NCCS

Vision
Under the hood
The “Platform”
Launch plan

A tale of two NCCS's

- What has changed?
  - Design of an integrated system (harmonized, standardized)
  - Modern data infrastructure, built for scale
  - Open-source workflow, built for sustainability

- Improved user experience

- Expanded data catalog

- New website ("data platform")

- A changing landscape: competitors, collaborations, opportunities and threats

https://lecy.github.io/nccs/

NOAA

THE CLIMATE CORPORATION

NOAA

THE CLIMATE CORPORATION

Climate Corporation hopes to transform the weather business. By pairing the tools of agronomics and climatology with **the power of "big data,"** the company has managed to interpret weather information more fully than any other organization.

Company scientists process **fifty terabytes of weather information every day**, roughly the equivalent of ten million songs, including eight years' worth of soil, moisture, and precipitation records for each of the twenty-nine million farm fields in the U.S.

The algorithm divides the country into nearly **half a million plots, then generates ten thousand daily weather scenarios for each of them**.

This information is used to create individualized insurance policies for corn, soybean, and wheat farmers covering major perils like drought, heat stress, and the risk of an early freeze. **No claims, forms, adjusters, or negotiations are required**.

NCCS IMPACT

- Data has tremendous value
  - We unlock that value when data is clean, well-documented, timely
  - Free data eliminates rent-seeking, forces innovation

NCCS IMPACT

- Data has tremendous **VALUE**
  - We unlock that value when data is clean, well-documented, timely
  - Free data eliminates rent-seeking, forces innovation
- We influence HOW people use data – **DISSEMINATION**
  - The new website is built for community and dissemination
- We can create **TOOLS** that provide insight
  - Climate Corporation example – what data driven management looks like
- We can help the sector develop **CAPACITY** to use data
  - Help nonprofits and foundations scale on the collection side
  - In a position to help them build their own capacity

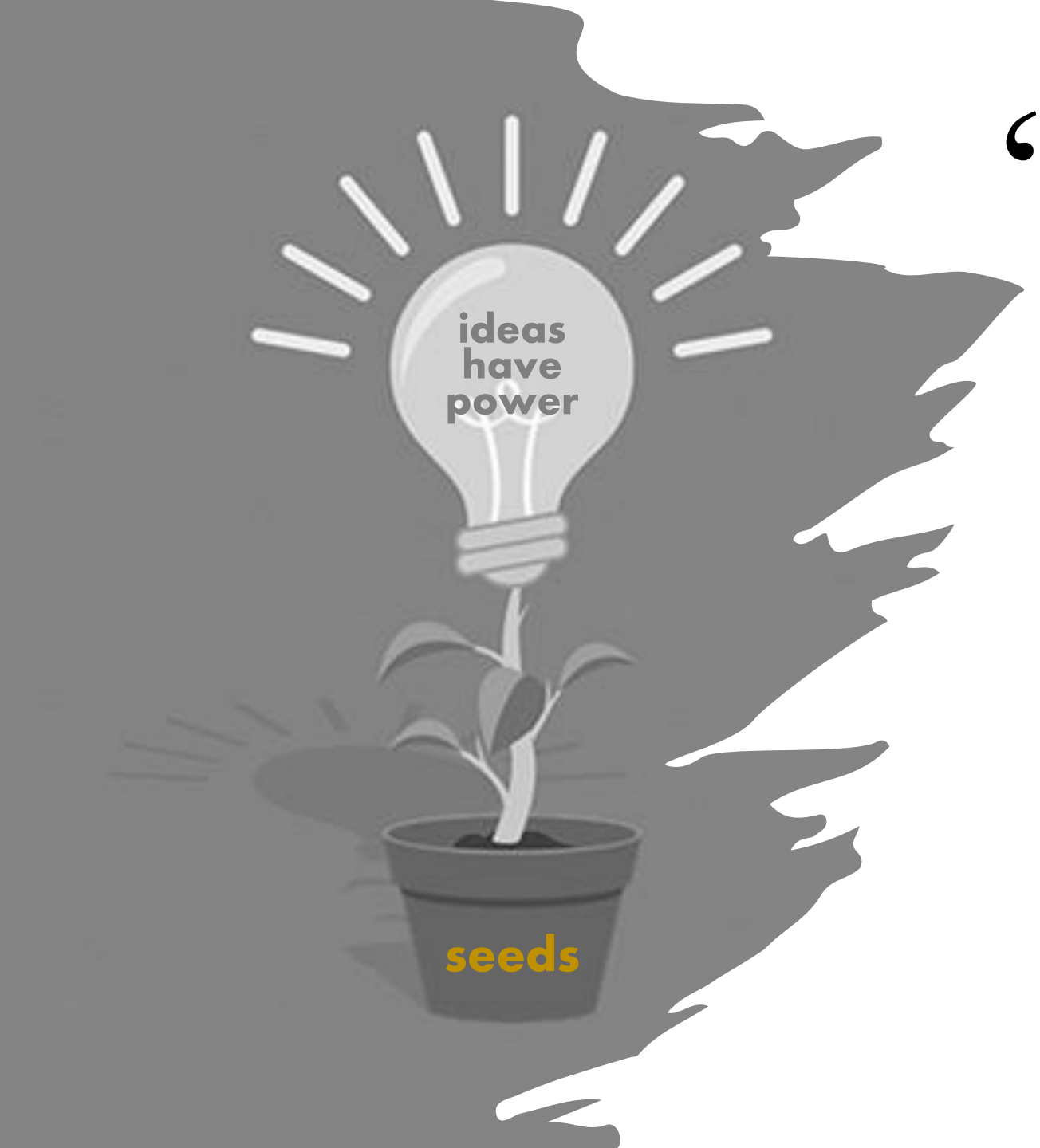
NCCS IMPACT

AUDIENCES

- Da
  - **PHASE 1: Pay the Tech Debt**
  - ented, timely
  - Free data eliminates rent-seeking, forces innovation
- We
  - **PHASE 2: Scale up**
  - dissemination
  - **Build the Community**
  - **Define products**
- We
  - ght
  - driven management looks li e
- We can help the sector develop capacity to use data
  - Help nonprofits and foundations scale on the collection side
  - In a position to help them build their own capacity

Wonks

Community partners



“

Practical men who believe themselves to be quite exempt from any intellectual influence, are usually the slaves of some defunct economist.

