

CS F363 Compiler Construction

Assignment-1 (Question-3)

Due date: 9 March 2023 11:59 PM

Marks : 5

You are given some data in the form of a text file (namely, *data.txt*). The data is a collection of records. Each record starts with a \$ followed by the customer id (*custId*) followed by a sequence of one or more transactions; each transaction contains the date in DD/MM format and the value of the transaction, and ends with a semicolon (;). The general format of a record is

$$\text{\$ } custId \text{ } DD/MM \text{ } val; \text{ } DD/MM \text{ } val; \text{ } \dots; \text{ } DD/MM \text{ } val;$$

where

1. *custId* is an alphanumeric string contains only upper case letters and digits 0 to 9; starts with upper case alphabets followed by at least one digit from 0,1,...,9. Ex. *ABC12*, *X123*, *PQR13*, ..., etc.
2. *DD* can be 01,02,...,30 and *MM* can be 01,02,...,12.
3. *val* is a non-negative integer that may also start with 0 i.e., 015, 0023, 000, are considered as integers.

Please note the following:

1. A record starts with \$ and is completely contained in the same line. Hence, the information before the \$ symbol in a line is not a part of any record. You can ignore such information.
For example: 04/12 8500 \$AB12 10/10 1000; 02/10 400;
In the above example "04/12 8500" is not part of any record.
2. Further, a line can contain more than one record.
For example: \$AB12 10/10 1000; 02/10 400; \$ BCD123 04/12 45000; contains two records.
3. If a record is missing *custId* (after \$) (call the record as **invalid record**), then just ignore the record.
For example: \$ 10/10 1000; 02/10 400; is missing the *cust_id* so ignore the record.
4. *data.txt* may contain some comments that start with //, just ignore the entire line.
Further, assume that // can appear at the start of the line, but not after any text/record i.e., \$AB123 10/20 100; //something is here never appear in data.txt.
5. You can assume that all the valid records are in the general form mentioned above. *If data.txt contain a date DD/MM with either DD is not in {01,02,...,30} or MM is not in {01,02,...,12}, then return error and stop your program.*

Write a LEX program that takes a date in the DD/MM format and outputs the following:

1. The number of valid transactions on the given date and
2. the *custId* of the customer with maximum transaction value on the given date.

General Instructions:

1. It is a **group assignment** and maximum size of a group is **two**.
2. Submit a single LEX (.l) file and one submission per group is sufficient.
3. Please enter the details of your group in <https://forms.gle/wBNsbf5iQzKZtK8L9>
4. Strictly follow the input and output formats mentioned below.

5. Due date is 9 March 2023 11:59 PM and request for the extension of the due date will not be considered.
6. **Late submission:** Each 1 hr delay fetch 2% penalty and late submission will not be accepted after 24 hours from the due date.
7. Your program must be compiled on Ubuntu 22.04 by following the sequence of commands:
 - `$lex filename.l`
 - `$gcc lex.yy.c -ll`
 - `$/a.out`

Input format: Two text files will be given as part of the input:

1. *data.txt*; the details about the file is given above.
2. *input.txt* contains a single line with an instance of DD/MM.

Output format: Generate a file, *output.txt* (do not use other names) with the output of the problem in the format ***k $custId$*** where *k* is the number of transactions on the given date and *$custId$* is the customer id of the customer with the maximum transaction value on the given date .

If your program is returning an error, just print 0 in output.txt.