The Effect of Gasoline Prices on Public Transit Ridership in New York City

Dominika Markowska-Desvallons

Orli Khaimova

CUNY School of Professional Studies

DATA 698

Nasrin Khansari

Abstract

Low supply and high demand for crude oil drove prices upwards in 2022. Additionally, the COVID- 19 pandemic made the MTA ridership go down to unseen numbers. Combining those two factors, we conducted research on what effect gasoline prices had on public transportation usage in NYC. By examining the data on modes of transportation: subway, buses, trains, and bicycles and the data on gas prices, different linear regression models were made to observe any relationships. The goal of the study was to find a correlation between gas prices and the usage of alternative modes of transportation in NYC. The models showed that commuter rail systems, the LIRR and Metro-North Railroad, had a higher correlation with gasoline prices compared to the others.

Introduction

Gasoline prices were on the rise during the pandemic, but fell down recently. The surge in gas prices was, in part, affected by the increase of oil prices. The supply of crude oil, which is used to make gasoline, has been affected since companies lowered oil production due to Covid-19. There was also a shutdown of a major pipeline in 2021. Then, in 2022, the Russian-Ukrainian conflict acted as another catalyst to push the prices higher, since Russia produces about 10% of the oil supply worldwide. Still, the economy sees elevated gas prices.

In our research project, we would like to find out how the changes in gas prices affected other modes of transportation in New York City from March 2020 to December 2022. We know that the MTA daily ridership declined due to the pandemic but it started increasing when NYC started reopening. We are interested in seeing if the gas prices had any effect on MTA daily ridership. We may also look at other modes of transportation, if data is available in NYC. This may include Citi Bike usage and sales of bicycles and electric scooters. If data is not available for NYC, we can extend our research to see if there is any correlation between gas price changes and the sales of electric cars, scooters, bicycles, and other modes of transportation.

We are interested to see what kind of effects the increase of gas prices had on consumers and if there are any new trends forming as a result. We wonder if consumers are using available travel alternatives when gasoline prices change. It would also be interesting to see if this is giving people a motive to become more sustainable and energy efficient. This can be an important subject to see if the demand of gasoline shifted, which would also affect the demand of crude oil. Since we are using oil reserves currently, it can help to forecast the future demand. It can also be important so that MTA, as well as other transportation agencies, can prepare for any changes when gasoline prices are affected. It may be of interest to see how a decrease in gas price affects consumer usage of public transportation, if there is any.

There has been other research done to see the relationship between transit ridership and gasoline prices. Mineta Transportation Institute looked at the ridership of ten different urban areas and what kind of short and long term effects followed the changes in gas price. Another research factored in how ridership apps, such as Uber, affected the elasticity of public transportation. They also looked at how gas tax changes had a significant impact on transit ridership. The other researchers used panel data analysis methods and time series analysis. They also tried to estimate the ridership in certain cities based on the gas price and other factors, such as weather and unemployment rate.

We believe that when gasoline prices increase, the usage of public transportation and other travel alternatives increase. We also believe that there is a correlation between gas prices and public transportation in NYC. Additionally, it seems that people are utilizing other alternatives such as bicycles and scooters because they are more financially attractive.

Our research will help NYC and other urban areas prepare for any changes in public transportation whenever there are changes in gasoline prices. They may also make other travel alternatives more accessible and available to others, while maximizing safety. This would also promote sustainability and help fight the climate crisis.

Keywords: *Bicycle Count, Elasticity, Fuel, Gasoline Prices, Mode of Transportation, MTA, Multiple Regression Model, Public Transportation, Regression Model, Ridership, Time-Series Analysis, Travel*

Literature Review

It is important for New York City policy makers to understand the effect that fuel prices have on transit ridership, bike sharing, and other modes of transportation. There are many factors that may influence a commuter's choice for their mode of transportation, in addition to fuel price. Some factors that were researched include travel times, the availability of alternative modes of transportation or routes, ride sharing, trip frequency, and congestion. Researchers also studied how significant price increases in fuel price affected ridership.

The relationship between fuel prices and transit ridership is not the same in every urban city. Some commuters may not consider public transit as an option due to it being inaccessible or it being constrained in a local area. Some cities are also affected by the popularity of ride-sharing applications and availability of electric vehicles. (Gershon, 2005)

Researchers have measured the cross-price elasticity of transit ridership, in respect to gasoline price changes. It was found that there was an elasticity of 0.12, meaning that for every 10% increase in fuel price, there was a 1.2% increase in the overall U.S. demand for transit (Currie & Phung, 2007). Another research showed an elasticity of 0.025, among 10 major urban areas in America between 2006 and 2018, which varied significantly between cities (Graham, 2020). A third research shows that there was a 2.4% increase in transit demand for every 10% increase in gasoline price (Haire & Machemehl, 2007). In Washington state, an elasticity of 0.17 was found and that there was an inverse relationship between transit ridership and fuel prices (Stover & Bae, 2011).

Sensitivity to fuel prices varied between different modes of transportation. This can be due to commuter travel time and the degree of accessibility. Carrie & Phung found that heavy rail was more susceptible to price changes in gasoline, compared to light rail and bus ridership in the United States (2007). During gasoline price surges, overall, transit ridership increased as found in various studies. However, some relationships were not statistically significant, such as the Atlanta bus and heavy rail transit system, San Francisco bus transit system, and the Virginia Railway Express commuter rail (Haire & Machemehl, 2007). In general, bus systems in U.S. cities did not see as big of an increase in transit ridership, compared to other modes of transportation which can be attributed to longer travel times. When gasoline prices went over a threshold of $4, light rail saw an increase of 9.34% and heavy rail saw a significantly higher rate of ridership (Iseki & Ali, 2014).

Travel time is a major factor for many commuters when they are deciding their method and route for transportation. Commuters look to minimize their travel time, fuel consumption, and travel expenses. When considering their fuel consumption, they also factor in the route, road congestion, trip frequency, and the distance. In Austin, Texas, those who valued their travel time less showcased the highest elasticities. They were more likely to switch to other modes of transportation, given an increase in fuel price. As the value of time increased, it was found that commuters would still prefer to drive alone and keep their travel time to a minimum. However, it was also found that road congestion increased in some areas due to alternate routes chosen (Levin et al., 2017).

Ride-sharing applications were taken into consideration when measuring the elasticity between fuel price and transit ridership. However, it was found to be inconclusive. Additionally, one researcher tried to determine if there was a difference between increase in fuel tax and fuel price and its effect on transit ridership. Although the results show insignificant impacts, the 2008 Minnesota gasoline tax increase had a significant effect on ridership (Graham, 2020).

Furthermore, the inverse relationship between fuel price and transit ridership may not be consistent throughout time. One outlier that is significant to the research is the start of the COVID-19 pandemic which saw a decrease in gas price and a decline in transit ridership. All nonessential businesses paused operations, causing employees to stay home and subway ridership to decrease by 90% (Halvorsen et al., 2021). Slowly, the numbers increased as the city reopened in phases.

Methodology

Data Sources

For our research that is based in New York City, we use data to collect the history of gasoline prices, bicycle counts throughout the city, and MTA ridership data.

There is a dataset that we are using for MTA ridership which is found from New York's Open Data Portal. The dataset describes the daily MTA ridership beginning in March 2020. It includes the above modes of transportation, as well as Access-A-Ride trips, and total Bridges and Tunnels traffic. The dataset also has percentages to compare ridership to pre-pandemic days.

We are also using a New York City gasoline prices dataset which is released by the U.S. Energy Information Administration. They publish data on a weekly, monthly, and annual basis and include all grades and formulations of gasoline.

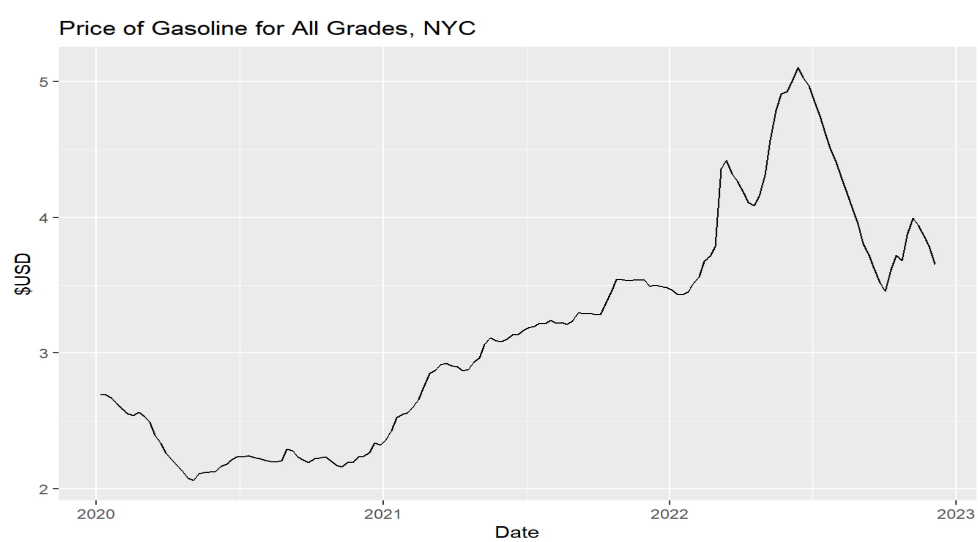
Another dataset that we are using is bicycle counts which is released by NYC Open Data. They conduct bicycle counts around the city at various key locations everyday. The counters update every 15 minutes.

For our analysis, we will conduct multiple time series regressions to see how gasoline prices affect ridership in New York City. The predictor variable will be the fuel price, while the dependent variables will be the ridership for each mode of transportation. The predictor variables will also include dummy variables that account for the trend and seasonality in the data. They can also mitigate any effects caused by outliers. The models will show the different relationships between gasoline prices and ridership in NYC which would then be used to compare and contrast.

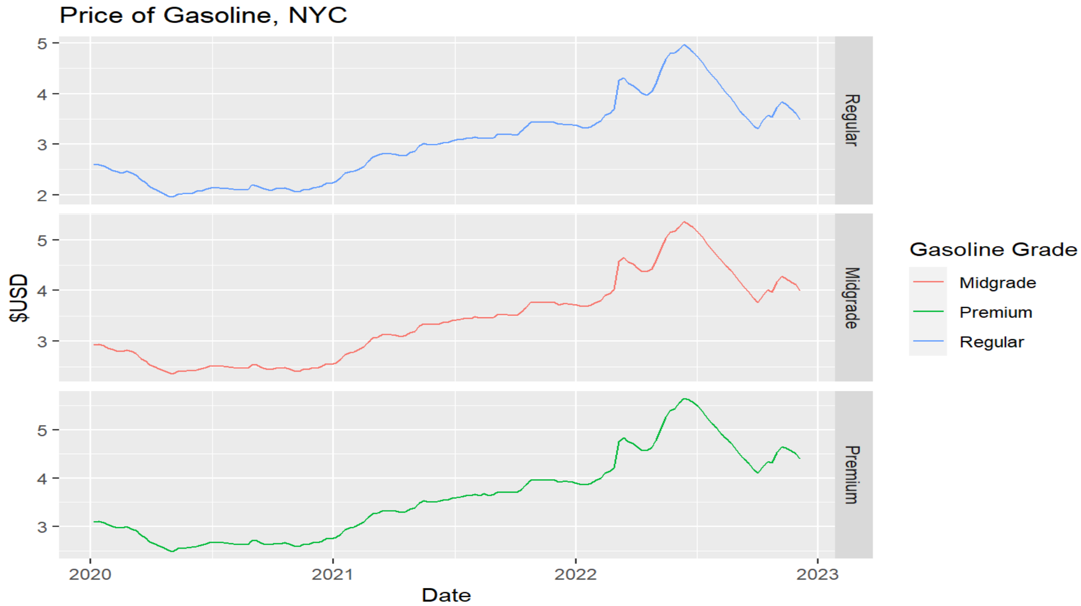
Summary Statistics

The price of gasoline was decreasing in the beginning of 2020, up until late February. Gasoline prices started increasing as the World Health Organization (WHO) declared the novel Coronavirus (COVID-19) outbreak as a pandemic on March 11, 2020. It continued increasing, reaching the ultimate peak in June 2022.

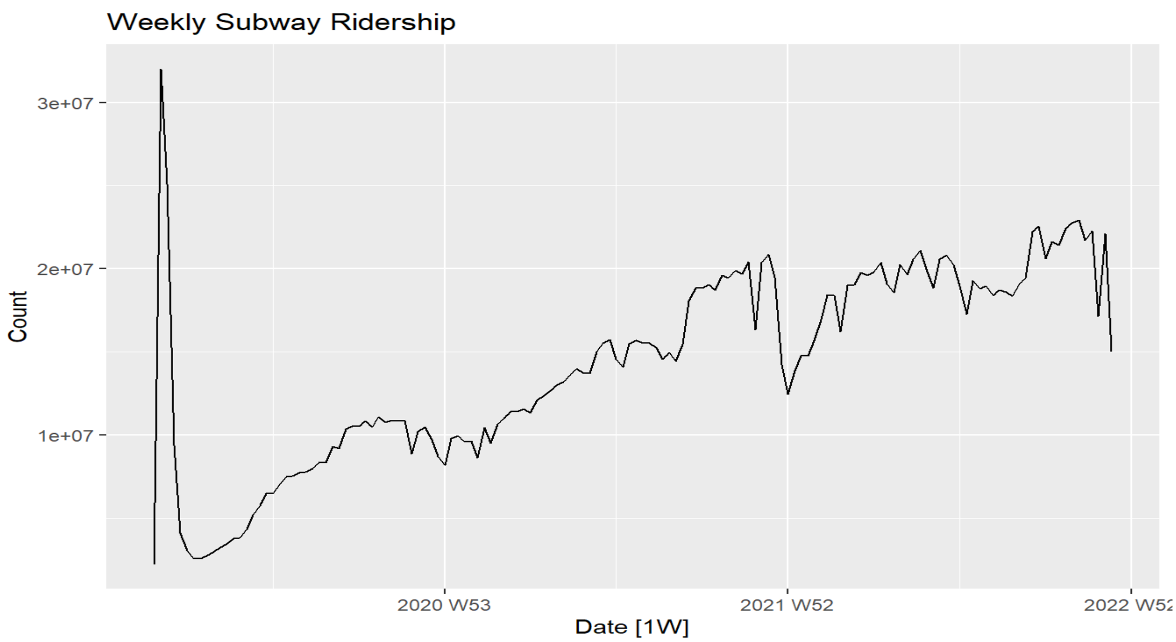
Some factors that affected the gasoline prices throughout the last few years were COVID-19 related supply disruptions, the Russian-Ukrainian conflict, and the Keystone XL pipeline cancellation.



The pattern in gasoline prices are similar across the different grades of gasoline.



We also explored the changes and patterns in ridership throughout the different modes of transportation in NYC. There was a significant decrease in March 2020 which is attributed to the start of the pandemic. There was a temporary decrease that was notable in December 2021, when there was a surge in COVID-19 cases due to the Omicron variant.



From March 2020 to December 2022, there is an increasing trend in ridership across all modes of transportation, except for bicycles. There is a seasonality component in the ridership, on a weekly and yearly basis. Ridership tends to decline during the weekends and increase during the weekdays which can be attributed to commuting to work and school.

Ridership increases around September each year, when the school year starts in NYC. There is also a slight decline every winter. There is a distinguished pattern in bicycle ridership as it decreases every winter and spikes during the summers.

Appendix A.7 shows the comparison of the price of gasoline versus each mode of transportation. It can be seen that for MTA subway, LIRR and Metro-North Railroad, that there is an increase in ridership as the price of gasoline increases.

Baseline Linear Regression with Mode of Transportation

A linear regression model observes the linear relationship between a single predictor variable and a dependent variable:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

where y_t is the ridership level of a mode of transportation at week *t*, x_t is the price of gasoline at week *t*, and coefficient β_1 is the rate of change in ridership resulting from a one unit increase in the price of gasoline. β_0 is the intercept and represents the ridership when $x = 0$. ε_t is the error term.

Linear-Log Regression with Mode of Transportation and Logarithmic Transformation of Gas

A logarithmic transformation is applied to the price of gasoline which helps to transform a skewed data into less skewed data. It spreads out the concentrations of data in the tails and makes the data more symmetric. It also helps meet the assumption associated with a linear regression model which is homoscedasticity, constant variance in the predictor variable. The model is specified as:

$$y_t = \beta_0 + \beta_1 \cdot log\, x_t + \varepsilon_t$$

In the model, $\beta_1$explained as the elasticity, or a 1% increase in x_t, price of gasoline, increases or decreases y_t, ridership, by $\beta_1/100$ units.

Linear Regression with Dummy Variables

There are other predictors that are advantageous in a time series regression model. A time series can observe a trend, which happens when there is an increase or decrease in the data overall. A predictor representing a trend in the dependent variable, where t = $x_{1,t}$.

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

Time series can also have seasonality, where there are patterns at predetermined intervals. This can be represented in the time series regression model as dummy variables, such as:

$$y_t = \beta_0 + \beta_1 t + \beta_2 x_{i,t} + . . . + \beta_I x_{I,t} + \varepsilon_t,$$

where $i = 2, . . . , I$ is the seasonal index. The first seasonal variable is not included in the model, as it is incorporated in the intercept, when the other dummy variables are equal to zero. The coefficients for each dummy variable can be interpreted as the difference between the first seasonal variable and that variable.

As the seasonal periods increase, it can be complicated to capture them as dummy variables, as it will make the interpretation difficult and tedious. Another approach would be to use Fourier terms as a replacement for the seasonal patterns. Fourier series is a summation of sinusoidal functions that can be used to estimate a function with periods. There is a weekly and yearly seasonality in the data, hence, Fourier series can be used to approximate the yearly seasonality as $m = 52$.

$$y_t = \beta_0 + \beta_1 t + f(t) + \varepsilon_t,$$

$$\text{where } f(t) = \sum_{t=0}^{n} (a_t \cdot cos\,(\frac{2\pi t}{m}) + b_k \cdot sin\,(\frac{2\pi t}{m})).$$

A linear regression that encompasses Fourier terms is also known as a harmonic regression.

Results

Baseline Linear Regression with Mode of Transportation

The table below shows the results of the baseline regression, prior to any transformations. It shows the R^2, which measures the proportion of the variance in the response variable that is explained by the predictor variables, and the adjusted R^2, which adjusts the R^2 for the amount of terms in the model.

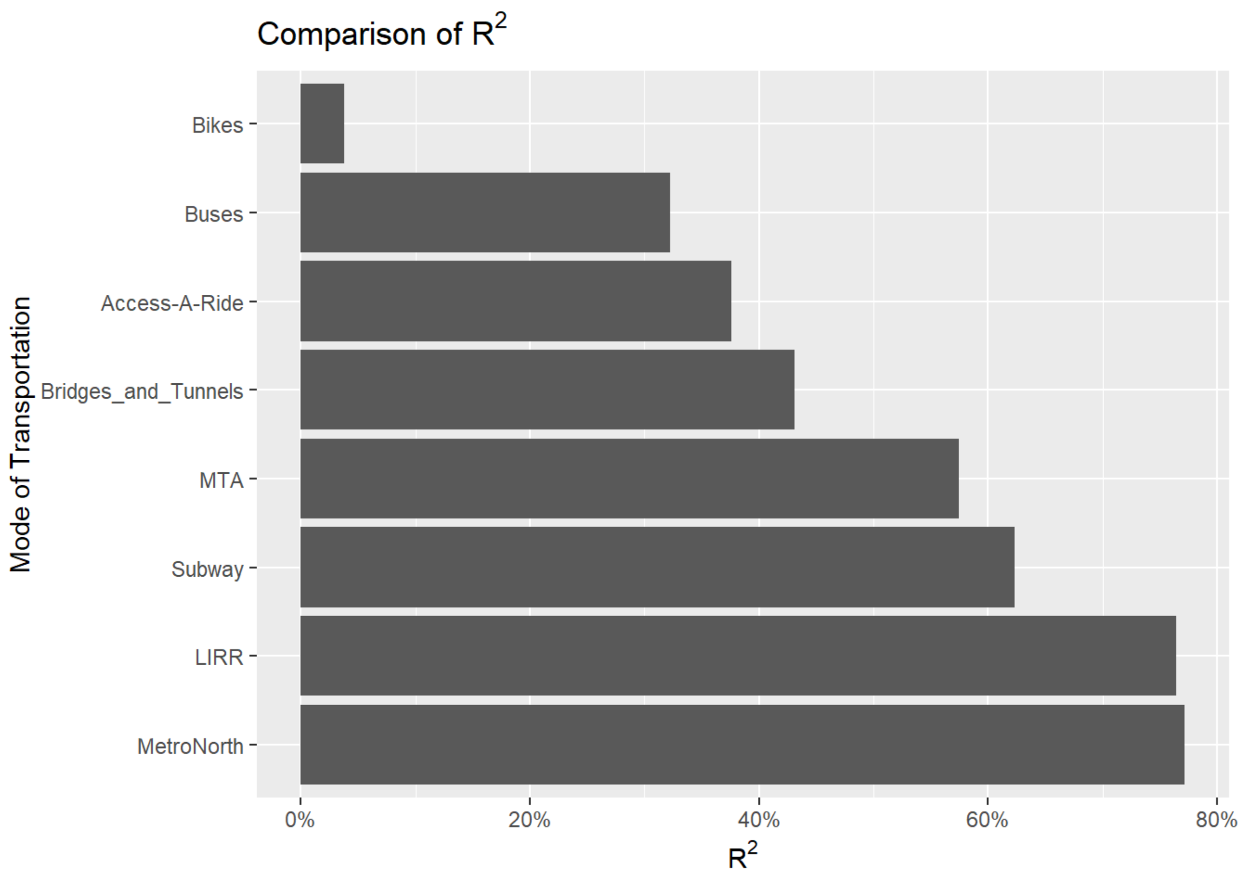
The table shows that the price of gasoline can explain the variation in LIRR and Metro-North Railroad ridership better than the other modes of transportation. For example, 77.2% of the variation in Metro-North ridership can be explained by the price of gasoline.

MTA subway ridership has 62.3% of its variation explained by the changes in the price of gasoline. The variable *MTA* is the ridership for subway and buses combined. Out of the four major MTA measures, bus ridership is the least affected by the price of gasoline.

Bridges_and_Tunnels registers vehicle crossings at seven bridges and two tunnels that are operated by the MTA. Only 43.1% of the variation in vehicle crossings can be explained by the price of gasoline fluctuations.

The least affected mode of transportation seems to be bicycles, with only 3.77% of its variation being explained by the changes in the price of gasoline.

Name	R^2	Adjusted R^2
Subway	0.623	0.620
Buses	0.323	0.318
LIRR	0.765	0.763
MetroNorth	0.772	0.770
Access-A-Ride	0.376	0.372
Bridges_and_Tunnels	0.431	0.427
MTA	0.575	0.572
Bikes	0.0377	0.0310



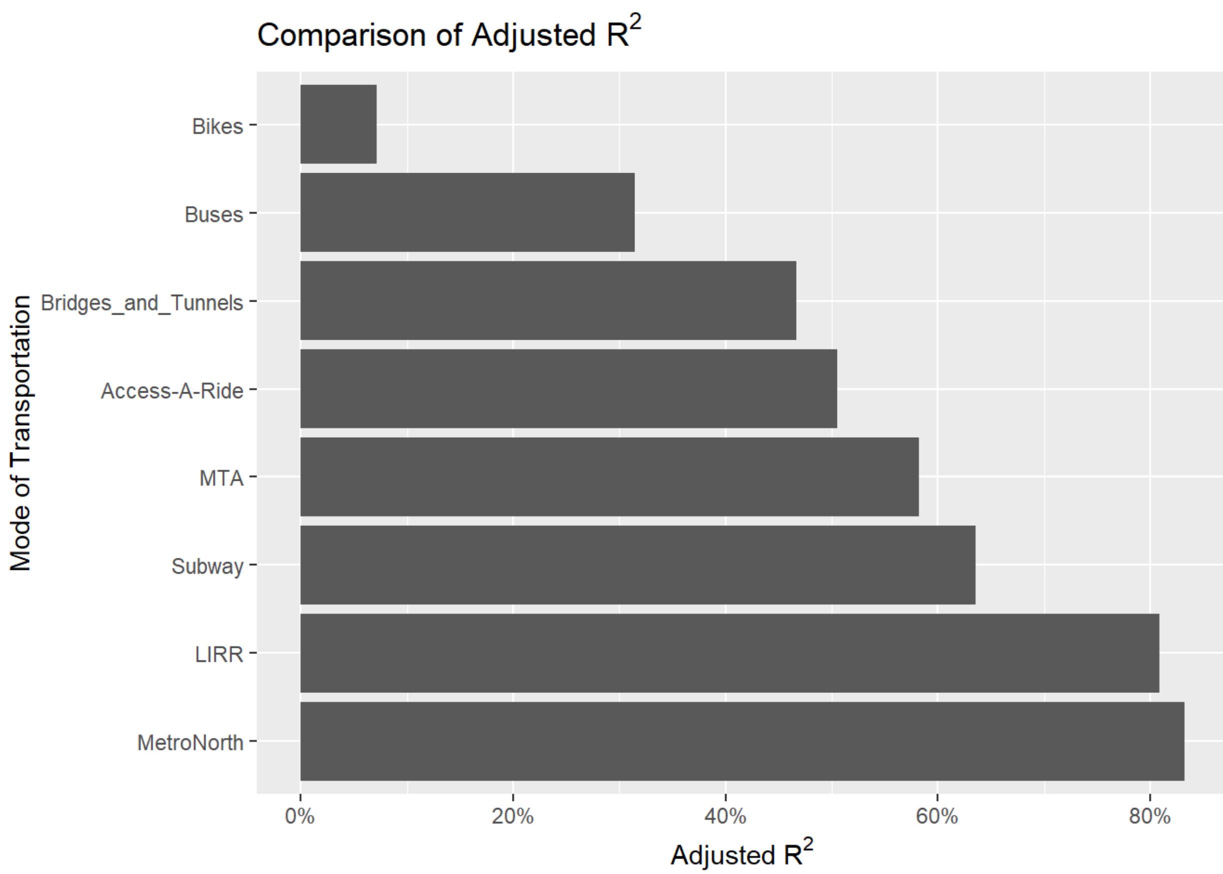
Baseline Linear Regression, Breakdown by Gasoline Grade

We examined how the different grades of gasoline affected the ridership. The adjusted R^2 seemed to improve for most of the modes of transit, except for bus ridership, as seen below in the table.

The appendix shows the results of each linear regression individually. It can be observed that regular and premium grades of gasoline had a positive effect on the ridership, while midgrade had a negative effect. However, the opposite can be said for bicycle ridership, as the price of midgrade gasoline has the only positive coefficient.

The pattern is almost the same, with LIRR and Metro-North Railroad ridership being affected the most by the price of different grades of gasoline, while bicycle ridership is the least affected. However, in the model, Access-A-Ride ridership has more of its variation explained by the subgrades of gasoline compared to Bridges and Tunnels vehicle crossings.

Name	R^2	Adjusted R^2
Subway	0.643	0.635
Buses	0.329	0.315
LIRR	0.812	0.808
MetroNorth	0.836	0.833
Access-A-Ride	0.516	0.505
Bridges_and_Tunnels	0.478	0.467
MTA	0.591	0.582
Bikes	0.908	0.0715

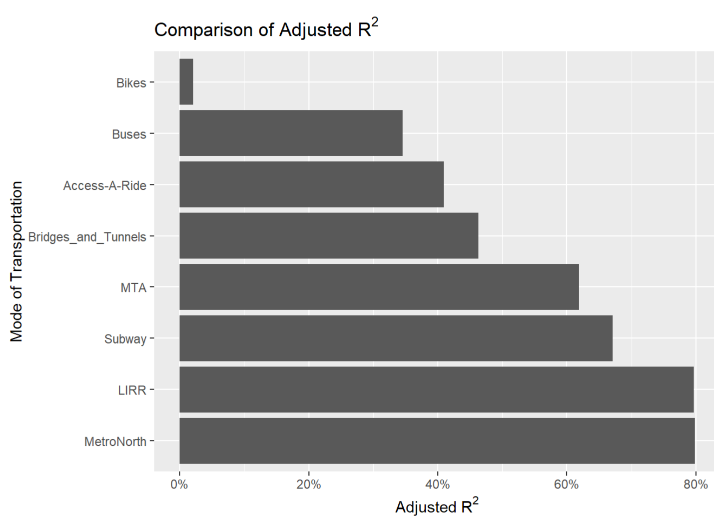


Linear-Log Regression with Mode of Transportation and Logarithmic Transformation of Gas

In the model, the price of gasoline was logarithmically transformed. There was an increase in the R^2, except in bicycle ridership. The log of the price of gasoline has a positive effect on the ridership for all modes of transit as seen in the appendix. For example, for every 1% increase in the price of gasoline, the MTA subway ridership increases by 180,767.35.

