Towards a Better LA Metro Rail: A Comprehensive Proposal for SMART

Viswadatta Chinta - Unstucker

Domonick Demas - Research god

Joseph Lee - Liaison of affairs

Nirbhay Singh Naruka - Content creator

Prit Shah - Supreme team leader

Our dataset encompasses a comprehensive array of variables relevant to public transportation and socioeconomic indicators. Spanning two years post-pandemic, this dataset is designed to offer accurate insights into the evolving dynamics of urban mobility and associated factors. The variables include:

1. Ridership Data: This encompasses the monthly counts of passengers using various modes of public transportation, specifically focusing on Metro, rail, and bus services. This variable is crucial for understanding public transit trends and their correlation with other socioeconomic factors.

2. Unemployment Rates: Monthly unemployment figures provide a window into the economic health of the region, offering a backdrop against which transportation trends can be analyzed.

3. Gas Prices: Fluctuations in gas prices are included as they significantly influence public transport usage. Higher gas prices often correlate with increased public transit ridership.

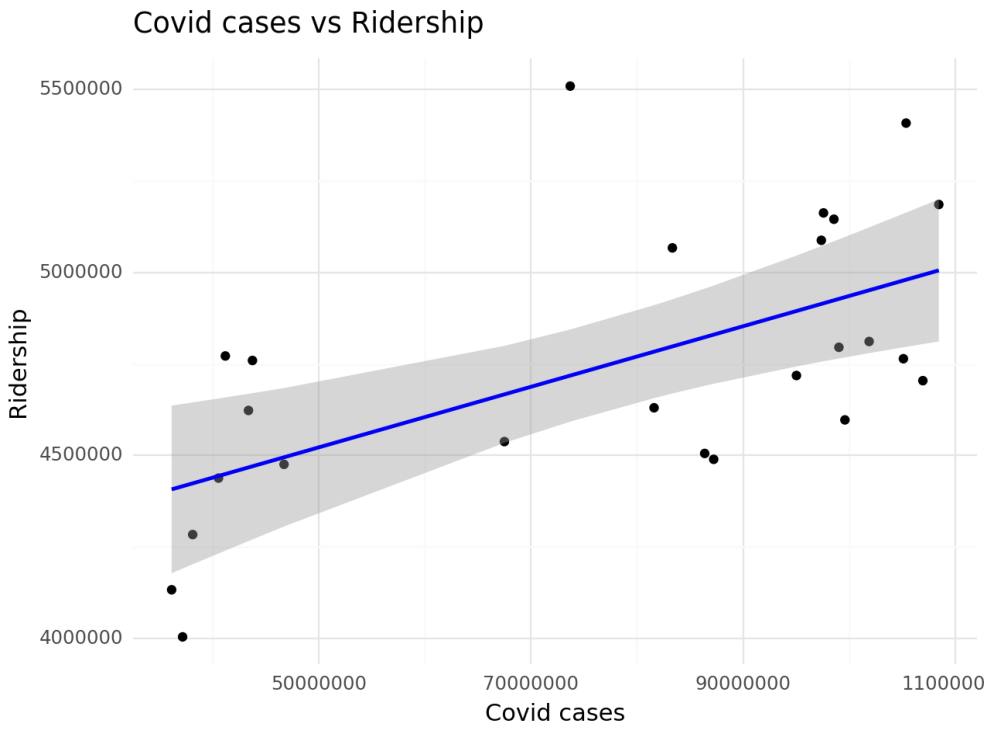
4. COVID-19 Case Counts: Monthly data on COVID-19 cases are crucial for understanding the pandemic's impact on public behavior, especially in the context of public transportation usage.

5. Crime Rates: Monthly crime statistics are included to assess safety perceptions and their potential impact on public transit usage.

6. Fare Price (LA MTA): Tracking the fare prices for Los Angeles Metro Transit Authority (LA MTA) provides insights into affordability and its influence on ridership patterns.

