Learning Plan for Cosine Similarity

This plan is designed to help you understand the concept of Cosine Similarity in the field of data science. The three-hour plan is divided into three phases, covering theoretical understanding, practical application, and deepening your knowledge with real-world examples.

# 09:30 - 10:00

## Activity:

Theoretical Understanding

## Details:

1. Review the concept of Cosine Similarity: Understand it as a measure used to evaluate the similarity between two vectors in a vector space. The closer the result is to 1, the more similar the vectors are.  
2. Why it is used: Commonly used in Natural Language Processing (NLP) for search, information retrieval, and determining the similarity between texts.

## Resources:

- Medium article on Cosine Similarity: https://medium.com  
- YouTube video on 'Cosine Similarity Explained' from educational channels like 'StatQuest' or '3Blue1Brown'.

# 10:00 - 11:00

## Activity:

Practical Application

## Details:

1. Apply Cosine Similarity using Python:  
 - Open Jupyter Notebook or your preferred development environment.  
 - Try calculating Cosine Similarity using the sklearn library:  
 ```python  
 from sklearn.metrics.pairwise import cosine\_similarity  
 import numpy as np  
  
 vector1 = np.array([[1, 2, 3]])  
 vector2 = np.array([[4, 5, 6]])  
  
 similarity = cosine\_similarity(vector1, vector2)  
 print(similarity)  
 ```  
 - Try applying it to texts using the TfidfVectorizer from the same library.

## Resources:

- scikit-learn documentation on `cosine\_similarity`: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.pairwise.cosine\_similarity.html  
- Data School's YouTube video on using Cosine Similarity with text data.

# 11:00 - 12:30

## Activity:

Deepening Understanding and Practical Experiments

## Details:

1. Apply Cosine Similarity to a real-world dataset:  
 - Download a text dataset from Kaggle or use a small collection of texts you compile yourself.  
 - Try applying Cosine Similarity between different text documents to analyze similarities between articles or sentences.  
2. Read practical examples:  
 - Read an article or two about using Cosine Similarity in NLP projects, such as recommendation systems or text search.

## Resources:

- Kaggle for datasets: https://www.kaggle.com  
- Stack Overflow for asking questions if you encounter problems.