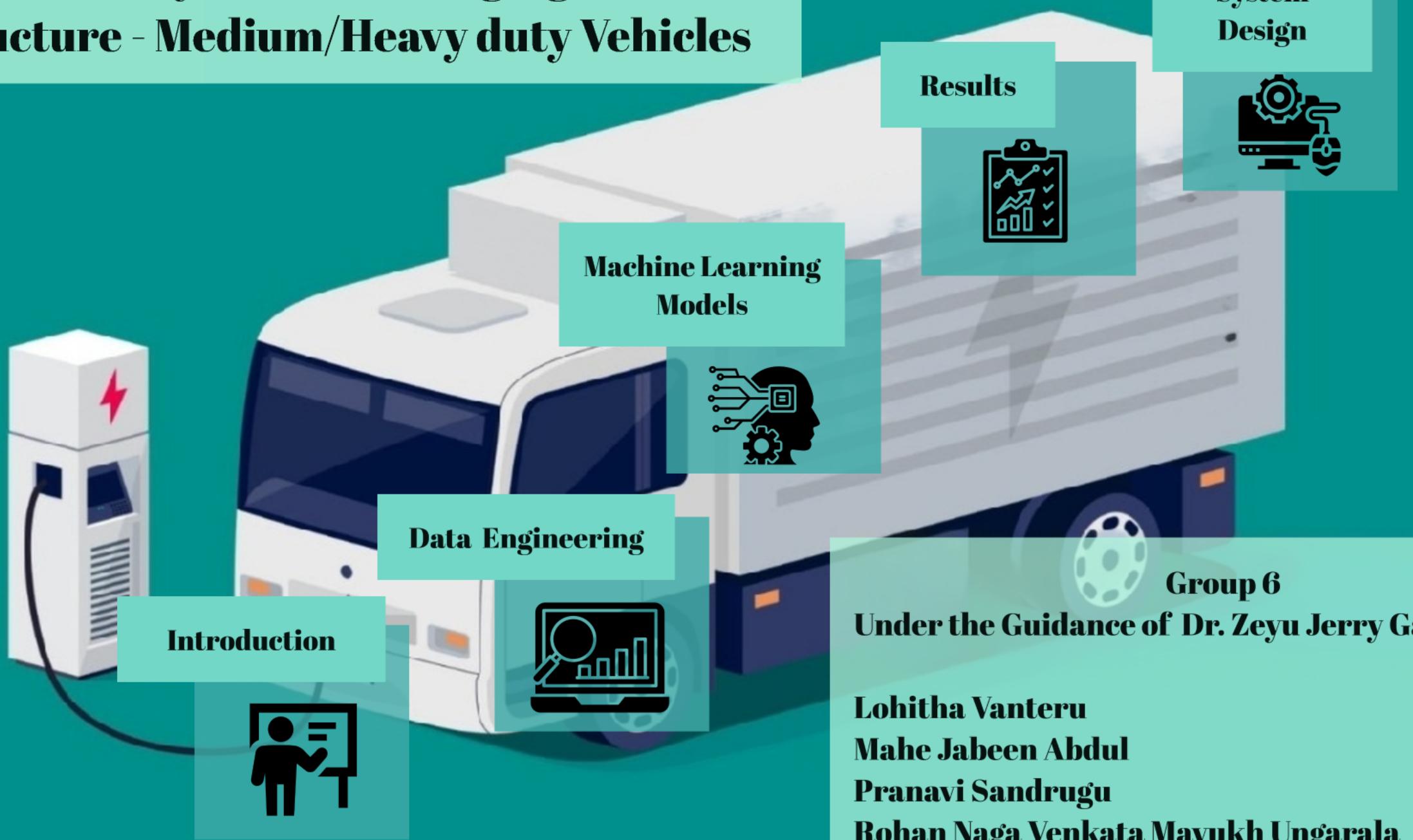


# Data Driven Analysis of EV Charging Infrastructure - Medium/Heavy duty Vehicles



## Setting the Stage:

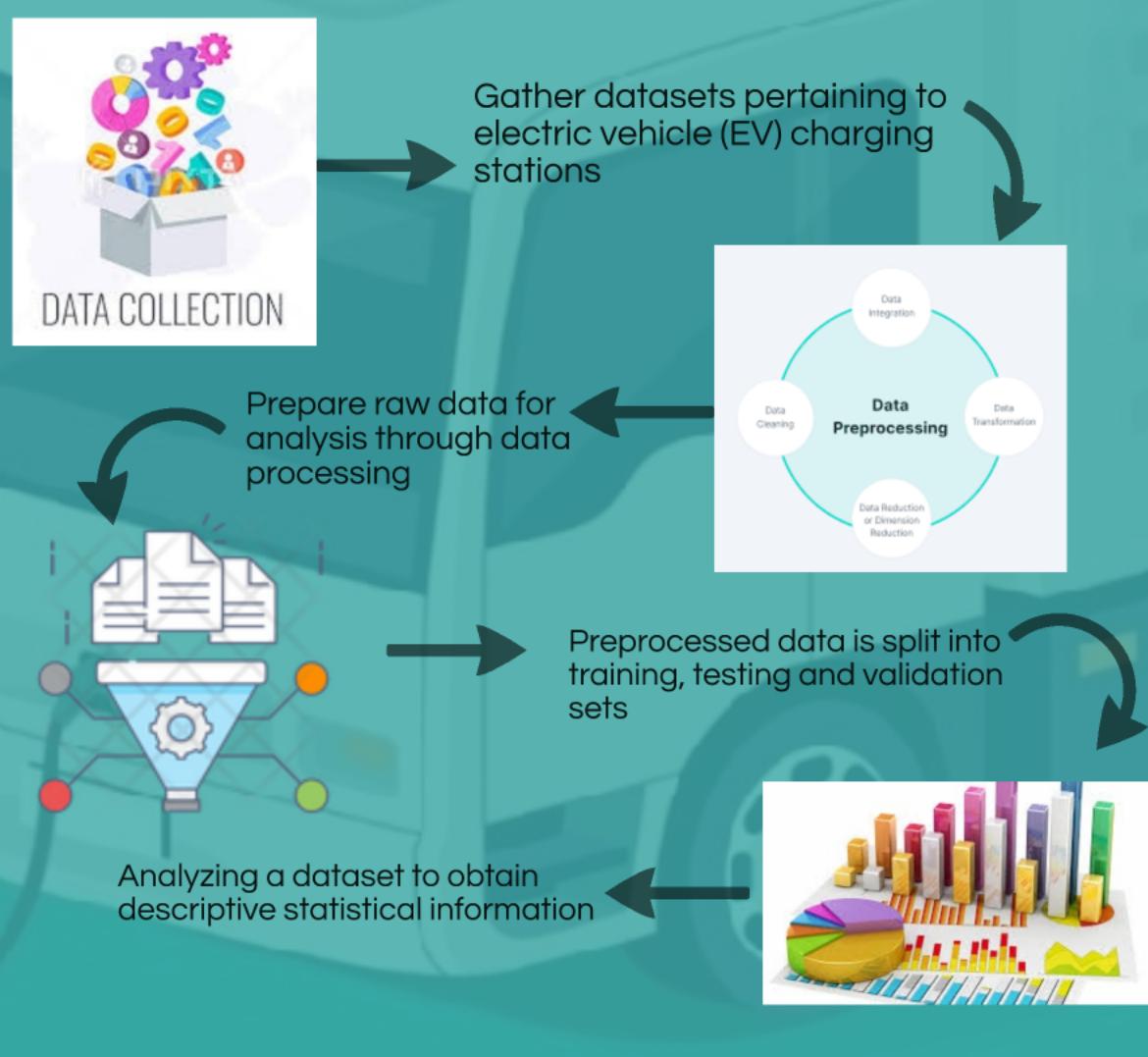
- MD/HD vehicles in the U.S. contribute to nearly one-third of the country's on-road transportation greenhouse gas emissions and are a major source of air pollution.
- The United States' Infrastructure Investment and Jobs Act (IIJA) allocates substantial funding, including \$5 billion for highway charging, \$2.5 billion for alternative fueling infrastructure, and \$5 billion for a Clean School Bus program, to advance electric vehicle infrastructure and sustainability initiatives.
- Heavy Duty Vehicle owners often face limited range due to the necessity of on-premises charging, and charging at regular passenger car stations is usually impractical due to infrastructure constraints and power requirements.

Targeted  
Problems

## **Addressing Key Challenges:**

- **Anticipating EV Market Growth:** *Forecasting the growing demand for EVs and the need for expanded charging infrastructure.*
- **Addressing Range Anxiety:** *Ensuring electric vehicle (EV) adoption by mitigating concerns related to limited driving range.*
- **Efficient Energy Resource Allocation:** *Optimizing energy resources by accurately forecasting short and long-term energy demand.*
- **Strategic Charging Infrastructure:** *Strategically placing charging stations to accommodate diverse vehicle types and urban mobility needs.*

# Data Engineering :Powering the EV Revolution



Data Collection

Data Preprocessing

Data Preparation

Data Statistics

# Data Collection

## Open Source Datasets :

- **CA Vehicle Registration Data**

*Number of vehicles registered each year for every zipcode and county in CA.*

- **CA EV Charging Stations Data**

*Real-time data of EV Charging stations for each zipcode in CA.*

- **Transit Bus Stops Data**

*Data of public bus stations locations for transit bus in San Jose.*

- **School Bus Terminals Data**

*Data of school bus terminal stations locations in San Jose.*

- **Heavy/Medium Vehicle Fleet Transactions Data**

*Real-world data of commercial fleet vehicle operating data for each weight class .*

Name	Vehicle Registration-Zip code	Charging Stations- Zip Code	Vehicle Fleet Transactions-Weight Class	Transit Bus Stations-City	School Bus Terminals-City
Source	California Department of Motor Vehicles	U.S. Alternate Fuel Data Center	National Renewable Energy Laboratory	CA Open Data Portal	CA Department of Education
Data Type	.xlsx	.csv	.xlsx	.geojson	.csv
Size	28.8 MB	16.7 MB	126.8 MB	1.6 MB	3.4 MB
Count	35,742	14,862	4,28,963	1289	3526
Yearly	Y	Y	Y	Y	Y
Weekly	N	N	Y	N	N
Monthly	N	N	Y	N	N
Duration	2010-2022	2010-2022	2018-2022	2022	2022

# Data Preprocessing

## Data cleaning:

- Replacing spaces with underscores in column names makes them more suitable for data analysis and avoids potential naming issues.
- Assigning appropriate data types to columns.

## Converting 'Year' to Datetime:

The 'Year' column is converted to datetime format using `pd.to_datetime`.

## Merging Data frames:

- Used the `merge` method to combine two data frames, `df` and `zc`, based on a common column, which is 'Zip\_Code'.
- Dropped the unnecessary columns, rearranges the columns in the resulting data frame to create a `vehicle_data` data frame.

## Data Aggregation:

Aggregated historical vehicle count data by grouping it based on the 'Year' column, summing the 'Vehicles' column for each year. This step organizes the data for analysis.

## Handling missing values:

- Removed rows from the Data Frame `charging_data` that contain missing (null) values.
- `SimpleImputer` is used to handle missing values in the resampled data. Missing values are filled with the mean value of the respective column using the `fit_transform` and `transform` methods for daily, weekly, and monthly forecasts, respectively.

## Data Transformation:

The new variables such as 'Charging\_Duration,' 'Average\_Power,' and 'Delta\_SOC' are created and added to the existing data to better represent and summarize aspects of the data.

## Data Preparation:

- Created `fit_and_forecast` function with the `yearly_vehicle_count` Data Frame and specified parameters for interval width.
- This prepares a Prophet model, fits it to the data, and forecasts EV vehicle counts until 2045.
- Feature engineering done calculating 'Delta\_SOC' and 'Charging\_Duration'), feature selection, and defining the target variable ('Range').
- The dataset is split into training and testing sets, preparing the data for subsequent machine learning modeling and evaluation.

## Raw Data Statistics

Dataset	Vehicle Registration-Zip code	Charging Stations- Zip Code	Vehicle Fleet Transactions-Weight Class	Transit Bus Stations-City	School Bus Terminals-City
<b>Data Type</b>	.xlsx	.csv	.xlsx	.geojson	.csv
<b>Size</b>	28.8 MB	16.7 MB	126.8 MB	1.6 MB	3.4 MB
<b>Count</b>	35,742	14,862	4,28,963	1289	3526
<b>Features</b>	15	65	48	12	16
<b>Duplicate Rows</b>	0	0	96	0	0
<b>Missing Values</b>	29	27	258	0	0

## Processed Data Statistics

Dataset	Vehicle Registration-Zip code	Charging Stations- Zip Code	Vehicle Fleet Transactions-Weight Class	Transit Bus Stations-City	School Bus Terminals-City
<b>Data Type</b>	.csv	.csv	.csv	.csv	.csv
<b>Size</b>	14.2 MB	8.9 MB	87.1 MB	89 KB	1.4 MB
<b>Count</b>	35,713	14,835	4,28,609	1289	3526
<b>Features</b>	7	15	14	5	4
<b>Duplicate Rows</b>	0	0	0	0	0
<b>Missing Values</b>	0	0	0	0	0

# Data Preparation

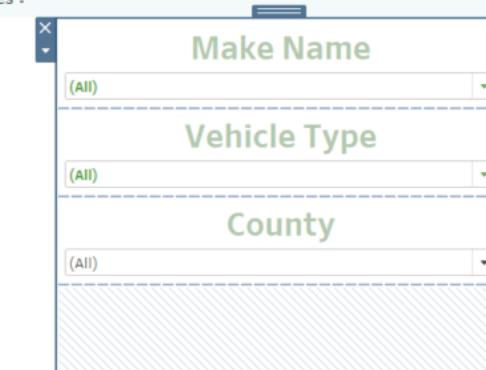
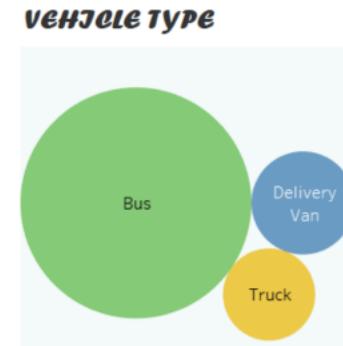
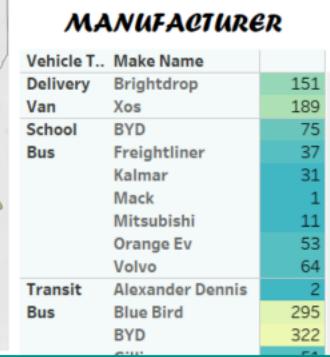
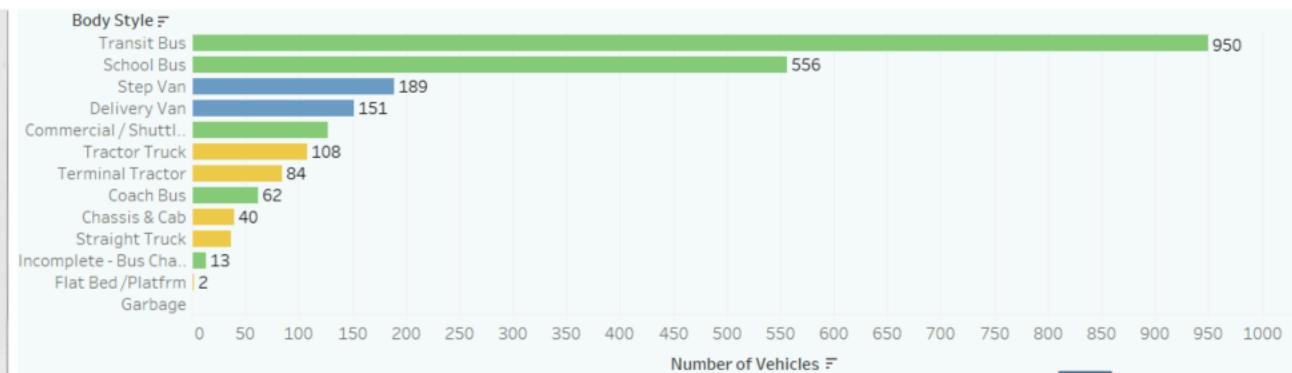
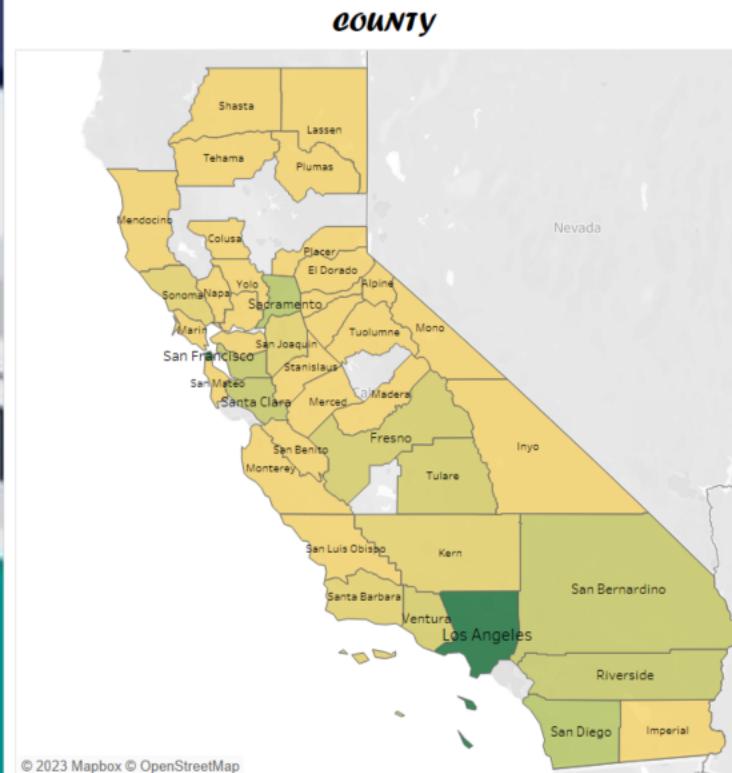
Dataset/Task	Raw	Preprocessed	Prepared			
			Total Rows	Train Rows	Test Rows	Validation Rows
CA Vehicle Registration Data	35,742	35713	35713	28571	3571	3571
CA EV Charging Station Data	14862	14835	14835	11869	1483	1483
Vehicle Fleet Transactions Data	4,28,963	4,28,609	4,28,609	3,42,887	42861	42861

*All the Datasets are split into 80:10:10 ratio*

# Data Statistics

## 2022 Medium & Heavy Duty ZEV Population

2,320



## Where are the Charging Stations ?

### Facility Wise EV Charging Stations Count

Others  
DC Charger Count : 4,833  
Level 2 Charger Count : 24,917

Shopping Centers  
DC Charger Count : 2,448  
Level 2 Charger Count : 177

Parking Places  
DC Charger Count : 1,007  
Level 2 Charger Count : 1,157

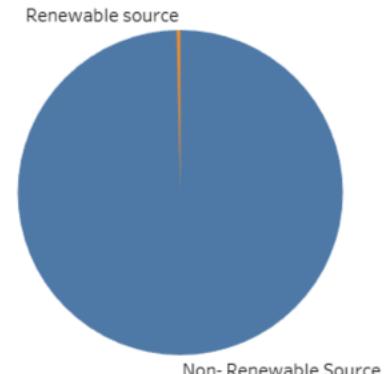
Recreation Services  
DC Charger Count : 956  
Level 2 Charger Count : 2,542

Car

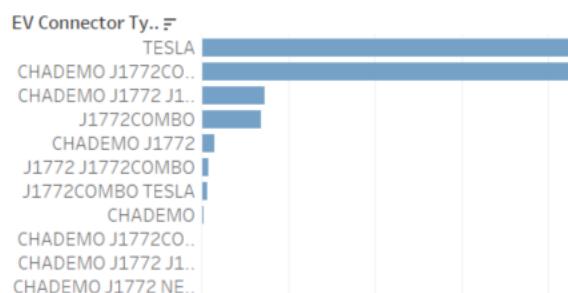
Facility Type

(All)

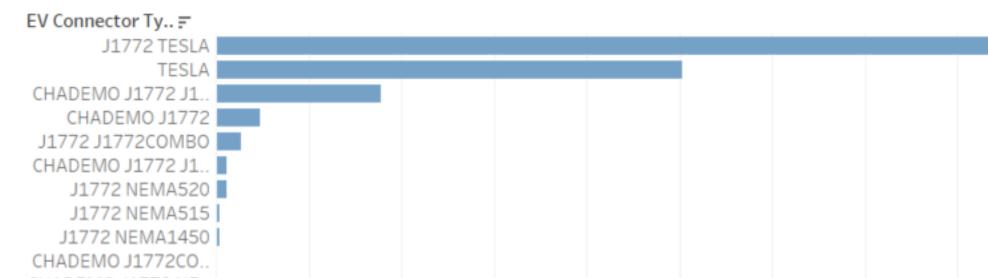
### Energy Sources of Charging Stations

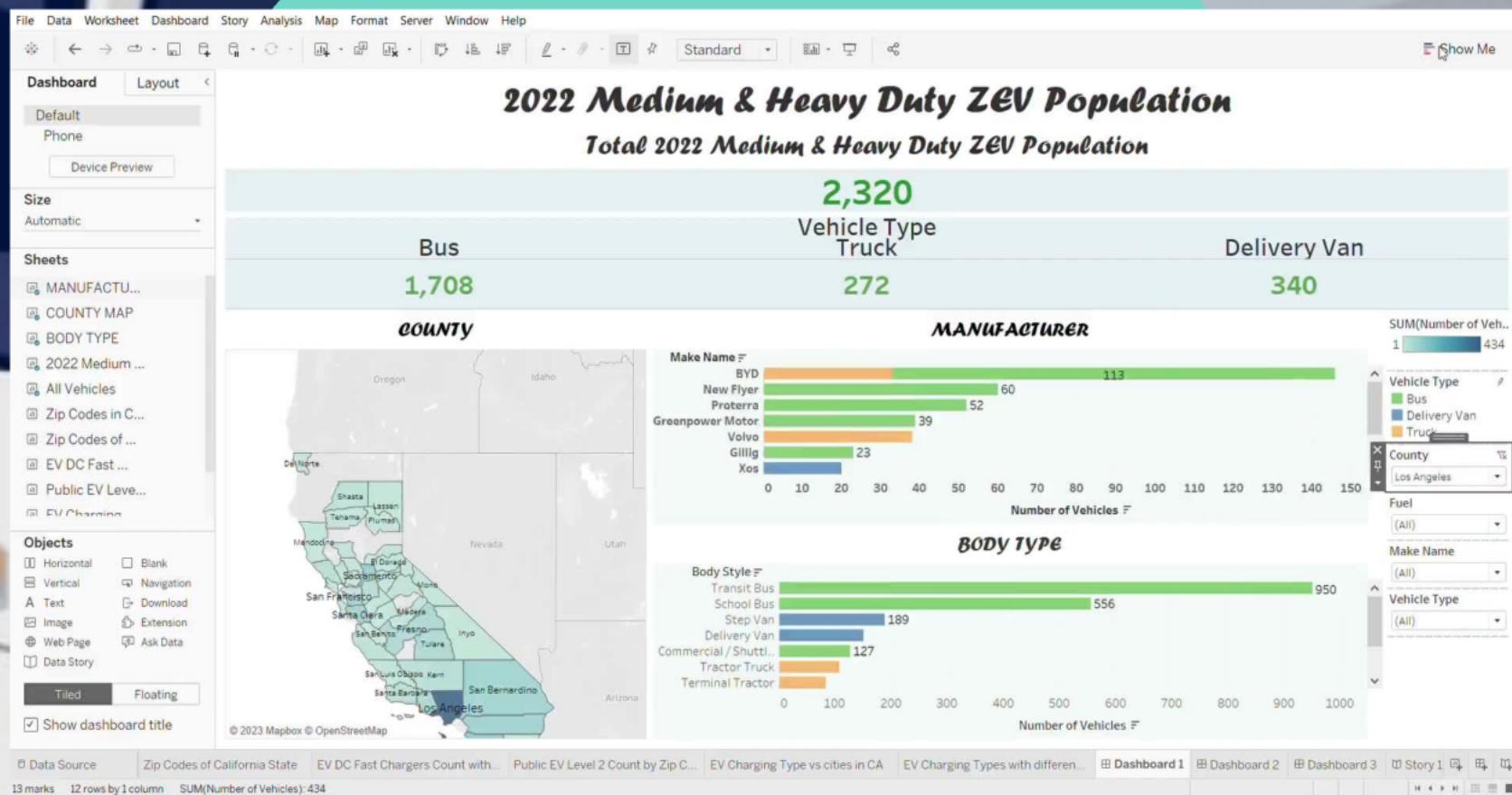


### DC Charger Connector Types



### Level 2 Charger Connector Types





Dashboard

Layout

Default

Phone

Device Preview

Size

Automatic

Sheets

MANUFACTU...

COUNTY MAP

BODY TYPE

2022 Medium ...

All Vehicles

Zip Codes in C...

Zip Codes of ...

EV DC Fast ...

Public EV Leve...

EV Charging

Objects

Horizontal

Blank

Vertical

Navigation

Text

Download

Image

Extension

Web Page

Ask Data

Data Story

Tiled

Floating

 Show dashboard title

# 2022 Medium & Heavy Duty ZEV Population

## Total 2022 Medium & Heavy Duty ZEV Population

2,320

Vehicle Type  
Truck

Bus

1,708

Delivery Van

272

340



Make Name

BYD	113
New Flyer	60
Proterra	52
Greenpower Motor	39
Volvo	38
Gillig	23
Xos	20

MANUFACTURER

Number of Vehicles

Body Style

Transit Bus	950
School Bus	556
Step Van	189
Delivery Van	165
Commercial / Shutt..	127
Tractor Truck	105
Terminal Tractor	95

BODY TYPE

Number of Vehicles

SUM(Number of Veh..)

1 434

Vehicle Type

Bus  
Delivery Van  
TruckCounty  
Los Angeles

Fuel

(All)

Make Name

(All)

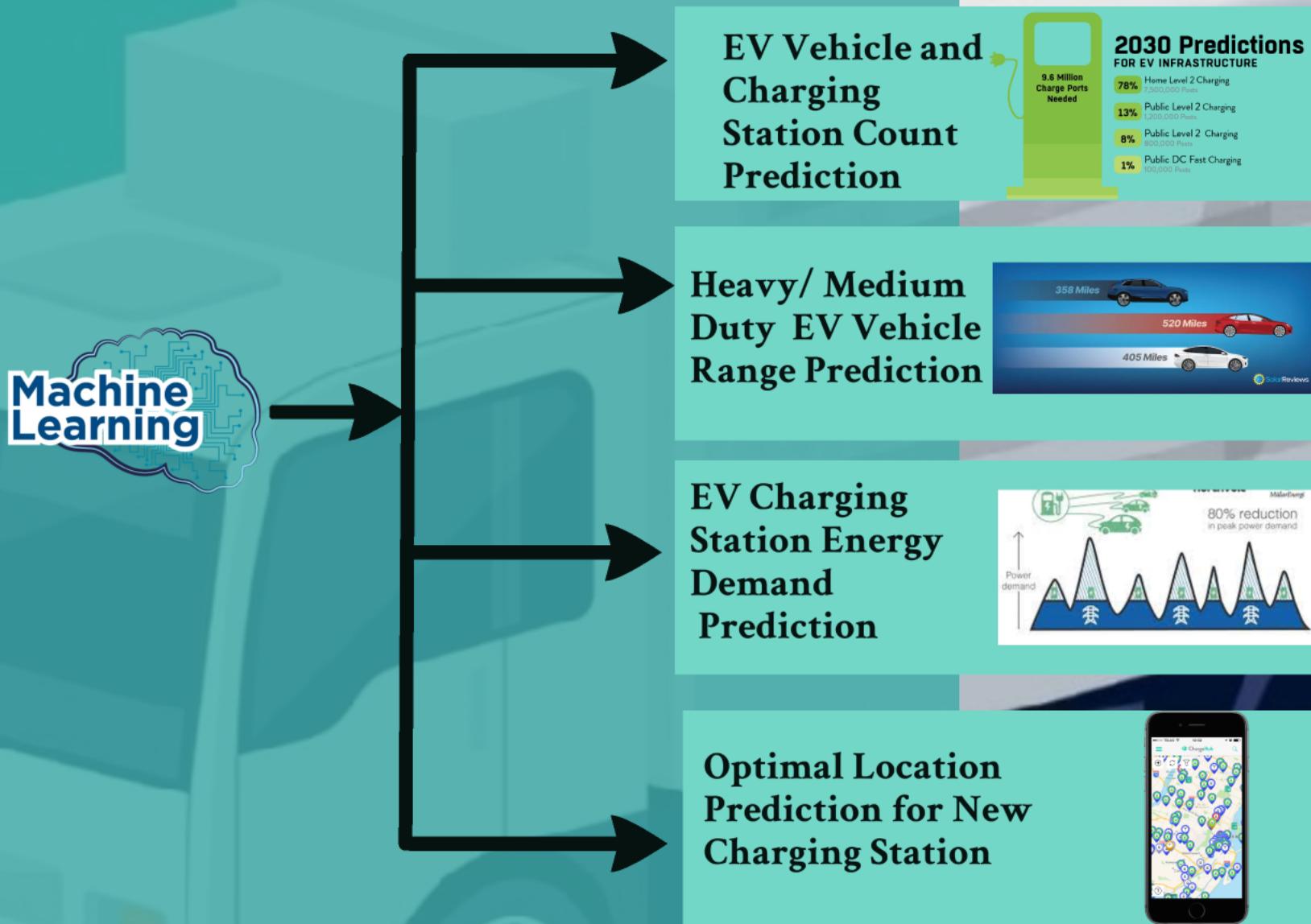
Vehicle Type

(All)

Story

1

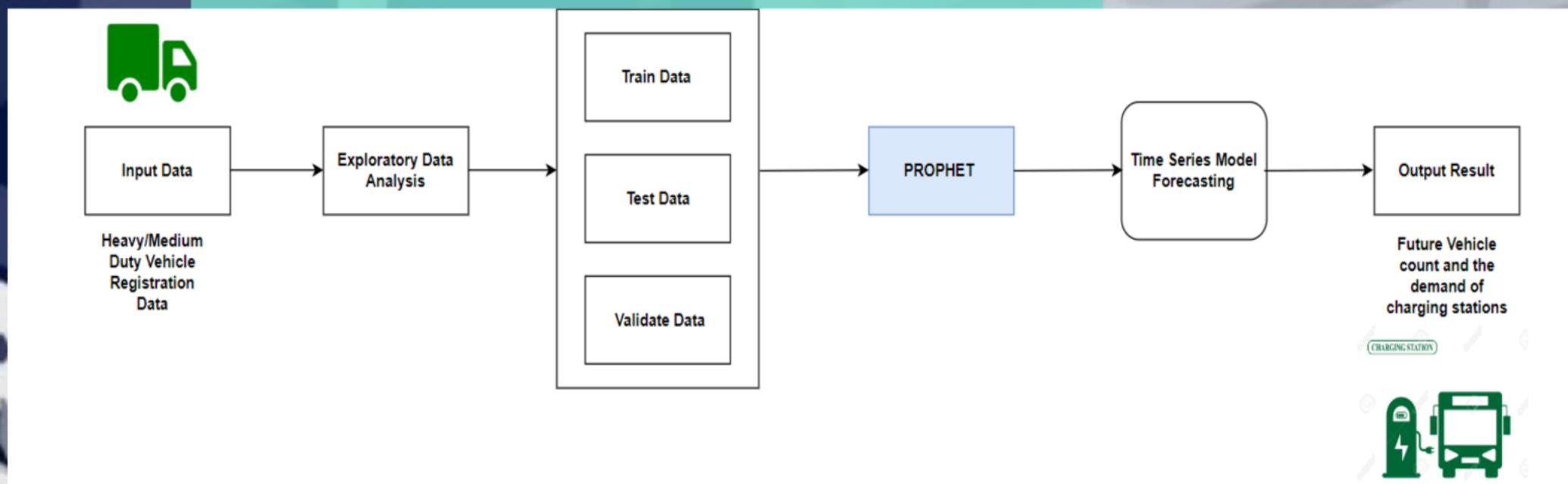
# Proposed Machine Learning Models



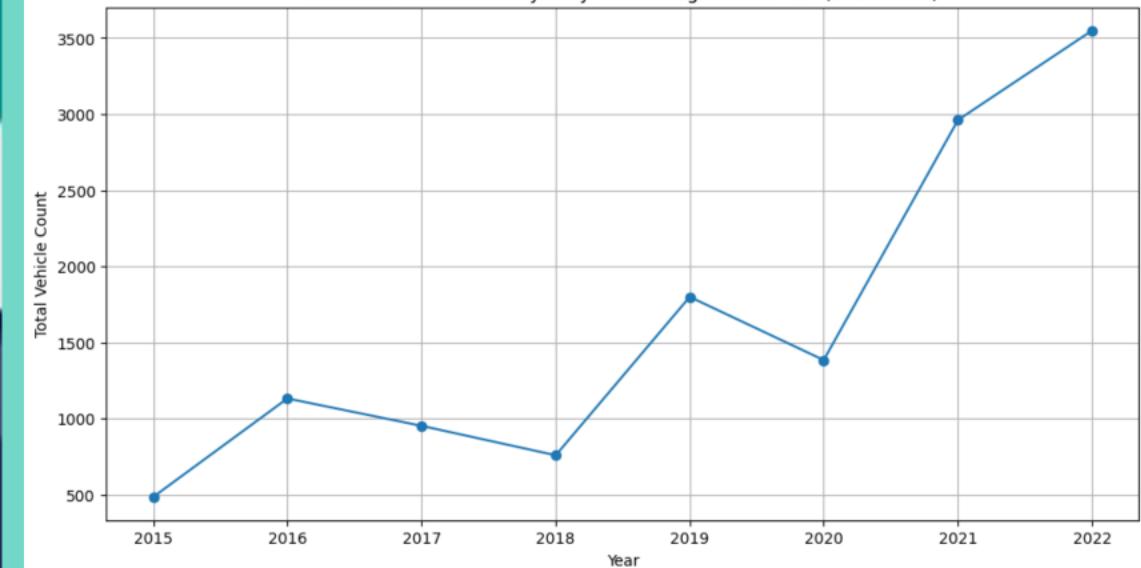
# FORECASTING HEAVY DUTY AND MEDIUM DUTY EV COUNT AND CHARGING STATIONS DEMAND

## PROPHET MODEL

*To Predict and forecast the future count of Heavy Duty and Medium Duty Vehicles and the required EV charging stations to cater the demand.*



Historical Count of Heavy Duty Vehicle registration - CA (2015-2022)



