Machine learning-Assignment5

Student id:700674089

Student Name: Jahnavi Danda

CRN: CS571013466

**GitHub Repository**

https://github.com/jxd40890/Machine-Learning-Assignment5.git

**Assignment5 video URL**

https://drive.google.com/drive/folders/1LyMLzmYnxdxPJqGzRUIP7ftS\_paDd6Tf?usp=sharing

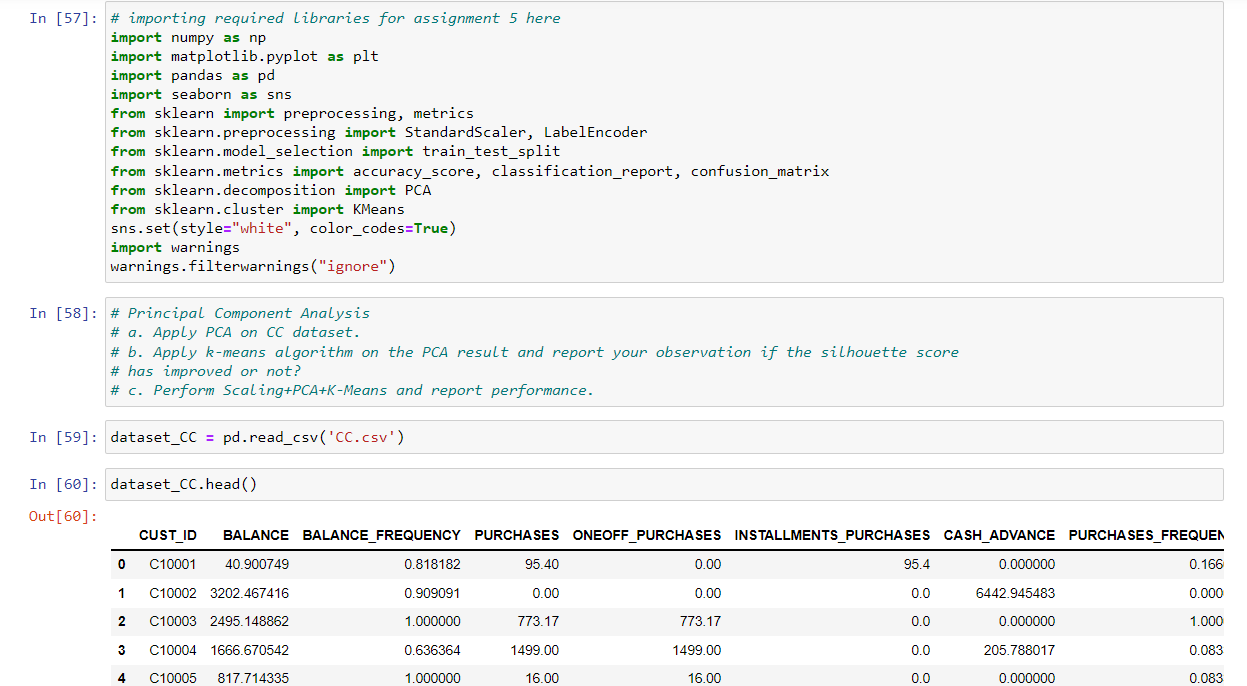
Question1

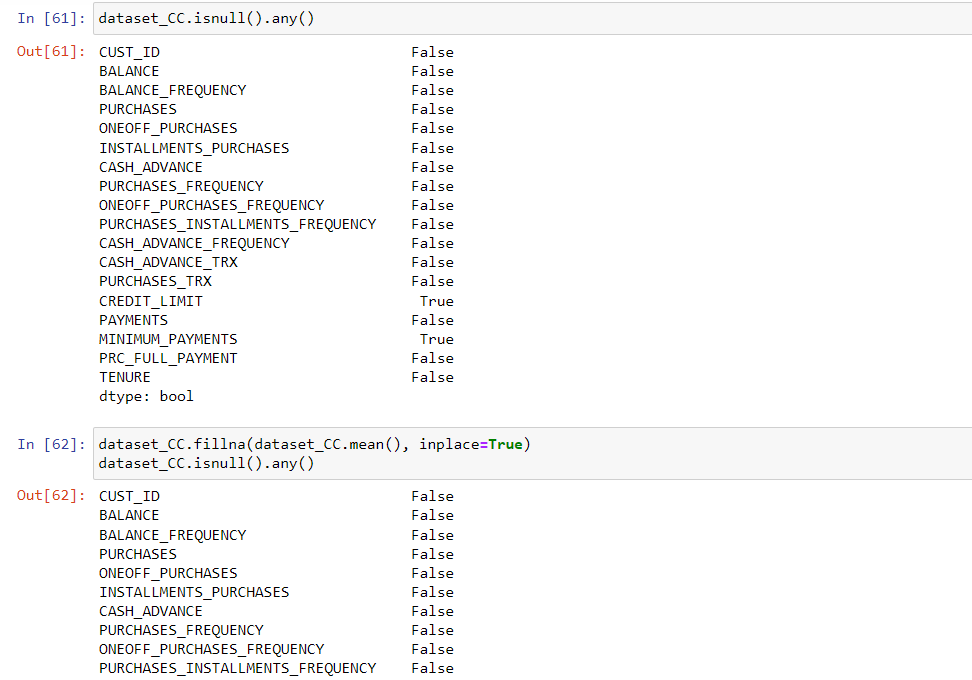
1. Principal Component Analysis

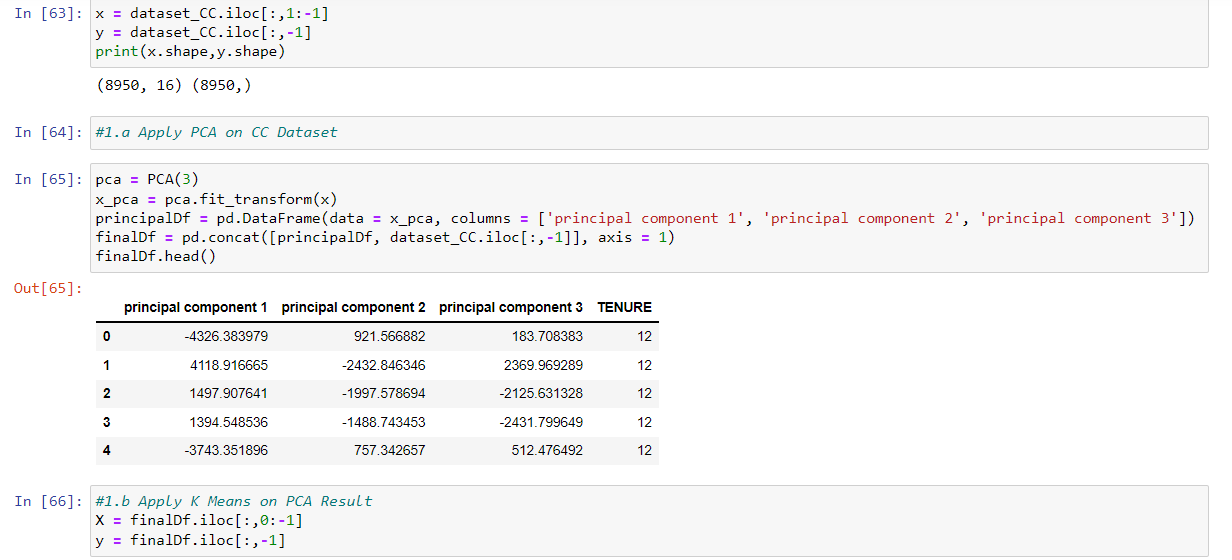
a. Apply PCA on CC dataset.

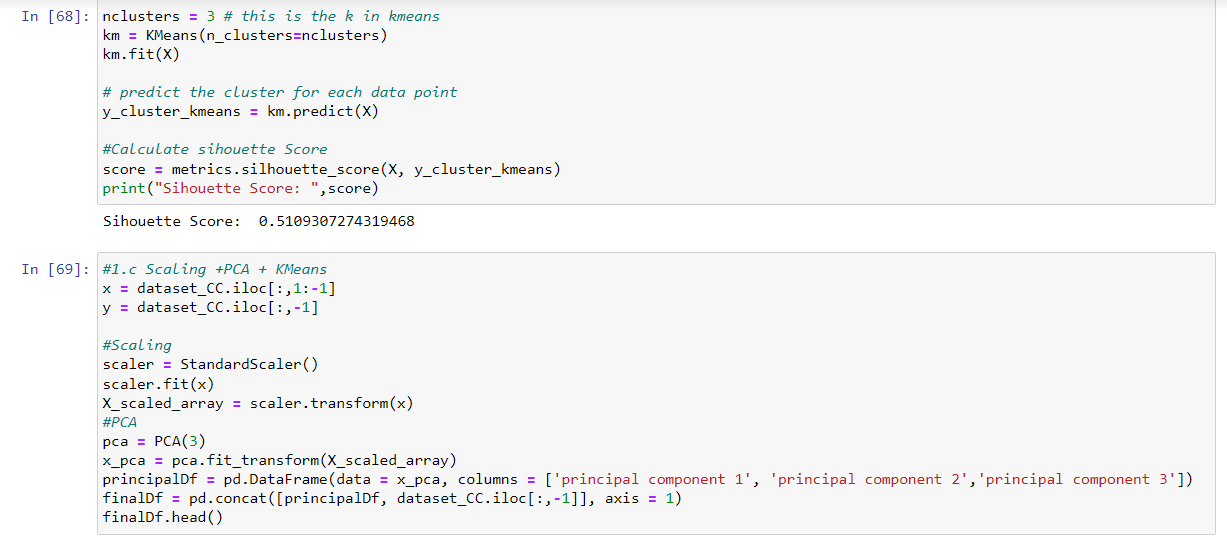
b. Apply k-means algorithm on the PCA result and report your observation if the silhouette score has improved or not?

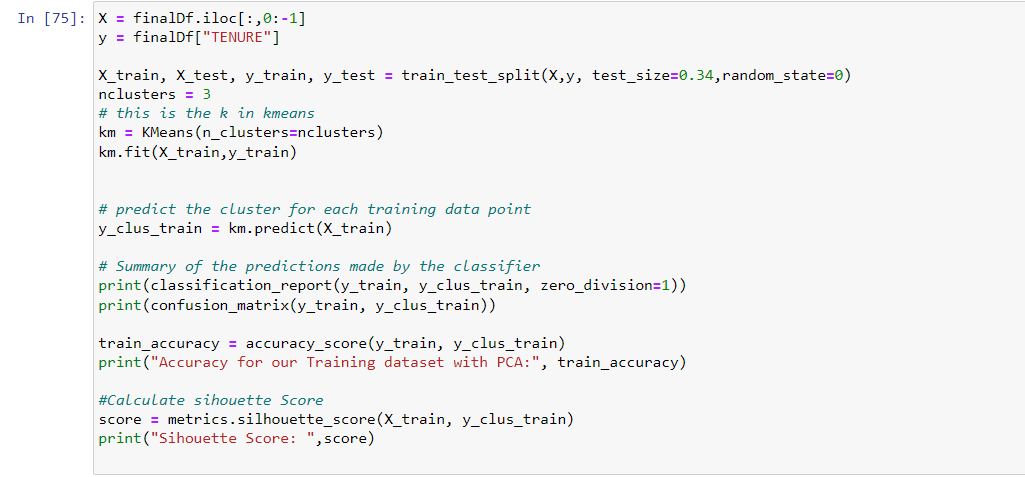
c. Perform Scaling+PCA+K-Means and report performance.

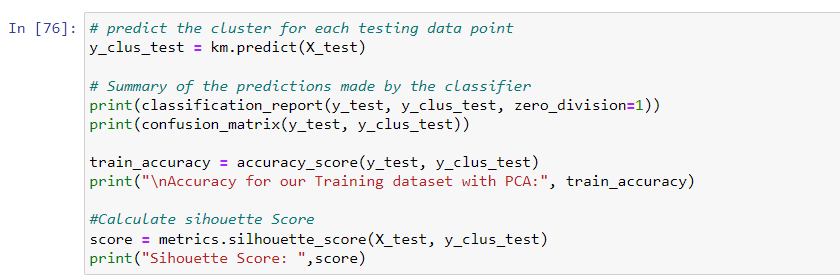






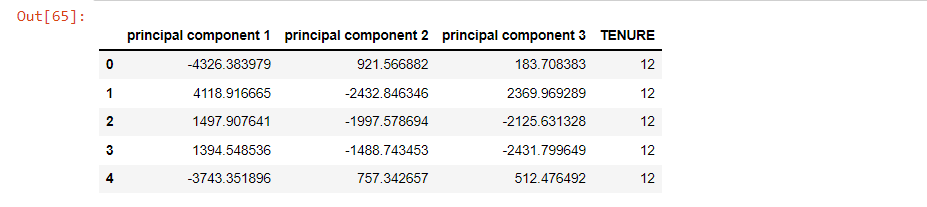




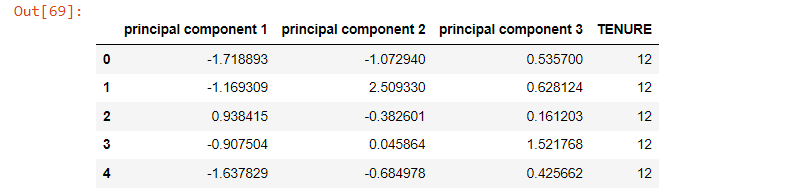


Output:

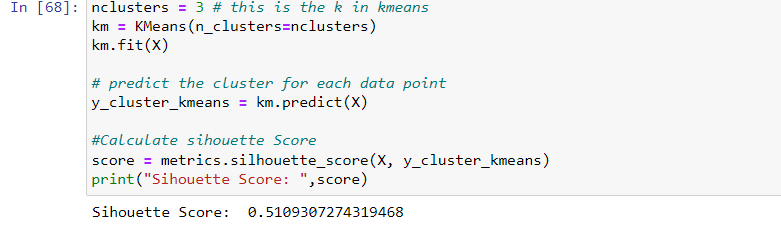
**Before scaling**



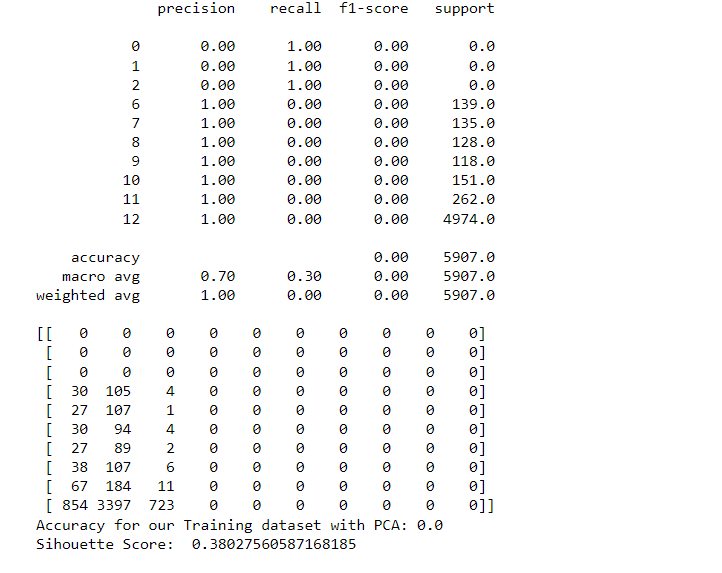
**After scaling**

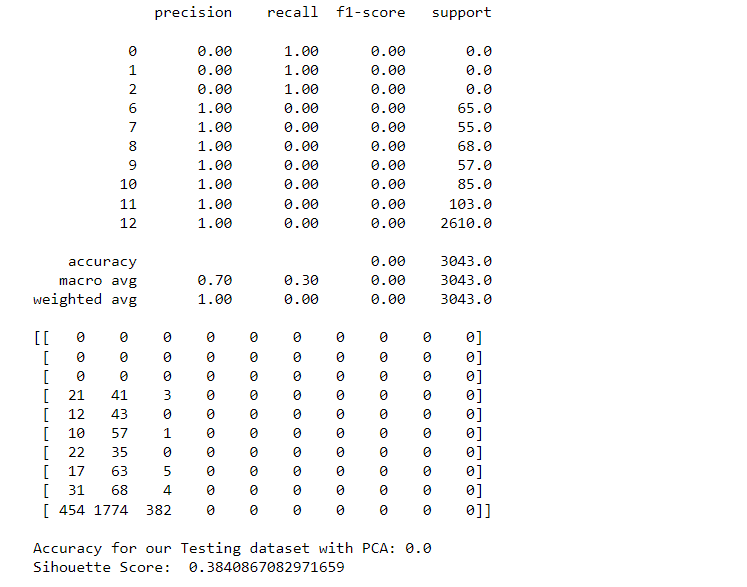


**Silhouette score before scaling:**

****

