Centre for Genomic Research

UKRI Natural Environment Research Council

NEOF NERC ENVIRONMENTAL OMICS FACILITY

Python for Bioinformatics

Reading & Writing Files

University of Liverpool

The University Of Sheffield.

Map

Reading / Writing Files:

Data streams

Write FASTQ data

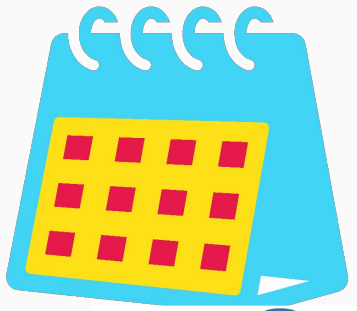
Calendar

Morning 1: Foundations of Python

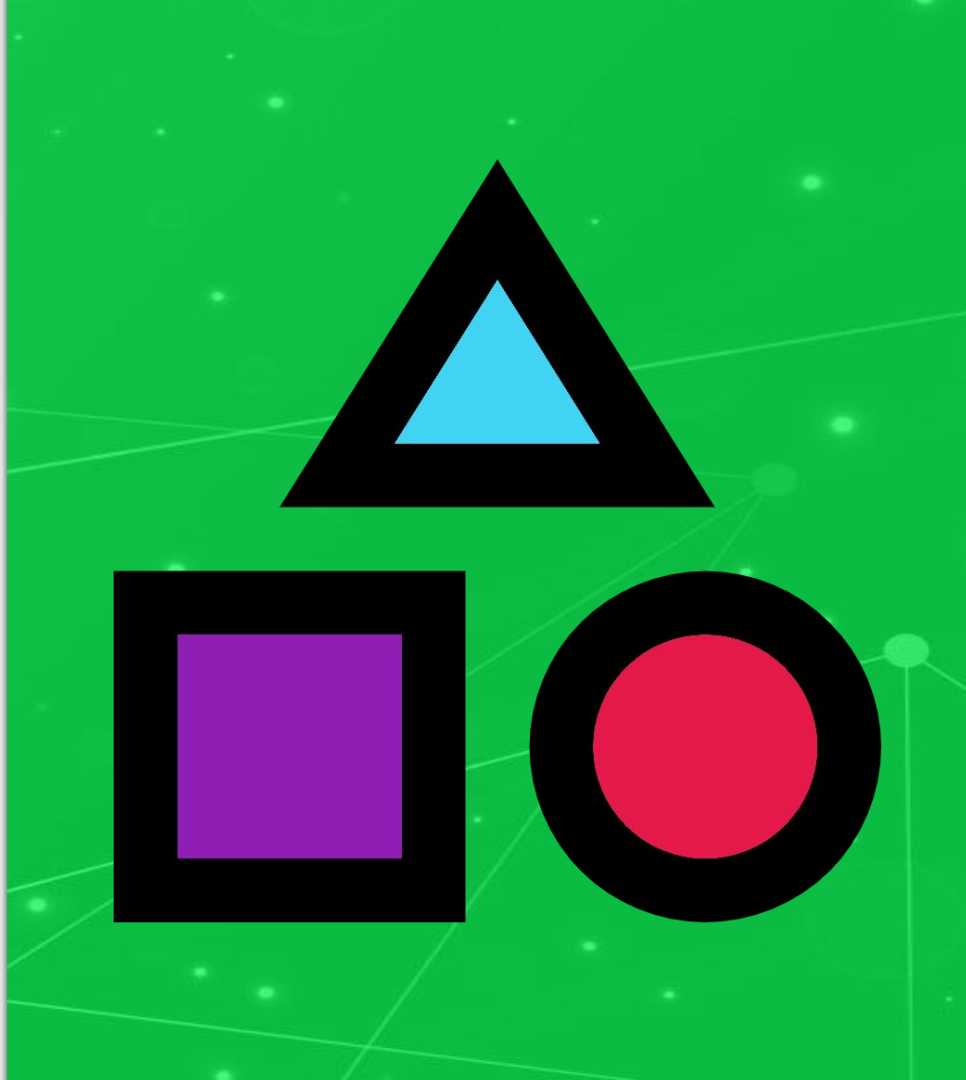
Afternoon 1: Data Structures, Conditions, & Loops

Morning 2: Reading & Manipulating DNA Data

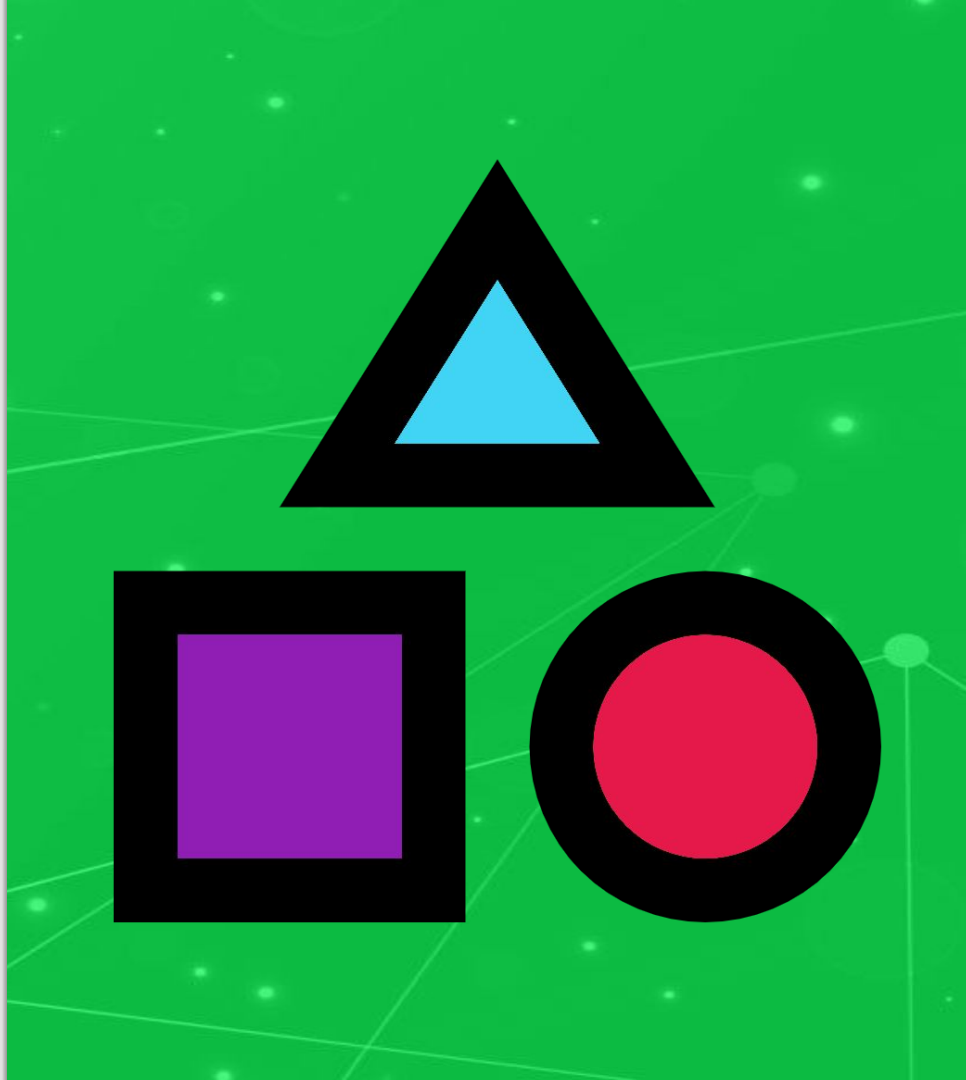
Afternoon 2: **Reading and Writing Files**



Reading and Writing Files



Reading and Writing Files



Opening files

open (file, “mode”)

“Open” establishes a link to the file

```
# read file using ‘with’ and ‘open’

with open(“/path/to/file”, ‘r’) as input_file:
```

Three options:

r = ‘read’

w = ‘write’

a = ‘append’

python™

Opening files

open (file, “mode”)

