

Lab 4: SF Food Safety

A **dataset** is a collection of information or data. Data scientists uncover patterns and make predictions from datasets.

In this lab, we will explore a dataset of restaurant food safety scores for restaurants in San Francisco. The scores and violation information have been made available by the [San Francisco Department of Public Health](#).

After this lab, you should be comfortable with:

- Exploring tabular data using Google Sheets,
- Recognizing missing data and outliers, and
- Thinking critically about why data is collected and who it affects.

Question 1: Accessing and Importing Dataset

To get started, you'll want to [access the dataset directly](#). You should be brought to the following page, which gives some background information on the data:

The screenshot shows the DataSF Open Data Portal interface. At the top, there's a blue header with the text "DataSF Newsletter - New datasets, updates, and more! [Sign up today!](#)". Below the header, the page title is "DataSF | Open Data Portal" with a search bar on the right. A navigation bar includes links for Home, Browse, Developers, SF.gov, DataSF, Blog, and Contact Us, along with a "Sign In" button. The main content area has tabs for "About", "Data", and "Related Content", with "About" selected. To the right of the tabs are "Actions" and "Export" buttons. The dataset title "Health Inspection Scores (2024-Present)" is displayed, with a sub-tab "Health and Social Services". A note states: "Note: As of August 8, 2025, updates to this dataset have been paused as the Department of Public Health resolves an ongoing file formatting error." The "Last Updated" date is "September 2, 2025". A "SUMMARY..." section with a "Read more" link is visible. Below this, the "About this Dataset" section provides more details. It includes an "Updated" date of "September 2, 2025", "Data Last Updated" and "Metadata Last Updated" both as "September 2, 2025", and a "Date Created" of "May 7, 2025". It also shows "Views" of 706 and "Downloads" of 1,072. The "Department Metrics" section shows "Publishing Department" as "Public Health". The "Detailed Descriptive" section shows "Geographic unit" as "Street address". The "Internal Management" section shows "Inventory ID" as "DPH-0060". The "Publishing Details" section is partially visible at the bottom.

About this Dataset	
Updated	September 2, 2025
Data Last Updated	September 2, 2025
Metadata Last Updated	September 2, 2025
Date Created	May 7, 2025
Views	706
Downloads	1,072

Department Metrics	
Publishing Department	Public Health

Detailed Descriptive	
Geographic unit	Street address

Internal Management	
Inventory ID	DPH-0060

Publishing Details	
--------------------	--

If you scroll down a bit, you can also read through what columns are included in the dataset. Feel free to take a couple minutes to read through the columns.

[About](#) [Data](#) [Related Content](#) [Actions](#) [Export](#)












What's in this Dataset?

Rows
18.7K

Columns
26

Each row is a
Inspection

Columns (26)

Column Name	Description	API Field Name	Data Type
 inspection_date	Date of inspection	inspection_date	Floating Timestamp
 inspector	Inspector	inspector	Text
 district	Inspection district	district	Text
 subdistrict	Inspection sub-district	subdistrict	Text
 subsector	Inspection Sub-sector	subsector	Text
 permit_number	Permit number (if applicable)	permit_number	Text
 dba	"Doing business as" business name	dba	Text
 permit_type	Type of permit	permit_type	Text
 street_address	Street Address	street_address	Text
 street_address_clean	Cleaned Street Address	street_address_clean	Text
 inspection_type	Type of Inspection	inspection_type	Text

Question 1.1

Based on the page you just accessed, what is the source of the data? How was the dataset created?

Insert your response here.

