

Jack Dando

UT DATA ANALYTICS

COURSE 2 TASK 2

Lessons Learned Report

I think one of the most important visualizations that I produced was a scatter plot placing the balance limit against any of the individual payment data sets out of the six different months worth per customer that we were provided with. From here we can clearly see a strong correlation between people with a higher balance limit being more prone to being paid in full or using revolving credit. It definitely makes sense and might be something that we would have already expected to begin with, but it was to see this laid out for us visually. In a sense, this information is a good sign for credit one: they are providing a higher credit limit to more responsible customers.

Two of the most useless features provided in the data according to some of the visuals that I created ending up being gender and marital status. There was really nothing of great importance here in terms of correlation between factors that could potentially really impact anything to do with credit score, at least compared to other, more financially relevant characteristics.

One of the first things that I wanted to accomplish in my EDA was to come up with a formula that would simplify the process of determining if a customer is someone who could be relied on to make their payments. With the data being a mixture of all the

different bill amounts and payment amounts, it ends up being nearly impossible to determine what the numbers really mean just by looking at them without comparing them very thoroughly. To come up with a precise formula seems to be something that would require a lot of manpower, but this is something that I would suggest the good people at Credit One to potentially look into and invest resources into discovering if they want the most informative solutions to come out of this data.