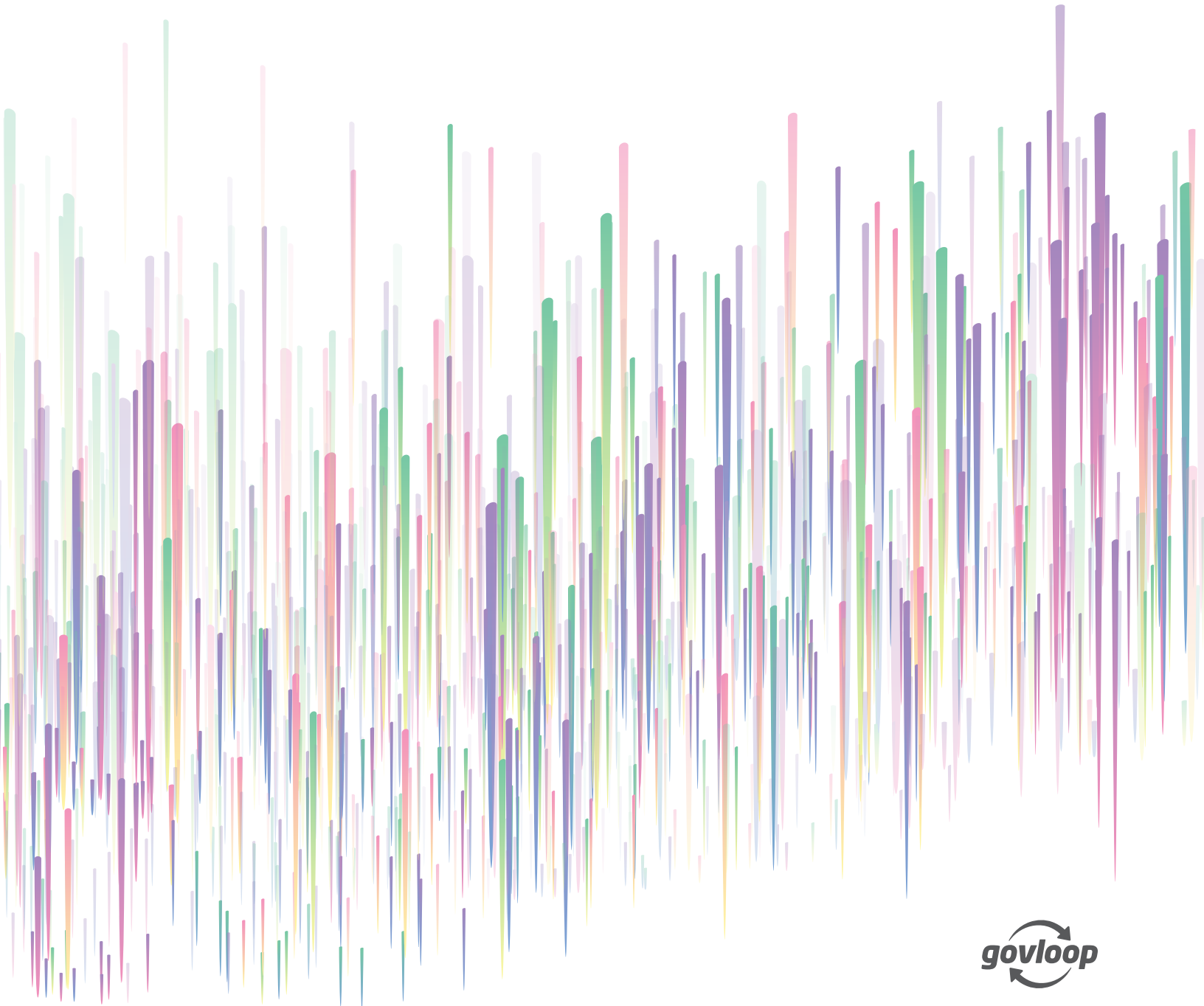


# How You Can Effectively Use Data Analytics



# CONTENTS

---

Executive Summary 3

**Play 1: Discovering Your Data** 4

**Play 2: Selling Your Data to the Boss** 6

3 Types of Visualizations That Can Bring  
Your Analytics to Life 8

Zeroing in on Data Analytics and  
Modernization 11

**Play 3: Implementing Your Project** 12

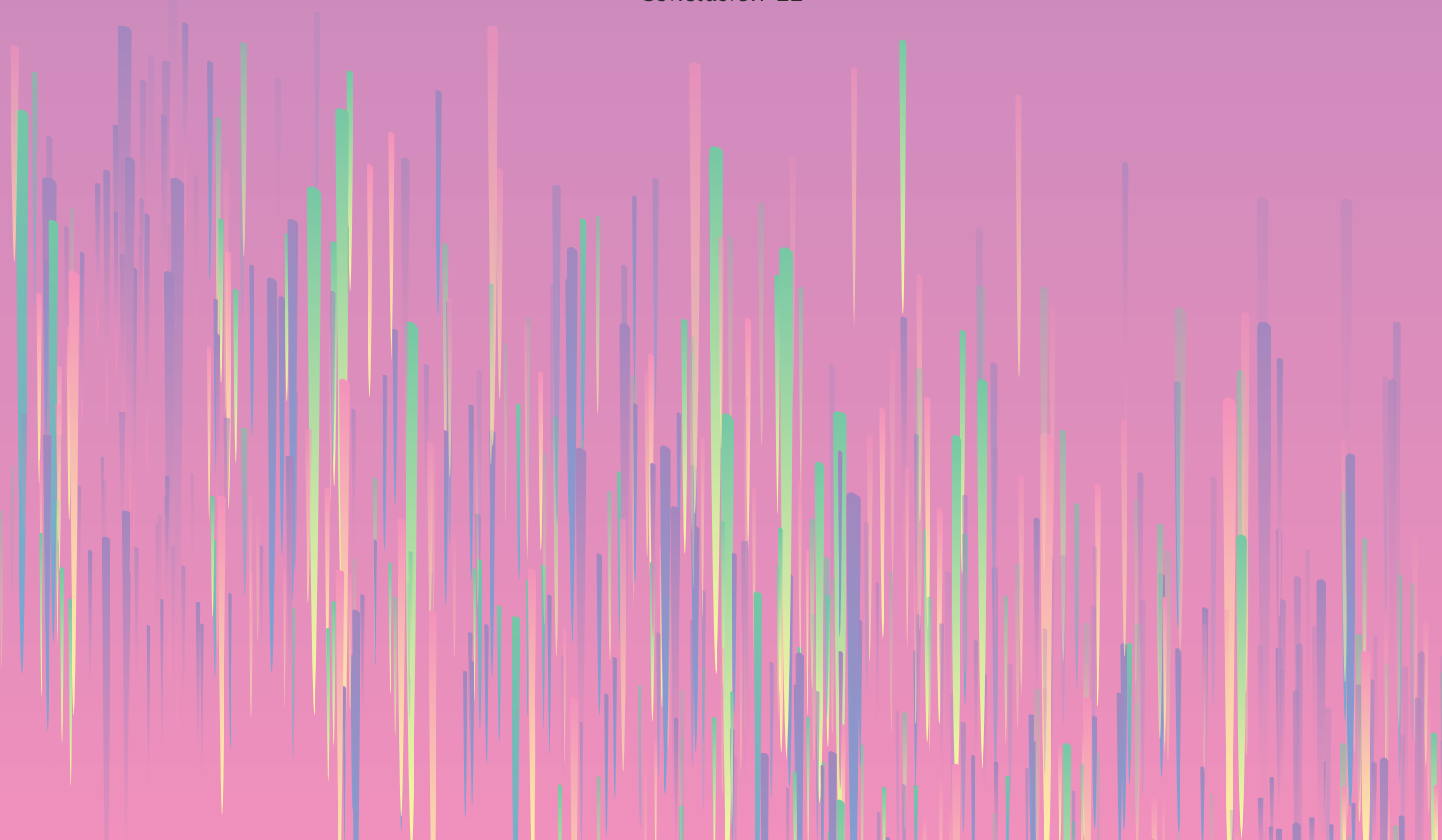
**Play 4: Translating Your Data into  
Mission Outcomes** 14

