

Electric Vehicle Market and Tesla Sales in Washington State

Frank Dong, Qiu Jingxi, Di Ziyang

Abstract

Over the past four years, the electric vehicle (EV) market has experienced significant growth, with Tesla Inc. leading this surge. Highlighting its market dominance, Tesla's stock value peaked at twenty times its lowest point in 2019, reflecting strong investor confidence in the EV sector and Tesla's innovative capabilities. Our analysis delves into electric vehicle trends from 2020 to 2024, focusing on overall sales increases, Tesla's market position, its competition with other EV manufacturers, the popularity of Tesla models, and the geographical distribution of EV sales. In this study, we utilized SVM prediction and K-means clustering to analyze processed data on vehicle holdings and transactions in Washington State to identify the characteristics of different markets and their influencing factors. We have identified the important features of the unpopular market, rapidly developing market, and highly developed market, as well as important indicators for determining the level of electric vehicle development in the region. Through this result, decision-makers can understand the characteristics of transportation and automobile markets in different counties, thereby more effectively positioning policies and resources.

Key words: EV market, SVM, K-means

Table of Contents

Abstract.....	0
1. Introduction.....	1
2. Data.....	3
2.1 Data Collection.....	3
2.2 Data Cleaning.....	4
3. Exploratory Data Analysis.....	4
3.1 Feature Extraction.....	4
3.2 Registration Types and Trends Analysis.....	4
3.3 Market Share Analysis.....	5
3.4 Electrification Level Distribution Analysis.....	6
3.5 Trends Analysis of Multiple Electrification Level Vehicles.....	7
3.6 Transaction Trend Analysis of Major Brands.....	8
3.7 Analysis of Tesla's Sales Trends.....	8
3.8 Analysis of WA State Market Share.....	9
3.9 Analysis for Tesla's Model Sales.....	10
4. Methods and Analysis.....	11
4.1 Variables Definition and Data Processing.....	12
4.2 Methodology.....	12
4.2.1 SVM.....	12
4.2.2 K-means Clustering.....	13
5. Results.....	13
5.1 SVM Results.....	13
5.2 K-means Clustering.....	14
6. Conclusion.....	15
References.....	17
Appendices.....	18
Appendix A.....	18
Appendix B.....	19
Appendix C.....	20

1. Introduction

This research paper delves deeply into the dynamics of the electric vehicle (EV) market in Washington State from 2020 to 2024, with a particular focus on the influential role of Tesla Inc. As the automotive industry shifts toward more sustainable practices globally, Tesla has emerged as a pivotal player, fundamentally changing how the market and consumers perceive electric vehicles. Through the application of advanced analytical methods such as Support Vector Machine (SVM) prediction and K-means clustering, this study conducted in-depth research on an extensive amount of vehicle ownership and transaction data. This analysis aims to help map prevalent market patterns, define distinct market segments, and determine the driving factors behind EV adoption in different counties.

The background of this study is marked by the rapid expansion of the electric vehicle sector, spurred by advancements in technology, heightened environmental concerns, and shifting consumer behaviors. Particularly in Washington State, which is known for its progressive environmental policies and tech-savvy population, there has been a notable increase in the adoption of electric vehicles. Tesla's strategy of continuous innovation and its aggressive marketing have not only heightened its brand visibility but have also set a high benchmark within the industry, prompting traditional automakers to hasten their own transitions towards electric mobility and thereby intensifying the competitive landscape.

Tesla Inc., known for its innovation in electric vehicles and sustainable energy solutions, has exhibited an extraordinary capacity to influence and propel market trends. The significant growth in Tesla's stock value, which peaked at twenty times its value in 2019, reflects strong investor confidence in Tesla's business model and its future prospects. Moreover, Tesla's substantial market share and the popularity of models like the Model Y and Model 3 underscore its success in meeting consumer demands for vehicles that are both affordable and environmentally friendly.

The primary objectives of this research are comprehensive. First, it aims to analyze the overall growth trajectory of the EV market within the state during the study period. It also evaluates

Tesla's position within the market against other EV manufacturers, exploring how Tesla's model preferences among consumers vary across the state, and using sophisticated machine learning techniques to identify different market segments, such as less affluent or rapidly developing markets. Additionally, this study examines the geographical distribution of EV sales to inform targeted marketing strategies and support business expansion plans.

In order to meet these objectives, the research employs a blend of predictive and descriptive analytical techniques. Support Vector Machine (SVM) prediction is used to anticipate market trends based on past vehicle ownership and transaction data, classifying counties into various categories based on their adoption rates and potential for growth. Meanwhile, K-means clustering is applied to segment the market based on similarities in market behaviors and consumer preferences, helping to categorize counties according to their EV adoption rates and socio-economic characteristics.

The dataset utilized in this study is sourced from the comprehensive vehicle registration records provided by the Washington State government. These records include detailed information about electric and non-electric vehicle registrations from 2019 to 2023. Originally encompassing approximately 6 gigabytes of data, the dataset was meticulously refined to focus on the most pertinent records, including attributes such as vehicle make, model, and county of registration, which are crucial for analyzing underlying market trends and consumer behavior.

This study offers insights into the EV market's dynamics in Washington State. Understanding Tesla's role and the broader competitive landscape will enable policymakers, business leaders, and investors to make informed decisions regarding resource allocation, policy making, and strategic planning in the EV sector. Additionally, the geographical analysis of EV sales will provide valuable information for crafting targeted marketing strategies and expanding business operations effectively.

2. Data

2.1 Data Collection

This study uses two main datasets, one is the Vehicle Registration Transactions by Department of Licensing and the other is the Electric Vehicle Population Data. Extra information has been included from [unitedstateszipcodes.org](https://www.unitedstateszipcodes.org/), to provide zipcode-county information to help with data cleaning.

Vehicle Registration Transactions by Department of Licensing

This dataset was provided by the Washington state government and includes detailed vehicle registration transaction data from 2020 to March 2024. This includes multiple columns, providing an in-depth understanding of the vehicle type, manufacturer's name, registration date, and vehicle type. This dataset contains approximately 6GB of data, with a total of 27,478,265 records and 22 attributes. This dataset provides a comprehensive view suitable for detailed analysis and research, including a deep understanding of vehicle brand, model, year, and specific registration details.

Electric Vehicle Population Data

This dataset was provided by the Washington state government and includes detailed Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) that are currently registered through Washington State Department of Licensing (DOL) until March 2024. It includes 181K rows and 17 columns, includes the information of county, city, postal code, model, model year, EV type, and other identifying information.

zipcode-county information

This dataset was provided by the United States Zip Codes website, the data includes information about zip codes and it corresponds to county name information. This information can be utilized to help with data cleaning.

2.2 Data Cleaning

We used the 'Electrification Level' attribute to mark records as EVs to identify electric vehicles, and extracted the 'Transaction Year' from the 'Transaction Month and Year' field for accurate time-stamping.

Then, we removed irrelevant or incomplete records, such as data from outside WA states, and maintained a strict focus on data relevant to Washington State and confirmed electric vehicles. We also use zipcode-county data to help filter out the wrong zip code data. Because many county columns are empty, this step helps us save plenty of data from being excluded.

We kept important attributes like county, transaction year, make, model, and vehicle type, creating a well-structured dataset ready for in-depth analysis. We also grouped the data by county and transaction year, allowing for a detailed examination of geographic and temporal trends in EV registrations.

3. Exploratory Data Analysis

3.1 Feature Extraction

To initiate our EDA, we first manipulate the dataset to extract relevant features that will allow us to focus specifically on electric vehicles. This involves creating a *“Transaction Year”* column derived from the *“Transaction Month and Year”* entries, which gives us a consolidated view of the data on an annual basis. Furthermore, leveraging the *“Electrification Level”* attribute, we ascertain the status of each vehicle as electric or not.

3.2 Registration Types and Trends Analysis

With these features at hand, we proceed to visualize the data through a bar chart that stacks electric and non-electric vehicle registrations by year, from 2020 to 2023. This visualization serves as a cornerstone of our EDA, offering a foundational understanding of the overall composition and trend of vehicle registrations within the state.

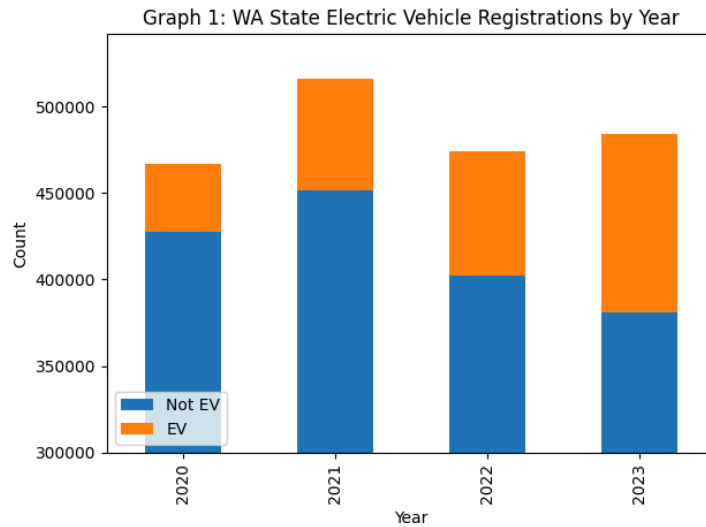


Figure 1 WA State Electric Vehicle Registrations by Year

The figure 1 prominently features the count of electric versus non-electric vehicle registrations across four years. The visualization clearly depicts a consistent increase in the number of registrations for electric vehicles, contrasting with the total registrations that exhibit some fluctuation across the years. This rise in electric vehicle registrations highlights a significant shift towards more sustainable transportation options, reflecting broader consumer acceptance and possibly the impact of incentivizing policies.

3.3 Market Share Analysis

Now we focused on all the vehicle manufactures and made a pie chart of all registered vehicles across Washington State. Due to the large amount of brands are hard to be shown, here are top 5 registered brands and their ratios :

Toyota as the market leader, holding a 31.9% share, followed by Tesla at 22.3%, and other notable brands like Honda, Ford, and Kia with 8.0%, 5.5%, and 5.1%, respectively. Based on the data on the distribution of vehicle registration brands, it is evident that Toyota and Tesla dominate the market share, both brands have so many well-known electric vehicle models. Therefore, in subsequent analysis, we plan to focus on examining the market share of these two brands and their respective models, as well as their trends over time.

A subsequent deep dive into the sales proportion of the main models from these brands provided us with a clear indication of consumer preferences within the EV sector. Tesla's Model Y and Toyota's RAV4 emerge as the most popular models, holding 21.3% and 20.5% of their respective brand's total registrations. Following are Tesla's Model 3 and Toyota's Prius, highlighting the market's inclination towards specific EV models due to factors which may include technological innovation, environmental impact considerations, and economic incentives.

After understanding the market share of different car brands and their models, we also want to explore other factors that may influence market share starting from the customers themselves. Given the limited attributes in the dataset, we hope to gain deeper insights through geographical data. For instance, we need to investigate the overall preference for electric vehicles across Washington State next.

We found that electric vehicles constituted approximately 54% of all registration vehicles, a strong indicator of the high rate of electric cars. However, counties can still have multiple preferences for different vehicle brands and types.

Therefore, we would like to gauge the impact of county-level preferences on vehicle selection. We constructed a heatmap contrasting counties with the manufacturers of registered vehicles, our findings revealed there is a wide preference for Toyota among most counties. However, Tesla's registration surpassed those of Toyota in King County, and this might indicate that Tesla can be a popular brand for some other reasons. This phenomenon underscores the various preferences in vehicles and the significant market penetration of electric vehicles.

3.4 Electrification Level Distribution Analysis

Because of the multiple sorts of electric vehicles, we generated another pie chart to elucidate the market share of each category. As figure 2 shown, the Battery Electric Vehicles (BEVs) lead with a 37.5% market share, making them as the most preferred type among electric vehicles.

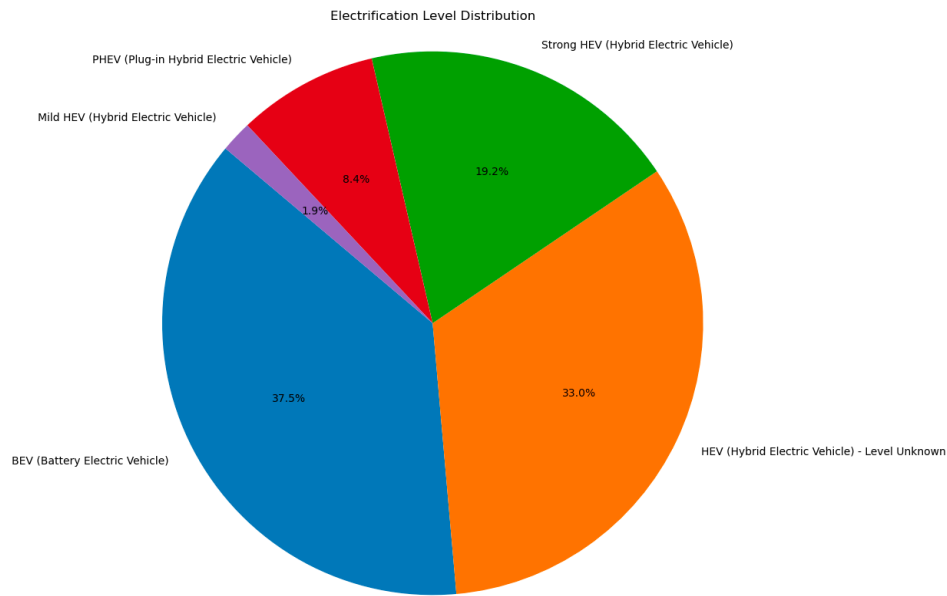


Figure 2

3.5 Trends Analysis of Multiple Electrification Level Vehicles

The time series chart of vehicle sales (figure 3) delineates the transaction trends for different types of electric vehicles, excluding ICE (Internal Combustion Engine) & non-powered vehicles. It's evident that BEVs (Battery Electric Vehicles) show a significant growth trajectory, signaling substantial market potential.

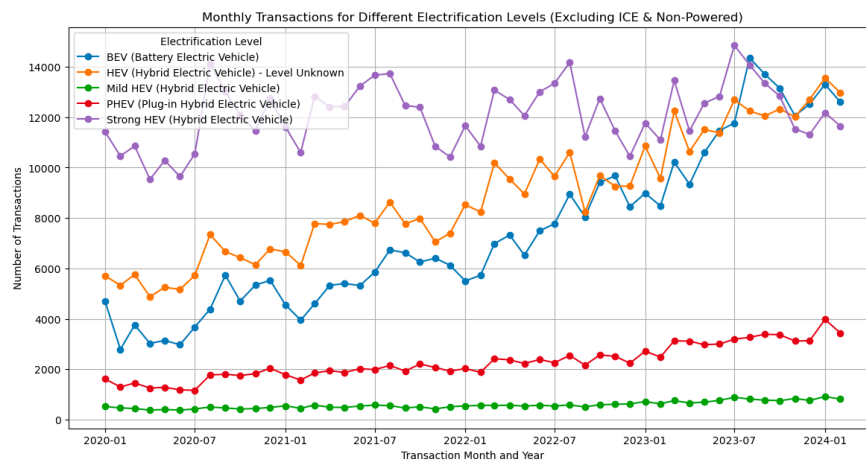


Figure 3 Monthly Transactions for Different Electrification Levels (Excluding ICE & Non-Powered)

3.6 Transaction Trend Analysis of Major Brands

To discern trends in vehicle sales, we plotted the annual transaction data for two major automotive manufacturers, Toyota and Tesla. The figure 4 delineates their sales trajectory from 2020 to 2023, revealing a significant, steady increase in Tesla's sales year over year. This upward trend for Tesla shows its better popularity as the year went, guiding our research focus toward understanding the factors contributing to Tesla's consistent sales growth.

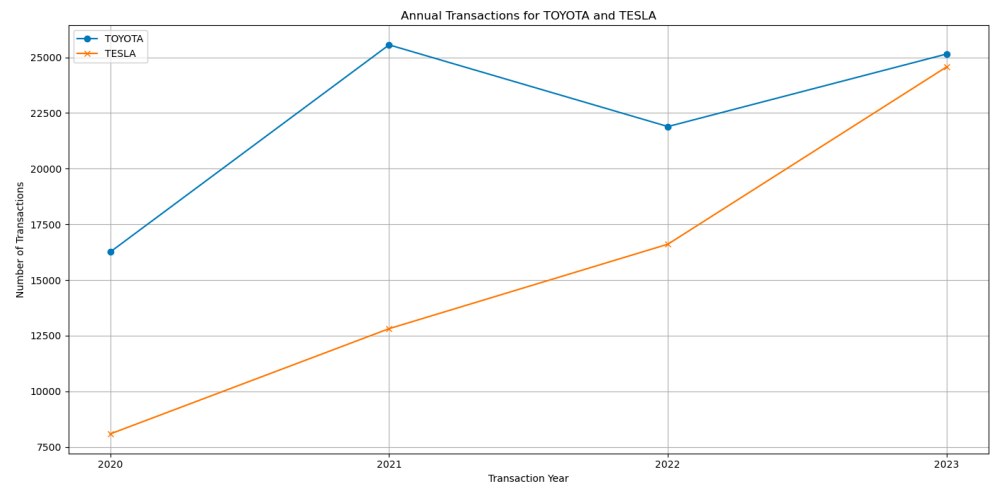


Figure 4 Annual Transactions for TOYOTA and TESLA

3.7 Analysis of Tesla’s Sales Trends

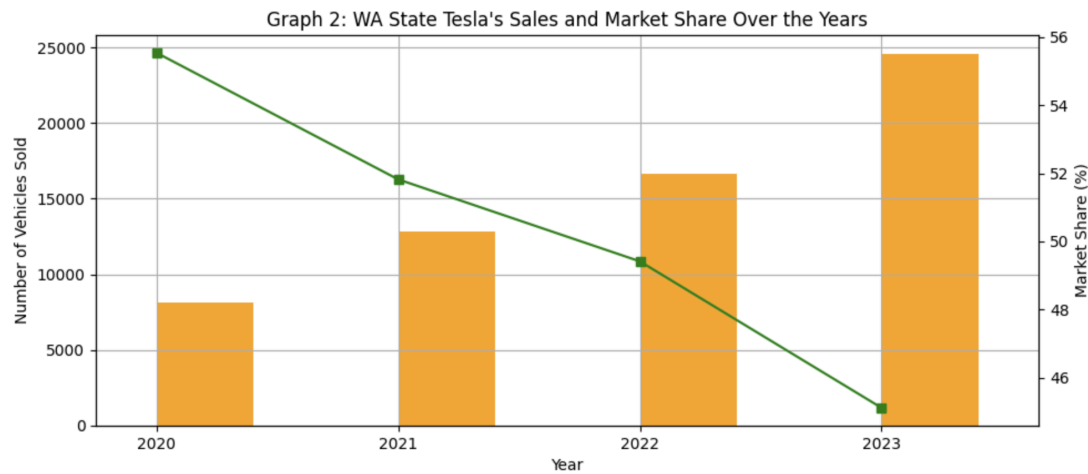


Figure 5 WA State Tesla’s Sales and Market Share Over the Years

The figure 5 offers a comprehensive view of Tesla's sales performance relative to its market share in the electric vehicle sector within Washington State from 2020 to 2023. This dual-axis bar and line chart captures the number of vehicles sold by Tesla each year (represented by orange bars) alongside Tesla's market share percentage (depicted by the green line).

Market Expansion: The increasing sales volume alongside a declining market share highlights the rapid expansion of the overall electric vehicle market in Washington State. This suggests that new entrants and existing competitors are intensifying efforts and possibly gaining traction against Tesla.

Brand Performance: The growth in Tesla's sales volume despite losing market share points to Tesla's strong brand presence and consumer loyalty. However, the decrease in market share also indicates that Tesla is facing significant challenges in maintaining its dominance as the market becomes more competitive.

3.8 Analysis of WA State Market Share

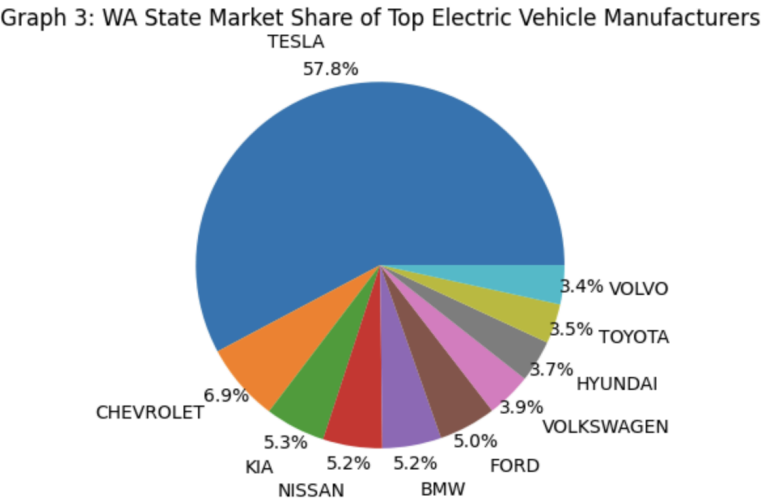


Figure 6 WA State Market Share of TOP Electric Vehicle Manufacturers

This competitive analysis reveals a diverse group of manufacturers competing for market share in the Washington State electric vehicle market. While Tesla holds a significant lead, the

presence of traditional automotive giants and newer entrants indicates a vibrant and competitive market. Each competitor brings different strengths to the table, such as Chevrolet’s affordability, BMW’s luxury appeal, and Ford’s performance orientation.

The spread of market shares among these manufacturers suggests that consumers in Washington State have a wide range of preferences and priorities when choosing electric vehicles. Tesla's dominance is significant, but the competition is robust, offering various options in terms of price, performance, style, and technology. This diversity fosters a healthy competitive environment that drives innovation and benefits consumers.

3.9 Analysis for Tesla’s Model Sales

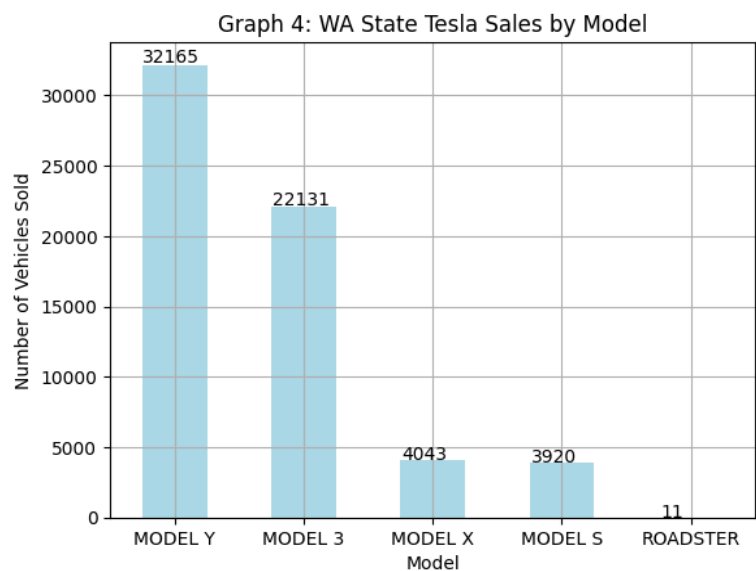


Figure 7 WA State Tesla Sales by Model

The sales data from Tesla's various models reveals significant consumer preferences within their lineup. Tesla's Model Y and Model 3 lead in popularity, demonstrating a robust preference for more affordable and practical electric vehicles. These models combine high performance with relative affordability, making them accessible to a broader demographic. In 2023, the Model Y and Model 3 accounted for the majority of Tesla's sales in Washington State, with the Model Y selling 32,165 units and the Model 3 selling 22,131 units. This reflects strong market demand for vehicles that balance cost with great features and sustainability.

In stark contrast, Tesla's luxury models, the Model X and Model S, along with the niche Roadster, exhibit significantly lower sales volumes. For example, the Model X and Model S sold only 4,043 and 3,920 units respectively in 2023, while the Roadster, aimed at a niche market, sold a mere 11 units. These figures underscore the limited appeal of high-priced models, which, despite offering advanced technology and superior performance, remain inaccessible to the average consumer due to their high cost.

The detailed EDA of electric vehicle (EV) market trends in Washington State, particularly focusing on Tesla's sales performance and market positioning relative to its competitors, provides critical insights for stakeholders within the automotive industry. Tesla's dominance in the market with a significant market share of 57.8% underlines its status as a market leader. However, the intricate dynamics of increasing sales alongside a decreasing market share indicate a rapidly evolving market landscape where competition is intensifying and consumer preferences are shifting.

The examination of Tesla's sales by model—the high sales figures for the Model Y and Model 3 versus the more modest sales of the luxury Model S and Model X, and the niche Roadster—highlights a clear consumer preference for more affordable and practical EVs. This trend suggests a broader market potential for vehicles that combine cost-effectiveness with the innovative features and sustainability that electric vehicles offer.

4. Methods and Analysis

Through walking through the visualization results, it can be found that different counties have different preferences for electric vehicles. A deeper analysis needs to be applied to better explain the phenomenon and discover the underlying logic. Therefore, this research plans to analyze the preferences of different counties through SVM and cluster analysis methods to get a more explainable result and discuss the potential problems electric vehicles still face in the market.

4.1 Variables Definition and Data Processing

The desired data will be organized in the directions which are detailed explained in [Appendix A]. There are three main categories of the data, electric vehicle holdings, preferences for brands and models, and transactions. These can be seen as important indicators of market preference for electric vehicles.

Since the original data are not formatted in the desired format, one is the transactions of all kinds of vehicles in WA, and the other one is the existing EVs in WA, we need to apply further aggregation to the data. The detailed steps are shown in [Appendix B].

After processing the data, we finally obtained the data processing results containing the following variables:

*'County', 'Total_Transactions', 'Electric_Vehicle_Count',
'Second_Electric_Vehicle_Count', 'Make_Model_Count', 'EV_Annual_Growth_Rate',
'Passenger_Ratio', 'Registration at Time of Transfer', 'Registration Renewal',
'Original_Registration', 'Major_Brand_Proportion_2023',
'Major_Model_Proportion_2023', 'BEV', 'PHEV'*

4.2 Methodology

To navigate the pattern of EV preference in each counties, we will apply SVM and K-means Clustering to the processed data.

4.2.1 SVM

Support Vector Machine (SVM) is a powerful supervised learning model used to solve classification and regression problems, optimizing model performance by constructing decision boundaries with maximum spacing between categories.

The target variable will be defined as “*High_EV_Adoption*”, which means whether a county has enough electric vehicle transactions in the overall transactions. We use 75% as the boundary, and those above 75% are defined as high EV adoption, and vice versa. The dependent variable is

defined as “*Total_Transactions*”, “*Passenger_Ratio*”, “*Major_Brand_Proportion_2023*”, “*Major_Model_Proportion_2023*”, and “*EV_Annual_Growth_Rate*”.

We will use SVM with L2 regularization to regress the data. We first do a train-test split with an 80-20 ratio, and given the hyper-parameter of $cv=5$, $penalty="l2"$, $solver="lbfgs"$. We use `StandardScaler` to normalize the data in order to eliminate the influence of order of magnitude.

4.2.2 K-means Clustering

K-means clustering is a widely used unsupervised learning algorithm that divides data points into K clusters to minimize the sum of distances from each point to its designated cluster center for data grouping. The clustering variables will be processed data without "county", and all other variables will be used. We have determined the optimal number of clusters based on elbow law as 3, which serves as a hyperparameter for clustering analysis of the data. Here, we will also use scaled data for training.

5. Results

5.1 SVM Results

In our analysis utilizing the Support Vector Machine model, our results are shown in Table 1. We have identified three key factors that can significantly affect the adoption of county-level electric vehicles in Washington State: total vehicle transactions, proportion of major brands in 2023, and proportion of major vehicle models in 2023.

Feature	Coefficient
0 Total_Transactions	0.6889
1 Passenger_Ratio	0.0686
2 Major_Brand_Proportion_2023	0.9321
3 Major_Model_Proportion_2023	0.9524
4 EV_Annual_Growth_Rate	-0.1290

Table 1 Coefficient of SVM

We found that higher transaction volumes are positively correlated with the popularity of electric vehicles, indicating that economically prosperous or populous counties are more inclined to adopt electric vehicles.

And there is also a high correlation between the major brands and models proportion. This phenomenon can be explained from two aspects. Firstly, in counties where electric vehicles are more accepted, people may have a stronger understanding of electric vehicles and have more money to support them in choosing mainstream brands and popular models. Brand promotion may be more effective in these areas; Secondly, it is also possible that the electric vehicle sales network is more widespread in these regions, and people may have better opportunities to purchase electric vehicles and receive better after-sales and service compared to other regions. These factors may affect people's acceptance of electric vehicles.

On the contrary, the annual growth rate of electric vehicles is negatively correlated, indicating that the growth rate of regions where electric vehicles have already been established has decreased due to base effects. However, in regions with relatively few initial electric vehicles and high growth rates, there are challenges in achieving significant market penetration.

The passenger ratio is weakly correlated with the adoption of electric vehicles. A higher proportion of passengers may indicate more household or individual vehicle usage, while this weak correlation indicates that people's needs are less affected by these factors, or that the base of electric vehicle transactions is small, which affects this coefficient.

5.2 K-means Clustering

Through K-means clustering analysis of the data, we obtained three types of center point data, as shown in [Appendix C Table 2 Central Points for K-means Clustering].

The clustering analysis results show that counties are divided into three main categories: unpopular, rapidly developing, and highly developed. The characteristics of each category are as follows:

Unpopular Market: Comprises 61 counties with an average electric vehicle growth of zero, likely reflecting remote or sparsely populated areas with minimal new car purchases, none of which include mainstream brands or models.

Rapidly Developing Market: Includes 37 counties, characterized by a high growth rate in electric vehicle adoption and a larger volume of new car purchases compared to the other types, predominantly of mainstream brands.

Highly Developed Market: Consists of only King County, where electric vehicles are widely accepted and have a significant existing base. Similar to the rapidly developing market, this category primarily purchases mainstream brands.

6. Conclusion

This comprehensive study has effectively mapped the dynamic landscape of the electric vehicle (EV) market in Washington State from 2020 to 2024, with a specific focus on Tesla Inc.'s substantial influence within the sector. Utilizing machine learning techniques - SVM prediction and K-means clustering, we analyzed vehicle ownership and transaction data to delineate various market patterns and identify crucial factors influencing EV adoption across different counties.

Tesla's market leadership position

Tesla Inc. has emerged as a predominant force in the EV market, exemplified by a remarkable surge in its stock value—peaking at twenty times its lowest point in 2019. This growth is not only a testament to Tesla's innovative capabilities but also reflects the high investor confidence that continues to propel the company's leadership in the EV industry. Tesla's clear dominance in the market is further highlighted by its substantial market share and the popularity of its various models, particularly the Model Y and Model 3, which cater effectively to consumer demand for affordability and sustainability.

Market segmentation insights

Our analysis identified three distinct market segments within Washington State: the Unpopular

Market, the Rapidly Developing Market, and the Highly Developed Market. Each of these segments exhibits unique characteristics that are crucial for tailoring marketing strategies and resource allocation. For instance, the Rapidly Developing Markets are ripe for targeted interventions to accelerate EV adoption, characterized by a surge in vehicle registrations and a growing consumer base that is increasingly leaning towards electric mobility solutions.

Highly Developed Markets, such as King County, show a mature EV market where Tesla has a significant footprint and consumer preferences are well-established. This market segment benefits from advanced infrastructure and greater consumer awareness, which supports sustained growth and deeper market penetration of new EV models.

Key factors for development

Through the use of SVM and K-means clustering in our methodology, we have been able to not only predict market trends but also cluster different counties based on their adoption rates and market characteristics. This has allowed us to provide actionable insights that can help decision-makers and businesses optimize resource allocation and enhance policy effectiveness in the emerging electric vehicle industry. For example, the region of Cluster 1 may be an ideal location for promoting new energy vehicle models and building charging infrastructure, while Cluster 0 may require more attention to activities that enhance residents' awareness and acceptance of electric vehicles.

In conclusion, this study provides a deep understanding of the electric vehicle market's evolution over the recent years and Tesla's pivotal role in shaping its trajectory. The insights gained from our analysis highlight the importance of strategic market segmentation and the need for adaptive strategies that cater to the varying levels of market development across different regions. As the EV market continues to evolve, these insights will be invaluable for stakeholders looking to navigate the complexities of an expanding and increasingly competitive market.

References

- Washington State Department of Licensing. (n.d.). *Vehicle registration transactions by Department of Licensing subagency*. Retrieved April 10, 2024, from <https://data.wa.gov/Transportation/Vehicle-Registration-Transactions-by-Department-of/b rw6-jymh/>
- Washington State Department of Licensing. (n.d.). *Electric vehicle population data*. Retrieved April 10, 2024, from https://data.wa.gov/Transportation/Electric-Vehicle-Population-Data/f6w7-q2d2/about_data
- United States Zip Codes. (n.d.). *ZIP code database*. Retrieved April 10, 2024, from <https://www.unitedstateszipcodes.org/zip-code-database/>

Appendices

Appendix A

Desired Data

I. Electric Vehicle Holdings

Electric vehicles count: The total number of registered electric vehicles in each county, which shows the popularity of electric vehicles in the county.

Electric vehicles ratio: The ratio between electric and total vehicle transactions

II. Preferences For Brands And Models

Manufacturer preferences: Which electric vehicle manufacturers are favored by each county (such as Tesla, Ford, etc.)

Model preference: The popularity of specific models, such as which specific models (such as Tesla Model 3, Nissan LEAF, etc.) are more popular in various counties.

Vehicle Type Distribution: The distribution proportion of different types of electric vehicles (such as pure electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs)).

III. Transactions

Transaction Velocity: Since we do not have detailed transaction data related to each specific car, the transaction velocity can be defined as the average number of transactions of vehicles within a specific time period. For the data of the entire county, we can calculate the average value of all vehicle transactions as the flow speed of the county.

Transaction Type: This includes the number of transactions per year in each county and the type of transactions (such as new registration, transfer, or renewal).

Appendix B

Data Processing Detail

This part will explain how adjustments were made to the original data to create the required dataset for analysis.

Based on the cleaned data from Chapter 2, group the data by column “*County*”, and get the aggregated sum of “*Total_Transactions*”, “*Electric_Vehicle_Count*”, “*Second_Electric_Vehicle_Count*”, and the mean of “*Make_Model_Count*”. Then, because the original data is recorded by transactions for each year, so that groupby data by “*County*” and “*Year*”, to get the aggregated sum of “*Electric_Vehicle_Count*”, which will be the EV transactions per year, and then we calculate the average percentage change of the EV transaction in each county to get “*EV_Annual_Growth_Rate*”, we fill the place with no EV transactions growth with 0.

The data from cleaning process includes detailed information of transactions in vehicle types and transaction types for each year, but here we would like to have the value in total. We calculate the overall vehicle transaction types in 2020-2024 for each county, like passenger car, MPV, truck, bus, motorcycle and other vehicles types. By calculating the proportion of household cars (passenger car and MPV), we get the “*Passenger_Ratio*”. In the same way, for EV cars we get “*EV_Passenger_Ratio*”. Then we can count on “*Vehicle_Transaction_Type_Distribution*” to get “*Registration at Time of Transfer*”, “*Registration Renewal*”, and “*Original Registration*”.

We define the top 10 brands and models as the “Major Brand”, we calculate the mean proportion “*Major_Brand_Proportion_2023*” and “*Major_Model_Proportion_2023*” for each county.

Finally, we calculate the average number of BEV (Battery Electric Vehicle) and PHEV (Plug-In Hybrid Electric Vehicles) each year, to get “*BEV*” and “*PHEV*”.

Appendix C

Table 2 Central Points for K-means Clustering

Cluster	0	1	2
Total_Transactions	1757.17	580190.91	6986939.00
Second_Electric_Vehicle_Count	31.89	18366.17	416266.00
Make_Model_Count	28.84	1630.45	3676.40
EV_Annual_Growth_Rate	0.00	0.10	0.05
Passenger_Ratio	0.73	0.68	0.84
Registration at Time of Transfer	33.75	14220.85	138250.60
Registration Renewal	289.93	93374.53	1115775.80
Original Registration	27.53	8098.55	115044.80
Major_Brand_Proportion_2023	0.04	0.85	0.84
Major_Model_Proportion_2023	0.02	0.61	0.64
BEV	1.05	1826.94	74942.00
PHEV	0.41	591.37	17798.00
EV_Proportion	0.00	0.01	0.04
EV_High_Low	0.00	0.69	1.00
Count	63	35	1