6. For this problem, we'll be using the *animals with attributes* data set. Go to

`http://attributes.kyb.tuebingen.mpg.de`

and, under "Downloads", choose the "base package" (the very first file in the list). Unzip it and look over the various text files.

This is a small data set that has information about 50 animals. The animals are listed in `classes.txt`. For each animal, the information consists of values for 85 features: does the animal have a tail, is it slow, does it have tusks, etc. The details of the features are in `predicates.txt`. The full data consists of a 50×85 matrix of real values, in `predicate-matrix-continuous.txt`. There is also a binarized version of this data, in `predicate-matrix-binary.txt`.

Load the real-valued array, and also the animal names, into Python. Now hierarchically cluster this data, using `scipy.cluster.hierarchy.linkage`. Choose Ward's method, and plot the resulting tree using the `dendrogram` method, setting the `orientation` parameter to `'right'` and labeling each leaf with the corresponding animal name.

You will run into a problem: the plot is too cramped because the default figure size is so small. To make it larger, preface your code with the following:

```
from pylab import rcParams
rcParams['figure.figsize'] = 5, 10
```

(or try a different size if this doesn't seem quite right).

(a) Show the dendrogram that you get.

(b) Ward's method of average linkage is essentially trying to minimize the k-means cost function. Let's see how well it does. Take $k = 10$ in what follows:

- Show the k-clustering returned by Ward's method. What is its cost?
- Run k-means on this data, 10 times (each time initializing with 10 centers chosen at random from the data). Pick out the best (lowest cost) solution and show it. What is its cost?