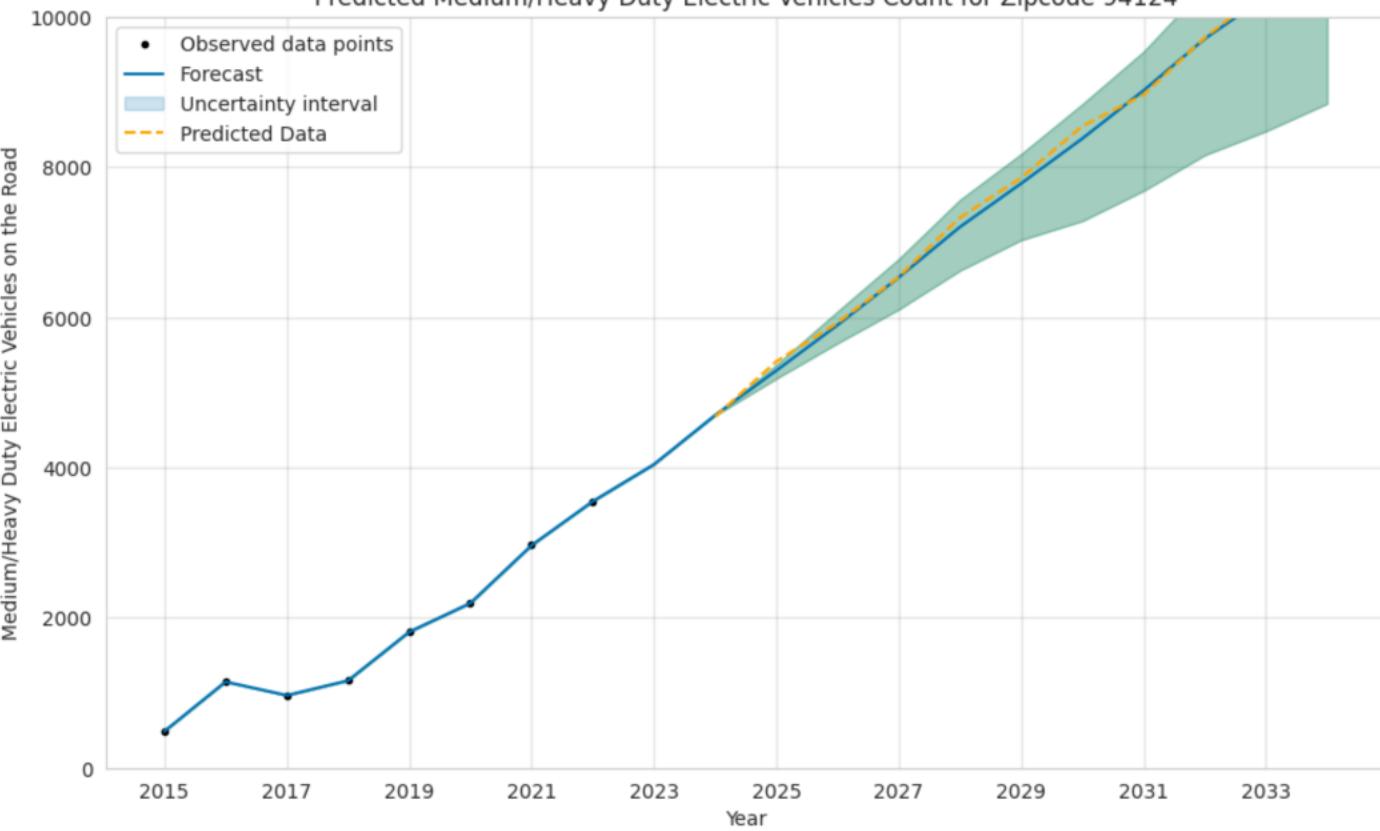
## Historical vehicle Count



## Forecasting Results for each Zip Code



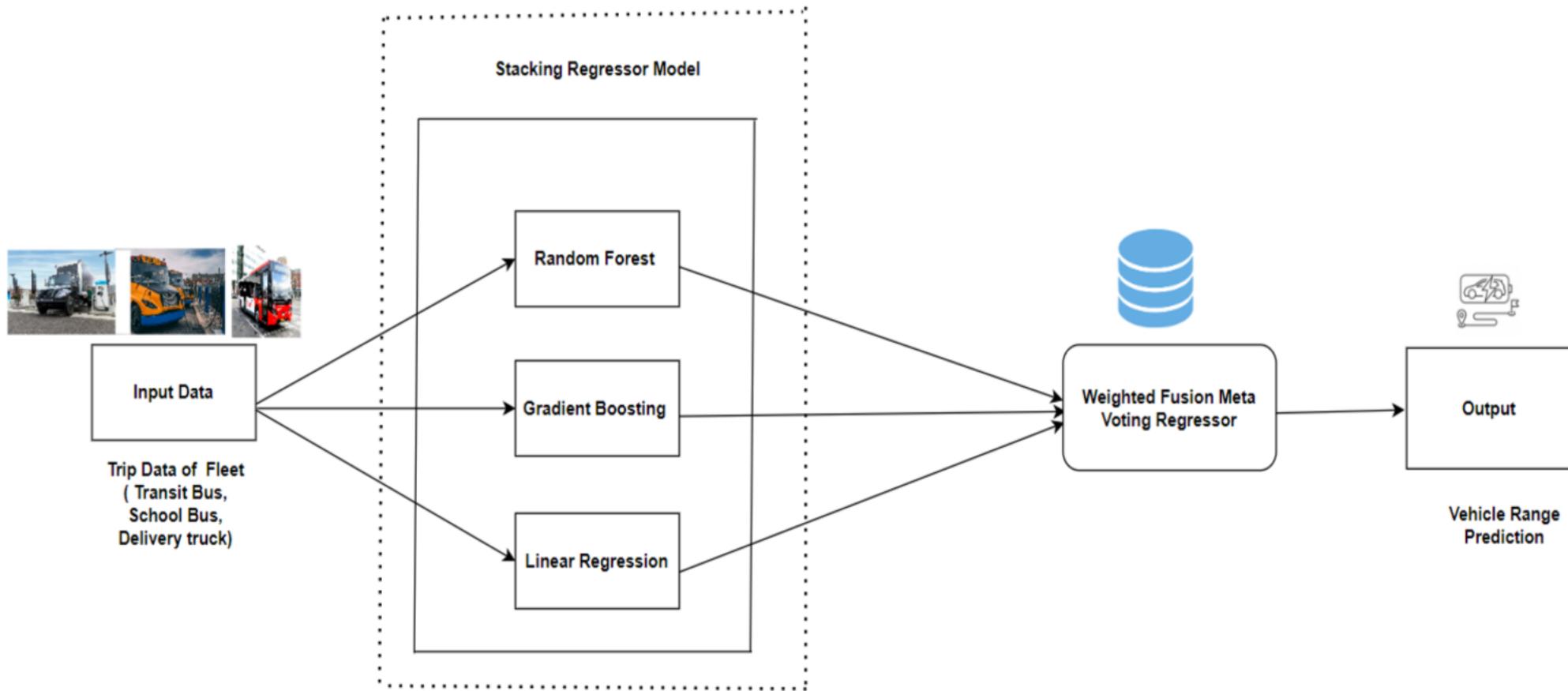
Predicted Medium/Heavy Duty Electric Vehicles Count for Zipcode 94124

