Experiment no. 9

Name: Sonali Dattatray Kaingade

PRN: 21620002

Title: Find correlation between items/entities.

code:

```
#include <iostream>

#include <fstream>

#include <vector>

#include <sstream>


using namespace std;


vector<int> readData(const string &filename)

{

   vector<int> data;

   fstream file(filename, ios::in);


   if (!file.is_open())

   {

      cerr << "Error in opening input file: " << filename << endl;

      exit(1);

   }


   string line, value;

   int lineCount = 0;
```

```
    while (getline(file, line))
    {
        if (lineCount > 0)
        {
            stringstream ss(line);
            getline(ss, value, ',');
            data.push_back(stoi(value));
        }
        lineCount++;
    }

    file.close();
    return data;
}

float calculateCorrelationCoefficient(const vector<int> &a, const vector<int> &b)
{
    int n = a.size();
    int a_plus = 0, b_plus = 0, ab_plus = 0;

    for (int i = 0; i < n; i++)
    {
        a_plus += a[i] == 1 ? 1 : 0;
        b_plus += b[i] == 1 ? 1 : 0;
```

```
        ab_plus += (a[i] == 1 && b[i] == 1) ? 1 : 0;
    }

    if (a_plus == 0 || b_plus == 0)
    {
        return 0.0; // To handle cases where division by zero may occur
    }

    return static_cast<float>(ab_plus) / (a_plus * b_plus);
}

void writeCorrelationCoefficient(const string &filename, float corr_coeff)
{
    ofstream file(filename, ios::out);

    if (!file.is_open())
    {
        cerr << "Error in opening output file: " << filename << endl;
        exit(1);
    }

    file << "Pearson Correlation Coefficient"
         << "," << corr_coeff << endl;
    file.close();
}
```

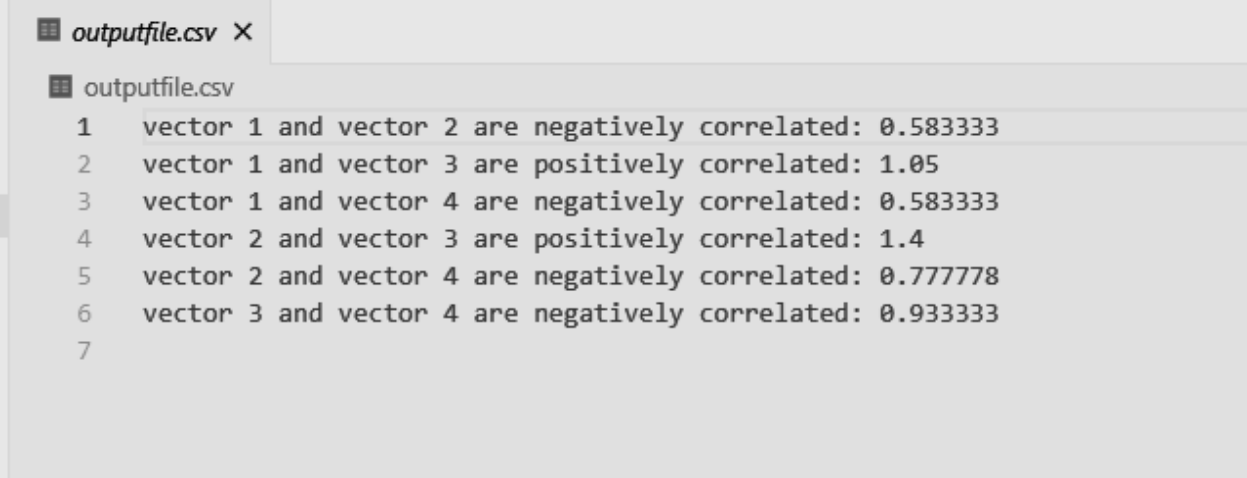
```
int main()
{
    string inputFileName = "correlation_input.csv";
    string outputFileName = "correlation_output.csv";

    vector<int> a = readData(inputFileName);
    vector<int> b = readData(inputFileName);

    float corr_coeff = calculateCorrelationCoefficient(a, b);

    writeCorrelationCoefficient(outputFileName, corr_coeff);

    cout << "Correlation coefficient calculated and saved in '" << outputFileName << "'." << endl;

    return 0;
}
```

Output:

Input.csv:

Tid	M	T	W	Th	F	S
1	1	1	0	0	1	0
2	0	1	1	0	0	1
3	1	1	1	0	1	1
4	0	0	0	1	1	1

Output.csv:


outputfile.csv
outputfile.csv
1 vector 1 and vector 2 are negatively correlated: 0.583333
2 vector 1 and vector 3 are positively correlated: 1.05
3 vector 1 and vector 4 are negatively correlated: 0.583333
4 vector 2 and vector 3 are positively correlated: 1.4
5 vector 2 and vector 4 are negatively correlated: 0.777778
6 vector 3 and vector 4 are negatively correlated: 0.933333
7