

# The Road to CLEAN

THE IMPACT ON CALIFORNIA'S NATURAL GAS EMISSION

Presented By:

Rose Strauss, Elissa Cheung, Lauren Wiebe  
& Robelio Medina-Vasquez



# AGENDA

- 01. Topic
- 02. Background
- 03. Sources
- 04. Tools
- 05. The Need for Supply
- 06. The Hunt For Demand
- 07. Machine Learning Model
- 08. Process and Results
- 09. Conclusion
- 10. Thought for the Future



# THE TOPIC

- With about 50% of greenhouse emission coming from Transportation coming from vehicles on the road in California
- The state has been moving toward Zero Emission Vehicles since 2012. This push will create a dynamic change in the energy infrastructure and strain to the current PowerGrid.



# BACKGROUND

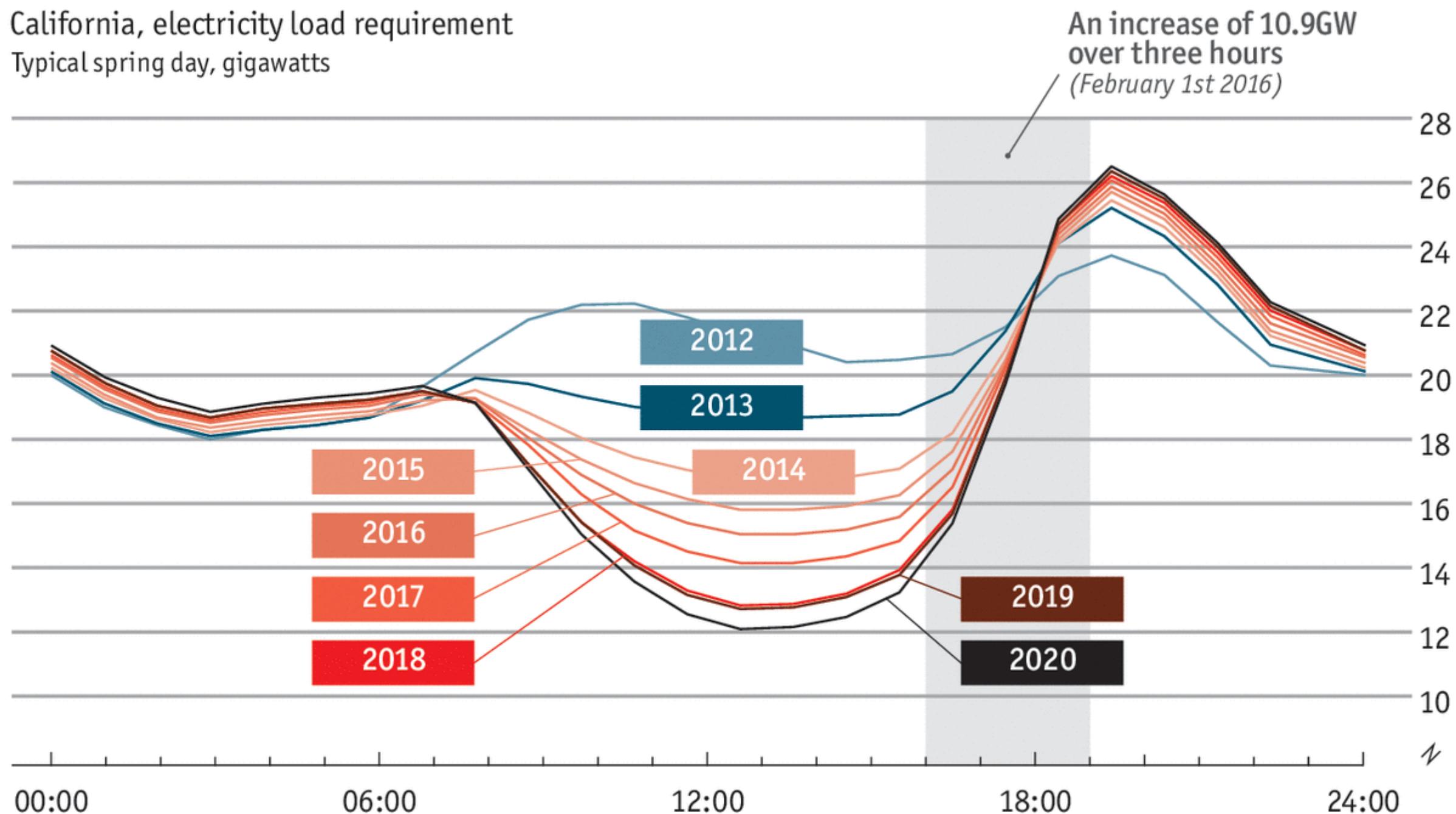
- Vehicle transportation attributes to 41% of greenhouse emission in California.  
*-California Air Resource Board - GHG Emission Report 2000-2019*
- Natural Gas makes up for approximately 43% of the state's energy source.  
*-EIA US Energy Information Administration (July 2021)*
- By 2030, California aims to have 5 million Zero Emission Vehicles on the road and 250K electric vehicle charging stations.  
*- Executive Order B-48-18*
- California is requiring all new vehicles sold in the state to be zero emission by 2035 and all commercial vehicles by zero emission by 2045.  
*- Executive Order N-79-20*

