PENDAHULUAN

Sekílas Sejarah R

"R" sebenarnya bukan bahasa pemrograman yang baru. Setidaknya R telah dikembangkan secara intensif sejak 10 tahun yang lalu, sebagai pengembangan bahasa pemrograman "S" di Bell Laboratories. Tepatnya R adalah bahasa pemrograman yang telah didisain ulang untuk memudahkan analisis statistika. Menurut situs **R project**, R adalah bahasa dan lingkungan untuk komputasi statistik dan grafis. R adalah proyek berjenis *open source* GNU. Entah apa yang dipikirkan oleh sang pembuat dengan memberi nama karyanya hanya dengan satu huruf. Tapi apalah arti sebuah nama.

Walaupun awalnya dikembangkan untuk analisis statistik, namun saat ini telah berkembang aplikasinya hingga dapat melakukan manipulasi data spasial serta menampilkannya secara dinamis dalam situs web. Ditambah lagi dengan *era data analysis* atau akrab disebut *big data*, maka perkembangan R menjadi tidak terbendung lagi.

Perintah dasar dalam bahasa R telah menyediakan berbagai *tool* untuk pemodelan statistik linear dan nonlinear, analisis *time-series*, klasifikasi, analisis klaster, dan analisis grafis. Kemampuan ini terus berkembang dengan adanya ribuan paket tambahan yang diunggah ke server CRAN tiap tahunnya.

R dan Markdown

Dari pemaparan ringkas di atas, sudah jelas apa itu R. Sekarang apakah "*Markdown*" itu? Nama ini diberikan oleh kreatornya, karena itu, John Gruber seorang programmer mengembangkan *markup language "Markdown*". Ia menyederhanakan berbagai perintah LaTeX agar dapat lebih mudah dipahami pemakai bagi pengguna yang bukan programmer dan bukan ahli matematika. Salah satu contohnya adalah *R markdown*. Bahasa *markup (markup language)* yang lebih mudah dari LaTeX atau html sekalipun. Dengan menggunakan *R markdown*, saat ini bisa digunakan untuk membuat *blogpost* atau naskah buku dengan R, seperti halnya naskah yang sedang anda baca saat ini.

Karakter R

Beberapa karakter R di antaranya:

⇒ **R gratis,** *Open Source*, **dan** *Cross Platform*. Karena gratis dan *open source*, maka kita dapat mengembangkan R sesuai kebutuhan kita, misalnya dengan membuat *add on package*. Karena bersifat *cross platform*, maka para pengguna yang menggunakan sistem operasi (OS) Linux, Mac dan Windows dapat saling bekerjasama. Peningkatan versi R akan selalu dilakukan bersamaan. Oleh karenanya menggunakan OS apapun, kita akan memiliki versi R yang setara.

- ➡ R Mendukung Prinsip Reproducibility. R adalah aplikasi berbasis command line, artinya setiap perintah harus diketik sebagai baris perintah, yang dapat diulang oleh orang lain hanya dengan meng-copy-paste kode perintahnya. Prinsip ini disebut sebagai reproducibility. Bila anda melakukan hal ini dengan SPSS, Statistica, atau Minitab, yang berbasis point and click serta drag and drop, maka anda harus menangkap (screen capture) untuk menggambarkan urutan langkah analisis yang anda lakukan.
- ⇒ **R Menghasilkan Visualisasi yang Berkualitas Tinggi**. R memiliki kemampuan plot yang tinggi. Plot sangat diperlukan untuk memvisualisasikan hasil analisis anda. Bentuknya sudah bukan lagi hanya *scatter plot* dan histogram, tapi R sudah dapat membuat peta *chloropleth* dalam *format spasial*.

Komunitas pengguna R

R seperti halnya piranti lunak *open source* lainnya memiliki basis komunitas pengguna yang sangat banyak. Daftar berbagai komunitas R sebagian dapat dilihat di Situs *R-evolution*. Mereka berkumpul secara rutin dalam pertemuan pengguna R (*R meet up*) di berbagai negara. Kegiatan tersebut saat ini telah diadakan 127 kota di 31 negara, menurut situs *R user group*.

Kebiasaan yang Dianjurkan

Sebagai pengguna R yang sampai saat ini masih belajar, maka kami menganjurkan tiga hal berikut ini:

- ⇒ Belajar dan Berbagi. Pada hari anda memutuskan untuk menggunakan R, maka di hari itulah anda berkomitmen untuk berkontribusi kepada para pengguna lainnya. Caranya mudah sekali, bagilah pengetahuan baru yang anda pelajari, posting kode anda di blog atau media sosial anda. Bila anda memiliki akun Twitter gunakan hashtag#rstats pada tweet anda tentang R. Bila anda memiliki akun Google Plus, Statistics dan R adalah komunitas pengguna R yang dapat diikuti. Kebiasaan berbagi kode juga dapat dilakukan melalui akun Github. Plaform ini adalah semacam media sosial khusus untuk para programmer. Uniknya semua materi yang diunggah seluruhnya berlisensi bebas untuk dibagikan. Biasanya lisensi yang digunakan adalah Creative Commons Attribution (CC-BY) atau Creative Commons Zero (CC-0). Anda dapat mengkopi (cloning), membuat varian (forking) dari kode atau materi lainnya dari para pengguna dan memodifikasinya tanpa khawatir dituduh melakukan plagiarisme. Riwayat penyuntingannya pun dapat dilacak (file versioning) dan diketahui oleh penulis aslinya. Demikian pula pengguna lainnya dapat melakukan hal yang sama.
- Belajar Markdown Syntax. Secara umum kode R dan umumnya Github akan bekerja baik bila anda menggunakan format teks dalam dokumen anda. Anda boleh tidak percaya, bahwa sekarang anda dapat menulis satu buku lengkap dengan syntax Markdown. Syntax ini adalah penyederhanaan dari syntax LaTeX tapi dengan format perintah yang lebih sederhana dan mudah diikuti. Oleh karenanya selain harus menginstalasi R dan R Studio IDE, anda harus menginstalasi distribusi LaTeX. Berikut tautannya untuk masing-masing OS dari Situs LaTeX project: LaTeX for Linux, LaTeX for Mac, LaTeX for Windows. Bila anda masih banyak berhubungan dengan file format doc atau docx, maka anda perlu menginstalasi Pandoc. Dengan

dapat melakukan konversi format dokumen apa saja, misalnya: Markdown (.md) atau .html ke format .doc/.docx, begitu pula sebaliknya. Dengan Pandoc, maka anda dapat menulis apa saja dalam format Markdown langsung dari jendela R atau R Studio anda. Menarik bukan.

⇒ Pantau Package Terbaru. Anda perlu memantau keberadaan *package* terbaru, karena sangat mungkin 10 baris perintah menggunakan fungsi dasar R dapat digantikan oleh satu baris perintah menggunakan *package* tersebut. Ingat bahwa R adalah *open source*, oleh karenanya pasti ada setidaknya satu orang di belahan dunia yang lain yang membuat *package* untuk satu kebutuhan yang belum terpikirkan oleh orang lain.

Referensi

- 1. Gio, P.U. dan E. Rosmaini, 2015. Belajar Olah Data dengan SPSS, Minitab, R, Microsoft Excel, EViews, LISREL, AMOS, dan SmartPLS. USUpress.
- 2. Github site, url: www.github.com, diakses 14 Feb 2016
- 3. John Gruber Wikipedia site, url: https://en.wikipedia.org/wiki/John_Gruber, diakses 14 Feb 2016
- 4. Markdown syntax site, url: https://daringfireball.net/projects/markdown/syntax, diakses 14 Feb 2016
- 5. LaTeX project official site, url: http://latex-project.org/ftp.html, diakses 14 Feb 2016
- 6. R-evolution Analytics site, url: http://www.revolutionanalytics.com/, diakses 14 Feb 2016
- 7. R user group site, url: http://blog.revolutionanalytics.com/local-r-groups.html, diakses 14 Feb 2016

BAB 1

SEKILAS R

Berikut diberikan kutipan penjelasan singkat terkait R.

⇒ Norman Matloff dalam bukunya yang berjudul "*The Art of R Programming*" menyatakan sebagai berikut (2009).

"R is a scripting language for statistical data manipulation and analysis. It was inspired by, and is mostly compatible with, the statistical language S developed by AT&T. The name S, obviously standing for statistics, was an allusion to another programming language developed at AT&T with a one-letter name, C. S later was sold to a small firm, which added a GUI interface and named the result S-Plus.

R has become more popular than S/S-Plus, both because it's free and because more people are contributing to it. R is sometimes called "GNU S."

⇒ Lebih lanjut, Norman Matloff dalam bukunya yang berjudul "*The Art of R Programming*" menyatakan sebagai berikut (2009).

"Why use anything else? As the Cantonese say, yauh peng, yauh leng -"both inexpensive and beautiful. Its virtues:

- → A public-domain implementation of the widely-regarded S statistical language; R/S is the de facto standard among professional statisticians.
- **Comparable**, and often superior, in power to commercial products in most sense.
- → Available for Windows, Macs, Linux.
- → In addition to enabling statistical operations, it's a general programming language, so that you can automate your analyses and create new functions.
- **→***Object-oriented and functional programming structure.*
- **→**Your data sets are saved between sessions, so you don't have to reload each time.
- →Open-software nature means it's easy to get help from the user community, and lots of new functions get contributed by users, many of which are prominent statisticians.

 I should warn you that one submits commands to R via text, rather than mouse clicks in a Graphical

I should warn you that one submits commands to R via text, rather than mouse clicks in a Graphical User Interface (GUI). If you can't live without GUIs, you should consider using one of the free GUIs that have been develop for R, e.g. R commander or JGR. Note that R definitely does have graphics – tons of it. But the graphics are for the output, e.g. plots, not for the input.

The term object-oriented can be explained by example, say statistical regression. When you perform a regression analysis with other statistical package, say SAS or SPSS, you get a mountain of output. By

contrast, if you call lm() regression function in R, the function return an object containing all the results – estimated coefficients, their standard errors, residual, etc. You then pick and choose which parts of that object to extract, as you wish."

⇒ Andrie de Vries dan Joris Meys dalam bukunya yang berjudul "*R for Dummies*" menyatakan sebagai berikut (2012).

"R is more a programming language than application. When you download R, you automatically download a console application that's suitable for your operating system.

The R Development Core Team has put a lot of effort into making R available for different types of hardware and software. This means that R is available for Windows, Unix systems (such as Linux), and the Mac.

R itself is a powerful language that performs a wide variety of functions, such as data manipulation, statistical modeling, and graphics."

⇒ Lebih lanjut, Andrie de Vries dan Joris Meys dalam bukunya yang berjudul "*R for Dummies*" menyatakan sebagai berikut (2012).

"In order to start working in R, you need to use an editing tool. Which editing tool you use depends to some extent on your operating system, because R does not provide a single graphical editor for all operating systems. The basic R install gives you the following:

→Windows: A basic editor called RGui.

 \rightarrow Mac OS X: A basic R editor called R.app.

→Linux: There is no specific R editor on Linux, but you can use any editor (like Vim or Emacs) to edit your R code.

Fortunately, there is an alternative called RStudio, third-party software that provides a consistent user interface regardless of operating system. In addition to demonstrating how to work with the Windows RGui, we also illustrate how to use RStudio.

The one thing that R is not is an application, which means that you have the freedom of selecting your own editing tools to interact with R. In this section we discuss the Windows R editor, RGui (short for R graphical user interface). Since the standard, basic R editors are so, well, basic, we also introduce you to RStudio. RStudio offers a richer editing environment than RGui and makes some common tasks easier and more fun.

As part of the process of downloading and installing R, you get the standard graphical user interface (GUI), called RGui. RGui gives you some tools to manage your R environment – most important, a console window. The console is where you type instructions, or scripts, and generally get R to do useful things for you.

⇒ Garrett Grolemund dalam bukunya "*Hands-On Programming with R*" menyatakan sebagai berikut(2014).

"Learning to program is important if you're serious about understanding data. There's no argument that data science must be performed on a computer, but you have a choice between learning a graphical user interface (GUI) or a programming language. Both Garrett and I strongly believe that programming is a vital skill for everyone who works intensely with data. While convenient, a GUI is ultimately limiting, because it hampers three properties essential for good data analysis:

- → Reproducibility, The ability to-recreate a past analysis, which is crucial for good science.
- → Automation, The ability to rapidly re-create an analysis when data changes (as it always does).
- → Communication, Code is just text, it is easy to communicate. When learning, this makes it easy to get help whether it's with email, Google, Stack Overflow, or elsewhere.

Don't be afraid of programming! Anyone can learn to program with the right motivation, and this book is organized to keep you motivated. This is not a reference book; instead, it's structured around three hands-on challenges. Mastering these challengeswill lead you through the basics of R programming and even into some intermediatetopics, such as vectorized code, scoping, and S3 methods. Real challenges are a greatway to learn, because you're not memorizing functions void of context; instead, you'relearning functions as you need them to solve a real problem. You'll learn by doing, not by reading."

⇒ Lebih lanjut, Garrett Grolemund dalam bukunya "*Hands-On Programming with R*" menyatakan sebagai berikut(2014).

"Not every programmer needs to be a data scientist, so not every programmer will findthis book useful. You will find this book helpful if you're in one of the following categories:

- 1. You already use R as a statistical tool but would like to learn how to write your ownfunctions and simulations with R.
- 2. You would like to teach yourself how to program, and you see the sense of learninga language related to data science.

One of the biggest surprises in this book is that **I** do not cover traditional applications of **R**, such as models and graphs; instead, I treat R purely as a programming language. Why this narrow focus? R is designed to be a tool that helps scientists analyze data. Ithas many excellent functions that make plots and fit models to data. As a result, manystatisticians learn to use R as if it were a piece of software—they learn which functions do what they want, and they ignore the rest.

Greg compares R to SPSS, but he assumes that you use the full powers of R; in other words, that you learn how to program in R. If you only use functions that preexist in R, you are using R like SPSS: it is a bus that can only take you to certain places."

⇒ Garrett Grolemund dalam bukunya "*Hands-On Programming with R*" menyatakan sebagai berikut (2014).

R isn't a program that you can open and start using, like Microsoft Word or InternetExplorer. Instead, R is a computer language, like C, C++, or UNIX. You use R by writing commands in the R language and asking your computer to interpret them. In the olddays, people ran R code in a UNIX terminal window—as if they were hackers in a moviefrom the 1980s. Now almost everyone uses R with an application called RStudio, and Irecommend that you do, too.

RStudio is an application like Microsoft Word—except that instead of helping you writein English, RStudio helps you write in R. I use RStudio throughout the book because itmakes using R much easier. Also, the RStudio interface looks the same for Windows, Mac OS, and Linux. That will help me match the book to your personal experience. You can download RStudio for free. Just click the "Download RStudio" button and follow the simple instructions that follow. Once you've installed RStudio, you can openit like any other program on your computer—usually by clicking an icon on your desktop."

W. John Braum dan Duncan J. Murdoch dalam bukunya "A First Course in Statistical Programming with R" menyatakan sebagai berikut (2007, 1:4).

"R is an open-source computing packagewhich has seen a huge growth in popularity in the last few years. Being opensource, it is easily obtainable by students and economical to install in ourcomputing lab. One of us (Murdoch) is a member of theRcore developmentteam, and the other (Braun) is a co-author of a book on data analysis using R. These facts made it easy for us to choose R, but we are both strongbelievers in the idea that there are certain universals of programming, and in this text we try to emphasize those: it is not a manual about programming in R, it is a course in statistical programming that uses R.

This book uses R, which is an open-source package for statistical computing. "Open source" has a number of different meanings; here the importantone is that R is freely available, and its users are free to see how it is written, and to improve it. R is based on the computer language S, developed by John Chambers and others at Bell Laboratories in 1976. In 1993 Robert Gentleman and Ross Ihaka at the University of Auckland wanted to experiment with the language, so they developed an implementation, and namedit R. They made it open source in 1995, and hundreds of people around the world have contributed to its development. S-PLUS is a commercial implementation of the S language. Because both R and S-PLUS are based on the S language, much of what is described in what follows will apply without change to S-PLUS.

R can be downloaded from http://cran.r-project.org!. Mostusers should download and install a binary version. This is a version thathas been translated (by compilers) into machine language for executionon a particular type of computer with a particular operating system. R isdesigned to be very portable: it will run on Microsoft Windows, Linux, Solaris, Mac OSX, and other operating systems, but different binary versionsare required for each. In this book most of what we do would be thesame on any system, but when we write system-specific instructions, wewill assume that readers are using Microsoft Windows."

Berdasarkan pemaparan sebelumnya terkait R, dapat ditarik informasi bahwa R merupakan bahasa pemrograman statistika yang dapat digunakan untuk analisis dan manipulasi data statistika (pemodelan statistika), dan grafik. R diciptakan (*was created*) oleh Ross Ihaka dan Robert Gentleman (nama depan sama-sama diawali dengan huruf R & R, **Ross** dan **Robert**) dari departemen statistika, di Universitas Auckland, New Zealand. Saat ini R dikembangkan oleh *R Development Core Team*. Munculnya R terinspirasi oleh bahasa statistika dengan nama S (*statistical language S*). Bahasa statistika S diciptakan oleh John Chambers dan rekan, di *Bell Labs* atau *AT&T Bell Laboratories*. Selain bahasa statistika S, bahasa pemrograman C juga dikembangkan di *AT&T Bell Laboratories*.

Beberapa *software* sepertiSPSS, Minitab, SAS, EViews, dan sebagainya, perlu membayar sejumlah uang, untuk mendapatkan izin menggunakan, sementara R gratis. Di samping itu, R juga tersedia untuk di-*install* di sistem operasi *Windows, Mac OS X*, dan *Linux*. Gambar 1.1 merupakan tampilan

Untuk Rotasi Tampilan:

Click Views --> Rotate View --> Clockwise/CounterClockwise

engenalan Software R

Yudiantri Asdi, M.Sc.

Apa itu R?

- didasarkan pada bahasa pemrograman S yang dikembangkan oleh R adalah suatu free software untuk analisis data dan grafik yang Rick Becker, John Chambers, dan Allan Wilks dari AT&T Bell Laboratories (1976)
- R merupakan bahasa tingkat tinggi (high level language)
- pertengahan tahun 1997, telah ada grup inti R ("R Development Core Auckland, Selandia Baru. Banyak kelompok orang telah berkontribusi pada R dengan mengirimkan kode dan melaporkan bug. Sejak R ditulis oleh Ross Ihaka dan Robert Gentleman di Universitas Team") yang dapat memodifikasi arsip kode sumber R

Apa Keunggulan dan Kekurangan R?

Keunggulan

- Cepat dan gratis
- Ahli statistika dapat mengembangkan metode dengan membuat package
- Kemampuan grafik yang baik
- Unggul untuk simulasi dan analisis yang membutuhkan pemrograman secara intensif
- Mendorong kita berpikir tentang analisis data

Kekurangan

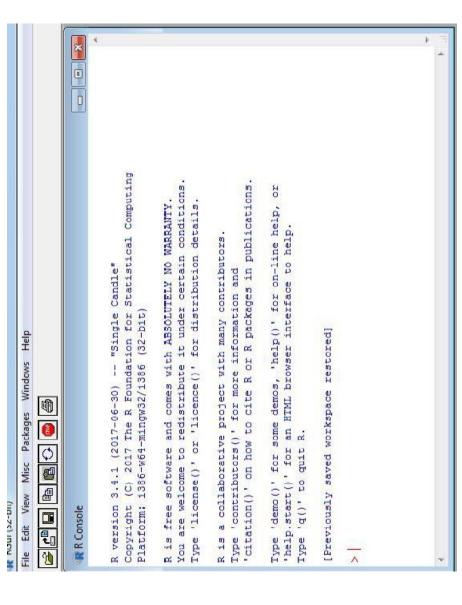
- Tidak user friendly
- Bagaimana menggunakan suatu fungsi dapat membuat kita frustasi
- Mudah membuat kesalahan.
 Error sudah menjadi hal yang biasa
- Penyiapan data dapat menjadi suatu yang sulit

Dimana mendapatkan R?

- R merupakan software yang bersifat open source (gratis) yang dapat diakses di alamat <u>https://www.r-project.org/</u>
- Versi R terakhir adalah 3.4.1 (30-06-2017)
- Untuk meningkatkan kemampuan, R juga dilengkapi dengan banyak package yang dapat digunakan untuk analisis statistik yang canggih

Bagaimana Cara Menginstal R?

- 1. Kunjungi <u>https://www.r-project.org/</u>
- . Pilih dan unduh versi terakhir R 3.4.1
- 3. Instal dan jalankan



Bagaimana Cara Kerja R?

- R dapat digunakan secara interaktif
- Hasil perhitungan segera dapat dilihat
- Untuk perhitungan yang lebih kompleks, perintah-perintah R ditulis lebih dahulu dalam text editor, kemudian dijalankan
- prompt ">" yang menunjukkan bahwa R siap menunggu perintah Setelah R terinstal pada komputer dan dijalankan, akan terlihat
- variabel, data, fungsi, hasil dan sebagainya, disimpan dalam memori aktif komputer dalam bentuk objek dan mempunyai sebuah nama R merupakan suatu bahasa berorientasi objek, artinya bahwa

- Pengguna dapat mengenakan aksi terhadap objek-objek tersebut melalui suatu operator (aritmatika, logika, dan relasi) dan fungsi (functions). Fungsi itu sendiri juga merupakan suatu objek
- Penamaan objek bersifat case sensitive (membedakan huruf besar dan kecil)

• R dapat digunakan seperti kalkulator. Ketik ekspresi dan tekan enter.

Untuk membuat variabel, gunakan <- atau =

a<-3

[1] 3

Membuat vektor x dan menghitung rata-ratanya

Apa itu Fungsi?

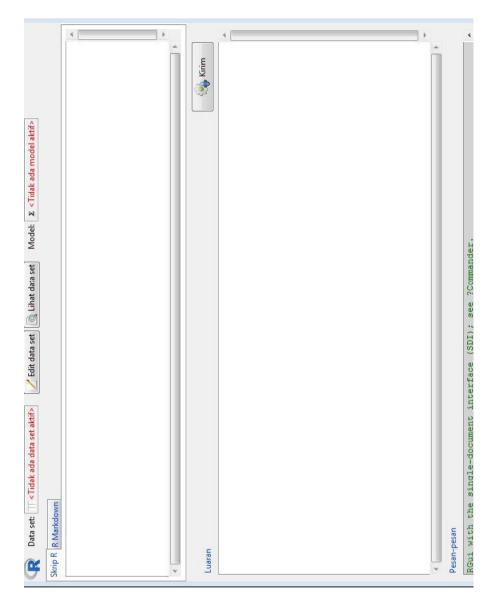
- Perintah dalam R merupakan fungsi
- Fungsi dituliskan dengan diakhiri oleh tanda "(" dan ")". Di dalam tanda kurung tersebut kadangkala diisi dengan satu atau lebih
- karena semua argumennya mempunyai nilai default (dapat diubah) atau karena tidak ada argumen yang didefinisikan pada fungsi tsb Beberapa fungsi ada yang tidak membutuhkan argumen. Hal ini
- Terdapat banyak fungsi yang tersedia dalam R
- Pengguna dapat membuat fungsi baru sesuai dengan kebutuhan

Apa itu package?

- analisis tertentu. Misalkan untuk analisis pengendalian mutu statistik, Package adalah kumpulan perintah-perintah yang digunakan untuk tersedia *package* 'qcc'. *Package* 'ggplot2' untuk penampilan visual data yang elegant.
- Terdapat ribuan package tersedia yang dapat diunduh secara gratis.
- Daftar semua package dapat dilihat di alamat https://www.rproject.org/
- Google Search Engine dengan kata kunci "..... in r". Misalnya "spatial Untuk menemukan package yang sesuai, dapat juga dicari melalui regression in r"

Package R commander

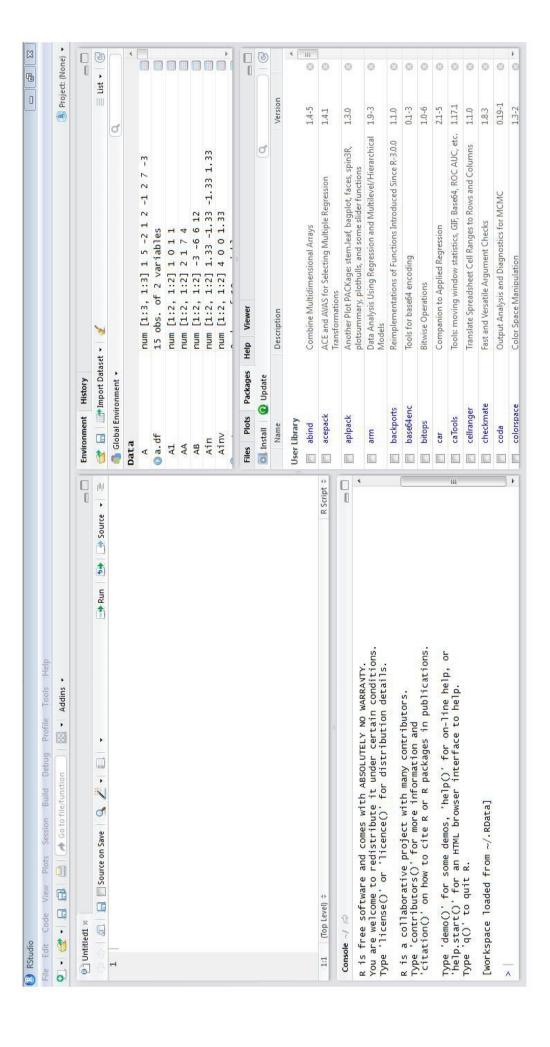
- Package ini digunakan untuk pemakaian R dengan tampilan menu sehingga pengguna dapat mengolah data dengan mudah
- Langkah-langkah menggunakan package ini:
- Package>InstallPackages>Cran Mirror Selection>Rcmdr
- Packages>Load Packages>Rcmdr



RStudio

- RStudio digunakan untuk lebih mempermudah pemakaian R
- Instalasi Rstudio dilakukan setelah R terinstal dengan langkah sbb:
- . Kunjungi alamat https://www.rstudio.com
- . Klik download
- 3. Klik pada Installers for Supported Platform,

Rstudio 1.0.143 - Windows Vista/7/8/10



Cara instal package dari RStudio

Klik Packages pada windows pada posisi kanan bawah

Klik Install

 Ketikkan nama package. Nama package harus tepat ejaan dan huruf besar kecilnya. Sebagai contoh

'qcc' : Quality Control Chart

'PolynomF' : Polynomials in R

Link Bermanfaat

- https://www.r-project.org/
- Rseek.org
- https://www.springer.com/series/6991
 A beginners' guide to R>>
- https://www.r-bloggers.com

Terima kasih