

Comma-separated values (CSV) files

“Comma Separated Values” is a common file format when exporting data from spreadsheets and databases. Each line of the file is a row, and each column is separated by a comma.

data.csv file contents:

```
Name,Location,URL,Students
Westminster College,Salt Lake City-UT,westminstercollege.edu,2135
Muhlenberg College,Allentown-PA,muhlenberg.edu,2330
University of Maine,Orono-ME,umaine.edu,8677
James Madison University,Harrisonburg-VA,jmu.edu,19019
Michigan State University,East Lansing-MI,msu.edu,38853
```

1. Write a function called `input_by_name` that takes in the name of a csv file and returns a dictionary organized by name like so:

```
1 college_dict = {  
2     'James Madison University': {'Location': 'Harrisonburg-VA',  
3                                   'Students': '19019',  
4                                   'URL': 'jmu.edu'},  
5     'Michigan State University': {'Location': 'East Lansing-MI',  
6                                    'Students': '38853',  
7                                    'URL': 'msu.edu'},  
8     'Muhlenberg College': {'Location': 'Allentown-PA',  
9                              'Students': '2330',  
10                             'URL': 'muhlenberg.edu'},  
11    'University of Maine': {'Location': 'Orono-ME',  
12                               'Students': '8677',  
13                               'URL': 'umaine.edu'},  
14    'Westminster College': {'Location': 'Salt Lake City-UT',  
15                            'Students': '2135',  
16                            'URL': 'westminstercollege.edu'}  
17 }
```


Note that transpose_by_URL is meant to me a mutator function (that is it changes the dictionary in place without creating another dictionary).

```
1 def transpose_by_URL(college_dict: dict) -> None:
2     keys = list(college_dict.keys())
3     for name in keys:
4         url = college_dict[name]['URL']
5         college_dict[name]['Name'] = name
6         del college_dict[name]['URL']
7         college_dict[url] = college_dict[name]
8         del college_dict[name]
```

4. With this new dictionary does the code for get_all_students need to change? Explain why.

No, still need to go through every key and search for the Students key inside the dictionary associated with it.

5. Write a function called get_min_students that takes in the dictionary organized by URL and returns the name of the college with the least number of students:

```
1 def get_min_students(college_dict: dict) -> str:
2     min = get_all_students(college_dict) #some large value
3     min_name = 'No name'
4     for name in college_dict:
5         if int(college_dict[name]['Students']) < min:
6             min = int(college_dict[name]['Students'])
7             min_name = college_dict[name]['Name']
8
9     return min_name
```

6. Write a function called `input_by_column` that takes in the filename and returns a dictionary organized by the column's name:

```
1 column_dict = {
2     'Location': ['East Lansing-MI',
3                 'Salt Lake City-UT',
4                 'Allentown-PA',
5                 'Orono-ME',
6                 'Harrisonburg-VA'],
7     'Name': ['Michigan State University',
8             'Westminster College',
9             'Muhlenberg College',
10            'University of Maine',
11            'James Madison University'],
12    'Students': ['38853', '2135', '2330', '8677', '19019'],
13    'URL': ['msu.edu',
14           'westminstercollege.edu',
15           'muhlenberg.edu',
16           'umaine.edu',
17           'jmu.edu']
18 }
```

```
1 def input_by_column(filename: str) -> dict:
2     file_dict = {}
3     with open(filename, 'r') as fileobj:
4         header_items = fileobj.readline().strip().split(',')
5         for item in header_items:
6             file_dict[item] = []
7         for line in fileobj:
8             entries = line.strip().split(',')
9             for sindex in range(len(entries)):
10                 file_dict[header_items[sindex]] += [entries[sindex]]
11
12     return file_dict
```

7. Re-write `get_min_students_column` that takes in the dictionary organized by column name and returns the name of the college with the least number of students:

```
1 def get_min_students_column(column_dict: dict) -> str:
2     min_index = 0
3     for index in range(len(column_dict['Students'])):
4         if int(column_dict['Students'][index])
5             < int(column_dict['Students'][min_index]):
6             min_index = index
7
8     return column_dict['Name'][min_index]
```