

**PEDOMAN PENGGUNAAN GUI-R PADA ANALISIS REGRESI  
PARAMETRIK DAN NONPARAMETRIK DENGAN MENGGUNAKAN  
R-SHINY**

Diajukan Guna Memenuhi Ujian Akhir Semester (UAS)  
Mata Kuliah Komputasi Statistika Lanjut

Dosen Pengampu : Dr. Budi Warsito, S.Si, M.Si.



---

---

**MODUL**

---

---

**Disusun Oleh:**

Haasya Wafdayanti

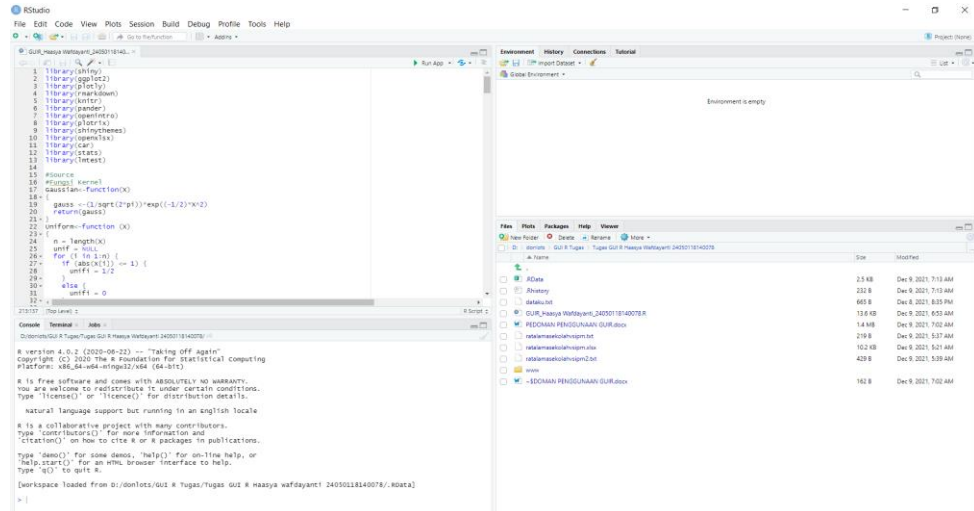
24050118140078

**DEPARTEMEN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
2021**

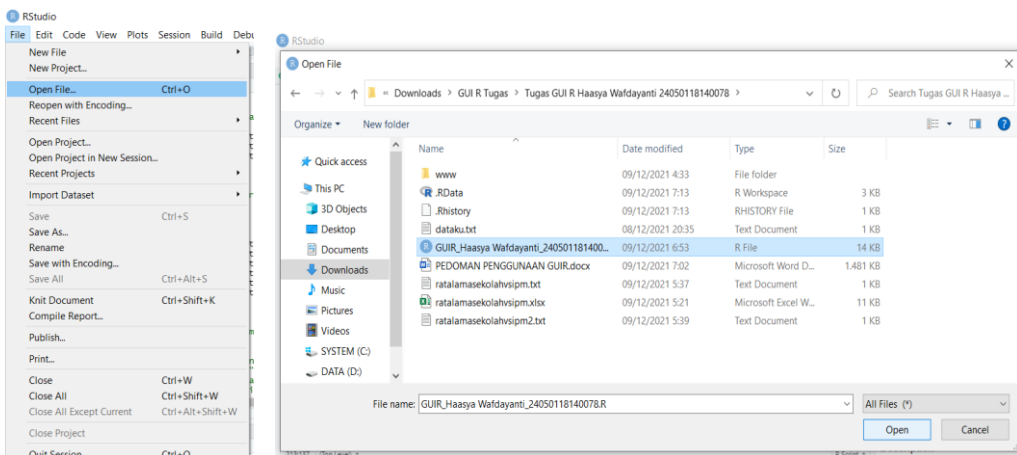
## PEDOMAN PENGGUNAAN GUI-R

1. Buka aplikasi R Studio yang sudah terinstall pada komputer. Untuk aplikasi R Studio dapat didownload di:

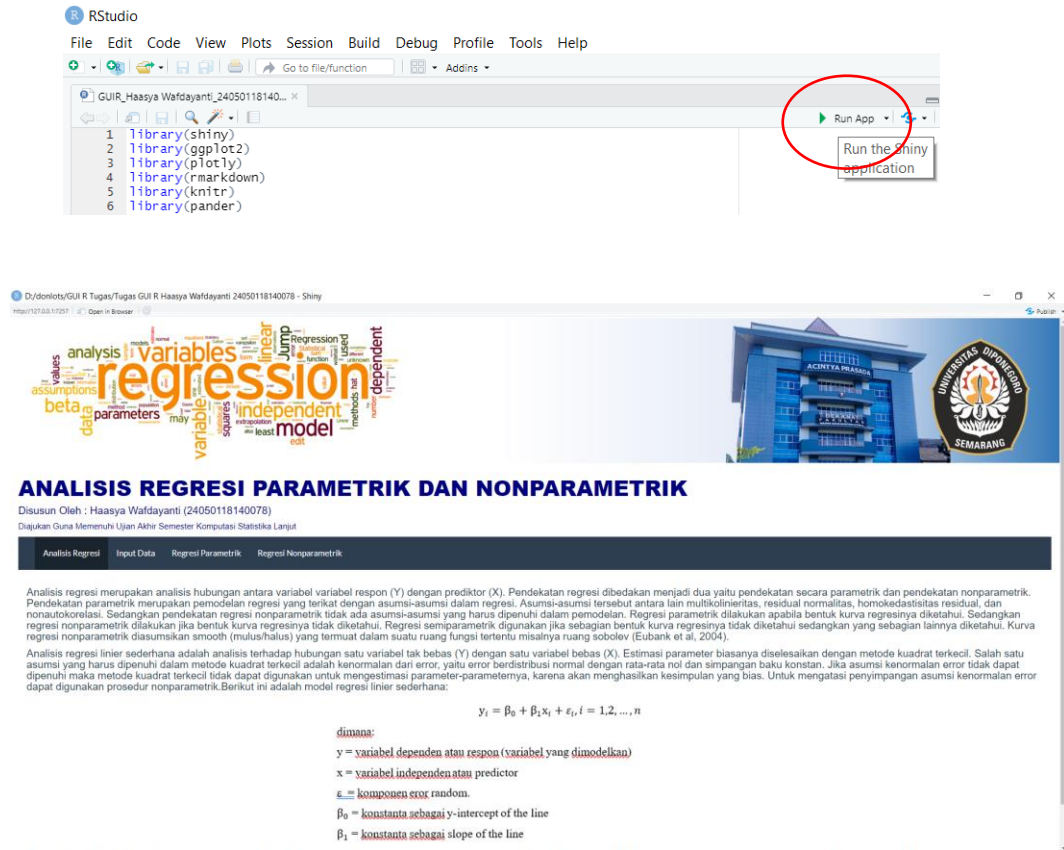
<https://www.rstudio.com/products/rstudio/download/#download>



