Lab: Parsing Json

imagine the following situation:

Your boss wants you to write Python code that performs the following steps:

- extracting an **amino acid** (AA) **sequence** that is located somewhere in a **.json** (JavaScript Object Notation) file
- once the sequence has been located, it has to be converted from a **three-letter sequence into a one-letter sequence**
- the code returns the sequence

What you need (see bcourses, repo)

Let's explore `ParsingJsonExtractAA_Students.py`

Lab: Parsing Json

exercise 1 (20min):

complete the `__init__`

the output should be something like

```
method: ratio

index and match:

411    sequenceCwu
dtype: object
------------------------------


------------------------------

method: partial_ratio

index and match:

411             sequenceCwu
412    sequenceListNewRules
413    sequenceListOldRules
414       sequencesListText
dtype: object
------------------------------


------------------------------

method: token_sort_ratio

index and match:

411    sequenceCwu
dtype: object
------------------------------


------------------------------

method: token_set_ratio

index and match:

411    sequenceCwu
dtype: object
------------------------------


------------------------------
```

Lab: Parsing Json

exercise 2 (15min):

complete the `ExtractSeq` method

Use the `dict` in order to translate the three-letter AA sequence into one-letter AA sequence

the output should be something like

Inde:	Type	Size	Value
0	str	16	QVQLQESGPGLVKPSQ
1	str	11	GSISSGGYYWS
2	str	14	WIRQHPGKGLEWIG
3	str	16	SIYYSGSTYYNPSLKS
4	str	16	RVTISVDTSKNQFSLK
5	str	11	AREGYHSGMDV
6	str	11	WGQGTTVTVSS
7	str	16	QVQLQESGPGLVKPSQ
8	str	16	LKSRVTISVDTSKNQF
9	str	16	DIQMTQSPSSLSASVG
10	str	11	QASQDISNYLN
11	str	15	WYQQKPGKAPKLLIY
12	str	7	DASNLAT
13	str	16	GVPSRFSGSGSGTDFT

Lab: Parsing Json

exercise 3 (10min):

take a look at the raw sequence and split it:

the output should be something like

based on that: How would you filter for the additional information?

Inde:	Type	Size	Value
0	str	19	225126PRTArtificial
1	str	10	SequenceVH
2	str	3	FR1
3	str	2	of
4	str	10	Antibodies
5	str	8	SIRPAB-1
6	str	2	to
7	str	8	SIRPAB-7
8	str	3	and
9	str	9	SIRPAB-17
10	str	2	to
11	str	9	SIRPAB-21
12	str	4	1Gln
13	str	3	Val
14	str	3	Gln

Inde:	Type	Size	
211	str	2	of
212	str	8	Antibody
213	str	8	SIRPAB-1
214	str	11	8caggtgcagc
215	str	10	tgcaggagtc
216	str	10	gggcccagga
217	str	10	ctggtgaagc
218	str	10	cttcacagac
219	str	10	cctgtccctc
220	str	12	60acctgtactg
221	str	10	tctctggtgg
222	str	10	ctccatcagc