

A Lead Generation Tool for Third-Party Data Demand via U.S. Public Job Postings

03 May 2025

Summary

This project is a tool designed to help identify demand for external data and target potential customers in the public sector using available job postings. The tool analyzes details from U.S. government job postings via USAJobs by first creating structured features, such as seniority, agency size, and industries, and secondly by implementing natural language processing (NLP) techniques to analyze the text in job descriptions to identify positions likely to engage in purchasing external data. Each of these potential data buyers is assigned a Data Buyer Score, which reflects the likelihood that the position is a data-buying one; this feature enables easy prioritization for future outreach and targeting. The tool is fully automated and easy to reuse and modify to target various industries and assess demand in real-time or expand to explore demand seasonalities.

Statement of need

The demand for external data in businesses has grown exponentially across industries. While we know this demand exists, public procurement records are often unavailable or do not identify the roles or departments responsible for purchasing. This lack of transparency leads to market inefficiencies where data vendors cannot reliably connect with or identify their clients.

This project addresses that inefficiency by creating a reproducible framework to estimate this demand through job descriptions in available position postings and outputs a ranked list of job titles and agencies by their likelihood of being data buyers- enabling researchers and data vendors to identify likely clients and then prioritize their outreach based on these scores.