Securing Data in a More Holistic Way 17

San Francisco Tackles Barriers to Data  
Science in State & Local Government 18

**Play 5: Showing & Telling Your Data** 20

Conclusion 22



# Executive Summary

So much data, so little time. We can all relate to the feeling of information overload, with data coming at us from all directions but not having the time, resources or know-how to make the most of it all. That's why data analytics is crucial for day-to-day operations and intelligent decision-making for a growing number of government agencies. Using analytics, organizations can understand all the information out there in context.

Data analytics allows agencies to glean important insights from massive amounts of data in real time and apply them to detecting fraud, responding to cyberthreats or designing better services and products for a particular audience. Additionally, policymakers are increasingly using analytics for evidence-based policymaking, which lets them better determine the most cost-effective programs to achieve desired outcomes.

For newcomers, however, actually implementing data analytics can be arduous and consuming. According to a recent poll of the GovLoop community, the top challenges are getting started on an analytics project (28 percent) and understanding the data (20 percent).

Getting started on the project - 28%



Understanding the data - 20%



Selling it to the boss - 19%



Sharing the data (visualization & metrics) - 18%



Relating the data back to our mission - 15%



To help, this GovLoop playbook breaks down the process of introducing data analytics to your organization. The guide focuses on five plays with step-by-step suggestions to make the most of your agency's data. Plus, you'll get firsthand insights and tips from data scientists in and around government, in addition to cheat sheets to help you implement these plays at your organization.

## How to Use This Playbook

You can use these plays in order, mix them up or apply them to the challenges giving you the biggest headaches. Each play is broken down into three practical steps and accompanied by a cheat sheet to help you implement those steps at your agency.

**Play 1: Discovering your data.** Decide what problem you're trying to solve and understand the different uses of data.

**Play 2: Selling your data to the boss.** Decide which metrics to use, how to frame the data and how to communicate all that to get leaders' and stakeholders' support.

**Play 3: Implementing your data analytics project.** Choose your tools and technologies, and evaluate your agency's data.

**Play 4: Translating the data into mission outcomes.** Connect your project to the mission and deliver insight-driven outcomes.

**Play 5: Showing and telling your data.** Decide which data to highlight and how to market it internally and externally using social media and other tools.

## PLAY #1

# Discovering Your Data

Discovering your data will be the first key step in identifying the data you need and what you want to accomplish with it. Ascertain the questions you need to answer, the type of analytics you'll use and how you'll use those questions to guide the project.

---

### STEP 1:

Explore the different uses of data.

Pinpointing [successful use cases](#) of data analytics at other agencies may inspire you to see how data can help meet your agency's mission needs.

There are [four main types](#) of data analytics to choose from. [Gartner](#) identifies them as:

- **Descriptive analytics**, which determines what happened, such as the number of visitors to a website;
- **Diagnostic analytics**, which determines why something happened, such as the cause of a drop in web traffic in the past month;
- **Predictive analytics**, which determines what will happen through algorithms; and
- **Prescriptive analytics**, which determines how your agency can make it happen, such as rerouting travel based on traffic predictions.

### STEP 2:

Understand what problem you're trying to solve.

Start by deciding the most important mission priorities your agency needs to address. Whether it's something as large as tracking the opioid crisis or smaller, such as updating your agency's website, determine what questions you need to ask and what problem you can feasibly address.

By knowing what problems you need to tackle, you can better determine the type of analytics to use: predictive, prescriptive, descriptive or diagnostic. Then, you can identify your units to measure, such as transaction, rate of occurrence of an event or names within constituent bases. Be sure to clearly define the data items, the units used and the meaning of each data point.

### STEP 3:

Decide how data can help you achieve your mission needs.

Finalize the discovery phase of your data by tying it all to your agency's mission. Ultimately, you want this analytics project to help reduce costs, increase efficiencies, improve services or do all of the above.

McKinsey and Co. [advises](#) making data analytics work for you by keeping the data purpose-driven, or focused on your organization's mission. Ask the right questions: How can we reduce costs? How can we increase revenues? Start with a smaller project that develops over time, so the possibilities won't overwhelm you as you discover data.



# DISCOVERING YOUR DATA CHEAT SHEET

Tyrone Grandison, a data scientist and former Deputy Chief Data Officer (CDO) at the Commerce Department, shared his [top steps](#) to help you discover your data while fostering a data-driven culture at your agency.

## Do you know where you're going?

Know the questions you are trying to answer with data. Do you want to make a process more efficient? Do you want to decrease the time it takes to complete a transaction? Do you want to increase the number of citizens you can serve? Know what you want to get answered from the data, then figure out the specific data items that need to be collected and stored.

### What question are you trying to answer?

---

---

---

---

---

### What type of analytics will you use? Descriptive, diagnostic, predictive or prescriptive?

---

---

---

---

---

### What data items will you need (i.e., the number of transactions with your agency website or demographic data)?

---

---

---

---

---

## Do you know your audience?

Once you know how you will interrogate the data, you need to understand the audience that will see this data. This means deciding who's going to have the final say on your data project. Will the decision-maker be yourself, a data scientist, a citizen? Knowing your intended audience forces you to start thinking about the actions you want people to take when you present the data to them. When they see the visualized data, should they reduce the steps of an existing organizational process? Should they increase the number of staff members on a particular task? Should they start monitoring a specific business area more closely?

### Who will be your key decision-makers? List them here.

---

---

---

---

---

### Who is your intended audience? Include key decision-makers and public audiences.

---

---

---

---

---

## Do you know what processes you're using for desired outcomes?

Now it is time to design and implement your ETL (Extract-Transform-Load) pipeline. This is where you create the process and supporting mechanisms (technical or otherwise) that allow you to get data from the desired data source, cleanse and massage it into the right form with the right semantics, and then store it in the data management system of your choosing. Start small. Pilot your ETL with your simplest use case. When you have it working, expand the scope of scenarios that your ETL pipeline can handle until it covers all your needs.

### Building an ETL pipeline means:

- Identifying your analysis code (label, combine and organize your data using software, manual processes or a combination of both);
- Identifying data sources (making sure you can easily see where your data is coming from, whether it's customer service feedback forms or directly from your agency's website); and
- Ensuring algorithmic randomness (making sure your data is being randomly sub-sampled for high-quality data).

## PLAY #2

# Selling Your Data to the Boss

You may be confident in your data's potential, but relaying the importance of your data project to executive leaders and other stakeholders – both internal and external – requires a clear and concise presentation of a data story. Here are a few steps to successfully sell your data project.

### STEP 1:

Decide which metrics to use and how to frame your data.

Use a mission statement to convey the story of your data. Your mission statement should include a hero, or affected members of your organization; the conflict, or reason for your proposed change; support, or data-driven evidence that makes a case for the change; and an epilogue, or proposed result.

In attempting to frame your data and choose the right metrics, it may help to work backward. Take note of the problem areas within your organization and draft a clear and achievable goal for improvement. From there, you can assign metrics by deciding what measurements will help the organization reach the goal.

### STEP 2:

Engage a support team to get more stakeholders on board.

Assembling a support group for your data project allows you to divide important tasks among your team. The more people you have on your team, the more you need to maintain consistent communication with numerous stakeholders. Create a calendar that team members can refer to so they know which stakeholders to reach out to and when.

If you are engaging with stakeholders who do not support some aspects of your project, don't be quick to shun their criticism. Having different perspectives within your team will help you create engaging counter arguments. When your boss sees that stakeholders are invested in your project, she will be more likely to support it.

### STEP 3:

Communicate the data to your executive leaders.

Now you must prove to executive leaders that your project will contribute to your organization's mission. First, acknowledge your organization's problem areas to create context and present the concrete, actionable ways your project can address these issues. To convince them even further, showcase the potential consequences of inaction.

When presenting data, remember that visual communication is effective. It's helpful to know your audience, too. You may need to present differently to your CEO than to outside stakeholders. By understanding your audience's familiarity with analytics, you can also choose to exclude or reveal the nitty-gritty details of your analysis.

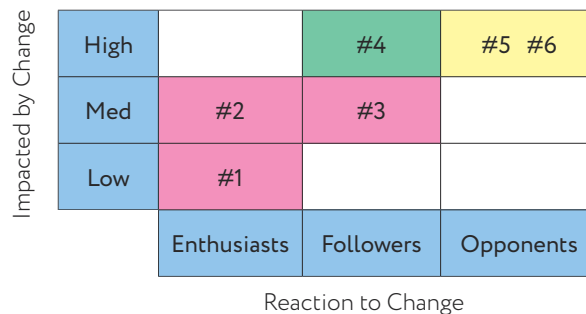
# SELLING YOUR DATA TO THE BOSS CHEAT SHEET

The Lean Methods Group created the [Stakeholder Map](#) that can be used to identify who internal and external stakeholders are and what their issues may be to better sell your analytics project.

1. Draw the Stakeholder Map.
2. List every key stakeholder in the center column of the stakeholder identification table.
3. Indicate the level of influence each stakeholder has over the change and/or other stakeholders in the right column of the table.
4. Meet with each stakeholder to discuss the data project, including its background and objective.
5. Solicit feedback from each stakeholder regarding perceptions and reactions to the project plan.
6. Reach a consensus among team members on each stakeholder's ratings for "impacted by change" and "reaction to change."
7. Indicate the "ID Code" when assigning each stakeholder to the appropriate block within the grid so you can assign a number to each stakeholder easily.
8. Use red to color block where stakeholders have a high influence, yellow for medium influence and green for low influence.
9. Add a footnote to the bottom of the diagram that shows the data it created and the name of one or more participants who can be contacted if additional information is needed about the process used to create the Stakeholder Map.

## SAMPLE STAKEHOLDER MAP

ID code	Stakeholder Description	Influence
#1	Bobby Jones, VP, Finance	High
#2	Martha Strudle, Comptroller	High
#3	Pierre de Vincent, Accounts Payable Supervisor	High
#4	Accounts Payable Workgroup Members	Low
#5	Contractors	Med
#6	Vendors	Med



## YOUR STAKEHOLDER MAP

ID code	Stakeholder Description	Influence

# FEDERAL SPOTLIGHT

## 3 Types of Data Visualizations That Can Bring Your Analytics to Life

The federal government spends almost \$500 billion a year on its contracts and nearly \$4 trillion overall – numbers so large that they’re almost abstract. It’s obviously a lot of money, but for citizens who are interested in what the government is spending, it can be hard to wrap their heads around those figures or what they might be going toward.

It was with those curious citizens in mind that the Treasury Department’s Bureau of the Fiscal Service launched the beta version of its new data visualization and analytics site, [DataLab.USASpending.gov](https://DataLab.USASpending.gov).

Taking into account users’ needs and the requirements of the Digital Accountability and Transparency (DATA) Act, which requires federal agencies to report financial information every quarter in a standardized way, the team behind the DataLab decided to present the wealth of federal spending data in a visually friendly and easy-to-manipulate way.

Justin Marsico, a Senior Policy Analyst at Treasury who leads the analytics team working to meet DATA Act requirements, explained the work and thought that went into [DataLab.USASpending.gov](https://DataLab.USASpending.gov) at a recent FCW citizen engagement event.

“Our mission was to discover insights within federal spending and encourage others to do the same,” he explained. “We wanted to explore visualization and tools the public could use and lead by example within the government.”

