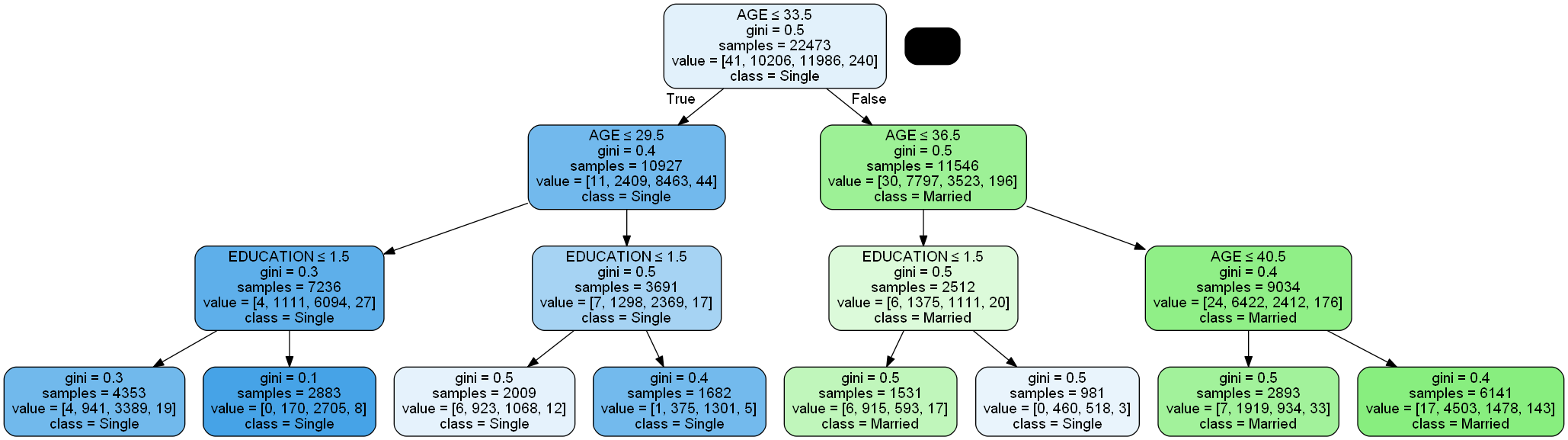
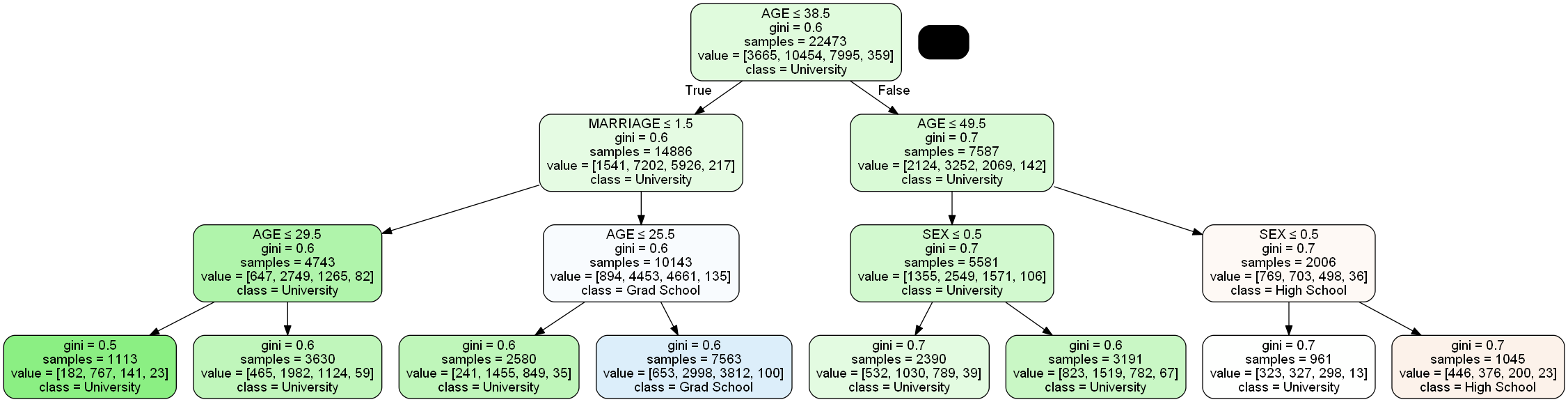
Capital One Findings

|  |
| --- |
| Company name |

|  |  |
| --- | --- |
| To: | Logan Crawford |
| From: | Alex Wallace |
| CC: |  |
| Date: | 05/31/2022 |
| Re: | Predictions Using Capital One Data |
| Comments: | Using the data provided, we were able to distinguish who would default on their loans and predict data from future customers. |

 As I had mentioned in the previous findings, I was surprised to find that university graduates were most likely to default on their loans out of all education levels. I was also surprised to find that single loan holders were fairly close to the number of married loan holders that defaulted on their loans. This went against my initial assumptions due to higher paying jobs from further education and tax breaks from married couples. Another factor that I noticed was age, with younger customers being more likely to default on their loans. Using predictive code, I was able to verify these trends for future customers, although no matter how hard I tried the accuracy for some of these tests didn’t come out perfect; while these results aren’t definitive proof, this data would still be helpful in making future decisions.

 In the model above, I used the data for customer’s marriage status against their age, education and whether or not they defaulted. While the results do show that younger married couples are more likely to default on their loans, especially those with a university level education, the accuracy on this tree was 73%. While that’s not ideal, it’s not too far off from 80% so I think the data would still be worthwhile to use. With the model below, it confirmed that university level graduates, especially those that are younger and married, are more likely to default on their loans. However, this test was only 51% accurate, and while that is a decent amount off from my goal it may still be useful information in the future.

While there is a trend of university graduates being more likely to default on their loans and marital status not playing as much of a role as expected, I think this points to a big flaw in the data: there was no record of salaries. Considering our loans should be very dependent on customer’s salaries, such as the extent we can give out and at what rate they will be able to pay back, I think further research using this info would be more helpful than the data that was provided.