1a)

"We want you to estimate the number of different words there are of 4, 5, 6, and so on up to 20 letters long."

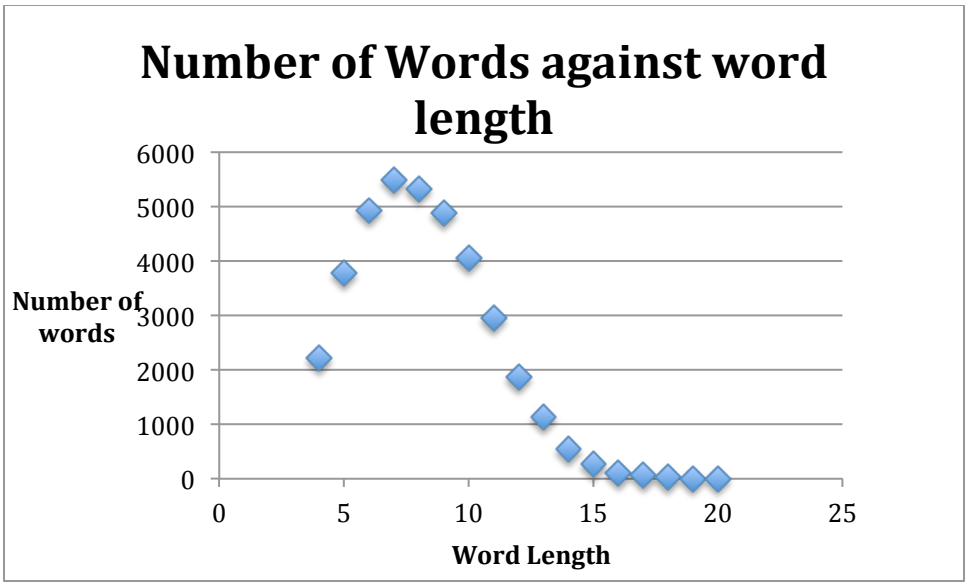
The code initially given in HangmanStats would simply iterate through the words in the text file lowerwords.txt, and print out the number of words of the length that you called using the getRandom method. In order to get an estimate of the number of words of length 4-20 inclusive, I modified the code I was given so that I would not have to call the method for a number of each given length individually, as this would be inefficient. To address this, I created a nested **for** loop us. The for loop would call the getRandomWord 10000 for words of a given length. I repeated this 10 times, and the highest value found was taken as the estimate of the number of words of a given length. I then printed the results, cleared the set, and then did same for other lengths from 4 – 20 inclusive, all with the use of nested for loops.

My code is shown below:

```
import java.util.HashSet;

// Author- Tanaka Jimha


public class HangmanStats {
    public static void main(String[] args) {
        HangmanFileLoader loader = new HangmanFileLoader();
        loader.readFile("lowerwords.txt");

        HashSet<String> set = new HashSet<String>();
        for(int i = 4; i < 21; i++ ) {
            int best = 0;
            for(int j = 0; j < 100; j++) {
                for(int k=0; k < 10000; k += 1) {
                    set.add(loader.getRandomWord(i));
                }
                best = Math.max(set.size(), best);
                set.clear();
            }
            System.out.printf("number of %d letter words = %d\n", i,  best);
        }


    }
}
```

The results I obtained are shown in the table below:

‘Word Length	Number of Words
4	2213
5	3783
6	4932
7	5485
8	5328
9	4880
10	4057
11	2959
12	1871
13	1137
14	545
15	278
16	103
17	57
18	23
19	3
20	3



From the bell shape of the graph, it can be seen that the highest number of words occurs for those length 7. There is a huge drop in the number of words as you increase from around 9 letter words, and the number of words also decreases as the word length becomes very small.

1b)

For this part, I decided to answer the question: *"What is an estimate of the average number of vowels contained in a word of length i, for* $i \in [4;20]$ *? ".*

To answer this question, I used an approach very similar to the one I used for part *a*.

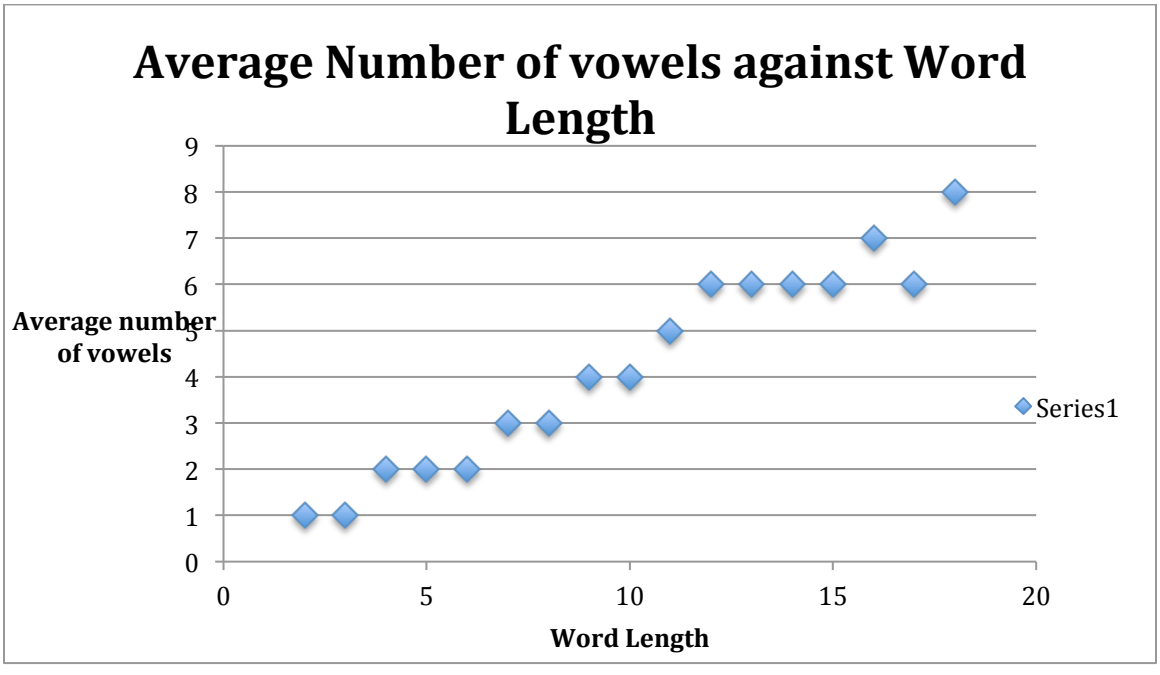
I used a nested ***for*** loop to call the getRandomWord function 10000 times for a word of length i , for $i \in [4;20]$ and stored these words in a set called **set**. Having gotten a set of all the words of that length, I then used another nested for loop to iterate through every letter of every word contained in the set, taking count of the number of vowels in the set. After getting the total number of vowels, I then divided the number by the size of the set to get the average number of vowels in a word of a given length.

My code is shown below:

```
import java.util.Arrays;

// Author- Tanaka Jimha

public class AverageNumberOfVowels {
    public static void main(String[] args) {
        HangmanFileLoader loader = new HangmanFileLoader();
        loader.readFile("lowerwords.txt");

        Character[] vowels = {'a', 'e', 'i', 'o', 'u'};

        HashSet<String> set = new HashSet<String>();

        for(int i = 4; i < 21; i++ ) {

            //populating the set with words of length i
            for(int k=0; k < 10000; k += 1) {
                set.add(loader.getRandomWord(i));
            }
            // iterating through every LETTER in the set and checking if it is a vowel
            int count = 0;

            for(String l : set) {

                for(int m = 0; m < l.length(); m++) {
                    if (Arrays.asList(vowels).contains(l.charAt(m))) {
                        count += 1;
                    }
                }
            }

            int average = Math.round(count /set.size());

            System.out.printf("Average number of vowels in %d letter words = %d\n", i, average);

            set.clear();
        }
    }
}
```

My results are shown below:

Word Length	Average Number of vowels
4	1
5	1
6	2
7	2
8	2
9	3
10	3
11	4
12	4
13	5
14	6
15	6
16	6
17	6
18	7
19	6
20	8



From the graph, the average number of vowels in words increases as the length of the word increases. Although the relationship is not perfectly linear, there is a general increase in the number of vowels with increasing word length.

Other questions worth studying and which can be explored I future studies include:

1) On average, how many times does a given letter appear in a word of a given length.
2) How many different words of a given length contain exactly the same letters but at different positions in the word.
3) What is the average number of consonants in a word of a given length.