Heinz Systems Project Proposal

Spring 2018

Project Title: Criteria and Recommendation of Frequent Transit Service Network Revision

Client: Port Authority of Allegheny County, Planning Department

Other partners:

- Allegheny Conference on Community Development
- Allegheny County Economic Development

Project Setup: Joint MSPPM-MISM project

Project Description:

The Port Authority of Allegheny County provides public transportation services in Allegheny County and the City of Pittsburgh via 97 bus routes, 2 light rail lines, and 2 inclined planes. Most of these routes are operated as a hub and spoke model, with the majority of the transit services connecting near downtown Pittsburgh.

There are two main ways a transit agency can "grow" service: coverage and frequency. "Coverage" refers to extending or adding routes to cover more area (often to lower-density areas), and "frequency" refers to adding more trips to existing routes so that the wait times are reduced. One important goal for all transit agencies is ridership, but there are many other factors as well, such as cost-per-rider, equitable service, and supporting other land use, development, neighborhood, or economic development goals, such as connecting to key destinations like a community college or a housing development, among many others.

American cities have been seeing a general downward trend in transit ridership, especially on bus routes, over the last few years. Agencies that have focused on maximizing access to high frequency transit options, however, are bucking this trend. The Port Authority would like to understand how future investment or reorganization of services towards routes with frequent service could affect its ridership

In this project, the systems team will consider the existing service, ridership, and cost information of the current system, the Port Authority's Transit Service Guidelines, development and mobility trends in the region, population and other demographic data, and best practices from around the country and world to provide recommendations to the Authority on how this strategy might affect ridership. To do this, the team will develop criteria to set frequency goals and then evaluate existing routes based on those criteria to create a recommended frequent transit service network for Allegheny County. This analysis would include estimated changes in costs of service provision, number of people and jobs which could be accessed by the new network, and projected changes to ridership from these shifts in frequency, as well as an overall recommendation as to whether this focus is something the Authority should pursue further.

Frequent Transit Service Network Evaluation and Design

Joint With

PortAuthority

PORT AUTHORITY OF ALLEGHENY COUNTY
ALLEGHENY CONFERENCE ON COMMUNITY DEVELOPMENT
CARNEGIE MELLON UNIVERSITY HEINZ COLLEGE
PUBLIC POLICY AND MANAGEMENT
NGANI NDIMBIE
XU HAN
ZHI WANG
CARNEGIE MELLON UNIVERSITY HEINZ COLLEGE
INFORMATION SYSTEM AND MANAGEMENT
SHUAIJUN YE
CHEN WEI
FEI WANG

DATE: MAY 11, 2018

Table of Contents

Table of Contents 1
Executive Summary 5
Introduction 6
Literature Review and Research Gap 8
Data Sources and Data Description 15
Data Source 1: Automatic Vehicle Location and Automatic Passenger Counter 15
Data Source 2: General Transit Feed Specification 16
Data Source 3: American Community Survey Data, Census Data 16
Factor Identification 19
Social-Demographic Factors 19
Land-Use 19
Senior Percentage 20
Population Density 20
Unemployment Rate 21
Factor: Income Level 21
Job Density 22
College Student Enrollment 22
Vehicle Ownership 23
Methodology: GIS Analysis 23
Bus Performance Factors 25
Data Separation 25
Peak Time Selection 26
Stop Skipping 29
On-time performance 29
Factor: Passenger Waiting time 30
Factor: Crowding Level 31
Regression Analysis Model Development 32
Regression Dataset Preparation 32
Add interaction factors: Collinear Analysis 32
Feature Selection 33
Feature Selection Method 1: LASSO 33
Feature Selection Method 2: Best-subset 35

Model Validation: Random Forest Model 37
Ridership Elasticity Estimation 38
Results And Findings 39
GIS Findings - Demographic Route Profile, Choropleth Maps 39
Below Middle-Class Households 41
Population Density 42
Vehicle Ownership 43
Unemployment Rate 43
College or Above Student Percentage by Block Group 45
Job Density 46
Land Use 47
Sample Results of 48 Regression Models 48
Route 51 Outbound Off-Peak 48
Route 82 Inbound Peak 49
Implementations And Recommendations 50
Ridership Elasticity Estimation of Route 51 50
Findings and Recommendations for Route 51 and 82 52
Route 51 Outbound Weekday Off-Peak Model 53
Route 82 Inbound Weekday Off-Peak Model 53
Recommendations for Route 51 and 82 53
Conclusion and Future Work 54
Work Flow 54
Future Work Suggestions 55
Appendix: 57
Peak Time 57
Using average load per trip as y value. 57
Using total load in half an hour: 62
Regression Results 68
Route 82 Inbound Weekday Peak 68
Route 82 Inbound Weekday Off-Peak 69
Route 82 Inbound Saturday 69
Route 82 Inbound Sunday 71
Route 82 Outbound Weekday Peak 72
Route 82 Outbound Weekday Off-Peak 73
Route 82 Outbound Saturday 74

Route 82 Outbound Saturday 75
Route 8 Inbound Weekday Peak..76
Route 8 Inbound Weekday Off-Peak 77
Route 8 Inbound Saturday 78
Route 8 Inbound Sunday..79
Route 8 Outbound Weekday Peak..80
Route 8 Outbound Weekday Off-Peak 81
Route 8 Outbound Saturday 82
Route 8 Outbound Sunday..83
Route 12 Inbound Weekday Peak..84
Route 12 Inbound Weekday Off-Peak 85
Route 12 Inbound Saturday 86
Route 12 Inbound Sunday..87
Route 12 Outbound Weekday Peak..88
Route 12 Outbound Weekday Off-Peak 88
Route 12 Outbound Saturday 89
Route 12 Outbound Sunday..89
Route 16 Inbound Weekday Peak..90
Route 16 Inbound Weekday Off-Peak 91
Route 16 Inbound Saturday 92
Route 16 Inbound Sunday..93
Route 16 Outbound Weekday Peak..94
Route 16 Outbound Weekday Off-Peak 94
Route 16 Outbound Saturday 95
Route 16 Outbound Sunday..95
Route 51 Inbound Weekday Peak..96
Route 51 Inbound Weekday Off-Peak 97
Route 51 Inbound Saturday 98
Route 51 Inbound Sunday..99
Route 51 Outbound Weekday Peak..99
Route 51 Outbound Weekday Off-Peak 100
Route 51 Outbound Saturday 101
Route 51 Outbound Sunday..102
Route 91 Inbound Weekday Peak..103
Route 91 Inbound Weekday Off-Peak 103

Route 91 Inbound Saturday 104
Route 91 Inbound Sunday..104
Route 91 Outbound Weekday Off-Peak 105
Route 91 Outbound Saturday 106
Route 91 Outbound Sunday...106

Executive Summary

American cities have been seeing a general downward trend in transit ridership, especially on bus routes, over the last few years. Agencies that have focused on maximizing access to high-frequency transit options, conversely, are seeing stable or growing ridership. Port Authority of Allegheny County and Allegheny Conference would like to understand how future investment or reorganization of services towards bus routes with frequent service could affect ridership.

Our team used historical bus-performance data from Automatic Passenger Counting (APC) and Automatic Vehicle Location (AVL) technologies, schedule data from the General Transit Feed Specification (GTFS), and socio-demographic data from Census Bureau, to offer a systematic way for decision-makers to understand how transit service ridership would change with respect to frequency change.

Through the process of factors selection, data separation, model build-up, and model utilization, our team analyzed six existing high-frequency routes – 8, 12, 16, 51, 82, and 91 – to create 48 regression models and a sensitivity analysis tool for further implementation.

Recommendations:

In general, we proposed that PAAC examine each of the 48 models and follow the ridership estimation approach to revise the current frequency setting and bus schedule. For instance, as for route 51, we recommend to add more trips during weekday off-peak time and shift the current peak time window an hour earlier to match the observed demand.

Introduction

According to the 2016 American Public Transportation Association Fact Book, Pittsburgh is the city with the 26th largest public transit system in America[1]. In Allegheny County, The Port Authority of Allegheny County (PAAC) plays an important role in connecting people to their destinations. PAAC provides public transportation within a 775 square mile region within Allegheny County and the City of Pittsburgh via 97 bus routes, 3 light rail lines, and 2 funiculars.

Over the last few years, American cities have been seeing a general downtrend in transit ridership, especially on bus services. According to the Annual Service Report of The Port Authority of Allegheny County, in the fiscal year 2016, the total ridership in Pittsburgh was decreased by 2.1 percent from that of the fiscal year 2015, mainly due to lower gasoline prices and increased use of ridesharing services[2]. Thus, increasing or stabilizing the ridership becomes an important goal for PAAC.

There are many factors that influence ridership but bus frequency is one of the few factors that the agency controls. But, the relationship between frequency increase and ridership is complex and service reorganization can require a significant investment of limited agency resources. Thus, the Port Authority of Allegheny County and the Allegheny Conference on Community Development collaborated to ask our team to develop a tool that would enable ridership estimation to assist in service planning.

Our clients assigned six Port Authority routes to examine. The six chosen routes – 8, 12, 16, 51, 82, and 91 – are all key corridor (20-minute headways at peak) or rapid routes (10-minute headways at peak) as defined by Port Authority service guidelines. Additionally,

---

[1] 2016 PUBLIC TRANSPORTATION FACT BOOK. (n.d.). Retrieved April 2, 2018, from http://www.bing.com/cr?IG=B452AAF7CA90443E966B8D72BC29AB18&CID=2878CDE661D366060990C620607C6778&rd=1&h=FwUzemQzXSiokbvG-LDwWablG8lquGbnXAbDrMQ-PuE&v=1&r=http://www.apta.com/resources/statistics/Documents/FactBook/2016-APTA-Fact-Book.pdf&p=DevEx,5068.1

[2] www.portauthority.org. (n.d.). Retrieved April 2, 2018, from http://www.portauthority.org/paac/portals/0/ServiceGuidelines/2016/2016%20Annual%20Service%20Report%20Final.pdf

unlike other peak and rapid routes, our six routes will not be impacted by the forthcoming bus rapid transit extension.

With these six, disparate routes in mind we used bus schedule data, actual bus performance data, the Port Authority service guidelines, and social demographic information to offer a systematic way for our clients to approach estimating bus frequency change's effect on ridership. Our process involved identifying the factors influencing ridership, selecting the most significant factors, building a model, and then estimating the influence of frequency on ridership.

Literature Review and Research Gap

At the start of this project, our team performed an in-depth review of past work related to the following topics:

- Investigating ridership elasticity with respect to bus frequency,
- Identifying demographic and performance factors that influence bus ridership,
- Using automatic passenger count (APC) and automatic vehicle location (AVL) data in ridership studies, and
- Using general feed transit service (GTFS) data in ridership studies.

One of the most useful pieces was an analysis of existing literature discussing the factors that influence transit ridership. **The Factors Influencing Transit Ridership: A Review and Analysis of the Ridership Literature** by Taylor and Fink breaks the analyzed papers into two initial groups, descriptive and causal; and divides the factors considered into two types, external and internal. Descriptive analyses are described as qualitative and usually related to marketing, while causal analyses typically attempt to discover causation for changes in ridership, typically by using multivariate regression. It further breaks causal analyses down to aggregate and disaggregates studies where aggregate studies focus on transit systems as units of study and use metropolitan level data, while disaggregate studies focus on travelers' individual mode choice decisions. It describes internal influences as the things that the transit agency itself controls like frequency, marketing, fare costs, and scheduling, while external analysis is broken into socio-economic factors, spatial factors, and public finance factors.

We found these classifications to accurately match the other papers we considered, giving us a new way to view our methodology in the context of other research on the topic. Specifically, our research – using multivariate regression to investigate and predict individual decisions – places it into Taylor and Fink's disaggregate, causal analysis category.

- **Factors considered:** This study concludes that 1. Vehicle access and utility (i.e. parking availability) are most linked to ridership, and 2. Economic factors including unemployment are the second most influential indicator of ridership.
- **Key Takeaways**
    - Our team's analysis is classified as a causal analysis with a disaggregate study of internal and external factors
    - Common pitfalls of causal analyses include:

        - Looking at unlinked rather than linked trips which describe the full origin-destination route,
        - The inaccuracy of the model due to collinearity among the independent variables, and
        - Endogeneity problems between the service supply and demand (i.e. demand is naturally higher when supply is greater and vice versa).

The **Factors Affecting Transit Ridership** report prepared by the City of Edmonton, Canada discusses the City's transit ridership based on literature and supplemented with local data from the Edmonton Transit System, where possible. The paper did a great job of describing the data that was readily available to transit agencies like rider surveys and other data on their existing ridership. The report gave our team a sense of what additional data sources Port Authority might have at its disposal and how it might be helpful.

Additionally, this paper underscored the fact that the most influential factors are often not within a transit agency's control.

- **Factors considered:** mode captivity and demographics, transit mode and right of way, scheduling and service hour changes, frequency, reliability, trip type (e.g. commuting, travel to school, recreational) land use, density, diversity, distance to transit, and natural environment.
- **Key Takeaways:**
    - While the factors most affecting transit ridership are outside of the control of the transit agency, reliability has a big impact on ridership.
    - Use of available survey data on current ridership can give clues to influential factors.

The 1999 study, **Using non-real-time Automatic Vehicle Location data to improve bus services**, uses historical automatic vehicle location data, passenger surveys, on bus observation, and on-bus passenger counts (gathered by observation) to calculate the passenger arrival rate at stops along a bus route. In the study, the passenger arrival rate is used to estimate annual patronage, length of each stop, and the speed of buses as they move between stops. Ultimately, the information was used to analyze which segment of a route would benefit from bus priority measures to improve on-time performance.

This study provided an illustration of how vehicle load and speed data was collected before automatic passenger counters.

- **Factors considered:** Data gathered from watching passengers board and alight

A study that specifically looks at bus frequency is Optimal fleet size, frequencies and vehicle capacities considering peak and off-peak periods in public transport by Jara-Díaz, Fielbaum, and Schwender. This paper determined the fleet size of a single bus line based on demand at peak and off-peak times. Specifically, the paper considers demand, trip length, and traffic speeds during the peak and off-peak periods in order to determine the size of the vehicle, and frequency of vehicles needed. Many of the key findings of this paper related to optimal bus size which we will not be commenting on in our research, but the design of the multi-period model is useful.

- **Key Takeaways**:
    - It is important to study both peak and off-peak frequencies in order to develop the most accurate model to represent the optimal frequency for a route.
    - Identifying the appropriate bus size plays a significant role in service quality and there is a benefit of looking at bus size in an analysis that considers peak and off-peak time periods separately, passenger arrival rate, and other characteristics.

One particularly appreciated piece of literature was the **FSUTMS Mode Choice Modeling: Factors Affecting Transit Use and Access** **(2011) study.** The revised Florida Standard Urban Transportation Model Structure takes transit level of service (LOS), regional accessibility, land use, and users' demographic and socioeconomic data into consideration to build multiple linear regression and forecast the future transit use.

- **Factors Considered:** The LOS data includes average bus headways, a total number of bus runs in a census tract, and the proportion of population and workers who are close to the bus stop, which is calculated by using GIS buffer to calculate. The larger proportion means the bus would be more accessible, which may positively influence the ridership. The regional accessibility considers the employment and travel time between zones. The land use includes factors such as population density, dwelling unit density, and parcel size. Other socioeconomic and demographic characteristics include variables such as gender and income.
- **Key Takeaways:**

    - Geographic Information Systems buffers can be used to create a more accurate picture of whose serviced by a toy.
    - Some bus users may access the bus stops by vehicles. The auto access distance is gained by geocoding park-n-ride locations. The auto access trip length is the shortest network distance between the residential points and the closest park-n-ride locations.
    - The research uses different variables to create several models and compare the R^2 and sum of squared prediction errors to compare the accuracy of the models.

Transit Ridership Growth Study (2015) had the main goal of determining the policies that could
increase transit ridership. The factors are evaluated separately in "high" and "low" scenarios that reflect different levels of implementation. Each factor is analyzed by the regional travel demand model.

- **Factors Considered:** The model shows that the CBD employment growth, jobs near the bus stop, pricing strategies such as variable fares, parking pricing, and the comfortableness of the bus will exert a strong effect on the transit use. Meanwhile, the sustainable ridership depends on transit capital investment and the strategies will have synergies.
- **Key Takeaways:**
    - Land use policies and other policies have a large effect on transit ridership.
    - The study provides a gain/loss ratio to show the benefit-cost analysis. The ratio is calculated by getting the result of the absolute value of the sum of increasing ridership divided by the decrease of the models on which ridership declines.

Optimal fleet size, frequencies and vehicle capacities considering peak and off-peak periods in public transport, (2017) determine the fleet size of a single bus line based on demand at peak and off-peak times. Specifically, the paper considers demand, trip length, and traffic speeds during the peak and off-peak periods in order to determine the size of the vehicle, and frequency of vehicles needed. Many of the key findings of this paper related to optimal bus size which we will not be commenting on in our research, but the design of the multi-period model is useful.

- **Key Takeaways:**

- Using a multi-period model can increase the accuracy of bus demand predictions.

Authors	**Year**	**Factors**	**Location**
Suwardo and Kamaruddin	2010	**travel time elasticity**, ticket fare, fuel price, per capita income, service frequency, and headway	Malaysia
Polzin	2004	**number of households with zero vehicles**, population in the zone, total employment in the zone, service employment in the zone	Florida
Not listed	2003	demand greater than frequency provided, transit supportive land uses, environmental justice concern, location with congested highways	Georgia
Taylor and Fink	2003	**access to a vehicle**, **vehicle utility** (i.e. parking availability), economic factors including unemployment	United States
City of Edmonton	2016	**automobile ownership**, transit mode, the right of way, scheduling and service hour changes, frequency, reliability, trip type, land use, density, distance to transit, natural environment	Edmonton, CA

Table 1: Summary of Literature Review Findings On Factors Impacting Transit Ridership

The second set of literature we reviewed considered ridership elasticity with respect to bus frequency. The paper that influenced us the most was the **Cross-Elasticities in Frequencies and Ridership for Urban Local Routes** by Totten and Levinson. While the paper aims to consider cross-elasticities, it begins with the following helpful literature review directed us to a number of papers that gave us a sense a reasonable range of ridership elasticity in response to frequency increase.

Author(s)	**Year**	**Ridership Elasticity**	**Locations Studied**
Evans	2004	0.30–1.03	North America, Europe
Currie and Loader	2009	0.17–0.38 weekday 0.80+ weekends	Melbourne, Australia
Suwardo and Kamaruddin	2010	.03 - .07	Malaysia
Litman	2015	0.50–0.70	North America, Western Europe

Table 2: Literature Review Findings On Ridership Elasticity with Respect to Frequency Increase

Existing Tools

After learning more about both the factors and demand elasticities found in other research, we considered an existing tool called STOPS. STOPS is the Federal Transit Administration's project that enables the user to speed up the process of four-step travel demand modeling. The tool is designed to estimate ridership on fixed guideways (bus rapid transit lines). Using 2000 Census data and General Transit Feed Service data, STOPS allows the user to input their agency data and fill in some fields and outputs a ridership estimate.

Research Gap

Having considered all of that literature and the Federal Transit Administration's STOPS tool, we knew how to proceed in a way that provided a meaningful addition to the field of study. Our research uniquely uses local, Port Authority APC/AVL data, recent American Community Survey data, and considers not only social demographic factors and performance factors but also interaction factors. in all, our study is unlike any other we had seen.

Data Sources and Data Description

Data Source 1: Automatic Vehicle Location and Automatic Passenger Counter

For data, we used information related to the bus's schedule and information about the ridership. Automated Vehicle Location (AVL) and Automated Passenger Counter (APC) data tell us the actual time and location of buses. Our APC/AVL data was given to us by Port Authority and it spanned March 2016 to July 2017, which is five quarters.

Field Name	Short Description	Data Type
DOW	Day of week code	Integer
dir	The direction of the trip along the route	Integer
ROUTE	Route Code	Integer
TRIPA	Trip Number	
BLOCKA	Block Number	
VEHNOA	Vehicle Number	Integer
daymoyr	Day/Month/Year of run	String
STOPA	stop sequential number	Integer
QSTOPA	PAAC stop alpha numeric ID number	String
ANAME	Stop Name	String
HR	Arrival Hour	Integer
Min	Arrival Min	Integer
SEC	Arrival Sec	Integer
DHR	Departure Hour	Integer
DMIN	Departure Min	Integer

DSEC	Departure Sec	Integer
ON	Observed Number of Passengers Boarding	Integer
OFF	Observed Number of Passengers Alighting	Integer
LOAD	Number of Passengers on Bus	Integer
DLMILES	Miles travelled from last stop	Float
DLMIN	Minutes travelled from last stop	Float
DLPMLS	Change in passenger miles from last stop	Float
DWTIME	Dwelling time (min)	Float
DELTA	Distance in feet from observed GPS coordinates of the record to GPS coordinates for the stop	Integer
SCHTIM	Scheduled arrival time	Integer
SCHDEV	Difference in arrival time with schedule time if a timepoint	Float
SRTIME	Scheduled run time from previous time point to current time point	Float
ARTIME	Actual travel time from previous time point to current time point	Float

Table 3: Dataset Description of AVL&APC

Data Source 2: General Transit Feed Specification

The General Transit Feed Specification (GTFS) data detail the scheduled bus departure time, location, and the size of the bus. When combined with AVL & APC data, we had a wealth of information about the bus performance.

Data Source 3: American Community Survey Data, Census Data

In the project, for our demographic data source, we used the United States Census' American Community Survey results. The American Community Survey has conducted annually, unlike the Census which is a decennial survey. We used the 5-year projections from the ACS survey to create our demographic factors. Using GIS were able to analyze the demographic characteristics of the Census block groups along the routes. We used the following GIS layers and tabular data:

1. A layer of bus routes of the Port Authority of Allegheny County (PAAC)[3]
2. A layer of bus stops of the Port Authority of Allegheny County (PAAC)[4]
3. Block Groups shape-file of Allegheny County, 2017 TIGER/Line® Shapefiles[5]
4. Water shape-file of Allegheny County, 2017 TIGER/Line® Shapefiles[6]
5. Unweighted Sample Count of the Population, 2012-2016 American Community Survey 5-Year Estimates[7]
6. Sex by Age, 2012-2016 American Community Survey 5-Year Estimates[8]
7. Means of Transportation to Work, 2012-2016 American Community Survey 5-Year Estimates[9]
8. School Enrollment by Detailed Level of School for the Population 3 Years and over, 2012-2016 American Community Survey 5-Year Estimates[10]
9. Employment Status for the Population 16 Years and over, 2012-2016 American Community Survey 5-Year Estimates[11]
10. Tenure by Vehicles Available, 2012-2016 American Community Survey 5-Year

[3] http://www.portauthority.org/generaltransitfeed/GIS/

[4] http://www.portauthority.org/generaltransitfeed/GIS/

[5] https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2017&layergroup=Block+Groups

[6] https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2017&layergroup=Water

[7] https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B00001#

[8] https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B01001#main_content

[9] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B08301&prodType=table

[10] https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B14007#main_content

[11] https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B23025#main_content

Estimates[12]

11. Median Household Income in the Past 12 Months (in 2016 Inflation-adjusted Dollars), 2012-2016 American Community Survey 5-Year Estimates[13]

12. 2015 Job Destinations Points of Allegheny County, The Longitudinal Employer-Household Dynamics (LEHD) On the Map Application[14]

---

12 https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B25044#main_content

13 https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B19013&prodType=table

14 https://onthemap.ces.census.gov/

Factor Identification

Generally, we could divide the factors affecting transit ridership into two groups, exogenous factors, and endogenous factors. Exogenous factors are the factors that transit agencies could not control, including vehicle ownership, fuel prices and availability, demographic factors (age, gender, household income, etc.), and population and employment distributions. As for endogenous factors that transit agencies could control, it includes fare, headway, route structure, and other performance measures.

Through our literature review, we would consider the following socio-demographic and bus performance factors as the independent variables in our ridership regression mode

Social-Demographic Factors

Demographic Factor	Definition
Household Percentage Under 80% AMI	$\frac{Households\ with\ Income\ Under\ 80\%\ AMI}{Total\ Households\ in\ the\ Block\ Group}$
Job Density	$\frac{Total\ Number\ of\ Jobs\ in\ the\ Block\ Group}{Block\ Group\ Area\ (sq\ mi)}$
University Enrollment	$\frac{Postsecondary\ Education\ Enrollment\ per\ Block\ Group}{Total\ Population\ in\ the\ Block\ Group}$
Land Use	*Dominant Land Use Type in the Block Group*
Senior Household Percentage	$\frac{Population\ Aged\ 65+}{Total\ Population}$
Unemployment Rate	$\frac{Unemployed\ Population}{Total\ Population}$
Vehicle Ownership	$\frac{Total\ Vehicle\ Owners}{Total\ Households\ per\ Block\ Group}$
Population Density	$\frac{Total\ Population}{Block\ Group\ Total\ Area\ (sq\ mi)}$

Figure 1: Factor Identification -- Social Demographic Factors

Land-Use

Land use and zoning controls are critical tools in the urban planning process. Especially, land use information is of great value in the transportation development. “Southwestern Pennsylvania Commission (SPC)’s Land Use/Land Cover (LULC) uses a three-tier hierarchical classification system. Six LULC types make up Level I: Urban-Built-Up, Agricultural, Rangeland, Forest, Water, and Barren Land. Levels II and III provide a more thorough classification of the land. Level III classes will be highlighted briefly following Level II definitions.

In our analysis, we use the block group level's major land use or land cover to quantify the correlation between the land use type and the ridership in Allegheny County block groups. Therefore, we use SPC's LULC shapefile spatial join the Allegheny County block groups feature to get the major type (with the largest area) with GIS. In this way, for each unique block group GEOID, we have the corresponding major type.

Data source: Allegheny County Land Use/Land Cover 2010 GIS shapefile, Southwestern Pennsylvania Commission[15]

Senior Percentage

People of different ages have different commuting needs. Elderly people are less likely to need to commute to work and their financial status can also influence their choices for transportation methods. We use the age of 65 as the cutoff for seniors and analyze the percentage of seniors at the block group level.

In the original dataset, it shows the age breakdown by sex of each block group. We firstly combined the data of both genders to get the total number of each age breakdown by block group, then we merged the data and separated all the people into the group above 65 and the group under 65. Finally, we used the formula of the number of people above 65/ the total population to get the eldership percentage of each block group.

Data Source: Sex by Age, Block Group level, Allegheny County, 2012-2016 American Community Survey 5-Year Estimates[16]

Population Density

The population distribution can affect the use of the bus service. The more people in a certain area, the more the bus riders can be. Instead of the simple population data, we use the population density which is measured by the people number per square mile.

We firstly get the total population and the land area of each block group. The original unit of the land area is square meter. To make it more readable, we transfer the unit to square

[15] http://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=1603

[16] Data Access and Dissemination Systems (DADS). (2010, October 05). Results. Retrieved from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B01001&prodType=table

meter. Then we use the formula of the total population of the block group / block group land area (sq mile) to get the population density.

Data Source: Total population Employment by Block Group, 2012-2016 American Community Survey 5-Year Estimates[17],Cartographic Boundary Shapefiles of Allegheny county[18]

Unemployment Rate

The employed and unemployed people may have different needs and preference for bus rides, and the bus performance (e.g. waiting time) may also have disparate effects on the bus choices of different groups. We put unemployment level into consideration in our model.

In the original dataset, it shows the total population and employed population of each block group. We firstly calculate the number of unemployed people by block group and then use the formula of unemployed population/total population to get the unemployment rate.

Data Source: Employment Status for The Population 16 Years and Over, Allegheny County, 2012-2016 American Community Survey 5-Year Estimates[19]

Factor: Income Level

The income level will influence the people's choice of the transportation method. We assume that the people with lower income tend to use more public transportation, especially the bus. In our research, we use the household income to measure the income level within each of the block group. According to the determination of income used in most federal and state, we define 80% AMI (Area Median Income) as the threshold of the middle class and upper class. In other words, we use the percentage of the household with the income below 80% AMI to represent the overall "below middle class" level in each block group and put this variable into our model.

The dataset provides us with the income breakdown and the number of the household in each income level by block group. We firstly find from the HUD(Department of Housing and Urban Development) website that the 80% AMI (Area Median Income) for the 4-people

[17]https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B01003&prodType=table

[18] https://www.census.gov/geo/maps-data/data/tiger-cart-boundary.html

[19] Data Access and Dissemination Systems (DADS). (2010, October 05). Results. Retrieved from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B01003&prodType=table

household of Allegheny County in the year 2017 is $58,100. We use this number as our standard. In the original dataset, we only have the income breakdown of $50,000 to $59,999, so we make the assumption that the households are averagely distributed and the household number of $50,000 to $59,999 is 80% of that of this income level. Then we add up all the numbers of the household under 80% AMI by block group. Finally, we divide this new figure by the total household number in each block group and then get the "below middle class" level variable.

Data Source: Household Income In The Past 12 Months (In 2016 Inflation-Adjusted Dollars) Universe: Households More Information of Allegheny County, 2012-2016 American Community Survey 5-Year Estimates[20]

Job Density

One of the basic function of the urban transit system is to connect people with jobs. In Pittsburgh region, workers are one of the major groups of public transit ridership. To figure out where are the destinations of the workers and to measure how many jobs the existing service network serves, it's necessary to include the job density factor in our research.

In our analysis, we select the job density as one of the independent variables in our regression analysis. To calculate the job density by block group, we spatial join the Allegheny County block groups feature with the job destinations points feature, summing up the total number of jobs (all kinds of jobs) by block group with GIS. Thus, for each unique block group GEOID, we have the corresponding job density.

Data Source: 2015 Job Destinations points in Allegheny County, the Longitudinal Employer-Household Dynamics (LEHD) program On the Map application[21]

College Student Enrollment

Connect people with education opportunities is one of the basic function of the urban transit system as well. In Pittsburgh region, students are also one of the major groups of public transit ridership. In our research, we assume that college students and grad school students

[20] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B01003&prodType=table

[21] https://onthemap.ces.census.gov/

are the major source of ridership of PAAC bus service. To represent the student's demand, we use the school enrollment of college and above among population 3 years and over.

To measure the student demand, we use the population of college and above enrollment census data to calculate the percentage of the total population by block group. Thus, for each unique block group GEOID, we have the corresponding percent population of college and above enrollment.

Data Source: School Enrollment by Detailed Level of School for the Population 3 Years and over, 2012-2016 American Community Survey 5-Year Estimates[22]

Vehicle Ownership

Vehicle available is an important socio-economic factor that could reflect the degree of car use. Although the relationship between transit use and car use could be interrelated, vehicle ownership is necessary to be included in any kind of ridership models.

In our analysis, to measure the vehicle available by block group, we use vehicles available data from census to calculate the vehicle ownership per household, with total owners/ total households in each block group. Thus, for each unique block group GEOID, we have the corresponding vehicle ownership per household.

Data Source: Tenure by Vehicles Available, 2012-2016 American Community Survey 5-Year Estimates[23]

Methodology: GIS Analysis

In our research, we used the following GIS techniques to demonstrate the spatial information of Allegheny County Block Groups and bus routes and stops of Port Authority of Allegheny County with Census Data and other GIS layers.

Proximity Analysis and Kernel Density Map

To show the areas of influence or the bus service coverage, we could create walksheds for bus routes. In ArcGIS, we could create bus stops point buffers and bus routes line buffers, or use Network Analyst to create service area around bus stops. In our analysis, we would use point buffers, to create ¼ mile service areas around the 7040 bus stops of PAAC in

---

[22] https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B14007#main_content

[23] https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=table&id=table.en.ACS_16_5YR_B25044#main_content

Allegheny County. In this way, we could determine how many jobs and people a specific route serves.

Kernel density map is a widely used method in statistics for smoothing data spatially and producing a “heat map”. For the Kernel Density Map, the input data is often centroids of a polygon or points of individual demands for goods or services. With the point or polyline features, we could calculate the density of specific population. In our analysis, we would use Kernel Density to create a heat map of the job destinations in Allegheny County.

Choropleth Map

Choropleth map is a thematic map with shaded polygons showing numeric data. Generally, we could use choropleth map to show the geographical distribution of a specific attribute with several classes by symbolizing with different colors or sizes. In our analysis, we use symbology of graduated colors to stand for different levels of socio-demographic data, for instance, population density, in Allegheny County by block groups. Each polygon stands for a census block group, each color stands for a class of population density. With the choropleth map, we could observe whether there is a significant geographical distinction of the socio-economic data among the 1100 block groups in Allegheny County. Thus, the visualization of socio-economic data could support the further regression analysis.

Bus Performance Factors

Performance Factor	Definition
On Time Performance	*The difference of the actual arrival time and the scheduled arrival time* $x = 1, \text{if difference} \in [-1, +6]$ $x = 0, \text{otherwise}$
Crowding Level	$Bus\ Load / Number\ of\ Seats$
Waiting Time	$Actual\ Headway / 2$
Stop Skipping	*A stop is skipped because the bus is fully loaded*

Figure 2: Factor Identification – Bus Performance Variables

Data Separation

To gain an accurate understanding of the bus performance factors we separated the data into 8 parts for each route. Each route has two directions, inbound and outbound. For each direction, we can observe four time periods, which are weekday peak hour, weekday off-peak hour, Saturday and Sunday. By separating each route data into 8 parts, we are able to capture the nuances among different time periods within different directions.

GTFS data is unchanged within a quarter, which means the bus scheduled headways are also unchanged within each quarter. Since we are going to capture the influence of bus frequency on ridership, we grouped our APC&AVL data by the same quarter definitions as the GTFS data.

Also, in our model, we introduced the influence of social demographic factors on ridership, and all the social demographic factors are based on block group level, there are different numbers of bus stops in each block. Thus, in order to coordinate the social demographic factors, we further grouped our APC&AVL data into block level.

In summary, in our project, we separated each route into 8 parts: inbound-weekday-peak, inbound-weekday-off-peak, inbound-Saturday, inbound-Sunday, outbound-weekday-peak, outbound -weekday-off-peak, outbound -Saturday, outbound -Sunday. And within each part, data for regression were grouped by block group and quarter time period.

Peak Time Selection

In order to determine the peak times and create the 8 periods for each route, we used a four-step method.

1. Load all the APC data and select by route and direction, as the basic datasets to be handled.
2. Select weekday, Saturday, Sunday, with DOW column.
3. Compare two methods:
    a. Use the trip as the x-axis and average load group by the same trip as the y-coordinate.
    b. Define half an hour as time range x, calculate total loads during that hour as y
4. Use the trend to determine peak hours distinguishably.

When defining peak hour using the average method (“a” in the above list) we found some trips to have unreasonable load numbers if the trip occurred when there was a frequency increase. In these cases using the average load was not showing the real peak. As you can see in the figure below, it causes the fluctuation and the time range and makes selecting a peak hour very difficult.

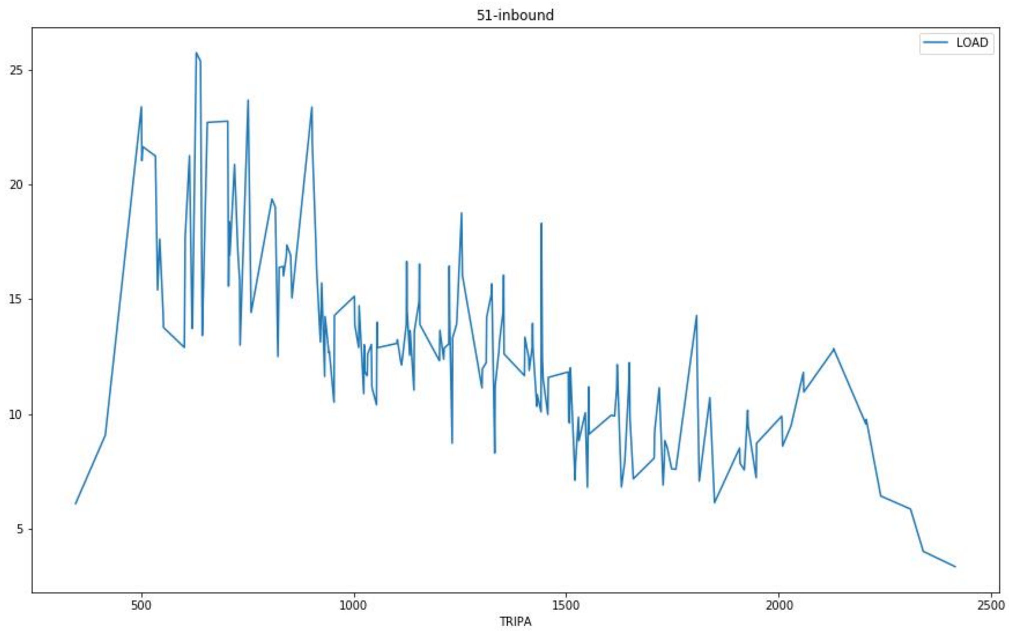


Figure 3: Original Methodology for Peak Time Selection

So we adjusted it by adding another criterion to help, that defines a half hour time range, and use total load to smooth the data (as described in 3b in the steps above). However, it also has another problem. If some half hour, for example, have 3 trips and the adjacent range only have two trips, this scenario may cause serrated plot, but we can still observe the trend to correct average selected results. Ultimately this is how we defined our peak hour time periods.

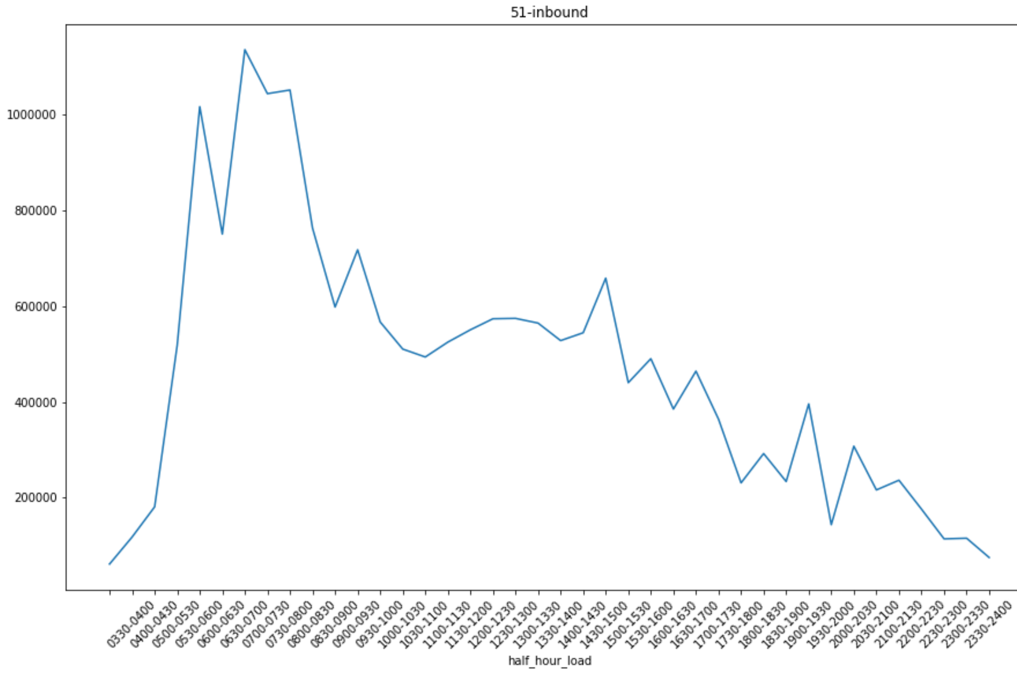


Figure 4: Updated Methodology for Peak Time Selection

The following table shows the result of the method of total load over the 5 quarters within the time frame:

Route	Inbound Period 1	Inbound Period 2	Inbound Period 3	Outbound Period 1	Outbound Period 2	Criteria
8	5:30 - 10:30	13:00 - 16:00		14:00 - 18:00		>= 300,000
12	7:30 - 10:00	13:30-14:30	20:30-22:00	7:30-16:00		>=200,000
16	5:00-8:30			14:00-18:00		>=400,000
51	5:30-8:00			14:30-17:00		>=700,000
91	6:30 - 8:00	13:00-17:30	>= 300,000	7:00-8:30	15:00 - 18:00	>=600,000
82	6:30 - 9:00			14:00-17:30		>=14

Table 4: Peak Time Selection Results For 8 Routes

Stop Skipping

In order to calculate stop skipping we considered three main conditions: the duration at a stop, the number of people on the bus, and the next trip's boarding number. We investigated a trip for stop skipping if it met the first two conditions below. If so, we proceeded to examine the third condition.

1. Duration (stop departure time - arriving time) = 0
2. Load number >= 1.2 * max seats of that vehicle type (max capacity)
3. Next trip's boarding number is greater than some threshold, which varies for each routes. Note: In the first filtering, we use 0 to eliminate last stops.

Then, based on the filtered dataset, to manually select and check with next trip's information:

Scenario 1: If next trip's boarding number is 0 and available seats (max capacity - actual load) is also 0, we define the previous trip at the stop as a skip.

Scenario 2: If next trip's boarding number is greater than twice of normal boarding (average number), then it is also a skip.

Scenario 3: If next trip's boarding number is small and between above thresholds, and the available seats is also approximately equal to the number, we regard it as a skip.

Problems Solving Notes:

The methodology of the stop skipping is not that perfect and has some potential pitfalls:
In Scenario 1 above, 0 boarding number on next trip could result from no passengers waiting at that stop. Actually, this condition takes up half of the final result, but we can't exactly distinguish it between double skip or false skip.

For Scenario 2 above, abnormally large boarding number could be from a bus transfer.

And in Scenario 3 the small boarding number on next trip could be new arriving passengers from that stop, not those remaining from the previous trip.

On-time performance

On-time performance is the most widely used transit reliability measure in North America (Sen, L., Majumdar, S. R., Highsmith, M., Cherrington, L., Weatherby, C, 2011). And it closely related to other factors, such as waiting time. However, different agencies have their own definition of "on time". And it varies a lot based on different situations and purposes. In this project, we defined a range of how early a bus arrived before the schedule and how late a bus arrived after the schedule as on time. This range was based on the PCCA definition. It can be counted as on-time within the 1 minute earlier than the schedule and 6 minutes later than the schedule.

Our definition of calculating on-time performance was as follows:

1. Calculate the arrival deviation using the difference between the scheduled stop arrival time and the real stop arrival time (arrival deviation = stop real arrival time - stop schedule arrival time).
    a. If -1 minute <= arrival deviation <= 6 minutes, we define it as on time.
    b. If arrival deviation < -1 minute or arrival deviation > 6 minutes, we define it is not on time.
2. Create a dummy variable.
    a. If the arrival deviation in the (-60, +360), the on time performance = 1.
    b. If the arrival deviation in the (-86,400, -60) or (+360,86400), the on time performance = 0.

Problems Solving Notes:

Some morning times were stored in over 24-hour format. For example, the 01:30 on the morning was stored in 25:30. So the first thing is to adjust the format in the uniform.

Factor: Passenger Waiting time

Waiting time is one important factor impacting the customers' experience. And it is also an important factor to measure if a transportation system can provide a reliable service worth customers to rely on.

The waiting time is divided into two important parts by the scheduled arrival time, waiting time before the bus arrives and waiting time after the bus arrives. The first part, waiting time before the bus arrival is mainly determined by the headway length. But the second part, waiting time after the bus arrival is highly correlated to the reliability of the transportation system ((Salek, M. D., Machemehl, R. B, 1999). But in fact, it is very difficult to capture the actual waiting time of each customer. So the waiting time in this report is calculated based on the uniform assumption, which is that all passengers will get to the bus stops randomly.

To calculate the waiting time we used the following we used the uniform distribution assumption, the wait time = headway/2.

Problems Solving Notes:

The first trip of each time period was ignored as the first trip did not have a headway.

Factor: Crowding Level

Crowding level is an elementary and impactful factor. First, it closely effects customers' riding experience. With the crowding level increases, customers feel more uncomfortable and inconvenient, such as standing for a long time, keep balancing and etc. The bad experience further impacts customers' frequency of taking buses. (KFH Group, 2013) Second, the more crowded the bus is, the longer its dwelling time is. It further impacts the service's regularity and reliability. So the operators should take it into account carefully.

Normally, the crowding level is defined as an acceptable level of the passenger loads against a standard. These standards can be defined relative to the seated capacity of the vehicle in question (e.g. 125% of seated capacity) (KFH Group, 2013). The acceptable levels of crowding are defined by the agency in question and can vary widely from agency to agency (Li, Z., Hensher, D. A, 2013). In this research, the crowding level was defined as a continuous relative factor. the crowding level equals loads of a trip at a station by the seats of the bus of this trip. And the data used in the model is calculated as the average of a crowing level per route per block per time period.

To calculate crowding level we used the following method: bus load divided by bus capacity, whereby the bus capacity is the number of seats per bus and the load the bus's passenger number of each trip at each station.

Regression Analysis Model Development

Regression Dataset Preparation

After we get all the cleaned datasets and both performance, social demographic parameters, we need to put together every parameter as well as ridership for building our models.

Step 1: Get the corresponding ridership

1). Load in the cleaned APC & AVL dataset for the route.

2). Select the corresponding time period, i.e. peak hour, off-peak hour, Saturday or Sunday.

3). Group the ridership by 5 quarters and the block groups passing by this route.

Step 2: Join the bus performance parameters, social demographic parameters, and the ridership together as a table for model building.

1). Prepare an identifier column, which is the combination of quarter name and block "GEOID", and prepare an empty table data frame based on the identifiers.

2). Read in bus performance parameter dataset, social demographic parameter dataset and ridership dataset from last step individually.

3). Put each variable and ridership based on identifier column by column.

4). Output the data frame as a CSV file for build model.

Add interaction factors: Collinear Analysis

From the collinearity table, choose the combination of bus performance variable and the social demographic variable has the correlation absolute value lower or equal to 0.1 as interaction parameters.

	Waiting Time	On time Performance	Crowding Level	Stop Skipping	Population Density	Job Density	% Households Uder 80% Percentage	Unemployment Rate	Vehicle Per Household	Elder Percentage	University Enrolled Students
Waiting Time	1										
On time Performance	0.62223	1									
Crowding Level	0.2799	0.44643	1								
Stop Skipping	-0.03214	0.14856	0.15597	1							
Population Density	-0.11821	-0.09302	0.18696	-0.08918	1						
Job Density	-0.33794	-0.52426	-0.4771	-0.13258	0.06313	1					
% Households Uder 80% Percentage	0.34004	0.5372	0.31599	0.20163	0.16843	-0.69175	1				
Unemployment Rate	0.01845	0.21472	-0.0511	0.15763	-0.28112	-0.33445	0.41342	1			
Vehicle Per Household	0.07795	-0.12671	-0.17589	-0.09269	-0.45211	-0.03528	-0.48916	-0.29093	1		
Elder Percentage	0.28341	0.26526	-0.03729	0.11754	-0.18571	-0.47665	0.49006	0.31665	-0.06392	1	
University Enrolled Students	-0.37991	-0.45422	0.04893	-0.06242	0.50782	0.2049	-0.05954	-0.44707	-0.33856	-0.35278	1

Table 5: Correlation Analysis, Route 82 Inbound Peak

As the table shown above, for route 82 inbound peak hour, we added in “waiting time * unemployment rate”, “waiting for time * vehicle per household”, “crowding level * unemployment rate”, “crowding level * elder percentage”, “crowding level * university enrolled students”, “stop skipping * vehicle per household” and “stop skipping * university enrolled students” as interaction variables.

Feature Selection

We introduced 3 kinds of variables in our model, but not all of them will cause a significant influence on our model. For different models, there will be different significant variables, so we introduced two following method to help us select the appropriate variables for each model.

Feature Selection Method 1: LASSO

Least **A**bsolute **S**hrinkage and **S**election **O**perator(LASSO) method estimates coefficients β_λ by minimizing the l_1 penalized RSS.

$$\underbrace{\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)}_{RSS}+\lambda\underbrace{\sum_{j=1}^{p}|\beta_j|}_{\text{penalty}}$$

Figure 5: LASSO Formula

Lasso measures model complexity according to $|\beta_j|$, the tuning parameter λ allows us to control the overall **complexity** of the model, λ = 0 takes us back to least squares, λ=∞ gives us β_λ =**0,** The Lasso has the amazing property that for intermediate values of λ, β_λ will have entries *shrunk* towards 0, and some will actually be estimated exactly as 0. i.e., Lasso automatically performs variable selection.

And in order to choose λ for LASSO, there are normally three ways:

1). Choose a sequence of λ values

2). Calculate the K-fold CV error at each λ

3). Use the minimum CV error or 1-SE rule to pick λ

In ore models, we applied 1-se rule, which is to choose the simplest model whose accuracy is comparable with the best model, to choose the λ.

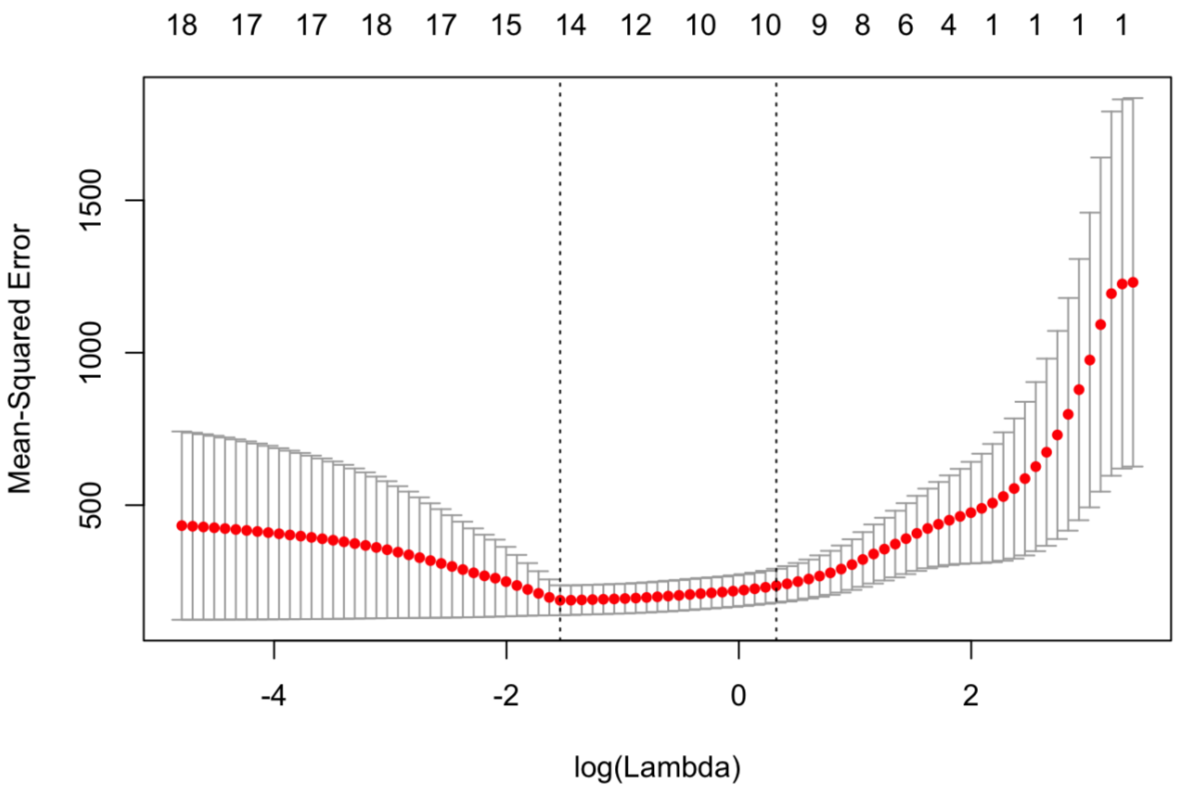


Figure 6: LASSO Method Sample Result, Route 12 Outbound Sunday

From the graph, the variable range between 10 to 14 have the lowest MSE. Since we choose 1-se rule to choose λ, LASSO will choose the simplest model within this lowest MSE range, which is 10 variables selected for this model.

Feature Selection Method 2: Best-subset

When we were selecting the features for model, LASSO method would always be the first choice, but, there are sometimes that LASSO cannot automatically select variables based on 1-se rule. For this occasion, we introduced a second feature selection approach -- best subset method.

The detailed approach of best subset method is as follows:

1). Let M0 denote the null model: The intercept-only model.

2). For k = 1,2,...,p

(a) Fit all k^p models that contain exactly k predictors

(b) Among these, pick the best model: The one having the smallest RSS, equivalently the *largest* R

3). Select the single best model from M0, M1, . . . M_p using both AIC and BIC as criterion.

- In Step 2)., we find the best model of each size
- In Step 3)., we put the models on *equal footing*, by looking at prediction error or explicitly adjusting for model complexity

And to be clear, the criterions “AIC” and “BIC” we talked about in step 3)., are just the different ways of setting penalty in model.

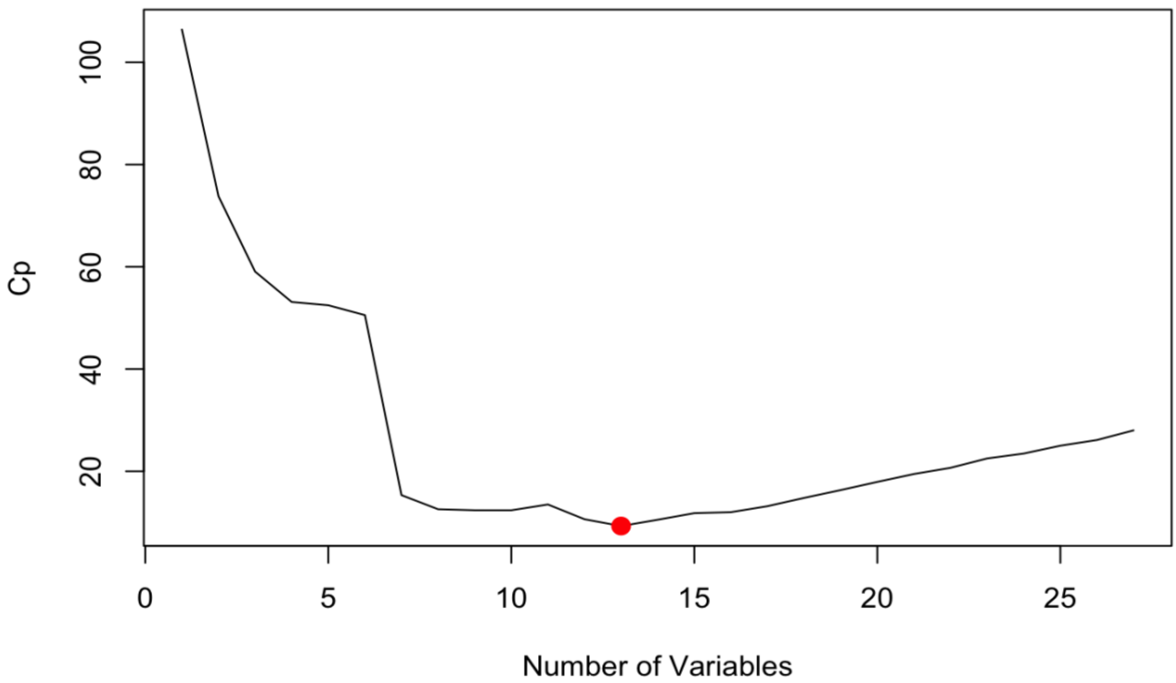


Figure 7: Best Subset Using AIC Sample Result, Route 12 Inbound Off-Peak

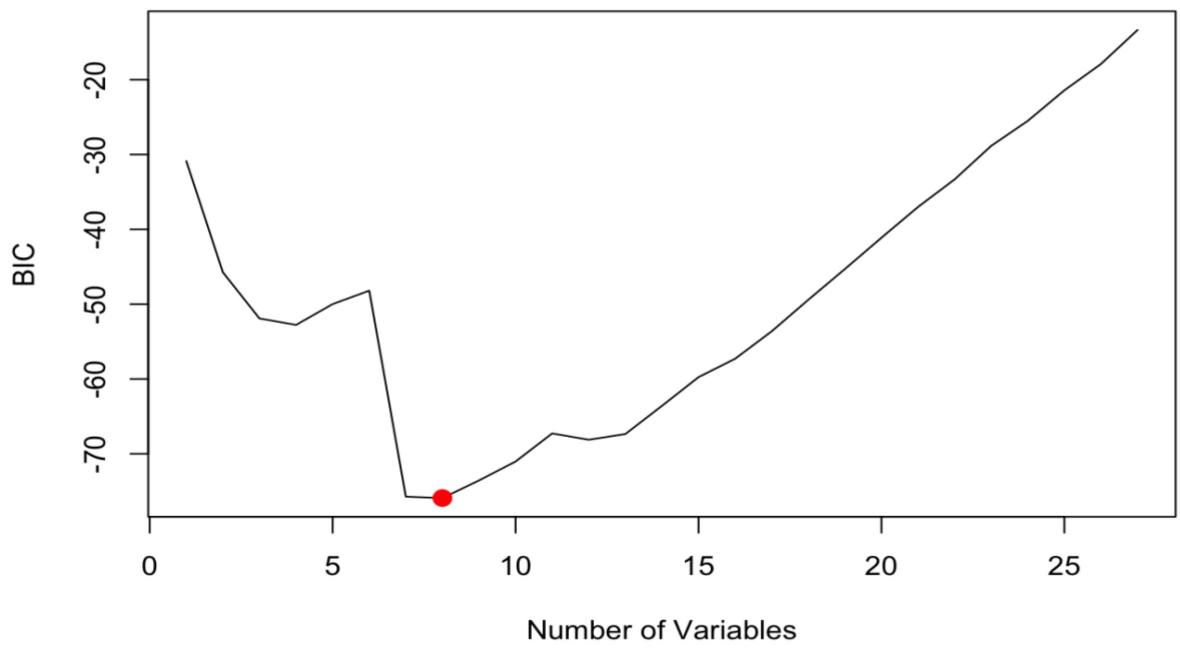


Figure 8: Best Subset Using BIC Sample Result, Route 12 Inbound Off-Peak

Model Validation: Random Forest Model

Random forest is a CART method(Classification and Regression Tree). It builds each tree on a bootstrapped training sample, each time a split in a tree is considered, the tree may only split between a predictor from a randomly selected subset of m predictors.

From the feature selection step above, we already have our linear regression model(using R for feature selection can also automatically generates the coefficients of each selected variables).   We are still uncertain about how this feature will perform in a non-linear model, thus we introduced our model validation process, building the random forest model based on the selected variables and perform 10 cross-validations to get the accurate R squared value and MSE for each model.

Ridership Elasticity Estimation

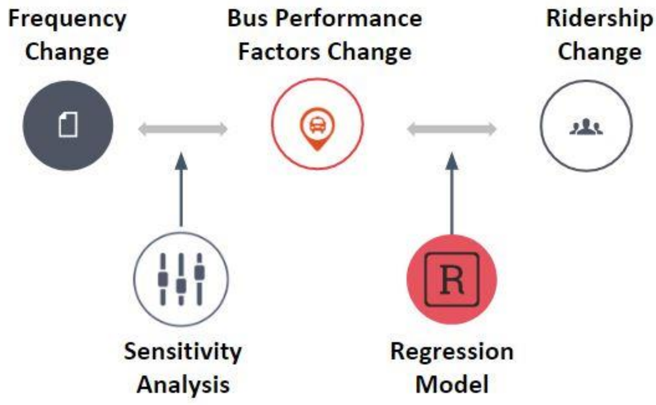


Figure 9: Flow Chart of Ridership Estimation

To measure the impacts of the change of four bus performance factors on ridership change, we've created 48 regression models. However, our ultimate goal is to estimate the ridership change with respect to the frequency change, which is exactly the ridership elasticity of frequency. Thus, we need a transition step to measure the quantitative relationship between frequency and our four bus performance factors.

Since there is no prevailing equations or models to quantify the relationship between frequency and our four performance factors, which are on-time performance, crowding level, passenger waiting time, and stop-skipping. We would use the coefficients from four separate correlation analysis as the sensitivities of the four bus performance factors.

In general, to estimate the ridership change, it would involve the following 3 steps:

1. Generate bus performance sensitivity matrix with correlation analysis.

For the route with frequency change ever, generate sensitivity matrix from correlation analysis.

For the route with no frequency change ever, import sensitivity matrix from other routes.

2. Import coefficients of bus performance factors and interaction terms from the 48 regression models.

3. Calculate the estimated ridership elasticities.

For model without interaction terms, use sumproduct of sensitivities and coefficients of bus performance factors as the estimated elasticity.

Results And Findings

GIS Findings - Demographic Route Profile, Choropleth Maps

Route	Population density	Percentage of households under 80% Area Median Income	Job Density	Senior (>65) Percentage	Unemployment Rate	Vehicle Ownership	Land Use	Whether go through major college district
8	Low	Medium	Low	Medium	Medium	Medium	Transportation Residential	No
12	Low	Low	Medium	High	High	High	Transportation Residential Non-Urban Built-up	Yes (a few colleges)
16	Medium	High	Low	Low	Medium	Medium	Residential Transportation Commercial	No
51	Mostly High	High	Low	High	High	High	Residential Non-Urban Built-up Transportation Commercial	Yes (a few colleges)
82	Mostly High	High	High	Low	High	Low	Residential Urban Built-up Transportation Commercial Non-Urban Built-up	Yes (most colleges)
91	Medium	Medium	High	Medium	Low	Medium	Transportation Residential Commercial	No

Table 6: Comparison Table of Characteristics

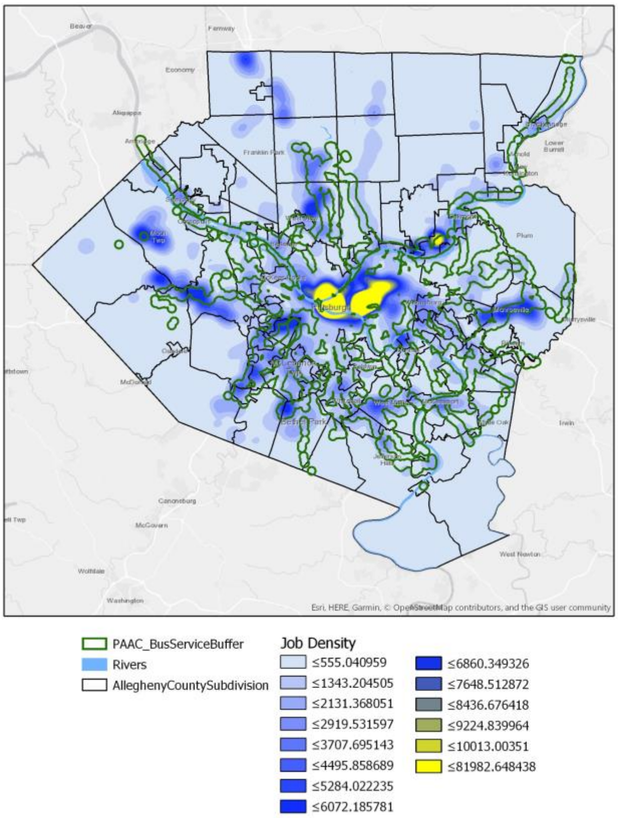


Figure 10: Job Density Heat Map

To better measure the effectiveness of PAAC transit service on connecting people to jobs, we created a kernel density map for job destinations and a multiple ring buffer for all the bus stops. As we can see in the figure, although the areas with the highest job density concentrate in the City of Pittsburgh, there are a large number of job centers distributed around the boundary of Allegheny County. Currently, the 7040 bus stops of PAAC covered 836 out of 1100 block groups. The PAAC transit service served 5858 out of 8849 destinations in Allegheny County. In total, the PAAC service covered 498,667 out of 710,479 jobs, which is a 70.19% coverage. It shows that there is around 30% of the job opportunities with no PAAC bus service coverage.

Below Middle-Class Households

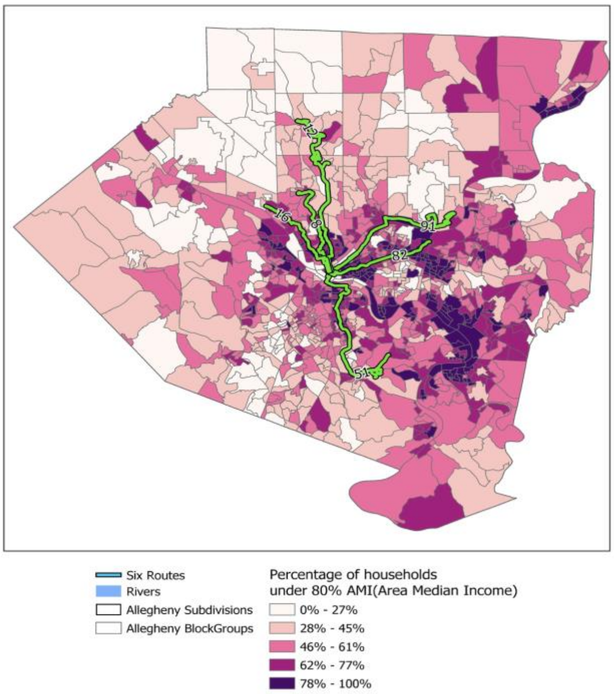


Figure 11: Map of Households Percentage under 80% Area Median Income

We define 80% Area Median Income as the threshold of the middle class, which means those whose income is less than 80% AMI are under middle class. The map shows the percentage of households under middle class in each block group. The Pittsburgh city, especially Pittsburgh CBD, and the Southeastern parts of Allegheny County near the edge of Pittsburgh have more block groups with a high percentage of under middle-class households.

Route 82 runs across Pittsburgh CBD and toward the East, whose block groups covered are those who have a large percentage of households below middle class. Route 16 goes from center Pittsburgh to the Northeastern of Allegheny County, which also covers block groups whose households are mostly those below middle class. Route 12 runs towards the North, which covers the block groups with comparatively low percentage of households below middle class.

Population Density

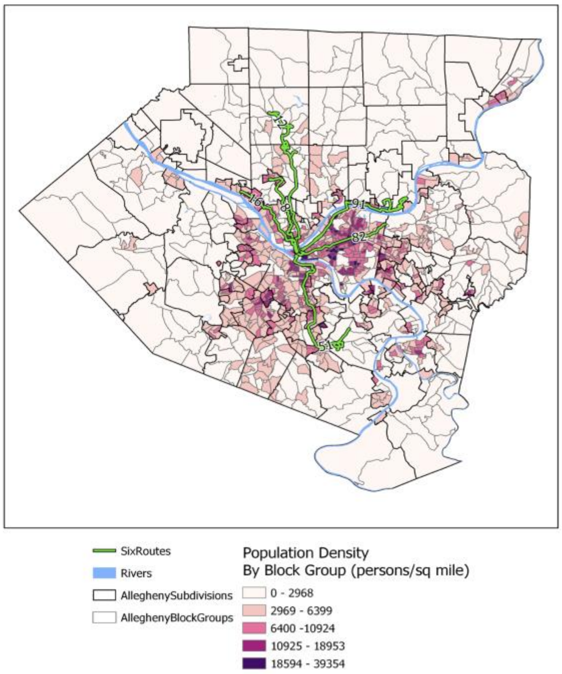


Figure 12: Map of Population Density by Block Group

We use the persons per square mile to measure the population density of each block group.

From the map, the most populated block groups lie in the East part of Pittsburgh city and the Southwestern part adjoining Pittsburgh. Route 82 runs towards Eastern Pittsburgh from the central business district, which covers the block groups with high population density. Route 51 runs to the Southern Allegheny, which also covers a few block groups with high population density. Route 12 goes from CBD Pittsburgh to the North, which passes the block groups with relatively low population density.

Vehicle Ownership

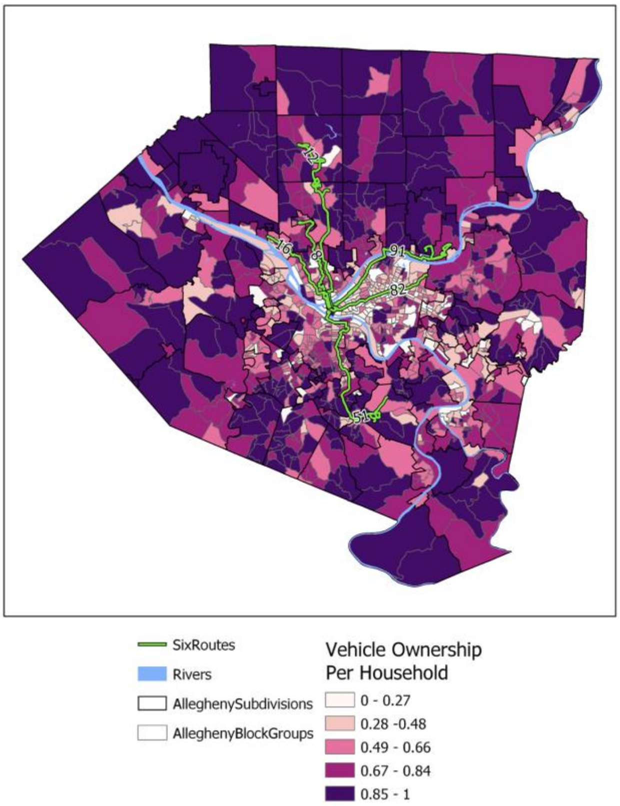


Figure 13: Map of Vehicle Ownership per Household by Group Block in Allegheny County

For the vehicle ownership, we use the number of the vehicle per household as the indicator. From the map, the block groups that locate in the edge of Allegheny County, which is away from Pittsburgh city have high vehicle ownership per household (0.67-1). The Southeastern, the Northwestern, and the Northeastern Allegheny, as well as the center of Pittsburgh city, have low vehicle ownership, which indicates the potential needs for public transportation.

Route 82 runs across the block groups whose vehicle ownership per households is quite low (basically below 0.49). Route 12 and Route 51 runs towards the North and the South separately. These two bus routes are long and run to the suburbs so that they cover some of the block groups whose vehicle ownership is high.

Unemployment Rate

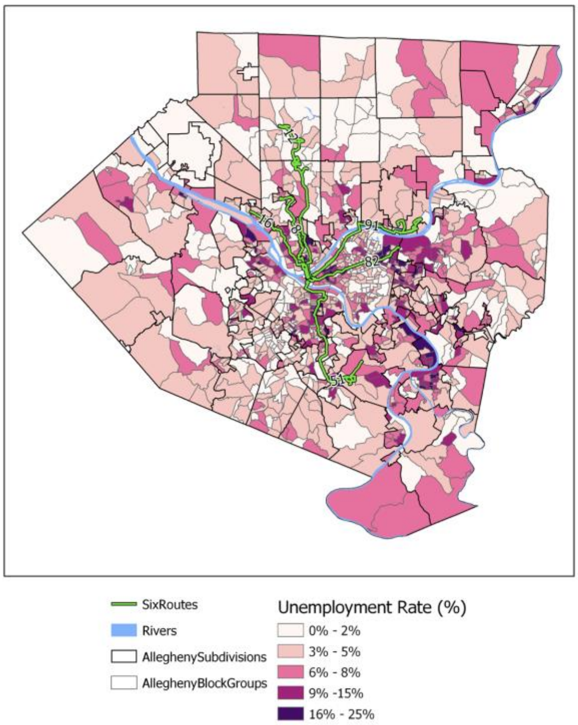


Figure 14: Map of Unemployment Rate by Group Block in Allegheny County

For the unemployment rate, we use the result of unemployed people divided by total population in each age group. We use the census data so the results indicate the unemployment of the residents. The areas with high unemployment rate locate in center Pittsburgh, the Northeastern and the Southeastern parts of Allegheny County who are near the edges of Pittsburgh. Other areas, such as the north side of Allegheny County, normally have a low unemployment rate. The map shows that Route 82, Route 51, and a few parts of Route 12 run across the block groups with high unemployment (over 16%).

College or Above Student Percentage by Block Group

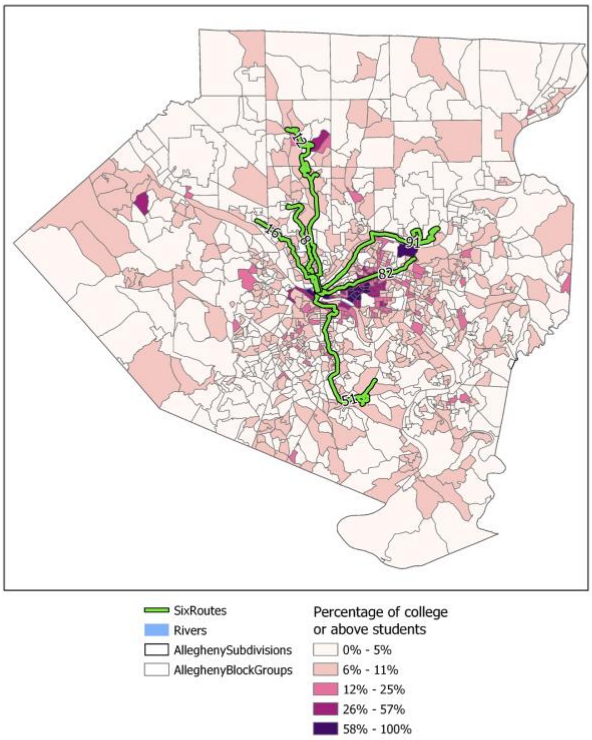


Figure 15: Map of College or Above Student Percentage by Block Group

To measure the effects of the college students who have free bus ID on the bus ridership, we use the data of college or above student percentage in every block group. The map shows that the main college school districts lie in the center and the Eastern parts of Pittsburgh. All the routes cover CBD Pittsburgh, where the college and above students account for high population percentage. Besides, Route 82 covers the areas near the main college school districts in Pittsburgh. Route 91 and Route 12 also run through a few school districts, while the other routes hardly go through the college school districts.

Job Density

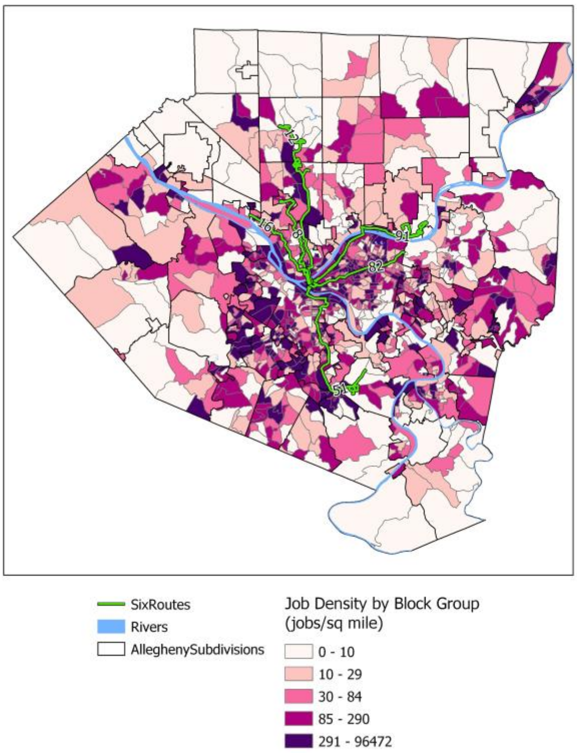


Figure 16: Map of Job Density by Block Group

This map shows the job density of Allegheny County, which means the actual job existing in the areas. From the results, Pittsburgh city has a high volume of job positions, especially the CBD and the eastern parts of the city. Combined with our bus route map, we can conclude that Route 82 and Route 91 covers the areas with high job density, while the other routes run through the block groups with the comparatively lower volume of jobs.

Land Use

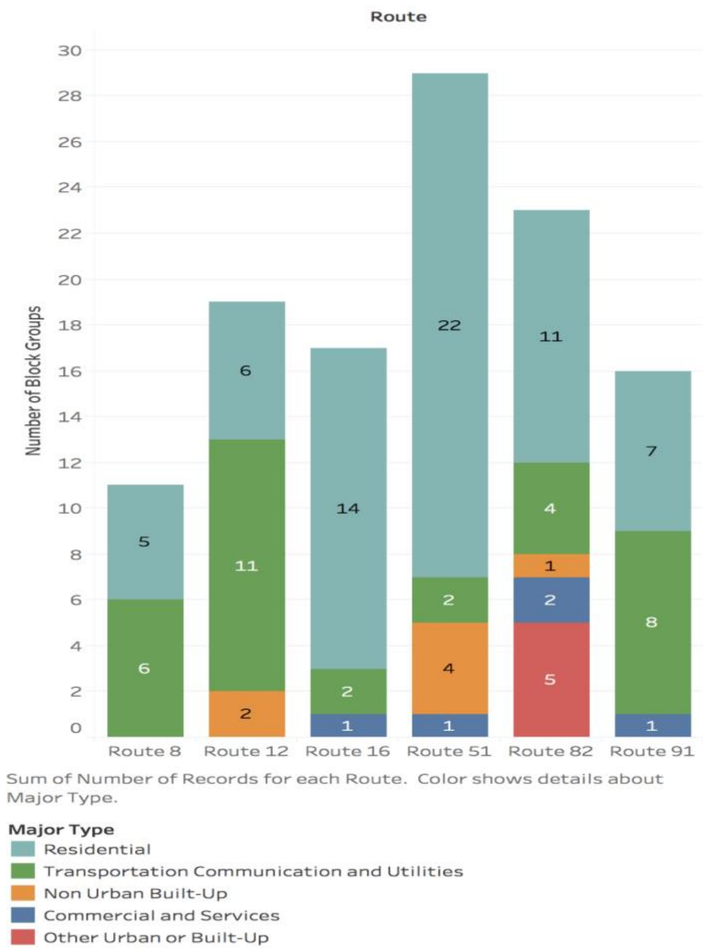


Figure 17: Summary of Main Land Use of the Block Groups by Route

This graph shows the summary of the land use types of the block groups covered by each of our six routes in the study. A block group could contain more than one land use types, and we regard the most significant one as the main type of such area. The results show that all routes run through the residential and transportation areas. The block groups covered by Route 8, Route 12, and Route 91 are mostly transportation areas, while the other routes cover the residential areas the most. The block groups covered by Route 82 has most varieties of land use, which also include non-urban built-up, commercial and services, and other urban or built-up.

Sample Results of 48 Regression Models

The parameters with positive coefficient will have a positive influence on the ridership, and so as the negative ones will negatively influence bus ridership. And performance variables will have a different influence on ridership under a different level of the social demographic variable if the corresponding interaction variable is selected by the model.

Route 51 Outbound Off-Peak

Factors	Coeffcient
Intercept	95.41112
Passenger waiting Time	-0.0113925
Crowding Level	84.57398
Population Density	5.02785E-05
Job Density	0.006595488
% Households Under 80% AMI	-34.58191
Unemployment Rate	-76.81234
Elder Percentage	-280.6613
On Time Performance * Unemployment Rate	-5.126918
Stop Skipping Rate * Population Density	-0.00042773
Stop Skipping Rate * % Households Under 80% AMI	-10.98373

Table 7: Regression Result, Route 51 Outbound Off-Peak

From the table above, we can see that the model tells us there are two performance variables, passenger waiting time and the crowding level have negatively and positively influence the ridership. And there are five social demographic variables are considered as significant and have a different influence on ridership. According to the interaction factors, stop skipping rate will have a different impact on ridership under a different level of population density and income level.

Route 82 Inbound Peak

Factors	Coeffcient
Intercept	9.348114
Crowding Level	-20.21496
Job Density	8.24763E-05
% Households Under 80% AMI	20.27948
Unemployment Rate	-60.34454
Elder Percentage	16.65789
% University Enrolled Student	-15.69872
Major Land Use Type: Mixed Urban or Built-Up	-1.351687
Major Land Use Type: Mixed Forest	-7.699552
Major Land Use Type: Residential	0.03590198
Stop Skipping Rate * % Households Under 80% AMI	39.08559

Table 8: Regression Result, Route 82 Inbound Peak

For route 82 inbound peak hour, only one performance factor is selected, the higher the crowding level, the lower the ridership. But as we can see the interaction factor, stop skipping rate also has an impact on ridership, just under different income level, has a different influence.

* All the 48 model results will be included in the appendix.

Implementations And Recommendations

To show the potential implementations of the 48 regression models and the ridership estimation tool, we take route 51 as an example to generate ridership elasticities.

Ridership Elasticity Estimation of Route 51

1. Generate bus performance sensitivity matrix with correlation analysis.

Performance Factor Sensitivity Matrix (Δfactors = E * Δfrequency)	WaitingTime	On-time Performance	Crowding Level	Stop-Skipping
Inbound Weekday Peak	-134.9	0.03679	0.01986	0.007193
Inbound Weekday Off-Peak	-228.4	0.06475	-0.02675	0.03049
Inbound Saturday	76.32	0.03806	-0.01535	
Inbound Sunday	8.57	-0.002176	0.004123	0.04181
Outbound Weekday Peak	29.67	0.05415	0.009683	0.01527
Outbound Weekday Off-Peak	-161	-0.0008242	0.05485	-0.01647
Outbound Saturday	-370.6	0.08387	-0.0006443	0.04919
Outbound Sunday	9.154	-0.007206	0.004667	0.04308

Table 9: Sensitivity Analysis of Bus Performance Factors for Route 82

In our estimation tool, the part in blue shown in the table is our inputs.

For every route with frequency change ever, we could use the numeric value of frequency to do four correlation analysis with four bus performance factors. With the results of correlation analysis, we could use the coefficients as the corresponding sensitivities. For the 8 scenarios in each route, we could get the corresponding four sensitivities. In this way, we could get a sensitivity matrix. However, since from March 2016 to July 2017, there was no frequency change for the route, we would use the sensitivity matrix of route 82 as an approximation.

We've created a 4 * 8 sensitivity matrix to measure the effects of frequency change on waiting time, on-time performance, crowding level, stop-skipping under the 8 scenarios.

The null value of sensitivity of stop-skipping in inbound Saturday is due to the fact that there is no stop-skipping in this particular model.

2. Import coefficients of bus performance factors and interaction terms from the 48 regression models.

Coefficient Matrix (From Regression Analysis)	WaitingTime	On-time Performance	Crowding Level	Stop-Skipping
Inbound Weekday Peak	1.10E-03	6.84E+00	-3.91E+00	3.59E+01
Inbound Weekday Off-Peak	2.24E-03			1.83E+01
Inbound Saturday	-2.30E-03	1.07E+01		
Inbound Sunday			3.99E+01	
Outbound Weekday Peak			1.32E+01	
Outbound Weekday Off-Peak	-1.14E-02		8.46E+01	
Outbound Saturday			4.50E+01	
Outbound Sunday	6.10E-03		4.01E+01	-3.98E+00

Table 10: Regression Coefficients for 8 Models of Route 51

In our estimation tool, the part in yellow shown in the table is our constants.

We've created a matrix of coefficients of the 8 regression models for route 51. For each bus performance factor and interaction terms selected in the model, there would be an effect on estimated ridership.

The null value of the coefficient in the table is due to the fact that, according to the regression analysis, the factor with a null value of coefficient has no significant effect on estimated ridership. In addition, the selected interaction terms are not shown in the above table but included in the analytic tool.

3. Calculate the estimated ridership elasticities.

Estimated Ridership Elasticity Matrix	ΔRidership (Min)	ΔRidership (Max)	Performance Factor Effects	Interaction Terms Effects(Min)	Interaction Terms Effects(Max)
Inbound Weekday Peak	0.28	0.15	0.283331565	0.00E+00	-1.31E-01
Inbound Weekday Off-Peak	0.05	0.05	0.046978796		
Inbound Saturday	0.23	-0.42	0.23275866	-1.15E-03	-6.48E-01
Inbound Sunday	0.17	0.23	0.164710675	5.72E-04	6.99E-02
Outbound Weekday Peak	0.13	0.26	0.127847207	2.45E-05	1.34E-01
Outbound Weekday Off-Peak	6.47	6.58	6.473075303	4.11E-04	1.08E-01
Outbound Saturday	-0.03	0.79	-0.02900245	2.85E-03	8.15E-01
Outbound Sunday	0.07	0.08	0.071275839	1.83E-06	9.99E-03

Table 11: Estimated Elasticities for 8 Models of Route 51

In our estimation tool, the part in yellow shown in the table is our results.

For each of the 8 scenarios, we've estimated a range of ridership elasticity. Each estimated elasticity reflects the combining effect of performance factors and interaction terms. For the effect of interaction terms, it would involve the effect of socio-demographic factors as well. Thus, the effect of interaction terms varies from selected socio-demographic factors in the regression model.

In general, compared to the benchmark elasticity range of 0.3 to 0.5, the estimated elasticities are within reasonable range. Except for the outbound weekday off-peak model, the estimated elasticity is relatively high, which means that there is a great potential to increase ridership by increasing frequency in this scenario.

Findings and Recommendations for Route 51 and 82

Through our analysis, we found that route 51 and route 82 have the greatest potential to increase ridership by increasing frequency. Especially, the elasticities of the Route 51

Outbound Off-Peak model and Route 82 Inbound Off-Peak model are relatively high among the examined models.

Route 51 Outbound Weekday Off-Peak Model

During the 18 hours and 10 minutes of off-peak operation, there are 85 trips in total, with a ridership of 1,010. The corresponding elasticity of the model is 6.47 to 6.58. Under this elasticity level, when we add one more trip or add one more service hour, there would be 76 to 78 more passengers. According to the estimation of PAAC, one hour of bus service would cost $188.09. Thus, for the increased 76 to 78 ridership, the cost per passenger served is estimated to be $2.41 to $2.47, which is lower than $6.23, the agency's average cost per passenger. In general, it is valuable to add more trips for route 51 outbound during weekday off-peak time.

Route 82 Inbound Weekday Off-Peak Model

During the 16 hour and 30 minutes of off-peak operation, there are 52 trips in total, with a ridership of 878. The corresponding elasticity of the model is 1.91 to 8.12. Under this elasticity level, when we add one more trip or add one more service hour, there would be 32 to 137 more passengers. According to the estimation of PAAC, one hour of bus service would cost $188.09. Thus, for the increased 32 to 137 ridership, the cost per passenger served is estimated to be $1.37 to $5.88, which is lower than $6.23, the agency's average. In general, it is valuable to add more trips for route 82 inbound during weekday off-peak time.

Recommendations for Route 51 and 82

In general, there are two ways to make adjustments for the current bus schedule. The first method is to increase frequency. The second method is to shift the time window of scheduled peak time (with the highest frequency).

Though our analysis of route 51 and route 82, we would make the following recommendations:

1. Add More Trips for Route 51 Outbound during Weekday Off-Peak
With one more trip or one more service hour, it's expected to add 76 to 77 more passengers, with an estimated cost per passenger of $2.41 to $2.47.

2. Add More Trips for Route 82 inbound during Weekday Off-Peak
With one more trip or one more service hour, it’s expected to add 32 to 137 more passengers, with an estimated cost per passenger of $1.37 to $5.88.

3. Shift the Peak Hour Time Window of Route 51 Inbound Earlier
Currently, the scheduled Peak Time starts from 6:10 am with a 10-min headway. However, through our analysis of observed peak time, the peak of observed ridership starts from 4:30 am to 5:00 am.

Conclusion and Future Work

Work Flow

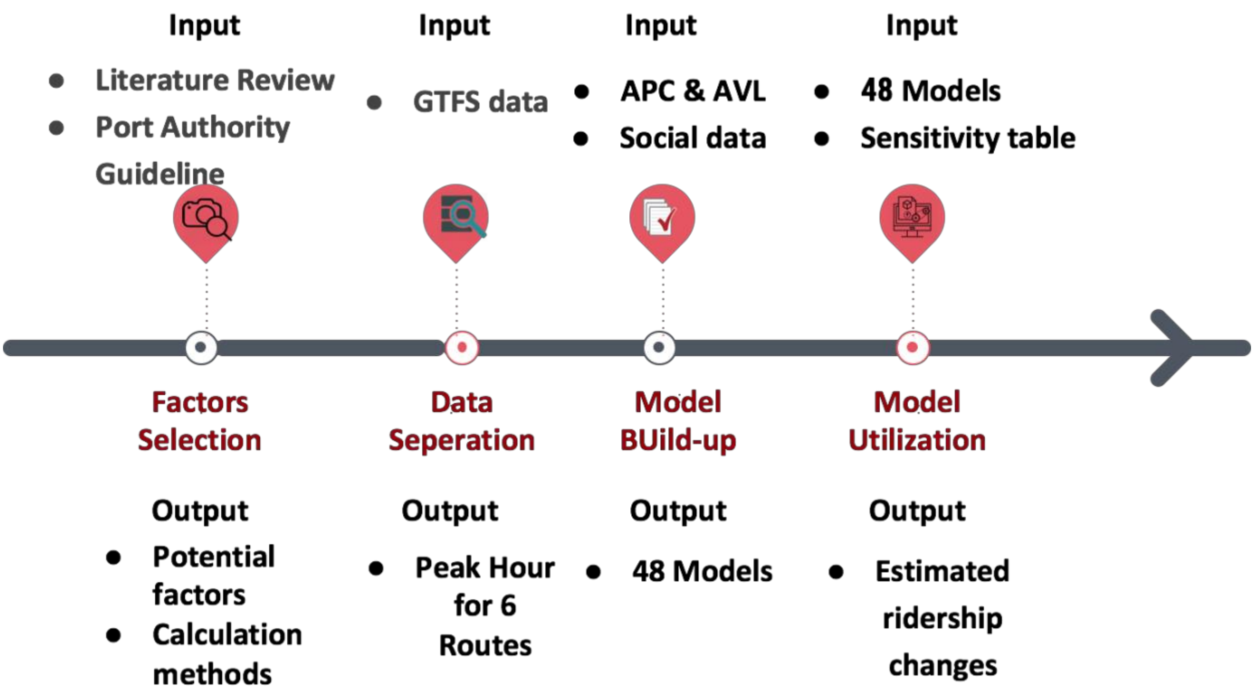


Figure 18: Overall Workflow Chart

Future Work Suggestions

Throughout our project, we found ample justification to create 8 models per route, a high level of detail compared to literature we reviewed. That said, we believe this to be the fewest number of models per route that anyone should consider if one wants to capture the differences between routes within different time periods.

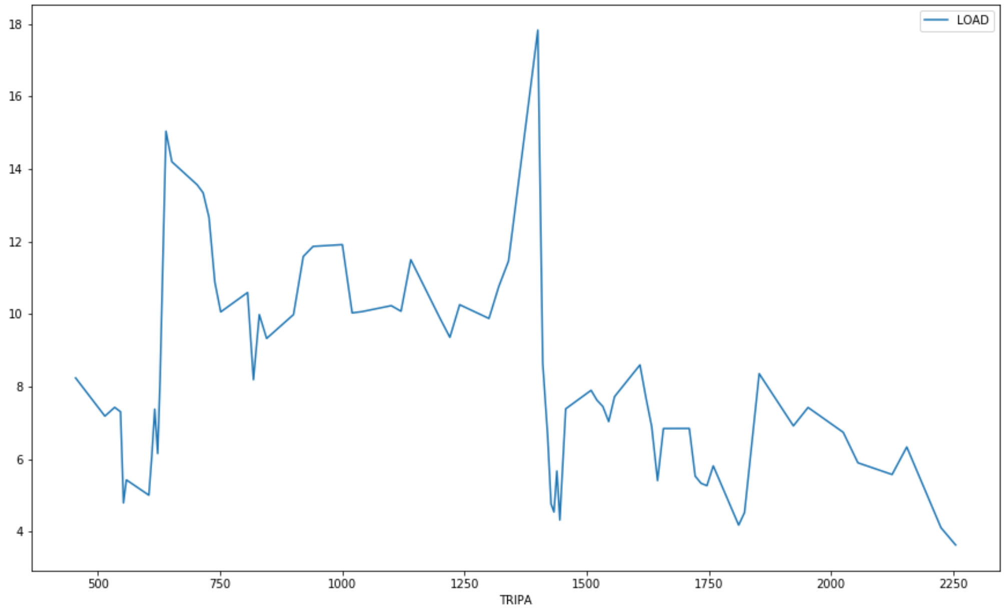
For the future work, we suggest that our clients:

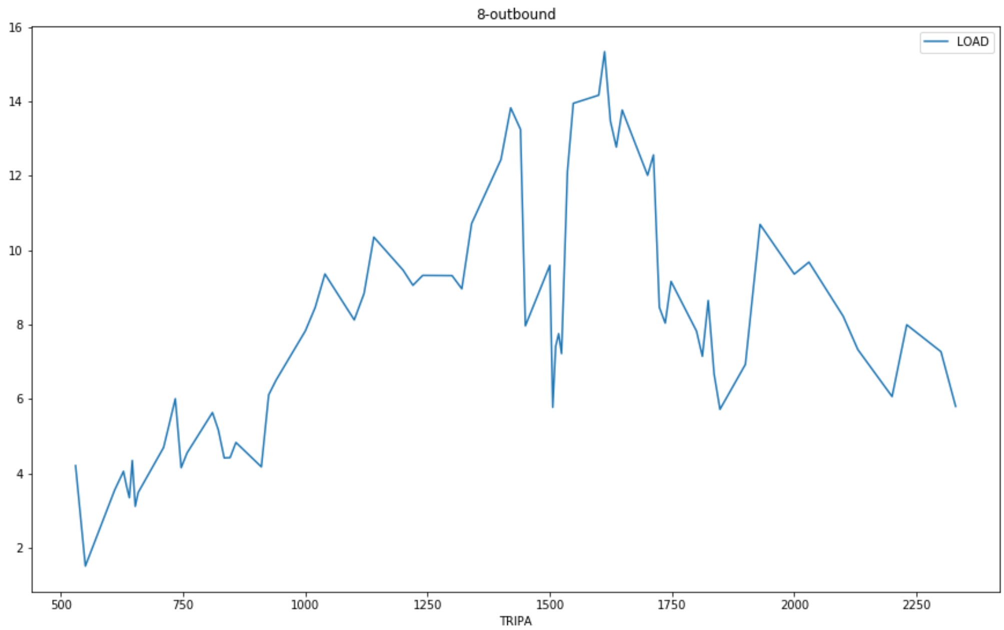
1. Review our 48 models
2. Follow our approach to building sensitivity tables for each route, thus having a more accurate understanding of how frequency changes can influence ridership change.
3. Follow our approach to building new models if needed for additional routes.

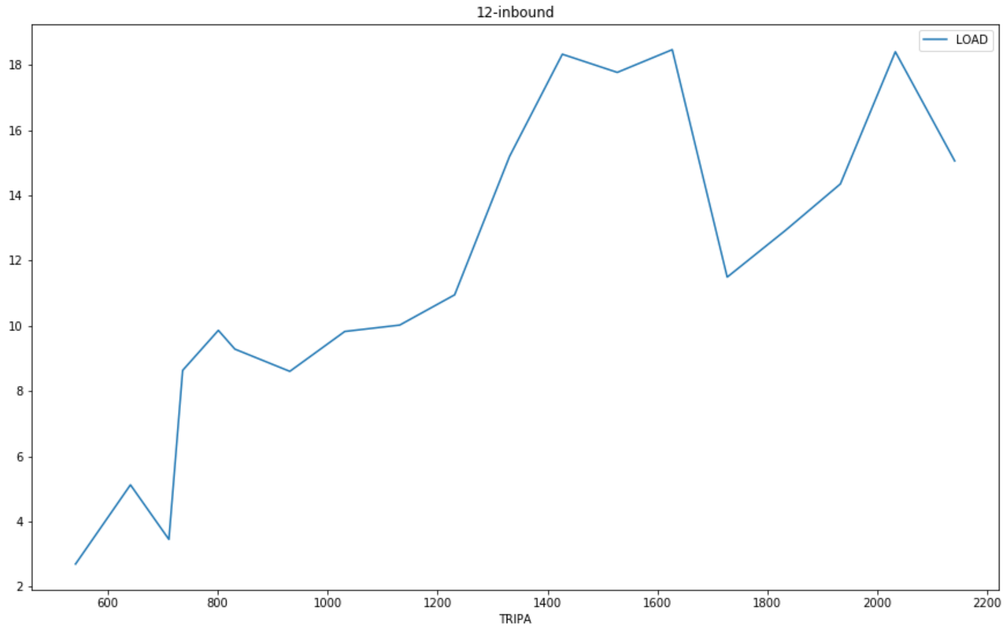
Appendix

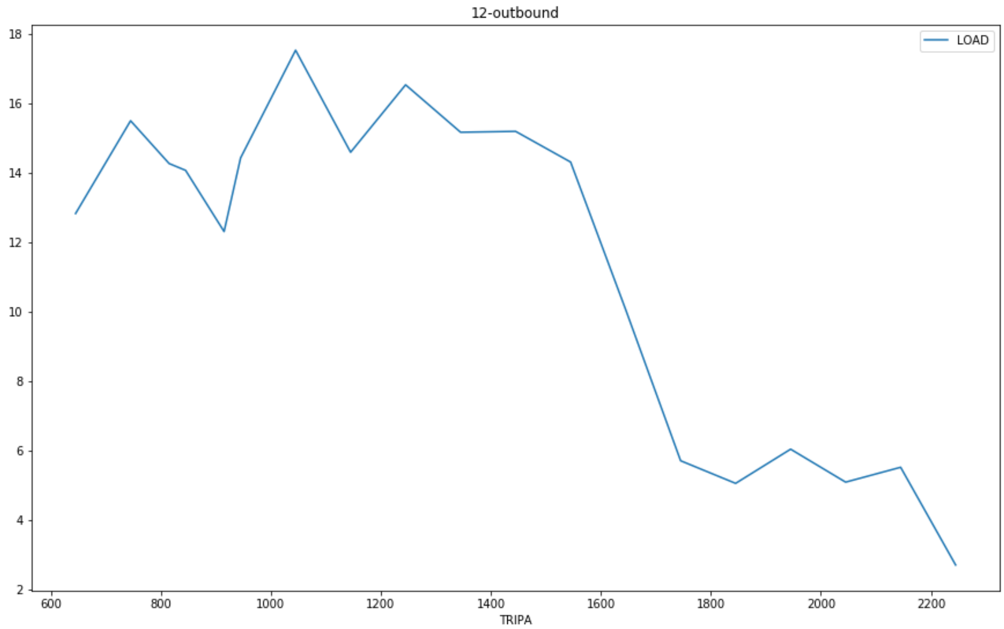
Peak Time

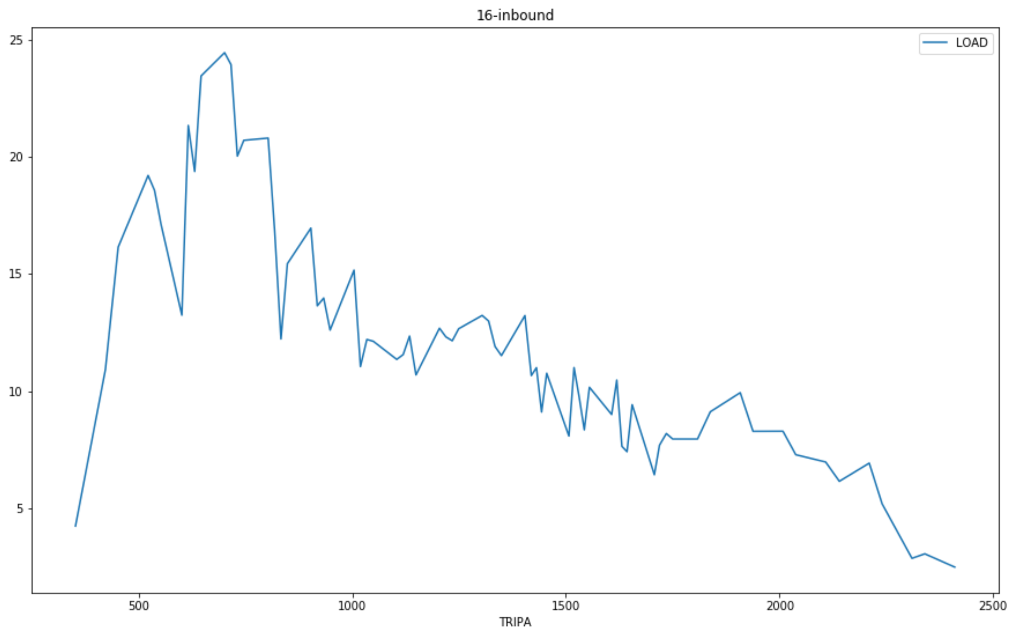
Using average load per trip as y value.

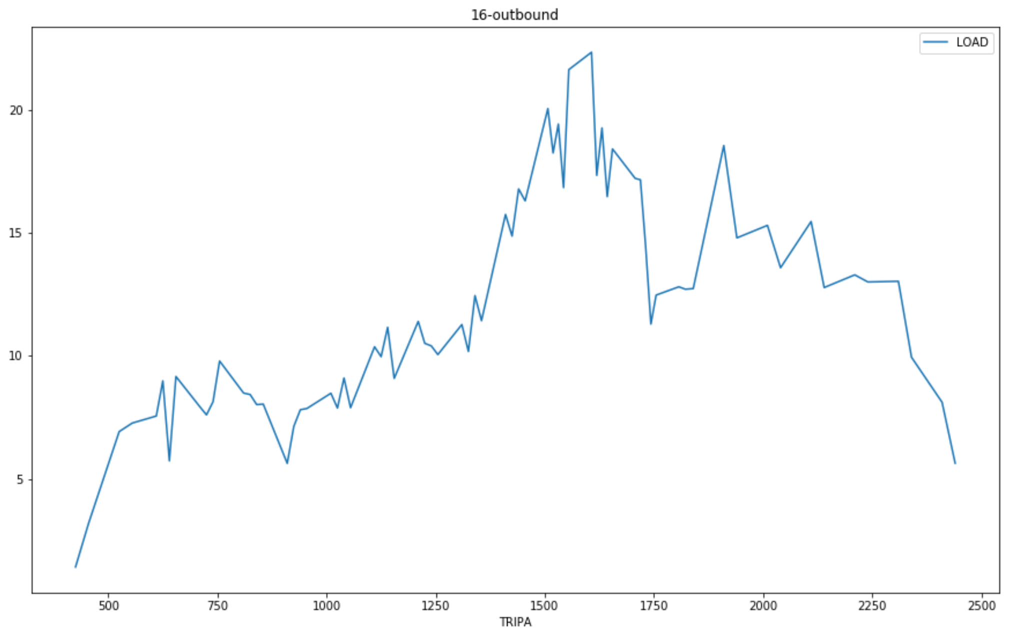


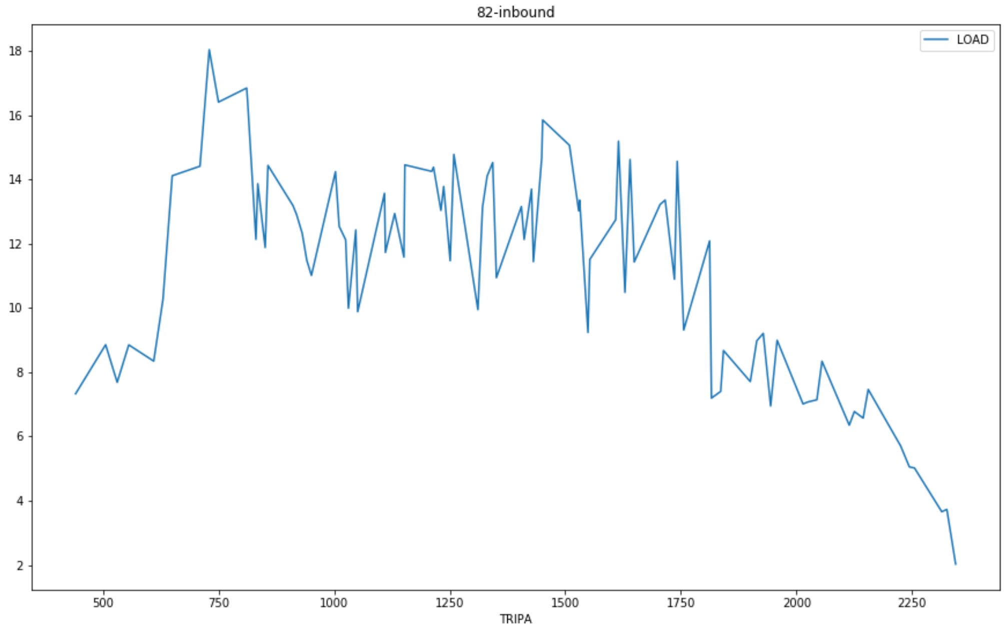


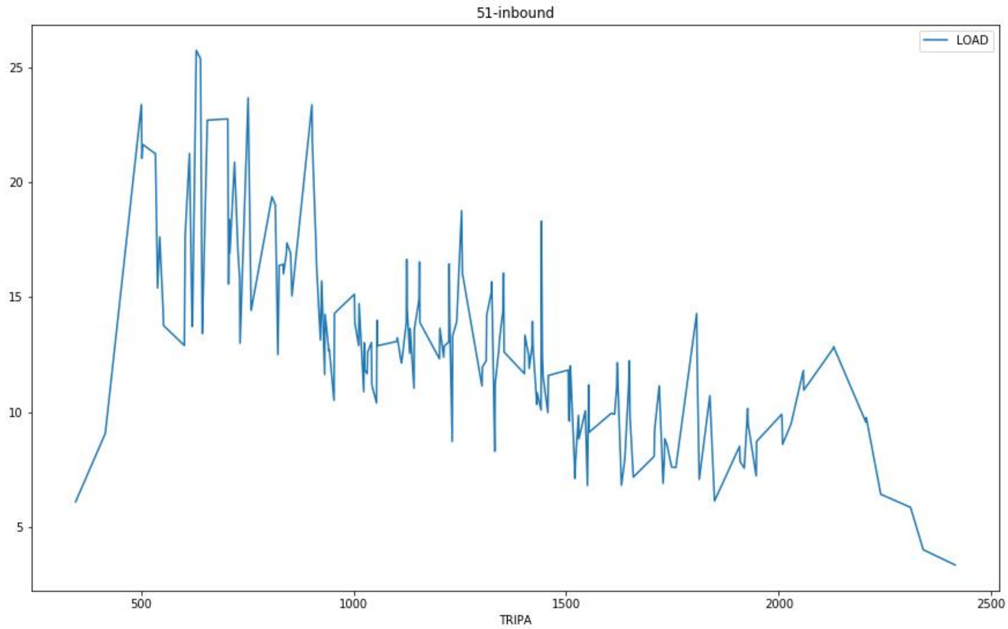

12-inbound
LOAD
TRIPA

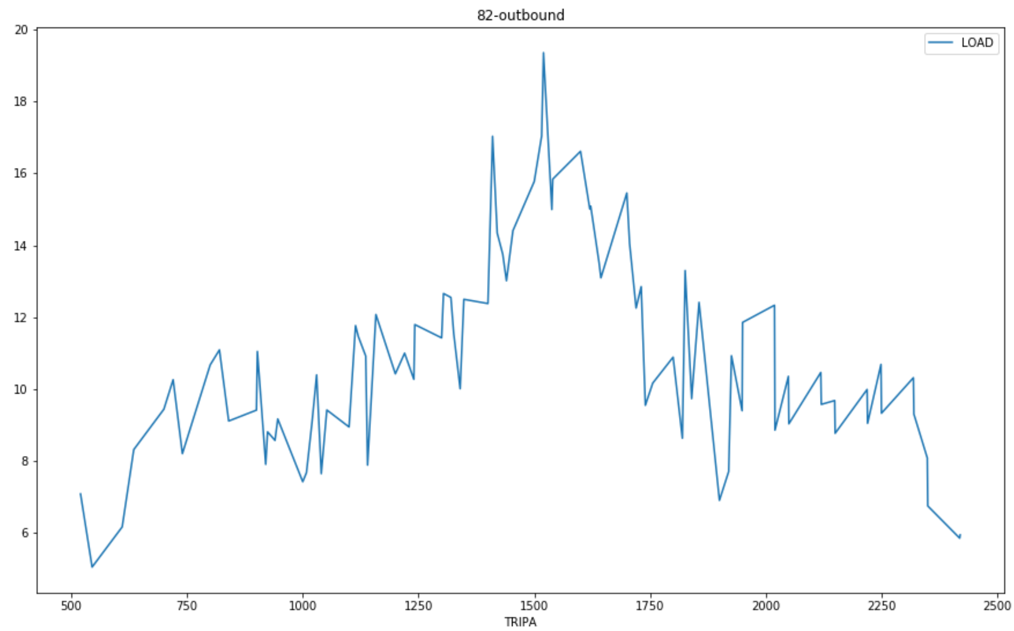

12-outbound
LOAD
TRIPA

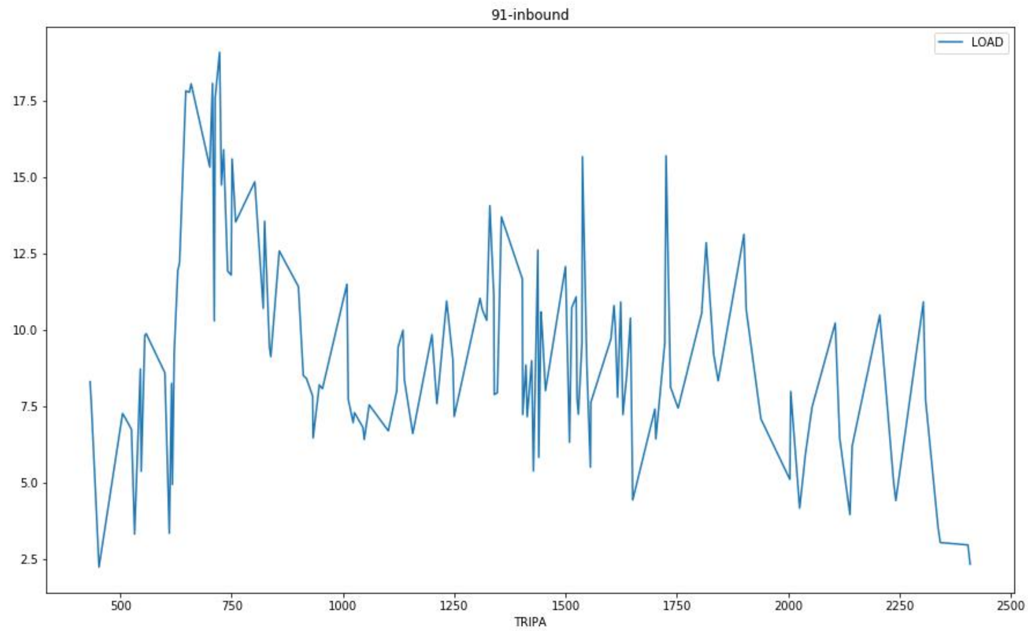

16-inbound
LOAD
TRIPA

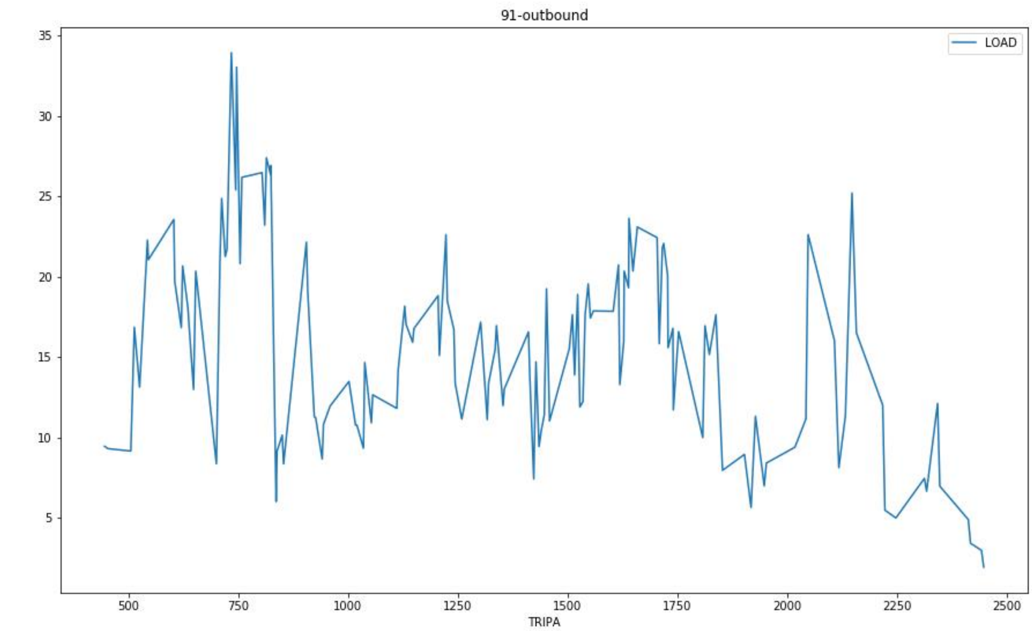

16-outbound
LOAD
TRIPA


82-inbound
LOAD
18
16
14
12
10
8
6
4
2
500
750
1000
1250
1500
1750
2000
2250
TRIPA

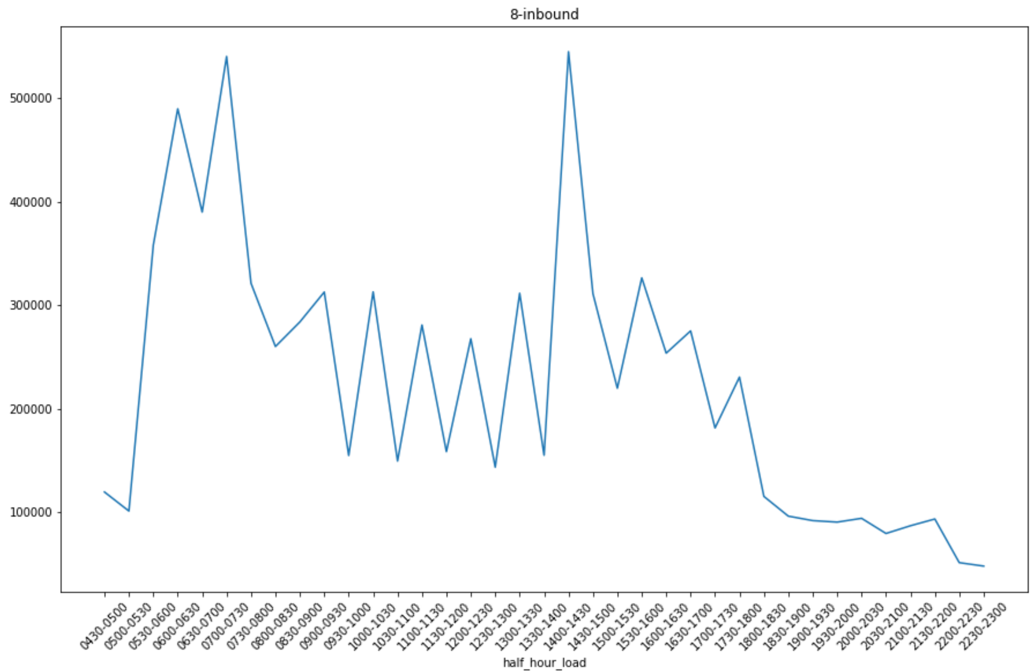

51-inbound
LOAD
25
20
15
10
5
500
1000
1500
2000
2500
TRIPA

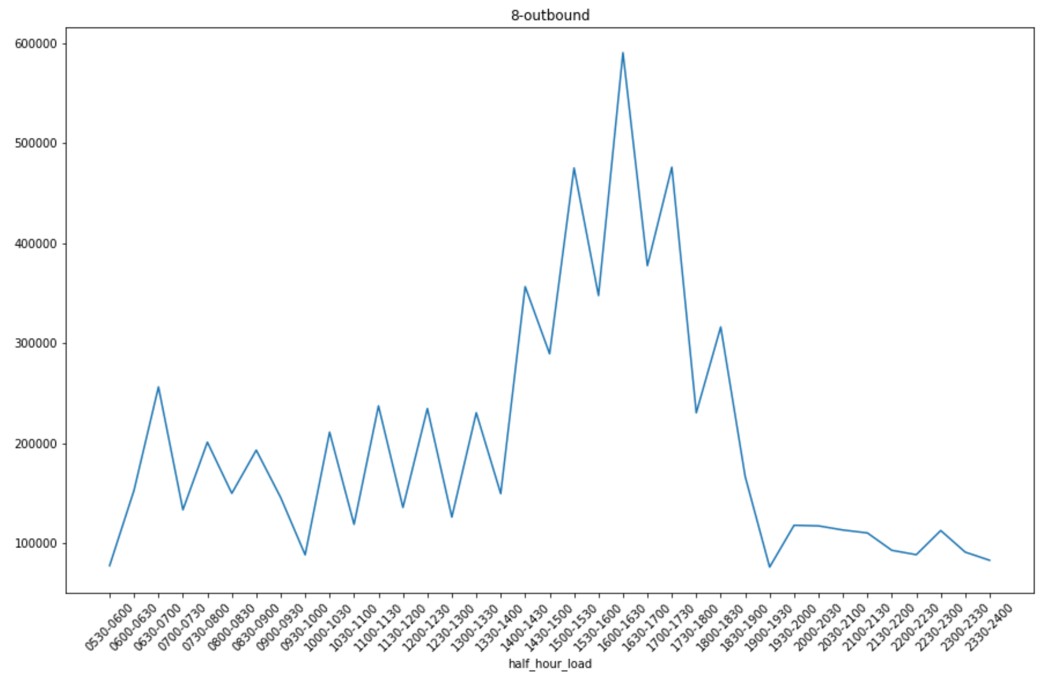

82-outbound
LOAD
TRIPA
20
18
16
14
12
10
8
6
500
750
1000
1250
1500
1750
2000
2250
2500

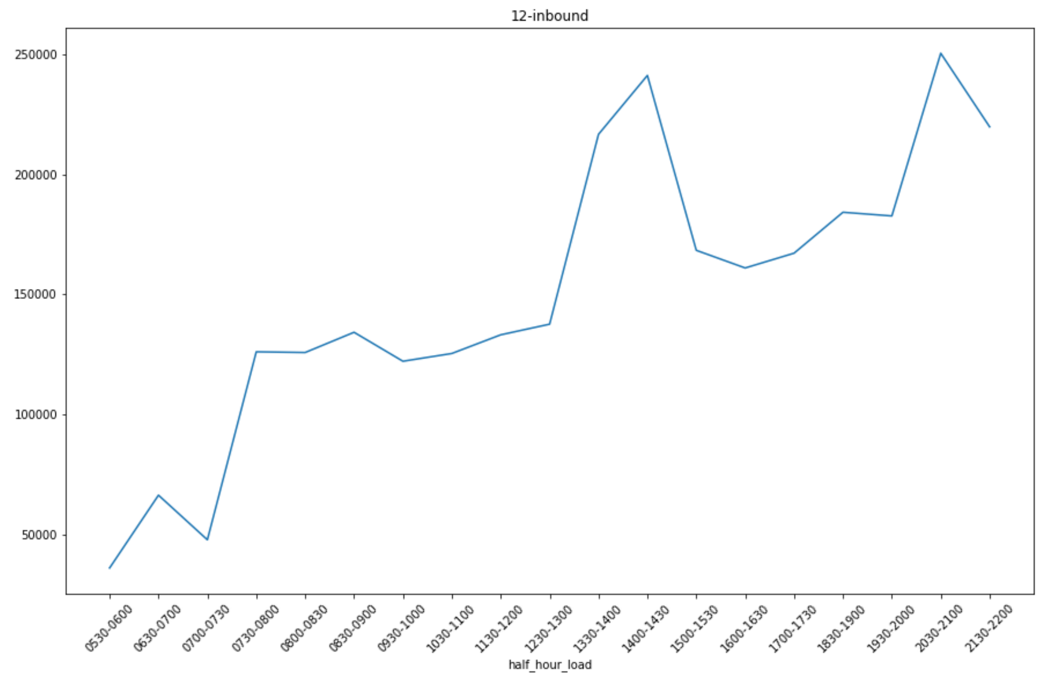

91-inbound
LOAD
TRIPA
17.5
15.0
12.5
10.0
7.5
5.0
2.5
500
750
1000
1250
1500
1750
2000
2250
2500

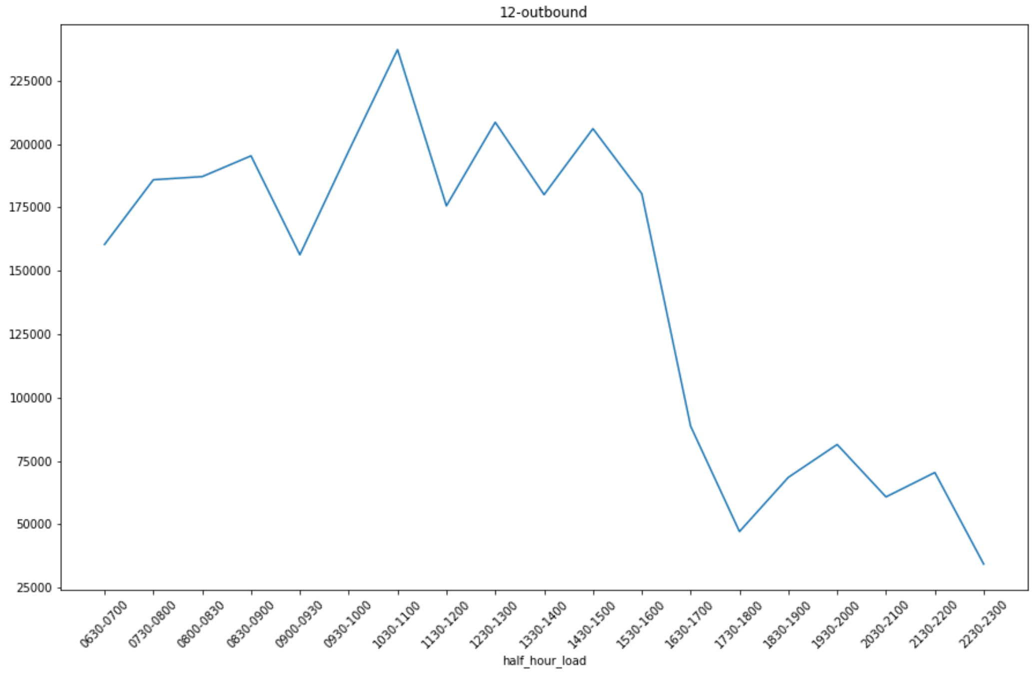


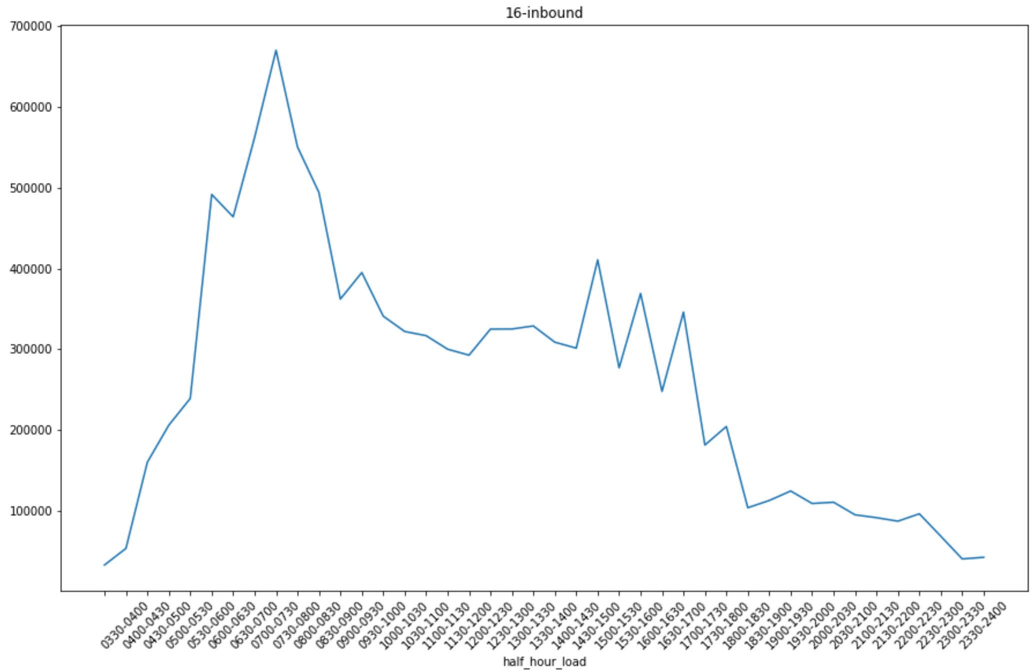
Using total load in half an hour:



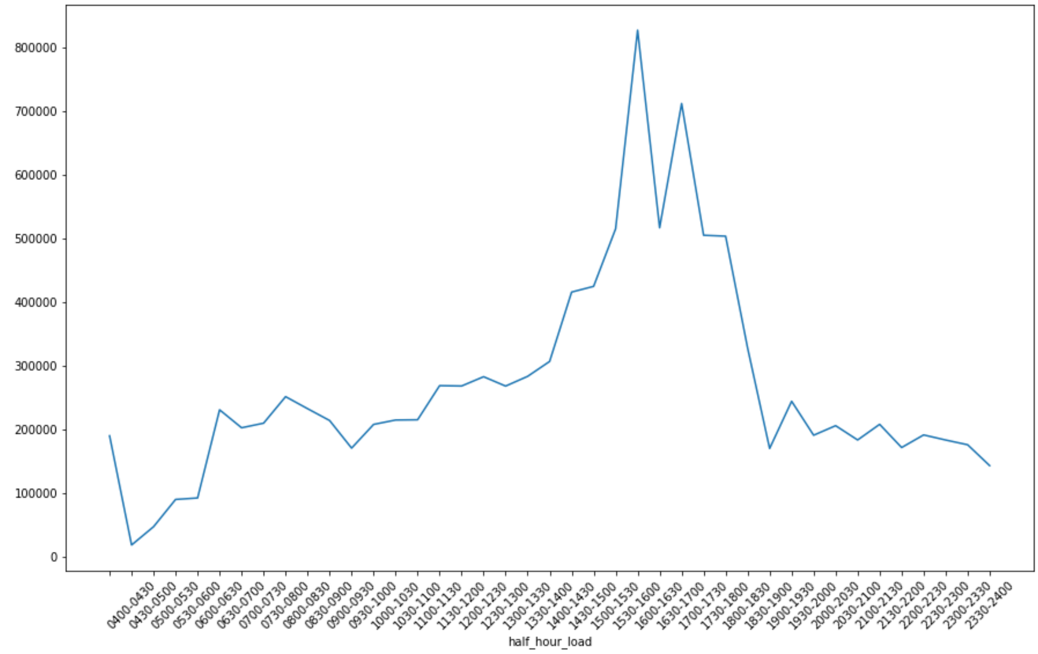

8-outbound
600000
500000
400000
300000
200000
100000
0530-0600
0600-0630
0630-0700
0700-0730
0730-0800
0800-0830
0830-0900
0900-0930
0930-1000
1000-1030
1030-1100
1100-1130
1130-1200
1200-1230
1230-1300
1300-1330
1330-1400
1400-1430
1430-1500
1500-1530
1530-1600
1600-1630
1630-1700
1700-1730
1730-1800
1800-1830
1830-1900
1900-1930
1930-2000
2000-2030
2030-2100
2100-2130
2130-2200
2200-2230
2230-2300
2300-2330
2330-2400
half_hour_load


12-inbound
250000
200000
150000
100000
50000
0530-0600
0630-0700
0700-0730
0730-0800
0800-0830
0830-0900
0930-1000
1030-1100
1130-1200
1230-1300
1330-1400
1400-1430
1500-1530
1600-1630
1700-1730
1830-1900
1930-2000
2030-2100
2130-2200
half_hour_load

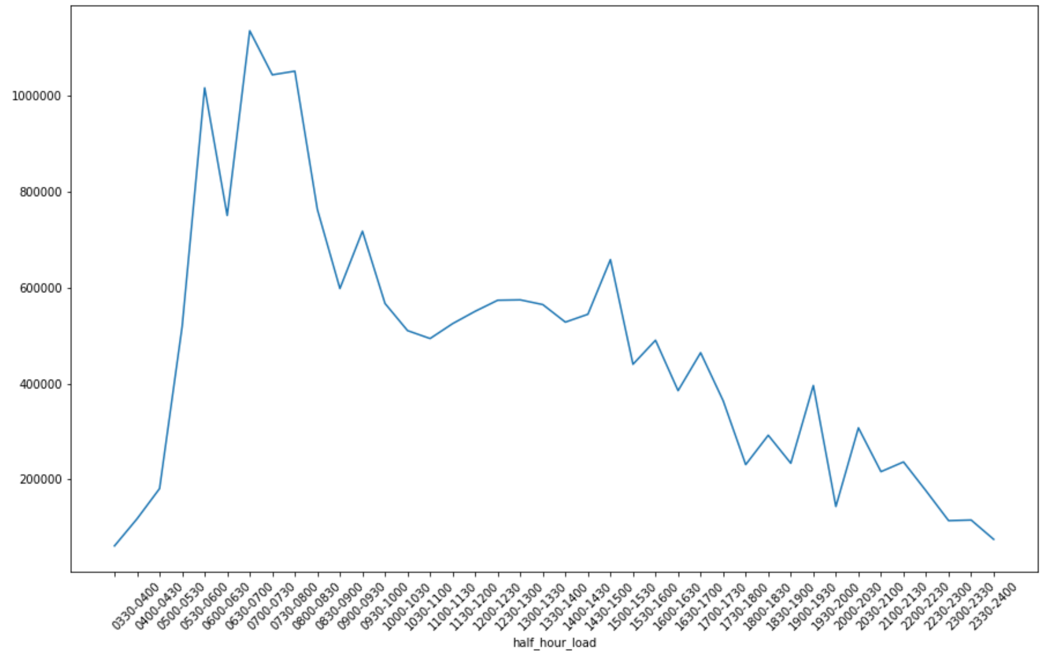

12-outbound
225000
200000
175000
150000
125000
100000
75000
50000
25000
0630-0700
0730-0800
0800-0830
0830-0900
0900-0930
0930-1000
1030-1100
1130-1200
1230-1300
1330-1400
1430-1500
1530-1600
1630-1700
1730-1800
1830-1900
1930-2000
2030-2100
2130-2200
2230-2300
half_hour_load

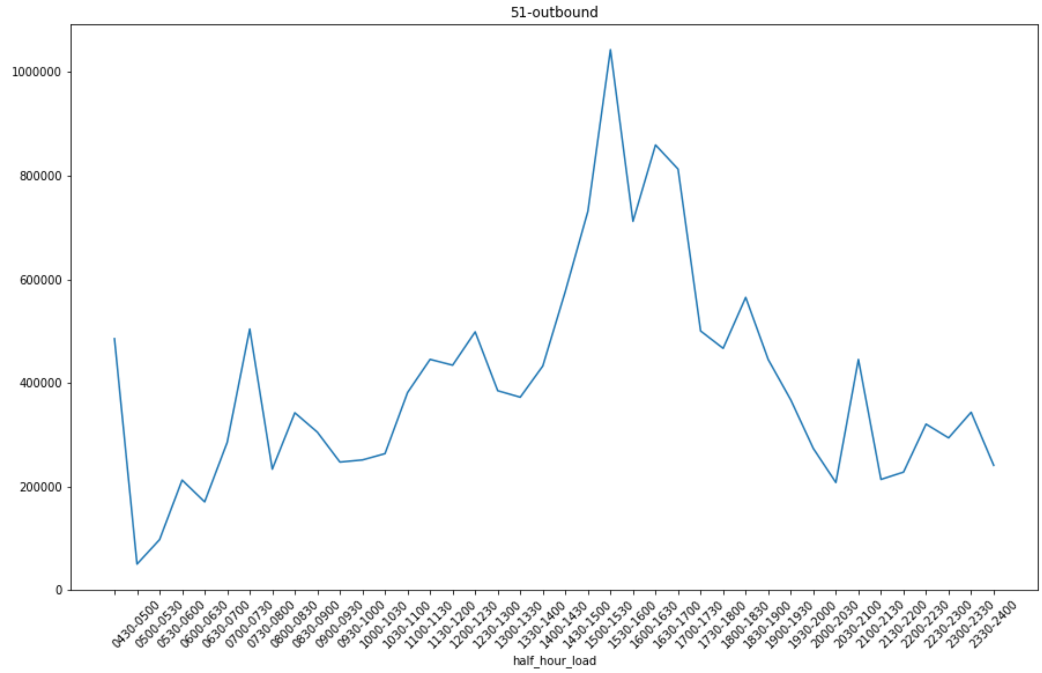

16-inbound
700000
600000
500000
400000
300000
200000
100000
half_hour_load

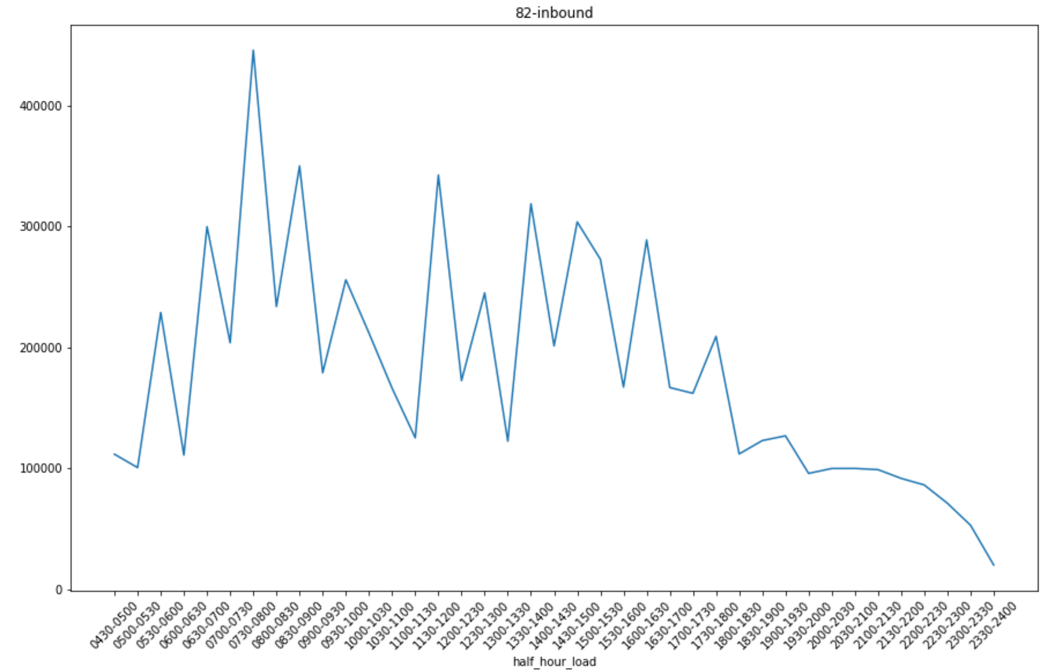
16-outbound

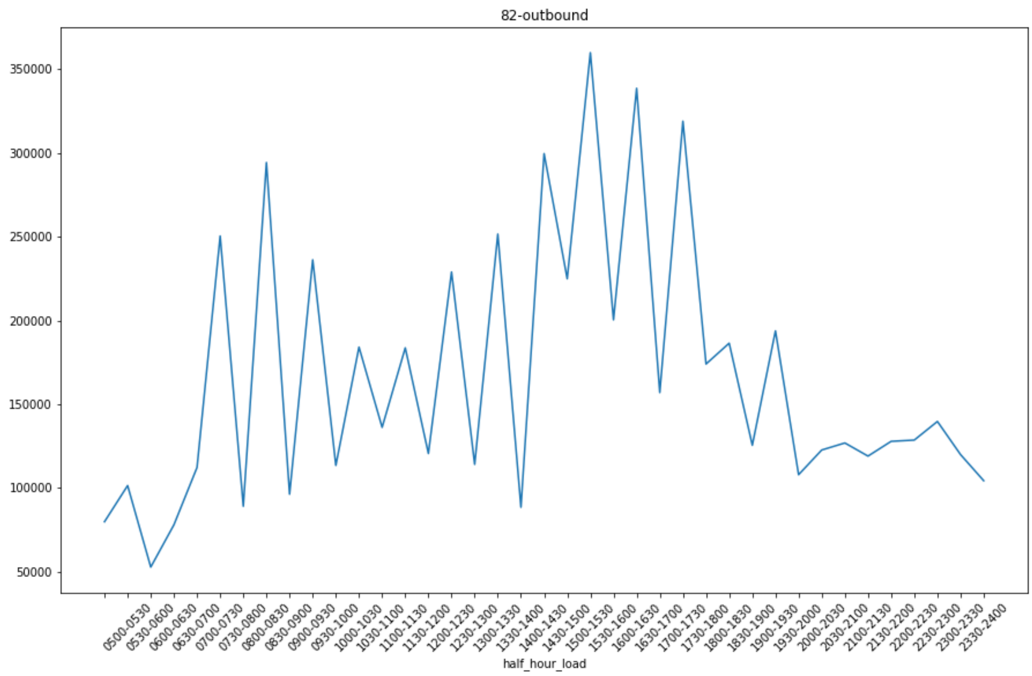


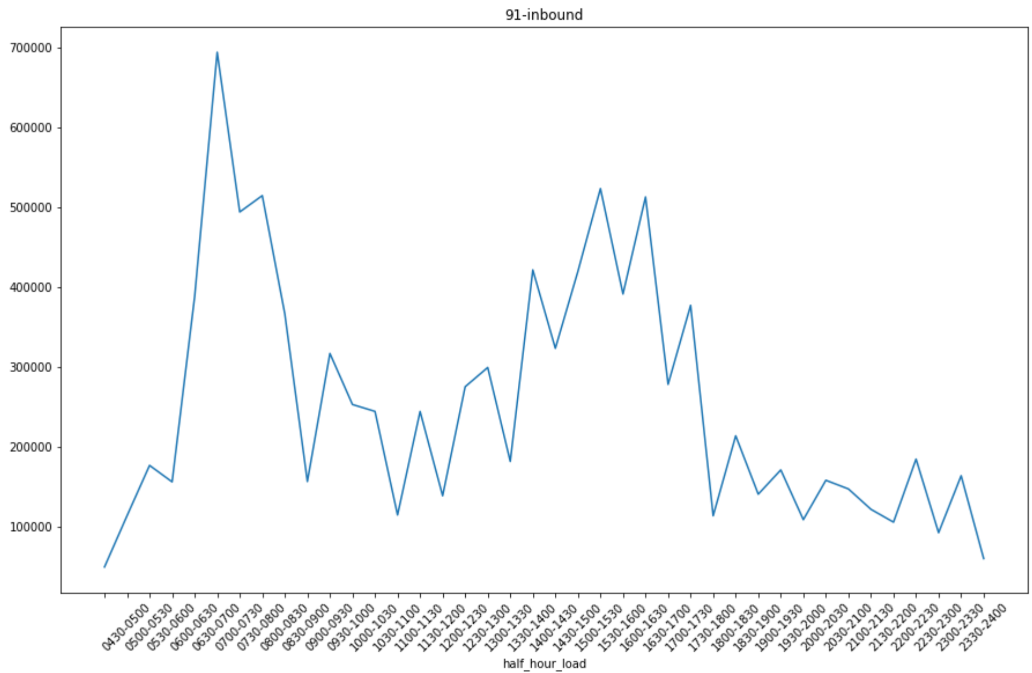
51-inbound

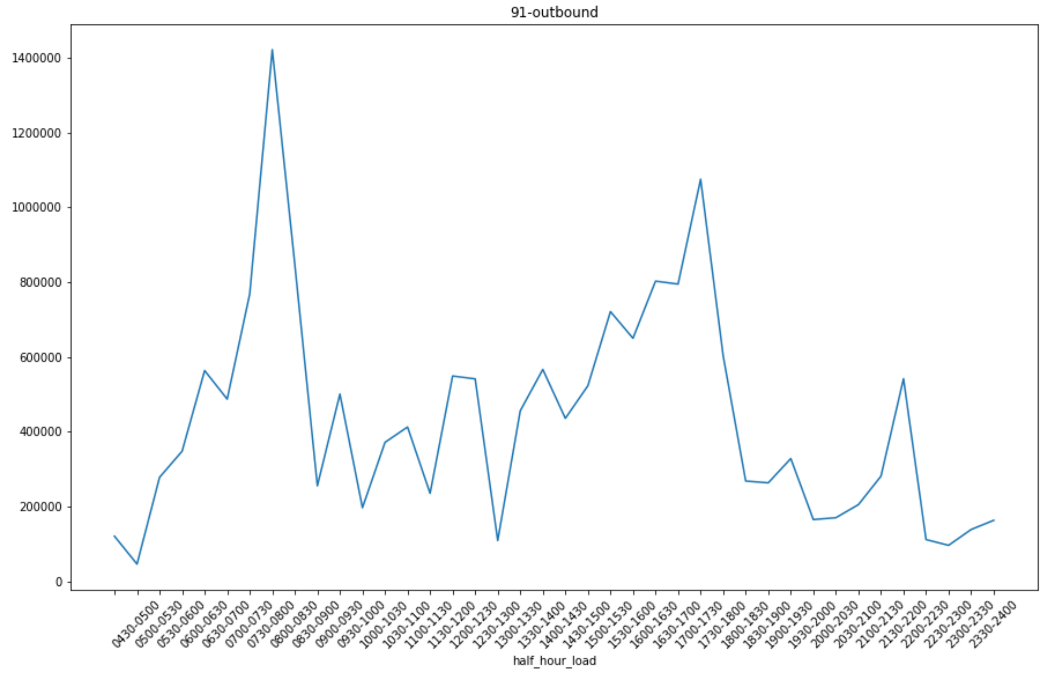



51-outbound
half_hour_load


82-inbound
half_hour_load


82-outbound
350000
300000
250000
200000
150000
100000
50000
0500-0530
0530-0600
0600-0630
0630-0700
0700-0730
0730-0800
0800-0830
0830-0900
0900-0930
0930-1000
1000-1030
1030-1100
1100-1130
1130-1200
1200-1230
1230-1300
1300-1330
1330-1400
1400-1430
1430-1500
1500-1530
1530-1600
1600-1630
1630-1700
1700-1730
1730-1800
1800-1830
1830-1900
1900-1930
1930-2000
2000-2030
2030-2100
2100-2130
2130-2200
2200-2230
2230-2300
2300-2330
2330-2400
half_hour_load


91-inbound
700000
600000
500000
400000
300000
200000
100000
0430-0500
0500-0530
0530-0600
0600-0630
0630-0700
0700-0730
0730-0800
0800-0830
0830-0900
0900-0930
0930-1000
1000-1030
1030-1100
1100-1130
1130-1200
1200-1230
1230-1300
1300-1330
1330-1400
1400-1430
1430-1500
1500-1530
1530-1600
1600-1630
1630-1700
1700-1730
1730-1800
1800-1830
1830-1900
1900-1930
1930-2000
2000-2030
2030-2100
2100-2130
2130-2200
2200-2230
2230-2300
2300-2330
2330-2400
half_hour_load



Regression Results

Route 82 Inbound Weekday Peak

(Intercept)	9.348114e+00
Bus Crowding Level	-2.021496e+01
Job Density	8.247632e-05
% Household under 80% AMI	2.027948e+01
Unemployment Rate	-6.034454e+01
Elder Percentage	1.665789e+01
University Enrolled Students	-1.569872e+01
Major Land Use Type: Mixed Urban or Built-Up	-1.351687e+00
Major Land Use Type: Mixed Forest	-7.699552e+00
Major Land Use Type: Residential	3.590198e-02
Stop Skipping Rate * % Household under 80% AMI	3.908559e+01

R^2

0.6441884

MSE

3.599473

Route 82 Inbound Weekday Off-Peak

(Intercept)	6.560912e+01
Bus Crowding Level	-7.157158e+01
% Household under 80% AMI	3.982462e+01
Vehicle Per Household	-3.548215e+01
Elder Percentage	4.704012e+01
University Enrolled Students	-7.351170e+01
Major Land Use Type: Mixed Forest	-1.365807e+01
Waiting Time * Unemployment Rate	-1.421541e-01
On-time Performance * Population Density	-1.442083e-04

R^2

0.6787185

MSE

32.65366

Route 82 Inbound Saturday

(Intercept)	6.784631e+01
Bus Crowding Level	-2.930322e+01
Population Density	-8.032872e-04
% Household under 80% AMI	5.275047e+01
Unemployment Rate	-3.049951e+02
Elder Percentage	7.658322e+01
University Enrolled Students	-5.092841e+01
Major Land Use Type: Mixed Forest	-1.332190e+01
Major Land Use Type: Residential	1.479922e+00
Major Land Use Type: Transportation Utilities	-1.167e+01
On-time Performance * Vehicle Ownership	-8.123234e+01
Crowding Level * Elder Percentage	-3.457228e+02
Crowding Level * University Enrolled Students	-1.474079e+02

R^2

```
0.7650725
```

MSE

```
24.22687
```

Route 82 Inbound Sunday

(Intercept)	1.891372e+01
Waiting Time	4.023924e-03
On-time Performance	-4.001198e+01
Bus Crowding Level	8.872973e+01
Job Density	-2.231309e-03
% Household under 80% AMI	-4.604919e+01
Unemployment Rate	-1.297010e+02
Vehicle Per Household	-5.425865e+01
Elder Percentage	-2.731520e+01
University Enrolled Students	-4.407355e+00
Major Land Use Type: Commercial and Services	1.738183e+02
Major Land Use Type: Mixed Urban or Built-Up	7.296053e+01
Major Land Use Type: Residential	9.380383e+01
Major Land Use Type: Transportation Utilities	4.539e+01
Waiting Time * Population Density	3.823431e-07
Waiting Time * Job Density	-7.489267e-07
On-time Performance * Vehicle Ownership	-1.701000e+01
On-time Performance * Elder Percentage	-7.261967e+01

R^2

```
0.4585308
```

MSE

```
46.19814
```

Route 82 Outbound Weekday Peak

(Intercept)	-1.167271e+00
On-time Performance	1.202550e+01
Stop Skipping Rate	1.506576e+01
Population Density	-1.895911e-04
Job Density	-2.097112e-03
% Household under 80% AMI	-2.812406e+01
Unemployment Rate	-9.706466e+01
Vehicle Per Household	-4.694868e+01
Elder Percentage	-4.850430e+01
Major Land Use Type: Commercial and Services	1.041310e+02
Major Land Use Type: Mixed Urban or Built-Up	5.863004e+01
Major Land Use Type: Residential	6.841700e+01
Major Land Use Type: Transportation Utilities	3.857569e+01
On-time Performance * Population Density	7.805264e-08
On-time Performance * University Enrolled Students	-4.962682e+00

R^2

```
0.4818866
```

MSE

```
31.86633
```

Route 82 Outbound Weekday Off-Peak

(Intercept)	2.196600e+01
Waiting Time	-1.147630e-02
On-time Performance	2.063963e+01
Job Density	-2.779212e-03
% Household under 80% AMI	-4.426015e+01
Unemployment Rate	-1.533342e+02
Vehicle Per Household	-7.565153e+01
Elder Percentage	-8.641614e+01
University Enrolled Students	3.320336e+00
Major Land Use Type: Commercial and Services	1.961891e+02
Major Land Use Type: Mixed Urban or Built-Up	8.039593e+01
Major Land Use Type: Residential	1.066210e+02
Major Land Use Type: Transportation Utilities	5.055785e+01
Waiting Time * Job Density	-1.495268e-07
On-time Performance * Job Density	-1.272834e-04
Stop Skipping Rate * Population Density	7.753968e-03

R^2

```
0.3997097
```

MSE

```
67.88565
```

Route 82 Outbound Saturday

(Intercept)	1.373711e+01
Waiting Time	2.831842e-03
On-time Performance	-2.707647e+01
Bus Crowding Level	4.979388e+01
Job Density	-3.232042e-03
% Household under 80% AMI	-5.882122e+01
Unemployment Rate	-1.238552e+02
Vehicle Per Household	-6.921714e+01
Elder Percentage	-8.576567e+01
University Enrolled Students	-2.121188e+00
Major Land Use Type: Commercial and Services	1.860277e+02
Major Land Use Type: Mixed Urban or Built-Up	8.396207e+01
Major Land Use Type: Residential	1.104501e+02
Major Land Use Type: Transportation Utilities	5.883e+01
Waiting Time * Population Density	3.507603e-07

R^2

```
0.6825003
```

MSE

```
52.95166
```

Route 82 Outbound Saturday

(Intercept)	2.137676e+01
Waiting Time	3.977333e-03
On-time Performance	-4.285683e+01
Bus Crowding Level	8.719459e+01
Job Density	-2.189942e-03
% Household under 80% AMI	-4.664091e+01
Unemployment Rate	-1.303458e+02
Vehicle Per Household	-6.632738e+01
Elder Percentage	-2.023455e+01
University Enrolled Students	-4.683998e+00
Major Land Use Type: Commercial and Services	1.734816e+02
Major Land Use Type: Mixed Urban or Built-Up	7.343850e+01
Major Land Use Type: Residential	9.452362e+01
Major Land Use Type: Transportation Utilities	4.585655e+01
Waiting Time * Population Density	3.892794e-07
Waiting Time * Job Density	-7.853191e-07
On-time Performance * Elder Percentage	-8.388101e+01

R^2

```
0.6267403
```

MSE

```
44.07142
```

Route 8 Inbound Weekday Peak

(Intercept)	1.008534e+01
Bus Crowding Level	-5.602439e+01
Population Density	3.115785e-03
Job Density	-1.148865e-01
% Household under 80% AMI	6.302666e+01
Unemployment Rate	-2.787170e+02
Vehicle Per Household	-1.384029e+02
Elder Percentage	2.697088e+02
University Enrolled Students	7.334213e+02
Major Land Use Type: Residential	2.884605e+01
Major Land Use Type: Transportation Utilities	-4.288094e+00
Stop Skipping Rate * Elder Percentage 3.061427e+02	

R^2

```
0.7236335
```

MSE

```
28.5125
```

Route 8 Inbound Weekday Off-Peak

(Intercept)	5.719240e+01
Bus Crowding Level	-3.722470e+02
Population Density	3.456396e-03
Job Density	-1.400743e-01
% Household under 80% AMI	1.245887e+02
Unemployment Rate	-3.795117e+02
Vehicle Per Household	-1.394770e+02
Elder Percentage	6.713427e+01
University Enrolled Students	4.035479e+02
Major Land Use Type: Residential	2.066355e+01
Major Land Use Type: Transportation Utilities	9.842824e+00

R^2

```
0.7394919
```

MSE

```
22.29985
```

Route 8 Inbound Saturday

(Intercept)	1.087245e+02
Bus Crowding Level	-5.768036e+02
Population Density	5.680390e-03
Job Density	-1.887476e-01
% Household under 80% AMI	2.161644e+02
Unemployment Rate	-4.213863e+02
Vehicle Per Household	-2.239074e+02
Elder Percentage	3.448090e+01
University Enrolled Students	3.942641e+02
Major Land Use Type: Residential	2.576904e+01
Major Land Use Type: Transportation Utilities	1.123130e+01
Stop Skipping Rate	8.660808e+01
Stop Skipping Rate * Job Density	1.927334e-04

R^2

0.8011891

MSE

26.41213

Route 8 Inbound Sunday

AIC Model

(Intercept)	-1.381288e+02
Waiting Time	6.347187e-02
Bus Crowding Level	5.251871e+01
Population Density	2.454414e-03
Job Density	3.744823e-04
% Household under 80% AMI	3.220410e+02
Unemployment Rate	-1.086343e+02
Elder Percentage	2.582820e+01
University Enrolled Students	-1.087646e+03
Major Land Use Type: Transportation Utilities	8.114e+00
Waiting Time * % Household Under 80% AMI	-1.682379e-01
Waiting Time * University Enrolled Students	5.514820e-01
On-time Performance * University Enrolled Students	1.138182e+00
Stop Skipping Rate * Elder Percentage	1.128339e+03
Stop Skipping Rate * University Enrolled Students	-3.133966e+03

R^2

0.7624918

MSE

25.12497

BIC Model

(Intercept)	2.916639e+00
Bus Crowding Level	9.369625e+01
Population Density	2.094845e-03
Job Density	2.306263e-04
% Household under 80% AMI	-2.857718e+01
Unemployment Rate	-1.508252e+02

R^2

0.5761661

MSE

26.21329

Route 8 Outbound Weekday Peak

(Intercept)	-1.447283e+01
On-time Performance	9.519140e+01
Bus Crowding Level	8.982825e+01
Job Density	1.716405e-04
Unemployment Rate	-4.835909e+01
Vehicle Per Household	-2.980445e+01
Elder Percentage	1.320379e+01
University Enrolled Students	8.094925e+01
Major Land Use Type: Transportation Utilities	1.416760e+00
Stop Skipping Rate	-1.396139e+02
Waiting Time * Population Density	-6.064086e-07
Waiting Time * Unemployment Rate	-3.238657e-02
Waiting Time * Elder Percentage	4.629702e-02
On-time Performance * Population Density	5.785692e-03
On-time Performance * % Household Under 80% AMI	-1.692437e+02
On-time Performance * Unemployment Rate	-5.279566e+01
Stop Skipping Rate * Job Density	8.161216e-04
Stop Skipping Rate * Vehicle Ownership	2.813885e+02

R^2

0.7205026

MSE

71.81174

Route 8 Outbound Weekday Off-Peak

(Intercept)	3.489546e+01
Bus Crowding Level	2.280773e+02
% Household under 80% AMI	-9.419216e+01
Unemployment Rate	-1.206454e+02
Vehicle Per Household	-1.604450e+01
Elder Percentage	1.309782e+01
University Enrolled Students	1.437628e+02
Waiting Time * Population Density	1.188696e-06
Waiting Time * University Enrolled Students	5.340429e-02
On-time Performance * Job Density	2.349202e-04
Stop Skipping Rate * Population Density	2.933478e-02

R^2

```
0.1381547
```

MSE

```
107.6857
```

Route 8 Outbound Saturday

(Intercept)	-5.139491e+01
On-time Performance	1.043894e+02
Bus Crowding Level	1.467827e+02
Job Density	4.042498e-04
% Household under 80% AMI	9.515482e+01
Unemployment Rate	-1.649707e+02
Vehicle Per Household	-2.470260e+01
Elder Percentage	1.103139e+02
University Enrolled Students	-3.445481e+02
Major Land Use Type: Transportation Utilities	3.736e+00
Waiting Time * % Household Under 80% AMI	-4.082344e-02
Waiting Time * Unemployment Rate	-2.304136e-05
Waiting Time * University Enrolled Students	3.345500e-01
On-time Performance * Population Density	6.027177e-03
On-time Performance * % Household Under 80% AMI	-1.924649e+02

R^2

```
0.6929018
```

MSE

```
64.38306
```

Route 8 Outbound Sunday

(Intercept)	1.784304e+01
Bus Crowding Level	6.225300e+01
Job Density	1.203684e-04
% Household under 80% AMI	-3.137536e+01
Unemployment Rate	-6.764459e+01
Vehicle Per Household	-1.335144e+01
Waiting Time * University Enrolled Students	4.118907e-02
On-time Performance * Population Density	1.610480e-03

R^2

```
0.4575572
```

MSE

```
51.12852
```

Route 12 Inbound Weekday Peak

AIC

(Intercept)	7.986142e+00
On-time Performance	-7.644858e+00
Bus Crowding Level	-7.363417e+01
Stop Skipping Rate	-1.306706e+02
Job Density	-7.923929e-02
% Household under 80% AMI	-3.425389e+00
Elder Percentage	-3.464196e+01
University Enrolled Students	1.278978e+02
Major Land Use Type: Residential	7.081142e+00
Major Land Use Type: Transportation Utilities	3.247e+01
Waiting Time * Job Density	1.453934e-05
Waiting Time * Unemployment Rate	1.675905e-02
Waiting Time * Vehicle Ownership	4.021206e-03
Waiting Time * University Enrolled Students	-3.546625e-02
On-time Performance * Job Density	5.357417e-04

R^2

0.8013854

MSE

2.177457

BIC

(Intercept)	1.932653104
Bus Crowding Level	-42.368673140
Major Land Use Type: Transportation Utilities	12.54500
Waiting Time * Unemployment Rate	0.017648662
Waiting Time * Vehicle Ownership	0.001636115
On-time Performance * Job Density	-0.025626427
Crowding Level * Job Density	0.122629929

R^2

0.674708

MSE

1.997354

Route 12 Inbound Weekday Off-Peak

(Intercept)	8.811441e+02
On-time Performance	-1.109624e+03
Population Density	2.681463e-01
Job Density	-5.412856e-01
% Household under 80% AMI	-3.351921e+03
Vehicle Per Household	-4.966732e+01
University Enrolled Students	2.723379e+02
Major Land Use Type: Residential	2.445137e+01
Waiting Time * Population Density	-6.804376e-05
Waiting Time * % Household Under 80% AMI	8.701222e-01
Waiting Time * University Enrolled Students	-2.372106e+00
On-time Performance * Job Density	6.916660e-01
On-time Performance * University Enrolled Students	1.190126e+04
Crowding Level * Job Density	7.952358e-01

R^2

0.7845079

MSE

68.5642

Route 12 Inbound Saturday

(Intercept)	4.425990e+01
Waiting Time	-1.131771e-02
On-time Performance	-1.822319e+01
Bus Crowding Level	-6.858369e+01
Stop Skipping Rate	6.679924e+00
Population Density	8.621623e-04
Job Density	-9.028497e-02
% Household under 80% AMI	-1.230146e+02
Unemployment Rate	1.598725e+02
Vehicle Per Household	1.575462e+01
Elder Percentage	-4.341745e+00
University Enrolled Students	-5.240958e+01
Major Land Use Type: Residential	7.392013e+00
Major Land Use Type: Transportation Utilities	2.943915e+01
Waiting Time * Job Density	4.270504e-06
Waiting Time * % Household Under 80% AMI	3.986253e-02
On-time Performance * Job Density	7.693389e-02
On-time Performance * University Enrolled Students	1.272745e+02
Crowding Level * Job Density	1.522265e-01
Stop Skipping Rate * % Household Under 80% AMI	-9.955293e+00
Stop Skipping Rate * University Enrolled Students	1.199069e+00

R^2

```
0.7199215
```

MSE

```
7.395775
```

Route 12 Inbound Sunday

(Intercept)	1.177529e+01
Job Density	2.553618e-02
% Household under 80% AMI	-4.031292e+01
Unemployment Rate	8.686153e+01
Elder Percentage	-3.608901e+01
University Enrolled Students	5.743893e+01
Waiting Time * Job Density	1.369097e-05
On-time Performance * Population Density	3.261137e-03
On-time Performance * Unemployment Rate	5.560002e+01
Stop Skipping Rate * Job Density	-5.427933e-02

R^2

```
0.6247414
```

MSE

```
74.58886
```

Route 12 Outbound Weekday Peak

(Intercept)	9.765329e+00
Population Density	2.809371e-04
Job Density	4.873629e-02
% Household under 80% AMI	-2.715068e+01
Elder Percentage	-2.499332e+01
Waiting Time * Job Density	2.918643e-06
Waiting Time * Unemployment Rate	3.788678e-02
On-time Performance * University Enrolled Students	8.184527e+01
Stop Skipping Rate * Job Density	-1.686892e-01

R^2

0.7106654

MSE

67.768

Route 12 Outbound Weekday Off-Peak

(Intercept)	2.721287344
Job Density	0.015809431
% Household under 80% AMI	-12.967889747
Vehicle Per Household	1.294046524
Elder Percentage	-9.931966345
University Enrolled Students	29.463507394
On-time Performance * Population Density	0.001744832
On-time Performance * % Household Under 80% AMI	-9.998629831
On-time Performance * Unemployment Rate	96.864275415

R^2

0.3266049

MSE

23.76495

Route 12 Outbound Saturday

(Intercept)	-12.700790140
On-time Performance	19.072771650
Job Density	0.065047389
% Household under 80% AMI	-4.666223125
Elder Percentage	-24.911644698
Waiting Time * University Enrolled Students	0.016156273
On-time Performance * Population Density	0.004684839

R^2

```
0.3995403
```

MSE

```
108.6355
```

Route 12 Outbound Sunday

(Intercept)	1.230119e+01
Job Density	2.674813e-02
% Household under 80% AMI	-4.331142e+01
Unemployment Rate	1.088935e+02
Elder Percentage	-3.576324e+01
University Enrolled Students	6.013136e+01
Waiting Time * Job Density	1.315172e-05
On-time Performance * Population Density	3.413005e-03
On-time Performance * Unemployment Rate	3.680784e+01
Stop Skipping Rate * Job Density	-5.653232e-02

R^2

```
0.5687569
```

MSE

```
74.73955
```

Route 16 Inbound Weekday Peak

(Intercept)	-5.660171e+01
On-time Performance	1.969340e+01
Bus Crowding Level	-4.046566e+01
Population Density	-2.331870e-04
% Household under 80% AMI	1.364304e+02
Elder Percentage	-9.992333e+01
University Enrolled Students	-3.015313e+02
Major Land Use Type: Rivers Streams Canals	1.355696e+01
Stop Skipping Rate	-2.105711e+00
Waiting Time * Job Density	-2.551737e-06
Waiting Time * % Household Under 80% AMI	5.325685e-04
On-time Performance * Job Density	2.489516e-03
On-time Performance * Unemployment Rate	-8.609088e+02
On-time Performance * Vehicle Ownership	7.969626e+01
Crowding Level * % Household Under 80% AMI	1.013984e+02
Stop Skipping Rate * Job Density	2.193129e-02

R^2

0.5844298

MSE

22.82339

Route 16 Inbound Weekday Off-Peak

(Intercept)	-2.299120e+02
Waiting Time	3.467594e-03
Bus Crowding Level	-2.996828e+01
Population Density	-3.805060e-03
Job Density	2.804581e-02
% Household under 80% AMI	2.914859e+02
Vehicle Per Household	2.799790e+00
University Enrolled Students	-1.827360e+03
Major Land Use Type: Rivers Streams Canals	9.911411e+01
Major Land Use Type: Transportation Utilities	-7.985e+00
Stop Skipping Rate	5.421755e+01
On-time Performance * % Household Under 80% AMI	4.608283e+02
On-time Performance * Unemployment Rate	-2.587654e+03
On-time Performance * Vehicle Ownership	7.418190e+01
Crowding Level * University Enrolled Students	3.750805e+03

R^2

0.5107866

MSE

43.6444

Route 16 Inbound Saturday

(Intercept)	-2.999686e+01
On-time Performance	1.264303e+02
Bus Crowding Level	-7.168375e+02
Population Density	-4.200723e-03
Job Density	4.721594e-02
% Household under 80% AMI	1.962960e+02
Unemployment Rate	-1.969044e+03
Vehicle Per Household	1.401745e+01
University Enrolled Students	-1.122055e+03
Major Land Use Type: Rivers Streams Canals	8.613750e+01
Major Land Use Type: Transportation Utilities	-3.78e+01
Waiting Time * % Household Under 80% AMI	-1.177744e-02
Crowding Level * % Household Under 80% AMI	1.195478e+03
Crowding Level * Elder Percentage	8.394386e+02

R^2

```
0.4667591
```

MSE

```
37.16859
```

Route 16 Inbound Sunday

(Intercept)	-3.324616e+01
Waiting Time	2.854759e-02
On-time Performance	1.396558e+02
Bus Crowding Level	-8.741904e+02
Population Density	-4.667272e-03
Job Density	5.547613e-02
% Household under 80% AMI	2.140897e+02
Unemployment Rate	-2.061549e+03
Vehicle Per Household	7.685367e+00
University Enrolled Students	-1.219636e+03
Major Land Use Type: Rivers Streams Canals	8.166493e+01
Major Land Use Type: Transportation Utilities	-4.18e+01
Waiting Time * % Household Under 80% AMI	-5.876264e-02
Crowding Level * % Household Under 80% AMI	1.403702e+03
Crowding Level * Elder Percentage	1.016317e+03

R^2

```
0.5078125
```

MSE

```
32.8588
```

Route 16 Outbound Weekday Peak

(Intercept)	-2.835507
On-time Performance	17.941782
Major Land Use Type: Transportation Utilities	13.212079
On-time Performance * % Household Under 80% AMI	15.771918

R^2

```
-0.03351879
```

MSE

```
102.6722
```

Route 16 Outbound Weekday Off-Peak

(Intercept)	-25.73108
On-time Performance	45.98901
University Enrolled Students	44.23493
Major Land Use Type: Transportation Utilities	27.08888
On-time Performance * % Household Under 80% AMI	18.88495

R^2

```
0.3063842
```

MSE

```
135.5204
```

Route 16 Outbound Saturday

(Intercept)	-26.85781
On-time Performance	50.11175
University Enrolled Students	41.22386
Major Land Use Type: Transportation Utilities	27.16647
On-time Performance * % Household Under 80% AMI	14.94120

R^2

```
-0.5442323
```

MSE

```
134.0025
```

Route 16 Outbound Sunday

(Intercept)	-3.990257e+01
Waiting Time	1.962742e-02
Population Density	4.457468e-04
University Enrolled Students	1.040589e+02
Major Land Use Type: Transportation Utilities	2.062e+01
Waiting Time * % Household Under 80% AMI	9.662737e-03
Crowding Level * % Household Under 80% AMI	8.716764e+00
Crowding Level * Unemployment Rate	1.400279e+01

R^2

```
0.3659108
```

MSE

```
113.8024
```

Route 51 Inbound Weekday Peak

(Intercept)	-5.880401641
Waiting Time	0.001104511
On-time Performance	6.844920100
Bus Crowding Level	-3.907295295
Population Density	0.002086437
Job Density	-0.004505540
Unemployment Rate	-10.419550022
Vehicle Per Household	-6.446395503
Elder Percentage	31.170749982
Major Land Use Type: Transportation Utilities	6.197328
Stop Skipping Rate	35.882715617
On-time Performance * Vehicle Ownership	-3.547867594

R^2

```
0.6731369
```

MSE

```
8.614487
```

Route 51 Inbound Weekday Off-Peak

(Intercept)	9.254952339
Waiting Time	0.002243363
Population Density	0.005056550
Job Density	-0.020771701
Vehicle Per Household	-47.324632401
Elder Percentage	185.456853775
Major Land Use Type: Transportation Utilities	35.222280619
Stop Skipping Rate	18.345782386

R^2

```
0.6017767
```

MSE

```
65.96008
```

Route 51 Inbound Saturday

(Intercept)	2.010065e+01
Waiting Time	-2.304589e-03
On-time Performance	1.073686e+01
Population Density	4.405254e-03
Job Density	-3.393613e-03
% Household under 80% AMI	3.500913e+00
Vehicle Per Household	-3.060743e+01
Elder Percentage	1.056346e+02
Major Land Use Type: Transportation Utilities	5.395172e+00
Waiting Time * Job Density	-3.877705e-06
On-time Performance * University Enrolled Students	-2.795257e+00
Stop Skipping Rate * Population Density	-5.151685e-04
Stop Skipping Rate * % Household Under 80% AMI	1.747435e+01

R^2

0.5184492

MSE

33.00957

Route 51 Inbound Sunday

(Intercept)	4.465405e+01
Bus Crowding Level	3.994923e+01
Elder Percentage	-1.418810e+02
Waiting Time * University Enrolled Students	6.850513e-03
On-time Performance * % Household Under 80% AMI	-4.345451e+01
On-time Performance * Unemployment Rate	-4.868617e+01
Stop Skipping Rate * % Household Under 80% AMI	-1.500399e+00

R^2

0.7750965

MSE

20.60196

Route 51 Outbound Weekday Peak

(Intercept)	17.430572889
Bus Crowding Level	13.203264180
% Household under 80% AMI	-7.518704812
Unemployment Rate	-8.490953445
Elder Percentage	-65.332471674
University Enrolled Students	3.462841429
Crowding Level * Job Density	0.006730845

R^2

0.8720505

MSE

18.84925

Route 51 Outbound Weekday Off-Peak

(Intercept)	9.541112e+01
Waiting Time	-1.139250e-02
Bus Crowding Level	8.457398e+01
Population Density	5.027845e-05
Job Density	6.595488e-03
% Household under 80% AMI	-3.458191e+01
Unemployment Rate	-7.681234e+01
Elder Percentage	-2.806613e+02
On-time Performance * Unemployment Rate	-5.126918e+00
Stop Skipping Rate * Population Density	-4.277298e-04
Stop Skipping Rate * % Household Under 80% AMI	-1.098373e+01

R^2

```
0.744551
```

MSE

```
71.80885
```

Route 51 Outbound Saturday

(Intercept)	8.227044e+01
Bus Crowding Level	4.501389e+01
% Household under 80% AMI	-3.181370e+01
Unemployment Rate	-3.006584e+01
Elder Percentage	-2.602395e+02
Waiting Time * Job Density	2.594889e-07
Waiting Time * Unemployment Rate	-2.204458e-02
On-time Performance * Population Density	5.654884e-04
On-time Performance * % Household Under 80% AMI	-2.581399e+01
Crowding Level * Job Density	6.641712e-03
Stop Skipping Rate * Job Density	6.446719e-03
Stop Skipping Rate * % Household Under 80% AMI	-2.506779e+00

R^2

```
0.7747412
```

MSE

```
30.32274
```

Route 51 Outbound Sunday

(Intercept)	5.544732e+01
Waiting Time	6.096628e-03
Bus Crowding Level	4.007914e+01
Population Density	2.338512e-04
% Household under 80% AMI	-4.222014e+01
Unemployment Rate	-6.834339e+01
Elder Percentage	-1.967962e+02
Stop Skipping Rate	-3.982870e+00
Crowding Level * Job Density	1.044101e-03

R^2

```
0.8288178
```

MSE

```
14.42932
```

Route 91 Inbound Weekday Peak

(Intercept)	3.502810e+01
On-time Performance	6.815915e+00
Population Density	-4.392181e-03
Job Density	-1.278806e-02
% Household under 80% AMI	-1.552796e+01
Unemployment Rate	1.603282e+02
Elder Percentage	-1.052719e+02
University Enrolled Students	4.212462e+02
Major Land Use Type: Rivers Streams Canals	2.390077e+01
Major Land Use Type: Transportation Utilities	7.172e+00
Waiting Time * % Household Under 80% AMI	-4.406653e-03
Crowding Level * University Enrolled Students	1.081652e+03

R^2

```
0.6684251
```

MSE

```
9.603251
```

Route 91 Inbound Weekday Off-Peak

(Intercept)	-9.744464e+00
On-time Performance	1.151884e+01
Population Density	-2.284958e-03
Job Density	-1.189995e-03
Unemployment Rate	2.585543e+02
Vehicle Per Household	-1.534959e
Elder Percentage	-9.547238e+01
University Enrolled Students	8.494098e+02
Major Land Use Type: Rivers Streams Canals	5.030966e+00
Major Land Use Type: Transportation Utilities	7.4548e+00
Stop Skipping Rate * University Enrolled Students	3.112817e+0 3

R^2

```
0.8363844
```

MSE

15.36771

Route 91 Inbound Saturday

(Intercept)	3.587843e+01
Waiting Time	-1.882434e-02
Population Density	-4.184395e-04
Job Density	-2.031343e-02
% Household under 80% AMI	-4.266374e+01
Unemployment Rate	3.723986e+02
Vehicle Per Household	-5.001225e+01
Elder Percentage	-7.330989e+01
University Enrolled Students	1.461071e+03
Major Land Use Type: Residential	1.798471e+01
Major Land Use Type: Transportation Utilities	1.858e+01
On-time Performance * University Enrolled Students	6.771213e+02
Crowding Level * University Enrolled Students	-3.831770e+03
Stop Skipping Rate * Population Density	1.159121e-03

R^2

0.9014989

MSE

50.1621

Route 91 Inbound Sunday

(Intercept)	12.57886160
Job Density	0.01043088
On-time Performance * Vehicle Ownership	-16.49860793
On-time Performance * University Enrolled Students	33.70377188
Stop Skipping Rate * University Enrolled Students	25.76242086

R^2

0.5276905

MSE

```
121.6257
```

Route 91 Outbound Weekday Peak

```
(Intercept)                                          6.902380646
Vehicle Per Household                               -4.571416597
University Enrolled Students                        23.160181880
Stop Skipping Rate * University Enrolled Students  403.420663966
```

R^2

```
0.7226108
```

MSE

```
4.085403
```

Route 91 Outbound Weekday Off-Peak

```
(Intercept)                                       8.990163e+00
Population Density                                7.834502e-04
 % Household under 80% AMI                        4.532931e+00
Vehicle Per Household                            -4.255233e+01
Elder Percentage                                  5.984379e+01
University Enrolled Students                      1.863865e+02
Major Land Use Type: Residential                  9.158462e-01
Major Land Use Type: Transportation Utilities     1.289608e+00
Stop Skipping Rate * % Household Under 80% AMI    5.061116e+02
```

R^2

```
0.7190686
```

MSE

```
9.131488
```

Route 91 Outbound Saturday

(Intercept)	-4.569149e+00
Crowding Level	2.627631e+01
Population Density	1.222725e-03
Job Density	3.291286e-04
Vehicle Per Household	-3.957718e+01
Elder Percentage	6.486651e+01
University Enrolled Students	1.159269e+02
Major Land Use Type: Residential	7.126404e+00
Major Land Use Type: Transportation Utilities	1.024072e+01
On-time Performance * % Household Under 80% AMI	1.114287e+00
Stop Skipping Rate * Job Density	3.273375e-02
Stop Skipping Rate * Vehicle Ownership	-1.460909e+01

R^2

```
0.5530474
```

MSE

```
4.901497
```

Route 91 Outbound Sunday

(Intercept)	9.261723e+00
Population Density	5.243457e-04
% Household under 80% AMI	6.205632e+00
Vehicle Per Household	-2.664091e+01
Elder Percentage	2.972514e+01
University Enrolled Students	7.728664e+01
On-time Performance * % Household Under 80% AMI	3.665342e-01
Stop Skipping Rate * Job Density	7.840316e-03
Stop Skipping Rate * Elder Percentage	3.692881e+01

R^2

```
0.7732426
```

MSE

```
12.62798
```