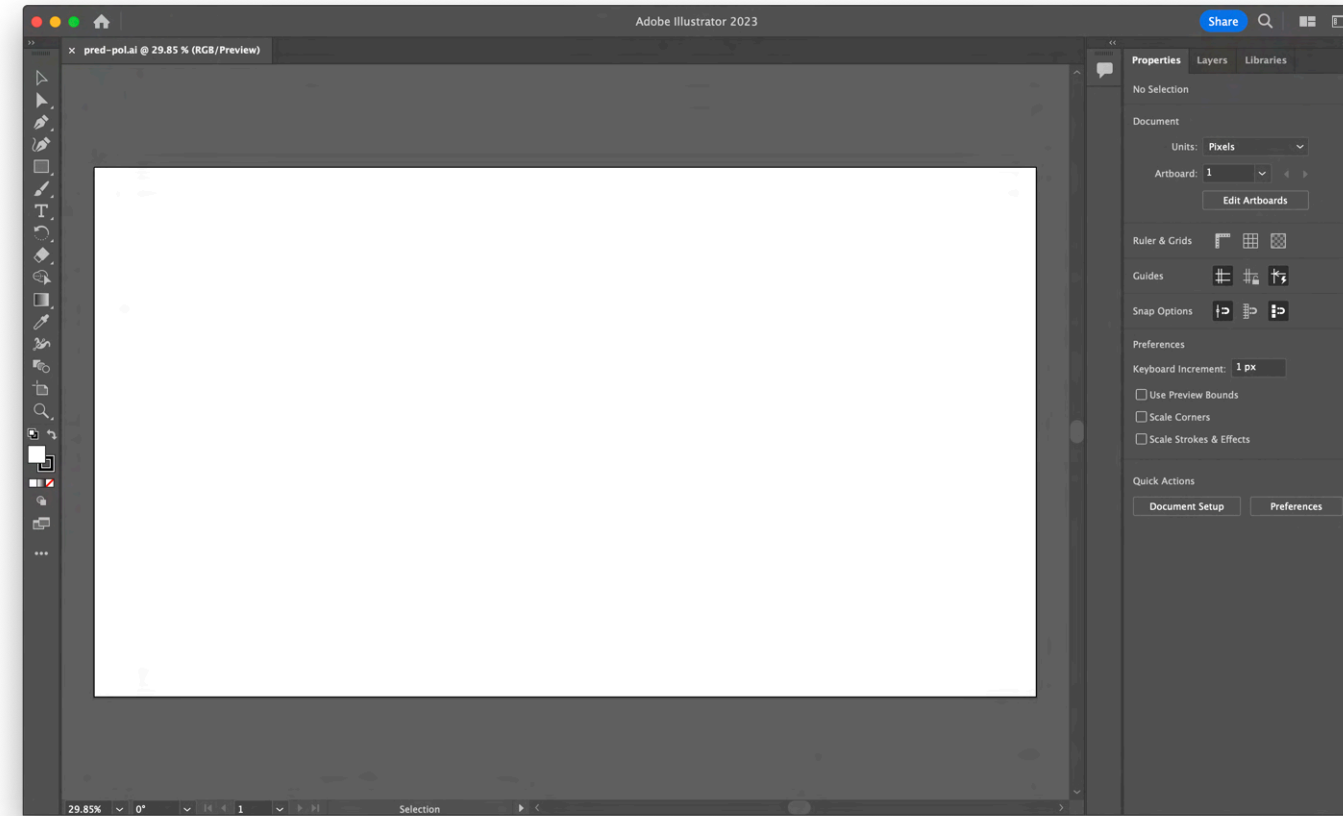


Many, Many (Many) Hours of Painstaking Work

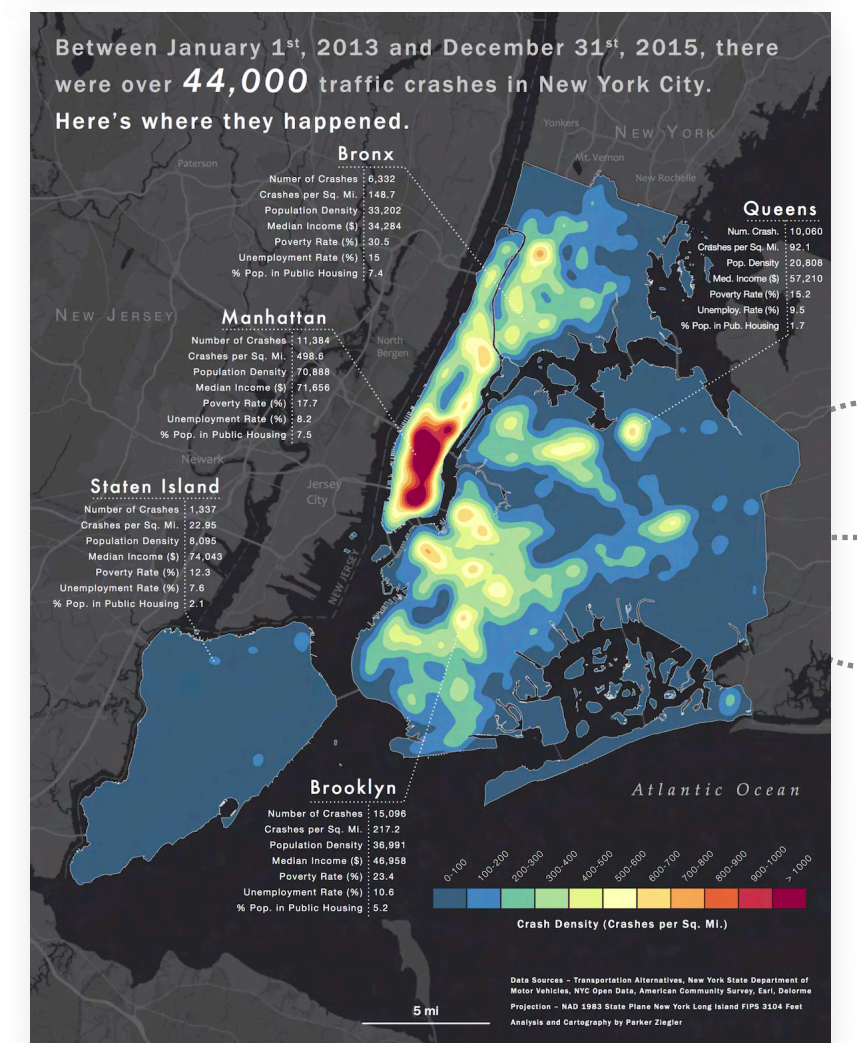
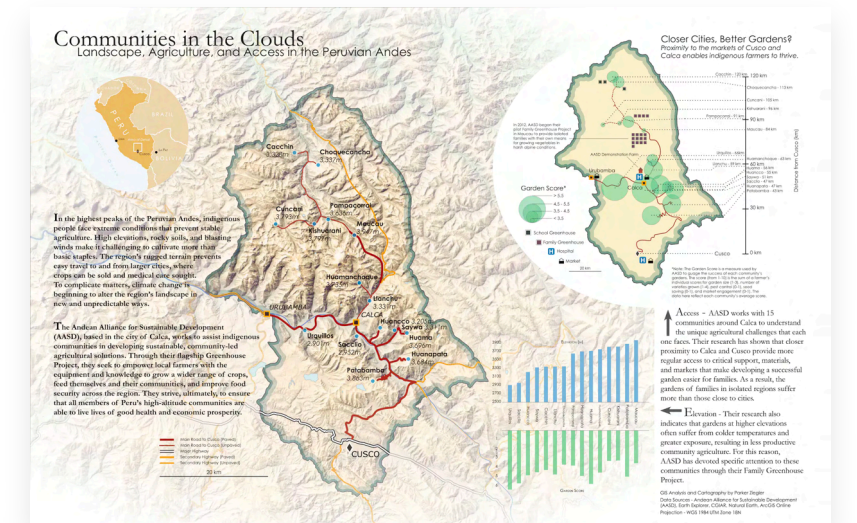
Examples



Many, Many (Many) Hours of Painstaking Work



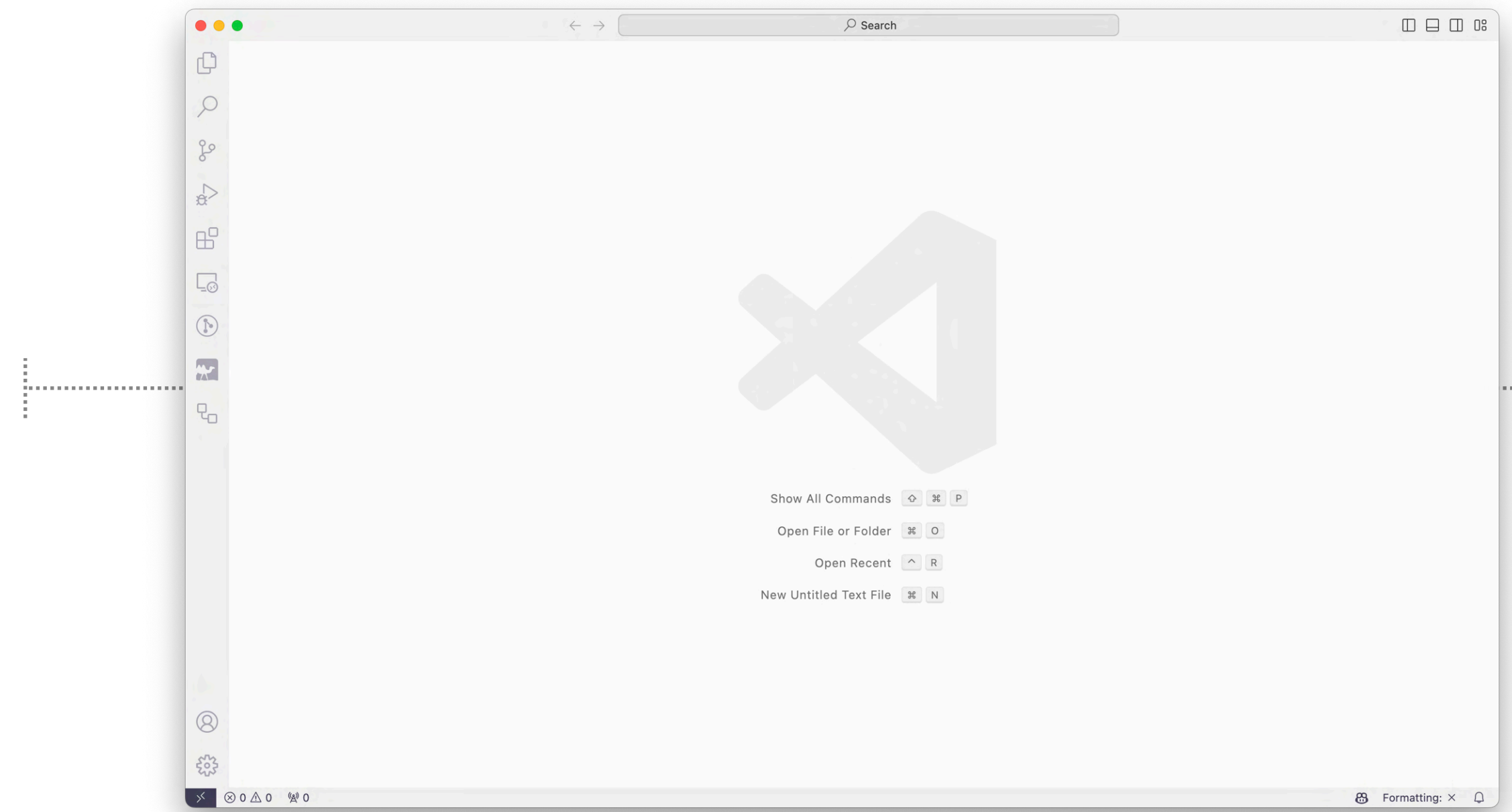
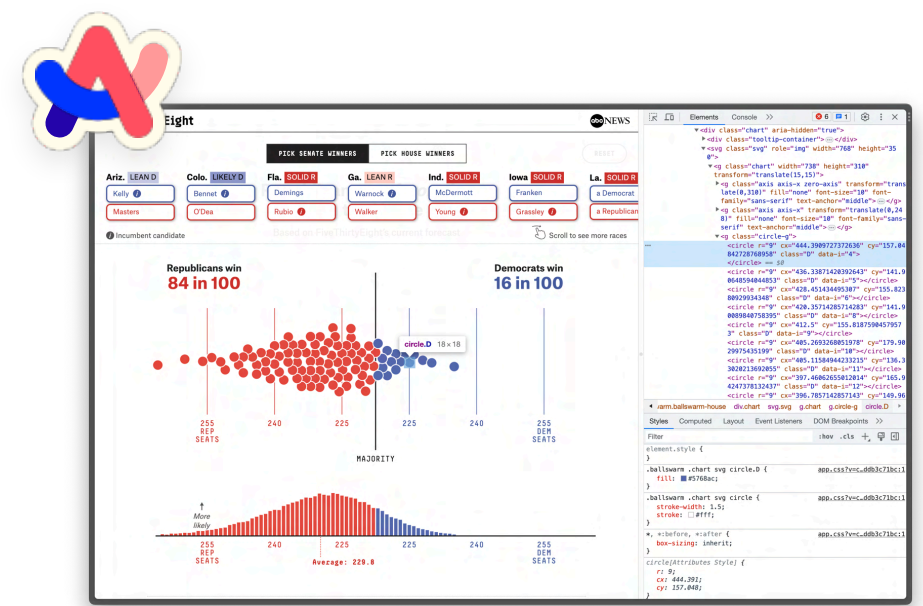
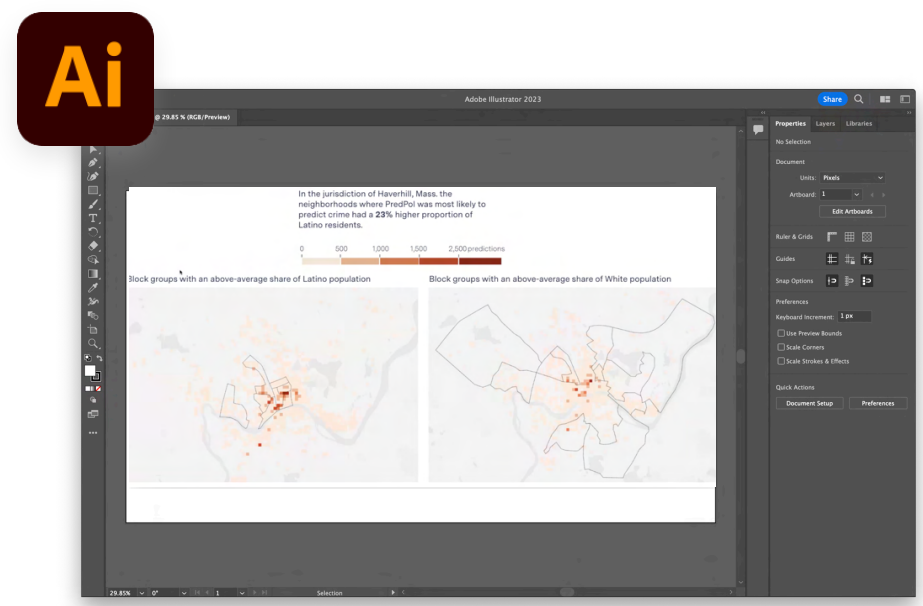
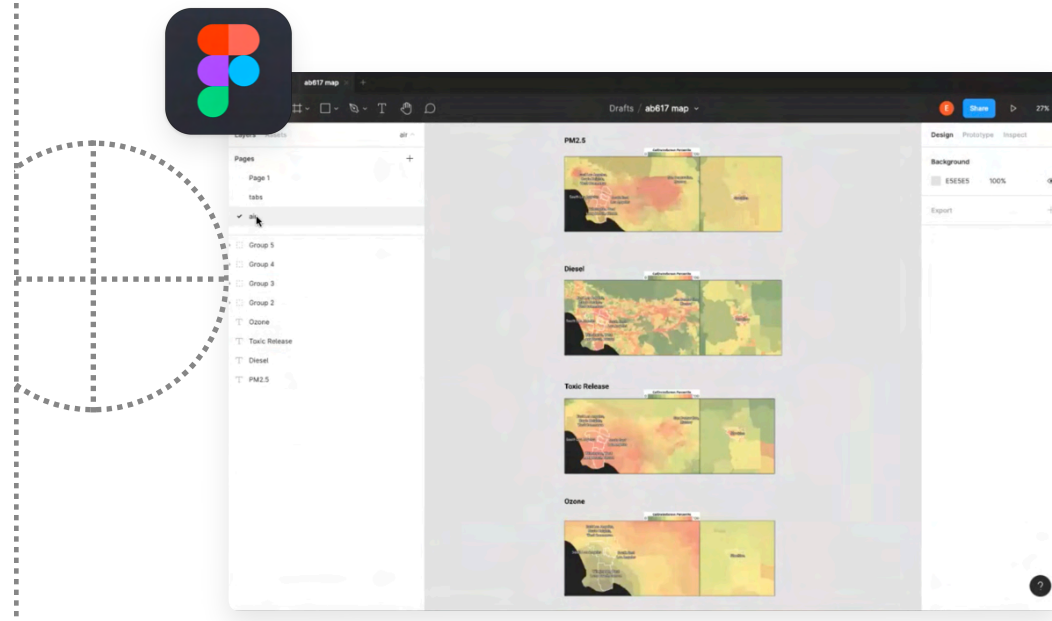
Outputs



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-...";

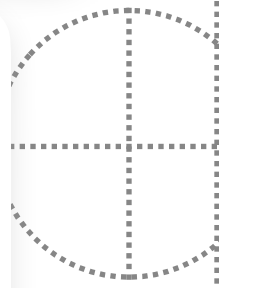
mapboxgl.accessToken = "pk.eyJ6J6fh2...";

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

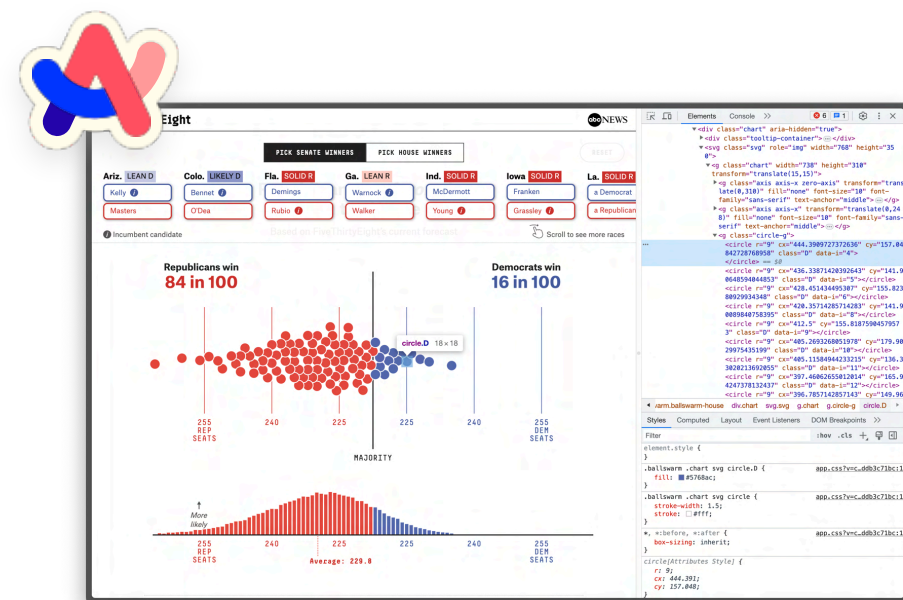
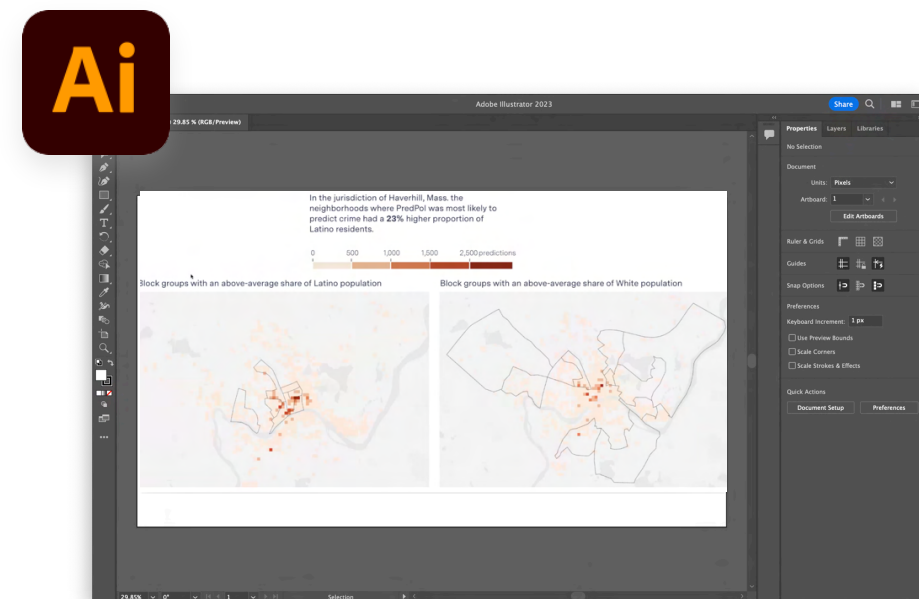
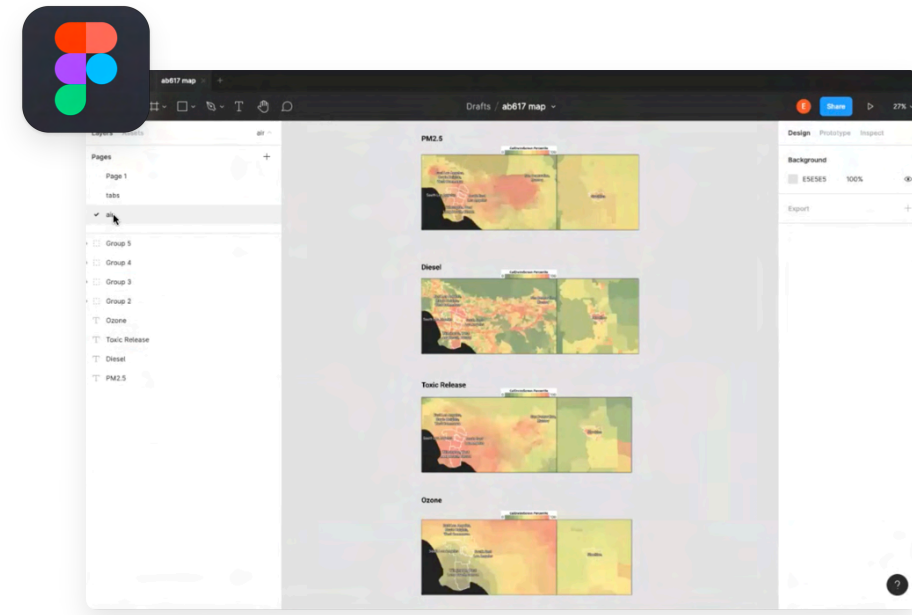
map.on("load", () => {
  map.addSource("cancer-regions", {
    type: "geojson",
    data: cancerRegions,
  });
});
```

```
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 "./canc-...";

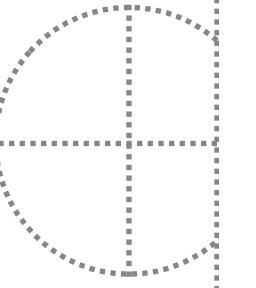
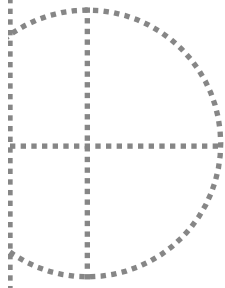
mapboxgl.accessToken = "pk.eyJ6f6fh2...";

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

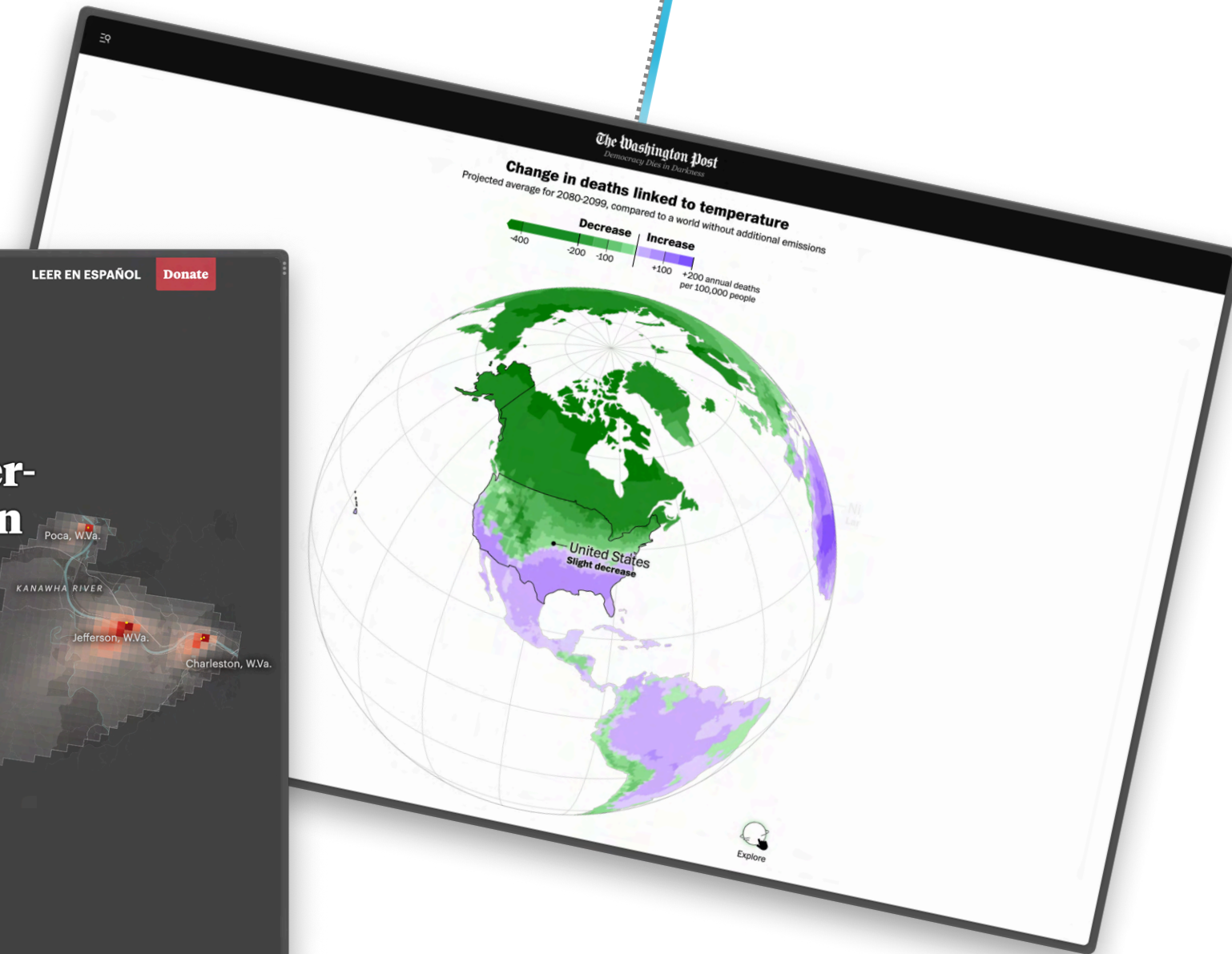
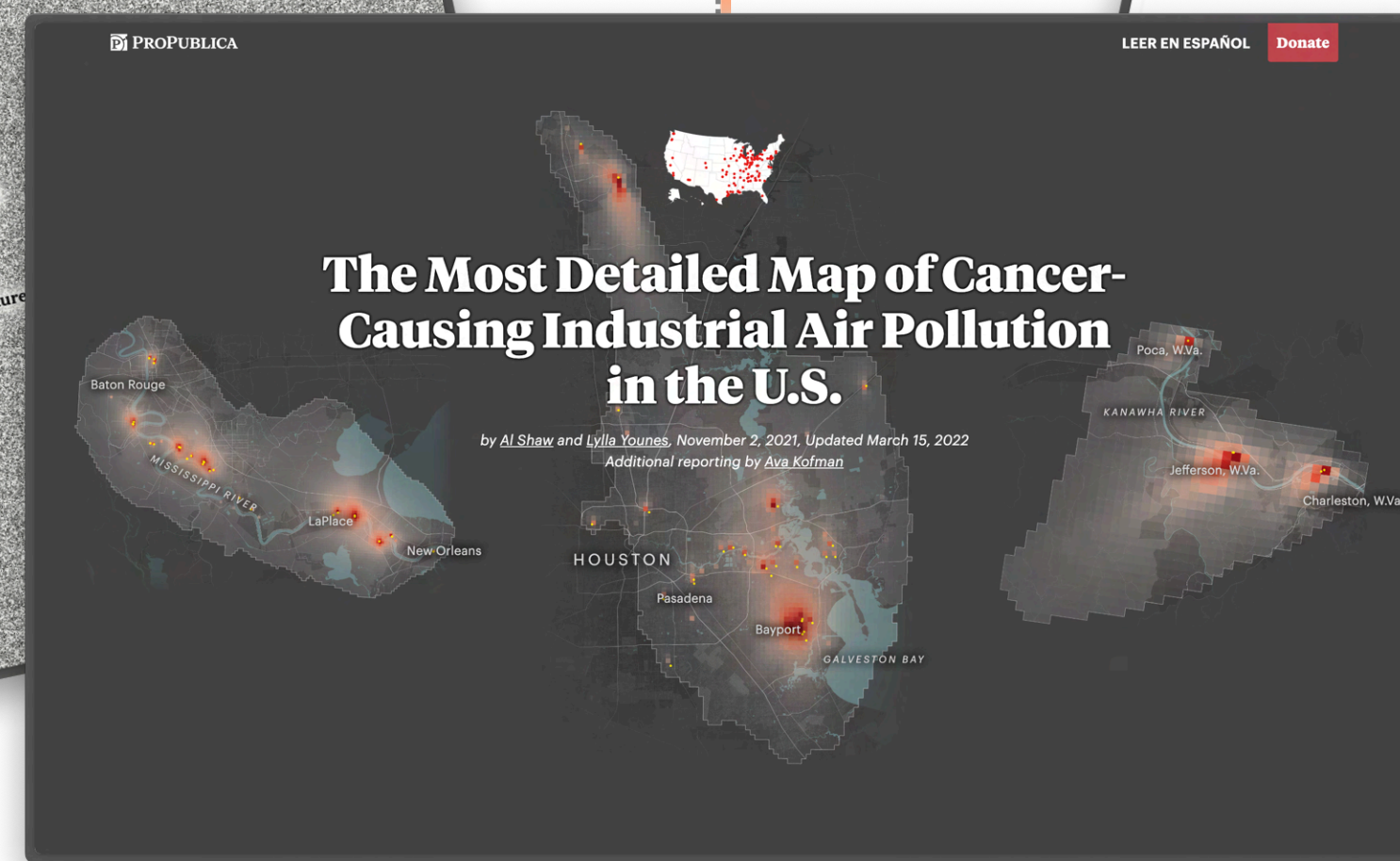
map.on("load", () => {
  map.addSource("cancer-regions", {
    type: "geojson",
    data: cancerRegions,
  });
});
```

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

