DS164 LABORATORY

UNLOCKING SALARY TRENDS: PREDICTIVE INSIGHTS FOR SOFTWARE PROFESSIONALS

CARREON | LUCAS | PERALTA | QUIPIT

Final Project

INTRODUCTION

This analysis explores salary trends for software professionals, focusing on factors such as company names, job titles, locations, and employment statuses. By preprocessing the dataset and applying machine learning models, we aim to **uncover key patterns and trends that will aid professionals and companies in making informed decisions**. The study includes visualizations of job distributions, salary averages, and top-paying companies, and evaluates model performance to enhance predictions. These insights are intended to support effective salary negotiations and strategic compensation planning in the competitive job market.

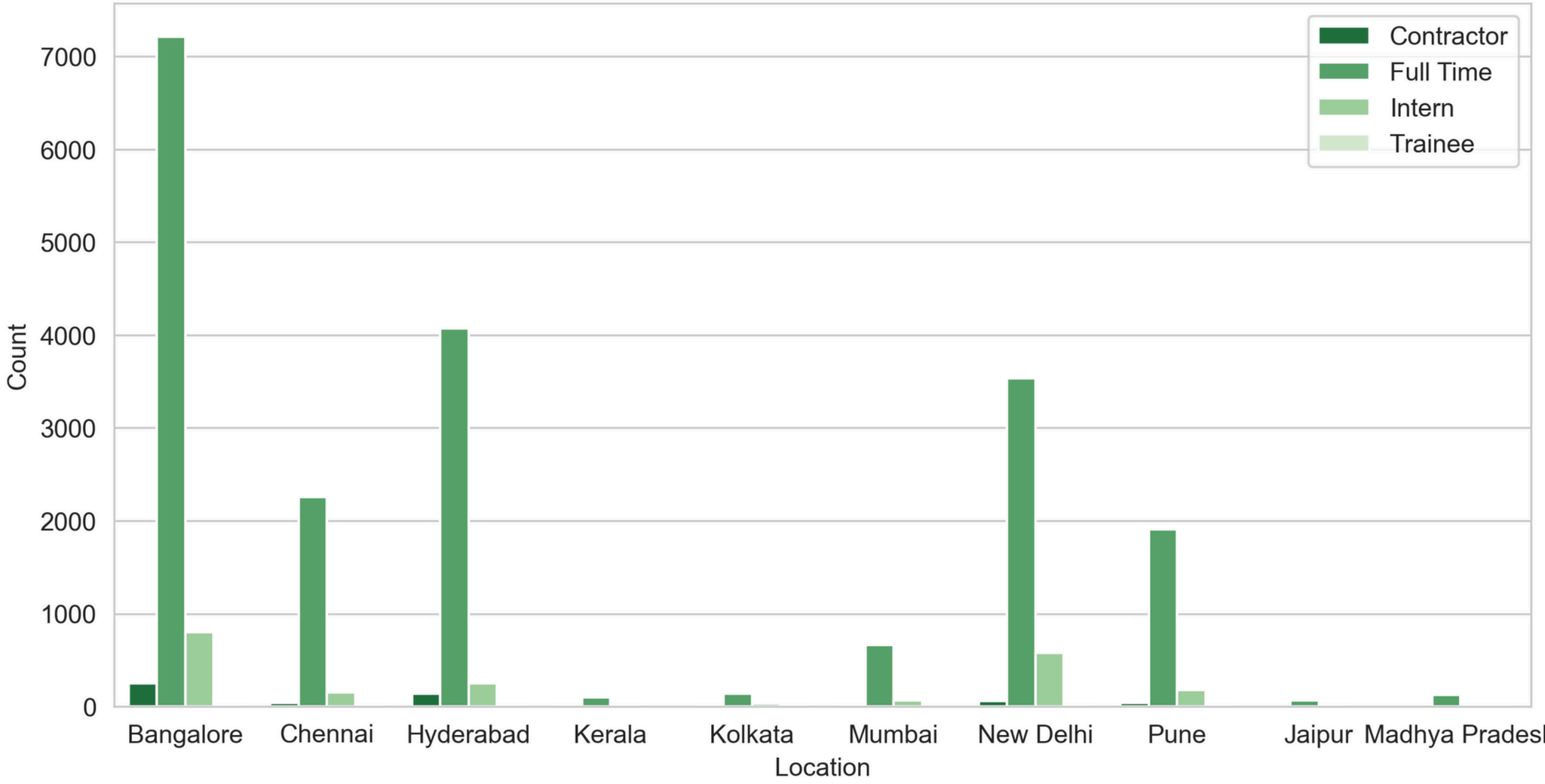
DATA DICTIONARY

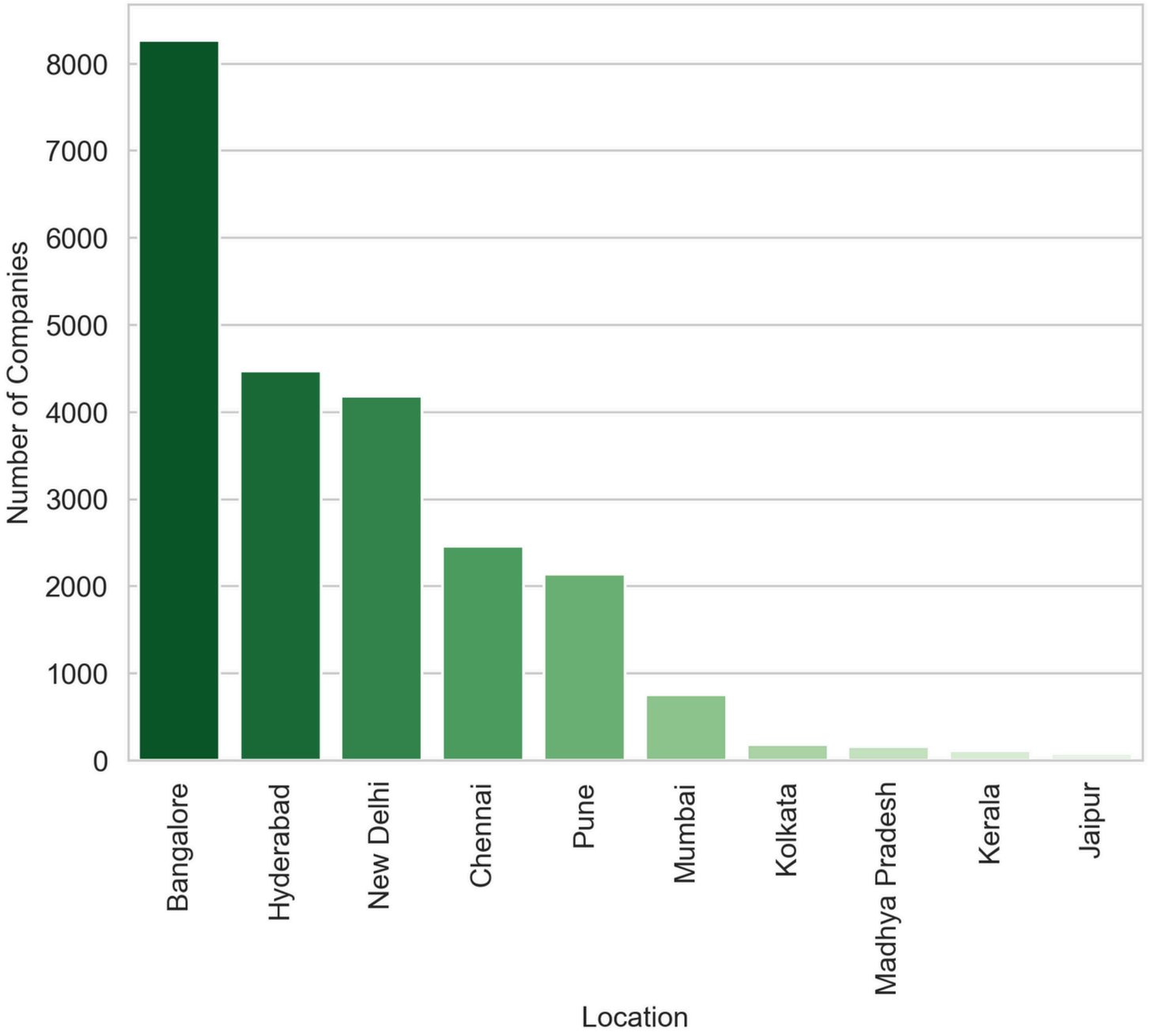
Attribute	Description
Rating	The rating or score given to the company or job title, reflecting overall satisfaction or reputation.
Company_Name	The name of the company where the job is located.
Job_Title	The title or position of the job within the company.
Salary	The salary amount for the job title at the company, expressed in INR (Indian Rupees).
Salaries_Reported	The number of individual salary reports for the job title at the company.
Location	The geographic location or city where the job is based.
Employment_Status	The type of employment, such as Full Time, Part Time, or Contractor.
Job_Roles	The specific roles or functions associated with the job title within the company.