# California's Duck Dilemma

## Who gets the bill?

California, electricity load requirement

Typical spring day, gigawatts



Source: California ISO

Economist.com

As solar provides more energy in California, the transition to natural gas energy for evening consumption create a surge and energy loss.



# THE QUESTION

How will the projected increase  
of electric vehicle usage in  
California effect the natural gas  
consumption?

# OUR SOURCES



CALIFORNIA  
ENERGY  
COMMISSION



INSIDE<sup>E</sup>EVs

# THE TOOLS

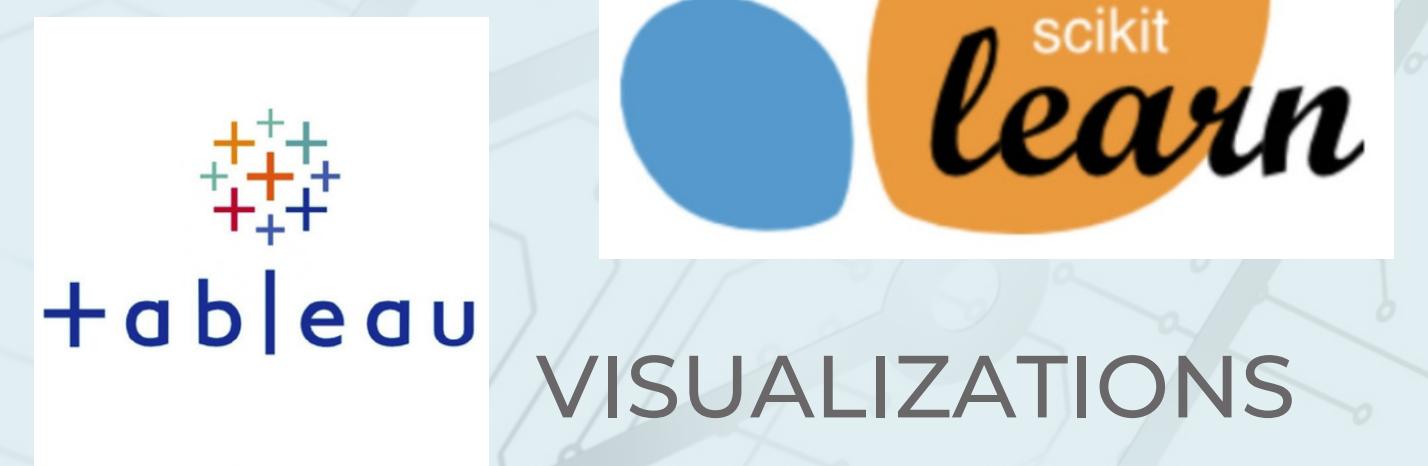
## 01 TECHNOLOGY



## 02 LANGUAGE



## 03 ALGORITHM



## 04 VISUALIZATIONS

# THE LIBRARIES

- Pandas
- Numpy
- Datetime
- Scikit-Learn
- OS
- Glob
- SQLAlchemy

# THE NEED FOR SUPPLY

## The RESEARCH:

- CAISO provided most accurate account of California's energy intake

## The DATA SET:

- Daily supply in megawatts (MW) for each 5 minute interval
- Broken down by energy source

## The DATA TRANSFORMATION:

- All cleaning and transforming in Pandas
- Size of end DF lead to storage in SQLite

