

CITY OF CAMBRIDGE

OPEN DATA WORKSHOP

APRIL 14, 2026

Who are we?

Alix Epstein

- Systems Data Analyst
- College librarian
- Public health background
- Running, trivia, crossword puzzles

Mohamud Suleyman

- Open Data Program Intern
- AI/Machine Learning Bootcamp Student
- Computer Science Undergraduate (Fall 2026)

Reinhard Engels

- Data Analysis & Open Data Program Manager
- Also an ex-librarian!
- Bioinformatics visualization at the Broad Institute
- Library Technology at Harvard and Cambridge Public Library

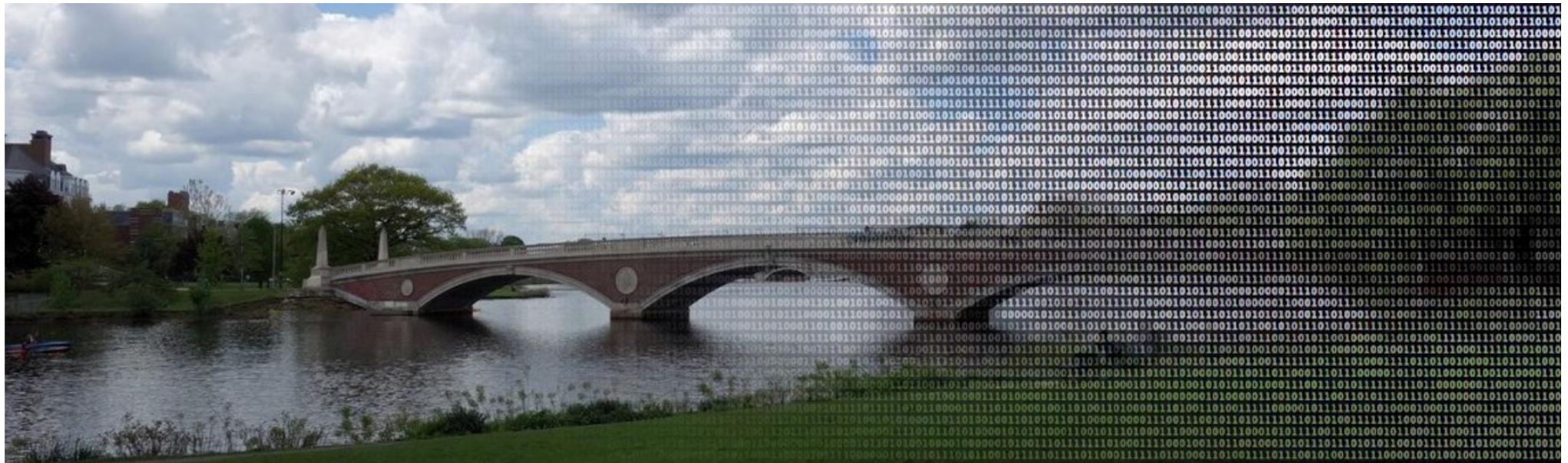
What is data?

- **Information in the form of facts and numbers, which can often be used to analyze something or make decisions**



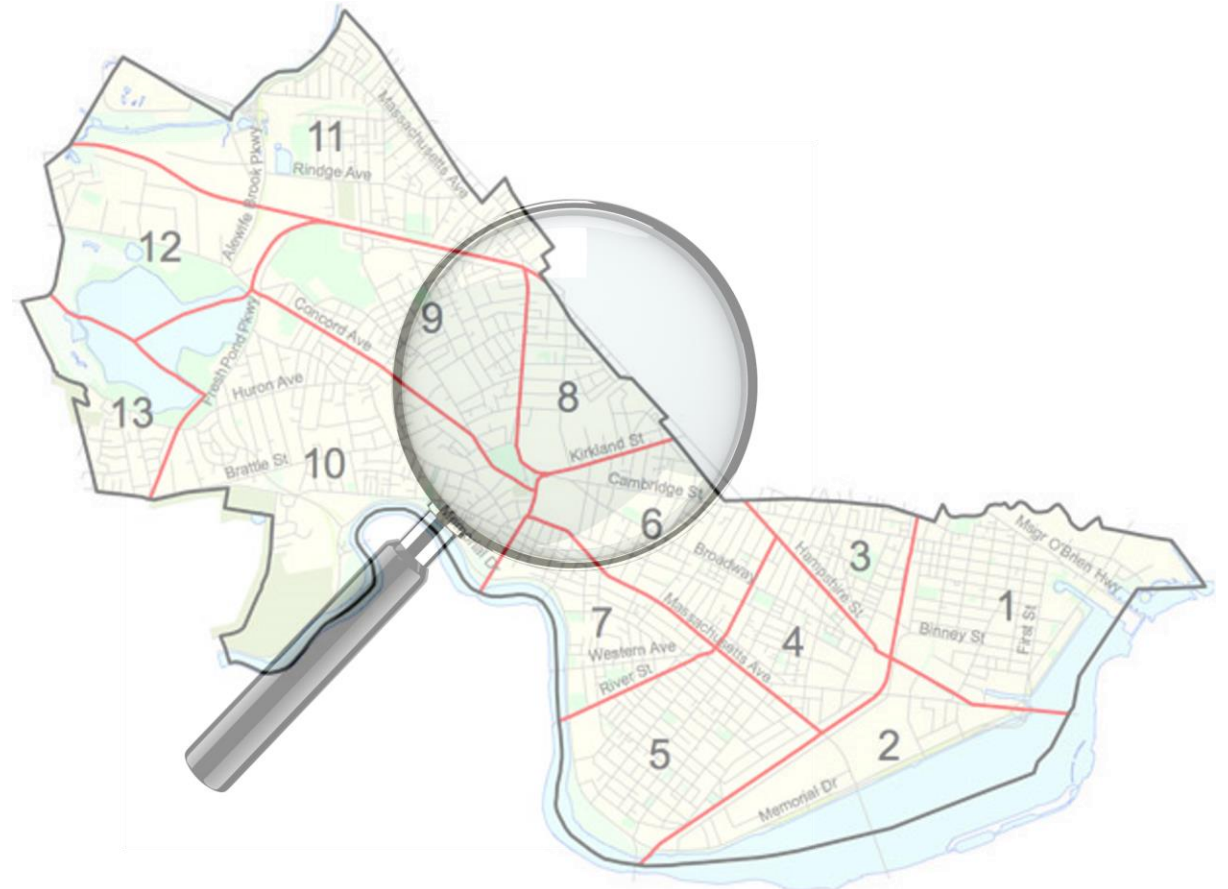
What is open data?

- **The idea that some data should be free and accessible to everyone, to use however they want, without restrictions**



The Cambridge Open Data Program

- **A set of systems, technologies, and policies for maximizing the value of Cambridge's data by sharing it, enabling city staff and the public to use it in various ways, making sure businesses have access to it free of cost, and removing barriers between city departments in data sharing**
- **Fostering transparency, encouraging community engagement, stimulating innovation – while prioritizing privacy and data security**



The Data Portal

data.cambridgema.gov

An official website of the United States government [Here's how you know](#) ✓

City of Cambridge

Quick Links ▾ Subscribe Translate ▾

Open Data Program

Search

Services ▾ I Want To ▾ Departments ▾

Open Data Portal
Explore hundreds of free datasets from across the City.

Cambridge 311 Performance Dashboard Now Available
It's the same tool our teams rely on every day to measure and improve how the City responds to SeeClickFix (311) service requests, and now we're giving you access to it too.

Open Data Workshop September 2025
Free, Hands-on course on Cambridge Open Data: all are welcome!

Gallery
Geographic Information Systems
Get Involved
Guides & Tutorials
Newsletter
Open Budget
Policy & Planning
Standards & Guidelines

The screenshot displays the City of Cambridge Open Data Portal interface. At the top, a dark blue header contains the site name, a search bar, and utility icons. A red circle highlights the search bar in the header. Below the header, a larger search bar is visible. On the left side, a 'Filters' sidebar is highlighted with a red circle. It includes sections for 'View Types', 'Categories' (with radio buttons for Assessing, Budget/Finance, General Government, Geographic Information (GIS), and Inspectional Services), and 'Tags' (with a 'Choose tag' dropdown). The main content area shows 405 results, sorted by 'Recently Added'. Three data sets are listed: 'Budget - Salaries- FY 2025' (Budget/Finance, 15 views), 'Outstanding Employee Awards' (General Government, 6 views), and 'Certificates of Occupancy' (Inspectional Services). Each entry includes a title, category, description, and metadata like creation date and view count.

Use the search box or the categories or tags to look for a data set

What kinds of data are on the Open Data Portal? What formats?

"Data Asset"

- Term for each individual result or item listed when you browse or search the portal
- Types of Data Asset:
 - Dataset
 - Chart
 - Files & Documents
 - Maps
 - Filtered Views

What kinds of data are on the Open Data Portal? What subjects?

Category

- City department producing or collecting the data
- There can be only one for each data asset
- Examples: Public Safety, Public Works, Traffic & Parking
- Exceptions: there are currently a couple of non-department categories for historical reasons

Tag

- Grab bag subject marker that can cut across departments or pinpoint something very specific within a department
- Each data asset can have multiple tags
- Examples: bicycles, permits, hvac

Metadata – data about the data

- Description
- Date last updated
- Date created
- Update frequency
- Category
- Tags
- License
- Number of rows
- Number of columns
- Rows
- Columns
- Source of data
- Limitations

The screenshot displays the metadata page for the 'Building Permits: Addition/Alteration' dataset on the City of Cambridge Open Data Portal. The page includes a navigation bar, a search bar, and a public access indicator. The main content area is divided into several sections: a description of the dataset, a 'Featured Content Using this Data' section, an 'About this Dataset' section with various statistics and details, and a 'What's in this Dataset?' section. The 'About this Dataset' section includes a table of metadata and a table of column headers.

Building Permits: Addition/Alteration Inspectional Services

List of permits related to the construction or the alteration of existing buildings in Cambridge. These permits cover construction that changes the structure, the footprint, or the volume of the building. This includes decks, porches, and additions, as well any reconstruction, renovation or repair to an existing structure. This dataset derives from...

Last Updated
June 9, 2024

Data Provided By
City of Cambridge Inspectional Services Department

[Read more](#)

Featured Content Using this Data

Choose to promote up to 3 assets that use this data, or choose to feature any relevant external resources. If left blank this section will be invisible. [Manage Featured Content](#)

About this Dataset

Updated June 9, 2024	Civic Innovation
Data Last Updated June 9, 2024	Problem Statements Related to this Data What can the permits' application dates reveal about patterns of economic activity in the city?
Metadata Last Updated March 4, 2024	Maintenance Plan
Date Created October 30, 2018	Estimated Update Frequency Weekly
Views 6,172	Specific Limitations
Downloads 1,131	Limitations Only completed applications are listed in this dataset.
Data Provided by City of Cambridge Inspectional Services Department	Topics
Dataset Owner Steven Espinoza	Category Inspectional Services
	Tags construction, alteration, housing and development, permits
	Licensing and Attribution
	License Open Data Commons Public Domain Dedication and License
	Source Link https://www.cambridgema.gov/inspection.aspx

What's in this Dataset?

Rows 10.6K	Columns 86	Each row is a Permit
----------------------	----------------------	--------------------------------

Columns (86)

Column Name	Description	API Field Name	Data Type
# Record Number		id	Number

Looking at the data

Each row is an item (in this case, a permit)
Each column is an attribute of that item

Building Permits: Addition/Alteration

itbuildi...	Number of Stories...	Building ...	Gross Square Foot...	Building Use	Owner-O...	Condo As...	Building Construction Type	Off-Stree...	Proposed...
	6	25	155355	Commercial / Mixed Use		False	Steel/Concrete, Non-Combustibl		
	3	45	50550	Commercial / Mixed Use		False	Steel/Concrete, Non-Combustibl		
	4	60	125200	Commercial / Mixed Use		False	Steel/Concrete, Non-Combustibl		
	6	62	25844	Commercial / Mixed Use		False	Steel/Concrete, Non-Combustibl		
	2	0	0	Commercial / Mixed Use		False	Masonry / Wood (Type III)		
	3	50	10000	Commercial / Mixed Use		False	Steel/Concrete, Non-Combustibl		

1 to 50 of 10,628

Querying (filtering) the data

Building Permits: Addition/Alteration Public Save As Edit ⋮

[← Back to Overview](#) [↔ Switch to Grid View](#) ↶ ↷ × Export

#	Build... build...	# Elect... elect...	# Plum... plum...	# Gas ... gas...	# HVA... hvac...	# T... t...	# Final... final...	# Fire ... fire_p...	T Desc... detail...	T Size ... size...	✓ Chan... floor...	✓ Chan... exter...
	\$3,119,225	\$1,063,837	\$197,130	\$0	\$973,730	\$5,700,292		\$346,370	Interior renov...	0	False	False
	\$170,000	\$36,000	\$35,000	\$5,000	\$35,000	\$296,000		\$15,000	Frame new i...	0	False	True
	\$28,000	\$0	\$0	\$0	\$0	\$28,000		\$0	Repair front ...	0	False	False
	\$26,100	\$0	\$0	\$0	\$0	\$26,100		\$0	Demoing offi...	0	False	False
	\$13,770	\$0	\$0	\$0	\$0	\$13,770		\$0	WEATHERIZ...	0	False	False

< 1 of 107 > Showing rows 1-100 of 10628

Filters | ✕ Clear all

Total Cost ▼ is greater than or equal to ▼ 5000 ×

+ AND ▼

Apply

Creating a visualization

Configure Visualization

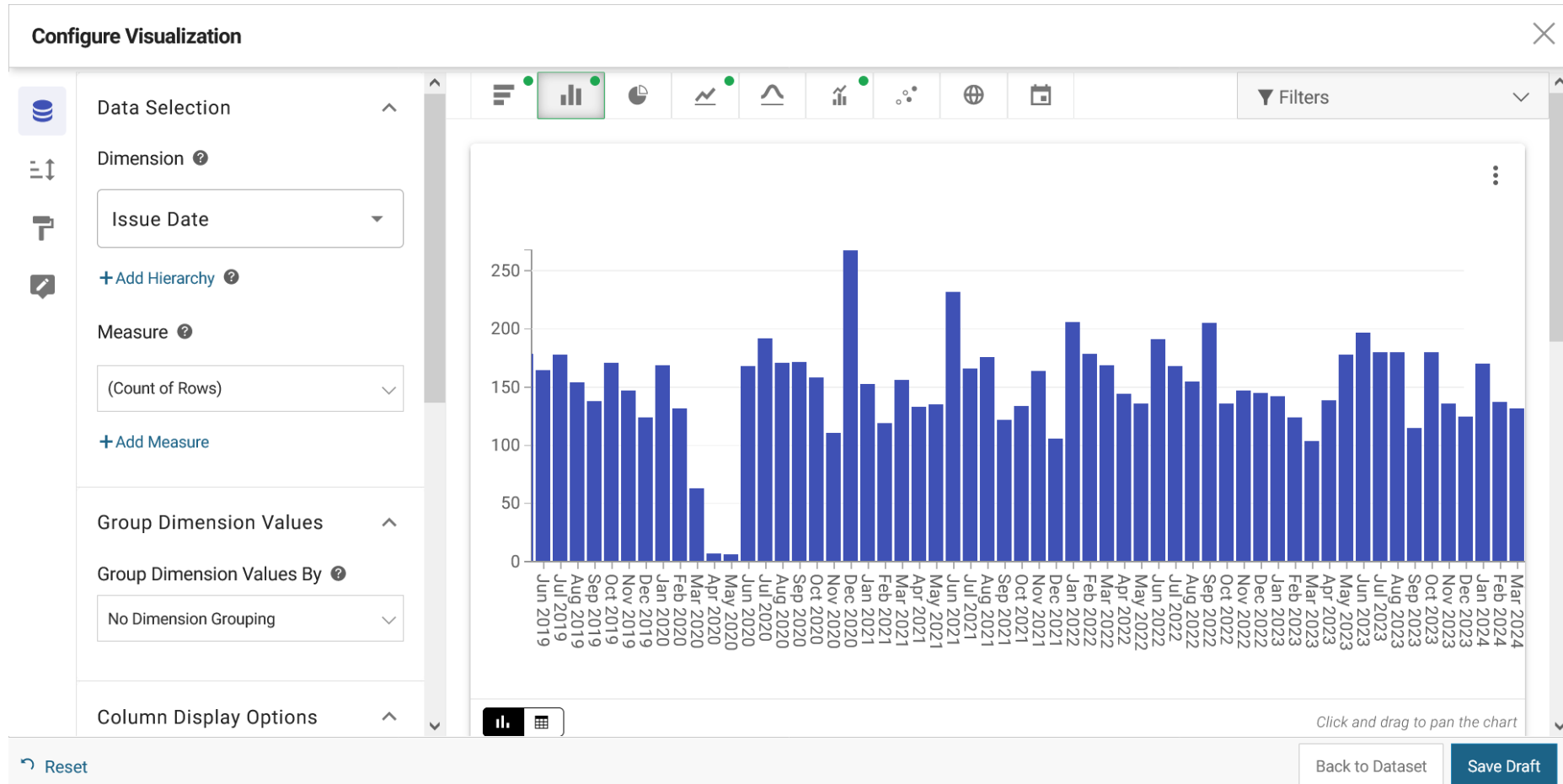
The interface is titled "Configure Visualization" and features a sidebar on the left with three main sections: "Data Selection" (with a list icon), "Dimension" (with a double-headed arrow icon and a "Select Column..." dropdown menu), and "Measure" (with a pencil icon). The top of the main area contains a toolbar with icons for various chart types: a list icon, a bar chart, a pie chart, a line chart, a line chart with a curve, a bar chart with a line, a scatter plot, a globe, and a calendar. A "Filters" dropdown is also present in the top right. The main area contains the following text:

Select data or a chart type to get started

There are two ways to get started. Select a chart type to see recommended dimensions and measures or select a dimension or measure to see recommended chart types.

At the bottom of the interface, there are three buttons: "Reset" (with a circular arrow icon), "Back to Dataset", and "Save Draft".

