

ANALÝZA DAT NA PC I.

I. Využití aplikací MS Office pro ukládání a základní sumarizace biologických dat

Analýza dat na PC I.

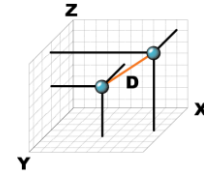
Obsah semináře

1. MS Office v analýze ekologických dat- databáze, algoritmování

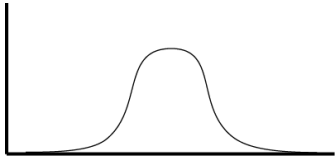


2. Algoritmy pro ekologická data I. Charakteristika společenstev, diverzita

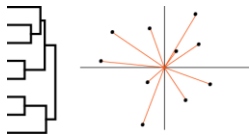
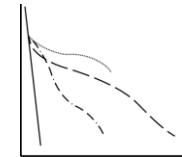
3. Algoritmy pro ekologická data II. Metriky podobností a vzdáleností



4. Algoritmy pro ekologická data III. Valence, nika

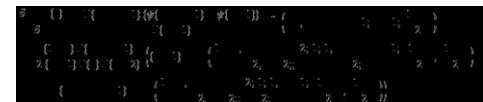


5. Algoritmy pro ekologická data IV. „Species-abundance“ modely, interpretace



6. Vícerozměrné metody v ekologii – shluková, ordinační analýza

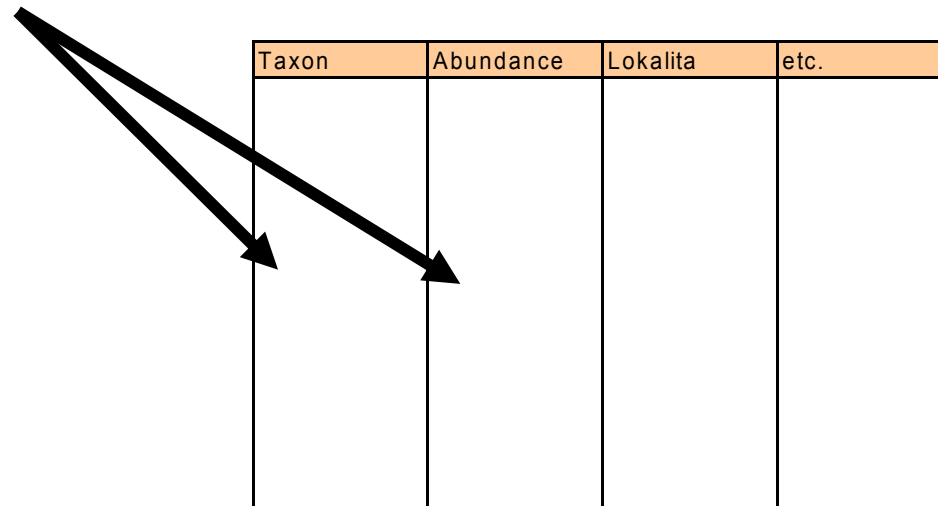
7. Užití diferenciálních rovnic v deterministických modelech v biologii a ekologii



Analýza dat na PC I.

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 (např. rozepsané taxonomické zařazení, abundance, místo a vlastnosti odběru atd.)



Taxon	Abundance	Lokalita	etc.

- ◆ Takto uspořádaná data je v tabulkových nebo databázových programech možné převést na libovolnou výstupní tabulku

Analýza dat na PC I.

Ukládání dat v MS Office

◆ MS Excel

- 📄 Kontingenční tabulky – rychlá sumarizace rozsáhlých tabulek
- 📄 Možnost výpočtů a grafových výstupů přímo v aplikaci
- 📄 Visual Basic – složitější aplikace
- Omezení tabulky na 256 × 65536 buněk
- Omezená kontrola chyb při zadávání



◆ MS Access

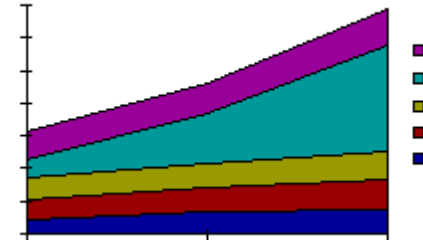
- 📄 Plnohodnotná databáze vhodná pro velké množství dat, řádky omezeny v podstatě jen dostupnou pamětí
- 📄 Kontrola typu dat
- 📄 Relace tabulek – omezení velikosti souboru
- 📄 Visual Basic a formuláře – složitější aplikace
- Omezení tabulky na 255 sloupců
- Výpočty a grafy jsou složitější než v Excelu



Analýza dat na PC I.

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)



P. bini	2		
Poččet z Délka	Pohlaví		
Číslo ryby2	Číslo rvt	Váha	?
1			
2			
3			
4			
5			
6			
7	26		
8	106		
9	121		
10	160		
11	34		
12	45		
13	70		
14	72		
15	87		
16	Celkový součet		
17			

10			
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			



Import a export dat

◆ Import dat

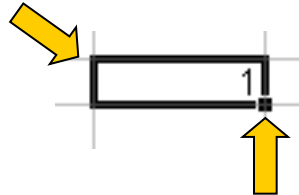
- ◆ Manuální zadávání
- ◆ import – podpora importu ze starších verzí Excelu, textových souborů, databází apod.
- ◆ kopírování přes schránku Windows – vkládání z nejrůznějších aplikací – MS Office, Statistica atd.
- ◆ využití textových souborů jako kompatibilního formátu pro přenos dat mezi různými aplikacemi

◆ Export dat

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

Tipy a triky

- ◆ **Výběr buněk**
 - ◆ 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 buněk**
 - ◆ 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).

Analýza dat na PC I.

Databázová struktura dat v Excelu

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

Jednotlivé záznamy
(taxon, lokalita 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	?	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	

Analýza dat na PC I.

Automatický zadávací formulář

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

The image illustrates the process of using the automatic data entry form in Excel. On the left, the 'Data' menu is open, and the 'Formulář...' option is highlighted. An arrow points from this menu item to the 'data' dialog box on the right. The dialog box contains several input fields with labels: 'Číslo ryby' (with a value of 1), 'Značka ryby' (with a value of 1), 'Společ číslo' (with a value of 1), 'Pohlaví' (with a value of m), 'Délka' (with a value of 27,5), 'Váha' (with a value of 23), 'P. anguillae' (with a value of 2), and 'P. bini' (with a value of 2). To the right of these fields are several buttons: 'Nový', 'Odstranit', 'Obnovit', 'Předchozí', 'Další', 'Kritéria', and 'Zavřít'. Yellow arrows point to these elements with labels: 'Nový záznam' points to the 'Nový' button, 'Vyhledávání' points to the 'Kritéria' button, 'Názvy sloupců' points to the labels of the input fields, and 'Obsah dané buňky - editovatelný' points to the input field for 'Číslo ryby'. The dialog box also shows '1 z 19' in the top right corner.

Analýza dat na PC I.

Automatické seznamy

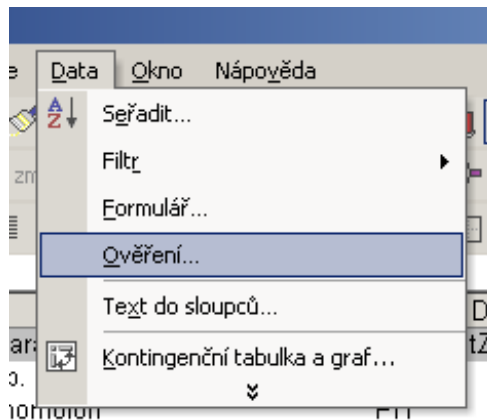
- ♦ Vytváří se z hodnot buněk v daném sloupci a umožňují vložit hodnotu výběrem ze seznamu již zadaných hodnot – usnadnění zadávání

Sloupec z něž je seznam vytvořen a pro který platí

The image shows a spreadsheet with columns labeled 'Taxon', 'Abundance', 'Lokalita', and 'etc.'. The 'Taxon' column is highlighted in light blue. A context menu is open over the 'Taxon' column, with the option 'Vybrat ze seznamu...' (Select from list...) highlighted. A dropdown list is shown below the menu, listing various taxonomic names such as 'Piscicola geometra (Linnaeus, 1761)', 'Acanthocephallus lucii (Müller, 1776)', and 'Argulus foliaceus (Linnaeus, 1758)'. A yellow arrow points from the text 'Buňka, do níž se vloží vybraná hodnota' (Cell into which the selected value will be placed) to the 'Taxon' column. Another yellow arrow points from the text 'Sloupec z něž je seznam vytvořen a pro který platí' (Column from which the list is created and for which it applies) to the 'Taxon' column header.

Automatická kontrola dat

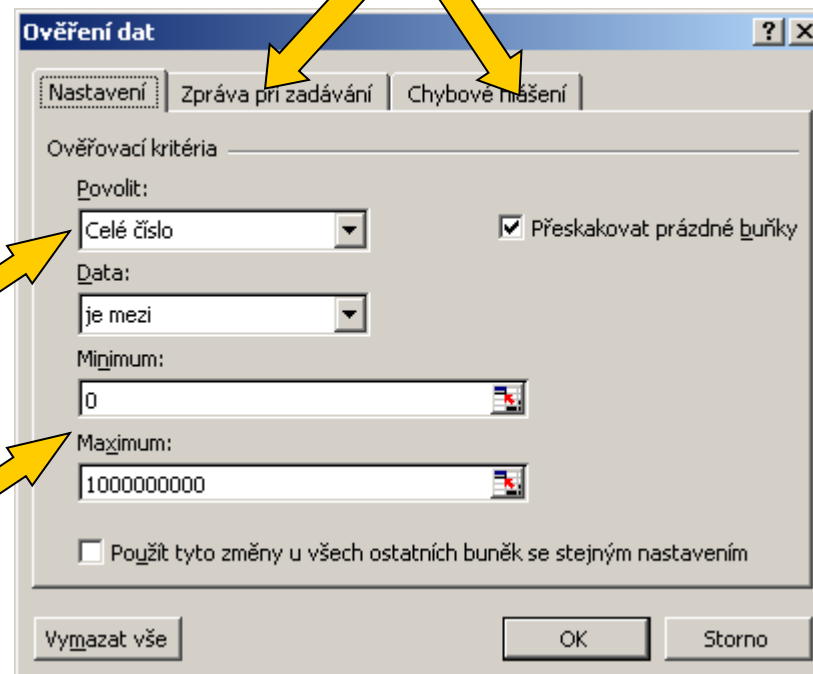
- Umožňuje ověřit typ, rozsah nebo povolit pouze určitý seznam hodnot zadávaných do sloupce databázové tabulky



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

Rozsahy hodnot, načtení seznamů apod.

komunikace s uživatelem



Analýza dat na PC I.

