



1. Implement DBSCAN in python.

Hint 1: you need:

- a. A function to calculate how many neighbours are in an epsilon environment of a point (regionQuery).
- b. A function that begins in a core point and expands it till it can't be expanded anymore (expandCluster) and
- c. A function that iterates over all points in the db and if they are core points expands them (DBSCAN).

Hints 2:

<https://en.wikipedia.org/wiki/DBSCAN>

<https://github.com/choffstein/dbscan/blob/master/dbscan/dbscan.py>

2. Check the algorithm on the Iris dataset.
3. Compare to scikit-learn implementation and answer the questions below:  
[https://github.com/scikit-learn/scikit-learn/blob/14031f6/sklearn/cluster/dbscan\\_.py#L156](https://github.com/scikit-learn/scikit-learn/blob/14031f6/sklearn/cluster/dbscan_.py#L156)
  - a. What can you see in the scikit-learn implementation is the difference between "fit" and "fit\_predict"?
  - b. Which of them calls the other
  - c. What does each of them return?