

3. Příprava dat



Klíčový význam korektního uložení získaných dat
Pravidla pro ukládání dat
Čištění dat před analýzou

Anotace



- Současná statistická analýza se neobejde bez zpracování dat pomocí statistického software. Předpokladem úspěchu je správné uložení dat v definované formě.
- Nejčastěji jde o databázové tabulky umožňující zpracování dat v celé škále různých aplikací.
- Neméně důležité je věnovat pozornost čištění dat předcházejícímu vlastní analýze. Každá chyba, která vznikne nebo není nalezena ve fázi přípravy dat, se promítne do všech dalších kroků a může zapříčinit neplatnost výsledků a nutnost opakování analýzy.

DATA – ukázka uspořádání datového souboru

Parametry (znaky)



Opakování

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Report	Country	Site	Matrix	ampl_met	Paramete	Year	Month	Day	Mean	Unit	Value	LOQ	Note	Page	Backgrou	in report
2	CEEC	Armenia	Sevan, Tsc	Air	pas	o,p-DDE	2008	7	29		pg/m3	3	0,5		REC	yes	no
3	CEEC	Armenia	Sevan, Tsc	Air	pas	PCB 118	2008	7	29		pg/m3	3,2	0,5		REC	yes	yes
4	CEEC	Armenia	Artashat, i	Air	pas	p,p-DDD	2008	7	29		pg/m3	3,7	0,5		REC	yes	yes
5	CEEC	Kazakhsta	Borovoe	Air	pas	PeCB	2008	7	29		pg/m3	3,9	0,5		REC	yes	yes
6	CEEC	Armenia	Yerevan, [Air	pas	PCB 138	2008	7	29		pg/m3	4,4	0,5		REC	no	yes
7	CEEC	Armenia	Yerevan, [Air	pas	PCB 153	2008	7	29		pg/m3	4,4	0,5		REC	no	yes
8	CEEC	Kazakhsta	Borovoe	Air	pas	gamma-H	2008	7	29		pg/m3	9,4	0,5		REC	yes	yes
9	CEEC	Armenia	Sevan, Tsc	Air	pas	PCB 28	2008	7	29		pg/m3	9,6	0,5		REC	yes	yes
10	CEEC	Armenia	Artashat, i	Air	pas	PCB 153	2008	7	29		pg/m3	9,9	0,5		REC	yes	yes
11	CEEC	Armenia	Amberd, r	Air	pas	o,p-DDE	2008	7	29		pg/m3	10	0,5		REC	yes	yes
12	CEEC	Armenia	Yerevan, [Air	pas	p,p-DDD	2008	7	29		pg/m3	10,2	0,5		REC	no	yes
13	CEEC	Armenia	Artashat, i	Air	pas	PCB 138	2008	7	29		pg/m3	10,5	0,5		REC	yes	yes
14	WEOG	USA	Eagle Hart	Air	active	Mirex	1990	11	16		pg/m3	0,03				yes	IADN
15	WEOG	Canada	Alert	Air	active	HCB	1995			60,8	pg/m3				A1_69		
16	WEOG	USA	Eagle Hart	Air	active	Gamma-H	1990	11	16		pg/m3	0,777				yes	IADN
17	WEOG	USA	Eagle Hart	Air	active	Alpha-HCl	1990	11	16		pg/m3	1,482				yes	IADN
18	WEOG	USA	Eagle Hart	Air	active	p,p-DDE	1990	11	16		pg/m3	2,428				yes	IADN
19	WEOG	USA	Eagle Hart	Air	active	Dieldrin	1990	11	16		pg/m3	3,993				yes	IADN
20	WEOG	USA	Eagle Hart	Air	active	PCB 101	1990	11	16		pg/m3	5,036				yes	IADN
21	WEOG	USA	Eagle Hart	Air	active	PCB 52	1990	11	16		pg/m3	6,764				yes	IADN
22	WEOG	USA	Eagle Hart	Air	active	p,p-DDD	1990	11	16		pg/m3	11,442				yes	IADN
23	WEOG	USA	Eagle Hart	Air	active	PCB 44	1990	11	16		pg/m3	12,613				yes	IADN
24	WEOG	USA	Eagle Hart	Air	active	Gamma-H	1990	11	16		pg/m3	24,33				yes	IADN
25	WEOG	Canada	Alert	Air	active	HCB	1998			70	pg/m3				A1_10		
26	WEOG	USA	Eagle Hart	Air	active	Alpha-HCl	1990	11	16		pg/m3	268,831				yes	IADN
27	WEOG	USA	Eagle Hart	Air	active	Aldrin	1990	11	16		pg/m3	<LoQ				yes	IADN
28	WEOG	USA	Eagle Hart	Air	active	Aldrin	1990	11	16		pg/m3	<LoQ				yes	IADN
29	WEOG	USA	Eagle Hart	Air	active	Dieldrin	1990	11	16		pg/m3	<LoQ				yes	IADN
30	WEOG	USA	Eagle Hart	Air	active	p,p-DDD	1990	11	16		pg/m3	<LoQ				yes	IADN
31	WEOG	USA	Eagle Hart	Air	active	p,p-DDE	1990	11	16		pg/m3	<LoQ				yes	IADN

Zásady pro ukládání dat



- Správné a přehledné uložení dat je základem jejich pozdější analýzy.
- Je vhodné rozmyslet si předem jak budou data ukládána.
- Pro počítačové zpracování dat je vhodné ukládat data v tabulární formě.
- Nejvhodnějším způsobem je uložení dat ve formě databázové tabulky:
 - každý sloupec obsahuje pouze jediný typ dat, identifikovaný hlavičkou sloupce;
 - každý řádek obsahuje minimální jednotku dat (např. pacient, měření apod.);
 - je nepřípustné kombinovat v jednom sloupci číselné a textové hodnoty;
 - komentáře jsou uloženy v samostatných sloupcích;
 - u textových (kategoriálních) dat je nezbytné kontrolovat překlepy v názvech kategorií;
 - specifickým typem dat jsou kalendářní data u nichž je nezbytné kontrolovat, zda jsou uložena v korektním formátu (dle aplikace).
- Takto uspořádaná data je v tabulkových nebo databázových programech možné převést na libovolnou výstupní tabulku.
- Pro základní uložení a čištění dat menšího rozsahu je možné využít aplikací MS Office.

MS Excel

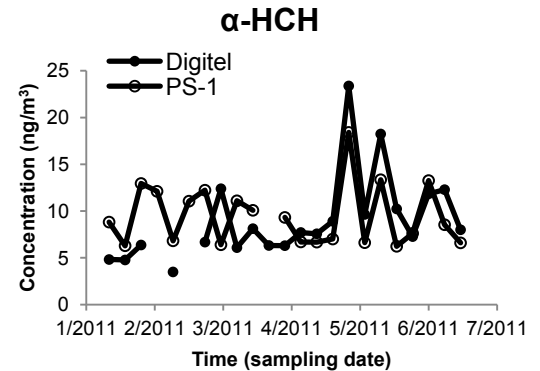


- Tabulkový procesor.
- První verze programu 30. 9. 1985 (Macintosh).
- Součást balíku kancelářských aplikací MS Excel.
- Aktualizace každé 2 až 3 roky; nové funkce, rozšíření počtu řádků a sloupců, změna formátu.
- Nejnovější formát Office XML je zazipovaný XML dokument, přípona .xlsx.
- Aktuální verze 2013 umožňuje ukládat tabulku až o 1 048 576 řádcích a 16 384 sloupcích.
- Maximální velikost buňky je 32 767 znaků.
- Excel umožňuje práci se širokou škálou dalších formátů.

Možnosti MS Excel



- Správa a práce s tabulárními daty.
- Řazení dat, výběry z dat, přehledy dat.
- Formátování a přehledné zobrazení dat.
- Zobrazení dat ve formě grafů.
- Různé druhy výpočtů pomocí zabudovaných funkcí.
- Tvorba tiskových sestav.
- Makra – zautomatizování častých činností.
- Tvorba aplikací (Visual Basic for Applications).



17	10	2
18	12	3
19	5	4
20	8	5
21	4	8
22	7	9
23	9	11
24	suma součinů řádků	310
25		

P. bini	2	Pohlaví			
Počet z	Délka				
Číslo	ryby2	Číslo	rv1	Váha	?
1	1				
2	2				
26					
106					
121					
160					
34					
45					
70					
72					
87					
Celkový součet					

OK Storno

Import a export dat



- **Import dat**
 - manuální zadávání;
 - import – podpora importu ze starších verzí Excelu, textových souborů, databází apod.;
 - kopírování přes schránku Windows – vkládání z nejrůznějších aplikací – MS Office, Statistica, přímo z HTML apod.;
 - využití textových souborů jako kompatibilního formátu pro přenos dat mezi různými aplikacemi.
- **Export dat**
 - ukládáním souborů ve formátech podporovaných jinými SW, časté jsou textové soubory, dbf soubory nebo starší verze Excelu;
 - přímé kopírování přes schránku Windows.

Import a export dat



- **Nejčastější datové formáty používané v MS Excel**
 - **.xlsx** – současný Office Open XML formát od verze MS Excel 2007;
 - **.xls** – starší binární varianta listů MS Excel (více verzí), stále používaná,
 - **.csv** – comma separated values, nejjednodušší tabulkový formát, 2 varianty,
 - **.dbf** – formát dBase, široce využívaný formát pro velké databáze;
 - **.db** – Paradox database, starší databázový systém;
 - **.slk** – SYmbolic LinK (SYLK) formát pro výměnu dat mezi aplikacemi Microsoft, neveřejný;
 - **.txt** – základní textový formát, často jediná možnost výměny dat s MS Excel.

Tipy a triky



• Výběr buněk

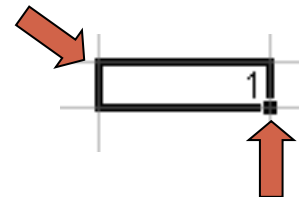
- CTRL+HOME – přesunutí na levý horní roh tabulky;
- CTRL+END – přesunutí na pravý dolní roh tabulky;
- CTRL+A – výběr celého listu;
- CTRL + klepnutí myší do buňky – výběr jednotlivých buněk ;
- SHIFT + klepnutí myší na jinou buňku – výběr bloku buněk;
- SHIFT + šipky – výběr sousedních buněk ve směru šipky;
- SHIFT+CTRL+END (HOME) – výběr do konce (začátku) oblasti dat v listu;
- SHIFT+CTRL+šipky – výběr souvislého řádku nebo sloupce buněk;
- SHIFT + klepnutí na objekty – výběr více objektů.

• Kopírování a vkládání

- CTRL+C – zkopírování označené oblasti buněk;
- CTRL+V – vložení obsahu schránky – oblast buněk, objekt, data z jiné aplikace;

• Myš a okraje buňky

- Chycení myší za okraj umožňuje přesun buňky nebo bloku buněk
- Při chycení čtverečku v pravém dolním rohu výběru je tažením možno vyplnit více buněk hodnotami původní buňky (ve vzorcích se mění relativní odkazy, je také možné vyplnění hodnotami ze seznamu – např. po sobě jdoucí názvy měsíců).



Ukotvení příček



- Umožňuje ukotvení libovolných řádků a sloupců pro pohodlné vkládání a prohlížení dat v tabulce.
- Umožňuje číst řádky/sloupce ze začátku tabulky i po přesunutí se dále.
- Záložka „Zobrazení“ → „Ukotvit příčky“.

- Nabízené možnosti:

- Ukotvit příčky – ukotví řádky nad označenou buňkou a sloupce vlevo od označené buňky.
- Ukotvit horní řádek.
- Ukotvit první sloupec.
- Ukotvení zrušíme opětovným odkliknutím možnosti ukotvení příček.

	F	G	H	
	poslední kontrola	pohlaví	nemocný	tíž
9	9.4.2010	muž		1
10	29.3.2010	muž		1

Databázová struktura dat v Excelu



Sloupce tabulky = parametry záznamů, hlavička udává obsah sloupce
– stejný údaj v celém sloupci

Jednotlivé záznamy
(taxon, lokalita,
měření, pacient atd.)



	A	B	C	D	E	F	G	H	I
1	Číslo	Značka	Společ	Pohlaví	Délka	Váha	P. anguillae	P. bini	
2	1	1	1	m	27,5	23,0	2	2	
3	2	2	2	f	34,0	62,5	0	2	
4	3	5	3	f	58,0	230,0	0	0	
5	4	6	4	f	42,0	155,0	0	0	
6	5	7	5	f	44,0	149,8	0	0	
7	6	8	6	f	56,0	323,0	0	1	
8	7	9	7	m	48,5	178,2	0	0	
9	8	10	8	f	30,5	47,7	4	6	
10	9	11	9	f	47,0	175,9	5	14	
11	10	12	10	f	40,0	85,1	5	9	
12	11	14	11	f	40,0	101,0	0	0	
13	12	15	12	f	31,0	84,0	15	9	
14	13	16	13	f?	22,0	9,0	0	0	
15	14	17	14	f	42,0	108,0	1	3	
16	15	18	15	f	44,0	130,0	0	0	
17	16	19	16	f	37,0	85,0	2	5	
18	17	20	17	f	50,0	212,0	1	8	

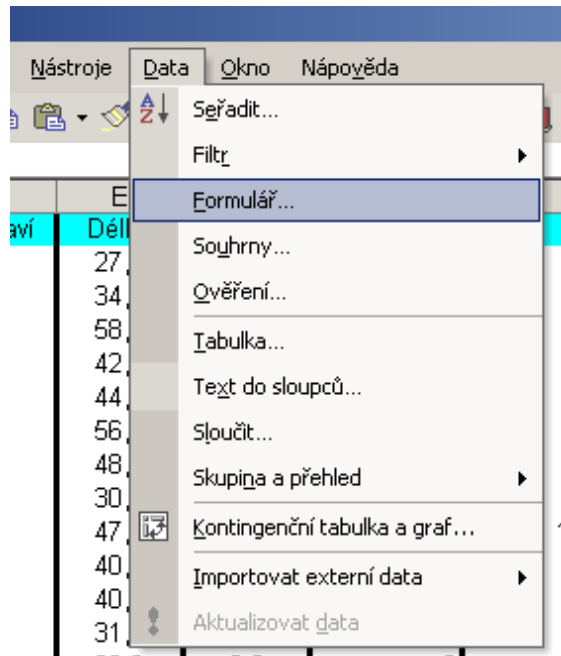
Excel neumožňuje pojmenování řádků a sloupců vlastními názvy.

Automatický zadávací formulář I.



- Slouží k usnadnění zadávání dat do databázových tabulek
- Načítá automaticky hlavičky sloupců jako zadávané položky

Microsoft Office 2003 a starší



Číslo ryby:	1	1 z 19
Značka ryby:	1	Nový
Společ číslo:	1	Odstranit
Pohlaví :	m	Obnovit
Délka:	27,5	Předchozí
Váha:	23	Další
P. anguillae:	2	Kritéria
P. lini:	2	Zavřít

Nový záznam

Vyhledávání

Názvy sloupců

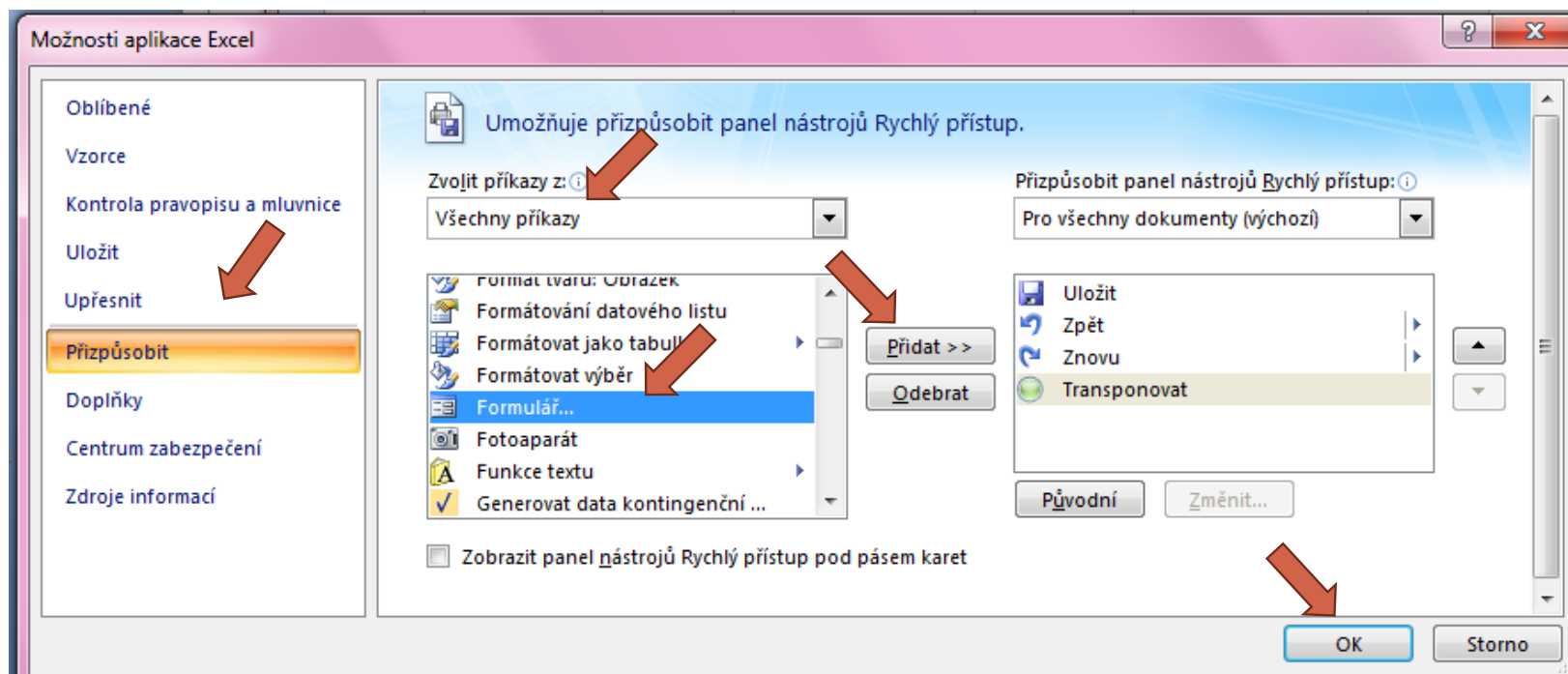
Obsah dané buňky - editovatelný

Automatický zadávací formulář II.

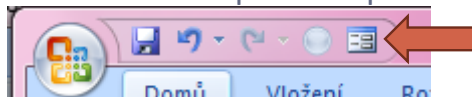
Microsoft Office 2007 a novější

- Aplikaci automaticky zadávaného formuláře je nutné aktivovat

- „Tlačítko Office“ → „Možnosti aplikace Excel“



- Automatický zadávací formulář spustíme pomocí nové ikonky na panelu nástrojů Rychlý přístup; dále stejné



Automatické seznamy



- Vytváří se z hodnot buněk v daném sloupci a umožňují vložit hodnotu výběrem ze seznamu již zadaných hodnot – usnadnění zadávání

Sloupec z něž je seznam vytvořen a pro který platí

Taxon	Abundance	Lokalita	etc.

Buňka, do níž se vloží vybraná hodnota

Glo

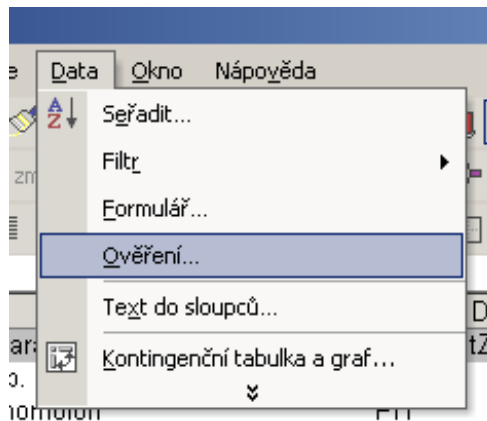
- Vyjmout
- Kopírovat
- Vložit
- Vložit jinak...
- Vložit buňky...
- Odstranit...
- Vymazat obsah
- Vložit komentář
- Formát buněk...
- Vybrat ze seznamu...**
- Přidat kukátko
- Hypertextový odkaz...

Caryophyllaeides fennica (Schneider, 1902) Ca
Piscicola geometra (Linnaeus, 1761) Pi
Acanthocephallus lucii (Müller, 1776) Pt
Apophallus mühlungi Jägerskiöld, 1899
Argulus foliaceus (Linnaeus, 1758)
Caryophyllaeides fennica (Schneider, 1902)
D. cabaleroi
D. crucifer Wagener, 1857
D. fallax Wagener, 1857
D. nanus Dogiel et Bychowsky, 1934

Automatická kontrola dat



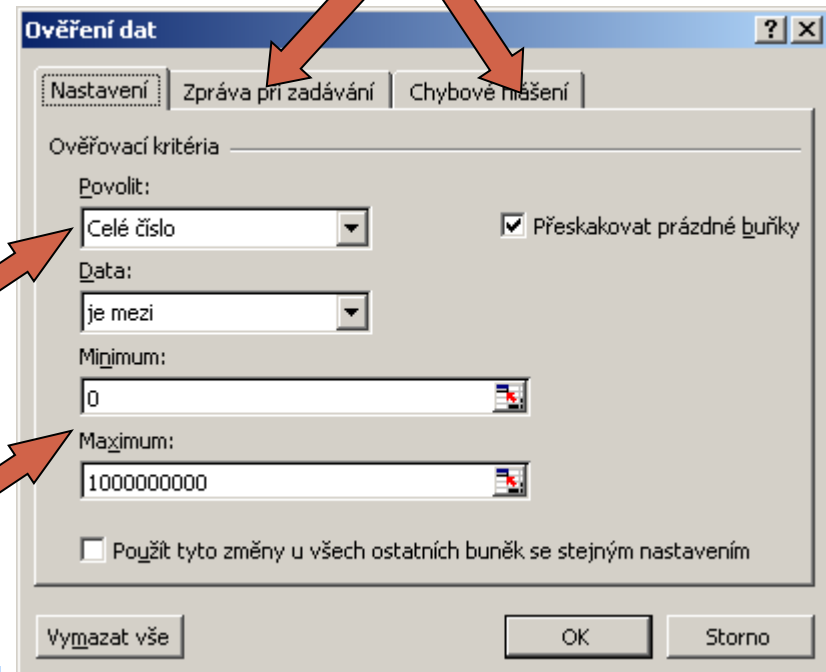
- Umožňuje ověřit typ, rozsah nebo povolit pouze určitý seznam hodnot zadávaných do sloupce databázové tabulky



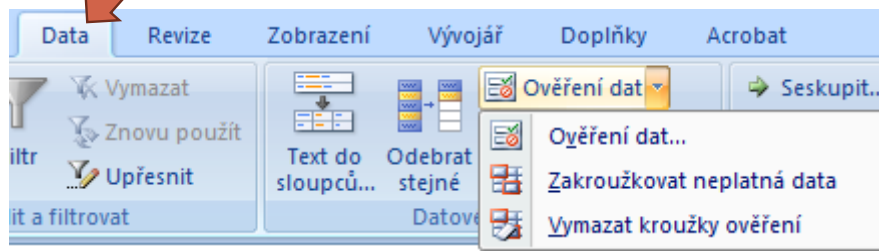
Co je povoleno – definiční obory čísel, seznamy, vzorce atd.

Rozsahy hodnot, načtení seznamů apod.

komunikace s uživatelem



Microsoft Office 2007

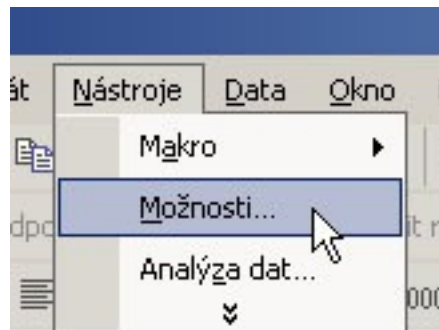


Seznamy I.

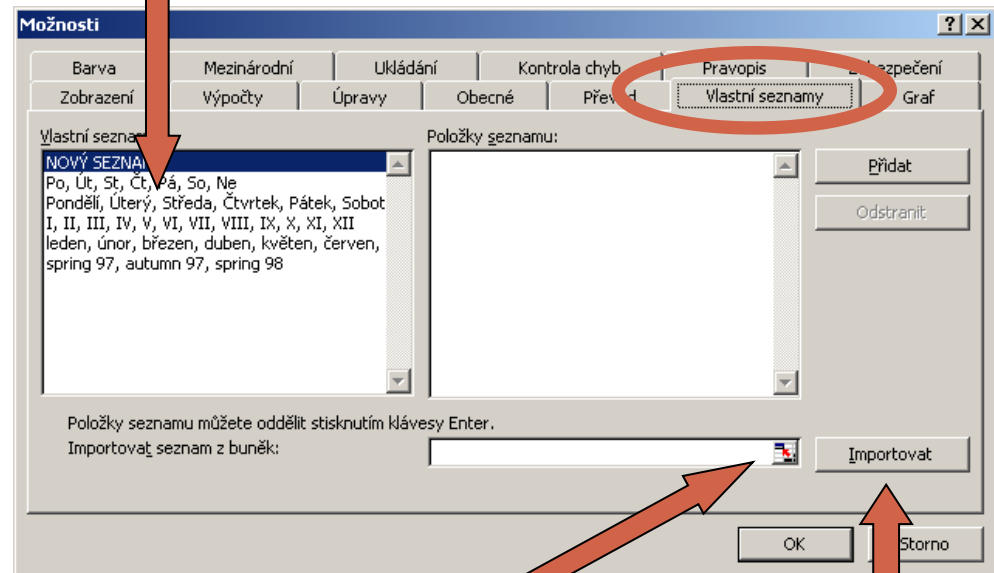


- Skupiny hodnot zachovávající logické pořadí, některé jsou zabudované (např. dny v týdnu, měsíce v roce), další je možné uživatelsky vytvořit, slouží pro účely řazení a automatického vyplňování dat

Microsoft Office 2003 a starší



Existující seznamy



Výběr buněk pro nový seznam

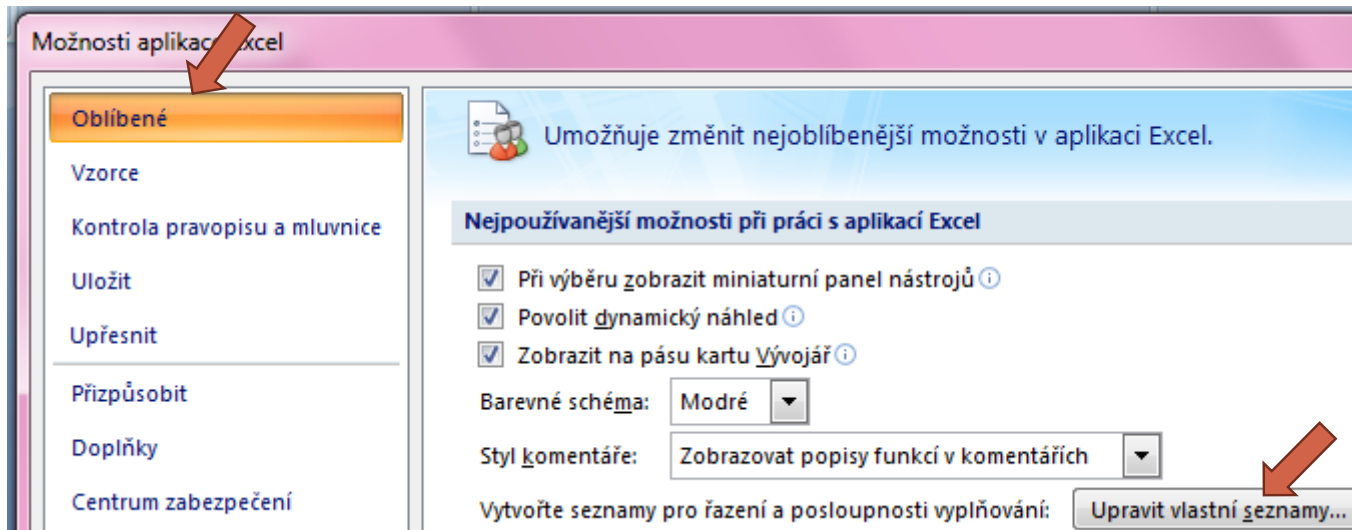
Načtení nového seznamu

Seznamy II.



Microsoft Office 2007

- „Tlačítko Office“ → „Možnosti aplikace Excel“



- Vlastní seznamy dále stejné (viz předchozí slide)

Řazení dat

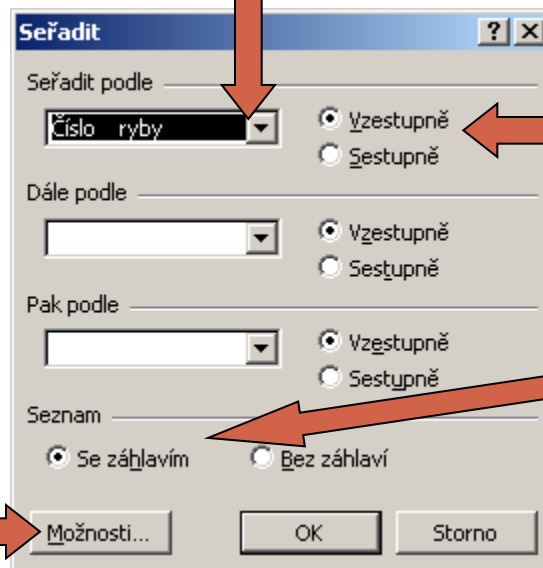
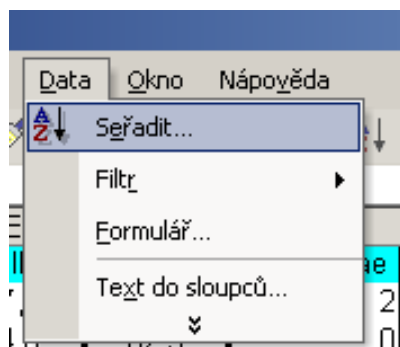


- Řazení dat je nejjednodušším způsobem jejich zpřehlednění, užitečným hlavně u menších/výsledkových tabulek



Zkontrolujte, zda seřazení nezničí vazby mezi buňkami = kontrola oblasti, kterou řadíte.

Podle čeho řadit



Směr řazení – vzestupně, sestupně

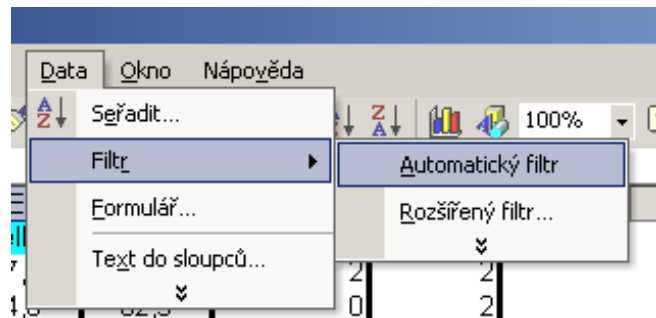
Využít první řádek oblasti jako záhlaví

Další možnosti – řazení řádků, řazení podle seznamu

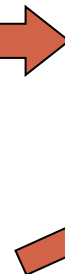
Automatický filtr



- Pomocí automatického filtru je snadné vybírat úseky dat pro další zpracování na základě hodnot ve sloupcích databázové tabulky, výběr je možný i podle více sloupců (např. určitá skupina pacientů).
- Funkce automaticky rozezná hlavičky sloupců v souvislé oblasti buněk.
- Čísla filtrovaných řádků jsou zobrazena modře.
- **Výhodné pro čištění dat (vyhledávání překlepů, kombinace textu a čísel).**



Rozbalení seznamu hodnot nalezených ve sloupci



Výběr hodnot pro filtraci

	A	B	C	D	E
	Číslo	Značka	Společ	ohlav	Délka
1		1	1	(Vše)	27,5
2		2	2	(Prvních 10...)	34,0
3		5	3	(Vlastní...)	58,0
4	3	6	f	f?	42,0
5	4	7	m		44,0
6	5	8		f	56,0
7	6	9		m	48,5
8	7				

Rozšířený filtr



- Funguje podobně jako automatický filtr, ale seznam povolených hodnot není nutné vybírat ručně – je uveden v oblasti jinde na listu (nebo i na jiném listu).
- Podmínkou jsou shodná záhlaví filtrované oblasti a oblasti povolených hodnot.
- Prázdné buňky odpovídají prázdné podmínce – tj. je-li v oblasti povolených hodnot nějaká buňka prázdná, splní podmínku libovolná buňka filtrované oblasti.
- Čísla řádků filtrované oblasti jsou zobrazena modře.

Tlačítko Upřesnit na kartě Data

Revize Zobrazení Vývojář Doplnky Acrobat

Seřadit a filtrovat

Seřadit Filtr Vymazat Znovu přetvořit Text do sloupců... Odebrat stejné

Upřesnit

	D
	ParazitZkratka
	Gsp
	DH

Upřesnit

Chcete-li omezit záznamy, které budou zahrnuty ve výsledné sadě dotazu, zadejte složitá kritéria.

Rozšířený filtr

Akce

Přímě v seznamu

Kopírovat jina

Oblast seznamu: []

Oblast kritérií: []

Kopírovat do: []

Bez duplicitních záznamů

OK Storno

Výběr oblasti cílových hodnot (přefiltrovaných)

Původní seznam včetně záhlaví

Oblast kritérií včetně záhlaví

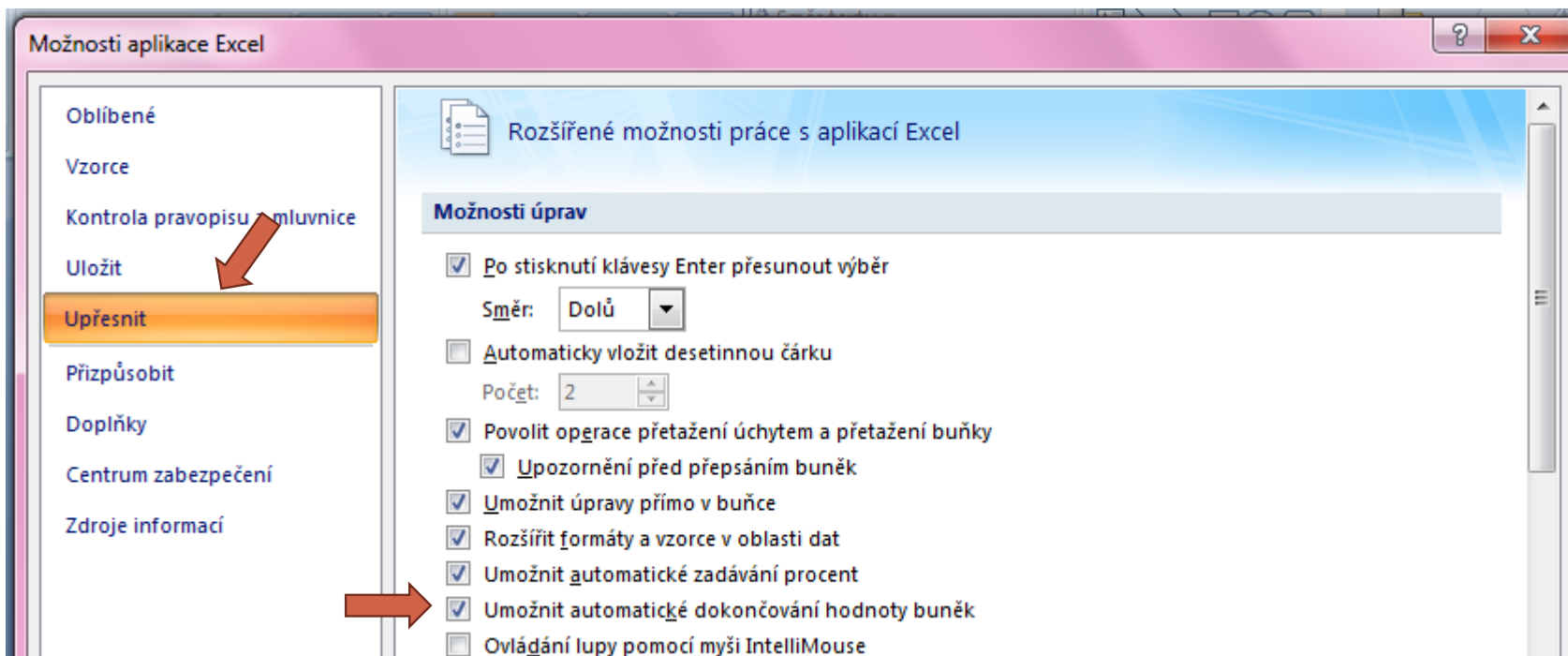
Automatické dokončování hodnot buněk



- Vhodné pro textová pole; následně není nutné vypisovat celé slovo či slovní spojení, ale jen zvolit nabízené, již dříve použité slovo či slovní spojení
- Automatické dokončování hodnot buněk je nutné nastavit
 - „Tlačítko Office“ → „Možnosti aplikace Excel“



→ „Možnosti aplikace Excel“



4. Vzorce v Excelu

Tipy pro práci s Wordem

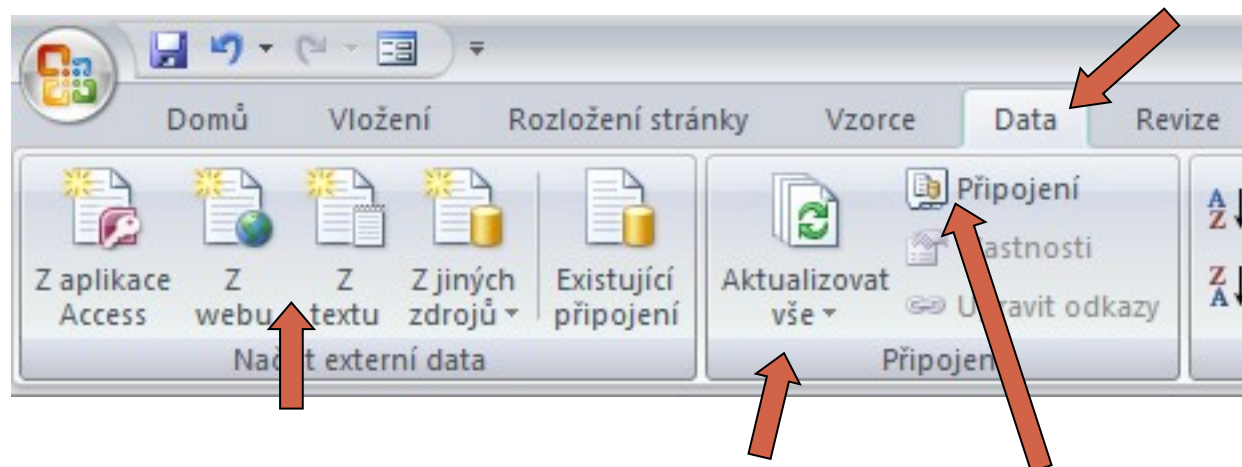
Kontingenční tabulky v Excelu



Zdroje dat Excelu



- Import dat z webu / MS Word pomocí schránky Windows.
- Excel umožňuje připojit externí zdroje dat.



- Propojení lze aktualizovat ručně/nastavit interval.
- Po zrušení propojení je třeba soubor odpojit.

Zdroje dat Excelu

Žlutý čtverec se šipkou u každé HTML tabulky.

The screenshot shows the Microsoft Excel interface with a web browser window open. The browser window displays the website of the Ministry of Environment (Ministerstvo životního prostředí) with a table of waste data. A red arrow points from a yellow square in the Excel grid (cell A1) to a yellow square on the web page table, indicating the source of the data.

Ministerstvo životního prostředí

Data získávána na základě zákona č. 185/2001 Sb. o odpadech a souvisejících prováděcích předpisů v platném znění. (Stav dat k 30.9.2013)

Přehled odpadů

Rok: 2012
Odpady: '200301'
Nakládání: nedefinováno
Kategorie: nedefinováno
Vykazované území: ORP: nedefinováno Okres: nedefinováno Kraj: nedefinováno
Území předání, převzetí: ORP: nedefinováno Okres: nedefinováno Kraj: nedefinováno
[Nové zadání](#)

Zobrazeny záznamy: 0-15 [další stránka](#)

Katalogové číslo odpadu	Kategorie odpadu	Kód nakládání	Množství (+) (t)	Množství (-) (t)
200301	N	A00	376.665568	
200301	N	AN3		664.363109
200301	N	BD10		0.315000
200301	N	BN30	0.315000	
200301	O	A00	2888663.988542	

Importovat Storno

Zdroje dat Excelu



Načtou se veškerá data v tabulce, často včetně balastu.

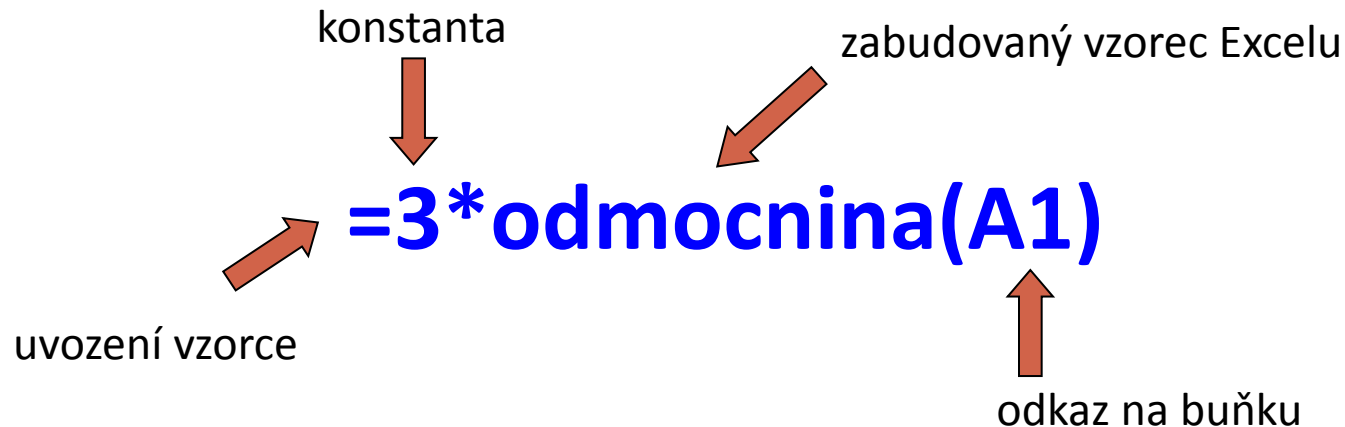
The screenshot shows an Excel spreadsheet with the following data:

	Kategorie odpadu	Kód nakládání	Množství (+) (t)	Množství (-) (t)
1	Přehled odpadů			
2	Rok: 2012			
3	Odpady: nedefinováno			
4	Nakládání: nedefinováno			
5	Kategorie: nedefinováno			
6	Vykazované území: ORP: nedefinováno Okres: nedefinováno Kraj: nedefinováno			
7	Území předání, převzetí: ORP: nedefinováno Okres: nedefinováno Kraj: nedefinováno			
8	Nové zadání			
9	Zobrazeny záznamy: 0-15 další stránka			
10	Katalogové			
11	číslo odpadu			
12	10101	O A00	7.880000	
13	10101	O AN3		7.880000
14	10102	O A00	10991.460000	
15	10102	O AN3		10991.460000
16	10102	O B00	16201.845000	
17	10102	O BD1		8918.910000
18	10102	O BD10		0.645000
19	10102	O BN12		2066.920000
20	10102	O BN13		788.440000
21	10102	O BN3		4426.930000
22	10304	N C00	0.130000	
23	10304	N CN5		0.130000
24	10306	O A00	3554.190000	
25	10306	O AN3		3545.030000
26	10306	O B00	6297.270000	
27	Součet množství na stránce:		37052.775000	30746.345000
28	Součet množství celkem:		77734711.841137	74905339.360736
29	Zobrazeny záznamy: 0-15 další stránka			
30				
31				

Vzorce v listu Excelu



- vepisují se do buněk sešitu
- vzorce jsou vždy uvozeny = (Ize též +, -).
- aritmetické operátory + zabudované funkce Excelu
- pro logické sčítání nečíselných položek se používá &
- výpočet je založen buď na číselných konstantách nebo odkazech na buňky

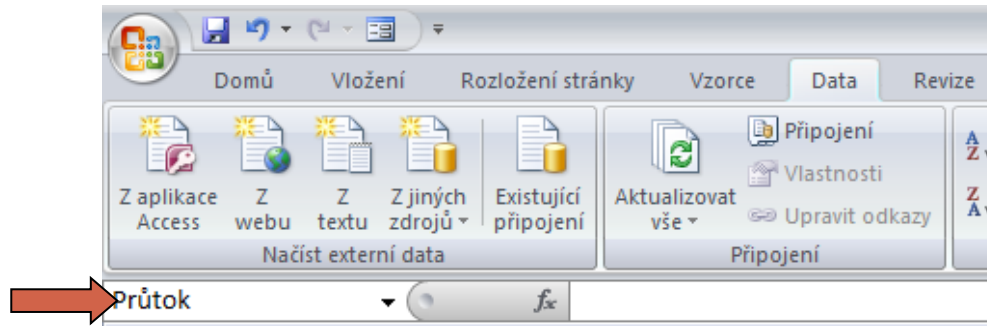


Vzorce – odkaz na buňku stylu A1



Relativní odkazy

- **A1** = buňka 1. řádku sloupce A
- **A1:B6** = blok buněk – levý horní roh je v 1. řádku, sloupec A, pravý dolní na řádku 6, sloupec B
blok lze pojmenovat vepsáním názvu do pole názvů:



- relativní odkaz se při automatickém vyplnění buněk vzorcem posune

Absolutní odkazy

- odkaz na buňku je pevně dán, při kopírování nebo automatickém vyplnění se nemění, lze uzamknout jak řádky, tak sloupce samostatně

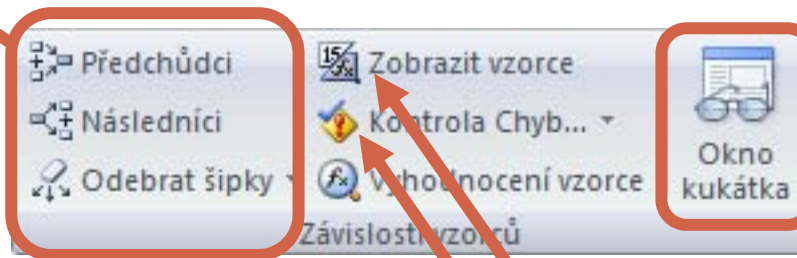
uzamčení sloupce → **\$A\$1** ← uzamčení řádku

Vzorce – tipy a triky I.



Závislosti vzorců – karta Vzorce

označení a odznačení předchozích a následných vzorců



kontrola chyb

sledování změn
hodnot i ve skrytých
a neviditelných sloupcích

Zpřehlednění vzorců

Vložit tabulátor CTRL+ALT+TAB
Vložit konec řádku ALT+ENTER

zobrazení vzorců namísto
hodnot v buňkách

Vzorce – tipy a triky II.



Vkládání komentářů, změny listu – karta Revize

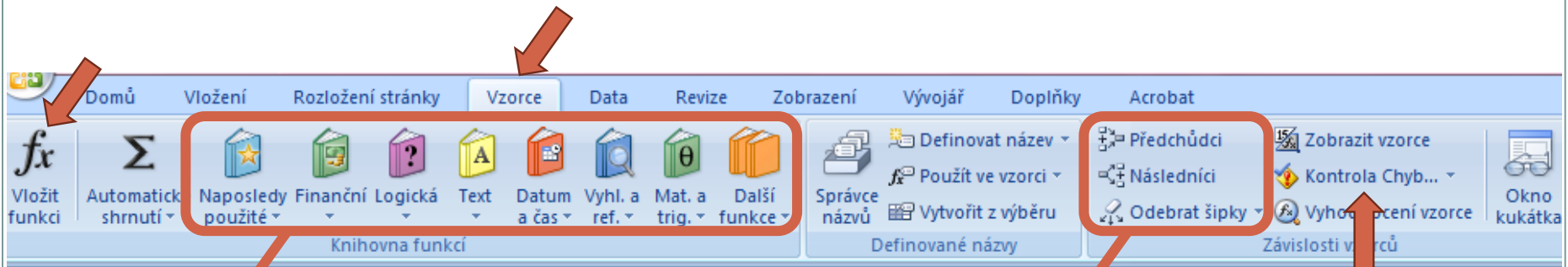
označení a odznačení předchozích a následných vzorců



možnost zamknout list či sešit dvojicí hesel

sledování změn a jejich schvalování nebo zamítní

Vzorce – tipy a triky III.

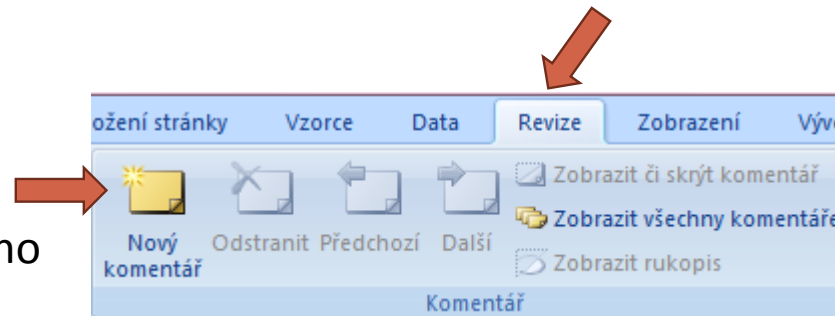


Výběr funkce z knihoven

Označení a odznačení předchozích a následných vzorců

Kontrola chyb

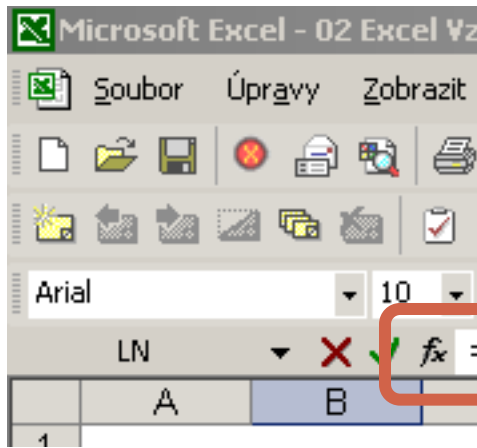
Vkládání komentářů a poznámek do vytvořeného souboru se vzorci



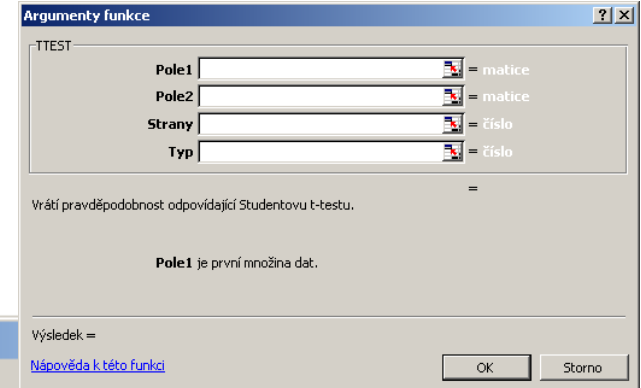
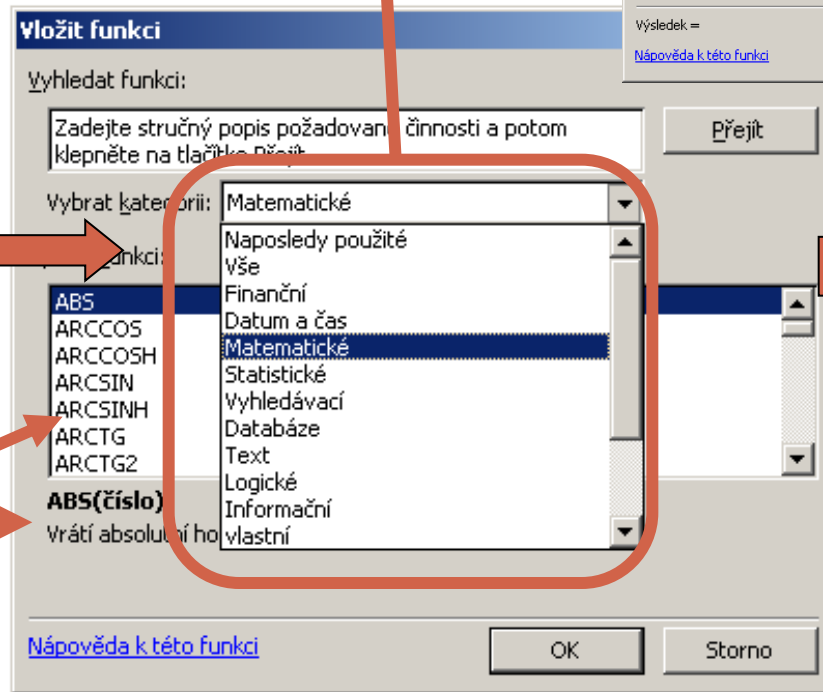
Vzorce – využití seznamu vzorců



Kategorie vzorců



Funkce a její stručný popis



průvodce funkcí

Vzorce – užitečné funkce



Celkem 408 funkcí ve verzi MS Excel 2010, ve verzi 2013 přidáno 50 nových funkcí

- **SUMA** – součet číselných hodnot oblasti;
- **SUMIF** – podmíněný součet (podmínky v doplňkové oblasti);
- **PRŮMĚR** – aritmetický průměr číselných hodnot oblasti;
- **GEOMEAN** – geometrický průměr číselných hodnot oblasti;
- **COUNTIF** – počet hodnot oblasti splňujících zadanou podmínku;
- **KDYŽ** – logická podmínka (if);
- **MAX, MIN** – maximum/minimum číselných hodnot oblasti;
- **MEDIAN** – výpočet mediánu;
- **PERCENTILE** – výpočet percentilů;
- **DATUM (ROK, MĚSÍC, DEN)** – práce s kalendářními daty;
- **ABS** – absolutní hodnota;

Statistické funkce v MS Excel



Od verze 2007 obsahuje MS Excel některé pokročilé statistické funkce

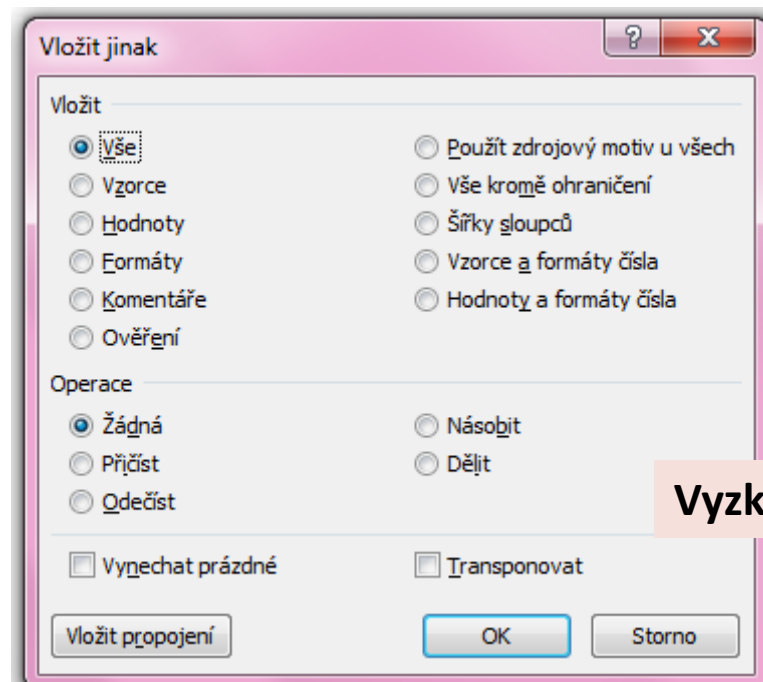
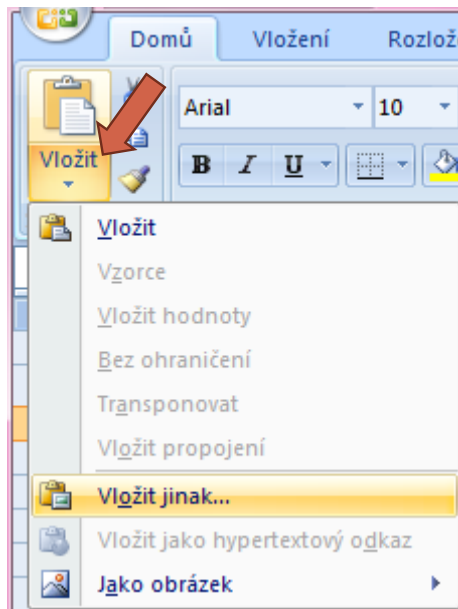
- **ZLEVA, ZPRAVA, ČÁST** – funkce pro ořezávání textových řetězců;
- **STANDARDNÍ MATEMATICKÉ FUNKCE (SIN, COS, LOG, LOGZ, EXP)** – a mnoho dalších.
- **CONFIDENCE** – výpočet intervalu spolehlivosti (při normálním rozdělení);
- **CORREL, PEARSON** – výpočet Pearsonova korelačního koeficientu;
- **COVAR** – výpočet kovariance dvou množin dat;
- **COUNTIF** – počet hodnot oblasti splňujících zadanou podmínku;
- **DEVSQ** – součet čtverců odchylek od výběrového průměru;
- **FDIST, GAMMADIST, CHIDIST, TDIST, NORMDIST** aj. – různá rozdělení pravděpodobnosti;
- **PRŮMODCHYLKA** – průměrná hodnota absolutních odchylek;
- **SLOPE** – směrnice lineárního modelu;
- **TTEST, ZTEST, CHITEST** – statistické testy shodnosti;

ŘADU DALŠÍCH FUNKCÍ VŠAK EXCEL POSTRÁDÁ A JE TŘEBA VYUŽÍT SILNĚJŠÍHO NÁSTROJE.

