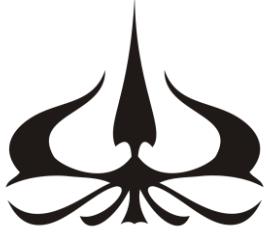


<p>Nama: Mohammad Afif R. Lingkeh</p> <p>NIM: 065002400002</p>	 <p>UNIVERSITAS TRISAKTI PRAKTIKUM STATISTIKA</p>	<p>MODUL 10 STATISTIKA</p> <p>Nama Dosen: Drs. Joko Riyono, M.Si.</p>
<p>Hari/Tanggal: Jumat, 13 Juni 2025</p>		<p>Nama Aslab:</p> <ol style="list-style-type: none">1. Tarum Widyasti P (064002200027)2. 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
1	Metode	Nilai
2	Metode A	77
3	Metode A	54
4	Metode A	67
5	Metode A	74
6	Metode A	71
7	Metode B	60
8	Metode B	41
9	Metode B	59
10	Metode B	65
11	Metode B	62
12	Metode C	49
13	Metode C	52
14	Metode C	69
15	Metode C	47
16	Metode C	56
17		

Script R:

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

```
> df_afif = read.delim("clipboard")
> View(df_afif)
> head(df_afif)
  Metode Nilai
1 Metode A    77
2 Metode A    54
3 Metode A    67
4 Metode A    74
5 Metode A    71
6 Metode 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(Nilai ~ Metode, data = df_afif)
> summary(model)
  Df Sum Sq Mean Sq F value Pr(>F)
Metode        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
>
```

Interpretasi

Karena p-value = 0.0699 > 0.05, maka:

Tidak ada perbedaan yang signifikan secara statistik antara rata-rata nilai pada ketiga metode pada tingkat signifikansi 5%.

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 = Nilai ~ Metode, data = df_afif)

$Metode
    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:

Dari hasil perbandingan antar metode, belum ditemukan perbedaan yang benar-benar signifikan. Perbedaan nilai rata-rata antara Metode C dan A hampir signifikan, tapi masih sedikit di atas ambang batas ($p = 0.0737$). Rentang kepercayaan untuk semua pasangan metode masih mencakup nol, artinya perbedaan yang ada bisa jadi hanya karena kebetulan. Secara keseluruhan, belum cukup bukti kuat untuk mengatakan salah satu metode jelas lebih unggul dari yang lain.

Screnshoot full screen

Mohammad Afif R. Lingkeh – 065002400002

The screenshot shows the RStudio interface with the following R code in the console:

```
> df_Afif = read.delim("clipboard")
> View(df_Afif)
> head(df_Afif)
  Metode Nilai
1   a    77
2   a    54
3   a    67
4   a    74
5   a    71
6   b    60
> model <- aov(Metode~Nilai,data=df_Afif)
>Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :
NA/NAN/Inf in 'y'
In addition: warning message:
In storage.mode(v) <- "double" : NAs introduced by coercion
> model <- aov(Nilai ~ Metode, data = data)
>Error in as.data.frame.default(data, optional = TRUE) :
cannot coerce class "function" to a data.frame
> model <- aov(Nilai ~ Metode, data = data)
>Error in as.data.frame.default(data, optional = TRUE) :
cannot coerce class "function" to a data.frame
> model <- aov(Nilai ~ Metode, data = df_Afif)
> summary(model)
  Df Sum Sq Mean Sq F value Pr(>F)
Metode   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 = Nilai ~ Metode, data = df_Afif)

$Metode
  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
```

The environment pane shows various objects defined in the workspace, including 'Afif' (10 obs. of 1 variable), 'chisq' (List of 9), 'con' (Formal class MySQLConnection), 'data_nama' (24 obs. of 6 variables), 'df' (3 obs. of 3 variables), 'df_afif' (15 obs. of 2 variables), 'eksp_pra&Afif' (16 obs. of 1 variable), 'f1_hw' (List of 10), 'fcast_dc' (List of 10), 'fit_dc' (List of 8), 'fit_hw' (List of 9), 'hargaber' (36 obs. of 16 variables), 'model' (List of 13), 'model_reg' (List of 12), 'Pra&Afif' (16 obs. of 1 variable), 'Prakonana' (16 obs. of 1 variable), 'rice_dec' (List of 6), 'ricesupply' (83 obs. of 13 variables), and 'tukey.test' (List of 1).

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_Afif = [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_Arif = [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}")

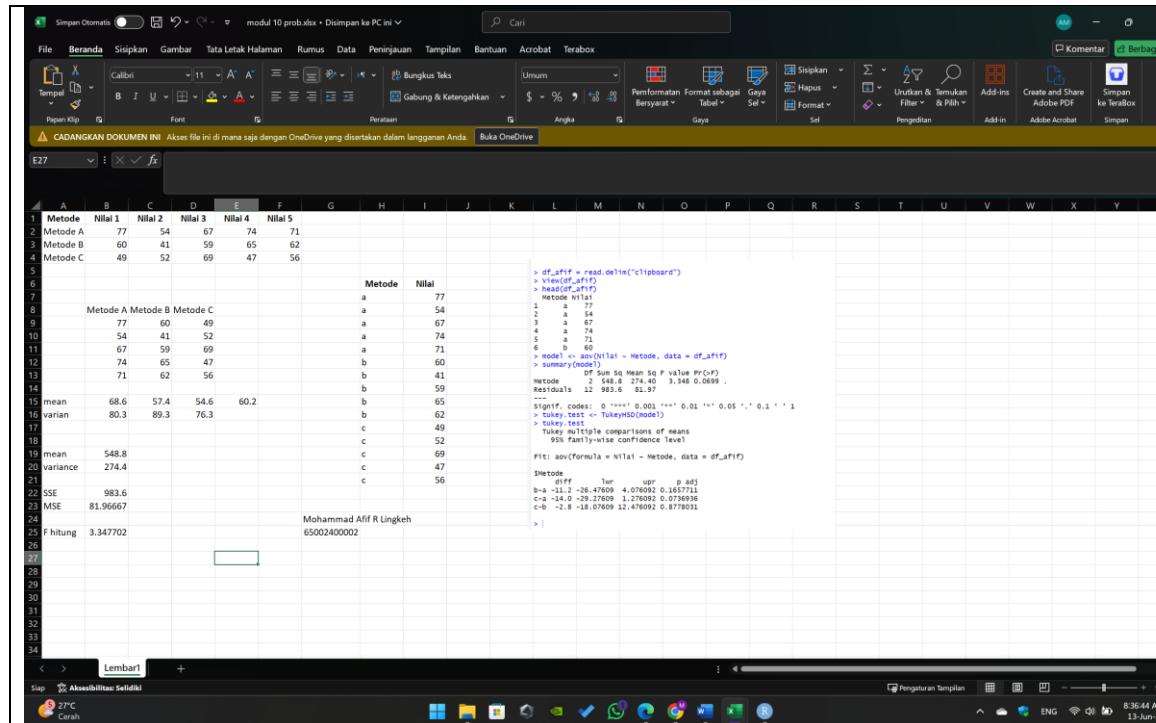
→ Nilai F: 3.347702318015454
Nilai p: 0.06993237223084403
```

Interpretasi (minimal 4 baris)

Hasil uji ANOVA menunjukkan nilai $F = 3.348$ dan $p\text{-value} = 0.0699$. Karena $p\text{-value}$ lebih besar dari 0.05, maka tidak ada perbedaan yang signifikan antar metode.

Namun, nilai p yang mendekati 0.05 memberi indikasi bahwa mungkin ada perbedaan yang perlu diselidiki lebih lanjut. Secara statistik, belum cukup bukti untuk menyimpulkan bahwa ketiga metode menghasilkan nilai yang berbeda.

Excel:



Interpretasi (minimal 4 baris)

Hasil uji ANOVA menunjukkan nilai F sebesar 3.348 dengan p-value 0.0699. Karena p-value lebih besar dari 0.05, maka secara statistik tidak ditemukan perbedaan yang signifikan antara rata-rata nilai dari ketiga metode. Namun, karena p-value mendekati 0.05, hal ini memberi indikasi bahwa mungkin terdapat perbedaan yang patut diselidiki lebih lanjut. Uji lanjutan dengan Tukey HSD mengonfirmasi bahwa tidak ada pasangan metode yang berbeda secara signifikan, meskipun perbandingan antara Metode C dan A menunjukkan p-value 0.0737, yang hampir signifikan. Secara keseluruhan, belum ada cukup bukti kuat untuk menyimpulkan bahwa salah satu metode secara jelas lebih unggul dari yang lain.

ELEMEN KOMPETENSI II

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

Script R:

```

> df_tarum=PlantGrowth
> View(df_tarum)
> head(df_tarum)

```

```
Loading required package: RMySQL
Loading required package: DBI
> df_afif=PlantGrowth
> View(df_afif)
> head(df_afif)
  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
> |
```

Analisis Anova

Script R:

```
> model <- aov(weight~group, data=df_tarum)
> summary(model)
```

```
> df_afif=PlantGrowth
> View(df_afif)
> head(df_afif)
  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
> model <- aov(weight~group, data=df_afif)
> summary(model)
      Df Sum Sq Mean Sq F value Pr(>F)
group       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)

Berdasarkan hasil uji ANOVA satu arah yang dilakukan, diperoleh nilai p sebesar **0,0159** dengan nilai **F sebesar 4,846**. Karena nilai p lebih kecil dari

tingkat signifikansi 0,05, maka dapat disimpulkan **Tolak H_0** : Terdapat perbedaan yang signifikan rata-rata pertumbuhan antar kelompok tanaman.

Analisis Tukey test

Source code:

```
> 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 = weight ~ group, data = df_afif)

$group
      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)

Hasil uji menunjukkan bahwa perbedaan rata-rata antara kelompok trt2 dan trt1 signifikan secara statistik dengan nilai p sebesar 0,0120, karena lebih kecil dari 0,05. Ini berarti **terdapat perbedaan nyata antara kelompok trt2 dan trt1**. Sementara itu, perbedaan antara trt1 dan ctrl ($p = 0,3909$) serta antara trt2 dan ctrl ($p = 0,1980$) tidak signifikan karena nilai p-nya lebih besar dari 0,05. Dengan demikian, dapat disimpulkan bahwa kelompok trt2 memiliki rata-rata yang secara signifikan berbeda dari trt1, namun tidak berbeda secara signifikan dengan kelompok ctrl.

Screnshoot full screen

Mohammad Affif R. Lingkeh – 065002400002

The screenshot shows the RStudio interface. In the top-left pane, there is an R script named 'Untitled.R' containing the following code:

```
1 df_affif<-plantGrowth
2 View(df_affif)
3 head(df_affif)
4 model <- aov(weight~group, data=df_affif)
5 summary(model)
6 tukey.test <- TukeyHSD(model)
7 tukey.test
8
```

In the bottom-left pane, the R console output is displayed:

```
R 4.4.3 --> 
> df_affif<-plantGrowth
> View(df_affif)
> head(df_affif)
  weight group
1 4.17  ctrl1
2 5.58  ctrl1
3 5.18  ctrl1
4 6.11  ctrl1
5 4.50  ctrl1
6 4.61  ctrl1
> model <- aov(weight~group, data=df_affif)
> summary(model)
   Df Sum Sq Mean Sq F value Pr(>F)
group  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
> tukey.test <- TukeyHSD(model)
> tukey.test
  Tukey multiple comparisons of means
  95% family-wise confidence level

Fit: aov(formula = weight ~ group, data = df_affif)

$group
    diff      lwr      upr p adj
trt1-ctrl -0.371 -1.0622161 0.3202161 0.3989711
trt2-ctrl  0.494 -0.1972161 1.1852161 0.1779860
trt1-ctrl  0.865  0.1737939 1.5501162 0.0120064
```

The top-right pane shows the environment browser with various objects listed. The bottom-right pane shows the system tray with icons for battery, signal, and date/time.

Python:

Source code:

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

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

# Gabungkan data ke dalam satu array
data_tarum = [plant_A, plant_B, plant_C]

# Hitung ANOVA menggunakan scipy.stats
f_statistic, p_value = stats.f_oneway(plant_A, plant_B, plant_C)

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



Output:

```
3d    import numpy as np
      import scipy.stats as stats

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

      # Gabungkan data ke dalam satu array
      data_tarum = [plant_A, plant_B, plant_C]

      # Hitung ANOVA menggunakan scipy.stats
      f_statistic, p_value = stats.f_oneway(plant_A, plant_B, plant_C)

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

→ Nilai F: 4.846087862380136
Nilai p: 0.0159099583256229
```

Interpretasi (minimal 4 baris)

Berdasarkan hasil uji ANOVA satu arah diperoleh nilai F sebesar **4.846** dan nilai p sebesar **0.0159**. Karena nilai p lebih kecil dari taraf signifikansi 0.05, maka dapat disimpulkan terima H1 bahwa terdapat **perbedaan rata-rata waktu pertumbuhan yang signifikan secara statistik** di antara ketiga tanaman (plant A, B, dan C).

Excel:

Mohammad Afif R. Lingkeh – 065002400002

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Metode	Nilai 1	Nilai 2	Nilai 3	Nilai 4	Nilai 5	Nilai 6	Nilai 7	Nilai 8	Nilai 9	Nilai 10				
2	ctrl	4.17	5.58	5.18	6.11	4.5	4.61	5.17	4.53	5.33	5.14	ctrl	4.17		
3	trt1	4.81	4.17	4.41	3.59	5.87	3.83	6.03	4.89	4.32	4.69	ctrl	5.58		
4	trt2	6.31	5.12	5.54	5.5	5.37	5.29	4.92	6.15	5.8	5.26	ctrl	5.18		
5												ctrl	6.11		
6												ctrl	4.5		
7		ctrl	trt1	trt2								ctrl	4.61		
8		4.17	4.81	6.31								ctrl	5.17		
9		5.58	4.17	5.12								ctrl	4.53		
10		5.18	4.41	5.54								ctrl	5.33		
11		6.11	3.59	5.5								ctrl	5.14		
12		4.5	5.87	5.37								trt1	4.81		
13		4.61	3.83	5.29								trt1	4.17		
14		5.17	6.03	4.92								trt1	4.41		
15		4.53	4.89	6.15								trt1	3.59		
16		5.33	4.32	5.8								trt1	5.87		
17		5.14	4.69	5.26								trt1	3.83		
18												trt1	6.03		
19	mean	5.032	4.661	5.526	5.073							trt1	4.89		
20	variance	0.339996	0.629921	0.195871								trt1	4.32		
21												trt1	4.69		
22	SSTR	3.76634										trt2	6.31		
23	MSTR	1.88317										trt2	5.12		
24												trt2	5.54		
25	SSE	10.49209	N-1 = 10-1 = 9									trt2	5.5		
26	MSE	0.388596	df= 30 - 3 = 27									trt2	5.37		
27												trt2	5.29		
28	F	4.846088										trt2	4.92		
29												trt2	6.15		
30												trt2	5.8		
31												trt2	5.26		
32															
33															

Interpretasi (minimal 4 baris)

Data yang digunakan adalah hasil pengamatan waktu produksi dari tiga kelompok mesin, yaitu ctrl, trt1, dan trt2, masing-masing terdiri dari 10 data. Rata-rata waktu produksi tertinggi terdapat pada kelompok trt2 (5.526), sedangkan yang terendah pada trt1 (4.661). Berdasarkan perhitungan manual di Excel, nilai SSTR (Sum of Squares for Treatment) sebesar 3.76634 dan SSE (Sum of Squares for Error) sebesar 10,49209, yang kemudian menghasilkan nilai F hitung = 4.846088.

Karena nilai p lebih kecil dari 0.05, maka keputusan yang diambil adalah **menolak hipotesis nol (H_0)** dan menyimpulkan bahwa terdapat **perbedaan rata-rata yang signifikan antar kelompok mesin**. Dengan kata lain, setidaknya satu mesin memiliki performa yang berbeda secara nyata dari yang lainnya.

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. (✓)

KESIMPULAN

Buatlah kesimpulan minimal 3 baris.

Dengan mengerjakan praktikum ini saya dapat memahami uji ANOVA digunakan untuk membandingkan rata-rata dari dua atau lebih kelompok sekaligus, dan ketika hasilnya signifikan, dibutuhkan uji lanjutan seperti **Tukey HSD** untuk mengetahui secara spesifik kelompok mana yang berbeda. Selain itu, praktikum ini juga melatih kemampuan dalam menghitung manual nilai-nilai statistik seperti **SSTR**, **SSE**, **MSTR**, dan **MSE**, serta membandingkan hasilnya dengan output dari perangkat lunak seperti Excel, Python, dan RStudio.

FORM UMPAN BALIK

Elemen Kompetensi	Tingkat Kesulitan	Tingkat Ketertarikan	Waktu Penyelesaian (menit)
Memahami analisis ragam.	Sangat mudah	Tertarik	50 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

Mohammad Afif R. Lingkeh – 065002400002

3: Tertarik

4: Sangat Tertarik