```
1 /Users/kavyajaganathan/Downloads/hw1/v/bin/python /
  Users/kavyajaganathan/Downloads/hw1/group_6.py
2 3.6.7
3 3.6.7
4 setting parameters
5 Other done
6 Years done
7 Decimals done
8 Integers done
9 Dates done
10 Tokenized
11 Train file created
12 Val file created
13 Test file created
----+
                i | ii | iii |
15 | data
        vi | Custom Metric 1- Average Sentence
  Length |
16 +========++=====++=====++====++====
  =====+
17 | Train | 417557 | 34730 | 28863 | 24017 | x
    | 137757 |
  17.1411
----+
19 | Validation | 52739 | 8724 | 7685 | 6415 |
  7685 | 17615 |
  18.4596 |
21 | Test
           | 52115 | 8650 | 7662 | 6342 |
  7662 | 17965 |
  20.2231
----+
```

```
23 vii - Train
----+
     NN | ( | : | , | JJ | VBD |
            DT |
                CC I
====+=====+
27 | 109166 | 5210 | 3826 | 23451 | 41335 | 19036 |
 12354 | 5214 | 34005 | 12383 |
----+
29 vii - Validation
30 +----+-----
  -+----+
31 | NNS | VBP | RB | NN | , | ( |
                                NNP
        . |
            DT I
32 +=====++=====++=====+====++=====++=====
 =+=====+=====+
33 | 2439 | 503 | 1253 | 13388 | 2810 | 556 | 1378
  | 555 | 261 | 4390 |
34 +----+-----
 -+----+
35 vii - Test
36 +----+----+-----
 -+----+
    NNS | VBP | RB | NN | , | ( | NNP
     ) |
        .
            DT I
38 +=====++====++====++=====++====++=====
 =+=====+=====+
39 | 2439 | 503 | 1253 | 13388 | 2810 | 556 |
  | 555 | 261 | 4390 |
40 +----+-----
  -+----+
41 setting parameters
42 Please enter a test sequence to encode and recover
  : alonso david kavya chennai
43
44 encode this sequence: alonso david kavya chennai
45 as a list of integers.
46
47 integer encoding: [10707, 1474, 10713, 10508]
```

```
48
49 encode this list [10707, 1474, 10713, 10508]
50 as a text sequence.
51
52 this is the encoded text: alonso david <UNK>
   chennai
53
54 Process finished with exit code 0
55
```