SplitByColumn

We have a data file in csv format contains 1000 row and 5 columns. We need to split the file depending on the value of one of the columns.

- All the lines that have the same combination of the specified column value will be written to the same split file.
- The column number used to split is an input from user, preferred to be a command line argument.
- The new filenames should be distinguished by the column values.
- Duplicate rows should be ignored.
- Ignore empty lines.
- Replace empty values with "NA".

Python is preferred but you can code using any other language.
Good Luck!

Example
Original file (file.csv):
1,2,3, ,5
6,7,2,8,9
10,11,2,12,13
14,15,3,16,17
14,15,3,16,17

user input is 3 (the third column).
The code will split the lines into new files depending on the values of the 3rd column.

Output: (two split files)
file_2.csv:
6,7,2,8,9
10,11,2,12,13

file_3.csv:
1,2,3,NA,5
14,15,3,16,17