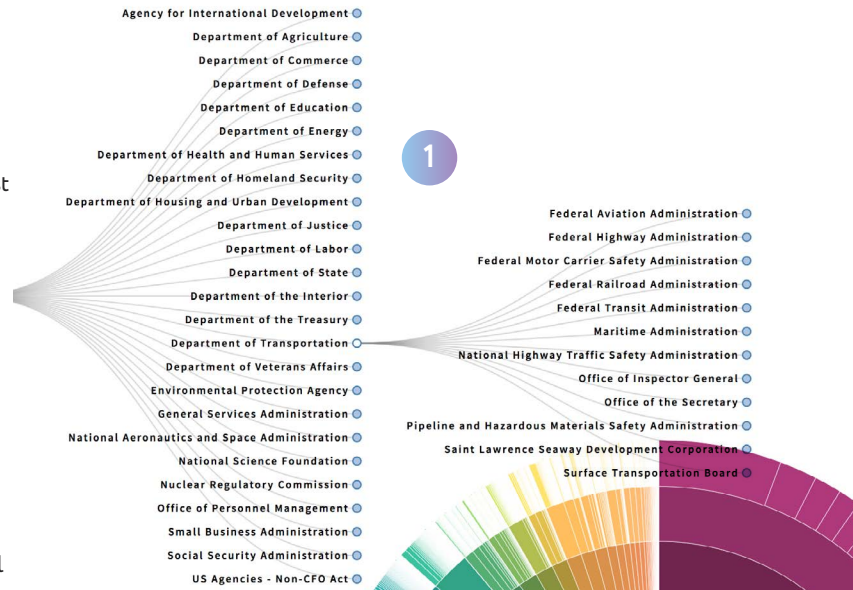
Doing so was a gargantuan effort that took multiple years, 91 agencies submitting data, and a team to test, work with end users, and eventually launch the site.

The result is a sleek, colorful site full of a variety of data visualizations that allow users to delve into data analytics about everything from which contractors each agency spends the most with to federal funding for specific community programs nationwide.

Marsico walked the audience through three types of data visualizations that the site used to engage citizens about federal spending data.

## #1: Federal Accounts Dendrogram

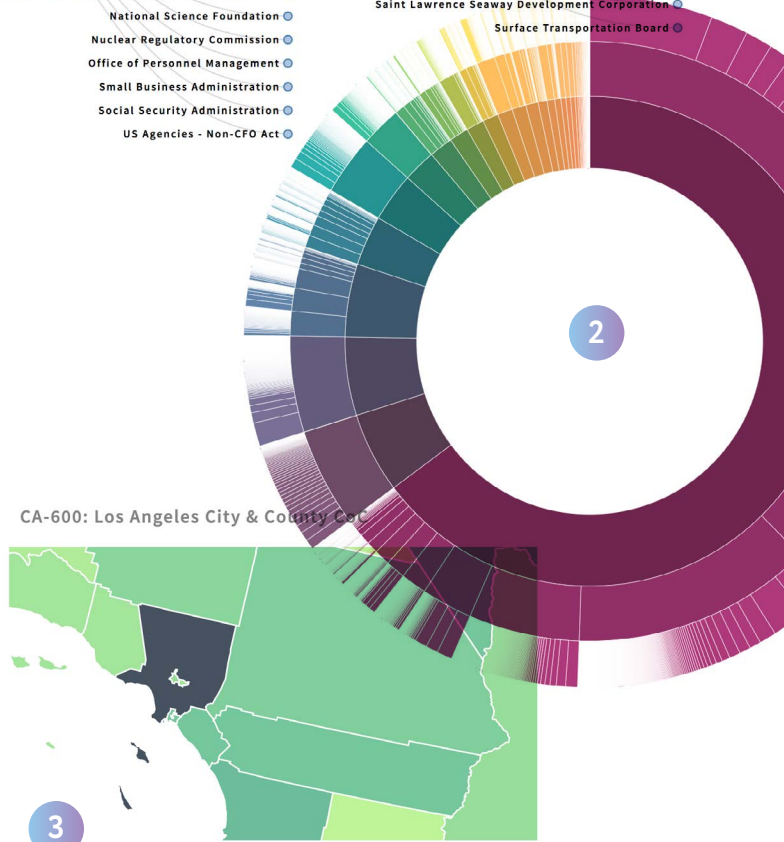
A dendrogram is a tree diagram that shows taxonomic relationships. The [DataLab's dendrogram](#) details an all-in-one view of federal accounts that the government uses to spend money and how each agency's accounts are organized. The first branch of the diagram shows agencies, the second shows sub-agencies and the third shows the federal accounts.



## #2: Federal Contracts Sunburst Data Visualization

A sunburst diagram or chart is a multilevel pie chart used to represent the proportion of different values found at each level in a hierarchy. The DataLab turned to a [sunburst diagram](#) to categorize and explore contract-related information for fiscal year 2017, including which organizations received federal contracts, amounts, awarding agencies, sub-agencies and types of goods or services contract recipients provided to the government.

The site explains how to use this sunburst visualization: "The diagram has three rings. The inner ring represents federal agencies, like the Defense Department, which are sized by the total dollar amounts they spent on contracts in FY 17. The middle ring represents sub-agencies, like the Air Force, which in most cases actually award contracts to contractors. The outer ring represents the contractors who receive awards from those sub-agencies."

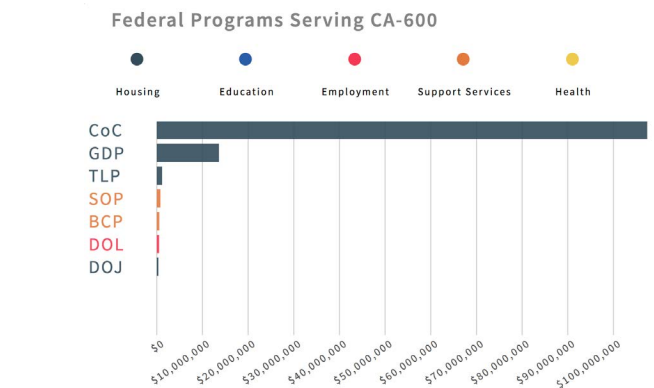


## #3: Cluster Analysis

A cluster analysis is a model that assigns observations to groups so that observations within each group are similar to one another with respect to variables or attributes of interest. The DataLab chose to analyze data around homelessness funding nationwide. The hope is that centralizing this sort of data could help cities and counties learn from one another about homelessness funding and programs. "We chose to conduct this analysis because it features data from across multiple federal agencies, which is all available in one dataset," the site noted.

DataLab.USASpending.gov has a variety of other datasets and visualizations, and Marsico said the next step is to sign up user groups for further testing and feedback before a formal launch.

"We really want to let the user take their own self-directed journey through our data to get them excited and educated about what government does," Marsico concluded.





# Seven transformation leaders. United.

We're committed to providing you with the intelligent technologies, services and partnership you need to set the pace in the digital business era.

**DELL**Technologies



DElLEMtC

Pivotal

RSA

Secureworks

virtustream

vmware

## INDUSTRY SPOTLIGHT

# Zeroing in on Data Analytics and Modernization

An Interview with Cameron Chehreh, Chief Technology Officer,  
Dell EMC's Federal Business

Agencies recognize that modern analytics has become an imperative. But in order to take advantage of these capabilities, they need access to computing platforms that support real-time decision-making as well as their mission priorities.

To fully leverage data analytics, agencies should modernize their data centers, to gain an edge in computing and securing data. At the same time, the workforce should have the ability to perform mission-critical duties from any device or location.

As a first step, agency leaders should incorporate automation and next-generation infrastructure into their data centers. Many agencies are streamlining compute, storage, networking and server virtualization into a converged or hyper-converged infrastructure. This holistic approach provides numerous advantages, particularly when it's part of an overall next-generation model. Agencies will be able to decrease their IT footprint, become more energy-efficient and deploy modern, cloud-based applications.

A modernized infrastructure also provides agencies increased flexibility in the data center, to decide whether and how to scale up or scale out. With the tedious, yet critical, tasks run by automation, the workforce is now more available and efficient, and has time to innovate.

By using modernized technologies, agencies can effectively combine hardware and software to push compute and storage resources out to the tactical edge thus bringing the data center closer to frontline users. For example, if a Federal Emergency Management Agency employee is responding to a disaster and needs to run applications but cannot get network

access or soldiers in the field require higher compute power than is available. Modern data centers allow agencies to quickly scale and allocate computing and networking resources as the mission requires. Agencies are also beginning to understand that cloud is not an end destination, but rather an operating model. As they seek to adopt more sophisticated analytics, agencies must work to properly secure data in the cloud while complying with IT policies in place.

Modernized infrastructure is also critical to protecting data. Robust security features can be built in from the silicon to the firmware, all the way up to the system level, such as enforcing hardware-based, root-of-trust for firmware upgrades. This includes encryption of data at rest, and the frequent rotation of user credentials, so they are only useful for short periods of time.

Data is a precious commodity because it powers decision-making, yet many agencies are struggling to develop comprehensive data protection strategies. Fortunately, through the application of software, they can create an envelope of security around data wherever it resides and wherever it might go.

Personal technology and an increasingly mobile workforce are also driving the demand for better and more comprehensive data security. Sixty percent of people work outside the nine-to-five schedule, and nearly 66 percent do some work from home. With 95 percent of security breaches initiating at the endpoint, that's a lot of opportunity for a problem.

Therefore, new security capabilities should utilize data analytics to look beyond signature-based issues at the endpoint and instead search for anomalies. To identify

a malicious link or file, for example, new security capabilities can evaluate patterns in how users work.

Furthermore, government's efforts to modernize IT and security align with a transformation of the workforce. Millennials aren't the only generation who want the latest devices to help them be productive wherever and whenever they want. Not only are technologies improving mobility (by helping employees access information and do work outside of their organizations) but they're also encouraging cultures to be more service-oriented.

Every agency is at a different stage in its IT modernization journey, but few have fully transformed. We've seen how technology can impact and transform entire industries – think of how ride-sharing changed the taxi cab industry, for example. Equipped with the right technologies, government agencies can undergo similar changes in the way they deliver citizen services and meet their mission. Modern infrastructure allows agencies to adopt more sophisticated analytics, solve their mission-critical problems, better secure their data and transform the public workforce for the better.

## PLAY #3

# Implementing Your Project

You've taken the time to discover your agency's data. You've discussed it with your teams and executives, and you have the buy-in you need. Now it's time to get to the fun part: getting started. Decide which technologies you will use, start evaluating the data and then implement the project.

### STEP 1:

Define your technical requirements.

The [technical requirements](#) phase involves examining the data available for your project and determining its quality. You also need to inventory the tools your agency currently uses and outline your agency's current architecture (i.e., hybrid cloud, on-premises data center or virtual data systems).

Consider the data sources you'll need to choose from, such as transactional data, survey data or web logins. If your agency uses purchased data, consider if it's supplemental data, such as demographics. If not, analytics from social media or news streams may complement your current data to create additional project value.

### STEP 2:

Choose your tools and technologies.

Decide what parts of the project you'll need to do manually with your team and what parts you can automate. For automation technologies, there are a variety of [tools](#), including visualization software, that leverage artificial intelligence, application programming interfaces and business intelligence to help you produce tables, charts and other visualizations. You may also consider statistical tools such as SPSS or SAS.

### STEP 3:

Implement your analytics project.

If you're in the federal government, the General Services Administration hosts a [Digital Analytics Program \(DAP\)](#) through its Technology Transformation Services organization that can help you get started. DAP JavaScript code is required for all federal public-facing websites and consists of web analytics tools, [trainings on YouTube](#), [implementation support and instructions](#), and ongoing desk help.

In addition to referring to DAP, state and local governments can define the [service change](#) they want to address. For example, this may entail upgrading a program your agency offers to citizens or improving employee productivity through a new technology. Defining a service change is preferable to producing a simple whitepaper or report and will strengthen your data project.



# IMPLEMENTING YOUR PROJECT **CHEAT SHEET**

The city of New Orleans' Office of Performance and Accountability houses an analytics team that partners with city departments to help them identify and deliver data projects. [NOLAytics](#) provides an easy-to-follow, comprehensive [checklist](#) that you can use to implement your analytics project.

<b>CRITERIA</b>	<b>QUESTION</b>	<b>RESPONSE</b>
Sponsorship (who)	Is there executive level sponsorship? Has a departmental contact been assigned? Has an office or department in your agency been assigned?	YES / NO YES / NO YES / NO
Definition of problem (what)	Has a problem statement been defined? What is the deliverable for this project? How will the deliverable be used - i.e., who will do what differently?	YES / NO
Impact (why)	Does the project align with Mayoral priorities? Potential operational impact? Policy outcome to residents? Measurability of impact?	YES / NO HIGH / MED / LOW HIGH / MED / LOW YES / NO
Feasibility (how/when)	Is data available for this project? Can we tap into existing processes? Is there a political timing consideration? Is there capacity to implement?	YES / NO YES / NO YES / NO YES / NO
Spillover	Is this project replicable/portable? Will this project augment the data infrastructure of the city? Is the project sustainable?	YES / NO YES / NO YES / NO
Data Readiness	Does data already exist? Is the data available on your agency's website? Is the data in a central repository? Is the data machine-readable, constantly updated and maintained?	YES / NO YES / NO YES / NO YES / NO

## PLAY #4

# Translating Your Data into Mission Outcomes

Analyzing the data, selling it to your boss and implementing the project may seem like only half the battle. You now have to account for the numbers and ensure that the project is on track to deliver the outcomes your agency needs.

