

ANALYSIS OF **EV** STATIONS ACROSS THE U.S.





Business Problem

- Drive growth by targeting underserved regions based on data- driven trends.
- Identify gaps in sustainability efforts and promote the adoption of eco-friendly energy solutions in the EV charging network.
- Are there enough fast charging stations to meet the growing demand?
- Are charging stations dominated by a particular category of owners?



Dataset Used

- Source: U.S. Department of Energy – Alternative Fuels Data Center.
- URL: <https://afdc.energy.gov>
- Records Count: 78,837

The screenshot displays the homepage of the Alternative Fuels Data Center (AFDC) website. The header features the U.S. Department of Energy logo and the text "Energy Efficiency & Renewable Energy". Below this is a green navigation bar with the title "Alternative Fuels Data Center" and several menu items: "FUELS & VEHICLES", "CONSERVE FUEL", "LOCATE STATIONS", "LAWS & INCENTIVES", "Maps & Data", and "Case Studies". The main content area shows a breadcrumb trail "EERE » AFDC » Tools" and a section titled "Data Downloads". A sub-header reads "To download data related to alternative fuels and advanced vehicles, follow the steps below." followed by "Step 1. Choose data to download". Below this, it says "Choose the dataset you want to download." and there is a dropdown menu labeled "Dataset" with "Alternative fuel stations" selected.

U.S. DEPARTMENT OF **ENERGY** | Energy Efficiency & Renewable Energy

Alternative Fuels Data Center

FUELS & VEHICLES CONSERVE FUEL LOCATE STATIONS LAWS & INCENTIVES Maps & Data Case Studies

[EERE](#) » [AFDC](#) » [Tools](#)

Data Downloads

To download data related to alternative fuels and advanced vehicles, follow the steps below.

Step 1. Choose data to download

Choose the dataset you want to download.

Dataset Alternative fuel stations



Project Workflow

- Data collection and cleaning.
- EDA and feature engineering.
- Dashboard design and visualization.
- Advanced analytics.
- Aesthetic Refinements.



ED

A

- Null Value Check for Key Columns
- Understanding the data distribution
- Duplicate Check
- Colab file link :
https://colab.research.google.com/drive/1oG_bJiAL5pFXMYQX18IJ_Mk7y0tL6lrln?usp=sharing#scrollTo=faJbRYog4uvo



Null Value Check for Key Columns

Missing Data Overview:		
	Missing Count	Missing Percentage
EV On-Site Renewable Source	78907	99.485595
EV Level1 EVSE Num	78617	99.119965
Federal Agency Name	78290	98.707685
Federal Agency ID	78290	98.707685
Federal Agency Code	78290	98.707685
Access Detail Code	74822	94.335246
Expected Date	73700	92.920633
Cards Accepted	71518	90.169577
Restricted Access	66811	84.235012
EV DC Fast Count	66799	84.219883
EV Pricing	63782	80.416063
Facility Type	59921	75.548131
Owner Type Code	58318	73.527076
EV Level2 EVSE Num	11389	14.359201
EV Network Web	9639	12.152808
Access Days Time	7986	10.068713
Open Date	441	0.556011
Date Last Confirmed	156	0.196684
Street Address	34	0.042867
EV Workplace Charging	11	0.013869
City	4	0.005043
Station Name	3	0.003782
Status Code	0	0.000000
Updated At	0	0.000000
State	0	0.000000
Access Code	0	0.000000
Longitude	0	0.000000
Latitude	0	0.000000
ID	0	0.000000
Geocode Status	0	0.000000
Fuel Type Code	0	0.000000
EV Network	0	0.000000
EV Connector Types	0	0.000000
Country	0	0.000000
ZIP	0	0.000000



Data Distribution (1)

Descriptive Statistics for Numerical Columns:

	EV Level2	EVSE Num	Latitude	Longitude	\
count	79315.000000	79315.000000	79315.000000		
mean	2.551291		37.870088	-96.199288	
min	1.000000		18.009854	-162.286348	
25%	2.000000		34.043346	-117.881102	
50%	2.000000		38.596430	-91.068251	
75%	2.551291		41.591639	-78.656218	
max	338.000000		64.852466	-65.756678	
std	3.357435		5.041741	19.380892	

	Date Last Confirmed	ID	\
count	79159	79315.000000	
mean	2024-09-07 19:18:28.346492416	220069.654038	
min	2019-12-12 00:00:00	1517.000000	
25%	2024-10-11 00:00:00	165452.500000	
50%	2024-12-02 00:00:00	205580.000000	
75%	2024-12-02 00:00:00	308834.500000	
max	2024-12-02 00:00:00	372237.000000	
std	NaN	85900.785636	

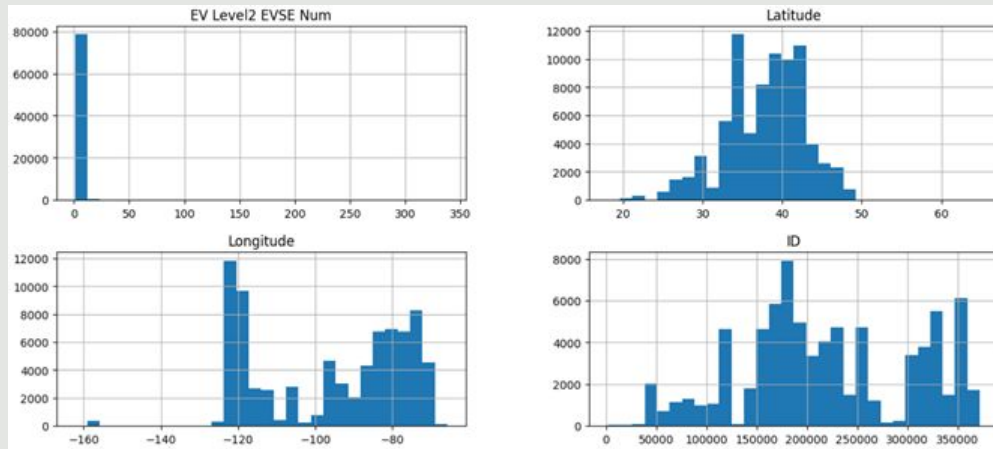
	Open Date	EV Workplace Charging
count	78874	79315.000000
mean	2021-03-28 11:03:17.631665408	0.019268
min	1995-08-30 00:00:00	0.000000
25%	2020-06-12 00:00:00	0.000000
50%	2021-10-23 00:00:00	0.000000
75%	2023-06-16 00:00:00	0.000000
max	2025-01-15 00:00:00	1.000000
std	NaN	0.137455

Data Distribution (2)



```
import matplotlib.pyplot as plt

# Plot histograms for all numerical columns
df_cleaned.select_dtypes(include=['float64', 'int64']).hist(figsize=(15, 10), bins=30)
plt.suptitle("Histograms for Numerical Columns", fontsize=16)
plt.show()
```





Duplicates Check

```
# Check for duplicate rows
print(f"Number of duplicate rows before removal: {df_cleaned.duplicated().sum()}")

# Remove duplicate rows
df_no_duplicates = df_cleaned.drop_duplicates()

# Display results
print(f"Shape before duplicate removal: {df_cleaned.shape}")
print(f"Shape after duplicate removal: {df_no_duplicates.shape}")
```

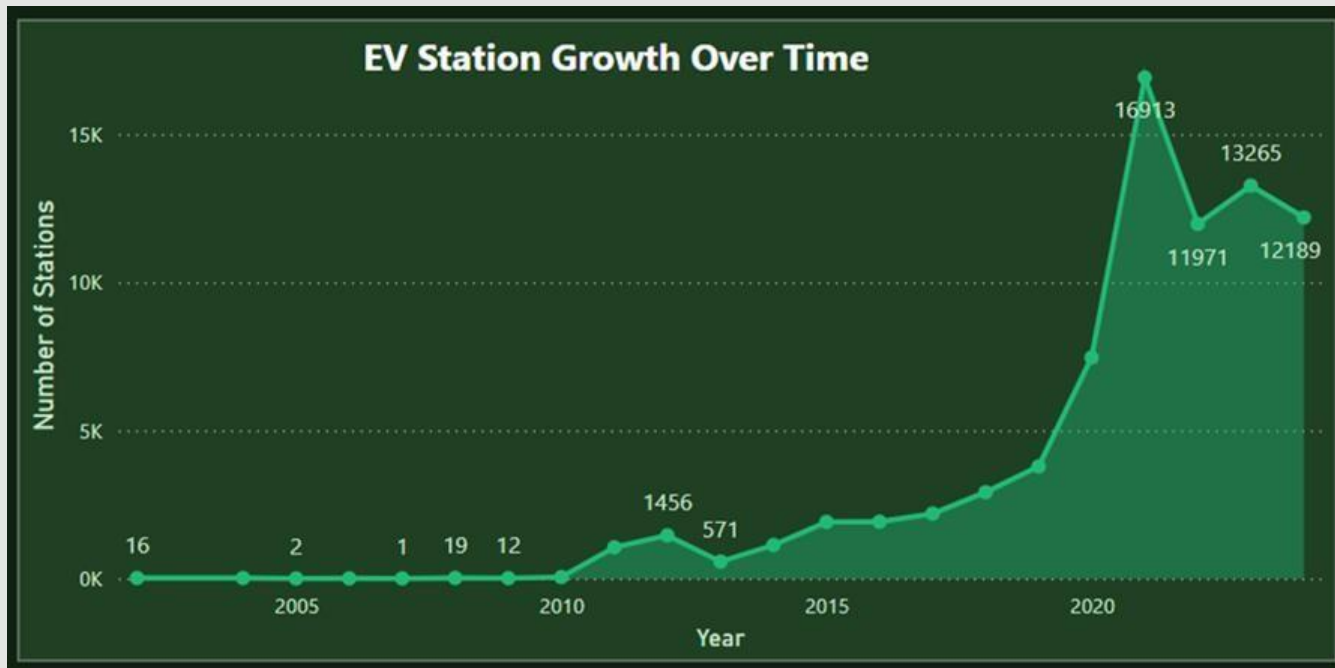
```
Number of duplicate rows before removal: 0
Shape before duplicate removal: (79315, 22)
Shape after duplicate removal: (79315, 22)
```



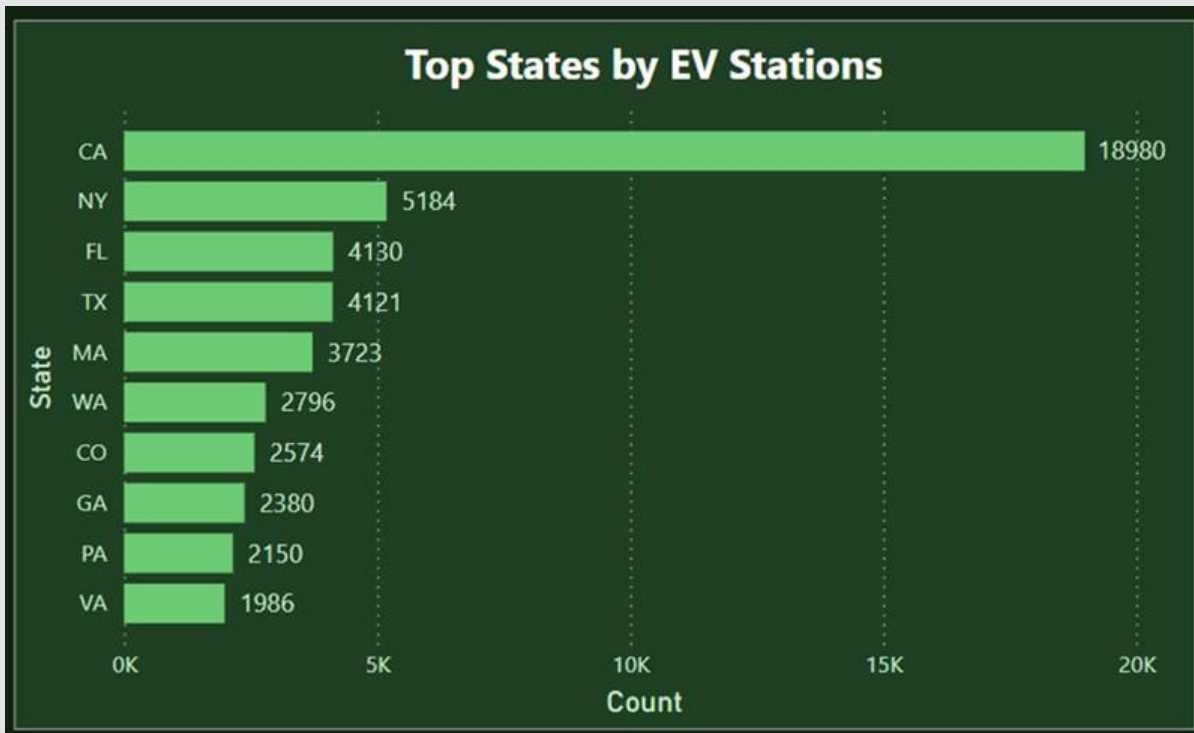
Analysis Performed

- Trend Analysis over Time.
- Access Type Distribution.
- Top States by EV Station Utilization
- Availability of Fast Chargers.
- Classification of Key Stakeholders.
- ML Forecasting of Future Expansion.

Area Chart



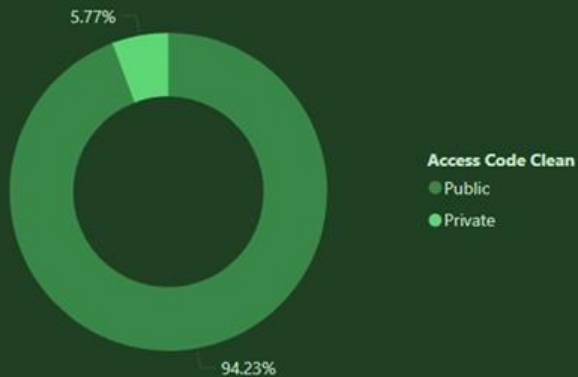
Bar Chart



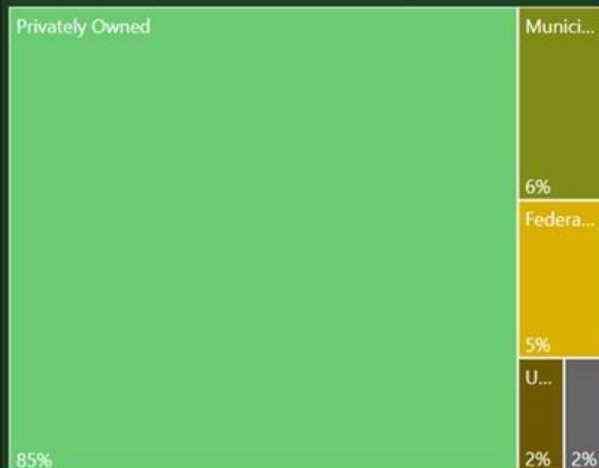
Donut Chart & Tree Map



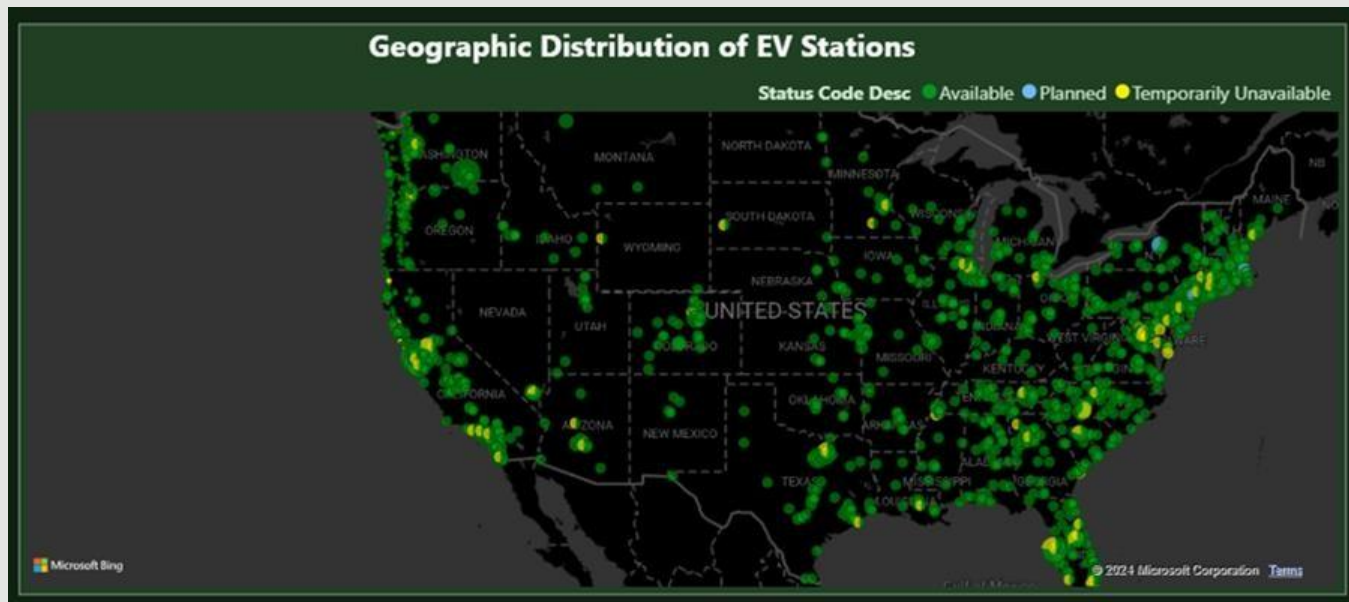
Access Type Distribution: Public vs Private



Distribution of Owner by Station



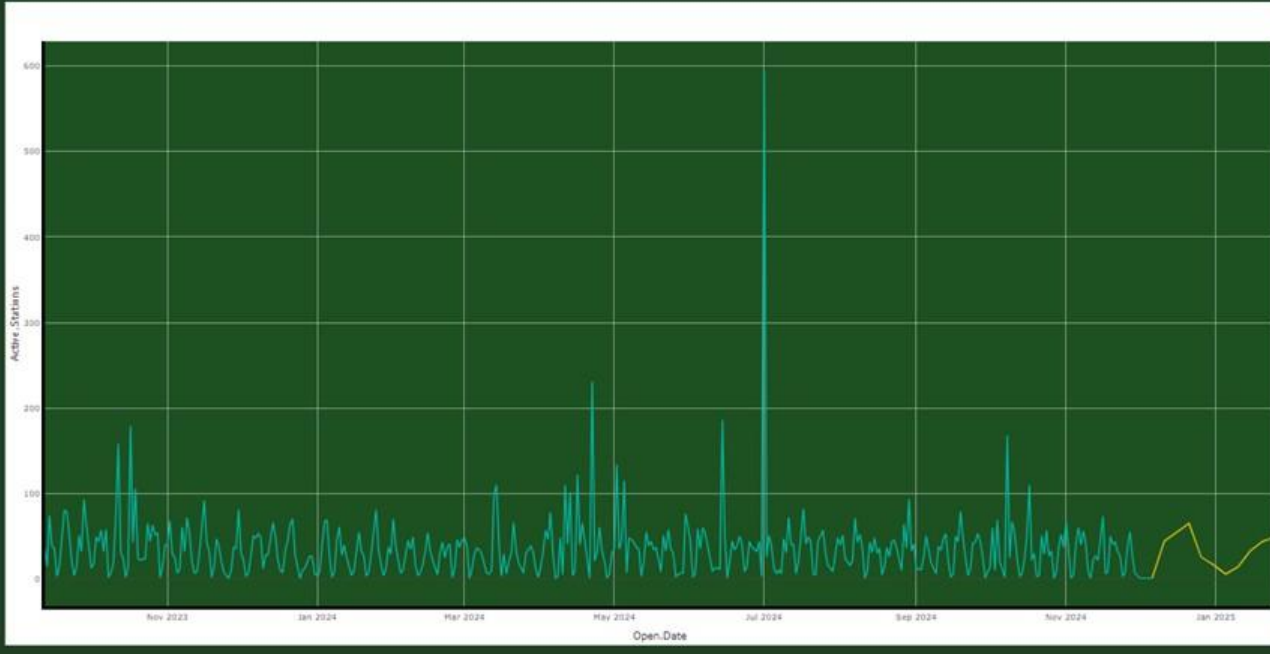
Geospatial Map



ML Prediction using MAQ Chart



▪ Number of Station Prediction using R





Dashboard View (Page 1)



Additional Features - Q & A



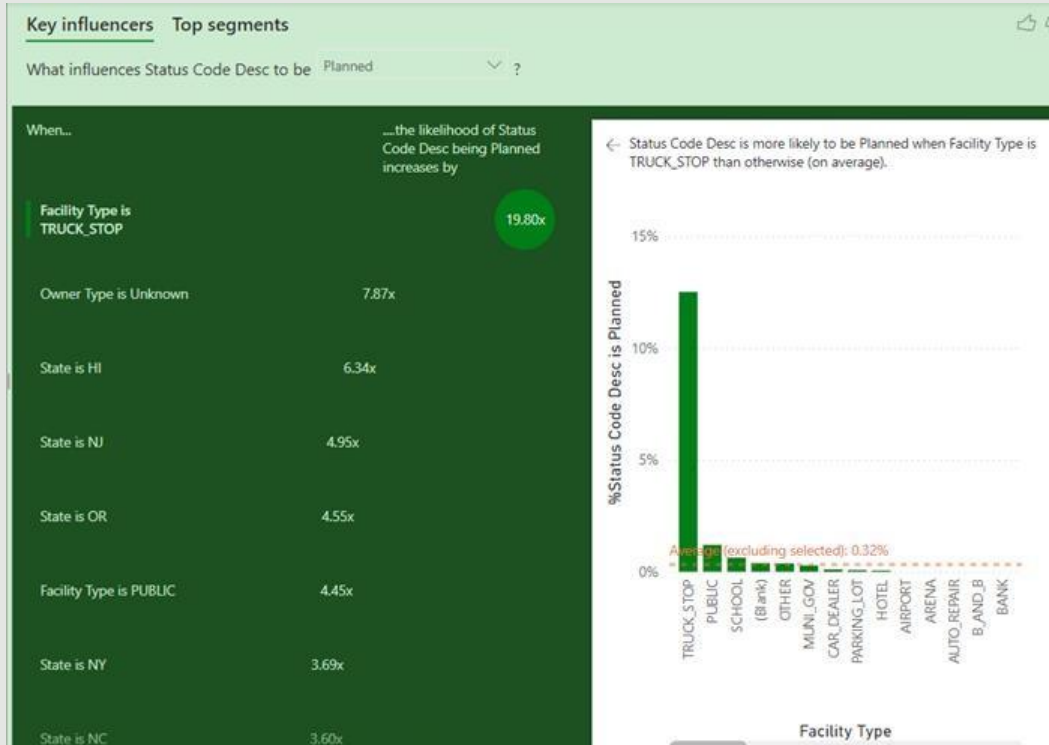
Which state has maximum EV network stations?

Showing results for Top state of EV stations data by active stations of those EV stations data

State

CA

Additional Features - Key Influencers





Additional Features - Decomposition Tree



Dashboard View (Page 2)



ADDITIONAL FEATURES

Which state has maximum EV network stations?

Showing results for Top states of EV stations data by active stations of those EV stations data

State
CA

Key Influencer

Key influencers Top segments

What influences Status Code Desc to be Planned

When...
...the likelihood of Status Code Desc being Planned increases by

Facility Type is TRUCK_STOP 19.80x

Owner Type is Unknown 7.87x

State is HI 6.34x

State is NJ 4.95x

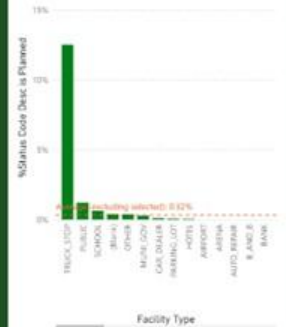
State is OR 4.55x

Facility Type is PUBLIC 4.45x

State is NY 3.88x

State is NC 3.68x

← Status Code Desc is more likely to be Planned when Facility Type is TRUCK_STOP than otherwise (on average)



☐ Only show values that are influencers

Decomposition Tree

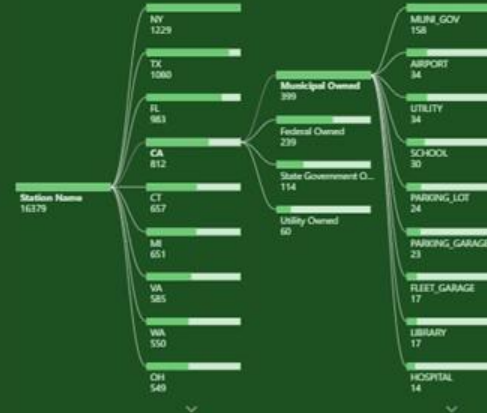
State

CA

Owner Type

Municipal Owned

Facility Type



Dashboard Link



<https://app.powerbi.com/groups/me/reports/2df8ca23-ccda-44ce-bf88-e7c20369b9b2/56dcbb3684d7de188dac?experience=power-bi>



Advanced Features

- DAX Metrics
- Q&A Feature
- Key Influencers
- Decomposition Tree



Conclusion

- California leads the U.S. in EV charging stations, showcasing its commitment to EV infrastructure.
- 95% of stations are publicly accessible, ensuring wide user access.
- EV charging stations have grown steadily since 2010, with a significant surge after 2019.
- Over 85% are privately owned, emphasizing the role of private entities in expansion.
- Only 18% offer DC Fast Charging, and less than 1% use renewable energy, highlighting opportunities for improvement in speed and sustainability.



Future Scope

- Integration with Real-Time Data to monitor station utilization and downtime.
- Use geo-spatial analysis and clustering techniques for recommending optimal locations for new stations.
- Incorporate user reviews for service improvements.