

Cvičení 0 – základy práce se systémem STATISTICA

Systém má modulární stavbu. V multilicenci pro Masarykovu univerzitu jsou k dispozici moduly: Základní statistiky/tabulky, Vícerozměrná regrese, ANOVA, Neparametrická statistika, Prokládání rozdělení, Rozdělení & simulace, Pokročilé lineární/nelineární modely, Vícerozměrné průzkumné techniky, Průmyslová statistika & Six Sigma, Analýza síla testu, Automatizované neuronové sítě.

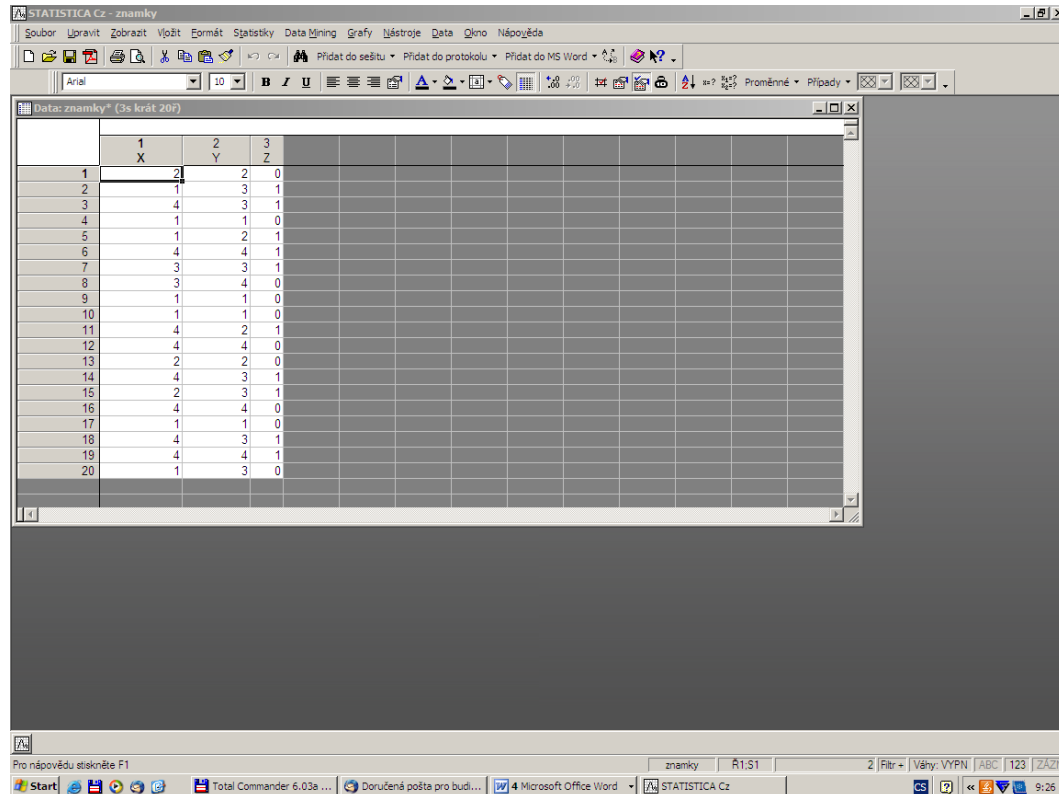
Velké množství informací o systému STATISTICA lze najít na webové stránce společnosti StatSoft, která je jejím distributorem v České republice (www.statsoft.cz). Z této stránky vede rovněž odkaz na elektronickou učebnici statistiky.

Instalace software STATISTICA je dostupná na adrese <https://inet.muni.cz/app/soft/licence>

STATISTICA má několik typů oken:

- **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).

Ukázka datového okna:



- **Pracovní sešit (workbook)** (má příponu stw). Do pracovního sešitu se 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.

Ukázka pracovního sešitu:

STATISTICA Cz - [PS 1* - Tabulka četností: známka z matematiky (znamky)]

Soubor Upravit Zobrazit Vložit Formát Statistky Data Účinný Grafy Nástroje Data Pracovní sešit Okno Nápověda

PS 1*
 Základní statistiky a tabulky (znamky)
 Dialog frekvenčních tabulek
 Tabulka četností: známka z matematiky (zna...
 Histogram: X: známka z matematiky

Kategorie	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
výborně	7	7	35.00000	35.00000
velmi dobře	3	10	15.00000	50.00000
dobře	2	12	10.00000	60.00000
neprospěl	8	20	40.00000	100.00000
ChD	0	20	0.00000	100.00000

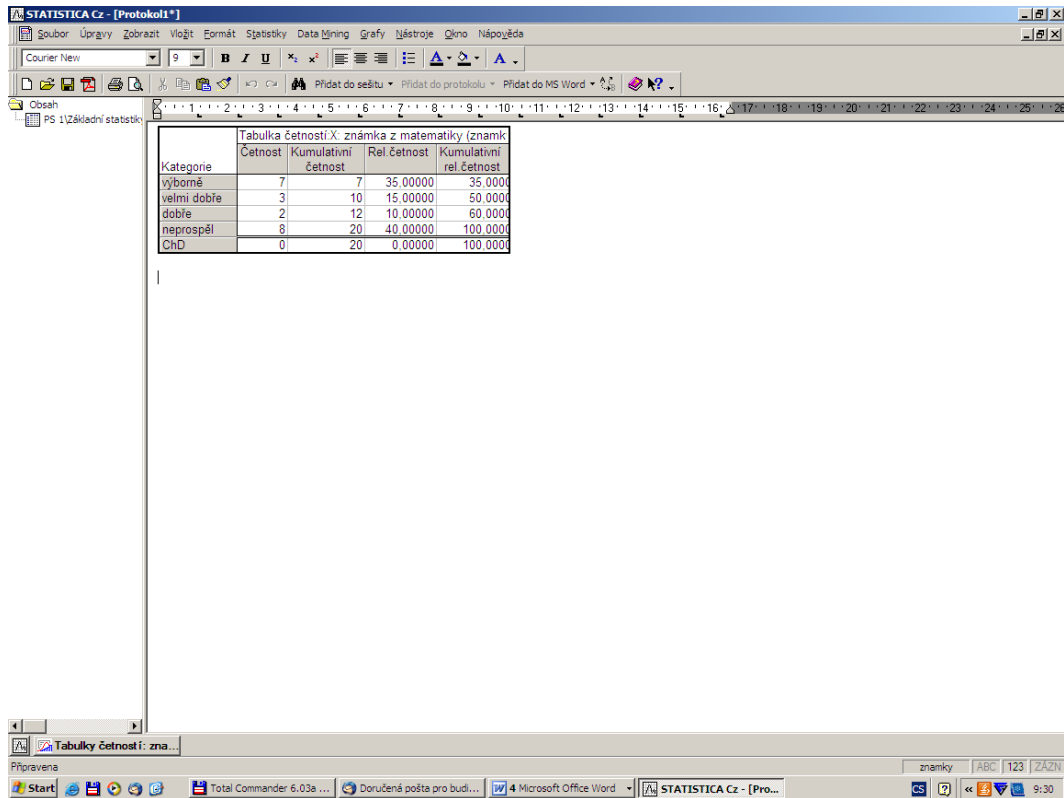
Tabulka četností: X: známka z matematiky (znamky) Histogram: X: známka z matematiky

znamenky Ř: 1: 51 7 Filt: Váhy: VYPN ABC 123 IZAZN

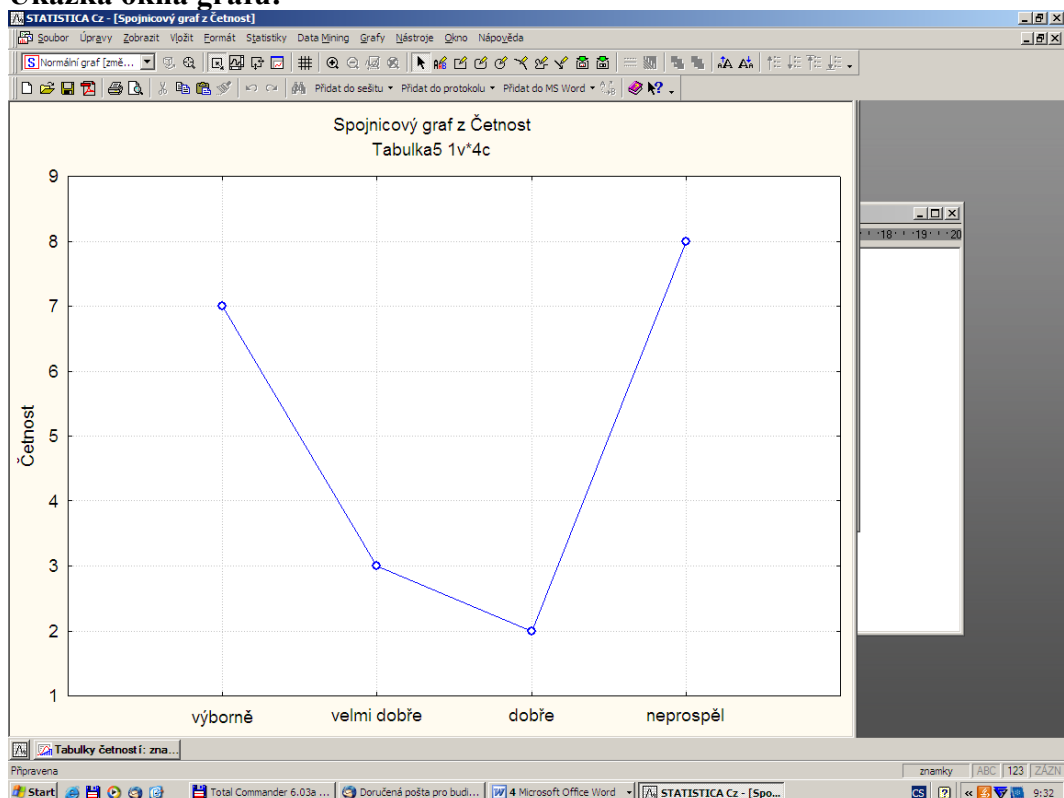
Start Total Commander 6.03a ... Doručená pošta pro bud... 4 Microsoft Office Word STATISTICA Cz - [PS ... 9:28

- **Protokol (report)** (má příponu str, lze ho uložit i ve formátu rtf, txt, htm či pdf). Pokud požadujeme, aby se výstupy ukládaly nejen do Pracovního sešitu, ale i do Protokolu či Wordu, postupujeme takto: Soubor – Správce výstupů – vybereme Výstup protokolu či Výstup MS Word – OK. Protokol se podobně jako Pracovní sešit skládá ze dvou oken. Do Protokolu můžeme vkládat vlastní text, vysvětlující komentáře, poznámky apod. Tabulky a grafy lze v Protokolu i Pracovním sešitu dále upravovat.

Ukázka protokolu:

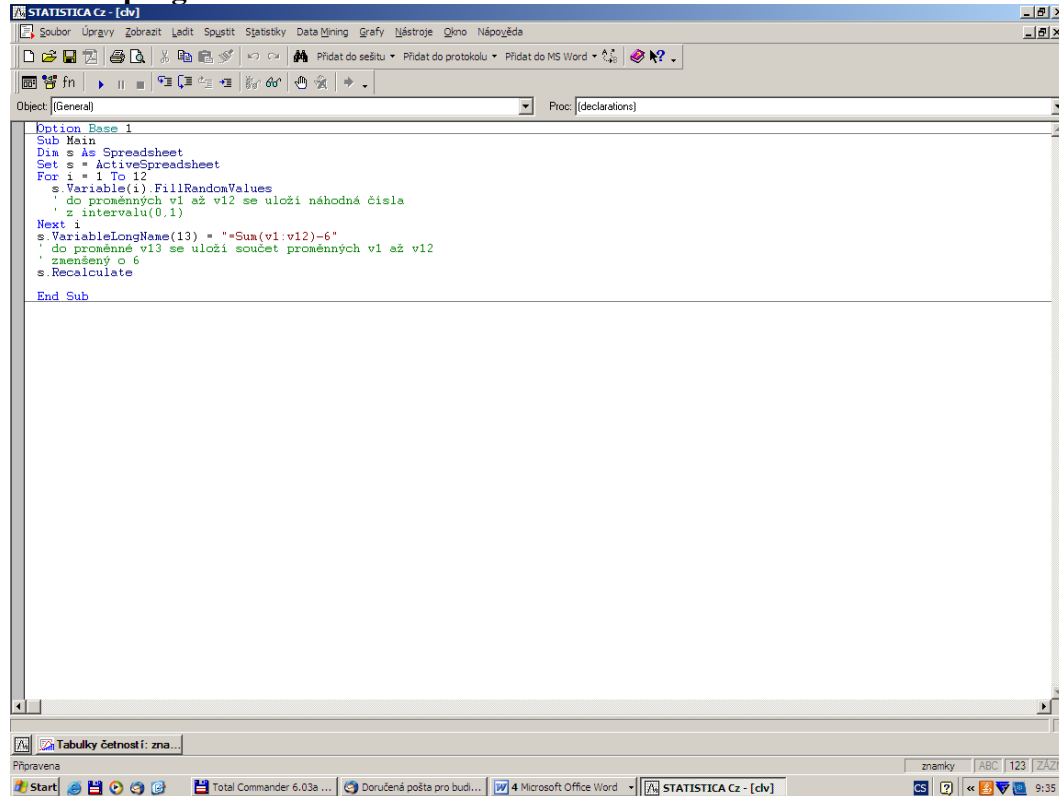


- **Okno grafů** (přípona stg, lze ho uložit i jako bmp, jpg, png, tif a wmf či pdf). Získá se tak, že v Pracovním sešitu klikneme pravým tlačítkem na graf a vybereme Klonovat graf.
Ukázka okna grafů:



- **Programovací okno** (přípona svb). Slouží pro zápis programů v jazyku STATISTICA Visual Basic. Vyvolá se z menu Nástroje – Makro.

Ukázka programovacího okna:



Mezi jednotlivými typy oken se přepínáme pomocí položky Okno v hlavním menu.

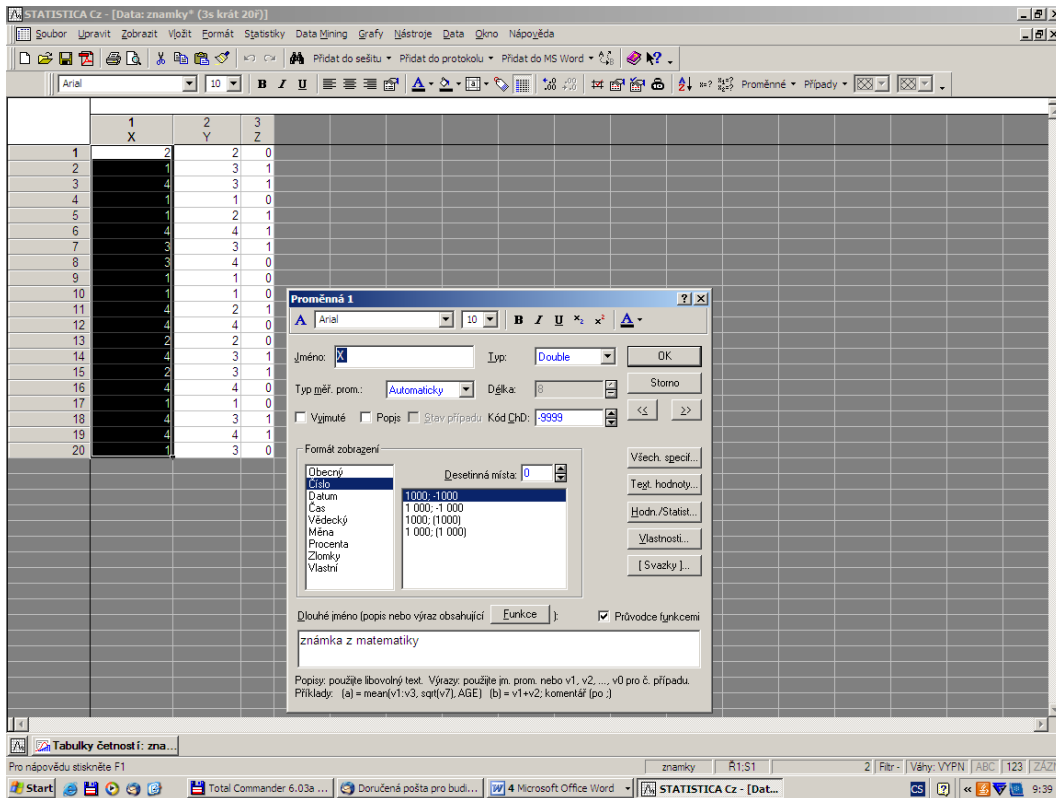
Od verze 8.0 je možno používat Projekty. Tento prvek umožňuje uložit soubor STATISTICA Projekt, a "zmrazit" tak doposud provedenou práci včetně umístění všech oken s objekty na obrazovce. Po otevření tohoto souboru se může přímo navázat na již provedenou část analýzy. Lze také uložit zaznamenané makro analýzy.

Vytvoření datového okna

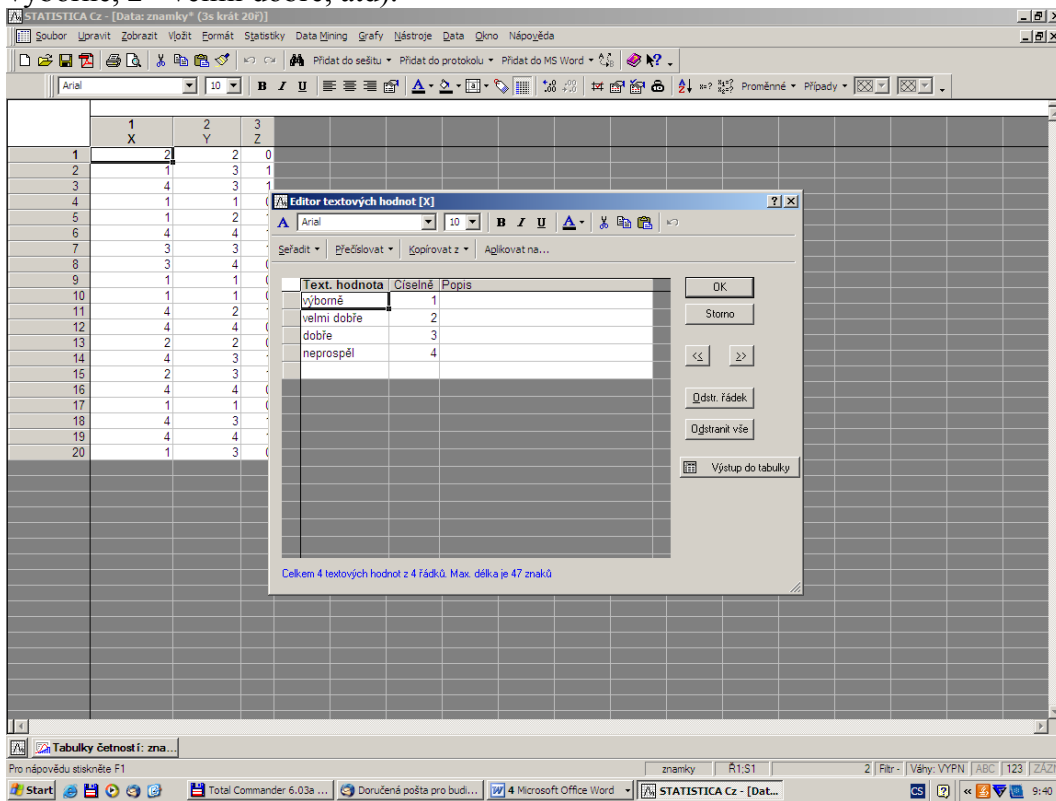
Po spuštění systému STATISTICA zavřeme všechna okna, která se objeví.

Vytvoření nového souboru: Soubor – Nový – na liště Tabulka dat zvolíme potřebný počet proměnných (sloupce) a případů (řádky). Zde lze také zvolit formát dat, délku jmen případů atd. (tyto vlastnosti můžeme upravit i později).

Nastavení vlastností proměnné: 2x klikneme na název proměnné. Objeví se nové okno. Zde je možnost změny jména proměnné, formátu dat, nastavení kódu pro chybějící data apod. V položce „Dlouhé jméno“ lze nastavit dlouhý popis (tzv. návěští proměnné, např. X – známka z matematiky) a informace vztahující se k proměnné, ale spíše se „Dlouhé jméno“ používá pro vytvoření vzorců a transformaci proměnné (například: $=v3/v2*100$). Některé funkce jsou předem vloženy – položka „Funkce“.

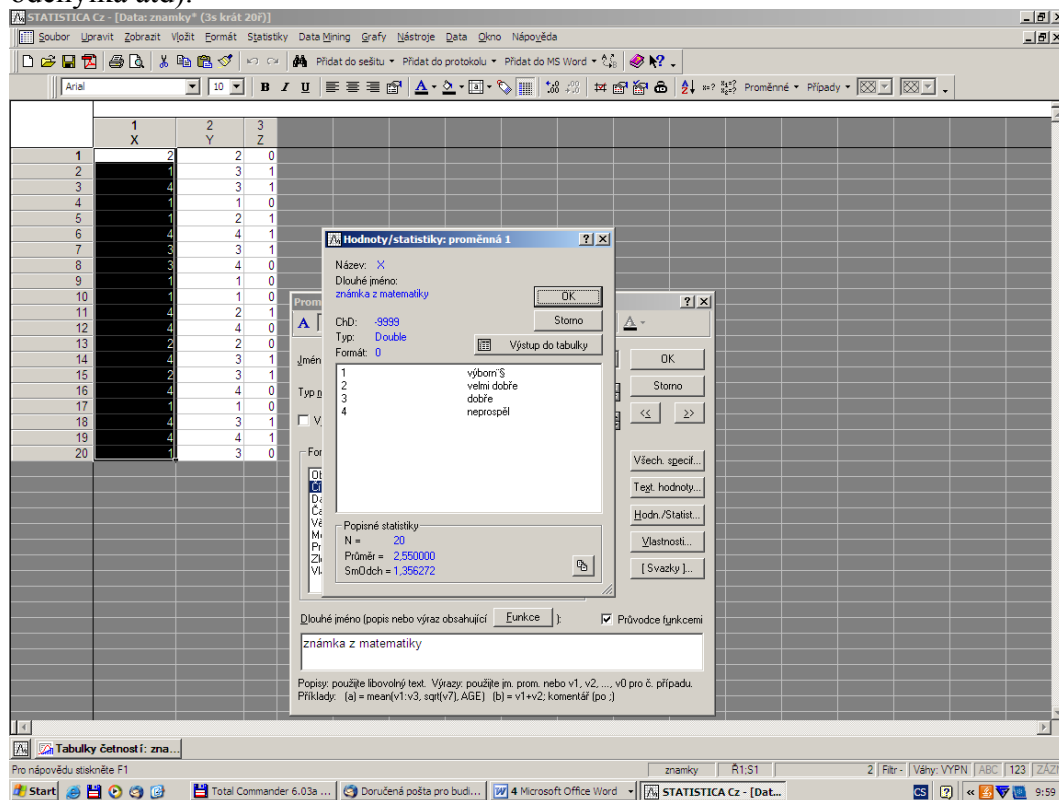


V položce „Textové hodnoty“ lze nastavit k variantám proměnné textový popis (např. 1 – výborně, 2 - velmi dobře, atd).



Poté můžeme přímo v datovém okně (ikona štítku v horní liště) měnit způsob zobrazení hodnot proměnné (buď číselný nebo slovní).

Položka Hodn/Statist udává celkové informace o proměnné (počet údajů, průměr, směrodatná odchylka atd).



Úkol 1.: Do datového okna programu STATISTICA načteme datový soubor lide.sta, který obsahuje následující údaje o 32 náhodně vybraných osobách:

Jmeno (křestní jméno osoby)

Sex (1 muž, 2 žena)

Vek (věk osoby v dosažených letech)

Vyska (výška osoby v cm)

Hmotnost (hmotnost osoby v kg)

Návod: Soubor – Otevřít – Lide.sta – Otevřít

Úkol 2.: Proměnným Jmeno, Sex, Vek, Vyska, Hmotnost vytvoříme návěští (Jmeno – křestní jméno osoby, Sex – pohlaví osoby, Věk – věk v letech, Výška – výška v cm, Hmotnost – hmotnost v kg). Popíšeme, co znamenají jednotlivé varianty u proměnné Sex (1 - muž, 2 – žena).

Návod: Kurzor nastavíme na proměnnou Jmeno – 2x klikneme myší – Dlouhé jméno křestní jméno osoby - OK. Kurzor nastavíme na Sex – 2x klikneme myší – Dlouhé jméno pohlaví osoby, Text. hodnoty – 1 muž, 2 žena (Přepínání mezi číselnými hodnotami a jejich textovým popisem se děje pomocí tlačítka s ikonou štítku.) Podobně vytvoříme návěští pro další proměnné.

Úkol 3.: Pomocí správce jmen případů převed'te proměnnou Jmeno na jména případů. Proměnnou Jmeno poté zrušte.

Návod: Data – Správce jmen případů – Přenést jména případů z proměnné Jmeno – OK – OK. Proměnnou Jmeno nyní zrušíme takto: Kurzorem se nastavme na proměnnou Jmeno – Proměnné – Odstranit – OK.

Úkol 4.: Pro každou osobu vypočtete BMI. (Body Mass Index se počítá podle vzorce

$$\text{BMI} = \frac{\text{hmotnost} \left[\begin{smallmatrix} \text{kg} \\ \text{m}^2 \end{smallmatrix} \right]}{\text{výška}^2 \left[\begin{smallmatrix} \text{m}^2 \end{smallmatrix} \right]}$$
 Osoby, které mají BMI pod 18,5, trpí podvýživou, BMI mezi 18,5 a

25 ukazuje na normální stav, hodnoty mezi 25 a 30 svědčí o nadváze a hodnoty nad 30 pak o obezitě.)

Návod: Za proměnnou Hmotnost vložíme novou proměnnou: Proměnné – Přidat – Za Hmotnost – Jméno BMI. Do Dlouhého jména napíšeme vzorec =10000*Hmotnost/Vyska^2

OK

Úkol 5.: Pro proměnné Vek, Vyska, Hmotnost, BMI zjistěte minimum, maximum a průměr.

Návod: Tento úkol lze splnit více způsoby, zatím si ukážeme postup pomocí funkce Statistika bloku dat, která slouží k umístování základních číselných charakteristik proměnných (např. průměr, směrodatná odchylka, medián, maximum, minimum atd.) přímo do datového okna. Kliknutím myši označíme proměnné, pro které chceme výpočet provést – stiskneme pravé tlačítko - Statistiky bloku dat - Blok sloupců - zvolíme danou charakteristiku.

Úkol 6.: Pro proměnnou BMI vytvořte sloupcový graf. Pod každý sloupec umístěte jméno příslušné osoby.

Návod: Grafy – 2D Grafy – Sloupcové/pruhové grafy – Proměnné BMI – OK – OK

Vytvoří se sloupcový graf, který upravíme tak, aby každý sloupec byl popsán jménem příslušné osoby.

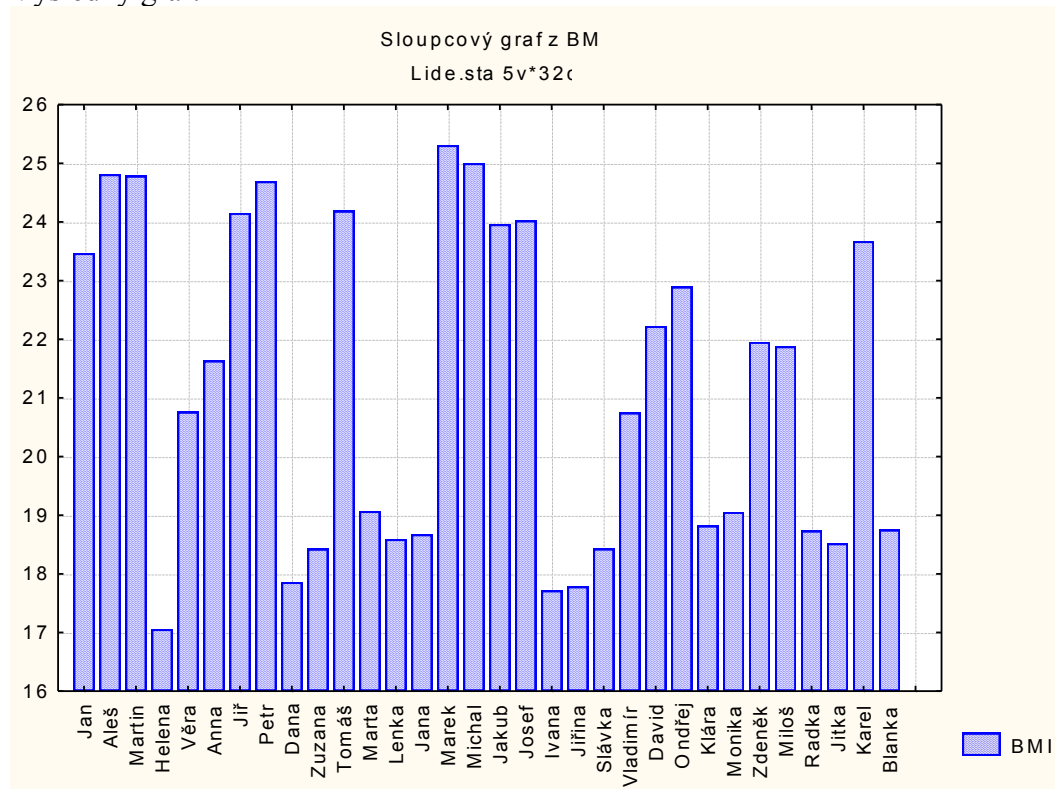
2x klikneme na některé jméno v popisu vodorovné osy. Otevře se okno „Vzhled osy“.

V Přesakovat hodnoty zaškrtneme Vypnuto a v Rozložení vybereme Kolmo na osu – OK.

Dále je vhodné na svislé ose znázorňovat hodnoty pouze od 16 do 26.

2x klikneme na některé číslo v popisu svislé osy. V okně „Vzhled osy“ vybereme Měřítko – Mód Ručně, Minimum 16, Maximum 26 – Upravit krok – Mód Ručně – Velikost kroku 1 – OK.

Výsledný graf:



Úkol 7.: Nyní vytvořte sloupcové grafy pro BMI zvlášť pro muže a zvlášť pro ženy.

Návod: Je zapotřebí nejprve vybrat pouze muže.

Grafy – 2D Grafy – Sloupcové/pruhové grafy – Filtr případů – zaškrtneme Zapnout filtr, některé, vybrané pomocí výrazu Sex = 1, OK.

Dále postupujeme analogicky jako v úkolu 6. Stejným způsobem vytvoříme graf pro ženy, kde ve filtru zadáme podmínku Sex = 2.