Rural Community Assistance Partnership

Federal funding map for water infrastructure to target outreach and assistance

Methodology report

April 4th, 2025

This project was funded by EPA under the National Water Infrastructure Environmental Finance Center Program.

For questions, please contact info@rcap.org.

Table of Contents

Introduction

The Rural Community Assistance Partnership Incorporated (RCAP) was selected by the U.S. Environmental Protection Agency (EPA) as a National Environmental Finance Center. Through this project, RCAP aims to help communities address their water, wastewater and financial resource challenges, and to effectively leverage EPA and other funding for small, rural, and Tribal communities, including those who have not received assistance related to water infrastructure in the past. RCAP analyzed past water infrastructure funding data to help prioritize outreach and technical assistance for small, rural, and Tribal communities across the country that most need support to access BIL and other infrastructure funding to assist with their most urgent water infrastructure needs. This document lays out the data analysis steps in order to create both understanding of the methodology and repeatability of the analysis at a later time.

Goal of the Analysis

Much of the water infrastructure within the U.S. and its territories is in dire need of maintenance and replacement, due to factors such as the aging-out of the physical infrastructure, a lack of funding/financial resources faced by many communities to perform necessary maintenance required by regulations and physical deterioration, and the need to keep up with new regulations. This is only further exacerbated in small, rural, and Tribal communities, that often do not have the means to do so and may rely on outside technical assistance providers, such as RCAP. To help prioritize and target outreach for these communities, RCAP (with input from EPA) developed an interactive map to help technical assistance providers identify the communities that may need the most assistance to access BIL and other water infrastructure funding sources. By using publicly available funding data from USA Spending, combined with spatial community data from The United States Geological Survey (USGS), this webmap can help identify places that have received little or no federal funding for water infrastructure alongside other data indices help technical assistance providers prioritize communities for outreach.

Program Selection

The first step was to identify which federal funding programs related to water infrastructure to include in our analysis. To search for applicable funding programs, we used the 2022 Annual Publication of Assistance Listings published on SAM.gov (which will be referred to as SAM APAL).[1] The 2022 SAM APAL was utilized as it was the most up-to-date list of all programs within one PDF and published via a federal government website. In this 3,600+ page document, the SAM APAL offers an in-depth overview of each federal program, such as the federal agency the funding falls under, objectives of the program, applicant and beneficiary eligibility, and important application deadlines.

[1] https://sam.gov/api/prod/fac/v1/programs/downloadPublication/2022?api_key=null

Each program is assigned a unique identification number, a 5-digit identification number in the form of ##.### which is called its "assistance listing." Although the publication refers to these identification numbers as "assistance listings,'" they are also referred to in other federal documents as "Catalog of Federal Domestic Assistance" (CFDA) numbers (which is how they will be addressed in this document). The programs that qualified as possible contenders for our analysis had to apply to water infrastructure in general – this could include drainage systems, drinking water systems, municipal and industrial water supply, wastewater disposal systems, water reuse and water storage, to name a few. Across five different U.S. agencies, a total of 48 programs were determined to meet the criteria of relating to water infrastructure.

Table 1 provides a short summary of these 48 programs. All programs on this list are federally funded programs, and do not include any state-level funding programs. It is also important to note that only federal programs with a CFDA number associated with them are included. This means that federal funding programs like the Disaster Recovery Fund are not listed on our program table, as they do not have a CFDA number and data was not available via USASpending.

The table below compiles the titles of each program, their associated CFDA number, and the agency they fall under. The fourth column provides a brief overview of the water infrastructure category the program relates to, while the fifth column indicates whether the program targets funding for drinking water (DW), wastewater (WW), or both. The sixth column indicates whether the program is a nationally available funding source or is specific to a region. If region-specific, the column also provides a short description of which region. The final column serves as a link to the SAM.gov listing of the program for further information such as funding program description, applicant eligibility, and deadlines for applying.

Table 1: *List of all Federally Funded Water Infrastructure Programs from 2022 SAM APAL*

Program	CFDA	Agency	Category of Funding Type	Drinking Water or Wastewater specific, or both eligible	Nationwide or Regional	Link
Part 1774 Special Evaluation Assistance for Rural Communities and Households (SEARCH)	10.759	USDA	Drinking water systems; Wastewater systems	Both	Nationwide	Link
Water and Waste Disposal Systems for Rural Communities	10.760	USDA	Water and wastewater disposal	Both	Nationwide	Link
Solid Waste Management Grants	10.762	USDA	Solid waste; Source water protection; Pollution reduction/elimination; Planning and Management	WW	Nationwide	Link
Emergency Community Water Assistance Grants	10.763	USDA	Drinking water systems	DW	Nationwide	Link
Water and Waste Disposal Loans and Grants (Section 306C)	10.770	USDA	Water and wastewater disposal	Both	Nationwide	Link
Rural Decentralized Water Systems Grant	10.862	USDA	Drinking water systems	Both	Nationwide	Link
Grant Program to Establish a Fund for Financing Water and Wastewater Projects	10.864	USDA	Drinking water systems; Wastewater systems	Both	Nationwide	Link
Watershed Protection and Flood Prevention	10.904	USDA	Wetlands/Watersheds/Estuaries; Municipal and industrial water supply	Both	Nationwide	Link
Investments for Public Works and Economic Development Facilities	11.300	EDA	Drinking water systems; Wastewater systems	Both	Nationwide	Link

Program	CFDA	Agency	Category of Funding Type	Drinking Water or Wastewater specific, or both eligible	Nationwide or Regional	Link
Community Development Block Grants/Entitlement Grants (CDBG)	14.218	HUD	Drinking water systems; Wastewater systems	Both	Nationwide	Link
Community Development Block Grants/Special Purpose Grants/Insular Areas	14.225	HUD	Drinking water systems; Wastewater systems	Both	Nationwide	Link
Community Development Block Grants/State's program and Non-Entitlement Grants in Hawaii	14.228	HUD	Drinking water systems; Wastewater systems	Both	Region - Hawaii	Link
Community Development Block Grants Section 108 Loan Guarantees	14.248	HUD	Drinking water systems; Wastewater systems	Both	Nationwide	Link
Hurricane Sandy Community Development Block Grant Disaster Recovery Grants (CDBG-DR)	14.269	HUD	Drinking water systems; Wastewater systems	Both	Region - NJ/NY	Link
Indian Community Development Block Grant Program	14.862	HUD	Drinking water systems; Wastewater systems	Both	Nationwide, but only for federally recognized tribes	Link
Small Surface Water and Groundwater Storage Projects	15.074	DOI	Water storage; Groundwater	Both	Region - 17 western states (including Alaska & Hawaii)	Link
Blackfeet Water Rights Settlement	15.075	DOI	Municipal and industrial water supply	Both	Region - Blackfeet Tribe, ND	Link

Program	CFDA	Agency	Category of Funding Type	Drinking Water or Wastewater specific, or both eligible	Nationwide or Regional	Link
Musselshell-Judith Rural Water System	15.076	DOI	Water Supply	Both	Region - citizens of Judith Basin, Wheatland, Golden Valley, Musselshell, and portions of Yellowstone and Fergus counties in the state of Montana	Link
Title XVI Water Reclamation and Reuse	15.504	DOI	Water Reuse	WW	Region	Link
WaterSMART (Sustain and Manage Americas Resources for Tomorrow)	15.507	DOI	Water conservation; Water efficiency; Water Planning	Both	Nationwide	Link
Colorado Ute Indian Water Rights Settlement Act	15.510	DOI	Municipal and industrial water supply	Both	Region - Ute Mountain Ute, Southern Ute Indian Tribe, the Navajo Nation and non-Tribal participants from the Animas-La Plata Project	Link
Fort Peck Reservation Rural Water System	15.516	DOI	Municipal and industrial water supply	Both	Region - Fort Peck Indian Reservation	Link
Garrison Diversion Unit	15.518	DOI	Municipal and industrial water supply	Both	Region - North Dakota	Link

Program	CFDA	Agency	Category of Funding Type	Drinking Water or Wastewater specific, or both eligible	Nationwide or Regional	Link
Lewis and Clark Rural Water System	15.520	DOI	Municipal and industrial water supply	Both	Region - Other private institutions/organizations; The Lewis and Clark Rural Water Supply System, Inc., and its member entities (rural water systems and municipalities the meet the requirements for membership in the Lewis and Clark Rural Water Supply System, Inc.).	Link
Lower Rio Grande Valley Water Resources Conservation and Improvement	15.521	DOI	Water Transportation (pipes)	Both	Region - Texas	Link
Mni Wiconi Rural Water Supply Project	15.522	DOI	Municipal and industrial water supply; Water and wastewater disposal	Both	Region - tribes in South Dakota	Link
Rocky Boy's/North Central Montana Regional Water System	15.525	DOI	Municipal and industrial water supply	Both	Region - Chippewa-Cree of the Rocky Boy's Indian reservation	Link
Equus Beds Division Aquifer Storage Recharge	15.539	DOI	Water storage; Groundwater	Both	Region - city of Wichita, Kansas	Link
Reclamation Rural Water Supply (Rural Water Program)	15.548	DOI	Water Supply	Both	Nationwide	Link
Navajo-Gallup Water Supply	15.552	DOI	Water Supply	Both	Region - Navajo-Gallup; New Mexico	Link

Program	CFDA	Agency	Category of Funding Type	Drinking Water or Wastewater specific, or both eligible	Nationwide or Regional	Link
Eastern New Mexico Rural Water System	15.553	DOI	Drinking water systems	DW	State- New Mexico	Link
Crow Tribe Water Rights Settlement	15.556	DOI	Municipal and industrial water supply	Both	Region - Crow Tribe of Montana	Link
White Mountain Apache Tribe Rural Water System	15.558	DOI	Drinking water systems	Both	Region - White Mountain Apache Tribe	Link
Water Storage Enhancement	15.574	DOI	Water storage	Both	Nationwide	Link
Lake Pontchartrain Basin Restoration Program (PRP)	66.125	EPA	Green Infrastructure; Wastewater systems;	WW	Region - Lake Pontchartrain Basin Watershed	Link
Geographic Programs - Southeast New England Coastal Watershed Restoration Program	66.129	EPA	Wastewater systems	WW	Region - SE New England	Link
Construction Grants for Wastewater Treatment Works	66.418	EPA	Wastewater systems	WW	Nationwide	Link
Water Infrastructure Improvements for the Nation Small and Underserved Communities Emerging Contaminants Grant Program	66.442	EPA	Drinking water systems; Wastewater systems	Both	Nationwide - Indian Tribal governments	Link
Reducing Lead in Drinking Water (SDWA 1459B)	66.443	EPA	Drinking water systems	DW	Nationwide	Link
Voluntary School and Child Care Lead Testing and Reduction Grant Program (SDWA 1464(d))	66.444	EPA	Drinking water systems	DW	Nationwide	Link

Program	CFDA	Agency	Category of Funding Type	Drinking Water or Wastewater specific, or both eligible	Nationwide or Regional	Link
Sewer Overflow and Stormwater Reuse Municipal Grant Program	66.447	EPA	Wastewater systems	WW	Nationwide	Link
Drinking Water System Infrastructure Resilience and Sustainability Program SDWA 1459A(l)	66.448	EPA	Drinking water systems	DW	Nationwide	Link
Capitalization Grants for Clean Water State Revolving Funds (CWSRF)	66.458	EPA	Wastewater Systems; Water Supply	WW	Nationwide	Link
Capitalization Grants for Drinking Water State Revolving Funds (DWSRF)	66.468	EPA	Drinking water systems	DW	Nationwide	Link
Disaster Relief Appropriations Act (DRAA) Hurricane Sandy Capitalization Grants For Clean Water State Revolving Funds	66.482	EPA	Wastewater systems	WW	Region - states of NY & NJ	Link
Disaster Relief Appropriations Act (DRAA) Hurricane Sandy Capitalization Grants for Drinking Water State Revolving Funds	66.483	EPA	Drinking water systems	DW	Region - states of NY & NJ	Link
Innovative Water Technology Grant Program	66.521	EPA	Drinking water systems	DW	Nationwide	Link
Water Infrastructure Finance and Innovation (WIFIA	66.958	EPA	Drinking water systems; Wastewater systems	Both	Nationwide	Link

