1. Use Case for HDFC SME Loans Portfolio:
   * HDFC aims to double its SME Loans portfolio by the end of 2027.
   * The current IT landscape includes a front-end website with API endpoints for loan applications in an AWS environment.
   * An RDS database stores data related to final loan applications and their processing.
   * A team of 1 manager and 7 employees manually enters paper-based applications, consuming approximately 150 man-hours per month.
2. Dataset Description:
   * The dataset contains gray-scale images of hand-drawn digits (0 to 9).
   * Each image is 28x28 pixels (784 pixels in total).
   * Pixel values range from 0 to 255, representing lightness or darkness.
   * The training dataset (train.csv) consists of 60,000 rows with 785 columns (784 for image information and 1 for digit labels).
   * Irrelevant columns were dropped, and null values were removed.
3. Data Cleaning and EDA:
   * Irrelevant columns (unnamed and extra index) were dropped.
   * Null values were removed.
   * Data was scaled using Min-Max Scaler.
4. Naive Bayes Algorithm:
   * Assumption: Features are conditionally independent.
   * Naive Bayes learners and classifiers are fast compared to more complex methods.
   * Decoupling class conditional feature distributions allows independent estimation.
   * This approach helps mitigate issues related to the curse of dimensionality.
5. Gauss BAyes Algorithm
   * Gaussian distribution is also called normal distribution.
   * Normal distribution is a statistical model that describes the distributions of continuous random variables in nature and is defined by its bell-shaped curve.
   * The two most important features of the normal distribution are the mean (μ) and standard deviation (σ).
   * The mean is the average value of a distribution, and the standard deviation is the “width” of the distribution around the mean
6. KNN Algorithm:
   * Step #1 - Assign a value to K.
   * Step #2 - Calculate the distance between the new data entry and all other existing data entries (you'll learn how to do this shortly). Arrange them in ascending order.
   * Step #3 - Find the K nearest neighbors to the new entry based on the calculated distances.
   * Step #4 - Assign the new data