

We're team Statistical Sleuths – Mao, Shelley, Zenia, and Raylin

Metrics for Edmunds

First, in order to determine which markets Edmunds should target, we estimated the “market share” of Edmunds by calculating the (# leads/ population) by designated marketing area (DMA). The population was obtained from summing populations in Zip Codes from U.S. Census data. Assuming that market size is proportional to population and that the largest market share ratio represents a market share of 100%, we then normalized by dividing by this largest market share ratio.

We then defined a metric called “market potential” = $(1 - \text{market share}) * \text{market size}$, where $(1 - \text{market share})$ is the part of the market not captured by Edmunds and market size is estimated from the total wages and salaries from U.S. Census data in the DMA. The resulting “market potential” is a measure of the potential value of a DMA not captured by Edmund's.

Helping Customers find Cars

Car shopping can be an overwhelming experience for consumers when they are faced with too many choice. Our objective was to help consumers narrow down their choices by using data analytics of their surrounding population to infer the most popular choices. In addition, this consumer insight will allow Edmunds to improve localized advertising to consumers in the tregon or through the website ads.

We looked at consumer preference for model, make, and style in three of the biggest markets for Edmunds.com. By region, we identified the breakdown of what makes, models, and styles of cars received the highest frequency of inquiry, and compared the demand across the three markets.

We found that the biggest differences lie in the style of the car, which include SUV, Sedan, Hatchback, hybrid, electric, and so on. The demand in SUV in NY is a lot greater than the demand for SUV in LA, while the demand for hybrid and electric is much greater in LA. The demand for trucks is slightly greater in Atlanta than LA and NY. Geographical may be correlated with demand because of weather or becaue of a factor we haven't yet thought of, so we then explore the variables that drive the difference in demand by incorporating demographic data through the 2010 Census by regressing the style of car on submitted leads on characteristics such as number of households, proportion of racial groups, median income, proportion of veterans, and other variables. Each observation is a visitor's submitted lead/car type. This can be done for every state, demonstrate the regression for Georgia.

Finding Associations between Preference and Demographics

We identified the characteristics of the visitor that tend to be associated with the types of car (SUV, hybrid, electric, etc) that they submit leads for. We want to characterize the consumer preference for car types through things like geographical location, median income of the area, population, racial composition, and age.

Finally, we suggest improving the current advertising algorithm at Edmunds. We were informed that Edmund's revenues mainly come from two areas: