

**grep, reguláris kifejezések (ismétlés)****1. Áttekintés, emlékeztető**

A `grep` szűrő esetén a következő reguláris kifejezéseket használhatjuk.

| reguláris kifejezés          | mire illeszkedik  |
|------------------------------|---|
| <code>david</code>           | a <code>david</code> szóra fog illeszkedni  |
| <code>[Dd]avid</code>        | David vagy david  |
| <code>.avid</code>           | bármilyen karakter ( <code>.</code> ), majd <code>avid</code>   |
| <code>^david\$</code>        | olyan sor, mely csak a <code>david</code> szót tartalmazza  |
| <code>d*avid</code>          | akárhány <code>d</code> , majd <code>avid</code> (pl.: <code>avid</code> , <code>david</code> , <code>ddavid</code> , <code>dddavid</code> , stb.)  |
| <code>[0-9]</code>           | egy számjegy (ekvivalens a <code>[0123456789]</code> reg. kifejezéssel)   |
| <code>[A-Gm-p]</code>        | ekvivalens az <code>[ABCDEFGmop]</code> reg. kifejezéssel   |
| <code>^[^abcef]avid\$</code> | A sorban 5 karakter szerepeljen, melynek a 2.-tól az 5.-ig terjedő része <code>avid</code> . Az 1. karakter nem lehet <code>a</code> , <code>b</code> , <code>c</code> , <code>e</code> , vagy <code>f</code> . |

Az alábbi karakterosztályok olyan gyakran fordulnak elő, hogy külön rövidítéseket vezettek be a számukra.

| reguláris kifejezés    | mire illeszkedik  |
|------------------------|---|
| <code>[:digit:]</code> | számjegy ( <code>[0-9]</code> )                                     |
| <code>[:alpha:]</code> | az ábécé egy betűje ( <code>[a-zA-Z]</code> )                       |
| <code>[:space:]</code> | egy „whitespace” karakter ( <code>space</code> , <code>TAB</code> ) |
| <code>[:upper:]</code> | nagybetű ( <code>[A-Z]</code> )                                     |
| <code>[:lower:]</code> | kisbetű ( <code>[a-z]</code> )                                      |

Ismétlések:

| reguláris kifejezés         | mire illeszkedik   |
|-----------------------------|--|
| <code>*</code>              | a megelőző reg. kif. 0-szor vagy többször fordul elő   |
| <code>\+</code>             | a megelőző reg. kif. 1-szer vagy többször fordul elő   |
| <code>\?</code>             | a megelőző reg. kif. 0-szor vagy 1-szer fordul elő (vagy szerepel vagy nem)                    |
| <code>\{n\}</code>          | a megelőző reg. kif. pontosan <code>n</code> -szer fordul elő                                  |
| <code>b[aeiou]\{2\}t</code> | <code>b</code> , majd 2 magánhangzó, végül <code>t</code>                                      |
| <code>\{n,\}</code>         | a megelőző reg. kif. legalább <code>n</code> -szer fordul elő                                  |
| <code>\{n,m\}</code>        | a megelőző reg. kif. legalább <code>n</code> -szer, legfeljebb <code>m</code> -szer fordul elő |

Csoportosítás: egy bizonyos mintát megjegyezzük, s azt a későbbiekben fel tudjuk használni.

| reguláris kifejezés     | jelentése  |
|-------------------------|--|
| <code>da\(\vid\)</code> | a <code>\(</code> és <code>\)</code> jelek közé eső részt megjegyeztük   |
| <code>\(abc\)1\1</code> | A <code>\(</code> és <code>\)</code> jelek közti részt beteszi az 1. regiszterbe. Vagyis az „abc”-t megjegyezzük. A <code>\1</code> előhívja az 1. regiszter tartalmát s behelyettesíti. Vagyis ez a minta a következő sztringre fog illeszkedni: „abcabcabc”. |
| <code>\(.\)bb1</code>   | mire illeszkedik: <code>abba</code> , <code>cbbc</code> , stb.   |

**2. Feladatok #1**

Az alábbi kérdések mindegyike az `/etc/passwd` állományra vonatkozik.

1. Írassuk ki az `/etc/passwd` állomány azon sorait, melyek „j”-vel kezdődnek. (Ez tulajdonképpen ekvivalens azzal, hogy írassuk ki a „j” betűvel kezdődő felhasználói nevekhez tartozó sorokat.)

---

2. Írassuk ki azon „j”-vel kezdődő sorokat, ahol a felhasználó login shell-je a `bash`.

---

3. Az `/etc/passwd` állományból írassuk ki a 0 és 99 közé eső csoportazonosítókat. A csoportazonosító a 4. oszlopban van.

---

4. Az `/etc/passwd` állományból írassuk ki a 0 és 99 közé eső csoportazonosítókat, illetve a hozzájuk tartozó felhasználói neveket is. A kimenet egy sora a következőképpen nézzen ki (példa):

```
usbmux:46
```

---

5. Az előző pont kimenetét csoportazonosító alapján csökkenő sorrendben írassuk ki.

---

6. A 4. pont kimenetén annyit módosítsunk, hogy csak a felhasználói azonosítókat írassuk ki, méghozzá rendezetten. Vagyis: mely felhasználók csoportazonosítója esik 0 és 99 közé?

---

**3. Feladatok #2**

A gyakorlat honlapjáról töltsük le a `grepdata.txt` fájlt az „ismetles” nevű könyvtárból. Az alábbi kérdések erre a fájlra fognak vonatkozni.

1. Írassuk ki mindazon telefonszámokat, melyek kiterjesztést is tartalmaznak (`x` vagy `X` után négy darab számjegy).

---

2. Írassuk ki mindazon sorokat, melyek három számjeggyel kezdődnek, majd utána szóköz következik. Használja a `\{` és `\}` jeleket az ismétlés leírásához.

---

3. Írassuk ki mindazon sorokat, melyekben szerepel egy dátum (a minta legyen egyszerű, nem kell a 2000 előtti dátumokra is működnie).

---

4. Írassuk ki mindazon sorokat, melyek tartalmaznak egy magánhangzót (`a`, `e`, `i`, `o`, `u`), utána szerepel egy tetszőleges karakter, majd ismét ugyanez a magánhangzó jön. Például „eve” és „adam” elfogadható, „vera” viszont nem. Tipp: `\(` és `\)`.

---

5. Írassuk ki mindazon sorokat, melyek nem `S`-sel (nagy `S`) kezdődnek.

---

#### 4. Feladatok #3

Továbbra is a `grepdata.txt` állománnyal dolgozunk. Az alábbi kérdések megválaszolásakor használjuk a `grep` parancssori kapcsolóit is.

1. Írassuk ki azokat a sorokat, melyek tartalmazzák a `CA` sztringet (akár kis- vagy nagybetűsen).

---

2. Írassuk ki azokat a sorokat, melyek tartalmazznak egy email címet (szerepel bennük a `@` jel). Írassuk ki a sorok sorszámát is.

---

3. Írassuk ki azokat a sorokat, melyek nem tartalmazzák a „Sep.” négybetűs szót (beleértve tehát a pontot is).

---

4. Írassuk ki azokat a sorokat, melyek tartalmazzák a „de” kétbetűs szót (mint *önálló* szót).

---

#### 5. Feladatok #4

Írjunk reguláris kifejezést, mely illeszkedik a következő sztringekre:

1. Dátum, mely a következőképpen néz ki (példa): `2012-03-26`

---

2. Dátum, mely a következőképpen néz ki (példa): `March 26, 2012`

---

3. VISA bankkártyaszám. Az első számjegy 4, s összesen 16 számjegyből áll.

---

4. A sorokban nevek szerepelnek a következő formában: családnév, vessző, tetszőleges számú szóköz és/vagy tab, majd a keresztnév első betűje, végül pont. Példa:

Kiss, I.  
Fekete, P.

---

5. Képet tartalmazó URL. Példa:

<http://i.imgur.com/D25PG.jpg>  
[http://upload.wikimedia.org/wikipedia/commons/0/09/Hoh\\_Rain\\_Forest\\_Maples.JPG](http://upload.wikimedia.org/wikipedia/commons/0/09/Hoh_Rain_Forest_Maples.JPG)

---