

ASTAc Biostatistika - cvičení



ASTAc/01 Vyučující: Mgr. Lucie Brožová; kontakt: brozova@iba.muni.cz

ASTAc/02 Vyučující: Mgr. Lucie Panáčková; kontakt: panackova@iba.muni.cz

ASTAc/03 Vyučující: Mgr. Petra Kovalčíková; kontakt: kovalcikova@iba.muni.cz

ASTAc/04 Vyučující: Mgr. Ivana Svobodová; kontakt: svobodova@iba.muni.cz

Důležité informace!!!



- Maximálně 1 absence
- Úspěšné splnění zápočtu na konci semestru
- Aktivita v hodinách
- Materiály v ISu
- Software: Microsoft Office - Excel 2010, Statistica 12 (EN)
- Možnost pracovat na vlastním počítači (složky se z univerzitních počítačů po vypnutí počítače mažou)

- Setkáme se: 24.9., 8.10., 22.10., 5.11., 19.11., 3.12., 17.12.
- Výuka: 16:00-17:40

Osnova



- Excel: příprava dat, základní vzorce a funkce v Excelu
- Základy popisné statistiky, kontingenční tabulky v excelu
- Základní rozdělení pravděpodobnosti, testování hypotéz
- Parametrické a neparametrické testy
- Analýza rozptylu
- Analýza kontingenčních tabulek, testy dobré shody
- Základy korelační analýzy a lineární regrese

Motivace



- Současná statistická analýza se neobejde bez zpracování dat pomocí statistických software. Předpokladem úspěchu je správné uložení dat ve formě „databázové“ tabulky umožňující jejich zpracování v libovolné aplikaci.
- Neméně důležité je věnovat pozornost čištění dat předcházející 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.

I. Příprava dat, MS Excel



Datová tabulka Zásady správné tvorby dat Představení MS Excel

DATA – ukázka uspořádání datového souboru

Parametry (znaky)

Opakování

Pařet	Číslo	a ₁	a ₂ %	a ₃ %	a ₄ %	a ₅ %	a ₆	a ₇	a ₈	a ₉	a ₁₀	a ₁₁	a ₁₂	a ₁₃	a ₁₄	a ₁₅	a ₁₆	
		cel.10 ³	%	%	%	%	cel.10 ³	cel.10 ³	cel.10 ³	cel.10 ³	%	m/s10 ³	m/s10 ³	m/s10 ³	m/s10 ³			
3	1	4										33	72		32			
4	2	76	8	38	66	24	06	44	50	18	33	95	19	48	10			
8	3	4	3	52	55	40	01	21	22	16	22	77	35	33	15			
11	4	61	5	39	64	35	03	36	39	21	33	103	26	49	13			
12	5	69	3	85	88	9	02	59	61	06	37	81	13	45	7			
14	6	59	15	55	70	19	09	33	41	11	32	137	33	61	15			
16	7	8	18	75	93	7	14	60	74	06	34	151	20	59	8			
20	8	96	3	72	75	23	03	69	72	22	40	77	11	33	5			
21	9	6	10	67	77	19	06	40	46	11	32	120	26	52	11			
22	10	33	4	55	59	39	01	18	20	13	28	81	42	24	12			
37	11	38	10	60	70	30	04	23	27	11	32	111	42	29	11			
38	12	64	2	76	78	17	01	49	50	11	25	366	73	115	23			
39	13	68	1	57	58	39	01	39	39	27	20	234	59	71	18			
49	14	85	7	67	74	26	06	57	63	22	30	156	25	108	17			
51	15	93	7	57	64	35	07	53	60	33	35	129	21	23	4			
52	16	22	10	55	66	34	02	12	15	07	33	45	30	12	8			
55	17	99	3	78	81	10	03	77	80	01	30	189	24	140	18			
56	18	5	2	80	82	13	01	40	41	07	25	101	25	54	13			
6	1	88	11	72	83	12	10	63	73	11	44	268	356	145	199			
9	2	92	2	66	68	28	02	61	63	26	42	168	269	76	122			
13	3	100	7	83	90	8	07	83	90	08	54	181	201	81	9			
15	4	96	1	75	76	23	01	72	73	22	45	343	47	124	169			
17	5	60									45	40		21				
19	6	72	2	78	80	18	01	56	58	13	44	103	178	63	109			
24	7	82	1	72	73	25	01	59	60	21	41	219	349	57	96			
26	8	103	1	85	86	3	01	88	89	03	41	364	411	112	126			
29	9	50	1	74	75	21	01	37	38	11	39	83	221	32	85			
30	10	119	1	51	52	47	01	61	62	56	33	83	134	52	84			
31	11	72	3	53	55	29	02	38	40	21	28	109	271	63	155			
32	12	108	36	50	76	8	39	54	93	09	27	146	157	106	114			
33	13	118	22	54	76	16	26	64	90	19	45	246	274	63	7			
34	14	170	1	82	83	16	02	139	141	27	34	440	312	119	84			
40	15	100	8	72	80	4	08	72	80	04	37	176	220	52	65			

Zásady pro ukládání dat

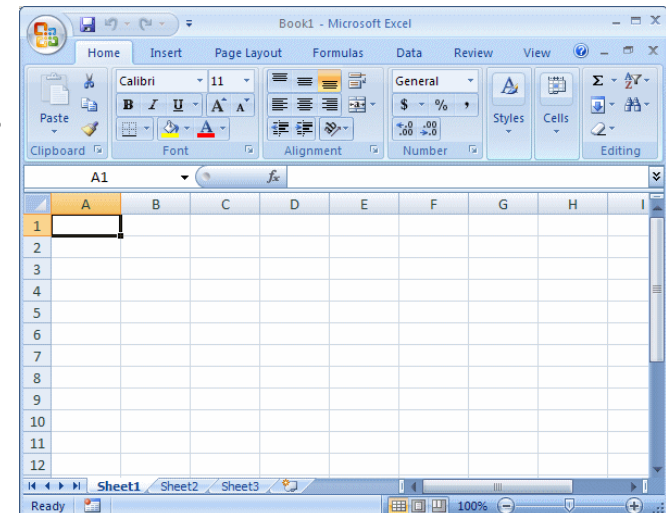


- Správné a přehledné uložení dat je základem jejich pozdější analýzy.
- Je vhodné rozmyslet si předem jak budou data ukládána.
- Pro počítačové zpracování dat je nezbytné ukládat data v tabulární formě.
- Nejvhodnějším způsobem je uložení dat ve formě databázové tabulky.
 - Každý sloupec obsahuje pouze jediný typ dat, identifikovaný hlavičkou sloupce;
 - Každý řádek obsahuje minimální jednotku dat (např. pacient, jedna návštěva pacienta apod.);
 - Je nepřípustné kombinovat v jednom sloupci číselné a textové hodnoty;
 - Komentáře jsou uloženy v samostatných sloupcích;
 - U textových dat je nezbytné kontrolovat překlepy v názvech kategorií;
 - Specifickým typem dat jsou data, u nichž je nezbytné kontrolovat, zda jsou uloženy v korektním formátu.
- 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 Excel.

