# Initial Analysis

As part of our initial analysis, we combed through the data to check for problematic data such as variables that are marked outside their available range, outliers, and other potential data issues. Thereafter, we did some simple analysis to get a better understanding of the data and discover any apparent patterns through some simple exploration and visualization.

To begin, we did a simple analysis of customer demographic patterns based on home ownership (H=0 for non-home owner or H=1 for home owner) and marriage (M=0 for non-married or M=1 for married):

	H0M0	H0M1	H1M0	H1M1
Total Spent	$175,151,310	$24,265,938	$160,012,518	$63,525,575
% $ Spent	41.4%	5.7%	37.8%	15.0%
Population Count	270,763	37,760	254,929	101,797
% Population	40.7%	5.7%	38.3%	15.3%
Mean	$647	$643	$628	$624
Median	$647	$642	$626	$622
Mode	$642	$639	$626	$607
Std Dev	45.92	44.93	43.92	44.58
Max	$922	$846	$912	$911
Min	$260	$356	$264	$280

There were few obvious findings based on this comparison alone due to the many similarities between the groups. One initial finding is that the groups are very similar except for the number of members in each group. Most significantly, M0 contains 79.2% of the sample population. We then made a visual to further explore the relationship of demographic and purchase decision:
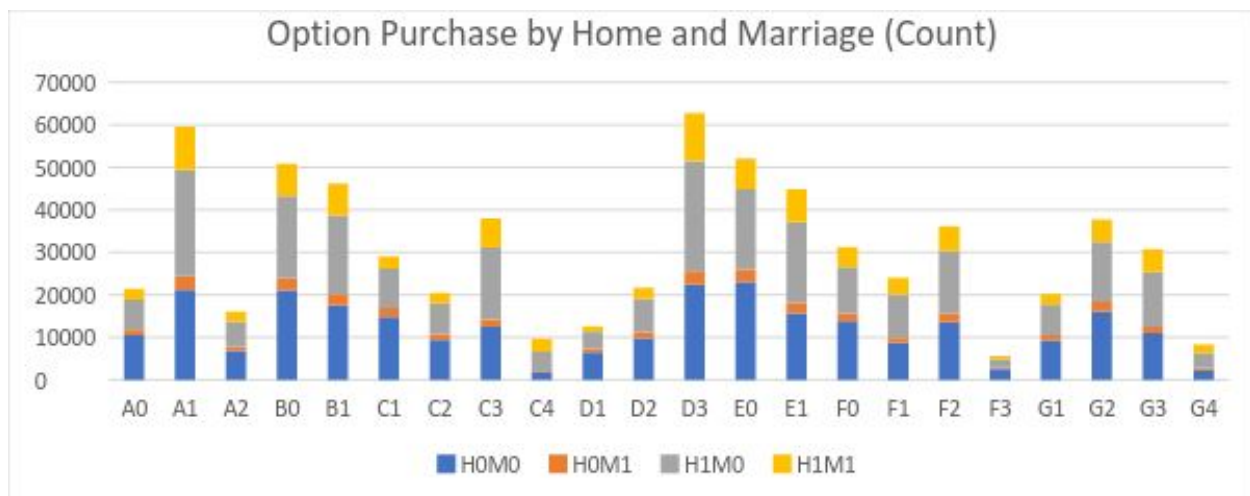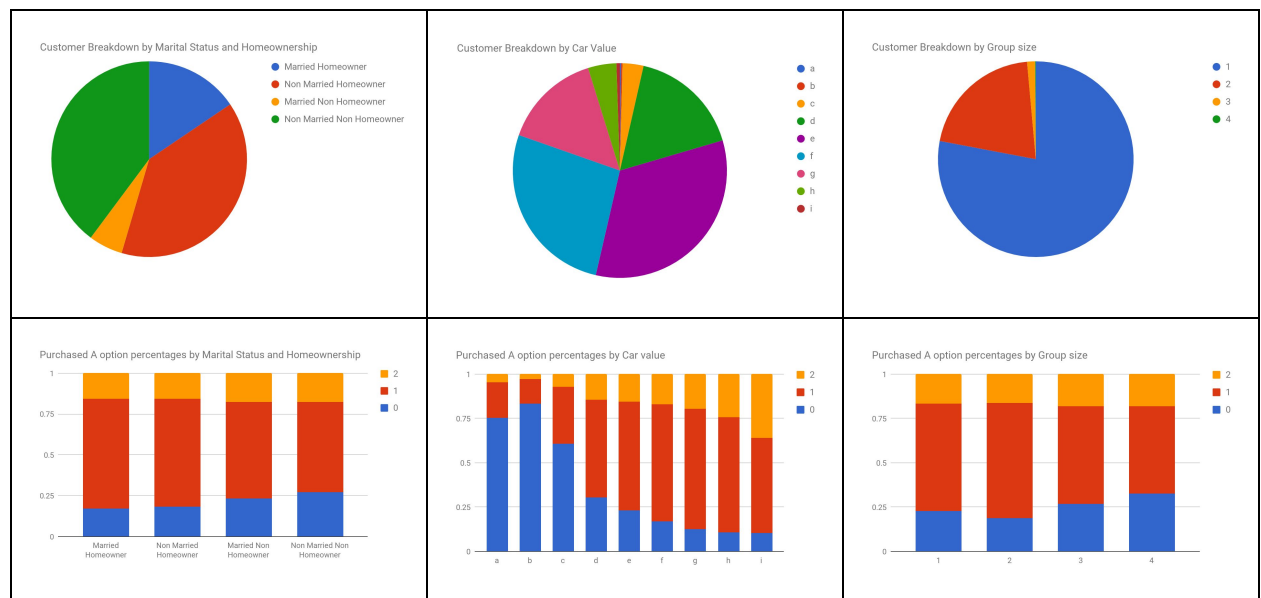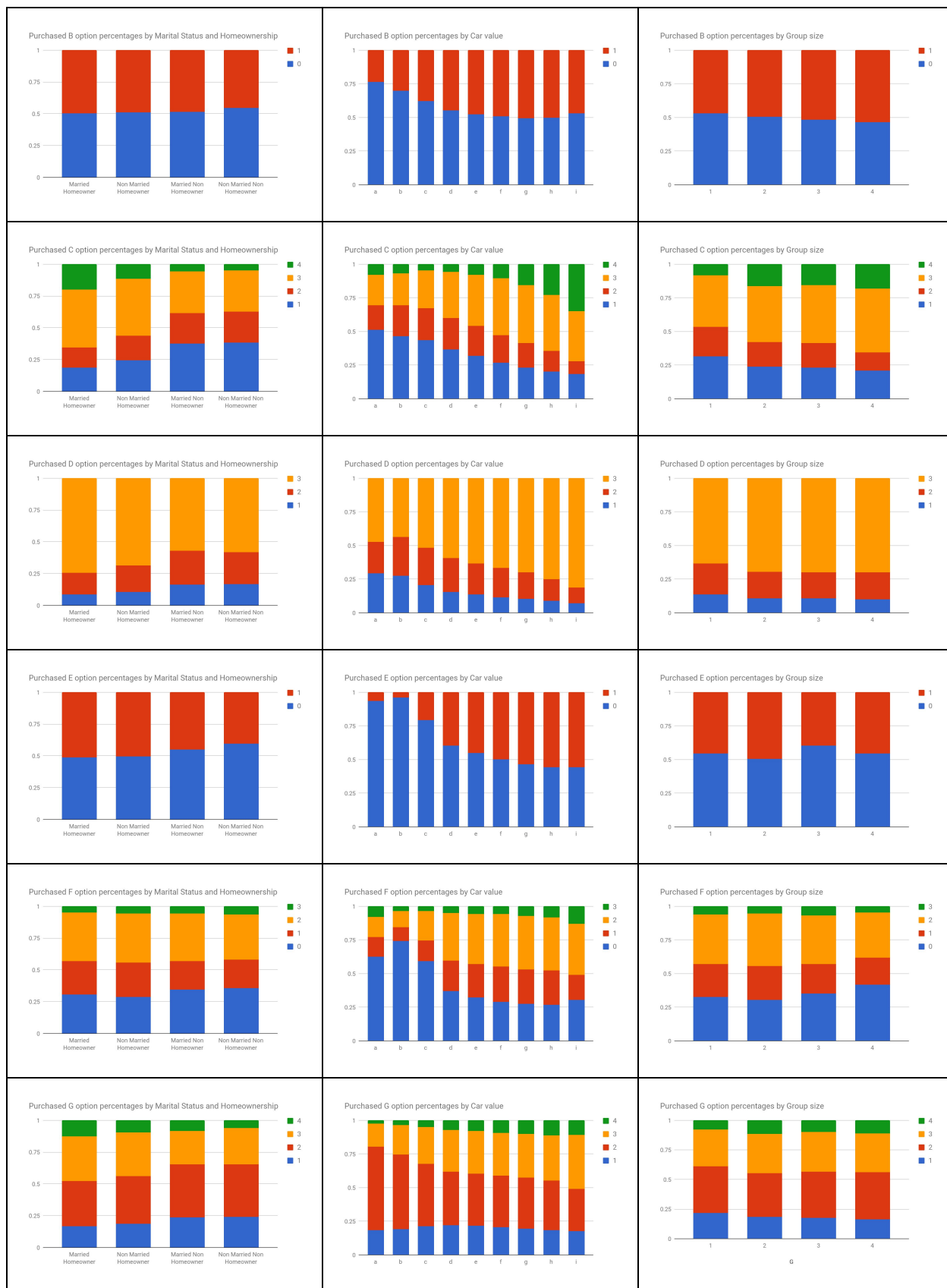

Due to the four demographics having differing populations, a visual representing choice by percentage was also made:

Option Purchase by Home and Marriage (%)
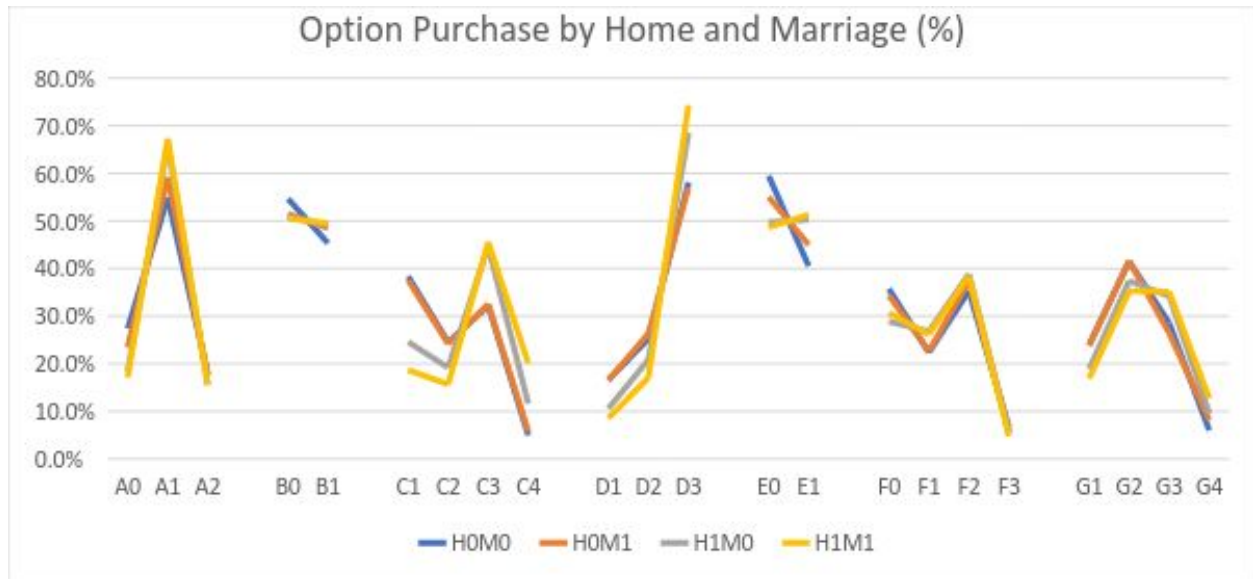
This visual provided some revealing data about which demographics tend to follow one another. For example, both H0 groups acted very similarly in all categories except B and E as well as both H1 groups are more likely to purchase a D3 than the H0 groups.

In an effort to get a clearer perspective on available patterns of data before initiating deeper exploration, we further explored customer demographic patterns to purchase behaviors as well as explored the patterns and relationships of purchase behavior based on car value and group size.

Purchased B option percentages by Marital Status and Homeownership

Purchased B option percentages by Car value

Purchased B option percentages by Group size

Purchased C option percentages by Marital Status and Homeownership

Purchased C option percentages by Car value

Purchased C option percentages by Group size

Purchased D option percentages by Marital Status and Homeownership

Purchased D option percentages by Car value

Purchased D option percentages by Group size

Purchased E option percentages by Marital Status and Homeownership

Purchased E option percentages by Car value

Purchased E option percentages by Group size

Purchased F option percentages by Marital Status and Homeownership

Purchased F option percentages by Car value

Purchased F option percentages by Group size

Purchased G option percentages by Marital Status and Homeownership

Purchased G option percentages by Car value

Purchased G option percentages by Group size

Through these analyses, we gain a high-level perspective on basic patterns (or lack thereof) of purchase behavior based on the demographics of home ownership, marriage, car value, and group size. The next step of our exploration is to do a regression analysis to further discover statistically significant patterns of purchase behavior based on all available data.