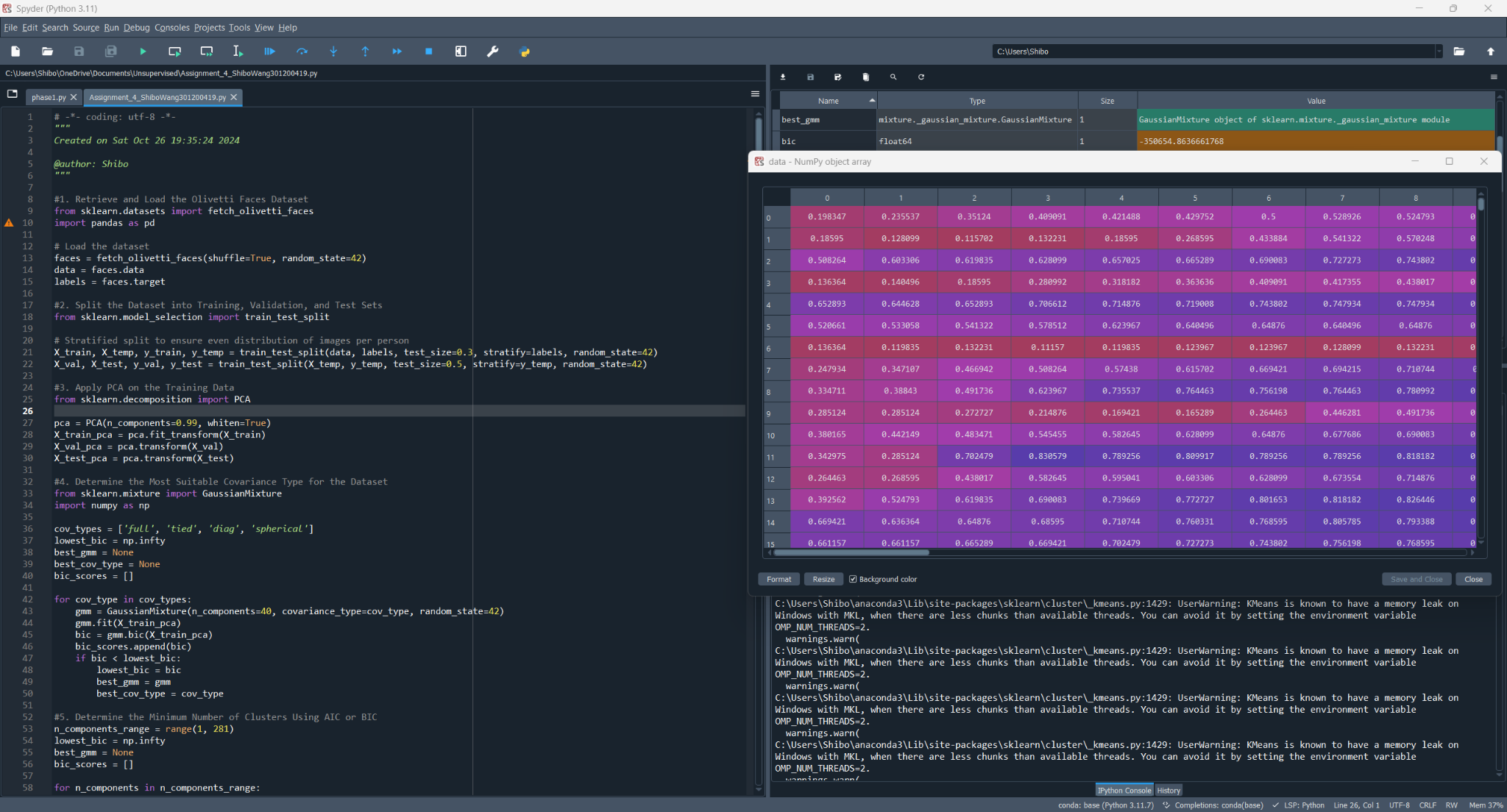
Assignment 4: Gaussian Mixture Models

