

1. d1: Plzeň je krásné město a je to krásné místo  
 d2: Ostrava je osklivé místo  
 d3: Praha je také krásné město Plzeň je hezčí

q1: krásné město

## 2. frequency

kolikrát se každé slovo  $t$  vyskytuje v dokumentu  $d$

d1: Plzeň:1	a:1	d3: Praha:1	q1: krásné:1
je:2	to:1	je:2	město:1
krásné:2	místo:1	také:1	
město:1		krásné:1	
		město:1	
d2: Ostrava:1	osklivé:1	Plzeň:1	
je:1	místo:1	hezčí:1	

## 3. term frequency

$$tf_{t,d} = \begin{cases} 1 + \log_{10}(f_{t,d}) & \text{kdysi } f_{t,d} > 0 \\ 0 & \text{jindy} \end{cases}$$

d1: Plzeň:1	a:1	d3: Praha:1	q1: krásné:1
je:1,301	to:1	je:1,301	město:1
krásné:1,301	místo:1	také:1	
město:1		krásné:1	
		město:1	
d2: Ostrava:1	osklivé:1	Plzeň:1	
je:1	místo:1	hezčí:1	

## 4. document frequency

$df_t$  = počet dokumentů se slovem  $t$

Plzeň:2	město:2	místo:2	Praha:1
je:3	a:1	Ostrava:1	také:1
krásné:2	to:1	osklivé:1	hezčí:1

## 5. inverse document frequency

$idf = \log_{10}\left(\frac{N}{df_t}\right)$ , kde  $N$  = počet dokumentů

Plzeň:0,176	místo:0,176
je:0	Ostrava:0,477
krásné:0,176	osklivé:0,477
město:0,176	Praha:0,477
a:0,477	také:0,477
to:0,477	hezčí:0,477

## 6. TF-IDF

$$= t_{t,d} \cdot idf_t$$

d1: Plzeň: 0,176

pe: 0

krajsné: 0,229

město: 0,176

a: 0,577

to: 0,577

místo: 0,176

d3: Praha: 0,577

pe: 0

také: 0,577

krajsné: 0,176

město: 0,176

Plzeň: 0,176

herci: 0,577

q1: krajsné: 0,176

město: 0,176

d2: Ostrava: 0,577

pe: 0

ošklivé: 0,577

místo: 0,176

## 7. Kosinová podobnost

$$\cos(\alpha) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

slovo +	TF-IDF d1	TF-IDF d2	TF-IDF d3	TF-IDF q1
Plzeň	0,176	<del></del>	0,176	0
pe	0	0	0	0
krajsné	0,229	<del></del>	0,176	0,176
město	0,176	<del></del>	0,176	0,176
a	0,577	<del></del>	<del></del>	0
to	0,577	<del></del>	<del></del>	0
místo	0,176	0,176	<del></del>	0
Ostrava	<del></del>	0,577	<del></del>	0
ošklivé	<del></del>	0,577	<del></del>	0
Praha	<del></del>	<del></del>	0,577	0
také	<del></del>	<del></del>	0,577	0
herci	<del></del>	<del></del>	0,577	0

$$d1 \times q1$$

$$A \cdot B = (0,176 \cdot 0) + (0 \cdot 0) + (0,229 \cdot 0,176) + (0,176 \cdot 0,176) + (0,577 \cdot 0) + (0,577 \cdot 0) + (0,176 \cdot 0)$$

$$A \cdot B = 0,040304 + 0,030976$$

$$A \cdot B = 0,07128$$

$$\|A\| = \sqrt{(0,176)^2 + 0^2 + (0,229)^2 + (0,176)^2 + (0,577)^2 + (0,577)^2 + (0,176)^2}$$

$$\|A\| = 0,7748$$

$$\|B\| = \sqrt{(0,176)^2 + (0,176)^2}$$

$$\|B\| = 0,2489$$

$$\cos(\alpha) = \frac{0,07128}{0,7748 \cdot 0,2489}$$

$$\cos(\alpha) = \frac{0,07128}{0,1928} = 0,3697$$

$$d_2 \times q_1$$

$$A \cdot B = (0,477 \cdot 0) + (0 \cdot 0) + (0,477 \cdot 0) + (0,176 \cdot 0) + (0 \cdot 0,176) + (0 \cdot 0,176)$$

$$A \cdot B = 0$$

$$\|A\| = \sqrt{(0,477)^2 + 0^2 + (0,477)^2 + (0,176)^2}$$

$$= 0,697$$

$$\|B\| = 0,2589$$

$$\cos(\angle) = \frac{0}{0,697 \cdot 0,2589}$$

$$\cos(\angle) = \frac{0}{0,1735} = 0$$

$$d_3 \times q_1$$

$$A \cdot B = (0,176 \cdot 0,176) + (0,176 \cdot 0,176)$$

$$= 0,061952$$

$$\|A\| = \sqrt{(0,176)^2 + (0,176)^2 + (0,176)^2 + (0,477)^2 + (0,477)^2 + (0,477)^2}$$

$$= 0,831$$

$$\|B\| = 0,2589$$

$$\cos(\angle) = \frac{0,061952}{0,831 \cdot 0,2589}$$

$$\cos(\angle) = \frac{0,061952}{0,21921} = 0,283$$

$$d_1 \times q_1 = 0,3697$$

$$d_2 \times q_1 = 0$$

$$d_3 \times q_3 = 0,283$$

↳ document 1 je nejrelevantnější vůči dotazu  $q_1$