PREPROCESSING.PY

In the document "Preprocessing.py" an action is defined which we will comment below in this file:

First, the File class is named. This class basically provides us with information about the files, their attributes, the directories, etc. The File class has three constructors:

• File (String path)
• File (String path, String name)
• File (File dir, String name)

The path parameter indicates the path to the directory where the file is located, and name indicates the name of the file. In this case the file "r" that is found in the "static / default of credit card clients.csv" directory is opened.

Immediately afterwards, the out object is used to display the "w" file of the "static / dataset.csv" directory.

Then, to read the file line by line to String, Readline is used. This method returns null when the end of the file is reached. In this document it has been called firstline.

The String class offers a split () method that is very convenient and sufficient for most scenarios. Simply divide the given string according to the delimiter, returning an array of strings. On the other hand, strip () removes the blank spaces at the beginning and end of the string.

We create a string: (string = ""). The strings are constant; their values can not be changed after they are created. Chain buffers support mutable strings. Because the String objects are immutable, they can be shared.

Next, we use the loops to put the values in the previous chain. The for statement provides a compact way to iterate over a range of values. Programmers often refer to it as the "for loop" because of the way it is repeated repeatedly until a particular condition is met.

A first loop, of rank 0 to 6, is introduced the values, along with a semicolon, in the chain, by means of the array args. In the same way it is done with the arrays from 12 to 24. Finally, once the number 24 is entered in the array, the string divided by n is shown on the screen.

In the next "for" the methods "Split" (to divide the string) and "Strip" are again used to eliminate the blanks. Only the values are entered in the chain of 25 elements if the condition that has been established is true. That is, as long as the range is from 12 to 24 and it is also of value 0, this condition will not be fulfilled and the values in the array are not inserted.

Finally, the file is closed and a 0 is returned. This action will be used when the name is "__main__".