

Name: Ipsa Mishra
Student ID: 1001759616

Data Mining: Assignment 1 Report

This report consists of solutions to k-means and text clustering in Data Mining.

In problem 1, a Python program (along with Numpy implementation and matplotlib implementation) has been developed for generating data and clustering that data into k clusters. The iteration is terminated when the Euclidean distance between the previous center and an updated center is ≤ 0.001 or when the number of iterations reaches 10000. Then the code is applied to the data generated with $k = 2$ and initial centers $c_1 = (10, 10)$ and $c_2 = (-10, -10)$ and data generated with $k = 4$ and initial centers $c_1 = (10, 10)$ and $c_2 = (-10, -10)$, $c_3 = (10, -10)$ and $c_4 = (-10, 10)$ for finding out the centers of each cluster.

In problem 2, a Python program (along with Numpy implementation) has been developed for finding the weight matrix of all the words present in Amazon_Reviews.csv file using tf-idf weight. Tf is text frequency and Idf is inverse document Frequency. Tf determines the number of times a term appears in a document and idf determines the importance of a term in a document. The tf-idf weight is the product of two statistics- Term Frequency and Inverse Document Frequency. Then we make our own list of five positive words and five negative words from the reviews file. Then we have represented each review in a vector space of these ten words (i.e., count matrix) as well as tf-idf weight matrix. Then for each review, we have summed up the frequency of positive words and negative words and then represented each review as a vector of length 2. Then I have applied code of Problem 1 to this 2D data to find out the centers of each cluster for $k=2,3,4$ and randomly initialized centers.

Problem 1

(k-means, 50pts) Generate 2 sets of 2-D Gaussian random data, each set containing 500 samples using parameters below.

$$\mu_1 = [1, 0] \quad \mu_2 = [0, 1.5] \quad \Sigma_1 = [[0.9, 0.4], [0.4, 0.9]] \quad \Sigma_2 = [[0.9, 0.4], [0.4, 0.9]]$$

1. Write a function `cluster = mykmeans(X, k, c)` that clusters data $X \in \mathbb{R}^{n \times p}$ (n number of objects and p number of attributes) into k clusters. The c here is the initial centers, although this is usually not necessary, we will need it to test your function. Terminate the iteration when the l2-norm between a previous center and an updated center is ≤ 0.001 or the number of iteration reaches 10000.

For generating Data:

```
def generate_data(mu1, sigma1, mu2, sigma2, n_samples):
```

```
    x1 = np.random.multivariate_normal(mu1, sigma1, n_samples)
```

```
    x2 = np.random.multivariate_normal(mu2, sigma2, n_samples)
```

```
    x = np.concatenate((x1, x2), axis = 0)
```

```
    return x
```

For K-means (Cluster function):

```
def mykmeans(X, k, c):

    thres = 0.001
    n_iter = 10000
    n_X = np.shape(X)[0]
    tags = np.zeros(n_X)
    iter_centers = []
    for nc in range(len(c)):
        iter_centers.append([])

    for nc in range(len(c)):
        iter_centers[nc].append(list(c[nc]))

    for i in range(n_iter):
        for j in range(n_X):
            dist = [np.linalg.norm(X[j] - c[n]) for n in range(len(c))]
            indx = np.where(dist == np.amin(dist))
            indx = list(indx[0])
            tags[j] = indx[0]

        check = 0
        upd_c = c * 0.0
        for j in range(len(c)):
            identify = (tags == j) + 0.0
            upd_c[j] = np.matmul(np.transpose(identify), X) / np.sum(identify)
            if np.linalg.norm(upd_c[j]-c[j]) <= thres:
                check = 1

        c = upd_c
        for nc in range(len(c)):
            iter_centers[nc].append(list(c[nc]))

    if check == 1:
        for j in range(n_X):
            dist = [np.linalg.norm(X[j] - c[n]) for n in range(len(c))]
            indx = np.where(dist == np.amin(dist))
            indx = list(indx[0])
            tags[j] = indx[0]
        break

    iter_centers = np.array(iter_centers)

    for nc in range(len(c)):
        plt.scatter(iter_centers[nc,:,0],iter_centers[nc,:,1],marker = '.')

    print('# of iterations = ', i)
    return c, tags
```

2. Apply your code to the data generated above with $k = 2$ and initial centers $c_1 = (10, 10)$ and $c_2 = (-10, -10)$. In your report, report the centers found for each cluster. How many iterations did it take? Show a scatter plot of the data and the centers of clusters found.

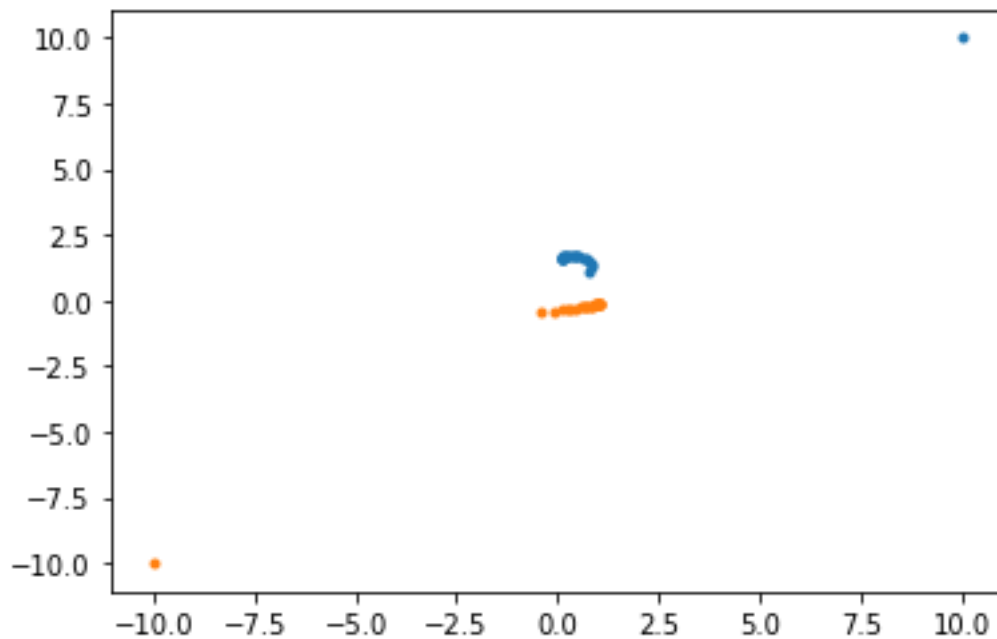
We have hard coded the value for k and the initial centers. When we need to report the centers for each cluster in this bit, we need to change the value of k and comment out the initial centers given in the third bit. Run in the terminal using following command

Python Problem_1.py

Number of Iterations: 19

The centers are: $\begin{bmatrix} 0.08920028 & 1.63562625 \\ 1.01262901 & -0.1616575 \end{bmatrix}$

Scatter Plot:



3. Apply your code to the data generated above with $k = 4$ and initial centers $c_1 = (10, 10)$ and $c_2 = (-10, -10)$, $c_3 = (10, -10)$ and $c_4 = (-10, 10)$. In your report, report the centers found for each cluster. How many iterations did it take? Show a scatter plot of the data and the centers of clusters found.

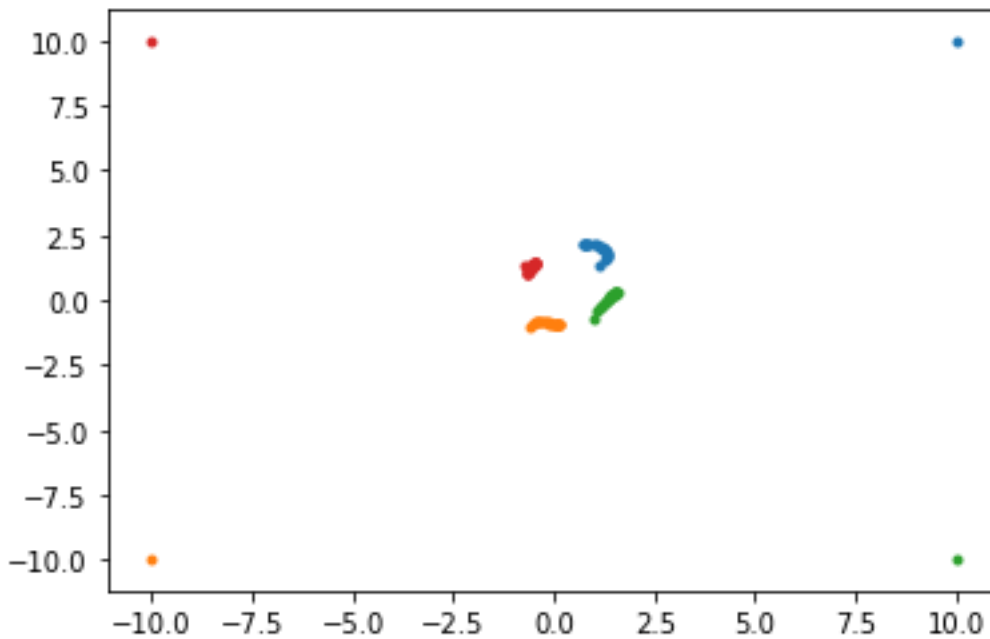
We have hard coded the value for k and the initial centers. When we need to report the centers for each cluster in this bit, we need to comment out the values of initial centers given in the second bit and change the value of k to 4. Run in the terminal using following command

Python Problem_1.py

Number of iterations: 19

The centers are: $\begin{bmatrix} 0.73117576 & 2.15659109 \\ 0.13478932 & -0.93907908 \\ 1.55459933 & 0.34268465 \\ -0.66593211 & 1.05488989 \end{bmatrix}$

Scatter Plot:



Problem 2

1. Let us focus on the reviews from the dataset without their labels. Build a weight matrix about all the words in the review part. Represent each review as a real-valued vector of tf-idf. Report each preprocessing step you applied and attach the corresponding part of your code. Visualize the matrix with color code (i.e., show the matrix as a 2D image where pixel intensity represents the weight).

Weight Matrix:

Review 0 : [0. 0. 0. ... 0. 0. 0.]
Review 1 : [0. 0. 0. ... 0. 0. 0.]
Review 2 : [0. 0. 0. ... 0. 0. 0.]
Review 3 : [0. 0. 0. ... 0. 0. 0.]
Review 4 : [0. 0. 0. ... 0. 0. 0.]
Review 5 : [0. 0. 0. ... 0. 0. 0.]

Review 6 : [0. 0. 0. ... 0. 0. 0.]
Review 7 : [0. 0. 0. ... 0. 0. 0.]
Review 8 : [0. 0. 0. ... 0. 0. 0.]
Review 9 : [0. 0. 0. ... 0. 0. 0.]
Review 10 : [0. 0. 0. ... 0. 0. 0.]
Review 11 : [0. 0. 0. ... 0. 0. 0.]
Review 12 : [0. 0. 0. ... 0. 0. 0.]
Review 13 : [0. 0. 0. ... 0. 0. 0.]
Review 14 : [0. 0. 0. ... 0. 0. 0.]
Review 15 : [0. 0. 0. ... 0. 0. 0.]
Review 16 : [0. 0. 0. ... 0. 0. 0.]
Review 17 : [0. 0. 0. ... 0. 0. 0.]
Review 18 : [0. 0. 0. ... 0. 0. 0.]
Review 19 : [0. 0. 0. ... 0. 0. 0.]
Review 20 : [0. 0. 0. ... 0. 0. 0.]
Review 21 : [0. 0. 0. ... 0. 0. 0.]
Review 22 : [0. 0. 0. ... 0. 0.032 0.]
Review 23 : [0. 0. 0. ... 0. 0. 0.]
Review 24 : [0. 0. 0. ... 0. 0. 0.]
Review 25 : [0. 0. 0. ... 0. 0. 0.]
Review 26 : [0. 0. 0. ... 0. 0. 0.]
Review 27 : [0. 0. 0. ... 0. 0. 0.]
Review 28 : [0. 0. 0. ... 0. 0. 0.]
Review 29 : [0. 0. 0. ... 0. 0. 0.]
Review 30 : [0. 0. 0. ... 0. 0. 0.]
Review 31 : [0. 0. 0. ... 0. 0. 0.]
Review 32 : [0. 0. 0.095 ... 0. 0. 0.]
Review 33 : [0. 0. 0. ... 0. 0. 0.]
Review 34 : [0. 0. 0. ... 0. 0. 0.]
Review 35 : [0. 0. 0. ... 0. 0. 0.]
Review 36 : [0. 0. 0. ... 0. 0. 0.]
Review 37 : [0. 0. 0. ... 0. 0. 0.]
Review 38 : [0. 0. 0. ... 0. 0. 0.]
Review 39 : [0. 0. 0. ... 0. 0. 0.]
Review 40 : [0. 0. 0. ... 0. 0. 0.]
Review 41 : [0. 0. 0. ... 0.033 0. 0.]
Review 42 : [0. 0. 0. ... 0. 0. 0.]
Review 43 : [0. 0. 0. ... 0. 0. 0.]
Review 44 : [0. 0. 0. ... 0. 0. 0.]
Review 45 : [0. 0. 0. ... 0. 0. 0.]
Review 46 : [0. 0. 0. ... 0. 0. 0.]
Review 47 : [0. 0. 0. ... 0. 0. 0.]
Review 48 : [0. 0. 0. ... 0. 0. 0.]
Review 49 : [0. 0. 0. ... 0. 0. 0.]
Review 50 : [0. 0. 0. ... 0. 0. 0.]
Review 51 : [0. 0. 0. ... 0. 0. 0.]
Review 52 : [0. 0. 0. ... 0. 0. 0.]
Review 53 : [0. 0. 0. ... 0. 0. 0.]
Review 54 : [0. 0. 0. ... 0. 0. 0.]
Review 55 : [0. 0. 0. ... 0. 0. 0.]
Review 56 : [0. 0. 0. ... 0. 0. 0.]
Review 57 : [0. 0. 0. ... 0. 0. 0.]
Review 58 : [0. 0. 0. ... 0. 0. 0.]

Review 59 : [0. 0. 0. ... 0. 0. 0.]
Review 60 : [0. 0. 0. ... 0. 0. 0.]
Review 61 : [0. 0. 0. ... 0. 0. 0.]
Review 62 : [0. 0. 0. ... 0. 0. 0.]
Review 63 : [0. 0. 0. ... 0. 0. 0.]
Review 64 : [0. 0. 0. ... 0. 0. 0.]
Review 65 : [0. 0. 0. ... 0. 0. 0.]
Review 66 : [0. 0. 0. ... 0. 0. 0.]
Review 67 : [0. 0. 0. ... 0. 0. 0.]
Review 68 : [0. 0. 0. ... 0. 0. 0.]
Review 69 : [0. 0. 0. ... 0. 0. 0.]
Review 70 : [0. 0. 0. ... 0. 0. 0.]
Review 71 : [0. 0. 0. ... 0. 0. 0.]
Review 72 : [0. 0. 0. ... 0. 0. 0.]
Review 73 : [0. 0. 0. ... 0. 0. 0.]
Review 74 : [0. 0. 0. ... 0. 0. 0.]
Review 75 : [0. 0. 0. ... 0. 0. 0.]
Review 76 : [0. 0. 0. ... 0. 0. 0.]
Review 77 : [0. 0. 0. ... 0. 0. 0.]
Review 78 : [0. 0. 0. ... 0. 0. 0.]
Review 79 : [0. 0. 0. ... 0. 0. 0.]
Review 80 : [0. 0. 0. ... 0. 0. 0.]
Review 81 : [0. 0. 0. ... 0. 0. 0.]
Review 82 : [0. 0. 0. ... 0. 0. 0.]
Review 83 : [0. 0. 0. ... 0. 0. 0.]
Review 84 : [0. 0. 0. ... 0. 0. 0.]
Review 85 : [0. 0. 0. ... 0. 0. 0.]
Review 86 : [0. 0. 0. ... 0. 0. 0.]
Review 87 : [0. 0. 0. ... 0. 0. 0.]
Review 88 : [0. 0. 0. ... 0. 0. 0.]
Review 89 : [0. 0. 0. ... 0. 0. 0.]
Review 90 : [0. 0. 0. ... 0. 0. 0.]
Review 91 : [0. 0. 0. ... 0. 0. 0.]
Review 92 : [0. 0. 0. ... 0. 0. 0.]
Review 93 : [0. 0. 0. ... 0. 0. 0.062]
Review 94 : [0. 0. 0. ... 0. 0. 0.]
Review 95 : [0. 0. 0. ... 0. 0. 0.]
Review 96 : [0. 0. 0. ... 0. 0. 0.]
Review 97 : [0. 0. 0. ... 0. 0. 0.]
Review 98 : [0. 0. 0. ... 0. 0. 0.]
Review 99 : [0. 0. 0. ... 0. 0. 0.]
Review 100 : [0. 0. 0. ... 0. 0. 0.]
Review 101 : [0. 0. 0. ... 0. 0. 0.]
Review 102 : [0. 0. 0. ... 0. 0. 0.]
Review 103 : [0. 0. 0. ... 0. 0. 0.]
Review 104 : [0. 0. 0. ... 0. 0. 0.]
Review 105 : [0. 0. 0. ... 0. 0. 0.]
Review 106 : [0. 0. 0. ... 0. 0. 0.]
Review 107 : [0. 0. 0. ... 0. 0. 0.]
Review 108 : [0. 0. 0. ... 0. 0. 0.]
Review 109 : [0. 0. 0. ... 0. 0. 0.]
Review 110 : [0. 0. 0. ... 0. 0. 0.]
Review 111 : [0. 0. 0. ... 0. 0. 0.]

Review 112 : [0. 0. 0. ... 0. 0. 0.]
Review 113 : [0. 0. 0. ... 0. 0. 0.]
Review 114 : [0. 0. 0. ... 0. 0. 0.]
Review 115 : [0. 0. 0. ... 0. 0. 0.]
Review 116 : [0. 0. 0. ... 0. 0. 0.]
Review 117 : [0. 0. 0. ... 0. 0. 0.]
Review 118 : [0. 0. 0. ... 0. 0. 0.]
Review 119 : [0. 0. 0. ... 0. 0. 0.]
Review 120 : [0. 0. 0. ... 0. 0. 0.]
Review 121 : [0. 0. 0. ... 0. 0. 0.]
Review 122 : [0. 0. 0. ... 0. 0. 0.]
Review 123 : [0. 0. 0. ... 0. 0. 0.]
Review 124 : [0. 0. 0. ... 0. 0. 0.]
Review 125 : [0. 0. 0. ... 0. 0. 0.]
Review 126 : [0. 0. 0. ... 0. 0. 0.]
Review 127 : [0. 0. 0. ... 0. 0. 0.]
Review 128 : [0. 0. 0. ... 0. 0. 0.]
Review 129 : [0. 0. 0. ... 0. 0. 0.]
Review 130 : [0. 0. 0. ... 0. 0. 0.]
Review 131 : [0. 0. 0. ... 0. 0. 0.]
Review 132 : [0. 0. 0. ... 0. 0. 0.]
Review 133 : [0. 0. 0. ... 0. 0. 0.]
Review 134 : [0. 0. 0. ... 0. 0. 0.]
Review 135 : [0. 0. 0. ... 0. 0. 0.]
Review 136 : [0. 0. 0. ... 0. 0. 0.]
Review 137 : [0. 0. 0. ... 0. 0. 0.]
Review 138 : [0. 0. 0. ... 0. 0. 0.]
Review 139 : [0. 0. 0. ... 0. 0. 0.]
Review 140 : [0. 0. 0. ... 0. 0. 0.]
Review 141 : [0. 0. 0. ... 0. 0. 0.]
Review 142 : [0. 0. 0. ... 0. 0. 0.]
Review 143 : [0. 0. 0. ... 0. 0. 0.]
Review 144 : [0. 0. 0. ... 0. 0. 0.]
Review 145 : [0. 0. 0. ... 0. 0. 0.]
Review 146 : [0.031 0. 0. ... 0. 0. 0.]
Review 147 : [0. 0. 0. ... 0. 0. 0.]
Review 148 : [0. 0. 0. ... 0. 0. 0.]
Review 149 : [0. 0. 0. ... 0. 0. 0.]
Review 150 : [0. 0. 0. ... 0. 0. 0.]
Review 151 : [0. 0. 0. ... 0. 0. 0.]
Review 152 : [0. 0. 0. ... 0. 0. 0.]
Review 153 : [0. 0.084 0. ... 0. 0. 0.]
Review 154 : [0. 0. 0. ... 0. 0. 0.]
Review 155 : [0. 0. 0. ... 0. 0. 0.]
Review 156 : [0. 0. 0. ... 0. 0. 0.]
Review 157 : [0. 0. 0. ... 0. 0. 0.]
Review 158 : [0. 0. 0. ... 0. 0. 0.]
Review 159 : [0. 0. 0. ... 0. 0. 0.]
Review 160 : [0. 0. 0. ... 0. 0. 0.]
Review 161 : [0. 0. 0. ... 0. 0. 0.]
Review 162 : [0. 0. 0. ... 0. 0. 0.]
Review 163 : [0. 0. 0. ... 0. 0. 0.]
Review 164 : [0. 0. 0. ... 0. 0. 0.]

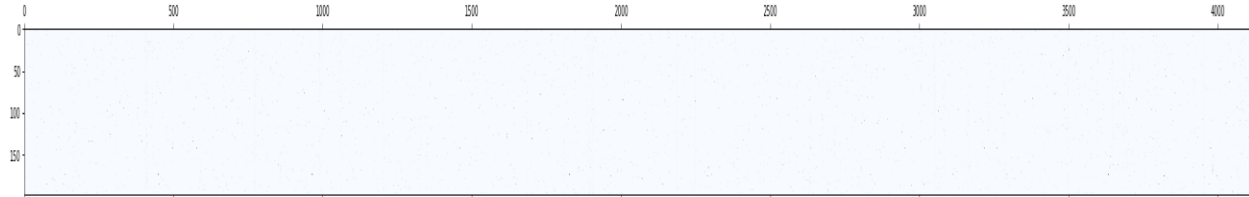
Review 165 : [0. 0. 0. ... 0. 0. 0.]
 Review 166 : [0. 0. 0. ... 0. 0. 0.]
 Review 167 : [0. 0. 0. ... 0. 0. 0.]
 Review 168 : [0. 0. 0. ... 0. 0. 0.]
 Review 169 : [0. 0. 0. ... 0. 0. 0.]
 Review 170 : [0. 0. 0. ... 0. 0. 0.]
 Review 171 : [0. 0. 0. ... 0. 0. 0.]
 Review 172 : [0. 0. 0. ... 0. 0. 0.]
 Review 173 : [0. 0. 0. ... 0. 0. 0.]
 Review 174 : [0. 0. 0. ... 0. 0. 0.]
 Review 175 : [0. 0. 0. ... 0. 0. 0.]
 Review 176 : [0. 0. 0. ... 0. 0. 0.]
 Review 177 : [0. 0. 0. ... 0. 0. 0.]
 Review 178 : [0. 0. 0. ... 0. 0. 0.]
 Review 179 : [0. 0. 0. ... 0. 0. 0.]
 Review 180 : [0. 0. 0. ... 0. 0. 0.]
 Review 181 : [0. 0. 0. ... 0. 0. 0.]
 Review 182 : [0. 0. 0. ... 0. 0. 0.]
 Review 183 : [0. 0. 0. ... 0. 0. 0.]
 Review 184 : [0. 0. 0. ... 0. 0. 0.]
 Review 185 : [0. 0. 0. ... 0. 0. 0.038]
 Review 186 : [0. 0. 0. ... 0. 0. 0.]
 Review 187 : [0. 0. 0. ... 0. 0. 0.]
 Review 188 : [0. 0. 0. ... 0. 0. 0.]
 Review 189 : [0. 0. 0. ... 0. 0. 0.]
 Review 190 : [0. 0. 0. ... 0. 0. 0.]
 Review 191 : [0. 0. 0. ... 0. 0. 0.]
 Review 192 : [0. 0. 0. ... 0. 0. 0.]
 Review 193 : [0. 0. 0. ... 0. 0. 0.]
 Review 194 : [0. 0. 0. ... 0. 0. 0.]
 Review 195 : [0. 0. 0. ... 0. 0. 0.]
 Review 196 : [0. 0. 0. ... 0. 0. 0.]
 Review 197 : [0. 0. 0. ... 0. 0. 0.]
 Review 198 : [0. 0. 0. ... 0. 0. 0.]

Preprocessing steps:

```

text= text.replace('.', '').replace('!', '').replace('""', '').replace('; ', ' ').replace(' ', ' ')    ## Removing puctuations &
double quotes
text = text.replace('(', '').replace(')', '').replace('[', '').replace(']', '').replace('{', '').replace('}', '')    ##Removing
parenthesis
lines = lines[1:-1]    ## Remove final empty line and initial line which has header
review_label = l.split('__label__')    ## Split each line into review and label using '__label__' as delimiter
  
```


Color Plot:



2. Pick your own 5 \positive" words and 5 \negative" words, which indicate if a product is good or bad, respectively. List the words you selected. Represent each review in a vector space of these ten words (i.e., count matrix) as well as tf-idf weight matrix.

```
bagofWords_pos = ['beautiful','good','awesome','great','amazing']  
bagofWords_neg = ['error','bad','beware','poor','incorrect']
```

```
## Positive Words  
## Negative Words
```

Count Matrix:

```
Review 0 : [1. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 1 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 2 : [0. 0. 0. 0. 1. 0. 0. 0. 0. 0.]  
Review 3 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 4 : [0. 1. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 5 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 6 : [0. 0. 0. 0. 0. 0. 1. 0. 0. 0.]  
Review 7 : [0. 0. 0. 0. 1. 0. 0. 0. 0. 0.]  
Review 8 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 9 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 10 : [0. 0. 0. 0. 0. 0. 0. 0. 1. 0.]  
Review 11 : [0. 0. 0. 1. 0. 0. 0. 0. 0. 0.]  
Review 12 : [0. 0. 0. 0. 0. 1. 0. 0. 0. 0.]  
Review 13 : [0. 0. 0. 0. 0. 0. 0. 1. 0. 0.]  
Review 14 : [0. 1. 0. 0. 0. 0. 0. 0. 2. 0.]  
Review 15 : [0. 0. 0. 0. 0. 0. 0. 0. 1. 0.]  
Review 16 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 17 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 18 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 19 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 20 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 21 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 22 : [0. 1. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 23 : [0. 0. 0. 1. 0. 0. 0. 0. 0. 0.]  
Review 24 : [0. 0. 0. 0. 1. 0. 0. 0. 0. 0.]  
Review 25 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 26 : [0. 0. 0. 0. 0. 0. 1. 0. 0. 0.]  
Review 27 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 28 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 29 : [0. 1. 0. 0. 0. 0. 0. 0. 0. 0.]  
Review 30 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 1.]  
Review 31 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
```

Review 32 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 33 : [0. 1. 0. 6. 0. 0. 0. 0. 0.]
Review 34 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 35 : [0. 0. 0. 0. 0. 0. 1. 0. 0.]
Review 36 : [0. 0. 0. 2. 0. 0. 0. 0. 0.]
Review 37 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 38 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 39 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 40 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 41 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 42 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 43 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 44 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 45 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 46 : [0. 0. 0. 1. 0. 0. 1. 0. 1.]
Review 47 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 48 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 49 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 50 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 51 : [0. 0. 0. 2. 0. 0. 0. 0. 0.]
Review 52 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 53 : [0. 2. 0. 2. 0. 0. 0. 0. 0.]
Review 54 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 55 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 56 : [0. 0. 0. 0. 1. 0. 0. 0. 0.]
Review 57 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 58 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 59 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 60 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 61 : [0. 0. 0. 0. 0. 0. 0. 0. 1.]
Review 62 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 63 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 64 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 65 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 66 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 67 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 68 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 69 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 70 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 71 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 72 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 73 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 74 : [0. 1. 0. 0. 0. 0. 1. 0. 0.]
Review 75 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 76 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 77 : [1. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 78 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 79 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 80 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 81 : [0. 0. 0. 2. 0. 0. 0. 0. 0.]
Review 82 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 83 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 84 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]

Review 85 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 86 : [0. 0. 0. 0. 0. 0. 1. 0. 0.]
Review 87 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 88 : [0. 0. 0. 0. 0. 0. 0. 1. 0.]
Review 89 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 90 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 91 : [0. 0. 0. 0. 1. 0. 0. 0. 0.]
Review 92 : [0. 0. 0. 0. 1. 0. 0. 0. 0.]
Review 93 : [0. 0. 0. 0. 0. 0. 1. 0. 0.]
Review 94 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 95 : [0. 1. 0. 0. 0. 0. 0. 0. 0.]
Review 96 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 97 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 98 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 99 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 100 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 101 : [0. 0. 0. 3. 0. 0. 0. 0. 0.]
Review 102 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 103 : [0. 0. 0. 0. 0. 0. 0. 0. 1.]
Review 104 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 105 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 106 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 107 : [0. 0. 0. 0. 0. 0. 0. 0. 1.]
Review 108 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 109 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 110 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 111 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 112 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 113 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 114 : [0. 0. 1. 0. 0. 0. 0. 0. 0.]
Review 115 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 116 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 117 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 118 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 119 : [0. 1. 0. 0. 0. 0. 0. 0. 0.]
Review 120 : [0. 0. 0. 3. 0. 0. 1. 0. 0.]
Review 121 : [0. 1. 0. 0. 0. 0. 0. 0. 0.]
Review 122 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 123 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 124 : [0. 1. 0. 0. 0. 0. 0. 0. 0.]
Review 125 : [0. 1. 0. 1. 0. 0. 0. 0. 0.]
Review 126 : [0. 0. 0. 0. 1. 0. 0. 0. 0.]
Review 127 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 128 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 129 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 130 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 131 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 132 : [0. 0. 0. 0. 0. 0. 0. 0. 1.]
Review 133 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 134 : [0. 0. 0. 1. 0. 0. 0. 0. 1.]
Review 135 : [0. 0. 1. 1. 0. 0. 0. 0. 0.]
Review 136 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 137 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]

Review 138 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 139 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 140 : [0. 1. 0. 0. 0. 0. 0. 0. 0.]
Review 141 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 142 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 143 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 144 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 145 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 146 : [0. 0. 0. 0. 0. 0. 2. 0. 0.]
Review 147 : [0. 0. 0. 0. 0. 0. 0. 0. 1.]
Review 148 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 149 : [0. 0. 0. 1. 0. 0. 1. 0. 0.]
Review 150 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 151 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 152 : [0. 0. 0. 0. 1. 0. 1. 0. 0.]
Review 153 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 154 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 155 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 156 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 157 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 158 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 159 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 160 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 161 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 162 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 163 : [1. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 164 : [0. 0. 0. 2. 0. 0. 0. 0. 0.]
Review 165 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 166 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 167 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 168 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 169 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 170 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 171 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 172 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 173 : [0. 0. 0. 1. 0. 0. 0. 0. 0.]
Review 174 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 175 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 176 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 177 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 178 : [0. 1. 0. 0. 0. 0. 0. 0. 0.]
Review 179 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 180 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 181 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 182 : [0. 1. 0. 0. 0. 0. 0. 1. 0.]
Review 183 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 184 : [0. 0. 0. 2. 0. 0. 2. 0. 0.]
Review 185 : [0. 2. 0. 0. 0. 0. 1. 0. 0.]
Review 186 : [0. 2. 0. 0. 0. 0. 0. 0. 0.]
Review 187 : [0. 1. 0. 0. 0. 0. 1. 0. 0.]
Review 188 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 189 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 190 : [0. 0. 0. 0. 0. 0. 1. 0. 0.]

Review 191 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 192 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 193 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 194 : [0. 0. 0. 1. 0. 0. 0. 0. 0. 0.]
Review 195 : [0. 0. 0. 2. 1. 0. 0. 0. 0. 0.]
Review 196 : [0. 1. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 197 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 198 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]

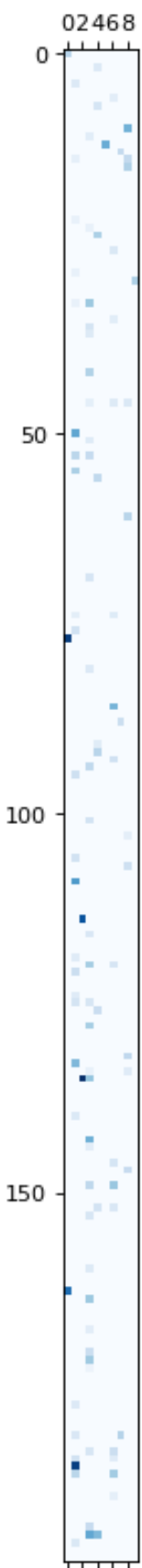
tf-idf weight matrix:

Review 0 : [0.052 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 1 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 2 : [0. 0. 0. 0. 0.024 0. 0. 0. 0. 0.]
Review 3 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 4 : [0. 0.024 0. 0. 0. 0. 0. 0. 0. 0.]
Review 5 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 6 : [0. 0. 0. 0. 0. 0. 0.019 0. 0. 0.]
Review 7 : [0. 0. 0. 0. 0.03 0. 0. 0. 0. 0.]
Review 8 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 9 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 10 : [0. 0. 0. 0. 0. 0. 0. 0.085 0.]
Review 11 : [0. 0. 0. 0.019 0. 0. 0. 0. 0.]
Review 12 : [0. 0. 0. 0. 0.09 0. 0. 0. 0.]
Review 13 : [0. 0. 0. 0. 0. 0. 0.045 0. 0.]
Review 14 : [0. 0.016 0. 0. 0. 0. 0. 0.045 0.]
Review 15 : [0. 0. 0. 0. 0. 0. 0. 0.056 0.]
Review 16 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 17 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 18 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 19 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 20 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 21 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 22 : [0. 0.013 0. 0. 0. 0. 0. 0. 0. 0.]
Review 23 : [0. 0. 0. 0.013 0. 0. 0. 0. 0.]
Review 24 : [0. 0. 0. 0. 0.058 0. 0. 0. 0. 0.]
Review 25 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 26 : [0. 0. 0. 0. 0. 0. 0.026 0. 0. 0.]
Review 27 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 28 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 29 : [0. 0.013 0. 0. 0. 0. 0. 0. 0. 0.]
Review 30 : [0. 0. 0. 0. 0. 0. 0. 0. 0.058]
Review 31 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 32 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 33 : [0. 0.013 0. 0.066 0. 0. 0. 0. 0. 0.]
Review 34 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 35 : [0. 0. 0. 0. 0. 0. 0.019 0. 0. 0.]
Review 36 : [0. 0. 0. 0.032 0. 0. 0. 0. 0. 0.]
Review 37 : [0. 0. 0. 0.023 0. 0. 0. 0. 0. 0.]
Review 38 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 39 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 40 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 41 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 42 : [0. 0. 0. 0.055 0. 0. 0. 0. 0. 0.]

Review 43 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 44 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 45 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 46 : [0. 0. 0. 0.016 0. 0. 0.023 0. 0.026 0.]
Review 47 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 48 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 49 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 50 : [0. 0.094 0. 0. 0. 0. 0. 0. 0.]
Review 51 : [0. 0. 0. 0.022 0. 0. 0. 0. 0.]
Review 52 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 53 : [0. 0.05 0. 0.044 0. 0. 0. 0. 0.]
Review 54 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 55 : [0. 0.059 0. 0. 0. 0. 0. 0. 0.]
Review 56 : [0. 0. 0. 0. 0.047 0. 0. 0. 0. 0.]
Review 57 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 58 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 59 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 60 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 61 : [0. 0. 0. 0. 0. 0. 0. 0. 0.052 0.]
Review 62 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 63 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 64 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 65 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 66 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 67 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 68 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 69 : [0. 0. 0. 0.029 0. 0. 0. 0. 0. 0.]
Review 70 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 71 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 72 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 73 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 74 : [0. 0.017 0. 0. 0. 0. 0.022 0. 0. 0.]
Review 75 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 76 : [0. 0.033 0. 0. 0. 0. 0. 0. 0. 0.]
Review 77 : [0.168 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 78 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 79 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 80 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 81 : [0. 0. 0. 0.024 0. 0. 0. 0. 0. 0.]
Review 82 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 83 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 84 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 85 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 86 : [0. 0. 0. 0. 0. 0. 0.083 0. 0. 0.]
Review 87 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 88 : [0. 0. 0. 0. 0. 0. 0. 0.039 0. 0.]
Review 89 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 90 : [0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 91 : [0. 0. 0. 0. 0.02 0. 0. 0. 0. 0.]
Review 92 : [0. 0. 0. 0. 0.052 0. 0. 0. 0. 0.]
Review 93 : [0. 0. 0. 0. 0. 0. 0.035 0. 0. 0.]
Review 94 : [0. 0. 0. 0.047 0. 0. 0. 0. 0. 0.]
Review 95 : [0. 0.036 0. 0. 0. 0. 0. 0. 0. 0.]

Review 96 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 97 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 98 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 99 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 100 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 101 : [0. 0. 0. 0.033 0. 0. 0. 0. 0. 0.]
Review 102 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 103 : [0. 0. 0. 0. 0. 0. 0. 0. 0.017 0.]
Review 104 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 105 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 106 : [0. 0.034 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 107 : [0. 0. 0. 0. 0. 0. 0. 0. 0.037 0.]
Review 108 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 109 : [0. 0.104 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 110 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 111 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 112 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 113 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 114 : [0. 0. 0.153 0. 0. 0. 0. 0. 0. 0. 0.]
Review 115 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 116 : [0. 0. 0. 0.027 0. 0. 0. 0. 0. 0. 0.]
Review 117 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 118 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 119 : [0. 0.021 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 120 : [0. 0. 0. 0.062 0. 0. 0.029 0. 0. 0. 0.]
Review 121 : [0. 0.04 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 122 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 123 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 124 : [0. 0.021 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 125 : [0. 0.031 0. 0.028 0. 0. 0. 0. 0. 0. 0. 0.]
Review 126 : [0. 0. 0. 0. 0.04 0. 0. 0. 0. 0. 0.]
Review 127 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 128 : [0. 0. 0. 0.061 0. 0. 0. 0. 0. 0. 0.]
Review 129 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 130 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 131 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 132 : [0. 0. 0. 0. 0. 0. 0. 0. 0.047 0. 0.]
Review 133 : [0. 0.08 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 134 : [0. 0. 0. 0.013 0. 0. 0. 0. 0.021 0. 0.]
Review 135 : [0. 0. 0.177 0.07 0. 0. 0. 0. 0. 0. 0. 0.]
Review 136 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 137 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 138 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 139 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 140 : [0. 0.025 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 141 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 142 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 143 : [0. 0. 0. 0.087 0. 0. 0. 0. 0. 0. 0.]
Review 144 : [0. 0. 0. 0.018 0. 0. 0. 0. 0. 0. 0.]
Review 145 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 146 : [0. 0. 0. 0. 0. 0. 0.03 0. 0. 0. 0.]
Review 147 : [0. 0. 0. 0. 0. 0. 0. 0. 0.043 0. 0.]
Review 148 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]

Review 149 : [0. 0. 0. 0.048 0. 0. 0.068 0. 0. 0.]
Review 150 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 151 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 152 : [0. 0. 0. 0. 0.032 0. 0.027 0. 0. 0.]
Review 153 : [0. 0. 0. 0.029 0. 0. 0. 0. 0. 0.]
Review 154 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 155 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 156 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 157 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 158 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 159 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 160 : [0. 0. 0. 0.023 0. 0. 0. 0. 0. 0.]
Review 161 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 162 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 163 : [0.135 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 164 : [0. 0. 0. 0.066 0. 0. 0. 0. 0. 0.]
Review 165 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 166 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 167 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 168 : [0. 0. 0. 0.018 0. 0. 0. 0. 0. 0.]
Review 169 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 170 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 171 : [0. 0. 0. 0.039 0. 0. 0. 0. 0. 0.]
Review 172 : [0. 0. 0. 0.065 0. 0. 0. 0. 0. 0.]
Review 173 : [0. 0. 0. 0.01 0. 0. 0. 0. 0. 0.]
Review 174 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 175 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 176 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 177 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 178 : [0. 0.024 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 179 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 180 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 181 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 182 : [0. 0.027 0. 0. 0. 0. 0. 0.054 0. 0. 0.]
Review 183 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 184 : [0. 0. 0. 0.028 0. 0. 0.039 0. 0. 0.]
Review 185 : [0. 0.034 0. 0. 0. 0. 0.021 0. 0. 0.]
Review 186 : [0. 0.166 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 187 : [0. 0.051 0. 0. 0. 0. 0.063 0. 0. 0. 0.]
Review 188 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 189 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 190 : [0. 0. 0. 0. 0. 0. 0.016 0. 0. 0. 0.]
Review 191 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 192 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 193 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 194 : [0. 0. 0. 0.044 0. 0. 0. 0. 0. 0.]
Review 195 : [0. 0. 0. 0.094 0.079 0. 0. 0. 0. 0. 0.]
Review 196 : [0. 0.026 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 197 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
Review 198 : [0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]



3. For each review, sum up the frequency of \positive" words and \negative" words. Represent each review as a vector of length 2. Now the reviews can be shown in 2D space, while one dimension is about \positive" and the other one is \negative". Apply your code from Problem 1 to this 2D data with $k = 2, 3, 4$ with randomly initialized centers. In your report, report the centers found for each cluster. How many iterations did it take? Show a scatter plot of the data and the centers of clusters found.

Review 0 : [0.013 0.]
Review 1 : [0. 0.]
Review 2 : [0.008 0.]
Review 3 : [0. 0.]
Review 4 : [0.011 0.]
Review 5 : [0. 0.]
Review 6 : [0. 0.007]
Review 7 : [0.01 0.]
Review 8 : [0. 0.]
Review 9 : [0. 0.]
Review 10 : [0. 0.029]
Review 11 : [0.01 0.]
Review 12 : [0. 0.017]
Review 13 : [0. 0.011]
Review 14 : [0.008 0.015]
Review 15 : [0. 0.019]
Review 16 : [0. 0.]
Review 17 : [0. 0.]
Review 18 : [0. 0.]
Review 19 : [0. 0.]
Review 20 : [0. 0.]
Review 21 : [0. 0.]
Review 22 : [0.006 0.]
Review 23 : [0.007 0.]
Review 24 : [0.019 0.]
Review 25 : [0. 0.]
Review 26 : [0. 0.01]
Review 27 : [0. 0.]
Review 28 : [0. 0.]
Review 29 : [0.006 0.]
Review 30 : [0. 0.011]
Review 31 : [0. 0.]
Review 32 : [0. 0.]
Review 33 : [0.042 0.]
Review 34 : [0. 0.]
Review 35 : [0. 0.007]
Review 36 : [0.017 0.]
Review 37 : [0.013 0.]
Review 38 : [0. 0.]
Review 39 : [0. 0.]
Review 40 : [0. 0.]
Review 41 : [0. 0.]
Review 42 : [0.03 0.]
Review 43 : [0. 0.]
Review 44 : [0. 0.]
Review 45 : [0. 0.]

Review 46 : [0.009 0.018]
Review 47 : [0. 0.]
Review 48 : [0. 0.]
Review 49 : [0. 0.]
Review 50 : [0.045 0.]
Review 51 : [0.012 0.]
Review 52 : [0. 0.]
Review 53 : [0.048 0.]
Review 54 : [0. 0.]
Review 55 : [0.029 0.]
Review 56 : [0.015 0.]
Review 57 : [0. 0.]
Review 58 : [0. 0.]
Review 59 : [0. 0.]
Review 60 : [0. 0.]
Review 61 : [0. 0.017]
Review 62 : [0. 0.]
Review 63 : [0. 0.]
Review 64 : [0. 0.]
Review 65 : [0. 0.]
Review 66 : [0. 0.]
Review 67 : [0. 0.]
Review 68 : [0. 0.]
Review 69 : [0.016 0.]
Review 70 : [0. 0.]
Review 71 : [0. 0.]
Review 72 : [0. 0.]
Review 73 : [0. 0.]
Review 74 : [0.008 0.008]
Review 75 : [0. 0.]
Review 76 : [0.016 0.]
Review 77 : [0.04 0.]
Review 78 : [0. 0.]
Review 79 : [0. 0.]
Review 80 : [0. 0.]
Review 81 : [0.013 0.]
Review 82 : [0. 0.]
Review 83 : [0. 0.]
Review 84 : [0. 0.]
Review 85 : [0. 0.]
Review 86 : [0. 0.032]
Review 87 : [0. 0.]
Review 88 : [0. 0.009]
Review 89 : [0. 0.]
Review 90 : [0. 0.]
Review 91 : [0.007 0.]
Review 92 : [0.017 0.]
Review 93 : [0. 0.014]
Review 94 : [0.026 0.]
Review 95 : [0.018 0.]
Review 96 : [0. 0.]
Review 97 : [0. 0.]
Review 98 : [0. 0.]

Review 99 : [0. 0.]
Review 100 : [0. 0.]
Review 101 : [0.018 0.]
Review 102 : [0. 0.]
Review 103 : [0. 0.006]
Review 104 : [0. 0.]
Review 105 : [0. 0.]
Review 106 : [0.016 0.]
Review 107 : [0. 0.012]
Review 108 : [0. 0.]
Review 109 : [0.05 0.]
Review 110 : [0. 0.]
Review 111 : [0. 0.]
Review 112 : [0. 0.]
Review 113 : [0. 0.]
Review 114 : [0.033 0.]
Review 115 : [0. 0.]
Review 116 : [0.015 0.]
Review 117 : [0. 0.]
Review 118 : [0. 0.]
Review 119 : [0.01 0.]
Review 120 : [0.034 0.011]
Review 121 : [0.019 0.]
Review 122 : [0. 0.]
Review 123 : [0. 0.]
Review 124 : [0.01 0.]
Review 125 : [0.03 0.]
Review 126 : [0.013 0.]
Review 127 : [0. 0.]
Review 128 : [0.033 0.]
Review 129 : [0. 0.]
Review 130 : [0. 0.]
Review 131 : [0. 0.]
Review 132 : [0. 0.016]
Review 133 : [0.038 0.]
Review 134 : [0.007 0.007]
Review 135 : [0.077 0.]
Review 136 : [0. 0.]
Review 137 : [0. 0.]
Review 138 : [0. 0.]
Review 139 : [0. 0.]
Review 140 : [0.012 0.]
Review 141 : [0. 0.]
Review 142 : [0. 0.]
Review 143 : [0.048 0.]
Review 144 : [0.01 0.]
Review 145 : [0. 0.]
Review 146 : [0. 0.012]
Review 147 : [0. 0.014]
Review 148 : [0. 0.]
Review 149 : [0.026 0.026]
Review 150 : [0. 0.]
Review 151 : [0. 0.]

Review 152 : [0.01 0.01]
Review 153 : [0.016 0.]
Review 154 : [0. 0.]
Review 155 : [0. 0.]
Review 156 : [0. 0.]
Review 157 : [0. 0.]
Review 158 : [0. 0.]
Review 159 : [0. 0.]
Review 160 : [0.013 0.]
Review 161 : [0. 0.]
Review 162 : [0. 0.]
Review 163 : [0.032 0.]
Review 164 : [0.036 0.]
Review 165 : [0. 0.]
Review 166 : [0. 0.]
Review 167 : [0. 0.]
Review 168 : [0.01 0.]
Review 169 : [0. 0.]
Review 170 : [0. 0.]
Review 171 : [0.021 0.]
Review 172 : [0.036 0.]
Review 173 : [0.005 0.]
Review 174 : [0. 0.]
Review 175 : [0. 0.]
Review 176 : [0. 0.]
Review 177 : [0. 0.]
Review 178 : [0.011 0.]
Review 179 : [0. 0.]
Review 180 : [0. 0.]
Review 181 : [0. 0.]
Review 182 : [0.013 0.013]
Review 183 : [0. 0.]
Review 184 : [0.015 0.015]
Review 185 : [0.016 0.008]
Review 186 : [0.08 0.]
Review 187 : [0.024 0.024]
Review 188 : [0. 0.]
Review 189 : [0. 0.]
Review 190 : [0. 0.006]
Review 191 : [0. 0.]
Review 192 : [0. 0.]
Review 193 : [0. 0.]
Review 194 : [0.024 0.]
Review 195 : [0.077 0.]
Review 196 : [0.013 0.]
Review 197 : [0. 0.]
Review 198 : [0. 0.]

Centers found for each cluster:

k = 2

Number of iterations = 2

The centers are: [[0.008 0.]

[0.005 0.015]]

k = 3

Number of iterations = 2

The centers are: [[0.043 0.001]

[0.003 0.]

[0.006 0.016]]

k = 4

Number of iterations = 2

The centers are: [[0. 0.001]

[0.01 0.019]

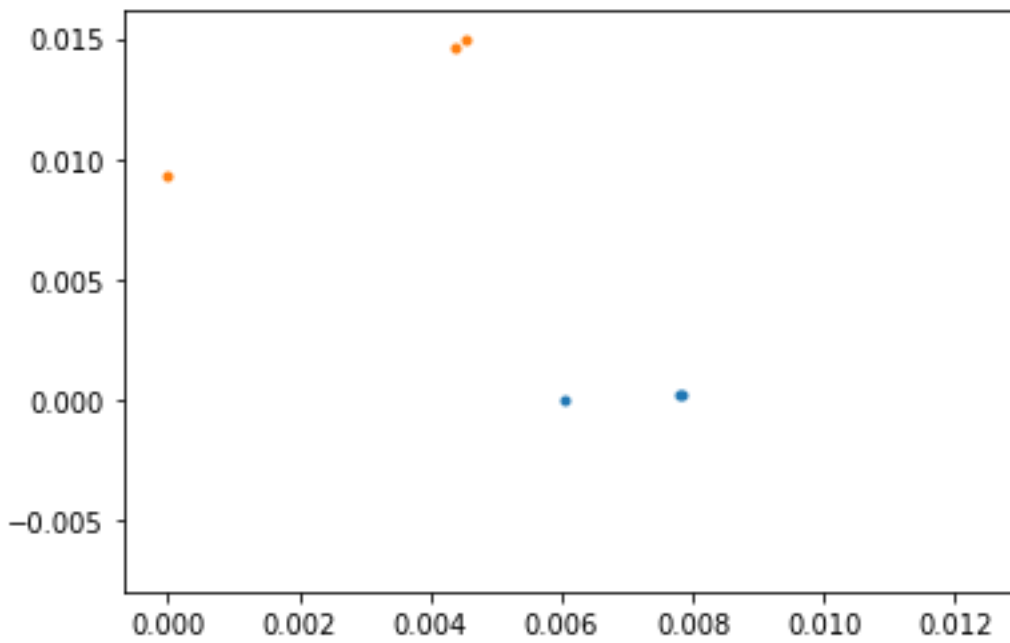
[0.044 0.001]

[0.014 0.]]

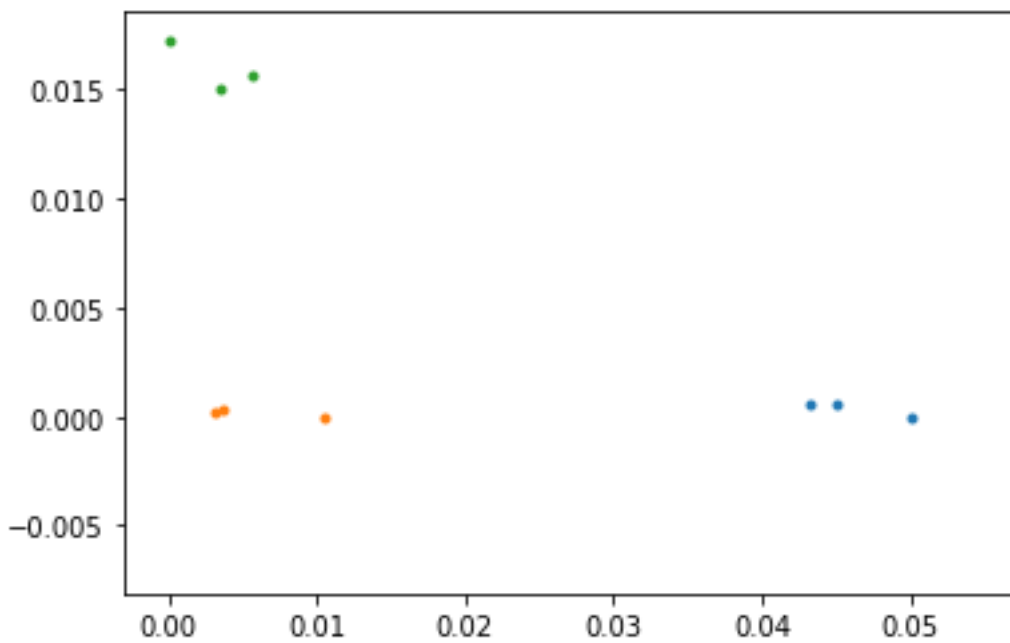
Scatter Plots (for k=2, k=3, k=4) on next page

Scatter Plots:

For k=2:



For k=3:



For k=4:

