Analyzing grants and other philanthropic transactions
About Candid Insights

This manual is written by Candid Insights. Candid's Insights team aims to make meaning out of Candid's data and share these insights with the field. These insights are often reflected in Candid’s blogs, research reports, infographics, and manuals. The recommendations in this manual on how to understand and analyze Candid data for research reflect decades of evolving institutional knowledge and standards.

Contributors

Anna Koob
Senior lead researcher

David Wolcheck
Research manager

Cathleen Clerkin
Senior director of Insights

Guy Mika
Research operations associate

Correspondence

Please contact Candid Insights at researchteam@candid.org with questions or to learn more about this data.

Citation

Candid (2023). Candid research manual: Analyzing grants and other philanthropic transactions. doi.org/10/gsfhbw

DOI

doi.org/10/gsfhbw

Copyright

© 2023, Candid. This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License, creativecommons.org/licenses/by-nd/4.0.

Date last updated: July 14, 2023
Contents

4 Overview

6 Part I: Grant/transaction data collection and sources
   7 Three primary sources of grant/transaction data
   9 Data collection timeline
   9 Proportion of data sources over time
  10 Minimum data requirements
  11 Data source hierarchy and de-duplication (i.e., survivorship)

13 Part II: Key data considerations for conducting research
   13 Dynamic data
   15 Timeline: Changes to the data over the years
   18 Limitations of recent data
   19 Possible double-counting of funds
  20 Why Candid may not have a particular transaction

22 Part III: Datasets for conducting research
   22 Transactions set
   26 Foundation 1000
  30 Side-by-side comparison of the two research datasets

32 Part IV: Key variables and data fields
   32 Transaction type
   33 Paid and authorized transactions
   35 Funder type
   37 Internal grants and pass-through funding
   39 Fiscal year
   39 Geographic location
   41 Source type
Overview

Over the past several decades, Candid has collected data on tens of millions of grants and other philanthropic transactions awarded by hundreds of thousands of funders located in the U.S. and around the world.\(^1\)\(^2\) This manual provides an overview of this data and includes guidelines and key considerations for its use in research. The document draws on the approaches and methods that Candid’s Insights team has developed over time to analyze transactions data. It is intended to increase transparency, help others better understand Candid’s data, and offer advice on how to incorporate this data into future research.

The manual is divided into four parts:

**Part I: Grant/transaction data collection and sources** outlines Candid’s sourcing methods for this data. The three main sources are: government (e.g. regulatory documents such as the IRS forms 990, 990-PF), funders themselves, and other public sources (e.g. organization websites). This section also discusses sourcing timelines, highlighting the fact that direct reporting by funders and public sources may offer very recent transactions, while government sources generally have significant time lags. This means that Candid won’t have comprehensive transactions data for 2-3 years after the calendar year has passed.

---

1. Grants represent more than 99 percent of all transaction types in Candid’s database, which also includes other types like program-related investments, in-kind gifts, and stock transfers. Given the dominance of grants as the primary transaction type we capture, the term is often used to refer collectively to all transactions awarded by philanthropic institutions. Throughout this guide, we’ve elected to use the broader terms “transactions” or “contributions” unless we are exclusively referring to grants.

2. In addition to philanthropic transactions, in 2016 Candid began collecting data on U.S. federal grants awarded since 2014. This specific type of funding data is not the primary focus of this research guide, but does appear in Candid products and tools and constitutes a large proportion of the total dollar value of funding that Candid collects.
Part II: Key considerations for conducting research focuses on central aspects of Candid's transaction data that researchers need to understand in order to accurately create, analyze, and interpret the data. This section includes an explanation of Candid's dynamic (“live”) data systems, and information about when data might be incomplete, double-counted, or missing. It also offers a timeline that details changes to Candid's sourcing and coding methodology over the years and emphasizes that changes to data collection (e.g. processing higher volumes of data around 2014-2015; increased collection of public sourced data in 2020) should not be conflated with changes in giving.

Part III: Data sets for conducting research presents the two main data sets that Candid Insights staff use for research: the Transactions set, which represents a cleaned version of all transactions data that Candid collects and processes; and the Foundation 1000, a subset of the former, limited to transactions of $10K or more awarded by 1,000 of the largest U.S. private and community foundations and created annually. The key use case for the Foundation 1000 is for trends analysis and this section describes the considerations and process used to create this standardized set each year. A side-by-side comparison also illustrates the main differences—e.g., in terms of the types of funders and transactions included—in the Transactions set and the Foundation 1000 subsets.

Part IV: Key variables highlights and describes important fields and variables that Candid researchers use to clean and analyze transactions data. In addition to defining these data fields, this section offers best practices for filtering, analyzing, and interpreting them. Identifying different funder types, understanding the difference between paid and authorized transactions (and how this shows up in Candid's data), and accounting for the double-counting of regranted funds are among the featured topics. The overview is oriented around Candid's Grants plus data extract and references specific fields included in that output format.
Part I: Grant/transaction data collection and sources

Why “transactions” and not “grants”?

Throughout this research guide, we intentionally use the term “transactions” to collectively refer to all types of contributions awarded by funders in Candid’s data. However, we recognize that the vast majority of these transactions (99%) are cash grants and conversationally the term “grants” is often used to more broadly mean “transactions”. Throughout this guide, we’ve elected to use the broader term “transactions” (or “contributions”) unless we are exclusively referring to grants.

Every year, Candid processes data on millions of grants and other philanthropic transactions representing hundreds of billions of dollars in funding. This data comes from a wide variety of sources, including data from government agencies (such as the Internal Revenue Service in the U.S.), from funders themselves and other data sharing partners, and—increasingly—from organizations’ websites and news sources. Data from these disparate sources is cleaned and coded according to Candid’s Philanthropy Classification System (PCS), and used to create structured and unified datasets that offer information about philanthropic giving by institutions and select high-net-worth individuals.

The most common type of transaction in Candid’s data is cash grants. However, Candid datasets also provide information about other types of transactions such as in-kind gifts, program-related investments, and pledges. For the full list of transaction types that Candid tracks, see: https://taxonomy.candid.org/transaction-type.

Transactions in Candid’s database go back as far as 2003 and are added and updated on a daily basis. Candid’s data sources and coding methods have evolved over the last 20 years, so a good understanding of these approaches is highly recommended before conducting research based on transactions data. This is especially important when examining trends over time and/or drawing conclusions about very recent funding.
Three primary sources of grant/transaction data

1. Government sources (including Internal Revenue Service (IRS) Form 990 and 990-PF).

U.S.-registered foundations are required by law to include lists of grants and contributions made in the given fiscal year on their IRS filings. Candid processes all available data on these transactions for private foundations, as listed on Part XIV on the IRS Form 990-PF. For public charities, including community foundations, Candid only processes transactions to domestic organizations (listed on Schedule I, Part II of the IRS Form 990). Other transactions—such as those to domestic individuals (listed on Schedule I, Part III), or organizations or individuals based outside of the U.S. (listed on Schedule F, Parts II and III)—have not historically been included, but will likely be in the future (this guidance will be updated accordingly at that time).

IRS Forms 990 and 990-PF are the most comprehensive sources of data about institutional philanthropic giving by the roughly 120,000 U.S. foundations. However, there are limitations to this data source, including: the slow rate at which these forms are made publicly available (typically 1-2 years after the given fiscal year ends—often made worse by delays); the limited descriptions that filers provide; and the fact that these sources are limited to giving by U.S. private foundations and public charity grantmakers only (and not direct corporate giving, for example).

Candid collects 990 data in multiple ways:

- For the increasing number of organizations that file electronically, Candid directly downloads filings data (saved in XML format and packaged as zipped files) from the IRS website, as they become available;

- For organizations that submit a paper filing, Candid works

3. Following the enactment of the Taxpayer First Act, organizations filing Forms 990, 990-EZ, and 990-PF are required to file electronically for tax years beginning after July 1, 2019. As a result of the full transition to efilling, we expect the need to outsource paper filings will be phased out entirely. In fiscal year 2009, 72 percent of the filings we received were submitted on paper. In 2019, that proportion was down to 22 percent.
with a third-party processor to convert the image files released by the IRS into machine readable XML documents³;

- Candid also sources IRS forms directly from funder websites. In these instances, the form has been posted by the organization but has not yet been processed or released by the IRS. As with paper filings, Candid works with a third-party processor to convert these files to machine-readable data.

Outside the United States, Candid also collects data from select government agencies (such as charity commissions) that release open data about the charitable sector or other philanthropic actors in their countries. (Details on specific sources by country are available here: https://candid.org/use-our-data/about-our-data/data-sources.)

2. Funder contributed data.
Hundreds of funders around the globe share their transactions data with Candid, either directly or via partner organizations.⁴ This funder contributed data is usually more recent than data received from a government agency (e.g. IRS) but is limited to those who elect to participate—and participation can’t be guaranteed year over year. Many of the very largest U.S. foundations (such as the Bill & Melinda Gates Foundation, Ford Foundation, and Hewlett Foundation, among others) regularly share their data with Candid directly. Funders choose how frequently they share data. A majority share their transactions data on an annual basis, but some share data monthly or as they make or announce awards. For a complete list of participating organizations, please see Candid’s data sharing partners page. This list is updated on a monthly basis.

3. Other public sources.

- **Organization websites.** Candid staff proactively search websites to look for transaction data specific to current issues. Exploration of methods for sourcing website data more systematically is currently underway. Data sourced

---

4. For example, since 2010, Candid (then-Foundation Center) has partnered with the Human Rights Funders Network (HRFN), who collects data from their members and other affiliated member organizations for inclusion in Candid’s databases. This data is featured in the annual research initiative, *Advancing Human Rights.*
from organization websites may include lists of transactions or searchable databases posted publicly by a funder or data compiled by another organization. The U.S. Chamber of Commerce’s Disaster Corporate Aid Trackers and 360Giving are two examples of the latter.

**News, social media, and press releases.** In 2020, Candid ramped up efforts to source data from news articles, social media, and press releases. Candid staff perform online searches of news sources and social media and also review funder newsletters for recent grantmaking announcements. Candid has also developed a news scraper that scours roughly 300,000 news and social media articles every day and identifies news about the social sector. When applicable, philanthropic transactions sourced from news, social media, and press releases are included in Candid’s datasets—though these make up a relatively small number of transactions in the database (see Figure 2 on page 11).

Efforts to source transactions data from the news, social media, and press releases tend to be very targeted and prioritize high profile issues of a given moment, such as Covid-19, racial equity, or the response to specific disasters. Although the majority of grants and other philanthropic transactions are not publicly announced, these public sources offer the most up-to-date insights about philanthropic transactions on critical topics.

**Data collection timeline**

It takes several years for Candid to collect a critical mass of philanthropic transactions for a particular year. For example, in mid-2022, Candid had collected most transactions data for 2018 but was still actively collecting data for years 2019-2021 (see Figure 1 on next page).

**Proportion of data sources over time**

For most recent years, funder-contributed data represents the largest source of grants and other transactions in Candid’s database by a substantial margin (based on the number of records). As Candid collects more data, government sources
(primarily the IRS) overtakes funder-contributed data as the largest source. Government sources are slower to arrive but eventually account for about 75-85 percent of the transactions in Candid’s database for a given year (see Figure 2 on next page).

Minimum data requirements

In order to load data about a transaction into our database, Candid seeks out the following information: funder name, recipient name, recipient location, fiscal year, amount (classified as either paid or authorized), and—ideally—a description. Whether any information beyond this—such as the population or geographic area served—is available, varies by source.

Transactions data from government sources, such as the 990-PF and 990 annual returns filed by funders, usually have

---

5. In some instances, recipient names may appear as “Multiple recipients” (where the data source reports a single amount but names more than one recipient) or “Unknown recipient” (where the data source specifies an amount but does not identify the recipient organization). In rare cases, funder name, recipient name, and/or location have been anonymized by the funder or by Candid in response to a security or privacy risk. In some exceptional circumstances, transactions amounts may be missing from a record. This often is indicative of a loading error, correctable by Candid staff.
very limited information. Funders who share their grants and other transactions data with Candid directly have the option of providing additional details about their contributions, including title, program area, and key characteristics, such as subject area, population served, and geographic area served (a complete list of data fields and their definitions can be found on the grants data template).

Data source hierarchy and de-duplication (i.e., survivorship)

Given that Candid’s datasets incorporate multiple data sources, there are inevitably cases in which information about the same grant (or other transaction) appears from multiple sources, often at different times. For example, a grant may first enter the Candid database based on information from a funder’s press release; and years later that same grant may appear again in the funder’s IRS Form 990-PF. To prevent transactions from being duplicated in the database, Candid uses a system of “survivorship” to determine which source of data gets precedence. Complete sources—that is a source that includes all transactions made by

Figures and notes:

Notes: Limited to transactions awarded by U.S.-based foundations. ‘Other public sources’ includes transactions sourced from 990 or 990-PF filings posted to a funder’s website in cases where Candid processed the form directly rather than waiting for the IRS to release it. These proactive efforts account for the relatively higher proportion of transactions from other public sources in 2019.
a funder in a particular year—are typically prioritized over partial ones. This system works as follows:

**First priority: Complete funder-contributed data.** Data that funders share directly with Candid overrides other sources—whether that funder is based in the U.S. or elsewhere. For example, if Candid receives a complete list of transactions from a funder for a particular year, it would replace the transactions coming from a government source (such as the funder’s IRS Form 990 or 990-PF) for that year. An exception occurs in cases where the list of philanthropic transactions shared with Candid does not represent a complete year of that organization’s giving.

**Second priority: Government sources.** After data shared directly from funders, a complete list of transactions from a government source, such as the IRS, is given second priority. Government-sourced transactions typically have less detail than funder-contributed data but may be considered more reliably “complete” (i.e., a source captures all transactions awarded by a funder in the year in question) than other public sources.

**Third priority: Other public sources.**

- **Organization websites.** Transactions sourced from organization websites might represent a full year of an organization’s funding or just a selection of transactions. If, upon manual review, the total value of transactions is comparable to giving in previous years, it is treated as a complete source. If the list of transactions appears to be partial, it is replaced with a more complete source if Candid receives one.

- **News, social media, and press releases.** Transaction information obtained from news sources often offer the most up to date insights about giving, however, that information also tends to reflect only one or two contributions made by a funder. For that reason, these data sources are replaced by a higher priority data source once a complete list of transactions is available for that particular funder in the year in question. An exception is the transaction type of pledges. A pledge remains in Candid’s data as originally entered even as Candid receives information about transactions paid out in fulfillment of that pledge (for more on our standard treatment of pledges for research purposes, see the section on transaction types under Part IV of this guide.)
There are several aspects of Candid’s data that are important to understand and account for to create, analyze, and interpret the data accurately. Such key considerations include understanding Candid’s “live” data collection processes; how the data has grown and changed overtime; limitations of recent data; when and how to identify possible double counting of funds; and when and why a particular transaction may be missing in Candid's datasets.

