A Cross-Domain Need-Finding Study with Users of Geospatial Data

Parker Ziegler
peziegler@cs.berkeley.edu

Sarah E. Chasins
schasins@cs.berkeley.edu

Background

- Geospatial data is playing an increasingly critical role in the work of **Earth and climate scientists**, **social scientists**, and **data journalists**
- Existing GIS software and programming tools for geospatial analysis and visualization are challenging to learn and difficult to use
- We **identified the unmet computing needs** of an expanding community of geospatial data users

Method

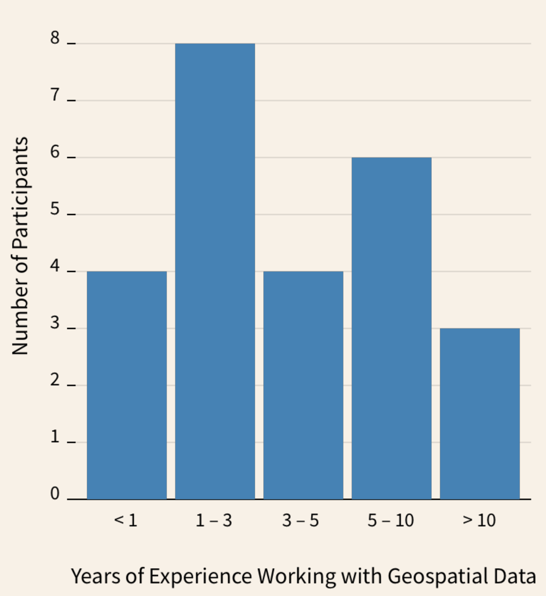
We conducted a **contextual inquiry** with 25 geospatial data users in **Earth and climate science**, **social science**, and **data journalism**.