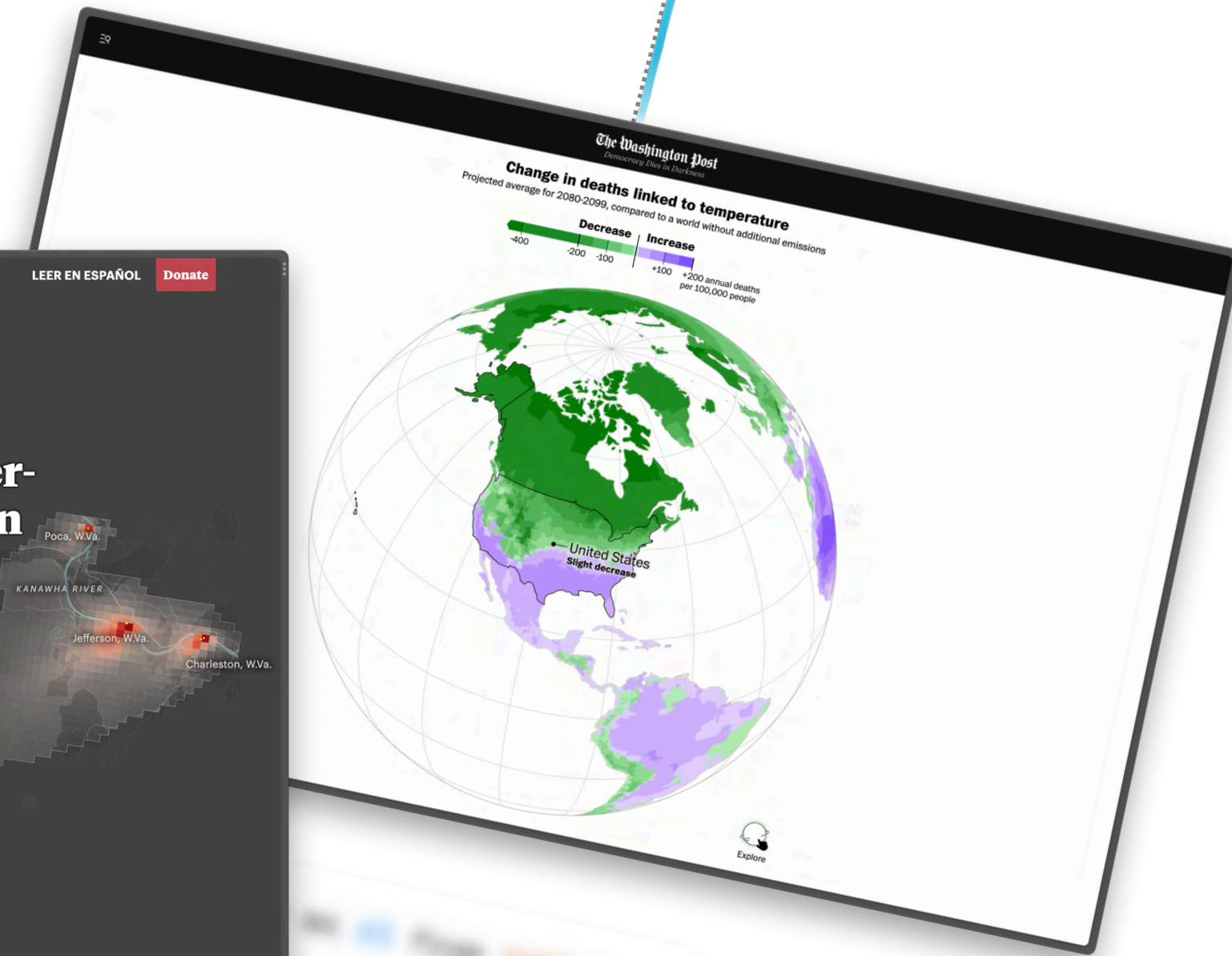
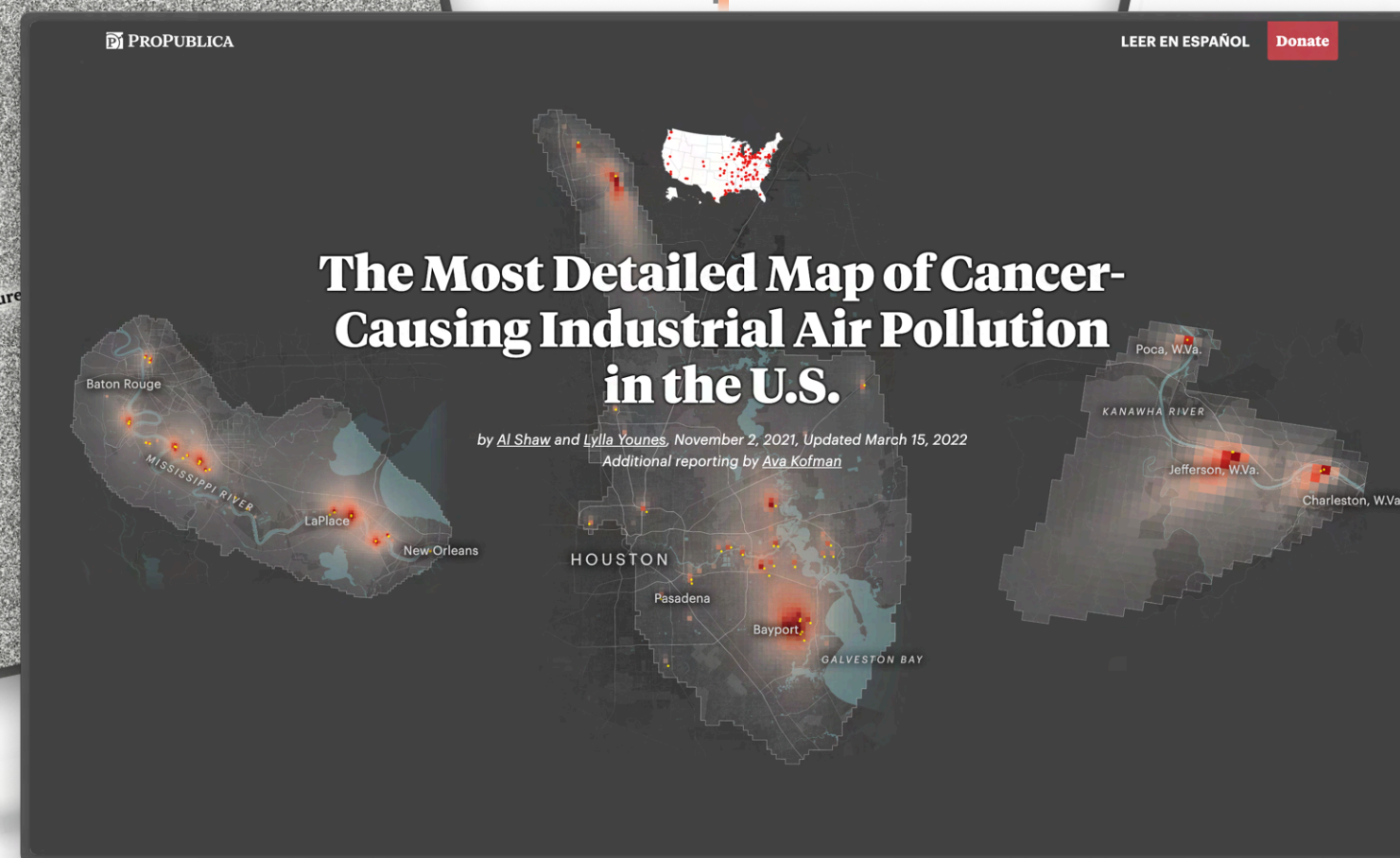
const plot = Plot.plot({
  y: { grid: true },
  marks: [
    Plot.line(covidDeaths,
      Plot.binX(
        { y: "sum" },
        { x: "date" },
      )
    ),
    Plot.ruleY([0])
  ]
});
```



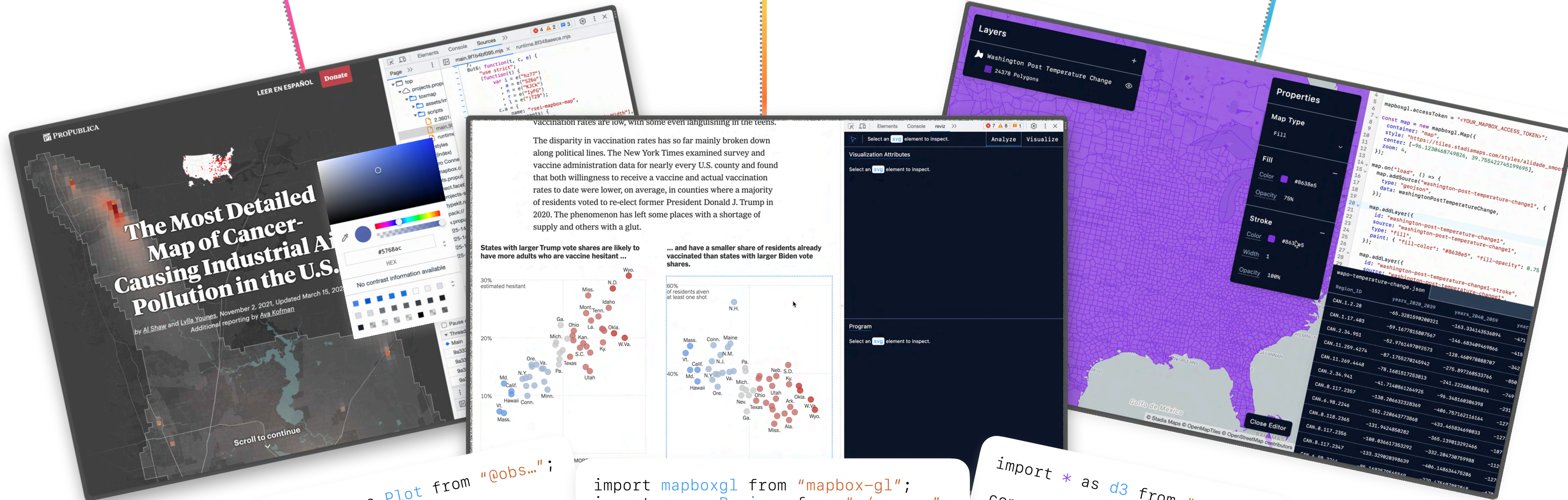
Fixed Entities



Fixed Entities



Live Objects



```
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.eyJ6fjh2...";  
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 d3 from "d3";  
const context =  
  canvas.getContext("2d");  
regionsGeo.features  
  .forEach((feature) => {  
    context.beginPath();  
    path(feature);  
    const c =  
      color(  
        props.years_2000_2009  
      );  
  });
```

Region_ID	years_2009_2009	years_2010_2009	year2
CAN_1.2.28	-65.328159280321	-163.33443336874	-671
CAN_1.17.463	-59.1677815887567	-166.48349469866	-671
CAN_2.36.951	-52.9761697892673	-129.46897888787	-645
CAN_11.269.4274	-87.1765278249942	-275.897268633766	-343
CAN_11.269.4668	-78.1681817253813	-241.92268684824	-858
CAN_3.34.943	-43.7148861264925	-90.34816886398	-769
CAN_8.117.2387	-138.286632328369	-486.757152136164	-231
CAN_6.98.2246	-182.228643773868	-633.46583469833	-127
CAN_8.118.2366	-131.942488282	-365.139813292466	-187
CAN_8.117.2386	-188.836617852992	-332.284738759988	-119
CAN_8.117.2347	-133.327828398639	-486.148634475286	-127

Live Objects

America is more diverse than ever — but still segregated

The United States is on track to be a majority-minority nation by 2044. But census data show most of our neighbors are the same race.

By Aaron Williams and Armand Emamdjomeh Updated May 10, 2018

- Black
- White
- Hispanic
- Asian/Pacific Islander
- Native American
- Multi-race and other

Chicago
Legacy segregation

Washington, D.C.

Each ■ represents 100 predictions

Birmingham, Ala. 0% White

Fort Meyers, Fla. 2% White

Elgin, Ill. 100% White

Haverhill, Mass. 97% White

Tacoma, Wash. 91% White

96% White

17% White

97% White

We analyzed more than five million predictions and neighborhoods with fewer predictions consistently had a higher population of White residents.

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

by Al Shaw and Lilla Younes, November 2, 2021, Updated March 15, 2022
Additional reporting by Ava Kafman

Scroll to continue

vaccination rates are low, with some even languishing in the teens.

The disparity in vaccination rates has so far mainly broken down along political lines. The New York Times examined survey and vaccine administration data for nearly every U.S. county and found that both willingness to receive a vaccine and actual vaccination rates to date were lower, on average, in counties where a majority of residents voted to re-elect former President Donald J. Trump in 2020. The phenomenon has left some places with a shortage of supply and others with a glut.

States with larger Trump vote shares are likely to have more adults who are vaccine hesitant ...

... and have a smaller share of residents already vaccinated than states with larger Biden vote shares.

Layers: Washington Post Temperature Change, 24378 Polygons

Properties: Map Type, Fill, Color #8638e5, Opacity 70%, Stroke, Color #8638e5, Width 1, Opacity 100%

```

const map = new mapboxgl.Map({
  container: "map",
  style: "https://tiles.stadiamaps.com/styles/allstate_smooth",
  center: [-96, 37],
  zoom: 4,
});
map.on("load", () => {
  map.addSource("washington-post-temperature-change1", {
    type: "geojson",
    data: washingtonPostTemperatureChange,
  });
  map.addLayer({
    id: "washington-post-temperature-change1",
    source: "washington-post-temperature-change1",
    type: "fill",
    paint: {
      "fill-color": "#8638e5", "fill-opacity": 0.70
    }
  });
  map.addLayer({
    id: "washington-post-temperature-change1-stroke",
    source: "washington-post-temperature-change1",
    type: "line",
    paint: {
      "line-color": "#8638e5", "line-width": 1, "line-opacity": 0.70
    }
  });
});

```

Region ID	years_2009_2009	years_2009_2009	years_2009_2009
CAN-1.2-28	-45.328159280321	-163.334143236874	-671
CAN-1.17-483	-59.1677815887567	-166.48349469866	-415
CAN-2.36-951	-52.9761697892573	-129.46897888787	-343
CAN-11.259-4274	-87.1765278249492	-275.897268633766	-858
CAN-11.269-4668	-78.1681817253813	-241.92268684824	-749
CAN-8.117-2387	-43.7148861264925	-90.34816836988	-231
CAN-8.98-2246	-138.286632328369	-484.757152136164	-127
CAN-8.118-2386	-181.942488282	-633.46583498833	-127
CAN-8.117-2386	-180.836617952992	-365.139812322466	-187
CAN-8.117-2347	-186.836617952992	-332.284738759988	-112
CAN-8.117-2347	-133.327828398639	-486.148634475286	-127

Average daily fog July–August 1996–2020

This delicate balance of forces creates a daily summertime ebb and flow. Late in the day, cool fog pulses inland. In the morning, it dissipates. Then the cycle repeats.

```

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.eyJ6f6fh2...";
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 d3 from "d3";
const context =
  canvas.getContext("2d");
regionsGeo.features
  .forEach((feature) => {
    context.beginPath();
    path(feature);
  });
const c =
  color(
    props.years_2000_2009
  );

```

FiveThirtyEight

Mexico 2nd REPS CONTROL

TISAN LEAN: D+3.9

CANDIDATE	CHANCE OF WINNING
Yvette Herrell (Rep.)	57 in 100
Gabriel Vasquez (Dem.)	43 in 100

FORECASTED VOTE SHARE

FiveThirtyEight Deluxe forecast, Oct. 5, 2022, at 8:20 PM.

Supporting Data Journalism through Compilers for Visual Inputs



 github.com/parkerziegler/reviz

 observablehq.com/@parkerziegler/hello-reviz

cartokit

 github.com/parkerziegler/cartokit

 alpha.cartokit.dev

Get in touch!

peziegler@cs.berkeley.edu

parkie-doo.sh

Parker Ziegler // @parker_ziegler

Ph.D. Student, UC Berkeley

 **Strange Loop**

St. Louis, MO • September 21, 2023