

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. At the top, there's a blue header bar with the text "DataSF Newsletter – New datasets, updates, and more! Sign up today!". Below the header, the main navigation bar includes links for Home, Browse, Developers, SF.gov, DataSF, Blog, Contact Us, and a "Sign In" button. The main content area has tabs for "About", "Data", and "Related Content". The "About" tab is selected, displaying information for the "Health Inspection Scores (2024-Present)" dataset. The dataset title is "Health Inspection Scores (2024-Present)" under the "Health and Social Services" category. 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." To the right, the last update is listed as "September 2, 2025". Below this, there's a "SUMMARY..." section with a "Read more" link, and a "Actions" dropdown menu with an "Export" option. The "Data" tab shows a detailed view of the dataset's metadata. It includes sections for "Updated" (September 2, 2025), "Department Metrics" (Publishing Department: Public Health), "Detailed Descriptive" (Geographic unit: Street address), "Internal Management" (Inventory ID: DPH-0060), and "Publishing Details". It also lists "Views: 706" and "Downloads: 1,072".

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.

The screenshot shows a dataset summary page with the following sections:

- 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.

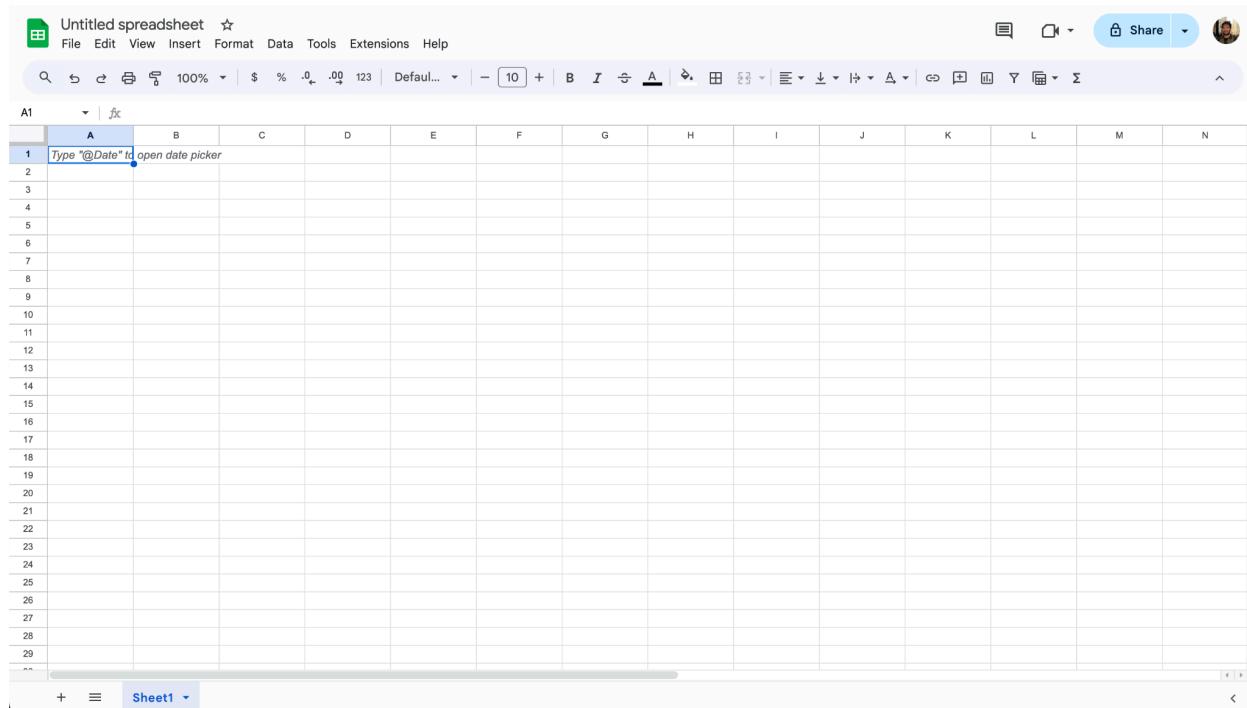
Question 1.2

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:

A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	inspection_date	district	subdistrict	subsector	permit_number	dba	permit_type	street_address	street_address_1	inspection_type	inspection_freq	total_time	facility_rating_st
2	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
3	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
4	2025/04/22 12:0	Zack Parsons	1	104	227	107465 PO'BOYS KITCH H24 - RESTAUR 317 CONNECTI	317 CONNECTI	Routine			2	55	Pass
5	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
6	2025/04/17 12:0	Abel Simon	1	101	232	6743970 MEXI-CALI H79 - MOBILE F 103 HORNE AVI 103 HORNE AVI New Ownership	103 HORNE AVI	New Ownership			1	30	Pass
7	2025/04/17 12:0	Roy Bwogi	1	103	609	6913 XTRA OIL COMI H07 - RETAIL M 3750 3RD ST	3750 3RD ST	Routine			1	40	Pass
8	2025/04/15 12:0	Zack Parsons	1	103	609	699 FYF119398 SUBWAY #7154 H25 - RESTAUR 3801 3RD ST	3801 3RD ST	Routine			2	65	Pass
9	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
10	2025/04/10 12:0	Danny Lee	1	107	117B	A118913 TRISARA REST H25 - RESTAUR 211 KEARNY ST 211 KEARNY ST	211 KEARNY ST	Reinspection			2	45	Pass
11	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
12	2025/04/10 12:0	Zack Parsons	1	104	227	88326 LIVE OAK SCH H91 - SCHOOL 1555 MARIPOS	1555 MARIPOS	Routine			2	50	Pass
13	2025/04/09 12:0	Amelia Castelli	1	106	257	6733435 RIZE UP BAKER H34 - PUSHCAF 147 FULTON ST	147 FULTON ST	Routine			1	15	Pass
14	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
15	2025/04/09 12:0	Farhan Khan	1	102	179	99835 COMPASS @ LI H85 - EMPLOYEE 222 2ND ST	222 2ND ST	Routine			1	15	Pass
16	2025/04/09 12:0	Roy Bwogi	1	103	233	6733729 SCO'S H24 - RESTAUR 1730 YOSEMITI	1730 YOSEMITI	Routine			2	65	Pass
17	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
18	2025/04/08 12:0	Roy Bwogi	1	103	607	96115 THE BUTCHER H24 - RESTAUR 57 DIVISION ST	57 DIVISION ST	Routine			2	70	Pass
19	2025/04/07 12:0	Danny Lee	1	108	176C	91805 REMEDE SPA H67 - MASSAGI 125 3RD ST	125 3RD ST	Plan Check (I)			30	Pass	
20	2025/04/03 12:0	Philips Ossai	1	103	607	97856 THE WATERINC H79 - MOBILE F 1379 4TH ST	1379 4TH ST	Structural			1	20	
21	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	
22	2025/04/02 12:0	Roy Bwogi	1	103	233	6733969 FRANK GRIZZL H25 - RESTAUR 5698 3RD ST	5698 3RD ST	Routine			2	70	Pass
23	2025/04/02 12:0	Zack Parsons	1	104	226	97842 UNGRAFTED H26 - RESTAUR 2419 3RD ST	2419 3RD ST	Routine			2	70	Pass
24	2025/04/01 12:0	Farhan Khan	1	102	179	103192 BON APPETIT H84 - EMPLOYEE 333 BRANNAN	333 BRANNAN	Site Visit			1	15	Pass
25	2025/04/11 12:0	Roy Bwogi	1	103	607	6743979 SOMISOM H25 - RESTAUR 1350 04TH ST	1350 04TH ST	New Ownership			1	60	Pass
26	2025/04/01 12:0	Philips Ossai	1	103	607	91166 SPARK SOCIAL H79 - MOBILE F 601 MISSON	601 MISSON	B/Structure			1	25	
27	2025/04/01 12:0	Roy Bwogi	1	103	609	111674 LADY M CAKE H33 - COMMIS 2045 MCKINNO	2045 MCKINNO	Routine			1	80	Pass
28	2025/03/28 12:0	Roy Bwogi	1	105	260	61016 PUPUSERIA EL H28 - TAKE-OU 785 MOSCOW	785 MOSCOW	Routine			2	70	Pass
29	2025/03/27 12:0	Roy Bwogi	1	103	230	6743661 SMOKE SOUL H07 - RETAIL M 4618 3RD ST	4618 3RD ST	Routine			1	85	Pass

