**Introduction**

**Accelerating Electric Vehicle Adoption Towards Emissions Reduction Goals**

In Washington, understanding and fostering electric vehicle (EV) adoption is more than a market trend; it's a crucial step towards achieving the state's ambitious goal of having 1 million EVs on the road by 2030, as part of its emissions reduction strategy. This analysis, using EV registration data, focuses on unraveling the trends in EV types, makes, models, and technological advancements. This exploration is key to understanding consumer preferences, technological progression, and overall market dynamics. These insights are vital in strategizing initiatives and policies that support the state's environmental objectives, by aligning with the factors driving EV adoption.

**Problem Statement**

Data-Driven Approach to Support Emissions Goals

The challenge lies in effectively leveraging EV registration data to inform strategies that contribute to the state's emissions reduction goals. Specific issues include:

Consumer Behavior and Preferences: Understanding the factors influencing consumer choices in EVs is essential for promoting wider adoption.

Technological Advancements Impact: Analysis of model years and electric ranges helps gauge progress in EV technology, crucial for meeting the emissions targets.

Strategic Alignment with Emissions Goals: Utilizing data to align marketing, product development, and policy-making with the state's environmental objectives.

BI Project Goals and Importance:-Leveraging Data for Environmental and Market Success

**BI project aims to harness EV registration data to:**

Influence Consumer Choices: By understanding preferences, this analysis can guide consumers towards more environmentally friendly options.

Track Technological Trends: Insights into technological trends will inform strategies to promote the adoption of newer, more efficient EV models.

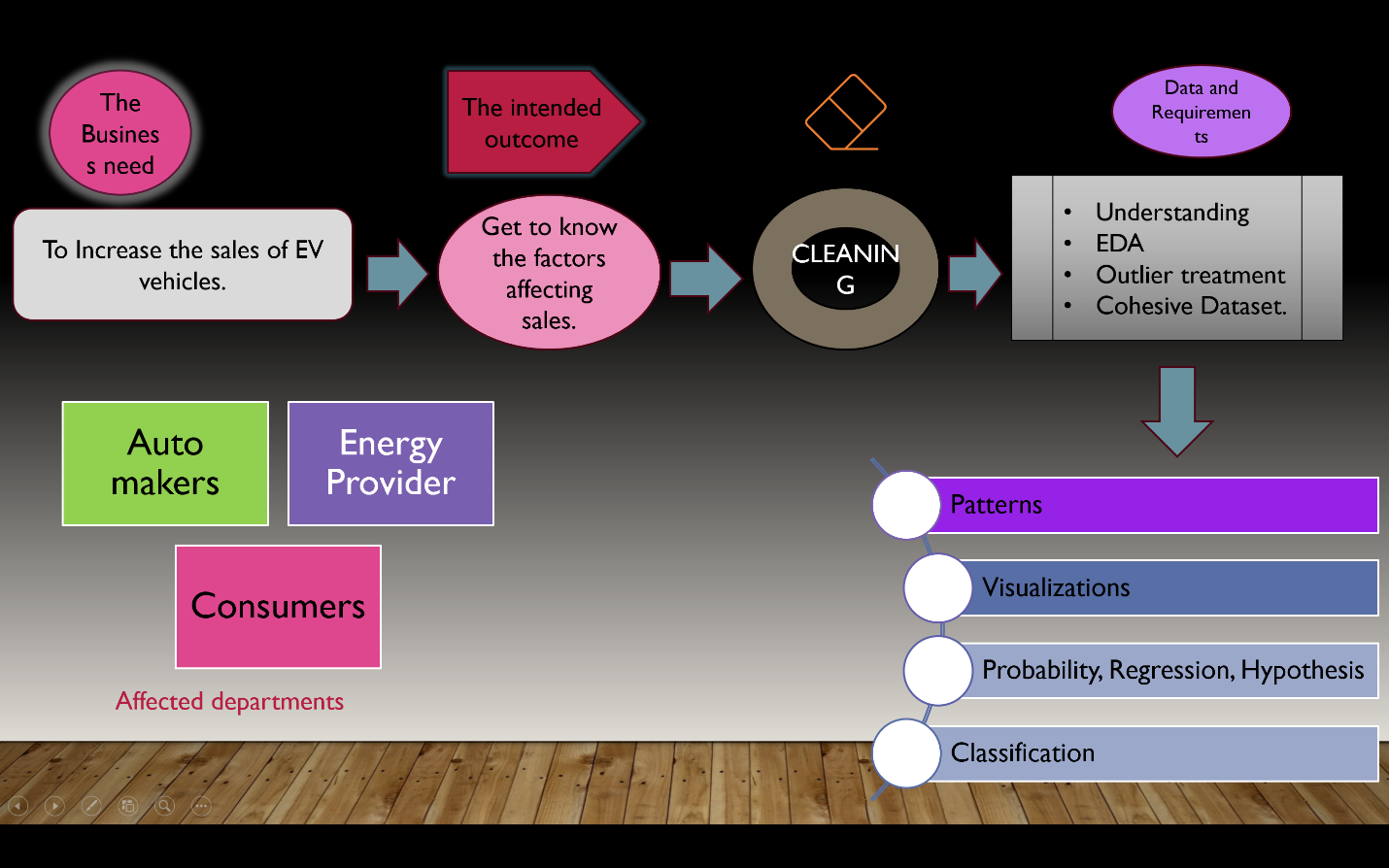
Align with State Emissions Goals: Data-informed strategies will help align business goals with the state's environmental objectives, creating a synergy between economic and ecological progress.

