Analyzing UK Data Science Job Market Trends through LinkedIn Job Postings for April & May 2025

Md Ashraful Alam Gazi
Umais Bhatti
University of Essex

May 2025

Abstract

Understanding real-time job market dynamics is essential for students and early-career professionals navigating competitive fields like data science. This study investigates data science job trends in the United Kingdom by analyzing job postings scraped from LinkedIn during April and May. The aim is to uncover hiring patterns, skill requirements, and geographical demand to support more informed career planning and project selection.

Due to LinkedIn's platform restrictions and limited public access to historical job data, job postings can only be reliably scraped within a narrow window—typically one or two months. As a result, this research focuses on a recent two-month period, providing a timely snapshot of current job market behavior rather than a long-term trend analysis.

Using Python tools such as Selenium and BeautifulSoup, data was collected from LinkedIn and processed using pandas, matplotlib, seaborn, and plotly. The dataset included job titles, locations, companies, post dates, and descriptions, enabling analysis across skills, experience levels, and hiring hubs.

The results show that R, Python, and SQL were among the most frequently mentioned technical skills, with machine learning and cloud expertise also in demand. London emerged as the top hiring location. Most roles favored candidates with postgraduate degrees and professional experience. This research provides valuable guidance for data science job seekers, helping them align their skills and portfolios with current employer expectations.

1   Introduction

In recent years, there has been a growing interest in data-driven approaches to career planning, especially in highly technical and competitive fields like data science. As the demand for skilled professionals continues to rise, students and job seekers are increasingly turning

to online platforms, real-time data, and analytics to guide their career development and job search strategies. Platforms like LinkedIn, which host millions of job postings and professional profiles, offer a rich—yet often underutilized—source of labor market insights.

Understanding hiring trends on LinkedIn can provide timely and practical information about the skills employers are looking for, which companies are actively recruiting, and where opportunities are concentrated geographically. For students and early-career professionals, this type of insight is especially useful in shaping academic choices, personal projects, and upskilling decisions.

However, due to limitations in LinkedIn's public data access, only recent job postings—usually within the past one or two months—are available for scraping and analysis. As a result, this study focuses on a short timescale (April and May) to provide an accurate and real-time snapshot of the UK data science job market.

The main objectives of this research are:

- To analyze seasonal and recent job posting trends in the UK

- To identify the most frequently demanded skills and qualifications

- To understand typical education and experience requirements

- To highlight top hiring companies and dominant industries

These objectives guide the core research questions:

- Which skills are currently most in-demand in UK data science job postings?

- What levels of education and experience are typically required?

- During which periods do most job postings appear, and how do they vary by location or role?

By answering these questions, the study aims to offer actionable insights to help aspiring data professionals make informed decisions in their job search and career planning.

2   Literature Review

2.1   Overview of Labor Market Analytics

Labor market analytics has become an essential area of research for understanding employment trends, skill demands, and industry shifts. Traditional labor statistics from government bodies provide macro-level insights, but they often lack granularity and timeliness. With the rise of digital job platforms, researchers are increasingly using online job postings to gain real-time, fine-grained insights into labor market dynamics.

Recent literature highlights how job postings reflect not only current demand but also evolving industry requirements. For example, data science and AI roles have seen exponential growth, prompting studies to track the emergence of new roles and hybrid skill sets.

2.2 Prior Work Using Online Job Postings

Several studies have explored labor market trends using data from platforms like Indeed, Adzuna, and LinkedIn. **(author?)** (1) analyzed 65 million UK job adverts from Adzuna, identifying 21 skill clusters and showing how employers increasingly seek multi-domain competencies. Similarly, **?**) used data from Reed.co.uk to assess how COVID-19 impacted sector-wise job postings, revealing a 60–70% decline in vacancies during lockdowns, with sharp differences across sectors such as hospitality vs healthcare.

LinkedIn, although less explored in academia due to scraping restrictions, has been studied internally by **(author?)** (4), who developed deep learning models to extract structured entities (like skills, companies, and job titles) from unstructured job postings, significantly improving recommendation systems on the platform.

2.3 Use of Web Scraping and Text Mining in HR Analytics

Web scraping and text mining are now widely used in human resource analytics. ISTAT (the Italian National Statistics Institute) demonstrated how scraping company websites can supplement ICT-related enterprise surveys. Skill extraction has also advanced through NLP techniques. **(author?)** (2) proposed a weak supervision method using the ESCO taxonomy to extract skills from job ads without human annotation, enabling scalable and flexible labor market insights.

