

**VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA**

**Hornicko – geologická fakulta**

**Katedra geoinformatiky**

## **PROSTOROVÁ ANALÝZA DAT**

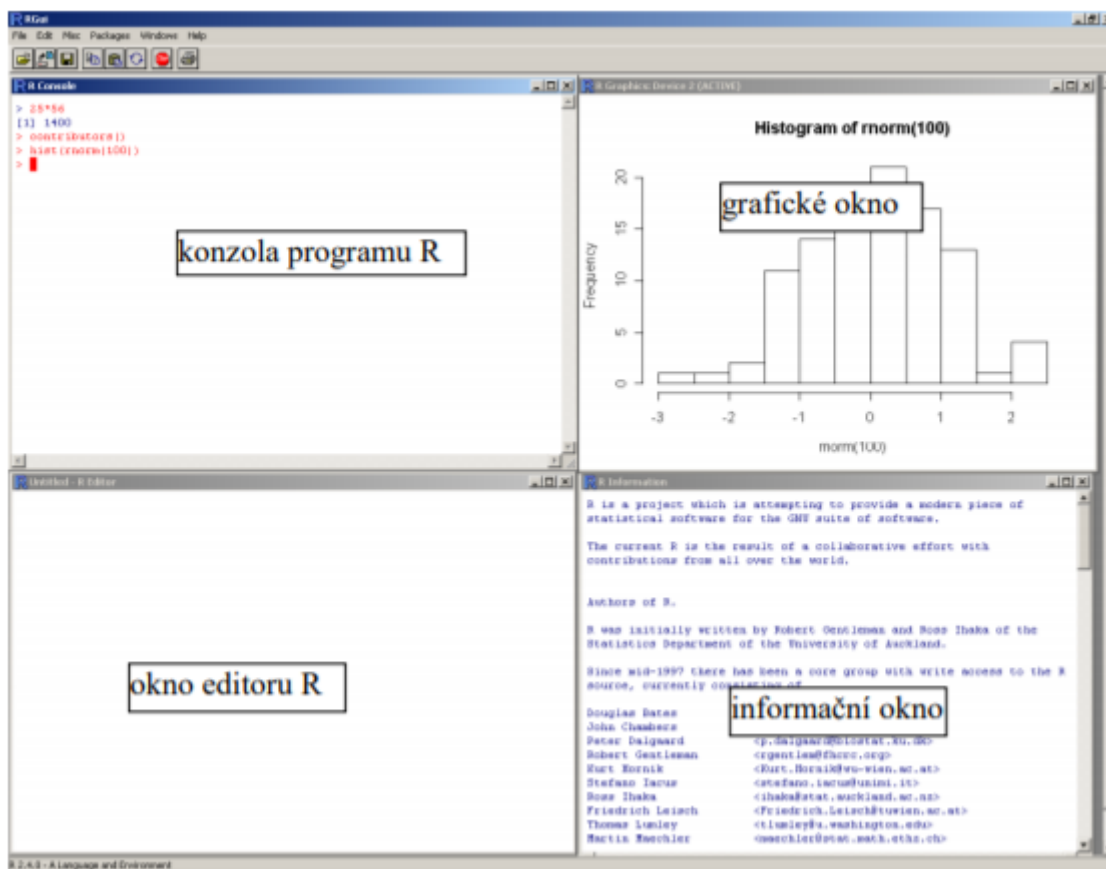
**Základy práce v prostředí R**

Lucie Orlíková

Ostrava, 2019

## Základní informace o prostředí a instalaci

Program lze získat z webových stránek projektu R (<http://www.R-project.org>), respektive z archivu CRAN (např. <http://cran.at.r-project.org>). K vlastnímu instalačnímu programu se dostanete v levém menu v oddílu Software volbou R Binaries. Dále se rozhodnete podle vašeho operačního systému (např. Windows) a následně zvolíte base, kde je stažitelný základní instalační program, např. R-2.5.1-win32.exe (27 MB). Vlastní instalace je poměrně jednoduchá a nevyžaduje žádné speciální znalosti. Balíček base obsahuje pouze základní knihovny (pro začátečníka jsou naprosto dostačující) – base (základní operace), datasets (knihovna příkladů dat), graphics, grDevices, grid (knihovny pro grafiku), methods (formálně definované metody a třídy objektů a programovací nástroje), stats, stats4, splines (statistické metody), tcltk, tools, utils (nástroje pro programování, vývoj a administraci). Kromě těchto základních knihoven však R obsahuje ohromné množství dalších nástrojů, které lze stáhnout z archivu CRAN (Packages) nebo přímo v menu programu Packages – Install packages.

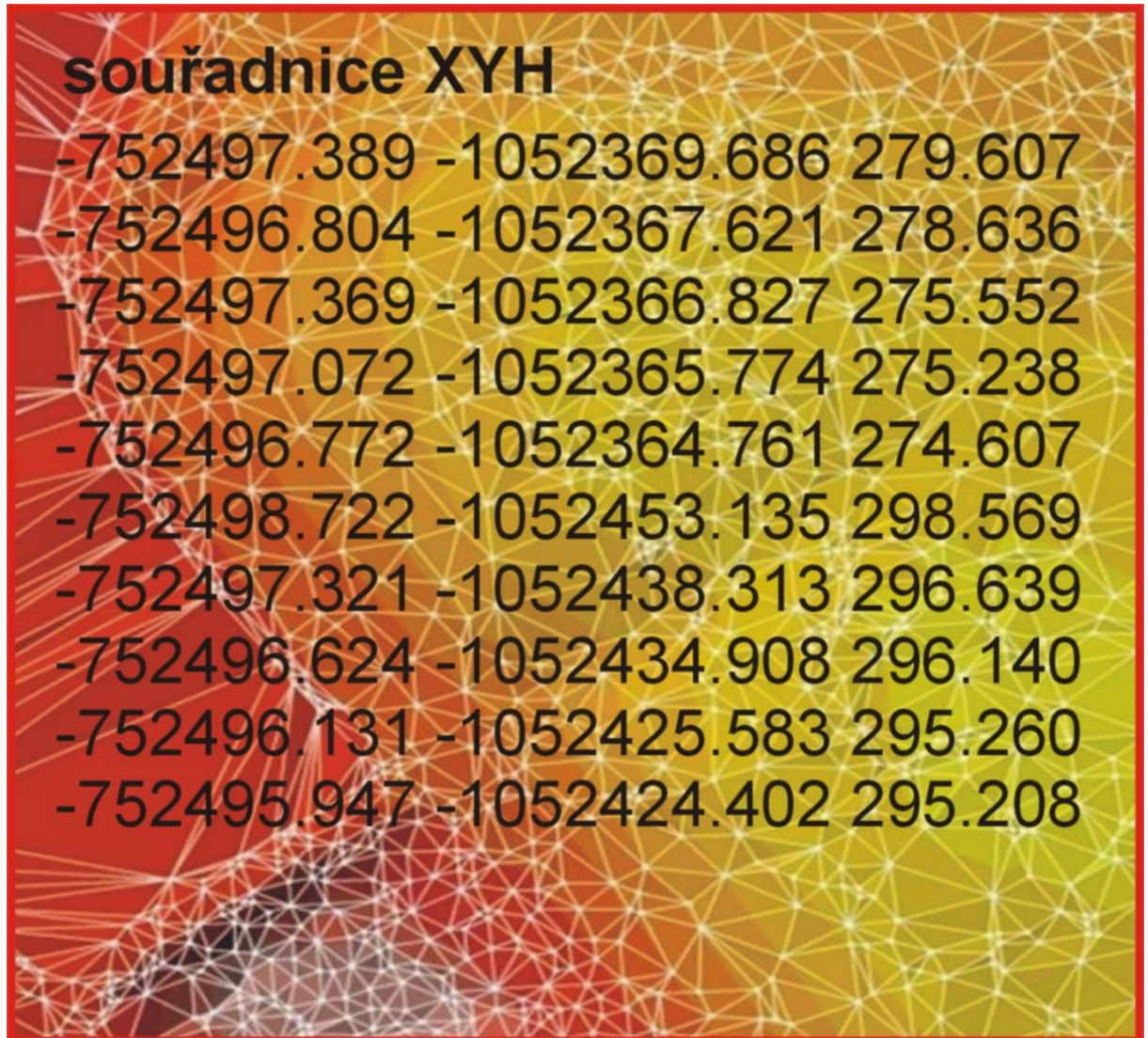


## Základní syntaxe

R je jako každý programovací jazyk velice striktní ve stylu zápisu. Při nedodržení přesného zápisu program buď hlásí chybu (syntax error) nebo může dojít k provedení odlišné operace. V tomto případě se jedná o sémantickou chybu, kterou počítač neodhalí, protože neví, že se po něm požaduje něco jiného.

### Cvičení na opakování práce s R studiem:

- Ze stránek ČÚZK využijte digitální model reliéfu DMR5G, ze kterého si každý vyextrahujte nadmořské výšky pro pravidelnou síť 50 bodů. Souřadnice bodů si vložte do excelu, přidejte sloupec s ID bodu. Takto vytvořený sloupec budete využívat pro práci v R studiu.



**souřadnice XYH**

-752497.389	-1052369.686	279.607
-752496.804	-1052367.621	278.636
-752497.369	-1052366.827	275.552
-752497.072	-1052365.774	275.238
-752496.772	-1052364.761	274.607
-752498.722	-1052453.135	298.569
-752497.321	-1052438.313	296.639
-752496.624	-1052434.908	296.140
-752496.131	-1052425.583	295.260
-752495.947	-1052424.402	295.208

**Zodpovězte následující otázky a nalezněte funkce pro řešení následujících úkolů:**

#Importujte si vytvořený soubor s body DMR5G.

#Odstraňte z dat sloupec Id a nechte si vypsát prvních pár záznamů z tabulky.

#Postupně spočítejte všechny základní charakteristiky pro sloupec s nadmořskými výškami - suma, průměr, medián, minimum, maximum, rozsah hodnot.

#Jak se od sebe liší hodnoty průměru a mediánu, vysvětlete rozdíl.

#Vypočítejte směrodatnou odchylku a rozptyl hodnot. Ověřte si správnost výsledku směrodatné odchylky přes odmocninu rozptylu.

#Co jsou to kvantily? Vypočítejte první, druhý a třetí kvartil.

#Vykreslete krabicový graf. Přidejte nadpis krabicový graf a popis osy y na nadmořská výška.

#Vykreslete histogram hodnot. Přidejte nadpis, upravte popisky osy x a y tak, aby pokrývaly celý rozsah hodnot. Změňte si počet tříd.