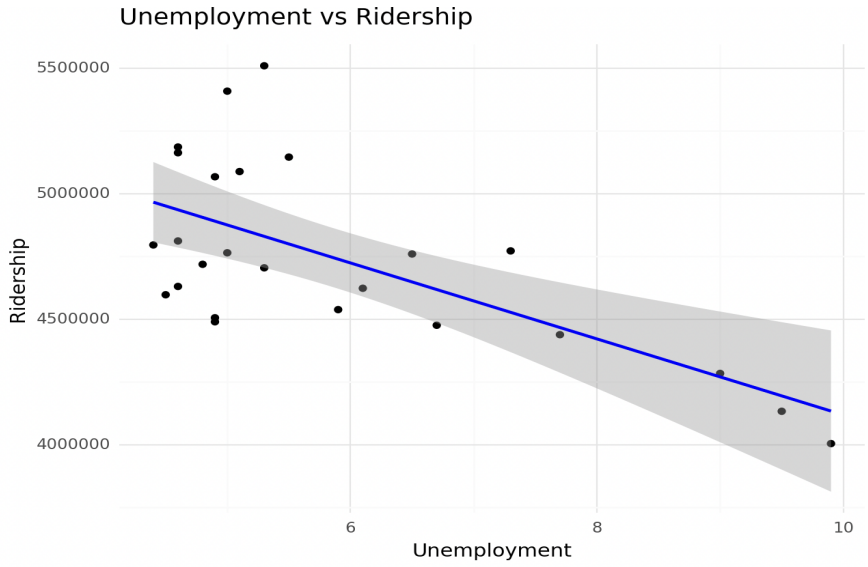
	Crime rate	Gas	Unemployment	Covid cases	Ridership	Rail Fare
count	25.000000	25.00000	25.00000	2.500000e+01	2.500000e+01	25.000000
mean	19017.960000	5.18512	5.86400	7.687504e+07	4.744449e+06	88.000000
std	877.619198	0.62890	1.62144	2.707232e+07	3.708916e+05	21.794495
min	17153.000000	4.38300	4.40000	3.616439e+07	4.004502e+06	50.000000
25%	18374.000000	4.68200	4.80000	4.377727e+07	4.505616e+06	100.000000
50%	18762.000000	5.02400	5.10000	8.638245e+07	4.718506e+06	100.000000
75%	19784.000000	5.59900	6.50000	9.902614e+07	5.067135e+06	100.000000
max	20416.000000	6.50600	9.90000	1.084390e+08	5.509095e+06	100.000000

In our dataset of 25 data points, we observe notable trends and variations across key metrics. The crime rate averages at around 19,018 with a standard deviation of 878, indicating a moderate fluctuation in these figures. Gas prices, with an average of approximately 5.19, also display a similar degree of variability. The unemployment rate, averaging 5.86, shows a wider range of variation, as evidenced by its standard deviation of 1.62. Significantly, COVID-19 case numbers average at about 76,875,040, highlighting the substantial impact of the pandemic during the observed time frame. Public transportation ridership averages around 4,744,449 with a standard deviation of 370,891, suggesting fluctuations in usage.

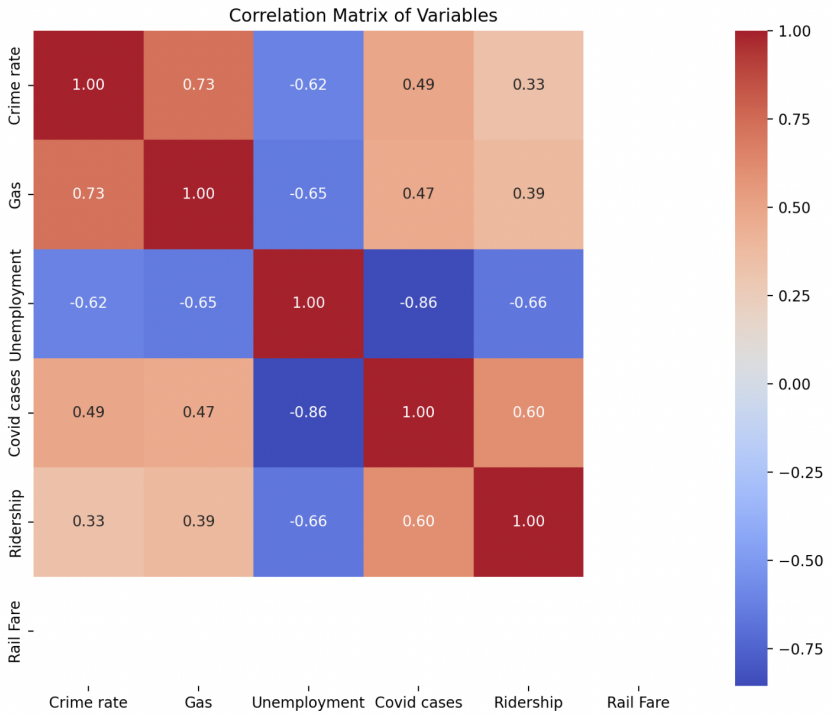


The scatter plot shows a favorable relationship between the number of COVID-19 cases and the use of public transportation. The trend line indicates that as COVID-19 instances grew, so did ridership, with the shaded region representing the confidence interval for this link.

People gradually resumed regular activities when COVID-19 limitations were relaxed and vaccination rates improved, boosting economic activity and mobility, which led to increased use of public transportation. This rise in riding corresponded with an increase in recorded COVID-19 cases.



The scatter plot shows an inverse relationship between unemployment and public transportation utilization, demonstrating that more unemployment is linked with decreased ridership. The confidence interval is shown by the shaded region surrounding the trend line, which provides a range for the strength of this connection.



The heat map is a correlation matrix that depicts the links between factors like crime rate, gas prices, unemployment, COVID-19 cases, ridership, and rail fare. Notably, unemployment has a large negative link with ridership and COVID-19 instances, but crime rate and gas costs have a strong positive correlation.

From June 2021 to December 2022, the regular one-way trip cost $1.75 and the 30-day pass cost $100.In late 2022, LA approved moving to a fare cap system that took effect in late 2023. From January 2023 to June 2023, all fares were 50% off, so the one-way trip cost $0.875 and the 30-day pass cost $50.

Research Questions and Models

Did the rail fare affect the demand for public transportation services in the LA metro?

Model Specifications: linear regression

Feature Sets: Ridership, Fare price (LA MTA)

There is a negative correlation between the Fare Price and ridership on LA Metro. This could be because higher fares make public transportation less affordable for many people, leading to decreased usage. For a more accurate understanding, a detailed study considering various other factors would be necessary.

How does the crime rate impact public transportation usage?

Model Specifications: linear regression

Feature Sets: Crime rate, Ridership

Interestingly, this graph shows a positive correlation as the crime rate increases, the ridership also increases. This is different from what one might expect, as higher crime rates could potentially deter people from using public transportation due to safety concerns. However, this could be due to various factors not shown in the graph, such as the effectiveness of law enforcement, the presence of security measures on public transportation, or socioeconomic factors.

How do gas prices influence public transportation usage?

Model Specifications: Multiple linear regression

Feature Sets: Gas prices, Number of passengers

The Gas Prices Impact Model reveals that changes in gas prices have a modest impact on public transportation usage, explaining about 18.8% of the variation in ridership. While there is a correlation, gas prices are not a dominant factor influencing ridership.

What is the impact of COVID-19 cases on public transportation usage?

Model Specifications: Random forest

Feature Sets: Number of COVID-19 cases, Number of passengers

The COVID-19 Impact Model indicates that COVID-19 case trends have a limited influence on public transportation usage, explaining only about 9.6% of the variation in ridership. Other factors not included in the model significantly affect ridership levels.

How does the unemployment rate affect public transportation usage?

Model Specifications: Linear regression

Feature Sets: Unemployment rate, Number of passengers

-Linear Regression: The linear regression model suggests a moderate relationship between the unemployment rate and public transportation ridership, explaining approximately 56.3% of the variation. As unemployment rates change, ridership tends to respond accordingly.

-Support Vector Machines (SVM): In contrast, the SVM model does not effectively explain changes in ridership based on unemployment rates, as its R^2 score is close to zero. It fails to establish a meaningful relationship.

Model	R^2 Score	Interpretation
COVID-19 Impact Model	0.096	Limited influence of COVID-19 cases
Unemployment - Linear	0.563	Moderate impact of unemployment rates
Unemployment - SVM	≈ 0	Ineffective in explaining unemployment impact

Gas Prices Impact Model	0.188	Modest impact of gas price fluctuations
Crime Rate Impact Model	0.196	Weak influence of crime rate changes

- Linear Regression was chosen as a starting point and we used the full dataset with all the variables included, and this yielded a root-mean-squared error (RMSE) value of 288833.44.

Model Usage

In developing a predictive model for SMART, we aim to harness key variables such as crime rate, gas prices, unemployment rate, and COVID-19 case trends to forecast public transportation ridership. This model offers valuable insights into the correlation between these factors and ridership, enabling informed decision-making in service planning, resource allocation, and policy formulation. By analyzing the impact of economic indicators and public health data, SMART can strategically adjust services to meet changing demands, enhance passenger safety, and optimize operational efficiency. The model's predictive capabilities also facilitate scenario planning, allowing SMART to proactively respond to potential shifts in ridership patterns. Continuous refinement and updating of the model with the latest data will ensure its relevance and accuracy, making it a vital tool for strategic planning and fostering a responsive, efficient public transportation system.

In guiding SMART's use of our predictive model, a key focus should be on ensuring equitable service delivery. Insights derived from variables like crime rates and unemployment should be utilized in a way that avoids disadvantaged and under-served communities. Decisions around service adjustments should be carefully weighed for their impact on accessibility. This will ensure that enhancements to efficiency align with SMART’s commitment to inclusivity and social equity.

Citations

https://www.metro.net/about/financebudget/

https://isotp.metro.net/MetroRidership/YearOverYear.aspx

https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8

https://ycharts.com/indicators/los_angeles_retail_price_of_regular_gasoline

https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/

https://ycharts.com/indicators/los_angeles_ca_unemployment_rate#:~:text=Los%20Angeles%2C%20CA%20Unemployment%20Rate%20is%20at%205.30%25%2C%20compared,but%20are%20without%20a%20job.