

Statistische Auswertung von Texten

Beim Knacken einer Caesar-Chiffre kann es hilfreich sein, das häufigste Geheimtextzeichen zu bestimmen. Im Folgenden soll daher ein Programm entstehen, das eine entsprechende Analyse des Textes unterstützt.

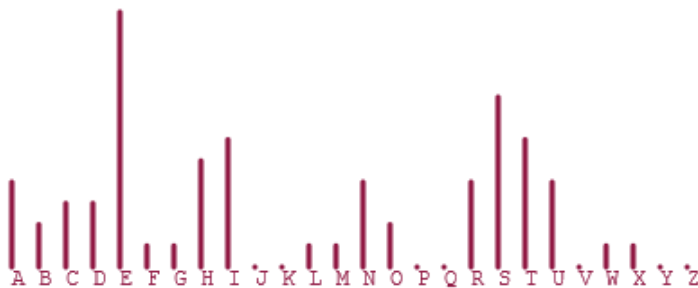
Aufgabe 1: Implementieren Sie ein Programm, das für einen Text und ein Zeichen, das der Anwender eingibt, zählt, wie oft das Zeichen im Text vorkommt.

Aufgabe 2: Um die Analyse eines Textes weiter zu automatisieren, ist es hilfreich für das Zählen eines beliebigen Zeichens in einem Text eine eigene Operation *zeichenZaehlen* zu erstellen.

- Implementieren Sie eine entsprechende eigene Operation *zeichenZaehlen*.
- Testen Sie die Operation *zeichenZaehlen*, indem Sie Ihr Programm aus Aufgabe 1 so verändern, dass die Operation *zeichenZaehlen* verwendet wird.

Aufgabe 3: Ergänzen Sie in Ihrem Programm die Möglichkeit, dass der Anwender nur einen Text eingibt und für jedes Zeichen des Alphabets die entsprechende Häufigkeit ausgegeben wird. Für die Darstellung der Ausgabe haben Sie verschiedene Möglichkeiten:

- Ausgabe der absoluten oder der relativen Häufigkeit
- Ausgabe als Text: Auflistung der Zeichen und der entsprechenden Häufigkeit
- Ausgabe als Histogramm, wie in der Abbildung:



Hinweis: Für die Ausgabe der relativen Häufigkeit ist es sinnvoll zu überlegen, ob als Bezugsgröße die Gesamtanzahl der Zeichen des Textes inklusive oder exklusive Leer- und/oder Satzzeichen verwendet werden soll und wie diese bestimmt wird.

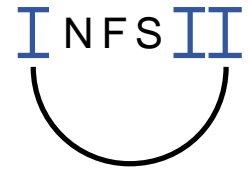
Aufgabe 4: Mit der Ausgabe aus Aufgabe 3 sollte es dem Anwender bereits möglich sein, das häufigste Zeichen zu bestimmen. Auch dieses Problem lässt sich aber algorithmisch lösen.

Ergänzen Sie in Ihrem Programm die automatisierte Ausgabe des häufigsten Zeichens eines Textes.

Aufgabe 5: Im Zusammenhang mit Texten können noch andere Fragestellungen interessant sein. Vielleicht mussten Sie schon einmal einen Text mit mindestens 100 Wörtern oder höchstens 500 Wörtern schreiben.

Implementieren Sie weitere statistische Auswertungen für den Eingabetext. Dabei können Sie eigene Ideen umsetzen oder aus den folgenden Vorschlägen wählen:

- Ausgabe der Anzahl der Wörter eines Textes
- Ausgabe der Häufigkeit einer bestimmten Buchstabenkombination, der der Anwender eingibt, z.B. ch, ei oder sch



Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung - Nicht-kommerziell - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](https://creativecommons.org/licenses/by-nc-sa/4.0/). Sie erlaubt Bearbeitungen und Weiterverteilung des Werks unter Nennung meines Namens und unter gleichen Bedingungen, jedoch keinerlei kommerzielle Nutzung.