

# Silhouette Homework

## Silhouette Analysis Using Manhattan Distance

Given the clustering:

- **Cluster 1:** {a, b}
- **Cluster 2:** {c, d, e}

**Data Points:**

Pattern	x	y
a	0.8	0.7
b	0.9	0.8
c	0.6	0.6
d	0.0	0.2
e	0.2	0.1

**Manhattan Distance Formula:**

$$d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$$

---

## a) Silhouette Score for Each Instance

The Silhouette score for each instance is calculated using the formula:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Where:

- $a(i)$  = Average distance to all other points in the same cluster.
- $b(i)$  = Average distance to all points in the nearest different cluster.

## Step 1: Compute Pairwise Manhattan Distances

	a	b	c	d	e
a	-	0.2	0.3	1.3	1.2
b	0.2	-	0.5	1.5	1.4
c	0.3	0.5	-	1.0	0.9
d	1.3	1.5	1.0	-	0.3
e	1.2	1.4	0.9	0.3	-

**Step 2: Calculate  $a(i)$  and  $b(i)$**

Pattern	$a(i)$	$b(i)$	$s(i)$
a	$d(a, b) = 0.2$	$\frac{0.3+1.3+1.2}{3} = 0.9333$	$\frac{0.9333-0.2}{0.9333} \approx 0.7857$
b	$d(b, a) = 0.2$	$\frac{0.5+1.5+1.4}{3} = 1.1333$	$\frac{1.1333-0.2}{1.1333} \approx 0.8235$
c	$\frac{1.0+0.9}{2} = 0.95$	$\frac{0.3+0.5}{2} = 0.4$	$\frac{0.4-0.95}{0.95} \approx -0.5789$
d	$\frac{1.0+0.3}{2} = 0.65$	$\frac{1.3+1.5}{2} = 1.4$	$\frac{1.4-0.65}{1.4} \approx 0.5357$
e	$\frac{0.9+0.3}{2} = 0.6$	$\frac{1.2+1.4}{2} = 1.3$	$\frac{1.3-0.6}{1.3} \approx 0.5385$

**Silhouette Scores for Each Instance:**

Pattern	$s(i)$
a	0.7857
b	0.8235
c	-0.5789
d	0.5357
e	0.5385

## b) Silhouette Score for Each Cluster

The Silhouette score for a cluster is the average of the Silhouette scores of its members.

- **Cluster 1:** {a, b}

$$s_{\text{Cluster 1}} = \frac{0.7857 + 0.8235}{2} = 0.8046$$

- **Cluster 2:** {c, d, e}

$$s_{\text{Cluster 2}} = \frac{-0.5789 + 0.5357 + 0.5385}{3} \approx 0.1651$$

### Silhouette Scores for Each Cluster:

Cluster	Silhouette Score
1	0.8046
2	0.1651

## c) Silhouette Score for the Entire Clustering

The overall Silhouette score is the average of all individual Silhouette scores.

$$s_{\text{Total}} = \frac{0.7857 + 0.8235 - 0.5789 + 0.5357 + 0.5385}{5} \approx 0.4209$$

**Overall Silhouette Score: 0.4209**

## d) Silhouette Visualization

Below is a sketch of the Silhouette plot for the given clustering. Each bar represents the Silhouette score of an instance.



