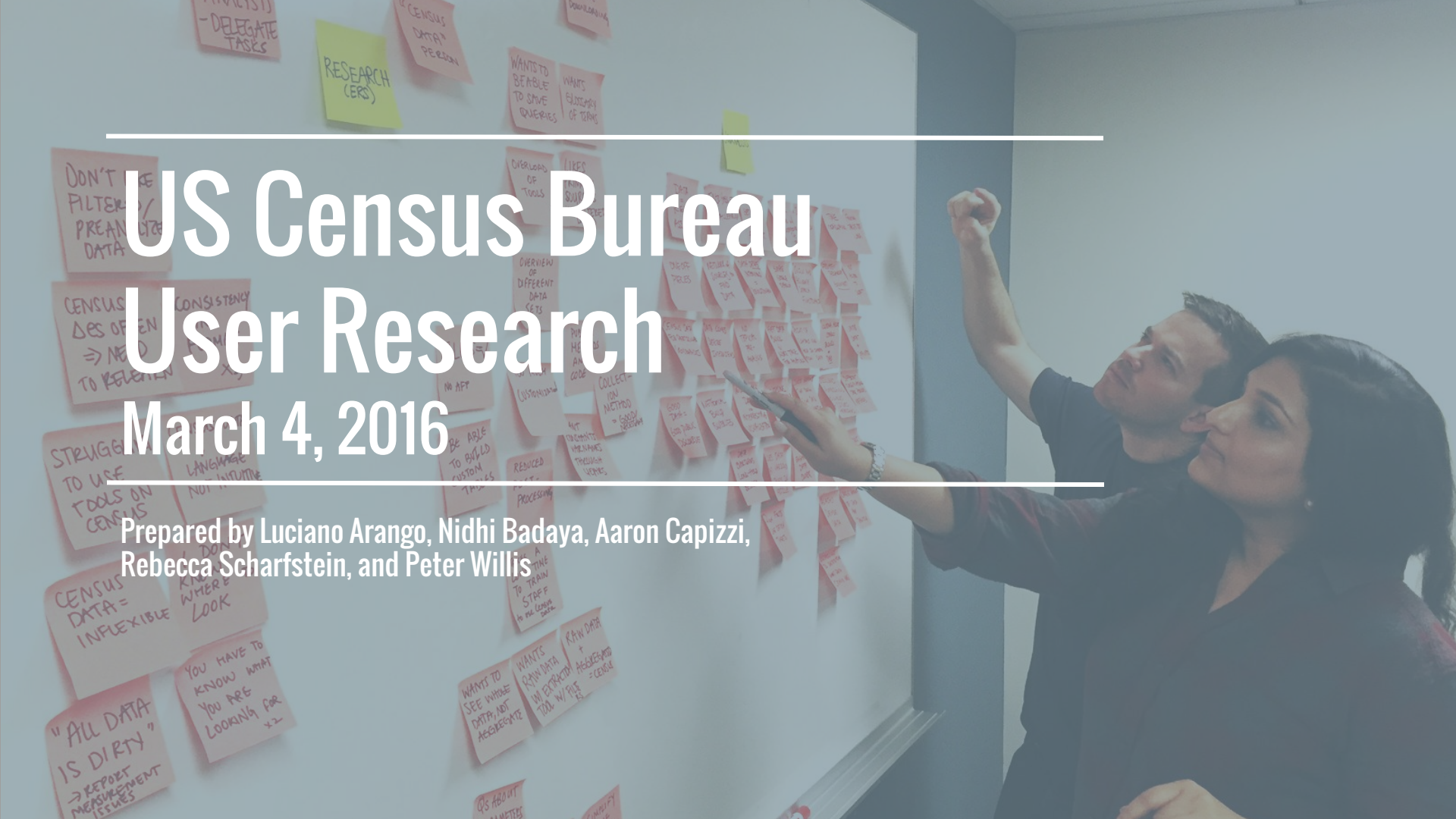


US Census Bureau User Research

March 4, 2016

Prepared by Luciano Arango, Nidhi Badaya, Aaron Capizzi, Rebecca Scharfstein, and Peter Willis

table of contents

client

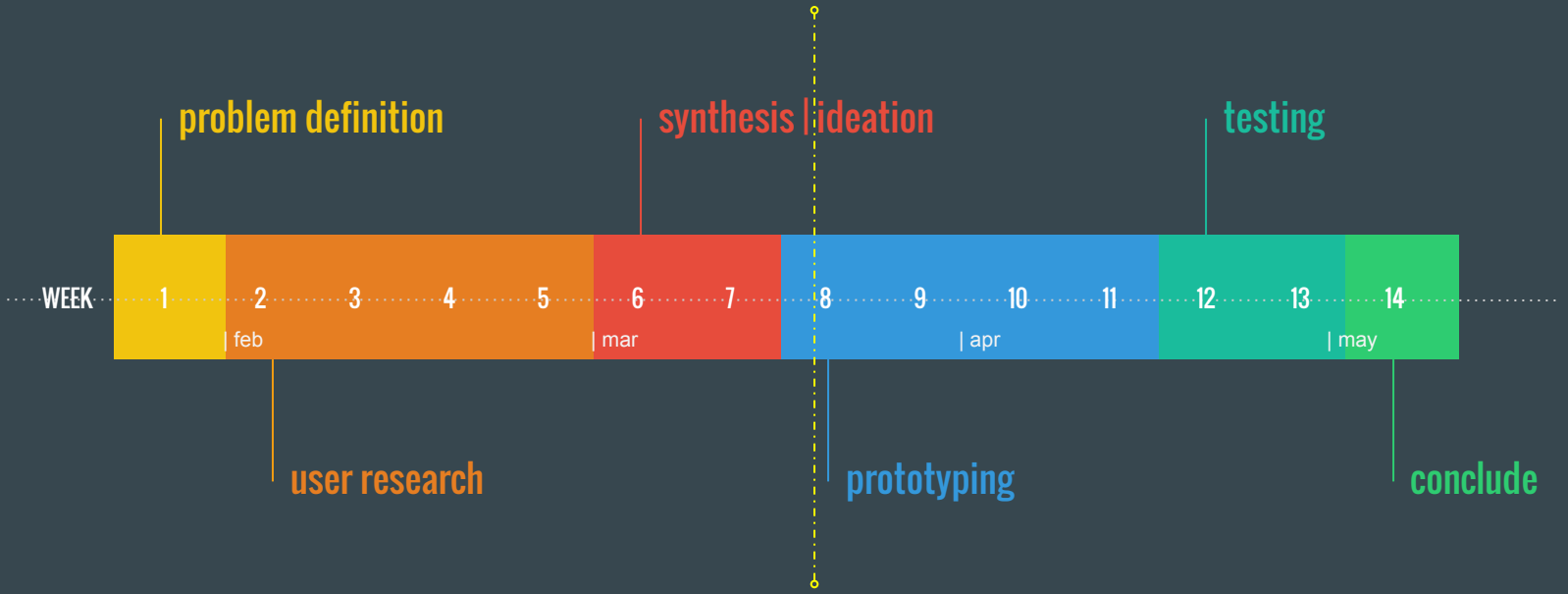
United States Census Bureau

The Census Bureau, part of the Department of Commerce, serves as the leading source of quality data about the nation's people and economy. We honor privacy, protect confidentiality, share our expertise globally, and conduct our work openly.

problem statement

How might we improve **user experience** - search, retrieval and incorporation - of US Census data, particularly **income distribution** data?

timeline


problem definition
synthesis | ideation
testing
WEEK
1
2
3
4
5
6
7
8
9
10
11
12
13
14
| feb
| mar
| apr
| may
user research
prototyping
conclude

METHODOLOGY

Through user interviews, we explored how individuals interact with Census data, what information they need to perform their jobs, and how their data search and retrieval process can be improved.

methods

Our research aimed to understand how users find, retrieve, and use data.

We engaged in a three step process in order to identify user personas:

(1) identify and interview user groups;

(2) synthesize findings into key insights; and

(3) articulate user personas that cut across user groups.

user groups

Who uses Census data?

media

data journalists

general journalists

influencers/ celebrities

research

research (academic)

research (private sector)

think tanks

government

government (federal)

government (state)

government (local)

politics

political campaigns

activists

community organizers

international

multilateral institutions

international organizations

private sector

financial institutions

consultants

real estate

developers

entrepreneurs

NGOs

individuals

concerned citizens

user groups

We evaluated user groups based on (1) practicality of research and (2) leverage, in order to determine which to target in our research.

1. Audience size, influence, reach
2. Access, likelihood of participation

user group focus

Our preliminary research focused on three user groups with high leverage and practicality.

sample interview questions

The Person

- How comfortable are you with technology?
- In which formats do you usually prefer to retrieve data?

The Need

- How do you use data in your [stories / research / grant applications]?
- How often do you search for data?

The Product

- How long, on average, does it take to incorporate data into your [stories / research / grant applications]?
- Do you perform operations on data, such as cleaning, filtering or analysis?

The Process

- What is your search process?
- Tell us about a time when it was really easy to find data.
- Tell us about a time when it was really difficult to find data.
- What are some common bottlenecks or issues in finding the data you need?
- How often do you use the US Census as a data source?
- How would you characterize your experience using the US Census website?

Income Inequality

- Do you search for income data?
- Are there particularly valuable data sources with respect to income or income inequality?
- Is income data unique from other data sets or topics?

design exercises

We used post-it notes and design exercises to synthesize our findings across user groups.

Identified user group characteristics

Reorganized notes based on thematic insights


NGO
FOCUSED ON QUALITY DATA
LOOKING FOR CREDIBLE DATA
CENSUS A GOOD SOURCE OF DATA
DATA ANALYSIS - DELEGATE TASKS
RESEARCH (ERS)
CLASSIFICATION NOT ALWAYS FIT FOR PURPOSE
NAVIGATION NOT INTUITIVE ON SITES
DOWNLOADS TO EXCEL
USES AMERICAN FACT FINDER
LIKE DATA PUSHED TO HER

FINDINGS

key insights

We generated 8 key insights from our user interviews:

Data

1. Users need data to do their jobs properly.
2. For users, data is personal and local.
3. Users go to extraordinary lengths to get the right data.
4. Manipulating data takes work, users don't expect shortcuts

Census

5. Census data can overwhelm and deter users.
6. In the minds of users, Census data exists in an ecosystem.

Trust and support

7. Users want assurance they are using data properly.
8. Users value personal support.

1 key insight – data

Users need data to do their jobs properly.

➡ Data allow users to perform the core responsibilities of their jobs, including

- Target spending better
- Write more engaging research papers
- Overturn popular assumptions
- Improve the quality of public debate
- Prove there is a need for their service or policy

➡ Users use data for a professional purpose. We did not interview users who explore data as a hobby.

> "[Data] makes the story more concrete"

> "There is evidence in the data that your hypothesis is right or wrong"

2 key insight – data

For users, data is personal and local.

- People want data that is important to *them*.
- Users look for data specific to their needs. Some are addressing a particular issue for the entire nation. Others want to collect all available data in a particular geographic area.
- Many users think in regional terms. They must align Census tract data with local neighborhood information to inform their activities within the community.

> "But I just want the data for Medford [City]..."

> "If you can hover over one region and get all the data for that area, that would be huge"

3 key insight – data

Users go to extraordinary lengths to get the right data.

- Many users spend hours searching for and extracting data.
- Other users sign up for webinars and develop personal relationships with data dissemination specialists who are employed by data providers
- Some users pay 3rd party providers for curated and difficult-to-find data that is provided for free on the Census website.

> To get one data point from census.gov the first time took me 12 hours. Now it only takes a couple

> If we can't get the data, we work around it

4 key insight – data

Manipulating data takes work, users don't expect shortcuts

- Some users like to play with the raw data. They want to understand the links between data sets and relish the challenge of overcoming the difficulties of the search process.
- Other users see data manipulation as a burden. When they want curated data they ask a colleague with analytical skills, such as a data journalist or Census employee, to manipulate the data for them.
- Despite their different attitudes towards the process, all users expect that they will have to do some work to get the right data. No-one expects it will be easy or straight-forward.

> "I would prefer an easy source but it's not always possible"

> "All data is dirty"

5 key insight – Census

Census data can overwhelm and deter users.

- The scope of Census collection can be daunting. Users may be discouraged from exploring Census data and finding if they don't know where to start.
- Too many tools complicates the issue by providing an excess of choice. Users unsure which of many choices is the right one, end up not choosing anything.

> daunting

> I wouldn't use Census as my first source because I don't know what they have!

> Why 5 mapping tools for the same data? I just want ONE place I can find what I need

6 key insight – Census

In the minds of users, Census data exists in an ecosystem.

- People use multiple data sources for their work.
- Going to multiple sources is not a cost itself; however, if it's consistently difficult to find data it becomes a very time consuming process.
- People keep downloads of data on their local hard drives and build their own databases. Their relationship with data sources is dynamic and customized for their own use.

> My first step is to ask another journalist [if looking for data]

> I never go straight to the Census website...instead I search through old articles to get context

7 key insight – trust and support

Users want assurance they are using data properly.

- Demographic data is complex and intricate. Professionals want to understand the methods, context, and quality of Census surveys to ensure the data supports their conclusions.
- Some users don't trust data until they inspect and exercise it. For them, accessing the data in raw form is best.
- Many users want to ensure data is searched and gathered comprehensively; missed data or geographic areas could undermine their purpose

> "If there's too many areas [where] I could get it wrong, I'll walk away"

> "They keep throwing datasets at you and don't explain what they do"

8 key insight – trust and support

Users value personal support.

- Data users seek out responsive, supportive providers. Even the most experienced data analysts want to establish personal relationships with data sources.
- Many users, when confronted with a data problem, reach out to a person for help by default. This can turn a potentially frustrating event into a positive experience.

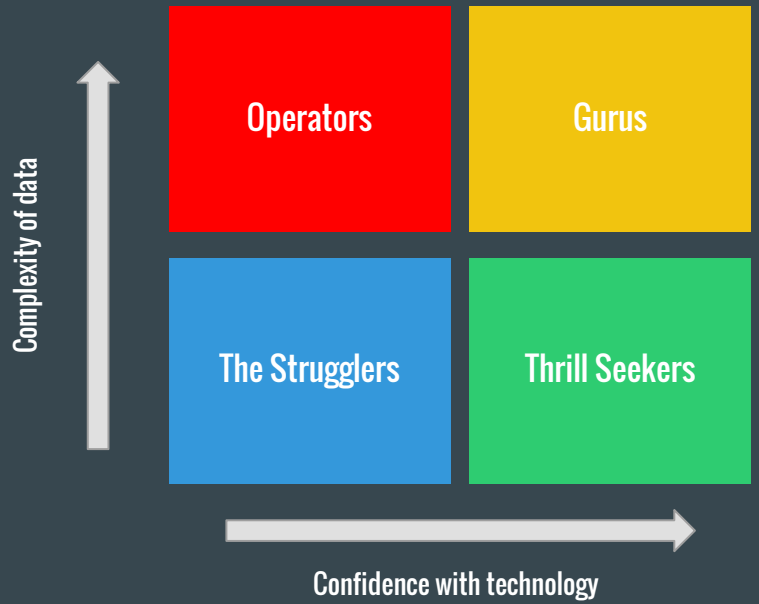
> "Alexandra is great!!!"

> "I'm very good at telling people what I need"

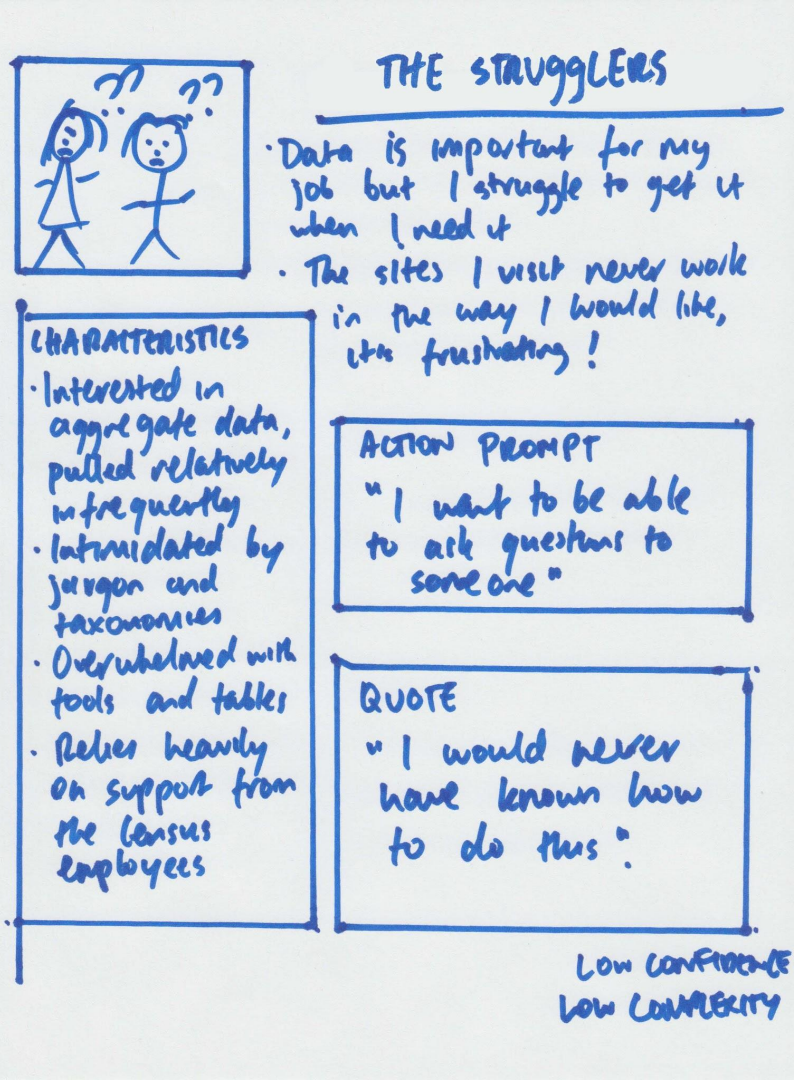
> "I'm dependent on the people at the Census to crunch the numbers for me"

personas

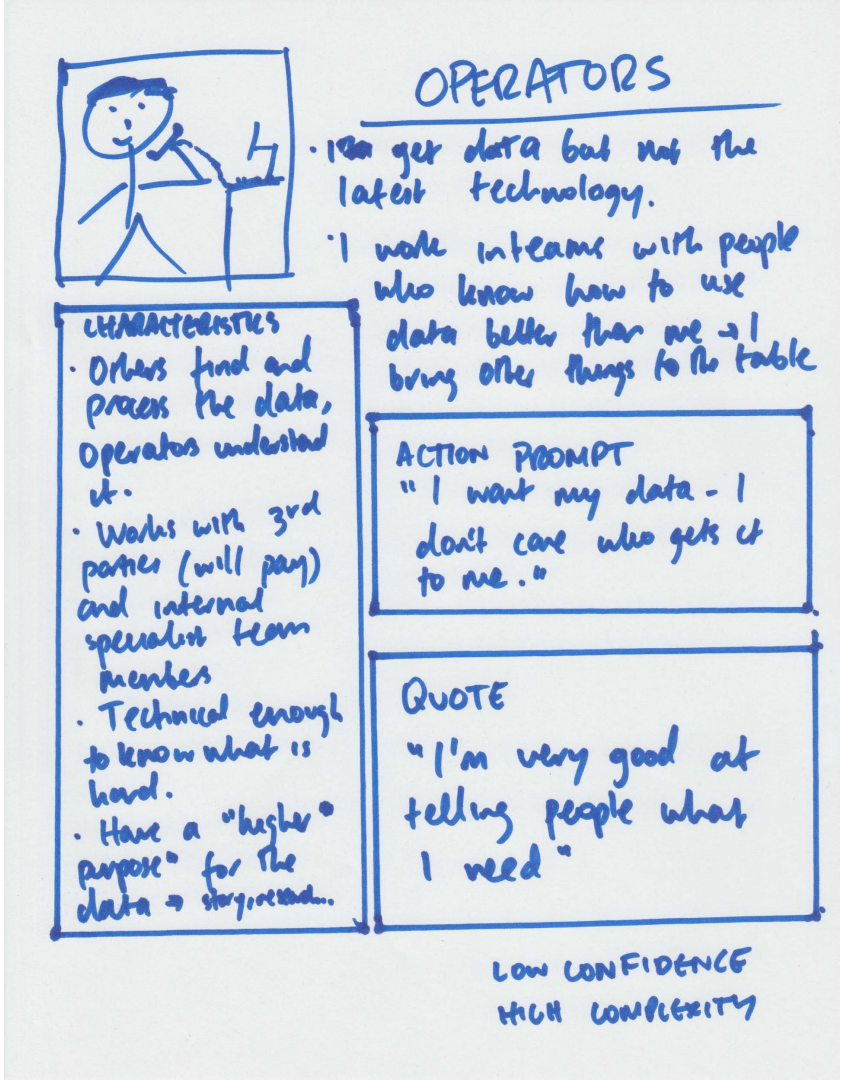
To create personas we segmented users according to their confidence with technology and the complexity of the data they use.



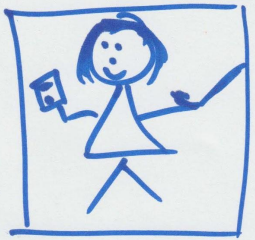
persona #1:
the strugglers



persona #2:
operators



persona #3:
thrillseekers



THRILL SEEKERS

- I've been working with data for years → it's fun and it's my job.
- I like to "play with the data." I enjoy the challenge of finding some thing new

CHARACTERISTICS

- Deals with large data sets
- Likes raw data, trusts it more than someone else's processed data.
- Loves finding unusual, quirky data
- Enjoys the challenge of using legacy interface websites

ACTION PROMPT

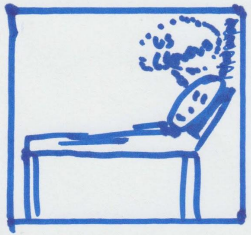
"I want direct access to all the raw data

QUOTE

"Why 5 mapping tools? I just want one place where I can find what I need."

