Case Study On Electric Vehicle Population

March 9, 2025

1 Case Study On Electric Vehicle Population By Monika Mahala

Connecting SQL Workbench to Jupyter Notebook:

[ ]:
```
! pip install sqlalchemy pymysql
! pip install cryptography
```

[2]:
```
from sqlalchemy import create_engine
import pandas as pd
```

[8]:
```
# Create database connection using SQLAlchemy
# engine = create_engine('mysql+pymysql://your_username:your_password@localhost/
 ↪your_database')

engine = create_engine('mysql+pymysql://root:####@localhost/ev')
```

Questions:

[9]:
```
# 1. Retrieve all records from the dataset where the State is 'Washington'.
query = """
SELECT *
FROM electric_vehicles
WHERE State = 'WA';
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[9]:
```
          VIN    County     City State Postal_Code  Model_Year   Make  \
0  3C3CFFGE4E    Yakima   Yakima    WA     98902.0        2014   FIAT
1  5YJXCBE40H  Thurston  Olympia    WA     98513.0        2017  TESLA
2  3MW39FS03P      King   Renton    WA     98058.0        2023    BMW

     Model                   Electric_Vehicle_Type  \
0      500          Battery Electric Vehicle (BEV)
1  MODEL X          Battery Electric Vehicle (BEV)
2     330E  Plug-in Hybrid Electric Vehicle (PHEV)

                        CAFV_Eligibility  Electric_Range  Base_MSRP  \
```

```
0  Clean Alternative Fuel Vehicle Eligible              87        0.0
1  Clean Alternative Fuel Vehicle Eligible             200        0.0
2    Not eligible due to low battery range              20        0.0

  Legislative_District DOL_Vehicle_ID                Vehicle_Location  \
0                 14.0        1593721   POINT (-120.524012 46.5973939)
1                  2.0      257167501     POINT (-122.817545 46.98876)
2                 11.0      224071816  POINT (-122.1298876 47.4451257)

                              Electric_Utility   Census_Tract Price_Category
0                                    PACIFICORP  53077000700.0            Low
1                         PUGET SOUND ENERGY INC  53067012331.0            Low
2  PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)  53033025803.0            Low
```

[10]:
```
#2. List distinct Electric Vehicle Types available in the dataset.
query = """
SELECT DISTINCT `Electric_Vehicle_Type`
FROM electric_vehicles;

"""

df = pd.read_sql(query, engine)
df.head(3)
```

[10]:
```
                    Electric_Vehicle_Type
0          Battery Electric Vehicle (BEV)
1  Plug-in Hybrid Electric Vehicle (PHEV)
```

[9]:
```
#3. Get all vehicles with an Electric Range greater than 200 miles, sorted in␣
 ↪descending order.
query = """
SELECT *
FROM electric_vehicles
WHERE `Electric_Range` > 200
ORDER BY `Electric_Range` DESC;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[9]:
```
           VIN     County           City State Postal_Code  Model_Year   Make  \
0  5YJSA1E49L  Snohomish        Bothell    WA     98021.0        2020  TESLA
1  5YJSA1E45L       King  Mercer Island    WA     98040.0        2020  TESLA
2  5YJSA1E44L      Mason        Belfair    WA     98528.0        2020  TESLA

     Model           Electric_Vehicle_Type  \
0  MODEL S  Battery Electric Vehicle (BEV)
```

```
1  MODEL S  Battery Electric Vehicle (BEV)
2  MODEL S  Battery Electric Vehicle (BEV)

                         CAFV_Eligibility  Electric_Range  Base_MSRP  \
0  Clean Alternative Fuel Vehicle Eligible             337        0.0
1  Clean Alternative Fuel Vehicle Eligible             337        0.0
2  Clean Alternative Fuel Vehicle Eligible             337        0.0

  Legislative_District DOL_Vehicle_ID                 Vehicle_Location  \
0                  1.0      241126777    POINT (-122.179458 47.802589)
1                 41.0        4987846   POINT (-122.2377542 47.582905)
2                 35.0      181430532  POINT (-122.8551647 47.4495785)

                                    Electric_Utility   Census_Tract  \
0                            PUGET SOUND ENERGY INC  53061051918.0
1      PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)  53033024500.0
2  BONNEVILLE POWER ADMINISTRATION||CITY OF TACOM…  53045960302.0

  Price_Category
0            Low
1            Low
2            Low
```

[11]:
```
#4. Find all vehicles with a Base MSRP between $30,000 and $60,000.

query = """
SELECT *
FROM electric_vehicles
WHERE `Base_MSRP`BETWEEN 30000 AND 60000;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[11]:
```
           VIN  County              City State Postal_Code  Model_Year   Make  \
0  WBA8E1C32H    King            Auburn    WA     98092.0        2017    BMW
1  LYVBR0DL2J  Kitsap          Bremerton    WA     98310.0        2018  VOLVO
2  KNDJX3AE6G  Kitsap  Bainbridge Island    WA     98110.0        2016    KIA

  Model                   Electric_Vehicle_Type  \
0  330E  Plug-in Hybrid Electric Vehicle (PHEV)
1  XC60  Plug-in Hybrid Electric Vehicle (PHEV)
2  SOUL           Battery Electric Vehicle (BEV)

                         CAFV_Eligibility  Electric_Range  Base_MSRP  \
0    Not eligible due to low battery range              14    44100.0
1    Not eligible due to low battery range              17    52900.0
```

```
2  Clean Alternative Fuel Vehicle Eligible              93    31950.0

  Legislative_District DOL_Vehicle_ID                Vehicle_Location  \
0                 47.0      178980651  POINT (-122.1820969 47.3198995)
1                 23.0      240710669    POINT (-122.611365 47.575195)
2                 23.0        8861044  POINT (-122.5235781 47.6293323)

                                Electric_Utility   Census_Tract Price_Category
0  PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)  53033031208.0            Mid
1                         PUGET SOUND ENERGY INC  53035080400.0            Mid
2                         PUGET SOUND ENERGY INC  53035090901.0            Mid
```

[12]:
```
#5. Count the number of electric vehicles for each Make.
query = """
SELECT Make, COUNT(*) AS vehicle_count
FROM electric_vehicles
GROUP BY Make
ORDER BY vehicle_count DESC;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[12]:
```
         Make  vehicle_count
0      TESLA          74834
1     NISSAN          13848
2  CHEVROLET          13072
```

[14]:
```
#6. Find the average Electric Range for each Model Year.
query = """
SELECT `Model_Year`, AVG(`Electric_Range`) AS avg_electric_range
from electric_vehicles
GROUP BY `Model_Year`
ORDER BY `Model_Year` asc;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[14]:
```
   Model_Year  avg_electric_range
0        1997                39.0
1        1998                58.0
2        1999                74.0
```

[15]:
```
#7. Get the total number of electric vehicles available in each City, showing␣
 ↪only cities with more than 100 vehicles.
query = """
```

```
select City, COUNT(*) AS vehicle_count
from electric_vehicles
group by  City
having COUNT(*) > 100
ORDER BY vehicle_count DESC;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[15]:

	City	vehicle_count
0	Seattle	27831
1	Bellevue	8364
2	Redmond	6032

[16]:
```
#8. Find the total Base MSRP of all electric vehicles in each Legislative␣
 ↪District, filtering districts where the total is above $10 million.
query = """
SELECT `Legislative_District`, SUM(`Base_MSRP`) AS total_base_msrp
from electric_vehicles
GROUP BY `Legislative_District`
having SUM(`Base_MSRP`) > 10000000
ORDER BY total_base_msrp DESC;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[16]:

	Legislative_District	total_base_msrp
0	41.0	14494310.0
1	48.0	14041575.0
2	45.0	13289530.0

[ ]:
```
#9. Assume you have a separate table Electric_Utility_Providers with columns␣
 ↪(Utility_ID, Electric_Utility, State). Write a query to fetch all electric␣
 ↪vehicles along with their Electric Utility Provider's State.
query = """
SELECT ev.*, eup.State AS utility_provider_state
FROM electric_vehicles ev
JOIN electric_utility_providers eup
ON ev.`Electric_Utility` = eup.`Electric_Utility`;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[18]:
```
#10. Retrieve all vehicle models that have the highest Electric Range in each␣
 ↪State using a subquery.
query = """
select ev.*
from electric_vehicles ev
JOIN ( SELECT State, MAX("Electric Range") AS max_range
    FROM electric_vehicles
    group by State  ) max_ev  ON ev.State = max_ev.State AND ev.
 ↪`Electric_Range` = max_ev.max_range;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[18]:

	VIN	County	City	State	Postal_Code	Model_Year	Make	Model	Electric_Vehicle_Type	CAFV_Eligibility	Electric_Range	Base_MSRP	Legislative_District	DOL_Vehicle_ID	Vehicle_Location	Electric_Utility	Census_Tract	Price_Category
0	7PDSGABA8P	Snohomish	Bothell	WA	98012.0	2023	RIVIAN	R1S	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b…	0	0.0	21.0	260084653	POINT (-122.1873 47.820245)	PUGET SOUND ENERGY INC	53061051927.0	Low
1	7SAYGDEE9P	Thurston	Olympia	WA	98502.0	2023	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b…	0	0.0	22.0	256162448	POINT (-122.92145 47.045935)	PUGET SOUND ENERGY INC	53067012002.0	Low
2	7FCTGAAL7N	Kitsap	Silverdale	WA	98383.0	2022	RIVIAN	R1T	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b…	0	0.0	23.0	221467284	POINT (-122.668076 47.665978)	PUGET SOUND ENERGY INC	53035091206.0	Low

[19]:
```
#11. Find the Make and Model of vehicles whose Base MSRP is higher than the␣
 ↪average Base MSRP of all vehicles.
query = """
```

```
SELECT Make, Model
from electric_vehicles
where `Base_MSRP` > (  SELECT AVG(`Base_MSRP`) FROM electric_vehicles );
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[19]:

	Make	Model
0	TESLA	MODEL S
1	BMW	330E
2	TESLA	MODEL S

[20]:
```
#12. Extract the first 3 characters from the Postal Code of each vehicle and␣
 ↪rename it as Postal_Region.
query = """
SELECT substring(`Postal_Code`, 1, 3) AS Postal_Region
from electric_vehicles;

"""

df = pd.read_sql(query, engine)
df.head(3)
```

[20]:

	Postal_Region
0	989
1	985
2	980

[22]:
```
#13. Retrieve all vehicles where the Model Name contains the word 'Tesla'␣
 ↪(caseinsensitive).
query = """
SELECT *
FROM electric_vehicles
WHERE LOWER(Make) LIKE '%%tesla%%';
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[22]:

	VIN	County	City	State	Postal_Code	Model_Year	Make	Model	Electric_Vehicle_Type
0	5YJXCBE40H	Thurston	Olympia	WA	98513.0	2017	TESLA	MODEL X	Battery Electric Vehicle (BEV)
1	5YJ3E1EB8L	King	Kent	WA	98031.0	2020	TESLA		
2	5YJ3E1EA2J	Kitsap	Port Orchard	WA	98366.0	2018	TESLA		

```
1  MODEL 3  Battery Electric Vehicle (BEV)
2  MODEL 3  Battery Electric Vehicle (BEV)

                         CAFV_Eligibility  Electric_Range  Base_MSRP  \
0  Clean Alternative Fuel Vehicle Eligible             200        0.0
1  Clean Alternative Fuel Vehicle Eligible             322        0.0
2  Clean Alternative Fuel Vehicle Eligible             215        0.0

  Legislative_District DOL_Vehicle_ID                Vehicle_Location  \
0                  2.0      257167501     POINT (-122.817545 46.98876)
1                 33.0      253771913  POINT (-122.2012521 47.3931814)
2                 26.0      280785123      POINT (-122.639265 47.5373)

                             Electric_Utility   Census_Tract Price_Category
0                       PUGET SOUND ENERGY INC  53067012331.0            Low
1  PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)  53033029305.0            Low
2                       PUGET SOUND ENERGY INC  53035092400.0            Low
```

[ ]:
```
#14. Create a new column Price_Category: 'Low' if Base MSRP < 30,000 'Mid' if␣
 ↪Base MSRP is between 30,000 and 60,000 'High' if Base MSRP > 60,000
query = """
alter table electric_vehicles
add column Price_Category VARCHAR(10);
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[ ]:
```
query = """
update electric_vehicles
set Price_Category =
    case
        when `Base_MSRP` < 30000 THEN 'Low'
        when `Base_MSRP` between 30000 AND 60000 THEN 'Mid'
        else 'High'
    END;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[23]:
```
query = """
select Price_Category from electric_vehicles;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[23]:
```
  Price_Category
0            Low
1            Low
2            Low
```

[ ]:
```
#15. Update all records where the State is NULL by replacing it with 'Unknown'.
query = """
UPDATE electric_vehicles
SET State = 'Unknown'
WHERE State IS NULL;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[ ]:
```
#16. Delete all records where Base MSRP is NULL or Electric Range is NULL.
query = """
DELETE FROM electric_vehicles
WHERE `Base_MSRP` IS NULL
   OR `Electric_Range` IS NULL;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[ ]:
```
#17. Create an index on the VIN column to improve query performance.
query = """
CREATE INDEX idx_vin
ON electric_vehicles (VIN);
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[27]:
```
#18. Use a Common Table Expression (CTE) to list all vehicles along with the␣
 ↪rank of their Electric Range within their Make.
query = """
with RankedVehicles AS (
    select *,
           rank () OVER (PARTITION BY Make ORDER BY `Electric_Range` DESC) AS␣
 ↪Range_Rank
    from electric_vehicles  ) select * from RankedVehicles;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[27]:
```
          VIN   County        City State Postal_Code  Model_Year        Make  \
0  ZASPATCW2R     King     Redmond    WA     98053.0        2024  ALFA ROMEO
1  ZASPATCW8R     King     Redmond    WA     98052.0        2024  ALFA ROMEO
2  ZASPATDWXR  Whatcom  Bellingham    WA     98226.0        2024  ALFA ROMEO

    Model                   Electric_Vehicle_Type  \
0  TONALE  Plug-in Hybrid Electric Vehicle (PHEV)
1  TONALE  Plug-in Hybrid Electric Vehicle (PHEV)
2  TONALE  Plug-in Hybrid Electric Vehicle (PHEV)

                         CAFV_Eligibility  Electric_Range  Base_MSRP  \
0  Clean Alternative Fuel Vehicle Eligible              33        0.0
1  Clean Alternative Fuel Vehicle Eligible              33        0.0
2  Clean Alternative Fuel Vehicle Eligible              33        0.0

  Legislative_District DOL_Vehicle_ID                Vehicle_Location  \
0                 45.0      259687116  POINT (-122.0222799 47.6958998)
1                 48.0      245634434      POINT (-122.12302 47.67668)
2                 42.0      251337306      POINT (-122.45493 48.76809)

                                  Electric_Utility   Census_Tract  \
0      PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)  53033032328.0
1      PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)  53033032330.0
2  PUGET SOUND ENERGY INC||PUD NO 1 OF WHATCOM CO…  53073000203.0

  Price_Category  Range_Rank
0            Low           1
1            Low           1
2            Low           1
```

[28]:
```
#   19. Use a Window Function to calculate the running total of electric␣
 ↪vehicles for each Model Year.
query = """
select `Model_Year`,
       COUNT(*) AS vehicle_count,
       SUM(COUNT(*)) OVER (PARTITION BY `Model_Year` ORDER BY `Model_Year`) AS␣
 ↪running_total
FROM electric_vehicles
group by `Model_Year`
ORDER BY `Model_Year`;
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[28]:
```
   Model_Year  vehicle_count  running_total
0        1997              1            1.0
```

```
1        1998              1            1.0
2        1999              3            3.0
```

[29]:
```
#20. Retrieve the top 5 most expensive vehicles and top 5 least expensive␣
 ↪vehicles (based on Base MSRP) in a single query using UNION.
query = """
(select *
 from electric_vehicles
 order by `Base_MSRP` desc
 LIMIT 5)
union
(SELECT *
 FROM electric_vehicles
 ORDER BY `Base_MSRP` ASC
 LIMIT 5);
"""

df = pd.read_sql(query, engine)
df.head(3)
```

[29]:

	VIN	County	City	State	Postal_Code	Model_Year	Make	Model	Electric_Vehicle_Type	CAFV_Eligibility	Electric_Range	Base_MSRP	Legislative_District	DOL_Vehicle_ID	Vehicle_Location	Electric_Utility	Census_Tract	Price_Category
0	WP0CA2A13F	King	Hunts Point	WA	98004.0	2015	PORSCHE	918	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	12	845000.0	48.0	100479039	POINT (-122.201905 47.61385)	PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)	53033024100.0	High
1	WP0AH2A73J	Clark	Vancouver	WA	98662.0	2018	PORSCHE	PANAMERA	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	14	184400.0	17.0	183245247	POINT (-122.5918493 45.6617058)	BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF C…	53011040709.0	
2	WP0AH2A76J	King	Seatac	WA	98188.0	2018	PORSCHE	PANAMERA	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	14	184400.0	33.0	277238377	POINT (-122.29179 47.43473)	PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)	53033028100.0	

```
1           High
2           High
```