COMMITEE AGENDA: 04/07/2022
FILE: CC 22-084
ITEM: (d)3.

CITY OF SAN JOSE
CAPITAL OF SILICON VALLEY

Memorandum

TO: SMART CITIES AND SERVICE IMPROVEMENTS COMMITTEE

FROM: Rob Lloyd

SUBJECT: **DATA INITIATIVES AND IMPACTS STATUS REPORT**

DATE: March 28, 2022

Approved  Date 3/30/2022

RECOMMENDATION

Accept the Data Initiatives and Impacts status report on work related to the data and analytics program, including use of data to improve services, measures for impact, support for departmental equity work, and capacity-building partnerships through 2022.

OUTCOME

The Committee will be aware of and provide feedback on the Data Initiatives and Impacts as presented by Department of Public Works (DPW), Department of Transportation (DOT), and Information Technology Department (ITD).

BACKGROUND

Data science remains one of the fastest growing segments of the technology sector, enabling applied use of data patterns and insights in all facets of customer service and service delivery.[1] The City of San José has recognized the need to harness data for effective, equitable, and efficient service delivery, but with dedicated precautions to ensure uses maintain the trust of the San Jose community.

The City of San José released its Open Data Policy in 2016[2]. City departments were committed to releasing public data sets via the City Open Data Portal[3] and Geospatial Information System

[1] Technavio (2021), *Data Science Platform Market by Component and Geography*: https://www.technavio.com/report/data-science-platform-market-industry-analysis
[2] City Policy 1.7.10: https://www.sanjoseca.gov/home/showpublisheddocument/17947/637141737451600000
[3] City of San Jose Open Data Portal: http://data.sanjoseca.gov

(GIS) portal[4] to maximize interoperability and transparency. The Open Data Portal currently maintains 184 open datasets, of which 145 are spatial datasets to provide a common means for the public to access information in and about the City. Datasets are shared with Creative Commons Zero licensing, making it freely downloadable for anyone to use as part of the public domain. Open Data is most often used by City staff while executing work, followed by City partners, the media, members of the academic and education community, community stakeholders, and general public. The Open Data Portal currently averages 80 active users per day with the following top five (5) most popular datasets:

1) Police Calls for Service[5]
2) Employee Compensation[6]
3) San José Fire Incidents[7]
4) Crash Data[8]
5) COVID-19 Food Distribution[9]

This pattern highlights public interest in and the value of ready-to-use, high quality data. The investment in a safe Open Data Portal was a key step in the City's efforts to make information on the City's activities transparent and accessible, while fueling process improvements both internally and with external partners. The wholistic approach activated data across departments to strengthen its City services along the lines of operations, equity, and other measures.

The City structures data discussions along concepts of data journalism for maximum engagement. Presentation modes span:

1) **Data Articles** – Small, cogent narratives extracted from the data. A data article provides a summarized view of a specific topic and is accessible by the majority of the City's public audience, which is more interested in a high-level understanding of specific topics. Data Articles are usually a few paragraphs long with 1-2 key data attributes with visualizations.
2) **Data Stories** – Data stories are detailed and thorough analyses of data in a topic area, created by and reviewed by experts. It is a combination of narrative, hard data, and visualizations in the form of an extended write-up. Data stories are targeted toward the partners, researchers, journalists, elected officials, and similar audiences that wish to gain some functional expertise in the topic areas. Members of this focused audience reach out to City staff more regularly.
3) **Raw Data** – Data in raw form are very useful for analysts, researchers, data scientists, and other deep data experts who are interested in working with data directly and possibly combining it with other datasets for analysis. This audience and use are best for multi-

[4] City of San José GIS Portal: http://csj.maps.arcgis.com
[5] City of San José Police Calls for Service: https://data.sanjoseca.gov/dataset/police-calls-for-service
[6] City of San José Employee Compensation: https://data.sanjoseca.gov/dataset/employee-compensation-plan
[7] City of San José San José Fire Incidents: https://data.sanjoseca.gov/dataset/san-jose-fire-incidents
[8] City of San José Crash Data: https://data.sanjoseca.gov/dataset/crashes-data
[9] City of San José COVID-19 Food Distribution: https://data.sanjoseca.gov/dataset/food-distribution

jurisdictional efforts and has been used by "Civic Hackers" and community data analysts on topics such as transportation safety, citations activity, and City spending.

Since 2019, the City has published six (6) showcases Data Article + Story + Data combinations via the Open Data Portal. This work has won or directly contributed to accolades including an American Planning Association Award of Merit for the Data Science behind Traffic Accidents and Fatalities[10]; the City twice placing as a Bloomberg What Works City[11]; and an IDC Smart 50 Award for Geo-Award Emergency Vehicle Preemption[12].

These successes are indicators of how the City's data initiatives have grown and are consistently multi-departmental when successful. Between 2017 and today, support for data initiatives have spanned the City Roadmap's Drive to Digital, City Workforce, and Digital Equity priorities; Vision Zero efforts; San José Smart City Vision initiatives; the spatial data that was key to completing the 2021 Redistricting[13]; 2020 Census Response efforts; and more. Just as important, in the challenges of COVID response and recovery, City staffs generated information from City data to guide status reporting, help shape communications and outreach efforts, decide Digital Inclusion investments, inform policy decisions, and drive resource allocation.

