

# PROJECT REPORT

## COMPUTER SCIENCE JOBS DATA ANALYSIS

A Web Scraping & Data Visualization Project using Python

Created By:

Somil Dogra

GF202347974

BTech CSE (Data Science)

2023-2027

Date: July 2025

GitHub: <https://github.com/somil-18>

# INDEX

<b>Serial Number</b>	<b>Title</b>	<b>Page Number</b>
1.	Project Title	3
2.	Objective	3
3.	Tools & Technologies Used	3
4.	Dataset Description	3 - 4
5.	Data Cleaning & Preprocessing	5 - 7
6.	Data Visualization & Insights	8 - 31
7.	Conclusion	32 – 34
8.	Limitations	35
9.	Reference	35

## **I. Project Title:**

Analysing computer science job trends on Internshala platform using web scrapping & data visualization.

## **II. Objective:**

- To scrape computer science job listings from Internshala.
- To clean and process the collected data.
- To perform visual analysis to undercover trends related to company, job types, locations, experience, salary.

## **III. Tools & Technologies Used:**

- Programming Language Used: Python.
- Python Libraries: requests, beautifulsoup, numpy, pandas, matplotlib, seaborn.
- Platform Used: Jupyter Notebook.
- Optional Tools: Excel (for data verification).

## **IV. Dataset Description:**

The dataset was created by scraping job listings from the "Computer Science" category on Internshala, followed by a data cleaning process to prepare it for analysis.

<u>Column Name</u>	<u>Description</u>
Title	Name of the job
Company	Name of the company
Location	Location of the job
Salary	Salary offered (if available)
Time (days ago)	Days after the job is posted
Experience (years)	Years of experience for the job

- Additional Columns:

<u>Column Name</u>	<u>Description</u>
Location Type	Known or Unknown location
Job Count	Total no. of jobs
Salary Type	Disclosed or Competitive salary
Field	Job field
Employee Type	Fresher or Experienced

## V. Data Cleaning & Preprocessing:

To ensure high-quality analysis, the raw scraped data was thoroughly cleaned and preprocessed using the following steps:

### 1. Column Inspection:

Checked available columns, data types, and summary statistics using:

- `df.columns`
- `df.dtypes`
- `df.describe()`

### 2. Handling Missing Values:

- Counted missing values: `df.isnull().sum()`
- Dropped rows with missing Salary values:  
`df.dropna(subset=["Salary"], inplace=True)`
- Filled missing values in the Location column with "Unknown":  
`df["Location"] = df["Location"].fillna("Unknown")`

### 3. Removing Duplicates:

- Checked for duplicates: `df.duplicated().sum()`
- Removed duplicate entries:  
`df.drop_duplicates(inplace=True)`

#### 4. Salary Parsing:

- Created a function `parse_salary()` to handle.
- Ranges like ₹ 2,50,000 - ₹ 3,00,000 → average of the two.
- Single values like ₹ 2,00,000 → parsed directly.
- Other formats → marked as "Competitive Salary".
- Applied the function:  

```
df["Cleaned_Salary"] = df["Salary"].apply(parse_salary)
```
- Dropped the original Salary column.

#### 5. Posting Time Cleaning:

- Used `clean_time()` to convert time descriptions like 2 weeks ago to equivalent days (e.g. 2 weeks ago → 14).
- Used `convert_to_days()` to convert cleaned strings into integers (e.g. 14, 0).
- Dropped intermediate columns after transformation.

#### 6. Experience Column Cleaning:

- Extracted numeric years from strings like "2 Years" using the `exp()` function (e.g. 2 year(s) → 2).
- Created a new column: "Cleaned\_Experience (Years)".
- Dropped the original Experience (Years) column.

## 7. Splitting Multiple Locations:

- Some postings had multiple comma-separated locations.
- Split these values and exploded them into separate rows:

```
df['Location'] = df['Location'].dropna().apply(lambda x:  
x.split(','))  
df = df.explode('Location').reset_index(drop=True)
```

## 8. Saving Cleaned Data:

- Exported the final cleaned DataFrame to a CSV file:

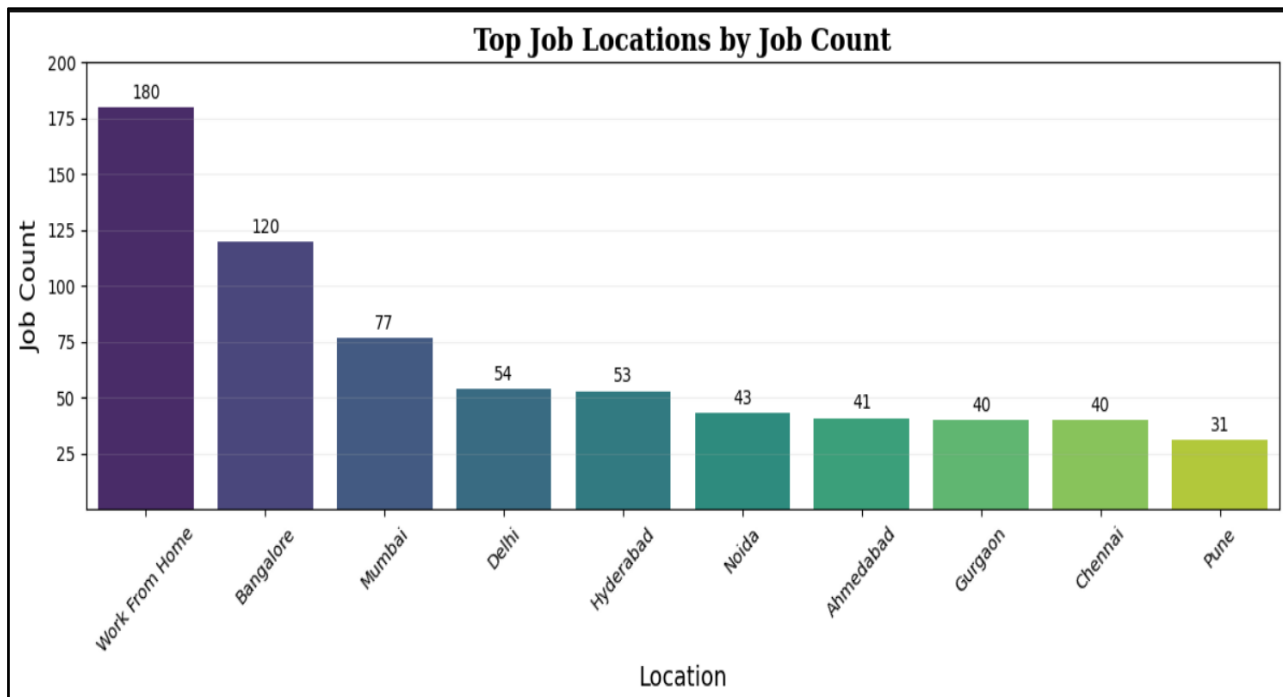
```
df.to_csv("cleaned_intershala_data.csv", index=False,  
encoding='utf-8')
```

## **VI. Data Visualizations & Insights:**

- After cleaning and preprocessing the scraped job data, visualizations were used to analyze the current trends in computer science jobs.
- These charts help uncover trends, patterns, and distributions in key features such as salary, job, title, location, experience etc.
- Visualizing the data allows me to better understand the structure and key patterns present in computer science job market on Internshala.
- Visual representations makes it easier to interpret the data and draw conclusions for decision making.



## 1. Location wise Job Analysis:

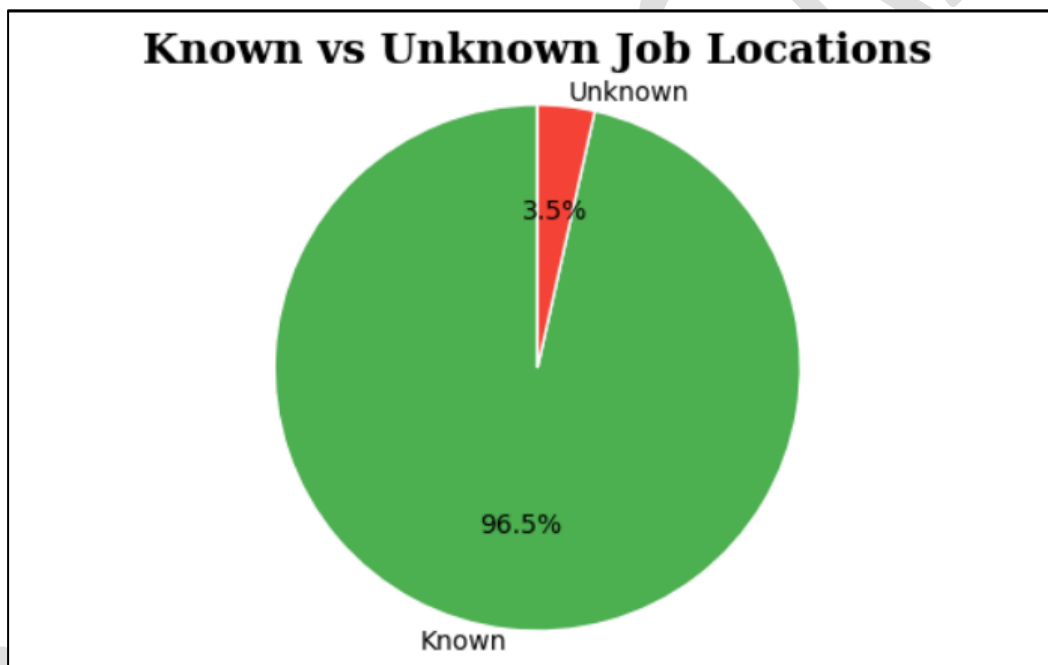


The bar chart represents the top 10 job locations based on job availability.

- Work From Home tops the chart with the highest number of job openings, indicating the rise in remote opportunities.
- Bangalore holds the second position, maintaining its status as India's top tech and startup hub.
- Mumbai ranks third, showing strong job demand in finance and corporate sectors.
- Delhi and Hyderabad follow closely, both being major employment centers in North and South India respectively.

- Noida, Ahmedabad, Gurgaon, and Chennai have moderate job availability.
- Pune concludes the top 10, contributing significantly to the tech and service sectors.
- The distribution reflects both traditional job cities and growing acceptance of remote work options.

## 2. Known vs Unknown Job Locations:



This pie chart illustrates the proportion of job listings with known vs unknown locations.

- A vast majority (96.5%) of job postings mention a specific location, helping job seekers target opportunities better.

- Only 3.5% of listings lack location details and are marked as Unknown, which may indicate missing or incomplete data.
- This highlights the high data completeness of the dataset in terms of geographical job distribution, making the analysis more reliable and location insights more actionable.

### 3. No. of Job Listings by Required Experience:

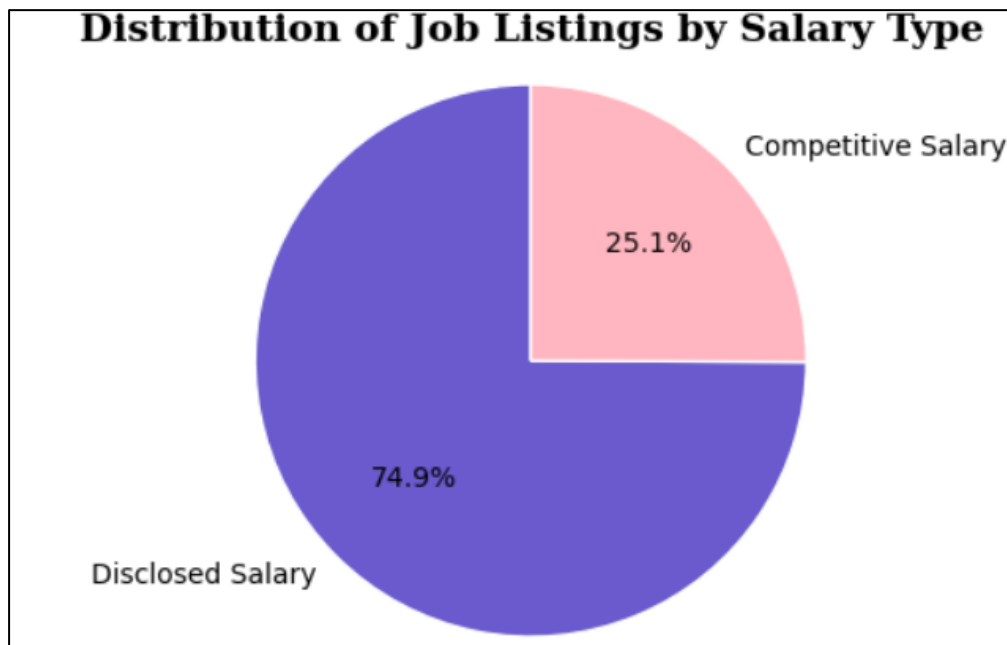


This bar chart shows the distribution of job listings based on the years of experience required:

- The highest demand is for 1-year experience roles (393 listings), followed by fresher roles (0 years) with 219 listings.
- This clearly indicates that companies are actively hiring entry-level candidates, making the job market favorable for fresh graduates and early professionals.
- There is a noticeable drop in job listings as experience requirements increase — only a few jobs demand over 10 years of experience.
- Roles needing 4 to 7 years of experience still maintain moderate demand, indicating steady opportunities for mid-level professionals.
- The data highlights a strong preference for early-career talent in the market, particularly those with 0 to 1 years of experience, which is encouraging for freshers and junior professionals in the field.

