



MODUL ANALISIS DATA STATISTIK SEKTORAL



DINAS KOMUNIKASI, INFORMATIKA,
STATISTIK DAN PERSANDIAN
KOTA SEMARANG

KATA PENGANTAR

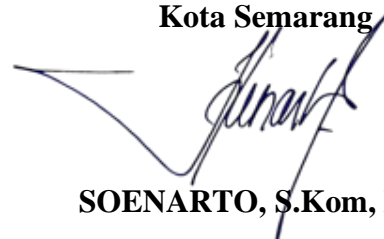
Penyelenggaraan kegiatan statistik sektoral merupakan urusan statistik yang termasuk ke dalam urusan wajib pemerintah non pelayanan publik yang harus dilaksanakan oleh Organisasi Perangkat Daerah di Kota Semarang khususnya Dinas Komunikasi, Informatika, Statistik dan Persandian Kota Semarang. Penyelenggaraan kegiatan statistik sektoral terdiri dari pengumpulan, pengolahan, analisis, dan diseminasi data statistik sektoral. Dalam pelaksanaannya, terdapat beberapa kendala terutama di dalam proses analisis data statistik sektoral. Masih banyak pengelola data yang belum memahami metode-metode analisis data beserta aplikasinya.

Selain itu, untuk mendukung misi Walikota Semarang yang pertama yang berbunyi “Meningkatkan kualitas & kapasitas Sumber Daya Manusia yang unggul & produktif untuk mencapai kesejahteraan & keadilan sosial” dan untuk meningkatkan kapasitas kemampuan sumber daya manusia di Dinas Komunikasi, Informatika, Statistik dan Persandian Kota Semarang yang menjalankan urusan pemerintahan bidang statistik khususnya penyelenggaraan statistik sektoral diperlukan adanya suatu modul yang mampu memberikan pengetahuan serta pemahaman pengelola data.

Modul ini berisi materi beberapa analisis data statistik beserta contoh dan dilengkapi dengan tutorial penggunaan aplikasi analisis data. Sehingga memudahkan pembaca untuk memahami modul ini.

Penyusunan modul ini pasti terdapat banyak kesalahan, sehingga diperlukan kritik dan saran dari pembaca agar modul ini dapat bermanfaat bagi kita semua dan mampu meningkatkan kompetensi pembaca.

Semarang, Desember 2022
**Kepala Dinas Komunikasi, Informatika,
Statistik dan Persandian
Kota Semarang**



SOENARTO, S.Kom, MM.

DAFTAR ISI

	Halaman
KATA PENGANTAR	i
DAFTAR ISI	ii
A. MODUL I : ANALISIS DESKRIPTIF	1
B. MODUL II : PENGENALAN SPSS DAN MINITAB	5
C. MODUL III : ANALISIS PARAMETRIK	21
D. MODUL IV : ANALISIS VARIANSI.....	36
E. MODUL V : ANALISIS REGRESI.....	45
F. MODUL VI : ANALISIS DATA KATEGORIK	56
G. MODUL VII: UJI VALIDITAS DAN RELIABILITAS	68
DAFTAR PUSTAKA	80

MODUL I

ANALISIS DESKRIPTIF

A. UKURAN PEMUSATAN

Sudaryono, Saefullah, dan Rahardja (2012) dalam jurnal : Martias, Lilih Deva. "Statistika deskriptif sebagai kumpulan informasi." *Fihris: Jurnal Ilmu Perpustakaan dan Informasi* 16.1 (2021): 40-59 menyebutkan bahwa ukuran pemusatan merupakan suatu bilangan yang menunjukkan kecenderungan memusatnya label angka atau bilangan pada suatu distribusi data. Ukuran pemusatan dapat dikatakan baik jika dapat menggambarkan atau mewakili suatu data. Pada modul I ini, akan dijelaskan secara singkat mengenai ukuran pemusatan yang terdiri dari *Mean*, Median, dan Modus.

1. Rata-Rata Hitung (*Mean*)

Rata-rata hitung (*Mean*) merupakan bagian dari ukuran pemusatan dari suatu data atau nilai pengamatan yang dipengaruhi oleh nilai ekstrim yang dapat diperoleh dari jumlah seluruh data atau nilai pengamatan dibagi dengan banyaknya data tersebut. Sebagai contoh yaitu rata-rata gaji karyawan swasta di Kota Semarang atau rata-rata usia anak yang bersekolah pada jenjang SMP. Rata-rata hitung data dapat langsung dihitung dengan cara menjumlahkan seluruh data kemudian dibagi dengan banyaknya data tersebut. Rumus untuk menghitungnya adalah sebagai berikut:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Dimana :

\bar{x} merupakan nilai rata-rata hitung

n merupakan banyaknya data

$\sum_{i=1}^n x_i$ merupakan jumlahan nilai dari seluruh data pertama sampai dengan data

ke-n

Sebagai contoh terdapat data tinggi badan anak di sebuah RT yang dinyatakan dalam cm.

Tinggi Badan Anak (cm)
150
152
158
154
147
160
155
154
149
151

Berdasarkan data tersebut, dapat diperoleh banyaknya data $n = 10$. Kemudian diperoleh juga nilai $\sum_{i=1}^n x_i = 150 + 152 + \dots + 151 = 1530$. Selanjutnya dapat dihitung nilai rata-rata hitung $\bar{x} = \frac{1530}{10} = 153$. Sehingga rata-rata tinggi badan anak pada RT tersebut sebesar 153cm.

2. Median

Median merupakan bagian dari ukuran pusat yang nilainya terletak di tengah-tengah dari sekumpulan data dan tidak dipengaruhi oleh nilai ekstrim. Untuk menghitung median, dimulai dari mengurutkan data dari yang terkecil sampai yang terbesar, kemudian ditentukan nilai median dengan rumus :

$$Med = X_{\frac{n+1}{2}} \text{ untuk } n \text{ ganjil}$$

$$Med = \frac{X_n + X_{n+2}}{2} \text{ untuk } n \text{ genap}$$

Sebagai contoh diketahui sampel nilai dari siswa kelas 6 SD sebagai berikut [80 75 79 90 65]. Untuk menghitung nilai mediannya, diurutkan terlebih dahulu data tersebut dari yang terkecil sampai yang terbesar, sehingga diperoleh data sebagai berikut [65 75 79 80 90]. Dapat diketahui juga nilai $n = 5$, sehingga $Med = X_{\frac{5+1}{2}} = X_3 = 79$. Jadi median dari nilai siswa tersebut adalah 79.

3. Modus

Modus merupakan nilai yang paling sering muncul dari sekumpulan data yang tidak dipengaruhi oleh nilai ekstrim. Misalkan terdapat sekumpulan data [4 8 7 9 10 18 7 6 2 7 1]. Jika diperhatikan dengan seksama, kita akan memperoleh nilai modus dari data tersebut yaitu [7] karena nilai tersebut muncul sebanyak tiga kali sedangkan data yang lain muncul tidak lebih banyak dari nilai [7]. Selanjutnya misalkan terdapat sekumpulan data [1 3 9 4 2 7], maka tidak ada modus pada data tersebut karena setiap nilai muncul dengan jumlah yang sama.

B. UKURAN SEBARAN

1. Rentang

Rentang merupakan bagian dari sebaran data yang dapat dihitung dengan cara menghitung selisih data yang terbesar dengan data yang terkecil dari sekelompok data. Rentang yang lebih kecil menunjukkan bahwa data tersebut bersifat homogen atau merata. Misalkan diketahui data [1 3 4 6 8 10], maka dapat diperoleh rentang dari data tersebut yaitu $\text{rentang} = 10 - 1 = 9$.

2. Variansi

Variansi merupakan ukuran yang menunjukkan seberapa jauh data tersebut tersebar. Variansi bernilai nol menunjukkan bahwa semua data memiliki nilai yang sama. Variansi dapat dihitung menggunakan rumus :

$$s^2 = \frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Dimana :

s^2 merupakan nilai variansi

\bar{x} merupakan nilai rata-rata hitung

n merupakan banyaknya data

x_i merupakan nilai dari tiap-tiap data

3. Standar Deviasi (Simpangan Baku)

Standar deviasi merupakan ukuran sebaran yang dapat dihitung dengan mengakarkan nilai dari variansi. Berikut adalah rumus untuk menghitung nilai dari standar deviasi :

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum (X_i - \bar{X})^2}$$

Dimana :

s^2 merupakan nilai variansi

\bar{x} merupakan nilai rata-rata hitung

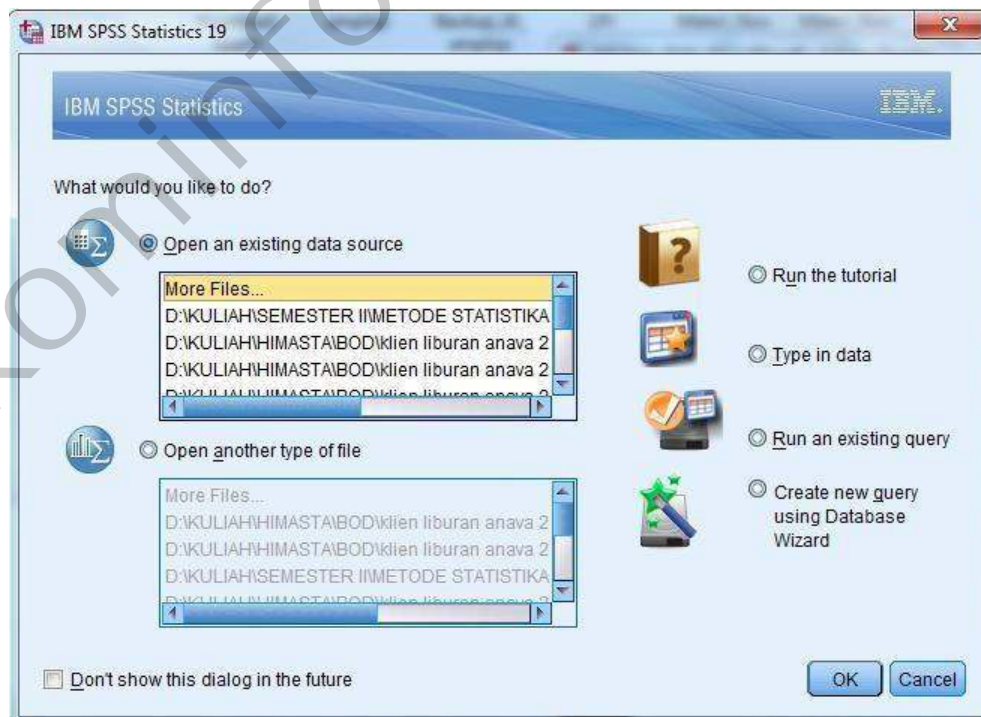
n merupakan banyaknya data

x_i merupakan nilai dari tiap-tiap data

PENGENALAN SPSS DAN MINITAB

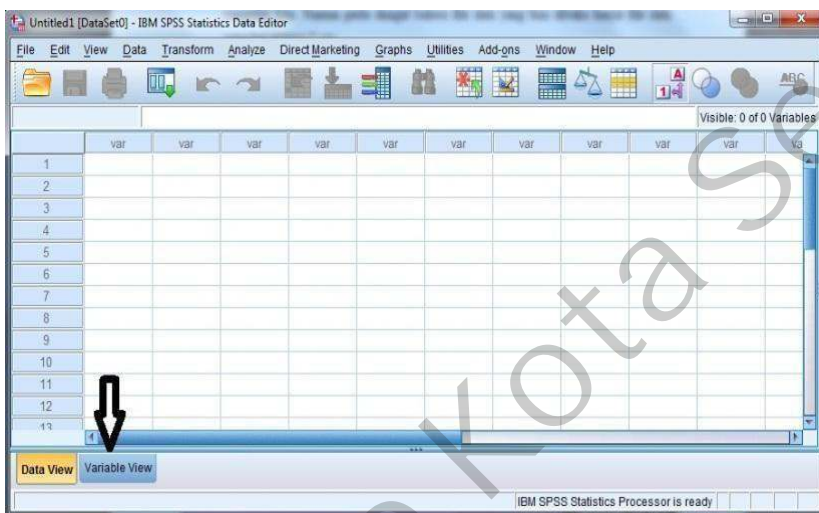
1. PENGENALAN SPSS

SPSS atau Statistical Product and Service Solution merupakan salah satu paket program yang sangat penting dalam pengolahan data yang sangat populer dan sangat mudah dalam pengoperasiannya mulai dari memasukkan data, mengedit data, mentransformasikan data, analisis data hingga menyajikan hasil analisis data. Fitur-fitur di dalam SPSS dinilai sangat lengkap dan selalu terupdate mengikuti perkembangan IPTEK. Tidak hanya statistisi saja yang menggunakan aplikasi ini, namun digunakan oleh banyak orang untuk berbagai kepentingan. Namun aplikasi ini tidak cocok digunakan untuk analisis terhadap data yang jumlahnya sangat besar (BIG DATA). Dalam melakukan input data tidak bisa secara langsung seperti Microsoft Excel. Namun harus menuliskan dan mendefinisikan variabel-variabel yang akan dianalisis. Langkah-langkahnya sebagai berikut



Pertama kita harus membuka aplikasi SPSS sampai muncul kotak dialog seperti gambar di atas. Selanjutnya kita dapat memilih beberapa pilihan awal yang kita inginkan, mulai dari menjalankan tutorial penggunaan SPSS (run the tutorial) sampai dengan pilihan untuk membuka file SPSS yang sudah tersimpan (open an existing data source). Jika kita ingin memulai SPSS dengan file kerja baru maka kita tinggal menutup kotak dialog awal tadi. SPSS akan langsung menampilkan lembar kerja kosong yang langsung bisa kita gunakan. Selanjutnya dalam memasukkan data terlebih dahulu kita harus mendefinisikan variabel yang akan kita analisis. Caranya adalah sebagai berikut :

1. Pilih variable view pada pojok kiri bawah, seperti pada gambar di bawah ini.



2. Kemudian muncul spreadsheet sebagai berikut :

Variable view ini berisi :

- Name, digunakan untuk mengisi nama variabel. Nama tidak boleh memuat tanda spasi, diawali dengan angka ataupun tanda operator.
- Type, digunakan untuk menentukan tipe data
 - a. Numeric, data berupa angka
 - b. Date, data berupa tanggal
 - c. Dollar dan Custom Currency, data berupa nilai mata uang
 - d. String, data berupa huruf
- Width, menentukan lebar kolom yang dikehendaki
- Decimals, menentukan banyaknya angka desimal yang dikehendaki
- Label, memberikan penjelasan rinci tentang nama variabel
- Values, memberikan penjelasan mengenai label berdasarkan nilai variabel

Jika kita menggunakan pilihan values ini, maka akan muncul kotak dialog

sebagai berikut :



Misal kita ingin mendefinisikan data kita sebagai berikut : 1=laki-laki, 2=perempuan, maka pada value kita ketik 1 dan pada label kita ketik laki-laki, lalu klik add, begitu seterusnya.

Contoh cara penginputan Data di SPSS :

Sebelum menggunakan	Setelah menggunakan	Jenis Kelamin
45	70	Laki-Laki
67	55	Perempuan
82	78	Perempuan
34	45	Laki-Laki

3. Terlebih dahulu kita definisikan data kita pada variable view sebagai berikut: Pada baris pertama kita definisikan untuk variabel “Sebelum Menggunakan Obat”.

- Name : Sebelum
- Type : Numerics - Decimals : 0
- Label boleh diisi boleh tidak diisi, label hanya untuk memperjelas nama variabelnya. Boleh diisi “Sebelum Menggunakan Obat”.
- Biarkan yang lainnya default.

Langkah yang sama kita gunakan untuk mendeskripsikan variabel “Sesudah Menggunakan Obat”. Untuk variabel Jenis Kelamin, langkah-langkahnya :

- Name : JK
- Type : Numeric
- Decimals : 0
- Label : Jenis Kelamin

- Values : 1 untuk laki-laki, dan 2 untuk perempuan. Sehingga diperoleh tampilan sebagai berikut:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	sebelum	Numeric	8	2	sebelum meng...	None	None	8	Right	Scale	Input
2	sesudah	Numeric	8	2	sesudah meng...	None	None	8	Right	Scale	Input
3	jenis_kelamin	Numeric	8	2	jenis kelamin	{1,00, laki-la...	None	8	Right	Scale	Input
4											

- Setelah semua Variabel kita deskripsikan, maka kita tinggal memasukkan data yang kita inginkan. Kembali ke worksheet data view, yaitu dengan memilih “data view” di pojok kiri bawah. Di samping variable view. Lalu isikan data, sehingga tampilannya menjadi :

	sebelum	sesudah	jenis_kela...	var
1	45,00	70,00	1,00	
2	67,00	55,00	2,00	
3	82,00	78,00	2,00	
4	34,00	45,00	1,00	
5				

Terlihat bahwa pada variabel JK, datanya masih berupa angka. Untuk memunculkan value dari variabel tersebut, maka klik View > Value Label. Maka pada data Variabel JK akan berubah sesuai dengan value yang telah kita definisikan sebelumnya (1=laki-laki, 2=perempuan).

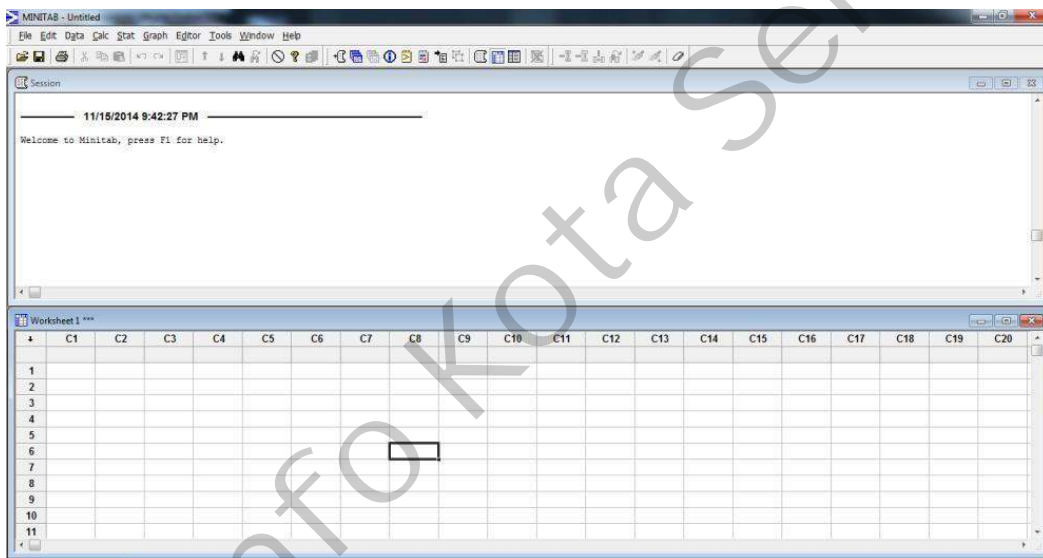
Tampilannya menjadi :

	sebelum	sesudah	jenis_kela...	var
1	45,00	70,00	laki-laki	
2	67,00	55,00	perempuan	
3	82,00	78,00	perempuan	
4	34,00	45,00	laki-laki	
5				

2. PENGENALAN MINITAB

Minitab adalah program komputer yang dirancang untuk melakukan pengolahan statistik. Minitab mengombinasikan kemudahan penggunaan layaknya Microsoft Excel dengan kemampuannya melakukan analisis statistik yang kompleks. Paket program Minitab merupakan salah satu software yang sangat besar kontribusinya sebagai media pengolahan data statistik. Software ini menyediakan berbagai jenis perintah yang memungkinkan proses pemasukan data, manipulasi data, pembuatan grafik, dan berbagai analisis statistik.

Tampilan minitab:



Minitab memiliki dua layar primer, yaitu worksheet (lembar kerja) untuk melihat dan mengedit lembar kerja dan Session atau command untuk menampilkan hasil.

Contoh penggunaan :

Data berat bayi setelah dilahirkan dan panjang bayi di Rumah Sakit.

Berat Lahir (kg)	2.75	2.15	4.41	5.52	3.21	4.32	2.31	4.30	3.71	3.60
Panjang Bayi (cm)	57.5	52.8	61.3	67.0	53.5	62.7	56.2	68.5	69.2	60.0

- Hitung nilai rata-rata, standar deviasi, nilai maksimum dan minimum, kuartil 1, median, dan kuartil 3 dari berat badan dan tinggi badan !
- Buatlah histogram dari Berat Lahir bayi!

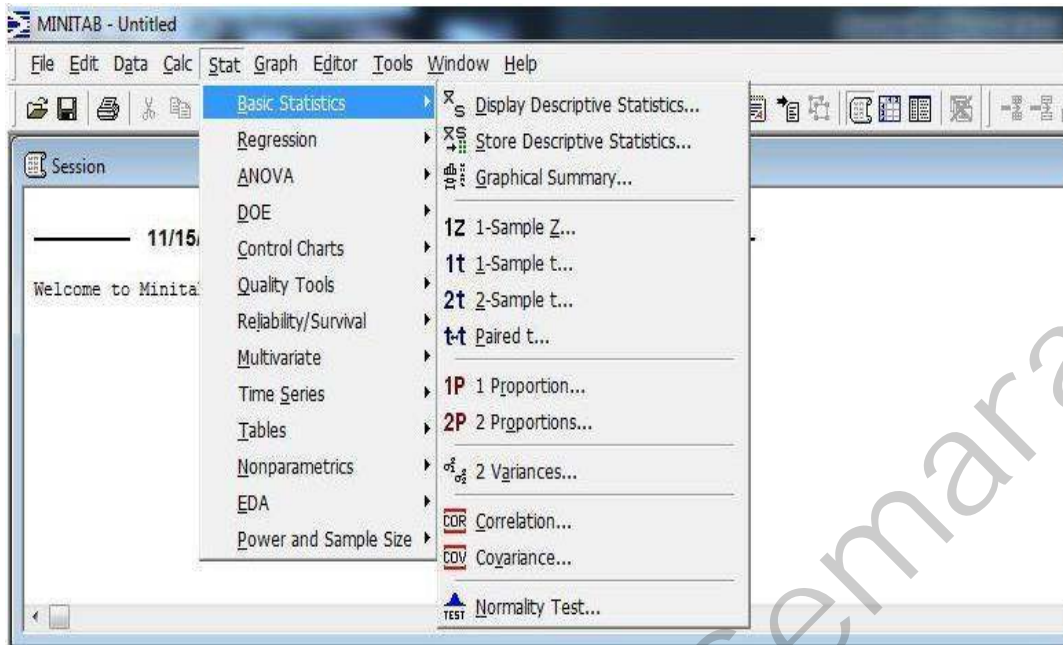
Langkah-langkah penyelesaian :

1. Input data berat lahir dan panjang bayi ke lembar kerja (Worksheet)

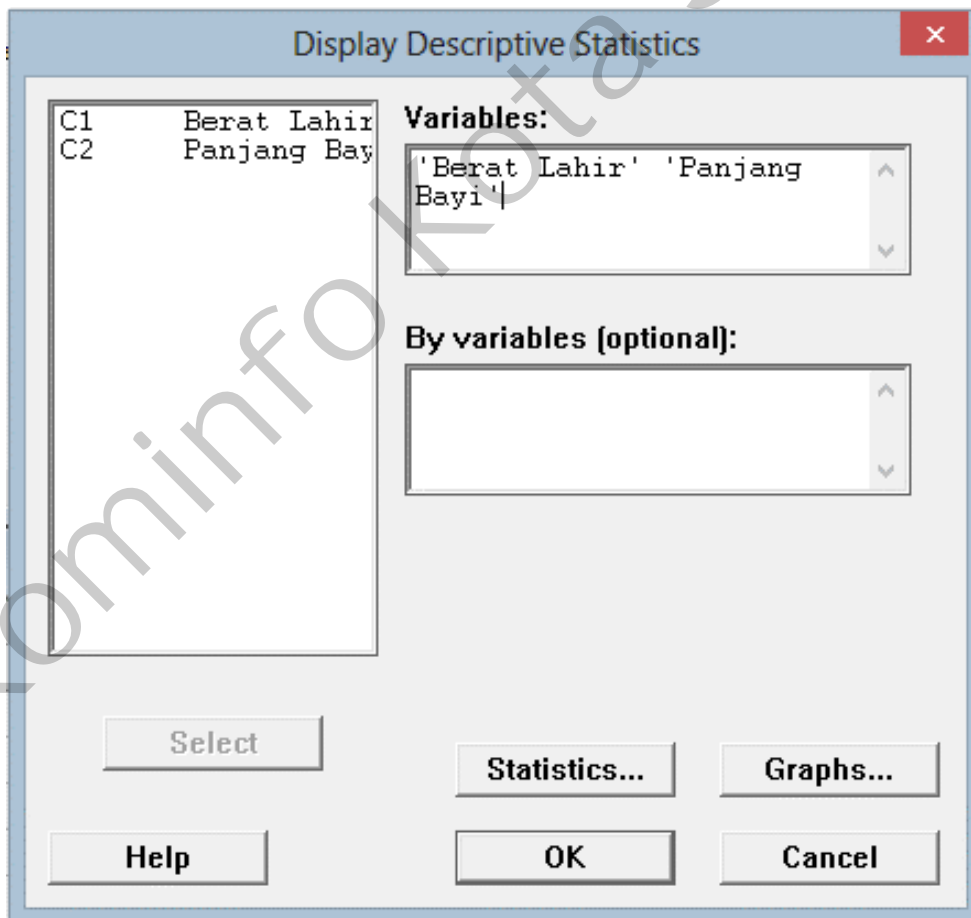
	C1	C2	C3	C4	C5	C6
	Berat Lahir	Panjang Bayi				
1	2.75	57.5				
2	2.15	52.8				
3	4.41	61.3				
4	5.52	67.0				
5	3.21	53.5				
6	4.32	62.7				
7	2.31	56.2				
8	4.30	68.5				
9	3.71	69.2				
10	3.60	60.0				
11						

Diskominfo Kota Semarang

- Klik menu Stat >> Basic Statistic >> Display Deskriptive Statistics

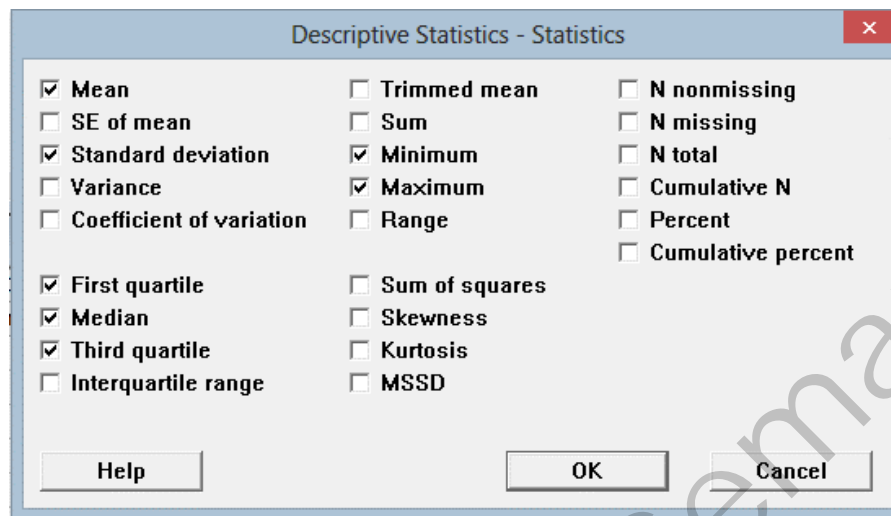


- Masukkan data berat lahir dan panjang bayi pada kolom variables



- Klik Statistics.

5. Check list pada nilai mean, standard deviation, first quartile, median, third quartile, minimum, maximum.



6. Klik OK
7. Hasil pada lembar Session

23/09/2015 7:55:59

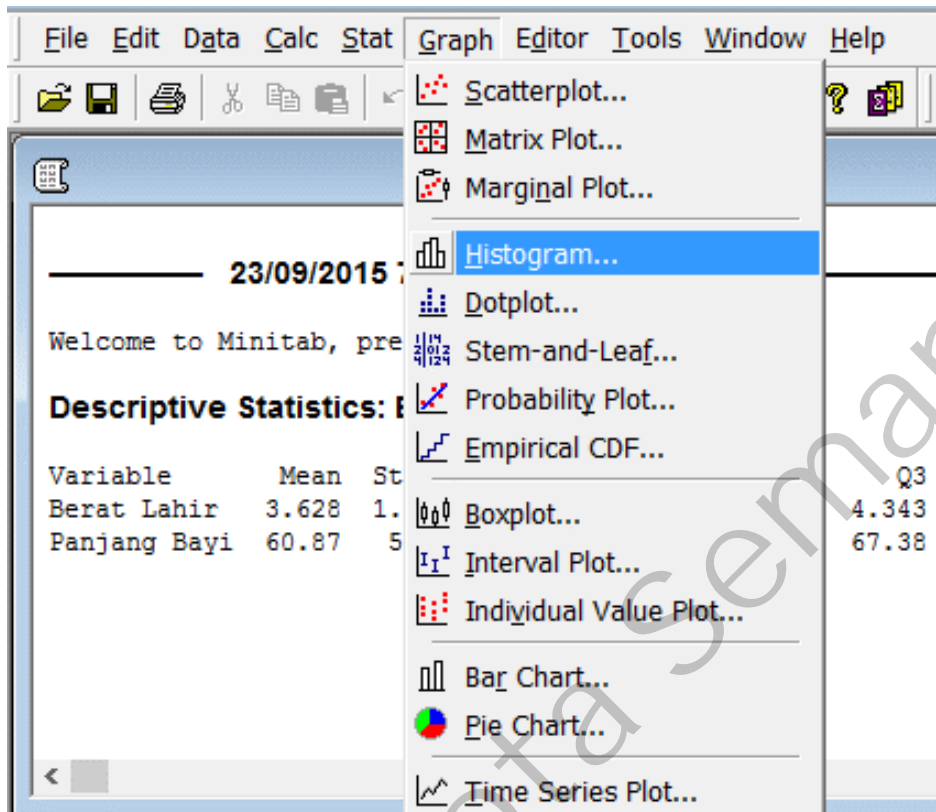
Welcome to Minitab, press F1 for help.

Descriptive Statistics: Berat Lahir; Panjang Bayi

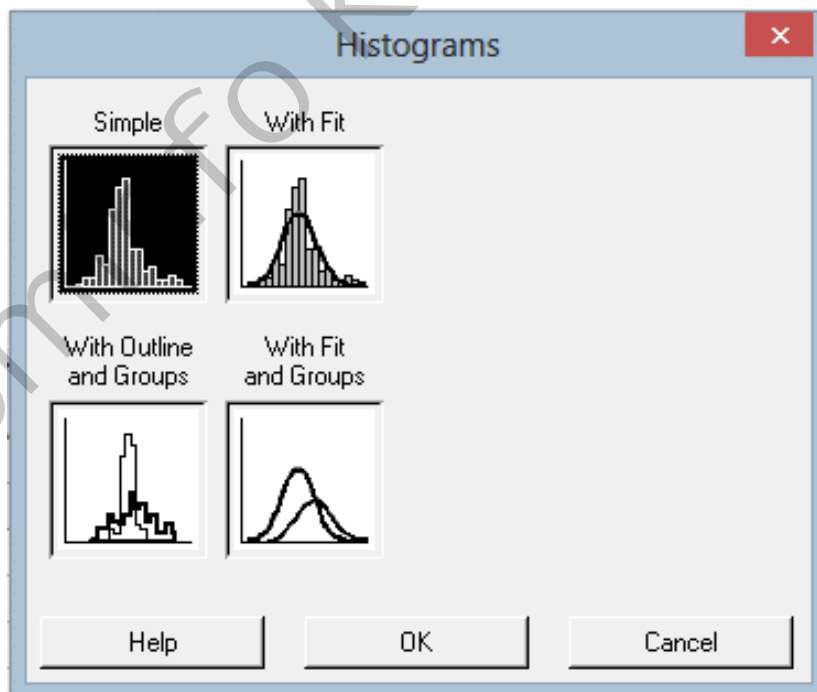
Variable	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Berat Lahir	3.628	1.054	2.150	2.640	3.655	4.343	5.520
Panjang Bayi	60.87	5.98	52.80	55.53	60.65	67.38	69.20

Membuat histogram berat badan:

