

## Lesson 01 Demo 01

### Data Augmentation Using GenAI

**Objective:** To augment a dataset through prompt engineering using GenAI

**Tools required:** Gemini or ChatGPT

**Prerequisites:** None

#### Steps to be followed:

1. Select a Use Case
2. Download the Dataset
3. Create a Prompt
4. Use a GenAI tool to Generate Output

#### Step 1: Select a Use Case

- 1.1 For this demo, we have selected a use case of sales analysis and forecasting on the Superstore dataset.

#### Step 2: Download the Dataset

- 2.1 Download the Superstore dataset from LMS under the reference material section
- 2.2 Upload the dataset on ChatGPT or Gemini AI

#### Step 3: Create a Prompt

- 3.1 Create a prompt that supports the use case and is effective in augmenting the given dataset

#### Prompt:

Using generative AI techniques, augment the given sales forecasting dataset with additional plausible data points to simulate potential sales forecasting. The provided dataset includes information on Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country, City, State, Postal Code, Region, Product ID, Category, Sub-Category, Product Name, Sales Quantity, Discount, and Profit. Aim to expand the dataset to 50 rows while maintaining the statistical properties and


patterns observed in the original data. Provide output in the form of a table.

**Note:** Copy and paste the data from the dataset along with the prompt

**Step 4: Use a GenAI tool to Generate the Output**

4.1 Use a GenAI tool such as ChatGPT or Gemini to augment the dataset

**Generated Data:**

**ChatGPT**

To augment the given sales forecasting dataset, I will generate additional data points while maintaining the statistical properties and patterns observed in the original data. Below is the augmented dataset with 50 rows:

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City
1	CA-2016-152156	08-11-2016	11-11-2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henders
2	CA-2016-152156	08-11-2016	11-11-2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henders
3	CA-2016-	12-06-	16-06-	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles