**Assignment -05**

**Github link:**

[**https://github.com/GottiparthiShreshta/Assignment\_05/tree/master**](https://github.com/GottiparthiShreshta/Assignment_05/tree/master)

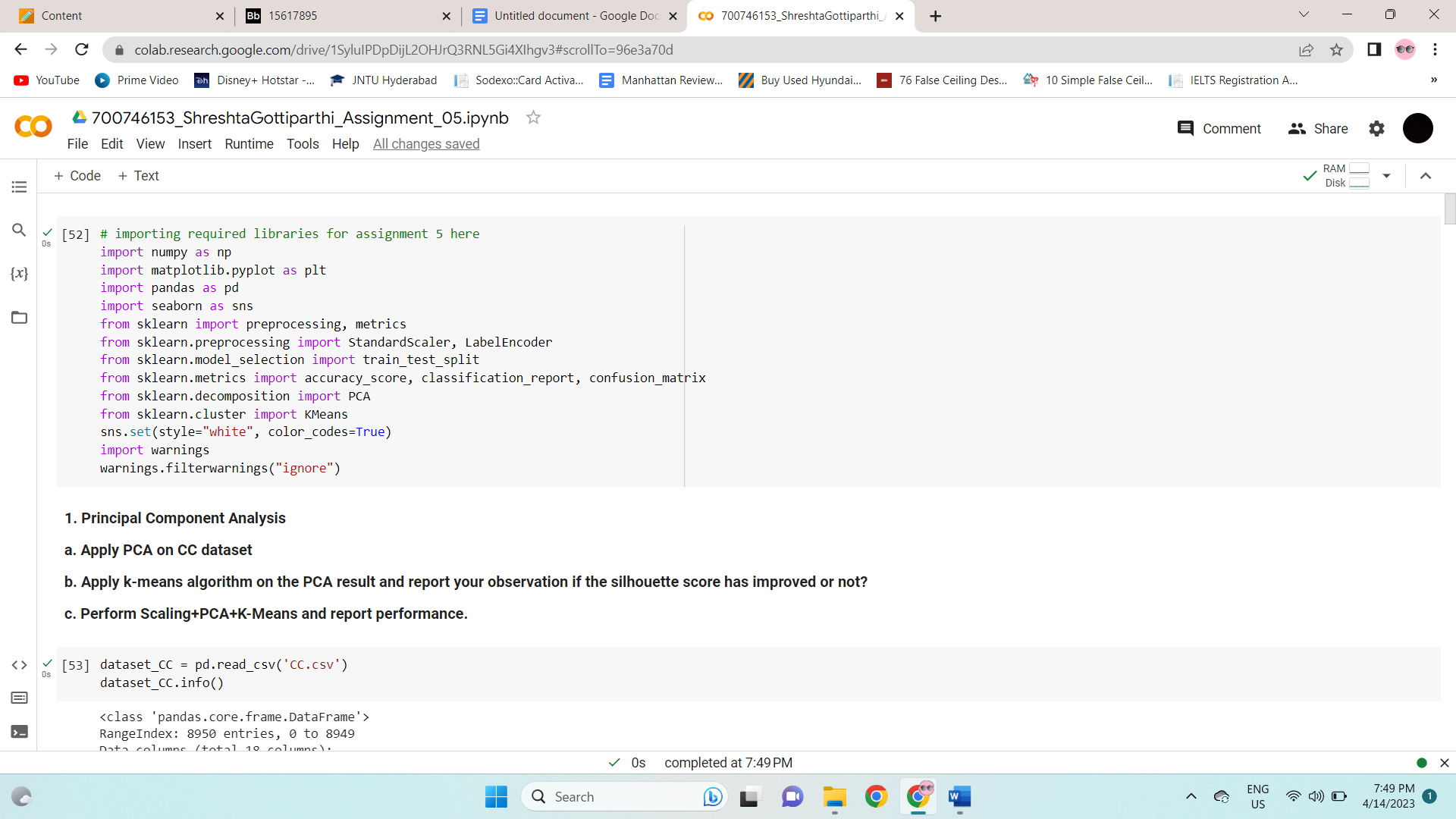
**Videolink:**

[**https://drive.google.com/file/d/1H7loQ6Bv289WzCteMVGp7oIJgjOhlTJF/view?usp=share\_link**](https://drive.google.com/file/d/1H7loQ6Bv289WzCteMVGp7oIJgjOhlTJF/view?usp=share_link)

**Shreshta Gottiparthi**

**700746153**

**Imported required libraries:**

****

**1. Principal Component Analysis**

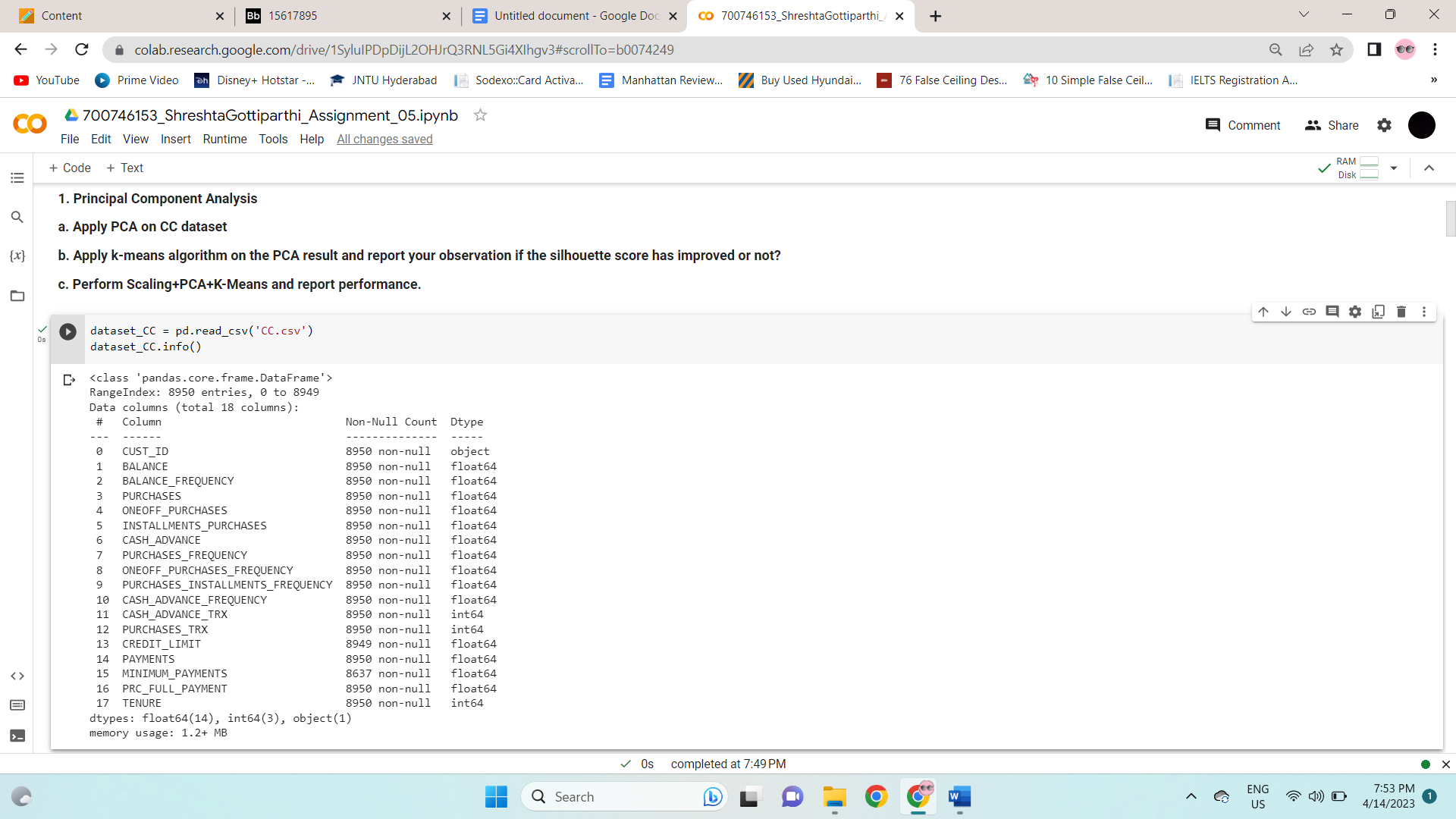
a. Apply PCA on CC dataset.

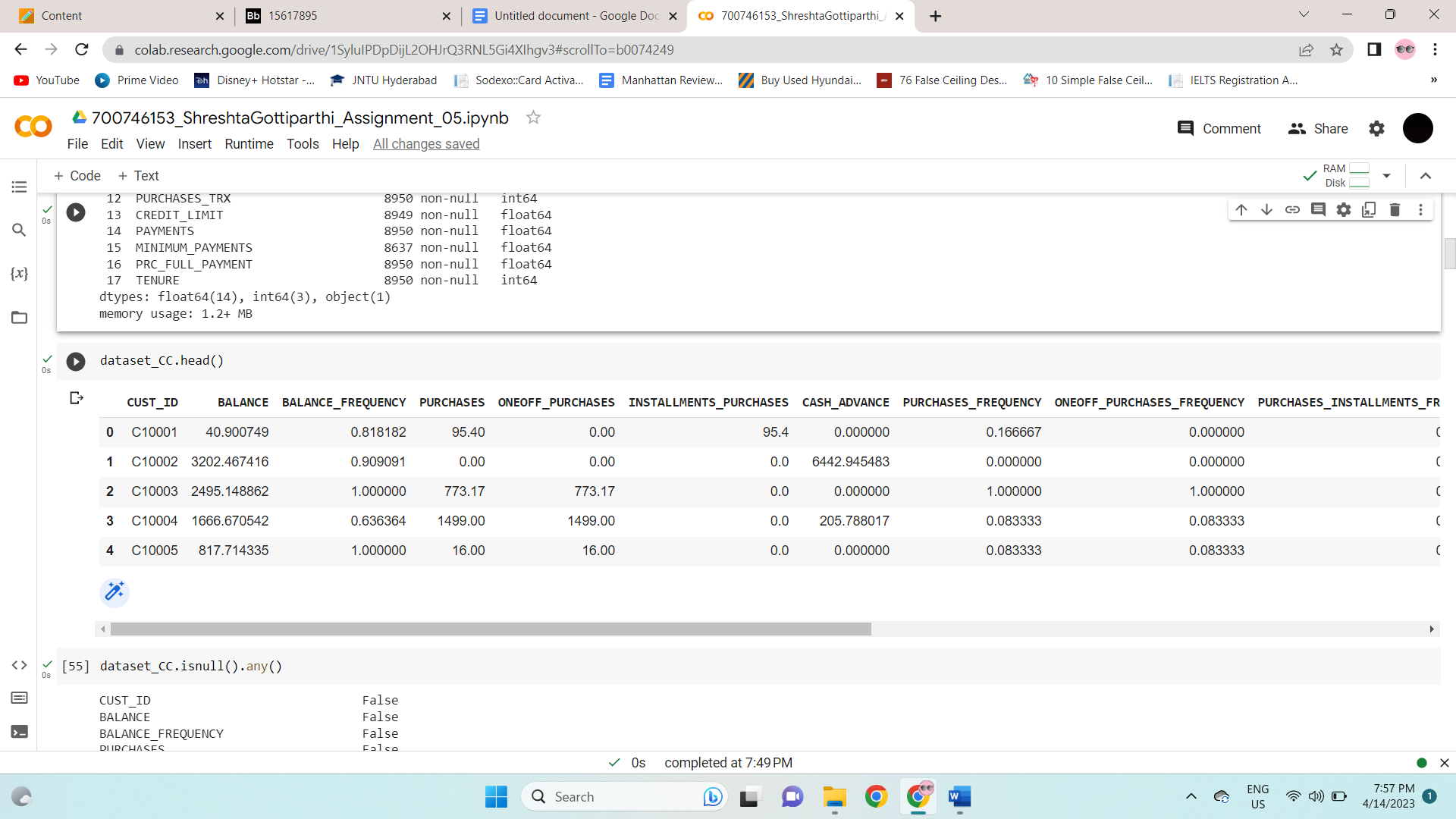
b. Apply k-means algorithm on the PCA result and report your observation if the silhouette score has improved or not?

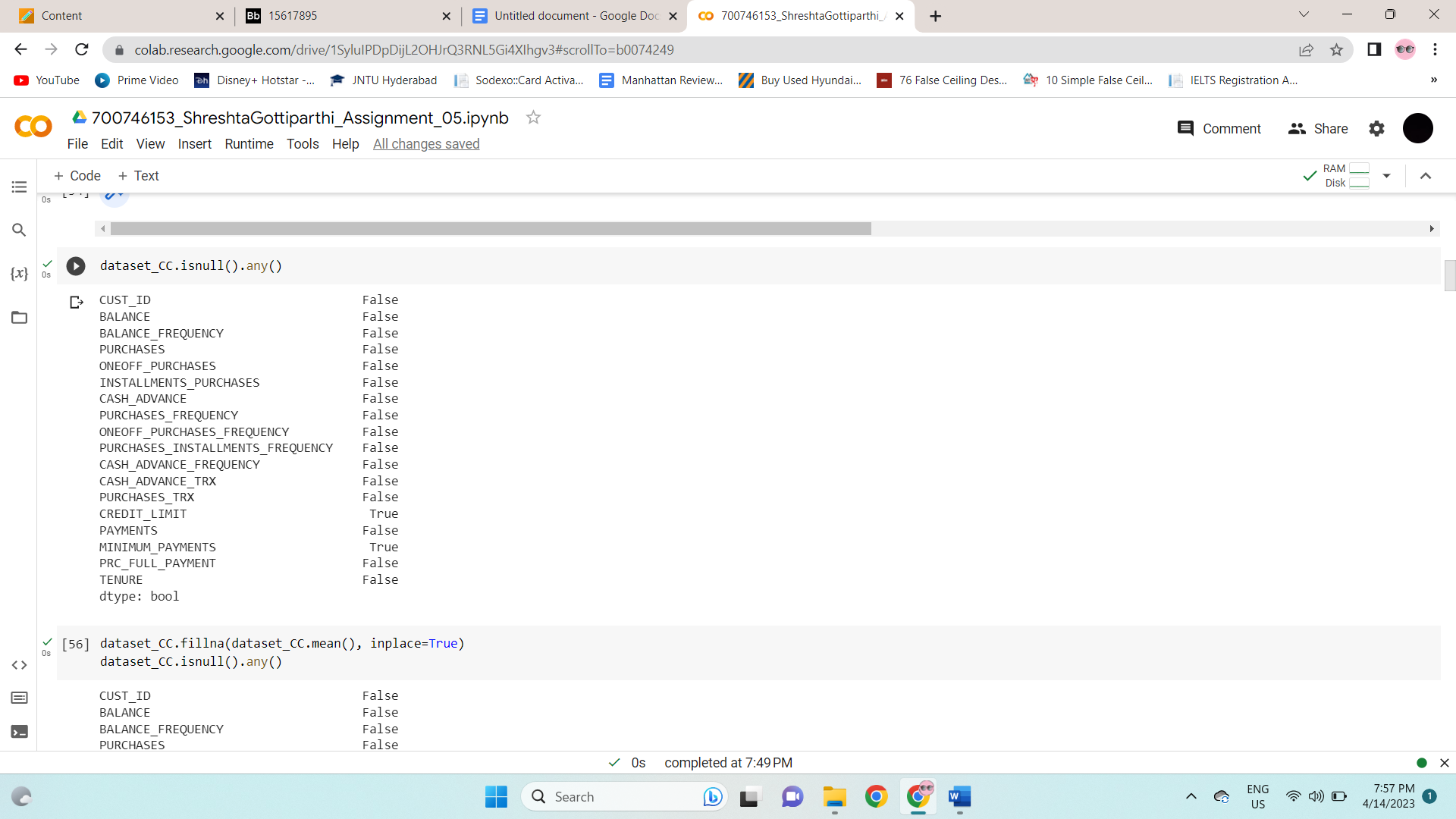
c. Perform Scaling+PCA+K-Means and report performance.

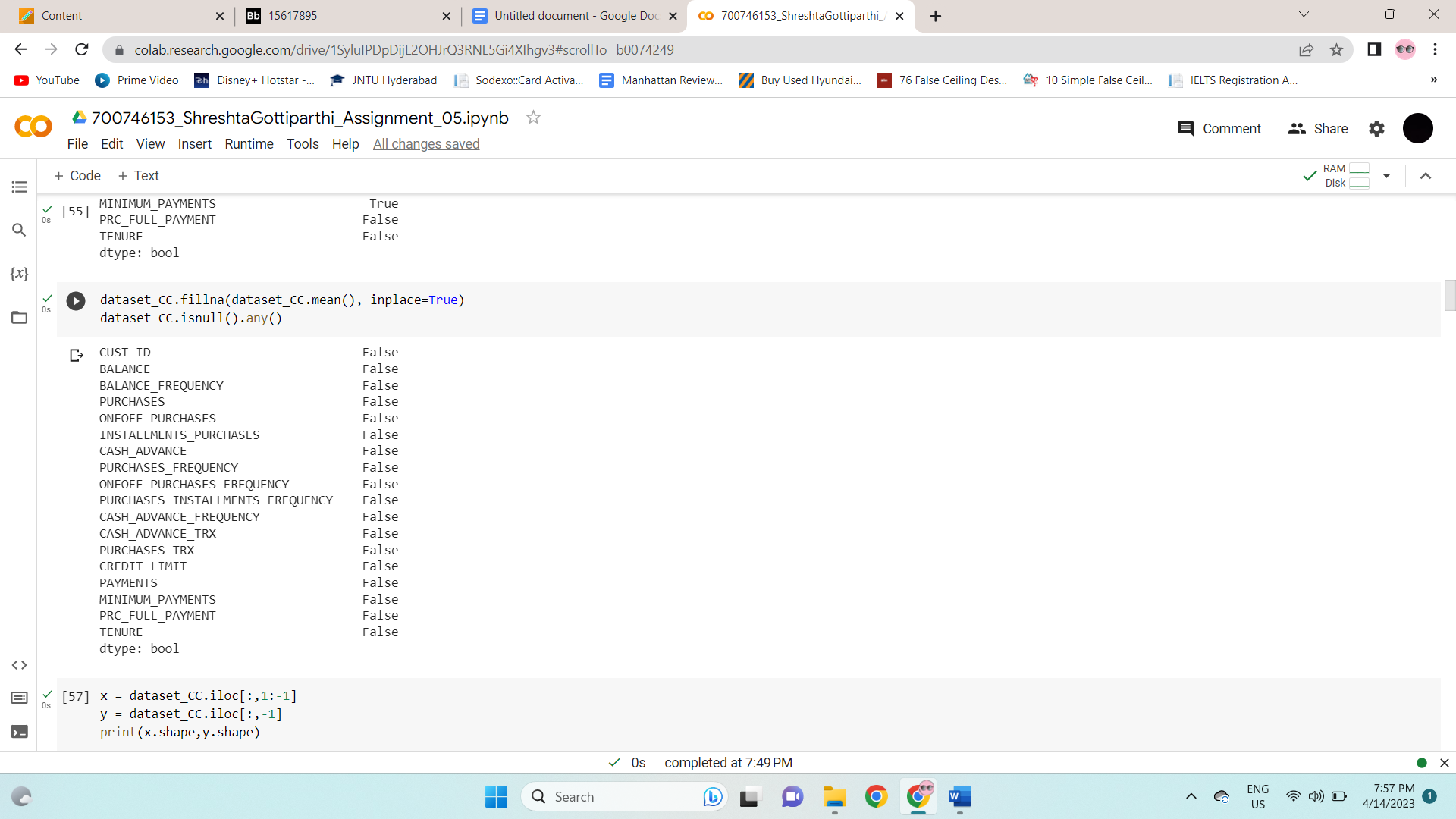
Read the file, checked if there are any null values. If any null values, filled them with mean.

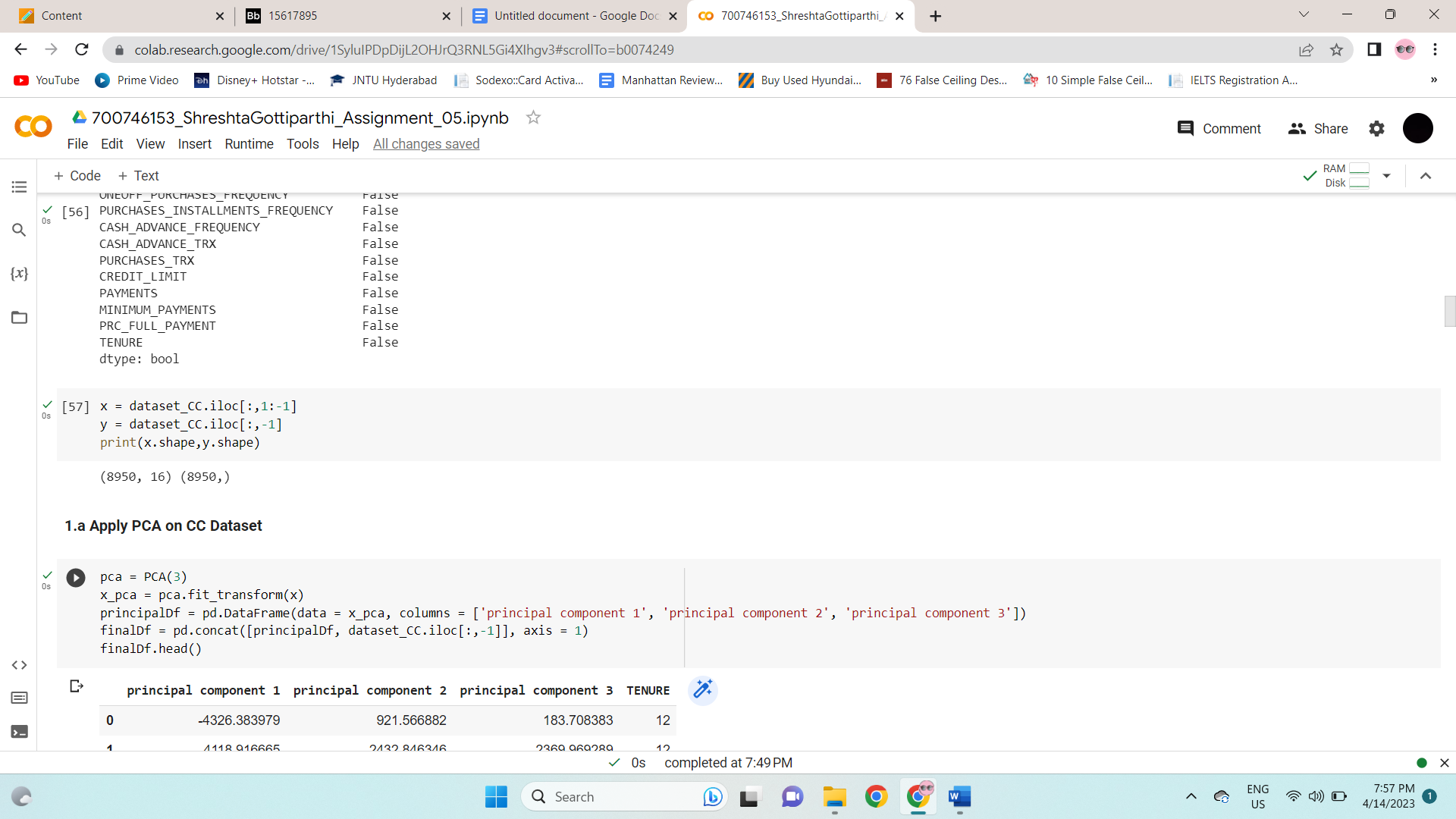
Applied PCA on cc dataset using “.fit transform” and showed the final dataset.

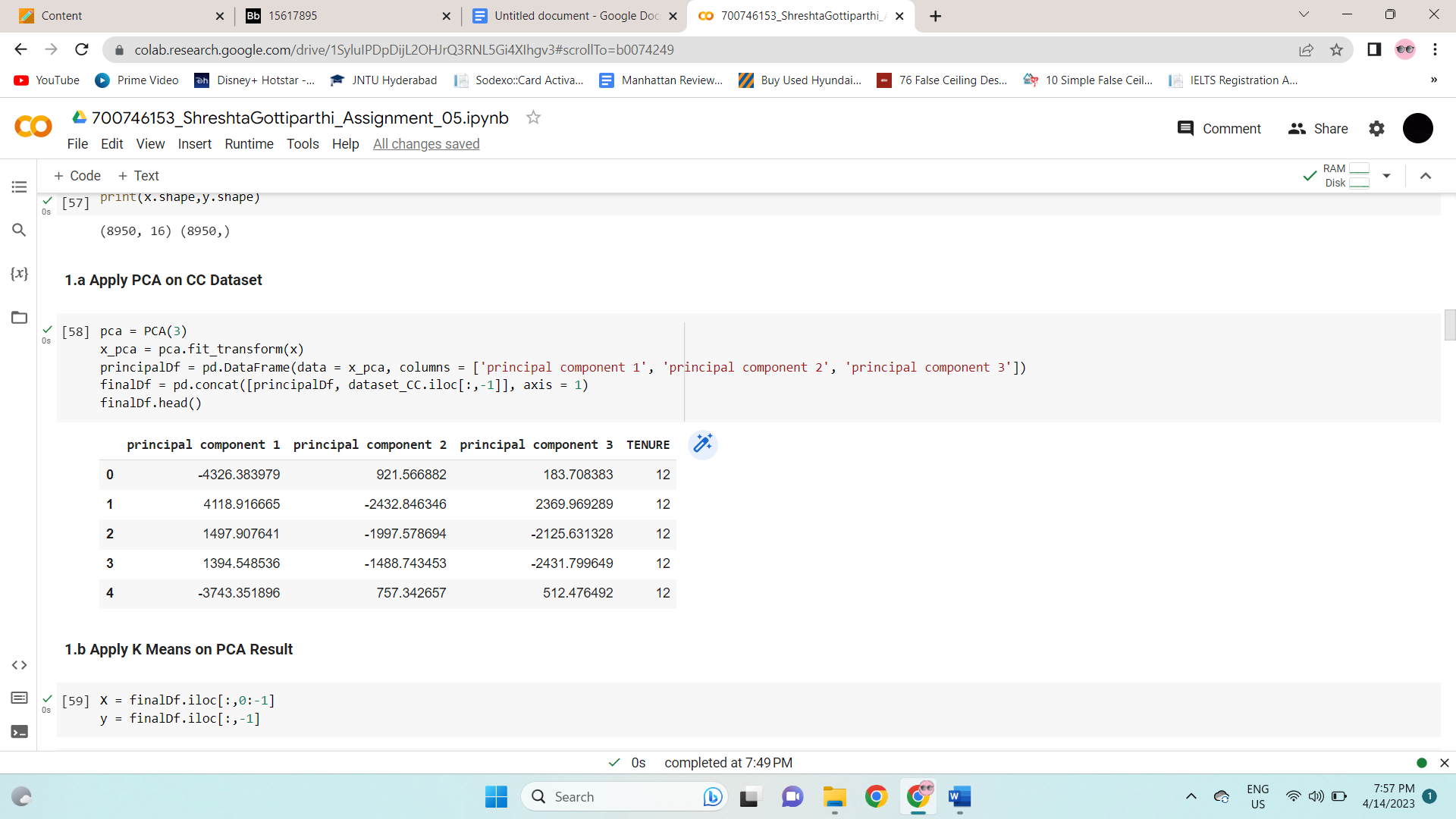


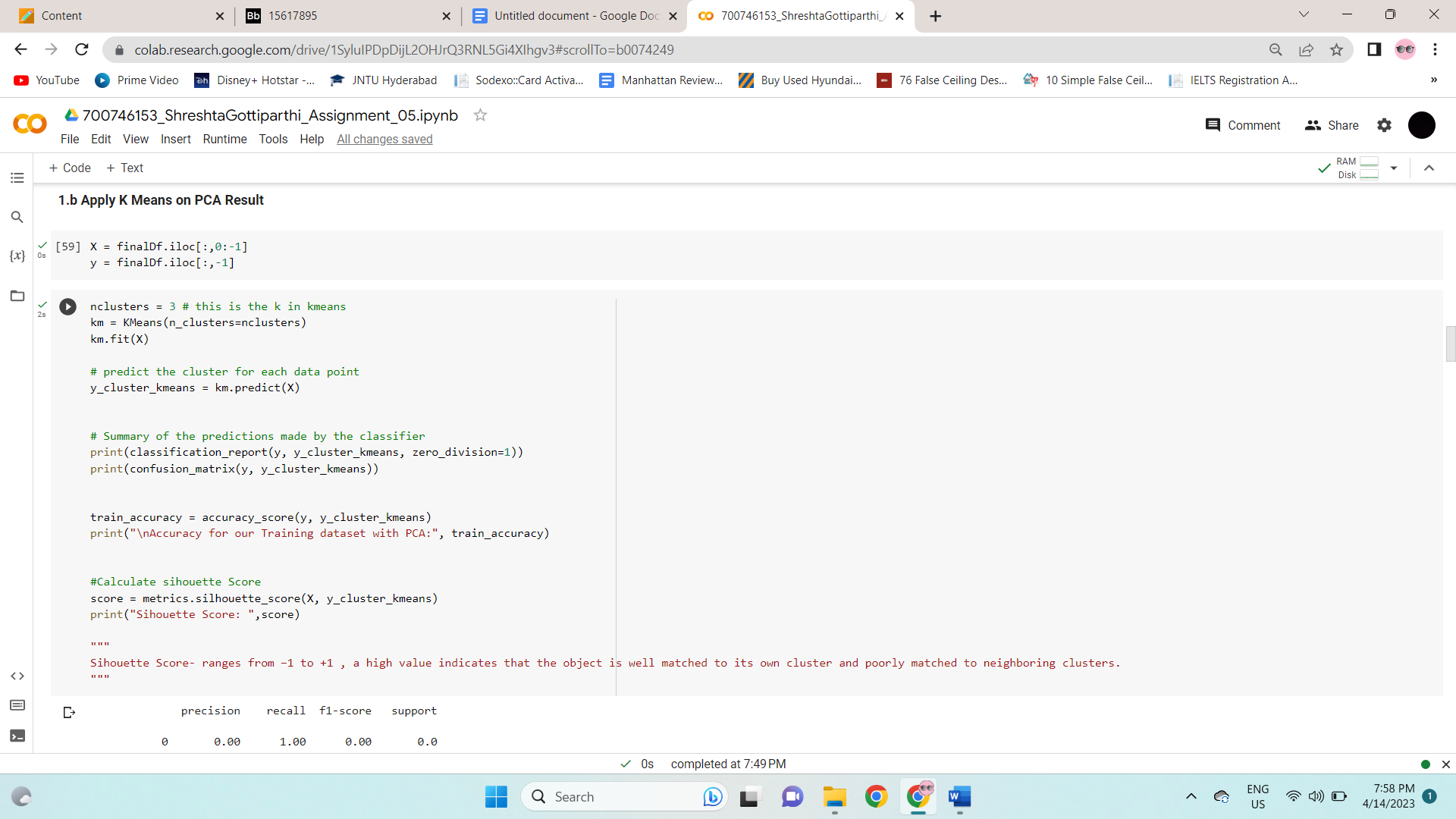


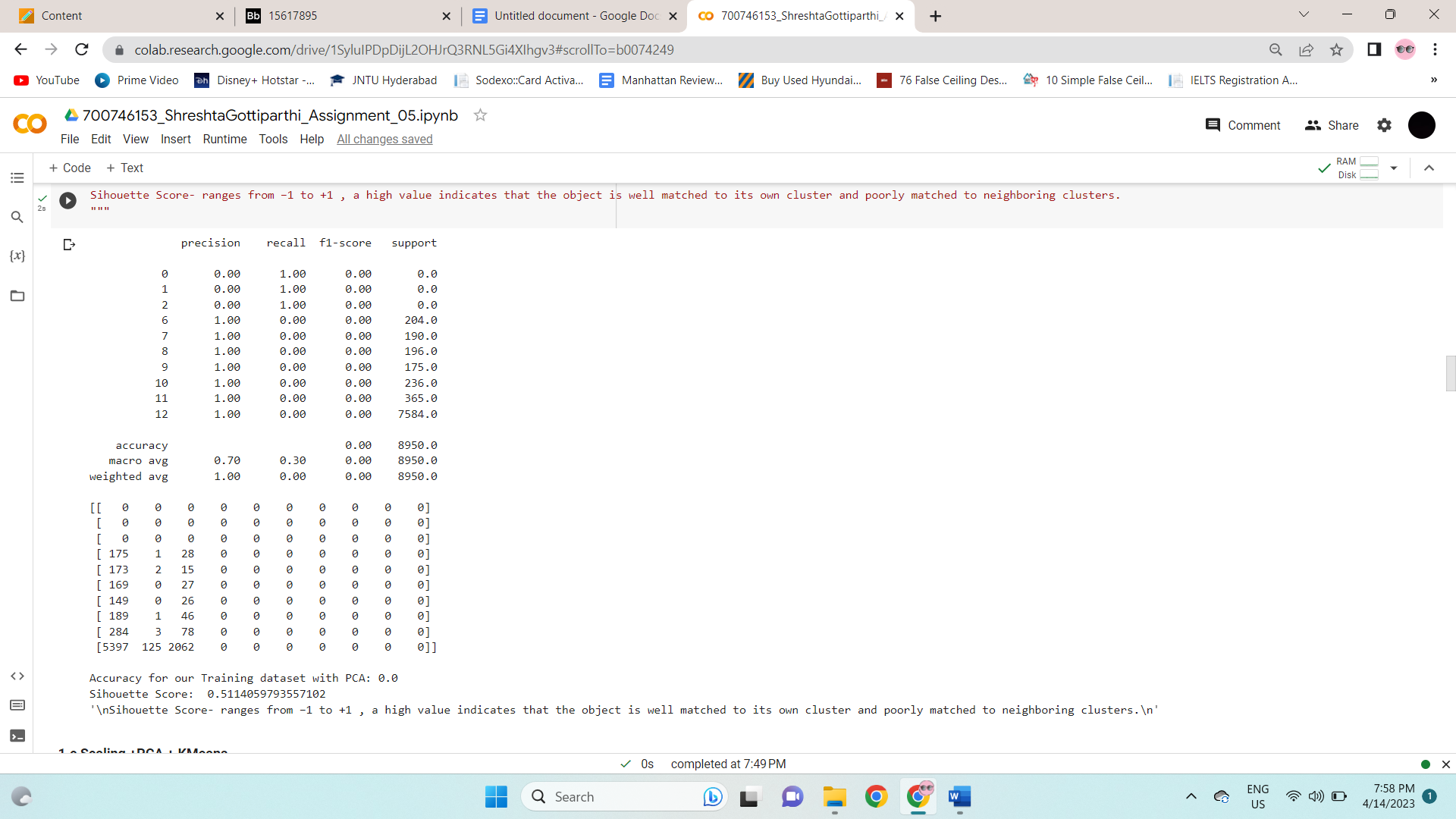


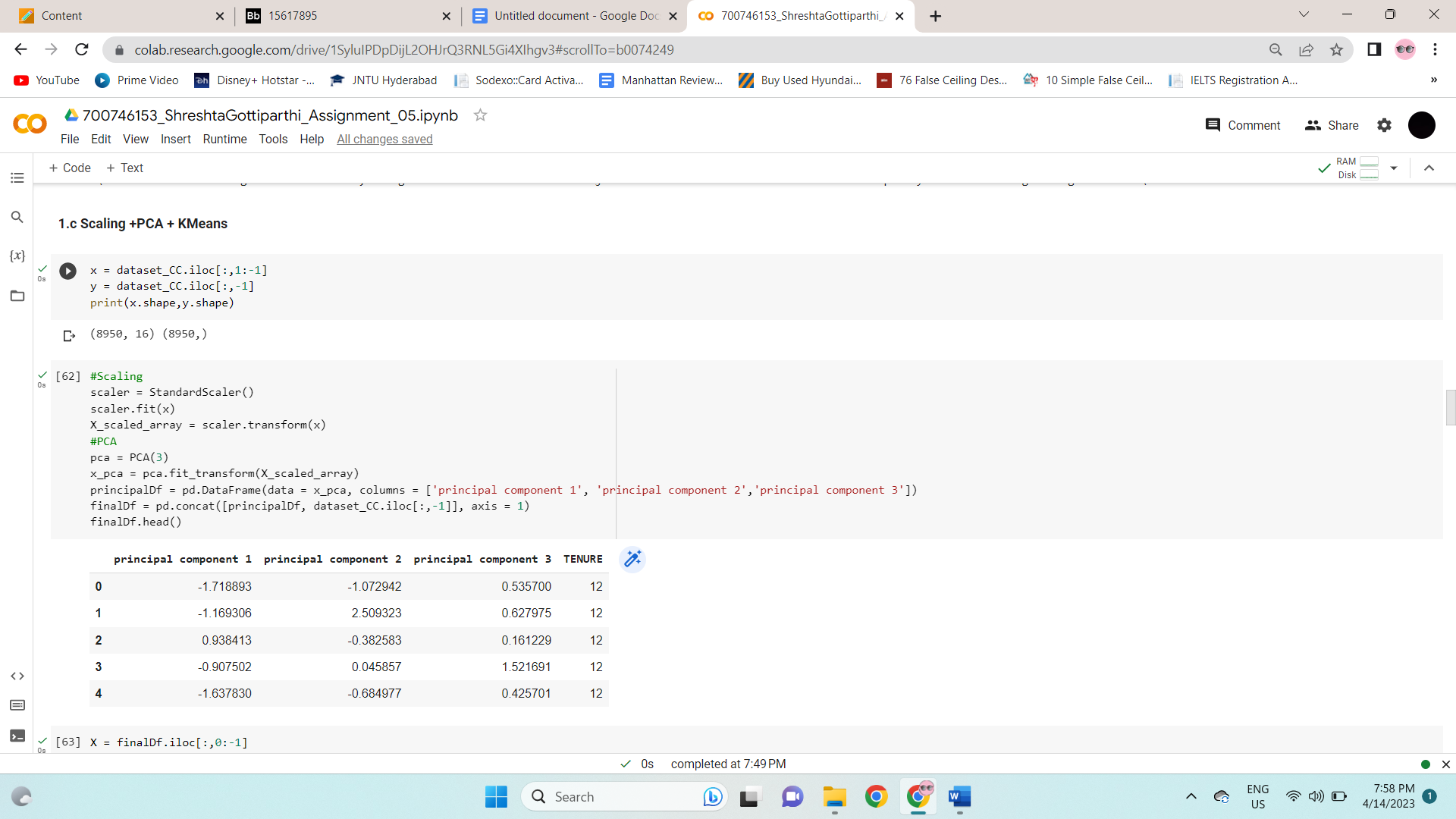


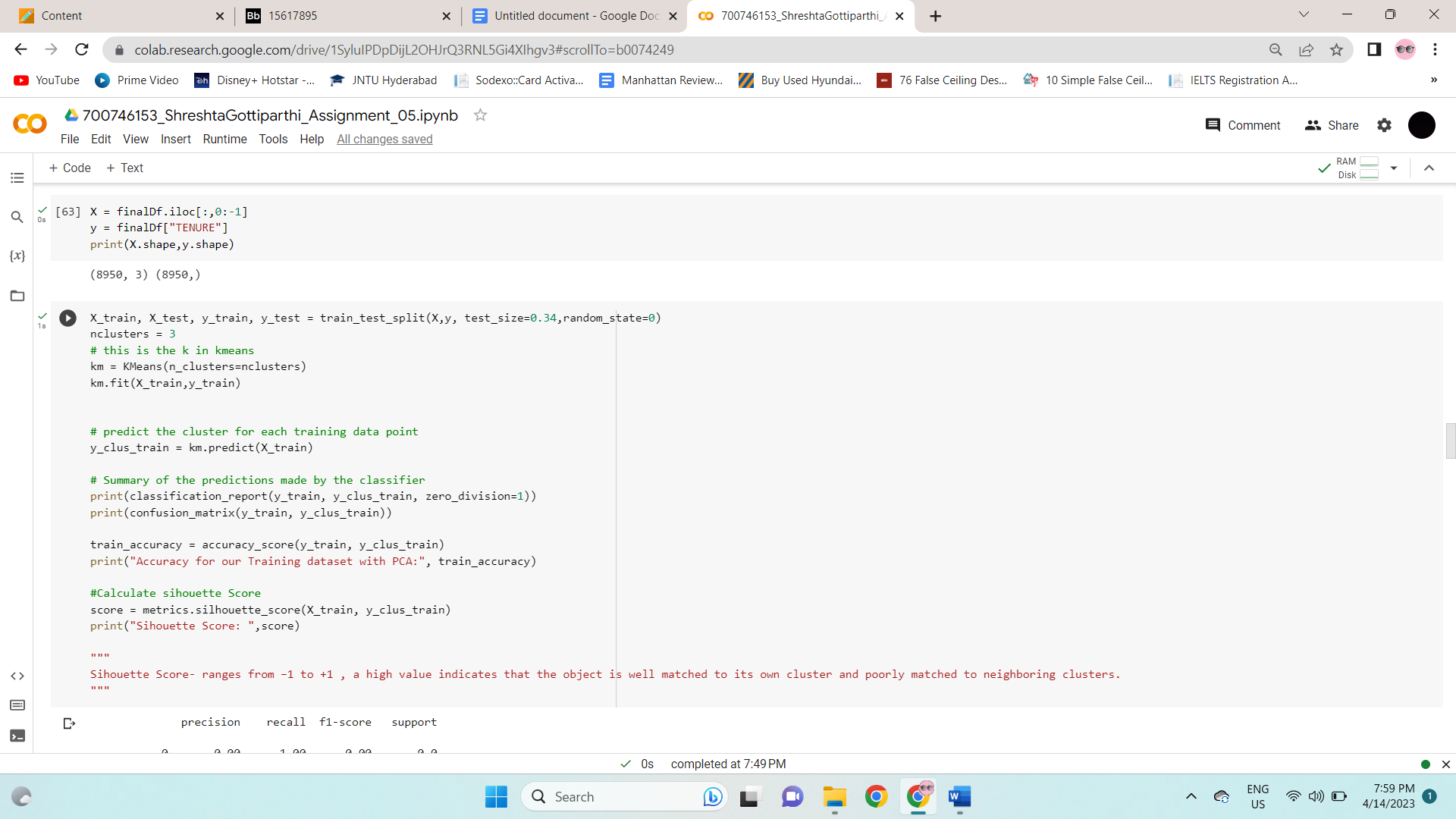


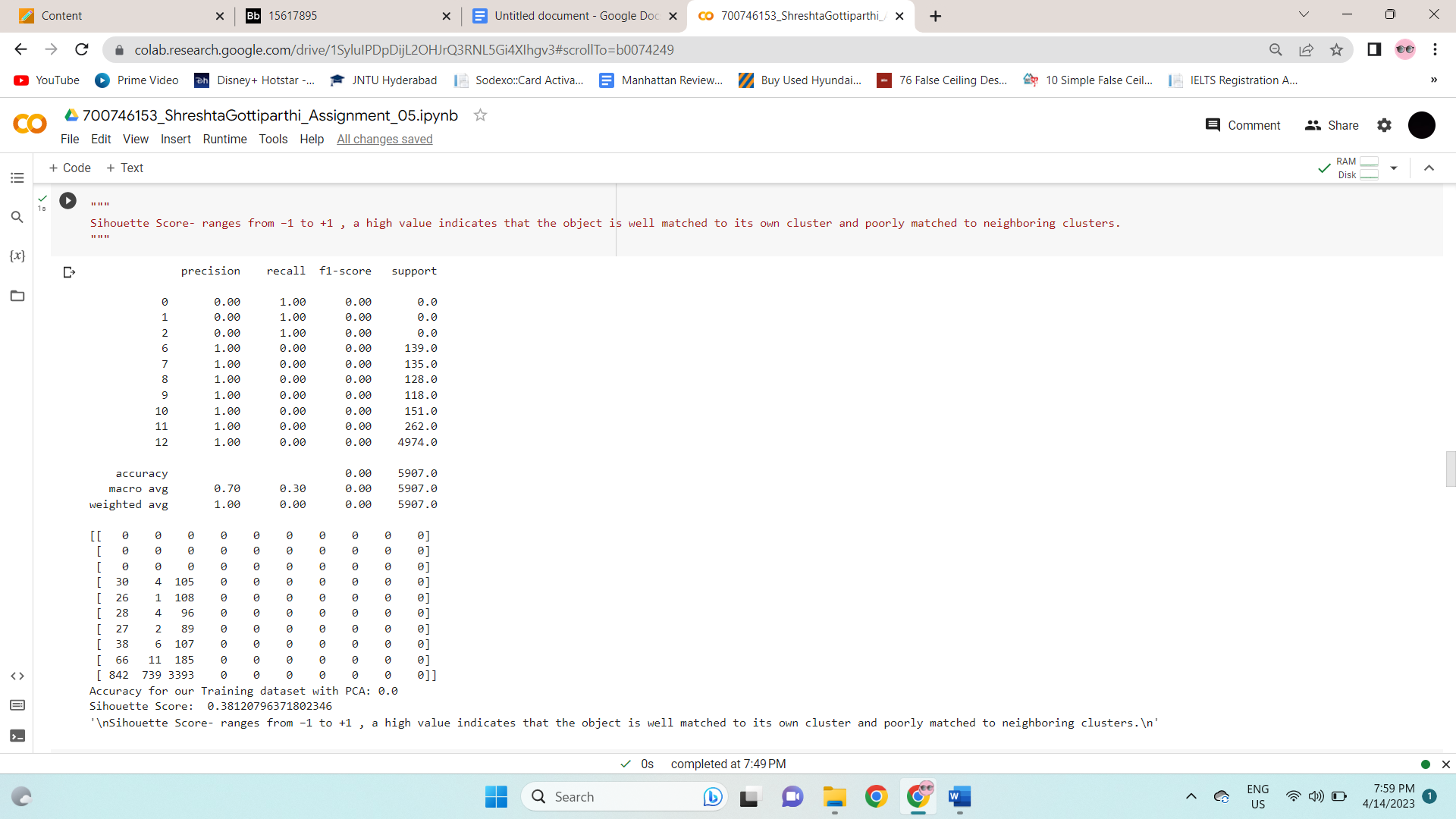


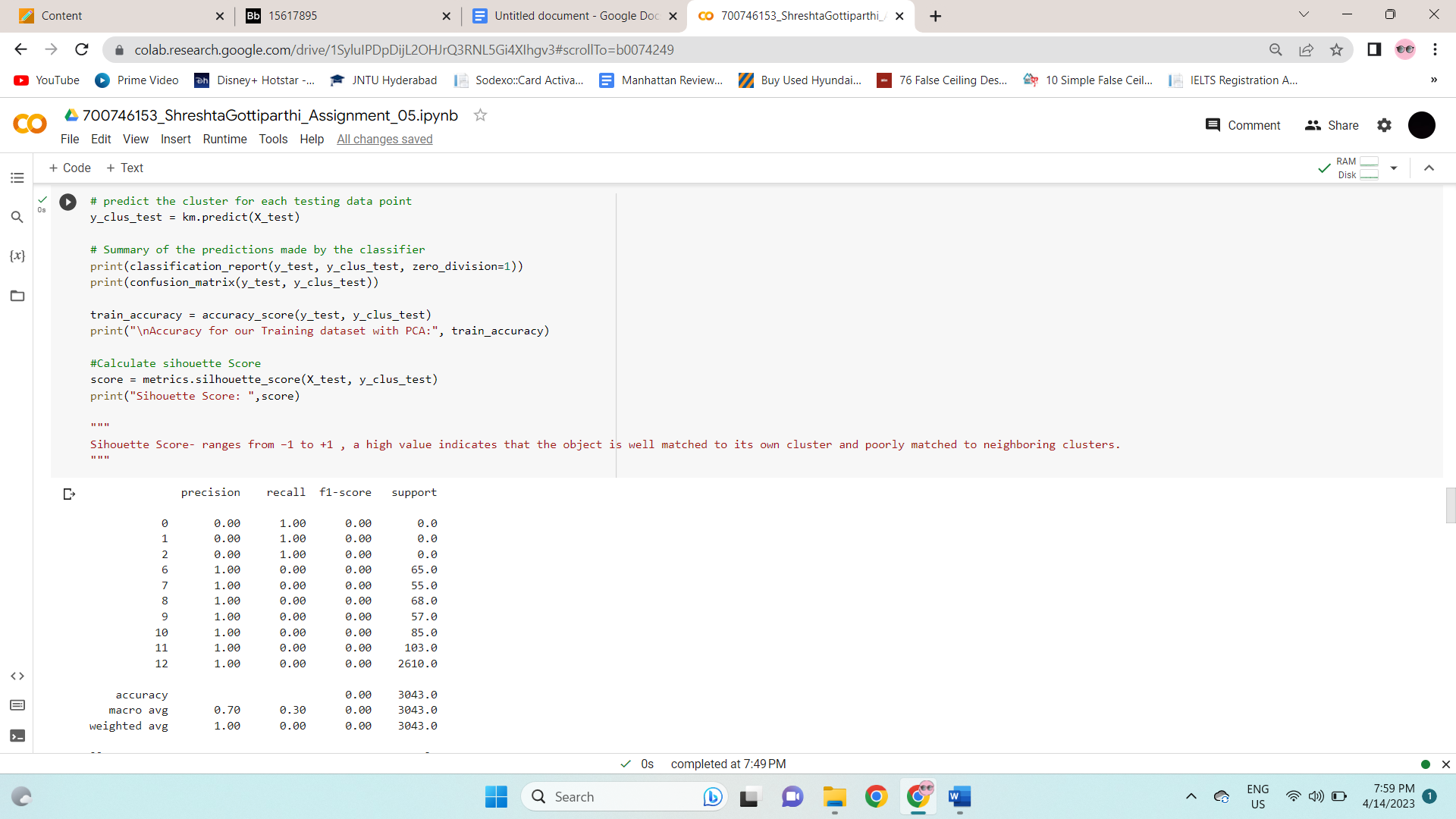


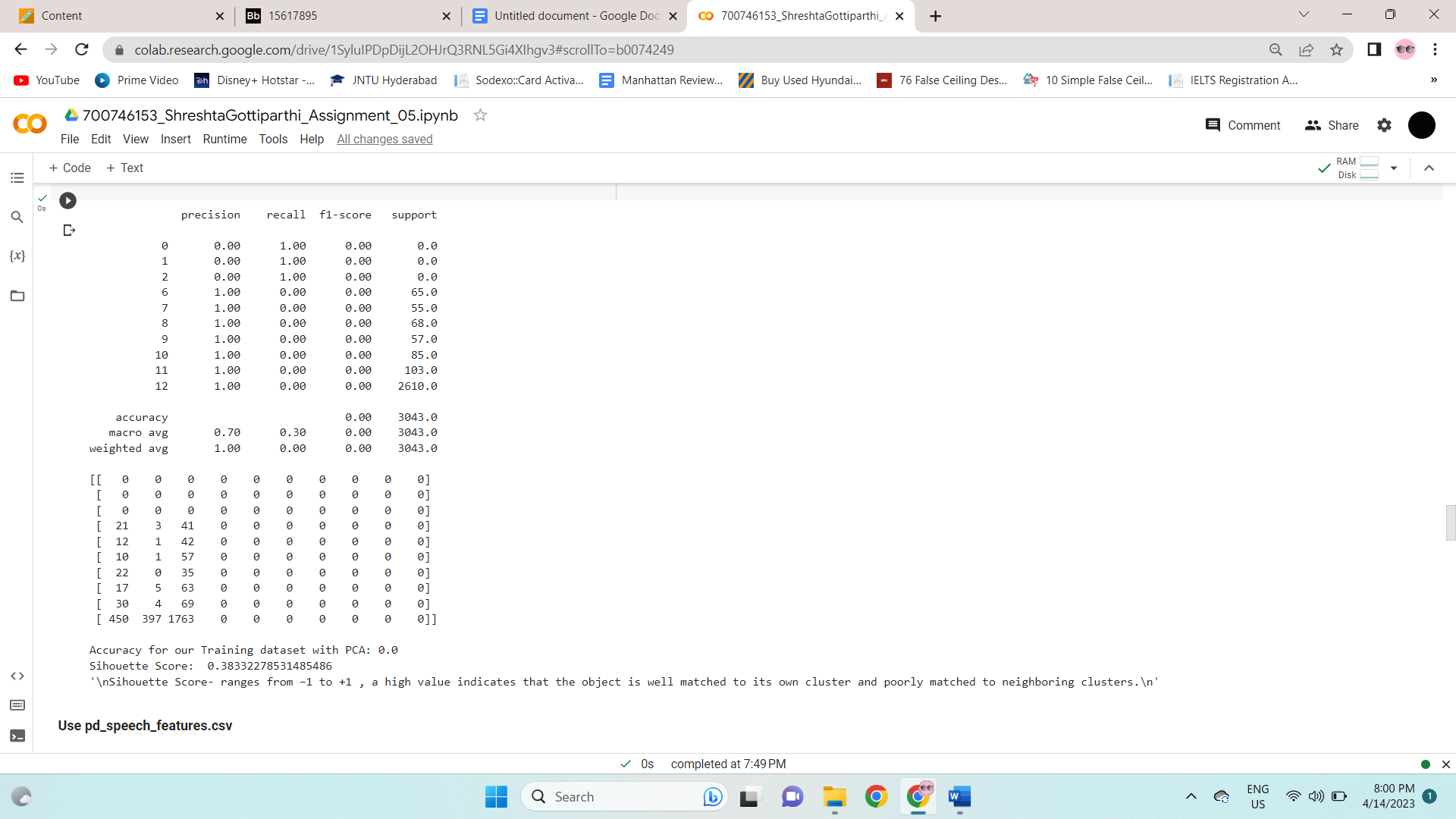










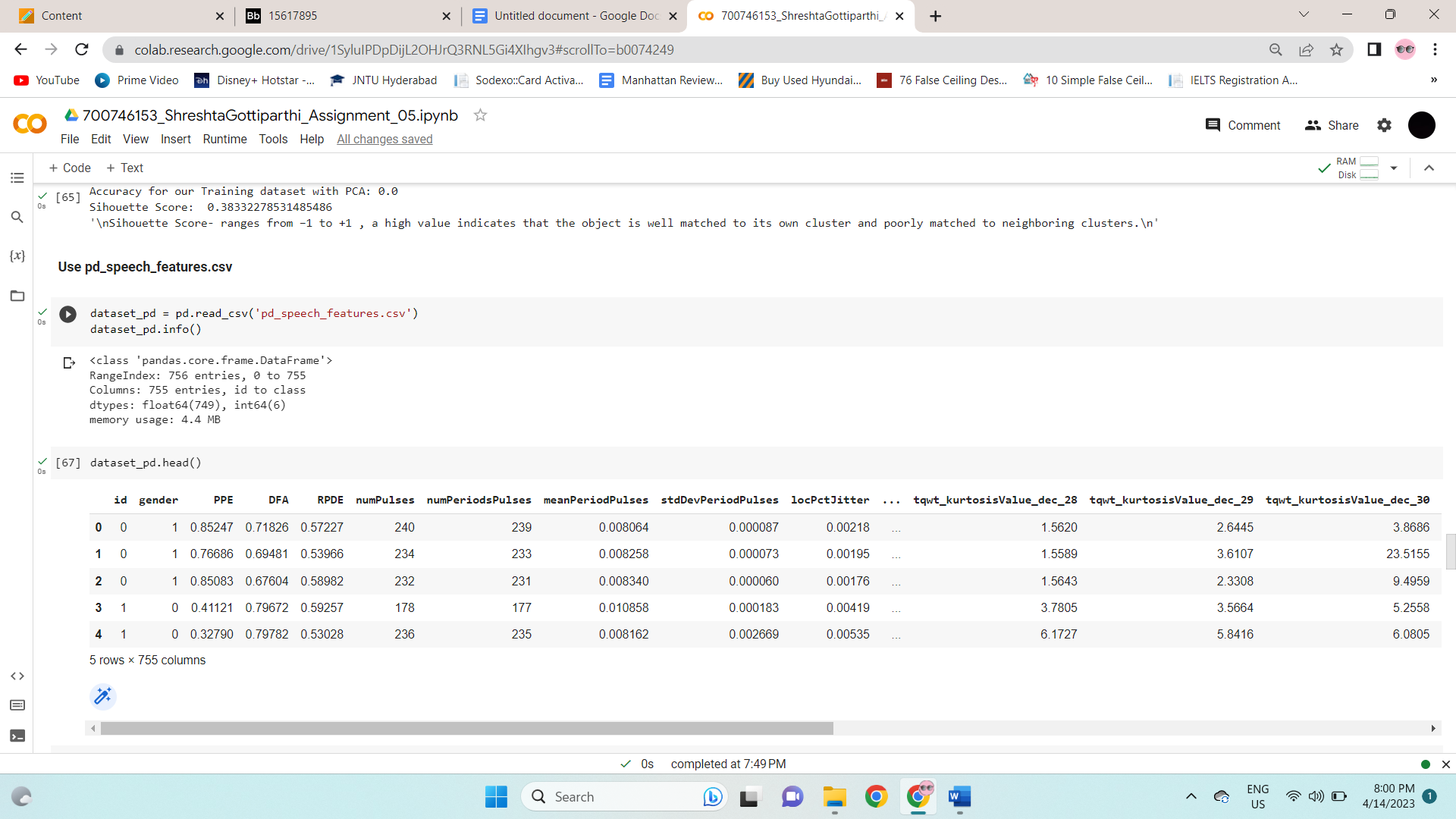


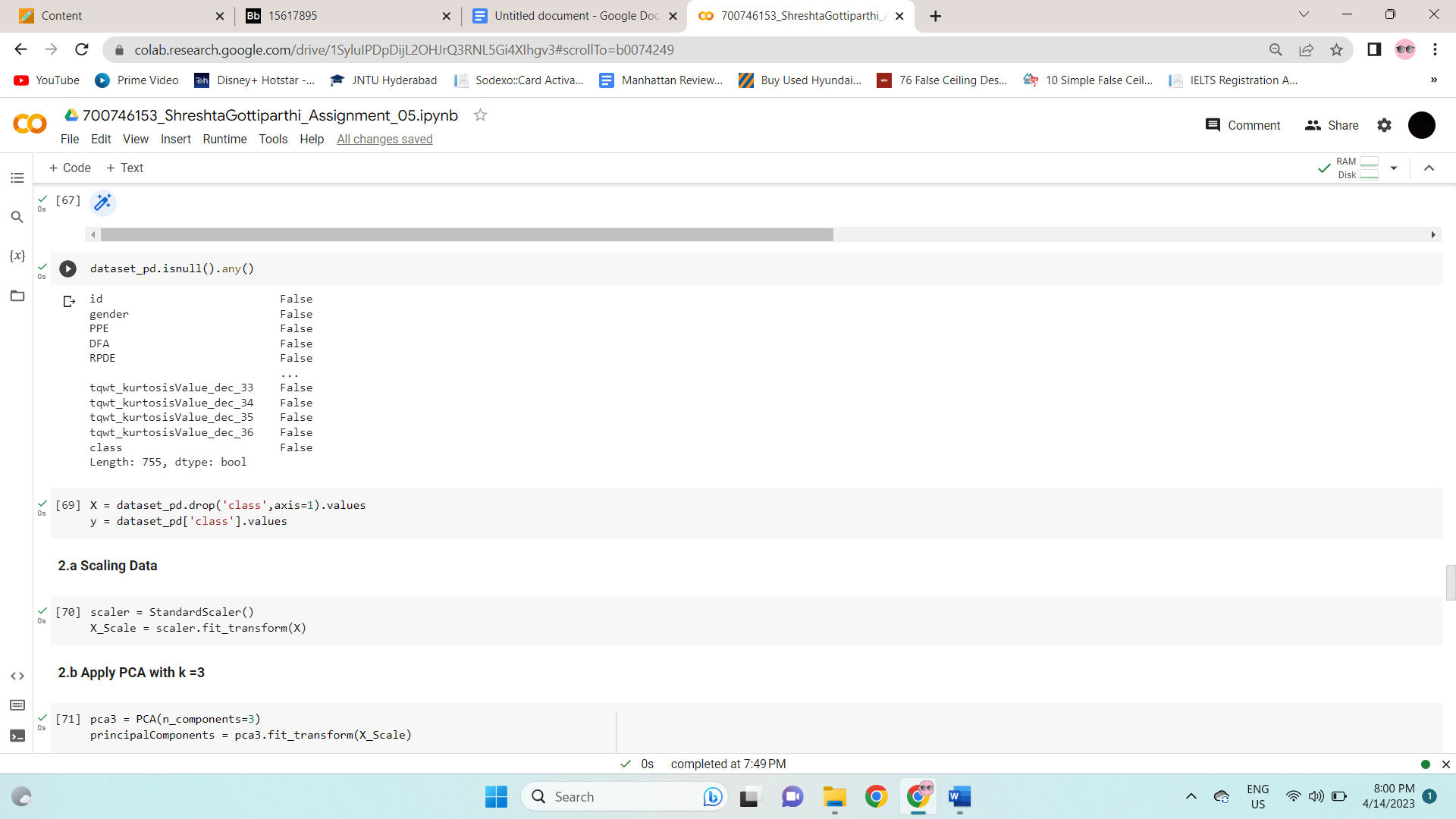
**2. Use pd\_speech\_features.csv**

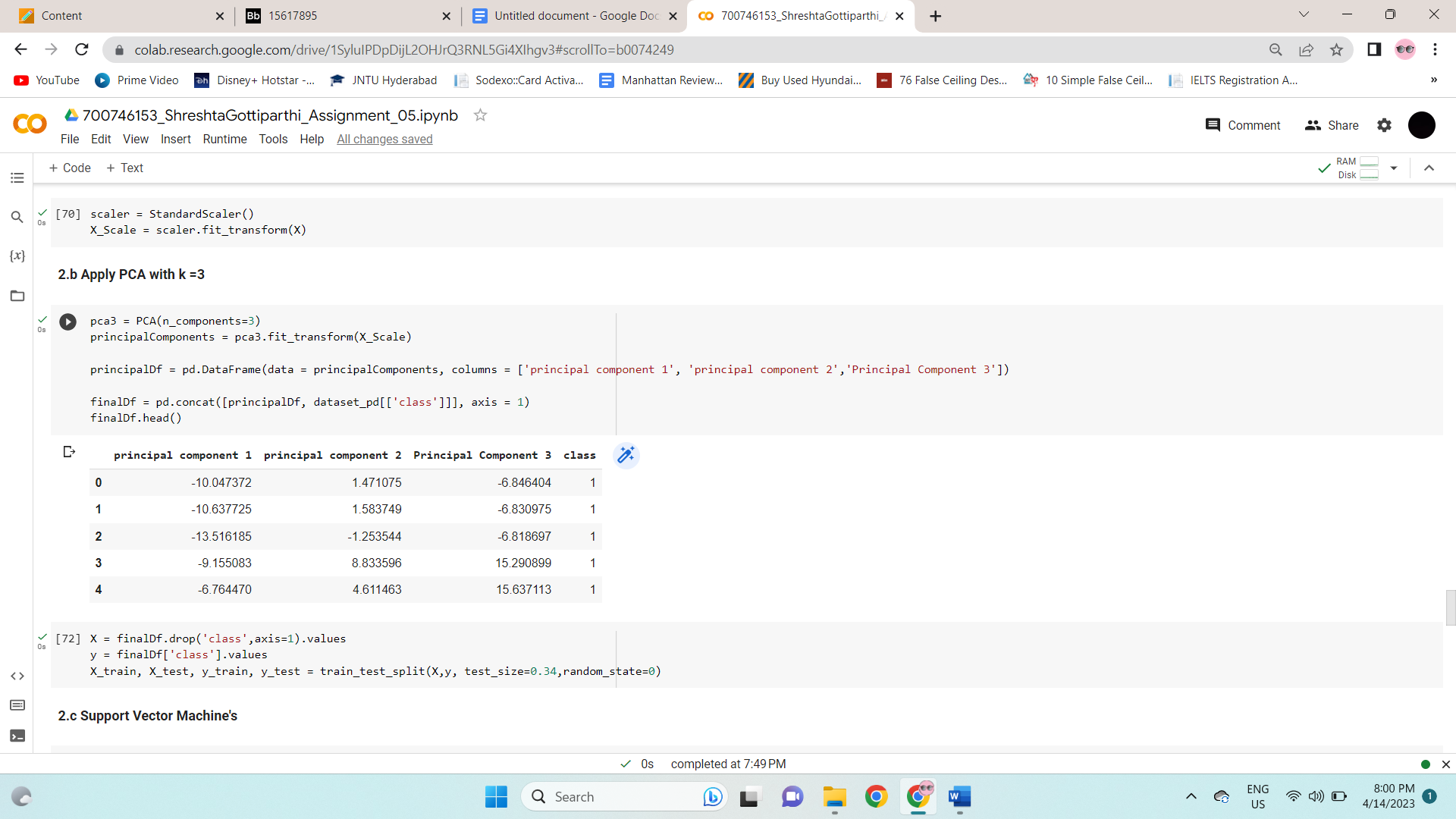
a. Perform Scaling

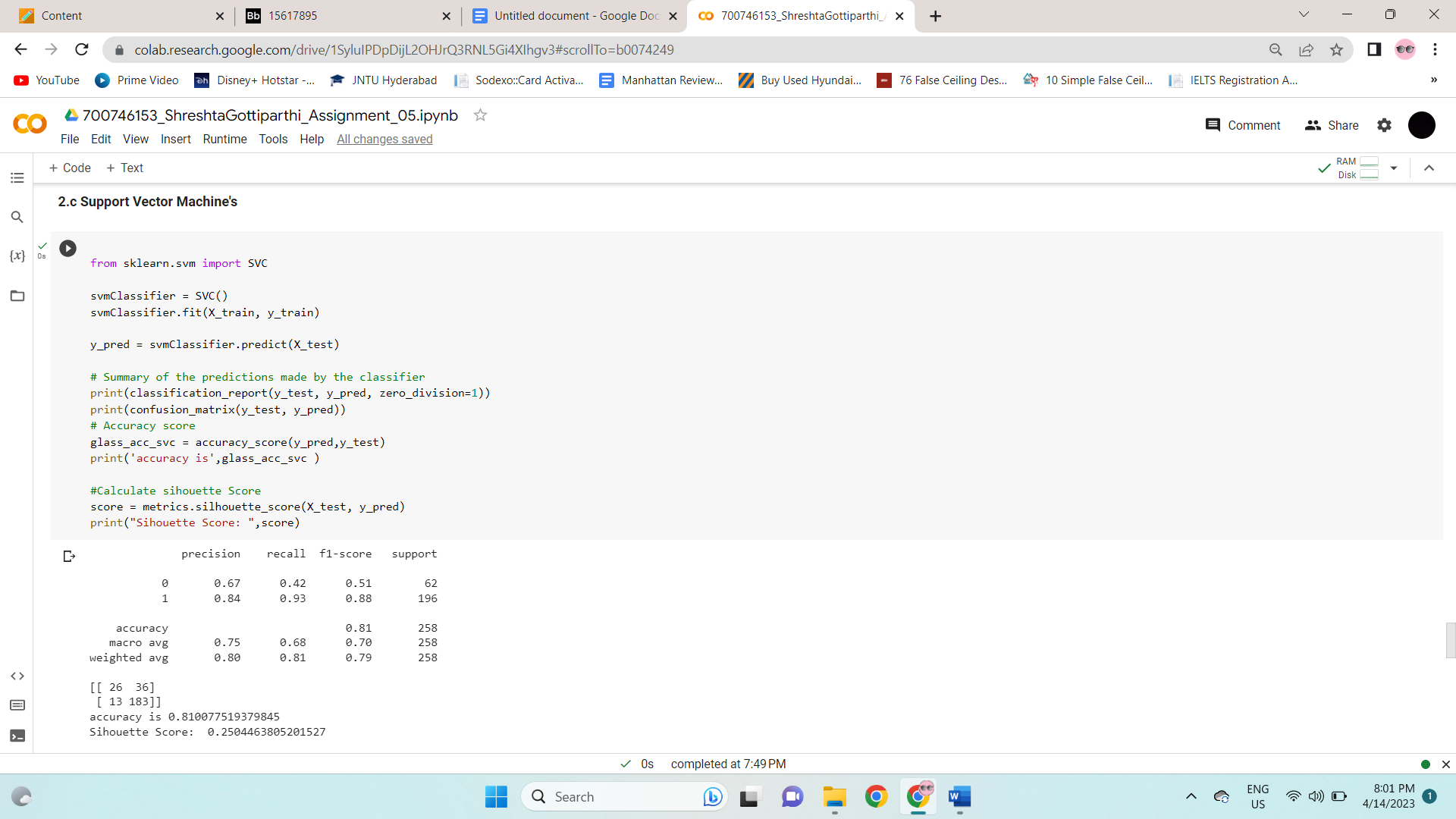
b. Apply PCA (k=3)

c. Use SVM to report performance

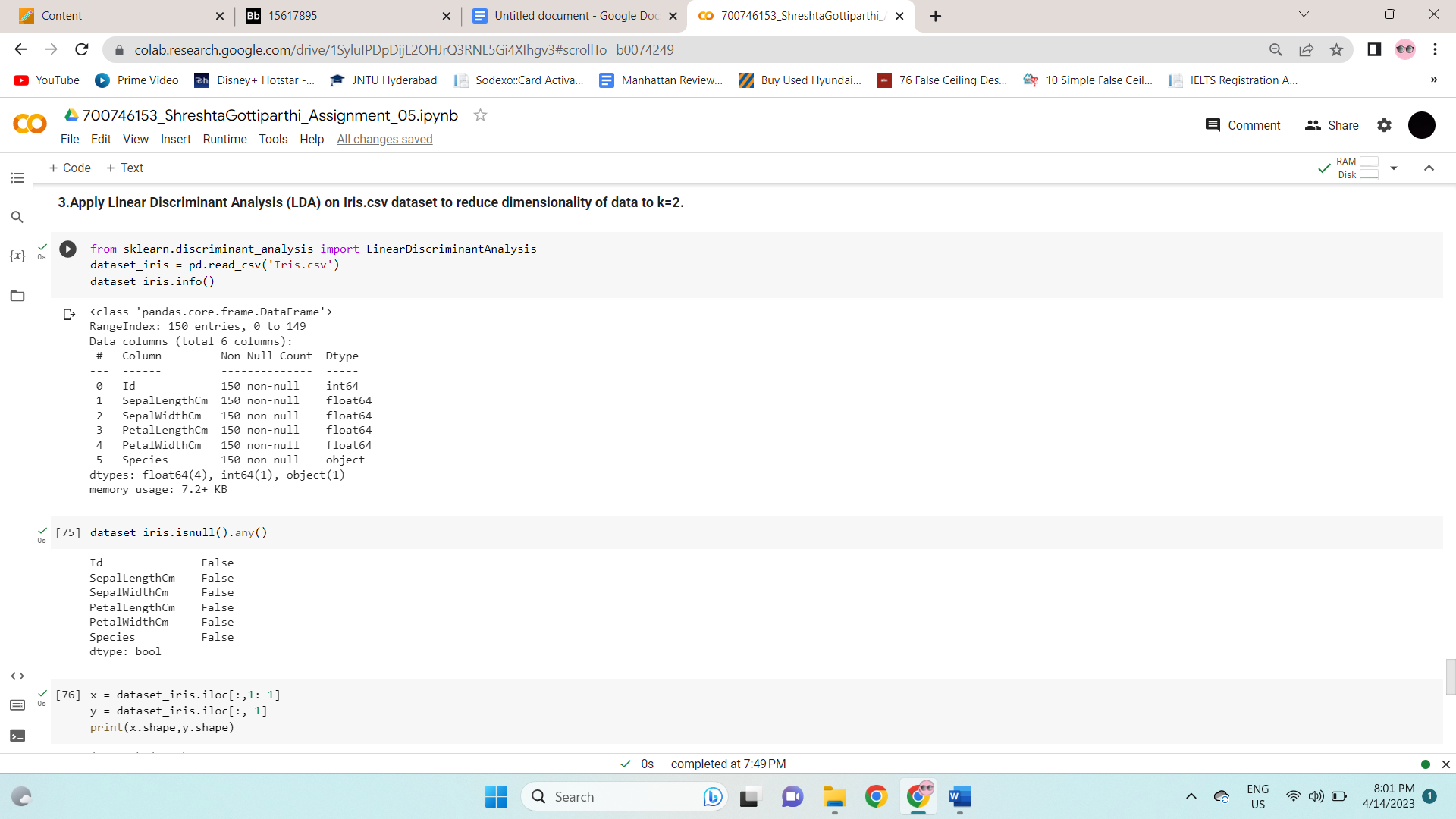








**3. Apply Linear Discriminant Analysis (LDA) on Iris.csv dataset to reduce dimensionality of data tok=2.**





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

To optimize variance in a lower dimension, LDA and PCA both use linear transformations. Unlike LDA, which uses supervised learning, PCA uses unsupervised learning. Accordingly, LDA discovers paths of maximum class separability while PCA discovers directions of maximum variance irrespective of class labels.

Principal components, which are linear combinations of the original variables, are used to condense the characteristics into a more manageable group of orthogonal variables. The highest amount of data variability is captured by the first component, followed by the second and third, and so on. To maximize the variation across the various categories while limiting the variance inside the class, LDA seeks out the linear discriminants.