# Task Title: Implement a Simple Retrieval-Augmented Generation (RAG) System Using PDF and Vector DB

## Objective:

Build a simple RAG system that extracts relevant information from a PDF, stores it in a vector database, and uses this database to retrieve contextually relevant data to generate a response.

## Requirements:

1. Extract text from a PDF document.  
2. Store the extracted text as embeddings in a vector database (e.g., FAISS, Pinecone).  
3. Retrieve relevant content from the vector DB based on a query.  
4. Use a language model to generate a response using the retrieved data.

## Deadline:

2 hours

# Task 2: Include Text and Word File

## Objective:

Include the text extracted from the PDF and the Word file generated in Task 1 as part of this task. Ensure that the content from the RAG system is properly incorporated and stored as part of the final deliverable.

## Requirements:

1. Extract text from the PDF as done in Task 1.  
2. Now use text and word file and extract the text and store the embeddings in the vector DB.

3. Retrieve relevant content from the vector DB based on a query.  
4. Use a language model to generate a response using the retrieved data.

## Deadline:

2 hours