CS102: Computer Programming II
Term 212
Course Programming Project

DNA signature

Sarah Aljurbua

220410528

Problem Statement:

➕ Write a Java application that will prompt the user to enter:

    1. The name of a text file containing a very long DNA signature.

    2. A String containing the DNA sequence to search for in the DNA file.

➕ Then, the program should save the following results on a file called **dna_stats.txt**:

    1. The DNA sequence and the number of times it is repeated in the DNA file.

    2. The start index of each sequence in the DNA file.

    Finally, the program should print on the screen:

    1. The DNA sequence and the number of times it is repeated in the DNA file.

    2. The execution time in seconds, try to make your code efficient

Sample Execution:

```
run:
Please enter the name of the DNA file: dna_sequence.txt
Please enter the DNA sequence you want to search for: aggc
DNA sequence search completed. Results saved in dna_stats.txt file!
Sequence : aggc found 12350 times in dna_sequence.txt file!
Execution time = 1250 seconds!

run:
Please enter the name of the DNA file: dnq_seq.txt
Error! File dnq_seq.txt not found!
Please enter the name of the DNA file: dna_sequence.txt
Please enter the DNA sequence you want to search for: aggc
DNA sequence search completed. Results saved in dna_stats.txt file!
Sequence : aggc found 12350 times in dna_sequence.txt file!
Execution time = 1250 seconds!
BUILD SUCCESSFUL (total time: 14 seconds)
```

```
run:
Please enter the name of the DNA file: dna_sequence.txt
Please enter the DNA sequence you want to search for: XaxX
DNA sequence search completed. Results saved in dna_stats.txt file!
Sequence : XaxX not found in dna_sequence.txt file!
Execution time = 1265 seconds!
BUILD SUCCESSFUL (total time: 13 seconds)
```
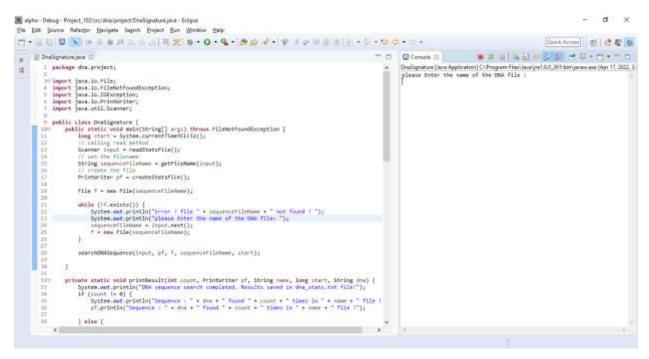
Solution:

1. *Code file:*

```java
package dna.project;

import java.io.File;
import java.io.FileNotFoundException;
import java.io.IOException;
import java.io.PrintWriter;
import java.util.Scanner;

public class DnaSignature {
    public static void main(String[] args) throws FileNotFoundException {
        long start = System.currentTimeMillis();
        // calling read method
        Scanner input = readStatsFile();
        // set the filename
        String sequenceFileName = getFileName(input);
        // create the file
        PrintWriter pf = createStatsFile();

        File f = new File(sequenceFileName);

        while (!f.exists()) {
            System.out.println("Error ! file " + sequenceFileName + " not
found ! ");

            System.out.println("please Enter the name of the DNA file : ");
            sequenceFileName = input.next();
            f = new File(sequenceFileName);
        }

        searchDNASequence(input, pf, f, sequenceFileName, start);

    }

    private static void printResult(int count, PrintWriter pf, String name, long
start, String dna) {
```

2

```java
                System.out.println("DNA sequence search completed. Results saved in
dna_stats.txt file!");
                if (count != 0) {
                        System.out.println("Sequence : " + dna + " found " + count + "
times in " + name + " file !");
                        pf.println("Sequence : " + dna + " found " + count + " times in "
+ name + " file !");

                } else {
                        System.out.println("Sequence : " + dna + "  not found " + " in "
+ name + " file !");
                        pf.println("Sequence : " + dna + "  not found " + " in " + name +
" file !");
                }

                pf.close();

                long end = System.currentTimeMillis();
                System.out.println("Execution time = " + ((end - start) / 1000) + "
seconds!");

        }

        private static void searchDNASequence(Scanner input, PrintWriter pf, File f,
String name, long start) {
                int count = 0;
                String dna = "";
                try {
                        Scanner file = new Scanner(f);
                        System.out.print("please enter the DNA sequence you want to
search for : ");
                        dna = input.next();
                        pf.println("DNA sequence " + dna);
                        String line;
                        int LineNumber = 0;

                        while (file.hasNextLine()) {
                                line = file.nextLine();
                                LineNumber++;

                                int loc = line.indexOf(dna);
                                if (loc != -1)
                                        pf.println("in line " + LineNumber + " found first
dna " + dna + " in index : " + loc);

                                while (loc != -1) {
                                        count++;

                                        line = line.substring(loc + dna.length());
                                        loc = line.indexOf(dna);
                                }

                        }
                        file.close();
```
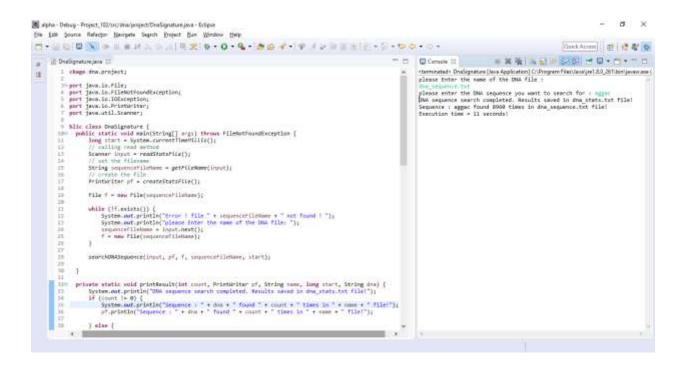
3

```java
                } catch (IOException e) {
                        System.out.println("Error in file");
                }

                printResult(count, pf, name, start, dna);

        }

        private static PrintWriter createStatsFile() throws FileNotFoundException {
                PrintWriter writer = new PrintWriter(new File("dna_stats.txt"));
                return writer;
        }

        private static String getFileName(Scanner input) {
                System.out.println("please Enter the name of the DNA file : ");
                String name = input.next();
                return name;
        }

        private static Scanner readStatsFile() throws FileNotFoundException {
                Scanner input = new Scanner(System.in);
                return input;
        }

}
```
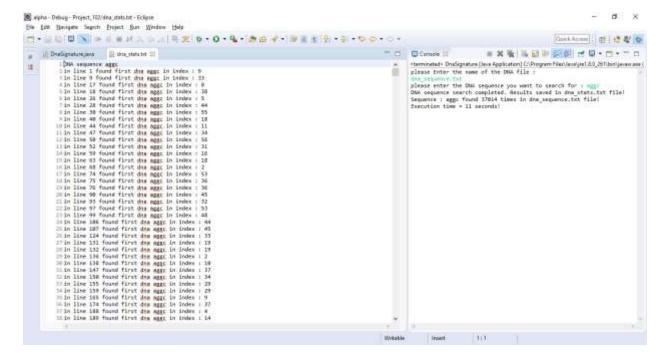
2. Screen shots:

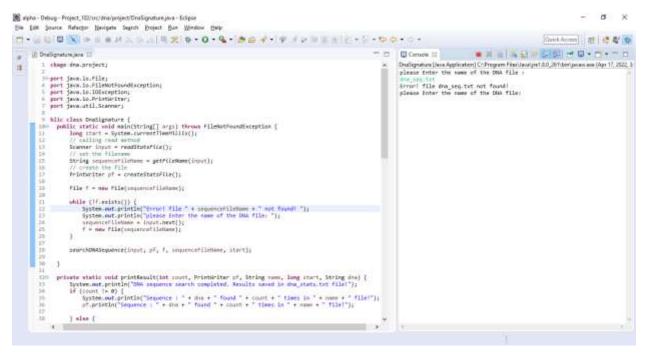```
please Enter the name of the DNA file :
dna_sequence.txt
```
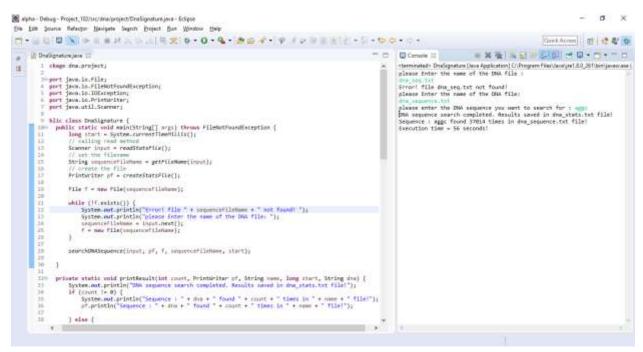
please Enter the name of the DNA file :
dna_sequence.txt
please enter the DNA sequence you want to search for : aggac
DNA sequence search completed. Results saved in dna_stats.txt file!
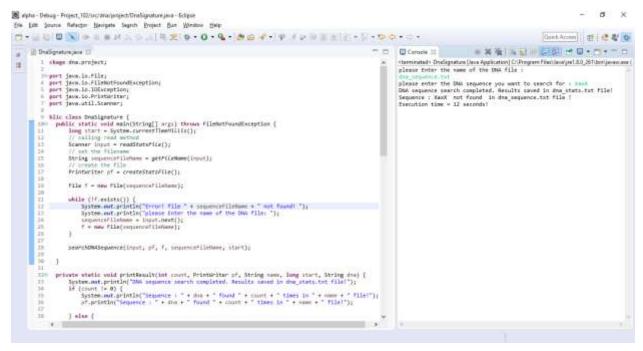Sequence : aggac found 8960 times in dna_sequence.txt file!
Execution time = 11 seconds!

please Enter the name of the DNA file :
dna_seq.txt
Error! file dna_seq.txt not found!
please Enter the name of the DNA file:



please Enter the name of the DNA file :
dna_seq.txt
Error! file dna_seq.txt not found!
please Enter the name of the DNA file:
dna_sequence.txt

please enter the DNA sequence you want to search for : aggc
DNA sequence search completed. Results saved in dna_stats.txt file!
Sequence : aggc found 37014 times in dna_sequence.txt file!
Execution time = 56 seconds!



please Enter the name of the DNA file :
dna_sequence.txt
please enter the DNA sequence you want to search for : XaxX
DNA sequence search completed. Results saved in dna_stats.txt file!
Sequence : XaxX  not found  in dna_sequence.txt file !
Execution time = 12 seconds!