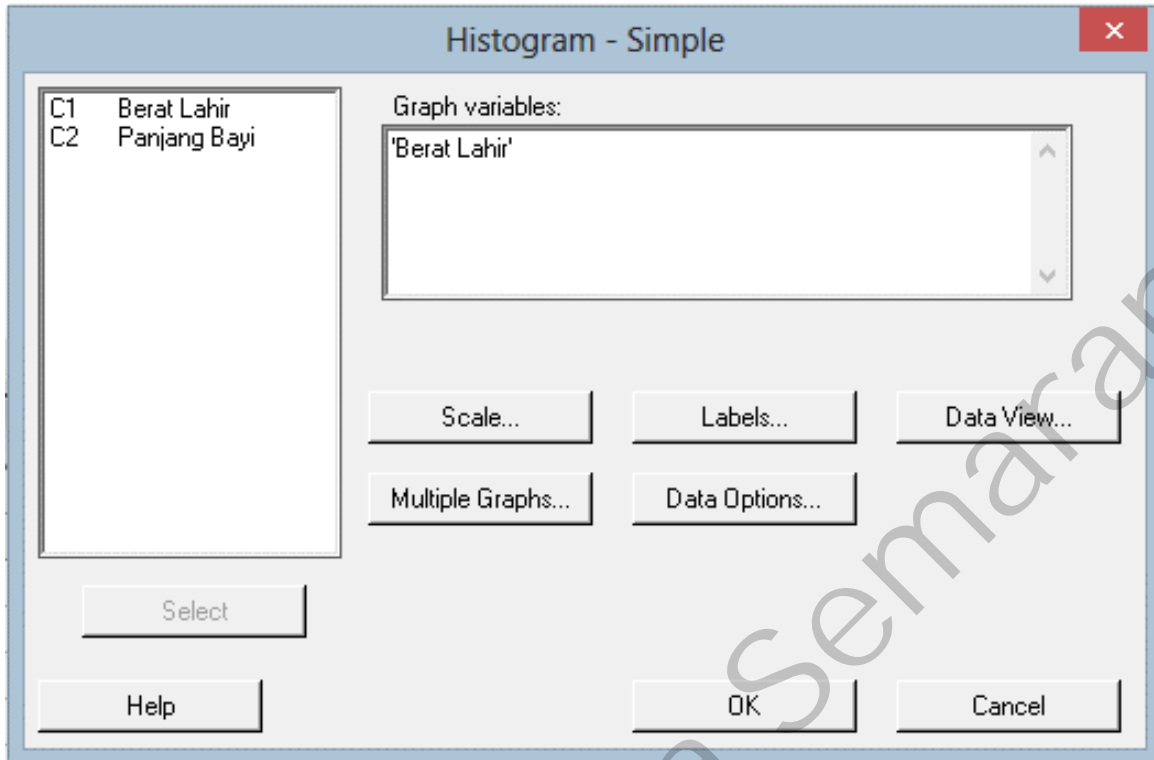
1. Klik menu Graph >> Histogram



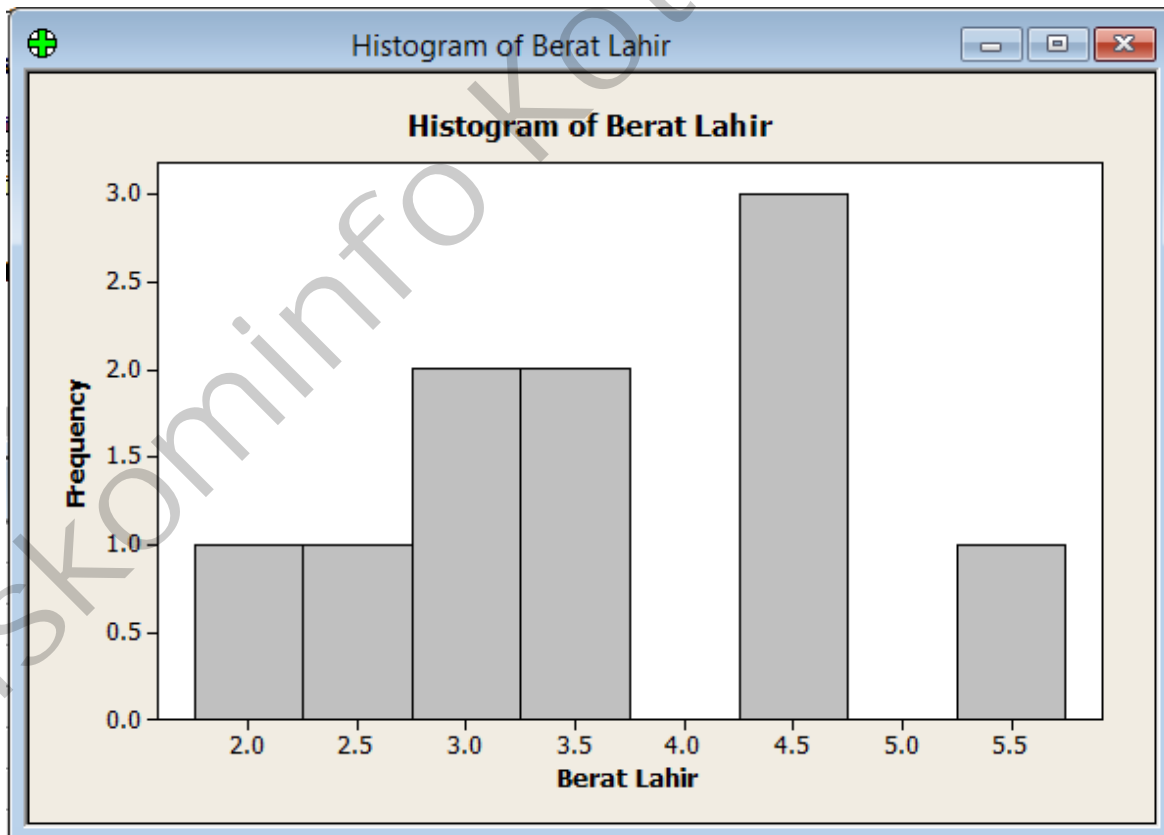
2. Klik simple >> OK



- Masukkan variabel berat lahir pada kolom Graph variables



- Klik Ok



MODUL III
ANALISIS PARAMETRIK

1. UJI MEAN 1 POPULASI DENGAN VARIANSI DIKETAHUI (UJI Z)

Uji Hipotesis Untuk Mean

Uji hipotesis yang digunakan untuk mean suatu populasi sama dengan harga tertentu dengan $n > 30$ atau variansi populasi diketahui maka digunakan statistik uji z (pada Minitab ; Uji z digunakan jika variansi populasi diketahui tanpa memperhatikan ukuran sampel).

Hipotesis yang diuji dapat berbentuk sebagai berikut :

1. Nyatakan H_0 dan H_1
 - a. $H_0 : \mu = \mu_0$ (uji dua sisi)
 $H_1 : \mu \neq \mu_0$
 - b. $H_0 : \mu \leq \mu_0$ (uji sisi kanan)
 $H_1 : \mu > \mu_0$
 - c. $H_0 : \mu \geq \mu_0$ (uji sisi kiri)
 $H_1 : \mu < \mu_0$
2. Tentukan taraf nyata α (tingkat signifikansi)
3. Tentukan statistik pengujian yang akan digunakan

Jika jumlah sampel lebih dari 30 (sampel besar), gunakan uji Z

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Jika jumlah sampel kurang dari 30 (sampel kecil), gunakan uji t.

4. Tentukan daerah kritik :

H_0 ditolak jika

- $Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$
- $Z > Z_{\alpha}$
- $Z < -Z_{\alpha}$

5. Ambil kesimpulan untuk menolak atau menerima

Berdasarkan langkah 4 dan hasil hitungan statistik pengujian langkah 3, diambil kesimpulan apakah ditolak atau tidak ditolak pada tingkat signifikansi α .

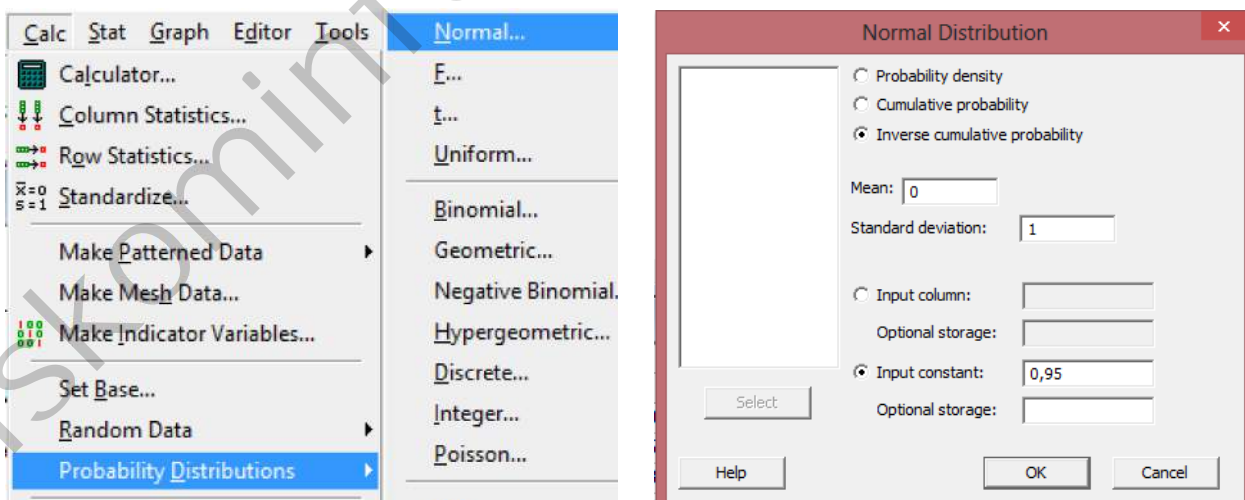
Contoh :

Suatu perusahaan asuransi jiwa di Yogyakarta mempunyai kantor cabang. Rata-rata waktu tunggu untuk klaim asuransi dari pengajuannya di cabang Jalan Kaliurang tahun lalu adalah 3,7 hari dengan standar deviasi 0,8 hari untuk sampel 45 klaim yang kemudian melakukan evaluasi. Berdasarkan data waktu tunggu di atas, apakah rata-rata waktu tunggu sekarang kurang dari 3 hari? Gunakan tingkat signifikansi 5% dan asumsikan bahwa sebaran umur tersebut normal!

Penyelesaiannya :

Diketahui : $n = 45$; $\mu_0 = 3$; $\bar{x} = 3,7$; $\sigma = 0,8$; $\alpha = 5\%$

Ditanyakan : Rata-rata waktu tunggu klaim sekarang kurang dari 3 atau tidak.



1. Hipotesis :

$$H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0$$

$$H_0 : \mu \leq 3 \text{ vs } H_1 : \mu > 3$$

2. Tingkat signifikansi :

$$\sigma = 5\% = 0,05$$

3. Uji statistik :

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$Z = \frac{3,7 - 3}{\frac{0,8}{\sqrt{45}}}$$

$$= 5,8697$$

4. Daerah kritik

H_0 ditolak jika $Z > Z_\alpha$ (menggunakan program Minitab)

$$5,8697 > 1,64485$$

Inverse Cumulative Distribution Function

Normal with mean = 0 and standard deviation = 1

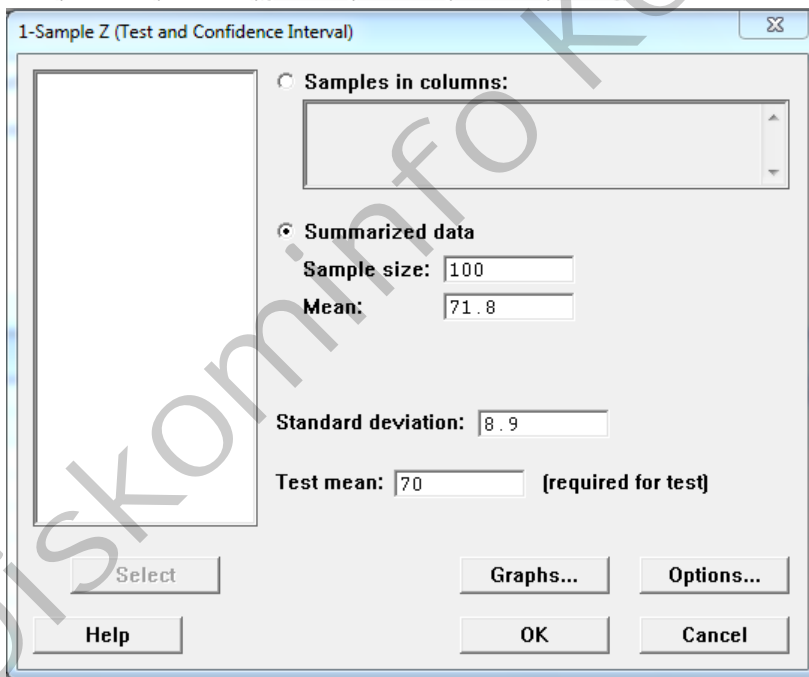
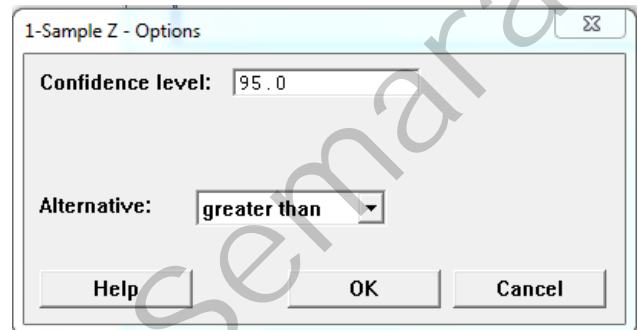
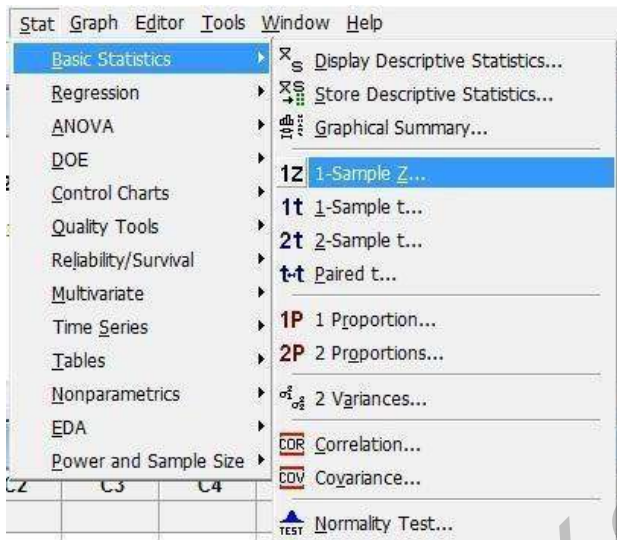
P(X <= x)	x
0,95	1,64485

5. Kesimpulan :

H_0 ditolak karena $Z > Z_\alpha$, maka benar bahwa rata-rata waktu tunggu klaim sekarang kurang dari 3 hari.

Cara lain dengan menggunakan Minitab :

Meng-klik Stat, memilih Basic Statistics, lalu memilih 1-Sample Z, mengisikan sample size(n), mean, standard deviation, dan test mean sesuai dengan data dalam soal, lalu klik Options untuk menentukan tingkat keyakinan (confidence level) dan alternative sesuai pertanyaan, selanjutnya klik OK, terakhir klik OK.



Hasil :

One-Sample Z					
Test of mu = 70 vs > 70					
The assumed standard deviation = 8.9					
95%					
Lower					
N	Mean	SE Mean	Bound	Z	P
100	71.8000	0.8900	70.3361	2.02	0.022

Kesimpulan :

H_0 ditolak karena p-value (0,000) < α (0,05), maka benar bahwa rata-rata waktu tunggu klaim sekarang kurang dari 3 hari.

2. UJI MEAN 1 POPULASI DENGAN VARIANSI TIDAK DIKETAHUI (UJI T)

Uji Hipotesis Untuk Mean

Uji hipotesis yang digunakan untuk mean suatu populasi sama dengan harga μ_0 tertentu dengan $n < 30$ atau variansi populasi tidak diketahui, maka digunakan statistic uji t (pada minitab uji t digunakan jika variansi populasi tidak diketahui tanpa memperhatikan ukuran sampel).

Hipotesis yang diuji dapat berbentuk sebagai berikut :

1. Nyatakan H_0 dan H_1
 - a. $H_0 : \mu = \mu_0$ (uji dua sisi)
 $H_1 : \mu \neq \mu_0$
 - b. $H_0 : \mu \leq \mu_0$ (uji sisi kanan)
 $H_1 : \mu > \mu_0$
 - c. $H_0 : \mu \geq \mu_0$ (uji sisi kiri)
 $H_1 : \mu < \mu_0$

2. Tentukan taraf nyata α (tingkat signifikansi)

3. Statistik uji

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

4. Tentukan daerah kritik

H_0 ditolak jika

- a. $t > t_{\alpha/2, n-1}$ atau $t < -t_{\alpha/2, n-1}$
- b. $t > t_{\alpha, n-1}$
- c. $t < -t_{\alpha, n-1}$

5. Ambil kesimpulan untuk menolak atau menerima Berdasarkan langkah 4 dan hasil hitungan statistik pengujian langkah 3, diambil kesimpulan apakah H_0 ditolak atau tidak ditolak pada tingkat tingkat signifikansi α .

Contoh :

Berikut adalah data mean jarak yang ditempuh oleh mobil per bulannya di Yogyakarta (dalam km). Diambil sampel sebanyak 12 mobil dan dicatat mean jarak tempuh per bulannya sebagai berikut :

2059	2255	1840
2161	1945	2503
2419	2028	1969
2109	1683	2252

Data tersebut dianggap berdistribusi normal. Selanjutnya, tentukan :

- a. Interval konfidensi 95% untuk mean jarak tempuh mobil dan simpulkan hasilnya!
- b. Berdasarkan data di atas, dapatkah kita simpulkan bahwa mean jarak yang ditempuh oleh mobil tersebut kurang dari 2200 km? Gunakan $\alpha = 5\%$.

Penyelesaian :

Diketahui : Data mean jarak tempuh mobil per bulan di Yogyakarta (dalam km)
2059, 2161, 2419, 2109, 2255, 1945, 2028, 1683, 1840, 2503, 1969, 2252.

Data diasumsikan berdistribusi normal.

Ditanyakan :

- a. Berdasarkan data di atas, dapatkah kita simpulkan bahwa mean jarak yang ditempuh oleh mobil tersebut kurang dari 2200 km? Gunakan $\alpha = 5\%$.

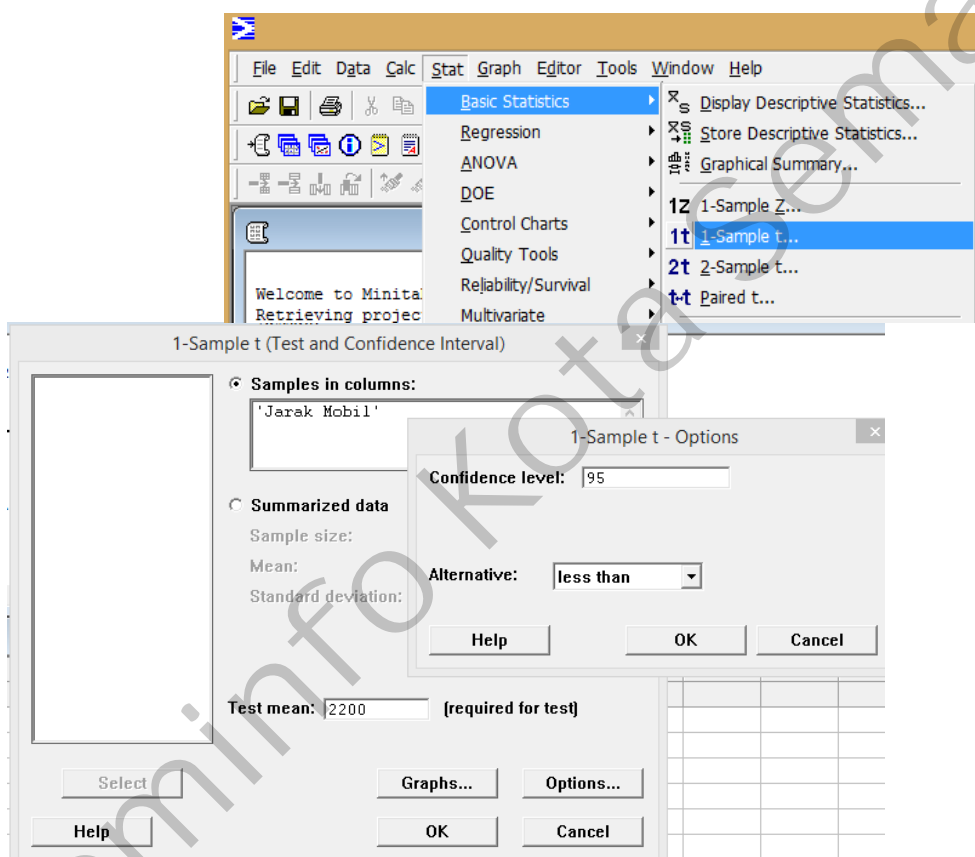
Jawab :

Data mean jarak tempuh mobil per bulan di Yogyakarta (dalam km) :
2059, 2161, 2419, 2109, 2255, 1945, 2028, 1683, 1840, 2503, 1969, 2252.

Karena data diasumsikan berdistribusi normal maka tidak perlu untuk melakukan uji kenormalan.

Dengan menggunakan program Minitab :

Stat >> Basic Statistics >> 1 Sample t >> mengisikan datanya >> Options >> mengisikan datanya dengan confidence level 95% dan alternative less than >> OK >> OK.



Hasil :

One-Sample T: Jarak Mobil							
Test of mu = 2200 vs < 2200							
					95%		
Variable	N	Mean	StDev	SE Mean	Upper Bound	T	P
Jarak Mobil	12	2101.92	234.44	67.68	2223.46	-1.45	0.088

1. Tentukan daerah kritik

H_0 ditolak jika $t < -t_{\alpha, n-1}$ (secara manual)

$$-1,45 > -1,796$$

H_0 ditolak jika $p\text{-value} < \alpha$

$$0,088 > 0,05$$

2. Kesimpulan

H_0 tidak ditolak. Jadi, mean jarak yang ditempuh oleh mobil tersebut per bulannya lebih dari atau sama dengan 2200 km.

3. UJI MEAN 2 POPULASI

Tujuan : Untuk mengetahui apakah terdapat kesamaan rata-rata dari data yang dikumpulkan 2 sampel random. Terdiri dari 2 macam:

- **Uji rata-rata 2 populasi Independen.** Berasal dari populasi yang berbeda
 Contohnya: mengetahui apakah ada beda hasil dari mengambil batubara dengan metode A dengan Metode B
- **Uji rata-rata 2 populasi dependen.** Berasal dari populasi yang sama
 Contohnya: Mengetahui tingkat keefektifan obat penurun berat badan, dengan data sebelum minum obat dan sesudah minum obat.

Asumsi :

- Normalitas
- Variasi data dari kedua kelompok sama

Note : Jika Asumsi normalitas tidak terpenuhi, maka digunakan uji mean 2 populasi nonparametrik (tidak dibahas di sini)

Contoh :

Diberikan data pendaftar sekolah sebagai berikut :

Sekolah Negeri	Sekolah Swasta
750	600
700	750
555	700
650	675
625	650
720	450
600	750
495	500

Modul Analisis Data Statistik

520	525
535	615
600	600
605	750
701	725
405	550
435	565

Apakah rata-rata pendaftar sekolah negeri lebih banyak daripada sekolah swasta? Gunakan $\alpha=5\%$!

Tahapan :

Uji Normalitas : Stat >> Basic Statistics >> Normality Test

- Pilih variabel yang akan diuji normalitasnya
- Percentile lines : none
- Test for normality :
Ryan-Joiner (Shapiro-Wilk) digunakan jika $n \leq 50$
Kolmogorov-Smirnov digunakan jika $n > 50$
- OK

Uji Kesamaan Variansi : Stat >> Basic Statistics >> 2 Variances

- Data : sample in different columns (jika kedua data sampel terletak pada dua kolom yang berbeda)
- Masukkan data sampel pada "first" dan "second"
- OK

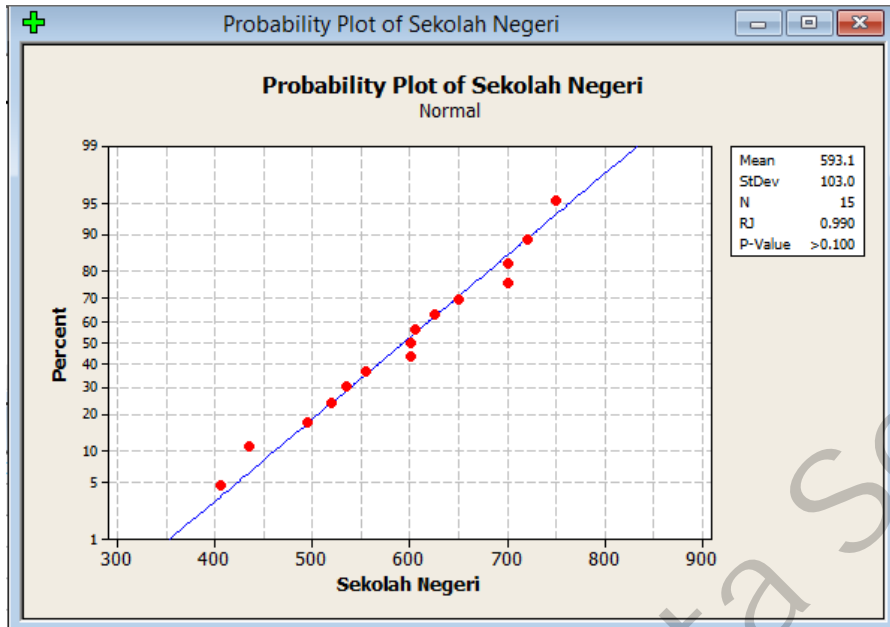
Uji Hipotesis Rata-Rata Dua Populasi Independen : Stat >> Basic Statistics >> 2-sample t

- Pilih sample in different columns jika kedua data sampel terletak pada dua kolom yang berbeda, kemudian masukkan datanya.
- Beri checklist pada "assume equal variances" (jika variansi kedua sampel sama)
- Options:
Confidence level : $1-\alpha$ (dalam persen)
Test difference : selisih $\mu_1-\mu_2$ yang ingin diuji
Alternative : H_1 dari hipotesis
- OK

Jawaban :

Diketahui : μ_1 = sekolah negeri μ_2 = sekolah swasta

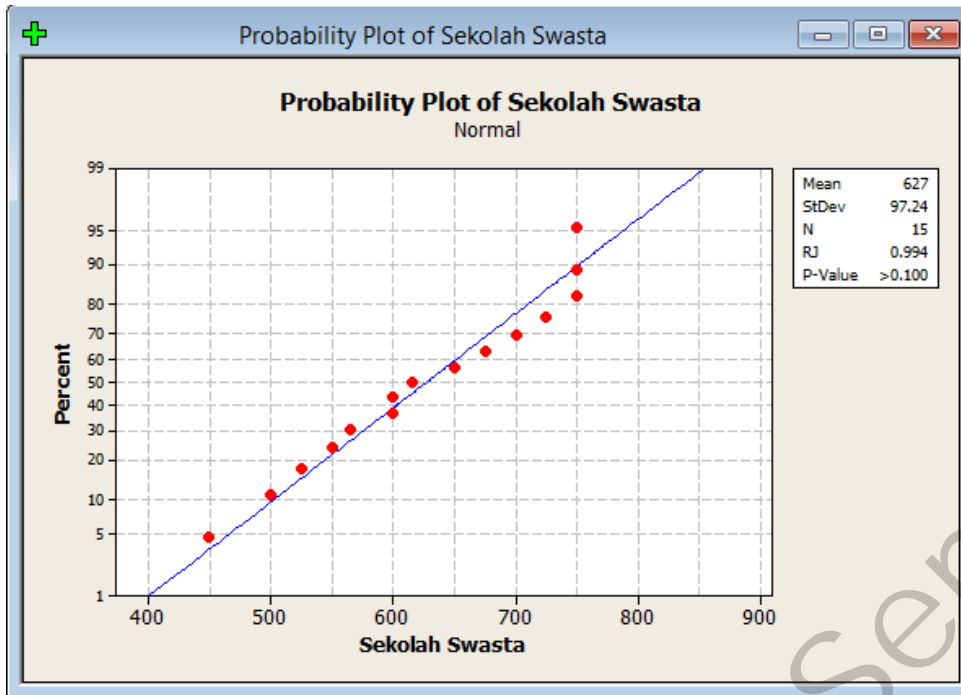
Uji Normalitas (Sekolah Negeri)



- Hipotesis :
 - H0: data berdistribusi normal
 - H1: data tidak berdistribusi normal
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = $> 0,100$
- Daerah kritik : H_0 ditolak jika p-value $< \alpha$
- Kesimpulan :

Karena p-value ($> 0,100$) $> \alpha$ (0,05), maka H_0 tidak ditolak, sehingga dapat disimpulkan bahwa data berdistribusi normal.

Uji Normalitas (Sekolah Swasta)



- Hipotesis :
 - H0: data berdistribusi normal
 - H1: data tidak berdistribusi normal
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = $> 0,100$
- Daerah kritik : H_0 ditolak jika p-value $< \alpha$
- Kesimpulan :

Karena p-value ($> 0,100$) $> \alpha$ (0,05), maka H_0 tidak ditolak, sehingga dapat disimpulkan bahwa data berdistribusi normal.

Uji kesamaan Variansi

Test for Equal Variances: Sekolah Negeri, Sekolah Swasta				
95% Bonferroni confidence intervals for standard deviations				
	N	Lower	StDev	Upper
Sekolah Negeri	15	72.2721	102.975	174.538
Sekolah Swasta	15	68.2497	97.244	164.824
F-Test (normal distribution)				
Test statistic = 1.12, p-value = 0.833				
Levene's Test (any continuous distribution)				
Test statistic = 0.00, p-value = 0.990				

- Hipotesis :
 $H_0: \sigma_1^2 = \sigma_2^2$ (variansi sama)
 $H_1: \sigma_1^2 \neq \sigma_2^2$ (variansi berbeda)
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = 0,990 (P-value Lavene's Test)
- Daerah kritik : H_0 ditolak jika p-value $< \alpha$
- Kesimpulan :
 Karena p-value (0,990) $> \alpha$ (0,05), maka H_0 tidak ditolak, sehingga dapat disimpulkan bahwa variansi dari kedua data tersebut sama.

Uji Hipotesis Rata-rata Dua Populasi Independen

Two-Sample T-Test and CI: Sekolah Negeri, Sekolah Swasta

Two-sample T for Sekolah Negeri vs Sekolah Swasta

	N	Mean	StDev	SE Mean
Sekolah Negeri	15	593	103	27
Sekolah Swasta	15	627.0	97.2	25

Difference = μ (Sekolah Negeri) - μ (Sekolah Swasta)
 Estimate for difference: -33.9333
 95% lower bound for difference: -96.1435
 T-Test of difference = 0 (vs >): T-Value = -0.93 P-Value = 0.819 DF = 28
 Both use Pooled StDev = 100.1508

- Hipotesis :
 $H_0: \mu_1 \leq \mu_2$
 $H_1: \mu_1 > \mu_2$
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = 0,819
- Daerah kritik : H_0 ditolak jika p-value $< \alpha$
- Kesimpulan :
 Karena p-value (0,819) $> \alpha$ (0,05), maka H_0 tidak ditolak, sehingga dapat disimpulkan bahwa rata-rata pendaftar sekolah negeri kurang dari atau sama dengan sekolah swasta atau dapat dikatakan bahwa pendaftar sekolah negeri tidak lebih banyak dibandingkan sekolah swasta.

Contoh :

Uji Hipotesis Rata-Rata Dua Populasi Dependen

Sebuah perusahaan obat-obatan mengiklankan bahwa obat pelangsing ABC mampu menurunkan berat badan orang yang meminumnya jika diminum secara teratur dalam waktu 3 minggu. Yayasan Lembaga Konsumen Indonesia mengadakan penelitian terhadap masalah ini dan diperoleh data sebagai berikut:

Wanita	1	2	3	4	5	6	7	8	9	10	11
Sebelum	78	77	73	65	59	53	61	62	62	60	61
Sesudah	72	78	69	61	53	49	62	56	57	61	53

Lakukanlah uji hipotesis dari data di atas dengan $\alpha = 5\%$. Apakah benar bahwa obat pelangsing tersebut dapat menurunkan berat badan?

Tahapan :

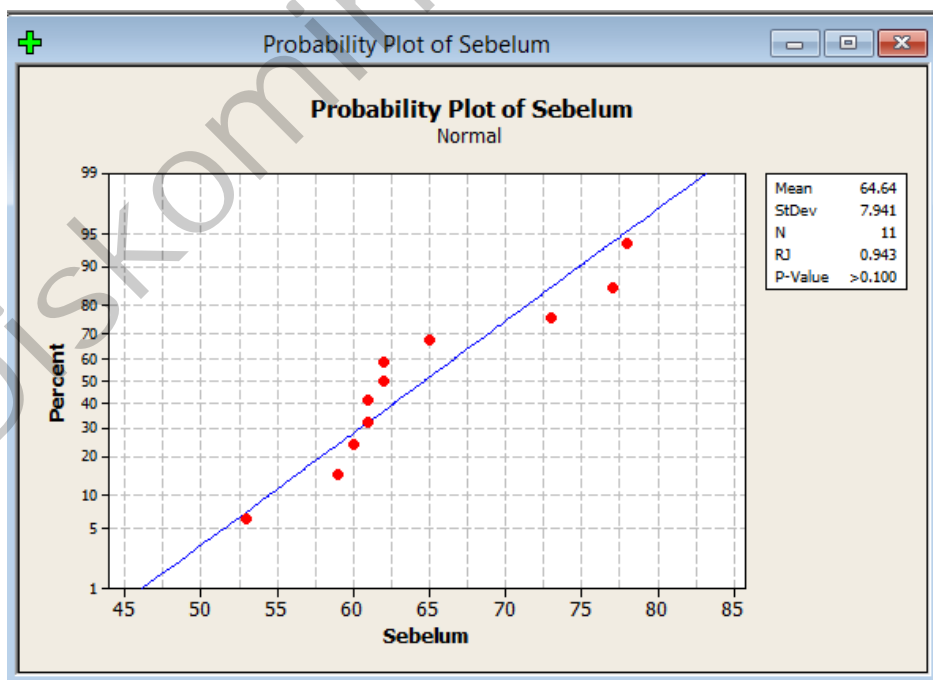
Uji Normalitas : Stat >> Basic Statistics >> Normality Test

Uji Hipotesis Rata-Rata Dua Populasi Dependen : Stat >> Basic Statistics >> Paired t

Diketahui : μ_1 = mean berat badan sebelum meminum obat

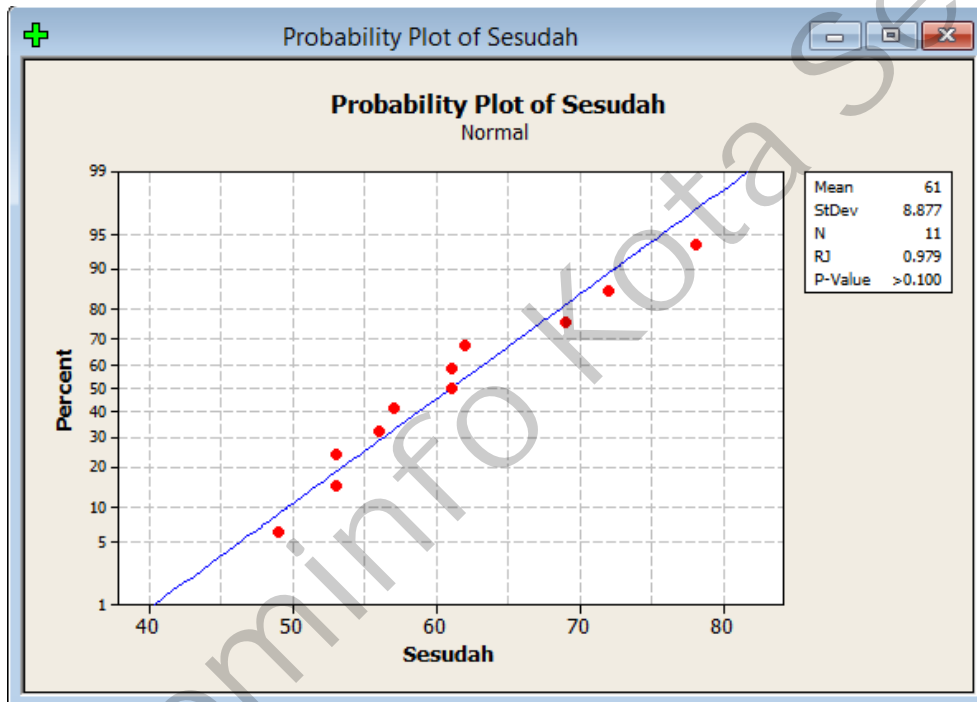
μ_2 = mean berat badan setelah meminum obat

Uji Normalitas (Berat Badan Sebelum Meminum Obat)



- Hipotesis :
H0: data berdistribusi normal
H1: data tidak berdistribusi normal
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = $> 0,100$
- Daerah kritik : H0 ditolak jika p-value $< \alpha$
- Kesimpulan :
Karena p-value ($> 0,100$) $> \alpha (0,05)$, maka H0 tidak ditolak, sehingga dapat disimpulkan bahwa data berdistribusi normal.

Uji Normalitas (Berat Badan Sesudah Meminum Obat)



- Hipotesis :
H0: data berdistribusi normal
H1: data tidak berdistribusi normal
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = $> 0,100$
- Daerah kritik : H0 ditolak jika p-value $< \alpha$
- Kesimpulan :
Karena p-value ($> 0,100$) $> \alpha (0,05)$, maka H0 tidak ditolak, sehingga dapat disimpulkan bahwa data berdistribusi normal

Uji Hipotesis Rata-rata Dua Populasi Dependen

Paired T-Test and CI: Sebelum, Sesudah				
Paired T for Sebelum - Sesudah				
	N	Mean	StDev	SE Mean
Sebelum	11	64.6364	7.9407	2.3942
Sesudah	11	61.0000	8.8769	2.6765
Difference	11	3.63636	3.20227	0.96552
95% lower bound for mean difference: 1.88639				
T-Test of mean difference = 0 (vs > 0): T-Value = 3.77 P-Value = 0.002				

- Hipotesis :
 - H0: $\mu_1 - \mu_2 \leq 0$
 - H1: $\mu_1 - \mu_2 > 0$
- Tingkat Signifikansi : $\alpha = 5\%$
- Statistik uji : p-value = 0,002
- Daerah kritik : H₀ ditolak jika p-value < α
- Kesimpulan :

Karena p-value (0,002) < α (0,05), maka H₀ ditolak, sehingga dapat disimpulkan bahwa obat pelangsing tersebut dapat menurunkan berat badan.

MODUL IV
ANALISIS VARIANSI

Uji mean yang dilakukan untuk k populasi ($k > 2$), memerlukan prosedur yang sedikit berbeda dengan 1 atau 2 populasi. Misal ada 4 populasi A,B,C dan D. Untuk mengetahui apakah ada perbedaan berarti diantara populasi tersebut, kita harus menguji apakah ada perbedaan antara pasangan AB,AC,AD,BC,BD, dan CD. Sehingga untuk kasus ini diperlukan 6 uji t atau uji Z yang terpisah. Hal ini sangat tidak efektif.

Oleh karena itu, diperlukan suatu metode yang tepat untuk mengatasi masalah tersebut, yaitu dengan menggunakan metode Analisis Variansi (ANOVA). Dalam kasus ini akan digunakan ANOVA 1 arah, yaitu analisis variansi yang hanya memperhitungkan satu faktor dalam penelitian. Jika ada 2 faktor maka disebut ANOVA 2 arah.

ANALISIS VARIANSI SATU ARAH

Contoh :

Akan diteliti pengaruh dari pemberian beberapa jenis obat yang berbeda terhadap lama kesembuhan pasien di salah satu rumah sakit. Jenis obat yang digunakan adalah obat A,B,C dan D.

- Faktor (variable independen) : Jenis Obat
- Level Faktor : A,B,C dan D
- Perlakuan/Level Faktor : A,B,C, dan D
- Variabel dependen : Lama Kesembuhan.

Tujuan ANOVA Satu Arah :

- Membandingkan rata-rata (mean) dari beberapa populasi (lebih dari 2)
- Melihat efek suatu faktor terhadap variable dependen

Asumsi dalam ANOVA

1. Data berdistribusi normal
2. Data mempunyai variansi yang relatif sama
3. Data bersifat independen

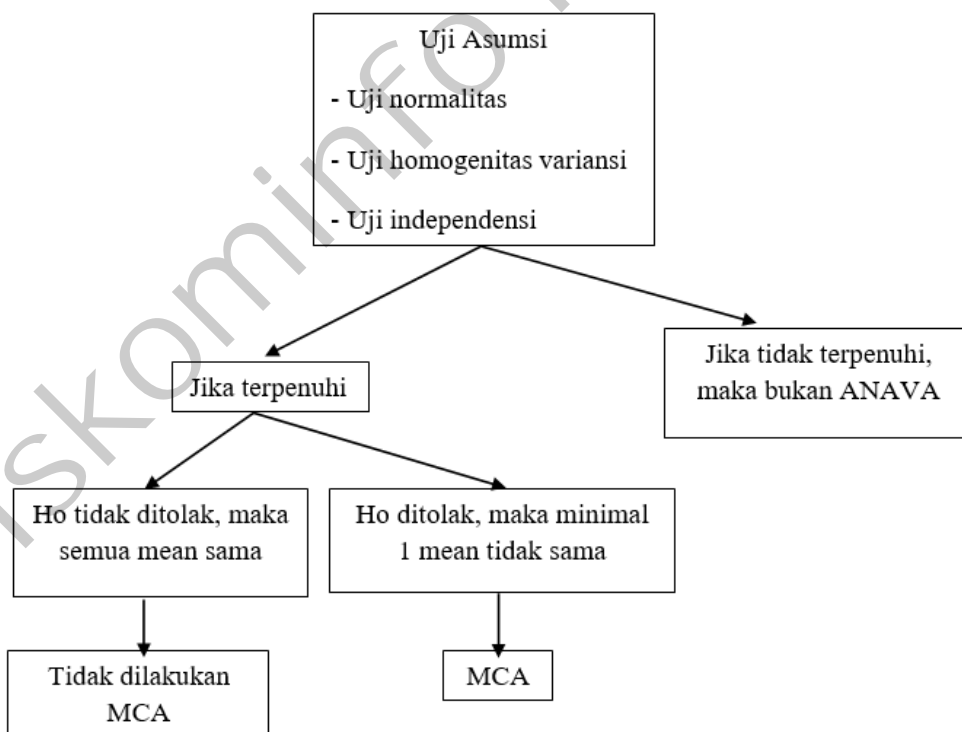
Jika asumsi di atas tidak terpenuhi, maka ada prosedur khusus yang harus dilakukan. Misal jika data yang dianalisis tidak berdistribusi normal, penanganan yang dapat dilakukan adalah dengan mentransformasikan data. Selain dengan transformasi data, cara lain ialah

dengan melakukan uji non-parametrik, seperti Mann-whitney, Wilcoxon, signed-Rank test, Kruskal Wallis, dll.

Analisis Perbandingan Ganda (Multiple Comparison Analysis)

1. Jika dalam ANAVA H_0 tidak ditolak, maka analisis selesai dengan kesimpulan semua rata-rata relatif sama.
2. Jika dalam ANAVA H_0 ditolak, maka selanjutnya kita lihat rata-rata populasi mana yang benar-benar berbeda dengan cara menggunakan Analisis Perbandingan Ganda atau MCA.
3. Syarat MCA ialah jumlah perlakuan lebih dari 2
4. Berbagai macam metode MCA :
 - Tukey
 - Bonferroni
 - Scheffe
 - Fisher (LSD=Least Square Different) : yang paling umum dilakukan

Bagan Langkah-Langkah Uji ANAVA 1 Arah



Contoh :

Dari sampel random yang terdiri dari 10 batang tembakau T1, 9 batang tembakau T2, dan 8 batang tembakau T3, diketahui kadar nikotin sebagai berikut (dalam mg).

