FINAL PROJECT REPORT

PROJECT ASSIGNMENT - 4

DATA ANALYTICS RESEARCH PROJECT

JAGADEESH VARMA GADIRAJU

G01465303

GUIDE: Prof. Alla Webb

Electric Vehicle Population Size History By County

Abstract:

This dataset provides a detailed monthly summary of vehicle registration statistics in Washington State. It is derived from a synergistic combination of data from the Environmental Protection Agency (EPA), the National Highway Traffic Safety Administration (NHTSA), and the Washington State Department of Licensing (DOL). The dataset provides a comprehensive repository for in-depth analyses, with an emphasis on county and vehicle type breakdowns, encompassing both passenger vehicles and trucks. Through real-time investigation of vehicle registration trends, fuel efficiency metrics, and related variables made possible by the integration, research and policy initiatives can benefit greatly. Our understanding of the dynamic landscape of vehicle registrations is enhanced by the granular dissection of data at the county and vehicle type levels, which adds significantly to our understanding of patterns in fuel efficiency.

Introduction:

In the rapidly evolving automobile industry, the state of Washington is leading the way in a revolutionary transition toward sustainability. Central to this development is a large dataset that was painstakingly assembled by the Washington State Department of Licensing (DOL), which combined data from the EPA, the NHTSA, and titling/registration records. This dataset, which is a comprehensive tool for analyzing and understanding the dynamics of the state's vehicular landscape, captures monthly registration statistics for vehicles, which are further categorised by county and vehicle type.

This paper explores the patterns present in this dataset, elucidating market trends among counties and charting the evolution of electric vehicle (EV) adoption. Our goal as we examine the finer points of vehicle registration is to glean insights that go beyond simple statistical analysis. Beyond data, these insights can impact industry strategies, shape policy decisions, and most importantly raise consumer awareness of the benefits and accessibility of electric vehicles. The dynamics of electric vehicle adoption, which are examined in these pages, are essential in guiding Washington State's transportation sector toward a future that is more ecologically conscious and sustainable.

Languages Used: Python, R and SQL.

Literature Survey:

➢ This journal "Analysis of Electric Vehicle Charging Station Usage and Profitability in Germany based on empirical data" Explains profit estimation and usage of electric vehicle charging stations. The Research show that while DC fast-charging stations are more common on weekends, AC charging is more common during the day and on weekdays. Because they have better profit margins, fast chargers serve about three times as many vehicles per charge point and make more money. Fast chargers deliver about 40 kWh of energy in a single charging event, whereas AC chargers typically deliver about 20 kWh. For AC chargers, energy transfer usually ends after 4 hours, and for fast chargers, it ends after 45 minutes. Additionally, for AC charging, the actual power rates seldom surpass 11 kW and are constantly lower than the rated station power. This gives a brief idea about the analysis. This paper makes it possible for other researchers to validate models or create simulations and test scenarios using the data that is provided[2]. This research is similar to the research question they are about trends tools place in Germany.

➢ Another similar work that I've found is "Market Penetration Analysis of Electric Vehicles in the German Passenger Car Market Towards 2030". Research indicates that external factors play a major role in the market viability of alternative powertrain vehicles. Alternative powertrains won't hit the market in the absence of purchase price incentives, rising oil prices, and cheap energy costs for electricity and hydrogen. On the other hand, if OEMs take on negative markups during the first ten years following product introduction, they may help to boost the market[3]. This research is similar to my research question in that they are about the first generation of electric vehicles which takes place in Germany.

➢ Another similar work that I've found "The impact of seasonal and geographical factors on the performance of electric car batteries" This article says Temperature, altitude, humidity, and other variables can all have an effect on how well electric car batteries perform. EV drivers can take steps to maximize their battery performance in different environments, even though these factors cannot be completely eliminated. Drivers of electric cars can enjoy their vehicles with confidence, no matter where their adventures take them, by being aware of how geography affects EV batteries and taking the necessary steps to mitigate these effects[4]. This will answer my research question in terms of seasonal or cyclic patterns in vehicle registrations.

Dataset:

The data presented here represents the monthly registration statistics for vehicles in Washington State, categorized by county and vehicle type, which includes passenger vehicles and trucks. This data is compiled through the integration of information from the National Highway Traffic Safety Administration (NHTSA) and the Environmental Protection Agency (EPA) fuel efficiency ratings with titling and registration data maintained by the Washington State Department of Licensing (DOL).

In essence, the Washington State DOL combines data from these various sources to provide comprehensive and up-to-date information on vehicle registration trends. This allows for the analysis and monitoring of vehicle registration patterns, fuel efficiency, and other related aspects, broken down by both county and vehicle type, to support various research and policy initiatives.

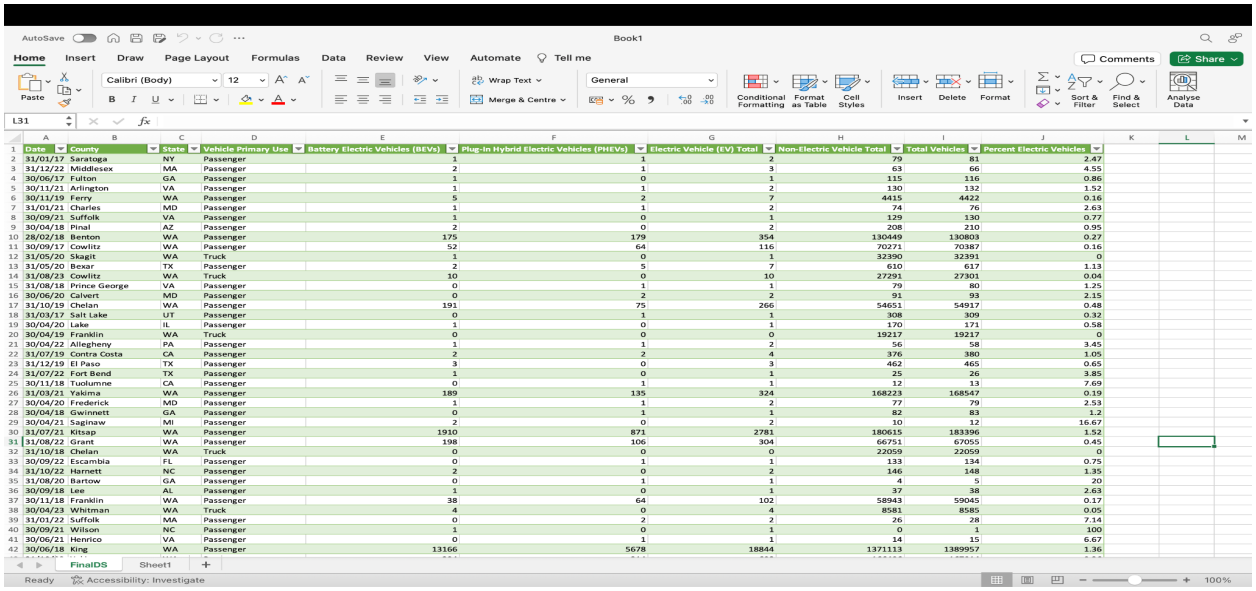
Metadata:

Resource Type	Dataset
Publisher	data.wa.gov
Maintainer	Department of Licensing
Identifier	https://data.wa.gov/api/views/3d5d-sdqb
Category	Transportation
Public Access Level	public
Metadata Context	https://project-open-data.cio.gov/v1.1/schema/catalog.jsonld
Metadata Catalog ID	https://data.wa.gov/data.json
Schema Version	https://project-open-data.cio.gov/v1.1/schema
Catalog Described by	https://project-open-data.cio.gov/v1.1/schema/catalog.json
Harvest Object Id	7c4ea445-6ffe-49cb-9ab8-00f47fd054f0
Harvest Source Id	1c33d1cb-2f69-4f1b-835e-453790f38dc7
Harvest Source Title	WA JSON

Homepage URL	https://data.wa.gov/d/3d5d-sdqb
License	http://opendatacommons.org/licenses/odbl/1.0/
Source Datajson Identifier	True
Source Hash	4a590b80cd326a060b0bc821cd5c7c9bfc5d4edafb76833b9d1b283e927663d6
Source Schema Version	1.1

Data Cleaning:

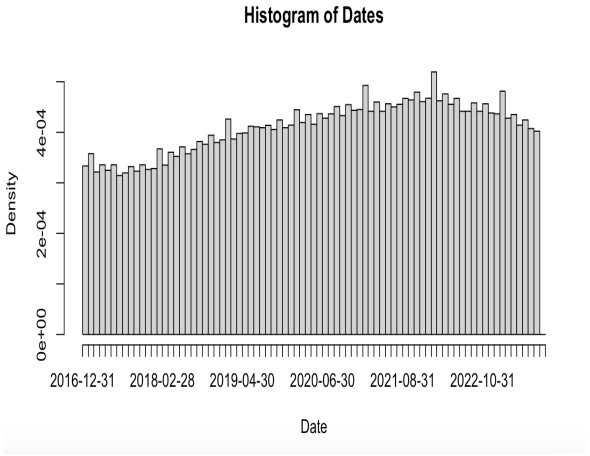
The first step before performing any actions on the dataset is data cleaning if the data is messy. This is one of the important steps. Before loading this data set to any tool I imported this dataset to Microsoft Excel and then started cleaning the data. I've removed null columns and filled empty spaces with N/A. I've confirmed that there are no nulls and unwanted empty spaces.

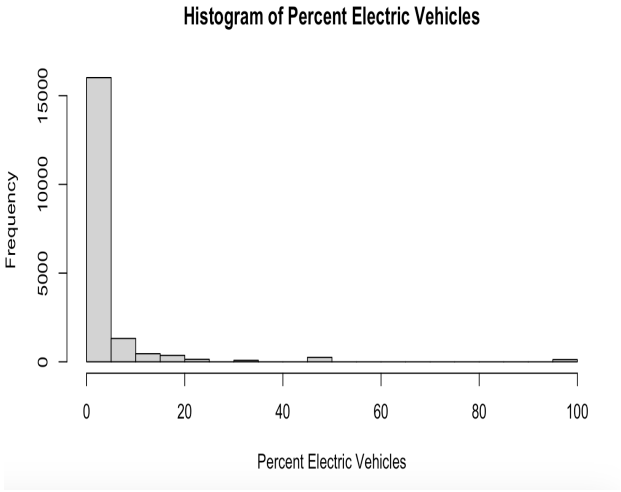


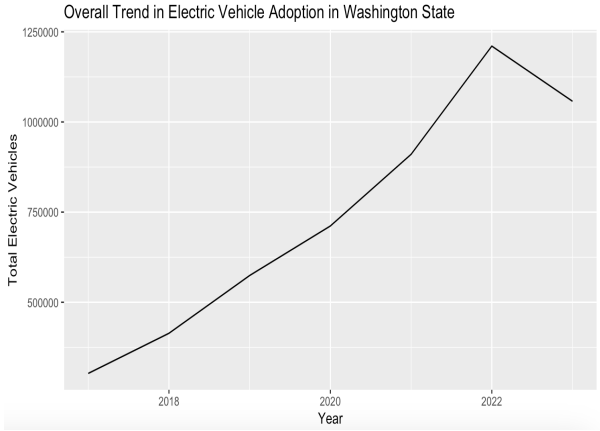
Date	County	State	Vehicle Primary Use	Battery Electric Vehicles (BEVs)	Plug-In Hybrid Electric Vehicles (PHEVs)	Electric Vehicle (EV) Total	Non-Electric Vehicle Total	Total Vehicles	Percent Electric Vehicles
31/01/17	Saratoga	NY	Passenger	1	1	2	79	81	2.47
31/12/22	Middlesex	MA	Passenger	2	1	3	63	66	4.55
30/06/17	Fulton	GA	Passenger	1	0	1	115	116	0.86
30/11/21	Arlington	VA	Passenger	1	1	2	130	132	1.52
30/11/19	Ferry	WA	Passenger	5	2	7	4415	4422	0.16
31/01/21	Charles	MD	Passenger	1	1	2	74	76	2.63
30/09/21	Suffolk	VA	Passenger	1	0	1	129	130	0.77
30/04/18	Pinal	AZ	Passenger	2	0	2	208	210	0.95
28/02/18	Benton	WA	Passenger	175	179	354	130449	130803	0.27
30/09/17	Cowlitz	WA	Passenger	52	64	116	70271	70387	0.16
31/05/20	Skagit	WA	Truck	1	0	1	32390	32391	0
31/05/20	Bexar	TX	Passenger	2	5	7	610	617	1.13
31/08/23	Cowlitz	WA	Truck	10	0	10	27291	27301	0.04
31/08/18	Prince George	VA	Passenger	0	1	1	79	80	1.25
30/06/20	Calvert	MD	Passenger	0	2	2	91	93	2.15
31/10/19	Chelan	WA	Passenger	191	75	266	54651	54917	0.48
31/03/17	Salt Lake	UT	Passenger	0	1	1	308	309	0.32
30/04/20	Lake	IL	Passenger	1	0	1	170	171	0.58
30/04/19	Franklin	WA	Truck	0	0	0	19217	19217	0
30/04/22	Allegheny	PA	Passenger	1	1	2	56	58	3.45
31/07/19	Contra Costa	CA	Passenger	2	2	4	376	380	1.05
31/12/19	El Paso	TX	Passenger	3	0	3	462	465	0.65
31/07/22	Fort Bend	TX	Passenger	1	0	1	25	26	3.85
30/11/18	Tuolumne	CA	Passenger	0	1	1	12	13	7.69
31/03/21	Yakima	WA	Passenger	189	135	324	168223	168547	0.19
30/04/20	Frederick	MD	Passenger	1	1	2	77	79	2.53
30/04/18	Gwinnett	GA	Passenger	0	1	1	82	83	1.2
30/04/21	Saginaw	MI	Passenger	2	0	2	10	12	16.67
31/07/21	Kitsap	WA	Passenger	1910	871	2781	180615	183396	1.52
31/08/22	Grant	WA	Passenger	198	106	304	66751	67055	0.45
31/10/18	Chelan	WA	Truck	0	0	0	22059	22059	0
30/09/22	Escambia	FL	Passenger	0	1	1	133	134	0.75
31/10/22	Harnett	NC	Passenger	0	2	2	146	148	1.35
31/08/20	Bartow	GA	Passenger	0	1	1	4	5	20
30/09/18	Lee	AL	Passenger	1	0	1	37	38	2.63
30/11/18	Franklin	WA	Passenger	38	64	102	58943	59045	0.17
30/04/23	Whitman	WA	Truck	4	0	4	8581	8585	0.05
31/01/22	Suffolk	MA	Passenger	0	2	2	26	28	7.14
30/09/21	Wilson	NC	Passenger	1	0	1	0	1	100
30/06/21	Henrico	VA	Passenger	0	1	1	14	15	6.67
30/06/18	King	WA	Passenger	13166	5678	18844	1371113	1389957	1.36

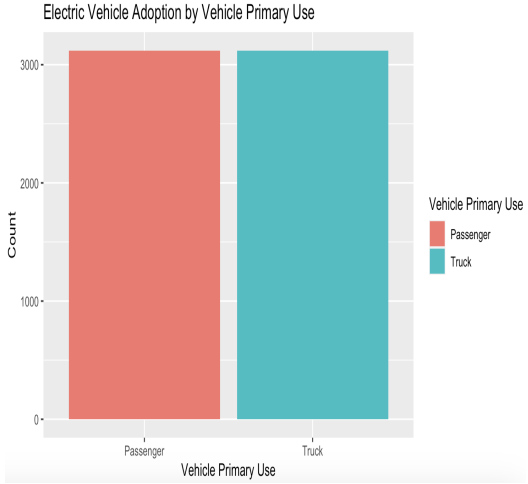
Exploratory Data Analysis:

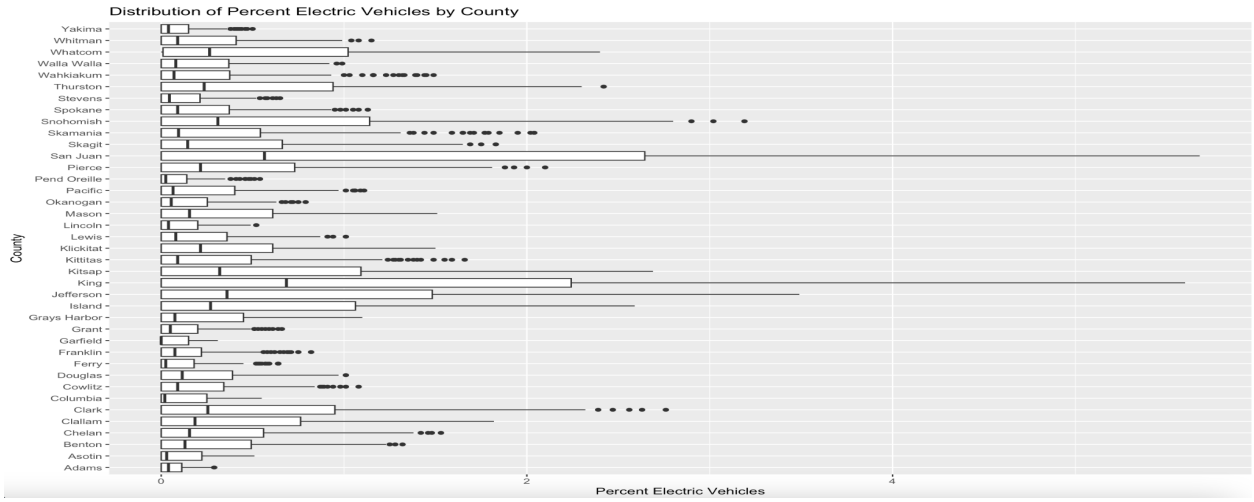
The exploratory data analysis is to analyze the data which can be used for analysis. By performing exploratory data analysis we will get some idea about the data simply so we can be familiar with this data. The figures shown below are the figures that I've done.











The above figures show/give basic information about the trends. The below figure shows a summary of statistics.

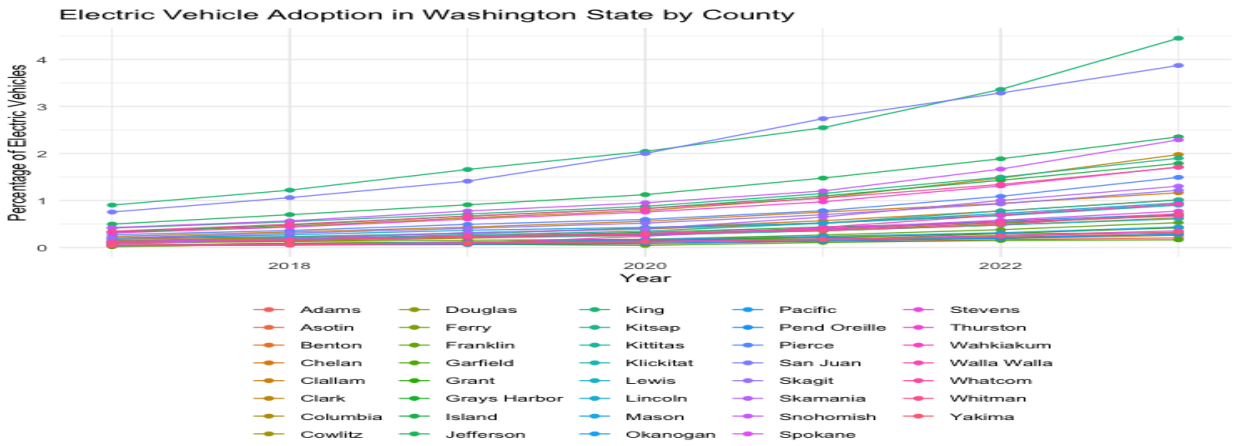
```
Battery.Electric.Vehicles..BEVs.  Plug.In.Hybrid.Electric.Vehicles..PHEVs.
Min.    :     0.0                 Min.    :     0.00
1st Qu.:     0.0                 1st Qu.:     0.00
Median :     1.0                 Median :     1.00
Mean    :   200.2                 Mean    :    77.17
3rd Qu.:     3.0                 3rd Qu.:     2.00
Max.    :61754.0                 Max.    :15166.00
Electric.Vehicle..EV..Total Non.Electric.Vehicle.Total Total.Vehicles
Min.    :     0.0           Min.    :       0        Min.    :       1
1st Qu.:     1.0           1st Qu.:      47        1st Qu.:      48
Median :     1.0           Median :     176        Median :     178
Mean    :   277.4           Mean    :   25936        Mean    :   26213
3rd Qu.:     4.0           3rd Qu.:    9188        3rd Qu.:    9200
Max.    :76920.0           Max.    :1400225        Max.    :1431343
Percent.Electric.Vehicles
Min.    :  0.000
1st Qu.:  0.350
Median :  1.150
Mean    :  3.897
3rd Qu.:  2.700
Max.    :100.000
```

I’ve done this EDA using R. This exploratory data analysis helps us to study this data.

Research Questions:

1) What is the overall trend in electric vehicle adoption in Washington State, and how does it vary by county?

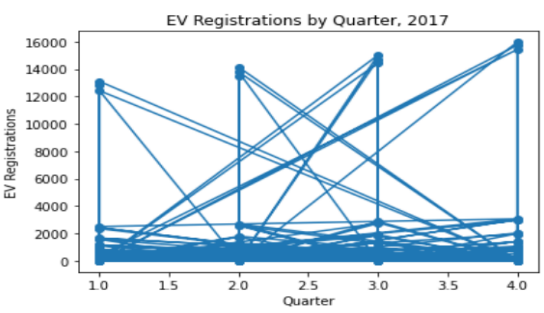
To Answer this Question I’ve used R and plotted the line graph with respect to all the counties in Washington state. The following figure is the line graph I've plotted.

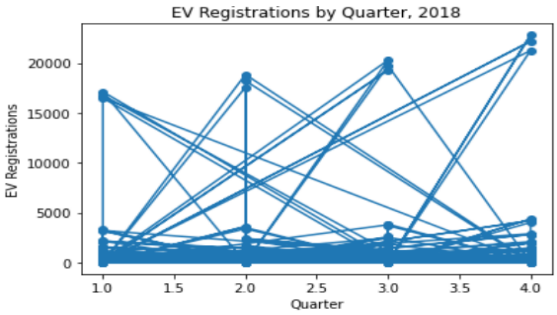


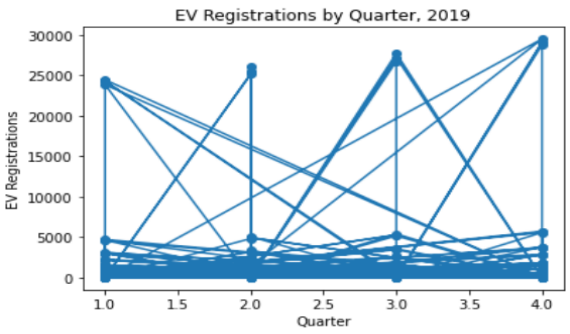
The overall trend is the electric vehicle population is increasing rapidly in each county located in Washington state. I've used R to answer this question.