2. Setelah aplikasi terbuka, klik menu *File* lalu klik *Open File* atau secara langsung dapat menggunakan *keyboard* dengan menekan Ctrl+O. Kemudian, buka file GUI-R yang sudah disiapkan sebelumnya sesuai dengan lokasi penyimpanan yaitu dengan membuka D:\donlots\GUI R Tugas\Tugas GUI R Haasya Wafdayanti 24050118140078 klik pada *file* GUIR\_Haasya Wafdayanti\_24050118140078 kemudian klik *Open*. Selain itu dapat pula dibuka dengan cara me-*double* klik file GUI pada folder penyimpanan yang sudah dibuka sebelumnya.



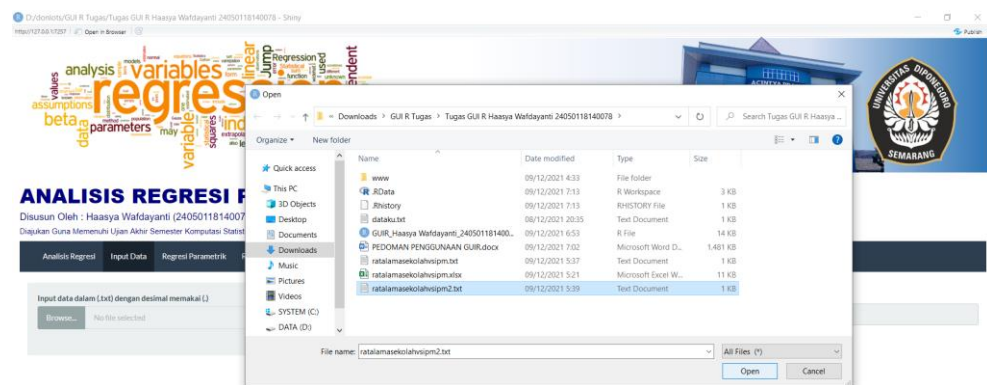
3. Setelah GUI berhasil terbuka, langkah selanjutnya adalah menjalankan GUI dengan meng-klik Run App, sehingga akan muncul tampilan awal GUI sebagai berikut :



