1. Load the dataset from the given file “1000_movies_data.csv” and perform the following operations:
    a. Retrieve only “Action” movies
    b. Retrieve only “Action + Sci-Fi” movies
    c. Retrieve the list of movies which are released after the year 2015
    d. Retrieve the movies which has the maximum runtime (in Minutes)
    e. Retrieve the list of movies which has got rating more than 7.5
    f. Retrieve the top 3 movies which has got more revenue (in Millions) in 2014, 2015 and 2016
    g. Retrieve the top 3 movies in which “Will Smith” acted and has the rating more than 7.0
    h. Retrieve the list of movies which is directed by “James Gray” and has metascore more than 70
    i. Find the total number of null entries in each and every column
    j. Retrieve the list of movies whose names are starting with “The”

2. Load the dataset from the given file “diabetes.csv” and perform the following operations:
    a. Assign the dummy column names to each column (for Eg: col_1, col_2, col_3, …, col_9). Note: the last column is the target label.
    b. Retrieve the count of each target label i.e., how many 0’s and how many 1’s
    c. Analyse the outliers on every input column and visualize the outliers via “box plot”.

3. Extract only the columns “col_2, col_3 and col_4” from the dataset “diabetes.csv” and store the dataset in the file named “cluster_data.csv”. Create 3 clusters on the dataset using “k-means clustering”.