

For this model, we will be using the Credit.csv data set available on the Introduction to Statistical Learning website: <http://www-bcf.usc.edu/~gareth/ISL>. It is also the same data set used in Chapter 6 of the book.

This dataset contains “Balance” which is the average credit card debt across people and 10 explanatory variables, of which 6 are quantitative and 4 are qualitative.

The quantitative variables are: Income, Credit Limit, Credit Rating, No. of Cards, Age and Education. The qualitative variables are: Gender, Student Status, Married and Ethnicity.