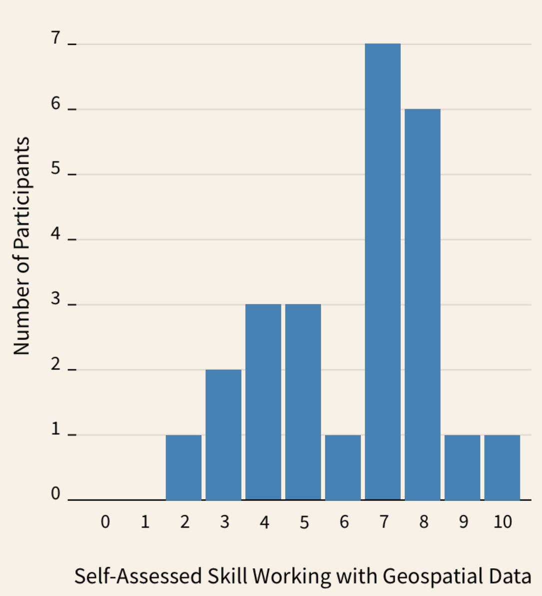
9 Earth and climate scientists

6 social scientists

8 data journalists

2 interdisciplinary





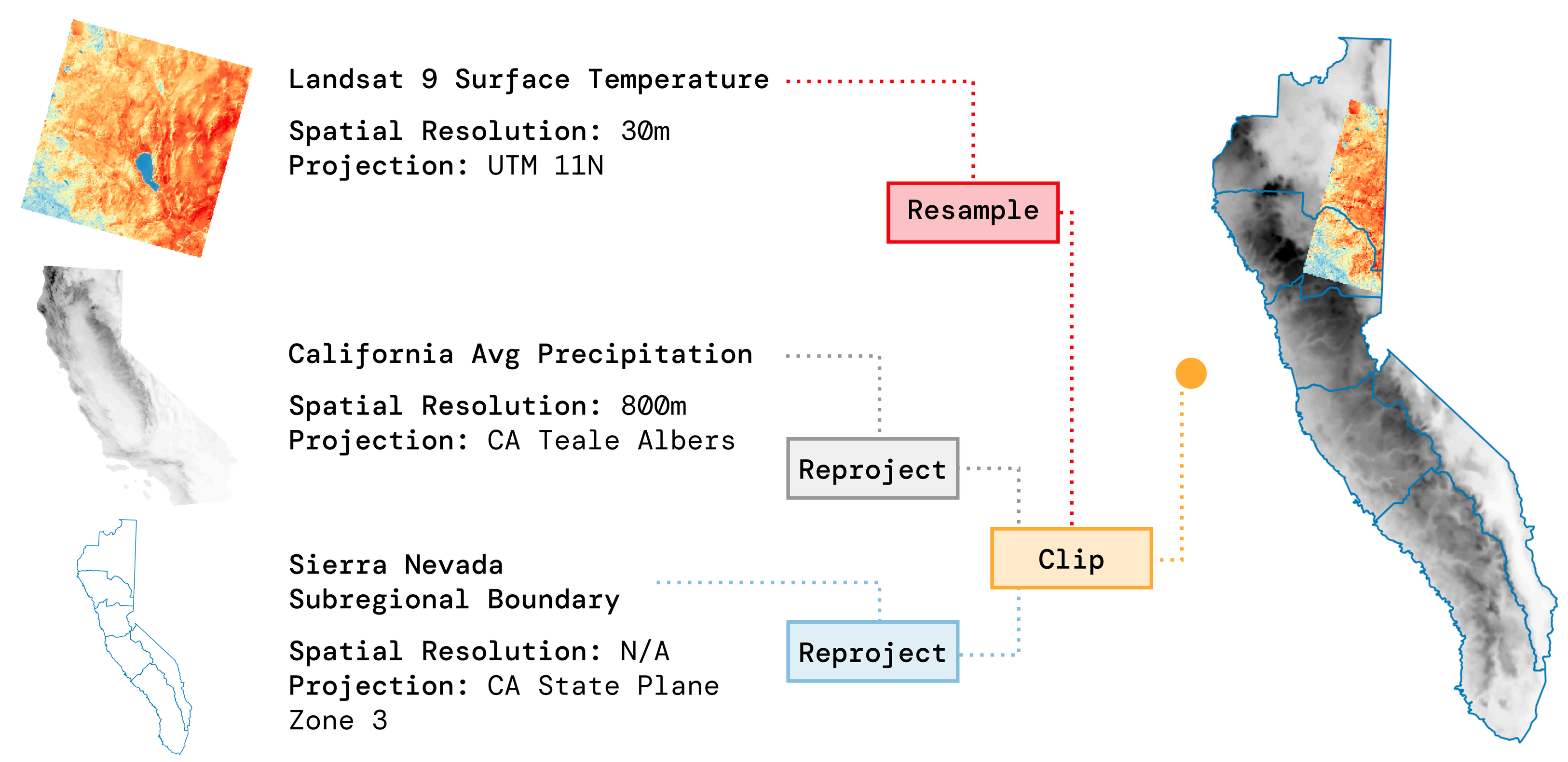
Design Opportunities

- Assistive tools for moving geospatial data between cartographic representations efficiently
  - → **Cartographic grammar of graphics** paired with **direct manipulation interfaces**
- Assistive tools for inferring geospatial analysis programs from constraints
  - → **Program synthesizers** using user demonstrations in map interfaces

EPIC DATA lab UC Berkeley · Berkeley UNIVERSITY OF CALIFORNIA

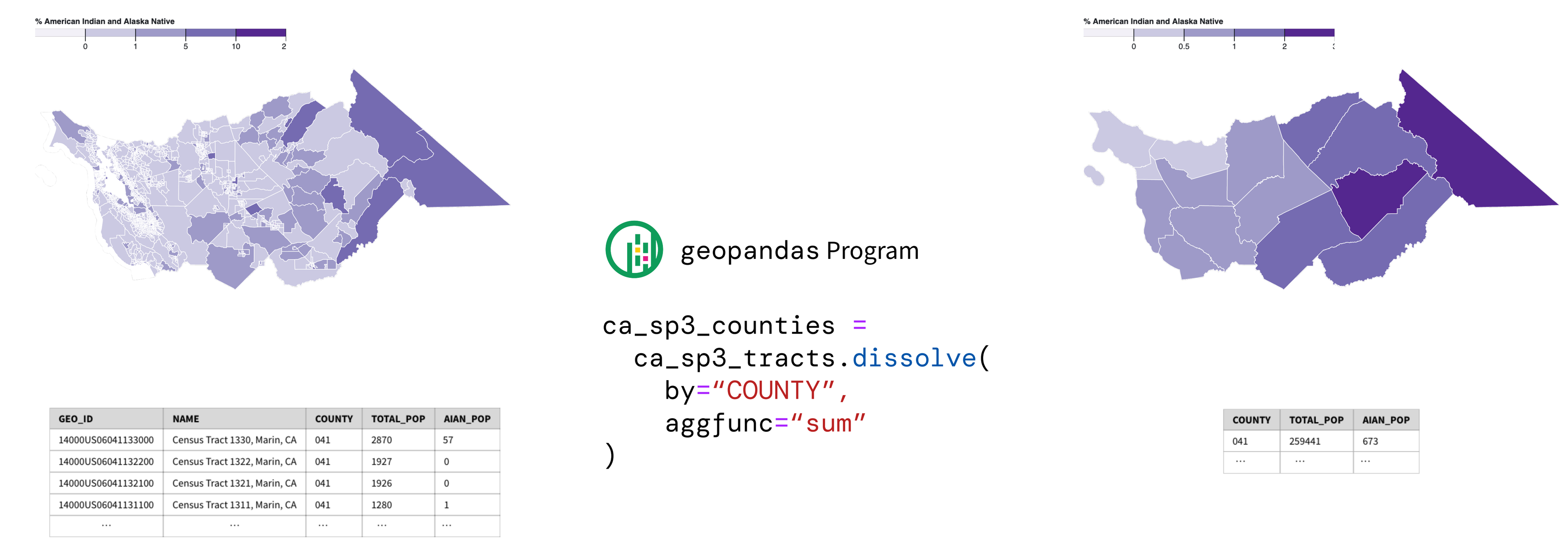
Domain experts using **geospatial data** face significant challenges in **data discovery**, **data transformation**, **analysis**, **analysis representation**, and **visualization**.

1. Finding and transforming geospatial data to satisfy spatiotemporal constraints



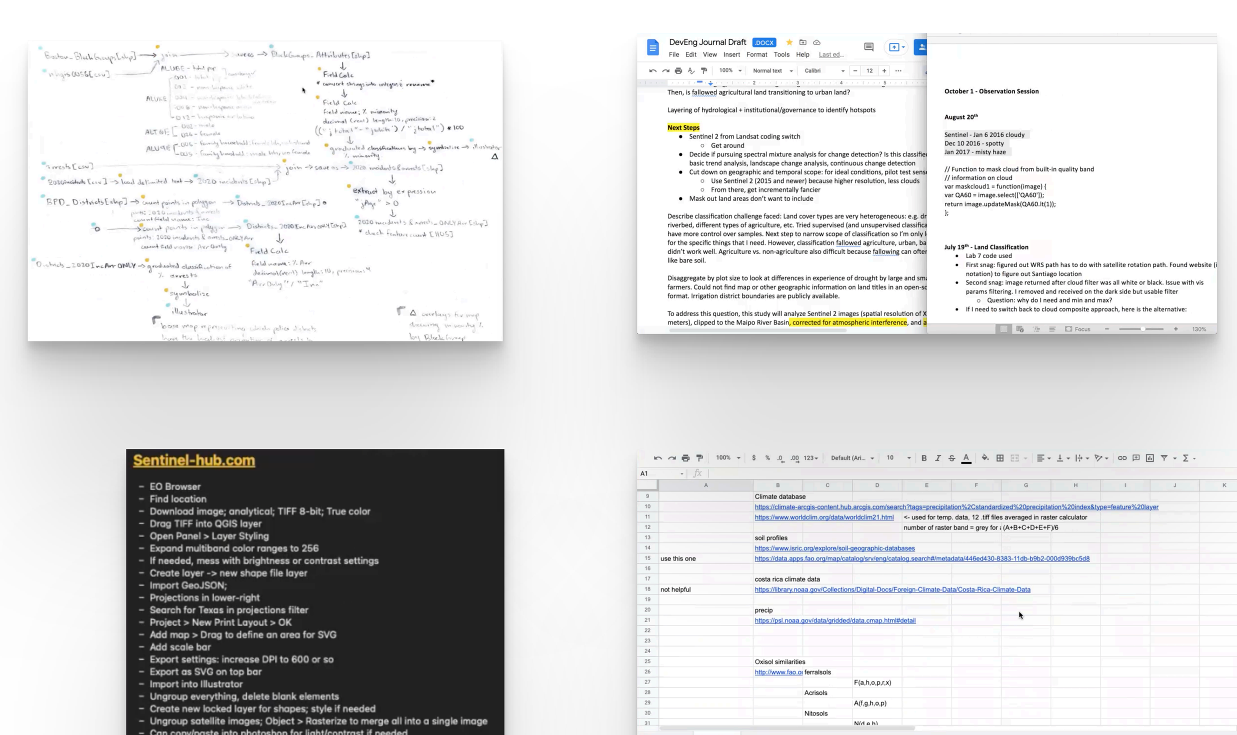
Transforming geospatial data. Geospatial data is captured at different spatial resolutions, extents, and projections and must be transformed to a common reference. Here, we resample a Landsat 9 surface temperature raster to match the resolution of a California precipitation raster, then clip to the Sierra Nevada range.

2. Identifying and understanding the semantics of geospatial operators



A plethora of geospatial operators. Modern GIS software and geospatial analysis libraries have hundreds of operators to choose from. Identifying and understanding the behavior of operators was difficult because it required reasoning about transformations to both the geometry and attributes of geospatial data.

3. Tracking geospatial data provenance



Informal program representations of geospatial analyses created by participants. These representations varied widely from bulleted lists tracking high-level analysis steps, spreadsheets and documents mixing code snippets with links to data sources, and even custom graphical notational systems.

4. Efficiently exploring the cartographic design space

A selection of draft maps made by a data journalist participant. This participant created over 20 drafts, experimenting with cartographic representations ranging from choropleth to dot density to gridded heat maps in small multiple layouts. Some drafts even attempted to show static UI states in response to user interaction.