📘 Lesson-01: What Is Data Science?

1. Summary

This lesson introduces the concept of data science from multiple expert perspectives. It defines data science as a process of exploring, analyzing, and understanding data to extract meaningful insights. Data science combines statistics, programming, and storytelling to help solve problems, validate models, and support decision-making. With the rapid growth of data and free tools, the field is more important and accessible than ever before.

2. Key Notes

Data science is a process, not a one-time event. It's about continuously learning from data.

- Involves creating hypotheses or models, and using data to test or validate them.
- It includes extracting and cleaning data from structured (e.g., tables) and unstructured (e.g., text, images) sources.
- Storytelling with data is a major part—turning numbers into understandable insights.
- Similar to other sciences (like biology or physics), data science studies data itself.
- The term "data science" emerged in the 1980s–90s from evolving statistical studies.
- It's not just about tools or coding it's about curiosity and the drive to explore and understand using data.
- Modern advantages of data science:
    1. Massive amounts of data now available.
    2. Algorithms are better and freely available.
    3. Open-source software (like Python, R) has replaced expensive tools.
    4. Data storage and processing are now cheap and scalable.

Key idea: if you have **data** and **curiosity**, you're already doing data science.

3. What I Learned

I learned that data science is more than just using software or making predictions. It's about asking smart questions, exploring data with curiosity, and finding patterns that can help people or organizations make better decisions.

4. How I Will Use This

I will approach every project with a question-first mindset, using data science as a tool to explore, validate, and explain. I'll focus on both the technical side (cleaning, modeling) and the communication side (storytelling and insight sharing). This mindset will guide how I build models, analyze data, and present results in real-world projects.

5. Analogy Table: Data Science vs Other Sciences

Science Type	Object of Study	Example Tool
Biology	Living organisms	Microscope
Physics	Physical forces	Equations/Simulations
Data Science	**Data**	Python, R, SQL, Jupyter

6. Vocabulary / Glossary

- Hypothesis – A testable idea or assumption about a problem.

- Unstructured Data – Data without a clear format (e.g., text, images, audio).

- Structured Data – Organized data (e.g., rows and columns in a database).

- Storytelling (in Data) – Turning insights into narratives that others can understand.

- Data Deluge – A massive and constant flow of data being generated daily.

8. Interview Insight
"When asked 'What is data science?', I'll describe it as a process of using data to test ideas, uncover trends, and tell stories that guide smarter decisions. It's a mix of logic, coding, statistics, and creativity."

9. Personal Notes
This lesson helped me see data science as both an art and a science. It starts with curiosity and grows into problem-solving. I feel more confident knowing that asking good questions is just as important as coding.