

JOB POSTING DASHBOARD



Contents

Dataset Overview2

Purpose and Goals2

Key Questions and Expectations3

Findings:4

Metrics and KPIs5

Data Model.....6

Conclusions6

Limitations 10

Dataset Overview

A brief overview of the dataset would be that it is a collection of job listings that have been scraped from LinkedIn. This dataset was retrieved from Kaggle and under the title on Kaggle it says 'Snapshot of the job market' which is an accurate summary. The dataset included details such as job titles, company names, locations, job descriptions, required skills, and possibly salary information and this information was scraped from LinkedIn. The frequency of updates for the database was set monthly. The job market is dynamic, so the data might not reflect the most current trends. This was an interesting database to work with as it gave us a good understanding of which sectors and industries are posting jobs and which skills are in demand each day. Overall, this LinkedIn job postings dataset offers a rich source of information for anyone interested in exploring job market trends, conducting labor market research, or young students aiming to hone their skills where there is a demand or developing job recommendation systems.

Purpose and Goals

One of the main purposes of this database was to analyze the LinkedIn job postings dataset was to gain insights into current job market, get the number of hirings, what the average salary in a particular role looks like and at which locations. The aim was to create a comprehensive dashboard that would provide valuable information for job seekers, employers, journalists and industry analysts. Our key goals included:

1. Analyze the job market: Just a snapshot of the past year at LinkedIn gave us an insight into what the job market was looking like. The data was from 2023-2024 to ensure our insights were up-to-date and relevant.
2. To identify which skills are in demand and jobs: LinkedIn job postings usually come with a list of skills required. We wanted to highlight which roles and skills are most sought after.
3. Exploring salary trends: Salary is a tricky thing when it comes to job postings. Some companies share a range, some share a number, and some do not share anything. Although there were challenges with incomplete salary data, the aim was to provide salary comparisons across different job titles where possible.
4. Geographical job distribution mapping: A good way to help identify which locations had the most job opportunities was through mapping, which could be valuable for job seekers considering relocation.

5. To show which companies were active in the job market: Some companies due to the industry they are in or due to their growth are more active than others when it comes to job posting and hiring. The purpose of this dashboard was to identify and showcase which companies were most actively hiring on LinkedIn.
6. Effectively visualize the data: One of the main goals was to present our findings in an easy to understand way using tools like Power BI, making the data accessible to a wide audience.

By working with this extensive dataset of over 125,000 job postings, we aimed to create a comprehensive picture of the current tech job market, offering valuable insights that could inform career decisions, hiring strategies, and industry analysis

Key Questions and Expectations

Which companies are most actively hiring in the industry?

Before we were beginning with our work, we opened all the files in the database to see the table structure and what type of data were in columns. We checked to see if there was any missing data, which we did run into as previously mentioned the data came from LinkedIn scraping, so we can say that what we saw is what we got. CompanyID and related tables gave us the thought that we can get a finding out which company hires the most based on how many times CompanyID has been repeated. So, this was the first key question.

What are the most frequent job titles and how do their salaries compare?

JobID and job titles were present in the database so we thought we could also get what which job title was recurrent, this would address which jobs are in high demand. And due to the salary columns, we would also compare the salaries based on job titles. So, this was one of our key questions.

How do job postings and salaries vary by location?

We also saw that despite having a lot of remote jobs, locations were clearly mentioned so we can create a heatmap of sorts to visually show that which jobs were located where.

So these were the expectations and key questions we had in mind going forward.

Findings:

In this section we will share the findings which we got from analyzing the database. Our analysis revealed that Liberty Healthcare was one of the most actively hiring companies on LinkedIn during the period of 2023-2024 covered by our dataset. Some small scale SMEs and startups, and rapidly growing tech firms also showed significant hiring activity. The dashboard we created allows users to visualize the top hiring companies. A clustered column graph to compare the company posting and tree map to show the top 10 companies.

In general, typically in-demand tech roles often include Software Engineers, Data Scientists, Product Managers, and UX Designers etc. We also showed the salary comparison graph which allows users to compare salaries between different job titles. For example, we might see that Senior Software Engineers generally command higher salaries than Junior Developers, or that Data Scientists in certain specializations earn more than general Data Analysts. It's important to note that salary data was sometimes incomplete, with many companies not disclosing salary ranges, which posed a challenge in creating comprehensive salary comparisons. To geographically show the distribution of job postings, which helps identify hotspots. Major tech hubs like San Francisco, New York, Seattle, and Austin likely show a high concentration of job postings. However, with the rise in remote work, we might see a more distributed pattern of job locations. Regarding salaries, they typically vary significantly based on location due to factors like cost of living and local market demand. Our dashboard allows users to filter job postings and salary data by location, providing insights into these geographical variations.

One of the main talking points in recent times is regarding remote work. This additional question is particularly relevant given the recent shifts in work arrangements. Our analysis of the LinkedIn dataset likely shows a significant increase in remote job postings. In terms of salaries, the rise of remote work might be leading to some salary equalization across different regions, although this trend would need careful analysis to confirm.

Metrics and KPIs

For the sake of structure and to track progress we had to introduce some KPIs. We defined the following KPIs to track our objectives:

1. Total number of job postings
2. Average salary job title
3. Most frequent skills
4. Job postings by location
5. Active hiring companies

For further details,

1. Total number of job postings:

We chose this KPI to get an overall picture of the job market activity. In order to find total postings, we included another primary key called Posting ID. Hence, we were able to find it and it's a fundamental metric that shows the volume of opportunities available. In Power BI, we likely created a simple measure `Total Jobs = COUNTROWS('Postings')`. This measure was then displayed using a card visual for easy reference. This gave an understanding of how the job market is behaving currently.

2. Average salary by job title:

One of the points we wanted to show was the average salary of the job title. This KPI provides insights into compensation trends across different sectors and roles. It's crucial for both job seekers and employers to understand market rates. Although some data was missing as mentioned there were some postings with no mention of salary. In Power BI, we created measures for average minimum and maximum salaries. The formula was: `Avg Max Salary = AVERAGE('Postings'[salary_max])`.

3. Most frequent skills:

Most postings had skills mentioned in the ad. This KPI was chosen to help identify which skills were highest in demand, reflecting current market needs. To visualize this in Power BI, we used a donut chart to show top 10 with skills in the legend sections. We used a measure like `Skill Count = COUNTROWS('Postings[skills]')` and then used the Top N filter to show only the most frequent titles.

4. Job postings by location:

We wanted to incorporate a heatmap. The distribution of jobs is valuable for understanding where opportunities are concentrated. In Power BI, we used a map

visual, with the size of dots representing the number of job postings in each location. In the column 'City' in the We might have created a measure for job count and then used it as the size value in the map visual.

5. Active hiring companies:

We also wanted to identify which companies are active. This KPI is useful for job seekers and for understanding market dynamics. In Power BI, we created a tree map showing companies. We used a measure like Company Name Count = `COUNTROWS('Postings'[company_name])` and then sorted the chart in descending order to highlight the most active companies.

These KPIs hopefully will give you a good picture of LinkedIn landscape and when we combined all these KPIs we were provided a comprehensive view of the job market, covering volume, geography, compensation, demand, and key players. This will allow for both high-level overview and detailed breakdown, making them valuable for various stakeholders in the job market.

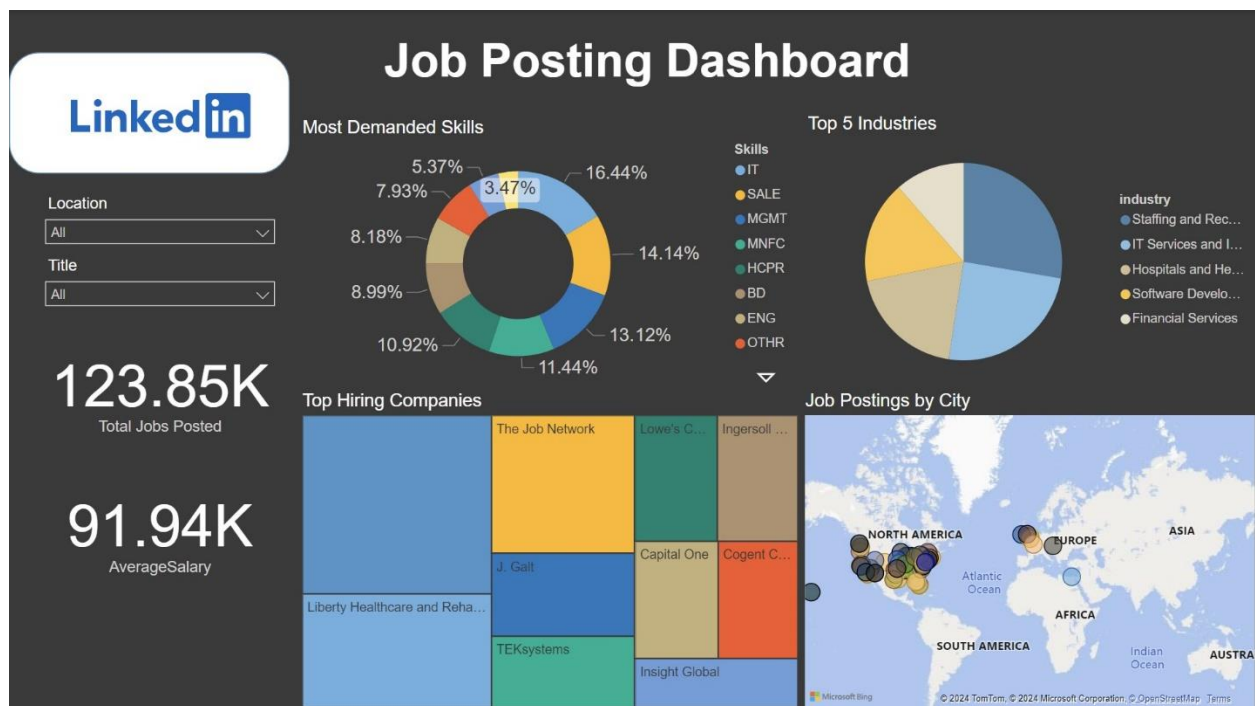
Dashboard

Data Model



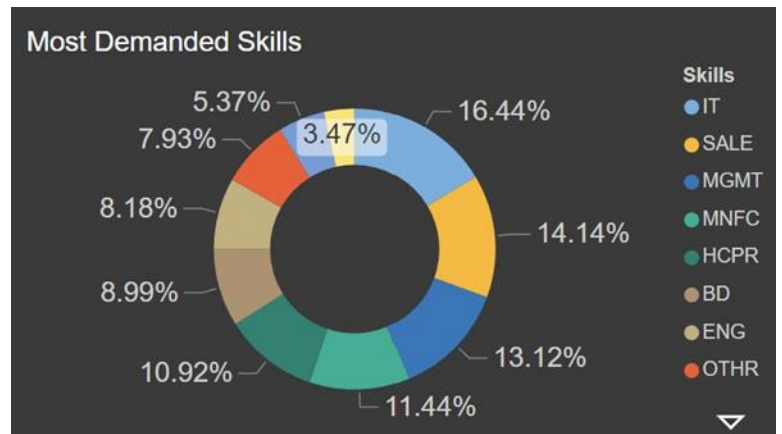
For the sake of simplicity, we tried normalization of the data model. The posting table was very extensive and large and served the role as the fact table. We still normalized the fact table so that it was easy for us to wrangle with the data. Also the database had a lot of tables other than postings table. Some had irrelevant data, we included the ones which were relevant with our objectives and were showing our KPIs. One of the industry table was extra so which was repeating the data from the postings table. We established relationships using primary keys. Also created a few extra table to run the group by function.

Visualization

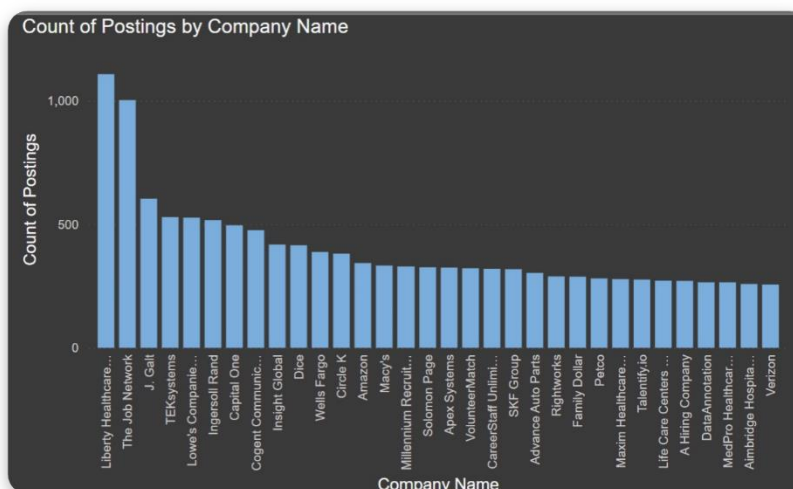
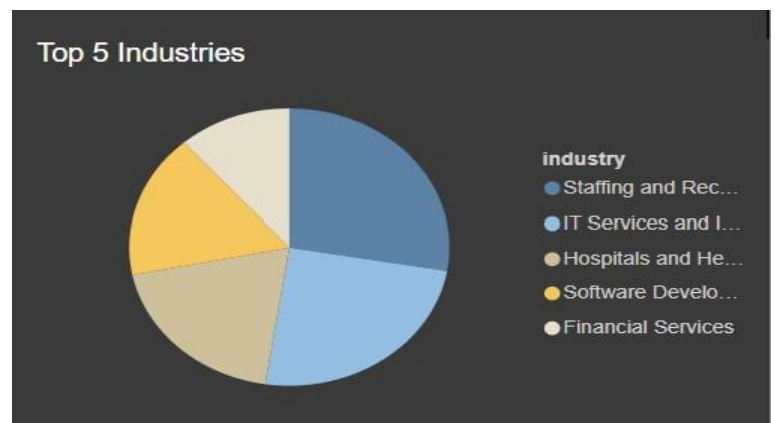


A LinkedIn logo was used to show more context regarding where the database was coming from. A splicer was included which either had the option to choose location or title to refine their findings. Top 5 industries was also a necessary addition which showed that HR sector was also very active followed by IT and Healthcare.

For this graph, we used a measure in the skills table, and it came back with the count of skills. We used a donut chart and in the legend section included skills column and in the values section we used the measure. And then we put a filter so that we would filter out the top 10 results. Which gave us this graph showing IT dominating the number one spot.



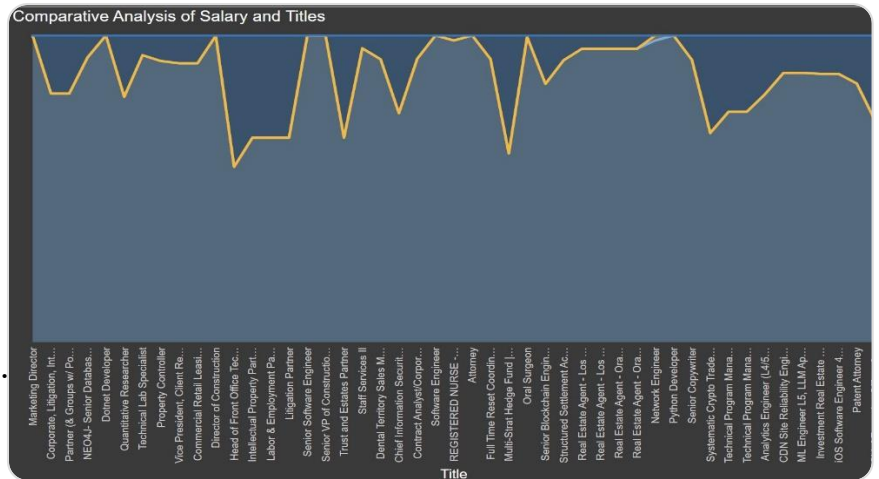
We also wanted to show which industries were more recurring for this we used the postings table and chose the pie chart. We included industries in the legend section and in values again included the industries and chose the count function for it. We then used the filter to make it in ascending order.



