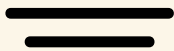
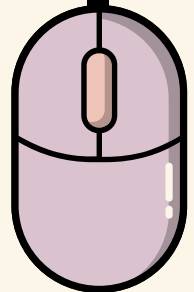


# **"Model Linear Mixed-Effects pada Kadar Gula di Negara Asia Tenggara"**

Kelompok 3 - Kapita Selekta Statistika 7B





# Anggota Kelompok 3



**01**

**Rosa Amalia  
Nursinta**

11190940000041

**02**

**Elviana  
Saputri**

11190940000043

**03**

**Meissy  
Astariva P**

11190940000063

## Nomor 1a



Model linear mixed effects yang telah kita diskusikan dapat dituliskan sebagai berikut:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i, \quad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{D}), \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i})$$

Tuliskan persamaan regresi linear mixed effects jika diasumsikan respon Y bergantung pada waktu secara linear, dimana setiap kelompok memiliki evolusi respon yang berbeda-beda, dengan random intercept dan random slope.

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 T_i + \beta_3 t_{ij} T_i + b_{0i} + b_{1i} t_{ij} + e_{ij}$$

## Nomor 1b



Jika diberikan subset data longitudinal dari 3 individu seperti tabel diatas, tuliskan matrik  $X_i$ ,  $\beta$ ,  $Z_i$ ,  $b_i$  dari persamaan regresi di 1(a) untuk setiap individu  $i = 1, 2, 3$

### Individu 1

$$X_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 6 & 0 & 0 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{bmatrix}$$

$$b_1 = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix}$$

### Individu 2

$$X_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 1 & 4 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

$$b_2 = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix}$$

### Individu 3

$$X_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 3 \\ 1 & 4 & 1 & 4 \\ 1 & 6 & 1 & 6 \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \end{bmatrix}$$

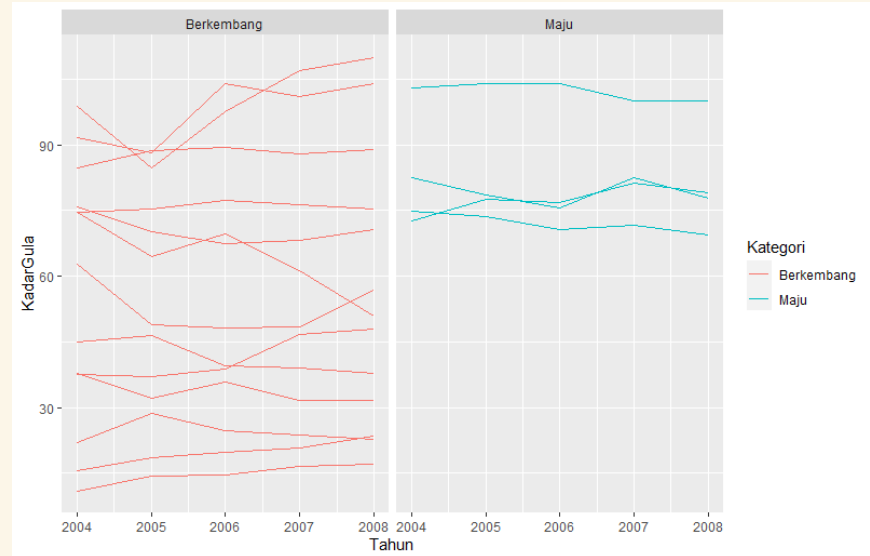
$$b_2 = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix}$$

## Nomor 2a – Plot Profil Subject



Tampilkan kembali plot spaghetti yang menampilkan profil subjek di dataset yang Anda gunakan. Berikan tanggapan Anda, apa bentuk efek tetap dan efek random yang akan Anda gunakan dalam model mixed effects? (contoh: apakah respon akan dimodelkan menggunakan trend linear saja atau trend kuadratik? apakah ada pengaruh kelompok? apakah ada indikasi menggunakan random intercept saja atau perlu random intercept dan slope?)

- Kami mengasumsikan data dari tahun ke tahun memiliki tren yang cenderung linear
- Berdasarkan letak intercept dan kemiringan garis, maka kami menetapkan efek random intercept dan random slope.

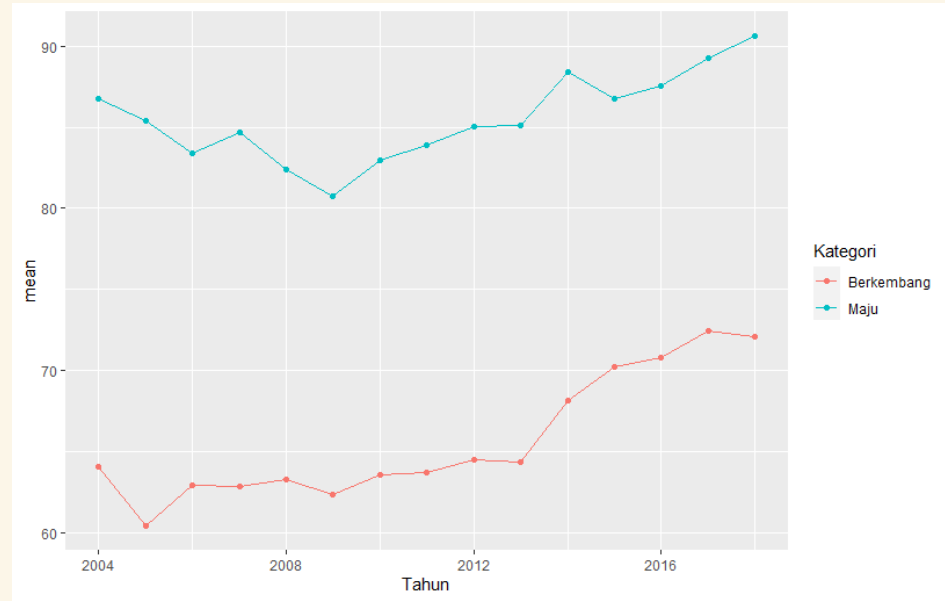


## Nomor 2a – Plot Profil Mean

### Interpretasi :

Mean dari waktu ke waktu pada kedua kelompok terlihat tidak parallel, maka kami mengasumsikan terdapat pengaruh interaksi tahun dan kelompok.

Sehingga model awal yang akan kami bentuk yaitu model tren linear dengan kovariat terdapat pengaruh interaksi tahun dan kelompok dan menggunakan efek random intercept dan random slope.



## Nomor 2b – Soal



Modelkan dataset Anda dengan menggunakan efek tetap yang Anda sebutkan di 2(a) dan dengan menggunakan efek random:

- I. Random intercept saja
- II. Random intercept dan random slope yang berkorelasi
- III. Random intercept dan random slope yang saling independent

## Nomor 2b (i, ii, dan iii)



### (I) Random intercept saja

```
> lme_fit3 <- lme(KadarGula ~ Tahun*Kategori, data = data,
+               random = ~ 1 | Negara)
> summary(lme_fit3)
Linear mixed-effects model fit by REML
Data: data
      AIC      BIC    loglik
588.9778 603.3445 -288.4889

Random effects:
Formula: ~1 | Negara
(Intercept) Residual
StdDev:    26.75317 4.854354

Fixed effects: KadarGula ~ Tahun * Kategori
              Value Std. Error DF t-value p-value
(Intercept) -587.7554 854.0970 66 -0.6881600 0.4938
Tahun        0.3208   0.4258 66 0.7534127 0.4539
KategoriMaju 1257.2654 1760.7661 15 0.7140445 0.4862
Tahun:KategoriMaju -0.6133 0.8777 66 -0.6987100 0.4872

Correlation:
      (Intercept) Tahun KategoriMaju
Tahun      -1.000
KategoriMaju -0.485 0.485
Tahun:KategoriMaju 0.485 -0.485 -1.000
```

### (II) Random intercept & random slope yang berkorelasi

```
> lme_fit6 <- lme(KadarGula ~ Tahun*Kategori, data = data,
+               random = ~ 1 + Tahun | Negara)
> summary(lme_fit6)
Linear mixed-effects model fit by REML
Data: data
      AIC      BIC    loglik
592.8883 612.0439 -288.4442

Random effects:
Formula: ~1 + Tahun | Negara
Structure: General positive-definite, Log-Cholesky parametrization
              StdDev Corr
(Intercept) 2.883392e-05 (Intrc)
Tahun       1.333689e-02 0.943
Residual    4.851042e+00

Fixed effects: KadarGula ~ Tahun * Kategori
              Value Std. Error DF t-value p-value
(Intercept) -587.7554 853.4820 66 -0.6886558 0.4935
Tahun        0.3208   0.4255 66 0.7538986 0.4536
KategoriMaju 1257.2654 1759.4983 15 0.7145590 0.4859
Tahun:KategoriMaju -0.6133 0.8772 66 -0.6991606 0.4869

Correlation:
      (Intercept) Tahun KategoriMaju
Tahun      -1.000
KategoriMaju -0.485 0.485
Tahun:KategoriMaju 0.485 -0.485 -1.000
```

### (III) Random intercept dan random slope yang saling independent

```
> lme_fit12 <- lme(KadarGula ~ Tahun*Kategori, data = data,
+               random = list(Negara = pdDiag(form = ~ Tahun)))
> summary(lme_fit12)
Linear mixed-effects model fit by REML
Data: data
      AIC      BIC    loglik
590.8883 607.6495 -288.4442

Random effects:
Formula: ~Tahun | Negara
Structure: Diagonal
              (Intercept) Tahun Residual
StdDev:    0.02511061 0.01333689 4.851042

Fixed effects: KadarGula ~ Tahun * Kategori
              Value Std. Error DF t-value p-value
(Intercept) -587.7554 853.4821 66 -0.6886558 0.4935
Tahun        0.3208   0.4255 66 0.7538986 0.4536
KategoriMaju 1257.2654 1759.4983 15 0.7145590 0.4859
Tahun:KategoriMaju -0.6133 0.8772 66 -0.6991606 0.4869

Correlation:
      (Intercept) Tahun KategoriMaju
Tahun      -1.000
KategoriMaju -0.485 0.485
Tahun:KategoriMaju 0.485 -0.485 -1.000
```

Model Fixed Effect (i)  $\rightarrow Y_{ij} = (-587.7554 + b_{0i}) + 0.3208t_{ij} + 1257.2654 \times T - 0.6133 t_{ij} \times T + e_{ij}$

Model Fixed Effect (ii)  $\rightarrow Y_{ij} = (-587.7554 + b_{0i}) + (0.3208 + b_{1i})t_{ij} + 1257.2654 \times T - 0.6133 t_{ij} \times T + e_{ij}$

Model Fixed Effect (iii)  $\rightarrow Y_{ij} = (-587.7554 + b_{0i}) + (0.3208 + b_{1i})t_{ij} + 1257.2654 \times T - 0.6133 t_{ij} \times T + e_{ij}$



## Nomor 2b (i, ii, dan iii)



- Memilih model

```
> anova(lme.fit3, lme.fit6, lme.fit12)
```

|                  | Model | df | AIC      | BIC      | logLik    | Test   | L.Ratio    | p-value |
|------------------|-------|----|----------|----------|-----------|--------|------------|---------|
| <u>lme.fit3</u>  | 1     | 6  | 588.9778 | 603.3445 | -288.4889 |        |            |         |
| <u>lme.fit6</u>  | 2     | 8  | 592.8883 | 612.0439 | -288.4442 | 1 vs 2 | 0.08949107 | 0.9562  |
| <u>lme.fit12</u> | 3     | 7  | 590.8883 | 607.6495 | -288.4442 | 2 vs 3 | 0.00000001 | 0.9999  |

- Perbandingan model fit12 dengan fit6

$$H_0 : \text{Model } lme.fit6$$

$$H_1 : \text{Model } lme.fit12$$

Perbandingan model fit12 dengan fit6 memiliki nilai p-value 0.9999, maka  $H_0$  diterima yang artinya model fit6 adalah model yang lebih baik.