Even madmen in authority who hear voices in the air are distilling their frenzy from some academic scribbler of a few years back.

”

— John Maynard Keynes

user experience

time from idea to insight:

the 10x goal

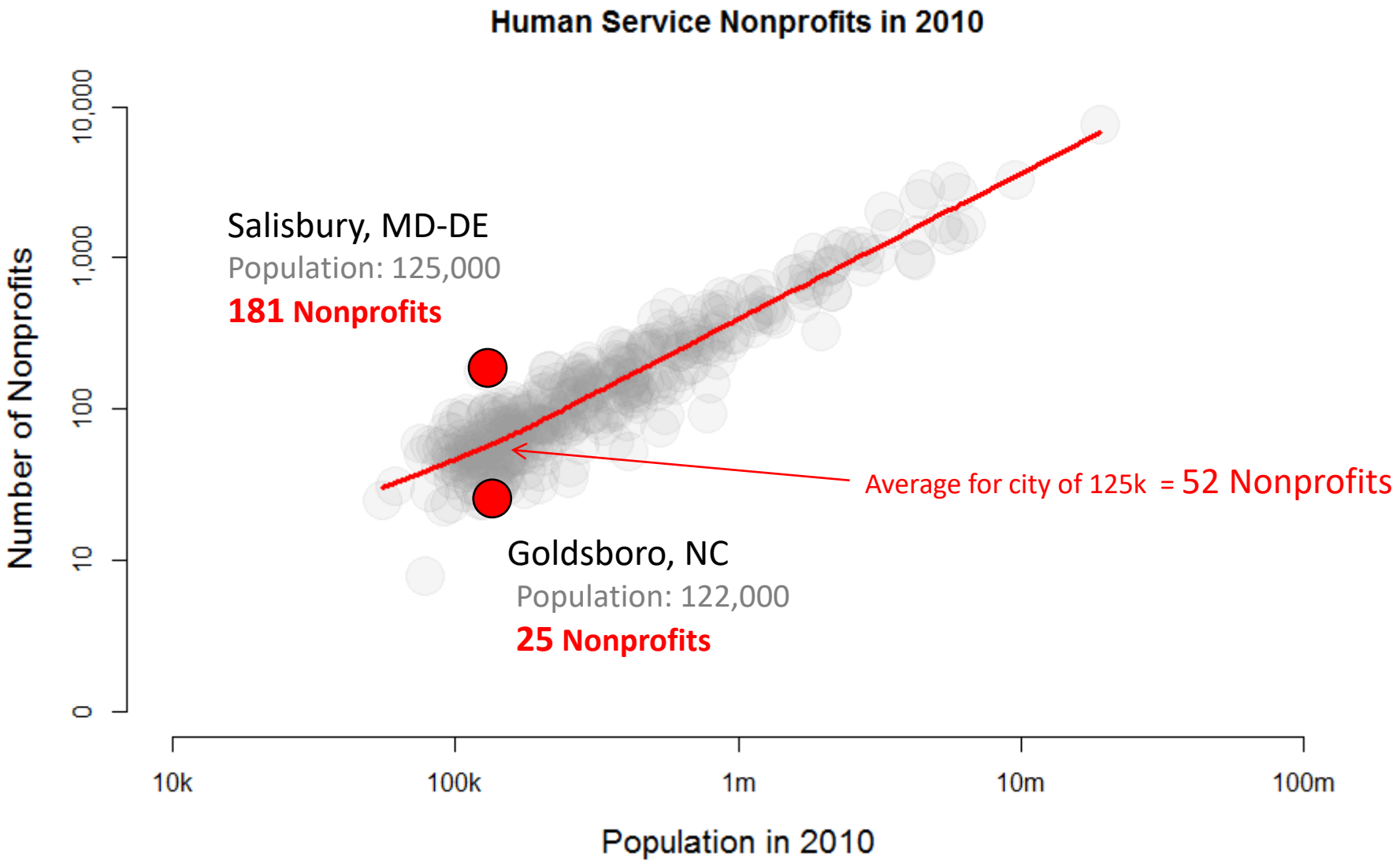
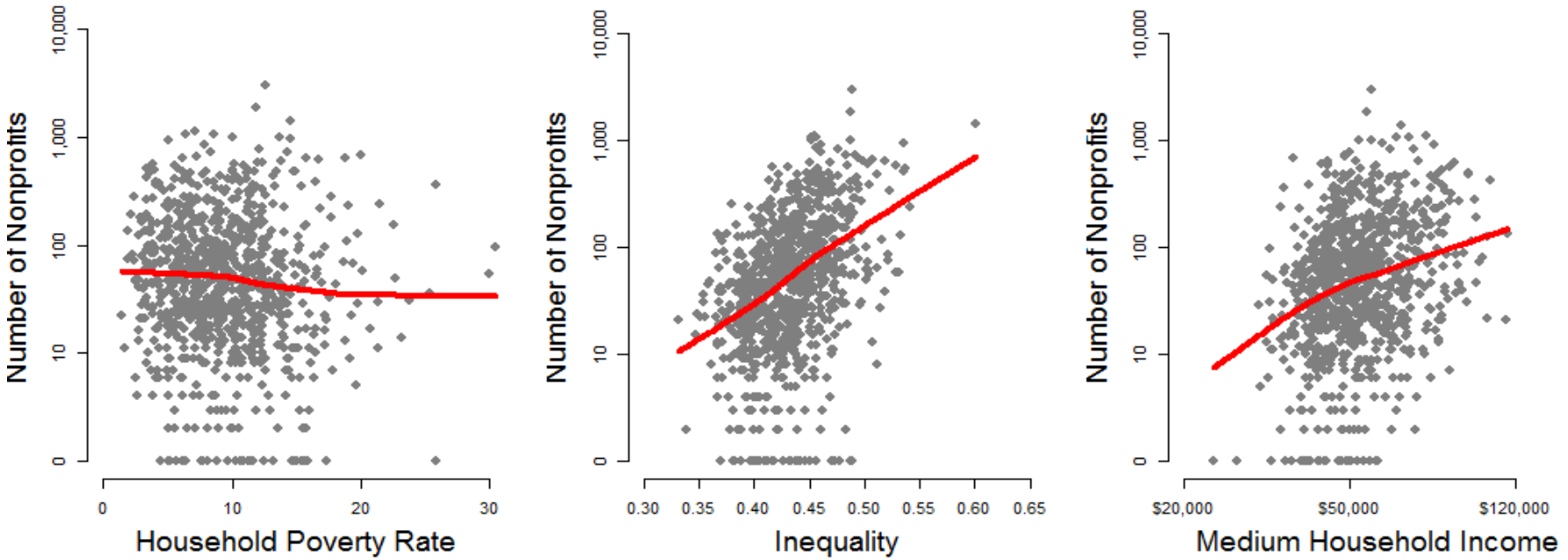
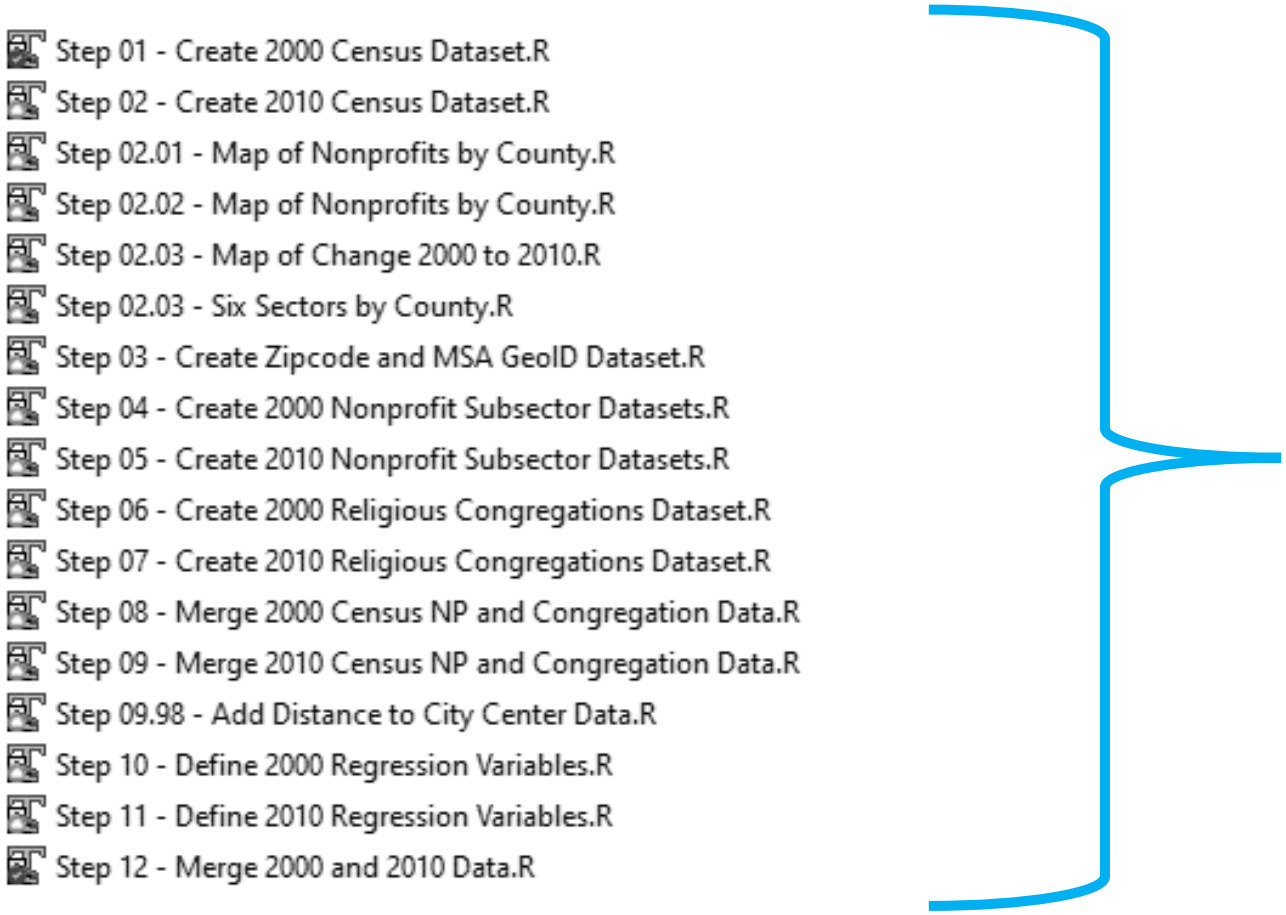