The table below shows the results of the linear regression models with a logarithmic transformation of the price of gasoline. The results are in the same order as the linear regression by subgrades of gasoline.

Name	R^2	Adjusted R^2
Subway	0.673	0.671
Buses	0.350	0.345
LIRR	0.798	0.797
MetroNorth	0.800	0.799
Access-A-Ride	0.414	0.409
Bridges_and_Tunnels	0.466	0.463
MTA	0.622	0.619
Bikes	0.0277	0.0209



Linear Regression with Dummy Variables

In these models, trend and seasonality predictors were included to explain the variation of a time series more appropriately. As seen in the table below, the adjusted R^2 seems to significantly improve compared to the previous models.

The table shows that the price of gasoline can explain the variation in LIRR and Metro-North Railroad ridership better than any of the other modes of transit. For example, 96.3% of the variation in Metro-North ridership can be explained by the log of price of gasoline, with respect to the trend and seasonality of the data.

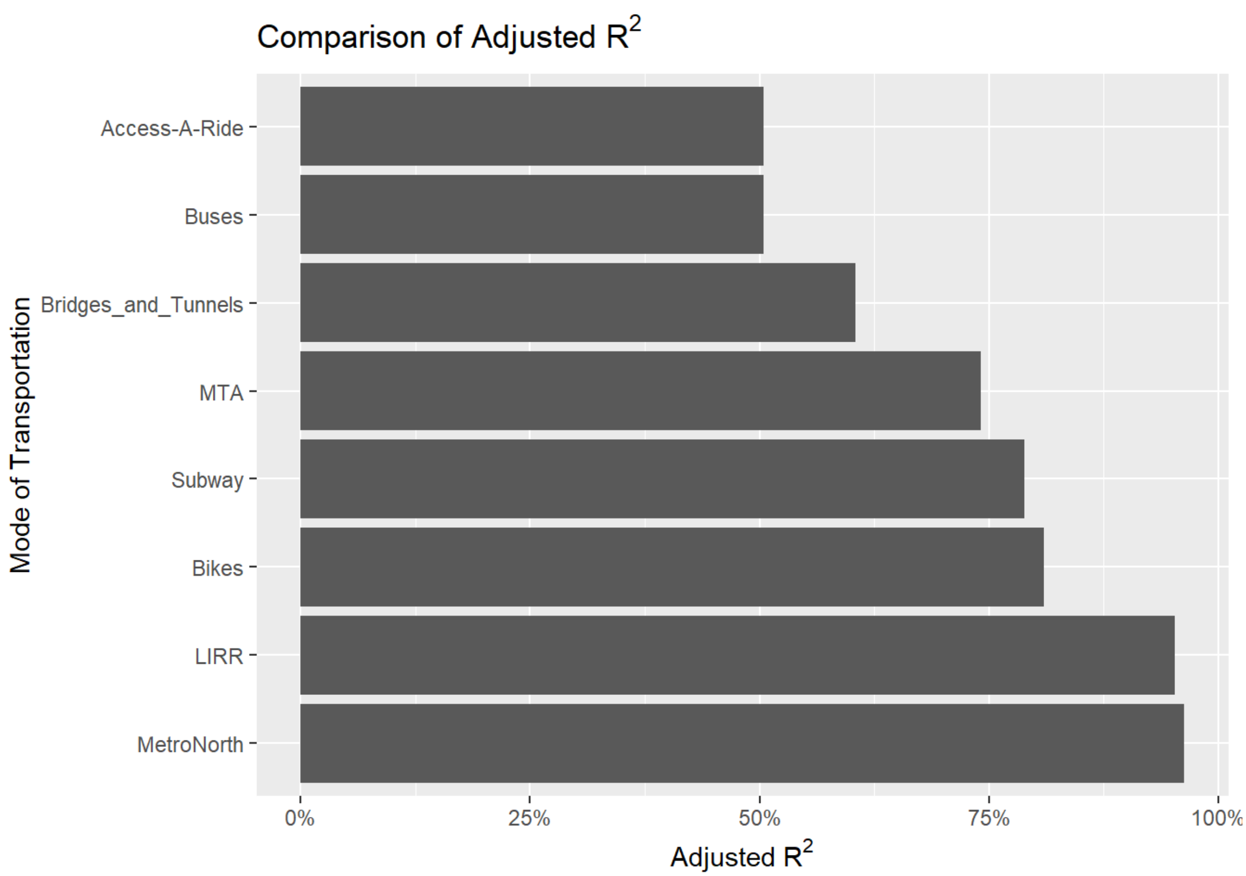
As seen before, the commuter rails, LIRR and Metro-North Railroad, are the most affected by the price of gasoline, even when the trend and seasonality of the data is accounted for.

Similarly, subway ridership is more sensitive to the changes of price of gasoline compared to bus ridership. MTA bus ridership continues to have the lowest sensitivity out of the four major MTA groups.

Unlike the previous models, bicycle ridership has 81.0% of its variation explained by the price of gasoline and seasonality. Once seasonal changes are considered, the bicycle ridership becomes more sensitive to the changes in the price of gasoline.

The least sensitive to the changes in the price of gasoline are observed in Access-A-Ride ridership, when seasonality is included in the model.

Name	R^2	Adjusted R^2
Subway	0.829	0.788
Buses	0.601	0.505
LIRR	0.962	0.952
MetroNorth	0.970	0.963
Access-A-Ride	0.598	0.504
Bridges_and_Tunnels	0.679	0.604
MTA	0.792	0.741
Bikes	0.845	0.810



Discussion

Our analysis of how the price of gasoline affected the alternative modes of transit ridership in NYC from March 2020 to December 2022 shows that there is a significant correlation between the price of gasoline and ridership levels. This supports our initial hypothesis that when gasoline prices increase, the usage of public transportation and other travel alternatives increase. It also supports that bicycle ridership increased when gasoline prices increased.

Commuter rails are the most sensitive to the price changes of gasoline, which is consistent with previous research. Commuters travel longer distances when they use commuter rails compared to light rails and buses. The LIRR and Metro-North Railroad connect NYC to northern suburbs in New York, Connecticut, and Long Island. Commuter rail users are more likely to utilize it when gasoline prices increase because they have a higher gas consumption and it is more financially attractive.

As found in other research, “U.S. light rail is particularly sensitive to gas prices,” with an elasticity between 0.27 to 0.38, and “bus ridership is only slightly sensitive to gas prices” (Currie & Phung, 2007). Graham found that the cross-price elasticity is the lowest for bus ridership (2020). This is consistent with our findings that bus ridership was the least sensitive out of the four major MTA groups.

Bicycle ridership in NYC did not have a significant increasing trend, but a strong seasonal pattern was observed. As shown in another research, bike sharing systems were less likely to see ridership drops outside of the seasonal factors (Teixeira & Lopes, 2020). It should also be noted that an increase in bicycle ridership is not directly caused by the gasoline price increases, as more users switched to bicycle usage to limit their risks of COVID-19.

We expected to see a decrease in vehicles on the road as the price of gasoline increased. However, our models for Bridges and Tunnels vehicle crossings, showed that there is a positive correlation between vehicle crossings and gasoline prices. This can be due to financial reasons, travel distances, among many other reasons.

Other researchers looked into other factors such as unemployment rate, household income, demographics of riders, and value of time. We were unable to find such data provided by the MTA or NYC Open Data. Demographic information may influence the behaviors of riders, in addition, to being affected by the changes in the price of gasoline. A researcher found that bus riders were less wealthier than commuter rail users and it would be interesting to see how that information influenced their ridership behavior (Graham, 2020). While we believe these limitations did not impact the primary results of the study, future studies may consider the demographic information, as well as the distance and time traveled during each trip.

Conclusion

As the price of gasoline fluctuates, it is important to understand how it can affect public transportation and bicycle ridership. An increase in the price of gasoline is found to be correlated with ridership levels. Commuter rail ridership is the most sensitive to the price, followed by subway ridership, and bus ridership is the least sensitive to the price. NYC bicycle users are seasonally affected by the changes in the price of gasoline. Future research in this topic can focus more on other factors that can affect riders, such as demographic information and travel time and distance. It can help NYC planners to prepare for any changes in public transportation and traffic congestion whenever there are changes in gasoline prices.

References

1. Baghestani, A., Tayarani, M., Allahviranloo, M., & Gao, H. O. (2020). Evaluating the traffic and emissions impacts of congestion pricing in New York city. *Sustainability*, *12*(9), 3655.
2. Currie, G., & Phung, J. (2007). Transit Ridership, Auto Gas Prices, and World Events: New Drivers of Change? Transportation Research Record, 1992(1), 3–10. https://doi.org/10.3141/1992-01
3. Dong, Xiaoxia (2019). Bye-Bye Bus, The Downside Of Ride-Hail In Philadelphia, https://kleinmanenergy.upenn.edu/wp-content/uploads/2020/08/Bye-Bye-Bus-1.pdf
4. Gershon, R.R.M. Public transportation: Advantages and challenges. *J Urban Health* 82, 7–9 (2005). https://doi.org/10.1093/jurban/jti003
5. Graham, J. (2020). The Effect of Gas Prices and Gas Taxes on Public Transportation Ridership. https://dataspace.princeton.edu/handle/88435/dsp01nc580q68z
6. Haire, A. R., & Machemehl, R. B. (2007). Impact of Rising Fuel Prices on U.S. Transit Ridership. Transportation Research Record, 1992(1), 11–19. https://doi-org.remote.baruch.cuny.edu/10.3141/1992-02
7. Halvorsen, A., Wood, D., Jefferson, D., Stasko, T., Hui, J., & Reddy, A. (2021). Examination of New York City Transit's Bus and Subway Ridership Trends During the COVID-19 Pandemic. *Transportation Research Record*, 03611981211028860.
8. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. https://otexts.com/fpp3/
9. Iseki, H., & Ali, R. (2014). Net effects of gasoline price changes on transit ridership in US urban areas (No. CA-MTI-14-1106). Mineta Transportation Institute.

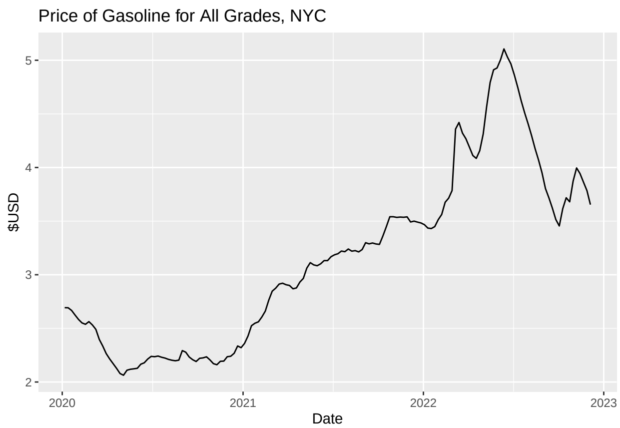
10. Levin, Jafari, E., Shah, R., & Boyles, S. D. (2017). Network-based model for predicting the effect of fuel price on transit ridership and greenhouse gas emissions. International Journal of Transportation Science and Technology, 6(4), 272–286. https://doi.org/10.1016/j.ijtst.2017.07.003

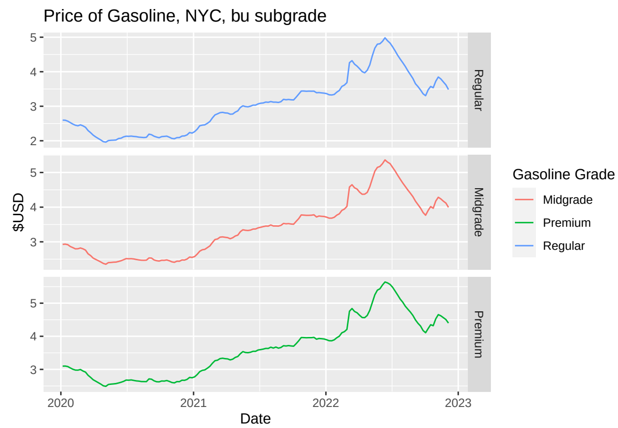
11. Stover, V. W., & Bae, C.-H. C. (2011). Impact of Gasoline Prices on Transit Ridership in Washington State. Transportation Research Record, 2217(1), 11–18. https://doi-org.remote.baruch.cuny.edu/10.3141/2217-02

12. Teixeira, J. F., & Lopes, M. (2020). The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York's Citi Bike. Transportation research interdisciplinary perspectives, 6, 100166. https://www.sciencedirect.com/science/article/pii/S2590198220300774

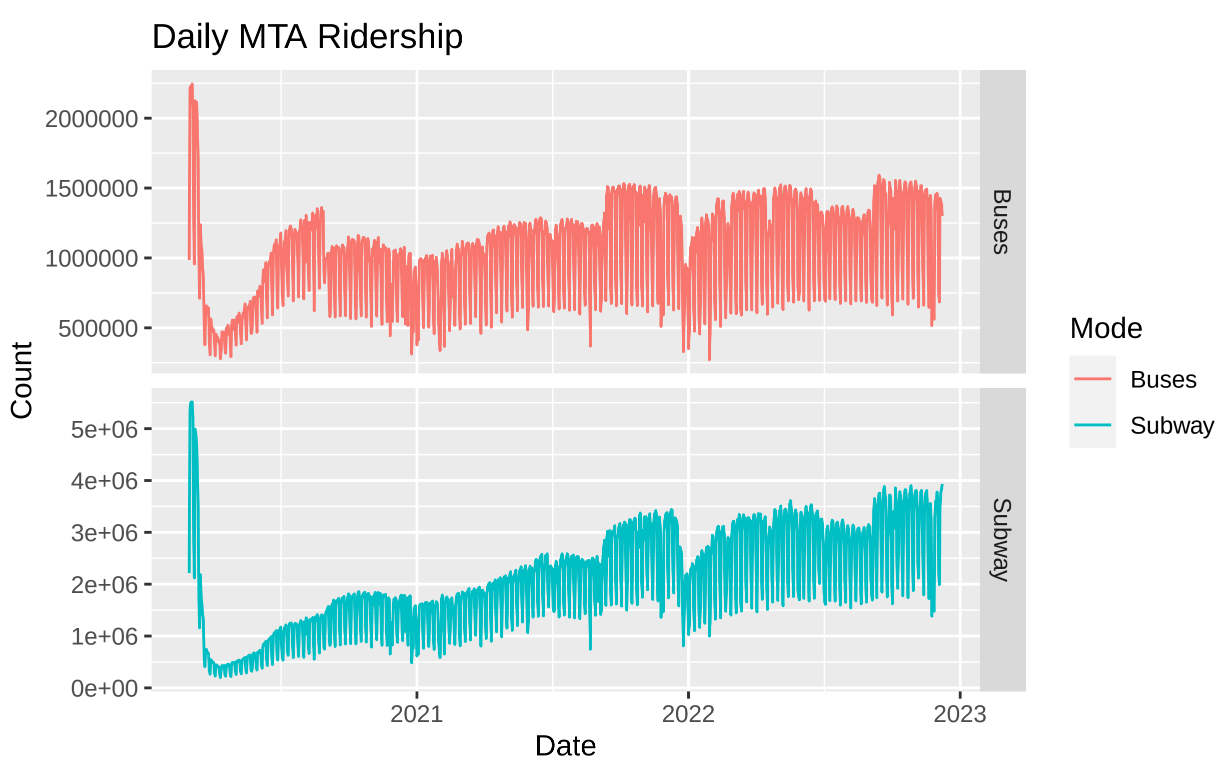
Appendix

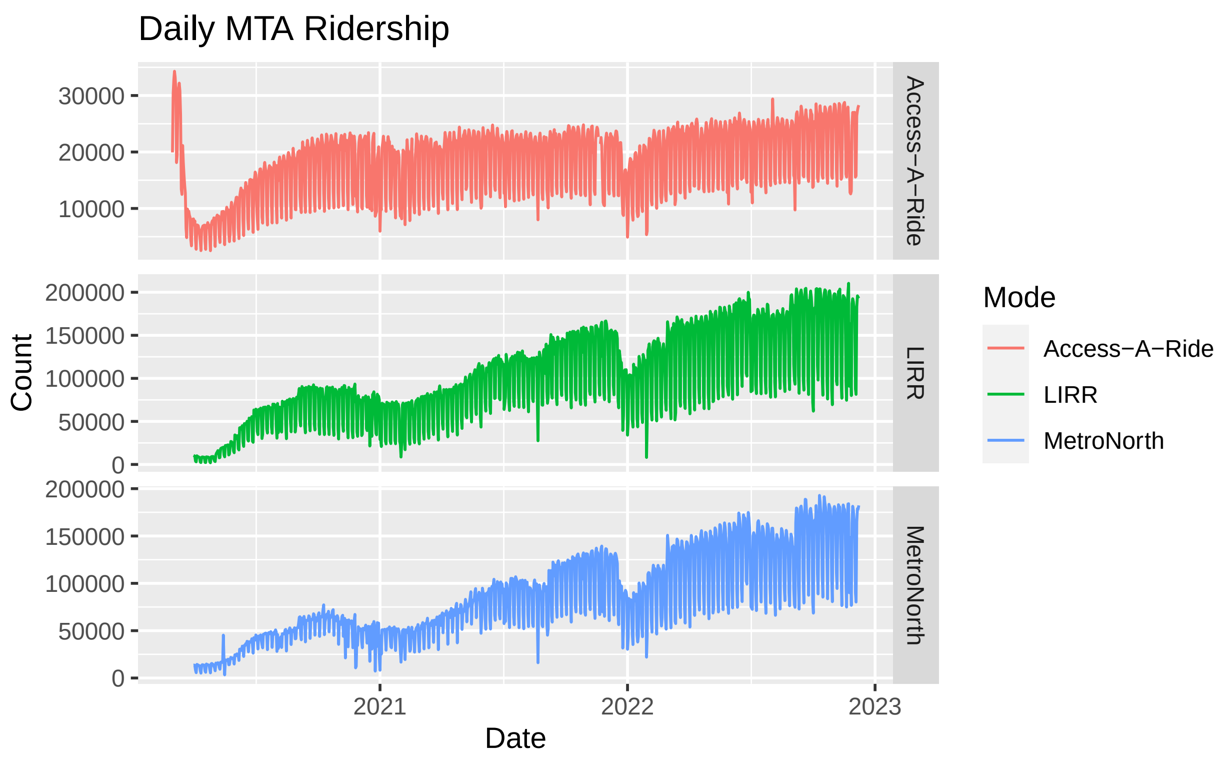
A.1 Weekly Price of Gasoline



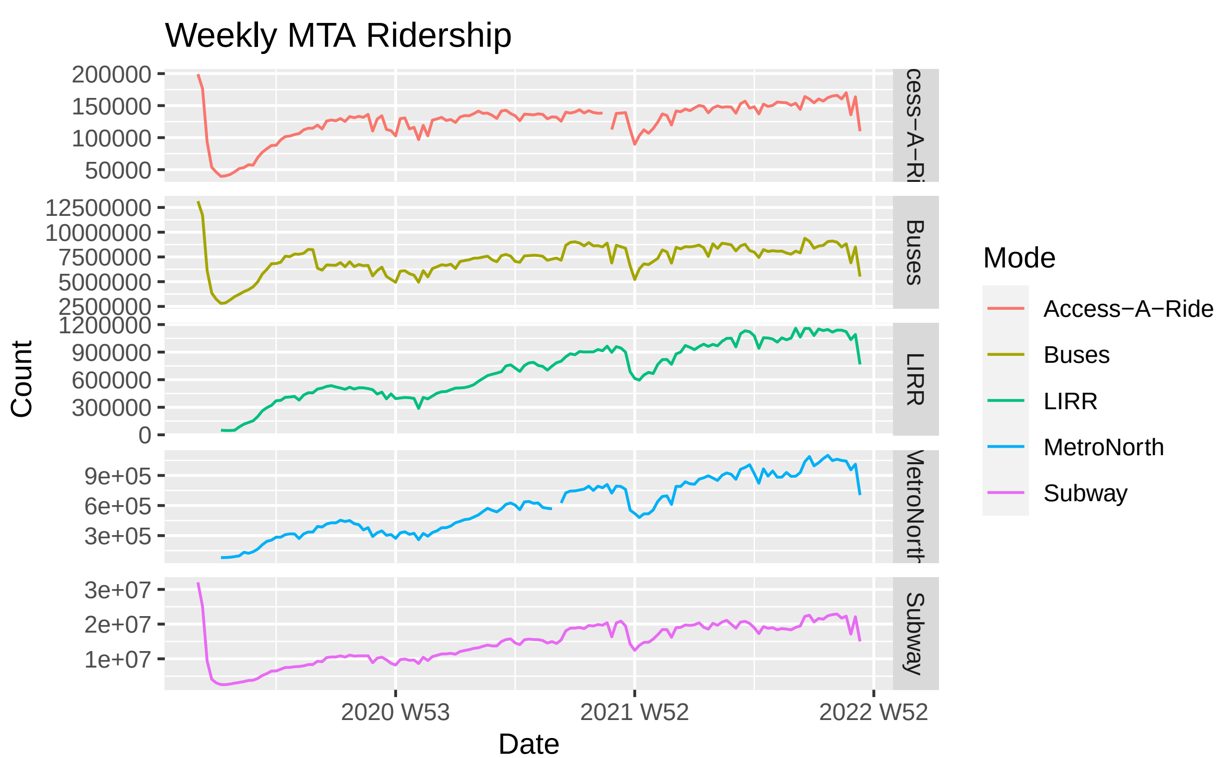


A.2 Daily MTA Ridership

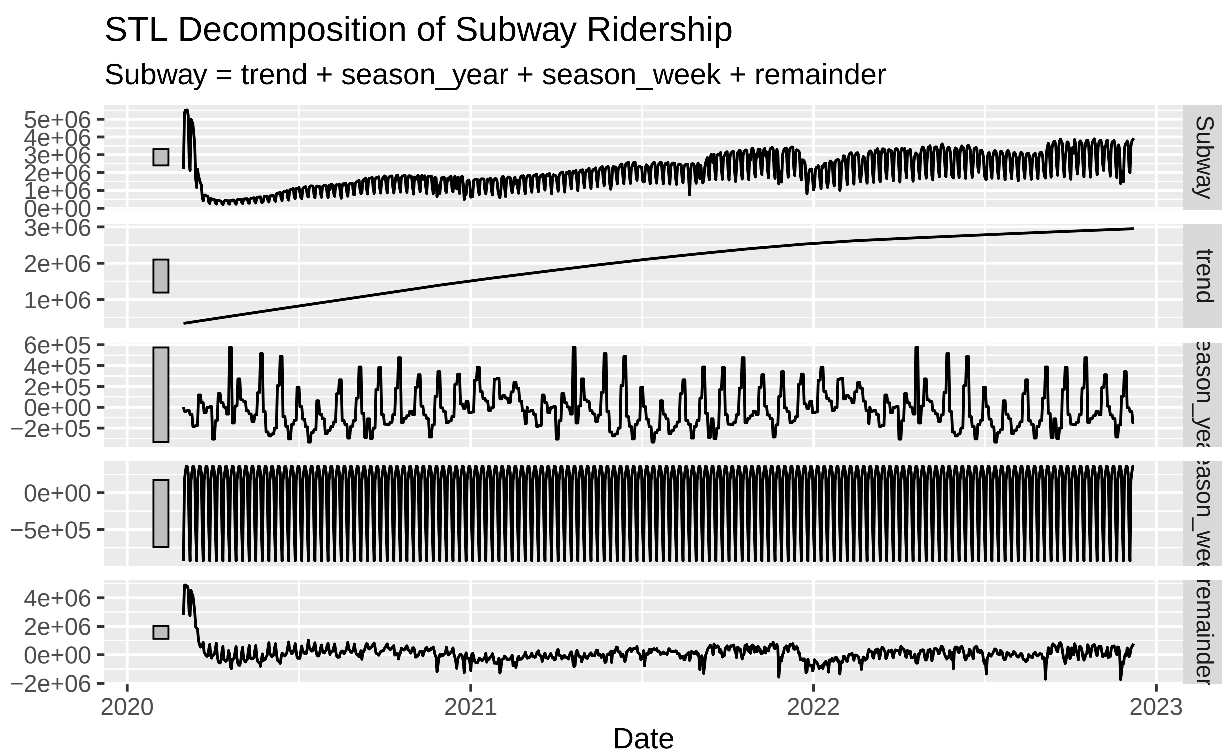




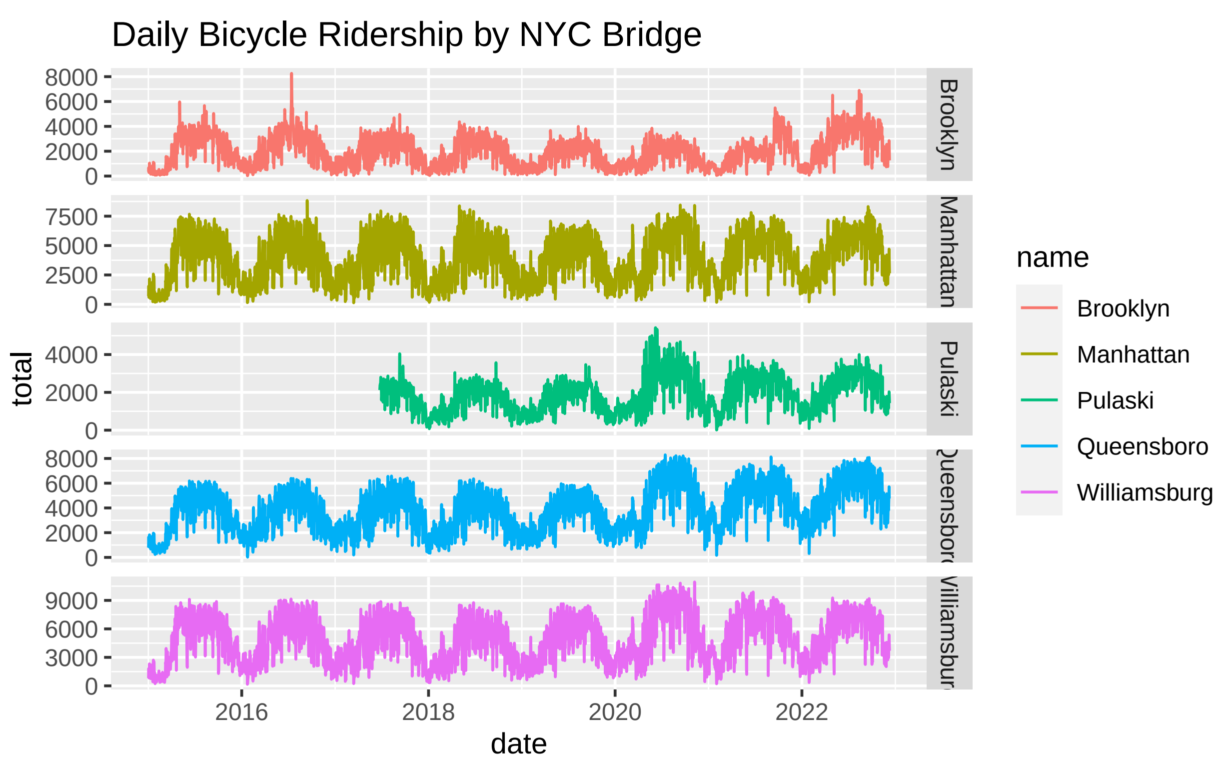
A.3 Weekly MTA Ridership



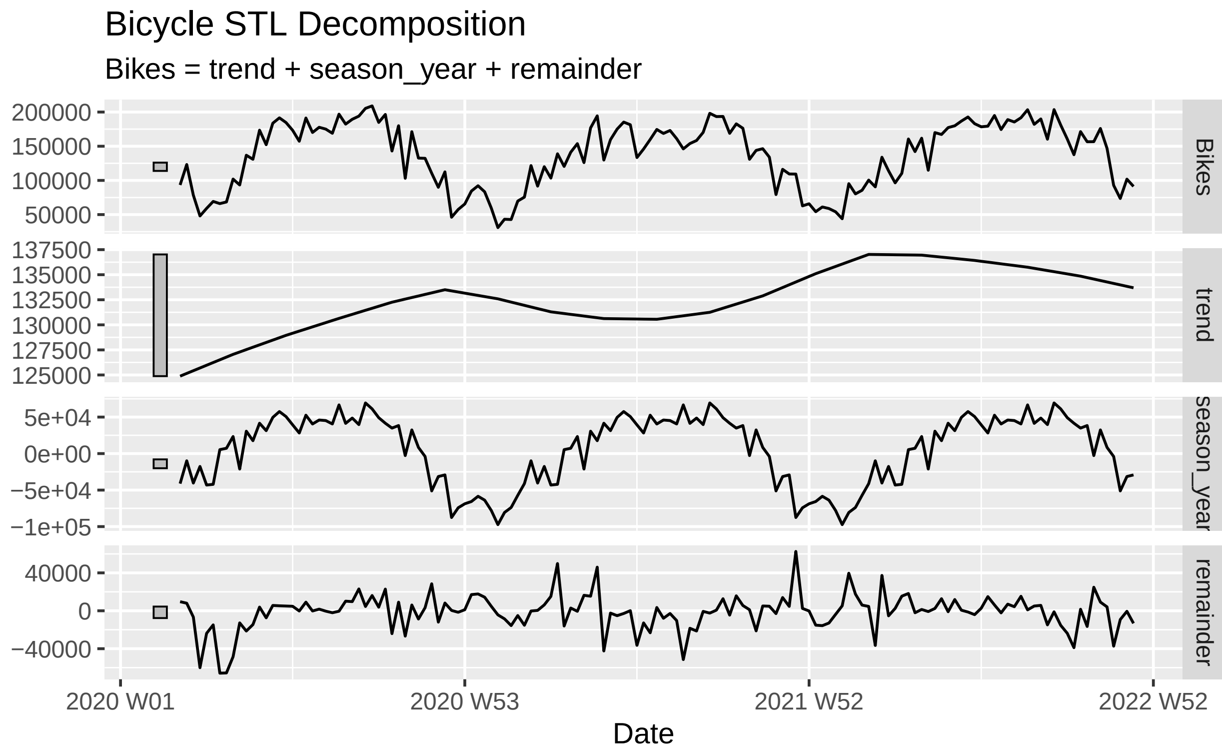
A.4 STL Decomposition of Subway Ridership



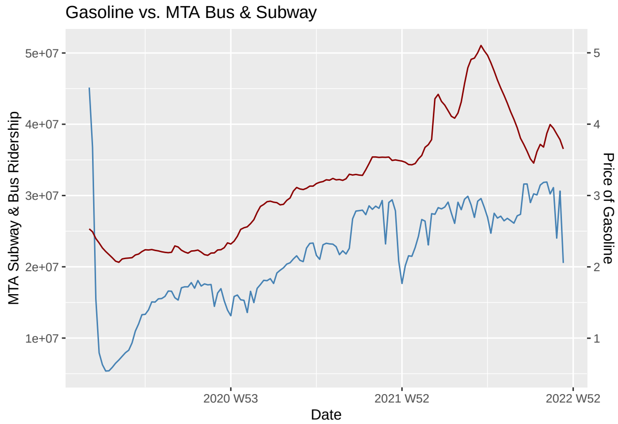
A.5 Daily Bicycle Ridership by NYC Bridge

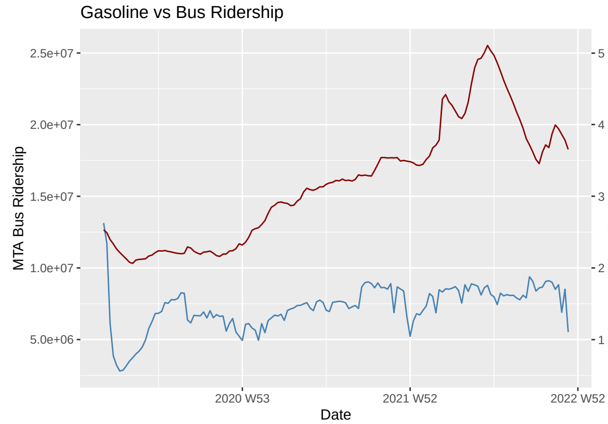


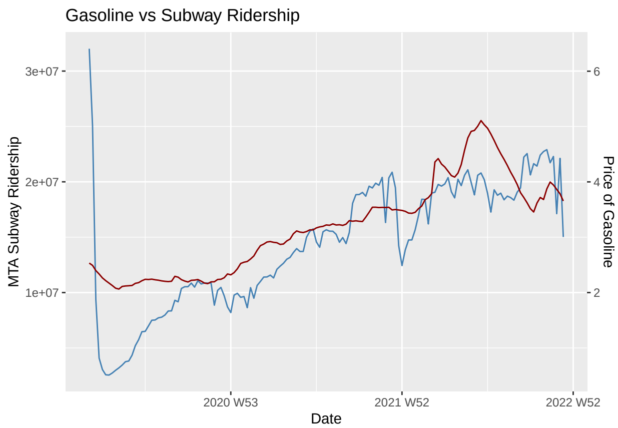
A.6 STL Decomposition of Weekly Bicycle Ridership

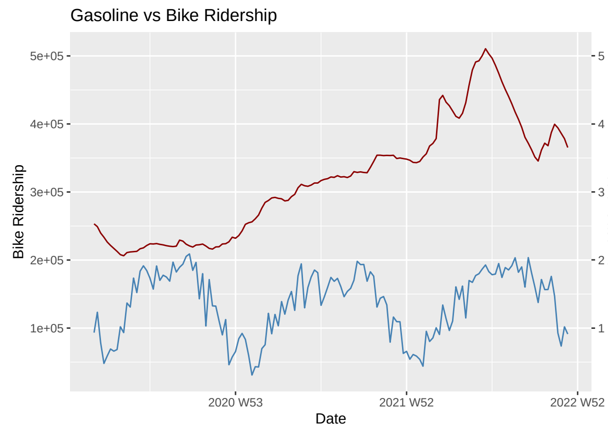


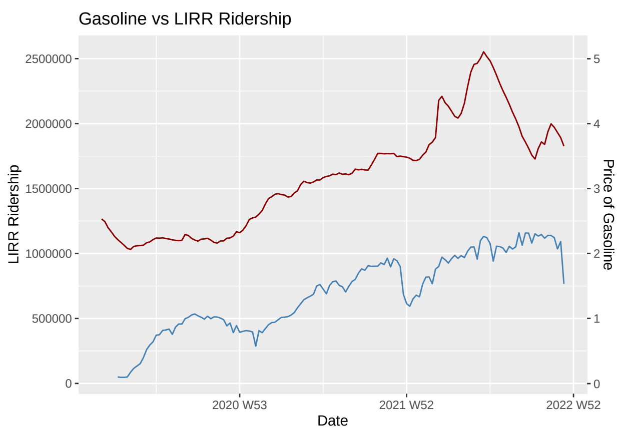
A.7 Gas vs. Mode of Transportation

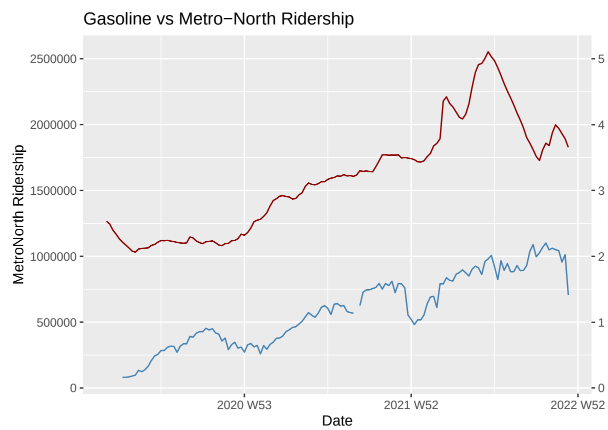


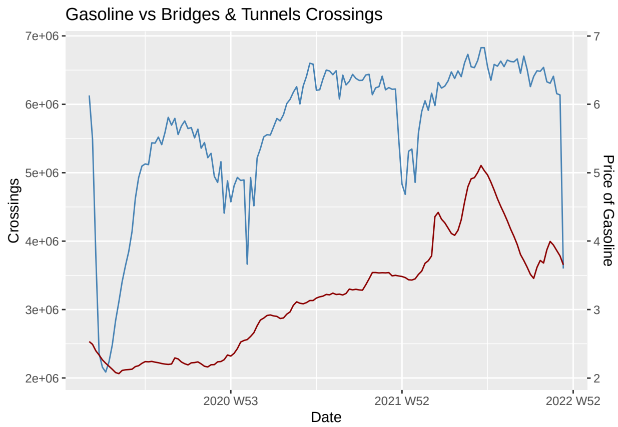




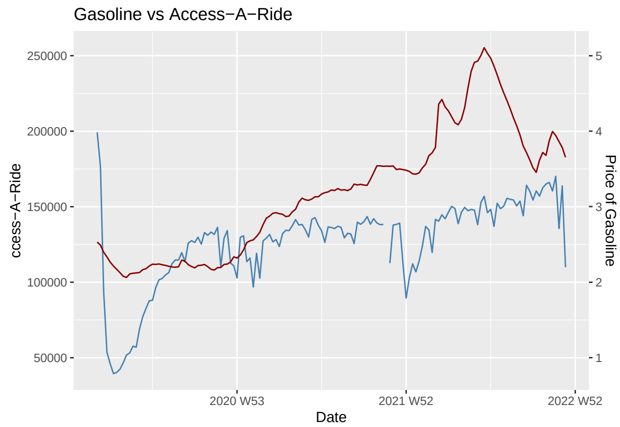








Gasoline vs Bridges & Tunnels Crossings
7e+06
6e+06
5e+06
4e+06
3e+06
2e+06
Crossings
7
6
5
4
3
2
Price of Gasoline
2020 W53
2021 W52
2022 W52
Date



Gasoline vs Access−A−Ride
250000
200000
150000
100000
50000
ccess−A−Ride
5
4
3
2
1
Price of Gasoline
2020 W53
2021 W52
2022 W52
Date

A.8 Baseline Regression Models with Mode of Transportation

Subway Linear

```
##
## Call:
## lm(formula = Subway ~ All_Grades, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6500633 -1706335  -250356  1491752 21217419
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3005632    1166011  -2.578    0.011 *
## All_Grades   5453226     354684  15.375   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3560000 on 143 degrees of freedom
## Multiple R-squared:  0.6231, Adjusted R-squared:  0.6204
## F-statistic: 236.4 on 1 and 143 DF,  p-value: < 2.2e-16
```

Buses Linear

```
##
## Call:
## lm(formula = Buses ~ All_Grades, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3434353  -581950    45629   571264  6560564
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3883422     422815   9.185 4.32e-16 ***
## All_Grades   1061775     128614   8.256 9.14e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1291000 on 143 degrees of freedom
## Multiple R-squared:  0.3228, Adjusted R-squared:  0.318
## F-statistic: 68.15 on 1 and 143 DF,  p-value: 9.136e-14
```

