

## **PENGUJIAN HIPOTESIS ASOSIATIF**

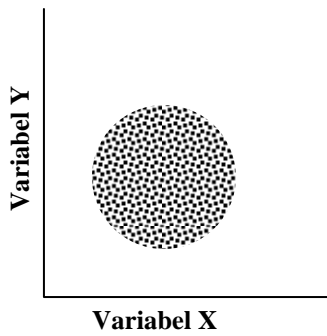
Hipotesis asosiatif merupakan dugaan adanya hubungan antar variabel dalam populasi, melalui data hubunga dalam sampel. Untuk itu, dalam langkah awal pembuktiannya, perlu dihitung terlebih dulukoefisien korelasi antar variabel dlam sampel, kemudian koefisien yang ditemukan tersebut diuji signifikansinya. Jadi menguji hipotesis asosiatif adalah menguji koefisien korelasi yang ada pada sampel untuk diberlakukan pada seluruh populasi tempat sampel diambil.

Terdapat tiga macam hubungan antar variabel, yaitu hubungan simetris, hubungan sebab akibat (kausal), dan hubungan interaktif (saling mempengaruhi). Untuk mencari hubungan antara dua variabel atau lebih dilakukan dengan menghitung koefisien korelasi antara variabel-variabel tersebut. Koefisien korelasi merupakan angka yang menunjukkan arah dan kuatnya hubungan antar variabel. Arah hubungan dinyatakan dengan tanda positif atau negatif, sedangkan kuatnya hubungan ditunjukkan dengan besarnya angka koefisien korelasi yang besarnya berkisar antara 0 sampai dengan  $\pm 1$ .

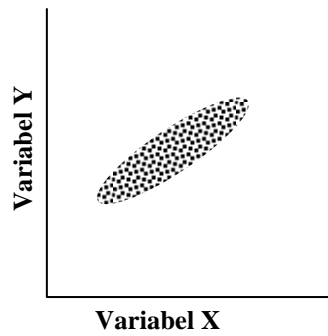
Hubungan positif antara dua variabel memberikan arti bahwa naiknya salah satu variabel akan menyebabkan naiknya variabel yang satunya. Sedangkan hubungan yang negatif mengandung arti bahwa ketika salah satu variabel nilainya naik maka variabel yang lain turun. Sebagai hubungan yang positif antara besarnya pendapatan dengan besarnya belanja bulanan, mengandung arti bahwa ketika pendapatan naik, maka belanja bulanan juga semakin naik. Sedangkan hubungan negatif terjadi misalnya dalam hubungan antara faktor usia dengan daya ingat, yang berarti bahwa semakin bertambah usia seseorang maka daya ingat akan semakin menurun. Demikian juga sebaliknya.

Angka koefisien korelasi yang berkisar antara 0 sampai dengan  $\pm 1$  menunjukkan kuat/lemahnya hubungan kedua variabel tersebut. Koefisien korelasi +1 menunjukkan bahwa antara kedua variabel tersebut terdapat hubungan positif sempurna. Sempurna disini mengandung arti bahwa naik atau turunnya salah satu variabel bisa dijelaskan dengan variabel yang lain dengan sepenuhnya tanpa kesalahan sedikit pun. Sedangkan koefisien korelasi sebesar nol, berarti bahwa antara kedua variabel tersebut sama sekali tidak terdapat hubungan. Artinya, naik atau turunnya variabel yang satu sama sekali tidak mempengaruhi variabel yang lain. Namun, dalam kehidupan sosial, korelasi sebesar nol dan satu ini jarang sekali terjadi (tidak akan pernah ada).

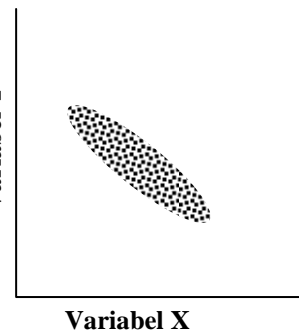
Dalam analisis statistik, besarnya koefisien korelasi bisa digambarkan dengan penyebaran titik data dalam kurva X-Y. Gambar-gambar yang menunjukkan koefisien korelasi adalah sebagai berikut:



Gambar 1



Gambar 2



Gambar 3

Gambar 1 menunjukkan persebaran hubungan antara variabel X dan variabel Y yang tidak menunjukkan pola tertentu. Artinya, pada saat variabel X rendah, variabel Y bisa rendah maupun tinggi. Demikian juga pada saat variabel X tinggi. Pola seperti ini menunjukkan tidak terdapat hubungan antara kedua variabel tersebut.

Gambar dua menunjukkan ketika variabel X rendah maka variabel Y juga rendah. Pada saat variabel X tinggi maka variabel Y juga tinggi. Hubungan seperti ini menunjukkan bahwa antara kedua variabel tersebut terdapat hubungan *positif* yang cukup kuat.

Gambar tiga menunjukkan ketika variabel X rendah maka variabel Y tinggi, dan pada saat variabel X tinggi maka variabel Y rendah. Hubungan seperti ini menunjukkan bahwa antara kedua variabel tersebut terdapat hubungan *negatif* yang cukup kuat.

Terdapat bermacam-macam teknik statistik korelasi yang dapat digunakan untuk menguji hipotesis asosiatif. Teknik koefisien yang mana yang akan dipakai tergantung pada jenis data yang dianalisis. Berikut adalah berbagai teknik statistik korelasi yang digunakan untuk menguji hipotesis asosiatif. Uji korelasi untuk data interval dan rasio menggunakan statistik parametris, sedangkan uji korelasi untuk data nominal dan ordinal menggunakan statistik nonparametris.

Teknik Uji Hipotesis Asosiatif untuk berbagai skala data

Skala data	Teknik Uji Statistik
Interval/Ratio	Pearson Product moment Korelasi Ganda Korelasi Parsial
Ordinal	Korelasi Rank Spearman Kendall Tau
Nominal	Koefisien Kontingency

## Korelasi Product Moment

Teknik korelasi ini digunakan untuk mencari hubungan dan membuktikan hipotesis hubungan antara dua variabel berbentuk interval dan rasio, yang berasal dari satu populasi. Rumus paling sederhana untuk menghitung korelasi adalah sebagai berikut:

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2 y^2)}}$$

Dimana:

$r_{xy}$  = Korelasiantara variabel x dengan variabel y

$x$  =  $(X_i - \text{rata-rata } X)$

$y$  =  $(Y_i - \text{rata-rata } Y)$

### Contoh kasus:

Ingin diketahui, benarkah terdapat hubungan antara pendapatan dengan pengeluaran. Untuk tujuan tersebut, data dikumpulkan terhadap 10 responden secara random. Data yang didapat adalah sebagai berikut:

X	800	900	700	600	700	800	900	600	500	500
Y	300	300	200	200	200	200	300	100	100	100

$H_0$ : Tidak terdapat hubungan antara pendapatan dengan pengeluaran

$H_a$ : terdapat hubungan antara pendapatan dengan pengeluaran

Dari perhitungan koefisien korelasi dengan rumus korelasi pada bagian atas, didapat:

Rata-rata  $x = 7$ ;

Rata-rata  $y = 2$ ;

$\sum X^2 = 20$ ;

$\sum Y^2 = 60$ ;

$\sum X^Y = 10$ ;

didapat  $r_{xy} = 0.9129$

Jadi, ada korelasi positif sebesar 0.9129 antara pendapatan dengan pengeluaran. Hal ini berarti bahwa semakin besar pendapatan maka pengeluaran juga akan semakin besar. Permasalahan adalah apakah angka korelasi tersebut signifikan (dapat digeneralisasikan) untuk dikatakan bahwa juga terdapat korelasi POPULASI. Untuk itu maka perlu diperbandingkan  $r$  hitung dengan  $r$  tabel ( $r$  tabel product moment) pada taraf signifikansi tertentu. Dengan melihat angka tabel  $r$  product moment, untuk tingkat signifikansi 5%, dengan  $N = 10$ , didapat  $r$  tabel = 0.632. Berartir hitung ( $=0.9129$ ) lebih besar dari  $r$  tabel, yang berarti kita harus menolak  $H_0$  dan menerima  $H_a$ . Dapat disimpulkan bahwa terdapat hubungan yang kuat dan signifikan antara pendapatan dengan pengeluaran.

Pengujian signifikansi korelasi, selain dilakukan dengan membandingkan koefisien korelasi (angka  $r$  hitung) dengan  $r$  tabel koefisien product moment, juga bisa dilakukan dengan membandingkan  $t$  hitung dengan  $t$  tabel.

Dengan cara ini, t hitung bisa dicari dengan rumus:

$$t = \frac{r \sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

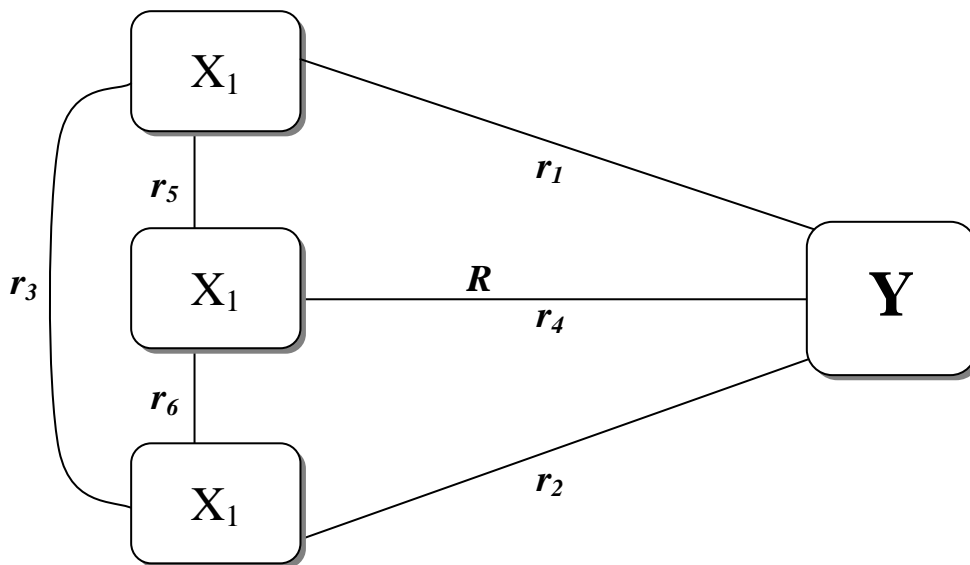
Untuk kasus, di atas, didapat t hitung = 6.33

Dari tabel t, dengan taraf signifikansi 5%, uji dua pihak, dengan dk = n-2 = 8, diperoleh harga t tabel = 2.306.

Karena t hitung lebih besar daripada t tabel, maka Ho ditolak dan Ha diterima, berarti terdapat hubungan yang kuat dan signifikan antara pendapatan dengan pengeluaran.

### **MULTIPLE CORRELATION (KORELASI GANDA)**

Korelasi ganda merupakan angka yang menunjukkan arah dan kuatnya hubungan antara dua variabel secara bersama-sama atau lebih dengan variabel lain. Pemahaman tentang korelasi ganda dapat dilihat melalui gambar berikut:



- $X_1$  = Kesejahteraan pegawai
- $X_1$  = Model kepemimpinan
- $X_1$  = Pengawasan
- $Y$  = Efektifitas Kerja

Dari gambar di atas, terlihat bahwa korelasi ganda R bukan merupakan penjumlahan dari korelasi sederhana yang ada pada setiap variabel (bukan  $r_1 + r_2 + r_3$ ). Namaun korelasi ganda  $R$  merupakan hubungan secara bersama-sama antara  $X_1, X_2, X_3$  dengan  $Y$ .

Rumus korelasi ganda ntuk dua variabel adalah sebagai berikut:

$$r_{y.x_1.x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$$

Dimana:

$r_{y.x_1.x_2}$  = Korelasi antara variabel X1 dan x2 secara bersama-sama dengan variabel Y

$r_{yx_1}$  = Korelasi product moment antara X<sub>1</sub> dengan Y

$r_{yx_2}$  = Korelasi product moment antara X<sub>2</sub> dengan Y

$r_{x_2x_1}$  = Korelasi product moment antara X<sub>1</sub> dengan X<sub>2</sub>

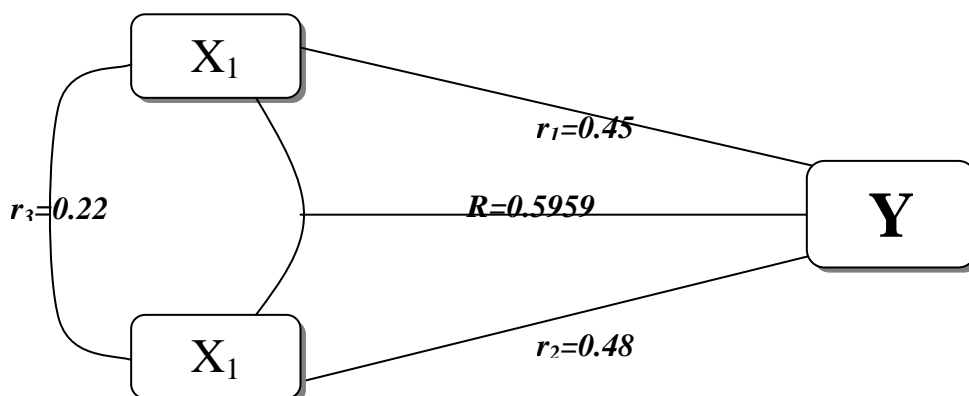
Contoh Kasusnya:

Untuk meneliti permasalahan model kepemimpinan dan tata ruang kantor dalam hubungannya dengan kepuasan kerja, dikumpulkan data yang relevan. Dari data tersebut kemudian dihitung korelasi sederhananya, dan didapat:

1. Korelasi antara model kepemimpinan dengan kepuasan kerja = 0.45
2. Korelasi antara tata ruang kantor dengan kepuasan kerja = 0.48
3. Korelasi antara model kepemimpinan dengan tata ruang kantor = 0.22

Dengan rumus korelasi ganda didapat:  $r_{y.x_1.x_2} = 0.5959$

Hasil perhitungan korelasi ganda dan sederhana tersebut dapat digambarkan sebagai berikut:



Dari perhitungan tersebut, ternyata besarnya korelasi ganda R harganya lebih besar dari korelasi individual  $r_{yx_1}$  dan  $r_{yx_2}$ .

Pengujian signifikansi terhadap koefisien korelasi ganda dapat menggunakan rumus uji F berikut:

$$F_h = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)}$$

Dimana:

R = Koefisien Korelasi ganda

K = Jumlah variabel independen

N = jumlah anggota sampel

Berdasarkan angka yang telah ditemukan, bila  $n = 30$ , maka harga  $F_h$  dapat dihitung dengan rumus di atas, dan didapat  $F_h = 7.43$

Harga F hitung tersebut selanjutnya dibandingkan dengan F tabel dengan dk pembilang =  $k = 2$ ; dan dk penyebut =  $n - k - 1 = 10 - 2 - 1 = 7$ . Dengan taraf signifikansi 5%, harga F tabel ditemukan sebesar 4.74.

Dengan membandingkan angka F tabel dengan F hitung, ternyata F hitung lebih besar daripada F tabel, berarti  $H_0$  ditolak dan  $H_a$  diterima. Jadi, koefisien korelasi ganda yang ditemukan adalah signifikan (dapat diberlakukan untuk populasi dimana sampel di ambil).

## KORELASI PARSIAL

Korelasi ini digunakan untuk menganalisis nilai peneliti bermaksud mengetahui pengaruh atau hubungan antara variabel independen dengan variabel dependen, di mana salah satu variabel independen yang lain dibuat tetap/dikendalikan. Jadi, korelasi parsial merupakan angka yang menunjukkan arah dan kuatnya hubungan antara dua variabel atau lebih setelah satu variabel yang diduga dapat mempengaruhi hubungan tersebut dikendalikan untuk dibuat tetap keberadaannya.

Contoh;

1. korelasi antara nilai tes IQ dengan prestasi sekolah = 0.58
2. Korelasi antara waktu belajar dengan prestasi sekolah = 0.10
3. Korelasi antara nilai tes IQ dengan waktu belajar = -0.40

Apabila ada pertanyaan, untuk siswa yang waktu belajarnya sama (diparsialkan) berapa korelasi antara IQ dengan prestasi kuliah? Jawaban pertanyaan tersebut bisa dicari dengan rumus korelasi parsial berikut:

$$R_{y_{x_1x_2}} = \frac{R_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{x_1x_2}^2) - \sqrt{(1 - r_{yx_2}^2)}}$$

Rumus di atas dapat dibaca: Korelasi antara Y dengan  $X_1$ , bila variabel  $X_2$  dikendalikan atau korelasi antara Y dengan  $X_1$ , bila  $X_2$  tetap.

Sedangkan bila  $X_1$  yang dikendalikan, maka rumusnya adalah:

$$R_{yx2x1} = \frac{R_{yx2} - r_{yx1} \cdot r_{x1x2}}{\sqrt{(1 - r_{x1x2}^2)} - \sqrt{(1 - r_{yx1}^2)}}$$

Sedangkan uji korelasi parsial dapat dihitung dengan rumus sebagai berikut:

$$t = \frac{r_p \sqrt{(n - 3)}}{\sqrt{(1 - r^2_p)}}$$

## KORELASI KONTINGENSI

Jenis korelasi ini digunakan untuk menghitung hubungan antara variabel bila datanya berbentuk nominal. Teknik ini mempunyai kaitan yang erat dengan Chi Kuadrat yang digunakan untuk menguji hipotesis komparatif k sampel independen. Oleh karena itu rumus yang digunakan mengandung nilai Chi Kuadrat. Rumustersebut adalah sebagai berikut:

$$t = \sqrt{\frac{X^2}{N + X^2}}$$

Harga Chi Kuadrat ( $X^2$ ) dicari dengan rumus:

$$X^2 = \sum \sum \frac{(Op_{ij} - E_{ij})^2}{EP_{ij}}$$

### CONTOH-e

Dilakukan penelitian untuk mencari tahu apakah terdapat hubunga antara profesi pekerjaan dengan jenis olah raga yang sering dilakukan. Profesi dikelompokkan menjadi : Dokter, Pengacara, Dosen, Bisnisman. (Dr, Pc, Ds, Bs). Sedangkan jenis olahraga dikelompokkan menjadi: Golf, Tennis, Bulutangkis, Sepak bola (Gf, T, Bt, Sp). Jumlah responden yang digunakan untuk mengumpulkan data adalah sebagai berikut:

Dr = 58

Pc = 75

Ds = 68

Bs = 81

Jumlah seluruhnya = 282

Perumusan hipotesisnya adalah:

Ho: Tidak ada hubungan yang signifikan antara profesi seseorang dengan jenis olahraga yang disenanginya

Ha: Ada hubungan yang kuat dan signifikan antara profesi seseorang dengan jenis olahraga yang disenanginya

Berdasarkan sampel empat kelompok profesiyang dipilih secara random, diperoleh data seperti data berikut:

Jenis Olah Raga	Profesi				Jumlah
	Dr	Pc	Ds	Bs	
Gf	17	23	10	30	<b>80</b>
Tn	23	14	17	26	<b>80</b>
Bt	12	26	18	14	<b>70</b>
Sp	6	12	23	11	<b>52</b>
Jumlah	<b>58</b>	<b>75</b>	<b>68</b>	<b>81</b>	<b>282</b>

Untuk mmenghitung f yang diharapkan (fh) pertama-tama dihitung berapa persen dari masing-masing sampel yang menyenangi olah raga Golf, Tenis, Bulu tangkis, dan sepakbola.

Dari sini, dapat dihitung prosentase:

Prosentasi yang menyenangi Golf =  $80/282 = 0.284$

Prosentasi yang menyenangi Tenis =  $80/282 = 0.284$

Prosentasi yang menyenangi Bulu tangkis =  $70/282 = 0.248$

Prosentasi yang menyenangi Sepak bola =  $52/282 = 0.184$

Selanjutnya, masing-masing fh (frekuensi yang diharapkan) kelompok yang menyenangi setiap jenis olah raga dapat dihitung

1. Yang menyenangi Golf:
  - a. Fh Dokter :  $0.284 \times 58 = 16.472$
  - b. Fh Pengacara :  $0.284 \times 75 = 21.300$
  - c. Fh Dosen :  $0.284 \times 68 = 19.312$
  - d. Fh Bisnisman :  $0.284 \times 81 = 23.004$
  
2. Yang menyenangi Tenis:
  - a. Fh Dokter :  $0.284 \times 58 = 16.472$
  - b. Fh Pengacara :  $0.284 \times 75 = 21.300$
  - c. Fh Dosen :  $0.284 \times 68 = 19.312$
  - d. Fh Bisnisman :  $0.284 \times 81 = 23.004$
  
3. Yang menyenangi Bulu tangkis:
  - a. Fh Dokter :  $0.248 \times 58 = 14.384$
  - b. Fh Pengacara :  $0.248 \times 75 = 18.600$
  - c. Fh Dosen :  $0.248 \times 68 = 16.864$
  - d. Fh Bisnisman :  $0.248 \times 81 = 20.088$



4. Yang menyenangkan Sepak Bola:

- a. Fh Dokter : 0.184 x 58 = 10.672
- b. Fh Pengacara : 0.184 x 75 = 13.800
- c. Fh Dosen : 0.184 x 68 = 12.512
- d. Fh Bisnisan : 0.184 x 81 = 14.902

Berdasarkan hasil perhitungan tersebut, selanjutnya dimasukkan ke dalam tabel sebagai berikut:

Olah Raga	Dr		Pc		Ds		Bs		JML.
	Fo	Fh	Fo	Fh	Fo	Fh	Fo	Fh	
Gf	17	16.472	23	21.300	10	19.312	30	23.004	80
Tn	23	16.472	14	21.300	17	19.312	26	23.004	80
Bt	12	14.384	26	18.600	18	16.864	14	20.088	70
Sp	6	10.672	12	13.800	23	12.512	11	14.904	52
JML	58		75		68		81		282

