**Summary Report: Scene-Based Video Segmentation**

**1. Introduction**

The project focuses on **scene-based video segmentation** using structured data from a CSV file containing information on game settings and crowd presence. The objective is to classify each scene into four categories:

1. **Indoor - With People**
2. **Indoor - Without People**
3. **Outdoor - With People**
4. **Outdoor - Without People**

The solution processes the dataset to categorize each scene effectively based on the available attributes.

**2. Approach**

The solution follows a structured approach:

1. **Data Collection & Preparation:**
   * A CSV file is used as input, containing fields like Game\_Setting, Number\_of\_Players, and Crowd\_Presence.
   * Data is preprocessed to remove inconsistencies and ensure proper formatting.
2. **Scene Classification Logic:**
   * Each row is classified based on:
     + **Indoor/Outdoor** determined by Game\_Setting.
     + **With/Without People** determined by Number\_of\_Players and Crowd\_Presence.
3. **File Storage & Output:**
   * The processed dataset is saved locally as a CSV file for further analysis.

**3. Challenges Faced**

**1. Data Consistency Issues**

* Some datasets had missing values or inconsistent formatting.
* Solution: Implemented preprocessing steps to clean the data before classification.

**2. Handling Empty or Corrupted Files**

* Uploading an empty or incorrectly formatted file caused errors.
* Solution: Added error handling to check for file validity before processing.

**3. Large Dataset Processing**

* Large datasets slowed down classification.
* Solution: Optimized pandas operations to improve performance.

**4. Solutions Implemented**

* **Data Preprocessing:** Cleaned and structured data for consistency.
* **Error Handling:** Added warnings and checks for missing or incorrect files.
* **Optimized Computation:** Used efficient pandas functions for fast classification.
* **Flexible File Saving:** Allowed users to modify the save path as needed.

**5. Potential Improvements**

1. **Machine Learning-Based Classification**
   * Train a model to automatically classify scenes based on extracted features.
2. **Enhanced Data Visualization**
   * Implement visualization tools to display scene distribution and trends.
3. **Cloud Integration**
   * Enable cloud-based file uploads and storage for remote access.
4. **Multi-File Processing**
   * Allow users to process multiple datasets in a batch.

**6. Conclusion**

This project successfully classifies scenes based on structured CSV data. The approach ensures **ease of use**, **efficiency**, and **scalability**.