Prakahr Saxena

Professor Shetty

INFO 250 – 001

October 2, 2020

Introduction:

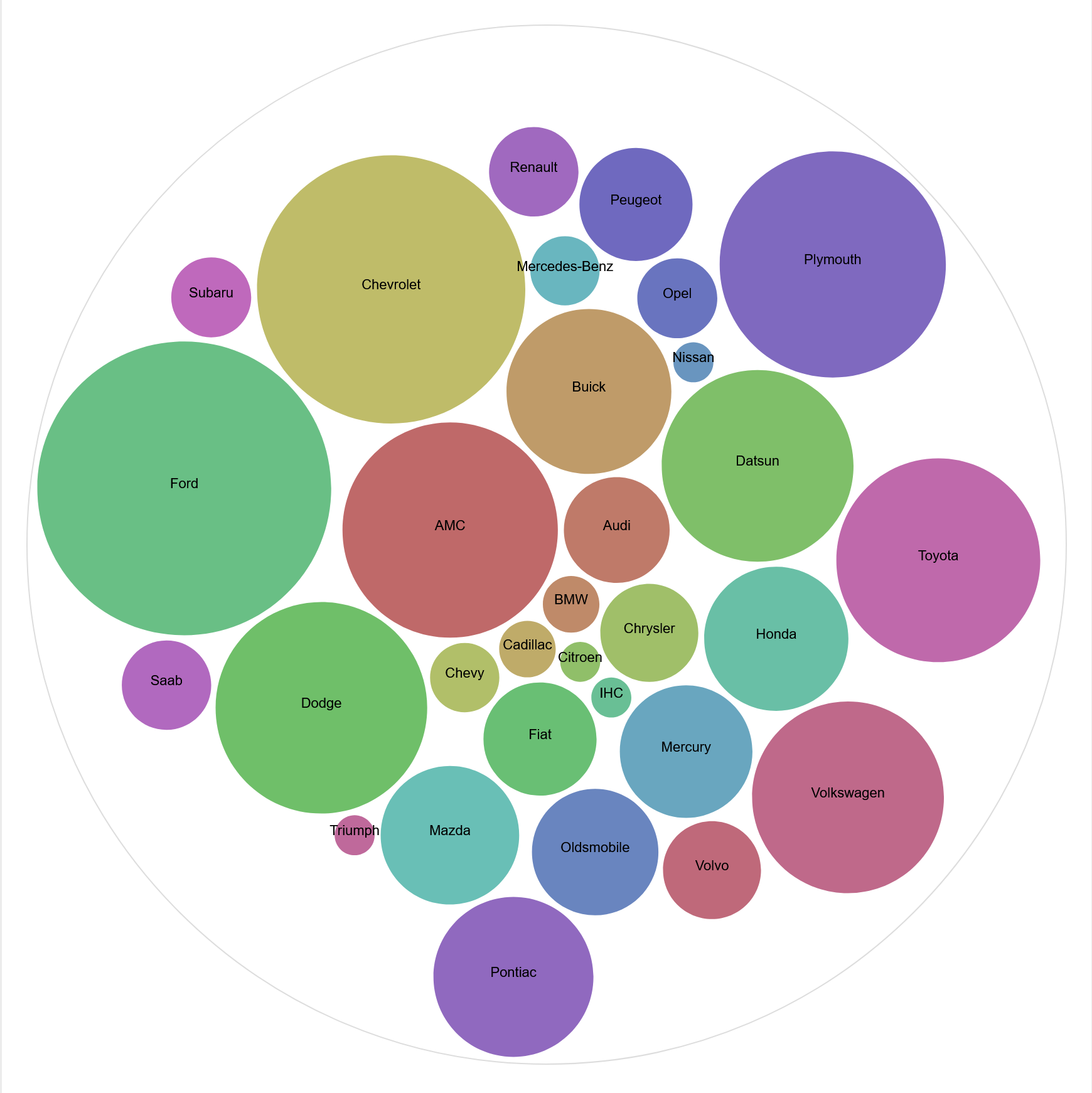
For this assignment, I chose the Cars dataset, because it was the most fascinating topic to me out of all the sample datasets. This dataset has \_ attributes:

name: String, brand: String, economy (mpg): Number, cylinders: Number, displacement (cc): Number, power (hp): Number, weight (lb): Number, 0-60 mph (s): Number, year: Number

The four visuals that I picked are Circle Packing, Beeswarm Plot, Scatter Plot

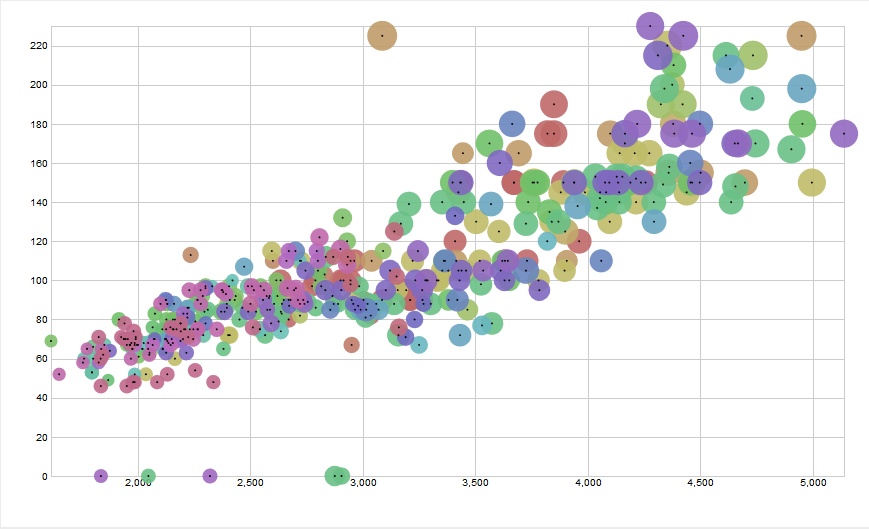
RAW has a straightforward UI, I don’t have much trouble figuring out the essentials. I did experiment with different attributes

Circle Packing



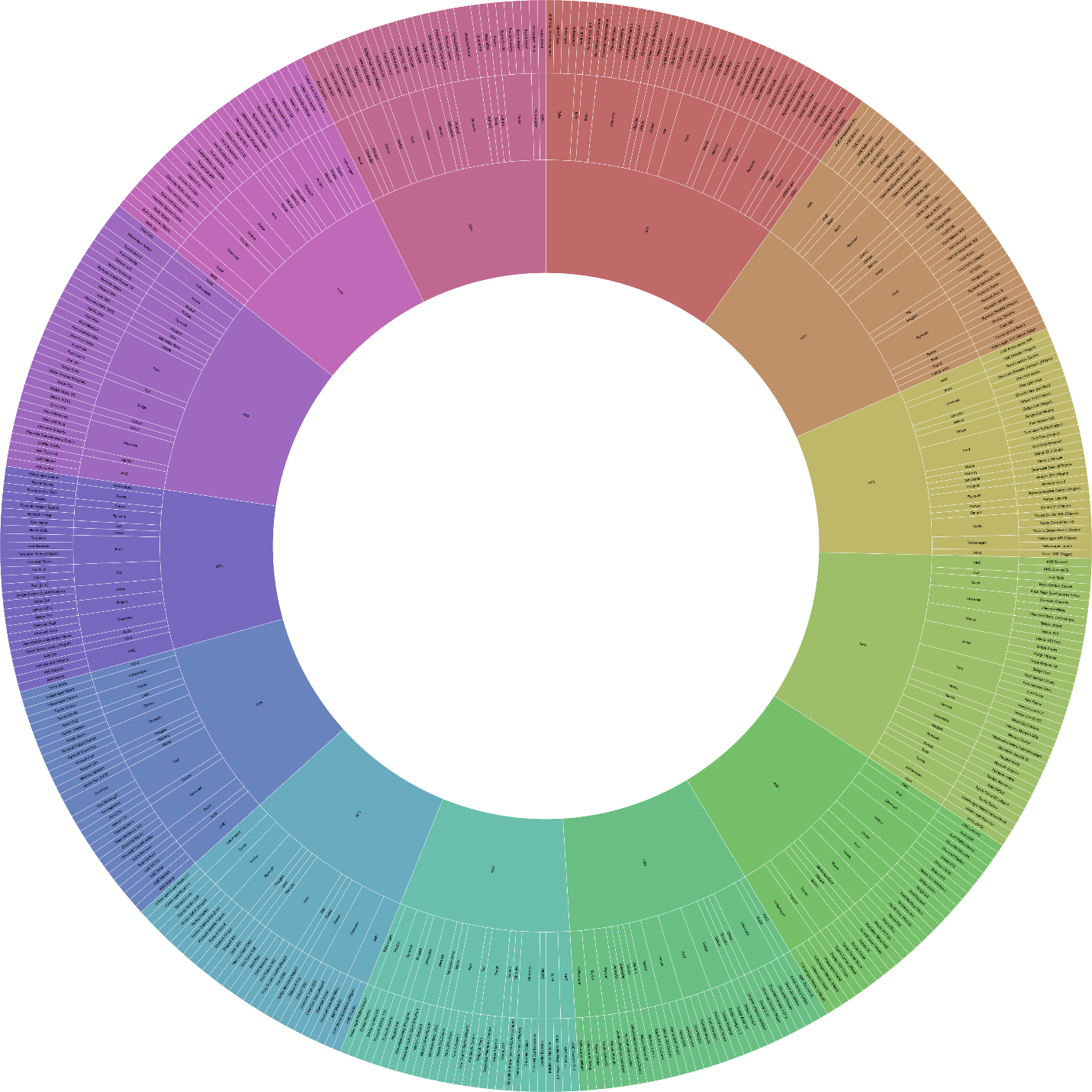
Beeswarm Plot



Scatter Plot

Scatter plot is a very basic graph, which depicts any polynomial relation between the X and the Y component. Here, I used Weight (lb) for the X Axis, and power (hp) for the Y Axis. Although that already gives us a fairly decent understanding into the relationship between the two dimensions, I thought it to be useful to add in the element of displacement (cc) of the engine as the Size component to all the scatter points. Additionally, I added the Brand to the Color component of the graph to improve visuals in understanding the varied range of cars made by different brands. I can infer from this graph, that generally, the power, the weight, and the displacement of the car have a manner of direct correlation with each other.

Sunburst



This visualisation, although not entirely appropriate for such a massive dataset, encompasses 4 patterns, Hierarchy, Size, Colour, and Label. With this visualisation, I wanted to explore the difference in the number of cars made by different brands by years. Therefore, I mapped year, brand and name to the Heirarchy pattern. And differentiated them based on year