

## SED Question 1

Amazon has suddenly got into an issue in their e-commerce application, due to which their database has been corrupted. They store the data of a customer in the format id, name, customerkey, gmail. Here id is the unique number to a person, name field consists of customer's name, customerkey is a special field amazon stores for product purchases and finally gmail field consists of the gmail of the customer.

Due to the data corruption most of the values in the id and name fields have been changed to some garbage values. But the id and name values can be obtained from the other two fields. The customerkey field has a value in the format "id:some random data " and the gmail field consists of values of format "[name@gmail.com](#)".

Task: Write a sed script in "task.sed" file to restore the corrupted data in the first two columns using the third and fourth columns. The input file is "customers.csv" and outputs the data into "output.txt".

```
sed -E 's/[^,]*,[^,]*,([0-9]+)(:[^,]*),([A-Za-z]+)@gmail.com/\1,\3,\1\2,\3@ gmail.com/g' customer.txt
```

Sample input file:

```
id,name,customerkey,gmail
1he,svdgh,23:sdfdf,markwilliams@gmail.com
45,john,45:pwmd,john@gmail.com
wpe,spencer,1:snfs,spencer@gmail.com
```

Required output file:

```
id,name,customerkey,gmail
23,markwilliams,23:sdfdf,markwilliams@gmail.com
45,john,45:pwmd,john@gmail.com
1,spencer,1:snfs,spencer@gmail.com
```

Usage:

```
sed -r -f task.sed customers.csv > output.txt
```