Bar chart



Map

Configure Visualization

← Layer List

Building Permits: Addition/Alteration

⚙️

Data Selection ^

Geo Column

Address for Mapping ▾

Resize Points by Value

None ▾

Style by Value

Status ▾

Point Aggregation ^

None

Heat Map

Region Map

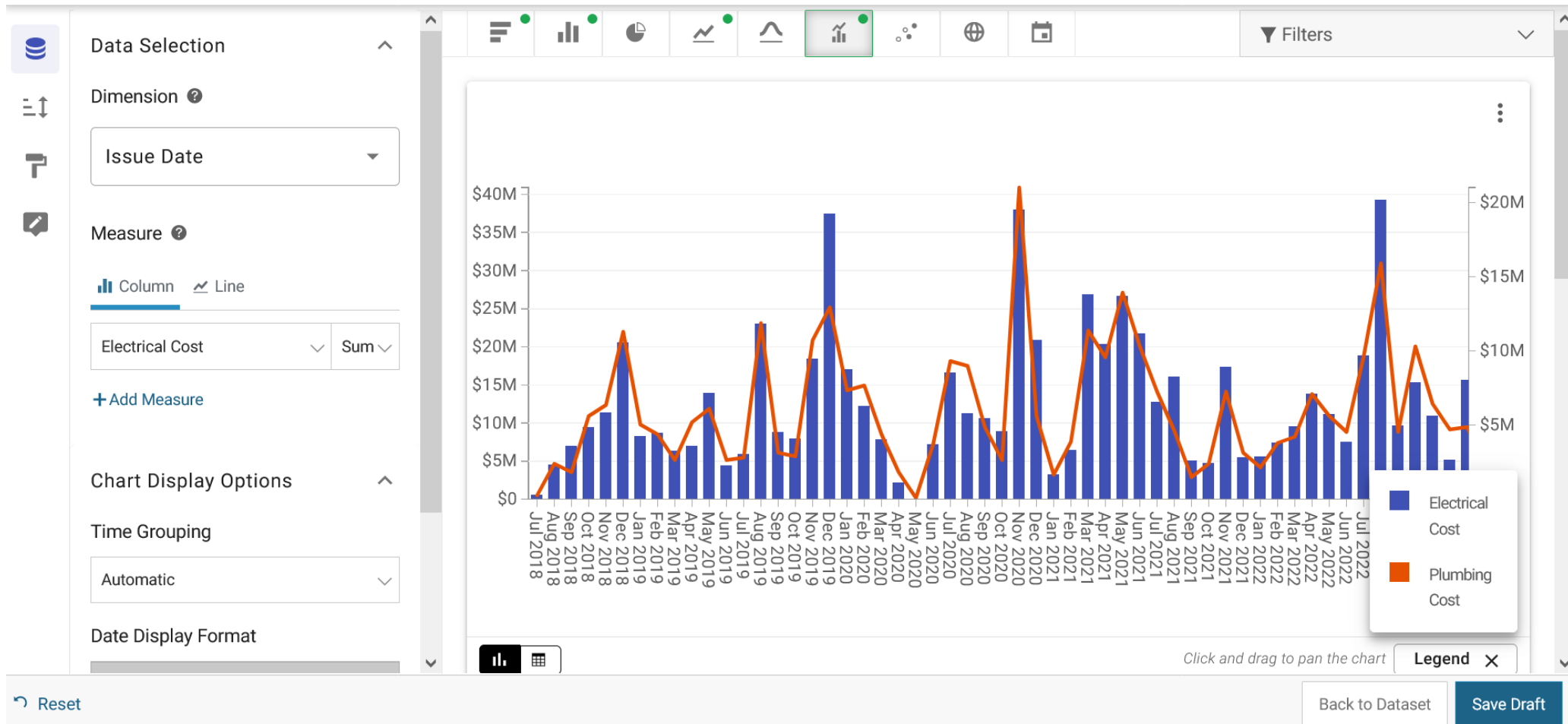
Reset

Back to Dataset Save Draft

The map displays a geographic area of Cambridge, Massachusetts, with numerous blue and orange circular markers representing building permits. The markers are densely clustered in the central and eastern parts of the city, particularly around the Avon Hill Historic District and the Cambridgeport area. The map includes street names such as Highland Ave, Main St, and Soldiers Field Rd. A search bar is visible in the top left of the map area, and a legend is located in the bottom right. The interface also features a toolbar with various visualization icons and a 'Filters' dropdown menu.

Combination chart

Configure Visualization



Saving a visualization

The screenshot displays the City of Cambridge Open Data Portal interface. At the top, there is a navigation bar with the text 'City of Cambridge Open Data Portal', a search bar, and several utility icons. Below the navigation bar, the main content area features a title 'Visualization - Based on Building Permits: Addition/Alteration' and an 'Export' button. A breadcrumb trail shows 'Inspectional Services' and a 'More Info' dropdown. A modal dialog box titled 'Name Your Visualization' is centered on the screen, containing a text input field with the text 'Plumbing vs electrical costs over time' and 'Cancel' and 'Save' buttons. The background visualization is a combined bar and line chart. The y-axis represents cost in millions of dollars, ranging from \$4M to \$40M. The x-axis represents time. The chart shows blue bars representing one data series and an orange line representing another. To the right of the chart is a 'Filters' panel with an 'Add Filter' button and a 'Clear All' link.

Publish for private use or publicly

City of Cambridge Open Data Portal Search

Plumbing vs electrical costs

Plumbing vs electrical costs

Inspectional Services

More Info

Publish this asset
Choose which audience can view the published version of this asset.

Private
Only individual collaborators and certain site roles can access

Public
The public can view

Cancel Publish

Save Publish... Export

Filters Add Filter Clear All

\$40M \$35M \$30M \$25M \$20M \$15M \$18M \$16M \$14M \$12M \$10M \$8M

Exporting Data: use Cambridge data with other tools

The screenshot displays the City of Cambridge Open Data Portal interface. At the top, the header includes the portal name, a search bar, and user navigation icons. The main content area is titled 'Crime Reports' and includes tabs for 'About', 'Data', and 'Related Content'. An 'Export' button is highlighted with a red box in the top right corner. A modal dialog box titled 'Export dataset' is open in the center, showing options for 'Download file' and 'API endpoint', an 'Export format' dropdown set to 'CSV', and a radio button selected for 'All data (95923 rows)'. A red arrow points from the 'Export' button to the dialog, and another red arrow points from the 'Download file' button to the 'Download' button at the bottom of the dialog.

APIs: use Cambridge data programmatically

Export dataset ✕

Crime Reports

[Download file](#) [API endpoint](#)

Data format
JSON

All data (95923 rows)

⚠ Default API limit exceeded

Our API has a default limit of providing 1,000 rows. [Learn more](#) about how you can modify the default limit.

API endpoint
`https://data.cambridgema.gov/resource/xuad-73uj.j`

[API documentation](#) [Developer portal](#)

[Cancel](#) [Copy to clipboard](#)

JSON Raw Data Headers

Save Copy Collapse All Expand All (slow) Filter JSON

```
0:
  file_number: "2024-04574"
  date_of_report: "2024-05-31T22:04:00.000"
  crime_date_time: "05/31/2024 22:00 - 22:03"
  crime: "Aggravated Assault"
  reporting_area: "1102"
  neighborhood: "North Cambridge"
  location: "0 CAMERON AVE, Cambridge, MA"

1:
  file_number: "2024-04572"
  date_of_report: "2024-05-31T22:02:00.000"
  crime_date_time: "05/31/2024 18:00 - 21:00"
  crime: "Larceny of Bicycle"
  reporting_area: "701"
  neighborhood: "Riverside"
  location: "0 PLEASANT ST, Cambridge, MA"

2:
  file_number: "2024-04573"
  date_of_report: "2024-05-31T21:33:00.000"
  crime_date_time: "05/31/2024 21:32 - 22:21"
  crime: "Missing Person"
  reporting_area: "504"
  neighborhood: "Cambridgeport"
  location: "200 FRANKLIN ST, Cambridge, MA"
```



<https://github.com/Chicago/RSocrata>



<https://github.com/socrata/socrata-py>

oData: live connections for excel, Tableau, PowerBI

The screenshot shows the 'Dogs of Cambridge' dataset page on the City of Cambridge data portal. The page includes a description of the dataset, a 'Last Updated' date of April 14, 2025, and a 'Featured Content Using this Data' section with two data stories. An 'About this Dataset' section provides metadata such as the update date (April 14, 2025) and the estimated update frequency (Annually). A red box highlights the 'Access via oData' option in the 'Actions' menu, which is also pointed to by a red arrow.

City of Cambridge Developers Help Q Sign In

Dogs of Cambridge General Government

This dataset displays the name, breed, and approximate location of dogs in Cambridge. It is based on dog license data collected by Cambridge's Animal Commission. All locations listed in this dataset have been obscured to protect privacy. Please see the limitations section below for more information.

Last Updated
April 14, 2025

Data Provided By
City of Cambridge Animal Commission

Featured Content Using this Data

- Dogs of Cambridge Data Story**
Public
November 3, 2023 26.7K Views
- Top Dog Names in Cambridge**
Public
April 14, 2025 1,276 Views

About this Dataset

Updated
April 14, 2025

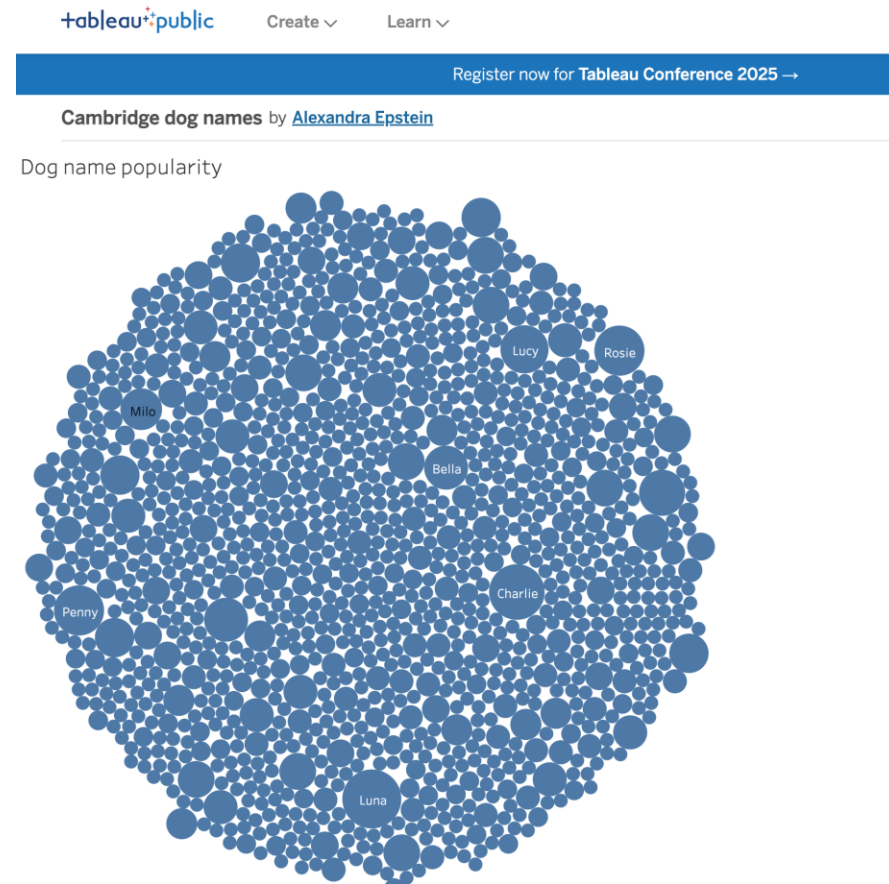
Data Last Updated: April 14, 2025
Metadata Last Updated: September 23, 2024
Date Created

Civic Innovation

Problem Statements Related to this Data: Where should new dog parks be located?

Maintenance Plan

Estimated Update Frequency: Annually



AI Tools: analyze Cambridge data with ChatGPT

Police_Department_Crash_Data_-_Updated_20...
CSV

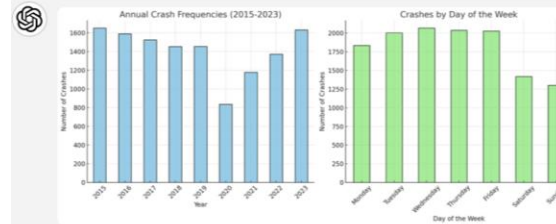
Analyze this Cambridge Police Department crash data. Do you notice any interesting or surprising trends? Please create charts to help me understand visually. Please exclude data from 2024 since it is incomplete.

The Cambridge Police Department crash data includes a vast array of columns, totaling 159, covering a wide range of details for each incident. Notable columns include:

- **Date Time:** The date and time of the incident.
- **Day of Week:** The day of the week on which the incident occurred.
- **Object 1 & Object 2:** Descriptions of the objects involved in the crash.
- **Street Number & Street Name:** The location of the incident.
- **May involve cyclist, May involve pedestrian, etc.:** Indicators of whether the crash involved cyclists, pedestrians, etc.
- Plus many more detailed attributes regarding the circumstances of each crash, involved parties, and outcomes.

Given the extensive nature of this dataset, a productive approach would be to identify key trends over time, such as annual crash frequencies, the involvement of cyclists or pedestrians, or crashes by day of the week. Since we should exclude data from 2024 due to its incompleteness, I will first filter the data accordingly and then proceed with analyzing trends and visualizing the results. (-)

Given these insights, I will now create visualizations for the annual crash frequencies and crashes by day of the week to provide a clearer understanding of these trends. The visualization for crashes involving cyclists and pedestrians will not be provided due to the unclear nature of the relevant data in the dataset. (-)



