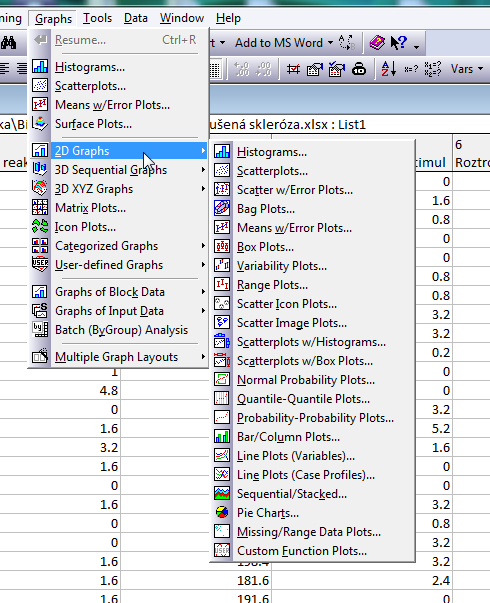
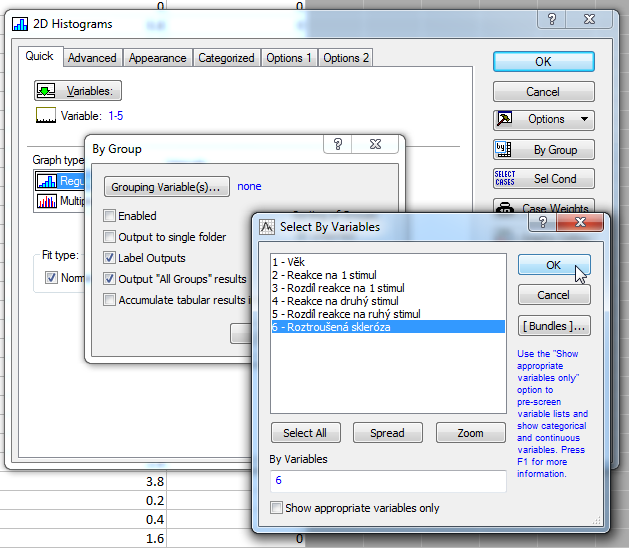
**Cvičení 5**

**1) Načtení datového souboru pro roztroušenou sklerózu z excelu   
(**Roztroušená skleróza.xlsx ve složce cviceni\_05**)**

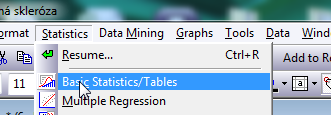
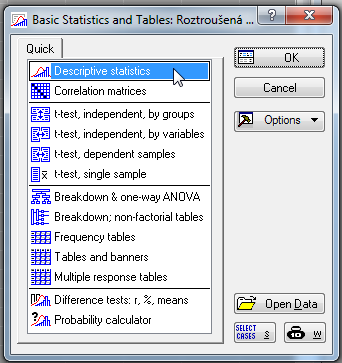
* Jedná se o data naměřená u pacientů s roztroušenou sklerózou a u kontrolní skupiny. Byly měřeny doby reakce (měřeno v milisekundách -ms) na dva odlišné podněty (stimuly). Popis jednotlivých proměnných:
  + 1 - Věk
  + 2 - Doba reakce obou očí na první podnět
  + 3 - Rozdíl reakce pravého a levého oka na první podnět
  + 4 - Doba reakce na druhý podnět
  + 5 - Rozdíl reakce pravého a levého oka na druhý podnět
  + 6 - Určení, zda se jedná o pacienta s roztroušenou sklerózou či kontrolu   
    (0-kontrola, 1- roztroušená skleróza)

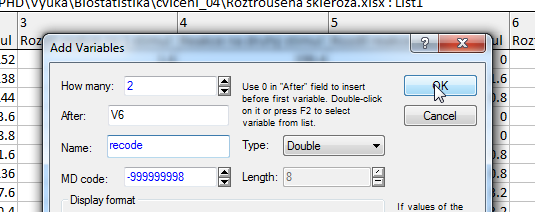
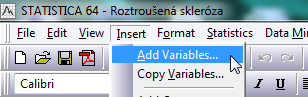
**2) Znázorněte si naměřená data. Pro každou proměnnou (*pro kterou to má smysl*) rozhodněte, zda by mohla pocházet z normálního rozdělení. Určete totéž pro tyto proměnné rozdělené dle pacientů s roztroušenou sklerózou a kontroly.**

* Nejlépe si pro posouzení vykreslit histogram a boxploty. Vzpoměňte si, jak lze dle boxplotu odhadnout, zda se jedná o normální/symetrické rozložení.
* U histogramu je třeba rozdělit soubor dle skupinové proměnné, viz obrázek.

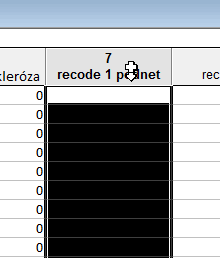
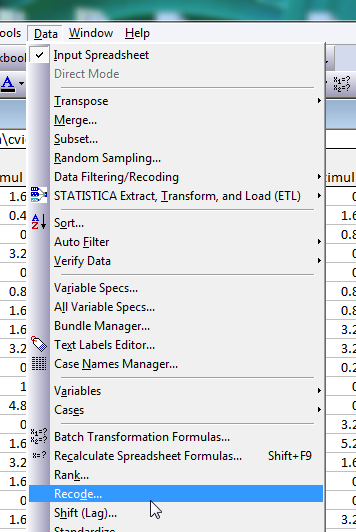
****

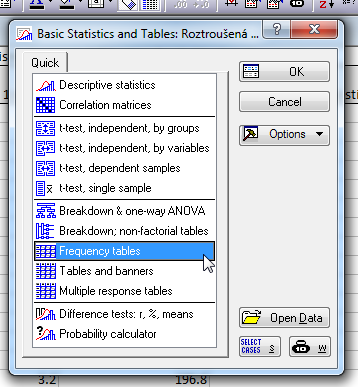
**3) Vypište si základní popisné statistiky pro jednolivé proměnné opět rozdělených dle pacientů s roztroušenou sklerózou a kontrol.**

* Pokud by byly proměnné z normálního rozdělení : průměr, směrodatná odchylka
* Pokud by nebyly z normálního rozdělení : medián + 9-95% percentil, popřípadě min-max
* Nevíme, zda jsou data z normálního rozdělení - vypíšeme si všechny tyto popisné statistiky + šikmost a špičatost  
  ****

**4) Vytvořte 2 nové kategoriální proměnné tak, že překódujete proměnné pro dobu reakce na méně jak 150ms a 150ms a více u prvního podnětu a 200ms a 200ms a více u druhého podnětu.**

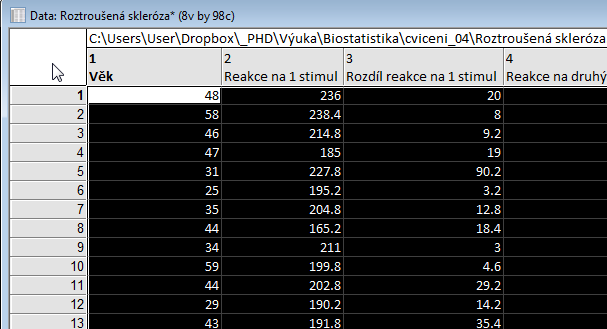
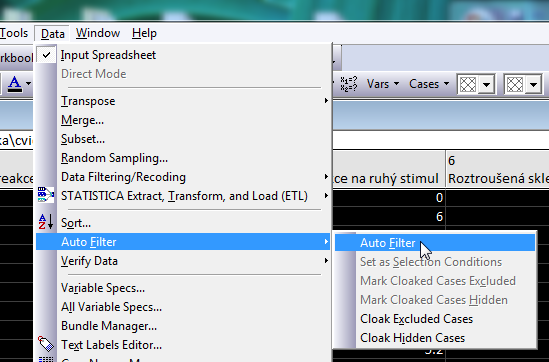
* Před překódováním nezapomeňte označit proměnnou, kam budou vloženy nové hodnoty!!

****

**5) Vytvořte frekvenční tabulky pro nově vytvořené proměnné. Zjistěte jaké procento pacientů a kontrol má odezvu více jak 200ms pro jednotlivé typy podnětů.  
**

**6) Kolik je mezi všemi pacienty kontrol, které mají Reakci na druhý stimul vyšší jak 225. Využijte k tomu *autofilter.***

* Nejprve označte celou tabulku kliknutím myší do levého horního rohu...

****

**7) Jak bychom mohli vizualizovat závislost délky reakce na první podnět a druhý podnět? Mají pacienti s delší dobou odezvy na první podnět také delší dobu odezvy při reakci na druhý podnět?**

* Vynášíme proti sobě dvě spojité proměnné...