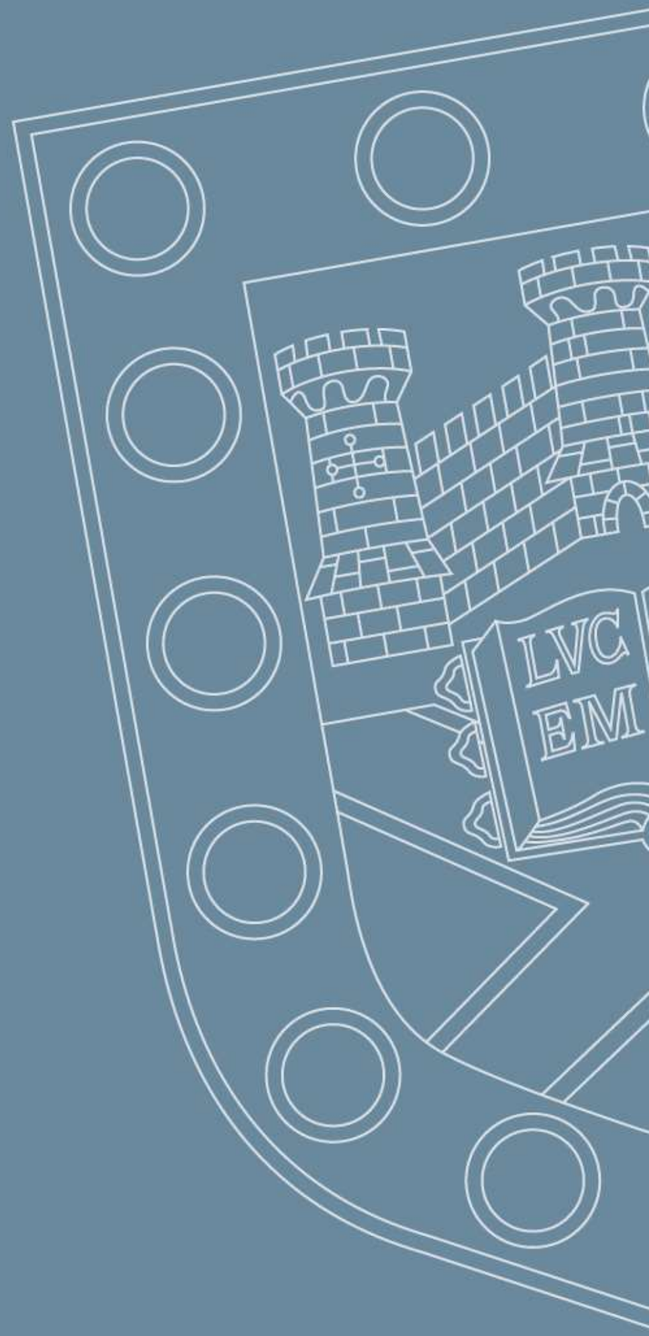
UNIVERSITY OF EXETER

Python programming

Files



institute of CODING

FILES

What are files?

Files are *persistent* storage.

Files can be local, or remote (networked).

Files are blocks of data on a storage device.

Usual to access files sequentially, i.e. reading characters from start to end.

FILES

Input

```
text = input("prompt> ")
num_str = input("type a number")
num = float(num_str)
# or maybe int() is more appropriate?
```

FILES

Reading text files

```
fileref = open('myfile.txt')
lineno = 1
for line in fileref:
    print(lineno, line)
```

FILES

Or read all lines at once

```
lines = open('myfile.txt').readlines()
```

FILES

“Binary” files

```
data = open('picture.jpg', 'b').read()
```

FILES

CSV files


```
"Country Name","Country Code","Indicator Name","Indicator
Code","1999","2000","2001","2002","2003","2004","2005","2006","2007","2008","2009","2010","2011","2012","2013","2014
","2015","2016","2017","2018","2019","2020","2021", "Euro Area","E19","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","1.9","2.4","1.9","1.6","1.5","1.3",
"Advanced economies","AME","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","1.7","2.3","2.2","2","1.6","1.5", "East Asia
and Pacific","EAA","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","6.3","6.6","6.3","6","6","5.8", "Europe and
Central Asia","ECH","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","1.7","4","3.1","2.3","2.7","2.9", "Latin
America and the Caribbean","LAP","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","-1.5","0.8","0.6","1.7","2.4","2.5", "Middle
East and North Africa","MNH","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","5.1","1.2","1.7","1.9","2.7","2.7", "South
Asia","SAP","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","7.5","6.2","6.9","7.1","7.1","7.1", "Sub-
Saharan Africa","SSP","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","1.3","2.6","2.7","3.4","3.6","3.7", "World
(WBG members)","WLT","GDP growth, constant 2010
USD","NYGDPMKTPKDZ","","","","","","","","","","","","","","","","","","","","","","2.4","3.1","3","2.9","2.8","2.8",
```

FILES

Reading CSV files

```
for row in open("datafile.csv"):
    cols = ",".split(row)
```

FILES

XML files

```
<wtext type="FICTION">
  <pb n="1"/>
    <div level="1">
      <head type="MAIN">
        <s n="1">
          <w c5="NN1" hw="prologue" pos="SUBST">PROLOGUE</w>
        </s>
      </head>
    </div>
  </pb>
</wtext>
```

FILES

JSON

JSON files look like literals.

```
["ant","spider","bee","wasp"]
```

FILES

JSON


```
[{"id":"3874","latitude":"50.807","longitude":
"-1.208","name":"Solent Mrsc"},{"id":"3882",
"latitude":"50.89","longitude":"0.319","name":
"Herstmonceux West End"}]
```

FILES

Reading and writing JSON

```
import json
data = json.load(open('datasample.json'))
for record in data:
    print(record)
# Modify the data
for record in data:
    record['checked'] = True
json.dump(data, open('newdata.json', 'w'))
```

FILES

Reading and writing JSON

Take a look at the new file newdata.json. Is there anything about it that surprises you?

```
[{"id": "3874", "latitude": "50.807", "longitude": "-1.208", "name": "Solent Mrsc", "checked": true}...
```

FILES

Terminology

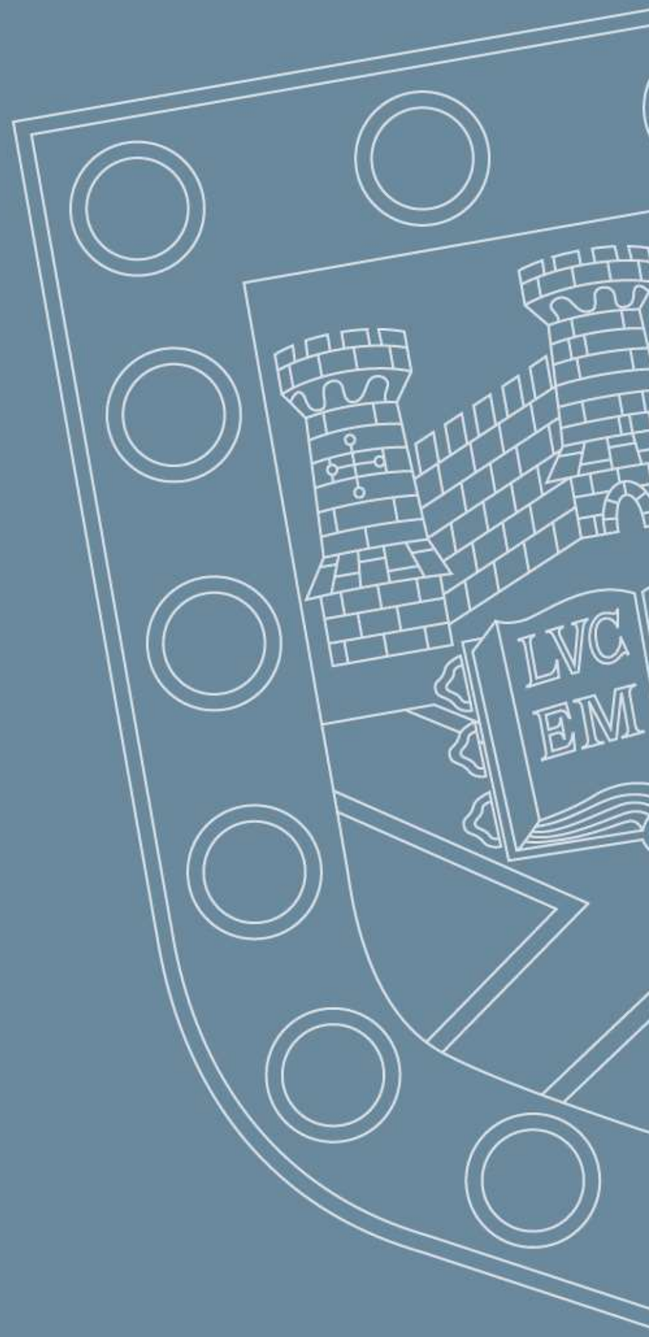
By now you've noticed that programmers tend to use common English words to mean particular things.

Read – nearly always means read from a file

Write – write to a file, i.e. store the data

Print – display on the screen (terminal)

UNIVERSITY OF EXETER



institute of CODING