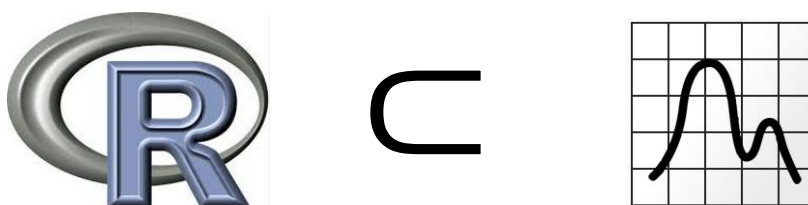


StatSoft

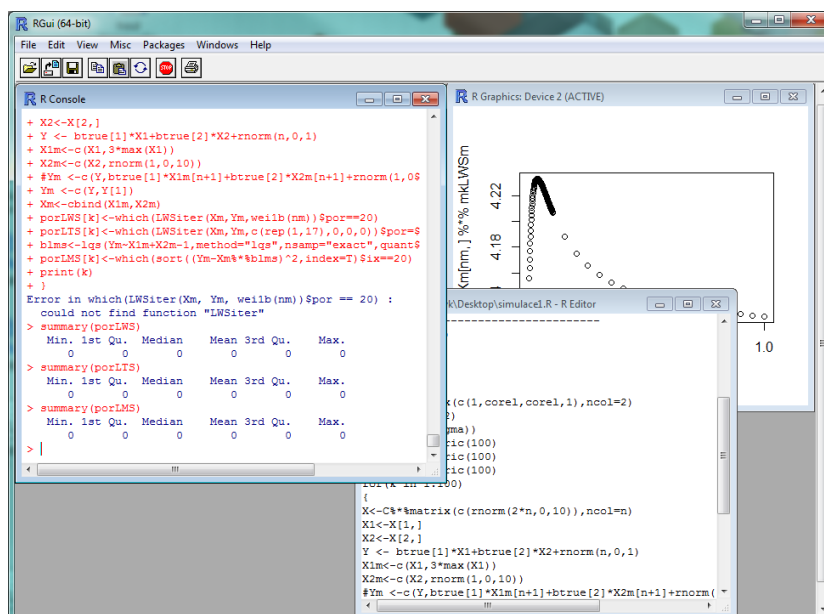
Propojení programů *STATISTICA* a R

Věděli jste, že je možné volat program R přímo ze *STATISTICY*? Nyní je to mnohem jednodušší s vlastním novým konektorem do R.



Něco málo o programu R

Program R (Rko) je programovací jazyk a prostředí pro statistické výpočty (www.r-project.org). Jedná se o open-source, který poskytuje širokou škálu nejrůznějších početních operací, statistických metod a grafických funkcionalit. To, co zrovna vaše instalace Rka umí, závisí na nainstalovaných balíčcích. Základní instalace má okolo základních 8 balíčků, nicméně dostupné funkcionality lze jednoduše doplnit instalací nejrůznějších dalších balíčků, kterých je více než 500, přičemž nové stále vznikají. Jedním z nich je i knihovna *COMadaptR*, která zprostředkovává právě propojení Rka a *STATISTICY*.

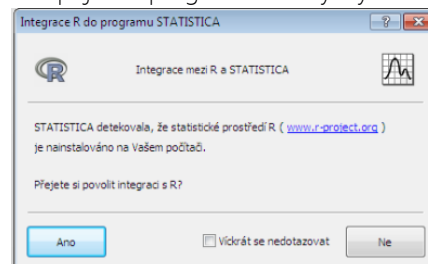


Proč se hodí toto spojení?

Jednoduše proto, že toto spojení Vám rozšíří možnosti Vaší statistické práce. S tímto spojením můžeme začít využívat unikátní schopnosti obou programů zároveň. Je jasné, že pro některé věci je lepší jeden program a pro jiné ten druhý. K Rku sáhnete především, pokud potřebujete něco velmi specializovaného, co ve *STATISTICE* není implementováno, případně můžete využít programovacího jazyka a naprogramovat si něco speciálního sami. *STATISTICU* využijete například pro přehlednější, flexibilnější a rychlejší pohled na data nebo výpočty nad velkými datovými soubory.

Jak zprovoznit propojení mezi programy

STATISTICA se dokáže propojit s mnoha programy, například s Microsoft Office prostřednictvím kanálu COM. Podobně lze ve *STATISTICE* spouštět i skripty (makra) programu R přes zvláštní aplikaci. Pro spojení s programem R byl vytvořen speciální konektor nazvaný COMadaptR. Ten je automaticky instalován se *STATISTICOU* od verze 12 SP3. Update české verze 12 na 12 SP3 naleznete na [našich internetových stránkách](#). Update anglické verze pak na [amerických stránkách](#). Pokud systém zjistí, že je v počítači přítomna instalace Rka (instalaci Rka je potřeba mít 2.13.0 nebo vyšší), pak *STATISTICA* nabídne možnost automaticky instalovat knihovnu COMadaptR do Rka, která je potřebná pro zprovoznění propojení. Propojení je tedy velmi jednoduché, nainstalujete si R i *STATISTICU* a *STATISTICA* Vám poté automaticky nabídne možnost propojení, kterou stačí odkliknout – viz obrázek vpravo.



