

ads1-hw-examples (/github/BenLangmead/ads1-hw-examples/tree/master)
/ hw1_naive_2mm.ipynb (/github/BenLangmead/ads1-hw-examples/tree/master/hw1_naive_2mm.ipynb)

```
In [1]: # Implement the naive_2mm function
from naive_2mm import naive_2mm
```

Example 1

```
In [2]: p = 'CTGT'
ten_as = 'AAAAAAAAAA'
t = ten_as + 'CTGT' + ten_as + 'CTTT' + ten_as + 'CGGG' + ten_as
occurrences = naive_2mm(p, t)
print(occurrences)

[10, 24, 38]
```

Example 2

```
In [3]: # Phi-X genome
!wget http://d396qusza40orc.cloudfront.net/ads1/data/phix.fa

--2015-08-24 21:55:39--  http://d396qusza40orc.cloudfront.net/ads1/data/phix.fa
Resolving d396qusza40orc.cloudfront.net... 54.230.39.133, 54.230.39.39, 54.230.38.100, ...
Connecting to d396qusza40orc.cloudfront.net|54.230.39.133|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 5528 (5.4K) [application/octet-stream]
Saving to: 'phix.fa'

phix.fa          100%[=====>]    5.40K  --.-KB/s   in 0.001s

2015-08-24 21:55:39 (5.09 MB/s) - 'phix.fa' saved [5528/5528]
```

```
In [4]: def readGenome(filename):
        genome = ''
        with open(filename, 'r') as f:
            for line in f:
                # ignore header line with genome information
                if not line[0] == '>':
                    genome += line.rstrip()
        return genome
```

```
In [5]: phix_genome = readGenome('phix.fa')
```

```
In [6]: occurrences = naive_2mm('GATTACA', phix_genome)
```

```
In [7]: print('offset of leftmost occurrence: %d' % min(occurrences))

offset of leftmost occurrence: 10
```

```
In [8]: print('# occurrences: %d' % len(occurrences))

# occurrences: 79
```

