

**The results are in!** See what nearly 90,000 developers picked as their most loved, dreaded, and desired coding languages and more in the 2019 Developer Survey.

Unix & Linux Stack Exchange is a question and answer site for users of Linux, FreeBSD and other Un\*x-like operating systems. Join them; it only takes a minute:

[Sign up](#)

**Here's how it works:**

Anybody can ask a question

Anybody can answer

The best answers are voted up and rise to the top

# UNIX & LINUX

## regex to match identifiers without double \_

[Ask Question](#)



2



I want a regex to match identifiers consisting of letters, digits and \_ , without double underscores. My current attempt:

```
^(?!_)(?!.*?_)$[a-zA-Z][a-zA-Z0-9_]+$
```

Example:

```
Abdfsfdfsf__  
1B2345_  
v1__23456  
__23456789b  
12345-6789  
123456789  
1_fsfdfsfdf  
v_fsfdfsfdf_fsfdf  
v_123  
v__123  
v134234_fsfdfsd  
123456  
a1b2c3d4e5
```

Matched:

```
v1__23456  
v_fsfdfsfdf_fsfdf  
v_123  
v__123  
123456789
```

By using our site, you acknowledge that you have read and understand our [Cookie Policy](#), [Privacy Policy](#), and our [Terms of Service](#).

How can I remove the rows  
v1\_\_23456 and v\_\_123 from the  
matches?

regular-expression

edited Jun 6 '13 at 23:33



Gilles

547k 130 1113  
1629

asked Jun 6 '13 at 8:18



cheng6588

13 1 3

**migrated** from [serverfault.com](#)

Jun 6 '13 at 16:32

This question came from our site for  
system and network administrators.

You want print all lines have two  
underscores, or you want print all  
lines have only one? Your match  
doesn't contain v1\_23456 –  
[cuonglm](#) Jun 6 '13 at 8:45

## 4 Answers



1

```
^(?!_)          # don't start wi
(?!.*?_$)      # don't end with
[a-zA-Z]       # the first char
[a-zA-Z0-9_]+$ # after that dig
```



You've expressed in two different  
ways that the first character must not  
be `_`, but nothing here says  
anything about `__` in the middle.

Using negative lookahead, it is  
simple to express “one or more  
alphanumeric characters or `_`, don't  
start with `_` or a digit, don't end with  
`_`, and don't allow `__` anywhere”:

```
^(?!([0-9_]|.*_.*_$|)[0-9A-Za-z]
```

Without negative lookahead (e.g. in  
`awk` or `grep -E`), you can split the  
pieces:

```
^[A-Za-z][0-9A-Za-z]*(_[0-9A-Za-z]
```

Start with a letter, then zero or more  
alphanumerics. then you can have

answered Jun 7 '13 at 1:06



Gilles

**547k** 130 1113

1629

thx. and thanks for understand what i want to express. – [cheng6588](#) Jun 7 '13 at 2:28



There is one way to do it:

0

```
perl -nle 'print if /( _ )*. * \1/' pa
```



Ouput:

```
1B2345_
12345-6789
123456789
1_fsdfdfsd
v_123
v134234_fsdfs
123456
a1b2c3d4e5
```

The regex `/ ( _ )*. * \1/` will match all line have two or more underscore `_` . Just negate the condition to print lines that don't match.

answered Jun 6 '13 at 8:38



cuonglm

**106k** 25 211 309

i think i shown the wrong meaning. I test this in the RegexBuddy to check my regex could match the username or not – [cheng6588](#) Jun 6 '13 at 8:53

Finally, you want to match line has double underscores `__` or line contains double underscores, i.e `__` or `_  
_<some char>_` ? – [cuonglm](#) Jun 6 '13 at 9:27



You want to find lines that DO NOT contain two consecutive underscores? Unless I'm missing something, this should work:

0

```
grep -v ' __ '
```



```
grep -v '___*'
```

or to remove lines that contain multiple adjacent underscores, but keep lines that contain only single underscores:

```
grep -v '___*'
```

edited Jun 6 '13 at 21:09

answered Jun 6 '13 at 21:02



CaffeineConnoisseur

330 1 3 9



0



```
$ grep -E '^(_?[A-Za-z0-9])*[A-Za-z0-9_]*$'
1_fsdfdfsdf
v_fsdfdfsdf_fsdfd
v_123
v134234_fsdfsd
a1b2c3d4e5
```

edited Jun 7 '13 at 2:05

answered Jun 7 '13 at 1:59



frostschutz

27.7k 1 57 90