

kb21_klausur

July 15, 2021

1 Probeklausur Korpusbearbeitung Sommersemester 21 (15.07.2021)

Anmerkungen: - Die Bearbeitungszeit beträgt 30 Minuten. - Die Probeklausur umfasst 5 Aufgaben.
- Alle Aufgaben geben jeweils maximal 2 Punkte. - Insgesamt können 10 Punkte erreicht werden (5*2=10). - Die Abgabe erfolgt entweder - direkt in diesem Jupyter-Notebook, - in einer separaten Textdatei oder - handschriftlich eingescannt als PDF.

1.1 Aufgabe 1

Die Datei `file.txt` enthalte folgenden Inhalt:

```
Zeile 1  
Zeile 2  
Zeile 3  
Zeile 4  
Zeile 5
```

Erläutern Sie folgenden Kommandozeilenaufruf: `cat file.txt | sed -n 2,4p` (Was ist die Ausgabe des Aufrufs und wie kommt das Ergebnis zustande).

Lösung:

1.2 Aufgabe 2

Gegeben seien drei Dateien `a.txt`, `b.txt`, `c.txt` in einem Verzeichnis `dir`:

```
$ ls dir  
a.txt b.txt c.txt
```

Wie können Sie Gesamtgröße aller drei Dateien im Verzeichnis `dir` berechnen (Kommandozeilenaufruf, Python-Programm)?

Lösung:

1.3 Aufgabe 3

Gegeben sei eine ASCII-kodierte Datei `corpus.txt`, die verschiedene Texte enthält. Geben Sie einen Kommandozeilenaufruf an, mit dem Sie die Anzahl aller Wörter in der Datei bestimmen können, die länger als 8 Zeichen sind.

Lösung:

1.4 Aufgabe 4

Betrachten Sie folgende beide Kommandozeilenaufrufe:

```
$ echo 'äße' | wc -c
6
$ echo 'äße' | wc -m
4
```

Erläutern Sie die beiden unterschiedlichen Ausgaben der beiden Aufrufe.

Lösung:

1.5 Aufgabe 5

In der Datei `tofahrenheit.py` sei folgendes Python-Programm implementiert:

```
import sys

for celsius in sys.argv[1:]:
    fahrenheit = (float(celsius) * 1.8) + 32
    print(f'{celsius}\u2103 is equal to {fahrenheit}\u2109')
```

Beschreiben Sie die Schritte, die man durchführen muss, um dieses Python-Programm direkt auf der Kommandozeile ausführbar zu machen (d.h. dass der Aufruf `$./tofahrenheit.py 31` auf der Kommandozeile funktioniert).