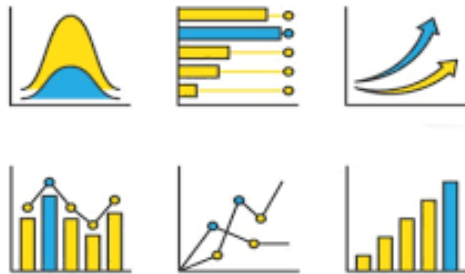


# KUANTITATIF DAN KUALITATIF LINGKUNGAN (STATISTIKA LINGKUNGAN)

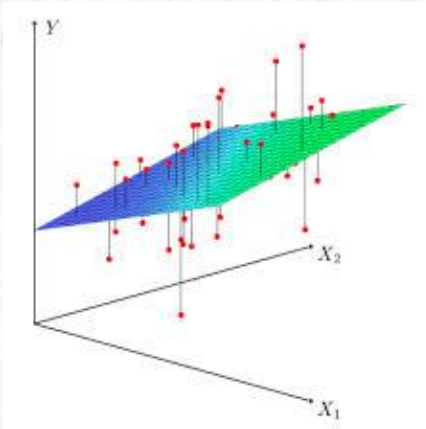
## STATISTICS FOR DATA SCIENCE



Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

PROGRAM STUDI MAGISTER ILMU LINGKUNGAN  
PASCASARJANA UNIVERSITAS JAMBI

# KUANTITATIF DAN KUALITATIF LINGKUNGAN (STATISTIKA LINGKUNGAN)



## STATISTIKA INFERENSIAL (Analisis Regresi Linear Berganda)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \varepsilon$$

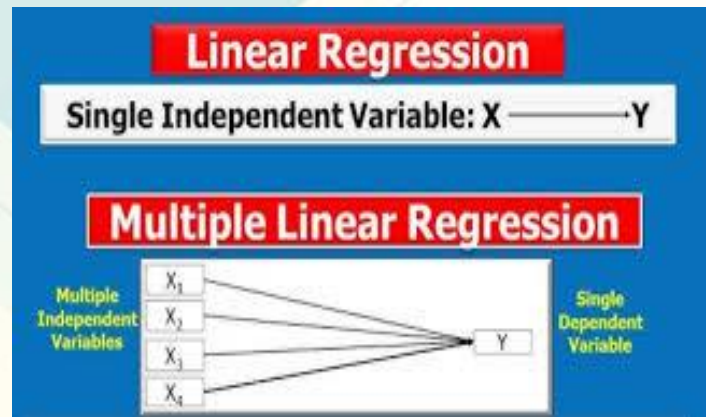
Diagram illustrating the components of the multiple regression equation:

- Dependent Variable (Response Variable) points to  $Y$ .
- Independent Variables (Predictors) points to  $X_1$  and  $X_2$ .
- Y intercept points to  $\beta_0$ .
- Slope Coefficient points to  $\beta_1$  and  $\beta_2$ .
- Error Term points to  $\varepsilon$ .

Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

PROGRAM STUDI MAGISTER ILMU LINGKUNGAN  
PASCASARJANA UNIVERSITAS JAMBI

## (Multiple Linear Regression)



Regresi linier berganda merupakan teknik statistik yang digunakan untuk memprediksi hasil suatu variabel berdasarkan nilai dua variabel atau lebih..

Model memperlihatkan hubungan antara satu variabel terikat (dependent variable) dgn **beberapa** variabel bebas (independent variables).

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i$$

dimana:  $i = 1, 2, 3, \dots, N$  (banyaknya pengamatan)

