


Nama: Naswa Aulia NIM: 065002300036	 UNIVERSITAS TRISAKTI PRAKTIKUM STATISTIKA	MODUL 10 STATISTIKA Nama Dosen: Dedy Sugiarto Nama Aslab: 1. Tarum Widyasti P (064002200027) 2. Kharisma Maulida S (064002200024)
Hari/Tanggal: Rabu, 5 Juni 2024		

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

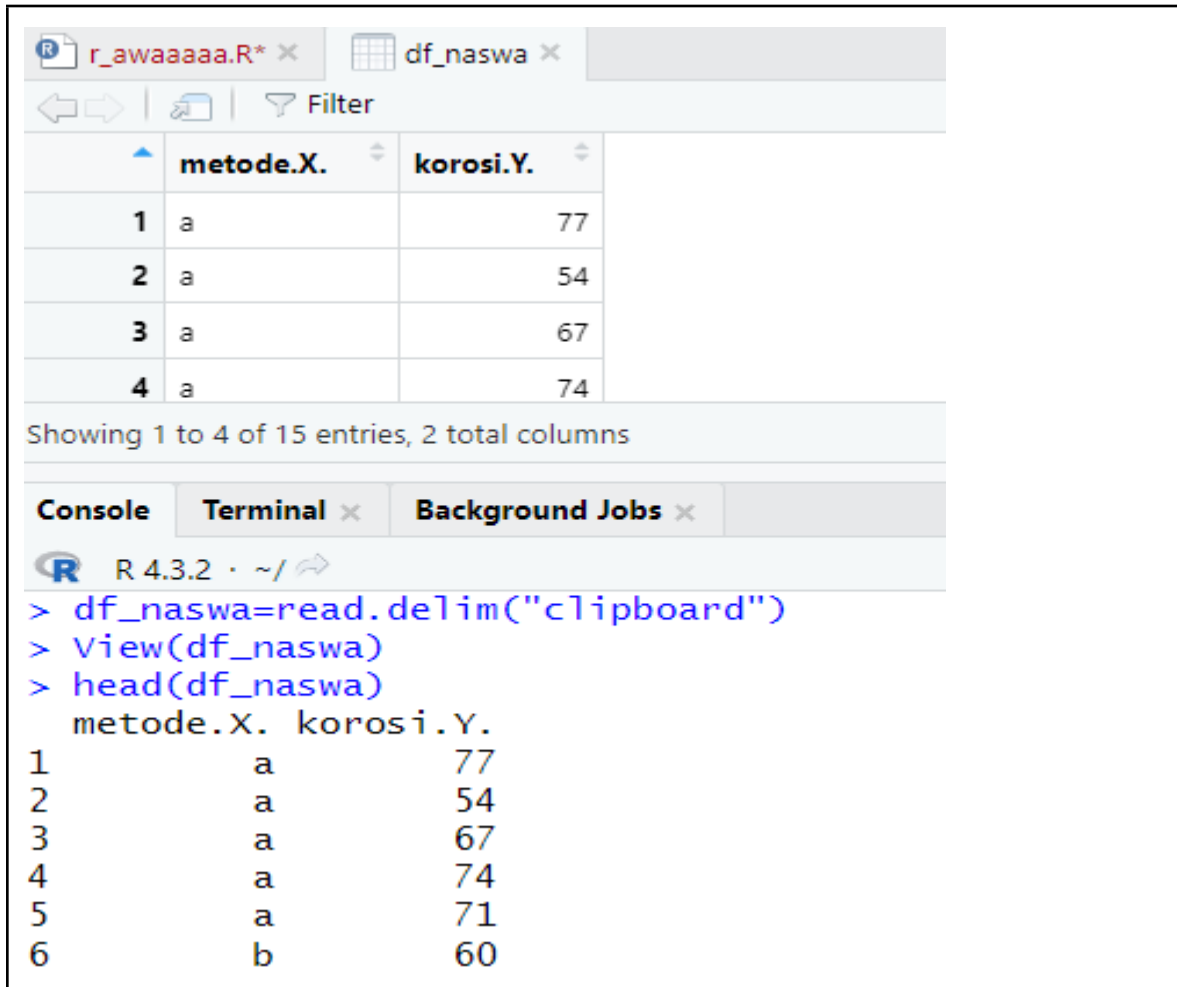
Dengan menggunakan alpha 0.05, ujilah bahwa ketiga metode tersebut mempunyai pengaruh yang sama terhadap pencegahan korosi :

Ubah data menjadi data bertumpuk pada excel

H	I
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)
```



The screenshot shows the RStudio interface. The top pane displays a data frame named `df_naswa` with two columns: `metode.X.` and `korosi.Y.`. The data is as follows:

	metode.X.	korosi.Y.
1	a	77
2	a	54
3	a	67
4	a	74

Below the data frame, it says "Showing 1 to 4 of 15 entries, 2 total columns". The bottom pane shows the R console with the following commands and output:

```
> df_naswa=read.delim("clipboard")
> View(df_naswa)
> head(df_naswa)
  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_naswa)
> 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
'***' indicates significance level < 0.001
```

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_naswa)

$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

Screenshot full screen

The screenshot shows the RStudio interface with a script editor and a console. The script editor contains the following R code:

```

6 df_naswa=read.delim("clipboard")
7 View(df_naswa)
8 head(df_naswa)
9 model <- aov(korosi.Y.~metode.X., data=df_naswa)
10 summary(model)
11 tukey.test <- TukeyHSD(model)
12 tukey.test
13

```

The console shows the output of the R code:

```

> df_naswa=read.delim("clipboard")
> View(df_naswa)
> head(df_naswa)
  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_naswa)
> 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_naswa)

$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 yang diperoleh adalah 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:

[illegible]

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_naswal)
head(df_naswa)
```

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_naswa=read.delim("clipboard")
> View(df_naswa)
> head(df_naswa)
  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_naswa)
summary(model)
```

Output:

```
> model <- aov(korosi.Y.~metode.X., data=df_naswa)
> 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
> TukeyHSD(model)
```

Interpretasi (minimal 4 baris)

ilai $p = 0.0159$ ($p < 0.05$). 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_naswa)

$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


```

30 df_naswa<-read.delim("clipboard")
31 View(df_naswa)
32 head(df_naswa)
33 model<- aov(koros1.Y.~metode.X., data=df_naswa)
34 summary(model)
35 tukey.test <- TukeyHSD(model)
36 tukey.test
37

```

Console

```

> df_naswa<-read.delim("clipboard")
> View(df_naswa)
> head(df_naswa)
  metode.X. koros1.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
> model<- aov(koros1.Y.~metode.X., data=df_naswa)
> 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
> tukey.test <- TukeyHSD(model)
> tukey.test
  Tukey multiple comparisons of means
  95% family-wise confidence level

Fit: aov(formula = koros1.Y. ~ metode.X., data = df_naswa)

$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

```

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:

[illegible]

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/NaswaAulia03/prak10probbat.git>

KESIMPULAN

(ANOVA) untuk menguji perbedaan rata-rata antar lebih dari 2 grup sampel. Jenis antara lain ANOVA satu arah (one-way ANOVA), ANOVA dua arah (two-way ANOVA), ANOVA tiga arah (three-way ANOVA). 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. 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.	sulit	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