#### 4. Distribution of Job Listings by Salary Type:

This pie chart represents how salary information is shared in job listings:



- 74.9% of job listings have a disclosed salary, making it easier for candidates to assess compensation expectations.
- 25.1% of listings mention only "Competitive Salary" without providing exact figures.
- The majority of companies are transparent with salary details, which reflects positively on job market clarity. However, the quarter of listings marked as "Competitive Salary" may indicate roles where salary is negotiable or varies based on experience.

- This visualization helps job seekers understand what to expect in terms of salary transparency during their job search.

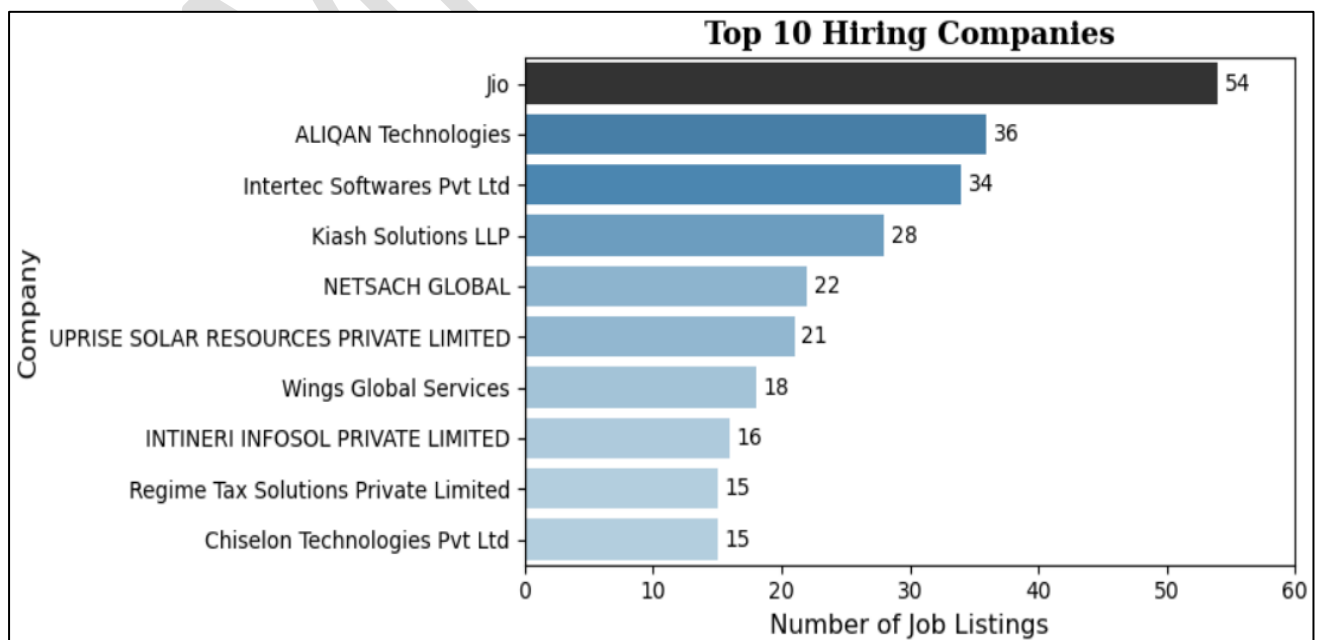
## 5. Salary vs Experience (Individual + Average):



- Blue dots represent individual job listings, showing the actual salary data posted for different experience levels.
- The red line indicates the average salary for each level of experience.
- There is a general upward trend — as experience increases, so does the average salary.

- From 0 to 3 years, salaries rise gradually.
- A notable jump occurs around 6 to 10 years, with some roles offering significantly higher packages (up to ₹80L).
- The variation suggests that salary depends on other factors too — such as job role, company, or location — not just experience.
- This plot highlights how experience positively influences salary, but also that individual cases can differ widely.

## 6. Top 10 Hiring Companies:

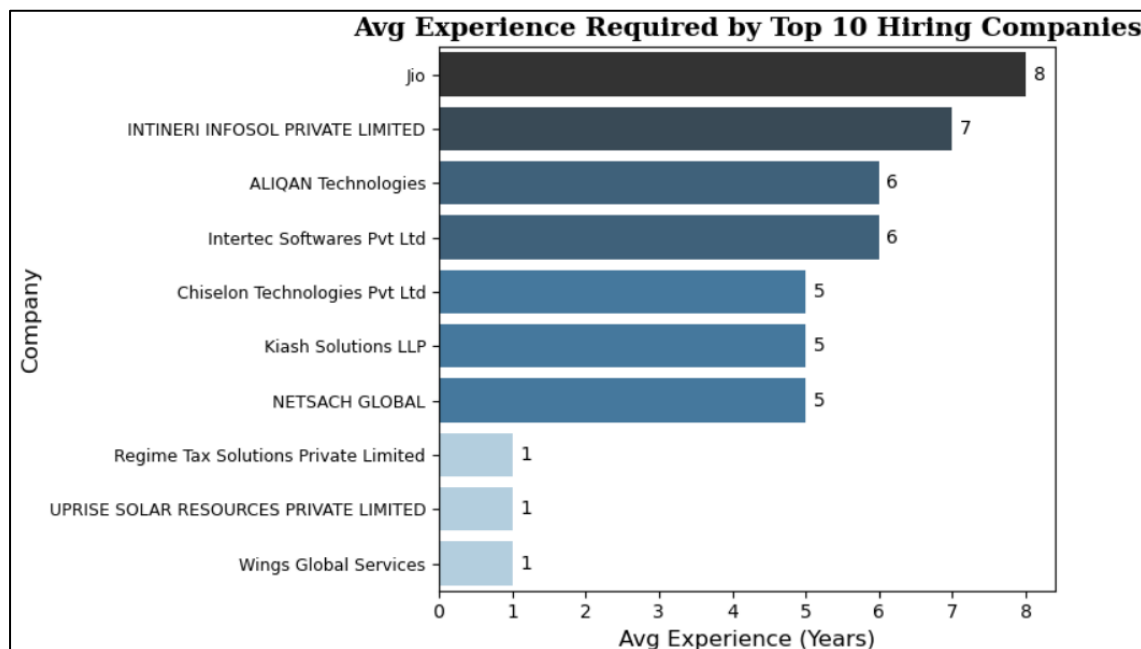


This horizontal bar chart displays the top 10 companies with the most job listings:

- Jio leads by a wide margin with 54 listings, showing a strong recruitment drive.
- Followed by: ALIQAN Technologies, Intertec Softwares Pvt Ltd, Kiash Solutions LLP etc.
- These companies are actively hiring, indicating strong growth.
- Job seekers can prioritize these firms when looking for opportunities, especially in tech and consulting roles.
- This chart helps identify the key players in the job market based on job posting frequency.



## 7. Avg. Experience Required by Top 10 Hiring Companies:



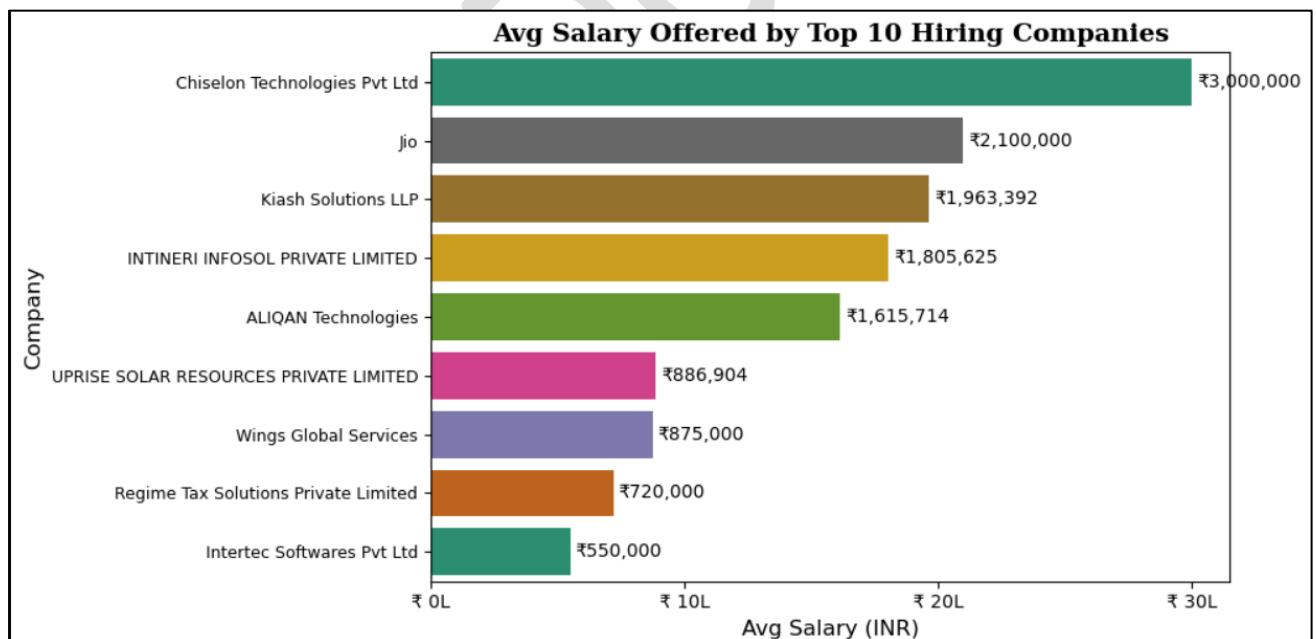
This chart highlights the average years of experience companies expect from candidates:

- Jio demands the highest, with an average of 8 years.
- INTINERI INFOSOL PRIVATE LIMITED follows with 7 years.
- Companies like ALIQAN Technologies, Intertec Softwares, and Chiselon Technologies require around 5–6 years.
- On the other hand, Regime Tax Solutions, UPRISE SOLAR, and Wings Global Services show an average

requirement of just 1 year, indicating openings for freshers or junior roles.

- The variation in experience requirements suggests that while some companies are hiring for senior or mid-level roles, others are open to fresh talent.
- This is useful for targeting companies based on your experience level.

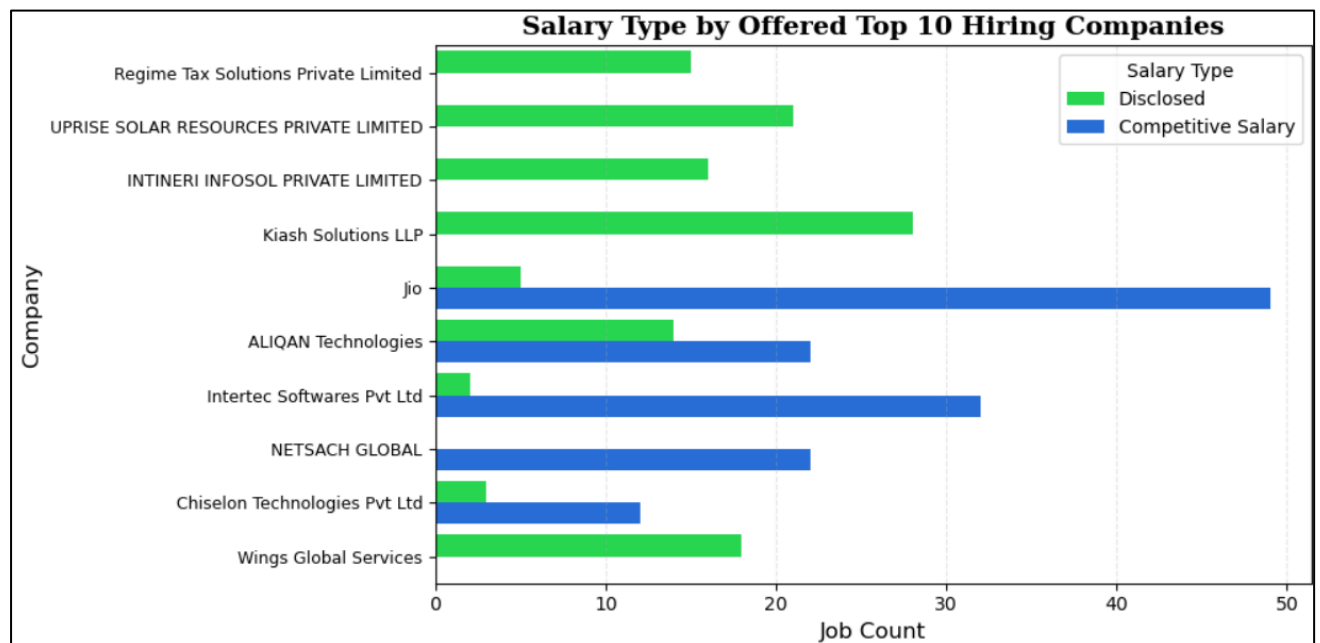
#### 8. Avg. Salary Offered by Top 10 Hiring Companies:



This chart displays the average salaries offered by the top hiring companies in INR:

- Chiselon Technologies Pvt Ltd leads with an impressive average salary of ₹30 LPA.
- Jio and Kiash Solutions LLP follow with salaries of ₹21 LPA and ₹19.6 LPA respectively.
- INTINERI INFOSOL and ALIQAN Technologies also offer attractive packages above ₹16 LPA.
- Companies like UPRISE SOLAR, Wings Global Services, and Regime Tax Solutions offer salaries in the range of ₹7–9 LPA.
- Intertec Softwares Pvt Ltd offers the lowest among the group at ₹5.5 LPA.
- There's a significant salary gap among companies, highlighting differences in role types, company scale, and experience level required.
- It's crucial to align salary expectations with both company standards and your skillset.

## 9. Salary Type Offered by Top 10 Hiring Companies:

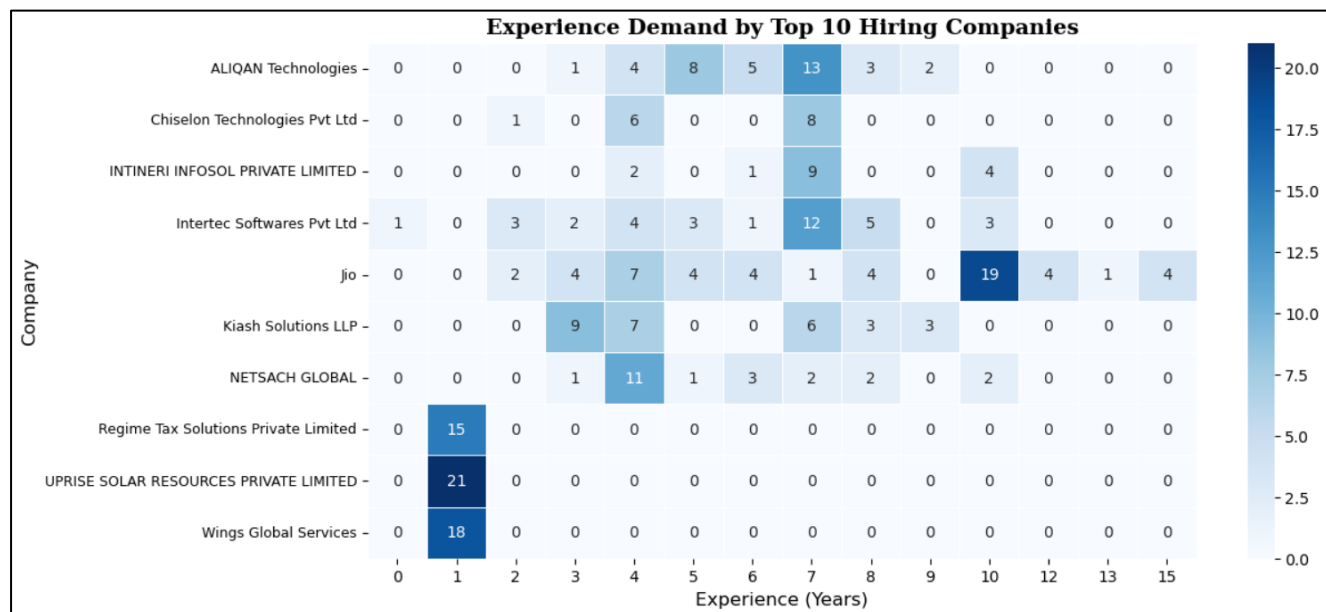


This horizontal bar chart compares the number of job listings by salary type — whether the salary is Disclosed (green) or marked as Competitive Salary (blue) — for the top 10 hiring companies.

- Jio has the highest count of jobs with competitive salary, showing they often avoid disclosing exact figures.
- Intertec Softwares Pvt Ltd and NETSACH GLOBAL also mostly post competitive salary roles, possibly indicating negotiable or variable pay structures.

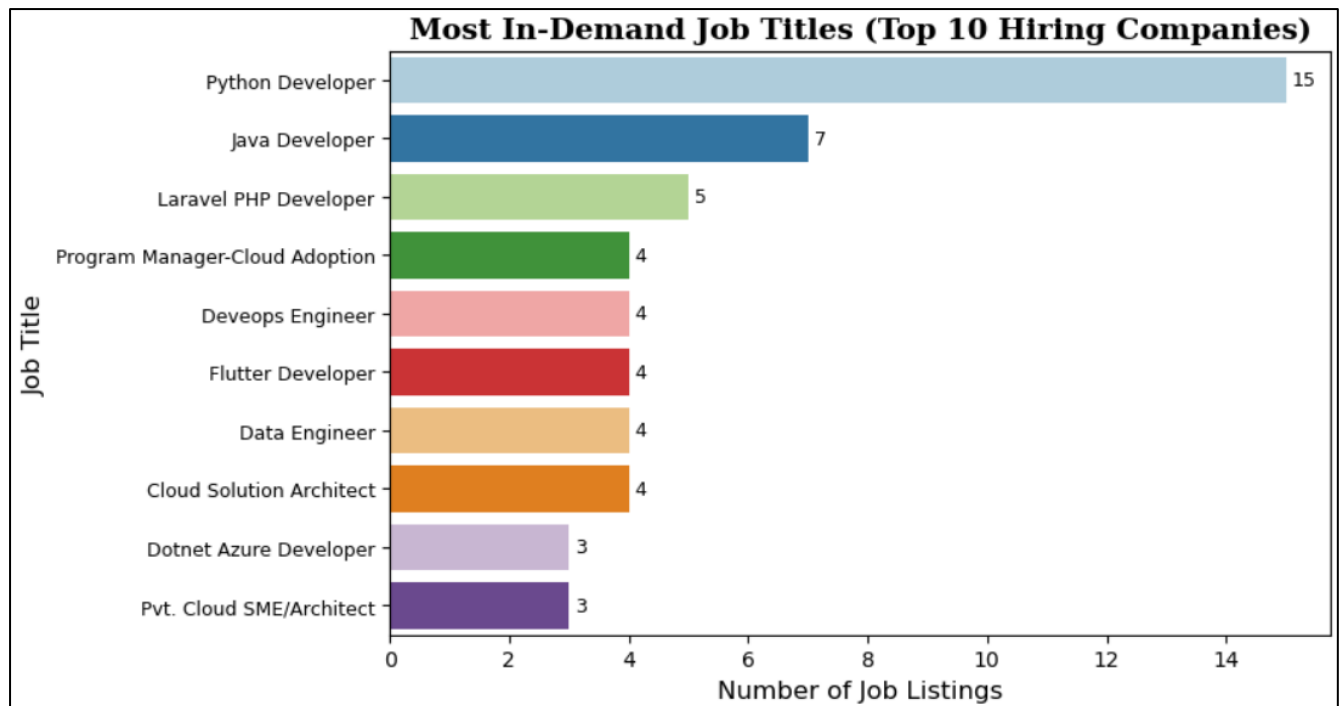
- On the other hand, companies like Kiash Solutions LLP, UPRISE SOLAR, and INTINERI INFOSOL tend to disclose salaries more transparently.
- Chiselon Technologies Pvt Ltd and Wings Global Services show a more balanced mix, offering both disclosed and competitive packages.
- Companies leaning toward competitive salaries may offer flexible or negotiable packages, often influenced by the candidate's profile.
- However, salary transparency can help candidates better evaluate and prepare, making companies with disclosed pay more attractive to job seekers.

