



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

Project Introduction

The goal of this project is to create a clean, analysis-ready dataset by merging two actual surveys of Kansas municipal governments. CivicPulse conducted a nationwide poll with a Kansas sample, whereas ARISE conducted a city study limited to Kansas. Although the questions are different, both surveys discuss capacity issues such as service delivery, personnel difficulties, and infrastructure condition. It is our responsibility to bring them into harmony so they can be studied together.

Project Overview

Main goal: Create a harmonized dataset + documentation + exploratory analysis that helps study local government capacity in Kansas.

Key idea: Different surveys often measure the same concept using different wording or scales. We will align them into shared constructs.

Deliverables: Harmonized dataset, data dictionary/notes on decisions, and a short written report with charts/tables and key findings.

Why we chose this project (as Data Science students)

- It is a real data-wrangling project, not a perfect 'clean' dataset. This is what industry and research data looks like.
- We get hands-on practice with codebooks, variable definitions, missing values, and inconsistent response scales.
- It strengthens core Data Science skills: data cleaning, merging, recoding, documentation, and communicating results.
- It has a clear policy impact: the analysis can help understand what Kansas local governments are struggling with (workforce, infrastructure, fiscal stress, etc.).
- It's a great portfolio project because it shows end-to-end work: raw data to final dataset to insights.

Datasets

ARISE Kansas City Survey (Kansas-only)

- File: Rachel_Krause_ARISEData_Combined General Dataset.csv
- Size: 309 responses (rows) and 91 variables (columns)
- What it contains: detailed information on infrastructure systems, governance capacity, grant challenges, fiscal stress, and more.

CivicPulse Survey (Kansas sample of a national survey)

- File: CivicPulse Codebook.xlsx (codebook)
- What it contains: respondent role/government type, contextual county measures, and survey questions including workforce challenges and capacity measures.
- Note: Our work will use the CivicPulse survey data (Kansas subset) along with the codebook to correctly interpret each variable.

10-week step-by-step plan

We will follow a simple weekly plan. The key is to move from understanding the surveys to building harmonized variables, and finally creating insights and a final report.

Weekly Plan	Focus
Week-1	Project setup, requirement review, and dataset inspection.
Week-2	Understanding survey structure and identifying overlapping variables
Week-3	Data Cleaning and handling missing or inconsistent values.
Week-4	Data harmonization and variable alignment

Week-5	Finalizing harmonized dataset and documentation
Week-6	Exploratory data analysis and visualization
Week-7	Comparative analysis across jurisdiction size and type
Week-8	Correlation analysis and pattern identification
Week-9	Interpretation of results and policy -relevant insights
Week-10	Final report, poster and presentation

Final Outcome

- A harmonized dataset that aligns key constructs across the ARISE city survey and CivicPulse Kansas sample.
- A clear crosswalk/data dictionary that explains every harmonization decision (so someone else can reproduce it).
- A short written report with: (1) project background, (2) methods for harmonization, (3) descriptive insights, and (4) simple charts/tables.
- a joined dataset with public context variables (county/city demographics) and a few simple visuals or maps.

Next Week Goal

We will study both surveys and their codebooks to understand the data layout. Then we will identify and align variables that capture similar concepts for use in later analysis.

