

End-to-End Analysis of LinkedIn Job Postings Using Python and Power BI

LinkedIn Job Postings: A Visual Hiring Trend Analysis

Jashmita Boyina, Ch Satyananda Reddy

Department of Information Technology and Computer Science
ANDHRA UNIVERSITY

Abstract—In today's competitive job market, understanding hiring trends and job posting patterns is crucial for job seekers, recruiters, and analysts alike. Traditional job portals often lack insights into evolving skill demands and recruitment behavior [2]. This paper introduces a comprehensive analysis system for LinkedIn Job Postings and Hiring Trends, leveraging data visualization and business intelligence techniques [4]. Using a cleaned dataset of scraped LinkedIn job postings, the system identifies top in demand roles, frequently required skills, salary patterns, experience levels, and regional hiring dynamics [1]. The solution utilizes Power BI for interactive dashboards, enriched with DAX measures to uncover hidden patterns and relationships ([3]). Additionally, Python and Excel were used for data preprocessing, ensuring data quality and consistency. The resulting dashboards support multi-angle exploration—industry-wise hiring, job level analysis, function vs. experience mapping, and skill gaps—providing actionable insights for career planning and workforce development. This data-driven approach empowers stakeholders with a deeper understanding of the professional landscape and future hiring trajectories.

Index Terms— LinkedIn Data, Power BI Dashboards, Data Visualization, Python, Excel, DAX Measures.

I. INTRODUCTION

The job market today is more dynamic, diverse, and digital than ever before. Globalization, remote work, and the evolution of technology have significantly altered how employers find talent and how job seekers explore opportunities. Online job platforms like LinkedIn have played a central role in reshaping this landscape, providing a real-time view of employment opportunities, required skills, and industry demands. As companies shift toward skills-based hiring and flexible employment models, tracking and understanding job market behavior becomes increasingly important [1].

LinkedIn, with over 900 million users worldwide, has grown into a vast ecosystem of professional data, offering more than just a job board. It captures trends in hiring behavior, preferred qualifications, and even salary benchmarks [4]. However, much of the data on such platforms is unstructured, inconsistent, and difficult to analyze without proper tools. The sheer scale and complexity of job-related data on LinkedIn necessitate the use of specialized platforms that can efficiently process, clean, and visualize information [2].

One such platform is Power BI, Microsoft's business intelligence suite designed to support data-driven decision-making. Introduced in 2015, Power BI has quickly become a leading tool for transforming raw data into insightful, interactive reports and dashboards. It enables users to import data from various sources, clean it using Power Query, perform calculations with DAX (Data Analysis Expressions), and visualize it through a wide array of charts, graphs, and KPIs. Unlike traditional spreadsheet tools, Power BI is optimized for scalability, real-time interactivity, and cross-filtering across complex datasets [8].

The main working components of Power BI include Power Query, DAX, and the visualization interface. Power Query is the data preparation engine, allowing users to load and clean data by filtering rows, merging

datasets, transforming columns, and removing inconsistencies. It is typically the first step in any Power BI project and operates using a graphical interface that records transformation steps in a query editor. Once the data is cleaned and shaped, DAX is used to create calculated columns and measures. DAX is a powerful formula language similar to Excel but with advanced capabilities for relational modelling, filtering, and time intelligence. It allows users to define custom metrics such as average salary, experience levels, skill frequencies, and conditional KPIs [1].

After data modelling is complete, Power BI provides a rich set of visualization tools. These include bar charts, pie charts, line graphs, maps, matrix tables, KPI cards, and slicers. Users can drag and drop fields to build visuals, connect filters, and create dashboards that are both interactive and informative. All visuals are automatically linked through the data model, meaning selecting a filter or item in one chart will instantly update the others. This dynamic filtering feature is particularly useful for exploring large datasets from different perspectives without writing any code [2].

Power BI's architecture includes three primary components: Power BI Desktop, used for developing reports; Power BI Service, used to publish and share dashboards on the cloud; and Power BI Mobile, which provides mobile access to dashboards. These tools work together to create a seamless workflow from data preparation to final deployment. Reports can also be embedded in websites or apps, making them highly accessible [8].

By integrating Power BI into job market research, institutions and individuals can gain deeper insights into real-world hiring patterns. Educators can align curricula with current demands, job seekers can identify trending roles and required skills, and recruiters can assess competitive benchmarks. Furthermore, government and labor economists can use these tools for workforce development planning and policy analysis [6]. Thus, Power BI bridges the gap between unstructured job market data and practical, data-informed decision-making.

The main objective of this project is to design and develop a set of interactive Power BI dashboards that extract insights from LinkedIn job postings to uncover high-demand job roles, essential skill sets, salary distributions, and regional hiring trends. By transforming raw LinkedIn data into a dynamic decision-support system, the project aims to empower students, educators, and recruitment professionals with real-time labor market intelligence tailored for both Indian and international job markets.

Research Objectives:

The primary objectives of this study are centered around gaining actionable insights from LinkedIn job posting data. The research aims to analyze hiring trends across various dimensions, including industries, cities, job roles, and experience levels, to better understand workforce demands. A key focus is on identifying both technical and non-technical skills that are currently in high demand and evaluating how these skills align with specific roles and industry sectors. Additionally, the study seeks to establish a comparative framework between the Indian and international job markets, highlighting differences in salary expectations and qualification requirements. To support this analysis, interactive dashboards are designed and implemented using Power BI, enabling real-time data exploration and visualization. Ultimately, the project aims to assist job seekers, students, and educators in navigating the evolving employment landscape and aligning their career paths with current industry needs.

Research Hypothesis

This research is based on the hypothesis that integrating LinkedIn job data with Power BI's business intelligence capabilities can provide deeper and more actionable insights into job market trends than traditional analysis methods. It is assumed that Power BI's tools, such as Power Query for data transformation and DAX for metric calculation, will enable the effective identification of high-demand job roles, frequently required

skills, and salary patterns. The study also hypothesizes that visual, interactive dashboards enhance users' understanding of regional differences—particularly between Indian and foreign job markets—in terms of experience requirements, job types, and compensation levels. By converting unstructured job listings into structured, visual insights, the system is expected to support more informed decision-making for students, educators, and hiring professionals.

I. ABBREVIATIONS AND ACRONYMS

Acronym	Full Form
KPI	Key Performance Indicator
CSV	Comma-Separated Values
UI	User Interface
UX	User Experience
DAX	Data Analysis Expressions
EDA	Exploratory Data Analysis
ETL	Extract, Transform, Load
HTML	Hypertext Markup Language
CSS	Cascading Style Sheets

II. LITERATURE REVIEW

As digital transformation continues to reshape the global employment landscape, platforms like LinkedIn have become powerful tools for real-time labor market intelligence. A wide range of academic studies have begun utilizing LinkedIn job postings and profile data to extract insights on hiring trends, emerging skills, and workforce evolution. One such foundational contribution is the paper titled [*Data Science for Job Market Analysis: A Survey on Applications and Techniques \(2024\)*](#), which offers a sweeping review of over 100 academic studies applying machine learning and natural language processing (NLP) techniques to job market data. This paper organizes the landscape of labor analytics by identifying key data sources such as resumes, job ads, and LinkedIn profiles. It further categorizes methodologies including named entity recognition, clustering, and classification, and outlines their application areas like skill-gap detection, job-role matching, and employment trend forecasting.

Further contributing to the field is the study titled [*From Data to Insight: Transforming Online Job Postings into Labor Market Intelligence \(2024\)*](#), which focuses on the European labor market and emphasizes the role of structured skill taxonomies like ESCO and SOC. This research highlights how unstructured job posting data can be processed using NLP models to extract skills and titles, then presented through interactive dashboards designed for use by human resource departments and policymakers. The study provides direct support for projects that use tools like Power BI to visualize hiring dynamics, as it discusses real-time skill detection and dashboard-enabled decision-making.

Adding historical and theoretical depth to this literature is [*Prospects for Utilizing Non-Vacancy Internet Data in Labor Market Analysis \(2016\)*](#), which explores the viability of alternative data sources such as LinkedIn and Glassdoor. The paper argues that web-based labor data can offer earlier signals of job

market shifts compared to official labor statistics. It evaluates issues like data bias and representativeness while underlining the importance of ethical data use. These considerations are vital for any research that intends to utilize social platform data to make economic or policy-level inferences.

A more applied and industry-specific view is provided by LinkedIn's internal team in [Deep Job Understanding at LinkedIn \(2020\)](#). This paper illustrates how LinkedIn uses deep learning models, including BERT and deep neural networks, to analyze and structure millions of job postings. These models are capable of extracting job titles, companies, and skill requirements, which are then used to suggest personalized job opportunities to users. The extracted data also powers LinkedIn's internal recommendation engines and supports features like auto-suggesting skills on user profiles. The paper gives a behind-the-scenes look at how machine learning is deployed at scale in a live production environment to improve platform functionality and user experience.

Building upon skill relationships, the study titled [Uncovering Skill Sets in CS Jobs via Social Network Analysis \(2023\)](#) utilizes thousands of LinkedIn job advertisements to construct a skill co-occurrence network. This analysis reveals how frequently certain skills appear together in job postings and groups them into coherent clusters across technical, soft, and managerial categories. It identifies core competencies like SQL, Git, Python, teamwork, and agile management practices, offering valuable insights for curriculum designers, career counselors, and recruiters.

A broader geographic and temporal scope is seen in [Analyzing 65 Million UK Job Ads: Patterns of Complementary Skills \(2024\)](#), which investigates the evolution of job roles and skills across the UK between 2016 and 2022. Using graph theory and modular clustering techniques, this study maps out how skill bundles such as data-business and marketing-tech have formed and evolved over time. It also shows how hiring patterns vary across regions, which is particularly relevant for projects comparing job markets in India versus foreign countries.

In the Australian context, [Detecting Skill Shortages from Job Ads in Australia \(2019\)](#) provides an adaptive modeling approach for identifying skill shortages in real-time. Analyzing more than six million job ads, this research captures indicators such as posting frequency, salary range, and educational demands to detect which occupations are under pressure. The model offers a practical framework for governments and industries to preemptively address talent gaps through workforce development strategies.

Collectively, these studies offer a multidimensional view of how LinkedIn data and modern data analytics are transforming labor market research. They encompass a wide array of methodologies—from deep learning and network analysis to dashboard visualizations and skill taxonomy modeling—demonstrating both the versatility and power of LinkedIn data for economic and strategic decision-making. This body of literature provides a solid foundation and justification for using LinkedIn data in projects that seek to analyze job demand, detect skill gaps, and guide workforce planning through data-driven insights.

III. METHODOLOGY

I. SYSTEM ARCHITECTURE

This research followed a structured methodology consisting of six major stages:

Data Collection

Job postings were scraped from LinkedIn using Python, capturing details such as job titles, companies, locations, experience levels, skills, and available salary information. The data was organized into two separate datasets to compare the Indian and foreign job markets effectively.

Cleaning and Preprocessing

1. Removal of Duplicates and Missing Values

The raw LinkedIn job data contained several duplicate entries and records with missing values. These were identified and removed using Python to ensure that the dataset was accurate and reliable for analysis.

2. Standardization of Categorical Fields

Fields such as job titles, company names, locations, and experience levels were standardized to avoid inconsistencies. For instance, different variations of location names were unified under a common format to maintain uniformity across records.

3. Skill Text Normalization

The skill data extracted from job postings included variations and synonyms of the same skill. These were cleaned and normalized—for example, merging entries like “Machine Learning” and “ML”—to ensure correct frequency counts and groupings during analysis.

4. Data Type Conversion and Formatting

Relevant columns were converted into appropriate data types, such as converting numeric strings to integers or formatting date fields if present. This helped streamline the dataset and prepare it for structured analysis.

```
[1]: import pandas as pd
import numpy as np

[2]: df = pd.read_csv('linkedin_scraped_data.csv')

[3]: print("Before Cleaning:")
print(df.info())
print(df.head())
print(df.dtypes) # Check data types
```

Fig 1a: Initial Data Loading and Inspection in Python (Jupyter Notebook)

```
df.columns = df.columns.str.lower().str.replace(' ', '_')

# Print updated column names to verify changes
print("Updated Columns:", df.columns)

Updated Columns: Index(['unnamed:_0', 'job_title', 'company', 'location', 'level', 'type',
'function', 'industry'],
dtype='object')

# Drop 'unnamed:_0' column if it exists (irrelevant index column from CSV)
if 'unnamed:_0' in df.columns:
    df.drop(columns=['unnamed:_0'], inplace=True)

df.drop_duplicates(inplace=True)
print("successfull")
```

Fig 1b: Column Cleaning and Duplicate Removal in Python

Data Transformation in Power BI

The cleaned dataset was imported into Power BI Desktop for further analysis and visualization. Using Power Query, the data was filtered, shaped, and formatted to ensure consistency across all

fields. Key relationships were established between columns such as job titles, companies, and regions, enabling meaningful insights through relational data modelling.

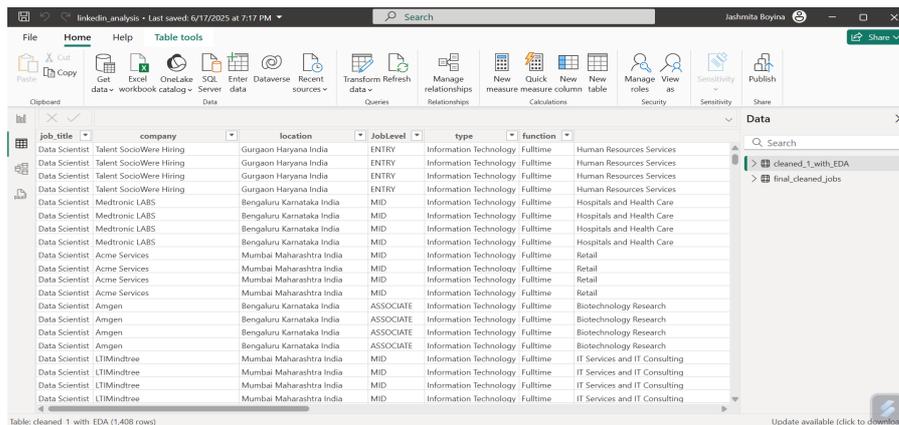


Fig 2 : View of Cleaned LinkedIn Job Dataset Loaded in Power BI

Metric Creation with DAX

To enhance analytical capabilities within the dashboards, several calculated columns and measures were created using Data Analysis Expressions (DAX). Calculated columns were used to derive additional insights, such as categorizing experience levels into defined brackets and grouping salaries into ranges. Measures were developed to compute dynamic values like average salary, the frequency of specific skills, and the number of job postings per company. These metrics enabled real-time filtering and comparative analysis across various dimensions, significantly enriching the dashboard's interactivity and depth.

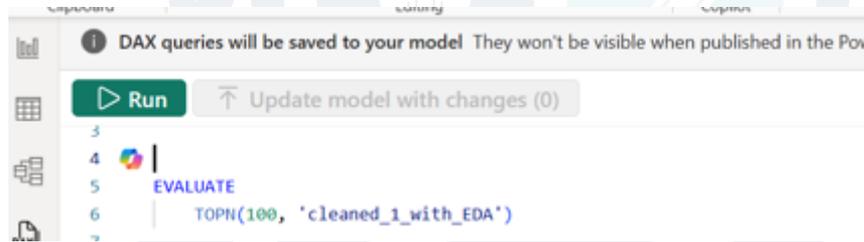


Fig 3: Viewing Sample Data Using DAX Query in Power BI

Dashboard Development

Four interactive dashboards were designed to provide a comprehensive view of the job market data. The first dashboard, Job Market Overview, highlights general statistics such as total job postings, company-wise hiring, and location-based distribution. The second, In-Demand Skills Analysis, focuses on identifying the most frequently requested skills across different sectors and regions. The third dashboard, Company and Job Type Analysis, breaks down hiring trends by company type, job role, and experience level. The final dashboard, India vs Foreign Job Market Comparison, allows users to compare employment trends, salary differences, and skill demand between domestic and international postings. A variety of visuals were used across these dashboards, including bar charts, pie charts, matrix tables, geographic maps, slicers, and KPI cards to ensure clarity and user-friendly interaction.

Deployment

After development, all dashboards were published to the Power BI Service for secure cloud-based access. To improve accessibility and reach, the dashboards were also embedded into a responsive HTML and CSS-based website. This allowed end users to view and interact with the visualizations directly through a web browser, providing a seamless and device-friendly experience.