LIRR Linear

```
##
## Call:
## lm(formula = LIRR ~ All_Grades, data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -340309 -102596   -1990  107723  375142
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -300636      48838  -6.156 7.64e-09 ***
## All_Grades    311894      14734  21.168  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 145500 on 138 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.7645, Adjusted R-squared:  0.7628
## F-statistic: 448.1 on 1 and 138 DF,  p-value: < 2.2e-16
```

Metro-North Linear

```
##
## Call:
## lm(formula = MetroNorth ~ All_Grades, data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -249956  -99470  -18356   70060  374233
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -330400      44427  -7.437 1.02e-11 ***
## All_Grades    288558      13401  21.532  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 132400 on 137 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.7719, Adjusted R-squared:  0.7702
## F-statistic: 463.6 on 1 and 137 DF,  p-value: < 2.2e-16
```

Access-a-Ride Linear

```
##
## Call:
## lm(formula = ‘Access-A-Ride‘ ~ All_Grades, data = weekly)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -66519 -10227   4187  12287  86397
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    58511       7606   7.692 2.20e-12 ***
## All_Grades     21432       2315   9.258 2.96e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 23220 on 142 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.3764, Adjusted R-squared:  0.372
## F-statistic: 85.71 on 1 and 142 DF,  p-value: 2.955e-16
```

Bridges and Tunnels Linear

```
##
## Call:
## lm(formula = Bridges_and_Tunnels ~ All_Grades, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2781982  -283038   184308   555527  1001804
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3052025     259023   11.78   <2e-16 ***
## All_Grades    819993      78791   10.41   <2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 790800 on 143 degrees of freedom
## Multiple R-squared:  0.431,  Adjusted R-squared:  0.427
## F-statistic: 108.3 on 1 and 143 DF,  p-value: < 2.2e-16
```

Subway and Buses Combined Linear

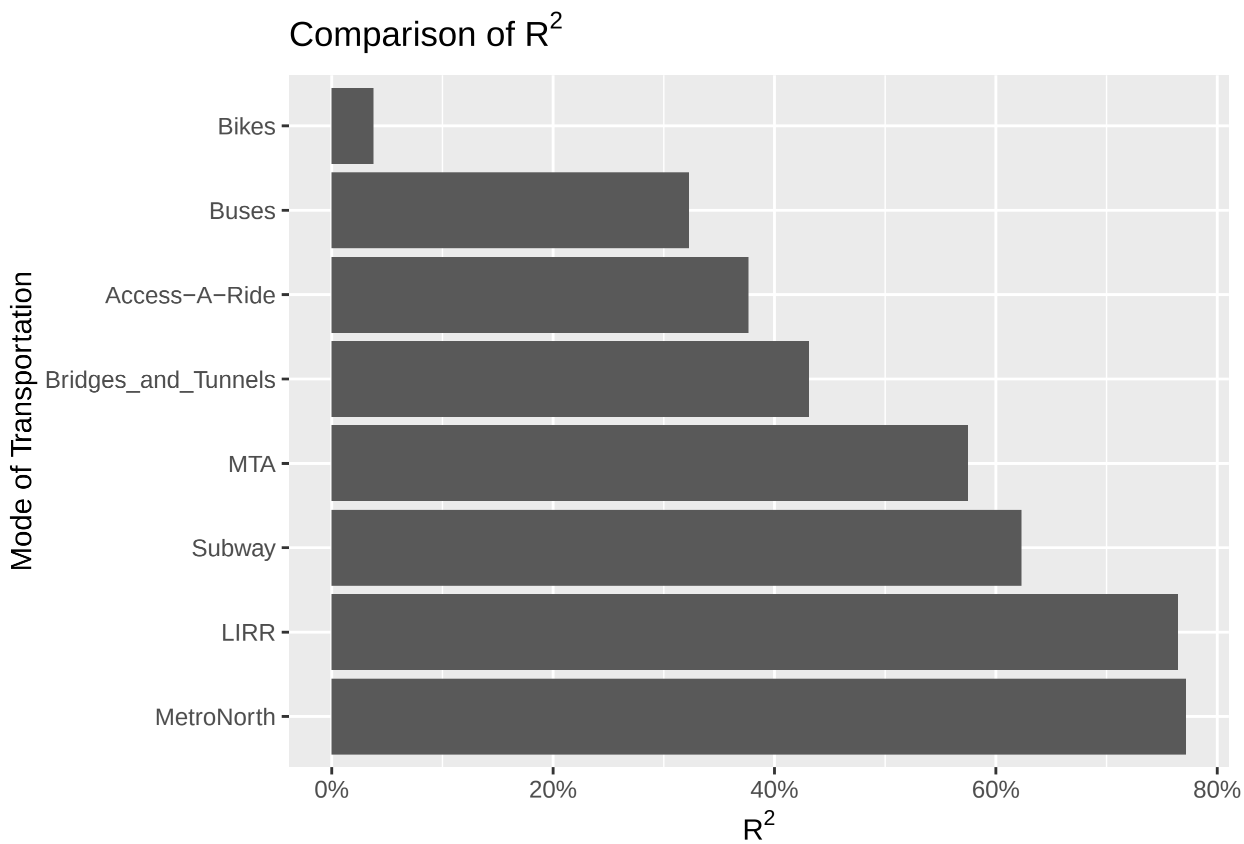
```
##
## Call:
## lm(formula = MTA ~ All_Grades, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9934986 -2063983  -185526  1931932 27777984
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   877790    1540355    0.57     0.57
## All_Grades   6515000     468554   13.90   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4702000 on 143 degrees of freedom
## Multiple R-squared:  0.5748, Adjusted R-squared:  0.5719
## F-statistic: 193.3 on 1 and 143 DF,  p-value: < 2.2e-16
```

Bikes Linear

```
##
## Call:
## lm(formula = Bikes ~ All_Grades, data = weekly)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -98947 -37539   9184  37377  82968
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   101684      15288   6.651 5.72e-10 ***
## All_Grades     11013       4650   2.368   0.0192 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 46670 on 143 degrees of freedom
## Multiple R-squared:  0.03774,    Adjusted R-squared:  0.03101
## F-statistic: 5.608 on 1 and 143 DF,  p-value: 0.01921
```

A.9 Linear Regression Comparison

```
## # A tibble: 8 x 13
##   rowname r.squared adj.r.squared  sigma statistic  p.value    df logLik   AIC
##   <chr>       <dbl>         <dbl>  <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl>
## 1 Subway     0.623         0.620  3.56e6    236.   4.23e-32     1 -2392. 4790.
## 2 Buses      0.323         0.318  1.29e6     68.2  9.14e-14     1 -2245. 4496.
## 3 LIRR       0.765         0.763  1.46e5    448.   3.55e-45     1 -1862. 3730.
## 4 MetroN~    0.772         0.770  1.32e5    464.   8.28e-46     1 -1835. 3677.
## 5 Access~    0.376         0.372  2.32e4     85.7  2.96e-16     1 -1651. 3308.
## 6 Bridge~    0.431         0.427  7.91e5    108.   3.12e-19     1 -2174. 4354.
## 7 MTA        0.575         0.572  4.70e6    193.   2.42e-28     1 -2432. 4871.
## 8 Bikes      0.0377        0.0310 4.67e4      5.61 1.92e- 2     1 -1764. 3533.
## # ... with 4 more variables: BIC <dbl>, deviance <dbl>, df.residual <int>,
## #   nobs <int>
```



A.10 Baseline Regression Models, Breakdown by Gasoline Grade

Subway Linear, by subgrade

```
##
## Call:
## lm(formula = Subway ~ Regular + Midgrade + Premium, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6801406 -2050592  -187081  1412487 22093578
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    910495    2693713   0.338  0.73586
## Regular      21392690    9100851   2.351  0.02013 *
## Midgrade    -41592320   16660778  -2.496  0.01370 *
## Premium      24860580    8649176   2.874  0.00468 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3489000 on 141 degrees of freedom
## Multiple R-squared:  0.6429, Adjusted R-squared:  0.6353
## F-statistic:  84.6 on 3 and 141 DF,  p-value: < 2.2e-16
```

Buses Linear, by subgrade

```
##
## Call:
## lm(formula = Buses ~ Regular + Midgrade + Premium, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3268729  -673516    57655   615968  6687775
##
## Coefficients:
##            Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4858430     999010   4.863 3.04e-06 ***
## Regular      4590464    3375206   1.360    0.176
## Midgrade    -6780560    6178934  -1.097    0.274
## Premium      3183393    3207695   0.992    0.323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1294000 on 141 degrees of freedom
## Multiple R-squared:  0.3288, Adjusted R-squared:  0.3145
## F-statistic: 23.02 on 3 and 141 DF,  p-value: 3.44e-12
```

LIRR Linear, by subgrade

```
##
## Call:
## lm(formula = LIRR ~ Regular + Midgrade + Premium, data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -330374 -103784   12247  105803  231467
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -264278     101162  -2.612  0.01000 *
## Regular       653681     346045   1.889  0.06102 .
## Midgrade    -2036904     634659  -3.209  0.00166 **
## Premium      1630259     328994   4.955 2.11e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 130800 on 136 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.8124, Adjusted R-squared:  0.8083
## F-statistic: 196.3 on 3 and 136 DF,  p-value: < 2.2e-16
```

Metro-North Linear, by subgrade

```
##
## Call:
## lm(formula = MetroNorth ~ Regular + Midgrade + Premium, data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -308891  -82475    7096   86803  231638
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -395226      87398  -4.522 1.33e-05 ***
## Regular       272938     298882   0.913  0.36277
## Midgrade    -1544542     548093  -2.818  0.00556 **
## Premium      1494163     284182   5.258 5.56e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 113000 on 135 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.8362, Adjusted R-squared:  0.8326
## F-statistic: 229.7 on 3 and 135 DF,  p-value: < 2.2e-16
```

Access-a-Ride Linear, by subgrade

```
##
## Call:
## lm(formula = `Access-A-Ride` ~ Regular + Midgrade + Premium,
##     data = weekly)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -52862  -9322   3320  11161  98631
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   124015      15939   7.781 1.43e-12 ***
## Regular       289083      53875   5.366 3.26e-07 ***
## Midgrade     -602330      98480  -6.116 9.04e-09 ***
## Premium       324907      51082   6.360 2.67e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 20610 on 140 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.5157, Adjusted R-squared:  0.5053
## F-statistic: 49.69 on 3 and 140 DF,  p-value: < 2.2e-16
```

Bridges and Tunnels Linear, by subgrade

```
##
## Call:
## lm(formula = Bridges_and_Tunnels ~ Regular + Midgrade + Premium,
##     data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2564716  -393721   163741   513218  1236044
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4893300     588706   8.312 7.12e-14 ***
## Regular       7692930    1988975   3.868 0.000167 ***
## Midgrade    -12593136    3641184  -3.459 0.000719 ***
## Premium       5600507    1890262   2.963 0.003579 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 762600 on 141 degrees of freedom
## Multiple R-squared:  0.4782, Adjusted R-squared:  0.4671
## F-statistic: 43.07 on 3 and 141 DF,  p-value: < 2.2e-16
```

Subway and Buses Combined Linear, by subgrade

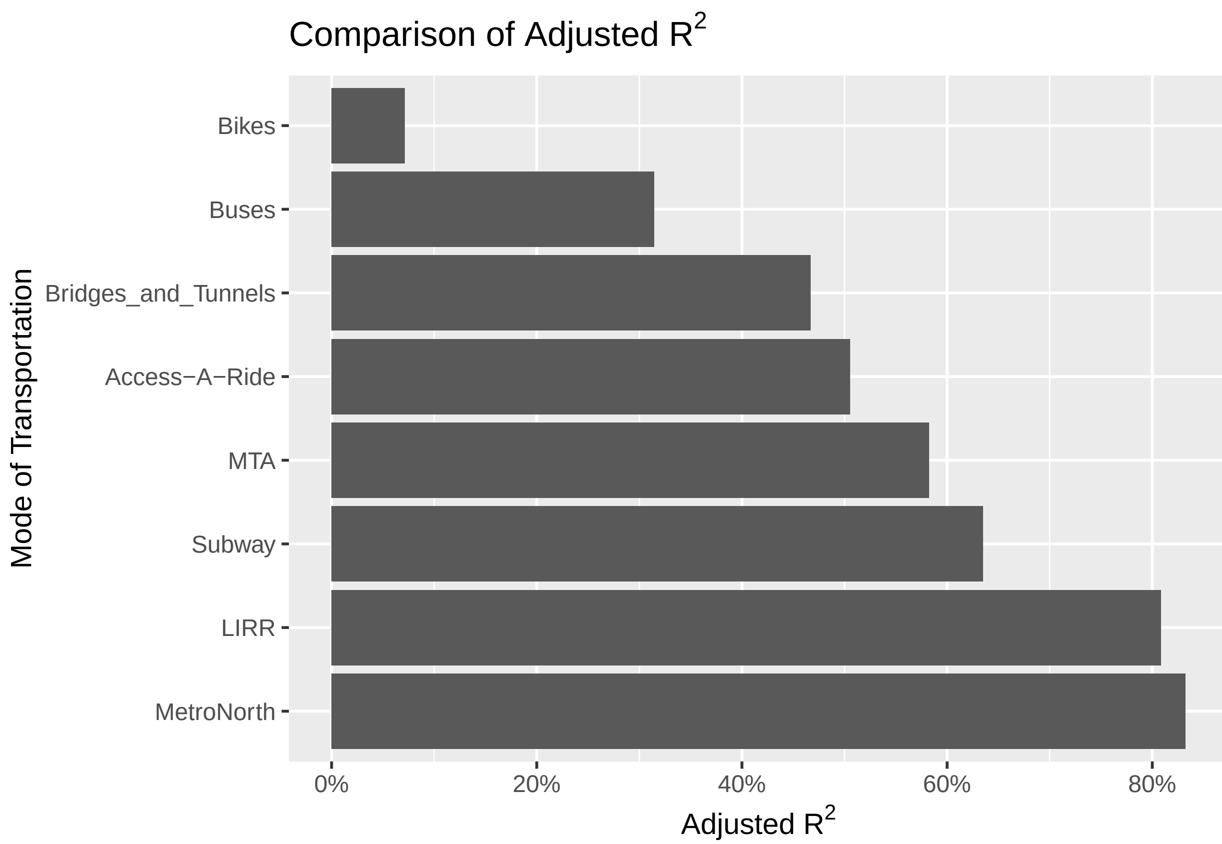
```
##
## Call:
## lm(formula = MTA ~ Regular + Midgrade + Premium, data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8854535 -2738598    57207  1960128 28781354
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5768925    3585763   1.609   0.1099
## Regular     25983154   12114686   2.145   0.0337 *
## Midgrade   -48372880   22178156  -2.181   0.0308 *
## Premium     28043973   11513435   2.436   0.0161 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4645000 on 141 degrees of freedom
## Multiple R-squared:  0.591,  Adjusted R-squared:  0.5823
## F-statistic:  67.9 on 3 and 141 DF,  p-value: < 2.2e-16
```

Bikes Linear, by subgrade

```
##
## Call:
## lm(formula = Bikes ~ Regular + Midgrade + Premium, data = weekly)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -97803 -36949  10733  32868  83949
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    12806      35269   0.363   0.7171
## Regular      -309816     119158  -2.600   0.0103 *
## Midgrade      455143     218140   2.086   0.0387 *
## Premium      -134138     113244  -1.185   0.2382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 45690 on 141 degrees of freedom
## Multiple R-squared:  0.0908, Adjusted R-squared:  0.07145
## F-statistic: 4.694 on 3 and 141 DF,  p-value: 0.003726
```

A.11 Linear Regression Comparison, by Gasoline Subgrade

```
## # A tibble: 8 x 13
##   rowname r.squared adj.r.squared  sigma statistic  p.value    df logLik   AIC
##   <chr>       <dbl>         <dbl>  <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl>
## 1 Subway     0.643         0.635  3.49e6     84.6  2.28e-31     3 -2388. 4786.
## 2 Buses      0.329         0.315  1.29e6     23.0  3.44e-12     3 -2244. 4499.
## 3 LIRR       0.812         0.808  1.31e5    196.   3.18e-49     3 -1846. 3702.
## 4 MetroN~    0.836         0.833  1.13e5    230.   7.89e-53     3 -1812. 3635.
## 5 Access~    0.516         0.505  2.06e4     49.7  6.23e-22     3 -1633. 3275.
## 6 Bridge~    0.478         0.467  7.63e5     43.1  8.08e-20     3 -2168. 4345.
## 7 MTA        0.591         0.582  4.64e6     67.9  3.13e-27     3 -2430. 4869.
## 8 Bikes      0.0908        0.0715 4.57e4      4.69 3.73e- 3     3 -1760. 3529.
## # ... with 4 more variables: BIC <dbl>, deviance <dbl>, df.residual <int>,
## #   nobs <int>
```



