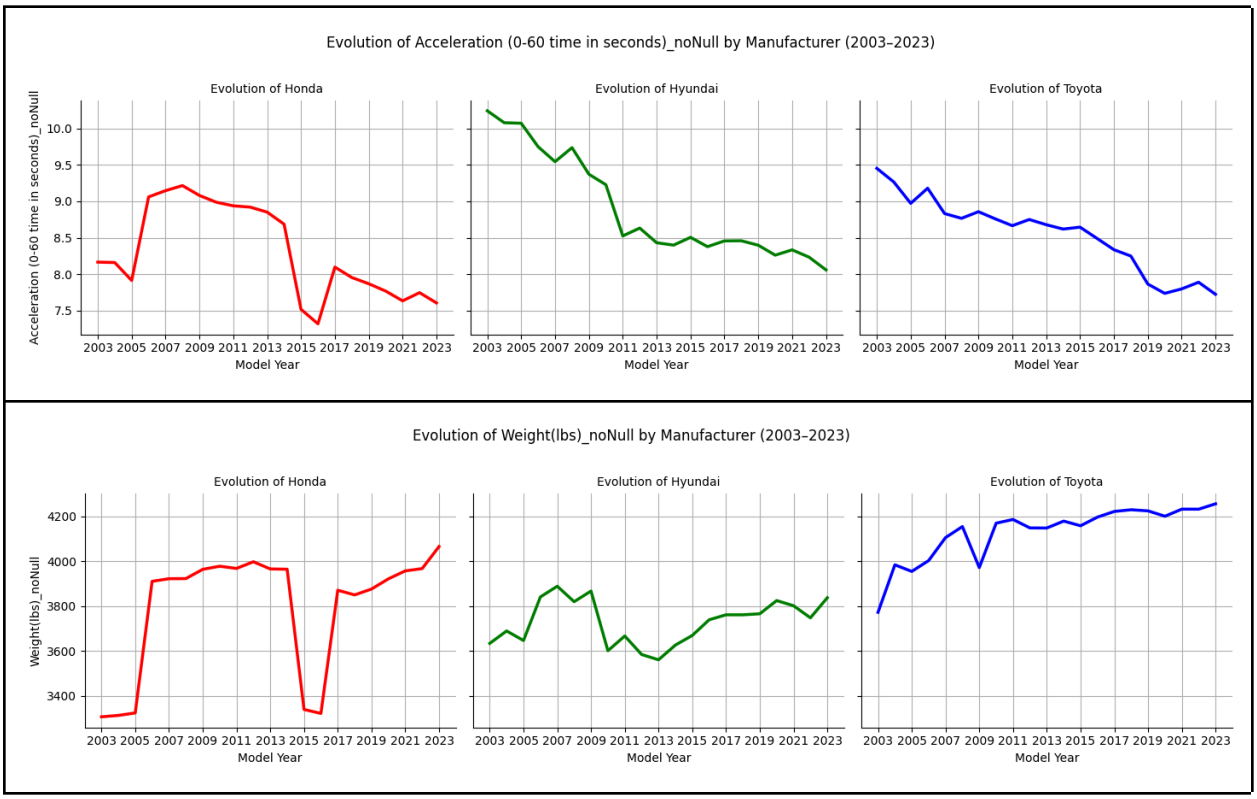
Inventado, Charles Fredric G.
Rodelas, John Vincent B.
Valles, James Vincent V.

4CSE - Tres Sigmas - LAB ACT - Time-Distribution

Time/distribution-based questions the Tres Sigmas group wanted to figure out.

- How different manufacturers benchmarks evolved over the years?
  (Time Chart focused)
- How are CO_2 emissions and fuel efficiency distributed across different vehicle types?
  (Distribution chart focused)
- How have CO_2 emissions per mile and fuel economy changed across model years?
  (Time Chart Focused)

CHART # 1 - Acceleration, Weight, and Horsepower (HP) of Honda, Hyundai, and Toyota from year 2003 to 2023



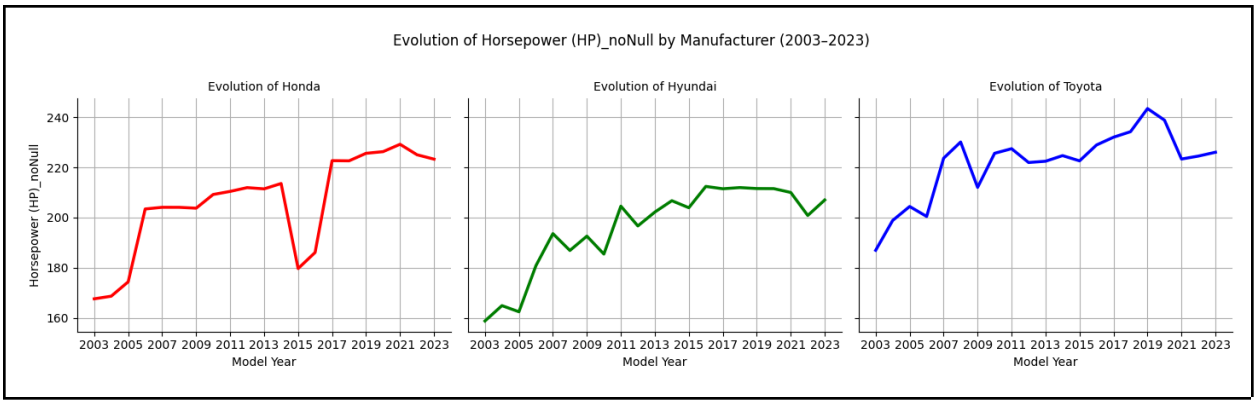


Figure 1.0 - Line Chart on Evolution of Acceleration, Weight, and Horsepower, respectfully of Honda, Hyundai, and Toyota from year 2003 to 2023

1. **Main takeaway in one sentence**

   Upward trend from 2003 to 2023 on Horsepower (HP), Downward trend on acceleration (means vehicles manufactured are getting faster on average), Weight has both ups and downs but trend continues to slowly rise up.

2. **One design decision and its benefit**

   Simple but elegant to provide necessary information on our situation, the trend line we can easily perceive going up after 20 years starting from the year 2003. Separating the three manufacturers instead of one single-combined line chart is also a wiser decision to have a clear view of the trend.

   Benefit of line charts here is the ease to monitor or track the behaviour of data points (specifically time-driven data) that makes communication clear and simple without creating confusion to the authors to the readers.

Additional Insights:

Limitations :

Since we see line charts frequently, the specific type of chart may not be the best option to choose if you aim to engage and gauge the interaction between you and the target audience.

Sometimes, you have to be specific on time points to deliver the message more clearly. Here is the example if we include from the 1980s to 2023;

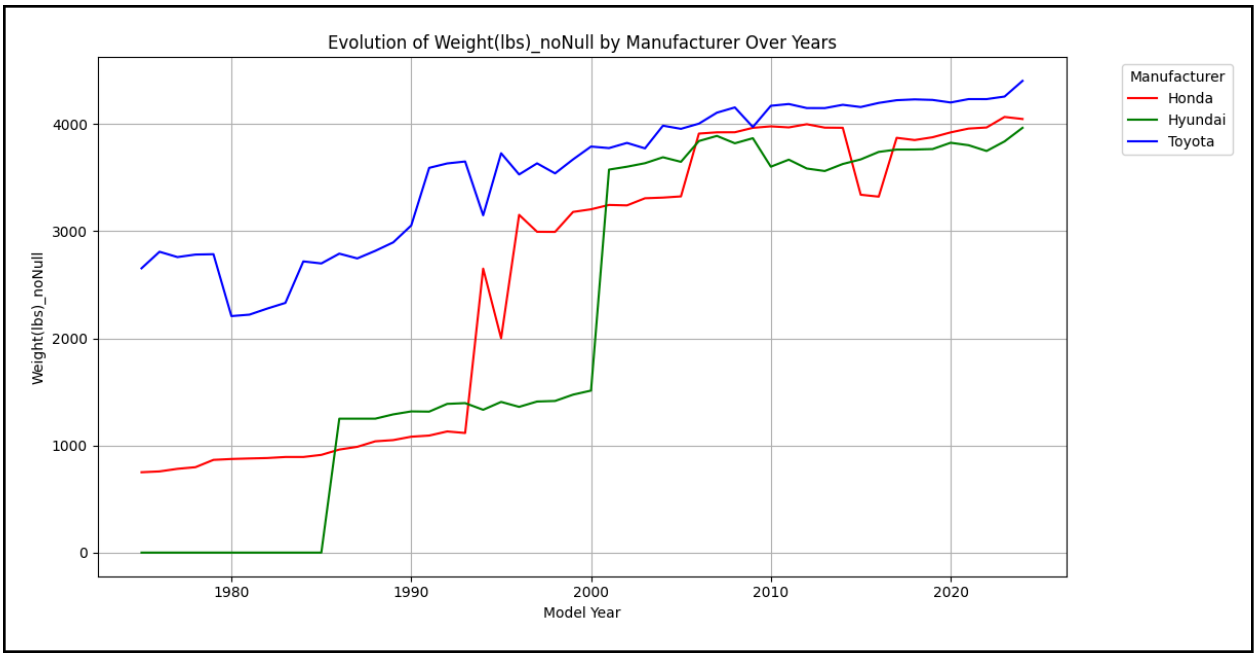


Figure 1.1 - Line Chart on Evolution of Weight of Honda, Hyundai, and Toyota from year 1975 to 2023

The timeline is honestly too long to give sensible interpretation since the 1975 to admit is too far away from the present already, so we decide to narrow it down to only the recent 20 years from 2003 to 2023, on which we still perceive the similar message of trend lines going upward.

We decide to limit from all 12 manufacturers down to 3 for simpler visualization as well, and repeatedly too dense of information is a visual clutter that will not gain the attention of any readers.

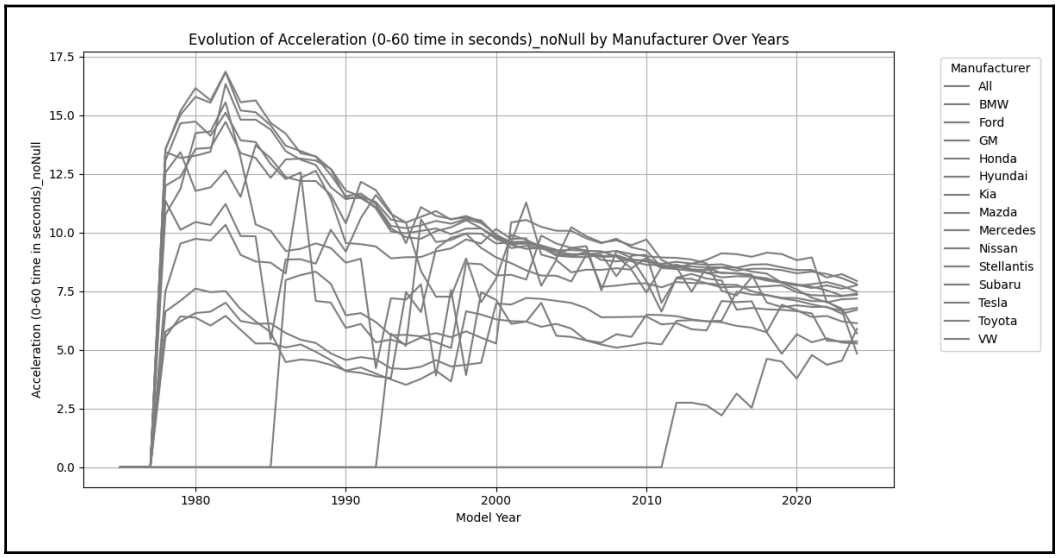
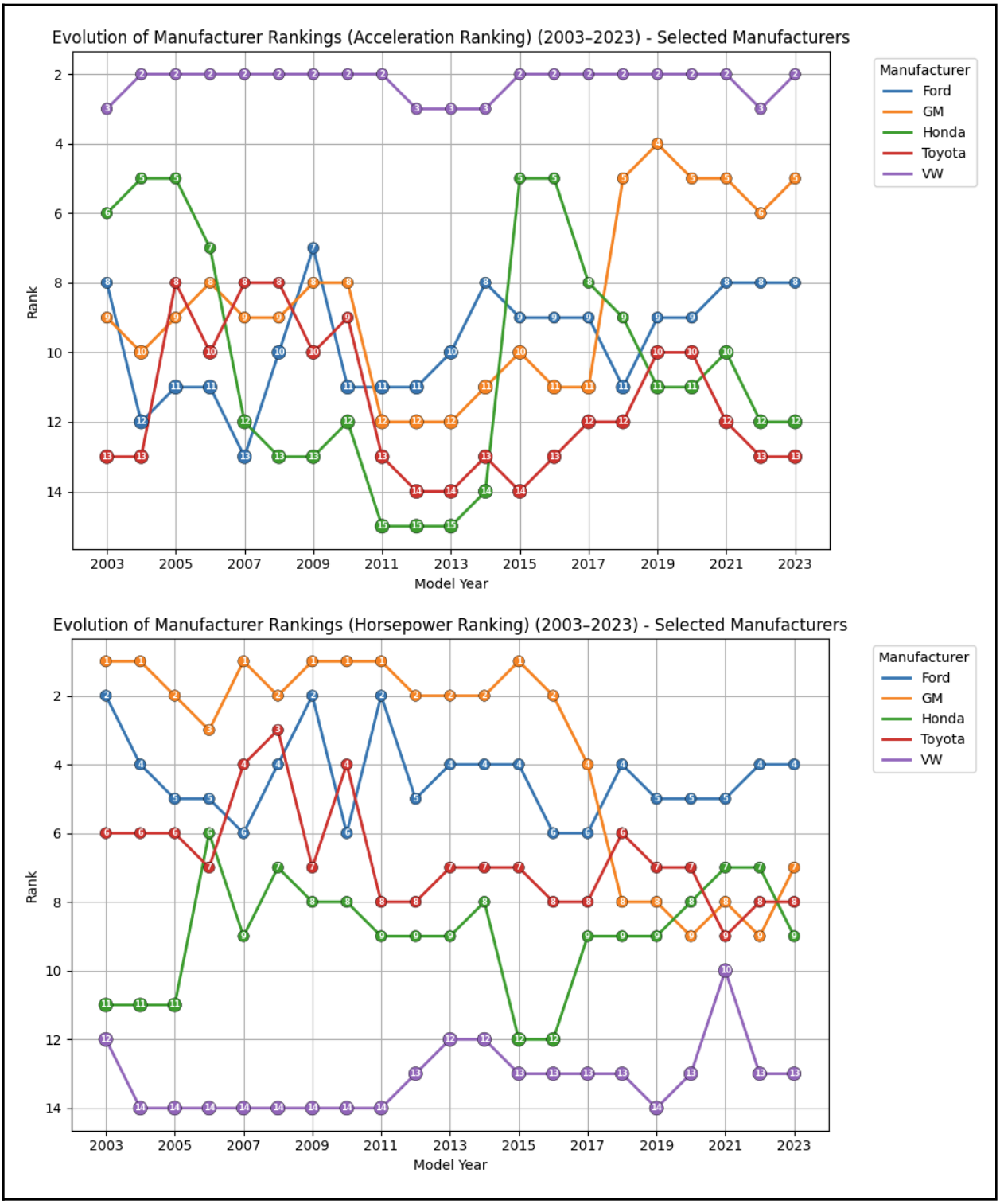