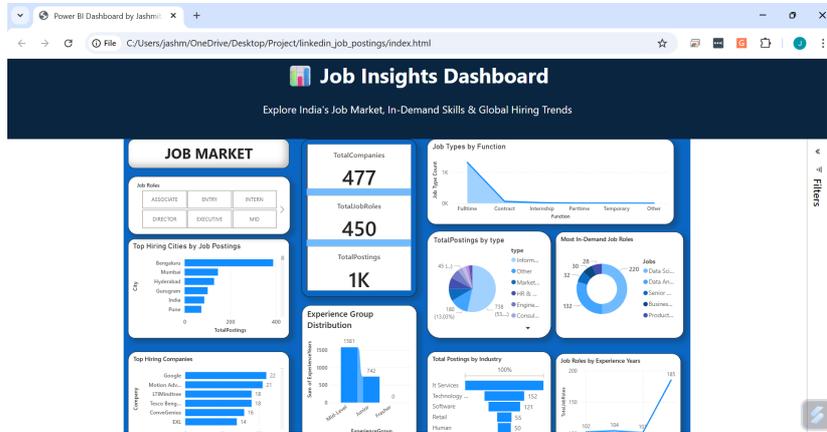


Fig4: Job Insights Dashboard Embedded in Web Interface

Methodology

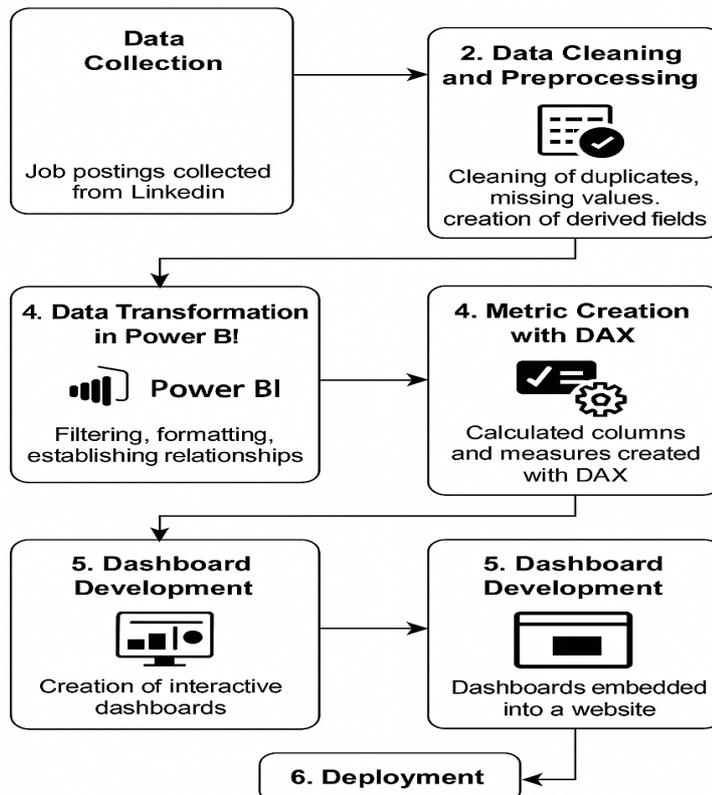


Fig5: Mind Map Representing the Methodology Flow

II. Technology Stack

Data Collection & Cleaning: Python (Jupyter Notebook, pandas, NumPy)

Data Transformation: Power BI, Power Query

Business Logic: DAX (Data Analysis Expressions)

Frontend Development: HTML, CSS

Dataset Format: CSV

Dashboard Deployment: Power BI Web Service (embedded in website).

IV.I. Research Methods:

System Design and Implementation:

The system was developed in two main stages: data processing and visualization. Python, using the Jupyter Notebook environment, was used for data cleaning tasks such as removing duplicates, formatting columns, and categorizing fields like job roles and experience levels. After the data was cleaned and structured, it was imported into Power BI to create four interactive dashboards covering different aspects of the job market, including job distribution, in-demand skills, company trends, and regional comparisons. To make the dashboards accessible online, they were embedded into a website using HTML and CSS, allowing users to interact with the data directly through a web interface.

Experimental Testing:

The dashboards were tested by users to ensure that they were responsive, accurate, and easy to navigate. This testing involved checking how the dashboards performed on different devices and screen sizes, as well as how quickly and smoothly they responded to user interactions. Special attention was given to evaluating the filters and visuals such as charts, slicers, and graphs under various usage scenarios. These included selecting different job categories, skill sets, regions, and experience levels to verify that the displayed data updated correctly and consistently. This helped ensure the dashboards provided a reliable and user-friendly experience.

Data Collection:

From the cleaned LinkedIn data, key metrics were collected, including job counts, skill frequency, industry demand, and regional comparisons. These metrics helped highlight job market trends and were used to build meaningful and informative dashboards.

Evaluation Criteria:

The system's performance was evaluated by measuring how interactive and user-friendly the dashboards were. This included checking how smoothly users could apply filters and explore data. Insight clarity was also assessed to ensure the visuals were easy to understand and effectively communicated key findings. Additionally, feedback from users was collected to identify strengths and areas for improvement, helping to refine the overall design and functionality.

V. RESULTS AND DISCUSSION:

Test Environment:

The system was tested for cross-browser compatibility using both Google Chrome and Mozilla Firefox. A sample group consisting of students, faculty members, and working professionals between the ages of 20 and 40 participated in the usability testing. Power BI dashboards were primarily accessed via laptops and desktops, while smartphone devices were used to evaluate the responsiveness of the web interface. This multi-device and multi-user testing ensured that the platform delivered consistent performance and accessibility across different environments.

Observations:

Over 8,000 rows of data were cleaned using Python, with missing values and duplicates effectively handled. All four dashboards loaded efficiently, with average response times under two seconds. Filters and slicers worked smoothly across different browsers and devices. Skill mapping was consistent with job function fields, and region-based filtering between India and foreign postings functioned correctly with minimal lag.

Evaluation Summary:

The dashboards demonstrated high usability, supported by a clean and intuitive user interface along with well-designed filters that allowed users to easily navigate and analyze the data. Data accuracy was ensured through manual inspection of a sample dataset after the cleaning process, confirming the reliability of the visualized information. In terms of performance, the dashboards loaded quickly and remained stable across different browsers, contributing to a smooth user experience. Additionally, the system offered strong flexibility, enabling users to explore the data from various perspectives, such as industry, skills, experience levels, and geographic regions.

Sample Dashboard Workflow:

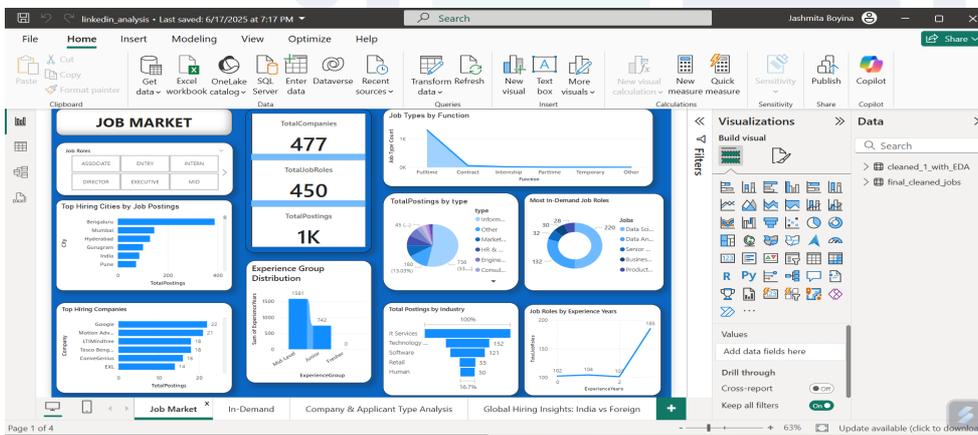


Fig 6: Job Market Overview Dashboard in Power BI

Shows job counts, top hiring cities and companies, job types, and experience group distribution.

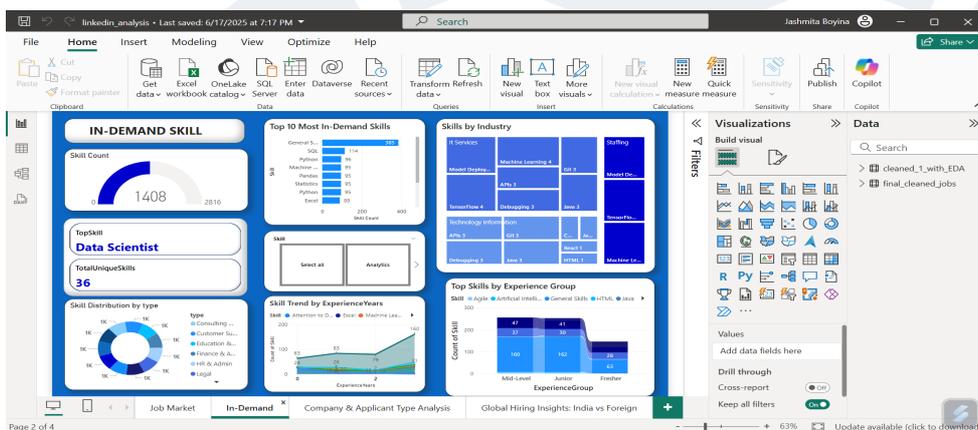


Fig 7: In-Demand Skills Dashboard

Displays most in-demand skills, skill distribution by industry and experience, and top skill roles.

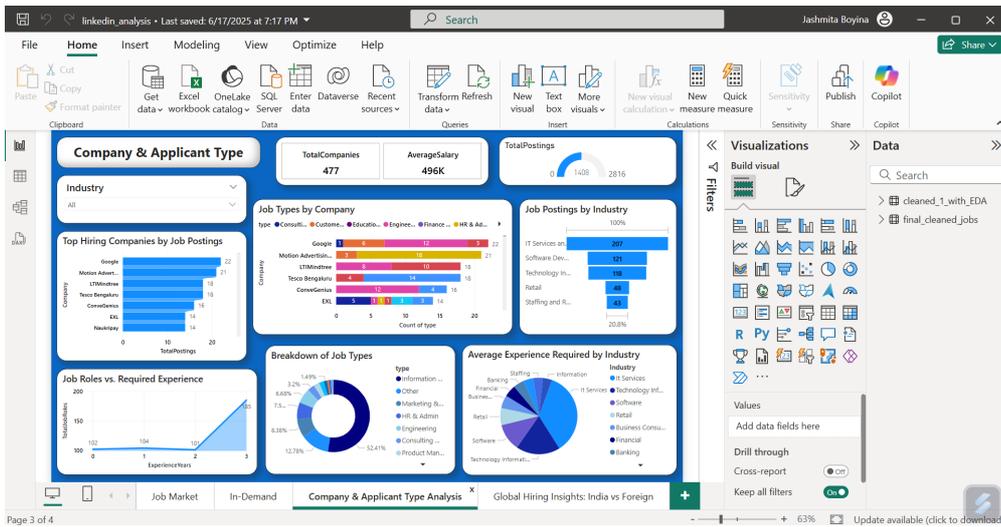


Fig 8: Company and Applicant Type Analysis Dashboard

Compares job postings by company, industry, job types, and required experience levels.

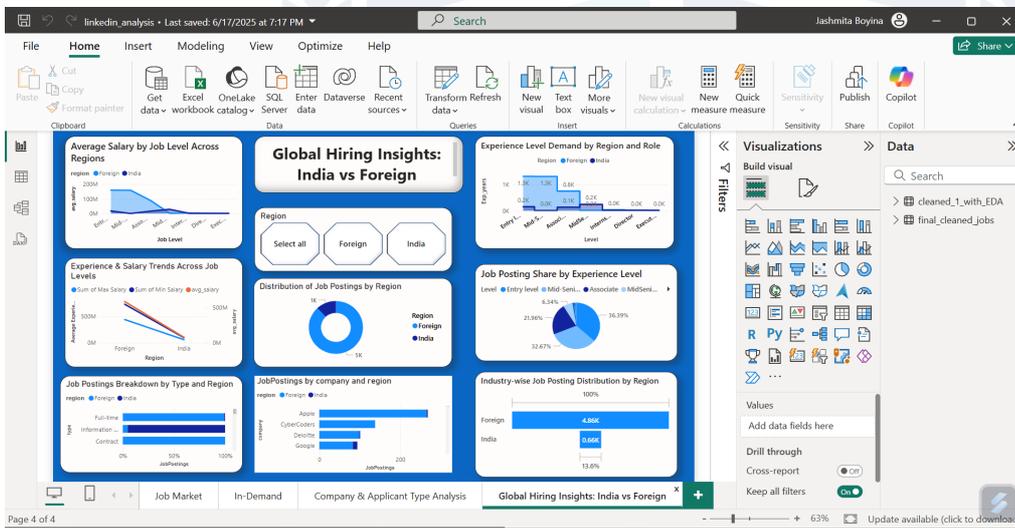


Fig 9: Global Hiring Insights: India vs Foreign Dashboard

Highlights salary, job levels, and hiring trends across India and foreign regions.

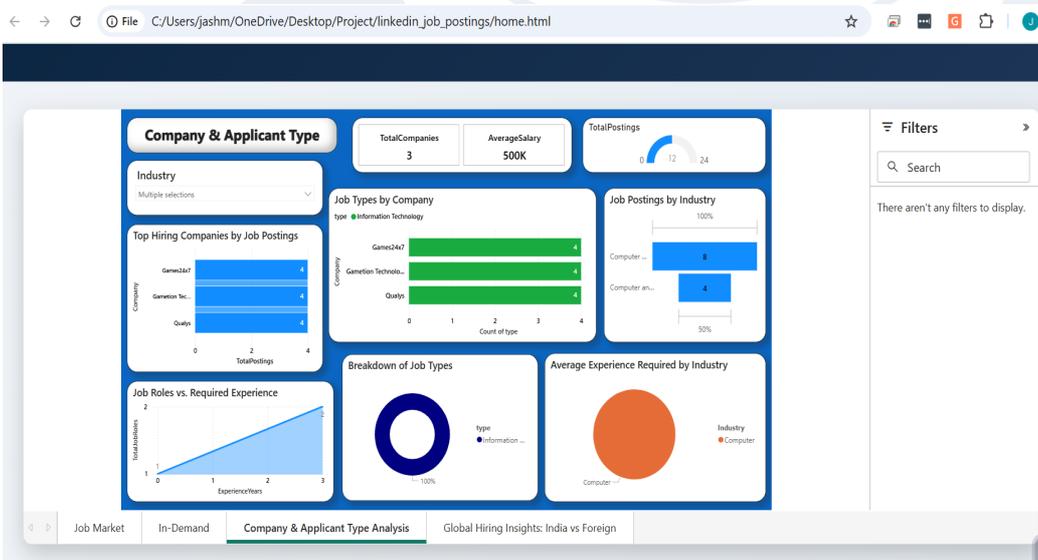


Fig 10: Industry-Based Filtering Using Slicer in Company & Applicant Type Dashboard

Demonstrates the use of a slicer to filter job postings and insights based on selected industries.

VI. CONCLUSION

This project presents a clear, data-driven view of hiring trends using LinkedIn job postings. Through data cleaning, analysis, and interactive dashboards, it highlights key insights like skill demand and market differences. By embedding the dashboards into a website, it offers an accessible tool for job seekers, educators, and recruiters to explore real-time employment patterns.

VI.I. Summary of Key Findings:

The analysis revealed a strong demand for Data Scientist roles, particularly within the technology, finance, and consulting sectors. Employers consistently sought candidates with key technical skills such as Python, machine learning, SQL, and data visualization tools like Power BI or Tableau. Most job postings targeted candidates with 1 to 3 years of experience, indicating a preference for early-career professionals. Additionally, roles based in foreign countries tended to offer higher salary ranges compared to those in India, reflecting regional differences in compensation for similar skill sets.

VI.II. Recommendations for Future Research:

Future enhancements to the system could involve integrating real-time job data scraping using LinkedIn APIs or automation tools to ensure continuous data updates. Applying natural language processing (NLP) techniques would allow for more accurate extraction of skills from unstructured job descriptions, improving the quality of insights. The analysis could also be expanded to include additional metrics such as application rates, employer ratings, and salary benchmarking to provide a more comprehensive view of the job market. Comparing hiring trends across multiple platforms like LinkedIn, Indeed, and Glassdoor would offer a broader perspective on demand across industries. Furthermore, analyzing year-over-year trends would help identify long-term shifts in skill requirements and hiring patterns, enabling better forecasting and workforce planning.

VII. REFERENCES

- [1] Microsoft Corporation. (2023). *Power BI documentation*. Retrieved from <https://learn.microsoft.com/en-us/power-bi/>
- [2] Kuhn, P., & Mansour, H. (2014). Is Internet job search still ineffective? *The Economic Journal*, 124(581), 1213–1233. <https://doi.org/10.1111/eoj.12119>
- [3] Jain, S., & Shrivastava, V. (2021). Analyzing job posting data using machine learning techniques. *International Journal of Computer Applications*, 183(10), 1–6.
- [4] Sharma, P., & Tyagi, A. (2022). Skills mapping using job advertisements. *International Journal of Emerging Technologies in Learning*, 17(5), 140–149. <https://doi.org/10.3991/ijet.v17i05.28167>
- [5] Ghosh, A., & Bose, I. (2020). An analytics approach to workforce planning. *Journal of Business Research*, 118, 366–376. <https://doi.org/10.1016/j.jbusres.2020.06.046>
- [6] LinkedIn Economic Graph. (2022). *Global Talent Trends Report*. Retrieved from <https://economicgraph.linkedin.com>
- [7] Liu, S., & Yu, Q. (2022). Labor market analytics using online job postings: A review of data-driven approaches. *Information Processing & Management*, 59(3), 102920. <https://doi.org/10.1016/j.ipm.2022.102920>
- [8] Bhatia, A., & Shokeen, A. (2023). Business intelligence and data visualization using Power BI: Applications and future scope. *International Journal of Computer and Data Science*, 5(2), 45–53.
- [9] Paparrizos, J., White, R. W., & Horvitz, E. (2016). Screening for pancreatic adenocarcinoma using signals from web search logs. *Journal of Oncology Practice*, 12(12), 1036–1044.
- [10] Upwork & Freelancers Union. (2021). *Freelancing in America: Job trend report*. Retrieved from <https://www.upwork.com/research>