The Commute Chronicles

An Analysis of Travel Efficiency, Weather Impacts, and Crowding

Opeoluwa Daniel Oyedeji

December 5, 2025

Contents

1 Introduction

2 Methods and Results

2.1 Exploratory Data Analysis

Load Data

```
df <- read.csv("bus_trip_data.csv")
df$date <- as.Date(df$date) # Convert date to a format suitable for R
```

The data has 10 columns and 28 rows. To ensure the table fits within the document margins, the data set is displayed below in two parts: the "To Campus" details and the "To Home" details.

date	day_of_week	dep_to_campus	arr_to_campus	dur_to_campus
2025-09-29	Monday	7:00:00 AM	9:18:00 AM	138
2025-10-01	Wednesday	7:00:00 AM	9:14:00 AM	134

dep_to_home	arr_to_home	dur_to_home	rain_to_home	crowded_unilink_to_home
3:08:00 PM	5:30:00 PM	142	FALSE	TRUE
12:00:00 PM	2:07:00 PM	127	TRUE	FALSE

Visual Inspection of Trends

Before running any complex statistical tests, we visualize the raw data to understand its shape. We are looking for two things: the **distribution** of the data (is it bell-shaped?) and any **trends** over time (is traffic getting worse?).

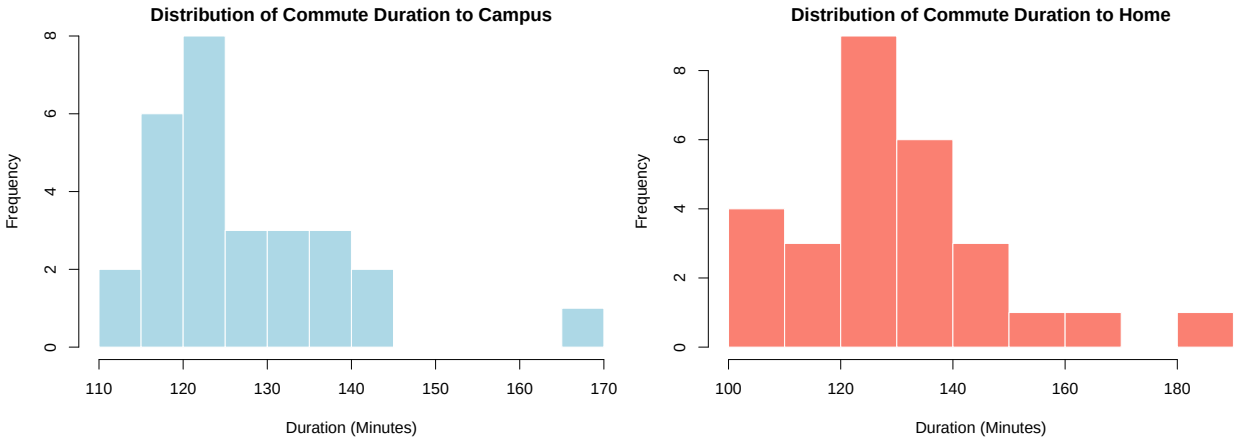

Figure 1: Histograms displaying the distribution of travel times for trips to campus (left) and trips to home (right) with 10-minute bin intervals.

An examination of the distributional shape highlights the following patterns:

- Both histograms exhibit a *positive skew* and a detached bar at the far right, indicating a *potential outlier* that may need to be addressed in our statistical tests.
- Both histograms are *unimodal* with the mode being around the *120-130 minute interval*.

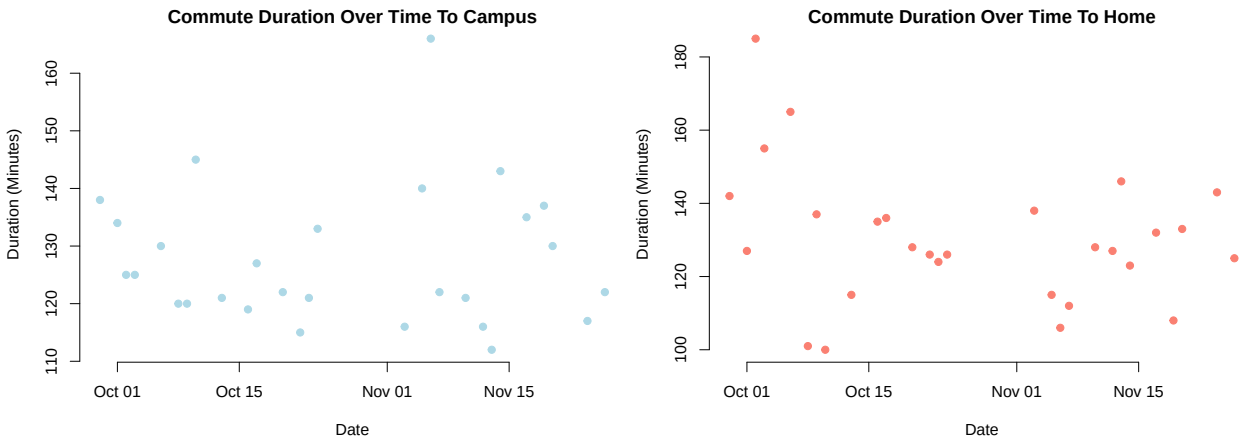

Figure 2: Scatter plots tracking daily commute durations for trips to campus (left) and return trips to home (right) over the observed period.

Plotting the commute durations over time highlights a stark contrast in stability between the two journeys:

- Trips to campus appear stable over time, showing only random noise without a noticeable increase or decrease in duration.
- Trips to home exhibit high variance, with widely dispersed data points, suggesting unpredictable delays that warrant further statistical testing.

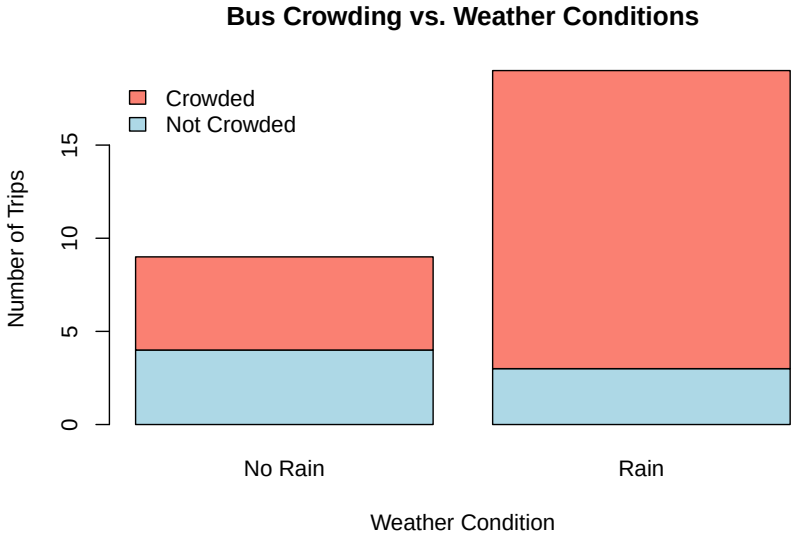

Figure 3: Stacked bar chart showing the proportion of crowded UniLink buses to home during clear weather versus rain.

An analysis of the interaction between weather and bus capacity suggests a strong correlation:

- **On clear days**, the distribution is balanced, with the bus being crowded roughly half the time.
- **On rainy days**, the vast majority of trips experience overcrowding, likely due to an increase in students choosing public transport over walking or cycling.