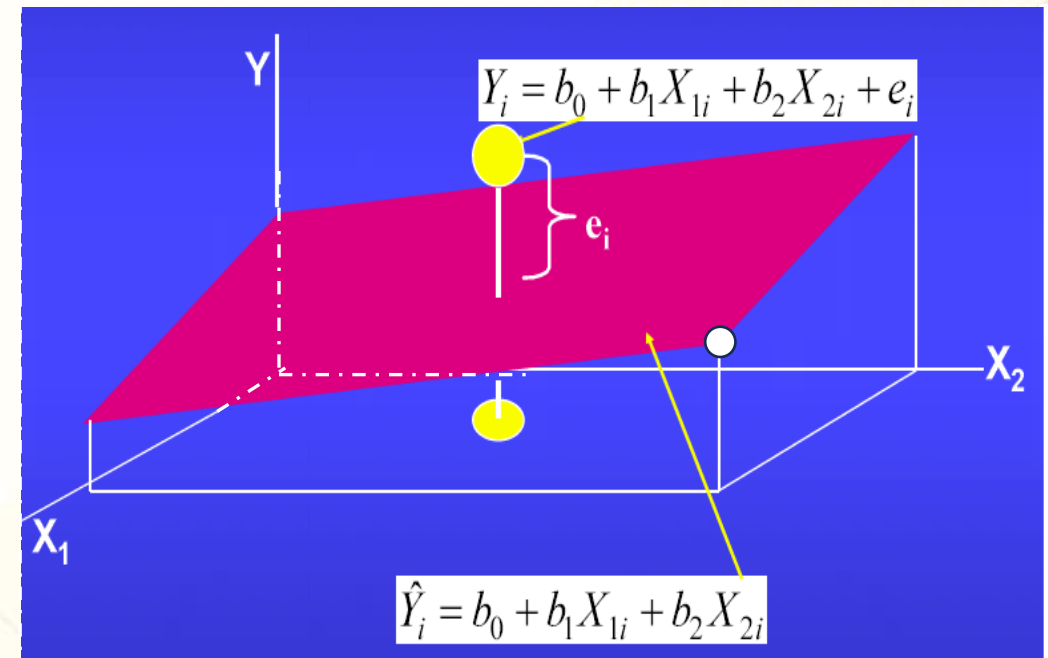
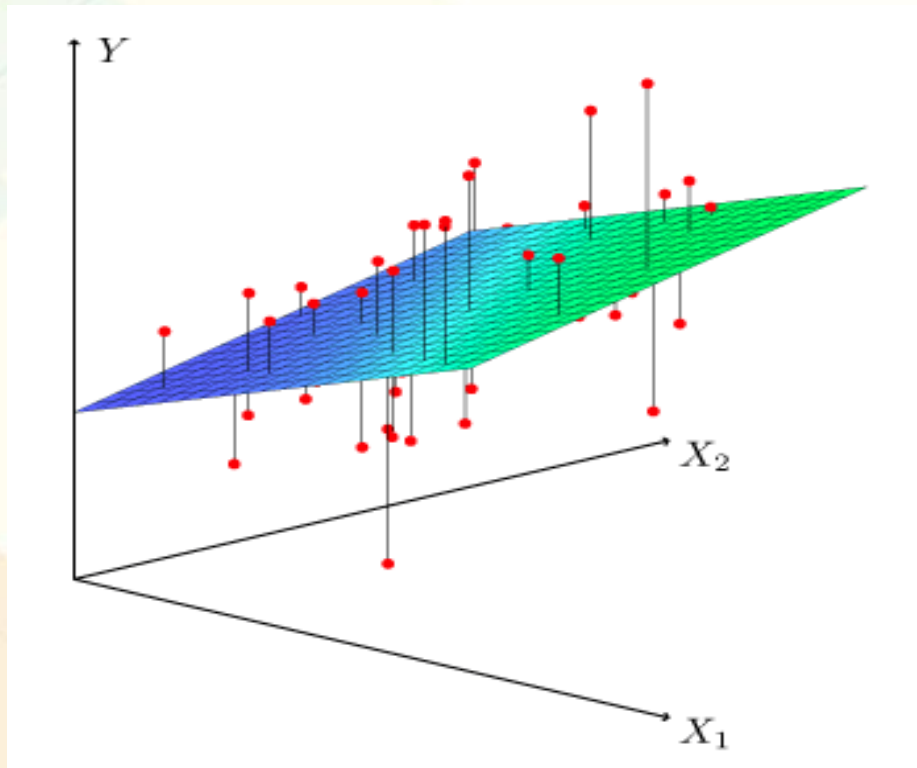
$\beta_0, \beta_1, \beta_2, \dots, \beta_k$  adalah parameter yang nilainya diduga melalui model:

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki} + e_i$$

dan hasil pendugaannya adalah

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki}$$

## PENDUGAAN vs HASIL PENDUGAAN SAMPEL



## ESTIMASI MODEL REGRESI LINIER BERGANDA

Model :  $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$

Model penduga :  $Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + e_i$

$b_0$ ,  $b_1$  dan  $b_2$  nilai penduga untuk  $\beta_0$ ,  $\beta_1$  dan  $\beta_2$ .

$$b_1 = \frac{(\sum y_i x_{1i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{1i} x_{2i})}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2}$$

$$b_2 = \frac{(\sum y_i x_{2i})(\sum x_{1i}^2) - (\sum y_i x_{1i})(\sum x_{1i} x_{2i})}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2}$$

$b_0 = \bar{Y}_i - b_1 \bar{X}_{1i} - b_2 \bar{X}_{2i}$   $\longrightarrow$   $b_0$  adalah intersep atau nilai Y bila X1 dan X2 sama dengan nol.

$$\text{var}(b_0) = \left[ \frac{1}{n} + \frac{\bar{X}_1^2 \sum x_{2i}^2 - \bar{X}_2^2 \sum x_{1i}^2 - 2 \bar{X}_1 \bar{X}_2 \sum x_{1i} x_{2i}}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2} \right] \sigma^2$$

$$\text{var}(b_1) = \left[ \frac{\sum x_{1i}^2}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2} \right] \sigma^2$$

$$\sigma^2 = \frac{\sum \mu_i^2}{n - 3}$$

$$\text{se}(b_i) = \sqrt{\text{var}(b_i)}$$

Utk  $i = 0, 1, 2.$

$$\sum \mu_i^2 = \sum y_i^2 - b_1 \sum y_i x_{1i} - b_2 \sum y_i x_{2i}$$

## Asumsi-asumsi Model Regresi Linier Berganda

(Agar hasil estimasi dapat diinterpretasikan dengan baik - BLUE)

- ✓ Nilai rata-rata disturbance term adalah nol,  
 $E(\mu_i) = 0$ .
- ✓ Tidak terdapat **serial korelasi (otokorelasi)** antar  $\mu_i$   
 $Cov(\mu_i, \mu_j) = 0$  untuk  $i \neq j$ .
- ✓ Sifat **homoskedastisitas**:  
 $Var(\mu_i) = \sigma^2$  sama untuk setiap  $i$
- ✓ Covarianse antara  $\mu_i$  dan setiap variabel bebas adalah nol.  $Cov(\mu_i, X_j) = 0$
- ✓ Tidak terdapat **multikolinieritas** antar variabel bebas.
- ✓ Model dispesifikasi dengan baik



- Variabel Dummy adalah variabel yg merepresentasikan kuantifikasi dari variabel kualitatif.  
Misal: jenis kelamin, pendidikan, lokasi, situasi, musim, & kualitas.
- Variabel dummy sering juga disebut variabel boneka, binary, kategorik atau dikotom.
- Jika data kualitatif tsb memiliki  $m$  kategori, maka jumlah variabel dummy yang dicantumkan di dalam model adalah  $(m-1)$ .
- Dummy memiliki nilai 1 ( $D=1$ ) utk salah satu kategori dan nol ( $D=0$ ) untuk kategori yang lain.
- Kesimpulan yang diambil dari keberadaan variabel dummy di dalam model adalah **perbedaan nilai** antar kategori yang bersangkutan.

- Variabel dummy digunakan sebagai upaya untuk melihat bagaimana kategori-kategori dalam sampel berpengaruh terhadap parameter pendugaan.
- Variabel dummy juga mencoba membuat kuantifikasi dari variabel kualitatif.

