

Analysis Statistics With R Commander

Hendra Perdana, S.Si., M.Sc

Program Studi Statistika Universitas Tanjungpura

hendra.perdana@math.untan.ac.id

0852-6984-2424

R Commander

Interaksi utama antara *user* dengan *R* adalah bersifat ***Command Line Interface*** (CLI). Sehingga perlu penyesuaian bagi *user* yang sudah terbiasa dengan fasilitas *Point and Click Graphical User Interface* (GUI).

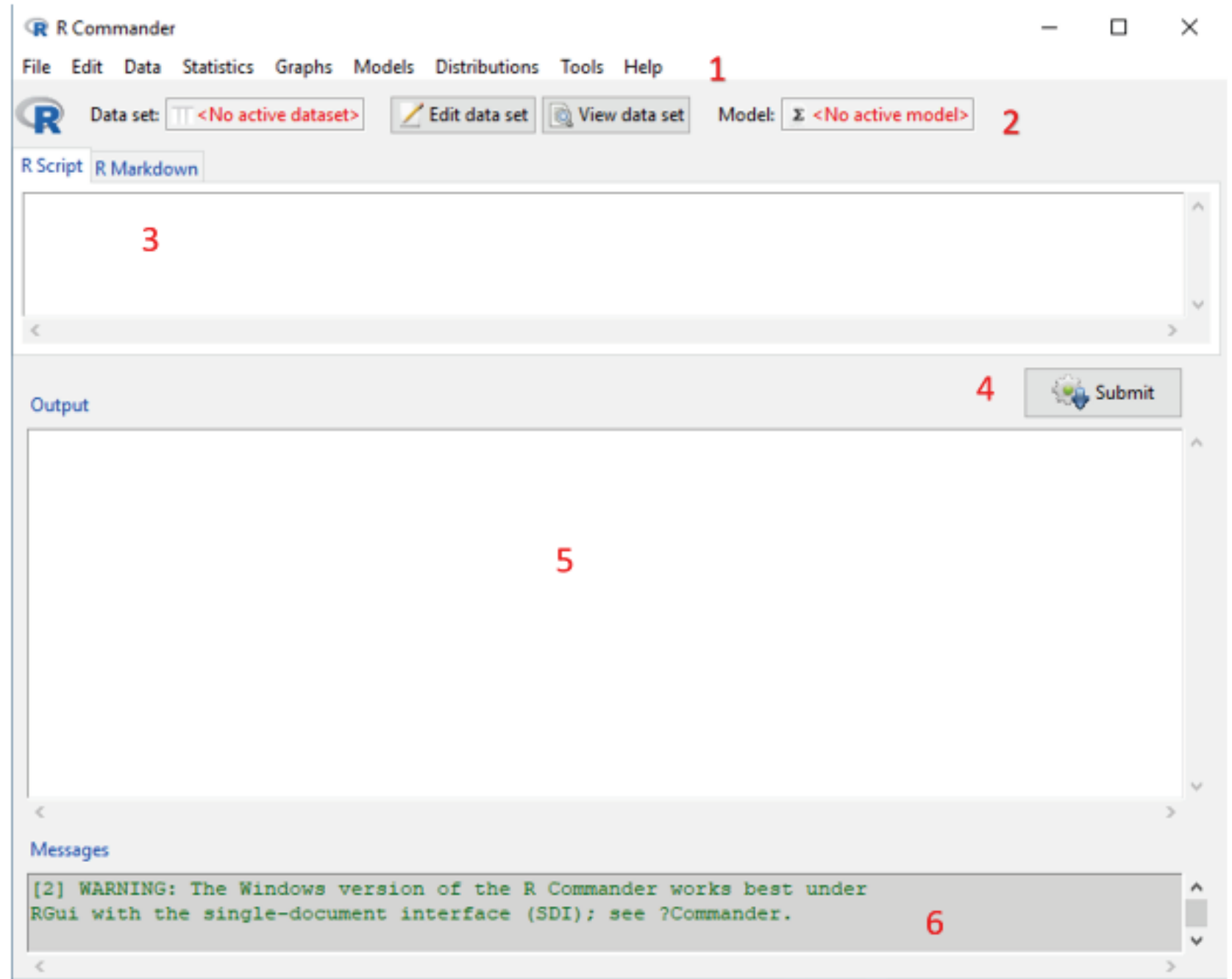
R Commander merupakan paket yang menyediakan tampilan antar muka atau ***Graphical User Interface*** (GUI) yang memungkinkan *user* berinteraksi dengan *R* melalui *Point and Click*.

Sehingga dapat dikatakan bahwa *Rcmdr* sebagai bagian yang '*friendly*' dalam *R*. *User* dapat mengklik berbagai analisis statistika dasar yang telah tersedia, **tanpa** perlu mengetikkan perintah atau sintaks program.

Selain itu, sintaks yang digunakan di tampilkan ke dalam *editor* yang tersedia pada jendela *Rcmdr* secara otomatis. Hal tersebut dapat mempermudah analisis dalam mempelajari kode apa saja yang perlu diketikkan.

TAMPILAN ANTAR MUKA JENDELA DIALOG RCMDR

1. Menu
2. Toolbar
3. Lembar Kerja
4. Tombol Submit
5. Output box
6. Message box



Bagian 1: Menu-menu

- **File:** Menu untuk memuat dan menyimpan file *script*, menyimpan *output* dan ruang kerja *R*; dan keluar dari *RCommander* atau dari *R* dan *RCommander*.
- **Edit:** Menu (*cut, copy, paste, dll.*) Untuk mengedit teks di berbagai *panel* dan *tab*. Mengklik kanan di salah satu *panel* atau *tab* ini juga memunculkan menu *edit* sesuai dengan *tab* yang digunakan.
- **Data:** Menu yang terdiri atas submenu untuk membaca dan mengolah data.
- **Statistics:** Menu yang terdiri atas berbagai *submenu* untuk melakukan berbagai analisis statistik.
- **Graphs:** Menu yang terdiri atas *submenu* yang berisikan berbagai metode visualisasi data.

(*Lanjutan*) Bagian 1

- **Models:** Menu dan submenu yang digunakan untuk memperoleh ringkasan numerik, rentang keyakinan, uji hipotesis, diagnosa, dan grafik dari model statistik, dan menambahkan hasil diagnosa pada *dataset* seperti menambahkan residu atau *error* model pada dataset.
- **Distribution:** Menu yang terdiri atas submenu yang digunakan untuk memperoleh probabilitas kumulatif, densitas probabilitas, kuantil, dan grafik dari distribusi statistika standar.
- **Tools:** Menu yang digunakan untuk mengakses paket *R* (contoh: memuat dataset dari paket lainnya, memuat paket untuk menambahkan metode analisis statistik, dll), untuk memuat paket *Rcmdr plug-in*, mengatur sebagian besar opsi pada *RCommander*.
- **Help:** Menu bantuan yang berguna untuk memperoleh informasi terkait *RCommander* dan paket terkait lainnya.

(*Lanjutan*) Bagian 1

Item pada menu **akan tidak aktif** (tulisan berwarna abu-abu) apabila tidak ada sesuai dengan konteks tertentu.

Misal: tidak ada dataset aktif maka sebagian besar *submenu Statistics* akan tidak aktif.

Contoh lainnya adalah tidak adanya data kategori pada dataset aktif maka submenu tabel kontingensi tidak akan aktif.

Bagian 2: *Toolbar*

- **Data set** : menampilkan nama dataset yang aktif dan memilih dataset yang akan diaktifkan.
- **Edit data set** : digunakan untuk melakukan proses *editing pada dataset* seperti: merubah nilai baris dan kolom, merubah nama kolom, merubah nama baris, dan menambahkan atau menghapus observasi.
- **View data set** : melihat observasi pada dataset aktif.
- **Model** : menampilkan dan memilih model statistik yang telah dibuat.

Bagian 3: Lembar Kerja

- **R Script:** menampilkan *script* perintah yang digunakan untuk menghasilkan *output*. Pembaca dapat melakukan proses *editing* pada *script* tersebut dan menjalankannya kembali untuk menambah kompleksitas pada luaran yang dihasilkan.
- **R Markdown:** membuat dokumentasi dari analisis yang telah dilakukan.

