DataScienceSF

Overview of Data Science Services from DataSF

Learn more at datasf.org/science

What is data science?

Data science is a way for you to harness the power of advanced analytics and applied statistics for the challenges you face in your department.

This new service from DataSF aims to help departments achieve more with their existing resources and processes.

Through a 4 month engagement, DataSF's Data Science team and your department will refine a problem, identify statistical methods to address it, and develop and institute a service change to improve your work.

Data Science
Applying advanced statistical tools to existing data to generate new insights

+

Service Change
Converting new data insights into (often small) changes to business processes

Smarter Work
More efficient and effective use of staff and resources

What types of problems can benefit from data science?

There are 5 basic types of civic problems data science can help address. (See reverse side for more details). Ask yourself if you or people in your department would want to:

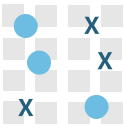
- Find the needle (in the haystack)

- Prioritize backlogs

- Flag "stuff" early

- A/B test something

- Optimize resources



How does data science fit in with other tools?

There are other great ways to bring about service improvement that complement data science:

Approach	Description
Performance management	Define, visualize (often via dashboards), and manage to metrics and key performance indicators
Policy Analysis	Define and assess alternatives using a range of tools and issue a report or memo with policy or program recommendations
Evaluation	Study the cost and benefit of a program to inform policy and program changes, including to inform investments
Open data	Make civic data public to facilitate data sharing, reporting and new tools

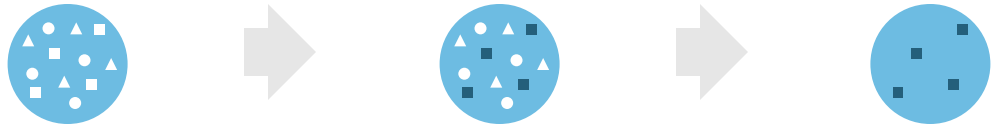
How to participate?

1. **Learn** more about the process at datasf.org/science. Use our office hours to help define your project.
2. **Apply** by deadline.
3. **Be available** for questions while DataSF works to define and select projects for the first cohort.
4. If selected, **partner** to refine and iteratively analyze the question.
5. **Implement & present** the results to your fellow cohort by December.

Data Science Project Types | **Submit a project at datasf.org/science**

Find the needle in the haystack

Who, what, or where to target? | Data science identifies candidates | Service changes to target candidates



Approach: Targets worth your department's resources are often difficult to identify (people, geographic areas, or categories). Data science identifies candidates to target based off analysis of past data.

Example: New York City wanted to improve the finding rate of its time consuming tax compliance audits. They identified filing patterns and increased the number of audits with findings from 63 to 78%.

Prioritize your backlog

What to prioritize in your backlog? | Data science identifies priorities | Service changes to order by priority

Approach: Department backlogs often grow because implementing a triage process is too staff intensive. Data science can help identify high, medium, and low priority cases by analysing existing data.

Example: In SF, the Assessor-Recorder must determine if a sale was done at fair market value through a time intensive manual process. DataScience SF instrumented a regression model that automatically flags outliers for closer inspection.

Flag "stuff" early

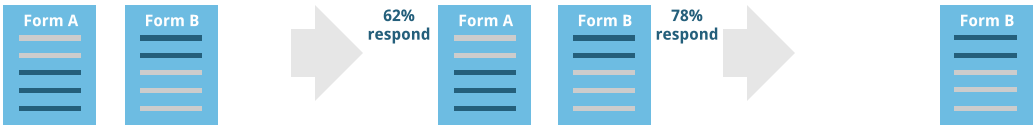
How do you flag stuff early? | Data science predicts events | Service changes to intervene

Approach: Many situations - good and bad - could be addressed more efficiently if caught early, even before they come to you. Data science identifies candidates for early intervention and engagement.

Example: Since 2011, SF's DPH-WIC program had seen an increase in mothers dropping out. DataScienceSF built a predictive model to help target outreach to clients who are at greatest risk for dropping out.

A/B test something

Which will get the best response? | Data science tests the forms | Service changes to use form B

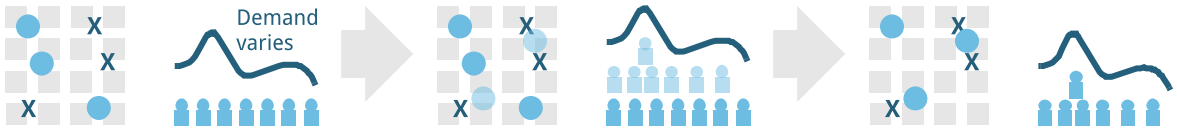


Approach: Departments want to maximize the efficacy of costly communication efforts. Data science can help identify and test various approaches to identify those that would be most successful.

Example: Forty percent cited for low-level violations in New York miss summons leading to costly arrests. A rapid redesign of the summons form and use of SMS is being tested on failure to appear rates.

Optimize your resources

How to distribute your resources? | Data science identifies alternative | Service changes to new distribution



Approach: Departments must decide how to distribute resources to minimize response time or queues and maximize services. Data science uses existing data to optimize distribution of services.

Example: Chicago's rodent program has trouble predicting outbreaks leading to spikes in 311 requests. Analytics identified leading indicators to help dispatch rodent team, which reduced 311 requests.