Seznamy

- ◆ 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

Existující seznamy

The screenshot shows the 'Možnosti' dialog box with the 'Vlastní seznamy' tab selected. The 'Vlastní seznamy' list contains 'NOVÝ SEZNAM' and several predefined lists. The 'Importovat' button is highlighted with a yellow arrow. A yellow arrow points from the 'Možnosti...' menu item to the dialog box. Another yellow arrow points from the 'Importovat' button to the text 'Načtení nového seznamu'.

Výběr buněk pro nový seznam

Načtení nového seznamu

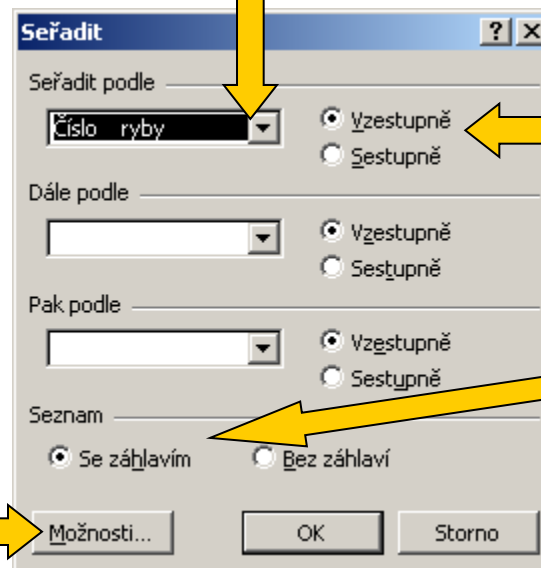
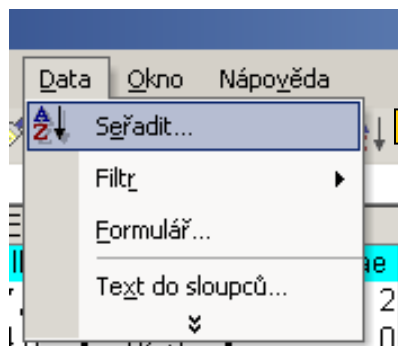
Řazení dat

- ◆ 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

Analýza dat na PC I.

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 lokalita v určité sezóně)
- ◆ Funkce automaticky rozezná hlavičky sloupců v souvislé oblasti buněk
- ◆ U sloupců použitých pro filtraci jsou rozbalovací seznamy zbarveny modře

Výběr hodnot pro filtraci

Rozbalení seznamu hodnot nalezených ve sloupci

	A	B	C	D	E
	Číslo	Značka	Společ	Ohlav	Délka
1					
2		1			27,5
3		2			34,0
4	3	5			58,0
5	4	6			42,0
6	5	7			44,0
7	6	8		f	56,0
8	7	9		m	48,5

Automatický filtr – vlastní filtrování

- Umožňuje filtrovat data databázové tabulky na základě rozsahů hodnot ve sloupcích

podmínka

hodnota

Vlastní automatický filtr

Zobrazit řádky

Délka

Je rovno 27,5

A Nebo

Je rovno

Zr
Není rovno

Zr
Je větší než

Je větší než nebo rovno

Je menší než

Je menší než nebo rovno

Má na začátku

17 Nemá na začátku 0 108,0 1

18 Má na konci 0 130,0 0

19 Nemá na konci 0 85,0 2

OK Storno

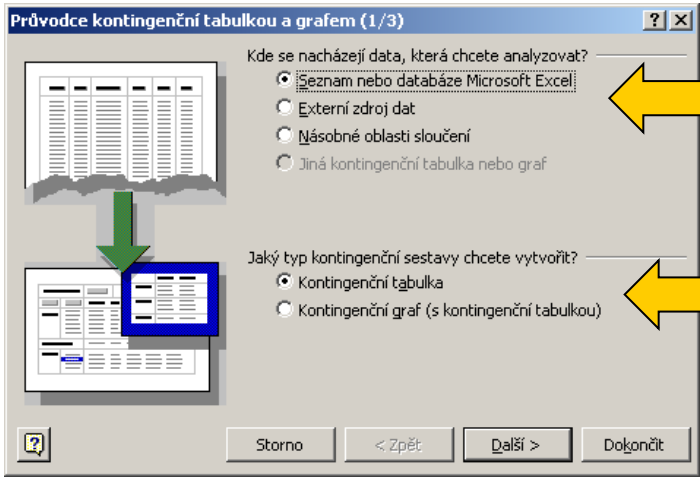
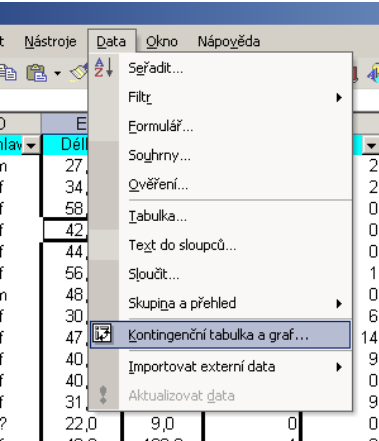
Logické operátory

Seznam podmínek

Analýza dat na PC I.

Kontingenční tabulka

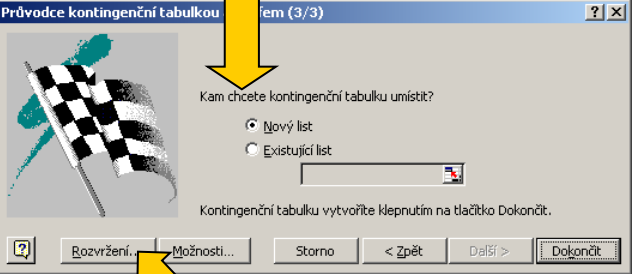
- ◆ Umožňuje snadno vytvářet sumarizace dat ve smyslu počty hodnot, průměry, minima, maxima atd. v kombinacích kategorií (např. počet jedinců různých druhů na různých lokalitách)
- ◆ Automaticky je vybrána souvislá oblast dat (obdobě jako v případě automatického filtru)



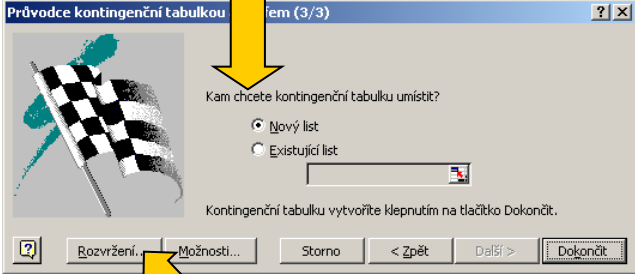
Zdroj dat (kromě Excelu i např. externí databáze)

Graf nebo tabulka

Umístění



Rozvržení a vlastnosti tabulek



Zdrojová oblast dat

Kontingenční tabulky - rozvržení

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

tzv. stránka = tabulky podle zde nastaveného kritéria parametry sloupců

parametry na řádcích

parametry dat a možnosti sumarizace

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

Pole kontingenční tabulky

Zdrojové pole: Délka

Název: Počet z Délka

Souhrn:

- Součet
- Počet
- Průměr
- Maximum
- Minimum
- Součin
- Počet čísel

OK Storno Skryt Číslo... Možnosti >>

Analýza dat na PC I.

Kontingenční tabulky - výsledek

- ◆ 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

Roletky položek tabulky

Automatický souhrn

Seznam polí tabulky

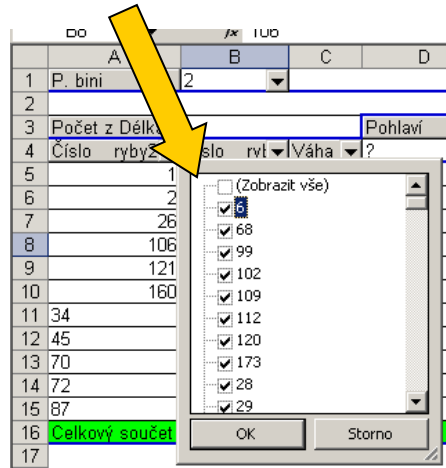
Panel nástrojů kontingenční tabulky

Počet z	Délka	Pohlaví	
Číslo ryby2	Číslo ryt	Váha	?
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

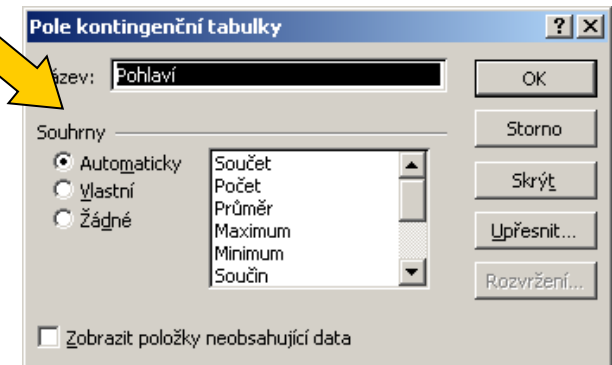
Kontingenční tabulky - nastavení

- ◆ Formát kontingenční tabulky je možné editovat obdobným způsobem jako jiné excelovské tabulky
- ◆ Lze editovat které z hlaviček sloupců a řádků budou zobrazeny, měnit jejich hierarchii a způsob řazení
- ◆ Zobrazení a způsob sumarizace dat, detailnost zobrazení různých úrovní dat

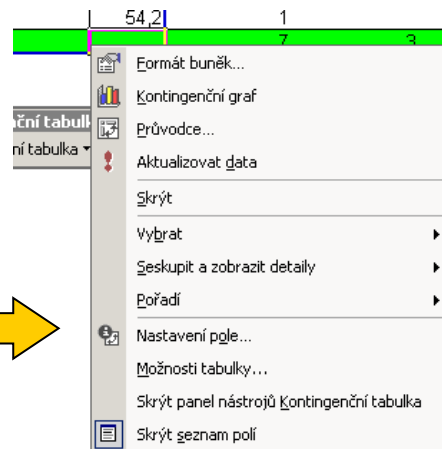
Výběr zobrazených položek



Název a způsob sumarizace položky



Kontextová nabídka buněk kontingenční tabulky

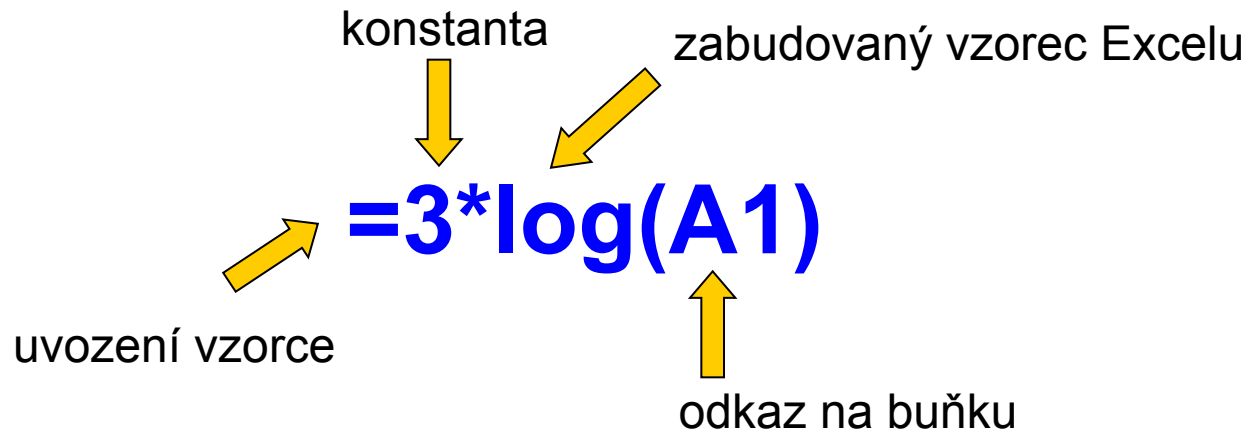


aktualizace dat
skrýt položku

možnosti tabulky

Vzorce v listu Excelu

- ◆ 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 stylu A1

- ◆ **Relativní odkazy**
- ◆ **A1** = buňka 1. řádku sloupci A
- ◆ **A1:B6** = blok buněk – levý horní roh je v 1. řádku, sloupec A, pravý dolní na řádku 6, sloupec B
- ◆ relativní odkaz se při automatickém vyplnění buněk vzorcem posune
- ◆ **Absolutní odkaz** – 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 – další možnosti odkazu na buňku

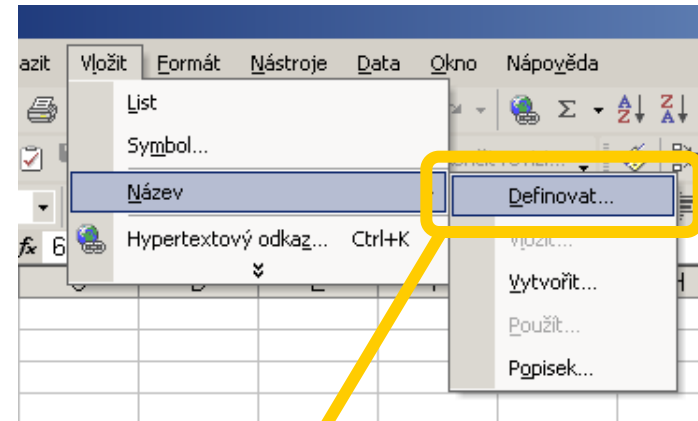
- ◆ **Pojmenované oblasti** – oblast buněk může být pojmenována a lze se na ni pomocí tohoto názvu odkazovat.

1) Vložení názvu oblasti



Pojmenovávaná oblast

2)

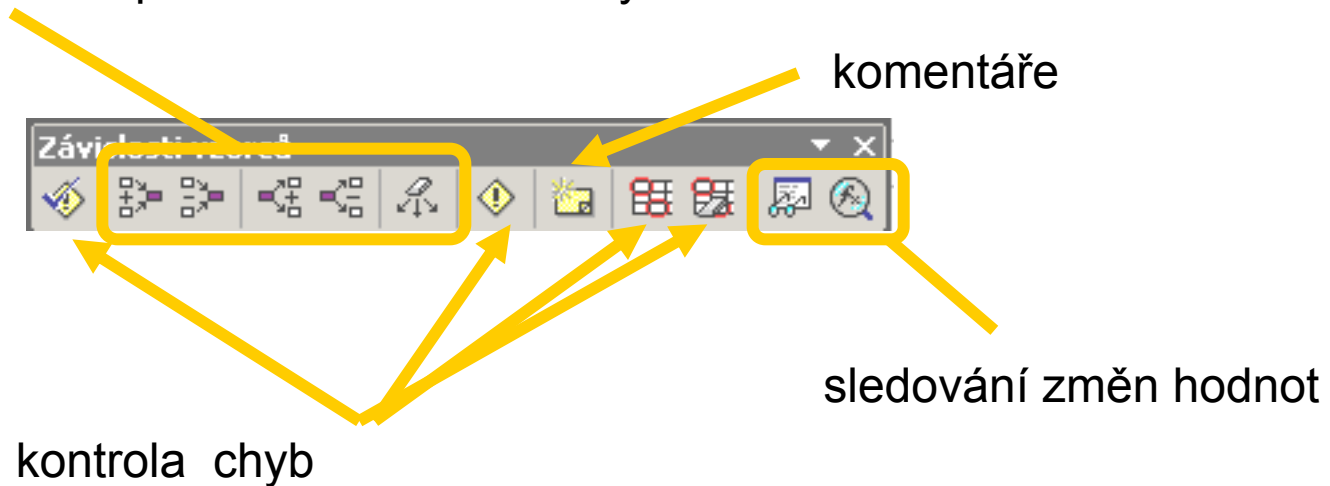


Mazání, úpravy atd. názvů oblastí.

Vzorce – tipy a triky

Závislosti vzorců – panel nástrojů

označení a odznačení předchozích a následných vzorců



Zpřehlednění vzorců

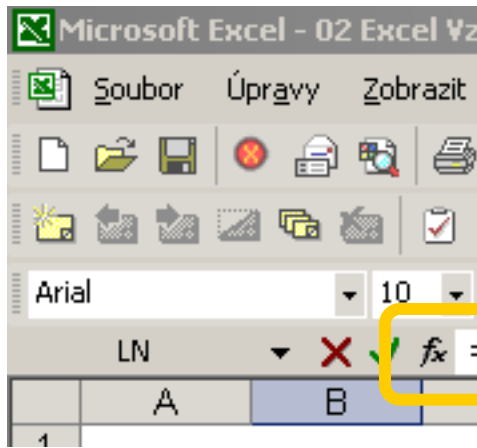
Vložit tabulátor CTRL+ALT+TAB

Vložit konec řádku ALT+ENTER

Analýza dat na PC I.

Vzorce – využití seznamu vzorců

Kategorie vzorců



Funkce a její stručný popis

The "Vložit funkci" (Insert Function) dialog box is shown. It has a search field "Vyhledat funkci:" with the instruction "Zadejte stručný popis požadované činnosti a potom klepněte na tlačítko Přejít". Below it, a list of categories is shown, with "Matematické" selected. A yellow box highlights this list. To the right, the "Argumenty funkce" (Function Arguments) dialog for the TTEST function is visible. It shows fields for "Pole1", "Pole2", "Strany", and "Typ", each with a dropdown arrow and a description: "Pole1 = matice", "Pole2 = matice", "Strany = číslo", "Typ = číslo". Below these fields, it says "Vrátí pravděpodobnost odpovídající Studentovu t-testu." and "Pole1 je první množina dat." The result field "Výsledek =" is empty. Buttons for "OK" and "Storno" are at the bottom.

průvodce funkcí

Maticové vzorce

- ♦ výpočty z matic dat
- ♦ zadávání je ukončeno stiskem CTRL+SHIFT+ENTER

16			
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			

Vzorec je založen na těchto dvou maticích dat

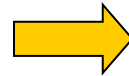
{=SUMA(A17:A23*B17:B23)}

Násobení řádků matic

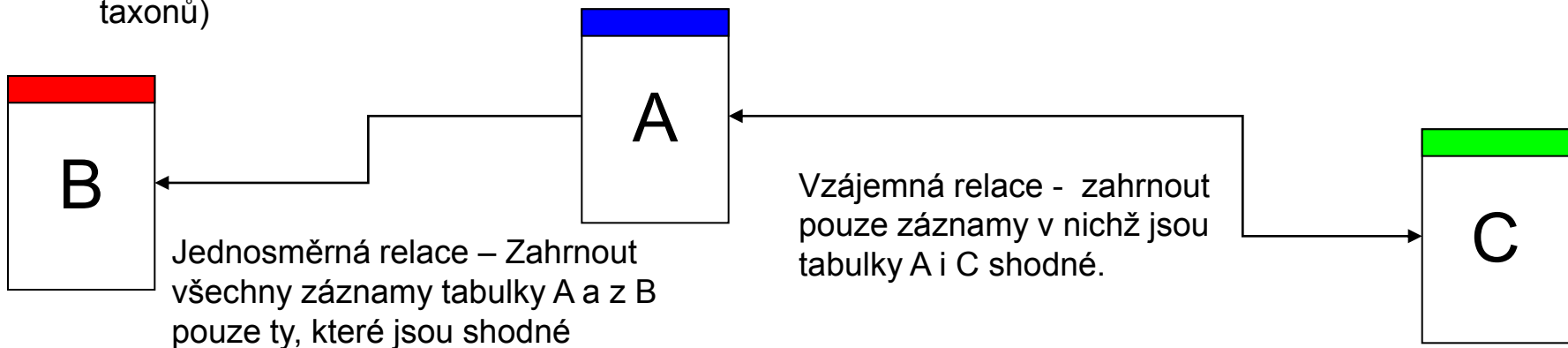
Celkové sečtení

MS Access – základní koncepce ukládání dat

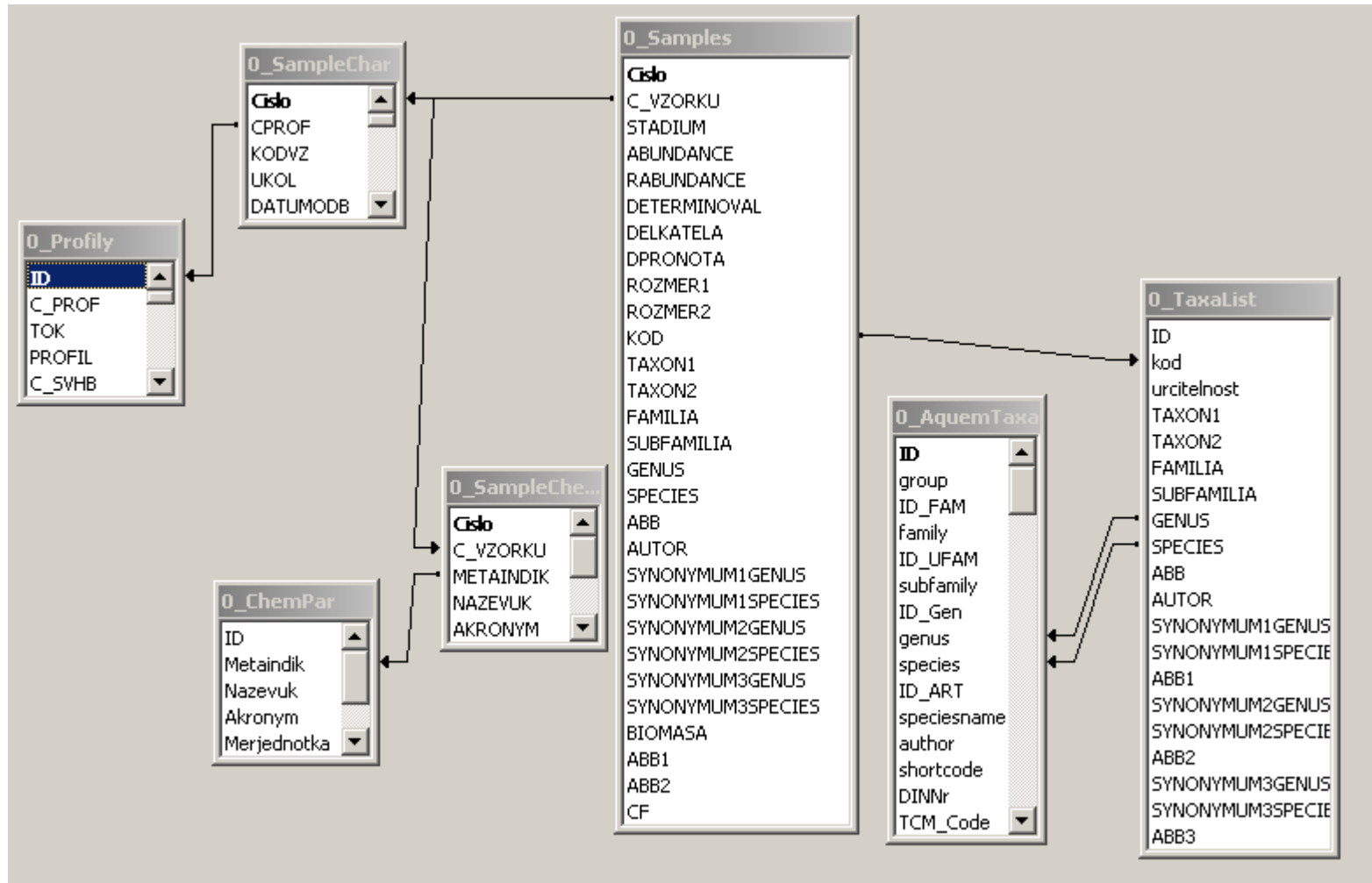
- ◆ Data jsou ukládána v databázové formě – Access je plnohodnotný databázový program
- ◆ Data jsou ukládána v tabulkách, kde každý sloupec představuje jeden definovaný typ dat, řádky tabulky se nazývají záznamy
- ◆ Je vhodné aby tabulky obsahovaly tzv. primární klíče - parametr nebo kombinace parametrů jehož/jejichž hodnoty jsou pro každý záznam jedinečné
- ◆ Tabulky spolu mohou být propojeny, tzv. relace tabulek, jsou vhodné např. pro často se opakující hodnoty (např. vlastnosti jednotlivých taxonů)



Taxon	Abundance	Lokalita	etc.

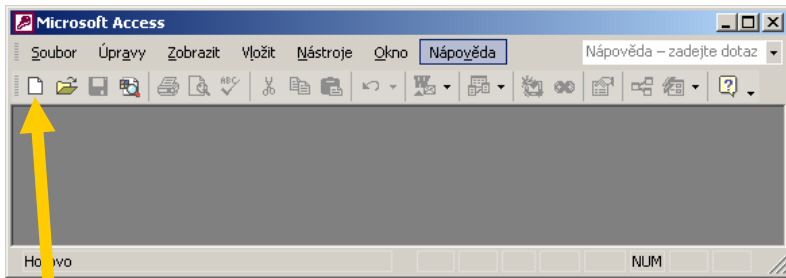


Příklad složitější relační struktury v MS Access

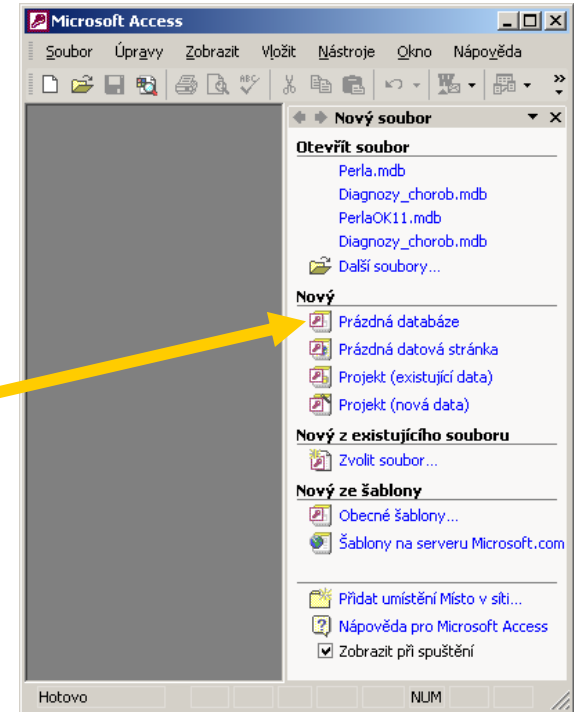


Analýza dat na PC I.

MS Access – první kroky

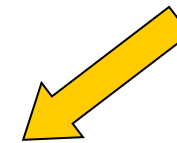
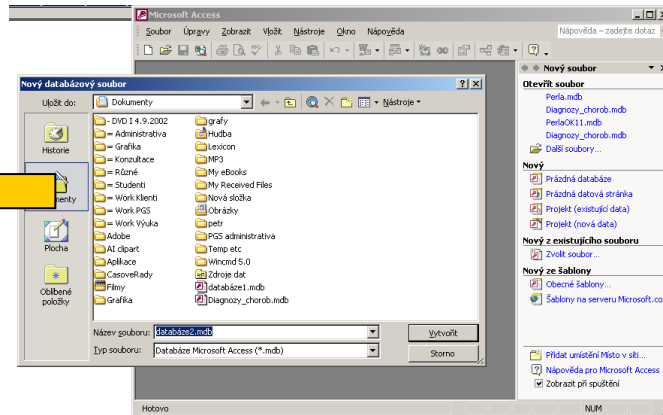
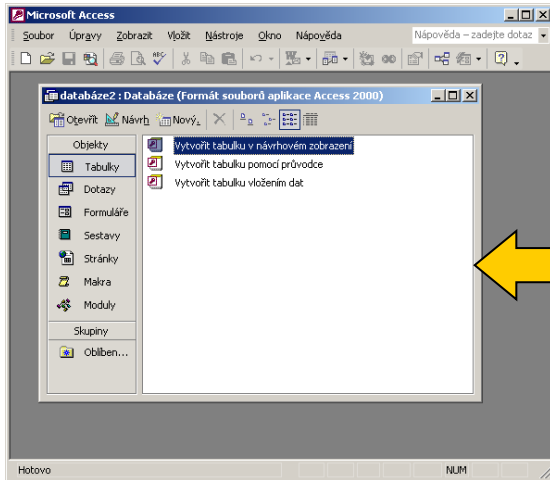


Nový soubor



Nová prázdná databáze

Uložení nové databáze (soubory v Accessu se ukládají při tvorbě souboru a pak již jen automaticky)



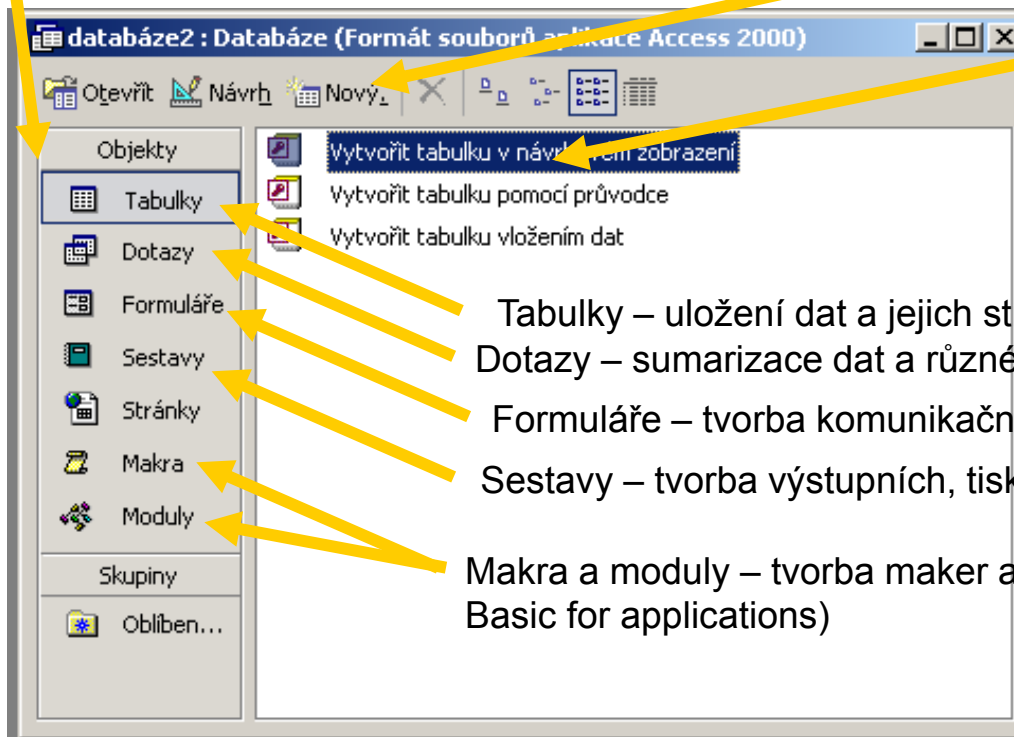
Analýza dat na PC I.

Správa databáze

Jednotlivé položky umožňují vytváření, editaci, správu a další úkony s částmi accessovské databáze

Kontextový panel nástrojů

Základní příkazy



Tabulky – uložení dat a jejich struktura

Dotazy – sumarizace dat a různé pohledy na data

Formuláře – tvorba komunikačních rozhraní až aplikací

Sestavy – tvorba výstupních, tiskových sestav

Makra a moduly – tvorba maker a programu ve VBA (Visual Basic for applications)

Analýza dat na PC I.

Tvorba nové datové tabulky