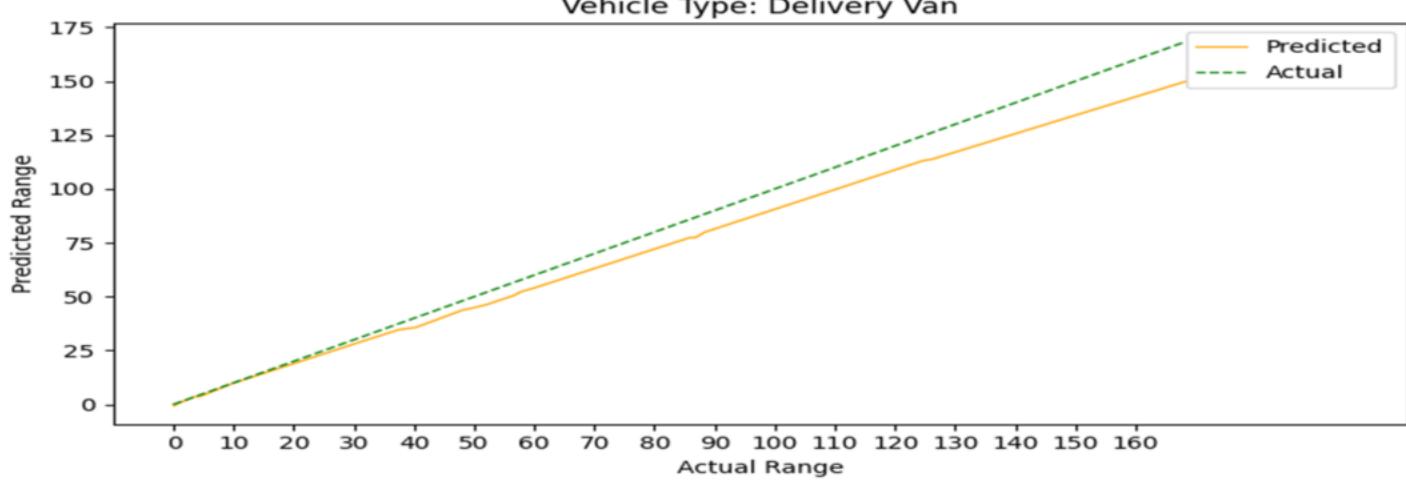
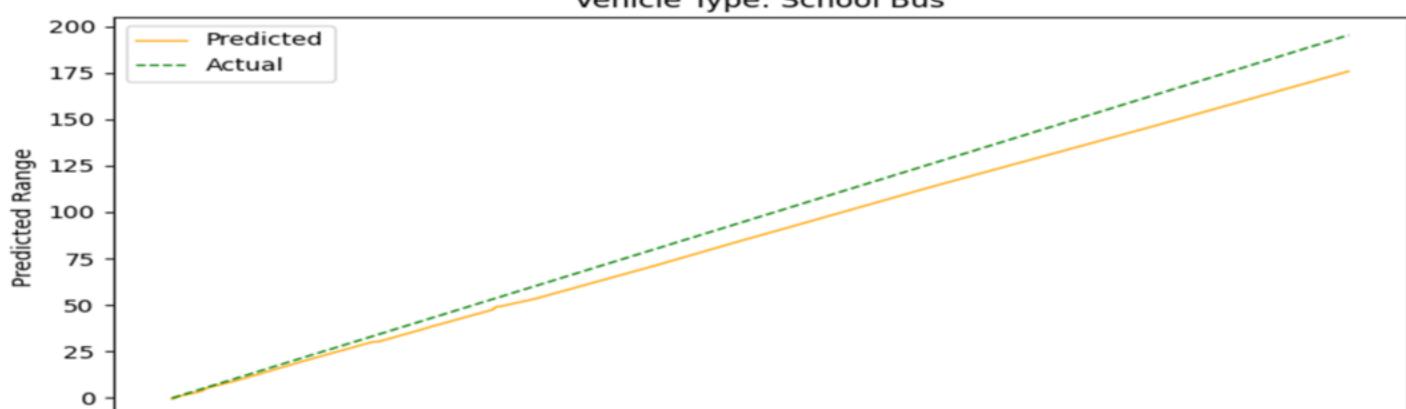
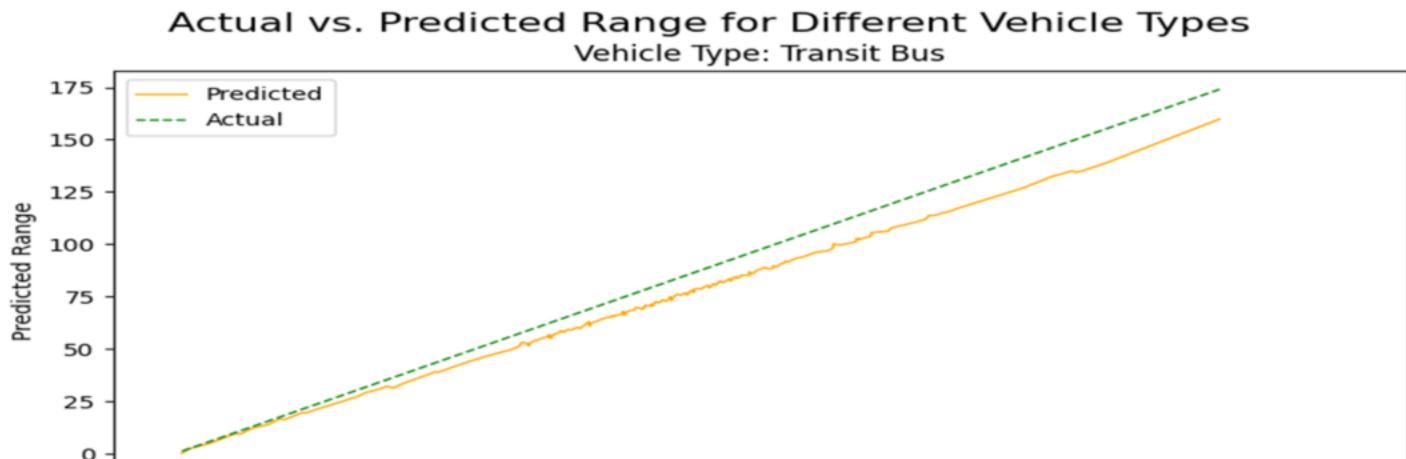


# PREDICTING VEHICLE RANGE OF HEAVY DUTY AND MEDIUM DUTY EV

## Stacking Regressor Model - Weighted Fusion Voting

To forecast range of heavy and medium duty vehicles of each weight class using the characteristics like charging duration, state of charge of the battery.