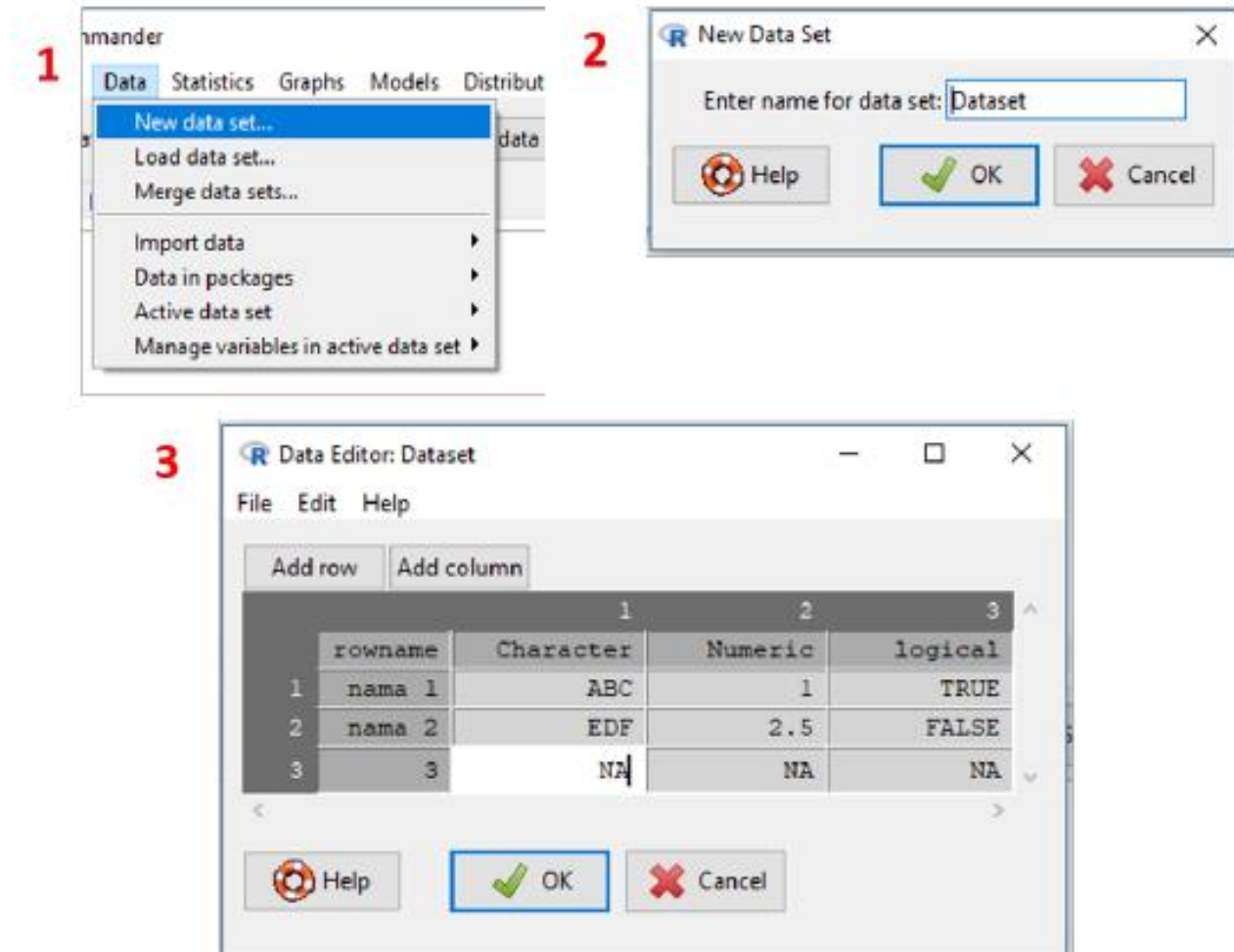
Bagian 4: Tombol *submit*. Untuk menjalankan kembali *R Script* yang telah dibuat, pembaca dapat meng-*highlight script* atau sintaks yang hendak diperoleh kembali hasilnya dan tekan tombol Submit

Bagian 5: *Output box*. Kotak ini berfungsi untuk menampilkan hasil perhitungan berdasarkan sintaks yang dimasukkan atau *disubmit*.

Bagian 6: *Messages box*. Menampilkan sejumlah pesan terkait operasi yang dilakukan. Pesan dapat berupa *error*, *warnings*, dan *note*.

Manajemen Data: Input Data

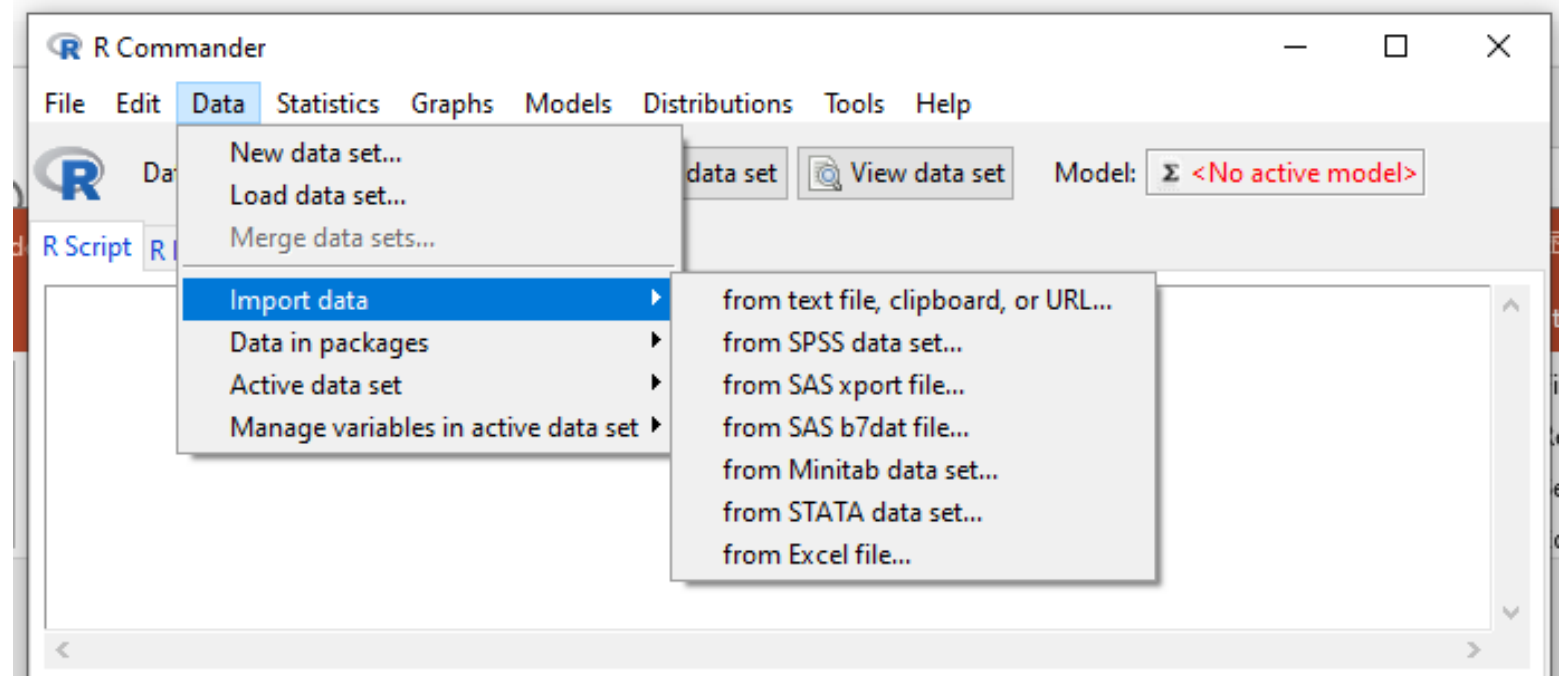
- Input data dapat dilakukan secara langsung pada R Commander. Input data secara langsung umumnya dilakukan jika jumlah data yang kita miliki relatif kecil



Manajemen Data:

Import Data

- Kita akan mempelajari bagaimana cara melakukan *import data* dari berbagai sumber seperti *plain text*, *spreadsheets*, SPSS, STATA, SAS, dan Minitab.



(Lanjutan) Import Data

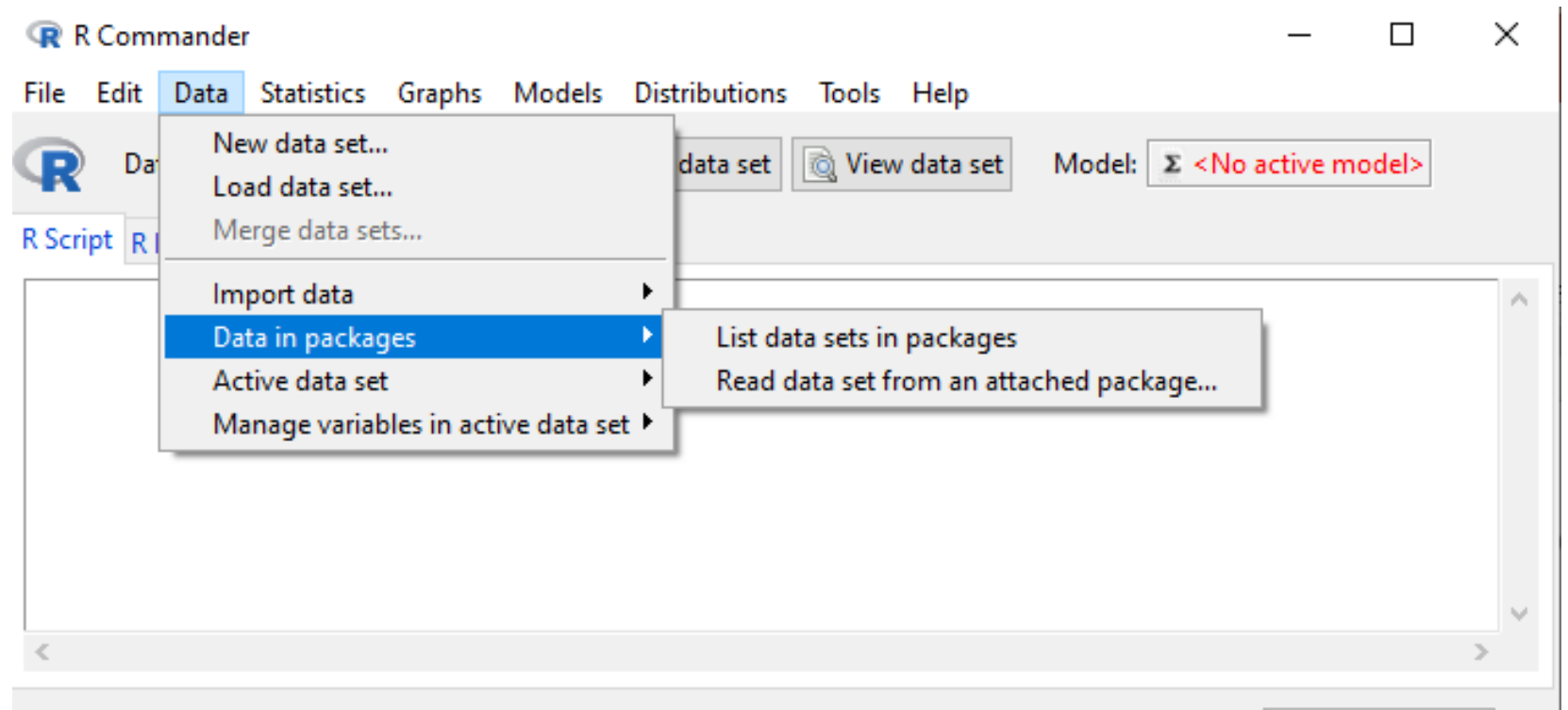
DEMO SIMULASI DENGAN RCommander

1. Import data file text, pasien.txt
2. Import data file SPSS, Nations.sav
3. Import Data file Excel, Datasets.xls
4. Import Data File CSV, GSS.csv

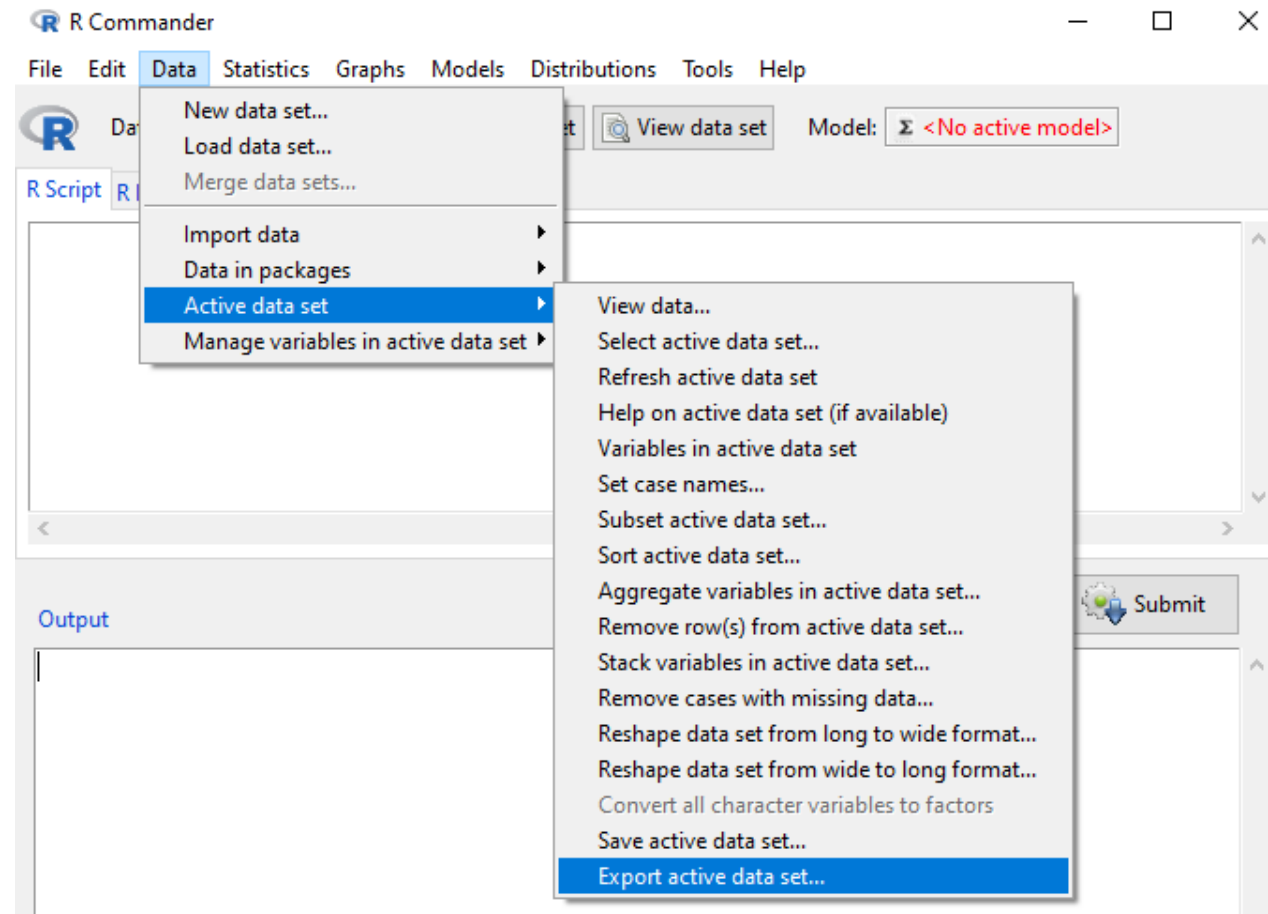
Manajemen Data:

Membaca data dari paket

Jika Anda ingin mengakses *dataset* dari paket yang Anda inginkan, Anda dapat menjalankan menu berikut:

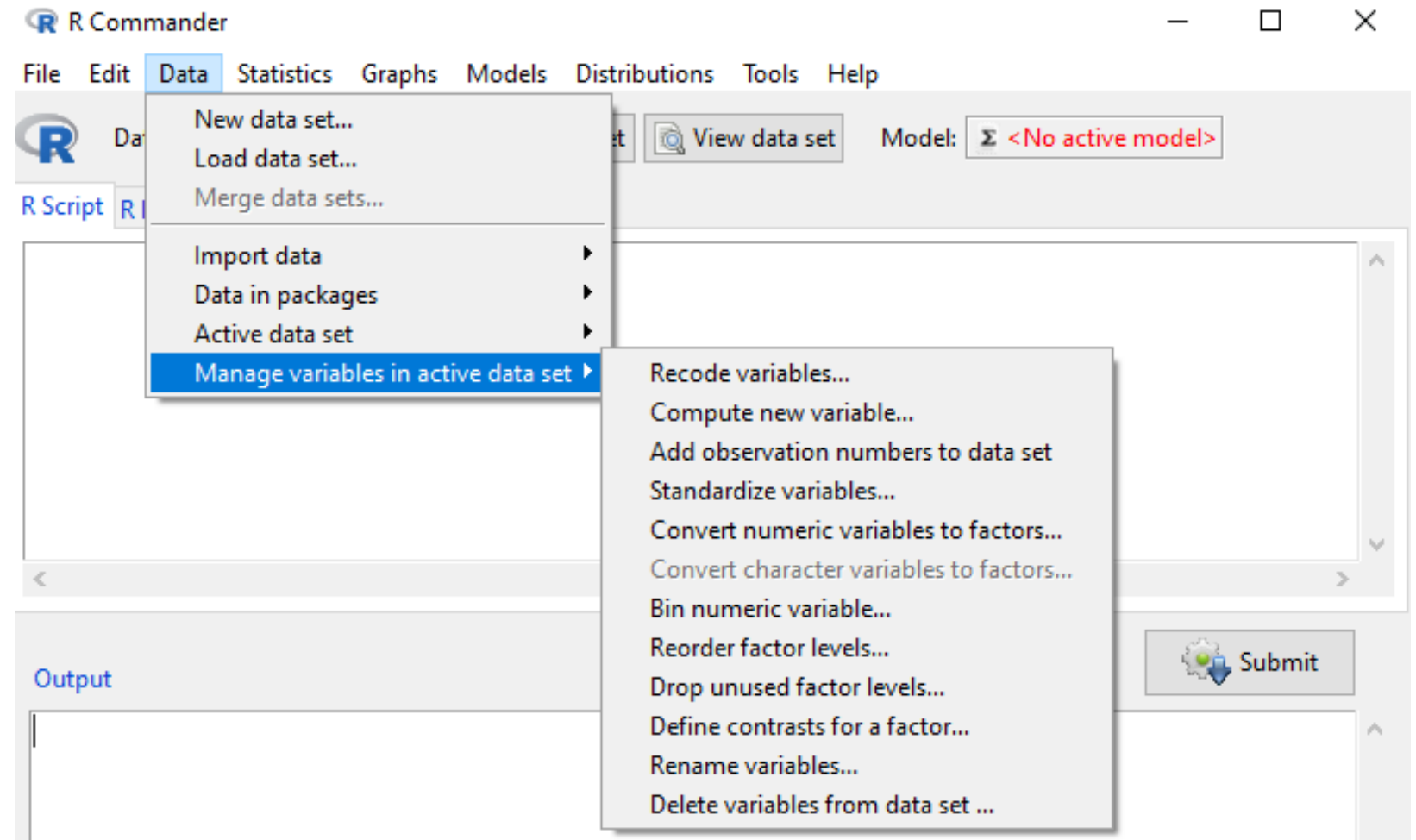


Manajemen Data: *Active Dataset*



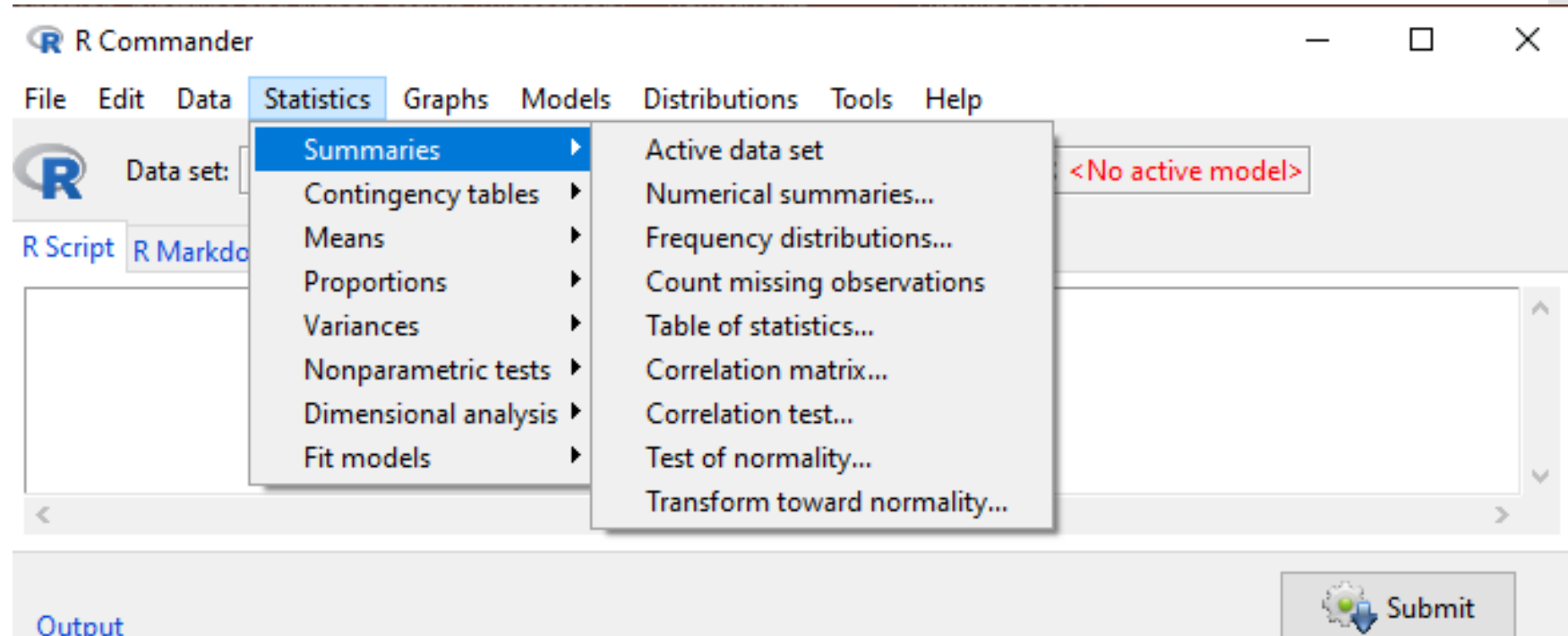
Manajemen Data:

Modifikasi Variabel pada data



Menu Statistics: *Summaries Data*

Menggunakan data dari *package carData* yaitu
AMSSurvey



Menu Rcmdr: Statistics >> Means >> Single Sample t-test

Data:
1. Atlit.txt

1. Pengujian rata-rata sampel terhadap nilai hipotesa yang telah ditentukan

Seorang pelatih lempar lembing melatih delapan atlit lempar lembing untuk mengikuti PORDA DIY. Dalam sebuah seleksi, dicatat kecepatan lemparan kedelapan atlit tersebut

Atlit ke	Kecepatan
1	3005
2	2925
3	2935
4	2965
5	2995
6	3005
7	2935
8	2905

Ujilah hipotesa bahwa rata-rata kecepatan lemparan lembing dari atlit tersebut adalah 3000

Menu Rcmdr:
Statistics >>
Means >> Paired
t-test

Data:
2. data kursus.txt

Sebuah perusahaan ingin meningkatkan kualitas pelayanan karyawannya dengan cara mengirimkan 10 karyawannya mengikuti kursus penambah keterampilan. Untuk melihat hasil dari pelatihan ini, kesepuluh karyawan ini dinilai dengan menggunakan metode tertentu sebelum dan sesudah tes, dan diperoleh data berikut:

Karyawan ke	Sebelum	Sesudah
1	54	60
2	56	59
3	50	57
4	52	56
5	55	56
6	52	58
7	56	62
8	53	55
9	53	54
10	60	64

Ingin dilihat apakah nilai dari karyawan akan naik setelah diberikan kursus (yakni alternatif uji adalah satu sisi, selisih dari nilai sebelum dengan sesudah latihan bernilai positif, hipotesa alternatif dipilih uji satu sisi $D > 0$.)

Data:

3. data padi.txt

Menu Rcmdr:

*Statistics >> Means
>> Independent
sample t-test*

dan

*Statistics>>
variances>>
Two Variance FTest*

3. Uji kesamaan variansi dan uji kesamaan mean dua sampel independen

Dua puluh bidang tanah ditanami padi, dimana 10 bidang menggunakan pupuk buatan baru. Hasil produksinya disajikan pada tabel dibawah ini (dalam ton). Asumsikan data mengikuti distribusi normal

Bidang	1	2	3	4	5	6	7	8	9	10
Pupuk Baru	6.1	5.8	7	5.8	6.1	6.4	6.1	6	5.9	5.8
Biasa	5.9	5.7	6.1	5.9	5.8	5.6	5.6	5.9	5.7	5.6

3.1. Uji kesamaan variansi menggunakan F-test

3.2. Uji mean

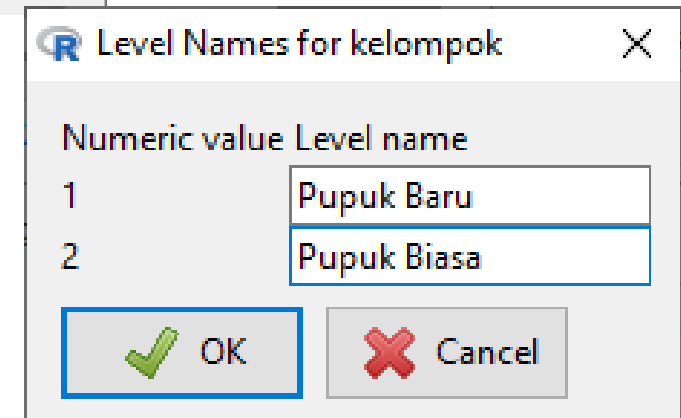
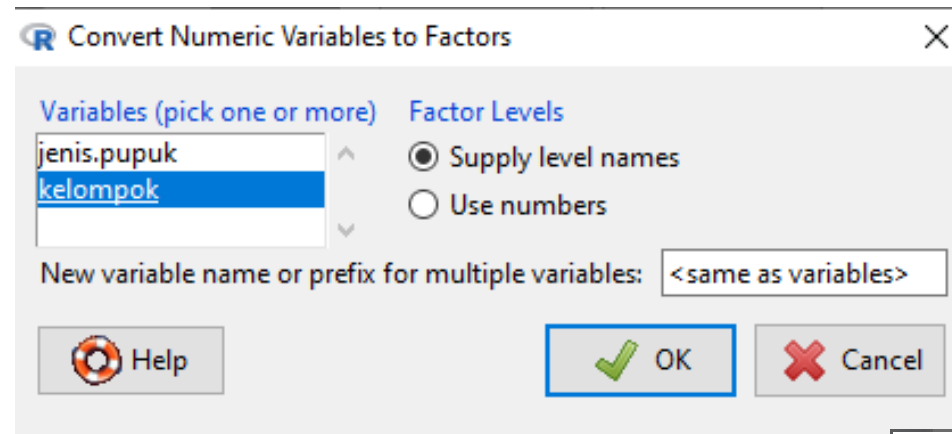
a. alternatif dua sisi dan

b. produksi pupuk jenis baru lebih besar dari pupuk biasa, maka dilakukan uji alternatif satu sisi, yakni dengan alternatif selisih produksi pupuk jenis baru dikurang jenis lama > 0 . Lakukan dengan cara yang sama seperti diatas tetapi sekarang dengan **Alternative hypothesis** bahwa **Difference >0**

(Lanjutan)

3. data padi.txt

- Setelah *Import data*, perlu disesuaikan terlebih dahulu yaitu merubah variable 'kelompok' yang semula bertipe numerik menjadi bertipe factor.
- Menu: *Data>>Manage Variable in active data set >> convert numeric variables to factors*



Menu Rcmdr:
Statistics >>
Means >> One
Way Anova

Data:
4. data
haryatmi.txt

4. Anava satu arahkan

Misalkan saja dimiliki data waktu pembekuan darah (dalam detik) dari 24 binatang yang secara random diberi 4 perlakuan (A,B,C, D) sebagai berikut (data dari Haryatmi, 1986, halaman 6.35):

A	62	60	63	59				
B	63	67	71	64	65	66		
C	68	66	71	67	68	68		
D	56	62	60	61	63	64	63	59

Ingin diketahui apakah perlakuan tersebut berbeda nyata, yakni uji hipotesa nol bahwa rata-rata kecepatan waktu pembekuan darah dari keempat perlakuan adalah sama (tidak ada efek perlakuan) , dengan hipotesa alternatif paling sedikit terdapat dua yang berbeda

(Lanjutan) Anova

MULTIPLE COMPARISON ANALYSIS, TUKEY

```
AnovaModel.1 <- aov(waktu.beku ~ kelompok, data=haryatmi)
TukeyHSD(AnovaModel.1, "kelompok", ordered = TRUE)
plot(TukeyHSD(AnovaModel.1, "kelompok"))
```

Lebih lengkap bila menggunakan package multcomp

Data:

5. data tahan hidup.txt

Menu Rcmdr:

Statistics >>

Means >> Multi

Way Anova

5. Analisa Variansi (Anava) dua arah

Data ini adalah daya tahan hidup (dalam 10 jam) dari suatu jenis binatang serangga terhadap pengaruh 3 macam racun (I,II,III) dan 4 macam perlakuan (A,B,C,D) yang dapat diberikan dalam tabel berikut:

Racun	Perlakuan			
	A	B	C	D
I	0.31	0.82	0.43	0.45
	0.45	1.1	0.45	0.71
	0.46	0.88	0.63	0.66
	0.43	0.72	0.7	0.62
II	0.36	0.92	0.44	0.56
	0.29	0.61	0.35	1.02
	0.4	0.49	0.31	0.71
	0.23	1.24	0.4	0.38
III	0.22	0.3	0.23	0.2
	0.21	0.37	0.25	0.36
	0.18	0.38	0.24	0.31
	0.23	0.29	0.22	0.33

Ingin diketahui apakah terdapat pengaruh dari kedua treatment ini?

Terima Kasih

Sampai jumpa lagi nanti di Webinar R
selanjutnya bulan November