

```
In [2]: import pandas as pd
data=pd.read_csv("Iris.csv")
data.head()
```

Out[2]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

```
In [3]: from sklearn.preprocessing import LabelEncoder
label_encoder=LabelEncoder()
data['species']=label_encoder.fit_transform(data['species'])
data
```

Out[3]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	2
146	6.3	2.5	5.0	1.9	2
147	6.5	3.0	5.2	2.0	2
148	6.2	3.4	5.4	2.3	2
149	5.9	3.0	5.1	1.8	2

```
In [4]: from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
data=sc.fit_transform(data)
data=pd.DataFrame(data)
data
```

Out[4]:

	0	1	2	3	4
0	-0.900681	1.032057	-1.341272	-1.312977	-1.224745
1	-1.143017	-0.124958	-1.341272	-1.312977	-1.224745
2	-1.385353	0.337848	-1.398138	-1.312977	-1.224745
3	-1.506521	0.106445	-1.284407	-1.312977	-1.224745
4	-1.021849	1.263460	-1.341272	-1.312977	-1.224745
...	...	...	...	...	...
145	1.038005	-0.124958	0.819624	1.447956	1.224745
146	0.553333	-1.281972	0.705893	0.922064	1.224745
147	0.795669	-0.124958	0.819624	1.053537	1.224745
148	0.432165	0.800654	0.933356	1.447956	1.224745
149	0.068662	-0.124958	0.762759	0.790591	1.224745

```
In [5]: from sklearn.decomposition import PCA
pca=PCA(n_components=2)
df_final=pca.fit_transform(data)
df_final=pd.DataFrame(df_final)
df_final
```

```
Out[5]:
```

	0	1
0	-2.576198	0.498650
1	-2.419715	-0.660518
2	-2.662302	-0.326611
3	-2.605790	-0.582925
4	-2.683088	0.666354
...	...	...
145	2.232927	0.385906
146	1.962246	-0.904731
147	1.928766	0.266655
148	1.808983	1.013922
149	1.445207	-0.028112

```
In [6]: df_final.columns=['A','B']
print("The principal comoponents of the data are")
df_final
```

Out[6]: The principal comoponents of the data are

	A	B
0	-2.576198	0.498650
1	-2.419715	-0.660518
2	-2.662302	-0.326611
3	-2.605790	-0.582925
4	-2.683088	0.666354
...	...	...
145	2.232927	0.385906
146	1.962246	-0.904731
147	1.928766	0.266655
148	1.808983	1.013922
149	1.445207	-0.028112

```
In [7]: import matplotlib.pyplot as plt
plt.scatter(df_final['A'],df_final['B'],color='red')
```

```
Out[7]: <matplotlib.collections.PathCollection at 0x15fdf0d10>
```

