

2SigmaSquared

Methodology

We did an initial analysis of the transactions and determined that about 60% of sales came from 20% of dealers, and the other 40% came from the 80% of the dealers. Charting a heat map of transactions, we observed a higher proportion of intra-state transactions compared to transactions between states. In an attempt to make sense of this phenomenon, we hypothesized that the distance between buyers and sellers inversely affected sales, while number of potential buyers and sellers directly affect sales. Specifically, we hypothesized that the transaction flow followed the gravity model of trade. We pulled out the zipcodes from the transaction data frame and the zipcodes from the visitors' dataframe and separated them according to the states of origin, and obtained the average distance between the buyers and sellers between two states as a variable. The number of potential buyers is the number of viewers on the website from each state, and number of sellers is the number of registered dealer on the website.

Insights

Calculating the coefficient of least square regression, we obtained a value of 0.78 for the gravity model, 0.39 for an exponential model. Ignoring insignificant transactions between states averaging under 5, we determined that transactions rarely occur when the distance between buyers and sellers exceed 200 miles. Additionally, the number of transactions is more deeply impacted when distance varies, as compared to mass varying. Potential suggestions to take into consideration for sale strategies could involve allowing an option for sellers in distant regions to pay a premium for their advertisements to be featured in all searches.