THOMAS MORE

WORKING WITH CSV FILES AND JSON DATA

Given by Yves Debruyn

SUMMARY

- CSV and JSON Files explained
- CSV module
- JSON module
- Quiz
- Q&A

CSV AND JSON FILES EXPLAINED

- CSV (Comma-Separated Values)
  - Simplified spreadsheets stored as plaintext files

```
r-nummer, naam, voornaam
r0000001, Doe, Jane
r0000002, Doe, John
```

r-nummer	naam	voornaam
r0000001	Doe	Jane
r0000002	Doe	John

CSV AND JSON FILES EXPLAINED

- JSON (JavaScript Object Notation)
  - Is a format that stores information as JavaScript source code in plaintext files

```
[ {
    "Naam": "JSON",
    "Type": "Gegevensuitwisselingsformaat",
    "isProgrammeertaal": false,
    "Zie ook": [ "XML", "ASN.1" ]
  },
  {
    "Naam": "JavaScript",
    "Type": "Programmeertaal",
    "isProgrammeertaal": true,
    "Jaar": 1995
  }
]
```

CSV MODULE

- Why use the CSV module?
  - Not every ‘,’ separates a cell
  - The use of escape characters won’t work
  - Using the CSV module is just easier ☺

READER OBJECTS

- What?
- Example:

```
1  import csv
2  exampleFile = open('test_files/studenten.csv')
3  exampleReader = csv.reader(exampleFile)
4  exampleData = list(exampleReader)
5  print(exampleData)
```

```
[['r-nummer', ' naam', ' voornaam'], ['r0000001', ' Doe', ' Jane'], ['r0000002', ' Doe'
, ' John'], ['r0000003', ' Doe', ' James'], ['r0000004', ' Doe', ' Jimmy']]
```

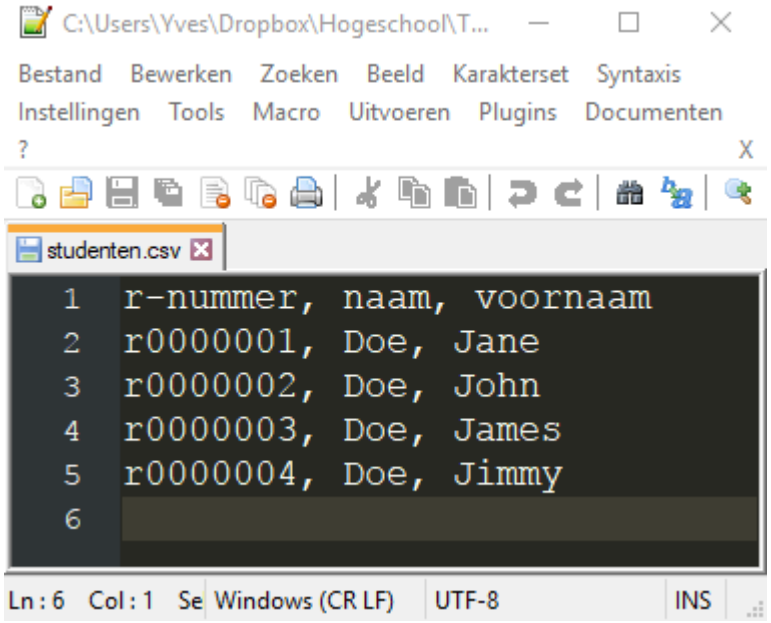
READER OBJECTS

- Retrieve data from list

```
>>> import csv
>>> exampleFile = open('example.csv')
>>> exampleReader = csv.reader(exampleFile)
>>> exampleData = list(exampleReader)
>>> exampleData
[['4/5/2015 13:34', 'Apples', '73'], ['4/5/2015 3:41', 'Cherries', '85'],
['4/6/2015 12:46', 'Pears', '14'], ['4/8/2015 8:59', 'Oranges', '52'],
['4/10/2015 2:07', 'Apples', '152'], ['4/10/2015 18:10', 'Bananas', '23'],
['4/10/2015 2:40', 'Strawberries', '98']]
>>> exampleData[0][0]
'4/5/2015 13:34'
>>> exampleData[0][1]
'Apples'
>>> exampleData[0][2]
'73'
>>> exampleData[1][1]
'Cherries'
>>> exampleData[6][1]
'Strawberries'
```

READER OBJECTS

- What is the output?




```python
import csv
exampleFile = open('test_files/studenten.csv')
exampleReader = csv.reader(exampleFile)
exampleData = list(exampleReader)

print(exampleData[0][1])
```

READER OBJECTS

- Reading data from Reader Objects in a for loop

```
>>> import csv
>>> exampleFile = open('example.csv')
>>> exampleReader = csv.reader(exampleFile)
>>> for row in exampleReader:
        print('Row #' + str(exampleReader.line_num) + ' ' + str(row))

Row #1 ['4/5/2015 13:34', 'Apples', '73']
Row #2 ['4/5/2015 3:41', 'Cherries', '85']
Row #3 ['4/6/2015 12:46', 'Pears', '14']
Row #4 ['4/8/2015 8:59', 'Oranges', '52']
Row #5 ['4/10/2015 2:07', 'Apples', '152']
Row #6 ['4/10/2015 18:10', 'Bananas', '23']
Row #7 ['4/10/2015 2:40', 'Strawberries', '98']
```

WRITER OBJECTS

- Write data to CSV file

```
>>> import csv
>>> outputFile = open('output.csv', 'w', newline='')
>>> outputWriter = csv.writer(outputFile)
>>> outputWriter.writerow(['spam', 'eggs', 'bacon', 'ham'])
21
>>> outputWriter.writerow(['Hello, world!', 'eggs', 'bacon', 'ham'])
32
>>> outputWriter.writerow([1, 2, 3.141592, 4])
16
>>> outputFile.close()
```

WRITER OBJECTS

- The delimiter and lineterminator Keyword Arguments

```
>>> import csv
>>> csvFile = open('example.tsv', 'w', newline='')
>>> csvWriter = csv.writer(csvFile, delimiter='\t', lineterminator='\n\n')
>>> csvWriter.writerow(['apples', 'oranges', 'grapes'])
24
>>> csvWriter.writerow(['eggs', 'bacon', 'ham'])
17
>>> csvWriter.writerow(['spam', 'spam', 'spam', 'spam', 'spam', 'spam'])
32
>>> csvFile.close()
```

JSON AND APIs

- Reading JSON with the loads() function

```
>>> stringOfJsonData = '{"name": "Zophie", "isCat": true, "miceCaught": 0,
"felineIQ": null}'
>>> import json
>>> jsonDataAsPythonValue = json.loads(stringOfJsonData)
>>> jsonDataAsPythonValue
{'isCat': True, 'miceCaught': 0, 'name': 'Zophie', 'felineIQ': None}
```

```
>>> jsonDataAsPythonValue['name']
'Zophie'
```

JSON AND APIs

- Writing JSON with the dumps() function

```
>>> pythonValue = {'isCat': True, 'miceCaught': 0, 'name': 'Zophie',
'felineIQ': None}
>>> import json
>>> stringOfJsonData = json.dumps(pythonValue)
>>> stringOfJsonData
'{"isCat": true, "felineIQ": null, "miceCaught": 0, "name": "Zophie" }'
```

QUIZ

1. What are some features Excel spreadsheets have that CSV spreadsheets don't?
2. What do you pass to csv.reader() and csv.writer() to create Reader and Writer objects?
3. What modes do File objects for reader and Writer objects need to be opened in?
4. What method takes a list argument and writes it to a CSV file?
5. What do the delimiter and lineterminator keyword arguments do?
6. What function takes a Python data structure and returns a string of JSON data?
7. What function takes a string of JSON data and returns a Python data structure?

Q&A

- Any questions?