**Dynamic data**

As the previous section on data source hierarchy survivorship suggests, Candid’s data is “dynamic”: i.e., it is constantly being updated. Every day, more complete data sources replace partial sources, new transactions are added, data details are updated, errors are corrected, and in some cases, transactions are removed. Candid staff is continually cleaning data entries and adding or updating codes to improve accuracy, which affects how transactions are classified. Organizational-level details may also change as staff merge duplicate records and update contact details, mission statements, and coding.

Notable exceptions to the ‘live data’ rule are the Foundation 1000 datasets. These exist as frozen sets, which reflect the data as it appeared in Candid’s database at the point the set was created. The Foundation 1000 sets are, however, periodically updated as major changes to Candid’s underlying data are introduced—for example, when Candid’s taxonomy, the Philanthropy Classification System (PCS), is revised and all transactions are re-coded to align with the new version. When a Foundation 1000 set is updated, all data in that set will be updated to reflect the
information currently in Candid’s database at that moment. The set will exist with that data until it is updated again. (See Part III for more information about the Foundation 1000).

**Evolving methods.** In addition to the minor daily adjustments that come with a dynamic dataset, Candid has introduced more substantial changes to data collection processes over the years, as technology and data availability have evolved. Some of these newer methods have resulted in large increases in data collection in recent years, as well as a re-coding of some older data. Please refer to the **Timeline** section of this document (starting on page 15) to learn more about these methods. It is advised that researchers consider these changes before conducting longitudinal analyses.

**Limited historical records.** Implicit in the discussion of dynamic data is the fact that older data gets overwritten. In other words, there are limited historical records in the database. For example, currently, when an organization changes addresses, the old address is overwritten in Candid’s database and replaced with the new one. This means that any transactions awarded by or to that organization in the past will reflect the updated address. Similarly, when changes are made to specific transactions—e.g., codes are added or removed—there is typically no historical record of previous codes.

In short, researchers should note that Candid data—at either the organization or transaction-level—reflects the most recent changes made to a record. This information may have changed since a particular data point initially appeared in Candid’s database.

**Managing sensitive data.** In rare cases, Candid may remove, anonymize, or otherwise modify publicly available data to prevent potential harm. For example, in 2021, following the Taliban’s takeover of Afghanistan, Candid removed details about funding to Afghanistan due to growing concerns about safety and retaliation. Because of this, on rare occasions, there might be inconsistencies in datasets pulled at different points in time (i.e. before/after sensitive data was removed). For more on the considerations that go into these decisions at Candid, see this blog post.
Timeline: Changes to the data over the years

Candid has tracked transaction-level institutional philanthropic giving for over 20 years. During this time, changes in IRS reporting, technology, and the field of philanthropy have prompted updates to our data collection and coding methodology to offer the most comprehensive, current, and useful data possible. Researchers should understand these changes and how they impact the data before conducting analyses—especially if comparing trends over time—so they do not mistake artifacts in the data for findings in the field.

Table 1: Timeline of changes to the data

<table>
<thead>
<tr>
<th>Year changes occurred</th>
<th>Years of data impacted</th>
<th>Nature of change &amp; research implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014 and prior</td>
<td>6</td>
<td>Transactions data collection largely limited to largest U.S. funders; annual set creation based on latest available data</td>
</tr>
</tbody>
</table>

Candid (then-Foundation Center) largely focused on collecting transactions of $10,000 or more awarded by a sample of the largest U.S. private and community foundations: the largest 800 nationally plus the top 25 in each state and Washington, D.C. as measured by total giving. Transactions data were cleaned and coded by an in-house team, with some use of rules-based coding.

To allow for year-over-year comparisons, these data sets were “frozen” each year to create an annual set known as the “Grants Indexing” or “GI” sets. The GI sets incorporated transactions from the most recent available fiscal year for each grantmaker at the time of set creation. For example, for the 2003 set, the majority of included grantmakers would have their fiscal year 2003 contributions represented. But in cases where 2003 contributions were not yet available for a particular funder, the set would include that funder’s fiscal year 2002 contributions. The actual number of foundations in the set varied each year, but generally there were between 900 and 1,300 organizations.

Additional individual transactions from organizations outside the GI sets were included for other grantmakers to serve as examples of the work those organizations

