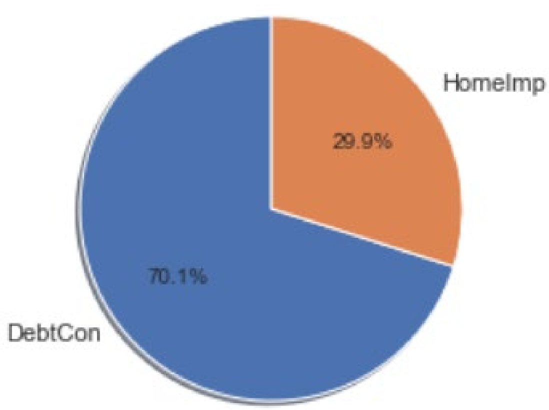
Assignment 1 – HMEQ Data Set Analysis

Here are some summary stats on the dataset that underwent analysis:

Total Loan Records	5960
% Defaulted	20%
Avg Loss Amt	13414.58
Avg Loan Amt	18607.97
Avg Mortgage Due	73001.04
Avg Home Value	101540.4
Avg Years on Job	8.76
% Derogatory	22.40%
% Delinquent	40.57%
Avg Credit Line Age (months)	179.44
Avg number of Inquiries	1.17
Avg number of credit lines	21.25
Avg debt to income ratio	34

Based on our analysis there appears to be a slightly greater chance of a loan defaulting if the loan was intended to be for home improvements (22%) rather than debt consolidation (19%). However, there is much more money lost when a loan for debt consolidation defaults (~$13600) as opposed to a loan for home improvement (~$5800). Even though loans for debt consolidation do not default as often as loans for home improvements, they are definitely very risky in that the consequences for the loan defaulting can result in more than twice as much money being lost. Below is a breakdown of the percentages of loan reason between Debt Consolidation (DebtCon) and Home Improvement (HomeImp). This is based on the 5960 records of data we have been provided in the data set.

LOAN REASON BREAKDOWN



There is also evidence that the borrower's occupation can be an indicator for how risky the loan can be. People working in sales have been found to have the highest loan default percentage at roughly 35%, with self-employed people being the next riskiest with a percentage of 30%. The least riskiest appear to be ProfExe (professional executive?) at roughly 17% and office workers at 13%. The losses for self-employed people have been found to be the highest (~$18500), with people in sales coming in second (~$15600). Borrowers whose occupations fell under 'Other' and office workers had the lowest amount loss at around $9600 and $10200 respectively.