

```

import pandas as pd
import numpy as np
from numpy import linalg as la
import pylab as pl
import matplotlib.pyplot as plt

data = pd.read_csv('/content/iris(For PCA Program).csv')

data

x = data.drop(columns='species')
x = np.array(x)
x

y = data['species']
y = np.array(y)
y

x_mean = np.mean(x)

y_mean = np.mean(y)

x_std = x-x_mean
y_std = y-y_mean

cov_mat = np.cov(x_std.T)
cov_mat

eig_val,eig_vec = la.eig(cov_mat)

eig_pairs = [(np.abs(eig_val[i]), eig_vec[:,i]) for i in range(len(eig_val))]

eig_pairs.sort(key=lambda x: x[0], reverse=True)

print('Eigenvalues in descending order:')
for i in eig_pairs:
    print(i[0])

total = sum(eig_val)
for i in eig_val:
    var = (i/total)*100
    print(var)

matrix_w = np.hstack((eig_pairs[0][1].reshape(4,1),eig_pairs[1][1].reshape(4,1)))
matrix_w

Y = x_std.dot(matrix_w)
print(Y[0:5])

# pl.figure()
# target_names = ['Setosa','Versicolor','Virginica']
# for c, i, target_name in zip("rgb", [1, 2, 3], target_names):
#     pl.scatter(Y[y==i,0], Y[y==i,1], c=c, label=target_name)
# pl.xlabel('Principal Component 1')
# pl.ylabel('Principal Component 2')
# pl.legend()
# pl.title('PCA of IRIS dataset')
# pl.show()

for i in range(len(y)):
    if y[i]==1:
        plt.scatter(Y[i,0],Y[i,1],c="red")
    elif y[i]==2:
        plt.scatter(Y[i,0],Y[i,1],c="green")
    elif y[i]==3:
        plt.scatter(Y[i,0],Y[i,1],c="blue")

```

[Colab paid products](#) - [Cancel contracts here](#)

