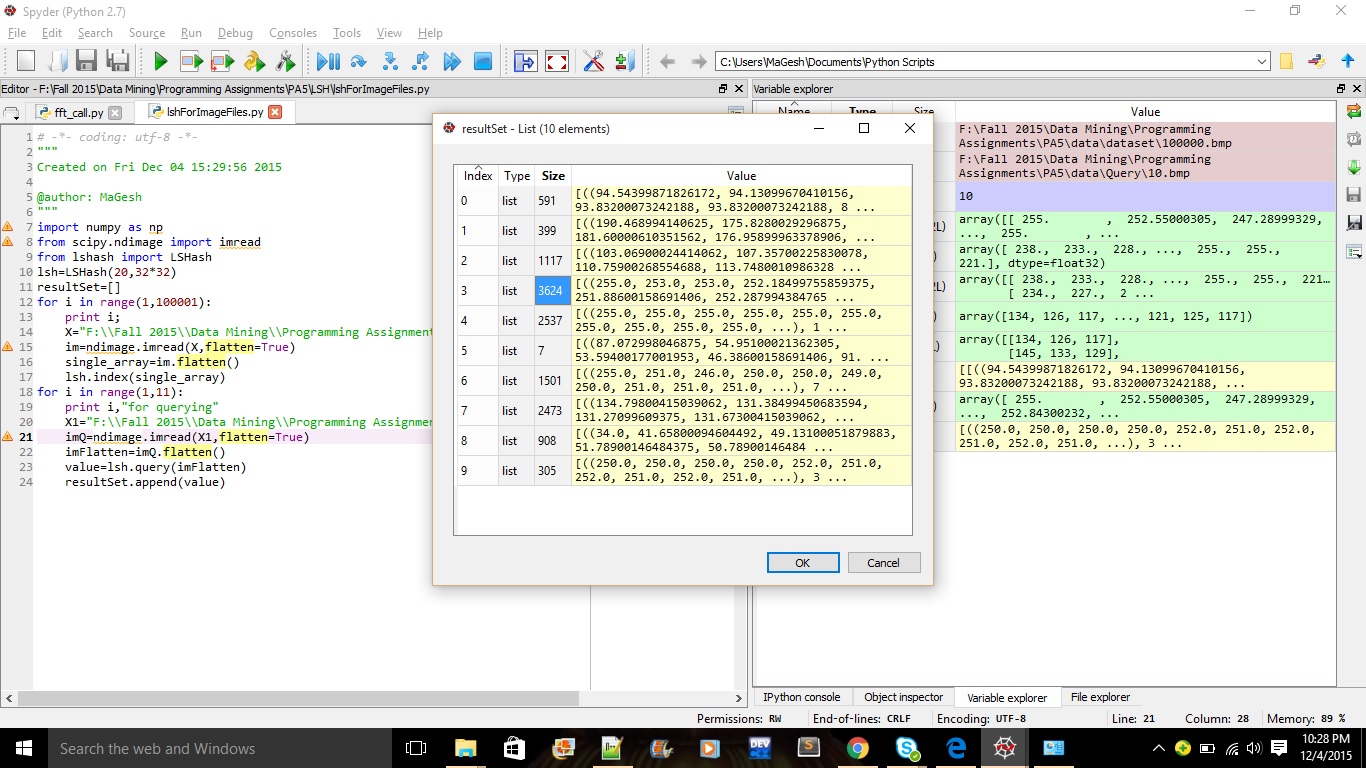
**CS521 Programming Assignment 4**

1. Attached is the image which will give the value of the query points since the resultSet is too huge I wasn’t able to write it to a file my program erred out so attaching the screen shot.



1. Parameters used:

* Reduced the size of three dimension to two by converting it to a grey scale image (32\*32\*3) has been changed to (32\*32).
* Used the hash bucket size of 20.
* Distance measure is Euclidean.

1. Python code:

imported <https://github.com/kayzh/LSHash> LSHash package (Github)

code:

# -\*- coding: utf-8 -\*-

"""

Created on Fri Dec 04 15:29:56 2015

@author: MaGesh

"""

import numpy as np

from scipy.ndimage import imread

from lshash import LSHash

lsh=LSHash(20,32\*32) #32\*32 is the dimension with 20 hash buckets

resultSet =[]

for i in range(1,100001):

print i;

X="F:\\Fall 2015\\Data Mining\\Programming Assignments\\PA5\\data\\dataset\\"+str(i)+".bmp"

im=imread(X,flatten=True)

single\_array=im.flatten()

lsh.index(single\_array)#hashing the each values in to the bucket

for i in range(1,11):

print i,"for querying"

X1="F:\\Fall 2015\\Data Mining\\Programming Assignments\\PA5\\data\\Query\\"+str(i)+".bmp"

imQ=imread(X1,flatten=True) #converting to grey scale

imFlatten=imQ.flatten()

value=lsh.query(imFlatten, distance\_func=”euclidean”) #querying the nearest points

resultSet.append(value)