Human Service Nonprofits in 2010
Number of Nonprofits
0
10
100
1,000
10,000
10k
100k
1m
10m
100m
Population in 2010
Salisbury, MD-DE
Population: 125,000
181 Nonprofits
Goldsboro, NC
Population: 122,000
25 Nonprofits
Average for city of 125k = 52 Nonprofits

Table 1: Nonprofit density and population measures in 2000 and from 2000-2010

	Min	Q25	Median	Mean	Q75	Max
Nonprofits Per 100k Residents in 2000						
All Subsectors	0	27.50	71.77	78.32	215.50	575.47
Change in Nonprofits Per 100k Residents 2000-2010						
All Subsectors	-20.71	17.69	26.96	29.15	37.75	171.02



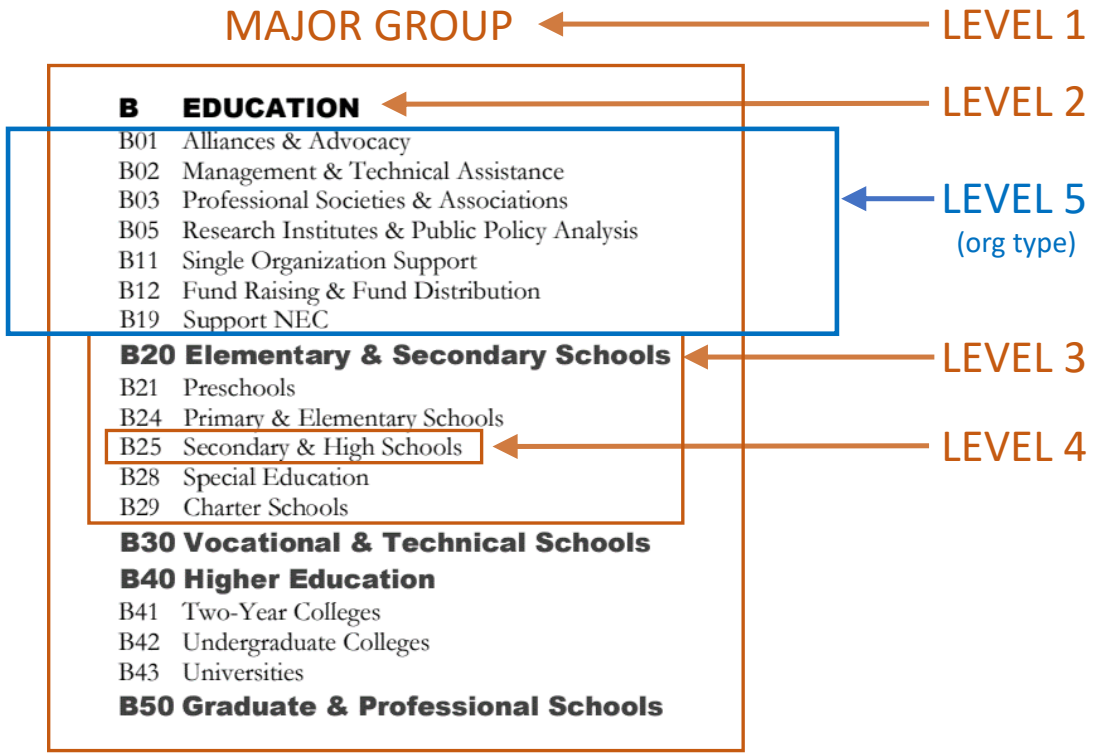


13,400
lines of code

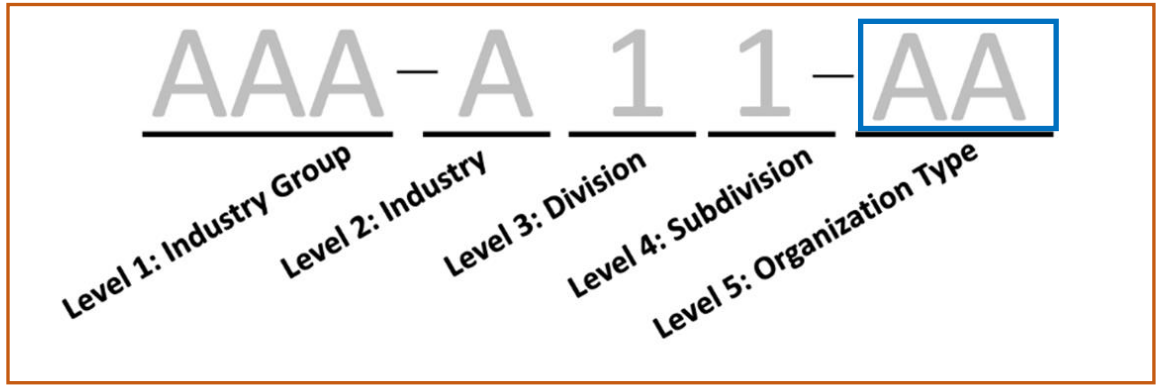
```
get_data(

    core,
    years=2010:2020,
    geo.region="southeast",
    ntee.lev1="human services" ) %>%

append_census( level="tract" )
```

```
get_table(

    data=       core,
    years=      2020,
    geo=        “metro”,
    ntee=       “level1”,
    summarize=  “expenses”,
    stat=       “mean”          )
```

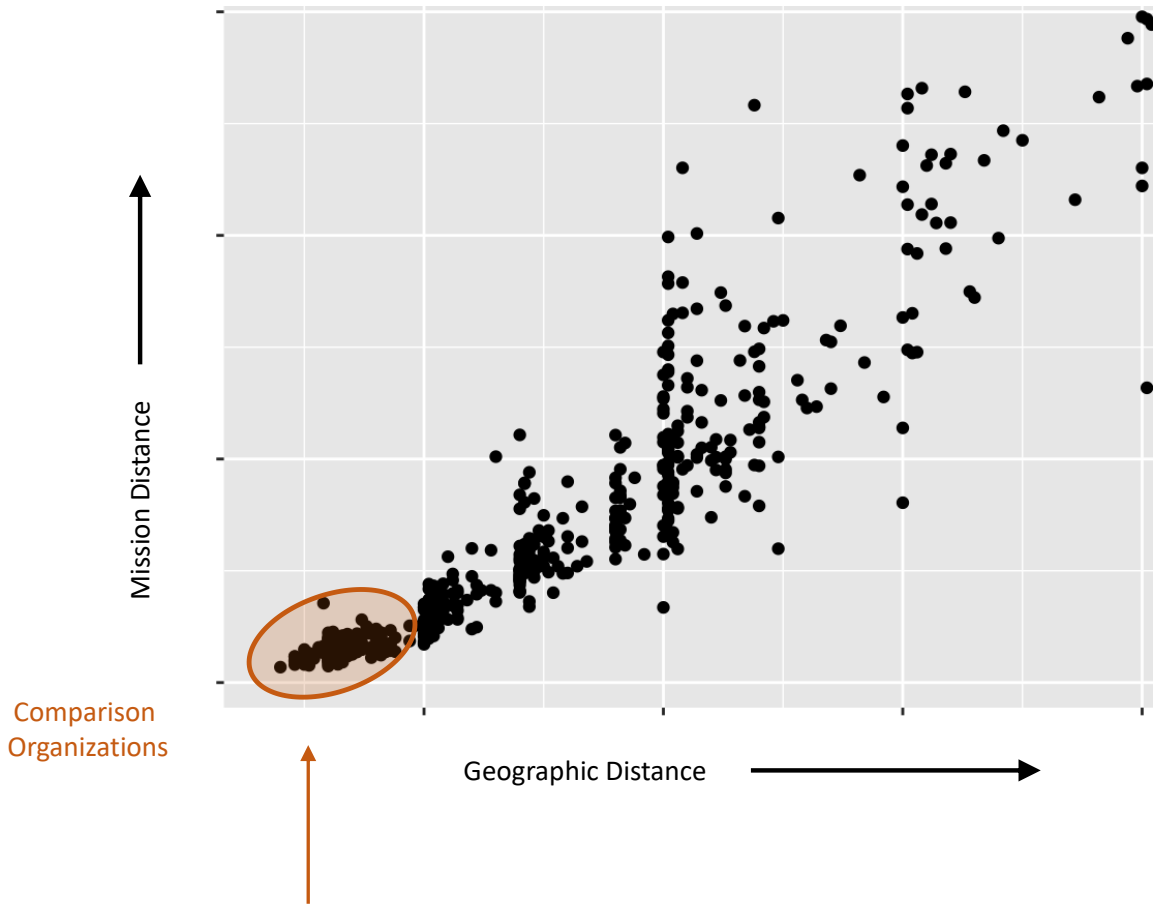
OLD NTEE



NEW NTEE



old	new	Level 1	Level 2	Level 3	Level 4	Level 5
C32	ENV-C32-RG	**ENV**: Environment and Animals	**ENV-C**: Environment	**ENV-C3**: Natural Resources Conservation & Protection	**ENV-C32**: Water Resources, Wetlands Conservation & Management	**ENV-C32-RG**: Regular Nonprofit
I31	HMS-I31-RG	**HMS**: Human Services	**HMS-I**: Crime & Legal-Related	**HMS-I3**: Correctional Facilities	**HMS-I31**: Half-Way Houses for Offenders & Ex-Offenders	**HMS-I31-RG**: Regular Nonprofit
Q36	IFA-Q36-RG	**IFA**: International, Foreign Affairs	**IFA-Q**: International, Foreign Affairs & National Security	**IFA-Q3**: International Development	**IFA-Q36**: International Science & Technology Development	**IFA-Q36-RG**: Regular Nonprofit
A0534	ART-A34-RP	**ART**: Arts, Culture & Humanities	**ART-A**: Arts, Culture & Humanities	**ART-A3**: Research Institute and/or Public Policy Analysis in arts, culture & humanities.	**ART-A34**: Radios	**ART-A34-RP**: Research Institute and/or Public Policy Analysis
X1121	REL-X21-MS	**REL**: Religion Related	**REL-X**: Religion-Related	**REL-X2**: Christianity	**REL-X21**: Protestant	**REL-X21-MS**: Monetary Support - Single Organization



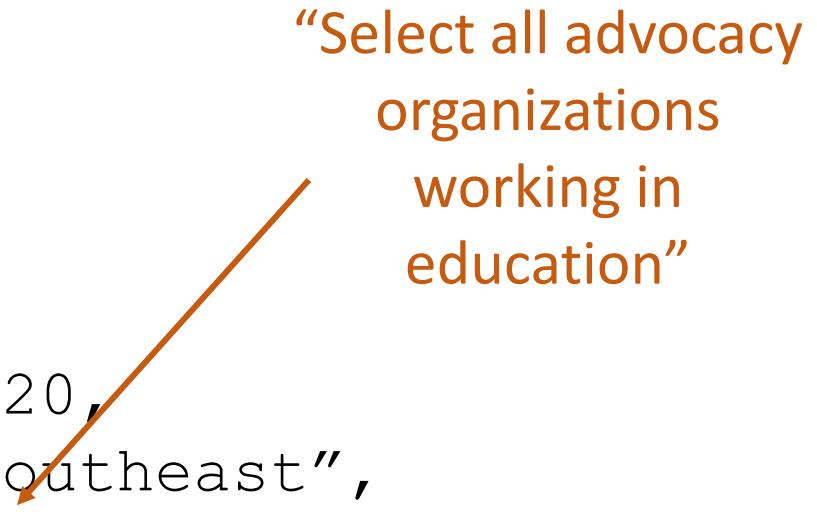
$$\sum_{i}^{4} \alpha_i \delta_i$$ DISTANCE

AAA-A 1 1-AA

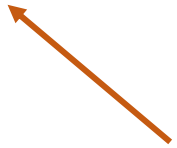
Data representation matters!

XXXV R V	35R5
VI | CCXV	6 | 215
−CLXXX	−180
XXXV	35
−XXX	−30
V	5

```
get_data(

    core,
    years=2010:2020,
    geo.region="southeast",
    ntee="B2x-AA" ) %>%

append_census( level="tract" )
```

