# Dataset Overview

**1. Dataset Used:**
- **CIFAR-10:** A dataset consisting of 60,000 color images (32x32 pixels) across 10 classes.
- **Classes:** Airplane, Automobile, Bird, Cat, Deer, Dog, Frog, Horse, Ship, Truck.
- The dataset is split into:
  - **Training set:** 50,000 images.
  - **Testing set:** 10,000 images.

**2. Shape of the Data:**
- **Training Data:**
  - Features (X_train): (50,000, 32, 32, 3).
  - Labels (y_train): (50,000, 1).
- **Testing Data:**
  - Features (X_test): (10,000, 32, 32, 3).
  - Labels (y_test): (10,000, 1).

## C. Key Steps in the Analysis

**1. Data Visualization**
**Random Image Grid:**
- A 10x10 grid of randomly selected images from the training set was displayed.
- Each image was labeled with its corresponding class name.

**Class Distribution:**
- Training and testing sets are perfectly balanced, with equal samples across all classes.

**2. Data Preprocessing**
**Scaling Features:**
- Pixel values of the images were scaled to the range [0, 1].
**One-Hot Encoding:**
- Labels were transformed into one-hot encoded vectors.

**3. Data Augmentation**
**Applied random transformations:**
- Horizontal flipping.
- Random width and height shifts (10% of image size).

# Model Training
**1. Transfer Learning:**

- Used DenseNet121 pre-trained on ImageNet.
- Modified the top layers for CIFAR-10 classification.

2. **Training Metrics:**
   - Loss: Categorical cross-entropy.
   - Optimizer: Adam.
   - Metrics: Accuracy, Precision, and Recall.

3. **Early Stopping:**
   - Early stopping was optionally used to prevent overfitting based on validation loss.

## Insights and Observations
1. **Balanced Dataset:**
   - Equal number of samples for each class ensures no bias in the dataset.
2. **High Variability:**
   - CIFAR-10 images exhibit significant variability, addressed with data augmentation.
3. **Transfer Learning:**
   - DenseNet121 provides a strong baseline for classification, leveraging ImageNet features.

## Conclusion
The CIFAR-10 dataset is well-structured and balanced. Scaling, one-hot encoding, and data augmentation ensure effective training. Transfer learning with DenseNet121 offers a robust approach to achieving high performance. The exploratory analysis confirms the quality, diversity, and readiness of the dataset for training.