

# CSE 355

## DATA SCIENCE PROFESSIONAL CERTIFICATION

E0422005  
ARCHIT BHATIA

### **Role: Data Scientist**

A **Data Scientist** is a professional responsible for extracting actionable insights from raw data using statistical methods, machine learning, and programming. Their core responsibility is to transform data into knowledge that informs business decisions.

### **Responsibilities:**

- Collect, clean, and preprocess structured and unstructured data.
- Perform exploratory data analysis (EDA) and hypothesis testing.
- Build predictive models using ML and AI.
- Communicate results via data visualizations and reports.
- Work cross-functionally with business and engineering teams.

### **Skills Required:**

- **Programming:** Proficiency in Python, R, and SQL.
- **Mathematics & Statistics:** Strong foundation in probability, linear algebra, and hypothesis testing.
- **Machine Learning:** Knowledge of supervised, unsupervised, and deep learning techniques.
- **Data Visualization:** Tools such as Tableau, PowerBI, or libraries like Matplotlib and ggplot2.
- **Communication:** Ability to translate technical insights into business recommendations.

### **Tools Commonly Used:**

- Jupyter Notebooks, RStudio for analysis.
- TensorFlow, PyTorch, Scikit-learn for ML.
- Big Data platforms (Hadoop, Spark).
- Git/GitHub for version control.

## Example LinkedIn Job Post (Highlights):

- Master's or Bachelor's in Computer Science/Statistics.
- 2–4 years of experience in ML model development.
- Strong Python/R and SQL knowledge.
- Experience with cloud platforms (AWS/Azure).
- Excellent problem-solving and communication skills.

A Data Scientist is a **blend of statistician, programmer, and business strategist**. Just as a scientist tests hypotheses in the lab, a data scientist tests data-driven hypotheses in real-world industries, shaping organizational strategy.

The screenshot shows a LinkedIn job search interface. The search bar at the top contains the text "data science". Below the search bar, there are filters for "Jobs", "Date posted", "Company", "Easy Apply", "Under 10 applicants", and "In my network". The search results show 249 results for "Associate Specialist Data Science" at MSD, located in Pune, Maharashtra, India (Hybrid), posted 2 weeks ago. The job details are expanded, showing the following information:

- Skills:** Machine Learning, Software Development, Generative AI, NLP, Prompt Engineering, Testing, Software Development Lifecycle (SDLC), Design Applications using AWS / GCP / Azure, Data Engineering.
- Specializations:** AI/ML (Machine Learning, Data Engineering, Data Science), Gen AI, NLP.
- Minimum Qualifications:** Bachelor's degree with 0-3 years industry experience, Proficiency in Python & SQL, 1-3 years' experience in healthcare analytics and/or consulting, Strong foundations of statistics and machine learning, Proficiency in Data Science Concepts, Microsoft Excel and PowerPoint.
- Preferred Qualifications:** (Section header visible, details not fully expanded).

At the bottom of the job details, there is a "Get job alerts for this search" toggle switch set to "Off".