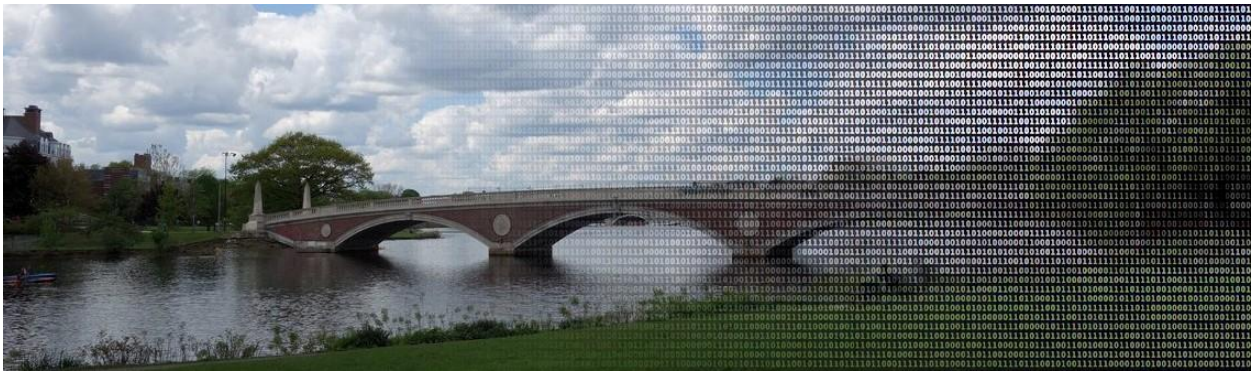
CITY OF CAMBRIDGE
OPEN DATA AND DATA ANALYTICS



STRATEGIC PLAN
FY2020-FY2022

CONTENTS

Executive Summary 3
Introduction 4
Part 1. Mission, Vision, and Strategic Framework 5
Mission 5
Vision 5
What It Means for Cambridge to be Data Driven 6
Strategic Framework 7
Part 2. Progress and Reflection 9
Data Supply 9
Data Analytics 11
Governance 12
Engagement 14
Capacity Building 18
Part 3. Moving Forward: Goals and Objectives 19
1. Evolve the Open Data Program into a Data Analytics and Open Data Program. 20
2. Focus on high value datasets when supplying open data 23
3. Tell stories with data 26
4. Involve the community in the data-driven problem solving and innovation 27
5. Build departmental capacity for data-driven problem solving 29
6. Prepare for the smart city and big data 31

EXECUTIVE SUMMARY

Data is an important tool for ensuring the wellbeing of Cambridge residents and stakeholders. Over the past three years, the City of Cambridge has made substantial progress in using data to be more transparent, efficient, and innovative. This strategic plan plots a three-year course for building on those initial successes and for evolving the Open Data Program into a Data Analytics and Open Data Program.

The plan envisions a program with a simple but powerful mission: Help the City of Cambridge leverage data to improve transparency, efficiency, and innovation so that Cambridge becomes an even better place to live, visit, work, and do business. This mission touches on the activities of nearly every city department and staff-person. It encapsulates much of the work the open data program has already done to build out Cambridge's data supply, implement data dashboarding across the City, bolster data governance processes, and improve public engagement with open data.

To continue our journey over the next three fiscal years, the plan suggests six program goals, each with multiple objectives:

1) *Evolve the Open Data Program into a Data Analytics and Open Data Program*. By creating a home for new urban analytics technologies, decision-support tools, and data training initiatives, this program will help the city leverage data-driven insights to streamline operations and drive innovation. In doing so, it will also help surface new open datasets to share with the public. This new goal helps formalize many of the duties already performed by the open data program.

2) *Focus on high value datasets when supplying open data*. By developing new practices for prioritizing potential open datasets, we can improve the efficiency of open data operations and focus limited resources on those datasets that would be most valuable to the public.

3) *Tell stories with data*. By engaging the public in open data through stories instead of just spreadsheets, we will broaden our audience and find new partners to help us innovate.

4) *Involve the community in data-driven problem solving and innovation*. By bolstering our outreach and partnership efforts, we will engage a more diverse and skilled team in tackling Cambridge's most pressing issues.

5) *Build departmental capacity for data-driven problem solving*. By developing new trainings and tools for using data in daily operations, we will help prepare departments for a future in which all municipal staff use data-driven insights to help them achieve their mission.

6) *Prepare for the era of the smart city and big data*. By assessing new data analysis tools and techniques, the Data Analytics and Open Data Program will help ensure that the City is well-prepared to leverage new sensors and data-streams.

These goals, described in greater detail in part 3 of this plan, are intended to work in concert as we continue pursuing the effective use of data and data-driven insights. Data and analytics are already changing every aspect of modern life. Municipal government is no exception. The City of Cambridge is well positioned to leverage data to improve the wellbeing of Cambridge's residents and stakeholders.

INTRODUCTION

The following strategic plan sets a course for Cambridge's open data and data analytics work for fiscal years 2020 through 2022. The plan should be read as a guide for growing and transforming the Open Data Program into a greater source of value for the City of Cambridge, rather than as a manual for the program's daily operations. In addition to pursuing the initiatives outlined in this strategic plan, Cambridge's data staff will remain focused on continuous improvements to the City's data quality, availability, and privacy policies.

To produce the plan, we first clarified the program's mission and vision, as well as a strategic framework for making long term decisions. These elements are described in Section 1 of the plan, and together they serve as a compass, pointing the direction in which Cambridge's staff should take the program. In Section 2, we take stock of how far we've come, detailing some of the Open Data Program's major accomplishments to date. Finally, in Section 3 we outline our roadmap for data analytics and open data during fiscal years 2020-2022. To create it, we compared our past progress against our strategic vision to identify major goals and objectives for the next three years.

Someone once said, "Plans are useless, but planning is indispensable." We are clear-eyed about the fact that parts of this plan will change over time. Some goals and objectives may fall away, while new ones might emerge. We can't predict exactly what Cambridge's data efforts will look like by the end of 2022, but we are confident that they will be stronger because we created this strategic plan.

PART 1. MISSION, VISION, AND STRATEGIC FRAMEWORK

MISSION

The mission of the Data Analytics and Open Data Program described by this strategic plan will be to help the City of Cambridge leverage data to improve transparency, efficiency, and innovation so that Cambridge becomes an even better place to live, visit, work, and do business.

VISION

The wellbeing of residents, visitors, workers, and organizations in Cambridge is improved through the use of data and analytics, which help make the delivery of city services more transparent, efficient, and innovative. Residents are better informed, engaged, and empowered because they have access to open data and data-driven insights. Data is governed wisely and strategically throughout the organization so that all datasets—whether open or internal—balance transparency and usability with privacy and security. City datasets are timely, accessible, usable, and understandable. Insights gleaned from data are clearly communicated, accurate, and actionable. City staff and leadership can effectively and ethically use data to track performance, investigate trends, and improve decision-making. The Data Analytics and Open Data Program not only provides City departments with analytical support, but also builds analytical capacity across City departments so that all municipal employees feel empowered to use data to achieve their mission.

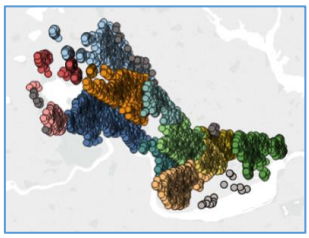
WHAT IT MEANS FOR CAMBRIDGE TO BE DATA DRIVEN

A Cambridge firefighter jumps into the passenger seat of a firetruck as her partner turns on the siren and pulls the vehicle into the street. They'll arrive at the blaze in three minutes. The firefighter opens the public safety portal on her tablet, and sees that a predictive analytics model has assigned this building and those next to it high-risk scores for fires. By the time she steps out of the truck, our firefighter has learned how many people live in the building, where each entrance is located, how many bedrooms are in each unit, and what materials were used during construction.

It's early October, the air is crisp, and a second-grade class is outside learning about trees. The classroom teacher is using Cambridge's interactive foliage map on her tablet to identify nearby tree species and to show her students the colors each tree will turn in a few weeks. She points at a sugar maple whose leaves are still green. Then she spins the screen toward the students to show them the deep red the maple's leaves will soon become.

Public comment begins at tonight's City Council meeting. A resident steps up to the microphone and asks the Council to consider making an improvement in his neighborhood. He confidently cites statistics he calculated from several datasets on the open data portal.

A Cambridge resident reports to her first day of work at a startup near Kendall Square. She got the job by impressing the company's founder with an interactive data dashboard she created during one of Cambridge's Staff & Residents Open Data Visualization trainings. A mile away, a city planner is using the skills he learned at that same training to build an affordable housing dashboard that will inform new development policies.



It's Tuesday evening, and a group of local software developers and undergraduate students is meeting to build open source software that helps solve problems for local governments. Four coders huddle around a laptop screen, pointing at interesting challenges listed on Cambridge's "Civic Innovation Challenge Inventory." They choose a challenge about analyzing municipal greenhouse gas emissions, open a related dataset, and begin working on a solution.

STRATEGIC FRAMEWORK

When Cambridge's City Council passed the Open Data Ordinance in September of 2015, it instructed the City to view data as a tool for achieving greater transparency, efficiency, and innovation, while instructing staff to seek balance between making data available and protecting privacy. Though the program has grown in scope and scale since then, these same values remain its guiding principles. One of the most important achievements of the last two years has been the creation of a framework for running and growing the Open Data Program in accordance with these principles. The framework defines a successful Open Data Program as one that moves the City of Cambridge forward in five areas: data supply and automation, data analytics, data governance, user outreach and community engagement, and internal capacity building. See Figure 1.

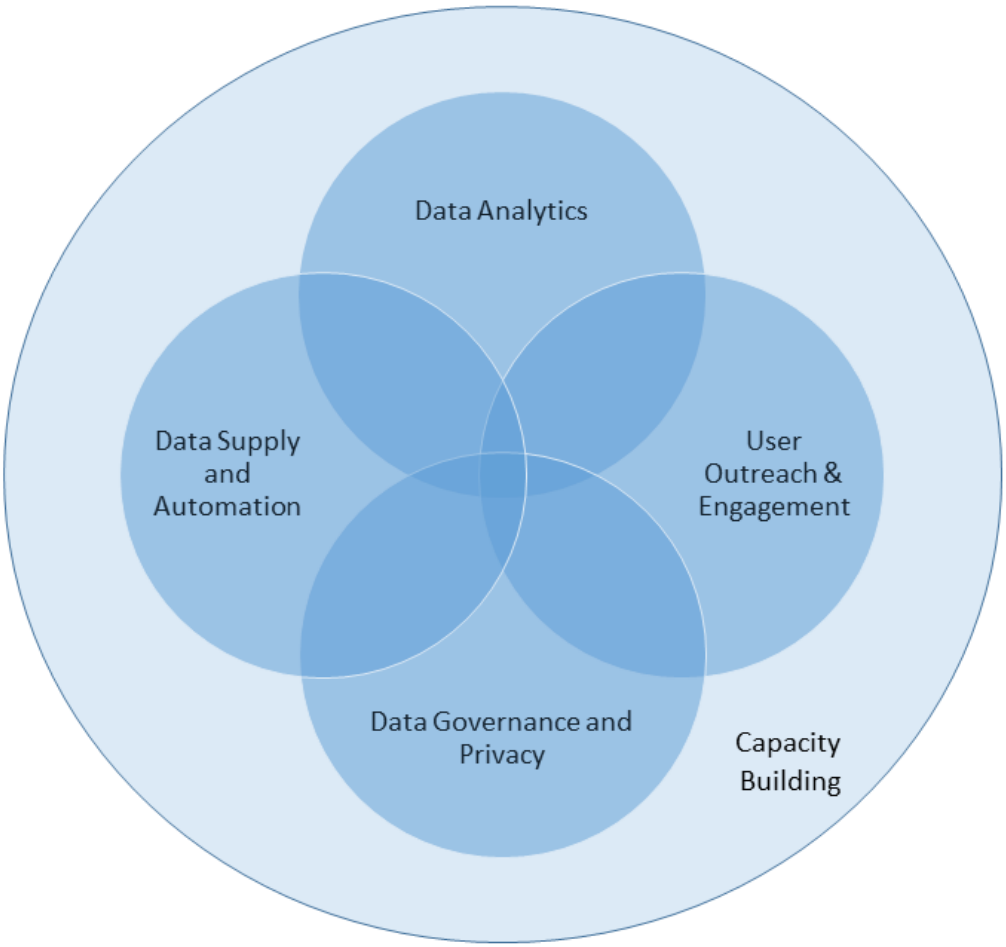


Figure 1: Framework underlying Cambridge's open data and data analytics work

Data supply and automation are efforts to share high quality data among departments and with the public. This fosters transparency and improves decision-making. Automation cuts the cost of supplying data while simultaneously improving data quality, thus increasing the net value of data.

Data analytics are techniques and technologies for shaping data into insights that help improve decision-making and make delivery of city services more efficient.

User outreach and community engagement are activities for marketing open data to the public and for empowering external stakeholders to participate in the innovation process. Outreach has multiple value propositions: transparency, economic development, civic engagement, and public-private partnership.

Data governance maximizes data's value while minimizing risks to privacy. Good data governance ensures that the correct data reaches the correct audience at the correct time in the correct format.

Capacity building constitutes efforts to spread data skills throughout all City departments. The Data Analytics and Open Data Program envisioned by this plan should be the tip of the spear, developing cutting edge tools and practices for data supply, analytics, community engagement, and data governance, while at the same time empowering departments to leverage data on their own.

PART 2. PROGRESS AND REFLECTION

This review is an important part of our strategic plan. It provides an overview of our major accomplishments and highlights since the passage of the open data ordinance in September 2015. Together, these achievements and opportunities for improvement point the way toward an even more successful program.

DATA SUPPLY

ACHIEVEMENTS

We've facilitated a steady increase in the number of open datasets added to our open data portal. Responsive supply of high quality data to the public remains the cornerstone of Open Data Program. See Figure 2.

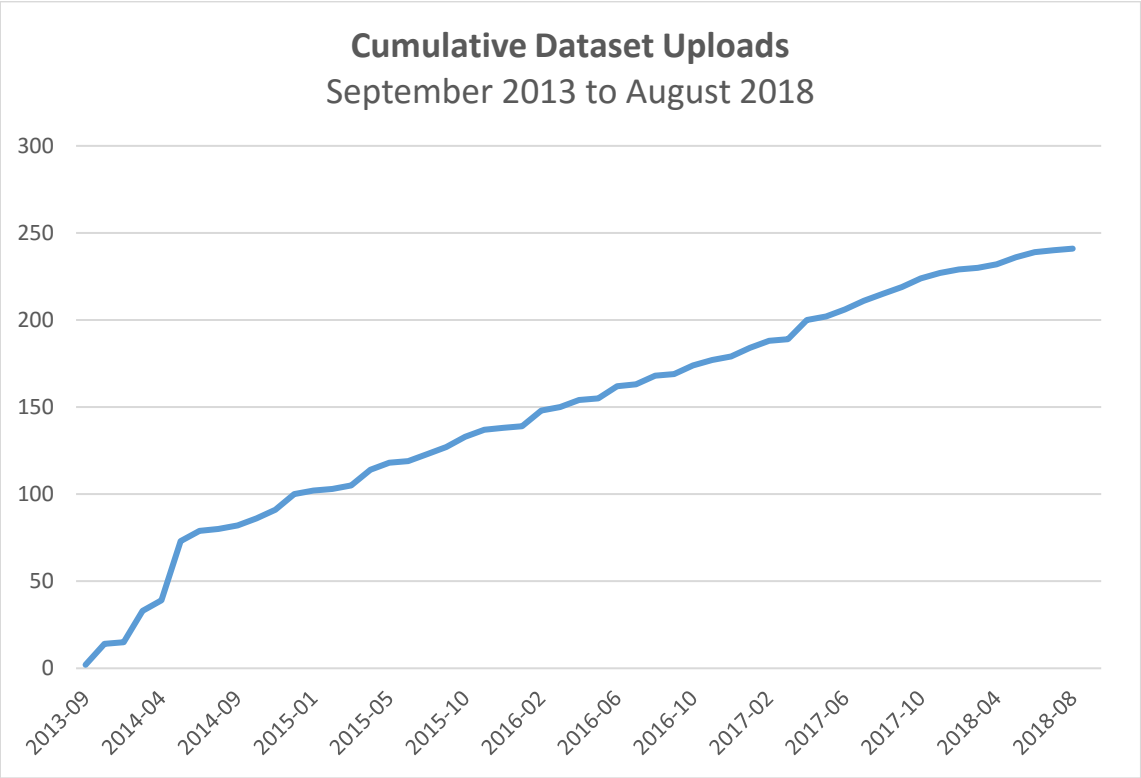
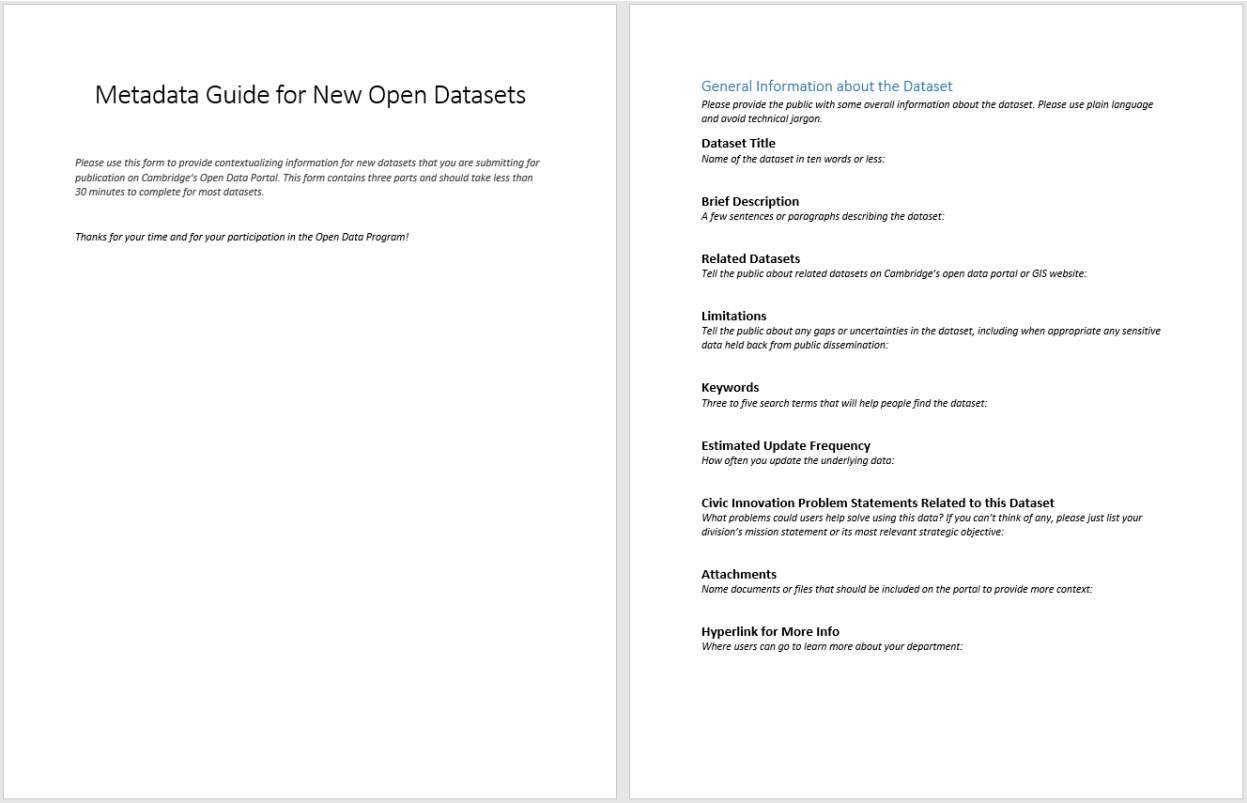


Figure 2: Cambridge has seen a steady increase in its supply of open data

We've improved our open data publishing processes. Subject matter experts can now more easily provide open data staff with prospective open datasets and metadata (contextual information about the data in an open dataset). The new process improves the speed with which we can provide new open datasets and reduces staff time requirements. More importantly, it allows subject matter experts to focus their efforts on improving data quality and surfacing new data instead of learning new software that's unrelated to their mission. See Figure 3.



Metadata Guide for New Open Datasets

Please use this form to provide contextualizing information for new datasets that you are submitting for publication on Cambridge's Open Data Portal. This form contains three parts and should take less than 30 minutes to complete for most datasets.

Thanks for your time and for your participation in the Open Data Program!

General Information about the Dataset

Please provide the public with some overall information about the dataset. Please use plain language and avoid technical jargon.

Dataset Title
Name of the dataset in ten words or less:

Brief Description
A few sentences or paragraphs describing the dataset:

Related Datasets
Tell the public about related datasets on Cambridge's open data portal or GIS website:

Limitations
Tell the public about any gaps or uncertainties in the dataset, including when appropriate any sensitive data held back from public dissemination:

Keywords
Three to five search terms that will help people find the dataset:

Estimated Update Frequency
How often you update the underlying data:

Civic Innovation Problem Statements Related to this Dataset
What problems could users help solve using this data? If you can't think of any, please just list your division's mission statement or its most relevant strategic objective:

Attachments
Name documents or files that should be included on the portal to provide more context:

Hyperlink for More Info
Where users can go to learn more about your department:

Figure 3: Updated data publishing procedures include a new metadata worksheet to expedite the upload process.

The implementation of automated updates has increased data quality and timeliness while drastically decreasing staff time spent on some open data updates. At present, approximately 20 percent of our tabular datasets and the majority of geospatial datasets use automated updates. See Figure 4.

```
url=paste(FTPsite,yesterday,Sensor,sep="")
destination="yesterday.csv"
download.file(url, destination)

### Read in and format csv file

data=read.csv("yesterday.csv",header=FALSE)
data=data[-c(1,2,3),1:5]  #Delete top three r
colnames(data)=c("Date","Time","Total","Westb
data$Date=as.Date(data$Date,"%d.%m.%Y")
data$Day=weekdays(data$Date)
data$DateTime=as.POSIXct(paste(data$Date,data
data=data[,c("DateTime","Day","Date","Time","
```

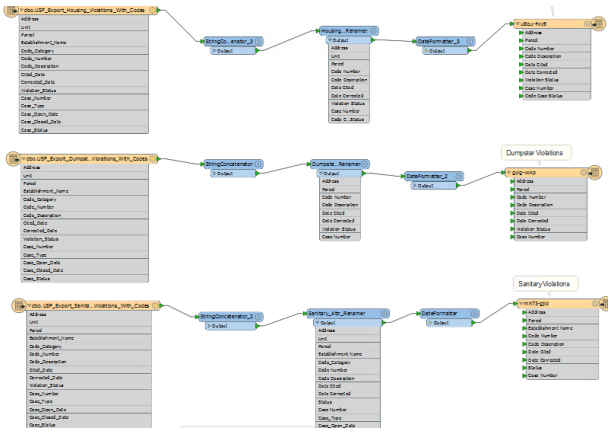


Figure 4: The City uses extract, transform, and load techniques to reduce the time demands of updating datasets.

OPPORTUNITIES FOR FUTURE IMPROVEMENTS

- A stronger emphasis on high value open datasets could help the City more efficiently supply open data while carving out time for new data analytics and community engagement activities.
- The implementation of systems for querying the public about data priorities would help the City more easily identify high value datasets.
- Standardized metadata keywords and categories could help residents search for datasets more effectively.

DATA ANALYTICS

ACHIEVEMENTS

The introduction of data dashboards has empowered City staff to more efficiently use data to drive decisions and improve public engagement efforts. Data dashboards are an information management tool for visualizing key insights from a dataset to help stakeholders improve decision making and transparency. The open data program led efforts to evangelize the use of dashboards for data visualization and exploration. Internal dashboards help department staff explore trends and track performance. Public-facing dashboards help engage residents in Cambridge open data. They also help the City protect privacy in the case of datasets that are too sensitive to be released in disaggregated, tabular format. Initial successes in this area include Cambridge's diversity dashboard, pay equity tool, and green buildings dashboard. See Figure 5.

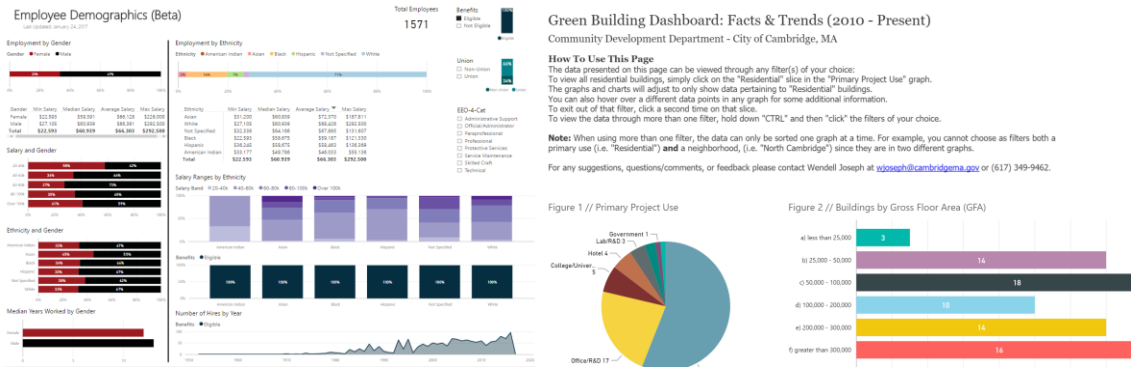


Figure 5: Two of our early dashboards. Employee Demographics (left) and Green Buildings (right).

The implementation of new technologies for creating analytics web apps sets the stage for Cambridge to rapidly deploy online tools to engage staff and residents. Our first app was our interactive foliage map, which was built with the City Trees Open Dataset (See Figure 6). As Cambridge begins conducting more sophisticated data analyses, these web-based tools will become crucial for communicating results and insights.

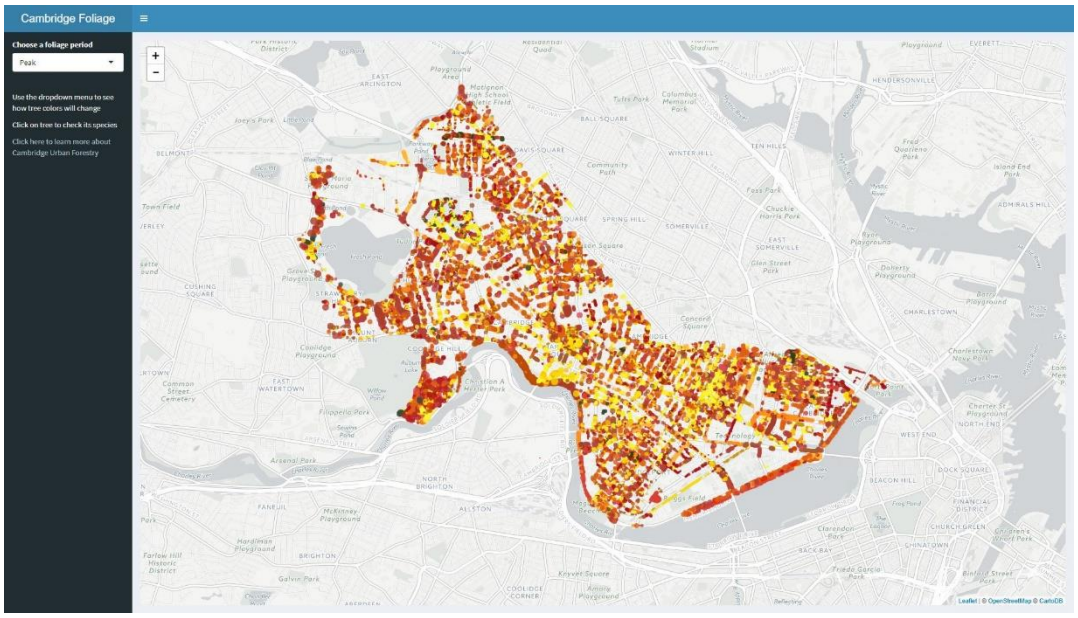


Figure 6: Cambridge's interactive foliage map.

Cambridge's open data program has become an important source of ad-hoc consulting for departments and divisions who don't have their own data analysts. Managers from other departments will often engage the open data program for help with data analysis tasks or for conversations about how to use data more effectively.

OPPORTUNITIES FOR FUTURE IMPROVEMENTS

- Better internal communications and marketing could help department staff and leadership more clearly understand the value proposition of data analytics.
- More rigorous project management procedures for analytics projects would help ensure that projects are as effective, efficient, and equitable as possible.
- Analytics projects should be thought of as a way to surface new open datasets.
- The formalization of a data analysis helpdesk service would enable departments to get help with small data engineering and analytics tasks such as connecting to APIs and conducting relational joins.

GOVERNANCE

ACHIEVEMENTS

The creation of the Open Data Review Board established an important new source of accountability for Cambridge's Open Data Program. It also created a new venue for stakeholders to discuss open data policies. The purpose of the review board is to advise the City Manager and program manager on policies and procedures for using data in Cambridge. The review board includes four city staff, representing Cambridge's Community Development Department, IT Department, Emergency Communications Department, and City Manager's Office; it also includes four city residents, representing academia, the startup community, and the civic technology community.

The development of a new open dataset review process created a formal but flexible approval process for releasing new datasets. The process, developed in coordination with Bloomberg's What Works Cities, was designed to more stringently assess datasets with a high privacy risk while expediting the release of low-risk datasets. See Figure 7.

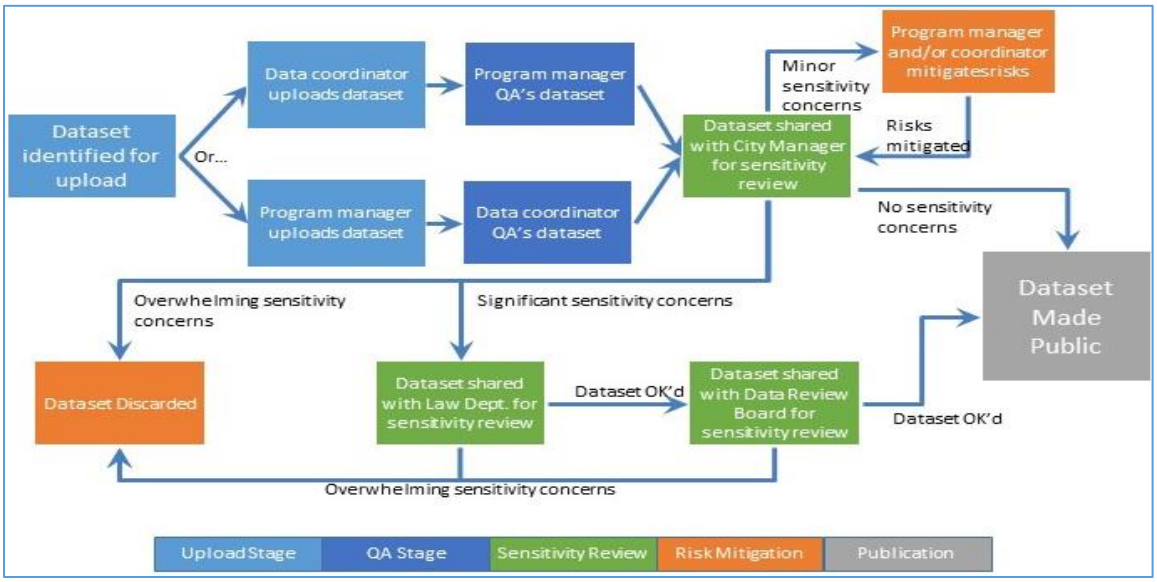


Figure 7: A flowchart illustrating the review process for new open datasets.

Cambridge was acknowledged for its contribution to Harvard Law School's [Open Data Privacy playbook](#). This guide lays a foundation for "risk-based privacy," a new field that will help open data cities like Cambridge make more informed decisions about data governance. See Figure 8.

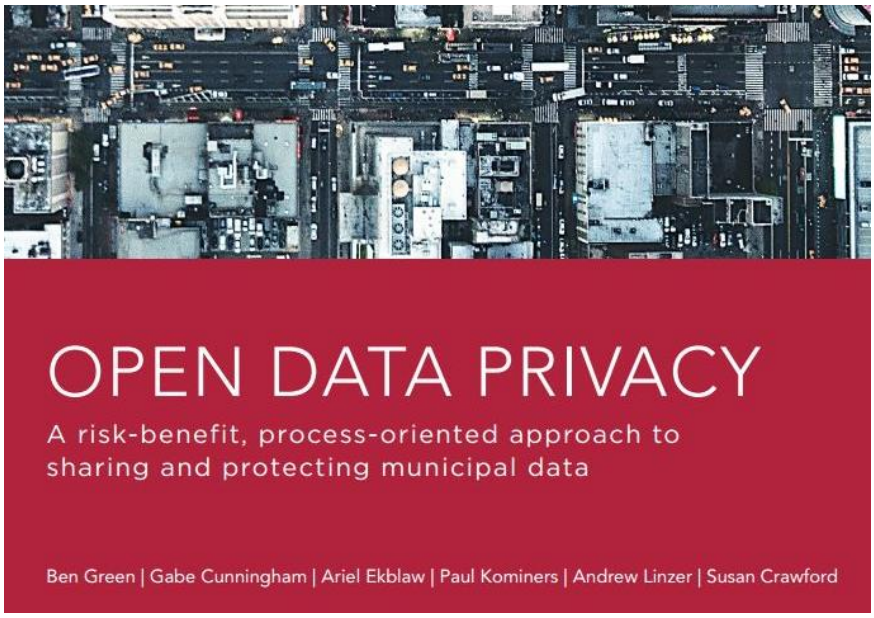


Figure 8: Cambridge contributed to a Harvard study on open data governance.

OPPORTUNITIES FOR FUTURE IMPROVEMENT

- Improving resources for internal data sharing would help break down information silos between departments and enable more data-driven collaboration.
- Research into privacy risks that emerge when multiple datasets are combined (the so-called "mosaic effect") could help improve data governance procedures in Cambridge and other open data cities.

ENGAGEMENT

ACHIEVEMENTS

Cambridge has deepened its partnership with Code for Boston over the last two years. This group is an important source of talent for using Cambridge open data to solve community issues. Volunteers at groups like Code for Boston also introduce city staff to new tools and techniques for analyzing data. Some of the work the group has done using Cambridge open data is available on the Code for Boston Github Repository.

Cambridge's Amplifying Civic Innovation report helped the City identify new ways to engage staff and residents around open data. During summer 2016, Cambridge hired an Open Data Research Fellow from Harvard's Kennedy School to investigate how we could better partner with the civic tech community to drive forward innovation using open data. Her report won national recognition. See Figure 9.

Figure 9: Amplifying Civic Innovation, by Jennifer Angarita

An improved open data website revamped the face of the Open Data Program and gave program staff more flexibility to market Cambridge's open data in new and interesting ways. The new design adds a project showcase, improved calls to action, and links to other data resources that complement the open data portal. See Figure 10.

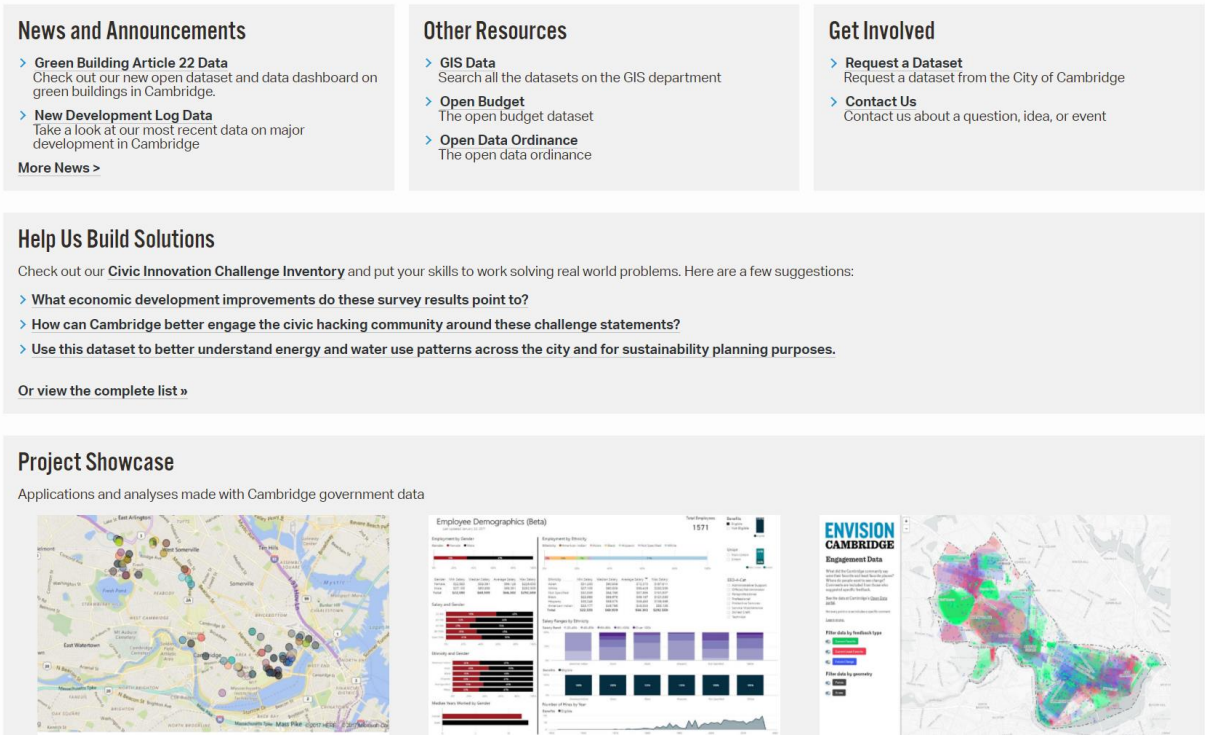


Figure 10: The new website includes a project showcase and multiple calls to action

We created Cambridge's first Civic Innovation Challenge Inventory to help facilitate collaboration between Cambridge's subject matter experts and the civic tech community. The inventory helps volunteers understand the most productive use cases for Cambridge open data. See Figure 11.

Figure 11: A State Scoop article about the civic innovation challenge inventory

The implementation of our first data story, Dogs of Cambridge, set the stage for Cambridge to broaden public engagement around open data. Data Stories intersperse open datasets, visualizations, and analyses with narrative

content such as text and video. By humanizing open data, Cambridge can attract a wider audience and show more residents the value of data. See Figure 12.

Figure 12: Our first data story, "Dogs of Cambridge"

Cambridge's first data visualization training for residents and staff was a great success. The training, which was conducted in partnership with Cambridge Community Television, taught participants how to use free software to transform Cambridge open data into interactive dashboards. The training empowered new open data champions in the community and created a novel venue for residents and city staff to meet each other and interact. See Figure 13.

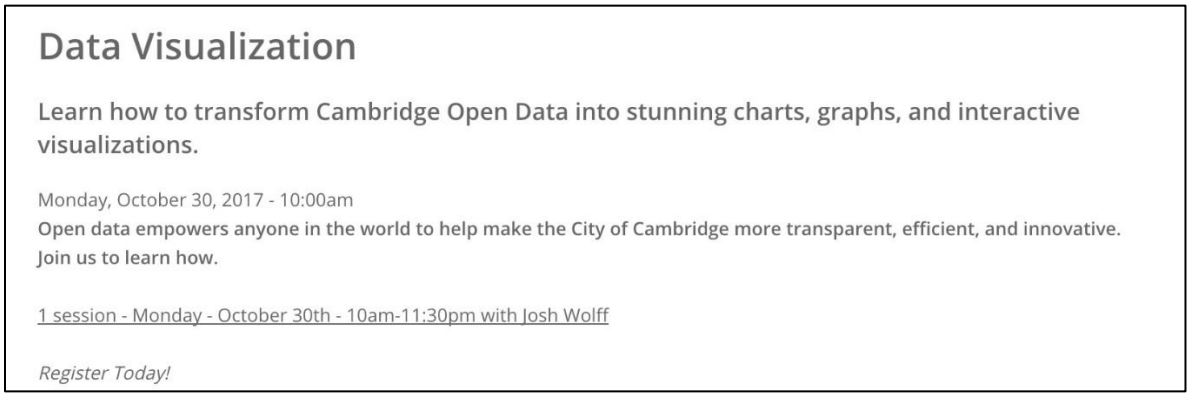


Figure 13: An ad for our first data visualization training event.

OPPORTUNITIES FOR FUTURE IMPROVEMENT

- The open data program can reach a more diverse audience through outreach to local schools and universities and by building on efforts such as public data visualization training.
- Initial efforts at data storytelling should be continued and expanded to engage a wider and more diverse online audience.
- The civic innovation challenge inventory was a successful first step for linking open datasets with the City's strategic initiatives; program staff should build on this effort across the open data website by more tactically using keywords and data collections.

CAPACITY BUILDING

ACHIEVEMENTS

Data trainings for municipal staff have helped executive leaders and project managers find new opportunities to use data and analytics to solve important issues. Cambridge's Open Data Program and City Manager's Office partnered with the Johns Hopkins University Center for Government Excellence in late 2018 to train staff from departments across the City on data governance, analytics, and storytelling.

Data visualization trainings have empowered staff to create their own data dashboards. Dashboards are a powerful tool for decision support and public engagement. By providing trainings that teach department staff to create their own dashboards, the Open Data Program free up time to work on more complex analytics projects.

Data Office Hours help engage city staff around open data. The open data program has had success hosting "office hours" sessions around the City to spur conversations about how data can be integrated into department activities. These informal meetings have helped spawn new data projects and give department staff a venue for asking data questions.

The convening of a city data analysts lunch group has provided a venue for analysts to come together to share best practices and peer review each other's work.

OPPORTUNITIES FOR FUTURE IMPROVEMENT

- Additional classes beyond data visualization training could more comprehensively train staff to work with data.
- By providing departments with more formal resources for hiring data talent and procuring data-related software, Cambridge's open data program could help build data talent across the city workforce.

PART 3. MOVING FORWARD: GOALS AND OBJECTIVES

The goals and objectives described in this section represent a three-year roadmap for building on the Open Data Program's initial successes and for evolving it into a Data Analytics and Open Data Program. Figure 14 illustrates the hierarchy we used during the planning process to conceptualize the relationship between the program's vision, strategic framework, goals, and objectives.

Each goal is a specific course of action that:

- Aligns with the mission and vision of Cambridge's Data Analytics and Open Data Program.
- Aligns with one or more facets of the strategic framework discussed in Section 1.
- Builds on the successes and opportunities outlined in the progress report from the previous section in this strategic plan.

Each goal comprises multiple objectives, which are discrete and well-defined projects intended to grow and transform the program. For each objective, we describe its purpose, status, and an action or deliverable that will indicate whether the objective has been achieved.

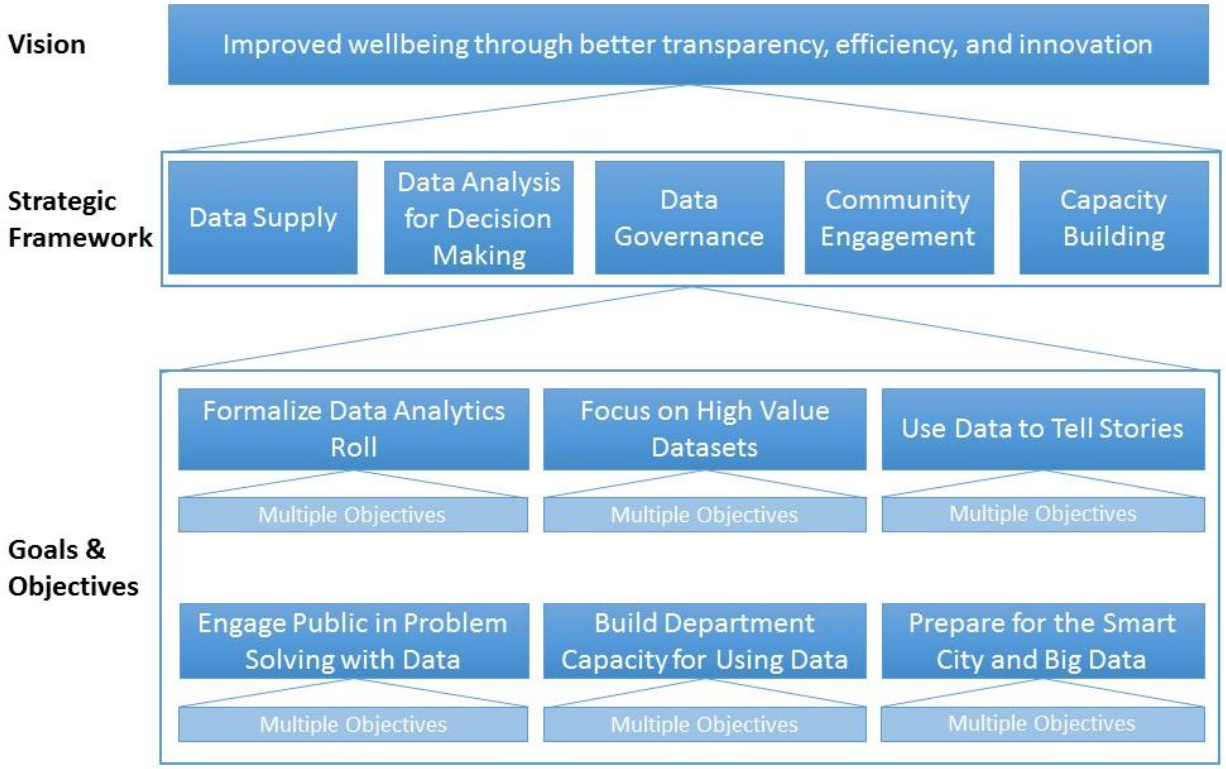


Figure 14: The strategic planning hierarchy.

1. EVOLVE THE OPEN DATA PROGRAM INTO A DATA ANALYTICS AND OPEN DATA PROGRAM.

Data analytics is the process of transforming datasets into useful insights that can improve decision-making and resource allocation. Some departments in Cambridge, and many cities around the country, have already begun using data analytics to make operations more efficient. Studies show that the use of data analytics and data-driven decision-making improves organizational productivity by 5-6 percent.[1] For Cambridge, a city with a $600 million annual budget and thousands of employees, this productivity gain implies a significant return on investment. Cambridge's Data Analytics and Open Data Program can help the City realize this potential by working with departments and external stakeholders to create tools such as data dashboards and predictive algorithms. These tools will help city staff and residents make more data-driven decisions and reduce wasted resources.

Data analytics isn't just about coding and statistics. Significant effort must go into managing projects, collaborating with subject matter experts, and communicating insights to stakeholders. The objectives below lay out a path for formalizing Cambridge's Data Analytics and Open Data Program.

Objective	**Purpose and Description**	**Action and Deliverable**	**Status**
1.1 Create a communications plan for marketing data analytics services to other departments.	A communications plan is necessary so that department managers and subject matter experts understand the value of data analytics and understand that analytics services are not limited to open datasets.	Collaborate with City communications managers on a written plan for internal communications and branding.	Planned 2020
1.2 Educate decision makers about data analytics and how to use it to drive better outcomes.	Data analytics is only useful if it leads to better decisions and service delivery. Decision makers must understand how to act on analytical insights and how to identify analytical limitations.	Create or procure trainings for the City's senior managers, possibly as part of a Cambridge data academy.	In Progress 2020
1.3 Implement software for creating interactive data web applications so that non-programmers can access data analytics.	To be widely used, data analytics must be accessible by those who do not know programming languages like R and Python.	Procure software for publishing interactive data web apps. Publish first app.	Complete

[1] Brynjolfsson, Erik and Hitt, Lorin M. and Kim, Heekyung Hellen, Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? (April 22, 2011). Available at SSRN: https://ssrn.com/abstract=1819486 or http://dx.doi.org/10.2139/ssrn.1819486

1.4 Implement systems for identifying, prioritizing, managing, and documenting new data science projects.	Much of the labor involved in data analytics goes into identifying, scoping, documenting, and managing projects. Business processes and forms are necessary to streamline these operations. A decision will need to be made about whether these projects fit under EGov.	Create project submission and documentation systems. Create project management templates and project dashboards. Align with EGov as necessary.	In Progress 2020
1.5 Implement business processes for linking data analytics and performance measurement.	Data analytics and performance measurement are complementary though not synonymous. A Data Analytics and Open Data Program could help create the statistical and technological infrastructure for a performance measurement program (e.g., quantitative metrics, dashboards, A/B tests) and in turn, ongoing performance tracking can surface problems and opportunities that need more rigorous analysis via data science techniques.	Begin collaborating with City Manager's Office and Budget Office to identify how data analytics and performance measurement might be linked but differentiated. This exercise should inform the communications plan from 1.1.	Planned 2021
1.6 Implement business processes for linking data analytics and open data.	Data analytics and open data are complementary. The value proposition of data analytics encourages departments to surface new data that can then be added to the open data portal. Conversely, the open data portal centralizes much of the City's data in one location for analysis. Most importantly, the open data portal allows us to "show our work" on data analysis when used in concert with tools like GitHub.	Ensure that analytics project management templates contain a section for stakeholders to identify new open datasets that will result from each project.	Planned 2020
1.7 Pilot the city's first data science project and create system for capturing lessons learned and marketing successes.	The new program will need a quick win that demonstrates the value proposition of data analytics to other internal customers.	Complete the first project of the official new program.	Planned 2020
1.8 Implement a data analytics help line for quick data analysis tasks.	Sometimes, department staff might not need a weeks-long data science engagement, but rather an hour of technical assistance calculating some statistics or connecting to an API.	Create and market a data analytics help line email group.	Planned 2021
1.9 Implement policies and practices aimed at harnessing the benefits of artificial intelligence (AI) while reducing associated risks.	AI and machine learning has the potential to significantly improve the performance and efficiency of city operations, but the City will need to carefully and deliberately work to mitigate bias and increase transparency from new AI tools.	Document policies and procedures for reducing the impact of demographic or geographic bias in new AI tools. Develop processes and venues for continuously updating AI practices based on stakeholder input.	Planned 2020

2. FOCUS ON HIGH VALUE DATASETS WHEN SUPPLYING OPEN DATA

Every open dataset requires substantial city resources to create and update, but only some datasets receive significant attention from the public. Because resources are limited, they should be spent on those datasets that would be most valuable to the public. The charts in Figure 15 clearly illustrate that nearly all public attention on the open data portal goes to a small number of datasets. Whether a dataset gets 100 views per month or none at all, however, it requires on average approximately ten hours of staff time to launch and two hours per year to maintain. Even the simplest dataset takes several hours of preparation before it is ready to be posted on the open data portal. Clearly, systems to prioritize new open datasets are necessary to ensure the that open data activities are operating efficiently. The objectives below will help the program better discern which potential datasets will provide the most value in terms of transparency, civic engagement, and innovation.

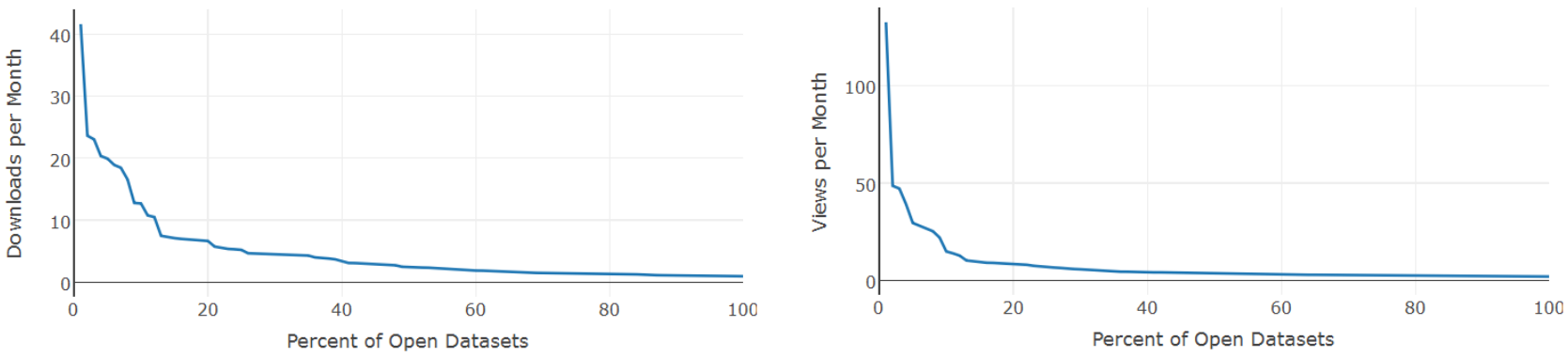


Figure 15: Only about 10 percent of open datasets receive substantial public attention.

Objective	Purpose and Description	Action and Deliverable	Status & Estimated Completion
2.1 Query the public to get feedback about data topic priorities.	The best way to identify high value open datasets is to let the public tell us what they want. By providing the public with a list of potential open datasets or data topics, we can get direct feedback from our customers about data priorities.	Publicly release a list of datasets or data topics that the city could offer as open data.	In progress 2021

2.2 Begin collaboration with public records office to better understand demand for City datasets	Public record requests are another important indicator of demand for different types of city data. By collaborating with the City's public records access officer, the Data Analytics and Open Data Program can better identify which new datasets will be particularly useful to the public.	Implement business processes that allow the sharing of insights between the public records access officer and the open Data Analytics and Open Data Program manager.	In progress 2020
2.3 Implement systems that position data analytics as a source of new open datasets.	Data used for internal analysis should be made open whenever possible. Datasets that are particularly useful to City staff are also likely to be valuable the public. Moreover, by providing analytical datasets as open data, we "show our work" and ensure transparency. At the same time, internal analytics customers should understand that data analytics services are not limited to open data—department staff should feel assured that confidential data will remain protected.	Create business processes for surfacing new open data through data analytics projects. This should likely be a section of the new project planning template (see 1.5 and 1.6)	In progress 2020
2.4 Align open data efforts with the City's strategic goals and initiatives.	Open data should be useful to staff as well as the public. Systems should be in place to prioritize datasets that relate to strategic initiatives, Council goals, and interdepartmental initiatives. Of importance is linking open data and performance measurement.	Create business processes for linking potential open datasets with strategic goals and performance metrics.	Planned 2020
2.5 Implement performance metrics that incentivize usefulness over quantity in Cambridge's Data Analytics and Open Data Program.	Municipal data programs are often judged based on the number and pace of tabular datasets made public. This creates perverse incentives to supply many low value datasets rather than focusing on high value datasets. A better strategy for evaluating open data efforts is to use a broad set of performance metrics, rather than only one or two.	Implement a balanced set of performance metrics for tracking progress of the city's data efforts.	In progress 2021
2.6 Create tools for analyzing and mitigating cross-dataset privacy risks from open data	Today's open data privacy risks arise from individual datasets; tomorrow's risks will arise when multiple datasets are combined and de-anonymized. Cambridge's Data Analytics and Open Data Program should continue developing and sharing tools for mitigating this risk. These tools could take the form of new technologies or business process improvements.	Create tools for mitigating inter-dataset privacy risks.	In Progress 2022

2.7 Identify gaps in Cambridge's open data offerings by surveying popular datasets in other cities with robust open data programs.	One way to identify which datasets might provide high value to Cambridge residents is to look for datasets that have been popular in other cities with large open data programs.	Perform gap analysis by comparing Cambridge's offerings to the top 10 most popular datasets in the 5-10 largest open data programs nationwide.	Planned 2020

3. TELL STORIES WITH DATA

In October 2017, Cambridge published a "data story" about dog licenses, and immediately shattered the City's record for single-day dataset viewership. Thousands of people who never knew Cambridge had an open data program were suddenly learning more about city services through numbers and interactive visualizations.

Most people don't want to pore over spreadsheets; they want to engage with a narrative about the meaning behind the numbers. That's why using data to tell stories is so important. Data storytelling helps us attract wider and more diverse public participation as we endeavor to increase transparency and data-driven decision making.

Objective	**Purpose and Description**	**Action and Deliverable**	**Status & Estimated Completion**
3.1 Implement software for telling interactive data stories	Few residents seek out tabular datasets to explore. Data is more engaging when presented along with a narrative about its meaning. By telling "data stories," Cambridge will attract new data users and demonstrate the process of transforming raw data into actionable insights.	Implement software for data storytelling.	Complete 2020
3.2 Create Cambridge's first data story	Departments across the City will need to be shown the value proposition of data storytelling. The first story should be a quick win and relate to a fun topic.	Release first data story	Complete
3.3 Release new data stories at a regular and reasonable cadence	Our initial assumption is that over the next three years we can reasonably release one data story every three months. During year one, we will test this assumption by tracking the time necessary to develop data stories.	Track staff time necessary to create data stories, and identify reasonable publication schedule.	In Progress 2021
3.4 Emphasize data visualization, data storytelling, and interactive content for new tabular datasets	Initial experience with data stories and visualization indicates that these tools have the potential to drastically increase engagement with open data. Department subject matter experts and Data Analytics and Open Data Program staff will need to collaborate to identify data story opportunities and craft the narratives. Not all new datasets will be good candidates for data stories. However, almost all new datasets should at least be published alongside simple plots or maps. This increased emphasis on content will require increased staff time per dataset. Therefore, a shift in focus to high value datasets is a prerequisite for effective data storytelling and engagement.	Create business processes for identifying and managing data storytelling opportunities and for facilitating collaboration between subject matter experts and data analysis staff.	In progress 2021

4. INVOLVE THE COMMUNITY IN THE DATA-DRIVEN PROBLEM SOLVING AND INNOVATION

A founder at a major technology company used to say: "No matter who you are, most of the smartest people don't work for you." That's no longer true in the age of open data. Nowadays, Cambridge can use our Data Analytics and Open Data Program to equip an army of community volunteers with the data they need to help us solve our toughest problems. In 2017 we became the first city to offer an "innovation inventory" to help marshal the efforts of community problem solvers. The steps below outline a plan for building on that early success.

Objective	**Purpose and Description**	**Action and Deliverable**	**Status & Estimated Completion**
4.1 Modify the Cambridge innovation challenge inventory (CICI) to include strategic initiatives and Council goals	The CICI has been an important tool for collaborating with the public to use open data to help Cambridge solve problems. With the release of Council goals and department objectives, the problem statements in the CICI should be standardized and aligned with those elements.	Create a pick list of Council goals and department objectives for staff filling out CICI metadata fields.	In Progress 2020
4.2 Develop and pilot public data visualization training	Data visualization is one of the most efficient ways for the public to use Cambridge open data to influence policy decisions. By training the public to visualize data, we increase the number of residents who can help the City solve problems and innovate.	Pilot a public data visualization course at CCTV.	Complete
4.3 Develop policies for posting community projects to a project showcase on the open data portal and for integrating civic innovation into operations.	Civic innovators should be incentivized to spend time using data to help the City solve problems. Publicity, appreciation, and change creation are powerful incentives. The public is more likely to use Cambridge open data if high value projects can be posted on the portal's project showcase.	Develop clear processes and review criteria for partnership and collaboration with data users for inviting community members to use public city data to solve pressing community issues.	In Progress 2020
4.4 Identify and pursue opportunities to reach new groups and demographics.	Cambridge is home to residents from every demographic and walk of life. As we try to engage residents in our mission of data-driven problem solving, we should seek a diversity of perspectives and life experience.	Collaborate with other city staff and stakeholders to identify new engagement opportunities.	Planned 2021

4.5 Reach out to new groups at local universities	Students and researchers at local universities represent an important open data user group. The Open Data Program has already sought to engage these groups, particularly those in computer science departments. Future efforts will be expanded to court potential data users in urban planning, public health, and public policy departments.	Begin marketing Cambridge open data to students and researchers in planning, public health, and policy departments at nearby universities.	Planned 2021
4.6 Add topical data categories to the portal homepage	The open data portal currently organizes data by department. However, the public could more easily find interesting data if we also organized datasets by popular topics (e.g., business, sustainability, transportation).	Reorganize data portal according to topical areas.	Planned 2020
4.7 Align dataset keyword tags with City goals and priorities	Datasets on the portal should relate to City goals. Metadata fields such as keyword tags should link datasets to City goals whenever possible.	Develop City goal keywords and assign them to all new and existing open datasets.	Planned 2020
4.8 Create dynamic “data hubs” for popular topics and initiatives (e.g., Vision Zero, Net Zero, Glocal Challenge)	Many of Cambridge's large scale initiatives—Vision Zero, Glocal Challenge, Net Zero, etc.—are associated with numerous open datasets. The portal should organize these datasets into easily accessible data hubs. Ideally, department staff who request a data hub should contribute new problem statements to the CICI.	Design a space on the homepage for temporary data hubs, and create a process for staff to request a new data hub.	Planned 2020

5. BUILD DEPARTMENTAL CAPACITY FOR DATA-DRIVEN PROBLEM SOLVING

Over the next decade, it will become increasingly important for every Cambridge department to be able to leverage data. IT's Data Analytics and Open Data Program could create lasting change across the municipal government by helping other departments learn new data skills, acquire new tools, and train great analysts. The following objectives represent a plan for strengthening our emphasis on training, collaboration, and capacity building in the coming years.

Objective	**Purpose and Description**	**Action and Deliverable**	**Status & Estimated Completion**
5.1 Create standing data analysis/visualization trainings for city staff.	Data visualization is one of the most efficient ways to use data to influence decision-making. Staff in every department around the City should be able to use modern software tools to create informative data visualizations and dashboards.	Finalize a schedule for standing data visualization trainings.	In Progress 2021
5.2 Create venues for data analysts and data governance staff to share best practices and peer review each other's work	Many departments around the City have only one data analyst. Solo analysis work is not conducive to best practice sharing or peer review of analyses. City data analysts and data governance stakeholders should meet regularly to collaborate, share best practices, and develop new standards.	Convene data analyst and data governance working groups.	In Progress 2020
5.3 Implement a data academy for Cambridge staff	A flexible but coherent set of course offerings and online course recommendations covering all aspects of basic data analysis should be Cambridge's gold standard for training data analysts. Moreover, only a few other employers offer such a data training academy, so such an offering could be a powerful employee retention tool.	Develop and pilot a Cambridge data academy.	Planned 2022
5.4 Develop a regular schedule for data office hours	Data office hours are an informal way for department staff to bring data projects and issues to the attention of Data Analytics and Open Data Program staff. They should be offered on a regular schedule.	Schedule and conduct one year of data office hours.	In Progress 2020
5.5 Draft resources for departments that want to train or recruit their first data analyst.	One of the best ways the Data Analytics and Open Data Program can build analytical capacity across the City is to help departments invest in analyst staff. The Data Analytics and Open Data Program should offer support to departments who want to build their own staff capacity for data analysis work.	Draft a "So You Need a Data Analyst" guide for department managers.	Planned 2021

5.6 Develop procurement advice for departments seeking machine learning software products.	Lots of software products these days advertise with buzzwords like "big data" and "machine learning." There can be real value in these tools, but staff need to understand objective methods for evaluating potential new products. Additionally, staff need to be able to anticipate and address data ownership questions when procuring new analytics software.	Work with purchasing department to outline data analytics procurement guidance and address data ownership issues.	Planned 2022
5.7 Begin publishing regular internal news items about data analytics and open data.	Learning to recognize opportunities for data-driven problem solving is a skill in itself. The Data Analytics and Open Data Program can help build this skill among subject matter experts by educating them about how their counterparts in other departments and cities are using data to deliver better services.	Identify internal newsletters such as PIO updates that might be good venues for data analytics news items, and begin regularly supplying the newsletter editors with articles about data activities.	Planned 2021
5.8 Create "Data Resources for Staff" section of the open data website or intranet.	The creation of a clearinghouse of resources related to data analytics will help staff more easily find the information or trainings they need. This clearinghouse could also help staff share data internally.	Create "Data Resources for Staff" section on the open data website.	Planned 2021
5.9 Work with partner departments to achieve What Works Cities certification.	What Works Cities (WWC) is a philanthropic organization aimed at helping cities become more data driven. Cambridge is currently pursuing a certification from WWC that will verify and communicate our commitment to data driven decision making. The Open Data Program is helping to lead this effort in conjunction with other departments.	Achieve WWC certification.	Planned 2022

6. PREPARE FOR THE SMART CITY AND BIG DATA

Smart city technologies present both a challenge and an opportunity for Cambridge. "Big data" is a term for datasets whose volume, variety, velocity, and veracity create processing difficulties. Despite their complexity, big datasets about topics like traffic patterns, energy usage, and air quality could help us make better decisions. Cambridge must begin planning today to integrate these new data sources into data-driven decision making processes. As smart city issues touch on all Cambridge departments, the following objectives will be pursued collaboratively with many other internal and external stakeholder groups. In some cases, the data program will take a leadership role and in others it will support other groups or departments. Throughout all smart city efforts, the data program will collaborate with its partners to navigate a quickly changing technology and policy landscape and to align smart city activities with other complex initiatives such as Cambridge's Surveillance Ordinance.

Objective	**Purpose and Description**	**Action and Deliverable**	**Status & Estimated Completion**
6.1 Develop options for internal data sharing of both streaming and static data	Data needlessly trapped within one department is as closed to the rest of the city staff as it is to residents. Open data shouldn't be a binary concept: Just because a dataset is too sensitive to be posted online doesn't automatically mean it should be viewable by only a single municipal employee. We need new technologies and business practices for ensuring that data from departments across the organization can be internally shared, combined, analyzed, and when appropriate, opened to the public.	Work with department-level analysts to identify specs for an internal data-sharing portal.	In Progress 2022
6.2 Develop tools for staff to more easily use APIs and streaming data	APIs are a rich source of real-time data and will become increasingly useful and available to staff with the advent of the smart city. However, analysts usually need specialized skills to access them. Cambridge should develop technical tools and business processes for giving staff better access to APIs.	Develop a Shiny system for turning open APIs into CSV data files, and provide staff with instructions for using the data analytics helpline to get API support.	Planned 2022
6.3 Assist in drafting a Cambridge smart city data playbook	Data analysis, governance, and privacy will change dramatically with the advent of the smart city. Decisions will need to be made quickly and often about whether and how to retain and aggregate data from a variety of new sensors. A living document should be begun to enshrine new policies and best practices about smart city data.	Work with departments and external stakeholders begin a new living policy document to enshrine new best practices for smart city data governance.	Planned 2022

6.4 Support the convening of a smart cities working group	Big data is only one smart city topic. Others include digital infrastructure, cybersecurity, and the internet of things. Data analytics should be part of a larger working group with a more global perspective on these issues.	Attend first smart cities working group meeting.	Planned 2021