Electric Vehicle Adoption Analysis for Business Decision-Making

Table of Content

Introduction

Electric vehicles (EVs) have gained significant popularity as a clean and sustainable means of transportation. To harness the

potential of this growing market, businesses need to understand EV adoption patterns, consumer preferences, and market trends.

This case study aims to analyze an electric vehicle population dataset to provide valauable insights for businesses in the EV

industry.

About the dataset

This dataset shows the Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) that are currently

registered through Washington State Department of Licensing (DOL). It features 124716 rows and 17 columns of EV related data

across the United States of America.

Objectives

The primary objectives of this data analysis project are:

a) Identify the geographic distribution of EVs: Determine which counties, cities and states have the highest concentration of electric vehicles.

b) Analyze EV model and manufacturers: Examine the market share and trends of different EV manufacturers and their respective models.

c) Investigate EV range to understand consumer preferences.

d) Analyze utility provider influence: Investigate the role of electric utility providers in promoting EV adoption.

e) Investigate CAFV eligibilty: Determine the proportion of EVs eligible for Clean Alternative Fuel Vehicle (CAFV) incentives.

Methodology

a) Data Cleaning: Remove any duplicates, missing values, or irrelevant columns from the dataset.

b) Descriptive Analysis: Generate descriptive statistics and visualizations and to gain an overall understanding of the dataset.

c) Geographic Analysis: Utilize geographical data to map the distribution of electric vehicles across counties, cities, and states.

d) Manufacturer and Model Analysis: Analyze the market share of different EV manufacturers and investigate the popularity of specific EV models.

e) Electric Range Analysis: Explore the electric range of EVs from popular manufacturers.

f) Electric Utility Analysis: Analyze the role of electric utility providers in promoting EV adoption by assessing their market penetration.

g) CAFV Eligibility Analysis: Calculate the percentage of EVs eligible for CAFV incentives to unerstand the potential impact on customer choices.

Expected Outcomes

a) Identification of hotspots for EV adoption: Pinpoint areas with high concentration of EVs, allowing businesses to prioritize marketing and infrastucture investments.

b) Insights into manufacturer and model preferences: Determine the most popular EV manufacturers and models, helping businesses understand consumer preferences and plan inventory accordingly.

c) Understanding the electric range and it's effect on the adoption of certain EVs: Determine the sensitivity of EV buyers based on electric range.

d) Evaluation of utility provider influence: Understand the role of electric utility providers in shaping EV adoption rates.

e) CAFV eligibility impact: Assess the influence of CAFV eligibilty on consumer choices, potentially indicating the effectiveness of incentives and suggesting improvements.

Exploratory Analysis

Importing the python packages to be used for analysis

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

Data Manipulation

Here we will take a first look at the dataset by loading it into our workspace and determining what parts of our dataset might

require a bit of cleaning or transformation.

In [2]:
```
# Load the dataframe using the pandas read_csv function
# Then assign the variable name 'evpd' to the dataframe
# Finally, we will print the dataframe

evpd = pd.read_csv('Downloads/Electric_Vehicle_Population_Data.csv')
evpd
```

Out[2]:

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Elect Vehi Ty
0	5YJ3E1EB4L	Yakima	Yakima	WA	98908.0	2020	TESLA	MODEL 3	Batt Elec Vehi (BE
1	5YJ3E1EA7K	San Diego	San Diego	CA	92101.0	2019	TESLA	MODEL 3	Batt Elec Vehi (BE
2	7JRBR0FL9M	Lane	Eugene	OR	97404.0	2021	VOLVO	S60	Plug Hyb Elec Vehi (PHE
3	5YJXCBE21K	Yakima	Yakima	WA	98908.0	2019	TESLA	MODEL X	Batt Elec Vehi (BE
4	5UXKT0C5XH	Snohomish	Bothell	WA	98021.0	2017	BMW	X5	Plug Hyb Elec Vehi (PHE
...	...	...	...	...	...	...	...	...	
124711	5YJ3E1EB6N	Snohomish	Monroe	WA	98272.0	2022	TESLA	MODEL 3	Batt Elec Vehi (BE
124712	KNDCM3LD2L	Pierce	Tacoma	WA	98406.0	2020	KIA	NIRO	Plug Hyb Elec Vehi (PHE
124713	7SAYGDEE0P	Whatcom	Bellingham	WA	98226.0	2023	TESLA	MODEL Y	Batt Elec Vehi (BE
124714	1G1FW6S03J	Pierce	Tacoma	WA	98444.0	2018	CHEVROLET	BOLT EV	Batt Elec Vehi (BE
124715	1G1RC6E47F	Benton	Benton City	WA	99320.0	2015	CHEVROLET	VOLT	Plug Hyb Elec Vehi (PHE

124716 rows × 17 columns

If you glance through, you will realize our dataframe contains 17 columns and 124,716 rows, Immediately we can establish that

certain columns will be requiring cleaning and transformation processes. The following columns will prove crucial to our

analysis.

- The counties, cities and state columns will prove essential to our geographical distribution of EVs.
- The Make and model columns, we will use for manufacturer and model Analysis.
- We will also be analyzing the electric vehicle type, CAFV eligibility, electric range and electric utility columns.

In [3]:
```
# Check for basic information regarding our dataset

evpd.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 124716 entries, 0 to 124715
Data columns (total 17 columns):
 #   Column                                             Non-Null Count   Dtyp
e
---  ------                                             --------------   ----
-
 0   VIN (1-10)                                         124716 non-null  obje
ct
 1   County                                             124714 non-null  obje
ct
 2   City                                               124714 non-null  obje
ct
 3   State                                              124716 non-null  obje
ct
 4   Postal Code                                        124714 non-null  floa
t64
 5   Model Year                                         124716 non-null  int6
4
 6   Make                                               124716 non-null  obje
ct
 7   Model                                              124535 non-null  obje
ct
 8   Electric Vehicle Type                              124716 non-null  obje
ct
 9   Clean Alternative Fuel Vehicle (CAFV) Eligibility  124716 non-null  obje
ct
 10  Electric Range                                     124716 non-null  int6
4
 11  Base MSRP                                          124716 non-null  int6
4
 12  Legislative District                               124419 non-null  floa
t64
 13  DOL Vehicle ID                                     124716 non-null  int6
4
 14  Vehicle Location                                   124687 non-null  obje
ct
 15  Electric Utility                                   124243 non-null  obje
ct
 16  2020 Census Tract                                  124714 non-null  floa
t64
dtypes: float64(3), int64(4), object(10)
memory usage: 16.2+ MB
```

From the results, our dataset does not include any null values. Also, all our columns have consistent data types.

In [4]:
```
# Check for unique values in each column of the dataset

evpd.nunique()
```

Out[4]:
```
VIN (1-10)                                             8340
County                                                  166
City                                                    651
State                                                    44
Postal Code                                             781
Model Year                                               21
Make                                                     35
Model                                                   120
Electric Vehicle Type                                     2
Clean Alternative Fuel Vehicle (CAFV) Eligibility         3
Electric Range                                          101
Base MSRP                                                31
Legislative District                                     49
DOL Vehicle ID                                       124716
Vehicle Location                                        768
Electric Utility                                         73
2020 Census Tract                                      2036
dtype: int64
```

The results above gives us a breakdown of the unique contents of each column.

In [5]:
```
# Check for the basic statistics and aggregations of the numerical columns in t

evpd.describe()
```

Out[5]:

	Postal Code	Model Year	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID
count	124714.000000	124716.000000	124716.000000	124716.000000	124419.000000	1.247160e+05
mean	98163.826740	2019.406339	79.471936	1556.068909	29.664481	2.040790e+08
std	2550.122515	2.976174	100.331969	10053.289929	14.749518	8.882569e+07
min	1730.000000	1997.000000	0.000000	0.000000	1.000000	4.385000e+03
25%	98052.000000	2018.000000	0.000000	0.000000	18.000000	1.541015e+08
50%	98121.000000	2020.000000	25.000000	0.000000	34.000000	1.995558e+08
75%	98370.000000	2022.000000	200.000000	0.000000	43.000000	2.275165e+08
max	99701.000000	2023.000000	337.000000	845000.000000	49.000000	4.792548e+08

The results above show the count, mean, standard deviation, minimum, percentile and maximum values of numeric columns in the

dataset. The results also show inconsistencies in the Base MSRP column where the minimum and percentile values up to the 75th

percentile is zero(0), as a result that column lacks the needed integrity for analysis.

Data Cleaning

For data cleaning we will carry out the following tasks:

- Drop the columns VIN (1-10), Postal Code, Base MSRP, Legislative District, DOL Vehicle ID, Vehicle Location and 2020 Census Tract, as they will not be needed for this analysis.
- Create a new column for CAFV eligibility that will show more consistent results than what currently is.
- Remove duplicate records from the dataset.
- Remove all rows with electric range as zero(0), as it is impossible to have electric cars with an electric range of zero.
- Rename all of the column headers, this time with consistent formatting.

Dropping columns not needed for analysis

In [6]:
```
# dropping unnecessary columns

evpd.drop(['VIN (1-10)', 'Postal Code', 'Base MSRP', 'Legislative District', '

# view columns to check for the effect of the changes made

evpd.head(0)
```

Out[6]:

County	City	State	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	Electric Utility

Creating a new column for CAFV eligibility

In [7]:
```
# creating a new column 'cafv_eligibility2' where cafv_eligibility options wil

conditions = [
    evpd['Clean Alternative Fuel Vehicle (CAFV) Eligibility'] == 'Clean Altern
    evpd['Clean Alternative Fuel Vehicle (CAFV) Eligibility'] == 'Not eligible
]

values = ['Eligible', 'Ineligible']
evpd.loc[:, 'cafv_eligibility2'] = np.select(conditions, values, default='Othe

# check for the effect of the changes made

evpd
```

Out[7]:

	County	City	State	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range
0	Yakima	Yakima	WA	2020	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	322
1	San Diego	San Diego	CA	2019	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	220
2	Lane	Eugene	OR	2021	VOLVO	S60	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	22
3	Yakima	Yakima	WA	2019	TESLA	MODEL X	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	289
4	Snohomish	Bothell	WA	2017	BMW	X5	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	14
...	...	...	...	...	...	...	...	...	...
124711	Snohomish	Monroe	WA	2022	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0
124712	Pierce	Tacoma	WA	2020	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	26
124713	Whatcom	Bellingham	WA	2023	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0
124714	Pierce	Tacoma	WA	2018	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	238
124715	Benton	Benton City	WA	2015	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38

124716 rows × 11 columns

Renaming column headers and introducing consistent naming formats

In [8]:
```
# making all columns lowercase

evpd.rename(columns = lambda x : x.lower(), inplace=True)
evpd.head(0)

# renaming mispelt columns

evpd.rename(columns = {'model year':'model_year', 'make':'manufacturer', 'elec
```

Removing duplicates from the dataset

In [9]:
```
# removing duplicate records from the dataset, we will assign 'df1' as the new

df = evpd
df1 = df.drop_duplicates()

# Checking the new copy of the dataset after removing duplicates

df1
```

Out[9]:

	county	city	state	model_year	manufacturer	model	electric_vehicle_typ
0	Yakima	Yakima	WA	2020	TESLA	MODEL 3	Battery Electri Vehicle (BEV
1	San Diego	San Diego	CA	2019	TESLA	MODEL 3	Battery Electri Vehicle (BEV
2	Lane	Eugene	OR	2021	VOLVO	S60	Plug-in Hybrid Electri Vehicle (PHEV
3	Yakima	Yakima	WA	2019	TESLA	MODEL X	Battery Electri Vehicle (BEV
4	Snohomish	Bothell	WA	2017	BMW	X5	Plug-in Hybrid Electri Vehicle (PHEV
...	...	...	...	...	...	...	.
124702	Clallam	Sequim	WA	2014	CHEVROLET	SPARK	Battery Electri Vehicle (BEV
124704	Chelan	Chelan	WA	2015	NISSAN	LEAF	Battery Electri Vehicle (BEV
124707	Grant	Ephrata	WA	2022	JEEP	GRAND CHEROKEE	Plug-in Hybrid Electri Vehicle (PHEV
124709	Whatcom	Bellingham	WA	2021	AUDI	E-TRON	Battery Electri Vehicle (BEV
124714	Pierce	Tacoma	WA	2018	CHEVROLET	BOLT EV	Battery Electri Vehicle (BEV

26317 rows × 11 columns

Removing all rows that have electric range as zero

In [10]:
```
# removing all rows that have electric range as zero and assign a new variable

df2 = df1[df1.electric_range !=0]
```