- Perbandingan model fit6 dengan fit3

$$H_0 : \text{Model } lme.fit3$$

$$H_1 : \text{Model } lme.fit6$$

Kemudian perbandingan model fit6 dengan fit3 memiliki nilai p-value 0.9562, maka  $H_0$  diterima yang artinya model fit3 adalah model yang lebih baik.

Maka berdasarkan nilai p-value dan nilai AIC, kami memilih model *lme.fit3* sebagai model terbaik.

## Nomor 2c



**Soal :** Dari efek random terpilih di soal 2(b), interpretasikan estimasi standard deviasi komponen efek random dan residualnya. Berapa besar kontribusi efek random (intercept/slope) dalam menjelaskan total variabilitas respon Y?

### Interpretasi Estimasi Standard Deviasi ;

- Estimasi standard deviasi dari efek-subjek (negara), atau  $\sigma_b$  dari  $b_i \sim N(0, \sigma_b^2)$ , yaitu  $\sigma_b = 26.75317$  yang dimana parameter  $\sigma_b$  menyatakan bagaimana variabel respon (Kadar Gula) bervariasi dari satu individu ke individu lainnya.
- Total variasi residual (variabilitas respon yang tidak dapat dijelaskan) adalah 4.854354.

Kita dapat memprediksi kadar gula berdasarkan trend waktu, tetapi setiap individu memiliki prediksi kadar gula tunggal (unik) dan  $\sigma_b$  merupakan rata-rata penyimpangan dari keseluruhan individu. Variabilitas karena efek-subjek berkontribusi sebesar 84% ( $26.75317 / (26.75317 + 4.854354)$ ) dari total variabilitas di variabel respon.

```
Summary lme.fit3
> summary(lme.fit3)
Linear mixed-effects model fit by REML
Data: data
      AIC      BIC    loglik
588.9778 603.3445 -288.4889

Random effects:
Formula: ~1 | Negara
      (Intercept) Residual
StdDev:    26.75317  4.854354

Fixed effects: KadarGula ~ Tahun * Kategori
              Value Std. Error  DF  t-value p-value
(Intercept) -587.7554   854.0970  66 -0.6881600  0.4938
Tahun        0.3208    0.4258   66  0.7534127  0.4539
KategoriMaju 1257.2654 1760.7661  15  0.7140445  0.4862
Tahun:KategoriMaju -0.6133    0.8777  66 -0.6987100  0.4872
Correlation:
              (Intr) Tahun_KtgrMi
Tahun        -1.000
KategoriMaju -0.485  0.485
Tahun:KategoriMaju 0.485 -0.485 -1.000

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.9235815 -0.5081764 -0.0366385  0.4306158  2.2941943

Number of Observations: 85
Number of Groups: 17
```

## Nomor 2d



i.) Model linear terdapat pengaruh interaksi tahun dan kelompok menggunakan efek random intercept saja.

```
> lme_fit3 <- lme(KadarGula ~ Tahun*Kategori, data = data,
+               random = ~ 1 | Negara)
> summary(lme_fit3)
Linear mixed-effects model fit by REML
Data: data
   AIC      BIC    logLik
588.9778 603.3445 -288.4889

Random effects:|
Formula: ~1 | Negara
(Intercept) Residual
StdDev:      26.75317 4.854354

Fixed effects: KadarGula ~ Tahun * Kategori
              Value Std.Error DF   t-value p-value
(Intercept) -587.7554  854.0970 66 -0.6881600  0.4938
Tahun        0.3208   0.4258 66  0.7534127  0.4539
KategoriMaju 1257.2654 1760.7661 15  0.7140445  0.4862
Tahun:KategoriMaju -0.6133  0.8777 66 -0.6987100  0.4872
Correlation:
      (Intr) Tahun KtrMj
Tahun -1.000
KategoriMaju -0.485 0.485
Tahun:KategoriMaju 0.485 -0.485 -1.000

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.9235815 -0.5081764 -0.0366385  0.4306158  2.2941943

Number of Observations: 85
Number of Groups: 17
```