MS Excel



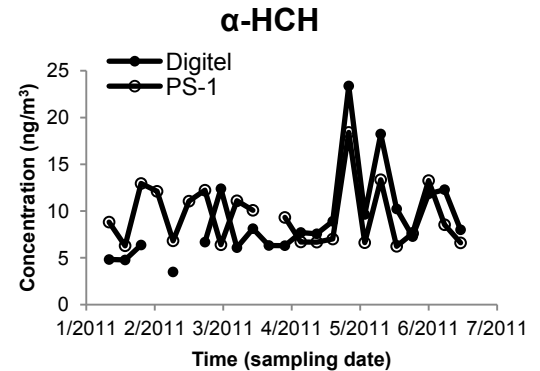
- Tabulkový procesor.
- Aktualizace každé 2 až 3 roky; nové funkce, rozšíření počtu řádků a sloupců, změna formátu.
- Starší formát: .xls, novější: .xlsx.
- Aktuální verze 2013 umožňuje ukládat tabulku až o 1 048 576 řádcích a 16 384 sloupcích.
- Maximální velikost buňky je 32 767 znaků.



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	rvl	Váha	?
1	1				
2	2				
26					
106					
121					
160					
34					
45					
70					
72					
87					
Celkový součet					

(Zobrazit vše)

OK Storno

II. Tvorba dat v MS Excel



Import a export dat

Struktura dat

Manuální zadávání nových dat

Tipy a triky jak se v datech pohybovat

Editace listů

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 atd.

- **Export dat**
 - Ukládáním souborů ve formátech podporovaných jinými SW, časté jsou textové soubory, dbf soubory nebo starší verze Excelu
 - Přímé kopírování přes schránku Windows

Databázová struktura dat v Excelu



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

Jednotlivé záznamy
(taxon, lokalita,
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	

Formáty buněk

- **Obecný** (bez formátu)
- **Číslo** (desetinná místa, oddělení 1000)
- **Měna** (desetinná místa, jednotky - symbol)
- **Datum** (různé formátování – 24.9.2015, 24. září 2015, 24-9-15, ...)
- **Čas**
- **Procenta**
- **Text**
- **Další** (vč. nastavení vlastního formátu)...

01_Excel_Data [Režim kompatibility] - I

Domů Vložení Rozložení stránky Vzorce Data Revize Zobrazení Vývojář PDF Acrobat

Arial 10

Obecný

Číslo

Formát buněk

Číslo Zarovnání Písmo Ohraničení Výplň Zámek

Druh:

- Obecný
- Číslo
- Měna
- Účetnický
- Datum
- Čas
- Procenta
- Zlomky
- Matematický
- Text
- Speciální
- Vlastní

Ukázka: CisloRyby

Desetinná místa: 0

Oddělovat 1000 (,)

Záporná čísla:

- 1234
- 1234
- 1234
- 1234

Číslo se používá pro obecné zobrazování čísel. Formát měny a účetnický formát poskytují speciální formáty pro peněžní hodnoty.

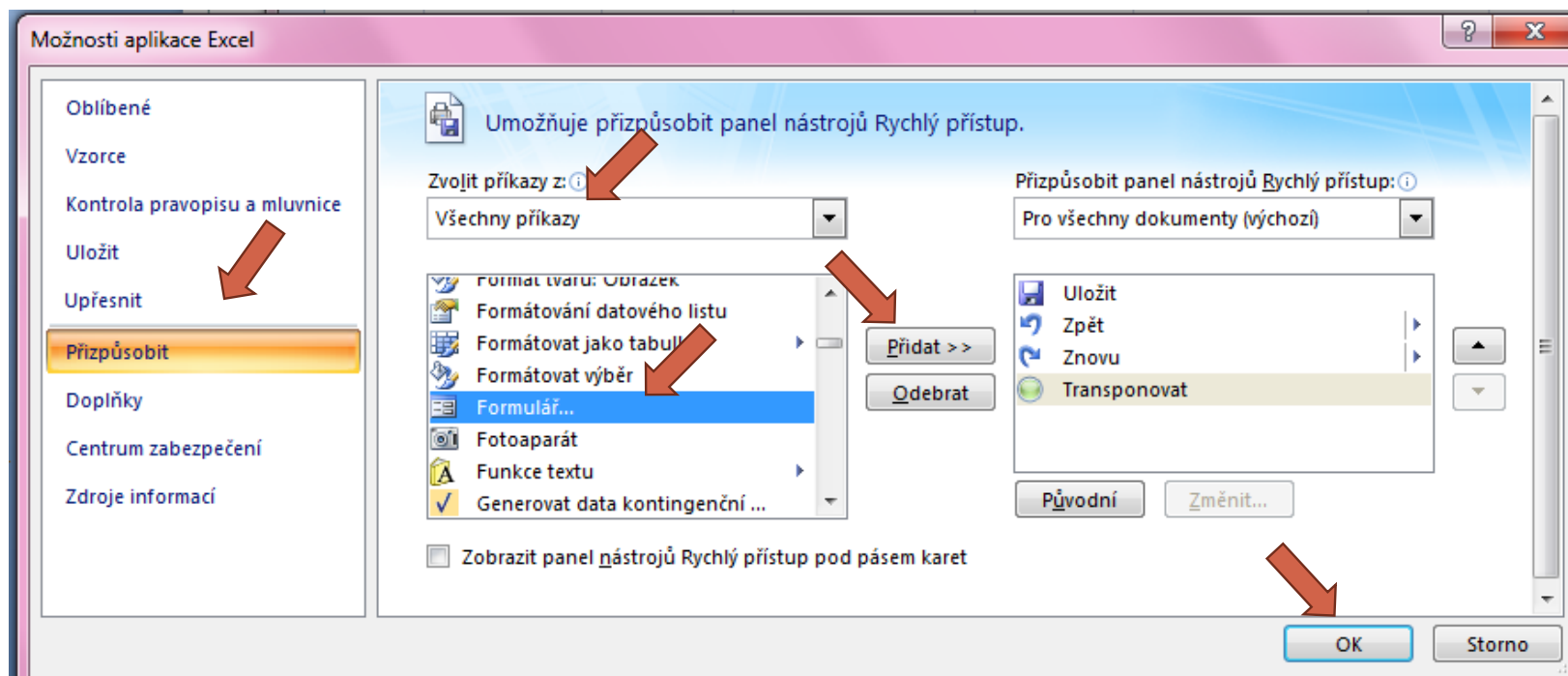
OK Storno

	A	B
1	CisloRyby	taxo
2	97	Monogenea
3	97	Monogenea
4	97	acanthocep
5	98	Crustacea
6	98	Monogenea
7	99	Monogenea
8	100	Monogenea
9	100	Monogenea
10	100	Monogenea
11	101	Monogenea
12	101	Monogenea
13	101	Monogenea
14	102	Monogenea
15	102	Monogenea
16	102	Monogenea
17	103	Monogenea
18	103	Crustacea
19	103	Monogenea
20	104	Monogenea
21	104	Monogenea
22	105	Monogenea
23	105	Monogenea
24	105	Monogenea
25	105	Monogenea
26	105	digenea
27	106	Monogenea
28	106	Monogenea

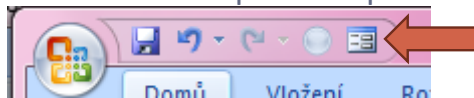
Automatický zadávací formulář I.



- Aplikaci automaticky zadávaného formuláře je nutné aktivovat
 - „Tlačítko Office“ → „Možnosti aplikace Excel“



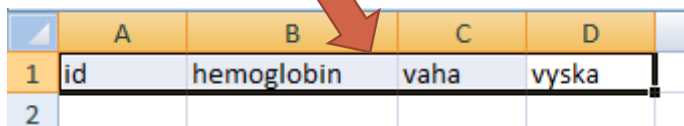
- Automatický zadávací formulář spustíme pomocí nové ikonky na panelu nástrojů Rychlý přístup



Automatický zadávací formulář II.

