Performance & Analytics

Open Data Update

Economic Development and
Intergovernmental Relations Committee

March 10, 2016

The City of SAN DIEGO

Why is open data important?

- Promote civic engagement
- Improve service delivery
- Increase opportunities for economic development
- Make communication with public more effective

How do we open our data?

- Release data in an open format onto a publicly accessible website without login requirements
- Describe the data
- Automate regular updates
- Promote data use both externally and internally

Policy milestones met

- **November 24, 2014:** Chief Data Officer joined the City
- **January 1, 2015:** Open Data Policy went into effect
- **March 9, 2015 - May 2015:** CDO and coordinators with 35 departments conduct first data inventory effort
- **July 1, 2015:** CDO published an implementation update in Gitbook
- **February 29, 2015:** Open Data team launches DataSD.org, which includes the results from the data inventory

Policy milestones to meet

- **July 1, 2016:** Portal is launched, and Open Data team begins publishing datasets
- **January 1, 2020:** All high-value public datasets are on the portal

Prioritized inventory

- Initially defined “dataset” broadly
- Allowed us to cast a wide net for city data
- The raw inventory results needed a lot of work to become a reliable list of City datasets

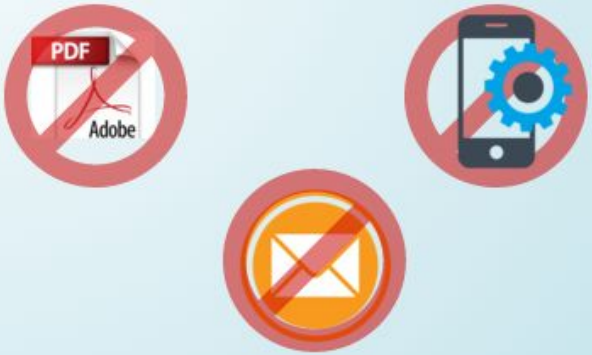
What is [not] a dataset?

A Map

A Dashboard or Chart

A Website

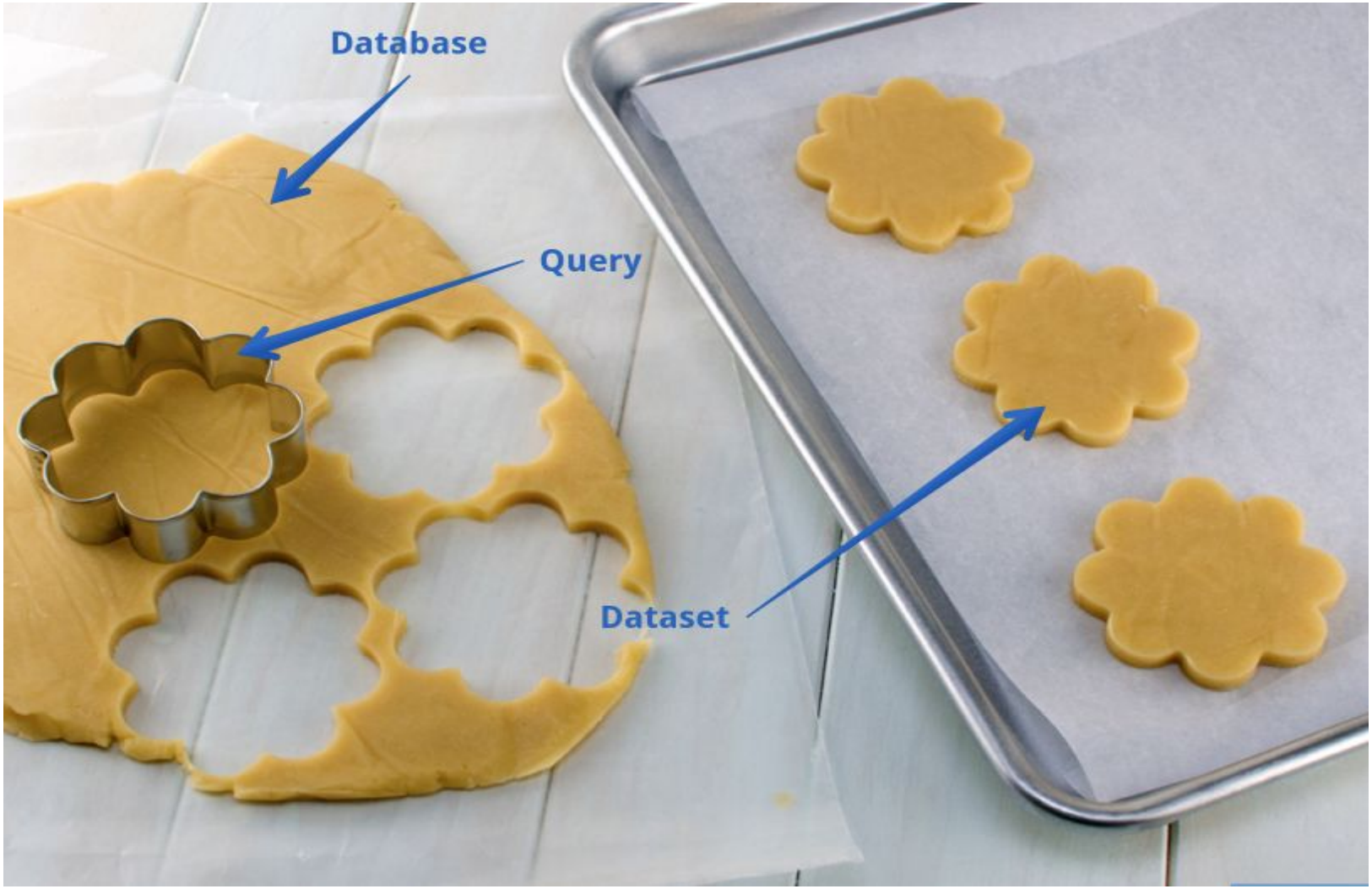
A PDF / Mobile App or E-Mail



What is a dataset?

C	D	E	F	G	H	I	J	K	L	M	N	
tics.ProcessThread												
d	IdealProcessor	PriorityBo	PriorityLevel	PrivilegedProcessorTime	StartAddress	StartTime	ThreadSta	TotalProcessorTim	UserProce	WaitReason	Processor	Site
7148		TRUE	Normal	00:09.0	0	11/7/2011 16:33	Wait	00:44.9	00:35.9	UserRequest		
8120		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
5320		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	00:00.0	UserRequest		
3080		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.1	00:00.1	UserRequest		
6556		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.1	00:00.1	UserRequest		
3400		TRUE	Normal	00:00.0	0	11/7/2011 16:33	Wait	00:00.2	00:00.1	UserRequest		
4612		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	0:00:00	UserRequest		
7252		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	0:00:00	UserRequest		
7140		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
7980		TRUE	Highest	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
5848		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
7844		TRUE	Normal	00:00.1	2005386496	11/7/2011 16:33	Wait	00:00.2	00:00.1	UserRequest		
7628		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	0:00:00	UserRequest		
6880		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	0:00:00	UserRequest		
6176		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
6152		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	00:00.0	UserRequest		
8176		TRUE	Normal	0:00:00	0	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
3008		TRUE	Normal	00:00.0	0	11/7/2011 16:33	Wait	00:00.2	00:00.1	EventPairLow		
2696		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	EventPairLow		
904		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	EventPairLow		
8004		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	UserRequest		
1640		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:33	Wait	0:00:00	0:00:00	EventPairLow		
7088		TRUE	Normal	00:00.0	2005386496	11/7/2011 16:33	Wait	00:00.0	0:00:00	UserRequest		
3244		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:34	Wait	00:00.5	00:00.5	UserRequest		
3384		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:34	Wait	0:00:00	0:00:00	UserRequest		
6612		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:42	Wait	0:00:00	0:00:00	UserRequest		
6496		TRUE	Highest	0:00:00	2005386496	11/7/2011 16:44	Wait	00:00.0	00:00.0	UserRequest		
5168		TRUE	Lowest	0:00:00	2005386496	11/7/2011 16:44	Wait	0:00:00	0:00:00	UserRequest		
7308		TRUE	Normal	0:00:00	2005386496	11/7/2011 16:46	Wait	0:00:00	0:00:00	UserRequest		
5612		TRUE	Normal	00:00.0	2005386496	11/7/2011 17:45	Wait	00:00.0	00:00.0	UserRequest		
6184		TRUE	Normal	0:00:00	2005386496	11/7/2011 18:03	Wait	0:00:00	0:00:00	UserRequest		
6836		TRUE	Normal	00:00.0	2005386496	11/7/2011 18:11	Wait	00:00.0	00:00.0	EventPairLow		
3756		TRUE	Normal	00:00.8	2005386496	11/7/2011 18:11	Wait	00:00.8	0:00:00	UserRequest		
4380		TRUE	Normal	0:00:00	2005386496	11/7/2011 18:13	Wait	0:00:00	0:00:00	EventPairLow		
3536		TRUE	Normal	0:00:00	2005386496	11/7/2011 18:26	Wait	0:00:00	0:00:00	EventPairLow		
7924		TRUE	Normal	0:00:00	2005386496	11/7/2011 18:27	Wait	0:00:00	0:00:00	EventPairLow		

What is a dataset?



What is a high-value dataset?

What is a high-value dataset?

- Value
- Security
- Quality
- Readiness

Example

- Calendar of Special Events (high value)
- List of designated chainsaw operators (not high value)

DataSD.org - Launched Feb 29

Help us prioritize datasets for release!

Show 10 entries Search:

Date Added	Category	Description	Vote
2/19/2016	Economy and community	Active and inactive business tax accounts	Vote
2/19/2016	Economy and community	Active businesses by Business Improvement District	Vote
2/19/2016	Culture and recreation	Amount of money each city-funded arts and culture organization annually spends in the local economy	Vote
2/19/2016	Culture and recreation	Amount of revenue contributed by entities other than the city annually to each arts and culture organization that also receives City funding	Vote
2/19/2016	City management and ethics	Appointments, reappointments, resignations in City boards and commissions	Vote
2/19/2016	Culture and	Athletic areas as a geographic dataset	Vote

Link

Traffic so Far

Sessions	Users	Pageviews
992	841	1,894

Pages / Session	Avg. Session Duration	Bounce Rate
1.91	00:01:49	47.38%

% New Sessions
84.78%

Demographics

- Language
- Country
- City

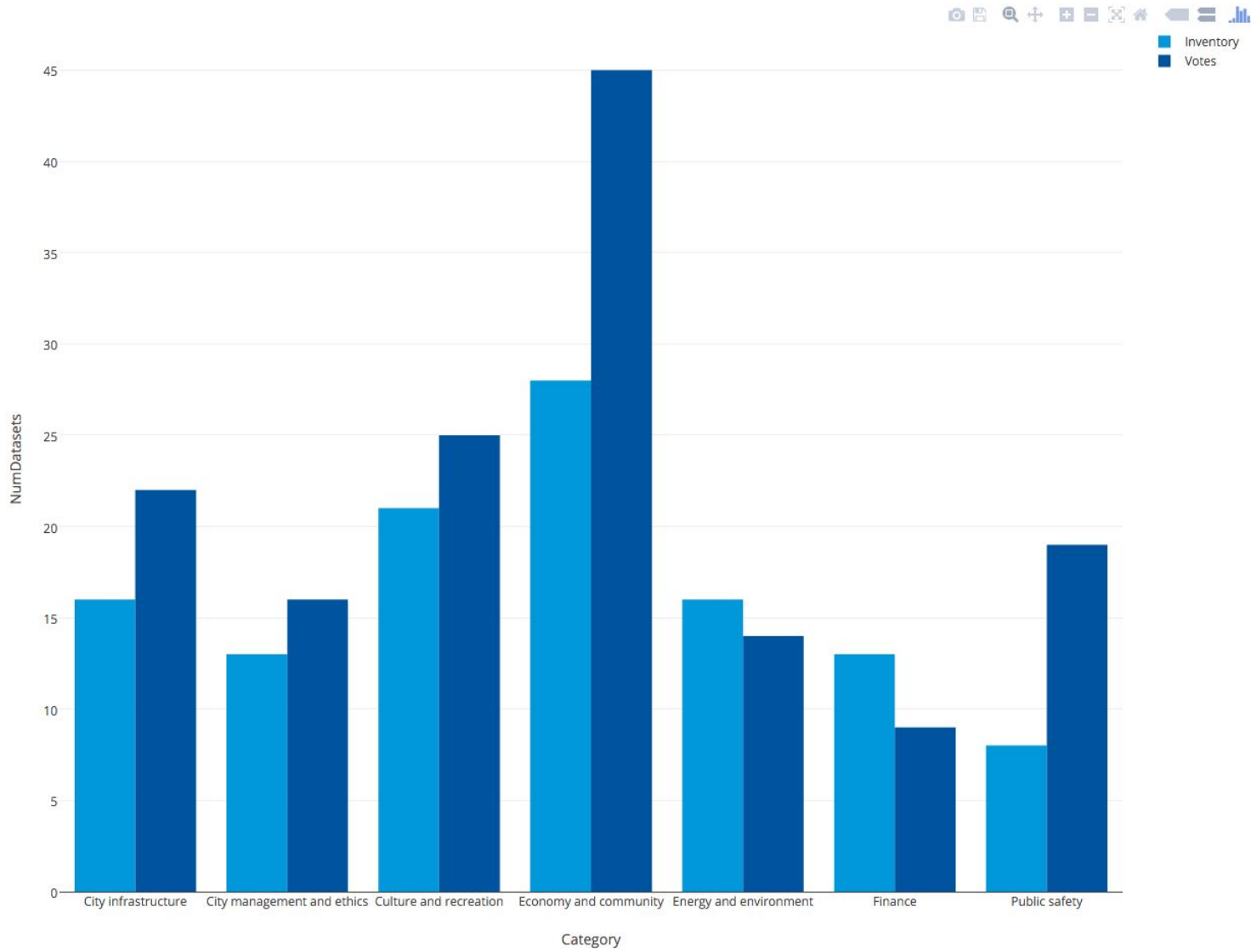
System

- Browser
- Operating System
- Service Provider

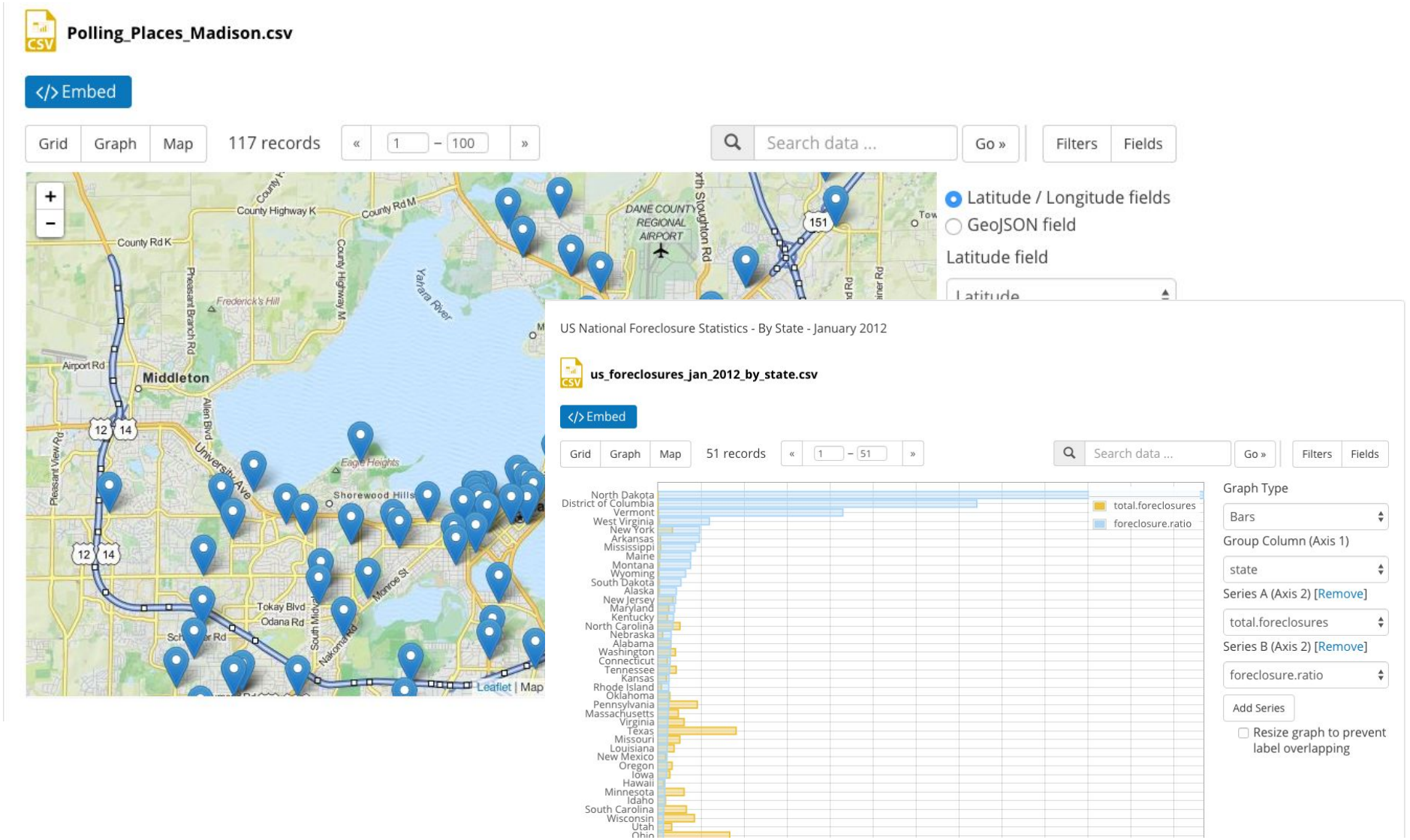
Mobile

	City	Sessions	% Sessions
1.	San Diego	696	70.16%
2.	Los Angeles	59	5.95%
3.	El Cajon	18	1.81%
4.	Chula Vista	17	1.71%
5.	Carlsbad	16	1.61%
6.	(not set)	10	1.01%
7.	Encinitas	9	0.91%
8.	Santee	9	0.91%

Votes by Category



Data Portal


