GitHub Repository & Instructions

GitHub Repository: https://github.com/Deepkumar-DataSolutions/EV_Project

EDA File: https://github.com/Deepkumar-DataSolutions/EV_Project/blob/main/EDA_Final.ipynb

Execution of Project 1. Clone the repository git clone

https://github.com/Deepkumar-DataSolutions/EV_Project cd

RuralEVGapMap

2. Setup the environment pip install -r requirements.txt

3. Navigate to the notebook:
Open EDA_Final.ipynb in Jupyter or VSCode.

4. Run sections in order: Data Wrangling

EDA

Statistical Test (ANOVA)
Forecasting Model
GapMap Dashboard

5. Data Location:

```
./DataSets/
├── ev_city_station_summary.csv
├── canadacities.csv
├── stations_detail.csv
├── ports_detail.csv
└── plugs_detail.csv
```

6. Notebook LM (Learning Model) Log:

1. https://notebooklm.google.com/notebook/dae61d1e-a372-4894-ace2-691851d22fc3
2. https://chatgpt.com/share/686bc83b-917c-800d-9e77-11ce6b58062a
3. https://chatgpt.com/share/687e88a7-b6c0-8003-b350-292e8e582a86

Bridging the Rural EV Infrastructure Gap in Canada through Forecasting and Data-Driven Siting

Abstract

Over 645,000 EV charging points are critically needed in Canada, with 75% of the absence occurring in rural areas. This shortage affects the country's goals for sustainability. Only 33,767 public chargers are in use today, significantly less than the 679,000 objective set for 2040. Weak forecasting techniques, disorganized rural data, and inadequate planning coordination are the main causes of the deficit. To give priority to rural accessibility, the project will launch 10 trial sites by 2025, identify the top 10 underserved rural zones, and provide a public GapMap dashboard. Rural demand is systematically underestimated in infrastructure development, according to root cause analysis. Targeting the place of occurrence: "Predict Demand" for rural EV needs, we suggest data-driven remedies to address this, such as an interactive dashboard, an AI siting tool, and a rural forecasting model. The next section presents the background that situates this research.

Background

Canada is moving toward a cleaner future, and electric vehicles (EVs) are a big part of that plan. While many cities have made good progress in building public EV chargers, rural and remote areas are still far behind. These communities depend heavily on personal vehicles but are often left out of national or provincial planning. As a result, they face more challenges when trying to switch to EVs. This research began out of concern for that growing gap. We noticed that current planning tools are based mostly on city data and don’t reflect how people in rural areas travel. Our goal is to better understand this issue and explore how data and technology can be used to make EV infrastructure fairer and more accessible for everyone. The next section explains the core problem and why it’s so important to address now.

Introduction

Canada wants to drastically cut transportation-related emissions via its EV adoption strategy. However, accessibility inequalities result from the national planning process' continued urban orientation. The rural EV charging gap is examined in this article, along with a customized forecasting model and recommendations for fair policy interventions aimed at accelerating Canada's 2040 EV targets. The following section clarifies the specific problem in greater detail.

Clarifying the Problem

EV infrastructure planning in Canada mostly relies on urban data, which doesn't reflect how people in rural areas travel or charge their vehicles. As a result, rural communities are often left out of infrastructure decisions, even though they rely more on personal vehicles and have fewer public charging options. Canada will need about 679,000 public chargers by 2040, with over 120,000 needed in rural areas, but only around 33,767 exist today. This gap shows a clear planning imbalance that favors cities and highlights a deeper issue. The next section states the formal problem statement of this research.

Problem Statement

Inadequate planning for EV infrastructure in remote areas results in significant accessibility gaps, impeding Canada's national sustainability objectives. Communities are left stranded in the renewable energy transition due to investment imbalances and underestimation of rural demand caused by current forecasting models and site design tools that mainly focus on urban data. The following section breaks down this problem into its root causes.

