

MBTA

ROUTE 1



ROUTE - I





HARVARD SQUARE 

BUS

TO

NUBIAN STATION 

BUS

-  Public Transportation Route in Massachusetts
-  6 miles and 25 stops
-  Service Optimization
-  Route Planning



N
E
X
T

OVERVIEW

WHAT CREATES DEMAND FOR ROUTE 1

Can we predict the average passenger load on MBTA buses based on factors such as time period (e.g., morning rush, afternoon peak), route characteristics (e.g., average users, stop sequence), and other relevant variables.



N
E
X
T

INDEPENDENT VARIABLES



stop_sequence: The order of the stop in the bus route.



direction_id: Indicates the direction of the bus



day_type_id: A categorical variable identifying the type of day like weekdays, weekends, or holidays



day_type_name: Text description of the day type, such as "weekday".



average_offs: The average count of passengers alighting vehicles per trip.



time_period_id: Identifier for the time period according to the service delivery policy.



time_period_name: Descriptive name for the time period



stop_id: Unique identifier for each bus stop.



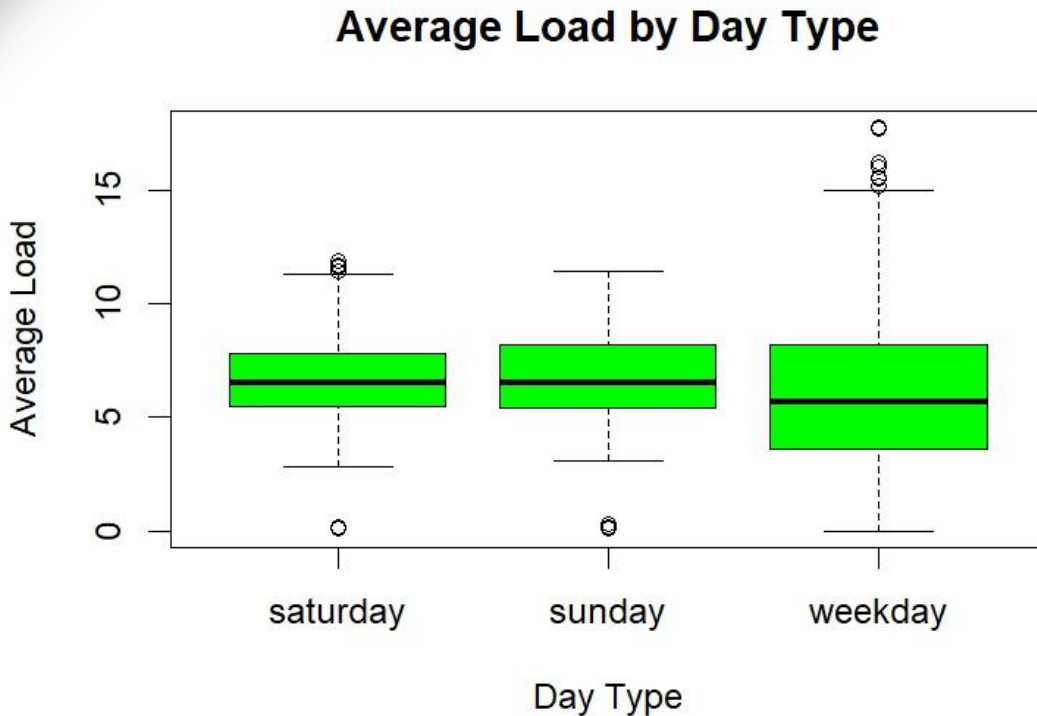
average_ons: The average count of passengers boarding vehicles per trip.



VISUALIZATION

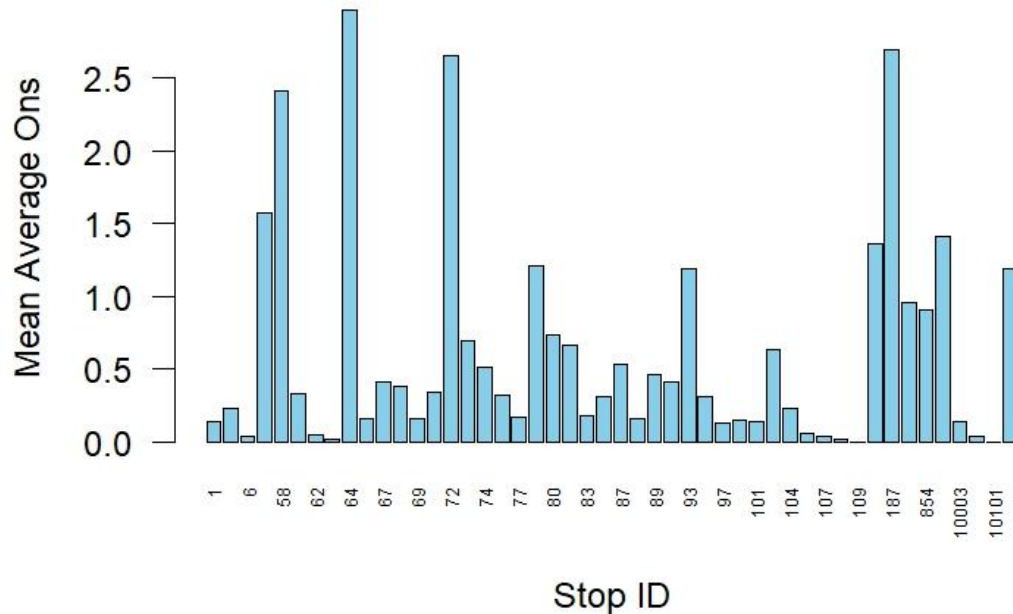
AVERAGE LOAD (DAY TYPE)

The "Average Load by Day Type" box plot reveals comparable median passenger counts across all days, with weekdays showing slightly higher variability, possibly due to commuter flux.



VISUALIZATION

Mean Average Ons by Stop ID



AVERAGE ONS (STOP ID)

The "Mean Average Ons by Stop ID" bar chart displays pronounced disparities in boardings at different stops, hinting at hubs of activity that might need enhanced transit services.

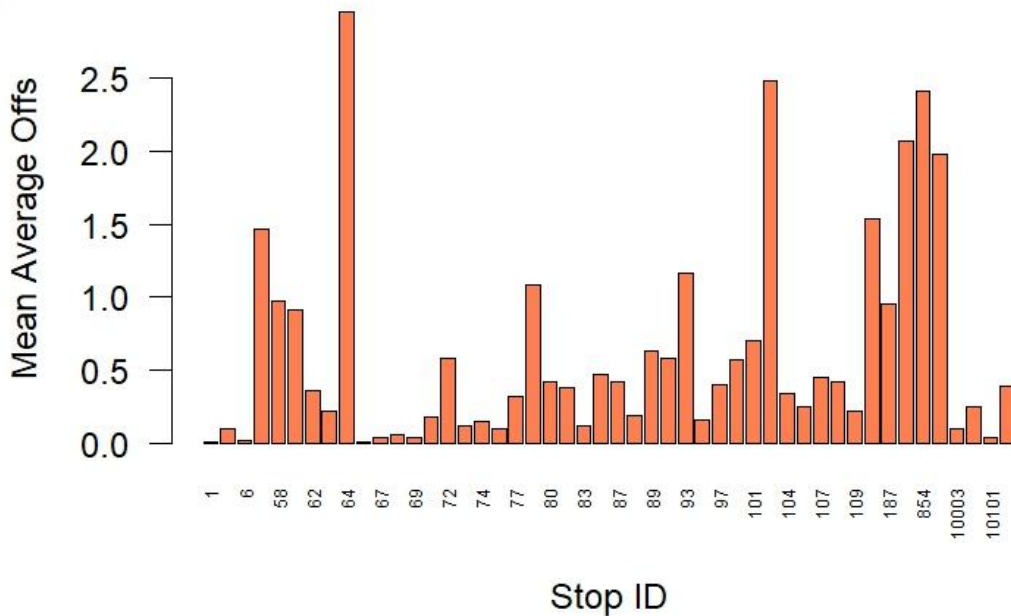


VISUALIZATION

AVERAGE OFFS (STOP ID)

