In this project, you'll be working with a dataset of submissions to [Hacker News](http://news.ycombinator.com/) from 2006 to 2015. Hacker News is a site where users can submit articles from across the internet (usually about technology and startups), and others can "upvote" the articles, signifying that they like them. The more upvotes a submission gets, the more popular it was in the community. Popular articles get to the "front page" of Hacker News, where they're more likely to be seen by others.

The dataset you'll be using was compiled by Arnaud Drizard using the Hacker News API, and can be found [here](https://github.com/arnauddri/hn). We've sampled 10000 rows from the data randomly, and removed all extraneous columns. Our dataset only has four columns:

* submission\_time -- when the story was submitted.
* upvotes -- number of upvotes the submission got.
* url -- the base domain of the submission.
* headline -- the headline of the submission. Users can edit this, and it doesn't have to match the headline of the original article.

You'll be writing scripts to answer some main questions:

* What words appear most often in the headlines?
* What domains were submitted most often to Hacker News?
* At what times are the most articles submitted?

You'll be answering these questions by writing command line scripts, instead of using IPython notebook. IPython notebooks are great for quick data visualization and exploration, but Python scripts are the way to put anything we learn into production. Let's say you want to make a website to help people write headlines that get as many upvotes as possible, and submit articles at the right time. To do this, you'll need scripts.