

VEHICLE ANALYSIS

“Gaining insights into vehicle distribution and trends across different factors”

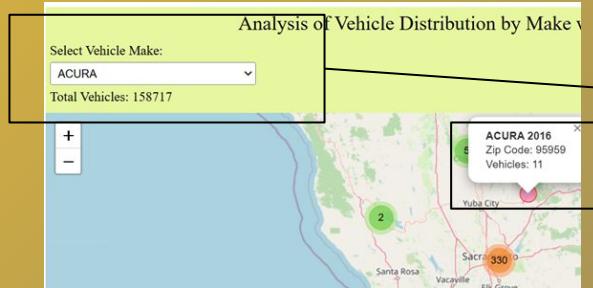
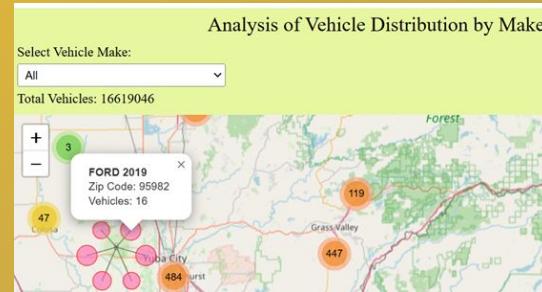
The project analyzes vehicle distribution in California to help policymakers and the public understand transportation trends and environmental impacts.

Vehicle Analysis Workflow: Cleaning, Integration, and Export

1. **Data Collection & Cleaning:**
 - o Read data from the CSV file vehicle-fuel-type-count-by-zip-code-20231.csv.
 - o Retrieve latitude and longitude for unique ZIP codes using the Geocode API and add to the CSV file.
 - o Clean data by removing invalid values and standardizing model_year.
2. **Database Integration:**
 - o Use SQLAlchemy to define the Vehicle table and map cleaned data into PostgreSQL.
 - o Insert cleaned records from vehicle_cleaned_output.csv into the database.
3. **Data Querying & Export:**
 - o Query the Vehicle table in PostgreSQL and load data into a pandas DataFrame.
 - o Export the DataFrame to a CSV file (vehicle_cleaned.csv).
4. **Outcome:**
 - o Successfully process, integrate, and export data for further analysis.
 - o The exported data file vehicle_cleaned.csv is used to generate various visualizations.

Data Visualization - Analysis of Vehicle Counts and Make for California State

Marker Cluster Map



Interactive Map:

- Select a vehicle make from the dropdown.
- Map updates to show data for the selected make.
- Default shows all makes.

Marker Clustering:

- Clusters group nearby markers for better performance.
- Zooming in reveals individual markers.

Vehicle Information:

- Markers represent vehicles by ZIP code, make and model year.
- Markers have popups that display vehicle count, model year, and ZIP code information
- Marker size indicates vehicle count.

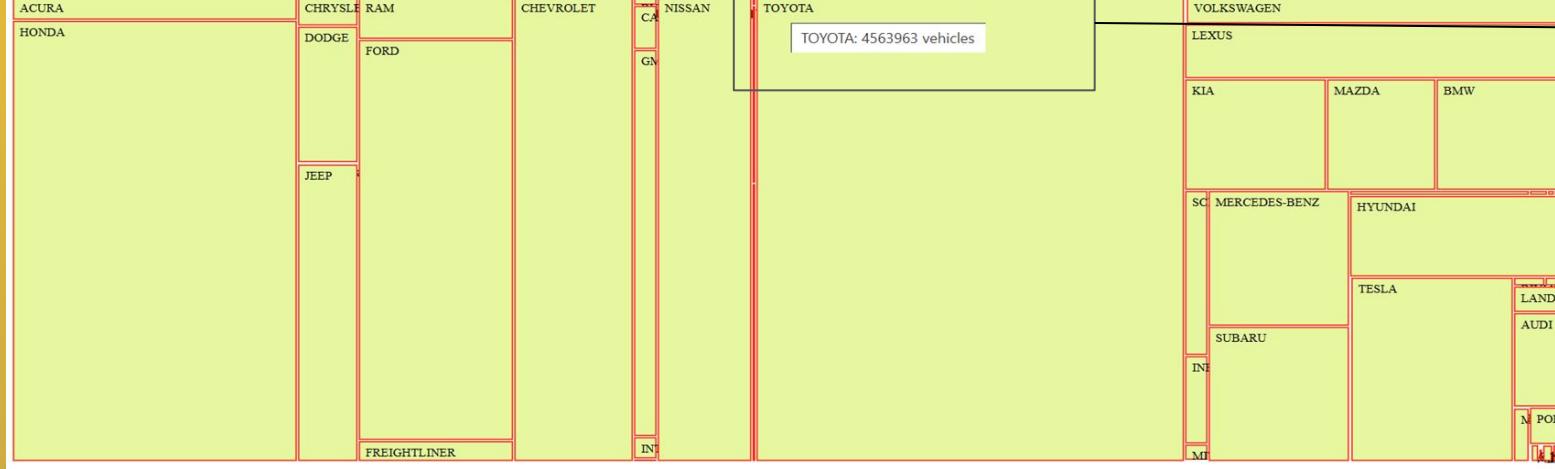
Selection of “Acura” make from dropdown shows a Total Vehicle Count of **158717**

Marker referring Zip Code: **95959**, Make: **ACURA**
2016, Vehicles:**11**

Data Insights: Treemaps, Bar and Donut Charts

Treemap

Vehicle Counts by Make (Treemap)



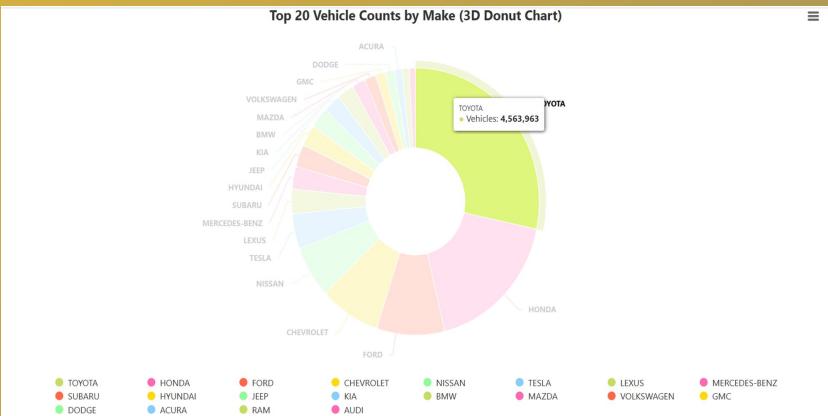
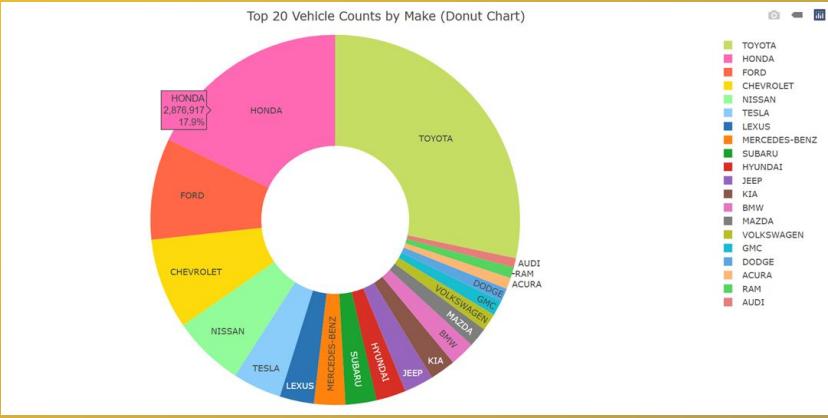
Bar Chart



- Displays a bar chart of vehicle counts for the top 20 and bottom 20 vehicle makes based on aggregated data.
- Provides insights into which vehicle makes dominate in terms of vehicle count.
- Interactive dropdown allows users to toggle between the top and bottom vehicle makes.

Interactive
DropDown Selection

Donut Chart (2D and 3D)



- Visualizes the top 20 vehicle makes as a donut chart, offering a clear comparison of vehicle counts across the most popular makes.
- Each slice represents a vehicle make with its corresponding count.
- On hovering over a specific slice, a popup shows the vehicle count for the make represented by that slice.

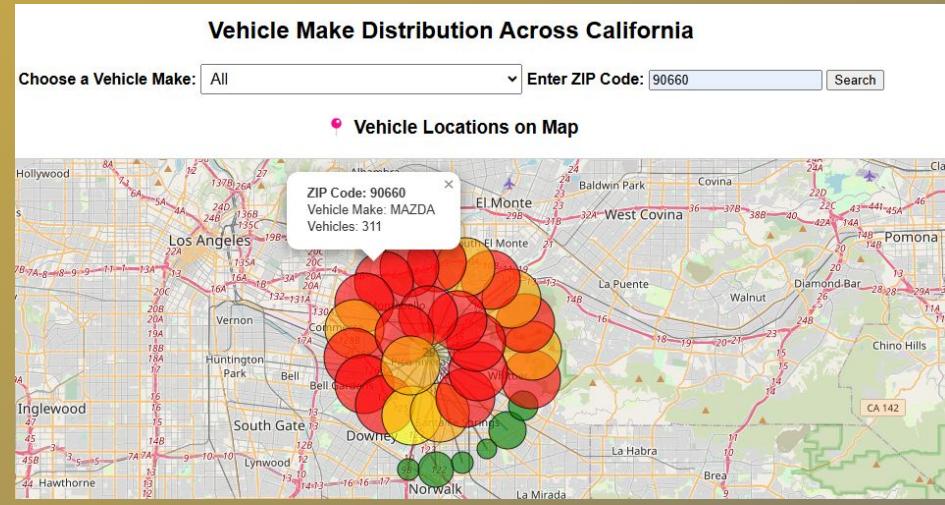
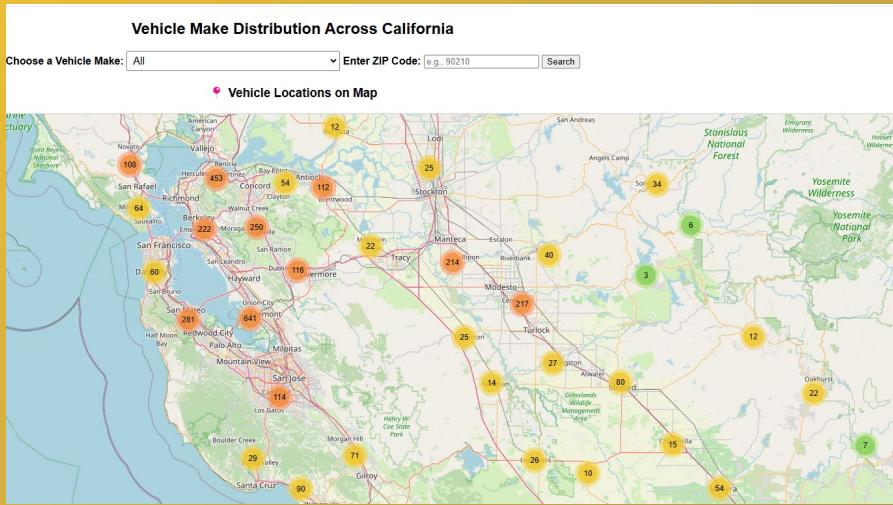
[GitHub Pages Link for Data Visualization](#)

Vehicle Age by Zip Code

- ❖ In summary, vehicle age is important for evaluating pollution and can help shape better policies, like:
 - Incentives for switching to cleaner cars
 - Stricter emission regulations

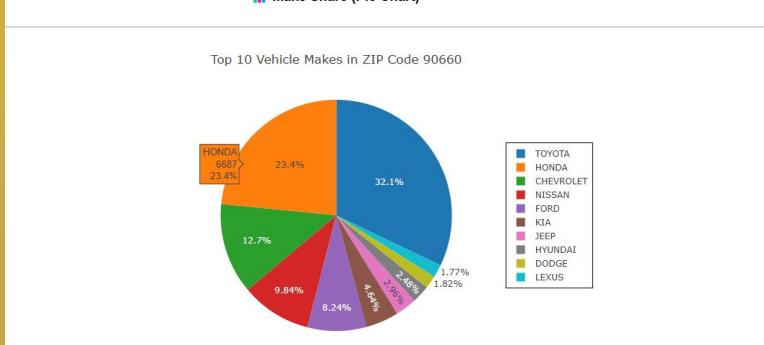
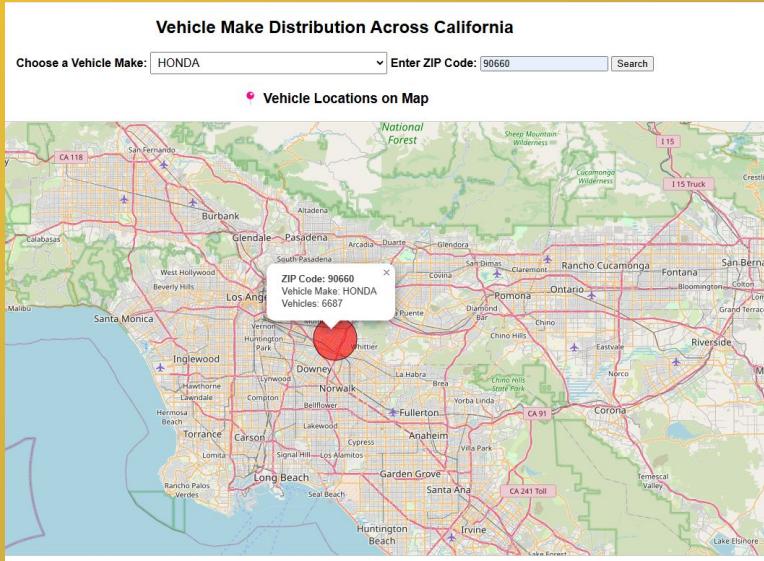


Interactive Marker Cluster Map: Vehicle Distribution Across California



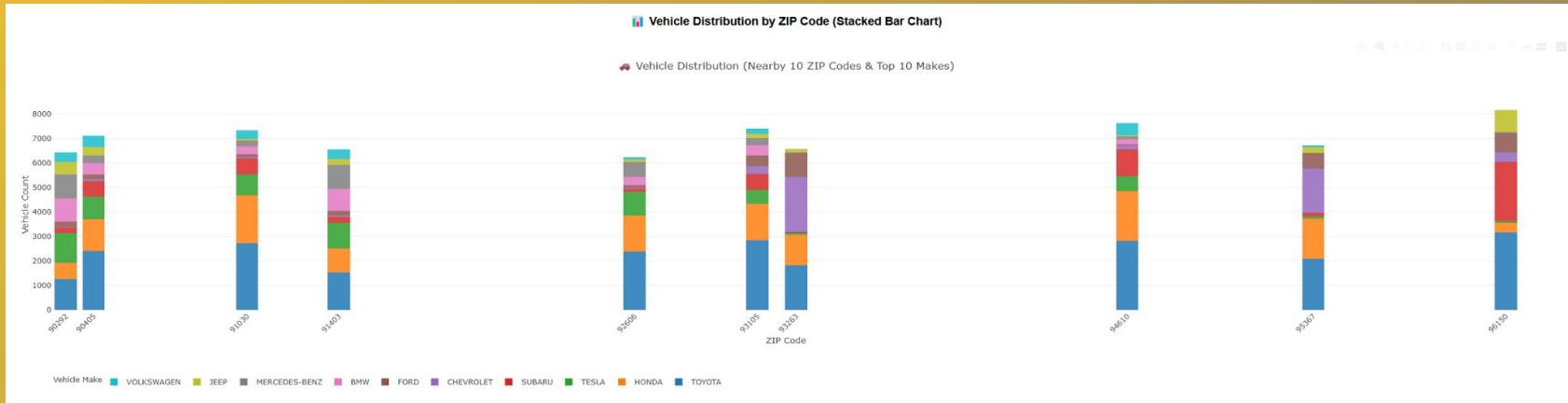
- **Left Image:** Statewide vehicle distribution with cluster markers.
- **Right Image:** Detailed vehicle distribution within a selected ZIP code, showing a breakdown of individual makes

Ensuring Data Consistency in Vehicle Make Distribution by ZIP Code



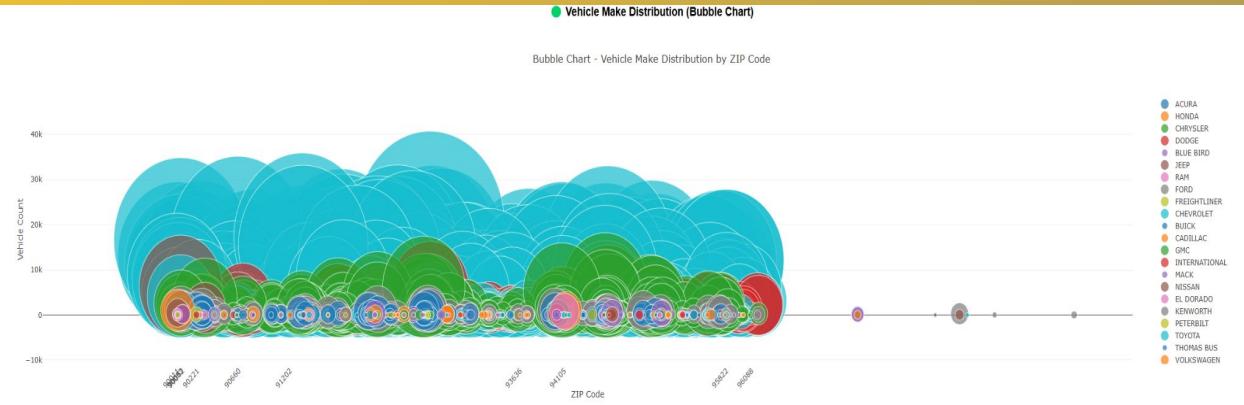
- **Selected ZIP Code:** 90660
- **Vehicle Make:** Honda
- **Vehicle Count:** **6,687** (Shown in both visuals)
- Top Vehicle Share: Honda represents **23.4%** of vehicles in ZIP 90660.
- **Map View (Top Image):** Shows Honda's total vehicle count in the selected ZIP code.
- **Pie Chart (Bottom Image):** Displays Honda's proportion compared to other vehicle makes in the same area.
- **Key Takeaway:** Cross-validation across different visualizations ensures accuracy and reliability in vehicle distribution analysis.

Comparing Vehicle Make Distribution Across Nearby ZIP Codes



- ❑ **Visualization Type & Purpose:** Stacked Bar Chart displays the distribution of the top **10** vehicle makes across **10** nearby ZIP codes.
- ❑ **Key Insights**
 - Each bar represents a ZIP code, showing how different car makes are distributed.
 - **Toyota** and **Honda** dominate multiple ZIP codes.
 - Vehicle density varies, with **92606** having the **lowest** count and **96150** the **highest**.
 - Some ZIP codes have a **higher concentration** of specific brands, showing **regional preferences**.
- ❑ **Use Case:** Helps analyze regional trends in vehicle ownership, which can be useful for dealerships, policymakers, and urban planners.

Analyzing Vehicle Make Distribution with Bubble Charts



Visualization Type: Bubble Chart

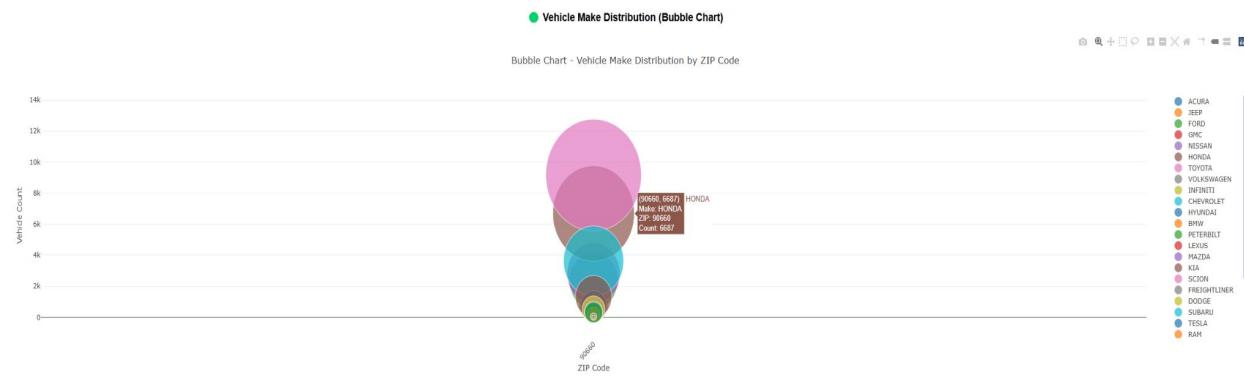
Purpose: Represents vehicle distribution across different ZIP codes. **Larger bubbles** represent higher vehicle counts, providing a clear comparative view of vehicle distribution.

Top Image: Shows the overall distribution of vehicle makes across all ZIP codes.

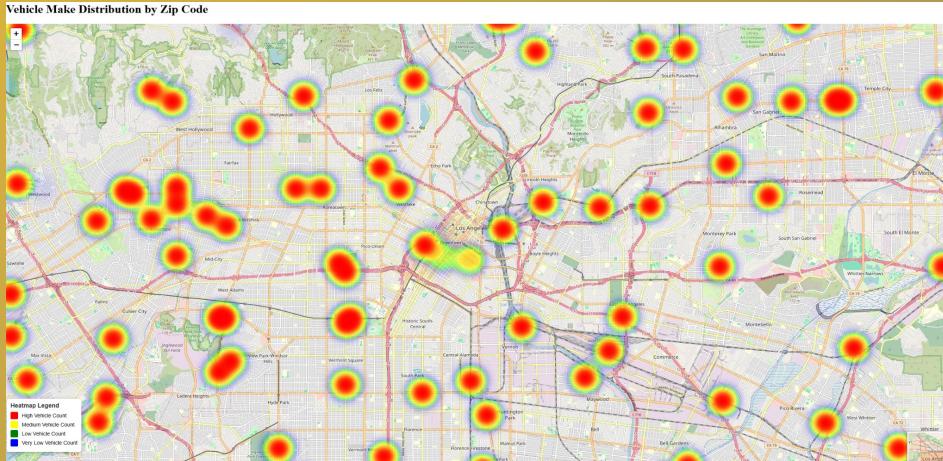
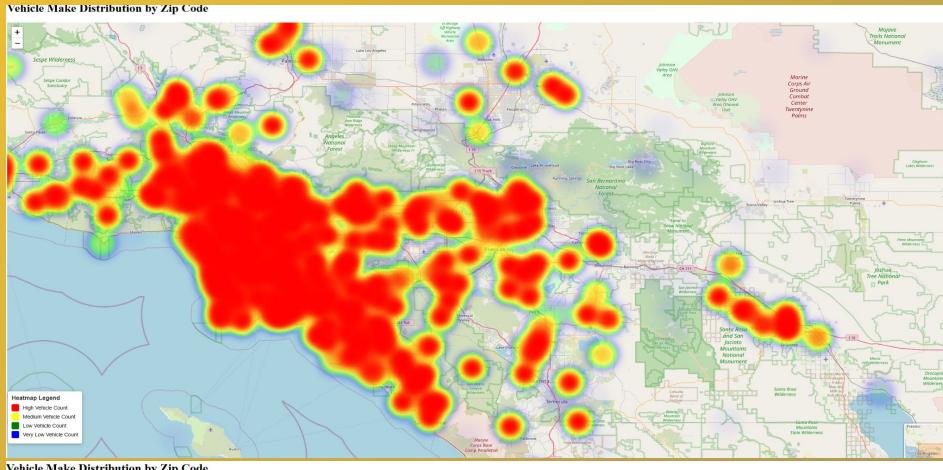
Bottom Image: Focuses on ZIP code **90660**, highlighting **Honda** with a count of **6,687** vehicles to showcase data consistency across different visualization methods.

Key Insight: The count remains consistent across multiple visualizations, reinforcing data accuracy and reliability.

✓ **Key Takeaway:** Cross-referencing results across multiple visualization formats ensures **data reliability** and **consistency** in vehicle distribution analysis.



Visualizing Vehicle Distribution Using Heatmaps



Purpose: Heatmaps highlight areas with high and low vehicle concentrations.

Top Image: Displays statewide vehicle distribution with intense red zones indicating high-density areas.

Right Image: Zoomed-in view of a metropolitan area, showing localized vehicle distribution.

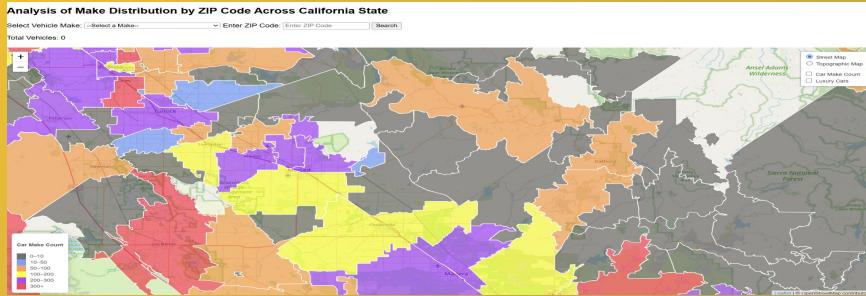
Color Meaning:

- **Red:** High vehicle density.
- **Yellow:** Moderate vehicle density.
- **Green/Blue:** Low vehicle density.

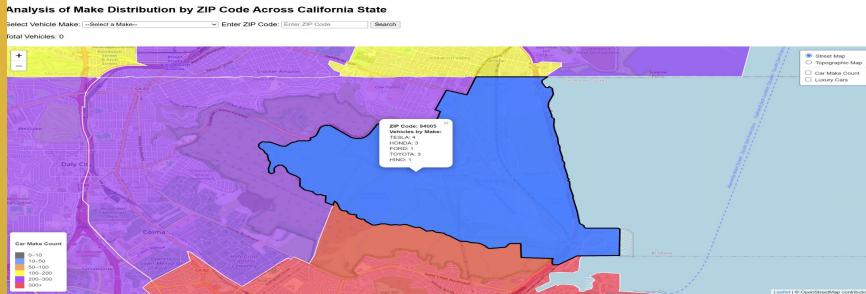
Small map: The small map offers a zoomed-in view of a metropolitan area, highlighting vehicle concentration and potential traffic congestion points.

Key Takeaway: Heatmaps offer valuable geographic insights into vehicle density and distribution, aiding urban planners and market analysts in decision-making.

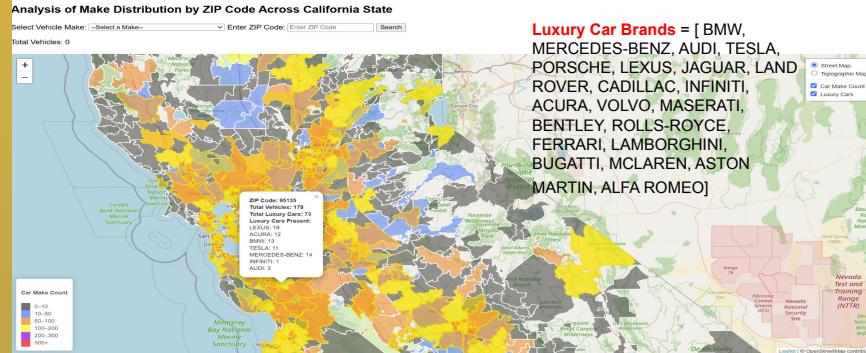
Analyzing Vehicle Distribution by ZIP Code Using Choropleth Maps



Purpose: Choropleth maps use color-coded regions to represent vehicle distribution patterns, providing insights into ZIP code divisions, vehicle density, and dominant car brands.



Top Image: Displays the ZIP code divisions across the state based on the predefined legend. Since no search key has been selected, it does **not** yet show vehicle distribution or dominant car makes.

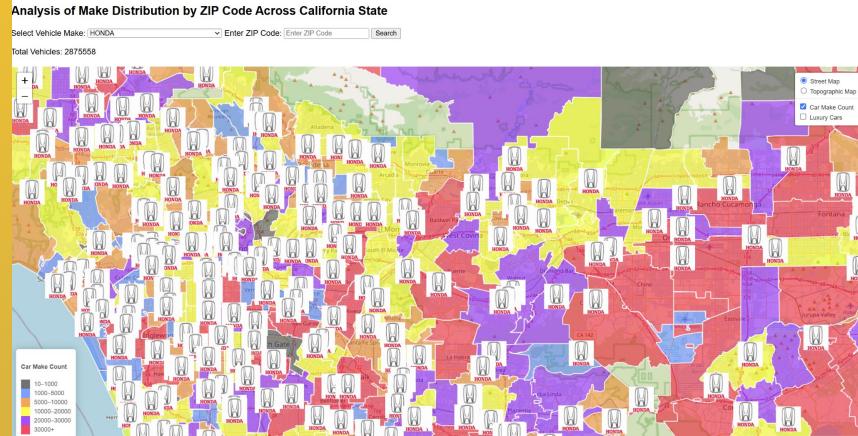


Middle Image: Focuses on a specific ZIP code that has been clicked, highlighting its boundaries in **black**. It provides a breakdown of various vehicle makes within that ZIP code, allowing for a more detailed analysis of local car preferences.

Bottom Image: Highlights the distribution of luxury vehicles, showcasing areas where premium and high-end cars are more prevalent.

Key Takeaway: These maps provide geographic insights into vehicle ownership, helping dealerships, manufacturers, and policymakers identify market trends and optimize regional strategies.

Searching for a Specific Vehicle Make Across ZIP Codes



Top Image:

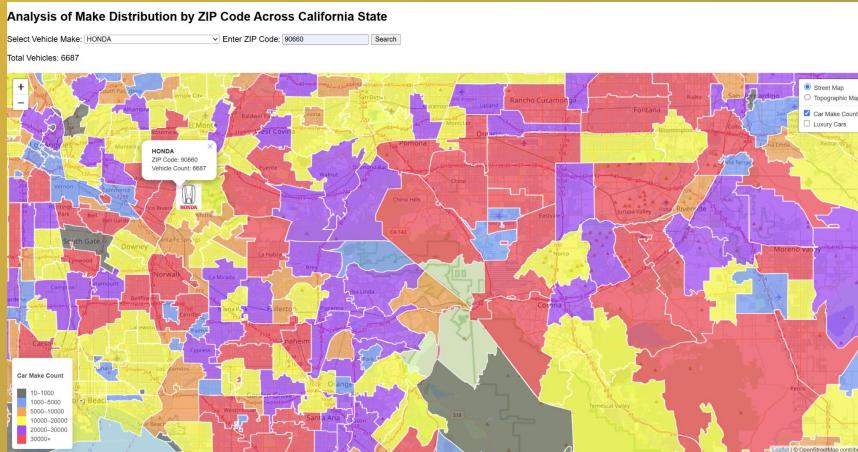
- Honda was selected as the vehicle make, displaying Honda's presence across all ZIP codes in the state.
- The **Honda logo** appears in ZIP codes where Honda vehicles are registered, providing a quick visual of its distribution.

Bottom Image:

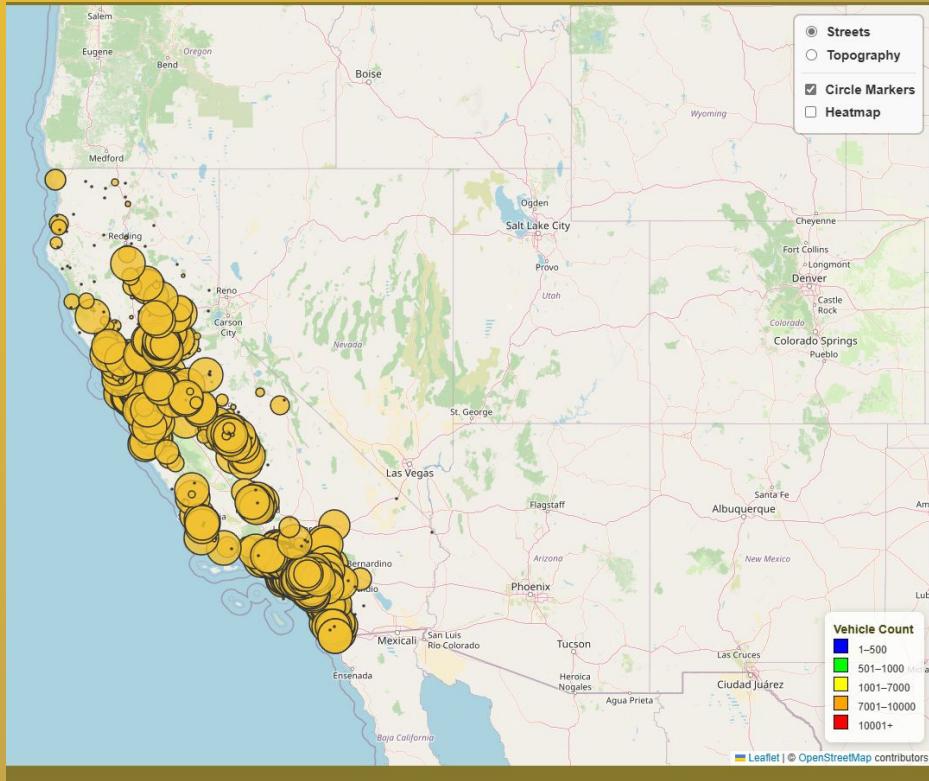
- A specific search was conducted for Honda in ZIP code **90660**.
- Clicking on the **Honda logo** within the selected ZIP code reveals the total Honda vehicle count (**6,687**).
- This result matches previous visualizations (Pie Chart, Bubble Chart, and Marker Cluster Map), reinforcing **data consistency**.

Key Takeaway:

- Filtering by vehicle make and ZIP code allows for detailed regional analysis.
- Cross-verifying results across multiple visualization formats ensures data reliability and accuracy.

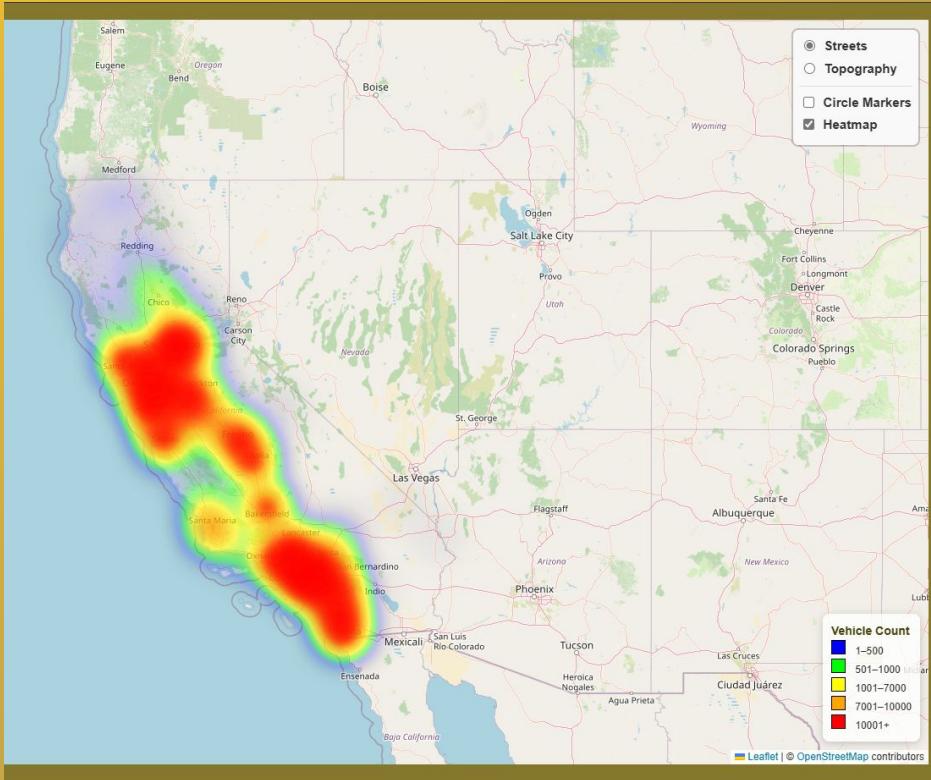


Vehicles by Zip Code - Circle Markers



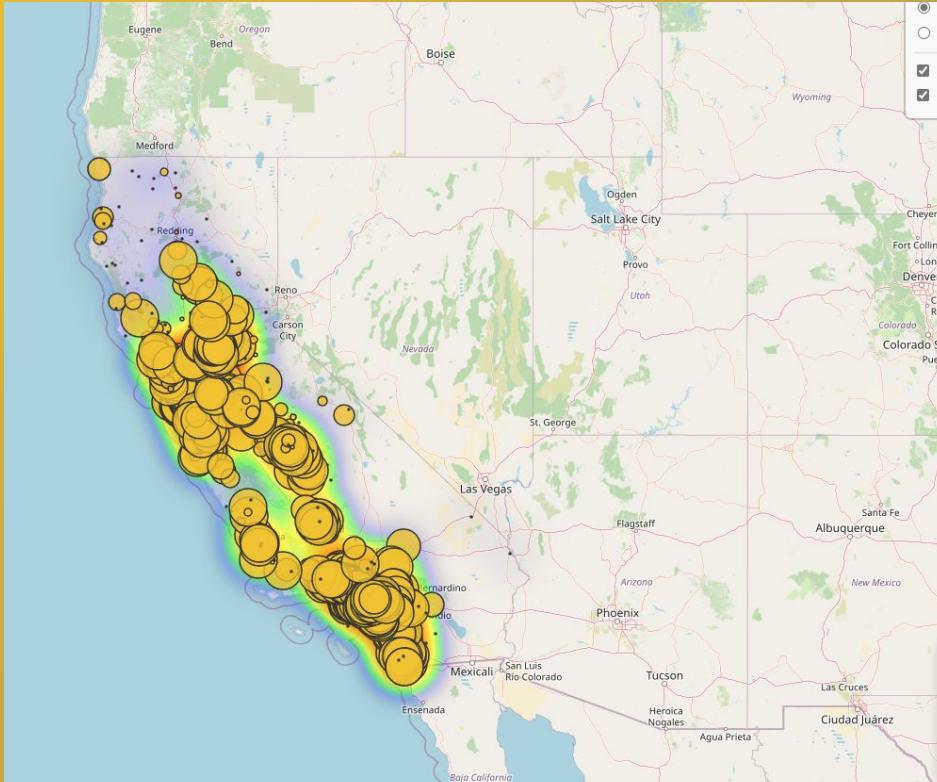
- The interactive circle markers show the representative size to the vehicle density of each zip code
- The bigger the circle, the more cars in that zip code
- Things to consider:
 - Geographical areas
 - Commutes

Vehicles by Zip Code - Heatmap



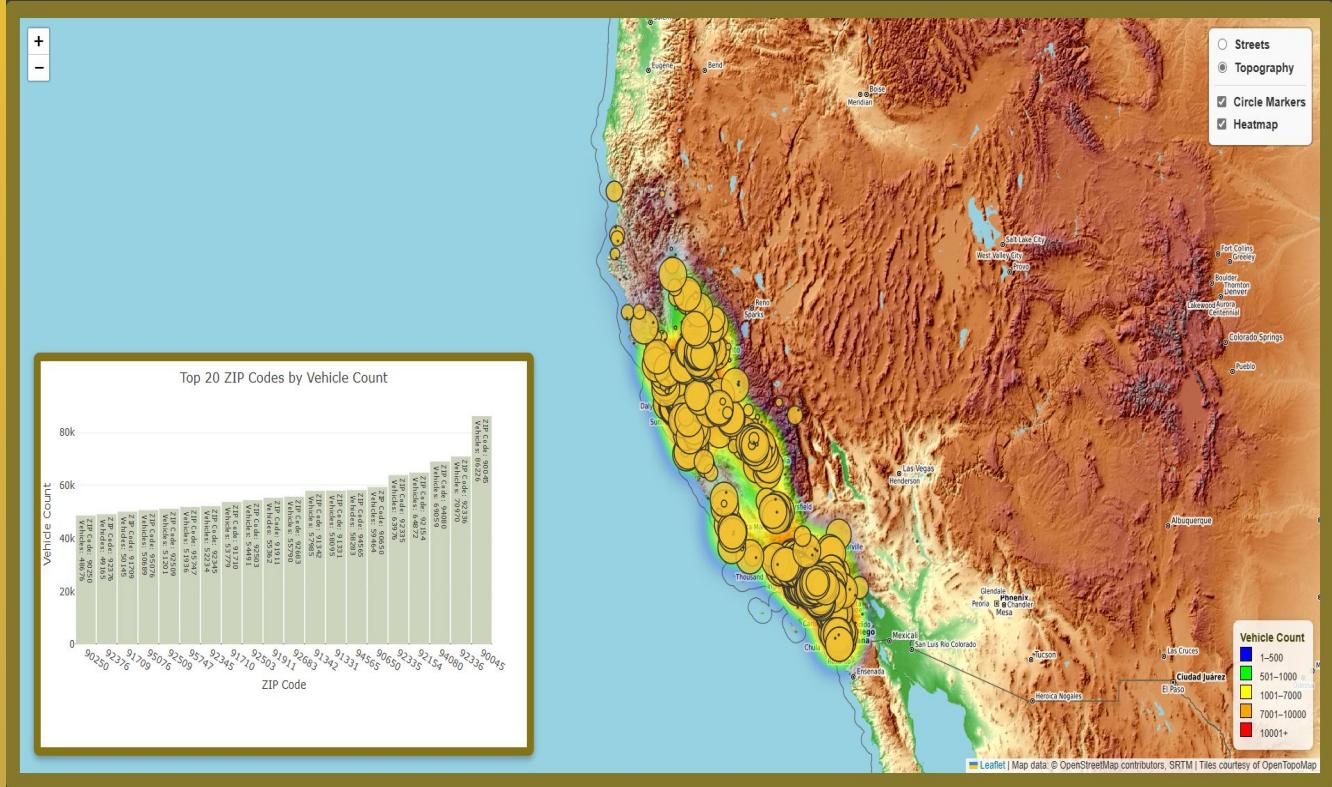
- See legend for this map
- Density is properly portrayed by the use of colors
- Things to consider:
 - Max zoom and Blur
 - Overlap in zip codes

Vehicles by Zip Code - Layers



- Is there a reason for the main location?
- Geography - Mountains!

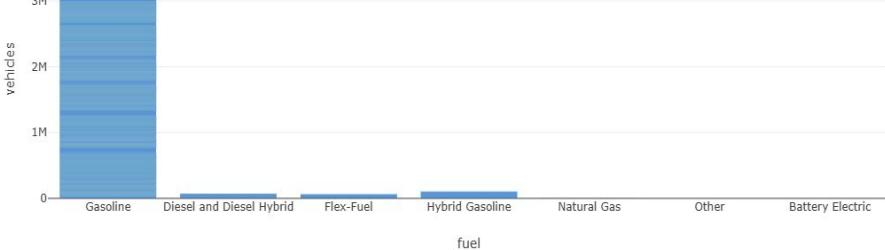
Vehicles by Zip Code - Satellite



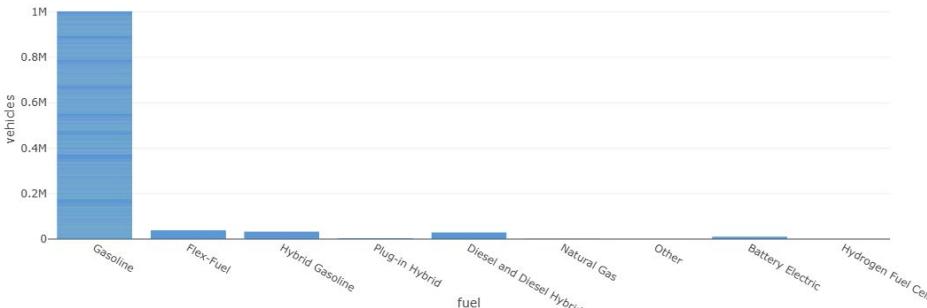
FUEL TYPE DISTRIBUTION BY MODEL YEAR

Fuel Type Distribution By Model Year (Bar Chart)

Vehicle Fuel Type Count By Model Year



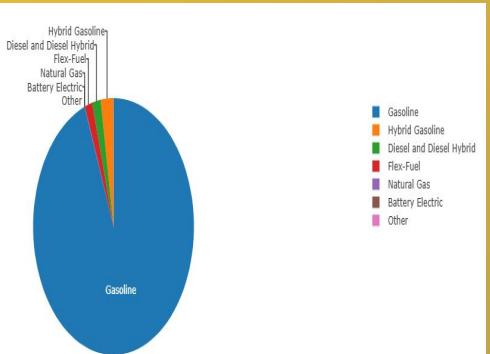
Vehicle Fuel Type Count By Model Year



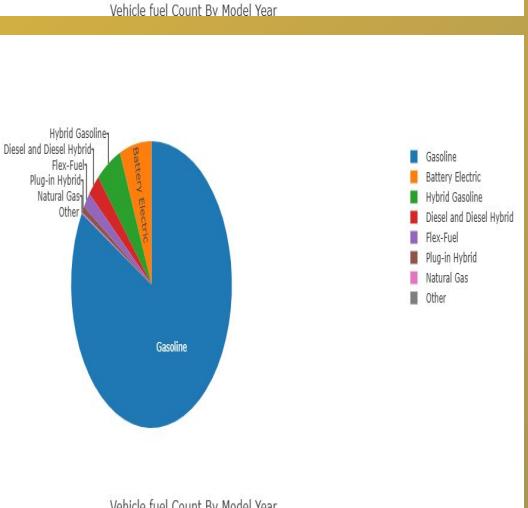
From the visualizations it's evident that gasoline has been the most dominant fuel type for all model years .

<https://baron14-coder.github.io/page/>

Fuel Type Distribution By Model Year(pie chart)



Hybrid gasoline cars ,diesel and diesel hybrid fuel types are more distributed with more recent car models And are strong alternative car fuel types according to our visualizations.



<https://baron14-coder.github.io/page/>

In summary gasoline is the most common fuel type in California this could be due to how economical it is and better performance . hybrid gasoline and diesel became more popular with the more recent car models .we can see that BATTERY electric cars were popular in 2020 models but had a huge decline with more recent models since gasoline and diesel hybrid fuel types became more in use flex fuel was really popular in 2016 models but shows evidence of engine damage so was most likely dropped in more recent car models

natural gas and hydrogen fuel cell, plug in hybrid are least used fuel types in California.

<https://baron14-coder.github.io/page/>

Conclusion

Analyzing vehicle age, count, make, and fuel type by ZIP code helps understand California's transportation trends and supports sustainable initiatives. The insights can promote cleaner transportation options.

<https://github.com/mctrashmoney/vehicle-analysis.git>