

## Regulární výrazy – příklady, použití

- např. pro filtry výběrů dat nebo filtry vlastních přehledů → můžete zvýšit flexibilitu pomocí regulárních výrazů

### Nejběžnější zástupné znaky a jejich význam:

- . jakýkoli jeden znak
- \* nulové či vícenásobné opakování předchozí položky
- + jedno- či vícenásobné opakování předchozí položky
- ? nulové či jednonásobné opakování předchozí položky
- () obsah v závorce se zadává jako položka
- [] odpovídá jedné položce v tomto seznamu
- vytvoření rozsahu v rámci seznamu
- | nebo
- ^ odpovídá počátku pole
- \$ odpovídá konci pole
- \ ukončení kteréhokoli z výše uvedených výrazů

### **Svislá čára (|) a závorky = NEBO**

*pra(bába|děd)*

*bude odpovídat výrazům prabába či praděd*

*Bož(en|k)a*

*bude odpovídat jménům Božena a Božka*

**Hranaté závorky a pomlčky [ - ]** = shoduje se jakákoli 1 položka ze seznamu v hranatých závorkách, pomlčkou zjednoduší po sobě jdoucí řadu položek

*[a,b,c] = [a-c]*

- při vyhledávání podobných klíčových slov, která se liší jen 1 částí, zadám do pole filtru toto: *PART[357]* → stejné jako *PART[3] + PART[5] + PART[7]*

*[a-lqtz]* = shoda s jakýmkoli jedním písmenem od a do l, případně s písmeny q, t nebo z

**Obrácené lomítko** = deaktivuje libovolný speciální znak na obyčejný písmenný znak, nejčastěji se používá ve spojení s tečkou

*mysite\.com*

IP adresa: [64\68\82\164](#)

**Tečka** = nahrazuje jakýkoli 1 znak (shoda s jakýmkoli 1 znakem)

[.es](#) – ve shodě s výrazy *pes, ves, 3es, les ...*

[homepage.com](#) – ve shodě s *homepagescom, homepagercom ..*

X [homepage\com](#) – ve shodě pouze s *homepage.com ;)*

**Stříška (^)** – pokud stříšce předchází jakýkoli další znak, řetězec se s tímto regulárním výrazem neshoduje

- dejme tomu, že máte například na webu dvě stránky:

<http://www.mysite.com/secondpage/contact/> a <http://www.mysite.com/contact/>  
Google Analytics tyto stránky označuje **/secondpage/contact/** a **/contact/**  
protože již zná doménu, kterou je *www.mysite.com*.

Pokud chcete vyhledat všechny řetězce, které začínají výrazem **/contact/** (druhá možnost), ale do řádku vložíte pouze regulární výraz **/contact/**, bude se s výrazem shodovat i možnost, kterou chcete vyloučit, tj. */secondpage/contact/*. Pokud chcete dosáhnout shody pouze se stránkou **<http://www.mysite.com/contact/>**, můžete použít následující regulární výraz:

[^/contact/](#)

- dejme tomu, že chcete vytvořit filtr zahrnující rozsah IP adres vaší firmy. Vaše IP adresy začínají dvoumístným číslem, například *64.xx.xx.xxx*. Vy proto chcete odfiltrovat cokoli, co by mohlo mít podobu *164.xx.xx.xx* apod. Tento problém řeší použití stříšky: **[^64](#)**

**Dolar \$** - odpovídá konci pole

- předpokládejme například, že některé vaše stránky končí písmeny „*htm*“ a jiné „*html*“. Chcete e-mailový přihlašovací formulář zapsat jako krok 1 Google Analytics (část cíle), přejete si ovšem zahrnout pouze novou verzi *.htm*. Váš výraz může vypadat takto:

[/email-form\.htm\\$](#)

= pokud za posledním písmenem „*m*“ ve výrazu „*htm*“ stojí v názvu stránky cokoli dalšího, nebude se jednat o shodu s výrazem. (Všimněte si obráceného lomítka před tečkou ve výrazu „[/email-form\.htm\\$](#)“, který zajišťuje, že tečku služba Google Analytics interpretuje jako tečku a nikoli jako zvláštní znak.)

**Plus +** - jedno- či vícenásobné opakování předchozí položky (znaku)

[Johan+a](#)

- shoduje se s výrazy *Johana, Johanna i Johannna*

*[abc]+*

- s tímto výrazem se shodují například výrazy: *a, ab, cab, c, b, bbbb* apod.

**Otazník ?** - přítomnost či nepřítomnost předchozího výrazu

*eth?anol*

- předchozí výraz (zde písmeno *h*) může i nemusí být

- shoduje se s *ethanol* i *etanol*

**Hvězdička \*** - nulové či vícenásobné opakování předchozí položky

- př. firma značí své produkty (např. díly) pětimístnými ID čísly ve formátu *ID00000*  
a my chceme zjistit, kolik lidí vyhledává díl číslo *84* → zápis bude takto:

*ID0\*84*

- hvězdička určuje, že předchozí položka (zde *nula*) nemusí vůbec být obsažena, ale zároveň může být obsažena jednou nebo i vícekrát

- výsledek tedy bude: *ID84, ID084, ID0084, ID00084, ID0000084* atd.