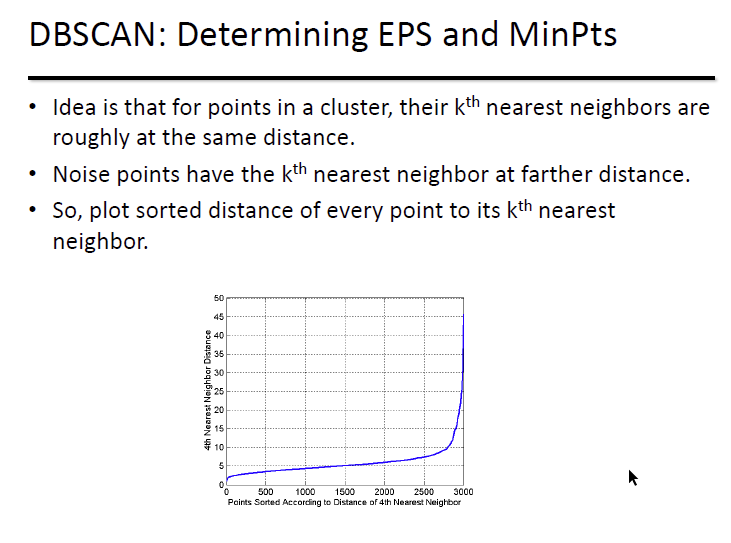
Labels go 1 through 7

Minpts and eps arguments

Each line in train.dat is a text file, Input is csr sparse matrix containing features associated with different feature ids



1. Deal with input data first..
   1. Each row is a text document
   2. Is each number representing a word followed by the occurences w/in document?

**Pip install argparse**

**Pip install ipykernel**

**Microsoft visual c++ 9.0 must be installed, get from** [**http://aka.ms/vcpython27**](http://aka.ms/vcpython27)

**Pip install sparsesvd**

**Pip install “dask[complete]”**