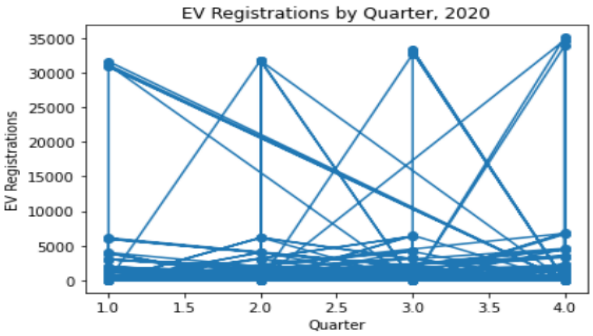
2) Can we identify any seasonal or cyclical patterns in electric vehicle registrations, and are these patterns consistent across counties?

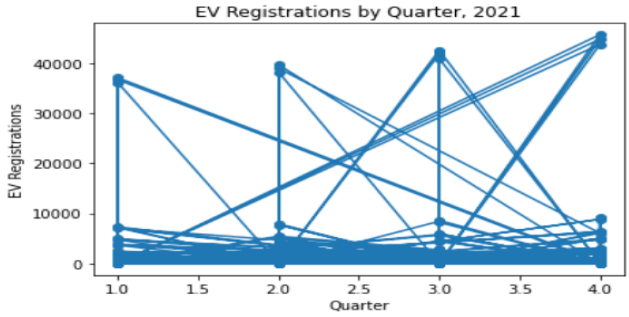
Initially, I generated a set of line graphs, each representing individual years segmented into quarters, to visualize the quarterly trends in Electric Vehicle (EV) registrations. Notably, irrespective of the specific year, the pattern emerged that EV registrations consistently increased as each year progressed, culminating in higher registration numbers by the end of each respective year. This observed trend underscores a consistent and positive trajectory in EV registrations, suggesting a year-end surge in adoption across the dataset.

