Nama:
Dewanto Maulana
Sukarno Putra
NIM:
065002300002

Hari/Tanggal: Rabu, 06 Juni 2024



PRAKTIKUM STATISTIKA

MODUL 10 STATISTIKA

Nama Dosen: Dedy Sugiarto

Nama Aslab:

- Tarum Widyasti P (064002200027)
- 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 vairabel 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

	5
Α	В
metode (x)	korosi (y)
a	77
a	54
а	67
a	74
a	71
b	60
b	41
b	59
b	65
b	62
С	49
С	52
С	69
С	47
С	56

Script R:

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

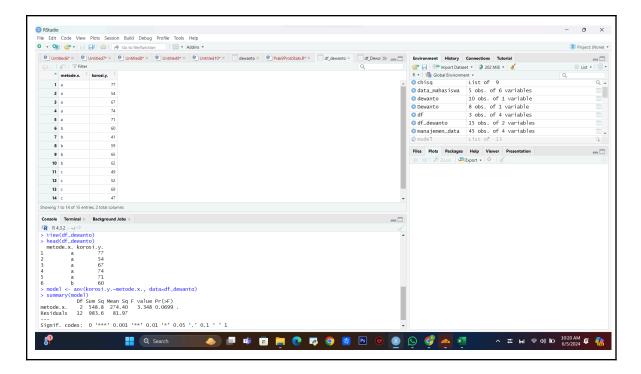
```
> df_dewanto = read.delim("clipboard")
> View(df_dewanto)
> head(df_dewanto)
  metode.x. korosi.y.
1
                     77
           a
2
                     54
           a
3
                     67
           a
4
                     74
           a
5
                     71
           a
6
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



Interpretasi

F Value: Nilai F adalah 3.348. Ini menunjukkan rasio antara variabilitas yang dijelaskan oleh model terhadap variabilitas yang tidak dijelaskan oleh model. Pr(>F): Nilai p (p-value) adalah 0.0699. Ini adalah probabilitas mendapatkan nilai

F setidaknya sebesar itu, dengan asumsi bahwa hipotesis nol adalah benar. Hipotesis nol dalam konteks ini adalah bahwa tidak ada perbedaan rata-rata korosi antara metode yang digunakan.

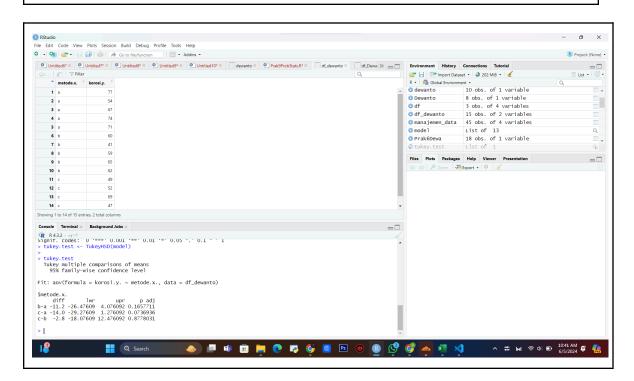
Secara keseluruhan, sementara tidak ada bukti yang cukup kuat untuk menyimpulkan perbedaan signifikan antara metode berdasarkan dataset yang ada, hasilnya menunjukkan kecenderungan yang bisa relevan dengan penelitian lebih lanjut.

Analisis Tukey test

Analisis ada atau tidaknya perbedaan antara 2 metode

Script R:

tukey.test <- TukeyHSD(model)
tukey.test</pre>



Interpretasi

- Perbandingan b-a:

Difference (Diff): Metode b memiliki rata-rata nilai korosi 11.2 unit lebih rendah dibandingkan dengan metode a.

Confidence Interval: Interval kepercayaan 95% untuk perbedaan rata-rata adalah dari -26.47609 sampai 4.076092.

P-value: Nilai p yang disesuaikan adalah 0.1657711, yang berarti perbedaan ini

tidak signifikan pada tingkat signifikansi 0.05. Tidak ada bukti yang cukup untuk menyatakan bahwa metode b dan a memiliki perbedaan rata-rata yang signifikan.

- Perbandingan c-a:

Difference (Diff): Metode c memiliki rata-rata nilai korosi 14.0 unit lebih rendah dibandingkan dengan metode a.

Confidence Interval: Interval kepercayaan 95% untuk perbedaan rata-rata adalah dari -29.27609 sampai 1.276092.

P-value: Nilai p yang disesuaikan adalah 0.0736936. Ini mendekati signifikan tetapi tidak cukup untuk melebihi ambang batas 0.05. Dengan nilai p yang lebih besar dari 0.05, kita tidak dapat menyatakan perbedaan ini signifikan, meskipun perbedaannya cukup besar dan mendekati signifikansi.

- Perbandingan c-b:

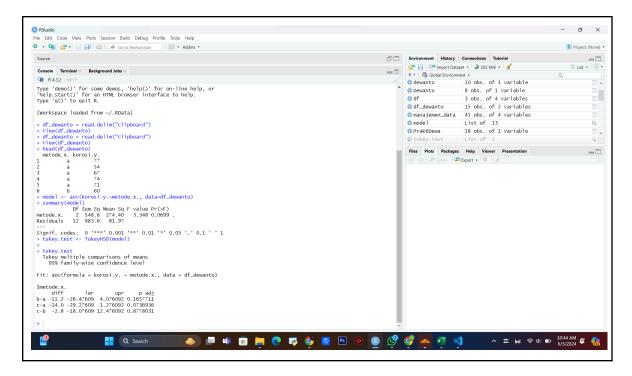
Difference (Diff): Metode c memiliki rata-rata nilai korosi 2.8 unit lebih rendah dibandingkan dengan metode b.

Confidence Interval: Interval kepercayaan 95% untuk perbedaan rata-rata adalah dari -18.07609 sampai 12.476092.

P-value: Nilai p yang disesuaikan adalah 0.8778031, yang jauh lebih besar dari 0.05. Tidak ada bukti signifikan bahwa metode c dan b memiliki perbedaan rata-rata yang berarti.

Secara keseluruhan, meskipun ada indikasi perbedaan rata-rata antara metode, hasil ini tidak cukup kuat untuk menyimpulkan perbedaan yang signifikan tanpa penelitian lebih lanjut.

Screnshoot 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:

```
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([69, 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^Nilal F: {f_statistic}")
print(f^Nilal p: {p_value}")

Nilai F: 3.347702318015454
Nilai p: 0.06993237223084404
```

Penjelasan Nilai F dan Nilai p:

Nilai F:

Nilai F adalah statistik yang digunakan untuk menentukan apakah ada perbedaan signifikan antara rata-rata kelompok.

Dalam konteks ini, nilai F sebesar 3.347702318015454 menunjukkan seberapa besar variabilitas antar kelompok dibandingkan dengan variabilitas dalam kelompok.

Nilai p:

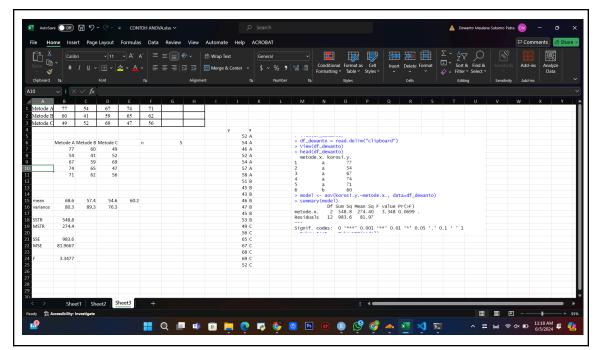
- Nilai p adalah probabilitas mendapatkan nilai F setidaknya sebesar itu jika hipotesis nol benar (yaitu, tidak ada perbedaan rata-rata antara kelompok).
- Nilai p sebesar 0.06993237223084404 menunjukkan bahwa ada sekitar 6.99% kemungkinan bahwa perbedaan rata-rata yang diamati terjadi secara kebetulan jika tidak ada perbedaan nyata antara metode.

Interpretasi Statistik:

- Hipotesis NoI (H0): Tidak ada perbedaan rata-rata waktu produksi antara metode A, B, dan C.
- Hipotesis Alternatif (H1): Setidaknya ada satu metode yang memiliki rata-rata waktu produksi berbeda dari yang lain.

Dengan nilai p sebesar 0.06993237223084404, yang lebih besar dari tingkat signifikansi umum 0.05, kita tidak dapat menolak hipotesis nol. Ini berarti bahwa berdasarkan data ini, tidak ada bukti yang cukup kuat untuk menyimpulkan bahwa ada perbedaan yang signifikan dalam rata-rata waktu produksi antara metode A, B, dan C pada tingkat signifikansi 0.05.

Excel:



1. Mean dan Variance:

Mean: Rata-rata waktu produksi untuk setiap metode.

Metode A: 68.6 Metode B: 57.4 Metode C: 54.6

Variance: Variansi dari waktu produksi untuk setiap metode.

Metode A: 80.3 Metode B: 89.3 Metode C: 76.3

- 2. SSTR (Sum of Squares for Treatment):
- SSTR: 548.8
- SSTR mengukur variabilitas antara rata-rata kelompok (metode). Ini menunjukkan seberapa jauh rata-rata kelompok berbeda dari rata-rata keseluruhan (grand mean).
- 3. MSTR (Mean Square for Treatment):
- MSTR: 274.4
- MSTR adalah SSTR dibagi dengan derajat kebebasan untuk treatment (k 1), di mana k adalah jumlah kelompok. Dalam hal ini, k = 3, sehingga derajat kebebasan adalah 2.
- 4. SSE (Sum of Squares for Error):
- SSE: 983.6
- SSE mengukur variabilitas dalam kelompok. Ini menunjukkan seberapa

jauh setiap pengamatan dalam kelompok berbeda dari rata-rata kelompoknya.

- 5. MSE (Mean Square for Error):
- MSE: 81.96667
- MSE adalah SSE dibagi dengan derajat kebebasan untuk error (N k), di mana N adalah jumlah total pengamatan dan k adalah jumlah kelompok.
 Dalam hal ini, N = 15 (5 pengamatan per kelompok), sehingga derajat kebebasan adalah 12.
- 6. F-Value:
- F: 3.347702318
