

Penerapan *Reproducible Research* pada Rstudio dengan Bahasa R dan Paket **knitr**

Weksi Budiaji

Abstract

Reproducible Research (RR) atau penelitian yang hasilnya dapat diulang sangat penting dalam sebuah penelitian karena peneliti lain dapat menjalankan ulang atau mengembangkan hasil penelitian sebelumnya dengan data dan kode yang telah tersedia. Rstudio sebagai *integrated development environment* (IDE) *software* R dan paket **knitr** menawarkan kemudahan untuk RR karena data dan kode dapat tersedia dengan mudah. Langkah pembuatan RR adalah pembuatan file `.Rmd`, penyesuaian **meta data**, penulisan teks dan kode (*chunk*), dan *rendering*. Penerapan teks-kode, teks-tabel, dan teks-gambar pada sebuah contoh kasus artikel dengan Rstudio, *software* R dan paket **knitr** memberikan keuntungan yaitu penyesuaian hasil dan pengecekan yang lebih menghemat waktu jika terjadi perubahan data, *cross reference* yang mudah pada tabel dan gambar, dan kemudahan peneliti lain memanfaatkan data dan kode yang telah tersedia. Pengadopsian *reproducible research* (data dan kode tersedia) menjadi penting bagi seorang peneliti agar kontribusi terhadap keilmuan lebih efektif.

Keywords: Reproducible Research, penelitian, kode, data

1 Pendahuluan

Reproducible research (RR) atau penelitian yang hasilnya dapat diulang merupakan elemen yang sangat penting dalam sebuah penelitian. Hasil penelitian dapat dikatakan *reproducible* jika peneliti lain dapat menghasilkan temuan yang sama pada data yang berbeda tetapi dengan prosedur yang sama (Gandrud 2015). RR sangat penting pada penelitian kuantitatif maupun kualitatif (King, Keohane, and Verba 1994), bahkan perubahan kebijakan tentang keharusan RR pada jurnal *Science* membuat 95% artikel yang diterbitkan dapat diproduksi ulang (*reproducible*) (V. Stodden, Seiler, and Ma 2018). Pada penelitian kuantitatif yang berdasarkan percobaan/ survei, RR berarti bahwa peneliti lain yang tidak terlibat dalam penelitian dapat meneliti ulang dengan percobaannya atau dengan samplingnya dan memvalidasi hasil temuannya yang sebanding dengan hasil penelitian terdahulu (Gandrud 2015).

Standar RR di statistika komputasi adalah data dan kode untuk menganalisis data tersedia bagi peneliti lain (R. D. Peng 2011), sehingga *tools* statistik yang berdasarkan menu akan sulit digunakan sebagai pendukung RR karena menu biasanya tidak didukung dengan kode. Stodden V., Leisch, and Peng (2014) menyebutkan beberapa *tools* yang dapat digunakan untuk RR antara lain *VisTralis*, *Sumatra*, dan *CDE*. Khusus *software* R, RR dapat menggunakan *sweave* (F. Leisch 2002) atau **knitr** (Yihui 2015) yang keduanya dapat menampilkan hasil

yang baik (Biecek and Kosinski 2017). **MATLAB** dan **Phyton** yang menggunakan *command line*/ kode juga dapat digunakan untuk **RR**, tetapi *software*/ bahasa **R** menjadi *tools* untuk **RR** yang paling mudah karena didukung oleh komunitas yang kuat dan kemunculan **Rstudio** sebagai *integrated development environment* (IDE) yang bebas (*free*) dan handal untuk bahasa **R**.

Software R adalah bahasa dan *software* yang dapat diperoleh dengan bebas untuk statistika komputasi dan grafik yang dikembangkan oleh R Core Team (2015). Salah satu kekuatan dari **R** adalah dukungan komunitas dalam bentuk paket-paket yang dapat diunduh dengan bebas dari [cran R](#). Ada banyak pilihan editor bahasa **R** untuk mempermudah penggunaan **R**, tetapi yang paling terkenal adalah [R studio](#) karena selain berfungsi sebagai editor, [R studio](#) juga merupakan IDE untuk bahasa **R**. **RR** pada [R studio](#) dapat menggunakan **sweave** yang berdasarkan pada bahasa **Tex** atau **knitr** yang menggunakan bahasa **rmarkdown**.

