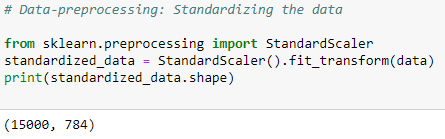
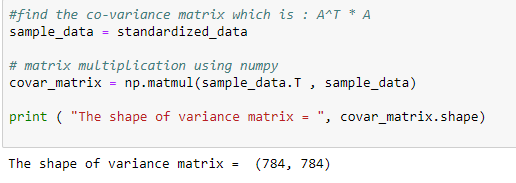
This will tell steps to perform PCA.

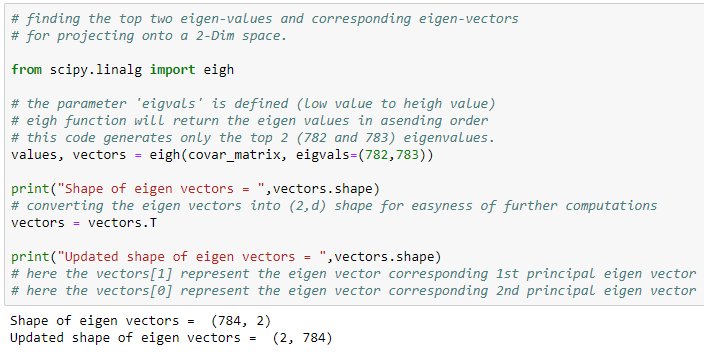
**Step 1:** Perform Standardization.



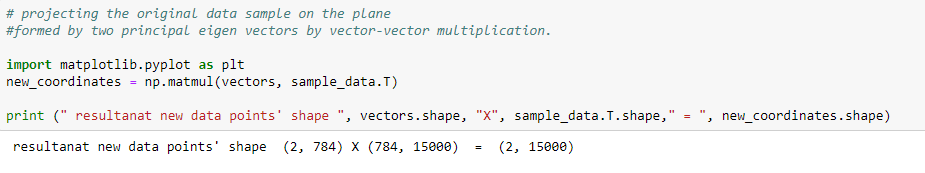
**Step 2:** Find co-variance matrix.



**Step 3**: Find eigen vector and eigen values:



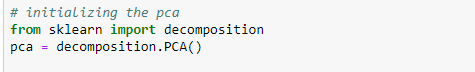
**Step 4:** Generate new components or features by projecting original data on eigen vectors.



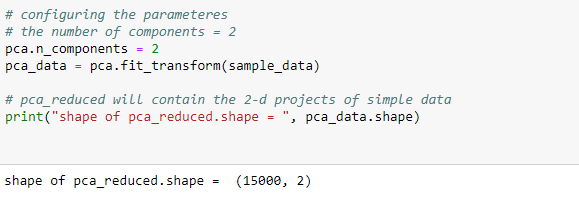
Following this all 4 steps you will get new features or components using PCA.

There is built in function available which will perform most of the work in **sklearn.decomposition**, following are the steps for this process:

**Step 1:** import decomposition from sklearn and create instance of PCA from it.



**Step 2:** configure the number of components or new features you want and apply fit\_transform, which will generate new data-matrix which contains our new components. And by using just two steps you can perform PCA.



There is always a question of how many new features we should generate or how much dimensionality reduction we should perform, to how many components to generate we can do following thing:

We can calculate percentage of variance explained for all d’ (where d’ = d, d is no of features of original dataset) components, then we can do cumulative sum to obtain till how many features we are obtaining the required percentage of variance explained.