Energy Source Usage Over Time



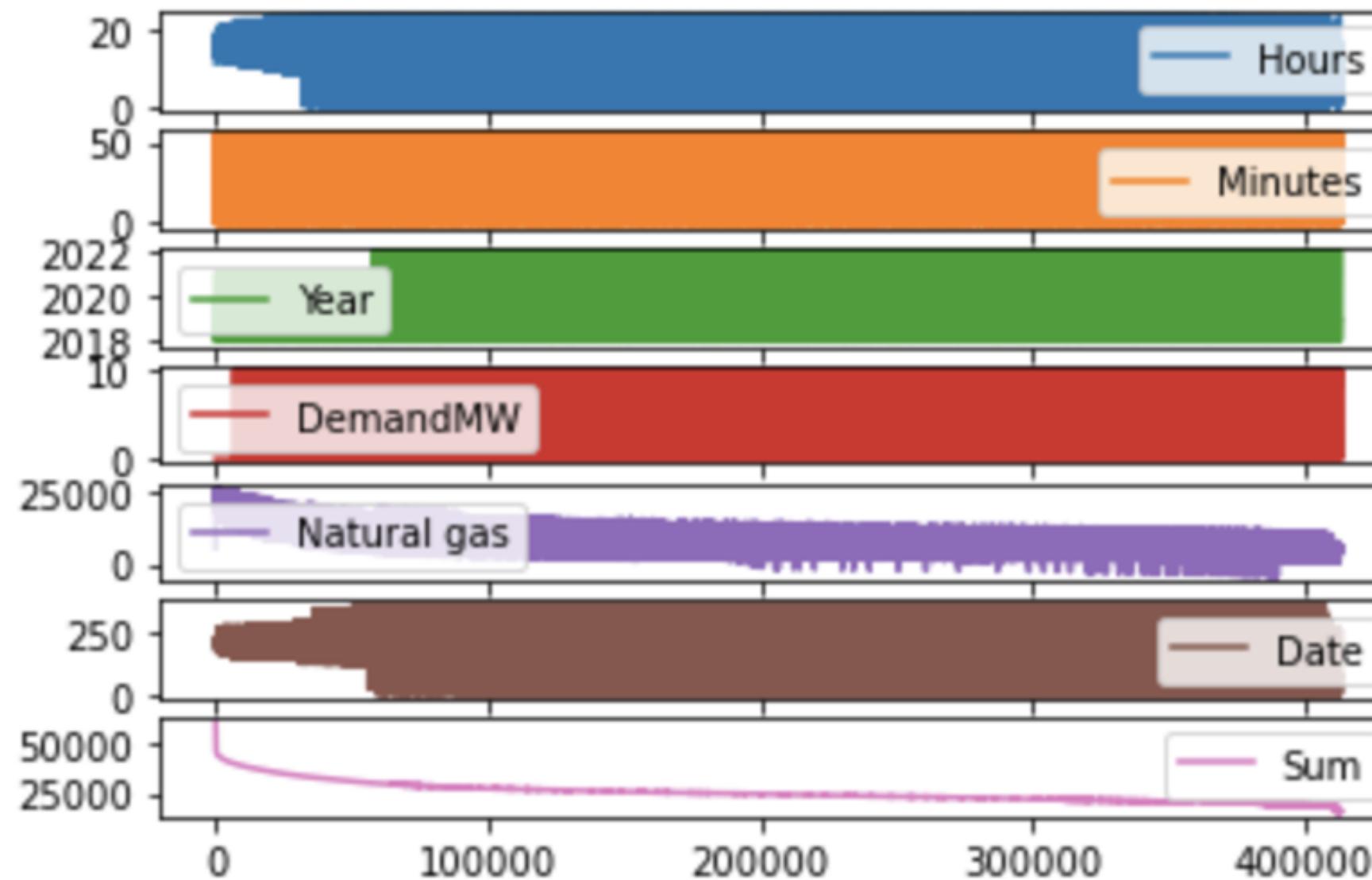
# THE HUNT FOR DEMAND

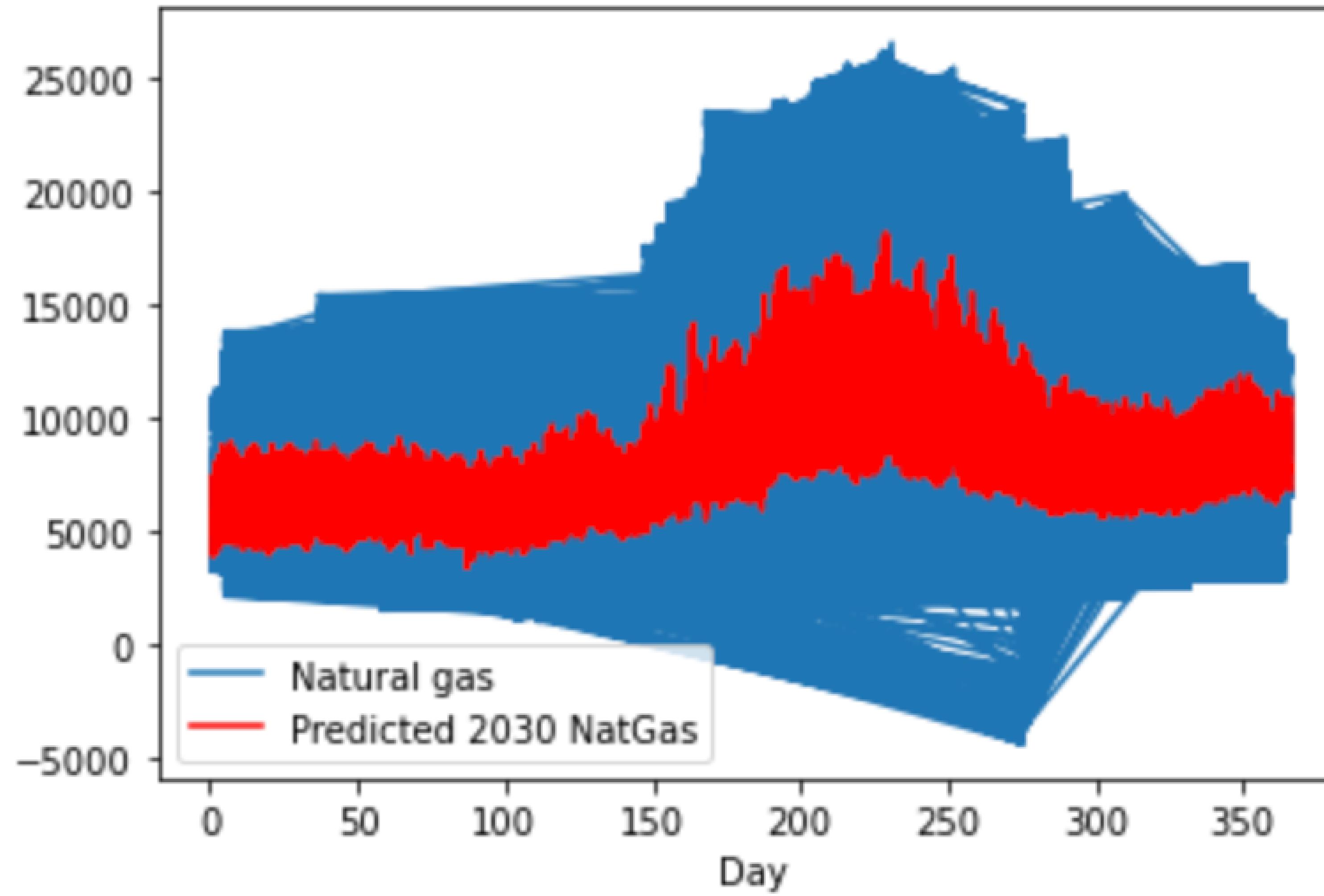
- Challenge: Searching for the demand of MW utilized by ZEVs currently does not exist.
- Assumptions: Creating an estimation on the variable made us select only certain factors for calculation.
- NBER - Based on the Working Papers we adopted their calculations and sample variables to utilize in our model.

# Machine Learning Model

- Multivariate Linear Regression
  - Limitations:
    - Limited to linear relationships
    - Only looks at the mean of the dependent variable
    - Sensitive to outliers
    - Data must be independent
  - Benefits:
    - Ability to determine the relative influence of one or more predictor variables to the criterion value
    - Ability to identify outliers

# MACHINE LEARNING DATA PRE-PROCESSING





# Our Conclusion

- Our Machine Learning Model stands inconclusive and with poor accuracy.
- Great Limitation in Data without proper time, planning and financing.
- High number of variables in the data that is being currently provided as open source - without specificity to ZEV energy consumption.

# **OUR RECOMMENDATION FOR THE SUCCESS FUTURE ANALYSIS**

- Planning for Accurate Data Collection
  - Possible sourcing and data collection
- Application for ZEV Data
  - Infrastructure Planning
  - Energy Storage System

## **WHAT IS HAPPENING TODAY...**