# Synthetic Data

# What is Synthetic Data?

Synthetic Data is Al-generated training data that mimics real-world conditions. It allows businesses to train Al models without needing massive real datasets.

💡 Think of it like this: Just as pilots train in a flight simulator, Al trains on Synthetic Data before handling real-world tasks.

# Why Businesses Use Synthetic Data

Without Synthetic Data:

- X Al models require large amounts of real-world data, which may take months to
- X Businesses struggle to train AI for rare events like fraud or supply chain disruptions.

#### With Synthetic Data:

- Al learns faster using Al-generated examples.
- Businesses save time and costs by reducing the need for real-world data collection.
- Al can train for rare or unpredictable situations.

## How Synthetic Data is Created

- 1 Al Simulations Al models real-world behaviors to generate new data.
- 2 Data Augmentation Al modifies existing data to create new variations.
- 3 Generative AI (GANs) AI produces entirely new datasets from scratch.

### 💡 Example: Al in Retail Forecasting

- The Problem: A retail company wants to train AI to predict Black Friday sales, but they only have one year of data.
  - The Solution: Al generates simulated shopping trends based on real data.
- The Outcome: The company gets accurate forecasts without waiting years for data collection.

## Real-World Use Cases

- **Retail:** All predicts **customer demand** using simulated shopping trends.
- \* Fraud Detection: Al creates artificial fraud cases to improve security models.
- Property Supply Chain: Al tests different crisis scenarios to prepare for disruptions.
- Marketing: Al simulates customer behavior to optimize ad targeting.
- Key Takeaway: Synthetic Data allows businesses to train Al faster, improve accuracy, and prepare for rare events.