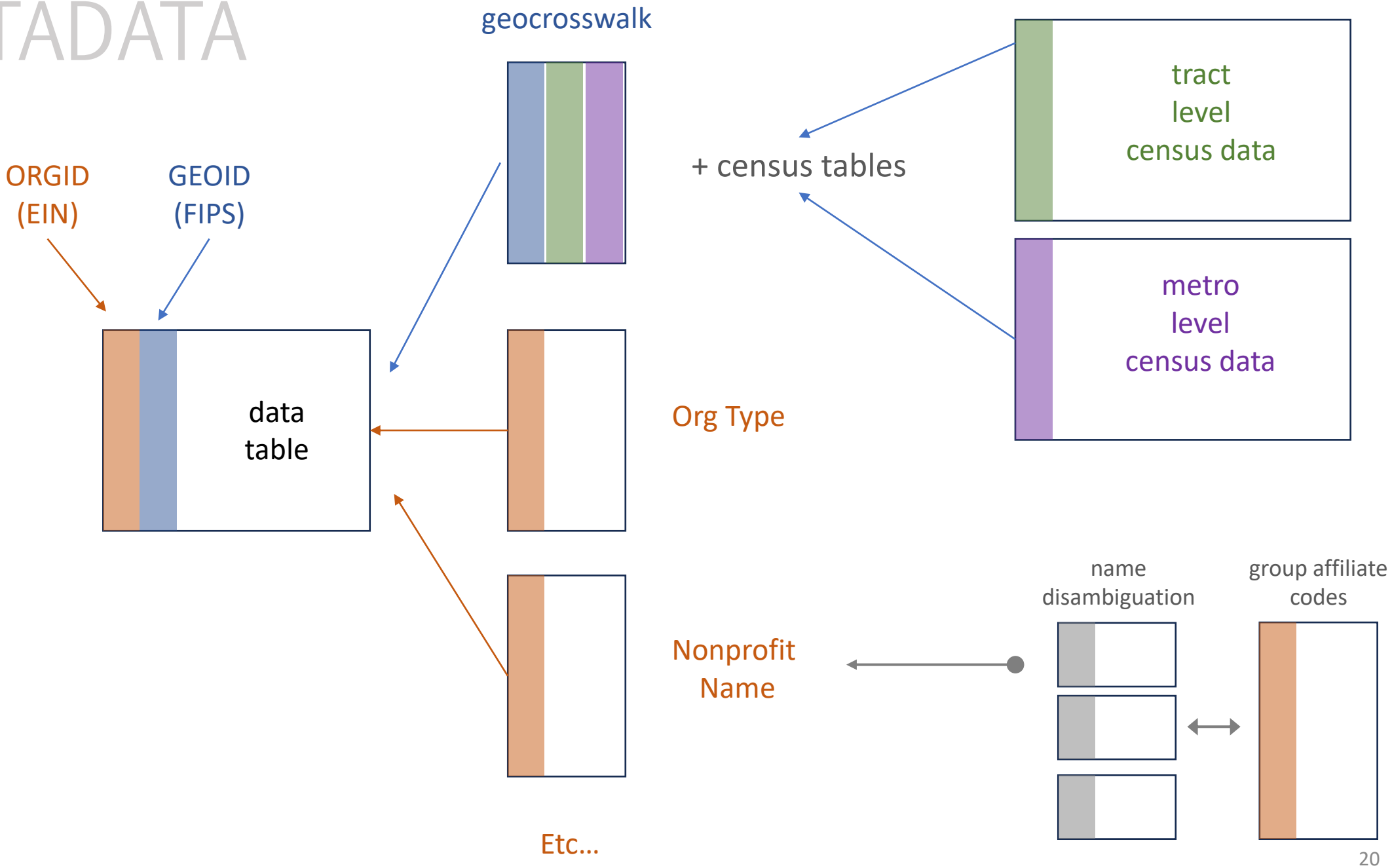


```
get_data(

    core,
    years=2010:2020,
    geo.region="southeast",
    ntee="B2x-AA" ) %>%

append_census( level="tract" )
```



Give me census data
for 2010-2020
at the tract level

METADATA


geocrosswalk
ORGID
(EIN)
GEOID
(FIPS)
data
table
+ census tables
tract
level
census data
metro
level
census data
Org Type
Nonprofit
Name
name
disambiguation
group affiliate
codes
Etc…

a tale of two NCCS's

What has changed?

	OLD	NEW
Raw Data	Hand entry of data	Automation / warehouse
Data Processing	Contractors	NCCS data pipeline
Data Cleaning	Interns	datagood workflow
Data Harmonization	Researchers	nccsdata packages
Data Analysis	Team of 3	templates / dashboards

NCCS Open-Source Library

Admin Tasks

- nccs-admin
- nccs website
- lambda tasks: automated data collectors

Data

- metadata standards: atomize and standardize nonprofit metadata
- master concordance file: crosswalk and data dictionary for nccs datasets
- nccstools: R package with basic data processing helpful across all nonprofit datasets
- nccs-data-package R package with core data query functionality
- nccs core data series
- **efile data series:**
  - data retrieval
  - build package
  - foundations dev
- **compensation data series:**
  - peopleparser
  - titleclassifier
- IRS Pub 78 Exempt Orgs Database
- IRS Nonprofit Business Master File
- IRS 990N Postcard Filers
- IRS Tax Exempt Revocations
- IRS 1023-EZ New Nonprofit Metadata
- 527 Political Organization Disclosures
- Historic Statistics of Income Nonprofit Microdata Samples
- Historic 990 data from 1982-1994 (pre-NCCS Core)

Tools

- **census data:**
  - nccs-geo: crosswalk for geographic conversions
  - geocrosswalk: package for building census data panels
- peopleparser 990 part vii name standardization and gender coding
- titleclassifier 990 part vii title standardization
- fiscal R package for creating a standard set of nonprofit fiscal health metrics
- compensator R package for automated compensation appraisal of NP executives
- npcompete tool to create standardized metrics of market competitiveness
- webscraper: package to assist in automation of web scraping, returns tidy table of text
- classifying nonprofit missions: demo project of how to build a mission classifier
- ProPublica API: task automation through ProPublica API

OLD

OLD STACK

- SQL + CSV
- Python + R
- AWS
- (Great Expectations)
- Drupal (website)
- PHP (data dictionary)

NEW

NEW STACK

- CSV
- R
- GitHub Pages

OLD

NEW

Raw Data

Hand entry of

Data Processin

not reproducible!