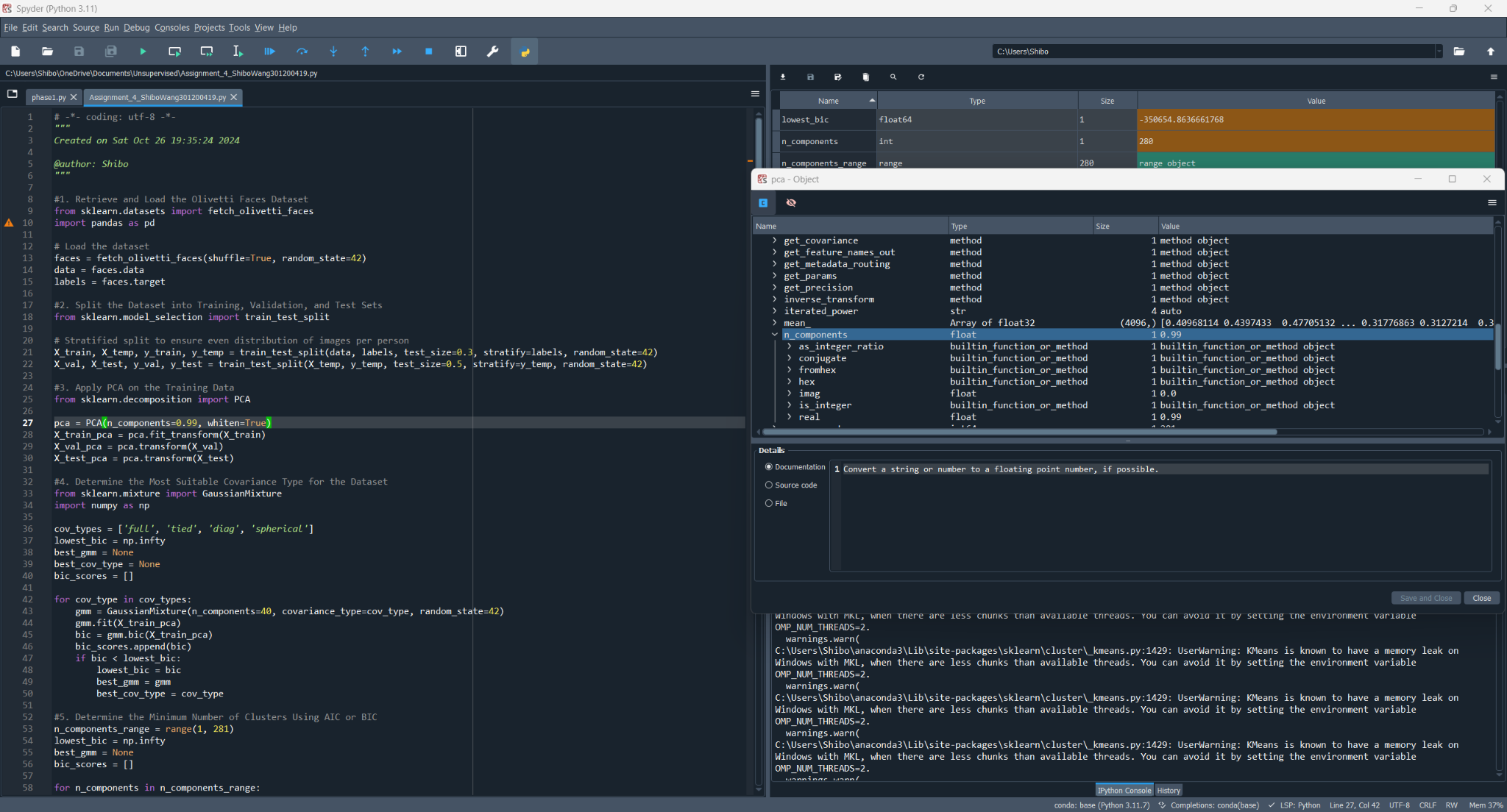
Written Report

1.Retrieve and load the Olivetti faces dataset. [0 points]

