Data & Analytics
Open Data
Implementation Update

City of San Diego
Performance & Analytics
FY 2018

ABOUT THIS REPORT 3

WHAT WE DO 4

HOW WE DO IT 5

OPEN DATA 5

DATA PRODUCTS 7
- StreetsSD 7
- PerformSD 7
- Budget.sandiego.gov 8
- Looking ahead to FY 2019 8

DATA INFRASTRUCTURE 9
- Current State 9
  - Poseidon 9
  - Data Ingestion 9
- Planned State 10

ANALYTICS ENABLEMENT / ALERTING 11
- Current State 11
  - External Data Catalog (data.sandiego.gov) 11
  - Internal Data Cataloguing Tool (Alation) 11
  - Data Analysis (PandaPack) 12
  - Data Analysis (DataLab) 12
  - Data Analysis (Carto) 13
- Planned State 13

ANALYTICS PROJECTS 14
- Current State 14
- Planned State 14

CONSULTING 15
- Current State 15
- Planned State 15

PROJECTS 16
- Project Status 16
- Project Overview 17

COMMUNITY INVOLVEMENT 29

AWARDS 30

OPEN SOURCE 30

OUR FUTURE 31

ABOUT THIS REPORT

We are dedicating the fourth annual report on the Data & Analytics team to Arnaud Vedy. We lost Arnaud in a motorcycle accident in early November. Many of the projects and work you see featured in this report could not have happened without Arnaud's passion, ingenuity, and ability to learn and try new things.

Arnaud was a key component of the Data & Analytics team, and he was in charge of spinning up our Data Science program. He helped us automate a multitude of datasets, implement geospatial processing pipelines, release data in a multitude of formats, and build amazing data science projects.

Today, cities across the country look to our program as an example of how to run a data science team and execute meaningful and impactful projects.

In FY 19, we will be expanding our analytics services, empowering our fellow city employees, building more products, releasing more data, and making data available in more formats to increase the number of people at all skill levels who can use our data.

Most importantly, just like Arnaud, we will focus on **doing** more than talking, **trying** more than planning, and trying new, exciting, and challenging things that have an impact on those around us.

Maksim Pecherskiy

Chief Data Officer, Performance & Analytics

Almis Udrys

Director, Performance & Analytics

WHAT WE DO

DataSD, or the Data & Analytics team in the Performance & Analytics Department, has the mission of helping people **get and use data.**

The City of San Diego Open Data Policy was adopted by the City Council in November 2014, creating the Open Data program, with the goal of using City data to increase transparency, efficiency and accountability. Through the process of understanding of what data the City has, and attempting to release it, we have found multiple opportunities to help residents and City staff get and use data.

Getting Data

To get data, City staff and residents have to know that it exists (**discover**) and that they are able to access it in the most raw format available (**acquire**). DataSD facilitates this process through our internal data catalog tool, Alation, and our public Open Data Portal (**data.sandiego.gov**). DataSD must maintain connections across hundreds of the City's primary data sources, such as mainframes and databases from different vendors, external SaaS providers, and spreadsheets on shared drives. The team also runs and maintains **Poseidon** - a data automation system that moves data from these less accessible data sources to more accessible places, such as the portal, to decrease the friction of acquiring and discovering data.

Using Data

Once a resident or City employee acquires data, she needs to be able to translate that data to information that drives a decision or changes a process. She needs to **understand** what is in the data file, including column definitions and the meanings of codes and abbreviations, as well as more complicated caveats and intricacies. She needs to be able to **ask the right question** that will influence her decision. She then needs to be able to get a **reliable, valid answer** to her question. Lastly, she needs to be able to translate that to a **decision.**

DataSD facilitates the use of data by applying the right tool or technique to the situation. Sometimes, we will partner with a department to perform an analysis or build a product such as **StreetsSD**. Other times, we'll help department staff plan their own project or attend vendor meetings to make sure the City acquires the right solution. Often, we can help City staff build a strategy to systematically solve a problem or automate alerts or reports for better monitoring.

HOW WE DO IT

DataSD's approach is to **define a painful problem and solve it in the simplest, most sustainable way**. We structure our work using a platform approach, which ensures that we never have to solve the same problem, from scratch, twice. We use modern, open-source technology to make our work reproducible and to avoid unnecessary licensing fees. Lastly, we work closely with the Department of Information Technology (DoIT) to coordinate and enable our work, while making it sustainable.

OPEN DATA

This past fiscal year, the Open Data program primarily focused on increasing the stability, usability, and quality of the Open Data Portal. Publishing data is certainly one of our core functions; but we also track whether people are actually able to use the data and remove any difficulties people face.

We collected feedback from various user groups consisting of City staff and residents, to determine how well they comprehend what the Portal is and how to use the data. At first, Portal users tended to have a higher rate of data literacy. They were City staff who knew about the Open Data Policy and had been following its progress, or they were people who had an interest in civic technology. Now that the Portal has been live for more than a year, the number of questions we field through our feedback channels has increased, and the type of questions have changed. The user group has expanded, and therefore the Portal has to be more usable for everyone.

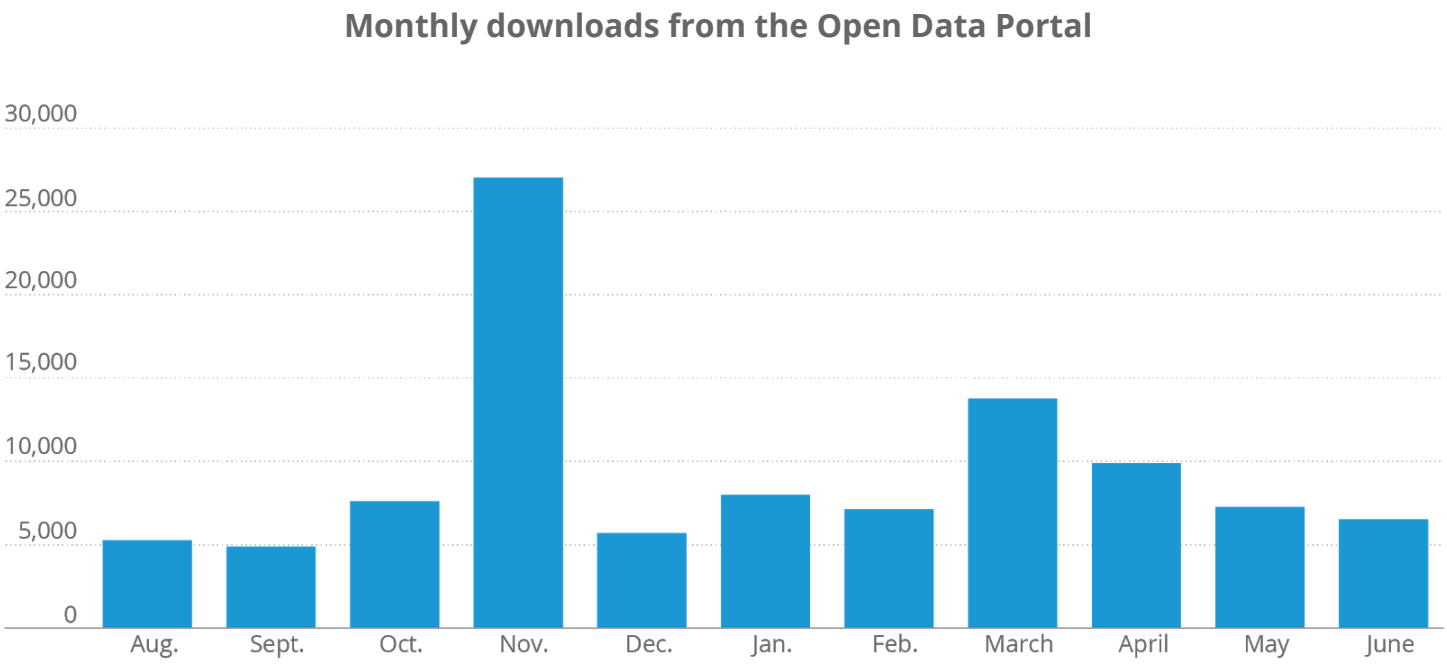
One enhancement has been offering more file types for geospatial data. The only file type available prior was a proprietary format developed by a company that sells expensive software. Now, geospatial data is available in three additional data formats that do not require a software license to use. We now offer geospatial data in GeoJSON, Shapefile, and PBF.

We also increased the amount of historical data for previously published, high demand data sets. One example is the Fire Incidents dataset, which includes 911 calls for Fire and EMS response. This data is requested frequently, and the Department was spending time responding to requests for historical data dating back several years. We now have Incidents by Year dated to 2007.

For more complicated topics, we added reference pages to help users navigate certain datasets. The City's budget data, for example, is broken out into seven different datasets, with four of those datasets serving as reference data. Someone unfamiliar with public sector budgeting might be confused by how the data is organized on the portal, and the reference page is meant to clarify.

Our next improvement that is already in progress is to create more opportunities for people to get value out of data without having to download a file and perform an analysis. We will accomplish this by offering map previews, expanding the number of data stories, highlighting popular and most frequently updated datasets, and by conducting user testing and incorporating feedback into the portal design and user interface.

Adding new data to the portal remains a major focus, and we will continue to work with City departments to get data published. Through Fiscal Year 2018, 68 percent of high-value, public datasets have been published to the open data portal. By the end of FY 2019, 80 percent will be published. Datasets scheduled to be published in FY 2019 include the City's public art collection, aggregated water usage, Capital Improvement Projects, the Sidewalk Overall Condition Index, and enhanced permitting information.

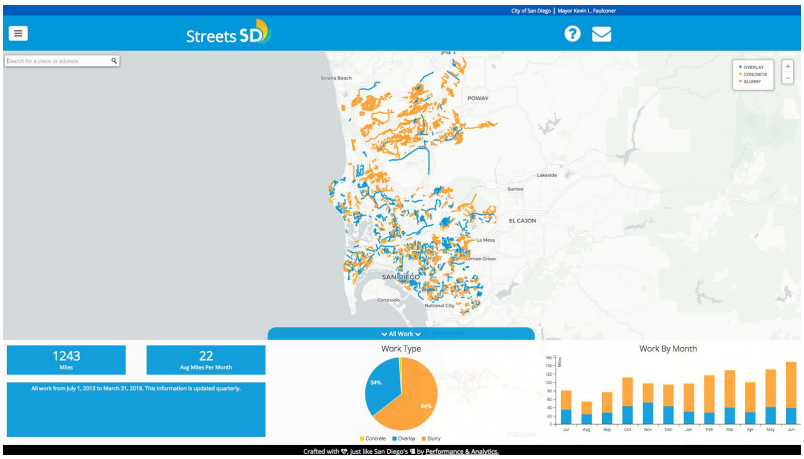


This may not demonstrate the full extent of data usage, since we syndicate our data to external sources and foundries, such as **data.world**

DATA PRODUCTS

As part of our mission to make data actionable and easy to use, we release and maintain various data products. Some of these are purchased through vendors and others we build and maintain in-house. The following are a few of our major data products:

StreetsSD

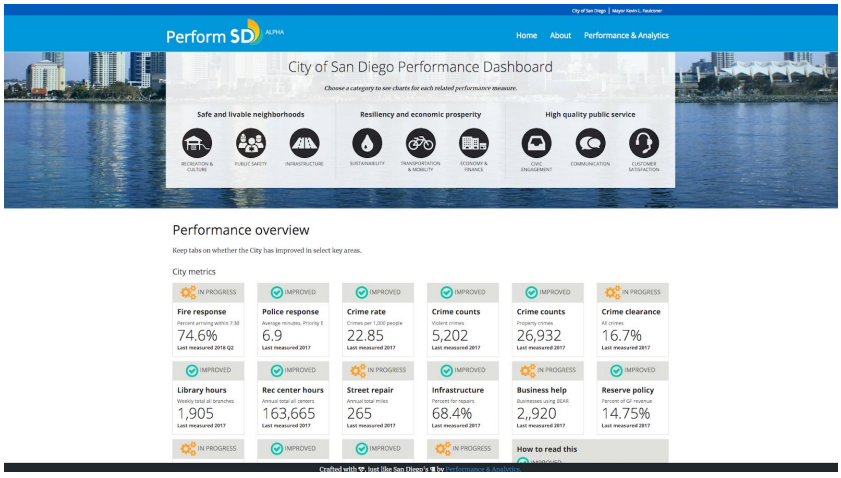


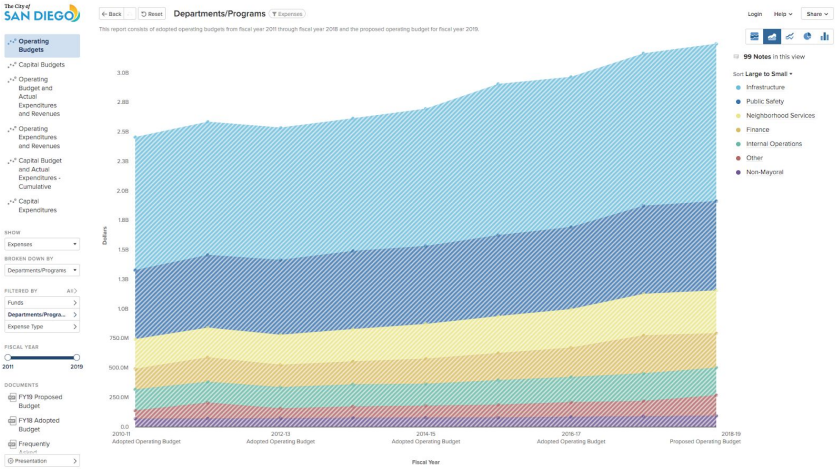
StreetsSD is a web map that highlights City streets that have been repaired since July 1, 2013 and those that are currently scheduled for repair. The Overall Condition Index for every street is also available as a layer.

This map is powered by data we extract directly from the operations department's database. Because we connect directly, we can send alerts and generate reports, which saves many hours of staff time.

PerformSD

PerformSD displays charts and other infographics showing City performance data. This online dashboard is organized around the City's Strategic Plan, Key Performance Indicators in the fiscal year budget, individual department goals, and inquiries from the public and decision-makers. These indicators cover a variety of areas, including recreation and culture, public safety, infrastructure, sustainability, transportation and mobility, economy and finance, civic engagement, communication and customer service.





Budget.sandiego.gov

Operating and capital expenses and revenue are displayed on this online, interactive version of our fiscal year budget. Viewers can see dollar amounts by department/program, fund type, or expense/revenue type and export data in a *.csv* format.

Looking ahead to FY 2019

Next steps in this area include: improving the design and functionality of existing data products; automating data collection and reporting for more of the indicators displayed on performance.sandiego.gov; publishing data behind performance metrics on the open data portal; and prototyping additional solutions with departments who have previously reached out for help.

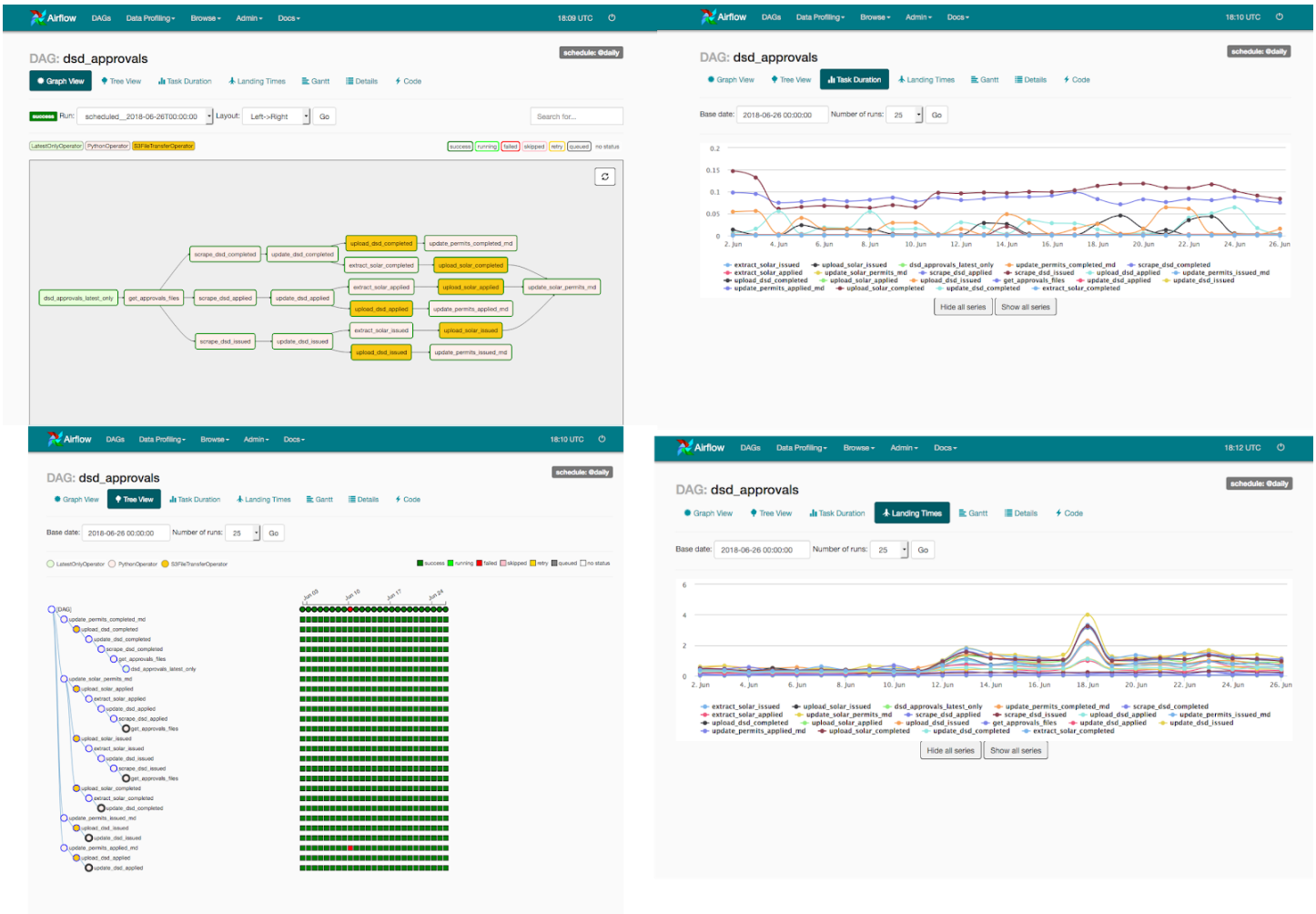
DATA INFRASTRUCTURE

The City has hundreds of primary data sources. We use hundreds of different technologies and data storage systems. We integrate with multiple vendors to move data in and out of our systems.

Current State

Poseidon

DataSD maintains Poseidon, a key piece of internal data infrastructure that connects to many of these systems. Poseidon enables all of our data products, the open data portal, our analytics projects and alerting. Poseidon consists of multiple jobs that continuously need updates and maintenance, since the City's systems and data are continually in flux.



Data Ingestion

DataSD provides FTP (File Transfer Protocol) locations for SDPD (San Diego Police Department) to upload their files, and for Poseidon to ingest them into the data portal. We provide FTP locations for vendors to upload data to our systems as well. Lastly, we provide and maintain multiple endpoints for vendors to send data to the City programmatically in close to real time.

Planned State

DataSD will continue to maintain Poseidon and the reliant Extract / Transform Load Pipelines. In the near future, we need to perform a significant version upgrade to Poseidon to increase system stability and security.

The next step is to augment the Poseidon system to normalize the data we currently bring in and move it to a central data storage location with granular access control. This will significantly assist in our analytics enablement practice across the organization.

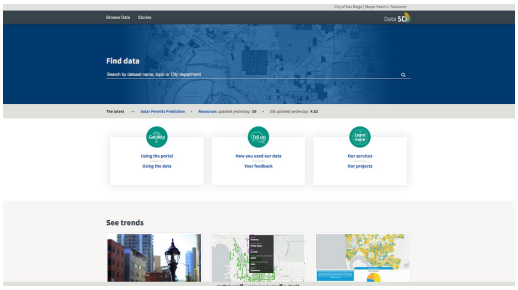
ANALYTICS ENABLEMENT / ALERTING

Analytics enablement is the work that DataSD does to empower employees across the organization to use the City's data to make decisions. We provide City staff with ways to discover data and understand the caveats and the field meanings. We also work to make sure the quality is high, and the data is updated as frequently as possible.

As part of our analytics enablement work, we provide products like StreetsSD to make data analysis and decision-making easier for common use cases.

Current State

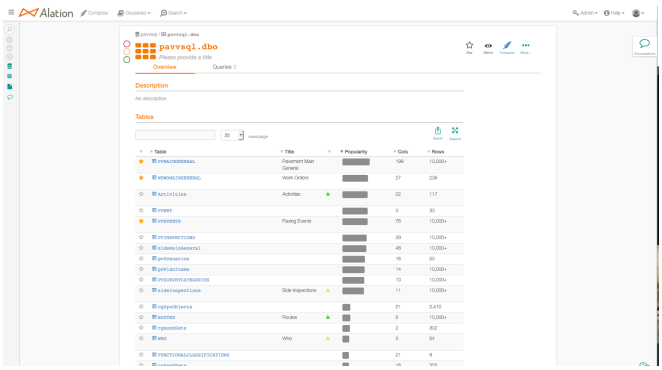
External Data Catalog (data.sandiego.gov)



DataSD facilitates data discovery and availability across the organization by providing data on data.sandiego.gov. When we integrate with a data source, we dive deep into department operations to truly understand the intricacies and caveats of the data. We use that knowledge to build our Poseidon Pipelines, populate data dictionaries (also available on the Portal), and perform analytics projects as needed.

*DataSD Owner**: Andrell B.***
*Status: **Support & Improve***

Internal Data Cataloguing Tool (Alation)



In an attempt to maintain information about the data that we have in a dynamic way, we maintain a data cataloguing tool. Alation monitors the City's databases, documenting the data in them. Alation is still in a prototype stage; once we have a centralized data storage location, we anticipate layering Alation on that.

*DataSD Owner**: Maksim P.***
*Status: **Support & Improve***

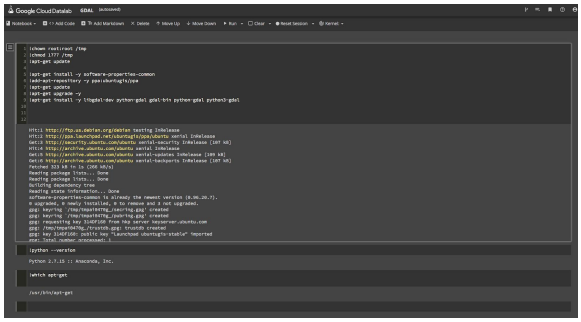
Data Analysis (PandaPack)

PandaPack is a set of technologies DataSD has developed to facilitate data analysis for City employees who only have access to limited systems. We use these technologies to enable people across the city to use tools like Python and QGIS to perform their analyses. Using this technology we have enabled Environmental Services Department staff to run python for solar panel placement analysis, and empowered our internal staff to work on PerformSD.

*DataSD Owner**: Maksim P. (Initial development by Arnaud V.)***
*Status: **Support & Improve***

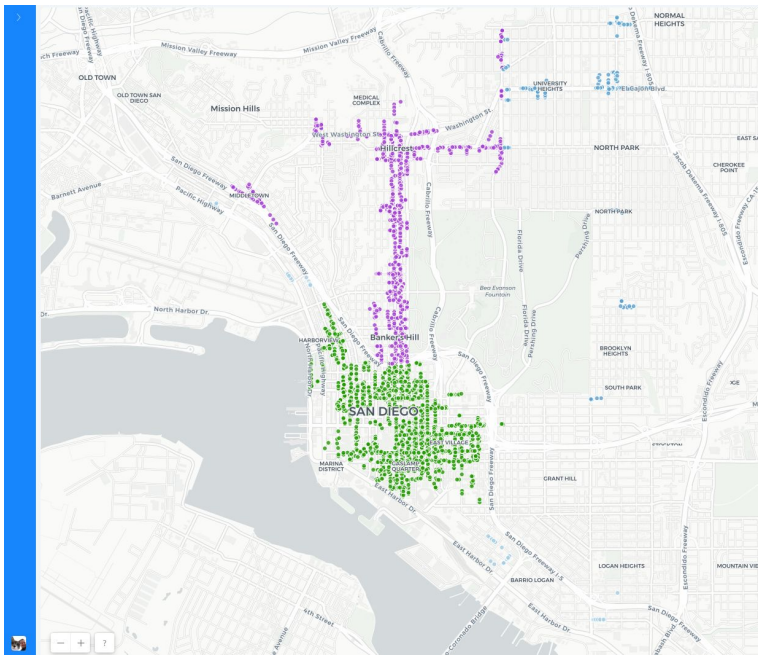
Data Analysis (DataLab)

DataLab is a cloud-based suite of tools that DataSD uses to analyze large datasets that cannot be run on our local machines. We have not provided this to any employees yet, but as we develop capacity, we expect to be able to.



*DataSD Owner**: Maksim P.***
*Status: **Active Development***

Data Analysis (Carto)



To empower employees across the city to do geospatial analytics, mapping and visualization in a simple manner, DataSD purchases and manages 17 Carto licenses that are used by departments across the organization. Carto provides staff with a simple GIS platform where they can visualize geospatial data, perform advanced geospatial operations, and publish the results to the City's web site or the data portal. A lot of the mapping projects we discuss below are using Carto as a platform.

*DataSD Owner**: Jon T.***
*Status: **Support & Maintain***

Planned State

DataSD will continue to run the data portal and data catalogs. In the coming year, we hope to expand Carto and DataLab licensing and usage throughout City departments and align it with our training program.

ANALYTICS PROJECTS

We regularly scope and engage in analytics projects to help departments streamline operations or provide access to data in a way that was impossible before.

Current State

We have a variety of supporting infrastructure that we have developed around performing analytics projects. DataLab and PandaPack are some of the front-facing examples. We also have worked with geocoders, advanced mapping applications, various pieces of Amazon Web Services and Machine Learning pipelines. This infrastructure allows us to execute quickly on projects with large datasets and advanced processing requirements.

Planned State

DataSD strongly believes in solving a problem once and replicating it across multiple projects. We will continue to develop and expand our data science infrastructure to successfully execute on more advanced analytics projects.

CONSULTING

DataSD often gets brought in to consult other departments on technology selection or purchasing. Sometimes we will also come in with a technological solution that is cheaper or simpler than the enterprise vendor offerings. For example, we worked with the Treasurer to provide them an address verification solution that is ⅓ of the cost of the one provided by our consultants.

Current State

Our consulting practice revolves around the Scope of Work. We work with the department stakeholders to develop a scope of work that we can hold each other to. From there, we work with the department to either consult on picking a vendor, solve a problem, or deliver a solution.

Planned State

The engagement process has been working extremely well, but can always be improved. We are also looking forward to partnering with Operational Excellence, Performance Management, Citywide Services and Get It Done to source and collaborate on projects.

PROJECTS

Our projects often fall into multiple categories of our practice. We have a wide range of projects that we engage with - from simple to very complex. Every project has a DataSD owner - who is the lead on the project. This person works closely with the Dept Stakeholder to execute on the project and pulls in additional PandA or department resources as needed.

Project Status

We classify projects in one of these statuses:

Research
We are working with the department to determine the feasibility, usefulness and impact of a project.

Active Development
A scope of work has been signed and the project is currently being worked on by the team.

On Hold
We are currently not working on this project due to capacity issues.

Support & Development
The project is complete and currently in production, but we're still working on it and adding features.

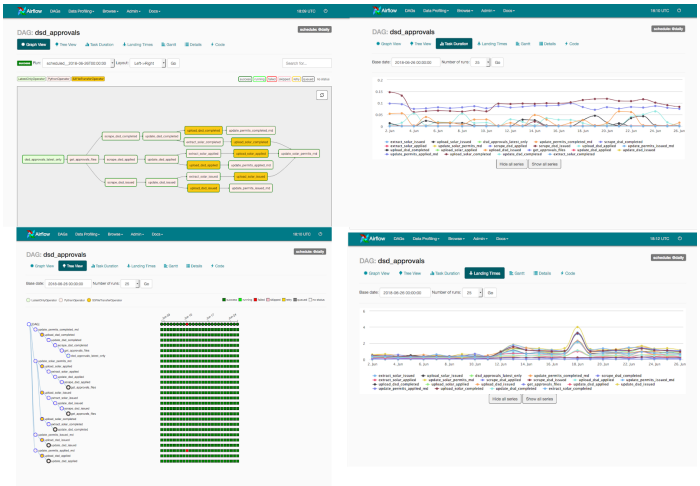
Support & Maintenance
The project is complete and currently in production. We are not adding features, but we are resolving bugs and handling issues as they come up.

Successful Handoff
We have completed the project, successfully handed it off to the owning department, and are no longer involved with it. We sometimes circle back on these to see if there are additional things we can do.

Project Overview

Poseidon

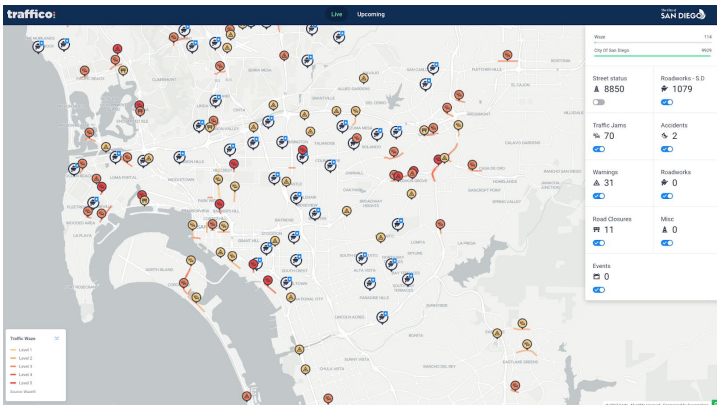
DataSD maintains Poseidon, a key piece of internal data infrastructure that connects to many of these systems. Poseidon enables all of our data products, the open data portal, our analytics projects and alerting. Poseidon consists of multiple jobs that continuously need updates and maintenance, since the City's systems and data are continually in flux.



DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Internal***
Status: ***Support and Development***
Tags: ***Data Products, Infrastructure, Analytics Enablement / Alerting, Analytics, Consulting.***

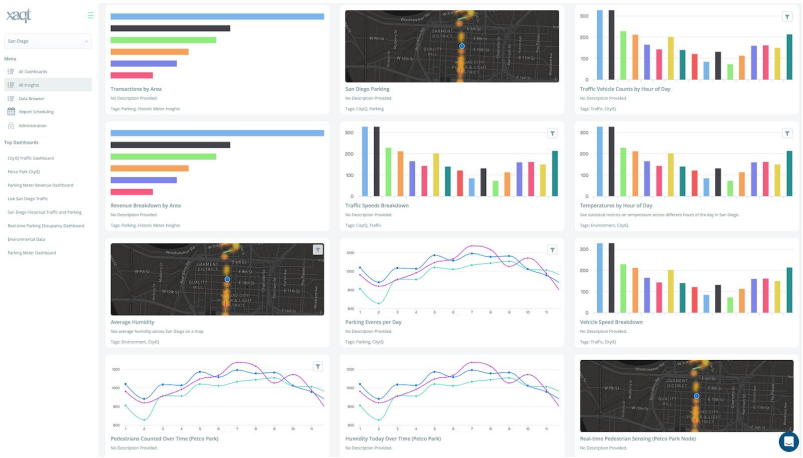
Carto + Waze

We are working with Carto to bring real-time road status, traffic light location and status, congestion information, incident information, and other key data points to improve traffic management at the traffic operations center.



DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Duncan H, Phil R.***
Status: ***Active Development***
Tags: ***Data Products, Analytics Enablement, Consulting***

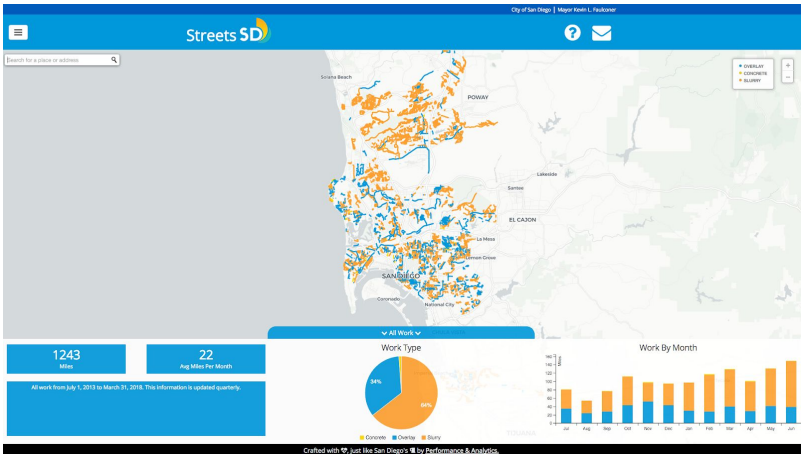
Smart Cities + Xaqt (Ongoing)



CityIQ data is currently only available via APIs, which are hard to work with even if you are a software engineer. We are working with Xaqt and ESD to determine use cases for centralizing the CityIQ data and creating interfaces for access and analysis.

DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Lorie Cosio Azar***
Status: ***Active Development***
Tags: ***Data Products, Analytics Enablement, Consulting***

Streets SD (Maintaining)



StreetsSD initially started off as a visualization tool. However, departments across the city now use it to view OCI conditions and plan street paving.

DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Joshua L, Chris H.***
Status: ***Support and Development***
Tags: ***Data Products, Analytics Enablement, Consulting, Analytics, Infrastructure***

Truck Routing

As part of our collaboration with the OpEx team, we have optimized routes for the City's delivery trucks that move across multiple city facilities.

DataSD Owner: ***Maksim P. (Initial work by Arnaud V).***
Dept Stakeholder: ***Thomas S.***
Status: ***Support & Develop***
Tags: ***Consulting, Analytics***

CIP/BID Performance Metrics

The Economic Development Department manages BIDs, or Business Improvement Districts. EDD identified a number of performance metrics they wanted to track per BID, including number of public improvement projects occurring in a BID and the amount of money spent on those projects. We helped EDD find and analyze the data about public improvement projects.

*DataSD Owner**: Andrell B.***
*Dept Stakeholder**: Liz S.***
*Status: **Successful Handoff***
*Tags: **Analytics, Consulting***

DSD/NETDU

To answer the question of how many dwelling units have been or will be built in San Diego, we helped extract data on construction permits from Development Services, and analyzed the data to determine net change in number of dwelling units.

*DataSD Owner**: Andrell B.***
*Dept Stakeholder**: Karen D.***
*Status: **Successful Handoff***
*Tags: **Analytics, Consulting***

Library Donation Fund Distribution Model

We developed a data model that is used by the Library Department to analyze various data sets for the recently approved Council Policy revision on library donation funds. This model incorporates demographics such as population and median household income to facilitate data-driven accounting by library service area. On an as-needed basis, the Library Department will update actual and historical information for all library locations used in the pool distribution model. Data sources include, but are not limited to:

- United States Census Bureau
- San Diego Association of Governments (SANDAG)
- GALE - Analytics on Demand

*DataSD Owner**: Jon T.***
*Dept Stakeholder**: Misty J.***
*Status: **Successful Handoff***
*Tags: **Analytics, Consulting***

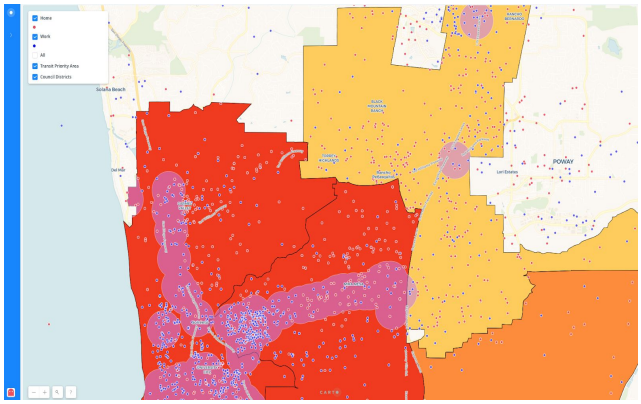
Commuter Survey

DataSD provided the development work and hosting for the first annual SD Commuter survey, which surveyed more than 4,000 people during February and March**.** The survey collected data about residents' commuting habits to help the City develop strategies for its Climate Action Plan.

*DataSD Owner****: Jon T., Andrell B.***
*Dept Stakeholder****: Cody H.***
*Status****: Successful Handoff***
Tags: ***Consulting***

Commuter Survey Mapping and Analysis

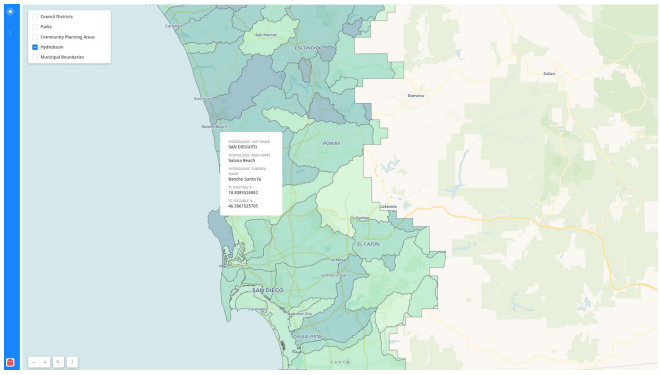
DataSD provided the visualization and analysis work for the Commuter Survey to present results to the Sustainability team on the current state of commuting in San Diego.



*DataSD Owner****: Jon T.***
*Dept Stakeholder****: Cody H.***
*Status****: Successful Handoff***
Tags: ***Analytics***

Tree Canopy Mapping and Analysis

DataSD provided visualization, data processing and analysis work to present results to the sustainability team on the current tree canopy status in San Diego.



*DataSD Owner****: Jon T.***
*Dept Stakeholder****: Cody H.***
*Status****: Successful Handoff***
Tags: ***Analytics***

Community Policing Event Tracking Tool

We are currently working on providing the Police Department with a tool to more consistently and efficiently track, broadcast, and analyze the community outreach work of their Community Relations Officers. This project is currently underway.

DataSD Owner: ***Jon T.***
Dept Stakeholder: ***Brent W.***
Status: ***Active Development***
Tags: ***Consulting***

DSD Permit Performance

We worked with the Development Services Department to analyze permit processing times since as early as 2003.

DataSD Owner: ***Maksim P (initial development by Arnaud V).***
Dept Stakeholder: ***David S.***
Status: ***Successful Handoff***
Tags: ***Analytics***

Real Estate Assets Property Database

We are working with READ to evaluate which of their properties are potential targets for sale to developers to build affordable housing.

DataSD Owner: ***Maksim P***
Dept Stakeholder: ***Mary C.***
Status: ***Research***
Tags: ***Analytics, Infrastructure***

RWQB Stormwater Settlement Solution

To comply with the Regional Water Quality Board's permitting requirements, the City needs to aggregate and make searchable inspection information across separate Public Works, Transportation & Stormwater, and Development Services project management systems. This project is currently underway.

*DataSD Owner**: Maksim P., Jon T.***
*Dept Stakeholder**: Alex H.***
*Status: **Active Development***
*Tags: **Consulting, Infrastructure***

Parking Meter Data Uptake

We provide a receiver endpoint to which IPS (our parking meter vendor) sends payloads when a transaction on a parking meter happens. This data then gets logged and stored. This lets us update the parking meters dataset on the portal, and provide near-real-time intelligence on transactions in the 5000+ parking meters in the city.

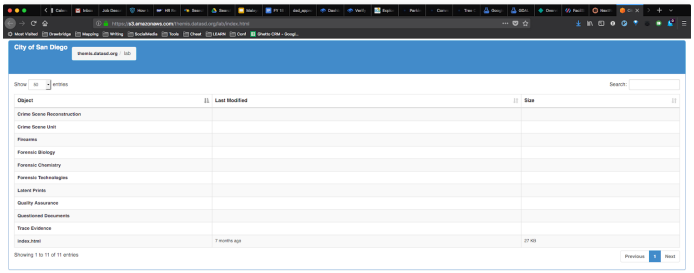
*DataSD Owner**: Maksim P.***
*Dept Stakeholder**: Jonathan C.***
*Status: **On Hold***
*Tags**: Infrastructure, Analytics Enablement / Alerting, Analytics, Consulting.***

FTP Server

DataSD runs an FTP server for external vendors to upload data to us.

*DataSD Owner**: Maksim P.***
*Dept Stakeholder**: Chris H, external vendors.***
*Status: **Support and Maintenance***
*Tags**: Infrastructure***

PD Lab File Uploads (Lab Documents)



The police department approached us for help uploading a few thousand files to the City of San Diego website. The current website architecture only allows one-by-one file uploads. We used our existing Amazon Web Services infrastructure to deliver them an embeddable, maintainable solution at a cost of less than $10 per month.

*DataSD Owner**: Maksim P.***
*Dept Stakeholder**: Danny P.***
Status: ***Support and Maintenance***
*Tags**: Infrastructure, Consulting.***

PD Lab File Uploads (Procedures)

The police department approached us for help uploading a few thousand files to the City of San Diego website. The current website architecture only allows one-by-one file uploads. We used our existing Amazon Web Services infrastructure to deliver them an embeddable, maintainable solution at a cost of less than $10 per month.

*DataSD Owner**: Maksim P.***
*Dept Stakeholder**: Danny P.***
Status: ***Support and Maintenance***
*Tags**: Infrastructure, Consulting.***

Public Art Data Ingestion

The Arts and Culture Commission curates a collection of public art around San Diego. The Commission wanted to make it easier for residents to browse the collection online, so they are launching a cloud-hosted collection management system with a public URL, where each piece of artwork has a landing page with details and photos. We are using the API for this system to get artwork details, including exact locations, and created a public art dataset and a map that shows where each piece is located.

*DataSD Owner**: Andrell B.***
*Dept Stakeholder**: Elise C***
Status: ***Support and Development***
*Tags**: Infrastructure, Consulting, Analytics Enablement / Alerting***

Get It Done Image Storage Support

Get It Done (GID) needs a place to store the images uploaded by users and display them in Salesforce and on the GID user-facing app and website. The GID team was considering using OpenText, which would have come with significant licensing fees and integration costs. DataSD worked with implementation vendor Deloitte to provide Get It Done with scalable image storage infrastructure on Amazon Web Services at a significantly-reduced cost to the City.

DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Alex H.***
Status: ***Support and Maintenance***
Tags: ***Infrastructure, Consulting***

TSW / Riva Sync Support

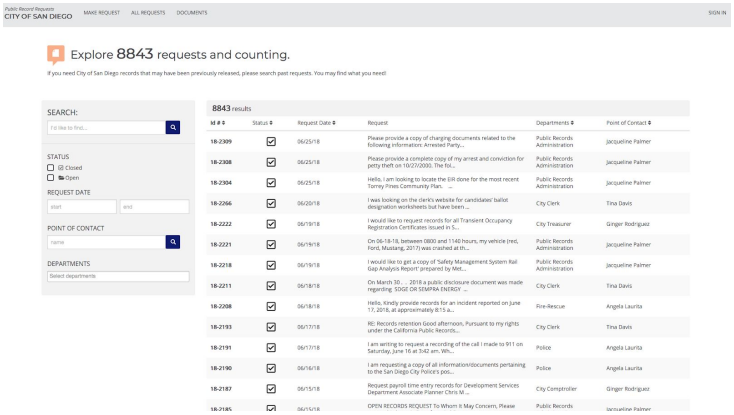
TSW needed to synchronize data from Cartegraph to the Riva system that is part of the IAM upgrade. As the team with the deepest knowledge of that datasource, DataSD provided expertise to write the necessary query for proper data integration between the two systems.

DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Chris H.***
Status: ***Successful Handoff***
Tags: ***Consulting***

Carto / Promise Zone / Sustainability

The Promise Zone program has many performance metrics, and staff needed simple analytics tools that would let them mix and view different datasets. We provided them training on Carto and Tableau to enable their analytics work.

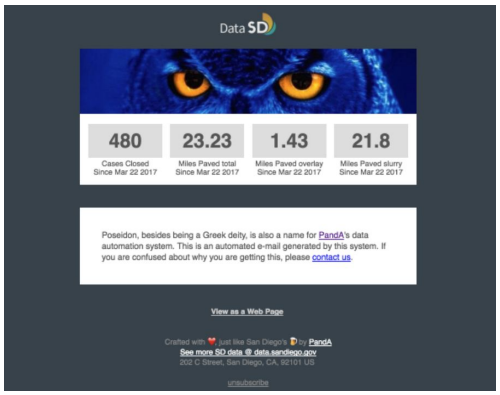
DataSD Owner: ***Maksim P.***
Dept Stakeholder: ***Catherine C.***
Status: ***Support & Maintenance***
Tags: ***Data Products, Analytics Enablement, Consulting, Analytics, Infrastructure***



NextRequest

DataSD continuously monitors NextRequest, a site where people can submit online Public Records Act requests to the City. By looking at requests, we identify data that is in high demand. This way, we can prioritize that data for release and decrease staff time spent responding to requests. For example, we worked with PD to automate the release of the Calls For Service data because of such high request volumes on NextRequest.

*DataSD Owner**: Maksim P., Andrell B.***
*Dept Stakeholder**: Jacqueline R.***
*Status: **Successful Handoff***
*Tags**: Consulting***

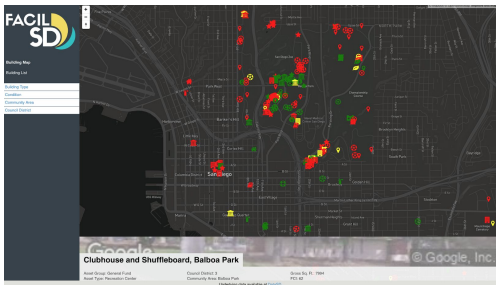


Sonar

Sonar is a prototype of a conditional KPI alerting interface. It can be programmed to send people different alerts based on the severity of the deviation.

*DataSD Owner**: Maksim P.***
*Dept Stakeholder**: Internal / Infrastructure***
*Status: **On Hold***
*Tags: **Data Products, Analytics Enablement, Consulting, Analytics, Infrastructure***

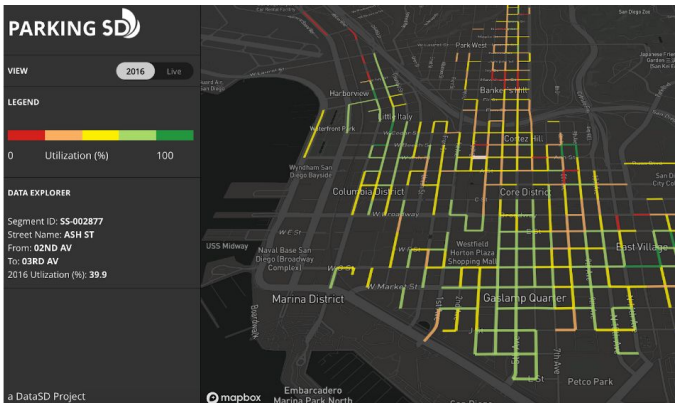
Facilities Condition Index



Public Works approached us to develop a map and web view of the Facilities Condition Index. The index rates various City buildings on a number of criteria and gives an idea of how much the City can expect to pay in upgrade and maintenance costs over the remaining life of the building.

*DataSD Owner**: Maksim P., Andrell B.***
*Dept Stakeholder**: Nevien A.***
*Status: **On Hold***
*Tags: **Data Products, Analytics Enablement, Consulting***

Parking Meter Utilization Based on Transaction



We used the parking meter transaction data published on the open data portal to find street segments that had the highest and lowest parking meter utilization in the city.

*DataSD Owner**: Maksim P. (Initial work by Arnaud V).***
*Dept Stakeholder**: Jonathan C.***
*Status: **Support & Develop***
*Tags: **Analytics, Data Products***

Parking Meter Utilization Based on Transaction / Meter Sensor / CityIQ Sensor (Ongoing)

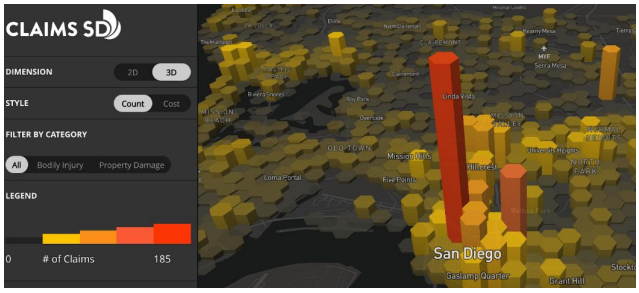
In order to understand accuracy of the multitude of data sources we have about parking utilization, we are in the process of designing an experiment to monitor several spots where all 3 technologies intersect. Then we will decide on error rates and deviations.

*DataSD Owner**: Maksim P. (Initial work by Arnaud V).***
*Dept Stakeholder**: Erik C.***
*Status: **Research***
*Tags: **Analytics, Consulting***

CityIQ Repositioning

We helped to reprioritize the position of the CityIQ sensors as they are going up around the city to be more effective for our Vision Zero work.

*DataSD Owner**: Maksim P. (Initial work by Arnaud V).***
*Dept Stakeholder**: Lorie C.***
*Status: **Successful Handoff***
*Tags: **Analytics, Consulting***



Risk Mapping

We analyzed claims cases across the city to find areas where there are highest occurrences of different types of claims cases.

*DataSD Owner**: Maksim P. (Initial work by Arnaud V.)***
*Dept Stakeholder**: Julio C.***
*Status: **On Hold***
*Tags: **Analytics, Consulting, Data Products***

EMS

To help make the City's dispatching for emergency medical services more efficient, the Fire/EMS Department partnered with doctors at the UCSD Department of Emergency Medicine to review dispatch codes. The codes are assigned to different categories of medical emergencies and control how quickly emergency vehicles are dispatched as well as the number and type of vehicle. We built a dataset to help the doctors and EMS staff compare the dispatch plan with the types of treatments required on-scene.

*DataSD Owner**: Andrell B.***
*Dept Stakeholder**: Anne J.***
*Status: **Successful Handoff***
*Tags: **Analytics, Consulting, Analytics Enablement***

Water Billing

The goal of this project is to analyze and explore trends in the City of San Diego water billing data to find anomalies within the billing system.

*DataSD Owner**: Andrell B.***
*Dept Stakeholder**: Vic B.***
*Status: **Active Development***
*Tags: **Analytics***

GID Proximity Search

Get It Done asked us to prototype a solution for searching the proximity for new cases as they are submitted. We developed the prototype quickly and at almost no cost, which saved the City significant cost, compared to a consultant's proposal.

*DataSD Owner**: Maksim P.***
*Dept Stakeholder**: Alex H.***
*Status**: Successful Handoff***
*Tags: **Consulting***

Insight

Two fellows from the Insight Data Science Fellows Program partnered with Performance & Analytics to analyze 911 calls for emergency medical services.

*DataSD Owner**: Andrell B.***
*Dept Stakeholder**: Anne J.***
*Status: **Successful Handoff***
*Tags: **Consulting***

COMMUNITY INVOLVEMENT

DataSD works extensively with regional community organizations. Below are highlighted events from the past fiscal year:

USD Training
Arnaud Vedy conducted a training for Bangladeshi researchers on geospatial data analysis

Community Lidar and Geospatial Trainings
Arnaud Vedy conducted various lidar and geospatial trainings with Open San Diego and ScaleSD.

Civic Analytics Network
City of San Diego is one of the inaugural members of Harvard's Civic Analytics Network, which shares our knowledge and expertise with cities across the country.

SDXD Emergency Response Hackathon
DataSD worked with San Diego UX Design Group along with Homeland Security and the Fire Department to put together the first ever emergency response design-a-thon.

San Diego Public Library Open Data Training
DataSD worked with the Library to offer an open data training that was run by the Library and used the data from the data portal.

ScaleSD
Scale is one of our partners that operates from Downtown Works, and puts on panels, discussions and trainings around City data. DataSD works regularly with this organization to host events with the community.

AWARDS

What Works Cities
Bloomberg Philanthropies' *What Works Cities* awarded the City of San Diego a Silver certification, largely due to our use of data and analytics.

Equipt To Innovate
Governing recognized the City of San Diego as the most data driven city in the country for our work in data management and automation.

Top 50 Data & Analytics Professionals
Corinium Intelligence recognized the CDO as one of the top 50 Data & Analytics professionals in North America. This is a strong reflection on the team and its accomplishments.

OPEN SOURCE

DataSD releases a majority of its code into the Open Source domain. Our notable open source projects include:

- The Open Data Portal (seaboard)
- StreetsSD
- PerformSD
- QA Map GL
- SD-Docker-Airflow

OUR FUTURE

DataSD has developed multiple relationships with UCSD, the Port of San Diego, and the County of San Diego. We look to build on these and create further regional data partnerships.

We will continue to enable and uplift staff citywide -- helping them be more advanced and more effective users of data. We will be partnering with OpEx to provide a data training component.

In the next year, we hope to focus deeper on our infrastructure. We have put a lot of focus into analytics projects this year, allowing our infrastructure work to lag behind a bit. However, since our analytics projects are impossible without a solid infrastructure supporting them, we plan to take on fewer analytics projects while we stabilize our infrastructure.

We also hope to determine the best way for vendors and IT to collaborate on helping us support our products and infrastructure. This will require some work and planning, but in the long term it will help us offset the temporal cost of support.

Lastly, we do consider City staff to be part of building strong data infrastructure in the city. We will be focusing on providing more data enablement and alerting tools to staff.