Project Proposal — Optimizing EV Charging Station Placement in New York City

The Problem:

Electric vehicle (EV) ownership is rising steadily in New York City, driven by city-wide climate goals, state-level incentives, and shifting consumer preferences. But while adoption is increasing, the charging infrastructure isn't keeping pace—especially not evenly across all neighborhoods.

Today, some areas—like parts of lower Manhattan or downtown Brooklyn—have more chargers than they need, while others—such as the South Bronx or Eastern Queens—have very few or none. This imbalance creates long wait times, underused chargers, and limited EV adoption in lower-income or renter filled communities that often rely on street parking. Without intentional planning, the city risks building an EV infrastructure that leaves large parts of the population behind.

Our project asks: How can firms and city planners make smarter, more equitable decisions about where to build new EV charging stations?

Instead of simply focusing on where EVs are already popular or concentrated, we aim to explore broader and deeper patterns across the city. Specifically, we'll examine how chargers are distributed by ZIP code and type, how usage patterns and traffic volume reflect actual demand, and how neighborhood characteristics—from land use to income levels—may influence charger access. Our analysis will take into account:

- Current charger distribution by ZIP code and charger type (Level 2 vs. fast charger)
- Utilization rates—how much the existing chargers are actually being used
- Traffic volume as a proxy for future demand
- Socioeconomic indicators like median income and renter percentage
- Walkability and land use—e.g., residential vs. commercial zones
- Distance to nearest charger—to identify true charger deserts

We also want to center social equity in our analysis. For example, we'll examine whether neighborhoods like East Harlem or Brownsville—areas with high EV potential but lower income levels—have fair access to charging infrastructure

Analysis & What We Expect to Learn

Our analysis will generate location-specific infrastructure recommendations based on neighborhood characteristics. For example, we may suggest installing fast chargers near commercial corridors in Flushing or placing Level 2 chargers in residential parts of Astoria with high overnight parking turnover. Beyond identifying underserved ZIP codes, we aim to uncover why certain areas lack sufficient access—whether due to low charger density, high renter populations without off-street parking, or concentrated demand with few available options. In parallel, we will generate a Demand Gap Score (A normalized metric (0-100) where higher values indicate greater urgency for infrastructure investment.) for each ZIP code, enabling planners and firms to rank locations by infrastructure need and prioritize investment accordingly. Classification outputs—such as charger type suitability—will further help align solutions with local usage patterns and land-use conditions.

Actionable Insights

The results of our analysis can help both public agencies and private firms make smarter, more equitable decisions about where to invest in new EV chargers. By identifying ZIP codes with high traffic, lower incomes, and limited charger access, their budgets could be spent more efficiently—targeting areas where the need is greatest. There are many potential applications of this data, but we've identified three key ways a private company could leverage the results to maximize impact and efficiency…

1. Expand to Underserved Neighborhoods with High Latent Demand

   If the model identifies ZIP codes with rising traffic volumes, low charger availability, and high public charger dependence, a company could prioritize these areas, often skipped by traditional investment models but hold long-term value as EV adoption spreads into those areas.

2. Tailor Charger Type to Neighborhood Use Patterns

   Different areas call for different charger types. A firm could install fast chargers near busy areas where drivers want a quick top-off. Meanwhile, slower Level 2 chargers could be better suited for more residential neighborhoods where cars are parked for longer stretches—especially overnight.

3. Plan for Grid Capacity and Maintenance

   If the data shows that certain areas have extremely high charger utilization, utilities can plan accordingly. This might include scheduling proactive maintenance, upgrading local grid capacity, or staggering loads during peak hours to reduce stress on infrastructure.

These recommendations are informed by both the coverage score rankings and classification labels produced by our models.

Type of Data Science Problem (regression, classification, unsupervised, etc) & Data Description

This project is fundamentally a supervised regression task, where we predict a continuous 'Charging Demand Gap Score' to quantify infrastructure needs across neighborhoods. Using historical data on charger locations, traffic patterns, socioeconomic factors, and EV adoption rates, we'll train regression models to output prioritized investment scores for each ZIP code.

To provide practical charger-type recommendations (fast vs. Level 2), we'll implement rule-based logic derived from empirical utilization patterns - for instance, automatically assigning fast chargers to high-traffic commercial zones where existing fast chargers show >80% utilization. This hybrid approach combines the predictive power of regression modeling with transparent, business-interpretable rules, avoiding the complexity of a full classification system while still delivering actionable type-specific guidance.

Datasets:

NYC Open Data - NYCHA Resident Data Book Summary as of February 12, 2025
This dataset contains demographic information for residents of the New York City Housing Authority (NYCHA). It includes data on resident counts, age distribution, household composition, total gross income, and other relevant attributes. This would be used in social fairness analysis.

NYC Open Data - Primary Land Use Tax Lot Output (PLUTO) as of July 17, 2025
This dataset contains information on every tax lot in New York City. It includes data on land use type, building characteristics, lot size, zoning, ownership, and location. This would be used to decide whether the land is suitable for building a charging station.

NYC Open Data - Automated Traffic Volume Counts as of September 3, 2024
This dataset contains traffic volume counts collected by automated sensors at various locations throughout New York City. It includes 15-minute vehicle counts, direction of travel, and other metadata.

NYC Open Data - NYC EV Fleet Station Network as of July 13, 2025
This dataset contains the locations and attributes of EV charging stations. It includes data on station addresses, types of the charger, number of charging points, and operational status.

NYSERDA - Map of EV Registrations as of July 7, 2025

This dataset provides information on registered electric vehicles across New York State. It includes registration date, DMV ID, ZIP code, and other information. This would be used to represent EV coverage and be used in regression analysis and predictive modeling.

Target Variable and Attributes in the Data

The core of our solution is a Charging Demand Gap Score generated through supervised regression. This continuous metric quantifies neighborhood-level infrastructure needs by analyzing:
- Current charging capacity (density, utilization)
- Demand indicators (traffic volume, EV adoption trends)
- Equity factors (income levels, renter percentages, spatial accessibility)

For operational recommendations, we apply transparent business rules:
- Fast chargers target high-traffic (>15k vehicles/day) commercial zones (>40% land use)
- Level 2 chargers prioritize residential areas with high overnight parking demand and multi-family housing density >60%

This dual-output system provides both strategic prioritization (via regression) and actionable installation guidance (via rules), avoiding the overhead of classification modeling while maintaining decision-ready outputs.

Key Attributes in the Data

For each geographic area or ZIP code, the principal attributes considered for the analysis will include:

Attribute Name	**Description**	**Type**
Number of Existing Chargers	Installed EV stations (by type)	Numerical
Population Density	Residents per square mile	Numerical
Median Household Income	Socioeconomic status indicator	Numerical
EV Ownership/Registrations	# of registered EVs in the area	Numerical
Traffic Volume	Average daily vehicles (potential charging demand proxy)	Numerical
Percentage of Renters	Neighborhood demographics	Numerical
Land Use Mix	Residential, commercial, or public land	Categorical
Distance to Nearest Charger	Mean distance residents must travel to reach a charging station	Numerical
Area Type	Borough classification (Manhattan, Brooklyn, etc.), residential vs. commercial area	Categorical
Charger Utilization Rates	Average usage (occupancy) of stations by location/type	Numerical

Example Entry:

ZIP Code	Existing Chargers	Median Income	Traffic Volume	% Renters	Land Use	Charger Distance	Borough	Utilization Rate	Underserved?
11201	12	$110,000	22,000	65%	Mixed	0.8miles	Brooklyn	70%	No

10453	3	$35,000	20,000	80%	Residential	2.0 miles	Bronx	90%	Yes

Additional attributes may be considered as available, such as public transportation access, availability of off-street parking, or local policies supporting EVs.

The resulting data science problem can be framed as:

- Regression (predicting a continuous “coverage score” or expected future demand)
- Classification (labeling areas as “adequately served” or “underserved”)