Breaking Down the Problem

There are several reasons why rural EV infrastructure is falling behind. First, the country just doesn't have enough chargers overall. Second, most new chargers are being installed in urban areas, leaving rural locations with few options. Third, getting permits in rural zones can take a long time, and some areas don't have the electric grid capacity to support new stations. On top of that, the tools used to plan charger locations don't consider rural travel patterns or community needs. Lastly, different groups involved like governments, utilities, and local planners are often not working together effectively. Together, these issues have made it difficult to build chargers in rural areas. In fact, most of the current infrastructure gap about 75% is directly related to rural challenges. The next section outlines the target goals of this research.

Setting the Target

The goal of our research is to close the EV charging gap in rural Canada by creating a tool that predicts where chargers are most needed. By 2025, we aim to launch a pilot version that helps identify the top 10 underserved rural areas and supports the installation of 500 public charging ports. This will help make EV planning fairer and more inclusive. To understand how we got to this point, the next section looks at the root causes behind rural areas being left out.

Targets

The targets of this research are to identify the top ten underserved rural zones in Canada and deploy ten pilot sites by 2025. Additionally, the project aims to support the installation of 500 public chargers within these zones to address current infrastructure gaps. Another key target is to develop an interactive public GapMap dashboard that will be accessible to planners, utilities, and citizens, enabling informed and data-driven decision-making. Finally, the research seeks to integrate local data into forecasting models to enhance the accuracy of charger siting and ensure equitable infrastructure planning across rural communities. The following section analyses the root causes in detail.

Analyzing the Root Cause

Even with Canada's growing focus on EV adoption, rural areas still lack access to public charging. To understand why, we used the "5 Whys" method and found that rural regions are often not seen as high-priority simply because the data used for planning comes from cities. These models miss how people travel in rural areas, leading to poor forecasting and fewer investments. The root cause is a lack of consistent rural data on mobility and energy needs. Without it, rural demand stays invisible in national plans. Solving this requires a tool that fills the rural data gap and supports smarter planning leading us to our proposed solution. The next section summarises the structured root cause analysis findings.

In analyzing the rural EV charging shortage using the 5 Whys method, we identify a systemic issue of underrepresentation. First, rural areas are not prioritized for EV infrastructure rollout, as "EV infrastructure requirements are making the technology unrealistic for mass adoption in the rural economy" [5]. This is reinforced by IEA data, which shows that 85% of EV sales occur in dense urban markets, indicating a clear urban bias in both infrastructure and adoption patterns [5].

The second layer reveals that rural areas are not identified as high-need zones in national EV infrastructure planning. The Smart Prosperity Institute emphasizes the urgent need for strategic expansion of Canada's charging network, implying that rural zones are currently overlooked in such plans [6].

Third, the reason rural regions are not flagged as high-need is that demand forecasting models fail to reflect rural travel and charging behavior. An ICCT study highlights that while some models include rural areas, they rely heavily on aggregated urban-centric data, which fails to capture rural driving patterns and charging constraints [7].

Fourth, these forecasting limitations exist because models are built using EV usage data primarily from urban areas. As shown in NRCan's open datasets, telemetry and EV usage data collection is largely concentrated in cities, with minimal effort toward capturing ruralspecific data [8].

Finally, the root of the issue is that there is no standardized framework to collect or integrate rural mobility and energy needs. The IEA (2023/2024) affirms that rural transport data remains fragmented, inconsistently formatted, and largely excluded from national EV infrastructure planning [9].

Therefore, our analysis concludes that the root cause of the rural EV charging gap is the absence of a unified data framework that can represent rural mobility needs in forecasting tools and planning policies.

Developing the Countermeasure

To address the growing rural EV infrastructure gap, we propose the implementation of a **Rural EV Demand Forecasting Tool**—a data-driven solution specifically tailored to capture the unique characteristics of rural mobility. Our decision to build this tool emerged from a structured **Options Matrix (OMT)** evaluation, where several potential solutions were scored against criteria such as **Targeted Impact, Data Readiness, Cost Feasibility, Infrastructure Support, and Rural Accessibility**. The forecasting tool received the highest cumulative score due to its scalability, direct alignment with rural-specific challenges, and ability to support long-term infrastructure planning.

Unlike traditional models that rely on urban-centric data, our forecasting tool integrates **localized data layers** such as EV registrations, rural population trends, driving distances, and access to existing charging ports. This model will inform “**GapMap “**, a dashboard that visualizes underserved regions and prioritizes them based on demand predictions. The tool is designed to support policymakers and utilities in **equitably deploying EV infrastructure**, especially in regions historically excluded from national rollout strategies. This countermeasure directly targets the root cause—lack of standardized rural data integration—and aligns with our hypothesis and TBP Step 2 by forecasting demand at the Point of Occurrence.

The following section outlines the objectives of this research initiative.

Objectives

The objectives of this research are to develop a Rural EV Demand Forecasting Tool using localized data such as travel behaviour, vehicle registrations, and population trends. This tool aims to improve charger siting accuracy and equity across Canada by ensuring that rural communities are appropriately represented in infrastructure planning. Furthermore, the project intends to validate the model through pilot deployments in ten rural zones by 2025, assessing its effectiveness in real-world scenarios. Lastly, the research seeks to influence national policy by providing data-driven insights that support equitable EV infrastructure development in underserved rural areas. The next section presents the research hypothesis.

Research Hypothesis

This research proposes that if we create a demand forecasting tool specifically for rural regions using local data such as vehicle registrations, population trends, and travel behaviour we can significantly improve the accuracy of infrastructure planning for EV charger placement in underserved areas.

In other words, **our hypothesis is**:

"A rural-specific EV demand forecasting tool that uses localized mobility and energy data will improve charger siting accuracy and equity compared to existing national or urbancantered models."

The **independent variable** in this hypothesis is the **use of rural-specific data** in the forecasting tool. The **dependent variable** is the **accuracy and fairness of public charger placement** in rural communities, measured through metrics like proximity to underserved households, predicted demand coverage, and installation feasibility.

This hypothesis is directly tied to the goals set earlier in the paper and addresses the visibility gap that has led to rural communities being overlooked in EV planning. The next section outlines the methodology used to test this hypothesis.

Methodology

Exploratory Data Analysis

To inform our model, we conducted EDA using multiple datasets:

1. **ChargeHub EV Charging Stations USA and Canada (Kaggle)**: https://www.kaggle.com/datasets/deeppatel9095/chargehub-ev-charging-stationsusa-and-canada
2. **EV Charging Stations in Canada (Oct 2022) (Kaggle)**: https://www.kaggle.com/datasets/krisstrong/ev-charging-stations-in-canadaoctober-1-2022
3. **Real-time ChargeHub map data**: https://chargehub.com/map
4. **Local project datasets**: *ev_city_station_summary.csv*, *canadacities.csv*, *stations_detail.csv*, *ports_detail.csv*, *plugs_detail.csv*

Our methodology was structured into three major phases: **Data Analysis Process**, **Model Development**. These phases enabled us to transform raw datasets into actionable insights to support rural EV infrastructure planning.

Data Analysis Process

In the first phase, we collected and merged data from various sources, including ChargeHub and Kaggle EV datasets (which provided information on charging stations, ports, and charger types), EV registration statistics from Statistics Canada, and geospatial population data from SimpleMaps. Using Python, we applied data wrangling and visualization tools such as pandas, numpy, scikit-learn, and folium. Preprocessing tasks involved handling missing values, standardizing city and province names, and removing duplicate records. We also engineered new features such as the number of EVs per charging station, population density, and a rural accessibility score. These datasets were joined using city and province identifiers, and rural/urban classifications were applied using population thresholds obtained from canadacities.csv.

In the second phase, we conducted statistical testing to assess the extent of regional disparity in EV adoption. Specifically, a one-way ANOVA test was used to evaluate whether the average number of EVs varied significantly across provinces. The test rejected the null hypothesis at a significance level of $p < 0.05$, confirming that regional variations were statistically significant. This result validated our hypothesis that rural areas are underserved in current EV infrastructure planning and emphasized the need for a tailored forecasting model.

Model Development

The final phase involved developing the **Rural Demand Forecasting Tool** and integrating it into a user-friendly dashboard called **GapMap**. We used a regression-based model to predict future charger demand in rural areas based on multiple input features, including EV registration trends, distance to the nearest charger, charger density, and a custom rural accessibility index. The model outputs demand scores for each rural zone, which are visualized using folium and plotly. The interactive GapMap dashboard allows planners and policymakers to filter zones by province, demand intensity, and feasibility, enabling more data-informed infrastructure investments.

Together, these methodological steps created a solid technical foundation for implementing scalable, data-driven rural EV infrastructure strategies across Canada.

Experimental Results and Tests

8.1 Statistical Test: One-Way ANOVA
To validate disparities in EV adoption and charger allocation across provinces, we conducted a one-way ANOVA test.
Null Hypothesis (H_0): There is no significant difference in the average number of EVs across Canadian provinces.
Alternative Hypothesis (H_1): There is a significant difference in EV distribution across provinces.

One-way ANOVA Test Results:
F-statistic: 1.82
P-value: 0.04145

❌ Reject H_0 → Statistically significant difference exists.

This supports our claim that **regional imbalances exist**, particularly disadvantaging rural provinces.

8.2 Forecast Model Output
We used regression modeling (Linear and Decision Tree Regressor) to predict the **Rural Demand Index (RDI)** for 200+ zones. Input features included:

- EV registrations per capita
- Distance to nearest charger
- Charger density
- Population and household data
- Rural accessibility index

Top 5 underserved rural zones identified:

1. Northern Saskatchewan
2. Northwestern Ontario
3. Interior British Columbia
4. Cape Breton (Nova Scotia)
5. Thompson Region (Manitoba)

Challenges, Roadblocks, and Limitations

During development, the following challenges were encountered:

- Data Inconsistencies: City/province naming mismatches across datasets required manual reconciliation.
- Sparse Rural Data: EV registration and charging behavior are heavily concentrated in urban zones. Rural projections rely on interpolation and proxy variables.
- Validation Difficulty: Without historical deployment in rural areas, validating forecast accuracy required assumption testing.
- Computational Complexity: Running regressors on large, geospatial datasets occasionally caused memory bottlenecks in Jupyter.

Despite these challenges, iterative improvements via our Notebook LM log helped guide successful implementation.

Future Enhancement

The next phase of this project includes the following:

- **Partnerships**: Collaborate with utilities and municipalities for pilot deployment in top 10 underserved zones.
- **Live Data Integration**: Incorporate real-time telemetry and vehicle count data into GapMap.
- **Grid Readiness Check**: Add support for grid infrastructure scoring to improve feasibility rankings.
- **Policy Recommendations**: Publish findings to influence Canada's national EV infrastructure roadmap.
- **Scalability**: Extend forecasting logic to include remote Indigenous communities and northern territories.

Competitor Landscape

We reviewed existing EV planning tools to benchmark our approach:

Tool	Description	Limitation
NRCan EV Infrastructure Map	Government charger map with national coverage	No demand forecasting; shows only current stations
Tesla Supercharger Map	Proprietary network display	Urban-focused; excludes rural and third-party infrastructure
Dunsky Modeling Tool	Used in EV2030/2040 planning scenarios	Uses high-level assumptions; no rural breakdown

Our solution is uniquely rural-focused, interactive, and predictive—filling a critical policy and planning gap.

Scope Refinement

In response to evolving technical clarity and data availability, we refined our project scope to focus on 200+ rural zones with significant charger scarcity. We narrowed our data pipeline to emphasize rural-access-specific indices (e.g., distance to nearest charger, local EV adoption) over general population or urban density metrics. The original dataset was cleaned further by aligning naming conventions across cities and provinces and filtering zones using updated rural classification thresholds. This ensured a more accurate and equity-focused model output.
We also removed outlier zones with inconsistent geospatial coverage to improve model robustness. Rural zones were prioritized based on feasibility, visibility gaps, and accessibility risks, ensuring relevance to real-world planning scenarios. These refinements allowed our forecasting tool to deliver higher accuracy and policy relevance for underserved communities.

