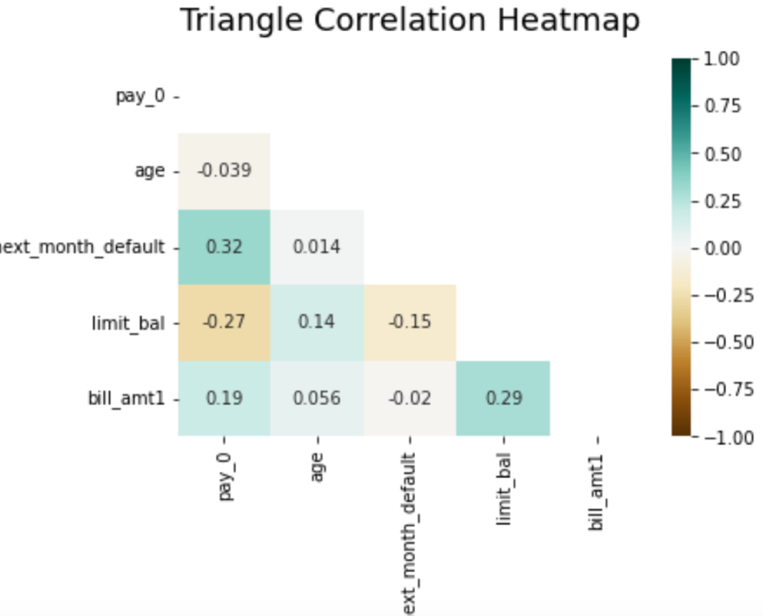
# Credit One’s business customers

1. Which attributes in the data can we deem to be statistically significant to the problem at hand?

The attributes that should be under close attention are limit balance (limit\_bal), the amount of bill statement in September (bill\_amt1), the payment status of customers in September (pay\_0), and the default.

There is a linear relationship between the limit balance ant the amount of bill statement in September. There is also a linear relationship between the payment status of customers in September and the default. The relationship can be appreciated in the map below.



1. What concrete information can we derive from the data we have?

The average limit balance is $167,484. Around 15% of the customers with limit balance bigger than the average defaulted. Around 27% of customers with limit balance lower than the average defaulted. Therefore, customers with a lower limit balance than the average tend to have a higher risk of default than the customers with a higher limit balance.

The customers with below average limit balance and that defaulted have on average at least one month of payment delay in September. The customers with above average limit balance and that have defaulted have on average close to no payment delays in September. Therefore, customers with a limit balance lower than average and at least one month of delay in September, which is the month before the customer defaulted, have a higher risk of default.

The age, sex, education level, and marital status of the customer have a non-existent or very weak relationship with the limit balance and the default of a customer.

1. What proven methods can we use to uncover more information and why?

We could uncover more information with classification algorithms because those can help predict a discrete variable, for example if a customer defaults or not, with the criteria of accuracy which is the number of times the model guessed right. The regression cannot use accuracy and is better used to predict a continuous quantity like the limit balance.

1. Recommendations

I really cannot provide how much credit balance should a customer have, but I highly recommend giving a lower balance or at least below average to customers that have one or more months of delay the month before because the risk of default is higher. Also, if the customer doesn´t have any delay payments the month before then give him a higher limit balance or at least above average because the risk of default is lower.