Phân tích số liệu và biểu đồ thường được tiến hành bằng các phần mềm thông dụng như SAS, SPSS, Stata, Statistica, và S-Plus. Đây là những phần mềm được các công ti phần mềm phát triển và giới thiệu trên thị trường khoảng ba thập niên qua, và đã được các trường đại học, các trung tâm nghiên cứu và công ti kĩ nghệ trên toàn thế giới sử dụng cho giảng dạy và nghiên cứu. Nhưng vì chi phí để sửdụng các phần mềm này tuơng đối đắt tiền (có khi lên đến hàng trăm ngàn đô-la mỗi năm). Do đó, các nhà nghiên cứu thống kê trên thế giới đã hợp tác với nhau để phát triển một phần mềm mới, với chủ trương mã nguồn mở, sao cho tất cảcác thành viên trong ngành thống kê học và toán học trên thếgiới có thể sử dụng một cách thống nhất và hoàn toàn miễn phí.

Năm 1996, trong một bài báo quan trọng vềtính toán thống kê, hai nhà thống kê học Ross Ihaka và Robert Gentleman [lúc đó] thuộc Trường đại học Auckland, New Zealand phát hoạmột ngôn ngữmới cho phân tích thống kê mà họ đặt tên là R. Nói một cách ngắn gọn, R là một phần mềm sử dụng cho phân tích thống kê và vẽ biểu đồ. Thật ra, về bản chất, R là ngôn ngữ máy tính đa năng, có thể sửdụng cho nhiều mục tiêu khác nhau, từ tính toán đơn giản, toán học giải trí (recreational mathematics), tính toán ma trận (matrix), đến các phân tích thống kê phức tạp. Vì là một ngôn ngữ, cho nên người ta có thể sử dụng R để phát triển thành các phần mềm chuyên môn cho một vấn đề tính toán cá biệt.

Sơ lược đề tài: Dùng ngôn ngữ R phân tích các tuyến xe buýt và hành trình của mỗi tuyến xe buýt của Thành phố Hồ Chí Mình. Với data được lấy từ hộp đen của mỗi xe giúp chúng ta cho biết được các vấn đề của các tuyến xe buýt, hành trình của các tuyến xe buýt sẽ như thế nào. Qua đó giúp chúng ta ra rút ra các nhận xét để có thể tìm ra một giải pháp dự đoán các điểm kẹt xe để thông báo kịp thời nhằm điều tiết giao thông một cách hiệu quả.