Artikel ini menggunakan **knitr** untuk menghasilkan contoh **RR** pada sebuah kasus data riil. Bab kedua menjelaskan metode/ cara **knitr** dengan baris-baris perintah/ kode dalam **Rstudio** yang dapat memunculkan hasil sesuai dengan standar **RR**. Bab ketiga menganalisis dan mendiskusikan data riil, sedangkan bab terakhir kesimpulan.

2 Metode

Paket **knitr** dibuat untuk memudahkan penggabungan teks dan kode **R** pada sebuah laporan agar jika terjadi perubahan pada data, peneliti tidak perlu lagi mengganti teks atau kode yang telah ditulis (Yihui 2017). Karakter ini yang membuat **Rstudio** dan **knitr** sangat cocok untuk **RR**. **knitr** dapat dipasang setelah **software R** dan **R Studio** terpasang dengan benar pada komputer yaitu dengan mengetikkan `install.packages(knitr)`. File baru yang harus dibuat untuk **RR** pada **Rstudio** adalah **R markdown**. Ekstensi dari file ini berupa `.Rmd`. Langkah-langkah untuk membuat sebuah file `.Rmd` agar dapat digunakan untuk **RR** adalah

1. Buka **Rstudio** dan pilih **File > New File > R Markdown** untuk membuat file berekstensi `.Rmd`.
2. Tulis **meta data** pada bagian atas (*front-matter*) dari file `.Rmd`.

meta data sangat penting karena berfungsi untuk mengatur teks yang akan dihasilkan misalnya judul, nama penulis, ukuran huruf, hasil teks (pdf, word, atau html) dan lain sebagainya. Bagian **meta data** berada pada blok paling atas dan diawali/ diakhiri dengan baris kode `---`. Artikel ini memiliki **meta data** sebagai berikut:

```
---
title: |
  | Penerapan *Reproducible Research* pada
  | `Rstudio` dengan Bahasa `R` dan
  | Paket `knitr`
```

```

author: "Weksi Budiaji"
bibliography: repro.bib
output:
  pdf_document:
    number_sections: yes
  html_document:
    highlight: pygments
    theme: journal
  word_document:
    reference_docx: style.docx
fontsize: 12pt
abstract: |
  *Reproducible Research* (RR) atau penelitian yang hasilnya dapat diulang
  sangat penting dalam sebuah penelitian karena peneliti lain dapat
  menjalankan ulang atau mengembangkan hasil penelitian sebelumnya.
  Software `R` menjadi alat RR termudah dibandingkan dengan software yang
  lain karena standar RR yang mengharuskan data dan kode tersedia. `Rstudio`
  sebagai *integrated development environment* (IDE) software `R` bahkan
  menyediakan fasilitas yang lengkap untuk RR, khususnya jika pengguna
  menggunakan paket `knitr`. Kemudahan `knitr` untuk RR diterapkan pada
  sebuah kasus data riil dengan hasil yang mirip dengan hasil penelitian
  asli. \par \textbf{Keywords: Reproducible Research, penelitian, kode, data}
---
```

3. Tulis teks dibawah meta data

Bagian teks berisi tulisan teks yang berada dibawah meta data. Teks yang dibuat dapat seperti teks pada MS Word, misalnya menebalkan atau memiringkan teks. Persamaan juga dapat ditulis, baik tersisip pada baris teks atau satu blok persamaan tersendiri. Jika sitasi disisipkan pada teks, sitasi tersebut akan muncul otomatis pada bagian akhir (daftar pustaka). Kode berikut adalah teks bagian Pendahuluan artikel ini:

Pendahuluan

Reproducible research (RR) atau penelitian yang hasilnya dapat diulang merupakan elemen yang sangat penting dalam sebuah penelitian. Hasil penelitian dapat dikatakan *reproducible* jika peneliti lain dapat menghasilkan temuan yang sama pada data yang berbeda tetapi dengan prosedur yang sama [Gandrud]. Gandrud menambahkan bahwa RR sangat penting pada penelitian kuantitatif maupun kualitatif. Pada penelitian kuantitatif yang berdasarkan percobaan/ survei, RR berarti bahwa peneliti lain yang tidak terlibat dalam penelitian dapat meneliti ulang dengan percobaannya atau dengan samplingnya dan memvalidasi hasil temuannya yang sebanding dengan hasil penelitian terdahulu [Gandrud].

4. Tulis kode bahasa R pada bagian kode (*chunk*)

Bagian kode/ *chunk* berisi kode/ bahasa R. Bagian kode dapat ditulis *inline*/ tersisip pada teks dengan menggunakan kode ‘**r ekspresi**’, misalnya angka 2 yang tertulis dalam kalimat ini menggunakan kode ‘**r 1+1**’ yang disisipkan (bukan menekan angka 2 di *keyboard*). *Chunk* dengan beberapa baris kode (blok kode) sebagai berikut

```
```{r }  
1+1
```
```

Saat blok kode digunakan, latar belakang Rstudio akan berubah (*highlight*). Jika tidak ada *highlight* dilatar belakang, blok kode akan dianggap teks biasa sehingga blok kode tidak dijalankan oleh R. Diantara dua kurung kurawal ({...}) dapat dituliskan pilihan-pilihan untuk mengatur hasil yang diinginkan, misalnya **echo** untuk menampilkan/ menyembunyikan kode R, **fig.cap** untuk memberi judul pada gambar yang dihasilkan, dan sebagainya. Perlu diperhatikan bahwa kedua tanda kurung tersebut harus dalam satu baris kode (jangan menekan **enter**) agar kode blok dapat bekerja saat dijalankan.

5. Jalankan file .Rmd (*rendering*)

File .Rmd dapat diubah menjadi format lain dengan memilih Knit HTML, Knit PDF, atau Knit Word. Komputer harus terpasang software pemroses **latex** dokumen seperti **Miktex**, **Livetex**, atau **Mactex** jika hasil yang diinginkan berupa **pdf**, atau cara lain dengan memasang paket **tinytex** di Rstudio yaitu dengan menjalankan `install.packages(tinytex)`.

Langkah-langkah pembuatan file untuk RR dapat diterapkan sejak dari awal penelitian, sehingga penyesuaian teks dan kode lebih mempersingkat waktu. Contoh kasus RR yang didiskusikan pada bagian berikutnya adalah kasus riil dari pemetaan komoditas sayuran di Kota Serang, Provinsi Banten yang menggunakan indikator produksi dan sosial-ekonomi (Budiaji, Pancawati J., and Suhera 2017). Pembahasan contoh kasus ini mirip dengan artikel asli (bukan menerjemahkan) dengan menitikberatkan pada penggunaan kode R pada paket **knitr**. Data dan kode R berekstensi .Rmd tersedia dan dapat diunduh secara gratis di laman penulis (<https://budiaji.info/publications/>) untuk memenuhi standar RR pada statistika komputasi (R. D. Peng 2011), sehingga peneliti lain dapat memproduksi ulang file .Rmd dan mendapatkan hasil yang sama dengan artikel ini.

3 Hasil dan Diskusi

Data set pada kasus pemetaan sayuran di Kota Serang (Budiaji, Pancawati J., and Suhera 2017) terdiri atas dua sumber yaitu data sosial ekonomi petani hasil survei dari 33 responden di enam kecamatan (Cipocok Jaya, Curug, Kasemen, Serang, Taktakan, dan Walantaka) dan data produksi sayuran yang berasal dari data sekunder (BPS 2012). Bab ini mengambil sebagian pembahasan dari artikel asli untuk diproduksi ulang yang menitikberatkan pada bagian (1) teks dan kode, (2) teks dan tabel, dan (3) teks dan gambar agar artikel ini menjadi contoh artikel RR yang lengkap.

3.1 Teks dan kode

Paragraf pertama dari *Respondent characteristics* dapat dijadikan sebagai pembahasan untuk teks dan kode. Sebelum menuliskan teks dan kode yang berdasarkan pada suatu data, *load data* (data berupa file `kuisisioner.txt` berada pada folder `C:/`) dijalankan terlebih dahulu agar data terbaca oleh sistem dan variabel-variabel didefinisikan untuk mempermudah penulisan kode *inline*. Variabel `survei` adalah data yang telah tersimpan disistem, sedangkan variabel-variabel lain, misalnya `responden`, `laki`, `usia40`, dan seterusnya, untuk membantu penulisan kode *inline*. Variabel-variabel yang akan digunakan dalam artikel dapat disimpan terlebih dahulu tanpa menampilkan pada artikel yaitu dengan cara menambahkan `echo = FALSE`. Penulisan di Rstudio dapat dilakukan seperti berikut:

```
```{r echo = FALSE}
survei <- read.table('C:/kuisisioner.txt', header=T)
responden <- nrow(survei)
laki <- (sum(survei[, 'JK'] == 1)/ responden)*100
usia40 <- (sum(survei[, 'Usia'] > 40)/responden)*100
pendrendah <- (sum(survei[, 'Pendidikan']!=3)/responden)*100
minut <- paste('Rp', prettyNum(min(survei[, 'PengeluaranUT']),
 big.mark = ','))
maxut <- paste('Rp', prettyNum(max(survei[, 'PengeluaranUT']),
 big.mark = ','))
utrendah <- (sum(survei[, 'PengeluaranUT']<6000000)/responden)*100
minrt <- paste('Rp', prettyNum(format(min(survei[-12, 'PengeluaranRT'])*30,
 scientific = FALSE), big.mark = ','))
maxrt <- paste('Rp', prettyNum(max(survei[, 'PengeluaranRT'])*30,
 big.mark = ','))
rtrendah <- (sum(survei[, 'PengeluaranRT']<2400000/30)/responden)*100
kembang <- (sum(survei[, 'Luaskembang']!=0)/responden)*100
nonkembang <- unique(survei[survei[, 'Luaskembang']==0, 'Kecamatan'])
minha <- min(survei[survei[, 'Luaskembang']==0, 'Luas'])/10000
maxha <- max(survei[survei[, 'Luaskembang']==0, 'Luas'])/10000
```
```

Kemudian satu paragraf yang berisi teks dan kode dapat ditulis. Berikut penulisan teks dan kode di Rstudio:

```
Ada `r responden` petani sayuran yang terlibat dalam survei.
`r round(laki,0)`\% dari responden berjenis kelamin laki-laki.
Usia responden paling muda adalah `r min(survei[, 'Usia'])` tahun
dan yang paling tua adalah `r max(survei[, 'Usia'])` tahun dan
sebanyak `r round(usia40,0)`\% usianya diatas 40 tahun. Mayoritas
responden memiliki latar belakang pendidikan yang rendah,
sekitar `r round(pendrendah,0)`\% adalah
```

```

lulusan sekolah dasar dan sekolah menengah pertama. Pengeluaran usaha tani
responden mempunyai variasi yang tinggi dari `r minat`
hingga `r maxut` sesuai dengan luas lahan sistem usaha
tani responden. `r round(utrendah,0)`\% responden mempunyai
pengeluaran untuk sistem usaha tani (SUT) kurang dari Rp 6.000.000.
Pengeluaran rumah tangga responden per bulan bervariasi dari `r minrt`
hingga `r maxrt` dengan sekitar
`r round(rtrendah,0)`\% responden mempunyai
pengeluaran rumah tangga per bulan kurang dari Rp 2.400.000. Ada
`r round(kembang,0)`\% responden berkeinginan untuk mengembangkan
SUT, dengan `r 100-round(kembang,0)`\% yang tidak berkeinginan
mengembangkan SUT berada di Kecamatan `r nonkembang` karena
mempunyai akses ke SUT sekarang hanya `r minha` sampai dengan
`r maxha` ha.

```

Hasil dari teks dan kode tersebut adalah:

Ada 33 petani sayuran yang terlibat dalam survei. 94% dari responden berjenis kelamin laki-laki. Usia responden paling muda adalah 25 tahun dan yang paling tua adalah 65 tahun dan sebanyak 79% usianya diatas 40 tahun. Mayoritas responden memiliki latar belakang pendidikan yang rendah, sekitar 91% adalah lulusan sekolah dasar dan sekolah menengah pertama. Pengeluaran usaha tani responden mempunyai variasi yang tinggi dari Rp 500,000 hingga Rp 45,000,000 sesuai dengan luas lahan sistem usaha tani responden. 70% responden mempunyai pengeluaran untuk sistem usaha tani (SUT) kurang dari Rp 6.000.000. Pengeluaran rumah tangga responden per bulan bervariasi dari Rp 600,000 hingga Rp 7,500,000 dengan sekitar 73% responden mempunyai pengeluaran rumah tangga per bulan kurang dari Rp 2.400.000. Ada 91% responden berkeinginan untuk mengembangkan SUT, dengan 9% yang tidak berkeinginan mengembangkan SUT berada di Kecamatan Taktakan karena mempunyai akses ke SUT sekarang hanya 0.2 sampai dengan 0.5 ha.

Jika dibandingkan teks dan kode antara Rstudio dan hasil akhir, terdapat perbedaan pada mayoritas angka/ bilangan yang dihasilkan pada hasil akhir. Angka yang tercetak pada hasil akhir merupakan hasil olahan dari kode R *inline*. Keuntungan yang diperoleh adalah jika terjadi perubahan pada data, misalnya salah *input* sehingga harus dijalankan/ dianalisis ulang, penggunaan teks dan kode ini sangat menguntungkan karena hanya memerlukan pengecekan yang lebih minimal (lebih menghemat waktu) dibandingkan tanpa kode *inline*. Keuntungan lain adalah peneliti yang ingin menggunakan kode juga lebih mudah, karena kode tersedia sehingga *Reproducible Research* terjamin.

3.2 Teks dan tabel

Tabel korelasi dan sidik ragam (*analysis of variance*) dari karakteristik responden terhadap akses usaha tani di masa depan (Budiaji, Pancawati J., and Suhera 2017) dijadikan pembahasan untuk teks dan tabel. Variabel yang akan dikorelasikan harus didefinisikan terlebih dahulu sebelum membuat tabel korelasi. Ada empat variabel yang akan dikorelasikan yaitu

usia, pendidikan, pengeluaran rumah tangga, dan jumlah tanggungan (jiwa). Blok kode yang dapat ditulis adalah sebagai berikut:

```
```{r echo = FALSE}
idx <- c(8:9,11,24)
korelasi <- round(cor(survei[,idx],method='spearman'),2)
korelasi[upper.tri(korelasi, diag = TRUE)] <- ''
knitr::kable(korelasi,
 caption = 'Korelasi antar variabel \\label{tbl:korelasi}')
```

Hasilnya adalah

Table 1: Korelasi antar variabel

	Usia	Pendidikan	PengeluaranRT	Jiwa
Usia				
Pendidikan	0.22			
PengeluaranRT	0	0.4		
Jiwa	0.09	0.1	-0.03	

Analisis ragam juga diterapkan pada empat variabel pada korelasi dengan variabel tak bebasnya adalah akses usaha tani dimasa depan. Salah satu variabel bebas yaitu pendidikan perlu dikategorikan dahulu, jika tidak maka nilai 0, 1, 2, dan 3 akan dianggap nilai kontinu oleh R. Variabel tak bebas masadepan merupakan penjumlahan dari akses usaha tani sekarang dan usaha tani yang ingin dikembangkan. Analisis sidik ragam pada artikel tersebut tidak menggunakan nilai asli dari variabel tak bebas tetapi menggunakan nilai logaritmanya, sehingga kode untuk tabel sidik ragam adalah