Figure 1.2 - Line Chart on Evolution of Weight of All Manufacturers from year 1975 to 2024

CHART # 2 - Rank of Acceleration, Weight, and Horsepower (HP) of Ford, GM, Honda, Toyota, and VW over the years from 2003 to 2023



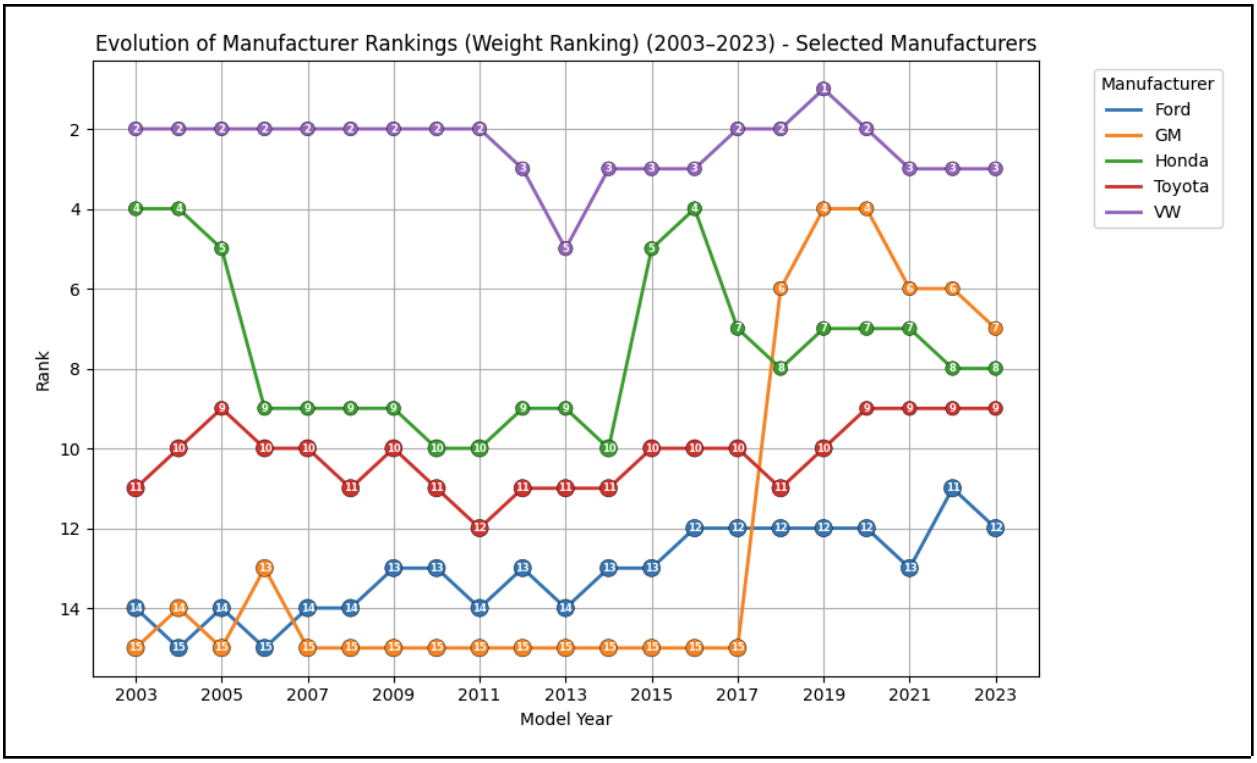


Figure 2.0 - Bump Chart of Manufacturer Ranking based on Acceleration, Horsepower, and Weight.

1. **Main takeaway in one sentence**
   For the selected manufacturers (Ford, GM, Honda, Toyota, and VW), the rankings in Acceleration, Horsepower, and Weight have fluctuated over the years from 2003 to 2023, with no single manufacturer consistently holding the top rank across all three metrics.

2. **One design decision and its benefit**
   A key design decision that made the chart easier to interpret was adding the rank number as text annotation within each circle on the plot. This allowed for quick identification of a manufacturer's exact rank in a given year without needing to refer to the y-axis grid lines.

Additional Insights:
Limitations : Since Bump Chart focuses on the relative ranks rather than the actual values, bump charts could not determine how dominant or different one manufacturer is compared to the one that is beneath or above their ranks. Additionally, there is difficulty in interpretation of these visuals including all manufacturers, too much visual clutter unless we removed some on the chart to focus on.

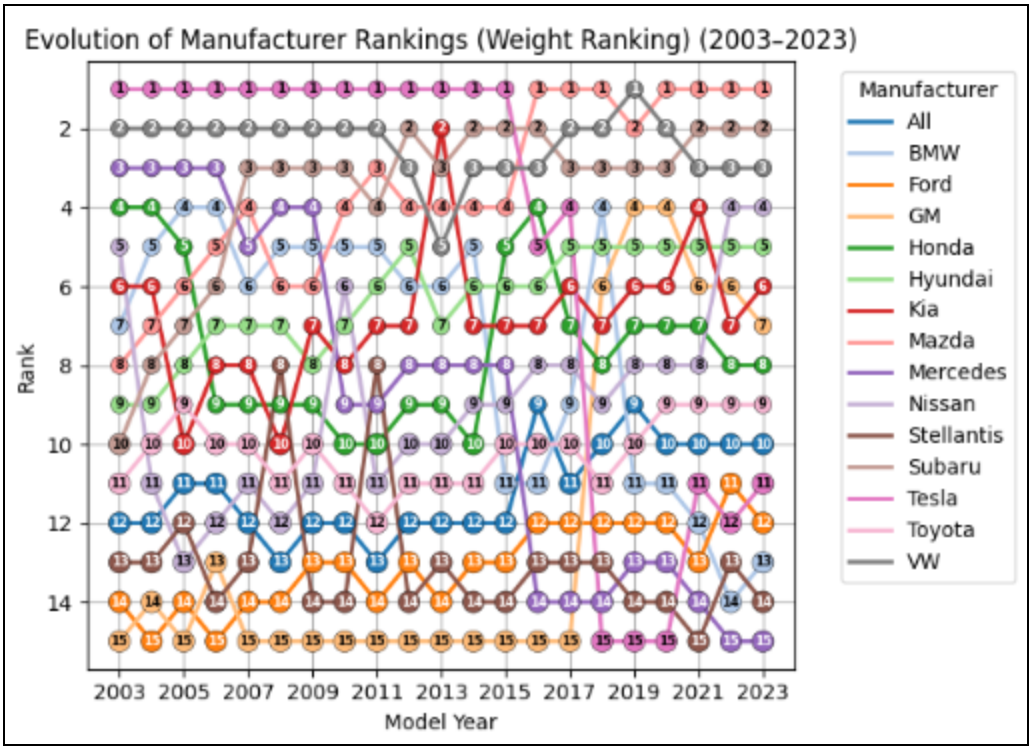


Figure 2.1 - Visual Clutter in a Bump Chart including All of the Manufacturers

Others : VW (Volkswagen) consistently maintained a high rank (low number) in Weight, suggesting their vehicles were generally lighter compared to the other selected manufacturers. Honda and Toyota showed more variability in their rankings across the three metrics compared to Ford and GM, indicating shifts in their performance benchmarks over time.

There are several instances of rank crossovers, particularly in Acceleration and Horsepower, highlighting the competitive nature of these performance metrics among manufacturers.

CHART # 3.0 - Beeswarm Plot of Distribution of Real-World CO_2 Emissions by Vehicle Type

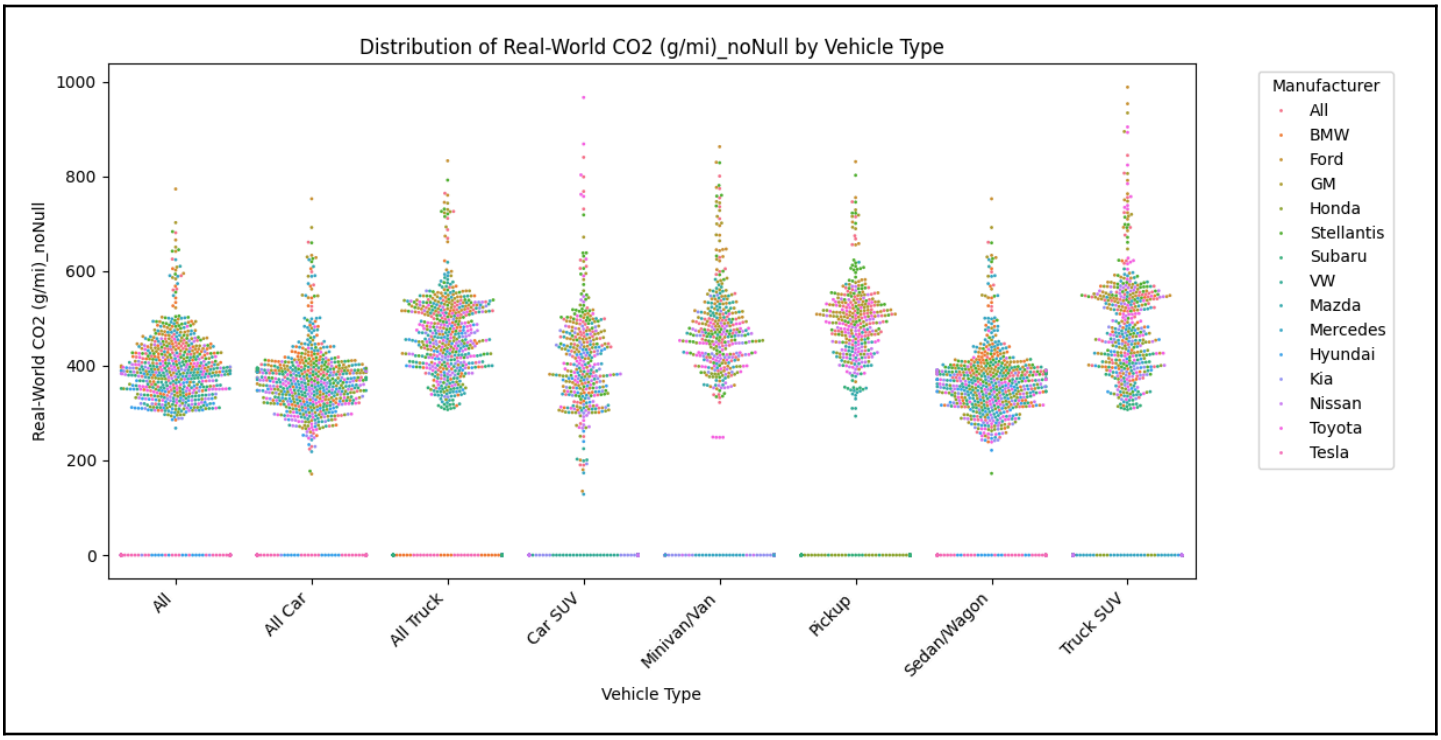


Figure 3.0 - Beeswarm Plot of Real-World CO_2 Emissions by Vehicle Type

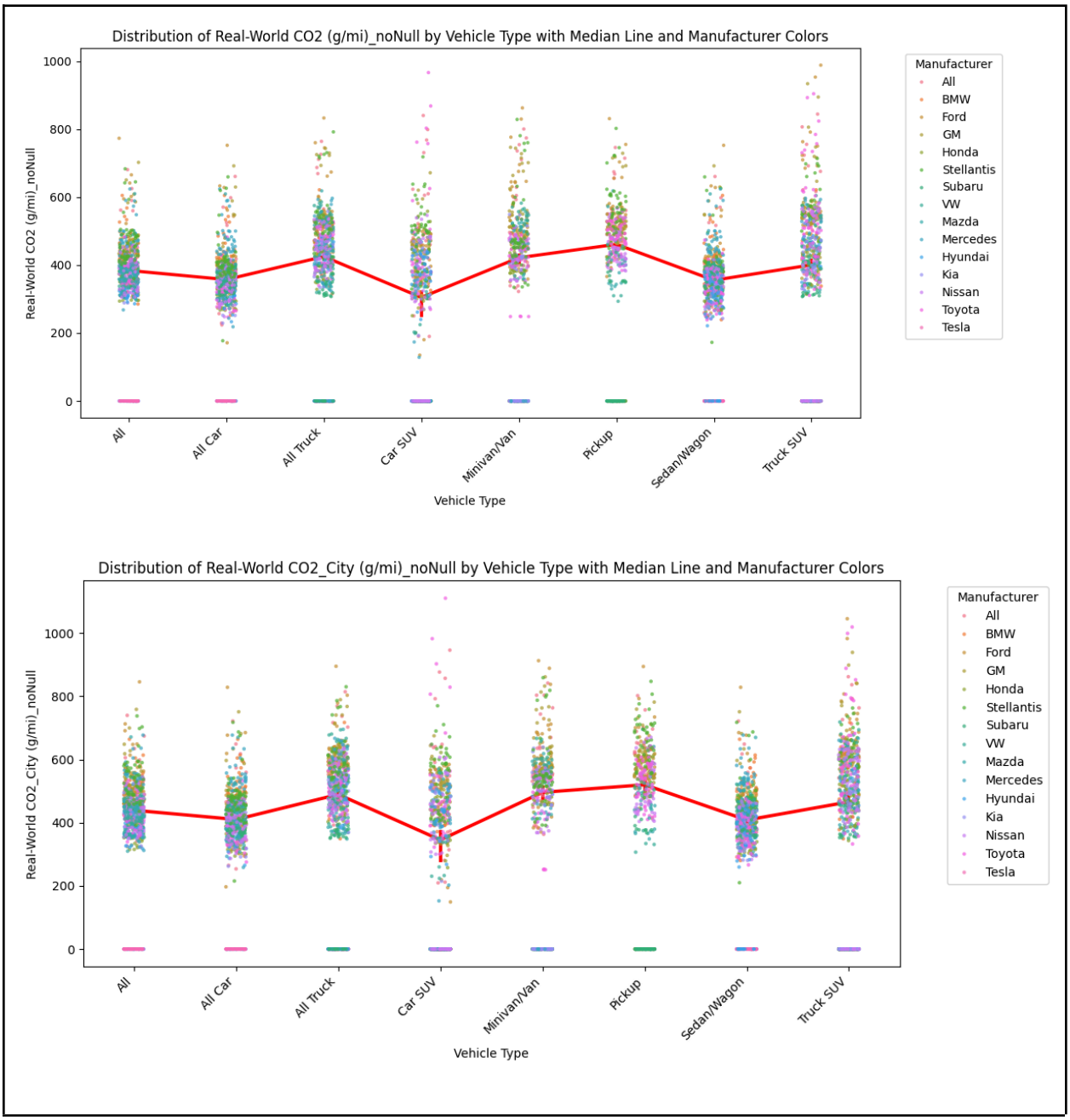
1. **Main takeaway in one sentence**

   The graph illustrates the distribution of real-world CO2 emissions by vehicle type, offering a comprehensive comparison across manufacturers such as BMW, Ford, and Tesla, and highlighting the varying environmental impacts associated with different vehicle categories.

2. **One design decision and its benefit**

   The design features a scatter plot that illustrates how CO2 emissions are distributed among different vehicle types. This visual representation method is beneficial because it helps us to easily compare emissions across various manufacturers and car models. It also effectively highlights trends and any outliers in emission levels. It enables a visual comparison that makes it easy to see which types of cars are greener and which aren't.

CHART # 3.1 - Strip Plot of Distribution of Real-World CO_2 Emissions by Vehicle Type



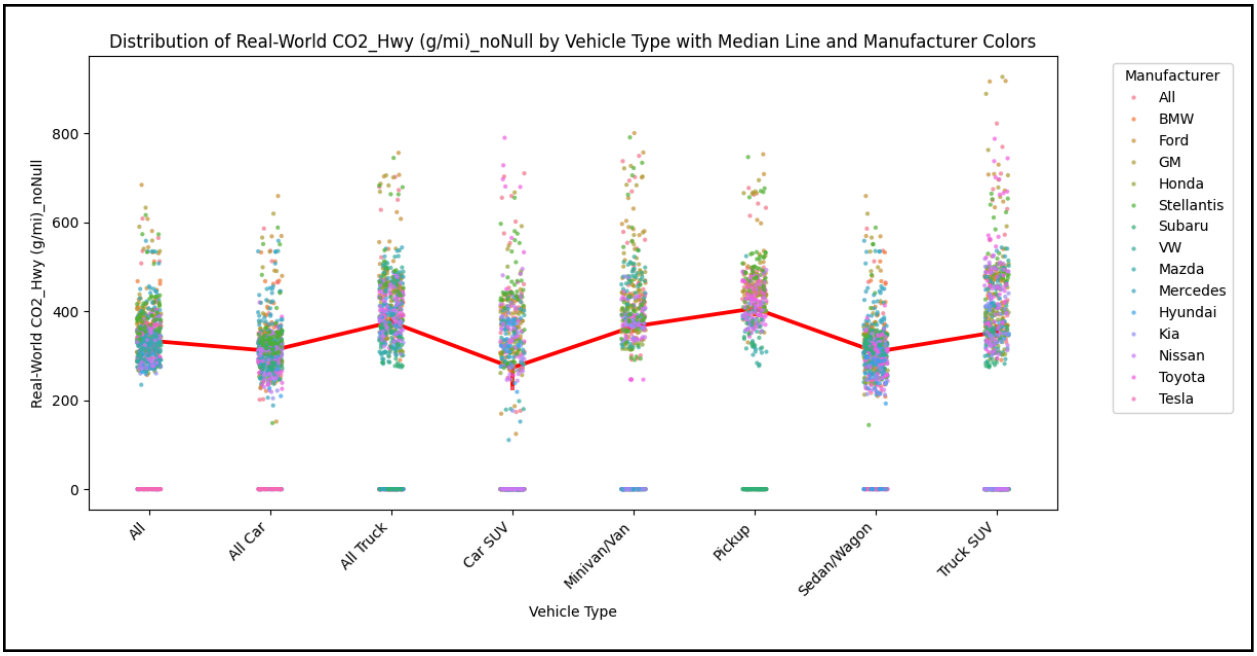


Figure 3.1 - Strip Plot of Real-World Distribution by Vehicle Type

The strip plots in the figure above take data points and align them along a single axis, which helps in reducing the overlapping of points. This makes it easier to see each individual point, unlike the beeswarm plot, which can get a bit crowded. Also, this plot is more structured and informative by adding a median line and uses different colors for each manufacturer. Each plot shows CO_2 emissions in broad conditions, city conditions, and highway conditions. This plot eliminates the error found in the beeswarm plot because it uses a strip plot visualization method that can handle overlapping data points more effectively. It also avoids the issues with points that can't be placed due to crowding, resulting in a cleaner and more readable plot.

Additional Insights:

Limitation : Not all points on the categories were all placed on the Beeswarm plot (not enough space to place these data points that occurred on all columns). Another specific limitation is that it doesn't show how many cars each manufacturer produces. So, a company with higher emissions might just be making more cars, not necessarily less eco-friendly ones. Strip plots can get cluttered when many data points overlap, which might obscure individual values and make it difficult to discern patterns, especially with large datasets. They don't inherently show the distribution shape (like being skewed or having multiple peaks), which can hide underlying trends. <See the image at the next page>

Proof of occurrence :

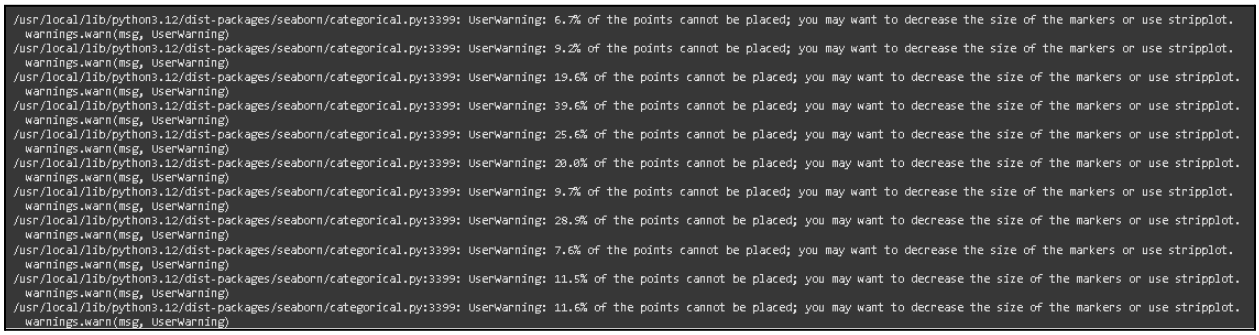
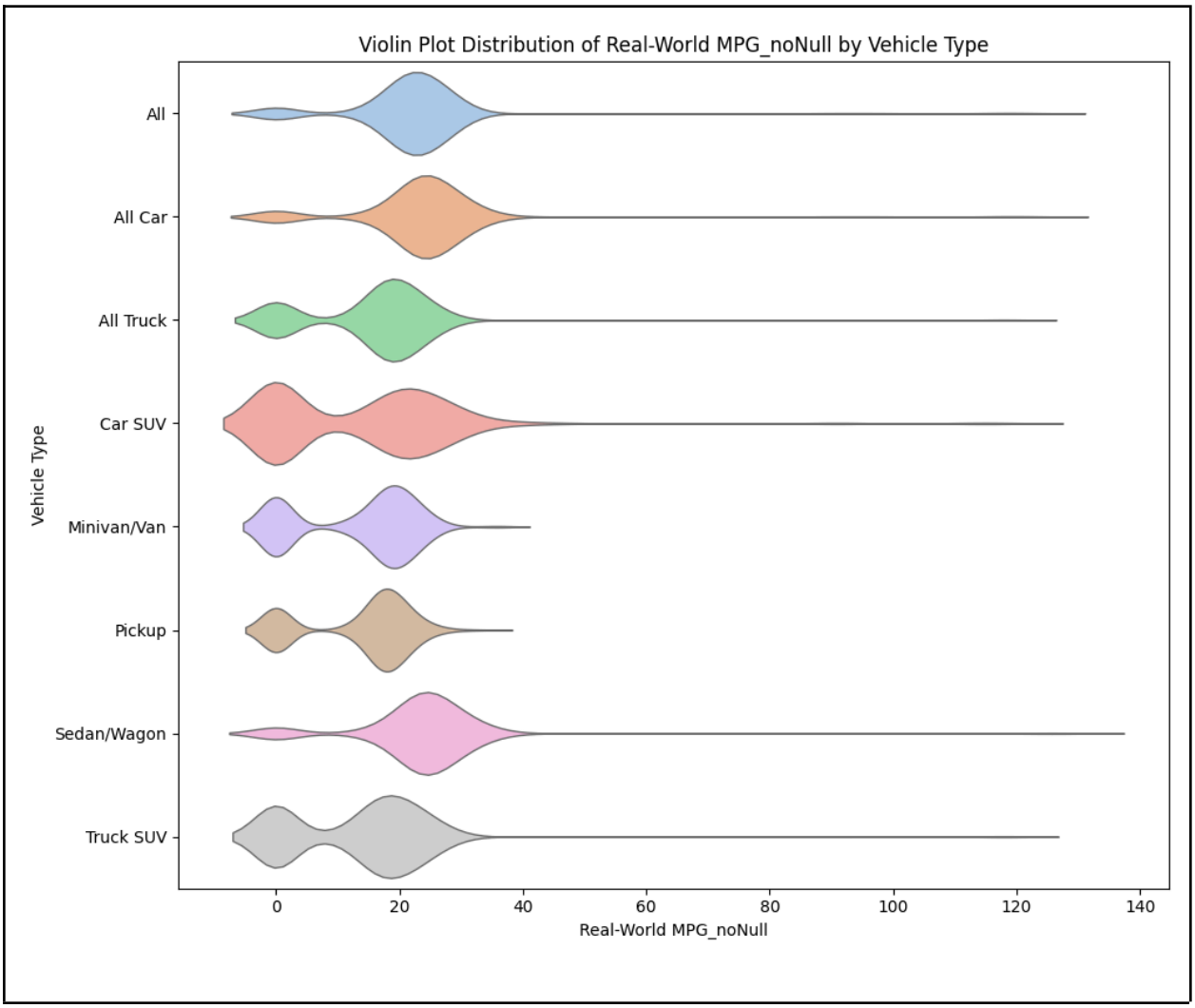
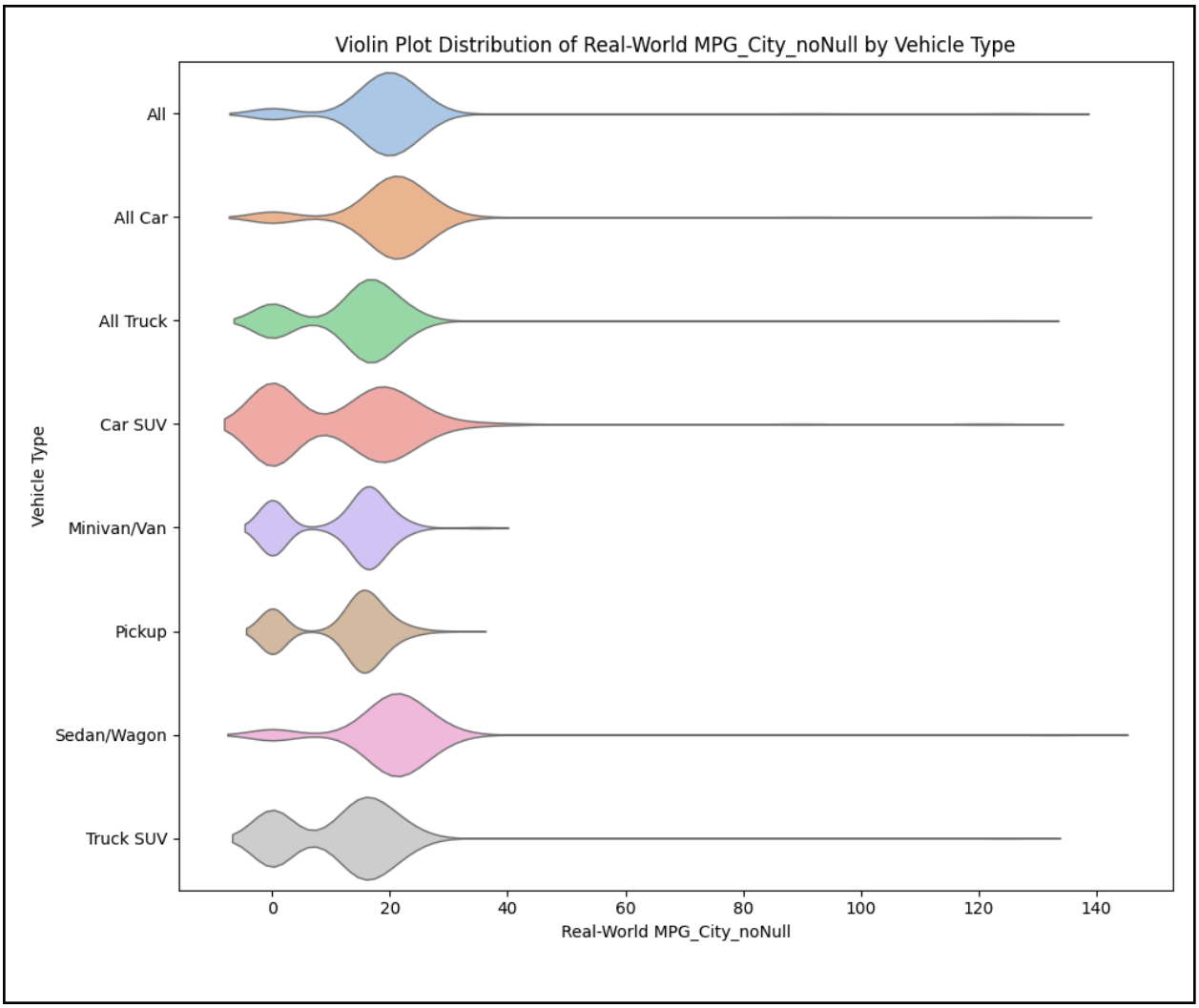


Figure 3.2 - Limitations of Beeswarm Plot on Many Records

As you can see in Figure 3.2 of a log, it indicates that a significant percentage of points in the beeswarm plot could not be placed, possibly leading to overlapping/overcrowding or unclear visualization. It suggests that we should reduce the marker size or use a different plot type like a strip plot, which will be expounded upon after this, for better clarity.

CHART # 4 - Violin Plot of Real-World Distribution of MPG by Vehicle Type




Violin Plot Distribution of Real-World MPG_City_noNull by Vehicle Type
Vehicle Type
All
All Car
All Truck
Car SUV
Minivan/Van
Pickup
Sedan/Wagon
Truck SUV
0
20
40
60
80
100
120
140
Real-World MPG_City_noNull

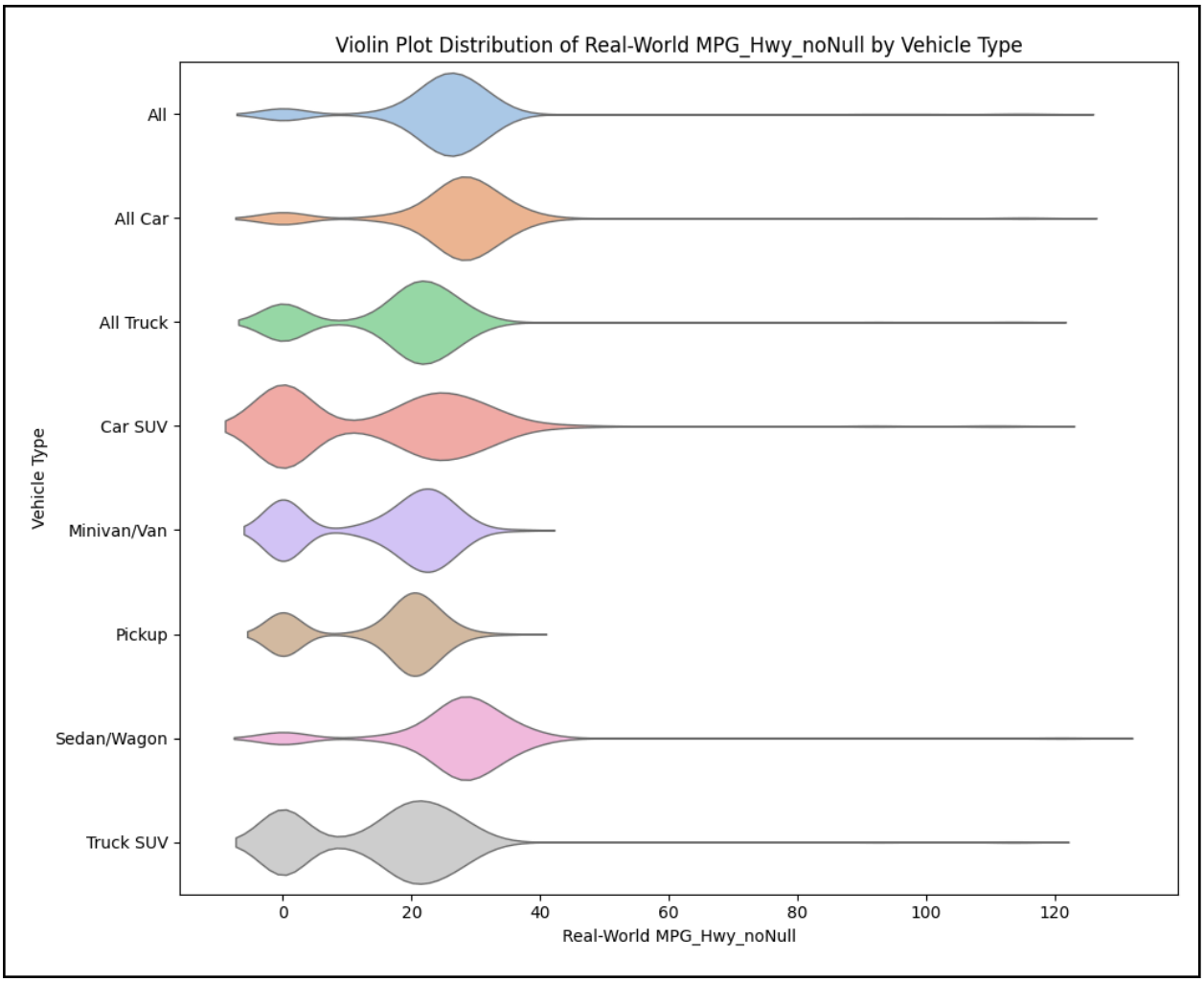


Figure 4.0 - Violin Plot of Real-World Distribution of MPG by Vehicle Type

1. **Main takeaway in one sentence**

   The violin plot in Figure 4.0 depicts that real-world MPG varies by vehicle type, with cars generally exhibiting higher fuel efficiency, while trucks and SUVs tend to have lower MPG, both in city and highway driving conditions.

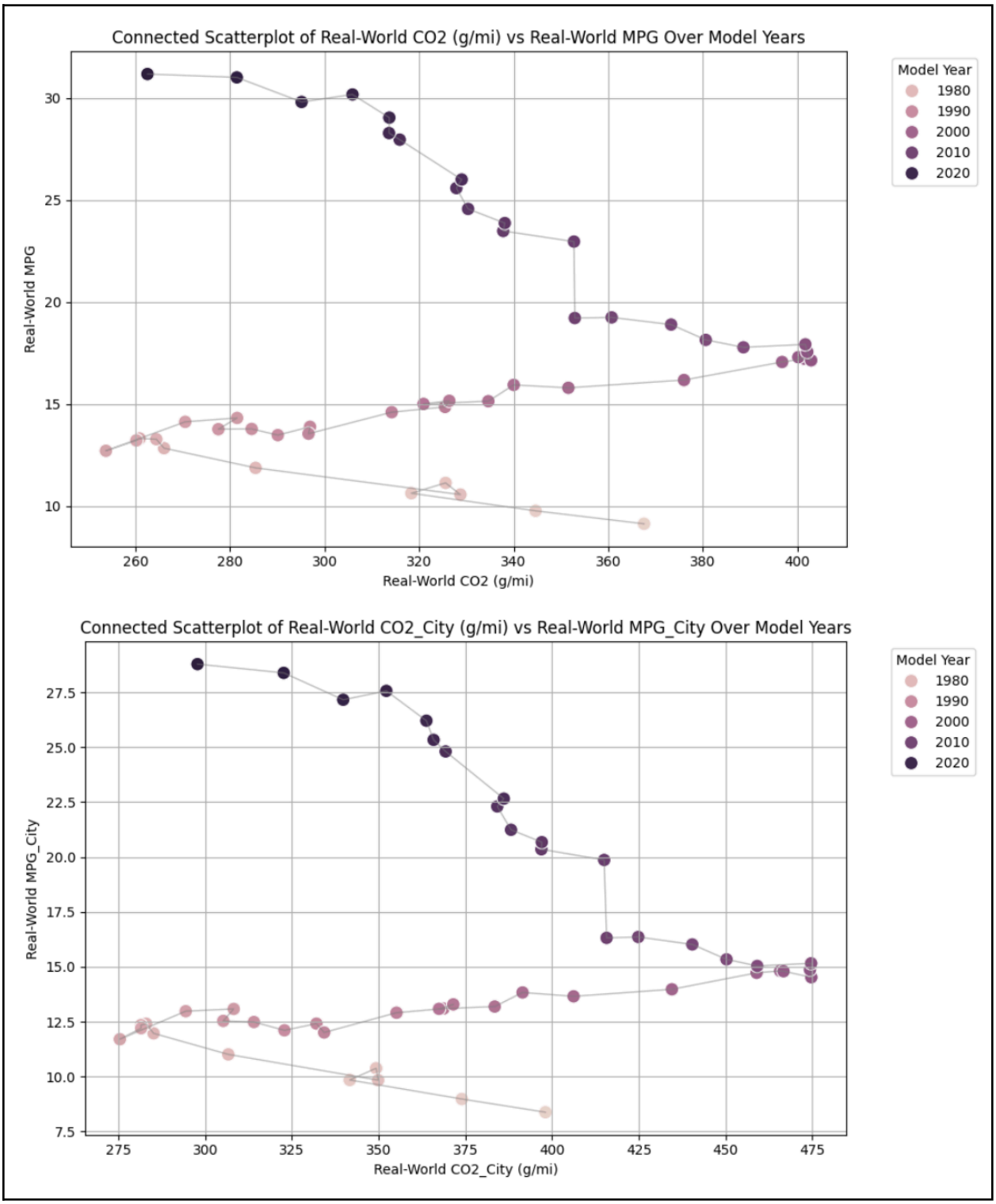
2. **One design decision and its benefit**

   Visually speaking, using different colors for each vehicle type makes it super easy to differentiate between categories at a glance, making data more accessible and the plot more visually engaging. The consistency in the scale of the plots also allows for easy comparison among the three plots.

Additional Insights:
Limitation: A limitation in these violin plots could be the overlapping of different vehicle types. In the violin plots shown above, however, there is little to no overlapping.

Others: In instances where it occurs, overlapping can make it quite difficult to distinguish between them when the distributions are similar. This could lead to some challenges in accurately interpreting the data. The overlapping in this plot involves the distributions blending together if they are close in value.

CHART # 5 - MPG to CO_2 Emission gets better



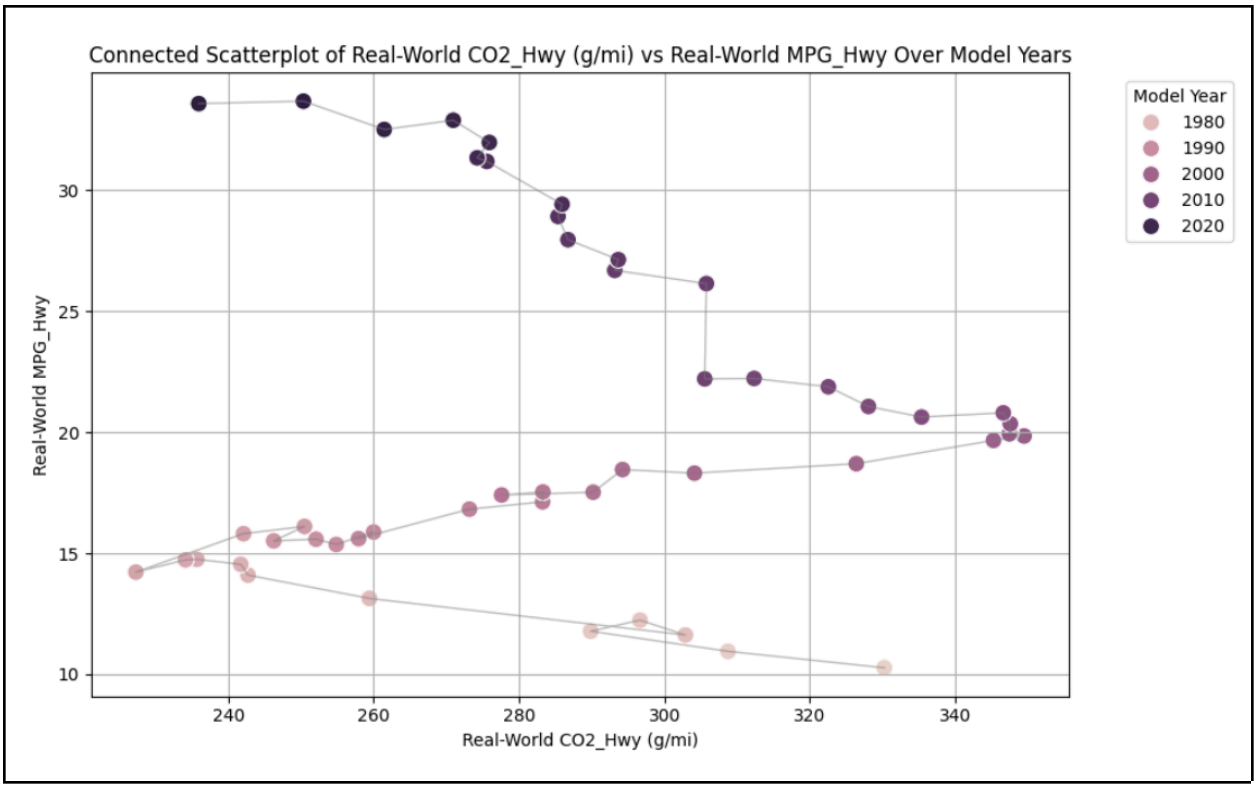


Figure 5.0 - Connected Scatterplot between MPG (Miles Per Gallon) and CO_2 Emissions from 1980 to 2020

1. **Main takeaway in one sentence**
   Over time the Miles Per Gallon (MPG) for all, City, and Highway (Hwy) increases while the Carbon Dioxide emissions went back and forth with a jaggedy decrease between 2000 to 2020.

2. **One design decision and its benefit**
   The use of the connected line chart easily helps interpret the changes of MPG and CO2 overtime. Additionally, Data points were colored to an associated year in order to better see when changes in the relationship occurred throughout the line.

Additional Insights:
Limitations : This chart shows the average for all cars in a given year. It doesn't tell you about the spread or variation within that year, like if there were some super efficient cars and some really bad ones. It's just the middle ground. Also, it doesn't show why these changes happened, like if it was new tech or just different types of cars being popular.

Others : You can see the general trend is up and to the left, meaning lower CO2 and higher MPG. The plots for city and highway driving show slightly different patterns, which makes sense because cars perform differently in those conditions.