Taxonomy

We can use Kohonen's Map (or SOM) as a way of organizing taxonomic data. For example, suppose that we want to see if we can organize animals by certain characteristics. We will categorize 16 animals based on 13 characteristics. The characteristics are given as follows. A zero means false, and a one means true.

1. Small
2. Medium
3. Large
4. Has 2 legs
5. Has 4 legs
6. Has hair
7. Has hooves
8. Has mane
9. Has feathers
10. Likes to hunt
11. Likes to run
12. Likes to fly
13. Likes to swim

For example, a duck and a wolf would be the following 2 rows (they are the 3d and 12th columns in X):

```
[1  0  0  1  0  0  0  0  1  0  0  1  1]
[0  0  1  0  1  1  0  0  0  1  1  0  0]
```

- Download the 4 files from our class website.
- Run `driverSOM` a couple of times to get the hang of it.
- In `driverSOM`, you'll see the line:

  ```
  net=newsom(minmax(X),[10,10],'gridtop');
  ```

  This creates 100 cluster centers, arranged as a 10×10 grid of points. Each cluster center now has a representation in the grid (in the plane $\mathbb{R}^2$), and a representation in $\mathbb{R}^{13}$. The code (`plotcell`) is currently set up to visualize the clustering in the grid- Each point of X is classified according to what point in the grid is its cluster center, and the animal label appears there. For example, if `horse` is at $(3, 1)$, then it corresponded to that cluster center.