USA Spending Analysis

USA Spending Overview

USASpending.gov is a U.S. Government website that provides an in-depth explanation of how federal funds are distributed among different agencies and programs. On a monthly basis, more than 100 federal agencies submit financial data to the website which is then made accessible and downloadable to the public.[2] Along with the capacity to download federal funding data, the website also provides an in-depth Data Dictionary and Glossary to help explain what each column represents in the downloadable data, which are in CSV format. Programs can be searched for on the website by information such as award type, awarding agency, or award ID. We used the CFDA numbers of the 48 selected programs that relate to water infrastructure as our program identifier. The output CSV included an entry on each federally funded project with the award ID, funding totals through obligations and outlays, fiscal years for when projects were funded, the associated CFDAs, and spatial information such as city, county, and state location. By utilizing this information, we can gather spatial data on which communities received these federal funds for our 48 selected programs from Federal Fiscal Year's 2000 to 2023.

Upon downloading the data for all 48 programs, some processing decisions had to be made before reaching the final output. This included deciding which column to use for information on "funding data" from USA Spending, as well as the spatial resolution that would be ideal for our analysis. These funding and spatial resolution decisions are described in greater detail below.

Funding Data Type

The raw data download from USA Spending includes many ways to measure funding. This includes information on the obligations, outlays, and/or loans that each grantee receives. An in-depth explanation on what each type of funding represents can be found here.[3] For this project, we decided to use "obligations" as our source of funding data, because obligations provide a general overview of the funding promised from the federal government to the recipient. USA Spending defines obligations as: "*...when awarding funding, the U.S. government enters a binding agreement called an obligation. The government promises to spend the money, either immediately or in the future. An agency incurs an obligation, for example, when it places an order, signs a contract, awards a grant, purchases a service, or takes other actions that require it to make a payment*."[4]

Within the obligations, USA Spending offers two possible fields to look at for spending data – "Federal Action Obligations" and "Total Obligated Amount." Between the two, "Total Obligated Amount" had fewer blank entries after an initial analysis and fewer negative numbers, and thus would likely provide a more informative allocation map.

[2] https://www.usaspending.gov/data-sources
[3] https://www.usaspending.gov/data-sources?section=federal-spending-overview
[4] https://www.usaspending.gov/?glossary

Spatial Data

Since the goal is to more effectively target outreach to communities in need, analyzing the highest spatial resolution possible is important. USA Spending provides information on the city, county, and state for information on the location of a project grantee. For city-level information, there are two relevant columns in USA Spending – "recipient_city_name" and "primary_place_of_performance_city_name." The recipient city name refers to "*Name of the city in which the awardee or recipient's legal business address is located*," while Primary Place of Performance (PoP) refers to "*The name of the city where the predominant performance of the award will be accomplished*."[5] While the PoP is preferable for understanding the specifics of where funding had been distributed, as it is possible that the recipient city will then distribute said funding to another community (such as to a smaller, nearby community), its use is highly inconsistent and 15% of the data entries under PoP in our dataset were left entirely blank for city names. However, "Recipient City Name" did not have any missing entries. Therefore, we prioritized not losing any data for communities and used "Recipient City Name" data to analyze which communities received funding from federal programs.

An important aspect to keep in mind throughout this project is that USA Spending represents only federal funding sources. When processing and understanding our data by communities and the funding totals they received through each CFDA, it is crucial to remember that these funding totals do not represent how communities fared with state-level or other funding sources.

Retrieving the datasets

Prime award data was downloaded in September of 2023 from the Award Data Archives page of USAspending.gov.[6] The data that was originally downloaded included funding information for all federally funded programs in the United States, so it required subsetting to only include the 48 CFDA numbers for programs relevant to this project. The prime awards are the base transactions tracked by USAspending.gov, tracking the flow of funds between federal agencies and various recipients across the nation. Prime awards may have associated subawards, if those prime award funds were then re-distributed by the recipient of the prime award. Subaward data is also available through the Custom Award Data page of USAspending.gov. This project analyzed subawards specific to SRF and CDBG programs, as detailed below.

Scripting and data processing

Data was combined, processed, and saved using Python scripts. The code used in this project can be found archived and documented at the following GitHub repository: https://github.com/blusky-rcap/RCAP-national-USASpending/tree/main. The Python Dask library (an open-source Python library for parallel computing) was used to handle the first step of filtering for prime transaction level data, as the dataset is too large for conventional Python-based data science tools (such as Pandas) at the initial stage. Once the original dataset had been filtered to reduce the data size, the

[5] https://www.usaspending.gov/data-dictionary
[6] https://www.usaspending.gov/download_center/award_data_archive

Pandas library was used for further processing. Final steps of formatting the output took place in MS Excel.

Before loading the data into GIS for mapping, an R-Script (also found on GitHub) was run on the final dataset to add a column for the number of awards each location received and for adding the full program names and agency housed under for each program that was awarded to a community. These two new attributes are useful for the webmap user to quickly understand the funding status of a community.

Subawards Data Processing

For some programs, the prime transaction data does not accurately represent the flow of funds to individual communities, because those funds are re-distributed by the prime recipient. This situation occurred within the CDBG, SRF, and EDA programs, meaning it was necessary to include data on the subaward recipients. The data download and filtering process was the same as for the prime transaction level data. Data was downloaded from USAspending.gov, decompressed, and filtered using Python. An additional step was required, however, as EDA and CDBG programs are not exclusively water related. In order to ensure non-water related spending was not included in the output, the "subaward description" field, which was available for subaward data, was analyzed to include only project descriptions with water-related keywords. The keywords used to process the subaward data is listed below:

Table 2: *List of water-related keywords used*

Water-related Keywords
'water'
'pipe'
'drain'
'wastewater'
'sewer'
'sewage'
'drinking'
'sink'
'pump'
'storm'
'filtration'
'faucet'
'fountain'
'sump'
'stormwater'
'treatment plant'
'hydrant'
'plumbing'
'sanitary'

Any entries that lacked water related keywords were discarded. Subsetting the project description to these keywords was a significant improvement over leaving large numbers of non-water related

entries in the data. Once the data was filtered and processed, subaward data was combined with prime transaction data to create a more accurate representation of the flow of funds spatially under each program.

National Dataset Results

After the USA Spending data was processed, the output is an Excel table that summarizes the total funding that each city/community received across all 48 CFDAs. The table was organized by community names from the "recipient_city_name" column to be best aligned with the goals of the project in identifying which communities should be targeted for technical assistance. The total number of communities that received funding from the federal government between Fiscal Years 2000 through 2023 was 10,485. Table 2 offers a short description of the output table's contents.

Table 3: *Overview of the resulting table columns and description*

Column Header in Output Table	Short Description
recipient_city_name	This was the basis of how the data was categorized – each community/city had a singular unique entry in the table
recipient_county_name/ primary_place_of_performance_state_name	Each community/city had 2 corresponding columns - one providing the name of the county and the other the state of the community/city
CFDA ##.###	Each CFDA number had its own column that provided a breakdown of the total funding received for the program in each community/city – this led to 43 out of 48 columns for CFDAs* *more explanation below on why 43/48 CFDAs
ALL	This column represented the total funding that the community/city received across all CFDAs
WW/DW/Both	Each CFDA was categorized as relating to Wastewater, Drinking Water, or both. These 3 columns are similar to 'ALL', in that they represent the total funding the community/city received across all CFDAs for each of their respective categorizations* *more explanation below on WW/DW/Both
nr_awards	This column is to help the user easily identify how many CFDAs total a community was awarded funding through

award_names	The full name and funding agency for each awarded CFDA

Distinguish funding by eligibility for drinking water, wastewater, or both

The decision to leave the WW/DW/Both Classification from Table 1 on a short overview of the selected programs in the output table after processing the USA Spending data was made to make the data more accessible and understandable for the user. For example, we want to guide users such as technical assistance providers as quickly and efficiently as possible on which program types communities may or may not have received.

Total number of CFDA programs downloaded from USA Spending

As mentioned previously, there were 48 total programs that matched the description of federal funding relating to water infrastructure in the 2022 SAM APAL publication. Although these 48 programs were initially included in the data processing, the result included only 43 out of these 48 programs. Investigation within USA Spending provided explanations for each of the five missing programs in our final dataset, listed in Table 3 below.

