Data Warehousing Project

Electric Vehicle Population Data

MIS 6309 : Business Data Warehousing

Group Members:

Sai Pranay Reddy Aare   SXA220072

Kumar Srinivasa   KXS220073

Sandeep D Arasappa SDA200000

CONTENTS

1. Abstract:

This report provides an exhaustive exploration of the Electric Vehicle (EV) population dataset sourced from the Research and Analysis Office of the Department of Licensing in Washington. Comprising 139,000 rows and 17 columns, the dataset offers an encompassing understanding of the EV landscape, with each column serving as a key to unlock insights into EV characteristics and distribution trends.

The dataset's intricate columns include:

- **Model:** Derived from Vehicle Identification Number (VIN) decoding, this column presents the specific model of each vehicle, aiding in the analysis of model preferences and market segmentation.
- **Electric Vehicle Type:** This column distin This column distinguishes between all-electric and plug-in hybrid vehicles, providing a clear classification of EVs based on their propulsion technology.
- **Clean Alternative Fuel Vehicle (CAFV) Eligibility:** Categorizing vehicles as Clean Alternative Fuel Vehicles (CAFVs), this column aligns with House Bill 2042, enacted in the 2019 legislative session, outlining fuel and electric-only range requirements for alternative fuel vehicle tax exemptions.
- **Electric Range:** Expressing the distance a vehicle can travel solely on electric power, this numeric column contributes valuable insights into the EV's operational range.
- **Base MSRP:** Providing the lowest Manufacturer's Suggested Retail Price (MSRP) for the vehicle's model, this numeric attribute aids in understanding the affordability of various EV options.
- **Legislative District:** Representing the vehicle owner's residing area within Washington State's legislative divisions, this numeric value offers insights into EV distribution across different legislative districts.
- **DOL Vehicle ID:** A unique identifier assigned by the Department of Licensing to each vehicle, this attribute facilitates streamlined data management and identification.
- **Vehicle Location:** Displaying the geographical center of the registered vehicle's ZIP Code, this point-based attribute adds a spatial dimension to the dataset, enabling geospatial analyses.
- **Electric Utility:** This column signifies the electric power retail service territories serving the vehicle's address. It encompasses diverse ownership types and territorial divisions, offering insights into power supply patterns influencing EV charging infrastructure.
- **2020 Census Tract:** Reflecting the United States Census Bureau's GEOID identifier, this alphanumeric entry connects the dataset to census tract demographics, enriching the analysis with socio-economic context.

Incorporating these diverse columns, the report offers an indispensable resource for stakeholders, policymakers, and researchers aiming to comprehend the intricacies of EV adoption, distribution, and relevant contextual factors. By leveraging the dataset's comprehensive attributes, this report contributes to the discourse on sustainable transportation solutions within the evolving landscape of electric mobility.

2. Data Source:

The dataset is open-source, and it is available to the public on the website below:
https://data.wa.gov/Transportation/Electric-Vehicle-Population-Data/f6w7-q2d2

3. Data Description:

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).
A Battery Electric Vehicle (BEV) is an all-electric vehicle using one or more batteries to store the electrical energy that powers the motor and is charged by plugging the vehicle in to an electric power source. A Plug-in Hybrid Electric Vehicle (PHEV) is a vehicle that uses one or more batteries to power an electric motor; uses another fuel, such as gasoline or diesel, to power an internal combustion engine or other propulsion source; and is charged by plugging the vehicle in to an electric power source.
Clean Alternative Fuel Vehicle (CAFV) Eligibility is based on the fuel requirement and electric-only range requirement as outlined in RCW 82.08.809 and RCW 82.12.809 to be eligible for Alternative Fuel Vehicles retail sales and Washington State use tax exemptions. Sales or leases of these vehicles must occur on or after 8/1/2019 and meet the purchase price requirements to be eligible for Alternative Fuel Vehicles retail sales and Washington State use tax exemptions.
Monthly count of vehicles for a county may change from this report and prior reports. Processes were implemented to more accurately assign county at the time of registration.
Electric Range is no longer maintained for Battery Electric Vehicles (BEV) because new BEVs have an electric range of 30 miles or more. Zero (0) will be entered where the electric range has not been researched.

Columns in this Dataset:

Column Name	Description	Data Type	Field Name
VIN (1-10)	The 1st 10 characters of each vehicle's Vehicle Identification Number (VIN).	Text	vin_1_10

County	This is the geographic region of a state that a vehicle's owner is listed to reside within. Vehicles registered in Washington state may be located in other states.	Text	county
City	The city in which the registered owner resides.	Text	city
State	This is the geographic region of the country associated with the record. These addresses may be located in other states.	Text	state
Postal Code	The 5 digit zip code in which the registered owner resides.	Text	zip_code
Model Year	The model year of the vehicle, determined by decoding the Vehicle Identification Number (VIN).	Text	model_year
Make	The manufacturer of the vehicle, determined by decoding the Vehicle Identification Number (VIN).	Text	make
Model	The model of the vehicle, determined by decoding the Vehicle Identification Number (VIN).	Text	model
Electric Vehicle Type	This distinguishes the vehicle as all electric or a plug-in hybrid.	Text	ev_type
Clean Alternative Fuel Vehicle (CAFV) Eligibility	This categorizes vehicle as Clean Alternative Fuel Vehicles (CAFVs) based on the fuel requirement and electric-only range	Text	cafv_type

	requirement in House Bill 2042 as passed in the 2019 legislative session.		
Electric Range	Describes how far a vehicle can travel purely on its electric charge	Number	electric_range
Base MSRP	This is the lowest Manufacturer's Suggested Retail Price (MSRP) for any trim level of the model in question.	Number	base_msrp
Legislative District	The specific section of Washington State that the vehicle's owner resides in, as represented in the state legislature.	Number	legislative_district
DOL Vehicle ID	Unique number assigned to each vehicle by Department of Licensing for identification purposes.	Text	dol_vehicle_id
Vehicle Location	The center of the ZIP Code for the registered vehicle.	Point	geocoded_column
Electric Utility	This is the electric power retail service territories serving the address of the registered vehicle.	Text	electric_utility
2020 Census Tract	The census tract identifier is a combination of the state, county, and census tract codes.	Text	_2020_census_tract

4. Logical Design and Physical Design:

Star Schema:

To comply with the STAR Schema, we normalized our main dataset and created 5 tables (i.e. 1 Fact table and 4 dimension tables).While performing EDA, we removed redundant columns to optimize memory and handled missing data in multiple fields by populating them with generic values to help with our analysis.

Data Cleaning and Normalization:

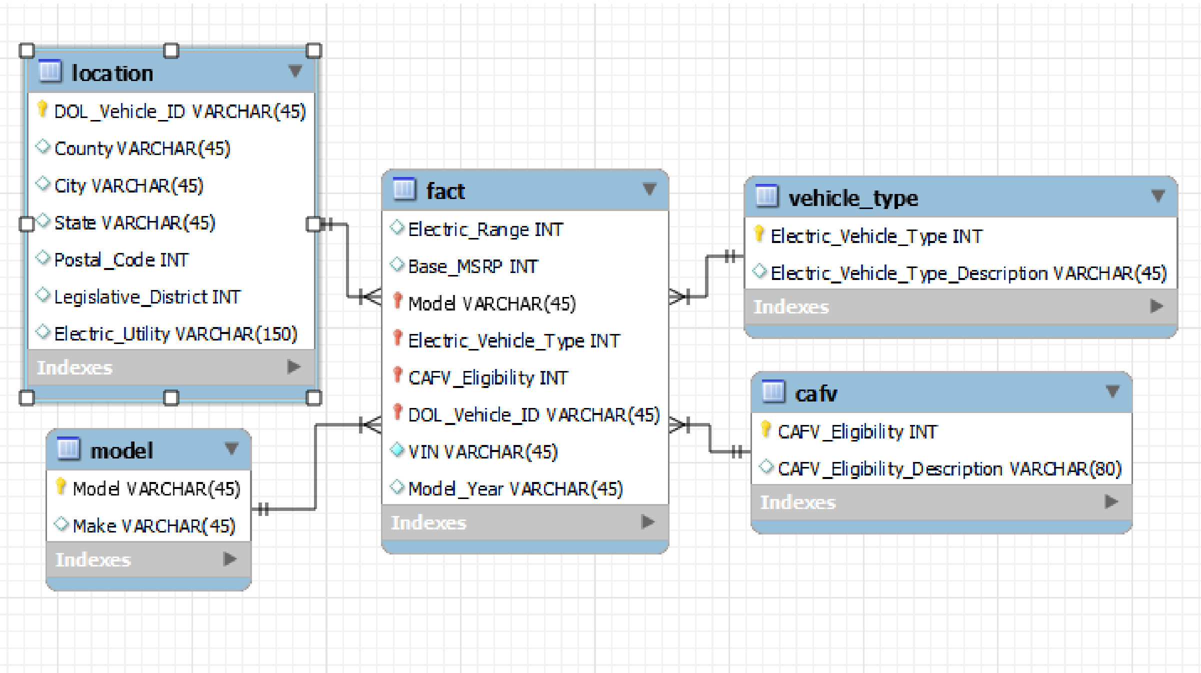
VIN (1-	County	City	State	Postal (	Model	Make	Model	Electric	Clean A	Electric	Base M	Legislat	DOL Ve	Vehicle	Electric	2020 Ce	us
1N4AZ0CF	Kitsap	Bremerto	WA	98310	2013	NISSAN	LEAF	Battery El	Clean Alte	75	0	23	2.14E+08	POINT (-1	PUGET SO	5.3E+10	
1N4AZ1CF	Kitsap	Port Orch	WA	98366	2019	NISSAN	LEAF	Battery El	Clean Alte	150	0	26	2.71E+08	POINT (-1	PUGET SO	5.3E+10	
5YJXCAE2	King	Seattle	WA	98199	2020	TESLA	MODEL X	Battery El	Clean Alte	293	0	36	8781552	POINT (-1	CITY OF SE	5.3E+10	
SADHC2S1	Thurston	Olympia	WA	98503	2019	JAGUAR	I-PACE	Battery El	Clean Alte	234	0	2	8308492	POINT (-1	PUGET SO	5.31E+10	

VIN (1-	County	City	State	Postal (	Model	Make	Model	Electric	Clean A	Electric	Base M	Legislat	DOL Ve	Vehicle	Electric	2020 Ce	us
1N4AZ0CF	Kitsap	Bremerto	WA	98310	2013	NISSAN	LEAF	Battery El	Clean Alte	75	0	23	2.14E+08	POINT (-1	PUGET SO	5.3E+10	
1N4AZ1CF	Kitsap	Port Orch	WA	98366	2019	NISSAN	LEAF	Battery El	Clean Alte	150	0	26	2.71E+08	POINT (-1	PUGET SO	5.3E+10	
5YJXCAE2	King	Seattle	WA	98199	2020	TESLA	MODEL X	Battery El	Clean Alte	293	0	36	8781552	POINT (-1	CITY OF SE	5.3E+10	
SADHC2S1	Thurston	Olympia	WA	98503	2019	JAGUAR	I-PACE	Battery El	Clean Alte	234	0	2	8308492	POINT (-1	PUGET SO	5.31E+10	

VIN (1-	County	City	State	Postal (	Model	Make	Model	Electric	Clean A	Electric	Base M	Legislat	DOL Ve	Electric	ility
1N4AZ0CF	NA	NA	AP	0	2013	NISSAN	LEAF	Battery El	Clean Alte	75	0	NA	1.12E+08	NA	
5YJ3E1EA5	NA	NA	BC	0	2019	TESLA	MODEL 3	Battery El	Clean Alte	220	0	NA	4.75E+08	NA	
5YJRE11B4	NA	NA	BC	0	2008	TESLA	ROADSTEF	Battery El	Clean Alte	220	98950	NA	1.44E+08	NA	

ERD is created in such a way that 3NF Normalization is achieved using STAR schema. The relationship between the fact table and all dimension tables is one to many.
Vehicle Location and 2020 Census Tract columns are removed because these columns determine the location, and we already have location columns. So, we removed these columns.

5. Entity Relationship Diagram(ERD):



6. Forward Engineering:

1. Schema Creation:

The script begins by dropping any existing schema named mydb and then creates a new one with the same name using the CREATE SCHEMA statement. The schema is set to use the utf8mb3 character set by default.

2.Table Creation:

The script defines four tables - cafv, location, model, and vehicle_type, using the CREATE TABLE statements. Each table is created with specific columns and data types. For example, the cafv table has columns CAFV_Eligibility (INT) and CAFV_Eligibility_Description (VARCHAR). Similarly, other tables are created with their respective columns.

3.Primary Keys and Indexes:

The script sets the primary keys and indexes for certain columns. For instance, the cafv table's primary key is set on the CAFV_Eligibility column, while indexes are defined for various columns in the fact table, like Model, Electric_Vehicle_Type, CAFV_Eligibility, and DOL_Vehicle_ID. Indexes help in optimizing search and retrieval operations.

4.Foreign Key Constraints:

The script establishes relationships between the tables using foreign key constraints. These constraints ensure referential integrity, meaning data in one table must correspond to data in another table. For example, the fact table has foreign keys referencing the model, vehicle_type, cafv, and location tables.

5.Storage Engine and Character Set:

The script explicitly specifies that all the tables will use the InnoDB storage engine and have the utf8mb3 character set by default. The storage engine determines how data is stored and managed within the tables, while the character set defines the encoding used for text data. Finally, the script resets the SQL mode and foreign key checks to their original values, ensuring that any changes made during the script execution do not affect subsequent queries.

Forward Engineering Code:

```
-- MySQL Workbench Forward Engineering

SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0;
 SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS,
FOREIGN_KEY_CHECKS=0;
```

```
 SET @OLD_SQL_MODE=@@SQL_MODE,
SQL_MODE='ONLY_FULL_GROUP_BY,STRICT_TRANS_TABLES,NO_ZERO_IN_DATE,
NO_ZERO_DATE,ERROR_FOR_DIVISION_BY_ZERO,NO_ENGINE_SUBSTITUTION';

-- -----------------------------------------------------
 -- Schema mydb
 -- -----------------------------------------------------
 DROP SCHEMA IF EXISTS `mydb` ;

-- -----------------------------------------------------
 -- Schema mydb
 -- -----------------------------------------------------
 CREATE SCHEMA IF NOT EXISTS `mydb` DEFAULT CHARACTER SET utf8mb3 ;
 USE `mydb` ;

-- -----------------------------------------------------
 -- Table `mydb`.`cafv`
 -- -----------------------------------------------------
 DROP TABLE IF EXISTS `mydb`.`cafv` ;

CREATE TABLE IF NOT EXISTS `mydb`.`cafv` (
  `CAFV_Eligibility` INT NOT NULL,
  `CAFV_Eligibility_Description` VARCHAR(80) NULL DEFAULT NULL,
  PRIMARY KEY (`CAFV_Eligibility`))
 ENGINE = InnoDB
 DEFAULT CHARACTER SET = utf8mb3;


-- -----------------------------------------------------
 -- Table `mydb`.`location`
 -- -----------------------------------------------------
 DROP TABLE IF EXISTS `mydb`.`location` ;

CREATE TABLE IF NOT EXISTS `mydb`.`location` (
  `DOL_Vehicle_ID` VARCHAR(45) NOT NULL,
  `County` VARCHAR(45) NULL DEFAULT NULL,
  `City` VARCHAR(45) NULL DEFAULT NULL,
  `State` VARCHAR(45) NULL DEFAULT NULL,
  `Postal_Code` INT NULL DEFAULT NULL,
  `Legislative_District` INT NULL DEFAULT NULL,
  `Electric_Utility` VARCHAR(150) NULL DEFAULT NULL,
  PRIMARY KEY (`DOL_Vehicle_ID`))
 ENGINE = InnoDB
 DEFAULT CHARACTER SET = utf8mb3;


-- -----------------------------------------------------
 -- Table `mydb`.`model`
```

```
 -- -----------------------------------------------------
 DROP TABLE IF EXISTS `mydb`.`model` ;

CREATE TABLE IF NOT EXISTS `mydb`.`model` (
  `Model` VARCHAR(45) NOT NULL,
  `Make` VARCHAR(45) NULL DEFAULT NULL,
  PRIMARY KEY (`Model`))
 ENGINE = InnoDB
 DEFAULT CHARACTER SET = utf8mb3;


-- -----------------------------------------------------
 -- Table `mydb`.`vehicle_type`
 -- -----------------------------------------------------
 DROP TABLE IF EXISTS `mydb`.`vehicle_type` ;

CREATE TABLE IF NOT EXISTS `mydb`.`vehicle_type` (
  `Electric_Vehicle_Type` INT NOT NULL,
  `Electric_Vehicle_Type_Description` VARCHAR(45) NULL DEFAULT NULL,
  PRIMARY KEY (`Electric_Vehicle_Type`))
 ENGINE = InnoDB
 DEFAULT CHARACTER SET = utf8mb3;


-- -----------------------------------------------------
 -- Table `mydb`.`fact`
 -- -----------------------------------------------------
 DROP TABLE IF EXISTS `mydb`.`fact` ;

CREATE TABLE IF NOT EXISTS `mydb`.`fact` (
  `Electric_Range` INT NULL DEFAULT NULL,
  `Base_MSRP` INT NULL DEFAULT NULL,
  `Model` VARCHAR(45) NOT NULL,
  `Electric_Vehicle_Type` INT NOT NULL,
  `CAFV_Eligibility` INT NOT NULL,
  `DOL_Vehicle_ID` VARCHAR(45) NOT NULL,
  `VIN` VARCHAR(45) NOT NULL,
  `Model_Year` VARCHAR(45) NULL DEFAULT NULL,
  PRIMARY KEY (`Model`, `Electric_Vehicle_Type`, `CAFV_Eligibility`, `DOL_Vehicle_ID`),
  INDEX `fk_Fact_table_Model_idx` (`Model` ASC) VISIBLE,
  INDEX `fk_Fact_table_Vehicle_type1_idx` (`Electric_Vehicle_Type` ASC) VISIBLE,
  INDEX `fk_Fact_table_CAFV1_idx` (`CAFV_Eligibility` ASC) VISIBLE,
  INDEX `fk_Fact_Location1_idx` (`DOL_Vehicle_ID` ASC) VISIBLE,
  CONSTRAINT `fk_Fact_Location1`
        FOREIGN KEY (`DOL_Vehicle_ID`)
        REFERENCES `mydb`.`location` (`DOL_Vehicle_ID`),
  CONSTRAINT `fk_Fact_table_CAFV1`
        FOREIGN KEY (`CAFV_Eligibility`)
```

```
        REFERENCES `mydb`.`cafv` (`CAFV_Eligibility`),
  CONSTRAINT `fk_Fact_table_Model`
        FOREIGN KEY (`Model`)
        REFERENCES `mydb`.`model` (`Model`),
  CONSTRAINT `fk_Fact_table_Vehicle_type1`
        FOREIGN KEY (`Electric_Vehicle_Type`)
        REFERENCES `mydb`.`vehicle_type` (`Electric_Vehicle_Type`))
 ENGINE = InnoDB
 DEFAULT CHARACTER SET = utf8mb3;


SET SQL_MODE=@OLD_SQL_MODE;
 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;
 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;
```

7. Insights:

Insight 1: Find the most popular electric vehicle type per year.

A. Electric Vehicle Adoption Trend: The output provides insights into the electric vehicle adoption trend over the years. By grouping the data by Model_Year, we can observe which years had the highest number of electric vehicles registered. The Max_Vehicle_Count column shows the peak number of electric vehicles for each year, indicating the periods of significant growth in electric vehicle adoption.
B. Popular Electric Vehicle Types: The Electric_Vehicle_Type_Description column in the output reveals the most popular types of electric vehicles for each Model_Year. By analyzing the maximum vehicle count for different electric vehicle types, we can identify which types were more prevalent during specific years. This information can help understand the preferences and demand for specific electric vehicle categories over time.
C. Demand Analysis: The output enables us to analyze the demand for electric vehicles over the years. By examining the highest vehicle counts for each year, we can identify trends in consumer interest and how the demand for electric vehicles has evolved. This insight can be valuable for businesses, policymakers, and researchers to understand the growth of the electric vehicle market and make informed decisions related to sustainability and infrastructure planning.

Query 1:

```
SELECT Model_Year, Electric_Vehicle_Type_Description, MAX(Vehicle_Count) AS
Max_Vehicle_Count
FROM (
```

```
   SELECT Fact.Model_Year, Vehicle_Type.Electric_Vehicle_Type_Description, COUNT(*)
AS Vehicle_Count
   FROM Fact
   JOIN Vehicle_Type ON Fact.Electric_vehicle_Type =
Vehicle_Type.Electric_Vehicle_Type
   GROUP BY Fact.Model_Year, Vehicle_Type.Electric_Vehicle_Type_Description
) AS SubQuery
GROUP BY Model_Year,Electric_Vehicle_Type_Description
Order by Model_Year;
```

<u>Output:</u>

Model_Year	Electric_Vehicle_Type_Description	Max_Vehicle_Count
1997	Battery Electric Vehicle (BEV)	1
1998	Battery Electric Vehicle (BEV)	1
1999	Battery Electric Vehicle (BEV)	4
2000	Battery Electric Vehicle (BEV)	9
2002	Battery Electric Vehicle (BEV)	2
2003	Battery Electric Vehicle (BEV)	1
2008	Battery Electric Vehicle (BEV)	19
2010	Battery Electric Vehicle (BEV)	21
2010	Plug-in Hybrid Electric Vehicle (PHEV)	3
2011	Battery Electric Vehicle (BEV)	710
2011	Plug-in Hybrid Electric Vehicle (PHEV)	76
2012	Battery Electric Vehicle (BEV)	786
2012	Plug-in Hybrid Electric Vehicle (PHEV)	865
2013	Battery Electric Vehicle (BEV)	2940
2013	Plug-in Hybrid Electric Vehicle (PHEV)	1630
2014	Battery Electric Vehicle (BEV)	1771
2014	Plug-in Hybrid Electric Vehicle (PHEV)	1810
2015	Battery Electric Vehicle (BEV)	3607
2015	Plug-in Hybrid Electric Vehicle (PHEV)	1314
2016	Battery Electric Vehicle (BEV)	3875
2016	Plug-in Hybrid Electric Vehicle (PHEV)	1783
2017	Battery Electric Vehicle (BEV)	4450
2017	Plug-in Hybrid Electric Vehicle (PHEV)	4105
2018	Battery Electric Vehicle (BEV)	10018
2018	Plug-in Hybrid Electric Vehicle (PHEV)	4351
2019	Battery Electric Vehicle (BEV)	8595

Insight 2: Find the top 3 vehicle makes with the most electric vehicles each year.

A. Top Vehicle Makes by Model Year: The output provides insights into the top vehicle makes for each Model_Year. By using the ROW_NUMBER() function with the PARTITION BY Fact.Model_Year clause, the query ranks the vehicle makes based on their counts (Vehicle_Count) within each Model_Year. The WHERE rn <= 3 condition restricts the results to only the top three vehicle makes for each year, based on their popularity (count of occurrences) in the dataset.
B. Popular Vehicle Makes Over Time: By analyzing the results, we can identify the vehicle makes that have consistently been among the top three most common in each Model_Year. This insight can provide valuable information about the sustained popularity and market presence of certain vehicle manufacturers across different years. It also indicates which makes have remained competitive and appealing to consumers over time.
C. Temporal Analysis of Vehicle Makes: The output enables a temporal analysis of vehicle makes' popularity. The data showcases how the top three vehicle makes change from one Model_Year to another. Changes in the ranking of vehicle makes

over time can reflect shifts in consumer preferences, the introduction of new models, or the impact of marketing and innovation strategies by different manufacturers.

Query:

```
SELECT Model_Year, Make, Vehicle_Count
FROM (
    SELECT Fact.Model_Year, Model.Make, COUNT(*) AS Vehicle_Count,
    ROW_NUMBER() OVER(PARTITION BY Fact.Model_Year ORDER BY
COUNT(*) DESC) as rn
    FROM Fact
    JOIN Model ON Fact.Model = Model.Model
    GROUP BY Fact.Model_Year, Model.Make
) AS SubQuery
WHERE rn <= 3;
```

Output:

Model_Year	Make	Vehicle_Count
1997	CHEVROLET	1
1998	FORD	1
1999	FORD	4
2000	FORD	9
2002	TOYOTA	2
2003	TOYOTA	1
2008	TESLA	19
2010	TESLA	21
2010	WHEEGO ELECTRIC CARS	3
2011	NISSAN	694
2011	CHEVROLET	76
2011	TESLA	7
2012	NISSAN	589
2012	CHEVROLET	475
2012	TOYOTA	387
2013	NISSAN	1940
2013	CHEVROLET	800
2013	TESLA	772
2014	CHEVROLET	719
2014	TESLA	654
2014	NISSAN	642
2015	NISSAN	1844
2015	TESLA	1074
2015	FORD	547
2016	TESLA	1623
2016	NISSAN	1172
2016	FORD	777
2017	CHEVROLET	2740
2017	TESLA	1641
2017	NISSAN	936
2018	TESLA	8045
2018	NISSAN	1227
2018	CHEVROLET	1154
2019	TESLA	4657

Insight 3: Find the average Base MSRP of electric vehicles per year, broken down by CAFV eligibility.

A. Average MSRP for CAFV Eligibility: The output provides insights into the average Manufacturer's Suggested Retail Price (MSRP) for vehicles categorized based on their CAFV (Clean Alternative Fuel Vehicle) eligibility. By joining the Fact table with the CAFV table on the CAFV_Eligibility column, the query groups the data by Model_Year and CAFV_Eligibility_Description. It then calculates the average MSRP for each group. The results are ordered in descending order of the average MSRP.

B. Comparing Average MSRP of Different CAFV Eligibilities: By analyzing the output, we can observe the average MSRP values for different CAFV eligibility descriptions across various Model_Years. This insight allows us to compare the price differences between clean alternative fuel vehicles with various eligibility criteria. Vehicles with higher average MSRP values might indicate premium offerings, advanced technology, or larger vehicle sizes, while lower average MSRP values could represent more affordable and accessible options.

C. Trends in Average MSRP: The output reveals trends in the average MSRP for CAFV-eligible vehicles over time. By examining the changes in average MSRP values for different Model_Years, we can identify whether the cost of clean alternative fuel vehicles has increased, decreased, or remained relatively stable over the years. These insights can help stakeholders in the automotive industry and policymakers understand the affordability and market dynamics of clean alternative fuel vehicles.

Query:

```
SELECT Fact.Model_Year, CAFV.CAFV_Eligibility_Description, AVG(Fact.Base_MSRP)
AS Average_MSRP
FROM Fact
JOIN CAFV ON Fact.CAFV_Eligibility = CAFV.CAFV_Eligibility
GROUP BY Fact.Model_Year, CAFV.CAFV_Eligibility_Description
Order by Average_MSRP DESC;
```

Output:

Model_Year	CAFV_Eligibility_Description	Average_MSRP
2010	Clean Alternative Fuel Vehicle Eligible	101205.6250
2008	Clean Alternative Fuel Vehicle Eligible	98950.0000
2014	Clean Alternative Fuel Vehicle Eligible	16314.9893
2013	Clean Alternative Fuel Vehicle Eligible	14428.5561
2018	Not eligible due to low battery range	13764.0544
2019	Not eligible due to low battery range	10436.5886
2012	Clean Alternative Fuel Vehicle Eligible	7282.0673
2016	Clean Alternative Fuel Vehicle Eligible	2986.1911
2017	Not eligible due to low battery range	2082.1834
2015	Not eligible due to low battery range	1251.8519
2020	Not eligible due to low battery range	1150.3546
2011	Clean Alternative Fuel Vehicle Eligible	970.7379
2017	Clean Alternative Fuel Vehicle Eligible	813.2538
2019	Clean Alternative Fuel Vehicle Eligible	515.0333
2016	Not eligible due to low battery range	346.0763
2018	Clean Alternative Fuel Vehicle Eligible	289.4329
2015	Clean Alternative Fuel Vehicle Eligible	0.0000
2020	Clean Alternative Fuel Vehicle Eligible	0.0000
2021	Clean Alternative Fuel Vehicle Eligible	0.0000
2022	Clean Alternative Fuel Vehicle Eligible	0.0000
2023	Clean Alternative Fuel Vehicle Eligible	0.0000
2024	Clean Alternative Fuel Vehicle Eligible	0.0000
1997	Clean Alternative Fuel Vehicle Eligible	0.0000
2022	Eligibility unknown as battery range ...	0.0000
2023	Eligibility unknown as battery range ...	0.0000
2021	Eligibility unknown as battery range ...	0.0000
2024	Eligibility unknown as battery range ...	0.0000
2020	Eligibility unknown as battery range ...	0.0000
2019	Eligibility unknown as battery range ...	0.0000
2003	Clean Alternative Fuel Vehicle Eligible	0.0000
2002	Clean Alternative Fuel Vehicle Eligible	0.0000
2021	Not eligible due to low battery range	0.0000
2022	Not eligible due to low battery range	0.0000
2023	Not eligible due to low battery range	0.0000

Insight 4: Find the county with the most electric vehicles for each electric vehicle type.

A. Top EV Types by County: The output shows the electric vehicle type (EV_Type) that has the highest number of vehicles (Max_Vehicle_Count) for each county. By joining the Fact table with the Vehicle_Type and Location tables, the query groups the data by electric vehicle type and county. It then calculates the count of vehicles for each combination and orders the results in descending order of the maximum vehicle count.
B. Popular EV Types in Counties: The results allow us to identify the most popular electric vehicle types in each county. For each county listed in the output, it shows the electric vehicle type with the highest number of vehicles. This information is valuable for understanding which types of electric vehicles are preferred or more prevalent in specific regions, indicating possible trends in consumer preferences or local policy incentives.
C. Geographical EV Adoption Patterns: The output provides insights into the geographical adoption patterns of different electric vehicle types. By examining the data by county, we can observe how the popularity of electric vehicle types varies across different regions. For instance, some counties might favor plug-in hybrids, while others might show a higher preference for battery electric vehicles or other types. These adoption patterns can be influenced by factors such as charging infrastructure availability, local environmental policies, and consumer preferences.

Overall, the output of this code offers valuable insights into the distribution of electric vehicles across different types and counties, shedding light on regional EV adoption patterns and helping stakeholders in the automotive industry, policymakers, and researchers better understand the dynamics of electric vehicle markets on a more localized level.

Query 4:

```
 SELECT EV_Type, county, MAX(Vehicle_Count) as Max_Vehicle_Count
FROM (
   SELECT Vehicle_Type.Electric_Vehicle_Type_Description AS EV_Type, Location.county,
COUNT(*) AS Vehicle_Count
   FROM Fact
   JOIN Vehicle_Type ON Fact.Electric_vehicle_Type =
Vehicle_Type.Electric_Vehicle_Type
   JOIN Location ON Fact.DOL_Vehicle_ID = Location.DOL_Vehicle_ID
   GROUP BY EV_Type, Location.county
) AS SubQuery
GROUP BY EV_Type, county
order by Max_vehicle_count DESC;
```

Output:

EV_Type	county	Max_Vehicle_Count
Battery Electric Vehicle (BEV)	King	58228
Plug-in Hybrid Electric Vehicle (PHEV)	King	14691
Battery Electric Vehicle (BEV)	Snohomish	12674
Battery Electric Vehicle (BEV)	Pierce	7844
Battery Electric Vehicle (BEV)	Clark	5912
Battery Electric Vehicle (BEV)	Thurston	3550
Battery Electric Vehicle (BEV)	Kitsap	3304
Plug-in Hybrid Electric Vehicle (PHEV)	Snohomish	3033
Plug-in Hybrid Electric Vehicle (PHEV)	Pierce	2789
Battery Electric Vehicle (BEV)	Whatcom	2534
Battery Electric Vehicle (BEV)	Spokane	2331
Plug-in Hybrid Electric Vehicle (PHEV)	Clark	2315
Plug-in Hybrid Electric Vehicle (PHEV)	Thurston	1424
Plug-in Hybrid Electric Vehicle (PHEV)	Kitsap	1278
Battery Electric Vehicle (BEV)	Benton	1147
Battery Electric Vehicle (BEV)	Skagit	1131
Battery Electric Vehicle (BEV)	Island	1101
Plug-in Hybrid Electric Vehicle (PHEV)	Spokane	1087
Plug-in Hybrid Electric Vehicle (PHEV)	Whatcom	875
Battery Electric Vehicle (BEV)	Chelan	666
Battery Electric Vehicle (BEV)	San Juan	636
Battery Electric Vehicle (BEV)	Jefferson	575
Battery Electric Vehicle (BEV)	Clallam	564
Plug-in Hybrid Electric Vehicle (PHEV)	Benton	558
Battery Electric Vehicle (BEV)	Yakima	505
Battery Electric Vehicle (BEV)	Cowlitz	479
Battery Electric Vehicle (BEV)	Mason	451
Plug-in Hybrid Electric Vehicle (PHEV)	Island	441
Plug-in Hybrid Electric Vehicle (PHEV)	Skagit	410
Battery Electric Vehicle (BEV)	Kittitas	363
Battery Electric Vehicle (BEV)	Lewis	330
Battery Electric Vehicle (BEV)	Franklin	307
Battery Electric Vehicle (BEV)	Grays Har...	303
Battery Electric Vehicle (BEV)	Grant	290

Insight 5: Find the number of Electric cars per year.

A. Model Year Car Count: The output provides a summary of the number of cars (Number_of_Cars) available in the dataset for each Model_Year. Each row in the result corresponds to a different Model_Year, and the corresponding Number_of_Cars represents the total count of cars associated with that particular year.
B. Temporal Car Distribution: By ordering the results in ascending order based on the Model_Year, the output presents a chronological distribution of cars over the years. It allows us to observe how the number of cars in the dataset has changed over time. By analyzing this temporal distribution, we can identify trends, periods of growth, or any other patterns in car availability for different model years.
C. Dataset Coverage: The output reveals the coverage of the dataset concerning different Model_Years. For instance, if the dataset covers a wide range of years, it may indicate the availability of historical car data. On the other hand, if certain years are missing or have relatively low car counts, it may highlight potential gaps or limitations in the dataset.

In summary, the output of this query provides valuable insights into the distribution of cars across different model years in the dataset, allowing for a temporal analysis of car availability and offering an overview of dataset coverage concerning model years. These insights can be beneficial for understanding the dataset's historical data, identifying trends over time, and assessing the dataset's completeness.

Query:

SELECT Model_Year, COUNT(*) as Number_of_Cars
FROM Fact
GROUP BY Model_Year
ORDER BY Model_Year;

Output:

Model_Year	Number_of_Cars
1997	1
1998	1
1999	4
2000	9
2002	2
2003	1
2008	19
2010	24
2011	786
2012	1651
2013	4570
2014	3581
2015	4921
2016	5658
2017	8555
2018	14369
2019	10506
2020	11056
2021	18258
2022	27697
2023	26986
2024	124

Insight 6: Average electric range of each vehicle type.

A. Average Electric Range by Vehicle Type: The output provides insights into the average electric range of different electric vehicle types. By joining the Fact table with the Vehicle_Type table on the Electric_vehicle_Type column, the query groups the data by Electric_Vehicle_Type_Description and calculates the average electric range (Electric_Range) for each vehicle type. This information allows us to compare the typical driving distances offered by various electric vehicle types.

B. Identifying High and Low Range EVs: By examining the output, we can identify which electric vehicle types offer the highest and lowest average electric ranges. Electric vehicle types with higher average ranges might be preferred by consumers seeking vehicles with longer driving ranges, suitable for long-distance travel without frequent charging. Conversely, electric vehicle types with lower average ranges may be more suitable for city driving and shorter commutes.

C. EV Type Performance Comparison: The output allows for a performance comparison of different electric vehicle types in terms of their electric ranges. This comparison is valuable for consumers, policymakers, and researchers interested in understanding the capabilities and limitations of various electric vehicles. It can influence decisions related to individual vehicle choices, fleet management strategies, and the development of charging infrastructure to support different electric vehicle types.

In summary, the output of this code offers insights into the average electric range of different electric vehicle types, facilitates a performance comparison, and aids in understanding which types of electric vehicles are better suited for specific driving needs. These insights are crucial for promoting sustainable transportation solutions, addressing range anxiety concerns, and guiding consumers toward electric vehicles that align with their driving habits and preferences.

Query:

```
SELECT Vehicle_Type.Electric_Vehicle_Type_Description, AVG(Fact.Electric_Range)
FROM Fact
JOIN Vehicle_Type ON Fact.Electric_vehicle_Type = Vehicle_Type.Electric_Vehicle_Type
GROUP BY Vehicle_Type.Electric_Vehicle_Type_Description;
```

<u>Output:</u>

Electric_Vehicle_Type_Description	AVG(Fact.Electric_Range)
Battery Electric Vehicle (BEV)	84.8033
Plug-in Hybrid Electric Vehicle (PHEV)	30.5813

Insight 7: Number of electric vehicles per legislative district

A. Legislative District-wise Vehicle Count: The output provides insights into the distribution of vehicles based on legislative districts. By joining the Fact table with the Location table on the DOL_Vehicle_ID column, the query groups the data by Legislative_District. Each row in the output represents a legislative district, and the corresponding count represents the total number of vehicles associated with that district. This information helps in understanding the concentration of vehicles in different legislative districts.

B. Identifying High and Low Vehicle Density Districts: By analyzing the results, we can identify legislative districts with a higher and lower number of vehicles. Legislative districts with a higher vehicle count might indicate densely populated urban areas or regions with high vehicle ownership rates. On the other hand, districts with a lower vehicle count may represent less populated or rural areas with lower vehicle ownership rates.

C. Insights for Policy and Infrastructure Planning: The output provides valuable information for policymakers and urban planners. It offers insights into vehicle distribution patterns across legislative districts, which can be used to inform decisions related to transportation infrastructure planning, traffic management, and policy initiatives. For example, areas with high vehicle density might require better public transit options or initiatives to promote electric vehicles, while regions with low vehicle density might need targeted transportation solutions.

In summary, the output of this code offers insights into the vehicle distribution across legislative districts, identifies high and low vehicle density areas, and provides useful data for policymakers and planners to make informed decisions regarding transportation policies, infrastructure development, and sustainable mobility solutions.

Query:

```
SELECT Location.Legislative_District, COUNT(*)
FROM Fact
JOIN Location ON Fact.DOL_Vehicle_ID = Location.DOL_Vehicle_ID
GROUP BY Location.Legislative_District;
```

Output:

Legislative_District	COUNT(*)
42	1933
40	3136
21	3336
25	1387
43	5414
23	3117
35	1910
48	7823
41	9255
31	2403
11	5030
36	6119
19	814
8	1440
34	4132
44	3419
46	5543
49	1913
29	872
32	3308
37	4206
12	1262
18	3713
15	323
16	713
9	749
1	5931
10	2458
5	5825
20	1182
47	2404
22	3395
24	1966
45	8516

Insight 8: Number of electric vehicles for each electric utility.

A. Electric Vehicle Adoption by Electric Utility Providers: The output provides insights into the adoption of electric vehicles based on the electric utility providers (Electric_Utility). By joining the Fact table with the Location table on the DOL_Vehicle_ID column, the query groups the data by Electric_Utility. Each row in the result represents an electric utility provider, and the corresponding count (count) indicates the total number of electric vehicles associated with that utility provider.
B. Identifying Leading Electric Utility Providers: By examining the output, we can identify the electric utility providers that have the highest number of electric vehicles associated with them. Providers with higher counts might indicate areas with stronger electric vehicle adoption rates or regions where the utility provider offers incentives or infrastructure to support electric vehicle charging.
C. Insights for Utility Providers and Government: The output offers valuable data for electric utility companies and government organizations. It provides insights into the current state of electric vehicle adoption based on the electric utility providers. Utility companies can use this information to assess the impact of electric vehicles on their infrastructure and plan for future developments in charging stations and electricity demand. Government organizations can use this data to monitor the success of electric vehicle incentives and formulate policies to promote sustainable transportation.

In summary, the output of this query provides insights into the distribution of electric vehicles across different electric utility providers, highlights the leading utility companies in terms of electric vehicle adoption, and offers valuable information for utility providers and government organizations to make informed decisions regarding infrastructure planning and policies related to electric vehicles.

Query:

```
SELECT Location.Electric_Utility, COUNT(*) as count
FROM Fact
JOIN Location ON Fact.DOL_Vehicle_ID = Location.DOL_Vehicle_ID
GROUP BY Location.Electric_Utility
Order by count DESC;
```

Output:

Electric_Utility	count
PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)	51021
PUGET SOUND ENERGY INC	27353
CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)	25379
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLARK COUNTY - (WA)	8031
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PENINSULA LIGHT COMPANY	6159
PUGET SOUND ENERGY INC||PUD NO 1 OF WHATCOM COUNTY	3203
BONNEVILLE POWER ADMINISTRATION||AVISTA CORP||INLAND POWER & LIGHT COMPANY	2065
BONNEVILLE POWER ADMINISTRATION||PUD 1 OF SNOHOMISH COUNTY	1248
PACIFICORP	1024
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF BENTON COUNTY	985
MODERN ELECTRIC WATER COMPANY	860
PUD NO 1 OF CHELAN COUNTY	826
BONNEVILLE POWER ADMINISTRATION||PUGET SOUND ENERGY INC||PUD NO 1 OF JEFFERSO...	814
BONNEVILLE POWER ADMINISTRATION||ORCAS POWER & LIGHT COOP	804
BONNEVILLE POWER ADMINISTRATION||CITY OF RICHLAND - (WA)	719
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF COWLITZ COUNTY	693
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLALLAM COUNTY	680
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PUD NO 3 OF MASON COU...	570
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF GRAYS HARBOR COUNTY	479
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF FRANKLIN COUNTY	414
PUD NO 2 OF GRANT COUNTY	414
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||ELMHURST MUTUAL POWE...	340
NON WASHINGTON STATE ELECTRIC UTILITY	312
AVISTA CORP	302
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PUD NO 1 OF LEWIS COUNTY	280
BONNEVILLE POWER ADMINISTRATION||VERA IRRIGATION DISTRICT #15	278
PUD NO 1 OF DOUGLAS COUNTY	270
BONNEVILLE POWER ADMINISTRATION||INLAND POWER & LIGHT COMPANY	260
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF KLICKITAT COUNTY	223
NO KNOWN ELECTRIC UTILITY SERVICE	197
CITY OF TACOMA - (WA)||TANNER ELECTRIC COOP	196
BONNEVILLE POWER ADMINISTRATION||PACIFICORP||PUD NO 1 OF CLARK COUNTY - (WA)	187
BONNEVILLE POWER ADMINISTRATION||CITY OF PORT ANGELES - (WA)	170

Result 9

Insight 9: Model year with the highest average base MSRP for electric vehicles.

A. Peak Average Manufacturer's Suggested Retail Price (MSRP) Year: The output identifies the specific Model_Year that had the highest average MSRP among all the years in the dataset. This insight highlights the year when vehicles were, on average, priced at their highest point in the market. It indicates a period when luxury or high-end vehicles might have been prevalent or when new, advanced, and more expensive models were introduced.

B. Trends in Vehicle Pricing: By analyzing the year with the highest average MSRP, we can gain insights into historical trends in vehicle pricing. It helps in understanding how the average pricing of vehicles has evolved over time and provides a glimpse into consumer preferences for premium features and luxury vehicles during that specific year.

C. Market Dynamics and Customer Preferences: The output sheds light on market dynamics and customer preferences during the year with the highest average MSRP. A surge in average MSRP might indicate strong demand for luxury or high-performance vehicles, signaling a willingness among consumers to pay a premium for certain vehicle features or brand names.

In summary, the output of this code provides valuable insights into the year with the highest average MSRP, highlighting historical trends in vehicle pricing, consumer preferences for premium vehicles, and market dynamics during that specific period. This information can be valuable for automotive manufacturers, researchers, and analysts seeking to understand historical market shifts and identify potential opportunities for high-end vehicle offerings.

Query:

```
SELECT Fact.Model_Year, AVG(Fact.Base_MSRP) as Average_MSRP
FROM Fact
GROUP BY Fact.Model_Year
ORDER BY Average_MSRP DESC
LIMIT 1;
```

Output:

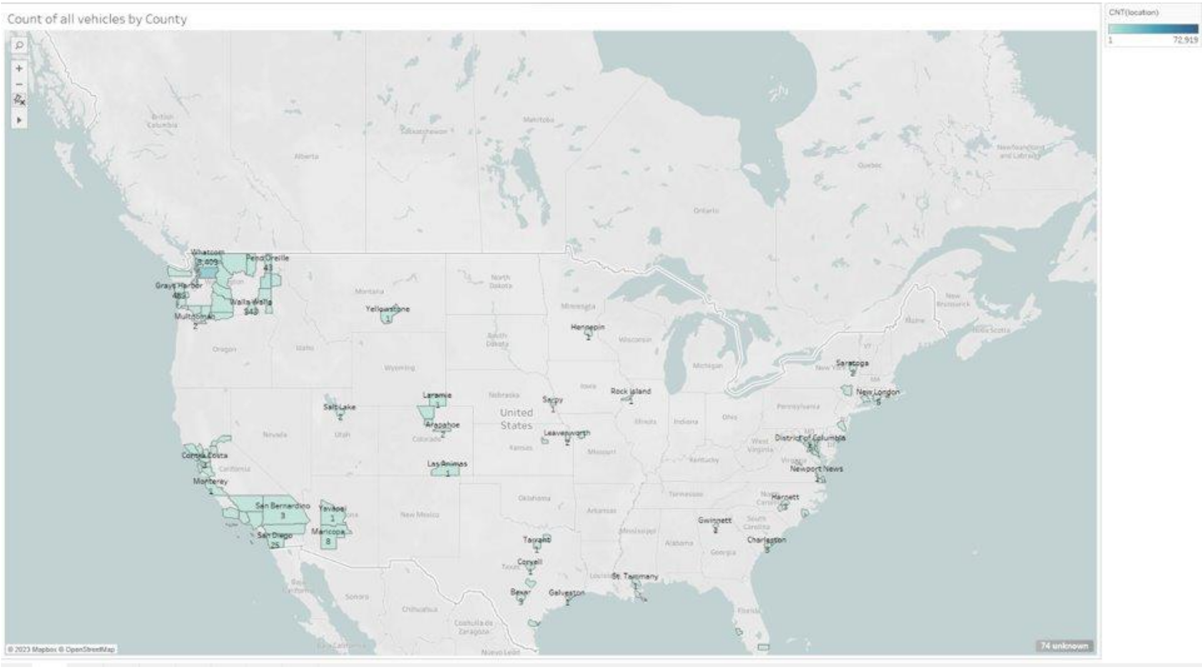
Result Grid | Filter Rows:

Model_Year	Average_MSRP
2010	101205.6250

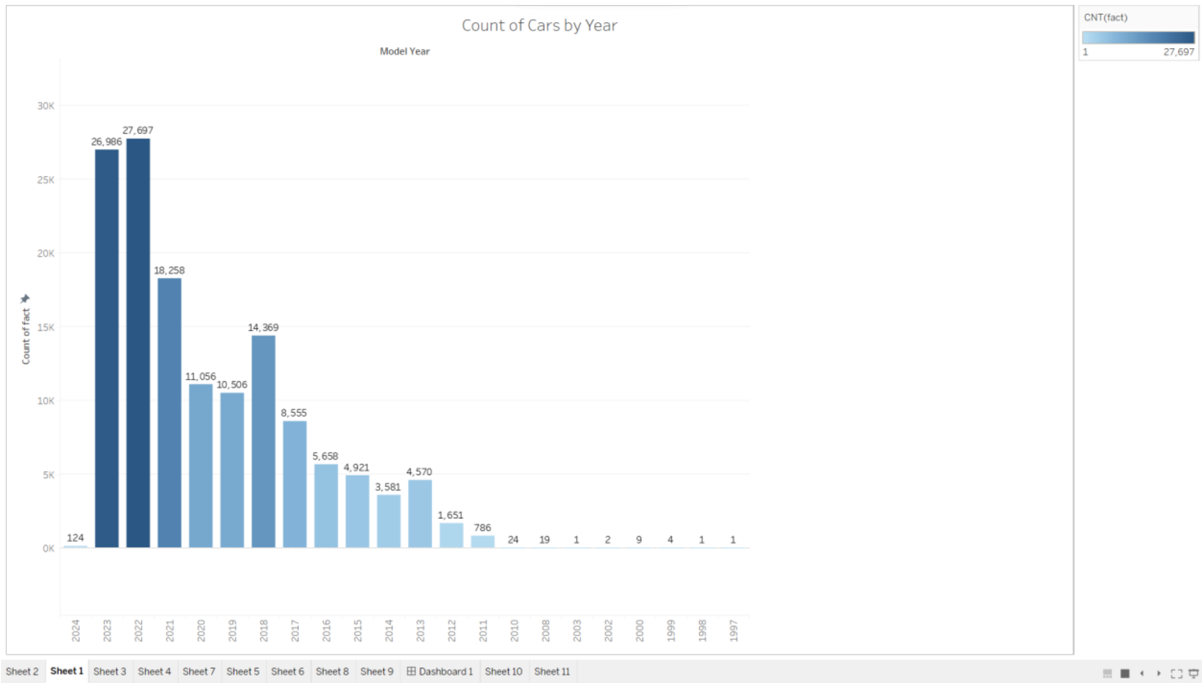
8. Visualizations:

