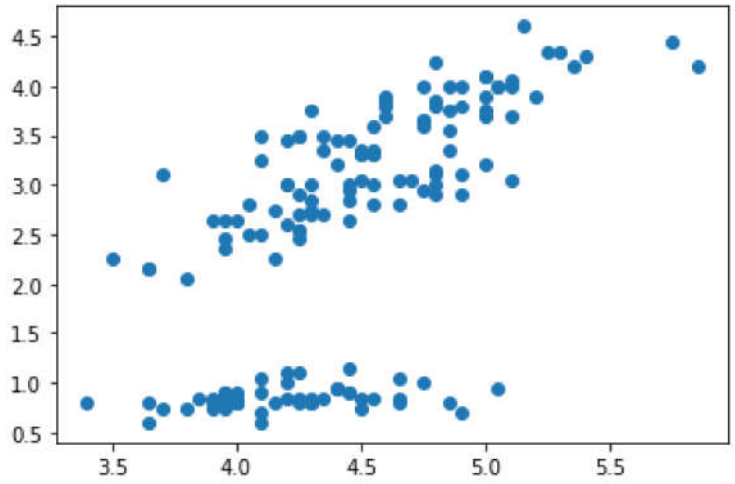
In [53]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import sklearn
from sklearn.cluster import KMeans
from sklearn import preprocessing
%matplotlib inline
```

In [49]:
```
df = pd.read_csv("ClusterPlot.csv", usecols=[1,2])
df

plt.scatter(df['V1'], df['V2'])
plt.show()
```

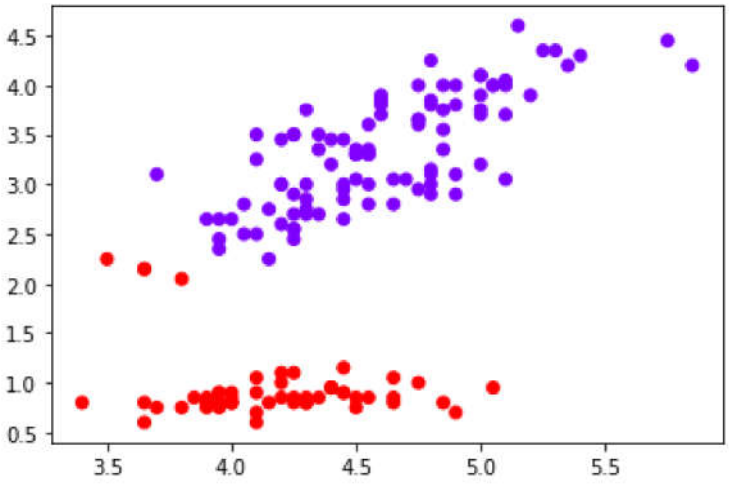


In [50]:
```
x=df.copy()
kmeans=KMeans(2)
kmeans.fit(x)
```

Out[50]:
```
KMeans(n_clusters=2)
```

In [51]:
```
cluster=x.copy()
cluster['cluster_pred']=kmeans.fit_predict(x)
```

In [52]:
```
plt.scatter(cluster['V1'], cluster['V2'], c=cluster['cluster_pred'], cmap='rainbow')
plt.show()
```



In [54]:
```
x_scaled= preprocessing.scale(x)
x_scaled
```

Out[54]:
```
array([[-3.34555875e-01, -1.34331530e+00],
       [-1.12218742e+00, -1.34331530e+00],
       [-1.12218742e+00, -1.38332667e+00],
       [-1.34722501e+00, -1.30330392e+00],
       [-3.34555875e-01, -1.34331530e+00],
       [ 4.53075670e-01, -1.14325841e+00],
       [-1.00966863e+00, -1.30330392e+00],
       [-5.59593460e-01, -1.30330392e+00],
       [-1.79730017e+00, -1.34331530e+00],
       [-1.00966863e+00, -1.34331530e+00],
       [ 2.28038086e-01, -1.30330392e+00],
       [-7.84631044e-01, -1.26329254e+00],
       [-1.23470621e+00, -1.38332667e+00],
       [-1.79730017e+00, -1.50336080e+00],
       [ 1.01566963e+00, -1.42333805e+00],
       [ 1.35322601e+00, -1.22328116e+00],
       [ 4.53075670e-01, -1.30330392e+00],
       [-3.34555875e-01, -1.30330392e+00],
       [ 6.78113254e-01, -1.18326979e+00],
       [ 3.00050113e-03, -1.26329254e+00],
       [-1.09518291e-01, -1.22328116e+00],
       [-1.09518291e-01, -1.22328116e+00],
       [-7.84631044e-01, -1.50336080e+00],
       [-5.59593460e-01, -1.10324703e+00],
       [-7.84631044e-01, -1.14325841e+00],
       [-1.00966863e+00, -1.26329254e+00],
       [-5.59593460e-01, -1.18326979e+00],
       [-2.22037083e-01, -1.30330392e+00],
       [-3.34555875e-01, -1.34331530e+00],
