Job Market Analysis Tool: Project Report

Prepared by: Subhankari Panda
Date: May 7, 2025

1 Introduction

The Job Market Analysis Tool is a Python-based application designed to analyze job postings and provide insights into skill demands. Built for job seekers and recruiters, it visualizes skill distributions and recommends job titles and complementary skills. The project evolved from a console-based script to a Streamlit web app, using a dataset (`raw_jobs.csv`) with 19 entries, where the `skills` column was empty, requiring programmatic skill assignment.

2 Objectives

- Process the job postings dataset, addressing the empty `skills` column.
- Create visualizations to understand skill demands.
- Develop a recommendation system for job titles and skills.
- Build an interactive web interface using Streamlit.

3 Methodology

3.1 Data Processing

The dataset was loaded using `pandas`. Skills were assigned programmatically:

- Used a dictionary (`skill_keywords`) to map job titles to skills (e.g., 'software engineer' to 'python', 'java').
- Applied a function to assign skills, with a default 'general' skill for unmatched titles.
- Exploded the data to create one row per skill, cleaned, and removed nulls.

3.2 Features

Implemented using `pandas`, `seaborn`, `matplotlib`, and `streamlit`:

- **Heatmap**: Top skills by job title, with shortened titles.
- **Bar Plot**: Top 10 skills across postings.
- **Pie Chart**: Skill distribution for a job title (e.g., 'Software Engineer': 33.3% 'python', 33.3% 'java').
- **Recommendations**: Job titles (e.g., 'python' yields 'Senior Software Engineer') and complementary skills (e.g., 'python' pairs with 'sql').

The Streamlit app offers a sidebar for navigation, text inputs, and downloadable CSV files.

4 Results

- **Visualizations**: Heatmap shows 'javascript' in developer roles; bar plot highlights 'javascript', 'sql' as top skills; pie chart for 'Software Engineer' shows balanced skills.
- **Recommendations**: 'python' suggests 'Senior Software Engineer' (3 postings); complementary skills include 'sql' (5 co-occurrences).
- **Interface**: Streamlit app runs at `http://localhost:8501`, with outputs in `outputs/` (e.g., `heatmap.png`).

5 Challenges and Solutions

- **Empty Skills**: Assigned skills using keywords.
- **All Locations Remote**: Focused on job titles.
- **Streamlit Errors**: Fixed `ScriptRunContext` by using `streamlit run`.
- **Deprecation Warning**: Resolved `groupby(...).apply` warning with `include_groups=False`.

6 Future Improvements

- Expand dataset for location-based analysis.
- Add skill co-occurrence matrix.
- Deploy on Streamlit Cloud.

7 Conclusion

The Job Market Analysis Tool provides insights into skill demands through visualizations and recommendations. The Streamlit interface enhances usability, making it a valuable tool for job market analysis.