These approaches show how modern techniques can uncover both technical and soft skill requirements, analyze job descriptions at scale, and even track how industries evolve over time.

2.4 Gaps Addressed by This Study

While existing work provides valuable contributions, several gaps remain:

- Most prior studies use broad, multi-year datasets, lacking localized, real-time analysis.

- Few studies focus specifically on LinkedIn, which is increasingly a primary platform for data-driven hiring.

- There is limited research offering UK-specific insights for data science roles, especially over short time windows.

- Many works analyze skill frequency but do not combine it with job titles, companies, locations, and education levels in a single study.

This research fills those gaps by:

- Scraping and analyzing UK data science job postings from LinkedIn in April–May 2025

- Extracting and visualizing skills, education levels, and company types together

- Providing a fine-grained snapshot of current hiring patterns, ideal for students and job seekers planning their next steps

3 Methodology

3.1 Data Collection

The dataset used in this study was collected from LinkedIn, one of the most widely used professional job platforms globally. Given LinkedIn's restrictions on public data access and scraping, only recent job postings—typically from the last one or two months—were available. Therefore, the focus was on job postings published during April and May 2025, offering a real-time snapshot of the current job market.

Data was scraped using Selenium for browser automation and BeautifulSoup for HTML parsing. The job search was conducted using targeted keywords such as *"Data Scientist"*, *"Machine Learning Engineer"*, *"AI Specialist"*, and related roles, filtered by United Kingdom as the geographic region. Data was stored and processed using `pandas`, a powerful data manipulation library in Python.

3.2 Data Preprocessing

After collection, the dataset was cleaned and structured to prepare it for analysis. The following preprocessing steps were applied:

- **Job Titles** were standardized to reduce redundancy (e.g., "Senior Data Scientist" and "Data Scientist" were grouped).

- **Posting Dates** were parsed and converted into `datetime` format to support time-based analysis.

- **Company Names** were extracted and used to group postings by employer.

- **Locations** were parsed to analyze city-wise trends.

- **Job Descriptions** were retained for keyword-based skill extraction.

3.3 Skill Extraction

To identify technical skill demand across job postings, a keyword-based matching approach was implemented. A curated list of common data science skills was compiled, including:

- **Programming:** Python, R

- **Data Technologies:** SQL, Excel

- **Machine Learning Tools:** TensorFlow, scikit-learn

- **Cloud Platforms:** AWS, Azure, GCP

- **Visualization:** Power BI, Tableau

- **Other:** Big Data, Statistics, NLP

These keywords were searched within job titles and descriptions to map each job to one or more skills.
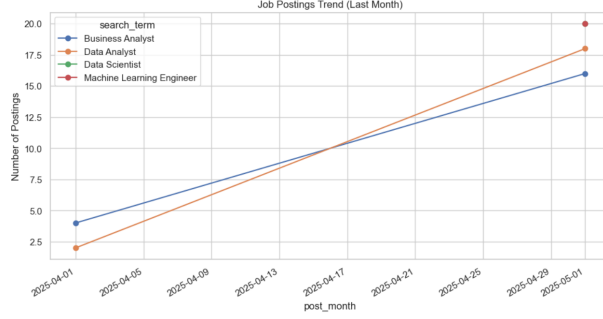
Figure 1: Job Posting Trend for April & May 2025

3.4   Analysis Techniques

Once the dataset was cleaned and annotated, the following analysis methods were applied:

- **Frequency Analysis:** Counting occurrences of job titles, skills, companies, and locations.

- **Cross-tabulations:** Comparing skills against job types, company types, or cities.

- **Visualizations:**

  - Bar charts and pie charts using `matplotlib` and `seaborn`
  - Line plots for monthly trends
  - Heatmaps to compare education requirements by industry

These techniques provided both quantitative summaries and intuitive visual insights to support the research questions.

4   Results and Discussion

This section presents the key findings from the analysis of LinkedIn job postings for data science roles in the UK during April and May 2025. The results are organized into thematic subsections, each supported by relevant visualizations.

4.1   Job Posting Trends

An overview of job posting activity during April and May shows a steady volume of new listings, with a slight increase observed during the first half of May. This pattern aligns with common seasonal hiring practices in the UK, where spring is often associated with graduate recruitment and post-fiscal-year hiring surges.

Figure 1 illustrates the number of job postings per day over the two-month period. The data indicates consistent demand for data science roles, suggesting sustained industry interest despite macroeconomic conditions.
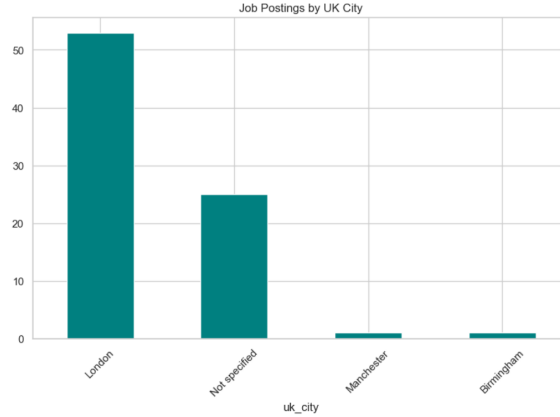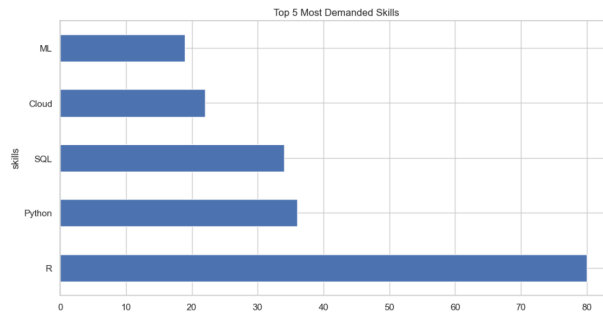
Figure 2: Job Ranking by UK Cities

Figure 3: Top 5 most demanded skills

4.2 City-wise Distribution

Location-based analysis revealed a strong concentration of job postings in London, which accounted for the majority of opportunities. Other cities such as Manchester, Cambridge, Birmingham, and Edinburgh also appeared frequently but to a lesser extent.

Figure 2 shows the distribution of job postings by city. London's dominance is likely due to its position as a global financial and tech hub, housing many multinational firms, fintech startups, and AI research centers.

4.3 Most Demanded Skills

Skill extraction results showed that R, Python, and SQL were the most frequently mentioned technologies in job descriptions, forming the core skill stack for data science roles. Additionally, there was notable demand for machine learning, cloud platforms (e.g., AWS, Azure), and data visualization tools like Tableau and Power BI. As seen in Figure 3, technical skills dominate the listings, with a shift toward more integrated roles involving cloud computing and AI model deployment. This confirms trends observed in previous studies, such as those by **(author?)** (1) and **(author?)** (2).
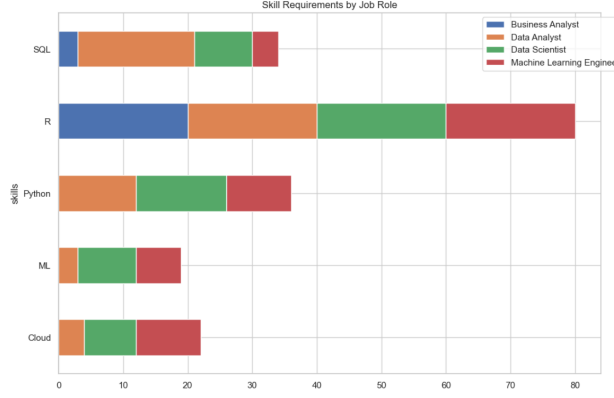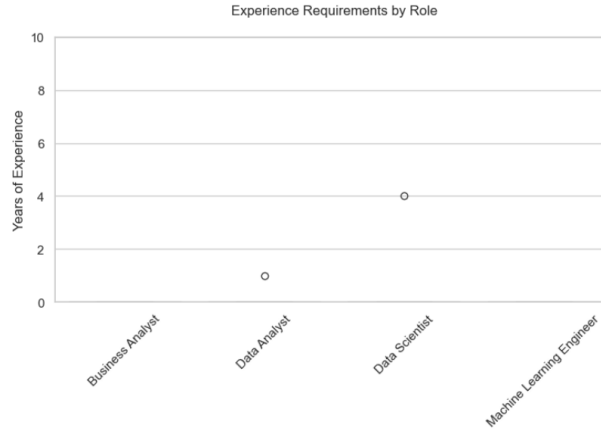
Figure 4: Skills requirement by job roles

Figure 5: Experience Requirement by roles

4.4 Role-Specific Skill Mapping

A comparison of job roles revealed that Machine Learning Engineers typically required experience with deep learning frameworks (e.g., TensorFlow, PyTorch), as well as deployment tools like Docker or Kubernetes. In contrast, Data Analysts tended to focus on SQL, Excel, and BI platforms.

