THE CASE FOR GOVERNMENT INVESTMENT IN ANALYTICS

WHY EVERY GOVERNMENT EXECUTIVE SHOULD CARE ABOUT DATA

SEPTEMBER 3, 2019

BY

JANE WISEMAN

Contents

Executive Summary

Government stands to gain $1 trillion globally from using data analytics.[1] Few government data teams have the resources to document their value, but those that do can show as much as eight-to-one return on their cost. There is significant non-financial benefit as well, as public faith in government may improve when saving time and money is paired with increased transparency and accountability.

Many leading governments are already innovating with data and improving results, but these exemplars remain in the minority. While there are over 30,000[2] units of local government, only two dozen local governments have a data leader who participates in the Civic Analytics Network, a peer network of data leaders hosted by Harvard Kennedy School. Similarly small numbers of state and federal data leaders are in place. And fewer still have been able to document the concrete value of their results.

How can other state and local governments tap into this potential for public value? And how can they measure their impact and demonstrate value? This paper documents successful data analytics efforts in government and describes approaches to calculating returns. The purpose of this paper is to enable jurisdictions to make the case for investment in data analytics with a goal of advancing the state of data-driven government.

While there are many possible ways to describe the value to government of using data, this paper addresses three types of value created:

- Financial return attributable to analytics efforts;
- Operational process improvements achieved due to data and analytic approaches; and
- Increased faith in government attributable to data and transparency efforts.

With increasing availability of low-cost tools and large volumes of data for analytics, now is an excellent time for further investment in government analytics capabilities. Low cost and user-friendly analytics tools such as visualization and dashboarding allow for pattern analysis. Advanced analytic models can identify and predict negative outcomes that would have been overlooked by human judgment alone. Internet of Things (IoT) sensors, drones, and modern mapping tools have rapidly increased the availability and speed of location-based data analysis. In this environment, government leaders should carefully examine the successful examples here of providing financial benefit, operational efficiency, and improved faith in government.

---

[1] Cunningham, Susan; Davis, Jonathan; and Dohrmann, Thomas, "The trillion-dollar prize: Plugging government revenue leaks with advanced analytics," January 2018, McKinsey & Company.

[2] United States Census Bureau, Census of Governments, 2017.

Acknowledgements

Introduction

While the application of big data analytics in government is not new, it is yet to be uniformly adopted. And yet, there is significant value to be achieved for those jurisdictions that adopt data analytics and related digital and technology approaches.

The consulting firm McKinsey estimates that globally, government stands to capture $1 trillion by using big data analytics to identify both revenue not collected and to recoup payments made in error.[3] Looking at the return on investment, the firm's research shows that efforts to apply data analytics to eliminating waste, fraud, and abuse in government can have returns as high as 10 to 15 times their cost.[4] There is significant non-financial benefit as well, because as governments improve efficiency and decrease waste, fraud, and abuse, their esteem among the public grows and faith in government improves.

A great deal of excellent work is being done by data leaders in federal, state, and local government, as the examples in this paper demonstrate. Often, data leaders such as chief data officers are under-resourced and juggling competing demands on their time. Precious little time is left for calculating and documenting the return on investment for their projects. And yet, setting aside time to evaluate and record the benefits of their work might result in government leaders realizing how powerful an asset they have and investing more resources.

Leaders in jurisdictions and agencies without a current data leader may see that creating such a position would bring value and help advance an agenda of not only better service to the public, but also improved public faith in government.

Many experts believe that the gap between analytics leaders and laggards is growing, not closing. As technology author Tom Davenport points out, advanced analytics tools are likely to help those already at an advantage as "big companies get bigger."[5] This is true in the public sector, as networking helps the leaders accelerate their achievements, while momentum to appoint new data officers is dissipating.

With this in mind, it is critical that state and local governments that have not already become leaders in analytics begin now to address this challenge, and an important first step is making the case for investment. This paper aims to help in that regard by providing evidence as well as

---

[3] Cunningham, Susan; Davis, Jonathan; and Dohrmann, Thomas, "The trillion-dollar prize: Plugging government revenue leaks with advanced analytics", January 2018, McKinsey & Company.

[4] Cunningham, Susan; McMillan, Mark; O'Rourke, Sara; and Schweikert, Eric, "Cracking down on government fraud with data analytics", October 2018, McKinsey & Company.

[5] Davenport, Thomas H. *The AI Advantage: How to Put the Artificial Intelligence Revolution to Work*, MIT Press, 2018, page 32.

an analytical framework for measurement of three types of value created: financial, operational and public trust in government.

Analytical Framework

The table below describes the three categories of public value: financial, operational, and public trust in government, and then provides examples and methods of measuring each type of result.

Analytical Framework for Public Value Realization For Government Data Analytics		
Type of gain achieved	**Example**	**Possible Methods for Measurement of Public Value**
FINANCIAL RETURN		
Fraud detection cost savings	• Fraudulent tax returns identified by matching across databases, probabilistic matching, anomaly detection, natural language processing, machine learning, network analysis, etc. • Fraudulent unemployment insurance claims detected by matching against other government databases • Fraudulent disability claims detected via matching data with internal and external sources • Fraudulent prescriptions detected by matching against vital records, etc.	• Dollars of increased revenue, number of fraudulent payments stopped • Number of prosecutions for fraud • Value of assets returned to proper usage • Increase in compliance, and contribution to fairness
Efficiency improvement cost savings for a service or function	• Temperature control accuracy improved while reducing power usage with sensors and analytics • Streetlight energy use lowered with smart sensors • Optimizing routes with real-time location and demand data, shortening routes or idling times via smart routing	• Cost savings on contracted service, product, or utility • Fuel cost saved • Environmental savings (e.g., reduced carbon emissions, etc.)
Accuracy improvement which results in cost savings	• Error checks performed by comparing across datasets or within systems with internal flags for exceeding expected threshold values • Compliance checks with automated flags for outlier data points	• Inaccurate payments prevented • Overpayments averted • Fees or fines for improper payments averted
Increased revenue capture	• Identifying missed opportunities for revenue capture with data and digital tools • Efficiency of parking improved with smart metering	• Increased tax receipts • Increased user fees • Revenue gains with additional transaction capture (e.g., parking)

Analytical Framework for Public Value Realization For Government Data Analytics		
Type of gain achieved	**Example**	**Possible Methods for Measurement of Public Value**
OPERATIONAL PROCESS IMPROVEMENT		
Efficiency improvement for a service or work process that results in improved outcomes or time savings	• Predictive algorithms that use machine learning for structured data, natural language processing for unstructured data, etc. • Geospatial analysis to identify hotspots and service gaps • Restaurant inspectors find violations of health code faster • Child welfare workers identify children at risk of harm faster • Child disease risk targeting (lead poisoning, etc.)	• Compliance increase with better targeting of inspections • Lower rates of bad outcomes (child maltreatment, etc.) • Labor hours saved in completing caseload, so that other tasks can be taken on, such as a backlog of cases • Lower cost to treat problems that are identified earlier
Safety improvement	• Improved infrastructure through prioritizing of risk (water mains, bridges, roads, etc.) • Data-driven decision making for pretrial detention • Crime prevented via predictive models and proactive problem-solving • Time reduced to identify suspects for crimes with sensor data or video and geographic analytics	• Lives saved • Injuries avoided • Property damage avoided • Lost work time avoided • Productivity improvement and tax revenue generated from individuals not detained pretrial due to data-driven decision-making
TRUST IN GOVERNMENT		
Transparency of government and public awareness	• Visibility for existing efforts increased by public awareness campaign and aggressive promotion of open data • Public engagement deepened through digital and in-person town hall or other fora • Civic engagement in government operations increased via participatory budgeting and the like	• Citizen satisfaction surveys • Customer feedback surveys for specific services • Social media sentiment mining • Open data and city APIs that show the breadth and types of data that governments collect and manage (privacy and ethical concerns)

The pages that follow describe selected success cases for government in using data and analytics to improve public outcomes. Jurisdictions seeking to establish or expand analytics programs should use examples provided here, along with locally-relevant data and examples, in making their case for investing in analytics and data-driven approaches.

Financial return on investment

Very few government performance or analytics teams have the available resources to measure the financial return on investment for their work. Two great examples are the cities of Louisville and Cincinnati:

- The Louisville Metro Government calculates a five-to-one return for every dollar of cost to the government for its innovation, data, analytics, and performance management efforts.[6]
- The Cincinnati Office of Performance and Data Analytics uses performance management, open data, and advanced geospatial analytics to drive process improvement and has achieved $6.1 million in value for the city in its first two years. Given a cost of approximately $700,000 over the two=year period, that is a return of nearly nine times the investment.[7]

Financial returns from analytics projects can take a variety of forms. Examples here address cost recovery and revenue gained through fraud detection, process efficiency improvement, improved data and service targeting accuracy, and revenue capture.

Fraud detection

Perhaps the easiest returns on investment for data analytics in government to quantify are returns from investigating fraud, which is a significant problem for government. McKinsey estimates that federal government agencies aren't even aware of the vast majority of the fraud that they suffer from, with $57 billion of fraud or abuse known by the government and another $91 billion going completely undetected.[8] McKinsey experts warn that fraud against government may only get worse in coming years, as private sector entities become more vigilant against outside attacks and as electronic transactions replace cash via growing digital commerce. These forces will drive more fraudsters and organized crime syndicates to turn to government targets instead.

The examples that follow demonstrate the range of ways government fraud detection with analytics can generate financial value.

Identifying Medicaid and Medicare fraud

---

[6] Results for America, "Louisville: Open Data, Performance Management, and Continuous Improvement," June 30, 2015.

[7] Theo Douglas, "5 Ways Data Analytics, Performance Agreements Saved Cincinnati $3.3 Million in Fiscal 2017," Government Technology, May 10, 2017.

[8] Cunningham, Susan; McMillan, Mark; O'Rourke, Sara; and Schweikert, Eric, Cracking down on government fraud with data analytics, October 2018, McKinsey & Company.

The mission of the United States Department of Health and Human Services (HHS) Office of the Inspector General (OIG) is to fight waste, fraud, and abuse in the $1 trillion spent on Medicare, Medicaid and other HHS programs. Among HHS OIG's 1,600 employees is a Chief Data Officer (CDO) who holds the rank of Assistant Inspector General, on peer status with the agency CIO and CFO. CDO Caryl Brzymialkiewicz works with inspectors to find ways to use data to advance their work and returns $5 for every $1 of cost by using data to find fraud.[9] The team uses data analytics to find fraudsters who "had figured out how to hide in the data," said Brzymialkiewicz. Some examples of the team's high-profile prosecutions of fraud include:

- A $1.3 billion anti-fraud takedown in 2017 brought to justice 57 physicians, 162 nurses and 36 pharmacists for fraudulent prescribing of opioids. This largest-ever takedown of fraud resulted from analysts working with investigators to uncover patterns of prescribing that indicated fraud.[10]

- Data analytics helped uncover $1 billion in fraud in 2016, charging 301 people with unnecessary treatment, bribes and kickbacks, identity theft, and false prescriptions. Data modelers and statistics experts examined more than a petabyte of Medicare payment data. Comparing it with external intelligence derived from field agents, the analysts were able to uncover the fraud.[11]
- A Detroit doctor was sentenced to 45 years in prison for fraudulently receiving $18 million for using cancer treatment drugs on hundreds of patients who didn't need them, including patients who didn't have cancer.[12] Data analytics caught him by identifying the pattern of his misuse.

Finding unemployment insurance fraud in Texas

The State of Texas has found and prevented $90 million in fraud using analytics.[13] Ed Kelly, Data Coordinator for the State of Texas Department of Information Resources, has led transparency

---

[9] United States Department of Health and Human Services, Office of Inspector General, Fiscal Year 2019 Justification of Estimates for Appropriations Committees.

[10] Ruth Ann Dorrill, *HHS OIG: Federal Healthcare Insights*, Lone Star Winter Institute (presentation), January 2018, http://lonestarhfma.org/wp-content/uploads/2015/06/180105-Dorrill.pdf.

[11] Frank Konkel, "Better Data Just Saved Taxpayers $900 Million in Medicare Fraud," *NextGov,* Government Executive Media Group, June 23, 2016, https://www.nextgov.com/analytics-data/2016/06/better-data-just-saved-taxpayers-900-million-medicare-fraud/129357/.

[12] Sonia Moghe, "Patients give horror stories as cancer doctor gets 45 years," *CNN*, July 11, 2015, https://www.cnn.com/2015/07/10/us/michigan-cancer-doctor-sentenced/index.html.

[13] Wood, Colin. "Data-sharing agreement saves Texas $90 million," Statescoop, January 11, 2019.

and open data, data sharing, and data literacy efforts in the state of Texas. One example of a high return on data investments in Texas is at the Texas Workforce Commission (TWC), which has identified $90 million in fraudulent unemployment benefits that would have been claimed by individuals incarcerated in the state's prisons and jails. By comparing unemployment claims data with data from the state's Department of Criminal Justice, the state can now ensure that unemployment benefits are not paid to the incarcerated. In the early years of the program over $25 million a year was identified, and in 2018 it was $18 million. This downward trend in averted fraud may be an indication that word is out that the analytics team will identify potential fraud, deterring additional attempts.

Finding tax fraud in Maryland

By comparing across multiple databases across different parts of government, for example business licenses, actual versus expected sales tax receipts, employee income reporting, owner income statements, and the timing of payments, government can predict and prevent problems like insolvency, failure to pay taxes, or underpayment of taxes.

The State of Maryland is more than ten times more efficient than before at identifying fraud with analytics.[14] Using data on risk profiles of individual returns and patterns of anomaly across returns allows the state to target which tax returns to audit. Before using data to prioritize, the state was auditing 110,000 tax returns a year and finding fraud in five to 10 percent of those audits. Now, the data-driven approach enables half as many audits to find fraud 60 percent of the time. The data-driven approach is over ten times more effective at finding fraud, which makes more judicious use of the auditing resources of the state, gains additional recouped taxes from those who perpetrate fraud, and increases fairness as fraudsters are forced to pay their fair share. The approach has recovered nearly $35 million, compared with an average of $10-20 million before implementing the data-driven approach.

Cost savings from improving efficiency

Most data analytics projects improve the efficiency of a government process or service by eliminating bottlenecks or redundancies, but as noted earlier, very few examples have been documented. Estimating the improved efficiency of a service requires starting with a baseline measurement in order to benchmark against that – something seldom done in government. It also requires taking a measurement of process efficiency at the end of the project for comparison purposes. Calculating the amount of time and money to achieve an outcome or complete a process in government is complicated by the many steps and stakeholders involved and because doing so is uncommon. This makes the examples shared here all the more valuable for their rarity.

---

[14] Shacklett, Mary, "Fighting tax return fraud with analytics," Tech Republic, April 14, 2016.

UK Digital Service saves billions in technology procurement costs

The UK Government Digital Service team was created in 2011 to help government agencies in the UK implement their government's "Digital by Default" strategy and to support government-wide digital transformation. The team works across the disciplines of digital, technology, and data to unlock efficiencies.

One project looked at technology spend across the largest agencies of national government and, using data analytics and mapping, identified how a small handful of large firms had locked in pricing that while advantageous to themselves left most government agencies overpaying compared to their peers.

Implementing a program of spend controls based on data insight, and a pipeline planning process to reduce redundancies, has saved over one billion pounds ($1.24 billion) by the third year of implementation, including 353 million pounds ($430 million) in the past year.[15] Increasing competition in the technology procurement portfolio also significantly increased quality and customer satisfaction among government buyers. The proactive approach to planning for technology projects in the pipeline has saved significant staff time in avoiding repetitive procurements and building on common tools. Agencies estimate their technology procurement process has been made more efficient with savings of 35-40 percent of staff time in the selection process with no increased risk of project failure through this collaborative approach.[16]

Savings in technology procurement are expected to continue and possibly increase as the Government Digital Service partners with the UK Government's Crown Commercial Service (CCS) on a government-wide shared Digital Marketplace. Further financial gains are expected as the Digital Marketplace goes global with its data-driven and user-centered anti-corruption program to identify improper spending across technology contracts.[17] The goal of these efforts is "smarter spending of taxpayers' money and a more effective digital transformation of key public services" according to Warren Smith, Global Digital Marketplace Programme Director in the UK Government Digital Service.

Boston saves money and lowers carbon footprint with data analytics monitoring

---

[15] Bond, Chad and Tyronne Fisher, "Spend controls: saving money and making things better," Blog post, UK Government Digital Service, 16 July 2019.

[16] Ibid.

[17] Smith, Warren, "How GDS is helping tackle global corruption," Blog post, UK Government Digital Service, January 22, 2019.

The City of Boston saves $5 million a year, along with eliminating 20,000 pounds of carbon emissions, with an algorithm to optimize school bus routing.[18] The city bore no cost to implement the solution as it was develop by graduate students from the Massachusetts Institute of Technology as part of a hackathon, hosted by the city and with prize money sponsored by a private funder.

The city also saves $1 million a year on city building energy costs with real-time monitoring and an energy manager who can strategically adjust consumption during peak cost times.[19] This savings includes $40,000 per year in savings at one building alone, the city's main library.

Savings from improved accuracy

When creating predictive analytics algorithms, most data scientists estimate that a significant portion of their time (up to 85 or 90 percent of time) will be spent cleaning their data. Government datasets are often riddled with incomplete records, inaccurate and sometimes conflicting information. Efforts to improve data completeness and accuracy can not only improve quality, they can also sometimes identify savings of time or money.

Improving compliance with business registration rules in Kansas City

Kansas City, MO improved business license registrations with plain text notices and found recipients 22 percent more likely to comply and renew their licenses, saving time and money for the city and businesses.[20] By law, all businesses in Kansas City, MO need to obtain and renew their occupational license every calendar year. But about half of businesses fail to renew each year, creating significant work for city staff in mailing out repeated follow up notices, and causing businesses to incur late fees that mount with each increasing deadline that passes.

To address this issue, the city data team partnered with the Behavioral Insights Team to test ways of getting more businesses to renew. In an experiment, the city changed the renewal notice from its current form to a plain language format. The new notice was sent to a random sample of over 6,000 businesses. Those receiving the plain language notice were 22 percent more likely to renew their business license than those in the control group.

Improving the accuracy of tax bills for the San Francisco Assessor's Office

---

[18] McGinty, Jo Craven. "How Do You Fix School Bus Routes? Call MIT." Wall Street Journal, August 12, 2017.

[19] Colin Wood, "How Boston plans to save $1 million by watching its power bill" Statescoop, March 30, 2017.

[20] Bo McCall, Aaron Dispenza and Trent Mordhorst, Increasing Registrations for Delinquent Business Licenses in Kansas City, Missouri, Kansas City Office of Performance Management.

