



**Team 7**

**Project Title: Local Governments Survey (Harmonizing ARISE  
City Survey + CivicPulse Kansas Sample)**

**Course Number:CS896**

**Team Names: Anudeepthi Chirumamilla- Y828R524  
Rupali Avhad - d798z925  
Teja Sree Mukkapati-Q389W584**

## **Week 2: Understanding Survey Structure & Identifying Overlapping Variables**

### **Objective**

Identifying variables that capture similar concepts across datasets and comprehending the structure of both surveys are the goals of Week 2. This step is critical to ensure accurate data harmonization in later analysis.

### **Surveys Used**

We are working with two independent survey datasets focused on Kansas local governments:

- **ARISE Kansas City Survey**

This city-focused dataset includes comprehensive questions about government capacity, infrastructure condition, workforce problems, and financial hardship. In addition to using numerical response scales, it contains several missing or irrelevant values.

- **CivicPulse Kansas Sample**

This dataset comes from a nationwide survey that was restricted to people in Kansas. It covers a variety of governmental entities, including counties, municipalities, and special districts, and emphasizes workforce-related and capacity-related metrics with response scales that are clearly identified.

### **Role of Codebooks**

- The purpose of the codebooks is to:
- Recognize the meaning of each variable.
- Analyze value labels and response scales.
- Determine the coding for missing or "not sure" responses.
- Make sure both datasets are interpreted consistently.

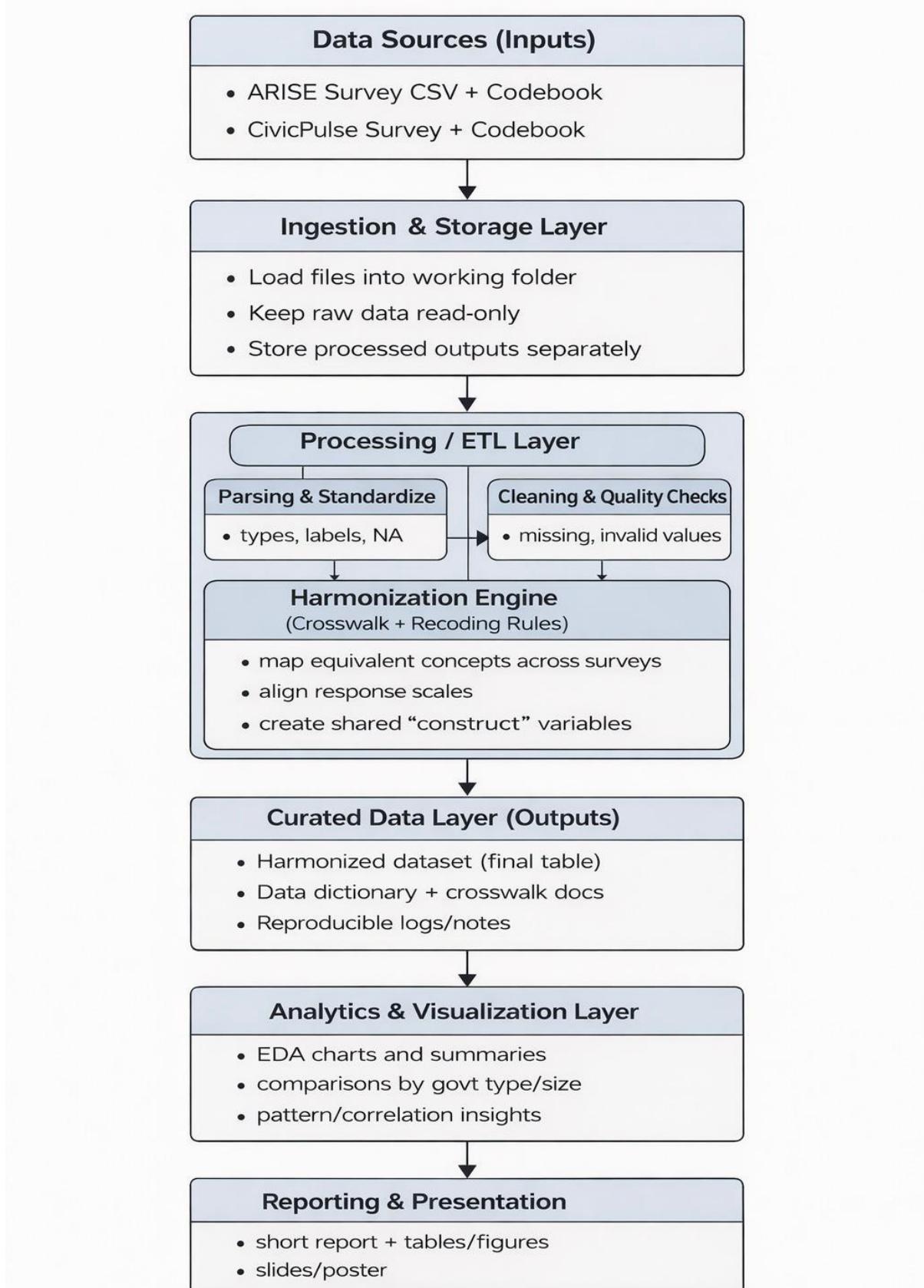
### **Identifying Overlapping Variables**

Using the surveys and their codebooks, we identified variables that measure similar underlying concepts, even if the wording or scale differs.

<b>Concept / Theme</b>	<b>ARISE Survey</b>	<b>CivicPulse Survey</b>	<b>Notes</b>
Workforce Challenges	Difficulty hiring and retaining enough staff	Challenge in maintaining a sufficient workforce	Same concept, different wording
Fiscal Condition	Overall fiscal condition of the local government	Overall fiscal condition rating	Same ordered scale
Infrastructure / Capital Assets	Condition of infrastructure systems and capital assets	Overall condition of capital assets and infrastructure	ARISE is more specific; CivicPulse is broader
Infrastructure Funding Constraints	Statements about unstable funding and unmet infrastructure needs	Agreement statements about infrastructure financing limitations	Aligned through agreement scales
Government Characteristics	City-level administrative information	Government type (municipality, county, district)	Used for comparison, not direct merging

These variables are noted and mapped conceptually so they can be aligned during the harmonization phase.

## System Architecture:



## Week 3 Plan: Data Cleaning and Handling Missing or Inconsistent Values

By removing inconsistent, invalid, and missing responses, we will clean up both datasets. This entails fixing coding errors, handling NA and "Not sure" values, and standardizing variable forms. The following step will involve preparing the cleansed datasets for trustworthy harmonization and analysis.