Tên: Huỳnh Lan Nhi

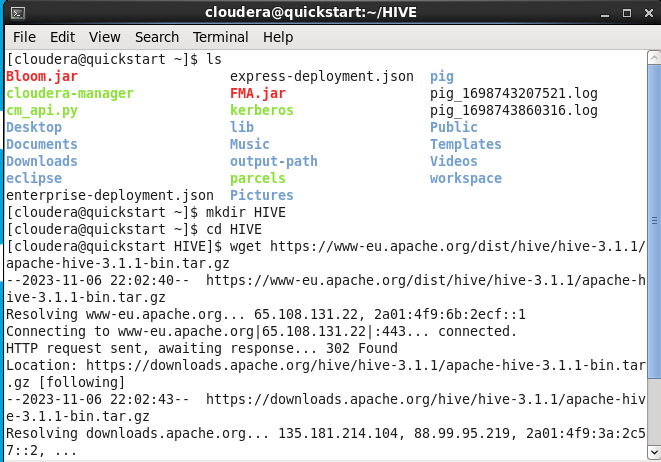
Lớp: 10\_ĐH\_CNTT1

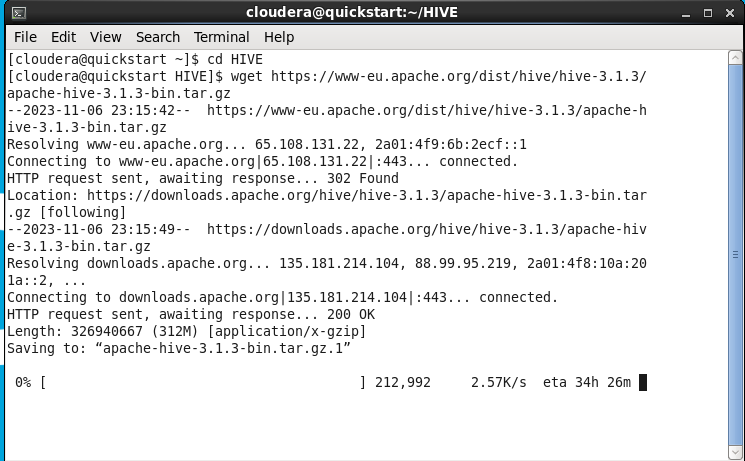
MSSV: 1050080028

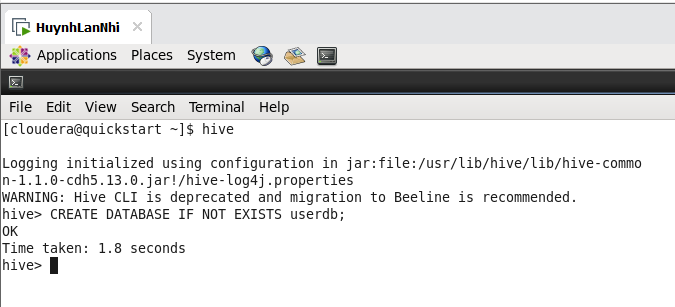
LAB 8

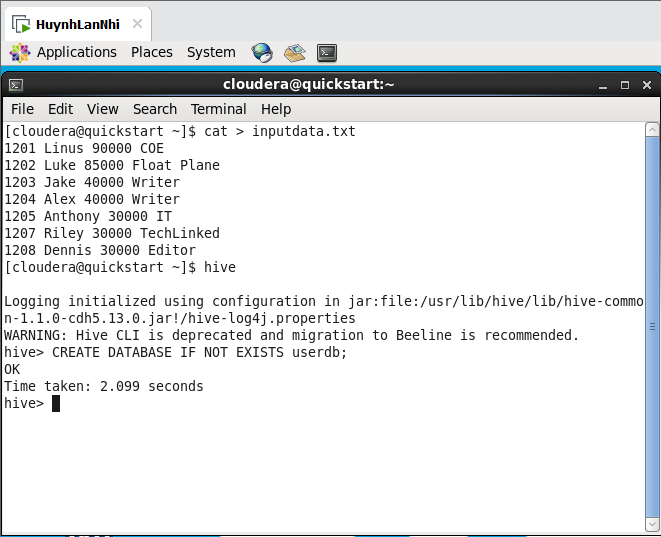
Mở terminal trên cloudera quickstart

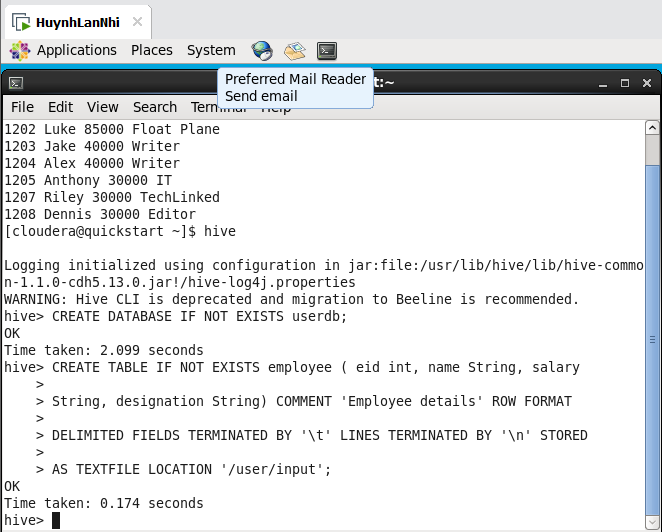
Sau đó thực hiện các câu lệnh để tải HIVE

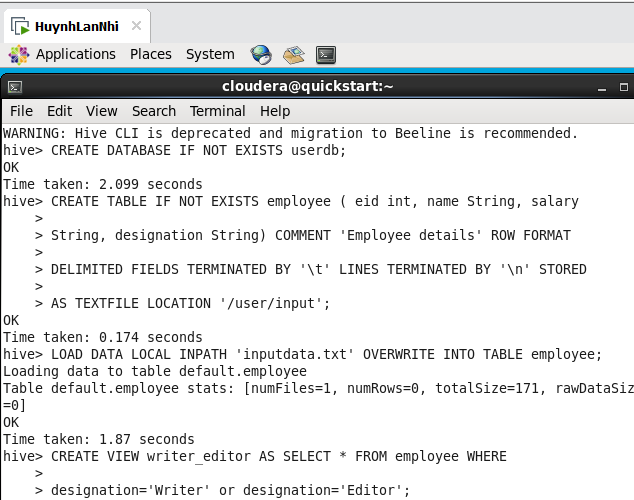






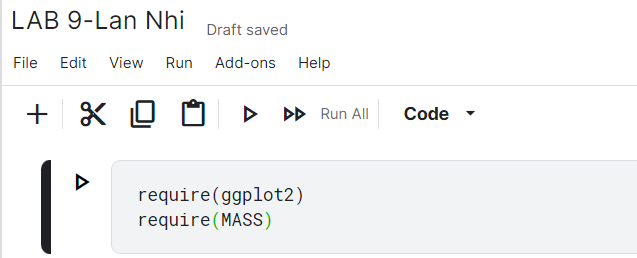




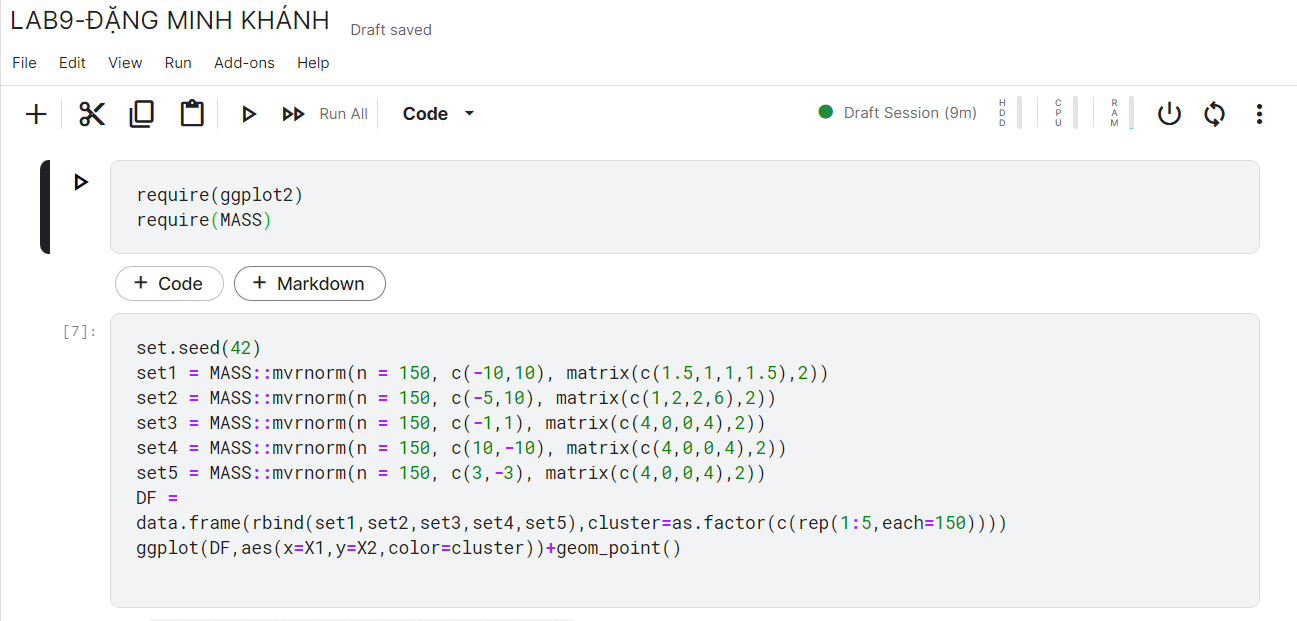


LAB 9

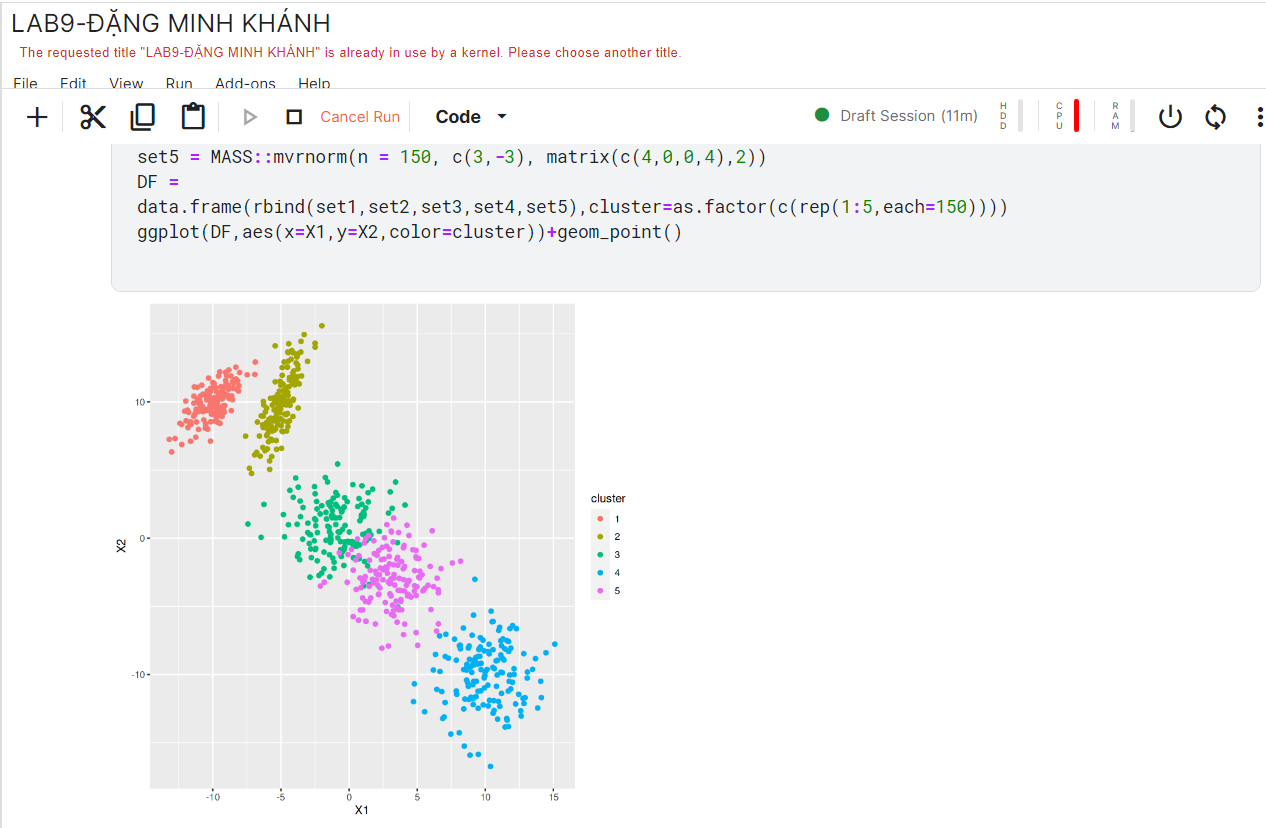
Chèn các gói thư viện vào môi trường hiện tại



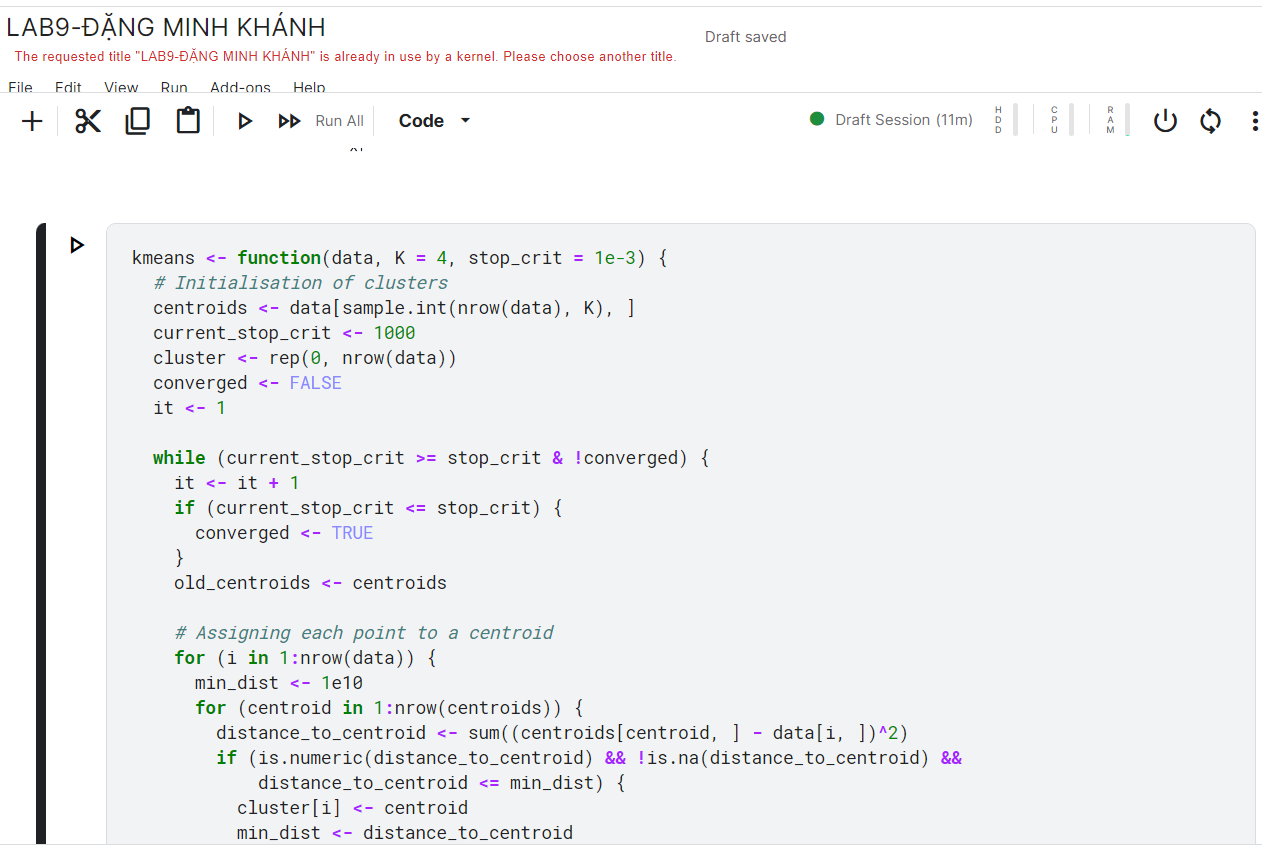
Tạo dữ liệu

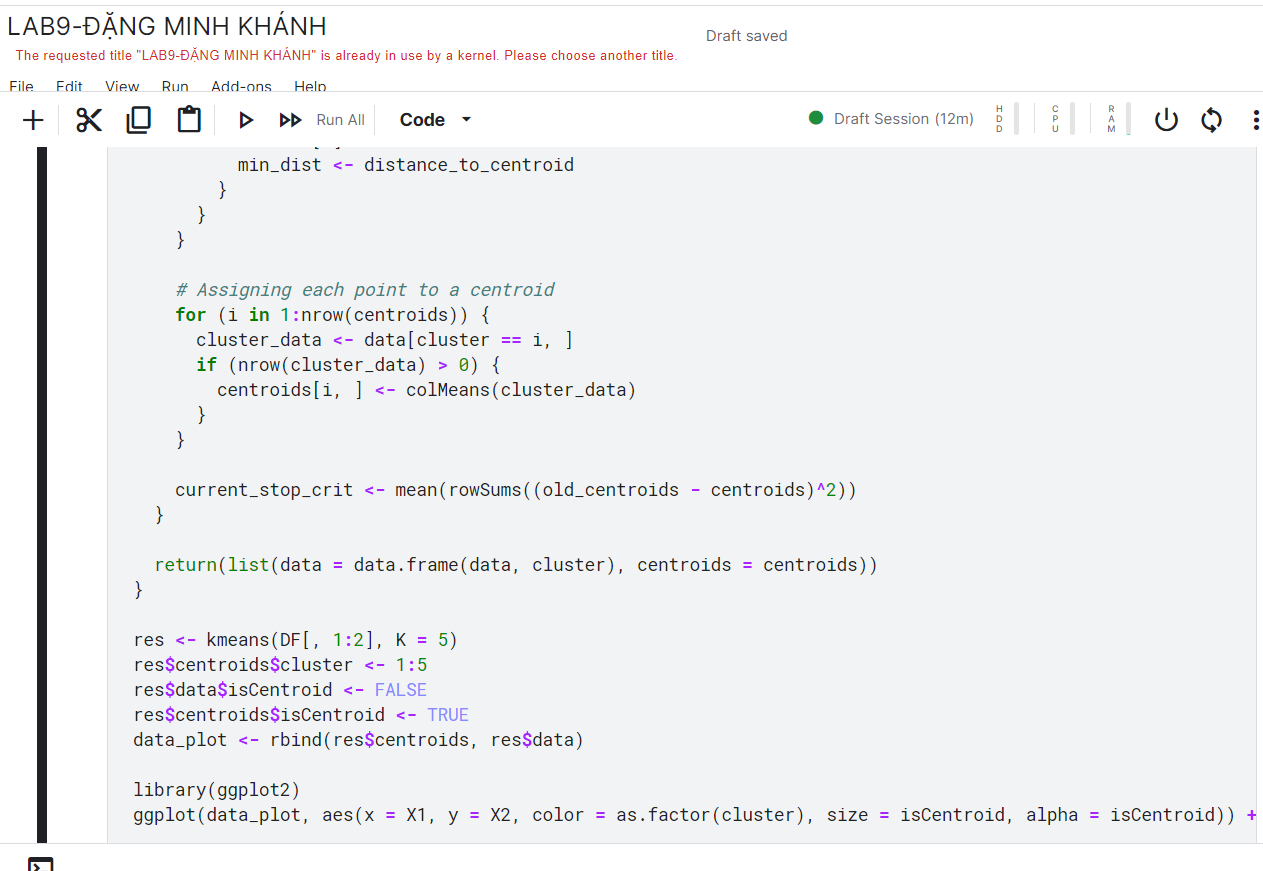


Sau khi thực hiện các câu lệnh trên sẽ ra kết quả



Thực hiện thuật toán K-Means clustering để phân cụm dữ liệu thành các nhóm dựa trên khoảng cách giữa các điểm và centroids.





Sau khi thực hiện được đoạn code trên sẽ hiện ra kết quả:

