

Assignment – Day 2 (Preprocessing in R)
Please use a new R script for the assignment

For the two customer .csv files you loaded:

1. Print the column names.
2. Find the number of columns.
3. Find the number of rows.
4. Look at the data types.
5. Convert the data types.

#####

Merge the two data frames and then

1. Check for the NA values in the overall dataset
 - a. using for loop and
 - b. colSums() functions.
2. Check the number of NA values in each column.
3. Impute the NA values if any.
4. Create a new column based on the following conditions
 - a. If Experience ≤ 5 = “Fresher”
 - b. If Experience > 5 = “Experienced”
5. Find the number of rows that fall under each of the levels in Education column.
6. Find the mean of each numeric column using apply functions.
7. Standardize the Income column.
8. Using for loop , create a new column which has “Y if Mortgage is more than 0, else “N”. and add it to the merged data frame.
9. Dummify the Education column and add it to the merged data frame, remove the Education column.
10. Find the average mortgage based on the education level of the employee.