



Coding challenge

If you're reading this, that means that Percolate is really interested in your engineering abilities. Congratulations!

This is intended to be a small coding exercise meant to gauge your programming style. Now's your chance to show off your design skills (simple but extensible), your testing habits (hey, hey unittest), and your keen attention to detail.

Your solution should be documented and tested to production standard. It should follow good Python conventions, and should be easy to get running.

Included is an input file that you will write a Python program to process. After you've engineered a solution to your satisfaction, package up your response in a tar.gz or .zip file, upload it to the URL (which should be in the e-mail you receive), and we'll get back to you shortly!

The problem

You're tasked with taking entries of personal information in multiple formats and normalizing each entry into a standard JSON format.

Python version

Please ensure your program runs correctly under Python 2.7 (running under Python 3 is fine, as long as it still runs under Python 2.7). In addition please only use modules available in the standard library.

Input

Your program will be fed an input file of n lines. Each line contains "entry" information, which consists of a first name, last name, a USA-style phone number, color, and a 5-digit ZIP code.

The order and format of these lines vary in three separate ways. The three different acceptable input formats are as follows:

```
Lastname, Firstname, (703)-742-0996, Blue, 10013
Firstname Lastname, Red, 11237, 703 955 0373
Firstname, Lastname, 10013, 646 111 0101, Green
```

Some input lines may not conform to any of the formats listed above, should be considered invalid, and should not interfere with the processing of subsequent valid lines. A line should be considered invalid if its phone number does not contain 10 digits (it may contain alphanumeric characters) and a ZIP code must contain, and only contain, 5 digits.

Output

The program should write a valid, formatted JSON object out to result.out. The JSON representation should be indented with two spaces and the keys should be sorted in ascending order.

Successfully processed lines should result in a normalized addition to the list associated with the "entries" key. For lines that were unable to be processed, a line number i (where $0 \leq i < n$) for each faulty line should be appended to the list associated with the "errors" key.

The "entries" list should be sorted in ascending alphabetical order by (last name, first name).

Sample

For the input:

```
Booker T., Washington, 87360, 373 781 7380, light blue
Chandler, Kerri, (623)-668-9293, pink, 123123121
James Murphy, yellow, 83880, 018 154 6474
asdfawefawea
```

We should receive the output:

```
{
  "entries": [
    {
      "color": "yellow",
      "firstname": "James",
      "lastname": "Murphy",
      "phonenumber": "018-154-6474",
      "zipcode": "83880"
    },
    {
      "color": "light blue",
      "firstname": "Booker T.",
      "lastname": "Washington",
      "phonenumber": "373-781-7380",
      "zipcode": "87360"
    }
  ],
  "errors": [
    1,
    3
  ]
}
```

Questions?

Feel free to email.

Good luck!
Your friends at Percolate