Descriptive Analysis

In [11]:
```
# Check the basic statistics and aggregations of the numerical columns in the 

df2.describe()
```

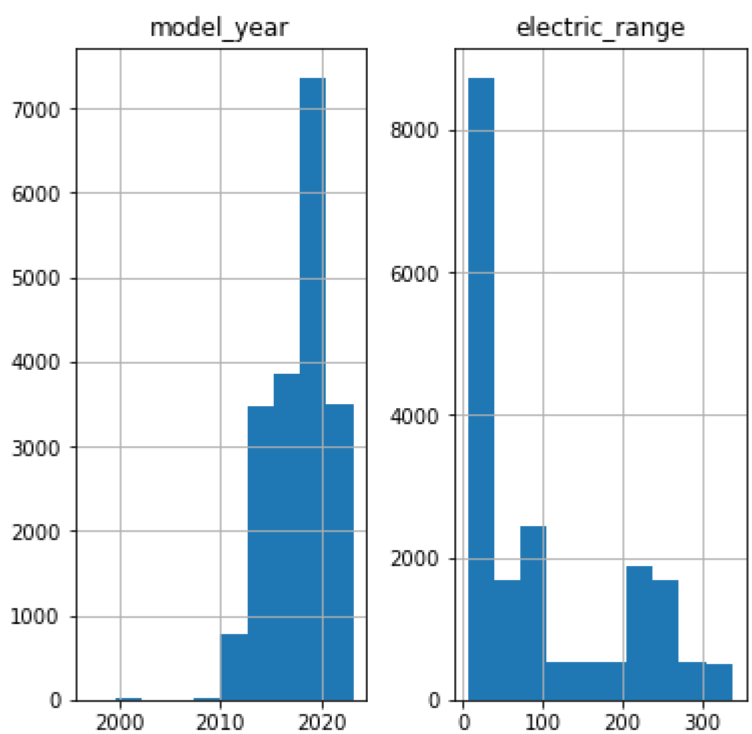
Out[11]:

	model_year	electric_range
count	18980.000000	18980.000000
mean	2017.784457	98.727134
std	2.933418	94.442398
min	1997.000000	6.000000
25%	2016.000000	22.000000
50%	2018.000000	53.000000
75%	2020.000000	203.000000
max	2023.000000	337.000000

After cleaning out all electric range values = 0, it's range of values now look okay to work with. Also the highest electric

range for an EV in this dataset is 337 miles.

In [12]:
```
# plotting a simple histogram of the numerical columns of the dataset using the

df2.hist(figsize=[6,6]);
```



The histogram shows that a greater proportion of the EV models in the dataset were manufactured between 2017 to 2020. Also a

large proportion of the vehicles in this dataset have an electric range below 100 miles.

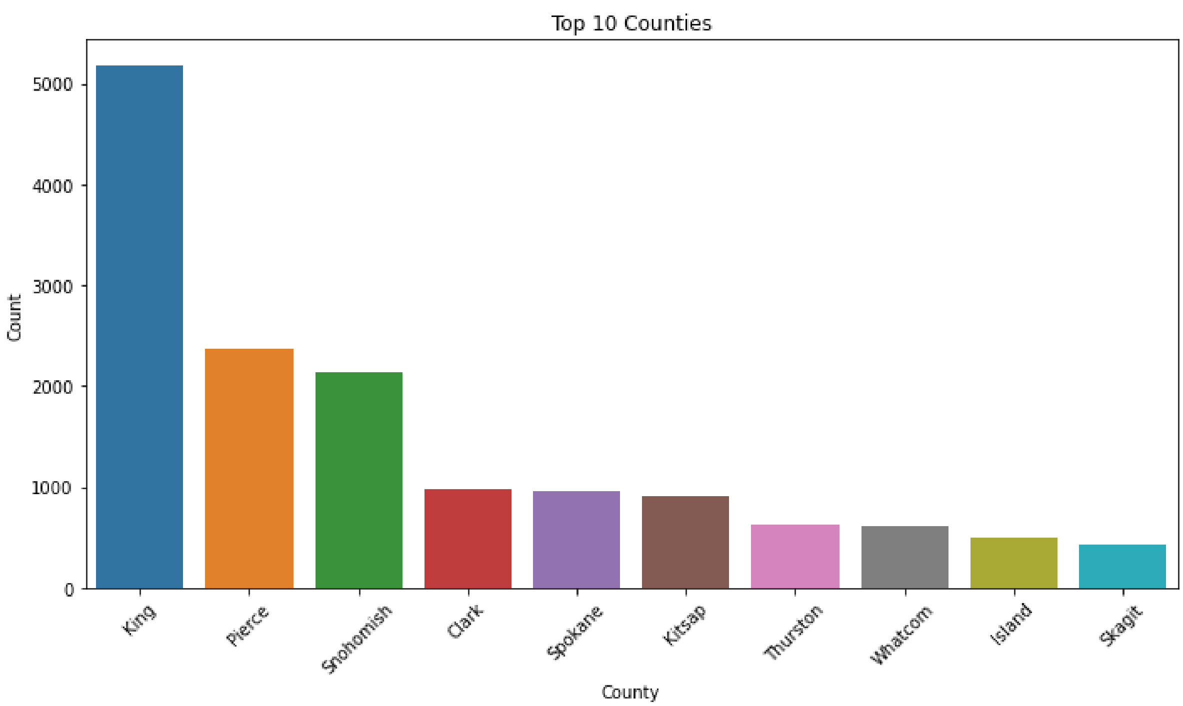
Geographical Analysis of the distribution of EVs

In this analysis we will, the counties, cities and states with the highest concentration of Evs in this dataset and represent

same using charts.

In [13]:
```
# Check to see the Top10 counties with the highest EV adoption

top_counties = df2['county'].value_counts().head(10)
```

In [14]:
```
# Top10 counties represented in a bar chart

plt.figure(figsize=(10, 6))
sns.barplot(x=top_counties.index, y=top_counties.values)
plt.xlabel('County')
plt.ylabel('Count')
plt.title('Top 10 Counties')
plt.xticks(rotation=45)
plt.tight_layout()

# Show the plot

plt.show()
```

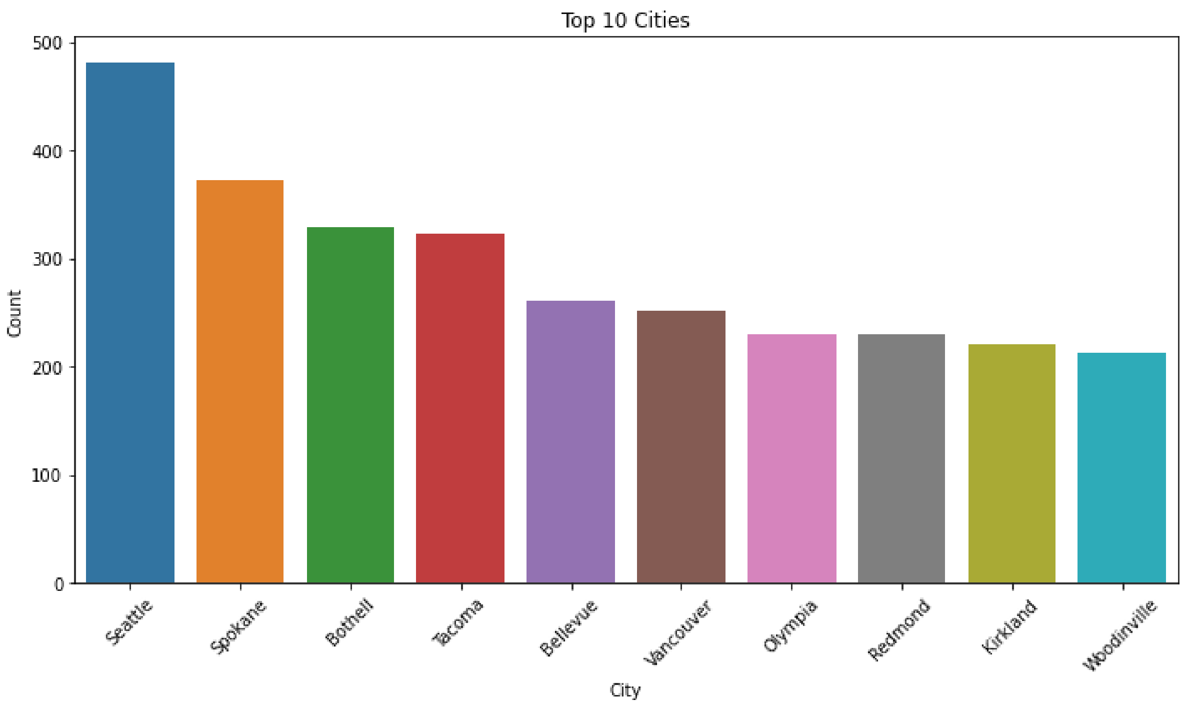


The results above shows that the King's, Pierce and Snohomish counties have the highest concentration of EVs in The United

States of America in this dataset.

In [15]:
```
# Check to see the Top10 cities with the highest EV adoption

top_cities = df2['city'].value_counts().head(10)
```

In [16]:
```
# Top10 cities represented in a bar chart

plt.figure(figsize=(10, 6))
sns.barplot(x=top_cities.index, y=top_cities.values)
plt.xlabel('City')
plt.ylabel('Count')
plt.title('Top 10 Cities')
plt.xticks(rotation=45)
plt.tight_layout()

# Show the plot
plt.show()
```

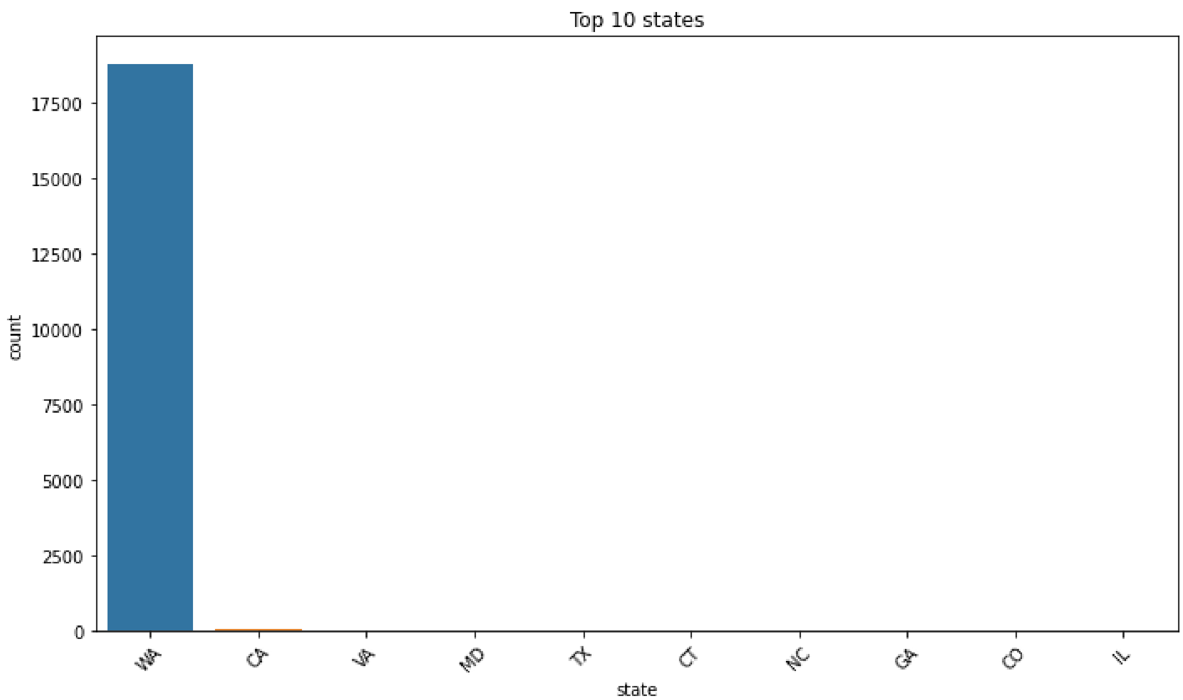


The results above shows that the cities of Seatle, Spokane and Bothell have the highest concentration of EVs in The United

States of America in this dataset.

In [17]:
```
# Check to see the Top10 states with the highest EV adoption

top_states = df2['state'].value_counts().head(10)
```

In [18]:
```
# Top10 states represented in a bar chart

plt.figure(figsize=(10, 6))
sns.barplot(x=top_states.index, y=top_states.values)
plt.xlabel('state')
plt.ylabel('count')
plt.title('Top 10 states')
plt.xticks(rotation=45)
plt.tight_layout()

# Show the plot
plt.show()
```



The results above shows that Washington DC has the highest concentration of EVs in The United

States of America in this dataset.