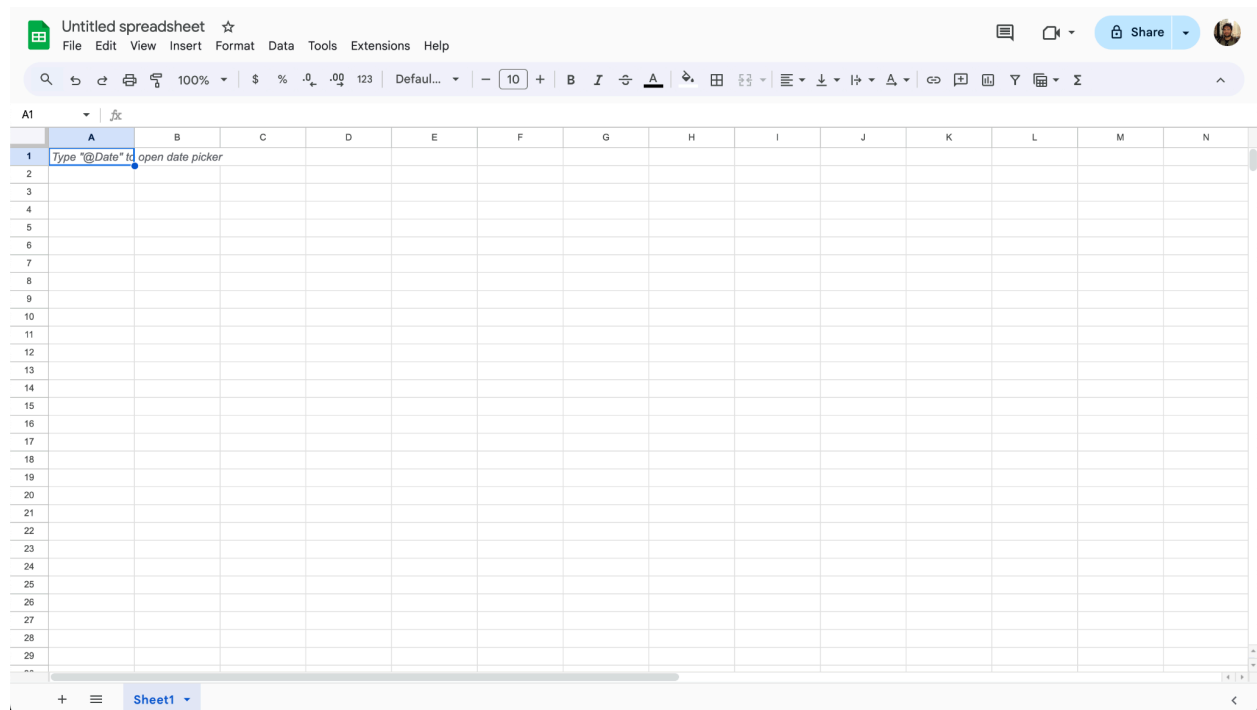
Question 1.2

What kinds of information does this dataset measure? Why might this data have been collected?

Insert your response here.

Question 2: Google Sheets

Once you've downloaded the .csv file, go to Google Sheets at <https://sheets.google.com/>. (Make sure you're logged into your Berkeley email.) Click "Blank Spreadsheet" to create a blank spreadsheet.



To import the .csv file, go to File → Import → Upload, then select the .csv file from your computer. Your imported spreadsheet should look something like this:

SF Health Inspection Scores (2024-)

File Edit View Insert Format Data Tools Extensions Help

100%

\$ % .00 123

Default...

- 10 +

B I T

U

A

Question 2.1

What do the rows represent? What do the columns represent?

Insert your response here.

Now, you'll want to create a [filter view](#), which is a named filter you can toggle on or off. To do so, click on Data → Create filter view, then click the three lines next to the column you want to filter to set the conditions.

SF Health Inspection Scores (2024-) ☆ 📁 🗑️ Help
File Edit View Insert Format Data Tools Extensions Help

100% | \$ % .00 .00 123 Default... | - 10 + | B I U A | 🗑️ 📄 📊 📈 📉 📌 📍 📎 📏 📐 📑 📗 📙 📚 📛 📜 📝 📞 📟 📠 📡 📢 📣 📤 📥 📦 📧 📨 📩 📪 📫 📬 📭 📮 📯 📰 📱 📲 📳 📴 📵 📶 📷 📸 📹 📺 📻 📼 📽 📾 📿 📠 📡 📢 📣 📤 📥 📦 📧 📨 📩 📪 📫 📬 📭 📮 📯 📰 📱 📲 📳 📴 📵 📶 📷 📸 📹 📺 📻 📼 📽 📾 📿

F3 | 91245

Temporary filter 2 Range: A1:Z18660 Save view X

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	inspection_dt	inspection_dt	district	subdistrict	subsector	permit_num	dba	permit_type	street_addr	street_addr	inspection_t	inspection_t	total_time	facility_ratin
1	2025/04/23 12:0	Zack Parsons	1	103	607	6734877	PREMIUM COM H36 - STADIUM 24 WILLIE MAY: 3RD ST & KING	Routine	1	150	Pass			
2	2025/04/22 12:0	Roy Bwogi	1	103	609	91245	TIFFANY'S CAF H24 - RESTAUR 3801 3RD ST	3801 3RD ST	Reinspection	2	25	Pass		
3	2025/04/22 12:0	Zack Parsons	1	104	227	107485	PO'BOYS KITCI H24 - RESTAUR 317 CONNECTI	317 CONNECTI	Routine	2	55	Pass		
4	2025/04/18 12:0	Roy Bwogi	1	103	607	6734023	HAPPY KING PI H26 - RESTAUR 227 KING ST	227 KING ST	Reinspection	2	95	Pass		
5	2025/04/17 12:0	Abel Simon	1	101	232	6743970	MEXI-CALI H79 - MOBILE F 103 HORNE AVI	103 HORNE AVI	New Ownership	1	30	Pass		
6	2025/04/17 12:0	Roy Bwogi	1	103	609	6913	XTRA OIL COM H07 - RETAIL M 3750 3RD ST	3750 3RD ST	Routine	1	40	Pass		
7	2025/04/15 12:0	Zack Parsons	1	103	609	FYF119398	SUBWAY #7154 H25 - RESTAUR 3801 3RD ST S	3801 3RD ST	Routine	2	65	Pass		
8	2025/04/14 12:0	Danny Lee	1	107 117B		115467	CAPITAL ONE C H24 - RESTAUR 101 POST ST	101 POST ST	Routine	1	60	Pass		
9	2025/04/10 12:0	Danny Lee	1	107 117B		A118913	TRISARA REST. H25 - RESTAUR 211 KEARNY ST	211 KEARNY ST	Reinspection	2	45	Pass		
10	2025/04/10 12:0	Farhan Khan	1	102	179	6733123	SAINT FRANK C H25 - RESTAUR 120 FOLSOM S	120 FOLSOM S	Complaint (I)	2	60	Pass		
11	2025/04/10 12:0	Zack Parsons	1	104	227	88326	LIVE OAK SCH H91 - SCHOOL 1555 MARIPOS	1555 MARIPOS	Routine	2	50	Pass		
12	2025/04/09 12:0	Amelia Castelli	1	106	257	6733435	RIZE UP BAKEFI H34 - PUSHCAF 147 FULTON ST	147 FULTON ST	Routine	1	15	Pass		
13	2025/04/09 12:0	Danny Lee	1	107 117B		80763	ROCK JAPANESE H26 - RESTAUR 614 PINE ST	614 PINE ST	Routine	1	75	Pass		
14	2025/04/09 12:0	Farhan Khan	1	102	179	99835	COMPASS @ LI H85 - EMPLOYE 222 2ND ST	222 2ND ST	Routine	1	15	Pass		
15	2025/04/09 12:0	Roy Bwogi	1	103	233	6733729	SCO'S H24 - RESTAUR 1730 YOSEMITI	1730 YOSEMITI	Routine	2	65	Pass		
16	2025/04/08 12:0	Danny Lee	1	107 117B		88498	MCDONALD'S H26 - RESTAUR 441 SUTTER ST	441 SUTTER ST	Routine	1	45	Pass		
17	2025/04/08 12:0	Roy Bwogi	1	103	607	98115	THE BUTCHER H24 - RESTAUR 57 DIVISION ST	57 DIVISION ST	Routine	2	70	Pass		
18	2025/04/07 12:0	Danny Lee	1	108 176C		91805	REMEDE SPA H67 - MASSAGE 125 3RD ST	125 3RD ST	Plan Check (I)	30	Pass			
19	2025/04/03 12:0	Philips Ossai	1	103	607	97856	THE WATERING H79 - MOBILE F 1379 4TH ST	1379 4TH ST	Structural	1	20			
20	2025/04/02 12:0	Danny Nguyen	1	103	607	A104207	RASA RASA H79 - MOBILE F 1379 4TH ST	1379 4TH ST	Reinspection	1	10			
21	2025/04/02 12:0	Roy Bwogi	1	103	233	6733969	FRANK GRIZZLI H25 - RESTAUR 5698 3RD ST	5698 3RD ST	Routine	2	70	Pass		
22	2025/04/02 12:0	Zack Parsons	1	104	226	97842	UNGRAFTED H26 - RESTAUR 2419 3RD ST	2419 3RD ST	Routine	2	70	Pass		
23	2025/04/01 12:0	Farhan Khan	1	102	179	103192	BON APPETIT H84 - EMPLOYE 333 BRANNAN	333 BRANNAN	Site Visit	1	15	Pass		
24	2025/04/11 12:0	Roy Bwogi	1	103	607	6743979	SOMISOMI H25 - RESTAUR 1350 04TH ST	1350 04TH ST	New Ownership	1	60	Pass		
25	2025/04/01 12:0	Philips Ossai	1	103	607	91166	SPARK SOCIAL H79 - MOBILE F 601 MISSION B	601 MISSION B	Structural	1	25			

+ ≡ Health_Inspection_Scores_(2024-Present)_20250912 <

SF Health Inspection Scores (2024-) ☆ 📁 🌐

File Edit View Insert Format Data Tools Extensions Help

100% | \$ % 0.00 123 | Default... | 10 | B I A

D1 | subdistrict

Temporary filter 1 Range: A1:Z18660 Save view

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	inspection_1	inspection_1	district	subdistrict	subsector	permit_num	dba	permit_type	street_addr	street_addr	inspection_1	inspection_1	total_time	facility_ratin
1	2025/04/23 12:0	Sort A to Z			607	6734877	PREMIUM COM H36 - STADIUM 24 WILLIE MAY: 3RD ST & KING	Routine				1	150	Pass
2	2025/04/22 12:0				609	91245	TIFFANY'S CAF H24 - RESTAUR 3801 3RD ST	Reinspection				2	25	Pass
3	2025/04/22 12:0	Sort Z to A			227	107465	PO'BOYS KITC H24 - RESTAUR 317 CONNECTI 317 CONNECTI	Routine				2	55	Pass
4	2025/04/18 12:0				607	6734023	HAPPY KING PI H26 - RESTAUR 227 KING ST	Reinspection				2	95	Pass
5	2025/04/17 12:0	Sort by color			232	6743970	MEXI-CALI H79 - MOBILE F 103 HORNE AVI 103 HORNE AVI	New Ownership				1	30	Pass
6	2025/04/17 12:0				609	69113	XTRA OIL COMI H07 - RETAIL M 3750 3RD ST	Routine				1	40	Pass
7	2025/04/15 12:0	Filter by color			609	FYF119398	SUBWAY #7154 H25 - RESTAUR 3801 3RD ST S 3801 3RD ST	Routine				2	65	Pass
8	2025/04/14 12:0				117B	115467	CAPITAL ONE C H24 - RESTAUR 101 POST ST	Routine				1	60	Pass
9	2025/04/10 12:0	Filter by condition			117B	A118913	TRISARA REST H25 - RESTAUR 211 KEARNY S1211 KEARNY S	Reinspection				2	45	Pass
10	2025/04/10 12:0	Filter by values			179	6733123	SAINT FRANK C H25 - RESTAUR 120 FOLSOM S 120 FOLSOM S	Complaint (I)				2	60	Pass
11	2025/04/10 12:0				227	88326	LIVE OAK SCH H91 - SCHOOL 1555 MARIPOS 1555 MARIPOS	Routine				2	50	Pass
12	2025/04/09 12:0	Select all 42 Clear			257	6733435	RIZE UP BAKEF H34 - PUSHCAF 147 FULTON ST 147 FULTON ST	Routine				1	15	Pass
13	2025/04/09 12:0				117B	80763	ROCK JAPANE H26 - RESTAUR 614 PINE ST	Routine				1	75	Pass
14	2025/04/09 12:0				179	99835	COMPASS @ LI H85 - EMPLOYE 222 2ND ST	Routine				1	15	Pass
15	2025/04/09 12:0				233	6733729	SCO'S H24 - RESTAUR 1730 YOSEMITT 1730 YOSEMITT	Routine				2	65	Pass
16	2025/04/08 12:0				117B	88498	MCDONALD'S H26 - RESTAUR 441 SUTTER S1441 SUTTER ST	Routine				1	45	Pass
17	2025/04/08 12:0				607	98115	THE BUTCHER H24 - RESTAUR 57 DIVISION ST 57 DIVISION ST	Routine				2	70	Pass
18	2025/04/07 12:0				176C	91805	REMEDE SPA H67 - MASSAG 125 3RD ST	Plan Check (I)				1	30	Pass
19	2025/04/03 12:0				607	97856	THE WATERINC H79 - MOBILE F 1379 4TH ST	Structural				1	20	