This memo highlights and builds on previous reports on specific initiatives. By focusing on capacity-building and cross-departmental coordination structures created by the Information Technology, Public Works, and Transportation departments, in partnership with the Mayor's Office of Technology and Innovation, the City's data efforts begin to transition from periodic to reproducible and consistent. This approach aims to maximize the following:

1. Improving City services;
2. Measuring community impact;
3. Supporting equity; and
4. Building community partnerships.

Building the foundations for City Data Efforts

A primary objective of the City's data efforts has been the development of highly usable data platforms that provide quality information of frequent need and use, and in a mode that is highly accessible to Citywide users and the public. With support in the annual City Budget Processes since 2016, staffs built an extensive geospatial data repository and implemented the City Open Data Environment (CODE) for analytics uses. Both are actively utilized by departments to store and process information. Recent focus has centered on growing the City's ability to make

[10] APA Award Merit: https://urbanlogiq.com/news/city-of-san-jose-department-of-transportation-receives-apa-smart-city-merit-award-for-data-driven-innovation/
[11] Bloomberg Silver Certification: https://medium.com/what-works-cities-certification/harnessing-silicon-valleys-genius-in-san-jose-96dfa76729a8
[12] IDC Smart 50 Award: https://spring.smartcitiesconnect.org/Smart50Awards/2020AwardRecipients.html
[13] San José Redistricting Portal: https://redistricting-csj.hub.arcgis.com/

analytical use of these data through standard tools, verification processes, and generating insightful visualizations, all to increase service efficiencies and to inform decision-making.

For analytics, applications, and spatially-enabled business systems, the City has adopted two core technical tools to help meet the growing and ongoing demand for data visualization and analytics. Supported by the City's Enterprise Geospatial Information Systems (GIS) team, these are:

1. The **Enterprise Extract-Transform-Load (ETL) Platform** – Provides a framework for integration of disparate data from various sources into a central repository, which receives timely updates in a consistent structure that feeds into multiple applications and dashboards. A core component of this is a standardized set of demographic data used by a growing number of departments for equity-based analyses.
2. A **Common Toolset** – Tools include ArcGIS, Tableau, and Power BI. These help staff process data for decisions, as well as create data visualization, dashboards, and analytical products that are accessible across business areas. Where these products are accessible and easy-to-use, staffs see maximum use and demand for further insights and drilldown capabilities.

ANALYSIS

Data and analytics investment have been at the heart of the work City Council and Council Committees have seen from staff in COVID, budget, Digital Inclusion, project, and status reporting work. The demand for geospatial and data analytics tools and support also continues to increase across City divisions and departments. Transitioning from anecdotal and intuition-based decisions to setting *data-driven direction* was identified as a key in the Equity Solutions pillar of the 2021-2023 IT Strategic Plan. This was based on the frequency of the theme in technology planning discussions with all City departments and offices.

Examples of initiatives driving new data and analytics capabilities are covered in the following sections. Critical in these efforts are the ability to provide consolidated, centralized platforms.

Improving City Services

Fuel Monitor Dashboard – Fleet Division, Department of Public Works

Police, Fire, and other City services rely on up-to-date information on fueling of their fleet and an evolving strategy for refueling. The Fuel Monitor Dashboard (FMD) supports this by providing a central interface to track fuel station outages, network communication issues, and current fuel levels for every tank managed by the City. Merging data from various sources into a single view, the FMD has become critical to running an efficient and effective Citywide fleet and ensuring continued fuel supply at fuel sites. In the Public Safety Power Shutoff emergency, the importance of this information and management platform was distinct.

The FMD is easy to use and staff have been able to adopt it quickly, with little to no training. It is accessed through a browser and available on any desktop or mobile device.

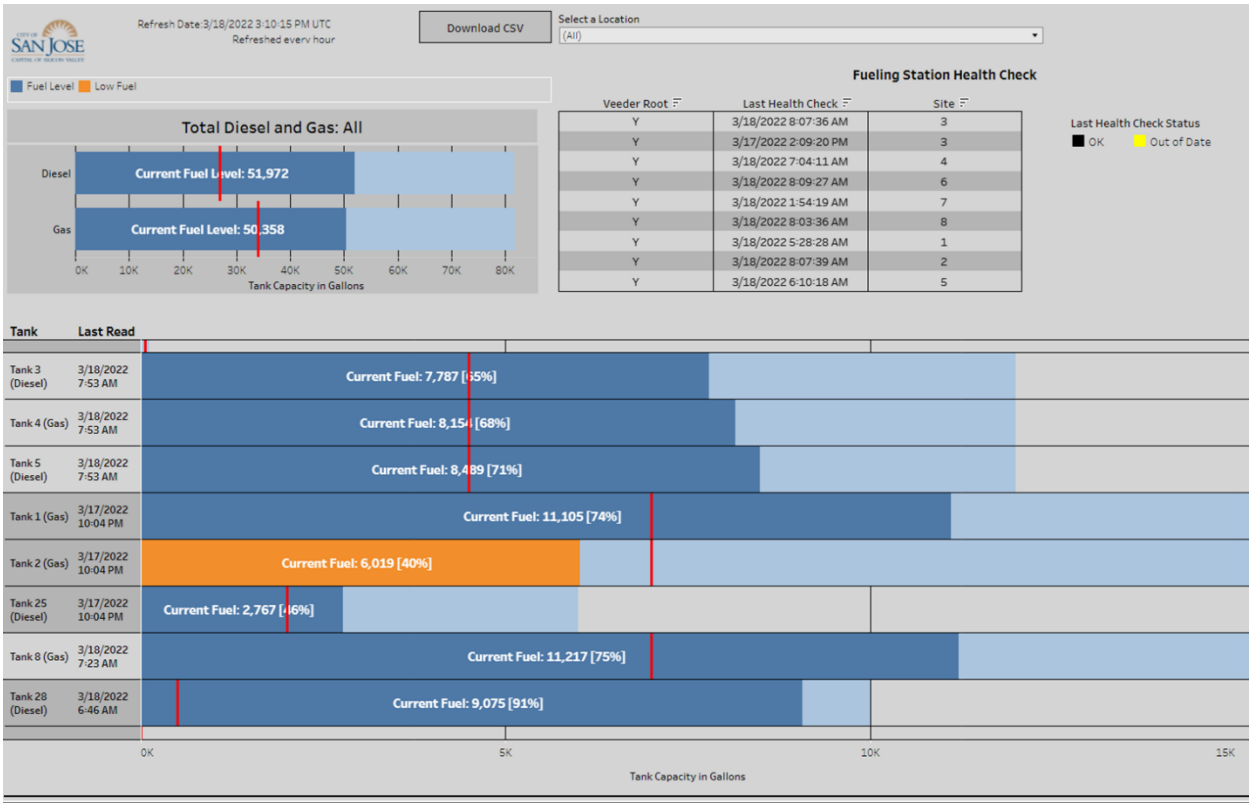

Figure 1: Fuel Monitor Dashboard displays current fuel level of fuel stations across the City

Prior to implementation of the FMD, staff relied on manual methods to track fuel usage and intuition to estimate fuel needs. With access to accurate, current data presented in easily understandable graphics, fuel resourcing and response times have improved, and staff time involved in the fuel management process has been reduced. The FMD is estimated to save 1.0 full-time equivalent (FTE) employee annually that was formerly dedicated to site visits and manual checking. Response times to outages have been reduced from approximately three days to real-time. More importantly, City departments have access to reliable fuel supplies in locations where they need it, reducing refueling time and increasing in-service time.

Animal Care Services Dispatch Map and Dashboard – ACS, Department of Public Works

The City's Animal Care and Services (ACS) Division responds to over 20,000 calls for service each year with a team of 13 field officers operating on a 24x7x365 response schedule. With a small team and large workload, efficient delivery of service is imperative. Historically, dispatchers relied on work lists and intuition to assign field cases. More recently, the ACS Dispatch Map was introduced and provided dispatchers with a map-based view of real-time calls

for service and field officer vehicle locations. This enabled the Division to dispatch the closest team member and thereby decrease response time and route travel.

The ACS Dispatch Map also aims to improve officer awareness and safety by displaying the location of active permits for "potentially dangerous", "dangerous", and "vicious" dogs. Verizon Connect telemetry was integrated into the ACS Dispatch Map, allowing dispatch and other officers to see the location of all Animal Services Trucks in near-real time.

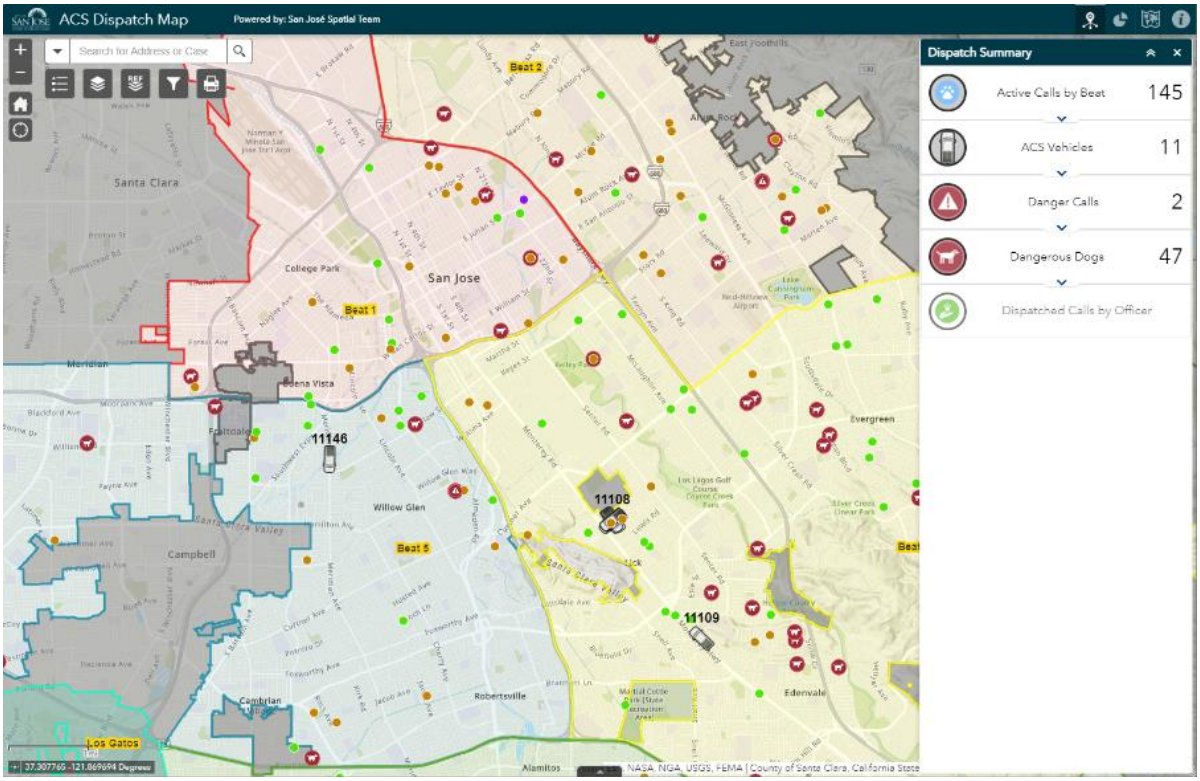

Figure 2: Animal Care Services Dispatch map provides Citywide visibility for response resources

With the success of the ACS Dispatch Map, a complementary dashboard is in development to provide access to historical calls for service since 2018, thereby allowing staff to plan and target activities in response to licensing and stray animal intake patterns, coupled with areas where call volumes trend high. The dashboard offers data filtering to recognize space-based and time-based patterns in call types and volumes to inform operational decisions and program-level practices.

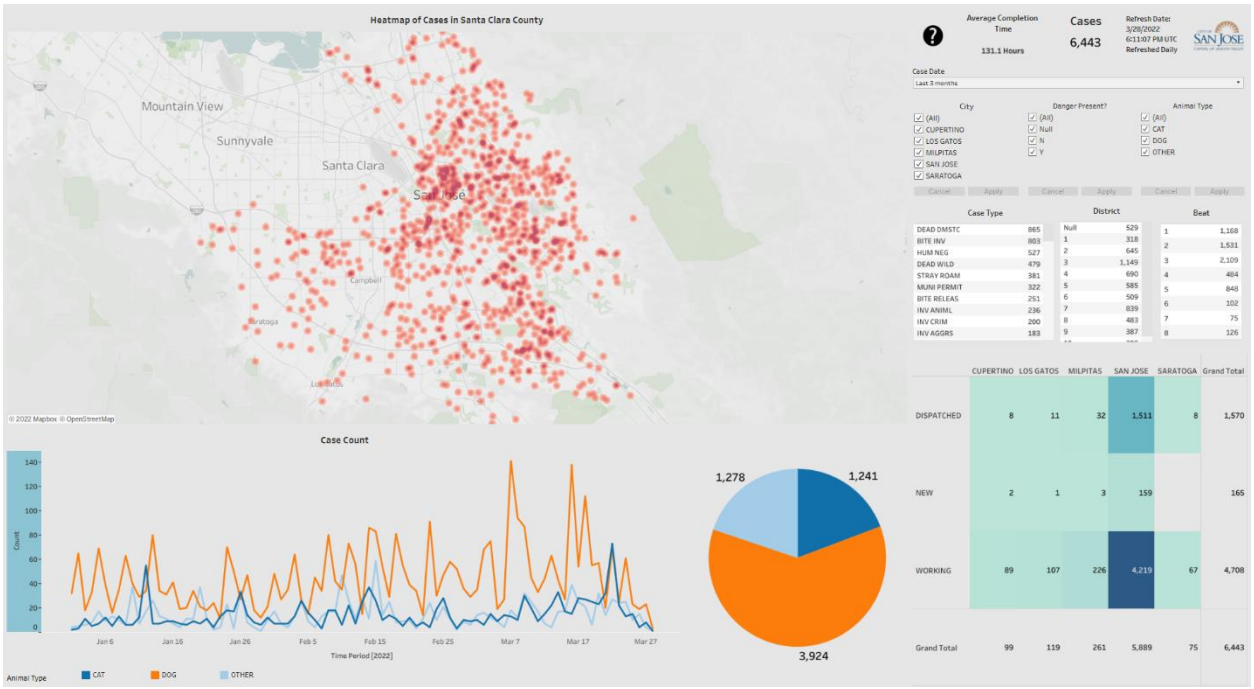

Figure 3: Animal Care and Services Dashboard, showing ACS case history

Applications for Sewer repairs and reporting

As highlighted in Figure 4, the City's Department of Transportation (DOT) received over 1,000 sewer repair requests from January 2021 to January 2022. The importance of this work is to maintain and improve the debris collection system infrastructure, minimize the number of sanitary sewer overflows (SSOs), and to safely move sewage to the regional Waste Water Treatment Plant. The data allow DOT staff to respond quickly, which is essential to life and safety, as well as regulatory compliance with the State Water Board. Previously, DOT used a mix of decentralized applications and systems, combined with paper and manual processes, to manage and report all sewer repair requests.

DOT has consolidated most applications and data into its Unity platform over the past seven years, enabling staffs to create more streamlined processes with accurate reporting. Staff created an automated backend system that can produce reports, visualizations, and analyses that pull from a single "source of truth" database, to maximize use and ease for City departments.

