

Modul Praktikum

Statistika Modern dengan R software



Oleh:

Yosie Syadza Kusuma, S.Hut., M.Si

Prof. Dr. Ir. Ashabul Anhar, M.Sc

Zulkarnain, S.Si., M.Si

Rizky Febriana Br. Lubis, M.Si

LEMBAR PENGESAHAN

Nama Mata Kuliah : Praktikum Statistika
Penyusun : 1. Yosie Syadza Kusuma, S.Hut., M.Si
2. Prof. Dr. Ir. Ashabul Anhar, M.Sc
3. Zulkarnain, S.Si., M.Si.
4. Rizky Febriana Br. Lubis, M.Si
Semester : III (Tiga)
Tahun : 2024

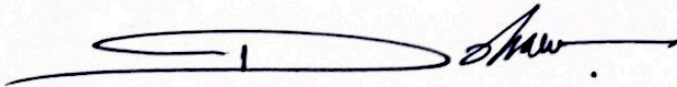
Disahkan pada:

Tanggal : 13 November 2024

Tempat : Banda Aceh

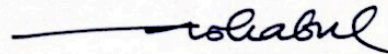
Banda Aceh, 13 November 2024

Mengetahui,
Koordinator Program Studi Kehutanan



Dr. Ir. Dahlan, S.Hut., M.Si., IPU
NIP. 197610062006041003

Koordinator/penanggung jawab



Prof. Dr. Ir. Ashabul Anhar, M.Sc
NIP. 196606291990031002

KATA PENGANTAR

Puji dan syukur kami panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya, sehingga Modul Statistika ini dapat diselesaikan dengan baik. Shalawat serta salam tak lupa kami haturkan kepada junjungan kita, Nabi Muhammad SAW, yang telah membawa umat manusia dari zaman kegelapan menuju cahaya ilmu pengetahuan.

Modul ini disusun sebagai panduan belajar bagi mahasiswa yang tengah mempelajari statistika dasar, baik untuk kebutuhan akademik maupun penerapan praktis di lapangan. Dengan pendekatan yang praktis, modul ini dirancang untuk memberikan pemahaman mendalam mengenai konsep statistik, teknik analisis data, dan penerapannya dalam berbagai bidang. Pembelajaran statistika diharapkan dapat menjadi landasan penting bagi mahasiswa dalam meningkatkan kemampuan analisis data dan mendukung pengambilan keputusan berbasis data yang sangat dibutuhkan di era modern.

Selain itu, modul ini menyertakan penggunaan *software* R sebagai alat bantu analisis data. Penggunaan R diharapkan dapat memperkaya wawasan mahasiswa dalam melakukan analisis data secara mandiri, serta membekali mereka dengan keterampilan yang diperlukan di dunia profesional.

Kami menyadari bahwa modul ini masih memiliki kekurangan, sehingga kami sangat terbuka untuk menerima saran dan masukan dari para pembaca demi penyempurnaan modul ini di masa mendatang.

Semoga modul ini dapat memberikan manfaat dan menjadi referensi yang baik dalam proses belajar mengajar statistika, serta membantu mahasiswa dalam mengembangkan wawasan yang berguna bagi kehidupan profesional mereka.

Banda Aceh, November 2024

Tim Penyusun



DAFTAR ISI

1. PENDAHULUAN.....	4
1.1 LATAR BELAKANG.....	4
1.2 TUJUAN PEMBELAJARAN.....	4
2. PENGANTAR R.....	5
2.1 INSTALASI R	5
2.2 MENJALANKAN PERINTAH R	6
A. Menjalankan perintah melalui R Script.....	6
B. Menjalankan perintah langsung melalui console	7
2.3 VARIABEL DAN FUNGSI	7
A. Variabel	7
B. Fungsi	8
2.4 PACKAGES R	9
3. PENGOLAHAN DATA DASAR	10
3.1 DATA KUALITATIF	10
A. Pengolahan Data Kualitatif.....	10
B. Pengujian Data Kualitatif.....	11
3.2 DATA KUANTITATIF.....	20
A. Pengujian Data Kuantitatif.....	20
B. Analisis dengan R.....	27



1. PENDAHULUAN

1.1 Latar Belakang

Menurut Gasperz (1989), statistika merupakan disiplin ilmu yang berhubungan dengan metode pengumpulan, pengolahan, dan analisis data, yang dilanjutkan dengan penarikan kesimpulan serta pengambilan keputusan yang didasarkan pada fakta-fakta yang relevan. Statistika merupakan disiplin ilmu yang sangat penting dalam pengambilan keputusan berbasis data, baik di bidang akademik maupun di dunia kerja.

Di era digital ini, data tersedia dalam jumlah yang sangat besar dan beragam, mulai dari data survei, data eksperimen, hingga data big data. Oleh karena itu, kemampuan untuk menganalisis dan menginterpretasikan data menjadi keterampilan yang tidak dapat diabaikan. Modul statistika ini dirancang untuk memenuhi kebutuhan tersebut, memberikan pemahaman yang mendalam mengenai konsep-konsep dasar statistika, teknik analisis data, serta aplikasinya dalam berbagai konteks.

Modul ini berfokus pada pengintegrasian perangkat lunak statistik, seperti R, untuk memperkuat pemahaman praktis mahasiswa dalam analisis data. Mengingat pentingnya keterampilan ini di dunia kerja, modul ini bertujuan untuk membekali mahasiswa dengan alat dan teknik yang relevan dan aplikatif untuk analisis statistika berbasis data dalam studi dan karier mahasiswa kedepannya. R adalah perangkat lunak *open-source* yang populer dalam bidang statistika karena fleksibilitas dan kapasitasnya untuk menangani data kompleks. Di samping itu, R didukung oleh komunitas pengguna dan pengembang yang luas, sehingga banyak tersedia paket (*packages*) yang terus diperbarui untuk berbagai kebutuhan analisis.

1.2 Tujuan Pembelajaran

Setelah mempelajari modul ini, mahasiswa diharapkan dapat:

- Memahami konsep dasar dalam statistika, seperti distribusi data, ukuran pemusatan, dan penyebaran data.
- Mengoperasikan perangkat lunak R untuk melakukan analisis statistika dasar.



- c. Menyusun dan menginterpretasi hasil analisis data untuk mendukung penelitian atau kajian yang lebih lanjut.

2. PENGANTAR R

2.1 Instalasi R

R dapat di unduh dengan mengunjungi situs <https://cran.r-project.org/mirrors.html> atau khususnya Indonesia dapat di unduh pada situs <https://cran.usk.ac.id/>. Tampilan pengunduhan R ditampilkan gambar 1.

Download and Install R

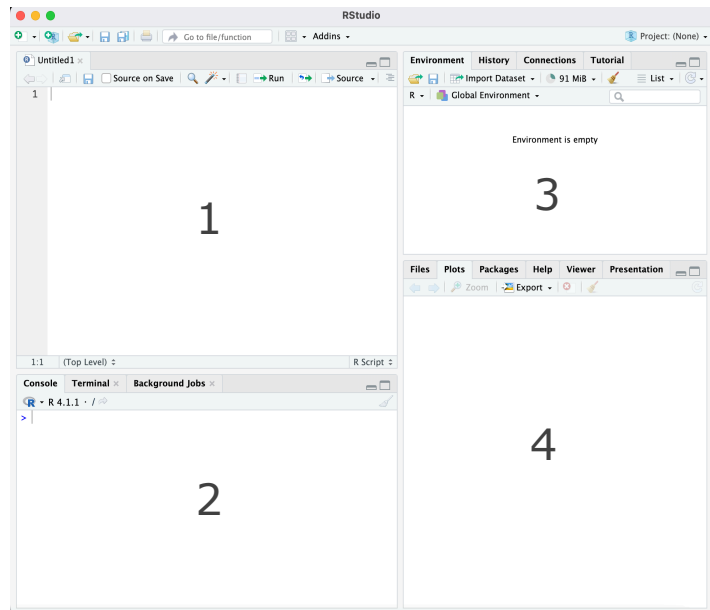
Precompiled binary distributions of the base system and contributed packages,
Windows and Mac users most likely want one of these versions of R:

- [Download R for Linux](#) ([Debian](#), [Fedora/Redhat](#), [Ubuntu](#))
- [Download R for macOS](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Kemudian pengunduhan Rstudio dapat dilakukan dengan mengunjungi situs <https://posit.co/download/rstudio-desktop/> - [download](#). RStudio adalah sebuah *user interface* yang memudahkan pengguna untuk bekerja dengan bahasa pemrograman R. Rstudio menyediakan fitur, seperti editor kode yang memudahkan penulisan dan pengeditan kode, konsol (*console*) untuk menjalankan perintah langsung, serta tampilan untuk melihat variabel yang sedang digunakan. RStudio juga memungkinkan pengguna membuat grafik dan visualisasi data dengan mudah, serta memudahkan mengelola proyek untuk mengatur file.

1. Panel script untuk untuk menulis, mengedit, dan menyimpan skrip R. Skrip dapat disimpan dalam file dengan ekstensi .R
2. Panel console, empat utama untuk mengeksekusi perintah R. Anda dapat memasukkan kode R secara langsung, menjalankan skrip, dan melihat output dari perintah yang dijalankan.
3. Panel Environment, menampilkan semua objek yang ada di dalam lingkungan kerja R, seperti variabel, fungsi, dan data frame.
4. Panel Files, Plots, dan Help digunakan untuk menavigasi dan mengelola file di direktori kerja, menampilkan grafik yang telah dihasilkan, dan memberikan akses ke dokumentasi dan bantuan terkait fungsi atau paket di R



Tampilan RStudio

2.2 Menjalankan Perintah R

A. Menjalankan perintah melalui R Script

1. Bukalah RStudio, buatlah skrip baru dengan cara klik **File > New File > R Script**. Kemudian di editor skrip cobalah ketik kode berikut:

```
# Penjumlahan sederhana
```

```
hasil <- 1 + 1
```

```
# Cetak hasil
```

```
print(hasil)
```

2. Jalankan kode dengan cara tekan **Ctrl + Enter** pada Windows atau **Cmd + Enter** pada macOS
3. Output akan muncul di panel Console di bagian bawah RStudio

```
[1] 2
```

B. Menjalankan perintah langsung melalui console

1. Ketik perintah berikut langsung pada console

```
print(1 + 1)
```

2. Output akan muncul di panel Console di bagian bawah RStudio

```
[1] 2
```

2.3 Variabel dan Fungsi

A. Variabel

Penggunaan variabel dalam R adalah sebagai berikut:

1. Memberi nama variabel, dapat dilakukan dengan menggunakan tanda `<-` atau `=`. Contohnya sebagai berikut:

```
# Menggunakan tanda '<-'
```

```
x <- 10
```

```
# Menggunakan tanda '='
```

```
y = 20
```

```
# Menampilkan nilai variabel
```

```
print(x) # Output: 10
```

```
print(y) # Output: 20
```

2. Menyimpan berbagai tipe data

```
# Vektor
```

```
angka <- c(1, 2, 3, 4, 5)
```

```
# Karakter
```



```
nama <- "John Doe"
```

```
# Logika
```

```
status <- TRUE
```

```
# Menampilkan variabel
```

```
print(angka) # Output: [1] 1 2 3 4 5
```

```
print(nama) # Output: [1] "John Doe"
```

```
print(status) # Output: [1] TRUE
```

B. Fungsi

Fungsi adalah sekumpulan kode yang melakukan tugas tertentu. Fungsi bisa menerima *input (argument)* dan menghasilkan *output (nilai kembali)*. Penggunaan fungsi dalam R adalah sebagai berikut:

1. Membuat fungsi baru, dapat dilakukan dengan menggunakan kata kunci `function()`. Berikut contoh fungsi yang menjumlahkan dua angka:

```
# Mendefinisikan fungsi penjumlahan
```

```
jumlah <- function(a, b) { return(a + b) }
```

```
# Memanggil fungsi
```

```
hasil <- jumlah(5, 10) print(hasil) # Output: 15
```

2. Fungsi bawaan R, contohnya sebagai berikut:

```
# Fungsi mean untuk menghitung rata-rata
```

```
data <- c(10, 20, 30, 40, 50)
```

```
rata_rata <- mean(data)
```

```
#menampilkan rata-rata
```

```
print(rata_rata) # Output: 30
```




```
# Fungsi sum untuk menghitung jumlah
```

```
total <- sum(data)
```

```
print(total) # Output: 150
```

2.4 Packages R

Package R adalah kumpulan fungsi, data, dan dokumentasi yang ditulis dalam bahasa pemrograman R, dirancang untuk memperluas fungsionalitas dasar R dalam analisis data, visualisasi, dan pemodelan statistik. Dengan menggunakan package, pengguna dapat mengakses metode dan teknik terbaru yang dikembangkan oleh komunitas R, sehingga mempermudah proses analisis. Package dapat diinstal dari CRAN menggunakan fungsi `install.packages()` dan harus dimuat dengan `library()`. Beberapa package populer meliputi `ggplot2` (Wickham, 2011) untuk visualisasi data, `dplyr` (Wickham et al., 2014) untuk manipulasi data, dan `caret` (Kun et al., 2020) untuk machine learning. Berikut adalah contoh dasar instalasi package:

```
# Menginstal package dari CRAN
```

```
install.packages("ggplot2") # Contoh package untuk visualisasi data
```

```
# Memuat package ke dalam sesi R
```

```
library(ggplot2)
```

```
# Menggunakan fungsi dari package
```

```
data(mpg) # Menggunakan dataset bawaan dari ggplot2
```

```
ggplot(mpg, aes(x = displ, y = hwy)) + geom_point() # Membuat plot sebar
```

3. PENGOLAHAN DATA DASAR

Data dalam statistika terbagi menjadi dua kategori utama yaitu kualitatif dan kuantitatif. Masing-masing jenis data ini memiliki karakteristik, metode pengumpulan, dan analisis yang berbeda.

3.1 Data Kualitatif

Data kualitatif adalah data yang bersifat deskriptif dan tidak dapat diukur dengan angka. Data ini biasanya diperoleh melalui wawancara, observasi, atau kuesioner yang tidak terstruktur, dan sering kali berupa kategori atau label. Data kualitatif dapat menggambarkan sifat, karakteristik, dan kategori dari subjek yang diteliti. Contohnya:

- Warna rambut: Hitam, coklat, pirang.
- Jenis kelamin: Laki-laki, perempuan.
- Kategori pendidikan: SMA, S1, S2.
- Preferensi makanan: Pedas, manis, asam.

A. *Pengolahan Data Kualitatif*

Perhitungan data kualitatif dapat dilakukan sebagai berikut:

1. Frekuensi: menghitung jumlah kemunculan setiap kategori.
2. Persentase: menghitung persentase masing-masing kategori dari total, dapat dihitung dengan rumus:

$$\text{Persentase} = (\text{frekuensi kategori} / \text{total observasi}) \times 100\%$$

Contoh soal:

Sasa melakukan survei terhadap 100 orang mengenai warna favorit mereka. Hasilnya adalah sebagai berikut:

- a. Merah: 30
- b. Biru: 50



c. Hijau: 20

Maka, Frekuensi dan Persentasenya:

- a. Merah: 30 orang, $(30/100) \times 100\% = 30\%$
- b. Biru: 50 orang, $(50/100) \times 100\% = 50\%$
- c. Hijau: 20 orang, $(20/100) \times 100\% = 20\%$

B. Pengujian Data Kualitatif

a. Pengujian Chi-Square

Pengujian Chi-Square (χ^2) adalah metode statistik yang digunakan untuk menentukan apakah ada hubungan yang signifikan antara dua variabel kategori. Ada dua jenis utama dari pengujian Chi-Square:

1. *Chi-Square Goodness of Fit*, digunakan untuk menguji apakah distribusi frekuensi dari satu variabel kategori sesuai dengan distribusi yang diharapkan.
2. *Chi-Square Test for Independence*, digunakan untuk menguji apakah dua variabel kategori independen satu sama lain dalam suatu tabel kontingensi.

Langkah-langkah Pengujian Chi-Square:

1. Menentukan Hipotesis:
 - H_0 (Hipotesis nol): Tidak ada hubungan antara variabel.
 - H_1 (Hipotesis alternatif): Ada hubungan antara variabel.
2. Menghitung nilai Chi-Square, menggunakan rumus:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

di mana O adalah frekuensi yang diamati dan E adalah frekuensi yang diharapkan.

3. Menentukan derajat bebas (df):



$$df=(r-1)(c-1)$$

di mana r adalah jumlah baris dan c adalah jumlah kolom dalam tabel kontingensi.

4. Membandingkan nilai Chi-Square dengan tabel Chi-Square:

Jika nilai χ^2 lebih besar dari nilai kritis pada tingkat signifikansi tertentu (misalnya $\alpha = 0,05$), maka H_0 ditolak.

b. Pengujian Fisher

Uji Fisher adalah metode statistik yang digunakan untuk menentukan apakah ada perbedaan yang signifikan antara dua kelompok kecil dalam data yang dikategorikan. Uji ini sangat berguna ketika ukuran sampel kecil dan tidak memenuhi asumsi distribusi normal.

Langkah-langkah Pengujian Fisher:

1. Menentukan Hipotesis:

- a. H_0 : Tidak ada perbedaan antara dua kelompok.
- b. H_1 : Ada perbedaan antara dua kelompok.

2. Membuat Tabel Kontingensi: Tabel 2x2 untuk menganalisis frekuensi dari dua kelompok.

3. Menghitung Probabilitas: Menggunakan rumus yang melibatkan kombinasi untuk menghitung probabilitas mendapatkan distribusi yang sama atau lebih ekstrem dari data yang diamati.

4. Membandingkan Nilai P: Jika nilai P yang diperoleh dari Uji Fisher kurang dari tingkat signifikansi (misalnya $\alpha = 0,05$), maka H_0 ditolak.

c. Kapan Uji Chi-Square dan Fisher Digunakan?

- Jika semua sel memiliki frekuensi ≥ 5 , gunakan uji Chi-Square
- Jika ada sel dengan frekuensi < 5 , gunakan uji Fisher

d. Analisis Chi-Square dengan R

Contoh soal:

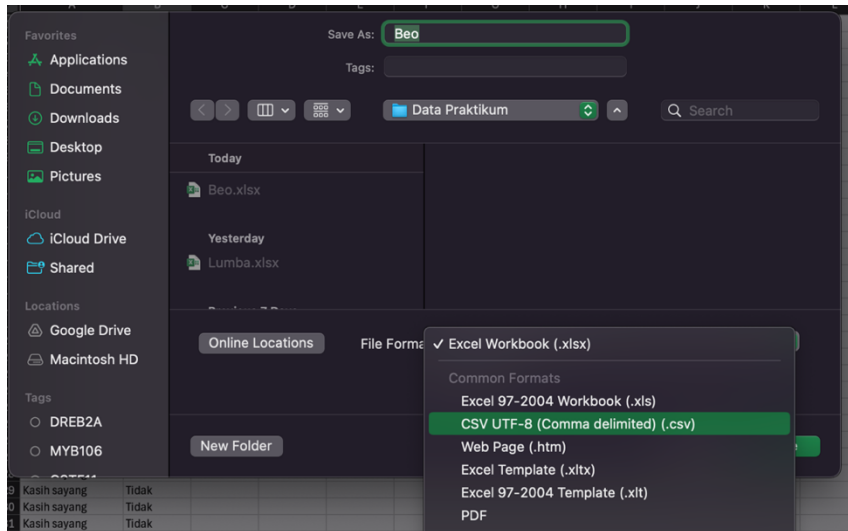
Seorang peneliti tertarik pada apakah burung beo dapat dilatih untuk berbicara. Ia mencoba melatih mereka berbicara dengan memberi mereka makanan atau kasih sayang (seperti mengelus dan memberi pujian) sebagai hadiah. Di akhir minggu dibuat catatan mengenai burung yang berbicara dan tidak. Dari perlakuan kasih sayang, dari 162 burung beo, hanya 48 yang berbicara, sedangkan perlakuan makanan dari 37 burung beo hanya 11 yang tidak bisa berbicara. Pertanyaannya adalah apakah ada pengaruh perlakuan pelatihan terhadap kemampuan burung beo dalam belajar bersiul?

Langkah-langkah visualisasi data dan analisis Chi-Square dengan R:

1. Input data pada Ms. Excel

	A	B	C	D	E	F	G	H	I	J
152	Kasih sayang	Ya								
153	Kasih sayang	Ya								
154	Kasih sayang	Ya								
155	Kasih sayang	Ya								
156	Kasih sayang	Ya								
157	Kasih sayang	Ya								
158	Kasih sayang	Ya								
159	Kasih sayang	Ya								
160	Kasih sayang	Ya								
161	Kasih sayang	Ya								
162	Kasih sayang	Ya								
163	Kasih sayang	Ya								
164	Makanan	Tidak								
165	Makanan	Tidak								
166	Makanan	Tidak								
167	Makanan	Tidak								
168	Makanan	Tidak								
169	Makanan	Tidak								
170	Makanan	Tidak								
171	Makanan	Tidak								
172	Makanan	Tidak								
173	Makanan	Tidak								
174	Makanan	Tidak								
175	Makanan	Ya								
176	Makanan	Ya								

2. Kemudian, Simpan dengan *file type* *.csv



3. Mengatur direktori kerja, tempat dimana data dan hasil analisis akan tersimpan. Gunakan perintah: `setwd("path")`:

```
# Mengatur direktori kerja
```

```
setwd("~/Documents/MATERI STATISTIK/Data Praktikum")
```

```
# Menampilkan direktori kerja
```

```
getwd()
```

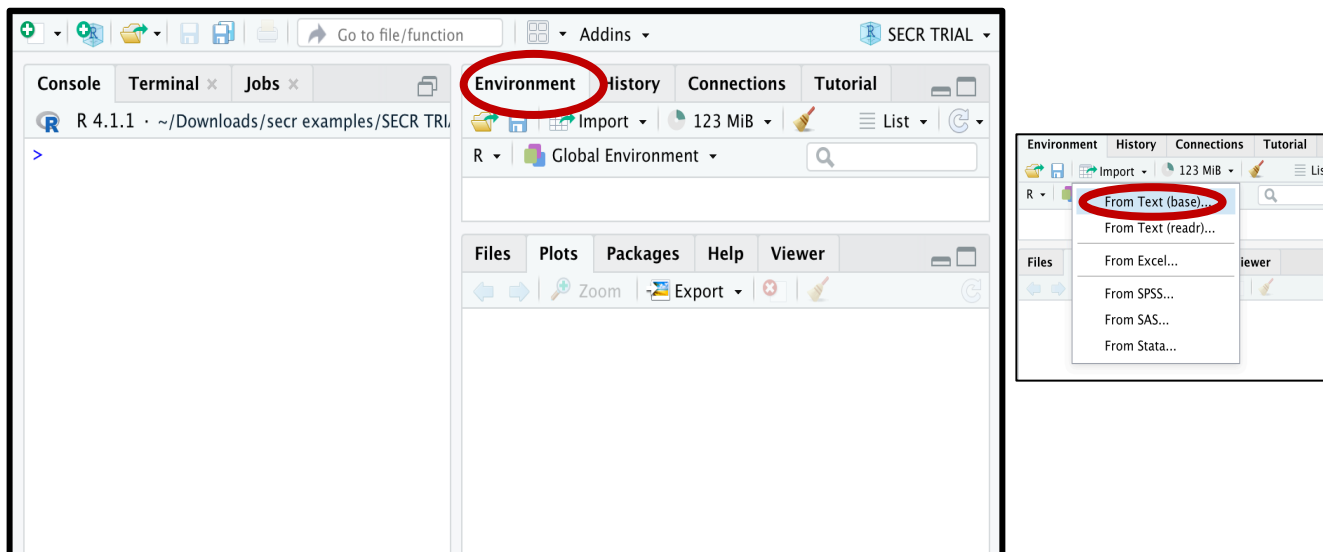
4. Memanggil data dapat dilakukan dua cara yaitu:

- a) Cara pertama, gunakan perintah: `read.csv("path")` pada console sebagai berikut:

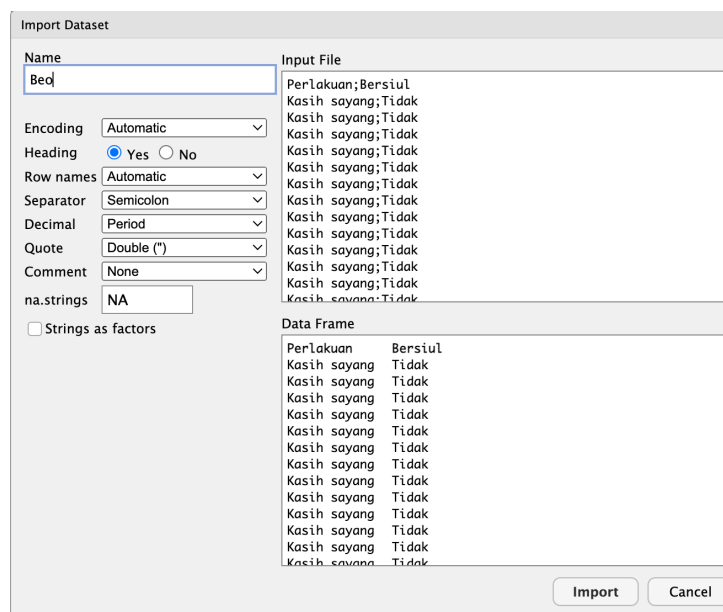
```
# Memanggil data
```

```
Beo <- read.csv("~/Documents/MATERI STATISTIK/Data Praktikum/Beo.csv",  
sep=";")
```

- b) Cara kedua, buka Rstudio klik "Environment", kemudian klik from "text (base)"



5. Pastikan tampilan data sebagai berikut, kemudian klik "import"



6. Menampilkan data

```
head(Beo)
```

```
View(Beo)
```

```
table(Beo)
```


7. Instalasi *packages*

Untuk instalasi package dapat dilakukan dengan instalasi beberapa package sekaligus. Namun, direkomendasikan untuk install satu per satu untuk menghindari kegagalan dalam instalasi.

```
#instalasi beberapa package sekaligus
```

```
install.packages(c("beanplot","gmodels","plotrix","car","pastecs","reshape2 "))
```

```
#instalasi package satu per satu
```

```
install.packages("beanplot")
```

```
install.packages("gmodels")
```

```
install.packages("plotrix")
```

```
install.packages("car")
```

```
install.packages("pastecs")
```

```
install.packages("reshape2")
```

Apabila gagal instalasi coba dengan penambahan script `type = "binary"` berikut:

```
install.packages("gmodels", type = "binary")
```

8. Setelah package terinstal, mengaktifkan package dilakukan dengan cara:

```
library(beanplot)
```

```
library(gmodels)
```

```
library(plotrix)
```

```
library(car)
```

```
library(pastecs)
```

```
library(reshape2)
```

9. Membuat persentase

a) Membuat tabel kontingensi

```
#membuat tabel kontingensi
```



```
tabel_kontingensi <- table(Beo$Perlakuan, Beo$Bersiul)
print(tabel_kontingensi)
```

Output akan ditampilkan pada bagian console R sebagai berikut:

	Tidak	Ya
Kasih sayang	114	48
Makanan	11	26

b) Menghitung persentase

```
#menghitung persentase
tabel_kontingensi100<- prop.table(tabel_kontingensi,1)
tabel_kontingensi100<- round(tabel_kontingensi100*100)

#print hasil
print(tabel_kontingensi100)
```

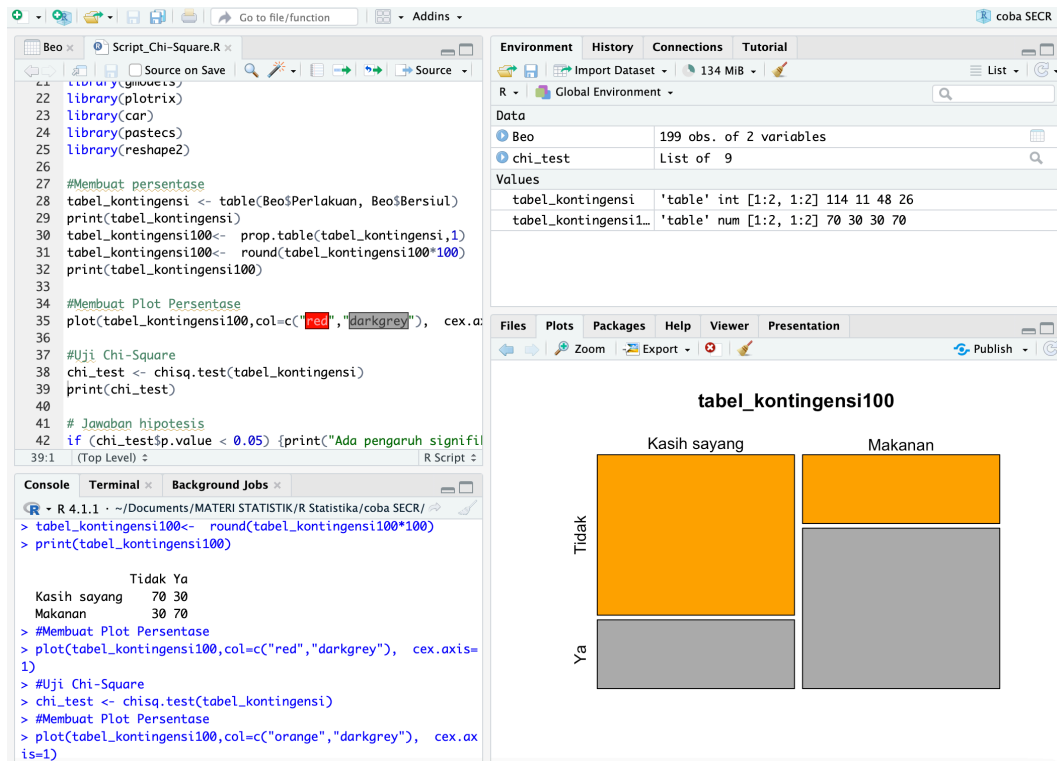
Output akan ditampilkan pada bagian console R sebagai berikut:

	Tidak	Ya
Kasih sayang	70	30
Makanan	30	70

10. Membuat plot persentase

```
plot(tabel_kontingensi100,col=c("red","darkgrey"), cex.axis=1)
```

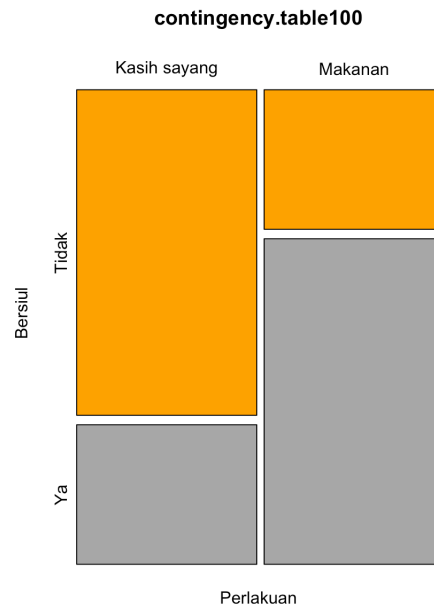
Plot akan selalu ditampilkan pada pojok kanan bawah sebagai berikut:



Catatan, jika ingin menyimpan gambar dapat dilakukan dapat dengan cara:

- Klik tombol **Export** di tengah atas panel Plots.
- Pilih format yang diinginkan (misalnya PNG, PDF, JPEG).
- Tentukan lokasi penyimpanan dan klik **Save**.

Tampilan plot setelah di export:



11. Uji Chi-Square

```
#uji chi-square  
chi_test <- chisq.test(tabel_kontingensi)  
  
#print hasil  
print(chi_test)
```

Output analisis Chi-Square adalah sebagai berikut:

```
Pearson's Chi-squared test with Yates' continuity correction  
data: tabel_kontingensi  
X-squared = 19.594, df = 1, p-value = 9.576e-06
```

12. Interpretasi hasil uji Chi-Square

```
if (chi_test$p.value < 0.05) {print("Ada pengaruh signifikan dari perlakuan terhadap  
kemampuan bersiul burung beo.")} else { print("Tidak ada pengaruh signifikan dari  
perlakuan terhadap kemampuan bersiul burung beo.") }
```

Output akan ditampilkan pada console sebagai berikut:

```
[1] "Ada pengaruh signifikan dari perlakuan terhadap kemampuan bersiul burung beo."
```

3.2 Data Kuantitatif

Data kuantitatif adalah jenis data yang dapat diukur dan dinyatakan dalam bentuk angka. Data kuantitatif sering digunakan untuk melihat pola, tren, atau hubungan antara variabel dan diambil melalui metode yang objektif, seperti survei, eksperimen, dan pengukuran. Data kuantitatif terbagi menjadi dua jenis utama:

- a. Data Diskrit, yaitu data yang hanya dapat mengambil nilai-nilai tertentu, umumnya bilangan bulat. Biasanya berasal dari hasil penghitungan.

Contoh:

- Jumlah siswa di kelas: 10, 15, 20 siswa.
- Jumlah mobil di parkir: 5, 10, 12 mobil.

- b. Data Kontinu, yaitu data yang dapat mengambil nilai apa pun dalam rentang tertentu, sering kali berasal dari hasil pengukuran.