**Silhouette score after scaling**

****

****

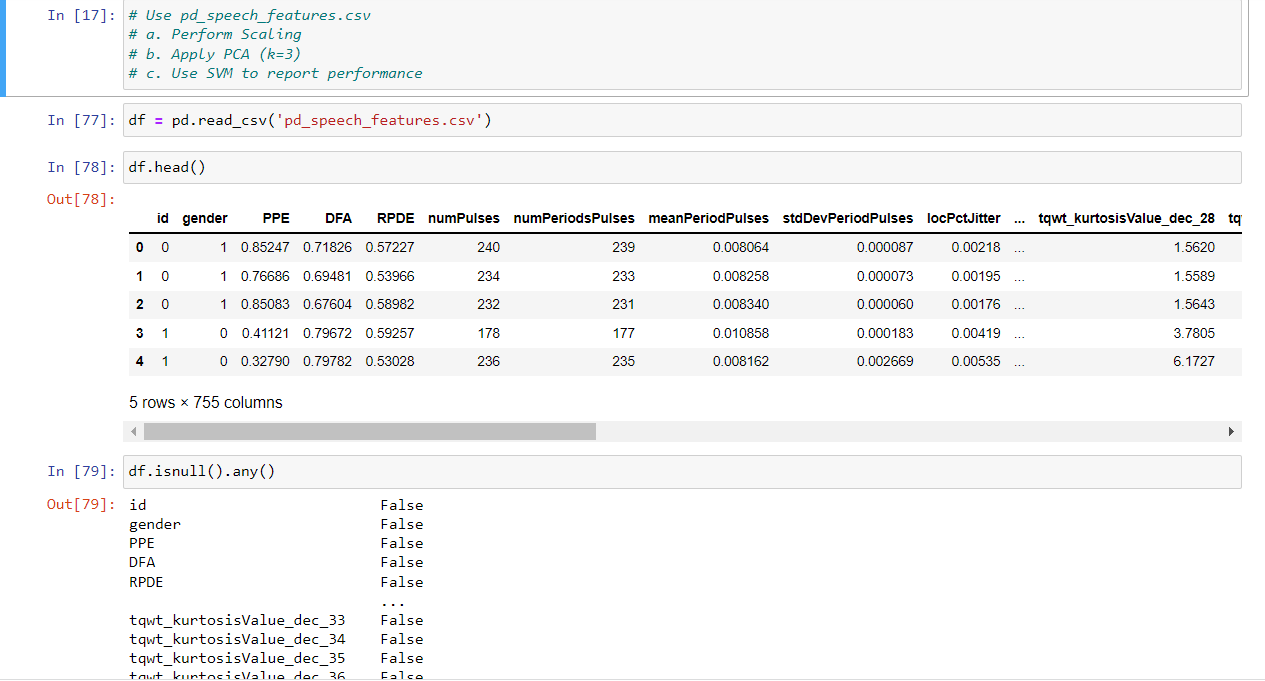
Question2: -

2. Use pd\_speech\_features.csv

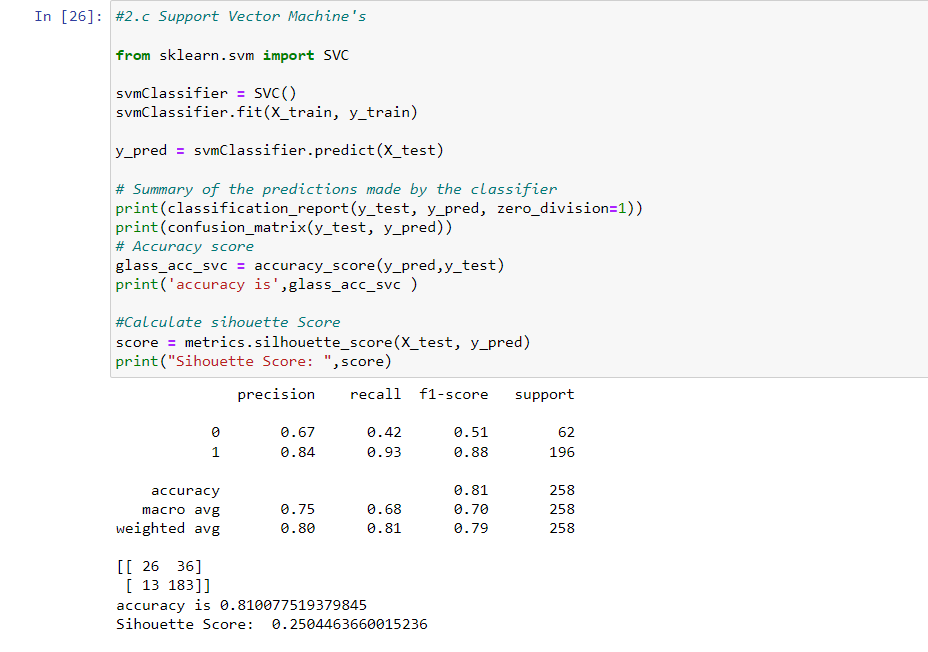
a. Perform Scaling

b. Apply PCA (k=3)

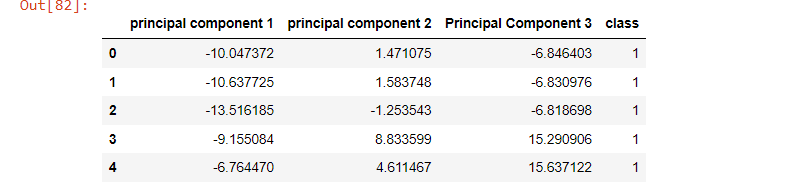
c. Use SVM to report performance

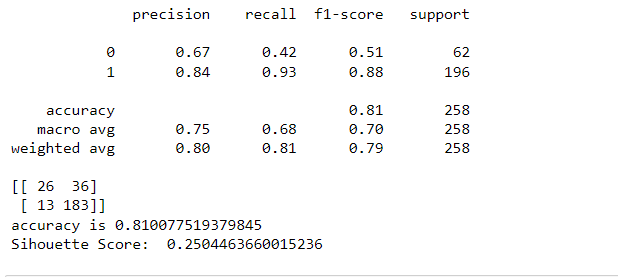






Output

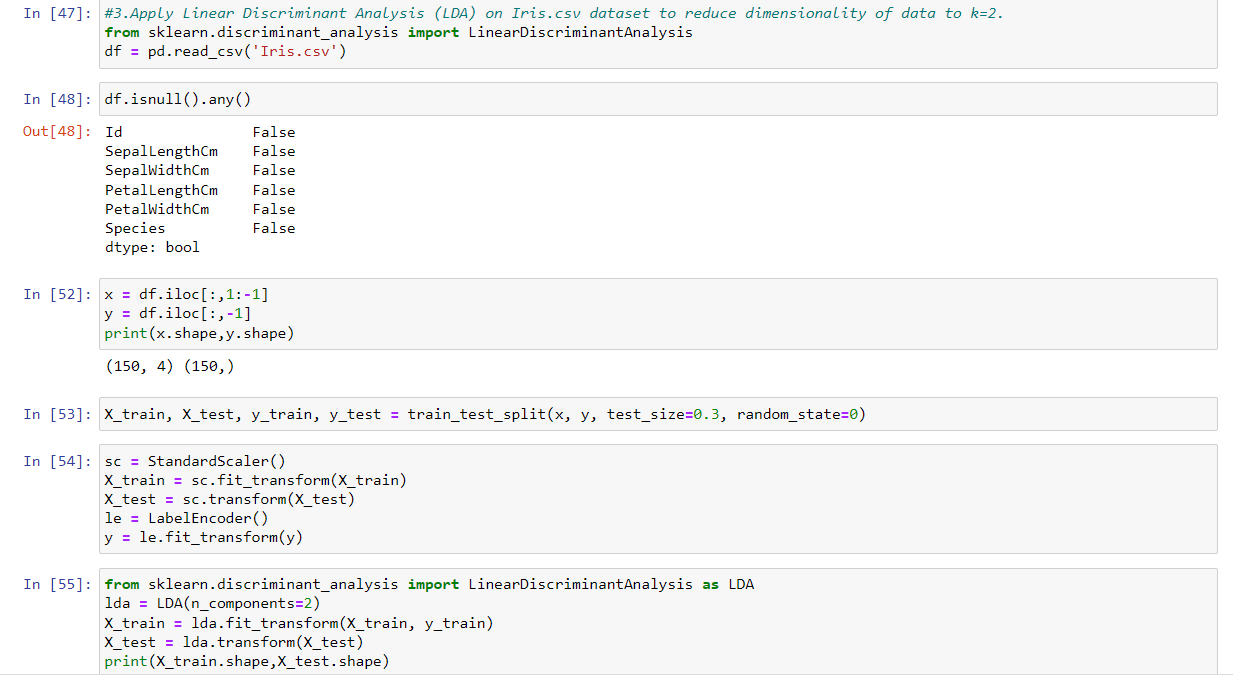




Question3: -

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

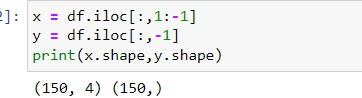
**Code Screenshot**



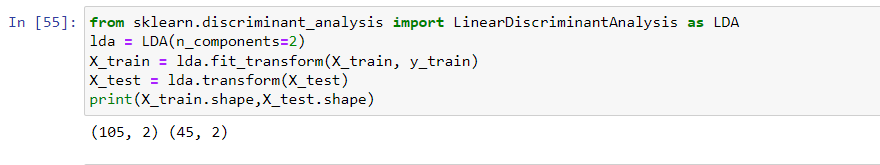


**Output Screenshot:**

Dimension before applying LDA



Dimension After applying LDA



**Question 4:**

Briefly identify the difference between PCA and LDA?

**Answer:**

PCA is an unsupervised learning algorithm while LDA is a supervised learning algorithm. This means that PCA finds directions of maximum variance regardless of class labels while LDA finds directions of maximum class separability.

PCA ignores class labels and focuses on finding the principal components that maximize the variance in a given data. Thus, it is an unsupervised algorithm. On the other hand, LDA is a supervised algorithm that intends to find the linear discriminants that represent those axes that maximize separation between different classes.

LDA performs better multi-class classification tasks than PCA. However, PCA performs better when the sample size is comparatively small. An example would be comparisons between classification accuracies that are used in image classification.

Both LDA and PCA are used in the case of dimensionality reduction. PCA is first followed by LDA.