


<b>Nama:</b> <b>Nadhira Aninidta</b> <b>Ralena</b> <b>NIM:</b> <b>065002300021</b>	 <b>UNIVERSITAS TRISAKTI</b>  <b>PRAKTIKUM</b> <b>STATISTIKA</b>	<b>MODUL 10</b> <b>STATISTIKA</b>  <b>Nama Dosen:</b> Dedy Sugiarto
<b>Hari/Tanggal: Rabu, 5</b> <b>Juni 2024</b>		<b>Nama Aslab:</b>  1. Tarum Widyasti P (064002200027)  2. Kharisma Maulida S (064002200024)

## **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

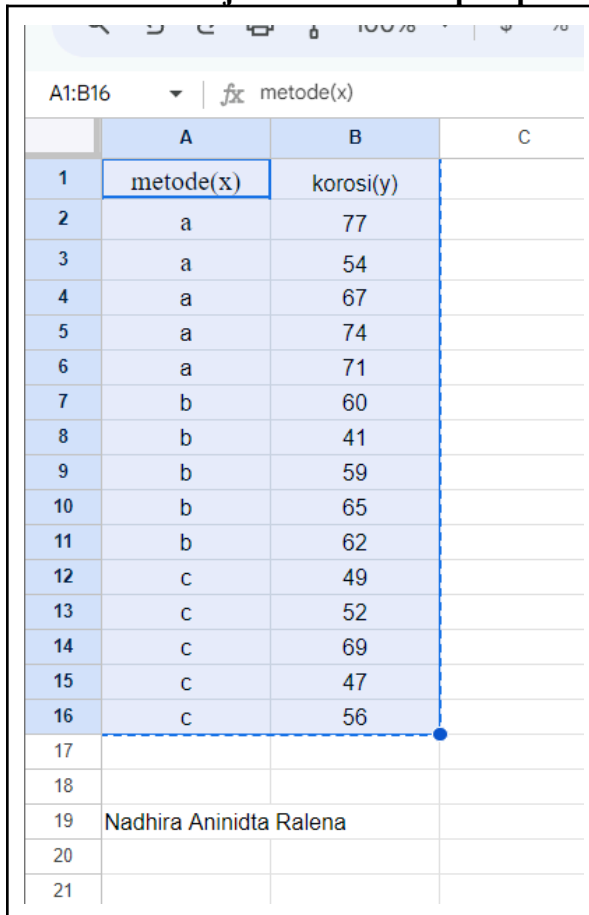
**ELEMEN KOMPETENSI I**

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 :

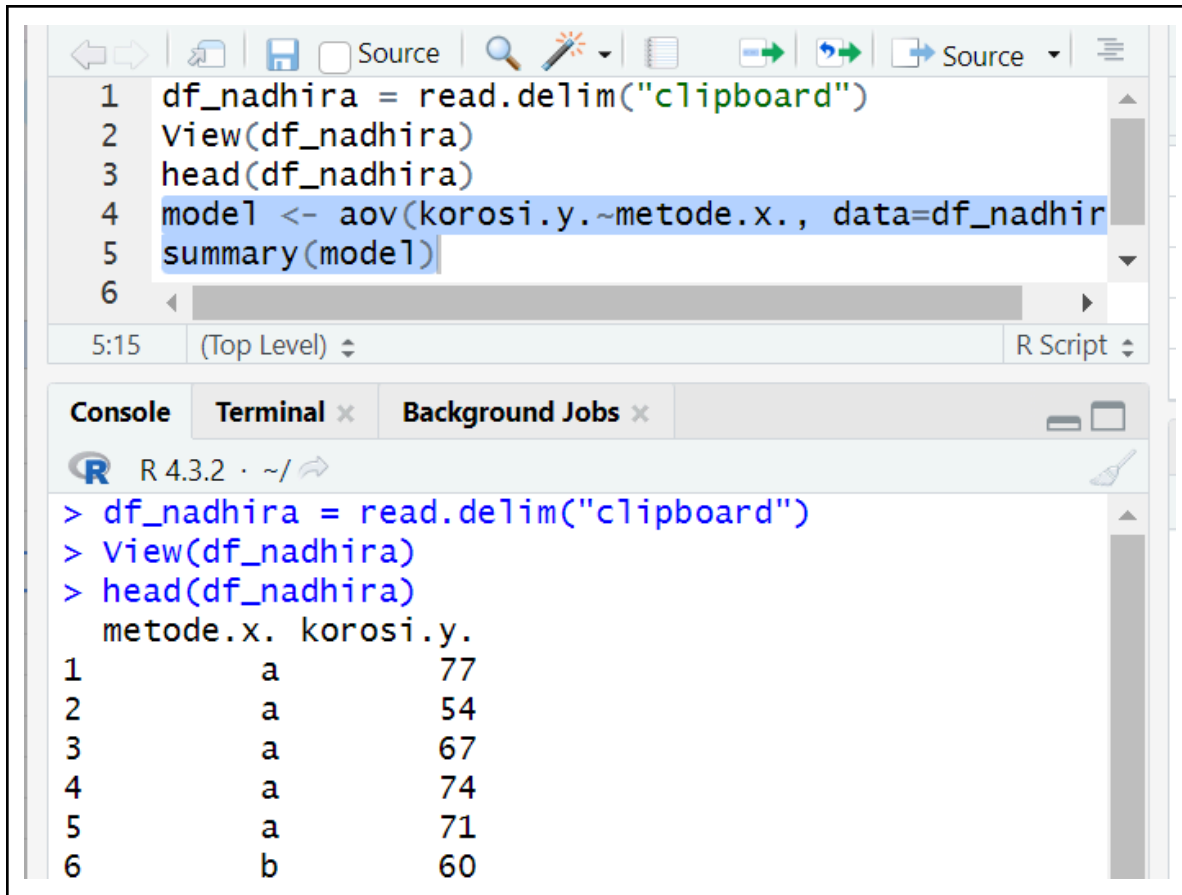
**Ubah data menjadi data bertumpuk pada excel**



	A	B	C
1	metode(x)	korosi(y)	
2	a	77	
3	a	54	
4	a	67	
5	a	74	
6	a	71	
7	b	60	
8	b	41	
9	b	59	
10	b	65	
11	b	62	
12	c	49	
13	c	52	
14	c	69	
15	c	47	
16	c	56	
17			
18			
19	Nadhira Aninidta Ralena		
20			
21			

**Script R:**

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



The screenshot shows the R Studio environment. The top pane is the Source editor, containing an R script with the following code:

```
1 df_nadhira = read.delim("clipboard")
2 View(df_nadhira)
3 head(df_nadhira)
4 model <- aov(korosi.y~metode.x., data=df_nadhira)
5 summary(model)
```

The bottom pane is the Console, showing the output of the executed script:

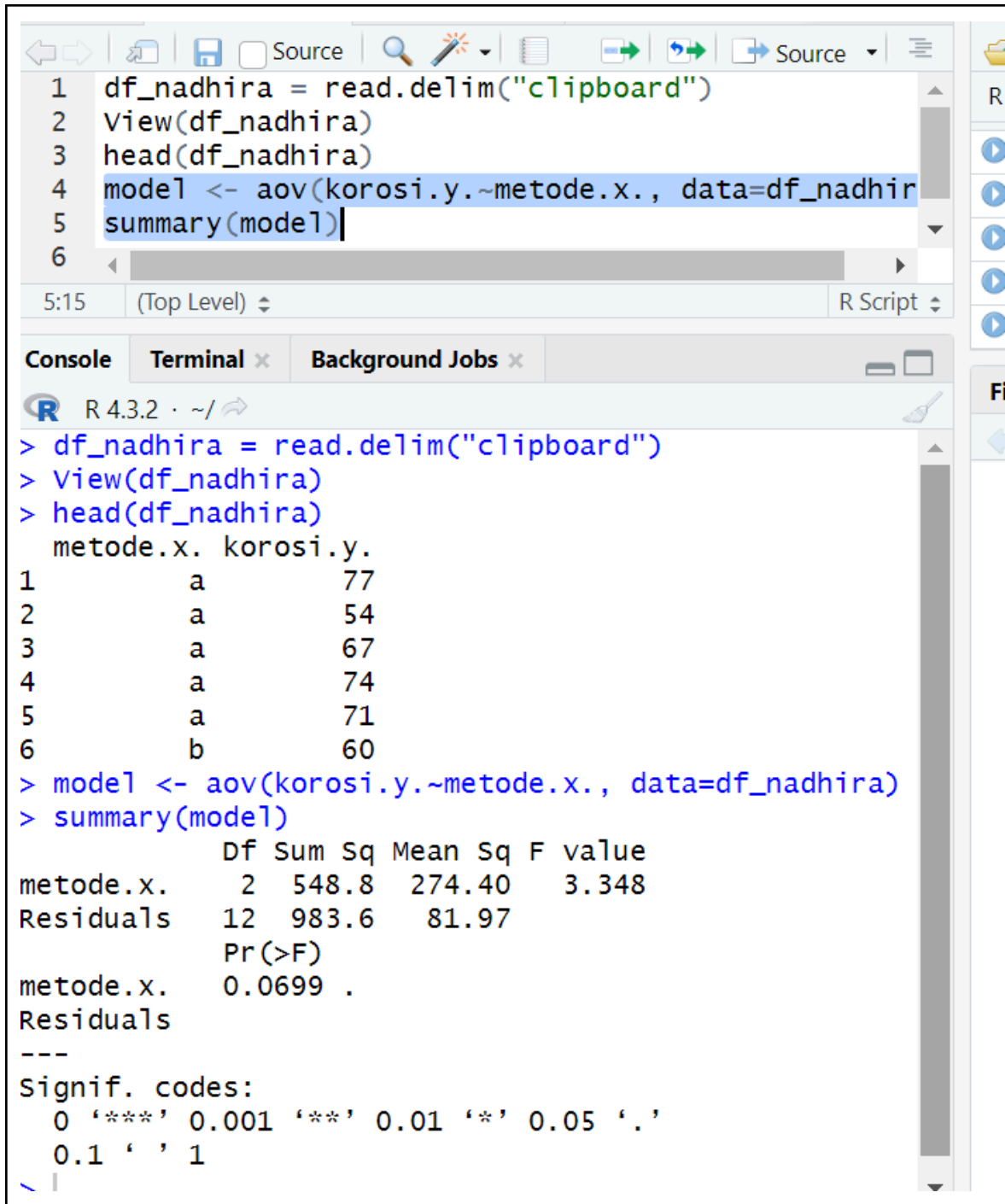
```
> df_nadhira = read.delim("clipboard")
> View(df_nadhira)
> head(df_nadhira)
  metode.x. korosi.y.
1         a         77
2         a         54
3         a         67
4         a         74
5         a         71
6         b         60
```

### Analisis Anova

Analisis keseluruhan terhadap ada atau tidaknya perbedaan pada metode

### Script R:

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



The screenshot shows the R Studio environment. The top pane displays a script with the following code:

```

1 df_nadhira = read.delim("clipboard")
2 view(df_nadhira)
3 head(df_nadhira)
4 model <- aov(korosi.y.~metode.x., data=df_nadhira)
5 summary(model)
6

```

The bottom pane shows the console output, which includes the execution of the script and the results of the ANOVA test:

```

> df_nadhira = read.delim("clipboard")
> View(df_nadhira)
> head(df_nadhira)
  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_nadhira)
> summary(model)

```

	Df	Sum Sq	Mean Sq	F value
metode.x.	2	548.8	274.40	3.348
Residuals	12	983.6	81.97	

```

      Pr(>F)
metode.x. 0.0699 .
Residuals
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
  0.1 ' ' 1

```

**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
```

The screenshot shows an R Studio interface with several open files: 'adhira.R\*', 'Untitled2\*', 'df\_nadhira', and 'uas probstat\_contoh.R'. The script editor displays the following code:

```
4 model <- aov(korosi.y ~ metode.x., data=df_nadhira)
5 summary(model)
6 tukey.test <- TukeyHSD(model)
7 tukey.test
8
9
```

The console output shows the results of the ANOVA and Tukey HSD test:

```
R 4.3.2 ~ /
metode.x.      2    348.8    274.40    3.348
Residuals    12    983.6    81.97
               Pr(>F)
metode.x.    0.0699 .
Residuals
---
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

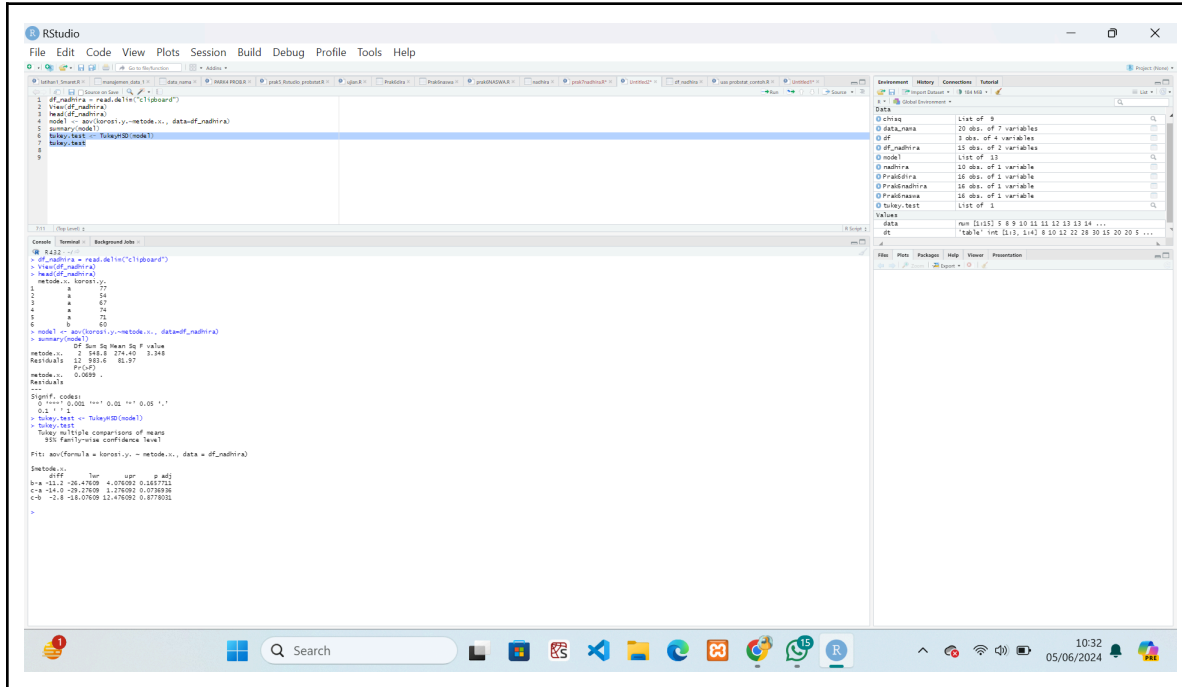
Fit: aov(formula = korosi.y. ~ metode.x., data = df_nadhira)

$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)**

variabel metode.x memiliki nilai p sebesar 0.0699, yang berarti pengaruh metode terhadap variabel korosi.y tidak signifikan pada level 5% ( $p > 0.05$ ). Ada perbedaan yang signifikan pada level 95% di antara beberapa metode, yang ditunjukkan oleh interval kepercayaan yang tidak melintasi nol. meskipun ANOVA tidak menemukan perbedaan signifikan secara keseluruhan, uji Tukey HSD menunjukkan adanya perbedaan tertentu di antara metode yang digunakan

## Screenshoot full screen



## 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:



```

R 4.3.2 ~ /
> df_nadhira = read.delim("clipboard")
> View(df_nadhira)
> head(df_nadhira)
  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_nadhira)
> summary(model)
          Df Sum Sq Mean Sq F value
metode.x.  2  548.8   274.40    3.348
Residuals 12  983.6     81.97
          Pr(>F)
metode.x.  0.0699 .
Residuals
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
  0.1 ' ' 1

```

### Interpretasi (minimal 4 baris)

Nilai F yang diperoleh adalah 3.347702318, menunjukkan adanya variabilitas antara metode yang diuji.

Nilai mean (rata-rata) waktu produksi untuk metode A, B, dan C berturut-turut adalah 68.6, 57.4, dan 54.6.

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_nadhira = read.delim("clipboard")
View(df_nadhira)
head(df_nadhira)

```

#### 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_nadhira = read.delim("clipboard")
> View(df_nadhira)
> head(df_nadhira)
  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_nadhira)
summary(model)

```

#### Output:

```

b         ctrl      5.17
> model <- aov(korosi.y~metode.x., data=df_nadhira)
> 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
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

#### Interpretasi (minimal 4 baris)

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. dengan nilai  $p = 0.0159$  ( $p < 0.05$ ). Hasil ini menunjukkan bahwa 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

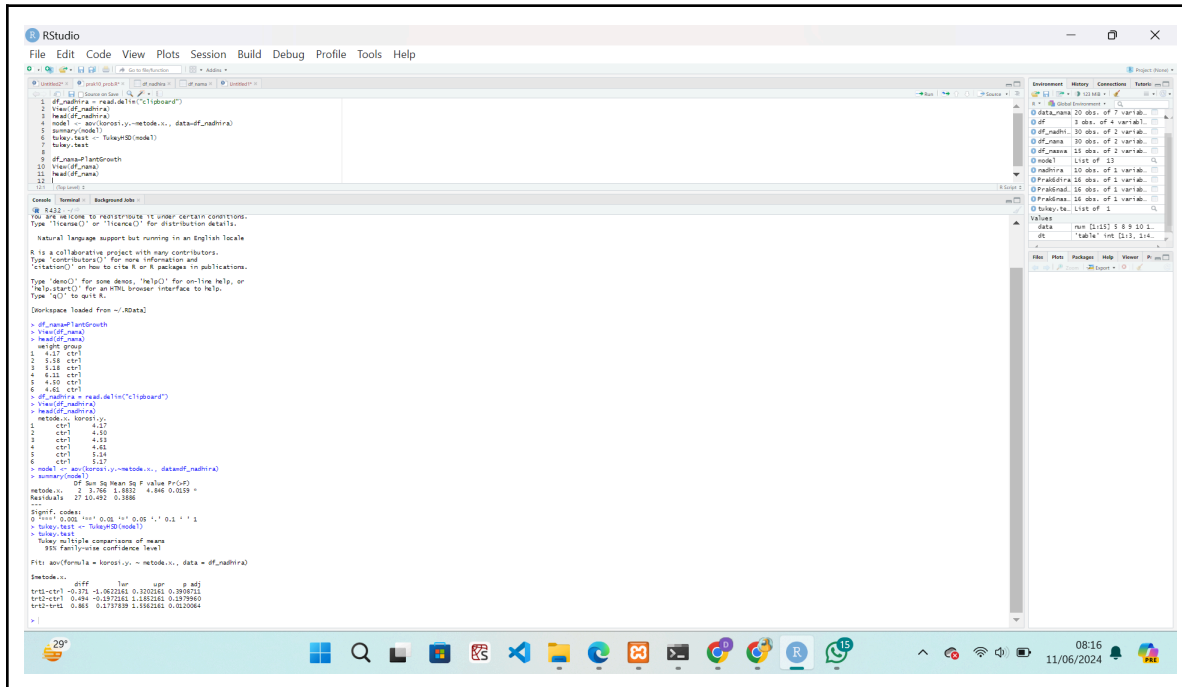
Fit: aov(formula = korosi.y. ~ metode.x., data = df_nadhira)

$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)

Tidak ada perbedaan signifikan antara metode trt1 dan ctrl1 ( $p\text{-adj} = 0.3908711$ ) serta antara trt2 dan ctrl ( $p\text{-adj} = 0.1979960$ ). Perbedaan signifikan ditemukan antara trt2 dan trt1 dengan  $p\text{-adj} = 0.0120064$ , menunjukkan trt2 berbeda secara signifikan dari trt1. Selang kepercayaan untuk perbedaan antara trt2 dan trt1 tidak mencakup nol (0.1737839 hingga 1.5562161), mendukung hasil signifikan. Metode 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)**

hasil uji ANOVA yang dilakukan pada data waktu produksi dari tiga mesin, didapatkan nilai F sebesar 4.846 dan nilai p sebesar 0.0159. Dengan nilai p kurang dari 0.05, kita menolak hipotesis nol yang menyatakan bahwa rata-rata waktu produksi ketiga mesin adalah sama. Artinya, 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)	korosil(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	0.388595259								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	5.50
28													trt2	5.54
29													trt2	5.80
30													trt2	6.15
31													trt2	6.31

**Interpretasi (minimal 4 baris)**

diperoleh nilai F sebesar 4.846 yang menunjukkan adanya variasi signifikan antar kelompok. Nilai rata-rata waktu produksi untuk mesin ctrl, trt1, dan trt2 masing-masing adalah 5.032, 4.661, dan 5.526. Dengan nilai p kurang dari 0.05, 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/NadhiraAninditaRalena/PROBSTAT10.git>

## KESIMPULAN

hasil uji ANOVA yang dilakukan pada data waktu produksi dari tiga mesin (ctrl, trt1, trt2), didapatkan nilai F sebesar 4.846 dan nilai p sebesar 0.0159. Dengan nilai p kurang dari 0.05, kita menolak hipotesis nol yang menyatakan bahwa rata-rata waktu produksi ketiga mesin adalah sama. Artinya, terdapat perbedaan signifikan pada rata-rata waktu produksi antara ketiga mesin tersebut.

## FORM UMPAN BALIK

Elemen Kompetensi	Tingkat Kesulitan	Tingkat Ketertarikan	Waktu Penyelesaian (menit)
Memahami analisis ragam.	biasa	sangat tertarik	30 menit

### 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