

Modul Praktikum Statistika dengan R

I Made Krisna Gupta

2020-08-09

Contents

| | | |
|----------|--|-----------|
| 1 | Pendahuluan | 5 |
| 1.1 | Modul apa ini? | 5 |
| 1.2 | Pengetahuan yang diperlukan | 5 |
| 1.3 | Kenapa R? | 5 |
| 1.4 | Alat dan bahan yang dibutuhkan | 6 |
| 1.5 | Tampilan R | 7 |
| 1.6 | update R | 8 |
| 2 | Memulai dengan R | 9 |
| 3 | Manipulasi Data | 11 |
| 4 | Visualisasi Data | 13 |
| 4.1 | Pendahuluan | 13 |
| 4.2 | Grafik | 14 |
| 4.3 | Latihan | 15 |
| 5 | Mean, Median, dan Modus | 17 |
| 6 | Tes Hipotesis | 19 |
| 7 | Analisis Regresi | 21 |

Chapter 1

Pendahuluan

1.1 Modul apa ini?

Modul ini merupakan panduan mahasiswa dalam mengaplikasikan teori tentang statistika dan analisis data ke dalam standar praktik di dunia usaha. Khususnya, modul ini bertujuan untuk membantu mahasiswa menggunakan R, sebuah bahasa pemrograman *open source* yang banyak digunakan oleh analis data untuk menganalisis data menggunakan statistik, dan melakukan visualisasi data. Lebih lengkap tentang R dapat ditemukan di <https://www.r-project.org/>

1.2 Pengetahuan yang diperlukan

Modul ini mengasumsikan bahwa mahasiswa telah mampu menggunakan komputer dan mengakses internet. Mengetik dan *point-and-click windows interface* dianggap tidak perlu diajarkan. Modul ini juga mengasumsikan pengetahuan mengenai statistik. Mahasiswa dianggap telah mengerti teori inferensi statistik (mean, median, dan modus), sampling, probabilitas, maupun regresi. Kemampuan menggunakan *spreadsheet* seperti Microsoft Excel atau Google Sheet merupakan pengetahuan yang sangat membantu, meskipun tidak diperlukan.

1.3 Kenapa R?

Ada beberapa aplikasi yang dapat digunakan untuk menganalisis data, seperti Microsoft Excel, EViews, SPSS, dan lain sebagainya. Beberapa alasan mengapa modul ini ditulis dengan menggunakan R:

1.3.1 Kebutuhan akan logika menulis kode

Beberapa aplikasi yang ada di pasaran saat ini seperti Microsoft Excel ataupun Google Sheet merupakan aplikasi yang cukup mudah dipelajari. Namun demikian, kebutuhan dunia usaha untuk tenaga kerja yang dapat menulis kode semakin meningkat. Menulis kode sendiri memiliki berbagai manfaat. Pertama, logika menulis kode akan mempermudah mahasiswa mempelajari bahasa lain yang lebih mendasar seperti Python. Bahasa-bahasa ini di masa depan dapat menjadi pilihan jika mahasiswa berniat meniti karir lebih khusus di dunia analisis data, di mana kebutuhan sumber daya manusia di sini masih sangat signifikan.

1.3.2 R termasuk cukup mudah dipelajari

R adalah bahasa yang memang ditulis untuk digunakan dalam analisis statistik. R adalah pintu masuk yang cukup baik dari statistisi menuju kemampuan lain yang berhubungan dengan *coding* seperti *text mining* dan lain sebagainya.

Selain modul ini, ada sangat banyak sumber belajar di internet. R memiliki komunitas pengguna yang cukup luas dan beragam. Komunitas-komunitas ini tidak segan-segan berbagi dan menjawab pertanyaan anda. Beberapa permasalahan yang anda temui mungkin sudah dijawab orang lain, dan forum-forum seperti ini akan sangat membantu ketika sudah bekerja.

1.3.3 Open Source

R adalah *open source*, dengan kata lain, *user* dapat menggunakan R secara cuma-cuma (iya, gratis).

1.4 Alat dan bahan yang dibutuhkan

Untuk menjalankan program-program yang ada di modul ini, anda membutuhkan sebuah komputer atau laptop, lalu anda akan memerlukan aplikasi bernama R dan Rstudio. Modul ini mengasumsikan anda menggunakan sistem operasi windows, namun melakukan instalasi di sistem operasi non-windows juga tidak kalah mudahnya. Langkah-langkah berikut ini juga dapat dengan mudah anda temukan di berbagai situs.

Anda juga membutuhkan kuota internet.

1.4.1 Menginstall R

R dapat didownload secara gratis di <https://cran.r-project.org/bin/windows/base/>. Kemudian anda dapat menetek pada tulisan “Download R X.x.x for windows (xx megabytes, 32/64 bit)” untuk memulai mengunduh R. Anda akan memulai mengunduh file dengan ekstensi .exe. Setelah unduhan selesai, silakan ketuk dua kali pada file .exe tersebut, dan klik next terus sampai instalasi dimulai.

1.4.2 Menginstall RStudio

Setelah R terinstall di komputer anda, silakan pergi ke <https://rstudio.com/products/rstudio/download/> untuk mengunduh RStudio. Pilih RStudio Desktop yang free, lalu ketuk “DOWNLOAD RSTUDIO FOR WINDOWS” untuk memulai pengunduhan. Setelah file selesai diunduh, silakan ketuk dua kali, klik next terus sampai instalasi dimulai.

1.5 Tampilan R

Setelah anda selesai menginstall RStudio, maka anda hanya perlu membuka RStudio dari desktop anda. Anda tidak perlu lagi membuka R. Karena itu, *shortcut* RStudio lebih penting daripada *shortcut* R.

RStudio memiliki tampilan utama berupa 4 jendela. Berikut adalah tampilan RStudio:

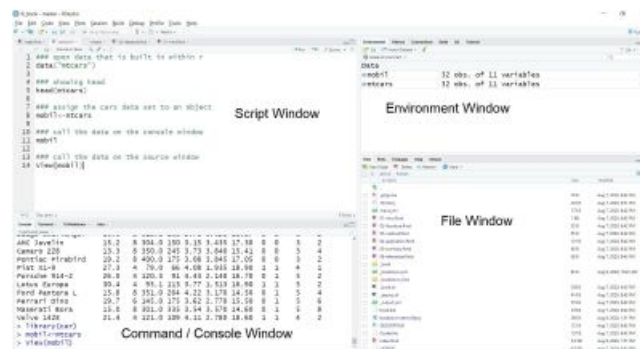


Figure 1.1: Tampilan RStudio

Secara garis besar, RStudio Memiliki 4(empat) jendela, yaitu script, console, environment dan file. kode yang ada di buku ini harus anda ketik di jendela ‘console’, sementara script merupakan kumpulan kode.

Untuk saat ini, mengetahui nama-nama dari keempat jendela ini sudah cukup. Bagaimana menggunakannya akan diperjelas di bab-bab berikutnya.

1.6 update R

R merupakan aplikasi yang cukup sering mendapatkan update. Karena itu, anda harus rajin-rajin ngecek update ketika menggunakannya. Jangan lupa juga bahwa anda akan memerlukan kuota internet untuk melakukan update.

Ada banyak cara untuk melakukan update terhadap R, tapi berikut ini akan disampaikan cara tercepat untuk melakukannya dengan menggunakan ‘console’ di RStudio.

pertama, instal paket bernama “installr”:

```
install.packages("installr")
```

Instalasi paket ini hanya perlu dilakukan pertama kali anda menginstall R. setelah sekali diinstal, paket itu akan selalu ada di komputer anda. Setelah menginstal paket “installr”, panggil paket tersebut dengan:

```
library(installr)
```

Fungsi library harus selalu dipanggil ketika akan menggunakan paket tersebut setiap kali anda memulai baru r. Setelah anda memanggil library tersebut, anda tinggal menggunakan fungsi “updateR()” pada console.

```
updateR()
```

Lalu di yes yes saja sampai update selesai.

Chapter 2

Memulai dengan R

under construction

Chapter 3

Manipulasi Data

under construction

Chapter 4

Visualisasi Data

4.1 Pendahuluan

Visualisasi data yang baik amat diperlukan guna mengkomunikasikan data yang kita miliki kepada orang lain. Visualisasi data dapat berbentuk tabel, grafik garis, grafik batang, dan lain sebagainya. Kali ini kita akan belajar cara melakukan visualisasi data dengan R.

Pada modul ini, kita akan menggunakan paket bernama `ggplot2` dan paket bernama `dplyr`. untuk itu, jika belum ada paket ini di komputer anda, maka harus diinstal dengan

Jangan lupa bahwa kita harus terkoneksi dengan internet untuk menginstall paket tersebut. Kita juga harus memanggil paket tersebut dengan

```
library(ggplot2)
library(dplyr)
library(tidyverse)
```

Kita akan menggunakan data yang telah disiapkan di situs buku ini.

```
dagang<-read.csv(url("https://imedkrisna.github.io/r/dataabi.csv"))
library(lubridate)
dagang$bulan<-dmy(dagang$bulan) # untuk membuat bulan menjadi waktu
head(dagang %>% select(bulan, udang, kopi, tembakau, sawit, tekstil, logam, kertas, crude))
```

| ## | | bulan | udang | kopi | tembakau | sawit | tekstil | logam | kertas | crude |
|------|------------|-------|-------|------|----------|--------|---------|--------|--------|-------|
| ## 1 | 2010-01-01 | 54958 | 36312 | 6425 | 565344 | 836334 | 593539 | 288466 | 718273 | |
| ## 2 | 2010-02-01 | 60874 | 36858 | 4728 | 805257 | 814565 | 548909 | 298346 | 826872 | |
| ## 3 | 2010-03-01 | 68049 | 39535 | 8512 | 928524 | 916195 | 1002479 | 350199 | 913640 | |

```
## 4 2010-04-01 69619 45247 10471 600508 870609 678313 357734 866674
## 5 2010-05-01 63006 60271 6909 788746 888548 654923 357008 903487
## 6 2010-06-01 83919 77593 7282 864754 981359 681485 345368 885780
```

data tersebut merupakan nilai ekspor per bulan Indonesia yang didapat dari Statistik Ekonomi dan Keuangan Indonesia (SEKI), Bank Indonesia, setelah sebelumnya diolah terlebih dahulu untuk mendapatkan tabel seperti di atas. Data dalam ribu USD.

4.2 Grafik

Data serial waktu (*time series*) seperti ini memang paling cocok divisualisasikan dengan menggunakan grafik garis. Kita akan mencoba menggambar grafik ekspor biji kopi. kode yang dapat digunakan adalah

```
grafik<-ggplot(dagang,aes(x=bulan,y=kopi))
grafik<-grafik+geom_line()+scale_x_date(date_breaks="1 year", date_labels = "%Y")
print(grafik)
```

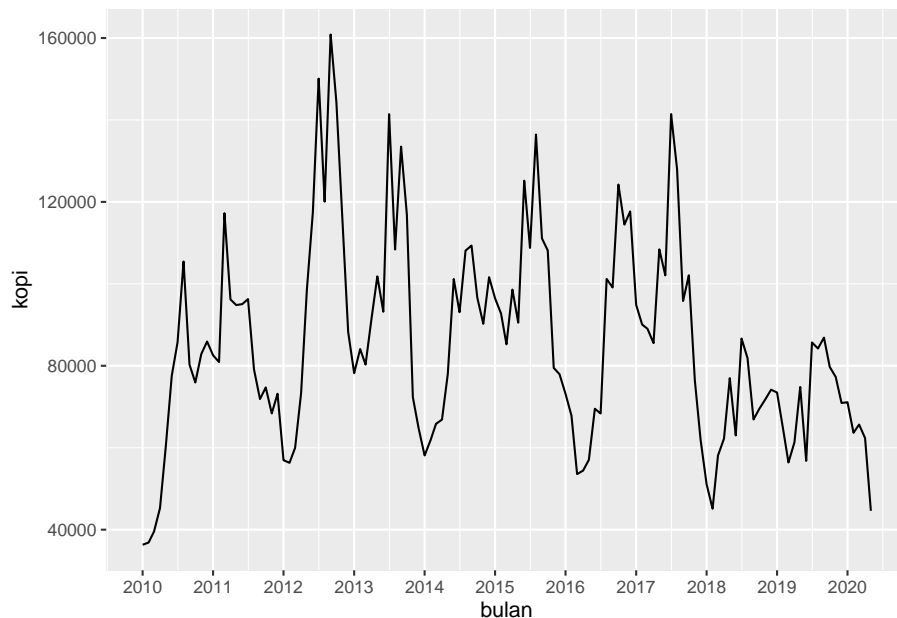


Figure 4.1: ekspor bulanan biji kopi

Grafik di atas menunjukkan perkembangan ekspor bulanan biji kopi Indonesia sejak Januari 2010 sampai Mei 2020.

4.3 Latihan

Apakah anda dapat menggambar grafik yang sama untuk ekspor kertas?
Tuliskan kode-nya di bawah ini

```
# kode untuk menggambar grafik ekspor kertas
```


Chapter 5

Mean, Median, dan Modus

under construction

Chapter 6

Tes Hipotesis

under construction

Chapter 7

Analisis Regresi

under construction