Contoh: -

- Tinggi badan: 160 cm, 170.5 cm.
- Berat badan: 55.5 kg, 60.3 kg.
- Suhu: 36.7°C, 38.2°C.

A. Pengujian Data Kuantitatif

Pengujian data kuantitatif biasanya melibatkan analisis statistik untuk menilai hubungan, perbedaan, atau distribusi dalam data. Berikut adalah beberapa metode pengujian yang umum digunakan adalah uji T (T-Test), ANOVA (Analisis of Varians), regresi, korelasi, dan uji Chi-Square.



1. Korelasi dan Regresi

a. Korelasi

Korelasi adalah metode statistik yang digunakan untuk menentukan hubungan antar dua variabel. Korelasi dapat diukur dengan menggunakan beberapa metode, yang paling umum adalah koefisien korelasi Pearson untuk data yang berdistribusi normal dan koefisien korelasi Spearman untuk data ordinal atau yang tidak memenuhi asumsi normalitas. Nilai koefisien korelasi berkisar antara -1 hingga 1, dengan $r = 1$ menunjukkan korelasi positif sempurna; $r = -1$ menunjukkan korelasi negatif sempurna; dan $r = 0$ menunjukkan tidak ada korelasi atau hubungan linear. Ada 3 jenis korelasi, yaitu:

- Korelasi positif, ketika satu variabel meningkat, variabel lainnya juga meningkat. Contohnya, korelasi antara kepadatan pohon terhadap keanekaragaman hayati. Seorang peneliti ingin mengetahui apakah ada hubungan antara kepadatan pohon dalam suatu hutan terhadap keanekaragaman spesies hewan. Jika koefisien korelasinya positif, berarti semakin padat hutan tersebut, semakin tinggi keanekaragaman hayatinya.
- Korelasi negatif, ketika satu variabel meningkat maka variabel lainnya menurun. Contohnya, analisis data tentang tingkat fragmentasi hutan dan ukuran populasi spesies tertentu, peneliti dapat menentukan apakah ada korelasi negatif, di mana lebih banyak fragmentasi berhubungan dengan penurunan populasi satwa liar.
- Korelasi nol, tidak ada hubungan linear antara kedua variabel. Contohnya: Peneliti ingin mengetahui apakah jenis tanah memiliki pengaruh terhadap pola penyebaran penyakit pada pohon-pohon dalam suatu kawasan hutan. Setelah menganalisis data, peneliti menemukan bahwa tidak ada pola yang jelas antara jenis tanah dan



penyebaran penyakit pada pohon, dengan koefisien korelasi mendekati nol.

Koefisien Korelasi Pearson dihitung sebagai:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

r: koefisien korelasi Pearson

n: Jumlah pasangan data

x: Nilai dari variabel pertama

y: Nilai dari variabel kedua.

$\sum xy$: Jumlah dari hasil kali setiap pasangan data ($x_i \times y_i$)

$\sum x$: Jumlah dari seluruh nilai variabel pertama

$\sum y$: Jumlah dari seluruh nilai variabel kedua

$\sum x^2$: Jumlah dari kuadrat seluruh nilai variabel pertama

$\sum y^2$: Jumlah dari kuadrat seluruh nilai variabel kedua

Koefisien Korelasi Spearman dihitung sebagai:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

ρ : Koefisien korelasi Spearman

d_i : Perbedaan antara peringkat dari setiap pasangan nilai ($x_i - y_i$)

$\sum d_i^2$: Jumlah dari kuadrat perbedaan peringkat.

n: Jumlah pasangan data.

b. Regresi

Persamaan regresi menggambarkan hubungan antara variabel bebas (independen) dan variabel terikat (dependen). Berikut adalah penjelasan dua jenis regresi yang umum digunakan:

- Regresi linier sederhana Digunakan untuk melihat hubungan antara satu variabel bebas (X) dan satu variabel terikat (Y). Regresi linear sederhana dapat dihitung dengan rumus:



$$Y = a + bX$$

Y: Variabel terikat (dependen)

X: Variabel bebas (independen)

a: Intersep (nilai Y saat X = 0)

b: Koefisien regresi (kemiringan garis)

$$b = \frac{n\sum(XY) - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$a = \frac{\sum Y - b\sum X}{n}$$

- Regresi linier berganda, digunakan ketika ada lebih dari satu variabel bebas ($X_1, X_2, \dots, X_k$).

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

$b_1, b_2, \dots, b_k$: Koefisien regresi untuk masing – masing variabel bebas

2. T-Test

T-Test adalah metode statistik yang digunakan untuk membandingkan rata-rata dari dua kelompok atau kondisi. Ada beberapa jenis T-Test yang digunakan tergantung pada situasi data dan tujuan analisis. Berikut adalah jenis-jenis T-Test yang umum digunakan:

a. Independent T-Test (Two-sample T-Test)

Digunakan untuk membandingkan rata-rata dari dua kelompok yang tidak berhubungan. Prinsip pengujian uji ini adalah melihat perbedaan variasi kedua kelompok data, sehingga sebelum dilakukan pengujian, terlebih dahulu harus diketahui apakah variannya sama (*equal variance*) atau variannya berbeda (*unequal variance*). Contoh: Ingin mengetahui apakah ada perbedaan signifikan antara nilai rata-rata kelompok eksperimen dan kelompok kontrol.



$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$\bar{x}$: rata – rata sampel
 n : ukuran sampel
 s^2 : variansi sampel

b. Paired T-Test

Uji t berpasangan digunakan untuk membandingkan perbedaan rata-rata ketika pengamatan diperoleh secara berpasangan, dan karenanya saling bergantung (Hsu & Lachenbruch, 2014). Biasanya, data diambil dari subjek yang sama pada dua waktu yang berbeda. Contoh: Menguji apakah ada perbedaan berat badan seseorang sebelum dan sesudah menjalani diet. Syarat jenis uji ini adalah:

- data berdistribusi normal.
- kedua kelompok data adalah dependen (saling berhubungan/berpasangan);
- jenis data yang digunakan adalah numerik dan kategorik (dua kelompok).

Rumus T-Test yang digunakan untuk sampel berpasangan (paired) adalah

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

dengan

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

s_p : standard deviasi gabungan
 $\bar{x}$: rata – rata sampel
 n : ukuran sampel
 s^2 : variansi sampel

c. One-sample T-Test

Digunakan untuk membandingkan rata-rata satu kelompok dengan nilai tertentu atau standar yang sudah diketahui. Contoh: Menguji apakah rata-rata tinggi siswa di satu sekolah sama dengan tinggi rata-rata nasional.

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

$\bar{x}$: rata – rata sampel

μ : rata – rata populasi atau hipotesis yang akan diuji

$s_{\bar{x}}$: standard error

s : standar deviasi sampel

n : jumlah data sampel

Asumsi T-Test:

- Data harus terdistribusi secara normal.
- Varian kedua kelompok harus homogen (untuk independent T-Test).
- Sampel yang digunakan harus independen (untuk independent T-Test).

Output T-Test:

- T-Value: Nilai statistik T yang menunjukkan perbedaan antara kelompok.
- P-Value: Nilai probabilitas yang menunjukkan apakah perbedaan yang ditemukan signifikan secara statistik. Biasanya, jika p-value < 0.05, maka perbedaan dianggap signifikan.

Perbedaan Utama Antara Jenis T-Test:

1. Independent Samples T-Test, digunakan ketika membandingkan dua kelompok yang berbeda dan tidak terkait satu sama lain.

2. Paired Samples T-Test, digunakan untuk membandingkan dua set data yang berasal dari kelompok atau individu yang sama.
3. One-sample T-Test, digunakan untuk membandingkan satu kelompok data dengan nilai tertentu (bukan kelompok lain).

Kapan Menggunakan Setiap T-Test:

- One-sample T-Test, jika Anda memiliki satu kelompok dan ingin membandingkan rata-rata dengan nilai populasi yang diketahui.
- Independent samples T-Test, jika Anda membandingkan dua kelompok yang berbeda dan tidak ada hubungan antara anggota kelompok.
- Paired samples T-Test, jika Anda membandingkan pengukuran dari kelompok atau individu yang sama dalam dua kondisi berbeda (misalnya, sebelum dan sesudah intervensi).

2. ANOVA (*Analysis of Variance*)

ANOVA adalah uji statistik yang digunakan untuk membandingkan rata-rata dari data yang lebih dari dua kelompok. ANOVA membantu menentukan apakah perbedaan rata-rata antar kelompok signifikan atau tidak. Beberapa jenis ANOVA adalah sebagai berikut:

a. *One-Way* ANOVA

Digunakan untuk menguji perbedaan rata-rata dari satu variabel independen (faktor) dengan beberapa kelompok (lebih dari dua). Contoh: Menguji apakah ada perbedaan signifikan antara nilai ujian dari tiga kelas yang berbeda.

b. *Two-Way* ANOVA

Digunakan ketika ada dua variabel independen (faktor) yang diuji terhadap satu variabel dependen. Contoh: Menguji pengaruh dua faktor (jenis pelatihan dan durasi pelatihan) terhadap kinerja kerja.



c. *Repeated Measures ANOVA*

Digunakan untuk menguji perbedaan rata-rata dari data yang dikumpulkan pada subjek yang sama pada beberapa waktu yang berbeda atau beberapa kondisi yang berbeda. Contoh: Mengukur efektivitas suatu obat pada sekelompok pasien pada beberapa waktu yang berbeda (sebelum, selama, dan setelah perawatan).

Asumsi ANOVA:

- Data terdistribusi normal.
- Homogenitas varian (varians dari kelompok-kelompok yang dibandingkan harus sama).
- Data diambil secara acak
- Kelompok harus independen, bahwa data yang berasal dari satu kelompok tidak boleh memiliki hubungan atau pengaruh terhadap data di kelompok lainnya

Output ANOVA:

- F-Value: Nilai statistik F yang mengukur perbedaan antar rata-rata kelompok.
- P-Value: Sama seperti T-Test, menunjukkan apakah perbedaan yang ditemukan signifikan. Jika $p\text{-value} < 0.05$, perbedaan dianggap signifikan.

Post-hoc Test:

Jika hasil ANOVA menunjukkan adanya perbedaan yang signifikan, post-hoc test (seperti Tukey HSD, Duncan, dsb,) digunakan untuk menentukan kelompok mana yang berbeda secara signifikan

B. Analisis dengan R

a. Korelasi

Contoh soal:



Sebuah penelitian dilakukan untuk mengevaluasi hubungan antara pertumbuhan tinggi pohon dan faktor lingkungan seperti kelembapan tanah. Sebanyak 20 pohon diukur tinggi dan kelembapan tanahnya. Data yang diperoleh menunjukkan variasi kelembapan tanah dari 15% hingga 30%.

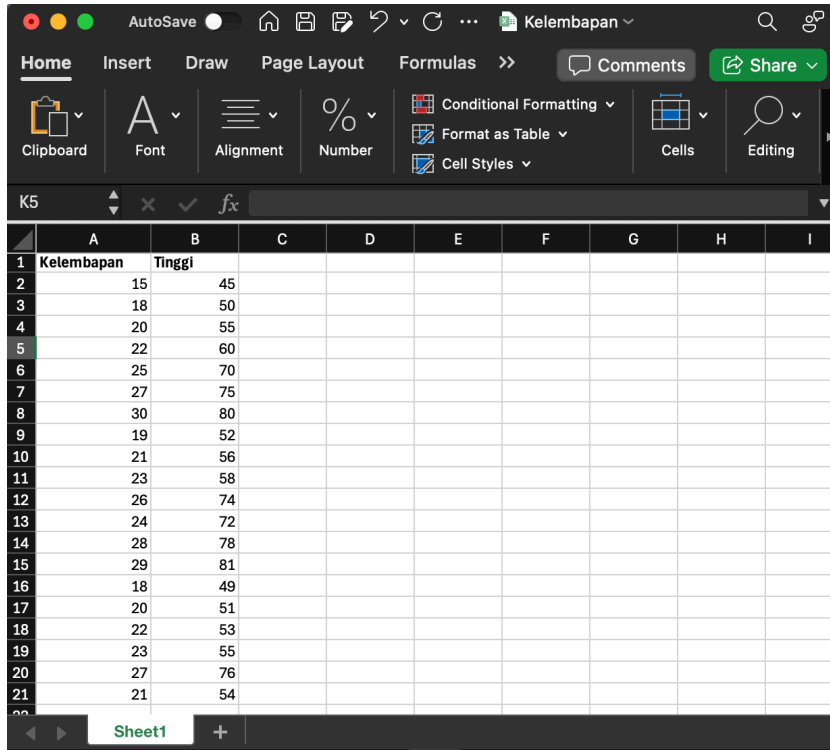
Kelembapan (%)	Tinggi (cm)
15	45
18	50
20	55
22	60
25	70
27	75
30	80
19	52
21	56
23	58
26	74
24	72
28	78
29	81
18	49
20	51
22	53
23	55
27	76
21	54

Pertanyaan:

1. Hitung koefisien korelasi antara tinggi pohon dan kelembapan tanah. Apakah ada korelasi yang signifikan antara kedua variabel tersebut?
2. Interpretasikan hasil korelasi yang Anda dapatkan. Apa arti dari nilai yang Anda peroleh?

Langkah-langkah Analisis Menggunakan R software

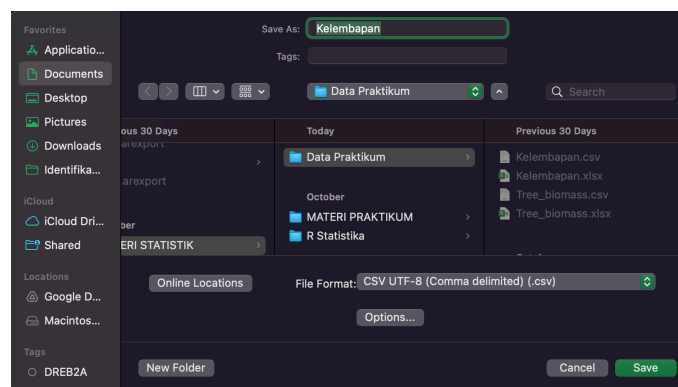
1. Input data di Ms. Excel dilakukan sebagai berikut:



The screenshot shows the Microsoft Excel interface with the 'Home' tab selected. The data is entered into the following table:

	A	B	C	D	E	F	G	H	I
1	Kelembapan	Tinggi							
2		15	45						
3		18	50						
4		20	55						
5		22	60						
6		25	70						
7		27	75						
8		30	80						
9		19	52						
10		21	56						
11		23	58						
12		26	74						
13		24	72						
14		28	78						
15		29	81						
16		18	49						
17		20	51						
18		22	53						
19		23	55						
20		27	76						
21		21	54						

2. Kemudian data di save dengan *file type* "tab delimited [*.txt]" atau "comma delimited [*.csv]"



3. Mengatur direktori kerja, tempat dimana data dan hasil analisis akan tersimpan. Gunakan perintah: `setwd("path")`

Mengatur direktori kerja

```
setwd("~/Documents/MATERI STATISTIK/Data Praktikum")
```

Menampilkan direktori kerja

```
getwd()
```

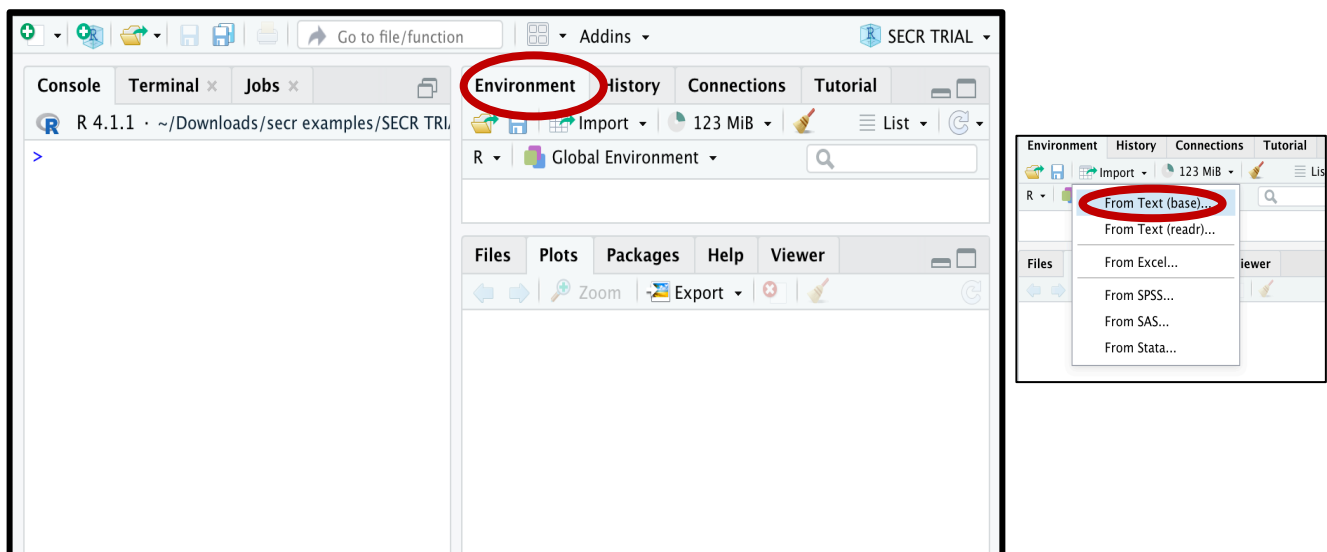
4. Memanggil data atau import data dapat dilakukan dua cara yaitu:

- a) Cara pertama, gunakan perintah: `read.csv("path")` pada console sebagai berikut:

Memanggil data

```
Kelembapan <- read.csv("~/Documents/MATERI STATISTIK/Data  
Praktikum/Kelembapan.csv", sep=";")
```

- b) Cara kedua, buka Rstudio klik "Environment", kemudian klik from "text (base)"



5. Pastikan tampilan data yang di import sebagai berikut:

Import Dataset

Name: Kelembapan

Input File: Kelembapan;Tinggi
15;45
18;50
20;55
22;60
25;70
27;75
30;80
19;52
21;56
23;58
26;74
24;72
28;78

Encoding: Automatic

Heading: ☒ Yes ☐ No

Row names: Automatic

Separator: Semicolon

Decimal: Period

Quote: Double (")

Comment: None

na.strings: NA

☐ Strings as factors

Data Frame

Kelembapan	Tinggi
15	45
18	50
20	55
22	60
25	70
27	75
30	80
19	52
21	56
23	58
26	74
24	72
28	78

Import Cancel

6. Mengecek struktur data

```
#Mengecek struktur data  
str(Kelembapan)
```

Output akan ditampilkan pada console sebagai berikut:

```
'data.frame':20 obs. of 2 variables:  
 $ Kelembapan: int 15 18 20 22 25 27 30 19 21 23 ...  
 $ Tinggi : int 45 50 55 60 70 75 80 52 56 58 ...
```

7. Instalasi package dan aktivasi package

```
#instalasi package  
install.packages("ggplot2")
```

8. Menghitung koefisien korelasi

```
# Menghitung koefisien korelasi
```




```
kor <- cor(Kelembapan$Kelembapan, Kelembapan$Tinggi)
print(paste("Koefisien Korelasi:", kor))
```

Output akan ditampilkan pada console sebagai berikut:

```
[1] "Koefisien Korelasi: 0.954613826604447"
```

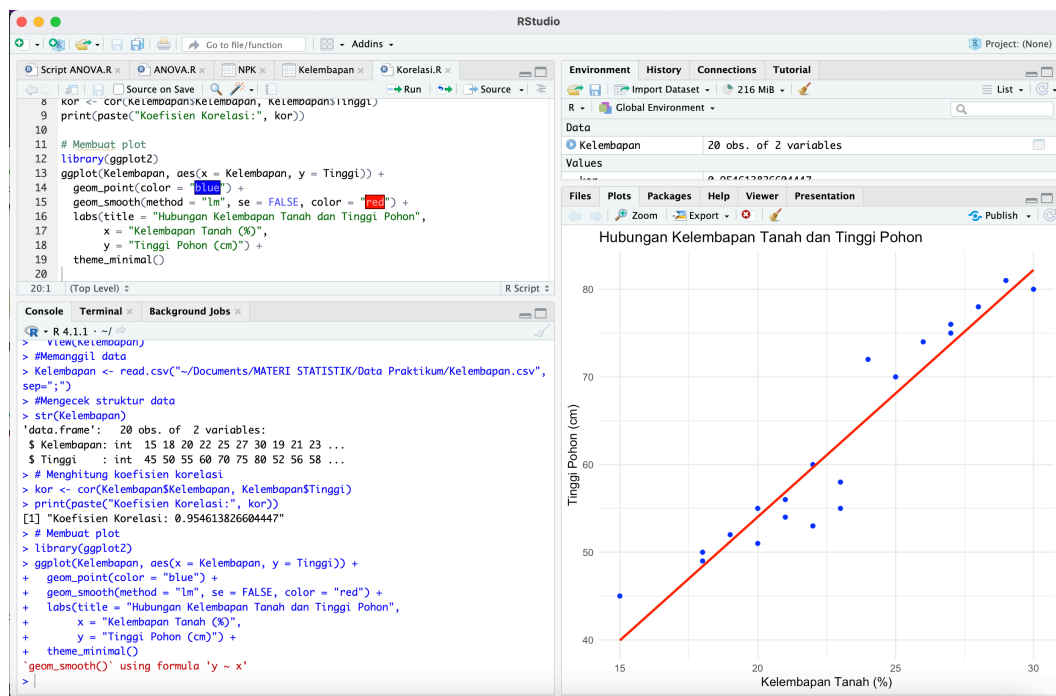
9. Membuat plot

Membuat plot

```
ggplot(Kelembapan, aes(x = Kelembapan, y = Tinggi)) +
  geom_point(color = "blue") +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  labs(title = "Hubungan Kelembapan Tanah dan Tinggi Pohon",
       x = "Kelembapan Tanah (%)",
       y = "Tinggi Pohon (cm)") +
  theme_minimal()
```

10. Hasil visualisasi plot

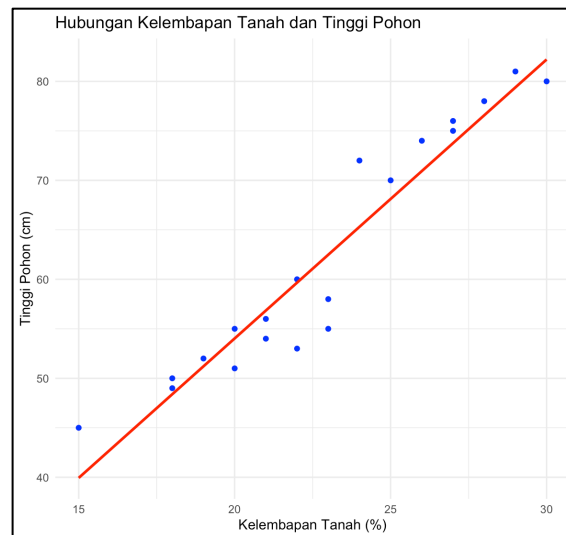
Plot akan selalu ditampilkan pada pojok kanan bawah sebagai berikut:



Catatan, jika ingin menyimpan gambar dapat dilakukan dapat dengan cara:

- Klik tombol **Export** di tengah atas panel Plots.
- Pilih format yang diinginkan (misalnya PNG, PDF, JPEG).
- Tentukan lokasi penyimpanan dan klik **Save**

Tampilan plot setelah di export:



b. Regresi

Contoh soal:

Seorang peneliti lingkungan yang mempelajari faktor-faktor yang memengaruhi biomassa pohon. Anda mengambil sampel pohon dari kawasan hutan dan mengukur beberapa variabel sebagai berikut:

Diameter	Tinggi	Biomassa
10	5	15
15	8	35
20	12	60
25	15	95
30	18	120
35	20	145

40	22	170
45	24	195
50	26	230
55	28	255

Pertanyaan:

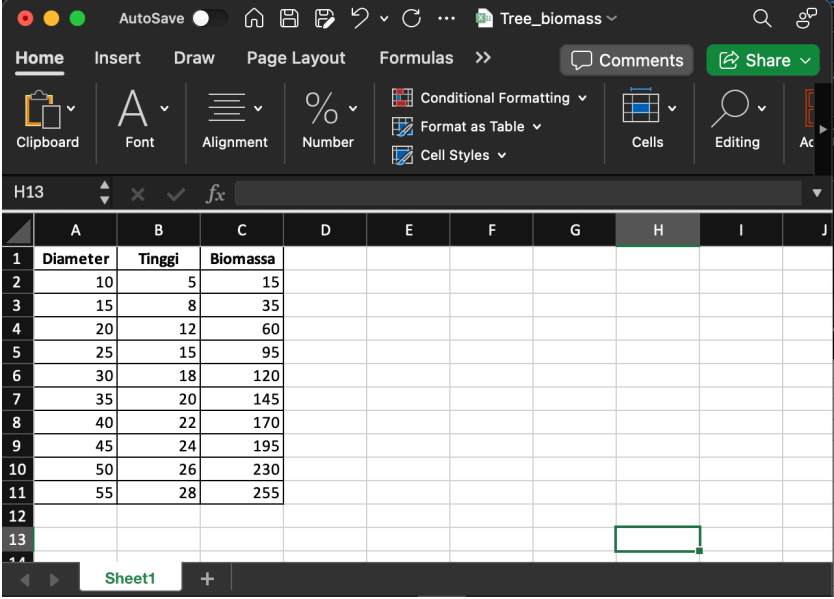
1. Gunakan analisis regresi linier untuk memodelkan hubungan antara diameter dan tinggi terhadap biomassa. Berdasarkan hasil regresi, jelaskan apakah variabel diameter dan tinggi signifikan dalam memengaruhi biomassa!
2. Interpretasikan koefisien regresi untuk diameter dan tinggi. Apa arti nilai koefisien ini dalam konteks biomassa pohon? Cek signifikansi setiap variabel (dengan p-value). Jika p-value untuk diameter dan tinggi di bawah 0,05, variabel tersebut dianggap
3. Lihat nilai R-squared dari model. Apa arti nilai ini dalam konteks data biomassa pohon?
4. Buatlah visualisasi data menggunakan scatter plot untuk menunjukkan hubungan antara diameter dan biomassa, serta tambahkan garis regresi linier pada grafik. Apakah ada pola atau tren yang terlihat dari grafik ini?

Jawaban Panduan:

1. Signifikansi (p-value): perhatikan nilai p-value untuk diameter dan tinggi. Jika p-value < 0.05, variabel tersebut signifikan memengaruhi biomassa. R-squared: periksa nilai r-squared untuk memahami proporsi variabilitas dalam biomassa yang dapat dijelaskan oleh diameter dan tinggi.
2. Interpretasi output:
 - jika koefisien diameter bernilai positif, maka setiap peningkatan satu unit diameter cenderung akan meningkatkan biomassa pohon.
 - Jika R-squared tinggi (misalnya >0.8), ini berarti model dapat menjelaskan sebagian besar variabilitas dalam Biomassa dengan menggunakan diameter dan tinggi.

Langkah-langkah Analisis Menggunakan R software

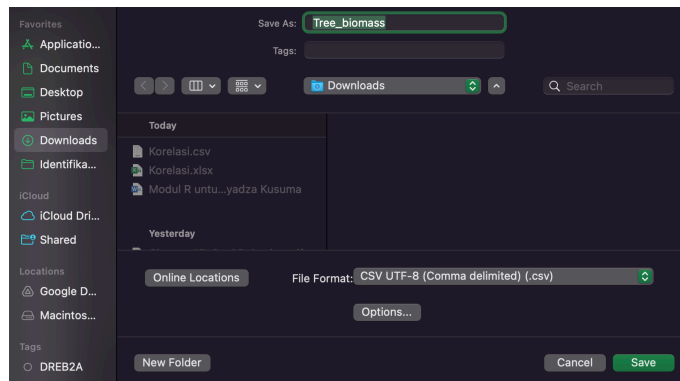
1. Input data di Ms. Excel dilakukan sebagai berikut:



The screenshot shows the Microsoft Excel interface with a file named 'Tree_biomass'. The data is entered into a table with the following structure:

	A	B	C	D	E	F	G	H	I	J
1	Diameter	Tinggi	Biomassa							
2	10	5	15							
3	15	8	35							
4	20	12	60							
5	25	15	95							
6	30	18	120							
7	35	20	145							
8	40	22	170							
9	45	24	195							
10	50	26	230							
11	55	28	255							
12										
13										
14										

2. Kemudian data di save dengan *file type* "tab delimited [*.txt]" atau "comma delimited [*.csv]"



3. Mengatur direktori kerja, tempat dimana data dan hasil analisis akan tersimpan. Gunakan perintah: `setwd("path")`

Mengatur direktori kerja

```
setwd("~/Documents/MATERI STATISTIK/Data Praktikum")
```

```
# Menampilkan direktori kerja
```

```
getwd()
```

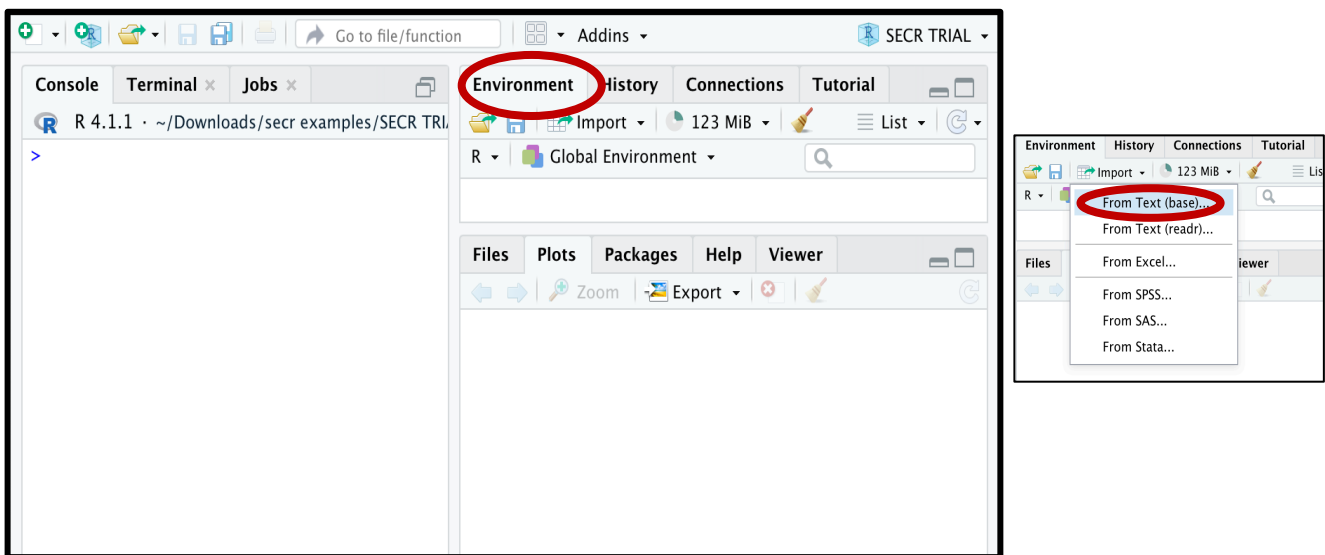
4. Memanggil data atau import data dapat dilakukan dua cara yaitu:

a) Cara pertama, gunakan perintah: `read.csv("path")` pada console sebagai berikut:

```
# Memanggil data
```

```
Tree_biomass <- read.csv("~/Documents/MATERI STATISTIK/Data  
Praktikum/Tree_biomass.csv", sep=";")
```

b) Cara kedua, buka Rstudio klik "Environment", kemudian klik from "text (base)"



5. Pastikan tampilan data yang di import sebagai berikut:

Diameter	Tinggi	Biomassa
10	5	15
15	8	35
20	12	60
25	15	95
30	18	120
35	20	145
40	22	170
45	24	195
50	26	230
55	28	255

6. Verifikasi struktur data

```
#Mengecek struktur data
```

```
str(Tree_biomass)
```

Output akan ditampilkan pada console sebagai berikut:

```
'data.frame': 10 obs. of 3 variables:
 $ Diameter : int 10 15 20 25 30 35 40 45 50 55
 $ Tinggi    : int 5 8 12 15 18 20 22 24 26 28
 $ Biomassa: int 15 35 60 95 120 145 170 195 230 255
```

7. Memastikan format data

```
#Memastikan Format Data:
```

```
Tree_biomass$Diameter <- as.numeric(Tree_biomass$Diameter)
```

```
Tree_biomass$Biomassa <- as.numeric(Tree_biomass$Biomassa)
```

8. Instalasi package dan aktivasi package

```
#instalasi package
```

```
install.packages("ggplot2")
```



```
#aktivasi package
```

```
library(ggplot2)
```

9. Menghitung koefisien korelasi

```
#Analisis regresi
```

```
model_pohon <- lm(Biomassa ~ Diameter + Tinggi, data = Tree_biomass)
```

```
# Menampilkan ringkasan hasil regresi
```

```
summary(model_pohon)
```

Output akan ditampilkan pada console sebagai berikut:

Call:

```
lm(formula = Biomassa ~ Diameter + Tinggi, data = Tree_biomass)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.4881	-2.9571	0.5119	3.0119	4.0633

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-42.4274	3.4225	-12.396	5.11e-06 ***
Diameter	5.7066	0.6307	9.048	4.12e-05 ***
Tinggi	-0.6201	1.2400	-0.500	0.632

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.759 on 7 degrees of freedom

Multiple R-squared: 0.9984, Adjusted R-squared: 0.9979

F-statistic: 2124 on 2 and 7 DF, p-value: 1.806e-10



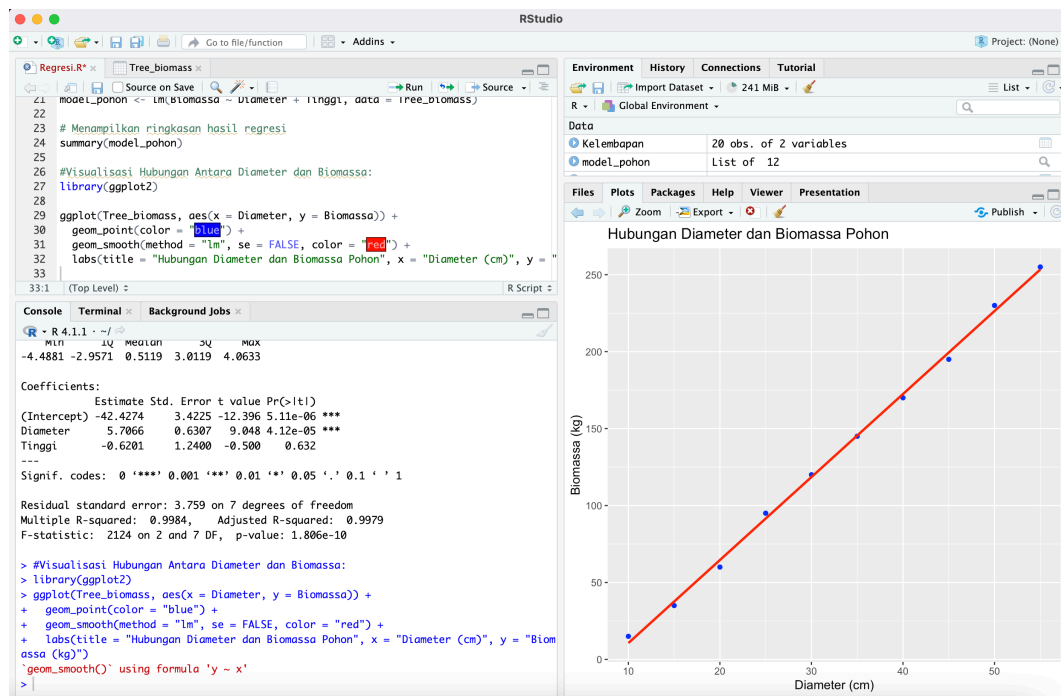
10. Membuat plot

Membuat plot

```
ggplot(Tree_biomass, aes(x = Diameter, y = Biomassa)) +  
  geom_point(color = "blue") +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Hubungan Diameter dan Biomassa Pohon", x = "Diameter (cm)", y =  
"Biomassa (kg)")
```

11. Hasil visualisasi

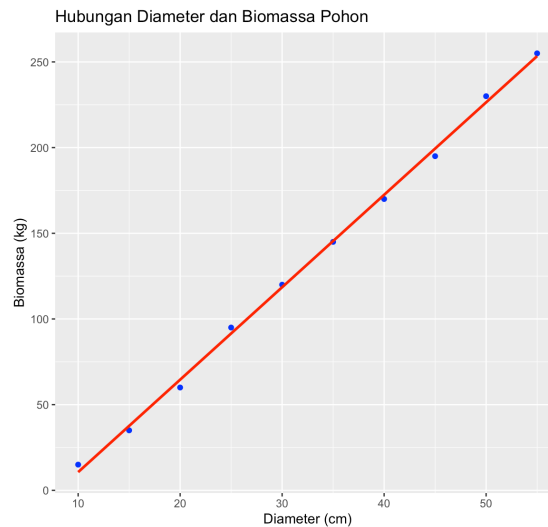
Plot akan selalu ditampilkan pada pojok kanan bawah sebagai berikut:



Catatan, jika ingin menyimpan gambar dapat dilakukan dapat dengan cara:

- Klik tombol **Export** di tengah atas panel Plots.
- Pilih format yang diinginkan (misalnya PNG, PDF, JPEG).
- Tentukan lokasi penyimpanan dan klik **Save**

Tampilan plot setelah di export:



c. Independent T-Test (Two-sample T-Test)

Contoh soal:

Sebuah penelitian dilakukan untuk membandingkan rata-rata berat harimau jantan dan harimau betina. Peneliti ingin tahu apakah ada perbedaan antara berat harimau jantan dengan berat harimau betina. Peneliti mengumpulkan data dari 40 sampel harimau jantan dan 40 sampel harimau betina. Berikut adalah data berat badan harimau (dalam kilogram):

Untuk jenis kelamin jantan, berat badan yang tercatat adalah sebagai berikut:

112.5, 117.4, 120, 121.4, 121.6, 122.3, 123.2, 123.5, 124.5, 124.9, 125, 125.6, 126.4, 126.5, 126.7, 127.6, 127.6, 128, 128.5, 129.4, 129.6, 129.7, 130, 130.5, 131.4, 132.5, 132.7, 134.5, 135.7, 135.9, 136.3, 136.3, 136.7, 137.8, 138.9, 138.9, 139.9, 140, 140.1, 130.6

Untuk jenis kelamin betina, berat badan yang tercatat adalah sebagai berikut:

100.2, 100.5, 101.2, 101.2, 103.4, 104.5, 105.2, 105.2, 105.6, 106.2, 106.2, 106.7, 107.4, 107.4, 108.7, 108.9, 109.2, 111, 111.3, 111.5, 111.9, 113.4, 113.4, 113.5, 116.2, 116.7, 117, 117.4, 118.4, 118.7, 119, 119.2, 119.3, 120, 122.5, 123.5, 124.3, 124.6, 121.4, 128

Pertanyaan:

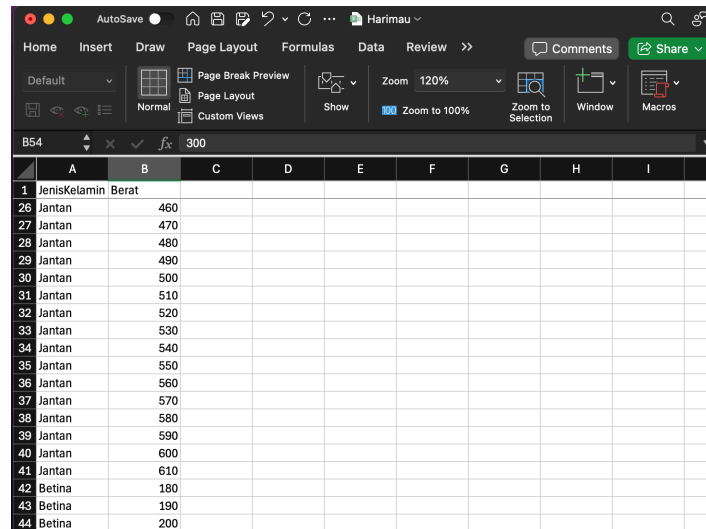
5. Tentukan hipotesis nol (H_0) dan hipotesis alternatif (H_1) untuk membandingkan rata-rata berat antara harimau jantan dan betina!
6. Hitung rata-rata dan variansi berat dari kedua kelompok!
7. Lakukan uji t dua sampel independen untuk menentukan apakah ada perbedaan signifikan dalam rata-rata berat antara harimau jantan dan betina? Gunakan tingkat signifikansi $\alpha = 0.05$!
8. Berdasarkan hasil uji t, apakah hasil menunjukkan peneliti menolak atau menerima hipotesis nol? Jelaskan kesimpulan!

Jawaban Panduan:

3. Hipotesis Nol (H_0): Tidak ada perbedaan rata-rata berat antara harimau jantan dan betina ($\mu_{\text{jantan}} = \mu_{\text{betina}}$).
Hipotesis Alternatif (H_1): Ada perbedaan rata-rata berat antara harimau jantan dan betina ($\mu_{\text{jantan}} \neq \mu_{\text{betina}}$)
4. Menghitung Rata-rata dan Variansi:
 - Rata-rata dan variansi dari kelompok harimau jantan.
 - Rata-rata dan variansi dari kelompok harimau betina.
5. Uji t dua sampel independen:
 - Hitung t-statistik menggunakan rumus T-Test dua sampel independen.
 - Bandingkan t-statistik dengan nilai kritis dari tabel distribusi t untuk $\alpha=0.05$, atau gunakan p-value untuk keputusan.
6. Kesimpulan:
 - Jika p-value < 0.05 , hipotesis nol ditolak, artinya ada perbedaan signifikan dalam berat antara harimau jantan dan betina.
 - Jika p-value > 0.05 , hipotesis nol diterima, artinya tidak ada perbedaan signifikan dalam berat antara harimau jantan dan betina.

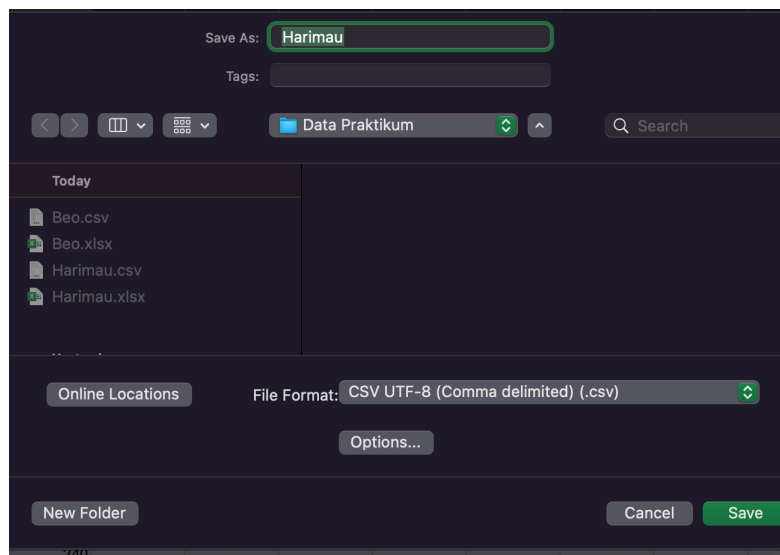
Langkah-langkah Analisis Menggunakan R software

1. Input data di Ms. Excel dilakukan sebagai berikut:



	A	B	C	D	E	F	G	H	I
1	JenisKelamin	Berat							
26	Jantan	460							
27	Jantan	470							
28	Jantan	480							
29	Jantan	490							
30	Jantan	500							
31	Jantan	510							
32	Jantan	520							
33	Jantan	530							
34	Jantan	540							
35	Jantan	550							
36	Jantan	560							
37	Jantan	570							
38	Jantan	580							
39	Jantan	590							
40	Jantan	600							
41	Jantan	610							
42	Betina	180							
43	Betina	190							
44	Betina	200							

2. Kemudian data di save dengan *file type* “tab delimited [*.txt]” atau “comma delimited [*.csv]”



3. Mengatur direktori kerja, tempat dimana data dan hasil analisis akan tersimpan.
Gunakan perintah: `setwd("path")`

Mengatur direktori kerja

```
setwd("~/Documents/MATERI STATISTIK/Data Praktikum")
```

```
# Menampilkan direktori kerja
```

```
getwd()
```

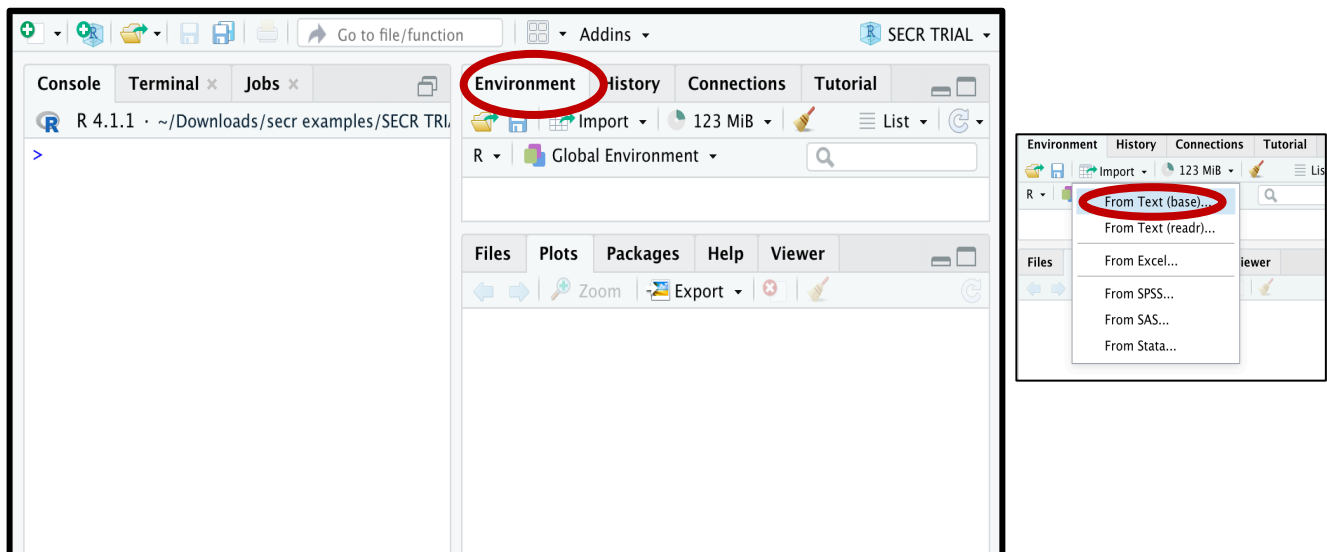
4. Memanggil data atau import data dapat dilakukan dua cara yaitu:

c) Cara pertama, gunakan perintah: `read.csv("path")` pada console sebagai berikut:

```
# Memanggil data
```

```
Harimau <- read.csv("~/Documents/MATERI STATISTIK/Data  
Praktikum/Harimau.csv", sep=";")
```

d) Cara kedua, buka Rstudio klik "Environment", kemudian klik from "text (base)"



5. Pastikan tampilan data yang di import sebagai berikut:

JenisKelamin	Berat
Jantan	220
Jantan	230
Jantan	240
Jantan	250
Jantan	260
Jantan	270
Jantan	280
Jantan	290
Jantan	300
Jantan	310
Jantan	320
Jantan	330
Jantan	340

6. Packages harus sudah diinstalasi terlebih dahulu, namun apabila sudah terinstal dapat dilanjutkan ke langkah aktivasi package. Apabila belum diinstalasi gunakan langkah berikut untuk menginstal package. Instalasi package dapat dilakukan dengan dua cara, yaitu instalasi beberapa package sekaligus dan instalasi satu per satu. Namun, direkomendasikan untuk melakukan instalasi satu per satu untuk menghindari kegagalan dalam instalasi.

#instalasi beberapa package sekaligus

```
install.packages(c("plotrix","car","pastecs", "beanplot", "reshape2"))
```

#instalasi package satu per satu

```
install.packages("plotrix")
```

```
install.packages("car")
```

```
install.packages("pastecs")
```

```
install.packages("beanplot")
```

```
install.packages("reshape2")
```

7. Setelah package terinstal, mengaktifkan package dilakukan dengan cara:

```
#aktivasi package
```

```
library(beanplot)
```

```
library(gmodels)
```

```
library(plotrix)
```

```
library(car)
```

```
library(pastecs)
```

```
library(reshape2)
```

8. Buat grafik *strip*

a) Membuat plot

```
#membuat grafik strip
```

```
stripchart(Harimau$Berat ~ Harimau$JenisKelamin, vertical = TRUE, method =  
"jitter", col = c("firebrick3", "darkcyan"), pch = 19, main = "Stripchart Berat Harimau  
Berdasarkan Jenis Kelamin", xlab = "Jenis Kelamin", ylab = "Berat (kg)", jitter = 0.1)
```

b) Menghitung rata-rata berat

```
#menghitung rata-rata
```

```
rata.berat <- tapply(Harimau$Berat, Harimau$JenisKelamin, mean)
```

```
rata.berat <- tapply(Harimau$Berat, Harimau$JenisKelamin, mean)
```

```
#Membuat garis rata-rata
```

```
loc.strip<- c(1,2)
```

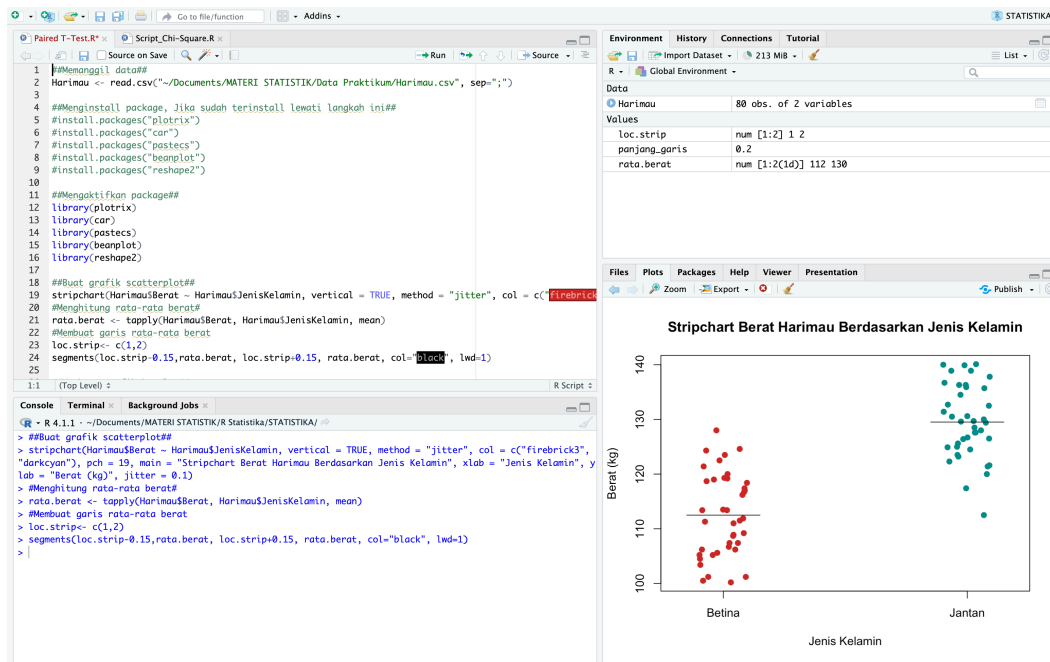
```
segments(loc.strip-0.15,rata.berat, loc.strip+0.15, rata.berat, col="black", lwd=1)
```

Catatan, penjelasan script sebagai berikut:

- *vertical = TRUE*: Menampilkan plot secara vertikal.
- *method = "jitter"*: Menggunakan metode jitter untuk menyebarkan titik-titik secara acak agar tidak tumpang tindih.
- *col = c("firebrick3", "darkcyan")*: Warna titik untuk kategori jenis kelamin.
- *pch = 19*: Bentuk titik (bulatan penuh).
- *main, xlab, ylab*: Judul plot dan label sumbu.
- *jitter = 0.1*: Menyebarkan titik-titik lebih merata dengan sedikit acak.
- *tapply()*: Ini menghitung rata-rata berat harimau berdasarkan jenis kelamin dan menyimpannya dalam variabel *rata.berat*.
- *segments()*: Ini digunakan untuk menggambar garis pada plot.
 - Parameter pertama (*loc.strip - 0.15*): Titik awal horizontal dari garis. Ini menggeser posisi garis sedikit ke kiri.
 - Parameter kedua (*rata.berat*): Titik vertikal dari garis (rata-rata berat).
 - Parameter ketiga (*loc.strip + 0.15*): Titik akhir horizontal dari garis. Ini menggeser posisi garis sedikit ke kanan.
 - *col="black"*: Menetapkan warna garis menjadi hitam.
 - *lwd=1*: Menetapkan ketebalan garis.

d. Plot grafik strip

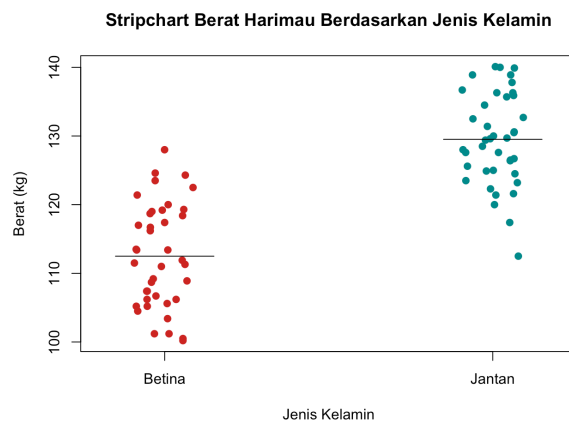
Plot akan selalu ditampilkan pada pojok kanan bawah sebagai berikut:



Catatan, jika ingin menyimpan gambar dapat dilakukan dengan cara:

- Klik tombol **Export** di tengah atas panel Plots.
- Pilih format yang diinginkan (misalnya PNG, PDF, JPEG).
- Tentukan lokasi penyimpanan dan klik **Save**

Tampilan plot setelah di export:

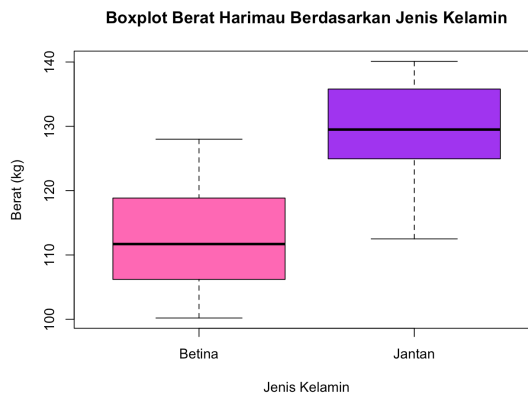


9. Membuat grafik boxplot

#Membuat boxplot

```
boxplot(Harimau$Berat ~ Harimau$JenisKelamin,col = c("hotpink", "purple"), xlab =  
"Jenis Kelamin", ylab = "Berat (kg)", main = "Boxplot Berat Harimau Berdasarkan  
Jenis Kelamin")
```

Boxplot akan ditampilkan pada panel plot sebagai berikut:

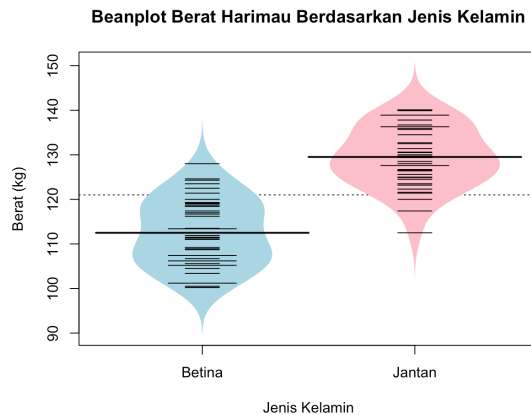


10. Buat grafik beanplot

#Membuat beanplot

```
beanplot(Harimau$Berat ~ Harimau$JenisKelamin, col = list("lightblue", "pink"), main =  
"Beanplot Berat Harimau Berdasarkan Jenis Kelamin", xlab = "Jenis Kelamin", ylab =  
"Berat (kg)", border = NA)
```

Beanplot ditampilkan pada panel plot sebagai berikut:



Catatan, jika ingin menyimpan gambar dapat dilakukan dapat dengan cara:

- `main, xlab, ylab`: Judul plot dan label sumbu.
- `border = NA`: Tidak menampilkan border untuk beanplot

11. Membuat plot histogram

```
# Mengatur agar 2 grafik tampil berdampingan
```

```
par(mfrow = c(1, 2))
```

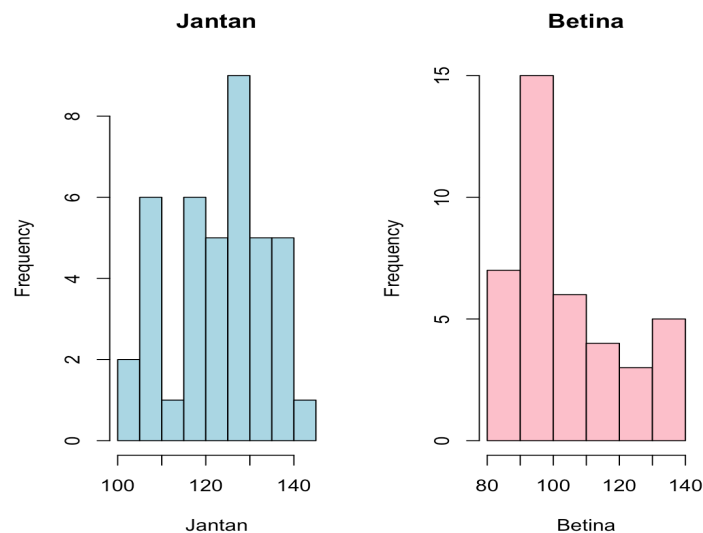
```
# Histogram untuk Harimau Jantan berdasarkan Berat
```

```
hist(Harimau[Harimau$JenisKelamin=="Jantan",]$Berat,main="Jantan",  
xlab="Jantan",col="lightblue")
```

```
# Histogram untuk Harimau Betina berdasarkan Berat
```

```
hist(Harimau[Harimau$JenisKelamin=="Betina",]$Berat,main="Betina",  
xlab="Betina",col="pink")
```

Histogram akan ditampilkan pada panel plot sebagai berikut:



12. Uji Normalitas

a) Uji normalitas pada harimau jantan

```
# Uji normalitas untuk panjang harimau jantan
```

```
shapiro.test(Harimau[Harimau$JenisKelamin=="Jantan",]$Berat)
```

Output akan ditampilkan pada console sebagai berikut:

```
Shapiro-Wilk normality test
```

```
data: Harimau[Harimau$JenisKelamin == "Jantan", ]$Berat
```

```
W = 0.96813, p-value = 0.3135
```

b) Uji normalitas pada harimau betina

```
#Uji normalitas untuk panjang harimau betina
```

```
shapiro.test(Harimau[Harimau$JenisKelamin=="Betina",]$Berat)
```

Output akan ditampilkan pada console sebagai berikut:

```
Shapiro-Wilk normality test
```

```
data: Harimau[Harimau$JenisKelamin == "Betina", ]$Berat
```

```
W = 0.89024, p-value = 0.001006
```

c) Menggunakan tapply untuk menghitung statistik deskriptif dan normalitas

```
tapply(Harimau$Berat, Harimau$JenisKelamin, stat.desc, basic=FALSE, desc=FALSE,  
norm=TRUE)
```

Output akan ditampilkan pada console sebagai berikut:

```
$Betina
```

skewness	skew.2SE	kurtosis	kurt.2SE	normtest.W	normtest.p
0.6959073	0.9308966	-0.8935652	-0.6098586	0.8902361	0.00100610

```
$Jantan
```

skewness	skew.2SE	kurtosis	kurt.2SE	normtest.W	normtest.p
-0.2195309	-0.2936606	-0.9042134	-0.6171261	0.9681281	0.3134607

Catatan, penjelasan script:

- *tapply(Harimau\$berat, Harimau\$Jeniskelamin, ...): Fungsi tapply digunakan untuk menghitung statistik terpisah berdasarkan jenis kelamin harimau.*
- *stat.desc(x, basic = FALSE, desc = FALSE, norm = TRUE): Fungsi stat.desc dari paket pastecs menghitung statistik deskriptif. Parameter basic = FALSE dan desc = FALSE digunakan untuk tidak menampilkan statistik dasar dan deskriptif yang lebih sederhana, sementara norm = TRUE menampilkan uji normalitas (Shapiro-Wilk test).*

13. Uji Homogenitas varians

Levene's Test digunakan untuk menguji kesamaan varians berdasarkan jenis kelamin

```
#Uji Levene
```

```
leveneTest(Harimau$Berat, Harimau$JenisKelamin,center=mean)
```

Output akan ditampilkan pada console sebagai berikut:

```
Levene's Test for Homogeneity of Variance (center = mean)
```

	Df	F value	Pr(>F)
group 1	1	1.5425	0.218
			78

Catatan, penjelasan script:

1. `leveneTest(response ~ group)`: Fungsi ini menguji apakah varians dari variabel `response` (dalam hal ini, `Harimau$berat`) sama di seluruh kelompok yang ditentukan oleh `group` (dalam hal ini, `Harimau$Jeniskelamin`).
2. Hasil dari `leveneTest()` akan memberikan nilai statistik dan *p-value*.
 - Jika *p-value* > 0.05, maka varians dianggap homogen (tidak ada perbedaan signifikan antar kelompok).
 - Jika *p-value* ≤ 0.05, maka varians dianggap tidak homogen (terdapat perbedaan signifikan antar kelompok).

14. Grafik Barplot

a) Mengatur tata letak plot

```
#mengatur tata letak plot
```

```
par(mfrow = c(1, 1))
```

b) Mengatur standar error

```
#menghitung standar error
```

```
Berat.se <- tapply(Harimau$Berat, Harimau$JenisKelamin, std.error)
```

```
#menampiltan standar error
```

```
Berat.se
```

c) Membuat Barplot

```
#membuat Barplot
```

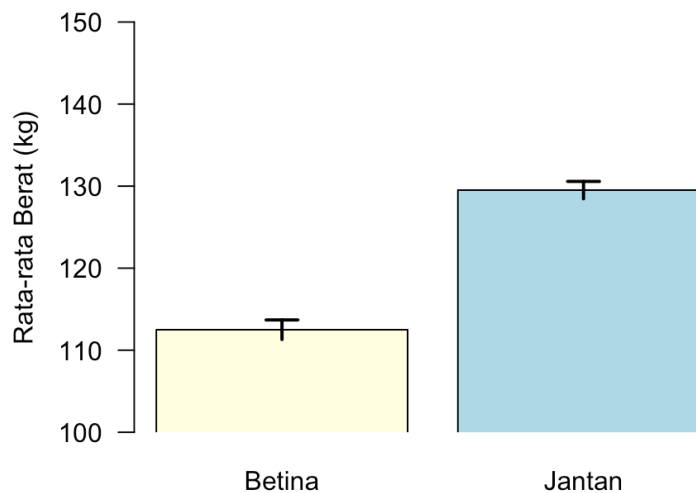
```
bar.berat <- barplot(rata.berat, col = c("lightyellow", "lightblue"), ylim = c(90, 140),  
beside = TRUE, width = 0.3, ylab = "Rata-rata Berat (kg)", las = 1, xpd = FALSE,  
names.arg = c("Betina", "Jantan"))
```

```
#membuat standar error pada bar
```



```
arrows(x0 = bar.berat, y0 = rata.berat - Berat.se, x1 = bar.berat, y1 = rata.berat +
Berat.se, angle = 90, length = 0.1, col = "black", lwd = 2)
```

Barplot akan di tampilkan pada panel plot sebagai berikut:



Catatan, penjelasan script:

- *par(mfrow = c(1, 1))*: Mengembalikan layout plotting ke satu plot per halaman.
- *length.means*: Ini adalah data rata-rata panjang (misalnya panjang harimau jantan dan betina).
- *length.se*: Ini adalah standar error dari panjang masing-masing kelompok.
- *barplot()*: Membuat barplot dengan panjang rata-rata, warna, dan label.
- *plotCI()*: Menambahkan error bars pada barplot, menggunakan standar error (*length.se*).

15. Uji T

```
#Uji T
t_test <- t.test(Harimau$Berat~Harimau$JenisKelamin, var.equal=T)
```

Output akan ditampilkan pada console sebagai berikut:

Two Sample T-Test

data: Harimau\$Berat by Harimau\$JenisKelamin

t = -10.657, df = 78, p-value < 2.2e-16

alternative hypothesis: true difference in means between group Betina and group Jantan is not equal to 0

95 percent confidence interval:

-20.19641 -13.83859

sample estimates:

mean in group Betina mean in group Jantan

112.4975

129.5150

Catatan, penjelasan script:

- *var.equal=T: Menunjukkan bahwa asumsi varians dari kedua kelompok dianggap sama.*

16. Kesimpulan

```
if (t_test$p.value < 0.05) {print("Ada perbedaan antara berat harimau jantan dengan harimau betina.")} else {print("Tidak ada perbedaan yang signifikan antara berat harimau jantan dengan harimau betina.")}
```

Output akan ditampilkan pada console sebagai berikut:

```
[1] "Ada perbedaan antara berat harimau jantan dengan harimau betina."
```

d. Paired T-Test

Contoh soal:

Seorang peneliti ingin mengetahui apakah pemberian pupuk baru dapat meningkatkan tinggi tanaman secara signifikan. Pengukuran tinggi tanaman dilakukan sebelum dan setelah pemberian pupuk pada kelompok tanaman yang sama.



Tanaman	Tinggi Sebelum (cm)	Tinggi Setelah (cm)
1	15	18
2	12	14
3	14	17
4	10	11
5	16	20
6	11	15
7	13	17
8	9	10
9	14	19
10	12	16
11	15	19
12	8	9
13	10	13
14	14	18
15	11	14

Pertanyaan:

Apakah pemberian pupuk secara signifikan meningkatkan tinggi tanaman menggunakan tingkat signifikansi 5%?

Jawaban Panduan:

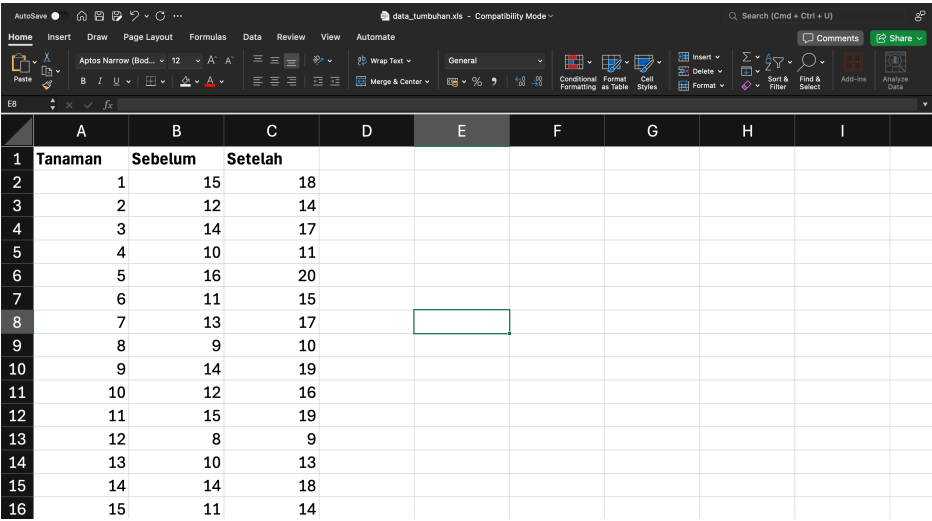
1. Hipotesis:

- H_0 : Tidak ada perbedaan signifikan antara tinggi tanaman sebelum dan sesudah pemberian pupuk (pupuk tidak meningkatkan tinggi tanaman).
- H_a : Terdapat peningkatan tinggi tanaman setelah pemberian pupuk (pupuk meningkatkan tinggi tanaman).

2. Uji Statistik: Menggunakan uji t berpasangan untuk data sebelum dan sesudah perlakuan dengan tingkat Signifikansi: $\alpha = 0,05$.

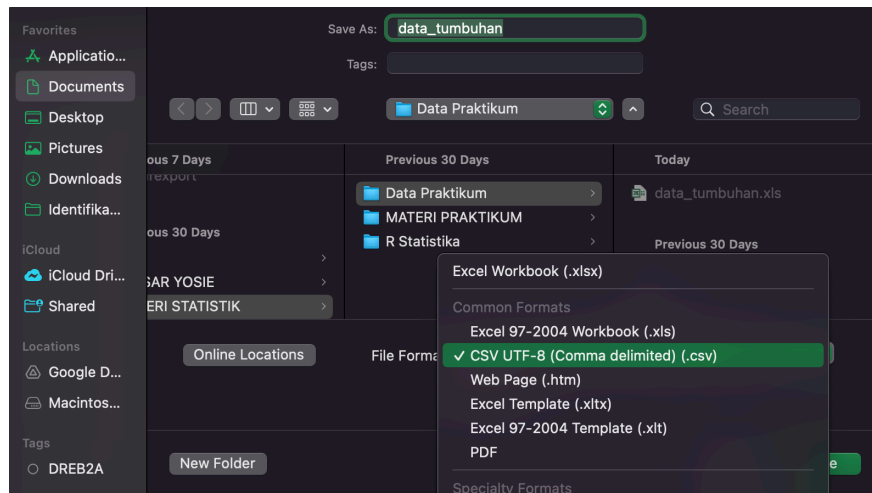
Langkah-langkah Analisis Menggunakan R software

1. Input data di Ms. Excel



	A	B	C
1	Tanaman	Sebelum	Setelah
2	1	15	18
3	2	12	14
4	3	14	17
5	4	10	11
6	5	16	20
7	6	11	15
8	7	13	17
9	8	9	10
10	9	14	19
11	10	12	16
12	11	15	19
13	12	8	9
14	13	10	13
15	14	14	18
16	15	11	14

2. Data di ketik pada excel dan di save dengan *file type* "tab delimited [*.txt]" atau "comma delimited [*.csv]"



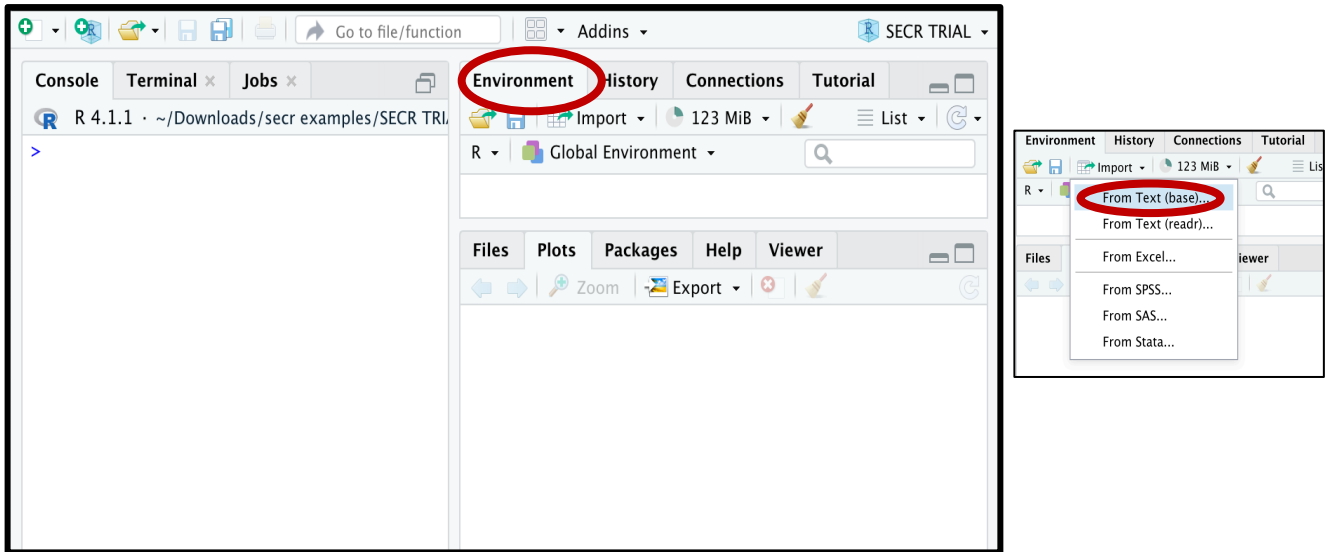
3. Memanggil data atau import data dapat dilakukan dua cara yaitu:

a. Gunakan perintah: `read.csv("path")`

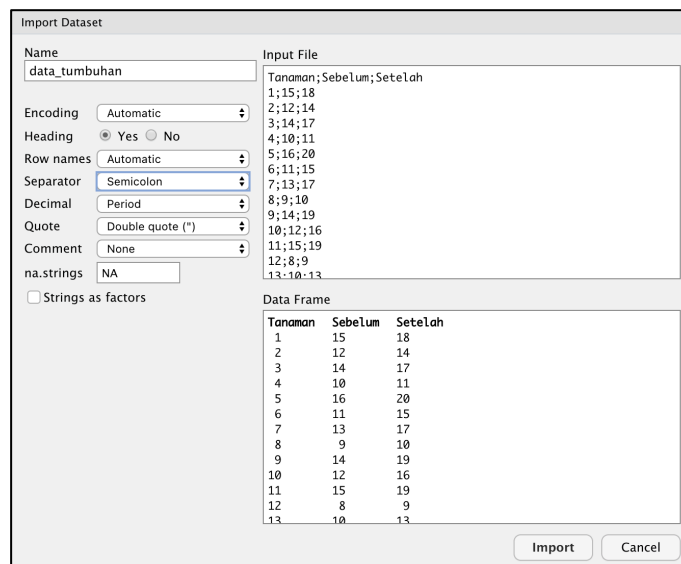
#Memanggil data

```
data_tumbuhan <- read.csv("~/Documents/MATERI STATISTIK/Data  
Praktikum/data_tumbuhan.csv", sep=";")
```

2. Buka Rstudio klik "Environment", kemudian klik "from text (base)"



4. Pastikan tampilan data yang di import sebagai berikut:



5. Packages harus sudah diinstalasi terlebih dahulu, namun apabila sudah terinstal dapat dilanjutkan ke langkah aktivasi package. Apabila belum diinstalasi gunakan langkah berikut untuk menginstal package. Instalasi package dapat dilakukan dengan instalasi beberapa package sekaligus. Namun, direkomendasikan untuk install satu per satu untuk menghindari kegagalan dalam instalasi.

#instalasi beberapa package sekaligus

```
install.packages(c("plotrix","car","pastecs", "reshape2", "fBasics"))
```

#instalasi package satu per satu

```
install.packages("plotrix")
```

```
install.packages("car")
```

```
install.packages("pastecs")
```

```
install.packages("reshape2")
```

```
install.packages("fBasics")
```

6. Setelah package terinstal, mengaktifkan package dilakukan dengan cara:

#aktivasi package

```
library(plotrix)
```

```
library(car)
```

```
library(pastecs)
```

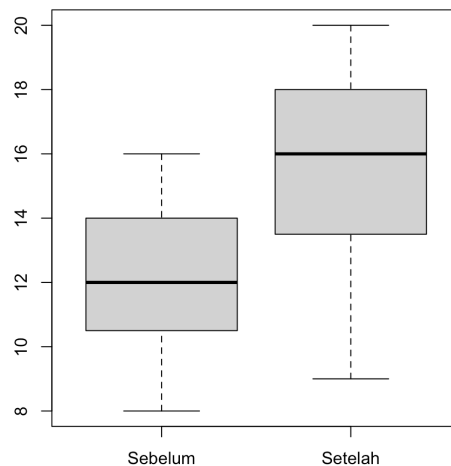
```
library(reshape2)
```

```
library(fBasics)
```

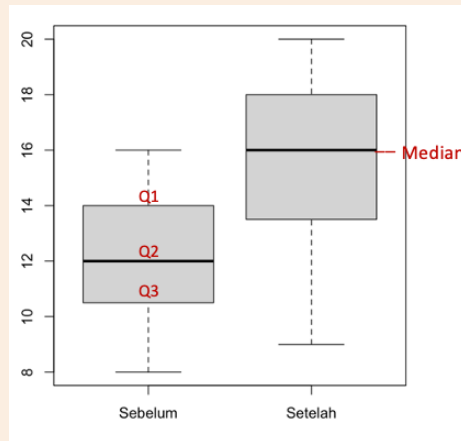
7. Membuat grafik boxplot

a) Membuat boxplot sederhana

```
boxplot(data_tumbuhan[2:3])
```



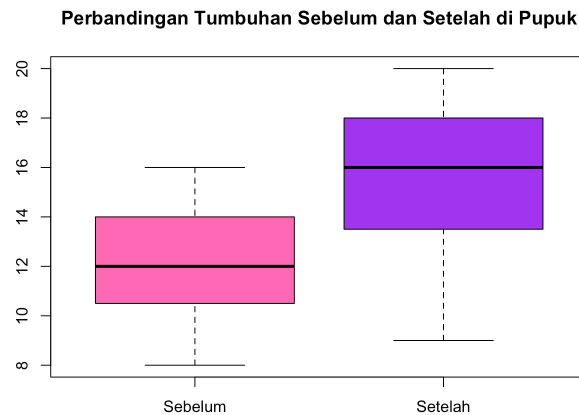
Catatan, penjelasan boxplot:



1. Kotak (box): Bagian kotak menggambarkan interkuartil (IQR), yaitu rentang dari kuartil pertama (Q1) hingga kuartil ketiga (Q3). Ini mencakup 50% tengah dari data.
 - Sebelum: Kotaknya berada dalam rentang kira-kira 10 hingga 13.
 - Setelah: Kotaknya berada dalam rentang kira-kira 14 hingga 17.
2. Garis tengah (median): Garis tebal di dalam kotak adalah median, yang merupakan nilai tengah data.
 - Sebelum: Median berada sekitar nilai 12.
 - Setelah: Median berada sekitar nilai 15.5.

b) Membuat boxplot dengan warna

```
boxplot(data_tumbuhan[2:3], col = c("hotpink", "purple"), main = "Perbandingan Tumbuhan Sebelum dan Setelah di Pupuk")
```



8. Lakukan uji normalitas dengan cara sebagai berikut:

```
#uji statistik deskriptif data tumbuhan sebelum perlakuan
stat.desc(data_tumbuhan$Sebelum, basic=F, desc=F, norm=T)
```

Output akan ditampilkan pada console sebagai berikut:

skewness	skew.2SE	kurtosis	kurt.2SE	normtest.W	normtest.p
-0.1488357	-0.1282802	-1.3313242	-0.5938655	0.9603206	0.6979687

```
#uji statistik deskriptif data tumbuhan sesudah perlakuan
stat.desc(data_tumbuhan$Setelah, basic=F, desc=F, norm=T)
```

Output akan ditampilkan pada console sebagai berikut:

skewness	skew.2SE	kurtosis	kurt.2SE	normtest.W	normtest.p
-0.4219932	-0.3637124	-1.2030784	-0.5366587	0.9401351	0.3840724

Catatan, penjelasan hasil analisis (pada data tumbuhan sebelum perlakuan, data setelah perlakuan menyesuaikan):

- *Skewness (-0,1488357): Skewness mengukur asimetri distribusi data. Nilai mendekati 0 menunjukkan bahwa distribusi data simetris. Skewness negatif berarti data sedikit condong ke kiri, artinya distribusi memiliki ekor kiri yang lebih panjang.*
- *Skew.2SE (-0,1282802): Skewness dibagi dengan dua standar error. Jika nilai ini lebih besar dari 1 atau lebih kecil dari -1, artinya skewness signifikan secara statistik. Dalam kasus ini, nilainya kecil, sehingga skewness tidak signifikan.*
- *Kurtosis (-1,3313242): Kurtosis mengukur bentuk ekor distribusi data. Nilai 3 menunjukkan distribusi normal (mesokurtik). Nilai di bawah 3 menunjukkan ekor lebih ringan (platykurtik). Nilai -1,33 menunjukkan bahwa distribusi ini memiliki ekor yang lebih ringan dari distribusi normal.*
- *Kurt.2SE (-0,5938655): Ini adalah kurtosis dibagi dengan dua standar error. Jika nilainya di luar rentang -1 hingga 1, maka kurtosis dianggap signifikan secara statistik. Nilai ini menunjukkan bahwa kurtosis tidak signifikan.*
- *Normtest.W (0,9603206): Ini adalah statistik uji dari uji normalitas*

9. Uji T Berpasangan (Paired T-Test)

#uji T berpasangan

```
t_test <- t.test(data_tumbuhan$Sebelum, data_tumbuhan$Setelah, paired=T)
```

Output akan ditampilkan pada console sebagai berikut:

Paired T-Test

data: data1\$Sebelum and data1\$Setelah

t = -9.2799, df = 14, p-value = 2.338e-07



alternative hypothesis: true difference in means is not equal to 0

-3.775441

95 percent confidence interval:

2.357892

sample estimates: mean of the differences

-3.066667

Catatan, penjelasan hasil analisis

1. t-value ($t = -9.2799$): Ini adalah nilai statistik t , yang menunjukkan seberapa besar perbedaan antara kedua kelompok dibandingkan dengan variasi di dalam kelompok. Nilai t yang sangat negatif menunjukkan perbedaan yang signifikan antara dua kelompok.
2. df (degrees of freedom, $df = 14$): Ini adalah derajat kebebasan, yang terkait dengan ukuran sampel. Karena Anda memiliki 15 pasangan data, $df = 14$ ($n - 1$).
3. p-value ($2.338e-07$): Ini adalah nilai p yang sangat kecil, yaitu 0.0000002338. Nilai p yang lebih kecil dari 0.05 berarti bahwa ada perbedaan yang signifikan secara statistik antara dua kelompok (Sebelum dan Setelah). Anda dapat menolak hipotesis nol (H_0) yang menyatakan bahwa tidak ada perbedaan antara kelompok *Sebelum* dan *Setelah*.
4. Alternative hypothesis: Hipotesis alternatif menyatakan bahwa perbedaan rata-rata antara kedua kelompok tidak sama dengan 0, dan ini didukung oleh hasil tes.
5. Confidence interval (CI): Interval kepercayaan 95% untuk perbedaan rata-rata adalah dari -3.775441 hingga -2.357892. Karena kedua batas interval kepercayaan negatif, ini menunjukkan bahwa nilai setelah

10. Kesimpulan hipotesis

#membuat kesimpulan

```
if (t_test$p.value < 0.05) {print("Ada perbedaan signifikan antara tinggi tanaman sebelum dan sesudah pemberian pupuk (pupuk meningkatkan tinggi tanaman).")} else {print("Tidak ada perbedaan signifikan antara tinggi tanaman sebelum dan sesudah pemberian pupuk (pupuk tidak meningkatkan tinggi tanaman).")}
```

Output akan ditampilkan pada console sebagai berikut:

```
[1] "Ada perbedaan signifikan antara tinggi tanaman sebelum dan sesudah pemberian pupuk (pupuk meningkatkan tinggi tanaman)."
```

e. ANOVA

Contoh soal:

Seorang peneliti ingin mengetahui apakah ada perbedaan yang signifikan pada laju perkecambahan biji tanaman berdasarkan lima perlakuan berbeda (A, B, C, D, E). Berikut adalah data jumlah biji yang berkecambah (dalam persentase) setelah 7 hari pada tiga ulangan untuk setiap perlakuan:

Perlakuan A	Perlakuan B	Perlakuan C	Perlakuan D	Perlakuan E
65	70	75	80	85
68	72	78	82	88
67	71	76	81	87
64	72	77	78	86
68	73	76	83	84
63	71	74	77	82

Pertanyaan:

Apakah perbedaan perlakuan berpengaruh signifikan terhadap persentase perkecambahan pada tingkat signifikansi 5%?

Jawaban Panduan:

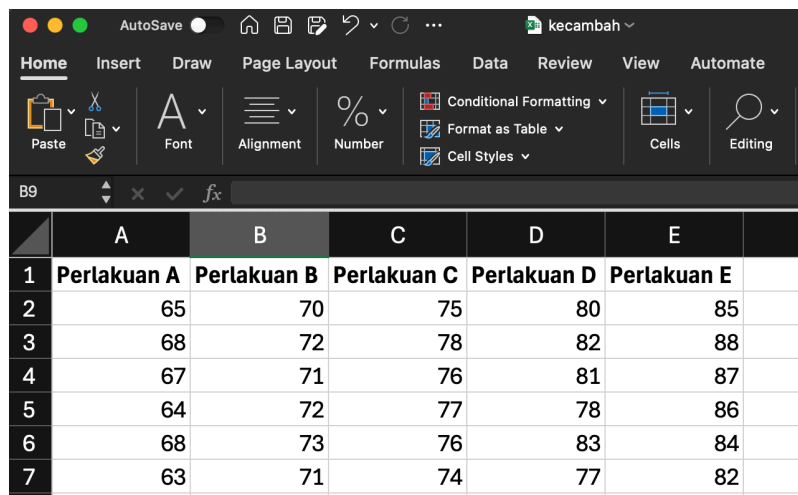
1. Hipotesis:

- H_0 : Tidak ada perbedaan signifikan dalam persentase perkecambahan antara perlakuan.
- H_1 : Ada perbedaan signifikan dalam persentase perkecambahan antara perlakuan.

2. Uji Statistik: menggunakan ANOVA perlakuan dengan tingkat signifikansi: $\alpha = 0,05$.

Langkah-langkah Analisis ANOVA Menggunakan R software

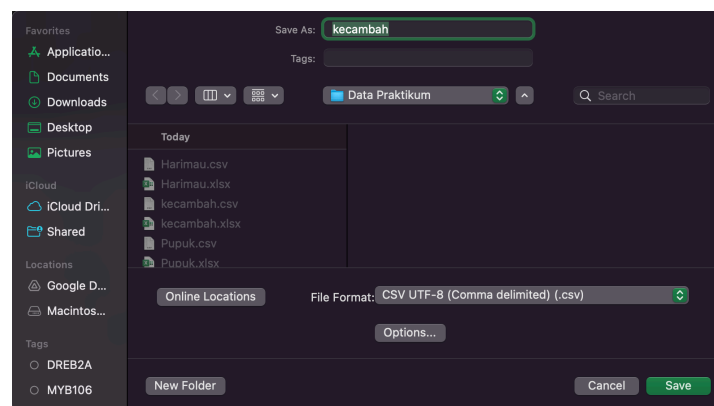
1. Input data di Ms. Excel



The screenshot shows the Microsoft Excel interface with a data table. The table has 7 rows and 6 columns. The columns are labeled 'Perlakuan A', 'Perlakuan B', 'Perlakuan C', 'Perlakuan D', and 'Perlakuan E'. The rows are numbered 1 to 7. The data values are as follows:

	A	B	C	D	E
1	Perlakuan A	Perlakuan B	Perlakuan C	Perlakuan D	Perlakuan E
2	65	70	75	80	85
3	68	72	78	82	88
4	67	71	76	81	87
5	64	72	77	78	86
6	68	73	76	83	84
7	63	71	74	77	82

2. Data di ketik pada excel dan di save dengan file type "tab delimited [*.txt]" atau "comma delimited [*.csv]"

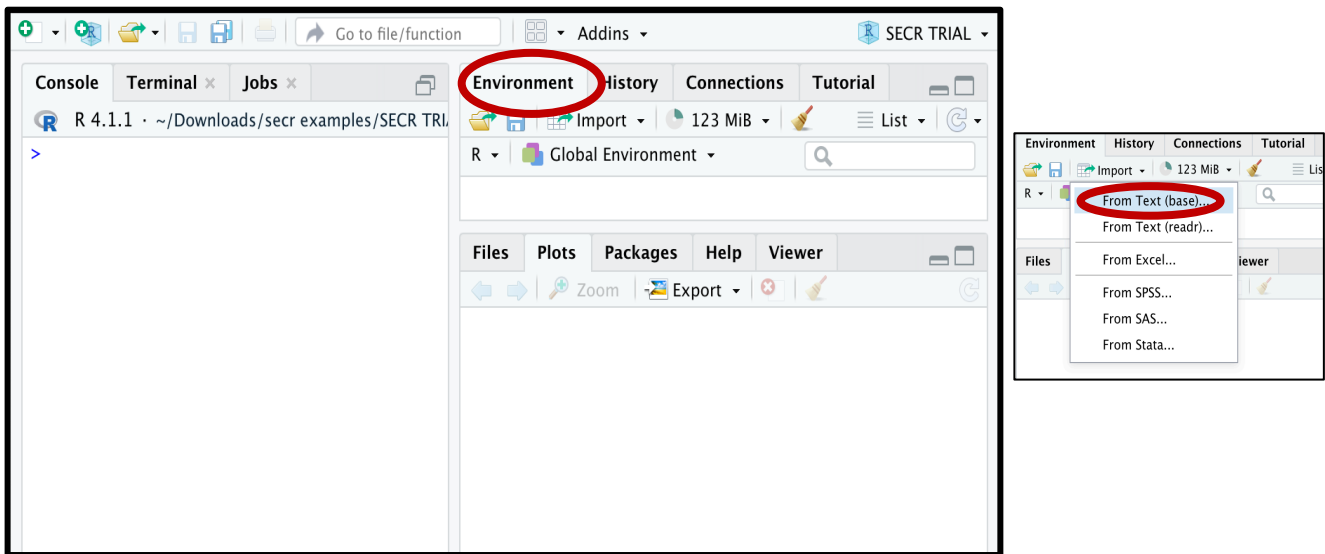


3. Memanggil data atau import data dapat dilakukan dua cara yaitu:

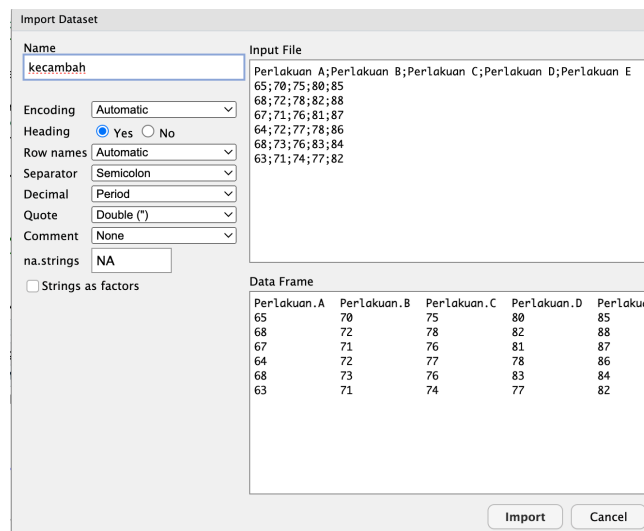
a) Gunakan perintah: `read.csv("path")`

```
kecambah <- read.csv("~/Documents/MATERI STATISTIK/Data  
Praktikum/kecambah.csv", sep=";")
```

b) Buka Rstudio klik "Environment", kemudian klik "from text (base)"



4. Pastikan tampilan data yang di import sebagai berikut:



5. Packages harus sudah diinstalasi terlebih dahulu, namun apabila sudah terinstal dapat dilanjutkan ke langkah aktivasi package. Apabila belum diinstalasi gunakan langkah berikut untuk menginstal package. Instalasi package dapat dilakukan dengan instalasi beberapa package sekaligus. Namun, direkomendasikan untuk install satu per satu untuk menghindari kegagalan dalam instalasi.

```
#instalasi beberapa package sekaligus
```

```
install.packages(c("plotrix","car","pastecs", "reshape2", "fBasics"))
```

```
#instalasi package satu per satu
```

```
install.packages("plotrix")
```

```
install.packages("car")
```

```
install.packages("pastecs")
```

```
install.packages("reshape2")
```

```
install.packages("fBasics")
```

6. Setelah package terinstal, mengaktifkan package dilakukan dengan cara:

```
#aktivasi package
```

```
library(plotrix)
```

```
library(car)
```

```
library(pastecs)
```

```
library(reshape2)
```

```
library(fBasics)
```

7. Merubah data format lebar (*wide*) menjadi format panjang (*long*)

```
#Mengubah format lebar menjadi format panjang
```

```
kecambah.melt <- melt(kecambah)
```

```
kecambah.melt
```

```
colnames(kecambah.melt) <- c("perlakuan", "persentase")
```

```
# Menghapus nilai NA dari kolom 'persentase'
```

```
kecambah.melt.clean <- kecambah.melt[!is.na(kecambah.melt$persentase), ]
```

```
# Menampilkan data yang sudah bersih dari NA
```

```
head(kecambah.melt.clean)
```

Output akan ditampilkan pada console sebagai berikut:

```
perlakuan persentase
```

```
1 Perlakuan.A      65
```

```
2 Perlakuan.A      68
```

```
3 Perlakuan.A      67
```

```
4 Perlakuan.A      64
```

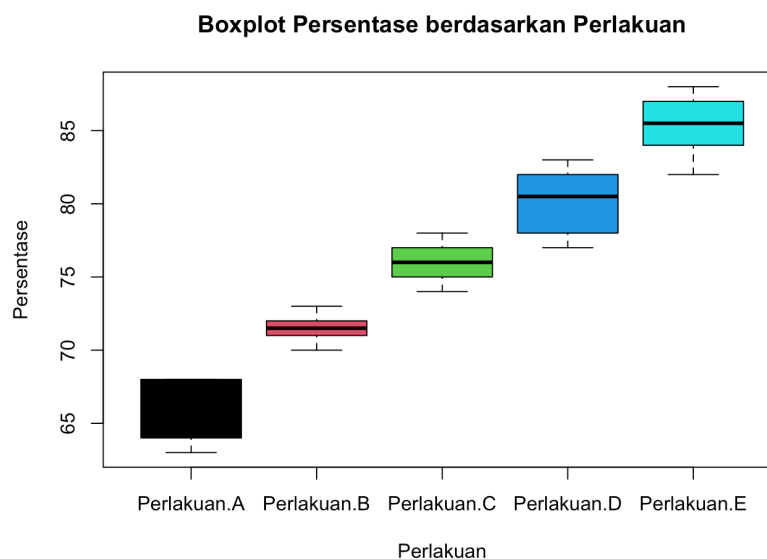
```
5 Perlakuan.A      68
```

```
6 Perlakuan.A      63
```

8. Membuat boxplot

```
#Membuat boxplot
```

```
boxplot(persentase ~ perlakuan, data = kecambah.melt.clean, main = "Boxplot  
Persentase berdasarkan Perlakuan", xlab = "Perlakuan", ylab = "Persentase", col =  
1:5)
```



9. Uji ANOVA

```
#Uji ANOVA
anova_result <- aov(persentase ~ perlakuan, data = kecambah.melt.clean)

#Menampilkan ringkasan hasil ANOVA
summary(anova_result)

#Mendapatkan nilai p
p_value <- summary(anova_result)[[1]][1, "Pr(>F)"][1, 1]

# Memeriksa apakah ada perbedaan yang signifikan
alpha <- 0.05 # Level signifikansi
if (p_value < alpha) { cat("Ada perbedaan yang signifikan antara kelompok perlakuan
(p =", p_value, ").\n")} else {cat("Tidak ada perbedaan yang signifikan antara
kelompok perlakuan (p =", p_value, ").\n")}
```

Output akan ditampilkan pada console sebagai berikut:

```
Ada perbedaan yang signifikan antara kelompok perlakuan (p = 8.03158e-15 ).
```

10. Uji Bonferroni

```
#Uji post-hoc dengan uji Bonferoni
bonferroni_result <- pairwise.t.test(kecambah.melt.clean$persentase,
kecambah.melt.clean$perlakuan, p.adjust.method = "bonferroni")

#Menampilkan hasil Uji Bonferroni
print(bonferroni_result)
```

Output akan ditampilkan pada console sebagai berikut:

```
Pairwise comparisons using t tests with pooled SD
```

```
data: kecambah.melt.clean$persentase and kecambah.melt.clean$perlakuan
```

	Perlakuan.A	Perlakuan.B	Perlakuan.C	Perlakuan.D
Perlakuan.B	0.00021	-	-	-
Perlakuan.C	1.2e-08	0.00343	-	-
Perlakuan.D	9.1e-12	2.5e-07	0.00755	-
Perlakuan.E	8.5e-15	2.0e-11	6.2e-08	0.00070

P value adjustment method: Bonferroni

11. Uji Tukey

```
#Uji Tukey
tukey_result <- TukeyHSD(anova_result)

#Menampilkan hasil Uji Tukey
print(tukey_result)
```

Output akan ditampilkan pada console sebagai berikut:

```
95% family-wise confidence level

Fit: aov(formula = persentase ~ perlakuan, data = kecambah.melt.clean)

$perlakuan
      diff      lwr      upr      p adj
Perlakuan.B-Perlakuan.A  5.666667  2.4764079  8.856925 0.0001915
Perlakuan.C-Perlakuan.A 10.166667  6.9764079 13.356925 0.0000000
Perlakuan.D-Perlakuan.A 14.333333 11.1430746 17.523592 0.0000000
Perlakuan.E-Perlakuan.A 19.500000 16.3097413 22.690259 0.0000000
Perlakuan.C-Perlakuan.B  4.500000  1.3097413  7.690259 0.0028950
Perlakuan.D-Perlakuan.B  8.666667  5.4764079 11.856925 0.0000002
```



Perlakuan.E-Perlakuan.B	13.833333	10.6430746	17.023592	0.0000000
Perlakuan.D-Perlakuan.C	4.166667	0.9764079	7.356925	0.0061798
Perlakuan.E-Perlakuan.C	9.333333	6.1430746	12.523592	0.0000001
Perlakuan.E-Perlakuan.D	5.166667	1.9764079	8.356925	0.0006164



DAFTAR PUSTAKA

- Gasperz, V. 1989. Metode perancangan.
- Hsu H & Lachenbruch, PA. 2014. Paired t test. *Wiley StatsRef: statistics reference online*
- Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A., & Team RC. 2020. Package 'caret'. The R Journal, 223(7), 48.
- Wickham H. 2011. ggplot2. *Wiley interdisciplinary reviews: computational statistics*, 3(2), 180-185.
- Wickham, H., Francois, R., Henry, L., & Müller, K. (2014). dplyr. *A Grammar of Data Manipulation 2020* [Last accessed on 2020 Aug 12] Available from, Rproject.



Modul Praktikum

Statistika Modern dengan R software

Modul ini disusun sebagai panduan belajar bagi mahasiswa yang tengah mempelajari statistika dasar, baik untuk kebutuhan akademik maupun penerapan praktis di lapangan. Dengan pendekatan yang praktis, modul ini dirancang untuk memberikan pemahaman mendalam mengenai konsep statistik, teknik analisis data dengan R software, dan penerapannya dalam bidang Kehutanan. Pembelajaran statistika diharapkan dapat menjadi landasan penting bagi mahasiswa dalam meningkatkan kemampuan analisis data dan mendukung pengambilan keputusan berbasis data yang sangat dibutuhkan di era modern.