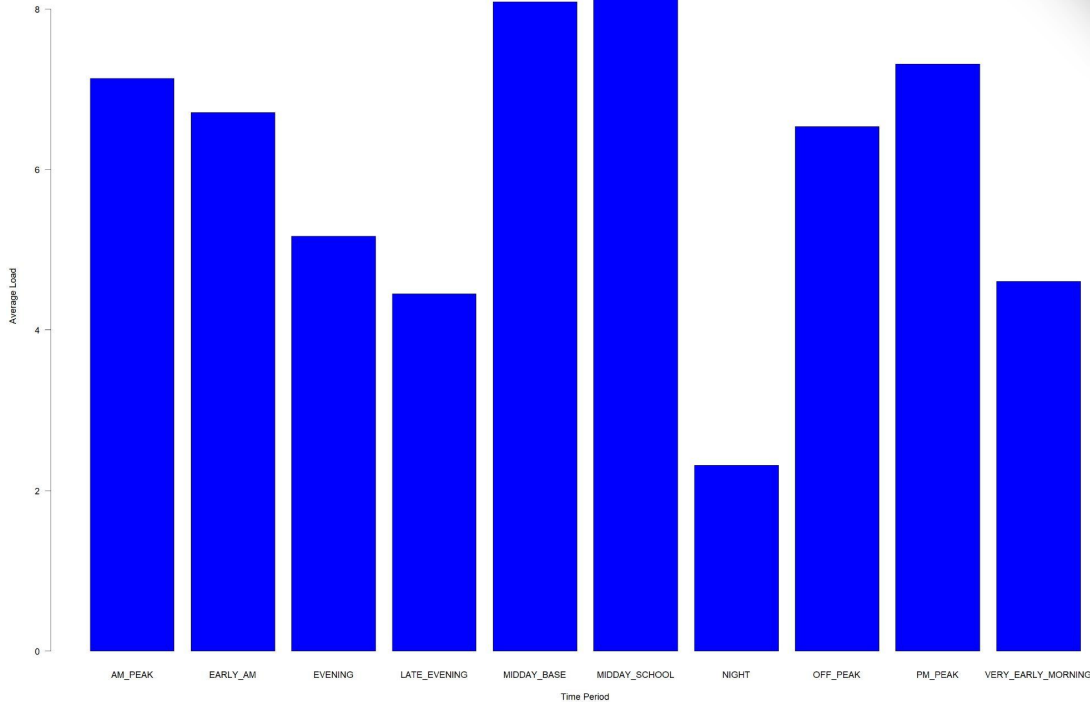
"Mean Average Offs by Stop ID" bar chart shows significant variance in alightings, with busier stops possibly aligning with key destinations, suggesting a need for improved infrastructure.

Mean Average Offs by Stop ID



VISUALIZATION

Average Load by Time Period



AVERAGE LOAD VS TIME PERIOD

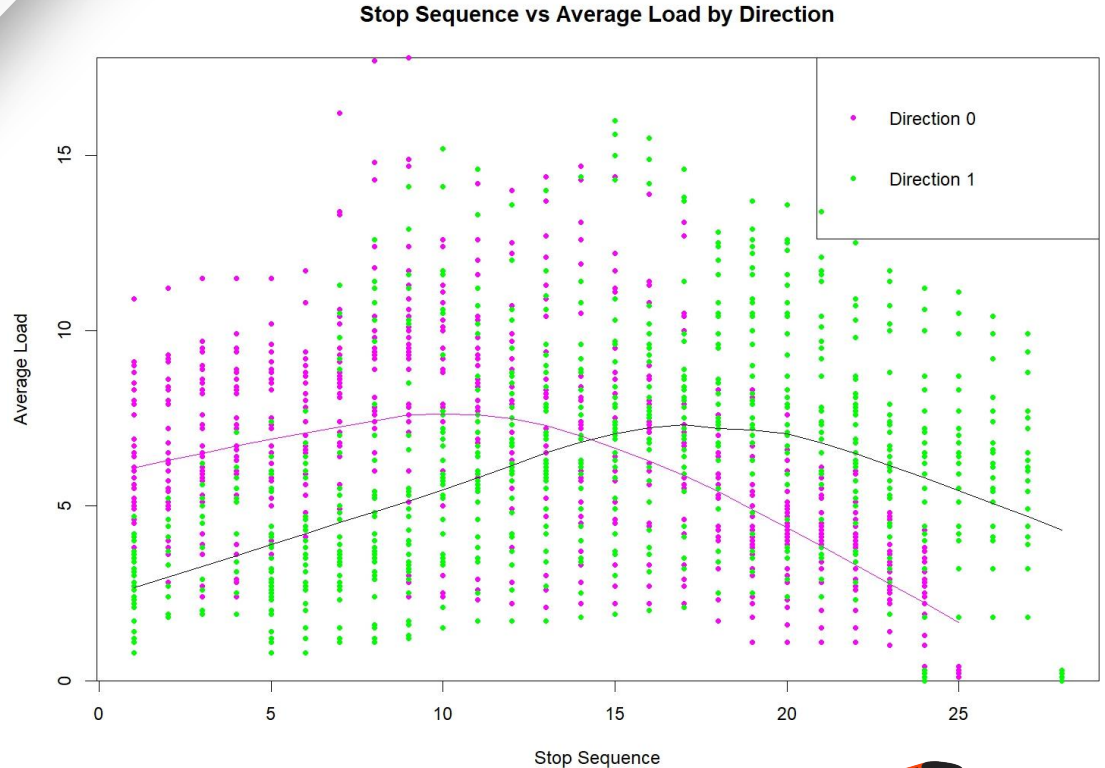
The "Average Load by Time Period" bar chart shows the passenger load's daily flow, peaking at Midday, which could inform targeted transit scheduling.



