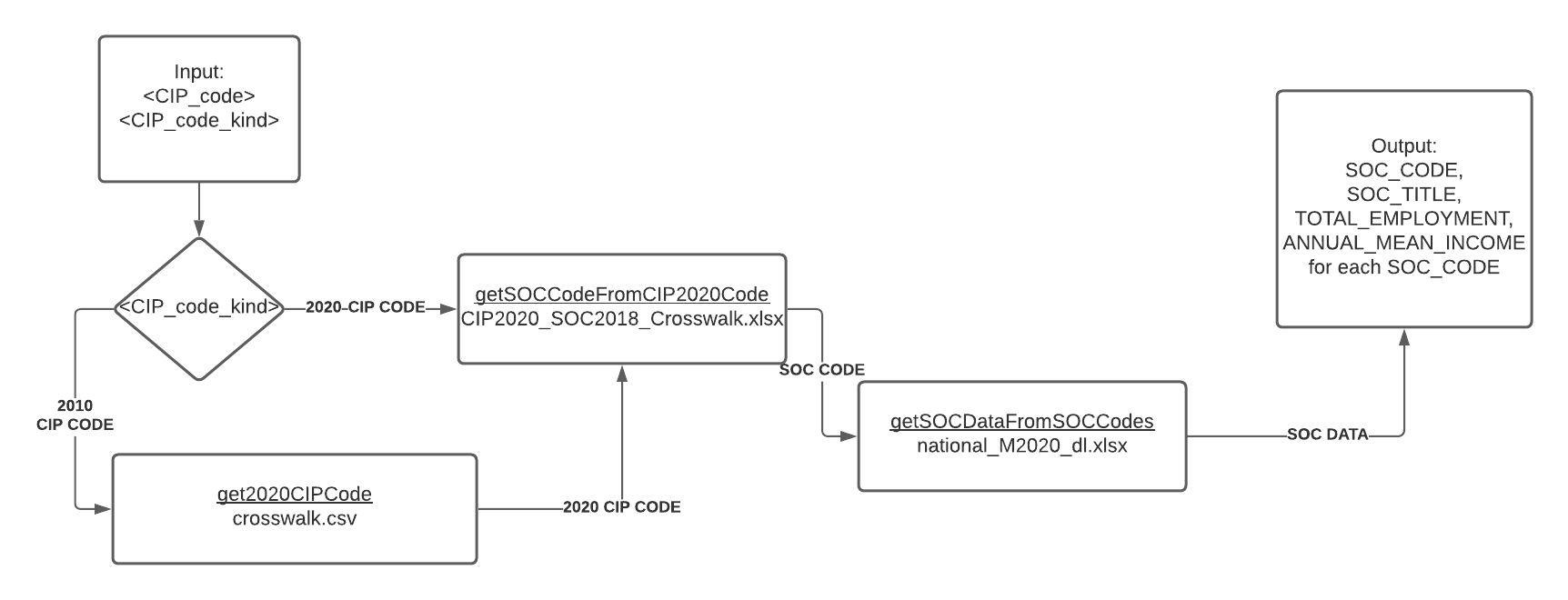
List of Deliverables:

1. Gather DoE and BLS data and created a generic crosswalk (many-to-many relationship)
2. Create system to convert CIP codes to SOC data/information
3. Gather school specific data from career services next steps after graduation
   1. Brown University
   2. Davidson College
   3. Georgia Institute of Technology
   4. Haverford College
   5. Northeastern University
   6. The Ohio State University
   7. Yale University
4. Compile data into psql database (stride\_db)
   1. bls2020
   2. bls2020description
   3. cip2010\_cip2020
   4. cip2020\_soc2018
   5. school\_career\_outcomes
5. Use DoE and BLS data to connect job names to SOC names using fuzzy matching techniques:
   1. Name matching
   2. Ngrams matching
   3. soundex()
   4. metaphone()
   5. TF-IDF
6. Create data visualization using data attached to SOC codes and percent industry from stride\_db
7. Create an API using FastAPI to help with CIP lookup and SOC matching

Next Steps:

1. Methods of analysis
   1. Establish percentage cutoff (reduce one-to-many matching)
   2. Weighting estimated earnings by job percentage
   3. Differences between school types (small vs. large, public vs. private, liberal arts vs. technical)