HIGH CONFIDENCE
LOW COMPLEXITY

persona #4:
gurus



GURUS

- I'm an experienced user, a specialist.
- I answer everyone's questions internally at my workplace
- I'm deep in specific areas, sources.

CHARACTERISTICS

- Uses advanced tools, coding skills
- Interested in metadata, the way classifications are made
- Does not use packaged tools, quick fact finders
- Maintains their own databases, matches data sets

ACTION PROMPT

"I want to engage with the people collecting and cutting the data."

QUOTE

"I want control... you lose freedom if you can't be completely in control of your [research] question"

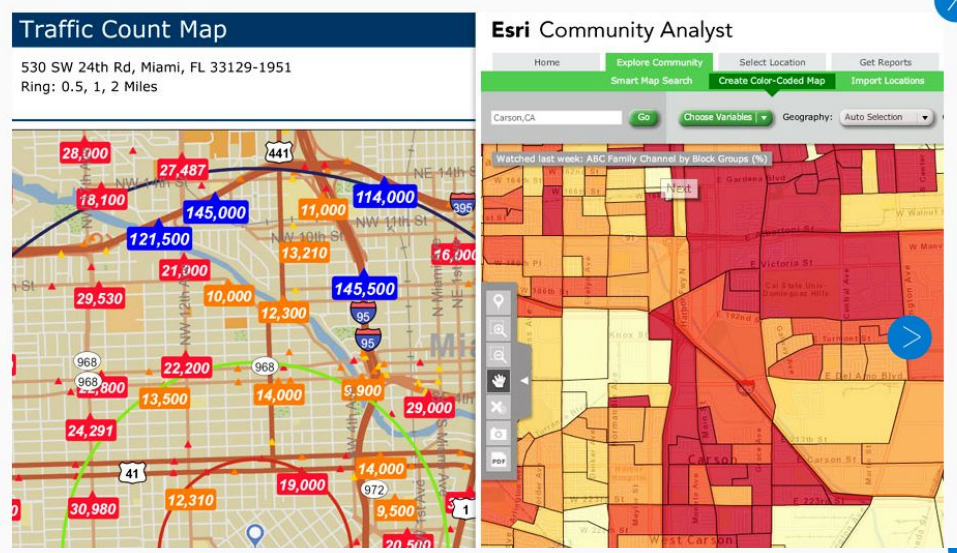
HIGH CONFIDENCE
HIGH COMPLEXITY

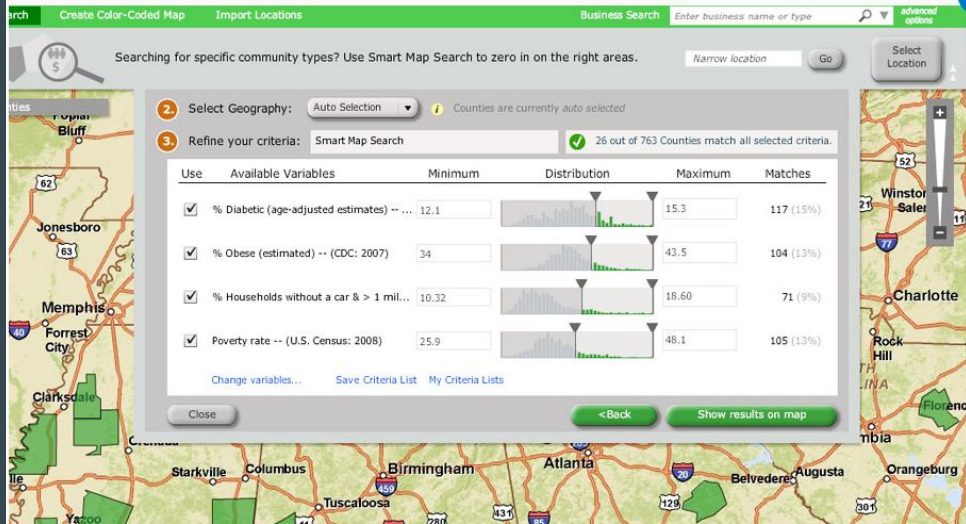
secondary

We asked users to share their experiences using other data sites. The following websites were positively endorsed by users.

ESRI

Community Analyst



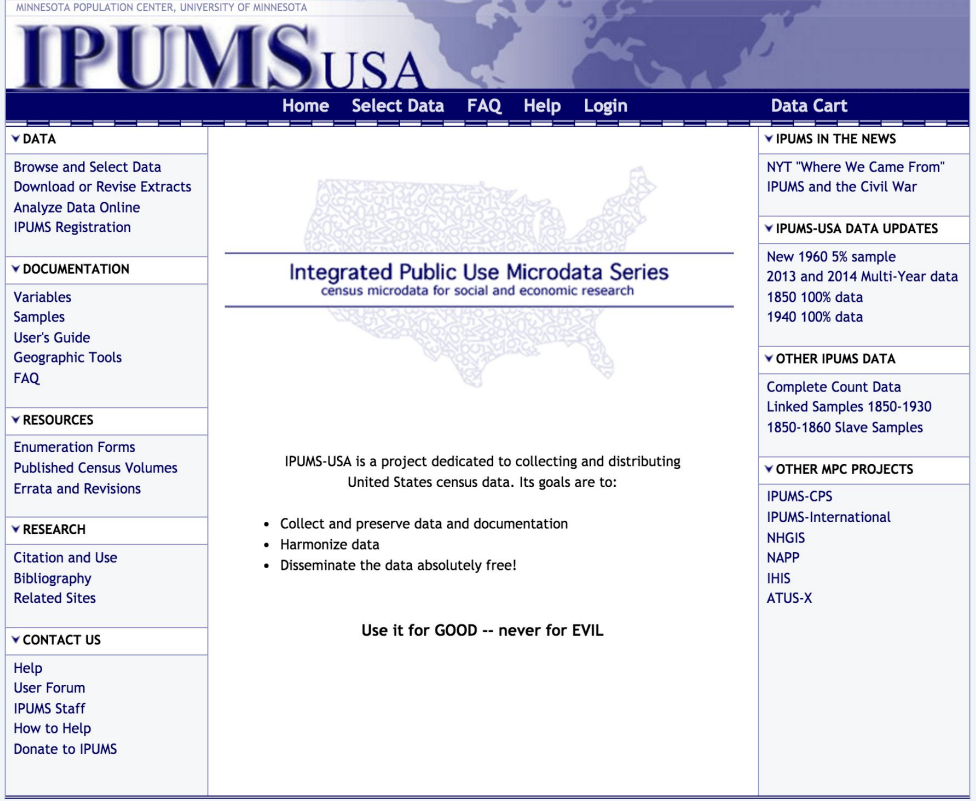


Minnesota

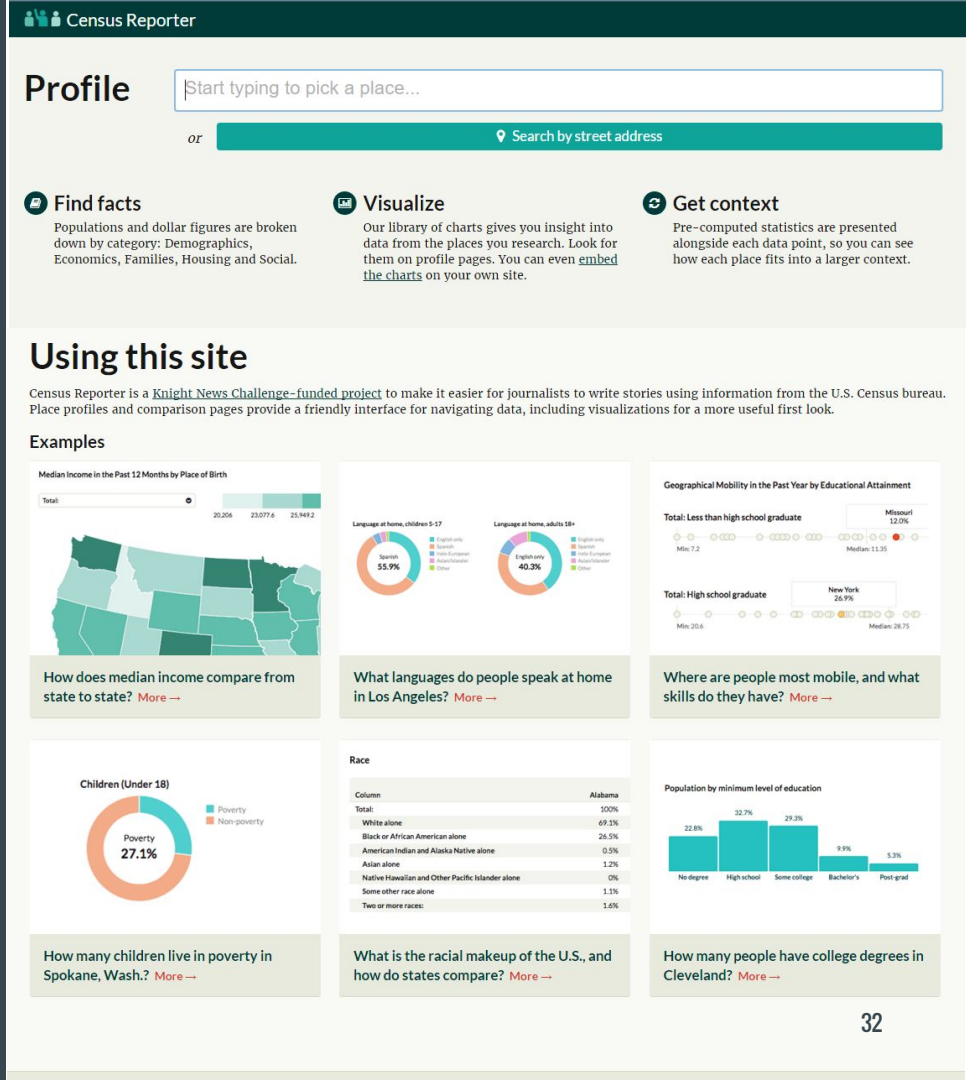
State Demographic Center

IPUMS

Integrated Public Use Microdata Series

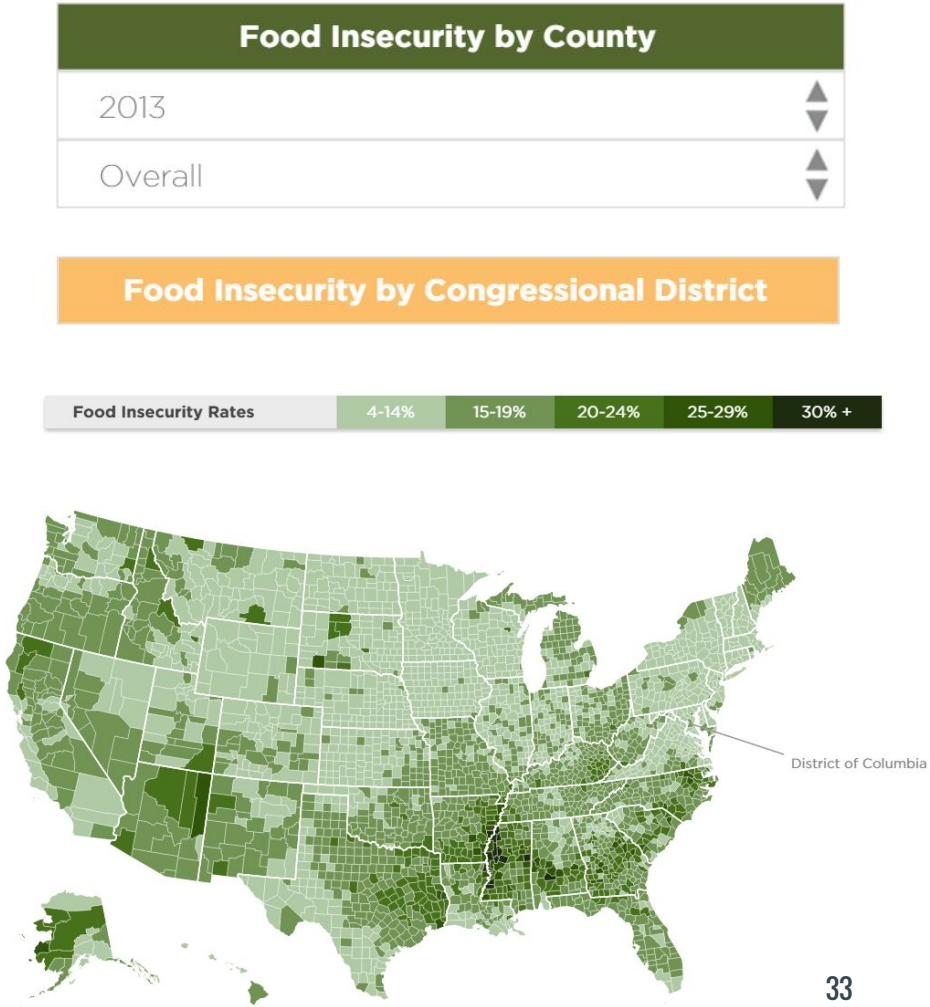


Census Reporter



Map the Meal Gap

Feeding America



Assets & Opportunity Scorecard

CFED