Priority	On Time	Missed	Pending
⊞	**0**	**0**	**126**
⊟ **A**	**107**	**5**	**0**
Manhole	8	0	0
Other	4	0	0
Sanitary Lateral	84	4	0
Sanitary Main	8	0	0
Storm Lateral	2	1	0
Storm Main	1	0	0
⊟ **B**	**53**	**6**	**0**
Manhole	2	0	0
Other	7	1	0
Sanitary Lateral	38	4	0
Sanitary Main	5	0	0
Storm Lateral	1	0	0
Storm Main	0	1	0
⊟ **C**	**623**	**55**	**77**
Manhole	23	5	7
Other	88	4	2
Sanitary Lateral	270	12	31
Sanitary Main	208	28	30
Storm Lateral	7	0	1
Storm Main	3	0	0
Undermine and Sink	24	6	6
Total	**783**	**66**	**203**

Figure 4: Summary dashboard of sewer repair requests, providing an immediate view of the City's performance

Consolidating systems and software development efforts has been a multi-year effort and has produced significant efficiencies. In this sewer repair service are, efforts have enabled the Sewers team to be proactive and make educated predictions. Thus, the Sewers team is now able to respond more efficiently to resident service requests for sewer items— reports that previously took 2-3 hours to generate now only take 15 minutes and are more reliable, leading to a quicker, and more proactive response to resident sewer service needs.

Measuring Community Impact

Visualizing impact – comparing socio-economic status with call response times

There is an increasing need for governments and systems to analyze the effectiveness and equity of its services at a large scale. In collaboration with Joy Hsu, a City Data Equity Fellow and PhD candidate at Stanford University, and Ramya Ravichandran and Edwin Zhang from DataKind,[14] a global nonprofit organization bringing expert data scientists to find solutions with social change organizations, the Mayor's Office of Technology & Innovation (MOTI) launched a

[14] Link to DataKind homepage: https://www.datakind.org/

publicly accessible system to understand the quality of emergency government services in San José by location and the demographics of people who live there.

Emergency response times are an important measure of quality of government service. This measure both affects decisions about future resourcing— e.g., where fire stations are placed— and can be used by community members to highlight both strengths and areas of potential improvement for City services.

On October 5, 2021, the team published their findings at the Association for Computing Machinery's Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (ACM EAAMO)[15], an academic conference with participation from technical experts as well as others in the public policy domain. The team then built a prototype[16] to demonstrate what this open data analysis might look like.

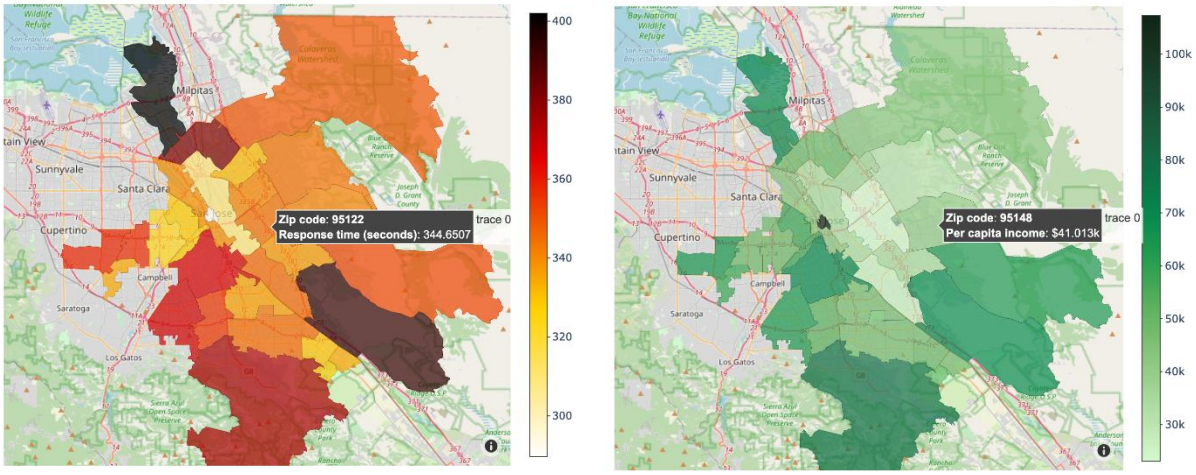

Figure 5: The left map shows emergency response time by zip code, and the right map shows per capita income from the Census.

Used well, this work can influence future data collection by the City of San José and other local governments to ultimately support greater local transparency, joint response models, station redistricting efforts, and performance monitoring for essential services like emergency response. There are many ways to generalize this work beyond just income and response times, and this is only a first step to more equitable service delivery in San José.

[15] Link to EAAMO homepage: https://eaamo.org/#home
[16] Link to prototype: https://sj-open-data-api.herokuapp.com/

Supporting Equity

City departments are asking for new data analytics to re-design services and outreach efforts. In conjunction with the City's Digital Privacy Officer, departments are defining new ways to improve services through data, while protecting civil liberties around individual privacy.[17] Used well, data and analytics can help COVID recovery efforts create opportunity for the San José community and guide how the City can dedicate resources based on clear priorities, defined outcomes, and common impact measures. This memo highlights foundational efforts to compare community needs with services provided throughout San José.