ii.) Model linear dengan kovariat tahun dan kelompok tanpa pengaruh interaksi

```
> lme_fit2 <- lme(KadarGula ~ Tahun + Kategori, data = data,
+               random = ~ 1 | Negara)
> summary(lme_fit2)
Linear mixed-effects model fit by REML
Data: data
   AIC      BIC    logLik
589.0411 601.0747 -289.5205

Random effects:
Formula: ~1 | Negara
(Intercept) Residual
StdDev:      26.75384 4.835778

Fixed effects: KadarGula ~ Tahun + Kategori
              Value Std.Error DF   t-value p-value
(Intercept) -298.29231 744.0370 67 -0.4009106  0.6898
Tahun        0.17647   0.3709 67  0.4758066  0.6358
KategoriMaju 27.04731 15.3470 15  1.7623854  0.0984
Correlation:
      (Intr) Tahun
Tahun -1.000
KategoriMaju -0.005 0.000

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.96510568 -0.51839417 -0.02779329  0.43893104  2.24324038

Number of Observations: 85
Number of Groups: 17
```

iii.) Model linear dengan kovariat tahun saja

```
> lme_fit1 <- lme(KadarGula ~ Tahun, data = data,
+               random = ~ 1 | Negara)
> summary(lme_fit1)
Linear mixed-effects model fit by REML
Data: data
   AIC      BIC    logLik
597.3193 606.9946 -294.6596

Random effects:
Formula: ~1 | Negara
(Intercept) Residual
StdDev:      28.47098 4.835778

Fixed effects: KadarGula ~ Tahun
              Value Std.Error DF   t-value p-value
(Intercept) -291.92824 744.0320 67 -0.3923598  0.6960
Tahun        0.17647   0.3709 67  0.4758066  0.6358
Correlation:
      (Intr)
Tahun -1

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.9795090 -0.5042945 -0.0340836  0.4604343  2.2343709

Number of Observations: 85
Number of Groups: 17
```

## Nomor 2d



```
> anova(lme.fit1, lme.fit2, lme.fit3)
```

|                 | Model | df | AIC      | BIC      | loglik    | Test   | L.Ratio   | p-value |
|-----------------|-------|----|----------|----------|-----------|--------|-----------|---------|
| <u>lme.fit1</u> | 1     | 4  | 597.3193 | 606.9946 | -294.6596 |        |           |         |
| <u>lme.fit2</u> | 2     | 5  | 589.0411 | 601.0747 | -289.5205 | 1 vs 2 | 10.278192 | 0.0013  |
| <u>lme.fit3</u> | 3     | 6  | 588.9778 | 603.3445 | -288.4889 | 2 vs 3 | 2.063244  | 0.1509  |

- Memilih Model Terbaik

- Perbandingan model fit3 dengan fit2

$H_0$  : Model *lme.fit2*

$H_1$  : Model *lme.fit3*

Perbandingan model fit3 dengan fit2 memiliki nilai p-value 0.1509, maka  $H_0$  diterima yang artinya model fit2 adalah model yang lebih baik.

- Perbandingan model fit2 dengan fit1

$H_0$  : Model *lme.fit1*

$H_1$  : Model *lme.fit2*

Kemudian perbandingan model fit2 dengan fit1 memiliki nilai p-value 0.0013, maka  $H_0$  ditolak yang artinya model fit2 adalah model yang lebih baik.

Maka berdasarkan nilai p-value, kami memilih model *lme.fit2* sebagai model terbaik yaitu Model linear dengan kovariat tahun dan kelompok tanpa pengaruh interaksi menggunakan efek random intercept saja

### Interpretasi Koefisien Regresi Model *lme.fit2*

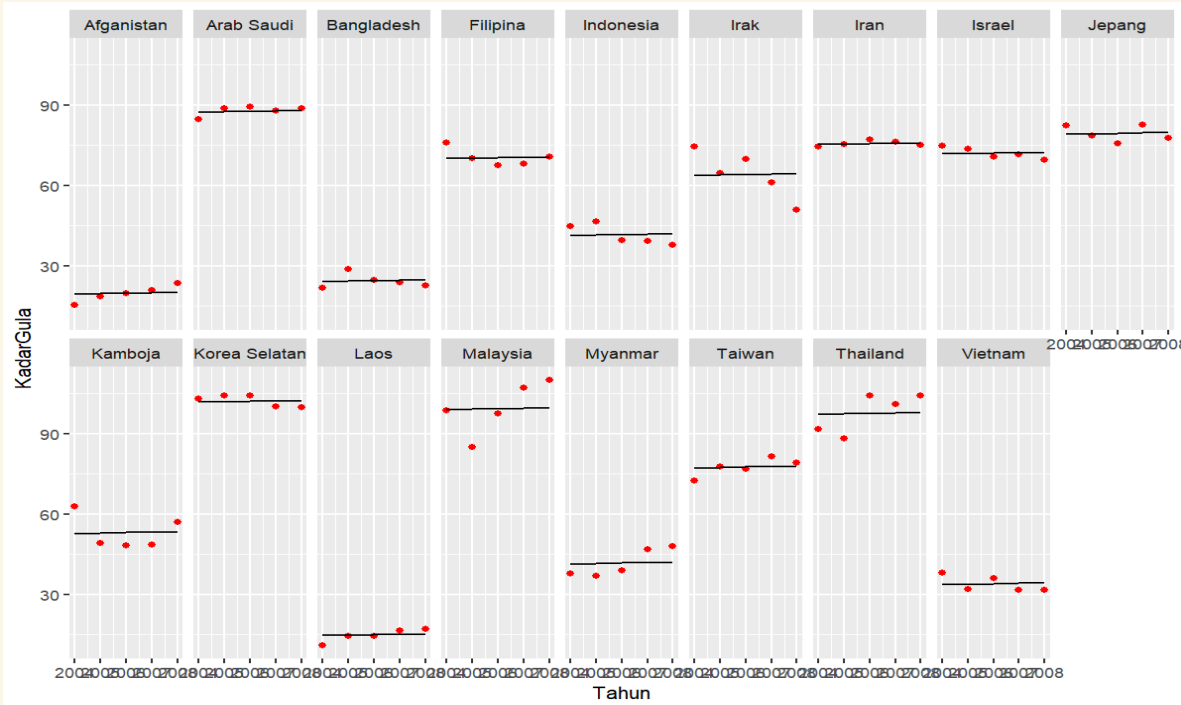
$$Y_{ij} = \beta_0 + \beta_1 \text{tahun}_{ij} + \beta_2 \text{KategoriMaju} + b_{0i} + b_{1i} \text{tahun}_{ij} + e_{ij}$$

$$Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) \text{tahun} + \beta_2 \text{KategoriMaju} + e_{ij}$$

$$\text{Kadar Gula}_{ij} = (-298.29231 + b_{0i}) + (0.17647 + b_{1i}) \text{tahun} + 27.04731 \text{KategoriMaju} + e_{ij}$$

- $\beta_0 + b_0$  artinya tingkat kadar gula pada waktu 0 tahun dan pada negara berkembang yaitu sebesar  $\beta_0 + b_0$
- $\beta_1 + b_1$  artinya setiap waktu bertambah 1 tahun, tingkat kadar gula meningkat sebesar  $\beta_1 + b_1$  untuk kategori negara yang sama.
- $\beta_2 = 27.04731$  artinya selisih tingkat kadar gula pada negara maju dan berkembang sebesar 27.04731 pada tahun yang sama.

## Nomor 2e - Visualisasi



**Merah : Data Aktual**  
**Hitam : Data Prediksi**

Perbandingan data aktual dengan prediksi spesifik-subjek (setiap garis mempunyai slope yang sama, tapi dengan intercept yang berbeda)



**Terima  
Kasih**

>>>>

~~~~~  
.....