# Exercise: Regular Expressions

Problems for exercise and homework for the Python Fundamentals Course @SoftUni.
Submit your solutions in the SoftUni judge system at https://judge.softuni.bg/Contests/1743

## 1. Capture the Numbers

Write a program that **finds all numbers** in a sequence of strings.

The **output** is all the numbers, **extracted** and **printed on a single line** – each separated by a **single space**.

### Examples

| Input | Output |
|---|---|
| The300<br>What is that?<br>I think it's the 3rd movie<br>Lets watch it at 22:45 | 300 3 22 45 |

| Input | Output |
|---|---|
| 123a456<br>789b987<br>654c321<br>0 | 123 456 789 987 654 321 0 |
| **Input** | **Output** |
| Let's go11!!!11!<br>Okey!1! | 11 11 1 |

## 2. Find Variable Names in Sentences

Write a program that finds all **variable names** in a given string. A variable name starts with an **underscore ("_")** and contains **only Capital and Non-Capital English Alphabet letters and digits**. Extract only their names, **without the underscore.** Try to do this **only with regular expressions**.

The **output** consists of all variable names, **extracted** and **printed on a single line**, each **separated** by a **comma**.

| Input | Output |
|---|---|
| The _id and _age variables are both integers. | id,age |

| Input | Output |
|---|---|
| Calculate the _area of the _perfectRectangle object. | area,perfectRectangle |
| **Input** | **Output** |
| __invalidVariable _evenMoreInvalidVariable_ _validVariable | validVariable |

SoftUni

# 3.  Find Occurrences of Word in Sentence

Write a program that finds, **how many times** a **given word**, is **used** in a **given sentence**. Note that letter case does not matter – it is **case-insensitive**.

The **output** is a single number indicating the **amount of times** the sentence contains the word.

## Examples

| Input | Output |
|---|---|
| The waterfall was so high, that the child couldn't see its peak.<br>the | 2 |

| Input | Output |
|---|---|
| How do you plan on achieving that? How? How can you even think of that?<br>how | 3 |
| **Input** | **Output** |
| There was one. Therefore I bought it. I wouldn't buy it otherwise.<br>there | 1 |

# 4.  Extract Emails

Write a program to **extract all email addresses from a given text**. The text comes at the only input line. Print the emails on the console, each at a separate line. Emails are considered to be in format **<user>@<host>**, where:

- **<user>** is a sequence of **letters** and **digits**, where '**.**', '**-**' and '**_**' can appear between them.
  - Examples of valid users: "**stephan**", "**mike03**", "**s.johnson**", "**st_steward**", "**softuni-bulgaria**", "**12345**".
  - Examples of invalid users: ''**--123**", "**.....**", "**nakov_-**", "**_steve**", "**.info**".
- **<host>** is a sequence of at least two words, separated by dots '**.**'. Each word is sequence of letters and can have hyphens '**-**' between the letters.
  - Examples of hosts: "**softuni.bg**", "**software-university.com**", "**intoprogramming.info**", "**mail.softuni.org**".
  - Examples of invalid hosts: "**helloworld**", "**.unknown.soft.**", "**invalid-host-**", "**invalid-**".
- Examples of **valid emails**: **info@softuni-bulgaria.org**, **kiki@hotmail.co.uk**, **no-reply@github.com**, **s.peterson@mail.uu.net**, **info-bg@software-university.software.academy**.
- Examples of **invalid emails**: **--123@gmail.com**, **...@mail.bg**, **.info@info.info**, **_steve@yahoo.cn**, **mike@helloworld**, **mike@.unknown.soft.**, **s.johnson@invalid-**.