2.Split the dataset into training, validation, and test sets using stratified sampling to ensure that each set contains the same number of images per person. [0 points]



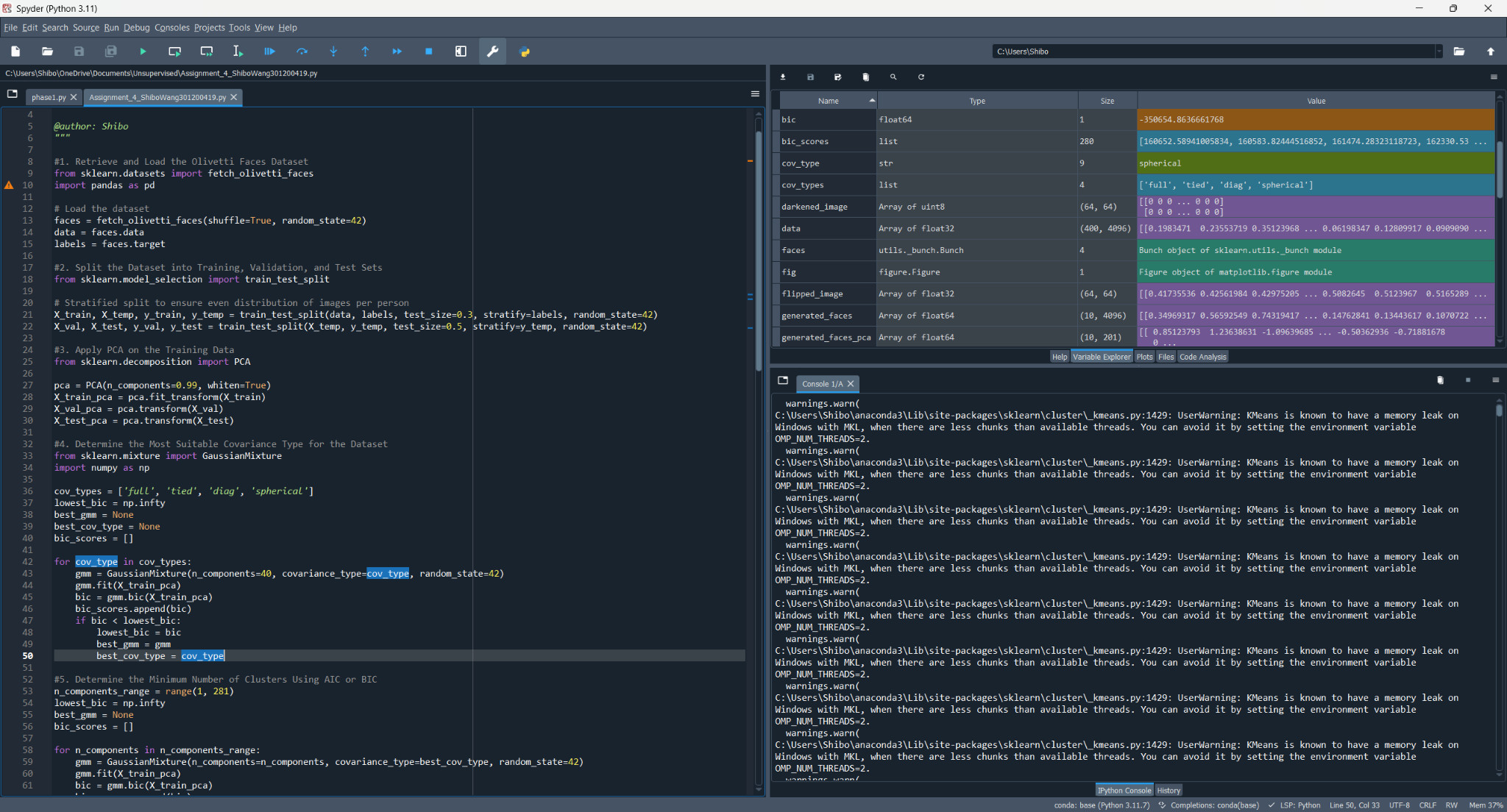
3.Apply PCA on the training data, preserving 99% of the variance, to reduce the dataset’s dimensionality.



4.Determine the most suitable covariance type for the dataset.

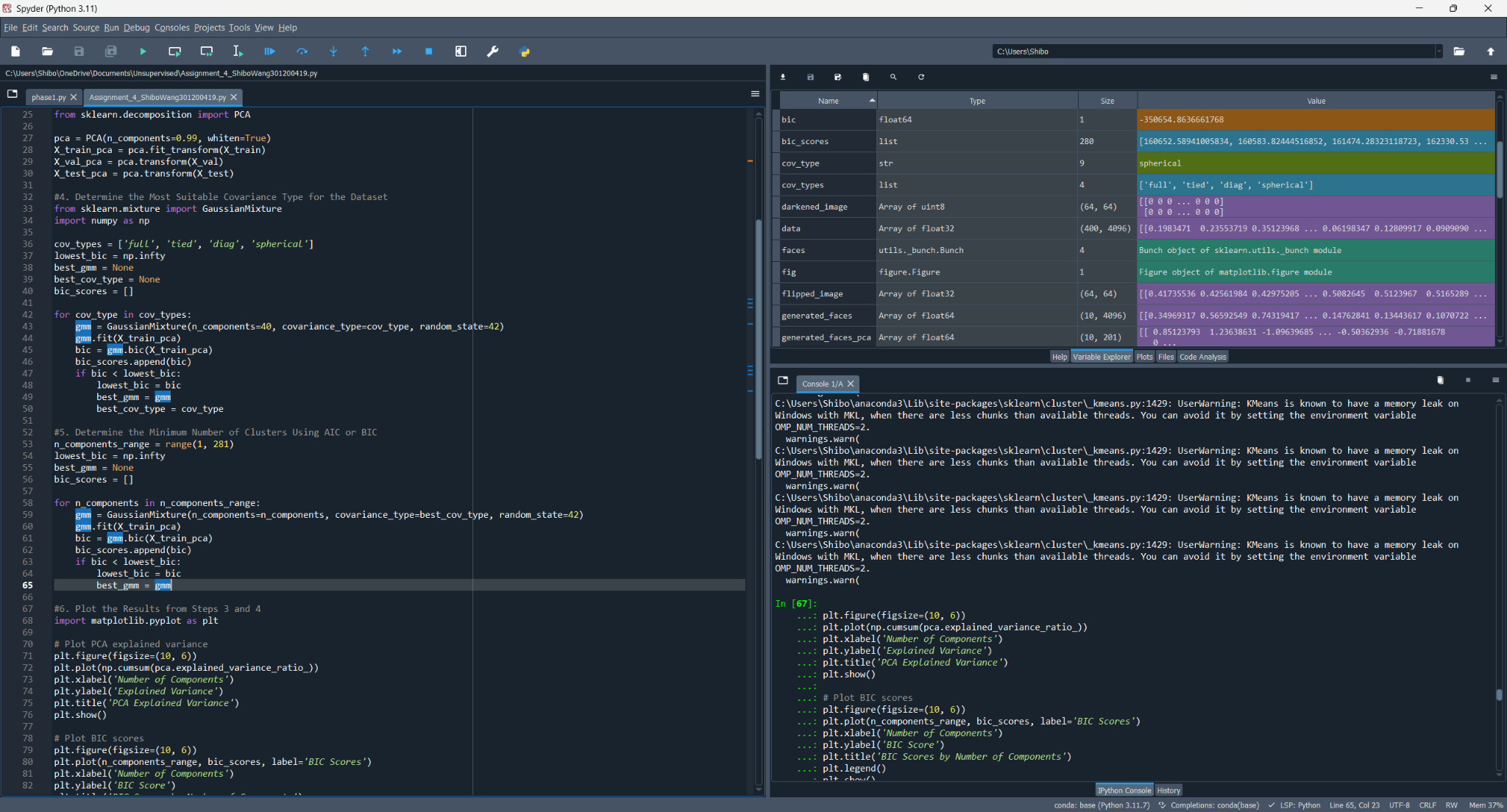
cov\_types = ['full', 'tied', 'diag', 'spherical']

The best one is 'spherical'.

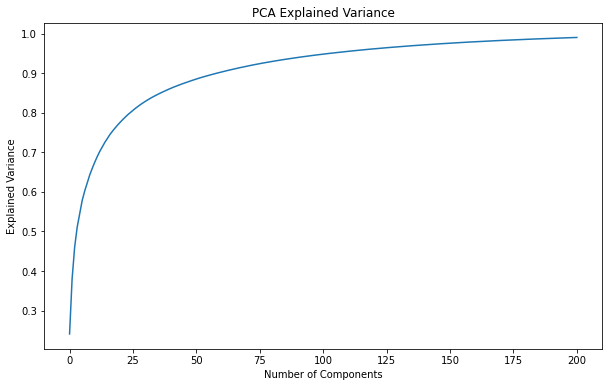


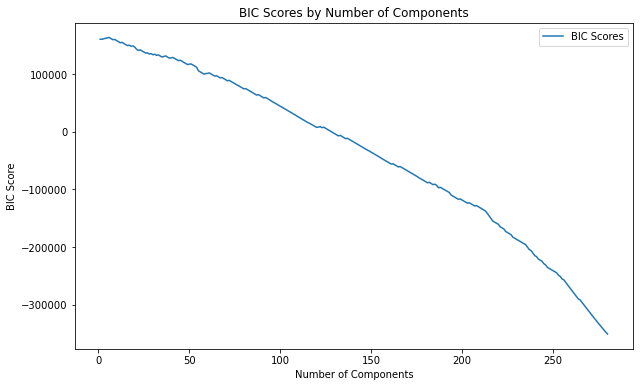
5.Determine the minimum number of clusters that best represent the dataset using either AIC or BIC.

I use BIC.



6.Plot the results from steps 3 and 4.



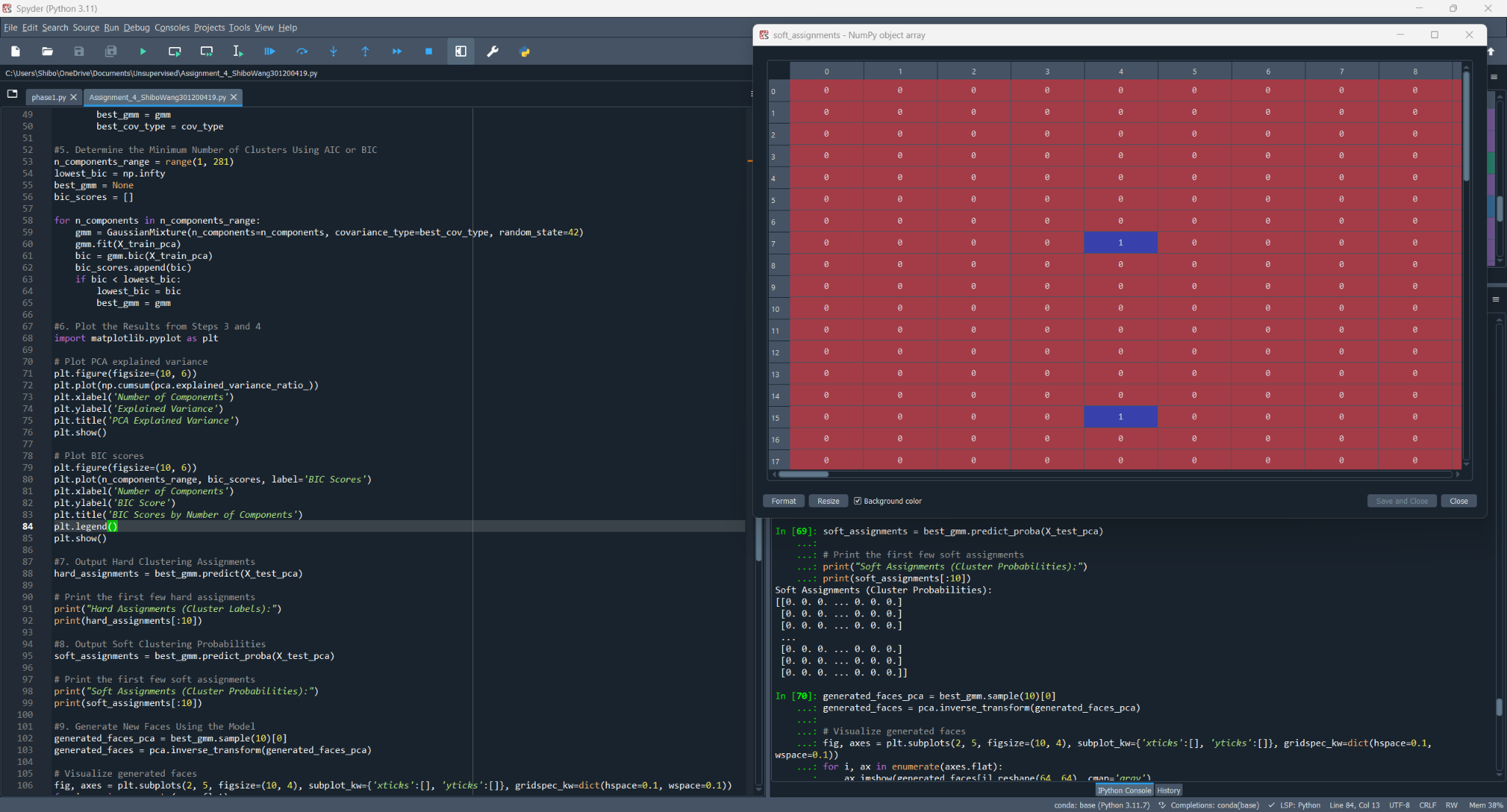


7.Output the hard clustering assignments for each instance to identify which cluster each image belongs to.

Hard Assignments (Cluster Labels):

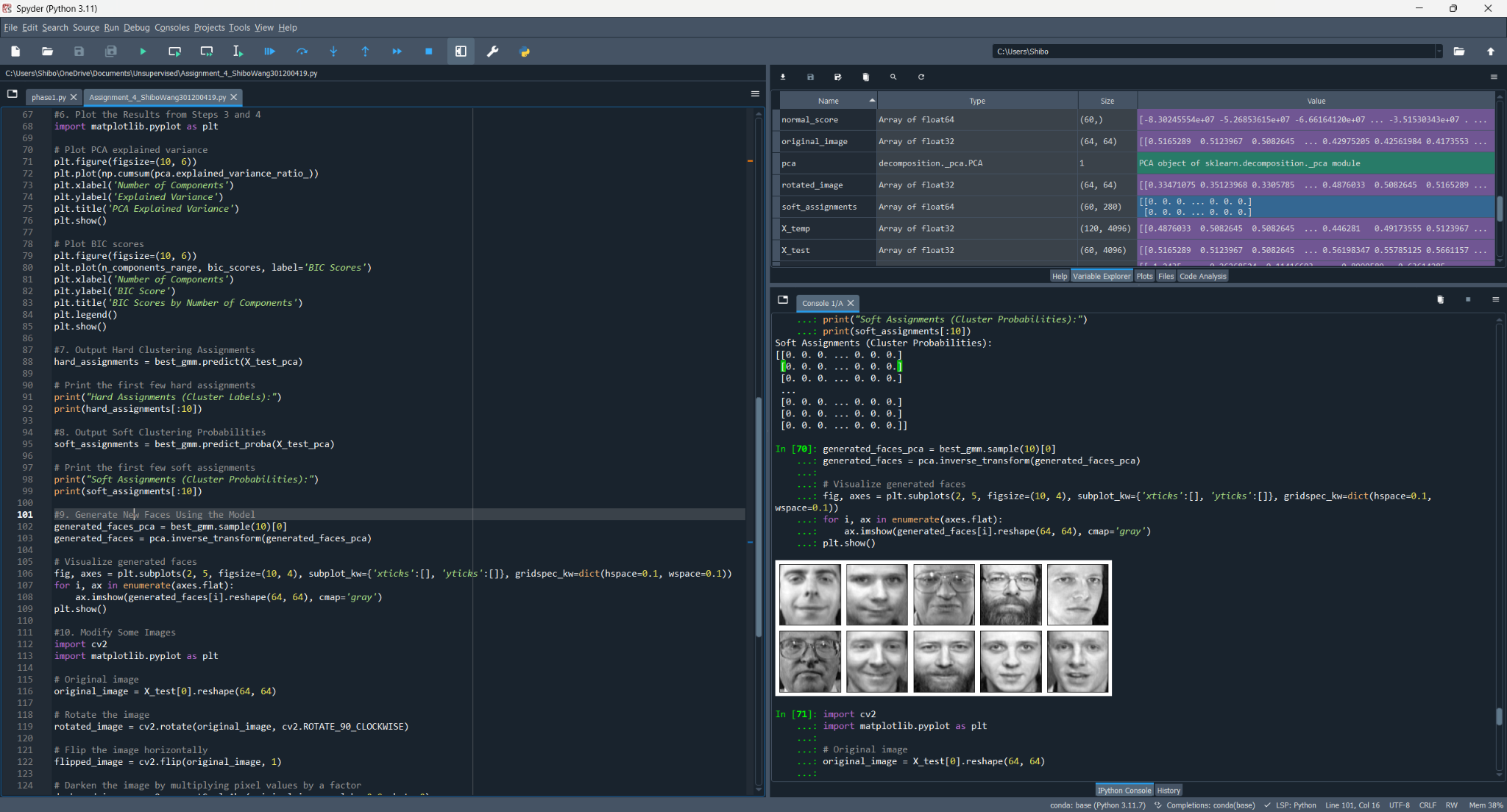
[ 25 257 242 274 192 189 50 4 129 257]

8.Output the soft clustering probabilities for each instance to show the likelihood of each image belonging to each cluster.

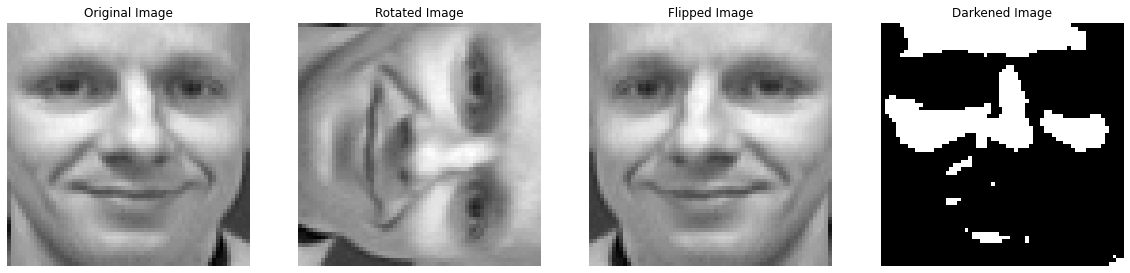


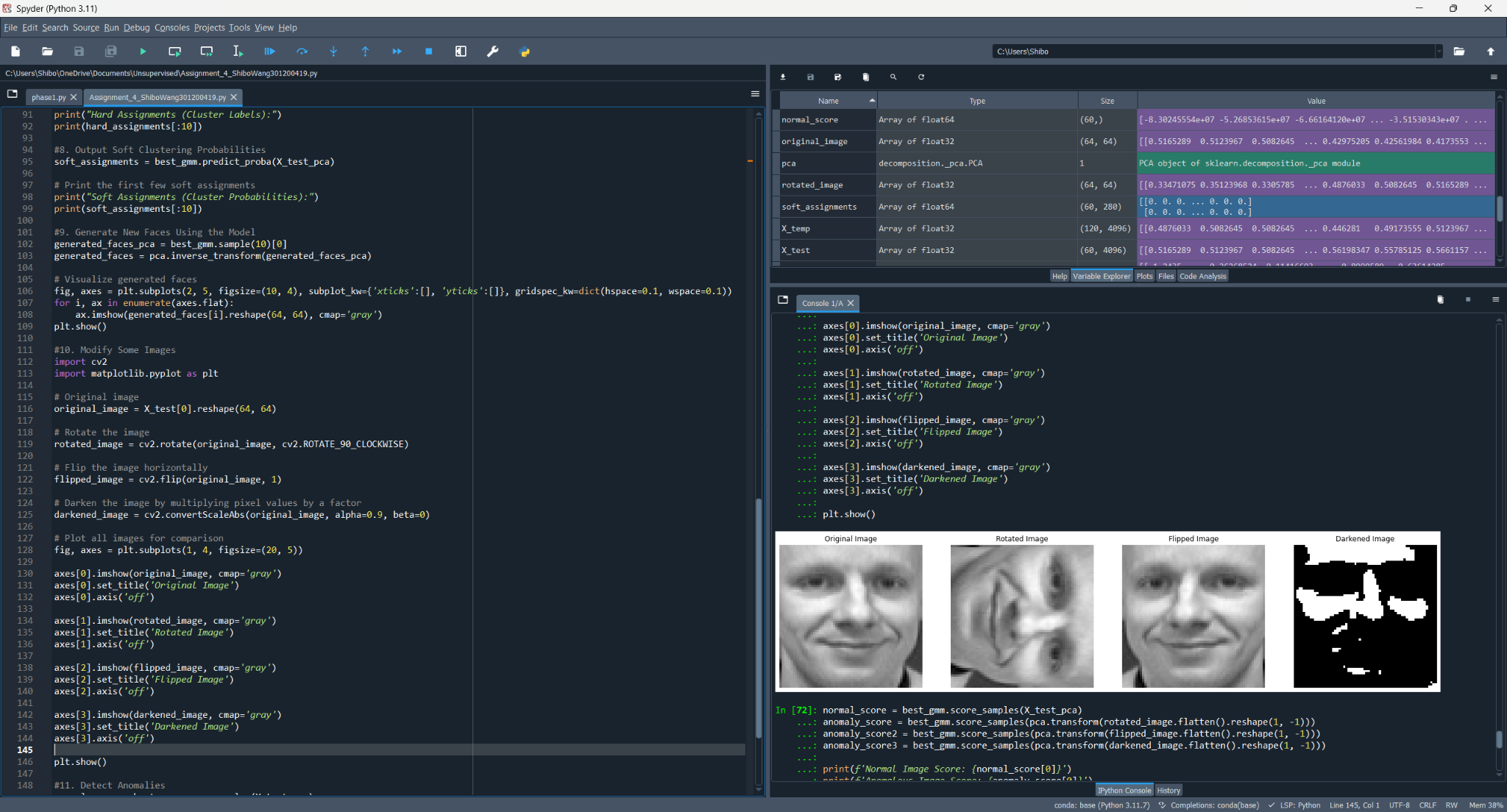
9. Use the model to generate some new faces (using the sample() method) and visualize them (use the inverse\_transform() method to transform the data back to its original space based on the PCA method used).





10.Modify some images





11.Determine if the model can detect the anomalies produced in step 10 by comparing the output of the score\_samples() method for normal images and for anomalies.

Normal Image Score: -83024555.39749917

This is the log-likelihood of a normal image from your test set. The higher the score, the better the image fits the model.

Anomalous Image Score: -181061625.74015063

This significantly lower score suggests that the rotated image is far less likely according to the GMM.

Anomalous Image2 Score: -60527820.848939136

This lower score compared to the normal image score indicates that the flipped image is also less likely but not as extreme as the rotated image.

Anomalous Image3 Score: -682107899.0173904

This extremely low score indicates that the darkened image is the least likely of all.

