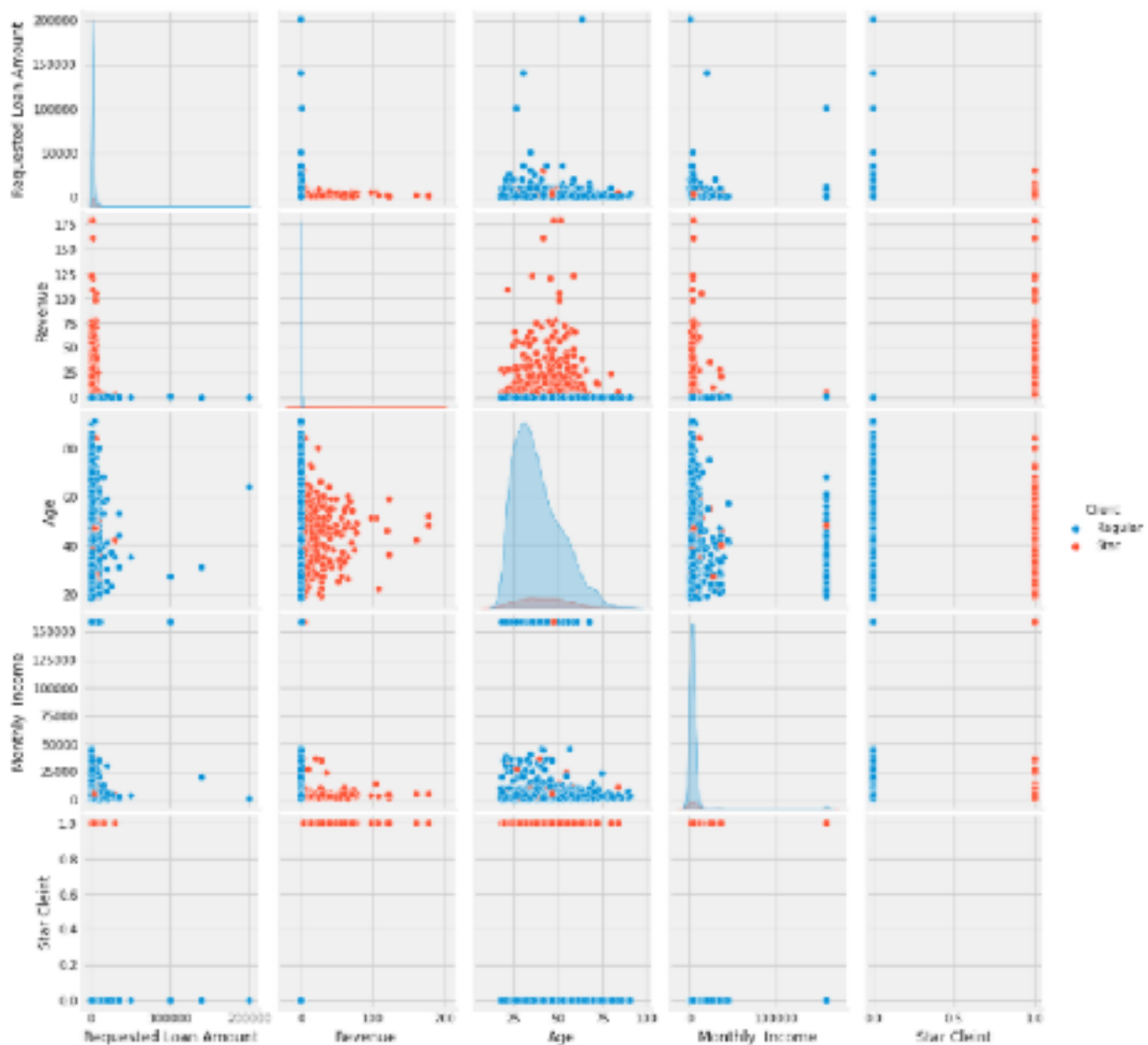# MVP Classifying Star Leads

The Goal of this project is to recognize and predict the the leads of the website as star or regular based on the owners recognition :
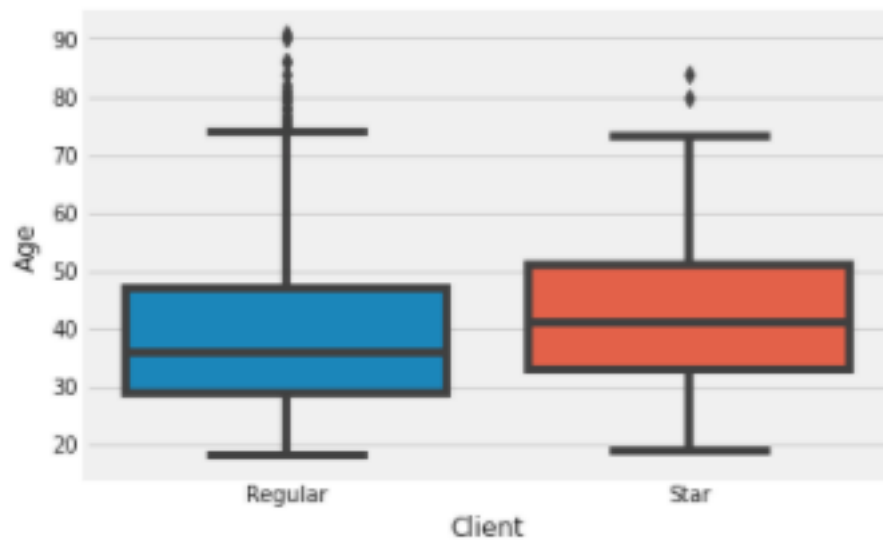
- Strat leads (Revenue >= $3)
- Regular Lead (Revenue < $3)

I'm planning to do the classification modeling with the limit being $3.

Extra:
If I have extra time, I am going to analyze different classifying limits in order to see what else I can be interpreted from the data and features.

.

After receiving the date from him, updated version included 6500 leads which I used google sheets to organize the data in the beginning and then used pandas for cleaning. I also had 405 star leads with the $3 classifying limit which makes it a 6% of my data and I have imbalanced data.

I'm in the final steps of my EDA and I'm going to introduce dummies to my data, butI played with some logistic regression models a bit and the plots that we I have attached are only with numerical features.

Next Steps:

• Introduce dummies
• Normalize data
• Undersample
• KNN,LR, Random Forest , Extra Tree
• Ensemble