4. Pada tampilan awal GUI, akan muncul 4 TabPanel yaitu “Analisis Regresi”, “Input Data”, “Regresi Parametrik”, dan “Regresi Nonparametrik”. Pada TabPanel “Analisis Regresi” akan berisi materi singkat mengenai analisis regresi parametrik dan nonparametrik serta metode yang akan digunakan sehingga pembaca dapat mengetahui mater dasar mengenai metode yang akan diangkat. Kemudian pada TabPanel “Input Data” merupakan sebuah wadah untuk mengimpor data yang akan dianalisis, selain itu juga pada bagian ini berisi statistik deskriptif tentang data kemudian persebaran data yang disajikan melalui *scatterplot*. Lalu, pada bagian “Regresi Parametrik” akan berisi output analisis regresi parametrik dengan metode regresi linier sederhana dimana dapat diketahui output dari model yang dihasilkan, nilai koefisien korelasi, serta uji asumsi. Kemudian, untuk bagian “Regresi Nonparametrik” akan dibatasi dengan menggunakan metode regresi kernel

dengan fungsi kernel yang digunakan adalah kernel gauss dan uniform yang akan menghasilkan output bandwidth optimum dan MSE dengan menggunakan CV dan GCV.

- Pertama-tama untuk memulai analisis pada data, kita akan melakukan *import* data dengan mengklik *Browse* kemudian buka *file* data yang akan dianalisis dalam format **\*.txt** dengan **pemisah desimal menggunakan tanda titik (.)** kemudian klik *Open*. Pada pedoman kali ini akan digunakan data pengaruh rata-rata lama sekolah terhadap IPM seluruh provinsi di Indonesia pada tahun 2020.

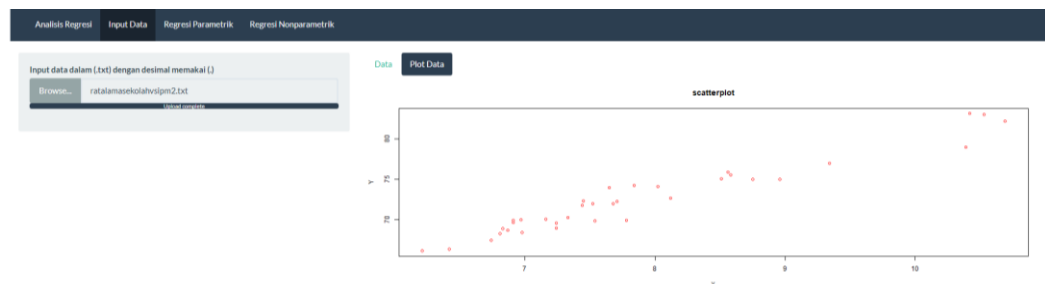
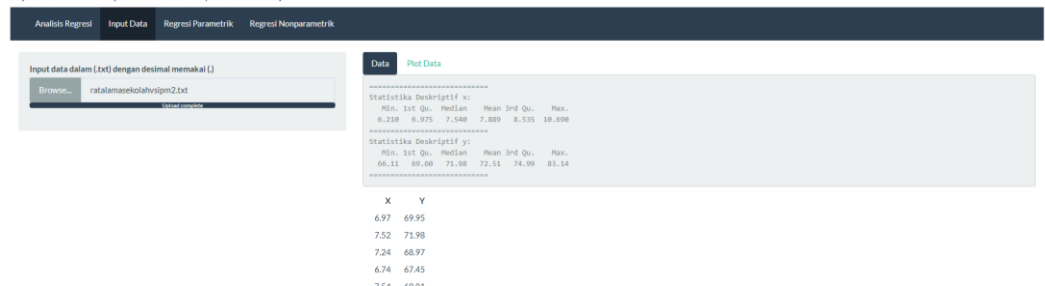


- Setelah upload data berhasil, maka pada akan keluar output yang berupa statistik deskriptif dari masing-masing variabel X dan Y kemudian akan muncul pula output *scatterplot* yang dapat dilihat dengan meng-klik bagian “Plot Data”.

## ANALISIS REGRESI PARAMETRIK DAN NONPARAMETRIK

Diusun Oleh : Haasya Wafdayanti (24050118140078)

Diajukan Guna Memenuhi Ujian Akhir Semester Komputasi Statistika Lanjut



7. Kemudian akan dilanjutkan pada tahap analisis regresi parametrik, dengan cara meng-klik TabPanel bagian “Regresi Parametrik”.

### ANALISIS REGRESI PARAMETRIK DAN NONPARAMETRIK

Disusun Oleh : Haasya Waldayanti (24050118140078)

Diajukan Guna Memenuhi Ujian Akhir Semester Komputasi Statistika Lanjut

Analisis RegresiInput DataRegresi ParametrikRegresi Nonparametrik

Analisis regresi linier sederhana haruslah memenuhi asumsi normalitas, homoskedastisitas, non autokorelasi

Regresi Linier SederhanaKorelasiUji NormalitasUji HomoskedastisitasUji Non Autokorelasi

Regresi Linier Sederhana  
-----  
Call:  
lm(formula = y ~ x)  
-----  
Residuals:  
Min 1Q Median 3Q Max  
-2.4912 -0.7004 0.1179 0.9256 2.2952  
-----  
Coefficients:  
Estimate Std. Error t value Pr(>|t|)  
(Intercept) 44.2347 1.3857 34.48 <2e-16 \*\*\*  
x 3.5848 0.1612 22.24 <2e-16 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
-----  
Residual standard error: 1.123 on 33 degrees of freedom  
Multiple R-squared: 0.9374, Adjusted R-squared: 0.9356  
F-statistic: 494.6 on 1 and 33 Df, p-value: < 2.2e-16  
-----

Pada bagian ini akan diperoleh output untuk model yang telah dihasilkan, nilai koefisien korelasi, dan juga output dari uji asumsi yang harus dipenuhi yang terdiri dari uji normalitas, uji homoskedastisitas, dan uji non autokorelasi. Selain itu, pada bagian ini kita juga dapat mengetahui nilai koefisien korelasi dengan menggunakan korelasi Pearson.

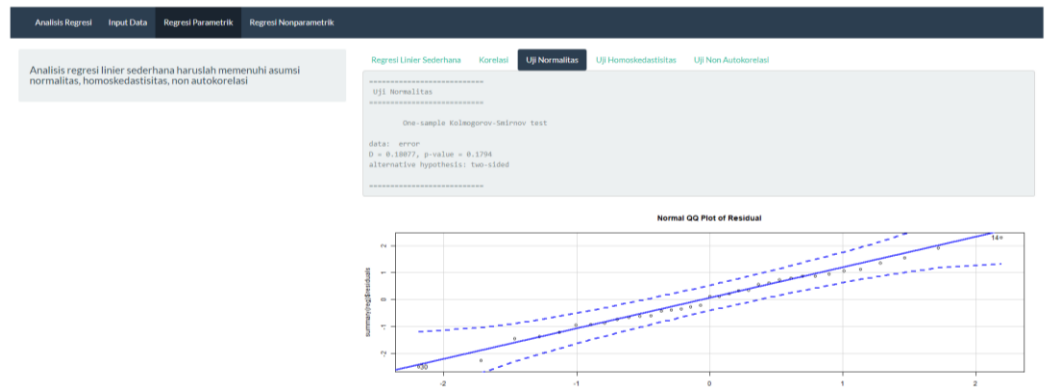
Analisis RegresiInput DataRegresi ParametrikRegresi Nonparametrik

Analisis regresi linier sederhana haruslah memenuhi asumsi normalitas, homoskedastisitas, non autokorelasi

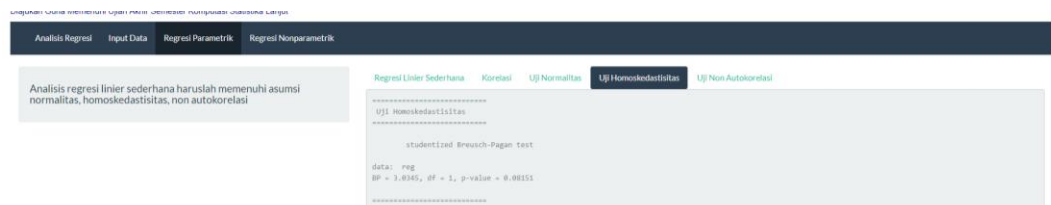
Regresi Linier SederhanaKorelasiUji NormalitasUji HomoskedastisitasUji Non Autokorelasi

Korelasi Pearson  
-----  
Pearson's product-moment correlation  
-----  
data: x and y  
t = 22.239, df = 33, p-value < 2.2e-16  
alternative hypothesis: true correlation is not equal to 0  
95 percent confidence interval:  
0.9374431 0.9839789  
sample estimates:  
cor  
0.9682185  
-----

Kemudian untuk uji asumsi normalitas residual diperoleh output dengan menggunakan uji Kolmogorov-Smirnov untuk pengujian secara formal serta *Normal Q-Q Plot of Residual* untuk pengujian secara visual yang mana dari output yang dihasilkan dapat diketahui apakah residual data berdistribusi secara normal atau tidak



Selanjutnya diperoleh output dari uji homoskedastisitas dengan menggunakan uji *Breusch Pagan Test* untuk melihat apakah varian dari data yang diuji konstan atau tidak.



Untuk pengujian selanjutnya, terdapat uji non autokorelasi dengan menggunakan nilai Durbin Watson dimana dari output yang dihasilkan dapat diketahui apakah terjadi autokorelasi atau tidak.



8. Sebagai salah satu analisis regresi parametrik, regresi linier sederhana harus memenuhi semua asumsi-asumsi yang ditetapkan sebelumnya, terutama asumsi distribusi yaitu asumsi residual berdistribusi secara normal. Apabila dari interpretasi output diperoleh informasi bahwa residual data tidak berdistribusi secara normal, maka data akan dianalisis menggunakan metode nonparametrik salah satunya adalah regresi kernel dengan menggunakan fungsi kernel gauss, uniform, epanechnikov, dan segitiga. Namun, dalam modul ini hanya akan digunakan fungsi kernel gauss dan juga uniform.
9. Kemudian untuk langkah selanjutnya, klik tabPanel bagian “ Regresi Nonparametrik”. Pada regresi nonparametrik dengan menggunakan regresi kernel, akan dicari bandwidth optimal dengan metode CV dan GCV, selain itu

juga akan dicari MSE dari model dengan ketentuan bandwidth yang berbeda untuk dicari model dengan nilai MSE terkecil sehingga dapat diperoleh model terbaik. Untuk langkah selanjutnya, pilih jenis fungsi kernel yang akan digunakan. Sebagai contoh akan digunakan fungsi kernel gauss dengan bandwidth minimum 0.1 dan bandwidth optimum 1 dengan penambahan bandwidth yang dicobakan 0.1. Kemudian klik “Run”.

Diajukan Guna Memenuhi Ujian Akhir Semester Komputasi Statistika Lanjut

Analisis Regresi   Input Data   Regresi Parametrik   Regresi Nonparametrik

Pencarian Bandwidth Optimal

Pilih Kernel

Gaussian

Gaussian

Uniform

U.1

Bandwidth Maksimum

1

Penambahan Bandwidth yang Dicobakan

0.1

Run

Bandwidth Optimal dengan CV   Bandwidth Optimal dengan GCV

Kemudian akan diperoleh output untuk CV dan GCV sebagai berikut :

Analisis Regresi   Input Data   Regresi Parametrik   Regresi Nonparametrik

Pencarian Bandwidth Optimal

Pilih Kernel

Gaussian

Bandwidth Minimum

0.1

Bandwidth Maksimum

1

Penambahan Bandwidth yang Dicobakan

0.1

Run

Bandwidth Optimal dengan CV   Bandwidth Optimal dengan GCV

Berikut 10 nilai CV minimal dan bandwidthnya:

```

=====
No h CV
=====
[1,] 0.1 1.841437
[2,] 0.2 1.841437
[3,] 0.3 1.841437
[4,] 0.4 2.810184
[5,] 0.5 2.898835
[6,] 0.6 2.343782
[7,] 0.7 2.528772
[8,] 0.8 2.879338
[9,] 0.9 3.295378
[10,] 1.0 4.338725
=====
nilai bandwidth optimal= 0.2
dengan nilai CV minimal= 1.841437

```

Diajukan Guna Memenuhi Ujian Akhir Semester Komputasi Statistika Lanjut

Analisis Regresi   Input Data   Regresi Parametrik   Regresi Nonparametrik

Pencarian Bandwidth Optimal

Pilih Kernel

Gaussian

Bandwidth Minimum

0.1

Bandwidth Maksimum

1

Penambahan Bandwidth yang Dicobakan

0.1

Run

Bandwidth Optimal dengan CV   Bandwidth Optimal dengan GCV

h.opt= 0.1 dengan GCV= 2.385777 dan MSE minimal= 0.8818151

Berikut 10 nilai GCV dan MSE terkecil beserta nilai bandwidth h:

```

=====
No h GCV MSE
=====
[1,] 0.1 2.385777 0.8818151
[2,] 0.2 1.881889 1.0708093
[3,] 0.3 1.716634 1.2298583
[4,] 0.4 1.985138 1.3452987
[5,] 0.5 2.138392 1.7042951
[6,] 0.6 2.448877 2.0043227
[7,] 0.7 2.786280 2.1082133
[8,] 0.8 3.288298 2.7373712
[9,] 0.9 3.638242 3.1737125
[10,] 1.0 4.183732 3.6746127

```

Dengan cara yang sama, dilakukan pula pada jenis fungsi kernel uniform :

Analisis Regresi

Input Data

Regresi Parametrik

Regresi Nonparametrik

Pencarian Bandwidth Optimal

Pilih Kernel

Uniform

Bandwidth Minimum

0.1

Bandwidth Maksimum

1

Penambahan Bandwidth yang Dicobakan

0.1

Run

Bandwidth Optimal dengan CV

Bandwidth Optimal dengan GCV

Berikut 10 nilai CV minimal dan bandwidthnya:

```

=====
No h CV
=====
[1,] [2,]
[1,] 0.4 1.041620
[2,] 0.5 1.922269
[3,] 0.6 1.956484
[4,] 0.7 2.083870
[5,] 0.8 2.421189
[6,] 0.9 2.556996
[7,] 1.0 2.693138
[8,] 0.3 1.70395684
[9,] 0.2 1.92326851
[10,] 0.1 1.1333495471
=====
nilai bandwidth optimal= 0.4
dengan nilai CV minimal= 1.04162

```

Analisis Regresi

Input Data

Regresi Parametrik

Regresi Nonparametrik

Pencarian Bandwidth Optimal

Pilih Kernel

Uniform

Bandwidth Minimum

0.1

Bandwidth Maksimum

1

Penambahan Bandwidth yang Dicobakan

0.1

Run

Bandwidth Optimal dengan CV

Bandwidth Optimal dengan GCV

h opt= 0.1 dengan GCV= 2.788294 dan MSE minimal= 0.8527299

Berikut 10 nilai GCV dan MSE terkecil beserta nilai bandwidth h:

```

=====
No h GCV MSE
=====
[1,] [2,] [3,]
[1,] 0.1 2.788294 0.8527299
[2,] 0.2 2.233190 1.0777298
[3,] 0.4 1.923640 1.1803080
[4,] 0.3 1.829732 1.1478251
[5,] 0.5 1.838880 1.3752249
[6,] 0.6 1.827070 1.4427851
[7,] 0.7 1.849508 1.5669929
[8,] 0.8 2.319777 1.9242877
[9,] 0.9 2.455424 2.0617807
[10,] 1.0 2.593177 2.2085524

```

Melalui perbandingan pemodelan secara parametrik maupun non parametrik, kita dapat mengetahui lebih luas analisis pada data dengan karakter yang beragam serta bagaimana model terbaik dari data tersebut. Dengan menggunakan GUI-R khususnya R-Shiny sangat membantu dalam melakukan analisis yang tepat dengan tampilan sederhana yang mudah untuk dipahami. Oleh karena itu, sangat perlu untuk dilakukan pengembangan lebih jauh untuk metode yang lebih beragam pada pembuatan GUI-R dengan R-SHINY.