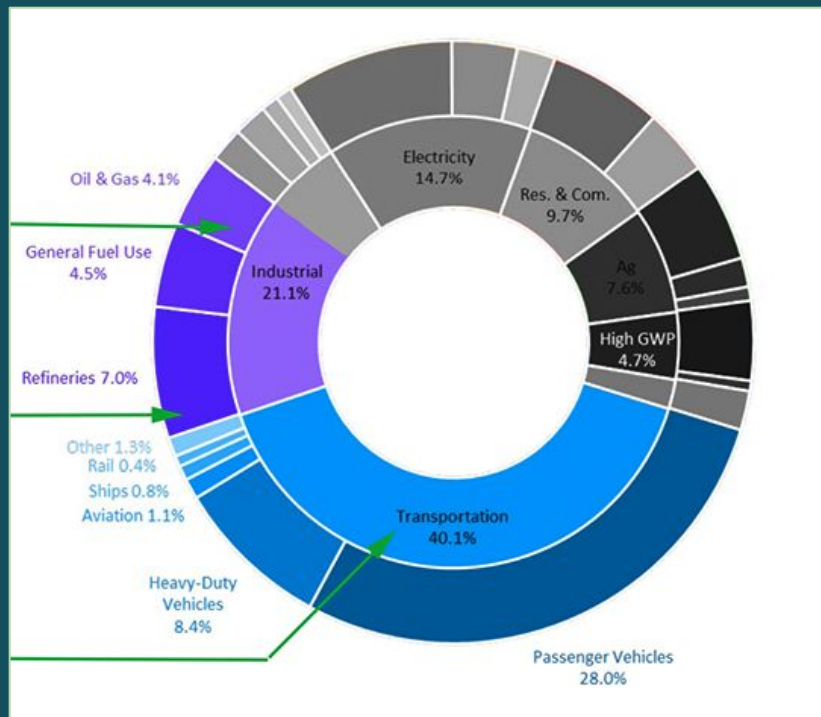


# Forecasting Transit Ridership

---

16 May 2022 | Capstone | GA DSIR 222

# MOTIVATION: Air Pollution Mitigation



Source: CA State Senate

- Passenger vehicles account for 28% (highest proportion) of emissions in CA
  - NOT diesel trucks
  - NOT refineries
  - NOT industrial processes
- Designing reliable mass transit systems that work for more addresses #1 source of emissions

Which *frameworks* in **Python** generate best models to forecast BART Transit Ridership?



FACEBOOK (2017): Prophet

2021 May: LINKED IN : Greykite,  
Silverkite (38s)

ARIMA: in Python, via Box and  
???

# EDA: BART Ridership

BART Monthly Ridership, 2010 - 2022



# Can additional regressors improve forecasts?

Monthly Average Gas Price (\$), California: 2010 - 2022



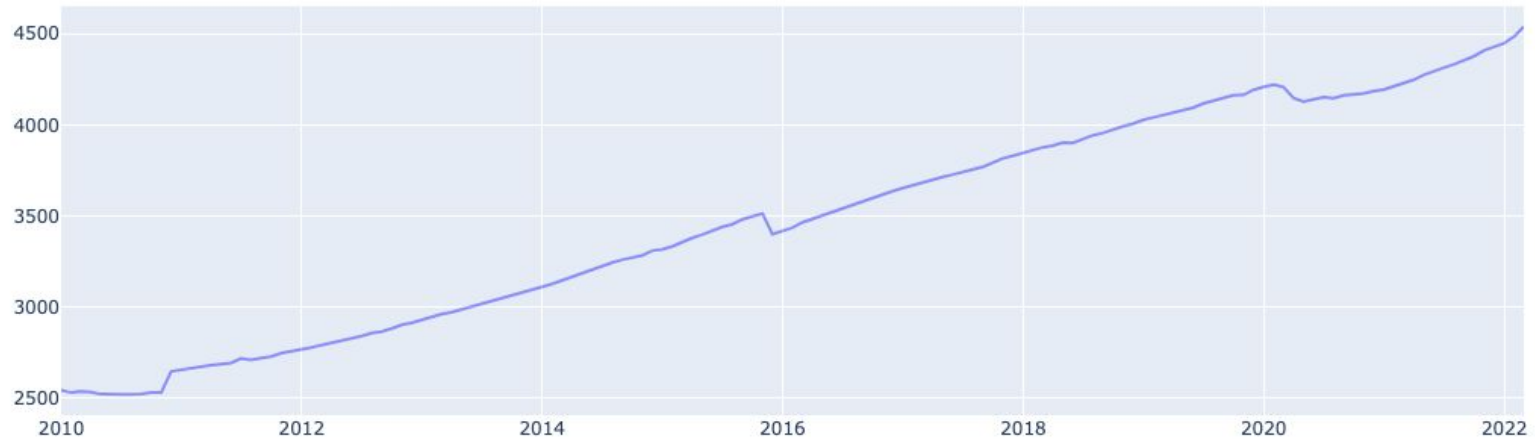
# Can additional regressors improve forecasts?

Estimated Count of Registered Cars CA: 2010 - 2021



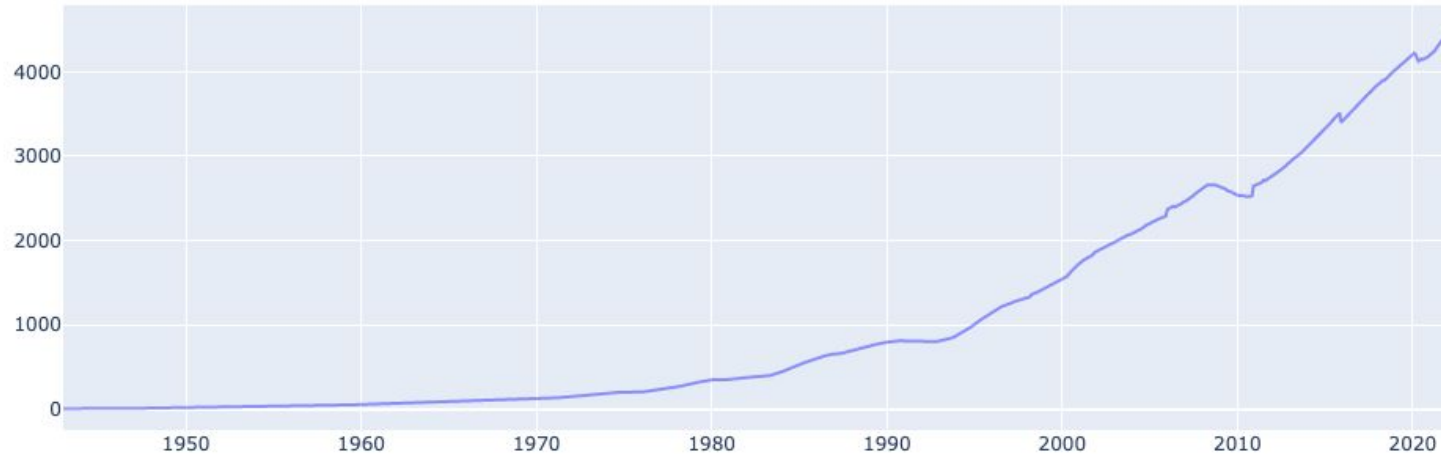
# Can additional regressors improve forecasts?

Consumer Debt (\$) 2010 - 2022 (not adjusted, Federal Reserve)



# Can additional regressors improve forecasts?

Consumer Debt(\$) 1943 - 2022 (not adjusted, Federal Reserve)



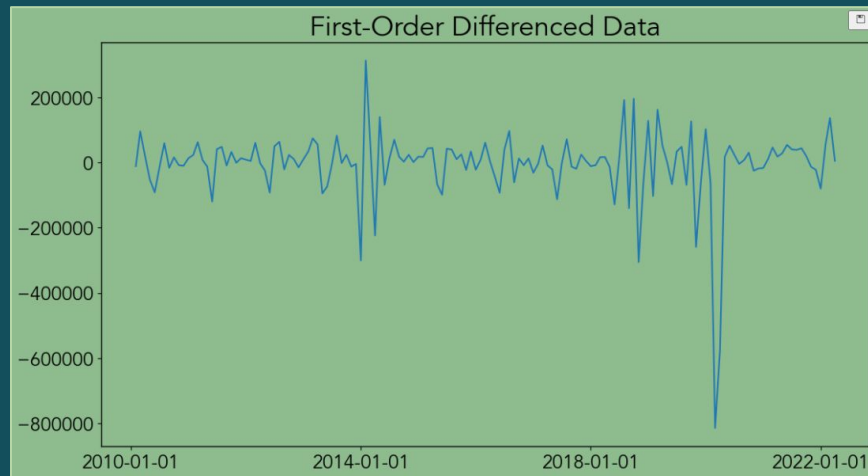


# Data Sources via API

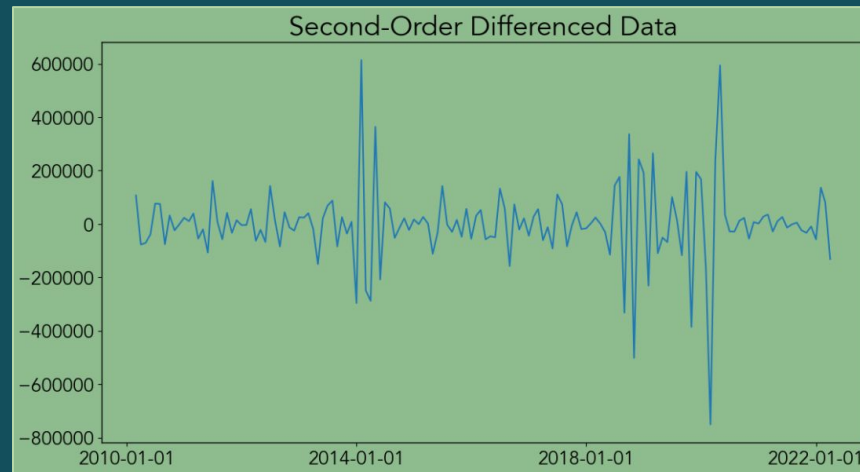


- BART: API + XLS
- EIA: weekly, monthly - API key required
- CA DMV: API + external data via e-mail
- Federal Reserve: consumer debt

# ARIMA $p = 0.79$

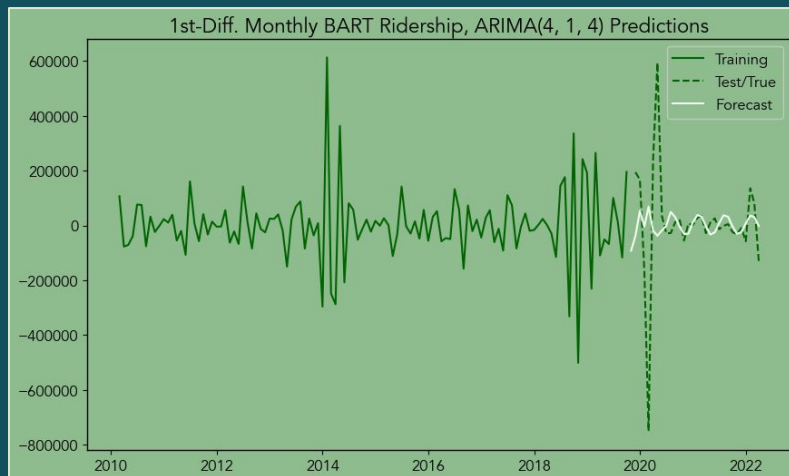


$p < 0.05$



$p < 0.05$

# ARIMA (4,1,4)



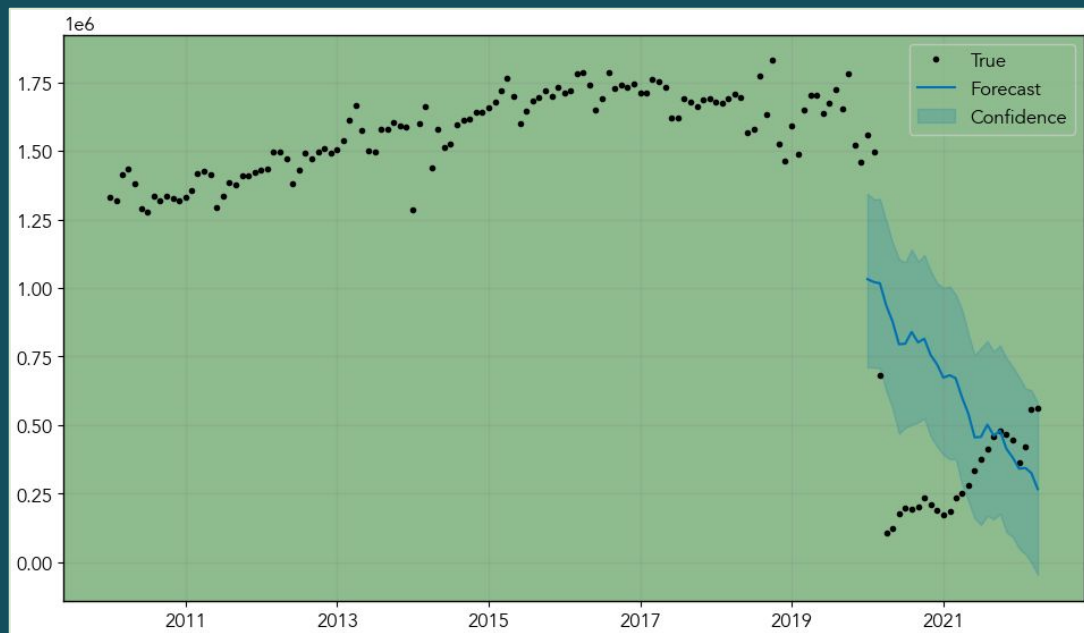
MAE: 106072

MSE: 39,094,230,679

RMSE: 197,722

AIC: 2913

# Prophet: Transit Ridership, Data 2010-2022

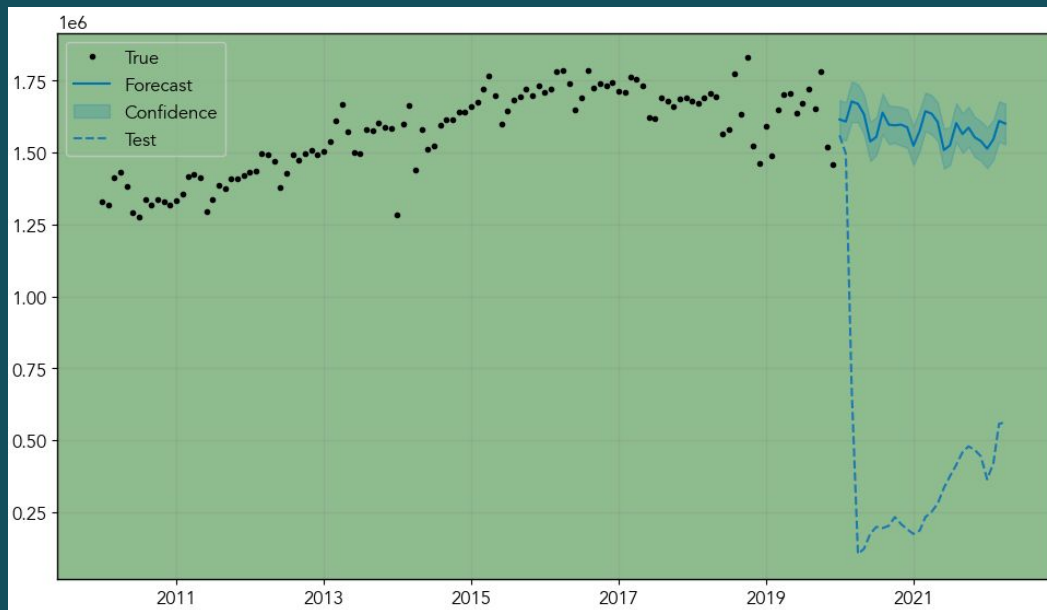


MAE: 1,181,450

MSE: 5 E11

RMSE: 1,243,269

# Prophet: Transit Ridership, Data 2010-2022

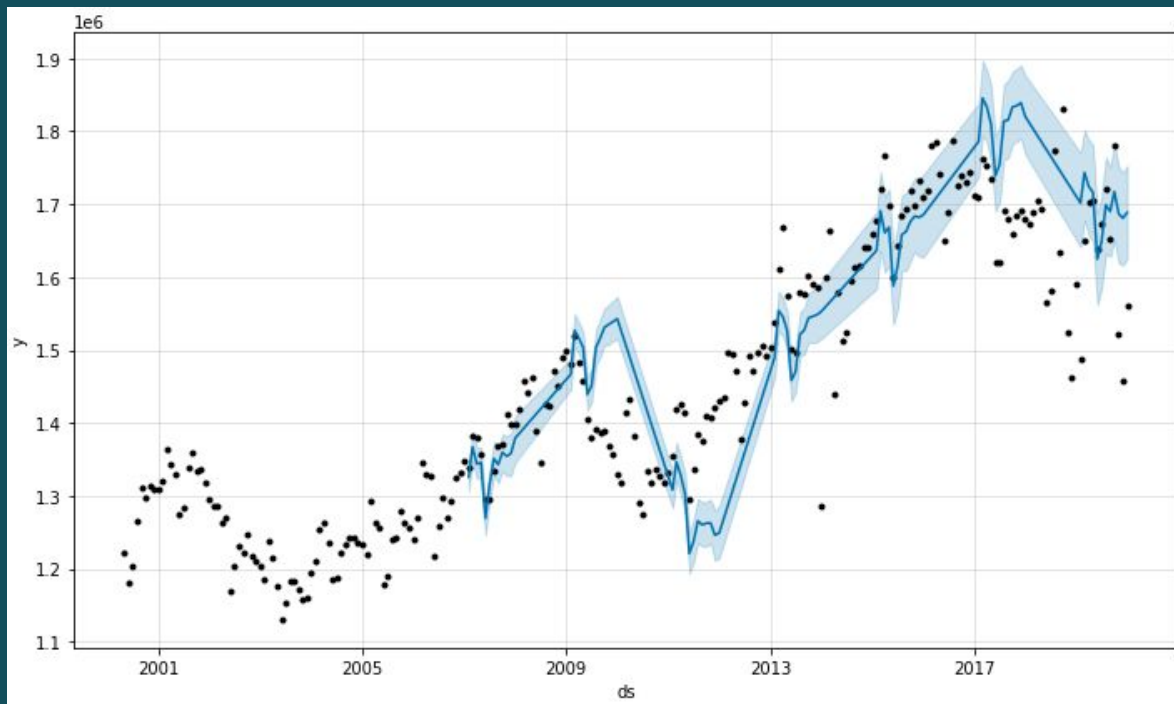


MAE: 375,686

MSE: 2 E13

RMSE: 1,243,269

## [2] Prophet: 2019-2022, CV



MAE: 1,181,450

MSE: 17294867277

RMSE: 131509

# Greykite: All Data: BAD



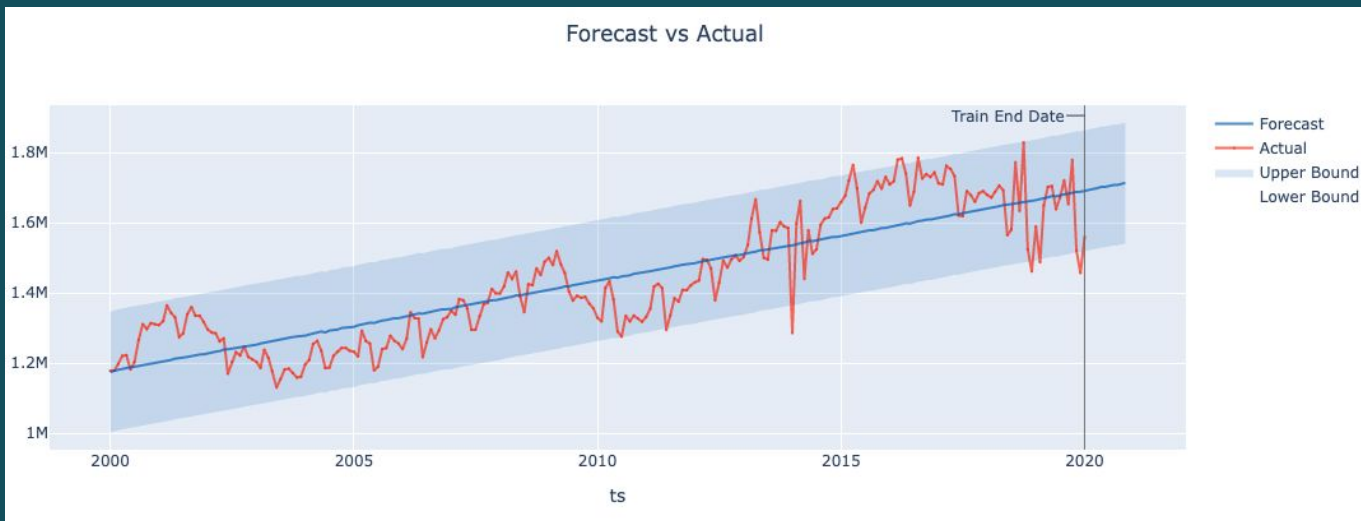
RMSE 446 1252

MSE 830 729 769 011

MAE 909 234

MAPE 205

# Greykite : 2020



RMSE 111,544

MSE 830 729 769 011

MAE 12,442,071,793

MAPE 5.25



# Greykite : Fuel

Forecast vs Actual



RMSE 111,544

MSE 830 729 769 011

MAE 12,442,071,793

MAPE 5.25

# Data Limitations



## yearly

- consumer debt
- vehicle counts
- *caltrain*

## monthly

- BART

## daily

- gas

# CONCLUSIONS



- Agencies must collect, or share better data
- Prophet, Greykite super easy to deploy, more difficult to customize
- Greykite great default graphics