6. Where applicable, we’ve distinguished between the calendar year and the data year, which recognizes that changes made in a given calendar year are applied to the data we’re processing at the time. Due to the time lag between when a contribution is made and when it becomes available as data for us to process, there may be 1-3 years difference between a given calendar year and the fiscal year associated with the data we’re processing at any given moment.
<table>
<thead>
<tr>
<th>Year changes occurred</th>
<th>Years of data impacted</th>
<th>Nature of change &amp; research implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014 and prior</td>
<td></td>
<td>funded. Toward the later end of this period, then-Foundation Center also began to collect transactions from non-U.S. funders, federal grantmakers, and other grantmaking entities besides community and private foundations. There was also some limited collection of less common transaction types, like pledges—typically to support special research projects. Due to data loss, only a subset of contributions collected prior to 2001 are available in Candid databases.</td>
</tr>
</tbody>
</table>
| 2013-2014              | New methodology starts with 2012 data; Data sets going back to 2002 were recreated to mirror new methods | **Annual set methodology changes to ensure consistent fiscal year is represented**  
Candid (then-Foundation Center) introduced the *Foundation Center 1000* (FC 1000) as a replacement for the GI sets. The FC 1000 annual set consisted of 1,000 of the largest U.S.-based private and community foundations nationally. Rather than combining transactions from different fiscal years for various organizations, each FC 1000 set included data from a consistent fiscal year. The set still focused on transactions of $10,000 or more. The initial set was created with FYE 2012 transactions; at that time historical sets from 2002-2011 were created with available transactions data to align with the new FC 1000 criteria. This methodological approach to creating an annual set is still in effect. In 2019, the name of these sets was simplified to the *Foundation 1000*. |
| 2015                  | Beginning with 2013/2014/2015 data | **Introduction of automated processes increases the volume of available transactions data**  
Starting in calendar year 2015, Candid (then-Foundation Center) introduced automated data loading processes and machine learning models, in part to take advantage of the IRS beginning to release tax return information in machine-readable XML format. In addition to speeding up the loading, cleaning, and structuring of data, automated processes were used to code transactions to relevant subject, population, support strategy and transaction type categories. Automation allowed Candid to process more transactions (below the $10,000 threshold) from more funders (beyond transactions from the largest U.S. private and community foundations) and substantially increased the number of transactions in Candid's database. To support automation and increased volume, a new data system—called the EDMS—was developed and introduced. Automated processes were initially applied to fiscal year 2013 data during the 2015 calendar year. These changes led to an increase in the number of yearly transactions available in Candid's database, which grew through 2014 data collection and into 2015 at which point the new, higher, volume of transactions began to stabilize. The overall increase in available transactions data around this time should not be conflated with an increase in actual giving by funders. |
<table>
<thead>
<tr>
<th>Year changes occurred</th>
<th>Years of data impacted</th>
<th>Nature of change &amp; research implications</th>
</tr>
</thead>
</table>
| 2015                  | Beginning with 2001 data | **Expanded taxonomy, the Philanthropy Classification System, is introduced and transactions are re-coded**  
In 2015, Candid (then Foundation Center) introduced a new taxonomy, called the Philanthropy Classification System or PCS, to capture more detail about grantmaking and nonprofit activity not offered by existing taxonomies at the time. The PCS and its over 1,200 codes are used to describe the work of grantmakers and recipient organizations, as well as the philanthropic transactions between those entities. The PCS is organized into categories of code or facets: subjects, populations, organization type, transaction type, and support strategy.  
When the PCS was introduced, transactions from 2001 on were re-coded to align with the new taxonomy.  
The predecessor to the PCS was the Grants Classification System, which Foundation Center had also developed and based on the National Taxonomy of Exempt Entities (NTEE). The NTEE system was developed by the IRS and the Urban Institute's National Center for Charitable Statistics and is still used to classify nonprofit organizations. For more on the PCS and other coding systems that Candid uses, see the forthcoming research guidance on this topic. |
| 2015/2016/2017       | Beginning with 2015 data | **New processes are developed to manually review coding of select transactions**  
Due to increased automated coding, including the use of machine learning algorithms to apply PCS codes to transactions in 2015 (see last two points above), new manual review processes were also developed. To introduce manageable quality control, staff began to review transactions of $250,000 or more from U.S. foundations included in the *Foundation 1000* set.  
Staff review additional transactions on an ad hoc basis, particularly on sets used to produce research with partner organizations—such as the Human Rights Funders Network and the Center for Disaster Philanthropy. |
| 2016                  | Beginning with 2014 data | **Federal grantmaking is added to database**  
Starting with 2014 data, Candid increased its collection of grantmaking by select U.S. government agencies. (For previous years going back to 2009, government grants collection had been limited to a few agencies like the National Science Foundation and Department of Education). Grantmaking by U.S. government agencies is sourced from usaspending.gov. Federal grants are processed and indexed to the PCS like all other transactions, but Candid does not perform any extensive manual review of this data before adding it to the database.  
Note: For years 2015-2018, federal grants account for 80-90% of all transactions dollars in Candid's database. Given the very large values and different nature of this funding, it is standard practice for Candid's Insights team to exclude federal grants from analyses about philanthropic funding. |
<table>
<thead>
<tr>
<th>Year changes occurred</th>
<th>Years of data impacted</th>
<th>Nature of change &amp; research implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2019</strong></td>
<td>Beginning with 2001 data</td>
<td><strong>PCS is revised and updated</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Candid released an updated version of the PCS in April 2019. Changes included: the introduction of new terms, the removal of others, updates to existing terms and definitions, and changes to the hierarchy/structure of the taxonomy. All transactions in the database were re-coded to align with the new taxonomy. To see the specific changes introduced, see: <a href="https://taxonomy.candid.org/resources/downloads/full-pcs-taxonomy-with-definitions">https://taxonomy.candid.org/resources/downloads/full-pcs-taxonomy-with-definitions</a></td>
</tr>
<tr>
<td><strong>2020</strong></td>
<td>Beginning with 2020 data</td>
<td><strong>Efforts to collect very recent data are ramped up and a new code for pledges is introduced</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>To improve data recency, in 2020 Candid expanded existing efforts to collect data about newly announced grants and other transactions. These methods include the manual review of funder websites, news sites, social media, and organization newsletters in addition to some automated scraping of similar sources. Data collected this way is more likely to capture “newsworthy” transactions (e.g. large donations or pledges from corporations, large foundations, or high net worth individuals; or transactions associated with a current event or crisis). Therefore, caution and due diligence should be taken before generalizing very recent data to trends in the field at large. Additionally, Candid began coding pledges as a distinct transaction type. Pledges reflect a funder’s public promise to dedicate a certain (typically large) amount of resources to address a crisis or issue, often over a set number of years. In contrast, grants and other transaction types reflect funding actually committed or paid to a specific organization or organizations. To avoid conflating the two, researchers should analyze pledges separately from grants and other transaction types.</td>
</tr>
</tbody>
</table>

**Limitations of recent data**

Due to the slow nature of IRS data collection, Candid typically won’t have a critical mass of transactions for a given year until 2-3 years after that calendar year has ended. Additionally, much of Candid’s most recent data (from the past 1-2 years) is based on what funders have shared with Candid or what’s been publicly announced (see **Figures 1** and **2** on pages 10 and 11).

Therefore, the most recent few years of Candid data should not be assumed to include the full universe of philanthropic transactions for those years. When examining transactions
awarded in the past 2-3 years, researchers should treat the data as a nonrepresentative sample with limited generalizability to all philanthropic giving. Transactions awarded by high net-worth individuals and corporations, as well as very large transactions or pledges will likely be overrepresented. Furthermore, researchers should avoid comparing recent years to older years or drawing conclusions about how funding has changed in recent years, unless sufficient research methods are used to test and address these limitations (e.g. using a smaller matched set of grantmakers and/or inferential statistics to assess differences).

Possible double-counting of funds

A side effect of having a comprehensive transactions database is that occasionally, the same money might be recorded twice. One example of this happens when funders publicly “pledge” a certain amount of money to a cause (e.g. “Foundation X pledges 1 million dollars towards ending issue Y”) and then later pays out this pledge through a number of smaller grants in subsequent years. To ensure such funds are not double counted, researchers should not sum “pledge” transactions and “cash grant” transactions when calculating total dollars for a particular area (see page 33 for further discussion).

Another case where double-counting is possible is when grants are awarded to other funders for re-granting purposes. An example is when a private funder contributes to a “regranter” like a community foundation, who may then award that money to a nonprofit such as a local food pantry. In cases where a data set includes transactions awarded both to and from a specific regranter, those dollars may be counted twice. For some research questions, it may not matter if both are included in a data set (e.g. a study focused on the number of organizations working on a particular topic). However, for research questions that require calculating exact dollar amounts granted to a specific cause and/or during a specific time period, it is important to exclude potentially double-counted transactions (see page 37 for methodological recommendations on this topic).
Why Candid may not have a particular transaction

It is impossible to guarantee that Candid’s data collection will capture all transactions by all funders within a given year. There are a number of reasons why a particular contribution may not be represented, for example, Candid may not have a particular transaction because:

- The funder is a non-U.S. funder who hasn’t shared their data with Candid. Transactions data from non-U.S. funders is typically limited to data that they’ve shared with Candid directly.

- The transaction was made recently and Candid hasn’t received it yet—either from the funder directly or from a government source (like the IRS in the U.S.)

- The transaction was announced publicly but it has not yet been identified by Candid’s data collection efforts (or it was missed)

- For U.S. funders, there may be an error in the IRS processing of the filing that cuts off the attached list of transactions. If Candid doesn’t have another data source for that funder, transactions may be missing for this reason.

- For U.S. grantmaking public charities including community foundations, the transaction may have been awarded to an individual or a non-U.S. organization. Currently, Candid limits processing of 990 transactions to domestic awards to organizations (listed on Schedule I, Part II).

- The transaction was awarded by a funder type whose grantmaking isn’t always made publicly available (e.g., high-net-worth individuals, direct corporate giving programs).

- The transaction was awarded through a donor advised fund (DAF). Grants awarded through DAFs aren’t required to be made publicly available (via a 990 filing or elsewhere). For community foundations, Candid may have transaction-level data about contributions awarded through the foundation’s discretionary fund, but not necessarily through their donor-
advised funds unless they elect to share that information with Candid.

In other cases, Candid may have *received* a transaction but it doesn’t appear in the datasets because:

- The organization receiving the contribution hasn’t yet been authorized or otherwise approved to appear in Candid’s data. Candid staff review seemingly new recipient organizations to ensure they’re not being duplicated in the database.

- A processing error is preventing the transaction from appearing in Candid databases.

- The transaction doesn’t meet inclusion rules for a particular data set, tool, or product. (e.g. see Table 2 for inclusion rules for transaction datasets).

- In rare cases, Candid may remove, anonymize, or otherwise modify data that Candid shares publicly to prevent potential harm (see page 14 for more details on how Candid manages sensitive data).
Part III: Datasets for conducting research

Depending on the type of research question and analysis, Candid’s Insights team utilizes either the full collection of transactions in Candid’s database or a specific subset of transactions, the Foundation 1000. The former, referred to as the Transactions set, includes the widest cut of Candid’s data on grants and other philanthropic transactions. The latter subset, the Foundation 1000, is an annual set of grantmaking by 1,000 of the largest U.S. funders to allow for trends analysis. Researchers should review the differences and similarities between these two datasets and determine which one is better suited for their project.

The standard view of these sets includes information about the funder, recipient organization, and the transaction itself. See here for a sample view of this data with included fields and a data dictionary. Note the same variables/data fields are available for all transactions regardless of whether they belong to the Foundation 1000 subset, or the Transactions set.

Transactions set

The Transactions set represents nearly all of Candid’s data on grants and other philanthropic transactions going back to 2001. For more recent years, this amounts to 2-4 million transactions captured each year from around 100,000 funders. This data set offers the broadest, most comprehensive view of the philanthropic funding landscape: including a wide variety of funders (e.g. smaller funders, non-U.S. funders, and in recent years, some high-net-worth individuals) and a broad range of transaction types (e.g. cash grants, pledges, in-kind gifts). The set is updated on a daily basis to capture the latest funding information available.
Considerations and recommendations for using this set

The Transactions set is the best source of up-to-date funding information. Because it is a large and frequently updated set, it’s particularly useful for research focused on producing descriptive statistical estimates of the amount of funding given by institutional philanthropy. It is also useful for research projects focused on giving by small funders, or examining emerging trends based on recent news headlines. In terms of data sources, the Transactions set is made up of a compilation of all data sources described on page 7, based on the survivorship methodology outlined on page 11. However, because this “kitchen sink” dataset includes nearly all transaction data collected by Candid, it is subject to idiosyncrasies as new data trickles in, and coding practices evolve. Of particular note:

Data is comprehensive, but not always complete, especially for recent years. The Transactions set does not capture contributions from all funders. Moreover, included transactions do not necessarily reflect each funder’s full year of grantmaking—especially for more recent years as data is still being collected. Data may come from “real time” sources—e.g., the news, social media, newsletters, funder websites, etc.—that only capture a partial list of funder’s awards for that year. Additionally, funders may only share a partial list when they share their data with Candid directly. Processing errors—either external or internal to Candid—as well as Candid’s data cleaning rules (such as requiring that transactions be awarded to a recipient organization that is ‘authorized’ or otherwise vetted) are other possible reasons for incomplete lists. Moreover, funding from the most recent 3 years of data should be considered “incomplete”, as data collection for recent years is on-going, and numbers may change from one day to the next as additional funding sources are identified and processed. For these reasons, a best practice is to include the disclaimer that analyses based on this set are “Based on available data”. It is also recommended that researchers perform their own completeness checks by comparing a funder’s total giving (e.g. as reported on a 990 filing) for a given year with the total value of transactions that Candid has available from that funder for that year—especially for large grants and funders that account for a large portion of giving within a particular data set.

The volume of data in the Transactions set varies year to year. This is due in part to Candid’s increasing capacity to collect and
index transactions. Since 2014, Candid has enhanced its ability to process transaction data, resulting in both a higher volume of smaller transactions and transactions from more organizations being loaded into the set. Another reason for year-to-year variation may be due to shifts in source availability: for example, a funder who shares data with Candid one year may not do so the next. If Candid isn’t able to source the transactions using another method, there will be gaps in that funder’s data.

**The Transactions set includes more than just grants.** As mentioned previously, Candid’s data includes many different transaction types that describe how a particular organization may provide or receive support (e.g. volunteer services, in-kind gifts, matching grant). In some cases, a transaction may receive more than one type: for example, a grant that’s being made to match funds provided by another donor may receive codes for both “matching grant” and “cash grant”. All transaction types are included in the Transactions set. However, most researchers are specifically interested in cash grants. Therefore, researchers using this dataset should make sure to filter out transaction types they are not interested in (see the section on transaction types under Part IV of this guide for more on this subject).

**Trends analysis in the Transactions set should be handled with care.** Although transactions in the Transactions set date back to 2001, changes in data collection methods mean that the volume of transactions and the nature of the funders and transaction types represented have changed over time. These changes are detailed in the timeline starting on page 15.

Of particular note, the automatic processes that Candid introduced starting with 2014 data enabled Candid to process additional types of funding (e.g., federal grants, public charity grantmaking, pledges, gifts from high net-worth individuals), which substantially increased the number of transactions in Candid’s database (see Figure 3 below showing the value and number of philanthropic transactions in the Transactions set since 2006 as copied from Foundation Maps.) Because these volumetric changes in the total number and value of the transactions in Candid databases reflect a change in data collection methods, rather than changes in giving, the Candid Insights team does not recommend using the Transactions set to compare funding before and after 2014 without additional data.
Researchers should also avoid including the most recent 2-3 years in longitudinal trends analysis as data collection for these years is considered incomplete. Moreover, pledges, corporate giving programs, and high net worth individuals should also be excluded from longitudinal trends analysis as methodological changes introduced to the set in 2020 increased the representation of these funders and transaction types (see the relevant section of the timeline on page 18 for more information).

Those interested in conducting a trends analysis inclusive of validation.7

Note: Figures exclude grantmaking by U.S. federal funders.

There is a large increase in the value and total number of transactions from 2014-2015 due to changes in Candid’s data collection methodology.

7. This chart includes all transactions in Candid’s Transactions set, including one-time deliveries of large data sets. For example, for 2015 and 2016 Candid received around 700,000 grants from Fidelity Investments Charitable Gifts Fund, which they voluntarily reported and which captured much of the giving made through the DAFs they host. In 2017, the number of grants that Candid received from this funder was closer to 35,000. This drop largely accounts for the large drop in transactions in 2017 and illustrates how the Transactions data set is very much subject to fluctuations based on how much data select grantmakers elect to share in a given year.
years before and after 2014 may want to consider using the Foundation 1000 dataset; or creating a limited ‘matched set’ of grantmakers for whom all transactions over a given period are included.

**Foundation 1000**

The Foundation 1000 is a subset of the Transactions dataset—in other words, all data in the Foundation 1000 can also be found in the Transactions set (but not vice versa). The Foundation 1000 is a curated dataset that consists of all grants (plus some additional transaction types detailed below) of $10,000 or more awarded by a set of 1,000 of the largest U.S. private and community foundations for a given year (based on the organization’s fiscal year end). While this dataset may at first seem quite small (e.g. it represents only 1% of the total number of funders in the Transactions set), the set includes a significant portion of total U.S. philanthropic dollars awarded each year. This is because institutional grantmaking data is incredibly skewed, with the largest foundations accounting for a disproportionately large amount of philanthropic dollars.

As Figure 4 illustrates, in recent years, the Foundation 1000 set represents between three fifths and two thirds of grantmaking by all U.S. private and community foundations that Candid collects in a typical year.

