


<p><b>Nama:</b> April Iesa Farsilis</p> <p><b>NIM:</b> 065002300005</p>	 <p><b>UNIVERSITAS TRISAKTI</b></p> <p><b>PRAKTIKUM STATISTIKA</b></p>	<p><b>MODUL 10 STATISTIKA</b></p> <p><b>Nama Dosen:</b> Dedy Sugiarto</p>
<p><b>Hari/Tanggal: Rabu, 5 Juni 2024</b></p>		<p><b>Nama Aslab:</b></p> <ol style="list-style-type: none"> <li>1. Tarum Widyasti P (064002200027)</li> <li>2. Kharisma Maulida S (064002200024)</li> </ol>

## MODUL 10

### Analysis of Variance (ANOVA)

---

#### Teori Singkat

Analisis ragam atau analysis of variance (ANOVA) merupakan teknik statistik yang dapat digunakan untuk menguji perbedaan rata-rata antar lebih dari 2 grup sampel. Teknik ANOVA sesungguhnya terbagi menjadi beberapa jenis antara lain ANOVA satu arah (one-way ANOVA), ANOVA dua arah (two-way ANOVA), ANOVA tiga arah (three-way ANOVA). Pada sesi ini hanya akan dibahas mengenai teknik ANOVA satu arah. Teknik ANOVA juga dipakai dalam kasus analisis data eksperimen untuk meneliti pengaruh dari baik dari satu faktor (variabel bebas) maupun beberapa faktor terhadap suatu variabel respon (variabel terikat).

#### Lab Setup

Hal yang harus disiapkan dan dilakukan oleh praktikan untuk menjalankan praktikum modul ini.

1. Menginstall library yang dibutuhkan untuk mengerjakan modul.
2. Menjalankan R Studio.
3. Menjalankan Jupyter
4. Menjalankan Excel

## ELEMEN KOMPETENSI I

Tiga macam metode pencegahan terhadap korosi dari suatu produk, dicoba efektivitasnya. Hasilnya berupa kedalaman korosi (dalam 0.001 inch) adalah sebagai berikut :

Metode A	77	54	67	74	71
Metode B	60	41	59	65	62
Metode C	49	52	69	47	56

Dengan menggunakan alpha 0.05, ujilah bahwa ketiga metode tersebut mempunyai pengaruh yang sama terhadap pencegahan korosi :

**Ubah data menjadi data bertumpuk pada excel**

metode(X)	korosi(Y)
a	77
a	54
a	67
a	74
a	71
b	60
b	41
b	59
b	65
b	62
c	49
c	52
c	69
c	47
c	56

**Script R:**

```
df_nama = read.delim("clipboard")
View(df_nama)
Head(df_nama)
```

```
> df_april=read.delim("clipboard")
> View(df_april)
> head(df_april)
  metode.X. korosi.Y.
1         a         77
2         a         54
3         a         67
4         a         74
5         a         71
6         b         60
```

### Analisis Anova

Analisis keseluruhan terhadap ada atau tidaknya perbedaan pada metode

#### Script R:

```
> model <- aov(korosi.y.~metode.x., data=df_nama)
> summary(model)
```

```
> model <- aov(korosi.Y.~metode.X., data=df_april)
> summary(model)
          Df Sum Sq Mean Sq F value Pr(>F)
metode.X.    2  548.8   274.40    3.348 0.0699 .
Residuals   12  983.6    81.97
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> tukey.test <- TukeyHSD(model)
```

#### Interpretasi (minimal 4 baris)

membaca data clipboard metode x dan korosi y, mengetahui hasil dengan menampilkan nilai df, sum, mean,value dari hasil hitung data data metode x korosi y dan residualsnya .

### Analisis Tukey test

Analisis ada atau tidaknya perbedaan antara 2 metode

#### Script R:

```
tukey.test <- TukeyHSD(model)
tukey.test
```

```

> tukey.test <- TukeyHSD(model)
> tukey.test
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = korosi.Y. ~ metode.X., data = df_april)

$metode.X.
      diff      lwr      upr      p adj
b-a -11.2 -26.47609  4.076092 0.1657711
c-a -14.0 -29.27609  1.276092 0.0736936
c-b  -2.8 -18.07609 12.476092 0.8778031

> |

```

### Interpretasi (minimal 4 baris)

nilai p sebesar 0.0699, pengaruh tidak signifikan pada level 5% ( $p > 0.05$ ). perbedaan 95% di antara beberapa metode, yang ditunjukkan oleh interval kepercayaan yang tidak melintasi nol. ANOVA tidak menemukan perbedaan signifikan secara keseluruhan, uji Tukey menunjukkan adanya perbedaan tertentu dimetode yang digunakan

### Screenshoot full screen

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
df_april df_nama
1 df_april=read.delim("clipboard")
2 View(df_april)
3 head(df_april)
4 model <- aov(korosi.Y.~metode.X., data=df_april)
5 summary(model)
6 tukey.test <- TukeyHSD(model)
7 tukey.test
8
1:1 (Top Level)
Console Terminal Background Jobs
R 4.3.2 ~./
> df_april=read.delim("clipboard")
> View(df_april)
> head(df_april)
metode.X. korosi.Y.
1 a 77
2 a 54
3 a 67
4 a 74
5 a 71
6 b 60
> model <- aov(korosi.Y.~metode.X., data=df_april)
> summary(model)
          Df Sum Sq Mean Sq F value Pr(>F)
metode.X.  2  548.8   274.40   3.348 0.0699 .
Residuals 12  983.6    81.97
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> tukey.test <- TukeyHSD(model)
> tukey.test
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = korosi.Y. ~ metode.X., data = df_april)

$metode.X.
      diff      lwr      upr      p adj
b-a -11.2 -26.47609  4.076092 0.1657711
c-a -14.0 -29.27609  1.276092 0.0736936
c-b  -2.8 -18.07609 12.476092 0.8778031

```

### Python

**Source code:**

```
import numpy as np
import scipy.stats as stats

# Data waktu produksi dari ketiga mesin
metode_A = np.array([77, 54, 67, 74, 71])
metode_B = np.array([60, 41, 59, 65, 62])
metode_C = np.array([49, 52, 69, 47, 56])

# Gabungkan data ke dalam satu array
data_tarum = [metode_A, metode_B, metode_C]

# Hitung ANOVA menggunakan scipy.stats
f_statistic, p_value = stats.f_oneway(metode_A, metode_B, metode_C)

# Tampilkan hasil
print(f"Nilai F: {f_statistic}")
print(f"Nilai p: {p_value}")
```

**Output:**

```
: import numpy as np
import scipy.stats as stats

# Data waktu produksi dari ketiga mesin
metode_A = np.array([77, 54, 67, 74, 71])
metode_B = np.array([60, 41, 59, 65, 62])
metode_C = np.array([49, 52, 69, 47, 56])

# Gabungkan data ke dalam satu array
data_tarum = [metode_A, metode_B, metode_C]

# Hitung ANOVA menggunakan scipy.stats
f_statistic, p_value = stats.f_oneway(metode_A, metode_B, metode_C)

# Tampilkan hasil
print(f"Nilai F: {f_statistic}")
print(f"Nilai p: {p_value}")

Nilai F: 3.347702318015454
Nilai p: 0.06993237223084404
```

**Interpretasi (minimal 4 baris)**

Nilai p sebesar 0.06993237223084404, Nilai F 3.3477023188015454, menunjukkan bahwa tidak ada perbedaan signifikan di antara ketiga metode pada level signifikansi 5% ( $p > 0.05$ ), tidak ada bukti yang cukup untuk menyatakan bahwa memiliki perbedaan signifikan.

### Excel:

K31									
	A	B	C	D	E	F	G	H	I
1	Metode A	77	54	67	74	71		metode(X)	korosi(Y)
2	Metode B	60	41	59	65	62		a	77
3	Metode C	49	52	69	47	56		a	54
4								a	67
5								a	74
6	Metode A	Metode B	Metode C					a	71
7		77	60	49				b	60
8		54	41	52				b	41
9		67	59	69				b	59
10		74	65	47				b	65
11		71	62	56				b	62
12	Mean	68,6	57,4	54,6	60,2			c	49
13	VAariance	80,3	89,3	76,3				c	52
14								c	69
15		SSTR	548,8	SSE	983,6			c	47
16		MSTR	274,4	MSE	81,96666667			c	56
17		F	3,348						
18									Nama: April Iesa Farsilis
19									
20									> model <- aov(korosi.Y~metode.X., data=df_april)
21									> summary(model)
22									Df Sum Sq Mean Sq F value Pr(>F)
23									metode.X.    2   548.8   274.40   3.348 0.0699 .
24									Residuals   12   983.6    81.97
25									---
26									Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
27									> tukey.test <- TukeyHSD(model)
28									

### Interpretasi (minimal 4 baris)

p-value yang telah dihitung sebelumnya (0.0699) menunjukkan bahwa tidak ada perbedaan signifikan di antara metode pada tingkat signifikansi 5% ( $p > 0.05$ ). meskipun ada perbedaan rata-rata waktu produksi antara metode, perbedaan ini tidak signifikan secara statistik pada tingkat 5%.

## ELEMEN KOMPETENSI II

Gunakan dataset plant growth yang telah tersedia di R untuk meneliti pengaruh beberapa treatment (perlakuan) terhadap tingkat pertumbuhan tanaman.

### Menampilkan Data

### Script R:

```
df_nama=PlantGrowth
View(df_nama)
head(df_nama)

df_aurel = read.delim("clipboard")
View(df_april)
head(df_april)
```

**Output:**

```

> df_nama=PlantGrowth
> View(df_nama)
> head(df_nama)
  weight group
1   4.17  ctrl
2   5.58  ctrl
3   5.18  ctrl
4   6.11  ctrl
5   4.50  ctrl
6   4.61  ctrl
> df_april=read.delim("clipboard")
> View(df_april)
> head(df_april)
  metode.X. korosi.Y.
1        ctrl      4.17
2        ctrl      4.50
3        ctrl      4.53
4        ctrl      4.61
5        ctrl      5.14
6        ctrl      5.17

```

**Analisis Anova****Script R:**

```

model <- aov(korosi.y~metode.x., data=df_april)
summary(model)

```

**Output:**

```

> model <- aov(korosi.Y~metode.X., data=df_april)
> summary(model)
              Df Sum Sq Mean Sq F value Pr(>F)
metode.X.      2  3.766   1.8832   4.846 0.0159 *
Residuals    27 10.492   0.3886

```

**Interpretasi (minimal 4 baris)**

Nilai  $p = 0.0159$  ( $p < 0.05$ ). Metode yang digunakan memiliki pengaruh yang signifikan terhadap tingkat korosi. Dengan F-value sebesar 4.846 dan tingkat kepercayaan 95%, kita menolak hipotesis nol yang menyatakan tidak ada perbedaan antar metode.

## Analisis Tukey test

### Source code:

```
tukey.test <- TukeyHSD(model)
tukey.test
```

### Output:

```
> tukey.test <- TukeyHSD(model)
> tukey.test
  Tukey multiple comparisons of means
    95% family-wise confidence level

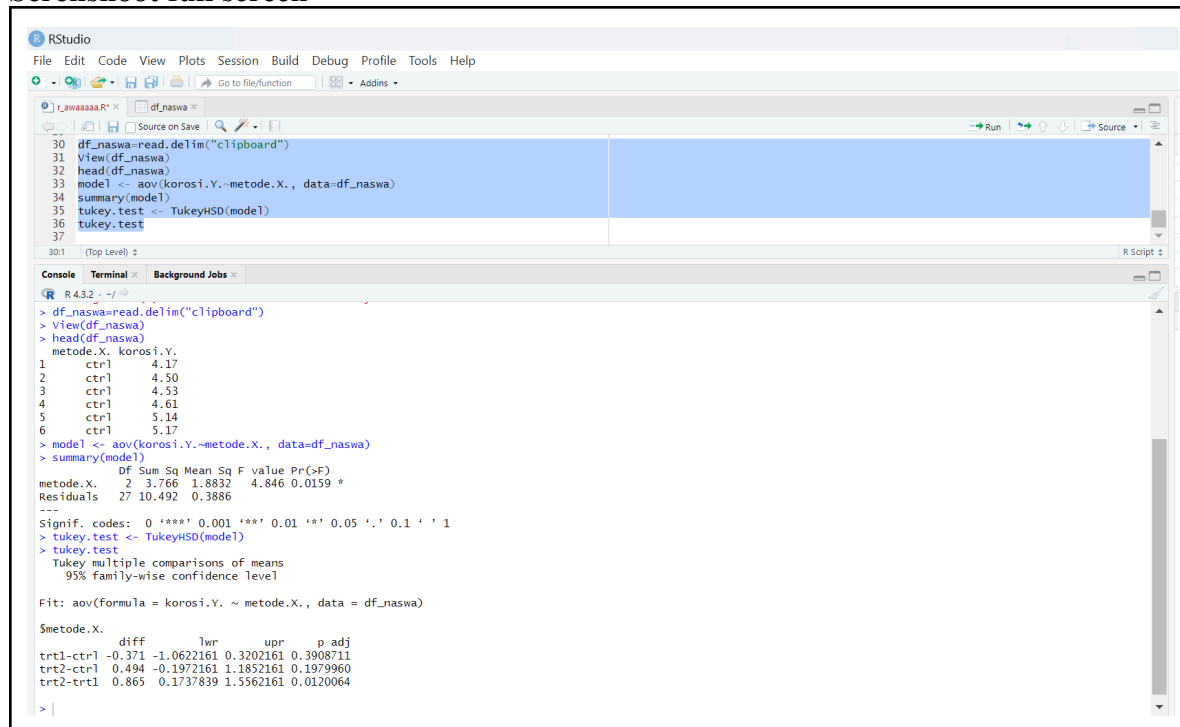
Fit: aov(formula = korosi.Y. ~ metode.X., data = df_april)