## 10. Experience Demanded by Top 10 Hiring Companies:



- High demand for 1-year experience: UPRISE SOLAR, Wings Global Services, Regime Tax Solutions
- Mid-level hiring (3–7 years): ALIQAN Technologies, Intertec Softwares, Kiash Solutions LLP, NETSACH GLOBAL
- Senior roles (10+ years): Jio: 19 jobs for 10 years, 4 jobs for 12–15 years
- Most companies are hiring for junior to mid-level roles, while Jio stands out with a strong demand for highly experienced professionals

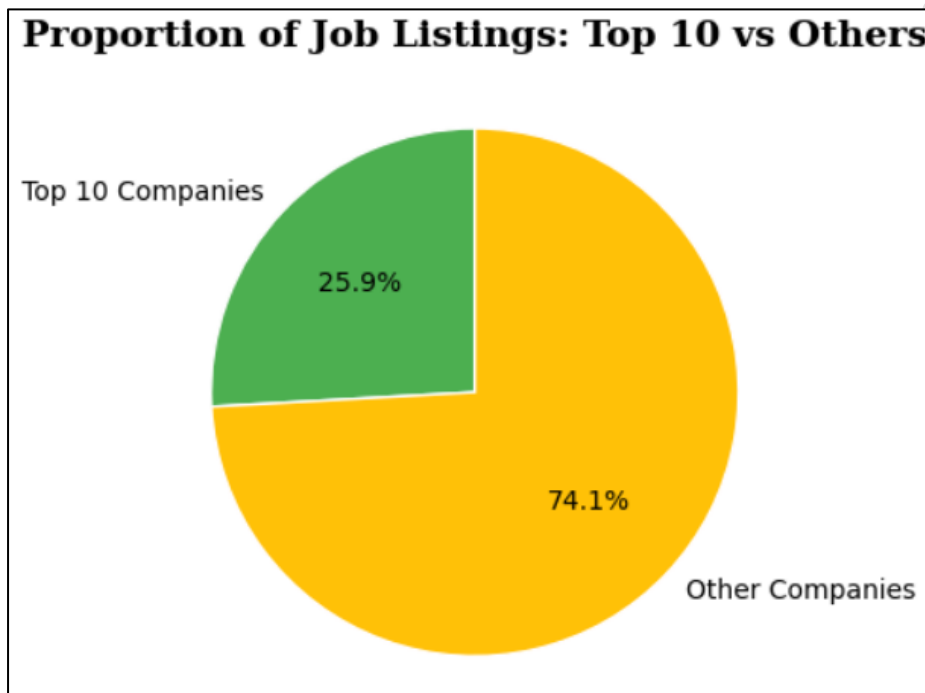
## 11. Most in Demand Job Titles in Top 10 Hiring Companies:



- Python Developer is the most in-demand.
- Java Developer and PHP Developer follows.
- Strong demand for programming (Python, Java) and cloud technologies.
- Mobile development roles like Flutter are also gaining traction.

- Python remains the top choice among employers, while cloud expertise and software development roles continue to drive hiring demand in the current tech landscape.

## 12. Proportion of Jobs – Top 10 vs Others:

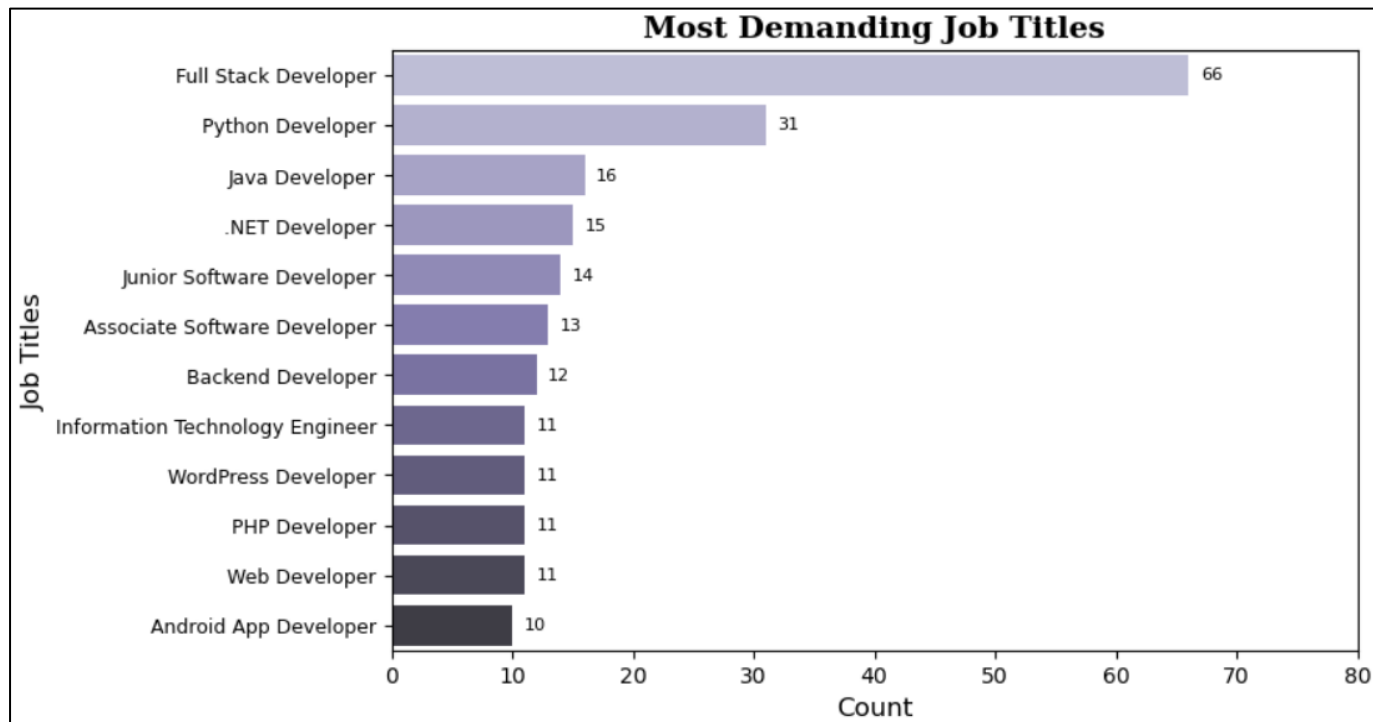


- The top 10 hiring companies account for 25.9% of all job listings. While this indicates their strong presence in the job market, they do not represent the majority.
- A significant 74.1% of job listings are contributed by other companies, reflecting a wide distribution of opportunities beyond the top recruiters.



- This trend suggests that the job market is not monopolized by a few large organizations; instead, it is shaped by numerous smaller or mid-sized companies actively hiring talent.
- The diversity in hiring organizations indicates that job seekers have access to a broader range of opportunities across different company sizes, industries, and roles.
- Job opportunities are not limited to top companies; smaller organizations play a bigger role in hiring.
- This trend gives job seekers more options across various companies.

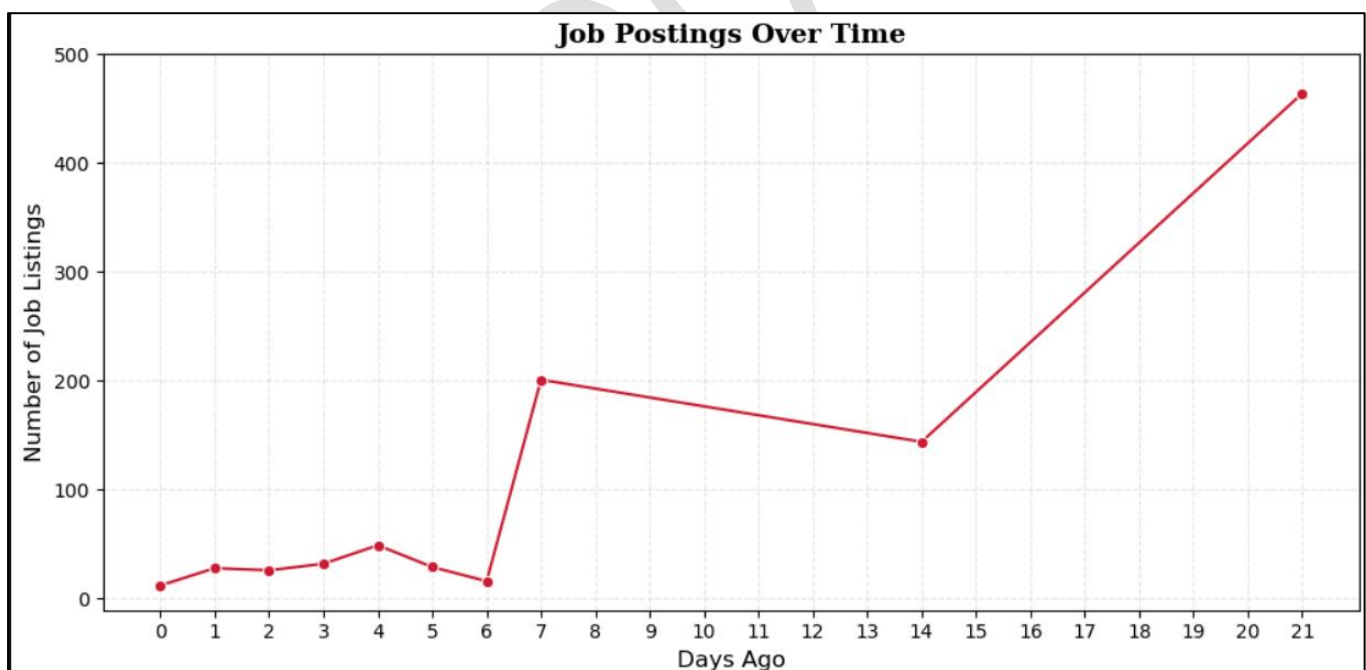
### 13. Most Demanding Job Titles:



- Full Stack Developer is the most demanded job with 66 listings, far ahead of other roles.
- Python Developer is second with 31 listings, showing high demand for programming skills.
- Java Developer (16), .NET Developer (15), and Junior Software Developer (14) are also in demand.
- Backend Developer (12) and Associate Software Developer (13) follow closely.

- Full stack and backend development roles dominate, followed by programming languages like Python and Java.
- There is balanced demand across both entry-level and specialized software development jobs.

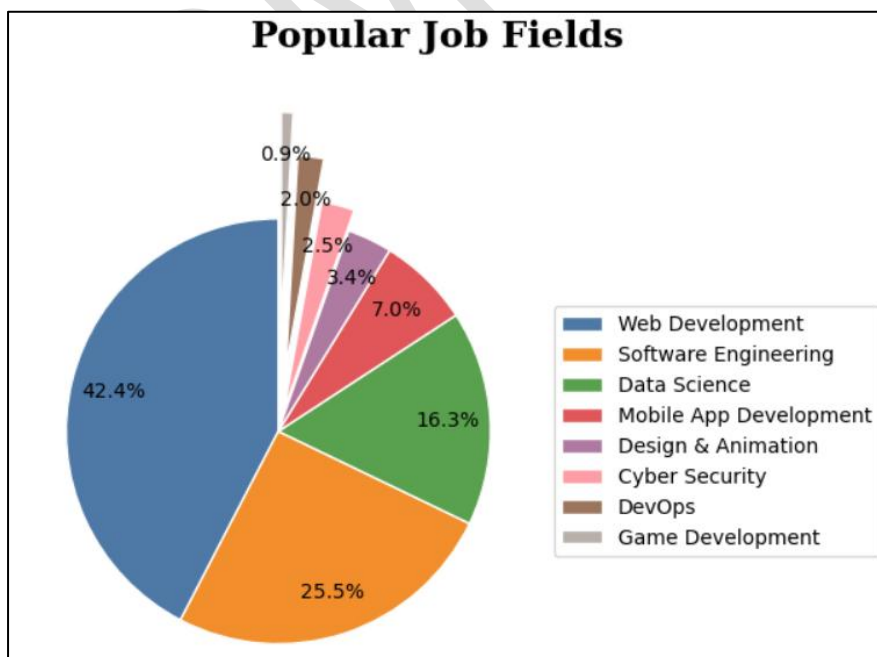
#### 14. Job Postings over Time:



- In the first 6 days, job postings were quite low, mostly below 50 listings per day.

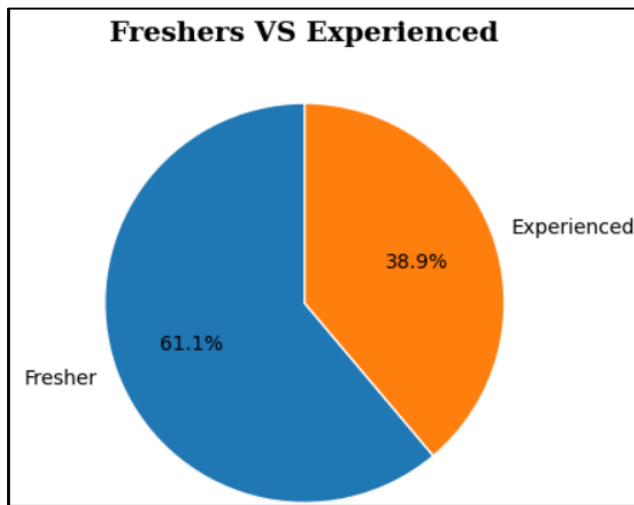
- On day 7, there was a big jump to 200 job listings, showing a sudden hiring push.
- After day 7, postings decreased slowly until day 14, when they were around 140 listings.
- From day 15 onwards, job postings increased sharply, reaching over 450 listings by day 21.
- Based on the line graph, it is observed that companies post significantly more job listings on weekends than on weekdays. This suggests that weekends may offer better opportunities for job seekers to find new openings.

## 15. Popular Job Fields:



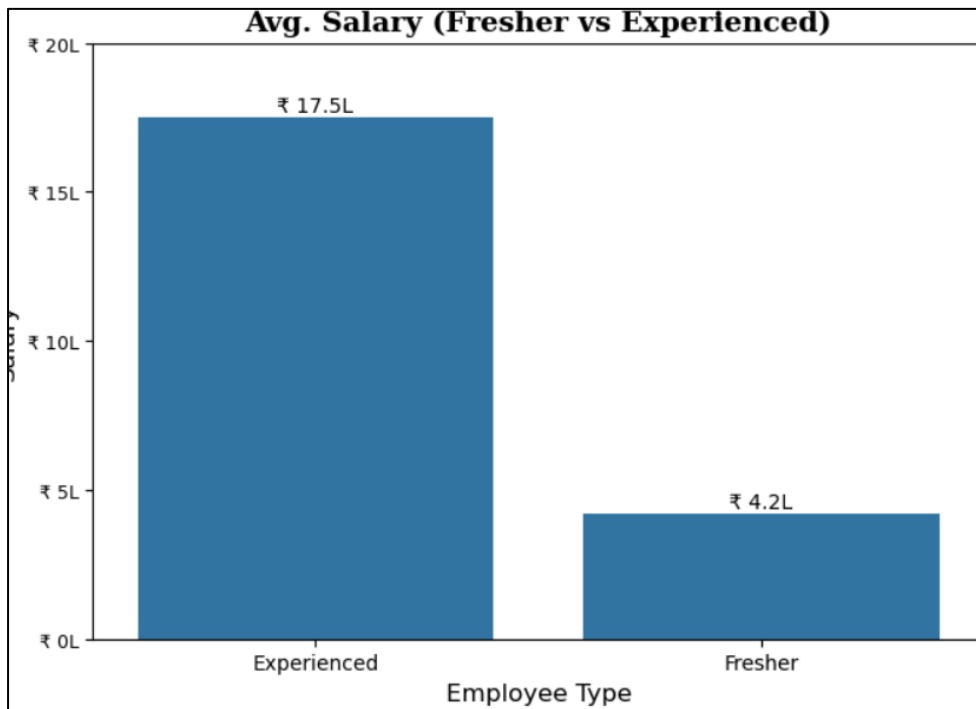
- Web Development accounts for 42.4% of all job listings, making it the most in-demand field in the tech industry.
- Software Engineering ranks second with 25.5% of job listings.
- Data Science comes third with 16.3%, and is showing a noticeable upward trend in demand, indicating its growing importance in modern tech roles.
- Fields like Mobile App Development (7.0%), Design & Animation (3.4%), and Cyber Security (2.5%) show moderate but steady hiring activity.
- DevOps (2.0%) and Game Development (0.9%) remain niche areas with limited but specialized job openings.
- The tech job market is largely driven by web and software development, while Data Science is emerging as a fast-growing field. Creative and security-related roles maintain a smaller but consistent presence, reflecting the industry's need for diverse skill sets.

## 16. Freshers vs Experienced:



- A majority of job listings (61.1%) are targeted at freshers, indicating strong demand for entry-level talent in the tech industry.
- While 38.9% of listings are for experienced professionals, this remains a significant share, highlighting ongoing opportunities for skilled and seasoned candidates.
- The data suggests a healthy job market for both freshers and experienced individuals, with a slight edge for freshers in terms of available opportunities.

## 17. Avg. Salary Freshers vs Experienced:



- Experienced professionals earn an average salary of ₹17.5L, which is significantly higher than the ₹4.2L average salary offered to freshers.
- The data highlights that experience plays a crucial role in boosting earning potential. While freshers begin their careers with lower salaries, there is a clear upward trajectory in compensation as they gain experience and develop their skills.

## **VII. Conclusion:**

### **1. Work from Home is Popular**

- Most job listings are for work from home roles. After that, **Bangalore** has the highest number of job postings, which makes sense as it's known as the tech hub of India.

### **2. Missing Location Reduces Attractiveness**

- Around **3.5% of job listings didn't mention the job location**, which can make those listings less attractive. Candidates usually prefer to know where they will work before applying.

### **3. Internshala Focuses on Freshers**

- The data was collected from the **Internshala website**.
- I have found that **around 600 jobs require only 0–1 years of experience**, showing that Internshala is a great platform for freshers looking for entry-level jobs.

### **4. Salary vs Experience**

- Generally, salary increases with more experience — this is clear in the graphs. However, it's not always the case. **Other factors like the job role, company, and location also affect the salary.**



## **5. Top Hiring Companies Prefer Junior to Mid-Level**

- Most of the **Top 10 hiring companies** are looking for candidates with **junior to mid-level experience**, except for **Jio**.

## **6. Jio Focuses on Experienced Roles**

- Jio mainly offers jobs for **senior-level professionals**.
- Also, many Jio listings don't show the salary upfront — it's discussed later.
- This can be good or bad depending on the candidate's expectations.

## **7. High Demand for Developers**

- Among the Top 10 companies, **Python Developers** are most in demand.
- But if we look at all listings, **Full Stack Developers** have the highest demand overall, followed by **Python** and **Java Developers**.

## **8. Best Time to Apply for Jobs**

- Most jobs are posted during **weekends**, not weekdays. So, it's a good idea to **check and apply during weekends** when there are more new job listings.

## 9. Top 3 Most In-Demand Job Fields

- The most popular job fields are:
  - **Web Development**
  - **Software Engineering**
  - **Data Science**
- Full Stack Development is preferred over just frontend or backend.
- **Data Science is growing fast** and is in demand for both freshers and experienced candidates.
- Fields like **Cyber Security, Game Development, and DevOps** have lower demand on Internshala, likely because they usually require more experience.

## 10. Market is Diverse, Not Monopolized

- The **Top 10 hiring companies account for 25.9%** of job listings.
- This means that **small and medium-sized companies also play a big role**, and the job market is not controlled by only big companies.

## 11. Salary Comparison

- On average:
  - Freshers earn about **₹4 LPA** (Lakhs Per Annum).
  - Experienced professionals earn up to **₹17 LPA**.

## **VIII. Limitations:**

- Data was collected only from Internshala's Computer Science jobs section, which mainly focuses on internships and entry-level roles.
- This may not fully represent the entire tech job market or experienced-level positions.

## **IX. Reference:**

- Job listings were sourced from:  
Internshala – Computer Science Jobs Section  
<https://internshala.com/jobs/computer-science-jobs>