20	2025/04/02 12:0				607	A104207	RASA RASA H79 - MOBILE F 1379 4TH ST	Reinspection				1	10	
21	2025/04/02 12:0				233	6733969	FRANK GRIZZL H26 - RESTAUR 5698 3RD ST	Routine				2	70	Pass
22	2025/04/02 12:0				226	97842	UNGRAFTED H26 - RESTAUR 2419 3RD ST	Routine				2	70	Pass
23	2025/04/01 12:0				179	103192	BON APPETIT H84 - EMPLOYE 333 BRANNAN 333 BRANNAN	Site Visit				1	15	Pass
24	2025/04/11 12:0				607	6743979	SOMISOMI H25 - RESTAUR 1350 04TH ST	New Ownership				1	60	Pass
25	2025/04/01 12:0				607	91166	SPARK SOCIAL H79 - MOBILE F 601 MISSON B 601 MISSON B	Structural				1	25	
26	2025/04/01 12:0				609	111674	LADY M CAKE H33 - COMMIS 2045 MCKINNO 2045 MCKINNO	Routine				1	80	Pass
27	2025/04/01 12:0	Roy Bwogi	1	103	260	61016	PUPUSERIA EL H33 - TAKE-OU 785 MOSCOW 785 MOSCOW	Routine				2	70	Pass
28	2025/03/28 12:0	Roy Bwogi	1	105	230	6743661	SMOKE SOUL H07 - RETAIL M 4618 3RD ST	Routine				1	85	Pass
29	2025/03/27 12:0	Roy Bwogi	1	103	226	77185	RAMP RESTAU H26 - RESTAUR 855 TERRY FR 855 TERRY FR	Reinspection				2	35	Pass
30	2025/03/26 12:0	Zack Parsons	1	104	179	6734379	COMPASS @ LI H85 - EMPLOYE 222 2ND ST	Routine				1	30	Pass
31	2025/03/25 12:0	Farhan Khan	1	102	227	18951	NEW POTEREC H03 - RETAIL M 1301 18TH ST					1	90	Pass
32	2025/03/25 12:0	Luz Brown	1	104										

+ Health_Inspection_Scores_(2024-Present)_20250912

Question 2.2

Using a filter view, how many inspections in Mission Bay received a Pass?

Insert your response here.

Question 3: Exploring Variables

Question 3.1

Identify two categorical variables and two numerical values, and explain your reasoning.

Insert your response here.

Question 3.2

Which column(s) might be most useful for evaluating restaurant safety, and why?

Insert your response here.

Question 3.3

Pick a unit of analysis, and compare two specific values within that unit. What differences do you notice? What confounding factors should be considered when comparing them?

Insert your response here.

Question 4: What's Missing?

Question 4.1

Find a column with blank entries or missing data, and state the name of the column. How might this affect our analysis?

Insert your response here.

Question 5: Reflection

Question 5.1

Thinking about what you've observed, how might this data shape public understanding of food safety in San Francisco? Who benefits from it, and who might be harmed?

Insert your response here.

Done!

To submit, download this doc as a PDF (File → Download → PDF Document), then submit to [Gradescope](#).