Data Cleaning

hard to sustain

Data Harmonizat

Data Analysis

Team of 3

OPEN SCIENCE

templates / dashboards

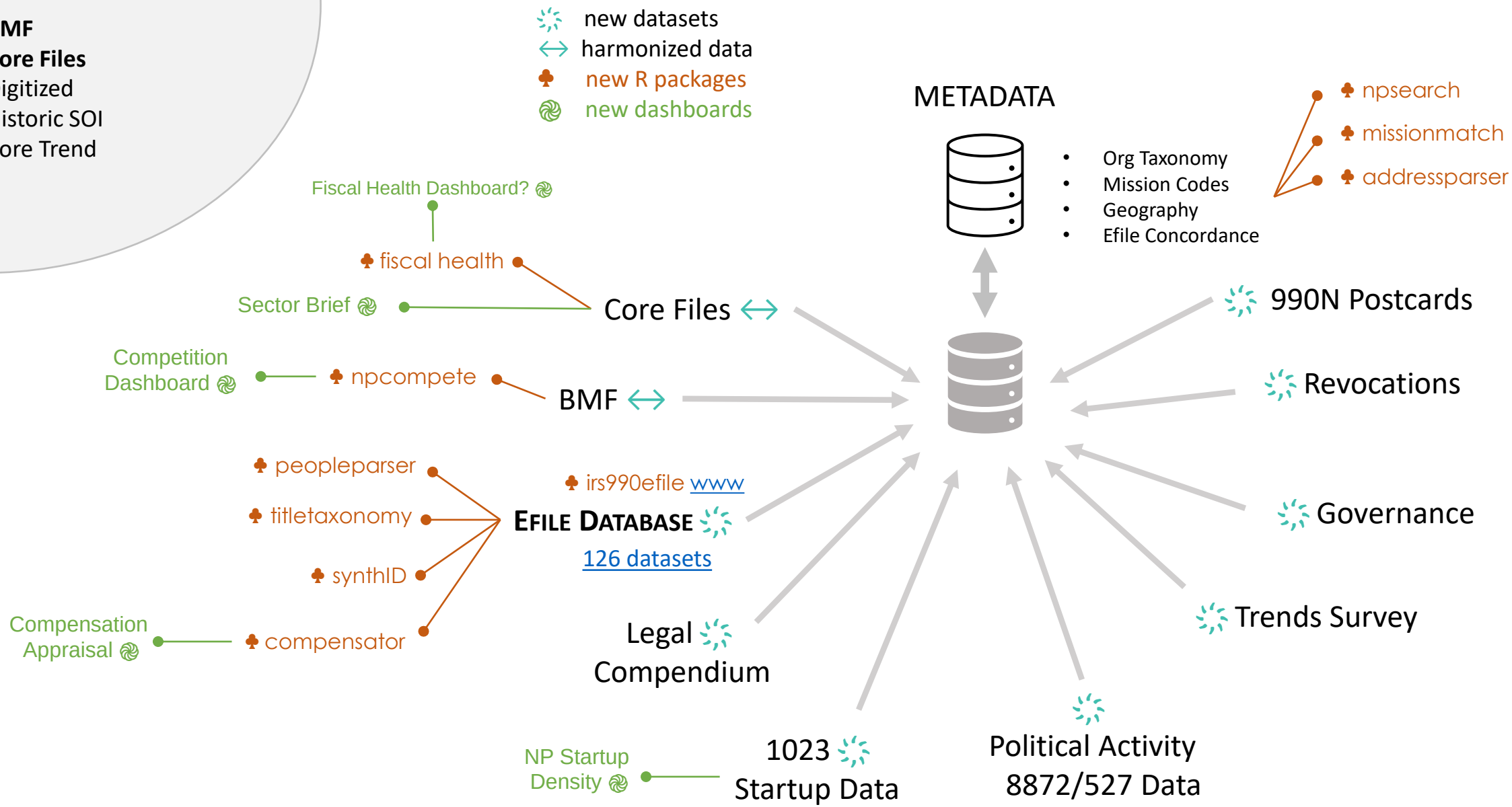
data catalog

2 old
10 new
efile data

New NCCS Data Platform

Current NCCS Catalog

- **BMF**
- **Core Files**
- Digitized
- Historic SOI
- Core Trend



NCCS website

data platform
tools
stories

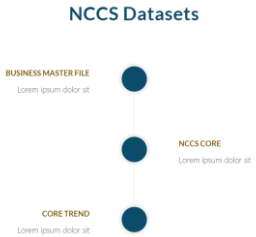
Welcome to the National Center for Charitable Statistics (NCCS) open data platform

Consectetur atque culpa ad quidem ex? Deleniti atque error necessitatibus cumque quisquam. Fugit necessitatibus commodi eveniet.

HERO CTA LINK

New Digital Infrastructure

New NCCS Website (Jekyll)



Blog / Vignette Templates

Each Dataset

data processing → documen-tation → API → data package → dashboard

- documen-tation: Data Validation Package
- API: plumbeR APIs
- data package: pkgdown
- dashboard: shiny template

Dataset Sample Full Fields

DOWNLOAD DATA

File size: 246kb

R PACKAGE

DATA HISTORY

How to Use this Data

USER GUIDE

GET HELP

Copy Citation

Choi. Y & Lee, Y.. 2023. "Dataset Sample Full Fields." Ednel: A large – scale hierarchical dataset in education. https://doi.org/10.5555/12345678. https://urbaninstitute.github.io/nccs/datasets/sample-dataset/

Overview

Sit delectus inventore architecto optio tempora Labore error cumque officiis culpa nesciunt Necessitatibus facilis excepturi eaque eveniet consectetur? Maiores numquam a natus aliquid nulla, voluptatum Error officiis amet eos porro

Sit harum necessitatibus excepturi suscipit tenetur? Animi iusto magnam illo porro aspernatur. Perspiciatis reprehenderit impedit quisquam aliquam ex, dicta. Perferendis similique perferendis fuga ipsum optio? Aspernatur rem distinctio odit explicabo

Lorem consequatur qui rem quibusdam repellendus in voluptatum Explicabo excepturi culpa laudantium minima exercitationem. Animi accusantium temporibus necessitatibus delectus nemo. Minus explicabo praesentium ratione debitis consectetur atque Itaque velit omnis.

