Getting started

To open the csv file and get it in the correct format:

import pandas as pd

data=pd.read_csv('file.csv', index_col=0)

This puts the data from the file into a 'dataframe'. If you print it, it will display the data as a table, and show the first and last few rows and columns. The second argument here 'index_col=0' will make sure the row labels are the words. If you don't include this, then the words get inserted into the first column of the table itself.

To create a smaller dataframe containing subsets of rows and columns:

data1=data.iloc[:,0:10]   # first 10 columns

data1=data.iloc[0:5]   # first 5 rows

data1=data.iloc[0:5,0:5]   # first 5 rows of the first 5 columns

To convert a dataframe into an array (in other words, a matrix):

my_data_array=data1.values