

#163 - Wortlängenanalyse

Bei diesem Beispiel geht es darum, die Buchstabenlängen der Worte in beliebigen deutschen Texten zu vergleichen.

Aufgabe 1

Erstelle mit **BeginInvoke/EndInvoke**-Threads, mit denen die Texte parallel eingelesen werden und von jedem Wort die Anzahl der Zeichen zählt. Zähle mit, wie oft welche Wortlänge vorkommt. Relevant für die Aufgabe sind Worte mit einer Länge von 2 bis 23 Buchstaben.

Beachte dabei:

Sonderzeichen wie Punkt, Beistrich und Strichpunkt zählen natürlich nicht als Zeichen von Wörtern und dürfen nicht mitgezählt werden.

Aufgabe 2

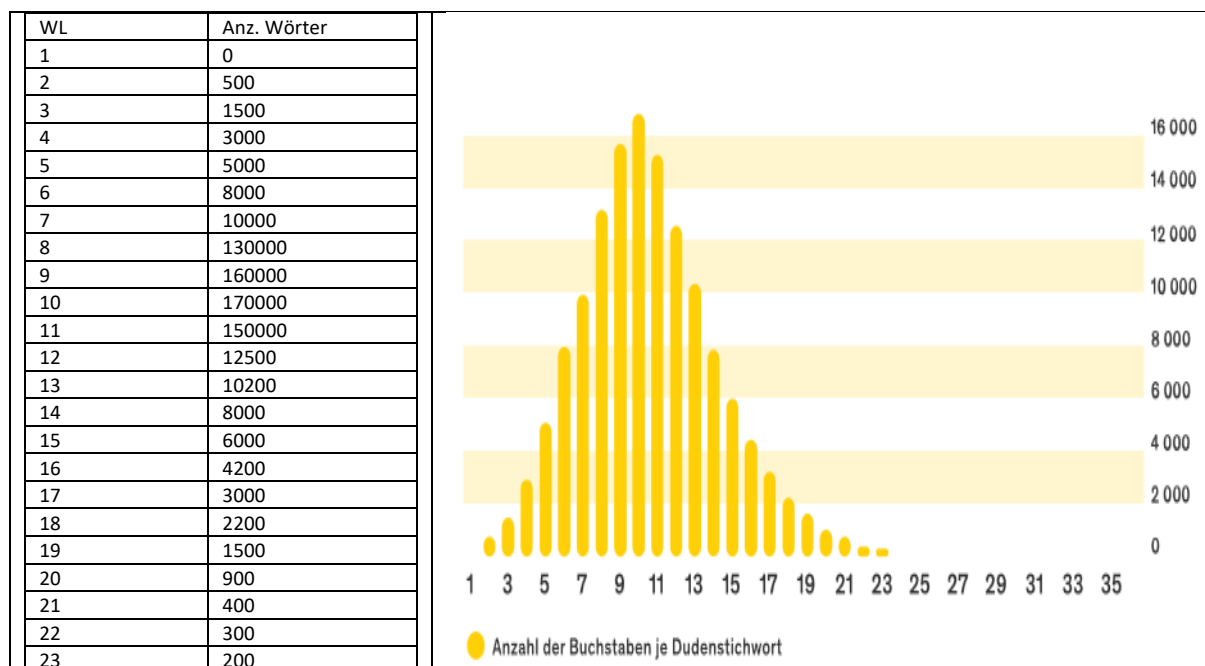
Führe die Ergebnisse der einzelnen Textdateien zusammen. Das bedeutet, jetzt soll ermittelt werden, wie oft ein Wort mit bestimmter Länge in allen 4 Texten vorkommt.

Gib das Ergebnis wie folgt auf der Konsole aus:

Text <1-4>: <Anzahl Worte mit 2 Buchstaben> <Anzahl Worte mit 3 Buchstaben> usw.

Aufgabe 3

Es gibt eine Statistik aus dem Rechtschreibduden, in welcher die Buchstabenlänge aller darin enthaltener Worte erfasst wurden.



<https://www.duden.de/sprachwissen/sprachratgeber/Durchschnittliche-Lange-eines-deutschen-Wortes>

Erstelle eine CSV-Datei, in welcher diese Werte, dem Ergebnis deiner Werte gegenübergestellt werden. Dabei sollen allerdings nicht die absoluten Zahlen herangezogen werden, sondern die prozentuelle Verteilung.

Erstelle anhand dieser CSV-Datei zwei Diagramme mittels Excel (muss nicht programmiert werden!) um das Ergebnis augenscheinlich zu beurteilen. Informiere Dich dabei, wie man in Excel csv-Dateien importiert.

Deliverables:

Results.docx: Screenshots der Konsole nach der Programmausführung und des Excel-Diagramms

Results.xlsx mit dem importierten CSV-Daten und dem Excel-Diagramm