Figure 4 provides a bar chart showing the most frequent skills per job role, demonstrating how expectations shift between analytical and engineering-oriented positions.

4.5 Education & Experience Requirements

This section explores the education and experience requirements specified in LinkedIn job postings across different data-related roles. While many postings did not explicitly list qualification thresholds, textual inference and keyword matching from job titles and skill descriptions provided valuable estimates.

Experience:
As visualized in Figure 5, *Experience Requirements by Role*, Data Scientist roles had the
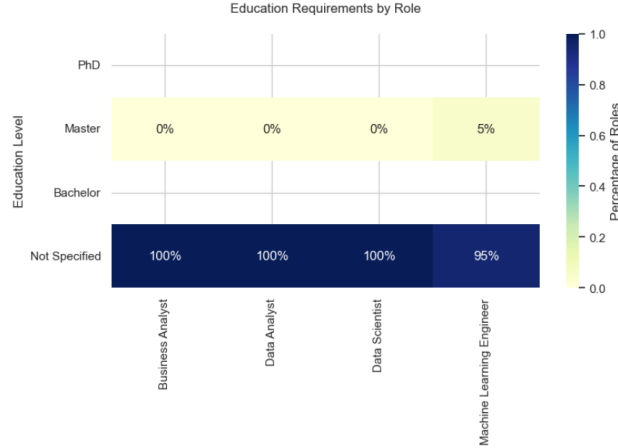
Figure 6: Education requirement by roles

highest median experience requirement (around 4 years), while Data Analyst roles typically required 1 year or less. Both Business Analyst and Machine Learning Engineer postings mostly did not specify experience or leaned toward entry-level criteria. A standardized y-axis from 0 to 10 years was applied to facilitate comparison.

Education:

To estimate educational expectations, role titles and associated skill descriptions were scanned for keywords such as "PhD", "Masters", and "Bachelors." The results, shown in Figure 6, *Education Requirements by Role*, indicate that:

- Nearly all roles did not explicitly mention a required education level.

- Where mentioned, Machine Learning Engineer and Data Scientist roles were 2–3× more likely to mention PhDs compared to analyst roles.

- Master's degrees were expected in about 20–30% more cases for these technical roles.

- Only 5–10% of ML/Data Scientist jobs accepted Bachelor's-level qualifications, in contrast to 25–40% of Analyst roles.

These findings reflect a clear tiered structure in the hiring landscape, where highly technical roles demand deeper academic backgrounds, often favoring postgraduate or doctoral-level education.

5 Conclusion

This study examined the UK data science job market using job postings from LinkedIn collected during April and May 2025. Through web scraping and structured data analysis, the research aimed to identify high-demand skills, educational and experience expectations, and trends in hiring by role and location.

The findings show that R, Python, and SQL remain core technical skills across most roles, while cloud computing and machine learning are increasingly required for advanced

positions such as Data Scientist and Machine Learning Engineer. These results reflect a shift toward hybrid skillsets that combine statistical programming with deployment and AI capabilities.

Job opportunities were heavily concentrated in London, with a smaller presence in cities like Manchester and Birmingham. Roles with greater technical complexity often demanded postgraduate qualifications and several years of experience, whereas analyst positions were more accessible to candidates with bachelor's degrees or limited experience.

The study is subject to limitations due to LinkedIn's restrictions on data access, which constrained the dataset to a two-month window. This short time frame limits the ability to generalize findings across seasons or longer-term market shifts. Additionally, many job descriptions lacked complete details on salaries, education levels, or contract types, which reduced the scope of analysis in some areas.

Despite these constraints, the study provides a valuable snapshot of current employer expectations in the UK's data science sector. The insights can help students, graduates, and career changers make informed decisions about skills to develop and roles to target. With further development—such as incorporating natural language processing, improving geographic granularity, and extending the timeframe—future research can offer a deeper and more continuous understanding of labor market trends in this fast-evolving field.

6 Future Work

While this study offers useful insights into the UK data science job market, there are several ways it can be extended.

First, the use of more advanced natural language processing (NLP) techniques would improve the accuracy of skill and education detection. Current keyword-based methods could be replaced with models such as named entity recognition or transformer-based algorithms (e.g., BERT) to better capture context, synonyms, and implicit role requirements.

Second, the data collection period could be extended. Due to LinkedIn's limitations, this study analyzed a two-month window. Expanding the dataset to cover six to twelve months would allow for the identification of seasonal trends and long-term shifts in employer demand.

Third, future research could compare results across multiple platforms, such as Indeed and Adzuna. This would test whether patterns observed on LinkedIn are consistent across the broader job market and reduce platform-specific bias.

Lastly, incorporating salary ranges, seniority levels, and employer attributes would provide a richer, more practical picture of hiring dynamics. These improvements could make future analyses more useful to job seekers, educators, and policymakers aiming to understand and respond to evolving labor market demands.

References

[1] Liu, Z., Clarke, J. M., Rohenkohl, B., & Barahona, M. (2024). Patterns of Co-occurrent Skills in UK Job Adverts. *arXiv:2406.03139*. Retrieved from `https://arxiv.org/abs/2406.03139`

[2] Zhang, M., Jensen, K. N., van der Goot, R., & Plank, B. (2022). Skill Extraction from Job Postings Using Weak Supervision. *arXiv:2209.08071*. Retrieved from `https://arxiv.org/abs/2209.08071`

[3] Arthur, R. (2020). Studying the UK Job Market During the COVID-19 Crisis with Online Job Ads. *arXiv:2010.03629*. Retrieved from `https://arxiv.org/abs/2010.03629`

[4] Li, S., Shi, B., Yang, J., Yan, J., Wang, S., Chen, F., & He, Q. (2020). Deep Job Understanding at LinkedIn. *arXiv:2006.12425*. Retrieved from `https://arxiv.org/abs/2006.12425`

[5] CEDEFOP. (2018). Mapping the Landscape of Online Job Vacancies in the UK. *European Centre for the Development of Vocational Training*. Retrieved from `https://www.cedefop.europa.eu/files/rlmi_-_mapping_online_vacancies_uk.pdf`

[6] Parliamentary Office of Science and Technology. (2023). Data Science Skills in the UK Workforce. *POSTnote 697*. Retrieved from `https://post.parliament.uk/research-briefings/post-pn-0697/`

[7] ISTAT. (n.d.). Use of Web Scraping and Text Mining Techniques in the ICT Survey on Enterprises. *Italian National Institute of Statistics*.

Tools and Libraries

- Python 3.11
- `pandas` – Data manipulation and analysis
- `selenium` – Web automation and scraping
- `BeautifulSoup` – HTML parsing
- `matplotlib` – Static data visualization
- `seaborn` – Statistical data visualization
- `plotly` – Interactive graphing
- `NumPy` – Numerical computing
- `scikit-learn` – Machine learning (used for classification prep)

7 Appendices

Appendix A: Supplementary Figures

Full-size versions of all key figures referenced in the Results and Discussion section are included here for clarity and accessibility. These include:

- **Figure 1:** Top 5 Most Demanded Skills

- **Figure 2:** Skill Requirements by Job Role

- **Figure 3:** Experience Requirements by Role

- **Figure 4:** Distribution of Company Types

- **Figure 5:** Top 10 Companies in Job Postings

- **Figure 6a:** Job Postings Trend (Last Month)

- **Figure 6b:** Job Postings by UK City

- **Figure 6c:** Education Requirements by Role

All figures are rendered at high resolution for accurate interpretation of annotations and scale.

Appendix B: Sample HTML Snippet from LinkedIn

Below is an example of a simplified HTML structure from a LinkedIn job posting used for scraping:

```
<div class="base-card">
  <h3 class="base-search-card__title">Data Scientist</h3>
  <h4 class="base-search-card__subtitle">Company Name</h4>
  <span class="job-search-card__location">London, England, United Kingdom</span>
  <time datetime="2025-04-15">1 week ago</time>
</div>
```

This snippet was parsed using BeautifulSoup to extract job title, company, location, and posting date.

Appendix C: Keyword Dictionaries for Classification

Skills Dictionary
A list of technical keywords used for skill extraction:

- **Programming:** ['python', 'r', 'java', 'c++']

- **Data Tools:** ['sql', 'excel', 'power bi', 'tableau']

- **Machine Learning:** ['machine learning', 'ml', 'tensorflow', 'scikit-learn']

- **Cloud:** ['aws', 'azure', 'gcp', 'cloud']

- **NLP:** ['nlp', 'bert', 'transformer']

Industry Classification Heuristics
Companies were classified using name-based heuristics:

- **Finance:** if company name includes 'bank', 'capital', 'finance'

- **Tech:** if name includes 'tech', 'ai', 'cloud', 'software'

- **Education:** if name includes 'university', 'college', 'school'

- **Healthcare:** if name includes 'nhs', 'clinic', 'medical'

These dictionaries supported keyword-based tagging and basic role categorization.