WHAT ARE THE KEY FACTORS INFLUENCING SOFTWARE PROFESSIONAL SALARIES ACROSS DIFFERENT LOCATIONS AND JOB TITLES, AND HOW CAN PREDICTIVE MODELS BE USED TO ESTIMATE SALARIES BASED ON THESE FACTORS?

EMPLOYMENT STATUS BY LOCATION

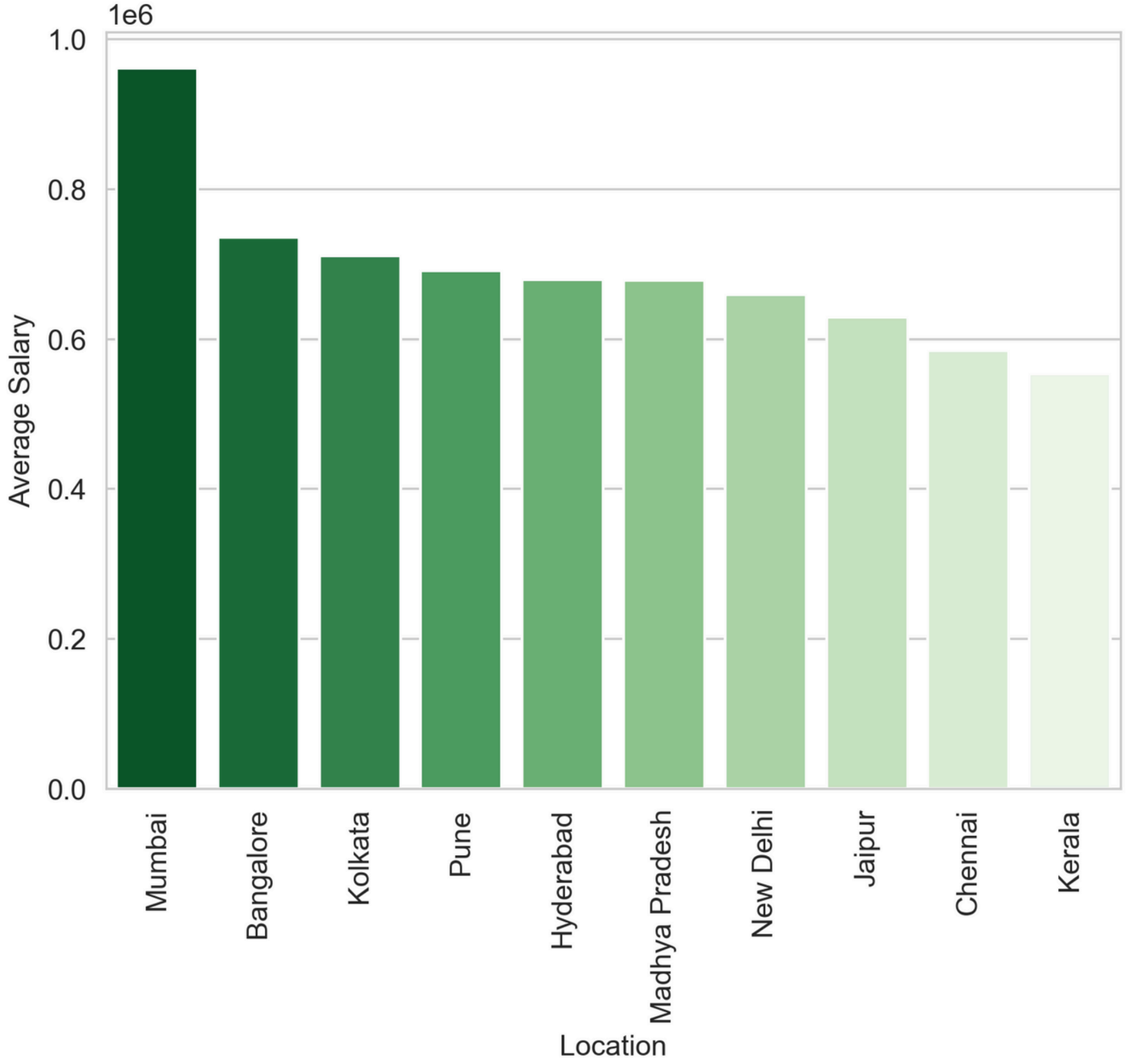
The bar chart reveals that **full-time employment is predominant in Indian cities**, with **Bangalore leading at about 6500 full-time employees**, followed by Hyderabad with 4500. Other cities like Chennai, New Delhi, and Pune also have substantial full-time numbers, while Mumbai has around 1000. Contractor, Intern, and Trainee roles are less common, with Bangalore having the most contractors. Overall, ***full-time roles are concentrated in major cities***, with other employment statuses being less frequent.





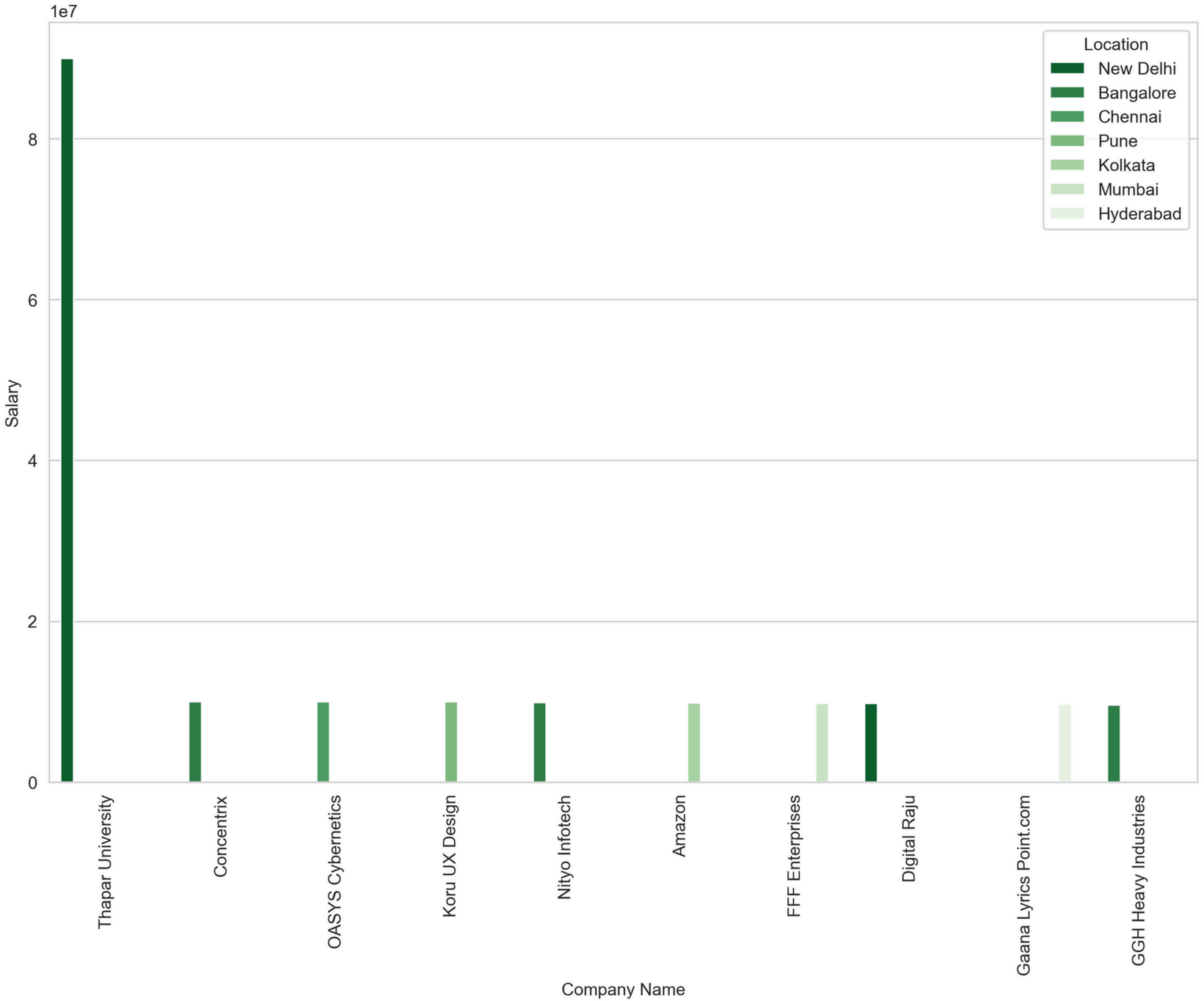
COMPANIES BY LOCATION

The bar graph shows **Bangalore as the leading city with nearly 8000 companies**, far surpassing others. Hyderabad and New Delhi follow with about 4500 and 4000 companies, respectively. Chennai has around 3000, Pune about 2500, and Mumbai around 1000. Other cities like Kolkata, Madhya Pradesh, Kerala, and Jaipur have fewer than 1000 companies each. ***Bangalore stands out as the primary business hub***, with Hyderabad and New Delhi also hosting significant numbers of companies.



AVERAGE SALARY BY LOCATION

The bar graph shows ***Mumbai with the highest average salary at nearly 1,000,000 units***. Bangalore and Kolkata follow with average salaries of around 800,000 units. Pune, Hyderabad, and Madhya Pradesh have average salaries of about 700,000 units. New Delhi and Jaipur average slightly above 600,000 units, while **Chennai and Kerala have the lowest** at around 500,000 units.

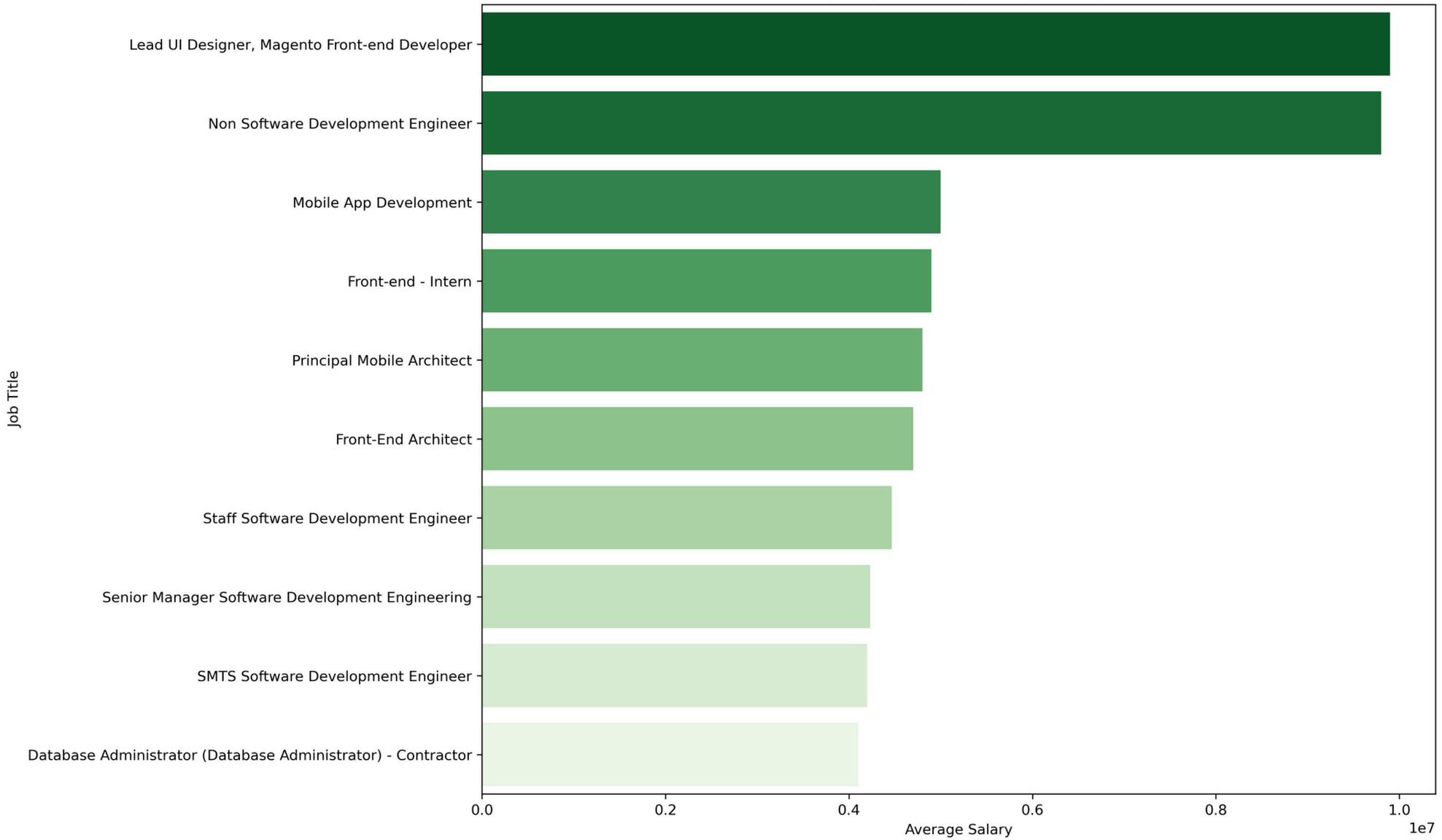


TOP 5 SALARIES BY COMPANY FOR EACH LOCATION

The bar graph shows that ***Thapar University in New Delhi offers the highest salary at 9 million***, while companies like Concentrix, OAYSIS Cybermetics, and others offer 1 million. **Thapar University stands out for its significantly higher compensation**, with other companies providing competitive but lower salaries. Bangalore appears frequently, indicating its strong employment market.

