

## interpretace dat: základní pojmy

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
2																	
3	xxxx	SONX	R/L/vokál	věta/slovo								JÁ					
4	<b>MLUVČÍ</b>	<b>KÓD MLUVČÍHO</b>	<b>MĚŘENÁ HLÁSKA</b>	<b>VĚTA/SLOVO</b>	<b>začátek hlásky v ms</b>	<b>konec hlásky v ms</b>	<b>hodnota F1 v Hz</b>	<b>hodnota F2 v Hz</b>	<b>hodnota F3 v Hz</b>	<b>hodnota F4 v Hz</b>		<b>R</b>	<b>průměr F1</b>	<b>průměr F2</b>	<b>průměr F3</b>	<b>průměr F4</b>	
5	Polomská	SON1	R	vrkat	553	641	580	1726	3183	4519			530	1771	2890	4318	
6	Polomská	SON1	R	brkat	561	649	547	1769	2651	4251							
7	Polomská	SON1	R	brkat	761	854	571	1835	2776	4133		<b>R</b>	<b>medián F1</b>	<b>medián F2</b>	<b>medián F3</b>	<b>medián F4</b>	
8	Polomská	SON1	R	prchat	669	760	487	1842	2986	4392			534	1774	2842	4322	
9	Polomská	SON1	R	hrkat	572	646	492	1746	2860	3998							
10	Polomská	SON1	R	Petr kope.	870	945	520	1793	2824	4215		<b>R</b>	<b>odchylka F1</b>	<b>odchylka F2</b>	<b>odchylka F3</b>	<b>odchylka F4</b>	
11	Polomská	SON1	R	Kopr poroste.	945	1015	560	1675	2692	4455			37	52	185	189	
12	Polomská	SON1	R	Macch poučuje.	602	696	486	1779	3145	4584			<b>přesnost (%)</b>	<b>93</b>	<b>97</b>	<b>93</b>	<b>96</b>
13	Polomská	SON1	L	včít	934	1026	539	1375	3108	4286		<b>L</b>	<b>průměr F1</b>	<b>průměr F2</b>	<b>průměr F3</b>	<b>průměr F4</b>	
14	Polomská	SON1	L	ilky	575	667	544	1367	3041	4169			515	1294	3085	4340	
15	Polomská	SON1	L	libec	534	615	455	1310	3073	4239							
16	Polomská	SON1	L	plkat	523	600	545	1330	3097	4515		<b>L</b>	<b>medián F1</b>	<b>medián F2</b>	<b>medián F3</b>	<b>medián F4</b>	
17	Polomská	SON1	L	hitat	556	629	505	1200	3080	4633			522	1304,5	3088,5	4280,5	
18	Polomská	SON1	L	Pekl kuře.	953	1035	497	1231	3147	4424							
19	Polomská	SON1	L	Kopl kámen.	737	811	540	1299	3203	4275		<b>L</b>	<b>odchylka F1</b>	<b>odchylka F2</b>	<b>odchylka F3</b>	<b>odchylka F4</b>	
20	Polomská	SON1	L	Máchl pákou.	1177	1239	493	1222	2933	4176			31	66	74	157	
21	Polomská	SON1	A	táta	665	891	794	1361	3186	4148			<b>přesnost (%)</b>	<b>94</b>	<b>95</b>	<b>98</b>	<b>96</b>
22	Polomská	SON1	E	léto	792	987	593	1904	2966	4280							
23	Polomská	SON1	I	týpek	336	467	288	2733	3561	4376							
24	Polomská	SON1	O	kófa	742	948	588	934	3180	4319							
25	Polomská	SON1	Ú	skútr	758	892	518	1860	3697	4531							

V tabulce vpravo jsou spočítané následující hodnoty:

- i) *průměrné* hodnoty pro formanty F1, F2, F3, F4 zvlášť pro hlásku /r/ a zvlášť pro hlásku /l/; funkce v Excelu:

CZ	en
=PRŮMĚR(.....)	=MEAN(.....)

- ii) *mediány* (střední hodnoty – když bychom všechny hodnoty uspořádali postupně podle velikosti do jedné řady, medián je hodnota, kterou najdeme uprostřed této řady) pro formanty F1, F2, F3, F4 zvlášť pro hlásky /r/ a /l/; funkce v Excelu:

CZ	en
=MEDIAN(.....)	=MEDIAN(.....)

- iii) *směrodatné odchylky* (o kolik se průměrně liší naměřené hodnoty od mediánu) pro formanty F1, F2, F3, F4 zvlášť pro hlásky /r/ a /l/, funkce v Excelu:

CZ	en
=SMODCH(.....)	=SMODCH(.....)

- iv) *přesnost* – pokud jsou odchylky měření malé, je naše měření vyrovnané (není rozkolísané) – to je pro nás dobře – číslo v řádku „přesnost %“ udává, kolik procent z mediánové hodnoty tvoří odchylka

– čím je přesnost větší, tím se jednotlivá měření odchylují od mediánu méně, tzn. všechny hodnoty se pohybují velmi blízko mediánové hodnoty --- naše měřené hodnoty nejsou rozkolísané;

– čím je přesnost menší, tím se jednotlivá měření odchylují od mediánu více, tzn. alespoň některé hodnoty se pohybují dost daleko od mediánové hodnoty --- naše měřené hodnoty jsou rozkolísané; funkce v Excelu:

vlastní funkce = kolik procent z mediánové hodnoty tvoří odchylka
= 100-(hodnota <sub>směrodatná odchylka</sub> / hodnota <sub>medián</sub> )*100

- A) na základě hodnot přesnosti z tabulky výše můžeme konstatovat, že je naše měření velmi málo rozkolísané (v této chvíli máme měření od jednoho mluvčího, u jednoho mluvčího většinou najdeme konzistentní hodnoty) – ve většině případů dosahuje hodnoty přes 95 %  
 B) u hodnot F1 a F3 u hlásky /r/ vidíme znatelnější rozkolísání – přesnost dosahuje 93 %

## interpretace dat II: z-score

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2	xxx	SONX	R/L/vokál	věta/slovo								
3	MLUVČÍ	KÓD MLUVČÍHO	MĚŘENÁ HLÁSKA	VĚTA/SLOVO	F1	z-score pro F1	F2	z-score pro F2	F3	z-score pro F3	F4	z-score pro F4
4	Polomská	SON1	R	vrkat	580	1,36	1726	-0,86	3183	1,58	4519	1,05
5	Polomská	SON1	R	krkat	547	0,45	1768	-0,05	2651	-1,29	4251	-0,36
6	Polomská	SON1	R	brkat	571	1,11	1835	1,24	2776	-0,61	4133	-0,98
7	Polomská	SON1	R	prchat	497	-1,19	1842	1,38	2986	0,52	4392	0,39
8	Polomská	SON1	R	hrkat	492	-1,05	1746	-0,47	2860	-0,16	3998	-1,69
9	Polomská	SON1	R	Petr kope.	520	-0,28	1793	0,43	2824	-0,35	4215	-0,55
10	Polomská	SON1	R	Kopr poroste.	560	0,81	1675	-1,84	2692	-1,07	4455	0,72
11	Polomská	SON1	R	Machr poučuje.	486	-1,21	1779	0,16	3145	1,38	4584	1,40
12	Polomská	SON1	L	víci	539	0,79	1375	2,65	3108	0,31	4285	-0,34
13	Polomská	SON1	L	kiky	544	0,96	1387	3,04	3041	-0,60	4169	-1,09
14	Polomská	SON1	L	blbac	455	-1,96	1310	0,52	3073	-0,17	4239	-0,64
15	Polomská	SON1	L	plkat	545	0,99	1330	1,17	3097	0,16	4515	1,12
16	Polomská	SON1	L	hlkat	505	-0,32	1200	-3,09	3080	-0,07	4633	1,87
17	Polomská	SON1	L	Pekl kuře.	497	-0,58	1231	-2,07	3147	0,84	4424	0,54
18	Polomská	SON1	L	Kopl kámen.	540	0,83	1299	0,16	3203	1,59	4275	-0,41
19	Polomská	SON1	L	Máchl pákou.	493	-0,71	1222	-2,37	2933	-2,06	4176	-1,04

	N	O	P	Q	R
JÁ					
R	průměr F1	průměr F2	průměr F3	průměr F4	
		530	1771	2890	4318
R	medián F1	medián F2	medián F3	medián F4	
		534	1774	2842	4322
R	odchylka F1	odchylka F2	odchylka F3	odchylka F4	
		37	52	185	189
L	průměr F1	průměr F2	průměr F3	průměr F4	
		515	1294	3085	4340
L	medián F1	medián F2	medián F3	medián F4	
		522	1304,5	3088,5	4280,5
L	odchylka F1	odchylka F2	odchylka F3	odchylka F4	
		31	31	74	157

### z-score:

- udává, kolikanásobek směrodatné odchylky musíme přičíst/odečíst k průměru, abychom dostali naměřenou hodnotu (pro F1 vrkat:  $(37 * (+1,36)) + 530 = 580$   
pro F1 hrkat:  $(37 * (-1,05)) + 530 = 492$ )
- to znamená, že z-score nám udává (vzhledem ke směrodatné odchylce), jak moc se hodnoty přibližují (popř. vzdalují) naší průměrné hodnotě
- z-score je způsob, jak ověřit rozkolísanost datového souboru, na základě posuzování jednotlivých naměřených hodnot

### PŘÍKLADY:

- pokud je z-score 0: naměřená hodnota se rovná [směrodatná odchylka\*0 + průměr] = průměr
- pokud je z-score 1: naměřená hodnota se rovná [směrodatná odchylka\*(+1) + průměr]
- pokud je z-score -1: naměřená hodnota se rovná [směrodatná odchylka\*(-1) + průměr]

### INTERPRETACE HODNOT z-score:

- z-score v intervalu (-1;1) – naměřená hodnota je velmi blízko průměru
- z-score v intervalu (-2;2) – naměřená hodnota ještě ujde, není až tak vzdálená průměru
- z-score v intervalu (-3;3) – naměřená hodnota už je dost daleko od průměru, zvažujeme odstranění outlierů (velmi vychýlených hodnot)
- z-score mimo výše uvedené intervaly – hodnoty vystřelily příliš daleko od průměru, určitě by se měly považovat za outliery a z datového souboru by měly být odstraněny

!!! z-score počítáme, když máme dostatečně velký objem dat, pro malé objemy dat se nehodí !!!

to platí i pro všechny dříve uvedené hodnoty: tzn. průměrování, mediánování, měření standardní odchylky i z-score se dělá, když mám dost dat! aplikováním statistiky na malé datové soubory zkresluje data!

V tabulce výše jsou červeně označeny ty z-score hodnoty, které jsou mimo interval (-2;2). To jsou hodnoty, které se oproti průměru vychylují výrazně a zvažovali bychom jejich odebrání. V našem datovém souboru se takto vychyluje celkem šest hodnot.