Selanjutnya, harga Chi Kuadrat (X<sup>2</sup>) dapat dihitung dengan rumus:

$$X^2 = \sum \sum \frac{(Op_{ij} + E_{ij})^2}{EP_{ij}}$$

Dalam hal ini, O (observation) = fo, dan E (Expectation) = fh

$$\begin{aligned}
 X^2 = & \frac{(17 - 16.472)^2}{16.472} + \frac{(23 - 21.300)^2}{21.300} + \frac{(10 - 19.312)^2}{19.312} + \\
 & \frac{(30 - 23.004)^2}{23.004} + \frac{(23 - 16.472)^2}{16.472} + \frac{(14 - 21300)^2}{21.300} + \\
 & \frac{(17 - 19.312)^2}{19.312} + \frac{(26 - 23.004)^2}{23.004} + \frac{(12 - 14.384)^2}{14.384} + \\
 & \frac{(26 - 18.600)^2}{18.600} + \frac{(18 - 16.864)^2}{16.864} + \frac{(14 - 20.088)^2}{20.088} + \\
 & \frac{(6 - 10.672)^2}{10.672} + \frac{(12 - 13.800)^2}{13.800} + \frac{(23 - 12.512)^2}{12.512} + \\
 & \frac{(11 - 14.904)^2}{14.904}
 \end{aligned}$$

$$X^2 = 0.007 + 0.136 + 4.490 + 2.218 + 2.587 + 2.502 + 0.277 + 0.390 + \\ 0.395 + 2.944 + 0.077 + 1.845 + 2.045 + 0.235 + 8.791 + 1.023$$

$$X^2 = \mathbf{29.881}$$

Jadi, harga Chi Kuadrat ( $X^2$ ) hitung = 29.881. Selanjutnya untuk menghitung koefisien kontingensi C, maka harga tersebut dimasukkan ke dalam rumus

$$C = \sqrt{\frac{X^2}{N + X^2}}$$

$$C = \sqrt{\frac{29.881}{282 + 29.881}}$$

$$C = 0.31$$

Jadi, besarnya koefisien antara jenis profesi dengan kesenangan olah raga = 0.31. Untuk menguji signifikansi koefisien C dapat dilakukan dengan menguji harga Chi Kuadrat ( $X^2$ ) hitung yang ditemukan dengan Chi Kuadrat ( $X^2$ ) tabel, pada taraf signifikansi dan dk tertentu. Harga  $dk = (k - 1) (r - 1)$ ; dimana  $K =$  jumlah sampel = 4;  $r =$  jumlah kategori olah raga. Jadi  $dk = (4 - 1) (4 - 1) = 9$ . Dengan  $dk = 9$  dan pada taraf signifikansi 0.05, maka harga Chi Kuadrat ( $X^2$ ) tabel = 15.51. Ketentuan pengujian kalau harga Chi Kuadrat ( $X^2$ ) hitung lebih besar dari Chi Kuadrat ( $X^2$ ) tabel, maka hubungannya signifikan. Pada kasus kita di atas, ternyata harga Chi Kuadrat ( $X^2$ ) hitung lebih besar daripada Chi Kuadrat ( $X^2$ ) tabel. ( $29.881 > 15.51$ ). Dengan demikian dapat disimpulkan bahwa  $H_0$  ditolak dan  $H_a$  diterima. Jadi, jenis profesi pekerjaan secara nyata mempunyai hubungan signifikan dengan jenis olahraga yang disenangi sebesar 0.31. Data yang ada pada sampel dan angka korelasi mencerminkan keadaan populasi di mana sampel diambil.

## ANALISIS REGRESI

Korelasi dan regresi keduanya mempunyai hubungan yang sangat erat. Setiap regresi pasti ada korelasinya, walaupun tidak setiap korelasi selalu dicati regressinya. Korelasi yang tidak dilanjutka dengan regresi adalah korelasi antara dua variabel yang tidak mempunyai hubungan kausal (sebab-akibat). Analisis regresi ini dilakukan bila hubungan dua variabel berupa hubungan kausal atau fungsional. Ntuk menetapkan kedua variabel mempunyai hubungan kausal ataukah tidak harus didasarkan pada teori atau konsep-konsep tentang dua variabel tersebut.

Hubungan antara model kepemimpinan dengan kepuasan kerja adalah hubungan kausal, karena secara teori modele kepemimpinan akan menyebabkan puas atau tidaknya pekerja dalam wilayah kepemimpinannya tersebut. Namun, hubungan antara kicau burung emprit dengan tamu yang datan bukan merupakan hubungan kasual, walaupun (konon katanya) kondisi seperti ini ada hubungannya.

Kita menggunakan analisis regresi apabila kita ingin mengetahui bagaimana variabel dependen (variabel kriteria) dapat diprediksikan dengan variabel independen (variabel prediktor). Hasil dari analisis regresi adalah suatualat proyeksi (prediksi) dan perencanaan, bagaimana naik atau turunnya suatu variabel bisa dikontrol/dikendalikan dengan menaikkan atau menurunkan variabel yang lain. Sebagai contohnya, apabila diketahui bahwa jumlah barang terjual mempunyai hubungan dan dipengaruhi oleh jumlah biaya iklan yang dibelanjakan, maka seoran pemasar bisa merencanakan berapa biaya iklan yang harus dibelanjakan apabila ingin menaikkan jumlah penualan menjadi dua kali lipatnya. Demikian dan seterusnya.

### REGRESI LINIER SEDERHANA

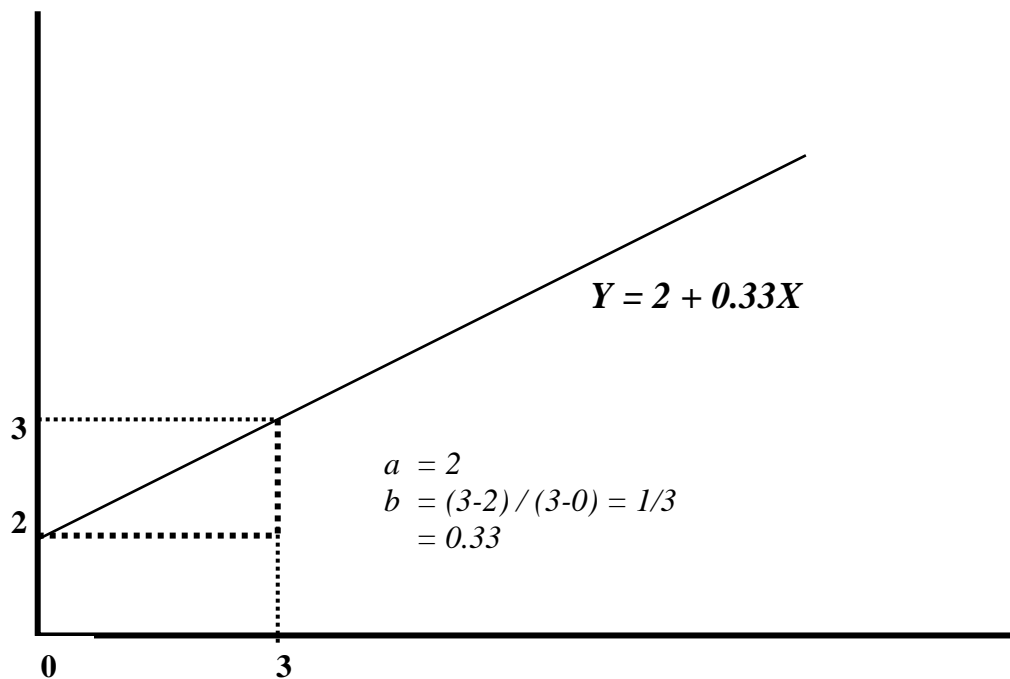
Regresi sederhana didasarkan pada hubungan fungsional ataupun kausal satu variabel independen dengan satu variabel independen. Persamaan umum regresi linier sederhana adalah;

$$Y' = a + bX$$

Dimana :

- Y' : Subyek dalam variabel dependen yang diprediksikan
- a : Harga Y bila X = 0 (harga konstan)
- b : Angka arah atau koefisien regresi, yang menunjukkan angka peningkatan ataupun penurunan variabel dependen yang didasarkan pada variabel independen.
- X : Subyek pada variabel independen yang mempunyai nilai tertentu.

Secara teknis harga b merupakan tangen dari (perbandingan) antara panjang garis variabel dependen setelah persamaan regresi ditemukan.



rumus untuk mendapatkan harga a dan b adalah sebagai berikut:

$$b = r \frac{S_y}{S_x}$$

$$a = Y - bX$$

dimana:

- r = koefisien korelasi product moment antara variabel X dengan variabel Y
- $S_y$  = simpangan baku variabel Y
- $S_x$  = simpangan baku variabel X

Jadi harga b merupakan fungsi dari koefisien korelasi. Bila koefisien korelasi tinggi, maka harga b juga besar, sebaliknya bila koefisien korelasi rendah maka harga b juga rendah (kecil). Selain itu bila koefisien korelasi negatif maka harga juga negatif.

Selain dengan menggunakan rumus di atas, harga a dan b juga bisa dicari dengan rumus berikut;

$$a = \frac{(\sum Y_i) (\sum X_i^2) - (\sum X_i) (\sum X_i Y_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

$$b = \frac{n \sum X_i Y_i - (\sum X_i) (\sum Y_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

### Contoh Perhitungan regresi linier sederhana

Data berikut adalah hasil pengamatan terhadap nilai kualitas layanan (X) dan nilai rata-rata penjualan (Y) balonpas:

Nomor	X	Y
1	54	167
2	50	155
3	53	148
4	45	146
5	48	170
6	63	173
7	46	149
8	56	166
9	52	170
10	56	174
11	47	156
12	56	158
13	55	150
14	52	160
15	50	157
16	60	177
17	55	166

Nomor	X	Y
18	45	160
19	47	155
20	53	159
21	49	159
22	56	172
23	57	168
24	50	159
25	49	150
26	58	165
27	48	159
28	52	162
29	56	168
30	54	166
31	59	177
32	47	149
33	48	155
34	56	160

Untuk menghitung persamaan regresi, maka kita membutuhkan alat bantu tabel, sebagai berikut:

Nomor	X	Y	XY	X <sup>2</sup>	Y <sup>2</sup>
1	54	167	9.018	2.916	27.889
2	50	155	7.750	2.500	24.025
3	53	148	7.844	2.809	21.904
4	45	146	6.570	2.025	21.316
5	48	170	8.160	2.304	28.900
6	63	173	10.899	3.969	29.929
7	46	149	6.854	2.116	22.201
8	56	166	9.296	3.136	27.556
9	52	170	8.840	2.704	28.900
10	56	174	9.744	3.136	30.276
11	47	156	7.332	2.209	24.336
12	56	158	8.848	3.136	24.964
13	55	150	8.250	3.025	22.500
14	52	160	8.320	2.704	25.600
15	50	157	7.850	2.500	24.649
16	60	177	10.620	3.600	31.329
17	55	166	9.130	3.025	27.556
18	45	160	7.200	2.025	25.600
19	47	155	7.285	2.209	24.025
20	53	159	8.427	2.809	25.281
21	49	159	7.791	2.401	25.281
22	56	172	9.632	3.136	29.584
23	57	168	9.576	3.249	28.224
24	50	159	7.950	2.500	25.281
25	49	150	7.350	2.401	22.500
26	58	165	9.570	3.364	27.225
27	48	159	7.632	2.304	25.281
28	52	162	8.424	2.704	26.244
29	56	168	9.408	3.136	28.224
30	54	166	8.964	2.916	27.556
31	59	177	10.443	3.481	31.329
32	47	149	7.003	2.209	22.201
33	48	155	7.440	2.304	24.025
34	56	160	8.960	3.136	25.600
<b>Σ</b>	<b>1.782</b>	<b>5.485</b>	<b>288.380</b>	<b>94.098</b>	<b>887.291</b>

Kita gunakan rumus untuk menghitung harga a dan b, sehingga kita dapatkan:

$$a = \frac{(5.485)(94.098) - (1.782)(288.380)}{(34).(94.098) - (1.782)^2} = 93.85$$

$$b = \frac{(34).(288.380) - (1.782).(5.485)}{(34).(94.098) - (1.782)^2} = 1.29$$

Dengan ditemukannya harga a dan b maka persamaan regresi bisa ditentukan, yaitu:

$$Y = 93.85 + 1.29 X$$

Persamaan regresi ini, selanjutnya bisa digunakan untuk melakukan proyeksi bagaimana individu-individu dalam variabel dependen akan terjadi bisa individu dalam variabel independen ditetapkan. Misalnya, nilai kualitas layanan ditetapkan = 64, maka nilai rata-rata penjualan bisa diprediksi sebesar:

$$Y = 93.85 + (1.29 \times 64) = 176.41$$

Dari persamaan regresi di atas, bisa diartikan bahwa apabila nilai kualitas pelayanan bertambah dengan satu satuan, maka nilai rata-rata penjualannya tiap bulan akan bertambah dengan 1.29 satuan. Bisa juga dikatakan bahwa setiap nilai kualitas layanan bertambah dengan 10 maka nilai rata-rata penjualan akan bertambah sebesar 12.9.

*Demikian, WaLLahu a'lam.*

## ANALISIS KORELASI DAN REGRESI DENGAN SPSS

Salah satu masalah umum dalam staitistika adalah ketika kita mengamati sifat-sifat suatu sampel muncul pertanyaan apakah perbedaan-perbedaan yang teramati menandakan bahwa populasi dari maa smpelitu diambil memang benar-benar berbeda dari anggapan yang kita buat atau berbeda antara dua sampel yang kita ambil.

Jika kita mengumpulkan dua kelompok angka skor secara acak, besar kemungkinan bahwa dia kelpok itu (sedikit) berbeda sastu sama lain. Pertanyaan yang muncul adalah benarkah perbedaan itu memang terjadi karena perbedaan yang terdapat dalam populais dari mana kedua sampel itu kita ambil, ataukah perbedaan itu semata terjadi karena kebetulan saja.

Prosedur-prosedur inferensi statistik memungkinkan kita menetapkan, dalam bentuk probabilitas, apakah perbedaan-perbedaan yang teramati itu termasuk hal yang mudah terjadi karena kebetulan, ataukah perbedaan-perbedaan yang terjadi sebegitu besarnya yang memang berasal dari populasi yang berbeda.

Dalam perkembangan metode-metode statistik modern, teknik-teknik inferensi yang pertama muncul adalah teknik-teknik yang membuat sejumlah besar asumsi mengenai sifat populasi dari mana skor-skor itu diambil. Karena nilai-nilai populasi adalah “parameter”, maka teknik-teknik statistik ini disebut “*parametrik*”. Teknik-teknik semacam ini menghasilkan kesimpulan yang disertai dengan persyaratan, m misalnya: “*Jika asumsi-asumsi mengenai bentuk populasinya sah, maka dapat disimpulkan bahwa.....*”

Dalam perkembangannya, munul teknik-teknik inferensi yang tidak membuat asumsi yang ketat mengenai parameter-parameter. Teknik-teknik “distribusi bebas” atau

*nonparametrik* ini menghasilkan kesimpulan-kesimpulan yang memerlukan kualifikasi yang lebih sedikit jumlahnya. Dengan menggunakan salah satu teknik nonparametrik ini, kita bisa menyatakan bahwa “*tanpa memperhatikan bentuk populasi (atau populasi-populasinya), kita dapat menyimpulkan bahwa, ..... ”*.

Dalam perhitungan pada test-test parametrik kita (dapat) menjumlahkan, membagi, dan melipatkan skor-skor sampel. Kalau proses aritmatik ini kita gunakan untuk skor-skor yang sebenarnya tidak berwujud angka (berwujud ranking atau ordering), maka akan sangat mungkin menimbulkan ketidaktepatan data dan dengan demikian kesimpulannya pun tidak dapat dapat dipakai. Ini sebabnya mengapa tes-tes parametrik hanya bisa digunakan untuk skor-skor yang memang dapat diangkakan.

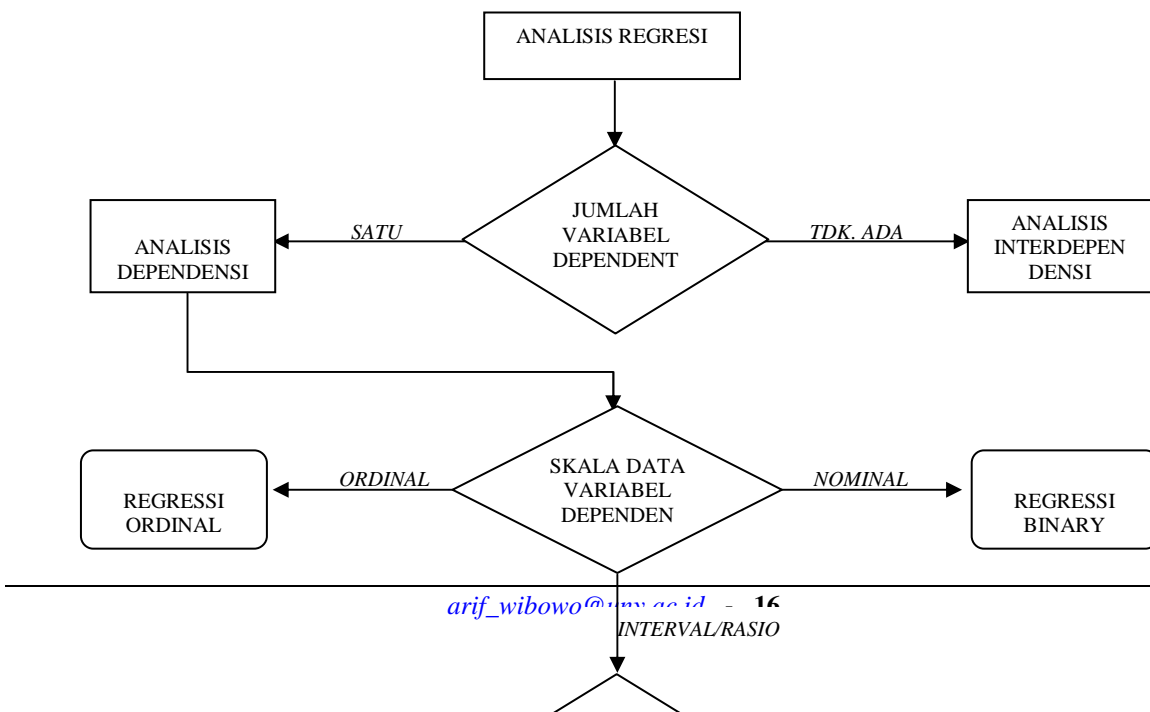
Sedangkan tes nonparametrik menitikberatkan pada urutan atau tingkatan skor-skor data, dan bukan pada nilai keangkaanya. Bahkan ttest nonparametrik lain bisa digunakan untuk data yang bahkan tidakmungkin dibuatkan urutannya (data bersifat klasifikasi).

### UJI KORELASI DAN REGRESI

Analisis Korelasi mempelajari *apakah terdapat hubungan* antara dua variabel atau lebih. Sedangkan analisis regresi memprediksi sejauh mana pengaruh tersebut ada.

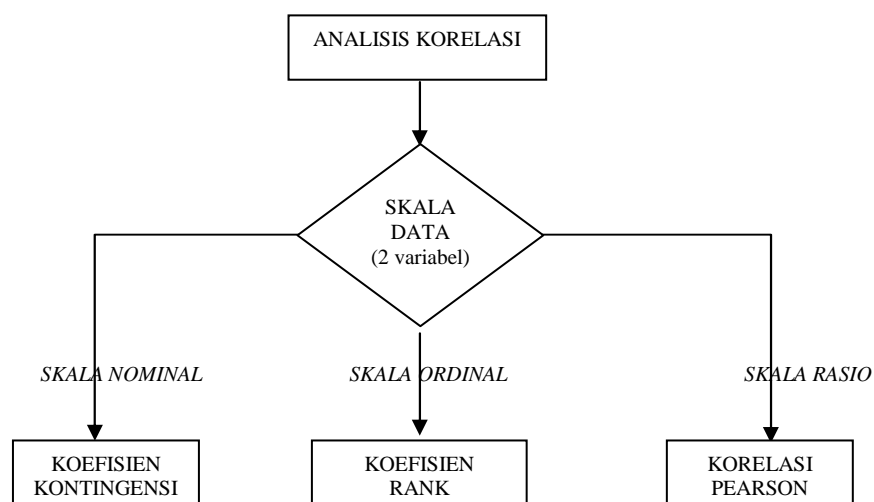
Berdasar skala datanya, analisis korelasi dan analisis regresi bisa dibedakan sebagai berikut:

Gambar 1.2.: Bagan analisis regresi berdasar jumlah variabel dan skala data





Gambar 1.1.: Bagan analisis korelasi berdasar skala data



Beberapa asumsi yang harus dipenuhi dalam analisis regresi:

1. **NORMALITAS:** nilai Y didistribusikan secara normal terhadap nilai X
2. **HOMOSKEDASTISITAS:** variasi di sekitar garis regresi konstan untuk setiap nilai X
3. **MULTIKOLINEARITAS:** Antar variabel independent tidak boleh saling berkorelasi
4. **AUTOKORELASI:** bebas dari outokorelasi, yaitu kesalahan prediksi bersifat random.

### *ANALISIS KORELASI*

Analisis ini digunakan untuk menguji apakah ada hubungan antara dua variabel atau lebih. Analisis ini juga dibedakan atas dua bagian besar tes statistik, yaitu tes parametrik dan tes nonparametrik.

Apabila kita mengkorelasikan data yang diukur dengan skala rasio, atau interval, kita bisa menggunakan tes korelasi parametrik. namun apabila data yang kita korelasikan datanya berskala ordinal atau nominal, kita akan menggunakan tes korelasi nonparametrik.

**KASUS 01:**

kita akan meneliti adakah hubungan antara tingkat gaji seorang pekerja, dengan usia dan lama pengamalan kerjanya.

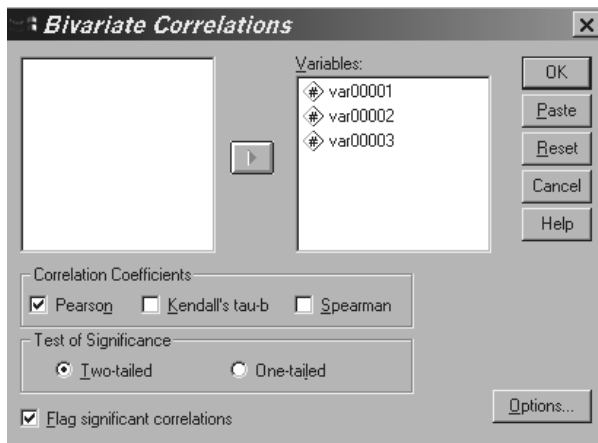
Untuk kepentingan ini dikumpulkan data dari sebuah perusahaan, dengan pengamatan pada variabel tingkat gaji, usia, dan lama kerja.

*Di dapat "DATA KARYAWAN", seperti yang ada di Lampiran Data Kasus 01:*

## PENGOLAHAN DENGAN SPSS

1. Masukkan data di atas, atau buka file (bila sudah ada).
2. Dari menu spss, klik: **ANALYZE → CORRELATE → BIVARIATE**

Akan muncul kotak dialog, isi kotak dialog menjadi sebagai berikut:



Klik OK dan di dapat output sebagai berikut:

**Correlations**

		Usia Pekerja	lama Bekerja	Gaji Karyawan
Usia Pekerja	Pearson Correlation	1	,438**	,682**
	Sig. (2-tailed)	.	,000	,000
	N	75	75	75
lama Bekerja	Pearson Correlation	,438**	1	,730**
	Sig. (2-tailed)	,000	.	,000
	N	75	75	75
Gaji Karyawan	Pearson Correlation	,682**	,730**	1
	Sig. (2-tailed)	,000	,000	.
	N	75	75	75

\*\* . Correlation is significant at the 0.01 level (2-tailed).

## ANALISIS OUTPUT SPSS:

### 1. Arti angka korelasi:

Dari tabel, angka korelasi dilihat pada baris “Pearson Correlation”

Korelasi antara Usia dengan Lama Bekerja	: 0.438
Korelasi antara Usia dengan Gaji karyawan	: 0.682
Korelasi antara Gaji Karyawan dengan Lama Bekerja	: 0.730

Nb: Korelasi positif menunjukkan arah hubungan (positif)

### 2. Signifikansi Hasil Korelasi

Ho : tidak ada hubungan antara variabel ..... dengan variabel .....

H1 : Ada hubungan (korelasi antara variabel .... dengan variabel .....

Dasar Pengambilan Keputusan

Jika  $\text{sig} > 0,05$  (atau 0.01) berarti Ho diterima

Jika  $\text{sig} \leq 0,05$  (atau 0.01) berarti H1 diterima (Ho ditolak)

Dari tabel, angka signifikansi dilihat pada baris Sig (2-tailed)

Signifikansi antara Usia dengan Lama Bekerja	: 0.0000
Signifikansi antara Usia dengan Gaji karyawan	: 0.0000
Signifikansi antara Gaji Karyawan dengan Lama Bekerja	: 0.0000

Karena semua angka sig dibawah .05, berarti dengan derajat kepercayaan 95% kesemua variabel memang secara nyata berkorelasi.

Keputusan ini juga bisa dilihat dengan adanya tanda (\*\*) pada angka korelasi, yang mengindikasikan bahwa angka korelasi secara nyata signifikan pada level 0.01 (apalagi pada 0.05)

### *Kata Kunci dalam membaca output SPSS: “SIG”*

Kata “sig” pada output SPSS berarti signifikansi, yang berarti **kemungkinan salah menolak Ho.**

3. Misalnya: terdapat angka sig = 0.02, berarti kemungkinan salah kita menolak Ho adalah 0.02 atau 2%. (*Kemungkinan salah menolak Ho maksudnya, ketika Ho sebenarnya adalah benar dan keputusan kita ternyata menolak Ho, berarti kita membuat keputusan yang salah, yaitu menolak sesuatu yang seharusnya kita terima (benar)... kemungkinan ini besarnya adalah 2%*)

SPSS selalu menggunakan tolok ukur signifikansi 5%, yang berarti resiko kesalahan mengambil keputusan dibatasi sampai dengan 5%, tidak boleh lebih.

Karena 2% < 5%, maka tolak saja Ho, karena resiko kesalahan kita menolak Ho hanya 2%, jauh di bawah 5%.

**Demikian. WaLLahu a’lam**

# ANALISIS REGRESI

Korelasi dan regresi keduanya mempunyai hubungan yang sangat erat. Setiap regresi pasti ada korelasinya. Tetapi korelasi belum tentu dilanjutkan dengan regresi. Korelasi yang tidak dilanjutkan dengan regresi adalah korelasi antara dua variabel yang tidak mempunyai hubungan kausal (sebab-akibat) atau hubungan fungsional.

Kita gunakan analisis regresi apabila kita ingin mengetahui bagaimana variabel dependen (=variabel kriteria) dapat diprediksikan melalui variabel independent (=variabel prediktor). Dampak dari penggunaan analisis regresi dapat digunakan untuk memutuskan apakah naik atau menurunnya nilai pada variabel dependen dapat dilakukan dengan menaikkan atau menurunkan keadaan variabel independent (variabel prediktor).

Beberapa Istilah:

1. Koefisien Korelasi =  $r$  = derajat keeratan hubungan antara dua variabel
2. Koefisien determinasi =  $r^2$  = yaitu suatu ukuran seberapa baik sebuah variabel independent menjelaskan variasi dalam variabel dependen
3. Garis regresi =  $Y = \beta_0 + \beta_1 X$

Fungsi regresi dibangun dengan seperangkat asumsi:

1. Normalitas, yaitu asumsi bahwa nilai-nilai Y untuk tiap X tertentu didistribusikan secara normal di sekitar meannya.
2. Homoskedastisitas, yaitu variasi nilai-nilai Y di sekitar nilai meannya tersebut adalah konstan untuk semua nilai X
3. Otokorelasi: Nilai rssidual pada suatu observasi tidak berkorelasi dengan data tersebut pada periode sebelumnya. Dengan kata lain, nilai eror untuk setiap nilai X harus bersifat acak (random).
4. Multikolinieritas: Antar variabel independent tidak boleh saling berkorelasi

## 1. Regresi linier Sederhana

Regresi ini didasarkan pada hubungan fungsional ataupun kausal satu variabel independent dengan satu variabel dependen.

Persamaan umum regresi sederhana adalah:

$$Y' = a + bX; \text{ dimana:}$$

- $Y'$  = Subyek dalam variabel dependent yang diprediksi  
 $a$  = harga Y bila  $X = 0$  (=harga konstan)

- $b$  = Angka arah (=koefisien regresi), yang menunjukkan angka peningkatan ataupun penurunan pada variabel dependen yang disebabkan oleh perubahan variabel independen.  
 $X$  = Subyek pada variabel independen yang mempunyai nilai tertentu.

## Memprediksi variabel dependen dengan menggunakan alat bantu program SPSS

### Pada dasarnya, tahapan penyusunan model regresi, meliputi:

1. Menentukan variabel bebas (independen) dan variabel terikat (dependen)
2. menentukan metode pembuatan model regresi (metode enter, stepwise, forward, backward)
3. Melihat ada tidaknya data outlier (ekstrim)
4. Menguji asumsi-asumsi pada regresi berganda (normalitas, ....)
5. menguji signifikansi model (uji t, Uji F, ...)
6. nterpretasi model

### Contoh kasus Regresi: Kasus 21

Manajer pemasaran Boneka mak Lampir ingin mengetahui kegiatan apakah yang menunjang penjualan perusahaan selama ini: apakah iklan di koran; iklan di radio; jumlah outlet penjualan; ataukah jumlah salesman.