- Nilai F adalah rasio MSTR terhadap MSE. Ini mengukur apakah variabilitas antara kelompok secara signifikan lebih besar daripada variabilitas dalam kelompok.

Interpretasi Statistik

- 1. Hipotesis Nol (H0):
- Tidak ada perbedaan rata-rata waktu produksi antara metode A, B, dan C.
- 2. Hipotesis Alternatif (H1):
- Setidaknya ada satu metode yang memiliki rata-rata waktu produksi berbeda dari yang lain.
- 3. Nilai F:
- Nilai F sebesar 3.347702318 menunjukkan rasio antara variabilitas antar kelompok terhadap variabilitas dalam kelompok.
- 4. Nilai p:
- Berdasarkan hasil sebelumnya, nilai p sebesar 0.06993237223084404. Ini lebih besar dari tingkat signifikansi umum 0.05.

Kesimpulan

- Dengan nilai p sebesar 0.06993237223084404, yang lebih besar dari 0.05, kita tidak dapat menolak hipotesis nol. Artinya, tidak ada bukti yang cukup kuat untuk menyatakan bahwa ada perbedaan signifikan dalam rata-rata waktu produksi antara metode A, B, dan C pada tingkat signifikansi 0.05.
- Meskipun ada perbedaan dalam rata-rata waktu produksi antara metode, perbedaan ini tidak cukup signifikan secara statistik untuk membuat kesimpulan definitif berdasarkan data yang ada.

Nilai p yang mendekati 0.05 menunjukkan bahwa ada kemungkinan perbedaan yang memerlukan investigasi lebih lanjut atau pengumpulan data tambahan untuk mendapatkan kesimpulan yang lebih pasti.

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_dewanto=PlantGrowth
> View(df_dewanto)
> head(df_dewanto)
```

Output:

```
> df_dewanto=PlantGrowth
> View(df_dewanto)
> head(df_dewanto)
  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_dewanto)
> summary(model)
```

Output:

Interpretasi

- F value: Nilai F yang tinggi menunjukkan bahwa variasi antar kelompok relatif besar dibandingkan dengan variasi dalam kelompok.
- Pr(>F): Nilai p (0.0159) menunjukkan probabilitas mendapatkan nilai F setinggi ini atau lebih tinggi jika tidak ada perbedaan nyata antara kelompok.

Kesimpulan

- Karena nilai p (0.0159) lebih kecil dari tingkat signifikansi umum (α = 0.05), kita menolak hipotesis nol.
- Ini berarti ada bukti yang cukup untuk mengatakan bahwa setidaknya ada satu perbedaan yang signifikan dalam berat antar kelompok.

Secara keseluruhan, hasil ANOVA menunjukkan bahwa berat rata-rata di setidaknya satu kelompok berbeda secara signifikan dari kelompok lain.

Analisis Tukey test

Source code:

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

Output:

Interpretasi

- 1. rt1 ctrl:
- Selisih rata-rata berat antara kelompok trt1 dan ctrl adalah -0.371.
- Interval kepercayaan 95% untuk perbedaan ini adalah dari -1.0622161 hingga 0.3202161.

- Nilai p yang disesuaikan adalah 0.3908711, yang lebih besar dari 0.05. Jadi, tidak ada perbedaan yang signifikan secara statistik antara kelompok trt1 dan ctrl.