- Slouží k usnadnění zadávání dat do databázových tabulek
- Po označení načítá automaticky hlavičky sloupců jako zadávané položky

1. Označíme názvy sloupců datové matice

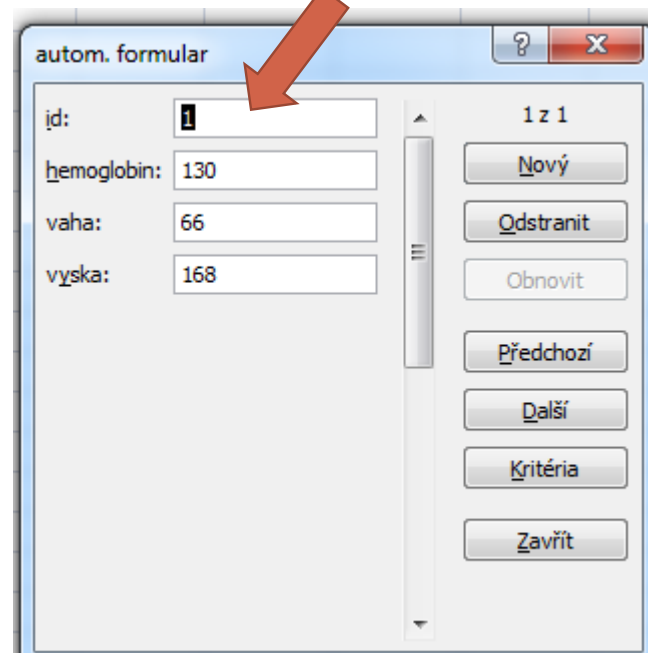


	A	B	C	D
1	id	hemoglobin	vaha	vyska
2				

2. Klikneme na novou ikonu Formulář v panelu nástrojů



3. Vyplníme údaje pro hodnocený subjekt



autom. formular

id:

hemoglobin:

vaha:

vyska:

1 z 1

Nový

Odstranit

Obnovit

Předchozí

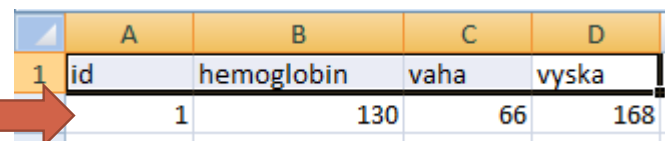
Další

Kritéria

Zavřít

Vyzkoušej !!!

4. Do datové tabulky se doplní zadané údaje



	A	B	C	D
1	id	hemoglobin	vaha	vyska
	1	130	66	168

Automatické seznamy



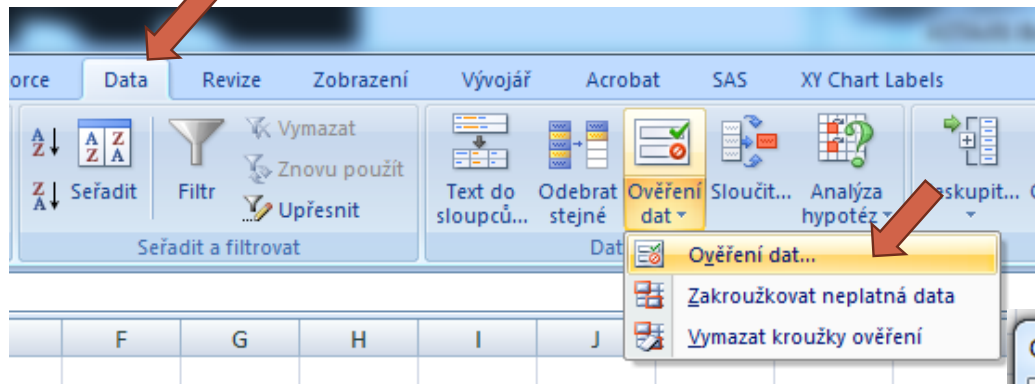
- Vytváří se z hodnot buněk v daném sloupci a umožňují vložit hodnotu výběrem ze seznamu již zadaných hodnot – usnadnění zadávání

Vyzkoušej !!!

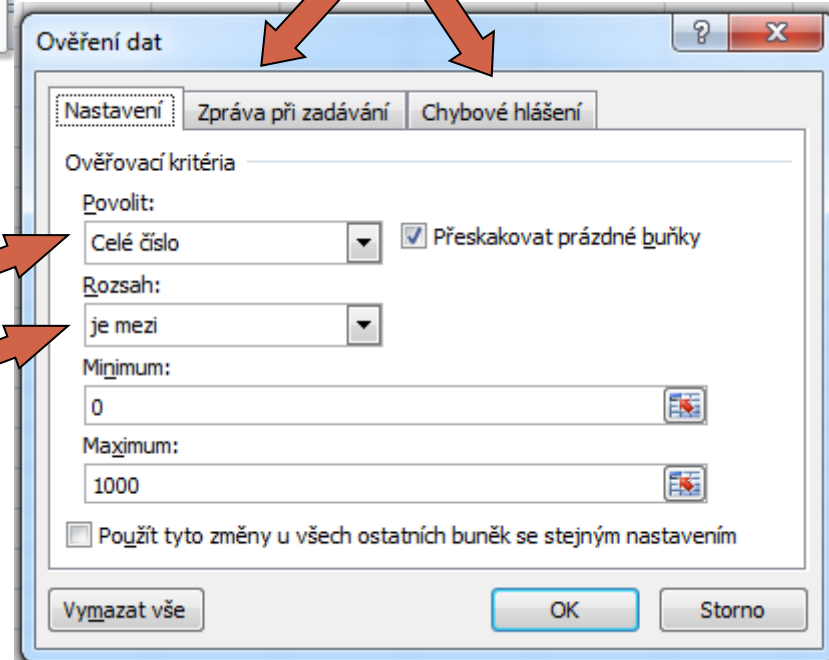
Automatická kontrola dat



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



komunikace s uživatelem



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

Rozsahy hodnot, načtení seznamů apod.

Vyzkoušej !!!

Seznamy I.

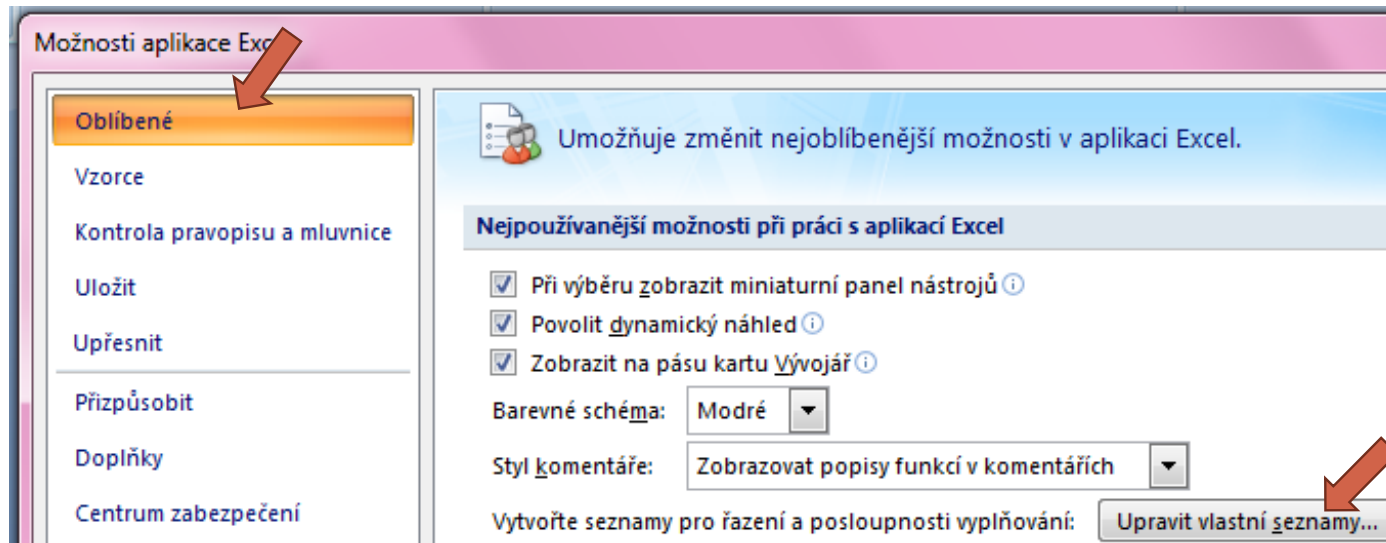


- Skupiny hodnot zachovávající logické pořadí, některé jsou zabudované (např. dny v týdnu, měsíce v roce), další je možné uživatelsky vytvořit, slouží pro účely řazení a automatického vyplňování dat

○ „Tlačítko Office“



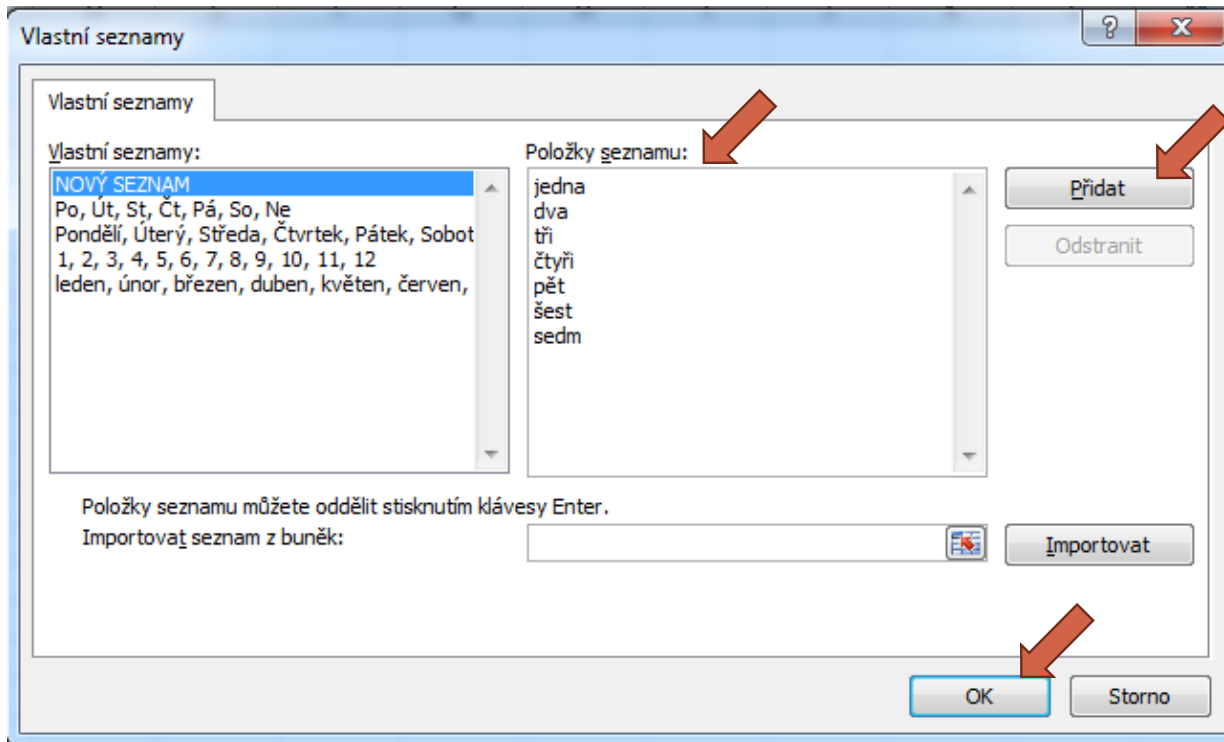
→ „Možnosti aplikace Excel“



Seznamy II.

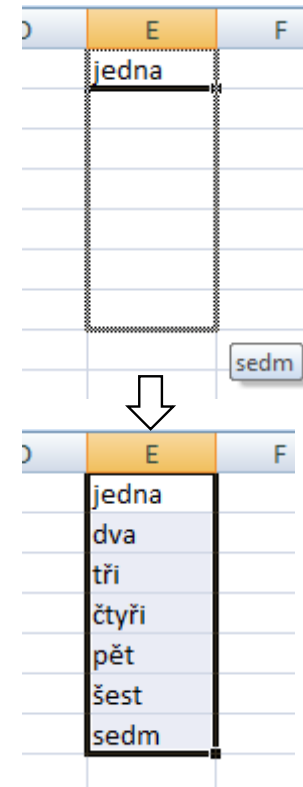
1) Definice seznamu

- Upravit vlastní seznamy →



2) Využití při tvorbě dat

Zápis jedné hodnoty ze seznamu a protažení do dalších buněk



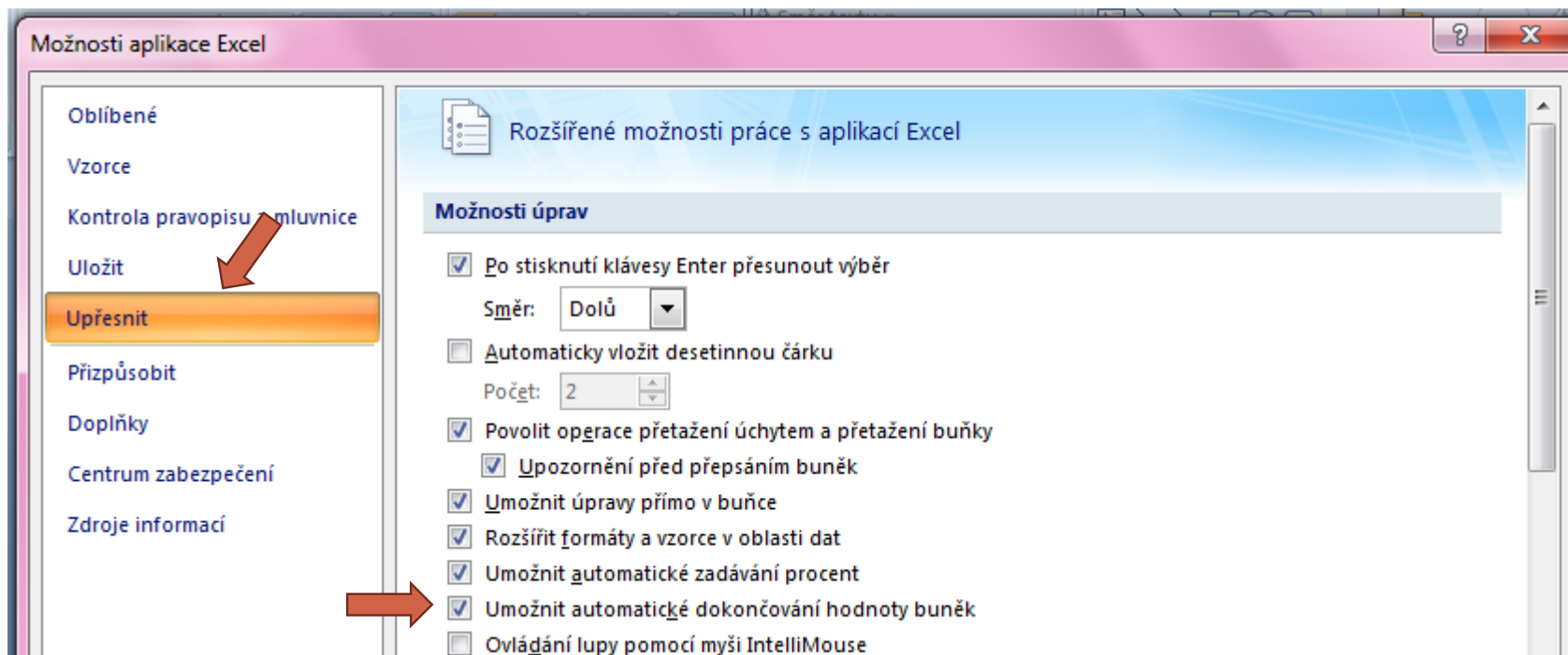
Automaticky byly doplněny následující složky seznamu

