Department of Computer Science and Engineering
PES University, Bangalore, India

Lecture Notes
Python for Computational Problem Solving
UE23CS151A

Lecture #50
File operations: Working with CSV files

By,
Prof. Sindhu R Pai,
Anchor, PCPS - 2023
Assistant Professor
Dept. of CSE, PESU

Verified by,
PCPS Team - 2023

Many Thanks to
Dr. Shylaja S S (Director, CCBD and CDSAML Research Centers, Former Chairperson, CSE, PES University)
Prof. Chitra G M(Asst. Prof, Dept. of CSE, PCPS Anchor – 2022)

Introduction to CSV files

Comma-separated values (CSV) is a text file format that uses commas to separate values. A CSV file stores tabular data (numbers and text) in plain text, where each line of the file typically represents one data record. Each record consists of the same number of fields, and these are separated by commas in the CSV file. If the field delimiter itself may appear within a field, fields can be surrounded with quotation marks.

To deal with **CSV files, python supports two modules – csv and pandas**. Let us try to understand the csv module and its functionality.

Consider matches.csv file. Reading this file using normal file read operations, we had to use splitlines() and split() multiple times to access particular field which we need in our application. Refer to Lecture #49. This can be avoided easily using csv module. This module is available by default. But pandas must be downloaded explicitly.

The csv module implements classes to read and write tabular data in CSV format. Programmers can also describe the CSV formats understood by other applications or define their own special-purpose CSV formats. The csv module's reader and writer objects read and write sequences. Programmers can also read and write data in dictionary form using the DictReader and DictWriter classes. We will be looking at only reader function as per the theory syllabus.

The reader function returns the lazy reader object which will iterate over lines in the given *csvfile*. Each row read from the csv file is returned as a list of strings.

The requirement is to find the number of matches in 2009 from the data file matches.csv.

```
import csv      #module must be imported explicitly
s1 = set()
c = 0
with open('matches.csv')as file:
    csvFile = csv.reader(file)    #Functions must be called using modulename.functionname
```

```
    #print(type(csvFile), csvFile)
    for lines in csvFile:
       if lines[1] == "2009":
                c = c + 1
    print(file.closed)
    print("The number of matches in 2009 are", c)
print(file.closed)
```

```
C:\Users\Dell>python practice_files_csv.py
False
The number of matches in 2009 are 57
True
```

Few programs for you to solve!

1. Find the number of matches in which the toss winner is same as the winner in 2016.

2. Print the winners of each season.

3. Find the man of the match awardee who received the award maximum times season wise.

4. Find the number of matches between RCB and KKR in this duration - 2008 and 2016.

-END-