Data yang dikumpulkan untuk analisis tersebut seperti pada contoh data kasus 21 berikut:

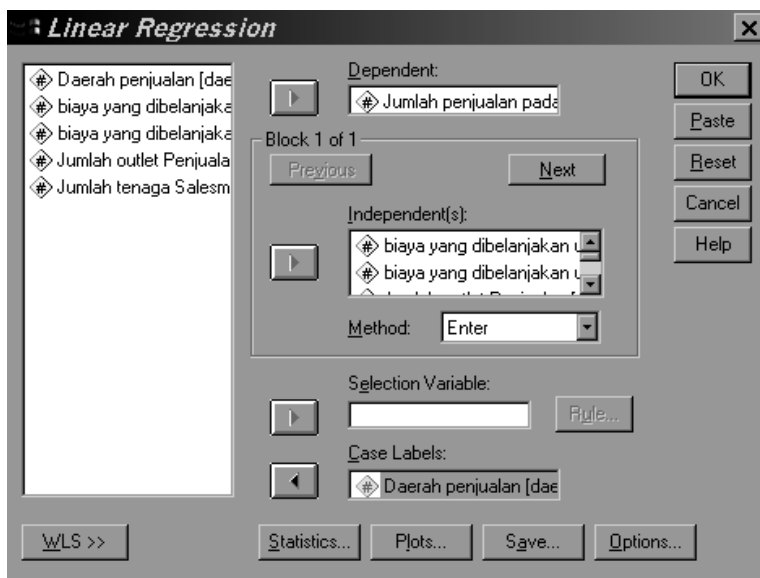
Data kasus 21:

No.	Penjualan	IklKor	IklRad	Outlet	Salesman
1	300,12	26,23	12,23	7	4
2	312,25	25,12	12,88	8	3
3	362,02	29,80	15,26	8	2
4	400,25	34,55	14,23	9	1
5	412,60	33,45	13,02	6	4
6	423,00	32,26	13,56	5	2
7	320,14	23,45	12,03	8	3
8	366,25	34,76	15,26	9	3
9	451,29	40,12	14,32	8	2
10	430,22	36,21	13,33	10	5
11	265,99	25,89	12,05	11	4
12	254,26	22,98	15,26	10	1
13	352,16	36,25	12,89	9	5
14	365,21	36,87	12,45	8	5
15	295,15	22,41	13,44	5	2
16	345,25	26,25	13,67	6	2
17	415,25	36,99	19,25	8	5
18	400,23	32,76	18,78	9	2
19	423,22	33,98	16,59	7	2
20	452,62	23,21	18,45	5	3

21	512,33	44,98	13,45	8	5
22	435,23	35,99	15,78	8	3
23	302,21	25,00	16,35	9	2
24	330,92	23,25	19,58	8	5
25	254,25	24,86	13,87	6	6
26	265,21	26,23	15,87	5	5
27	215,36	20,98	13,23	7	4
28	235,26	24,88	15,69	9	3
29	222,32	25,87	18,97	8	6
30	323,45	28,94	18,29	9	5

Langkah analisis regresi dengan SPSS

1. masukkan data (buka file data)
2. Klik **ANALYZE** → **REGRESSION** → **LINEAR** .....
3. Isi kotak dial;og menjadi sebagai berikut:



Didapat output sebagai berikut:

**Model Summary<sup>a</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,869 <sup>a</sup>	,756	,717	41,55040

a. Predictors: (Constant), Jumlah tenaga Salesman, Jumlah outlet Penjualan, biaya yang dibelanjakan untuk iklan di radio, biaya yang dibelanjakan untuk iklan di kurani

b. Dependent Variable: Jumlah penjualan pada derah itu

*Analisis:*

Korelasi  $\rightarrow R = 0.869 \rightarrow$  terdapat hubungan yang kuat antara var dependen dengan keempat var independe

Koefisien Determinasi  $\rightarrow R$  Square Adjusted = 0.717  $\rightarrow$  nilai var dependent, sebesar 71.7% bisa dijelaskan (memang dipengaruhi) oleh keempat variabel dependent

ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	133500,0	4	33374,997	19,332	,000 <sup>a</sup>
	Residual	43160,901	25	1726,436		
	Total	176660,9	29			

a. Predictors: (Constant), Jumlah tenaga Salesman, Jumlah outlet Penjualan, biaya yang dibelanjakan untuk iklan di radio, biaya yang dibelanjakan untuk iklan di koran

b. Dependent Variable: Jumlah penjualan pada daerah itu

Analisis:

Nilai Uji- F sebesar 19.332 dengan sig = 0.000, yang berarti signifikansi < 0.05. maka model regresi bisa dipakai untuk memprediksi penjualan. Atau bisa dikatakan: Iklan di koran, iklan di radio, jumlah outlet, dan jumlah salesman secara bersama-sama berpengaruh terhadap besarnya penjualan.

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	95,984	71,362		1,345	,191
	biaya yang dibelanjakan untuk iklan di koran	10,930	1,279	,871	8,548	,000
	biaya yang dibelanjakan untuk iklan di radio	5,038	3,313	,151	1,521	,141
	Jumlah outlet Penjualan	-13,067	4,965	-,266	-2,632	,014
	Jumlah tenaga Salesman	-13,800	5,259	-,262	-2,624	,015

a. Dependent Variable: Jumlah penjualan pada daerah itu

Analisis:

Dengan melihat kolom B pada Unstandardized Coefficient, kita bisa menulis persamaan regrssi menjad:

$$\text{Penjualan} = 95.985 + 10.930\text{Iklkor} + 5.038\text{Iklrad} - 13.067\text{Outlet} - 13.800\text{Salesman}$$

- ❖ Konstanta sebesar 95.985 menyatakan bahwa jika tidak ada iklan, outlet, dan salesman, maka penjualan adalah sebesar 95.985
- ❖ Koefisien Iklkor sebesar 10.930 menyatakan bahwa setiap penambahan (tanda+) Rp 1 pada biaya iklan di koran akan menyebabkan meningkatnya penjualan sebesar 10.930
- ❖ Demikian juga untuk ketiga variabel independen lain

Masih pada tabel tersebut terdapat kolom *t* yang (sebenarnya) digunakan untuk menguji signifikansi konstanta untuk tiap variabel.

Juga terdapat kolom *sig* yang menunjukkan tingkat signifikansi setiap uji *t* tersebut, dengan kaidah:

Jika probabilitas (=sig)  $\leq 0.05 \rightarrow H_0$  diterima atau tidak ada berpengaruh yang signifikan

Jika probabilitas (=sig)  $> 0.05 \rightarrow H_0$  ditolak atau ada berpengaruh yang signifikan

*Keputusan:*

Dengan melihat nilai probabilitas pada kolom *sig*, maka:

Variabel *iklkor*, *outlet*, dan *salesman* mempunyai angka *sig* di bawah 0.05. Berarti ketiga variabel independent tersebut memang berpengaruh secara signifikan terhadap variabel dependen

Variabel *iklrad* dan konstanta regresi mempunyai angka signifikansi di atas 0.05. Karena itu, kedua variabel tersebut sebenarnya tidak mempengaruhi penjualan.

Dengan demikian, variabel *iklrad* harus dikeluarkan dari model, kemudian prosedur pencarian model regresi diulang sekali lagi.

**Demikian, Wallahu a'lam.**