Table 4: *The missing CFDAs from our analysis and explanation for why*

CFDA	Program Title	Why missing?
14.862	Indian Community Development Block Grant Program	All CDBG subawards were run against a water-related keyword list, and there were no matches
15.074	Small Surface Water and Groundwater Storage Projects	All have start dates of Oct./Nov. 2023, which is past the date that the data was initially downloaded
15.075	Blackfeet Water Rights Settlement	There were no results on USA Spending for this CFDA
66.448	Drinking Water System Infrastructure Resilience and Sustainability Program SDWA 1459A(l)	There were no results on USA Spending for this CFDA
66.958	Water Infrastructure Finance and Innovation (WIFIA)	There were no results on USA Spending for this CFDA

Non-positive Funding Sums: $0 Transaction Communities & Negative Funding Communities

When looking at the data, there may be some instances in which a community/city has $0 listed as their funding for each CFDA and does not appear to have any funding in the ALL and WW/DW/Both categories. There could be several reasons for the $0 transactions for obligations. As mentioned previously in the "*Funding Data Type"* section, there are multiple ways in which USA Spending reports funding information. For this project, the funding numbers per community were calculated

by looking at the "Total Obligated Amount," and there are instances in which the total for a community's obligations within this column is entered in USA Spending as equaling $0. Funding types can also include loans and subsidy costs, and not just grants. Every $0 entry is thus either representative of $0 entries in the raw data, or of positive and negative entries that precisely cancel. Extensive spot checks on $0 entries in the aggregated data confirm this - most $0 entries are the result of positive and negative prime transactions for a particular community canceling each other out, and the remainder are for awards where the "Total Obligated Amount" column value is simply recorded as zero. Understanding the data presentation and specific accounting practices of USAspending.gov is outside the scope of this project, so the choice to focus on a single general spending category was made.

While we are not looking specifically at those funding transactions, we included the communities that may appear to be a $0 transaction for Total Obligated Amount because we believe it is still important information, by demonstrating how the community may have already begun to create a connection with the federal government through funding mechanisms. This sets these communities apart from the ones that have had no prior relationship with the federal government for funding.

There are also instances where a community may have negative funding totals within the individual CFDAs or their overall funding total. This is because there are circumstances where funds are offered as a loan and require repayment to the awarding agency. This would appear as a negative total for the community in the prime transaction data used for the analysis. Similarly to the communities with unknown funding totals and represented as $0, the negative funding total communities were also preserved in the dataset, since a loan is still a form of financial assistance which the community received for water infrastructure.

From funding analysis results to a webmap

Formatting USA Spending in Excel for GIS Incorporation

Once the final output table was created for the USA Spending data, the next step was to begin to find the spatial locations of the communities. Using the Geography data type in Excel, which allowed us to pull information on each location from Bing Maps using the city and state information provided from USA Spending, we were able to get the latitude and longitude for most communities to convert them to spatial points in GIS.[7] For the communities that the Excel function did not assign coordinates to, locations were manually entered by searching for them in Google Maps. Some community names were entered with typos in the original data source, USA Spending, and the correct name was assigned by the Geography function in most cases. Similar to the missing coordinates, any typos that were not corrected by the Geography function were then searched for on Google to validate the correct spelling. Any duplicate entries for a community due to erroneous entries were consolidated into one entry, any obvious errors were removed, and all typos were

[7] https://support.microsoft.com/en-us/office/get-geographic-location-data-287b4cf2-3d7d-4bc1-b412-3d00f45dbbd6

corrected to the greatest possible extent. An example of an erroneous entry is if the community is mislabeled in either the city or state entry. For example, the USA Spending data had an entry for funding data in "Arlington, Utah". The Geography function on Excel was not able to locate coordinates for this listing, and after being unable to find this community with a search on Google, it was concluded that "Arlington, Utah" was an erroneous entry on USA Spending. There are multiple communities with the name 'Arlington' in the United States, and it would not be possible to identify what the correct 'city, state' combination would be for this entry. Therefore, to prevent inaccurate funding representation, the data entry for 'Arlington, Utah' was removed.

After cleaning the data up, there were a total of 10,485 communities that received federal funding within the 48 targeted CFDAs.

The USGS Populated Places dataset

Although it is important to know which communities have received funding and under which programs, it is also crucial for streamlining outreach and technical assistance to know what communities have not qualified for federal funding.

In order to compare the USA Spending point data to all communities, it was necessary to find a detailed dataset, developed by the federal government, of all communities in the U.S. that included very small communities. The "National Map" on the United States Geological Survey (USGS) website is an online interactive mapping tool which includes the "Populated Places" data layer – a robust, detailed layer that includes spatial points on more than 176,000 communities and available for download.[8] We use the USGS dataset in conjunction with the USA Spending dataset to get an idea of which communities have not had any federal funding in the timeframe analyzed. This can be accomplished utilizing both datasets in ArcGIS Pro by "merging" and "unmerging" the two datasets, which is explained further below.

Dataset Preparation in ArcGIS Pro – "City, County, State" Naming Convention

In preparation for the merging of the two datasets, it is necessary that they both have naming conventions in the form of "city, county, state" to ensure as much accuracy as possible. For the USGS data, this was simpler to accomplish as the dataset already had all three listed in their attribute table and were concatenated together in ArcGIS Pro.

However, an extra step was necessary for the USA Spending data to achieve the same goal of the 'city, county, state' naming convention as the dataset did not include reliable county information in the raw data file, and the Geography data type process in Excel did not provide consistent county names. Therefore, to get the county for each USA Spending community, an ESRI-published data layer called "USA Counties Generalized Boundaries" was used.[9] After the USA Spending data was brought into ArcGIS and converted into point data using the latitude/longitude, it was overlaid with the ESRI county boundary layer. From there, a spatial join was done so that the county names could be transferred from the ESRI data, to the USA Spending data. Where there were any potential gaps

[8] https://apps.nationalmap.gov/viewer/

[9] https://www.arcgis.com/home/item.html?id=3c164274a80748dda926a046525da610

in the county names, as the ESRI boundary layer sometimes did not extend fully to the edge of the coastlines, were manually inputted. From there, the data was concatenated in the same manner as the USGS data into a 'city, county, state' format.

Dataset Merge in ArcGIS Pro

Once both datasets have the communities listed in the "city, county, state" format, the geoprocessing tool "Join Field" is used.[10] This joins the data from the two datasets together by a similar input, which in this case is the same naming conventions ('city, county, state') in each of the USA Spending and USGS datasets. Joining the datasets allows us to identify which communities did NOT get a match between the two datasets, indicating that those communities did not receive federal funding according to USA Spending. After the two datasets were joined, the "Select by Attributes" tool was used to select the communities that had a "null" value in the funding columns. We then "unmerge" these non-funded communities by creating a separate data layer that only includes the communities that did not receive any funding. The final GIS output are two point data layers: 1) all of the USA Spending communities that received funding, and 2) the list of communities that originated from the USGS dataset but did not receive any federal funding. By creating these two data layers, it accomplishes both identifying where federal funding has been distributed as well as the crucial step of helping identify which communities may need focus for technical assistance in the future.

As mentioned previously, there are situations when the total funding for a community was listed as $0 in the prime transaction history but was still representative of an existing relationship between the community and federal government. Due to the nature of how it is represented in the raw data file for prime transaction data, it is unknown under which program(s) this "$0" funding is relevant to. Therefore, on the finalized webmap, we included the $0 total funded communities in the USA Spending point layer, as "Funding Received, in/out adds to $0," and marked "unknown" in the award names section of the popup. Although on a smaller scale it might be possible to investigate the specifics of $0 transaction communities to in order to understand what those records represent, at the scale of this project it was not feasible.

A very important disclaimer is that any inconsistencies in the data (such as spelling differences, typos, inaccuracies) were corrected to the best of our knowledge and ability, but there may be some remaining. The easily-identifiable errors (such as common typos and special characters) were addressed and the data was joined to the best of our ability without comparing each individual entry of the 10,000+communities on USA Spending to the USGS dataset of 176,000+ communities.

Final Webmap Product

To enable technical assistance providers to locate communities based on their past funding status, the federally funded communities as well as the not-federally funded communities for the analyzed timeframe were added to a webmap application, found here. This webmap has additional layers that the user can toggle on and off. The user can also filter communities by size and number of awards received. Additional widgets allow the user to add data layers from their own source or

[10] https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/join-field.htm

https://livingatlas.arcgis.com/en/home/, print the map on the screen to a PDF, and save bookmarks for a geographic area.