**Importance:-**

This BI initiative is pivotal in aligning organizational strategies with Washington's environmental goals. By adopting a data-driven approach, this can contribute significantly to the state's mission of increasing EV adoption, thereby playing a vital role in reducing emissions and fostering sustainable development.

**1. The Business Process**

**BI Roadmap**

****

**Readiness Assessment**

Before beginning analysis of electric vehicle data for Washington State, the following readiness checks have conducted:

Ensured access to required datasets from data.gov and other public data sources

Verified data quality and completeness

Confirmed software tools are available for data analytics (Excel, R)

**Anticipated Outcomes**

The key anticipated outcomes from analysis of EV data for Washington State include:

Manufacturers: Offers insights into vehicle types and model popularity by region, helping to tailor production and marketing.

Policymakers: Provides data on the adoption rates of EVs, supporting the development of targeted incentives and infrastructure investments.

Stakeholders: Enables data-driven strategy alignment with consumer adoption trends and infrastructure developments.

**2. Data and Data Structures**

**Overview of the Dataset**

This analysis utilizes a dataset detailing attributes of electric vehicles (EVs) in Washington State. Key data points include County, City, Model Year, Make, Model, EV Type (BEV or PHEV), Electric Range, and CAFV Eligibility.

**Data Cleaning and Transformation**

**To enhance data quality:**

* CAFV Eligibility Adjustment: Simplified into 'Eligible' and 'Not Eligible/Unknown' for clear analysis.
* Data Pruning: Filtered to include only models up to the year 2023 and limited to Washington State.
* Column Removal: Extraneous columns like 'access\_days\_time', 'Base.MSRP', and others were removed for data precision.

**Data Enrichment and Imputation**

* Electric Range Imputation: Updated zero range values based on make and model, ensuring accurate representation of EV capabilities.
* Descriptive Statistics: Calculated key statistics like mean, median, and standard deviation for the electric range to understand the distribution pattern.

**Visual Analysis and Consumer Insights**

* Popularity of EV Makes: Bar plot visualization to identify trends in EV makes.
* Technological Advancements Impact: Line plot depicting the progression in electric range over model years.
* Emissions Goals Alignment: Heatmap showcasing average electric range across counties and model years.

**Advanced Visualizations**

* 3D Scatter Plot: Explored the interplay between model year, electric range, and EV make.
* Electric Range Distribution: Histogram to analyze the spread of electric range values.

**Data Summary**

A comprehensive overview of the dataset was provided, highlighting the diversity in EV makes, models, and electric range distribution.

**Applied Probability Analysis**

Probability Distribution of Electric Vehicle Types Across Counties: Utilized dplyr in R to group data by county and electric vehicle type, calculating the frequency and probability distribution. This method was chosen for its simplicity and effectiveness in handling categorical data.

Likelihood Estimation Based on Make and Electric Range: Again leveraging dplyr, likelihoods of CAFV eligibility were calculated for different makes and electric ranges, providing insights into factors influencing CAFV eligibility.

**Statistical Analysis**

Descriptive Statistics: Calculated average electric range using the mean function in R, both overall and grouped by city. This fundamental statistical approach was crucial to understand the central tendency of data.

Comparative Analysis Across Cities: Grouped data by city and calculated average electric ranges. This step offered a geographical perspective on electric range variability.

**Hypothesis Testing**

T-Tests for Electric Range: Conducted t-tests to compare the electric range between different vehicle types and CAFV eligibility statuses. The t-test was selected for its ability to compare means between two groups, suitable for binary categorical variables (Electric Vehicle Type, CAFV Eligibility).

**Regression Analysis**

Linear Regression for Electric Range and Model Year: Performed a linear regression to examine the relationship between electric range and model year. The linear model was chosen for its efficiency in exploring relationships between a continuous dependent variable and one or more independent variables.

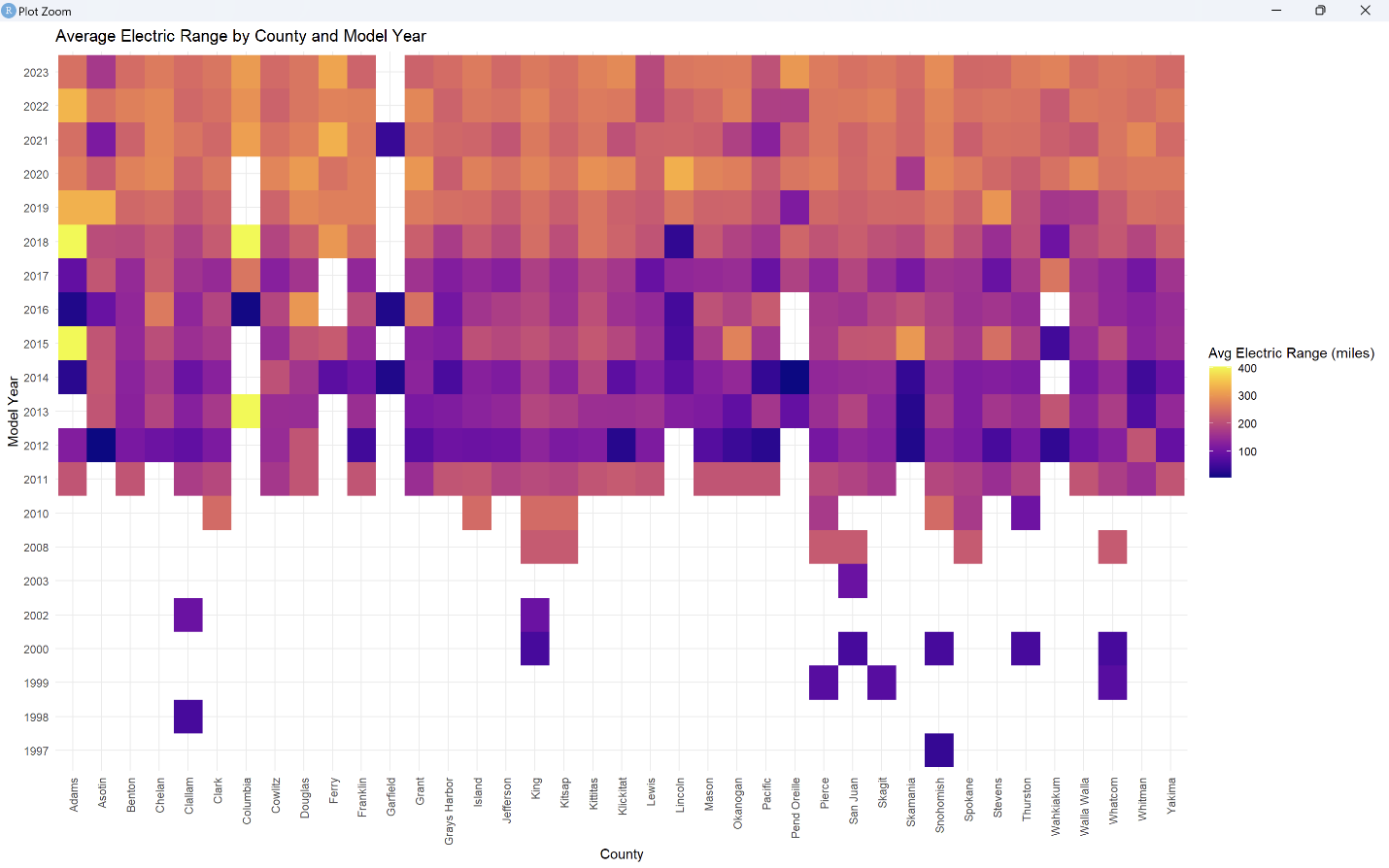
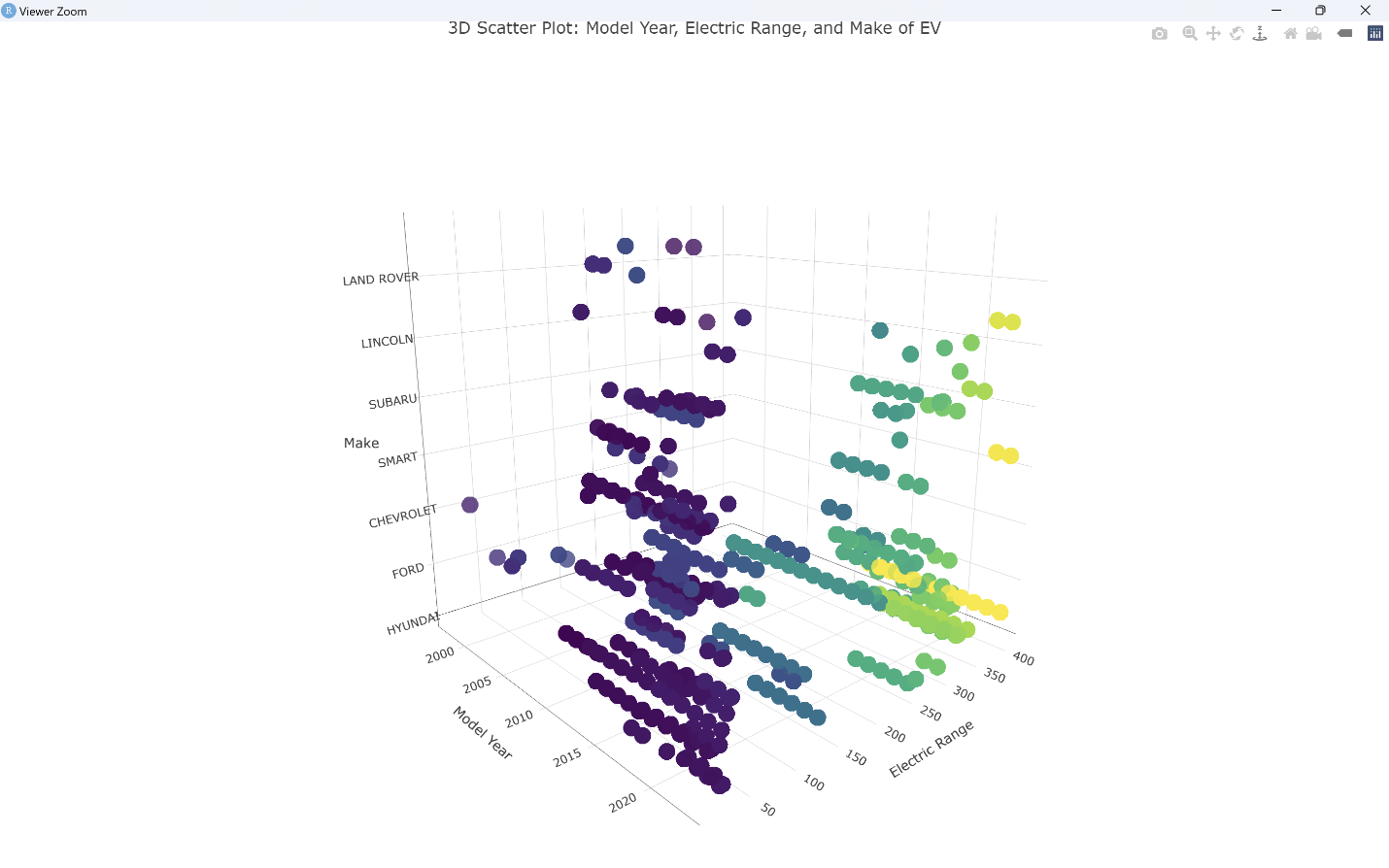
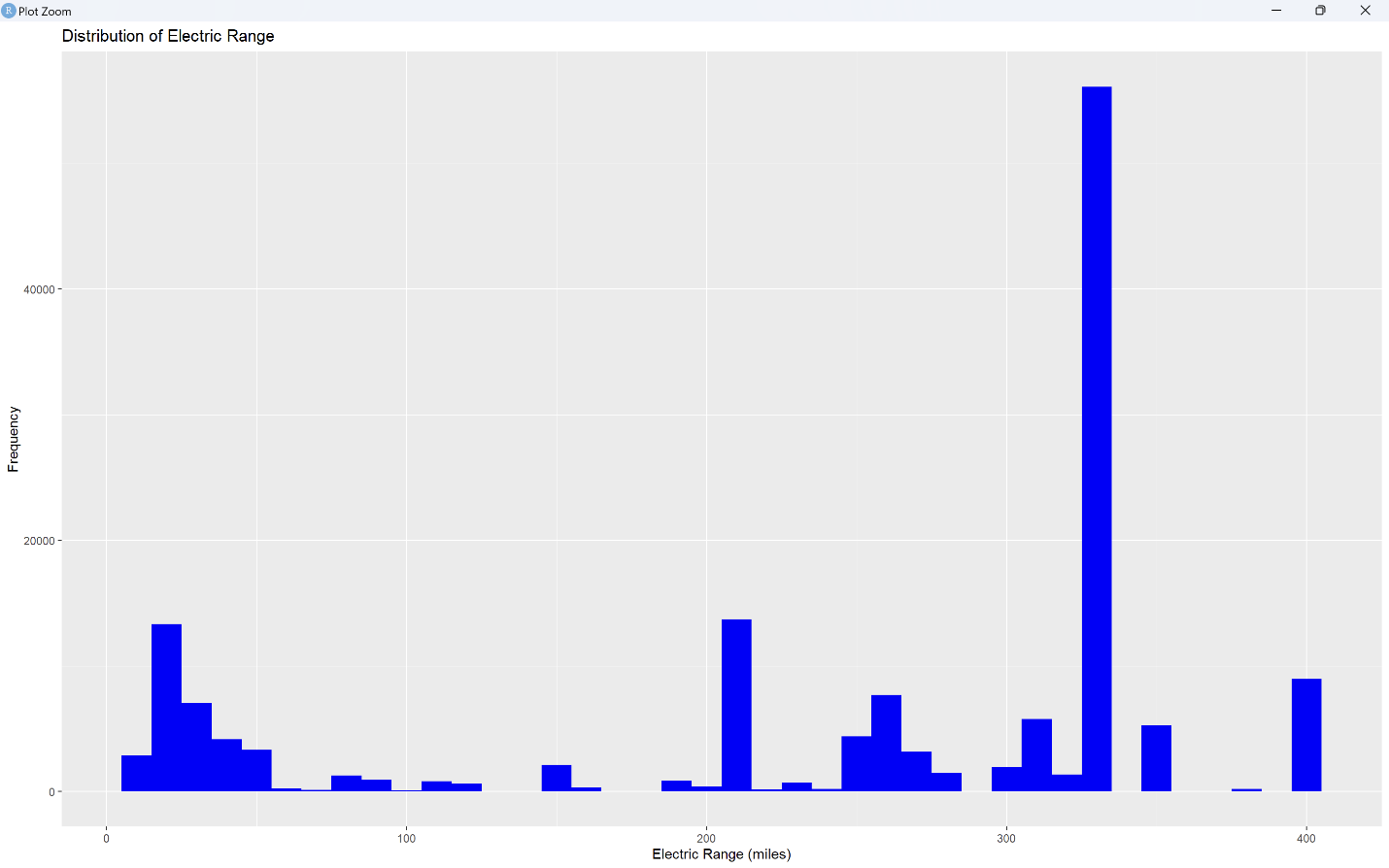
Visualization

