Letter Frequency Analysis

Description.

This application counts the frequency of each letter on a message. This program is useful for cryptanalysis.

Files enclosed for demonstration.

The folder containing this program in the repository, has a text file called message1.txt. This text file can be used to demonstrate the functionality of the program

User's guide.

In order to execute the program is necessary to compile and run it through command line. For executing it, it is necessary to input an argument through command line. This argument represents the name of the text file whose message will be analyzed by the letter frequency analysis.

Input example:

```
N BMJJQJI RD HFWY YTBFWI HZXYTRX FSI XFB RDXJQK UJ
WKJHYQD UWTKNQJI YMWTZLM YMJ HTRUZYJW'X JDJX. XNQN
HTS RNSN-RNQQNTSFNWJ, XYTHP TUYNTSX WJYNWJJ, QZHPD
NSMJWNYTW TK YMJ WJATQZYNTS UWTHQFNRJI TS YJQJANX
NTS GD YMJ BTRFS NS WZSSNSL XMTWYX.
```

Output example:

```
Input Message:

GQYOT NCQEO FYFES DRRME LEINQ RKRLR MTHMB HGTHN MNVRQ TFRKE
STTWN FUMAQ RAYRE QSART RGTHV RBHGT HNMFE SDRRM EOEQT HGUKE
QKYBR QTHKR CQNUM ABNQT FHSOK NTRKR LRMTS TEQTH MCWHT FRACE
QEKKE MONRS TFRCN KADUC

Table of Frequencies:

A 6
B 4
C 6
D 3
E 14
```

F 8
G 5
H 11
I 1
J 0
K 10
L 3
M 11
N 11
O 5
P 0
Q 13
R 22
S 7
T 18
U 4
V 2
W 2
X 0
Y 4
Z 0

Histogram of Frequencies

A XXXXXX
B XXXX
C XXXXXX
D XXX
E XXXXXXXXXXXXX
F XXXXXXXX
G XXXXX
H XXXXXXXXXX
I X
J
K XXXXXXXXX
L XXX
M XXXXXXXXXX
N XXXXXXXXXX
O XXXXX
P
Q XXXXXXXXXXXX
R XXXXXXXXXXXXXXXXXXXX
S XXXXXXX
T XXXXXXXXXXXXXXXXX

```
U XXXX
V XX
W XX
X
Y XXXX
Z
```

Program's Design.

Data Structures.

The program uses a HashMap called frecuency that takes characters as keys in order to relate them to integers. This data structure is used to store the frequencies of the characters in the message.

Methods.

analyzeToken(String).
This method takes a string and traverses through it in order to record the occurrences of its Alphabetic characters. These occurrences are stored in the HashMap frecuency.

printFrecuency().
This method prints the one-to-one relations of the HashMap in order to show the letter frequencies of the characters.

histogram().
This method prints a histogram that represents the frequencies of a character using X's as each occurrence. This method uses the one-to-one relations from the HashMap.

main().
The main method initializes the HashMap and reads in the input. It parses every line of the input and uses analyzeToken() to study the frequencies of the characters of each line. Once this happens, then it uses printFrecuency() and histogram() to print the frequencies of the characters and a histogram that represent them. This method also manages the possibility of getting an argument with an incorrect or non-existing file name by giving a message explaining the error and ending the program's execution.