**Exploratory Data Analysis (EDA) Report**

**Data set: Glassdoor\_Job\_Postings**

**1. Data Overview**

* The dataset contains **900 job postings** with 18 attributes, including company details, job roles, salary estimates, and ratings.
* Missing values exist in **company ratings, salary estimates, and company-founded year**.
* Data cleaning involved converting salary estimates to numeric, handling missing values, standardizing revenue categories, and mapping job locations.

**2. Key Findings**

**2.1 Company Ratings & Salaries**

* Most companies have ratings between **3.0 and 4.5**, with a slight right-skew.
* Salary estimates are **right-skewed**, with most jobs offering **₹3-5 lakhs per year**.
* Larger companies (10,000+ employees) tend to offer **higher median salaries**, while smaller firms show more variation.
* Some industries, such as **finance and technology**, consistently offer higher salary estimates.
* Companies with higher ratings tend to have **lower employee turnover**, making them attractive for long-term career growth.

**2.2 Job Roles & Locations**

* **Top Job Titles**: Data Scientist, Data Analyst, and AI/ML-related roles are in high demand.
* **Top Locations**: Bengaluru leads in job postings, followed by other Indian tech hubs like Mumbai and Hyderabad.
* Remote job postings are increasing but still form a small portion of the dataset.
* The demand for AI and Machine Learning specialists is growing, with salaries for these roles outpacing traditional data analyst roles.

**2.3 Industry & Revenue Trends**

* **Top Industries Hiring**: Banking & Lending, IT Services, Healthcare, and Consulting.
* **Revenue Distribution**: Most postings are from **large enterprises (10B+ revenue)**, indicating that established firms dominate hiring.
* Mid-sized companies (100M-1B revenue) are expanding their hiring efforts but still lag behind larger enterprises.
* **Startups in emerging tech fields**, such as AI and blockchain, are increasing hiring efforts despite having lower revenue compared to established firms.

**2.4 Correlation Analysis**

* **Salary & Company Ratings**: No strong correlation, suggesting **high-rated companies do not necessarily offer higher salaries**.
* **Career Growth & Culture Ratings**: Strong correlation (~0.75), implying **companies with better career opportunities tend to have a positive work culture**.
* **Work-Life Balance & Compensation**: Moderate correlation, meaning **higher-paying jobs may offer better work-life balance**.
* **Company Size vs. Salary**: Larger firms tend to offer more competitive salaries, but startups offer greater role flexibility and growth potential.
* **Employee Satisfaction & Career Growth**: Companies investing in employee training and development have **higher job satisfaction ratings**, making them more desirable employers.

**3. Business Recommendations**

* **For Job Seekers**:
  + **Consider larger companies** for higher salaries.
  + Look at companies with **high career opportunities ratings**, as they align with better work culture.
  + **Finance & Tech sectors** have the most job openings.
  + Consider startups for diverse role exposure and career growth opportunities.
  + If location flexibility is a priority, explore **remote job options** where available.
  + **Target roles in AI and emerging tech fields**, as they offer competitive salaries and career growth.
* **For Employers & Recruiters**:
  + Competitive salaries are important, but **career growth and work culture matter more** to attract top talent.
  + **Expanding hiring outside Bengaluru** could balance job demand across regions.
  + Small-to-mid-sized firms need to **differentiate through benefits & work-life balance** since they compete with high-revenue firms.
  + Investing in **remote work policies** could help attract a broader talent pool.
  + Increasing transparency in salary ranges and growth prospects could improve applicant engagement.
  + **Providing upskilling opportunities** and professional development can lead to higher employee satisfaction and retention.

**Initial Findings from the Dataset**

The dataset contains **900 job postings** with **18 columns**, covering aspects such as **company details, job information, salary estimates, ratings, and industry sector**.

**Key Observations:**

1. **Missing Values:**
   * company\_rating, salary\_avg\_estimate, and company\_size have **missing values**.
   * job\_description is missing for **12 entries**.
   * Ratings (career\_opportunities\_rating, comp\_and\_benefits\_rating, etc.) are missing in some rows.
2. **Data Types:**
   * Most columns are **strings (object type)**, but ratings and company-founded year are numerical.
   * salary\_avg\_estimate is stored as a string instead of numeric.
3. **Potential Issues:**
   * salary\_avg\_estimate needs cleaning (removing currency symbols and converting to numeric).
   * company\_founded contains "--", which needs conversion to NaN or appropriate handling.
   * revenue column contains text like **"$10+ billion (USD)"**, which may need transformation.

**Next Steps:**

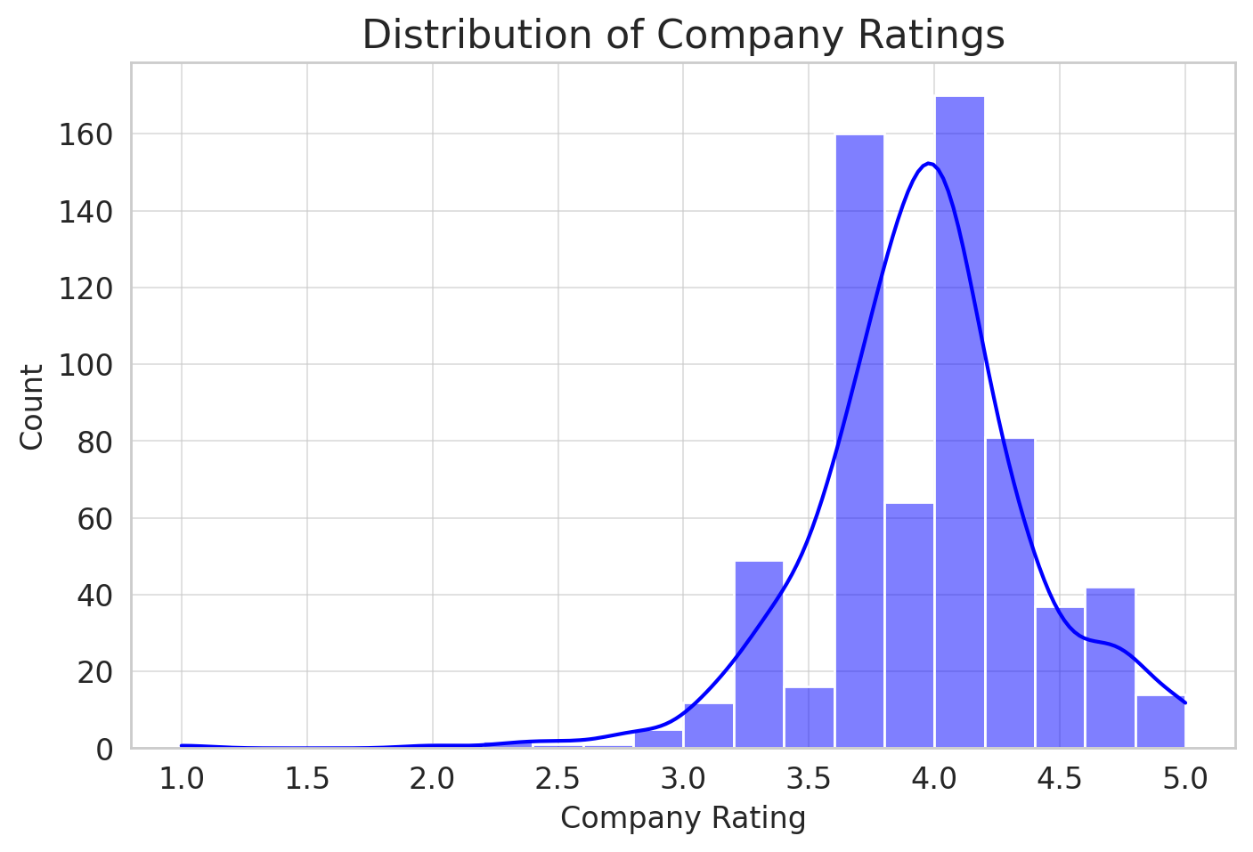
✅ Handle missing values and clean the data.  
✅ Convert numerical columns stored as strings into the correct format.  
✅ Perform exploratory visualizations (e.g., salary distribution, company ratings, job locations).

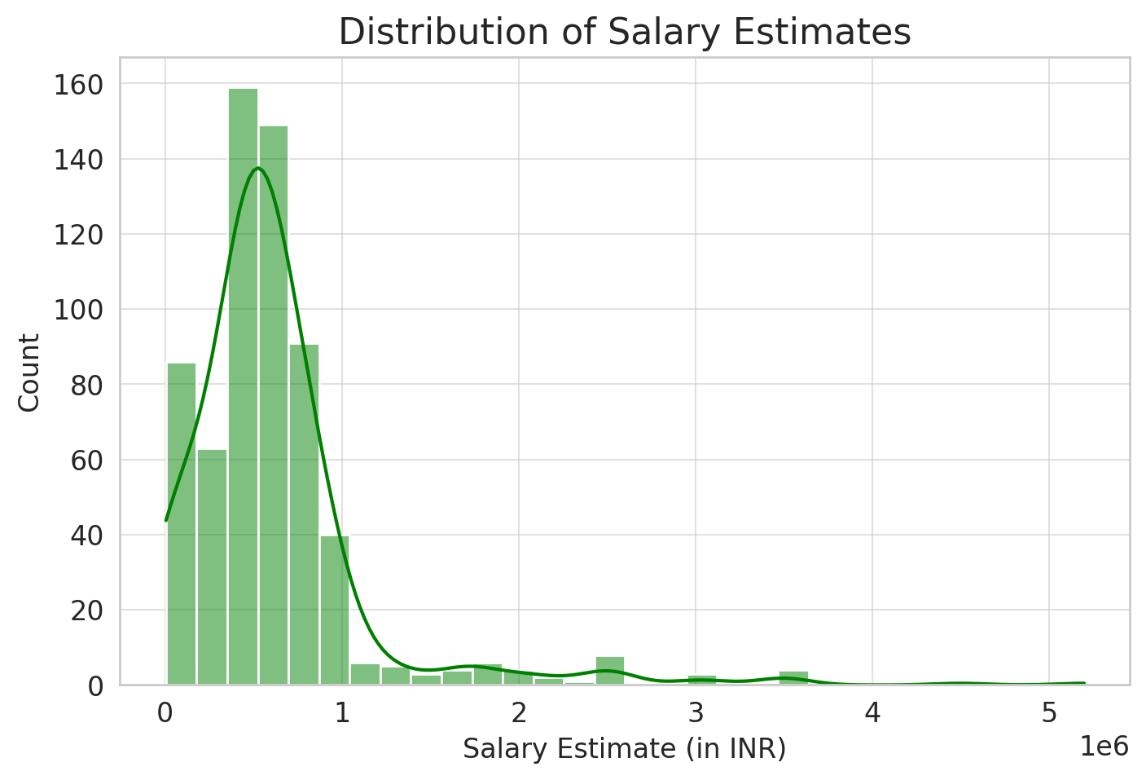
**Data Cleaning Summary**

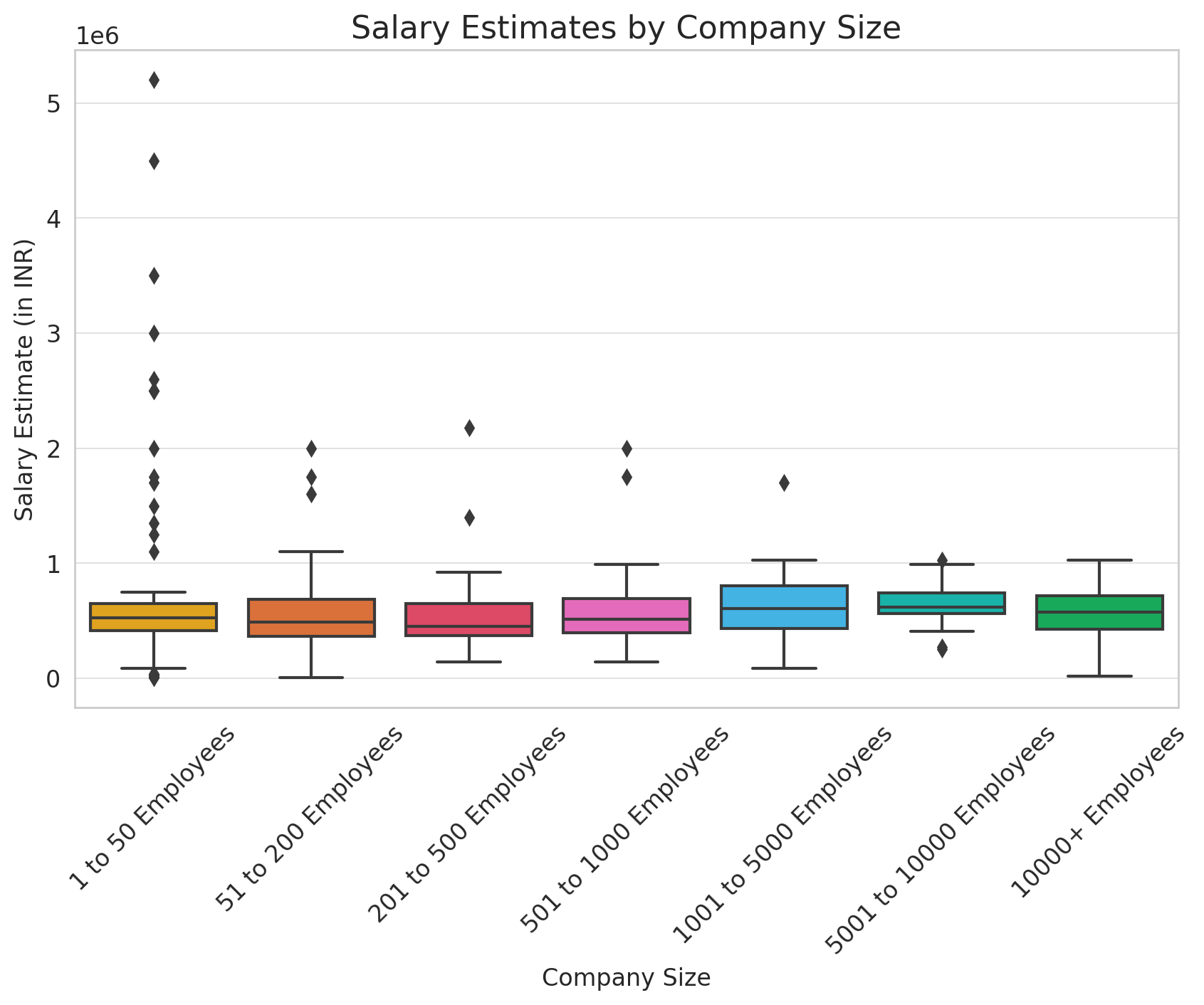
✅ **Salary Data Cleaned**: Converted salary\_avg\_estimate into numeric format.  
✅ **Company Founded Year**: Replaced invalid values ("--") with NaN.  
✅ **Revenue Column Standardized**: Mapped revenue ranges into categories (e.g., "10B+", "1-5M").

**Summary statistics and visualizations! 🎨📊 ​​**

* **Visual Analysis Insights**







📌 **Company Ratings Distribution**

* Most companies have ratings **between 3.0 and 4.5**, with a slight right-skew.
* Few companies have extremely low or high ratings.

📌 **Salary Estimates Distribution**

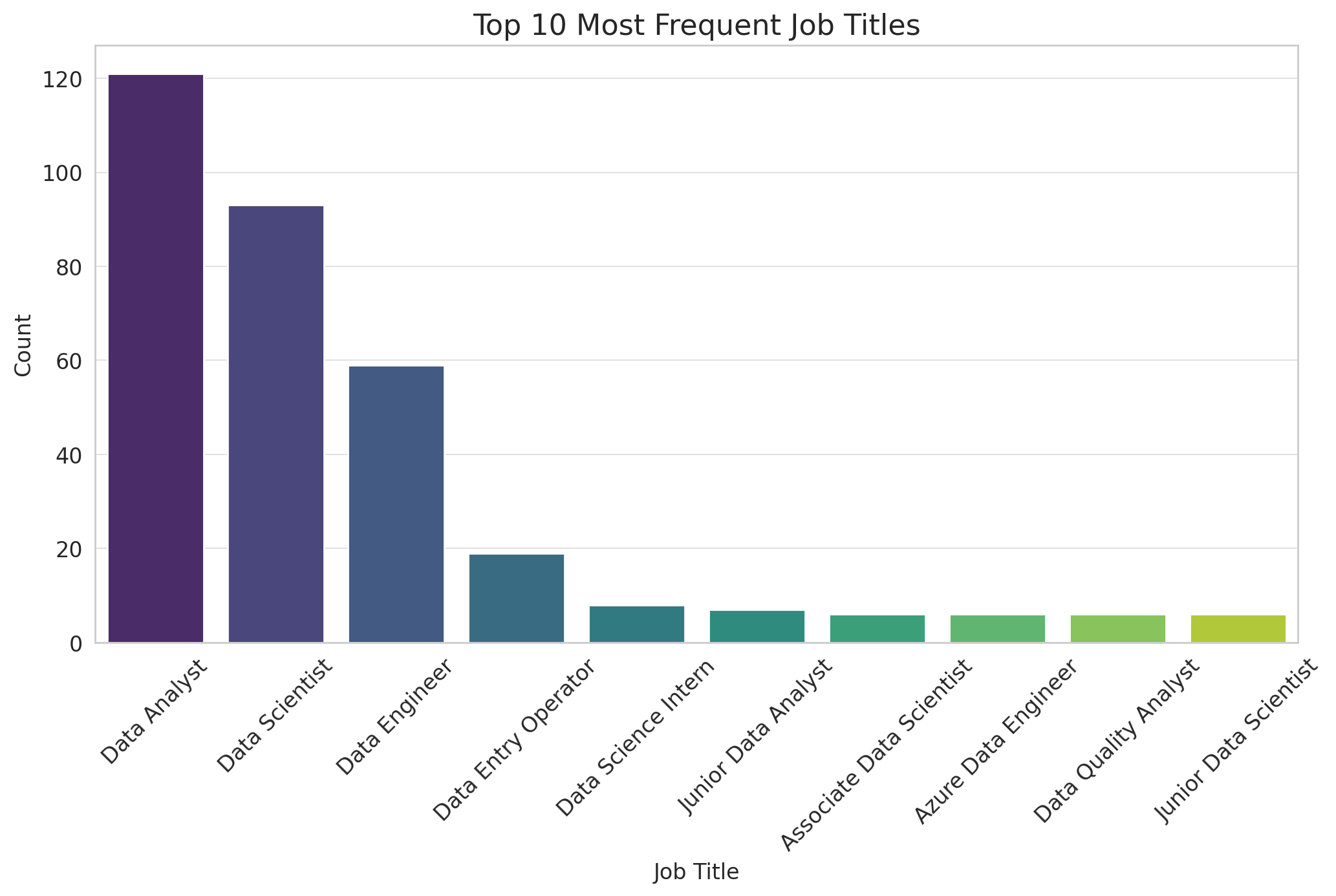
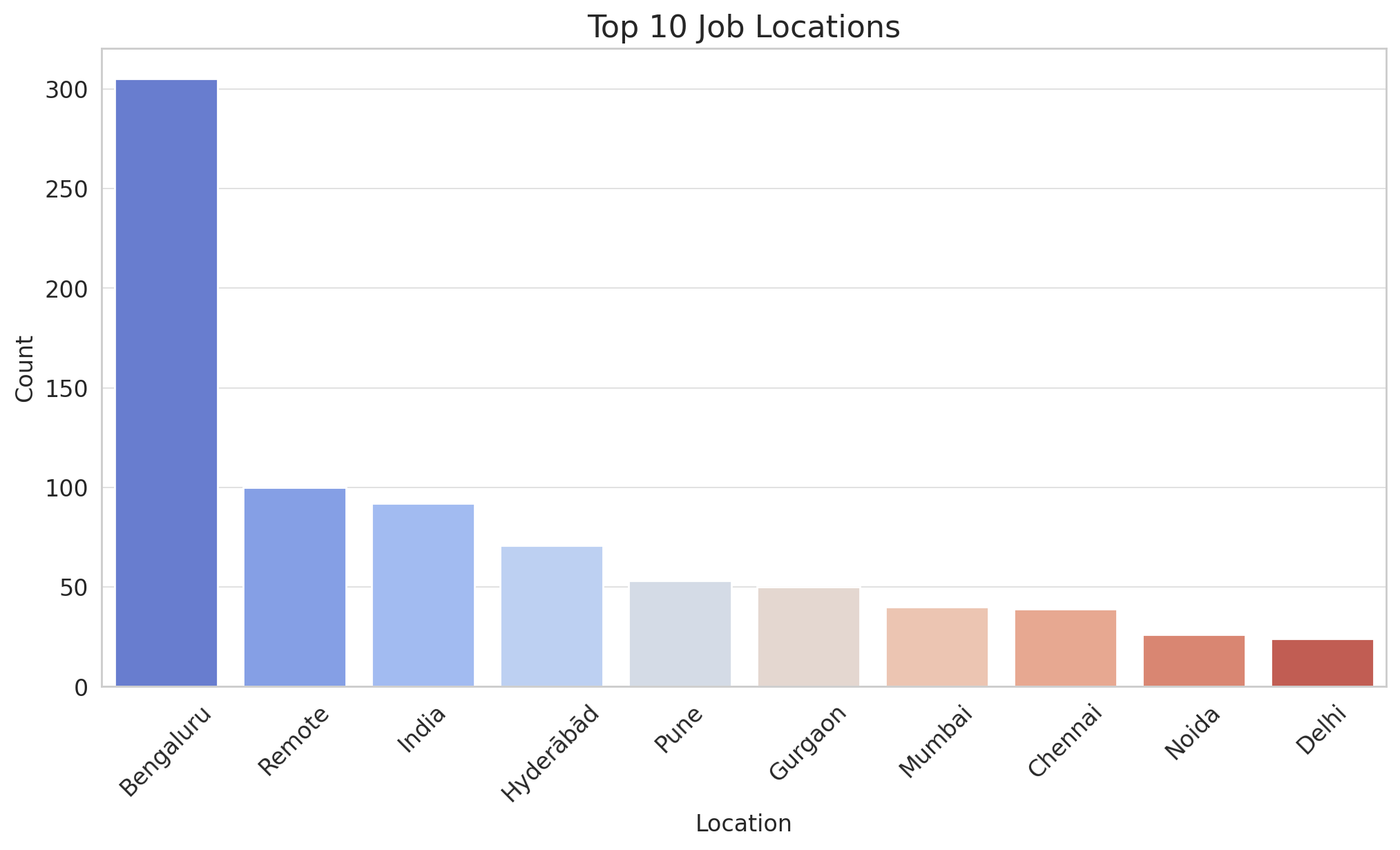
* Salary estimates are **right-skewed**, meaning most job postings have lower salaries, with fewer high-paying roles.
* A peak is observed in the **3-5 lakh INR range**.

📌 **Salaries by Company Size**

* **Larger companies (10000+ employees)** tend to offer **higher median salaries**.
* Smaller companies (1-50 employees) have **wider variation** in salaries.

​​

* **Job Roles & Location Insights**

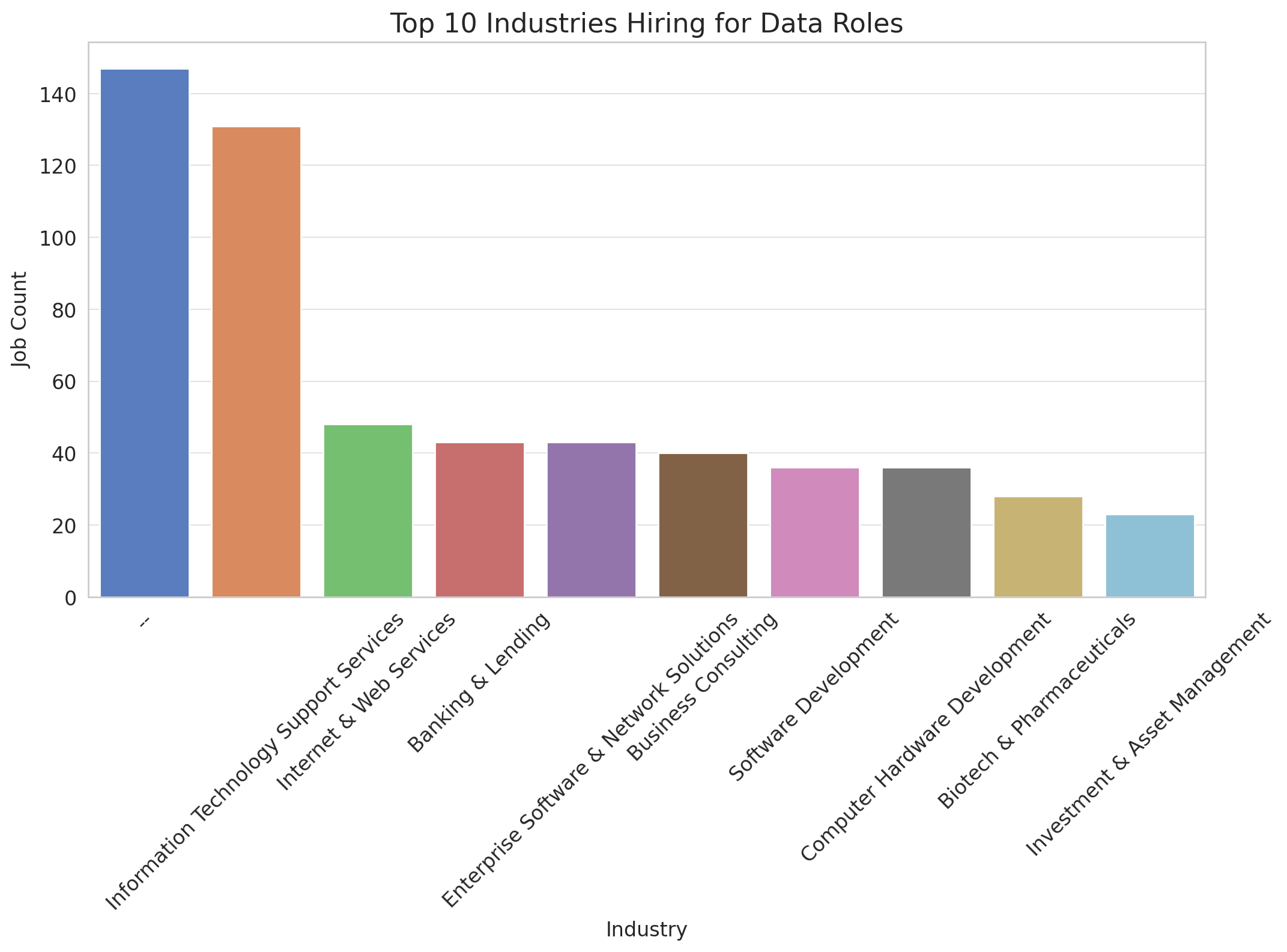
****

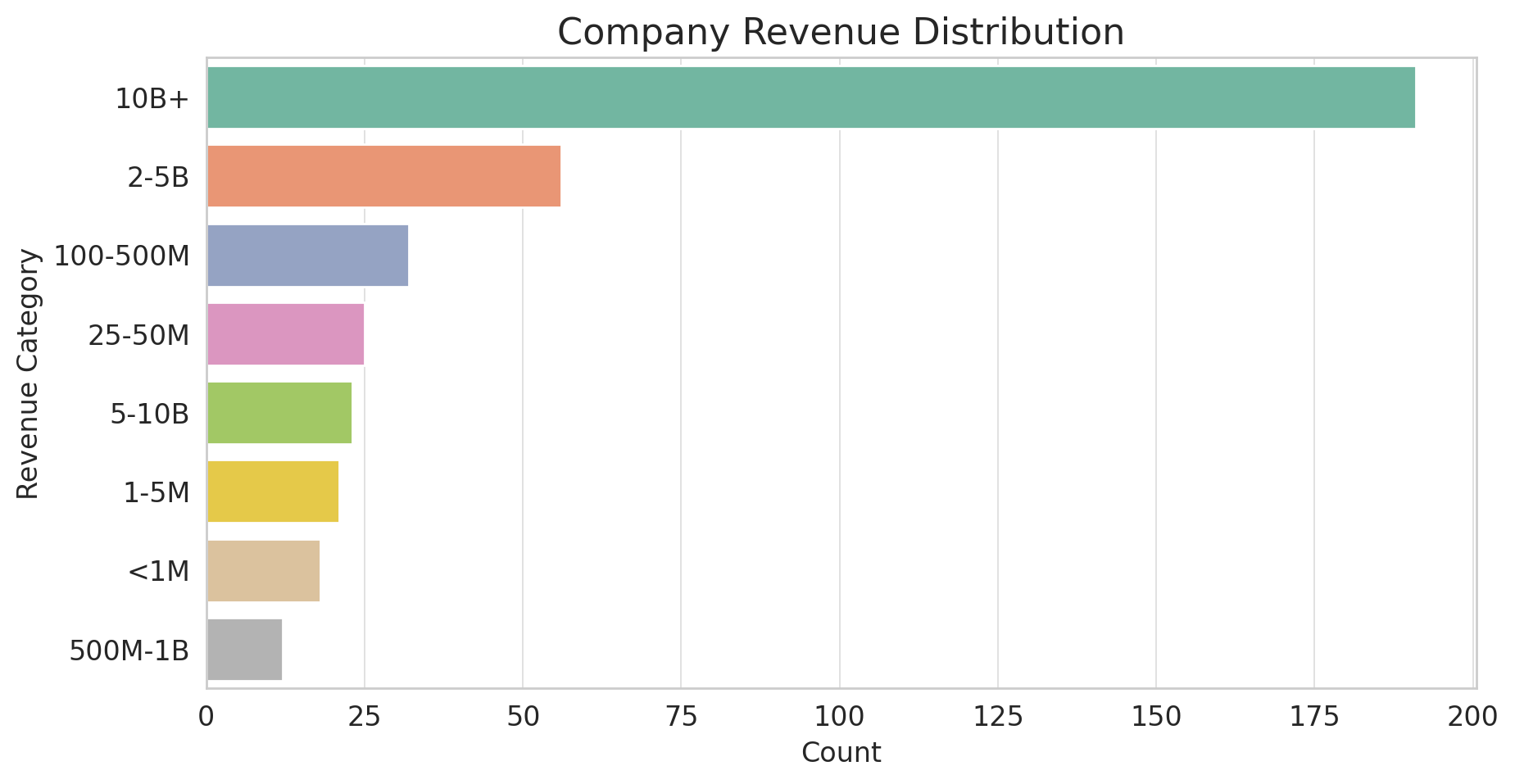
📌 **Top Job Titles**

* The most common roles include **Data Scientist, Data Analyst, and AI/ML roles**.
* There is a high demand for analytical and machine learning-related positions.

📌 **Top Job Locations**

* The highest number of job postings are in **Bengaluru**, followed by other major Indian cities.
* **Bengaluru** being the tech hub dominates the job market.
* **Industry & Revenue Insights**

****

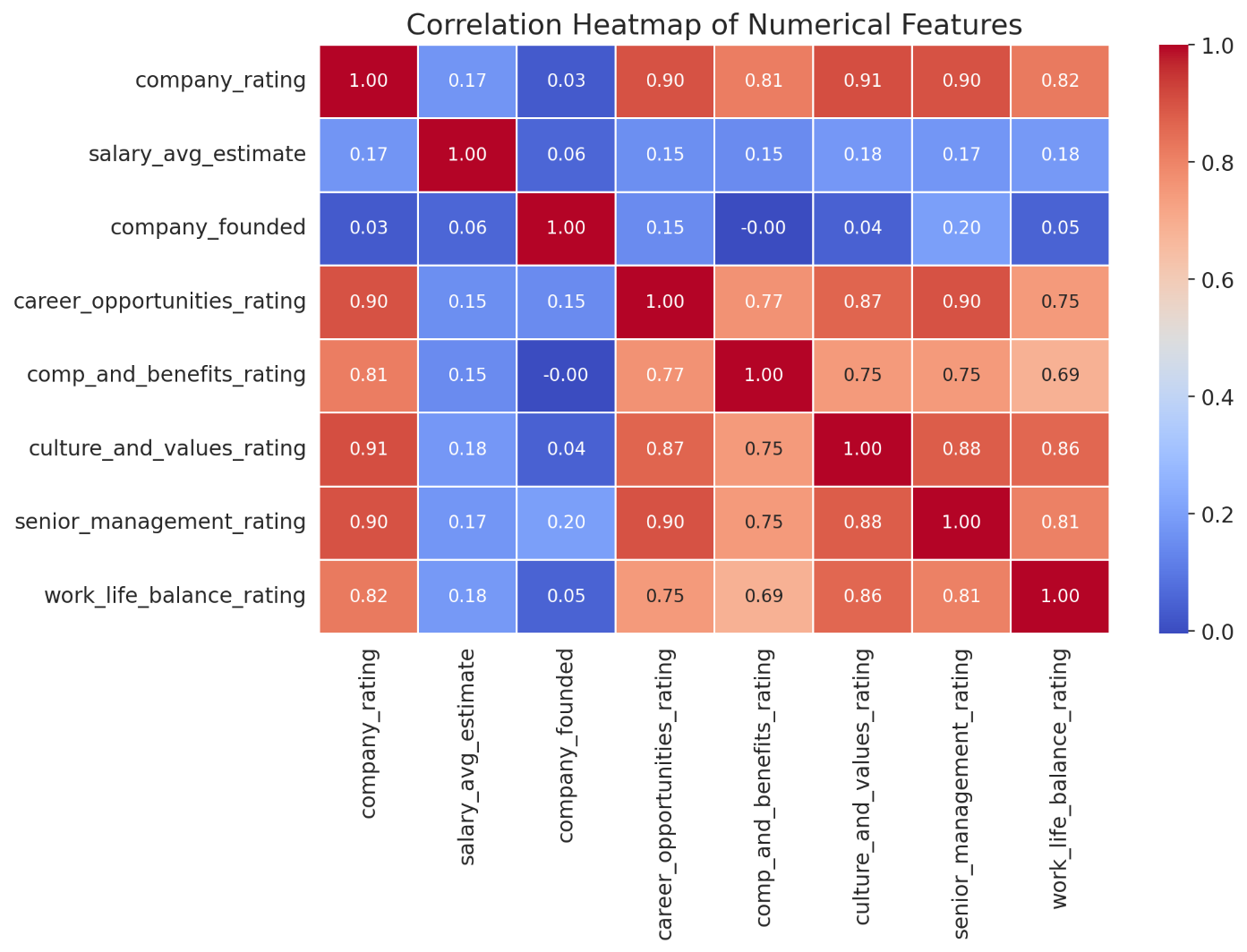
****

📌 **Top Hiring Industries**

* **Banking & Lending, IT Services, Healthcare, and Consulting** are the dominant industries for data roles.
* **Finance & Tech sectors** have the highest demand for data-related positions.

📌 **Company Revenue Distribution**

* Most job postings come from companies with **$10B+ in revenue**, indicating **large enterprises** dominate hiring.
* There are fewer postings from **small to mid-sized companies (<$50M revenue)**.
* **Correlation Analysis Insights**

****

📌 **Key Observations from the Heatmap**

* **Salary Estimate vs. Company Ratings**: Weak correlation, meaning **higher-rated companies do not necessarily pay more**.
* **Career Opportunities & Culture Ratings**: Strong correlation (~0.75), indicating that **companies offering career growth also have a strong workplace culture**.
* **Work-Life Balance & Compensation**: Moderate positive correlation, suggesting that **better compensation may align with better work-life balance**.

**4. Conclusion**

The analysis of Glassdoor job postings reveals key trends in hiring, salaries, and company culture. Larger companies tend to dominate job postings, especially in the finance and tech sectors, offering more stability and competitive pay. However, startups in emerging tech fields are growing, providing exciting opportunities for innovation and career growth. While salary remains a key factor for job seekers, career growth and work-life balance play an equally important role in employee satisfaction. Companies that invest in career development and foster a strong workplace culture tend to attract and retain top talent. As the job market evolves, both job seekers and employers must adapt by embracing technological advancements, remote work opportunities, and continuous learning initiatives.