

The US Government's Alternative Fuels Data Center collects records of electric vehicle (EV) charging infrastructure, including charging ports and station locations, as well as sales of electric vehicles. With the EV market rapidly evolving, understanding trends in charging facilities and sales is essential to inform strategic planning.

As a data scientist working for a leading EV charging network operator, you recognize the potential in this data and start wrangling and visualizing the aggregated yearly data.

This yearly data captured in December of each year encompasses a record of EV charging port installations and station localities spanning roughly ten years, capturing both public and private charging environments.

The Data

private_ev_charging.csv

Variable	Description
year	Year of data collection
private_ports	The number of available charging ports owned by private companies in a given year
private_station_locations	The number of privately owned station locations for EV charging

public_ev_charging.csv

Variable	Description
year	Year of data collection
public_ports	The number of available charging ports under public ownership in a given year
public_station_locations	The number of publicly owned station locations for EV charging

The sales information is available for each model and year in the ev_sales.csv file:

Variable	Description
Vehicle	Electric vehicle model
year	Year of data collection

Variable	Description
<code>sales</code>	The number of vehicles sold in the US

```
# Import required libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Start coding here
private_ev_charging = pd.read_csv("private_ev_charging.csv")
public_ev_charging = pd.read_csv("public_ev_charging.csv")
ev_sales = pd.read_csv("ev_sales.csv")

df_combined = private_ev_charging.merge(public_ev_charging, on="year", how="outer",
indicator=True)
df_temp = df_combined[df_combined['_merge'] == 'both']

df_temp = df_temp.drop(columns=['_merge'])

ev_total_sales = ev_sales.groupby('year')['sales'].sum().reset_index()

print(ev_total_sales)
ev_sales_2018 = 361315

df_complete = df_temp.merge(ev_total_sales, how='left', on='year')

df_complete = df_complete.dropna(subset="sales")

fig, ax = plt.subplots()

sns.lineplot(data=df_complete, x='year', y='private_ports', label='Private Ports')
sns.lineplot(data=df_complete, x='year', y='public_ports', label='Public Ports')
sns.lineplot(data=df_complete, x='year', y='sales', label='Total Sales', linestyle=':')

ax.set_title('EV Ports and Sales Over Time')
ax.set_xlabel='Year', ylabel='Count')

ax.legend(loc='upper left')

plt.show()

trend = "same"
```

	year	sales
0	2011	17763.0
1	2012	53171.0
2	2013	97102.0
3	2014	118882.0
4	2015	114023.0
5	2016	159616.0
6	2017	195581.0
7	2018	361315.0
8	2019	326644.0

