✅ Module 1 Overview

- Introduces key foundational concepts: **population vs. sample**, **probability**, and **sample space**
- Emphasizes that **probability** is the foundation of **statistics** and **data science**

📊 What is Statistics?

- **Definition**: The science of using data to gain knowledge
- Involves **ethical collection and analysis** (based on test assumptions, not morality)
- **Population**: Full set of individuals/objects of interest
- **Sample**: Subset drawn from the population (often due to size or inaccessibility)

🔁 Statistics vs. Probability

- **Probability**: Assume population characteristics → infer about samples
- **Statistics**: Use sample characteristics → infer about populations

---

🎲 Sample Spaces and Events

- **Probability** provides a framework to handle **randomness and uncertainty**
- Examples:
    - Coin flips
    - Purchase behavior
    - Stock price predictions

Key Terms:

- **Experiment**: Any process that generates observations
- **Sample Space (S)**: Set of all possible outcomes
- **Event**: A single or collection of outcomes
- **Cardinality**: Number of outcomes (denoted |S|)

🔢 Sample Space Examples

1. **Single coin flip**:
    - S = {0, 1} where 0 = head, 1 = tail
    - |S| = 2
2. **Two coin flips**:
    - S = {00, 01, 10, 11}
    - |S| = 4
3. **Flip until first tail**:
    - S = {T, HT, HHT, HHHT, ...}
    - |S| = ∞ (infinite outcomes)
4. **Car inspection for 3 defects** (engine, seatbelt, paint):
    - Each outcome: 3-digit binary string (e.g., 100 = engine only)
    - |S| = 8

🧮 Set Notation and Events

- **Union (A ∪ B)**: Outcomes in A or B (or both)
- **Intersection (A ∩ B)**: Outcomes in both A and B
- **Complement (A^c)**: Outcomes not in A
- **Mutually Exclusive**: A ∩ B = ∅ (no overlap)

🚗 Car Defects Example – Event Definitions

- **Event A**: Engine problem → outcomes with 1 in first digit
  - A = {100, 110, 101, 111}
- **Event B**: Exactly one defect
  - B = {100, 010, 001}
- **Event C**: Exactly two defects
  - C = {110, 101, 011}

Combining Events:

- **A ∩ B** = {100}
- **A^c** = {000, 010, 001}
- **A^c ∪ B** = {000, 010, 001, 100}
- **B ∩ C** = ∅ (no overlap)

🔵 Venn Diagrams

- Used for **visualizing** unions, intersections, and complements

- Each circle represents an event (e.g., A, B, C)
- Every outcome from S appears exactly once in the diagram

Example Placement:

- $A \cap B = \{100\}$
- $A \cap C = \{110, 101\}$
- $B = \{010, 001\}$, $C = \{011\}$, $A = \{111\}$, Outside $A \cup B \cup C = \{000\}$

📘 Key Takeaways

- Understand **population vs. sample**
- Learn the **structure of probability** (experiments, sample spaces, events)
- Use **set operations** and **Venn diagrams** to describe relationships between events
- Foundation for future topics: **prediction, inference, modeling**