```
```{r echo = FALSE}
survei$PendidikanF <- as.factor(survei$Pendidikan)
survei$masadepan <- survei$Luas + survei$Luaskembang
model.1 <- lm(log(masadepan)~Usia + PendidikanF +
              PengeluaranRT + Jiwa, data=survei)
knitr::kable(anova(model.1), digits = 2,
              caption = 'Analisis sidik ragam \\label{tbl:anova}')
```

Hasil yang diperoleh adalah

Table 2: Analisis sidik ragam

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Usia	1	0.79	0.79	1.11	0.30
PendidikanF	3	2.85	0.95	1.34	0.28
PengeluaranRT	1	3.18	3.18	4.49	0.04
Jiwa	1	1.81	1.81	2.55	0.12
Residuals	26	18.43	0.71	NA	NA

Saat menampilkan tabel, `cross reference` dari tiap tabel dapat digunakan pada teks, misalnya di Rstudio

Tabel `\ref{tbl:korelasi}` menunjukkan bahwa korelasi antar variabel sangat kecil, sehingga regresi linear sederhana dapat diterapkan untuk mendeteksi variabel yang berpengaruh terhadap akses ke usaha tani dimasa depan. Tabel `\ref{tbl:anova}` menunjukkan bahwa hanya variabel pengeluaran rumah tangga yang berpengaruh terhadap akses SUT dimasa depan.

Hasil yang ditampilkan adalah

Tabel 1 menunjukkan bahwa korelasi antar variabel sangat kecil, sehingga regresi linear sederhana dapat diterapkan untuk mendeteksi variabel yang berpengaruh terhadap akses ke usaha tani dimasa depan. Tabel 2 menunjukkan bahwa hanya variabel pengeluaran rumah tangga yang berpengaruh terhadap akses SUT dimasa depan.

Tabel-tabel yang sering digunakan untuk memperjelaskan hasil penelitian dapat dengan mudah dihasilkan dengan menggunakan `knitr::kable`. Hasil analisis dengan R dapat langsung ditampilkan jika kelas (*class*) variabel adalah *matrix* atau *data frame*. Selain pengecekan yang mudah/ minimal saat data berubah, keuntungan lain yang diperoleh adalah kemudahan pembuatan tabel tanpa harus mengubah/ menyesuaikan ukuran isi/ konten dari tabel. *Cross reference* untuk tabel juga dapat dilakukan mirip dengan *cross reference* pada persamaan.

3.3 Teks dan gambar

Gambar biplot (*Figure 1a* (Budiaji, Pancawati J., and Suherna 2017)) dapat dihasilkan ulang untuk contoh teks dan gambar. Gambar ini adalah gambar biplot dari enam kecamatan di Kota Serang dan tujuh komoditas sayuran berdasarkan data sekunder BPS (2012). Sebelum proses pembuatan biplot, *load data* dari data sekunder tersebut diperlukan agar data tersimpan dalam sistem. Analisis komponen utama juga dibutuhkan untuk pembuatan biplot, sehingga kode yang ditulis adalah

```
```{r echo = FALSE, out.width='75%',
```

```

fig.cap = 'Biplot indikator produksi \\label{gbr:biplot}',
fig.align = 'center', fig.pos = 'h'}
data1a <- read.table ('C:/data.txt', sep = '',header=T)
datM <- as.matrix(data1a[,-1])
jumbar <- apply(datM,2,sum)
dataprod <- datM[,jumbar != 0]
rownames(dataprod) <- data1a[,1]
pca <- prcomp(dataprod[,,-1],scale=T)
hasilpca <- summary(pca)
dim1 <- hasilpca$importance[2,1]*100
dim2 <- hasilpca$importance[2,2]*100
biplot(pca, xlim = c(min(as.numeric(hasilpca$rotation[,1:2]))-0.5,
 max(as.numeric(hasilpca$rotation[,1:2]))+0.5),
 xlab = paste('Dimension 1 (', round(dim1,1),'%', sep=''),
 ylab = paste('Dimension 1 (', round(dim2,1),'%', sep=''))
...

```

Perhatikan bahwa data yang akan digunakan yaitu `data.txt` telah disimpan dalam folder `C:/`. Blok kode diantara dua tanda kurung kurawal (`{...}`) terdiri atas tiga baris (lihat tiga baris pertama). Selama tidak menekan **enter** blok kode tersebut dapat dijalankan yang ditandai dengan *highlight* dilatar belakang, tetapi blok kode hanya akan dianggap teks jika latar belakangnya tidak ada *highlight* (gambar biplot tidak dihasilkan). Biplot yang dihasilkan dengan blok kode tersebut adalah

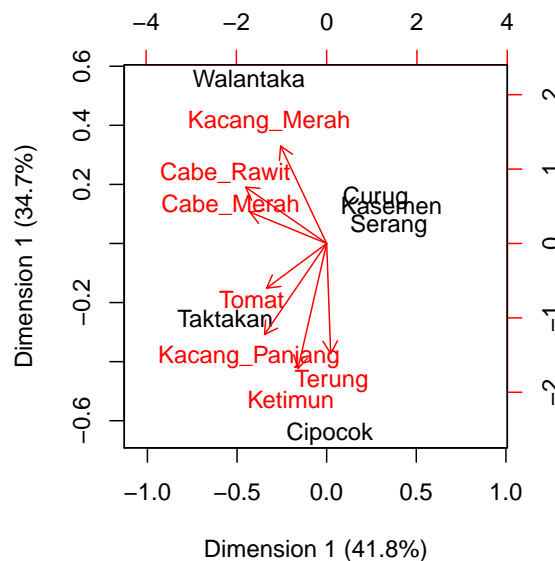


Figure 1: Biplot indikator produksi

Saat gambar ditampilkan, *cross reference* dari tiap gambar dapat juga digunakan pada teks seperti *cross reference* pada tabel, jika di Rstudio ditulis

Gambar \ref{gbr:biplot} menunjukkan jenis komoditas yang berpotensi menjadi komoditas unggulan berdasarkan produksi komoditas di setiap kecamatan. Terung dan ketimun menjadi komoditas unggulan di Kecamatan Cipocok, tomat dan kacang panjang di Kecamatan Taktakan, cabe merah dan cabe rawit di Kecamatan Walantaka. Tiga kecamatan yaitu Kecamatan Curug, Kasemen, dan Serang belum memiliki komoditas unggulan. Keragaman total yang dapat diterangkan oleh biplot sebesar  $\text{round}(\text{dim1}+\text{dim2},1)\%$ , sehingga biplot pada Gambar \ref{gbr:biplot} dapat mendeskripsikan hubungan antara komoditas dan kecamatan pada dua dimensi dengan baik.

Hasil yang ditampilkan adalah

Gambar 1 menunjukkan jenis komoditas yang berpotensi menjadi komoditas unggulan berdasarkan produksi komoditas di setiap kecamatan. Terung dan ketimun menjadi komoditas unggulan di Kecamatan Cipocok, tomat dan kacang panjang di Kecamatan Taktakan, cabe merah dan cabe rawit di Kecamatan Walantaka. Tiga kecamatan yaitu Kecamatan Curug, Kasemen, dan Serang belum memiliki komoditas unggulan. Keragaman total yang dapat diterangkan oleh biplot sebesar 76.5%, sehingga biplot pada Gambar 1 dapat mendeskripsikan hubungan antara komoditas dan kecamatan pada dua dimensi dengan baik.

Seorang peneliti seharusnya mengadopsi *reproducible research* (data dan kode tersedia) sehingga kontribusi terhadap keilmuan lebih efektif (Ivie and Thain 2018) karena mendapatkan dukungan yang cukup baik dan proses yang mudah dengan R, Rstudio, dan paket knitr. Penelitian melibatkan banyak jenis data yang berbeda-beda dan banyak peneliti/ kontributor memerlukan skema alur kerja secara lengkap dari pemrosesan data mentah hingga analisis (Finak et al. 2018). Tujuannya agar data dan kode yang beragam (karena dari banyak kontributor) dapat dengan mudah disebarkan dan dipublikasikan untuk kepentingan RR.

## 4 Kesimpulan

Artikel ini memberikan contoh RR dengan menggunakan Rstudio, software R dan paket knitr. Data dan kode yang digunakan dalam artikel ini dapat diunduh dan dijalankan ulang pada Rstudio sebagai syarat standar sebuah RR. Hasil reproduksi ulang pada sebagian hasil penelitian dalam sebuah artikel dapat dilakukan dengan baik dan berhasil menggunakan Rstudio, software R dan paket knitr. Keunggulan penggunaan Rstudio, software R dan paket knitr yaitu teks yang dapat disisipi kode *inline*, *cross reference* yang mudah terutama pada tabel dan gambar yang sering digunakan untuk menampilkan hasil penelitian, hasil akhir berupa teks, tabel dan gambar yang dapat menyesuaikan jika terjadi perubahan data, dan peneliti lain yang ingin menggunakan kode lebih mudah sehingga RR terjamin. Peneliti yang melaporkan hasil penelitiannya seharusnya menggunakan konsep RR agar peneliti lain dapat menggunakan hasil penelitian tersebut untuk dijalankan ulang atau untuk pengembangan keilmuan lebih lanjut.

## Daftar Pustaka

- Biecek, P., and M. Kosinski. 2017. “Archivist: An R Package for Managing, Recording and Restoring Data Analysis Results.” *Journal of Statistical Software* 82. <https://doi.org/10.18637/jss.v082.i11>.
- BPS. 2012. *Kota Serang Dalam Angka 2011*. Badan Pusat Statistik Kota Serang.
- Budiaji, W., Pancawati J., and Suhera. 2017. “Vegetables Mapping Using Production and Socioeconomic Indicators Approach.” In *ICoSI 2014 Proceedings of the 2nd International Conference on Sustainable Innovation*, edited by Taufik T. et al., 109–18. Springer, Singapore. [https://doi.org/10.1007/978-981-287-661-4\\_13](https://doi.org/10.1007/978-981-287-661-4_13).
- Finak, G., B. Mayer, W. Fulp, P. Obrecht, A. Sato, E. Chung, D. Holman, and R. Gottardo. 2018. “DataPackageR: Reproducible Data Preprocessing, Standardization and Sharing Using R/Bioconductor for Collaborative Data Analysis [Version 2; Referees: 2 Approved, 1 Approved with Reservations].” *Gates Open Research* 2 (31). <https://doi.org/10.12688/gatesopenres.12832.2>.
- Gandrud, C. 2015. *Reproducible Research with R and RStudio*. 2nd ed. Chapman; Hall.
- Ivie, P., and D. Thain. 2018. “Reproducibility in Scientific Computing.” *ACM Computing Surveys* 51 (3). New York, NY, USA: ACM: 63:1–63:36. <https://doi.org/10.1145/3186266>.
- King, G., Keohane R., and S. Verba. 1994. *Designing Social Inquiry*. Princeton University Press.
- Leisch, F. 2002. “Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis.” In *Proceedings in Computational Statistics*, edited by W. Härdle and B. Rönz, 575–80. Compstat 2002. Physika Verlag, Heidelberg, Germany.
- Peng, R. D. 2011. “Reproducible Research in Computational Science.” *Science*, 1226–27.
- R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Stodden, V., J. Seiler, and Z. Ma. 2018. “An Empirical Analysis of Journal Policy Effectiveness for Computational Reproducibility.” *Proceedings of the National Academy of Sciences* 115 (11). National Academy of Sciences: 2584–89. <https://doi.org/10.1073/pnas.1708290115>.
- Stodden, V., Leisch F., and R. D Peng. 2014. *Implementing Reproducible Research*. Chapman; Hall.
- Yihui, X. 2015. *Dynamic Documents with R and Knitr*. Chapman; Hall.
- . 2017. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*.