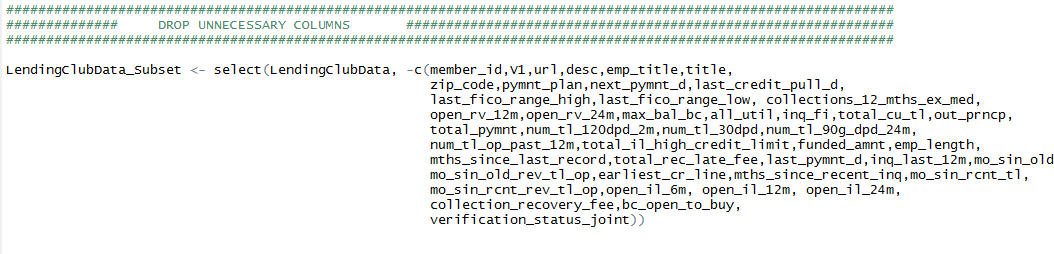
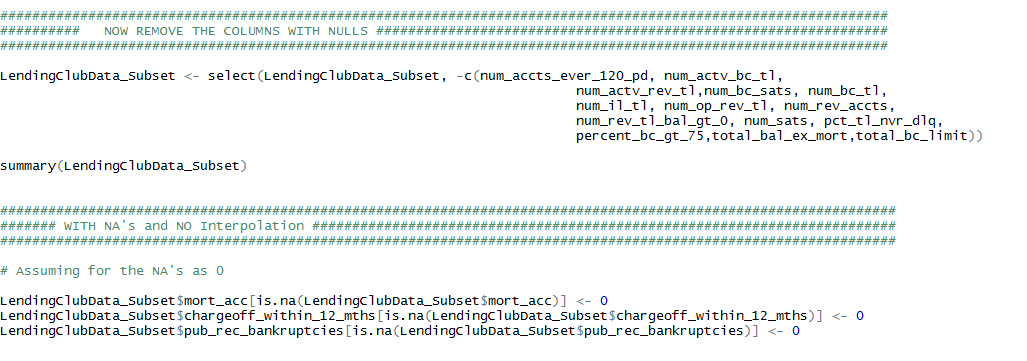
**Unsupervised Learning**

Unsupervised Learning is performed when we are trying to predict labels for the unlabelled dataset.

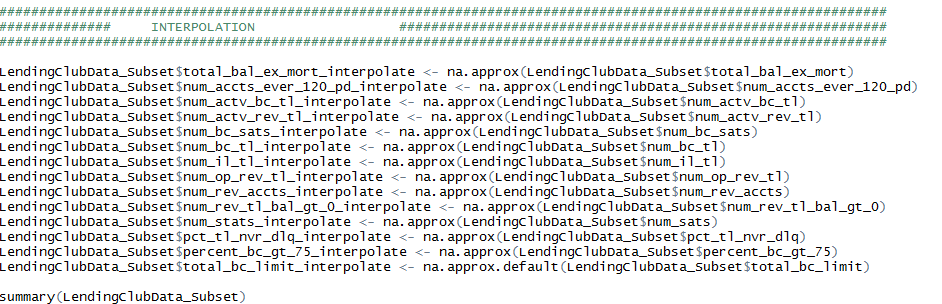
**Feature Engineering:**

1. In Unsupervised learning first we did feature engineering, in which we first validated the entire file and removed the missing values

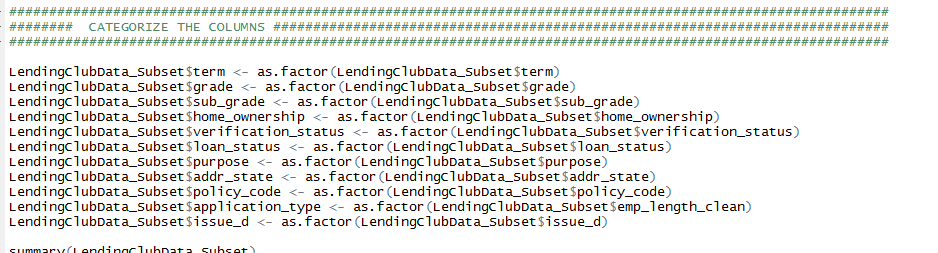




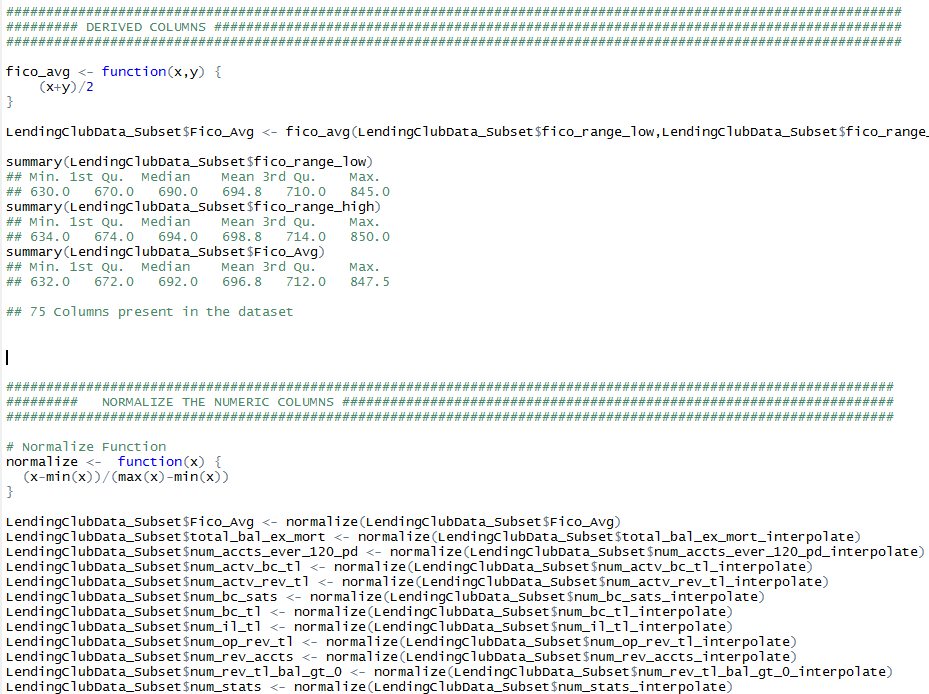
1. Some columns that seemed relevant to us, we interpolated those values



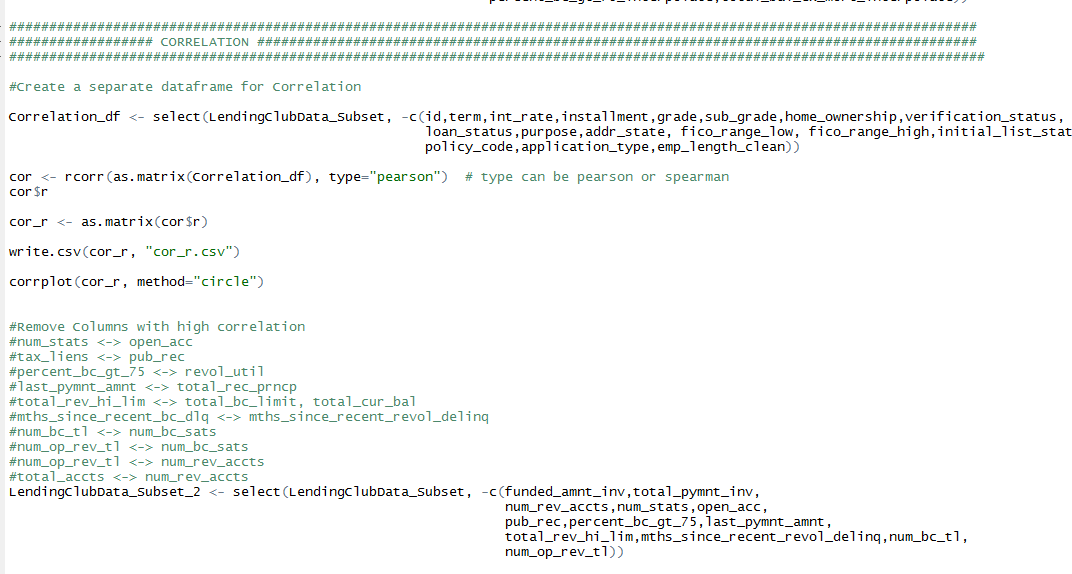
1. Then we chose to convert the categorical columns in the categorical values

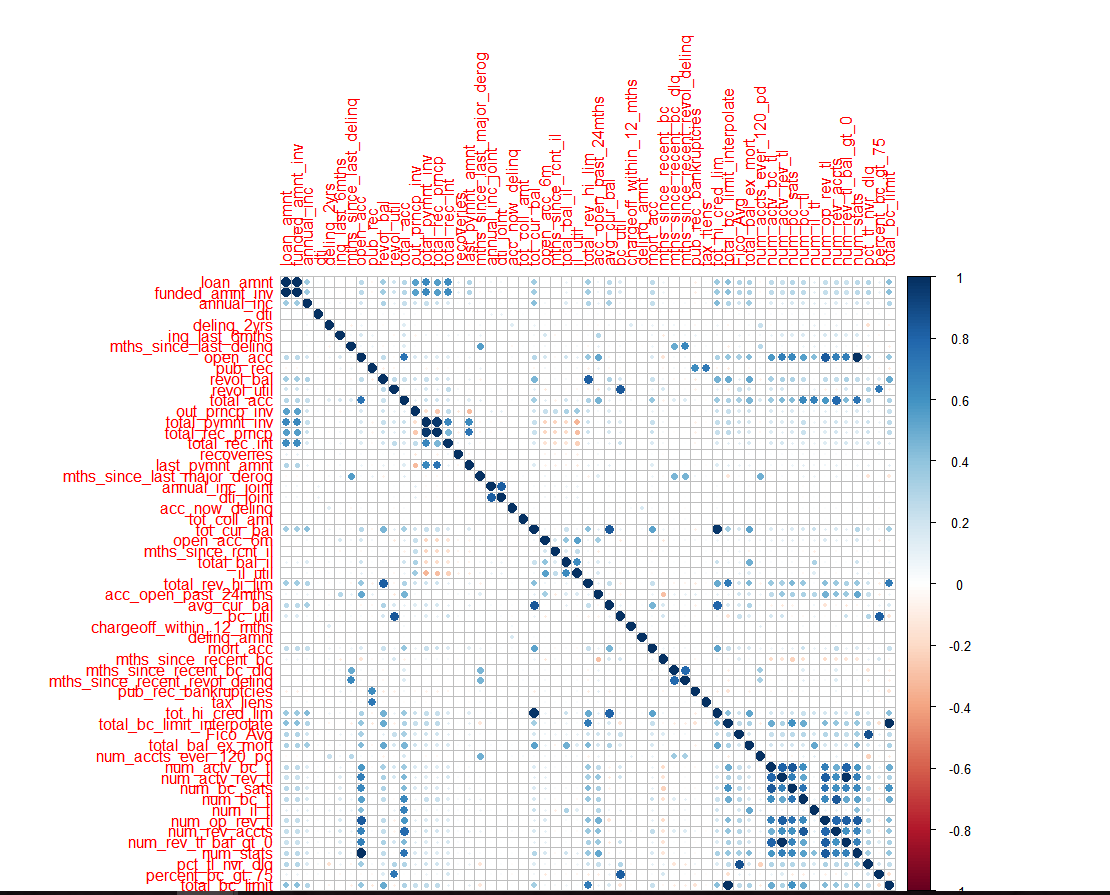


1. We, then chose to normalize the numeric column to put them one scale

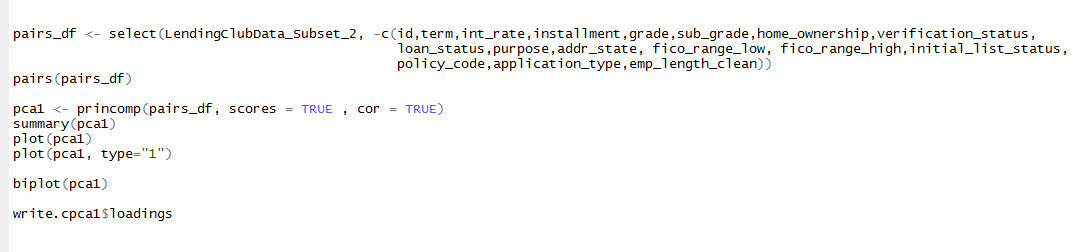


1. After which we created a correlation plot and removed the most correlated columns which weren’t relevant to us.

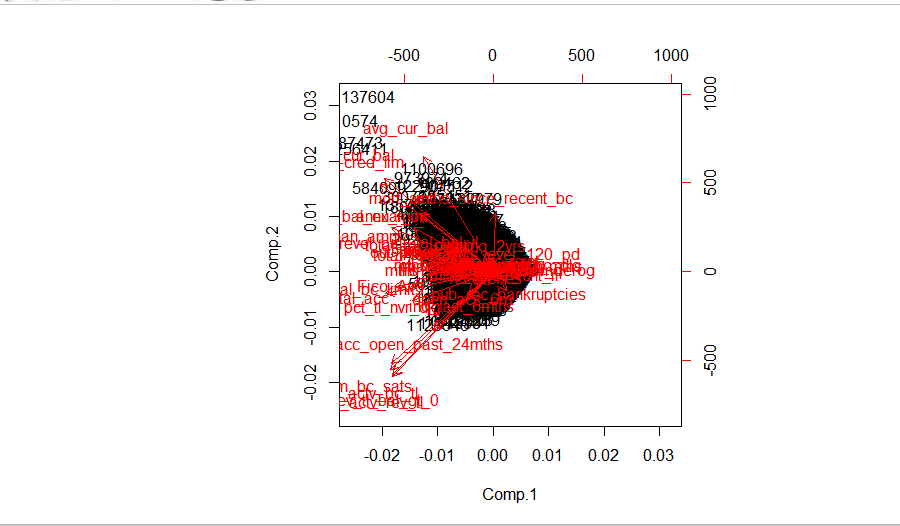




1. We then performed PCA – that is principal component analysis and analysed the loadings to check the variables that cause the maximum variance in the dataset and reduce the dimensionality, so that we get the best features



PCA Graph



PCA Loadings

