**CS 5710**

**MACHINE LEARNING**

**Name:** Shiva Kandagatla

**Student Id:** 700745245

**E-mail:** SXK54250@ucmo.edu

**Course:** CS 5710

**Assignment:** Assignment #5

**GitHub Link:**

**https://github.com/Shiva-Kandagatla98/In-Class-Assignment-5.git**

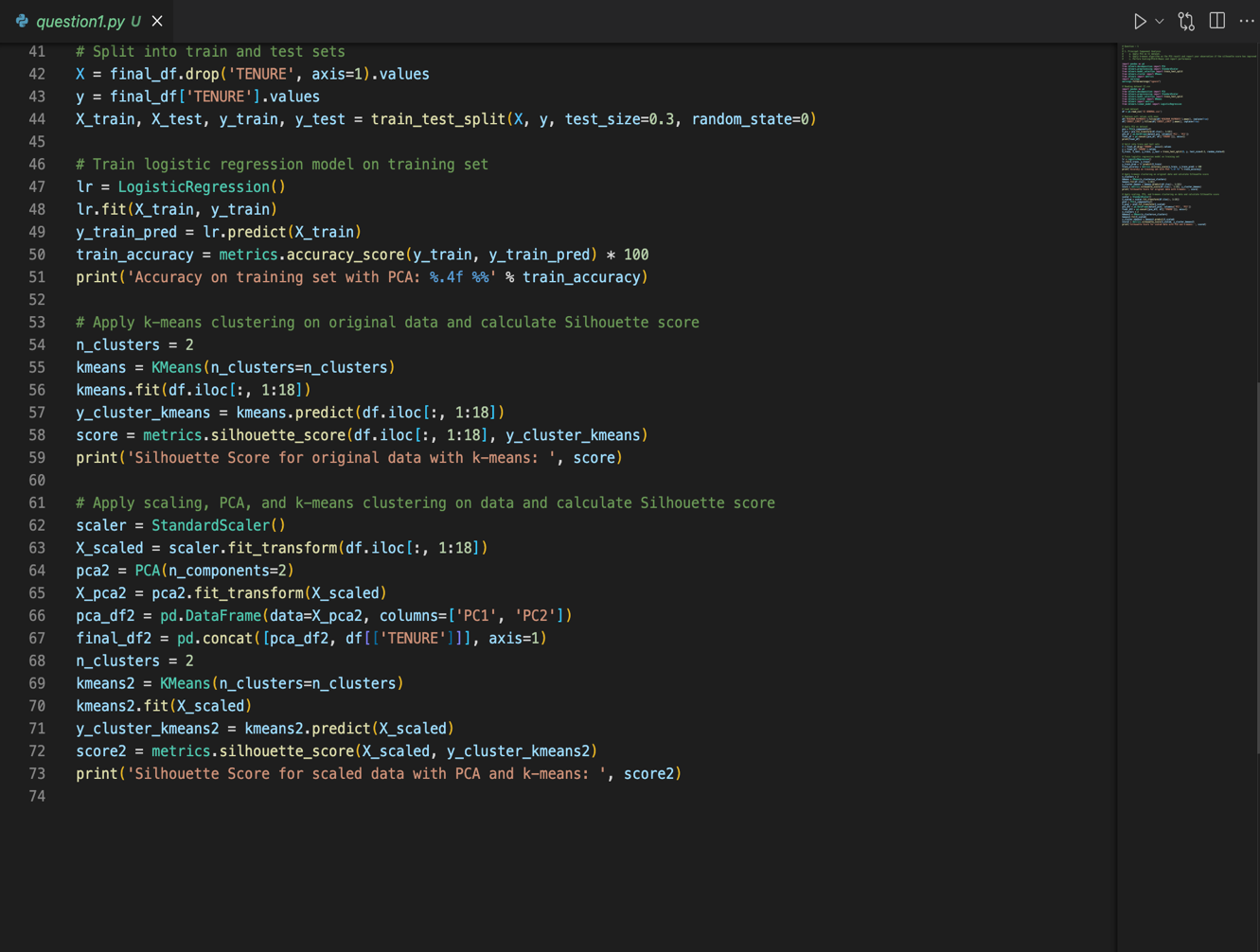
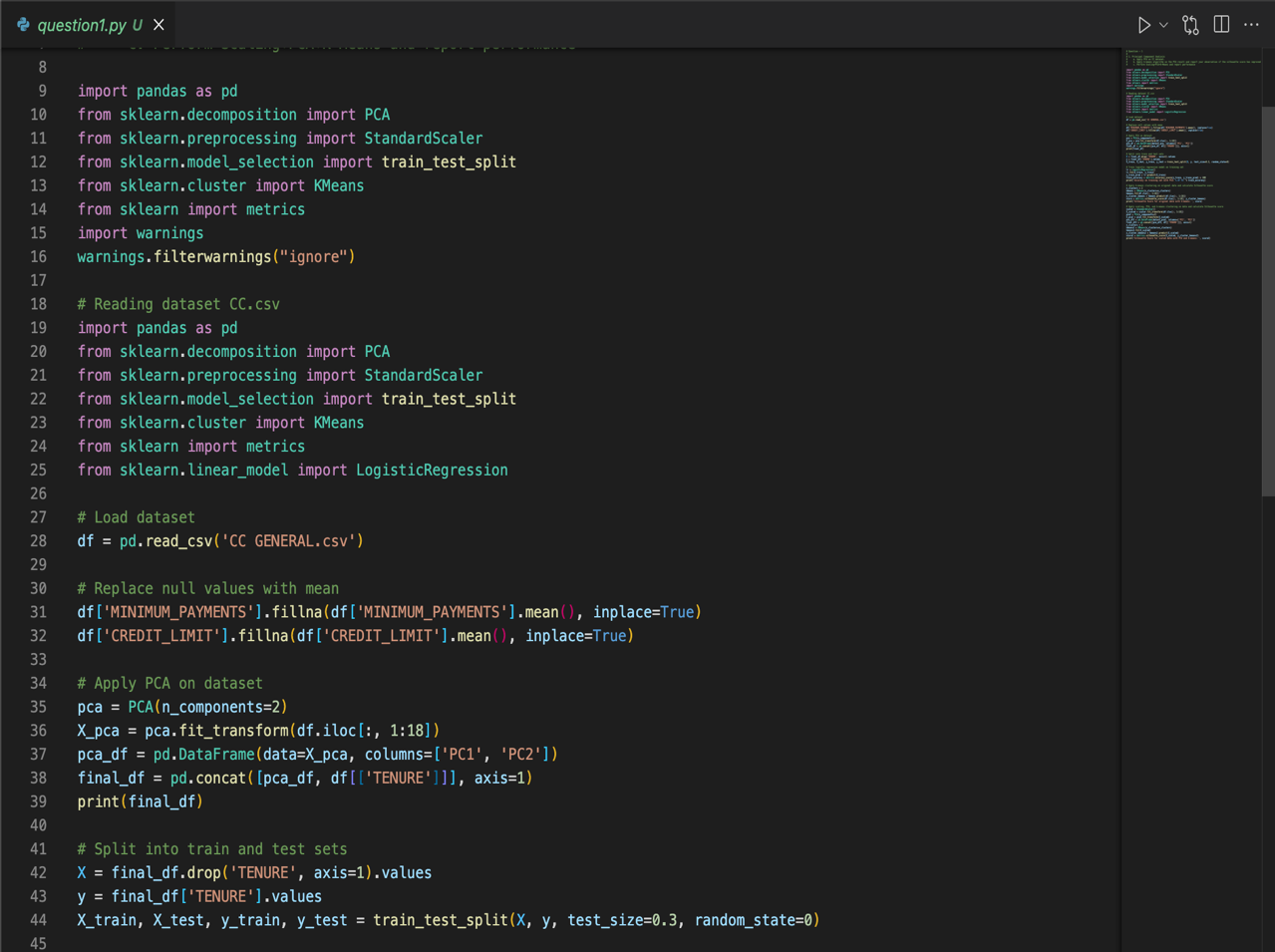
**Video Demo Link:**

**1. Principal Component Analysis**

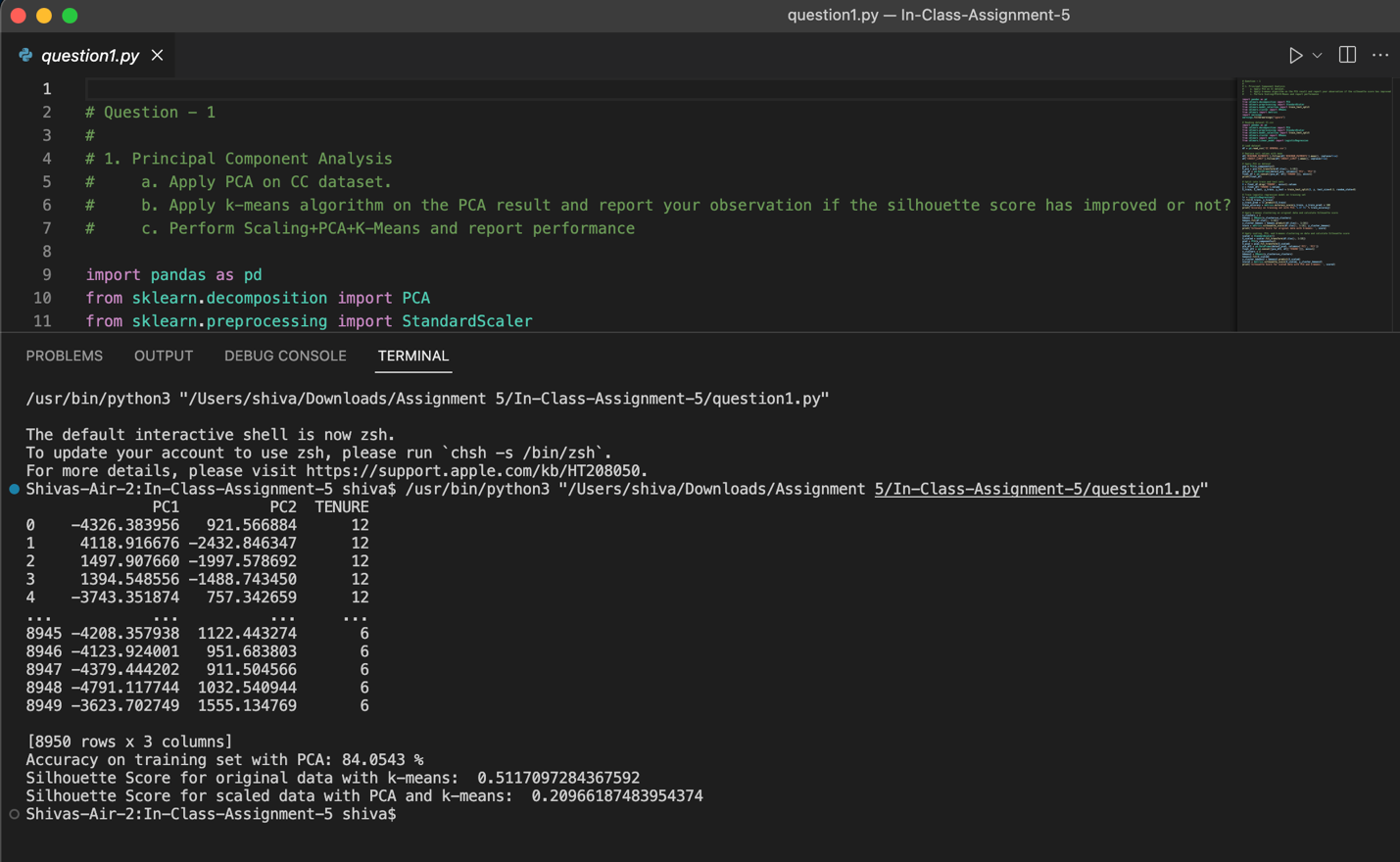
**a. Apply PCA on CC dataset.**

**b. Apply k-means algorithm on the PCA result and report your observation if the silhouette score has improved or not?**

