Python for Computational Problem-Solving

LABORATORY MANUAL

Week 9

Semester: I

Course Code: UE23CS151A

Course Anchor: Prof. Sindhu R Pai

Lab Anchor: Dr. Divyashree N

Session: September 2023 – December 2023

Instructions to Faculty

1. All faculty members shall be in the Lab on time.
2. Faculty members should not leave lab sessions unattended, when students are present, and Lab sessions should be engaged fully.
3. Faculty should explain all the lab programs during the lab hours only and make students execute all the lab programs given for that week.
4. Faculty should not share the faculty copy of the lab manual to students.
5. Insist students to use lab systems only and not their laptops.

To Learn and Solve	• Program questions on reading strings from CSV files
	• Program solutions for mandatory questions should be submitted for evaluation. • Additional programs are only for practice, and not for grading

13-11-2023 (Monday) to 17-11-2023(Friday):

Sections - All Sections:

You are given a CSV file called "Articles.csv" that contains news articles' information in rows, including the article, date, heading, and news type. Write Python programs for the following questions.

1) **Problem Statement:** For the given file, write programs to:
   a) Write all the rows which have “Cape Town” in the headings onto a csv file named as “Output.csv.
   b) Open the same file in append, append those rows which have date > 2/1/2016.

Solution:

```
input_csv_file = "News_Articles.csv"
output_txt_file = "Output.csv"

cape_town_rows = []
filtered_rows = []

with open(input_csv_file, "r") as csv_file:
  lines = csv_file.readlines()
  #print(lines)
  cape_town_rows.append(lines[0])
  for line in lines[1:]:
    parts = line.strip().split(',')
    article = parts[0]
    date = parts[1]
    day, month, year = date.split('-')
    day = int(day)
    month = int(month)
    year = int(year)
    #print(day, month, year)
```

```
    if "CAPE TOWN" in article.upper():
        cape_town_rows.append(line)
    elif day > 2 and month > 1 and year > 2015:
        filtered_rows.append(line)

with open(output_txt_file, "w") as txt_file:
    txt_file.writelines(cape_town_rows)

with open(output_txt_file, "a") as txt_file:
    txt_file.writelines(filtered_rows)
```

Output:

	Article	Date	Heading	NewsType
2	Cape Town: Ben Stoke	01-03-2016	England d	sports
3	CAPE TOWN: Ben Stok	01-03-2016	Stokes bat	sports
4	CAPE TOWN: Jonny Ba	01-04-2016	Tearful Ba	sports
5	CAPE TOWN: England	01-05-2016	Bavumas r	sports
6	CAPE TOWN: Captain	01-05-2016	Amla lead	sports
7	CAPE TOWN: Captain	01-04-2016	Amla mak	sports
8	London: Pakistan Yasi	16-07-2016	Yasir stars	sports
9	LONDON: Yasir Shah t	16-07-2016	Yasir Shah	sports
10	KARACHI: State Bank (	30-07-2016	State Bank	business
11	LONDON: Pakistan cap	16-07-2016	Misbah ou	sports
12	LONDON: Pakistan Ya	16-07-2016	Yasir Shah	sports

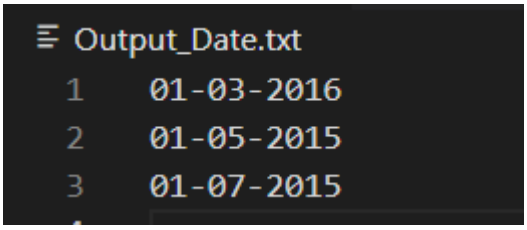
2) **Problem Statement:** From the "Articles.csv" file, write a Python code to extract the "Date" column for rows 7,8 and 9 and store them in a "Output_Date.txt" file.

Solution:

```
# Specify the input CSV file and output text file
```

```
input_csv_file = "News_Articles.csv"
output_txt_file = "Output_Date.txt"

# Initialize a list to store the extracted articles
extracted_articles = []

with open(input_csv_file, "r") as csv_file:
    lines = csv_file.readlines()

    for line_number, line in enumerate(lines, start=1):
        if 7 <= line_number <= 9:
            # Assuming the CSV uses a comma as the delimiter
            columns = line.split(',')

            if len(columns) >= 2:
                # Assuming the "Date" column is at index 1
                article = columns[1].strip()
                extracted_articles.append(article)

if extracted_articles:
  # Write the extracted articles to the output text file
  with open(output_txt_file, "w") as txt_file:
    for article in extracted_articles:
      txt_file.write(article + "\n")
```

Output:



