

- 1. Query both CMS API and U.S. Census Bureau API for necessary data
- 2. Feed this data into Kafka Producer with each dataset as a unique topic
- 3. Feed Data into Azure Data Lake from Kafka Consumer
- 4. Run ETL on data in Data Lake
  - 4.1. Drop unnecessary rows
  - 4.2. Impute/remove rows with missing data
  - 4.3. Prepare for ML and visualizations
- 5. Store this data in Azure SQL Database
- 6. Use data in database to train and tune a machine learning model
- 7. Create visualizations with data from database and results of machine learning model
- 8. Create dashboard in PowerBI with visuals