**c. Perform Scaling+PCA+K-Means and report performance.**

**Source Code:**

**Output:**

****

**----------------\*\*\*\*\*\*\*----------------**

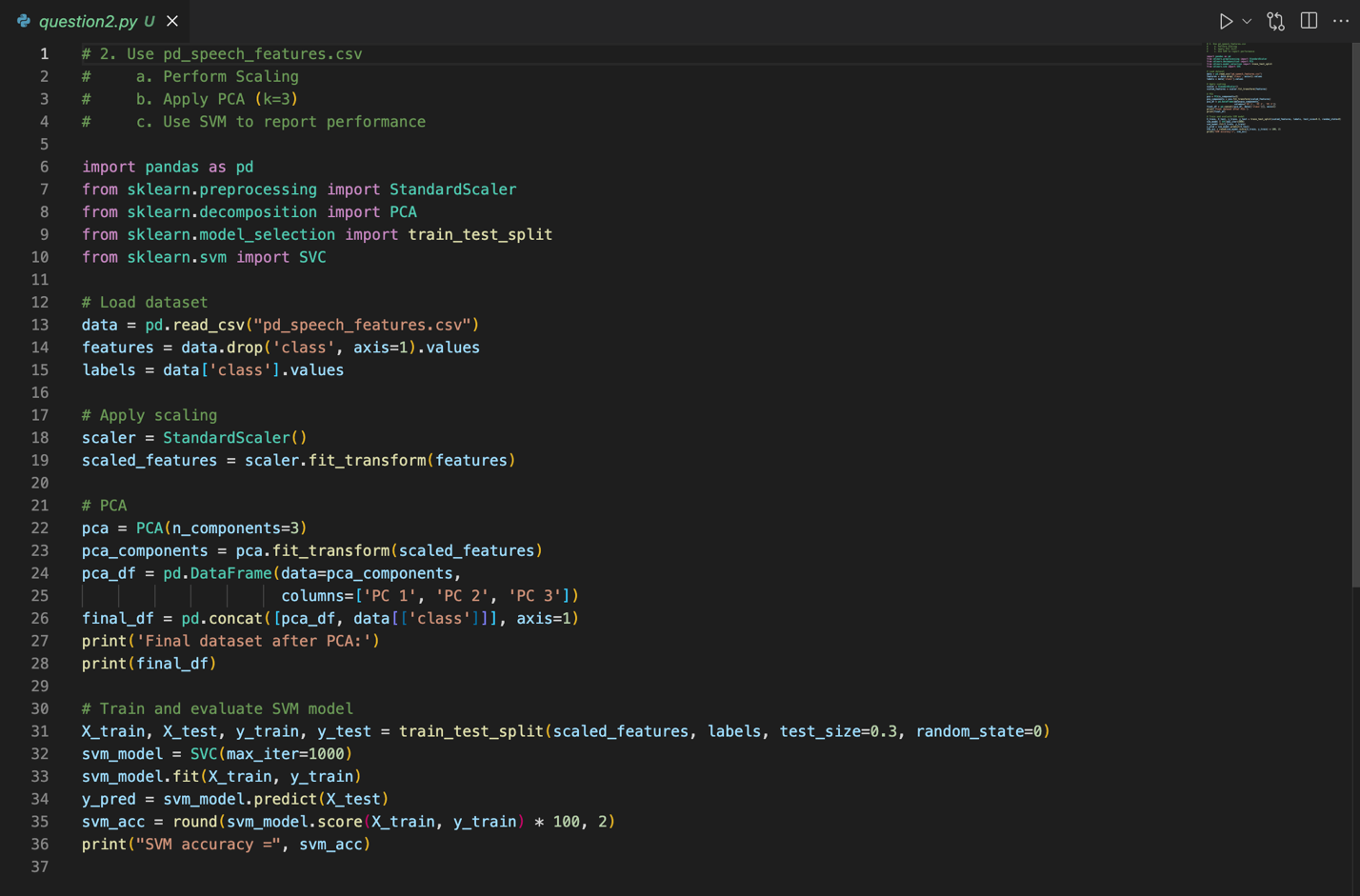
**2. Use pd\_speech\_features.csv**

**a. Perform Scaling**

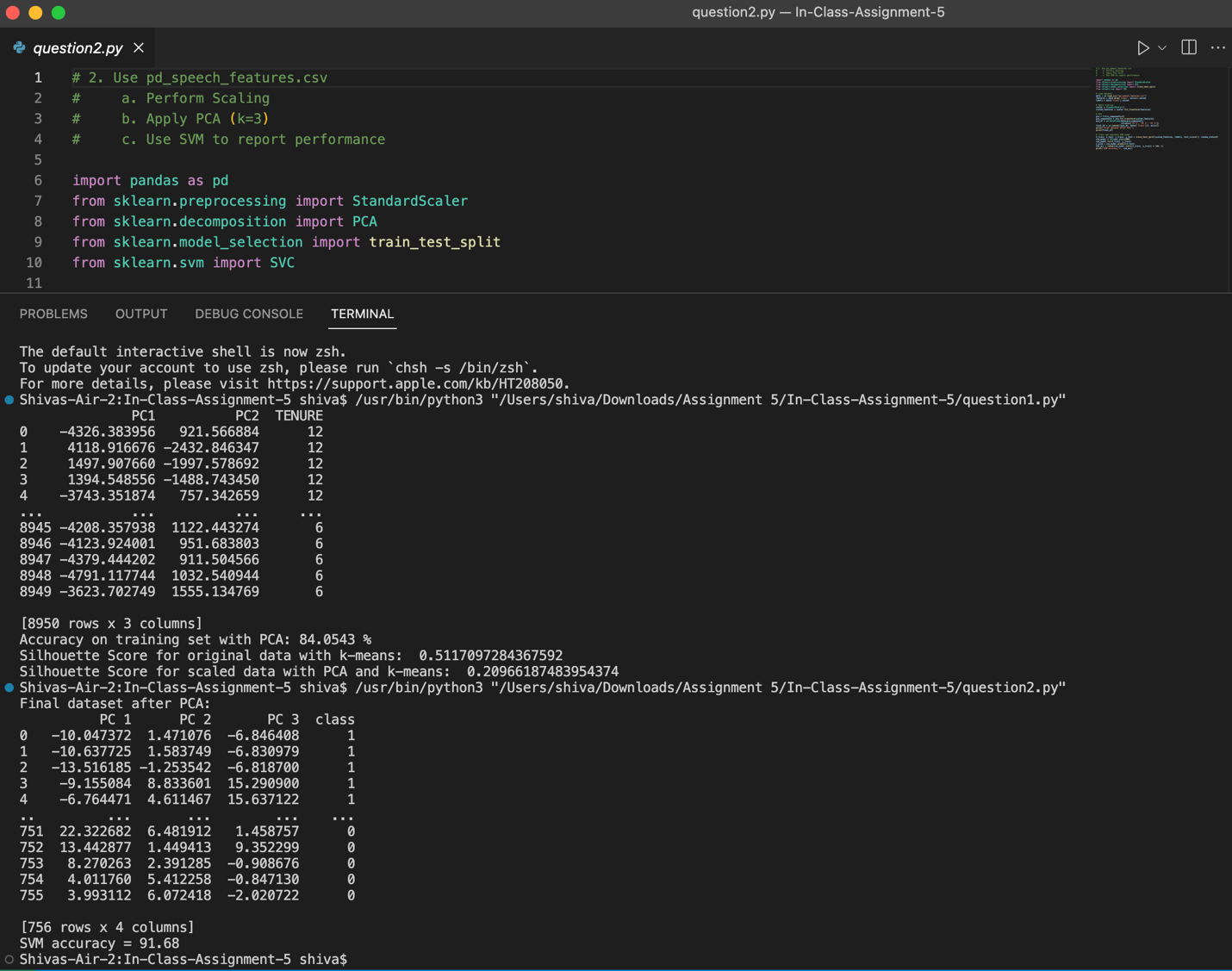
**b. Apply PCA (k=3)**

**c. Use SVM to report performance**

**Source Code:**

****

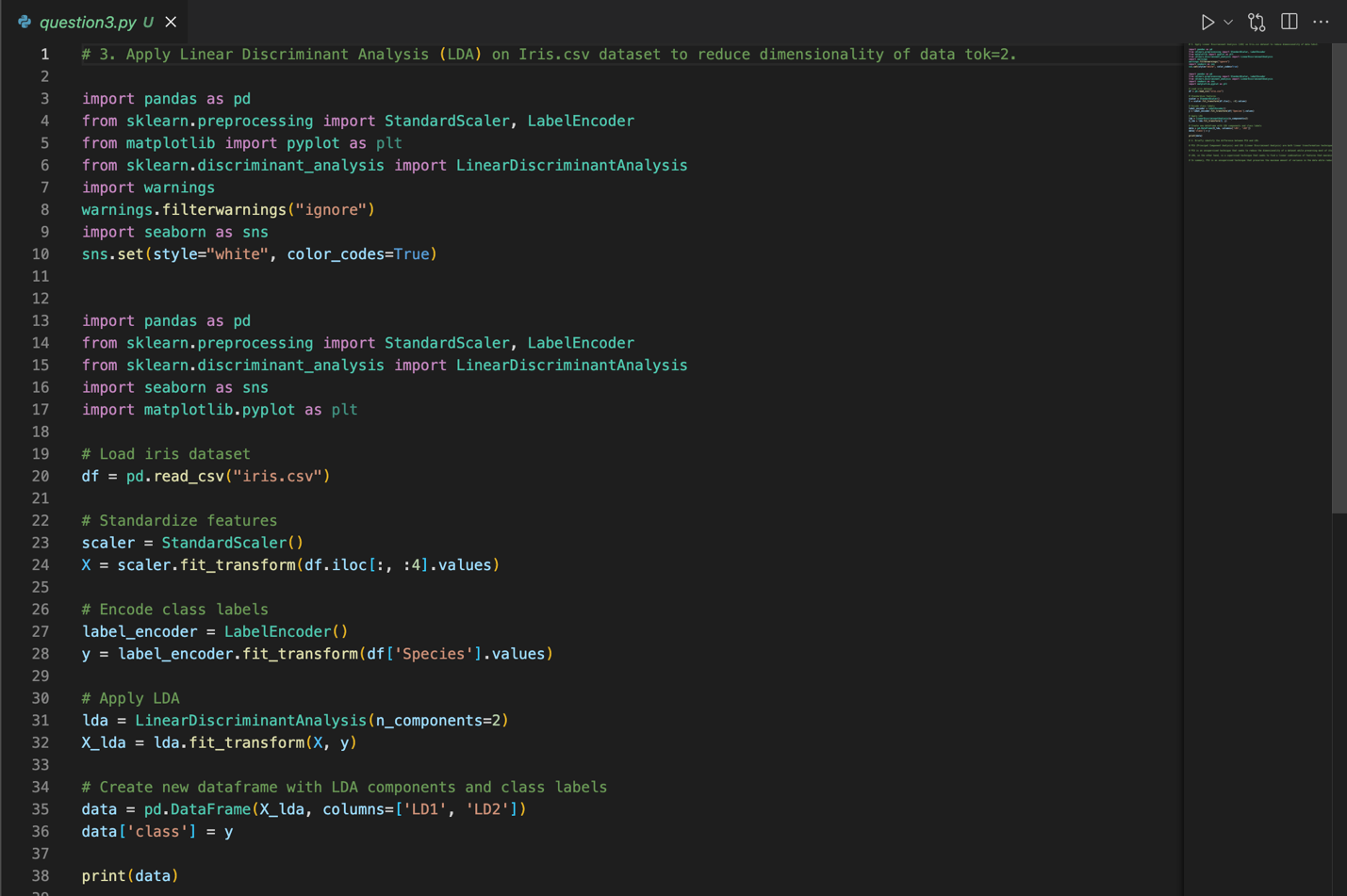
**Output:**

****

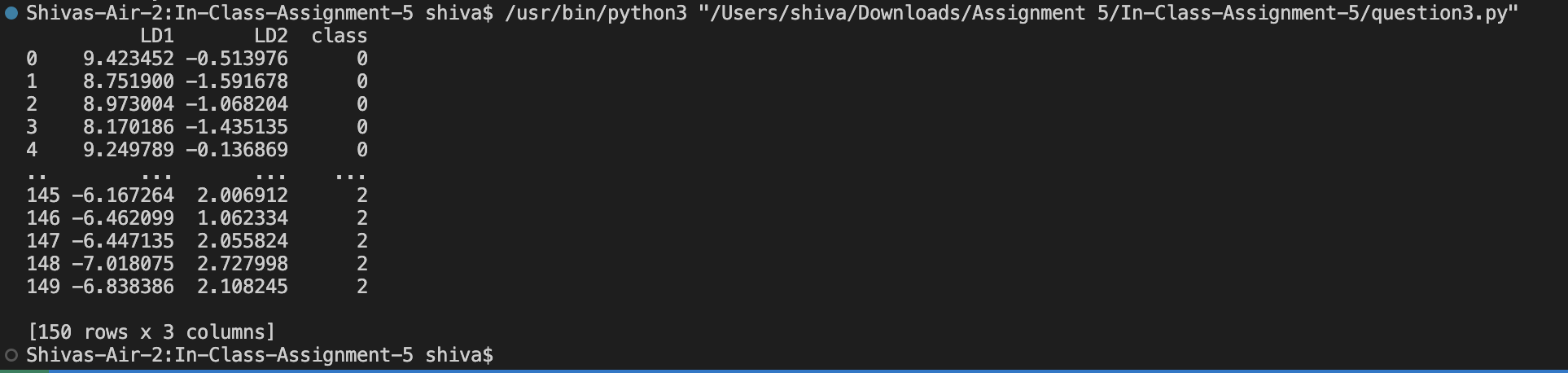
**----------------\*\*\*\*\*\*\*----------------**

**3. Apply Linear Discriminant Analysis (LDA) on Iris.csv dataset to reduce dimensionality of data to k=2.**

**Source Code:**

****

**Output:**

****

**----------------\*\*\*\*\*\*\*----------------**

**4. Briefly identify the difference between PCA and LDA**

PCA is an unsupervised technique that seeks to reduce the dimensionality of a dataset while preserving most of its variance. It does so by transforming the data into a new coordinate system where the new axes (principal components) are orthogonal and ordered by the amount of variance they capture. PCA is often used for data visualization, data compression, and feature extraction.

LDA, on the other hand, is a supervised technique that seeks to find a linear combination of features that maximizes the separation between different classes of data. It does so by finding a projection that maximizes the between-class scatter and minimizes the within-class scatter. LDA is often used for classification and feature extraction.

In summary, PCA is an unsupervised technique that preserves the maximum amount of variance in the data while reducing dimensionality, while LDA is a supervised technique that seeks to maximize the separation between classes by finding a projection that preserves the class structure of the data.

**----------------\*\*\*\*\*\*\*----------------**