VISUALIZATION

AVERAGE LOAD VS STOP SEQUENCE (BY DIRECTION)

the "Stop Sequence vs Average Load by Direction" scatter plot suggests a mid-route rise and subsequent fall in passenger load, likely due to the density of the traversed areas.



FIRST ORDER MODEL

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n + \epsilon$$

- First_Order_Model <- *lm(average_load ~ time_period_name + day_type_id + average_ons + average_offs + stop_id, data = mbta_data_clean)*
- Multiple R-squared: 0.8046 Adjusted R-squared: 0.7973
- F-statistic: 109.4 p-value: < 2.2e-16
- Residual standard error: 1.44 on 1568 degrees of freedom

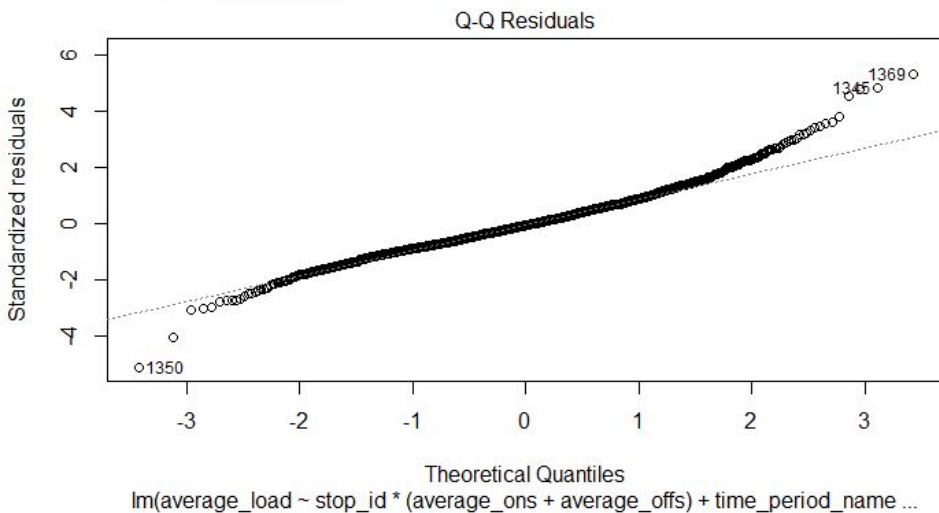
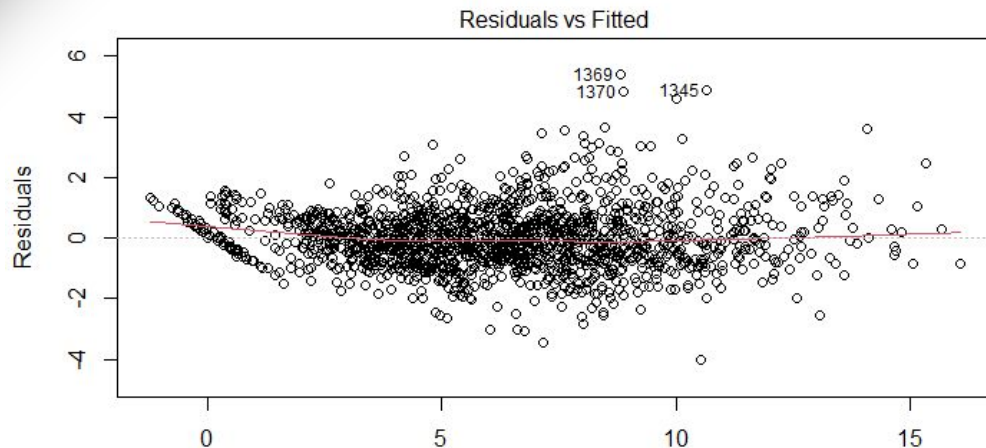


RECOMMENDED MODEL

- ❑ `Rec_Model <- lm(average_load ~ stop_id * (average_ons + average_offs) + time_period_name + day_type_id, data = mbta_data_clean)`
- ❑ Multiple R-squared: 0.9036
- ❑ Adjusted R-squared: 0.8938
- ❑ F-statistic: 91.64
- ❑ p-value: < 2.2e-16
- ❑ Residual standard error: 1.043 on 1476 degrees of freedom



MODEL PLOTS



Fitted values
lm(average_load ~ stop_id * (average_ons + average_offs) + time_period_name ...)



N
E
X
T

FUTURE DIRECTIONS

- ❑ Adjust Bus Frequencies
- ❑ Strategic Planning for Routes
- ❑ Allocate Resources Efficiently
- ❑ Infrastructure Investments



CLOSING

Takeaway : MBTA Users are more time based users, and we also understand that only certain stops along the route add up to the average passenger load.

