Apéro Technique de l'OREME #02 - 09/10/14

LES APEROS TECHNIQUES






- Dicussions techniques bimensuelles autour de la “mesure” et de la donnée

- Apéro Technique #01, juin 2014

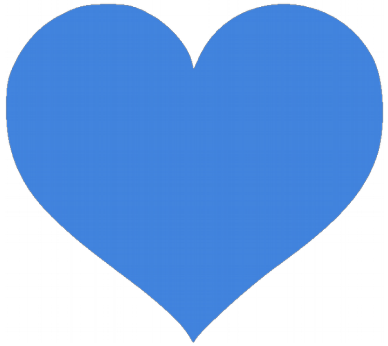
  Méthodes de rapatriement de données
  Méthodes d'alimentation de bases de données

- Apéro Technique #02, oct 2014

  R pour les données de la recherche #1

- Apéro Technique #03, déc 2014

  Outils d'analyse en géochimie

I  R

R pour les néophytes






R

- Langage et environnement pour l'analyse, la visualisation et la modélisation de données
- Standard chez les statisticiens … mais pas que !

http://cran.r-project.org/

Caractéristiques






Libre, open source, multi plateformes

Large communauté

Hyper extensible (~ 6000 packages !)

Langage interprété et interactif

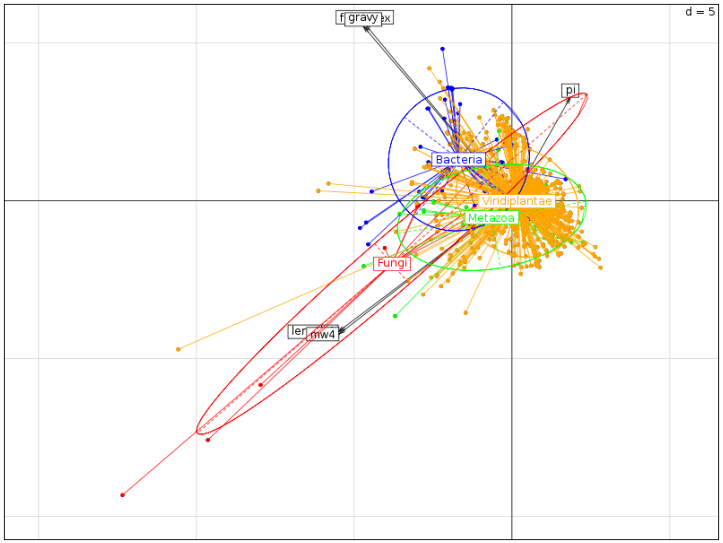
Script : automatisation des traitements, reproductibilité

Analyses de données / calcul






- Analyses statistiques :
  - Linéaire, non linéaire, séries temporelles
  - Clustering, classification
  - Spatial, ..



- Calcul sur les tableaux / matrices

Traitements






- Des packages toujours à la pointe
  - Génétique : genetics, qtl, seqinr ..
  - Données spatiales : plotKML, RgoogleMaps ..
  - Web : twittR, Shiny, Rgoogle-analytics
  - ..
- Liste des packages par tâche Non exhaustif !
- Top 100 des packages (2013)
- Top 50 des plus utilisés (2012)
- Liste des packages utiles (2010)

Graphiques






- Facilités graphiques
  - Tout est modifiable
  - Formules et symboles mathématiques
- Export : PDF, PS, JPG, ….

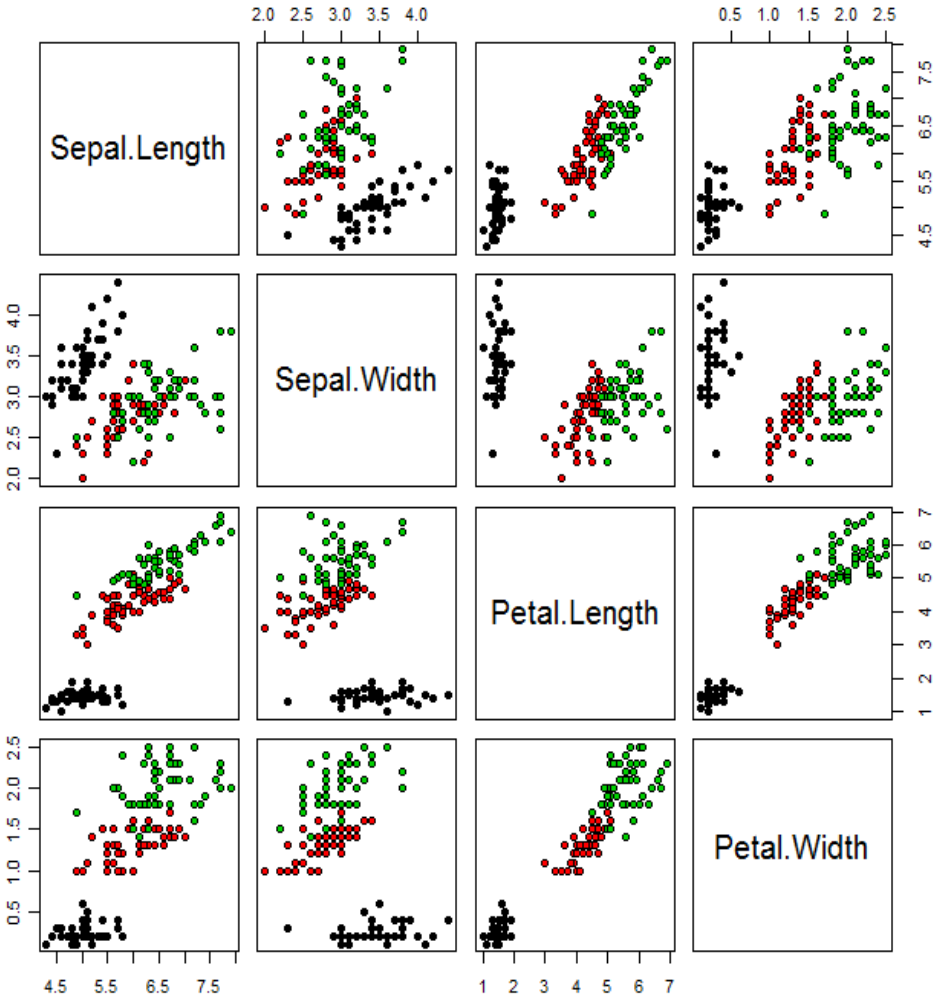
LES APEROS TECHNIQUES






Graphiques

Nuages de point



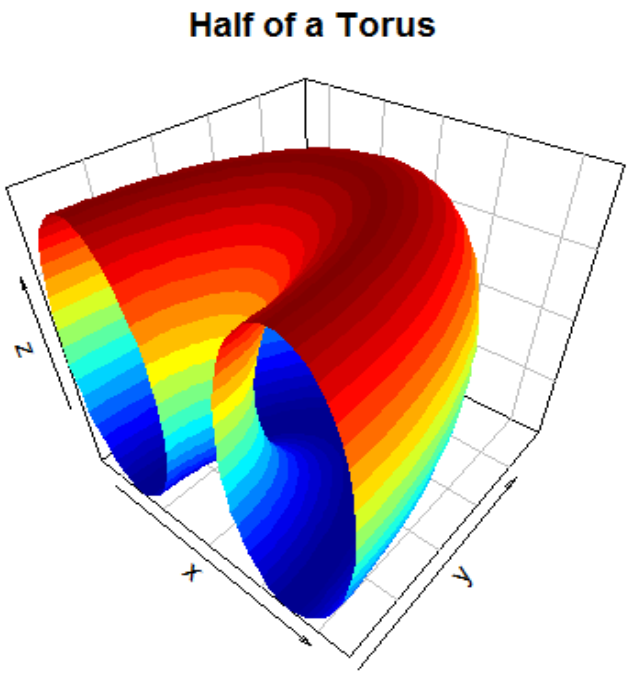
http://www.r-bloggers.com/graphically-analyzing-variable-interactions-in-r/






Graphiques

Perspective - 3D



http://blog.revolutionanalytics.com/2014/02/3d-plots-in-r.html

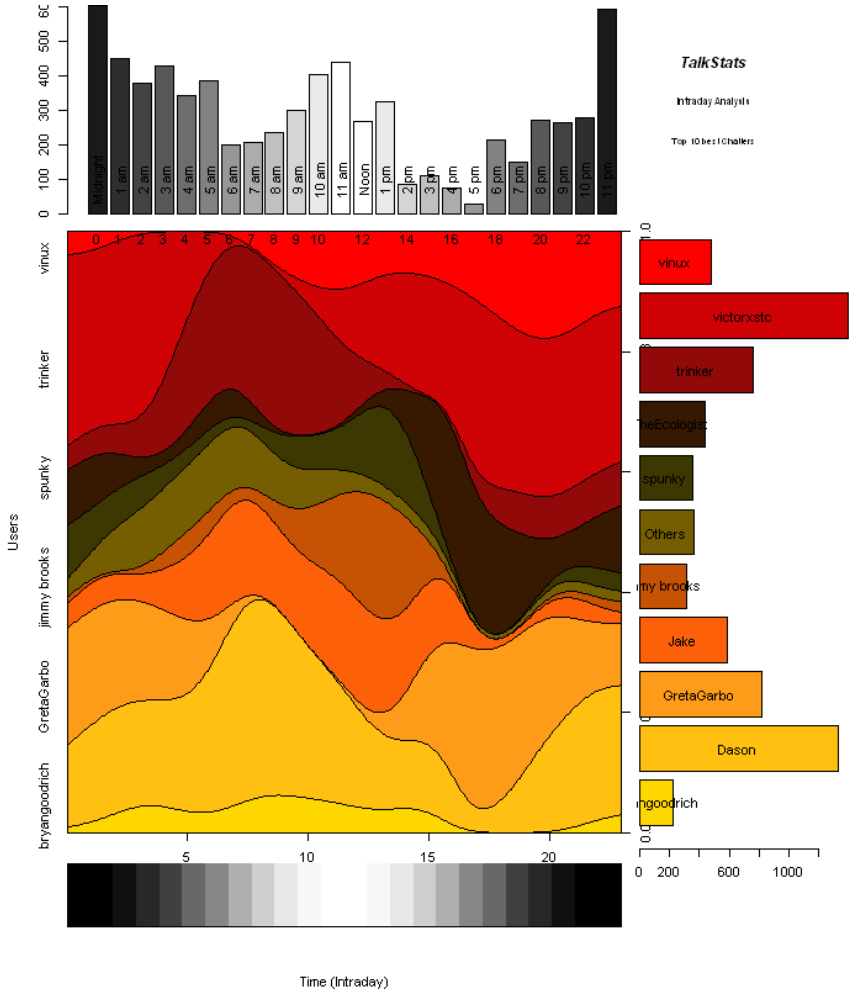
LES APEROS TECHNIQUES






Graphiques

Graphs conditionnels – multi graphs



http://www.talkstats.com/showthread.php/29320-R-Graphics-Beautiful-graphics-thread

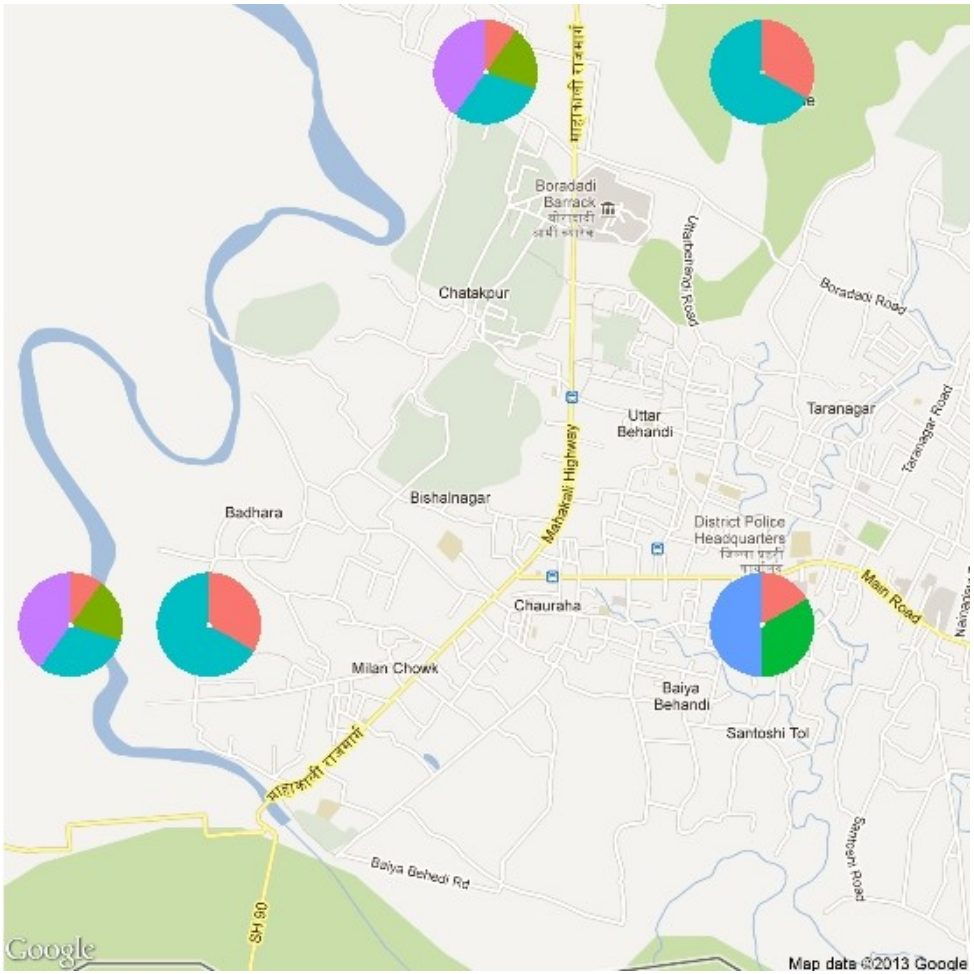
LES APEROS TECHNIQUES






Graphiques

Cartes

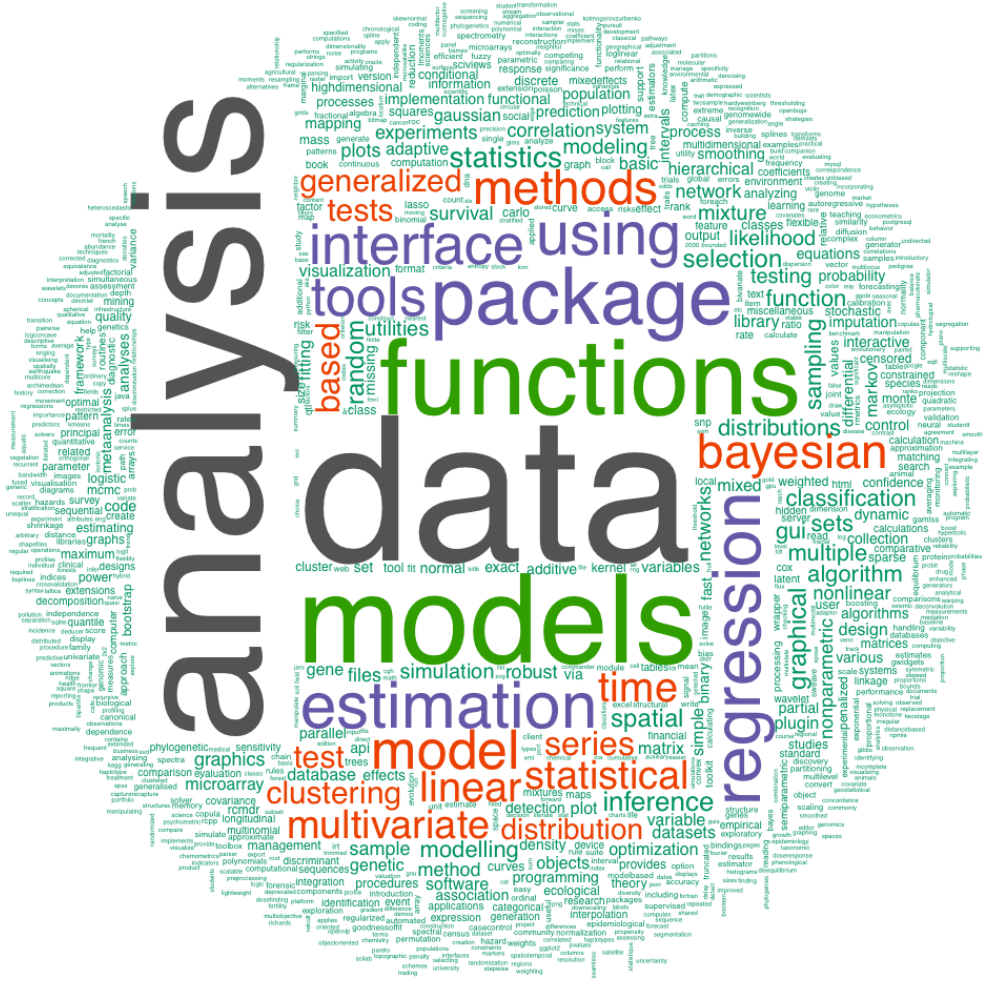


http://rgraphgallery.blogspot.fr/search/label/map

LES APEROS TECHNIQUES

Graphiques

Nuages de mots



http://onertipaday.blogspot.fr/2011/07/word-cloud-in-r.html

R et les données de la recherche






- Import / Export
  - Fichiers : CSV, texte, XLS, NetCDF, ..
  - Bases de données
  - Via le réseau (HTTP, FTP, web services, ...)
- Le type dataframe = variables x observations
- Traitement : Exploration, transformation, analyses, modélisation, visualisation
- Diffusion : Web, rapports dynamiques, ...

R et les données de la recherche






- Import / Export
  - Fichiers : CSV, texte, XLS, NetCDF, ..
  - Bases de données
  - Via le réseau (HTTP, FTP, web services, ...)
- Le type dataframe = variables x observations
- Traitement : Exploration, transformation, analyses, modélisation, visualisation
- Diffusion : Web, rapports dynamiques, ...

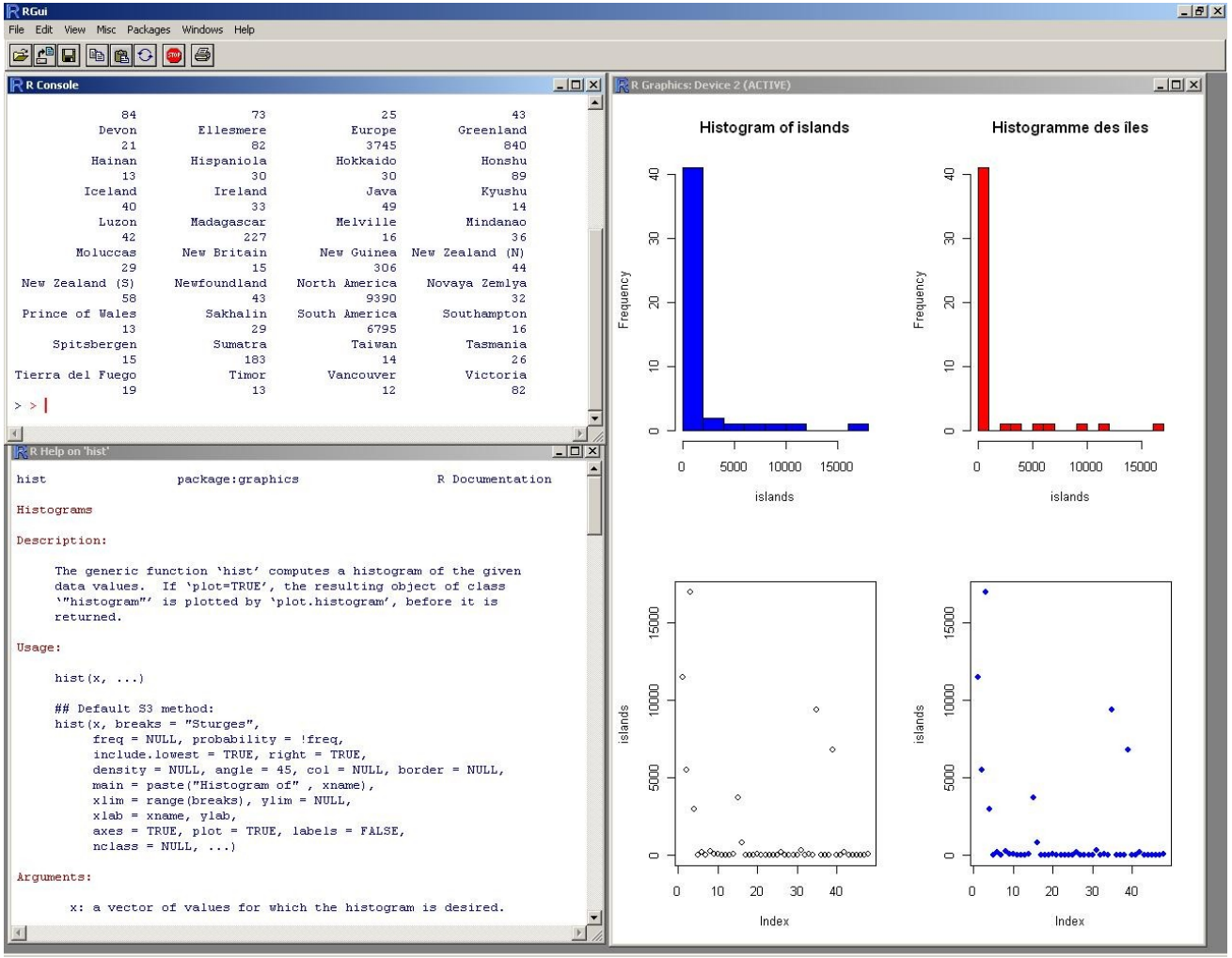
REPRODUCTIBLE

Interfaces






- RGui (Windows, Mac OS)



Interfaces






- Console

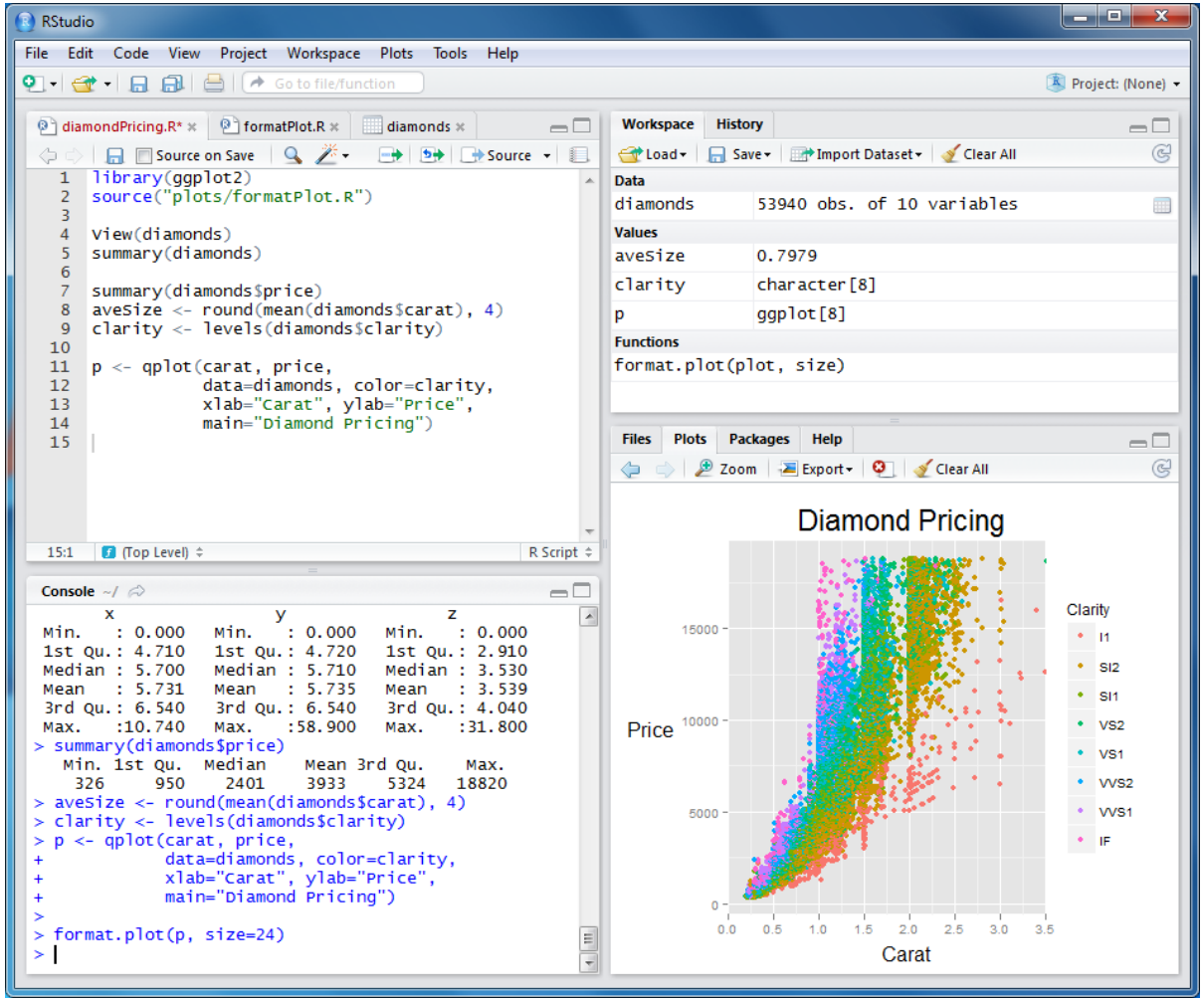
```
juliette@Port-OSU:~$ R

R version 3.1.1 (2014-07-10) -- "Sock it to Me"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> library(XLConnect)
XLConnect 0.2-7 by Mirai Solutions GmbH
http://www.mirai-solutions.com ,
http://miraisolutions.wordpress.com
> wb <-loadWorkbook('Bureau/OREME/so_keywords.xls')
> data <- readWorksheet(wb, 1)
> head(data)
  SO.short.name              SO.long.name               Keyword.FR
1        ECOPOP  Ecologie des populations écologie des populations
2          <NA>                      <NA>                évolution
3        EVOPOP Evolution des populations               populations
4          <NA>                      <NA>                      <NA>
5        EVOPOP Evolution des populations                     <NA>
6          <NA>                      <NA>                      <NA>
                                  GEMET.URI
1 http://www.eionet.europa.eu/gemet/concept/6520
2 http://www.eionet.europa.eu/gemet/concept/3026
3 http://www.eionet.europa.eu/gemet/concept/6519
4                                      <NA>
5                                      <NA>
6                                      <NA>
                                 ENVTHES.URI
1          http://vocabs.lter-europe.net/EnvThes3/82
2 http://vocabs.lter-europe.net/EnvThes3/USLterCV_186
3 http://vocabs.lter-europe.net/EnvThes3/USLterCV_439
4                                          <NA>
5                                          <NA>
6                                          <NA>
>
```

RStudio






- Environnement multi plateformes



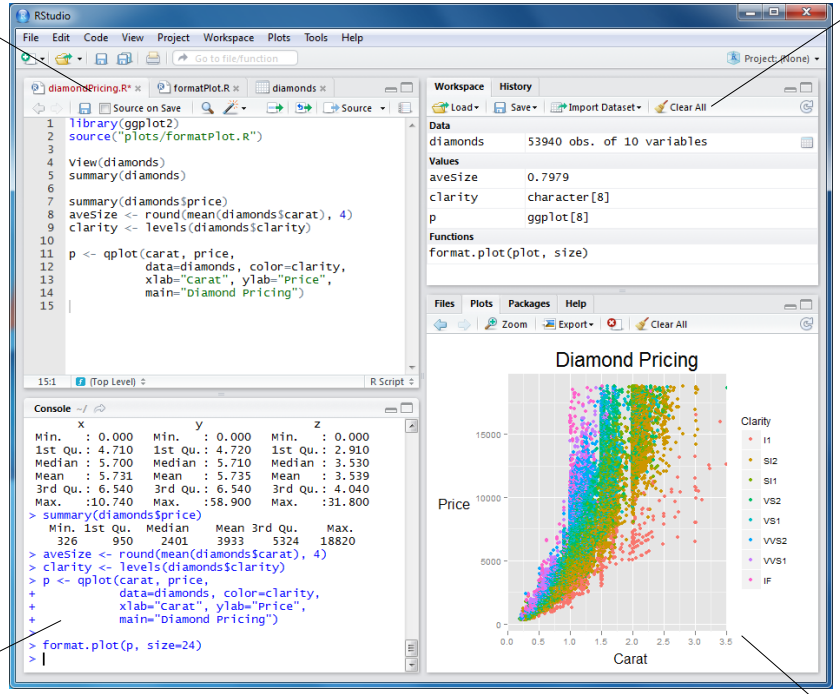





RStudio

Editeur : coloration syntaxique, aide à l'exécution

Objets, Historique, Versioning (Git et SVN)



Console

Système de fichiers, Graphiques, Librairies, Aide

RStudio






- Débuggage
- Création de “notebook” de code
- Mode serveur
- …

medycyss_datacite_metadata_export.R

juliette

Tue Oct 7 11:55:54 2014

```
# MEDYCYSS METADATA EXPORT INTO DATACITE SCHEMA
# author : Juliette Fabre
# creation : 09/14
# last update : 09/14
```

```
# The script reads MEDYCYSS metadata tables, and for specific station and data types it formats and ex
ports metadata into an XML file matching the Datacite format.

# Get argument
args <- commandArgs()
cleanargs <- sub("-.*=", "", args)
#type = cleanargs[regexpr("-t=", args)>0]

# Load functions file
script_path <- dirname(sub("--file=", "", args[4]))
source(paste(script_path, "/../oreme/oreme_db_v1.2_private.R", sep = ""))

# Load librairies
library(XML)

# Connect to OSU database
connexion <- connect_osu()
```

```
## Loading required package: DBI
```

http://www.rstudio.com/products/rstudio/






Liens

- Blogs, forums, …
  - Forum R du Cirad
  - R-bloggers
  - R graphs gallery
  - Inside-R
- Conférences :
  - France : Rencontres R (2014 Montpellier, 2015 Grenoble)
  - Monde : useR! (2014 LA, 2015 Danemark)