$metode.X.
      diff      lwr      upr      p adj
trt1-ctrl -0.371 -1.0622161 0.3202161 0.3908711
trt2-ctrl  0.494 -0.1972161 1.1852161 0.1979960
trt2-trt1  0.865  0.1737839 1.5562161 0.0120064
```

### Interpretasi (minimal 4 baris)

Selang kepercayaan untuk perbedaan antara trt2 dan trt1 tidak mencakup nol (0.1737839 hingga 1.5562161), mendukung hasil signifikan. trt2 memiliki efek yang berbeda secara signifikan dibandingkan trt1 dalam pengujian korosi, sementara perbedaan dengan kontrol tidak signifikan.

### Screenshoot full screen





**Python:****Source code:**

```
import numpy as np
import scipy.stats as stats

# Data waktu produksi dari ketiga mesin
ctrl = np.array([ 4.17, 4.50, 4.53, 4.61, 5.14, 5.17, 5.18, 5.33, 5.58, 6.11])
trt1 = np.array([3.59, 3.83, 4.17, 4.32, 4.41, 4.69, 4.81, 4.89, 5.87, 6.03])
trt2 = np.array([4.92, 5.12, 5.26, 5.29, 5.37, 5.50, 5.54, 5.80, 6.15, 6.31])

# Gabungkan data ke dalam satu array
data_tarum = [ctrl, trt1, trt2]

# Hitung ANOVA menggunakan scipy.stats
f_statistic, p_value = stats.f_oneway(ctrl, trt1, trt2)

# Tampilkan hasil
print(f"Nilai F: {f_statistic}")
print(f"Nilai p: {p_value}")
```

**Output:**

```
import numpy as np
import scipy.stats as stats

# Data waktu produksi dari ketiga mesin
ctrl = np.array([ 4.17, 4.50, 4.53, 4.61, 5.14, 5.17, 5.18, 5.33, 5.58, 6.11])
trt1 = np.array([3.59, 3.83, 4.17, 4.32, 4.41, 4.69, 4.81, 4.89, 5.87, 6.03])
trt2 = np.array([4.92, 5.12, 5.26, 5.29, 5.37, 5.50, 5.54, 5.80, 6.15, 6.31])

# Gabungkan data ke dalam satu array
data_tarum = [ctrl, trt1, trt2]

# Hitung ANOVA menggunakan scipy.stats
f_statistic, p_value = stats.f_oneway(ctrl, trt1, trt2)

# Tampilkan hasil
print(f"Nilai F: {f_statistic}")
print(f"Nilai p: {p_value}")

Nilai F: 4.846087862380133
Nilai p: 0.01590995832562293
```

**Interpretasi (minimal 4 baris)**

ANOVA yang dilakukan nilai p sebesar 0.0159, nilai p kurang dari 0.05, kita menolak hipotesis nol yang menyatakan bahwa rata-rata waktu produksi ketiga mesin adalah sama. Bahwa terdapat perbedaan signifikan pada rata-rata waktu produksi antara ketiga mesin.

**Excel:**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	CTRL	4.17	4.50	4.53	4.61	5.14	5.17	5.18	5.33	5.58	6.11	metode(X)	korosi(Y)	
2	TRT1	3.59	3.83	4.17	4.32	4.41	4.69	4.81	4.89	5.87	6.03	ctrl	4.17	
3	TRT2	4.92	5.12	5.26	5.29	5.37	5.50	5.54	5.80	6.15	6.31	ctrl	4.50	
4												ctrl	4.53	
5												ctrl	4.61	
6		ctrl	trt1	trt2								ctrl	5.14	
7		4.17	3.59	4.92								ctrl	5.17	
8		4.50	3.83	5.12								ctrl	5.18	
9		4.53	4.17	5.26								ctrl	5.33	
10		4.61	4.32	5.29								ctrl	5.58	
11		5.14	4.41	5.37								ctrl	6.11	
12		5.17	4.69	5.50								trt1	3.59	
13		5.18	4.81	5.54								trt1	3.83	
14		5.33	4.89	5.80								trt1	4.17	
15		5.58	5.87	6.15								trt1	4.32	
16		6.11	6.03	6.31								trt1	4.41	
17	mean	5.032	4.661	5.526	5.073							trt1	4.69	
18	variasi	0.339995556	0.629921111	0.195871111								trt1	4.81	
19												trt1	4.89	
20	SSTR	3.76634		SSE	10.49209							trt1	5.87	
21	MSTR	1.88317		MSE	0.388595926							trt1	6.03	
22	F	4.846										trt2	4.92	
23												trt2	5.12	
24												trt2	5.26	
25												trt2	5.29	
26												trt2	5.37	
27												trt2	5.5	
28												trt2	5.54	
29												trt2	5.8	
30												trt2	6.15	
31												trt2	6.31	

### Interpretasi (minimal 4 baris)

Nilai p kurang dari 0.05, nilai F sebesar 4.846 yang menunjukkan adanya variasi signifikan antar kelompok. kita menolak hipotesis nol bahwa rata-rata waktu produksi ketiga mesin sama.terdapat perbedaan yang signifikan pada waktu produksi antara ketiga mesin.

### Sumber :

<http://www.sthda.com/english/wiki/one-way-anova-test-in-r>

<https://rpubs.com/aaronsc32/post-hoc-analysis-tukey>

### CEK LIST (✓)

1. Memahami analisis ragam.

(✓)

### GITHUB

<https://github.com/lesa04/prob10.git>

### KESIMPULAN

Teknik ANOVA juga dipakai dalam kasus analisis data eksperimen untuk meneliti pengaruh dari baik dari satu faktor (variabel bebas) maupun beberapa faktor terhadap suatu variabel respon (variabel terikat). Jenis ANOVA satu arah (one-way ANOVA), ANOVA dua arah (two-way ANOVA), ANOVA tiga arah (three-way ANOVA). Nilai p

kurang dari 0.05, nilai F sebesar 4.846 adanya variasi signifikan, menolak hipotesis nol bahwa rata-rata waktu produksi ketiga mesin sama. Terdapat perbedaan yang signifikan pada waktu produksi antara ketiga mesin. nilai F sebesar 4.846 dan nilai p sebesar 0.0159 jadi nilai p kurang dari 0.05, menolak hipotesis nol yang menyatakan bahwa rata-rata waktu produksi ketiga mesin adalah sama. Terdapat perbedaan rata-rata waktu ketiga mesin tersebut.

### FORM UMPAN BALIK

<b>Elemen Kompetensi</b>	<b>Tingkat Kesulitan</b>	<b>Tingkat Ketertarikan</b>	<b>Waktu Penyelesaian (menit)</b>
<b>Memahami analisis ragam.</b>	<b>sulit</b>	<b>sangat tertarik</b>	<b>30 menit</b>

#### Keterangan Tingkat Kesulitan

- 1: Sangat Mudah
- 2: Mudah
- 3: Biasa
- 4: Sulit
- 5: Sangat Sulit

#### Keterangan Tingkat Ketertarikan

- 1: Tidak Tertarik
- 2: Cukup Tertarik
- 3: Tertarik
- 4: Sangat Tertarik