Scatter Plot with Regression Line: Utilized ggplot2 to create a scatter plot with a regression line, visually depicting the relationship between model year and electric range. This was instrumental in illustrating trends and patterns in the data.

Classification Analysis

Classification of Vehicles Based on CAFV Eligibility: Applied logistic regression using the caret package for classifying vehicles as eligible or not for CAFV based on various attributes. Logistic regression was appropriate here due to the binary nature of the response variable (CAFV eligibility).

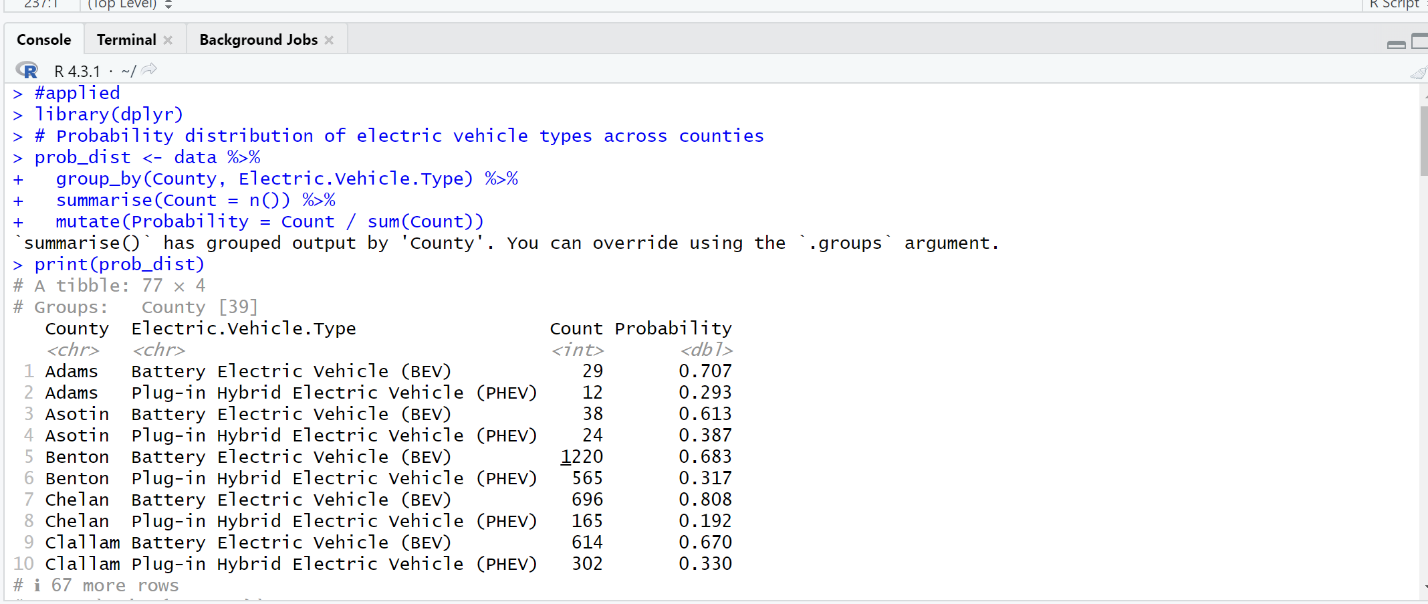
**Visual Output:-**

****

**Results**

The results indicate that there is a difference, in the distribution of Battery Electric Vehicles (BEVs) and Plug in Hybrid Electric Vehicles (PHEVs), among counties. For example in Adams County BEVs made up 70.7% of the total while PHEVs accounted for 29.3%. On the hand Asotin County had a split of 61.3% BEVs and 38.7% PHEVs.

Visual Output:



Significance:- This data is vital for targeted marketing strategies and infrastructure development that cater to specific county preferences.

**Likelihood Estimation Based on Make**

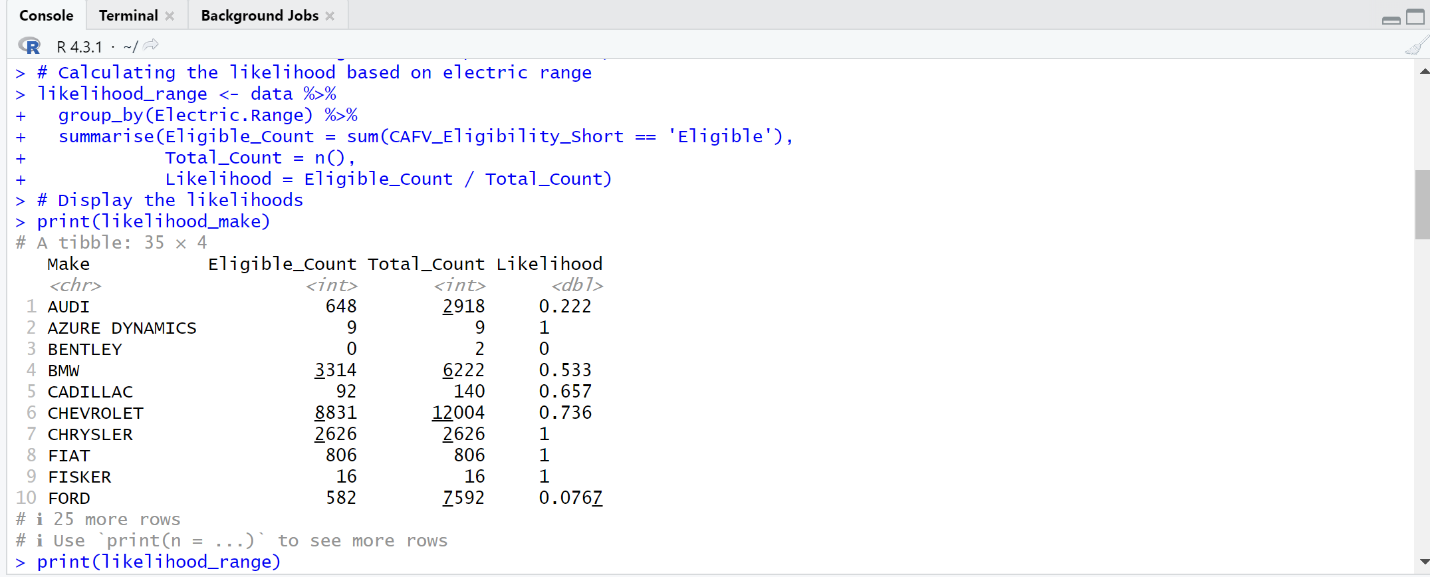
Findings:

* The likelihood of CAFV eligibility varied by vehicle make. Chevrolet showed a 73.6% likelihood, while Ford's was significantly lower at 7.67%. Chrysler, Fiat, and Fisker exhibited 100% eligibility.

Significance:

* These differences are key for manufacturers in understanding CAFV eligibility criteria and for policymakers in evaluating the program's inclusiveness.

**Visual output:**

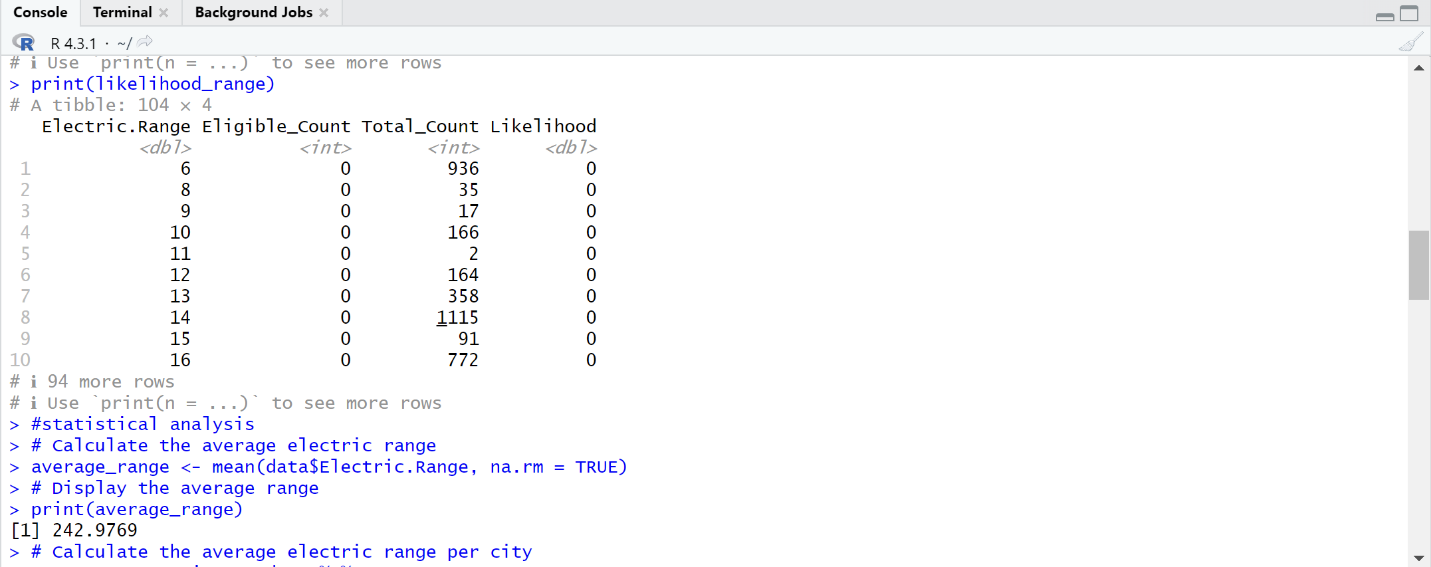


**Likelihood Estimation Based on Electric Range**

Findings:

Certain electric ranges demonstrated a zero likelihood of CAFV eligibility, indicating specific range thresholds for program participation.

**Visual Output:**



Significance:-This variation is crucial for strategic decisions in EV infrastructure development and consumer choice, particularly in urban planning.

**Average Electric Range Analysis**

Findings:

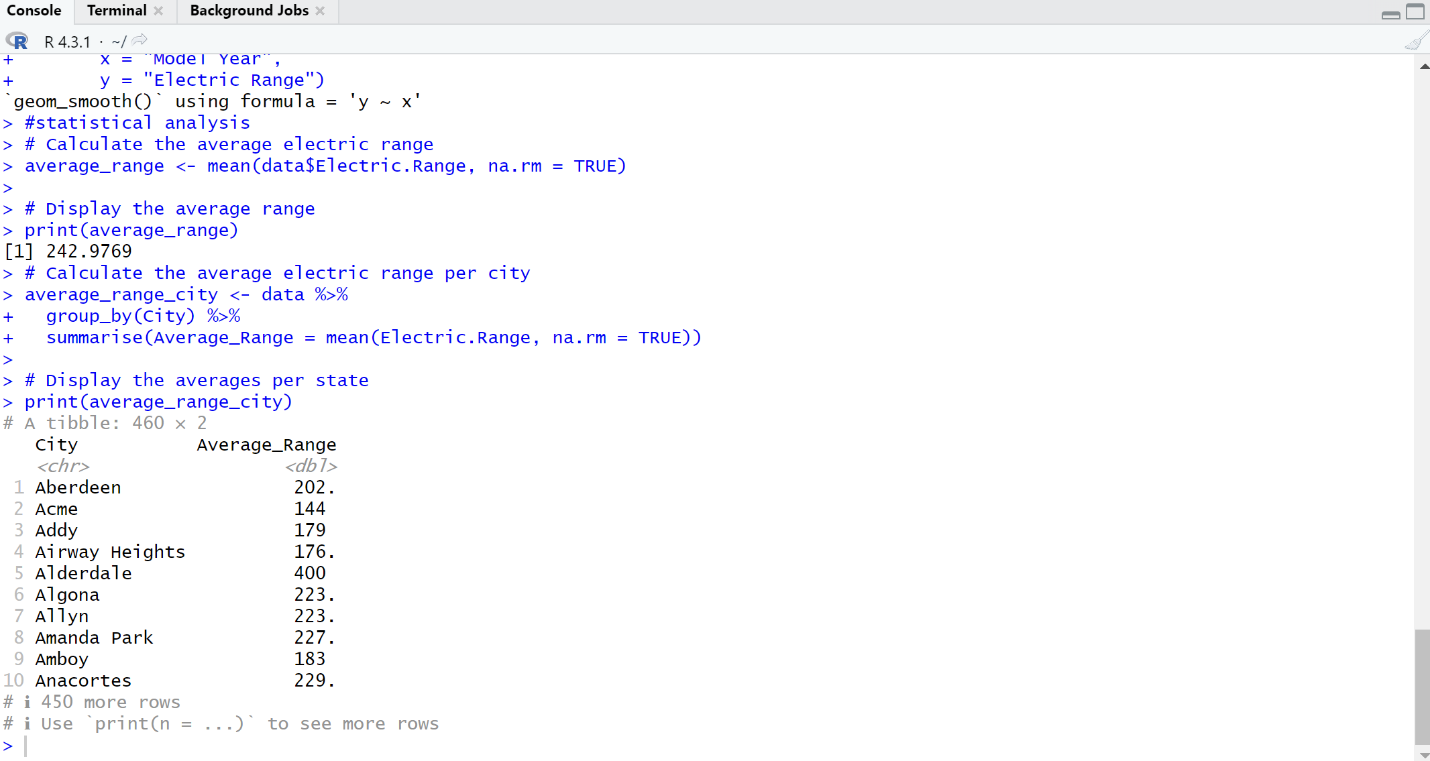
The overall average electric range for electric vehicles (EVs) was calculated to be approximately 243 miles.

Significant variations were observed in average electric ranges across different cities. For example, Aberdeen averaged 202 miles, while Alderdale notably averaged 400 miles.

Interpretation:

This analysis underscores the diversity in EV capabilities across urban locations, reflecting factors such as model availability and geographic preferences.

**Visual Output:**



**Hypothesis Testing on Electric Range**

Hypothesis Testing, on Electric Range

T Test; Types of Electric Vehicles (Rejecting the Null Hypothesis)

I’ve conducted a Welch Two Sample t test to determine if there is a difference in the range between Battery Electric Vehicles (BEV) and Plug in Hybrid Electric Vehicles (PHEV).

After analyzing the data rejected the hypothesis (H0) that suggests no difference in electric range between BEVs and PHEVs.

On the hand findings supported the hypothesis (H1) indicating that there is indeed a significant difference in electric range between these two types of vehicles.

The results demonstrate a difference (p value < 2.2e 16) between the groups.

The average electric range for BEVs stands at 300.68 miles, which's significantly higher compared to PHEVs with an average of 40.73 miles.

Furthermore calculated a 95% confidence interval for the difference in means, which ranged from 259.32 to 260.59 miles.

T Test; CAFV Program Eligibility (Rejecting Null Hypothesis)

Also performed another Welch Two Sample t test to evaluate whether there exists a discrepancy in range between vehicles eligible for the Clean Alternative Fuel Vehicle (CAFV) program and those that are not eligible or have unknown eligibility status.

Upon analysis rejected the hypothesis (H0) which posits no difference, in electric range based on CAFV program eligibility.

The evidence supported the hypothesis (H1) stating that there is a variation, in electric range depending on eligibility, for CAFV.

The results indicate a difference (p value < 2.2e 16), between the two groups.

On average vehicles that meet the eligibility criteria have a range of 227.34 miles while those that are not eligible or have an unknown eligibility have a range of 254.25 miles.

The 95% confidence interval for the difference in means ranged from 28.19 to 25.63 miles.

This discovery offers insights for consumers regarding their range expectations. Can assist policymakers in refining the criteria, for Clean Air Friendly Vehicles (CAFV).

**Regression Analysis: Electric Range and Model Year**

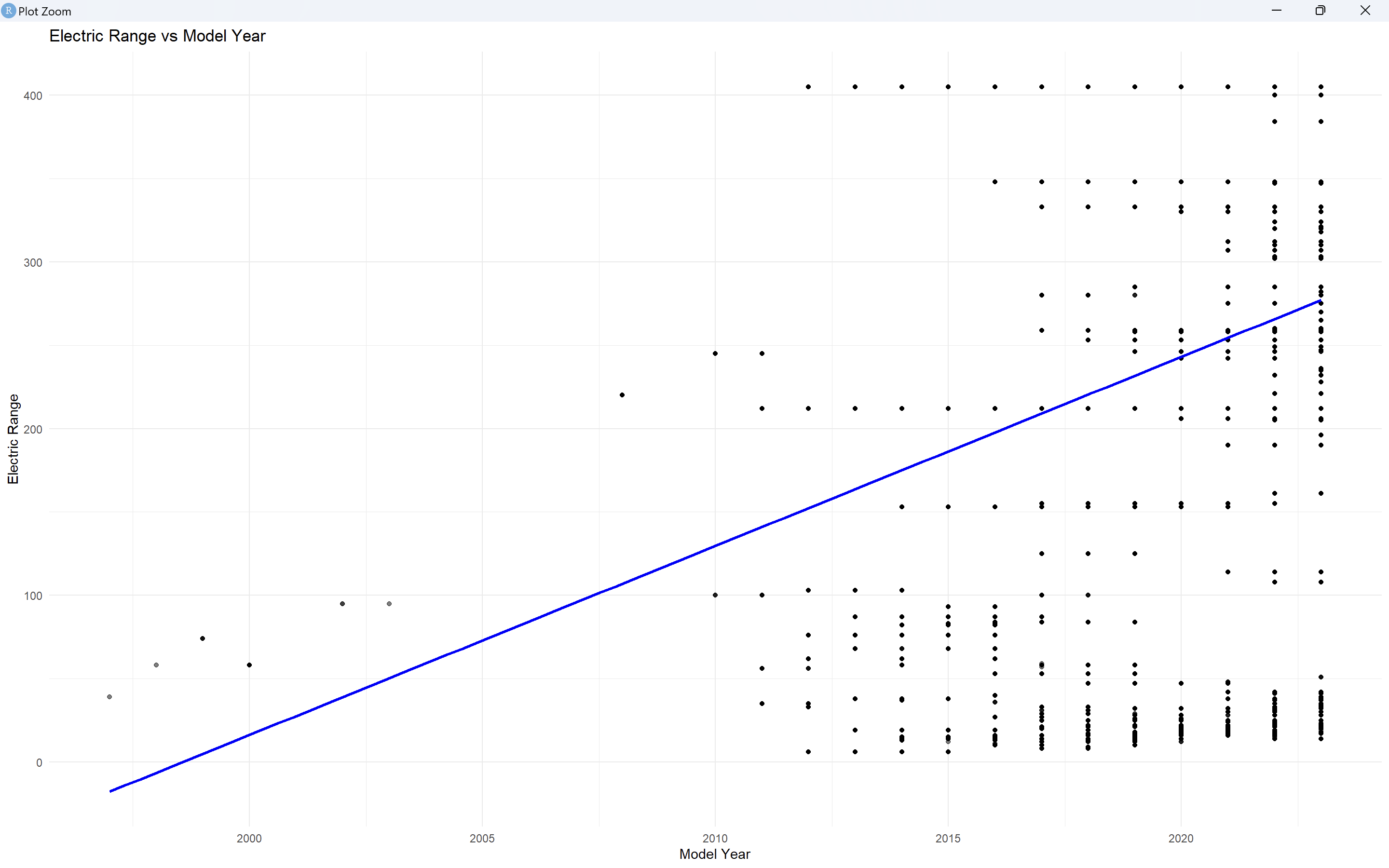
A study was conducted to analyze the connection, between the year a model was released and the range of vehicles. The study examined data from over 149,000 vehicles. The results of the study showed that for each model year there was an estimated increase in range of about 11.3 miles.

The analysis revealed a positive relationship between the model year and electric range (p < 0.001). This was indicated by a t value (t = 109.9) and a small p value for the coefficient related to the model year.

The adjusted R squared value of 0.075 indicates that the model year alone can account for 7.5% of the variation in range within the dataset. It is important to note that other factors like battery capacity EV make, etc. are likely to have an impact on range.

The residuals plots demonstrate a distribution and consistent variance meeting all assumptions required for regression analysis. No notable patterns were found when examining residuals, against fitted values.

This analysis uncovered a definite positive trend between vehicle model year and electric range over time, with range increasing by over 11 miles with each newer model year on average. The model year factor explains 7.5% of the variation in range, with other technical factors also playing a role.

**Visual Output:**

Significance:-This trend underscores the rapid advancements in EV technology, vital for manufacturers and consumers alike.

**Conclusions**

Leveraging Model Insights for Business Optimization: The linear regression analysis has revealed a significant positive relationship between electric vehicle range and model year. To meaningfully improve business performance, incorporate this insight into product development cycles and consumer marketing strategies. Understanding the over 11 mile per year increase in range empowers strategic alignment with evolving consumer preferences and technological advancements.

Addressing Analysis Limitations: Acknowledging analytical limitations is crucial for continuous improvement. The linear regression relies on assumptions like linearity and normal distributions that may not perfectly hold. Its relatively low 7.5% R-squared value indicates other unexplored factors like battery capacity also influence range. Additionally, potential data issues like outliers may impact accuracy. To build on these findings, future analyses should explore nonlinear models and additional relevant variables.

Justifying the Analysis: Despite limitations, a linear regression approach was justified to examine the relationship between range and model year given its interpretability and ability to quantify statistically significant trends. Its simplicity provides a foundation for strategic decision making even amid a complex, dynamic market.

Broader Ecosystem Perspectives: The electric vehicle market is part of a rapidly evolving industry shaped by technological innovations, policy changes, consumer preferences and more. While focused on model year and range, this analysis serves as a stepping stone. Future initiatives should adopt a multifaceted perspective across an array of variables to support comprehensive, agile business intelligence strategies that captialize on market opportunities.

Informing Decisions & Paving the Road Ahead: These insights inform product planning, inventory, and consumer marketing. More broadly, this analysis promotes a data-driven culture. Moving forward, expanded analyses, advanced techniques and continuous learning will help the organization nimbly navigate the electric vehicle landscape.

**Appendices**

Appendix A: R Syntax for Data Analysis

Analyzed Few websites for R code and layering them and Reffered to Textbook as well