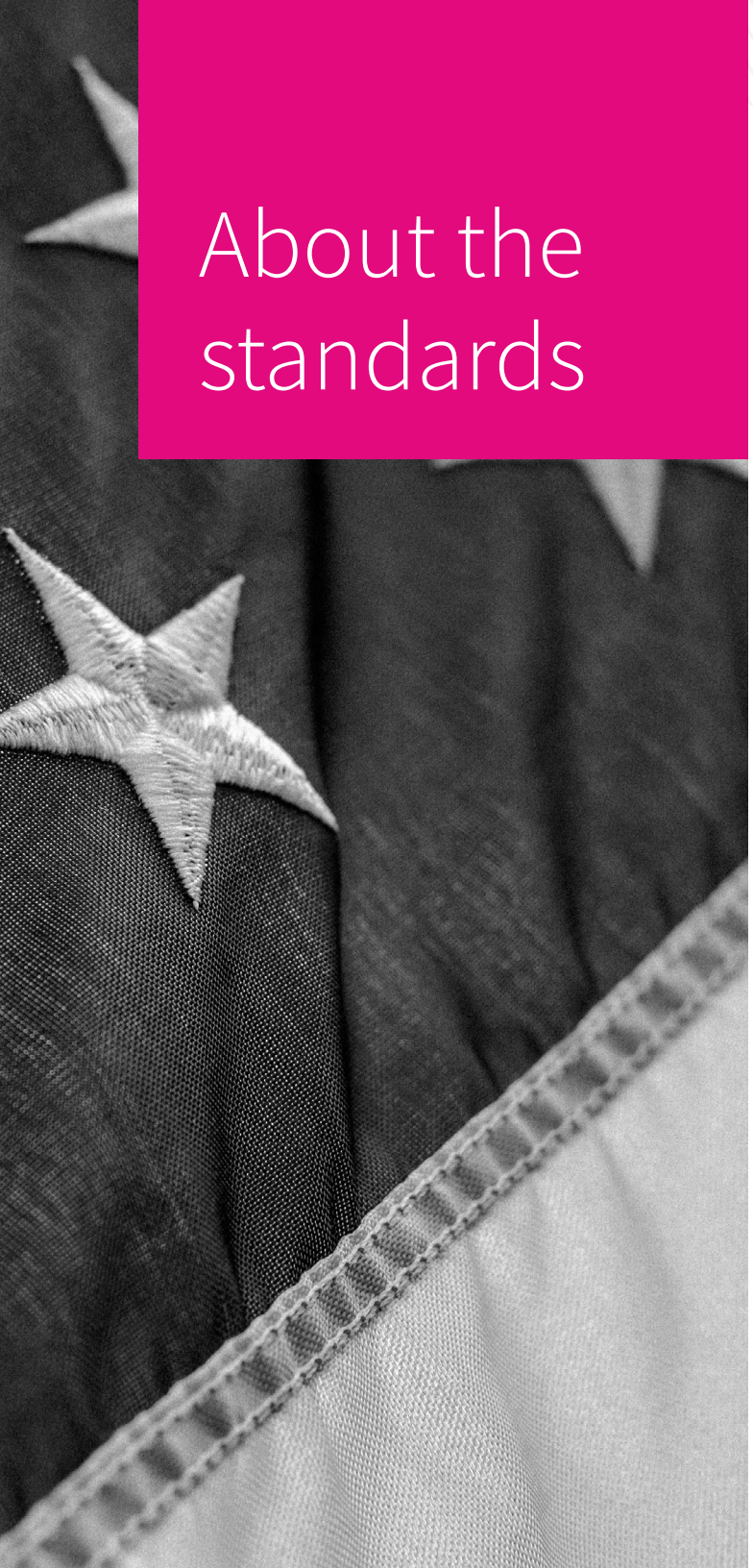
USAFACTS

Standards for excellent data products

DETAILED USER GUIDE



About the standards

Introduction

This document contains standards to help government agencies share data with the public by building data products that are user-friendly and transparent. A data product is anything powered by data: a map, a series of data tables, a report, a dashboard, or another interactive tool. These standards describe characteristics of excellent data products and are applicable to the many kinds of data shared by the government.

The Federal Data Excellence (FDE) program — developed by USAFacts and the Partnership for Public Service — has three priorities:

- Providing a set of standards to government data experts seeking to maximize the impact of their work by making it more useful to the public
- Building a community of public data stewards who can share successes and challenges of data product development, support inter-agency collaboration, and inspire new ideas for supporting the user experience
- Recognizing best practices and excellent data products published by agency teams

We hope that having a set of best practices will help reduce the burden of deciding how best to present a data product that is focused on user needs.

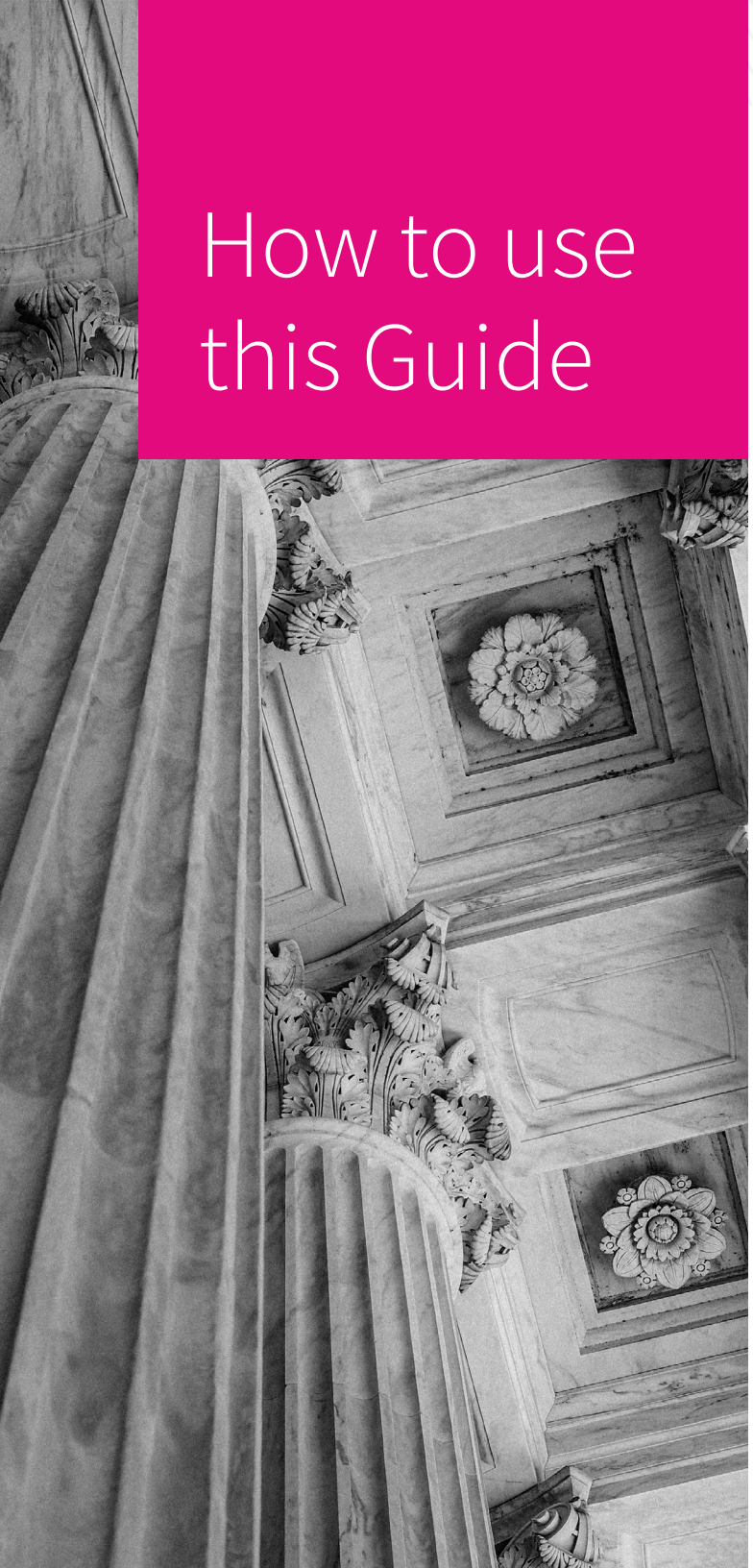
Stay in touch

Please reach out to fde@usafacts.org to ask questions about the program, share your experiences implementing the standards, and provide feedback on how to refine or make them more universally applicable. We have many ideas for modifying and expanding the program in 2025 and want your input.

Also sign up for our recurring newsletter about the Federal Data Excellence program. The newsletter will feature:

- Additional findings gleaned from our inaugural evaluation of 55 data products from 40 agencies or bureaus
- Updates on the 2025 program
- Opportunities to provide feedback and learn from others confronting similar challenges at remote or in-person meetings.

How to use this Guide



Frequently Asked Questions

Q: Do these standards consider data quality?

A: While these standards address the quality of data presentation, accessibility, and usability, they *do not* address quality of data collection or statistical analysis. We'll leave that to the expert civil servants who ensure that data are methodologically sound and ready for publication. These standards are only concerned with the last mile of data delivery to the public: how that high-quality data is packaged up and shared.

Q: Why did USAFacts and the Partnership for Public Service create the Federal Data Excellence program?

A: As a not-for-profit, nonpartisan organization focused on making government data more accessible, USAFacts works with federal data products daily. We've developed a deep understanding of which products model best practices and which need improvement. While agency data is obligated to be open by default, we believe that data products must take additional steps to truly be useful tools for the public.

The Partnership for Public Service brings decades of deep expertise in good governance and building communities of practice within the federal government. Leveraging our combined understanding of government data and community of data experts, we're creating these standards to provide a basis for data product development across domains.

Q: How did you develop these standards?

A: We began by reviewing existing resources inside and outside of government to see if standards existed that dictated attributes of good data products. We then kicked off a series of engagements in 2024 with agency leadership, data product owners, and other federal data experts to verify the need for standards and understand use cases. Through small group sessions and two larger convenings, we presented five different versions of these standards to gather and incorporate feedback, test, and make needed adjustments. The result is a set of standards that are broad enough to be widely applicable but specific enough to guide decision-making.

Q: Are the criteria in the standards equally important?

A: We believe that a truly excellent data product would be able to answer "yes" to each of the questions posed by these standards. Part of our engagement with the federal data community included identifying weighting for each standard. Each of the questions presented in these guidelines are weighted based on these expert opinions. We use these weights to evaluate nominations for FDE recognition, but they do not appear in the standards themselves.

How to use this Guide

Frequently Asked Questions

Q: Some of these are difficult to implement!

A: While that's not a question, we do understand. Federal data producers face constraints related to competing priorities, resource availability, and the nature of the data itself. Part of our engagement with the federal data community focused on this question; these standards account for the tradeoffs between difficulty of implementation and importance for the user experience. While some standards will be harder to implement and others easier, all are worthwhile. Ultimately, the teams creating these data products have the best perspective on how to incorporate these standards, whether in full or in part, or all at once or over time.

Q: Who do you hope uses these guidelines?

A: Any government data expert seeking to maximize the impact of their work by making it more useful to the public. The standards are focused on the federal government, but data practitioners in state and local government, non-profits, and even corporations can benefit from incorporating the concepts presented here.

Q: How and where would you hope people implement these guidelines?

A: Use this guide when preparing to package and share data with the public. Every new data product or refresh is an opportunity to consider what elements are included, how they connect to each other, and what form they take. We hope these standards provide a functional framework and a fresh perspective for evaluating and improving data products.

Q: Why is it important to have these standards?

A: To increase transparency and trust in data and by extension, the government agencies producing the data products.

Evaluation criteria

Data accessibility

Offers data formats that are easy to download and interact with, and includes clear descriptions that orient the user to the data product and collection program.

JUMP TO PAGE →

Helpfulness to the user

Provides features that meet common user needs.

JUMP TO PAGE →

Interoperability

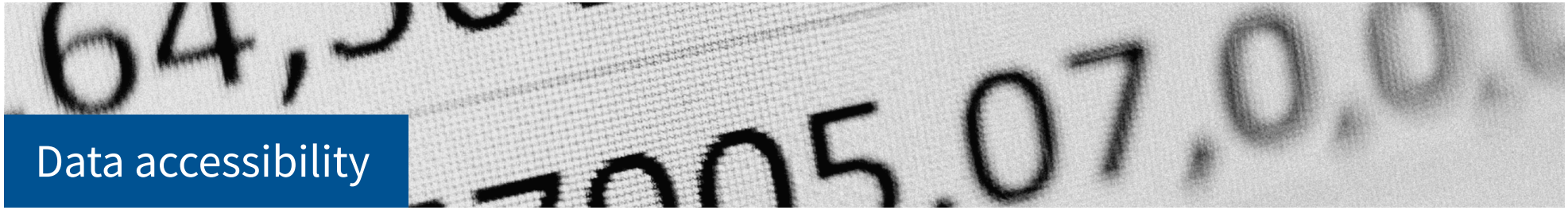
Includes metadata and variables that are both human-understandable and machine-readable to facilitate integration with other data sources.

JUMP TO PAGE →

Timeliness of data

Appropriately notes and discusses the temporal characteristics of the data and timelines of the data collection program.

JUMP TO PAGE →



Data accessibility

01 Can you use filters or similar tools in a dashboard, viewer, or other interactive environment to look at pieces of, rather than the whole, dataset?

☐ LEARN MORE →

02 Is there a landing page with information about the data product that is clearly linked from the data product?

☐ LEARN MORE →

03 Is the writing on the data landing page "clear, simple, meaningful, and jargon-free?"

☐ LEARN MORE →

04 Is the data available for download in rows and columns (e.g., .csv, .xlsx)?

☐ LEARN MORE →

05 Is API access to the data available and clearly indicated on the data product or landing page?

☐ LEARN MORE →

06 Is an FTP site available to support bulk downloads and clearly indicated on the data product or landing page?

☐ LEARN MORE →

01 | Can you use filters or similar tools in a dashboard, viewer, or other interactive environment to look at pieces of, rather than the whole, dataset?

RETURN TO FRAMEWORK

Additional directions and context

This question assesses the interactivity of the data product, which gives the user control over their experience while providing constraints. Interactivity can be relatively simple, as with filters, or much more complex, as with data dashboards, maps, or other visualizations. No matter the level of complexity, an interactive data product stands in contrast to static data products, which only allow the user a single view of the data.

Motivating intent behind the question

Allowing the user to interact with the data is valuable. Shaping, querying, zooming to areas or indicators of interest, or other options that allow a user to customize the experience to focus on what is important to them is a good practice.

Example

The California State Water Resources Control Board provides an example of what a simple interactive data product might look like. Note the ability to filter by water system type/status/source water type and the ability to produce county-wide results. This lets the user view only the statistics of interest, rather than the entire dataset.

An example of a data product that meets the standard

Source: Safe Drinking Water Information System, California State Water Resources Control Board

See more here

02 | Is there a landing page with information about the data product that is clearly linked from the data product?

← RETURN TO FRAMEWORK

Additional directions and context

A landing page can take different forms, but fundamentally it is a central hub that provides information about the data product and links to important features and resources that are useful for understanding and working with the data. A landing page may contain the data product itself, or clearly link to it.

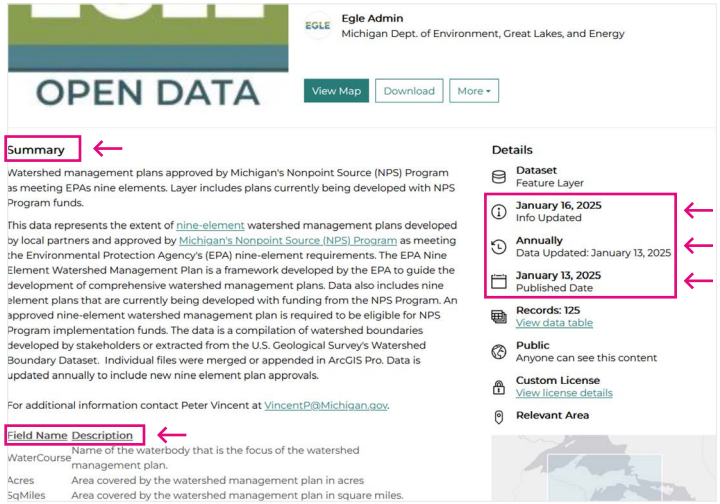
Motivating intent behind the question

Collecting key information in one easy-to-navigate spot helps the user. Data products — whether a simple .csv download or a feature-laden data dashboard or viewer — can benefit from having a hub from which the user can learn about and access the data.

Example

Michigan's EGLE department has a landing page that summarizes the data contained in the watershed management plan map. It describes the data, available fields, and attributes. It also provides external links about watershed management plans, who to contact with questions, and information on data recency. The landing page links to the map ("View Map") where users can explore the data and download data.

An example of a data product that meets the standard



Clicking "VIew Full Details" in the data product takes the user back to the landing page, making navigation simple. This two-way navigation between the data and landing page is particularly important because users who find the data through search (rather than navigating an agency site) may start with the map data and eventually need information from the landing page.

Source: Approved and Pending Nine-Element Watershed Management Plans, EGLE Maps and Data

See more here

03 | Is the writing on the data landing page "clear, simple, meaningful, and jargon-free?"

← RETURN TO FRAMEWORK

Additional directions and context

If there is no landing page, leave the box unchecked.

Motivating intent behind the question

This question is assessing for compliance with the Plain Writing Act of 2010, and borrows language from an OMB memo describing how agencies should communicate with the public. Since "clear, simple, meaningful, and jargon-free" are subjective terms, finding people to review draft language who are not working with the data product in question can be a helpful practice. Following these principles helps to educate the broadest possible audience and extend the impact of the data product.

Example

The Healthy Kids Colorado Survey Dashboard starts the user on a "Welcome!" landing page that clearly describes (without jargon) the survey program and the kinds of data the people can find in the data product. It is also clear to the user where they can interact with and download the data.

An example of a data product that meets the standard

Source: Healthy Kids Colorado Survey Dashboard, Colorado Department of Public Health and the Environment
See more here

04 | Is the data available for download in rows and columns (e.g., .csv, .xlsx)?

RETURN TO FRAMEWORK

Additional directions and context

Data structured as rows and columns (i.e., a table) is the most recognizable and useful download format for most people. Tables are easy to work with for programmers and non-programmers alike.

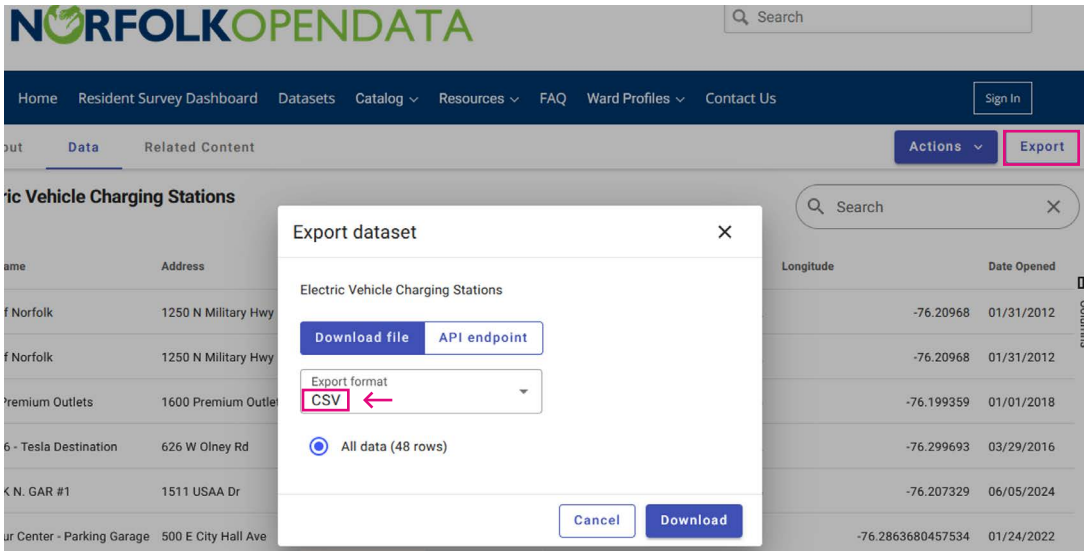
Motivating intent behind the question

Data products, whether static or interactive, should always provide access to the underlying data so that people can work with the numbers — .csv and .xlsx formats are analysis-ready and useful to a broad user base. While there are techniques to convert .pdfs, images, and .txt files to tabular formats, some users won't be familiar with them, and all require extra work.

Example

Norfolk Open Data shows one way to provide data in rows and columns. The user viewing the data is given a prominent "Export" button that offers the opportunity to select formats — including .csv — to support their needs. Of note here, deselecting columns in the data view does not prompt the interface to offer a download of the dataset that includes only the variables of interest. Offering filtered datasets for download (e.g., removing unwanted columns or filtering them for certain values) can be a helpful feature, particularly for large datasets.

An example of a data product that meets the standard



Source: Electric Vehicle Charging Stations, Norfolk Open Data
See more here

05 | Is API access to the data available and clearly indicated on the data product or landing page?

← RETURN TO FRAMEWORK

Additional directions and context

Check the box only if AI access to the data is available and also visibly linked to the data product or landing page. Clearly noting when the data in a product is available through an API is best practice.

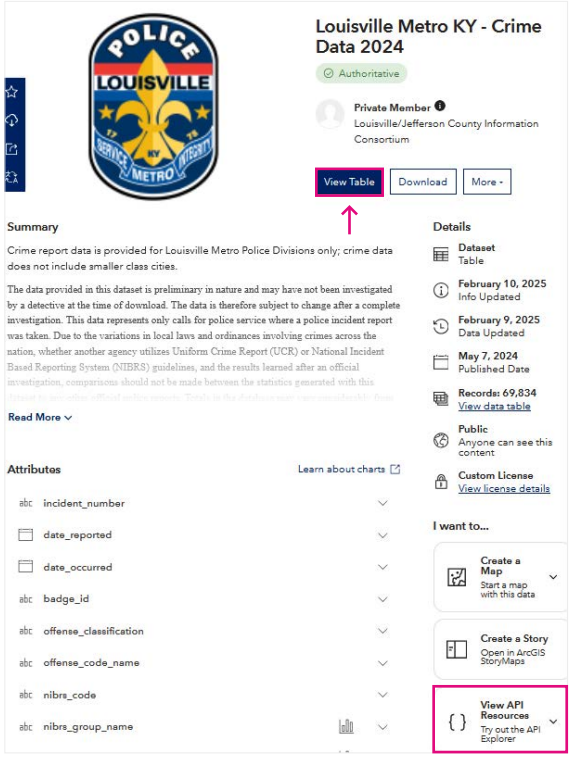
Motivating intent behind the question

Though an API can be challenging to build, it adds value for users who want to access data programmatically. Typically, an API is built to provide access to more data than is available in a specific data product. Often, an API will exist on an agency or program level, and clearly signaling that the data is also available via an agency API is helpful to a user who may not know to look to the parent agency or program for API access.

Example

Louisville Metro Open Data makes API access clearly visible from the landing page that describes the data. Users can do initial data explorations by clicking “View Table” and can also launch the API Explorer, which provides a link to detailed documentation as well as a build-your-own query feature.

An example of a data product that meets the standard



Source: Louisville KY Metro Crime Data 2024, Louisville Metro Open Data
See more here

06 | Is an FTP site available to support bulk downloads and clearly indicated on the data product or landing page?

← RETURN TO FRAMEWORK

Additional directions and context

Secure File Transfer Protocol (SFTP) sites provide a similar user experience to FTP sites but offer enhanced security. Only check the box if an FTP or SFTP site for data transfer is present and also visibly linked to the data product or landing page.

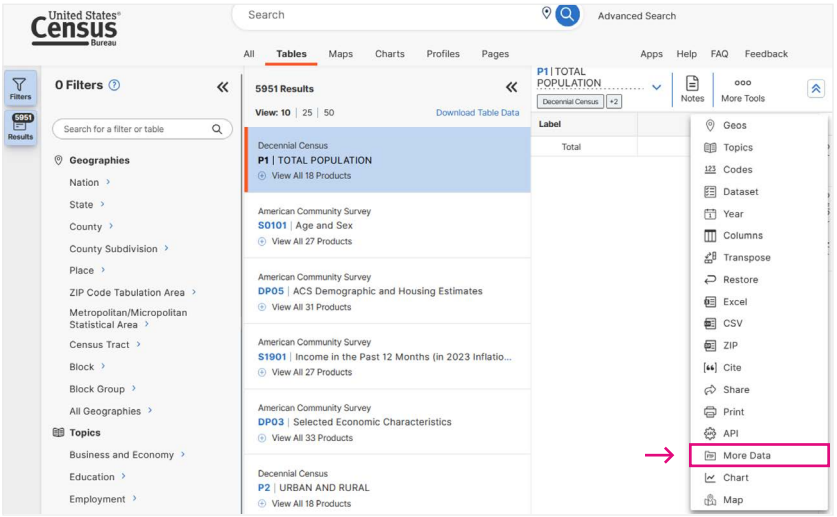
Motivating intent behind the question

A File Transfer Protocol (FTP) site is another way to acquire large amounts of data. The familiar file structure of an FTP site allows non-technical users to download data in bulk.

Example

The US Census Bureau links to its FTP site from the table explorer at data.census.gov.

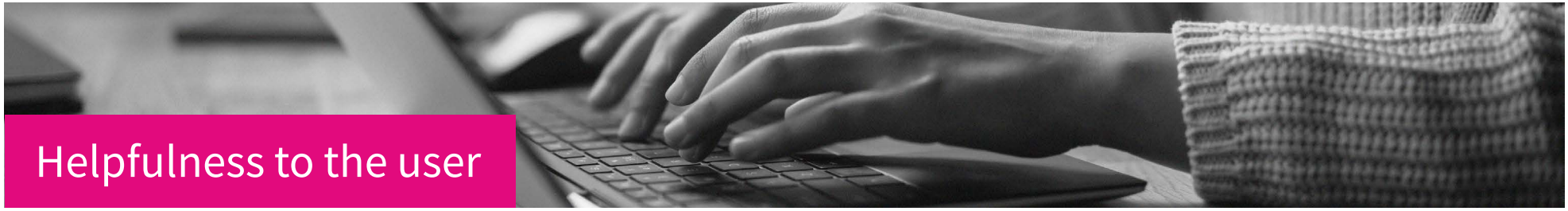
An example of a data product that meets the standard



The FTP directory looks like this:

Source: US Census Bureau Table Explorer

See more here



Helpfulness to the user

01 Are variable definitions provided via codebook, data dictionary, glossary, or key terms list?

☐ LEARN MORE →

02 Are variable definitions clearly linked or displayed?

☐ LEARN MORE →

03 Does the data product address uncertainty through statistical methods or discussion?

☐ LEARN MORE →

04 Are popular or useful cross-sections, takeaways, or views of the data provided?

☐ LEARN MORE →

05 Are there accompanying reports that include relevant data tables?

☐ LEARN MORE →

06 Is a point of contact provided?

☐ LEARN MORE →

07 Is a suggested citation provided?

☐ LEARN MORE →

01 | Are variable definitions provided via codebook, data dictionary, glossary, or key terms list?

← RETURN TO FRAMEWORK

Additional directions and context

This is a specific piece of metadata that belongs with every data product. Defining variables and programs lets the user understand the data being presented without making assumptions.

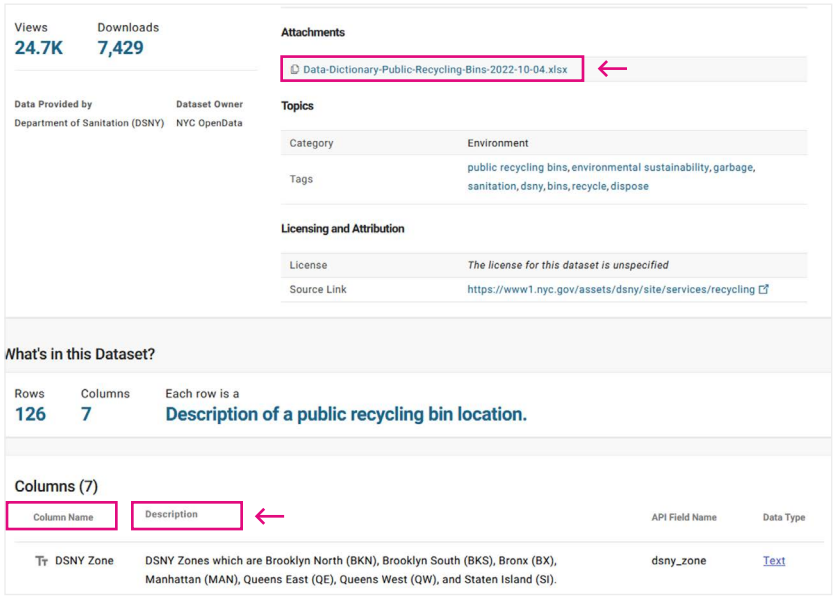
Motivating intent behind the question

Users need unambiguous explanations of variables to understand the data provided. Providing definitions is good, specifying the types of data each variable can take (e.g., text/string, numeric, categorical) is better, and being specific about those values (e.g., a two decimal number, three categories named 'X,' 'Y', and 'Z' whose definitions are 'X1,' 'Y1,' and 'Z1') is a best practice.

Example

This dataset from the NYC Department of Sanitation describes public recycling bins in the City. On the landing page, the seven columns (variables) are defined and mapped to the field name and assigned a data type. All are listed as "text" data, though some would more appropriately be characterized as alphanumeric, numeric, or categorical. From the landing page, there is a separate file with the dictionary and more information about the data product, including expected or allowed values (but this is not filled in).

An example of a data product that meets the standard



Source: Public Recycling Bins, Department of Sanitation (DSNY)

See more here

02 | Are variable definitions clearly linked or displayed?

RETURN TO FRAMEWORK

Additional directions and context

Check the box if variable definitions are obviously available on the data product or landing page, or if they are clearly linked from the landing page.

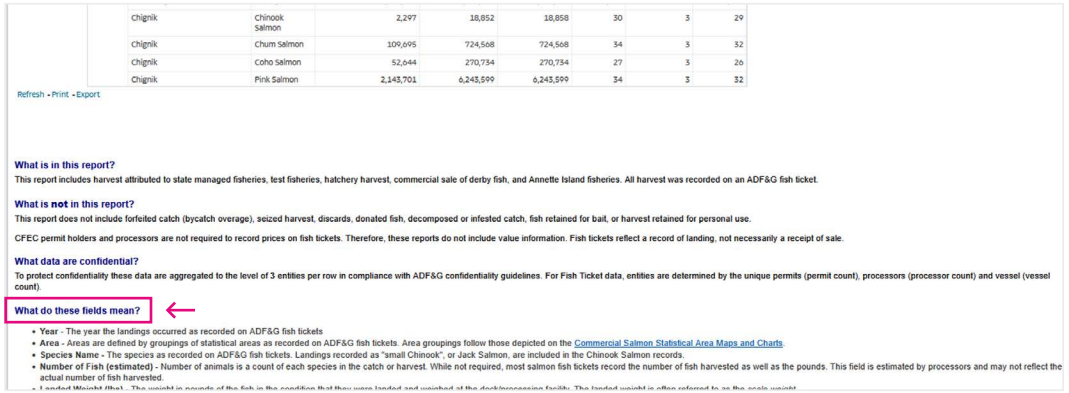
Motivating intent behind the question

Variable definitions are not helpful if they can't be found. Putting definitions in multiple places helps increase findability, for example, by adding them to the data download. However, relying solely on the data download to share variable definitions hides key information from users who prefer interacting with the data product over working with the source data. Including this information in the data download can also affect findability: a clearly named, separate file is best, while a clearly named, separate tab in the data file is sufficient.

Example

The Alaska Department of Fish and Game places information describing key variables just below the data product at the bottom of the page.

An example of a data product that meets the standard



Chignik	Chinook Salmon	2,297	18,852	18,858	30	3	29
Chignik	Chum Salmon	109,695	724,568	724,568	34	3	32
Chignik	Coho Salmon	52,644	270,734	270,734	27	3	26
Chignik	Pink Salmon	2,143,701	6,243,599	6,243,599	34	3	32

Refresh - Print - Export

What is in this report?
This report includes harvest attributed to state managed fisheries, test fisheries, hatchery harvest, commercial sale of derby fish, and Annette Island fisheries. All harvest was recorded on an ADF&G fish ticket.

What is not in this report?
This report does not include forfeited catch (bycatch overage), seized harvest, discards, donated fish, decomposed or infested catch, fish retained for bait, or harvest retained for personal use.

CFEC permit holders and processors are not required to record prices on fish tickets. Therefore, these reports do not include value information. Fish tickets reflect a record of landing, not necessarily a receipt of sale.

What data are confidential?
To protect confidentiality these data are aggregated to the level of 3 entities per row in compliance with ADF&G confidentiality guidelines. For Fish Ticket data, entities are determined by the unique permits (permit count), processors (processor count) and vessel (vessel count).

What do these fields mean?

- **Year** - The year the landings occurred as recorded on ADF&G fish tickets
- **Area** - Areas are defined by groupings of statistical areas as recorded on ADF&G fish tickets. Area groupings follow those depicted on the Commercial Salmon Statistical Area Maps and Charts.
- **Species Name** - The species as recorded on ADF&G fish tickets. Landings recorded as "small Chinook", or Jack Salmon, are included in the Chinook Salmon records.
- **Number of Fish (estimated)** - Number of animals is a count of each species in the catch or harvest. While not required, most salmon fish tickets record the number of fish harvested as well as the pounds. This field is estimated by processors and may not reflect the actual number of fish harvested.

Source: Statewide Salmon Landings by Area, Alaska Department of Fish and Game
See more here

03 | Does the data product address uncertainty through statistical methods or discussion?

← RETURN TO FRAMEWORK

Additional directions and context

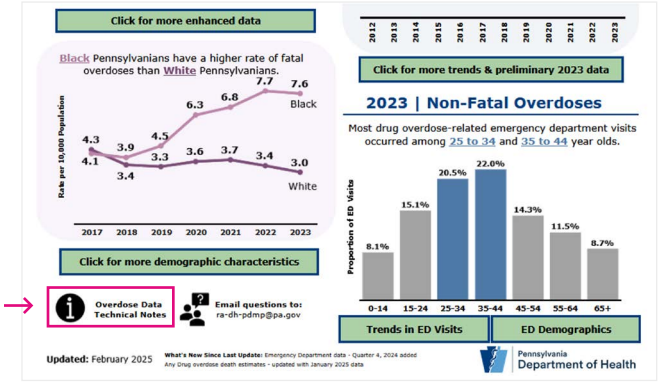
Data quality considerations emerge during data collection and cleaning. This question assesses whether these considerations are carried through and made visible in the data product itself. Signals of uncertainty can appear directly in the data product, data download, or supporting materials accessed via the landing page.

Motivating intent behind the question

Ultimately, the data product user wants to understand how confident they can be in the data presented and the conclusions they draw from it. Every dataset — statistical or otherwise — has limitations, uncertainty, and methodological decisions that need explaining. Uncertainty can take many forms and can be explored through statistical measures like confidence intervals, margins of error, or other significance testing, or through discussion of data completeness or methodological limitations.

Example

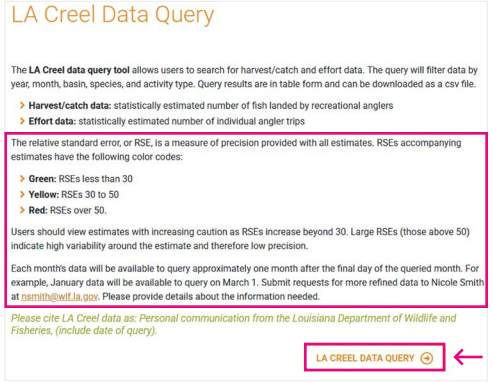
An example of a data product that meets the standard



The Pennsylvania Department of Health provides the user with a limitations section accessed from the “Overdose Data Technical Notes” link in the data product that discusses how medical coding influences overdose counts.

Source: Fatal and Non-Fatal Drug Overdose Surveillance, Office of Drug Surveillance and Misuse Prevention, Pennsylvania Department of Health

See more here



The Louisiana Department of Fish and Game uses a statistical approach to describe uncertainty, but groups the relative standard errors by color to indicate how much confidence the user should have in the results.

LA Creel Data Query, Louisiana Department of Fish and Game

See more here

04 | Are popular or useful cross-sections, takeaways, or views of the data provided?

← RETURN TO FRAMEWORK

Additional directions and context

There are many ways to orient the data product user to what might be most interesting in the data. These can include summary paragraphs, highlights, popular views, or cross-sections of the data. Note that this is different from features like “browse by topic” or “sort by popular,” which are common tools in data catalogs that help users find data products but do not highlight key insights within a data product itself.

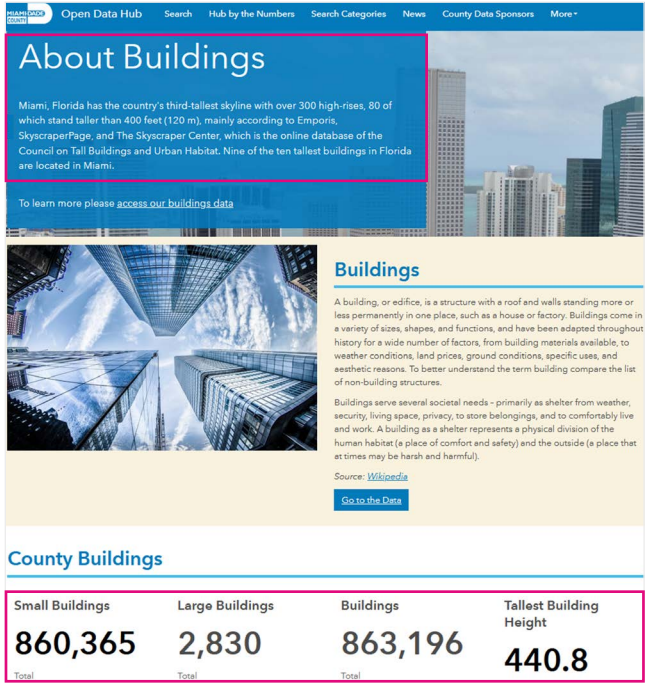
Motivating intent behind the question

The data product publisher has the best sense for highlights or insights in the data — signaling this to the user is a helpful jumping-off point for exploration. This type of data curation can help answer basic research questions and spur further exploration.

Example

In Miami-Dade County, the landing page for its buildings data takes a simple “fast facts” approach to data curation. It provides links to the buildings data (the data catalog filtered to buildings), plus a short, quantitative description of the building stock in the county and highlights from a notable dataset — Building Footprint 2D — in the form of four statistics about county buildings.

An example of a data product that meets the standard



Source: Building Footprint 2D, Miami-Dade County Open Data

See more here

05 | Are there accompanying reports that include relevant data tables?

← RETURN TO FRAMEWORK

Additional directions and context

A summary report, memo, executive summary, blog post, or newsletter are examples of supplementary materials that can pair key tables with explanatory text to help the user of the data product engage more deeply with the data being presented.

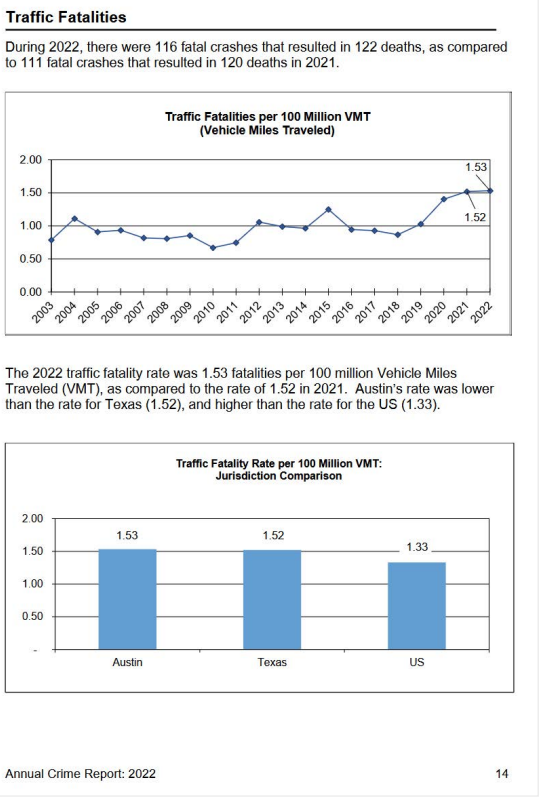
Motivating intent behind the question

Providing charts, text, and interpretation of the raw numbers can guide a data product user by showing what stands out from the data and invite deeper inquiry.

Example

This data is available via the City of Austin's open data portal, and this report provides additional value to the data product user by sharing charts and trends that explain what is interesting and important from this large public safety dataset. One weakness: it is difficult to find from the source data.

An example of a data product that meets the standard



Traffic Fatalities

During 2022, there were 116 fatal crashes that resulted in 122 deaths, as compared to 111 fatal crashes that resulted in 120 deaths in 2021.

The 2022 traffic fatality rate was 1.53 fatalities per 100 million Vehicle Miles Traveled (VMT), as compared to the rate of 1.52 in 2021. Austin's rate was lower than the rate for Texas (1.52), and higher than the rate for the US (1.33).

Annual Crime Report: 2022 14

Source: Annual Crime and Traffic Report 2022, Austin Police Department

See more here

06 | Is a point of contact provided?

RETURN TO FRAMEWORK

Additional directions and context

Common strategies include providing a program or department-specific helpline (email or telephone), contact information for individuals involved with the product, or a feedback form where the user provides their question and contact information.

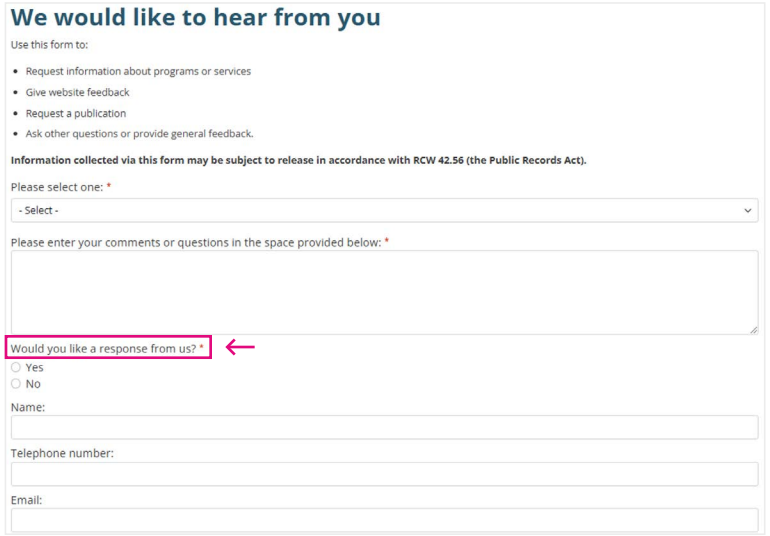
Motivating intent behind the question

While it can be difficult to monitor a helpline or shared inbox, the data product user should be offered some way to get questions answered should the provided documentation prove insufficient.

Example

Here is an example of a multipurpose feedback form from the Washington State Department of Health that can accommodate questions about data products. It features a dropdown menu to select the type of question being posed, solicits contact information, and asks if a repsonse is needed. Linking to this style of form on a data product provides the user with a plausible path to getting questions answered without needing to provide agency contact information.

An example of a data product that meets the standard



Source: Feedback Form, Washington State Department of Health
See more here

07 | Is a suggested citation provided?

RETURN TO FRAMEWORK

Additional directions and context

Providing the information necessary to create a citation (e.g., authors, publication date, title, etc.) is not sufficient to check the box.

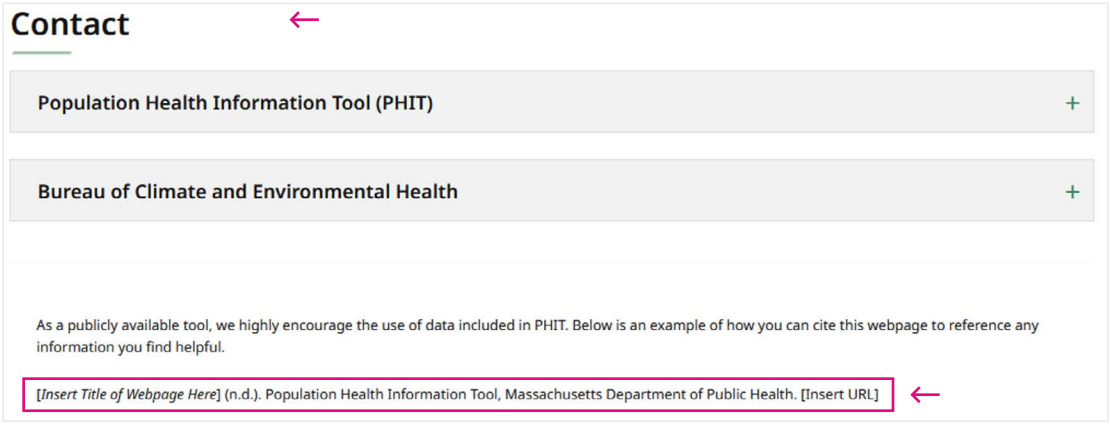
Motivating intent behind the question

Ready-made citations help users understand which program within an agency produced the data product, and remove the burden of assembling a citation.

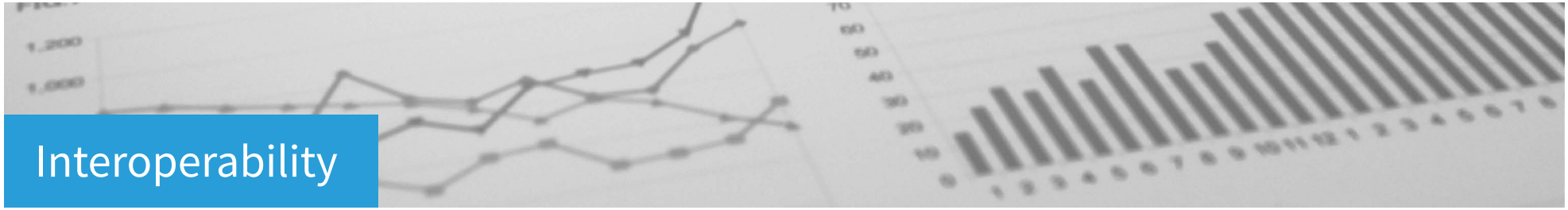
Example

The Massachusetts Department of Public Health provides a flexible citation on the landing page of its data product that can be used for different information contained in the Population Health Information Tool.

An example of a data product that meets the standard



Source: Environmental Hospitalization Data, Population Health Information Tool, Massachusetts Department of Public Health
See more here



Interoperability

01 **Is the methodology provided?**

☐ LEARN MORE →

02 **Is the methodology clearly linked or displayed?**

☐ LEARN MORE →

03 **Is other metadata provided and clearly linked or displayed?**

☐ LEARN MORE →

04 **Are variable names both short and meaningful (to facilitate quick scanning and human readability)?**

☐ LEARN MORE →

05 **Are variable names free of special characters (to facilitate machine readability)?**

☐ LEARN MORE →

01 | Is the methodology provided?

RETURN TO FRAMEWORK

Additional directions and context

Methodology documentation explaining how the data was collected and transformed is a specific piece of metadata that should be provided with every data product. Every data product is different, and the methodology documentation will necessarily vary in length and detail.

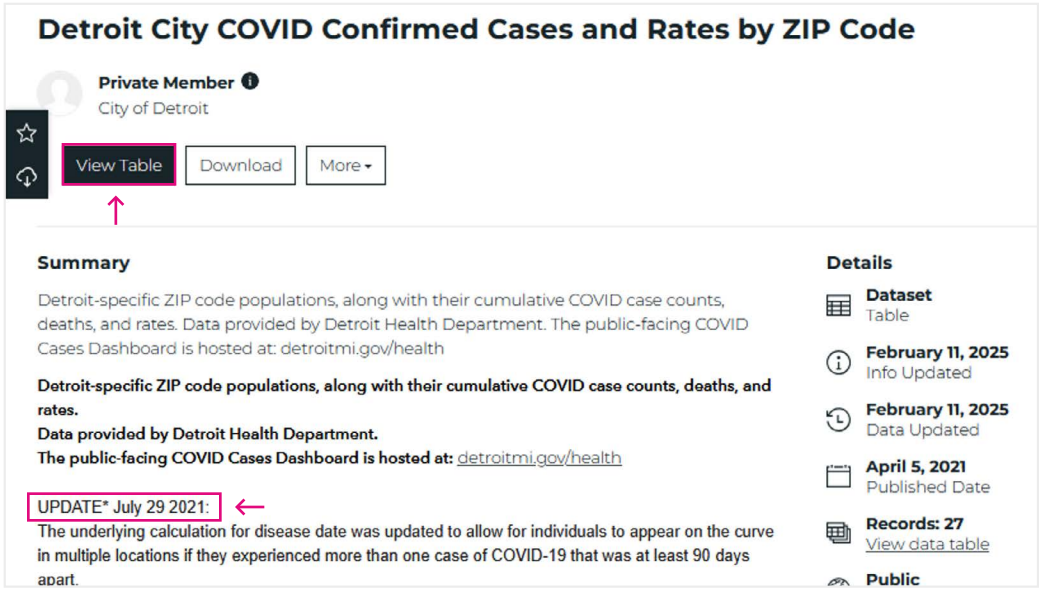
Motivating intent behind the question

Showing the data product user how the data came to be demonstrates transparency and is a best practice.

Example

The Detroit Health Department provides a landing page for their COVID data, which can be accessed via the “View Table” button. In this example, the methodology is sufficiently short to be included directly on the landing page. The landing page describes case counts, where those counts originate from, improving the accuracy of case counts within Detroit ZIP codes, excluded/suppressed data, circumstances under which an individual may be counted multiple times, and other technical matters.

An example of a data product that meets the standard



Source: Detroit City COVID Confirmed Cases and Rates by ZIP Code, City of Detroit

See more here

02 | Is the methodology clearly linked or displayed?

RETURN TO FRAMEWORK

Additional directions and context

Check the box if variable definitions are obviously available on the data product or landing page, or if they are clearly linked from the landing page.

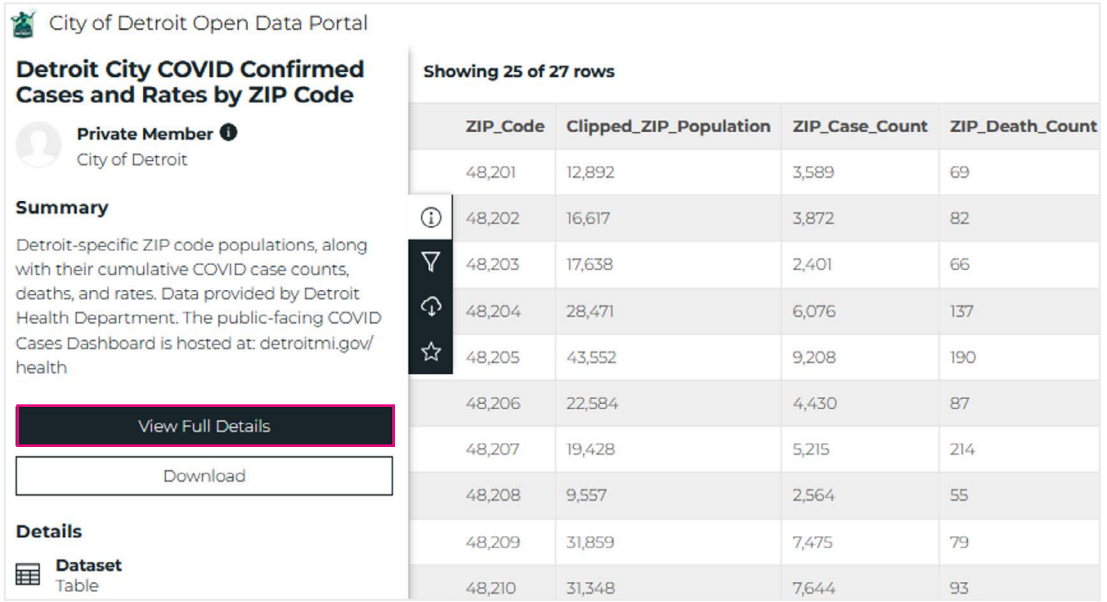
Motivating intent behind the question

Methodology documentation is not helpful if it can't be found. Putting the methodology in multiple places helps increase findability, for example by adding it as a separate file to the data download or including it in an accompanying report. However, relying solely on the data download to share methodology hides key information from users who prefer interacting with the data product over working with the source data.

Example

From the dataset put together by the Detroit Health Department, clicking "View Full Details" brings the user to a landing page with information about the data product, including key methodological notes.

An example of a data product that meets the standard



Source: Detroit City COVID Confirmed Cases and Rates by ZIP Code, City of Detroit

See more here

03 | Is other metadata provided and clearly linked or displayed?

RETURN TO FRAMEWORK

Additional directions and context

Check the box if something besides methodology and variable definitions that helps with the interpretation and interoperability of the data — like an FAQ section, a crosswalk table, or a changelog — is provided and obviously available on the data product or landing page, or if they are clearly linked from the landing page.

Motivating intent behind the question

There are nearly always opportunities to provide additional context or details that support machine readability or the user's understanding of the data product.

Example

The NYC Workforce Data Portal provides notes for each data release, discussing what's new or different from the prior release and setting expectations for future releases.

An example of a data product that meets the standard

Source: Workforce Data Portal, Mayor's Office for Economic Opportunity

See more here

04 | Are variable names both short and meaningful (to facilitate quick scanning and human readability)?

← RETURN TO FRAMEWORK

Additional directions and context

This category is subjective. Check the box if there is an attempt to truncate variable names while still retaining meaning. Often, this takes the form of dropping vowels (e.g., “year” to “yr”) and/or using common abbreviations (e.g., “population” to “pop”).

Motivating intent behind the question

Done right, condensing variable names while retaining meaning makes it possible for the user to understand the data *in situ* without heavily depending on a data dictionary. The best data products are also consistent with respect to spacing and capitalization of variable names, making them easier for humans to read.

Example

The Georgia Governor’s Office of Student Achievement reports on ACT scores in the state. They have abbreviated long words while retaining meaning by removing vowels (“COUNT” becomes “CNT”) and abbreviating common words (“NUMBER” becomes “NUM”). They also use a uniform approach to naming, connecting text with underscores, and adopting all caps; this reduces variation and makes it easy to scan.

An example of a data product that meets the standard

INSTN_NUMBER	INSTN_NAME	SUBGRP_DESC	TEST_CMPNT_TYP_CD	NATIONAL_NUM_TESTED_CNT
103	Appling County	All Students	Composite	1374791
103	Appling County	All Students	English	1374791
103	Appling County	All Students	Mathematics	1374791
103	Appling County	All Students	Reading	1374791

Source: ACT Scores, Georgia Governor’s Office of Student Achievement

See more here

05 | Are variable names free of special characters (to facilitate machine readability)?

RETURN TO FRAMEWORK

Additional directions and context

Only check the box if variable names use numbers, letters, and underscores. Anything other than underscores — like $, ~, *, or superscript numerals for footnotes/endnotes — interferes with machine readability.

Motivating intent behind the question

Computer code will often use special characters — by including these in variable names, it becomes more difficult to programmatically work with data.

Example

The Ohio Department of Natural Resources Division of Wildlife packages its data in a .csv format with machine-readable variable names. Multi-word variables are joined with underscores, no spaces or other special characters are present, and the (non)use of capital letters is consistent in the variable names.

An example of a data product that meets the standard

id	date	location	zone	species	species_category	count	query_run_date
4	12/4/1990	Alum Cre	North	American	Dabblers	3	1/31/2025
5	12/17/1990	Alum Cre	North	American	Dabblers	3	1/31/2025
7	1/14/1991	Alum Cre	North	American	Dabblers	58	1/31/2025
14	12/3/1991	Alum Cre	North	American	Dabblers	40	1/31/2025

Source: Ohio Bi-weekly Waterfowl Survey, ODNR Division of Wildlife, DataOhio

See more here

Timeliness of data

01 **Is the update or publication cadence posted?**

☐ LEARN MORE →

02 **Is the publication date or date of last update posted?**

☐ LEARN MORE →

03 **Is the amount of historical data available explained?**

☐ LEARN MORE →

04 **Is the appropriateness of comparing data over time discussed?**

☐ LEARN MORE →

01 | Is the update or publication cadence posted?

RETURN TO FRAMEWORK

Additional directions and context

Signaling how often the data available in the product will be refreshed can be accomplished different ways: a brief program description (e.g., “a quarterly report”) or publication calendar are common approaches.

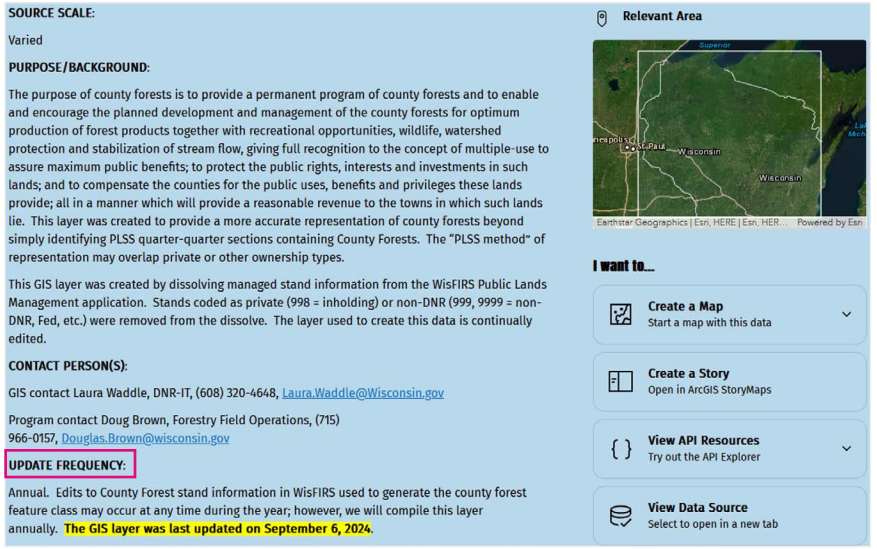
Motivating intent behind the question

The user will want to know how often to expect fresh data. How frequently will the data update, if at all?

Example

Here is an example from Wisconsin Department of Natural Resources Open Data, where the landing page for the data includes a section titled “Update frequency.” It provides an expected update/publication cadence (annual), a note about irregular updates or edits to the data, and states when the data was last updated.

An example of a data product that meets the standard



SOURCE SCALE:

Varied

PURPOSE/BACKGROUND:

The purpose of county forests is to provide a permanent program of county forests and to enable and encourage the planned development and management of the county forests for optimum production of forest products together with recreational opportunities, wildlife, watershed protection and stabilization of stream flow, giving full recognition to the concept of multiple-use to assure maximum public benefits; to protect the public rights, interests and investments in such lands; and to compensate the counties for the public uses, benefits and privileges these lands provide; all in a manner which will provide a reasonable revenue to the towns in which such lands lie. This layer was created to provide a more accurate representation of county forests beyond simply identifying PLSS quarter-quarter sections containing County Forests. The “PLSS method” of representation may overlap private or other ownership types.

This GIS layer was created by dissolving managed stand information from the WisFIRS Public Lands Management application. Stands coded as private (998 = inholding) or non-DNR (999, 9999 = non-DNR, Fed, etc.) were removed from the dissolve. The layer used to create this data is continually edited.

CONTACT PERSON(S):

GIS contact Laura Waddle, DNR-IT, (608) 320-4648, Laura.Waddle@Wisconsin.gov

Program contact Doug Brown, Forestry Field Operations, (715) 966-0157, Douglas.Brown@wisconsin.gov

UPDATE FREQUENCY:

Annual. Edits to County Forest stand information in WisFIRS used to generate the county forest feature class may occur at any time during the year; however, we will compile this layer annually. **The GIS layer was last updated on September 6, 2024.**

Relevant Area

I want to...

Create a Map
Start a map with this data

Create a Story
Open in ArcGIS StoryMaps

View API Resources
Try out the API Explorer

View Data Source
Select to open in a new tab

Source: Wisconsin County Forests, Wisconsin DNR Open Data

See more here

02 | Is the publication date or date of last update posted?

RETURN TO FRAMEWORK

Additional directions and context

A common pitfall is adding a date to a data product or landing page, but not being specific about what has been updated. Be precise when using a timestamp to signal when the data in the product last updated: e.g., “data updated on DATE” is clear. A date by itself or “updated/refreshed on DATE” might indicate a data update, or it might indicate updates to language or other features of the data product or landing page.

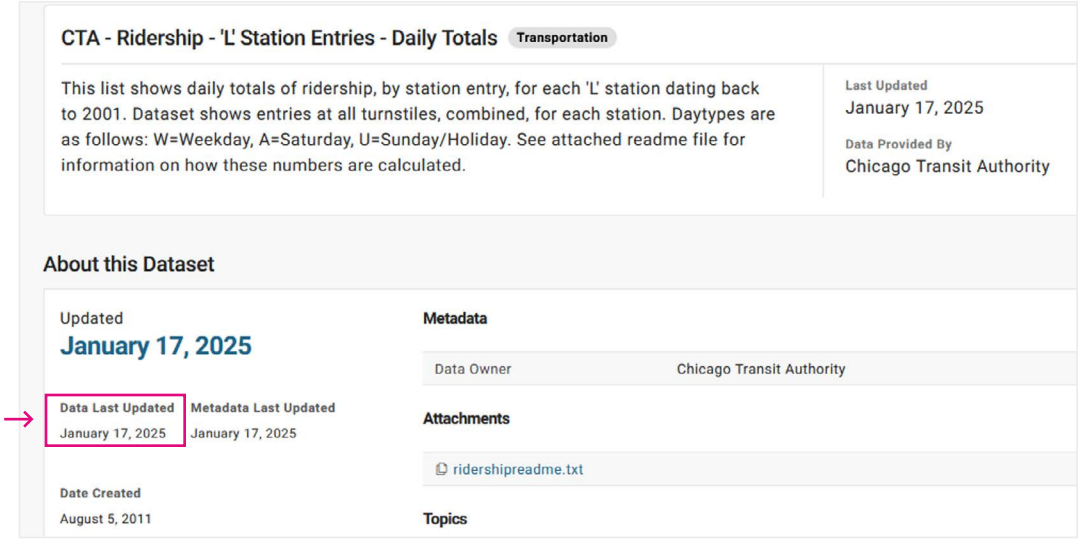
Motivating intent behind the question

The user will want to know if the data shown by the product is up to date. Often, it takes time to collect, clean, and publish data, so an up-to-date data product might show data from a prior calendar year. Without explanation, a data product user might be uncertain that they’re looking at current information.

Example

The Chicago Transit Authority provides data counting daily entries to its ‘L’ stations. The “About” section of the landing page explicitly provides the date that the data were last updated, and separately provides a date for when the metadata were updated.

An example of a data product that meets the standard



Source: Chicago Transit Authority, ‘L’ Station Entries —Daily Totals
See more here

03 | Is the amount of historical data available explained?

RETURN TO FRAMEWORK

Additional directions and context

A sentence or two describing the program's history, changes, and/or publication cadence is often sufficient to explain the scope of data presented in the data product. Footnotes or endnotes explaining why charts start when they start, or indications of where older data is archived and accessible can also be helpful.

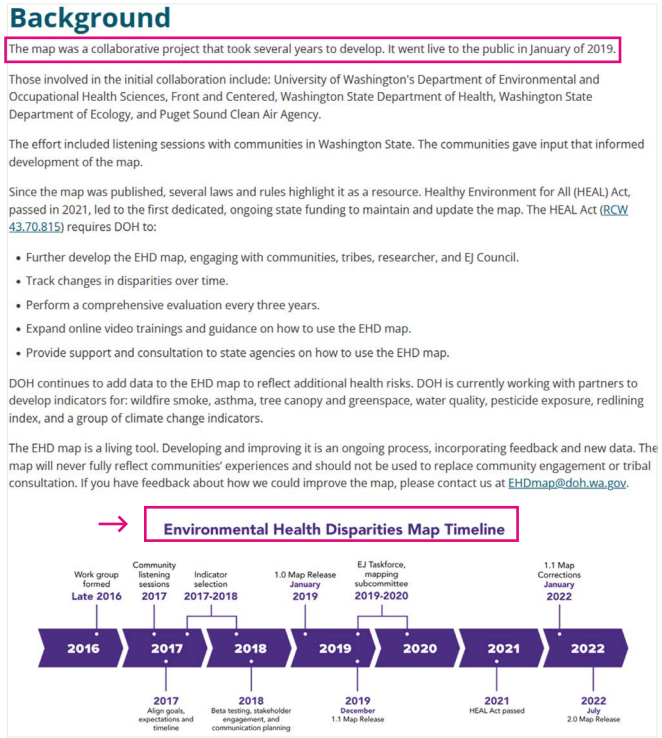
Motivating intent behind the question

Data product users may wonder why they're looking at data for a certain number of years/months/quarters and not others. Are they being shown a five-year view of the data because no more exists? If more exists, why is it not being shown?

Example

The Washington State Department of Health explains the history of the Environmental Health Disparities Map, and that each map release is unique.

An example of a data product that meets the standard



Background

The map was a collaborative project that took several years to develop. It went live to the public in January of 2019.

Those involved in the initial collaboration include: University of Washington's Department of Environmental and Occupational Health Sciences, Front and Centered, Washington State Department of Health, Washington State Department of Ecology, and Puget Sound Clean Air Agency.

The effort included listening sessions with communities in Washington State. The communities gave input that informed development of the map.

Since the map was published, several laws and rules highlight it as a resource. Healthy Environment for All (HEAL) Act, passed in 2021, led to the first dedicated, ongoing state funding to maintain and update the map. The HEAL Act (RCW 43.70.815) requires DOH to:

- Further develop the EHD map, engaging with communities, tribes, researcher, and EJ Council.
- Track changes in disparities over time.
- Perform a comprehensive evaluation every three years.
- Expand online video trainings and guidance on how to use the EHD map.
- Provide support and consultation to state agencies on how to use the EHD map.

DOH continues to add data to the EHD map to reflect additional health risks. DOH is currently working with partners to develop indicators for: wildfire smoke, asthma, tree canopy and greenspace, water quality, pesticide exposure, redlining index, and a group of climate change indicators.

The EHD map is a living tool. Developing and improving it is an ongoing process, incorporating feedback and new data. The map will never fully reflect communities' experiences and should not be used to replace community engagement or tribal consultation. If you have feedback about how we could improve the map, please contact us at EHDmap@doh.wa.gov.

Environmental Health Disparities Map Timeline

Source: Environmental Health Disparities Map, Washington State Department of Health

See more here

04 | Is the appropriateness of comparing data over time discussed?

RETURN TO FRAMEWORK

Additional directions and context

Check the box if not applicable (e.g., if the data represent a snapshot in time, if the comparison is provided for the user, or if comparability is otherwise not a concern). Flagging reasons why a user might need to be careful with comparisons over time can be accomplished with footnotes, discussions of limitations, or visual cues on a chart or interactive environment.

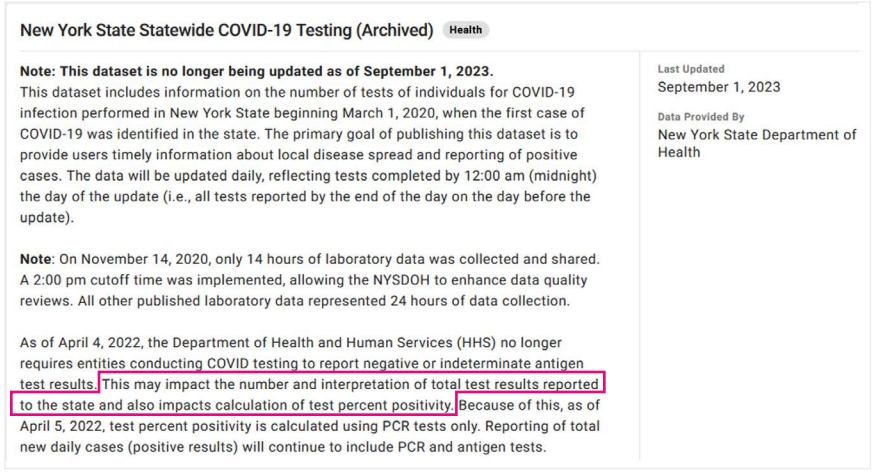
Motivating intent behind the question

Seeing whether a data point has increased or decreased is a common way for data product users to learn about an issue. If it is acceptable to compare changes over time, it is helpful to provide the time series data that shows this so users do not have to assemble it themselves. Changes to programs, definitions, or data collection methods might affect year-over-year comparability; providing these kinds of warnings to explain a choice to not provide a time series to support multi-year analysis is helpful to the user.

Example

Here, the New York State Department of Health describes the dataset, explaining a meaningful change in data collection that would caution the user against overinterpreting a decrease in April of 2022 when reporting mandates expired.

An example of a data product that meets the standard



New York State Statewide COVID-19 Testing (Archived) Health

Note: This dataset is no longer being updated as of September 1, 2023.
This dataset includes information on the number of tests of individuals for COVID-19 infection performed in New York State beginning March 1, 2020, when the first case of COVID-19 was identified in the state. The primary goal of publishing this dataset is to provide users timely information about local disease spread and reporting of positive cases. The data will be updated daily, reflecting tests completed by 12:00 am (midnight) the day of the update (i.e., all tests reported by the end of the day on the day before the update).

Note: On November 14, 2020, only 14 hours of laboratory data was collected and shared. A 2:00 pm cutoff time was implemented, allowing the NYSDOH to enhance data quality reviews. All other published laboratory data represented 24 hours of data collection.

As of April 4, 2022, the Department of Health and Human Services (HHS) no longer requires entities conducting COVID testing to report negative or indeterminate antigen test results. This may impact the number and interpretation of total test results reported to the state and also impacts calculation of test percent positivity. Because of this, as of April 5, 2022, test percent positivity is calculated using PCR tests only. Reporting of total new daily cases (positive results) will continue to include PCR and antigen tests.

Last Updated
September 1, 2023

Data Provided By
New York State Department of Health

Source: COVID-19 Testing Data, New York State Department of Health
See more here