

# Guess the Flipkart Query

**Flipkart** is a popular Indian e-commerce portal. One of the most common actions performed by users of the portal, is to use the search box and query for a brand, product or product-line. The search box then returns the best matching products which it can find - along with their prices, details, descriptions, etc.

We tried out **twenty** different search queries (specified below), and made a list of some of the product names which were returned in response to them. You are provided with a list of **N** names of products from this list. Your task is to guess, **which search query** each of them was returned in response to.

The twenty search queries we made were:

```
axe deo
best-seller books
calvin klein
camcorder
camera
chemistry
chromebook
c programming
data structures algorithms
dell laptops
dslr canon
mathematics
nike-deodorant
physics
sony cybershot
spoken english
timex watch
titan watch
tommy watch
written english
```

## Here's a small example of the task at hand:

*In response to which of these queries, was the product 'Dell Vostro 2520 Laptop (2nd Gen PDC/ 2GB/ 320GB/ Linux...' (most likely) returned?*

Answer:**dell laptops**

*In response to which of these queries, was the product 'Calvin Klein One Eau de Toilette - 200 ml' (most likely) returned?*

Answer:**calvin klein**

## Input Format

The first line contains an integer **N**.

This is followed by **N** lines each containing the name of a product.

## Input Constraints

$1 \leq N \leq 200$

The product names will not exceed **200 characters** in length. Sometimes, when the product name is

long and descriptive, after the first 55 characters, there are likely to be a series of dots, such as the examples below. Please handle them appropriately (strip them off, or ignore them).

```
Laptops: AMD Mobile Platform, AMD Vision, Barebook, Cen...
Dell Vostro 2520 Laptop (2nd Gen PDC/ 2GB/ 320GB/ Linux...
Dell Inspiron 15R 5521 Laptop (3rd Gen Ci7/ 8GB/ 1TB/ W...
```

## Output Format

The output should contain exactly **N** lines.

The  $i^{\text{th}}$  line should contain the query (your best guess) which returned the  $i^{\text{th}}$  product name in the input file. The query should strictly be from one of the twenty queries specified above, **as is**. Please do not add any leading or trailing spaces or any extra punctuation. Also ensure that the case remains the same.

## Sample Input, Output and Training Files

The sample input, output and training files can be accessed at the following links:

[Training File](#)

[Sample Input](#)

[Sample Output](#)

The training file can also be opened from your program, during execution. It can be opened using the name "training.txt" and is available in the same directory where the program is being run.

## Sample Input

```
60
Data Structures and Algorithms with Object- Oriented Design Patterns in C++ 1 Edition (Paperback)
God Moments: Stories That Inspire, Moments to Remember (Paperback)
The Ultimate C: Concepts, Programs and Interview Questions (Paperback)
Canon EOS 1100D SLR (Black, with Kit (EF S18-55 III))
A Textbook of Organic Chemistry for JEE Main & Advanced and Other Engineering Entrance Examinations
(Paperback)
Test your C ++ Skills 1 Edition (Paperback)
IIM Ahmedabad Business Books: Day to Day Economics (Paperback)
Calvin Klein One Eau de Toilette - 200 ml
.....
```

## Sample Output

```
data structures algorithms
written english
c programming
dslr canon
chemistry
c programming
best-seller books
calvin klein
.....
```

## Explanation

The first product in the sample input is a book 'Data Structures and Algorithms with Object- Oriented Design Patterns in C++ 1 Edition (Paperback)' which was returned in response to the query '**data structures algorithms**'. The second product in the sample input is a paperback book 'God Moments: Stories That Inspire, Moments to Remember (Paperback)' which was returned in response to the query '**written english**'. Please note, that as in the real world, there are always cases like the second one, where it is nearly impossible - to identify which is the most appropriate search query which led to this product: that is fine - you can answer with your best guess in such situations.

## Training File

A small training file with a few examples of products returned for the various search queries is available. Please note, that this is only a **small** training file, and it is expected that a mix of simple and creative ideas from machine learning, string matching and information retrieval will be used in the submitted solution.

*The format of the training file is as specified:*

The first line contains an integer **N**.

This is followed by **N** lines each containing the product name, and the search query, **separated by a tab** character.

```
N
productName_1    query_1
productName_2    query_2
productName_3    query_3
....
```

Here's a quick look at what the training file looks like:

```
111
Calvin Klein IN2U Eau de Toilette - 150 ml (For Men)  calvin klein
For The Love of Physics (Paperback) physics
Nike Fission Deodorant Spray - 200 ml (For Men)    nike-deodrant
Spoken English (With CD) 2nd Edition (Paperback)   spoken english
The C++ Programming Language 3 Edition (Paperback) c programming
.....
```

## Scoring

Your score for a test case =  $C/N * M$  where:

M = Maximum Score for the test case C = search queries correctly identified N = Total number of product names in the test case

The hidden test case carries **thrice** as much weightage as the sample test case (which is visible on hitting 'Compile and Test').

## Libraries

Libraries available in our Machine Learning/Real Data challenges will be enabled for this contest and are listed [here](#). Please note, that occasionally, a few functions or modules might not work in the constraints of our infrastructure. For instance, some modules try to run multiple threads (and fail). So please try importing the library and functions and cross checking if they work in our online editor in case you plan to develop a solution locally, and then upload to our site.

