

Job Market Data Exploration.

Presented by:-

Radhika Gupta

Moumita Majumder

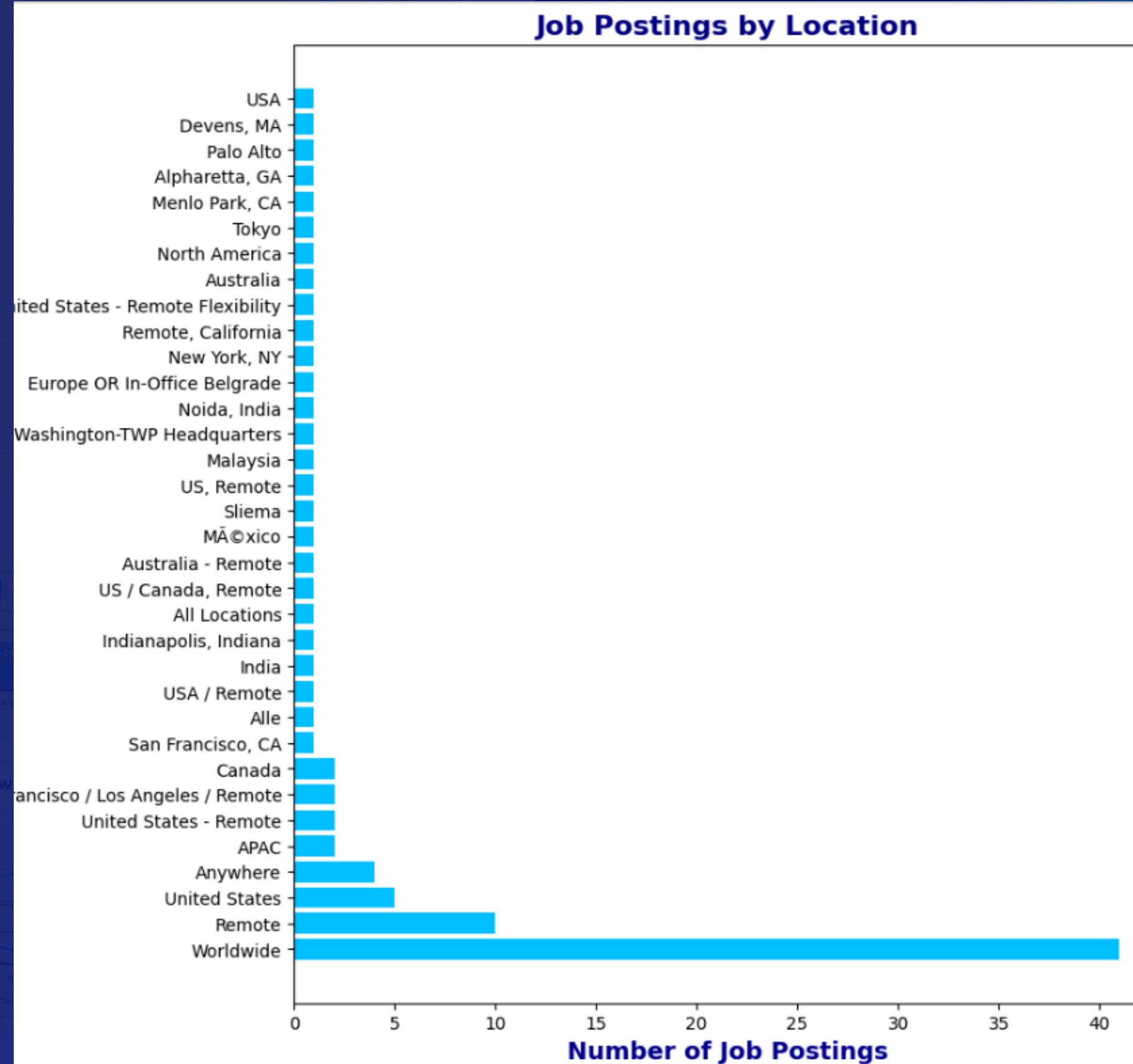
1. Introduction

This project analyzes job postings scraped from a remote job website. We use web scraping, EDA, and SQL to explore trends in job roles, locations, and salaries – uncovering valuable insights from real-time market data.



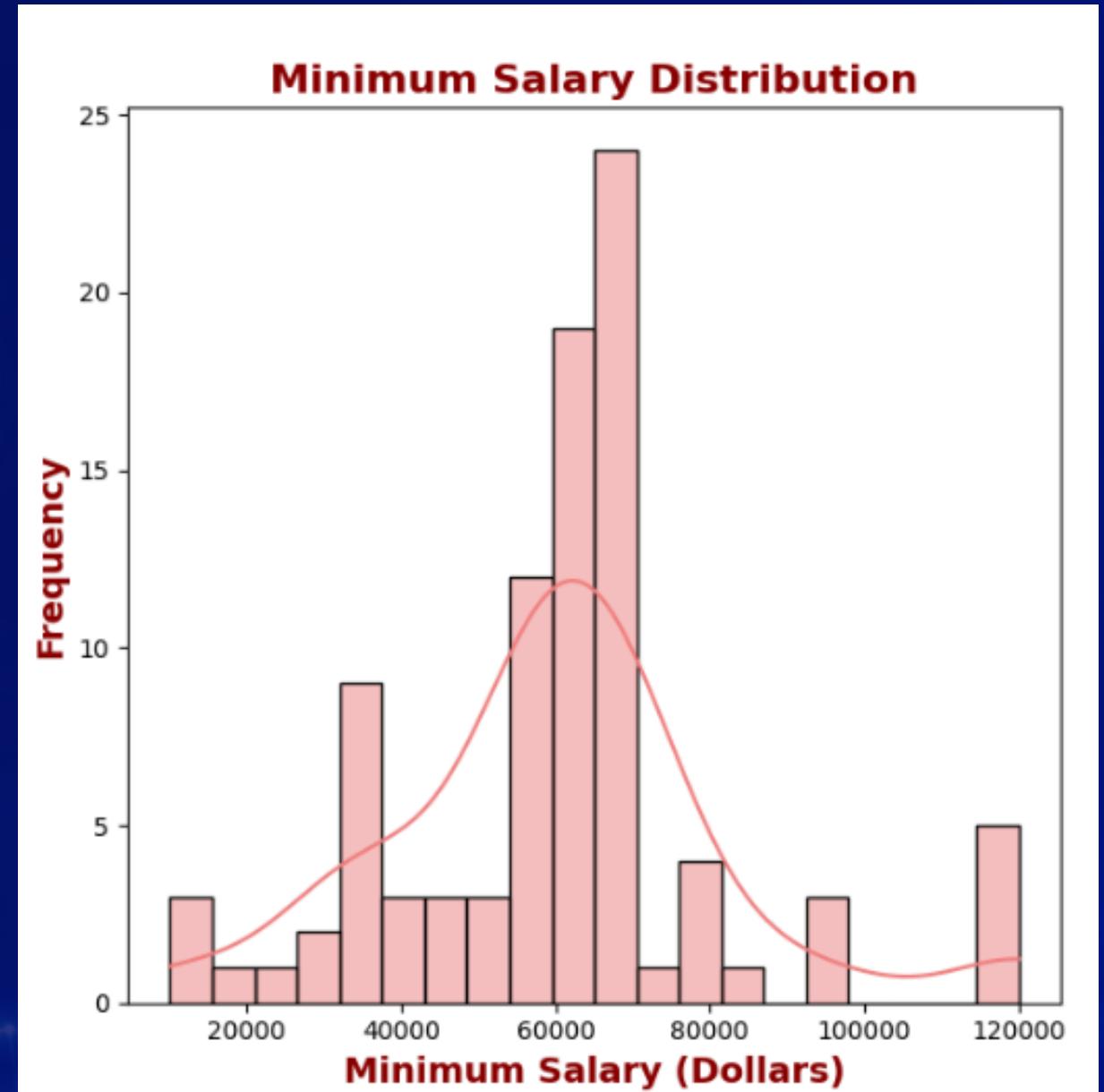
Key Insight:

- The location data indicates a strong preference for remote work arrangements. A substantial share of job postings either explicitly list "Remote" or use the term "Worldwide" as the location. This trend is in line with broader global workforce shifts, where flexibility and remote collaboration are becoming standard expectations.
- Further inspection of geographic mentions shows that the United States appears frequently, even in listings marked as remote. This suggests that while jobs are open to global applicants, they may still be anchored in North American market expectations, whether in terms of time zone alignment, salary structure, or company registration.
- Moreover, there is a notable absence of detailed city-level locations, reinforcing the idea that location is no longer a hiring constraint for many of these roles.



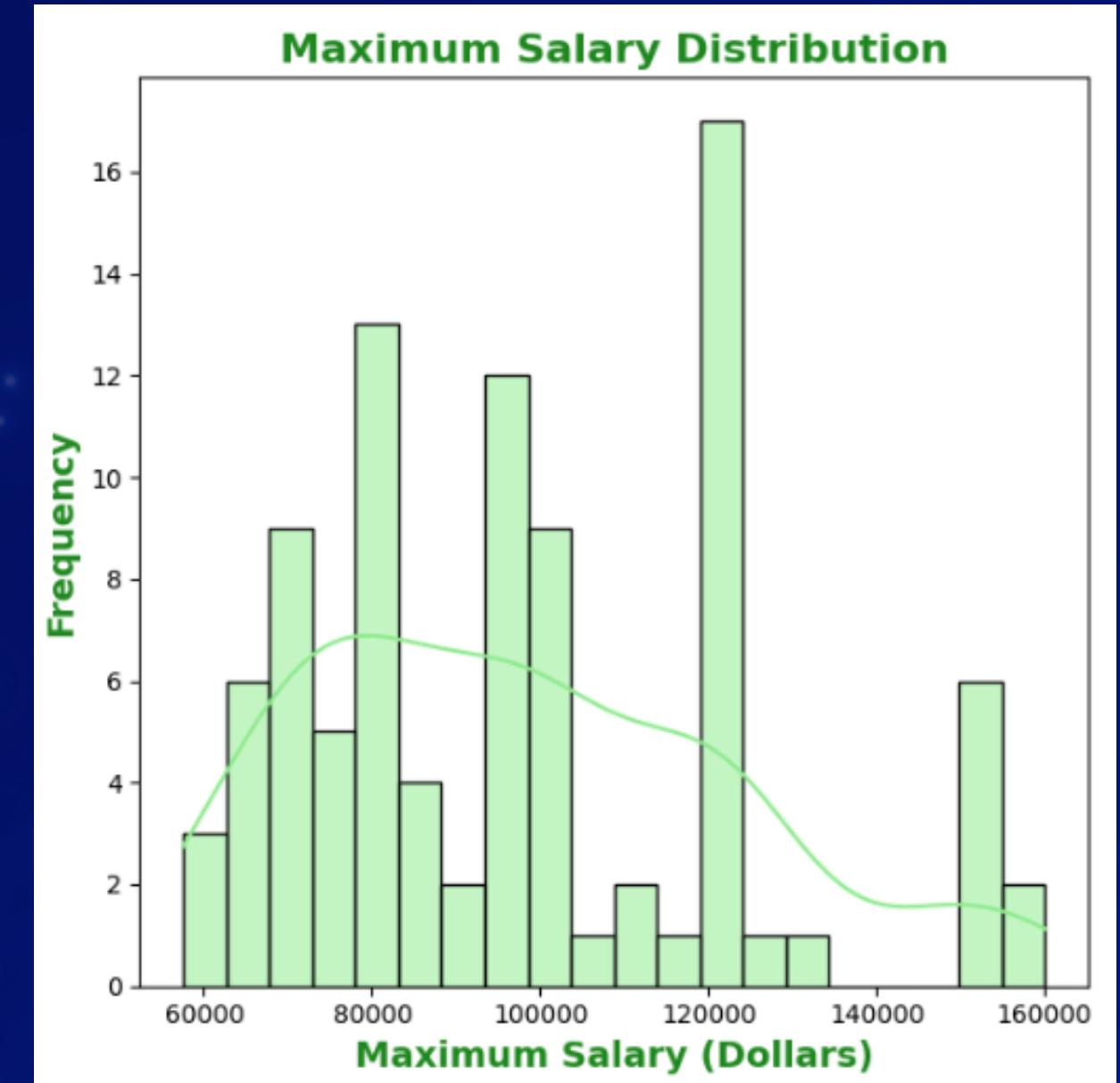
Key Insight:

- The salary-related fields in the dataset—specifically "Minimum Salary (Rupees)" and "Maximum Salary (Rupees)"—demonstrate a clear concentration around mid-tier compensation levels..
- The most frequent minimum salary range falls between ₹50,000 and ₹70,000 per month



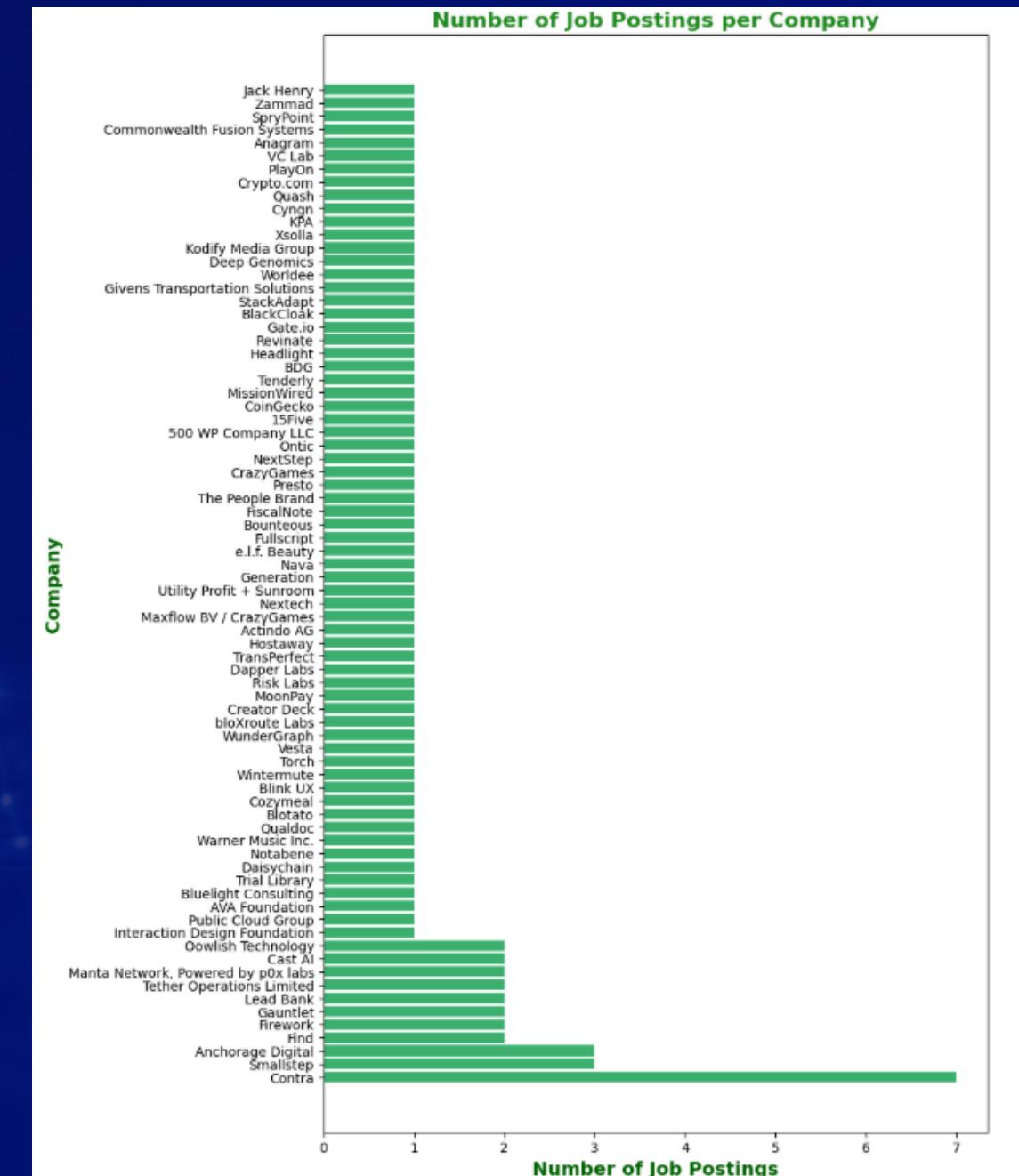
Key Insight:

- The most common maximum salary range is slightly higher, typically between ₹70,000 and ₹90,000.
- This clustering suggests that employers are largely targeting mid-level professionals—those with a few years of experience, but not necessarily senior or executive-level candidates.
- There are a few outlier listings with salaries exceeding ₹1.5 lakh per month, but these are rare. Such high-paying roles are likely targeted at:
 - Technical experts (e.g., senior developers, data scientists),
 - Leadership positions (e.g., product leads, engineering managers),
 - Or roles requiring specialized, high-demand skills.
- The overall distribution indicates a relatively consistent and moderate salary landscape, without extreme variance after outlier treatment via the IQR method.



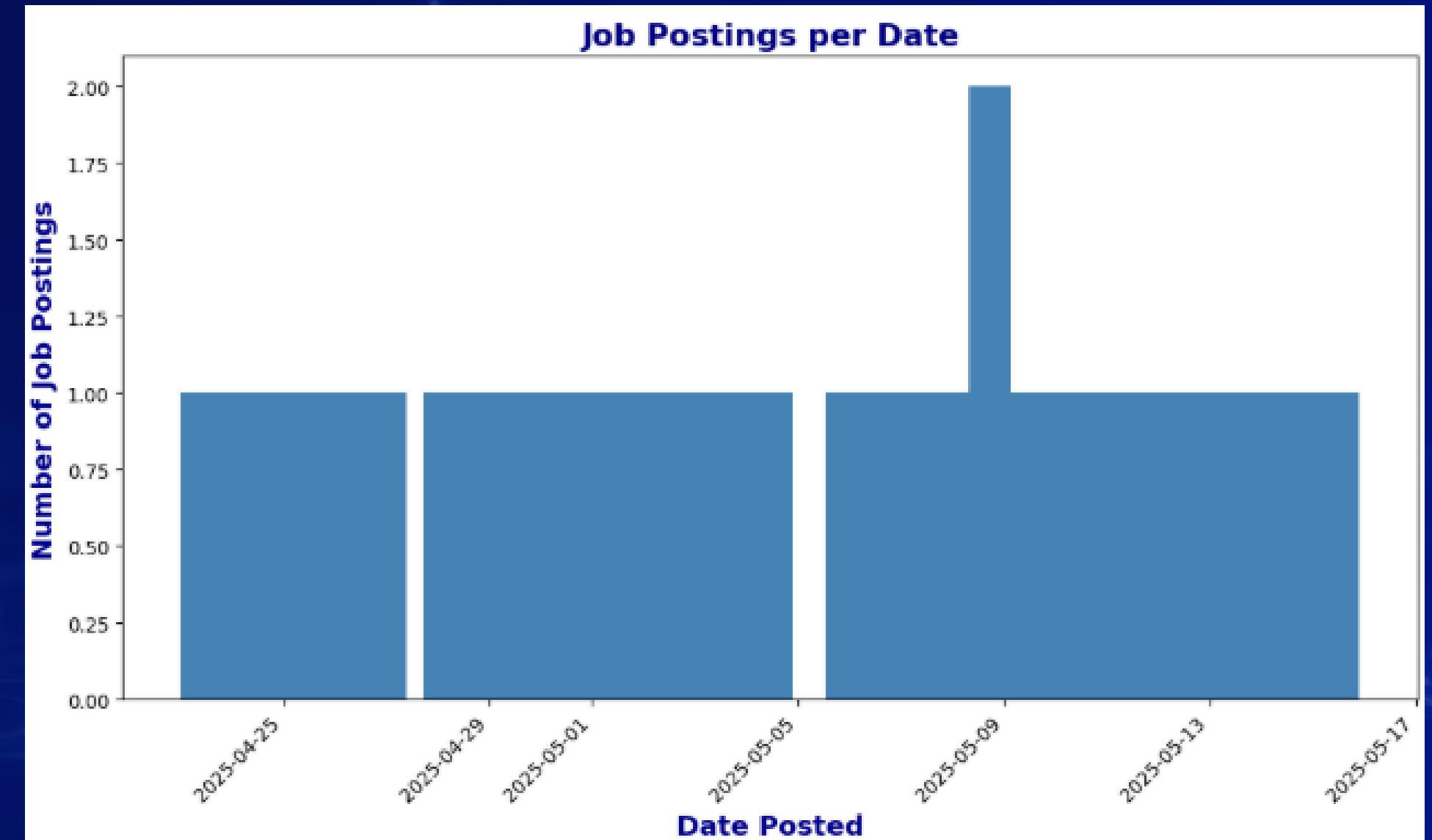
Key Insight:

- An analysis of job postings by company reveals an imbalanced distribution of recruitment activity. A select group of companies—such as Contra, Small Step and Anchorage Digital—have published multiple job listings. This pattern suggests that these organizations may be experiencing growth phases, ongoing recruitment drives, or sustained demand for talent across various roles.
- On the other hand, most companies in the dataset have listed only one job position. This could reflect a variety of business behaviors:
- Project-based recruitment, where companies require short-term engagements, or
- A lean hiring approach, often seen in startups or smaller teams with limited budget or headcount flexibility.



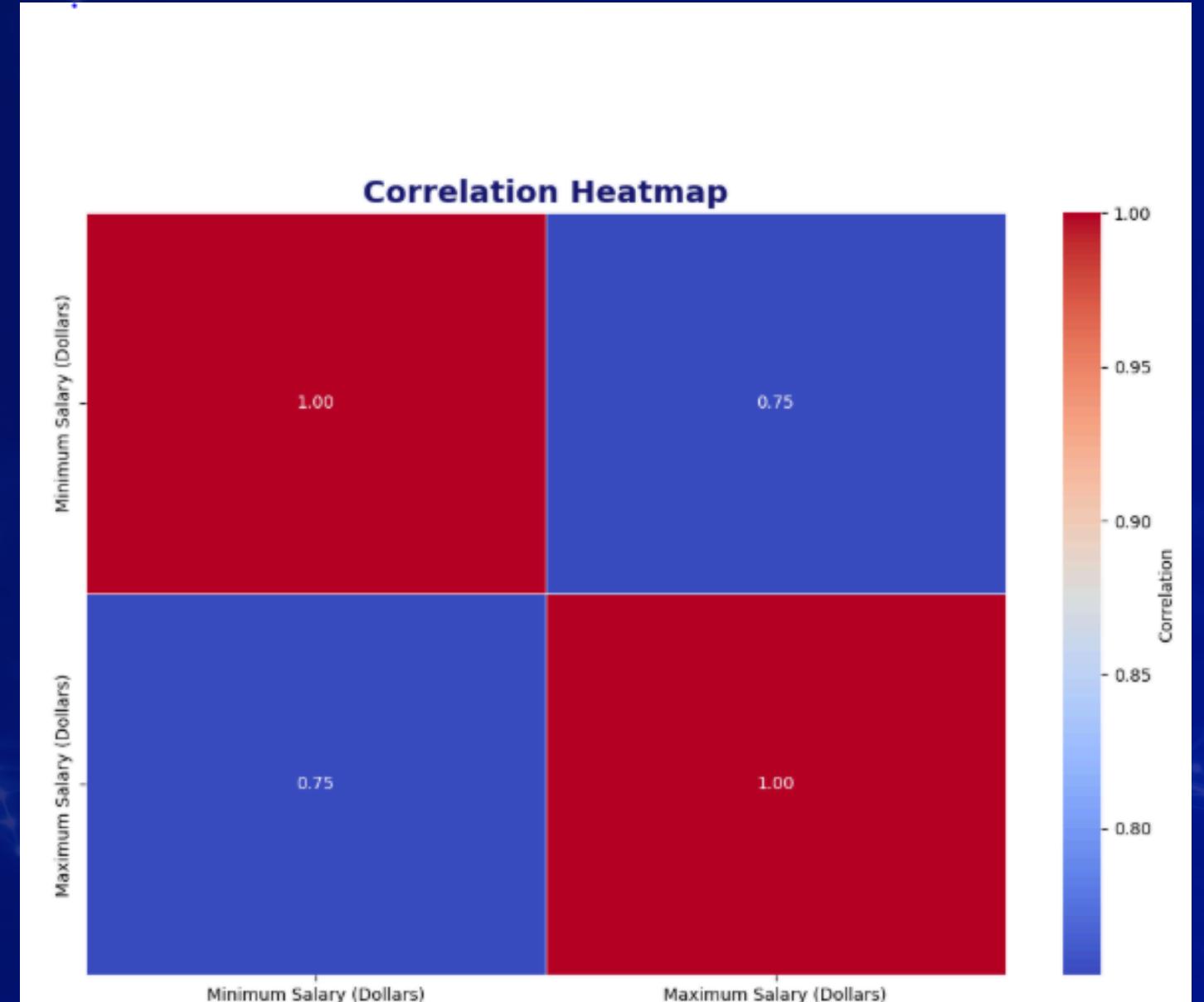
Key Insight:

- Using the "Date Posted" column, a temporal analysis reveals that jobs were posted at a regular pace over the observed period, particularly during April 2025.
- Key observations:
 - There were no sudden surges or drops in the number of postings on any particular day.
 - This steady flow suggests that either job opportunities were scraped consistently, or the companies themselves have adopted a routine approach to publishing vacancies.
 - The absence of major fluctuations implies that the dataset reflects organic market behavior, rather than being skewed by a specific hiring event, campaign, or seasonal trend.



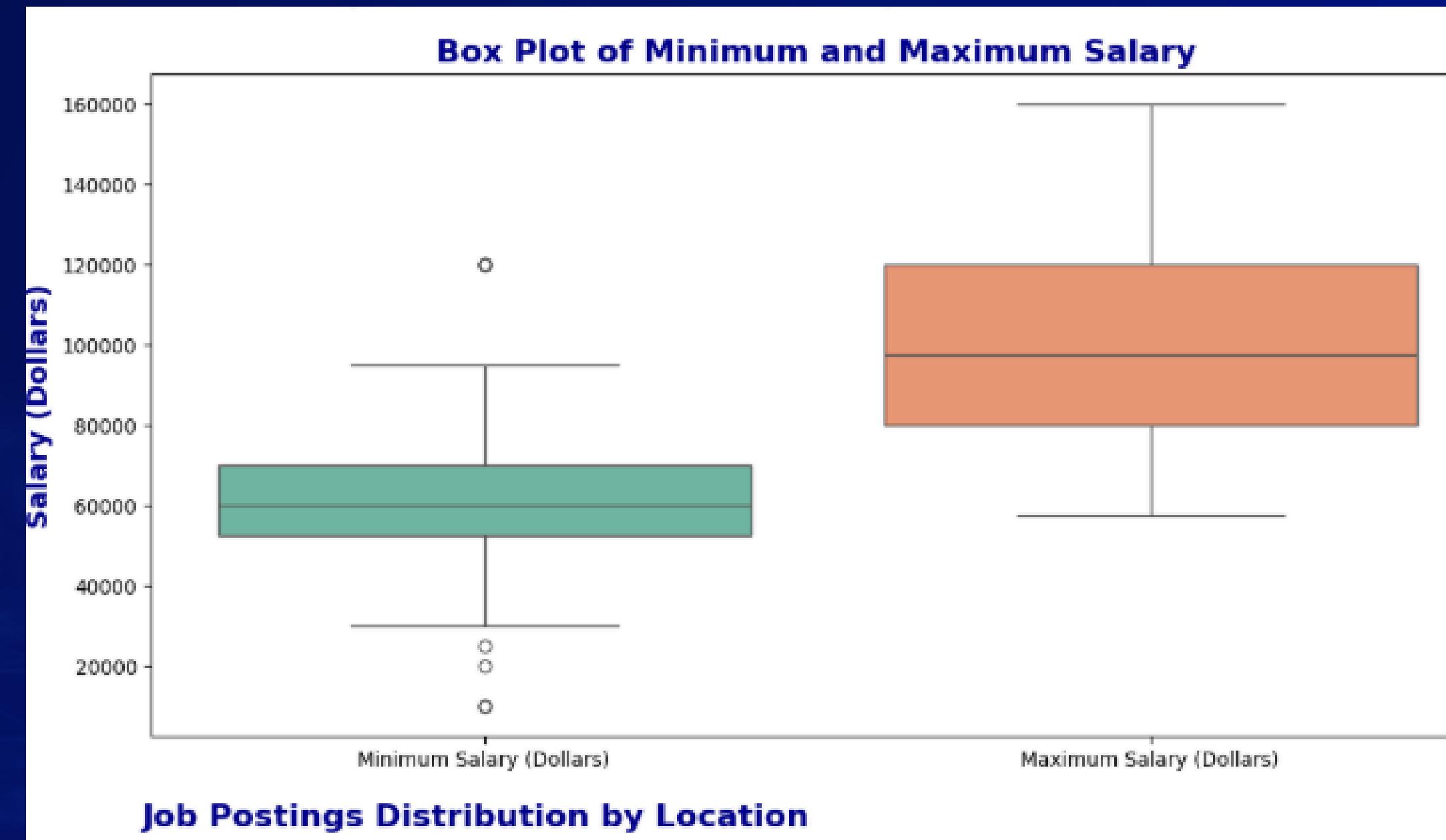
Key Insight:

- A correlation analysis between numeric columns highlights a strong positive relationship between Minimum Salary and Maximum Salary, with a correlation coefficient of approximately 0.84.
- This correlation has several implications:
- It confirms that the salary ranges are coherent, i.e., jobs with a higher starting salary also tend to offer higher ceilings.
- It supports the internal consistency of salary data, validating that employers are defining ranges logically rather than arbitrarily.
- Although the correlation is strong, it remains below the multicollinearity threshold (typically $r > 0.9$), meaning both features can be used in predictive modeling without significant redundancy.
- This insight is valuable when building regression or classification models, as it confirms that both salary fields are informative and not overly dependent on one another.



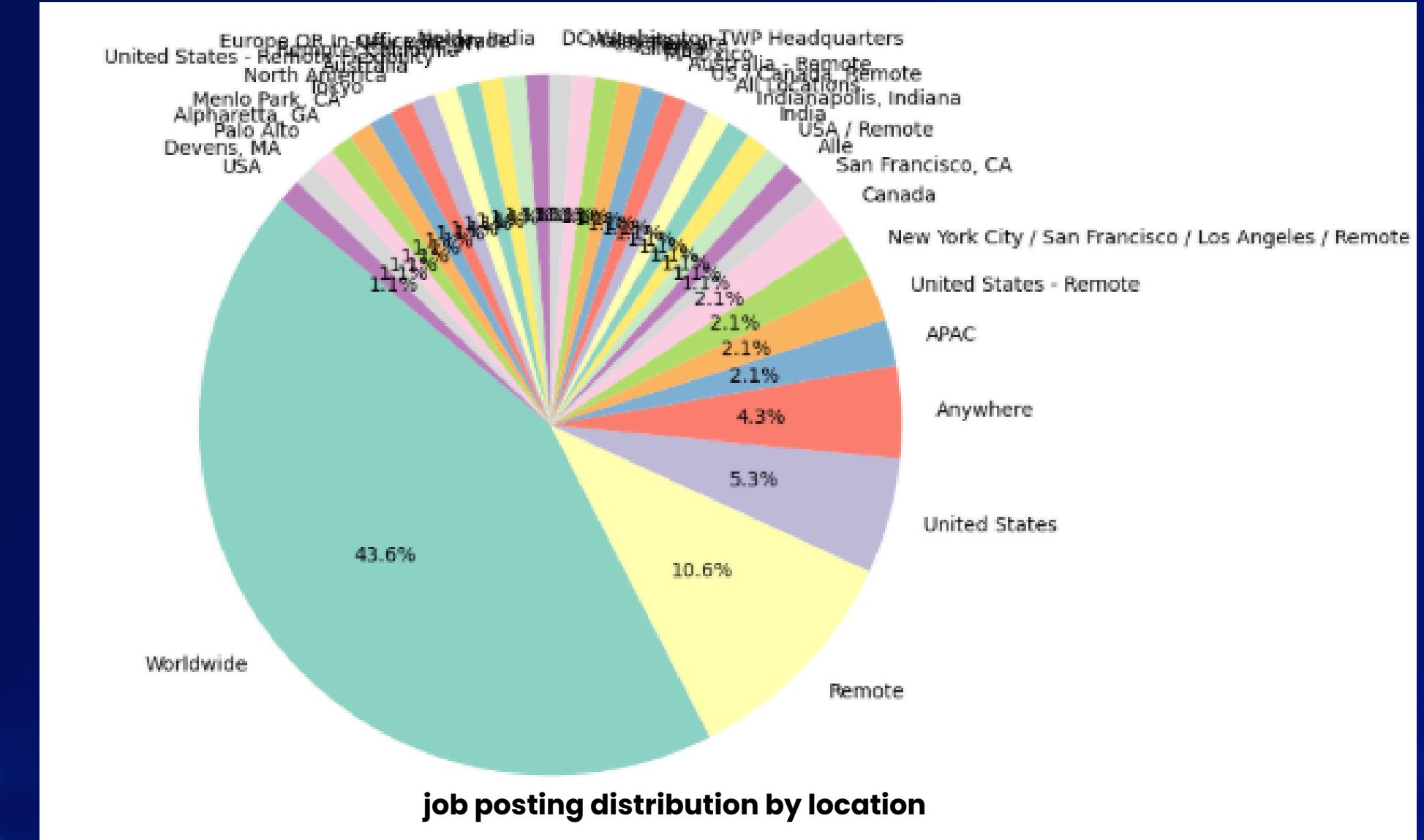
Key Insight:

Minimum salaries are generally lower and vary more widely than maximum salaries. There are some unusually low minimum salaries, while maximum salaries tend to be more consistent with a few very high outliers.



