VJ Davey 3/3/17

**Executive Summary**

The insurance dataset was analyzed with the motive of finding interesting tidbits of information that may be useful for making business decisions as an insurance company. The data was explored using a combination of Hue/Impala and Tableau.

Major Insights

* Students and blue collar workers are more likely on average to crash (*approx. 37% and 34% crash probability)* and have below average income.
* People with high school education or less are more likely to crash (*approx. 32-34% crash probability*) and have below average income.
* People with PhD education are less likely to crash *(approx. 17% crash probability*) and have above average income.
* People who fall into the 20-29 age group have particularly high crash rates (*approx. 49% crash probability*), this is especially pronounced when looking at the 20-24 age subgroup.
* People in the age groups of 40-49 or 50-59 have particularly lower crash rates (*approx. 23% and 21% crash probabilities respectively*).

Interesting Insights

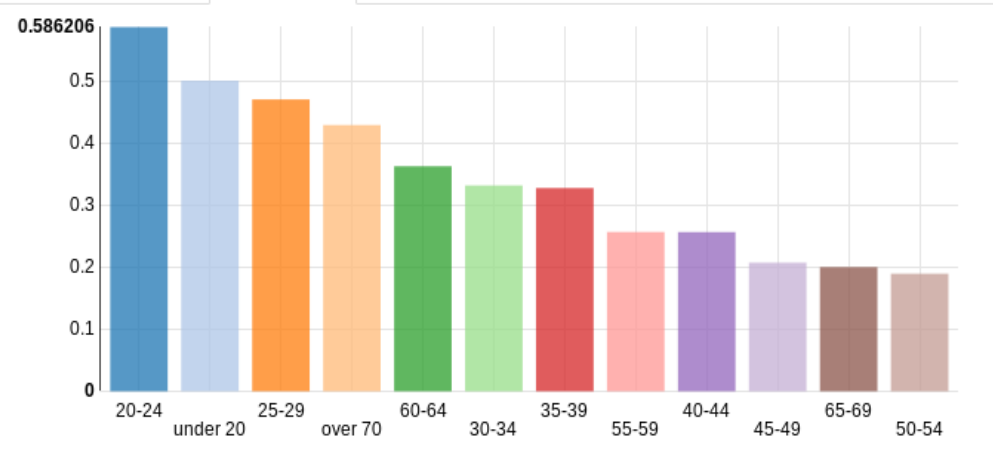
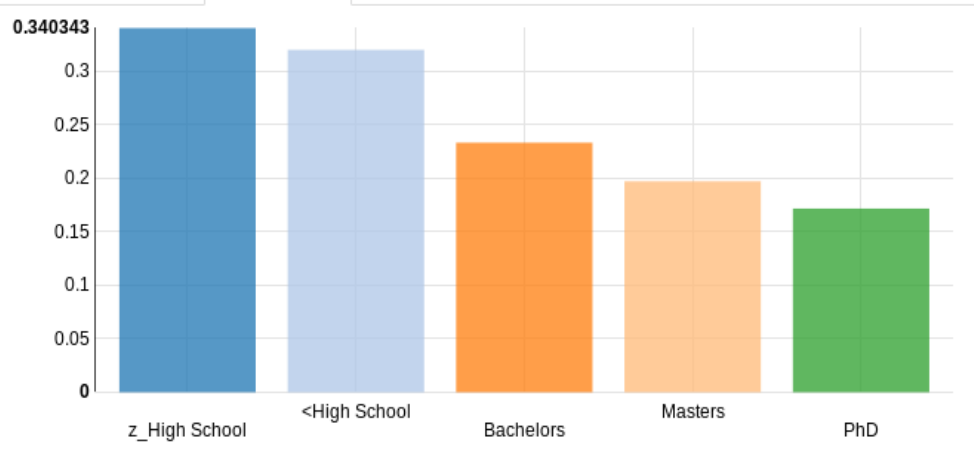
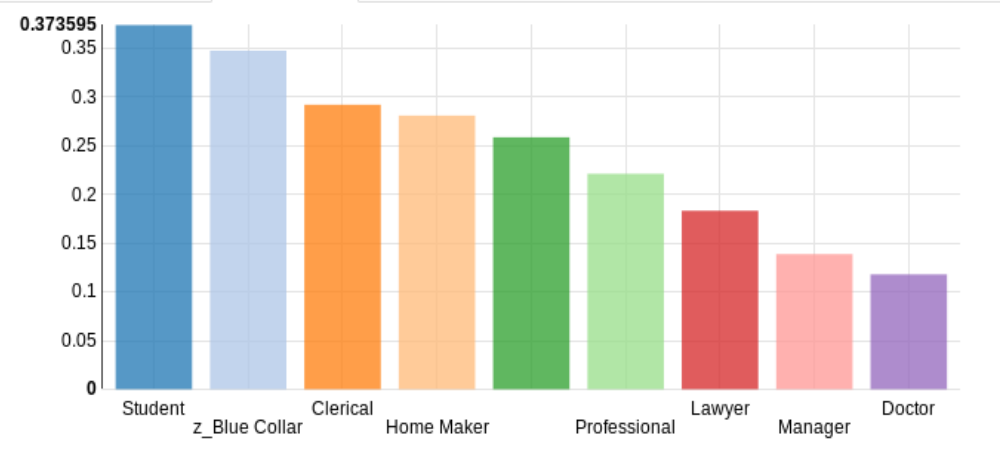
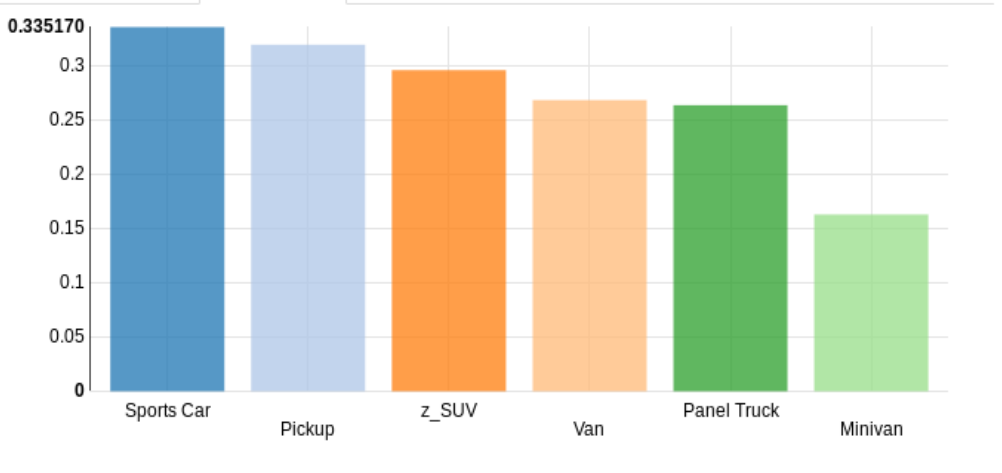
* Those driving minivans are less likely on average to crash *(approx. 16% crash probability*).
* Those driving sports cars, pickups, or SUVs are more likely on average to crash (*approx. 33%, 31% and 30% crash probabilities respectively*).
* Unmarried people are more likely to crash, but there may not be much reason to vary their rates or avoid their business due to their group making above average income.
* Revoked license drivers tend to crash more, but there isn’t any large accompanying income discrepancy.
* Some factors, such as sex, are not worth using to evaluate the merit of doing business with a customer due to there not being as large of a gap in crash rate as seen with other factors.

