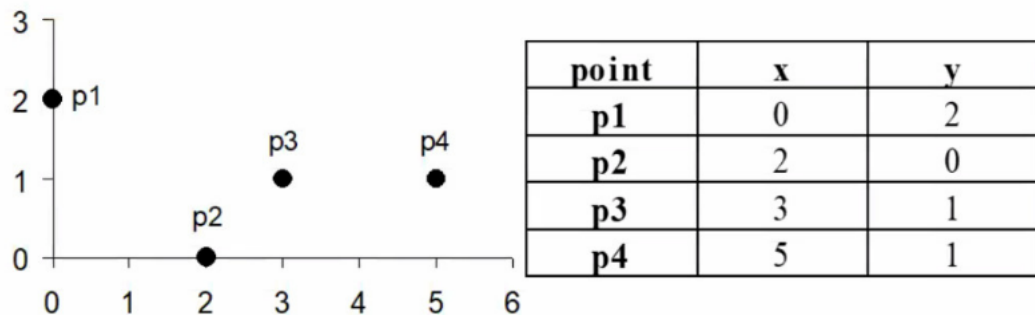


| Họ tên | MSSV | Lớp |
|----------------|----------|-----------|
| Hoàng Tiến Anh | 21068521 | ĐHKHDL17A |

Dựa vào bộ dữ liệu được xác định bởi các điểm như sau:



Sử dụng các phép đo khoảng cách Euclid và Minkowski với $r=1$ để tính toán khoảng cách giữa các điểm. (Minkowski với $r=1$ là Manhattan distance)

Link tham khảo công thức : [tại đây](#)

```
In [30]: # point x y
point = {
    'p1': [0, 2],
    'p2': [2, 0],
    'p3': [3, 1],
    'p4': [5, 1],
}
```

```
In [41]: import pandas as pd

# chuyển dict thành dataframe với 3 cột là point, x, y
df = pd.DataFrame.from_dict(point, orient='index', columns=['x', 'y'])
df.index.name = 'point'
df
```

```
Out[41]:
```

| | x | y |
|----|---|---|
| p1 | 0 | 2 |
| p2 | 2 | 0 |
| p3 | 3 | 1 |
| p4 | 5 | 1 |

Cách 1 : Sử dụng một số thư viện tính khoảng cách có sẵn

```
In [51]: # Tính khoảng cách giữa các điểm theo công thức Euclid
# import euclidean_distances
from sklearn.metrics.pairwise import euclidean_distances
print("Khoảng cách giữa các điểm theo công thức Euclid")
pd.DataFrame(euclidean_distances(df), index=df.index, columns=df.index)
```

Khoảng cách giữa các điểm theo công thức Euclid

```
Out[51]:
```

| point | p1 | p2 | p3 | p4 |
|-------|----------|----------|----------|----------|
| p1 | 0.000000 | 2.828427 | 3.162278 | 5.099020 |
| p2 | 2.828427 | 0.000000 | 1.414214 | 3.162278 |
| p3 | 3.162278 | 1.414214 | 0.000000 | 2.000000 |
| p4 | 5.099020 | 3.162278 | 2.000000 | 0.000000 |

```
In [53]: # Tính khoảng cách giữa các điểm theo công thức Manhattan
from sklearn.metrics.pairwise import manhattan_distances
print("Khoảng cách giữa các điểm theo công thức Manhattan")
pd.DataFrame(manhattan_distances(df), index=df.index, columns=df.index)
```

Khoảng cách giữa các điểm theo công thức Manhattan

```
Out[53]:
```

| point | p1 | p2 | p3 | p4 |
|-------|-----|-----|-----|-----|
| p1 | 0.0 | 4.0 | 4.0 | 6.0 |
| p2 | 4.0 | 0.0 | 2.0 | 4.0 |
| p3 | 4.0 | 2.0 | 0.0 | 2.0 |
| p4 | 6.0 | 4.0 | 2.0 | 0.0 |

Cách 2 : Viết hàm tính khoảng cách

```
In [34]: # Sử dụng Euclidean distance để tính khoảng cách giữa 2 điểm
def euclidean_distance(point1, point2):
    return ((point1[0] - point2[0])**2 + (point1[1] - point2[1])**2)**0.5

# Sử dụng Minkowski distance để tính khoảng cách giữa 2 điểm r = 1
def minkowski_distance(point1, point2, r):
    return (abs(point1[0] - point2[0])**r + abs(point1[1] - point2[1])**r)**(1/r)

# Tính khoảng cách giữa các điểm
print('Euclidean distance')
for i in range(4):
    for j in range(i+1,4):
        print('p'+str(i+1)+'-p'+str(j+1)+':', euclidean_distance(point['p'+str(i+1)], point['p'+str(j+1)]))

print('Minkowski distance')
for i in range(4):
    for j in range(i+1,4):
        print('p'+str(i+1)+'-p'+str(j+1)+':', minkowski_distance(point['p'+str(i+1)], point['p'+str(j+1)], 1))

# Tạo thành bảng khoảng cách
distance_euclidean = []
for i in range(4):
    distance_euclidean.append([])
    for j in range(4):
        distance_euclidean[i].append(euclidean_distance(point['p'+str(i+1)], point['p'+str(j+1)]))

distance_minkowski = []
for i in range(4):
    distance_minkowski.append([])
    for j in range(4):
        distance_minkowski[i].append(minkowski_distance(point['p'+str(i+1)], point['p'+str(j+1)], 1))
```

```
distance_minkowski.append([])
for j in range(4):
    distance_minkowski[i].append(minkowski_distance(point['p'+str(i+1)], point[
```

```
Euclidean distance
p1-p2: 2.8284271247461903
p1-p3: 3.1622776601683795
p1-p4: 5.0990195135927845
p2-p3: 1.4142135623730951
p2-p4: 3.1622776601683795
p3-p4: 2.0
Minkowski distance
p1-p2: 4.0
p1-p3: 4.0
p1-p4: 6.0
p2-p3: 2.0
p2-p4: 4.0
p3-p4: 2.0
```

```
In [54]: import pandas as pd

euclidean = pd.DataFrame(distance_euclidean, columns=['p1', 'p2', 'p3', 'p4'], index=
minkowski = pd.DataFrame(distance_minkowski, columns=['p1', 'p2', 'p3', 'p4'], index=
euclidean.index.name = 'euclidean'
minkowski.index.name = 'minkowski'
```

```
In [36]: euclidean
```

```
Out[36]:
```

| | p1 | p2 | p3 | p4 |
|------------------|----------|----------|----------|----------|
| euclidean | | | | |
| p1 | 0.000000 | 2.828427 | 3.162278 | 5.099020 |
| p2 | 2.828427 | 0.000000 | 1.414214 | 3.162278 |
| p3 | 3.162278 | 1.414214 | 0.000000 | 2.000000 |
| p4 | 5.099020 | 3.162278 | 2.000000 | 0.000000 |

```
In [37]: minkowski
```

```
Out[37]:
```

| | p1 | p2 | p3 | p4 |
|------------------|-----|-----|-----|-----|
| minkowski | | | | |
| p1 | 0.0 | 4.0 | 4.0 | 6.0 |
| p2 | 4.0 | 0.0 | 2.0 | 4.0 |
| p3 | 4.0 | 2.0 | 0.0 | 2.0 |
| p4 | 6.0 | 4.0 | 2.0 | 0.0 |

Cách 3 : Tính toán trực tiếp không code

Khoảng Cách Euclid

Công thức khoảng cách Euclid giữa hai điểm $A(x_1, y_1)$ và $B(x_2, y_2)$ là:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Khoảng Cách Minkowski với r=1 (Khoảng Cách Manhattan)

Công thức Minkowski tổng quát :

$$d(A, B) = \sqrt[r]{(|x_2 - x_1|)^r + (|y_2 - y_1|)^r}$$

Công thức khoảng cách Manhattan giữa hai điểm $A(x_1, y_1)$ và $B(x_2, y_2)$ là:

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

Áp dụng công thức cho các cặp điểm p1, p2, p3, p4 :

| point | x | y |
|-------|---|---|
| p1 | 0 | 2 |
| p2 | 2 | 0 |
| p3 | 3 | 1 |
| p4 | 5 | 1 |

1. Khoảng cách giữa p1 và p2:

- Euclid: $\sqrt{(2-0)^2 + (0-2)^2} = \sqrt{4+4} = \sqrt{8} \approx 2.83$
- Manhattan: $|2-0| + |0-2| = 2+2 = 4$

2. Khoảng cách giữa p1 và p3:

- Euclid: $\sqrt{(3-0)^2 + (1-2)^2} = \sqrt{9+1} = \sqrt{10} \approx 3.16$
- Manhattan: $|3-0| + |1-2| = 3+1 = 4$

3. Khoảng cách giữa p1 và p4:

- Euclid: $\sqrt{(5-0)^2 + (1-2)^2} = \sqrt{25+1} = \sqrt{26} \approx 5.10$
- Manhattan: $|5-0| + |1-2| = 5+1 = 6$

4. Khoảng cách giữa p2 và p3:

- Euclid: $\sqrt{(3-2)^2 + (1-0)^2} = \sqrt{1+1} = \sqrt{2} \approx 1.41$
- Manhattan: $|3-2| + |1-0| = 1+1 = 2$

5. Khoảng cách giữa p2 và p4:

- Euclid: $\sqrt{(5-2)^2 + (1-0)^2} = \sqrt{9+1} = \sqrt{10} \approx 3.16$
- Manhattan: $|5-2| + |1-0| = 3+1 = 4$

6. Khoảng cách giữa p3 và p4:

- Euclid: $\sqrt{(5-3)^2 + (1-1)^2} = \sqrt{4+0} = \sqrt{4} = 2$
- Manhattan: $|5-3| + |1-1| = 2+0 = 2$

Như vậy, ta có bảng khoảng cách giữa các điểm như sau:

Theo khoảng cách Euclid:

| point | p1 | p2 | p3 | p4 |
|-------|------|------|------|------|
| p1 | 0 | 2.83 | 3.16 | 5.10 |
| p2 | 2.83 | 0 | 1.41 | 3.16 |
| p3 | 3.16 | 1.41 | 0 | 2.00 |
| p4 | 5.10 | 3.16 | 2.00 | 0 |

Theo khoảng cách Manhattan:

| point | p1 | p2 | p3 | p4 |
|-------|----|----|----|----|
| p1 | 0 | 4 | 4 | 6 |
| p2 | 4 | 0 | 2 | 4 |
| p3 | 4 | 2 | 0 | 2 |
| p4 | 6 | 4 | 2 | 0 |

In []: