

Úvod do programu STATISTICA

Geografický ústav PřF MU

(pracovní materiál určený pro studenty předmětu
Z0069 [Statistické metody a zpracování dat](#))

1 Dostupnost, instalace a spuštění programu

1.1 Dostupnost

- MU vlastní multilicenci, každý student může pořídit vlastní instalaci. Instalační soubory a informace potřebné k instalaci na vlastní počítač lze získat přes IS: <https://inet.muni.cz/app/soft/licence> případně v Univerzitním počítačovém centru MU, Komenského nám. 2. Oficiální stránky: www.statsoft.com , www.statsoft.cz

1.2 Instalace

- Podrobný postup instalace zde: http://www.statsoft.cz/page/index2.php?install_instructions_v71#enterprise
- Po nainstalování programu budete požádáni o registraci své licence u společnosti StatSoft. Instalace bude funkční 14 dní, plnou licenci obdržíte po odeslání registračních údajů. Údaje o registraci mohou být automaticky zaslány společnosti StatSoft pomocí emailu přímo při instalaci, případně později v registračním souboru po dokončení instalace na adresu license@statsoft.com
- Po obdržení odpovědi e-mailem uložte přiložený soubor s registrací (s příponou .lic) na plochu. Spusťte STATISTICU. Poté v menu Help→ About STATISTICA→Správa licencí→Update licence→vložit soubor

1.3 Spuštění

- Start-Programy-Statistica Cz 8.0-Statistica.cz

1.4 Základní typy oken programu STATISTIKA

Datové okno (spreadsheet) má příponu sta, jeho obsah však lze exportovat i v jiných formátech). Do datového okna lze načítat datové soubory nejrůznějších typů (např. z tabulkových procesorů, databázové soubory, ASCII soubory). Datové okno má analogickou strukturu jako list v EXCELU. Sloupce tvoří **proměnné**, řádky tvoří **případy**. Většina ovládacích prvků je stejná či podobná jako v EXCELU.

Pracovní sešit (workbook) má příponu stw. Do pracovního sešitu ukládají výstupy, tj. tabulky a grafy. Skládá se ze dvou oken, v levém okně je znázorněna stromová struktura výstupů, v pravém jsou samotné výstupy. V levém okně se lze pohybovat myší nebo kurzorem, mazat, přesouvat, editovat apod. Výstupy mohou sloužit jako vstupy pro další analýzy a grafy.

Protokol (report) má příponu str, lze ho uložit i ve formátu rtf, txt či htm. Výstupy lze ukládat od pracovního sešitu i do protokolu. Podobně jako pracovní sešit se skládá ze dvou oken. Do protokolu můžeme vkládat vlastní text, vysvětlující komentáře, poznámky apod. Tabulky a grafy lze v reportu i v pracovním sešitu dále upravovat.

Okno grafů přípona stg, lze ho uložit i jako bmp, jpg, png a wmf). Získá se tak, že v pracovním sešitu klikneme pravým tlačítkem na graf a vybereme Klonovat graf.

Programovací okno přípona svb. Slouží pro zápis programů v jazyku STATISTICA Visual Basic

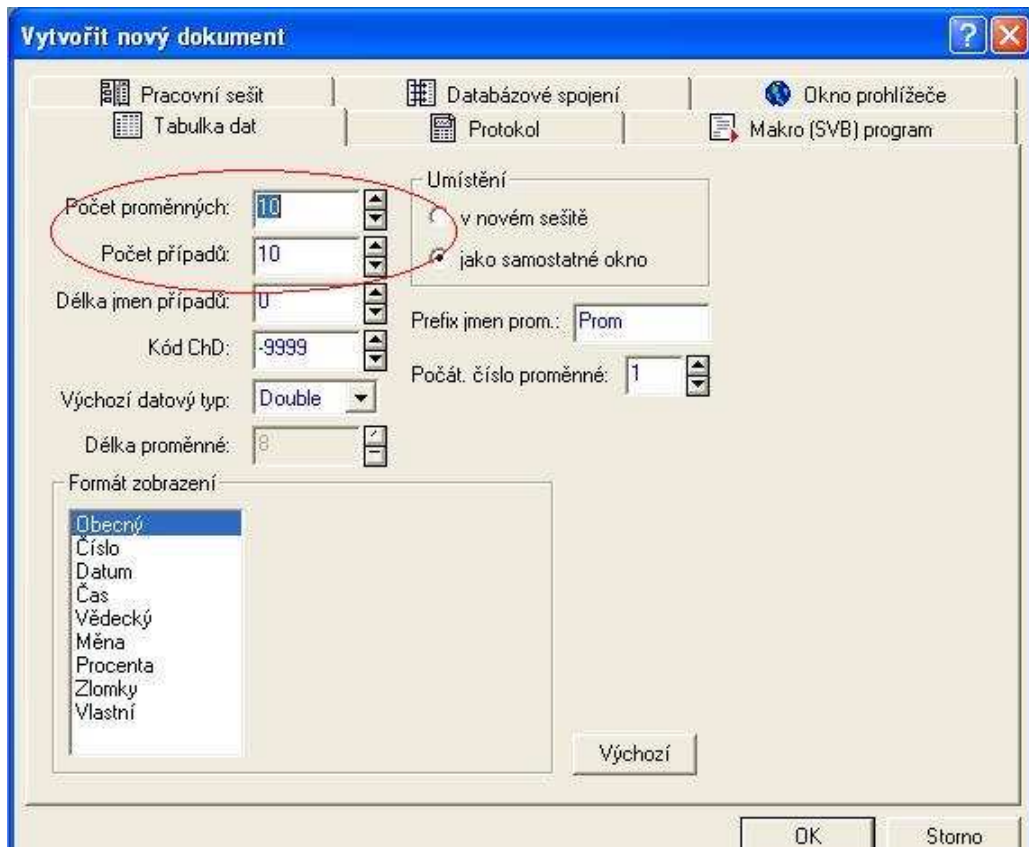
Výstupy, které směřují do pracovního sešitu lze poslat také do Protokolu: Posloupností Nástroje – Možnosti – Správce výstupů zatrhněte Také poslat do okna protokolu.

Vytvoření Okna grafu: Po vytvoření grafu v pracovním sešitě na něm klikněte pravým tlačítkem myši a z kontextové nabídky zvolte „Klonovat graf“

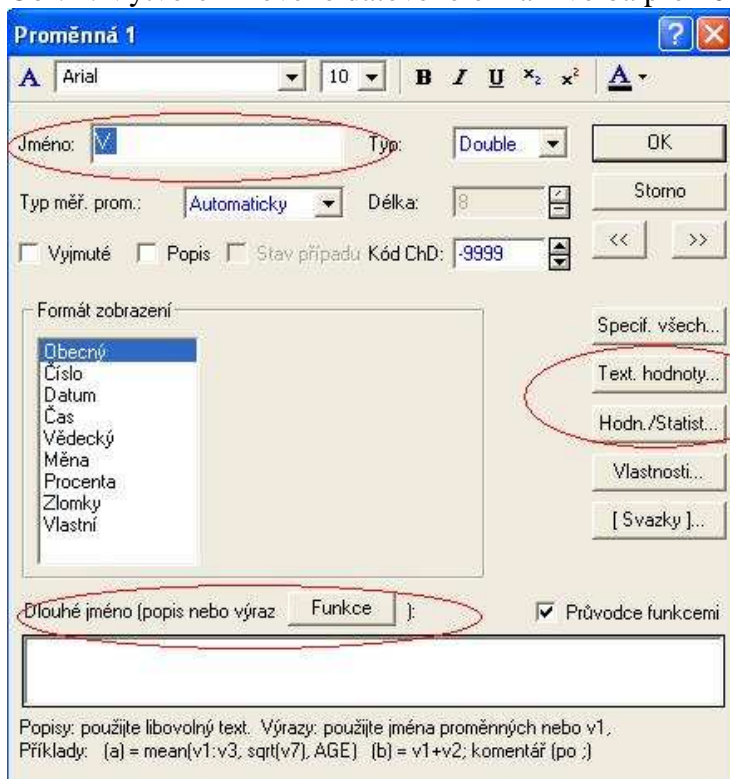
2 Základy práce s daty

2.1 Vytvoření datového okna

- Spustě program a zavřete všechna okna, která vyskočí vyjma hlavního okna aplikace Statistica
- soubor→nový→lišta Tabulka dat→zvolte potřebný počet proměnných (sloupce) a případů (řádky) – viz. obr. 1. Dále se zde dá navolit formát dat, délka jmen případů atd. (lze upravit i později).
- nastavení vlastností proměnné: kliknout na proměnnou→vyskočí nové okno→zde je možnost změny jména proměnné, formátu dat. V položce „dlouhé jméno“ lze nastavit buď dlouhý popis a informace k proměnné, ale spíše se používá pro vytvoření vzorců a transformaci proměnné. Funkce je podobná jako v EXCELU, tedy zadání vzorce (například: $=\sqrt{3}/\sqrt{2}*100$). Program Statistica má některé funkce už předem vložené – položka „Funkce“. V položce „textové hodnoty“ lze nastavit k číslům slovní hodnocení (například: u známek- 1...vynikající 2...velmi dobře 3.....dobře atd). Poté můžete přímo v datovém okně (ikona štítku v horní liště) měnit druh popisu (buď číselný nebo slovní). Položka Hodn/Statist udává celkové informace o proměnné (počet údajů, průměr, směrodatná odchylka atd). – viz. obr. 2.



Obr. 1. Vytvoření nového datového okna – volba proměnných a případů



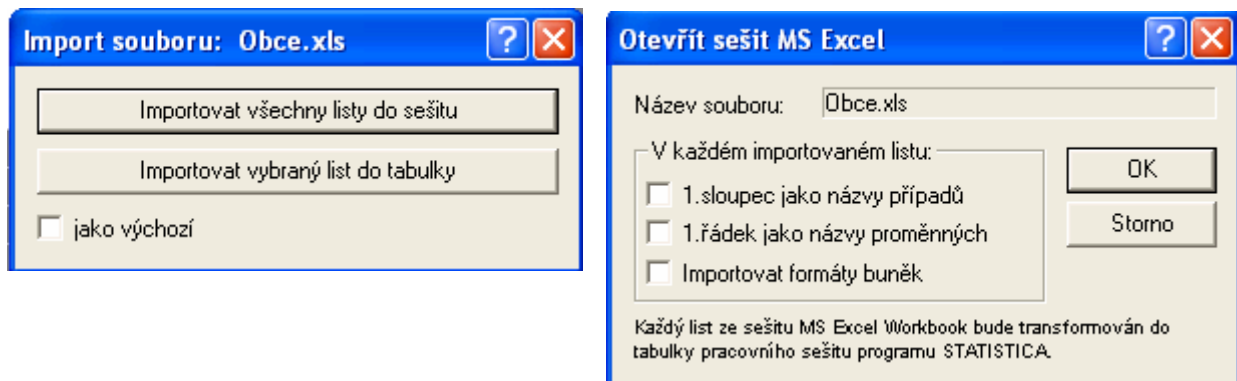
Obr. 2. Úprava proměnné

2.1.1 Úkol 1

- Vytvořte nový datový soubor o 5 proměnných a 7 případech.
- Proměnná 1: změňte jméno na „stanice“. Do datového okna vypište tyto údaje: Dukovany, Brno, Znojmo, Bítov, Jemnice, Zlín, Holešov
- Proměnná 2: nazvěte ji „index“. Do datového okna vypište tyto hodnoty: -3, 2, 0,1,-2, -1, 3. Ke každé hodnotě přiřaďte textovou hodnotu -3...extra chladný, -2...velmi chladný, -1...mírně chladný, 0...normální, 1...mírně teplý, 2...velmi teplý, 3...extra teplý. V datovém okně si pomoci „štitku“ vyzkoušejte přepínání mezi textovou hodnotou a číselnou.
- Proměnná 3: nazvěte ji „t1“. Vepište hodnoty: 10,5; 16,5; 13,9; 14,4; 11,3; 12,1; 18,9;
- Proměnná 4: nazvěte ji „t2“. Vepište tyto hodnoty: 11,3; 17,9; 14,3; 15,2; 12,6; 11,6; 20,9;
- Proměnná 5: nazvěte ji „t_prumer“. Vypočítejte průměr hodnot proměnných t1 a t2 pomocí vzorce vloženého do „dlouhého jména“
- Soubor uložte jako „pokus1.sta“

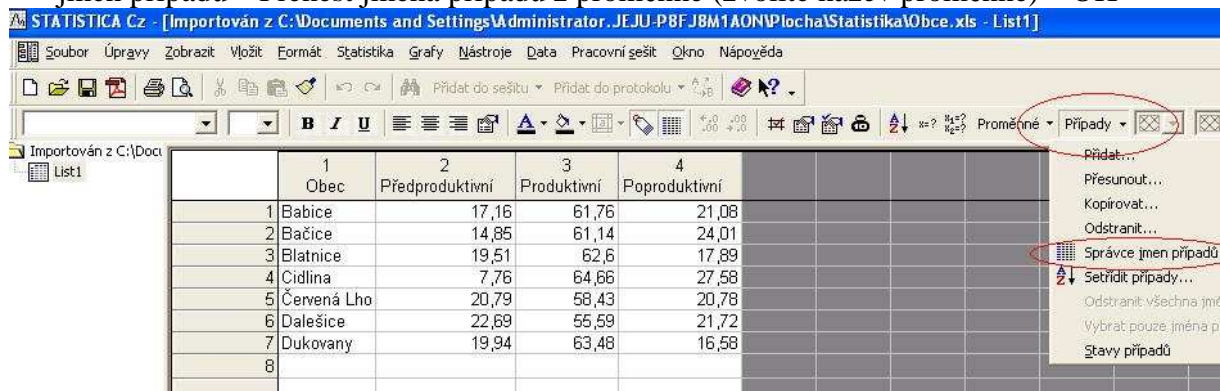
2.2 Import dat z EXCELU

- možnost 1: klasické přepokopírování dat pomocí schránky a funkce KOPÍROVAT a VLOŽIT z EXCELU do STATISTIKY
- možnost 2: SOUBOR→OTEVŘÍT→zvolit soubory s příponou xls. → můžete vybrat import všech listů nebo pouze vybraného listu→ zvolíte, chcete-li 1.řádek/sloupec převést na jména proměnných/případů (viz. obr. 3)



Obr. 3. Import dat z EXCELU

- Jména případů lze upravit přepokopírováním údajů z některé z proměnných. Na horní liště se nachází položka Případy→Správce jmen případů→Nastavit potřebnou délku jmen případů→Přenést jména případů z proměnné (zvolíte název proměnné)→OK



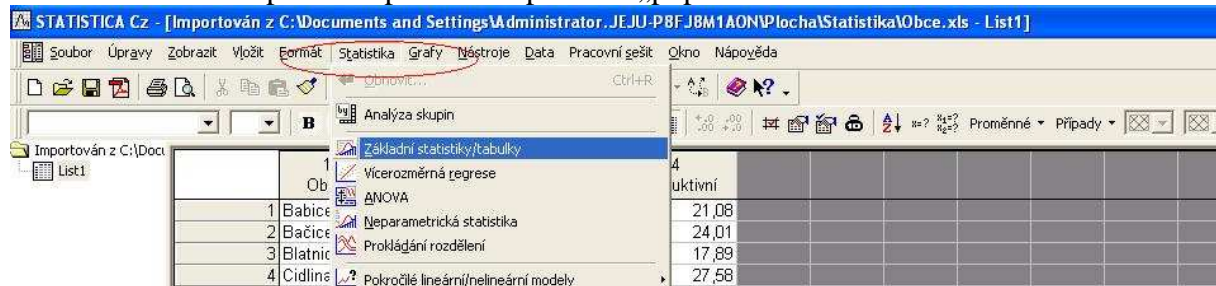
Obr. 4. Správce jmen případů

2.2.1 Úkol 2

- Importujte soubor Obce.xls do programu Statistica
- Převed'te 1. řádek do názvu proměnných
- Pomocí správce jmen případů převed'te proměnou 1 na jména případů

3 Výpočet statistických charakteristik

Pro výpočet statistických charakteristik se využívá položka „Statistika“ v horní liště programu. Zde se využívá funkce „Základní statistiky/tabulky“ → v dalším okně lze vybrat výpočet popisné statistiky, korelace, různých t-testů, kontingenčních tabulek, četnostních tabulek atd. Budete používat především položku „popisná statistika“.



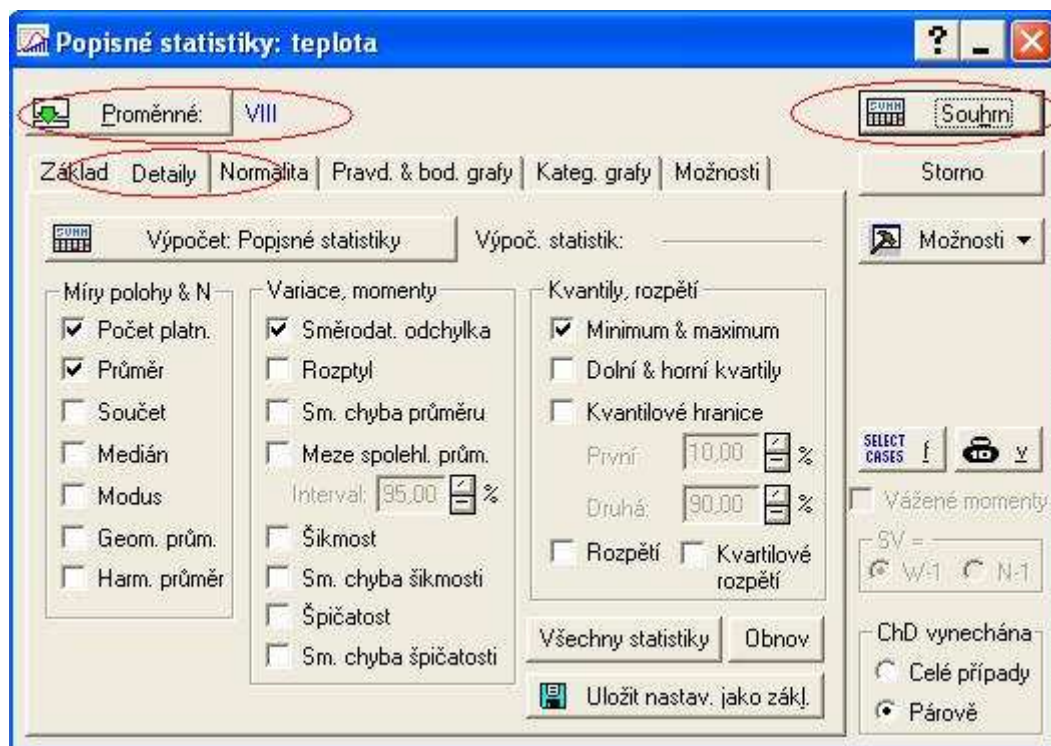
Obr. 5. Funkce Statistika

3.1 Popisná statistika



Obr. 6. Funkce základní statistiky a tabulky

- V okně popisná statistika zvolíte proměnnou/proměnné, pro kterou chcete vypočítat statistické charakteristiky
- Na kartě „Detaily“ zvolíte charakteristiky, které chcete vypočítat
- Pak zvolíte „souhrn“
- Viz. obr. 7
- Otevře se tzv. pracovní sešit (workbook, přípona .stw) a v něm výsledný výpočet (tabulka).
- Tabulku lze překopírovat do WORDU → označte hodnoty → pravé tlačítko myši → kopírovat se záhlavími



Obr. 7. Funkce Popisná statistika

3.1.1 Statistika bloku dat

- Funkce „statistika bloku dat“ slouží k umístění základních statistických charakteristik (např. průměr, směrodatná odchylka, medián, maximum, minimum atd.) přímo do datového okna.
- Kliknutím myši označte proměnné, pro které chcete výpočet provést → pravé tlačítko myši → Statistiky bloku dat → Blok sloupců → zvolte danou charakteristiku
- Analogicky se postupuje u případů

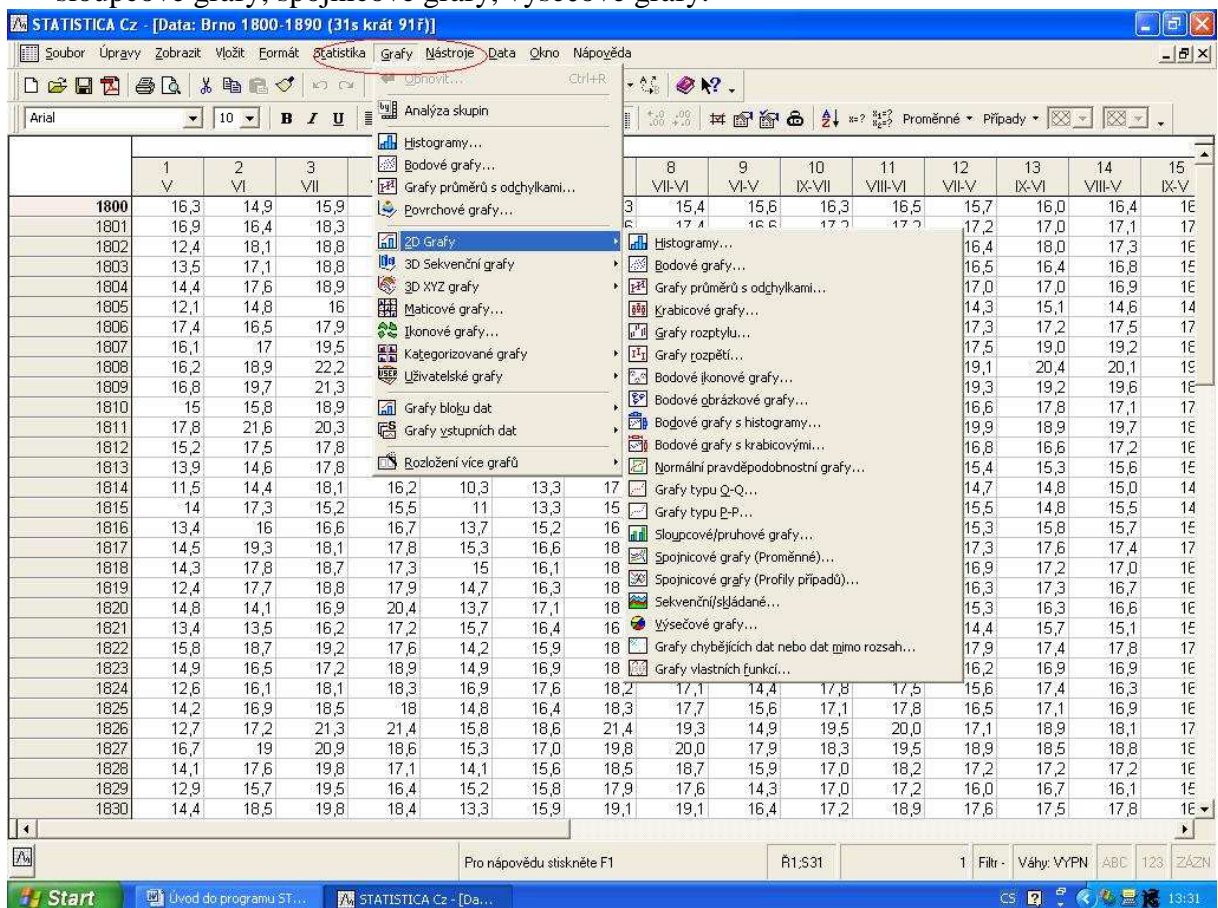
	1 I	2 II	3 III	4 IV	5 V	6 VI	7 VII	8 VIII	9 IX	10 X
1961	-3,6	2,1	6,4	11,6	11,8	18,1	17,3	18,1	16,9	10,4
1962	-1	-1,1	0							
1963	-7,9	-6,1	1							
1964	-6	-1,2	-0,3							
1965	-0,9	-3,6	1,9							
1966	-5	5	3,5							
1967	-2	1	5,4							
1968	-3,4	0,5	4,8							
1969	-3,1	-2,3	0,6							
1970	-3,4	-2,2	1,7							
1971	-3,4	1,1	1							
1972	-3,5	1,6	5,8							
1973	-1,5	0,8	4,6							
1974	0,8	3,3	6,8							
1975	2,2	-0,1	4,7							
1976	0,1	-1,3	1							
1977	-1,6	1,1	6,9							
1978	-0,6	-2,7	5							

Obr. 8. Statistika bloku dat
3.2 Úkol 3

- Do programu STATISTICA importujte list 1 ze souboru teplota.xls. Soubor obsahuje průměrné měsíční teploty z vybrané stanice za období 1961-2000.
- 1. řádek převedte na jména proměnných.
- 1. proměnou převedte na jména případů a upravte na adekvátní délku.
- Přidejte jednu proměnnou, kterou nazvete „rok“ a pomocí vhodné funkce vložené do „dlouhého jména“ vypočtete průměrnou roční teplotu.
- Vypočtete pro všechny měsíce průměr, směrodatnou odchylku, minimum a maximum. Výslednou tabulku z pracovního sešitu přepokopírujte do WORDU.
- Do datového okna vložte průměr za všechny měsíce.
- Do datového okna vypočtete maximum za každý rok bez použití vzorce vloženého do „dlouhého jména“.

4 Grafy

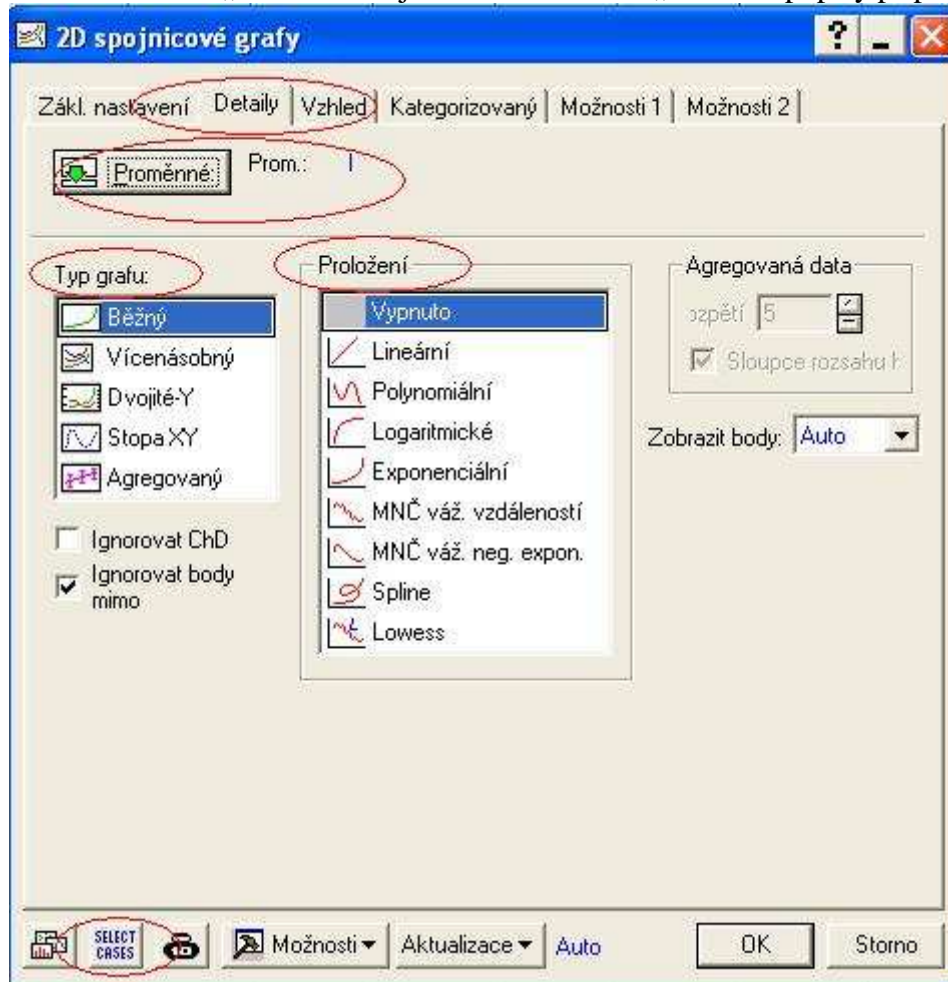
- Program STATISTICA umožňuje vytvářet různé druhy grafů a tyto následně dále upravovat.
- K vytváření grafů slouží položka „Grafy“ v horní liště.
- Nejčastěji se využívá funkce 2D Grafy – histogramy, bodové grafy, krabicové grafy, sloupcové grafy, spojnicové grafy, výsečové grafy.



Obr. 9. Funkce grafy

4.1 Spojnicový graf

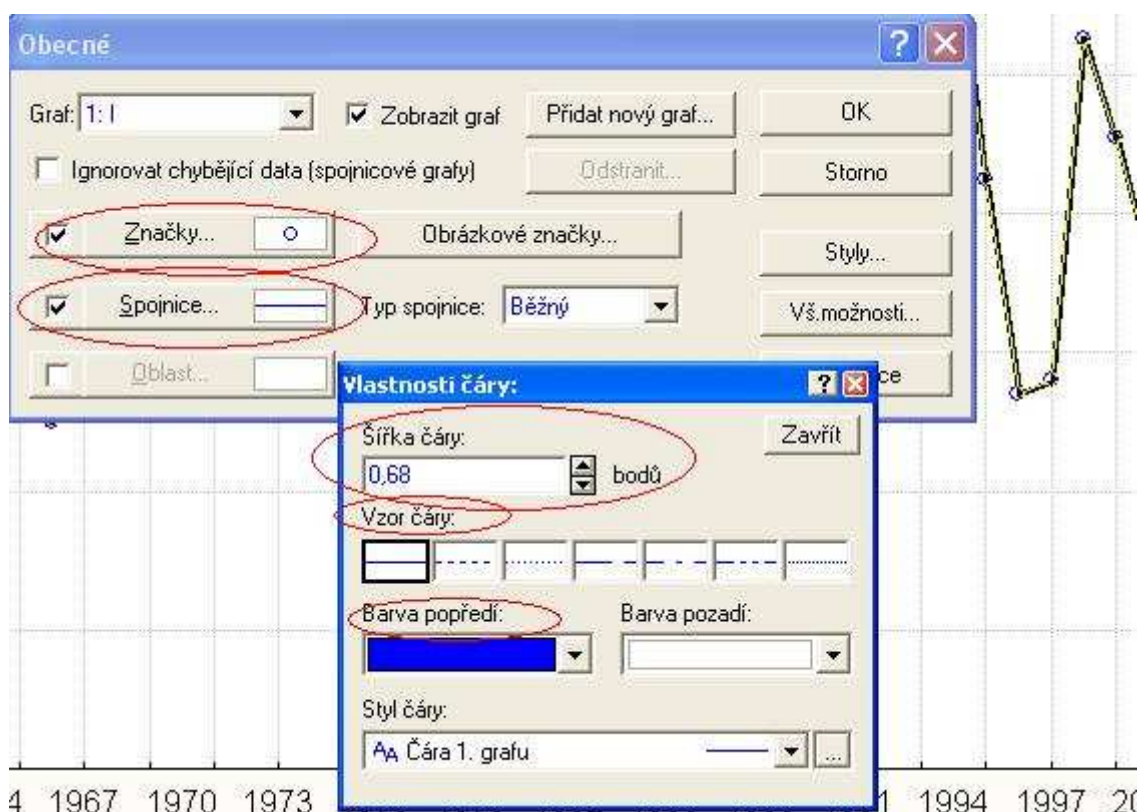
- V otevřeném okně grafu zvolíte prvně proměnnou, pro kterou chcete vytvořit graf.
- Na kartě „Detaily“ lze navolit typ grafu (běžný, vícenásobný), proložení (vypnuté, lineární atd.).
- Funkce „Select cases“ slouží k zapnutí filtru, který umožňuje vybrat případy, pro které chcete aby se graf vytvořil. Pokud tento filtr nezapnete, graf se vytvoří pro všechny případy (nejčastěji používané).
- Na kartě „možnosti 1“ je dobré zaškrtnout „Zobrazit popisy případů“.



Obr. 10. Vytvoření spojnicového grafu

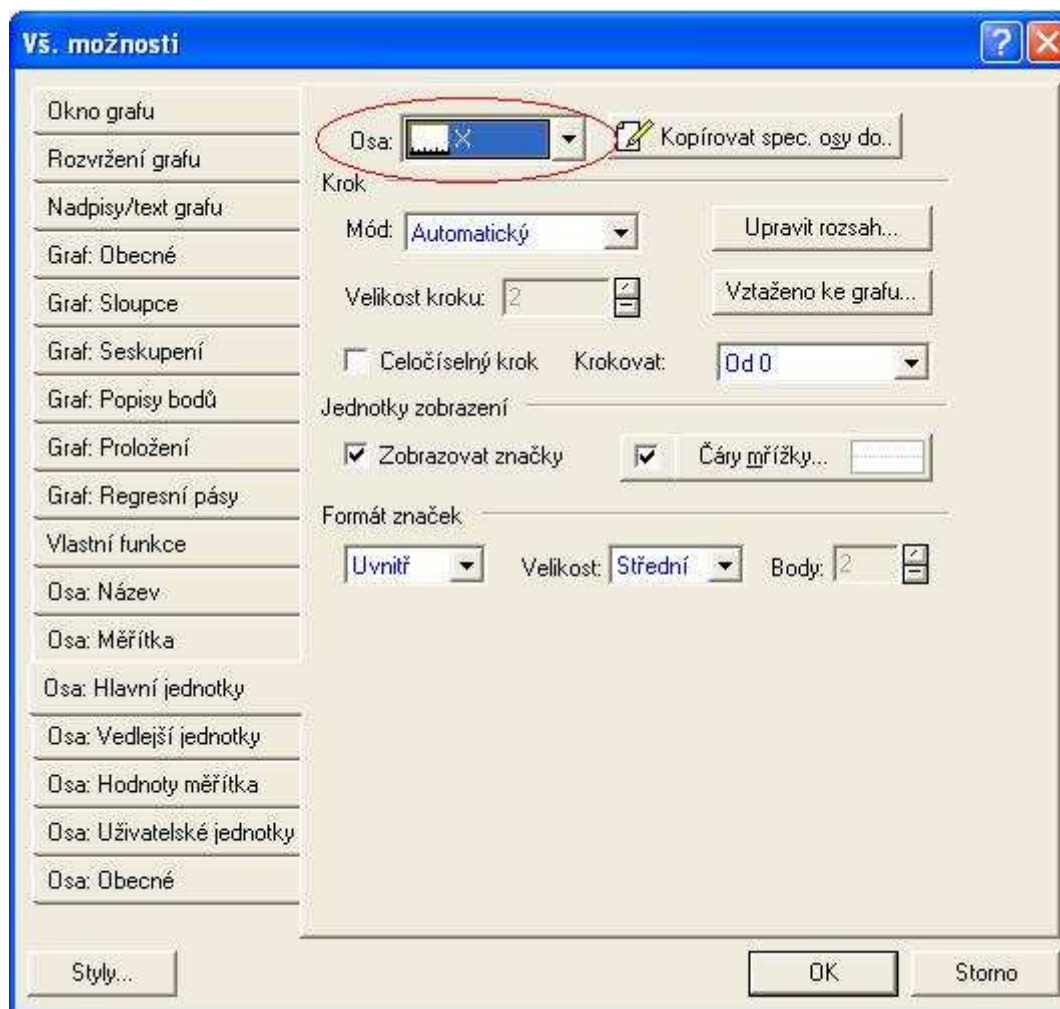
4.1.1 Editace grafu

- Poklepáním na spojnici bodů v grafu se otevře okno, ve kterém lze editovat značky pro body (lze vypnout nebo změnit za jiný znak) a upravit vlastnosti spojnice (šířku, barvu, vzor). Pomocí tohoto okna je možné i vložit nový graf („přidat nový graf“) – viz. obr. 11



Obr. 11. Úprava vlastností spojnice grafu

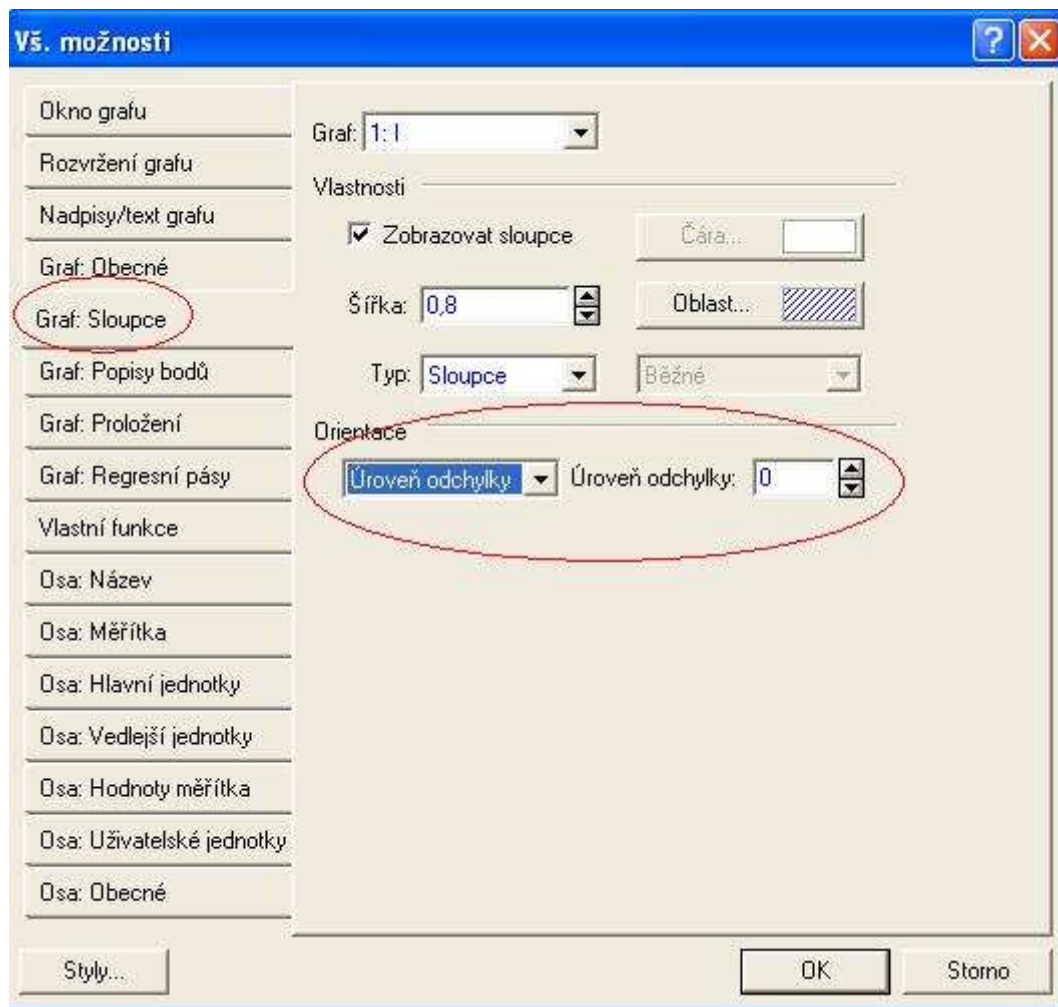
- Hlavní editační okno se otevře poklepáním na pole grafu.
- V levé části okna jsou karty, ve kterých lze přepínat a upravovat vlastnosti grafu. V následující části budou vysvětleny nejčastěji používané karty.
- Nadpisy/text grafu: mění se název grafu, velikost i typ písma,
- Graf: Obecné: stejná funkce, jako při poklepání na spojnici grafu-mění se druh bodů, velikost a barva spojnice.
- Graf: Proložení: tato karta umožňuje vložení proložení a zvolení jejího typu (lineární, polynomické atd.)
- Osa: Název: vypíšete název osy. Tu si volíte v horní části okna (X,Y). viz. obr 12
- Osa: Měřítko: v horní části opět měníte osu, pro kterou chcete změnit nastavení. Poté je tam důležitá funkce- „mód“: změníte nastavení z automatického na manuální a můžete změnit rozsah osy (např.: osa x je delší než vlastní spojnice, proto upravíte maximální délku na počet případů, které chcete aby bylo na ose vyneseno).
- Osa: Hodnoty měřítka: nastavení formátu hodnot na ose a také upravuje množství popisu osy (například každý 5 popis).
- Osa: Uživatelské jednotky: zde můžete měnit popis jednotek osy (například potřebujete změnit rok 1960 na 60+). Do sloupce Pozice vepíšete číslo pozice údaje, který chcete zaměnit, na ose a do sloupce Text napíšete váš nový údaj.



Obr. 12. Okno pro úpravu grafu.

4.2. Sloupcové grafy

- Grafy→2D Grafy→Sloupcové/pruhové grafy
- Postup je velmi podobný jako u spojnicového grafu
- Výsledný sloupcový graf v programu Statistica je odlišný od EXCELU. Pokud jsou v datech obsaženy záporné i kladné hodnoty, tak se osa x neprotíná s osou y v 0, ale v nejnižší hodnotě (například -10). Proto je nutná úprava→otevřete editační okno→Graf: Sloupec→Orientace→Úroveň odchylky: 0. viz. obr. 13
- Na kartě Graf: Sloupec: lze také upravovat šířku a barvu sloupců.
- Ostatní editační funkce jsou stejné jako u spojnicového grafu (společné pro všechny grafy)



Obr. 13. Nastavení úrovně odchylky při sloupcovém grafu

4.3 Kopírování grafu

- Právě tlačítko na myši → kopírovat graf → vložit například do WORDU
- Právě tlačítko myši → uložit graf → uložit jako typ: např. JPEG

4.4 Úkol 4

- Importujte do programu Statistica soubor teplota.xls
- Vytvořte spojnicový graf pro měsíc leden za období 1961-2000.
- Popište osy grafu: osa x: rok, osa y: teplota [°C].
- Upravte popis osy x tak, aby byl zobrazen každý 5 rok.
- Z grafu odstraňte značky bodů
- Spojnici zvětšete na tloušťku 1,5 a změňte barvu na červenou.
- Do grafu přidejte lineární trend-tloušťka 1,1, barva zelená, čárkovaně
- Graf překopírujte do WORDU a dokument uložte jako pokus2.doc
- Pro každý měsíc za období 1961-2000 vypočtete průměr.
- Pro měsíc leden vypočtete difference teploty od průměru za období 1961-2000
- Pro difference měsíce ledna vytvořte sloupcový graf.
- Popište osy x,y: osa x: rok, osa y: difference teploty [°C]
- Sloupcový graf upravte tak, aby osa x a osa y se protínaly v 0.
- Upravte šířku sloupce na hodnotu 0,5 a změňte barvu sloupců na žlutou.
- Graf uložte jako JPEG obrázek, vložte do WORDU a uložte jako pokus3.doc