2. trt2 - ctrl:

- Selisih rata-rata berat antara kelompok trt2 dan ctrl adalah 0.494.
- Interval kepercayaan 95% untuk perbedaan ini adalah dari -0.1972161 hingga 1.1852161.
- Nilai p yang disesuaikan adalah 0.1979960, yang lebih besar dari 0.05.
 Jadi, tidak ada perbedaan yang signifikan secara statistik antara kelompok trt2 dan ctrl.

3. trt2 - trt1:

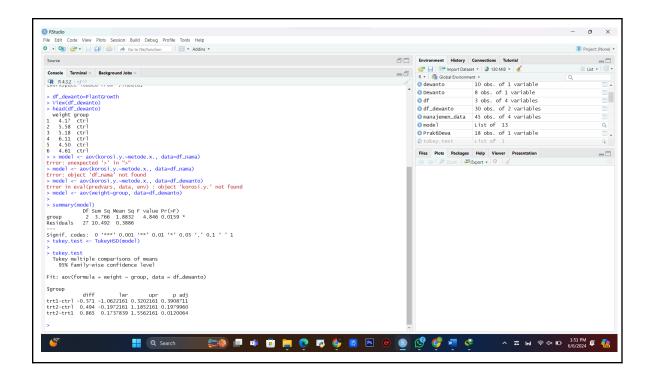
- Selisih rata-rata berat antara kelompok trt2 dan trt1 adalah 0.865.
- Interval kepercayaan 95% untuk perbedaan ini adalah dari 0.1737839 hingga 1.5562161.
- Nilai p yang disesuaikan adalah 0.0120064, yang lebih kecil dari 0.05. Jadi, ada perbedaan yang signifikan secara statistik antara kelompok trt2 dan trt1.

Kesimpulan

- trt2 trt1: Ada perbedaan signifikan dalam berat antara kelompok trt2 dan trt1.
- trt1 ctrl dan trt2 ctrl: Tidak ada perbedaan signifikan dalam berat antara kelompok trt1 dengan ctrl dan trt2 dengan ctrl.

Uji Tukey menunjukkan bahwa perbedaan berat yang signifikan hanya ditemukan antara trt2 dan trt1, sedangkan perbandingan lain tidak menunjukkan perbedaan yang signifikan.

Screnshoot full screen



Python:

Source code:

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

# Data waktu produksi dari ketiga mesin ctrl0 = np.array([4.17, 5.58, 5.18, 6.11, 4.50, 4.61, 5.17, 4.53, 5.33, 5.14]) ctrl1 = np.array([4.81, 4.17, 4.41, 3.59, 5.87, 3.83, 6.03, 4.89, 4.32, 4.69]) ctrl2 = 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_dewanto = [ctrl0, ctrl1, ctrl2]

# Hitung ANOVA menggunakan scipy.stats f_statistic, p_value = stats.f_oneway(ctrl0, ctrl1, ctrl2)

# 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
ctrl0 = np.array([4.17, 5.58, 5.18, 6.11, 4.50, 4.61, 5.17, 4.53, 5.33, 5.14])
ctrl1 = np.array([4.81, 4.17, 4.41, 3.59, 5.87, 3.83, 6.03, 4.89, 4.32, 4.69])
ctrl2 = 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_dewanto = [ctrl0, ctrl1, ctrl2]

# Hitung ANOVA menggunakan scipy.stats
f_statistic, p_value = stats.f_oneway(ctrl0, ctrl1, ctrl2)

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

Nilai F: 4.846087862380136
Nilai p: 0.0159099583256229
```

Hasil ANOVA

- 1. Nilai F (F-statistic): 4.846
- Nilai F mengukur proporsi variasi antara kelompok terhadap variasi dalam kelompok. Nilai F yang lebih besar menunjukkan bahwa variasi antar kelompok lebih besar dibandingkan variasi dalam kelompok.
- 2. Nilai p (p-value): 0.0159
- Nilai p mengukur probabilitas mendapatkan nilai F setinggi ini atau lebih tinggi jika tidak ada perbedaan nyata antara kelompok. Nilai p yang lebih kecil dari 0.05 menunjukkan bahwa hasilnya signifikan secara statistik pada tingkat kepercayaan 95%.

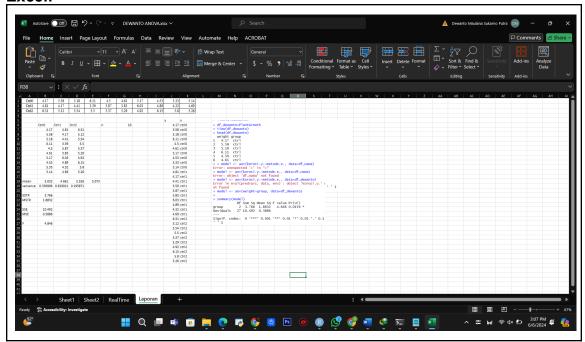
Interpretasi

- Karena nilai p (0.0159) lebih kecil dari tingkat signifikansi umum (α = 0.05), kita menolak hipotesis nol.
- Ini berarti ada bukti yang cukup untuk menyatakan bahwa setidaknya ada satu perbedaan yang signifikan dalam waktu produksi antara ketiga mesin tersebut.

Kesimpulan:

Hasil ANOVA menunjukkan bahwa ada perbedaan yang signifikan dalam waktu produksi antara mesin-mesin (ctrl0, ctrl1, dan ctrl2). Langkah selanjutnya adalah melakukan uji post-hoc (seperti Uji Tukey) untuk menentukan pasangan mana dari kelompok yang berbeda secara signifikan.

Excel:



Interpretasi

- 1. Mean (Rata-rata):
- Kelompok Ctrl0 memiliki rata-rata sebesar 5.032.
- Kelompok Ctrl1 memiliki rata-rata sebesar 4.661.
- Kelompok Ctrl2 memiliki rata-rata sebesar 5.526.
- Rata-rata total *x* adalah 5.073.
- 2. Variance (Varian):
- Varian untuk kelompok Ctrl0 adalah 0.339995556.
- Varian untuk kelompok Ctrl1 adalah 0.629921111.
- Varian untuk kelompok Ctrl2 adalah 0.195871111.
- 3. SST (Total Sum of Squares) atau Jumlah Kuadrat Total:
- Jumlah kuadrat total adalah 3.766.
- 4. MST (Mean Square Total) atau Kuadrat Rata-rata Total:
- Kuadrat rata-rata total adalah 1.8832.
- 5. SSE (Error Sum of Squares) atau Jumlah Kuadrat Kesalahan:
- Jumlah kuadrat kesalahan adalah 10.492.

- 6. MSE (Mean Square Error) atau Kuadrat Rata-rata Kesalahan:
- Kuadrat rata-rata kesalahan adalah 0.3886.
- 7. F-statistic (F-Statistik):
- Nilai F adalah 4.846.

Interpretasi:

- Analisis ini bertujuan untuk mengetahui apakah terdapat perbedaan yang signifikan antara rata-rata kelompok-kelompok tersebut.
- Nilai F-statistik yang signifikan menunjukkan bahwa setidaknya satu dari rata-rata kelompok tersebut berbeda secara signifikan dari yang lain.
- Berdasarkan nilai F-statistik yang cukup besar (4.846) dan signifikan (p
 0.05), kita dapat menyimpulkan bahwa ada perbedaan yang signifikan antara setidaknya satu pasang kelompok.
- Untuk mengetahui pasangan kelompok mana yang berbeda secara signifikan, perlu dilakukan uji perbandingan pasangan (post hoc test) atau kontrast tertentu.

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

Pada praktikum ke 10 ini kita belajar mengenai ANOVA di mana kita menggunakan aplikasi R studi, Bahasa Python, dan Excel.

Pada data set realtime:

- Dengan nilai p sebesar 0.06993237223084404, yang lebih besar dari 0.05, kita tidak dapat menolak hipotesis nol. Artinya, tidak ada bukti yang cukup kuat untuk menyatakan bahwa ada perbedaan signifikan dalam rata-rata waktu produksi antara metode A, B, dan C pada tingkat signifikansi 0.05.
- Meskipun ada perbedaan dalam rata-rata waktu produksi antara metode, perbedaan ini tidak cukup signifikan secara statistik untuk membuat kesimpulan definitif berdasarkan data yang ada.

Pada data set Laporan:

- Analisis ini bertujuan untuk mengetahui apakah terdapat perbedaan yang signifikan antara rata-rata kelompok-kelompok tersebut.
- Nilai F-statistik yang signifikan menunjukkan bahwa setidaknya satu dari rata-rata kelompok tersebut berbeda secara signifikan dari yang lain.
- Berdasarkan nilai F-statistik yang cukup besar (4.846) dan signifikan (p < 0.05), kita dapat menyimpulkan bahwa ada perbedaan yang signifikan antara setidaknya satu pasang kelompok.
- Untuk mengetahui pasangan kelompok mana yang berbeda secara signifikan, perlu dilakukan uji perbandingan pasangan (post hoc test) atau kontrast tertentu.

FORM UMPAN BALIK

Elemen Kompetensi	Tingkat Kesulitan	Tingkat Ketertarikan	Waktu Penyelesaian
			(menit)

Memahami analisis	 	
ragam.		

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