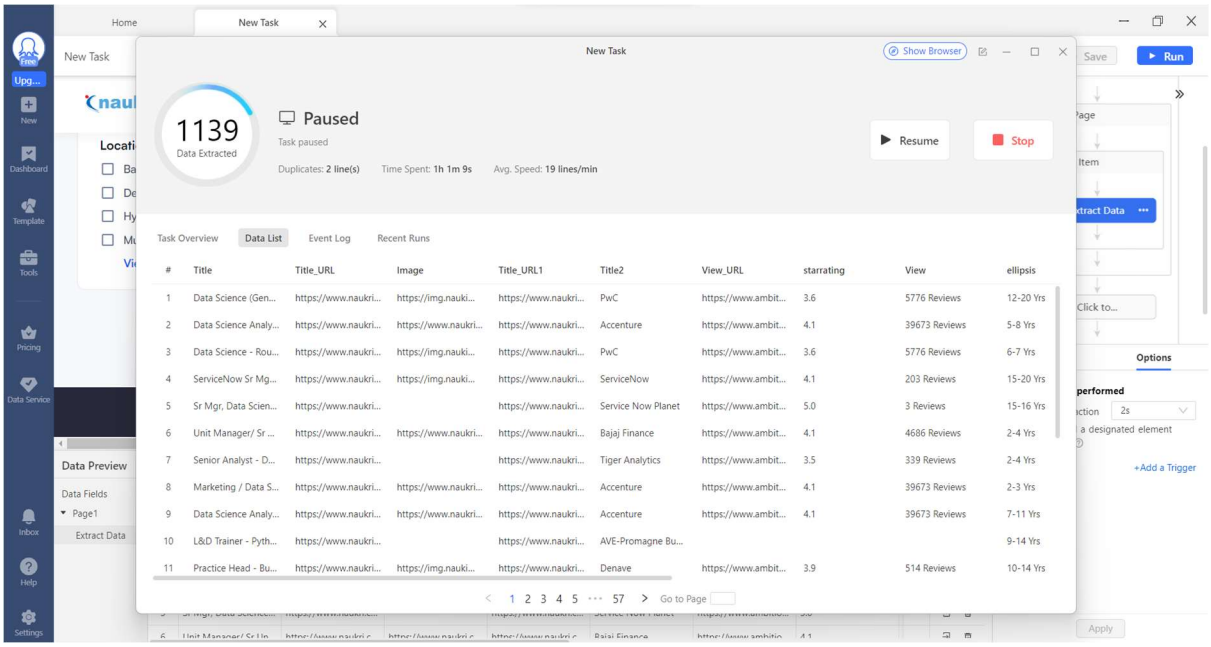
Analysis of job market data to explore the open positions related to the jobs in the field of data scientist or machine learning in a particular region.

job market data was extracted from the website https://www.naukri.com for the position of **"Data Scientist"** in India.

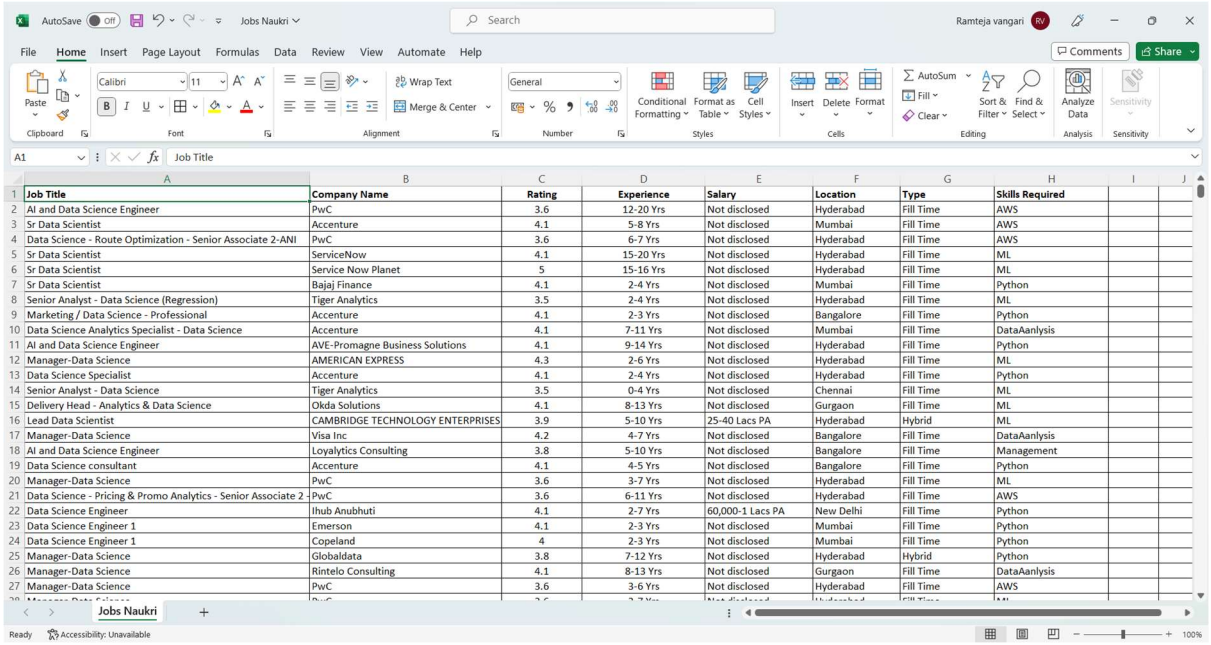
a. Data Collection Method:

I used a tool called OctoParse Web Scraper to get job details from Naukri.com. This tool automatically collected all the needed information.

I gathered details like Job Title, Company Name, Rating, Experience, Salary, Location, Type, and Skills Required.



The dataset included 1,139 job openings. I saved this information in a CSV file and then cleaned it up to make sure job skill names, title names, and location names were all consistent.



Job Title	Company Name	Rating	Experience	Salary	Location	Type	Skills Required
AI and Data Science Engineer	PwC	3.6	12-20 Yrs	Not disclosed	Hyderabad	Fill Time	AWS
Sr Data Scientist	Accenture	4.1	5-8 Yrs	Not disclosed	Mumbai	Fill Time	AWS
Data Science - Route Optimization - Senior Associate 2-ANI	PwC	3.6	6-7 Yrs	Not disclosed	Hyderabad	Fill Time	AWS
Sr Data Scientist	ServiceNow	4.1	15-20 Yrs	Not disclosed	Hyderabad	Fill Time	ML
Sr Data Scientist	Service Now Planet	5	15-16 Yrs	Not disclosed	Hyderabad	Fill Time	ML
Sr Data Scientist	Bajaj Finance	4.1	2-4 Yrs	Not disclosed	Mumbai	Fill Time	Python
Senior Analyst - Data Science (Regression)	Tiger Analytics	3.5	2-4 Yrs	Not disclosed	Hyderabad	Fill Time	ML
Marketing / Data Science - Professional	Accenture	4.1	2-3 Yrs	Not disclosed	Bangalore	Fill Time	Python
Data Science Analytics Specialist - Data Science	Accenture	4.1	7-11 Yrs	Not disclosed	Mumbai	Fill Time	DataAanlysis
AI and Data Science Engineer	AVE-Promagne Business Solutions	4.1	9-14 Yrs	Not disclosed	Hyderabad	Fill Time	Python
Manager-Data Science	AMERICAN EXPRESS	4.3	2-6 Yrs	Not disclosed	Hyderabad	Fill Time	ML
Data Science Specialist	Accenture	4.1	2-4 Yrs	Not disclosed	Hyderabad	Fill Time	Python
Senior Analyst - Data Science	Tiger Analytics	3.5	0-4 Yrs	Not disclosed	Chennai	Fill Time	ML
Delivery Head - Analytics & Data Science	Okda Solutions	4.1	8-13 Yrs	Not disclosed	Gurgaon	Fill Time	ML
Lead Data Scientist	CAMBRIDGE TECHNOLOGY ENTERPRISES	3.9	5-10 Yrs	25-40 Lacs PA	Hyderabad	Hybrid	ML
Manager-Data Science	Visa Inc	4.2	4-7 Yrs	Not disclosed	Bangalore	Fill Time	DataAanlysis
AI and Data Science Engineer	Loyalytics Consulting	3.8	5-10 Yrs	Not disclosed	Bangalore	Fill Time	Management
Data Science consultant	Accenture	4.1	4-5 Yrs	Not disclosed	Bangalore	Fill Time	Python
Manager-Data Science	PwC	3.6	3-7 Yrs	Not disclosed	Hyderabad	Fill Time	ML
Data Science - Pricing & Promo Analytics - Senior Associate 2 -	PwC	3.6	6-11 Yrs	Not disclosed	Hyderabad	Fill Time	AWS
Data Science Engineer	Ihub Anubhuti	4.1	2-7 Yrs	60,000-1 Lacs PA	New Delhi	Fill Time	Python
Data Science Engineer 1	Emerson	4.1	2-3 Yrs	Not disclosed	Mumbai	Fill Time	Python
Data Science Engineer 1	Copeland	4	2-3 Yrs	Not disclosed	Mumbai	Fill Time	Python
Manager-Data Science	Globaldata	3.8	7-12 Yrs	Not disclosed	Hyderabad	Hybrid	Python
Manager-Data Science	Rintelo Consulting	4.1	8-13 Yrs	Not disclosed	Gurgaon	Fill Time	DataAanlysis
Manager-Data Science	PwC	3.6	3-6 Yrs	Not disclosed	Hyderabad	Fill Time	AWS

b. Market Data Visualization:

I gathered the following information from the **Naukri.com** website:

Information Gathered	Details
Job Title	Data Scientist in India
Opening Positions	1139
Salary	Mostly 'Not Disclosed'
Location	Bangalore, Hyderabad, Chennai, New Delhi, Mumbai
Type	Mostly 'Full-Time' for Data Scientist positions
Skill Required	Python, ML, SQL, AWS
Rating	The average Rating is 4.1

I cleaned the data to remove unnecessary and disorganized information that wouldn't be useful for further analysis. I got rid of empty data points that didn't contribute any meaningful insights. Initially, there were 1,139 data points, but after cleaning, that number was reduced to around 985.

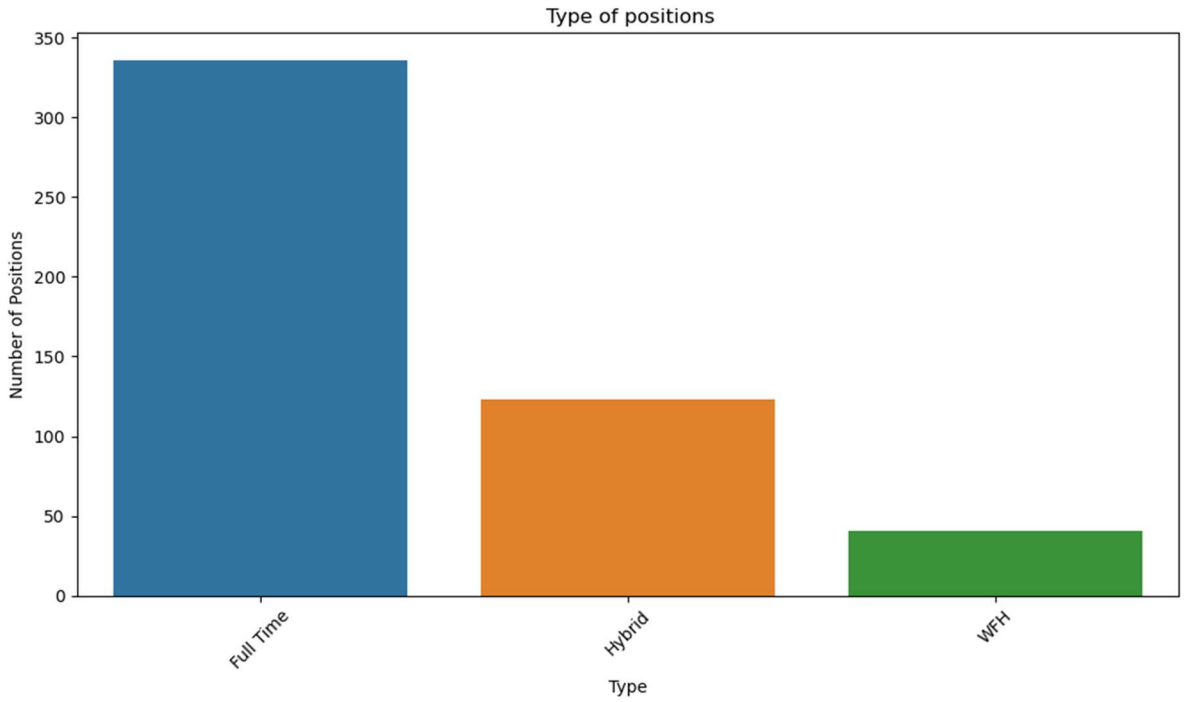
From the cleaned data, I randomly selected a sample of about 500 data points for creating graphs and charts to visualize the information.

1. Analysis of Employment Types:

Looking at the raw data from Naukri website, most of the Data Scientist jobs in India were marked as requiring a Full-Time commitment. It seems that companies mostly need employees who would work full-time. There were only a few jobs that allowed you to Work-From-Home (WFH) or had a mix of in-office and remote work (Hybrid).

While Full-Time jobs were more common, the number of opportunities for WFH and Hybrid positions was quite small. This tells us that jobs where you can work from home or hybrid are not very common in the Data Scientist field.

Companies might be trying to find a balance between having their employees fully involved in their work and giving them some flexibility to work from home. This is why there aren't many jobs that offer WFH or Hybrid arrangements.

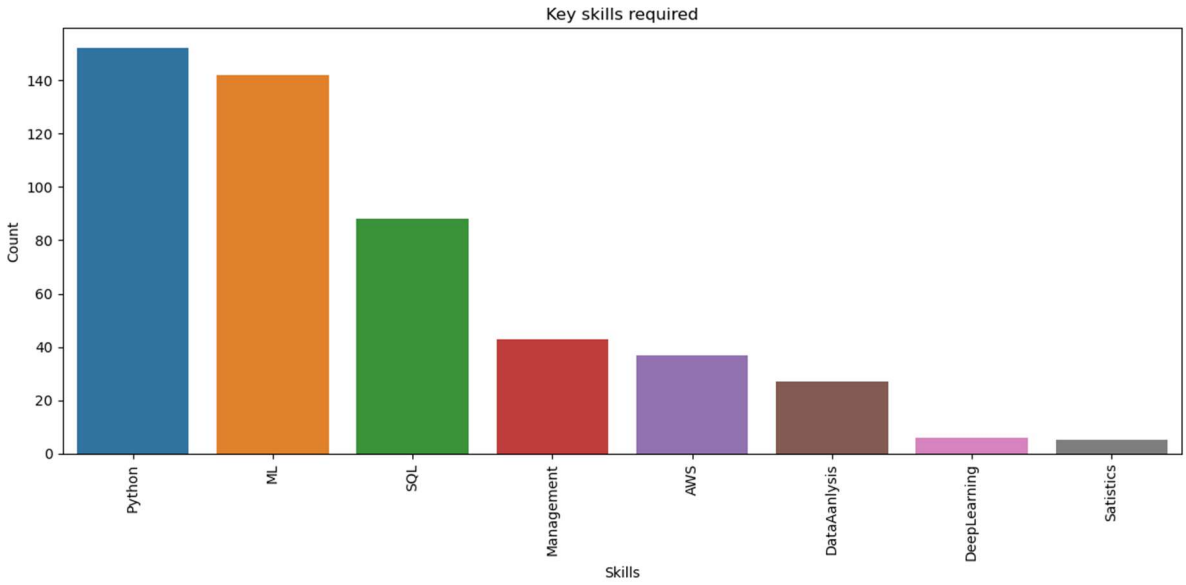


2. Analysis of Skill Sets:

The graphs are clear and easy to understand. Python stands out as a fundamental skill required for Data Scientist roles.

To work effectively with large datasets used in ML algorithms, it's important to have cloud-based computing knowledge like AWS or Azure. This helps handle decentralized and heavy computing tasks.

