

# DataSF Guidebook:

## *Detailed Inventory Guide Steps 2 & 3*



**City and County of San Francisco**

**(Version 1.1)**

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

# Table of Contents

[Table of Contents](#)

[Introduction](#)

[Purpose of this document and Version Notes](#)

[Getting Started](#)

[Department Profiles](#)

[Low inventory complexity](#)

[Medium inventory complexity](#)

[High inventory complexity](#)

[Inventory Guidance](#)

[Low inventory complexity](#)

[Step 2: Brainstorm](#)

[Step 3: Inventory](#)

[Medium inventory complexity](#)

[Step 2: Brainstorm](#)

[Step 3: Inventory](#)

[High inventory complexity](#)

[Step 2: Brainstorm](#)

[Step 3: Inventory](#)

# Introduction

## Purpose of this document and Version Notes

The purpose of this document is to provide more detailed guidance on completing the dataset inventory to Data Coordinators in the City and County of San Francisco. Data Coordinators should use this as a starting point and reference to completing the last two steps of the dataset inventory: identification of datasets, and final inventory of datasets.

Date	Version	Description of changes
September 23, 2014	1.0	Creation of guidebook including: <ul style="list-style-type: none"><li>• Background and motivation for Open Data in San Francisco</li><li>• Roles and responsibilities</li><li>• Guidance on creating a data catalog (inventory)</li></ul>
October 21, 2014	1.1	Revised guidebook based on feedback <ul style="list-style-type: none"><li>• Removed references to the survey so that this document can stand separately as a reference</li><li>• Moved deadline to the front of the document</li></ul>

## Getting Started

Now that you've completed Step 1 of the inventory process (where you identified data sources), you are ready for Steps 2 and 3.

For Steps 2 and 3, we provide slightly different guidance depending on your type of department. While there are unique challenges for each department, below are three general profiles that we think cover the majority of departments.

Please keep in mind that **these profiles relate only to the inventory process**. Implementation of your open data publishing processes will come in a later phase and the complexity of implementation may differ from the complexity of completing the inventory.

1. Please review the three profiles in the next section
2. Review the specific guidance for that department category
3. Complete the inventory process by **March 13th, 2015**

# Department Profiles

## Low inventory complexity

You are a department where a single division/group/person has a good handle on all the data that flows in and out of the department. If you have questions, it's easy enough to get them answered by walking down the hall or making a quick visit to someone's office.

**Take action:** If this profile closely matches your department, [review the guidance](#).

## Medium inventory complexity

You are a department where, while multiple divisions produce and use data, a known group of people or a handful of individuals has a good handle on who produces what data and for what purpose. Even if the department is large (and you don't know all of your colleagues), you can quickly identify the right person to talk to about the department's data. Questions can be answered easily over email, a quick phone call or a quick jog around the building.

**Take action:** If this profile closely matches your department, [review the guidance](#).

## High inventory complexity

You are a department where nearly every division/group uses, maintains, or reports on data, and there is no readily defined group or set of individuals who have a good handle on all of the data in the department. Even if each division/group has an individual that understands their respective data, and there may still be blindspots to explore. Your department may be spread over different buildings or many floors of a building and getting answers to questions may take many phone loops or emails to find the "right" person.

**Take action:** If this profile closely matches your department, [review the guidance](#).

# Inventory Guidance

## Low inventory complexity

### Step 2: Brainstorm

1. **Identify who should lead the first attempt at brainstorming a list of datasets.** In many cases, it may just be you, the Data Coordinator, but it could be another colleague you feel may have a good 30,000 foot view of the department's data.
2. **That person will download your department's copy of the inventory spreadsheet (MS Excel).** You should have the link to download in your email, if you cannot find it, please email [jason.lally@sfgov.org](mailto:jason.lally@sfgov.org).
3. **Then, brainstorming can begin in the tab labeled "Step 2 Brainstorm,"** begin listing the datasets as you think of them. This is literally just a list of all the datasets one can think of. When brainstorming datasets, that person should consider:
  - a. What data populates your monthly or quarterly reports?
  - b. What departmental data is currently publicly available on DataSF or elsewhere online?
  - c. What data does your department use for internal performance and trend analysis?
  - d. What information is published as a performance metric?
  - e. What data is reported to federal, state or local agencies?
  - f. Talk with your Public Information Officer (PIO) - what data has been requested under Sunshine?
  - g. What data do other departments ask for?
  - h. What kinds of open data are similar agencies across the country publishing?

#### Quick tip

You can review quick instructions in the first tab of your Excel inventory workbook labeled **Instructions**. This is a quick way to remind yourself of the steps without having to return to this document.

4. **Have another colleague review the brainstorm if you'd like.** As they say, two heads are better than one. Of course, if you've got a small enough department where you are the sole steward of data, than this may not be necessary.
5. When your brainstorm is complete, it's time to fill out the inventory!

### Step 3: Inventory

6. **In your Excel spreadsheet on the tab "Step 2 Brainstorm," click on the button**

**labeled “Prepare my Inventory.”** This will take your awesome brainstorm list and pre-fill in the top rows of the worksheet “Step 3 Dataset Inventory” with the dataset names and your department.

- 7. Now you can start populating across your datasets one by one in the tab labeled “Step 3 Dataset Inventory.”** If this list isn’t too big, and you have good knowledge of the datasets, it’s probably something you can start in on your own. Each field has directions included as a comment. Just mouse over each one to see what you should provide for each field.

#### Quick tip

There are a number of fields to cover for each dataset in the inventory and we know it could be overwhelming. It may help to **start with identifying the business contact (Data Steward) and technical contact (Data Custodian) first** for each dataset. That way if you’re stuck, you’ll have a handy reference on who to follow up with.

- 8. For datasets you may not have deep enough knowledge on, follow up with the business or technical contact for the dataset to fill in any remaining fields.** Our recommendation is to try to get as far as possible with the inventory yourself and then pull in your colleagues to fill in the gaps.
- 9. On completion you can attach your inventory to an email and send to [datasf@sfgov.org](mailto:datasf@sfgov.org), subject: Dataset Inventory Complete: <Your Department Name>.**

## Medium inventory complexity

### Step 2: Brainstorm

1. **Download your department's copy of the inventory spreadsheet (MS Excel).** You should have the link to download in your email, if you cannot find it, please email [jason.lally@sfgov.org](mailto:jason.lally@sfgov.org).
2. **Identify two or more of your colleagues that can brainstorm an initial list of datasets along with you.** These individuals will be first completing their own independent brainstorm within department's inventory spreadsheet.
3. **Each person on your team should complete an individual brainstorm within the Step 2 Brainstorm worksheet in your data inventory Excel workbook.** Each person will list all of the datasets as they think of them. This is literally just a list of all the datasets one can think of. Don't worry about overlapping a colleague's list. In the next couple of steps you'll all get together and remove duplicates. When brainstorming datasets, that person should consider:
  - a. What data populates your monthly or quarterly reports?
  - b. What departmental data is currently publicly available on DataSF or elsewhere online?
  - c. What data does your department use for internal performance and trend analysis?
  - d. What information is published as a performance metric?
  - e. What data is reported to federal, state or local agencies?
  - f. Talk with your Public Information Officer (PIO) - what data has been requested under Sunshine?
  - g. What data do other departments ask for?
  - h. What kinds of open data are similar agencies across the country publishing?

#### Quick tip

You can review quick instructions in the first tab of your Excel inventory workbook labeled **Instructions**. This is a quick way to remind yourself of the steps without having to return to this document.

4. **Each person on your brainstorming team should then send you their Excel workbooks.** You'll be in charge of compiling and creating a master workbook.
5. **Copy each of the lists (columns B & C) from the Step 2 spreadsheet into your own master Excel workbook (the one you'll eventually send back to the DataSF team).** Don't worry about duplicates quite yet. If there are exact duplicates, you'll see highlighting to indicate that. You can remove exact duplicates, but keep everything else even if you think it may be a duplicate. You'll review these with your team in the next

step.

6. **Send out your master spreadsheet to your team to review and set up a time to get together for 30 minutes to an hour.** You shouldn't need more than an hour.
7. **Review your brainstorm list together.** Spend a couple of minutes going through the list. Review your programs and divisions together and make sure you haven't left anything out.
8. **Assign the inventory.** At your meeting with your team, assign each person a group of datasets that they should inventory. This could simply break out by division, program or some other organizational theme that makes sense for your department. Make sure everyone knows exactly what they'll be doing next by going over Step 3 Dataset Inventory with them.

#### Quick tip

Use Column C (3. *Assigned to Complete*) in the **Step 2 Brainstorm** tab of your inventory workbook to assign responsibility for completing the inventory for particular datasets.

### Step 3: Inventory

9. **In your inventory workbook, click on the button "Prepare my Inventory".** This will take your awesome brainstorm list and pre-fill in the top rows of the worksheet "Step 3 Dataset Inventory" with the dataset names, your department, and the person assigned to complete the inventory for that dataset.
10. **Send the master spreadsheet to each of your brainstorming team members.** They should find their section
11. **Now you and your team can start populating across the datasets one by one.** If this list isn't too big, and you have good knowledge of the datasets, it's probably something you can start in on your own. Each field has directions included as a comment. Just mouse over each one to see what you should provide in each field.

#### Quick tip

There are a number of fields to cover for each dataset in the inventory and we know it could be overwhelming. It may help to **start with identifying the business contact (Data Steward) and technical contact (Data Custodian) first** for each dataset. That way if you're stuck, you'll have a handy reference on who to follow up with.

12. **Compile the individual inventories into your master after you've received these**



back from your assigned stakeholders.

13. On completion you can attach your inventory to an email and send to [datasf@sfgov.org](mailto:datasf@sfgov.org), subject: **Dataset Inventory Complete: <Your Department Name>**. You can also quickly do this by clicking on the link to send your inventory in the *Instructions* tab of your dataset inventory.

## High inventory complexity

### Step 2: Brainstorm

1. **Identify your key data stakeholders across the department.** These people should collectively have broad knowledge of data related to programs across the department. With big departments, it may be helpful to start with your org chart or other existing lists of programs to help organize your list of stakeholders.
2. **Download your department's copy of the inventory spreadsheet (MS Excel).** You should have the link to download in your email, if you cannot find it, please email [jason.lally@sfgov.org](mailto:jason.lally@sfgov.org).
3. **Reach out to your stakeholders to schedule a workshop and ask them to do some prep work ahead of time.** The workshop can be a couple of weeks out, it's important to get everyone comfortable. The prep work should be for them to compile a list of all the datasets they can think of. Give them the following questions to consider as they brainstorm:
  - a. What data populates your monthly or quarterly reports?
  - b. What departmental data is currently publicly available on DataSF or elsewhere online?
  - c. What data does your department use for internal performance and trend analysis?
  - d. What information is published as a performance metric?
  - e. What data is reported to federal, state or local agencies?
  - f. Talk with your Public Information Officer (PIO) - what data has been requested under Sunshine?
  - g. What data do other departments ask for?
  - h. What kinds of open data are similar agencies across the country publishing?
4. **Let the CDO and DataSF team know you're scheduling a workshop with the list of the stakeholders.** This will allow us to provide extra guidance and additional support to make your workshop successful.
5. **At the workshop, engage your stakeholders in an agenda that: builds understanding and generates as many possible datasets to inventory.** The workshop can vary in format, but generally, you should try to break out the brainstorming around sensible categories. You could organize brainstorming around programs, divisions, or data sources and systems.
6. **Before the end of your workshop, assign individuals responsibility of finishing the inventory.** Again, this could simply shake out by program area or division, but generally you'll want to have engaged and interested volunteers that can get you the completed inventory before the deadline.

### Step 3: Inventory

7. **During or shortly after your workshop, you should compile the master list of brainstormed datasets in your master inventory workbook.** You'll list the datasets

and the person responsible in columns B and C of the Step 2 Brainstorm tab.

8. **In your inventory workbook, click on the button “Prepare my Inventory.”** This will take your awesome brainstorm list and copy your department, dataset names and responsible party into *Step 3 Dataset Inventory*.
9. **Save your master inventory and send around to your stakeholders to complete their assigned datasets in Step 3 Dataset Inventory.**
10. **Compile the individual inventories into your master after you’ve received these back from your assigned stakeholders.**
11. **On completion you can attach your inventory to an email and send to [datasf@sfgov.org](mailto:datasf@sfgov.org), subject: Dataset Inventory Complete: <Your Department Name>.** You can also quickly do this by clicking on the link to send your inventory in the *Instructions* tab of your dataset inventory.