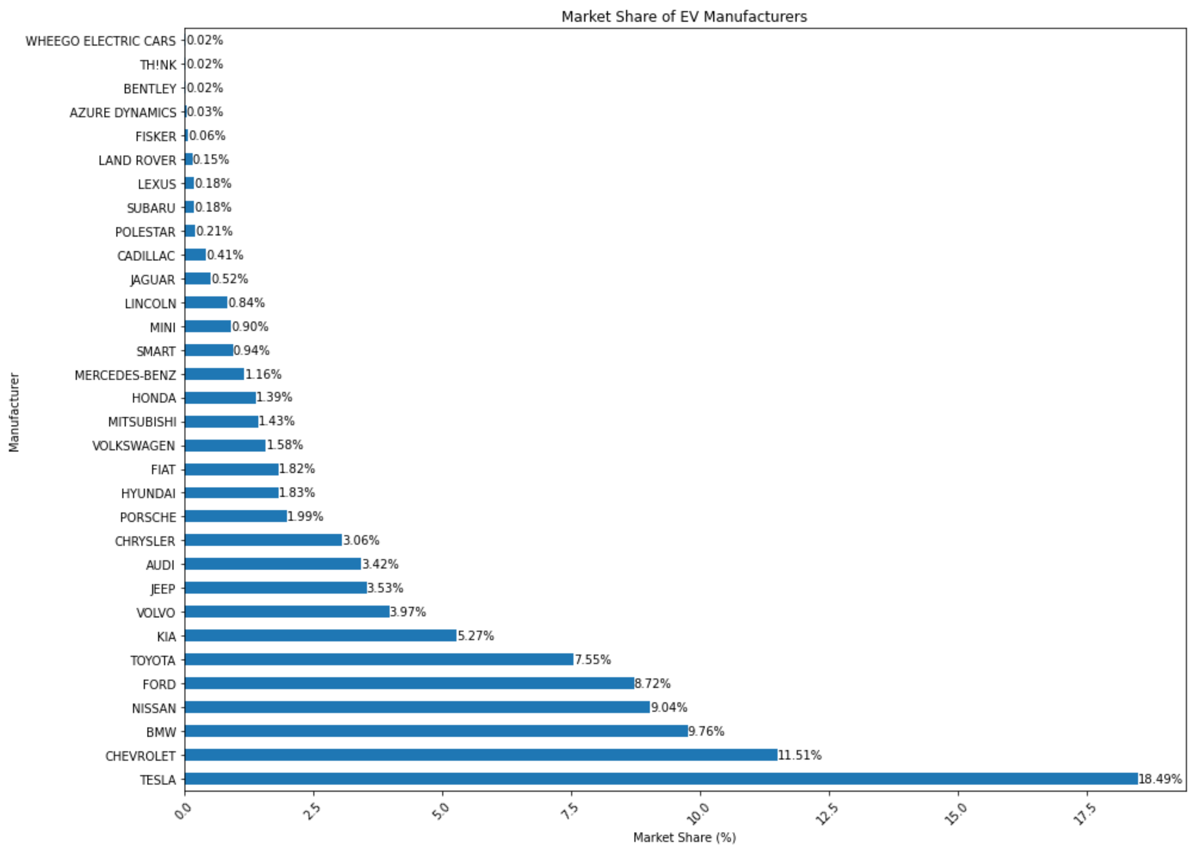
Manufacturer and model market share analysis

In this analysis we will explore the market share of the various manufacturers, we will also look at the models that are highly

sort after by consumers. These details will be represented using charts.

In [19]:
```
# manufacturer market share Analysis

manufacturer_counts = df2['manufacturer'].value_counts()
total_count = manufacturer_counts.sum()
manufacturer_market_share = (manufacturer_counts / total_count) * 100
```

In [20]:
```
# Sort manufacturers by market share in descending order
manufacturer_market_share_sorted = manufacturer_market_share.sort_values(ascen

# Plotting the market share of EV manufacturers
plt.figure(figsize=(14, 10))
manufacturer_market_share_sorted.plot(kind='barh')
plt.xlabel('Market Share (%)')
plt.ylabel('Manufacturer')
plt.title('Market Share of EV Manufacturers')
plt.xticks(rotation=45)

# Add market share percentage as text on each bar
for i, v in enumerate(manufacturer_market_share_sorted.values):
    plt.text(v, i, f'{v:.2f}%', va='center')

plt.tight_layout()
plt.show()
```

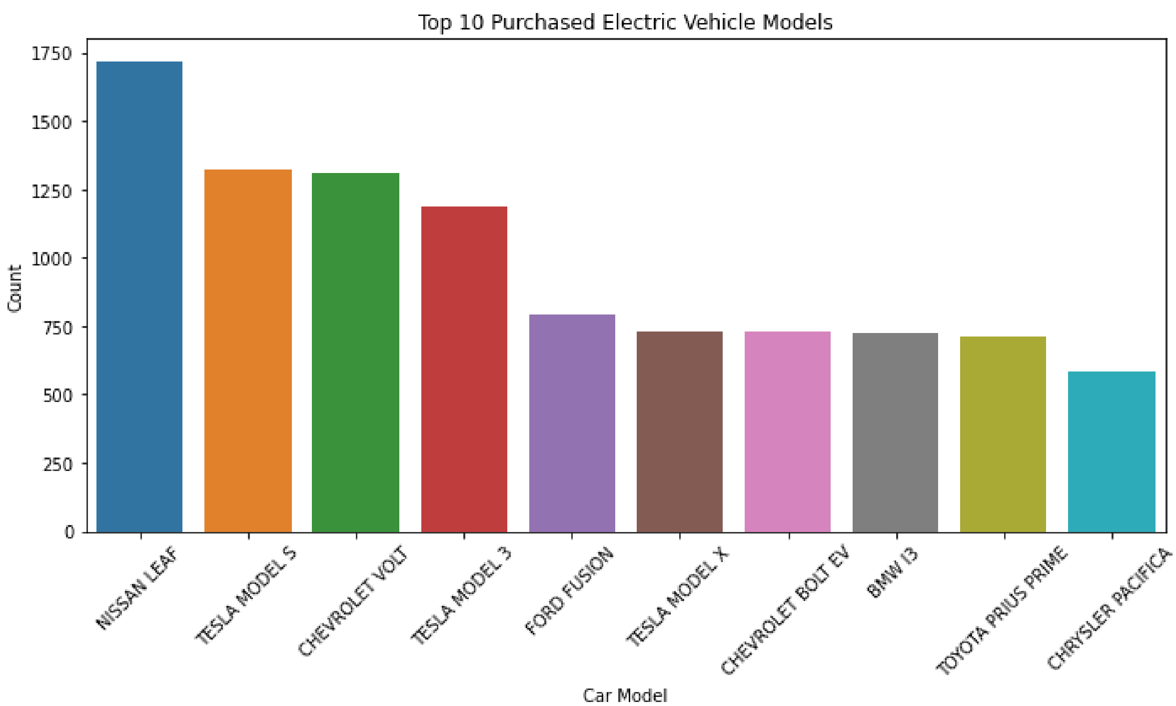


In this dataset, the results above shows that Tesla, Chervolet and BMW are the manufacturers with the highest market share of

EVs in The United States of America. Nissan, Ford, Toyota, Kia, Volvo, Jeep and Audi make up the top 10 manufacturers with the highest market shares.

In [21]:
```
# Top10 EV models in use

top_models = df2.groupby(['manufacturer', 'model']).size().reset_index(name='co
top_models = top_models.sort_values('count', ascending=False).head(10)
```

In [22]:
```
# Top10 EV models in use represented in a bar chart

# Set up the data for plotting
models = top_models['manufacturer'] + ' ' + top_models['model']
counts = top_models['count']

# Create the bar chart
plt.figure(figsize=(10, 6))
sns.barplot(x=models, y=counts)
plt.xticks(rotation=45)
plt.xlabel('Car Model')
plt.ylabel('Count')
plt.title('Top 10 Purchased Electric Vehicle Models')
plt.tight_layout()

# Show the chart
plt.show()
```



The chart shows the top10 EV models that are clearly sort after by consumers.

Manufacturer and Electric Range Analysis

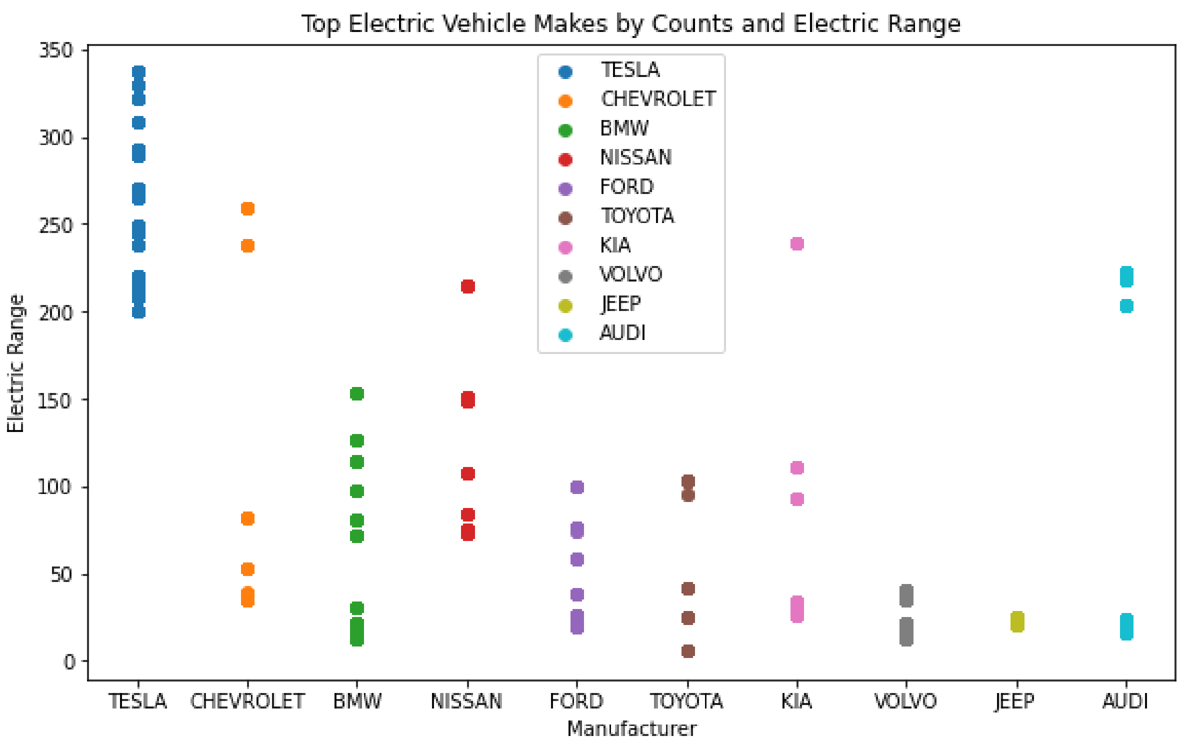
In this section, we will explore the various manufacturers and the electric range of their vehicles, we will check the minimum, maximum and mean values of their electric range and also show the distribution for the top10 manufacturers using scatterplots

In [23]:
```
#showing the maximum, minimum and mean values for the range of EVs

ev_range = df2.groupby(['manufacturer']).aggregate({'electric_range': ['mean',
ev_range
```

Out[23]:

	electric_range		
	mean	min	max
manufacturer			
AUDI	84.021538	16	222
AZURE DYNAMICS	56.000000	56	56
BENTLEY	18.666667	17	21
BMW	47.243389	13	153
CADILLAC	35.371795	31	40
CHEVROLET	113.086957	35	259
CHRYSLER	32.333907	32	33
FIAT	85.421965	84	87
FISKER	33.000000	33	33
FORD	27.418731	19	100
HONDA	45.984848	13	48
HYUNDAI	101.547550	27	258
JAGUAR	234.000000	234	234
JEEP	22.758209	21	25
KIA	88.961000	26	239
LAND ROVER	19.000000	19	19
LEXUS	37.000000	37	37
LINCOLN	23.157233	21	28
MERCEDES-BENZ	31.520362	8	87
MINI	45.087719	12	110
MITSUBISHI	29.944853	22	62
NISSAN	108.141691	73	215
POLESTAR	233.000000	233	233
PORSCHE	68.550265	12	203
SMART	62.882022	57	68
SUBARU	17.000000	17	17
TESLA	246.966667	200	337
TH!NK	100.000000	100	100
TOYOTA	26.728542	6	103
VOLKSWAGEN	104.140000	83	125
VOLVO	20.269588	13	41
WHEEGO ELECTRIC CARS	100.000000	100	100

