The main purpose of the project is to find movies similar to each other. The project contains one dataset from Wikipedia and IMDb to investigate and find movies similar to each other. To answer this question, I used a TFIDF (Term Frequency–Inverse Document Frequency) that determines the most important words in each movie plot and it ignores the stopwords like “the”. The TFIDF vectorizer also contained tokenization and a stemmer function that tokenizes and transforms the words into their roots. Also, KMeans were used to create clusters and see the relationships between movie summaries. Moreover, the cosine similarity method was used to calculate how close each sentence is close to each in every movie plot. Lastly, I used the dendrogram plot to represent similar movies connected. The main result of the project is to have a diagram that shows similar movies connected. These results can be used while building a recommender system for movie recommendations websites.