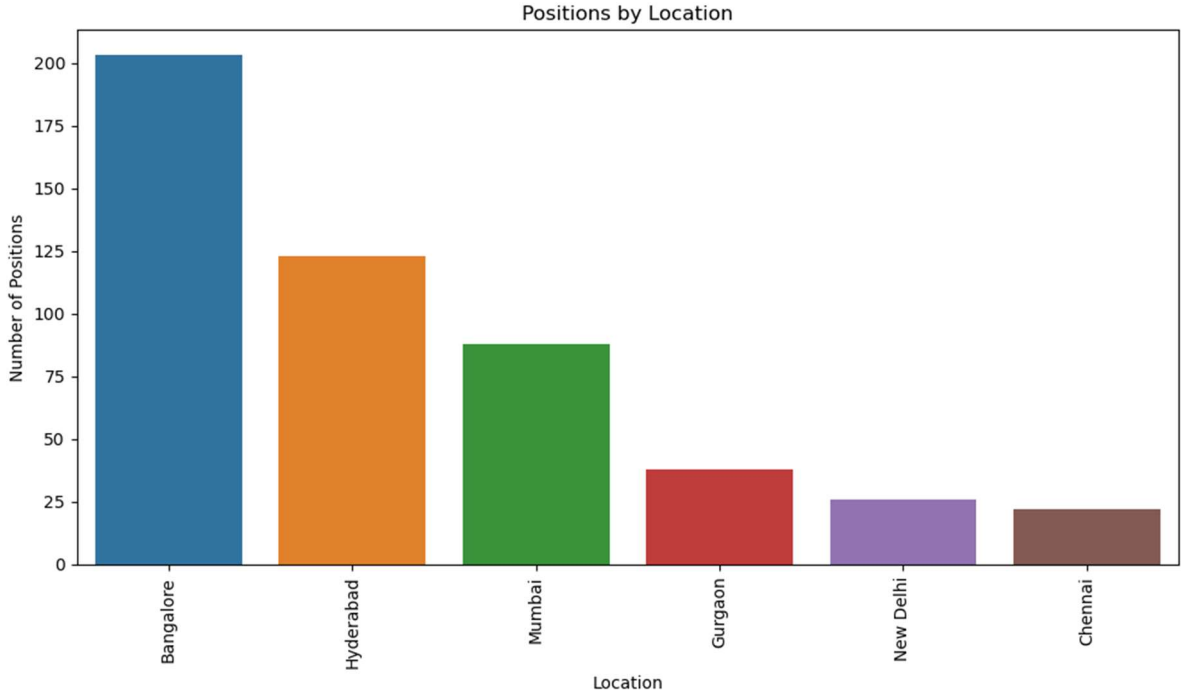
Dealing with numerous datasets involves using Database Management Systems (DBMS). This requires knowing languages like SQL, MySQL, or NoSQL.



3. Analysis of Location:

I used a bar graph to show where Data Scientist job openings are available across different places. It's not surprising that Bengaluru is a major hub for Data Scientist roles in India. Among the jobs in the sample, Bengaluru had the most Data Scientist openings, which is a trend seen throughout the dataset.

Bengaluru's significance matches its reputation as a tech and innovation centre. Many Data Scientist job opportunities are concentrated there.



Hyderabad's position is noteworthy. It shows how the city is growing as a tech and research destination, attracting companies looking for Data Science expertise.

The job openings distribution in these cities reflects how Data Science jobs are spread across India. It's not limited to one place.

These cities are known for their strong tech communities. They create an environment for innovation and attract both companies and skilled professionals. The ranking of these cities shows how the job market for Data Scientists is changing. Some cities are becoming more important, while others continue to contribute consistently.

The presence of tech companies, research institutions, and supportive policies could be why these cities have more Data Scientist jobs. The distribution shows that Data Scientist roles are available in different regions of India, not just in one place.

As the Data Science field grows, other cities might join the list of tech hubs, showing that there's room for growth and innovation.

Job seekers can think about matching their skills with what these cities are looking for in Data Scientists to get the best chances of landing a job.

4. Analysis of positions available:

The word cloud graph in the dataset points out that "Data Scientist," "Data Analyst," "Data Science," and "Senior Data Scientist" are key roles. They show that businesses want to make smart choices based on data. While "Data Scientist" is the big winner, the mix of roles highlights how vital it is to understand data at different levels for decision-making.

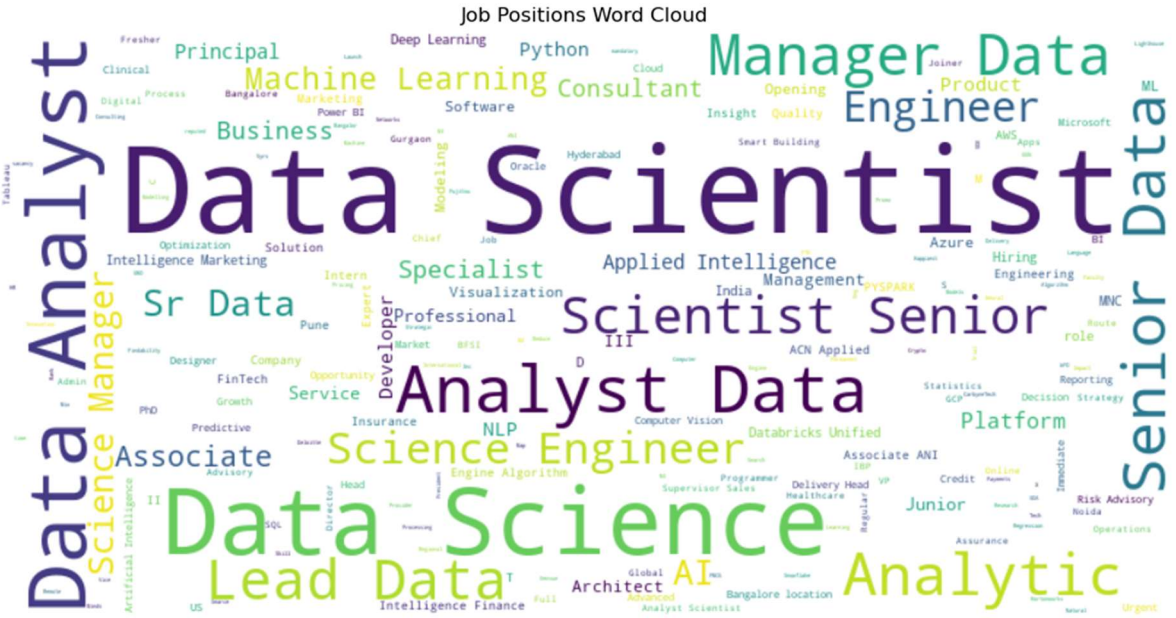
"Data Scientist" is the most common role in the dataset. This shows that there's a big need for experts who can handle complex data analysis.

"Data Analyst" is also quite popular in the dataset. This suggests that there's a demand for people who can find valuable insights into data.

The appearance of "Data Science" shows that the whole field is a big deal in these job openings. It's about using data for smart decisions.

Seeing "Senior Data Scientist" means companies are looking for experienced pros. They want people who know their way around data and can lead.

These roles all revolve around working with data, whether it's analysing, science, or leading as a senior. The order from "Data Scientist" to "Senior Data Scientist" suggests different levels of expertise. Companies might want both beginners and seasoned experts.



The word cloud also indicates that there are several related positions in the dataset. These positions might have slight variations in their titles but share common themes and skill requirements.

c. Your ideal job:

In Hyderabad, there are many opportunities for Data Scientist positions because there are lots of jobs in India that involve working with data. Since I'm from Hyderabad and I know there are plenty of job options, this could influence my decision about which job to choose.

Because I live in Hyderabad, this job opportunity makes sense for me geographically.

Taking all these things into account, the job I think would be best for me is a Data Scientist. This job matches what I'm good at, fits with my career plans, and is in line with what I want since I'm from Hyderabad.

skills I need for this job:

- Data Analysis
- Machine Learning
- Statistical Modelling
- Python Programming

Based on all of this, I believe the Data Scientist role at **"Innominds Software"** in Hyderabad is the right choice for me and my career.