

Python Projekttag 2

1) Schreibe einen Webcrawler, der eine Twitter-handle (z.B. @BTA_Berlin) bekommt und damit die Follower dieses Users herausfinden soll.

(Ergebnis bei @BTA_Berlin: 75 Follower (23.01.2020 19:56))

2) Schreibe ein Script, welches die Daten zur Europawahl aus zwei csv Dateien (2014, 2019) ausliest.

2014:

https://www.statistik-berlin-brandenburg.de/opendata/EU2014_BE_EndgErg_Wahlbezirke.csv

2019:

https://www.statistik-berlin-brandenburg.de/opendata/EU2019_BE_EndgErg_Wahlbezirke.csv

Wichtig: Die csv Dateien haben eine seltsame Codierung. Dies kann man in der read_csv funktion schon umgehen mit dem Parameter: encoding z.B. für diese Dateien:

```
df = pd.read_csv('filename.csv', encoding = "cp1252", delimiter = ';')
```

Die einzelnen DataFrames zu 2014 und 2019 könnten so aussehen:

	Wahlbezirk	CDU	GRÜNE	SPD
0	1A	201	174	225
1	1B	172	102	159
2	1C	139	280	219

	Wahlbezirk	CDU	GRÜNE	SPD
0	100	81.0	219.0	90.0
1	101	87.0	277.0	93.0
2	102	72.0	194.0	76.0

Wahlbezirke, die nicht in beiden CSV Dateien vorkommen können ignoriert werden.

2.1)

Für jeden Wahlbezirk sollen die Stimmen für: CDU, SPD und GRÜNE verglichen werden. Es soll für jeden Wahlbezirk die absolute Veränderung vom Jahr 2014 zum Jahr 2019 berechnet werden. (Differenz)

2.2)

Berechne für jeden Wahlbezirk die Gesamtzahl der Stimmen für die Parteien CDU, SPD und GRÜNE und vergleiche sie für die Jahre 2014 und 2019. (Differenz)

2.3)

Ermittle die gesamten Stimmen für jede der drei Parteien in den jeweiligen Jahren.