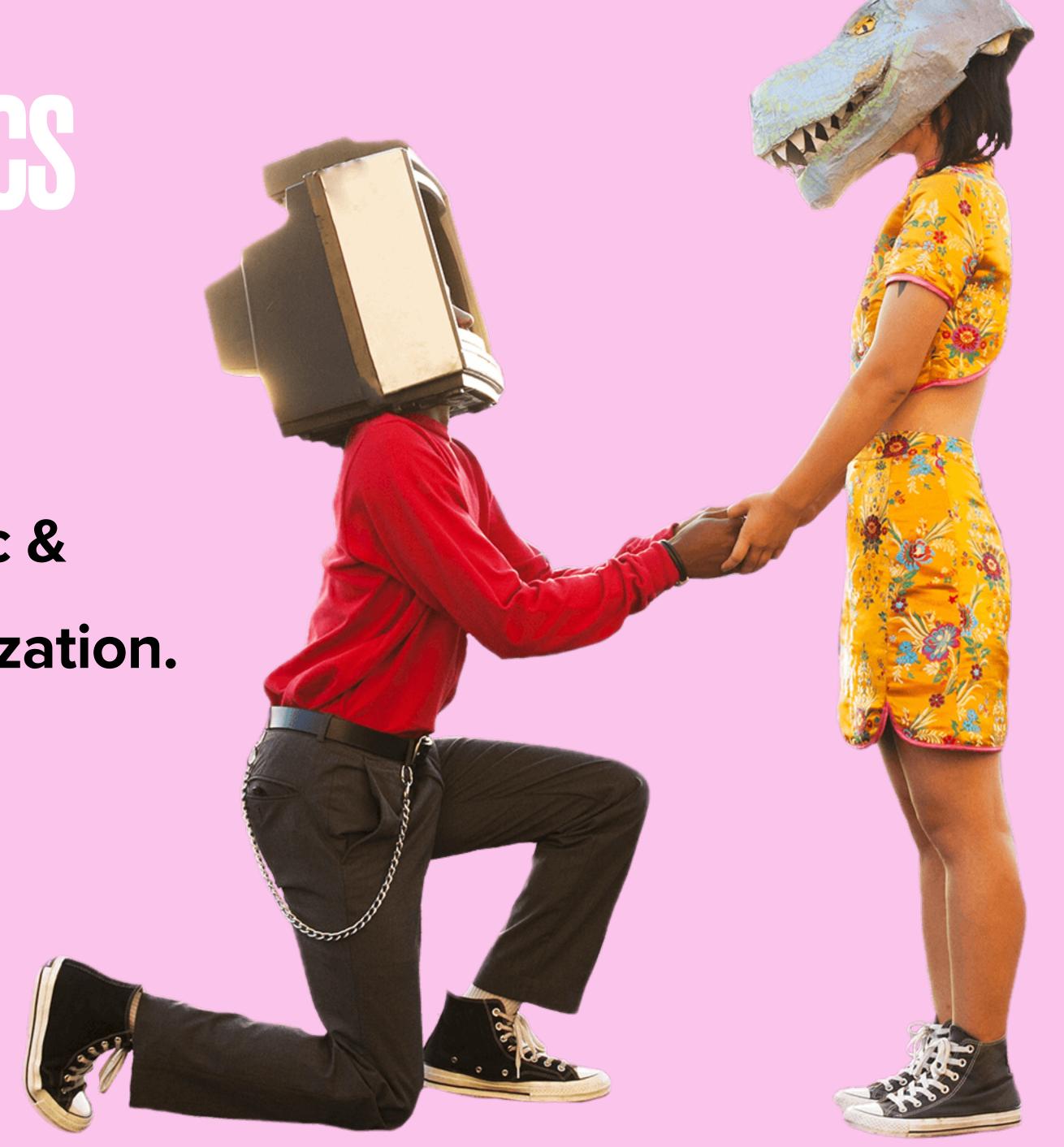
## DATA SOURGE AND ETHICS

Scraped from user profiles.

▶ Data approved for academic & business use after anonymization.

**Source:** ResearchGate





## DATA CLEANING & PRE-PROCESSING

- ▶ Loaded profiles, LatLon, addr, and sharedCensus datasets.
- Merged datasets on the 'location' column.
- Removed age outliers (e.g., 109, 110).
- ▶ Corrected data points like "drinks desperately".
- Handled missing values.
- Recategorized 'education' for clarity.
- Set boundaries for US geographic data.