Kopírování / Vkládání



- Kopírování vzorců, textů, celých sloupců (zkopírování pomocí CTRL+C; dále „Vložit jinak...“)

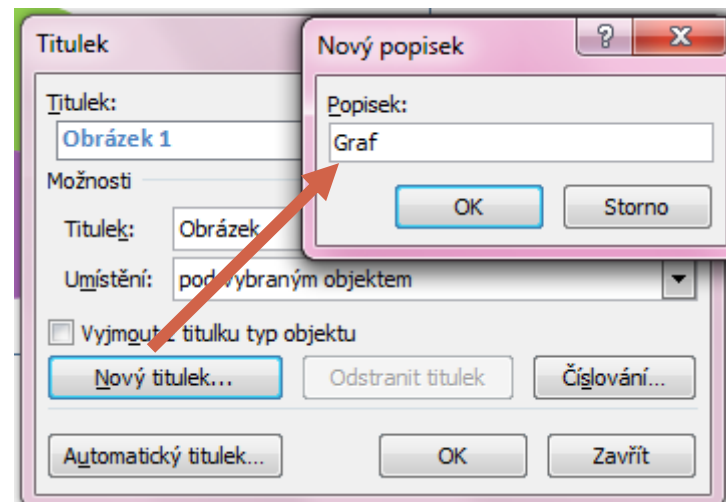
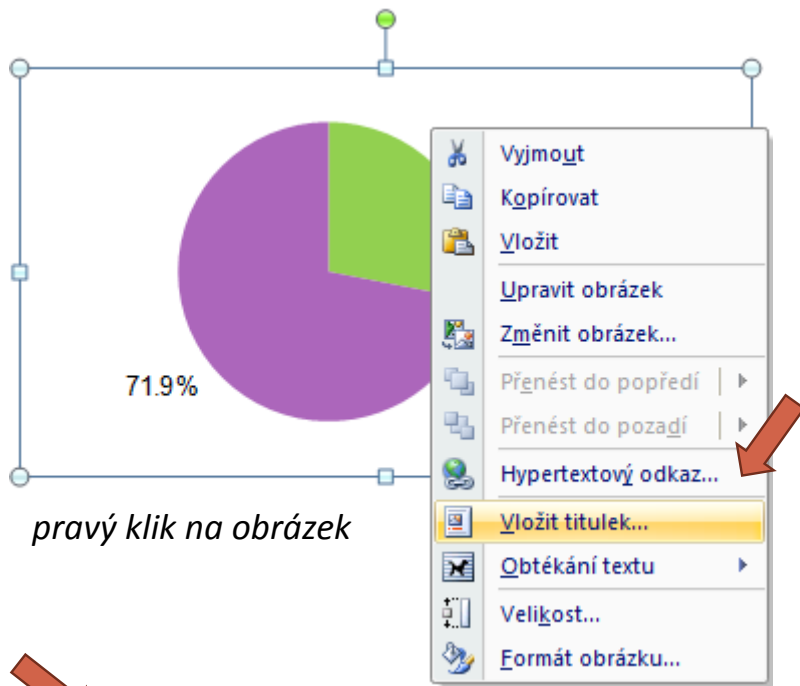


- Kopírování grafů z Excelu do Wordu:
Vložit jinak → Typ: Obrázek (rozšířený metasoubor)

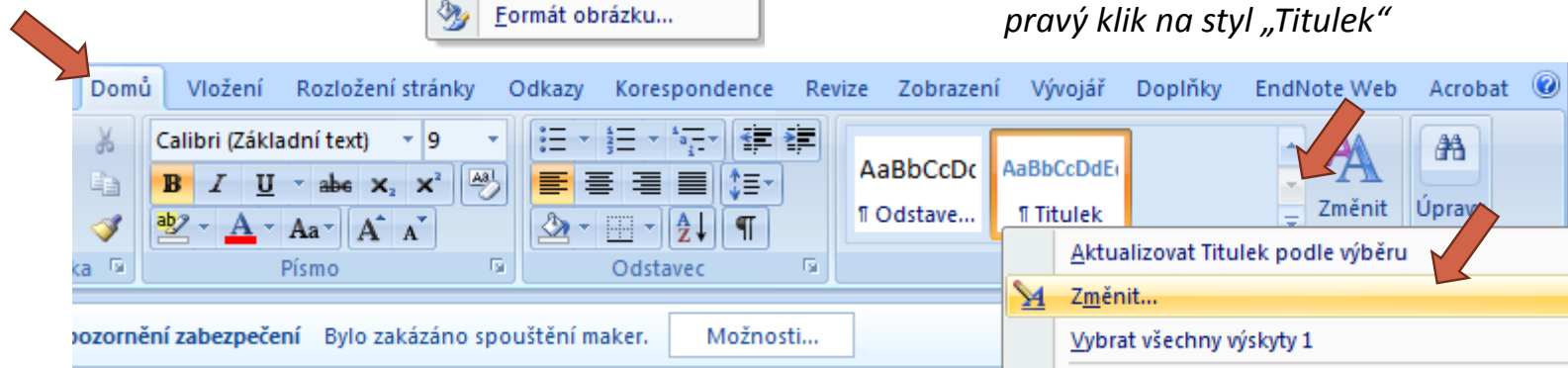
II.b Tipy pro práci s Wordem

Automatické titulky ve Wordu

- Automatické titulky ve Wordu pro snazší úpravy



pravý klik na styl „Titulek“



úprava stylu dle potřeby

Automatické seznamy ve Wordu

- Máme-li vytvořené automatické titulky grafů a tabulek, pak lze vytvořit automatické seznamy

71.9%

Graf

!

Každý objekt, který chceme zahrnout do automatického seznamu, musí mít automatický titulek

Aktualizace čísel titulků – pravý klik na označený titulek (obdobně aktualizace již vytvořeného seznamu)

Kontingenční tabulky v Excelu, 1. část

Ukázka kontingenční tabulky



Kontingenční tabulka vztahu pohlaví a onemocnění

	Nemocný	Zdravý	Celkem
Muž	a	b	a + b
Žena	c	d	c + d
Celkem	a + c	b + d	a + b + c + d = N

Ukázka kontingenční tabulky



Kontingenční tabulka vztahu pohlaví a onemocnění

	Nemocný	Zdravý	Celkem
Muž	45	11	56
Žena	25	6	31
Celkem	70	17	87

Ukázka kontingenční tabulky



Kontingenční tabulka vztahu pohlaví a onemocnění

	Nemocný	Zdravý	Celkem
Muž	45	11	56
Žena	25	6	31
Celkem	70	17	87

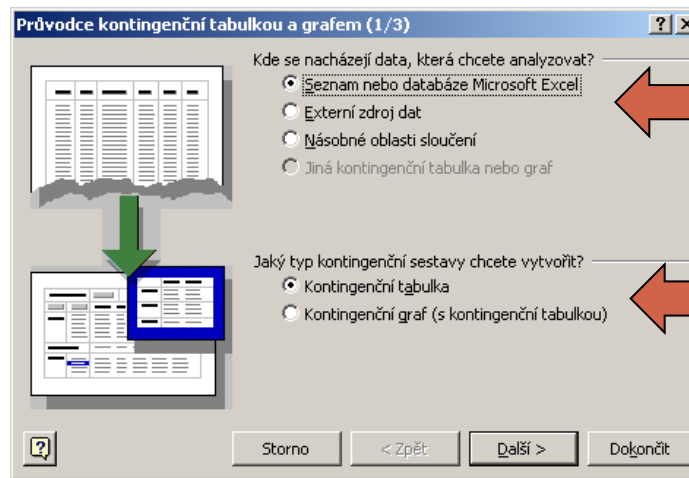
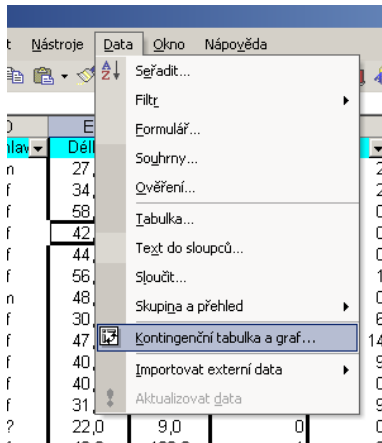
Hodnocení **nesmyslného** vztahu: dosažené vzdělání a doba strávená v nemocnici

	do 1 týdne	1 – 2 týdny	nad 2 týdny	Celkem
Základní vzdělání	10	9	5	24
Středoškolské vzdělání	32	18	6	56
Vysokoškolské vzdělání	4	2	2	8
Celkem	46	29	13	88

Kontingenční tabulka I.

- Umožňuje snadno vytvářet sumarizace dat ve smyslu počty hodnot, průměry, minima, maxima atd. v kombinacích kategorií (např. počet jedinců různých druhů na různých lokalitách)
- Automaticky je vybrána souvislá oblast dat (obdobně jako v případě automatického filtru)

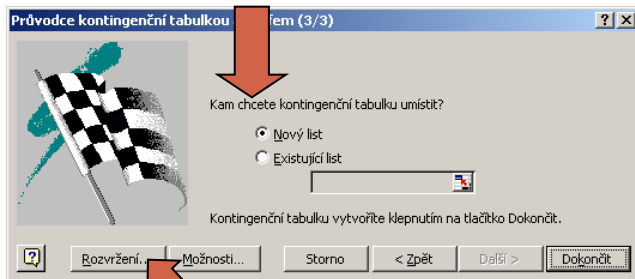
Microsoft Office
2003 a starší



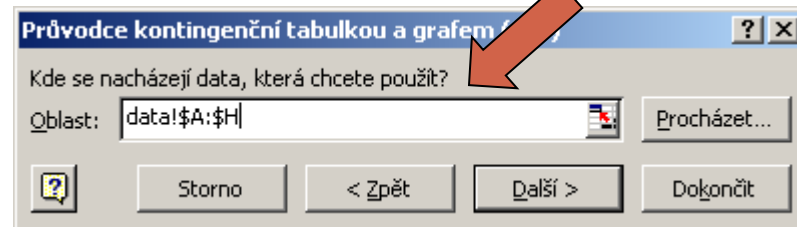
Zdroj dat (kromě
Excelu i např. externí
databáze)

Graf nebo tabulka

Umístění



Rozvržení a vlastnosti tabulek



Zdrojová oblast dat

Kontingenční tabulky – rozvržení I.

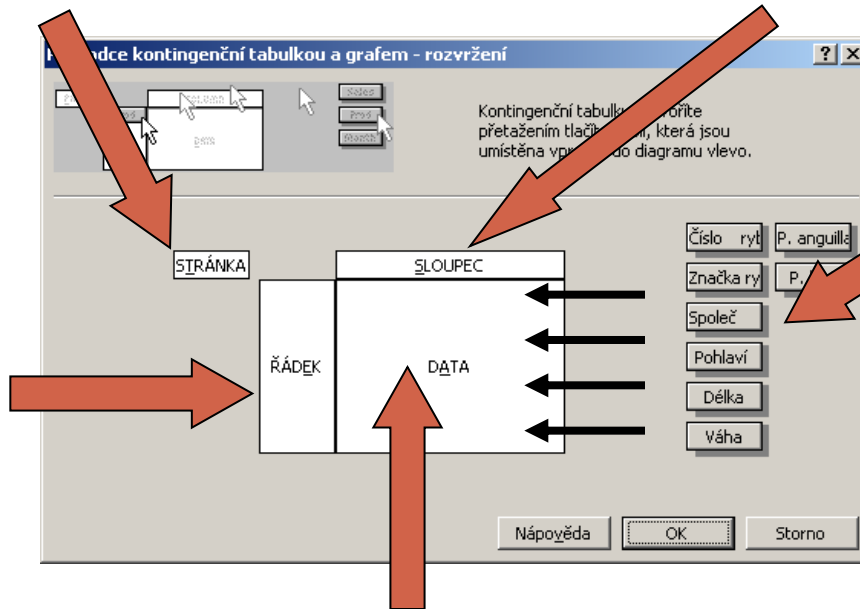
- Nastavit rozvržení kontingenčních tabulek je možné dvěma způsoby, zde představený postup je obsažen v Excel 97, 2000, 2007 i 2010 (speciální dialog), druhou možností je obdobná specifikace přímo v listu Excelu (2000, XP)

Microsoft Office
2003 a starší

tzv. stránka = tabulky podle zde nastaveného kritéria

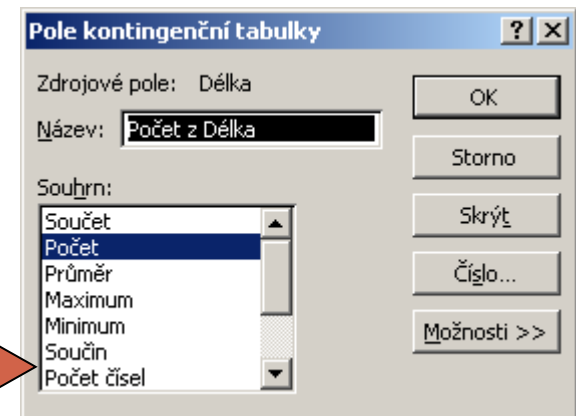
parametry sloupců

parametry
na řádcích



parametry, které je možné
zobrazit (hlavičky sloupců
databázové tabulky)

parametry dat
a možnosti sumarizace



Kontingenční tabulky – výsledek I.



- Výsledkem analýzy je tabulka vynášející proti sobě hodnoty řádkových a sloupcových parametrů kontingenční tabulky (např. taxony proti lokalitám, jde o seznamy hodnot obsažených v jednotlivých sloupcích), na průsečíku je zobrazena vybraná sumární charakteristika vybraných dat (průměr, suma, počet atd.)
- Tabulku v této formě je možné nadále editovat co se týče formátu i obsažených dat

Microsoft Office
2003 a starší

Roletky položek tabulky

Automatický souhrn

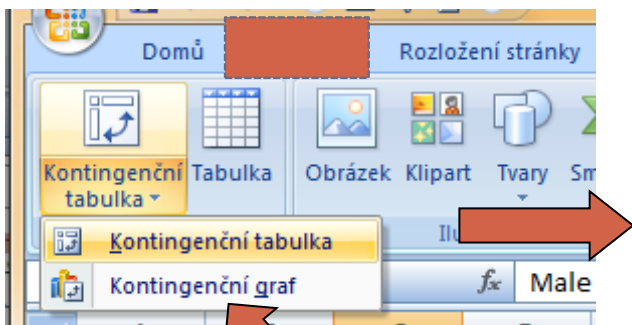
Počet z	Délka	Pohlaví					
Číslo ryby2	Číslo ryt	Váha	?	f	m		Celkový součet
1	1	23					
2	2	62,5			1		
26	26	72,4	1				
106	106	170,6	1				
121	121	190,2	1				
160	160	81,7		1			
34	34	43,1		1			
45	45	3,8	1				
70	70	12	1				
72	72	6,4	1				
87	87	54,2	1				
Celkový součet			7	3			

Seznam polí tabulky

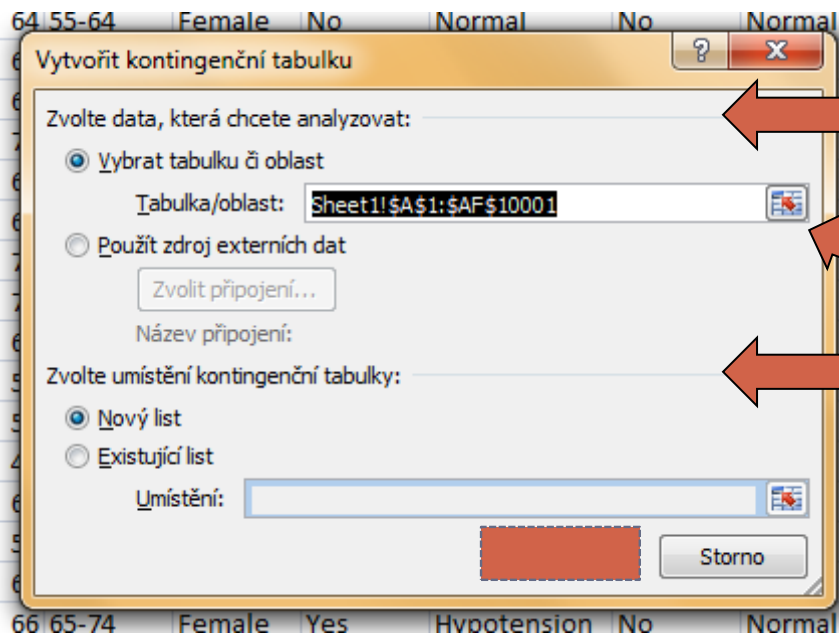
Panel nástrojů kontingenční tabulky

Kontingenční tabulka II.

Microsoft Office 2007



Graf nebo tabulka



Zdroj dat (kromě Excelu i např. externí databáze)

Zdrojová oblast dat

Umístění tabulky

Kontingenční tabulky – rozvržení II.

Microsoft Office 2007

Seznam polí kontingenční tabulky

Zvolte pole, které chcete přidat do sestavy:

- age
- agecat
- gender
- diabetes
- bp
- smoker
- choles
- active
- obesity
- angina
- mi
- nitro
- antidot

Přetáhnout pole mezi následujícími oblastmi:

- Filtr sestavy
- Popisky sloupců
- Popisky řádků
- Σ Hodnoty

Odložit aktualizaci rozlo... Aktualizovat

parametry, které je možné zobrazit v kontingenční tabulce

filtr

parametry ve sloupcích

parametry na řádcích

parametry dat

Seznam polí kontingenční tabulky

Zvolte pole, které chcete přidat do sestavy:

- age
- gender
- diabetes
- bp
- choles

Přetáhnout pole mezi následujícími oblastmi:

- Filtr sestavy
- Popisky sloupců
- Popisky řádků
- Σ Hodnoty

smoker

agecat

Počet z agecat

Kontingenční tabulky – nastavení II.

Seznam polí kontingenční tabulky

Zvolte pole, které chcete přidat do sestavy:

- age
- agecat**
- gender
- diabetes
- bp
- smoker**
- choles

Přetáhnout pole mezi nás

Filtr sestavy

Popisky řádků

agecat

Počet z agecat

Popisky sloupců

Popisky řádků

	No	Yes	Cell
45-54	1694	501	
55-64	3015	863	
65-74	2200	661	
75+	816	250	
Celkový součet	7725	2275	

Přesunout nahoru

Přesunout dolů

Přesunout na začátek

Přesunout na konec

Přejít k filtru sestavy

Přejít k popiskům řádků

Přejít k popiskům sloupců

Přejít k hodnotám

Odstranit pole

Nastavení polí hodnot...

Kontingenční tabulka

Microsoft Office 2007

Nastavení polí hodnot

Název zdroje: agecat

Vlastní název: Počet z agecat

Kritéria shrnutí pole hodnoty

Zvolte typ kalkulačky, který chcete použít pro shrnutí dat z vybraného pole:

- Součet
- Počet**
- Průměr
- Maximum
- Minimum
- Součin

Formát čísla

OK

Způsob sumarizace položky

Možnosti tabulky

Microsoft Excel - Přítent_loš

Nástroje kontingenční tabulky

Domů Vložení Rozložení stránky Vzorce Data Revize Vizualizace Vývojář Doplnky Acrobat Možnosti Návrh

Kontingenční tabulka

Aktivní pole: Počet z agecat

Nastavení pole

Výběr skupiny

Oddělit...

Skupinové pole

Seřadit

Aktualizovat

Změnit zdroj dat

Vymazat

Vybrat

Přesunout

Kontingenční graf

Vzorce

Nástroje OLAP

Seznam polí

Tlačítka +/-

Záhlaví polí

Zobrazit či skrýt