

## wk5 ex1

October 2, 2020

```
[48]: Seqs =  
→ ['CCGCAACTCTGGCTTGACCCAAGGACCCGGCCACCCCCTCAGGACAACCTCGCATCTCAGCAGAGCAGCCCCTGGCCCAG',  
→ 'FTD', 'CCACCCATAGGCAGATGGCTCCGCCCCACCCCGGGAGGATTTCTTAATGGG',  
→ 'TCCGGGACCCCGGGGCCGAGCTGGGCGCGGGAGCCCCGCAGGGGCGGAGAGACCCGGCGAGCCTGAAGAA',  
→ 'GGAGGAGAGATTACACAACCTTCA',  
→ 'TCACCCTCCTCTGCCCCCAACTCCATCTCTGAGACCTCCTGCCCCCAAAAAAGAA']
```

```
[144]: def GC_content(dna_list=[], *args):  
    seq_org = {}  
    for dna in dna_list:  
        for nt in dna:  
            if nt not in ['A', 'T', 'C', 'G']:  
                seq_org[dna] = 'not nt'  
                break  
            else:  
                length = len(dna)  
                g_content = dna.upper().count('G')  
                c_content = dna.upper().count('C')  
                gc_content = (100*(g_content + c_content))/ length  
                round_gc_content = round(gc_content, 1)  
            if gc_content > 65:  
                seq_org[dna] = 'High GC content- ' + str(round_gc_content)  
            elif gc_content < 45:  
                seq_org[dna] = 'Low GC content- ' + str(round_gc_content)  
            else:  
                seq_org[dna] = 'Medium GC content- ' + str(round_gc_content)  
    return(seq_org)
```

```
[145]: GC_content(Seqs)
```

```
[145]: {'CCGCAACTCTGGCTTGACCCAAGGACCCGGCCACCCCCTCAGGACAACCTCGCATCTCAGCAGAGCAGCCCCTGGCCCA  
G': 'High GC content- 67.1',  
'FTD': 'not nt',  
'CCACCCATAGGCAGATGGCTCCGCCCCACCCCGGGAGGATTTCTTAATGGG': 'Medium GC content-  
64.2',  
'TCCGGGACCCCGGGGCCGAGCTGGGCGCGGGAGCCCCGCAGGGGCGGAGAGACCCGGCGAGCCTGAAGAA':  
'High GC content- 78.9',
```

```
'GGAGGAGAGATTACACA ACTTCA': 'Low GC content- 43.5',  
'TCACCTCCTCTGCCCCAACTCCATCTCTGAGACCTCCTGCCCCCAAAAAAGAA': 'Medium GC  
content- 56.9'}
```

[ ]:

[ ]: