ISM 4403 Homework Week 10

### Tasks:

Create a new Excel spreadsheet from the following table and export it as a CSV file without quotes.

**Chart 10.1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| id | Height (inches) | gender | Hair color | Eye Color | Age |
| 1 | 67 | male | brown | brown | 25 |
| 2 | 64 | female | brown | green | 23 |
| 3 | 74 | male | blond | blue | 27 |
| 4 | 73 | Male | brown | brown | 35 |
| 5 | 60 | female | red | green | 40 |
| 6 | 61 | female | brown | green | 45 |
| 7 | 73 | female | blond | blue |  |
| 8 | 70 | female | brown | blue | 50 |
| 9 | 56 | female | blond | brown | 60 |
| 10 | 57 | male | blond | brown | 18 |
| 11 | 64 |  | brown | brown | 25 |
| 12 | 69 | male | brown | green | 23 |
| 13 | 69 | female | blond |  | 27 |
| 14 | 70 | female | brown | brown | 35 |
| 15 | 71 | female | red | green | 40 |
| 16 | 60 | female |  | green | 45 |
| 17 | 80 | male | blond | blue | 41 |
| 18 | 75 | male | brown | blue | 50 |
| 19 | 78 | male | blond | brown | 60 |
| 20 | 69 | male | blond | brown | 18 |
| 21 |  | female | brown | brown | 25 |
| 22 | 66 | male | brown | green | 23 |
| 23 | 74 | female | blond | blue | 27 |
| 24 | 72 | male | brown | brown | 35 |
| 25 | 68 | female | red | green | 40 |
| 26 | 64 | female | brown | green | 45 |
| 27 | 63 | female | blond | blue | 41 |
| 28 |  | male | brown | blue | 50 |
| 29 | 70 | male | blond | brown | 60 |
| 30 | 62 | male | blond | brown | 18 |

Using Jupyter Notebooks perform the following tasks:

Open the CSV file in Python and read each line one at a time storing each in a list.

Print the length of the list.

Print the list to the console using the print() function.

Enumerate the list storing each line in a variable named “line” using a for loop.

While enumerating the list print the length of each line.

Lowercase the text of each line to make sure all alpha characters are lowercase and store this in the list replacing the existing value.

**Paste your code here**

import pandas as pd

import csv

f = pd.read\_csv('lab9\_flat\_file\_try3.csv')

reader = csv.reader(f)

headers = next( reader, None )

print( headers )

column = {}

for h in headers:

column[h] = []

print( column )

for row in reader:

for h, v in zip( headers, row ):

column[h].append(v)

# How do we skip a row???

print( "Columns: ", len( column ) )

print( "Rows: ", len( column['id'] ) )

print( "Entire sheet: \n", column )

for c in column:

print( c.lower(), "," , end="", sep='' )

print()

for x in range(0, len( column['id'] )):

line = ""

for c in column:

print( column[c][x].strip().lower(), ",", end='', sep='' )

line += column[c][x].strip() + ","

print("")

**End of Paste**

**Paste your results here**

Python 3.7.3 (default, Apr 24 2019, 15:29:51) [MSC v.1915 64 bit (AMD64)]

Type "copyright", "credits" or "license" for more information.

IPython 7.6.1 -- An enhanced Interactive Python.

import pandas as pd

import csv

pandas.read\_csv("lab9\_flat\_file.csv)

File "<ipython-input-2-0a1f9b418283>", line 1

pandas.read\_csv("lab9\_flat\_file.csv)

^

SyntaxError: EOL while scanning string literal

pandas.read\_csv("lab9\_flat\_file.csv")

Traceback (most recent call last):

File "<ipython-input-3-8d358244e18d>", line 1, in <module>

pandas.read\_csv("lab9\_flat\_file.csv")

NameError: name 'pandas' is not defined

pd.read\_csv("lab9\_flat\_file.csv")

Traceback (most recent call last):

File "<ipython-input-4-6983abb2c3ac>", line 1, in <module>

pd.read\_csv("lab9\_flat\_file.csv")

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 705, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

FileNotFoundError: [Errno 2] File b'lab9\_flat\_file.csv' does not exist: b'lab9\_flat\_file.csv'

Removing all variables...

pd.read\_csv("lab9\_flat\_filecsv")

Traceback (most recent call last):

File "<ipython-input-5-1bdabacae76f>", line 1, in <module>

pd.read\_csv("lab9\_flat\_filecsv")

NameError: name 'pd' is not defined

import pandas as pd

import csv

pd.read\_csv("lab9\_flat\_filecsv")

Traceback (most recent call last):

File "<ipython-input-7-1bdabacae76f>", line 1, in <module>

pd.read\_csv("lab9\_flat\_filecsv")

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 705, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

FileNotFoundError: [Errno 2] File b'lab9\_flat\_filecsv' does not exist: b'lab9\_flat\_filecsv'

pd.read\_csv("lab9\_flat\_file.csv")

Traceback (most recent call last):

File "<ipython-input-8-6983abb2c3ac>", line 1, in <module>

pd.read\_csv("lab9\_flat\_file.csv")

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 705, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

FileNotFoundError: [Errno 2] File b'lab9\_flat\_file.csv' does not exist: b'lab9\_flat\_file.csv'

Removing all variables...

pd.read\_csv("lab9\_flat\_file.csv")

Traceback (most recent call last):

File "<ipython-input-9-6983abb2c3ac>", line 1, in <module>

pd.read\_csv("lab9\_flat\_file.csv")

NameError: name 'pd' is not defined

import pandas as pd

import csv

pd.read\_csv("lab9\_flat\_file.csv")

Traceback (most recent call last):

File "<ipython-input-10-fd3793661c62>", line 4, in <module>

pd.read\_csv("lab9\_flat\_file.csv")

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 705, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

FileNotFoundError: [Errno 2] File b'lab9\_flat\_file.csv' does not exist: b'lab9\_flat\_file.csv'

pd.read\_csv("C:\\Users\\Owner\\OneDrive\\Desktop\\coding\\ISM4403\_advanced\_business\_analytics\_local\\homework")

Traceback (most recent call last):

File "<ipython-input-11-a245c9c9fb22>", line 1, in <module>

pd.read\_csv("C:\\Users\\Owner\\OneDrive\\Desktop\\coding\\ISM4403\_advanced\_business\_analytics\_local\\homework")

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 709, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

OSError: Initializing from file failed

pd.read\_csv('C:\\Users\\Owner\\OneDrive\\Desktop\\coding\\ISM4403\_advanced\_business\_analytics\_local\\homework')

Traceback (most recent call last):

File "<ipython-input-12-48fed67d6e6a>", line 1, in <module>

pd.read\_csv('C:\\Users\\Owner\\OneDrive\\Desktop\\coding\\ISM4403\_advanced\_business\_analytics\_local\\homework')

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 709, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

OSError: Initializing from file failed

pd.read\_csv('C:\\Users\\Owner\\OneDrive\\Desktop\\coding\\ISM4403\_advanced\_business\_analytics\_local\\homework')

Traceback (most recent call last):

File "<ipython-input-13-48fed67d6e6a>", line 1, in <module>

pd.read\_csv('C:\\Users\\Owner\\OneDrive\\Desktop\\coding\\ISM4403\_advanced\_business\_analytics\_local\\homework')

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 702, in parser\_f

return \_read(filepath\_or\_buffer, kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 429, in \_read

parser = TextFileReader(filepath\_or\_buffer, \*\*kwds)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 895, in \_\_init\_\_

self.\_make\_engine(self.engine)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1122, in \_make\_engine

self.\_engine = CParserWrapper(self.f, \*\*self.options)

File "C:\Users\Owner\Anaconda3\lib\site-packages\pandas\io\parsers.py", line 1853, in \_\_init\_\_

self.\_reader = parsers.TextReader(src, \*\*kwds)

File "pandas/\_libs/parsers.pyx", line 387, in pandas.\_libs.parsers.TextReader.\_\_cinit\_\_

File "pandas/\_libs/parsers.pyx", line 709, in pandas.\_libs.parsers.TextReader.\_setup\_parser\_source

OSError: Initializing from file failed

pd.read\_csv('lab9\_flat\_file\_try3.csv')

Out[14]:

id height gender hair\_color eye\_color age

0 1 67.0 male brown brown 25.0

1 2 64.0 female brown green 23.0

2 3 74.0 male blond blue 27.0

3 4 73.0 Male brown brown 35.0

4 5 60.0 female red green 40.0

5 6 61.0 female brown green 45.0

6 7 73.0 female blond blue NaN

7 8 70.0 female brown blue 50.0

8 9 56.0 female blond brown 60.0

9 10 57.0 male blond brown 18.0

10 11 64.0 NaN brown brown 25.0

11 12 69.0 male brown green 23.0

12 13 69.0 female blond NaN 27.0

13 14 70.0 female brown brown 35.0

14 15 71.0 female red green 40.0

15 16 60.0 female NaN green 45.0

16 17 80.0 male blond blue 41.0

17 18 75.0 male brown blue 50.0

18 19 78.0 male blond brown 60.0

19 20 69.0 male blond brown 18.0

20 21 NaN female brown brown 25.0

21 22 66.0 male brown green 23.0

22 23 74.0 female blond blue 27.0

23 24 72.0 male brown brown 35.0

24 25 68.0 female red green 40.0

25 26 64.0 female brown green 45.0

26 27 63.0 female blond blue 41.0

27 28 NaN male brown blue 50.0

28 29 70.0 male blond brown 60.0

29 30 62.0 male blond brown 18.0

reader = csv.reader(f)

headers = next( reader, None )

Traceback (most recent call last):

File "<ipython-input-15-bddb5a638499>", line 1, in <module>

reader = csv.reader(f)

NameError: name 'f' is not defined

f = pd.read\_csv('lab9\_flat\_file\_try3.csv')

reader = csv.reader(f)

headers = next( reader, None )

print( headers )

['id']

column = {}

for h in headers:

column[h] = []

print( column )

for row in reader:

for h, v in zip( headers, row ):

column[h].append(v)

# How do we skip a row???

print( "Columns: ", len( column ) )

print( "Rows: ", len( column['id'] ) )

print( "Entire sheet: \n", column )

for c in column:

print( c.lower(), "," , end="", sep='' )

print()

for x in range(0, len( column['id'] )):

line = ""

for c in column:

print( column[c][x].strip().lower(), ",", end='', sep='' )

line += column[c][x].strip() + ","

print("")

{'id': []}

Columns: 1

Rows: 5

Entire sheet:

{'id': ['height', 'gender', 'hair\_color', 'eye\_color', 'age']}

id,

height,

gender,

hair\_color,

eye\_color,

age,

**End of Paste**

**Rubric**

Printing the length of each line: 20%

Opening the file: 20%

Lowercasing the line: 20%

Storing the line in a variable named line: 20%

Enumerating the list using a for loop: 20%