The first file (train.csv) contains information about the credit card accounts and the loyalty score for each.

test.csv contains the same information, but about different accounts, and it does not contain loyalty scores - because those are the accounts for which we are trying to predict the loyalty! merchants.csv has information about the merchants where the credit cards are being used to make transactions.

Finally, the historical\_transactions.csv and new\_merchant\_transactions.csv files contain lists of all the individual transactions which were made between the credit cards and the merchants, and some features of those transactions (time and date, purchase amount, etc).

Eventually, we'll want to **build a predictive model,** train it on the cards data in train.csv and the information in the merchants and transactions datasets associated with those cards, and then predict the loyalty scores for the cards. But first, we need to load and clean the data!