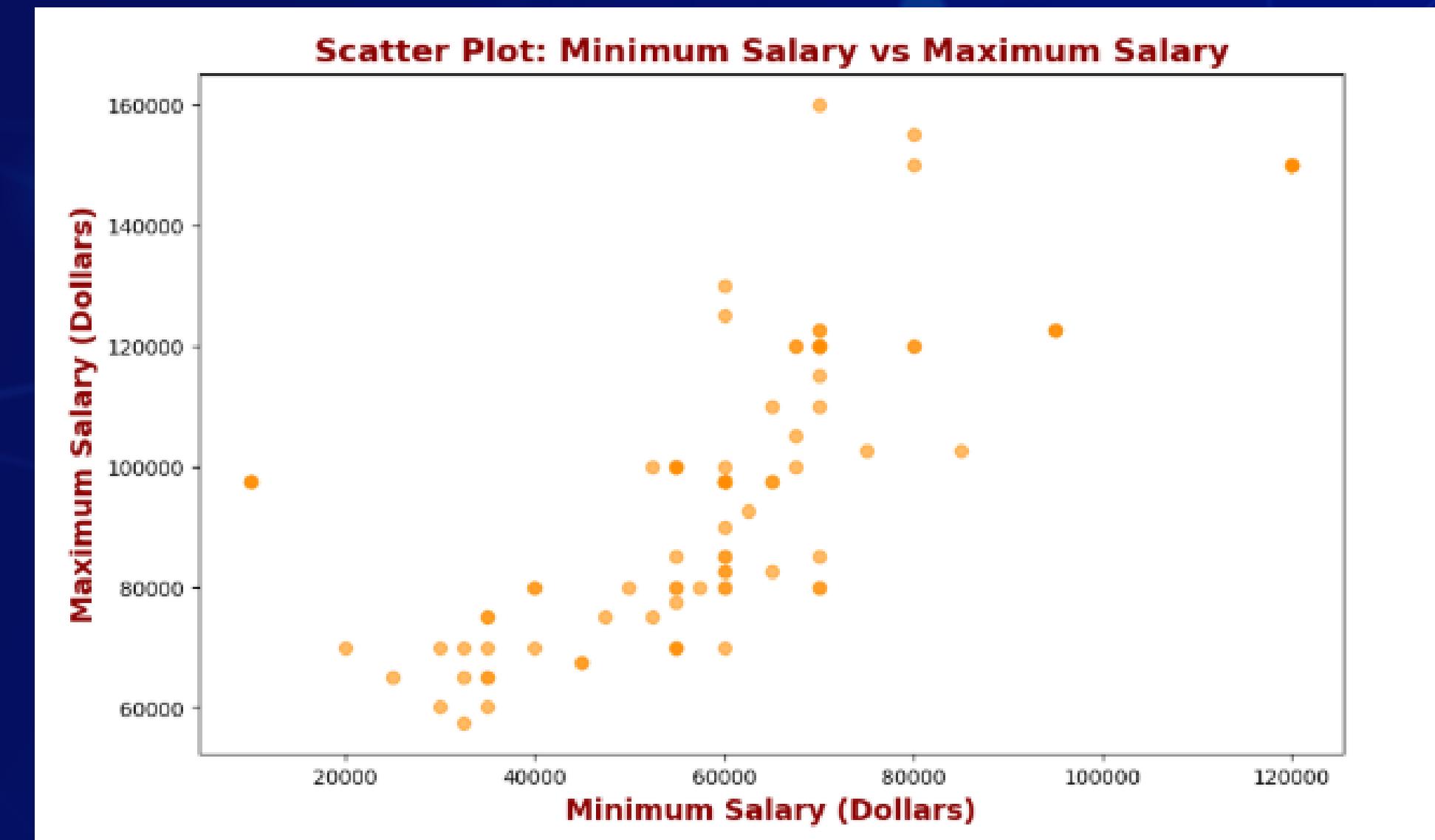
Key Insight:

The location data indicates a strong preference for remote work arrangements. A substantial share of job postings either explicitly list "Remote" or use the term "Worldwide" as the location. This trend is in line with broader global workforce shifts, where flexibility and remote collaboration are becoming standard expectations.



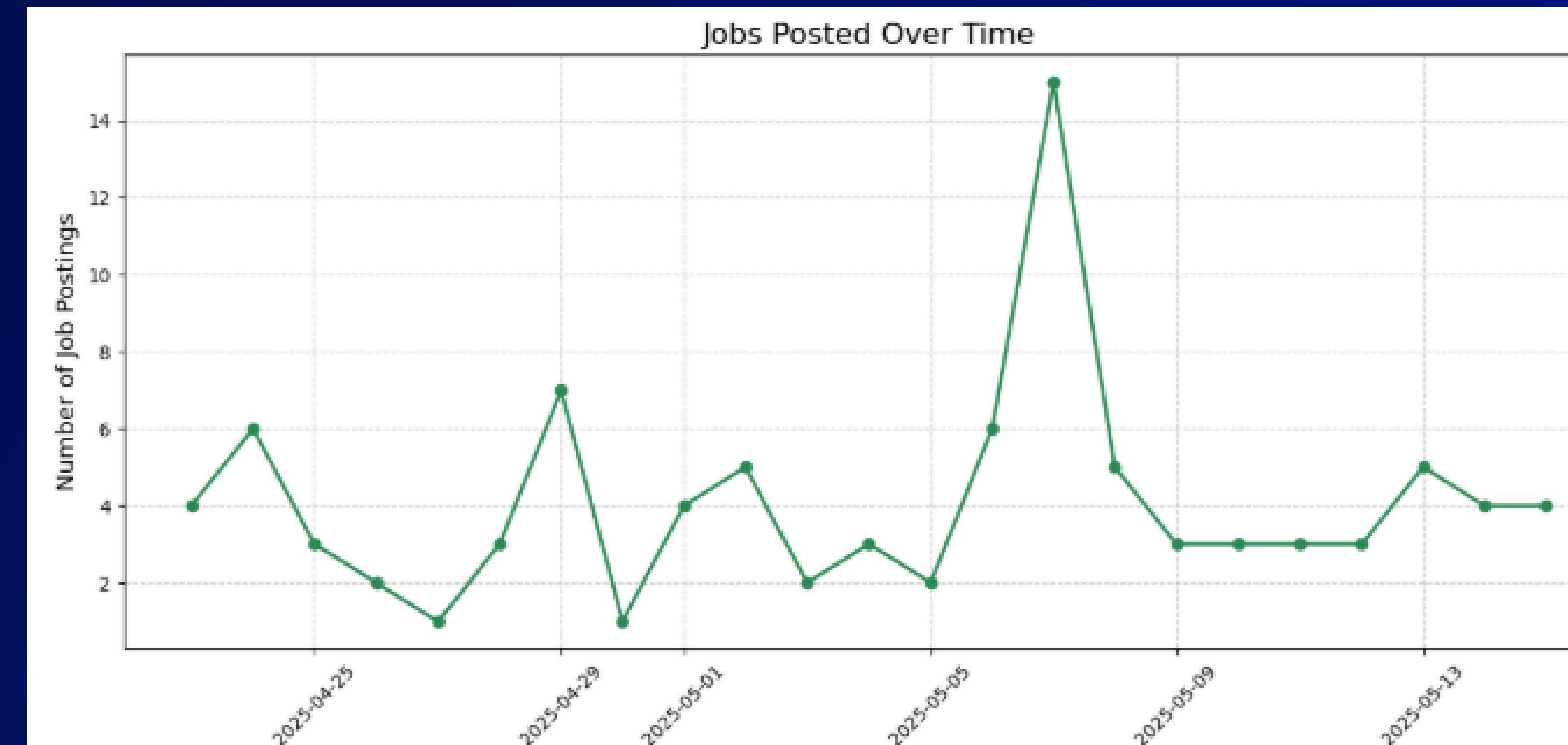
Key Insight:

- The difference between minimum and maximum pay varies, and might be larger for higher-paying jobs.
- There are some unusual cases with surprisingly high or low maximum salaries for their minimum.
- Many of the represented jobs fall within a mid-range salary band.