A.12 Linear-Log Regression with Logarithmic Transformation of Gas

Subway Linear-Log

```
##
## Call:
## lm(formula = Subway ~ log(All_Grades), data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5842698 -1769687  -545468  1809551 21186135
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -5960240    1213512  -4.912 2.44e-06 ***
## log(All_Grades) 18076735    1052659  17.172  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3313000 on 143 degrees of freedom
## Multiple R-squared:  0.6734, Adjusted R-squared:  0.6712
## F-statistic: 294.9 on 1 and 143 DF,  p-value: < 2.2e-16
```

Buses Linear-Log

```
##
## Call:
## lm(formula = Buses ~ log(All_Grades), data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3304422  -555300     -157   648910  6555555
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3301877     463164   7.129 4.61e-11 ***
## log(All_Grades)  3525226     401771   8.774 4.69e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1265000 on 143 degrees of freedom
## Multiple R-squared:   0.35,  Adjusted R-squared:  0.3454
## F-statistic: 76.99 on 1 and 143 DF,  p-value: 4.691e-15
```

LIRR Linear-Log

```
##
## Call:
## lm(formula = LIRR ~ log(All_Grades), data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -308005  -97987  -12208  104345  342203
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -450624      50553  -8.914 2.62e-15 ***
## log(All_Grades)  1016695      43526  23.358  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 134800 on 138 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.7981, Adjusted R-squared:  0.7967
## F-statistic: 545.6 on 1 and 138 DF,  p-value: < 2.2e-16
```

Metro-North Linear-Log

```
##
## Call:
## lm(formula = MetroNorth ~ log(All_Grades), data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -219177  -83177  -18060   71372  347781
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -465334      46473  -10.01   <2e-16 ***
## log(All_Grades)   937441      40017   23.43   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 123900 on 137 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.8002, Adjusted R-squared:  0.7988
## F-statistic: 548.8 on 1 and 137 DF,  p-value: < 2.2e-16
```

Access-a-Ride Linear-Log

```
##
## Call:
## lm(formula = `Access-A-Ride` ~ log(All_Grades), data = weekly)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -63748  -8805   2892  11866  86376
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)        46225       8249   5.604 1.05e-07 ***
## log(All_Grades)    71661       7161  10.008  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22520 on 142 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4136, Adjusted R-squared:  0.4095
## F-statistic: 100.2 on 1 and 142 DF,  p-value: < 2.2e-16
```

Bridges and Tunnels Linear-Log

```
##
## Call:
## lm(formula = Bridges_and_Tunnels ~ log(All_Grades), data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2682394  -254291   170355   511775  1062116
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2605502     280432   9.291 2.32e-16 ***
## log(All_Grades)  2720169     243260  11.182  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 765700 on 143 degrees of freedom
## Multiple R-squared:  0.4665, Adjusted R-squared:  0.4628
## F-statistic:   125 on 1 and 143 DF,  p-value: < 2.2e-16
```

Subway and Buses Combined Linear-Log

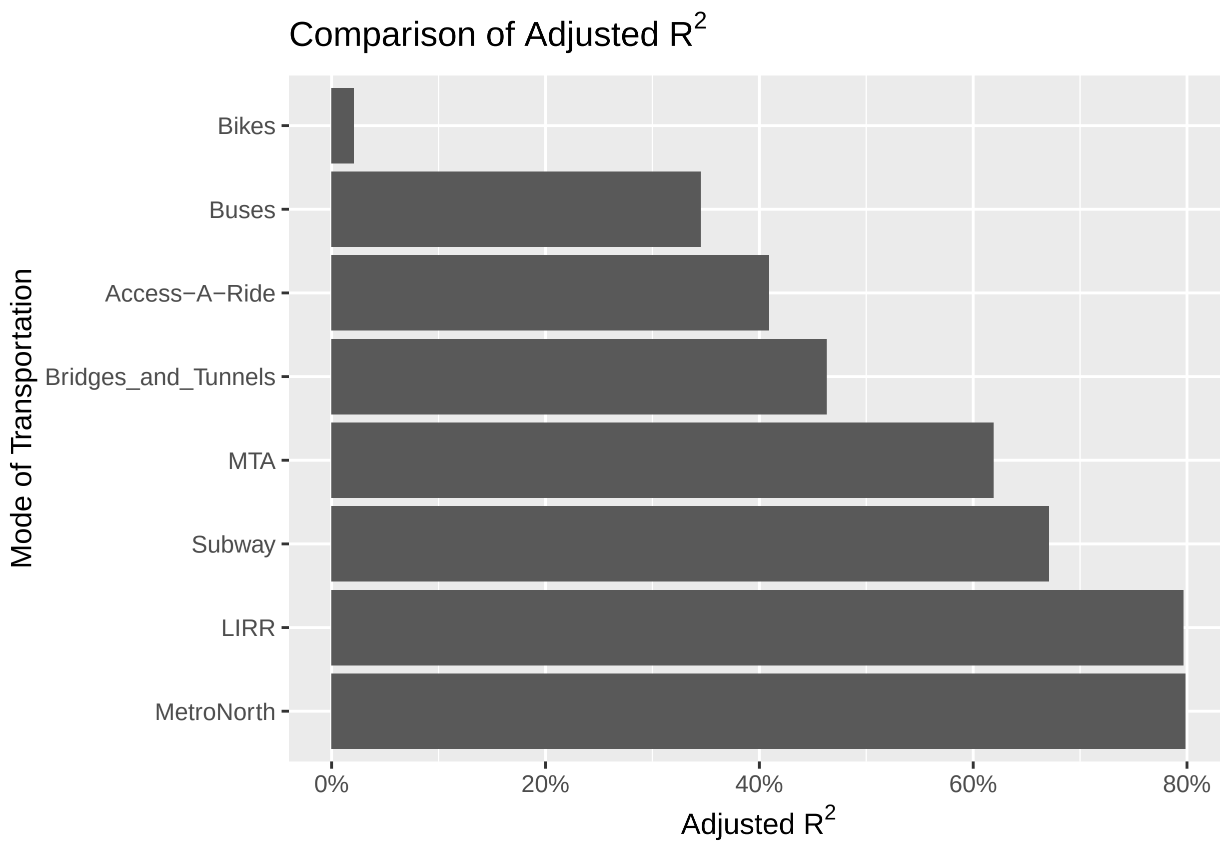
```
##
## Call:
## lm(formula = MTA ~ log(All_Grades), data = weekly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9147120 -2386471  -357666  2245694 27741690
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -2658363    1624774  -1.636    0.104
## log(All_Grades) 21601961    1409408  15.327   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4436000 on 143 degrees of freedom
## Multiple R-squared:  0.6216, Adjusted R-squared:  0.619
## F-statistic: 234.9 on 1 and 143 DF,  p-value: < 2.2e-16
```

Bikes Linear-Log

```
##
## Call:
## lm(formula = Bikes ~ log(All_Grades), data = weekly)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -100274  -37558    8166   38003   82217
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       102907      17182   5.989 1.63e-08 ***
## log(All_Grades)    30103      14904   2.020   0.0453 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 46910 on 143 degrees of freedom
## Multiple R-squared:  0.02773,    Adjusted R-squared:  0.02094
## F-statistic: 4.079 on 1 and 143 DF,  p-value: 0.04528
```

A.13 Linear-Log Regression Comparison

```
## # A tibble: 8 x 13
##   rowname r.squared adj.r.squared  sigma statistic  p.value    df logLik   AIC
##   <chr>       <dbl>         <dbl>  <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl>
## 1 Subway     0.673         0.671  3.31e6    295.   1.43e-36     1 -2382. 4769.
## 2 Buses      0.350         0.345  1.26e6     77.0  4.69e-15     1 -2242. 4490.
## 3 LIRR       0.798         0.797  1.35e5    546.   8.50e-50     1 -1851. 3708.
## 4 MetroN~    0.800         0.799  1.24e5    549.   9.26e-50     1 -1826. 3659.
## 5 Access~    0.414         0.409  2.25e4    100.   3.58e-18     1 -1646. 3299.
## 6 Bridge~    0.466         0.463  7.66e5    125.   2.99e-21     1 -2169. 4345.
## 7 MTA        0.622         0.619  4.44e6    235.   5.59e-32     1 -2424. 4854.
## 8 Bikes      0.0277        0.0209 4.69e4      4.08 4.53e- 2     1 -1764. 3535.
## # ... with 4 more variables: BIC <dbl>, deviance <dbl>, df.residual <int>,
## #   nobs <int>
```



A.14 Linear Regression with Dummy Variables

Subway Linear with log(Gas) + Trend + Season

```
## Series: Subway
## Model: TSLM
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6441451 -1083652  -139580  1210269 15878276
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -8596036    2186556  -3.931 0.000144 ***
## log(All_Grades)       21970549    3192232   6.883 3.19e-10 ***
## trend()                 -24419      20431  -1.195 0.234448
## fourier(K = 13)C1_52   2435610     436703   5.577 1.62e-07 ***
## fourier(K = 13)S1_52    371308     333058   1.115 0.267221
## fourier(K = 13)C2_52    873333     320661   2.724 0.007457 **
## fourier(K = 13)S2_52   -888498     314427  -2.826 0.005556 **
## fourier(K = 13)C3_52    988170     324421   3.046 0.002872 **
## fourier(K = 13)S3_52   -146845     318396  -0.461 0.645517
## fourier(K = 13)C4_52   -257150     319913  -0.804 0.423150
## fourier(K = 13)S4_52    412583     318455   1.296 0.197694
## fourier(K = 13)C5_52   -417787     321647  -1.299 0.196554
## fourier(K = 13)S5_52   -567000     317463  -1.786 0.076707 .
## fourier(K = 13)C6_52    191356     316929   0.604 0.547167
## fourier(K = 13)S6_52    147405     319413   0.461 0.645314
## fourier(K = 13)C7_52   -628585     316465  -1.986 0.049361 *
## fourier(K = 13)S7_52    118517     316402   0.375 0.708659
## fourier(K = 13)C8_52    517017     317199   1.630 0.105826
## fourier(K = 13)S8_52   -278258     315925  -0.881 0.380262
## fourier(K = 13)C9_52    372167     315804   1.178 0.241019
## fourier(K = 13)S9_52    438178     317315   1.381 0.169967
## fourier(K = 13)C10_52  -275831     316698  -0.871 0.385576
## fourier(K = 13)S10_52   172021     317296   0.542 0.588757
## fourier(K = 13)C11_52   410817     317312   1.295 0.198003
## fourier(K = 13)S11_52   -59139     314933  -0.188 0.851376
## fourier(K = 13)C12_52  -244216     315173  -0.775 0.439997
## fourier(K = 13)S12_52   317385     315522   1.006 0.316555
## fourier(K = 13)C13_52  -514290     313260  -1.642 0.103353
## fourier(K = 13)S13_52  -287869     315513  -0.912 0.363458
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2659000 on 116 degrees of freedom
## Multiple R-squared: 0.8293,  Adjusted R-squared: 0.7881
## F-statistic: 20.13 on 28 and 116 DF, p-value: < 2.22e-16
```

Buses Linear with log(Gas) + Trend + Season

```
## Series: Buses
## Model: TSLM
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2555289  -611400    23943   628375  4566947
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            1814916     904396   2.007 0.047100 *
## log(All_Grades)        5855225    1320359   4.435 2.11e-05 ***
## trend()                 -16099       8450  -1.905 0.059250 .
## fourier(K = 13)C1_52    725307     180627   4.015 0.000106 ***
## fourier(K = 13)S1_52   -360876     137758  -2.620 0.009979 **
## fourier(K = 13)C2_52    118382     132631   0.893 0.373937
## fourier(K = 13)S2_52   -349839     130052  -2.690 0.008200 **
## fourier(K = 13)C3_52    436003     134186   3.249 0.001514 **
## fourier(K = 13)S3_52     10802     131694   0.082 0.934772
## fourier(K = 13)C4_52    -38218     132321  -0.289 0.773226
## fourier(K = 13)S4_52    197964     131718   1.503 0.135573
## fourier(K = 13)C5_52    -52804     133038  -0.397 0.692167
## fourier(K = 13)S5_52   -238280     131308  -1.815 0.072159 .
## fourier(K = 13)C6_52     49594     131087   0.378 0.705881
## fourier(K = 13)S6_52     66475     132115   0.503 0.615806
## fourier(K = 13)C7_52   -288908     130895  -2.207 0.029271 *
## fourier(K = 13)S7_52      5136     130869   0.039 0.968760
## fourier(K = 13)C8_52    177146     131199   1.350 0.179577
## fourier(K = 13)S8_52    -88336     130672  -0.676 0.500377
## fourier(K = 13)C9_52    124004     130622   0.949 0.344422
## fourier(K = 13)S9_52    161555     131247   1.231 0.220840
## fourier(K = 13)C10_52   -57028     130992  -0.435 0.664111
## fourier(K = 13)S10_52    94051     131239   0.717 0.475039
## fourier(K = 13)C11_52   166945     131246   1.272 0.205917
## fourier(K = 13)S11_52    -5866     130262  -0.045 0.964160
## fourier(K = 13)C12_52  -133921     130361  -1.027 0.306411
## fourier(K = 13)S12_52    53579     130505   0.411 0.682158
## fourier(K = 13)C13_52  -169675     129569  -1.310 0.192942
## fourier(K = 13)S13_52  -163281     130501  -1.251 0.213385
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1100000 on 116 degrees of freedom
## Multiple R-squared: 0.601,   Adjusted R-squared: 0.5047
## F-statistic: 6.241 on 28 and 116 DF, p-value: 6.4776e-13
```

LIRR Linear with log(Gas) + Trend + Season

```
## Series: LIRR
## Model: TSLM
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -262752  -33553    8992   37419  122348
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -205648.06   54901.56  -3.746 0.000287 ***
## log(All_Grades)        590505.35   81606.78   7.236 6.36e-11 ***
## trend()                  3005.79     532.59   5.644 1.29e-07 ***
## fourier(K = 13)C1_52    96741.37   10805.16   8.953 9.11e-15 ***
## fourier(K = 13)S1_52   -73107.97    8411.17  -8.692 3.60e-14 ***
## fourier(K = 13)C2_52    38882.02    7898.89   4.922 2.99e-06 ***
## fourier(K = 13)S2_52   -13197.36    8001.25  -1.649 0.101890
## fourier(K = 13)C3_52    25910.97    8045.66   3.220 0.001679 **
## fourier(K = 13)S3_52     5174.01    7968.43   0.649 0.517477
## fourier(K = 13)C4_52   -13264.00    7991.57  -1.660 0.099788 .
## fourier(K = 13)S4_52     3037.37    7927.61   0.383 0.702350
## fourier(K = 13)C5_52    -9942.31    8019.91  -1.240 0.217700
## fourier(K = 13)S5_52     3264.50    7926.41   0.412 0.681243
## fourier(K = 13)C6_52    -7024.73    7953.88  -0.883 0.379046
## fourier(K = 13)S6_52       21.84    7938.63   0.003 0.997810
## fourier(K = 13)C7_52    -3398.20    7961.77  -0.427 0.670342
## fourier(K = 13)S7_52     -798.02    7900.94  -0.101 0.919730
## fourier(K = 13)C8_52     6486.69    7908.91   0.820 0.413874
## fourier(K = 13)S8_52     4946.19    7943.85   0.623 0.534795
## fourier(K = 13)C9_52     -981.28    7882.42  -0.124 0.901152
## fourier(K = 13)S9_52    -6457.10    7958.78  -0.811 0.418919
## fourier(K = 13)C10_52   -4023.83    7922.99  -0.508 0.612554
## fourier(K = 13)S10_52    6511.31    7933.81   0.821 0.413574
## fourier(K = 13)C11_52    1083.13    7934.94   0.137 0.891672
## fourier(K = 13)S11_52    4921.85    7899.55   0.623 0.534527
## fourier(K = 13)C12_52   -3151.81    7911.84  -0.398 0.691126
## fourier(K = 13)S12_52  -10000.04    7897.50  -1.266 0.208082
## fourier(K = 13)C13_52   -1406.23    7863.45  -0.179 0.858396
## fourier(K = 13)S13_52    -815.53    7857.71  -0.104 0.917525
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 65180 on 111 degrees of freedom
## Multiple R-squared: 0.962,   Adjusted R-squared: 0.9524
## F-statistic: 100.4 on 28 and 111 DF, p-value: < 2.22e-16
```

MetroNorth Linear with log(Gas) + Trend + Season

```
## Series: MetroNorth
## Model: TSLM
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -229462  -33600    1265   28719   98358
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -120950.1    44942.6  -2.691  0.00823 **
## log(All_Grades)        349232.5    66846.6   5.224 8.36e-07 ***
## trend()                  4115.9      436.2   9.436 7.67e-16 ***
## fourier(K = 13)C1_52    56954.9     8870.9   6.420 3.55e-09 ***
## fourier(K = 13)S1_52   -70309.6     6891.3 -10.203  < 2e-16 ***
## fourier(K = 13)C2_52    57165.4     6458.2   8.852 1.66e-14 ***
## fourier(K = 13)S2_52   -11997.7     6576.6  -1.824  0.07082 .
## fourier(K = 13)C3_52    27537.7     6622.1   4.158 6.37e-05 ***
## fourier(K = 13)S3_52     1502.9     6515.5   0.231  0.81801
## fourier(K = 13)C4_52    -1704.4     6550.1  -0.260  0.79518
## fourier(K = 13)S4_52    -6097.4     6498.5  -0.938  0.35016
## fourier(K = 13)C5_52   -10471.5     6553.2  -1.598  0.11293
## fourier(K = 13)S5_52     -687.8     6519.9  -0.105  0.91618
## fourier(K = 13)C6_52    -6583.2     6533.1  -1.008  0.31582
## fourier(K = 13)S6_52    -2618.9     6494.0  -0.403  0.68752
## fourier(K = 13)C7_52    -2573.9     6537.2  -0.394  0.69454
## fourier(K = 13)S7_52      237.2     6465.2   0.037  0.97080
## fourier(K = 13)C8_52     8729.0     6463.1   1.351  0.17960
## fourier(K = 13)S8_52     4662.0     6535.8   0.713  0.47717
## fourier(K = 13)C9_52     2192.6     6465.6   0.339  0.73517
## fourier(K = 13)S9_52    -2036.5     6522.3  -0.312  0.75545
## fourier(K = 13)C10_52   -2735.5     6512.7  -0.420  0.67529
## fourier(K = 13)S10_52    1689.3     6485.9   0.260  0.79499
## fourier(K = 13)C11_52    7651.7     6489.9   1.179  0.24093
## fourier(K = 13)S11_52    3168.5     6494.2   0.488  0.62660
## fourier(K = 13)C12_52   -7470.8     6479.9  -1.153  0.25144
## fourier(K = 13)S12_52   -9016.5     6481.3  -1.391  0.16698
## fourier(K = 13)C13_52   -6570.7     6472.8  -1.015  0.31227
## fourier(K = 13)S13_52   -4050.9     6420.7  -0.631  0.52941
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 53260 on 110 degrees of freedom
## Multiple R-squared: 0.9703,  Adjusted R-squared: 0.9628
## F-statistic: 128.6 on 28 and 110 DF, p-value: < 2.22e-16
```

Access-A-Ride Linear with Gas + Season

```
## # A mable: 1 x 1
##   `TSLM(\`Access-A-Ride\` ~ All_Grades + fourier(K = 13))`
##                                                      <model>
## 1                                                     <TSLM>
```

Bridges and Tunnels Linear with log(Gas) + Season

```
## Series: Bridges_and_Tunnels
## Model: TSLM
##
## Residuals:
##       Min       1Q   Median       3Q      Max
## -2188818  -252140    48552   362500  1178489
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            2670345     242383  11.017  < 2e-16 ***
## log(All_Grades)        2632547     210707  12.494  < 2e-16 ***
## fourier(K = 13)C1_52    365408      75515   4.839 4.02e-06 ***
## fourier(K = 13)S1_52   -480393      80997  -5.931 3.13e-08 ***
## fourier(K = 13)C2_52    -11087      78779  -0.141   0.8883
## fourier(K = 13)S2_52    -92949      77541  -1.199   0.2331
## fourier(K = 13)C3_52    170436      78932   2.159   0.0329 *
## fourier(K = 13)S3_52    -41680      77665  -0.537   0.5925
## fourier(K = 13)C4_52    -51446      77806  -0.661   0.5098
## fourier(K = 13)S4_52    116037      78589   1.477   0.1425
## fourier(K = 13)C5_52    -25337      78086  -0.324   0.7462
## fourier(K = 13)S5_52    -74895      78217  -0.958   0.3403
## fourier(K = 13)C6_52     79699      78206   1.019   0.3103
## fourier(K = 13)S6_52     34544      78081   0.442   0.6590
## fourier(K = 13)C7_52   -120652      78138  -1.544   0.1253
## fourier(K = 13)S7_52     40965      78113   0.524   0.6010
## fourier(K = 13)C8_52     26529      78243   0.339   0.7352
## fourier(K = 13)S8_52    -53617      77991  -0.687   0.4931
## fourier(K = 13)C9_52     19200      77976   0.246   0.8059
## fourier(K = 13)S9_52      3483      78276   0.044   0.9646
## fourier(K = 13)C10_52   -60446      77962  -0.775   0.4397
## fourier(K = 13)S10_52    24523      78248   0.313   0.7545
## fourier(K = 13)C11_52    21441      78341   0.274   0.7848
## fourier(K = 13)S11_52   -31375      77715  -0.404   0.6872
## fourier(K = 13)C12_52   -18099      77818  -0.233   0.8165
## fourier(K = 13)S12_52   -50220      77904  -0.645   0.5204
## fourier(K = 13)C13_52    17368      77346   0.225   0.8227
## fourier(K = 13)S13_52   -39232      77888  -0.504   0.6154
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 656700 on 117 degrees of freedom
## Multiple R-squared: 0.6789,  Adjusted R-squared: 0.6048
## F-statistic: 9.162 on 27 and 117 DF, p-value: < 2.22e-16
```

Combined Subway and Bus Linear with log(Gas) + Trend + Season

```
## Series: MTA
## Model: TSLM
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8349762 -1694400   -91096  1820192 20445223
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -6781121    3005443  -2.256  0.02593 *
## log(All_Grades)       27825775    4387754   6.342 4.53e-09 ***
## trend()                 -40517      28082  -1.443  0.15177
## fourier(K = 13)C1_52   3160917     600252   5.266 6.49e-07 ***
## fourier(K = 13)S1_52     10432     457791   0.023  0.98186
## fourier(K = 13)C2_52    991715     440751   2.250  0.02633 *
## fourier(K = 13)S2_52  -1238337     432183  -2.865  0.00495 **
## fourier(K = 13)C3_52   1424173     445920   3.194  0.00181 **
## fourier(K = 13)S3_52   -136044     437639  -0.311  0.75647
## fourier(K = 13)C4_52   -295368     439723  -0.672  0.50310
## fourier(K = 13)S4_52    610547     437720   1.395  0.16573
## fourier(K = 13)C5_52   -470591     442107  -1.064  0.28935
## fourier(K = 13)S5_52   -805280     436357  -1.845  0.06752 .
## fourier(K = 13)C6_52    240950     435622   0.553  0.58125
## fourier(K = 13)S6_52    213879     439037   0.487  0.62707
## fourier(K = 13)C7_52   -917493     434984  -2.109  0.03707 *
## fourier(K = 13)S7_52    123653     434897   0.284  0.77667
## fourier(K = 13)C8_52    694163     435993   1.592  0.11407
## fourier(K = 13)S8_52   -366594     434242  -0.844  0.40029
## fourier(K = 13)C9_52    496172     434076   1.143  0.25537
## fourier(K = 13)S9_52    599734     436153   1.375  0.17177
## fourier(K = 13)C10_52  -332859     435305  -0.765  0.44603
## fourier(K = 13)S10_52   266072     436127   0.610  0.54300
## fourier(K = 13)C11_52   577761     436149   1.325  0.18788
## fourier(K = 13)S11_52   -65005     432879  -0.150  0.88089
## fourier(K = 13)C12_52  -378137     433208  -0.873  0.38453
## fourier(K = 13)S12_52   370964     433688   0.855  0.39411
## fourier(K = 13)C13_52  -683965     430578  -1.588  0.11490
## fourier(K = 13)S13_52  -451150     433675  -1.040  0.30037
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3655000 on 116 degrees of freedom
## Multiple R-squared: 0.7916,  Adjusted R-squared: 0.7413
## F-statistic: 15.74 on 28 and 116 DF, p-value: < 2.22e-16
```

Bikes Linear with Gas + Season

```
## Series: Bikes
## Model: TSLM
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -49204 -13645   1114  14096  38322
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            111508.06    6862.47  16.249  < 2e-16 ***
## All_Grades               6383.19    2095.60   3.046  0.00287 **
## fourier(K = 13)C1_52     7239.23    2383.01   3.038  0.00294 **
## fourier(K = 13)S1_52   -59914.07    2559.95 -23.404  < 2e-16 ***
## fourier(K = 13)C2_52    12048.64    2481.47   4.855 3.75e-06 ***
## fourier(K = 13)S2_52    -3981.63    2445.33  -1.628  0.10616
## fourier(K = 13)C3_52     4480.24    2487.67   1.801  0.07428 .
## fourier(K = 13)S3_52    -5275.39    2447.11  -2.156  0.03315 *
## fourier(K = 13)C4_52    -4101.43    2451.48  -1.673  0.09699 .
## fourier(K = 13)S4_52      276.45    2475.79   0.112  0.91128
## fourier(K = 13)C5_52     2846.03    2460.11   1.157  0.24968
## fourier(K = 13)S5_52    -2104.75    2464.17  -0.854  0.39477
## fourier(K = 13)C6_52     -350.44    2463.83  -0.142  0.88714
## fourier(K = 13)S6_52     3320.21    2459.92   1.350  0.17971
## fourier(K = 13)C7_52    -4113.51    2461.63  -1.671  0.09738 .
## fourier(K = 13)S7_52     3068.13    2460.84   1.247  0.21497
## fourier(K = 13)C8_52     3914.60    2464.93   1.588  0.11496
## fourier(K = 13)S8_52    -1198.94    2456.97  -0.488  0.62648
## fourier(K = 13)C9_52     3520.93    2456.47   1.433  0.15443
## fourier(K = 13)S9_52     1727.70    2465.93   0.701  0.48493
## fourier(K = 13)C10_52      13.56    2456.08   0.006  0.99560
## fourier(K = 13)S10_52    2553.64    2465.11   1.036  0.30238
## fourier(K = 13)C11_52    2548.59    2468.03   1.033  0.30390
## fourier(K = 13)S11_52    3403.84    2448.34   1.390  0.16709
## fourier(K = 13)C12_52   -6508.25    2451.59  -2.655  0.00904 **
## fourier(K = 13)S12_52   -2220.57    2454.26  -0.905  0.36744
## fourier(K = 13)C13_52   -2792.50    2436.67  -1.146  0.25412
## fourier(K = 13)S13_52     310.19    2453.73   0.126  0.89962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 20690 on 117 degrees of freedom
## Multiple R-squared: 0.8453,  Adjusted R-squared: 0.8096
## F-statistic: 23.68 on 27 and 117 DF, p-value: < 2.22e-16
```

A.15 Linear with Trend and Seasonality Regression Comparison

Linear with Trend and Seasonality

```
##                                                            .model r_squared
## 1     TSLM(Subway ~ log(All_Grades) + trend() + fourier(K = 13)) 0.8293352
## 2      TSLM(Buses ~ log(All_Grades) + trend() + fourier(K = 13)) 0.6010397
## 3       TSLM(LIRR ~ log(All_Grades) + trend() + fourier(K = 13)) 0.9620075
## 4 TSLM(MetroNorth ~ log(All_Grades) + trend() + fourier(K = 13)) 0.9703479
## 5                            TSLM(Bikes ~ All_Grades + fourier(K = 13)) 0.8452933
## 6              TSLM(`Access-A-Ride` ~ All_Grades + fourier(K = 13)) 0.5976953
## 7  TSLM(Bridges_and_Tunnels ~ log(All_Grades) + fourier(K = 13)) 0.6789031
## 8        TSLM(MTA ~ log(All_Grades) + trend() + fourier(K = 13)) 0.7915920
##   adj_r_squared
## 1     0.7881403
## 2     0.5047390
## 3     0.9524238
## 4     0.9628001
## 5     0.8095918
## 6     0.5040555
## 7     0.6048038
## 8     0.7412866
```