### STEP 1:

Focus on the connection between the data and your agency's mission.

Deloitte defines mission analytics as a way to align data insights with agency objectives. You have to make your mission measurable by defining it in ways that make it quantifiable. Break down the list of potential measures into inputs, outputs and outcomes. Inputs are factors such as funding and resources, while outputs are products of government activity itself, such as changes to programs. Outcomes are the consequences of direct relevance to citizens such as citizen services. You can also refer to Play 1 to make sure you're aligning the project to mission outcomes early on.

### STEP 2:

Institutionalize your data project.

Once your data program or service change is in place, the next phase may be institutionalizing it. This is when resources are allocated to the data project or program and roles and responsibilities are standardized.

Using evidence-based decision-making can also help you reduce wasteful spending, expand innovative programs and strengthen accountability. Make sure you're tracking costs in terms of what your agency was spending or losing in revenue before the project. Don't forget to also keep track of costs that your project is incurring.

### STEP 3:

Deliver insight-driven outcomes.

Acquire the right tools to pull meaning out of the data and glean new business insights. Although tools can help, it's equally important to ask the right questions:

- What is the current organizational performance?
- What does current performance mean for the mission?
- What if we applied resources or solutions differently, and what effect would that have on the mission?

These questions help connect operational data to mission outcomes and separate the mission analytics framework from more generic business intelligence tools.

# TRANSLATING YOUR DATA INTO MISSION OUTCOMES **CHEAT SHEET**

Deloitte has a [visual map](#) that can help you use mission analytics to bridge appropriate business decisions and actions to your agency's mission.

## WHAT?

Describe your agency's current performance. Identify what's going well, what could be improved.



## SO WHAT?

Identify what your agency is currently spending in dollar amounts. What's your agency's current performance? What are the goals for return on investment? Identify dollar amounts for how you want to decrease spending or increase revenue.



## WHAT IF?

Put together what you identified in the previous two categories.  
What would your desired performance look like with your desired return on investment?



# Enabling the Public Sector with Hortonworks

A Modern Data Architecture  
for the Government



100% Open Source



Improve Efficiency



Reduce Costs



Secure



[hortonworks.com](http://hortonworks.com)



## INDUSTRY SPOTLIGHT

# Securing Data in a More Holistic Way

An Interview with Henry Sowell, Chief Information Security Officer,  
Federal Sector, Hortonworks

As the volume of data increases every day, government will need new and innovative ways to identify, prioritize and, most importantly, secure their data. This includes managing access to government systems and databases and ensuring data isn't exposed to malicious threats.

Part of the challenge for agencies is they've traditionally relied on older, rule-based policies for network access control. But these policies lead to fragmented security approaches rather than a holistic one for securing data.

To better secure their data and stay ahead of cyberthreats, agencies need a comprehensive platform that combines behavior analytics and cybersecurity. By comprehensive, this means platforms should provide a centralized view of all relevant data, real-time processing to stay ahead of threats and cost-effective data storage for improved long-term visibility. In an interview with GovLoop, Henry Sowell, Chief Information Security Officer of the Federal Sector at Hortonworks, discussed how agencies can combine their cybersecurity and analytics needs into a comprehensive platform to better secure their data.

Getting a well-rounded analytics platform is only the first step to assessing what data agencies own. One concern is that by bringing their data together, agencies create a bigger target for threat actors to easily access their most sensitive asset: data. "When you start collating your data together into one spot, you have a larger set of risks because you're making a very desirable target for malicious actors," Sowell said. "And if you're not approaching security in a consistent manner, you're risking these large targets."

So how does a comprehensive platform deliver on improved analytics and security?

Sowell explained that it boils down to interoperability, consistent security implementations and holistically viewing data.

A comprehensive platform combines the power of behavior analytics and cybersecurity. With Hortonworks' cybersecurity platform, for example, agencies can take advantage of interoperability by using an enterprise open source solution. "Partners in the open source community provide flexibility and the ability to improve interoperability," Sowell said. This means agencies have access to insights that are based on long-term, historical data collection from a community of users. They can then share and receive ongoing feedback and insights from this community of open source users over time.

By taking a data science approach to cybersecurity, agencies can practice consistent security implementations. This approach to security involves the use of behavior analytics and machine learning to identify anomalous activity and combat threats as they are escalating. Alerts and visualization dashboards help provide actionable intelligence to stay one step ahead of bad actors and identify malicious intent.

Such threat detection is based on dynamic rules and constant machine learning. "By combining different datasets like threat data with machine learning, network events that may appear innocuous when examined individually, become detected threat events, enabling investigation and prevention of incidents occurring on their network," Sowell said. "Additionally, machine learning on a capable platform can enable this detection and prevent threats as they're occurring in real-time."

Lastly, a comprehensive cybersecurity platform consolidates threat streams into a single, integrated view and toolset. "Hortonworks' cybersecurity platform, for example, allows agencies to collect all of their network traffic together from many different threat sources, quickly ingesting a variety of data sources and creating a holistic view of the threats within their networks," Sowell said.

To ensure that agencies fully harness the benefits of a comprehensive analytics and cybersecurity platform, Sowell recommended keeping abreast of current best practices for cyber posture. First, agencies should ensure their policies are updated and consistent with their security needs.

"Develop appropriate policies that allow for continued flexibility while maintaining a secure posture as you enter into the next phase within your organization," Sowell said.

Additionally, it's important to adjust capabilities as cyberthreats continue to evolve. "You have to be flexible," Sowell said. "Data sources and threats are continually changing. Your tools have to be able to adapt with that. So you need a platform that allows you to remain flexible, adapt to different data sources and develop new approaches to achieve agency missions."

For government, the volume of data and cyberthreats they face will only continue to expand. As attacks grow increasingly sophisticated and multifaceted, existing security tools only provide one facet of cybersecurity information. But by leveraging data from multiple systems in one comprehensive platform, agencies can assess the true nature of a single threat across enterprises and, ultimately, better secure their data.

## STATE & LOCAL SPOTLIGHT

# San Francisco Tackles Barriers to Data Science in State and Local Government

Accumulation of data is common for any organization — especially those in state and local governments — and can be difficult to translate into a data science project that is well-defined and properly implemented.

Through its [DataSF program](#), the city of San Francisco strives to break down barriers to accessing important data by focusing on the structure, design and visual aspects of data programs. DataSF assists agencies within the city with using analytics to improve operations and offer better services. In an interview with GovLoop, Joy Bonaguro, the first Chief Data Officer (CDO) for the city and County of San Francisco, identified common challenges for data leaders in harnessing a data science program and best practices to combat those challenges.

When Bonaguro joined DataSF in 2014, the CDO role was poorly defined in the public and private sectors. To better refine the focus of San Francisco's data program, Bonaguro turned the emphasis to overcoming obstacles to data use. "Our data strategy has been to systematically attack barriers to data use," she said. "We didn't jump into a data science program, because if we did that, we would have been met with skepticism."

Instead, Bonaguro stressed that CDOs should understand their local context. "San Francisco has many analysts distributed across our departments," Bonaguro said. "We have 52 or more and many of those have strong policy, planning and analysis skills."

To fully harness those skills and understand the everyday barriers to data use, DataSF instituted an annual survey, beginning in 2014.

One barrier is differences in the level of core data skills among city analysts. To reduce these discrepancies, DataSF created a training program called [Data Academy](#). "We co-founded Data Academy with our Controller's Office to lift and level data literacy skills across the city," Bonaguro said. "We offer about 20 peer-taught classes that are free and that we regularly teach."

Bonaguro found a way to make these classes accessible for any analyst. "We don't charge for them because competency with use of data shouldn't be a function of your department's budget," she said. DataSF minimizes expenses through the participation of volunteer teachers, using consistent training materials, and light administration and management of the system. "We're saving about \$1.7 million a year in saved staff time due to improved skills," Bonaguro said.

"We have a gateway to additional work and projects," Bonaguro said. "It gives you an opportunity to work with and touch lots of different people across the organization. When they see you as someone who's creating value for them, they want to work with you on other things." As Data Academy gained traction, other jurisdictions started to create their own versions of the school and reached out to DataSF for help.

Another common obstacle to data is what Bonaguro called "the Sisyphean reporting cycle," or a barrier to an organization's capacity for more advanced data use given the opportunity cost.



“You’re constantly writing and authorizing reports, and when you have analysts exhausting their time on them, they’re not doing high-level analysis,” Bonaguro said. “An open data portal can help with some of that. With our self-service dashboarding strategy, we can empower analysts to automate the reporting requirements in a way that’s flexible and can address ad hoc data requests.”

Once analysts free up time by automating certain processes, they have more availability for advanced tasks like active performance management and high-level analytics.

To support improved metrics and performance management, Bonaguro’s team collaborated with the Controller’s Office to create tools and guidebooks about what it means to have useful and valuable metrics. “We put together tools and guidebooks into a resource collection on performance management and metrics,” Bonaguro said.

When considering whether to offer data science services, Bonaguro noticed that a common issue in the field was implementing changes based on data science insights. “A pattern that our team observed was that a lot of projects struggle to get implemented and completed,” she said. “Anyone can model and run cool stats on a large dataset. The hard part is meaningfully answering an operational question and then helping that group implement the data insight you provided.”

At DataSF, the emphasis is on user-centered design in the data science engagement. Before assisting clients, such as county

and city departments and agencies, DataSF takes the time to understand their current business processes to help fuel ideas about implementing change. “It’s not about an academic report. You’re trying to implement a service change,” Bonaguro said. “We develop a project charter and constantly talk about the service change.”

Bonaguro and her team ensure that through the project charter and other solicitation materials, clients know from day one that the DataSF team will work with them to implement a service change based on the data science insights.

Lastly, DataSF helps clients implement service and data changes in a way that is understandable and feasible. “While we may use advanced statistics and modeling in the project, we tailor the end product to meet the needs of the end users. Sometimes that’s a script or model and sometimes it’s a user-friendly workbook that [the clients] can use in their business process.”

Following the implementation process, DataSF hosts a showcase where they present the work they completed with their clients and client testimonials using easy visuals such as Microsoft PowerPoint presentations and online articles.

Overall, Bonaguro believes that creating a product that is understandable and usable is key. “Being able to translate your insights into something that’s a useful tool or a useful business process is really essential,” Bonaguro concluded.

## PLAY #5

# Showing & Telling Your Data

Now that you've done all the hard work, it's time to present your data in the most comprehensive and user-friendly way. Whether you're using bar charts, pie graphs or new visualization software, you need to be mindful of your audience – internal or external – and clearly show and tell your data.

### STEP 1:

Decide how you want to visualize the data.

Tables can help you show the numbers in a straightforward manner, while bar graphs can convey a greater depth of analysis. When it comes to relaying a purpose, the data cannot speak for itself. Collect your graphs and organize them into a presentation to showcase the full data story.

Remember that you're aiming to convey the results as clearly as possible to users and stakeholders. Try to keep it simple. The right visual display will depend on the amount of data collected and if users want to quickly understand the driving factors of your project.

### STEP 2:

Assemble a publicity campaign.

Determine the goal of your campaign effort. Then, map your goals to the marketing and campaign metrics. For example, if your objective is to sell a new service change, you may want to help citizens understand the change and get them excited about it.

Having a clear, well-defined objective for a campaign will help specify what types of analysis you can complete to determine the success of your campaign. Each public relations or marketing campaign should have goals with measurable outcomes that match the objectives of your agency.

### STEP 3:

Find creative ways to share results.

Whether internally or externally, continue sharing the results your data project and marketing campaign yield. Did citizens enjoy the service change? If so, how? How many more views did your agency's website receive?

A weekly newsletter can be a fun way to engage your organization in sharing the data metrics. GovLoop's marketing team, for example, shares a weekly email called "Metrics Monday" that highlights our top-trending email subject lines, top-performing blog posts and downloads for content marketing campaigns.



# SHOWING & TELLING YOUR DATA **CHEAT SHEET**

Dave Uejio, Chief Strategy Officer at the Consumer Financial Protection Bureau shared [practical tips with GovLoop](#) on how to communicate data to your audiences.

---

## 1. Storyboarding sets the stage for your talk.

Don't think about how wed you are to any one slide in your presentation. And don't set out to simply design nice slides. Be a ruthless editor and make each one count.

## 2. Tables provide a baseline comparison.

Beyond bar and column graphs, there are many ways to visualize data. Show your audience the numbers or give the impression you are showing the numbers through tables.

## 3. Pie charts have clear limits in conveying insights.

They are terrible for showing proportionality or comparing datasets. They only thing worse than one pie chart is two pie charts.

## 4. Column and bar graphs convey greater depth of analysis.

You should have clear axis labels on your graph, and they should be relevant to what you are trying to communicate. Include a title, don't have hanging decimals, and provide a source and legend for clarity.

## 5. Waterfalls can demonstrate proportionality and magnitude.

These are a combination of bar or chart graphs. They are useful for financial analysis, and they provide a clean way to tell consistent stories. Plus, they don't take up a ton of real estate.

## 6. Infographics can be framed to tell a story.

It isn't fair yet to assume everyone in an analyst role is a graphic designer. But there are plugins online that you can use to create infographics.

## 7. False proportionality creates abstract perspectives.

You have to balance ruthless storytelling with a baseline level of rigor and integrity.

## 8. Y even have an axis?

Bottom line: Don't screw around with axes and tell misleading stories with data. Proportionality matters, and it's on you to conduct analyses responsibly.

## 9. Draw on open source design communities to develop a unique palette.

Using color palettes gives your presentation a unified look and feel. Also, fonts set the tone for your assertions, so do what you can to differentiate your work. Whatever you do, don't let work go out of your shop that is not pristine.

## 10. Pictures directly evoke emotions.

Whether it's photo albums from your own agency or one of the many free picture sites out there, there are many neat visuals you can use.

# Conclusion

It may seem there's too much data in government to analyze and too little time to do it, but you don't have to let the numbers overwhelm you. Data analytics is feasible once you break it down and simply get started. By using these five plays: discovering your data, selling your data to your boss, implementing your project, translating your data to mission outcomes, and showing and telling your data, you can effectively and successfully harness data analytics for your organization.

---

## ABOUT GOVLOOP

GovLoop's mission is to inspire public sector professionals by serving as the knowledge network for government. GovLoop connects more than 270,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to the public sector.

For more information about this report, please reach out to [info@govloop.com](mailto:info@govloop.com).

[govloop.com](http://govloop.com) | [@govloop](https://twitter.com/govloop)

## THANK YOU


Thank you to Carahsoft, Dell and Hortonworks for their support of this valuable resource for public-sector employees.

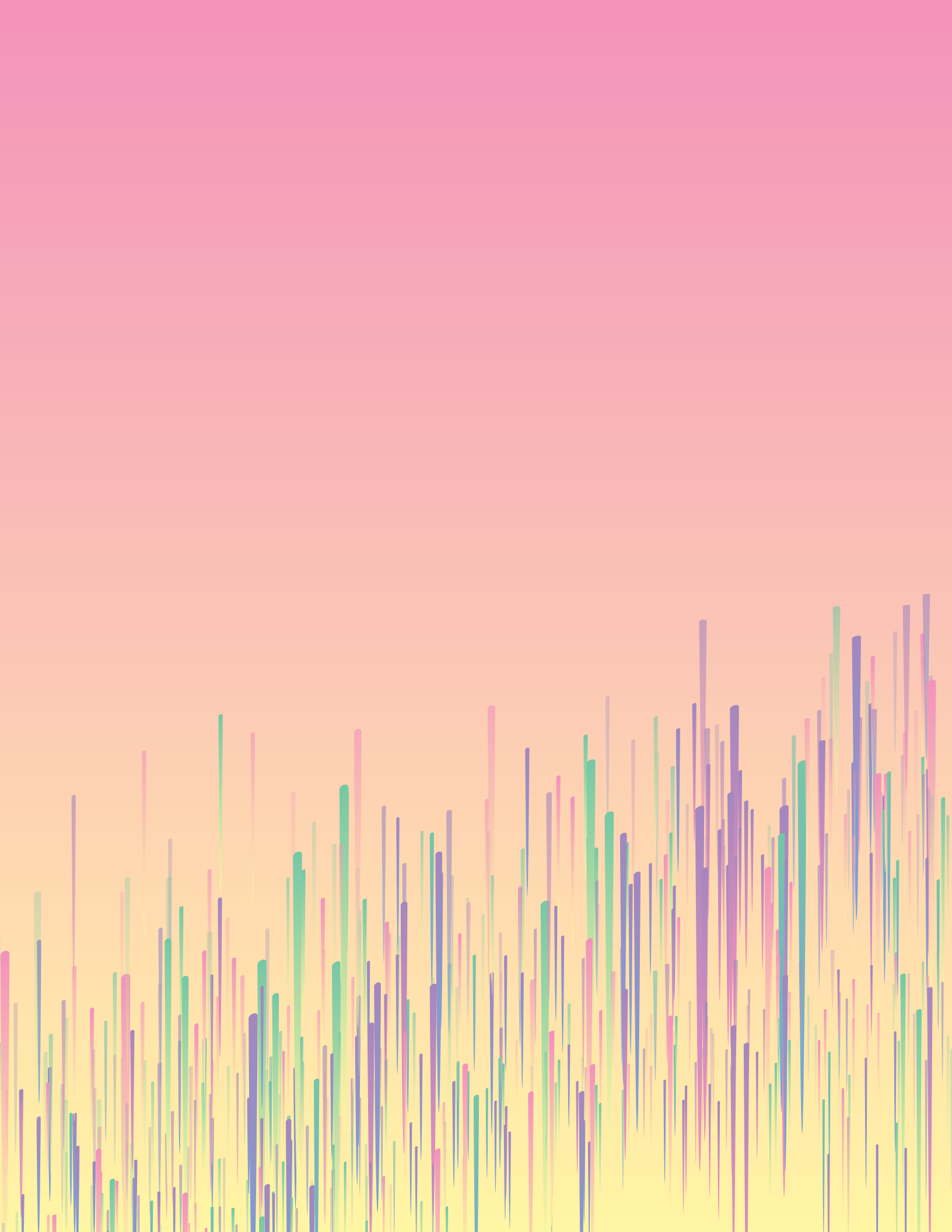
## AUTHORS

Francesca El-Attrash, Senior Staff Writer  
Danielle Poindexter, Editorial Fellow

## DESIGNER

Kaitlyn Baker, Creative Lead







1152 15th St. NW Suite 800  
Washington, DC 20005

P: (202) 407-7421  
F: (202) 407-7501

[www.govloop.com](http://www.govloop.com)  
[@GovLoop](https://twitter.com/GovLoop)