## Examples

| Input | Output |
|---|---|
| Please contact us at: support@github.com. | support@github.com |
| Just send email to s.miller@mit.edu and j.hopking@york.ac.uk for more information. | s.miller@mit.edu j.hopking@york.ac.uk |
| Many users @ SoftUni confuse email addresses. We @ Softuni.BG provide high-quality training @ home or @ class. –- steve.parker@softuni.de. | steve.parker@softuni.de |

# 5.       Furniture

Write a program to calculate the total cost of different types of furniture. You will be given some lines of input until you receive the line "Purchase". For the line to be valid it should be in the following format:

**">>{furniture name}<<{price}!{quantity}"**

The price can be floating point number or whole number. Store the names of the furniture and the total price. At the end print the each bought furniture on separate line in the format:

**"Bought furniture:**
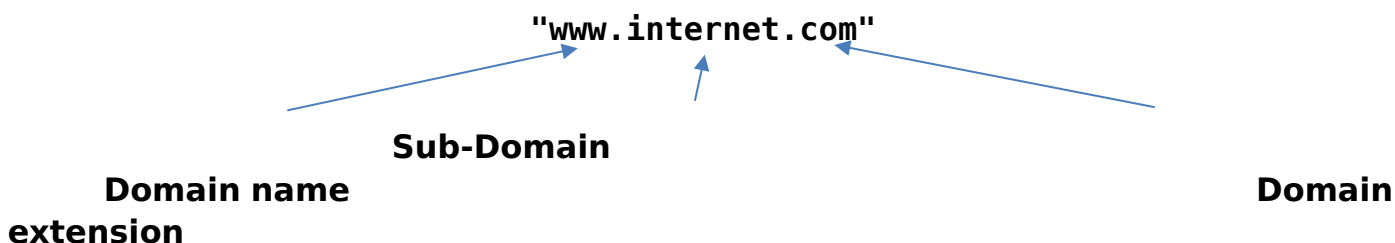
**{1st name}**

**{2nd name}**

**…"**

And on the last line print the following: **"Total money spend: {spend money}"** formatted to the second decimal point.

## Examples

| Input | Output | Comment |
|---|---|---|
| >>Sofa<<312.23!3 >>TV<<300!5 >Invalid<<!5 Purchase | Bought furniture: Sofa TV Total money spend: 2436.69 | Only the Sofa and the TV are valid, for each of them we multiply the price by the quantity and print the result |

# 6.  *Extract the Links

Write a program that **extracts links** from a **given text**. The text will come in the form of strings, each representing a sentence. You need to extract **only the valid links** from it. Example:

"www.internet.com"

**Sub-Domain**

**Domain name**

**Domain**

**extension**

SoftUni

The **Sub-Domain** must always be "**www**". The **Domain name** can consist of English alphabet letters (**uppercase** and **lowercase**), digits and dashes ("**–**"). The **Domain extension** consists of one or more **domain blocks**, a **domain block** consists of a **dot** followed by **one or more lowercase** English alphabet **letters**, a **Domain extension** must have at least **one domain block** in order to be **valid**. The Sub-Domain and Domain name must be separated by a single **dot**. Any link that **does NOT follow** the specified above rules should be treated as **invalid**.

**Example incorrect links:**

- "**ww**.justASite.bg"
- "**lel**.awesome.com"
- "www.weird_site.hit.bg"
- "www.no-symb**#^**ols-allow**%**ed.com"
- "www.mark**.12**"
- "www.web-site**.**"
- "www.example-site**._*^#**"

**Example correct links:**

- "Some textwww.softuni.bg"
- "Just a link in a www.sea-of-text.bg-swimming around"
- "Instruments www.Instruments.rom.com.trombone2000 Instrument here"
- "All your ice cream flavors-www.scream.for.ice.cream(We  also have squirrels)"

The **output** is all valid links you've found, printed – each on a new line.

## Examples

| Input | Output |
|---|---|
| Join WebStars now for free, at www.web-stars.com<br>You can also support our partners:<br>Internet - www.internet.com<br>WebSpiders - www.webspiders101.com<br>Sentinel - www.sentinel.-ko | www.web-stars.com<br>www.internet.com<br>www.webspiders101.com |

| Input | Output |
|---|---|
| Need information about cheap hotels in London?<br>You can check us at www.london-hotels.co.uk !<br>We provide the best services in London.<br>Here are some reviews in some blogs:<br>London Hotels are awesome! -<br>www.indigo.bloggers.com<br>I am very satisfied with their services -<br>ww.ivan.bg<br>Best Hotel Services! - www.rebel21.sedecrem.moc | www.london-hotels.co.uk<br>www.indigo.bloggers.com<br>www.rebel21.sedecrem.mo<br>c |

SoftUni