**Creating the annual Foundation 1000 dataset.** Unlike the Transactions set, which includes nearly all transaction data available in real time, the Foundation 1000 set is a curated dataset that is systematically created by the Candid Insights team on an annual basis. To create the set, the team starts with a list of about 1,500 of the largest funders, including those that have appeared in past F1000 sets as well as those with the largest total giving (as reported on their latest 990 filing). Then, throughout the year, staff track Candid’s receipt and processing of contributions awarded by these 1,500 funders. In addition to reviewing transactions data to ensure that they represent a critical mass of a particular funder’s annual giving, Candid staff manually review the coding of transactions that are $250,000 or more to ensure that the largest are coded accurately. Once complete grant lists are received and reviewed for 1,000 funders, the dataset is considered complete and the
Foundation 1000 subset is frozen for the year (i.e., it is not dynamically updated and doesn’t reflect any day-to-day changes made in the underlying database). The annual Foundation 1000 sets go back to 2002.

Considerations and recommendations for using this set.
The purpose of the Foundation 1000 dataset is twofold. First, because donations by large foundations have such an outsized impact on overall donation trends, it is important to identify these foundations and review the data for accuracy. To this end, transactions in the Foundation 1000 subset undergo the most targeted cleaning and validation among all transactions in the broader Transactions set. Second, the Foundation 1000 dataset is designed to allow for analyzing giving trends over time. Because the set always includes 1,000 organizations whose sources have been reviewed for completeness, this set is not subject to the same volumetric variations as the Transactions

Notes: “All transactions” are limited to those awarded by US private and community foundations. Pledges are excluded. The value of known transactions may not always equal reported total giving.

8. While this level of meticulous manual review is ideal, it is impossible to conduct for all 100,000+ grantmakers Candid tracks; therefore, we prioritize those for which errors would most likely lead to erroneous conclusions about overall giving.
set. In other words, an increase in year over year giving within the Foundation 1000 set (whether for a specific cause, or in giving overall) is more likely to accurately reflect changes in giving among large foundations (rather than an increase in the number of foundations included in the dataset). Therefore, it is recommended that the Foundation 1000 set be used in research focusing on longitudinal giving trends at this time.

The Foundation 1000 data sets are organized by year (i.e., each set represents a year of grantmaking based on each included foundation's fiscal year end), so trends analysis requires the use of multiple sets. However, there are nuances and idiosyncrasies that should be considered when conducting research or drawing inferences from this dataset as well. Notably:

**The Foundation 1000 is not a matched set.** In other words, the Foundation 1000 does not consist of the exact same 1000 funders year over year. There are annual fluctuations in the funders included in the set based on whose data Candid is able to receive in a timely manner and how complete the source documents are. From one year’s set to the next, at least three-quarters of the funders included remain the same and special attention is taken to ensure that the very largest funders (based on total giving for a particular year) are included for that year since these represent the vast majority of giving within the set. Therefore, while the Foundation 1000 set provides a sense of sector-wide trends, it is not always possible to look at a specific grantmaker year over year within this set. Researchers planning analyses that require examining changes within the same grantmakers (e.g. case studies, repeated measures ANOVA/ regressions or autocorrelations) should first create a matched dataset.

**It can take years to close a Foundation 1000 set.** Gaining a complete record of grantmaking for 1,000 foundations is a lengthy process. A key constraint in developing the Foundation 1000 is the IRS’ schedule in making the forms 990 and 990-PF publicly available. More than half the transactions that make up any Foundation 1000 dataset are sourced from IRS filings

9. Italicized field names refer to specific data fields in Candid’s standard extract (a sample of which can be accessed here along with a data dictionary.) This is a standard format for transactions data requested through Candid’s custom data services (aka, the ‘Grants Plus’ extract) and it’s also how Candid’s Insights team typically interacts with transactions data.
(the remainder largely come from funders who report their contributions to Candid directly). Because of the slow nature of collecting IRS filings\textsuperscript{10}, and the need for complete grantmaking data, the Foundation 1000 set for a given year takes 2-3 years to create. (e.g. the Foundation 1000 set of grantmaking for the year 2018 was considered closed in October 2020). Since then, the time lag has increased such that as of early 2023, Candid does not have enough data to finalize a 2020 Foundation 1000 set. For more on these ongoing data sourcing issues, see this blog post.

**The Foundation 1000 is a frozen dataset.** As mentioned previously, unlike the rest of Candid’s data, the annual Foundation 1000 dataset is frozen (or locked) each year upon completion. However, subsequent updates do occur on rare occasions (e.g., when major changes are made to Candid’s taxonomy). Because the Foundation 1000 datasets update less frequently, they may offer a ‘time capsule’ view on foundations. This can be helpful if a researcher is trying to track down historical data (such a previous address or CEO of a given organization) that may otherwise be overwritten in Candid’s main datasets. However, this also means that there may occasionally be discrepancies between a given funder’s grantmaking record in the Transactions set and the Foundation 1000 set (e.g. if updates are made after a year’s dataset is closed).

The Foundation 1000 focuses primarily on grants—to the exclusion of some other transaction types. In addition to excluding transactions under $10,000, the set also excludes a number of transaction types distinct from cash grants. Specifically, the Foundation 1000 excludes costs associated with foundation-administered programs, employee-related scholarships, employee volunteer services, promissory notes, grants to individuals, loans to individuals, contracts, use of facilities, program-related investments (PRIs), mission-related investments (MRIs), pro bono services, and pledges. The set includes cash grants, matching grants, employee matching gifts, 

\textsuperscript{10} Typically, foundations report their grantmaking to the IRS at the end of their fiscal year, although it is common practice to request extensions (e.g. funds awarded in 2019 might be filed with the IRS as late as fall of 2020). It may take the IRS another year to review, process and release these filings. Therefore, Candid may not receive IRS data about 2019 transactions until late 2021 or early 2022.
officers and trustees’ discretionary grants, stock transfers and certificates, and most in-kind gifts (except for use of facilities). The reason for the exclusion of some transaction types in this dataset is to focus on those most commonly used in research and analysis. In this way, the Foundation 1000 is ‘cleaner’ for research purposes. However, it is worth noting that occasionally researchers may be interested in a type of transaction not covered by this dataset.

Side-by-side comparison of the two research datasets

For researchers trying to determine whether to use the Foundation 1000 set or the Transactions set in their research, the table below offers a side-by-side comparison of the details of each dataset.

<table>
<thead>
<tr>
<th></th>
<th>Transactions set</th>
<th>Foundation 1000 subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of transactions (per year of data)</td>
<td>2-4 million transactions</td>
<td>~150,000-200,000 transactions</td>
</tr>
<tr>
<td>Dollar value of transactions (per year of data)</td>
<td>$124.6 billion (2019)</td>
<td>$36 billion (2019)</td>
</tr>
<tr>
<td>No. of funders (per year of data)</td>
<td>&gt;100,000</td>
<td>1,000</td>
</tr>
<tr>
<td>Update frequency</td>
<td>Daily</td>
<td>Annually (with adjustments made periodically thereafter to reflect updates to underlying data)</td>
</tr>
<tr>
<td>Funder types included</td>
<td>Private foundations of all sizes (including company sponsored, operating, and independent); community foundations; corporate giving programs; government grantmakers; individuals; grantmaking public charities; and grantmaking NGOs.</td>
<td>A set of the largest private foundations (including company sponsored, operating, and independent) and community foundations</td>
</tr>
<tr>
<td></td>
<td>Transactions set</td>
<td>Foundations set</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Funder types excluded</strong></td>
<td>None (i.e., Includes all grantmaker types that Candid collects)</td>
<td>Corporate giving programs, government grantmakers, individuals, other grantmaking NGOs and nonprofits.</td>
</tr>
<tr>
<td><strong>Includes non-U.S. funders?</strong></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Transaction types included</strong></td>
<td>All (see full list) except grants to individuals</td>
<td>Cash grants, matching grants, employee matching gifts, officers and trustees discretionary grants, stock transfers and certificates, and most in-kind gifts (with the exception of ‘Use of facilities’)</td>
</tr>
<tr>
<td><strong>Transaction types excluded</strong></td>
<td>Grants to individuals</td>
<td>Foundation-administered program awards, employee-related scholarships, employee volunteer services, promissory notes, grants to individuals, loans to individuals, contracts, use of facilities, program-related investments (PRIs), mission-related investments (MRIs), pro bono services, and pledges</td>
</tr>
<tr>
<td><strong>Further exclusions</strong></td>
<td>Excludes transactions between $1 and $9</td>
<td>Excludes transactions under $10,000</td>
</tr>
</tbody>
</table>
Part IV: Key variables and data fields

