Introduction

The job market is highly dynamic, with new skills rising in demand and traditional skills evolving. Professionals need to adapt quickly to remain competitive, while recruiters must anticipate future requirements. LinkedIn, as the world's largest professional networking site, provides valuable job posting data that can reveal these trends.

This project leverages web scraping and data analysis techniques to examine LinkedIn job postings, highlighting which skills are most sought after in different cities and how they correlate with specific job roles. The goal is to offer a data-driven approach for decision-making in career planning, recruitment, and workforce development.

Abstract

In the digital job market, identifying emerging skill demands and analyzing location-specific job trends is critical for professionals and recruiters alike. This project focuses on scraping LinkedIn job postings using Python and analyzing them to uncover demand patterns in various cities and roles. By extracting job titles, skills, and locations, the data is processed into structured form and used to generate insights such as top skill demand by city and skill–role relationships. The results are visualized using heatmaps to provide an intuitive understanding of where specific skills are in demand and how they align with different roles. The outcome of this project is a comprehensive skill–demand analysis that can guide job seekers, employers, and training providers.

Tools Used

1. **Python** – Primary programming language.
2. **BeautifulSoup (bs4)** – Used for scraping job postings from LinkedIn HTML pages.
3. **Pandas** – For data cleaning, wrangling, and analysis of extracted job data.
4. **Matplotlib & Seaborn** – For data visualization, creating heatmaps and trend graphs.
5. **Excel (Optional)** – For storing and exporting the cleaned datasets for further manual exploration.

Steps Involved in Building the Project

1. **Data Collection (Web Scraping)**
    - Using Python requests and BeautifulSoup, job postings are extracted from

LinkedIn search results.

- Extracted features include job title, city/location, and required skill tags.

2. **Data Cleaning & Preprocessing**
    - Raw scraped data often contains duplicates, HTML tags, or incomplete records.
    - Skills are parsed into structured format (lists) for easier aggregation.
    - Missing values are handled, and irrelevant entries are removed.
3. **Data Analysis**
    - Frequency counts of skills are calculated for each city.
    - A pivot table is created to summarize the top 10 skills across cities.
    - Another matrix is prepared showing the relationship between skills and roles.
4. **Visualization**
    - Heatmap of **Top 10 Skills by City** to illustrate geographic distribution of demand.
    - Heatmap of **Skills vs Roles** to highlight skill requirements for various job roles.
    - These visuals provide a clear view of both location-based and role-based skill trends.
5. **Trend Recommendation**
    - Identify cities with high demand for specific skills (e.g., Python in Bangalore, Data Analysis in Delhi).
    - Highlight skill gaps where certain roles consistently require emerging skills.

Conclusion

This project demonstrates how web scraping and data visualization can transform unstructured LinkedIn job postings into actionable insights. By analyzing skills across cities and roles, we uncover demand patterns that are valuable for professionals seeking to upskill, recruiters planning hiring strategies, and training institutes designing courses.

The results reveal that job demand is not uniform—certain skills dominate in specific cities, while others are linked to particular job roles. The approach used in this project is flexible and can be extended to track changes in skill demand over time, providing a powerful tool for future workforce analysis.