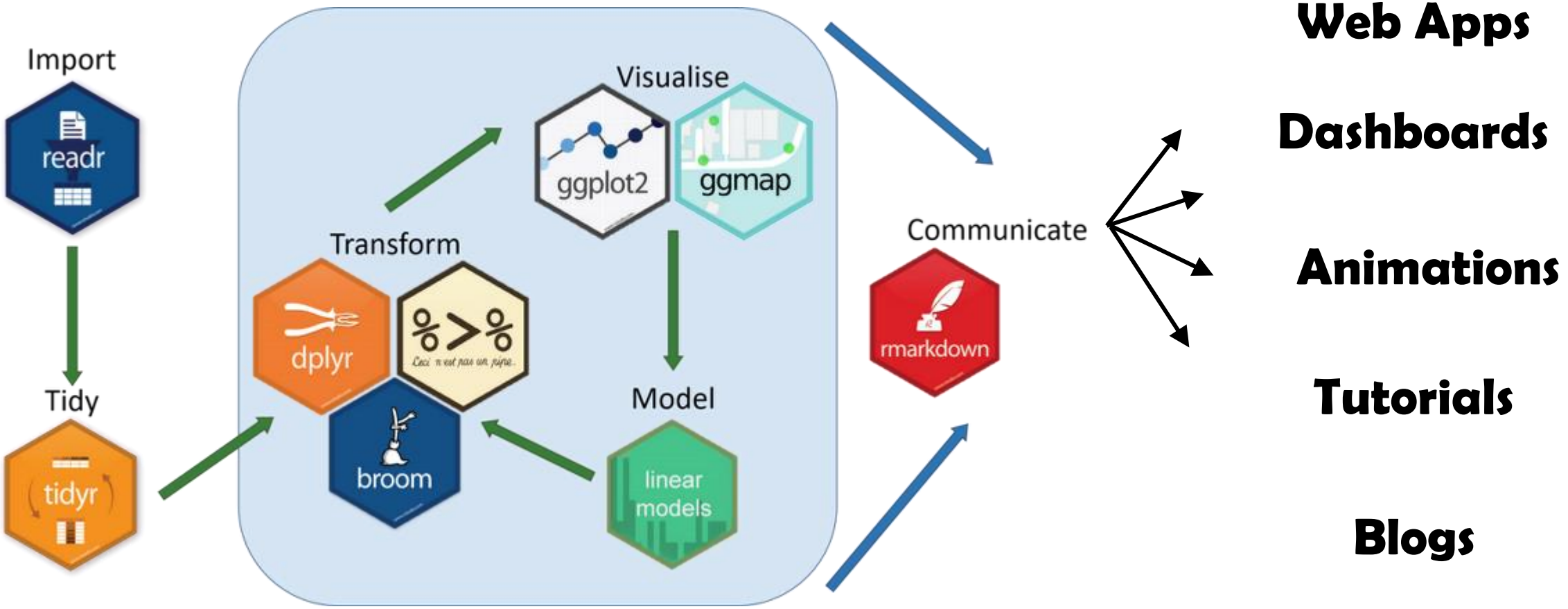
- Donec id elit non mi porta gravida at eget metus.
- Nullam quis risus eget urna mollis ornare vel eu leo.
  - Donec ullamcorper nulla non metus auctor fringilla.
  - Aenean lacinia bibendum nulla sed consectetur.
- Cras mattis consectetur purus sit amet fermentum.

Lorem Ipsum

Elit aut impedit alias deserunt blanditiis Placeat libero laboriosam quas iste nisi. Ut dicta nulla tempore beatae aliquid repudiandae soluta? Magnam rerum magni minus voluptate ea Eligendi harum obcaecati error

MORE DATASETS

Story-telling is built into the platform



Drill-Down Principle

Tweet / News Story

Curiosity

NCCS BLOG

Context

NCCS Ecosystem

IMPACT

VIGNETTE / PACKAGE

Content

Data & Scripts

A changing landscape:

competitors,
collaborations,
opportunities and
threats

https://x4i.org/about-us

Sponsored

x4i.org
https://www.x4i.org

NTEE Code by Source - Nonprofit NTEE Code Finder

Classification by Focus Area and under the UN SDGs. Nonprofit EIN, **NTEE**, and UN SDG Classification. Ads Free Search Engine. Ideas and Funding. Free data intelligence. Tech for NPOs Marketplace. Types: Geographic Differences, Benchmarks by Org Size.

Free Research and Reports · Free Interactive Reports · Systemic Problems · Tech Solutions

---

The NTEE code is **a three-digit code that specifies an exempt organization's purpose.** The National Center for Charitable Statistics (NCSS) developed the National Taxonomy of Exempt Entities (NTEE) coding system as a way to categorize nonprofit organizations.

Legal for Good
https://www.legalforgood.com › post › ntee-codes-what-t...

NTEE Codes: What They Are and How to Change Them

About featured snippets • Feedback

National Center for Charitable Statistics
https://nccs.urban.org › project › national-taxonomy-e...

National Taxonomy of Exempt Entities (NTEE) Codes

Jan 22, 2019 — **NTEE** Core **Codes** (**NTEE**-CC) Overview ... The National Taxonomy of Exempt Entities (**NTEE**) system is used by the IRS and **NCCS** to classify nonprofit ...

You've visited this page many times. Last visit: 4/28/2023

https://nccs.urban.org › publication › irs-activity-codes

Shahar Brukner

Simon Dickson

Tom Huberman

PEOPLE

Team

Dan Lev Almagor
Head of Product

Alex Nikiforov
Lead UX/UI Designer

Sergiy Kalinchuk
Senior Developer

Gregory Koshelenko
Senior Developer

Serhii Popov
Senior Developer

Jeff Poulos
Head of Partnerships

David Fox-Estrin
Partnership Director

Laurel Chausse
Customer Success Manager

https://impala.digital/