       [-1.12218742e+00, -1.26329254e+00],
       [-1.12218742e+00, -1.26329254e+00],
       [-1.09518291e-01, -1.22328116e+00],
       [ 4.53075670e-01, -1.34331530e+00],
       [ 9.03150839e-01, -1.34331530e+00],
       [-1.00966863e+00, -1.34331530e+00],
       [-7.84631044e-01, -1.42333805e+00],
       [ 1.15519293e-01, -1.38332667e+00],
       [-1.00966863e+00, -1.34331530e+00],
       [-1.68478138e+00, -1.38332667e+00],
       [-4.47074668e-01, -1.30330392e+00],
       [-4.47074668e-01, -1.34331530e+00],
       [-2.35989414e+00, -1.34331530e+00],
       [-1.45974380e+00, -1.38332667e+00],
       [-4.47074668e-01, -1.10324703e+00],
       [ 3.00050113e-03, -1.06323566e+00],
       [-1.23470621e+00, -1.30330392e+00],
       [ 3.00050113e-03, -1.26329254e+00],
       [-1.23470621e+00, -1.34331530e+00],
       [ 1.15519293e-01, -1.30330392e+00],
       [-6.72112252e-01, -1.34331530e+00],
       [ 1.46574480e+00,  4.57196667e-01],
       [ 7.90632047e-01,  4.17185290e-01],
       [ 1.24070722e+00,  5.77230798e-01],
       [-1.23470621e+00,  1.37105652e-01],
       [ 4.53075670e-01,  4.57196667e-01],
       [-4.47074668e-01,  3.37162536e-01],
       [ 7.90632047e-01,  5.37219421e-01],
```

```
[-1.79730017e+00, -2.63008118e-01],
[ 6.78113254e-01,  3.77173913e-01],
[-1.12218742e+00,  1.37105652e-01],
[-2.13485655e+00, -1.82985364e-01],
[ 3.00050113e-03,  2.97151159e-01],
[-7.84631044e-01,  1.70715208e-02],
[ 1.15519293e-01,  4.57196667e-01],
[-4.47074668e-01, -2.29398561e-02],
[ 1.01566963e+00,  3.37162536e-01],
[-3.34555875e-01,  4.17185290e-01],
[-4.47074668e-01,  5.70828978e-02],
[-5.59593460e-01,  4.17185290e-01],
[-8.97149837e-01,  1.70715208e-02],
[ 2.28038086e-01,  6.57253552e-01],
[ 3.00050113e-03,  1.37105652e-01],
[-1.09518291e-01,  5.77230798e-01],
[ 3.00050113e-03,  3.77173913e-01],
[ 4.53075670e-01,  2.57139782e-01],
[ 7.90632047e-01,  3.37162536e-01],
[ 7.90632047e-01,  4.97208044e-01],
[ 9.03150839e-01,  6.97264929e-01],
[ 3.00050113e-03,  4.17185290e-01],
[-6.72112252e-01, -1.82985364e-01],
[-1.12218742e+00, -2.29398561e-02],
[-1.12218742e+00, -1.02962610e-01],
[-4.47074668e-01,  5.70828978e-02],
[-2.22037083e-01,  6.97264929e-01],
[-5.59593460e-01,  4.17185290e-01],
[ 5.65594462e-01,  4.57196667e-01],
[ 1.01566963e+00,  4.97208044e-01],
[-3.34555875e-01,  2.97151159e-01],
[-3.34555875e-01,  1.77117029e-01],
[-1.00966863e+00,  1.37105652e-01],
[-8.97149837e-01,  2.57139782e-01],
[ 2.28038086e-01,  4.17185290e-01],
[-5.59593460e-01,  9.70942747e-02],
[-1.79730017e+00, -2.63008118e-01],
[-6.72112252e-01,  2.17128406e-01],
[-2.22037083e-01,  1.77117029e-01],
[-3.34555875e-01,  2.17128406e-01],
[ 2.28038086e-01,  2.57139782e-01],
[-1.45974380e+00, -3.43030872e-01],
[-4.47074668e-01,  1.77117029e-01],
[ 7.90632047e-01,  1.41746971e+00],
[-4.47074668e-01,  8.17299060e-01],
[ 1.35322601e+00,  1.21741283e+00],
[ 3.40556878e-01,  9.77344567e-01],
[ 6.78113254e-01,  1.21741283e+00],
[ 1.91581997e+00,  1.49749247e+00],
[-1.68478138e+00,  4.97208044e-01],
[ 1.46574480e+00,  1.25742421e+00],
[ 3.40556878e-01,  1.05736732e+00],
[ 2.14085755e+00,  1.45748109e+00],
[ 9.03150839e-01,  8.57310437e-01],
[ 2.28038086e-01,  8.97321814e-01],
[ 1.01566963e+00,  1.05736732e+00],
[-7.84631044e-01,  8.17299060e-01],
```

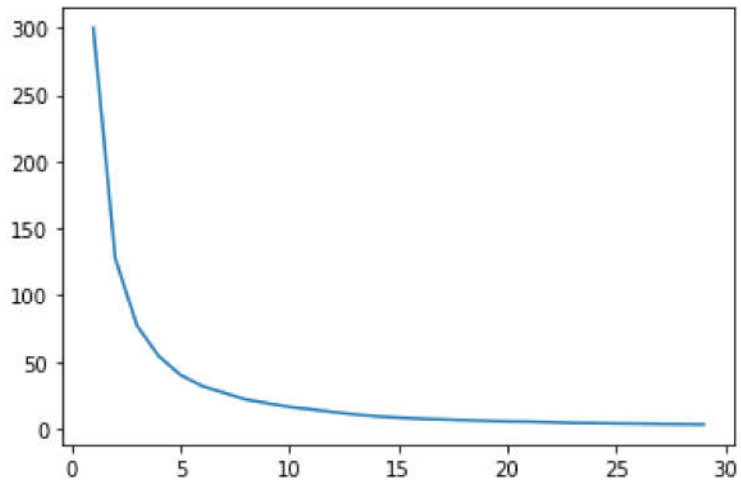
```
[-3.34555875e-01,  1.01735594e+00],
[ 7.90632047e-01,  1.05736732e+00],
[ 6.78113254e-01,  9.37333190e-01],
[ 2.92848910e+00,  1.57751522e+00],
[ 1.57826359e+00,  1.69754935e+00],
[-7.84631044e-01,  6.17242175e-01],
[ 1.35322601e+00,  1.21741283e+00],
[-5.59593460e-01,  7.77287683e-01],
[ 1.80330118e+00,  1.49749247e+00],
[ 1.15519293e-01,  6.97264929e-01],
[ 1.24070722e+00,  1.13739008e+00],
[ 1.69078238e+00,  1.13739008e+00],
[ 1.15519293e-01,  6.57253552e-01],
[ 2.28038086e-01,  6.97264929e-01],
[ 3.40556878e-01,  1.09737870e+00],
[ 1.46574480e+00,  9.77344567e-01],
[ 1.46574480e+00,  1.21741283e+00],
[ 3.15352668e+00,  1.37745834e+00],
[ 3.40556878e-01,  1.13739008e+00],
[ 2.28038086e-01,  6.57253552e-01],
[-2.22037083e-01,  8.17299060e-01],
[ 2.02833876e+00,  1.37745834e+00],
[ 9.03150839e-01,  1.21741283e+00],
[ 6.78113254e-01,  9.37333190e-01],
[ 1.15519293e-01,  6.57253552e-01],
[ 1.24070722e+00,  1.01735594e+00],
[ 1.01566963e+00,  1.21741283e+00],
[ 1.24070722e+00,  9.77344567e-01],
[-4.47074668e-01,  8.17299060e-01],
[ 1.24070722e+00,  1.29743558e+00],
[ 1.24070722e+00,  1.29743558e+00],
[ 9.03150839e-01,  1.01735594e+00],
[-1.09518291e-01,  7.77287683e-01],
[ 6.78113254e-01,  8.97321814e-01],
[ 7.90632047e-01,  1.09737870e+00],
[ 3.00050113e-03,  7.77287683e-01]])
```

In [55]:
```
wcss=[]

for i in range(1,30):
    kmeans=KMeans(i)
    kmeans.fit(x_scaled)
    wcss.append(kmeans.inertia_)
wcss
```

Out[55]:
```
[300.0,
 128.1313015023252,
 77.06914260423744,
 54.152376736977665,
 40.118991086518434,
 31.814834974518323,
 26.789102027884116,
 21.800865131669852,
 18.959864636770273,
 16.254466184641725,
 14.431789209617444,
 12.428719243873426,
 10.664685394048988,
 9.237580491185577,
 8.246354949925522,
 7.450670820606328,
 6.931529533675638,
 6.376718181634388,
 5.898706380637183,
 5.4485735463288965,
 5.315293897881143,
 4.800584026750359,
 4.298391956991633,
 4.130547438657414,
 3.895956352545999,
 3.7340879179510496,
 3.474939887527766,
 3.3621713317429536,
 3.1630450976059046]
```

In [56]:
```
plt.plot(range(1,30),wcss)
plt.show()
```



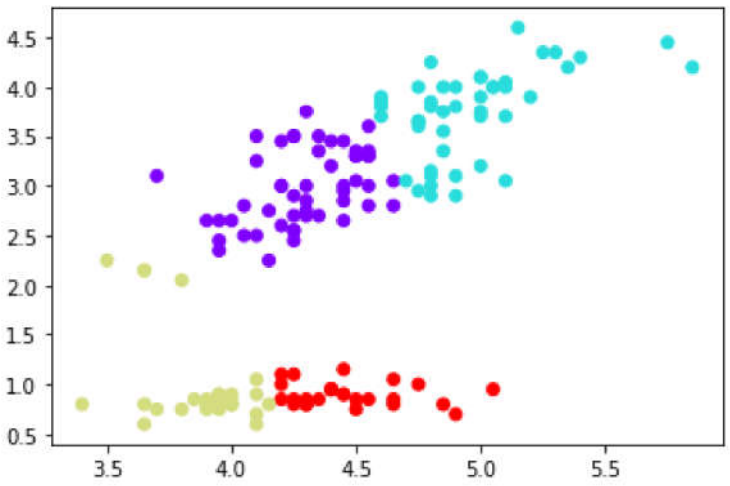
In [57]:
```
kmeans_new=KMeans(4)
kmeans.fit(x_scaled)
cluster_new=x.copy()
cluster_new['cluster_pred']=kmeans_new.fit_predict(x_scaled)
cluster_new
```

Out[57]:

	V1	V2	cluster_pred
0	4.30	0.80	3
1	3.95	0.80	2
2	3.95	0.75	2
3	3.85	0.85	2
4	4.30	0.80	3
...	...	...	...
145	4.85	3.75	1
146	4.40	3.45	0
147	4.75	3.60	1
148	4.80	3.85	1
149	4.45	3.45	0

150 rows × 3 columns

In [58]:
```
plt.scatter(cluster_new['V1'], cluster_new['V2'], c=cluster_new['cluster_pred'
], cmap='rainbow')
plt.show()
```



In [ ]:
```

```