In [24]:
```
# Get the top 10 manufacturers by count
top_10_manufacturers = df2['manufacturer'].value_counts().nlargest(10).index

# Filter the DataFrame to include only the rows with the top 10 manufacturers
filtered_df = df2[df2['manufacturer'].isin(top_10_manufacturers)]

# Increase the size of the scatter plot
plt.figure(figsize=(10, 6))

# Create a scatter plot for each manufacturer
for make in top_10_manufacturers:
    make_data = filtered_df[filtered_df['manufacturer'] == make]
    plt.scatter(make_data['manufacturer'], make_data['electric_range'], label=

# Set labels and title for the plot
plt.xlabel('Manufacturer')
plt.ylabel('Electric Range')
plt.title('Top Electric Vehicle Makes by Counts and Electric Range')
plt.legend()

# Show the scatter plot
plt.show()
```



The results show the amongst the ten most purchased brands, Tesla offers the best electric range for a greater proportion of its EVs

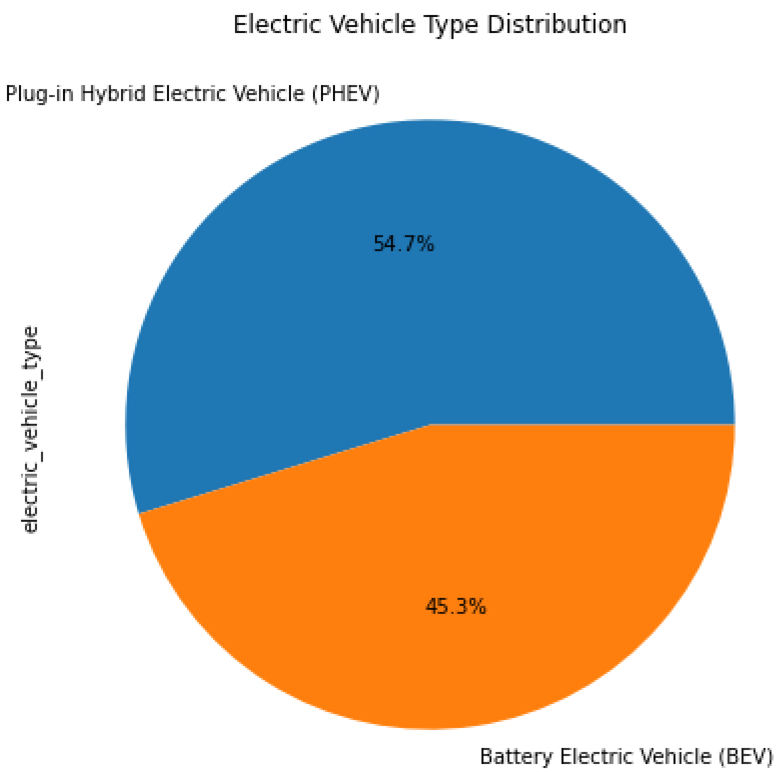
Battery Vehicle Type Analysis

In this analysis we will explore the battery vehicle type preference of the consumers.

In [25]:
```
# checking the count of the different types of electric vehicle type

df2['electric_vehicle_type'].value_counts()
```

Out[25]:
```
Plug-in Hybrid Electric Vehicle (PHEV)    10386
Battery Electric Vehicle (BEV)             8594
Name: electric_vehicle_type, dtype: int64
```

In [26]:
```
# Representing this in a pie chart

# Generate the value counts for 'caf_eligibility2'
value_counts = df2.electric_vehicle_type.value_counts()

# Set the size of the figure
fig, ax = plt.subplots(figsize=(7, 7))

# Create the pie chart with percentage values
value_counts.plot(kind='pie', autopct='%1.1f%%', ax=ax)

# Add a title to the chart
plt.title('Electric Vehicle Type Distribution')

# Display the chart
plt.show()
```



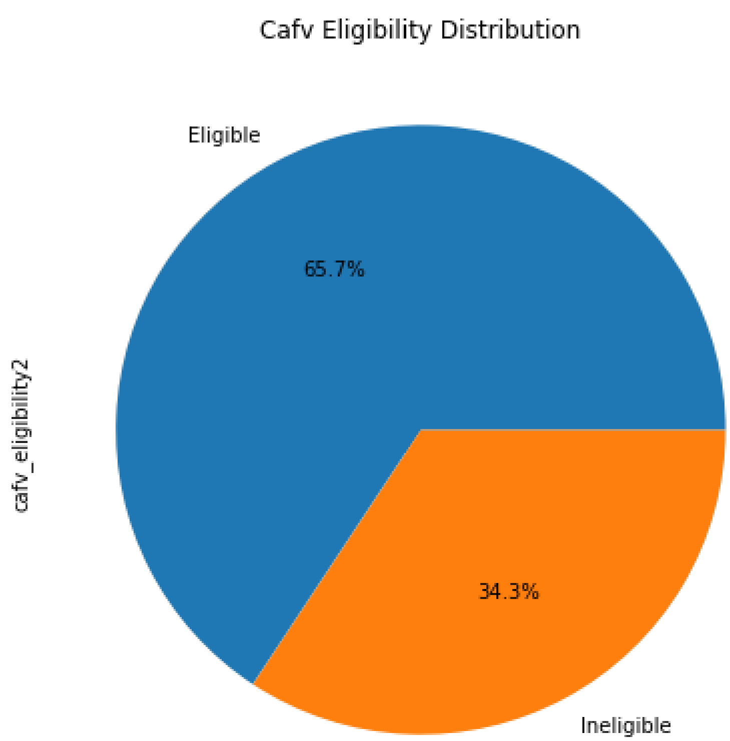
The results show 54.7% of the EV population use plug-in Hybrid Electric Vehicle and 45.3% use battery electric vehicles.

CAFV Eligibility Analysis

In [27]:
```
# checking the count of eligible EVs compared with those that are ineligible

df2['cafv_eligibility2'].value_counts()
```

Out[27]:
```
Eligible      12474
Ineligible     6506
Name: cafv_eligibility2, dtype: int64
```

In [28]:
```
# Representing this in a pie chart

# Generate the value counts for 'caf_eligibility2'
value_counts = df2.cafv_eligibility2.value_counts()

# Set the size of the figure
fig, ax = plt.subplots(figsize=(7, 7))

# Create the pie chart with percentage values
value_counts.plot(kind='pie', autopct='%1.1f%%', ax=ax)

# Add a title to the chart
plt.title('Cafv Eligibility Distribution')

# Display the chart
plt.show()
```



The results show 65.7% of the EV population are eligible and 34.3% are ineligible clean alternative fuel vehicle.

Electric Utility Analysis

In this analysis we will explore the market share of electric utility providers, to know which ones are most active in the EV business.

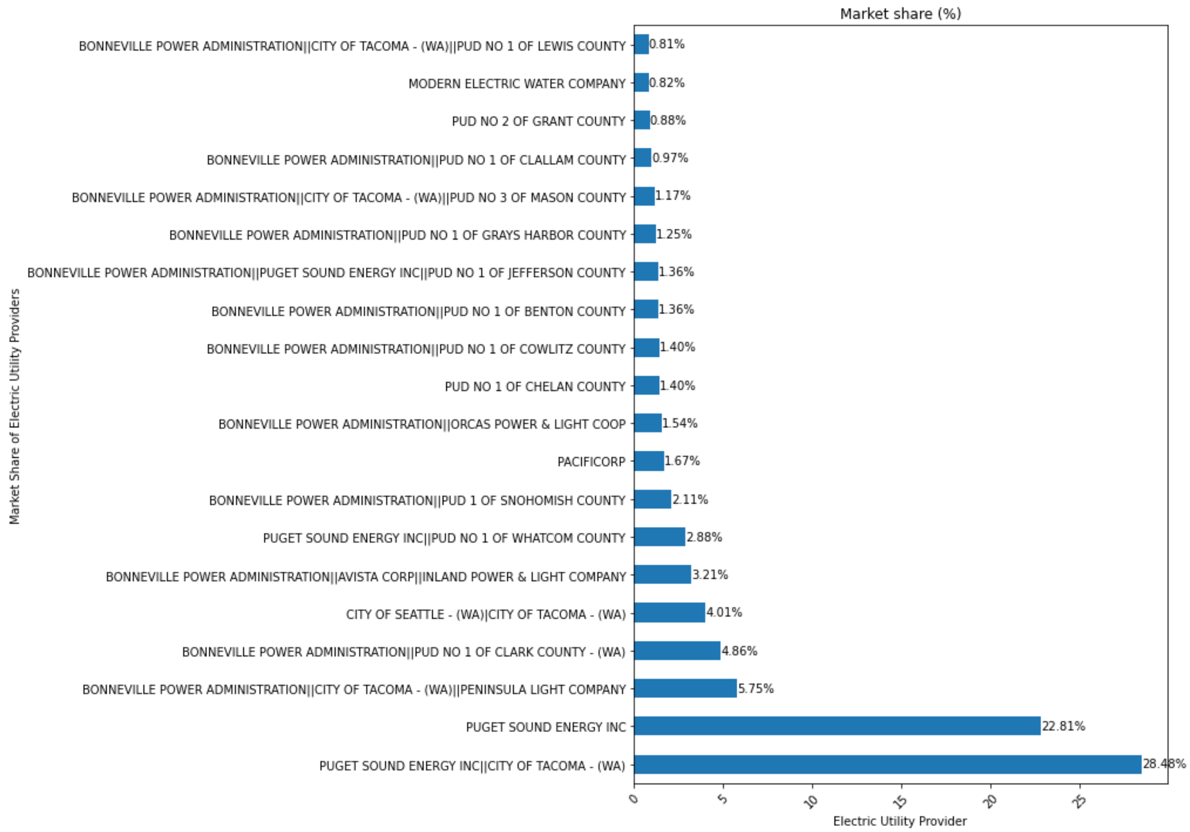
In [29]:
```
# Analysis to ascertain the market share of each EV utility provider
utility_counts = df2['electric utility'].value_counts()
total_count = utility_counts.sum()
utility_market_share = (utility_counts / total_count) * 100
utility_market_share
```

Out[29]:
```
PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)
28.481114
PUGET SOUND ENERGY INC
22.812751
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PENINSULA LIGHT COMPA
NY                5.754085
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLARK COUNTY - (WA)
4.859362
CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)
4.012858

...
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF ASOTIN COUNTY||INLAND POWER & LI
GHT COMPANY     0.010715
BONNEVILLE POWER ADMINISTRATION||NESPELEM VALLEY ELEC COOP, INC
0.010715
BONNEVILLE POWER ADMINISTRATION||PENINSULA LIGHT COMPANY
0.005358
PORTLAND GENERAL ELECTRIC CO
0.005358
CITY OF SEATTLE - (WA)
0.005358
Name: electric utility, Length: 71, dtype: float64
```

This result shows that there are 71 different utility providers across the United States. We will go on to

In [30]:
```
top_ums = utility_market_share.head(20)
```

In [31]:
```
# Plotting the market share of the Top20 electric utility providers
plt.figure(figsize=(14, 10))
top_ums.plot(kind='barh')
plt.xlabel('Electric Utility Provider')
plt.ylabel('Market Share of Electric Utility Providers')
plt.title('Market share (%)')
plt.xticks(rotation=45)
# Add market share percentage as text on each bar
for i, v in enumerate(top_ums.values):
    plt.text(v, i, f'{v:.2f}%', va='center')
plt.tight_layout()
plt.show()
```



From the results, Puget sound energy INC has the highest market share for a single outlet, nevertheless, Bonneville power

administration has a greater spread of outlets across counties and cities.

Conclusion

By conducting a comprehensive analysis of the electric vehicle population dataset, this case study was aimed at providing

valuable insights for businesses operating in the electric vehicle industry. The outcomes will support decision-making

processes, allowing businesses to tailor their strategies to target specific geographic areas, understand consumer preferences, optimize electric range and pricing, and collaborate with utility providers. These insights should contribute to the sustainable

growth and development of the electric vehicle market.

However, a few issues were encountered that may likely skew our results in a way, 98.96% of the EVs used in this analysis are in

the US state of Washington DC, as a result our analysis will mostly support business decision-making in that state and not

necessarily the rest of the United States.

Also the base MSRP column had a majority of its values as zero, hence it could not be used for analysis of product pricing and

its effect in the EV market.