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.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N		
1	inspection_i	=	inspection_i	=	district	subdistrict	subsector	permit_num	dba	permit_type	street_addr	street_addr	inspection_i	=		
2	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	
3	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	
4	2025/04/22	12:0	Zack Parsons	1	104	227	107465	PO'BOYS KITCH H24 - RESTAUR 317 CONNECTI	317 CONNECTI	Routine			2	55	Pass	
5	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	
6	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		
7	2025/04/17	12:0	Roy Bwogi	1	103	609	6913	XTRA OIL COMI H07 - RETAIL M	3750 3RD ST	3750 3RD ST	Routine		1	40	Pass	
8	2025/04/15	12:0	Zack Parsons	1	103	609	FYF119394	SUBWAY #7154 H25 - RESTAUR 3801 3RD ST	3801 3RD ST	Routine			2	65	Pass	
9	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	
10	2025/04/10	12:0	Danny Lee	1	107	117B	A118913	TRISARA REST H25 - RESTAUR 211 KEARNY S1211 KEARNY S1	Reinspection			2	45	Pass		
11	2025/04/10	12:0	Farhan Khan	1	102	179	6733123	SAINT FRANK (H25 - RESTAUR 120 FOLSOM S' 120 FOLSOM S' Complaint (I)					2	60	Pass	
12	2025/04/10	12:0	Zack Parsons	1	104	227	88326	LIVE OAK SCH: H91 - SCHOOL	1555 MARIPOS	1555 MARIPOS	Routine		2	50	Pass	
13	2025/04/09	12:0	Amelia Castelli	1	106	257	6733435	RIZE UP BAKER H34 - PUSHCAF 147 FULTON ST	147 FULTON ST	Routine			1	15	Pass	
14	2025/04/09	12:0	Danny Lee	1	107	117B	80763	ROCK JAPANEH H26 - RESTAUR 614 PINE ST	614 PINE ST	Routine			1	75	Pass	
15	2025/04/09	12:0	Farhan Khan	1	102	179	99835	COMPASS @ LI H85 - EMPLOYEE 222 2ND ST	222 2ND ST	Routine			1	15	Pass	
16	2025/04/09	12:0	Roy Bwogi	1	103	233	6733729	SCO'S	H24 - RESTAUR 1730 YOSEMITI 1730 YOSEMITI	Routine			2	65	Pass	
17	2025/04/08	12:0	Danny Lee	1	107	117B	88498	MCDONALD'S	H26 - RESTAUR 441 SUTTER S1441 SUTTER S1	Routine			1	45	Pass	
18	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	
19	2025/04/07	12:0	Danny Lee	1	108	176C	91805	REMEDI SPA	H67 - MASSAGI 125 3RD ST	125 3RD ST	Plan Check (I)			30	Pass	
20	2025/04/03	12:0	Philips Ossai	1	103	607	97856	THE WATERINC	H79 - MOBILE F 1379 4TH ST	1379 4TH ST	Structural			1	20	
21	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	
22	2025/04/02	12:0	Roy Bwogi	1	103	233	6733969	FRANK GRIZZL	H25 - RESTAUR 5698 3RD ST	5698 3RD ST	Routine			2	70	Pass
23	2025/04/02	12:0	Zack Parsons	1	104	226	97842	UNGRAFTED	H26 - RESTAUR 2419 3RD ST	2419 3RD ST	Routine			2	70	Pass
24	2025/04/01	12:0	Farhan Khan	1	102	179	103192	BON APPETIT H48 - EMPLOYEE 333 BRANNAN	:333 BRANNAN : Site Visit					1	15	Pass
25	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
26	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	

SF Health Inspection Scores (2024-) ⚡ 📁 🌐

File Edit View Insert Format Data Tools Extensions Help

D1 | fx subdistrict

Temporary filter 1 Range: A1:Z18660

Save view X

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	inspection_id	inspection_id	district	subdistrict	subsector	permit_num	dba	permit_type	street_addr	street_addr	inspection_id	inspection_id	total_time	facility_rating
2	2025/04/23 12:00		Sort A to Z		607	6734877	PREMIUM COM H36 - STADIUM 24 WILLIE MAY: 3RD ST & KING	Routine			1	150	Pass	
3	2025/04/22 12:00				609	91245	TIFFANY'S CAF H24 - RESTAUR 3801 3RD ST	3801 3RD ST	Reinspection		2	25	Pass	
4	2025/04/22 12:00		Sort Z to A		227	107465	PO'BOYS KITCH H24 - RESTAUR 317 CONNECT	317 CONNECT	Routine		2	55	Pass	
5	2025/04/18 12:00				607	6734023	HAPPY KING PI H26 - RESTAUR 227 KING ST	227 KING ST	Reinspection		2	95	Pass	
6	2025/04/17 12:00		Sort by color		232	6743970	MEXI-CALI H79 - MOBILE F 103 HORNE AVI	103 HORNE AVI	New Ownership		1	30	Pass	
7	2025/04/17 12:00				609	6913	XTRA OIL COM H07 - RETAIL M 3750 3RD ST	3750 3RD ST	Routine		1	40	Pass	
8	2025/04/15 12:00		Filter by color		609	FYF11939	SUBWAY #7154 H25 - RESTAUR 3801 3RD ST	S'3801 3RD ST	Routine		2	65	Pass	
9	2025/04/14 12:00				117B	115467	CAPITAL ONE H24 - RESTAUR 101 POST ST	101 POST ST	Routine		1	60	Pass	
10	2025/04/10 12:00		▶ Filter by condition		117B	A118913	TRISARA REST H25 - RESTAUR 211 KEARNY S1211 KEARNY S1	Reinspection			2	45	Pass	
11	2025/04/10 12:00		▼ Filter by values		179	6733123	SAINT FRANK (H25 - RESTAUR 120 FOLSOM S' 120 FOLSOM S'	Complaint (I)			2	60	Pass	
12	2025/04/10 12:00		Select all 42 - Clear	Displaying 42	227	88326	LIVE OAK SCH H91 - SCHOOL 1555 MARIPOS	1555 MARIPOS	Routine			2	50	Pass
13	2025/04/09 12:00				257	6733435	RIZE UP BAKER H34 - PUSHCAF 147 FULTON ST	147 FULTON ST	Routine		1	15	Pass	
14	2025/04/09 12:00				117B	80763	ROCK JAPANESE H26 - RESTAUR 614 PINE ST	614 PINE ST	Routine		1	75	Pass	
15	2025/04/09 12:00				179	99835	COMPASS @ LI H85 - EMPLOYEE 222 2ND ST	222 2ND ST	Routine		1	15	Pass	
16	2025/04/09 12:00				233	6733729	SCO'S H24 - RESTAUR 1730 YOSEMITI	1730 YOSEMITI	Routine		2	65	Pass	
17	2025/04/08 12:00		✓ (Blanks)		117B	88498	MCDONALD'S H26 - RESTAUR 441 SUTTER S1441 SUTTER ST	S1441 SUTTER ST	Routine		1	45	Pass	
18	2025/04/08 12:00		✓ 101		607	98115	THE BUTCHER H24 - RESTAUR 57 DIVISION ST	57 DIVISION ST	Routine		2	70	Pass	
19	2025/04/07 12:00				176C	91805	REMDE SPA H67 - MASSAGI 125 3RD ST	125 3RD ST	Plan Check (I)			30	Pass	
20	2025/04/03 12:00		✓ 102		607	97856	THE WATERINC H79 - MOBILE F 1379 4TH ST	1379 4TH ST	Structural		1	20		
21	2025/04/02 12:00		✓ 103		607	A104207	RASA RASA H79 - MOBILE F 1379 4TH ST	1379 4TH ST	Reinspection		1	10		
22	2025/04/02 12:00				233	6733969	FRANK GRIZZL H25 - RESTAUR 5698 3RD ST	5698 3RD ST	Routine		2	70	Pass	
23	2025/04/02 12:00				226	97842	UNGRAFTED H26 - RESTAUR 2419 3RD ST	2419 3RD ST	Routine		2	70	Pass	
24	2025/04/01 12:00				179	103192	BON APPETIT (H84 - EMPLOYEE 333 BRANNAN : 333 BRANNAN : Site Visit	333 BRANNAN :			1	15	Pass	
25	2025/04/11 12:00				607	6743979	SOMISOMI H25 - RESTAUR 1350 04TH ST	1350 04TH ST	New Ownership		1	60	Pass	
26	2025/04/01 12:00				607	91166	SPARK SOCIAL H79 - MOBILE F 601 MISSION BA 601 MISSION BA	601 MISSION BA	Structural			25		
27	2025/04/01 12:00 Roy Bwogi	1	103		609	111674	LADY M CAKE I H33 - COMMIS 2045 MCKINNO	2045 MCKINNO	Routine			1	80	Pass
28	2025/03/28 12:00 Roy Bwogi	1	105		260	61016	PUPSERIA EL H28 - TAKE-OUT 785 MOSCOW	785 MOSCOW	Routine		2	70	Pass	
29	2025/03/27 12:00 Roy Bwogi	1	103		230	6743661	SMOKE SOUL H07 - RETAIL M 4618 3RD ST	4618 3RD ST	Routine		1	85	Pass	
30	2025/03/26 12:00 Zack Parsons	1	104		226	77185	RAMP RESTAUH26 - RESTAUR 855 TERRY FR/855 TERRY FR/	855 TERRY FR/	Reinspection		2	35	Pass	
31	2025/03/25 12:00 Farhan Khan	1	102		179	6734379	COMPASS @ LI H85 - EMPLOYEE 222 2ND ST	222 2ND ST	Routine		1	30	Pass	
32	2025/03/25 12:00 Luz Brown	1	104		227	18951	NEW POTEREC H03 - RETAIL M 1301 18TH ST	1301 18TH ST			1	90	Pass	

+ = 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?

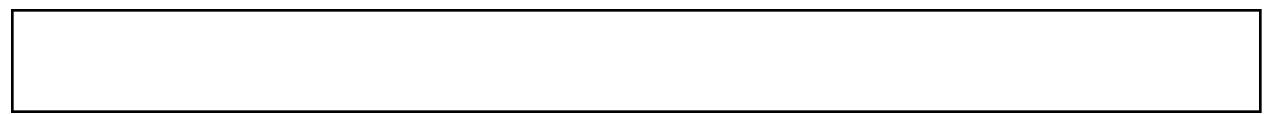
Insert your response here.

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](#).