Vyzkoušej !!!

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“



Tipy a triky jak se v datech pohybovat



• Výběr buněk

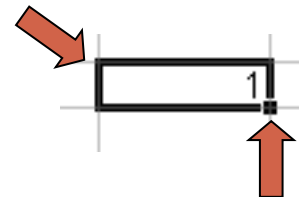
- CTRL+HOME – přesunutí na levý horní roh tabulky
- CTRL+END – přesunutí na pravý dolní roh tabulky
- CTRL+A – výběr celého listu
- CTRL + klepnutí myši do buňky – výběr jednotlivých buněk
- SHIFT + klepnutí myši na jinou buňku – výběr bloku buněk
- SHIFT + šipky – výběr sousedních buněk ve směru šipky
- 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

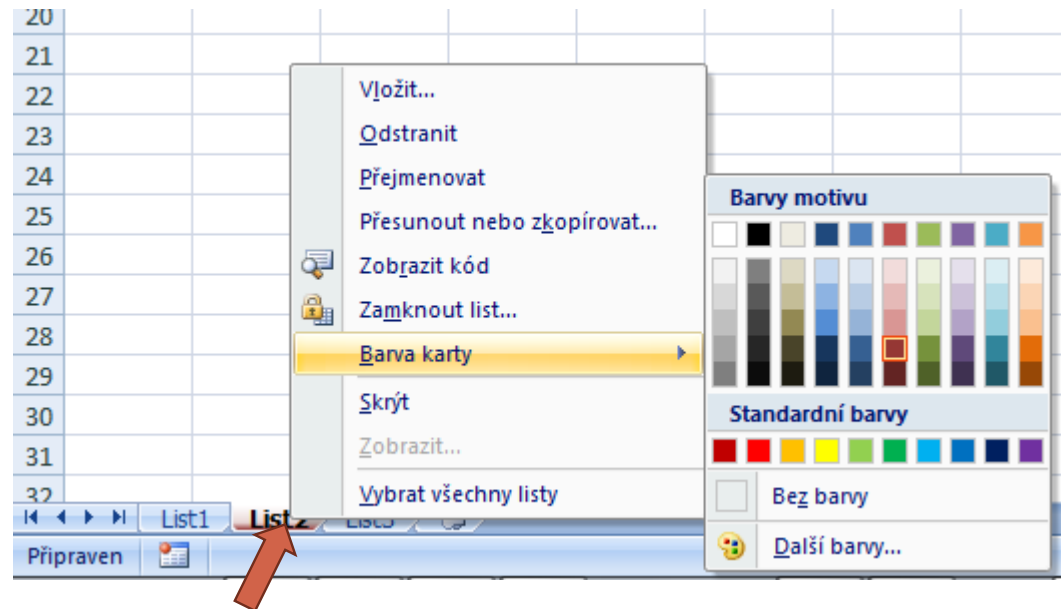
- Chycení myši za okraj umožňuje přesun buňky nebo bloku buněk
- Při chycení čtverečku v pravém dolním rohu výběru je tažením možno vyplnit více buněk hodnotami původní buňky (ve vzorcích se mění relativní odkazy, je také možné vyplnění hodnotami ze seznamu – např. po sobě jdoucí názvy měsíců).



Editace listů



- Excelovský soubor (sešit) se skládá z více listů, kde každý list může obsahovat na sobě nezávislá data a výstupy.
- Možnosti editace listu: vytvářet nový list/kopii existujícího listu, mazat, pojmenovat, obarvit záložku listu, uzamknout



Klik pravým tlačítkem
myši na záložku listu

III. Práce s daty v MS Excel aneb jak se v datech vyznat

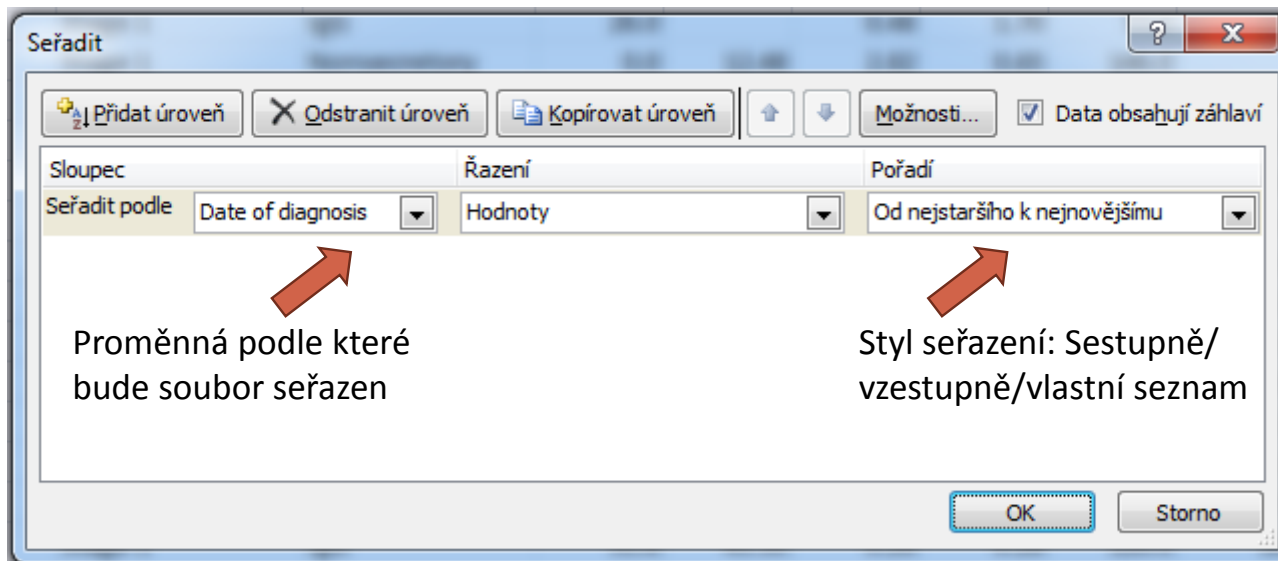
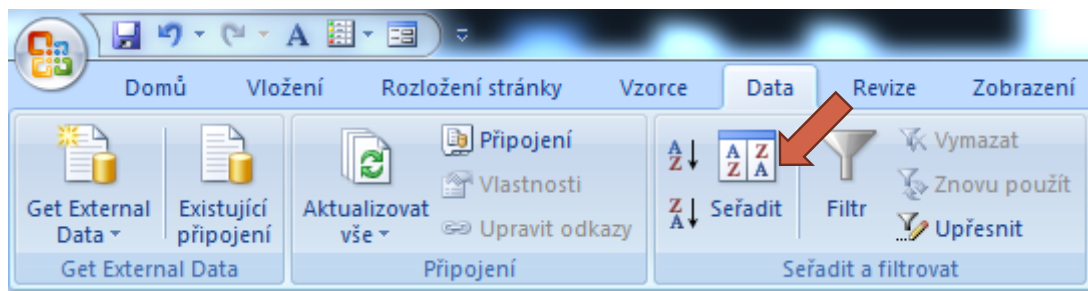


Seřazení dat
Filtrování
Ukotvení příček
Podmíněné formátování

Řazení dat



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



Proměnná podle které bude soubor seřazen

Styl seřazení: Sestupně/vzestupně/vlastní seznam



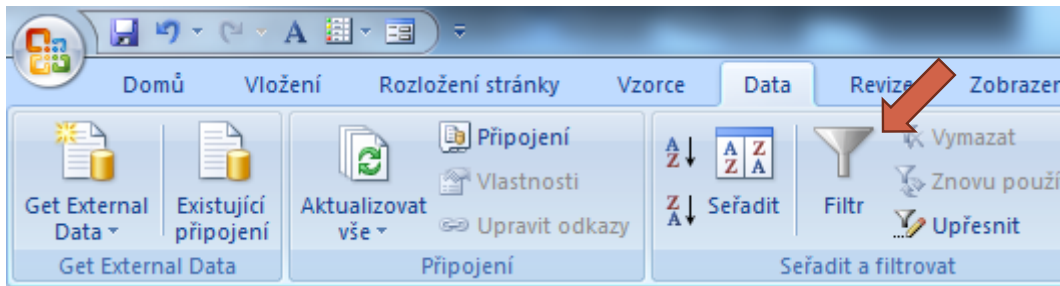
Zkontrolujte, zda seřazení nezničí vazby mezi buňkami = kontrola oblasti, kterou řadíte.

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
- **Výhodné pro čištění dat (vyhledávání překlepů, kombinace textu a čísel)**

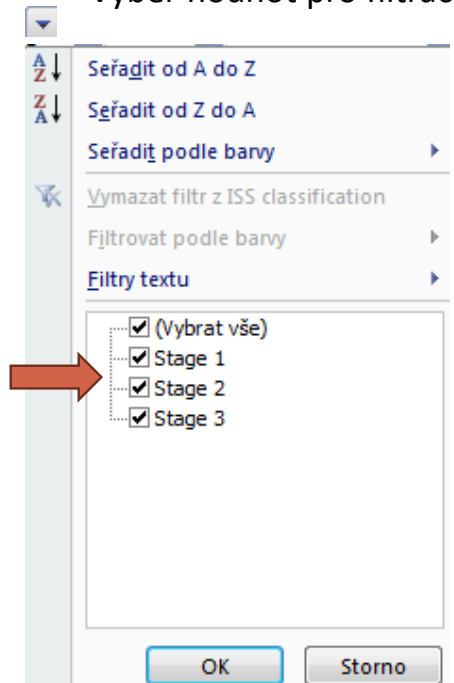
1. Zapnutí filtru (alternativa klávesová zkratka CTRL+SHIFT+L)



2. Objeví se rozbalovací šipka s výčtem všech unikátních hodnot v daném sloupci dat

	A	B	C	D	E
1	ID	Date of birth	Date of diagnosis	Sex	ISS classification
2	189	26.12.1959	29.6.1994	F	Stage 1
3	5	9.8.1945	3.3.1997	M	Stage 1
4	192	19.7.1935	2.10.1997	M	Stage 2

Výběr hodnot pro filtraci

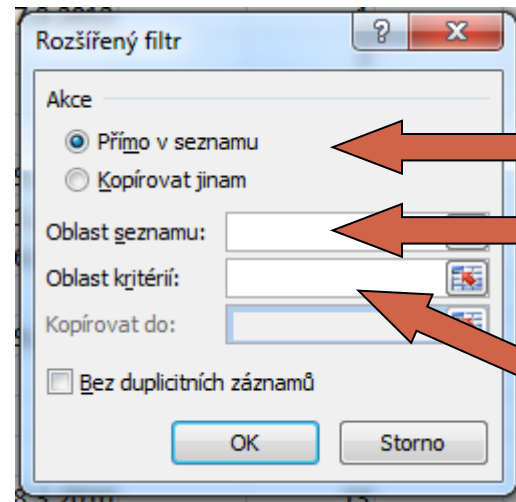
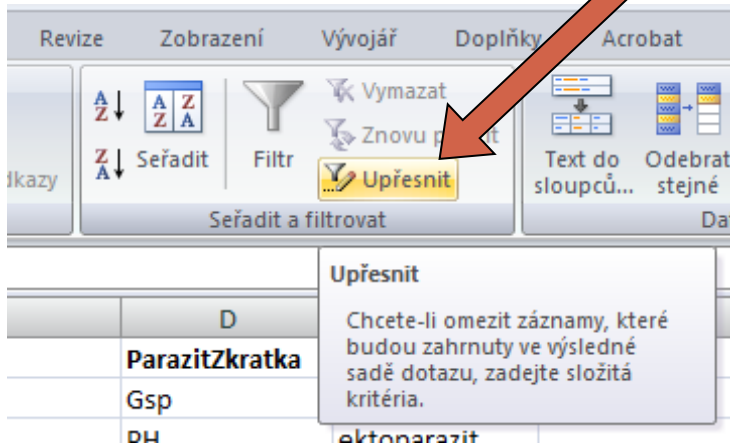


Rozšířený filtr



- 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



Výběr oblasti cílových hodnot (přefiltrovaných)

Původní seznam včetně záhlaví

Oblast kritérií včetně záhlaví

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:

	F	G	H	
	poslední kontrola	pohlaví	nemocný	tíž
9	9.4.2010	muž		1
0	29.3.2010	muž		1

Ukotví řádky nad označenou buňkou a sloupce vlevo od označené buňky

- Odstranění ukotvení: Po ukotvení příček se automaticky možnost „Ukotvit příčky“ změní na „Uvolnit příčky“

Podmíněné formátování

- Záložka „Domů“ → „Podmíněné formátování“.
- Barevné označení buněk nebo výplň buňky symbolem podle námi zadaných kritérií, např:
 - numerická hodnota větší/menší než průměr
 - datum z konkrétního období
 - podobná slova
 - duplicitní údaje
- Co s barevnými buňkami?
- Použijeme filtr!

	123.0	320	2.35	41.1
0.45	129.0	218	2.36	48.5
	96.0	191	2.37	45.2

IV. Vzorce v Excelu

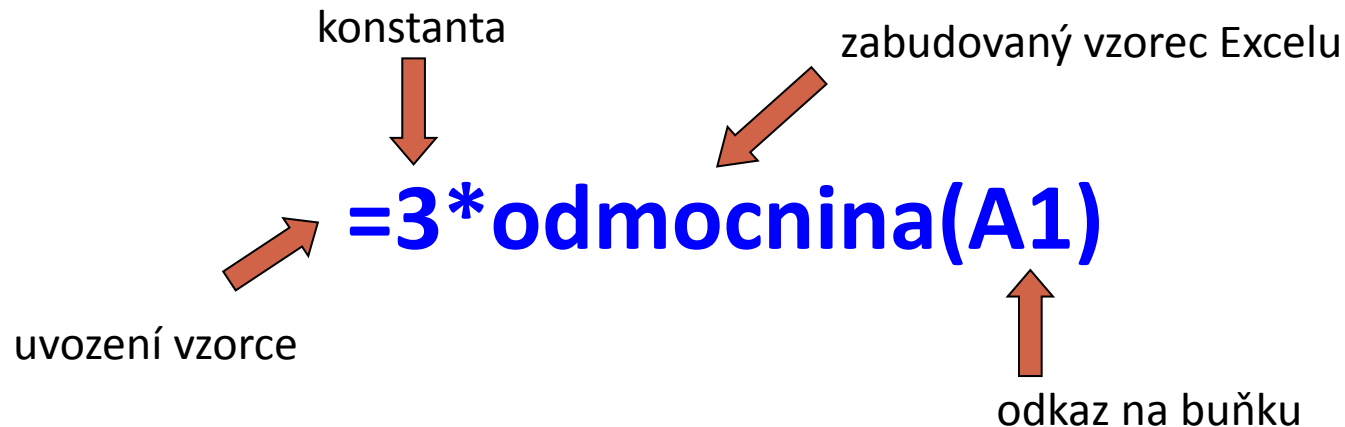


Tvorba vzorců
Absolutní a relativní odkazy
Kopírování vzorců

Vzorce



- vpisují se do buněk sešitu
- vzorce jsou vždy uvozeny = (lze též + -)
- aritmetické operátory + zabudované funkce Excelu
- pro „sčítání“ nečíselných položek se používá &
- výpočet je založen buď na číselných konstantách nebo odkazech na buňky



Vzorce – odkaz na buňku



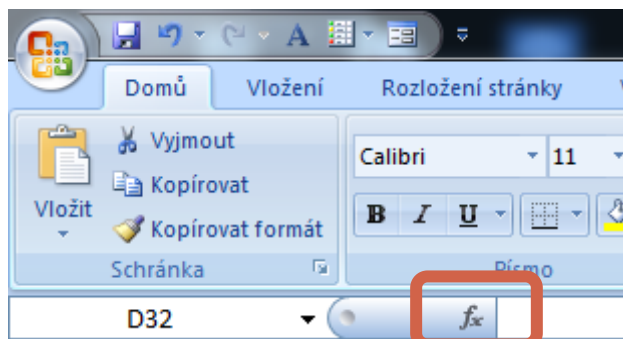
- **Relativní odkazy**
- **A1** = sloupec A, 1. řádek
- **A1:B6** = blok buněk – levý horní roh: buňka A1, pravý dolní: buňka B6
- relativní odkaz se při automatickém vyplnění buněk vzorcem posune
- mění se s kopírováním, při vložení a odstranění řádku nebo sloupce

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

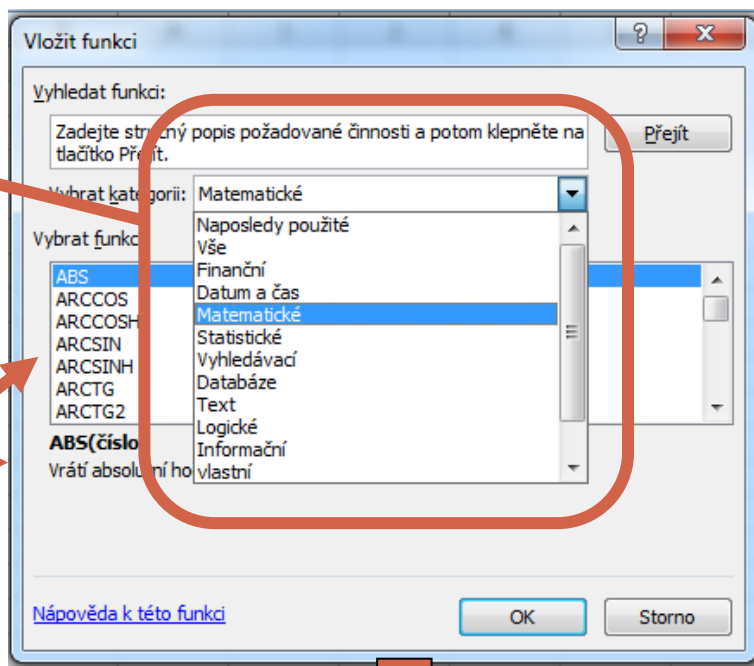
uzamčení řádku → **\$A\$1** ← uzamčení sloupce

Pamatuj: Adresu upevníme pomocí znaku \$

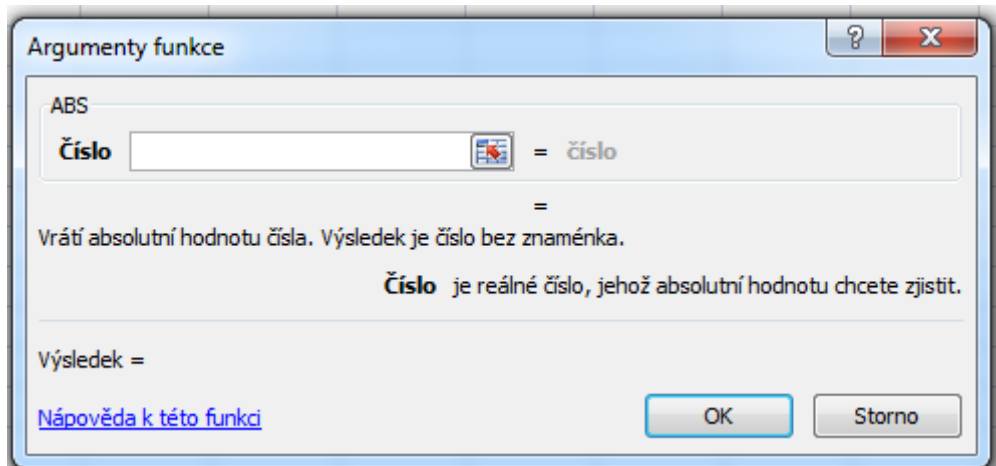
Vzorce – využití seznamu vzorců



Kategorie vzorců

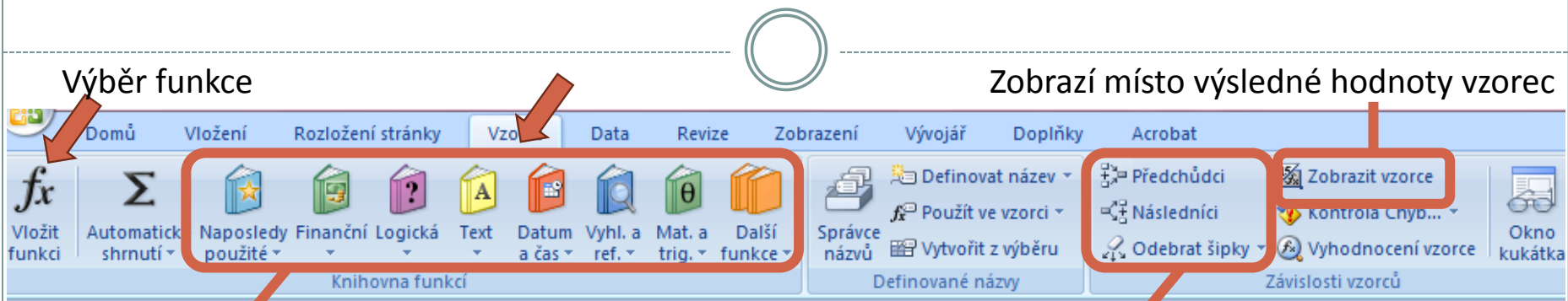


Funkce a její stručný popis



průvodce funkcí

Tvorba a závislosti vzorců



Výběr funkce z konkrétních knihoven

Označení a odznačení předchozích a následných vzorců

	A	B	C	D
1	Hemoglobin level (g/l)			Průměr
2	77.0			114.1
3	146.0			
4	87.0			
5	106.0			
6	125.0			
7	133.0			
8	117.0			
9	130.0			
10	86.4			
11	122.0			
12	123.0			
13	129.0			
14	96.0			
15	120.0			
16				

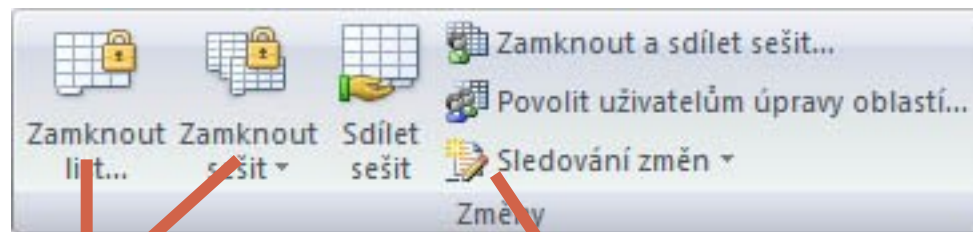
Zpřehlednění vzorců: ALT+ENTER

Komentáře, sledování změn



- karta „Revize“

Komentáře k jednotlivým buňkám



možnost zamknout list či sešit dvojicí hesel

sledování změn a jejich schvalování nebo zamítní

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



Vyzkoušej !!!

Od verze 2007 obsahuje MS Excel některé pokročilé statistické funkce

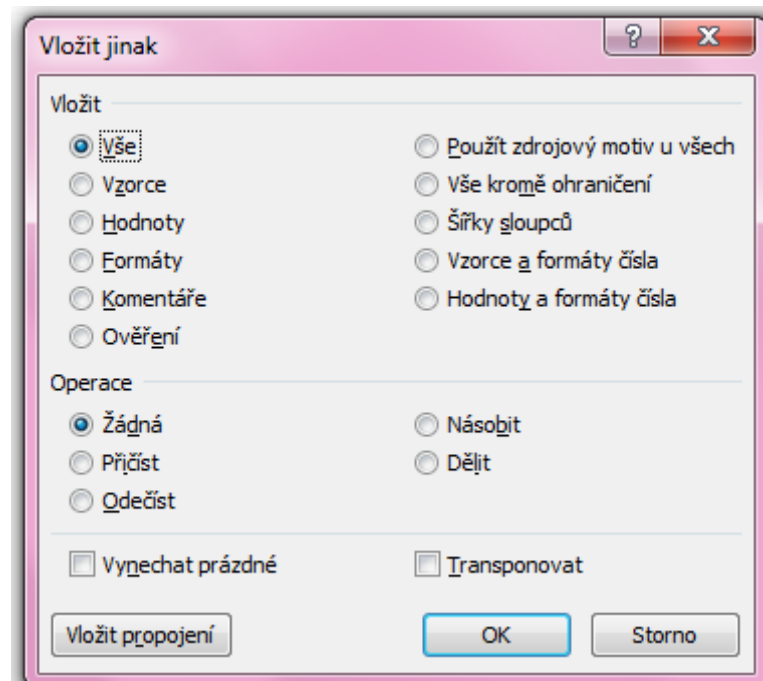
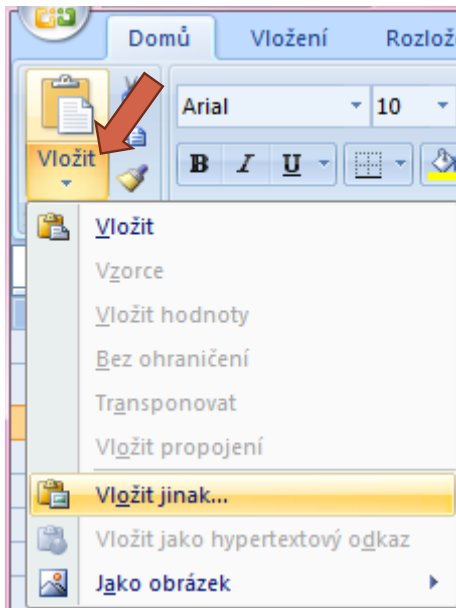
- **ZLEVA, ZPRAVA, ČÁST** – funkce pro ořezávání textových řetězců;
- **STANDARDNÍ MATEMATICKÉ FUNKCE (SIN, COS, LOG, LOGZ, EXP)** – a mnoho dalších.
- **CONFIDENCE** – výpočet intervalu spolehlivosti (při normálním rozdělení);
- **CORREL, PEARSON** – výpočet Pearsonova korelačního koeficientu;
- **COVAR** – 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)

V. Samostatné cvičení



Úkoly



- I. Vytvořte kopii listu zadání a nazvěte ji výsledky. (**nápověda**: přesunout nebo zkopírovat list → vytvořit kopii)
- II. Ukotvěte ID pacientů a názvy proměnných ve sloupcích. (**nápověda**: funkce ukotvení příček)
- III. Pomocí filtru zkontrolujte kvalitu zadaných dat. Nalezněte alespoň 3 překlepy. Kolik pacientů má jedno osteolytické ložisko („**Osteolytic lesions X-ray**“ = 1 *osteolytic lesion*)?
- IV. Seřadte datový soubor dle proměnné „**Date of diagnosis**“.
- V. Pomocí podmíněného formátování nalezněte duplicitní záznamy. (**nápověda**: označ sloupec ID → karta „Domů“ → podmíněné formátování → zvýraznit pravidla buněk → duplicitní hodnoty → filtrovat podle barvy). Duplicitní údaj smažte.
- VI. Spojte „**M-protein type**“ a „**Light chain type**“ do nového sloupce s názvem „**Heavy-light chain type**“ tak, aby bylo odděleno pomlčkou (např. IgG-kappa). (**nápověda**: využijte např. funkce & a pro oddělení slov pomlčkou “-”)
- VII. Spočítejte stáří pacienta při diagnóze (z „**Date of diagnosis**“ a „**Date of birth**“).
- VIII. Z data „**Date of diagnosis**“ vyberte pouze rok. (**nápověda**: využijte funkci z Knihovny – Datum a čas)
- IX. Sloupec „**Hemoglobin level (g/l)**“ překódujte pomocí funkce „když“ následovně: $>120 \text{ g/l} = 0$, $\leq 120 \text{ g/l} = 1$.
- X. Převeďte „**Hemoglobin level (g/l)**“ na „**Hemoglobin level (g/dl)**“.
- XI. Spočítejte minimální, maximální a průměrnou hodnotu „**Hemoglobin level (g/dl)**“