Pertimbangkan model berikut ini:

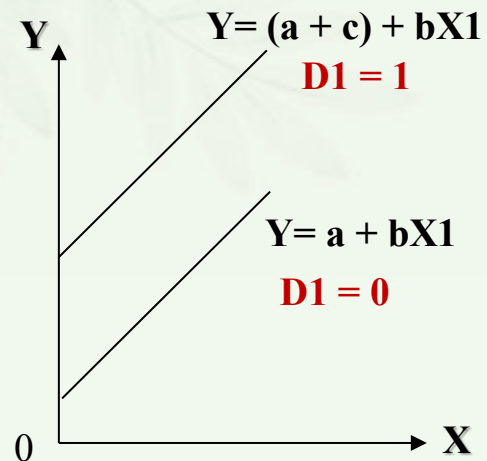
I.  $Y = a + bX + cD$  (**Model Dummy Intersep**)

II.  $Y = a + bX + c(DX)$  (**Model Dummy Slope**)

III.  $Y = a + bX + c(DX) + dD$  (**Kombinasi**)

### Model Dummy Intersep

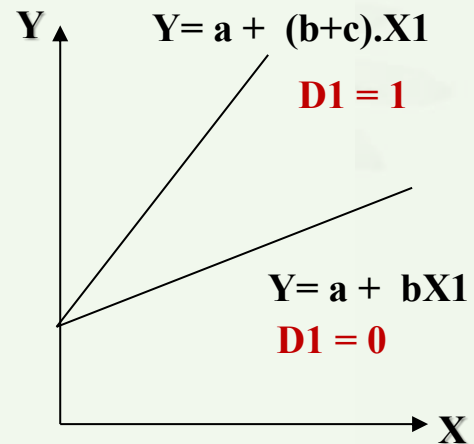
$$Y = a + bX_1 + cD_1$$



Dummy Intersep

### Model Dummy Slope

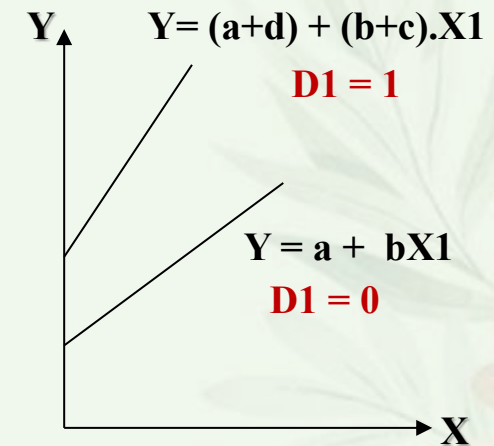
$$Y = a + bX_1 + cD_1X_1$$



Dummy Slope

### Model Dummy Kombinasi

$$Y = a + bX_1 + cD_1X_1 + dD_1$$



Dummy Kombinasi

## Dummy sebagai Variabel Bebas:

ANOVA Model:

$$Y_i = \alpha + \beta D_i + \mu$$

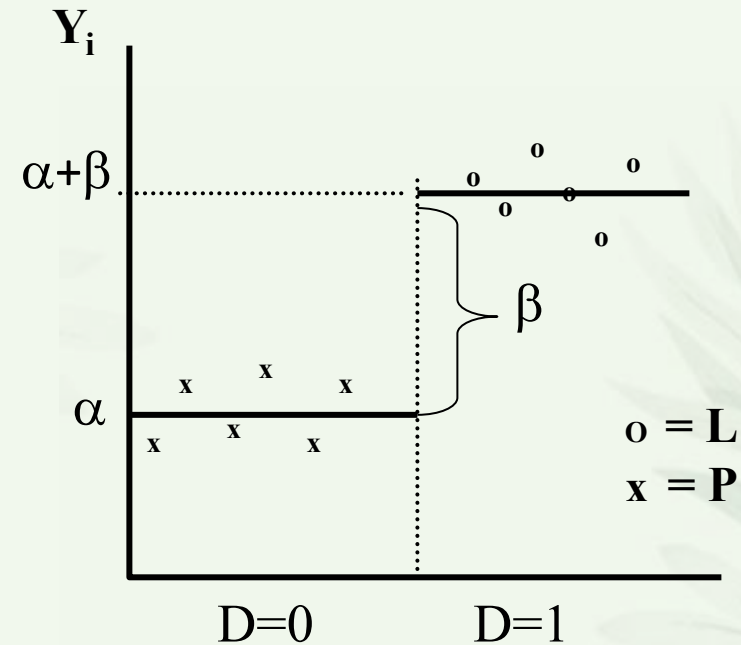
Misal :

$Y_i$  = Penghasilan Karyawan

$D_i = 1$  untuk laki-laki  
= 0 untuk wanita

$$E(Y_i | D_i=0) = \alpha$$

$$E(Y_i | D_i=1) = \alpha + \beta$$



Interpretasi:

Apakah jenis kelamin berpengaruh thdp penghasilan. Berapa perbedaan penghasilan antara laki2 dan wanita.

## Dummy sebagai Variabel Bebas:

(...lanjutan)

**ANCOVA Model:** (gabungan kuantitatif & kualitatif)

1. Satu kuantitatif, satu kualitatif dg 2 kategori.

$$Y_i = \alpha_1 + \alpha_2 D_i + \beta X_i + \mu$$

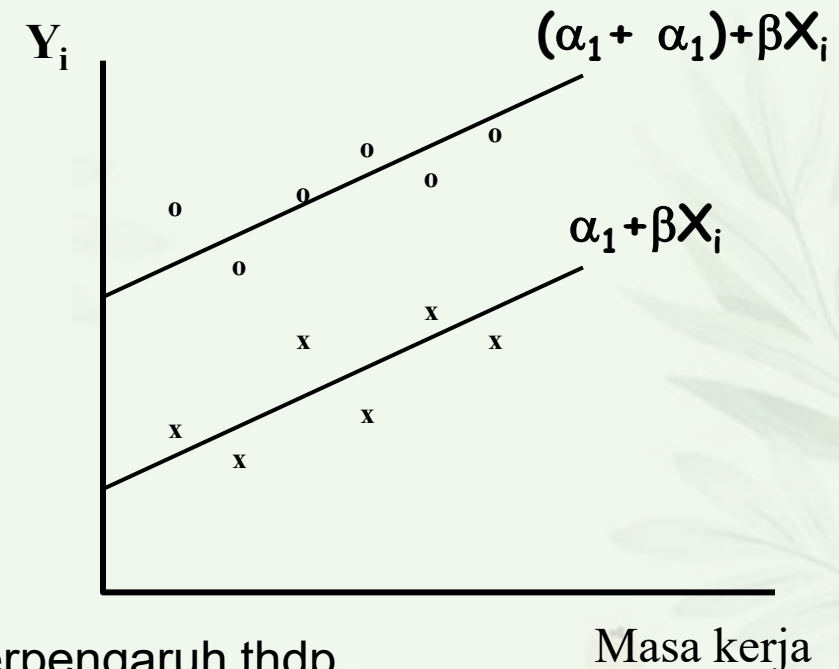
$X_i$  = Masa kerja

$$E(Y_i | X_i, D_i=0) = \alpha_1 + \beta X_i$$

$$E(Y_i | X_i, D_i=1) = (\alpha_1 + \alpha_2) + \beta X_i$$

**Interpretasi:**

Apakah jenis kelamin dan masa kerja berpengaruh thdp penghasilan. Pada masa kerja tertentu, brp perbedaan penghasilan antara Laki dan wanita.



## Dummy sebagai Variabel Bebas: (...lanjutan)

### 2. Satu kuantitatif, satu kualitatif dg 3 kategori.

Misal: Selain masa kerja, penghasilan karyawan juga dipengaruhi oleh tingkat pendidikan (tdk tamat SMU, tamat SMU, PT)