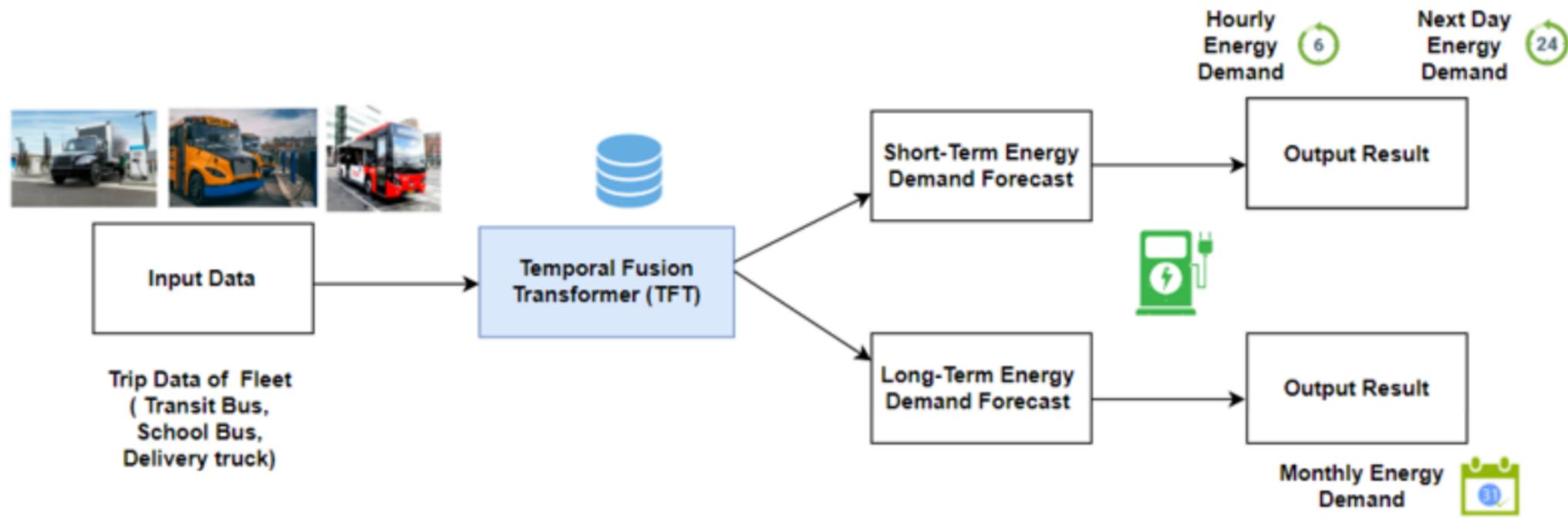




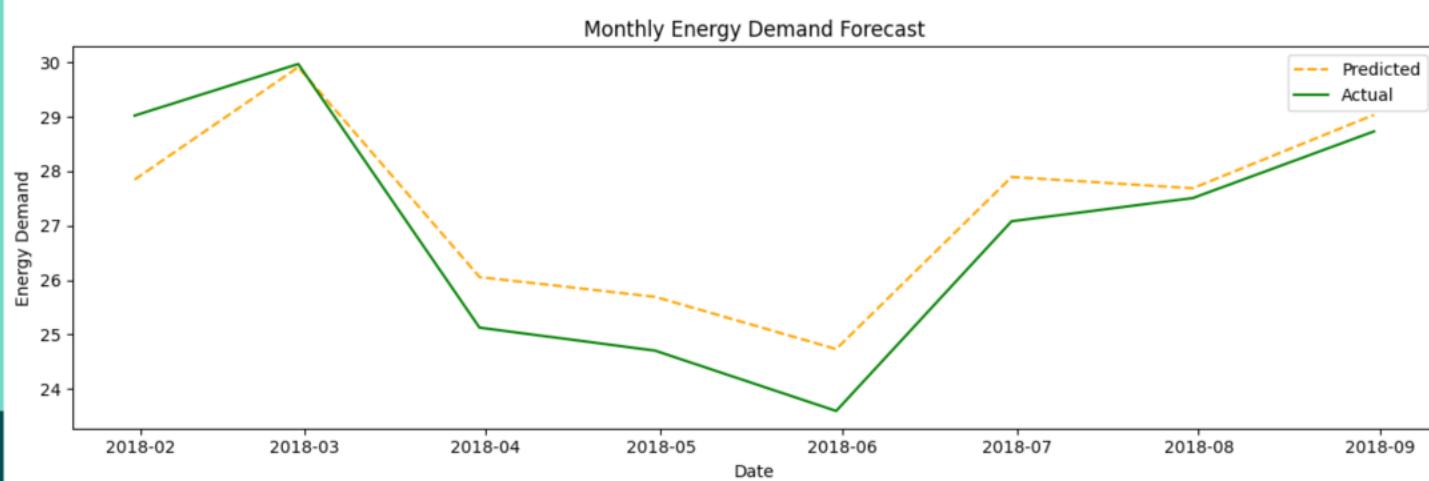
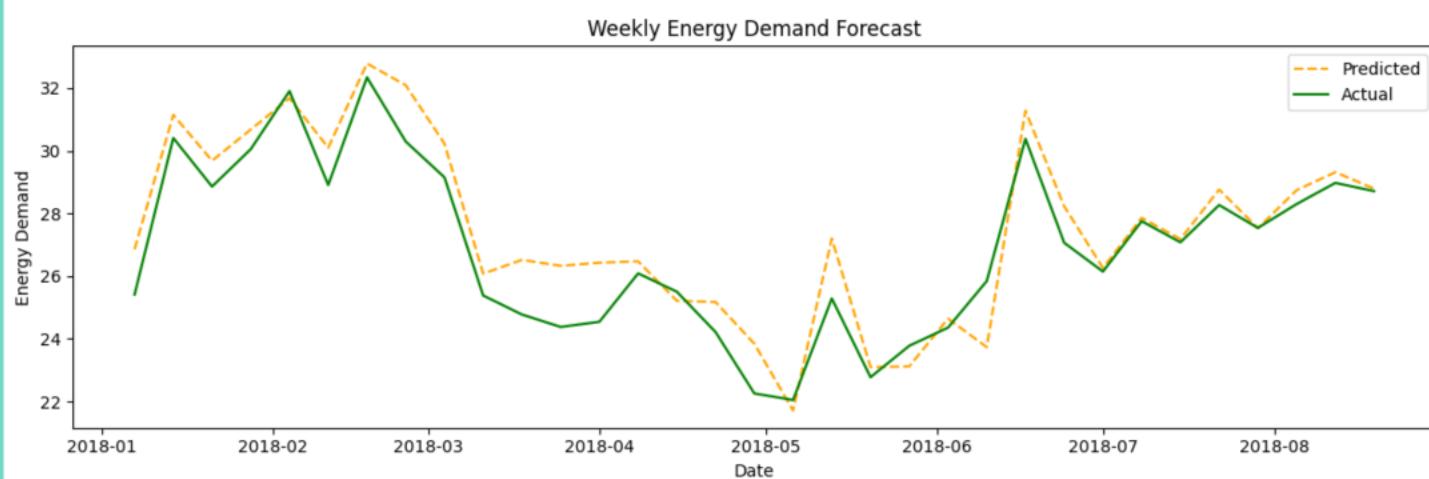
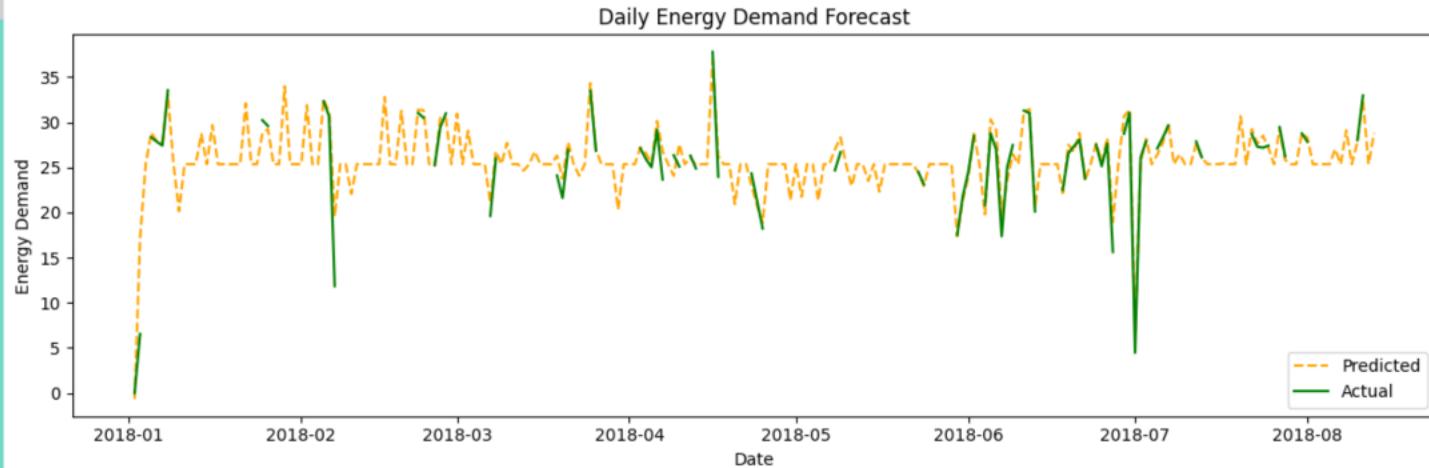
# PREDICTING SHORT AND LONG-TERM ENERGY DEMAND

## Temporal Fusion Transformer Model (TFT)

To Predict short term energy demand like hourly and long term energy demand like weekly, monthly for each charging station.



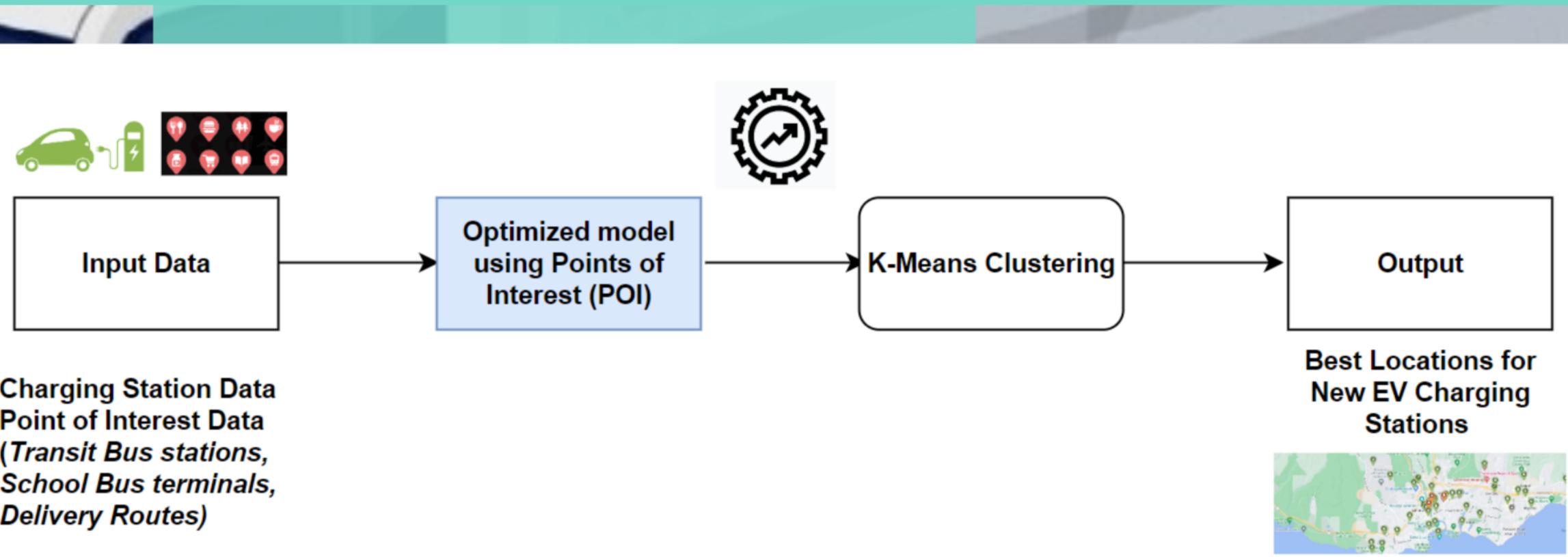
# Energy Demand Prediction for a Charging Station



# OPTIMAL PLACEMENT OF NEW CHARGING STATIONS FOR EACH VEHICLE TYPE

## Linear Programming and Clustering Model

To identify optimal locations for placing new charging stations for each type of vehicle using the point of interest related to them utilizing Pulp library linear programming and K-Means Clustering model.



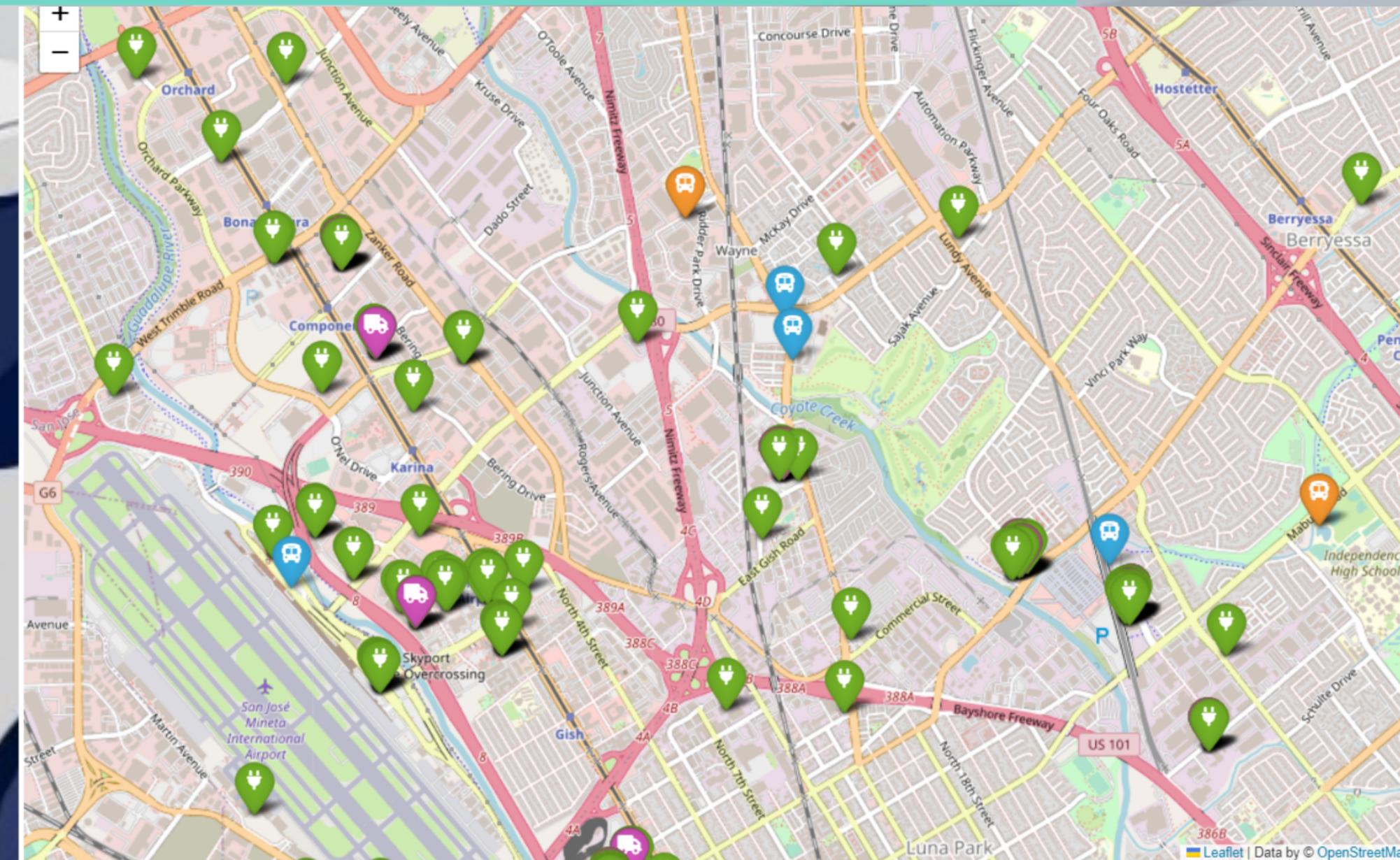
# Optimal Locations for New EV Charging Stations- San Jose

**Green**- Existing EV charging Stations

**Blue**- Suggested Charging Station Location for Transit Bus

**Orange**- Suggested Charging Station Location for School Bus

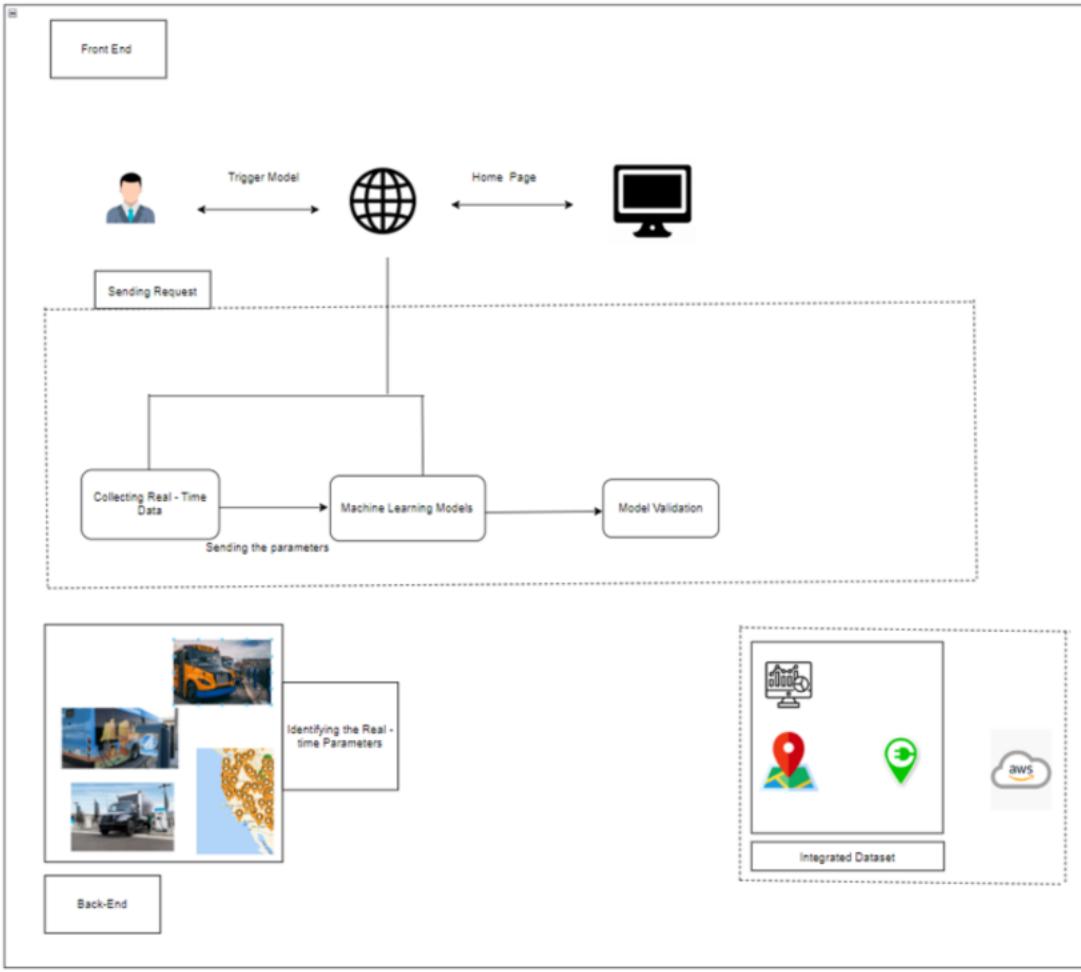
**Purple**- Suggested Charging Station Location for Delivery Truck



# Ongoing Modelling Results-Evaluation Metric Values

Feature	Vehicle Type	Time Frame	Model	MSE	MAE	RMSE	R <sup>2</sup>
EV Demand Forecast	Heavy/Medium Duty Vehicles	Yearly	Prophet	12.753	11.695	26.874	0.87
EV Range Prediction	Transit Bus	On Demand	Weighted Fusion Meta Regressor Model	9.063	9.874	88.113	0.75
	School Bus	On Demand		8.947	9.113	87.248	0.76
	Delivery Vehicle	On Demand		11.345	12.278	83.903	0.82
EV Energy Demand prediction	Transit Bus	Daily	Temporal Fusion Transformer	0.828	0.957	4.36	0.93
		Weekly		1.794	1.897	9.36	0.91
		Monthly		0.984	0.932	7.64	0.91
	School Bus	Daily		0.865	0.847	5.96	0.92
		Weekly		0.897	0.828	4.98	0.92
		Monthly		1.895	1.952	9.68	0.90
	Delivery Truck	Daily		2.969	2.643	12.36	0.88
		Weekly		2.453	2.387	11.97	0.89
		Monthly		2.874	2.775	13.01	0.88

# Outline of the System Design



Thank You



THANK YOU

# Data Driven Analysis of EV Charging Infrastructure - Medium/Heavy duty Vehicles

