

Aufgabenstellung zur Coding Challenge

{Coding Girls gesucht!}

Du bist Studentin oder Auszubildende in einem IT-nahen Studiengang / Ausbildungsprogramm und möchtest Dich einer spannenden Coding-Herausforderung stellen? Dann bist Du herzlich willkommen! Hier findest Du alle Informationen zur Aufgabenstellung unseres Wettbewerbs.

{Worum geht es genau?}

Wir stellen Dir einen Datensatz (CSV) zur Verfügung, der Produktbewertungen aus unserem Online-Shop enthält. Diese CSV-Datei kannst Du unter der folgenden URL herunterladen: <https://www.bonprix.de/corporate/karriere/coding-challenge/>

In der CSV-Datei findest Du die Bewertungen für 10 unterschiedliche bonprix Produkte.

- ✓ In **Spalte A** ist die Artikelnummer ausgewiesen. So kannst Du nachvollziehen, welche Bewertungen zusammengehören.
- ✓ In **Spalte B** findest Du die Kundenbewertungen in Textform.
- ✓ In **Spalte C** ist das Rating abgebildet, das von 1 bis 5 reicht. Dieses Rating entspricht der Vergabe von Bewertungssternchen, die der bonprix Kunde jedem gekauften Produkt zuordnen kann. Die Skala reicht von 1 (schlechteste Bewertung) bis 5 (beste Bewertung).

Deine Aufgabe besteht darin, eine Tag-Cloud zu erstellen, die ein Filtern der Rezensionen nach bestimmten Bewertungskategorien (Tags) ermöglicht.

Eine Tag-Cloud (auch „Wort-Wolke“ genannt) bezeichnet eine Methodik zum Aggregieren und Visualisieren von Informationen, die – je nach vorliegender Gewichtung – besonders hervorgehoben werden.

Die Tag-Cloud sollte für jedes Produkt die 10 häufigsten Bewertungskategorien (Tags) abbilden. Für jeden Tag sollten zudem die passenden Rezensionen dargestellt werden. Ob es am Ende ein Programm, ein Service oder eine REST-API wird, ist Dir überlassen. Wir stellen nur eine einzige Bedingung an Dich:

Verwende bitte **Java** oder **Python** für Deine Entwicklung. Beim Web-Frontend hast Du wiederum alle Freiheiten, die Du brauchst.

Aufgabenstellung zur Coding Challenge

Unser Wettbewerb besteht aus einem obligatorischen und einem kreativen Part. Für das „Pflichtprogramm“ kannst Du Dich gern an dem folgenden Vorgehen orientieren:

- 1) **Sentencesplitting & Tokenization:** Zerlege die Produktbewertungen in Satz- und Worteinheiten als Grundlage für alle folgenden Arbeitsschritte.

Beispiel

Der Satz „*The quick brown fox jumps over the lazy dog*“ wird von einem Tokenizer zurückgegeben als:

```
<sentence>
  <word>The</word>
  <word>quick</word>
  <word>brown</word>
  <word>fox</word>
  <word>jumps</word>
  <word>over</word>
  <word>the</word>
  <word>lazy</word>
  <word>dog</word>
</sentence>
```

- 2) **Part-of-speech-Tagging:** Identifiziere die Wortarten innerhalb der Produktbewertungen (z.B. Substantiv, Adjektiv, Verb, Pronomen, Präposition).

Beispiel

Für den Satz „*The quick brown fox jumps over the lazy dog*“ sollte das Part-of-speech-Tagging das folgende Resultat ergeben:

The	Artikel
quick	Adjektiv
brown	Adjektiv
fox	Substantiv
jumps	Verb
over	Präposition
the	Artikel
lazy	Adjektiv
dog	Substantiv

Aufgabenstellung zur Coding Challenge

- 3) **Stemming:** Reduziere verschiedene Varianten eines Wortes auf den gemeinsamen Wortstamm (Stem), um unterschiedliche Variationen desselben Wortes gleichwertig zu behandeln.
- 4) **Tag-Cloud:** Zähle die Stems pro Produkt und zeige für die 10 häufigsten Stems jedes Produktes eine Tag-Cloud an. Für jeden Tag sollten zudem die entsprechenden Rezensionen ausgespielt werden.

{Ein kleiner Tipp von uns}

Es gibt bereits gut dokumentierte Bibliotheken, auf die Du zurückgreifen kannst. Du musst das Rad also nicht neu erfinden, sondern kannst Dich an dieser kleinen Auswahl an Bibliotheken für Java und Python orientieren:

- ✓ DKPro Core
- ✓ ClearTK
- ✓ OpenNLP
- ✓ NLTK

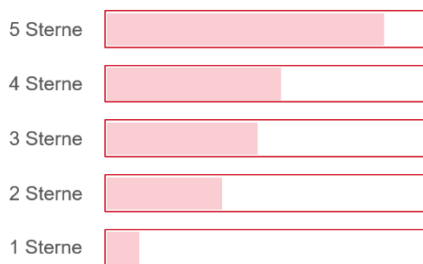
{Nach der Pflicht, kommt die Kür}

Im kreativen Part unseres Wettbewerbs sind Deinen Coding-Skills keinerlei Grenzen gesetzt. Du kannst Dir weitere Visualisierungsmöglichkeiten für Deine Tag-Cloud überlegen oder Deinen Service mit Zusatzfeatures ausstatten. Hier kannst Du groß denken und visionäre Features entwickeln. Erlaubt ist alles, was Dein Developer-Herz höher schlagen lässt. Ein Blick über den Tellerrand lohnt sich hier auf alle Fälle. Schau Dir also gerne an, wie andere Online-Shops das Thema „Automatisierte Produktbewertungen“ umsetzen.

Aufgabenstellung zur Coding Challenge

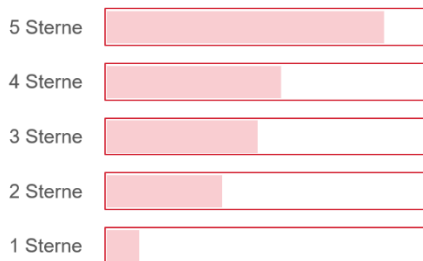
Hier findest Du ein Beispiel für eine Tag-Cloud. Das ist allerdings nur eine kleine Orientierungshilfe. Welchen Lösungsweg Du wählst, bleibt Dir und Deiner Kreativität überlassen.

Kundenbewertungen 1



Material	Passform	Muster
Farbe	Größe	Pflege
Preis-Leistung	Länge	Länge

Kundenbewertungen 2



Material	Passform	Muster
Farbe	Größe	Pflege
Preis-Leistung	Länge	Länge

★★★★★ Anja / 19.09.2018

(Gr. 36/38) / Weite: Passt genau, Länge: Passt genau, Körpergröße: 175-179 cm

Material ist etwas schwerer als gedacht, sitzt perfekt und sieht einfach schick aus

1 von 1 Kunden fanden diese Bewertung hilfreich

Hat diese Rezension Dir weitergeholfen?

[Rezension melden](#)

★★★★★ Sabrina / 17.09.2018

(Gr. 36/38) / Weite: Viel zu weit, Länge: Zu lang, Körpergröße: 160-164 cm

Farbe war sehr schön. Material war toll. Kleid war jedoch an der Taille viel zu groß und obenrum ebenfalls leider zu groß. Leider wieder zurück

0 von 0 Kunden fanden diese Bewertung hilfreich

Hat diese Rezension Dir weitergeholfen?

[Rezension melden](#)

Aufgabenstellung zur Coding Challenge

{Wie bewerten wir Deinen Code?}

Wenn wir Deinen Code bewerten, stellen wir uns die folgenden Fragen:

- ✓ **Funktionalität**
Bekommen wir Dein Programm zum Laufen?
- ✓ **Set-up**
Schickst Du uns eine detaillierte Doku (README) mit?
- ✓ **Qualität**
Weist Dein Code eine sinnvolle Struktur auf und ist in sich schlüssig?
Bitte keinen Spaghetti-Code!
- ✓ **Usability**
Wie nutzerfreundlich ist Deine Anwendung?
Gibt es Barrieren in der Bedienbarkeit?
- ✓ **Zusätzliche Features**
Wir lieben Kreativität!
Deshalb erhältst Du für geniale Zusatzfeatures Bonuspunkte!

{Unser Versprechen an Dich!}

Dein Code und Deine Ideen gehören Dir! Für immer und ewig! Du hast unser Wort. Uns geht es in erster Linie um Nachwuchsförderung und Empowerment, nicht darum geniale Ideen von angehenden IT-Talenten zu kopieren. Copycat ist einfach uncool! Deinen Quellcode brauchen wir lediglich, um die Qualität Deiner eingereichten Lösung bewerten zu können.

{Du hast eine geniale Lösung! Was passiert als nächstes?}

Dein Ergebnis schickst Du als Zip-File an wettbewerb@bonprix.net. Vergiss nicht eine aktuelle Immatrikulations- bzw. Ausbildungsbescheinigung Deiner Mail anzuhängen.

{Hast Du noch Fragen?}

Wir freuen uns auf Deine E-Mail an wettbewerb@bonprix.net