Conclusions

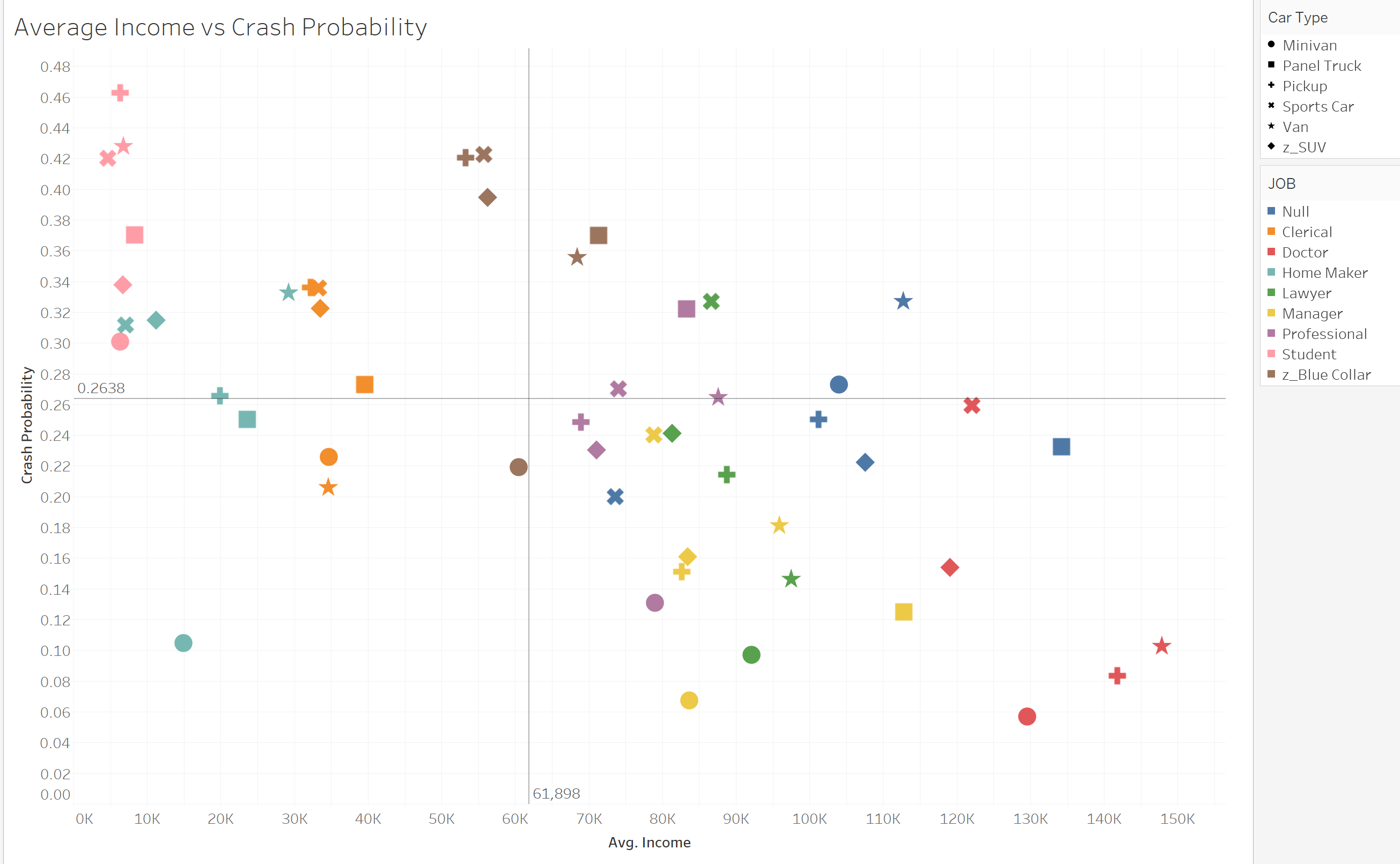
Given some of the base insights and deeper analysis on various combinations of factors via scatterplots, trends suggested that a company **might feel comfortable doing business with a manager or doctor driving any type of car**. They are less likely to crash, and they make above average income, suggesting they may be a good revenue source via their insurance payments and less of a liability due to their more infrequent claims. **Students should be evaluated closely when doing business, as should homemakers and clerical workers** (especially if they don’t drive minivans), due to their average income levels compared to their crash rates.

Though a business should be cautious with the 20-29 age group, they should **feel more comfortable if a potential 20-29 year old client has master’s level education** due to above average income level and below average claim amounts among crashers in that age group.

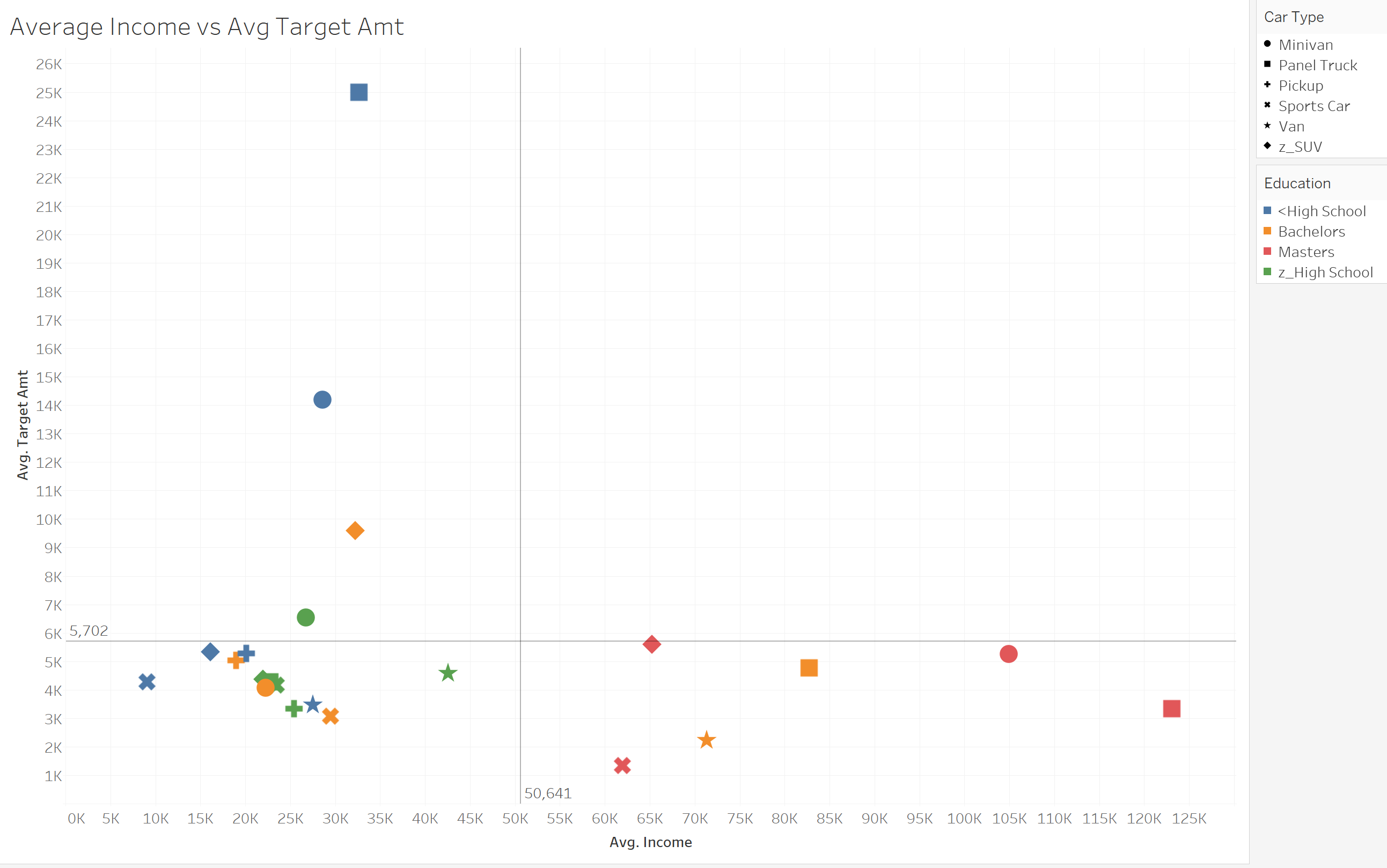
Among non-crashers, using bluebook value of vehicle as a proxy for claim amount, **it is a good idea to remain alert when doing business involving panel trucks or vehicles of any type owned by doctors**. These vehicles typically have higher value and may require larger payouts in the event of a crash.   
A handful of charts are included on the next couple of pages. Not all the charts made in Hue and Tableau are included for the sake of brevity. If you wish to get a closer look at the queries used to make these charts or if you would like to see the full assortment of charts made in Hue and Tableau, email me at [akdavey@email.wm.edu](mailto:akdavey@email.wm.edu).



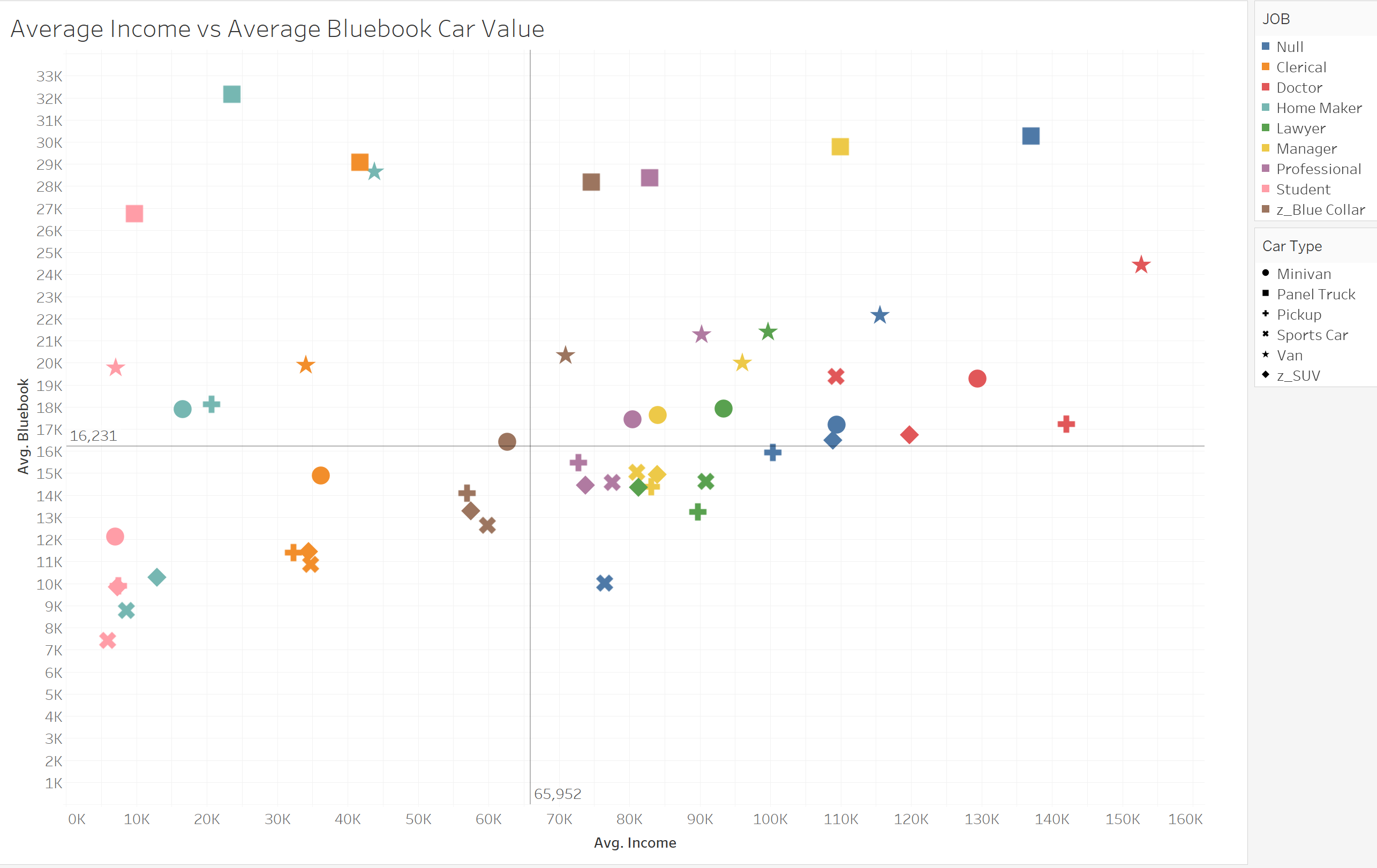
*Impala Charts made using SQL language queries to get the probability of a car crash (y-axis). The x-axes vary as follows: top left: car type, top right: job type, bottom left: education level, bottom right: age group.*



*Above: Tableau scatter plot chart showing Income vs Crash Probability separated by Car Type and Job Type. A company would ideally want to go after customers in the lower right quadrant, who make an above average income with lower crash rates. They should be safe revenue sources, and they should make claims infrequently.*



*Above: Tableau scatter plot chart showing Avg. Income vs Avg. Target Amount separated by Car Type and Education Level filtered to only include clients who have crashed in the 20-29 age group. There appears to be a trend where those clients with Masters level education tend to have higher income and below average claim amounts. These customers would likely be more reliable in making insurance payments and they would tend to make lower claims when they do crash, thus making them more valuable as clients than their similarly aged counterparts.*



*Above: Tableau scatter plot chart showing Avg. Income vs Avg. Bluebook separated by Car Type and Job Type filtered to only include clients who have not crashed. Bluebook here is used as a proxy for Target Payment as cars with a higher bluebook value would tend to require higher claim payouts. A company should be wary when agreeing to insure panel trucks from customers working all kinds of jobs, and they should notably be wary when insuring the cars of doctors, who car types which all fall above average bluebook value.*