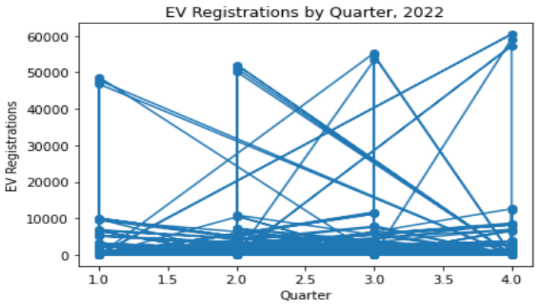


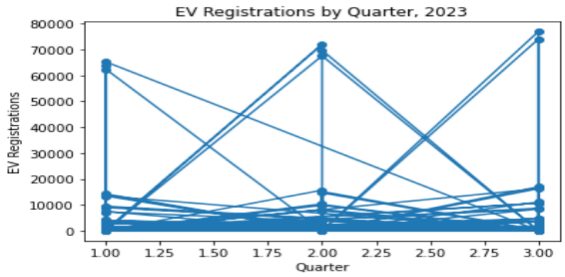




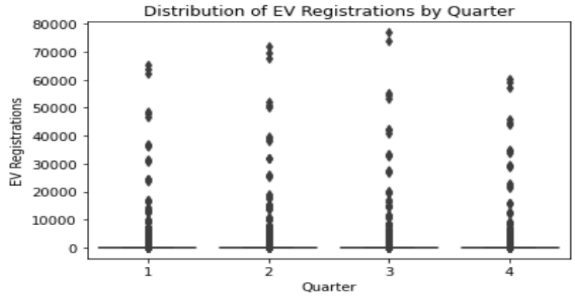




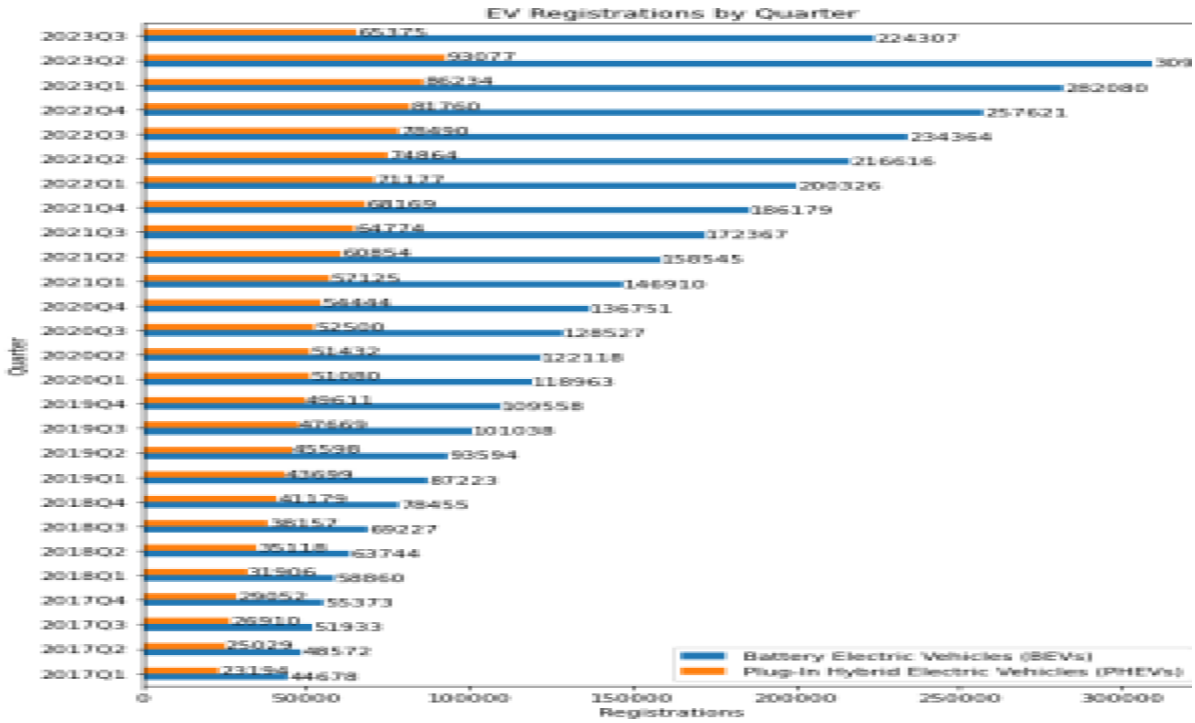




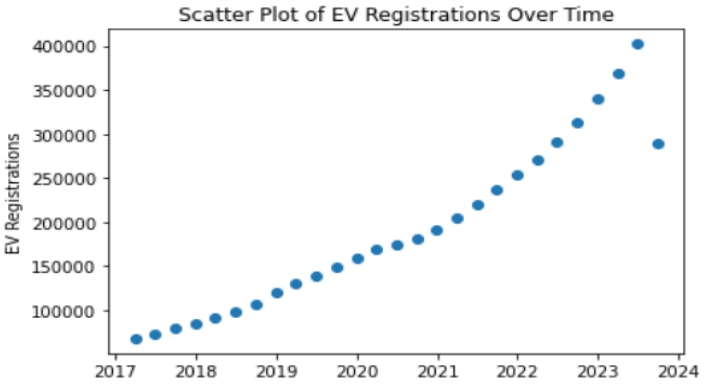
Subsequently, I created a comprehensive visualization by plotting box plots that encapsulate all the years on a single graph, organized quarterly. This graphical representation vividly highlights notable patterns in Electric Vehicle (EV) registrations throughout the quarters. Particularly, the box plots reveal a distinct surge in EV registrations during the third quarter across all the years. This observation suggests a recurring trend, signifying a concentration of EV adoption during the specific period encompassing the third quarter. The prevalence of higher registration numbers during this quarter underscores a noteworthy aspect of the dataset, inviting further exploration into the factors contributing to this recurring surge in the third quarter.



As part of the analysis, I generated insightful bar graphs for each year, categorized by quarter, to provide a detailed comparison of registrations between Plug-In Hybrid Electric Vehicles (PHEVs) and Battery Electric Vehicles (BEVs). Remarkably, the visual examination of these bar graphs unveils a compelling narrative – while there might not be discernible cyclic patterns throughout the quarters, a consistent upward trajectory is evident in each and every quarter.



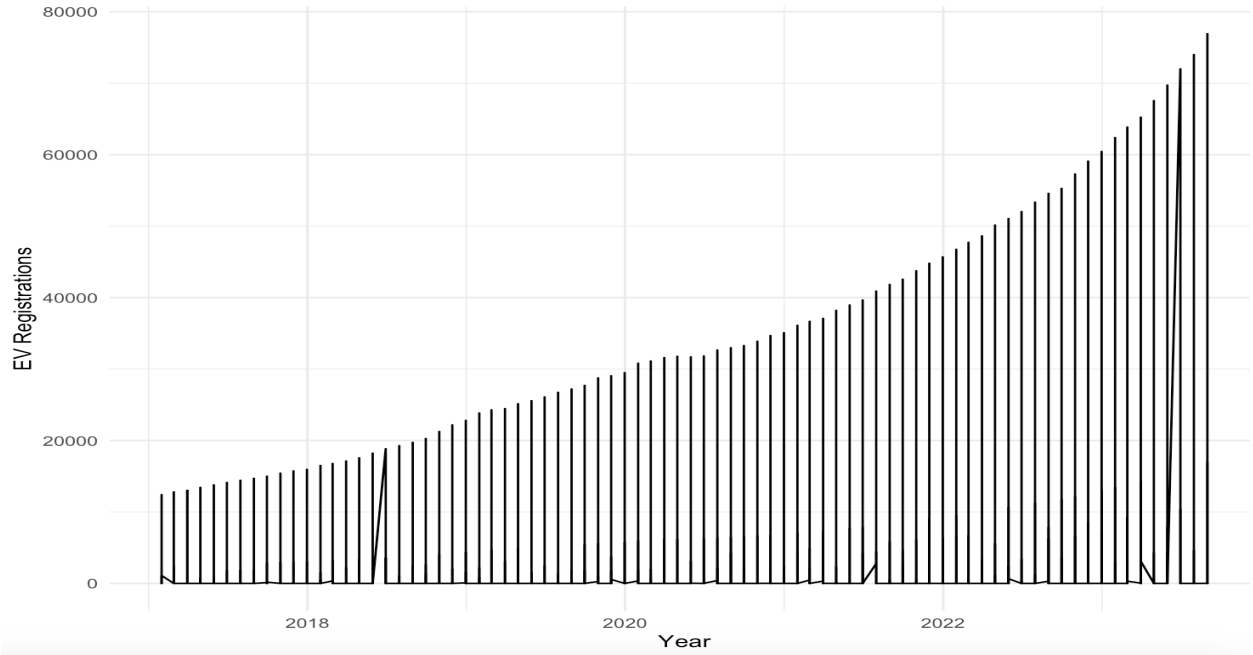
Finally, I've plotted a scatter plot of each year separately to make sure there are really any cyclic patterns in Electric Vehicle registrations.



Overall, there are no cyclic patterns in EV registrations but there is an increase in EV registrations quarter by quarter and year by year and all these patterns are consistent across counties. I've used Python to answer this question.

3) Is there any relationship between the rise/fall vehicle registrations and specific county levels?

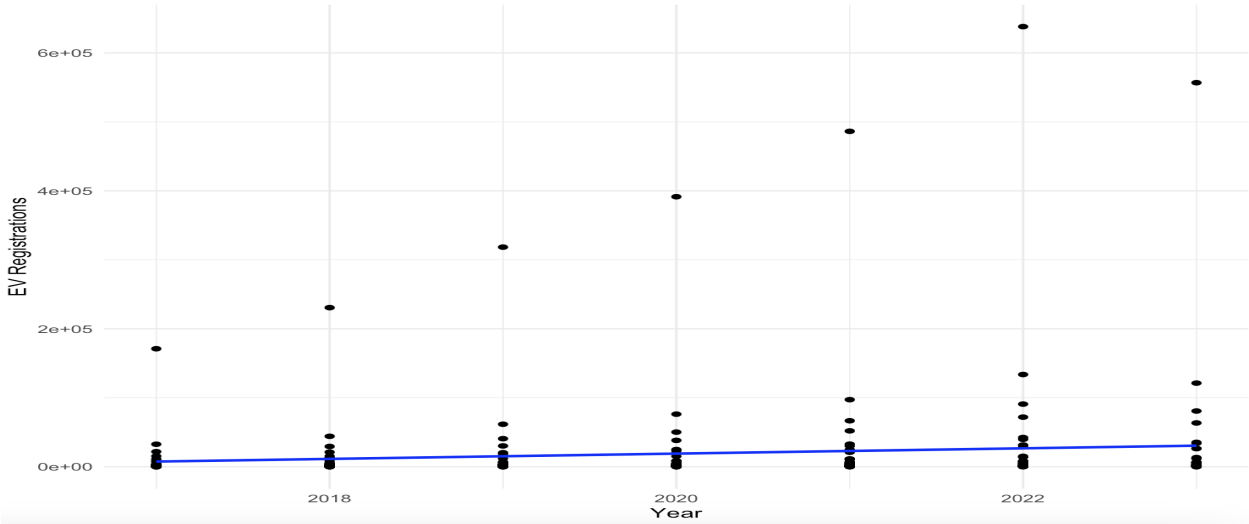
Firstly, I've plotted a line graph of registrations respective to the year to identify the Rise/Fall of the EV registrations where I can't specifically say about counties.



So I’ve printed the table for each and every county for each year separately In this keen observation I got that there is no specific Rise/Fall of registrations with respect to a specific county.

	County	2017	2018	2019	2020	2021	2022	2023
1	Adams	68	116	133	182	295	367	300
2	Asotin	121	202	246	348	422	519	439
3	Benton	3574	4690	6582	8427	11215	14999	13242
4	Chelan	1216	1950	2809	3860	5356	7281	6323
5	Clallam	1968	2850	3794	4850	6599	8161	6647
6	Clark	15372	21083	30130	38185	51994	71878	63424
7	Columbia	21	43	53	80	103	91	83
8	Cowlitz	1302	1668	2443	3172	4478	6203	5372
9	Douglas	708	927	1137	1393	1753	2376	2103
10	Ferry	32	57	77	111	159	284	197
11	Franklin	891	1177	1479	1713	2653	3795	3507
12	Garfield	5	20	19	13	31	42	30
13	Grant	567	928	1239	1669	2442	3428	3189
14	Grays Harbor	807	1319	2017	2695	3726	4534	3824
15	Island	3137	4831	6834	8483	11239	14508	11958
16	Jefferson	2005	2885	3818	4701	6196	7854	6447
17	King	171050	230671	318415	391473	486301	638167	556856
18	Kitsap	11282	15344	20273	24881	32731	42307	35280
19	Kittitas	529	969	1426	1870	2786	4180	3568
20	Klickitat	551	653	848	1160	1489	1916	1694
21	Lewis	1024	1619	2344	2651	3714	4979	4171
22	Lincoln	80	94	127	183	244	315	277
23	Mason	1492	2141	2990	3527	4331	5980	5124
24	Okanogan	250	360	572	881	1234	1625	1488
25	Pacific	223	450	640	784	1141	1628	1334
26	Pend Oreille	75	111	131	163	255	345	337
27	Pierce	21750	29346	40660	50158	66622	90798	80767
28	San Juan	1749	2550	3479	4889	6801	8165	6328
29	Skagit	2943	4508	6122	7785	10519	13774	11847
30	Skamania	194	282	488	603	969	1473	1251
31	Snohomish	32601	44072	61569	76200	97215	133676	121219
32	Spokane	5705	7977	11786	15111	20970	30009	26366
33	Stevens	356	441	617	841	1069	1507	1380
34	Thurston	8992	12928	18503	23115	29955	39749	33957
35	Wahkiakum	22	29	50	138	253	408	358
36	Walla Walla	600	939	1335	1699	2338	3286	2647
37	Whatcom	7735	11318	15638	19378	24762	31429	26364
38	Whitman	370	573	919	1113	1322	1813	1769
39	Yakima	1299	1861	2521	3137	4599	6591	5926

In response to the absence of distinct patterns in Electric Vehicle (EV) registrations across specific counties, linear regression was employed to identify the overall trend. This analytical approach allowed for a succinct quantification of the rise or fall in EV registrations over time, transcending localized variations. The resulting insights contribute to a comprehensive understanding of the collective dynamics influencing the trajectory of EV adoption.



Overall, there is no specific Rise/Fall in registrations with respect to a specific county and all the counties that are analyzed are identical. I've used R to answer these question.

Working with SQL:

Firstly I've imported the interior data set on My SQL workbench.

Date	County	State	Vehicle Primary Use	Battery Electric Vehicles (BE...	Plug-In Hybrid Electric Vehicles (PHE...	Electric Vehicle (EV) T...	Non-Electric Vehicle T...	Total Vehicles	Percent Electric Vehicl...
30/06/23	King	WA	Passenger	57391	14609	72000	1317849	1389849	5.18
31/12/22	King	WA	Passenger	48187	12254	60441	1337897	1398338	4.32
30/06/18	King	WA	Passenger	13166	5678	18844	1371113	1389957	1.36
31/07/21	Kitsap	WA	Passenger	1910	871	2781	180615	183396	1.52
31/05/20	Benton	WA	Passenger	354	321	675	134156	134831	0.5
31/05/20	Skagit	WA	Passenger	400	216	616	91167	91783	0.67
31/10/22	Yakima	WA	Passenger	394	214	608	166406	167014	0.36
30/04/22	Yakima	WA	Passenger	318	183	501	167410	167911	0.3
28/02/21	Jefferson	WA	Passenger	295	168	463	24787	25250	1.83
28/02/23	Kittitas	WA	Passenger	321	92	413	29520	29933	1.38
28/02/18	Benton	WA	Passenger	175	179	354	130449	130803	0.27
31/01/20	San Juan	WA	Passenger	265	83	348	14400	14748	2.36
31/03/21	Yakima	WA	Passenger	189	135	324	168223	168547	0.19
31/08/22	Grant	WA	Passenger	198	106	304	66751	67055	0.45
31/10/19	Chelan	WA	Passenger	191	75	266	54651	54917	0.48
31/05/17	Island	WA	Passenger	113	129	242	59916	60158	0.4
30/06/21	Franklin	WA	Passenger	123	97	220	62011	62231	0.35
31/03/22	Douglas	WA	Passenger	125	56	181	28529	28710	0.63

Performed a few SQL queries.

The first Query is to perform grouping by "County" and "Registration_Date" while calculating the total electric vehicle registrations.

County	Registration_Da...	Total_EV_Registratio...
Alameda	31/08/17	1
Allegheny	30/04/22	2
Allegheny	31/05/22	2
Arlington	30/11/21	2
Asotin	30/06/19	0
Asotin	30/06/20	0
Asotin	31/05/18	0
Bartow	31/08/20	1
Benton	28/02/18	354
Benton	31/05/20	675
Bexar	31/01/17	6
Bexar	31/05/20	7
Broward	30/09/20	2
Broward	31/08/20	3
Calvert	29/02/20	2
Calvert	30/06/20	2
Cape May	30/04/22	1
Carroll	31/01/22	1
Carson…	30/11/18	1
Charles	31/01/21	2
Chelan	30/04/20	0
Chelan	31/03/18	0
Chelan	31/07/17	0

The Second Query is to perform grouping by "County" and "Registration_Date" while calculating the percentage of electric vehicles.

	County	Registration_Da...	EV_Percent
	Alameda	31/08/17	0.28
	Allegheny	30/04/22	3.45
	Allegheny	31/05/22	3.51
	Arlington	30/11/21	1.52
	Asotin	30/06/19	0
	Asotin	30/06/20	0
	Asotin	31/05/18	0
	Bartow	31/08/20	20
	Benton	28/02/18	0.27
	Benton	31/05/20	0.5
	Bexar	31/01/17	0.48
	Bexar	31/05/20	1.13
	Broward	30/09/20	1.77
	Broward	31/08/20	2.63
	Calvert	29/02/20	2.15
	Calvert	30/06/20	2.15
	Cape May	30/04/22	12.5
	Carroll	31/01/22	16.67
	Carson...	30/11/18	6.25
	Charles	31/01/21	2.63
	Chelan	30/04/20	0
	Chelan	31/03/18	0
	Chelan	31/07/17	0

Future work:

We plan to explore the commercial landscape of electric vehicle (EV) sales and the spread of EV charging stations in future analyses. We aim to examine market trends, consumer preferences, and factors that influence the adoption of electric vehicles (EVs) from a commercial perspective. In addition to offering businesses insightful information, this investigation will support well-informed decision-making in the rapidly changing electric vehicle sector. Furthermore, one of our next standards is to apply machine learning models to improve prediction abilities. We hope to forecast trends, predict consumer behaviour, and streamline strategic decision-making procedures by utilizing advanced analytics. The dataset's intricate patterns will be largely unlocked by machine learning algorithms, allowing us to find hidden correlations and make data-driven predictions. Beyond projections, we will investigate additional aspects, including the need for infrastructure, the effects of regulations, and how EV adoption interacts with larger sustainability programs. We hope to make a significant contribution to the current conversation on sustainable transportation through these upcoming projects and help different sectors make well-informed decisions.

Conclusion:

Overall, This study Explains about irrespective of the county the population of electric vehicles is increasing rapidly in all counties, There are no cyclic patterns in EV registrations but there is an increase in EV registrations quarter by quarter and year by year and all these patterns are consistent across counties and There is no specific county levels that can impact on rise/fall of electric vehicles.

List of terms:

EVs- Electric vehicles
BEVs - Battery Electric Vehicles
PHEVs- Plug-In Hybrid Electric Vehicles
OEMs - Original equipment manufacturer
EDA- Exploratory Data Analysis

References:

[1] Electric Vehicle Population Size History By County - Catalog. (2023b, November 17).**https://catalog.data.gov/dataset/electric-vehicle-population-size-history-by-county**

[2]Hecht, C., Figgener, J., & Sauer, D. U. (2022). Analysis of electric vehicle charging station usage and profitability in Germany based on empirical data. iScience, 25(12), 105634. https://doi.org/10.1016/j.isci.2022.105634

[3] Propfe, B., Kreyenberg, D., Wind, J., & Schmid, S. (2013b). Market penetration analysis of electric vehicles in the German passenger car market towards 2030. International Journal of Hydrogen Energy, 38(13), 5201–5208. https://doi.org/10.1016/j.ijhydene.2013.02.049

[4] Energy, E. C. (2023, June 13). The impact of seasonal and geographical factors on the performance of electric car batteries. Energy5. https://energy5.com/the-impact-of-seasonal-and-geographical-factors-on-the-performance-of-electric-car-batteries