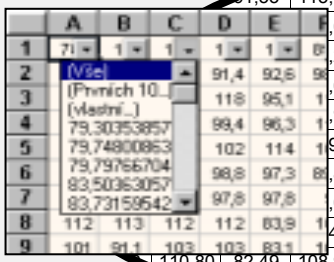
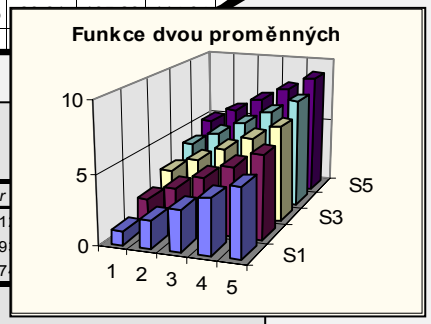
KURZ PRAKTICKÉHO POUŽITÍ PROGRAMU *Microsoft* ® EXCEL PRO ZPRACOVÁNÍ A ANALÝZU DAT

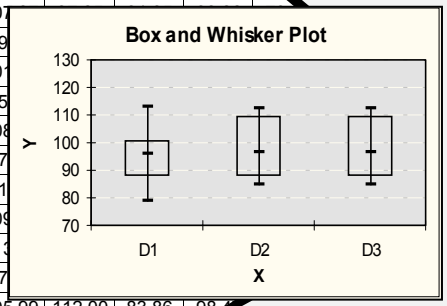
Příručka





F-test

	Soubor 1	Soubor 2
stř. hodnota	99,4911173	97,6875634
rozptyl	122,617713	119,575281
pozorování	39	40
rozdíl	38	39
F	1,02544365	
P(F<=f) (1)	0,46857906	
F krit (1)	1,70873449	



Korelace

	sloupec 1	sloupec 2
sloupec 1	1	
sloupec 2	0,3401632	1

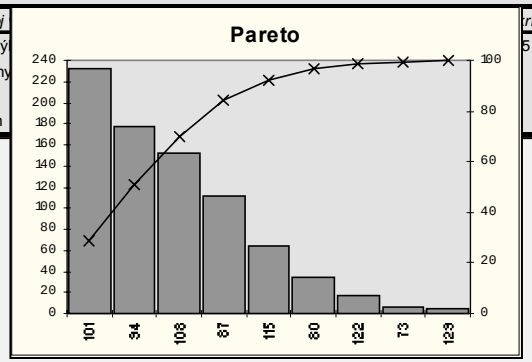
Anova: jeden faktor

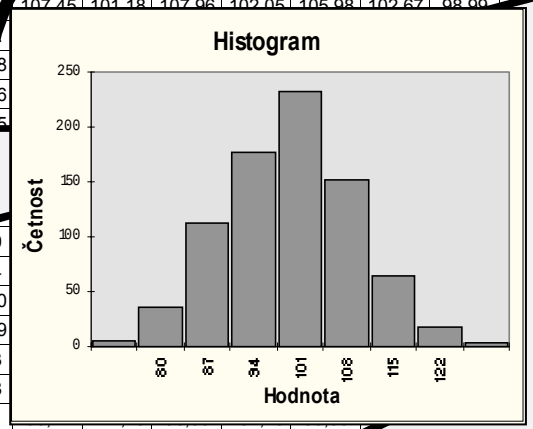
faktor

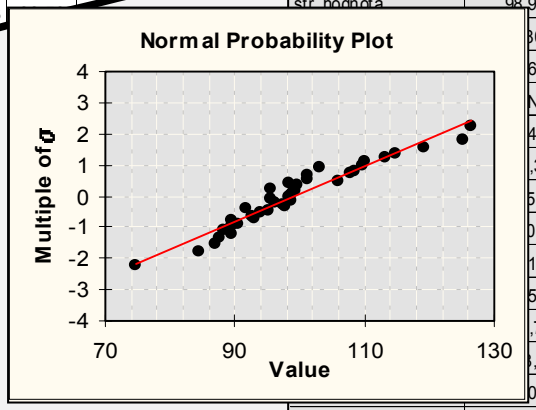
výběr	počet	součet	průměr
sloupec 1	39	3880,154	99,4911
sloupec 2	39	3795,477	97,3199
sloupec 3	39	3882,206	99,5437

ANOVA

Zdroj	
mezi vý	
všechny	
celkem	







Popisná statistika

stř. hodnota	98,9687
	3055
	6059
	N/A
	4188
	,3900
	5087
	0137
	1525
	5949
	,7474
	,7484
	0000
věrohodnost (95,0%)	3,6519

UFUnkce užitečné 2.2

VPS s.r.o.

Úvod

Microsoft, Visual Basic a Windows jsou registrované ochranné známky společnosti Microsoft Corporation.

Tato příručka, a soubory na CD "Štatistické kurzy" jsou určeny pouze pro účastníky kurzu. Jiné použití je možné po dohodě s firmou VPS s.r.o. .

Příručka byla připravena pro potřeby kurzu praktického použití programu Microsoft® EXCEL pro zpracování a analýzu dat. Příručka není učebnicí ani uživatelským manuálem programu *Microsoft*® EXCEL. Kurz je koncipován pro EXCEL 2002. Většina zpracovaných témat je platná také pro nižší verze EXCELu.

V příručce jsou stručně a většinou názorným způsobem popsaná jednotlivá témata probírána v kurzu, má tedy sloužit účastníkům kurzu pro rychlé připomenutí postupů použití některých nástrojů EXCELu, nebo interpretace jejich výsledků.

Každé téma je zpracováno na jednom listu. V patě listu jsou informační údaje. Je zde odkaz na vzorová data a také na soubor uživatelských funkcí, pokud si jej popisované téma vyžaduje. Pata listu **Úvod** slouží pouze k vysvětlení významu informací.

Probírána témata jsou rozdělena do tří typů:

- informace, vysvětlení, teorie, nastavení
- standardní nástroje EXCELu
- nástroje vyžadující instalaci souboru *.XLA nebo speciální přípravu

V úvodě každého listu je navozen problém, který se týká vzorových dat a pak nástin řešení. Obvykle v levé části listu je popsán postup použití příslušného nástroje a v pravé části další možnosti použití, anebo interpretace výsledků (z prostorových důvodů může být toto rozdělení na některých listech jiné).

Použití symbolů:

(1) (2) pořadová čísla kroků postupu

! důležitá upozornění

Použití písma:

Arial	Text, vzorce
Arial, Bold	Názvy nástrojů, nabídek, dialog. oken, funkcí a pod.
`Courier New`	Hodnoty v textu

Tato příručka byla připravena pomocí programu *Microsoft*® EXCEL 2002 pod operačním systémem WINDOWS 98. !

Pata tohoto listu slouží jako příklad !

Značka pro typ listu / tématu

Zde je uveden název souboru, pokud popsaný nástroj vyžaduje jeho instalaci.

Zde je uvedeno číslo listu, kde je návod jako popsaný nástroj připravit (pokud není standardním nástrojem).

Název listu / tématu

V případě, že je pro nástroj / téma běžně používaný anglický název, je uveden zde.

Název souboru se vzorovými daty, v závorkách je název příslušného listu.

Čísla listů se souvisejícími tématy

Číslo listu / tématu

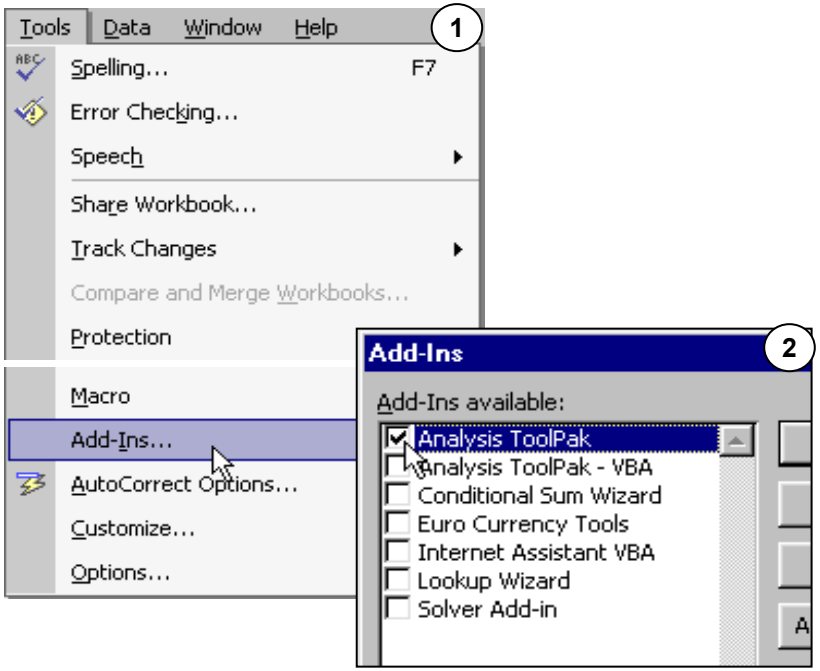
Užitečná nastavení

Problém:
Zejména na pomalejších počítačích mohou některé akce EXCELu probíhat neúnosně dlouho.

Řešení:
Rychlost startu EXCELu se dá výrazně ovlivnit počtem otevřených doplňků.
Při použití velkého množství funkcí nebo rozsáhlých vstupních dat je vhodné nastavit manuální ovládání výpočtů.

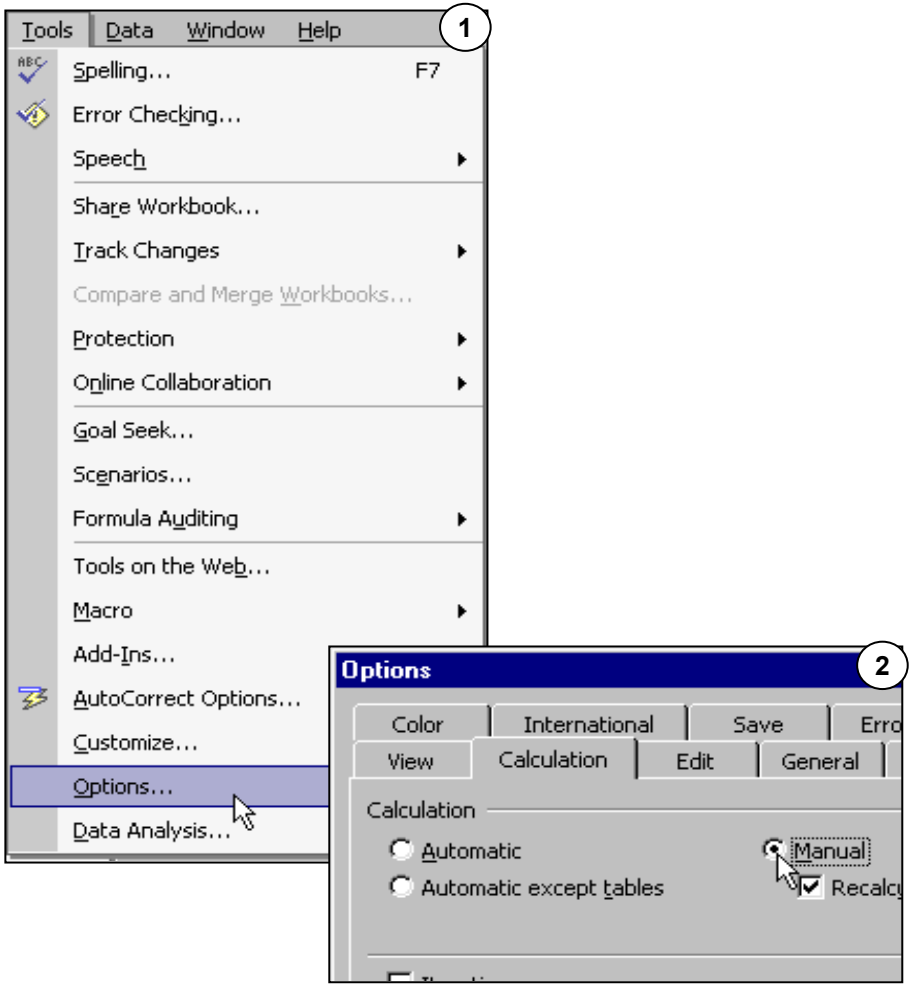
Otevření (zavření) doplnků (Add-Ins)

Přes nabídku **Tools - Add-Ins** otevřete dialogové okno **Add-Ins**. V okénkách zaškrtněte jen ty doplňky, které opravdu potřebujete k práci. Zbytečně otevřené doplňky zpomalují start EXCELu a také zabírají operační paměť počítače. Toto nastavení můžete kdykoli změnit.



Přepočet automaticky/ručně

Přes nabídku **Tools - Options** otevřete dialogové okno **Options - Calculation** a označte **Calculation - Manual**. Přepočet listu proběhne až po stisknutí klávesy **F9**. Nastavení se ukládá se sešitem. Po startu nového sešitu EXCELu je vždy nastaveno **Calculation - Automatic**.



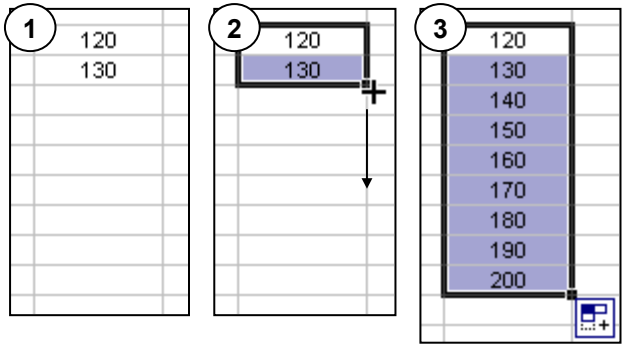
Automatické řady a seznamy - přetažením v listu

	A	B	C	D
1				
2	120	pondělí		
3	130	úterý		
4	140	středa		
5	150	čtvrtek		
6	160	pátek		
7	170	sobota		
8	180	neděle		
9	190			
10	200			
11				
12				
13				

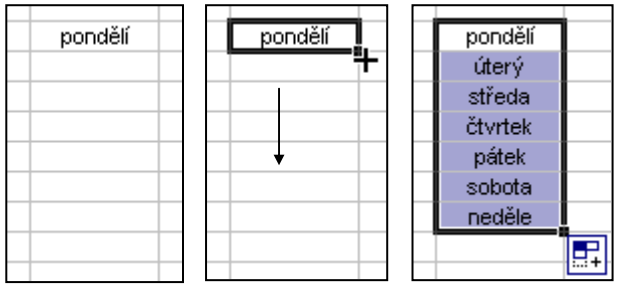
Problém:
Časté vytváření různých seznamů bývá zdlouhavé.

Řešení:
EXCEL umí automaticky vytvářet seznamy podle prvního členu, nebo aritmetickou řadu podle prvních dvou členů.

Vytvoření automatického seznamu v listu

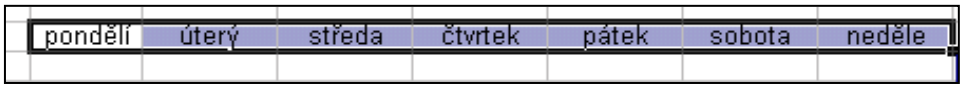


Napište první dva členy aritmetické řady a označte je. Pak umístěte kurzor myši na značku v pravém dolním rohu označené plochy, až se objeví křížek. Stiskněte levé tlačítko myši a táhněte směrem dolů. Po uvolnění tlačítka myši bude označena oblast vyplněna členy aritmetické řady.



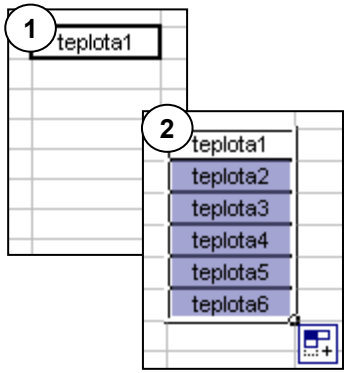
Podobně vytvoříte i jiné řady například názvy dnů týdne, nebo názvy měsíců. V těchto případech stačí napsat první člen řady. Pokud se seznam nevytvoří, doplňte jej do rejstříku seznamů (list 4).

Podobně můžete vytvářet řady a seznamy také v řádcích

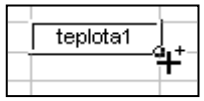


Další možnosti

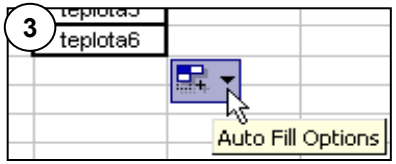
EXCEL se při těchto operacích chová inteligentně. Z hodnoty `teplota1` dostanete řadu `teplota1`, `teplota2`, ... Příklady jsou na obrázcích.

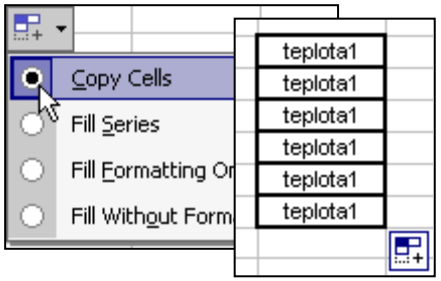


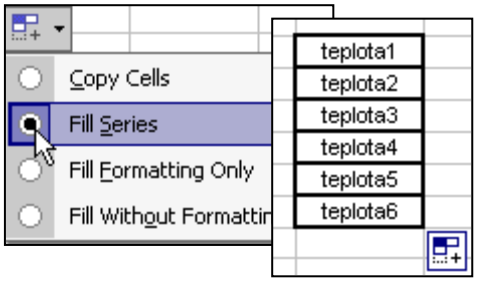
Pokud nechcete vytvářet automatický seznam, ale pouze kopírovat, podržte při tažení myší stisknuté tlačítko **Ctrl** (objeví se malé znaménko +).

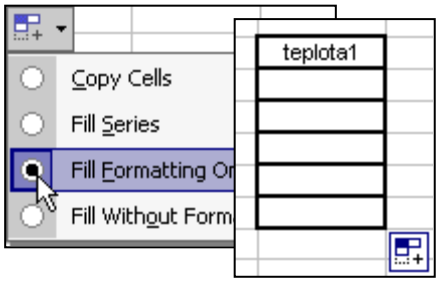


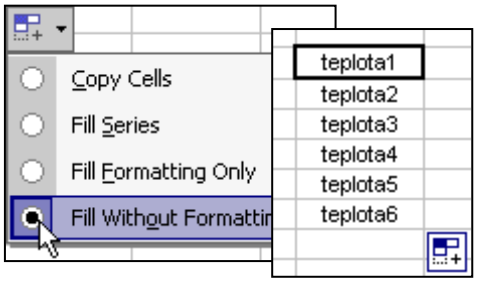
Ikona v dolní části řady nabízí menu s dalšími možnostmi pro vytvořenou řadu.











Přidání seznamu do rejstříku seznamů

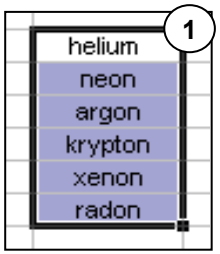
	A	B	C
1			
2		helium	
3		neon	
4		argon	
5		krypton	
6		xenon	
7		radon	
8			
9			
10			
11			
12			
13			

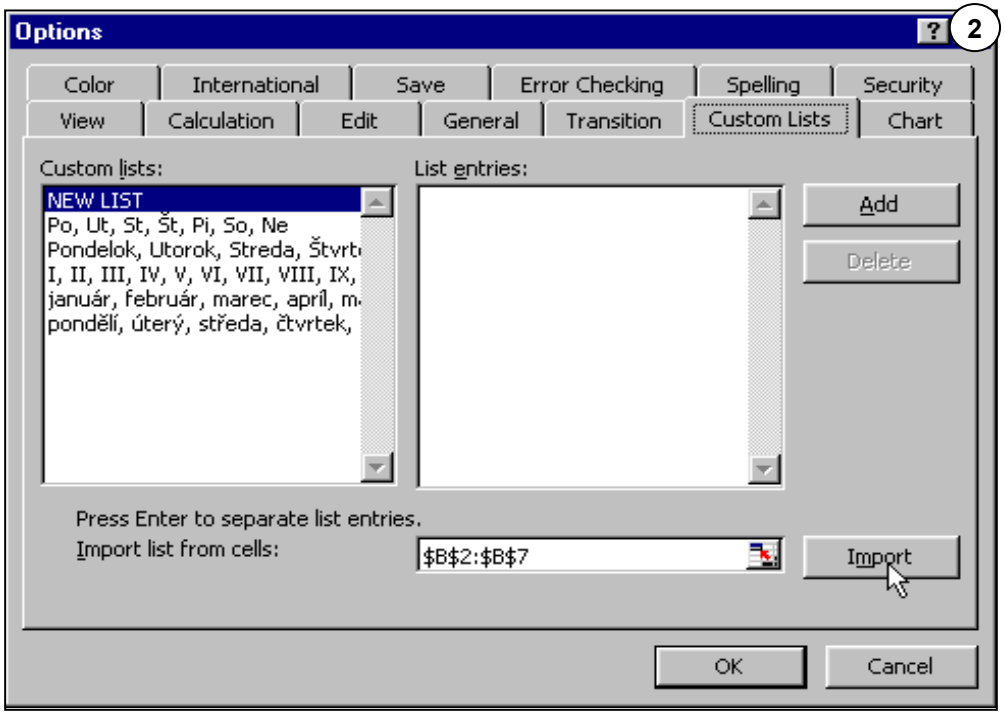
Problém:
Časté vytváření různých seznamů které nejsou v EXCELu standardní.

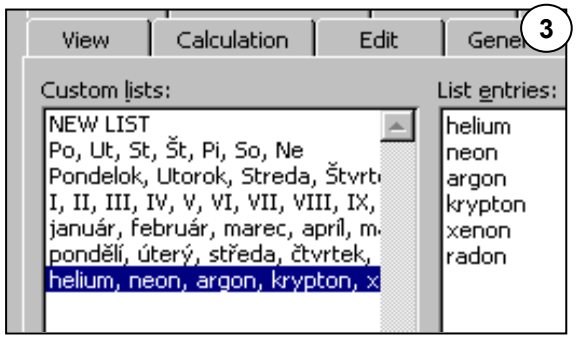
Řešení:
Přidejte potřebné seznamy do rejstříku seznamů.

Postup

Napište seznam na libovolné místo v listu a označte jej. Pak přes nabídku **Tools - Options** otevřete dialogové okno **Options**. V něm zvolte záložku **Custom Lists**. Kliknutím na tlačítko **Import** se seznam z listu přenese do rejstříku seznamů.







Nový seznam nyní můžete používat standardním způsobem.

Automatické řady a seznamy - nástrojem Series

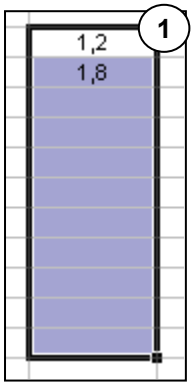
	A	B	C	D
1				
2	1,2		12.9.02	
3	1,8		13.9.02	
4	2,7		14.9.02	
5	4,05		15.9.02	
6	6,075		16.9.02	
7	9,1125		17.9.02	
8	13,6688		18.9.02	
9	20,5031		19.9.02	
10	30,7547		20.9.02	
11	46,132		21.9.02	
12	69,198		22.9.02	
13				

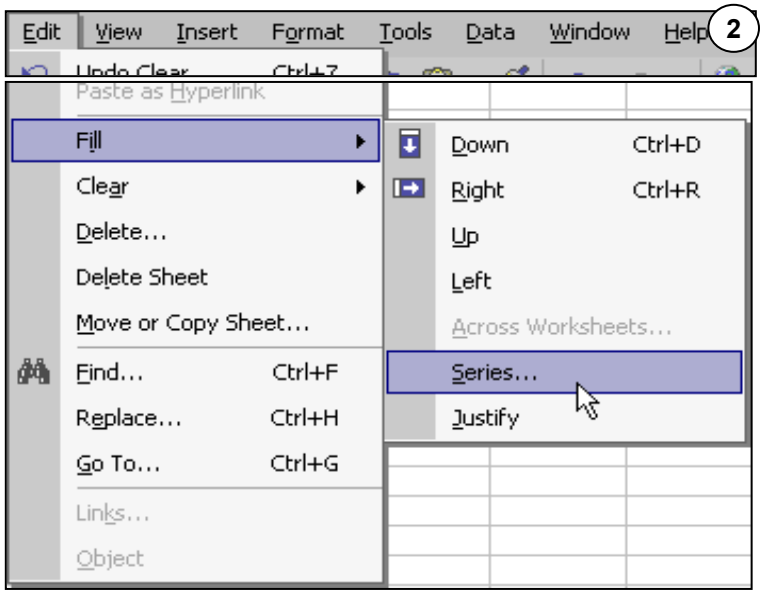
Problém:
Vytvořit geometrickou řadu s prvními dvěma členy `1,2` a `1,8`.
Vytvořit seznam pracovních dnů od `12.9.02`

Řešení:
Pomocí nástroje **Series** můžete vytvářet aritmetické, geometrické a kalendářní řady.

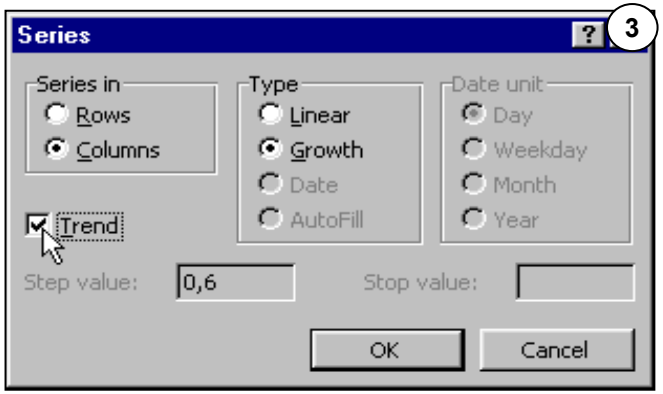
Vytvoření geometrické řady

Napište první dva členy řady, pak označte oblast, která má být vyplněná, včetně prvních členů. Přes nabídku **Edit - Fill - Series** otevřete dialogové okno **Series**.



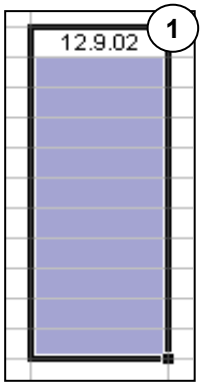


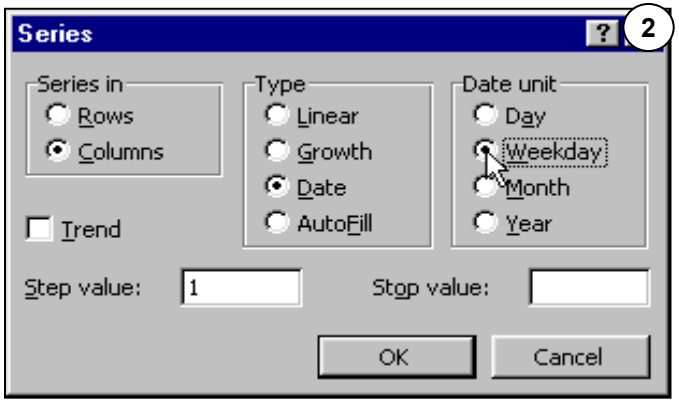
Zvolte **Type - Growth** a zaškrtněte **Trend** (v tom případě parametr **Step value** nemá vliv).



Vytvoření kalendářní řady

Napište první datum řady, pak označte oblast, která má být vyplněná, včetně prvního data. Přes nabídku **Edit - Fill - Series** otevřete dialogové okno **Series**. Zvolte **Type** - **Date**, **Date unit - Weekday**, **Step value** 1.





Pozor, **Date unit** `Weekday` v tomto případě znamená vyplnění oblasti pouze pracovními dny, ne krok jeden týden. Ostatní jednotky znamenají velikost kroku. !

Pozor, při tomto postupu vytváření řad se zkopírují také formáty prvních buněk. !

Databáze a tabulky

	A	B	C	D	E	F	G	H	I
1	**Reaktor A**				**Databáze**				
2	**DATUM**	**TYP**	**R**		**DATUM**	**TYP**	**R**	**REAKT**	
3	12.8.02	MC78L	3,40		12.8.02	MC78L	3,40	A	
4	13.8.02	MC78L	3,35		13.8.02	MC78L	3,35	A	
5	13.8.02	MA78	3,25		13.8.02	MA78	3,25	A	
6	15.8.02	MA78	3,00		15.8.02	MA78	3,00	A	
7	16.8.02	MA78	3,15		16.8.02	MA78	3,15	A	
8	16.8.02	TL431	3,30		16.8.02	TL431	3,30	A	
9	18.8.02	TL431	3,30		17.8.02	TL431	2,80	B	
10	19.8.02	MA78	2,80		18.8.02	TL431	3,30	A	
11	19.8.02	MA78	3,40		18.8.02	TL431	2,90	B	
12					18.8.02	TL431	3,20	B	
13	**Reaktor B**				18.8.02	TL431	3,55	B	
14	**DATUM**	**TYP**	**R**		18.8.02	MC78M	3,20	B	
15	17.8.02	TL431	2,80		18.8.02	MC78M	3,35	B	
16	18.8.02	TL431	2,90		18.8.02	MC78M	3,55	B	
17	18.8.02	TL431	3,20		18.8.02	MC78M	3,50	B	
18	18.8.02	TL431	3,55		19.8.02	MA78	2,80	A	
19	18.8.02	MC78M	3,20						
20	18.8.02	MC78M	3,35						
21	18.8.02	MC78M	3,55						
22	18.8.02	MC78M	3,50						
23									

Problém:
Uspořádat data co nejuniverzálnějším způsobem.

Řešení:
Vytvořit databázi.

Na obrázku, ve sloupcích **A**, **B**, **C**, jsou dvě tabulky. Jedna pro reaktor A druhá pro reaktor B. Hodnoty z tabulek můžete použít pro vyhodnocení reaktoru A nebo B.

Pro komplexní analýzu je výhodnější uspořádání dat tak, jako ve sloupcích **E**, **F**, **G**, **H**. Zde jsou stejná data, ale je přidaný další sloupec REAKT, ve kterém je uvedeno zda data pocházejí z reaktoru A anebo B. Z takto uspořádaných dat je možné získat ty samé výsledky, jako z předchozích tabulek, navíc, je možné získat další informace.

Ve sloupcích **E**, **F**, **G**, **H** je vytvořena DATABÁZE (DATBASE).

Vytváření databází je výhodné, protože z databáze lehko vytvoříte libovolné tabulky (naopak je to obtížné). Z databází je možné získat i další informace, které v separovaných tabulkách zanikají. Pokud tedy chcete řešit specifický problém a data jsou získána pouze pro tento účel, vytvářejte specifické tabulky. Ale pokud chcete data analyzovat, vytvořte databázi.

Pravidla

Existují jisté nesrovnanosti v pojmech TABULKA (TABLE), SEZNAM (LIST) a DATABÁZE (DATABASE). V nápovědě EXCELu nejsou někdy rozlišeny. V této příručce budeme používat pojmy TABULKA a DATABÁZE v tom smyslu, jako byly uvedeny zde, SEZNAM pak bude mít význam podle listů **Automatické řady a seznamy - ...** . !

	A	B	C	D
1	**DATUM**	**TYP**	**R**	**REAKT**
2	12.8.02	MC78L	3,40	A
3	13.8.02	MC78L	3,35	A
4	13.8.02	MA78	3,25	A
5	15.8.02	MA78	3,00	A
6	16.8.02	MA78	3,15	A
7	16.8.02	TL431	3,30	A
8	17.8.02	TL431	2,80	B
9	18.8.02	TL431	3,30	A
10	18.8.02	TL431	2,90	B
11	18.8.02	TL431	3,20	B
12	18.8.02	TL431	3,55	B
13				

Záhlaví (header) databáze

Záznamy (records) databáze

sloupec - Pole (field) databáze

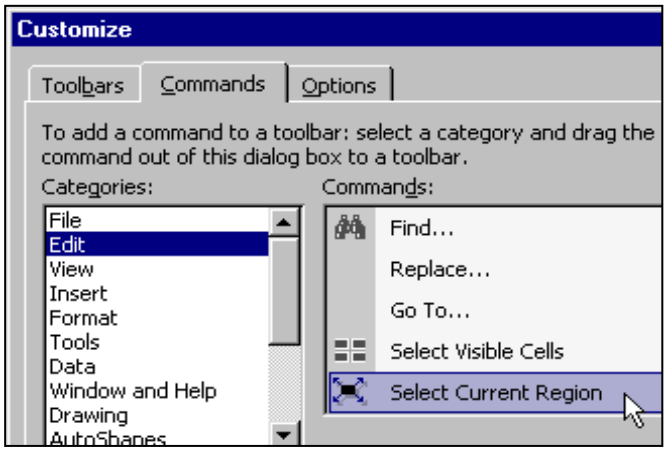
Databázi vytvářejte vždy ve svislém směru. Tedy nahoře záhlaví, pod ním záznamy.

Pokud máte na jednom listu více databází, umístěte je pod sebe.

Dbejte na to, aby databáze a tabulky byly vždy ohraničeny prázdnými buňkami.

Používejte záhlaví.

Při dodržení těchto pravidel vám budou dobře sloužit nástroje EXCELu určené na zpracování databází, jako jsou **Sort**, **Filter**, **Form**, **Subtotal, Pivot Table** atd. .
Pro práci s databázemi a tabulkami je užitečné přidat do panelu nástrojů ikonu nástroje Výběr oblasti. Pomoci této ikony vyberete rychle celou oblast tabulky nebo databáze.



Přes nabídku **View - Toolbars** - z nabítky vyberte **Customize**. Získáte dialogové okno **Customize**. Označte záložku **Commands**. Zvolte **Edit** a pak z plochy **Commands** přesuňte ikonu **Select Current Region** na panel nástrojů.

Automatický filtr (AutoFilter)

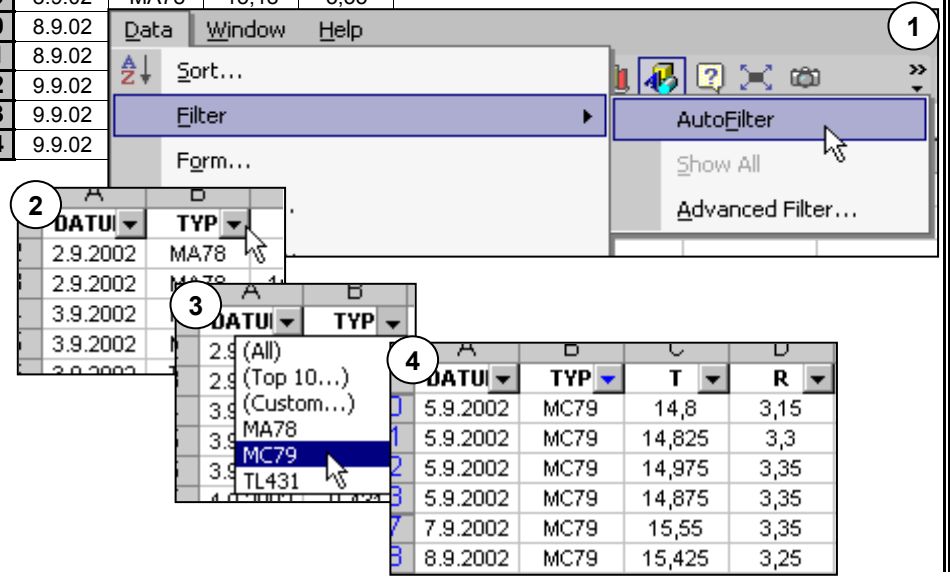
	A		C	D
1	DATUM	TYP	T	R
2	2.9.02	MA78	15,00	3,60
3	2.9.02	MA78	15,33	3,25
4	3.9.02	MA78	14,60	3,30
5	3.9.02	MA78	14,88	3,45
6	3.9.02	TL431	14,98	3,30
7	4.9.02	TL431	14,90	3,30
8	4.9.02	MA78	15,08	3,25
9	4.9.02	MA78	14,90	3,35
10	5.9.02	MC79	14,80	3,15
11	5.9.02	MC79	14,83	3,30
12	5.9.02	MC79	14,98	3,35
13	5.9.02	MC79	14,88	3,35
14	6.9.02	MA78	15,10	3,55
15	6.9.02	MA78	15,03	3,35
16	7.9.02	TL431	14,65	2,90
17	7.9.02	MC79	15,55	3,35
18	8.9.02	MC79	15,43	3,25
19	8.9.02	MA78	15,18	3,35
20	8.9.02			
21	8.9.02			
22	9.9.02			
23	9.9.02			
24	9.9.02			

Problém:
Z databáze vybrat pouze řádky s typem MC79.

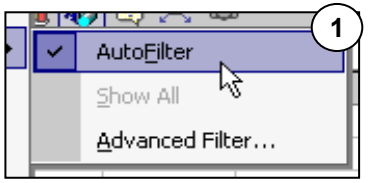
Řešení:
Použití automatického filtru umožňuje rychle skrýt řádky nevyhovující zadané podmínce.

Spuštění automatického filtru

Označte libovolnou buňku uvnitř databáze. Pak přes nabídku **Data - Filter - AutoFilter** zapněte režim filtrování.
V nejvyšším řádku databáze, nebo v záhlaví se objeví tlačítka. Klikněte na tlačítko ve sloupci Typ. Otevře se nabídka pro podmínky filtrace. Zvolte hodnotu MC79. Filtr skryje všechny řádky, ve kterých hodnoty buněk nevyhovují podmínce. Podobně můžete přidávat podmínky do dalších sloupců.

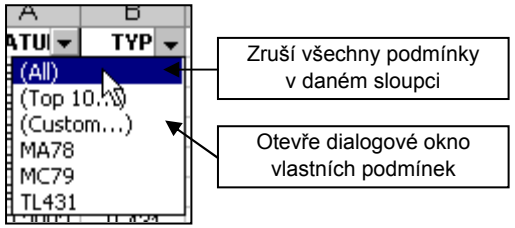


Další možnosti

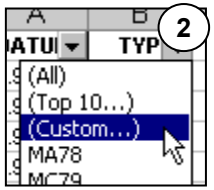


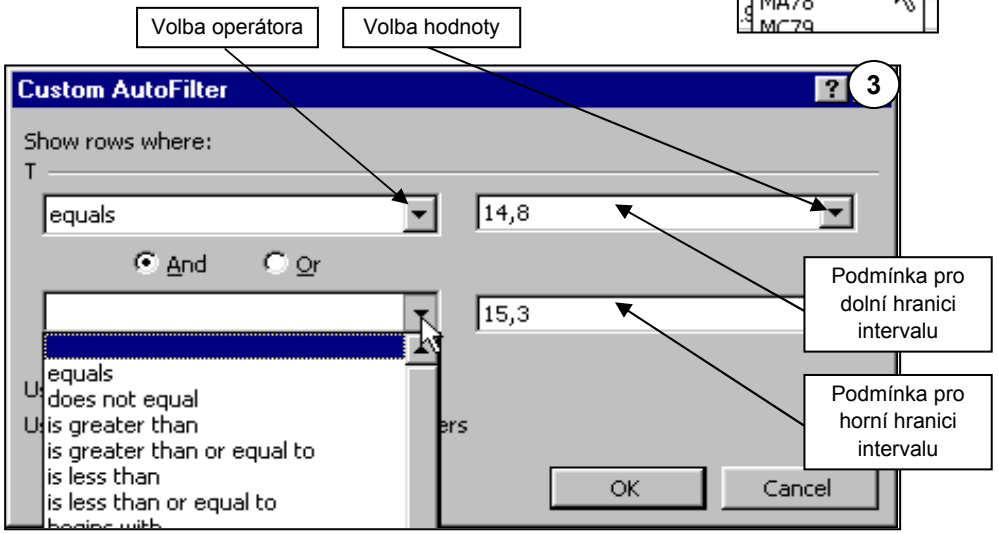
Automatický filtr můžete zrušit opět přes nabídku **Data - Filter - Autofilter**.

Volba položky **Show All** zruší všechny podmínky filtrace, ale nezruší filtr.



Pokud chcete zobrazit data v určitém intervalu, nastavte podmínky filtrace v dialogovém okně **Custom...**.





Pokud před spuštěním filtru označíte nějakou oblast databáze (více jako jednu buňku), bude filtr působit pouze na řádky označené oblasti a níže. !

Automatický filtr skryje celé řádky, tedy i buňky mimo oblast databáze. !

Rozšířený filtr (Advanced Filter)

	A	B	C	D	E
1	TYP	Odpor			
2	MC79	=(C5+D5)/2>3,3			
3					
4	DATUM	TYP	R1	R2	
5	2.9.02	MA78	3,40	3,80	
6	2.9.02	MA78	3,30	3,20	
7	3.9.02	MA78	3,40	3,20	
8	3.9.02	MA78	3,50	3,40	
9	3.9.02	TL431	3,30	3,30	
10	4.9.02	TL431	3,40	3,20	
11	4.9.02	MA78	3,20	3,30	
12	4.9.02	MA78	3,40	3,30	
13	5.9.02	MC79	3,30	3,00	
14	5.9.02	MC79	3,40	3,20	
15	5.9.02	MC79	3,50	3,20	
16	5.9.02	MC79	3,50	3,20	
17	6.9.02	MA78	3,60	3,50	
18	6.9.02	MA78	3,40	3,30	
19	7.9.02	TL431	3,00	2,80	
20	7.9.02	MC79	3,60	3,10	
21	8.9.02	MC79	3,10	3,40	
22	8.9.02	MA78	3,50	3,20	
23	9.9.02	MA78	3,70	3,30	
24	9.9.02		3,40	3,30	

Při použití rozšířeného filtru je vhodné ponechat nad databází několik volných řádků (3 a více) pro umístění podmínek filtrace. !

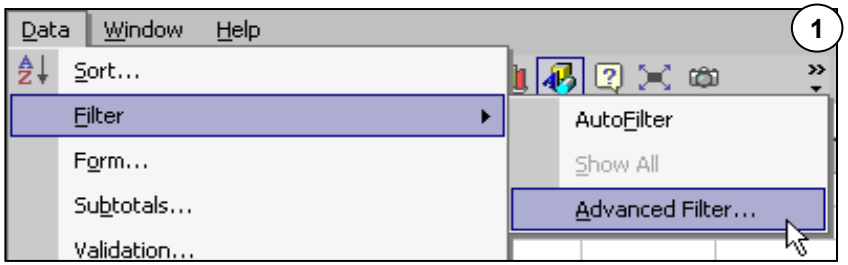
Při použití rozšířeného filtru musí mít databáze i oblast podmínek záhlaví. !

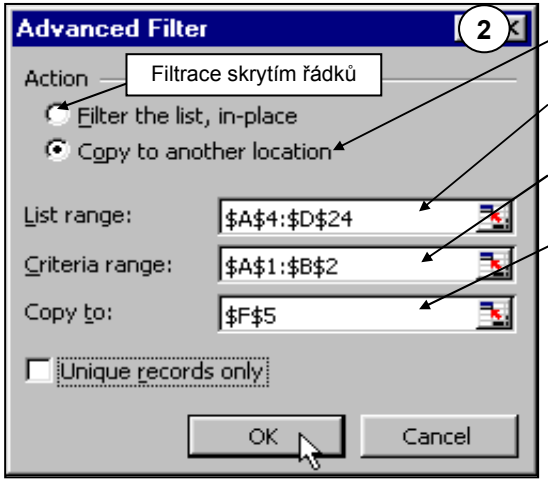
Problém:
Z databáze vybrat pouze řádky s typem `MC79` a průměrem `R1` a `R2` větším `3,3`.

Řešení:
Rozšířený filtr EXCELu umožňuje použití složitých podmínek obsahujících matematické funkce. Umožňuje také provést filtraci skrytím řádek, jako automatický filtr, nebo překopírování dat vyhovujících podmínce na jiné místo listu.

Spuštění rozšířeného filtru

Napište podmínky filtrace (v příkladu buňky **A1:B2**). Označte libovolnou buňku uvnitř databáze, pak přes nabídku **Data - Filter - Advanced Filter** otevřete dialogové okno.





Filtrace skrytím řádků

Filtrace zkopírováním hodnot - extrakce

Oblast databáze včetně záhlaví databáze

Oblast podmínek včetně záhlaví podmínek

Oblast do které bude umístěn výsledek (pouze v případě extrakce)

Výsledek v oblasti **F5:I8**

DATUM	TYP	R1	R2
5.9.02	MC79	3,50	3,20
5.9.02	MC79	3,50	3,20
7.9.02	MC79	3,60	3,10

3

Další možnosti

Porovnávací podmínka
Záhlaví stejné jako záhlaví příslušného sloupce databáze

Odkazovací podmínka
Záhlaví musí být odlišné od záhlaví databáze.

	A	B	C	D
1	TYP	Odpor		
2	MC79	=(C5+D5)/2>3,3		
3				

Vzorec musí obsahovat aspoň jeden odkaz na buňku databáze (**C5**, **D5**).

Hodnota, nebo hodnota s logickým operátorem. Podmínka `<>MC79` by zobrazila typy jiné jako `MC79`.

Mezi hodnotami v jednom řádku platí operátor **AND**

Mezi hodnotami v různých řádcích platí operátor **OR**

	D	E	F	G	H
1	DATUM	TYP		DATUM	TYP
2	4.9.02	MC79		7.9.02	
3					MC79

Zobrazí všechny typy dne `4.9.02`, v ostatních dnech pouze typ `MC79`.

Typ `MC79` dne `4.9.02`

	J	K	L	M
1	DATUM		DATUM	DATUM
2	<4.9.02		>=4.9.02	<=7.9.02
3	>7.9.02			

Zobrazí dny `4.9.02` až `7.9.02`

Vyloučí dny `4.9.02` až `7.9.02`

Filtr v módu filtrace skrytím řádků skryje celé řádky, tedy i buňky mimo oblast databáze. !

Práce s filtrovanými daty

	A	B	C	D
1	DATUM	TYP	R	
2	4.9.02	TL431	3,20	
3	4.9.02	MA78	3,15	
4	4.9.02	MA78	3,35	
5	5.9.02	MC79	3,15	
6	5.9.02	MC79	3,30	
7	5.9.02	MC79	3,35	
8	5.9.02	MC79	3,35	
9			3,26	
10				
11				
12				
13				
14				
15				

=AVERAGE(**C2:C8**)

Automatický filtr

Problém:
Chcete zobrazit pouze den `4.9.02` a vypočítat průměr ze sloupce `R` také pouze pro tento den. Při aplikaci automatického filtru na data bude skryta i buňka **C9** se vzorcem, protože ji filtr považuje za data nevyhovující podmínce. V případě, že oddělíte buňku se vzorcem od dat prázdným řádkem, buňka se vzorcem skryta nebude, ale protože funkce **AVERAGE** počítá i se skrytými buňkami, výsledek nebude závislý na podmínce filtru.

	A	B	C
1	DATU	TYP	R
2	4.9.02	TL431	3,20
3	4.9.02	MA78	3,15
4	4.9.02	MA78	3,35
9			

Chybí průměr

Řešení:
SUBTOTAL(funkce, odkaz) je funkce umožňující provádět výpočty s filtrovanými daty. Ignoruje všechny skryté buňky a buňky s funkcí **SUBTOTAL**.

	A	B	C
12	DATUM	TYP	R
13	4.9.02	TL431	3,20
14	4.9.02	MA78	3,15
15	4.9.02	MA78	3,35
16	5.9.02	MC79	3,15
17	5.9.02	MC79	3,30
18	5.9.02	MC79	3,35
19	5.9.02	MC79	3,35
20			3,26
21			
22			3,00
23			5,00
24			8,00
25			4,00
26			7,00
27			19,00

=SUBTOTAL(1;**C13:C19**)

=SUBTOTAL(9;**C22:C23**)

=SUBTOTAL(9;**C22:C26**)

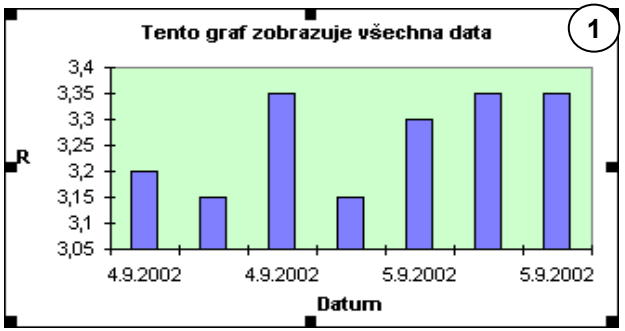
SUBTOTAL(funkce, odkaz):
Funkce je číslo od 1 do 11, které určuje typ funkce použité pro výpočet. Například 1 znamená průměr, 9 součet (nápověda funkce).

Funkce **SUBTOTAL** v buňce **C20** bude počítat pouze s hodnotami zobrazenými filtrem.

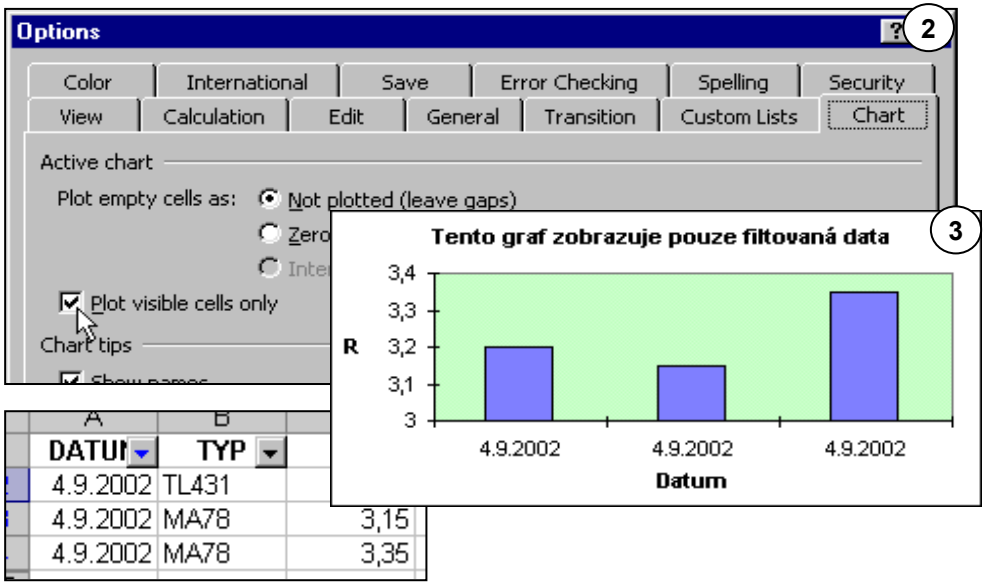
Funkce **SUBTOTAL** v buňce **C27** ignoruje hodnotu 8 v buňce **C24**, protože je výsledkem jiné funkce **SUBTOTAL**.

Nástrojem **Kopírovat - Vložit** přenesete pouze filtrem zobrazená data! !

Graf s filtrovanými daty



Pokud chcete, aby graf zobrazoval pouze hodnoty, které jsou výsledkem filtrace, skontrolujte označení políčka **Plot visible cells only** v záložce **Chart**. Označte graf a přes nabídku **Tools - Options** otevřete dialogové okno **Options**. Objeví se záložka **Chart**. Políčko **Plot visible cells only** musí být označené.



Toto nastavení platí pro každý graf individuálně! !

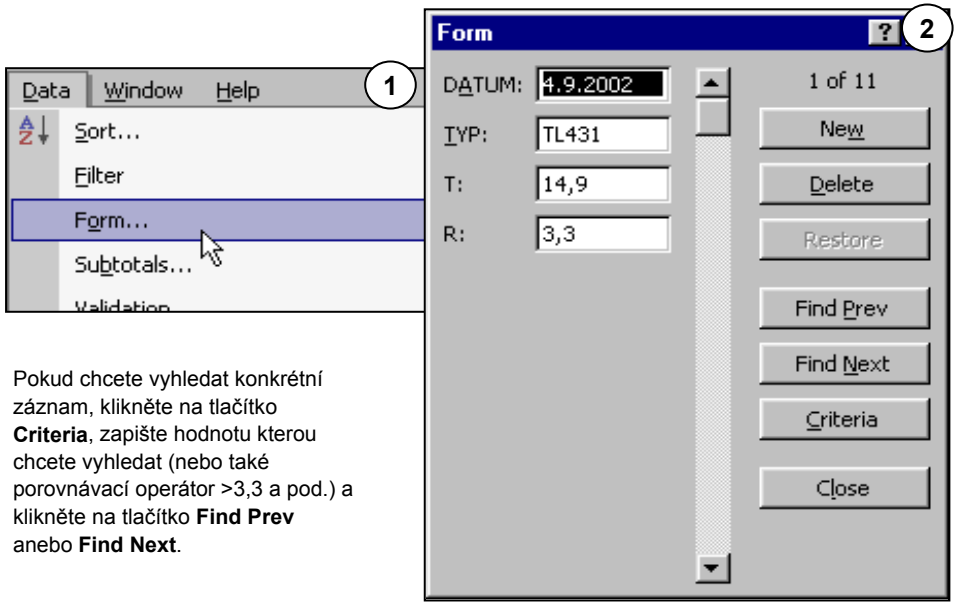
Formuláře a souhrny (Form, Subtotal)

	A	B	C	D
1	DATUM	TYP	T	R
2	4.9.02	TL431	14,90	3,30
3	4.9.02	MA78	15,08	3,25
4	4.9.02	MA78	14,90	3,35
5	6.9.02	MA78	15,10	3,55
6	6.9.02	MA78	15,03	3,35
7	7.9.02	TL431	14,65	2,90
8	7.9.02	MC79	15,55	3,35
9	8.9.02	MC79	15,43	3,25
10	8.9.02	MA78	15,18	3,35
11	8.9.02	MA78	15,13	3,35
12	8.9.02	MA78	15,48	3,30
13				

Problém:
Zjednodušit zadávání nebo úpravy dat. Rychle vypočítat průměry v jednotlivých dnech.

Řešení:
Nástroj **Form** slouží ke zjednodušení úprav databází.
Nástroj **Subtotals** vytvoří potřebné souhrnné výpočty.

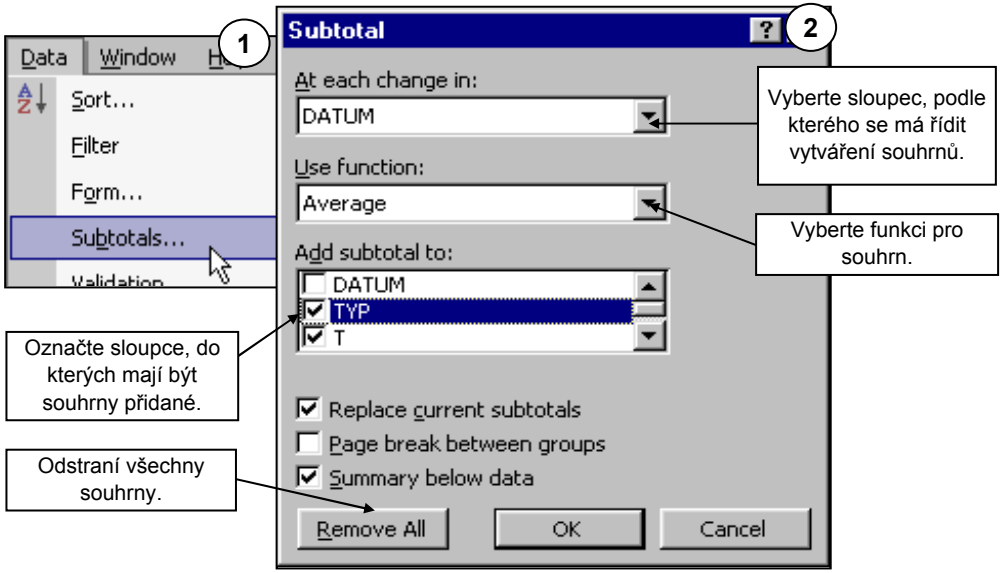
Označte některou buňku databáze. Přes nabídku **Data - Form** vytvoříte formulář pro editaci databáze. Pomocí formuláře můžete měnit údaje v jednotlivých záznamech, přidávat záznamy, nebo je odstraňovat.



Pokud chcete vyhledat konkrétní záznam, klikněte na tlačítko **Criteria**, zapište hodnotu kterou chcete vyhledat (nebo také porovnávací operátor >3,3 a pod.) a klikněte na tlačítko **Find Prev** anebo **Find Next**.

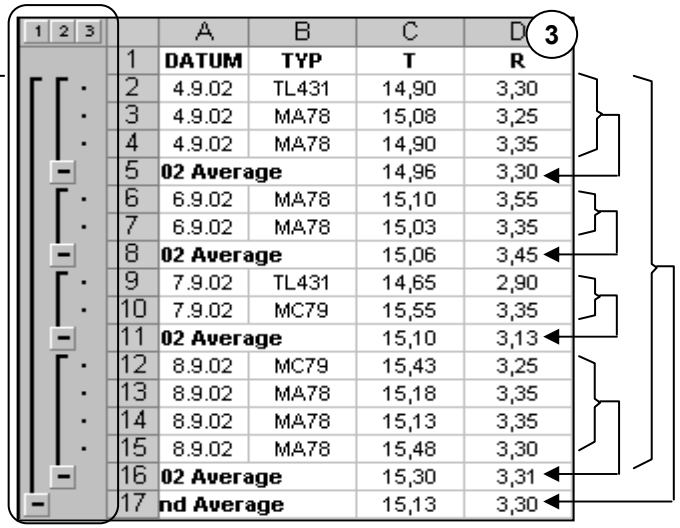
Souhrn (Subtotal)

Označte některou buňku databáze. Přes nabídku **Data - Subtotals** otevřete okno **Subtotal**.



Ovládací lišta přehledů

Pomocí ovládacích prvků přehledů můžete skrývat nebo zobrazovat části tabulky.



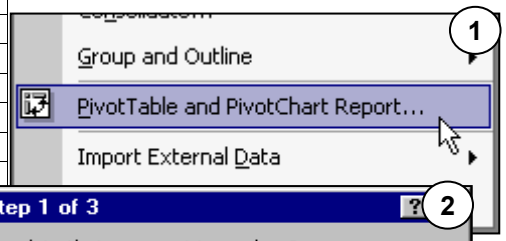
Kontingenční tabulka (Pivot Table)

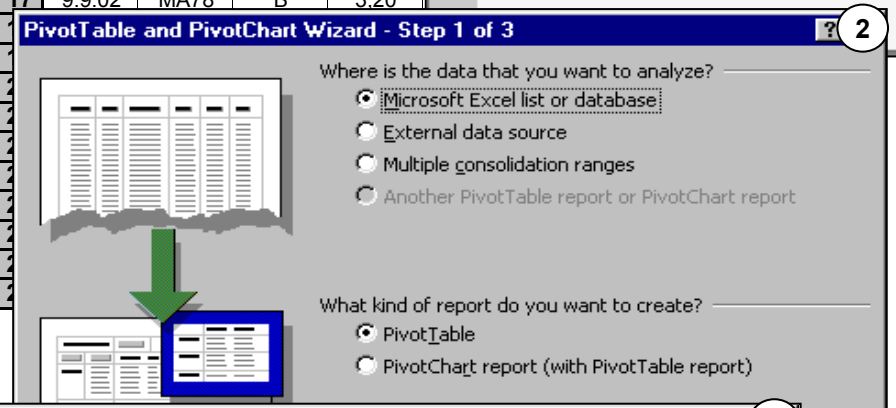
	A	B	C	D
1	DATUM	TYP	REAKT	R
2	6.9.02	MA78	A	3,55
3	6.9.02	MA78	A	3,45
4	6.9.02	MA78	A	3,35
5	6.9.02	MA78	A	3,25
6	6.9.02	MA78	A	3,35
7	7.9.02	MC79	A	3,20
8	7.9.02	MC79	A	3,35
9	8.9.02	MC79	A	3,25
10	8.9.02	MA78	A	3,35
11	8.9.02	MA78	A	3,30
12	9.9.02	MA78	A	3,50
13	9.9.02	MA78	A	3,30
14	9.9.02	MA78	A	3,55
15	9.9.02	MA78	A	3,35
16	9.9.02	MA78	B	3,35
17	9.9.02	MA78	B	3,20

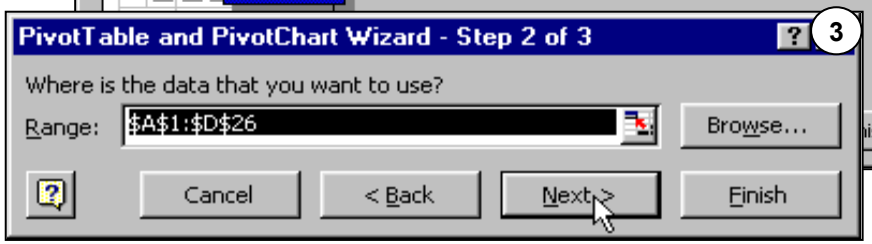
Problém:
Sestrojit tabulku přehledu zpracování typů v jednotlivých dnech.

Řešení:
Nástroj **Pivot Table** umožňuje vytvářet z databází křížové tabulky.

Označte některou buňku databáze. Přes nabídku **Data - Pivot Table** otevřete okno **PivotTable and PivotChart Wizard**.

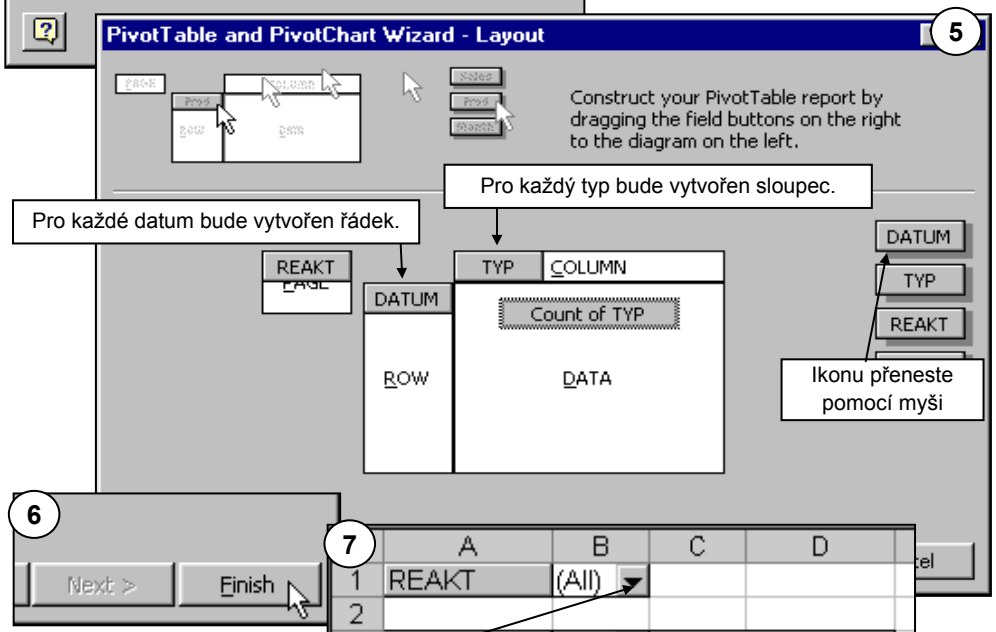


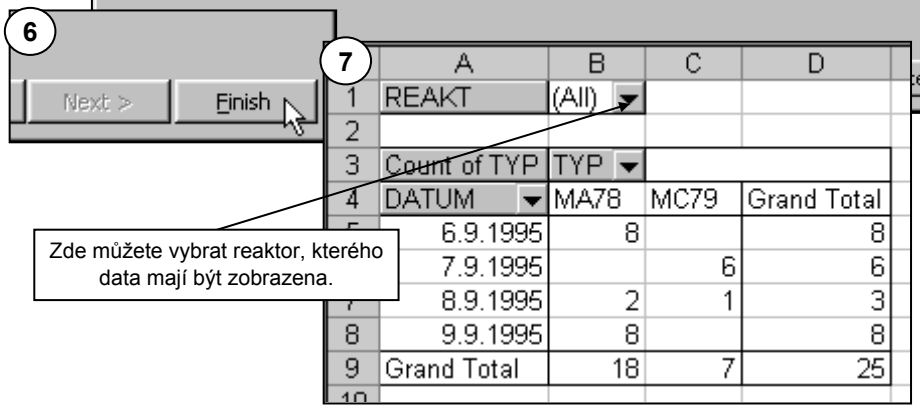




Kontingenční tabulka může být vytvořena na novém listu.

Klikněte na tlačítko **Layout** a vyberte oblast, ve které se má označený výběr zobrazit nebo na tlačítko **Finish** a pak vyberte oblasti zobrazení výběrů v otevřeném okně **Pivot Table Field List**.





Použití ovládacích prvků

	A	B	C	D
1	DATUM	TYP	T	R
2	2.9.02	MA78	15,00	3,60
3	2.9.02	MA78	15,33	3,25
4	3.9.02	MA78	14,60	3,30
5	3.9.02	MA78	14,88	3,45
6	3.9.02	TL431	14,98	3,30
7	4.9.02	TL431	14,90	3,30
8	4.9.02	MA78	15,08	3,25
9	4.9.02	MA78	14,90	3,35
10	5.9.02	MC79	14,80	3,15
11	5.9.02	MC79	14,83	3,30
12	5.9.02	MC79	14,98	3,35
13	5.9.02	MC79	14,88	3,35
14	6.9.02	MA78	15,10	3,55
15	6.9.02	MA78	15,03	3,35
16	7.9.02	TL431	14,65	2,90
17	7.9.02	MC79	15,55	3,35
18	8.9.02	MC79	15,43	3,25
19	8.9.02	MA78	15,18	3,35
20	8.9.02	MA78	15,13	3,35
21	8.9.02	MA78	15,48	3,30
22	9.9.02	MA78	15,30	3,50
23	9.9.02	MA78	15,25	3,55
24				

Problém:

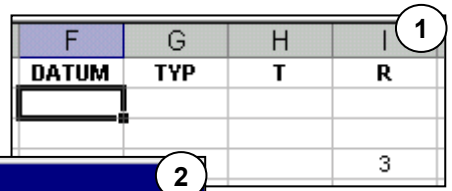
Rychle zobrazit záznam databáze podle aktuální požadavky.

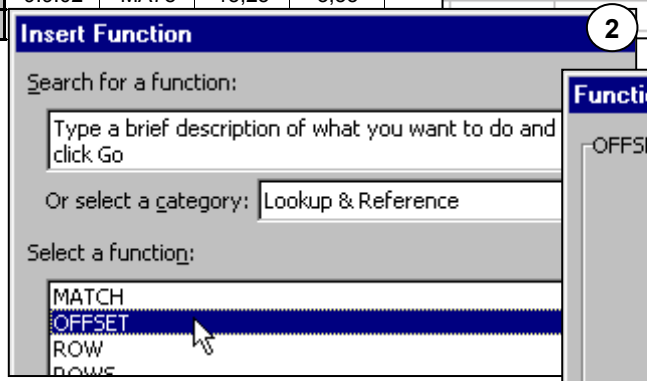
Řešení:

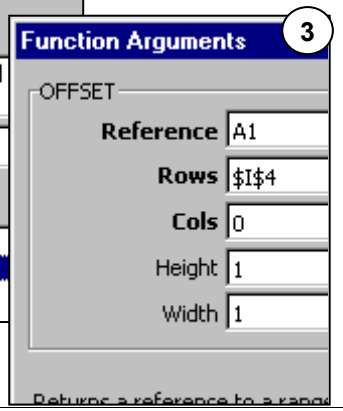
Pomocí funkce **OFFSET** vytvořte odkaz na databázi, který umožní vyhledat a v definované oblasti zobrazit určený počet řádků a sloupců z této databáze.

Postup

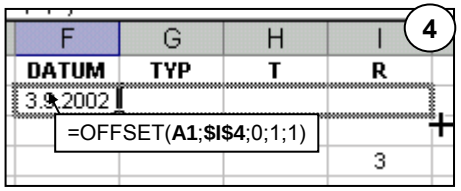
Do buňky **I4** napište číslo udávající počet řádku, o kolik se posune začátek zobrazení hledaného záznamu od začátku databáze. Označte a zkopírujte hlavičku databáze a vložte ji do buňky **F1**. Do buňky **F2** vložte funkci **OFFSET**(v kategorii **Lookup & Reference**) a zadejte vstupní hodnoty podle obrázku.



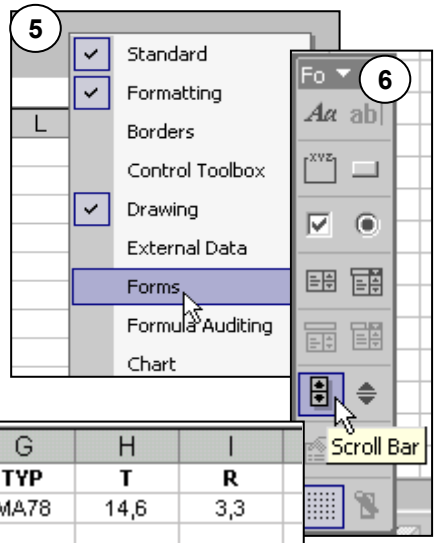


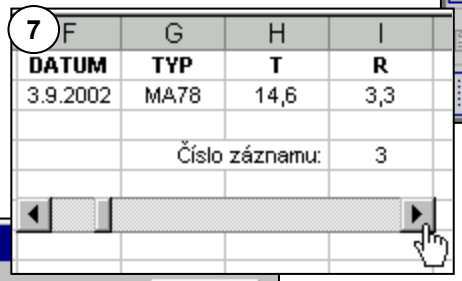


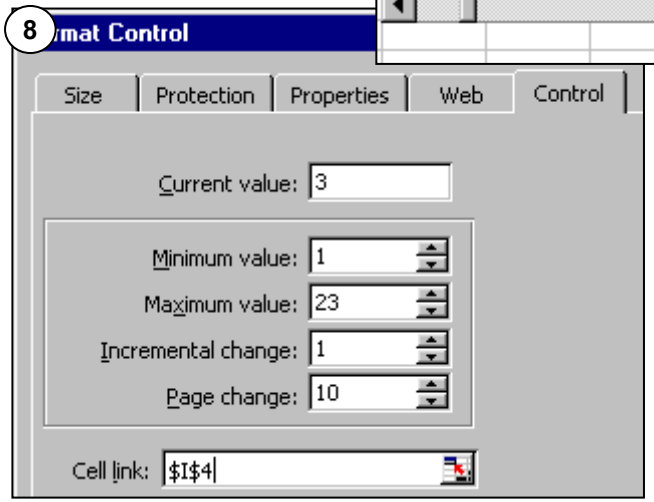
Označte buňku **F2**. Tažením za značku v pravém dolním rohu směrem vpravo zkopírujte vloženou funkci i do buněk **G2** - **I2**. Upravte formát buňky **F2** na "Date". Změnou hodnoty v buňce **I4** se v poli **G2** - **I2** zobrazí odpovídající záznam.



Pro snadnější ovládání této funkce vložte ovládací prvek **Scroll Bar** z panelu nástrojů **Forms** do listu. Přes pravou myš umístěnou na tomto prvku vyberte z nabídky **Format Control** a v dialogovém okně zadejte hodnoty podle obrázku.







Databázové funkce EXCELu

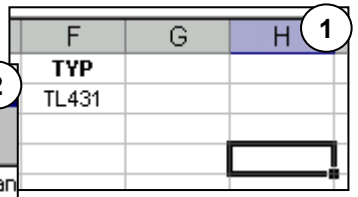
	A	B	C	D
1	DATUM	TYP	T	R
2	2.9.02	MA78	15,00	3,60
3	2.9.02	MA78	15,33	3,25
4	3.9.02	MA78	14,60	3,30
5	3.9.02	MA78	14,88	3,45
6	3.9.02	TL431	14,98	3,30
7	4.9.02	TL431	14,90	3,30
8	4.9.02	MA78	15,08	3,25
9	4.9.02	MA78	14,90	3,35
10	5.9.02	MC79	14,80	3,15
11	5.9.02	MC79	14,83	3,30
12	5.9.02	MC79	14,98	3,35
13	5.9.02	MC79	14,88	3,35
14	6.9.02	MA78	15,10	3,55
15	6.9.02	MA78	15,03	3,35
16	7.9.02	TL431	14,65	2,90
17	7.9.02	MC79	15,55	3,35
18	8.9.02	MC79	15,43	3,25
19	8.9.02	MA78	15,18	3,35
20	8.9.02	MA78	15,13	3,35
21	8.9.02	[illegible]	[illegible]	[illegible]
22	9.9.02	[illegible]	[illegible]	[illegible]
23	9.9.02	[illegible]	[illegible]	[illegible]
24	9.9.02	[illegible]	[illegible]	[illegible]

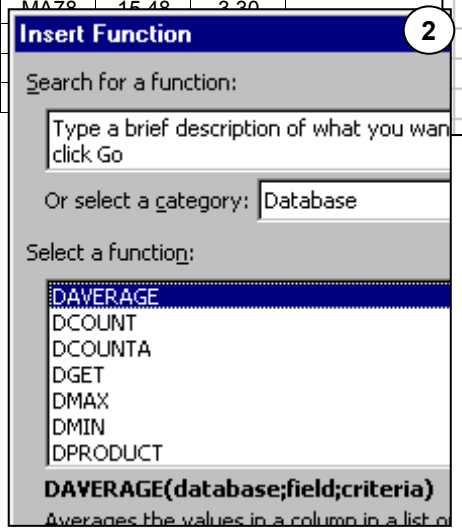
Problém:
Získat průměr hodnot T pro typ TL431 z celé databáze.

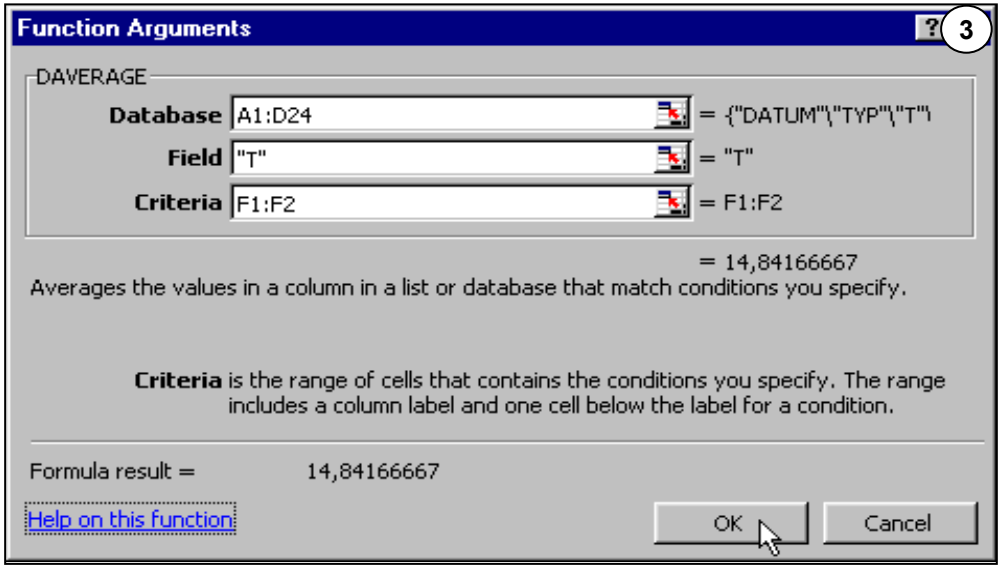
Řešení:
Funkce **DAVERAGE** umožňuje výpočet průměru pouze hodnot vybraných z databáze podle kritérii specifikovaných v podmínce výběru.

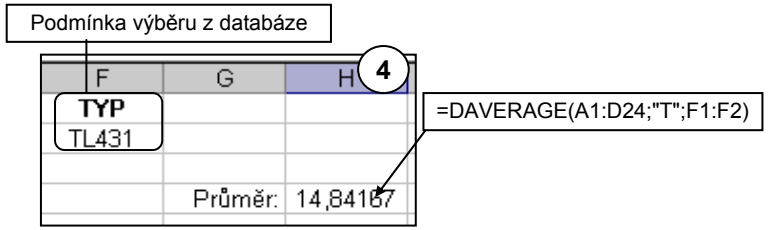
Postup

Do buněk **F1** a **F2** vložte podmínku výběru z databáze. Do buňky **H4** vložte funkci **DAVERAGE** (v kategorii **Database**) a zadejte vstupní hodnoty podle obrázku.









Podmínky výběru pro databázové funkce se tvorí stejným způsobem jako pro rozšířený filtr (Advanced Filter) (list 8).

Stejný postup lze použít pro všechny funkce z kategorie Database. !

Výpočty a závislosti (Calculations, Auditing)

Přesnost výpočtů

	A	B	C	D	E	F
1	T1	T2	Rozp.		Hodn.	Počet
2	944,6	944,5		0	0,1	0
3	946,5	946,8	0,30	0		
4	890,7	890,7	0,00	0		
5	894,6	894,6	0,00	0		
6	880,7	880,8	0,10	0		
7	880,6	880,8	0,20	0		
8	945,5	945,4	0,10	0		
9	940,4	940,5	0,10	0		
10	950,4	950,3	0,10	0		
11	946,7	946,6	0,10	0		
12	934,2	933,7	0,50	0		
13	931,4	930,6	0,80	0		
14	890,8	890,7	0,10	0		
15	894,3	894,8	0,50	0		
16	881,0	881,0	0,00	0		

=SUM(**D2**:**D16**)

=IF(**C2**=**E2**;1;0)

=ABS(**A2**-**B2**)

C
Rozp.
0,100000000000023
0,299999999999955
0,000000000000000
0,000000000000000
0,099999999999909
0,199999999999932
0,100000000000023
0,100000000000023

Problém:

Sčítat, kolik rozpětí `0,1` se nachází ve sloupci `Rozp`. Standardní postup vede k výsledku, že hledaná rozpětí neexistují. Výsledek v buňce **F2** je `0`, i když ve sloupci `Rozp.` je sedm rozpětí `0,1`.

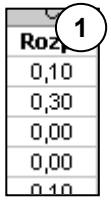
Příčina:

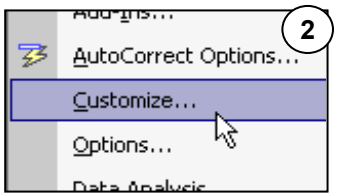
Roztáhněte sloupec **C** tak, aby do něj vešlo alespoň 15 desetinných míst, pak nastavte počet zobrazených desetinných míst na 15.
Je vidět, že vypočtená rozpětí nejsou přesné hodnoty, ale jsou zatížená malou chybou, která vzniká při výpočtu v pohyblivé řádové čárce, nicméně při porovnávání s přesnou hodnotou je zjištěna nerovnost hodnot.

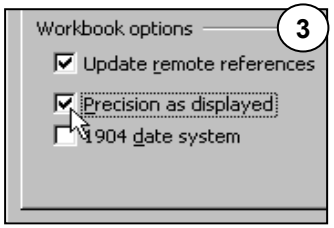
Řešení:

Jednou z možností je vhodné zaokrouhlení výpočtů pomocí funkce **ROUND**(číslo; číslice). Další možností je nastavení **Precision as displayed**.

V oblasti C2:C16 nastavte číselný formát 0,00. Přes nabídku **Tools - Options** otevřete dialogové okno **Options - Calculation**.



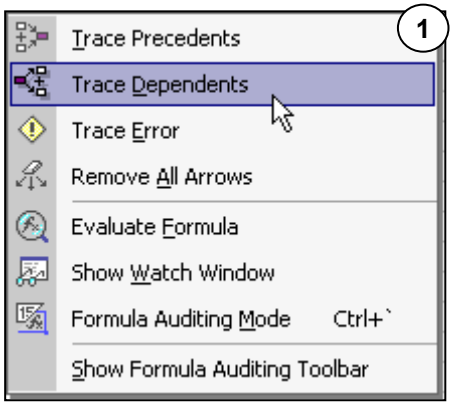


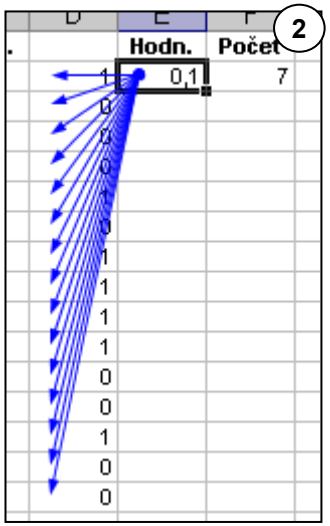


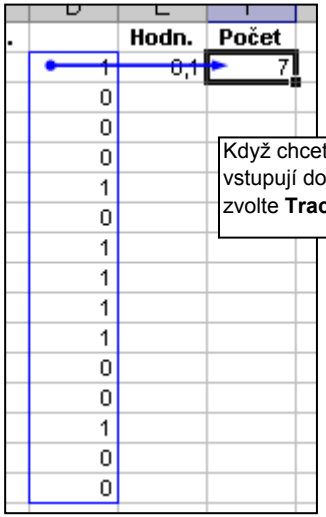
Nastavení **Precision as displayed** platí pro celý sešit a ukládá se s ním. Nastavení je aktuální pouze pro buňky s číselným formátem, který má pevný počet desetinných míst. !

Hledání závislostí (Auditing)

V některých případech jsou výpočty v listu tak složité, že ztrácíte přehled. Jindy se může vyskytnout chyba. Přehled v závislostech mezi buňkami získáte pomocí nástroje **Formula Auditing**. Označte buňku o níž chcete zjistit, kde všude její hodnota vstupuje. Pak přes nabídku **Tools -Formula Auditing - Trace Dependents** získáte znázornění závislostí přímo v listu.

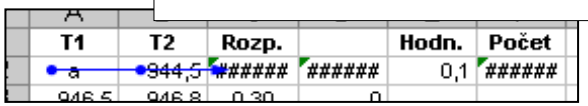






Když chcete získat představu o tom, které hodnoty vstupují do výpočtu, označte buňku se vzorcem a zvolte **Trace Precedents**.

Pokud v některé buňce vznikne výpočtová chyba, označte ji a zvolte **Trace Error**. Získáte znázornění postupu výpočtu a šíření chyby.



Všechny šipky zrušíte volbou **Remove All Arrows**. !

Funkce a pole

	A	B	C	D	E
1	81,45	95,54	114,2		
2	114,27	101,64	110,07		84
3	96,33	88,62	113,62		88
4	90,29	88,13	106,84		92
5		108,48			96
6	107,69	86,28			100
7	94,62	90,35	79,18		104
8	100,42	108,88	124,11		108
9	90,4		86,85		112
10	101,17	86,32	127,87		116
11	B	114	94,16		
12	A	94,49	99,38		
13	104,76	87,43			
14	89,17	89,3	113,75		
15	87,89	105,87	98,59		
16					

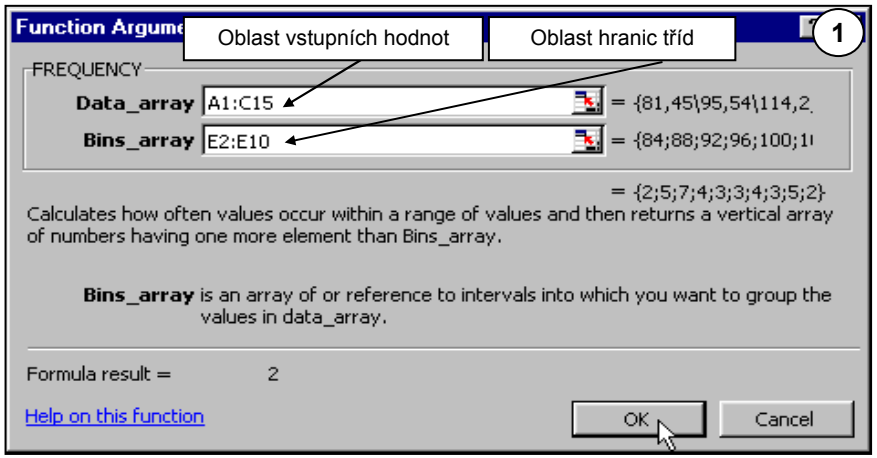
Problém:
Připravit data pro sestrojení histogramu. Hranice tříd jsou v sloupci **E**.

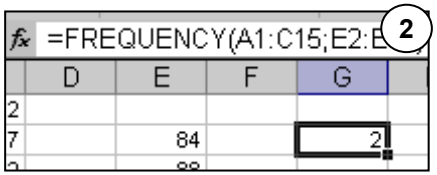
Řešení:
Funkce **FREQUENCY** provede rozdělení dat do zadaných tříd. Její výstupem je pole.

Funkce **FREQUENCY** vrátí pole v rozsahu jeden sloupec a počtem řádků rovným počtu hranic tříd + 1. V první třídě jsou spočteny všechny hodnoty od - nekonečna po první hranici. V poslední třídě pak hodnoty od poslední hranice po nekonečno.
Výstupem funkce FREQUENCY - v tomto příkladu - bude tedy vertikální 10 prvkové pole.

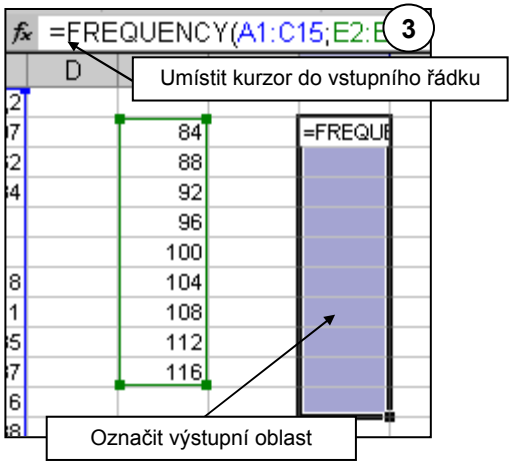
Postup

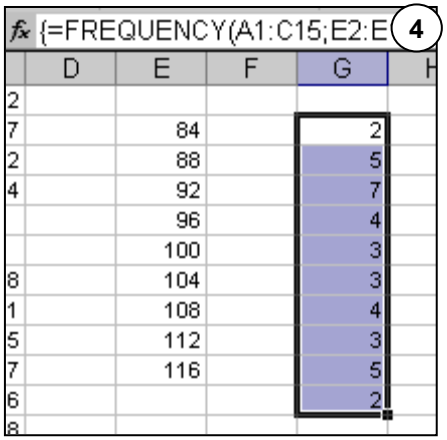
Označte buňku ve které má být první četnost. Vložte funkci **FREQUENCY** (v kategorii **Statistical**). Zadejte vstupní hodnoty podle obrázku.





V buňce **G2** je nyní vložena funkce. Označte oblast výstupního pole **G2:G11**, pak stiskněte **F2** (nebo umístěte kurzor do vstupního řádku) a pak kombinaci kláves **Ctrl - Shift - Enter**. Označená oblast pak bude vyplněna výstupními hodnotami funkce.

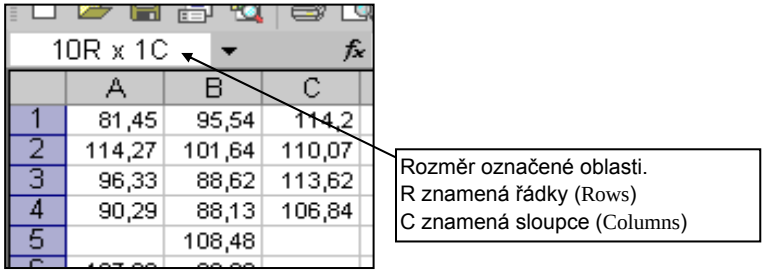




Podobně se vkládají i další funkce jejichž výstupem jsou pole.

Pole jsou také vstupem funkce **FREQUENCY**. Ve vstupních polích dat jsou ignorovány nečíselné hodnoty.

Během označování výstupního pole sledujte okénko v levém horním rohu sešitu, kde je zobrazen rozměr označené oblasti. Nemusíte pak počítat buňky.



Datová tabulka (Table)

	A	B	C	D
1	P =	105	kW	
2	F_H =	270	l/min	
3				
4				
5	h =	16,78	µ	
6				
7				
8				
9				

=14,38+0,064***B1**-0,016***B2**

Problém:

Pro tloušťku vrstvy byla experimentem zjištěna rovnice h=14,38+0,064*P-0,016*F_H, kde h je tloušťka v µ, P je výkon v kW a F_H je průtok vodíku v l/min. Je třeba sestavit tabulku tloušťky, pro výkon P v rozsahu 105 až 115 kW a průtok vodíku F_H od 260 do 380 l/min.

Řešení:

Sestrojte datovou tabulku se dvěma vstupy.

Sestrojení datové tabulky

Podle obrázku sestrojte záhlaví tabulky a levý sloupec. Vzorec musí být umístěn v levém horním rohu tabulky a nesmí být navázán na žádnou buňku budoucí tabulky. Pak označte oblast **B5:I16** a přes nabídku **Data - Table** otevřete dialogové okno **Table**. Odkaz pro **Row input cell** bude **B2** a pro **Column input cell** bude **B1**.

1

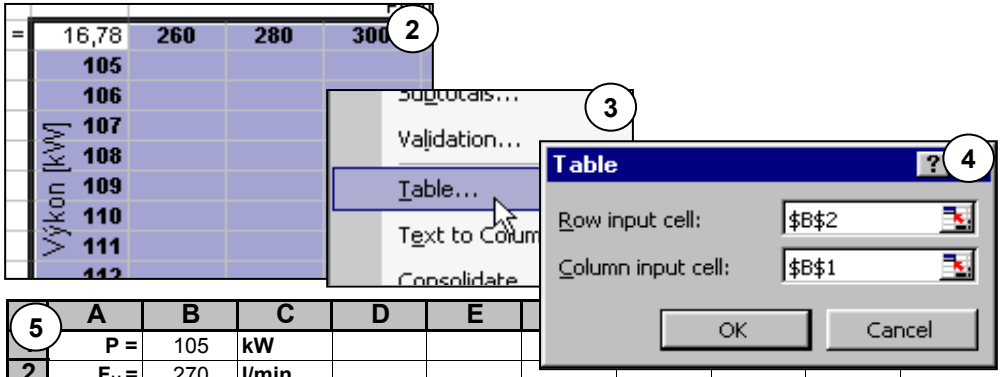
	A	B	C	D	E	F	G	H	
1	P =	105	kW						
2	F_H =	270	l/min						
3									
4			Průtok H2 [l/min]						
5	h =	16,78	260	280	300	320	340	360	380
6		105							
7		106							
8		107							
9		108							
10		109							
11		110							
12		111							
13		112							
14		113							
15		114							
16		115							
17									

=14,38+0,064***B1**-0,016***B2**
vzorec v levém horním rohu tabulky

Hodnoty pro průtok vodíku

Výkon [kW]

Hodnoty pro výkon



5

	A	B	C	D	E				
1	P =	105	kW						
2	F_H =	270	l/min						
3									
4			Průtok H2 [l/min]						
5	h =	16,78	260	280	300	320	340	360	380
6		105	16,94	16,62	16,30	15,98	15,66	15,34	15,02
7		106	17,00	16,68	16,36	16,04	15,72	15,40	15,08
8		107	17,07	16,75	16,43	16,11	15,79	15,47	15,15
9		108	17,13	16,81	16,49	16,17	15,85	15,53	15,21
10		109	17,20	16,88	16,56	16,24	15,92	15,60	15,28
11		110	17,26	16,94	16,62	16,30	15,98	15,66	15,34
12		111	17,32	17,00	16,68	16,36	16,04	15,72	15,40
13		112	17,39	17,07	16,75	16,43	16,11	15,79	15,47
14		113	17,45	17,13	16,81	16,49	16,17	15,85	15,53
15		114	17,52	17,20	16,88	16,56	16,24	15,92	15,60
16		115	17,58	17,26	16,94	16,62	16,30	15,98	15,66

Výkon [kW]

Další možnosti

	L	M	N	O
5	Úhel	Sin	cos	
6	0	0	1	
7	10			
8	20			
9	30			
10	40			
11	50			
12	60			
13	70			
14	80			
15	90			

=COS(RADIANS(**L6**))

=SIN(RADIANS(**L6**))

Můžete vytvořit také datovou tabulku s jedním vstupem, ale pro více vzorců.

Označte oblast **L6:N15**, přes nabídku **Data - Table** otevřte okno **Table** a zadejte **L6** jako **Column input cell**.

Makra

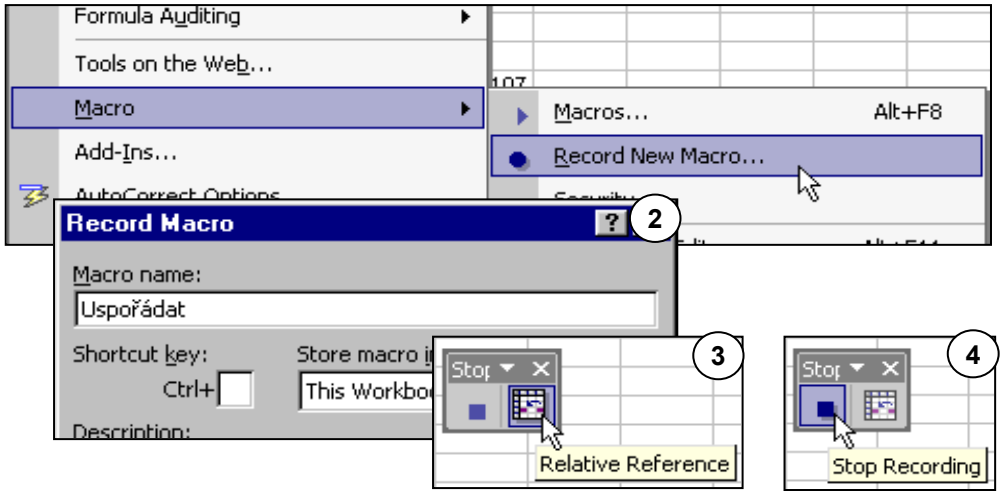
	A	B	C	D	E	F
1	A	0,34214			A	0,66274
2	0,44584	B			B	0,59844
3	C	0,49217			C	0,71783
4	0,97729	D			D	0,68654
5	E	0,41337			E	0,97461
6						
7						
8	A	0,55531			A	0,50395
9	B	0,18882			B	0,34608
10	C	0,75652			C	0,50295
11	D	0,9613			D	0,46086
12	E	0,37165			E	0,12815

Problém:
Ve všech oblastech uspořádat čísla a písmena tak, jako v oblasti **A1:B5**.

Řešení:
Nástroj **Makro** slouží k automatizování opakované činnosti.

Záznam makra

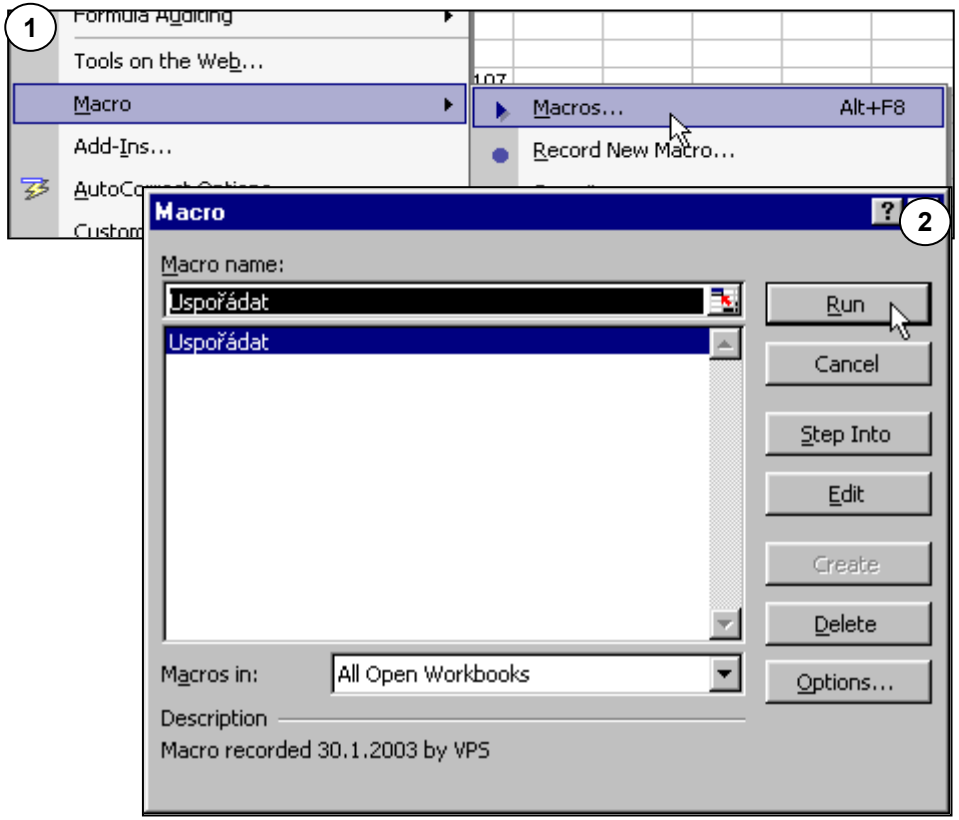
Označte buňku **E1**. Přes nabídku **Tools - Macro - Record New Macro...** otevřete dialogové okno **Record Macro**. Napište **Macro name:** - `Uspořádat` a klikněte na **OK**. Objeví se tlačítka pro ukončení makra a pro použití **Relative Reference**. Klikněte na tlačítko **Relative Reference** a upravte buňky v oblasti **E1:F5** podle vzoru v **A1:B5**. Záznam makra ukončete kliknutím na tlačítko **Stop Recording**.



Spuštění makra

Nyní, když máte připravený nástroj, můžete jej aplikovat na každou oblast.

Označte buňku v levém horním rohu oblasti (**A8**). Přes nabídku **Tools - Macro** otevřete dialogové okno **Macro**. Zvolte makro **Uspořádat** a klikněte na tlačítko **Run**.



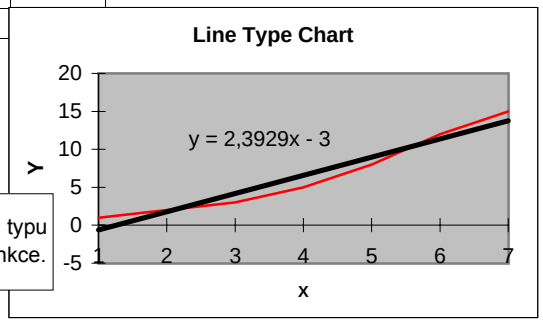
Při použití relativních odkazů se činnost makra provádí vždy vzhledem k označené buňce před spuštěním makra. Pokud chcete aby činnost makra byla nezávislá, nesmí být zapnuta možnost **Relative Reference** před spuštěním záznamu.

Graf

	A	B	C	D
1	Hodnoty pro zobrazení přímky y=x			
2	X	Y		
3	1	1		
4	2	2		
5	3	3		
6	5	5		
7	8	8		
8	12	12		
9	15	15		
10				

Problém:
Správně zobrazit přímku a rovnici spojnice trendu. Zobrazení uvedených dat grafem Line nevede k správnému výsledku.

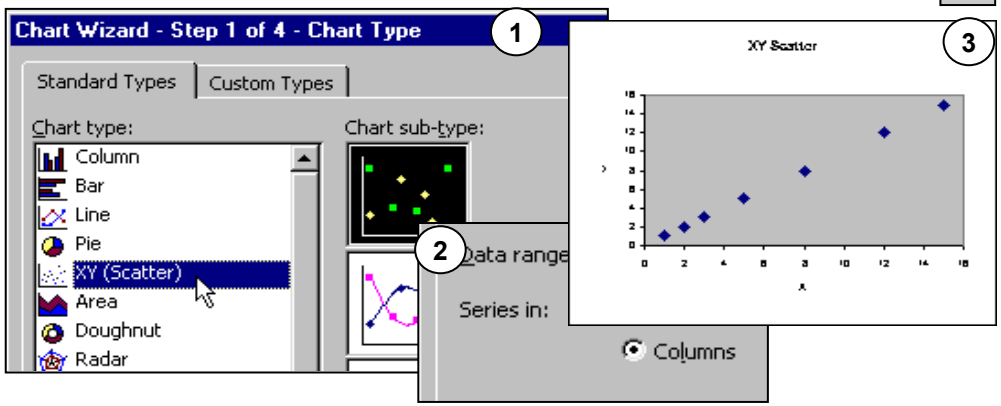
Řešení:
Pro správné zobrazení funkce použijte graf **XY (Scatter)**.



Nesprávné použití grafu typu "**Line**" pro zobrazení funkce.

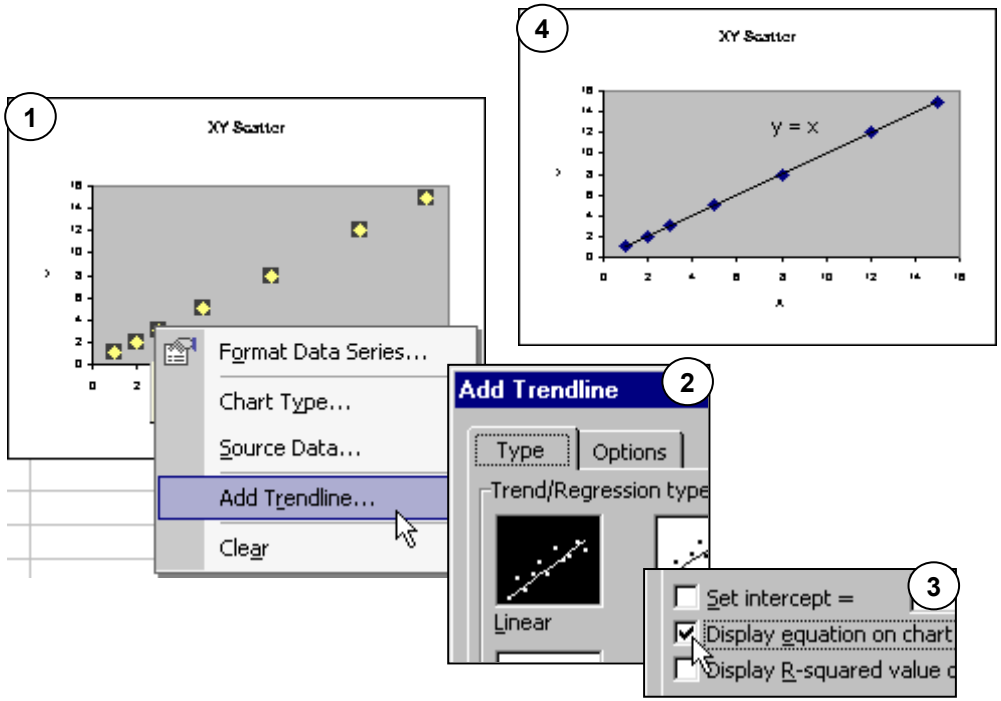
XY (Scatter)

Označte oblast A3:B9 a pomocí **Chart Wizard** sestrojte **XY (Scatter)** graf.



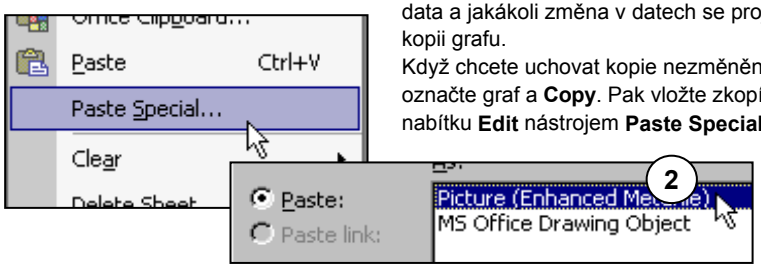
Vložení spojnice trendu

Klikněte na některý bod zobrazené řady údajů a pak pravým tlačítkem myši vyvolejte místní nabídku. Zvolte **Add Trendline**, vyberte záložku **Type** a označte **Linear**, a vyberte záložku **Options** a označte položku **Display equation on chart**.



Kopírování grafů

Při kopírování grafu bude jeho kopie napojena na původní data a jakákoli změna v datech se projeví jak v originálu, tak v kopii grafu.
Když chcete uchovat kopie nezměněné, stiskněte **Shift**, označte graf a **Copy**. Pak vložte zkopírovaný graf přes nabítku **Edit** nástrojem **Paste Special** a zvolte **Picture**.



Uživatelem definovaný typ grafu

	A	B	C	D
1	DATUM	R	Dolní	Horní
2	10.8.02	3,63	2,6	4
3	11.8.02	3,28		
4	12.8.02	3,4		
5	13.8.02	3,3		
6	15.8.02	3,18		
7	16.8.02	3,23		
8	18.8.02	3,27		
9	19.8.02	3,31		
10	20.8.02	3,33		
11	21.8.02	3,31		
12	22.8.02	3,36		
13	23.8.02	3,41		
14	24.8.02	3,48		
15	25.8.02	3,33		
16	26.8.02	3,38		
17	27.8.02	3,34		
18	28.8.02	3,33		
19	29.8.02	3,31		
20	31.8.02	3,31	2,6	4
21				

Nastavte graf tak, aby se meze zobrazovaly jako vodorovné čáry. Upravte formát, zrušte legendu, přizpůsobte velikost písma.

Problém:

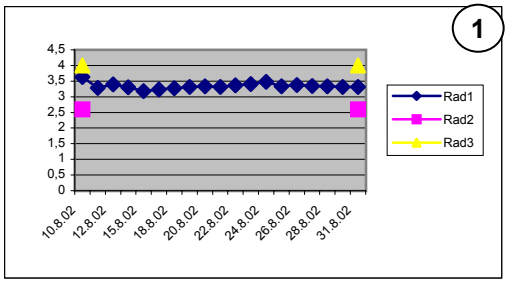
Každý měsíc zobrazit pohyb hodnoty `R` v mezích `Dolní` a `Horní`. Je jednoduché vytvořit potřebný graf. Soustavné nastavování formátu grafu je však zdlouhavé.

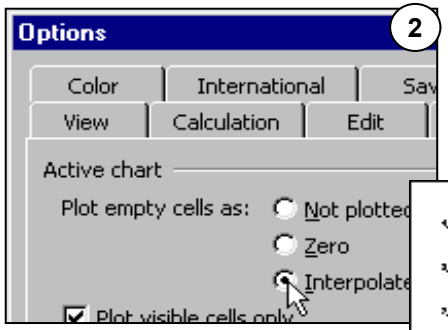
Řešení:

Vytvořte potřebný formát a pak jej uložte jako typ grafu definovaný uživatelem. Ten pak můžete kdykoli vyvolat.

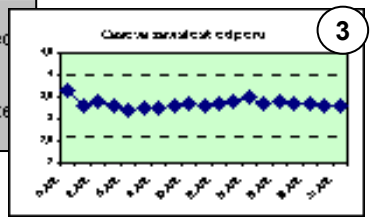
Postup

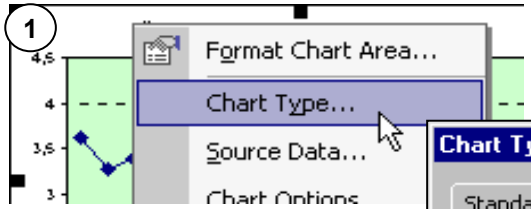
Označte oblast **A1:D20**, pak pomocí **Chart Wizard** sestrojte Line chart, sub-type 4.





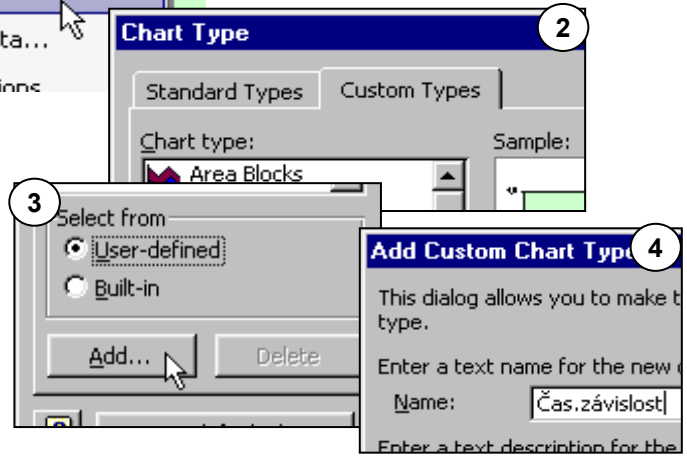
Označte sestrojený graf. Přes nabídku **Tools - Options** otevřete okno **Options** a záložku **Chart**. Označte volbu **Plot empty cells - Interpolated**. Toto nastavení umožňuje zadávat pouze krajní body úseček (v tomto případě mezí v buňkách **C2**, **D2**, **C20**, **D20**).





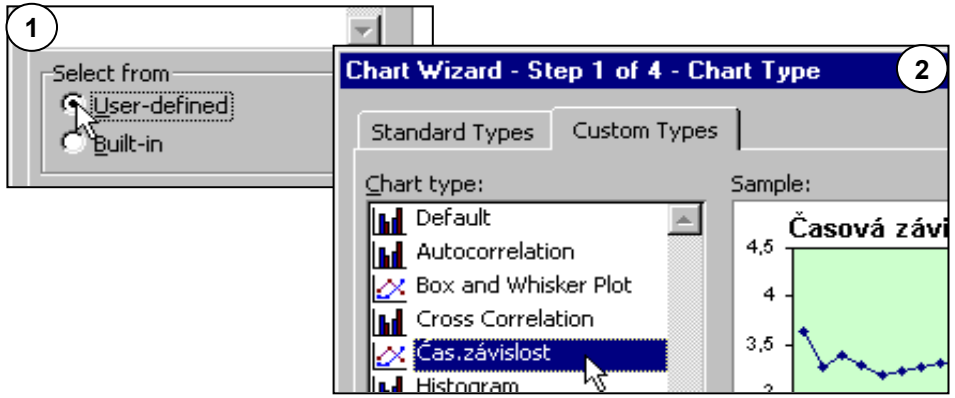
Nyní, když máte vytvořený formát jaký potřebujete, uložte jej jako **User-defined Type**.

Označte graf, pak pravým tlačítkem myši vyvolejte místní nabídku. Zvolte **Chart Type**, čímž otevřete okno **Chart Type**. Zde vyberte **Custom Types,** pak **User-defined** a **Add...**. Napište název nového vlastního typu grafu a zvolte **OK**.

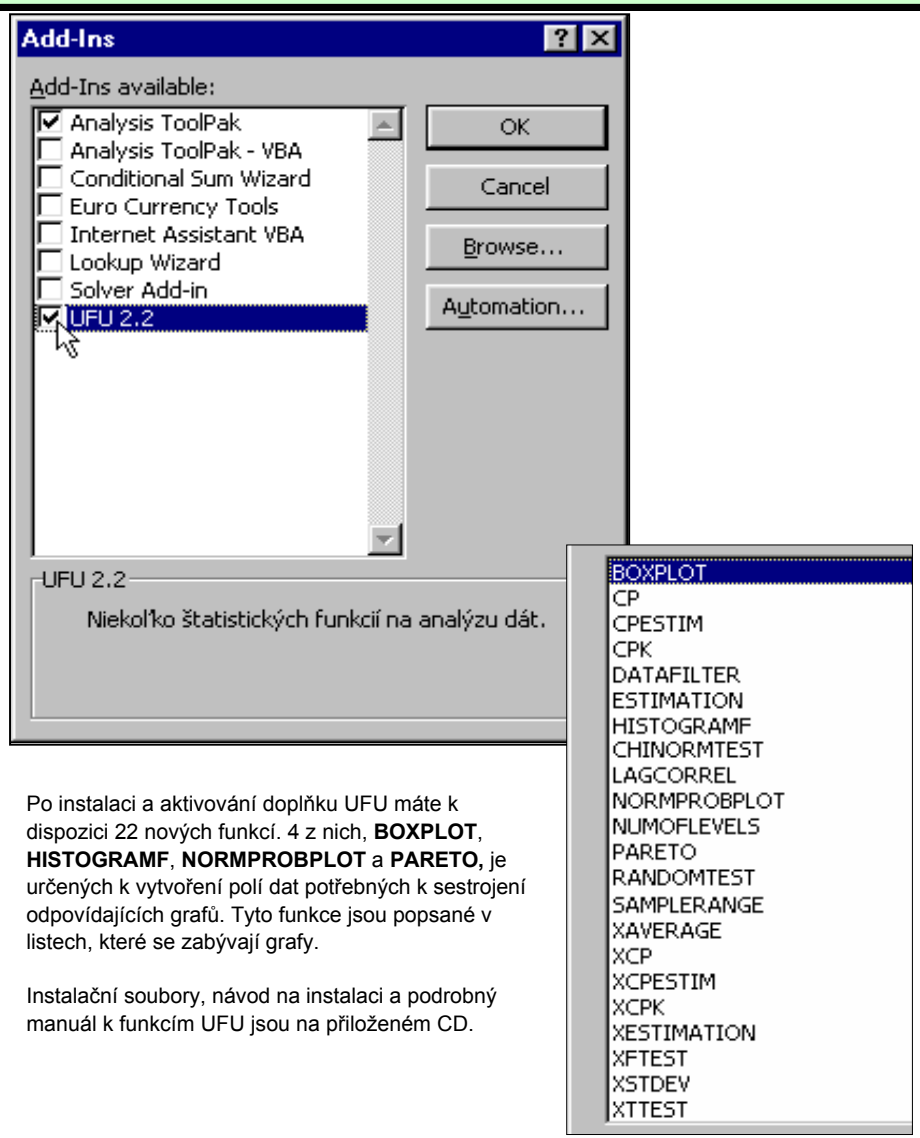


Použití vlastního typu grafu

Označte oblast vstupních dat. Pomocí ikony **Chart Wizard** otevřete okno **Chart Wizard**. Zde vyberte **Custom Tytes**, pak **User-defined** a vyberte vhodný typ grafu.

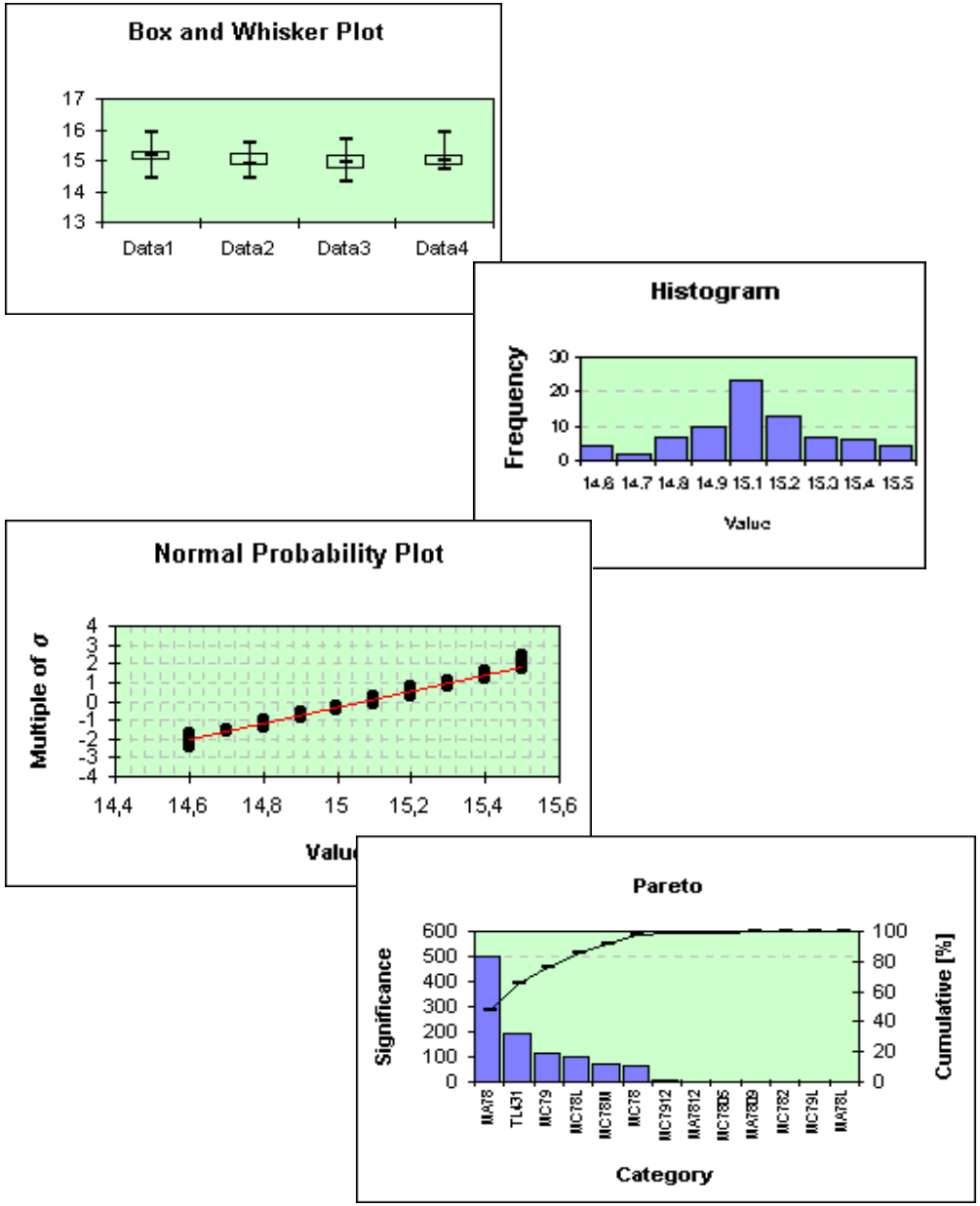


UFUnkce užitečné 2.2 UFU - užitečné funkce



Po instalaci a aktivování doplňku UFU máte k dispozici 22 nových funkcí. 4 z nich, **BOXPLOT**, **HISTOGRAMF**, **NORMPROBPLOT** a **PARETO,** je určených k vytvoření polí dat potřebných k sestrojení odpovídajících grafů. Tyto funkce jsou popsané v listech, které se zabývají grafy.

Instalační soubory, návod na instalaci a podrobný manuál k funkcím UFU jsou na přiloženém CD.

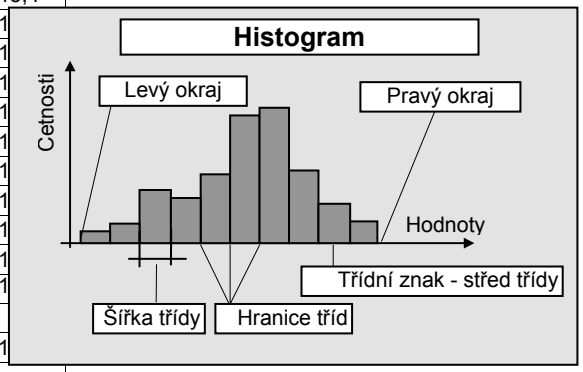


Histogram

	A	B	C	D
1	**Tloušťka, náh. výběr z 12.02**			
2	15,1	15	15,2	15,4
3	14,9	15,1	15,2	15,5
4	15	15	15,1	15,4
5	14,7	14,9	14,9	15,1
6	15	15,1	15,2	15,4
7	14,9	15	15,2	15,3
8	14,9	14,9	15,1	15,4
9	15	15,1	15,2	1
10	15,3	15,1	14,8	1
11	15,1	15,2	15,3	1
12	15	15,2	15,2	1
13	15,2	14,6	15,1	1
14	15,1	14,6	15,2	1
15	15,1	14,9	15,2	1
16	15,2	14,9	15,2	1
17	15,1	14,8	14,8	1
18	15,1	14,8	15,3	1
19	15,4	15	15,5	
20	15,3	14,9	15,3	1
21				

Problém:
Znázornit rozložení hodnot tloušťky.

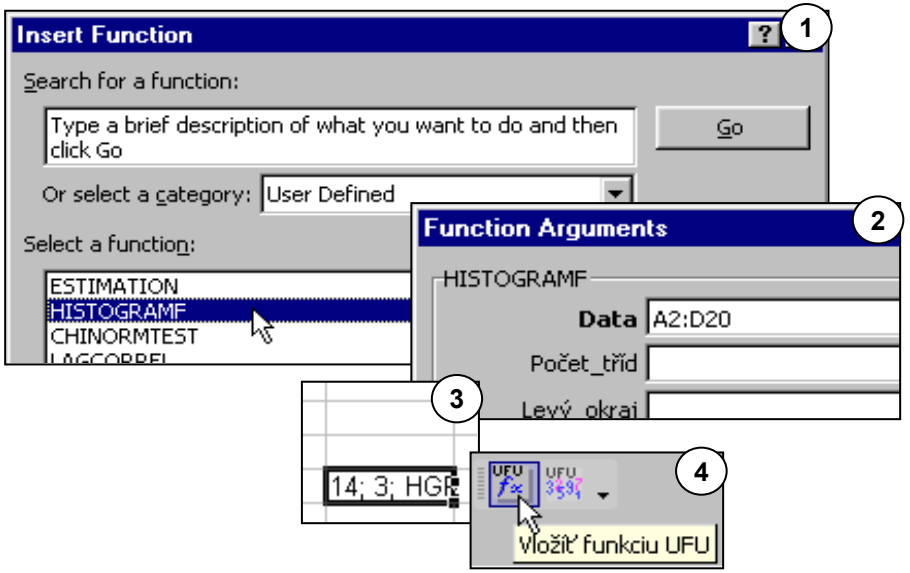
Řešení:
Rozložení hodnot lze velmi dobře znázornit histogramem.



Histogram znázorňuje rozložení hodnot prostřednictvím četností jejich výskytu v jistém počtu stejně širokých intervalů - **tříd**. Počet tříd je zaokrouhlenou odmocninou z počtu hodnot, nejméně však 5 a nejvíce 20. Do každé vnitřní třídy padnou hodnoty > dolní hranice a <= horní hranice třídy. Do levé krajní třídy padnou všechny hodnoty <= její horní hranice, do pravé krajní pak hodnoty > její dolní hranice. Histogram může být sestrojen jako kumulativní, kdy do každé třídy padnou hodnoty <= její horní hranice. Pravá krajní třída pak obsahuje všechny hodnoty. Zobrazení četností může být relativní, kdy se zobrazuje poměr počtu hodnot v každé třídě k celkovému počtu hodnot. **Levý okraj** a **Pravý okraj** nemají matematický smysl. Existují pouze pro jednodušší zadávání do funkce **Histogram**.

Histogram můžete sestrojit pomocí nástroje EXCELu **Analýza dat** - **Histogram**. Dále můžete použít funkce **FREQUENCY**, popsané na listu 15. V obou případech musíte předem připravit hranice tříd. Pokud máte nainstalovaný doplněk **UFU**, můžete použít jeho funkci **HISTOGRAMF**. Výstupem funkce **HISTOGRAMF** je pole.

Sestrojení histogramu



14; 3; HGR		
Avg	15,0776	
Std	0,23071	
Count	76	
Value	Abs. freq.	Bord.
14,6	4	14,6563
14,7125	2	14,7688
14,825	7	14,8813
14,9375	10	14,9938
15,05	23	15,1063
15,1625	13	15,2188
15,275	7	15,3313
15,3875	6	15,4438
15,5	4	

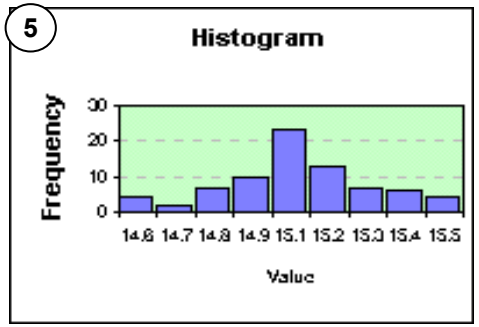
Počet vstupních hodnot

Směrodatná odchylka

Průměr ze vstupních hodnot

V této buňce bude v případě kumulativního histogramu označení **Cumul.**

V záhlaví sloupce četností bude v případě relativních četností označení **Rel. freq.**, anebo **Dens.** v případě hustoty pravděpodobnosti



Grafický test normality

	A	B	C	D
1	**Tloušťka, náh. výběr z 12.02**			
2	15,10	15	15,20	15,40
3	14,90	15,1	15,20	15,50
4	15,00	15	15,10	15,40
5	14,70	14,9	14,90	15,10
6	15,00	15,1	15,20	15,40
7	14,90	15	15,20	15,30
8	14,90	14,9	15,10	15,40
9	15,00	15,1	15,20	15,50
10	15,30	15,10	14,80	14,70
11	15,10	15,20	15,30	15,50
12	15,00	15,20	15,20	15,40
13	15,20	14,60	15,10	14,60
22				
23	**Tloušťka, náh. výběr z 1.03**			
24	14,80	14,80	15,00	15,20
25	14,90	14,90	15,10	15,20
26	14,80	15,10	15,10	15,40
27	15,00	14,90	15,10	15,30
28	14,90	14,90	15,10	15,20
29	14,70	14,90	14,90	15,20
30	14,80	14,80	15,00	15,20
31	14,80	14,80	15,00	15,10
32	14,80	14,90	15,00	15,20
33	15,00	14,90	15,1	15,20

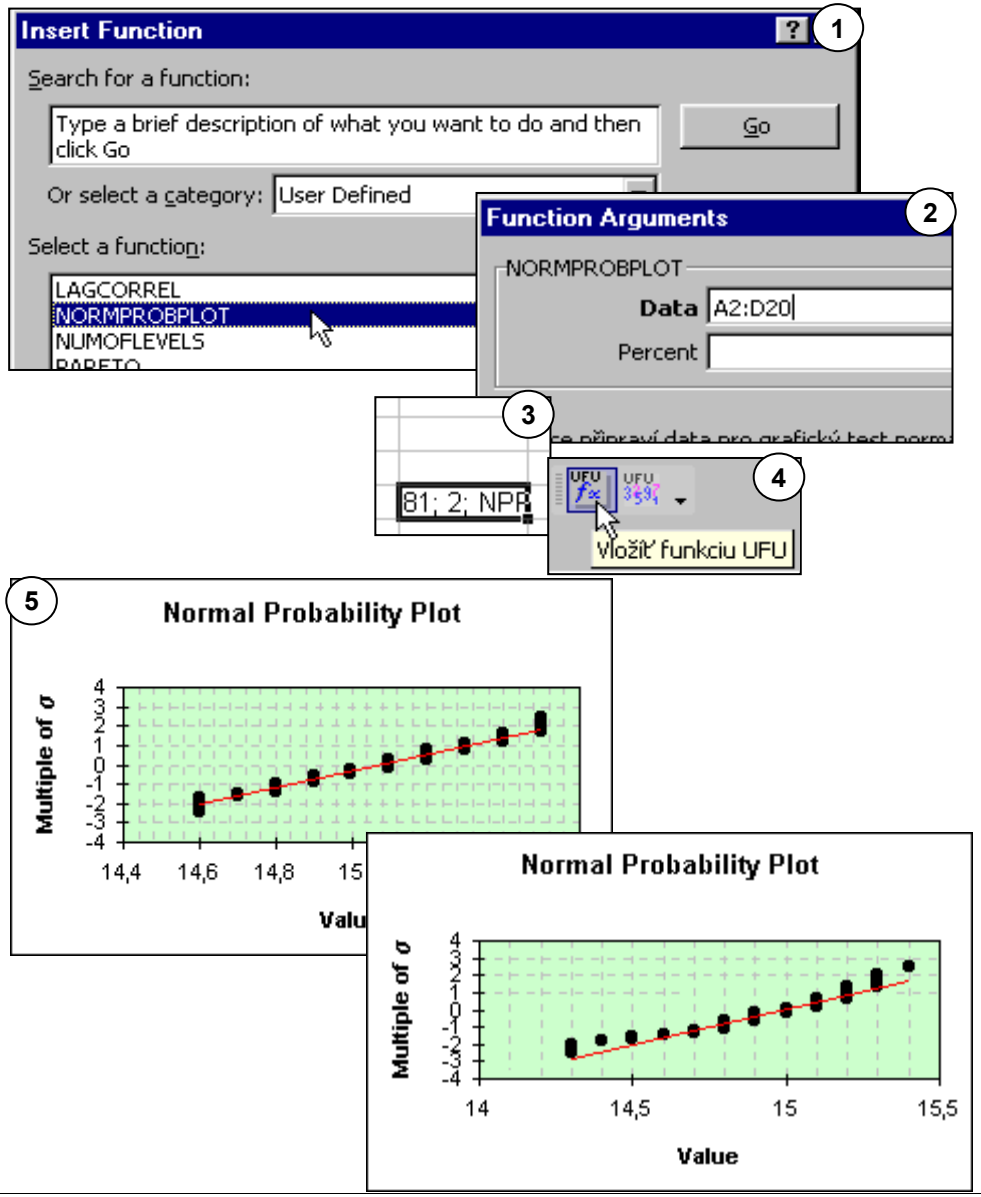
Problém:

Je předpoklad, že data z 12.02 a z 1.03 pocházejí z normálního rozdělení. Je možné tento předpoklad zamítnout, nebo ne?

Řešení:

Použijte funkci UFU **NormProbPlot**.

Sestrojení grafu

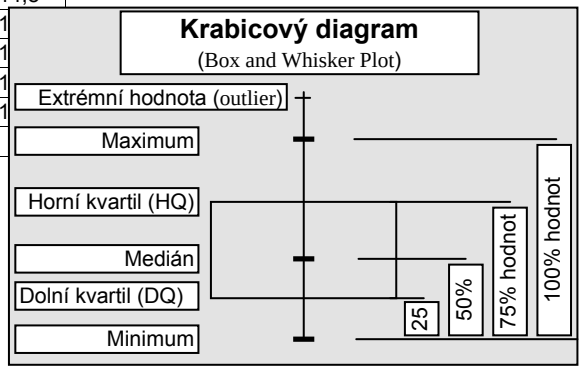


Krabicový diagram (Box and Whisker Plot)

	A	B	C	D
1	T1	T2	T3	T4
2	15,0	14,4	15,4	15,1
3	15,3	14,8	14,8	15,0
4	15,1	14,8	14,7	14,8
5	14,6	15,0	14,8	15,2
6	15,2	15,4	15,0	15,3
7	14,6	14,8	15,1	14,8
8	15,0	14,9	14,9	14,9
9	15,1	14,9	14,7	15,9
10	15,4	15,3	15,1	15,2
11	15,2	15,0	14,8	14,7
12	15,1	14,6	15,2	15,0
13	14,4	15,3	15,2	15,1
14	15,2	14,9	15,3	15,2
15	14,9	15,2	14,3	15,0
16	15,4	14,5	14,7	15,0
17	15,3	15,1	14,6	14,8
18	15,3	14,9	15,7	1
19	15,9	15,3	14,7	1
20	15,0	14,8	15,3	1
21	15,3	15,6	15,0	1
22				

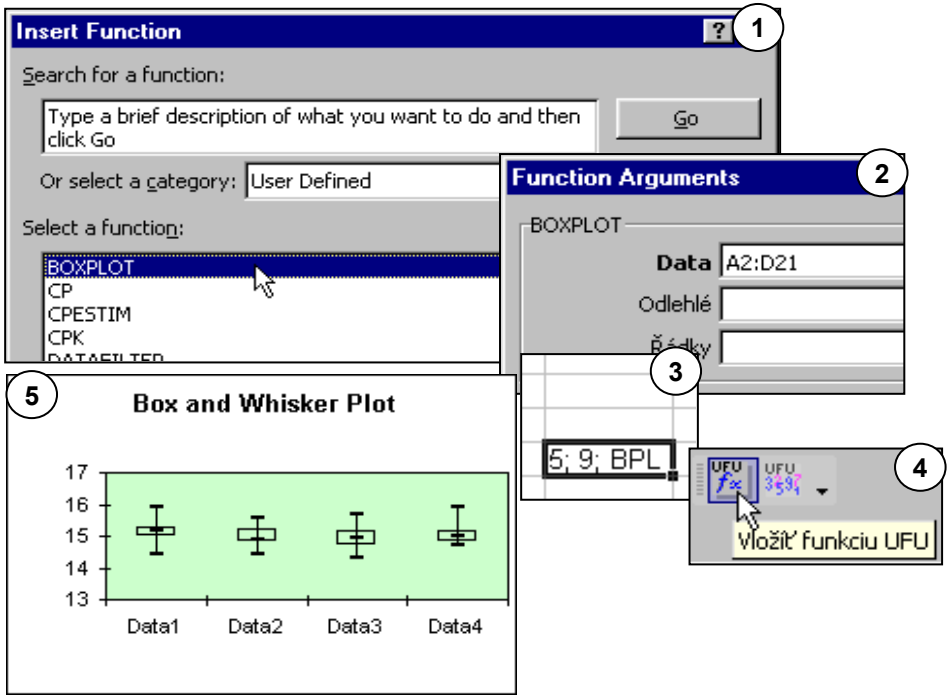
Problém:
Čtyři skupiny dat z tabulky je třeba zobrazit takovým způsobem, aby byl z grafu patrný jednak vzájemný vztah skupin, jednak rozložení hodnot v každé skupině.

Řešení:
Každou skupinu dat zobrazit krabicovým diagramem. Ten je obrazem rozdělení dat ve skupině, vzájemná poloha diagramů pak vyjadřuje vztah mezi skupinami. Pro sestrojení grafu použijte funkci UFU **BoxPlot**.



Krabicový diagram zobrazuje rozložení hodnot ve čtyřech částech po 25% hodnot. Hranice těchto oblastí jsou kvartily. Krabice (box) znamená dvě střední části a krabici přesahující úsečky (whiskers - vousy) zobrazují dvě krajní 25% části. Nesymetrické uspořádání hranic částí - kvartilů ukazuje na nesymetrické rozdělení. Extrémní hodnoty (outliers) jsou hodnoty větší jako HQ + 1,5 x (HQ - DQ), anebo menší jako DQ - 1,5 x (HQ - DQ). Hodnoty pro sestrojení grafu jsou pak přepočítány bez těchto extrémních hodnot. Funkce **BoxPlot** nevyhodnocuje extrémní hodnoty.

Sestrojení grafu



5; 13	25%	Min.	Max.	50%	OutMinOr	OutMaxO	75%	AVG	STD	Count	Outli	
T1	15					15,9	15,3	15,1111	0,2			19
T2	15	14,4	15,6	14,9	14,9	14,9	15,225	14,975	0,30758	20		
T3	15	14,3	15,7	14,95	14,95	14,95	15,2	14,965	0,32811	20		
T4	15	14,7	15,4	15	15	15,9	15,2	15,0158	0,21412	19	9	

Počet řádků a sloupců výstupní oblasti

Četnosti skupin

Názvy skupin

Hodnoty kvartilů

Průměry skupin

Oblast hodnot pro graf

Extrémní hodnoty (outliers), pokud se vyskytují, jinak medián

Směrodatné odchylky skupin

Pořadová čísla extrémních hodnot (outliers) ve skupinách. Záporná čísla znamenají extrémní hodnoty na straně minima.

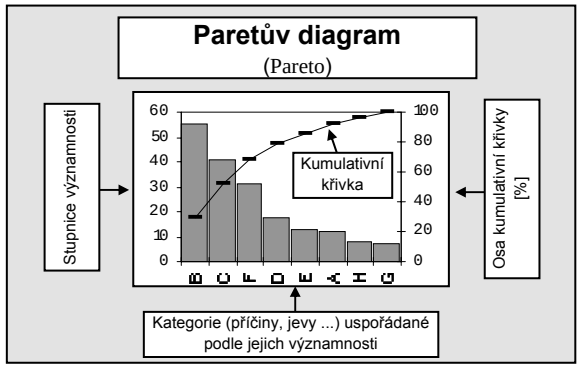
Z extrémních hodnot (otliers) jsou zobrazeny grafem pouze největší anebo nejmenší. !

Paretův diagram (Pareto Chart)

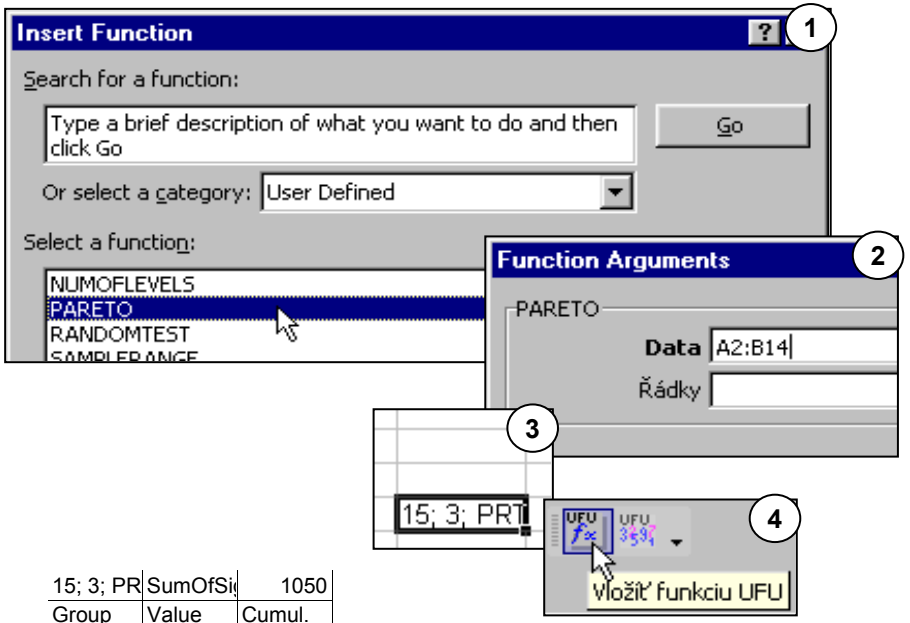
	A	B	C
1	TYP	POČET	
2	MA78	500	
3	MA7809	2	
4	MA7812	3	
5	MA78L	1	
6	MC78	63	
7	MC7805	3	
8	MC782	2	
9	MC78L	97	
10	MC78M	68	
11	MC79	112	
12	MC7912	7	
13	MC79L	2	
14	TL431	190	
15			

Problém:
Znázornit přehled typů podle jejich jejich podílu na vytěžení výrobní linky.

Řešení:
Na zobrazení je vhodné použít **Paretův diagram**. Pro přípravu dat použijte funkci UFU **Pareto**.



Sestrojení grafu

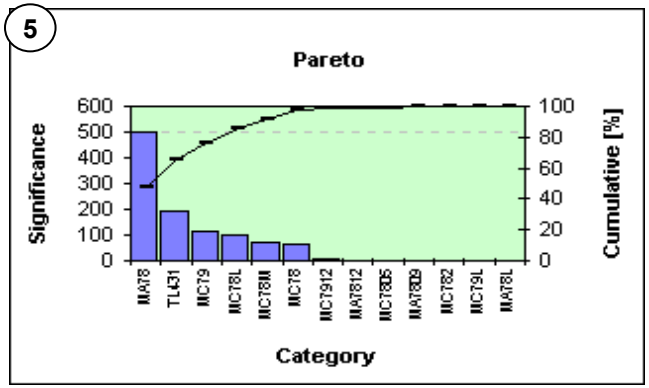


15; 3; PR	SumOfSi	1050
Group	Value	Cumul.
MA78	500	47,619
TL431	190	65,7143
MC79	112	76,381
MC78L	97	85,619
MC78M	68	92,0952
MC78	63	98,0952
MC7912	7	98,7619
MA7812	3	99,0476
MC7805	3	99,3333
MA7809	2	99,5238
MC782	2	99,7143
MC79L	2	99,9048
MA78L	1	100

Názvy kategorií

Významnosti

Kumulativní hodnoty



Zobrazení funkce dvou proměnných

Problém:
Experimentem byla pro tloušťku vrstvy zjištěna rovnice h=14,38+0,064*P-0,016*F_H , kde h je tloušťka v µ, P je výkon v kW a F_H je průtok vodíku v l/min. Je třeba sestrojit graf závislosti tloušťky, pro P v rozsahu 105 až 115 kW a průtok vodíku od 260 do 380 l/min.

Řešení:
Připravte datovou tabulku hodnot. Pak hodnoty zobrazte 3D grafem.

Sestrojení datové tabulky

	A	B	C	D	E	F	G	H	I
1	**P =**	105	**kW**						
2	**F_H =**	270	**l/min**						
3									
4						Průtok H2 [l/min]			
5	**h =**	16,78	**260**	**280**	**300**	**320**	**340**	**360**	**380**
6		**105**							
7		**106**							
8		**107**							
9		**108**							
10		**109**							
11		**110**							
12		**111**							
13		**112**							
14		**113**							
15		**114**							
16		**115**							
17									

=14,38+0,064***B1**-0,016*****B2**
vzorec v levém horním rohu tabulky

Hodnoty pro průtok vodíku

Výkon [kW]

Hodnoty pro výkon

Další postup sestrojení datové tabulky je popsaný na listu 16.

Pro účel vytvoření 3D grafu funkce dvou proměnných je důležité, aby řady hodnot pro obě proměnné byly tvořeny aritmetickými řadami. 3D graf v EXCELu umí na osách x, y zobrazit pouze pořadí hodnot a ne jejich polohu. Hodnoty v záhlaví tabulek budou zobrazeny pouze jako popisky os.

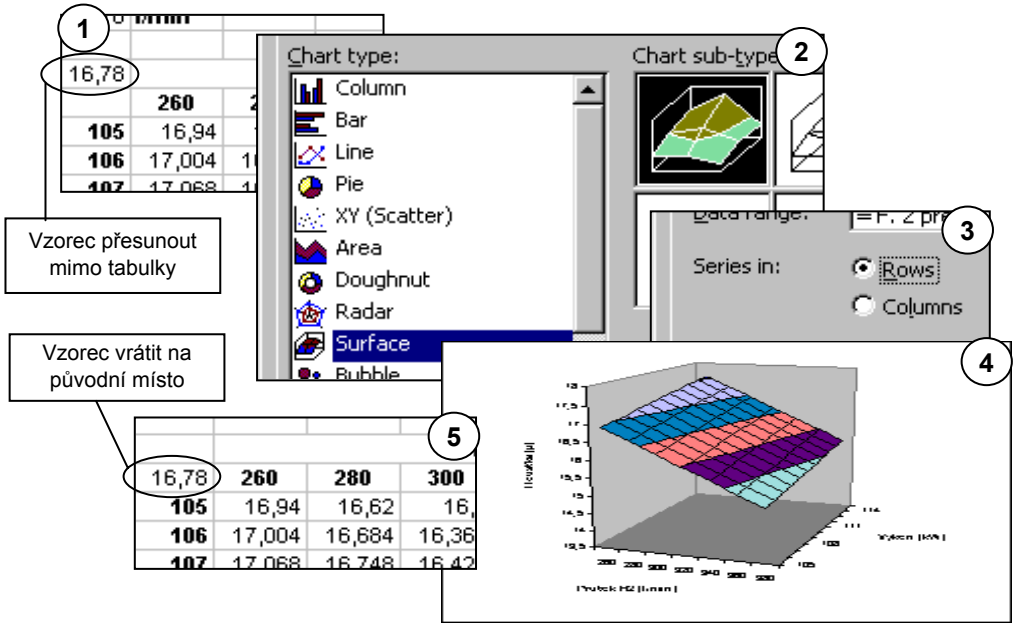
Pro jednoduché sestrojení grafu je důležité, aby v levém horním rohu tabulky vstupních dat byla během sestrojování grafu prázná buňka. !

	Průtok H2 [l/min]							
16,78	**260**	**280**	**300**	**320**	**340**	**360**	**380**	
105	16,94	16,62	16,30	15,98	15,66	15,34	15,02	
106	17,00	16,68	16,36	16,04	15,72	15,40	15,08	
107	17,07	16,75	16,43	16,11	15,79	15,47	15,15	
108	17,13	16,81	16,49	16,17	15,85	15,53	15,21	
109	17,20	16,88	16,56	16,24	15,92	15,60	15,28	
110	17,26	16,94	16,62	16,30	15,98	15,66	15,34	
111	17,32	17,00	16,68	16,36	16,04	15,72	15,40	
112	17,39	17,07	16,75	16,43	16,11	15,79	15,47	
113	17,45	17,13	16,81	16,49	16,17	15,85	15,53	
114	17,52	17,20	16,88	16,56	16,24	15,92	15,60	
115	17,58	17,26	16,94	16,62	16,30	15,98	15,66	

Výkon [kW]

Sestrojení grafu

Obsah buňky **B5** přesuňte mimo tabulky, označte oblast **B5:I16** a pomocí **Chart Wizard** sestrojte graf typu **Surface**. Po dokončení grafu přesuňte vzorec na původní místo.



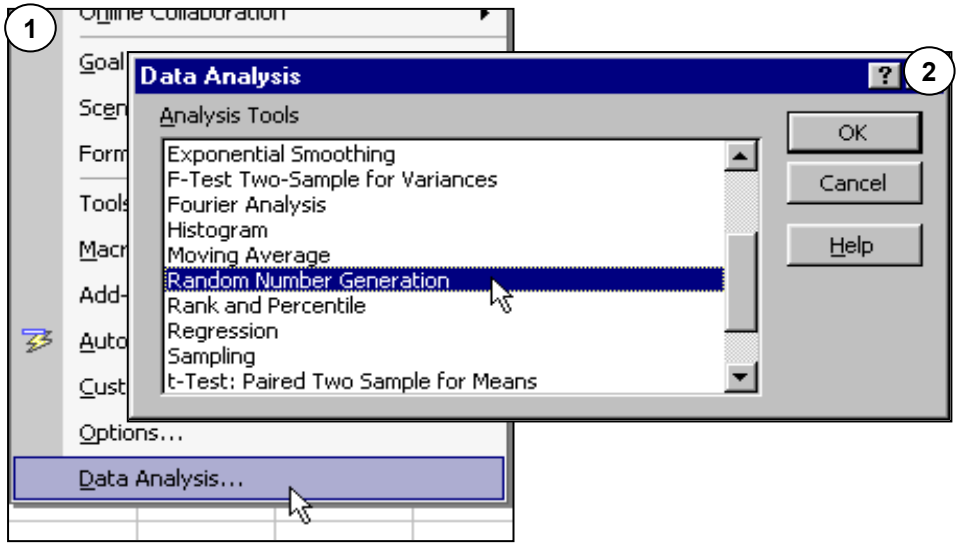
Generátor pseudonáhodných čísel (Random ...)

Problém:
Pro ověření funkce statistických nástrojů vytvořit množinu 100 čísel, která bude reprezentovat výběr z normálního rozdělení se střední hodnotou 100 a směrodatnou odchylkou 10.

Řešení:
Na vytvoření množiny náhodných čísel s definovaným rozdělením použijte nástroj **Random Number Generation**.

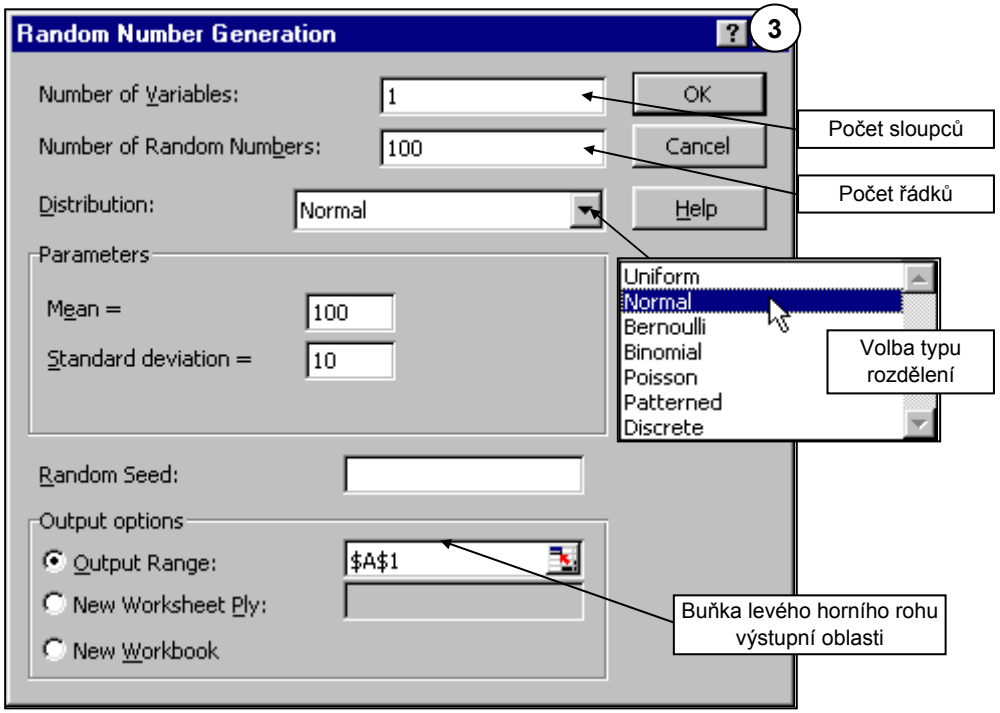
Použití generátoru

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Random Number Generation**.



V dialogovém okně **Random Number Genaration** zadejte vstupní hodnoty podle obrázku:

Number of Variables - počet sloupců výstupu
Number of Random Numbers - počet řádků výstupu
Mean - střední hodnota
Standard Deviation - směrodatná odchylka
V rámečku **Output Options** zvolte **Output Range** a zadejte buňku levého horního rohu výstupní oblasti.



Po klepnutí na **OK** bude oblast **A1:A100** vyplněna pseudonáhodnými čísly, která reprezentují náhodný výběr z normálního rozdělení se střední hodnotou 100 a směrodatnou odchylkou 10.

Random Seed (základ generátoru) může nabývat hodnot 1 - 32767 a ovlivňuje start generátoru. Při stejném základu generátoru dostanete stejná pseudonáhodná čísla.

Vzorkování (Sampling)

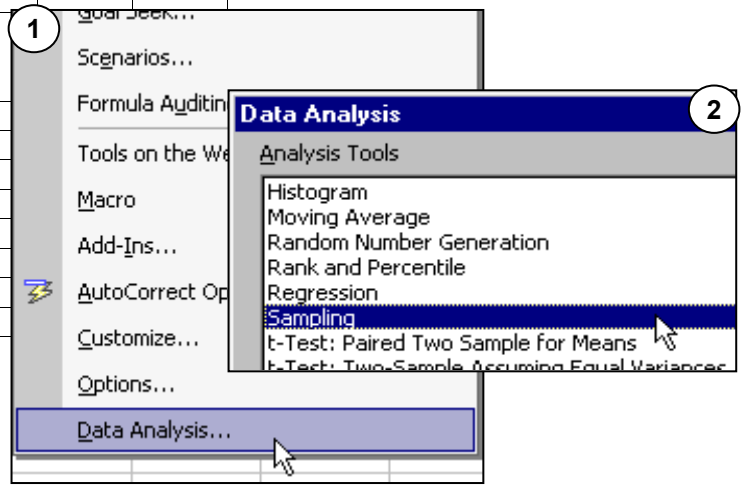
	A	B	C	D
1	T			
2	15,5			
3	15,6			
4	16			
5	15,5			
6	15,7			
7	15,5			
8	15,5			
9	15,5			
10	15,5			
11	15,4			
12	15,4			
13	15,4			
14	14,9			
15	15,2			
993	15,1			
994	15,1			
995	15,1			
996	15,3			
997	15			
998	15			
999	15,1			
1000	15			
1001	15,5			

Problém:
Z 1000 hodnot získaných z technologického deníku náhodně vybrat 100 hodnot pro test normality

Řešení:
Nástroj **Sampling** umožňuje provést náhodný výběr.

Použití nástroje Vzorkování

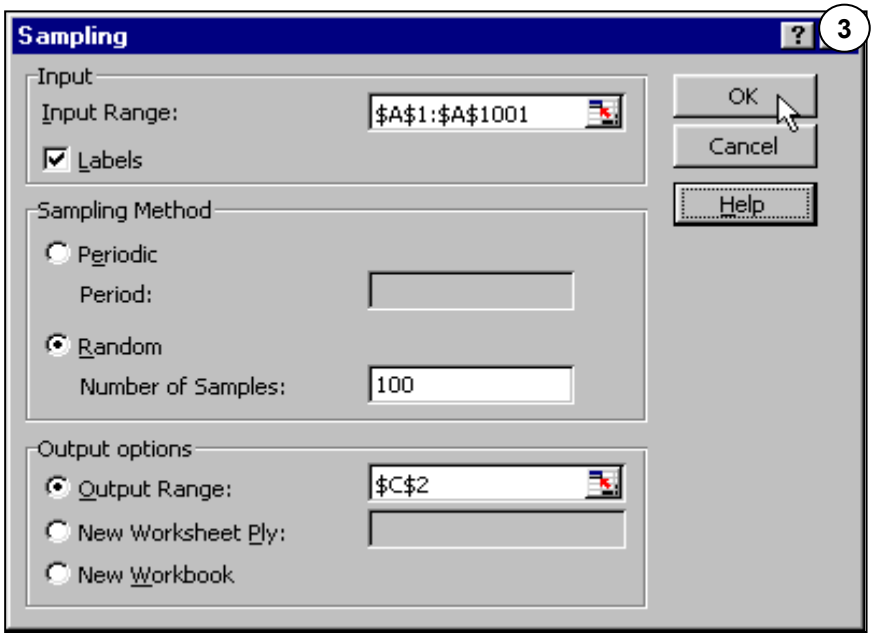
Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Sampling**.



V dialogovém okně **Sampling** zadejte vstupní hodnoty podle obrázku:

Input
Labels (označte, pokud jste zadali vstupní oblast včetně záhlaví)
Sampling Method (zvolte Random)

V rámečku **Output Options** zvolte **Output Range** a zadejte buňku levého horního rohu výstupní oblasti.



Nástroj **Sampling** provede vzorkování s opakováním. !

Jiná možnost

	A	B	C
1	T		
2	15,5	0,70002	
3	15,6	0,87532	
###	15	0,70997	
###	15,5	0,36081	

=RAND()

=RAND()

Do sloupce **B** vedle čísel vložte funkci **RAND**(). Pak označte oblast **A1:B1001** a pomocí nástroje **Seřadit** uspořádejte hodnoty podle sloupce **B**. Nyní vyberte ze sloupce **A** prvních 100 hodnot.

Tímto způsobem provedete náhodný výběr bez opakování.

Pořadová statistika a percentily (Rank ...)

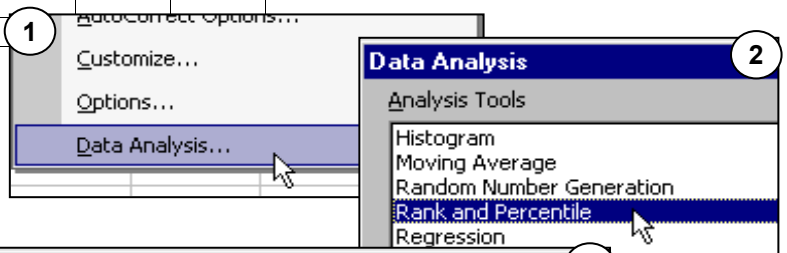
	A	B	C	D
1	T			
2	15,1			
3	14,9			
4	15			
5	14,7			
6	15			
7	14,9			
8	14,9			
9	15			
10	15,3			
11	15,1			
75	14,8			
76	15			
77	14,8			
78				

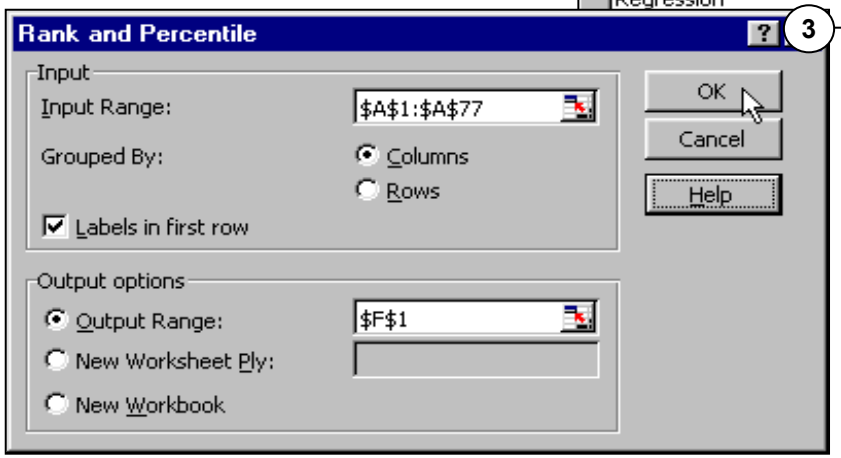
Problém:
Graficky znázornit distribuční funkci hodnot získaných náhodným výběrem.

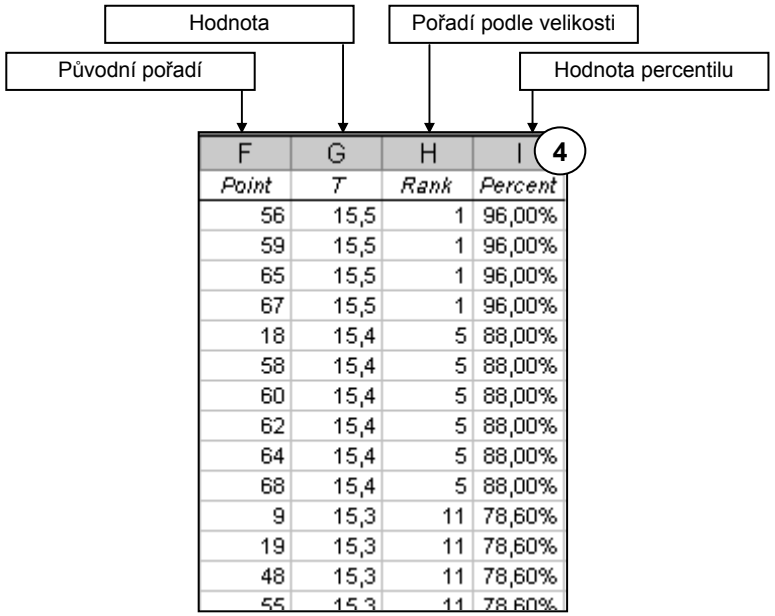
Řešení:
Pomocí nástroje **Rank and Percentile** můžete vypočítat hodnoty výběrové distribuční funkce a pak ji graficky znázornit.

Použití nástroje Pořadová statistika a percentily

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Rank and Percentile** a zadejte hodnoty podle obrázku.

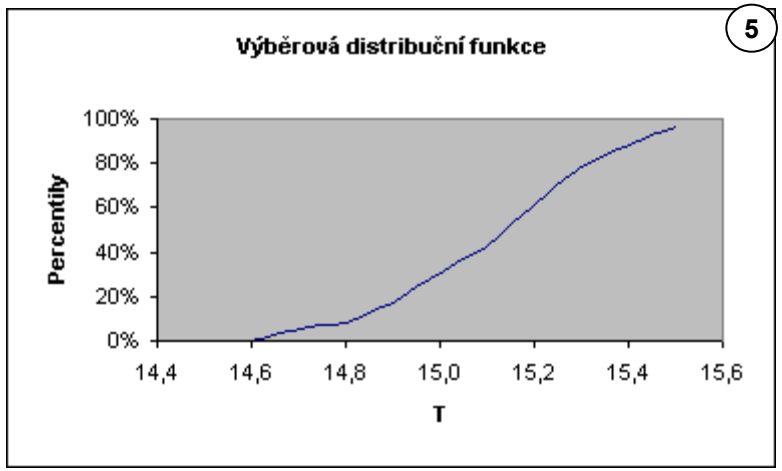






F	G	H	I
Point	*T*	*Rank*	*Percent*
56	15,5	1	96,00%
59	15,5	1	96,00%
65	15,5	1	96,00%
67	15,5	1	96,00%
18	15,4	5	88,00%
58	15,4	5	88,00%
60	15,4	5	88,00%
62	15,4	5	88,00%
64	15,4	5	88,00%
68	15,4	5	88,00%
9	15,3	11	78,60%
19	15,3	11	78,60%
48	15,3	11	78,60%
55	15,3	11	78,60%

Označte oblasti **G2:G77** a **I2:I77**. Pomocí **Chart Wizard** sestrojte **XY (Scatter)**.



Popisná statistika (Descriptive Statistics)

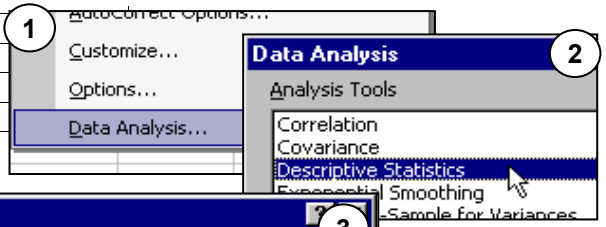
	A	B	C	D
1	T1	T2		
2	15,1	15,1		
3	15,3	15,3		
4	15,6	15,5		
5	15	15		
6	15,1	15,2		
7	15,2	15,1		
8	15,1	15,1		
9	15,1	15,2		
10	15,1	15,10		
11	15,1	15,00		
12	14,9	14,80		
13	14,9	14,70		
14	14,4	14,50		
15	14,8	14,90		
16				

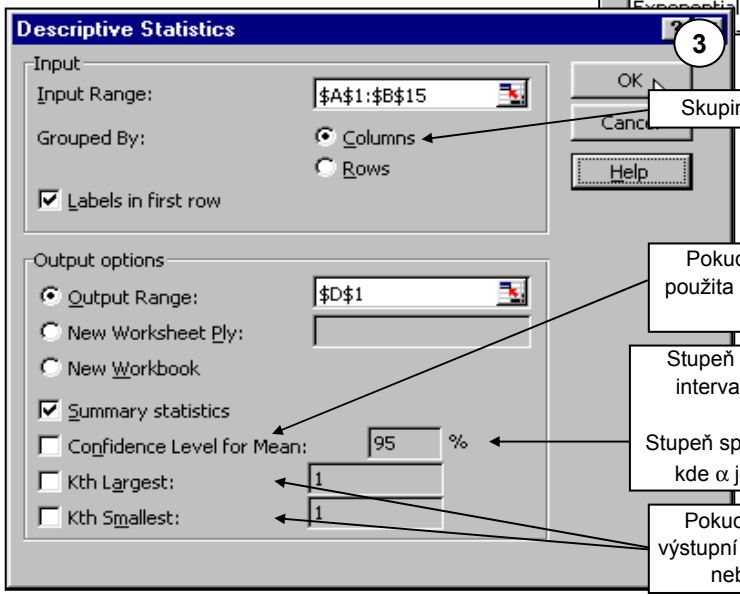
Problém:
Při výpočtech statistických charakteristik může být vkládání funkcí, označování oblastí atd. zdlouhavé.

Řešení:
Nástrojem **Descriptive Statistics** můžete najednou získat základní statistické charakteristiky jedné, nebo více skupin dat.

Použití nástroje Popisná statistika

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Descriptive Statistics**.





Skupiny dat jsou ve sloupcích

Pokud není zaškrtnuto, bude použita hodnota 95%, jinak podle zadání

Stupeň spolehlivosti pro určení intervalu spolehlivosti střední hodnoty.
Stupeň spolehlivosti = 100(1 - α)%, kde α je hladina významnosti

Pokud je zaškrtnuto, bude ve výstupní tabulce také k-tá největší, nebo nejmenší hodnota

$\frac{\text{sm. odchylka}}{\sqrt{\text{počet}}}$

Výsledky pro T1

Výsledky pro T2

4

D	E	F	G
T1		T2	
Mean	15,05	Mean	15,035714
Standard Error	0,0716754	Standard Error	0,0676181
Median	15,1	Median	15,1
Mode	15,1	Mode	15,1
Standard Deviation	0,2681848	Standard Deviation	0,2530039
Sample Variance	0,0719231	Sample Variance	0,064011
Kurtosis	2,7686496	Kurtosis	0,6799335
Skewness	-0,5164449	Skewness	-0,4366418
Range	1,2	Range	1
Minimum	14,4	Minimum	14,5
Maximum	15,6	Maximum	15,5
Sum	210,7	Sum	210,5
Count	14	Count	14

Kurtosis (Špičatost)

Vyjadřuje relativní špičatost rozdělení, vhledem k normálnímu rozdělení. Kladná **Špičatost** znamená, že rozdělení je strmější, jako normální (hodnoty jsou více soustředěny kolem střední hodnoty). Naopak, záporná **Špičatost** znamená příliš ploché rozdělení. Stejnou hodnotu vrátí funkce **KURT**.

Skewness (Šikmost)

Vyjadřuje asymetrii rozdělení. Kladná **Šikmost** znamená posun k větším hodnotám, záporná posun k menším hodnotám. Stejnou hodnotu vrátí funkce **SKEW**.

V jiných programech (nebo literatuře) mohou být výsledky pro špičatost a šikmost poněkud odlišné. Většinou se liší pouze o konstantu. !

Korelace (Correlation)

	A	B	C	D
1	t1	t2	T	
2	97,5	1000,4	1054,1	
3	93,6	997,3	1000,0	
4	93,1	1004,0	1001,3	
5	93,2	997,8	1011,7	
6	95,7	1000,7	1027,9	
7	97,7	997,6	1037,2	
8	92,1	999,1	995,4	
9	93,7	997,8	1021,0	
10	97,1	1001,2	1029,3	
11	94,9	998,9	1003,0	
12	92,8	1001,6	993,6	
13	96,1	999,0	1013,0	
14	93,1	999,9	1014,3	
15	94,7	994,5	1022,4	
16	92,5	997,6	993,5	
17				

Problém:
Zjistit, zda je velká variabilita tloušťky `T` způsobená změnou teploty `t1` nebo `t2` anebo ani jednou z nich.

Řešení:
Použít nástroje **Correlation** k vytvoření korelační matice. Podle velikostí korelačních koeficientů posoudit závislosti. Korelační koeficient dvou skupin dat vypočítá také funkce **CORREL**.

Korelační koeficient **R** vyjadřuje vzájemnou závislost dvou skupin dat.

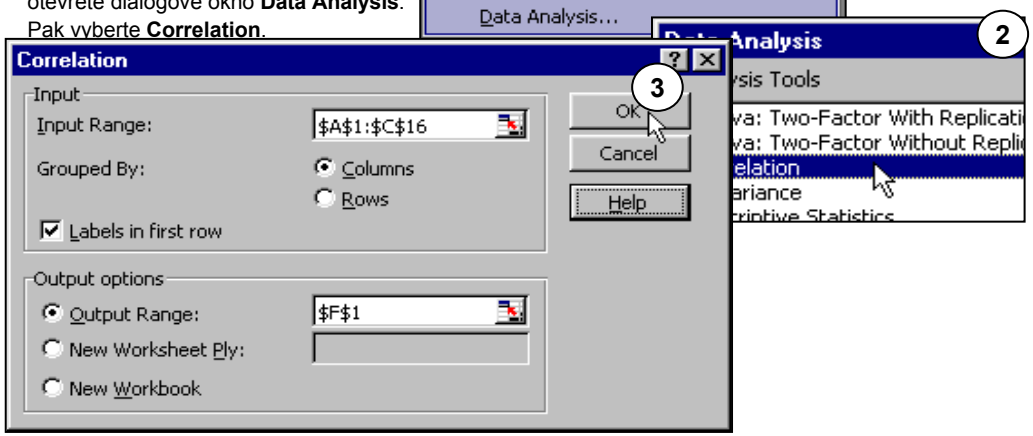
R=0 skupiny jsou nezávislé
R=1 nebo **R**=-1 skupiny jsou lineárně závislé

Test hypotézy H_0, že dvě skupiny dat jsou vzájemně nezávislé (pro rozsah výběru n > 10, α=0,05).

1. Vypočítat **R** a pak hodnotu z´=**FISHER**(**R**)
2. Hypotézu H_0 zamítáme pokud **SQRT**(n-3)*z´ > 1,96
3. Hypotézu H_0 obvykle zamítáme když $\mathbf{R}^2$>0,8

Použití nástroje Korelace

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Correlation**.



	F	G	H	4
		t1	t2	T
t1		1		
t2		0,007	1	
T		0,8455	-0,0505	1

Na obrázku vidět, že pouze hodnota 0,8455 může znamenat významnou závislost.
$\mathbf{R}^2$=0,7148, proveďte tedy test hypotézy H_0, že `T` a `t1` jsou nezávislé.

z´=**FISHER**(0,8455)=1,24 **SQRT**(15-3)*z´=4,3
4,3 > 1,96 proto se H_0 zamítá.
Výsledek potvrdil, že variabilitu tloušťky výrazně ovlivňuje právě změna teploty `t1`.

Další možnosti

	I	1
1	T	
2	13,69	
3	14,02	
4	14,49	
100	13,93	
101	14,56	

Problém:
Zjistit, zda v časové řadě není závislost mezi předcházejícími a následujícími hodnotami - autokorelace.

Řešení:
Zkopírujte hodnoty z oblasti **I2:I101** postupně do **H3:H102**, **G4:G103**, ... **A10:A109**. Potom pomocí nástroje **Correlation** vytvořte korelační matici pro vstupní oblast **A10:I101** do výstupní oblasti začínající v buňce **K1**. (tentokrát nezaškrtávejte políčko **Záhlaví v prvním řádku**).

2	A	B	C	D	E	F	G	H	I
1									T
2									13,69
3								13,69	14,02
4							13,69	14,02	14,49
5									
6					13,69				
7				13,69	14,02				
8			13,69	14,02	14,49				
9		13,69	14,02	14,49	14,77				
10	13,69	14,02	14,49	14,77	15,08				
11	14,02	14,49	14,77	15,08	16,14				
12									

3	K	L	M	N	O
		Column 1	Column 2	Column 3	Colum
	Column 1	1			
	Column 2	0,29354	1		
	Column 3	-0,201	0,28019	1	
	Column 4	-0,3898	-0,2138	0,27385	
	Column 5	-0,4381	-0,3981	-0,2169	0,27
	Column 6	-0,2072	-0,4406	-0,3999	-0,2
	Column 7	0,26421	-0,1867	-0,4309	-0,3
	Column 8	0,76168	0,27226	-0,1781	-0,4
	Column 9	0,27321	0,75598	0,26105	-0,1

Ve výslední matici je vidět že je výrazná korelace mezi sloupci 1 a 8, 2 a 9 atd.

Regrese (Regression)

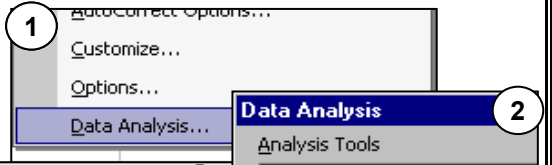
	A	B	C	D
1	t1	T		
2	97,5	1054,1		
3	93,6	1000,0		
4	93,1	1001,3		
5	93,2	1011,7		
6	95,7	1027,9		
7	97,7	1037,2		
8	92,1	995,4		
9	93,7	1021,0		
10	97,1	1029,3		
11	94,9	1003,0		
12	92,8	993,6		
13	96,1	1013,0		
14	93,1	1014,3		
15	94,7	1022,4		
16	92,5	993,5		

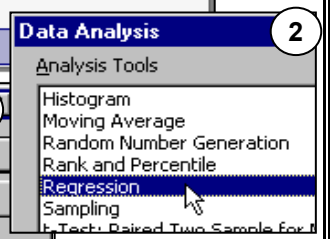
Problém:
Vytvořit lineární model závislosti T na `t1` a zjistit jeho statistickou významnost.

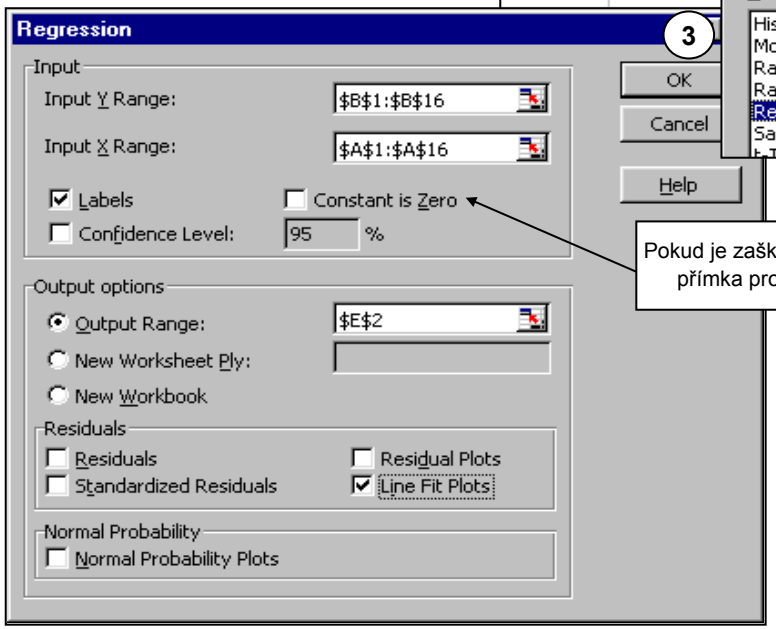
Řešení:
Použijte statistický nástroj **Regrese**.

Použití nástroje Regrese

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Regression**.







Pokud je zaškrtnuto, bude regresní přímka procházet počátkem

4

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0,84546
R Square	0,7148
Adjusted	0,69286
Standard	9,77599
Observati	15

Korelační koeficient **R**

R^2

Upravené R^2 (Adj R^2)

Reziduální odchylka

Počet stupňů volnosti (dF)

p - hodnota Fisherova rozdělení pro F

ANOVA	*df*	*SS*	*MS*	*F*	*Significance F*
Regressi	1	3113,79	3113,79	32,5813	7,2E-05
Residual	13	1242,41	95,5699		
Total	14	4356,2			

$MS_{Regrese}$

$MS_{Rezidua}$

	Coefficients	*Standard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	267,892	130,827	2,04768	0,06135	-14,743	550,526	-14,743	550,526
t1	7,89909	1,38386	5,708	7,2E-05	4,90943	10,8887	4,90943	10,8887

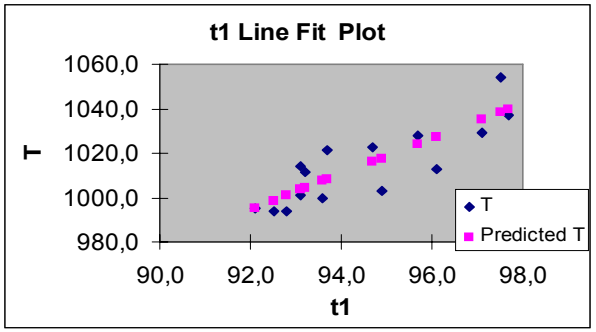
a

b

regresní přímka: $y = \alpha + \beta x$
odhad α = a odhad β = b

Hodnoty pro zadanou úroveň věrohodnosti, pokud bylo zaškrtnuto políčko **Úroveň věrohodnosti**.

Hranice intervalů spolehlivosti pro odhad parametrů regresní přímky



Statistická významnost modelu

Hypotéza H_0, že model je statisticky nevýznamný, se zamítá jestliže vypočtená p - hodnota (**Significance F**) je menší, jako předem zvolená hladina významnosti (0,05).

Vypočtený model platí pro průměry, přičemž konstanty a, b jsou odhadem parametrů regresní přímky (α, β). Hranice intervalů spolehlivosti pro tyto hodnoty jsou ve sloupcích *Lower 95%* a *Upper 95%*.

F - Test

	A	B	C	D
1	T	T1	T2	
2	14,5	14,8	14,7	
3	14,5	14,8	14,4	
4	14,9	14,4	14,4	
5	14,7	15,9	15,8	
6	14,8	15,5	15,4	
7	15,2	15,1	14,2	
8	13,5	13,6	15,0	
9	14,3	14,9	15,5	
10	13,9	14,3		
11	15,4			
12	13,3			
13	15,1			
14	16,2			
15	15,6			
16	14,6			
17	16,6			
18	14,3			
19	15,9			
20				
21	0,878	0,675		

Problém:

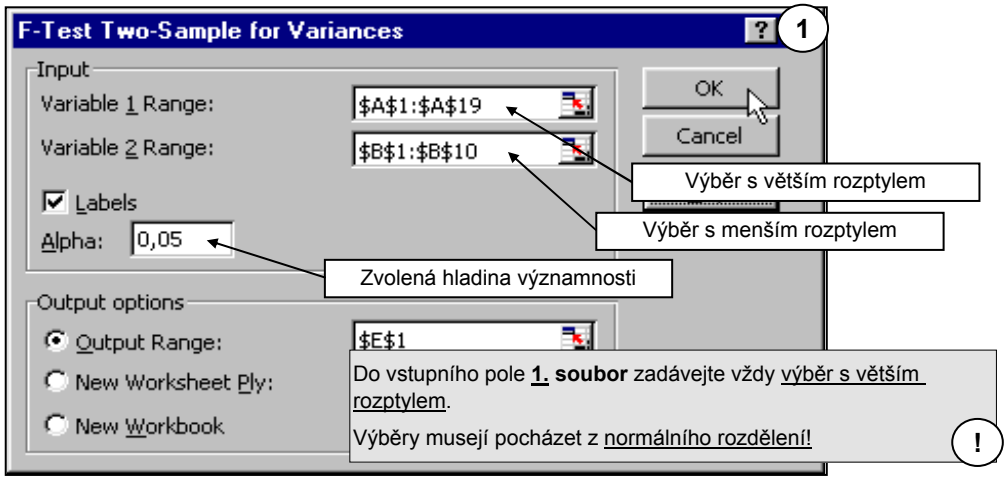
Ve sloupci **A** jsou hodnoty tloušťky a v buňce **A:21** je směrodatná odchylka. Protože pro výrobní účely je příliš velká, byly provedeny různé úpravy zařízení a pak technologické zkoušky. Výsledky jsou ve sloupcích **B** a **C**. Směrodatné odchylky v řádku **21** naznačují, že by mohlo jít o zlepšení. Nyní je nutno ověřit, jestli změny směrodatné odchylky jsou výsledkem technologických úprav nebo jsou pouze náhodné.

Řešení:

Statistický nástroj **F - Test** umožňuje provést test hypotézy, že dva výběry pocházejí z rozdělení se stejným rozptylem. Postavte hypotézu H_0: výběry T a T1 pocházejí z rozdělení se stejným rozptylem a hypotézu H_1: výběr T1 pochází z rozdělení s menším rozptylem, jako výběr T. Podobně pro T2. Když pomocí F - testu zamítnete hypotézu H_0, můžete považovat zlepšení směrodatné odchylky za výsledek technologických úprav.

Použití nástroje F - Test

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **F-Test Two -Sample for Variances**.



F-Test Two-Sample for Variances

2

	T	T1
Mean	14,85	14,8111
Variance	0,77088	0,45611
Observati	18	9
df	17	8
F	1,69012	
P(F<=f) o	0,22828	
F Critical	3,1867	

Odhady středních hodnot

Odhady rozptylů

Počty hodnot

Počty stupňů volnosti

Hypotéza H_0 se zamítá, když F > F krit (1), anebo P(F<=f) (1) je menší jako zvolená hladina významnosti.
Dvoustranní p-hodnotu vrátí také funkce **FTEST**, anebo UFU funkce **XFtest**.
P(F<=f)(1) = P(F<=f)(2) / 2

Hodnota testového kritéria

Jednostranní p-hodnota

Jednostranní kritická hodnota F-rozdělení

1,6901 < 3,1867, tedy hypotézu, že výběry T a T1 pocházejí z rozdělení se stejným rozptylem nemůžete zamítnout. Podobný výsledek je také pro T a T2 (2,5115 < 3,1867). Zdá se tedy, že snížení směrodatných odchylek pro T1 a T2 je pouze náhodné.

Na následujícím obrázku byly malé výběry T1 a T2 nahrazeny doplněnými výběry T1d a T2d. Nyní s těmito výběry opět proveďte F - test.

	A	B	C
24	T	T1d	T2d
25	14,5	14,8	14,7
26	14,5	14,8	14,4
27	14,9	14,4	14,4
28	14,7	15,9	15,8
29	14,8	15,5	15,4
30	15,2	15,1	14,2
31	13,5	13,6	15,0
32	14,3	14,9	15,5
33	13,9	14,3	14,9
34	15,4	17,0	14,5
35	13,3	13,5	14,9
36	15,1	15,7	14,7
37	16,2	14,9	14,7
38	15,6	15,5	15,4
39	14,6	14,9	15,3
40	16,6	14,5	14,4
41	14,3	16,1	14,7
42	15,9	15,5	15,3
43			
44	0,878	0,862	0,458

F-Test Two-Sample for Variances

	T	T1d
Mean	14,85	15,05
Variance	0,7709	0,7438
Observa	18	18
df	17	17
F	1,0364	
P(F<=f) c	0,4711	
F Critical	2,2719	

V případě T1d se hypotéza H_0 opět nezamítá. Také směrodatná odchylka odhadnuta z výběru T1d (0,862) je blízká původní hodnotě 0,878. Je patrné, že rozdílná hodnota odhadnutá z výběru T1 (0,675) byla pouze náhodní (protože výběr T1 je malý).

F-Test Two-Sample for Variances

	T	T2d
Mean	14,85	14,9
Variance	0,7709	0,2094
Observa	18	18
df	17	17
F	3,6812	
P(F<=f) c	0,0052	
F Critical	2,2719	

V případě T2d se hypotéza H_0 zamítá (3,6812>2,2719).
V tomto případě můžete pokládat zmenšení směrodatné odchylky za výsledek úpravy zařízení.

t - test

	A	B	C
1	T	T1	
2	15,7	15,6	
3	14,9	15,2	
4	14,5	16,0	
5	14,5	15,7	
6	14,8	15,0	
7	14,9	15,5	
8	15,1	15,2	
9	14,7	16,2	
10	15,0	15,6	
11	15,0	15,4	
12	14,6	16,1	
13	15,3	15,1	
14	15,8	15,5	
15	14,6	15,4	
16	15,5	15,1	
17			
18	14,993	15,507	
19			

Problém:

Po poruše zařízení vzniklo podezření, že byla zvětšená střední hodnota. Stav před poruchou je reprezentovaný výběrem T, stav po poruše pak výběrem T1. V řádku **18** jsou odhady středních hodnot. Je třeba ověřit, zda skutečně došlo ke zvýšení střední hodnoty, anebo je výsledek pouze náhodný.

Řešení:

Pro testování hypotézy o rovnosti středních hodnot, nebo jejich definovaném rozdílu slouží statistický nástroj **t - test**.

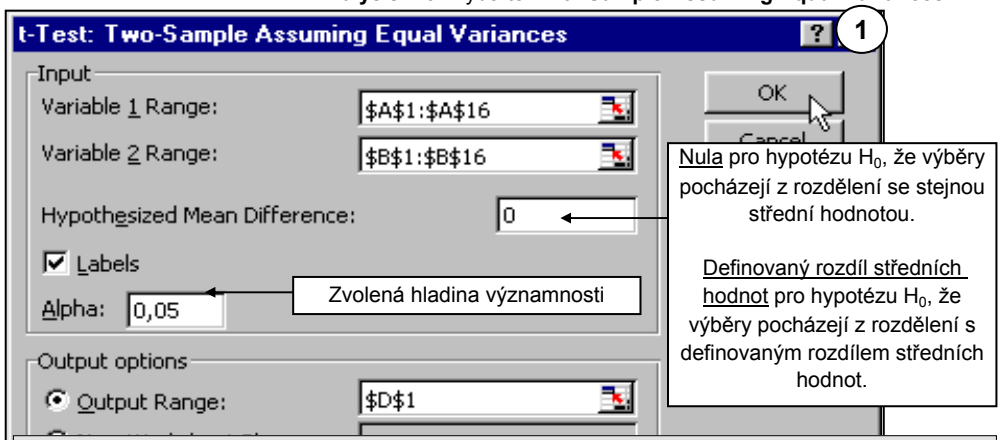
Postavte hypotézu H_0: výběry T a T1 pocházejí z rozdělení se stejnými středními hodnotami

a hypotézu H_1: výběr T1 pochází z rozdělení s větší střední hodnotou jako výběr T.

Když pomocí t - testu zamítnete hypotézu H_0, můžete prohlásit, že vlivem poruchy byla zvětšena střední hodnota.

Použití nástroje T - Test

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Two -Sample Assuming Equal Variances**.



Pokud bude F-testem zamítnuta hypotéza o rovnosti rozptylů, použijte nástroje **Two - Sample Assuming Unequal Variances** (to však vede ke snížení počtu stupňů volnosti).

Výběry musejí pocházet z normálního rozdělení! !

t-Test: Two-Sample Assuming Equal Variances (2)

	T	T1
Mean	14,9933	15,5067
Variance	0,17495	0,13781
Observati	15	15
Pooled V	0,15638	
Hypothes	0	
df	28	
t Stat	-3,555	
P(T<=t) o	0,00068	
t Critical	1,70113	
P(T<=t) tv	0,00137	
t Critical t	2,04841	

Odhady středních hodnot; Odhady rozptylů; Počty hodnot; Počet stupňů volnosti

Sdružený rozptyl (pooled variance)

Definovaný rozdíl mezi středními hodnotami

Hodnota testového kritéria

Jednostranní p-hodnota
p-hodnotu vrátí také funkce **TTEST**.

Jednostranní kritická hodnota T-rozdělení

Dvoustranné hodnoty

Hypotéza H_0 se zamítá, když t stat > t krit (t stat < -t krit), anebo P(T<=t) je menší jako zvolená hladina významnosti.

Když je H_1 definována tak, že jeden výběr pochází z rozdělení se střední hodnotou pouze větší (menší) jako střední hodnota druhého rozdělení pak se použijí jednostranné hodnoty. V případě definice H_1, že výběry pocházejí z rozdělení s různými středními hodnotami, se použijí dvoustranné hodnoty.

p-hodnoty vrátí také funkce **TTEST**, anebo UFU funkce **XTtest**.

V tomto případě použijte jednostranné hodnoty. `-3,555 < -1,7011`, tedy hypotézu, že výběry T a T1 pocházejí z rozdělení se stejnými středními hodnotami můžete zamítnout.

Byla provedena oprava zařízení. Po opravě by měla být střední hodnota snížena přibližně o 0,5. Stav po opravě je reprezentován výběrem T2.

Postavte hypotézu H_0: výběry T1 a T2 pocházejí z rozdělení se stejnými středními hodnotami a hypotézu H_1: výběry T1 a T2 pocházejí z rozdělení s různými středními hodnotami.

Když pomocí t - testu zamítnete hypotézu H_0, můžete prohlásit, že opravou byl dosažen posun hodnoty.

	A	B	C
21	T1	T2	
22	15,6	14,7	
23	15,2	15,7	
34	15,5	15,5	
35	15,4	14,5	
36	15,1	14,7	
37			
38	15,507	15,027	

	T1	
Mean	15,507	15
Variance	0,1378	0
Observa	15	
Pooled V	0,1664	
Hypothes	0	
df	28	
t Stat	3,2227	
P(T<=t) c	0,0016	
t Critical	1,7011	
P(T<=t) t	0,0032	
t Critical t	2,0484	

Položku **Hypotesized Mean Difference** zadejte 0.

Je vidět, že `3,2227 > 1,7011` ale také `3,2227 > 2.0484`. Hypotézu, že rozdíl mezi středními hodnotami je 0 můžete zamítnout, tedy oprava byla účinná. Sestrojením konfidenčních intervalů (funkce UFU **Estimation**) pro průměry ve výběrech T1 a T2 se přesvědčte, jakého posunu střední hodnoty bylo dosaženo.

Anova

	A	B	C	D
1	T1	T2	T3	T4
2	14,5	15,1	15,7	16,5
3	16,8	14,6	14,7	15,3
4	16,8	14,8	16,7	14,7
5	16,0	15,4	14,4	13,5
6	14,8	16,2	14,4	13,8
7	14,8	15,6	15,4	13,6
8	14,7	14,4	14,5	15,2
9	15,4	14,8	14,6	14,4
10	16,7	13,6	13,8	15,5
11	14,0	13,7	14,2	13,9
12	13,7	15,5	15,3	14,3
13	15,7	16,6	14,8	14,9
14	15,3	16,8	14,5	14,0
15	14,6	13,5	14,5	14,2
16	15,0	15,1	15,5	13,4
17				
18	15,253	15,047	14,867	14,480
19				

Problém:

Výběry T1, T2, T3, T4 pocházejí ze čtyř různých zařízení. Otázka zní, zda jsou stejné střední hodnoty produkce z těchto čtyř strojů.

Řešení:

Problém by bylo možno rozřešit testováním všech dvojic pomocí t-testu. To by ale bylo značně nepohodlné a také nespolehlivé. EXCEL poskytuje vhodnější nástroj - **ANOVA** (Analysis of Variance).

Postavte hypotézu H_0: výběry pocházejí z rozdělení se stejnými středními hodnotami a hypotézu H_1: výběry pocházejí z rozdělení s různými středními hodnotami.

Použití nástroje Anova

Přes nabídku **Tools - Data Analysis** otevřete dialogové okno **Data Analysis**. Pak vyberte **Anova: Single Factor**.

2

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
T1	15	228,8	15,2533	0,95838
T2	15	225,7	15,0467	1,03552
T3	15	223	14,8667	0,53238
T4	15	217,2	14,48	0,756

ANOVA

Source of Vari	SS	df	MS	F	P-value	F crit
Between	4,84983	3	1,61661	1,9701	0,12892	2,76943
Within Gr	45,952	56	0,82057			
Total	50,8018	59				

Hypotéza H_0 se zamítá, když F > F krit, anebo hodnota P je menší jako zvolená hladina významnosti.

Počty stupňů volnosti (dF)

Testové kritérium F - testu = $\frac{MS_{\text{Mezi výběry}}}{MS_{\text{Všechny výběry}}}$

Jednostranní kritická hodnota F - rozdělení

Součty čtverců

Střední kvadratické odchylky

Jednostranní p - hodnota testového kritéria

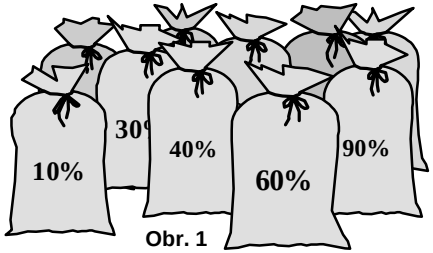
`1,9701 < 2,7694`, tedy hypotézu, že výběry `T1` až `T4` pocházejí z rozdělení se stejnými středními hodnotami nemůžete zamítnout.

ANOVA zkoumá, jestli je možné rozlišit rozdíl mezi odhady středních hodnot jednotlivých skupin, vzhledem k rozptylům hodnot uvnitř skupin.

Výsledek testu s výběry T1 až T4 tedy říká, že mezi nimi není možné zjistit rozdíl, protože je velký rozptyl uvnitř výběrů - každé zařízení pracuje s velkým rozptylem. Jiná situace by nastala, kdyby se snížily rozptyly (seřízením). Pak by se mohly projevit rozdíly.

Vysvětlení k testům hypotéz

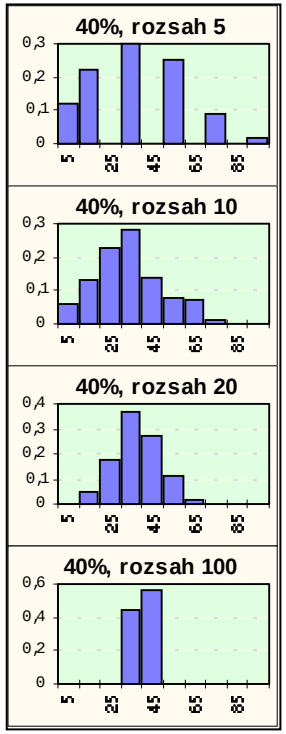
Proveďme myšlenkový pokus. Máme několik pytlů, v nichž jsou černé a bílé kuličky. Pytle jsou označeny podle procenta černých kuliček. Někdo přinese 100 kuliček a řekne že je vybral z pytle 40%. Jeho hypotéza tedy je: Přinesené kuličky pocházejí z pytle 40%. Označme ji H_0. Pak někdo jiný tvrdí, že kuličky byly vybrány z jiného pytle. Máme další hypotézu: Kuličky pocházejí z jiného pytle jako 40%. Tuto hypotézu označme H_1.



Obr. 1

Nyní máme dvě hypotézy postavené proti sobě. Dokážeme nějakým způsobem získat důkazy o tom, že první tvrzení je velice málo pravděpodobné a naopak, že pravdivost druhého tvrzení je vysoce pravděpodobná ? Dokážeme tedy zamítnout hypotézu H_0 ve prospěch hypotézy H_1?

Taková rozhodnutí jsou možná, musíme ale mít k dispozice vhodné nástroje. Připravme si je.



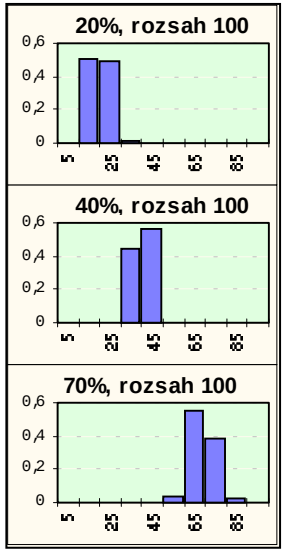
Obr. 2

Zkusme zjistit, zda skupina náhodně vybraných kuliček z jednoho pytle nese informaci o svém původu.

Z pytle 40% vybereme náhodně pět kuliček a zjistíme procento černých. Toto provedeme 100 krát. Ze zjištěných hodnot sestrojíme histogram relativních četností. Je na Obr.2 označený 40%, rozsah 5. Je vidět, že nejčastěji se vyskytovaly výběry, které obsahovaly něco mezi 30% a 40% černých. Hodně bylo také 5 procentních výběrů ale i 70 procentních.
Nyní zvětšíme rozsah výběru na 10, pak na 20 a 100. Příslušné histogramy jsou na Obr.1.
Zdá se že se zvyšováním rozsahu výběru, se informace zlepšuje. Je to tím, že ve větším výběru, pokud je náhodný, je větší počet prvků, které mohou mít různé hodnoty. Každý má volnost nabýt libovolnou hodnotu, ale má hodnotu charakterizující právě místo jeho původu. Při větším rozsahu výběru je více prvků majících volnost nabývat různých hodnot - máme větší **počet stupňů volnosti**.

Histogramy na Obr. 2 reprezentují rozdělení pravděpodobnosti výskytu hodnot procenta černých kuliček ve výběru pro 40% pytel. Procento černých kuliček ve výběru je tedy parametr obsahující informaci o tom, z kterého pytle pochází a může sloužit jako **testové kritérium** (též testová statistika) pro odhad místa původu výběru.

Máme testové kritérium a poznáme pro něj rozdělení pravděpodobnosti s různým počtem stupňů volnosti. Jinými slovy, když poznáme velikost výběru (počet stupňů volnosti) a procento černých ve výběru, můžeme určit pravděpodobnost s jakou se vyskytne právě tato hodnota.



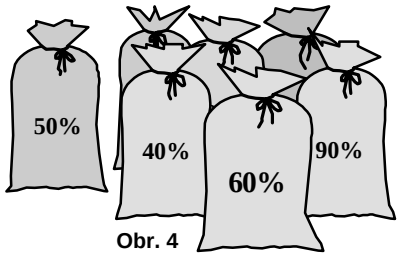
Obr. 3

Nyní zkusme jiné pytle. Na Obr. 3 jsou histogramy pro pytle 20% a 70% porovnány s histogramen pytle 40%, vše pro rozsah výběru 100. Je patrné, že histogramy neboli rozdělení se liší. Jinak, hodnoty procenta černých (testového kritéria), které jsou málo pravděpodobné pro výběry z pytle 40%, mohou být naopak vysoce pravděpodobné pro některé jiné pytle.

Máme vyhráno. Provedeme výběr z pytle, zjistíme procento černých a z rozdělení pravděpodobnosti pro pytel 40% zjistíme zda hodnota testového kritéria není pro pytel 40% málo pravděpodobná. Když víme, že pro některý jiný pytel je tato hodnota vysoce pravděpodobná, můžeme zamítnout předpoklad, že výběr pochází z pytle 40%. Zůstává určit co je málo pravděpodobné, tedy určit takovou hodnotu pravděpodobnosti, že všechny menší už jsou to "málo". Této hodnotě říkáme **hladina významnosti** (označovaná α).

Na Obr.1 jsou znázorněny pytle s kuličkami, jsou tam 10%, 30% tedy menší jako 40% a také větší. Pokud bude hodnota testového kritéria pro nějaký výběr například 2 procenta, je malá pravděpodobnost, že pochází z pytle 40%, ale velká, že pochází třeba z pytle 5%. Podobně pro hodnotu testového kritéria 85. Na obou stranách rozdělení pro 40% pytel tedy nacházíme rozdělení pro jiné pytle které nám dávají vysokou pravděpodobnost, že výběr pochází z nich, i když je pravděpodobnost pro 40% pytel malá. Na každé straně rozdělení pro 40% pytel můžeme zamítnout hypotézu H_0, že výběr pochází ze 40% pytle ve prospěch hypotézy H_1, že výběr pochází z jiného pytle. Protože to jde na obou stranách, hovoříme o **dvoustranném testu**.
Na Obr. 4 je jiná situace. Zde chybějí pytle 10%, 20%, 30%, jsou jen 40% a větší. Kdyby byla hodnota testového kritéria jakkoli nízká a pravděpodobnost pro 40% pytel jakkoli malá, nenajdeme pytel, který by poskytoval větší pravděpodobnost. Na levé straně rozdělení pro 40% pytel tedy není možné hypotézu H_0 zamítnout. Hypotéza H_1 může mít v tomto případě tvar: Výběr pochází z pytle, kde je podíl černých kuliček větší jako 40%. Zde se jedná o **jednostranný test**.
Všimněte si formulace hypotéz H1 pro dvoustranný a jednostranný test. V prvním případě se vyskytuje slovo "jiný" ve druhém "větší".

Při takovémto statistickém přístupu k testování se můžeme dopustit dvou chyb. **Chyba prvního druhu** vznikne když zamítneme H_0 a ona je pravdivá. **Chyby druhého druhu** se dopustíme, když H_0 pravdivá není a my ji nezamítneme. Možnost vzniku chyby prvního druhu omezíme snížením hladiny významnosti. Možnost vzniku chyby druhého druhu omezíme zvyšováním počtu stupňů volnosti.



Obr. 4