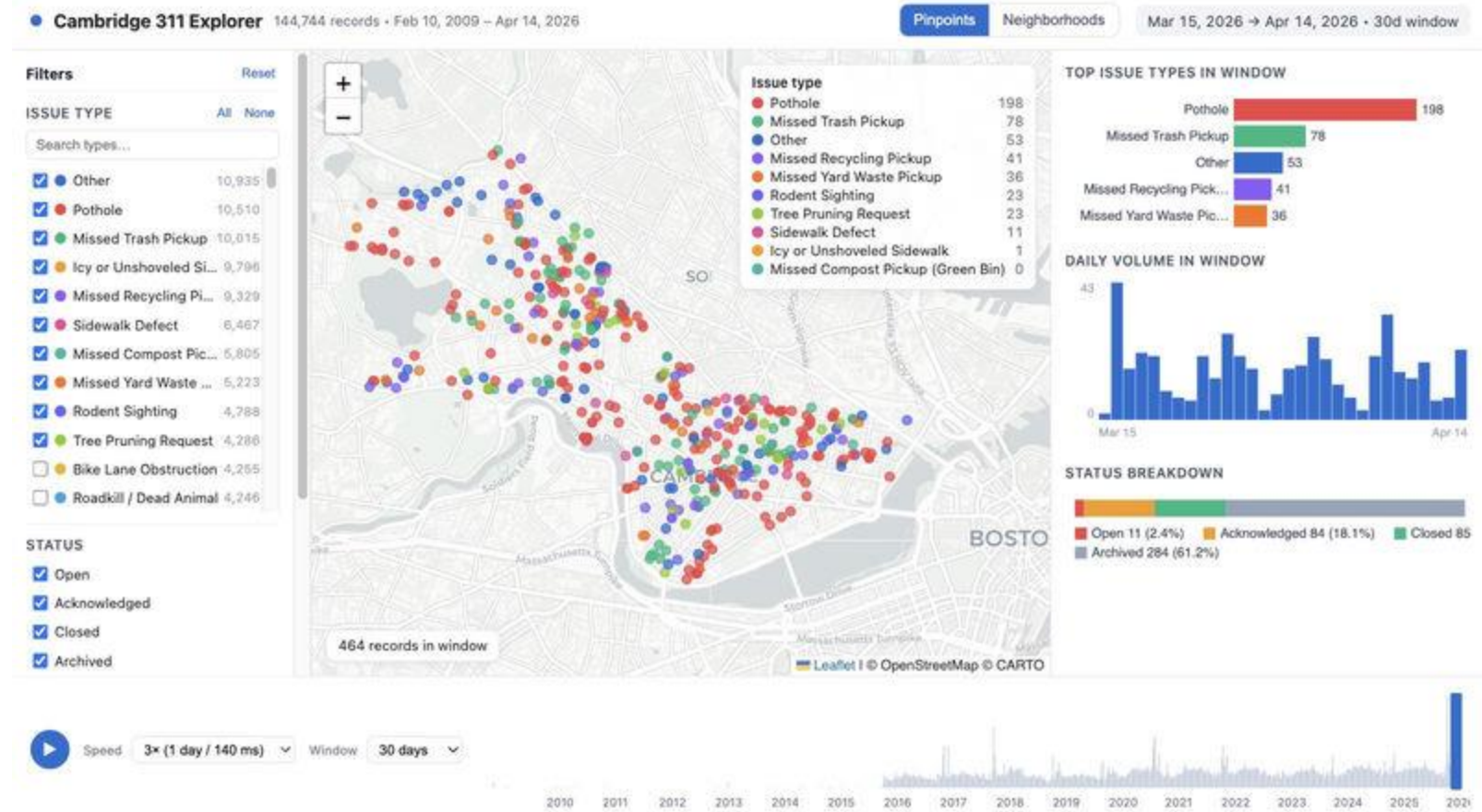
The visualizations offer a clear depiction of the identified trends in the Cambridge Police Department crash data:

1. **Annual Crash Frequencies (2015-2023):** This chart shows the number of crashes per year, highlighting the significant dip in 2020. Following this dip, there's a noticeable rebound in crash numbers, approaching pre-pandemic levels by 2023.
2. **Crashes by Day of the Week:** The frequency of crashes is higher on weekdays, with a peak on Wednesday, followed closely by Thursday and Friday. The number of crashes decreases notably on weekends, with Sunday having the fewest crashes.

These trends suggest a strong correlation between traffic patterns and the occurrence of crashes, with factors such as the COVID-19 pandemic affecting annual variations and typical weekly routines influencing daily fluctuations. (-)

AI Tools: Use Claude Code to "vibe code" an interactive dashboard

"Please create a graphical Cambridge 311 explorer that contains both a map, a timeline and charts and lets you set various filter criteria. You should be able to color code different types of feature. By default it should show previous 30 days of data, but you can drag the timeline around to show whatever the selected range is. It should be delightfully easy to move the time window around. As the timeline and filter criteria move around, the map and charts change accordingly. There should be a pinpoint mode on the map showing individual records, or an area mode showing densities by census blocks or cambridge neighborhoods. YThere should be a button to animate the timeline, which will then do a slowmo crawl over the data through time, updating the maps and filters accordingly. You should be able to set the speed of playback. ou can find the data here -- grab it all up front with a "loading data" indicator so that subsequent interactions are zippy. Eventually it will be GH pages site, but for now let's just do it locally



https://data.cambridgema.gov/Public-Works/Commonwealth-Connect-Service-Requests/2z9k-mv9g/about_data

Thank you!

- **Feedback:** <https://cambridgema.jotform.com/260984887814880> or use QR code
- **You will receive an email link to these slides**
- **Email us:** OpenData@CambridgeMA.gov
- **Thanks to our partners, the Cambridge Public Library**