Data analytics added $2.8 million in immediate property tax revenue to the City and County of San Francisco by reducing backlogs in the Assessor's Office, and uncovered potential property tax avoidance by flagging lower than market rate housing sales.[21] This is significant, since nearly one third of city revenue comes from property taxes, and in a city with burgeoning wealth and also inequality it is helpful to have everyone pay their fair share.

The San Francisco Assessor's Office must assess how much a new owner should pay in property taxes every time a home is sold in that city. Due to California's Proposition 13, every time a property is sold, the Assessor must reset the property's assessed value to market value. Most of the time the sale price reflects the market value, but sometimes, when there is a property transfer at lower than market price to a friend or family member, the city loses out on tax revenue if the new property tax assessment is set based on the sale price, not the market value.

The challenge for the Assessor's Office was not just to find possible cases of avoided property taxes, but also to prioritize their work, as they had a significant backlog, as much as three years in some cases. The data team helped the Assessor's Office predict instances when property prices did not reflect market rates by building a predictive analytics model using data on when and where the property sold and property details like the year built and square feet to generate a prediction. The first rollout of the predictive model helped remove 166 cases from the backlog and generated $2.8 million in tax revenue for the city.

Increased revenue capture

Using data analytics to improve a process can optimize fairness, as in cities using data to implement demand-based pricing for parking. Data and digital tools can also identify areas where revenue capture or user fees were inefficient and point to areas for increased fee capture.

Using process improvement to create a more efficient tax collection process in Kansas City

Kansas City, MO tripled tax revenue using data analytics and lean process improvement. Under the leadership of Chief Data Officer Eric Roche, the data team (DataKC) worked with the city's legal department to examine the process for collection of overdue taxes owed to the city. Insights provided by the data led to a new way of case processing and the decision to add two dedicated staff for pursuing late tax payments. Annual collections rose from $1.1 million in FY15 to $3.2 million in FY18. Counting staff salaries, this is a return of eight times the investment.[22]

---

[21] Hicks, Kim, Data Scientist, and Dana Cano & Chris Castle, Principal Real Property Appraisers, City and County of San Francisco, "Streamlining property tax appraisals."

[22] Eric Roche, Chief Data Officer, Kansas City, Missouri, email correspondence April 17, 2019.

Operational process improvements

Data analytics and digital transformation can significantly improve the efficiency of government operations. Consulting firm Deloitte has estimated that applying automation technologies such as machine learning, natural language processing, and robotics could save 1.2 billion hours of effort by federal government workers and save $41.1 billion.[23]

Applying data analytics can turbo-charge a government efficiency effort. The consulting firm McKinsey surveyed state, local, and national government officials in 18 countries and found that most government transformation efforts fail, with less than one in five "very or completely" successful in achieving their goals.[24] **However, the government transformation efforts that used analytics were twice as likely to succeed as those that did not.** As an example, the report references a McKinsey client that saved $10 million a year when it used analytics to optimize the efficiency of its fleet of 7,000 vehicles, in part by cutting the use of short-term car rentals by 70 percent.

The City and County of San Francisco has leveraged data and process improvement efforts to great effect, and has started to quantify this value. San Francisco's Data Academy, at the cost of about one FTE, saves $1.7 million a year for the city by reaching 700 employees who can then implement time-saving and money-saving technology innovations.[25]

The examples below describe success in achieving results in the two main categories of operational value from data analytics: efficiency gains and safety improvements.

Efficiency gains

Applying data and technology to service delivery brings about efficiency that can be quantified in either time saved, which can be redeployed to other tasks, or in revenue generated or cost avoided. The examples here demonstrate the wide array of ways government is achieving real benefit from analytics.

Massachusetts serves needy families without interruption from unanticipated absences

In Massachusetts, a central human resources analytics center of excellence supports departments in answering questions such as: Are certain demographic groups turning over at higher rates than others? Can I do smarter workforce planning by better understanding

---

[23] William Eggers; Schatsky, David; and Viechnicki, Peter; "AI-Augmented Government: Using Cognitive Technologies to Redesign Public Sector Work," Deloitte University Press, April 2017.

[24] Allas, Tera; Dillon, Roland; and Gupta, Vasudha; "A Smarter Approach to Cost Reduction in the Public Sector." June 2018, McKinsey & Company.

[25] DataSF blog post, "The Results are in: Data Academy Makes a Big Impact!" June 26, 2017.

patterns of attrition and new entrants? And, can I understand how the use of overtime compares across agencies, locations and job types?

In one example of the central analytics hub helping a department, at the Department of Transitional Assistance (DTA), which provides supports to low income families, leadership was concerned about the impact of unanticipated staff absences on vulnerable client families. Exploring type of absence, seasonality, times of month, days of week, types of caseworkers, and different area offices led to insights that allowed, for the first time, advance planning for shortages of workers. As a result, during the heavy vacation season, instead of falling behind in processing 7,000 cases, DTA engaged other staff to fill the gap, and was able to process an additional 11,500 cases —maintaining progress in critical timeliness performance measures while giving the agency a jump start heading into the new year.[26]

Chicago improves efficiency of rodent abatement with analytics

The city's rodent abatement program is now 20 percent more efficient based on the predictive model used to identify and prevent problems before they occur.[27] This algorithm is based on a variety of types of 311 data ranging from stray animal calls to vacant and abandoned buildings to restaurant complaints. The predictive analytics project helps more effectively allocate scarce resources and to target the locations most in need of preventive baiting. In the first weeks of the pilot, an address with no prior rodent calls was identified by the predictive algorithm. When the rodent baiting team arrived, they found an infestation so severe the unit had to be condemned.

The data program in the City of Chicago has been successful in not only increasing transparency and data sharing across agencies, it has also resulted in tangible manpower efficiencies from the publication of open data. For example, the city's Department of Public Health noticed a 40 percent decrease in Freedom of Information Act requests after the launch of the city's open data portal. This has allowed that department to deploy its resources on higher value-add activities rather than photocopying documents and mailing them to requestors.

Improving efficiency of fraud inspections at the USPS

The chief data officer at the United States Postal Service (USPS) Office of Inspector General (OIG) has returned significant public value using analytics models to find waste and abuse of funds, with over $920 million returned via analytics in 2016 alone.[28]

---

[26] Accenture Consulting, "Measured and Treasured: Massachusetts Puts HR Analytics to Work for its People," 2017.

[27] Sean Thornton, Using Predictive Analytics to Combat Rodents in Chicago, Data-Smart City Solutions, July 12, 2013.

[28] CDO Club, *Kelly Tshibaka,* CDO Summit Speakers, July 2017, http://dc.cdosummit.com/speakers/kelly-tshibaka/.

One recent project involved creating a predictive analytics model and visualization tool to help USPS OIG auditors and investigators identify high-risk contracts so that they could prioritize their work, with data-driven leads based on known fraud, waste, and abuse schemes rather than relying solely on tips for their work. The tool enabled analysts and investigators to access and visualize data on high-risk contracts and automated the identification of risk with anomaly detection. Leads generated by the new tool proved to be valuable, with 74 percent of the highest scored contracts showing evidence of fraud, waste, or abuse.[29]

Applying the contracts tool to health care fraud investigation, the number of hours worked per case decreased by 30 percent and dollars saved increased by 35 percent. Using the tool, investigators found over $11 million in recoveries, restitutions, and cost avoidance.[30]

Safety improvements

Safety improvements enabled by data analytics span a wide range from preventing childhood lead paint exposure and predicting and preventing child maltreatment, to averting foodborne illness through targeted inspections, to preventive maintenance of infrastructure that averts water main breaks and bridge collapses. The examples below provide attempts to quantify immediate results such as savings from preventive versus emergency repairs. More difficult to quantify and not addressed here are the harder to measure effects such as acute and chronic diseases prevented or the injury or loss of life averted when safety prevails.

Data-driven safety improvement for aging infrastructure in Syracuse

Syracuse, New York is a city with aging infrastructure, with some water mains as much as 100 years old. When one breaks there is not only great inconvenience for the public who go without water and need to avoid the location of the break, but the emergency repairs can also be very costly to city coffers. So, the city's Chief Data Officer Sam Edelstein set out to solve the problem. His motto is, "First, figure out how to count things accurately, which turns out not to be that easy in government." And as he pointed out, only with accurate data are advanced analytics possible.

Partnering with the University of Chicago through their Data Science for Social Good program, the city has saved over $1 million on emergency water main repairs, using predictive analytics to target preventive maintenance. The predictive model analyzes over 400 miles of water main and prioritizes the risk of a break using data about road ratings, soil characteristics, paving history, weather and traffic among other variables. By identifying the riskiest water mains and prioritizing them for replacement first, the city saves money and minimizes public inconvenience. The predictive algorithm is far more objective than repaving based on staff

---

[29] Elder Research, Reducing Fraud, Waste, and Abuse at the United States Postal Service, case study.

[30] Ibid.

subjective assessment, and mapping the results enables a proactive approach to integrate with ongoing city sewer and road paving efforts.[31]

Two other notable projects in Syracuse delivered tangible results to taxpayers. In using a "nudge" to get more delinquent taxpayers to pay overdue amounts, the city's finance commissioner put hand-written notes on the outsides of the envelopes of letters to delinquent taxpayers and generated $1.5 million in addition revenue. Another project analyzed data on road ratings by the repaving staff and was able to prioritize preventive work to keep roads in good repair rather than waiting until they were in such bad shape that they required more expensive measures.

Kansas City reduces unsafe blight conditions by getting to lagging inspections 11 times faster

Kansas City, MO reduces blight conditions and improves housing safety with data analytics. The DataKC team worked with the city's Neighborhoods and Housing Services department to analyze inspection times for city housing code violations in an effort to help prioritize workload and improve outcomes.

While the data analysis found that most inspections were completed in a timely manner (50 percent of inspections were done within nine days), it took 131 days to complete 90 percent of the inspections.[32] This meant that there were some inspections lagging significantly, with housing conditions deteriorating for as much as four and a half months before an inspection was done. Data analysis uncovered this issue, and process improvements have been designed so that now 90 percent of initial inspections are completed in just 12 days – a 119 day improvement. The city is now mitigating potential blight and unsafe conditions far faster than before the data analytics project.

New Orleans speeds ambulance response time, improves fire safety, and reduces blight with data

In New Orleans, the data analytics team works with city agencies on high priority projects and has improved public safety in three ways:

- Both the efficiency and equity of ambulance response time was improved with data analytics and mapping, by determining the optimal locations for ambulances to wait in between calls. Faced with increasing demand for ambulance services, a decrease in ability to quickly respond to the most urgent calls, and disparity in response times by neighborhood, the city EMS team turned to the Performance and Analytics team for help making sense of the data and to find a new deployment map. By mapping five

---

[31] Sam Edelstein, Medium blog post, https://medium.com/@samedelstein/chief-data-officer-7bfc16b2401c.

[32] Eric Roche, Chief Data Officer, Kansas City, Missouri, email correspondence April 17, 2019.

years of 911 call data, the team ran multiple simulations and developed a map of where to place ambulances to cover calls most efficiently. Analysis of existing response times by neighborhood, geographic and temporal demand data, ambulance placement, and calls for service data enabled more strategic deployment of ambulances to reduce response time for emergency medical services.[33] Night shift response time decreased and the response time in the two neighborhoods with the slowest baseline response times improved by nine percent and 20 percent.

- In its first year of implementation, the city's data-driven smoke detector project literally saved eleven lives because the family was able to flee a burning building when their smoke alarm went off. The family had been given a free smoke alarm in a new data-driven distribution program. The analytics team optimized the city's smoke alarm distribution program in two ways – they predicted which homes were most likely to have a fire, and predict which homes most likely lacked smoke detectors. Those two predictions pinpointed where the fire department could prioritize distribution of alarms, block by block.[34]
- The city's blight reduction project addressed 15,000 blighted properties and cut in half the inspection time (from 160 to 80 days), while reducing the backlog of appeals, all by applying data analytics and a stat process to managing the effort.[35]

Chicago improves safety of beaches and restaurants with data analytics

The City of Chicago developed sophisticated predictive analytics models that have resulted in major quality of life improvements for city residents, including increased effectiveness of restaurant inspections so that the most risky inspections are now done seven days sooner,[36] and improved the timeliness and accuracy of estimates of unsafe swimming conditions at city beaches.[37]

---

[33] Linda Gibbs and Maia Jachimowicz, "In New Orleans, Using Administrative Data to Save Lives," Route 50, May 30, 2018.

[34] Katherine Hillenbrand, "Predicting Fire Risk: From New Orleans to a Nationwide Tool," Data-Smart City Solutions, June 9, 2016.

[35] Charles Chieppo, New Orleans' Winning Strategy in the War on Blight, Data-Smart City Solutions, March 18, 2014.

[36] Julian Spector, "Chicago Is Predicting Food Safety Violations. Why Aren't Other Cities?" Citylab, January 7, 2016.

[37] Sean Thornton, "Making Chicago's Beach Water Safer With Analytics," Data-Smart City Solutions, September 12, 2017.

Improved trust in government

Not all value that data leaders deliver is easily quantified. In fact, some types of public value are very difficult to measure or are measured very rarely, such as the contribution of government services to the wellbeing of individuals or society as a whole, or to environmental sustainability or economic mobility.

Some promising efforts are bringing the voice of the customer into government operations, via 311 system customer satisfaction surveys, social media sentiment mining, and online feedback forms, but most of these efforts are narrowly focused on one task or department. A small number of state and local governments regularly measure public satisfaction with their results, but they are the exception rather than the rule.

However, one common measure is public sentiment about government. Public trust in government has reached historic lows in the past decade, and while consistently measured on a national level, very few state and local governments regularly measure public confidence in their work. And yet public trust in government is a key area on which data and analytic approaches can help make improvements.

Improving trust in government with greater transparency of operations

Private-sector research shows that customer satisfaction improves when the customer can see work being done on their behalf.[38] This is why the travel site Kayak shows users a rotating screen of airline names while the search for flights goes on. Research showed that customers had higher satisfaction with the results just by "seeing" this work done on their behalf, even when the prices weren't better. This is the reason we now wait and watch the barista at Starbucks individually steam our lattes, and why the automated Apple voice response system has pre-recorded typing sounds to create the illusion that our query is being typed into their system – because even the illusion of work on our behalf improves customer satisfaction.

Two excellent examples of government applying this approach of "showing the work" are in Buenos Aires, Argentina and in Kansas City, Missouri.

- The City of Buenos Aires, Argentina measured the impact on resident's trust in city government of providing information to the public via its open data portal. An online random assignment experiment of 2,000 residents showed a ten percent increase in public trust among those who saw positive results from city efforts. Interestingly, this experiment also showed the challenge of reaching the public with government performance information as more than 40 percent of respondents had never seen the website or heard about the mayor's public commitments, which covered issues like

---

[38] Norton, Michael and Ryan Buell, "Think Customers Hate Waiting? Not So Fast…" *Harvard Business Review*, May 2011.

access to health care and efficiency of public transportation.[39] The results of this experiment are consistent with a Harvard Business School case study showing that as the public learns more about government operations and has more visibility into activity, trust and satisfaction levels increase.

- In Kansas City, MO, after seeing low citizen satisfaction ratings on snow removal, the city went on a media blitz. It provided information to help educate residents on what to expect — curb-to-curb plowing on main arteries, one lane of travel on residential roads, and the expected timeframe for plowing. In addition, the city manager did a "Tweet-along" while driving in a snowplow, providing real-time updates on the city's progress. Newly-enabled GPS data allowed the public to track the location and progress of snowplows. Survey data showed an improvement in customer satisfaction. Nothing about core operations changed — what changed was the amount of information the public had and their expectations of city performance. Satisfaction went from 50 percent (three years prior) to 62 percent based on these efforts.[40]

Improving faith in government via high-visibility data and technology projects

High profile government engagement in local civic technology and digital inclusion efforts is another way to boost public confidence in government. A good example comes from the City of Louisville, where Chief Data Officer Michael Schnuerle, now chair of the Civic Analytics Network, was himself a prominent part of the civic tech community before joining city government.

Schnuerle says, "As Chief Data Officer my mission is to make government better and improve the lives of residents using data, and a key component of that mission is improving equity in the city. We're working hard to make open data valuable to the public, and giving everyone access and ways to engage. We want to create an active dialog and integrate analysis on top of the data, to create a platform, and work with the community for their needs and to track outcomes."

The spirit of data, tech, and innovation in Louisville is supported both inside city hall and outside of it. A strong Code for America brigade of local volunteers has donated over 15,000 hours to the community, with over 50 meetups, 18 hackathons, and 50 civic projects in their portfolio. A related organization called Code Louisville provides free coding and programming courses and has over 125 graduates that have been hired locally.[41]

As part of Louisville's digital inclusion efforts, the city's Civic Innovation and Technology teams have refurbished $82,000 worth of donated computers and distributed them to low-income

---

[39] Martin Alessandro, Carlos Scartascini, Bruno Cardinale Lagomarsino, Jerónimo Torrealday, "Transparency and trust in government: Evidence from a survey experiment," Inter-American Development Bank, February 2019.

[40] Wiseman, Jane. "Customer Driven Government," *Data-Smart City Solutions*, August 2015.

[41] Michael Schnuerle, Chief Data Officer, Louisville, Kentucky, "Louisville Annual Open Data Report — 2018," Oct 29, 2018.

individuals, signed up over 300 families for reduced-cost high-speed Internet connections, and provided digital skills training.[42]

Additional ways of advancing public trust in government with data and technology

Other methods to improve public trust in government including increasing transparency via open data, providing easier and more convenient ways for the public to engage with government such as digital transactions, and offering the public a voice in their government's decisions via efforts such as participatory budgeting. All of these methods are data-driven or data-enabled.

Internal trust and collaboration can be improved when data leaders break down the silos of government and establish a data culture through formal or informal communities of practice. Training is one powerful way to do this, and as one city CDO said, "it's the most important thing I do because it breaks down silos." Creating standard data sharing agreements so that it is easier to share data across agencies is another way to gradually build relationships of trust.

---

[42] Ibid.

Conclusion

With increasing availability of low-cost tools and large volumes of data for analytics, now is an excellent time for further investment in government analytics capabilities. Increasing digital commerce means fewer cash transactions and lower levels of data entry error, which is producing both higher data quality and lower levels of untraceable transactions. Low-cost and user-friendly analytics tools such as visualization and dashboarding allow pattern analysis. Advanced analytic models can identify and predict negative outcomes, such as health and safety problems or compliance risks, that would have been overlooked by human judgment alone. **Government leaders should carefully examine the successful examples described here of financial benefit, operational efficiency and improved faith in government.**

An important way that a data leader advances public value is by serving as a data evangelist or advocate for data-driven government, and this value is nearly impossible to quantify. However, as a proxy for this, one way that a data leader delivers return on investment is by serving as the point person for pursuing pro bono partnerships, and going after grant funding opportunities and supporting grant pursuits with accurate data. The cities of Chicago and New Orleans have aggressively courted pro bono and low-cost partnerships with private sector and academic partners, to great success.

The city that has been successful with a wide range of entrepreneurial grant funding is Louisville, where the Office of Civic Innovation has brought its city $13 million in grant and in-kind funding over the past seven years. Chief Data Officer Michael Schnuerle has helped since his appointment three years ago. Grants come from funders such as Bloomberg Philanthropies, Living Cities, PNC Bank, The US Department of Transportation, and in-kind contributions from the University of Louisville, University of Pennsylvania, IBM, Amazon Web Services, and Johns Hopkins University. One of the projects pursued via a collaborative partnership allows the city to save between $50,000 and $300,000 a year on traffic studies.[43]

A final recommendation about calculating returns is to use a portfolio approach. Jurisdictions with existing analytics programs should consider calculating their return on investment both on individual projects, and as a portfolio of work. A portfolio approach to documenting return on investment aligns with the way private sector investors approach their portfolio of investments. Further, a portfolio approach allows for the fact that some efforts to gain efficiency do not realize gains right away, but will in the future return the investment.

Science relies on experimentation, and the best organizations allow for creativity and have a willingness to take on risk and embrace failure as a part of the process. In fact, many successful

---

[43] Michael Schnuerle, Chief Data Officer, Louisville, Kentucky, "Louisville Annual Open Data Report — 2018," Oct 29, 2018.

analytics efforts today have achieved their results based on failing first with large "moonshot" projects and later learning and adapting via smaller pilot projects – IBM's Watson is a good example of learning from early failure when their much-hyped early efforts at M. D. Anderson did not quickly find solutions to cancer. Failure is often regarded as the best teacher, as long as we are open to learning and adapting.

The findings shared in this report of positive returns for government analytics efforts in organizations that have a data leader such as Chief Data Officer are consistent with research that shows private sector companies with an executive responsible for data analytics outperform peers on financial metrics.[44] As government leaders consider the returns that can be achieved by investing in data skills and tools, it is helpful to consider the wisdom of one of our founding fathers, Ben Franklin, who once said, "An investment in knowledge pays the best interest."

---

[44] Yang Lee, and Stuart Madnick, Richard Wang, Forea Wang, and Hongyun Zhang. "A Cubic Framework for the Chief Data Officer: Succeeding in a World of Big Data," Sloan School of Management, Massachusetts Institute of Technology. March 2014, page 2.

Helpful sources

As demonstrated in this paper, very few of the 30,000 existing units of local government have experience quantifying the value of their data and analytics projects. No standard tools and techniques have been adopted and no common terminology or methodology exists. As jurisdictions begin to create their own locally-relevant case for investing in analytics, the following resources may be of value for guidance or inspiration. In the annotated list that follows, government sources are described first, followed by academic sources.

United States Office of Management and Budget (OMB)

The United States Office of Management and Budget, part of the Executive Office of the President, provides guidance on management to federal agencies and to state and local government as recipients of federal funds. One helpful resource for governments seeking to conduct return on investment analysis is OMB Circular A-94. This 22-page document provides general guidance for conducting benefit-cost and cost-effectiveness analyses. This document was first issued in 1992, and is regularly updated to provide new discount rates for calculating net present value. This resource is useful for anyone in government seeking to understand the long-term return on investment of an initiative or program, but will be far less useful in estimating short term or small scale effects. Most state and local government analytics efforts do not need the complexity of net present value calculations provided in this circular. The circular can be found at:

https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A94/a094.pdf

United States Government Accountability Office (GAO)

The United States Government Accountability Office (GAO) provides reference tools to the public to help perform audits in the methodical and thorough fashion that the GAO does. The Yellow Book, one of the key reference documents used in a GAO audit, is provided on their website for public reference. This document outlines the requirements for audit reports, professional qualifications for auditors, and audit organization quality control. This may be of interest to state and local government in conducting self-assessments.

The tool can be found at: https://www.gao.gov/yellowbook/overview

The Washington State Institute for Public Policy (WISIPA)

The Washington State Institute for Public Policy (WSIPP) is a nonpartisan public research group staffed by a team of multidisciplinary researchers who conduct applied policy research for the state legislature in a creative and collaborative environment. Created in 1983, WSIPP carries out practical, non-partisan research at the direction of the state legislature or for its Board of Directors. To achieve transparency about their methods, and to help others mimic their

methods, WSIPP publishes a comprehensive manual on their methodology. This 220-page document describes the computational procedures used in the Washington State Institute for Public Policy's Benefit-Cost Model. It can be found at:

http://www.wsipp.wa.gov/TechnicalDocumentation/WsippBenefitCostTechnicalDocumentation.pdf

Harvard School of Public Health Guidelines for Benefit-Cost Analysis

A project at the Chan School of Public Health at Harvard has created a detailed set of guidelines for benefit-cost analysis to help non-profits and state and local governments measure the return on the investments they make in health and development projects. The goal of the project is to provide common methods and tools so that across projects, returns can be compared. The website also offers a rich resource of case studies and methodology references. The website is found at:

https://sites.sph.harvard.edu/bcaguidelines/guidelines/

Urban Institute Pay for Success

The Urban Institute, a Washington, DC-based non-profit policy research and think tank, has a dedicated team supporting Pay for Success programs with tools and resources that may be of interest to state and local government staff seeking to document return on investment. The premise of Pay for Success is that government programs which are deemed successful gain a payback for their investors, while unsuccessful ones do not. This makes measurement of success critically important to the investor, and to all involved.

The Pay for Success website includes a series of publications on evaluations, controlled experiments and other ways of measuring program success. There are also reports on evaluations already conducted by Urban Institute and guidance on how to best set up a program for effective measurement. Insights from this portal may be valuable to state and local government analysts seeking to be more confident of their data and research skills.

The website can be found at www.pfs.urban.org.

Sources

Accenture Consulting, "Measured and Treasured: Massachusetts Puts HR Analytics to Work for its People," 2017.

Allas, Tera; Dillon, Roland; and Gupta, Vasudha; "A Smarter Approach to Cost Reduction in the Public Sector." June 2018, McKinsey & Company.

Alessandro, Martin, Carlos Scartascini, Bruno Cardinale Lagomarsino, Jerónimo Torrealday, "Transparency and trust in government: Evidence from a survey experiment," Inter-American Development Bank, February 2019.

Bloomberg Cities, "How a handwritten note helped Syracuse collect $1.5 million in back taxes," December 12, 2018.

Bond, Chad and Tyronne Fisher, "Spend controls: saving money and making things better," Blog post, UK Government Digital Service, 16 July 2019.

Chieppo, Charles. "New Orleans' Winning Strategy in the War on Blight," Data-Smart City Solutions, March 18, 2014.

Cunningham, Susan; McMillan, Mark; O'Rourke, Sara; and Schweikert, Eric, "Cracking down on government fraud with data analytics," October 2018, McKinsey & Company.

Cunningham, Susan; Davis, Jonathan; and Dohrmann, Thomas, "The trillion-dollar prize: Plugging government revenue leaks with advanced analytics," January 2018, McKinsey & Company.

Davenport, Thomas H. *The AI Advantage: How to Put the Artificial Intelligence Revolution to Work*, MIT Press, 2018.

Diaz, Alejandro; Rowshankish, Kayvaun; and Saleh, Tamim. "Why Data Culture Matters." McKinsey Quarterly. September 2018.

Dorn, Stan; Milner, Justin, and Eldrige, Matthew. "More Than Cost Savings: A New Framework for Valuing Potential Pay for Success Projects." Urban Institute, May 5, 2017.

Dorrill, Ruth Ann. HHS OIG: Federal Healthcare Insights, Lone Star Winter Institute (presentation), January 2018, http://lonestarhfma.org/wp-content/uploads/2015/06/180105-Dorrill.pdf.

Douglas, Theo. "5 Ways Data Analytics, Performance Agreements Saved Cincinnati $3.3 Million in Fiscal 2017," Government Technology, May 10, 2017.

Eggers, William; Schatsky, David; and Viechnicki, Peter; "AI-Augmented Government: Using Cognitive Technologies to Redesign Public Sector Work," Deloitte University Press, April 2017.

Edelstein, Sam, Chief Data Officer Medium blog post, June 13, 2019. https://medium.com/@samedelstein/chief-data-officer-7bfc16b2401c

Elder Research, "Reducing Fraud, Waste, and Abuse at the United States Postal Service," case study.

Hicks, Kim, Data Scientist, and Dana Cano & Chris Castle, Principal Real Property Appraisers, City and County of San Francisco, "Streamlining Property Tax Appraisals"

Hillenbrand, Katherine. "Predicting Fire Risk: From New Orleans to a Nationwide Tool," Data-Smart City Solutions, June 9, 2016.

Hirsch, Michael. Washington State Institute for Public Policy, interview by author, April 4, 2019.

Kelly, Ed. Data Coordinator, State of Texas Department of Information Resources, interview by author March 8, 2019.

Konkel, Frank. "Better Data Just Saved Taxpayers $900 Million in Medicare Fraud," NextGov, Government Executive Media Group, June 23, 2016, https://www.nextgov.com/analytics-data/2016/06/better-data-just-saved-taxpayers-900-million-medicare-fraud/129357/.

Lee, Yang, and Stuart Madnick, Richard Wang, Forea Wang, and Hongyun Zhang. "A Cubic Framework for the Chief Data Officer: Succeeding in a World of Big Data," Sloan School of Management, Massachusetts Institute of Technology, March 2014.

McCall, Bo, Aaron Dispenza and Trent Mordhorst, "Increasing Registrations for Delinquent Business Licenses in Kansas City, Missouri," Kansas City Office of Performance Management.

McGinty, Jo Craven. "How Do You Fix School Bus Routes? Call MIT," Wall Street Journal, August 12, 2017.

Moghe, Sonia. "Patients give horror stories as cancer doctor gets 45 years," *CNN*, July 11, 2015, https://www.cnn.com/2015/07/10/us/michigan-cancer-doctor-sentenced/index.html.

Norton, Michael and Ryan Buell, "Think Customers Hate Waiting? Not So Fast…" *Harvard Business Review*, May 2011.

Results for America, "Louisville: Open Data, Performance Management, and Continuous Improvement," June 30, 2015.

Roche, Eric, Chief Data Officer, Kansas City, Missouri, email correspondence April 17, 2019.

Schnuerle, Michael, Chief Data Officer, Louisville, Kentucky, "Louisville Annual Open Data Report — 2018," Oct 29, 2018.

Shacklett, Mary, "Fighting tax return fraud with analytics," Tech Republic, April 14, 2016.

Smith, Warren, "How GDS is helping tackle global corruption," Blog post, UK Government Digital Service, January 22, 2019.

Smith, Warren, Global Digital Marketplace Programme Director in the UK Government Digital Service, interview by author July 29, 2019.
Spector, Julian. "Chicago Is Predicting Food Safety Violations. Why Aren't Other Cities?" Citylab, January 7, 2016.

Thornton, Sean. "Using Predictive Analytics to Combat Rodents in Chicago," Data-Smart City Solutions, July 12, 2013.

Thornton, Sean. "Making Chicago's Beach Water Safer With Analytics," Data-Smart City Solutions, September 12, 2017.

Tshibaka, Kelly. CDO Summit Speakers, http://dc.cdosummit.com/speakers/kelly-tshibaka/, CDO Club, July 2017.

UK National Audit Office, " Digital transformation in government" Report by the Comptroller and Auditor General, March 30, 2017.

Washington State Institute for Public Policy, Benefit-Cost Model Technical Documentation, December 2018.

Wiseman, Jane. "Discovering the True Value of City Data Experts." Data-Smart City Solutions. Ash Center at Harvard Kennedy School, November 2017.

Wiseman, Jane. "Customer Driven Government," *Data-Smart City Solutions*, August 2015.

Wiseman, Jane. Data-Driven Government: The Role of Chief Data Officers, IBM Center for the Business of Government, September, 2018.

Wood, Colin. "Data-sharing Agreement Saves Texas $90 million," Statescoop, January 11, 2019.

Wood, Colin. "How Boston plans to save $1 million by watching its power bill" Statescoop, March 30, 2017.

United States Department of Health and Human Services, Office of Inspector General, Fiscal Year 2019 Justification of Estimates for Appropriations Committees.