San José 311 Analytics

The San José 311 (SJ311) team engaged with data scientists with partners at Dell Technologies to gain insights into the General Requests/Other Issues category of Service Requests (SRs) in the SJ311 platform. "Other issues" constitute more than 30% of the total SRs created through the platform. Data Scientists used Natural Language Processing on the free text "Description" field to help identify the top 10 issues that are reported through the catch-all "Other Issues" category. The intent for the analysis was to (1) inform decisions for future SJ311 services in discussion with department teams and City Council, and (2) look for insights that could help increase community equity in those service development decisions.

Internally, the City's Enterprise GIS team assembled a dashboard for departmental and Council District staff to see visualizations of SJ311 SR data. SRs can be filtered by timeframe, type of service requested, department receiving the request, and with various geographic boundaries, including council districts, neighborhoods, and census tracts. By overlaying neighborhood, the Equity Score, Language Score, and Census data on the SJ311 SR data, the SJ311 team and City Manager's Communications Office can better understand usage patterns and tailor outreach program for communities where SJ311 is underutilized. This dashboard provides both a mechanism to better identify "missed" communities and the ability to understand demographics to improve outreach.

[17] The City's approach to individual privacy can be found in the City's Digital Privacy Principles and Policy, which can be viewed on the City website: https://www.sanjoseca.gov/your-government/departments-offices/information-technology/digital-privacy

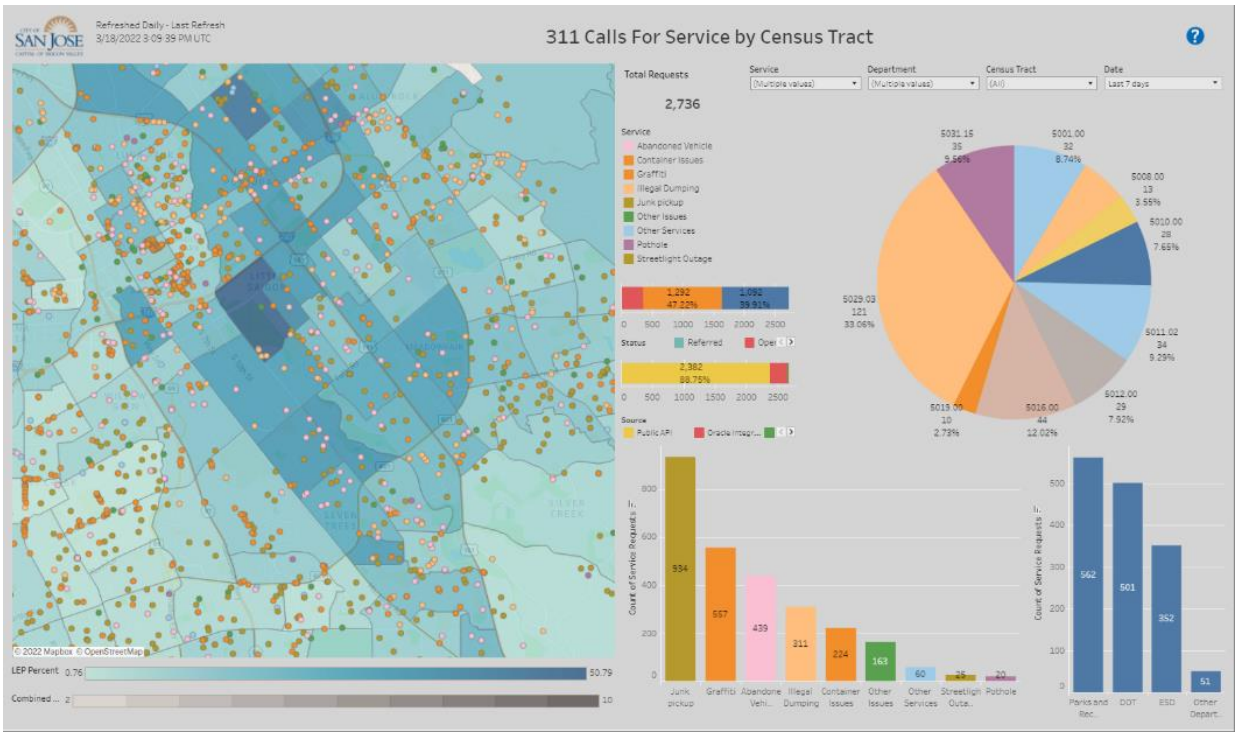

Figure 6: San José 311 (SJ311) Dashboard compares calls for service (shown by the color of the dots) with the census tract percent of population with Limited English Proficiency (LEP Percent)

San José Equity Atlas – Citywide

Currently in preliminary release, the San José Equity Atlas adopts a framework to assign equity scores to communities based on demographic factors, including race and income. The composite score is easily interpretable and usable for prioritization of services and investments. The Equity Atlas has been used in recent budget activities to present and evaluate proposals from an equity perspective. Several City programs have also used this index as a factor in project prioritization, including green infrastructure and sanitary sewer improvements. In addition, this methodology for a simple composite score-based index has been ported to other initiatives, including COVID-19 vaccination. The Equity Atlas score informed decisions on the deployment of vaccination sites and services.

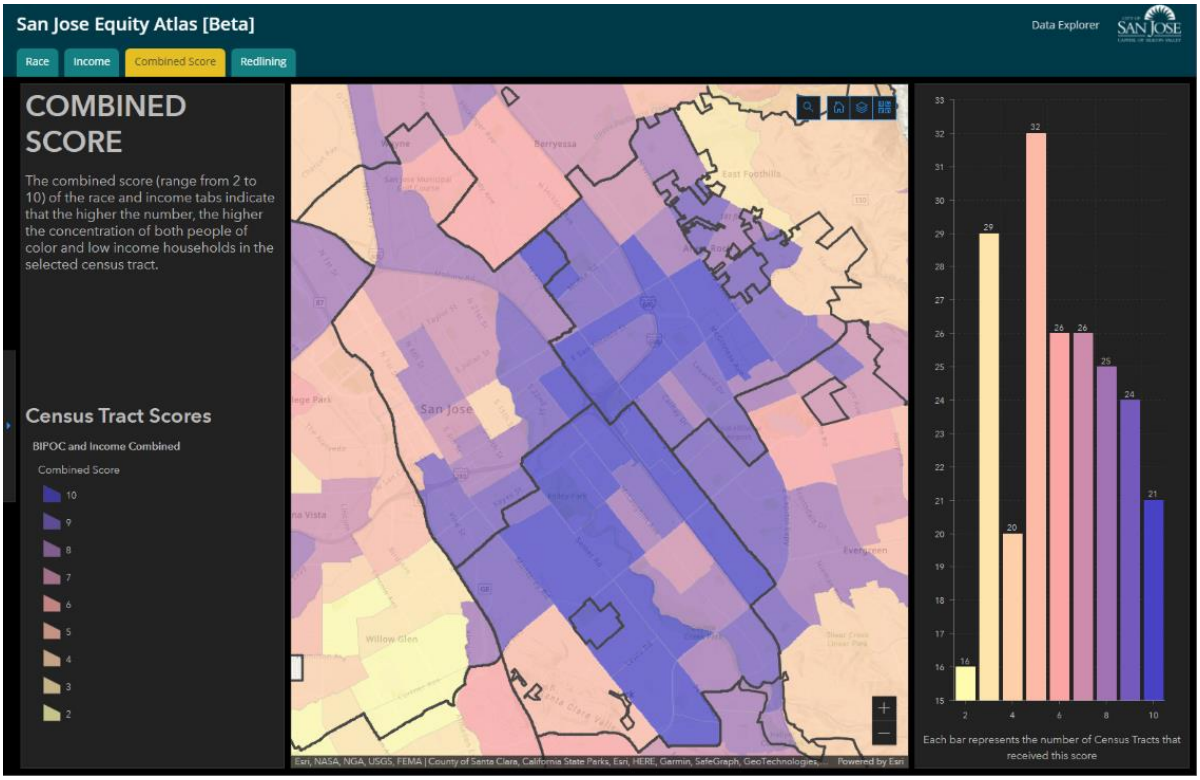

Figure 7: San José Equity Atlas provides a prioritization score based on median income and racial composition of census tracts in the City

Many of the dashboards rely on a core set of community demographic data to help tailor services based on need and providing the equity lens departments seek. In response to this, the Enterprise GIS team is working to build a standardized data structure with commonly used Census attributes, available for all dashboards and demographic analytical products. This model will begin with Census 2020 data and be updated annually as new American Community Survey (ACS) data become available.

Existing dashboards will be refactored to utilize this standardized demographic data. Moving forward, these dashboards will be automatically updated with the most current available demographic data when the base data model is refreshed. This includes updates to prioritization indices that are based on this demographic data. This structure is aimed at establishing a consistent approach to equity analysis, reducing duplication of effort, and ensuring that current and accurate data is available for Citywide analytics.

Next steps – Building community partnerships

The projects outlined in this memo present direct impacts through data and digital services. To further unlock the value of data, the City is partnering across departments and with external stakeholders to build on the people, processes, and tools for data-driven decision-making.

People

Through collaboration across the Mayor's Office of Technology and Innovation (MOTI), the Information Technology Department (ITD), and Parks, Recreation & Neighborhood Services (PRNS), the City has partnered with universities including San José State, Stanford, and Harvard in accessing rare talent for necessary data projects. Beginning in Summer 2022, two graduates in the Stanford Impact Labs Collaborative Research Fellowship Program will join the City as full-time Equity through Data Fellows, focused on advancing equity through data analytics and data privacy.

To further grow capacity, ITD has created a dedicated team focused on supporting City and departmental equity goals through data and digital privacy. The Equity through Data and Privacy (EDP) program partners with the Office of Racial Equity, Library, PRNS, Public Works, MOTI, and others to develop a sustainable data infrastructure that will improve policy and program development so that ultimately outcomes are improved for low-income and communities of color.

Processes

The Equity through Data and Privacy (EDP) program will work with the Office of Racial Equity (ORE) and the COVID-19 Recovery Taskforce (CRT) in defining performance measures and outcomes for COVID recovery, monitoring performance, and identifying long-term impact.

While many data projects have proven valuable, the City has observed that data projects often fail to create an impact when they are started (1) without a clear goal(s), (2) with limited access or access to only low-quality data, and/or (3) with teams unprepared with the skills to collect, manage and analyze data.