Funkčnost spojení otestujete spuštěním nějakého makra pro Rko (můžete použít například příklady v *Soubor-Otevřít příklady-složka R*).

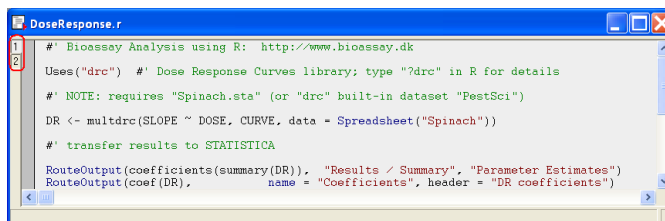
Samotné spojení

Co tedy toto propojení umožňuje?

1. Je možné spouštět skripty Rka přímo ze *STATISTICY*.
2. Oboustranný přenos dat a prezentace výstupů mezi programy (provedeno skrz makro spuštěné ve *STATISTICE*).
 - a. To zahrnuje zaslání výstupů z Rka do protokolů, sešitů a grafů *STATISTICA*.
 - b. Zpracovávání datových souborů ze *STATISTICY*Rkem.
 - c. Přenášení tabulkových výsledků z R do *STATISTICA* tabulek.
 - d. Volání Rka prostřednictvím *STATISTICA Visual Basic*
3. Využití Rka ve workspace, *STATISTICA Server* i *STATISTICA Enterprise*

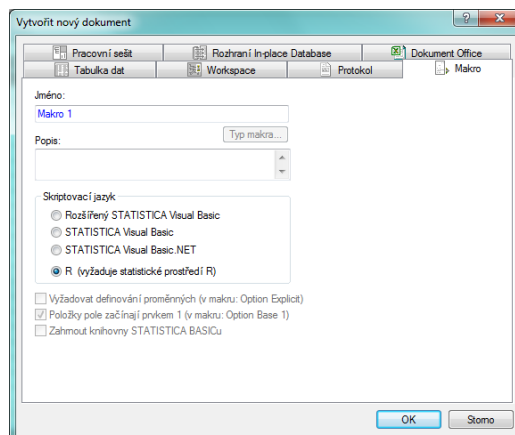
Jak to funguje

Při volání Rka se spustí spolu s makrem i podpůrný kód, který přenesou do Rka některé objekty a funkce potřebné pro dobrou komunikaci. (Vidíme, že makro má dva listy a jednom z nich jsou tyto instrukce – podpůrný kód.)



Vyzkoušejme si jednoduše prakticky, které objekty to jsou. Nejdříve si otevřeme nový skript Rka ve *STATISTICE*, to provedeme následovně, dejte *Soubor-Nový-Nový dokument-Makro* a zaškrtněte možnost *R* (viz dialog vpravo).

Do tohoto makra můžeme vložit či napsat přímo kód Rka s jeho příkazy a spustit jej pomocí tlačítka *Spustit makro* v záložce *Makro*.



Po spuštění makra je celý výstup konzole R zkopírován do protokolu.

Vyzkoušejme si to na příkladu, kdy budeme chtít vidět objekty, které se nakopírovaly do Rka se spuštěním makra, jak jsme již naznačovali výše. Příkaz Rka pro vypsání všech objektů uložených v paměti je `ls()`. Napišeme tedy do makra (záložky 1) pouze jednoduchý příkaz `ls()` a spustíme makro. Výsledek vidíte v protokolu *STATISTICA*. Výřez toho, co vidíme je vpravo. Je to přesně stejný vzhled výsledku, jako bychom makro spustili přímo v Rku.

Vidíme mnoho objektů, které se zde vytvořily. Slouží hlavně pro rozumnou komunikaci obou programů.

Vypíchněme si některé nahrané objekty a funkce.

Jako první bychom vyzdvihli objekt *ActiveDataSet* – toto je jednoduše aktivní tabulka, která byla přenesena do prostředí R ve formátu „data frame“.

Pokud tedy máme otevřený třeba soubor *Cars.sta* (najdete v příkladech programu *STATISTICA*) a spustíme následující příkazy (makro):

```

1 #' To transfer data between STATISTICA and R, use the
2 #' ActiveDataSet : R variable allocated by STATISTICA
3 #' Spreadsheet("[pathto/somefolder/]some[.sta]") : loads an
4 #' existing spreadsheet into R
5 #' RouteOutput(X, name = NULL, header = NULL) : route X to
6 #' STATISTICA as spreadsheet (graphs are automatically retrieved)
7 #' X - dataframe, vector, or matrix-type object; name, header -
8 #' string
9
10 X<-ActiveDataSet
11 summary(X)
12 plot(X[,31],X[,41])

```

```

> ls()
[1] "%+%" "ActiveDataSet"
[3] "as.StaSpreadsheet" "Assign"
[5] "dev.off" "Eval"
[7] "EvalMsgBox" "EvalMsgBoxR"
[9] "GetComError" "getSourceWithComments"
[11] "getSourceWithoutComments" "getSourceWrapper"
[13] "graphics.off" "help"
[15] "IIF" "MACRO"
[17] "MAS" "MAS.Data"
[19] "MAS.MonitorID" "MAS.ReturnCodes"
[21] "MsgBox" "MsgBoxR"
[23] "myTraceback" "myTry"
[25] "q" "quit"
[27] "RouteOutput" "Spreadsheet"
[29] "srcDF" "STATISTICA"
[31] "trim" "trim.left"
[33] "trim.right" "Uses"
[35] "Uses.repos" "WEB"
[37] "win.graph" "win.metafile"
[39] "win.print" "windows"
[41] "x11" "X11"

```

Poznámka k makru: trochu popíšeme něco málo ze znalostí Rka použitých v makru: komentář se uvozuje znakem mřížky. Přičažení objektu jsou znaky <-, *summary* je funkce pro výpočet shrnutí, tedy jakási analogie popisných statistik. Funkce *plot* je funkce pro vykreslení grafu. Rko je „case sensitive“, je tedy potřeba si dávat pozor na velká a malá písmena.

Objeví se následující výstup:

```

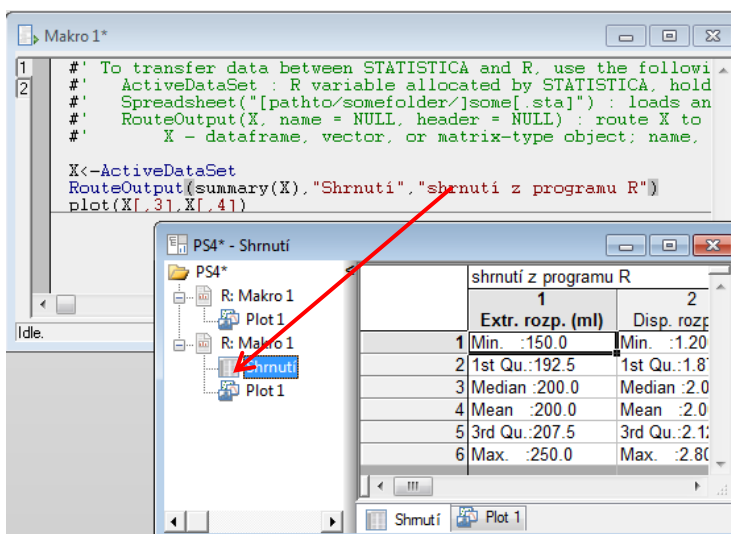
PS4* - R: Makro 1
# To transfer data between STATISTICA and R, use the following
# extensions:
# ActiveDataSet : R variable allocated by STATISTICA, holds the
# active dataset [as a data frame]
# Spreadsheet("[pathto/somefolder/]some[.sta]") : loads an
# existing spreadsheet into R
# RouteOutput(X, name = NULL, header = NULL) : route X to
# STATISTICA as spreadsheet (graphs are automatically retrieved)
# X - dataframe, vector, or matrix-type object; name, header -
# string
> X<-ActiveDataSet
> summary(X)
      PRICE      ACCELERATION      BRAKING
HANDLING
Min. -0.7984 Min. -2.2149 Min. -4.198894 Min. -4.23006
1st Qu.-0.6135 1st Qu.-0.5996 1st Qu.-0.008571 1st Qu.-
0.09137

```

Grafy generované skriptem jsou automaticky převedeny do protokolu ve formě škálovatelného grafu, tyto grafy jsou také replikovány jako grafy *STATISTICA* (škálovatelné obrázky jsou uloženy dovnitř grafu – jako příklad klikněte ve vygenerovaném protokolu na graf *Plot 1*). Protokoly, stejně jako grafy, mohou být vytisknuty nebo uloženy stejně jako by se jednalo o klasické výstupy.

Další důležité funkce: *Spreadsheet()* – slouží k otevření souboru s koncovkou .sta. Více detailů o této funkci i s jejími parametry najdete v dokumentu o integraci s R (str. 15), který najdete [zde](#).

Velmi důležitou funkcí je *RouteOutput()* – tato funkce dovolí vrátit objekt R jako tabulku *STATISTICA*. Uvedeme příklad: do již existujícího makra přidejte řádek pro výstup shrnutí do tabulky (viz obrázek vpravo):



Další informace

Tento článek byl jen o opravdu zásadních informacích, se kterými můžete ihned začít propojení Rka a *STATISTICA* využívat. Pokud by někdo bažil po detailnějších informacích a možnostech, pak ho odkážeme opět na skvěle sepsaný [White paper o integraci s R](#).