

Trade & AI: Report 1

Nardeen Abdulkareem
The University of Toronto
January 18, 2026
nardeen.abdulkareem@mail.utoronto.ca

1. Data

1.1 *Datasets*

Our data contains eleven sets of information which includes the unified individual data set. The unified individual data set is of high importance in the subsequent analysis. We observe ID's for the job postings, companies, titles, categories, skill sets, locations, expected salaries, expected hires, and where the job posting was sourced from. The datasets are:

- Company reference list
- Individual user positions
- Individual user on-line profiles
- Individual user education
- Individual user skills
- Lay-offs data
- Unified postings data
- Sentiments and reviews of jobs
- Three sets of dynam data sets by demographics, geography, and role. These contain inflows and outflows.
- I have used population data from 2010 to 2024 from the US census which I use to derive the jobs per capita in the USA (population.csv).

1.2 *Data derived from unified job postings*

I include a script which checks for corrupt data and in the unified data I found one file (postings_unified_individual_0324_part_04.parquet) which cannot be opened. The following is a list of data I used and pulled from the unified dataset:

- Included in the outputs/data file is a data sparsity table which measures the number of job openings per year and state. We find the data is well populated from the 2016 to 2024 (data_sparsity.csv).
- A list of the 1500 job categories which are present in the dataset (jobs_catagories.csv).
- A sample of unified data in csv format for you to look at (sample.csv).
- I also provide a count of total job postings and job postings in AI related fields per capita (jobs_by_year_state_population.csv, ai_per_cap.csv)

Next I outline the scripts I have used and the Illustrate the plots I have created

2. Analysis

I have Four scripts which I have used. The first, is a list of packages. The second, is a corruption check script. The third, counts the total number of job postings per state, and year; and measures these counts on per capita terms. I include the code for the maps and the percent rankings in the third script. The final script is nearly identical to the third, but it filters the total job postings by job categories related to AI. Here is a list of the job categories I choose which I believe to be related to AI:

- Data Scientist
- Data Engineer
- Data Science
- Statistician
- Automation Engineer
- Research
- Researcher
- Research Analyst
- Research Scientist
- Research Engineer
- Research & Development
- Research & Development Engineer
- Scientist
- Scientist I
- Scientific
- Computer Scientist
- Innovation
- Software Engineer
- Software Engineer I
- Software Engineering
- Software Development
- Software Development Engineer
- Software Development Engineer in Test
- Software Engineer Analyst
- Software Engineering Analyst
- Software Architect
- Advisory Software Engineer
- Software Designer
- Software Consultant
- Software Programmer
- Developer
- Programmer
- Programmer Analyst
- Computer Programmer
- SDE
- SDET
- SW Engineer
- Python Developer
- Java Developer
- Java Software Developer
- Java Software Engineer
- Android Developer
- Web Developer
- Web Application Developer
- Frontend Developer
- UI Developer
- .NET Developer
- Embedded Software Engineer
- System Software Engineer
- DevOps
- DevOps Engineer
- Site Reliability Engineer
- Cloud Engineer
- Cloud Architect
- Infrastructure Engineer
- Infrastructure Architect
- Infrastructure
- IT Infrastructure
- Systems Engineer
- Systems Engineering
- System Engineering
- Systems Architect
- System Architect
- Storage Engineer
- Data Center
- Deployment
- ETL Developer
- Database Engineer
- Database Analyst
- Database Developer
- Database Administrator
- Data Architect
- QA Automation Engineer
- Test Automation Engineer

3. Maps

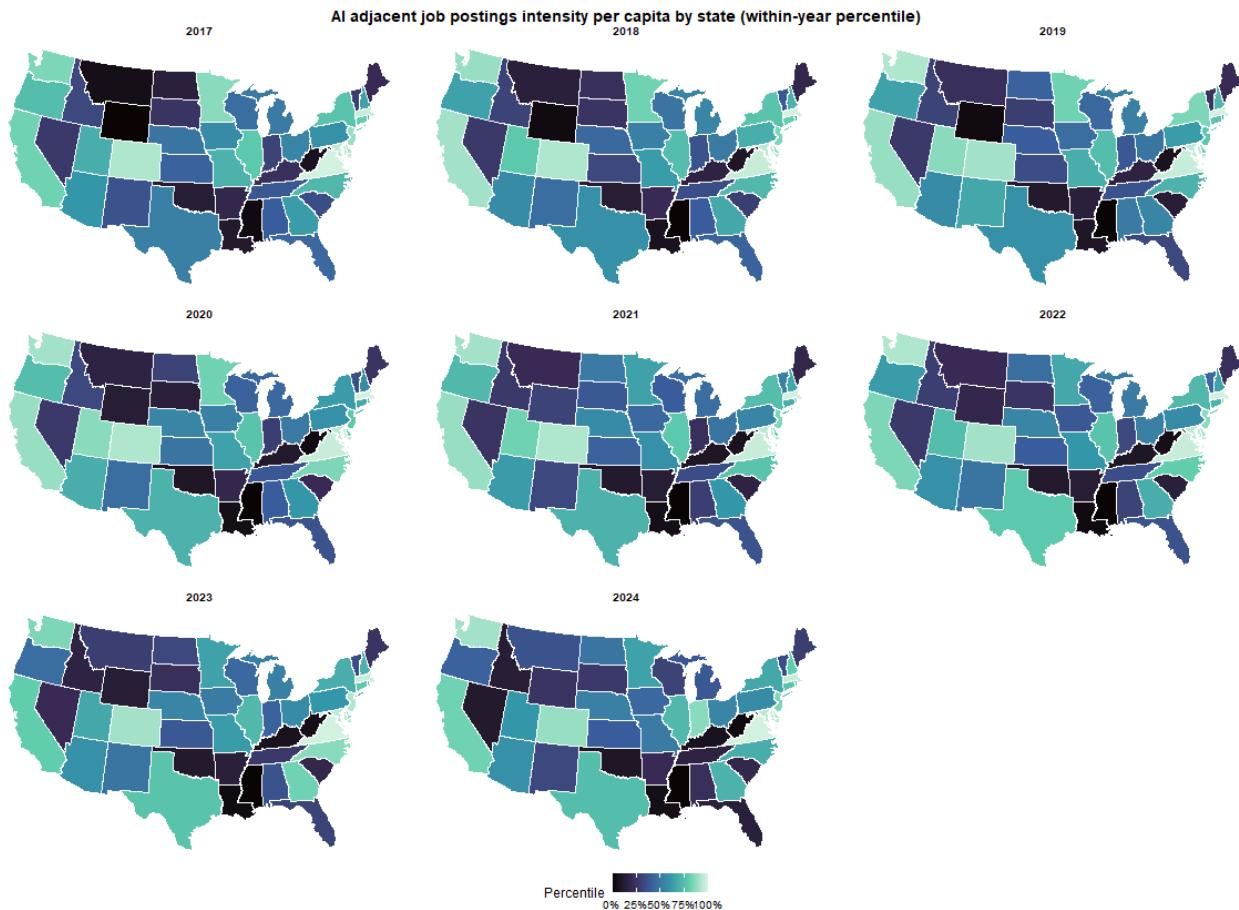


Figure 1: AI Adjacent Job Postings per Capita

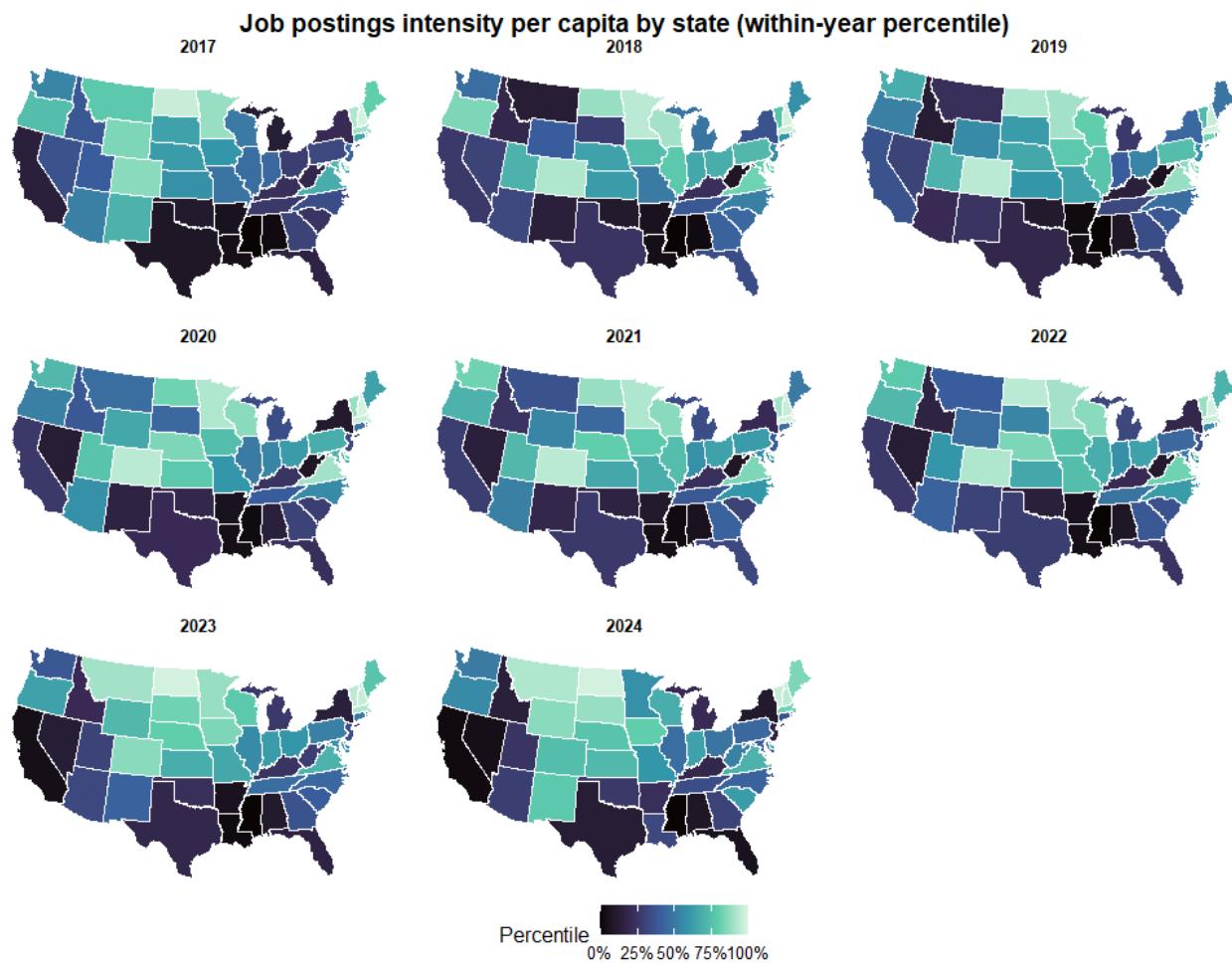


Figure 2: Total Job Postings per Capita