

BAB 6

ANALISIS HUBUNGAN VARIABEL

6.1. TIPE-TIPE HUBUNGAN VARIBEL

Pertanyaan-pertanyaan penelitian di bidang bisnis sering berkisar dalam hal pengkajian hubungan antara 2 atau lebih variabel. Beberapa tujuan pengkajian ini diantaranya adalah mengetahui kuat atau tidaknya hubungan (strength), arah hubungannya (direction), bentuk hubungan (shape), dan sifat-sifat hubungan lainnya. Atau menyangkut pertanyaan taktis atau strategis yang bisa dijawab melalui perkiraan nilai satu variabel berdasarkan variabel lain. Beberapa contoh kasus di bidang manajemen atau bisnis yang menggambarkan keterkaitan antar variabel tersebut dapat dilihat sebagai berikut:

Kasus 1

Seorang tenaga penjualan berbagai jenis produk berusaha meningkatkan omzet penjualannya dengan memfokuskan penjualan setiap jenis produk kepada kelompok pelanggan yang tepat. Dia mensinyalir bahwa jenis produk yang dibutuhkan pelanggan berkaitan dengan jenis pekerjaan pelanggan tersebut. Bagaimana membuktikan dugaan tersebut?

Kasus 2

Seorang tenaga pengajar menghadapi persoalan yang berkaitan dengan efektifitas metoda pengajaran di Universitas Gunadarma. Selama ini dia secara bergantian menggunakan metode deep aproach, surface approach, atau kombinasinya. Dia ingin menentukan metode pengajaran mana yang disukai untuk masing-masing fakultas berdasarkan tanggapan dari mahasiswanya. Apakah ada kaitan antara fakultas dengan metode pengajaran?

Kasus 3

Seorang praktisi di bidang pengelolaan sumber daya manusia mempunyai dugaan bahwa produktifitas kerja karyawan berkaitan dengan motivasi kerja karyawan tersebut. Masalahnya adalah bagaimana perusahaan bisa mengukur tingkat motivasi kerja karyawan tersebut dan menghubungkannya dengan produktifitas karyawan?

Kasus4

Sebuah perusahaan asuransi mencatat bahwa total premi yang berhasil dikumpulkan cenderung meningkat dalam 12 bulan terakhir. Secara teoritis, kemampuan pembayaran premi ini terkait dengan tingkat

pendapatan masyarakat. Dapatkah bagian pemasaran perusahaan asuransi tersebut memasang target premi untuk bulan depan berdasarkan perkembangan tingkat pendapatan per kapita?

Semua kasus di atas bisa diselesaikan dengan mengukur hubungan antara variabel yang terlibat. Berbagai teknik pengukuran hubungan bisa digunakan tergantung skala pengukuran dan tujuan pertanyaan penelitiannya. Kasus 1 sampai 3 melibatkan variabel-variabel dengan tipe skala pengukuran nominal, ordinal, dan interval; sedangkan kasus 4 bisa dikaji dengan analisis regresi sederhana.

Dua istilah statistika yang populer dan sering digunakan dalam mengukur hubungan antar variabel adalah korelasi dan regresi. Dengan korelasi, kita menghitung suatu indeks yang mengukur hubungannya antara variabel. Sedangkan dengan regresi, suatu persamaan dikembangkan untuk memprediksi nilai suatu variabel tak bebas (*dependent variable*) berdasarkan nilai variabel bebas (*independent variable*). Yang perlu diperhatikan secara seksama adalah kapan kita menggunakan berbagai teknik mengukur tersebut. Sekali lagi, pemilihannya tergantung tipe skala dan tujuan pertanyaan penelitiannya. Tabel 5.1 mencantumkan berbagai cara pengukuran hubungan dan penggunaannya.

Tabel 5.1 menunjukkan bahwa setiap skala pengukuran mempunyai teknik statistika tertentu untuk mengukur hubungan variabelnya. Dua kemungkinan kesalahan yang dibuat jika kita tidak tepat menggunakan teknik analisis sesuai dengan tingkat skalanya adalah (1) kesalahan teknis dan (2) kehilangan informasi. Jika skala lebih rendah menggunakan teknik analisis untuk skala yang lebih tinggi maka terjadi kesalahan teknis. Sedangkan jika skala lebih tinggi dianalisis dengan teknik untuk skala yang lebih rendah maka kita kehilangan informasi. Gambar 5.1 menjelaskan posisi ketepatan teknik analisis dan kemungkinan kesalahan yang dibuat seorang peneliti.

		TEKNIK ANALISIS			
		Nominal	Ordinal	Interval	Rasio
SKALA	Nominal	Tepat			
	Ordinal		Tepat		
	Interval			Tepat	
	Rasio				Tepat
		Kesalahan Teknis			
		Kehilangan Informasi			

Tabel 5.1 Teknik pengukuran asosiasi antara variabel untuk setiap skala pengukuran

Pengukuran	Koefisien	Penggunaan
Interval dan rasio	Pearson product moment correlation ratio (eta)	Untuk variabel yang berhubungan secara linear dan kontinyu
	Biserial	Untuk data non linear atau menghubungkan pengaruh utama terhadap variabel dependent kontinyu
	Partial correlation	Satu variabel kontinyu dengan satu variabel dikotomi berdasarkan asumsi distribusi normal
	Multiple correlation	Tiga variabel: mengkaitkan dua variabel dengan pengaruh variabel ketiga
	Bivariate linear regression	Tiga variabel: mengkaitkan satu variabel dengan dua variabel lainnya
Ordinal	Gamma	Memprediksi satu variabel berdasarkan nilai variabel yang lainnya
	Kendall's tau -b	Berdasarkan pasangan konkordan dan diskordan: (P-Q); interpretasi proportional reduction in error (PRE)
	Kendall's tau -c	Berdasarkan P-Q; penyesuaian terhadap ranking yang sama
	Sommer's d	Berdasarkan P-Q; penyesuaian terhadap dimensi tabel
	Spearman's rho	Berdasarkan P-Q; perluasan asimetris dari Gamma
Nominal	Phi	korelasi produk moment untuk data yang diranking
	Cramer's V	Berdasarkan chi-square (CS) untuk tabel 2x2
	Contingency coefficient	Berdasarkan CS; penyesuaian jika satu dimensi tabel > 2
	Lambda	Berdasarkan CS; Asumsi data dan distribusi yang fleksibel
	Goodman & Kruskal's tau	Interpretasi berdasarkan PRE
	Uncertainty coefficient	Berdasarkan PRE dengan menekankan pada marginal tabel
Kappa	Berguna untuk tabel multidimensional	
		Mengukur kesepakatan

6.2. HUBUNGAN ANTAR VARIABEL NOMINAL

Mengukur keterkaitan antara variabel nominal menggunakan tabel klasifikasi silang. Beberapa teknik yang digunakan secara umum terdiri dari dua kelompok, yaitu koefisien yang berdasarkan pada chi square dan koefisien berdasarkan proportional

reduction in error (PRE). Tiga koefisien yang berdasarkan chi square adalah sebagai berikut:

1. Phi (ϕ)

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

Soal Latihan

Seorang mahasiswa Universitas Gunadarma melakukan penelitian mengenai hubungan antara jenis kelamin konsumen (laki-laki atau perempuan) dengan preferensinya (suka atau tidak suka) terhadap suatu produk baru yang akan dipasarkan. Informasi dari penelitian tersebut dibutuhkan untuk penentuan segmen pasar dan strategi promosinya. Dengan menggunakan instrumen peneltian yang sudah dirancangnya, data hasil survey terhadap 40 responden dapat dilihat pada tabel berikut:

Respon den ke	Jenis Kelami n	Preferensi	Respon den ke	Jenis Kelamin	Preferensi
1	L	tidak suka	21	P	suka
2	P	Suka	22	P	suka
3	L	Suka	23	L	tidak suka
4	P	suka	24	L	suka
5	L	tidak suka	25	P	tidak suka
6	L	tidak suka	26	P	suka
7	L	suka	27	P	suka
8	P	suka	28	L	tidak suka
9	L	tidak asuka	29	P	suka
10	P	suka	30	L	tidak suka
11	L	tidak suka	31	P	suka
12	P	suka	32	L	tidak suka
13	L	tidak suka	33	P	suka
14	P	suka	34	L	tidak suka
15	P	tidak suka	35	L	suka
16	P	suka	36	L	suka
17	L	tidak suka	37	P	suka
18	L	tidak suka	38	P	tidak suka
19	P	suka	39	P	suka
20	P	suka	40	L	tidak suka

Apakah ada hubungan antara jenis kelamin dengan preferensinya?

SPSS

Data (open) : Soal6_1
 Menu : **Statistics Summarize Crosstabs**
 Option : **Phi and Cramer's V** untuk *statistic*
 row, coloumn dan total untuk cell
 Output :

GENDER Jenis Kelamin by PREFEREN preferensi konsumen

		PREFEREN Page 1 of 1		
Count		tidak	su	suka
Row Pct				
Col Pct		ka		Row
Tot Pct		0	1	Total
GENDER				
0		3	18	21
wanita		14,3	85,7	52,5
		17,6	78,3	
		7,5	45,0	
1		14	5	19
laki		73,7	26,3	47,5
		82,4	21,7	
		35,0	12,5	
Column		17	23	40
Total		42,5	57,5	100,0

Statistic	Value	ASE1	Approximate Val/ASE0	Significance
Phi	,60003			,00015 *1
Cramer's V	,60003			,00015 *1

*1 Pearson chi-square probability

2. Crammer's Phi

$$V = \sqrt{\frac{\chi^2}{N(k-1)}}$$

dimana k adalah kolom atau baris yang lebih kecil

3. Contingency Coefficient C

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

Soal Latihan

Sebuah penelitian dilakukan untuk mengetahui keterkaitan antara jenis perusahaan dengan sikap pemilik perusahaannya terhadap masuknya investor dari luar perusahaan. Kategori jenis perusahaannya adalah Manufaktur (M), Perdagangan (D), atau Jasa (J); sedangkan kategori sikapnya adalah setuju atau menolak. Adakah keterkaitan antara kedua variabel tersebut jika diketahui hasil penelitian terhadap 50 perusahaan sebagai berikut.

No	Perusahaan	Sikap	No	Perusahaan	Sikap
1	M	Setuju	26	J	Setuju
2	D	Setuju	27	J	Tidak Setuju
3	D	Tidak setuju	28	D	Setuju
4	J	Tidak setuju	29	D	Tidak setuju
5	M	Setuju	30	D	Setuju
6	J	Setuju	31	M	Setuju
7	M	Tidak setuju	32	D	Tidak setuju
8	M	Setuju	33	J	Tidak setuju
9	D	Setuju	34	J	Tidak setuju
10	J	Tidak setuju	35	M	Tidak setuju
11	J	Tidak setuju	36	D	Tidak setuju
12	M	Setuju	37	M	Setuju
13	M	Tidak setuju	38	D	Tidak setuju
14	D	Tidak setuju	39	D	Setuju
15	M	Setuju	40	J	Tidak setuju
16	J	Tidak setuju	41	J	Setuju
17	J	Setuju	42	J	Tidak setuju
18	M	Setuju	43	M	Setuju
19	D	Setuju	44	J	Tidak setuju
20	M	Setuju	45	D	Tidak setuju
21	J	Tidak setuju	46	D	Setuju
22	J	Setuju	47	M	Setuju
23	J	Tidak setuju	48	M	Setuju
24	M	Tidak setuju	49	J	Setuju
25	M	Setuju	50	D	Tidak setuju

SPSS

Data (open) : Soal6_2

Menu : **Statistics Summarize Crosstabs**

Option : **contingency coefficient** dan *phi* dan *cramer's V* untuk **statistic row, coloumn** dan **total** untuk **cell**

Output :

SIKAP by J_USAHA

		J_USAHA			
		D	J	M	Row Total
SIKAP 0 TIDAK SETUJU	Count	8	12	4	24
	Row Pct	33,3	50,0	16,7	48,0
	Col Pct	53,3	66,7	23,5	
	Tot Pct	16,0	24,0	8,0	
1 SETUJU	Count	7	6	13	26
	Row Pct	26,9	23,1	50,0	52,0
	Col Pct	46,7	33,3	76,5	
	Tot Pct	14,0	12,0	26,0	
Column Total	15	18	17	50	
Total	30,0	36,0	34,0	100,0	

Statistic	Value	ASE1	Val/ASE0	Approximate Significance
Phi	,36776			,03401 *1
Cramer's V	,36776			,03401 *1
Contingency Coefficient	,34516			,03401 *1

Berdasarkan nilai ketiga ukuran hubungan antara variabel nominal tersebut dapat disimpulkan bahwa terdapat hubungan antara jenis usaha dengan sikap para pemiliknya terhadap investor dari luar. Ketiga statistik tersebut hanya menunjukkan kuatnya hubungan tanpa mengetahui arahnya. Hal ini sesuai dengan karakteristik skala ordinal yang tidak mempunyai sifat urutan diantara kategori-kategorinya. Nilai ketiga statistik tersebut berkisar antara 0 dan 1. 0 menunjukkan tidak ada hubungan sama sekali sedangkan 1 menunjukkan hubungan yang sempurna antara dua variabel nominal

4. Lambda (λ)

Statistik lain yang bisa digunakan untuk mengukur hubungan antara dua variabel yang keduanya berskala nominal adalah lambda. Rumus yang digunakan untuk menghitungnya adalah:

$$\lambda = \frac{\sum f_i - F_d}{n - F_d}$$

dimana f_i adalah modus frekuensi dalam setiap kategori variabel bebas, F_d adalah modus frekuensi diantara total variabel tak bebas, dan n adalah banyaknya satuan pengamatan. Nilai lambda berkisar antara 0 dan 1 yaitu dari tidak ada hubungan sama sekali sampai hubungan kuat.

Misalkan kita akan mengukur hubungan antara jenis perusahaan dan sikap terhadap investor dari luar pada soal 6.2 diatas. Langkah pertama adalah menyajikan kembali hasil survei dalam bentuk tabel silang berikut:

	Jenis Perusahaan			Jumlah
	Dagang	Jasa	Manufaktur	
Tidak Setuju	8	12	4	24
Setuju	7	6	13	26
Jumlah	15	18	17	50

Misalkan berdasarkan teori sebelumnya, Anda menduga bahwa sikap terhadap kehadiran investor dipengaruhi oleh jenis perusahaan. Jadi jenis perusahaan disini merupakan variabel bebas sedangkan sikap terhadap investor merupakan variabel tak bebas.

Langkah selanjutnya adalah menentukan nilai f_i untuk setiap kolom, yaitu $f_1 = 8$ (karena untuk jenis perusahaan dagang frekuensi yang sering muncul atau modusnya adalah tidak setuju sebesar 8), $f_2 = 12$, dan $f_3 = 13$. Sedangkan untuk F_d adalah sebesar 26, yaitu dilihat frekuensi terbesar pada kolom jumlah variabel tak bebas. Jadi nilai lambdanya bisa dihitung sebagai berikut:

$$\begin{aligned}\lambda &= \frac{(8 + 12 + 13) - 26}{50 - 26} \\ &= 0.2917\end{aligned}$$

SPSS

Data (open) : Soal6_2
Menu : *Statistics Summarize Crosstabs*
Option : *Lambda* untuk *statistic*
Output :

Statistic	Value	ASE1	Val/ASE0	Approximate Significance
Lambda :				
symmetric	,25000	,12877	1,80619	
with SIKAP dependent	,29167	,20145	1,23705	
with J_USAHA dependent	,21875	,12040	1,64900	

HUBUNGAN ANTAR VARIABEL ORDINAL

1. Gamma G

$$\gamma = \frac{P - Q}{P + Q}$$

P adalah banyaknya pasangan yang *concordant* sedangkan Q adalah banyaknya pasangan yang *discordant*. Gamma berkisar antara -1 (hubungan tidak searah sempurna) dan +1 (hubungan searah sempurna).

Soal Latihan

Seorang peneliti sedang melakukan penelitian untuk mengetahui hubungan antara daya saing perusahaan dengan tingkat teknologi yang diterapkan di perusahaan. Dugaannya adalah semakin tinggi tingkat teknologi maka daya saingnya akan meningkat. Variabel daya saing skalanya adalah rendah, cukup dan tinggi; sedangkan tingkat teknologinya adalah rendah, menengah, dan tinggi. Perusahaan yang menjadi sampel penelitian adalah sebanyak 65 perusahaan dengan hasil pengamatan sebagai berikut:

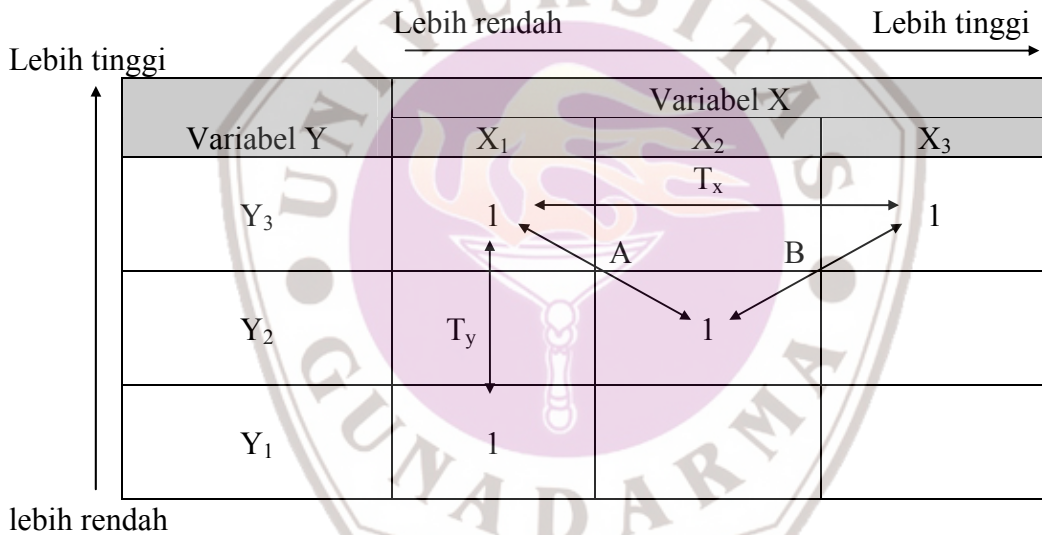
Tingkat Teknologi	Daya saing perusahaan			Total baris
	rendah	Cukup	tinggi	
Tinggi	2	6	20	28
Menengah	3	5	9	17
Rendah	12	6	2	20
Total kolom	17	17	31	65

Apakah ada hubungan antara kedua variabel tersebut? Seberapa kuat hubungan tersebut dan bagaimana arah hubungannya?

Penentuan pasangan concordant dan discordant

Jika sebuah subyek mempunyai urutan lebih tinggi pada satu variabel dan lebih tinggi juga pada variabel kedua dibandingkan subyek lain, maka pasangan observasi (dua subyek) tersebut dikatakan concordant. Jika sebuah subyek mempunyai urutan lebih tinggi dibandingkan subyek lain pada satu variabel tetapi mempunyai urutan lebih rendah pada variabel kedua, maka pasangan subyek tersebut disebut discordant. Pengertian tersebut akan lebih jelas dengan menggunakan ilustrasi sebagai berikut:

Misalkan diketahui tabel silang dengan variabel 1 dan variabel 2 yang sama-sama mempunyai skala ordinal (ingat, skala ordinal menunjukkan adanya urutan). Untuk memudahkan pemahaman, sel yang terisi hanya 3 buah dengan frekuensi masing-masing sebesar 1, seperti terlihat pada Tabel berikut:



Pasangan A adalah pasangan discordant karena X₁ **lebih rendah** dari X₂ tetapi Y₃ **lebih tinggi** dari Y₂, atau bisa juga dinyatakan X₂ **lebih tinggi** dari X₁ tetapi Y₂ **lebih rendah** dari Y₃. Pasangan B adalah pasangan concordant karena X₂ **lebih rendah** X₃ dan Y₂ **lebih rendah** dari Y₃, atau bisa juga dinyatakan X₃ **lebih tinggi** dari X₂ dan Y₃ **lebih tinggi** dari Y₂. Jadi bisa dikatakan pasangan concordant menunjukkan perubahan ranking pada variabel X **searah** dengan perubahan ranking pada Variabel Y, sedangkan pasangan discordant menunjukkan perubahan ranking pada variabel X **tidak searah** dengan perubahan ranking pada Variabel Y. Banyaknya pasangan searah (concordant) dan tidak searah (discordant) itulah yang menjadi dasar perhitungan statistik Gamma. Pasangan T_x dan T_y akan dijelaskan pada perhitungan statistik berikutnya.

Contoh ilustrasi diatas hanya melibatkan 3 satuan pengamatan sehingga jumlah pasangan yang bisa dibentuk relatif sedikit, yaitu hanya 3 pasangan. Bagaimana jika semua sel terisi dengan frekuensi yang lebih banyak. Hal ini bisa dijelaskan dengan

menjawab secara lengkap soal latihan diatas mengenai hubungan tingkat teknologi dengan daya saing perusahaan.

$$\begin{aligned} \text{Total pasangan yang bisa dibentuk} &= n(n-1)/2 \\ &= 65(64)/2 = 2080 \text{ pasang} \\ \text{Jumlah pasangan concordant (P)} &= 20(5+3+6+12) + 6(3+12) + 9(12+6) + 5(12) \\ &= 832 \\ \text{Jumlah pasangan discordant (Q)} &= 2(5+9+6+2) + 6(9+2) + 3(6+2) + 5(2) \\ &= 144 \end{aligned}$$

Jadi:

$$\begin{aligned} \gamma &= \frac{P - Q}{P + Q} \\ &= \frac{832 - 144}{832 + 144} \\ &= 0.7049 \end{aligned}$$

Interpretasinya adalah variabel daya saing berhubungan cukup erat dengan variabel tingkat teknologi. Hubungan tersebut bersifat serarah (positif), atau bisa dikatakan semakin tinggi tingkat teknologi komputer yang digunakan maka semakin tinggi daya saing perusahaan.

2. Kendall's Tau -b (τ_b)

Statistik ini merupakan penyempurnaan gamma yang masih memperhitungkan pasangan *pair*, yaitu pasangan yang mempunyai nilai yang sama pada variabel X, variabel Y, atau keduanya.

$$\tau = \frac{P - Q}{\sqrt{\left(\frac{n(n-1)}{2} - T_x\right)\left(\frac{n(n-1)}{2} - T_y\right)}}$$

n adalah banyaknya satuan pengamatan, Tx adalah banyaknya pasangan ties pada variabel kolom, dan Ty adalah banyaknya pasangan ties pada variabel baris (lihat contoh pasangan Tx dan Ty pada tabel diatas. τ juga berkisar antara -1 (hubungan negatif sempurna) dan +1 (hubungan positif sempurna).

Kita akan menghitung statistik τ untuk contoh yang sama. Terlebih dahulu kita menghitung Tx dan Ty dengan rumus berikut:

$$T_x = \frac{17(16) + 17(16) + 31(30)}{2}$$

$$= 737$$

$$T_y = \frac{28(27) + 17(16) + 20(19)}{2}$$
$$= 704$$

Nilai statistik τ bisa dihitung sebagai berikut:

$$\tau = \frac{832 - 144}{\sqrt{(2080 - 737)(2080 - 704)}}$$
$$= 0.5061$$

3. Kendall's Tau -c (τ_c)

$$\tau_c = \frac{2m(P - Q)}{N^2(m - 1)}$$

m adalah ukuran kolom atau baris yang terkecil. Statistik ini lebih teliti dibandingkan statistik sebelumnya terutama untuk ukuran tabel yang lebih besar. Nilainya berkisar antara -1 dan +1. Jika hubungan variabel pada soal diatas diukur dengan statistik ini, maka hasilnya adalah sebagai berikut:

$$\tau_c = \frac{2m(P - Q)}{N^2(m - 1)}$$
$$= \frac{2(3)(832 - 144)}{65^2(3 - 1)}$$
$$= 0.4885$$

4. Sommers's d

Sommers's d juga merupakan statistik untuk mengukur hubungan dua variabel berdasarkan pasangan concordant dan discordant. Kelebihannnya statistik ini adalah kemampuannya untuk menghilangkan pengaruh pasangan pair dan disesuaikan pada arah variabel bebas. Statistik ini juga bisa digunakan untuk hubungan simetris dan asimetris, sehingga ada tiga bentuk rumus perhitungannya, yaitu:

$$d_{sym} = \frac{(P - Q)}{n(n - 1) - T_x T_y / 2} \quad (1)$$

$$d_{y-x} = \frac{(P - Q)}{\frac{n(n - 1)}{2} - T_x} \quad (2)$$

$$d_{x-y} = \frac{(P - Q)}{\frac{n(n - 1)}{2} - T_y} \quad (3)$$

Persamaan (1) adalah simetris, persamaan (2) adalah asimetris dimana y sebagai variabel bebas, dan persamaan (3) adalah asimetris dimana x sebagai variabel bebas. Nilai statistik ini untuk soal diatas adalah sebagai berikut:

$$d_{sym} = \frac{832 - 144}{65(64) - 737(704)/2} =$$
$$d_{y-x} = \frac{832 - 144}{2080 - 737} = 0.5123$$
$$d_{x-y} = \frac{832 - 144}{2080 - 704} = 0.5$$

SPSS

Kita akan melakukan analisis antara dua variabel ordinal contoh soal yang sudah dihitung secara manual diatas, dengan menggunakan SPSS. Menu yang digunakan untuk menganalisis hubungan variabel nominal dan ordinal selalu melalui menu crosstabs karena dasar perhitungannya adalah distribusi frekuensi pada tabel silne tersebut.

Data (open) : Gamma.sav

Menu : **Statistics Summarize Crosstabs**

Pilih variabel **saing** sebagai **coloum variable** dan **tekno** sebagai **row variable**. Klik **statistics** dan pada kotak dialog yang muncul, pilih **Gamma, Kendall tau-b, Kendall-tau c, dan Sommers'd**.

Output :

TEKNO TINGKAT TEKNOLOGI by SAING DAYA SAING PERUSAHAAN

Count	SAING			Row Total
	1	2	3	
TEKNO 1	12	6	2	20 30,8
2	3	5	9	17 26,2
3	2	6	20	28 43,1
Column Total	17 26,2	17 26,2	31 47,7	65 100,0

Statistic	Value	ASE1	Val/ASE0	Approximate Significance
Kendall's Tau-b	,50611	,08700	5,73032	
Kendall's Tau-c	,48852	,08525	5,73032	
Gamma	,70492	,09800	5,73032	
Somers' D :				
symmetric	,50607	,08700	5,73032	
with TEKNO dependent	,51229	,08850	5,73032	
with SAING dependent	,50000	,08682	5,73032	

Berdasarkan nilai-nilai statistik diatas maka dapat disimpulkan bahwa ada hubungan positif antara tingkat teknologi dengan daya saing perusahaan. Untuk populasi sasaran yang diamati, semakin tinggi tingkat teknologi yang diterapkan perusahaan maka semakin meningkat daya saing perusahaan tersebut.

5. Spearman's rho

Spearman's rho merupakan statistik yang populer digunakan untuk mengukur hubungan antara dua variabel ordinal. Statistik ini, yang merupakan bentuk khusus dari korelasi Pearson's product moment, mengkaitkan ranking 2 buah variabel. Teknik ini juga biasa digunakan untuk skor variabel yang tadinya interval atau rasio tetapi skor tersebut terlalu banyak ketidaknormalannya sehingga skor tersebut diturunkan tingkatannya dalam bentuk urutan nilainya saja dan dihitung nilai spearman's rho-nya, yaitu dengan rumus sebagai berikut:

$$r_s = 1 - \frac{6\sum d^2}{n^3 - n}$$

n adalah banyaknya subyek yang diranking sedangkan d adalah perbedaan ranking antara kedua variabel.

Statistik ini memang mudah dihitung tetapi mengandung kelemahan utama, yaitu sangat sensitif terhadap ranking yang sama. Jadi semakin banyak ranking yang sama untuk satu variabel akan menyebabkan statistik ini tidak akurat untuk mengukur hubungan variabel. Contoh pengukuran hubungan variabel dengan statistik ini adalah sebagai berikut:

Contoh soal

Sepuluh bank orang calon pegawai menjalani proses penilaian oleh 2 buah tim, yaitu tim manajemen yang terdiri dari 4 orang eksekutif perusahaan dan tim psikologi lembaga pengembangan sumber daya manusia dari luar perusahaan. Penilaian yang dilakukan oleh kedua tim tersebut adalah dalam bentuk ranking, yang dapat dilihat pada kolom kedua dan ketiga berikut:

Calon Pegawai	Ranking menurut tim:		d	d ²
	Manajemen	Psikologi		
A	3.5	6	-2.5	6.25
B	10	5	5	25
C	6.5	8	-1.5	2.25
D	2	1.5	0.5	0.25
E	1	3	-2	4
F	9	7	2	4
G	3.5	1.5	2	4
H	6.5	9	-2.5	6.25
I	8	10	-2	4
J	5	4	1	1
				57

Berdasarkan hasil perhitungan yang disajikan kolom keempat dan kelima pada tabel di atas maka nilai statistik sparman rhonya adalah sebagai berikut:

$$\begin{aligned}
 r_s &= 1 - \frac{6\sum d^2}{n^3 - n} \\
 &= 1 - \frac{6(57)}{(10)^3 - 10}
 \end{aligned}$$

$$= 0.654$$

Hasil tersebut menunjukkan bahwa hubungan kedua penilaian tersebut adalah positif dan relatif cukup kuat. Atau bisa dikatakan kedua tim memberikan hasil penilaian yang cukup konsisten satu sama lain.

6.4. HUBUNGAN VARIABEL INTERVAL DAN RATIO

Mengukur hubungan antar variabel interval dan ratio merupakan pokok bahasan utama dalam analisis statistik bivariate. Analisis asosiasi bivariate berbeda dengan analisis regresi dan analisis asosiasi non parametrik dalam 2 hal. Pertama, korelasi parametrik membutuhkan 2 variabel yang kontinyu dengan skala pengukuran interval dan rasio. Kedua, koefisien tidak membedakan variabel bebas dan variabel tak bebas. Teknik-teknik analisis yang akan dibahas disini adalah koefisien produk momen Pearson, matriks korelasi, dan analisis regresi

1. Koefisien *Pearson's product moment*

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{(N-1)s_x s_y}$$

dimana

N adalah jumlah pasangan kasus
 s_x dan s_y adalah standar deviasi x dan y, yaitu

$$s_x = \frac{\sqrt{\sum x^2}}{\sqrt{N}} ; s_y = \frac{\sqrt{\sum y^2}}{\sqrt{N}}$$

Koefisien korelasi produk momen Pearson berkisar antara -1 sampai +1. Koefisien ini menunjukkan besar dan arah keterkaitan antara variabel. Hubungan dikatakan lemah jika mendekati 0 dan kuat jika mendekati +1 atau -1, sedangkan arah hubungan adalah searah jika positif dan berlawanan arah jika negatif.

Soal Latihan

Seorang peneliti akan mengkaji hubungan antara keuntungan bersih dengan aliran kas di sepuluh perusahaan dengan data tercantum pada Tabel berikut

Perusahaan	Keuntungan bersih (Jutaan Rp)	Aliran Kas (Jutaan Rp)
1	82.6	126.5
2	89.0	191.2
3	176.0	267.0
4	82.3	137.1

5	413.5	806.8
6	18.1	35.2
7	337.3	425.5
8	145.8	380.0
9	172.6	326.6
10	247.2	355.5

Hitunglah besar dan arah hubungan antara variabel keuntungan dengan aliran kas dengan menggunakan koefisien produk momen Spearman?

2. Matriks Korelasi

Matriks korelasi adalah tabel yang digunakan untuk menampilkan koefisien-koefisien korelasi untuk lebih dari dua variabel.

Soal Latihan

Seorang mahasiswa Fakultas Ekonomi Universitas Gunadarma meneliti 10 perusahaan dengan mengukur enam variabel yang menjadi perhatiannya. Apakah ada keterkaitan antara keenam variabel tersebut jika diketahui hasil pengamatannya sebagai berikut:

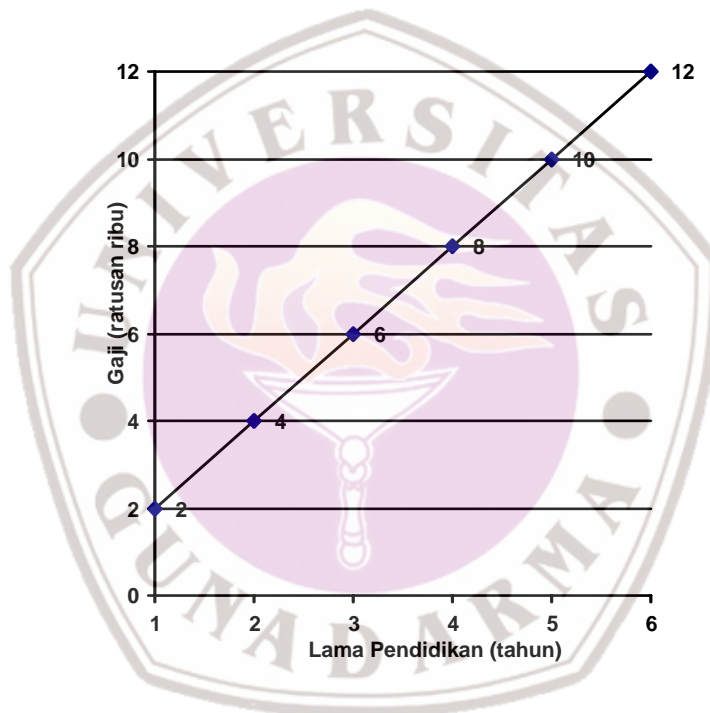
Perusahaan	Assets	Sales	Market value	Net profit	cash flow	Jumlah karyawan
1	1034.00	1510.00	697.00	82.60	126.50	16.60
2	956.00	785.00	1271.00	89.00	191.20	5.00
3	1890.00	2533.00	1783.00	176.00	267.00	44.00
4	1133.00	532.00	752.00	82.30	137.10	2.10
5	11682.00	3790.00	4149.00	413.50	806.80	11.90
6	6080.00	635.00	291.00	18.10	35.20	3.70
7	31044.00	3296.00	2705.00	337.30	425.50	20.10
8	5878.00	3204.00	2100.00	145.80	380.00	10.80
9	1721.00	981.00	1573.00	172.60	326.60	1.90
10	2135.00	2268.00	2634.00	247.20	355.50	21.20

3. Analisis Regresi Linear Sederhana

Seorang calon sarjana Universitas Gunadarma melakukan penelitian pada bidang manajemen sumber daya manusia. Variabel yang diamati adalah lama pendidikan dan besarnya gaji pertama yang diterima karyawan. Hasil pengamatan yang dilakukan terhadap 6 orang karyawan yang menjadi satuan pengamatan disajikan pada Tabel berikut:

Individu	Lama Pendidikan	Gaji (ratusan ribu)
A	1	2
B	2	4
C	3	6
D	4	8
E	5	10
F	6	12

Data diatas dapat digambarkan dalam suatu grafik dengan variabel bebas (independent) diletakkan pada sumbu X dan variabel tak beba (dependent) diletakkan pada sumbu Y. Setiap individu diwakili oleh satu titik dalam grafik seperti terlihat pada Gambar berikut ini.



Garis tersebut memiliki persamaan $Y=2X$. Jadi kita bisa meramalkan nilai Y secara pasti bila diketahui nilai variabel X. Angka dua dari persamaan garis disebut koefisien regresi, yang artinya setiap perubahan satu unit dalam variabel X akan diikuti oleh perubahan dua unit Y. Variabel-variebel ini dihubungkan secara lurus (linear), yaitu sebuah garis lurus yang memberikan gambaran tentang hubungan kedua varibel tersebut. Garis ini disebut garis regresi linear, regresi dari variabel tak bebas Y yang didasarkan atas variabel bebas X.

Jika jumlah variabel yang dianalisis dengan regresi linear ini ada 2 buah variabel maka disebut regresi linear sederhana. Bentuk umum dari setiap garis lurus dinyatakan dalam persamaan sebagai berikut:

$$Y = a + bX$$

a adalah konstanta atau *Y-intercept* dan b adalah koefisien arah atau koefisien regresi. Konstanta dan koefisien tersebut bisa dicari dengan menggunakan menggunakan kriteria kuadrat terkecil, yang rumus akhirnya adalah sebagai berikut:

$$a = \bar{y} - b\bar{x} \quad \text{dimana } y \text{ rata-rata adalah } (\Sigma y)/n \text{ dan } x \text{ rata-rata adalah } (\Sigma x)/n$$

$$b = \frac{n \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)}{n \Sigma x_i^2 - (\Sigma x_i)^2}$$

Perhitungan selengkapnya secara manual untuk contoh kasus diatas adalah sebagai berikut:

X_i	Y_i	$X_i Y_i$	X_i^2
1	2	2	1
2	4	8	4
3	6	18	9
4	8	32	16
5	10	50	25
6	12	72	36
$\Sigma X_i = 21$	$\Sigma Y_i = 42$	$\Sigma X_i Y_i = 182$	$\Sigma X_i^2 = 91$
rata ² x = 3.5	rata ² y = 7		

y rata-rata adalah 21/6 atau 7 sedangkan x rata-rata adalah 21/6 atau 3. Perhitungan nilai koefisien b adalah:

$$b = \frac{n \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)}{n \Sigma x_i^2 - (\Sigma x_i)^2}$$

$$= \frac{6 \times 182 - 21 \times 42}{6 \times 91 - 21^2}$$

$$= 2$$

sedangkan nilai konstanta a adalah:

$$a = y - bx$$

$$= 7 - 2 \times 3.5 = 0$$

Seorang peneliti tertarik untuk meneliti hubungan antara besarnya Loan to Deposit Ratio (LDR) dengan nilai Return on Asset bank sehingga bisa diperkirakan perkiraan nilai LDR sebuah bank dengan mengetahui nilai LDR-nya. Data LDR dan ROA 24 buah bank yang menjadi sampel penelitian adalah sebagai berikut:

Nama Bank	LDR (persen)	ROA (persen)
A	106.43	1,16
B	108.64	1,05
C	105.61	1,07
D	119.87	,62
E	107.52	0,95
F	107.94	1,09
G	124.86	,02
H	113.72	,29
I	81.30	2,07
J	123.54	,25
K	128.44	,11
L	124.94	,06
M	107.69	,95
N	108.35	1,01
O	107.57	,92
P	100.43	1,04
Q	113.77	,28
R	97.90	2,03
S	109.71	1,01
T	108.03	1,01
U	115.92	,21
V	118.32	,51
W	104.58	1,05
X	100.77	1,21

SPSS

Data (Open) : linear.sav

Menu : Statistics Regression Linear

Pilih LDR sebagai **independent variable** dan ROA sebagai **dependent variable**. Klik **statistics** dan pada kotak dialog yang muncul, pilihlah **estimate** dan **Model Fit**. Jika sudah selesai klik **continue** atau **OK**

Output :

R Square ,84173

Adjusted R Square	,83454		
Standard Error	,22215		
Analysis of Variance			
	DF	Sum of Squares	Mean Square
Regression	1	5,77413	5,77413
Residual	22	1,08567	,04935
F =	117,00724	Signif F =	,0000

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
LDR	-,048704	,004503	-,917461	-10,817	,0000
(Constant)	6,201441	,498449		12,441	,0000

4. Analisis Regresi Linear Berganda

Model

Analisis regresi merupakan teknik statistika yang digunakan untuk mengidentifikasi dan membuat model mengenai kaitan antar variabel. Kaitan tersebut adalah antara satu variabel tak bebas dengan satu atau lebih variabel bebas. Jika model regresi tersebut melibatkan lebih dari satu variabel bebas dengan kaitan bersifat linear maka disebut model *regresi liner berganda*. Model umum matematisnya adalah sebagai berikut:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k + \varepsilon$$

yang terdiri dari satu *variabel tak bebas* (y) dan k buah *variabel bebas* (x_i). Parameter-parameter β_j dimana j =0, 1, ..., k disebut *koefisien regresi*. Pengertian *linear* digunakan untuk menunjukkan fungsi linear dari parameter β_j -nya. Sedangkan ε menunjukkan pengaruh variabel lain yang tidak diamati. Model regresi tersebut akan diduga oleh model regresi yang diformulasikan berdasarkan contoh data yang diamati.

Perhitungan yang akan digunakan dalam analisis regresi linear berganda disini melibatkan 1 variabel bebas dan 4 buah variabel tak bebas dengan contoh data sebanyak 24 buah item data untuk bidang kajian manajemen teknologi. Analisis yang dilakukan mencakup pengujian apakah datanya memenuhi asumsi distribusi normal, estimasi parameter model, pendugaan selang kepercayaan untuk populasinya, pengujian signifikansi hubungan linear pada model regresi yang diformulasikan, serta diagnosis regresinya yang terdiri dari evaluasi kecocokan model regresi, identifikasi dan pengaruh data pencilan, dan pengujian adanya multikolinearitas.

Pendefinisian Variabel

Analisis regresi ini adalah untuk melihat hubungan antara daya saing bank (Y) dengan Kualifikasi Sumber Daya Manusia di bidang komputer(X_1), tingkat restrukturisasi organisasi (X_2), Tingkat integrasi Teknologi informasi yang diimplementasikan(X_3), dan Tingkat diversifikasi produk dan jasa perbankan berbasis teknologi informasi (X_4). Dengan analisis regresi diharapkan dapat diketahui variabel mana yang terkait linear dan seberapa besar signifikansinya terhadap daya saing bank. Jenis dan fungsi masing-masing variabel tersebut selengkapnya dapat dilihat pada tabel 1 berikut.

Tabel 1. Jenis dan fungsi variabel yang akan dianalisis dengan model regresi linear berganda

Notasi	Fungsi variabel	Label
Y	tak bebas	tingkat daya saing bank yang telah mengimplementasikan teknologi informasi
X_1	bebas	Kualifikasi SDM di bidang teknologi informasi
X_2	bebas	Tingkat restrukturisasi organisasi
X_3	bebas	Tingkat integrasi teknologi yang diimplementasikan
X_4	bebas	Diversifikasi produk dan jasa bank berbasis teknologi informasi

Soal 8.1

Taraf pengukuran variabel-variabel yang digunakan adalah ordinal sehingga perlu dikonversi ke skala interval supaya bisa dianalisis dengan menggunakan model regresi. Data hasil transformasi ke skala interval selengkapnya dapat dilihat pada Tabel 2 berikut ini.

Tabel 2. Data pengamatan setelah dikonversi ke skala interval

ITEM	X_1	X_2	X_3	X_4	Y
1.	106.433	81.023	27.302	33.407	136.673
2.	108.636	82.700	32.700	49.550	129.085
3.	105.611	86.293	33.856	57.636	156.384
4.	119.871	75.461	27.829	72.004	117.553
5.	107.520	74.381	32.344	64.077	140.681
6.	107.939	84.675	28.439	70.027	154.653
7.	124.857	90.150	30.036	69.616	160.000
8.	113.725	88.795	29.988	66.741	118.320
9.	81.300	40.800	25.584	36.300	98.400

10.	123.541	70.600	27.202	57.661	137.100
11.	128.440	76.973	31.850	57.131	160.355
12.	124.944	84.874	30.665	50.286	137.780
13.	107.695	70.300	27.634	64.424	135.881
14.	108.347	83.910	27.329	62.126	134.872
15.	107.569	82.204	32.227	65.430	131.937
16.	100.435	73.246	26.075	48.082	110.000
17.	113.772	90.778	28.087	70.027	156.323
18.	97.900	82.980	27.538	55.662	118.014
19.	109.709	83.772	30.144	62.227	130.084
20.	108.032	77.143	27.781	57.850	142.554
21.	115.916	87.070	33.463	70.027	159.621
22.	118.319	96.701	32.240	56.800	154.118
23.	104.575	71.182	29.329	50.755	132.008
24.	100.768	74.661	32.380	55.474	130.000

Asumsi Kenormalan Data

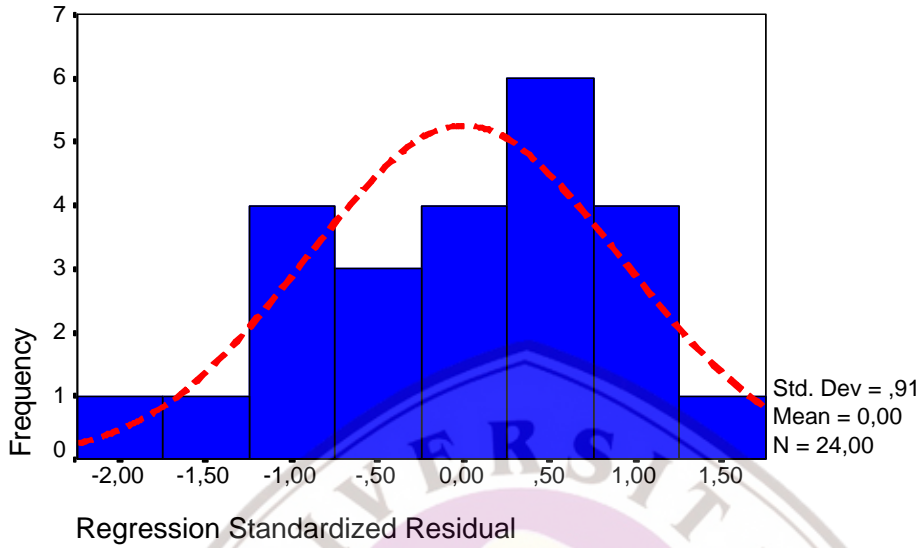
Analisis regresi mengandung asumsi-asumsi yang meliputi variabel tak bebas berhubungan secara linear dengan variabel bebas, variabel bebas tidak berkorelasi, semua variabel dalam skala interval, variabel dependent harus berdistribusi normal, dan variasi variabel tak bebas harus konstan. Pengujian yang akan dilakukan disini adalah uji normal saja.

Uji normal berdasarkan analisis grafik dengan menggunakan 2 jenis grafik, yaitu (1) histogram distribusi normal untuk data residual dan (2) Grafik yang memplot probabilitas kumulatif data observasi dengan dengan probabilitas kumulatif hasil regresinya. Kedua grafik yang dapat dihasilkan oleh program SPSS berdasarkan data pengukuran dapat dilihat pada Gambar 1 dan 2 berikut. Berdasarkan gambar 1, terlihat bahwa distribusi frekuensi kumulatif untuk data residu yang sudah distandarkan mendekati kurva normal (ditunjukkan oleh garis putus-putus berwarna merah).

Sedangkan berdasarkan Gambar 2 diatas maka data observasi juga mendekati distribusi normal, yaitu ditunjukkan dengan titik-titik yang secara acak dan tidak berpola berada disekitar garis diagonal (garis dimana probabilitas kumulatif observasi sama dengan probabilitas kumulatif regresi). Jadi asumsi mengenai kenormalan data terpenuhi atau data hasil pengamatan tersebut bisa dianalisis dengan menggunakan regresi linear.

SPSS

Menu : Statistic Regression
Option :
Output :



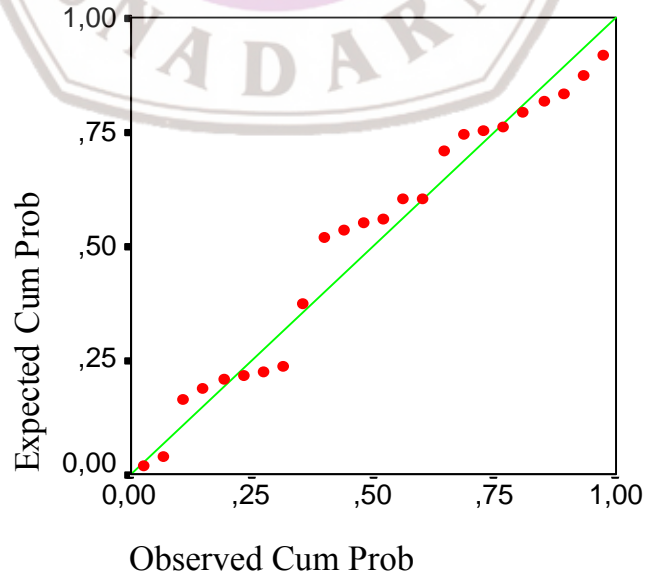
Gambar 1. Histogram distribusi frekuensi kumulatif

SPSS

Menu : Statistics Regression
Option :

Output

Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Y



Gambar 2. Plot probabilitas normal untuk standar residual

Pendugaan dan Uji Hipotesis

Model persamaan regresi linear berganda yang akan digunakan dalam perhitungan adalah sebagai berikut :

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Penentuan nilai-nilai koefisien persamaan regresi, yaitu β_0 , β_1 , β_2 , β_3 , dan β_4 , menggunakan metoda jumlah kuadrat terkecil. Perhitungan selengkapnya untuk penentuan nilai-nilai koefisien tersebut dengan menggunakan perangkat lunak SPSS adalah sebagai berikut :

SPSS

Menu : Statistics Regression
 Option :
 Output :

Tabel 3. Hasil perhitungan untuk koefisien regresi dengan SPSS

Variable	B	SE B	Beta	T	Sig T
X1	,570786	,325377	,349131	1,754	,0955
X2	,386225	,332248	,248765	1,162	,2594
X3	1,787881	1,201046	,263680	1,489	,1530
X4	,130809	,307471	,079052	,425	,6753
(Constant)	-17,597833	37,135920		-,474	,6410

Nilai-nilai estimasi untuk koefisien β_j adalah nilai-nilai yang terlihat pada tabel yaitu nilai-nilai B untuk setiap variabel, yaitu $\beta_0 = -17,597833$, $\beta_1 = 0,570786$, $\beta_2 = 0,386225$, $\beta_3 = 1,787881$ dan $\beta_4 = 0,130809$, atau dengan persamaan regresi selengkapnya adalah :

$$y = -17,597833 + 0,570786x_1 + 0,386225x_2 + 1,787881x_3 + 0,130809x_4$$

Interpretasi model regresi tersebut untuk contoh data sebanyak 24 buah adalah sebagai berikut:

1. Koefisien regresi yang positif pada setiap variabel bebas menunjukkan bahwa daya saing bank akan **semakin tinggi** jika kualifikasi SDM di bidang komputer, tingkat restrukturisasi organisasi, tingkat integrasi teknologi informasi, dan diversifikasi produk perbankan berbasis teknologi informasi **semakin tinggi**,
2. Jika untuk satu variabel bebas berubah sebesar 1 pada saat variabel bebas lainnya tetap maka pengaruhnya terhadap daya saing bank adalah meningkat 0,570786 satuan jika kualifikasi SDM meningkat 1 satuan, meningkat 0,386225 satuan jika tingkat restrukturisasi meningkat 1 satuan, meningkat 1,787881 satuan jika tingkat integrasi teknologi meningkat 1 satuan, dan meningkat sebesar 0,130809 satuan jika diversifikasi produk meningkat 1 satuan.

Jika berdasarkan model regresi untuk contoh data tersebut selanjutnya akan dilakukan generalisasi untuk menyimpulkan kaitan antara daya saing bank dengan variabel-variabel yang mempengaruhinya, maka diperlukan analisis lebih lanjut yang meliputi pendugaan parameter populasi, dan pengujian hipotesis untuk mengukur ketepatannya.

a. Selang Kepercayaan

Kisaran nilai parameter populasi untuk masing-masing β_j untuk tingkat kepercayaan sebesar $100(1 - \alpha)$ persen dapat dihitung dengan menggunakan rumus :

$$\beta_j - t_{\alpha/2, n-p} \sqrt{\sigma^2 C_{jj}} \leq \beta_j \leq \beta_j + t_{\alpha/2, n-p} \sqrt{\sigma^2 C_{jj}}$$

dengan hasil perhitungan interval selengkapnya untuk selang kepercayaan sebesar 95 persen adalah sebagai berikut- $0,110235 \leq \beta_1 \leq 1,251808$

$$-0,309178 \leq \beta_2 \leq 1,081628$$

$$-0,725937 \leq \beta_3 \leq 4,301698$$

$$-0,512735 \leq \beta_4 \leq 0,774353$$

$$-95,324206 \leq \beta_0 \leq 60,128541$$

b. Uji Hipotesis

Uji hipotesis merupakan uji sigifikansi untuk menentukan apakah terdapat hubungan linear antara variabel tak bebas y dengan variabel bebas x1, x2, x3, dan x4. Uji hipotesis ini terdiri dari dua bagian, yaitu (1) uji signifikansi *secara umum* mengenai hubungan linear antara variabel tak bebas dengan variabel bebas dengan menggunakan *statistik uji F* dan (2) uji signifikansi *individual* untuk hubungan variabel tak bebas dengan masing-masing variabel bebas dengan menggunakan *statistik uji t*.

c. Uji F untuk uji signifikansi persamaan regresi

Uji signifikansi yang digunakan untuk menentukan hubungan anatar variabel tak bebas dengan variabel bebas secara umum ini menggunakan formulasi hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 : \beta_j \neq 0 \text{ untuk paling sedikit satu } j$$

statistik uji yang digunakan adalah Uji F dengan rumus perhitungannya adalah;

$$F = \frac{\text{Mean Square of Regresion}}{\text{Mean Square of Error}} = \frac{\text{MSR}}{\text{MSE}}$$

Hasil perhitungan dengan menggunakan perangkat lunak SPSS, nilai F hitungnya adalah sebagai berikut:

-----Output SPSS untuk Nilai F -----

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	4	3500,42009	875,10502
Residual	19	3005,70043	158,19476

F = 5,53182 Signif F = ,0040

Karena *Signif. F-nya relatif kecil* atau *F hitung lebih besar dari $F_{0,05, 4,19}$* ($5,53182 > 2,90$) maka dapat disimpulkan bahwa *ada kaitan linear yang signifikan* antara daya saing bank dengan variabel-variabel bebas yang meliputi kualifikasi SDM, Tingkat strukturisasi organisasi, tingkat integrasi teknologi informasi, dan diversifikasi produk berbasis teknologi informasi.

d. Uji t untuk masing-masing koefisien regresi

Hipotesis untuk menguji signifikansi masing-masing koefisien regresi, yaitu β_j , adalah sebagai berikut:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

Jika $H_0 : \beta_j = 0$ tidak ditolak atau artinya tidak ada kaitan antara satu variabel x_i , maka variabel tersebut dapat dihilangkan dari model regresinya. Statistik ujinya adalah :

$$t_0 = \frac{\beta_j}{\sqrt{\sigma^2 C_{jj}}}$$

Hasil perhitungan dengan menggunakan SPSS adalah :

-----Output SPSS untuk statistik Uji t-----

Variable	B	SE B	Beta	T	Sig T
X1	,570786	,325377	,349131	1,754	,0955
X2	,386225	,332248	,248765	1,162	,2594
X3	1,787881	1,201046	,263680	1,489	,1530
X4	,130809	,307471	,079052	,425	,6753
(Constant)	-17,597833	37,135920		-,474	,6410

Berdasarkan nilai statistik uji t untuk masing-masing koefisien regresi di atas maka dapat disimpulkan sebagai berikut :

1. Berdasarkan nilai signif T-nya maka X_1 relatif paling signifikan hubungan linearnya dengan variabel tak bebas. Urutan berikutnya berdasarkan tingkat signifikansi hubungan linearnya dengan variabel tak bebas adalah X_3 , X_2 , dan terakhir X_4 .
2. Jika menggunakan perbandingan dengan nilai $t_{0,05, 19}$ sebesar 1,729 maka pada $\alpha = 0.10$, hanya *kualifikasi sumber daya manusia (x_1) saja yang menunjukkan adanya kaitan linear yang signifikan dengan daya saing bank (y)*.
3. Jika t hitung untuk masing-masing koefisien regresi dibandingkan dengan nilai $t_{0,1, 19}$ sebesar 1,328 maka pada $\alpha = 0.2$, *kualifikasi SDM (X_1) dan tingkat intergrasi teknologi informasi (X_3) menunjukkan kaitan linear yang signifikan dengan variabel daya saing bank (y), sedangkan tingkat restrukturisasi organisasi (X_2) dan Diversifikasi Produk berbasis teknologi informasi (X_4) tidak signifikan*.

Pengujian signifikansi untuk masing-masing koefisien regresi di atas dapat diringkaskan dalam bentuk Tabel 4 berikut.

Tabel 4. Hasil uji signifikansi untuk masing-masing koefisien regresi

No	Koefisien regresi	Variabel	$\alpha=0.1$	$\alpha=0.2$
1.	β_1	x_1	signifikan	signifikan
2.	β_2	x_2	tidak signifikan	tidaksignifikan
3.	β_3	x_3	tidak signifikan	signifikan
4.	β_4	x_4	tidak signifikan	tidak signifikan

Uji Kecocokan Model

Mengevaluasi kecocokan model merupakan bagian terpenting dari masalah regresi berganda. Beberapa metoda yang digunakan untuk evaluasi tersebut meliputi koefisien determinansi, plot residual, metoda pengukuran residual, dan sebagainya. Penggunaan metoda-metoda tersebut disini akan lebih dititikberatkan pada menilai kecocokkan model, identifikasi data pencilan, dan pengaruh data pencilan serta cara penanganannya.

Kecocokan model regresi dengan menggunakan nilai R^2 dan *Adjusted R*²

Untuk menguji apakah model regresi linear berganda yang telah diformulasikan tersebut cocok untuk contoh data yang dianalisis, digunakan koefisien korelasi (R), koefisien determinan (R^2) dan *adjusted R*². Nilai-nilai tersebut yang dihitung dengan menggunakan SPSS adalah sebagai berikut:

-----Ouput SPSS untuk R, R2, dan adjusted R2 -----

Multiple R ,73350
R Square ,53802
Adjusted R Square ,44076
Standard Error 12,57755

R dengan nilai sebesar 0,73 merupakan koefisien korelasi berganda antara variabel tak bebas y dengan himpunan variabel bebas x_1, x_2, \dots, x_k yang menunjukkan ukuran hubungan linear diantara kedua jenis variabel tersebut. Nilai R^2 sebesar 0.53802 menunjukkan bahwa sekitar 53,8 persen keragaman variabel tak bebas dapat diterangkan oleh keragaman variabel-variabel bebasnya. Atau jika dihubungkan dengan jenis variabel yang diamati, keragaman daya saing bank hanya bisa diterangkan 53,8 persen oleh keragaman variabel kulaifikasi SDM, Tingkat strukturisasi organisasi, tingkat integrasi teknologi informasi, dan diversifikasi produk berbasis teknologi informasi.

Nilai *adjusted* R^2 sebesar 0,44076 menunjukkan cara interpretasi yang sama tetapi dengan persentase yang lebih kecil yang secara umum menunjukkan bahwa model tersebut relatif *over specified* atau terlalu banyak variabel bebas yang kaitan linearnya dengan variabel tak bebas relatif kurang signifikan. Pertanyaannya adalah kira-kira variabel mana yang bisa dieliminasi sehingga model menjadi lebih sederhana dan bisa menunjukkan kecocokan yang lebih baik? Hal ini bisa dikaitkan dengan hasil pengujian hipotesis sebelumnya.

Berdasarkan ukuran-ukuran di atas maka dapat disimpulkan bahwa model regresi berganda yang sudah diformulasikan tersebut *kurang cocok* digunakan untuk melihat kaitan antara variabel bebas y dan 4 buah variabel tak bebas X_i . Jika dikaitkan dengan hasil pengujian hipotesis individual untuk masing-masing koefisien regresi yang pada $\alpha = 0.1$ hanya menunjukkan X_1 saja yang signifikan mempunyai kaitan linear dengan variabel tak bebas y, maka nilai *adjusted* R^2 sebesar 0,44076 yang relatif lebih kecil dibandingkan R^2 disebabkan penambahan variabel X_2 dan X_4 yang tidak signifikan kaitan linearnya terhadap y.

Pengaruh data pencilan (Outlier)

Data pencilan pada dasarnya merupakan data individual yang nilainya relatif ekstrim terhadap variasi umum keseluruhan data. Metoda yang digunakan untuk mengidentifikasi adanya data pencilan adalah dengan mengukur *standardized residuals* (d_i) dan *studentized residuals* (r_i) dengan rumus sebagai berikut :

$$d_i = \frac{e_i}{\sqrt{MS_E}}, \quad i = 1, 2, \dots, n$$

dan

$$r_i = \frac{e_i}{\sqrt{MS_E(1 - h_{ii})}}, \quad i = 1, 2, \dots, n$$

Data pencilan adalah data yang nilai d_i atau r_i -nya lebih besar dari 2 atau lebih kecil dari -2. Nilai-nilai d_i dan r_i untuk masing-masing item data selengkapnya dapat dilihat pada Tabel 5 berikut ini.

Tabel 5. Nilai-nilai residual, standarized dan studentized residual

	Y	\hat{y}	RESIDU	Z_RES	S_RES	KETERANGAN
1	136.673	127.62843	9,04457	,71910	,97823	
2	129.085	141,29621	-12,21121	-,97087	-1,07036	
3	156.384	144,08180	12,30220	,97811	1,10955	
4	117.553	139,14153	-21,58853	-1,71643	-1,98090	
5	140.681	138,70998	1,97102	,15671	,17333	
6	154.653	136,72158	17,93142	1,42567	1,55799	
7	160.000	151,29421	8,70579	,69217	,74977	
8	118.320	143,95499	-25,63499	-2,03815	-2,13710	Data Outlier
9	98.400	95,05459	3,34541	,26598	,44233	
10	137.100	136,36168	,73832	,05870	,07070	
11	160.355	149,86012	10,49488	,83441	1,03020	
12	137.780	147,90218	-10,12218	-,80478	-,93233	
13	135.881	128,85816	7,02284	,55836	,60344	
14	134.872	133,64093	1,23107	,09788	,10512	
15	131.937	141,72719	-9,79019	-,77839	-,83553	
16	110.000	120,92708	-10,92708	-,86878	-,95355	
17	156.323	141,77878	14,54422	1,15636	1,28708	
18	118.014	126,84685	-8,83285	-,70227	-,79314	
19	130.084	139,41114	-9,32714	-,74157	-,76374	
20	142.554	131,09633	11,45767	,91096	,94378	
21	159.621	151,18207	8,43893	,67095	,73946	
22	154.118	152,35661	1,76139	,14004	,15729	
23	132.008	128,66038	3,34762	,26616	,27741	
24	130.000	133,90319	-3,90319	-,31033	-,34332	

Keterangan :

- Y : Data observasi
- \hat{y} : Hasil regresi
- Residu : Y - \hat{y}
- Z_res : Standarized residual
- S_res : Studentized residual

Berdasarkan tabel tersebut maka *data ke 8 adalah data pencilan (outlier) karena nilai standarized residualnya lebih dari -2, yaitu -2,03815 dengan nilai studentized residualnya -2,13710.*

Pengaruh data pencilan terhadap parameter model dapat diketahui dengan menggunakan ukuran *Cook's Distance (D_i)* dan *DFBETA_{j,i}* dengan rumus sebagai berikut:

$$D_i = \frac{(\beta_{(i)} - \beta)' X'X (\beta_{(i)} - \beta)}{pMS_E}$$

$$DFBETA_{j,i} = \frac{\beta_j - \beta_{j(i)}}{pMS_E}$$

$$\sqrt{S^2_{(i)}C_{ij}}$$

Nilai *Cook's Distance* (D_i) digunakan untuk melihat pengaruh data pencilan terhadap model regresi secara keseluruhan. Jika Nilai $D_i > F_{0.5,p,n-p}$ maka data pencilan tersebut berpengaruh terhadap model regresinya. Sebaliknya jika nilai D_i relatif kecil dan mendekati nilai $F_{0.5,p,n-p}$, nilai $F_{0.5,p,n-p} \approx 1$, maka data pencilan tersebut relatif tidak berpengaruh terhadap persamaan regresinya. Sedangkan nilai $DFBETA_{j,i}$ digunakan untuk melihat pengaruh data pencilan ke i terhadap masing-masing parameter β_j . Jika nilai $|DFBETA_{j,i}| > 2/\sqrt{n}$ maka data pencilan tersebut berpengaruh terhadap koefisien regresi β_j . Nilai D_i dan $DFBETA_{j,i}$ selengkapnya untuk masing-masing item data dapat dilihat pada Tabel 6 berikut.

Tabel 6. Nilai *Cook's Distance* (D_i) dan $DFBETA_{j,i}$

	D_i	$DFBETA_{0,i}$	$DFBETA_{1,i}$	$DFBETA_{2,i}$	$DFBETA_{3,i}$	$DFBETA_{4,i}$
1	,16278	10,84949	,02705	,14313	-,34956	-,24215
2	,04937	6,02362	,01341	-,02289	-,38116	,08568
3	,07062	-7,41534	-,08622	,03857	,52450	-,01797
4	,26047	1,24062	-,16122	,18255	,44149	-,21029
5	,00134	-,97090	-,00418	-,01502	,07014	,01108
6	,09430	8,50574	-,10015	,06816	-,33984	,13824
7	,01949	-4,76041	,05068	,01054	-,08859	,02374
8	,09085	1,47433	,04588	-,10431	,17067	-,07638
9	,06909	13,47863	-,04032	-,10752	,02062	-,01197
10	,00045	-,11734	,01274	-,00820	-,01783	-,00104
11	,11130	-17,02343	,20460	-,13612	,31124	-,05581
12	,05947	8,50403	-,13047	-,00293	-,03626	,11326
13	,01223	3,26567	,00144	-,04169	-,08491	,04714
14	,00034	,82453	-,00384	,00697	-,03546	,00264
15	,02125	3,36386	,04582	,01041	-,23830	-,04522
16	,03722	-13,13860	,03319	-,04257	,33101	,04275
17	,07914	5,57283	-,04319	,11070	-,45584	,07953
18	,03467	-10,47384	,09451	-,09127	,22866	,00115
19	,00708	-,18169	,02180	-,01946	-,00597	-,01552
20	,01307	5,40455	-,00499	,00787	-,18050	,00660
21	,02347	-8,04794	-,00793	-,01760	,27288	,04541
22	,00129	-1,19406	,00130	,01724	,01646	-,01227
23	,00133	,83242	-,00048	-,00966	,02015	-,00781
24	,00528	,69724	,02340	,01279	-,14652	-,00253

Pada Tabel 4 diatas, terlihat bahwa data ke 8 yang merupakan data pencilan mempunyai nilai D_i relatif kecil yaitu sebesar 0,09085 atau lebih kecil dari 1. Hal itu menunjukkan bahwa *data pencilan tersebut secara umum tidak berpengaruh terhadap persamaan regresinya*. Untuk melihat pengaruh untuk masing-masing β_j , terlihat bahwa nilai-nilai

DFBETA_{j,i}-nya relatif kecil dibandingkan $2/\sqrt{24}$, kecuali $DFBETA_{0,8}$ sebesar 1,47433 yang lebih besar dari nilai pembanding tersebut. Jadi kesimpulannya, *data pencilan ke 8 tersebut relatif tidak mempengaruhi β_1 , β_2 , β_3 , dan β_4 , dan hanya mempengaruhi konstanta (β_1) persamaan regresinya.*

Multikolinearitas

Pengujian multikolinearitas bertujuan untuk melihat apakah ke 4 variabel bebas tidak berkorelasi satu sama lain (atau saling bebas). Multikolinearitas dapat dideteksi dengan pengukuran matrix korelasi atau dengan analisis eigensystem yang mengukur *condition indices* dari matrix $X'X$. Matrix korelasinya adalah sebagai berikut:

-----Output SPSS untuk matrix korelasi -----
Correlation, 1-tailed Sig:

	Y	X1	X2	X3	X4
Y	1,000	,616	,614	,514	,436
X1	,616	1,000	,580	,324	,474
X2	,614	,580	1,000	,470	,492
X3	,514	,324	,470	1,000	,262
X4	,436	,474	,492	,262	1,000

Dari matrix korelasi tersebut dapat dilihat bahwa hubungan antara variabel bebas X_i adalah relatif lemah yang ditunjukkan oleh p-value yang relatif kecil. Atau dapat disimpulkan bahwa tidak ada multikolinearitas. Cara lain untuk melihat multikolinearitas adalah dengan mengukur *condition indices* pada sistem Eigen. Jika nilai index tersebut kecil (lebih kecil dari 100), maka tidak terdapat multikolinearitas. Nilai-nilai index untuk setiap variabel bebas pada contoh data ini adalah relatif kecil yaitu kurang dari 100 seperti terlihat di hasil perhitungan dengan SPSS. Dengan tidak adanya multikolinearitas ini diharapkan variansi jumlah kuadrat untuk estimasi nilai β_j relatif kecil sehingga presisinya tinggi.

Kolinieritas pada setiap variabel dalam analisis regresi linier berganda dengan SPSS ditandai dengan beberapa indikator berikut:

- nilai signifikansi pada uji t bernilai > 0.05 (atau $>$ nilai α lainnya)
- nilai Variation Inflation Factor (VIF) > 5 di mana $VIF = \frac{1}{TOLERANCE}$ dan $TOLERANCE = 1 - R^2$.
- Eigen Value mendekati 0 (nol)

- nilai Condition Index > 15 (terdeteksi adanya kolinieritas), Condition Index > 30 (kolinieritas serius)

Salah satu cara penanganan kolinieritas adalah dengan mengeliminasi variabel yang terbukti memiliki kolinieritas. SPSS menyediakan fasilitas eliminasi variabel dalam menu METHOD. Pilihan BACKWARD pada METHOD akan memberikan beberapa model persamaan regresi berganda. Model terakhir adalah model yang kolinieritasnya sudah minimal. Model ini dapat memberikan gambaran variabel-variabel yang paling berpengaruh.

Interpretasi

Walaupun secara umum ada kaitan antara variabel daya saing bank (y) dengan variabel bebas yang terdiri dari kualifikasi SDM (x1), restrukturisasi organisasi (x2), integrasi teknologi informasi (x3), dan diversifikasi jasa berbasis TI (x4), secara individual hanya variabel kualifikasi SDM dan tingkat integrasi teknologi informasi saja yang menunjukkan hubungan linear yang signifikan. Hubungan linear antara variabel tak bebas dan variabel bebasnya adalah positif, yaitu jika kualifikasi SDM dan tingkat integrasi teknologi semakin tinggi maka daya saing bank akan semakin tinggi.

Model regresi linear berganda yang digunakan untuk menganalisis hubungan daya saing dengan faktor-faktor yang terkait, secara umum kurang cocok untuk menggambarkan bentuk kaitan yang tepat sesuai dengan contoh data yang digunakan. Selain berdasarkan hasil pengujian hipotesisnya, juga dihubungkan dengan nilai koefisien determinan yang menunjukkan bahwa keragaman daya saing bank tidak terlalu banyak diterangkan oleh keragaman variabel bebasnya, serta nilai *adjusted R²* yang jauh lebih kecil dibandingkan *R²* menunjukkan bahwa pada model regresi ada beberapa variabel yang dimasukkan ke dalam model tersebut relatif lemah kaitan linearnya dengan daya saing bank (atau modelnya *overspecified*).

Soal 7.2

Data yang akan diolah dengan menggunakan analisis regresi linier berganda terdiri dari 5 buah variabel, yaitu transformasi kewirausahaan (X_1), kompetisi (X_2), sistem imbalan dan hukuman (X_3), fungsi-fungsi manajemen (X_4), dan kinerja perusahaan (Y) dari 100 perusahaan. Data selengkapnya dapat dilihat pada Tabel 4. Lakukan pengujian kenormalan distribusi data dan kolinieritas, hingga didapat variabel-variabel yang paling berpengaruh terhadap kinerja perusahaan (Y)

Tabel 4. Data transformasi kewirausahaan, kompetisi, sistem imbalan dan hukuman, fungsi-fungsi manajemen, dan kinerja 100 perusahaan.

NO.	Y	X ₁	X ₂	X ₃	X ₄
1.	106,430	27,302	33,407	85,806	81,023
2.	108,640	34,503	49,550	85,102	64,127
3.	105,610	33,856	57,636	88,874	86,293

4.	119,870	27,829	72,004	89,596	75,461
5.	107,520	32,344	64,077	81,300	74,381
6.	107,940	28,439	70,027	84,838	84,675
7.	124,860	30,036	69,616	82,950	90,150
8.	113,730	29,988	66,741	84,001	88,795
9.	78,704	25,584	55,110	65,795	59,130
10.	123,540	27,202	57,661	100,250	89,806
11.	128,440	22,048	57,131	86,659	76,973
12.	124,940	30,665	50,286	96,823	84,874
13.	107,700	27,634	64,424	86,597	96,528
14.	108,350	27,329	62,126	89,400	83,910
15.	107,570	32,227	65,430	92,899	82,204
16.	100,440	26,075	48,082	83,390	73,246
17.	113,770	28,087	70,027	85,322	90,778
18.	88,512	27,538	55,662	80,817	82,980
19.	109,710	30,144	62,227	86,074	83,772
20.	108,030	27,781	57,850	87,646	77,143
21.	115,920	33,463	70,027	87,139	87,070
22.	118,320	27,055	73,106	91,972	96,701
23.	104,580	29,329	50,755	83,277	71,182
24.	100,770	32,380	55,474	82,944	74,661
25.	119,380	27,273	59,701	80,208	85,471
26.	113,180	22,998	58,642	108,580	87,064
27.	126,340	28,440	55,553	88,608	67,306
28.	117,630	29,160	57,600	88,738	88,179
29.	113,380	30,913	70,027	78,516	95,944
30.	121,570	28,297	52,771	93,298	70,609
31.	124,200	30,589	54,417	86,514	84,280
32.	120,730	29,163	70,027	99,116	92,720
33.	114,720	38,171	59,514	91,413	89,156
34.	108,260	32,718	41,672	82,693	81,553
35.	105,120	32,793	56,378	84,490	74,177
36.	118,050	30,198	63,485	85,065	87,503
37.	118,780	32,089	70,027	96,817	97,760
38.	113,900	29,619	66,054	88,818	84,889
39.	111,710	37,515	57,968	83,314	84,855
40.	115,990	33,646	68,159	86,671	76,138
41.	119,180	28,978	65,670	76,915	95,460
42.	134,860	27,196	91,054	95,935	112,590
43.	121,350	28,640	55,512	89,838	69,889
44.	113,740	24,877	80,991	83,670	91,044
45.	84,243	29,622	42,499	79,092	38,359
46.	120,200	33,452	70,027	88,704	90,897

47.	146,430	25,409	85,184	99,610	103,700
48.	165,560	31,871	95,105	114,860	128,130
49.	114,150	30,593	70,027	89,818	92,229
50.	103,770	29,598	61,231	75,529	77,224
51.	85,237	29,622	42,499	79,092	43,619
52.	140,810	27,103	44,030	93,959	62,719
53.	119,370	30,541	58,562	92,990	70,418
54.	133,900	27,062	62,158	102,860	87,976
55.	134,370	27,962	81,665	93,094	103,090
56.	123,690	28,873	53,361	93,122	91,024
57.	111,560	30,728	57,737	85,559	91,024
58.	105,310	34,841	65,479	79,151	77,391
59.	104,760	31,234	57,966	83,105	77,801
60.	107,840	27,274	64,855	76,005	78,531
61.	108,070	33,641	66,675	86,916	83,393
62.	123,280	39,197	64,663	87,344	89,678
63.	103,680	29,195	53,147	85,066	88,454
64.	115,390	25,819	80,911	95,900	94,224
65.	113,210	35,193	63,866	98,519	82,331
66.	97,088	27,969	33,749	82,755	43,958
67.	111,740	31,813	70,027	80,529	91,596
68.	115,340	28,785	67,785	87,448	85,884
69.	110,710	31,701	67,989	84,768	88,834
70.	107,080	24,733	57,561	89,634	66,572
71.	105,370	30,152	48,913	72,897	71,615
72.	113,790	33,860	61,113	103,640	100,740
73.	109,290	24,639	55,207	87,051	76,355
74.	119,380	32,155	70,027	87,337	92,720
75.	107,510	27,976	42,144	89,444	77,900
76.	104,470	31,622	59,233	89,061	87,777
77.	111,460	32,547	52,577	90,548	73,461
78.	124,580	33,594	55,163	95,020	77,428
79.	105,360	28,691	57,483	84,073	71,665
80.	114,050	26,267	49,752	94,566	68,016
81.	116,850	34,356	66,397	90,963	87,369
82.	109,500	25,678	36,398	86,936	78,475
83.	122,330	23,945	50,555	86,291	61,523
84.	113,180	33,255	60,921	99,612	98,540
85.	113,460	27,880	60,295	91,709	86,601
86.	107,910	32,648	55,567	81,704	67,854
87.	118,840	43,235	70,131	101,780	99,686
88.	113,710	27,763	66,735	86,744	87,163
89.	110,820	29,983	70,027	78,187	89,054

90.	122,110	31,528	62,547	90,855	78,836
91.	121,970	32,663	70,027	94,287	102,030
92.	118,440	26,594	71,340	90,410	92,720
93.	112,640	34,831	59,112	93,550	83,508
94.	127,410	27,847	67,785	104,920	87,603
95.	115,870	26,575	67,888	84,282	85,767
96.	116,920	34,206	64,753	88,024	82,772
97.	136,450	33,503	88,163	100,460	111,950
98.	115,740	28,338	63,453	94,005	82,689
99.	111,720	28,042	52,180	85,487	64,770
100.	107,310	22,707	49,821	87,985	95,559

