# Leveraging Data to Reduce $CO_2$ Emissions and Enhance Public Transit Efficiency

By:  Owen Barberie, Will Kapner, Daniel Chamberlin, JD Singh,  Adarsh Parsekar

## Data

Our data included information on public transport across different regions in the US. Specifically, the variables included expenditure per capita, revenue per capita, transportation mode percentages, and energy consumption.



Expenditure vs. Public Transportation Percentage by Region



CO2 Emissions Over Time by Region
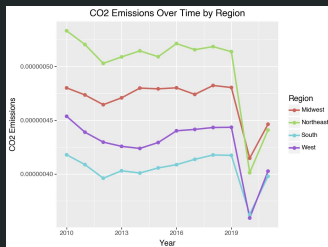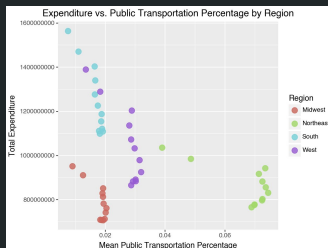
## Strategy

We decided to look at the following variables as predictors for emissions: expenditure, revenue, region, and percent of population in region who use public transport or drive.

We standardized expenditure and revenue on population to find a comparable statistic between locations.

**Expenditure** ↑

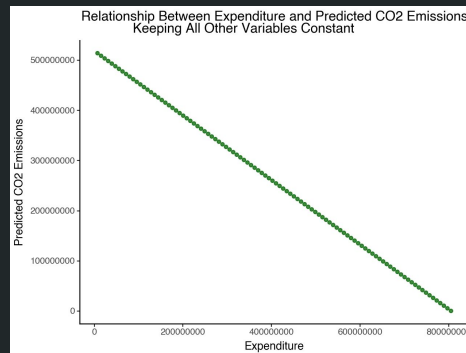**Emissions** ↓

## Findings

Final model with highest $R^2$ and lowest RMSE:
**Elastic Net (alpha = 0.1, L1 = 0.65)**
(Also tried: Logistic, Linear, Ridge, LASSO, and Decision Tree models)

Critical predictor variables:
1. Public Transportation Percentages
2. Personal Vehicle Use Percentages
3. Expenditures per Capita
4. Revenue per Capita

These variables were responsible for 13.24% of the variation in $CO_2$ emissions.



Relationship Between Expenditure and Predicted CO2 Emissions Keeping All Other Variables Constant

## Recommendations

- Increase public transportation usage as that variable had the largest effect on reducing emissions.

- In order to lower the emissions, SMART needs to target Public Transport Expenditure in effect increasing Revenue, allowing for further advances in reducing our carbon footprint.

- Model public transportation after the Northeast region which has the highest mean public transportation percentages, lowest mean personal vehicle percentages ,and most importantly the best emissions impact.

- Launch public awareness campaigns to educate the public on the benefits of public transportation.

- Use predictive analytics to enhance the user experience of public transportation by reducing wait times and optimizing schedules.