

1. Reading file data

- (a) Write a function that takes as input the name of a text file and returns the total number of words in the file. You can test your function with:

http://stanford.edu/~danfrank/cme193/data/course_description.txt

- (b) Write a function that takes as input the name of a text file and returns the most frequently occurring word in the file. You can test your function with:

http://stanford.edu/~danfrank/cme193/data/bill_of_rights.txt

Hint: The Python `sorted()` function might be useful.

- (c) Suppose that we have scraped some web data from reddit. We have a text file where each row contains the base web page of the post, the number of comments, and the number of “upvotes” the post has received. Suppose that the data for each post is separated by `|`. For example, here are 5 data points:

```
imgur.com | 345 | 5060
i.imgur.com | 1228 | 4215
quickmeme.com | 434 | 4312
scientificamerican.com | 185 | 2412
bbc.co.uk | 191 | 726
```

Write a function that computes the average number of comments (number of comments is the right-most column) of the data points. Assume that the name of the data file is the function argument.

The above data is available at:

http://stanford.edu/~arbenson/cme193/data/reddit_data.txt

- (d) Write a function that computes the average number of upvotes for posts on imgur and average number of upvotes for posts not on imgur.

2. Cartesian points

- (a) In the second homework, you represented (x, y, z) points using tuples and lists. Write a class called `CartPoint` that contains `x`, `y`, and `z` points as member variables. The constructor for the class should take an (x, y, z) tuple.
- (b) Add a `magnitude()` function to `CartPoint` that computes the magnitude of the point $(\sqrt{x^2 + y^2 + z^2})$.
- (c) Create a subclass of `CartPoint` called `CartPointTime` that represents an (x, y, z, t) vector, t representing time. The constructor for the class should take an (x, y, z, t) tuple.
- (d) Add a `magnitude_and_time()` to `CartPointTime` function that returns a $(\text{magnitude}, \text{time})$ tuple.

3. Exploring NumPy

- (a) What does the function `numpy.arange()` do? How is it different from `range()`?
- (b) Create an identity matrix with `I = np.eye(3)`. What is the difference between the following two sequence of commands:

```
>>> x = I[0, [0, 1]]
>>> x[0] = 3.0
>>> x[1] = 2.0
```

```
>>> x = I[0, 0:2]
>>> x[0] = 3.0
>>> x[1] = 2.0
```

In the latter sequence, x is called a *view*. Why?

- (c) Using what you learned in part (b), write a function that takes as input a 2-D ndarray and scales the last row and column by 2.5 (the last entry should be scaled by 2.5^2).
- (d) Explain what the following command does: `A = np.random.rand(6, 3) * np.arange(3)`.