

Analisis Data Eksploratif & Statistika Inferensia Dasar Pertemuan 13

Korelasi

Korelasi

Penelitian ilmiah adalah aktivitas yang bertujuan untuk mengetahui hubungan antara satu fenomena dengan fenomena yang lain. Hasil pengukuran terhadap fenomena yang diteliti dalam penelitian ilmiah disebut variabel penelitian. Karena itu hubungan antara satu fenomena dengan fenomena yang lain disebut juga dengan hubungan antara satu variabel dengan variabel lain. Hubungan antara satu variabel dengan variabel lain disebut korelasi. Dengan kata lain korelasi merupakan istilah yang digunakan untuk mengukur kekuatan hubungan antar variabel.

Analisa korelasi pada garis besarnya dibedakan kepada dua, yaitu:

1. Korelasi antara dua variabel (*bivariate correlation*) misalnya hubungan antara rambu-rambu lalu lintas dengan kecelakaan lalu lintas.
2. Korelasi antara tiga variabel atau lebih (*multivariate correlation*) misalnya hubungan antara kecerdasan dan motif belajar dengan prestasi belajar.

Penggunaan Korelasi

Statistik mempersiapkan bermacam-macam teknik analisa korelasi yang dapat digunakan untuk mengetahui apakah terdapat hubungan antara dua variabel. Sekalipun hasil menunjukkan adanya hubungan antara dua variabel atau lebih, peneliti tidak bisa secara langsung menarik kesimpulan bahwa antar variabel tersebut bersifat sebab-akibat (kausal). Menurut Walter R. Bog dan Meredith D, Gall, penelitian korelasi tidak dapat digunakan untuk meneliti hubungan sebab akibat antara satu variabel dengan variabel yang lain.

Arah Korelasi

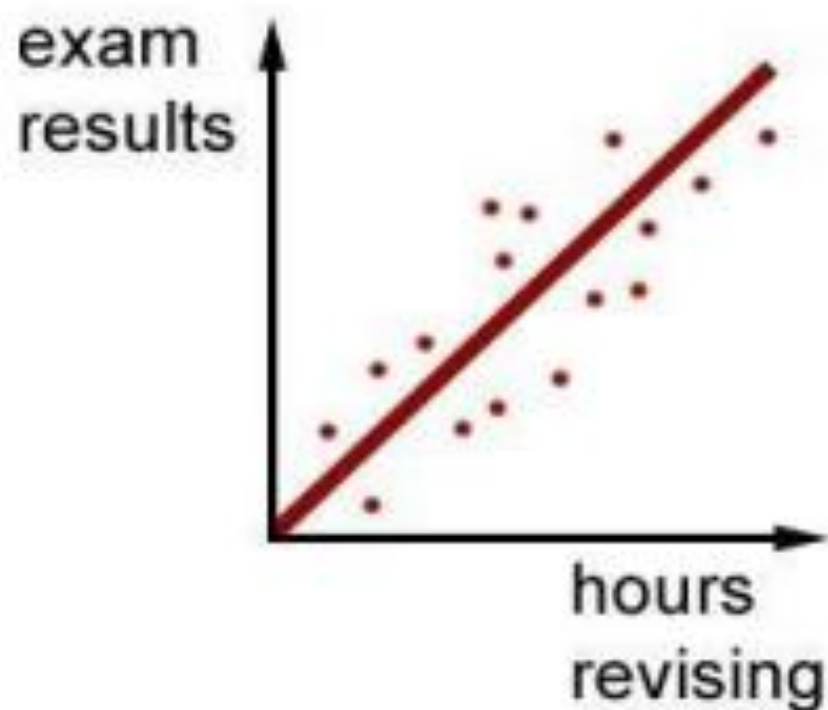
1. Korelasi searah atau korelasi positif (+)

Yaitu apabila dua variabel atau lebih berkorelasi secara paralel, kenaikan satu variabel disertai dengan kenaikan pada variabel yang lain atau penurunan nilai satu variabel disertai dengan penurunan pada variabel yang lain. Misalnya, kenaikan harga bahan bakar minyak (BBM), diikuti kenaikan ongkos-ongkos angkutan. Sebaliknya jika harga BBM turun, maka ongkos angkutan menjadi turun. Dalam dunia pendidikan misalnya, terdapat korelasi positif antara nilai-nilai belajar matematika dengan nilai hasil belajar Fisika, Kimia, Biologi dan sebagainya.

Arah Korelasi (2)

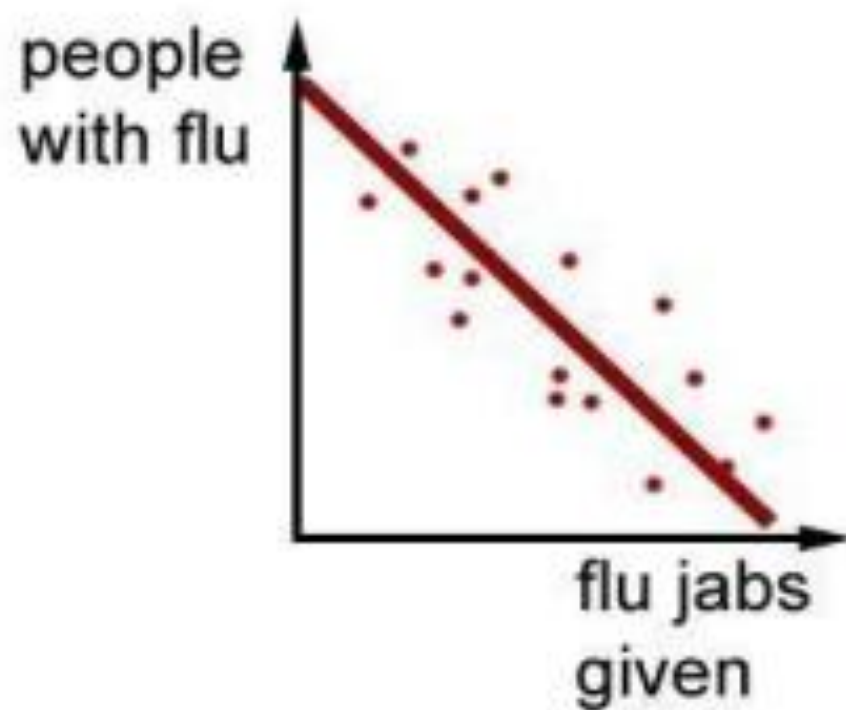
2. Korelasi berlawanan arah atau korelasi negatif (-)

Yaitu apabila dua variabel atau lebih berkorelasi secara berlawanan arah, kenaikan nilai satu variabel disertai dengan penurunan nilai variabel lainnya. Misalnya meningkatnya kesadaran hukum dikalangan masyarakat diikuti dengan menurunnya jumlah kejahatan yang dilakukan anggota masyarakat, atau sebaliknya. Penurunan hasil belajar siswa dalam bidang studi seni suara disertai dengan peningkatan hasil belajar bidang studi matematika, kimia, biologi, dan sebagainya.



POSITIVE CORRELATION

- people who do more revision get higher exam results.
- revising increases success.



NEGATIVE CORRELATION

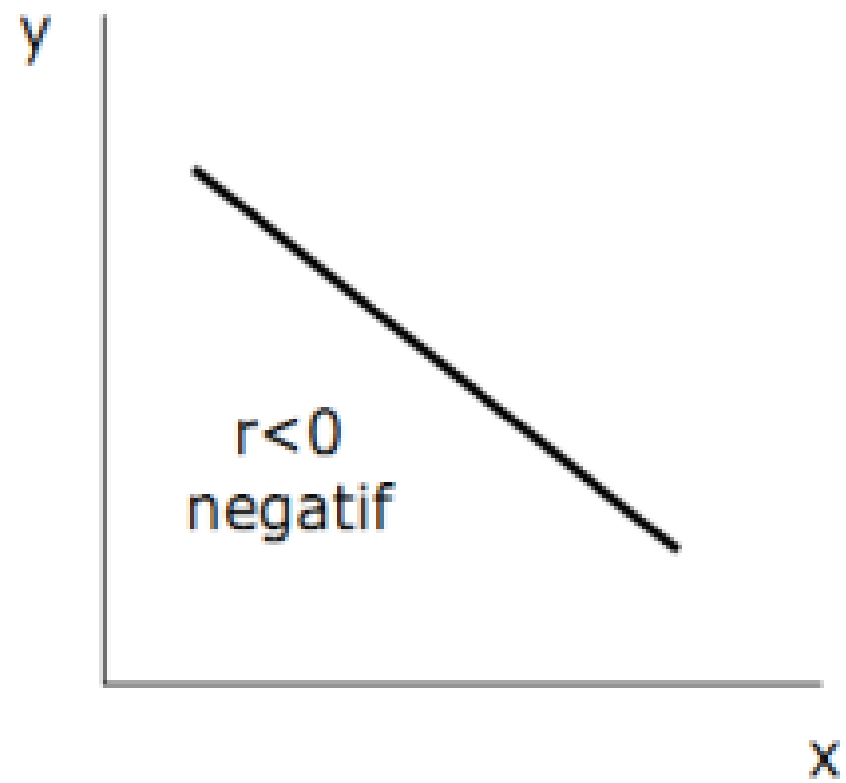
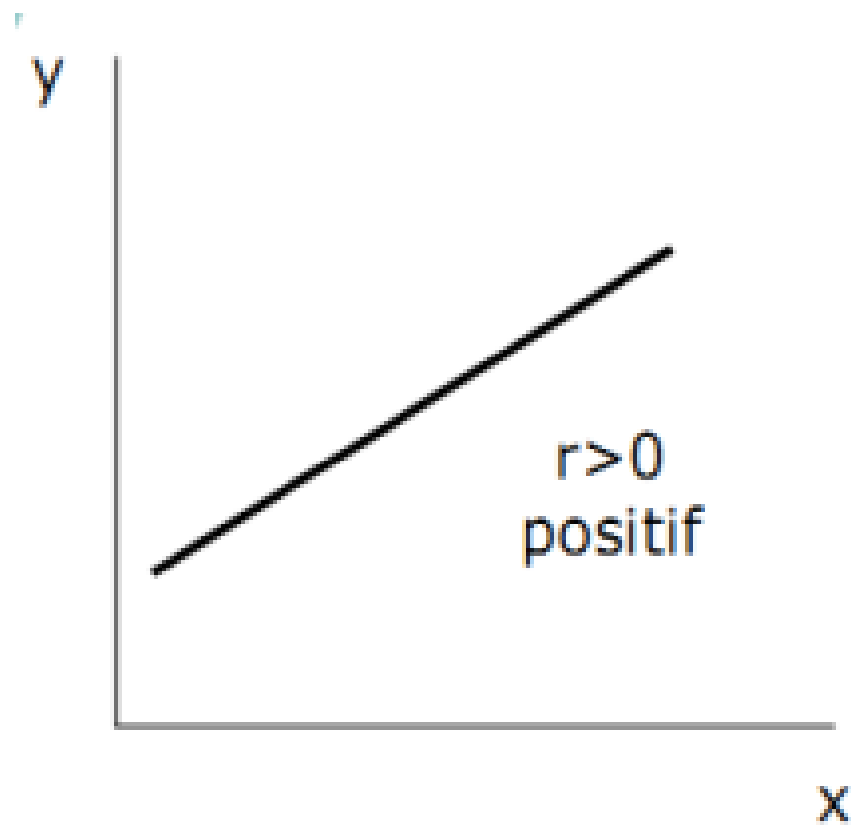
- when more jabs are given the number of people with flu falls.
- flu jabs prevent flu.

Koefisien Korelasi

Angka korelasi atau *coeffcient of correlation* adalah angka (koefisien) yang dapat dijadikan sebagai petunjuk untuk mengetahui tinggi rendahnya, kuat lemahnya atau besar kecilnya korelasi antara variabel yang sedang diselidiki. Besarnya angka korelasi berkisar antara 0 (nol) sampai dengan $\pm 1,00$. Ini berarti bahwa angka korelasi tidak bisa lebih dari $+1,00$. Apabila dalam suatu perhitungan diperoleh angka korelasi lebih besar dari 1,00 hal itu merupakan petunjuk adanya kesalahan dalam perhitungan.

Koefisien Korelasi

Tanda + (plus) atau - (minus) di depan angka korelasi merupakan petunjuk arah korelasi. Bila di depan angka korelasi bertanda plus (+), misalnya; $r_{xy} = +0,325$ ini berarti bahwa korelasi antara x dan y merupakan korelasi searah atau korelasi positif. Sebaliknya apabila di depan angka korelasi bertanda minus (-) misalnya; $r_{xy} = -0,785$ ini berarti korelasi antara variabel x dan y merupakan korelasi berlawanan arah atau korelasi negatif.



Jenis Korelasi

Pearson

- Data Interval/Rasio (Numerik)

Rank
Spearman

- Data Ordinal

Kontingensi

- Data Nominal

Koefisien
Determinasi

- Pengaruh Antar Variabel

Derajat Kekuatan/ Keeratan Hubungan

Menurut Sugiyono (2007) :

0,00 - 0,199 = Sangat rendah

0,20 - 0,399 = Rendah

0,40 - 0,599 = Sedang

0,60 - 0,799 = Kuat

0,80 - 1,000 = Sangat kuat

Korelasi Product Moment /Pearson

$$r_{xy} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

Contoh

Diberikan data tinggi badan (X) dan tinggi lompatan (Y) seperti table berikut.

No	Tinggi Badan (X)	Tinggi Lompatan (Y)
1	168	154
2	165	151
3	166	152
4	170	155
5	161	151
6	162	152
7	160	148
8	165	155
9	165	156
10	162	150

Tentukan derajat hubungan
antara Tinggi Badan dengan
Tinggi Lompatan!

Penyelesaian

$$\sum X = 1644, \sum Y = 1524$$

$$\sum X^2 = 270364$$

$$\sum Y^2 = 232316$$

$$\sum XY = 250599$$

No	Tinggi Badan (X)	Tinggi Lompatan (Y)	X ²	Y ²	XY
1	168	154	28224	23716	25872
2	165	151	27225	22801	24915
3	166	152	27556	23104	25232
4	170	155	28900	24025	26350
5	161	151	25921	22801	24311
6	162	152	26244	23104	24624
7	160	148	25600	21904	23680
8	165	155	27225	24025	25575
9	165	156	27225	24336	25740
10	162	150	26244	22500	24300
Σ	1644	1524	270364	232316	250599

Koefisien Korelasi :

$$r_{XY} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}} = \frac{10 \cdot 250599 - (1644)(1524)}{\sqrt{(10 \cdot 270364 - (1644)^2)(10 \cdot 232316 - (1524)^2)}} = 0,7349$$

Jadi, terdapat hubungan yang kuat antara Tinggi Badan dengan Tinggi lompatan dengan koefisien korelasinya 0,7349.

Uji Signifikansi Koefisien Korelasi Pearson

1 **Hipotesis :**

$H_0 : \rho_{XY} = 0$ (tidak ada hubungan yang signifikan antara X dan Y)

$H_1 : \rho_{XY} \neq 0$ (ada hubungan yang signifikan antara X dan Y)

2 **Tingkat Signifikansi : α**

3 **Statistik Uji :**

$$t = \frac{r_{XY}\sqrt{n-2}}{\sqrt{1-r_{XY}^2}}$$

4 **Kriteria Penolakan :**

H_0 ditolak jika $t_{hitung} > t_{\alpha,df}$ atau $t_{hitung} < -t_{\alpha,df}$ dengan $df = n - 2$

5 **Kesimpulan**

Contoh (Gunakan Contoh Sebelumnya)

1 **Hipotesis :**

$H_0 : \rho_{XY} = 0$ (tidak ada korelasi yang signifikan antara X dan Y)

$H_1 : \rho_{XY} \neq 0$ (ada korelasi yang signifikan antara X dan Y)

2 **Tingkat Signifikansi :** $\alpha = 0,05$

3 **Statistik Uji :**

$$t = \frac{r_{XY}\sqrt{n-2}}{\sqrt{1-r_{XY}^2}} = \frac{0,7349\sqrt{10-2}}{\sqrt{1-(0,7349)^2}} = 3,065$$

4 **Kriteria Penolakan :**

H_0 ditolak jika $t_{hitung} > t_{\alpha,df}$. Karena $t_{hitung} = 3,065 > 2,306 = t_{0,05;8}$ maka H_0 ditolak.

5 **Kesimpulan :** Jadi, ada korelasi signifikan antara tinggi badan dengan tinggi lompatan

t Table

cum. prob	<i>t</i> _{.50}	<i>t</i> _{.75}	<i>t</i> _{.80}	<i>t</i> _{.85}	<i>t</i> _{.90}	<i>t</i> _{.95}	<i>t</i> _{.975}	<i>t</i> _{.99}	<i>t</i> _{.995}	<i>t</i> _{.999}	<i>t</i> _{.9995}
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646

Latihan

Hasil penjualan tahunan dan biaya iklan suatu perusahaan disajikan dalam table berikut :

No	Penjualan (X)	Biaya Iklan (Y)
1	2100	100
2	2200	150
3	2000	100
4	2300	200
5	2400	400
6	2500	500
7	2600	600
8	2725	700
9	2900	800
10	3000	900

- Tentukan koefisien korelasinya
- Seberapa kuatkah perubahan hasil penjualan tiap tahun dipengaruhi oleh biaya iklan yang dikeluarkan
- Apakah ada hubungan yang signifikan antara besarnya penjualan dengan biaya iklan yang dikeluarkan? Gunakan $\alpha = 0,05$

Koefisien Rank Spearman

- Digunakan untuk mengukur derajat hubungan antara 2 variable yang datanya berskala ordinal (bertingkat).
- Digunakan untuk mengurutkan objek/tingkatan dari yang paling rendah sampai yang paling tinggi atau sebaliknya
- Contoh : peringkat kelas, skala likert,dsb

Rumus Korelasi Rank Spearman

- Jika Tidak Ada Nilai Pengamatan Yang Sama

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

r_s : Koefisien Korelasi Rank Spearman

n : Jumlah Sampel

d : selisih rangking X dan Y

- Jika Ada Nilai Pengamatan Yang Sama

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum d_i^2}{2\sqrt{\sum x^2 \sum y^2}}$$

$$\sum x^2 = \frac{n^3 - n}{12} - \sum T_X$$

$$\sum y^2 = \frac{n^3 - n}{12} - \sum T_Y$$

$$T_X = T_Y = \frac{t^3 - t}{12}, \text{ dengan } t \text{ adalah banyaknya nilai observasi yang sama}$$

Contoh (Data Tidak Ada Yang Sama)

Diberikan data yang menggambarkan pengalaman usaha tani dengan penerapan teknologi sbb :

No	Pengalaman Usaha Tani	Penerapan Teknologi
1	12	88
2	15	74
3	19	78
4	13	90
5	9	85
6	14	87
7	16	94
8	18	98
9	11	81
10	17	91
11	10	76
12	8	70

Penyelesaian

No	X	Y	Rank X	Rank Y	$d = \text{Rank X} - \text{Rank Y}$	d^2
1	12	88	5	8	-3	9
2	15	74	8	2	6	36
3	19	78	12	4	8	64
4	13	90	6	9	-3	9
5	9	85	2	6	-4	16
6	14	87	7	7	0	0
7	16	94	9	11	-2	4
8	18	98	11	12	-1	1
9	11	81	4	5	-1	1
10	17	91	10	10	0	0
11	10	76	3	3	0	0
12	8	70	1	1	0	0
Σ						140

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

$$= 1 - \frac{6(140)}{12^3 - 12}$$

$$= 1 - \frac{840}{1716} = 1 - 0,4895$$

$$= 0,5105$$

Karena $r_s = 0,5105$ maka tingkat keeratannya adalah sedang

Uji Signifikansi Koefisien Korelasi Rank Spearman

1 **Hipotesis :**

$H_0 : \rho_S = 0$ (tidak ada hubungan Yang signifikan antara X dan Y)

$H_1 : \rho_S \neq 0$ (ada hubungan yang signifikan antara X dan Y)

2 **Tingkat Signifikansi : α**

3 **Statistik Uji :**

$$t = \frac{r_S \sqrt{n - 2}}{\sqrt{1 - r_S^2}}$$

4 **Kriteria Penolakan :**

H_0 ditolak jika $t_{hitung} > t_{\alpha, df}$ atau $t_{hitung} < -t_{\alpha, df}$ dengan $df = n - 2$

5 **Kesimpulan**

Contoh (Dengan Contoh Sebelumnya)

1 **Hipotesis :**

$H_0 : \rho_S = 0$ (tidak ada hubungan Yang signifikan antara X dan Y)

$H_1 : \rho_S \neq 0$ (ada hubungan yang signifikan antara X dan Y)

2 **Tingkat Signifikansi : $\alpha = 5\%$**

3 **Statistik Uji :**

$$t = \frac{r_S \sqrt{n-2}}{\sqrt{1-r_S^2}} = \frac{0,5105 \sqrt{12-2}}{\sqrt{1-(0,5105)^2}} = 1,8774$$

4 **Kriteria Penolakan :**

H_0 ditolak jika $t_{hitung} > t_{\alpha,df}$ dengan $df = n - 2 = 10$. Karena $t_{hitung} = 1,8774 < 2,228 = t_{0,05;10}$ maka H_0 tidak ditolak.

5 **Kesimpulan :** Jadi, tidak ada hubungan yang signifikan antara pengalaman usaha tani dengan penggunaan teknologi.

Contoh (Data Observasi Ada Yang Sama)

Diberikan data yang menggambarkan pengalaman usaha tani dengan penerapan teknologi sbb :

No	Pengalaman Usaha Tani	Penerapan Teknologi
1	12	85
2	10	74
3	10	78
4	13	90
5	11	85
6	14	87
7	13	94
8	14	98
9	11	81
10	14	91
11	10	76
12	8	74

Penyelesaian

No	X	Y	$Rank\ X$	$Rank\ Y$	$d=Rank\ X- Rank\ Y$	d^2
1	12	85	7	6.5	0.5	0.25
2	10	74	3	1.5	1.5	2.25
3	10	78	3	4	-1	1
4	13	90	8.5	9	-0.5	0.25
5	11	85	5.5	6.5	-1	1
6	14	87	11	8	3	9
7	13	94	8.5	11	-2.5	6.25
8	14	98	11	12	-1	1
9	11	81	5.5	5	0.5	0.25
10	14	91	11	10	1	1
11	10	76	3	3	0	0
12	8	74	1	1.5	-0.5	0.25
						22.5

<i>Rank X</i>	<i>t</i>	T_X	<i>Rank Y</i>	<i>t</i>	T_Y
3	3	2	1.5	2	0.5
5.5	2	0.5	6.5	2	0.5
8.5	2	0.5			
11	3	2			
Σ		5	Σ		1

$$\sum X^2 = \frac{n^3 - n}{12} - \sum T_X = \frac{12^3 - 12}{12} - 5 = 138$$

$$\sum Y^2 = \frac{n^3 - n}{12} - \sum T_Y = \frac{12^3 - 12}{12} - 1 = 142$$

$$T_X = \frac{t^3 - t}{12} = \frac{3^3 - 3}{12} = \frac{24}{12} = 2$$

$$T_X = \frac{t^3 - t}{12} = \frac{2^3 - 2}{12} = \frac{6}{12} = 0,5$$

$$T_X = \frac{t^3 - t}{12} = \frac{2^3 - 2}{12} = \frac{6}{12} = 0,5$$

$$T_X = \frac{t^3 - t}{12} = \frac{3^3 - 3}{12} = \frac{24}{12} = 2$$

$$T_Y = \frac{t^3 - t}{12} = \frac{2^3 - 2}{12} = \frac{6}{12} = 0,5$$

$$T_Y = \frac{t^3 - t}{12} = \frac{2^3 - 2}{12} = \frac{6}{12} = 0,5$$

Dengan demikian

$$r_s = \frac{\sum X^2 + \sum Y^2 - \sum d^2}{2\sqrt{\sum X^2 \sum Y^2}} = \frac{138 + 142 - 22,5}{2\sqrt{138 \cdot 142}} = 0,9197$$

Karena $0,8 < 0,9197 < 1,0$ maka terdapat hubungan yang sangat kuat antara pengalaman usaha tani dengan penerapan teknologi

Uji Signifikansi

1 **Hipotesis :**

$H_0 : \rho_S = 0$ (tidak ada hubungan Yang signifikan antara X dan Y)

$H_1 : \rho_S \neq 0$ (ada hubungan yang signifikan antara X dan Y)

2 **Tingkat Signifikansi : $\alpha = 0,05$**

3 **Statistik Uji :**

$$t = \frac{r_S \sqrt{n-2}}{\sqrt{1-r_S^2}} = \frac{0,9197 \sqrt{12-2}}{\sqrt{1-(0,9197)^2}} = 7,408$$

4 **Kriteria Penolakan :**

H_0 ditolak jika $t_{hitung} > t_{\alpha,df}$ dengan $df = n - 2 = 10$. Karena $t_{hitung} = 7,408 > 2,228 = t_{0,05;10}$ maka H_0 ditolak.

5 **Kesimpulan :** Jadi, ada hubungan yang signifikan antara pengalaman usaha tani dengan penggunaan teknologi.

Latihan

Lakukan analisis data berikut ini menggunakan teknik analisis korelasi tata jenjang. Hitunglah koefisien korelasi antara hasil Tes Keterampilan dan Hasil Prestasi Pertandingan. uji signifikansinya, dan tentukan kesimpulannya.

Data Hasil Tes Keterampilan Teknik dan Hasil Prestasi Pertandingan

Kasus	X	Y
1.	22	41
2.	27	51
3.	25	46
4.	42	66
5.	21	35
6.	31	53
7.	38	60
8.	46	72
9.	45	69
10.	19	39