

# Báo cáo tiến độ

Ngày 12/10-18/10

GVHD: Huỳnh Xuân Phụng

Sinh Viên thực hiện:

17110213 Trần Cao Quyền

17110181 Nguyễn Trọng Luật

## **Kế Hoạch Báo Cáo:**

1. Nội dung thực hiện trong tuần
  - Tiếp tục Xây dựng Giao Diện
  - Cải tiến thuật toán
  - Tìm hiểu thuật toán hierarchical clustering
2. Nội dung thực hiện trong tuần tới
  - Tiếp tục hoàn thiện giao diện app

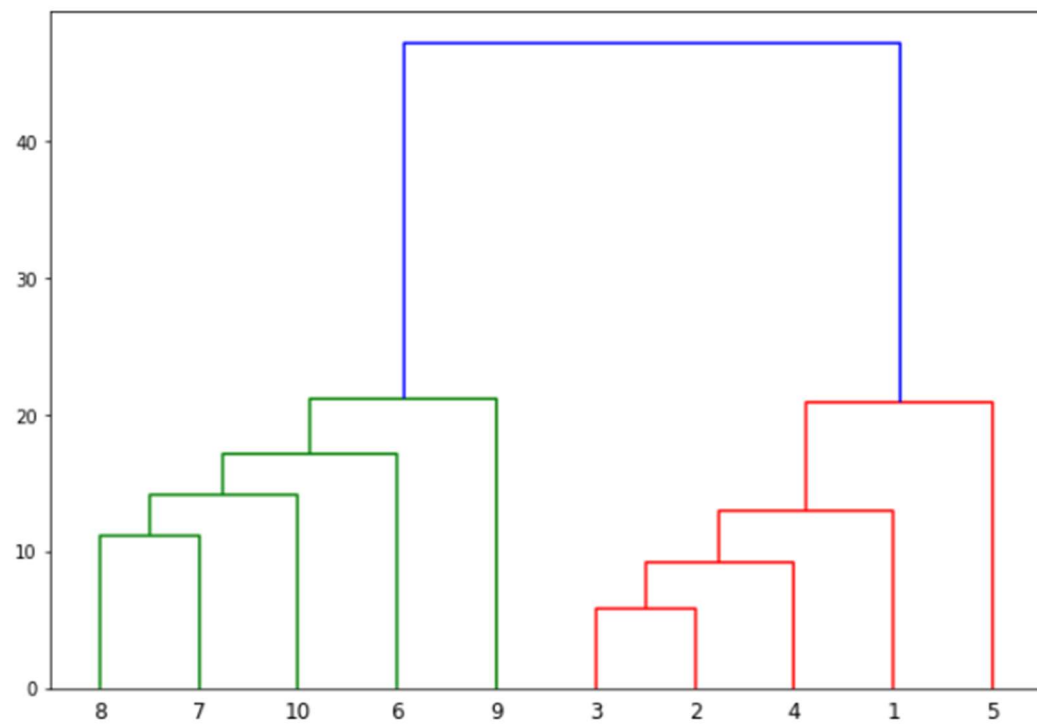
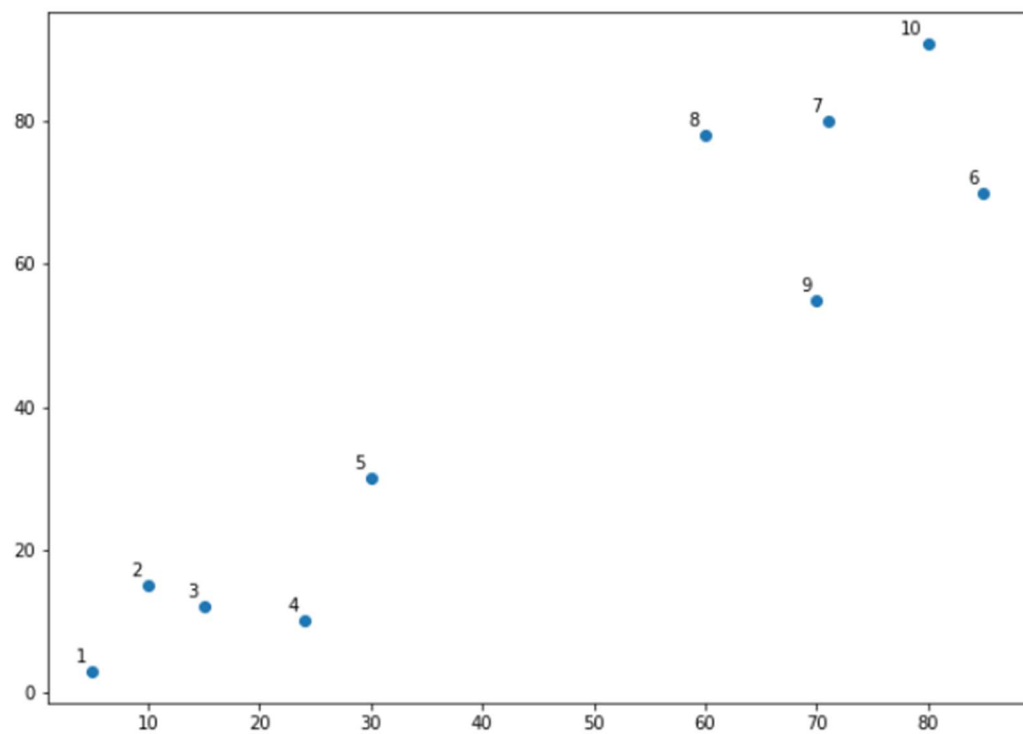
# Nội dung báo cáo:

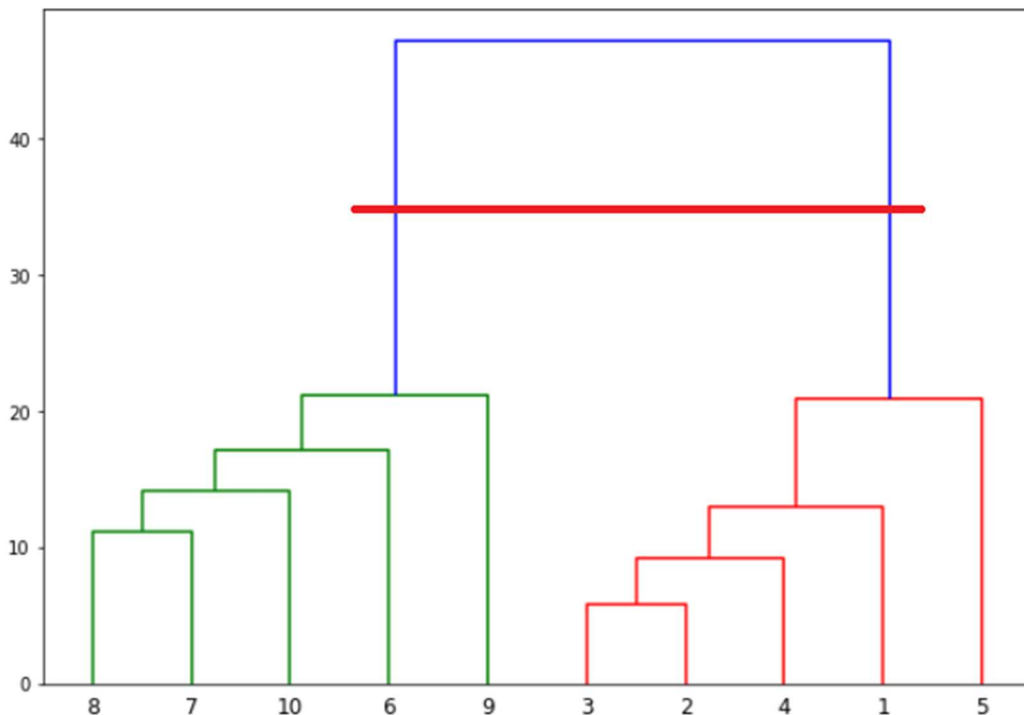
## 1. Thuật Toán hierarchical clustering

- là bài toán gom nhóm các đối tượng dữ liệu vào thành từng cụm (cluster) sao cho các đối tượng trong cùng một cụm có sự tương đồng theo một tiêu chí nào đó.
- Đặc điểm
  - Số cụm dữ liệu không được biết trước
  - Có nhiều các tiếp cận, mỗi cách lại có vài kỹ thuật
  - Các kỹ thuật khác nhau thường mang lại kết quả khác nhau.
- Độ đo khoảng cách

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$$

- Vd:  $v1(0, 1, 0, 1, 0)$  và  $v2(1, 1, 0, 1, 0)$   
 $\Rightarrow$  vậy  $d(v1, v2) = 1$
- Ứng dụng:
  - Dựa vào đây ta có thể ứng dụng vào bài toán giao hàng, dữ liệu đưa vào là đường, khu phố, phường, quận để làm dữ liệu đưa về giá trị 1 và 0, hoặc dựa vào kinh độ và vĩ độ của tọa độ. Từ đó ta tính được khoảng cách giữa tương đối giữa cái địa chỉ tạo thành 1 cụm tất cả đơn hàng. Từ cụm lớn đó ta có thể cắt nhỏ thành từng cụm sao cho phù hợp với tiêu chí hợp lý





- Ưu điểm:
  - Tổng hợp được tất cả các đơn mà không bỏ sót đơn hàng
  - Dựa vào thống kê có thể phân bố đơn hàng phù hợp nhất, tránh trường hợp có shipper giao quá nhiều đơn và có shipper giao ít đơn.
- Nhược điểm
  - Mỗi đơn hàng chỉ tồn tại trong 1 nhánh, trong trường hợp nhánh có nhiều đơn hàng mà quá số lượng vận chuyển thì phải cắt nhỏ cụm ra thêm 1 lần nữa dẫn đến các nhánh khác cũng phân cụm theo.