“Open” establishes a link to the file

```
# read file using ‘with’ and ‘open’

with open(“/path/to/file”, ‘r’) as input_file:
```

Three options:

r = ‘read’

w = ‘write’

a = ‘append’

python™

Opening files

a)

with open( )

b)

as variable

```
# read file using 'with' and 'open'

with open("/path/to/file", 'r') as input_file:
```

Opening files

a)

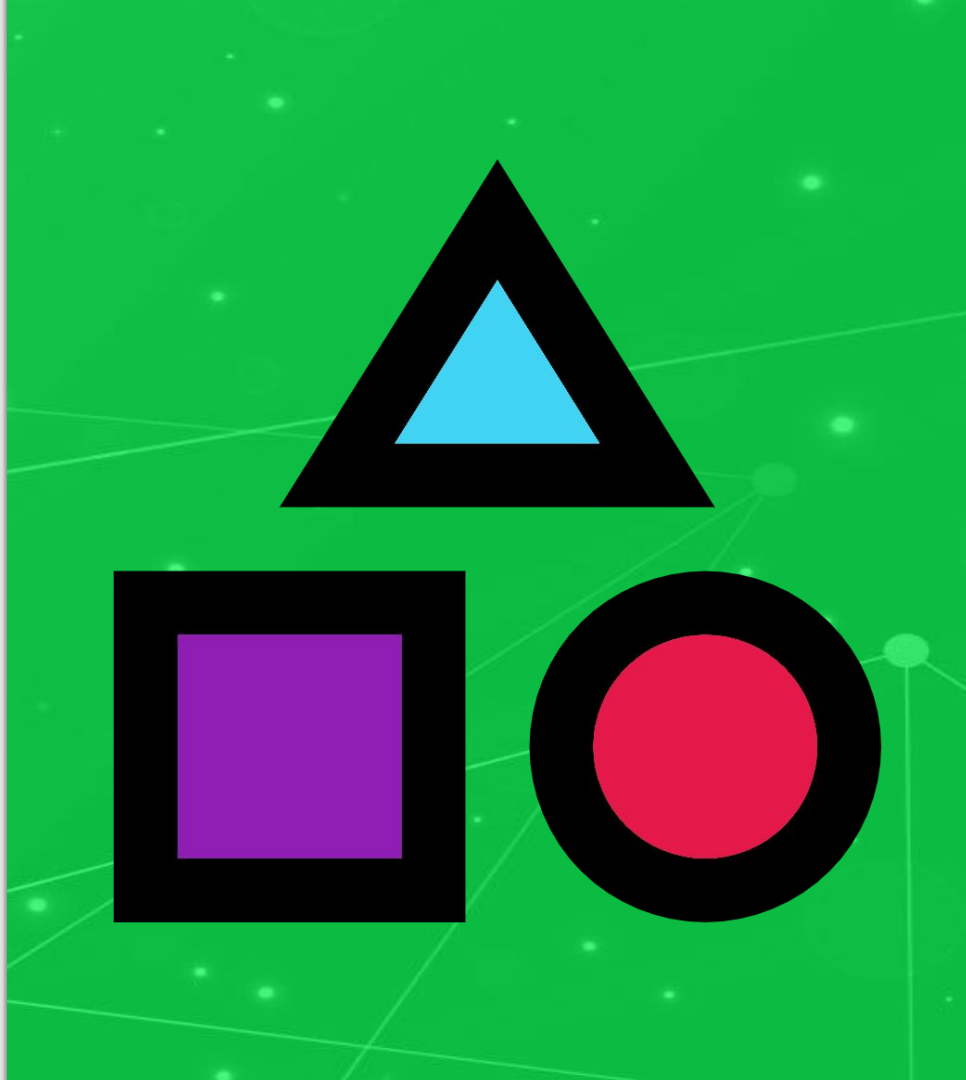
with open( )

b)

as variable

```
# read file using 'with' and 'open'
with open("/path/to/file", 'r') as input_file:

    # loop through the file
    for row in input_file:

        # do something
        print(row)
```

Reading and Writing Files

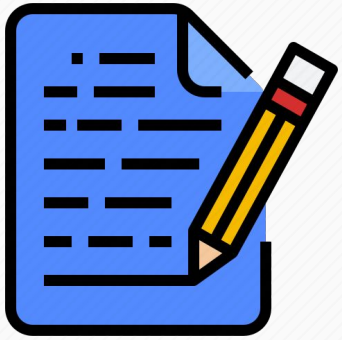


Writing files

Use the

.write()

method



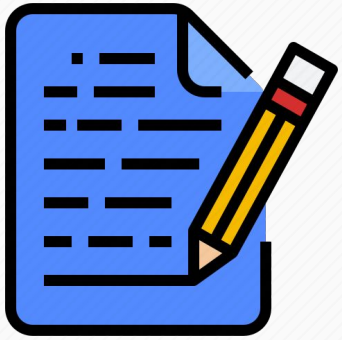
```
# write file using ‘with’ and ‘open’
with open(“/path/to/file”, ‘w’) as output_file:
```

Writing files

Use the

.write()

method



```
# write file using ‘with’ and ‘open’

with open(“/path/to/file”, ‘w’) as output_file:
```

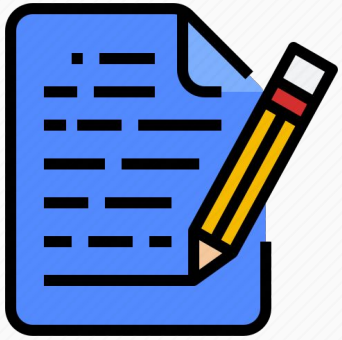
python™

Writing files

Use the

.write()

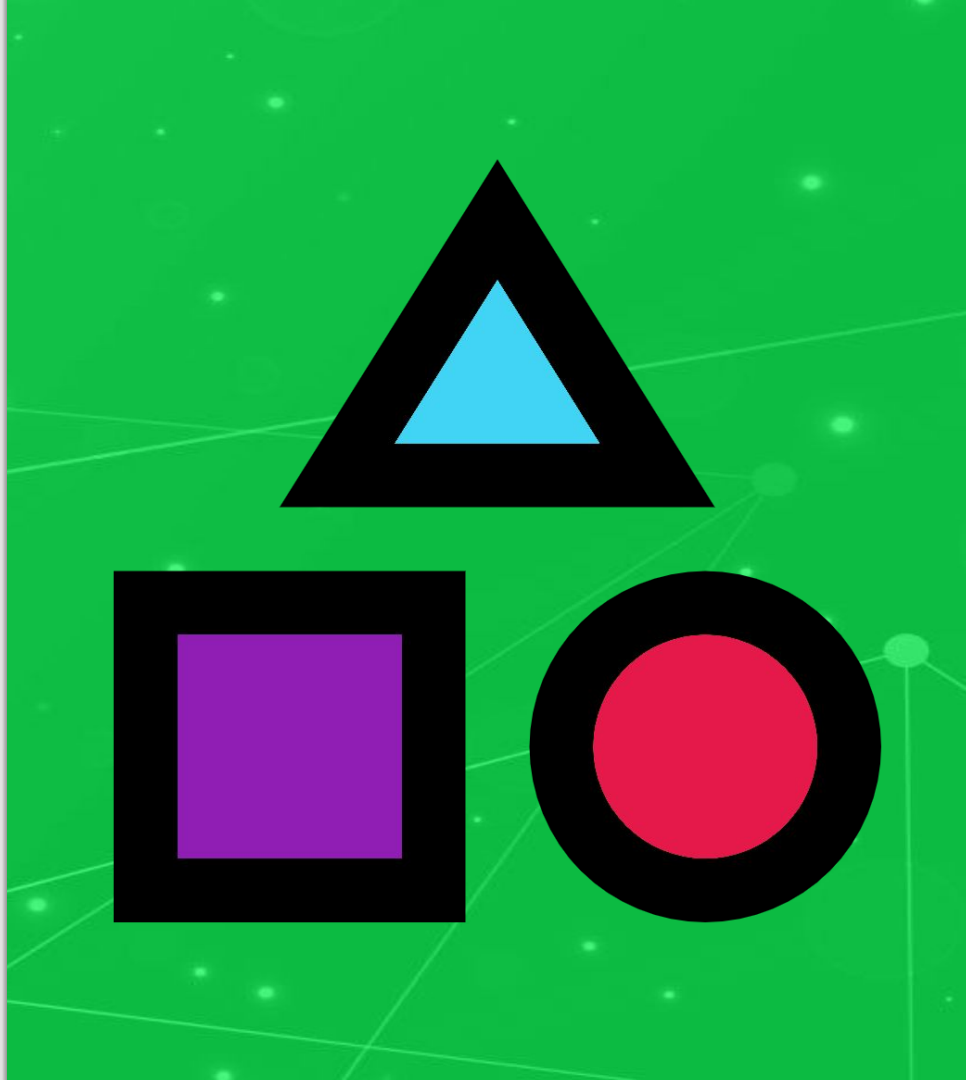
method



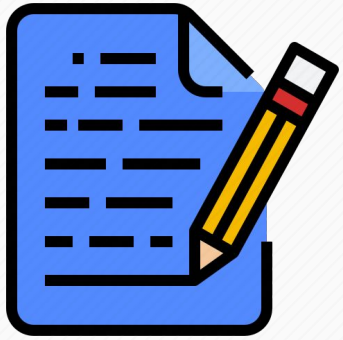
```
# write file using 'with' and 'open'
with open("/path/to/file", 'w') as output_file:

    output_file.write("text_to_write")

    output_file.write("CGTGATGAG")
```

python™

Writing DNA Data to File



Writing DNA Sequence files



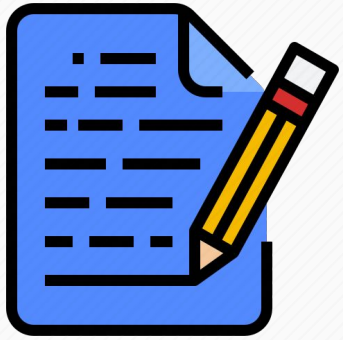
Use the SeqIO.write() method

```
seqIO.write(list_of_sequence_records,
            output_filename,
            format)
```

Format options: “fastq”, “fasta”

python™

Writing DNA Sequence files



Use the SeqIO.write() method

```
seqIO.write(list_of_sequence_records,
            output_filename,
            format)

seqIO.write(sequence_records,
            "output.fastq", "fastq")
```

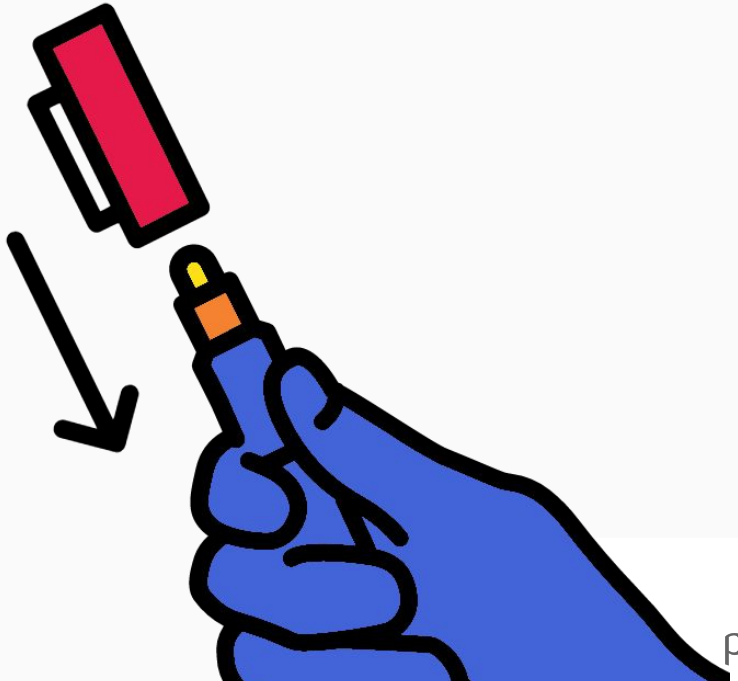
python™

Recap

with & open

read, write, append

SeqIO.write()



Questions?