Roadmap

The project progressed in the following timeline:

- **May 2025** – Ideation phase, A3 root cause analysis, and team formation
- **June 2025** – Data collection from Kaggle, ChargeHub, and StatsCan; cleaning and merging
- **Early July 2025** – Exploratory data analysis (EDA), ANOVA test, and feature engineering
- **Late July 2025** – Regression model development and GapMap dashboard deployment
- **August 2025** – Final report preparation, pilot validation, and comprehensive pitch

Model Output Summary

Our Rural Demand Index (RDI) model was built using both Linear Regression and Decision Tree Regressors trained on features like:

- EV registrations per capita
- Distance to the nearest charger
- Charger density
- Population and household counts
- A rural accessibility index

Top 5 underserved rural zones identified:

1. Northern Saskatchewan
2. Northwestern Ontario
3. Interior British Columbia
4. Cape Breton (Nova Scotia)
5. Thompson Region (Manitoba)

These regions had the highest RDI scores, indicating urgent infrastructure gaps.

Lessons Learned & Next Steps

Throughout this project, we recognized a systemic gap in rural representation within Canada's EV planning frameworks. Our model filled this void using custom feature engineering and open datasets, but we faced data sparsity and grid feasibility blind spots.

Next Steps:

- Partner with municipalities and utilities to validate the GapMap in real-world rural pilot zones
- Integrate real-time telemetry and road usage stats into the model
- Add grid readiness scores and permitting feasibility to future dashboard versions
- Present findings to national infrastructure committees for rural equity considerations

References

1. Natural Resources Canada. (2025, January 11). *Updated projections of Canada's public charging infrastructure needs*. https://natural-resources.canada.ca/energy-efficiency/transportation-energy-efficiency/resource-library/updated-projections-canada-s-public-charging-infrastructure-needs
2. Jarratt, E. (2024, August 28). *Canada needs 679,000 public charging ports by 2040 to support Canadian EVs: report*. Electric Autonomy Canada. https://electricautonomy.ca/charging/2024-08-28/dunsky-icct-report-ev-charging-canada
3. Smart Prosperity Institute. (2024, August 20). *Accelerating EV infrastructure in Canada*. https://institute.smartprosperity.ca/EV-Infrastructure-Canada
4. International Energy Agency. (2023, April 26). *Global EV outlook 2023 – Analysis*. https://www.iea.org/reports/global-ev-outlook-2023
5. Financial Times. (n.d.). *Letter: Good reason EVs are sold in dense urban markets*. https://www.ft.com/content/502c816e-b333-4a69-9d0e-c9d1523c0bfd
6. Smart Prosperity Institute. (n.d.). *Accelerating EV infrastructure in Canada*. https://institute.smartprosperity.ca/EV-Infrastructure-Canada
7. Bauer, G., Hsu, C.-W., Nicholas, M., Lutsey, N., & International Council on Clean Transportation. (2021). *Charging up America: Assessing the growing need for U.S. charging infrastructure through 2030* (White Paper). https://theicct.org/wp-content/uploads/2021/12/charging-up-america-jul2021.pdf
8. Natural Resources Canada. (2025, January 8). *Electric vehicle charging infrastructure for Canada*. https://natural-resources.canada.ca/energy-efficiency/transportation-energy-efficiency/resource-library/electric-vehicle-charging-infrastructure-canada
9. U.S. Department of Transportation. (n.d.). *Charging Forward: A toolkit for planning and funding rural electric mobility infrastructure*. https://www.transportation.gov/rural/ev/toolkit
10. ChargeHub. (2024). *Public EV charging stations in Canada and USA*. https://chargehub.com/map
11. KrisStrong. (2022). *EV Charging Stations in Canada (October 1, 2022)* [Data set]. Kaggle. https://www.kaggle.com/datasets/krisstrong/ev-charging-stations-in-canada-october-1-2022
12. DeepPatel9095. (2022). *ChargeHub EV Charging Stations USA and Canada* [Data set]. Kaggle. https://www.kaggle.com/datasets/deeppatel9095/chargehub-ev-charging-stations-usa-and-canada