

```

1  # Übungsblatt 5 -- Dateien und Zeichenketten
2  # M. Endejan, V. Engels 2021
3
4  # 1a) Öffnen Sie die bereitgestellte Datei NASA_access_log_01Jul95 und lesen
5  # Sie den Inhalt der gesamten Datei in eine neue Liste 'log'; die einzelnen
6  # Zeilen der Datei sollen jeweils in einem separaten Listenelement
7  # gespeichert werden.
8
9  dateiname = "/Users/fom/daten/NASA_access_log_01Jul95"
10 with open(dateiname) as datei:
11     log = datei.readlines()
12
13 print("type(log):", type(log))
14
15
16 # 1b) Geben Sie die Anzahl der insgesamt gelesenen Zeilen sowie die Inhalte
17 # der ersten zehn Zeilen (Listenelemente) auf dem Bildschirm aus.
18
19 print("Anzahl gelesener Zeilen:", len(log))
20 print("Inhalt der ersten zehn Zeilen:\n", *log[:10])
21
22
23 # 2) Erstellen Sie eine neue Liste log_info mit der folgenden Struktur:
24 # [[wer, wann, was], [wer, wann, was], ...]
25 # wer: IP-Adresse oder Domain-Name
26 # wann: Uhrzeit (ohne Angabe des -- immer gleichen -- Datums)
27 # was: Name der aufgerufenen Datei bzw. des aufgerufenen Verzeichnisses
28 # Bsp.: [['199.72.81.55', '00:00:01', 'history/apollo/'], ...]
29
30 log_info = []
31
32 for zeile in log:
33     log_zeile = zeile.split() # teilt die Zeile in einzelne Elemente auf
34     ip = log_zeile[0] # die ip bzw. der Domain-Name befindet sich in Element 0
35     zeit = log_zeile[3][13:] # Form von Element 3: '[01/Jul/1995:hh:mm:ss'
36     datei = log_zeile[6] # Datei- bzw. Verzeichnisinformation
37     log_info.append([ip, zeit, datei])
38
39
40 # 3) Geben Sie die ersten zehn Einträge von log_info nummeriert in der
41 # folgenden Form auf dem Bildschirm aus:
42 # 1 ['199.72.81.55', '00:00:01', '/history/apollo/']
43 # 2 ['unicomp6.unicomp.net', '00:00:06', '/shuttle/countdown/']
44
45 for nr, eintrag in enumerate(log_info[:10], 1):
46     print(nr, eintrag)
47
48 # Alternative (kein Kopieren der Einträge notwendig, dafür aber zus. Abfrage)
49 for nr, eintrag in enumerate(log_info, 1):
50     print(nr, eintrag)
51     if nr == 10:
52         break
53
54
55 # 4) Die Log-Informationen der ersten zehn Zeilen sollen nun in Form
56 # einer Tabelle (s.u.) ausgegeben werden.
57 # Die Spaltenbreite soll sich am jeweils längsten Element der Spalte
58 # orientieren (so dass keine 'Verschiebungen' auftreten können.
59 # Beispiel:
60 # -----
61 # | wer | wann | was |
62 # |-----|-----|-----|
63 # | 199.72.81.55 | 00:00:01 | /history/apollo/ |
64 # | unicomp6.unicomp.net | 00:00:06 | /shuttle/countdown/ |
65 # | ... | ... | ... |
66 # | 129.94.144.152 | 00:00:13 | / |
67 # -----
68
69 # Ermittlung der maximalen Breiten (also der längsten Elemente pro Spalte)
70 # Lösungsbeispiel
71 max_breite_wer = 0

```

```

72 max_breite_wann = 8 # die Uhrzeit benötigt immer acht Zeichen
73 max_breite_was = 0
74
75 auswahl = log_info[:10] # die ersten zehn Einträge
76
77 for eintrag in auswahl:
78     max_breite_wer = max(max_breite_wer, len(eintrag[0]))
79     max_breite_was = max(max_breite_was, len(eintrag[2]))
80
81 # Lösungsbeispiel 2
82 # Alternative Lösung zur Ermittlung der maximalen Breiten direkt über max
83 max_breite_wer = max([len(e[0]) for e in auswahl])
84 max_breite_was = max([len(e[2]) for e in auswahl])
85
86
87 # Weiterer Hinweis:
88 # Noch effizienter als die Verwendung einer Liste Comprehension wäre die Ver-
89 # wendung eines Generators: max_breite_wer = max((len(e[0]) for e in auswahl))
90
91
92 # Definition des Tabellenformates
93 format_überschrift = \
94     " | {:^{breite_wer}} | {:^{breite_wann}} | {:^{breite_was}} | "
95
96 tabelle = " | {:{breite_wer}} | {:{breite_wann}} | {:{breite_was}} | "
97
98 # Überschrift ausgeben
99 überschrift = format_überschrift.format("wer", "wann", "was",
100     breite_wer=max_breite_wer,
101     breite_wann=max_breite_wann,
102     breite_was=max_breite_was))
103 print('-' * len(überschrift))
104 print(überschrift)
105 print('-' * len(überschrift))
106
107 # Einträge ausgeben
108 for eintrag in auswahl:
109     print(tabelle.format(*eintrag,
110         breite_wer=max_breite_wer,
111         breite_wann=max_breite_wann,
112         breite_was=max_breite_was))
113
114 print('-' * len(überschrift))
115
116
117 # 5) Speichern Sie die Informationen der Liste log_info im folgenden
118 # Format in einer neuen Datei log_info.csv:
119 # wer/wann/was
120 # ip/zeit/dateiname
121 # Bsp.:
122 # wer/wann/was
123 # 199.72.81.55;00:00:01;/history/apollo/
124 # ...
125
126 dateiname = '/Users/fom/daten/log_info.csv'
127
128 with open(dateiname, "w") as datei:
129     datei.write("wer/wann/was\n")
130     for eintrag in log_info:
131         datei.write(";".join(eintrag)+'\n')
132

```