

Python

Sandhya Srinivasa

2020-05-11

```
library(reticulate)
```

```
nt <- "CAGCCCAATCAGGCTCTACTGCCACTAACTTACGCAGGATATATTTACGCCGACGTACT"
```

```
def kmer(sequence, k):  
    counts={}  
    find = len(sequence)-k+1  
    for i in range(find):  
        rf=sequence[i:i+k]  
        if rf not in counts:  
            counts[rf]=0  
        counts[rf]+=1  
    return counts
```

```
kmer(r.nt, 3)
```

```
## {'CAG': 3, 'AGC': 1, 'GCC': 3, 'CCC': 1, 'CCA': 2, 'CAA': 1, 'AAT': 1, 'ATC': 1, 'TCA': 1, 'AGG': 2,
```