Key Insight:

- There were no sudden surges or drops in the number of postings on any particular day.
- This steady flow suggests that either job opportunities were scraped consistently, or the companies themselves have adopted a routine approach to publishing vacancies.



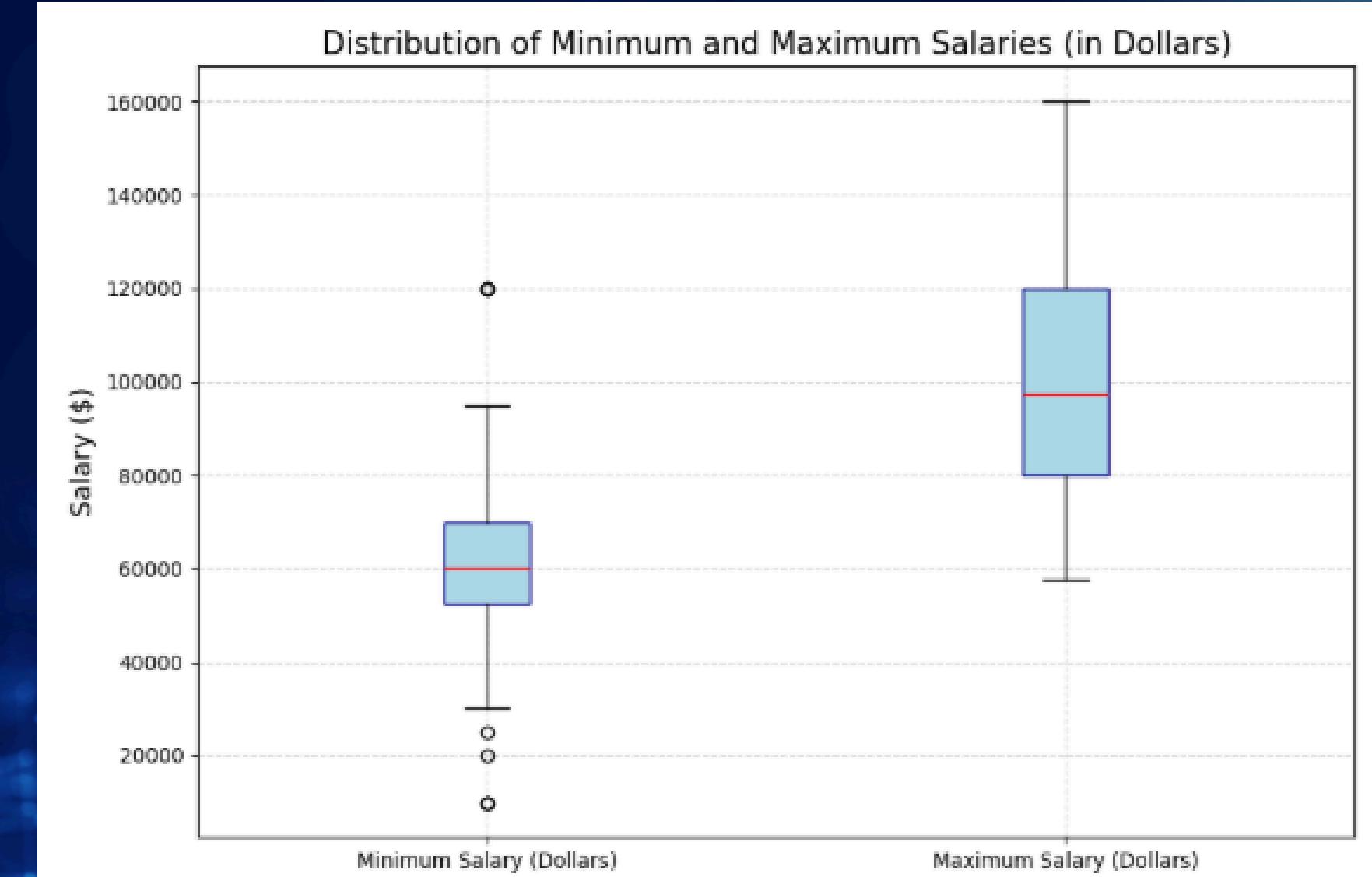
Key Insight:

- The location data indicates a strong preference for remote work arrangements. A substantial share of job postings either explicitly list "Remote" or use the term "Worldwide" as the location. This trend is in line with broader global workforce shifts, where flexibility and remote collaboration are becoming standard expectations.
- Further inspection of geographic mentions shows that the United States appears frequently, even in listings marked as remote. This suggests that while jobs are open to global applicants, they may still be anchored in North American market expectations, whether in terms of time zone alignment, salary structure, or company registration.



Key Insight:

- The most frequent minimum salary range falls between ₹50,000 and ₹70,000 per month.
- The most common maximum salary range is slightly higher, typically between ₹70,000 and ₹90,000.



CONCLUSION

This project analyzed remote job listings from the RemoteOK portal using EDA, Python, and SQL. We identified top job roles, hiring companies, salary trends, and location patterns. Most jobs were in tech, with high demand for software and data-related roles. Companies like Contra and Smallstep emerged as frequent recruiters. While many listings were location-flexible, salary data showed inconsistencies, underlining the importance of data cleaning. Overall, the analysis offered meaningful insights into remote job trends and highlighted the power of combining SQL and Python for real-world data exploration.

Thank You