The EDP program is working with departments across the City to introduce a Data Chartering process that structures the creation of data projects and facilitates identifying the value from data. Data Chartering helps identify the goals for data analysis upfront, makes it easy to organize resources, and ensures the utility and sustainability of the analysis. Among other key initiatives, EDP is working with ITD and the Library Department on refining digital inclusion and improving SJ311 services.

Identifying impact, or how an initiative or program effects an outcome (e.g., unemployment), can take months or years due to the need for a backlog of data to study. By forging partnerships and making investments now the City can see exponential returns in understanding the impact of City services to ultimately improve those services. In the meantime, departments will continue to leverage existing data to equitably service distribution, identify urgent or emerging needs for services, and study impact on programs that have already collected high-quality data.

Tools

The City is leveraging partnerships with vendors to develop the Open Data Community Architecture (ODCA), a system for housing shared City data to enable secure cross-department and external collaboration on data projects. The City continues to combine the ongoing technical work with department outreach to facilitate Citywide adoption for storing and sharing metrics on City programs.

Internally, the City uses a set of industry-standard tools for data analysis and visualization. Department staffs have used Power BI, Tableau, and ArcGIS for creating dashboards like the examples shown in this memo. The City already has robust support for Power BI and ArcGIS, but limited support in Tableau. A growing number of departments are looking for a more effective tool at merging interactive maps with their own data to show a complete, interactive data story. Departments are working together to defines, procure, and support the need for a GIS-compatible data visualization tool with centralized support.

CONCLUSION

The City's ongoing investments into its data systems have produced insights and tools that improve how the City serves the community in San José. Along the goals for these efforts that were defined in 2017, the City continues to grow how it:

Improves service delivery through data with increased visibility, better communications, and low code/no code automated processes to save time and money while improving service quality. The Fuel Monitor Dashboard has reduced the time needed to identify fuel needs of public vehicles from days to hours, reducing refueling time and increasing in-service time of emergency response engines. The Sewer repairs data consolidation project has reduced reporting time for accurate information from hours to minutes, and reduced response time for emergency sewer repair.

Measures the impact of City services on our communities, allowing the City to refine our long-term strategy to create more value with the same resources. The prototype dashboard comparing emergency response time with income across zip codes highlights where the City can improve services in the future. As the City continues to grow its backlog of usable data, additional studies on service impact will become possible.

Supports equity work by comparing the measured needs of our communities with the services provided to identify who, where, and how the City needs to serve better. The SJ311 analytics dashboard provides an essential tool in identifying where residents are requesting support, and what demographics need additional outreach to increase service utilization. Paired with the San José Equity Atlas, the City is defining its base understanding of racial and other forms of equity across the City to identify general need for services.

To continue growing its capacity for valuable analytics, the City is forging partnerships across private industry, philanthropic foundations, and academic institutions. One major example: the Knight Foundation is providing San José $750,000 through a three-year grant to start the City's Equity through Data and Privacy (EDP) program. The EDP program will support the City's leadership and capacity to make data-driven decisions, measure outcomes, protect digital privacy, and drive results that incorporate equity into the planning and delivery of City services. This program will coordinate with the Office of Racial Equity, Parks, Recreation and Neighborhood Services, and participating departments to improve City investments in racial equity, data analytics, and digital privacy protections.

The investments the City has made in data and platforms are beginning to show measurable and noticeable results. The cross-department data efforts from Mayor's Office of Technology and Innovation, Department of Public Works, Department of Transportation, Information Technology Department, Parks Recreation and Neighborhood Services, Office of Racial Equity, and others will continue to push for a smarter, safer, and more equitable City in the future.

EVALUATION AND FOLLOW-UP

The City Manager's Office, Office of Racial Equity, Department of Public Works, Department of Transportation, Mayor's Office of Technology and Innovation, and Information Technology Department will report progress on Data Initiatives and Impacts Status at future Smart Cities and Service Improvements Committee meetings as approved in the Committee Workplan.

CLIMATE SMART SAN JOSÉ

The recommendation in this memo has no effect on Climate Smart San José energy, water, or mobility goals.

PUBLIC OUTREACH

This memorandum will be posted on the Smart Cities and Service Improvements Committee agenda website for the April 7, 2022, meeting.

COORDINATION

This memorandum has been coordinated with the City Manager's Office, City Attorney's Office, Office of Racial Equity, Department of Transportation, Department of Public Works, Mayor's Office of Innovation and Technology, and Information Technology Department.

COMMISSION RECOMMENDATION/INPUT

This report was not presented at a Commission.

CEQA

Not a Project, File No. PP17-009, Staff Reports, Assessments, Annual Reports, and Informational Memos that involve no approvals of any City action. (Information Technology)

/s/
ROB LLOYD
DEPUTY CITY MANAGER/
CHIEF INFORMATION OFFICER
CITY MANAGER'S OFFICE

For questions, please contact Arti Tangri, Information Technology Department Data Architect, at arti.tangri@sanjoseca.gov or (408) 793-6862; Harsh Gautam, Public Works Division Manager, at Harsh.Gautam@sanjoseca.gov or (408) 535-8464; or Vince Pereira, Department of Transportation Department Information Technology Manager, at Vince.Pereira@sanjoseca.gov or (408) 975-3233; Albert Gehami, Information Technology Department Digital Privacy Officer, at albert.gehami@sanjoseca.gov or (408) 793-6878.