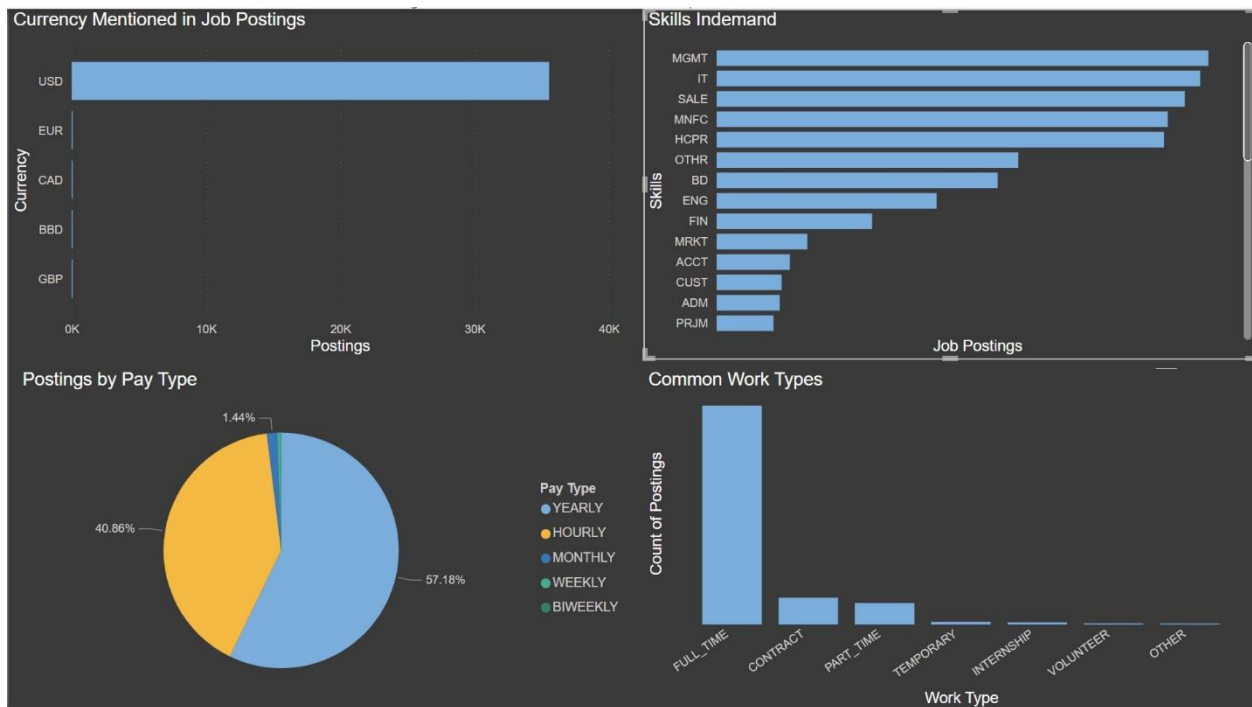
For which companies that were active on LinkedIn we chose clustered column chart we made a count of postings by counting Posting ID. And putting it on y-axis. And company name on the x-axis, which gave us the following.

Raja Muhammad Uzair Nasir

We also wanted to compare the salaries across different job titles. At default it looks convoluted. But we can choose which jobs to compare with which ones. The difference between the peaks are meant to be the percentage of difference between the max salaries.



The cards were used to show that which average salary of the specific job chosen. The top 5 industry graph and job posting graphs are meant to be static and not change. While the rest of the dash board is supposed to be interactive. Some additional graphs were made just for any additional research one may want to do. Which is the following.



Conclusions

In conclusion, the job market shows significant activity, with certain companies leading in hiring volume. The dashboard was aimed to be interactive, insightful and pragmatic in its approach to analysis. Geographical location plays a significant role in job availability and potentially in salary ranges. Certain job titles are in higher demand, reflecting current industry trends and skill requirements. Metrics such as the total number of job postings, average salaries, and top hiring companies offered a structured approach to understanding job market dynamics. Overall, this analysis provided actionable information for various stakeholders, including job seekers, industry analysts, and employers, helping them make informed decisions in a rapidly changing employment landscape.

Limitations

Some of the limitations which we ran into were that many companies did not provide salary information, limiting the comprehensiveness of our salary analysis. the dataset is specific to LinkedIn postings and may not represent the entire job market but since LinkedIn is a big factor in the job market so it can still count.