Field names for key variables

Throughout this section, italicized names refer to specific data fields in Candid’s standard “Grants Plus extract” (a sample of which can be accessed here along with a data dictionary.) This is the standard format for transactions data requested through Candid’s custom data services and it’s also how Candid’s Insights team typically works with transactions data. Note that the field names mentioned in this section applies to both the Transactions set and the Foundation 1000 subset.

Candid’s transaction datasets include hundreds of data fields and variables. Which data fields to include and how to use them largely depends on the research questions being addressed. However, there are a few key variables that are particularly important to understand when cleaning and creating data files, running analyses and/or interpreting results of Candid data. This section includes a brief description of these variables and best practices for filtering, analyzing, and interpreting them.

Transaction type

Cash grants are the most common type of transaction represented in Candid’s data, accounting for more than 99 percent of all transactions. Candid tracks sixteen additional types in the “transaction type” facet of the PCS (Philanthropy Classification System), including program- or mission-related investments (PRIs and MRIs, respectively), in-kind gifts, pro bono services, and pledges. Some transaction types—like MRIs—are only identifiable when that information has been shared directly with Candid and are very rare in the database.

Researchers may want to exclude specific transaction types to avoid combining very different types of gifts. In the data file, the field grant_transaction_tran lists the transaction type. The full list of transaction types, along with the associated alphanumeric code for each are listed on this webpage.

As noted in detail in Part III, the Transactions and the Foundation 1000 datasets differ in which transaction types are included. The Transactions set includes all transaction types except for grants to individuals, while the Foundation 1000 set only includes cash grants, matching grants, employee matching gifts, officers and trustees’ discretionary grants, stock transfers and certificates, and most in-kind gifts (except for use of facilities.)
When the Candid Insights team conducts research on transactions using the Transactions set, the typical approach is to remove PRIs, MRIs, foundation-administered programs, and pledges. This is done to avoid combining very different types of transactions in a single analysis, though exceptions are sometimes made. Researchers should specify which transactions they’ve elected to include as there is no clear standard practice—though it’s always best practice to analyze pledges separately from other transactions types.

**Research recommendation: Separate pledges**

The Insights team’s standard practice and strong recommendation is to analyze pledges separately from grants and other transaction types. Pledges reflect a funder’s public promise to dedicate a certain amount of resources to address a crisis or issue, often over a set number of years. In contrast, grants reflect funding formally committed or paid to a specific organization.

As noted in the timeline on page 18, Candid increased efforts to capture pledges in 2020. Pledges can be an important early signal of funding trends and therefore provide insight into future giving. However, pledges should not be interpreted as money in the bank, as there is no contractual obligation to fulfill these promises. Researchers should also be aware that pledges always remain pledges in Candid’s database. In other words, pledges do not get removed from the system if/when the pledge-fulfilling-grant(s) get recorded. This is because Candid’s coders have found that attempting to do so is virtually impossible. Pledges tend to be broad and vague (e.g. “Funder X pledges 5 million dollars to support equity over the next decade”), while cash grants tend to be narrow and specific (e.g. “Funder X awards 50k to Nonprofit Y in 2022 to conduct a research study on gender parity”). This mismatch makes systematic deduplication efforts very difficult to do reliably and accurately. At the same time, summing pledges and other transaction types risks double counting, as some grants may indeed be pledges fulfilled. (Note: pledges have a grant_transaction classification of “TV00”).

**Paid and authorized transactions**

After filtering or cleaning data to include the desired
transactions, researchers should be aware of another transaction-related variable: paid versus authorized. By default, Candid’s data sets combine transactions reported as “paid”—that is, money that has been given to the recipient—and those reported as “authorized”—meaning that a funder has issued a gift to be awarded over a set period but has not necessarily paid it in full.

These two transaction types can largely be traced back to process differences in respective data sources. Specifically, the IRS requires that organizations report their paid amounts for a particular fiscal year, so transactions sourced via tax forms will reflect the amount paid in the corresponding year. However, funders who report their transactions to Candid directly often report authorized amounts. Note that authorized grants differ significantly from pledges in that, unlike pledges, authorized grants are not broad promises to support a cause, but specific monetary amounts already awarded to a given organization. They are essentially checks signed but not yet cashed. For example, a funder might report $1.2 million awarded for a 3-year grant though they may have only paid out $400,000 of the grant to date. Multi-year authorized transactions are only counted once—in the full amount and in the year of authorization—to ensure that funding is not double counted.

Transactions are flagged as either “paid” or “authorized” under the data field payment_status. There is also a field for duration as measured in months, to indicate the length of the grant/transaction period. However, Candid only has “duration” information when funders supply it (i.e., it is not available via IRS filings, and not always self-reported by funders). Moreover, duration is most useful for authorized transactions since a paid transaction may be part of an unspecified multi-year commitment which had previously been authorized but not reported as such.

Candid asks data sharing partners to be consistent year over year in whether they share paid or authorized transactions. However, Candid does not enforce this policy. In rare cases, transactions reported as authorized in one year may be duplicated in subsequent years if the funder’s reporting method changes to include only paid transactions.
Research recommendation: Acknowledge whether transactions are paid and/or authorized

Because funders may report their transactions as either paid or authorized, the Transactions and Foundation 1000 sets include both. For analysis, Candid’s Insights team allows a combination of both paid and authorized transactions and recommends that researchers using Candid data use this approach as well. Limiting the data to one or the other would exclude a considerable number of funders and significantly reduce the size of the data set. Regardless of approach, researchers should note in their methodology whether they include both paid and authorized transactions in their analysis. Moreover, as always, a best practice is to manually review extremely large grants in longitudinal analyses to ensure there is no double counting.

Funder type

Candid’s data represents transactions awarded by a number of different funder types, including private foundations (subdivided into independent, company-sponsored, and operating foundations), community foundations, individuals, federal funders, grantmaking public charities, and corporate giving programs—among others. (In the standard view of the transaction data in Excel, the field listing the grantmaker type is gm_organization_tran). As can be seen in Figure 5 below, the majority of funders in Candid’s database are classified as independent foundations or public charities.

Transactions from some funder types are more comprehensive than others. Specifically, Candid will eventually have a fairly comprehensive database of funding for organizations that file IRS 990 forms (once all IRS data has been collected and processed): community foundations, independent foundations, public charities, company-sponsored foundations, and operating foundations. However, the remaining funder types are not required to disclose grantmaking to the IRS, therefore, information is limited to what they choose to share publicly (e.g. through press releases, websites, etc.). The “individual donor” funder type in particular should not be assumed to be comprehensive. While individual donors provide a great deal of contributions to nonprofits, Candid selectively captures and tracks public announcements of large donations by high-
Independent foundations and public charities are the two most common grantmaker types, representing 90% of funders among Candid’s transactions data.

For Candid’s definition of each grantmaker type see taxonomy. candid.org/organization-type.

net-worth individuals (as information about most individual donations are not public information).

The Foundation 1000 dataset is limited to two specific types of funders: U.S. private (i.e., Independent, operating, and corporate) and U.S. community foundations, whereas the Transactions set also includes all other funder types (i.e., Federal funders, individuals, grantmaking public charities, and corporate giving programs).

Research recommendation: remove certain funder types if interested in private institutional funding

A common use-case for analyzing Candid’s transaction datasets is to better understand private institutional philanthropy (i.e. funding by foundations and similar institutions). In such cases, researchers leveraging the Transactions set should remove high net-worth individuals as a funder type to focus on institutions, as well as federal funders and public charities to isolate private
giving. This is particularly important as occasional large donations from individual givers (e.g. MacKenzie Scott) or large federal grants can skew a data set. In particular, while federal funders make up just .1% of funders in Candid’s databases (see Figure 5), they may account for 80-90% of the overall value of transactions in some years. For this reason, Candid’s Insights team rarely includes this funding in analysis.

**Research recommendation: remove certain funder types for trends analysis and/or to isolate a known universe of funding.**

Two other common use-cases for analyzing transactions are to: 1) examine trends over time; or 2) summarize a specific universe of funding (e.g. “total dollars awarded in 2018 by community foundations”). In both these cases, researchers will most likely want to isolate transactions to funder types for which there is comprehensive data. For any given year, Candid will eventually (i.e., 2-3 years after the calendar year has passed) have funding data from nearly all U.S. private and community foundations, while other funder types—like individuals, corporate giving programs, and non-U.S. donors—may only make their data available selectively. Researchers using the Transactions set to get a sense of all possible grantmaking in a particular space shouldn’t necessarily limit the set based on funder type. But for those looking at year-over-year changes or calculating total giving by funder type, it’s recommended to look at funding by U.S. independent, operating, company-sponsored, community foundations and grantmaking public charities only gm_organization_tran of “Independent foundations”, “Operating foundations”, “Company-sponsored foundations”, “Community foundations”, and “Public charities”, respectively.

**Internal grants and pass-through funding**

As mentioned on page 19, Candid’s data may include transactions awarded to other funders for re-granting purposes. An example is when a private funder contributes to a “regranter” like a community foundation, who may then in turn grant money to a nonprofit such as a direct services public charity (see Figure 6 below). To check whether regranted funds may be included in a specific subset of the data, researchers should review the
Double counting of funding dollars may occur when there is an organization in the dataset that is receiving and awarding philanthropic transactions. There are two types of organizations involved: the initial funder and the regrantee. The initial funder, such as a private foundation, awards funds to a community foundation, which in turn awards funds to a direct services public charity. This can lead to double counting of funding dollars, which is illustrated in Figure 6.

To address this issue, the research recommendation is to remove internal grants. The Candid Insights team updates their accounts for regranted funds by removing transactions awarded to regranting organizations while keeping any that the regrantee themselves has awarded to other organizations. Researchers can simply remove or filter out records with a value of “1” under the internal transaction field. This approach is only relevant where a particular dataset includes both transactions being awarded to and awarded from the same regrantee. If a dataset only includes transactions awarded to or from a known regrantee (but not both) then deduplication is not necessary.

It’s also important to note that this methodology serves as an estimation for regrants because Candid’s data doesn’t provide sufficient detail to guarantee that the funding from an initial funder was in fact regranted. In Figure 6 for example, it is possible that the private foundation awarded the money to the community foundation for their own capacity building efforts, and that those funds were not regranted.
Moreover, when producing a list of funders along with funding amounts, researchers may want to include internal (regranted) grants in order to represent the full extent of a single organization’s giving.

One final note is that, while Candid’s PCS has a code for “Regranting” (UB0000) under the support strategy facet, it is not a replacement or supplement for this method. We find that this code is very rarely applied to transactions because it’s difficult to independently identify regranted dollars or funds intended to be regranted without any explicit indication from the funders sharing data.

**Fiscal year**

The year assigned to transactions in Candid’s database in field `fiscal_year` reflects the awarding organization’s fiscal year end during which the transaction was paid or authorized. For funders whose fiscal year does not align with the calendar year, transactions awarded in one calendar year may be associated with the next year in Candid’s data. For example, if a funder’s fiscal year spans from July 2022 to June 2023, a grant awarded in December 2022 would be categorized as a 2023 transaction because the funder’s fiscal year ends in 2023.

It’s also worth noting that Candid rarely receives information about the specific day or month that a transaction was awarded; therefore, time series analyses using Candid data should focus on tracking changes in funding by year.

**Geographic location**

Researchers may also want to consider the geographic area associated with a given transaction. There are three types of geographic area data that may be represented: 1) grantmaker location, 2) recipient location, or 3) geographic area served by the transaction (for example, a grantmaker in New York (1) could give a nonprofit based in San Francisco (2) a grant to support women leaders in India (3)).

Information about the funder’s location is in the fields `gm_city`, `gm_state`, and `gm_country`. Candid’s transactions data is most
comprehensive when looking at grantmakers based in the U.S. By definition, the Foundation 1000 set limits transactions to those awarded by (the largest) U.S. funders (where gm_country = “United States”). The Transactions set, in contrast, includes transactions from funders outside of the United States. Therefore, researchers interested in using the Transactions set to conduct research on US funders should identify and remove non-US-based grantmakers.

Transactions data from funders based outside the U.S. is sourced directly from funders, from publicly available sources—like 360Giving in the U.K., or the news. This selective sourcing method means that Candid does not currently have comprehensive data from non-U.S. grantmakers. Because of this, Candid’s data on grantmaking by international funders shouldn’t be used as-is to derive year-over-year trends or to draw conclusions about overall funding by all grantmakers in a particular country or region.

Candid also tracks recipient location (recip_city, recip_state, or recip_country) as well as a transaction’s geographic area served—where that information has been provided. Candid utilizes the GeoNames database to identify and apply eligible place names to the data. Specific geographic areas can be viewed in Candid’s tools that include transactions-level data, like Foundation Maps. For ease of analysis, in the standard view of the transaction data (in Excel) this information is rolled up to the international region, country, and/or U.S. state-level, where applicable, under the fields intl_regions, intl_countries_tran, and dom_states_tran, respectively.

It should be noted that US domestic funders can fund non-U.S.-based recipients and/or international areas served. Therefore, non-U.S. recipients and areas served may still be represented in the Foundation 1000 set, even though the funders are limited to those located in the U.S.

Some research questions center the general geographic focus of a particular transaction. In such cases, the Candid Insights team’s standard practice is to first consider the transaction’s geographic area served —either at the country or U.S. state-level. If this information isn’t available, the recipient location is considered as a secondary way to categorize geographic focus area.
Source type

As outlined in Part 1, Candid’s transaction data relies on multiple data sources, and there is an extensive hierarchy and de-duplication methodology deployed to increase the chances that the data included in Candid’s data draws from the most recent, reliable, and complete sources. The data source for a particular transaction appears in the data set under the field source. The categories in this field offer details about the precise sources included—see this mapping for interpretation of these specific values.