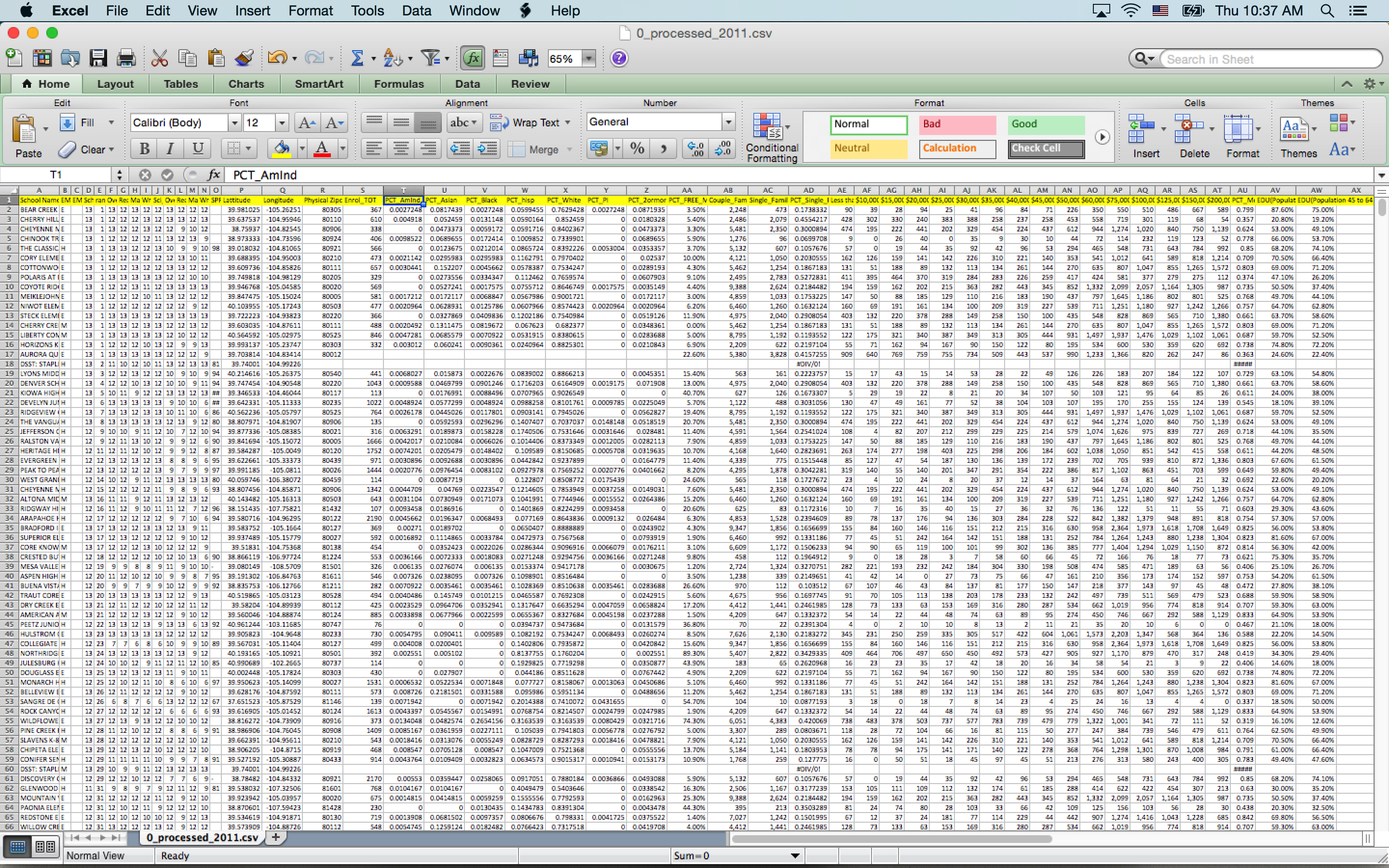
**Data Preprocessing**

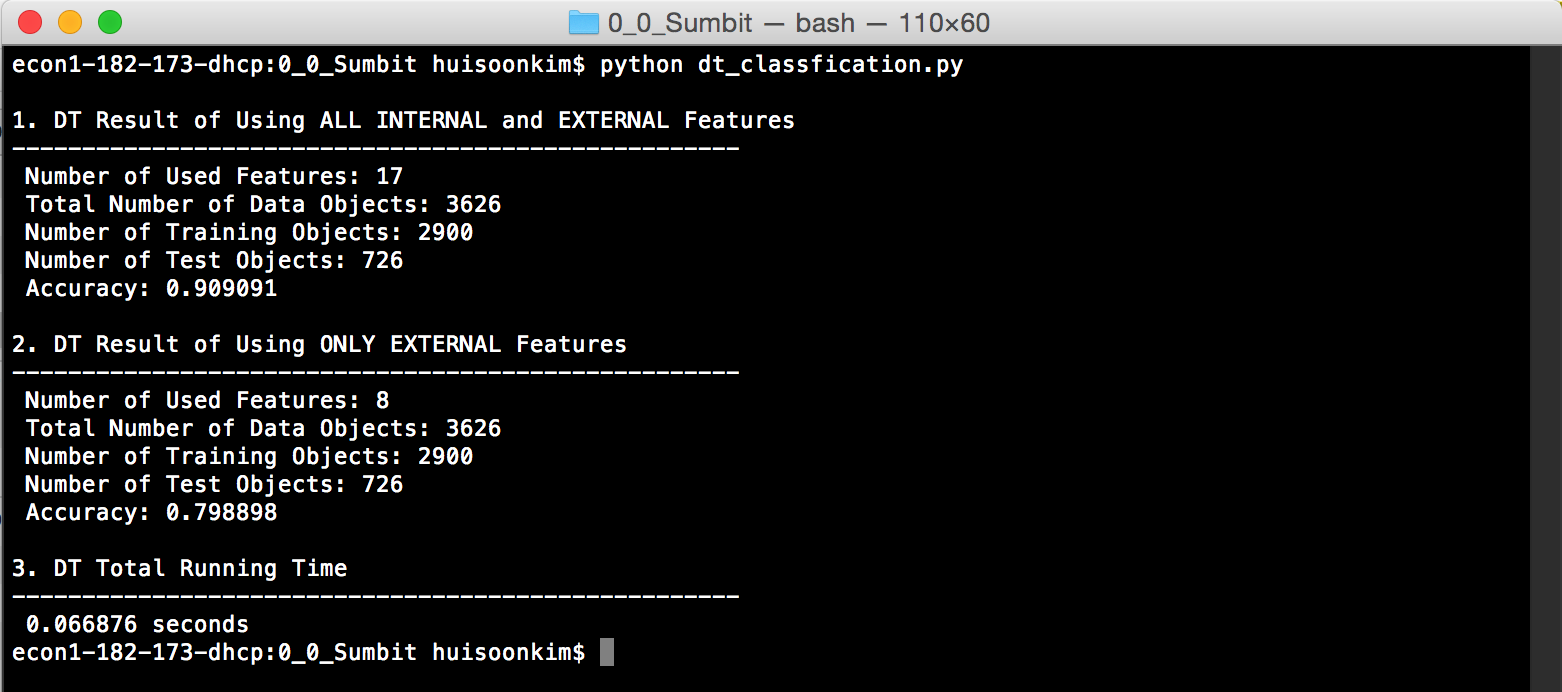
We used Python Script (data\_preprocessing.py) and Excel program for data integration, cleaning, reduction and discretization for school and census data sets. The following snapshot shows the result of data Integration. The result of all of the preprocessing consists of 2,016 objects for year 2011 data set and 1,962 objects for year 2012 data set. And each year’s data sets have total 49 attributes.



**Classification**

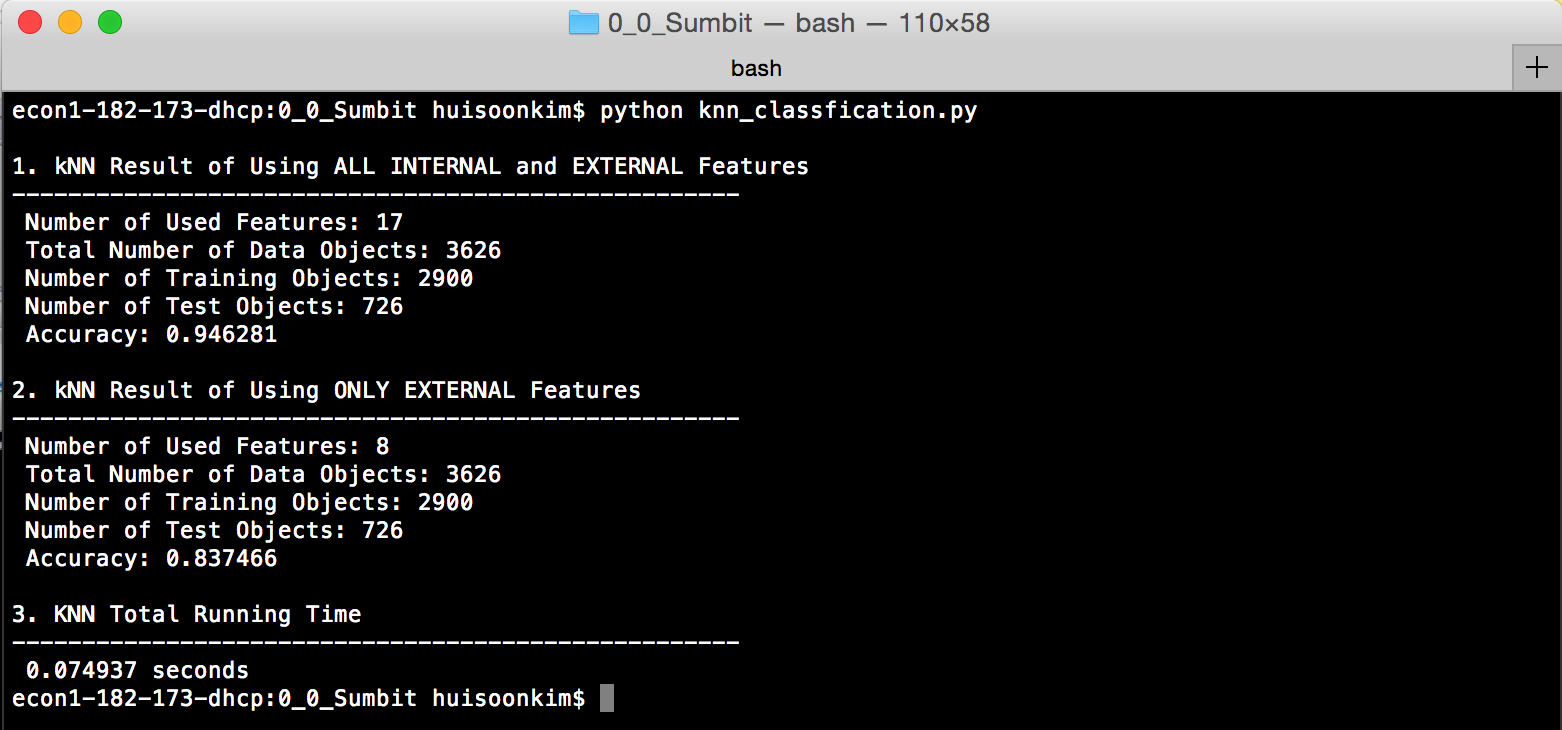
The DT, k-NN and SVM classifier have been applied on two kinds of predictor sets : both direct and external attributes data set, only external attributes data set. And the accuracy is estimated like following figures.

1. Result of Decision Tree Classifier (Source code: “dt\_classification.py”)



\* See also the tree picture of DT from attached file(“DT\_picture.pdf”)

1. Result of k-NN(k=3) Classifier (Source code: “knn\_classification.py”)



1. Result of SVM Classifier (Source code: “svm\_classification.py”)