Polling_Places_Madison.csv
Embed
Grid
Graph
Map
117 records
1 – 100
Search data ...
Go »
Filters
Fields
Latitude / Longitude fields
GeoJSON field
Latitude field
Latitude
Middleton
Shorewood Hills
Eagle Heights
Frederick's Hill
DANE COUNTY REGIONAL AIRPORT
Yahara River
Leaflet | Map
US National Foreclosure Statistics - By State - January 2012
us_foreclosures_jan_2012_by_state.csv
Embed
Grid
Graph
Map
51 records
1 – 51
Search data ...
Go »
Filters
Fields
North Dakota
District of Columbia
Vermont
West Virginia
New York
Arkansas
Mississippi
Maine
Montana
Wyoming
South Dakota
Alaska
New Jersey
Maryland
Kentucky
North Carolina
Nebraska
Alabama
Washington
Connecticut
Tennessee
Kansas
Rhode Island
Oklahoma
Pennsylvania
Massachusetts
Virginia
Texas
Missouri
Louisiana
New Mexico
Oregon
Iowa
Hawaii
Minnesota
Idaho
South Carolina
Wisconsin
Utah
total.foreclosures
foreclosure.ratio
Graph Type
Bars
Group Column (Axis 1)
state
Series A (Axis 2) [Remove]
total.foreclosures
Series B (Axis 2) [Remove]
foreclosure.ratio
Add Series
Resize graph to prevent label overlapping

The portal will be a data catalog

CITY DATA

Database tables ›	Map files ›	Spreadsheets ›	Reports ›
City water use	ESRI Map Gallery	Tax revenue	Budget
Water testing	SanGIS layers	Parks water use	CIP

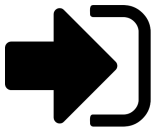
Residents | City staff | Businesses | Data scientists

Defined data release process

We want to make sure the data we release is well-described and of high-quality.

To achieve this, we created a set of steps for departments and our team to follow for each dataset that will be released onto the portal.

Meeting with data stewards



Review data sample

Remove PII

Stage dataset on portal

Get approvals

Plan for updates

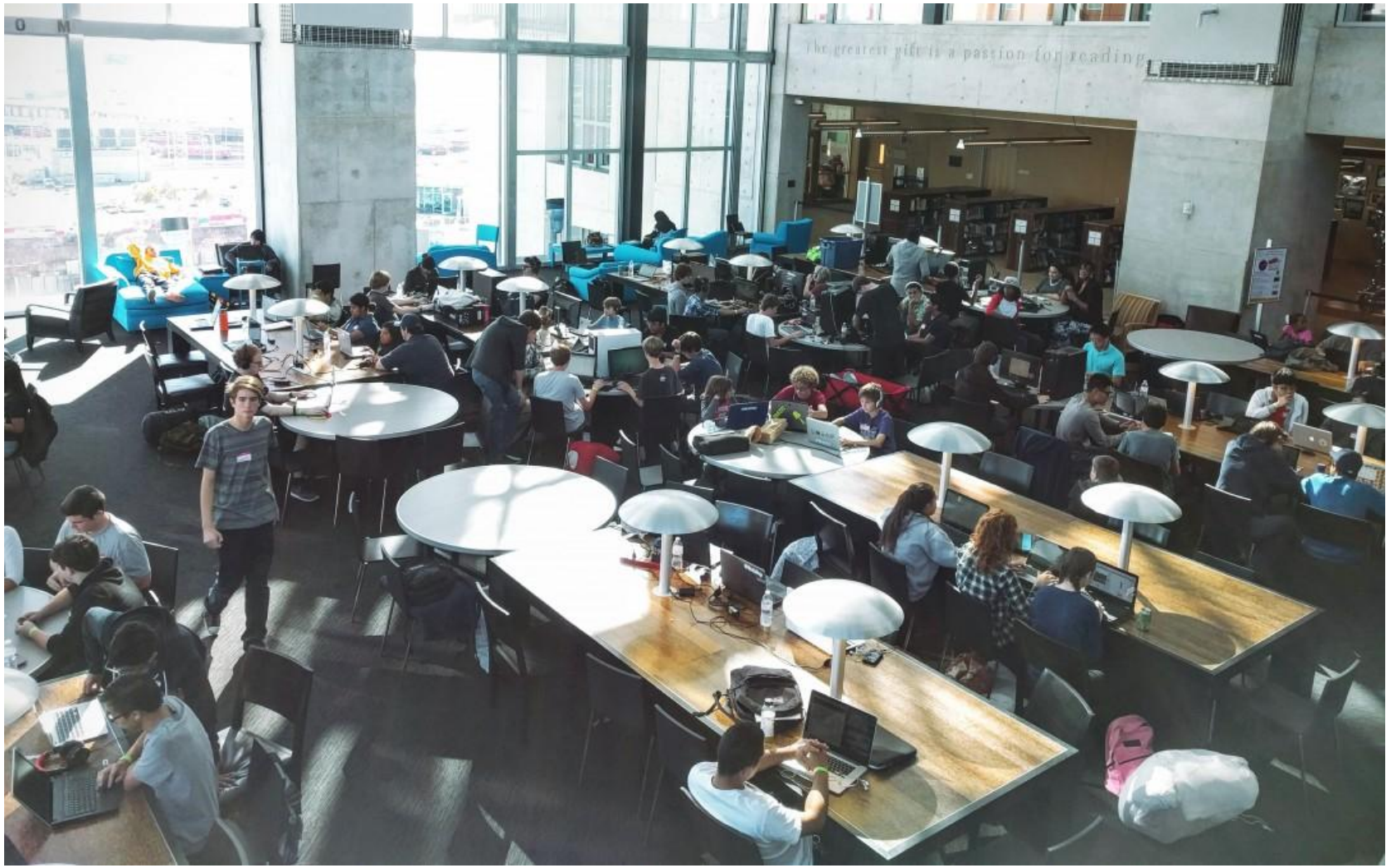
Next steps

- Get datasets ready for the portal
  - Work with Info Coordinators on describing data
  - Put technical processes in place for upload and updating
- Get portal ready for launch
  - Stage datasets as they become ready
  - Include open licensing and legal disclaimers
  - Develop features for the portal front page
- Keep community updated
- Review, refresh Open Data Policy as needed

Ongoing work besides portal

- Promote community code projects based on City data
- Offer technical advice, assistance and data toward other City efforts
- Advising other cities interested in open data

Community



Thank You!

Questions?