

FINAL EXAM

MATA KULIAH MANAJEMEN DAN ANALISIS DATA DENGAN R



ASTI OKTOVIANTI SUNMAYA ANANDA PUTRI

NPM. 131520220005

PROGRAM STUDI EPIDEMIOLOGI

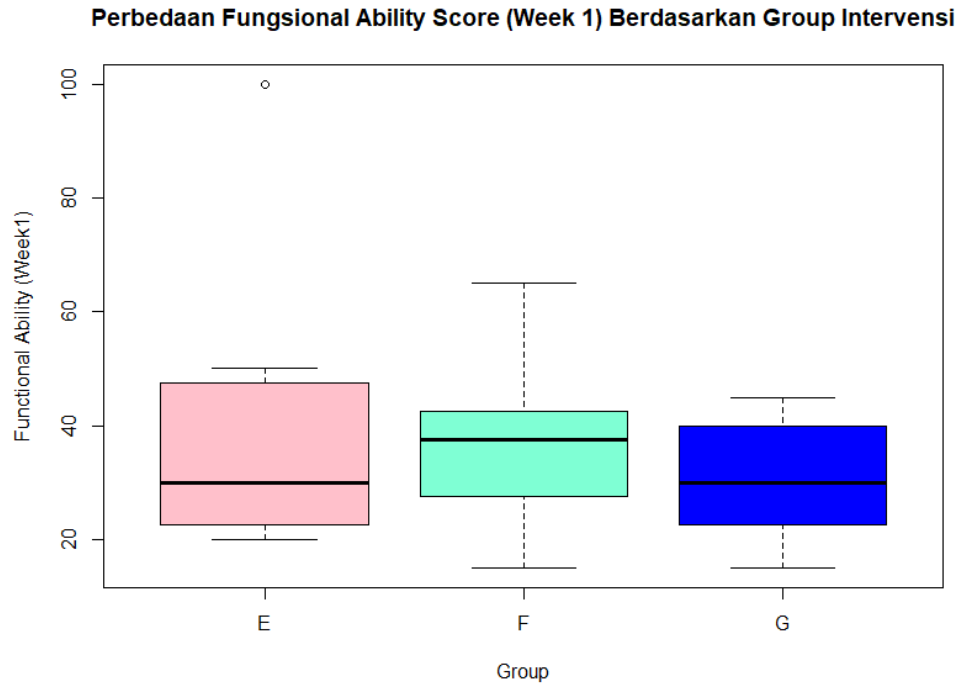
FAKULTAS KEDOKTERAN

UNIVERSITAS PADJADJARAN

2023

1. Menganalisis perbedaan dari Functional Ability pada pekan pertama (Bart1) berdasarkan grup intervensi (Group) dengan visualisasi boxplot.

Membuat grafik boxplot dan menginterpretasi dan menyimpulkan hasilnya.



Interpretasi:

- Median functional ability score week 1 pada group F lebih besar dibandingkan group E maupun group G
- Nilai minimum terendah berada pada group F dan G yaitu 15
- Terdapat 1 outlier pada group E dengan score 100 (Nilai maksimum tertinggi dari seluruh observasi)

2. Mencek normalitas data dari Functional Ability pekan pertama (Bart1) dengan uji statistik yang sesuai dan menginterpretasikannya.

Hasil:

p-value = 0.0007 → <0.05 = Data tidak terdistribusi normal

3. Menghitung perubahan Functional Ability dari pekan pertama (Bart1)sampai pekan terakhir (Bart8) dan membuatnya menjadi variable baru (Bart_diff).

E5	LE6	LE7	LE8	Bal1	Bal2	Bal3	Bal4	Bal5	Bal6	Bal7	Bal8	Bart1	Bart2	Bart3	Bart4	Bart5	Bart6	Bart7	Bart8	Bart_diff
22	23	24	24	8	9	10	10	10	10	10	10	45	45	45	45	80	80	80	90	45
12	13	18	18	2	2	3	3	5	7	8	8	20	25	25	25	30	35	30	50	30
7	10	12	12	4	5	8	10	10	10	10	10	50	50	55	70	70	75	90	90	40

4. Mengecek normalitas data dari perubahan Functional Ability (Bart_diff) dari pekan pertama (Bart1) sampai pekan terakhir (Bart8) dengan uji statistik yang sesuai dan menginterpretasikannya.

Hasil:

p-value = 0.08875 \rightarrow >0.05 = Data terdistribusi normal

5. Mengecek kesamaan variance dari perubahan Functional Ability (Bart_diff) antara grup intervensi (Group) dengan uji statistic yang sesuai dan menginterpretasikannya.

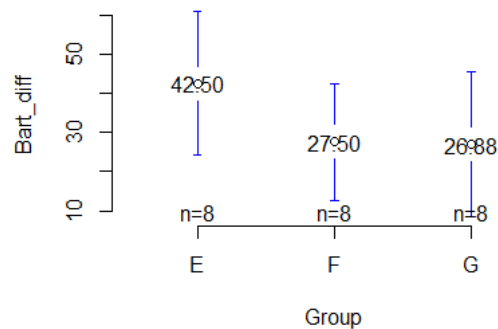
Hasil:

Bartlett test of homogeneity of variances

data: stroke\$Bart_diff and stroke\$Group

Bartlett's K-squared = 0.39433, df = 2, p-value = 0.8211 \rightarrow >0.05 = Varians data adalah sama atau data bersifat homogen.

6. Memplot mean dan 95% Confidence Interval dari nilai perubahan Functional Ability (Bart_diff) berdasarkan grup intervensi (Group) dalam 1 grafik.



7. Melakukan uji anova untuk membandingkan rata-rata(mean) nilai perubahan Functional Ability (Bart_diff) antara 3 grup intervensi (Group) dan menginterpretasikannya.

Hasil:

	Df	Sum Sq	Mean Sq	F	value	Pr(>F)
Group	2	1252	626.0	1.461		0.255
Residuals	21	8997	428.4			

Interpretasi:

p-value $> 0.05 \rightarrow$ tidak terdapat perbedaan yang signifikan pada nilai functional ability antar grup intervensi

8. Melakukan analisis model linear regresi dengan Functional Ability (Bartlet) sebagai outcome(y) dan explanatory variables meliputi: waktu(time/week), grup intervensi (group), dan interaksi waktu dan grup intervensi.

```

Call:
lm(formula = ability ~ as.numeric(time) + Group + as.numeric(time) *
    Group, data = stroke_long)

Residuals:
    Min       1Q   Median       3Q      Max
-47.812 -13.560  -5.305   13.337   63.854

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    29.82143    5.77401   5.165 6.16e-07 ***
as.numeric(time)  6.32440    1.14342   5.531 1.07e-07 ***
GroupF          3.34821    8.16569   0.410  0.6823
GroupG         -0.02232    8.16569  -0.003  0.9978
as.numeric(time):GroupF -1.99405    1.61705  -1.233  0.2191
as.numeric(time):GroupG -2.68601    1.61705  -1.661  0.0984 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.96 on 186 degrees of freedom
Multiple R-squared:  0.2612,    Adjusted R-squared:  0.2413
F-statistic: 13.15 on 5 and 186 DF,  p-value: 5.695e-11

```

Interpretasi:

- Variabel yang berhubungan dengan functional ability score dengan hasil yang signifikan (p-value < 0.01) hanya terdapat pada variabel time → Jika jumlah minggu pemeriksaan bertambah, maka nilai functional ability dapat bertambah sebesar 6.32440.
- Nilai F-Statistic = 13.15 dengan p-value < 0.05, berarti model ini signifikan secara statistik

9. Melakukan ulang Langkah no 8 tanpa variable interaksi di dalam model

```

Call:
lm(formula = ability ~ as.numeric(time) + Group, data = stroke_long)

Residuals:
    Min       1Q   Median       3Q      Max
-49.332 -13.907  -4.532   15.043   58.394

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    36.8415    3.9712   9.277 < 2e-16 ***
as.numeric(time)  4.7644    0.6619   7.198 1.42e-11 ***
GroupF         -5.6250    3.7147  -1.514  0.13164
GroupG        -12.1094    3.7147  -3.260  0.00132 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 21.01 on 188 degrees of freedom
Multiple R-squared:  0.2494,    Adjusted R-squared:  0.2374
F-statistic: 20.82 on 3 and 188 DF,  p-value: 1.084e-11

```

Interpretasi:

- Variabel yang berhubungan dengan functional ability score dengan hasil yang signifikan (p-value < 0.05) terdapat pada variabel time dan group G.
- Nilai F-Statistic = 20.82 dengan p-value < 0.05, berarti model ini signifikan secara statistik

10. Menghitung AIC model no 8 dan 9, serta menginterpretasikan perbandingan nilai AIC nya.

Hasil:

Model selection based on AICc:

	K	AICc	Delta_AICc	AICcwt	Cum.wt	LL		
ability.time.group	5	1720.50	0.00	0.65	0.65	-855.09	➡	No.9
ability.time.group.timegroup	7	1721.74	1.24	0.35	1.00	-853.56	➡	No.8

Interpretasi:

Nilai AIC pada model no 9 (1720.50) lebih kecil dibandingkan nilai AIC pada model no 8 (1721.74)

11. Model no 8 dan 9, manakah yang terbaik? Pilih salah satu kemudian interpretasikan hasil dari analisisnya dari model yang dipilih (hubungan antara variable explanatory dengan outcome)
Model yang terbaik adalah model no 9.

```
call:
lm(formula = ability ~ as.numeric(time) + Group, data = stroke_long)

Residuals:
    Min       1Q   Median       3Q      Max
-49.332 -13.907  -4.532  15.043   58.394

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    36.8415     3.9712   9.277  < 2e-16 ***
as.numeric(time)  4.7644     0.6619   7.198 1.42e-11 ***
GroupF         -5.6250     3.7147  -1.514  0.13164
GroupG        -12.1094     3.7147  -3.260  0.00132 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 21.01 on 188 degrees of freedom
Multiple R-squared:  0.2494,    Adjusted R-squared:  0.2374
F-statistic: 20.82 on 3 and 188 DF,  p-value: 1.084e-11
```

Berdasarkan hasil diatas, terdapat dua variabel explanatory (time and group G) dengan p-value yang signifikan (<0.05) terhadap outcome functional ability.

12. Melakukan analisis mixed model (random intercept) menggunakan package nlme. Functional Ability (Bartlett) sebagai outcome(y) dan explanatory variables meliputi: waktu(time/week), grup intervensi (group), dan Random intercept

```

Linear mixed-effects model fit by REML
Data: stroke_long
      AIC      BIC    logLik
1467.559 1486.978 -727.7796

Random effects:
Formula: ~1 | Subject
(Intercept) Residual
StdDev:      20.1068  8.960882

Fixed effects: ability ~ as.numeric(time) + as.factor(Group)
              Value Std.Error DF   t-value p-value
(Intercept)  36.84152  7.307749 167   5.041432  0.0000
as.numeric(time)  4.76438  0.282241 167  16.880541  0.0000
as.factor(Group)F -5.62500  10.177433  21  -0.552693  0.5863
as.factor(Group)G -12.10937  10.177433  21  -1.189826  0.2474
Correlation:
              (Intr) as.n() a.(G)F
as.numeric(time) -0.174
as.factor(Group)F -0.696  0.000
as.factor(Group)G -0.696  0.000  0.500

Standardized Within-Group Residuals:
      Min       Q1       Med       Q3      Max
-2.17717122 -0.62871793  0.02125076  0.65225196  3.00241837

Number of Observations: 192
Number of Groups: 24

```

13. Melakukan ulang analisis dengan Functional Ability (Bartlett) sebagai outcome(y) dan explanatory variables meliputi: waktu(time/week), grup intervensi (group) dengan General Estimating Equation (GEE) dengan correlation structure:

- Exchangeable

```

Generalized least squares fit by REML
Model: ability ~ as.factor(Group) + as.numeric(time) + as.factor(Group) * as.numeric(time)
Data: stroke_long
      AIC      BIC    logLik
1452.715 1478.521 -718.3573

Correlation Structure: Compound symmetry
Formula: ~1 | Subject
Parameter estimate(s):
      Rho
0.84671

Coefficients:
              Value Std.Error   t-value p-value
(Intercept)  29.821429  7.497378  3.977581  0.0001
as.factor(Group)F  3.348214  10.602894  0.315783  0.7525
as.factor(Group)G -0.022321  10.602894 -0.002105  0.9983
as.numeric(time)   6.324405  0.467228  13.536016  0.0000
as.factor(Group)F:as.numeric(time) -1.994048  0.660760 -3.017809  0.0029
as.factor(Group)G:as.numeric(time) -2.686012  0.660760 -4.065033  0.0001

Correlation:
              (Intr) as.(G)F as.(G)G as.n() a.(G)F:
as.factor(Group)F -0.707
as.factor(Group)G -0.707  0.500
as.numeric(time) -0.280  0.198  0.198
as.factor(Group)F:as.numeric(time)  0.198 -0.280 -0.140 -0.707
as.factor(Group)G:as.numeric(time)  0.198 -0.140 -0.280 -0.707  0.500

Standardized residuals:
      Min       Q1       Med       Q3      Max
-2.1857469 -0.6199072 -0.2425206  0.6097030  2.9190912

Residual standard error: 21.87467
Degrees of freedom: 192 total; 186 residual

```

- Auto regressive

```
Generalized least squares fit by REML
Model: ability ~ as.factor(Group) + as.numeric(time) + as.factor(Group) * as.numeric(time)
Data: stroke_long
      AIC      BIC    logLik
1320.321 1346.127 -652.1607

Correlation Structure: AR(1)
Formula: ~1 | Subject
Parameter estimate(s):
  phi
0.9495754

Coefficients:
              Value Std.Error   t-value p-value
(Intercept)  33.39312   7.937178   4.207178  0.0000
as.factor(Group)F -0.11518  11.224865  -0.010262  0.9918
as.factor(Group)G -6.22568  11.224865  -0.554632  0.5798
as.numeric(time)  6.07484   0.843600   7.201091  0.0000
as.factor(Group)F:as.numeric(time) -2.14085   1.193030  -1.794467  0.0744
as.factor(Group)G:as.numeric(time) -2.23826   1.193030  -1.876112  0.0622

Correlation:
              (Intr) as.(G)F as.(G)G as.n() a.(G)F:
as.factor(Group)F -0.707
as.factor(Group)G -0.707  0.500
as.numeric(time) -0.478  0.338  0.338
as.factor(Group)F:as.numeric(time)  0.338 -0.478 -0.239 -0.707
as.factor(Group)G:as.numeric(time)  0.338 -0.239 -0.478 -0.707  0.500

Standardized residuals:
      Min      Q1      Med      Q3      Max
-2.1430431 -0.5861291 -0.2259572  0.6532219  2.8251592

Residual standard error: 21.42606
Degrees of freedom: 192 total; 186 residual
```

- Unstructure

```
Generalized least squares fit by REML
Model: ability ~ as.factor(Group) + as.numeric(time) + as.factor(Group) * as.numeric(time)
Data: stroke_long
      AIC      BIC    logLik
1338.118 1451.019 -634.0591

Correlation Structure: General
Formula: ~1 | Subject
Parameter estimate(s):
Correlation:
  1  2  3  4  5  6  7
2  0.931
3  0.868 0.931
4  0.789 0.875 0.952
5  0.708 0.819 0.892 0.913
6  0.576 0.731 0.815 0.855 0.965
7  0.426 0.606 0.693 0.782 0.886 0.945
8  0.319 0.522 0.609 0.707 0.840 0.908 0.975

Coefficients:
              Value Std.Error   t-value p-value
(Intercept)  35.71491   7.944761   4.495404  0.0000
as.factor(Group)F -5.51048  11.235588  -0.490449  0.6244
as.factor(Group)G -11.30440  11.235588  -1.006124  0.3157
as.numeric(time)  6.69319   1.166119   5.739713  0.0000
as.factor(Group)F:as.numeric(time) -3.23684   1.649141  -1.962742  0.0512
as.factor(Group)G:as.numeric(time) -3.85733   1.649141  -2.338991  0.0204

Correlation:
              (Intr) as.(G)F as.(G)G as.n() a.(G)F:
as.factor(Group)F -0.707
as.factor(Group)G -0.707  0.500
as.numeric(time) -0.760  0.537  0.537
as.factor(Group)F:as.numeric(time)  0.537 -0.760 -0.380 -0.707
as.factor(Group)G:as.numeric(time)  0.537 -0.380 -0.760 -0.707  0.500

Standardized residuals:
      Min      Q1      Med      Q3      Max
-2.47104066 -0.50835338 -0.03544121  0.82609508  2.70723746

Residual standard error: 21.27331
Degrees of freedom: 192 total; 186 residual
```

14. Mengingat GEE tidak dapat mengeluarkan AIC, dengan menggunakan statement gls, menghitung AIC dari model GLS dengan ketiga struktur korelasi di atas (Exchangeable, Auto regressive, dan Unstructure).

	df	AIC
gls.exch	8	1452.715
gls.ar1	8	1320.321
gls.un	35	1338.118

15. Membuat tabel untuk Membandingkan AIC dari model dengan korelasi struktur Exchangeable, Auto regressive, dan Unstructure, dengan AIC linear regresi model (Model dari instruksi no 9). Interpretasikan dan simpulkan.

	df	AIC
gls.exch	8	1452.715
gls.ar1	8	1320.321
gls.un	35	1338.118
Model no.9	188	1720.50

Berdasarkan hasil diatas, nilai AIC yang terkecil adalah pada auto regresive (gls.ar1) dengan nilai AIC 1320.321. Maka ini merupakan yang paling optimal untuk digunakan/dipilih.