$$Y_i = \alpha_1 + \alpha_2 D1_i + \alpha_3 D2_i + \beta X_i + \mu$$

$D1_i = 1$  untuk tamat SMU  
 $= 0$  Lainnya

$D2_i = 1$  untuk PT  
 $= 0$  Lainnya

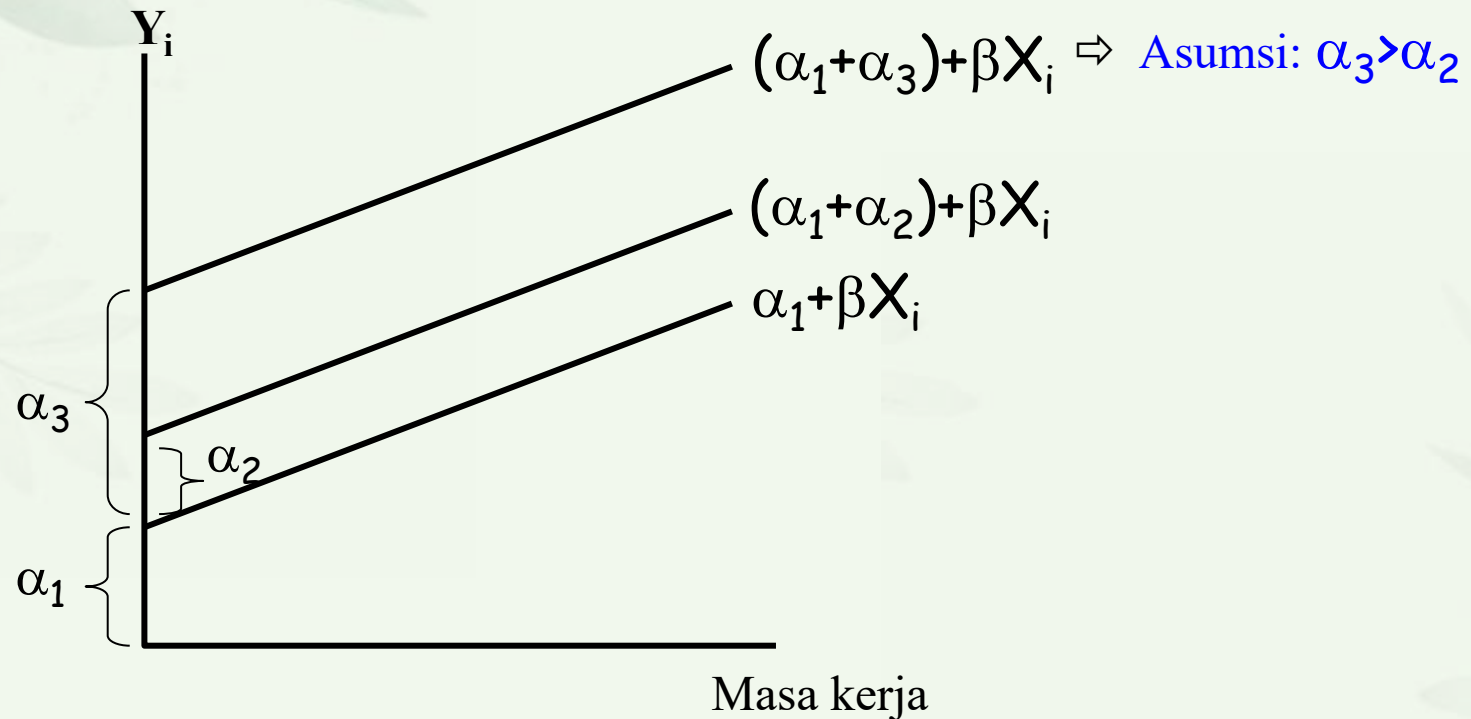
Sebagai kategori dasar adlh tidak tamat SMU

$$E(Y_i | X_i, D1_i=0, D2_i=0) = \alpha_1 + \beta X_i \quad (\text{tdk tamat SMU})$$

$$E(Y_i | X_i, D1_i=1, D2_i=0) = (\alpha_1 + \alpha_2) + \beta X_i \quad (\text{Tamat SMU})$$

$$E(Y_i | X_i, D1_i=0, D2_i=1) = (\alpha_1 + \alpha_3) + \beta X_i \quad (\text{PT})$$

Satu kuantitatif, satu kualitatif dg 3 kategori.



### Interpretasi:

Apakah Masa kerja dan tkt pendidikan berpengaruh thdp penghasilan?  
Brp besar perbedaan penghasilan menurut tkt pendidikan pd masa kerja tertentu?.

## Dummy sebagai Variabel Bebas: (...lanjutan)

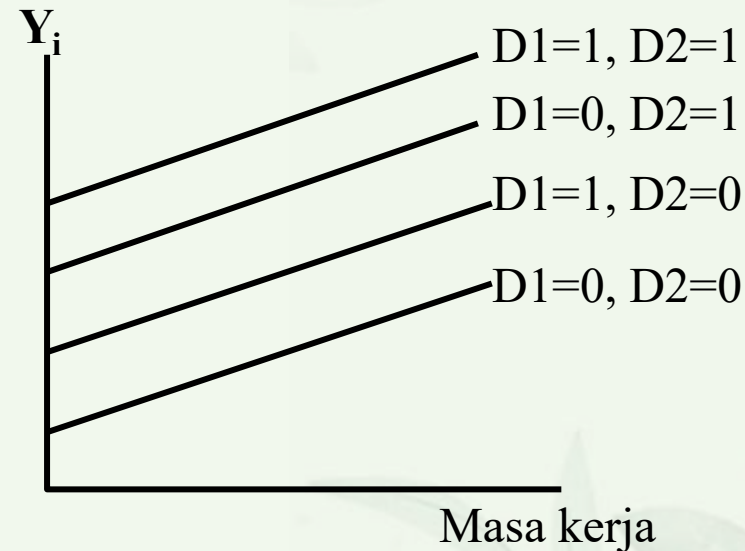
### 3. Satu kuantitatif, dua kualitatif masing-masing dg 2 kategori.

Misal: D1 adalah dummy jenis kelamin (laki/wanita), dan D2 adlh dummy tempat kerja (kota/desa).

$$Y_i = \alpha_1 + \alpha_2 D1_i + \alpha_3 D2_i + \beta X_i + \mu$$

$D1_i = 1$  untuk Laki-laki  
 $= 0$  untuk wanita

$D2_i = 1$  untuk kota  
 $= 0$  untuk desa



No.	Produktivitas (ton/ha/thn)	Luas kebun (ha)	Umur Tan. (thn)	Keaktifan Petani	Pendidikan Petani
1	23.8	3.0	16	Aktif	SMA
2	18.9	2.5	15	Aktif	SMA
3	14.7	2.0	12	Tdk Aktif	SMP
4	18.4	2.0	18	Aktif	SD
5	15.5	4.0	12	Aktif	SD
6	16.0	4.0	15	Aktif	SD
7	17.9	1.5	16	Aktif	SMP
8	15.3	2.0	15	Tdk Aktif	SMP
9	17.9	2.5	18	Aktif	SMP
10	22.4	3.0	16	Aktif	SMA
11	16.0	2.0	15	Aktif	SMP
12	13.3	1.5	10	Tdk Aktif	SD
13	14.5	2.0	12	Tdk Aktif	SD
14	15.2	5.0	16	Tdk Aktif	SMP
15	15.7	3.0	12	Aktif	SD
16	17.9	2.5	12	Aktif	SMA
17	14.8	2.0	10	Tdk Aktif	SD
18	19.5	3.0	18	Aktif	SMA

### Keaktifan Petani

(Kualitatif, 2 kategori) →  
1 variabel dummy:

D1 → D1 = 1 aktif

D1 = 0 tdk aktif

### Pendidikan Petani

(Kualitatif, 3 kategori)  
→ 2 variabel dummy

D2 → D2 = 1 SMA

D2 = 0 lainnya

D3 → D3 = 1 SMP

D3 = 0 lainnya



No.	Produktivitas (ton/ha/thn)	Luas kebun (ha)	Umur Tan. (thn)	Keaktifan Petani (D1)	Pendidikan D2	Pendidikan D3
1	23.8	3.0	16	1	1	0
2	18.9	2.5	15	1	1	0
3	14.7	2.0	12	0	0	1
4	18.4	2.0	18	1	0	0
5	15.5	4.0	12	1	0	0
6	16.0	4.0	15	1	0	0
7	17.9	1.5	16	1	0	1
8	15.3	2.0	15	0	0	1
9	17.9	2.5	18	1	0	1
10	22.4	3.0	16	1	1	0
11	16.0	2.0	15	1	0	1
12	13.3	1.5	10	0	0	0
13	14.5	2.0	12	0	0	0
14	15.2	5.0	16	0	0	1
15	15.7	3.0	12	1	0	0
16	17.9	2.5	12	1	1	0
17	14.8	2.0	10	0	0	0
18	19.5	3.0	18	1	1	0

Petani yang nilai nya 0 pada D2 dan D3 berarti Pendidikan nya adalah SD

