Record Filter Service Requirements

## Base Requirements

Write a program (suggested: C# console application) that processes all files in a directory with file names matching [customer]\_[filetype]\_[date].csv (ex: Eliza\_MemberRecords\_20160229.csv). The input CSV file will contain a list of records all with the same data fields. Once imported, the program will produce one or more output files based on a set of rules.

Here are the rules to implement:

* Duplicate record with value or combination of values
  + This rule identifies a record that is a duplicate of a previously seen record. The definition of what constitutes a duplicate is configurable as part of the rule. Examples:
    - Duplicate records where FirstName, LastName and PhoneNumber are the same.
    - Duplicate records where only MemeberNumber is the same.
* Check if value is within range
  + This rule tells us if a value in a given record field is in between 2 numbers. Examples:
    - Score is between 0 and 10.
    - Age is between 18 and 60.
* Check if value is one of allowed options
  + This rule tells us if a value is one of a list of possible values. Examples:
    - Color is one of the following values: red, green, blue.
    - Gender is Male.

During processing and rule evaluation, the program needs to take action based on the result of the rule evaluation. There are two possible actions the program may take:

* + Remove record
  + Assign record to a named group
    - Ex: Assign to Group A

If records are assigned to a “group” (the action should allow for the naming of the group), the program will collect all same-group records and write them to a corresponding output file. The name of the resulting output file will follow the convention of [input file name]\_group[groupname].csv.

Rules and actions must be decoupled. So while a rule will specify that a particular record matched rule criteria, what happens to the record as a result of the rule matching is not specified by the rule itself. As such, the actions must be configured outside of the rules themselves.

Rule configuration must be configurable external to the application (i.e. it should not take a recompile to change the rule set).

Files for different customers and file types may have different rule configurations.

Program should be optimized to process large files in the least possible amount of time.

## Sample Use Case

Process a file such that:

* Records with matching FirstName, LastName and Gender are sorted out to Duplicates Group
* Records of males between ages of 18 and 100 to Adult Men group
* Records of females between ages of 18 and 100 to Adult Women group
* Records of everyone under 18 to Children Group

## Nice-to-Haves

(done) Instead of processing all input files in a directory and exiting, continue to monitor the directory for changes and additional files.

(not done) While in this monitoring state allow the rule set and actions map to change without restarting the application.