3) **Problem Statement:** Write a program that reads the data from the file, adds a new column called "Location" and writes it to a new csv file named “With_Locations.csv”. **Note - The location can be extracted from the article. The location and article content must be in two separate columns.**

   For example, given the following row in the CSV file:

KARACHI: Wholesale market rates for sugar dropped.,13-01-2015,sugar prices drop,business

The new row in the new CSV file should be:

KARACHI,Wholesale market rates for sugar dropped.,13-01-2015,sugar prices drop,business

Solution:

```
head = 1  # Flag to check for column names
with open("News_Articles.csv", "r") as in_file:
    with open("With_Locations.csv", "w") as out_file:
        for row in in_file:
            if head:  # Modifying column names
                out_file.write("Location," + row)
                head = 0
                continue
            data = row.split(",")
            loc_art = data[0].split(": ")  # Splitting location and article
            new_row = ",".join(loc_art + data[1:])
            out_file.write(new_row)
```

Output:

	Location	Article	Date	Heading	NewsType
2	KARACHI	The Sindh	01-01-2015	sindh govt	business
3	HONG KONG	Asian mar	01-02-2015	asia stock	business
4	KARACHI	Strong bul	01-09-2015	bullish kse	business
5	Singapore	Oil fell fur	01-12-2015	oil falls fu	business
6	KARACHI	Wholesale	13-01-2015	sugar pric	business
7	Cape Town	Ben Stoke	01-03-2016	England d	sports
8	HONG KONG	Hong Kon	01-05-2015	hong kong	business
9	New York	Oil prices	01-07-2015	oil hits ne	business
10	CAPE TOWN	Ben Stoke	01-03-2016	Stokes bat	sports
11	SYDNEY	Carlos Bra	01-04-2016	Brathwait	sports
12	HONG KONG	Asian mar	01-06-2015	asian stoc	business
13	NEW YORK	US oil pric	01-06-2015	us oil pric	business
14	CAPE TOWN	Jonny Bair	01-04-2016	Tearful Ba	sports
15	CAPE TOWN	England g	01-05-2016	Bavumas i	sports
16	MUMBAI	Maharash	01-05-2016	Indian sch	sports
17	SYDNEY	Australia l	01-05-2016	Australia v	sports
18	London	Pakistan Y	16-07-2016	Yasir stars	sports
19	LONDON	Yasir Shah	16-07-2016	Yasir Shah	sports
20	KARACHI	State Bank	30-07-2016	State Bank	business
21	Karachi	Pakistan s	08-01-2016	Pakistan s	business
22	LONDON	Pakistan c	16-07-2016	Misbah ou	sports
23	LONDON	Pakistan Y	16-07-2016	Yasir Shah	sports
24	Islamabad	Internatio	08-01-2016	Pakistan v	business
25	LAHORE	Pakistan C	01-05-2016	Yasir Shah	sports
26	CAPE TOWN	Captain H	01-05-2016	Amla lead	sports
27	ISLAMABAD	The Jewel	08-01-2016	Pakistan J	business
28	Tokyo	Tokyo sto	08-01-2016	Tokyo sto	business
29	SYDNEY	Australia s	01-03-2016	Australias	sports
30	CAPE TOWN	Captain H	01-04-2016	Amla mak	sports

4) **Bonus Question**

Problem Statement: Write a program to read a csv file and filter out the rows which belong to News Type "sports" and were published between March 2016 and April 2016. Write these rows on to a new csv file named "sports.csv", sorted in ascending order by the published date.

Solution:

```
with open('News_Articles.csv', 'r') as file:
   data = file.readlines()
header = data[0]
data = [row.split(',') for row in data[1:] if row]

filtered_rows = []
for row in data:
   columns = row
   if len(columns) == 4 and columns[3].lower().strip() == "sports":
         date = columns[1].strip().split('-')
         if len(date) == 3:
            year = int(date[2])
            month = int(date[1])
            if year == 2016 and 3 <= month <= 4:
               filtered_rows.append(','.join(columns))

filtered_rows.sort(key=lambda x: x.split(',')[1].strip())
filtered_rows.insert(0,header)
with open('sports.csv', 'w') as file:
   file.write(''.join(filtered_rows))
```

Output:

	Article	Date	Heading	NewsType
2	Cape Town: Ben Stol	01-03-2016	England d	sports
3	CAPE TOWN: Ben Stc	01-03-2016	Stokes bat	sports
4	SYDNEY: Australia sp	01-03-2016	Australias	sports
5	SYDNEY: Carlos Brath	01-04-2016	Brathwait	sports
6	CAPE TOWN: Jonny E	01-04-2016	Tearful Ba	sports
7	CAPE TOWN: Captair	01-04-2016	Amla mak	sports

****