

In [2]:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

In [40]:

```
data = pd.read_csv('C:/Users/DELL/Desktop/AI ML _ SCM/Assignment 2/Dataset.csv')
```

In [41]:

```
data.head()
```

Out[41]:

| S.No | District | Number of Positive Cases | Positivity Rate | Achievement towards 1st Dosage Covishield 18+ | Achievement towards 2nd Dosage Covishield 18+ | Total Achievement towards 1st Dose of Covishield and Covaxin | Total Achievement towards 2nd Dose of Covishield and Covaxin |
|------|----------|--------------------------|-----------------|---|---|--|--|
| 0 | 1 | Ariyalur | 16 | 1.6 | 93878 | 6169 | 196727 |
| 1 | 2 | Chennai | 183 | 0.8 | 662095 | 89809 | 2542245 |
| 2 | 3 | Coimbatore | 205 | 1.8 | 418387 | 19173 | 1170289 |
| 3 | 4 | Cuddalore | 57 | 1.3 | 310026 | 6333 | 612752 |
| 4 | 5 | Dharmapuri | 21 | 1.1 | 229475 | 4172 | 446398 |

In [79]:

```
X = data.iloc[:, [3,5]].values
```

In [80]:

X

Out[80]:

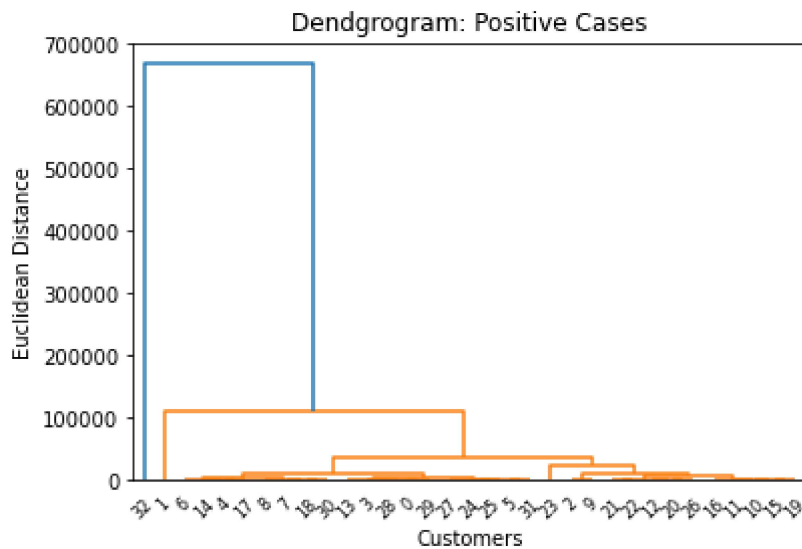
```
array([[1.60000e+00, 6.16900e+03],
       [8.00000e-01, 8.98090e+04],
       [1.80000e+00, 1.91730e+04],
       [1.30000e+00, 6.33300e+03],
       [1.10000e+00, 4.17200e+03],
       [2.00000e-01, 8.83000e+03],
       [1.50000e+00, 3.81400e+03],
       [7.00000e-01, 2.98200e+03],
       [5.00000e-01, 2.47800e+03],
       [7.00000e-01, 1.90620e+04],
       [6.00000e-01, 1.36020e+04],
       [1.00000e+00, 1.43210e+04],
       [2.00000e-01, 1.05880e+04],
       [1.80000e+00, 5.51700e+03],
       [2.70000e+00, 4.45200e+03],
       [1.50000e+00, 1.34870e+04],
       [1.40000e+00, 1.53580e+04],
       [1.00000e+00, 4.10600e+03],
       [1.30000e+00, 3.36600e+03],
       [1.20000e+00, 1.33800e+04],
       [1.30000e+00, 1.11240e+04],
       [2.50000e+00, 1.19760e+04],
       [3.00000e-01, 1.24710e+04],
       [1.00000e+00, 2.99450e+04],
       [8.00000e-01, 8.45900e+03],
       [1.30000e+00, 8.50400e+03],
       [1.00000e+00, 1.12280e+04],
       [7.00000e-01, 7.54700e+03],
       [1.30000e+00, 6.48000e+03],
       [1.10000e+00, 5.90500e+03],
       [1.60000e+00, 3.22200e+03],
       [2.00000e-01, 9.03400e+03],
       [1.00000e+00, 4.91982e+05]])
```

In [81]:

```
import scipy.cluster.hierarchy as sch
```

In [82]:

```
ddgrm = sch.dendrogram(sch.linkage(X, method='ward'))  
plt.title('Dendrogram: Positive Cases')  
plt.xlabel('Customers')  
plt.ylabel('Euclidean Distance')  
plt.show()
```



In [83]:

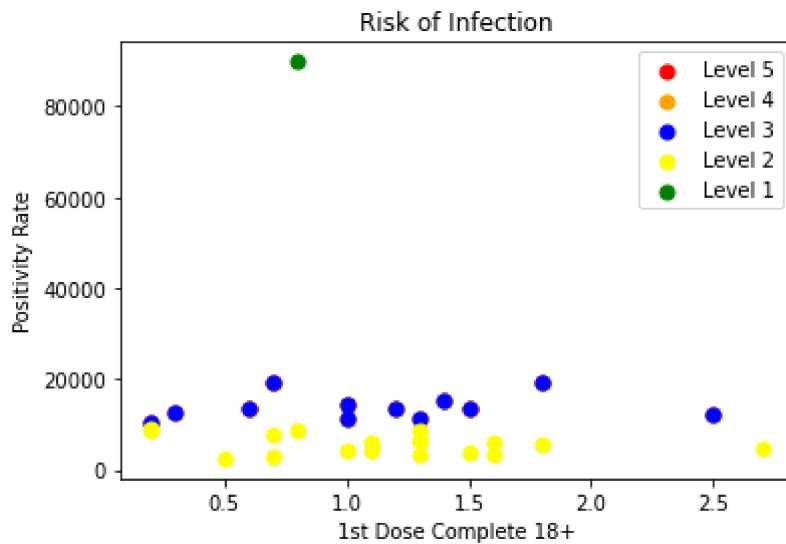
```
from sklearn.cluster import AgglomerativeClustering  
hc = AgglomerativeClustering(n_clusters=5, affinity='euclidean', linkage='ward')  
y_hc = hc.fit_predict(X)
```

In [85]:

```

plt.scatter(X[y_hc==0, 0], X[y_hc==0, 1], s=50, c= 'red', label='Level 5')
plt.scatter(X[y_hc==0, 0], X[y_hc==0, 1], s=50, c= 'orange', label='Level 4')
plt.scatter(X[y_hc==0, 0], X[y_hc==0, 1], s=50, c= 'blue', label='Level 3')
plt.scatter(X[y_hc==1, 0], X[y_hc==1, 1], s=50, c= 'yellow', label='Level 2')
plt.scatter(X[y_hc==2, 0], X[y_hc==2, 1], s=50, c= 'green', label='Level 1')
plt.title('Risk of Infection')
plt.xlabel('1st Dose Complete 18+')
plt.ylabel('Positivity Rate')
plt.legend()
plt.show()

```



In []: