

Cvičení 3:

Výpočet základních statistických charakteristik

Statistické metody a zpracování dat 1 (podzim 2016)

Martin Lulák

Import dat z txt do xls

- Otevřeme si excel a přes něj soubor námi vybrané stanice ve formátu txt
- Vyběhne nám okénko pro import, klikneme dokončit a máme hodnoty

Průvodce importem textu (1/3)

Průvodce převodem textu zjistil, že data jsou pevné šířky.

Zvolte datový typ, který datům odpovídá nejlépe, a potom klikněte na tlačítko Další.

Zdrojový datový typ

Vyberte typ souboru, který datům nejlépe odpovídá:

Oddělovač - Pole jsou oddělena speciálními znaky (čárka, tabulátor).

Pevná šířka - Pole jsou zarovnána do sloupců a jsou oddělena mezerami.

Začátek importu na řádku: 1 Typ souboru: 852 : Středoevropské jazyky (DOS)

Data obsahují záhlaví.

Náhled souboru C:\Users\Martin\Desktop\cviceni_03_data\cviceni_03_data\AS.TXT.

	Rok	1	2	3	4	5	6	7	8	9	10	11	12	
1														
2	1961	-38	12	44	97	81	144	131	142	159	88	15	-35	
3	1962	-18	-31	-25	64	77	124	133	149	112	78	3	-60	
4	1963	-101	-76	5	69	105	140	163	144	125	65	49	-63	
5	1964	-51	-36	-23	66	120	164	170	141	120	52	20	-30	

Storno < Zpět Další > **Dokončit**

Import dat z txt do xls

- Můžeme si odmazat poslední dva sloupce, které nebudeme potřebovat
- Označíme sloupce, pravým tlačítkem na názvy sloupců klikneme a dáme odstranit

The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Domů' (Home). The data table has columns labeled A through S. The last two columns, N and O, are highlighted in green. A context menu is open over column O, with the 'Odstranit' (Delete) option highlighted in red.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Rok	1	2	3	4	5	6	7	8	9	10	11	12	01.XII	04				
2	1961	-38	12	44	97	81	144	131	142	159	88	15	-35	70	1				
3	1962	-18	-31	-25	64	77	124	133	149	112	78	3	-60	51	1				
4	1963	-101	-76	5	69	105	140	163	144	125	65	49	-63	53	1				
5	1964	-51	-36	-23	66	120	164	170	141	120	52	20	-30	60	1				
6	1965	-19	-51	-9	46	90	138	133	136	110	70	-11	-1	53	1				
7	1966	-58	24	7	77	115	154	136	135	114	95	-5	-15	65	1				
8	1967	-30	-5	25	43	109	129	175	149	119	94	15	-29	67	1				
9	1968	-44	-21	20	70	90	143	145	142	112	82	4	-50	58	1				
10	1969	-33	-52	-16	47	120	129	170	143	122	92	26	-75	57	1				
11	1970	-58	-39	-15	31	87	154	142	150	117	65	32	-26	54	1				
12	1971	-30	-19	-14	71	121	116	169	167	100	72	5	6	64	1				
13	1972	-56	-13	37	45	92	134	163	140	90	47	13	-24	56	1				
14	1973	-27	-21	17	26	108	141	152	166	133	50	3	-34	60	1				
15	1974	-1	0	41	58	84	119	130	156	112	25	16	14	63	1				
16	1975	10	-7	17	45	106	125	164	167	144	62	-4	-20	68	125				
17	1976	-20	-24	-15	47	113	155	176	142	108	71	19	-41	61	124				
18	1977	-27	-1	41	32	100	138	144	142	93	81	20	-18	62	109				
19	1978	-28	-43	17	49	95	124	137	126	96	67	8	-18	53	104				
20	1979	-69	-36	13	40	110	155	130	139	109	57	4	13	56	114				
21	1980	-56	-2	8	31	82	125	126	145	118	53	1	-26	51	105				
22	1981	-48	-32	46	57	110	138	140	145	125	59	16	-49	60	119				
23	1982	-51	-30	22	41	107	143	170	153	148	76	31	-3	68	127				

Import dat z txt do xls

- Poté si vypočtené hodnoty překopírujeme na nový list (pravým tlačítkem vložit hodnoty), který si nějak můžeme pojmenovat
- Na první řádek si ještě vložíme řádek s popisem proměnných, tedy rok a pak jednotlivé měsíce (stačí číselně)
- Poté excel můžeme zavřít (samozřejmě si sešit nezapomeňte uložit jako sešit excelu, tedy xls nebo xlsx)

AS - Excel

Soubor Domů Vložení Rozložení stránky Vzorce Data Revize Zobrazení Řekněte mi, co chcete udělat...

Vymout Kopírovat Kopírovat formát Vložit

Calibri 11

Zalamovat text

Obecný

Schránka Písmo Zarovnání Číslo Podmíněné formátování Formátovat jako tabulku

A1 1961

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1961	-3,8	1,2	4,4	9,7	8,1	14,4	13,1	14,2	15,9	8,8	1,5	-3,5	
2	1962	-1,8	-3,1	-2,5	6,4	7,7	12,4	13,3	14,9	11,2	7,8	0,3	-6	
3	1963	-10,1	-7,6	0,5	6,9	10,5	14	16,3	14,4	12,5	6,5	4,9	-6,3	
4	1964	-5,1	-3,6	-2,3	6,6	12	16,4	17	14,1	12	5,2	2	-3	
5	1965	-1,9	-5,1	-0,9	4,6	9	13,8	13,3	13,6	11	7	-1,1	-0,1	
6	1966	-5,8	2,4	0,7	7,7	11,5	15,4	13,6	13,5	11,4	9,5	-0,5	-1,5	
7	1967	-3	-0,5	2,5	4,3	10,9	12,9	17,5	14,9	11,9	9,4	1,5	-2,9	
8	1968	-4,4	-2,1	2	7	9	14,3	14,5	14,2	11,2	8,2	0,4	-5	
9	1969	-3,3	-5,2	-1,6	4,7	12	12,9	17	14,3	12,2	9,2	2,6	-7,5	
10	1970	-5,8	-3,9	-1,5	3,1	8,7	15,4	14,2	15	11,7	6,5	3,2	-2,6	
11	1971	-3	-1,9	-1,4	7,1	12,1	11,6	16,9	16,7	10	7,2	0,5	0,6	
12	1972	-5,6	-1,3	3,7	4,5	9,2	13,4	16,3	14	9	4,7	1,3	-2,4	
13	1973	-2,7	-2,1	1,7	2,6	10,8	14,1	15,2	16,6	13,3	5	0,3	-3,4	
14	1974	-0,1	0	4,1	5,8	8,4	11,9	13	15,6	11,2	2,5	1,6	1,4	
15	1975	1	-0,7	1,7	4,5	10,6	12,5	16,4	16,7	14,4	6,2	-0,4	-2	
16	1976	-2	-2,4	-1,5	4,7	11,3	15,5	17,6	14,2	10,8	7,1	1,9	-4,1	
17	1977	-2,7	-0,1	4,1	3,2	10	13,8	14,4	14,2	9,3	8,1	2	-1,8	
18	1978	-2,8	-4,3	1,7	4,9	9,5	12,4	13,7	12,6	9,6	6,7	0,8	-1,8	
19	1979	-6,9	-3,6	1,3	4	11	15,5	13	13,9	10,9	5,7	0,4	1,3	
20	1980	-5,6	-0,2	0,8	3,1	8,2	12,5	12,6	14,5	11,8	5,3	0,1	-2,6	
21	1981	-4,8	-3,2	4,6	5,7	11	13,8	14	14,5	12,5	5,9	1,6	-4,9	
22	1982	-5,1	-3	2,2	4,1	10,7	14,3	17	15,3	14,8	7,6	3,1	-0,3	
23	1983	0,7	-4,6	1,8	6,9	9,7	14,6	18,7	16	11,6	6,8	0,7	-2,4	
24	1984	-2,5	-3,3	-0,3	4,6	8,7	11,5	13,8	14,9	9,9	8,2	2	-1,7	
25	1985	-7,6	-6,1	0,5	5,5	11,8	11,2	15,6	14,4	11,7	6,9	-1,7	1,1	
26	1986	-2,7	-8,9	0,8	4,7	13	13,6	15	14,6	9,5	7,4	3,1	-1,3	
27	1987	-8,7	-2,9	-3,9	7,1	7,9	11,9	15	12,7	13	6,7	1,9	-0,8	
28	1988	-0,1	-1,2	-0,5	6,4	12,4	12,9	15,2	15,2	11	7,6	-0,4	-0,1	
29	1989	-0,8	0,1	5,1	5,6	11,8	13,2	15,7	14,8	12,9	8,3	0	0,2	
30	1990	-1,7	2,9	4,6	5,1	12,3	13,6	14,8	17,1	9,3	7,7	1,8	-2,7	

AS aš

Vberte cíl a stiskněte klávesu F11 nebo zvolte příkaz Vložit.

Základní popisné statistiky

- Otevřeme si vytvořený excel sešit ve statistice, nezapomeneme si vybrat správný list, a zaškrtnout oba názvy, abychom měli tabulku hezky přehlednou
- Poté karta statistiky (nahore), základní statistiky, popisné statistiky a zaškrtnáme všechny popisné charakteristiky, které dle zadání potřebujeme, samozřejmě nezapomeneme vybrat proměnné, tedy všechny měsíce.

The screenshot shows the STATISTICA software interface. The main window displays a data table with the following structure:

Rok	1	2	3	4	5	6	7	8	9
1961	-5	0	3,4	8,7	7,2	14,1	12,5	13,6	14,4
1962	-2,9	-4,2	-3,6	6,2	7	11,4	12,5	14,1	10,1
1963	-10,1	-7,9	-0,3	5,7	9,8	13,5	16,4	14,1	11,9
1964	-5,9	-5	-3,8	5,9	11	15,6	16,5	13	11,1
1965	-2,8	-6,3	-1,8	3,9	7,9	12,8	13	13	10,5
1966	-7	1,5	-0,5	6,8	10,5	14,3	13,3	13	10,3
1967	-4,2	-1,5	1,6	3,6	10,1	12,4	16,4	14	11,6
1968	-5,1	-2,8	1,1	6,5	8,2	13,7	13,6	13,6	10,6
1969	-4,3	-5,6	-3	3,7	11,6	12,5	15,9	13,8	11,1
1970	-6,6	-5,1	-2,3	2,4	8	14,3	14	14,2	10,2
1971	-3,6	-2,8	-2,4	5,4	11,4	11	15,6	16,4	8,9
1972	-6,6	-2,2	2,4	3,8	9	12,8	16,1	13,1	8,5
1973	-3,3	-2,5	1,2	2	10,2	13	14,5	15,7	12,4
1974	-1,1	-0,8	3,1	5,2	8,5	11,3	12,5	15,3	10,7
1975	0,1	-1,6	1	3,6	9,5	12,1	15,8	16,2	14
1976	-3,5	-3,1	-2,9	4,3	10,1	14,5	16,8	13,3	9,7
1977	-3,7	-0,8	3,1	2,5	9,5	13,3	13,4	13,5	8,4
1978	-3,5	-5,1	1,2	4,2	8,8	12,2	12,7	12,4	8,7
1979	-7,6	-3,9	0,6	3,3	10,6	15,3	11,6	13,4	10,7
1980	-6,7	-1,5	-0,9	2,2	7,4	11,9	12,1	13,7	10,8
1981	-4,6	-3,4	3,4	4,2	10,6	13,2	13,4	13,8	11,6
1982	-5,3	-4,2	1,3	2,7	10,4	13,8	16,4	15,4	14,4
1983	-0,1	-5,5	1	6,2	9,9	13,8	18,4	15,6	11,4
1984	-3,6	-4,4	-1,5	3,8	8,5	10,7	12,3	14,5	9,2
1985	-8,7	-7,2	-0,7	4,7	11	10,2	14,5	14,2	10,7

The 'Popisné statistiky: MILESOVK' dialog box is open, showing the 'Výpočet' (Calculation) tab. The 'Výpočet statistik' (Calculate statistics) section is checked, and the 'Všechny statistiky' (All statistics) button is visible. The 'Kvantily, rozpětí' (Quantiles, range) section is also checked, with 'První' (First) set to 10,00 and 'Druhá' (Second) set to 90,00. The 'Výpočet' button is highlighted.

Základní popisné statistiky

- Výslednou tabulku si zkopírujeme se záhlavími do excelu

The screenshot shows the STATISTICA software interface. A dialog box titled "PS4* - Popisné statistiky (AS)" is open, displaying a table of descriptive statistics for 12 categories (I to XII). The table includes columns for Mean, Median, Mode, Modality, Minimum, Maximum, Lower Quartile, Upper Quartile, Decile, and Skewness. A context menu is open over the table, with the option "Kopírovat se záhlavími" (Copy with headers) highlighted.

Proměnná	Průměr	Medián	Modus	Četnost modu	Minimum	Maximum	Dolní kvartil	Horní kvartil	Kvantil 10,00000	Kvan
I	-3,60667	-3,00000	-2,70000	3	-10,1000	1,00000	-5,60000	-1,90000	-7,25000	-0,10
II	-2,47667	-2,65000	Vícenás.	2	-8,9000	2,90000	-3,90000	-0,50000	-5,65000	0,65
III	1,08000	1,05000	1,70000	3	-3,9000	5,10000	-0,90000	2,50000	-1,95000	4,50
IV	5,37000	5,00000	4,70000	3	2,6000	9,70000	4,50000	6,60000	3,15000	7,10
V	10,32667	10,65000	Vícenás.	2	7,7000	13,00000	9,00000	11,80000	8,15000	12,20
VI	13,52333	13,60000	Vícenás.	3	11,2000	16,40000	12,50000	14,30000	11,75000	15,40
VII	15,12333	15,00000	17,00000	3	12,6000	18,70000	13,70000	16,40000	13,05000	17,20
VIII	14,72000	14,50000	14,20000	4	12,6000	17,10000	14,20000	15,80000	14,60000	16,60
IX	11,58333	11,50000	11,20000	3	9,0000	15,90000	10,80000	13,80000	11,80000	13,80
X	6,99000	7,05000	Vícenás.	2	2,5000	9,50000	6,40000	8,10000	7,25000	9,00
XI	1,18000	1,40000	2,00000	3	-1,7000	4,90000	0,30000	2,50000	1,30000	3,10
XII	-2,20333	-2,20000	Vícenás.	2	-7,5000	1,40000	-3,40000	-1,10000	-2,25000	0,80

Základní popisné statistiky

- Tabulku si zkopírujeme do excelu, celou jí označíme, opět zmáčkne ctrl c a pravým tlačítkem dáme vložit a transponovat, viz obrázek
- A poté, stejně jako v minulém cvičení, tabulku zkrášlíme a upravíme dle zadání

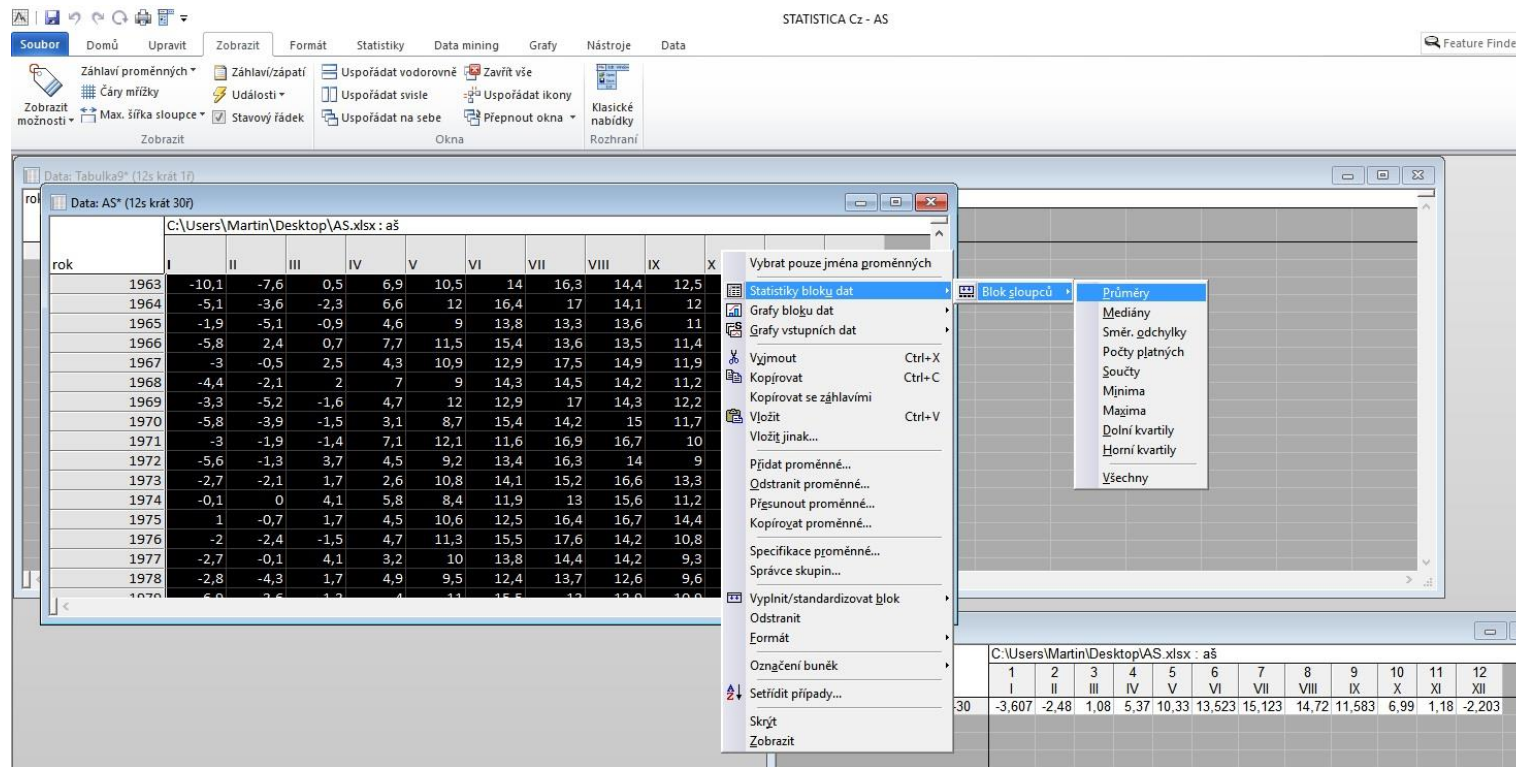
The screenshot shows the Microsoft Excel interface with a data table. A context menu is open over the table, and the 'Vložit jinak...' (Paste Special) option is highlighted with a red box. The 'Vložit jinak...' submenu is also open, showing the 'Vložit jako tabulku' (Paste as Table) option, which is used to transposing the data.

	F	G	H	I	J	K	L	M	N	O	
4	0000	3	-10,1000	1,00000	-5,60000	-1,90000	-7,25000	-0,10000	2,697244	-0,436035	-0,02
5	inás.	2	-8,9000	2,90000	-3,90000	-0,50000	-5,65000	0,65000	2,690172	-0,212732	0,27
6	0000	3	-3,9000	5,10000	-0,90000	2,50000	1,95000	4,50000	2,406084	-0,062783	-0,77
7	0000	3	2,6000	9,70000	4,50000	6,60000	3,15000	7,10000	1,588352	0,515303	0,43
8	inás.	2	7,7000	13,00000	9,00000	11,80000	8,15000	12,20000	1,550291	-0,150823	-1,23
9	inás.	3	11,2000	16,40000	12,50000	14,30000	11,75000	15,45000	1,331756	0,210514	-0,55
10	0000	3	12,6000	18,70000	13,70000	16,40000	13,05000	17,25000	1,646874	0,316210	-0,90
11	0000	4	12,6000	17,10000	14,20000	15,20000	13,55000	16,65000	1,093366	0,449312	0,24
12	0000	3	9,0000	15,90000	10,80000	12,50000	9,40000	13,85000	1,650096	0,658980	0,57
13	inás.	2	2,5000	9,50000	6,20000	8,10000	5,10000	9,00000	1,529785	-0,723226	1,23
14	0000	3	-1,7000	4,90000	0,30000	2,00000	-0,45000	3,10000	1,423788	0,326792	0,47
15	inás.	2	-7,5000	1,40000	-3,40000	-0,30000	-5,50000	0,85000	2,266622	-0,425208	-0,12

	III	IV	V	VI	VII	VIII	IX	X	XI	XII			
19	1,08000	5,37000	10,32667	13,52333	15,12333	14,72000	11,58333	6,99000	1,18000	-2,20			
20	1,05000	5,00000	10,65000	13,60000	15,00000	14,50000	11,50000	7,05000	1,40000	-2,20			
21	1,700000	4,700000	Vícenás.	Vícenás.	17,00000	14,20000	11,20000	Vícenás.	2,000000	Více			
22	Cetnost (modu)	3	2	3	3	2	3	3	4	3	2	3	
23	Minimum	-10,1000	-8,9000	-3,9000	2,6000	7,7000	11,2000	12,6000	12,6000	9,0000	2,5000	-1,7000	-7,5
24	Maximum	1,00000	2,90000	5,10000	9,70000	13,00000	16,40000	18,70000	17,10000	15,90000	9,50000	4,90000	1,40
25	Dolní (kvartil)	-5,60000	-3,90000	-0,90000	4,50000	9,00000	12,50000	13,70000	14,20000	10,80000	6,20000	0,30000	-3,40
	Horní	1,00000	2,90000	5,10000	9,70000	13,00000	16,40000	18,70000	17,10000	15,90000	9,50000	4,90000	1,40

Grafické znázornění ročního chodu průměrné teploty vzduchu

- Stejně jako ve cvičení 1 si označíme všechny proměnné, klikneme pravým tlačítkem a dáme možnost viz obrázek
- Tím dostaneme tabulku průměrů



The screenshot shows the STATISTICA software interface. The main window displays a data table with the following structure:

rok	I	II	III	IV	V	VI	VII	VIII	IX	X
1963	-10,1	-7,6	0,5	6,9	10,5	14	16,3	14,4	12,5	
1964	-5,1	-3,6	-2,3	6,6	12	16,4	17	14,1	12	
1965	-1,9	-5,1	-0,9	4,6	9	13,8	13,3	13,6	11	
1966	-5,8	2,4	0,7	7,7	11,5	15,4	13,6	13,5	11,4	
1967	-3	-0,5	2,5	4,3	10,9	12,9	17,5	14,9	11,9	
1968	-4,4	-2,1	2	7	9	14,3	14,5	14,2	11,2	
1969	-3,3	-5,2	-1,6	4,7	12	12,9	17	14,3	12,2	
1970	-5,8	-3,9	-1,5	3,1	8,7	15,4	14,2	15	11,7	
1971	-3	-1,9	-1,4	7,1	12,1	11,6	16,9	16,7	10	
1972	-5,6	-1,3	3,7	4,5	9,2	13,4	16,3	14	9	
1973	-2,7	-2,1	1,7	2,6	10,8	14,1	15,2	16,6	13,3	
1974	-0,1	0	4,1	5,8	8,4	11,9	13	15,6	11,2	
1975	1	-0,7	1,7	4,5	10,6	12,5	16,4	16,7	14,4	
1976	-2	-2,4	-1,5	4,7	11,3	15,5	17,6	14,2	10,8	
1977	-2,7	-0,1	4,1	3,2	10	13,8	14,4	14,2	9,3	
1978	-2,8	-4,3	1,7	4,9	9,5	12,4	13,7	12,6	9,6	
1979	-5,0	-2,5	1,2	4	11	15,5	12	12,0	10,0	

A context menu is open over the table, showing options for data analysis. The 'Průměry' (Averages) option is selected, which has opened a sub-menu with the following options:

- Průměry
- Mediány
- Směr. odchylky
- Počty platných
- Šoučty
- Mjníma
- Magima
- Dolní kvartily
- Horní kvartily
- Všechny

Grafické znázornění ročního chodu průměrné teploty vzduchu

- Graf vytvoříme stejně jako v předchozích cvičeních (obrázek...)
- Graf poté samozřejmě upravíme do krásné odevzdávatelné podoby jako v předchozích cvičeních

The screenshot shows the STATISTICA software interface. The main window displays a data table with columns labeled 'I' and 'II' and rows for years from 1963 to 1978. A context menu is open over a selected data block, showing options for creating various types of graphs. The option 'Spojnicový graf: blok řádků' (Line graph: block of rows) is highlighted with a red box.

rok	I	II
1963	-10,1	-7,6
1964	-5,1	-3,6
1965	-1,9	-5,1
1966	-5,8	2,4
1967	-3	-0,5
1968	-4,4	-2,1
1969	-3,3	-5,2
1970	-5,8	-3,9
1971	-3	-1,9
1972	-5,6	-1,3
1973	-2,7	-2,1
1974	-0,1	0
1975	1	-0,7
1976	-2	-2,4
1977	-2,7	-0,1
1978	-2,8	-4,3

Context menu options:

- Statistiky bloku dat
- Grafy bloku dat
- Grafy vstupních dat
- Vymout (Ctrl+X)
- Kopírovat (Ctrl+C)
- Kopírovat se záhlavími
- Vložit (Ctrl+V)
- Vložit jinak...
- Vložit se záhlavím
- Vyplnit/standardizovat blok
- Smazat
- Formát
- Označení buněk

Graph creation options:

- Histogram: blok sloupců
- Histogram: celé sloupce
- Spojnicový graf: celé sloupce
- Spojnicový graf: blok řádků**
- Krabicový graf: blok sloupců
- Normální pravděpodobnostní graf: blok sloupců
- Vlastní graf bloku podle sloupce
- Vlastní graf bloku podle řádku
- Vlastní graf celého sloupce
- Vlastní graf celého řádku
- Vlastní seznam...