Business Intelligence Tool: Tableau

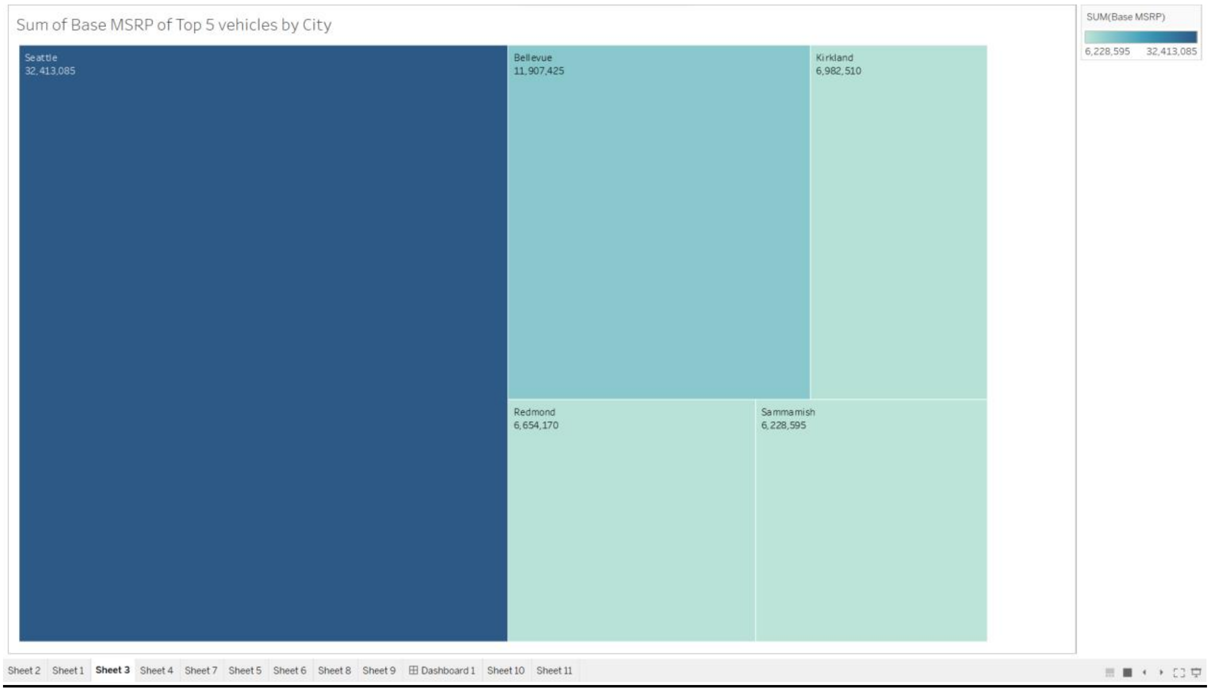
1. Number of Electric Vehicles by County:



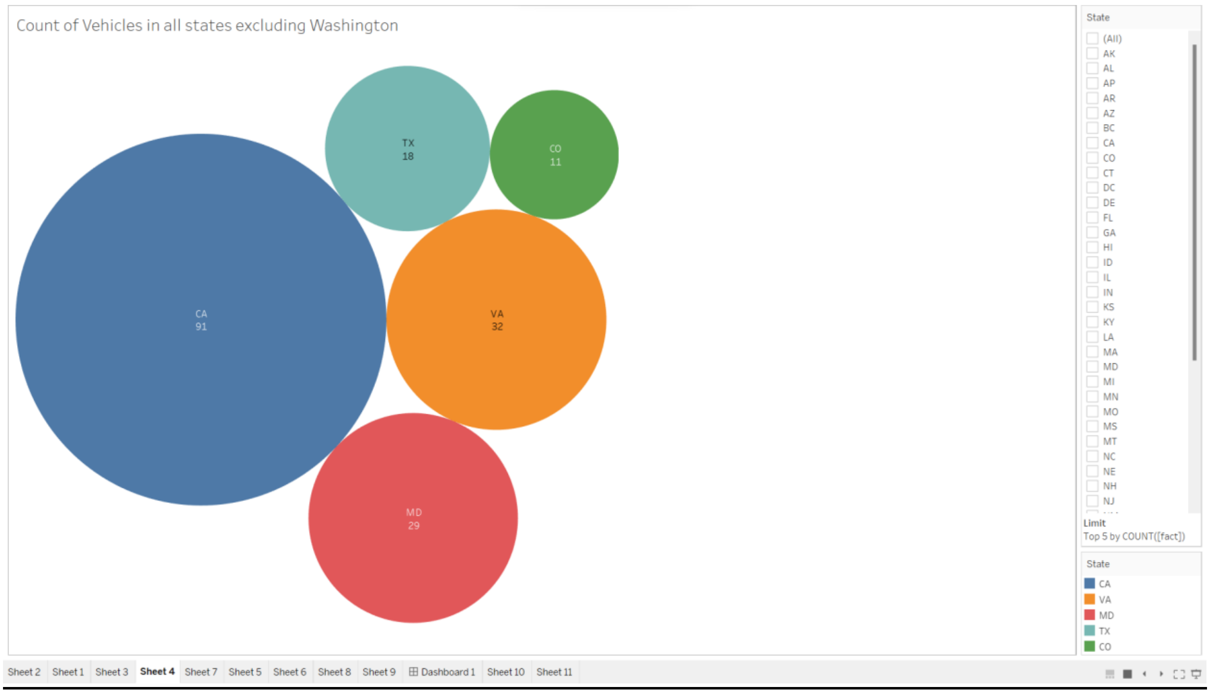
2. Number of EV's by year:



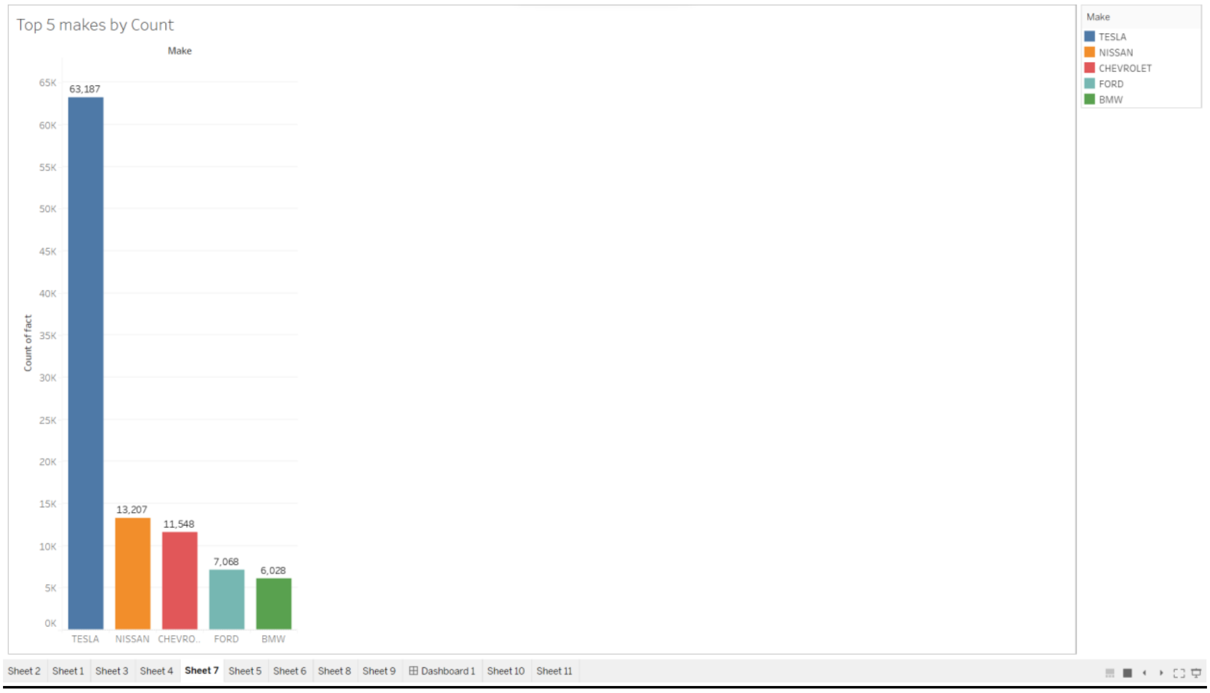
3. Top 5 Cities by Sum of Base MSRP:



4. Number of EV's in all states excluding Washington:



5. Top 5 Manufacturing Companies:



6. Count of Vehicles by EV Type:

7. Dashboard of Tesla Vehicles: