

SEW I - Übung: Wort Statistik

Ziel

Implementierung eines simplen Programms zur Analyse eines Satzes und der darin enthaltenen Wörter.

Lernziele

- Arbeiten mit Methoden.
- Arbeiten mit *Strings*, *chars* und der *ASCII*-Tabelle.
- Arbeiten mit *Arrays*.

Abgaberrichtlinien

- Ihre implementierte Lösung als `.java`-Datei. **Vergessen Sie dabei nicht auf Kommentare und Kommentarkopf!**
- Achten Sie auf **saubere Variablenbenennung** und **Nutzung von Konstanten** wenn sinnvoll!

Aufgabe

Zu implementieren ist ein Programm `WordStatistics`, in dem ein Satz eingelesen, anhand eines konstanten Trennzeichens (in unserem Fall ein Leerzeichen) unterteilt und verschiedene Informationen über den Satz und die enthaltenen Wörter berechnet. Dabei sollen folgende Methoden implementiert werden:

- `printlnLine` gibt eine einfache Trennlinie in der Konsole aus.
- `countSymbolInText` berechnet wie oft ein übergebenes Zeichen in einer übergebenen Zeichenkette vorkommt und gibt das Ergebnis als Ganzzahl zurück.
- `countCharactersWithoutSplitSymbols` berechnet die Anzahl von Zeichen in einer übergebenen Zeichenkette, die **nicht** das übergebene Trennzeichen sind, und gibt das Ergebnis als Ganzzahl zurück.
- `splitIntoWords` trennt eine übergebene Zeichenkette anhand eines übergebenen Zeichens in Wörter auf und gibt diese als *Array* von Zeichenketten zurück.
- `printWords` akzeptiert ein *Array* von Zeichenketten und gibt dieses in *Mengenschreibweise* (i.e. `{blau, weiss}`) aus.
- `countUniqueWords` akzeptiert ein *Array* von Zeichenketten und berechnet die Anzahl von Unikaten. **Groß- und Kleinschreibung soll dabei ignoriert werden.**
- `markDuplicates` ersetzt die in einem übergebenen *Array* vorkommenden Duplikate durch einen konstanten Wert (bspw. `DUPLICATE!`).
- `getShortestWord` und `getLongestWord` finden das kürzeste bzw. das längste Wort in einem übergebenen *Array* von Zeichenketten und geben es zurück.

- `isNumber` überprüft ob eine übergebene Zeichenkette eine Zahl ist. Eine Zahl darf nur aus Ziffern bestehen, wobei das erste Zeichen ein *Minus* sein kann.
- `isNegativeNumber` überprüft ob eine übergebene Zeichenkette eine negative Zahl ist. Eine Zahl darf nur aus Ziffern bestehen, wobei das erste Zeichen ein *Minus* sein muss.
- `isHexColor` überprüft ob eine übergebene Zeichenkette der hexadezimalen Farbdefinition folgt. Sie muss also mit einer Raute beginnen, auf die 3 oder 6 Ziffern oder Buchstaben zwischen *A* und *F* (Groß- und Kleinschreibung egal!) folgen.
- `isPalindrome` überprüft ob eine übergebene Zeichenkette ein *Palindrom* - also von vorne und hinten gelesen identisch - ist.
- `isSoccerScore` überprüft ob eine übergebene Zeichenkette ein gültiges Fußball-Ergebnis sein könnte. Ein gültiges Fußball-Ergebnis besteht immer aus Zahlen die von exakt einem Doppelpunkt (der nicht am Anfang oder Ende stehen darf!) getrennt sind.

Beispiele

Beispiel 1

Please enter a sentence: Hannah hat 50 Euro Schulden bei Otto - Der hat -100 euro am Konto

There is a total of 52 characters in 14 words.

The words in the sentence are: {Hannah, hat, 50, Euro, Schulden, bei, Otto, -, Der, hat, -100, euro, am, Konto}.

The shortest word is: -.

The longest word is: Schulden.

Number of unique words: 12.

The words with marked duplicates are: {Hannah, hat, 50, Euro, Schulden, bei, Otto, -, Der, DUPLICATE!, -100, DUPLICATE!, am, Konto}.

Word 1: Hannah.

Is a number: false.

Is a negative number: false.

Is a hex color: false.

Is a palindrome: true.

Is a soccer score: false.

Word 2: hat.

Is a number: false.

Is a negative number: false.

Is a hex color: false.

Is a palindrome: false.

Is a soccer score: false.

```

-----
Word 3: 50.
Is a number: true.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 4: Euro.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 5: Schulden.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 6: bei.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 7: Otto.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: true.
Is a soccer score: false.
-----

Word 8: -.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: true.
Is a soccer score: false.
-----

Word 9: Der.
Is a number: false.
Is a negative number: false.

```

```

Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----
Word 10 is a duplicate.
-----
Word 11: -100.
Is a number: true.
Is a negative number: true.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----
Word 12 is a duplicate.
-----
Word 13: am.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----
Word 14: Konto.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

```

Beispiel 2

```

Please enter a sentence: Polen Slowakei 1:2 polen Spanien 1:1 #fFf #ff0000 #abcd
-----
There is a total of 47 characters in 9 words.
The words in the sentence are: {Polen, Slowakei, 1:2, polen, Spanien,
1:1, #fFf, #ff0000, #abcd}.
The shortest word is: 1:2.
The longest word is: Slowakei.
Number of unique words: 8.
The words with marked duplicates are: {Polen, Slowakei, 1:2, DUPLICATE!,
Spanien, 1:1, #fFf, #ff0000, #abcd}.
-----
Word 1: Polen.
Is a number: false.
Is a negative number: false.

```

```

Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 2: Slowakei.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 3: 1:2.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: true.
-----

Word 4 is a duplicate.
-----

Word 5: Spanien.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 6: 1:1.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: true.
Is a soccer score: true.
-----

Word 7: #fFf.
Is a number: false.
Is a negative number: false.
Is a hex color: true.
Is a palindrome: false.
Is a soccer score: false.
-----

Word 8: #ff0000.
Is a number: false.
Is a negative number: false.
Is a hex color: true.
Is a palindrome: false.

```

```
Is a soccer score: false.
-----
Word 9: #abcd.
Is a number: false.
Is a negative number: false.
Is a hex color: false.
Is a palindrome: false.
Is a soccer score: false.
-----
```