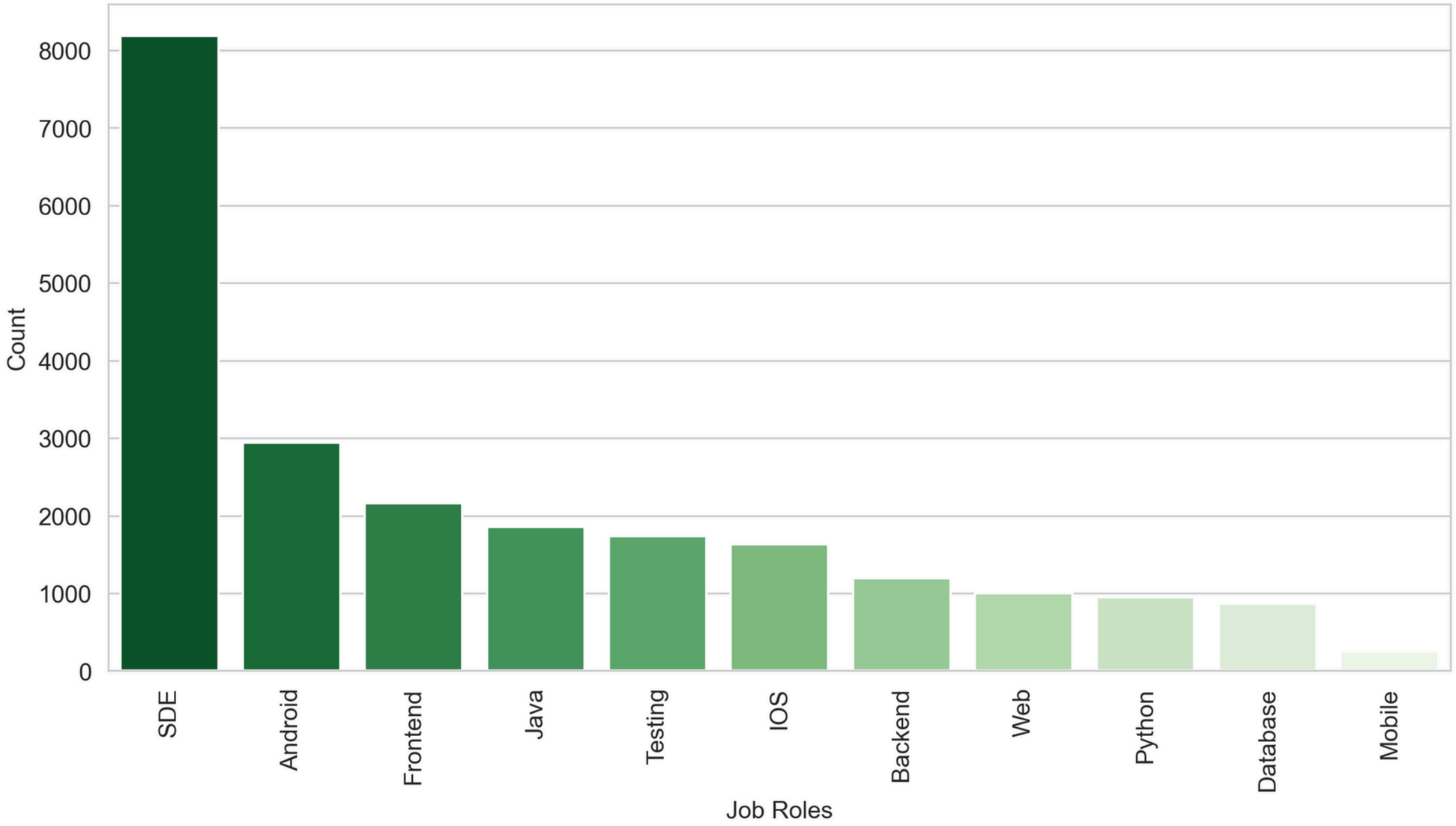
TOP 10 JOB TITLES BY AVERAGE SALARY

Lead UI Designers and Magento Front-end Developers top the salary list at 10 million, with Non-Software Development Engineers close behind at 9.8 million. SDEs are the most common job, with nearly 8000 positions, followed by 3000 Android developers.



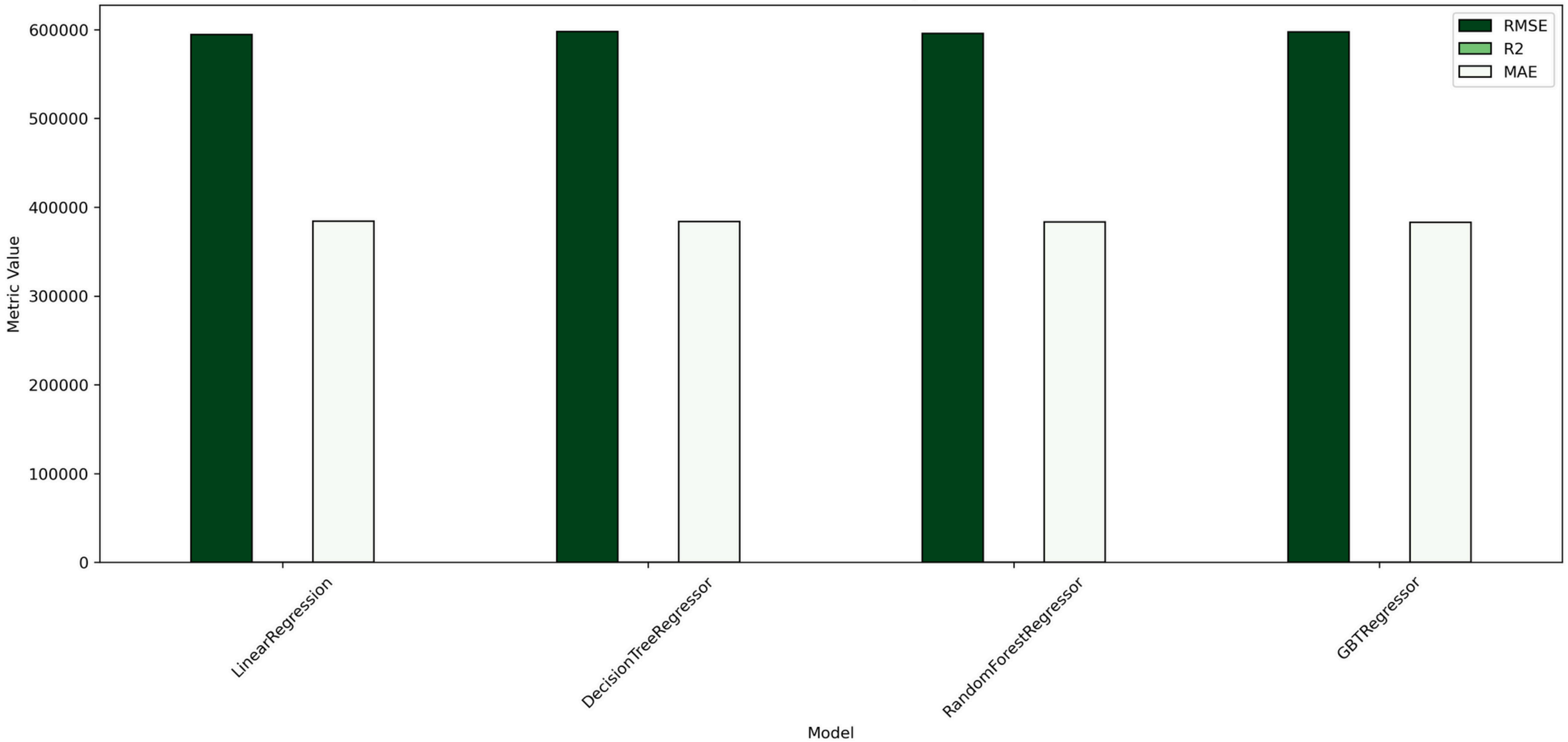
COUNT OF JOB ROLES

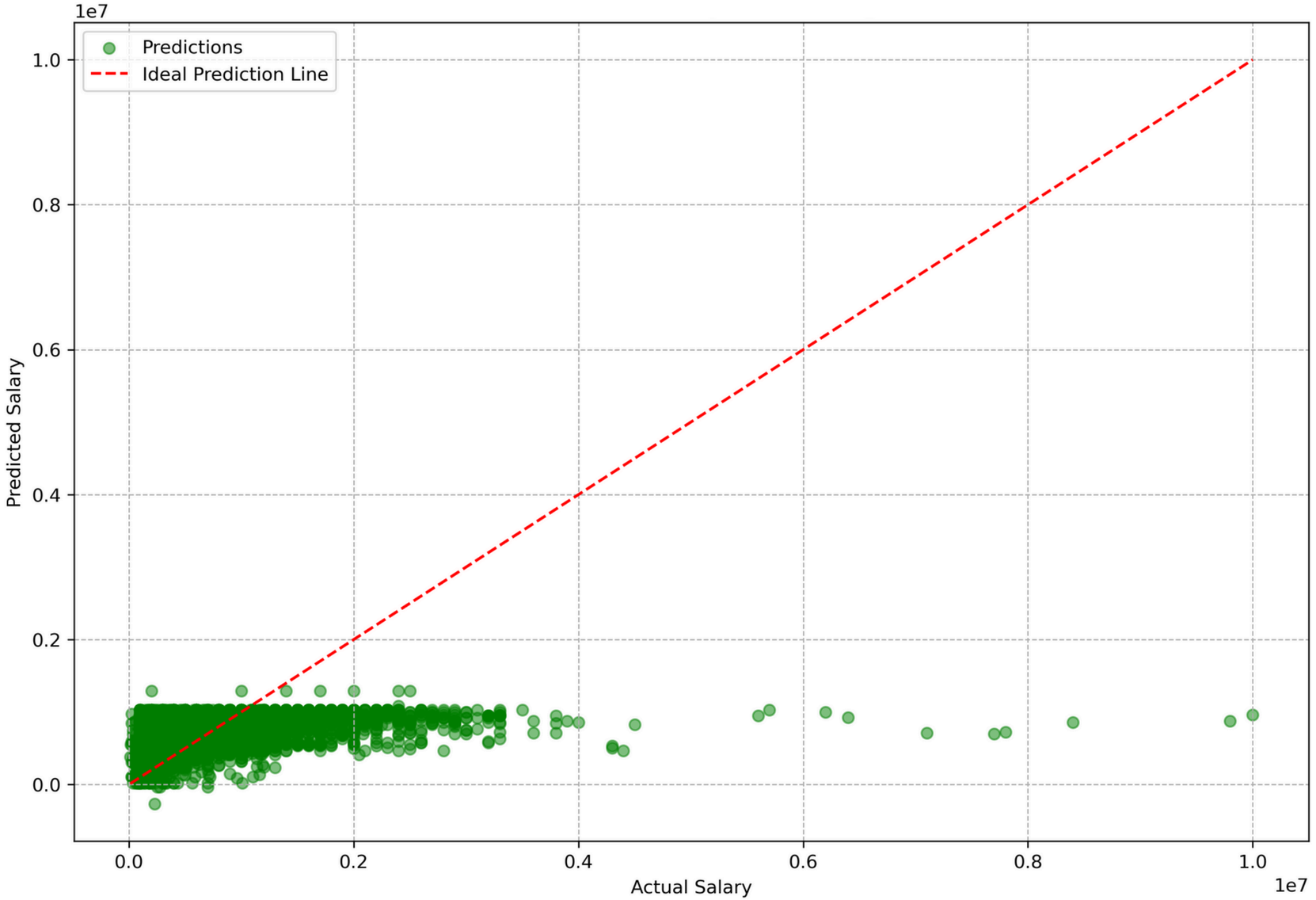
It is evident that the job role, ***SDE, has gained the highest count of job roles with over 8000***. Meanwhile, the other job roles obtained around three thousand and lower in count. This means that the job role SDE has the most employee roles.



MODEL PERFORMANCE COMPARISON

The graph shows that Linear Regression, Decision Tree, Random Forest, and GBRT Regressor have similar performance. **All models have comparable RMSE and MAE** values around 600,000 and 400,000, respectively, indicating **no significant differences in accuracy**. R-squared values are lower, suggesting none of the models excel in explaining data variability. Overall, ***the models perform similarly in terms of prediction accuracy.***





ACTUAL VS PREDICTED SALARY FOR LINEAR REGRESSION

The scatter plot shows that the Linear Regression model underestimates salaries. The red dashed line indicates ideal predictions, but the green dots (predictions) are clustered below this line, especially for higher salaries. This deviation suggests that the ***model fails to accurately predict higher salary values and may be underfitting the data***.

CONCLUSION

In summary, the analysis shows full-time jobs dominate major cities, with Bangalore, Hyderabad, and New Delhi leading. Mumbai offers the highest salaries, while Chennai and Kerala have lower pay. Specialized roles in UI/UX and front-end development are the highest paying. Regression models are accurate but struggle with high salaries, suggesting a need for improved predictions. Recommendations include adjusting compensation strategies and refining models to better predict higher salaries.

THANK YOU

Chauhan, A. (2023, July 9). 22700+ software Professional Salary Dataset. Kaggle. https://www.kaggle.com/datasets/whenamancodes/software-professional-salary-dataset/code