Tembakau jenis manakah yang memiliki kadar nikotin paling tinggi?

T1	T2	T3
19	27	32
18	34	37
22	36	44
18	31	42
24	30	40
11	28	36
14	27	30
17	33	29
16	26	
20		

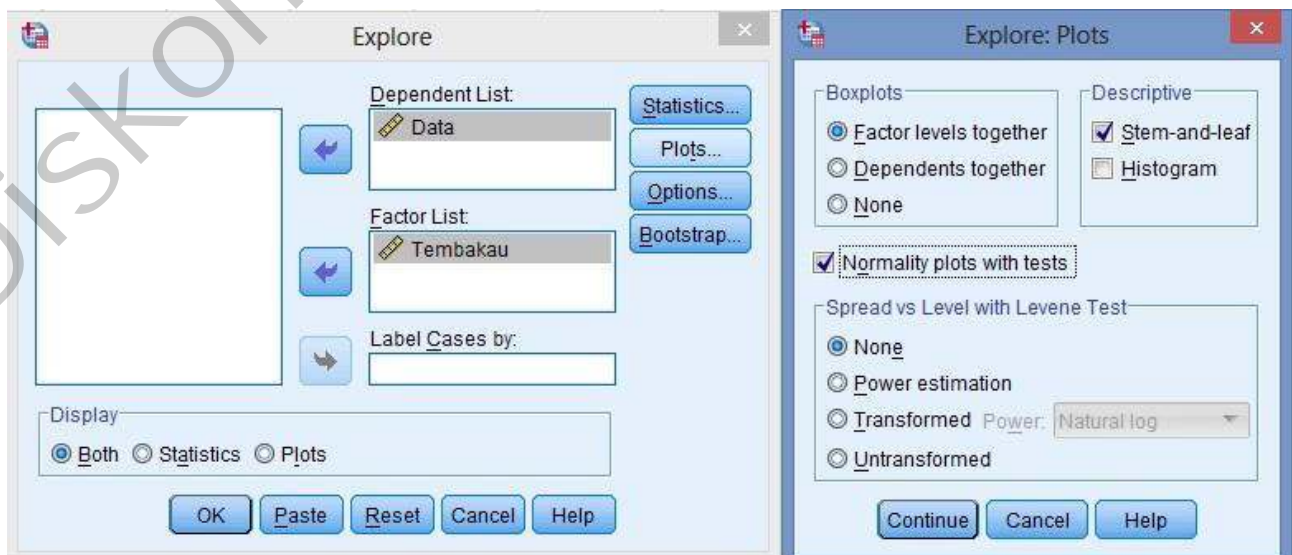
Penyelesaian :

1. Uji Asumsi

a. Uji Normalitas

Analyze → Descriptive Statistics → Explore

- Masukkan data yang kita miliki ke dalam kotak Dependent List. Masukkan perlakuan ke factor list.
- Klik Plot.
Pilih normality plot with test dan power estimation.



- Hipotesis
 - H_0 : Sampel tersebut diambil dari populasi yang berdistribusi normal.
 - H_1 : Sampel tersebut diambil dari populasi yang berdistribusi tidak normal.
- Tingkat signifikansi
 - $\alpha = 0.05$
- Statistik Uji

Tests of Normality

	Tembakau	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Data	1.00	.111	10	.200*	.987	10	.992
	2.00	.180	9	.200*	.932	9	.498
	3.00	.152	8	.200*	.944	8	.651

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Pada kasus ini, P-value yang digunakan adalah P-value pada kolom Shapiro-Wilk karena banyak data pada masing-masing perlakuan kurang dari atau sama dengan 50.

P-value tembakau 1 = 0,992

P-value tembakau 2 = 0,498

P-value tembakau 3 = 0,651

- Daerah Kritik
 - H_0 ditolak jika $P\text{-value} < \alpha$

Bahan tembakau 1

$0,992 > 0,05$

Bahan tembakau 2

$0,498 > 0,05$

Bahan tembakau 3

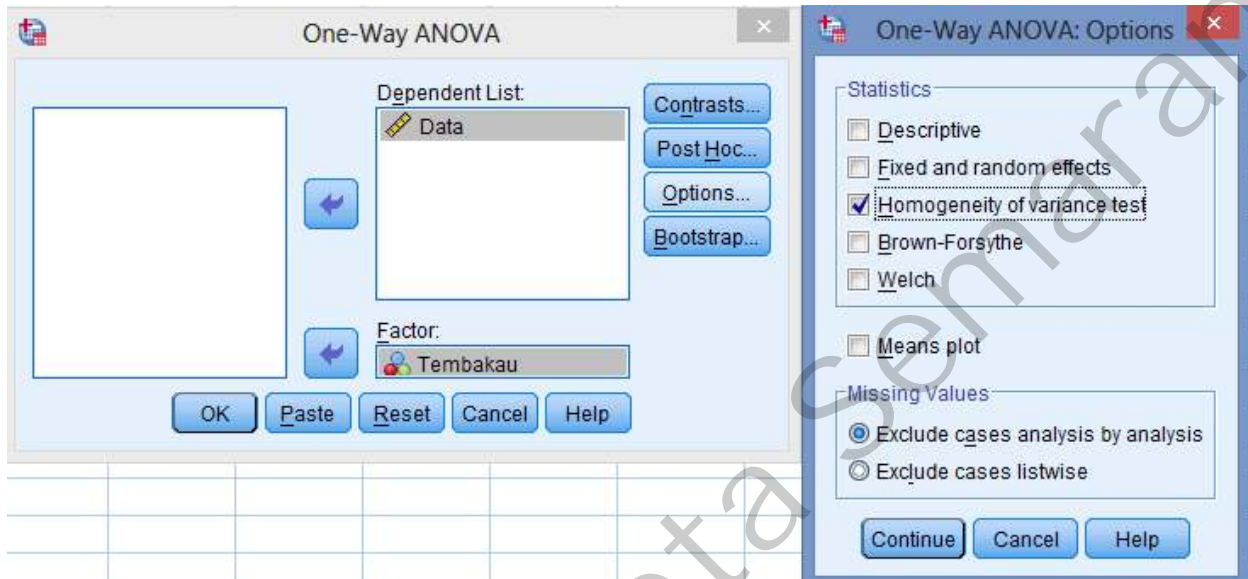
$0,651 > 0,05$

- Kesimpulan

Karena $P\text{-value} > \alpha$, maka H_0 tidak ditolak. Oleh karena itu, dapat disimpulkan bahwa sampel diambil dari populasi yang berdistribusi normal.

Analyze → Compare Means → One way ANOVA

- Masukkan data yang kita miliki ke dalam kotak Dependent List.
- Masukkan variabel group ke dalam kotak Factor.
- Klik Options dan pilih Homogeneity of variance test.
- Klik Continue → OK.



- Uji Hipotesis

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 \text{ (variansinya sama)}$$

$$H_1: \text{tidak semua } \sigma_i^2$$
- Tingkat Signifikansi
 $\alpha = 0,05$
- Statistik Uji
 sama (variansinya tidak sama)

Test of Homogeneity of Variances

Data

Levene	df1	df2	Sig.
Statistic			
1.480	2	24	.248

P-value = 0,248

- Daerah Kritik
 H_0 ditolak jika $P\text{-value} < \alpha$
 Karena $0,248 > 0,05$ maka H_0 tidak ditolak

- Kesimpulan

Karena H_0 tidak ditolak maka variansi dari ketiga jenis tembakau tersebut sama dengan tingkat signifikansi 0,05.

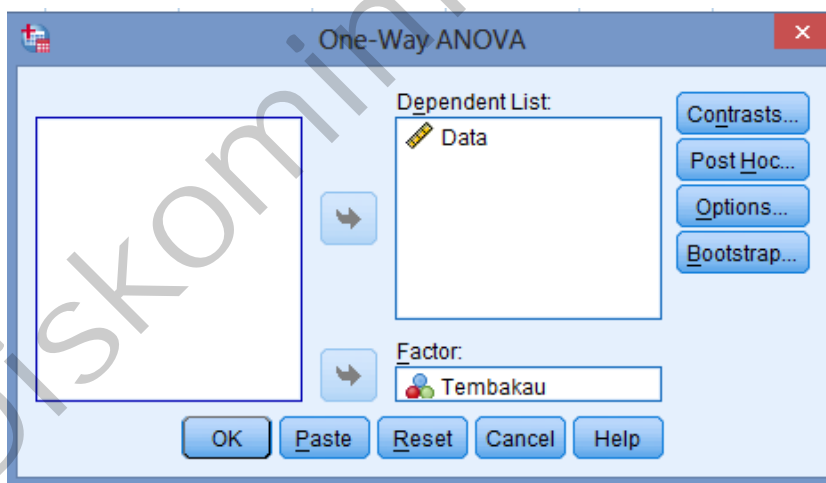
b. Uji Independensi

Asumsi ini dapat diupayakan untuk dipenuhi dengan melakukan randomisasi (keacakan) pada saat pengambilan sampel. Namun, perlu diingat bahwa randomisasi sebenarnya tidak menjamin asumsi keindependenan observasi pasti akan terpenuhi. Tetapi randomisasi memperkecil kemungkinan terjadinya dependensi pada setiap observasi. Di samping itu, ketika data yang kita peroleh adalah data sekunder kita harus pastikan bahwa si peneliti memperoleh datanya dengan menggunakan prinsip keacakan. Bila keacakan data tidak dapat dipastikan, maka kita tidak dapat melakukan analisis variansi karena dapat menghasilkan kesimpulan yang keliru.

2. Uji ANAVA 1 Arah

Analyze → Compare Means → One Way ANOVA

- Masukkan data yang kita miliki kedalam kotak Dependent List.
- Masukkan variabel grup ke dalam Factor.
- Klik OK.



Uji hipotesis ANAVA 1 arah

- Hipotesis

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

H_1 : Tidak semua μ_k sama

- Tingkat signifikansi

$$\alpha = 0,05$$

- Statistik uji

ANOVA

Data

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1600.711	2	800.356	43.267	.000
Within Groups	443.956	24	18.498		
Total	2044.667	26			

$$P\text{-value} = 0,000$$

- Daerah kritik

H_0 ditolak jika $P\text{-value} < \alpha$

Karena $0,000 < 0,05$ maka H_0 ditolak

- Kesimpulan

Karena H_0 ditolak, maka dapat disimpulkan bahwa terdapat perbedaan kadar nikotin pada ketiga jenis tembakau tersebut.

3. Uji MCA

Karena terdapat perbedaan pada rata-rata populasi maka akan dilakukan uji MCA.

Berikut adalah langkah-langkahnya :

Analyze → Compare Means → One Way ANOVA

- Masukkan data yang kita miliki kedalam kotak Dependent List
- Masukkan variabel grup ke dalam Factor.
- Klik Post Hoc.

Pada kotak dialog ini akan muncul beberapa metode MCA yang dapat digunakan.

- Untuk kali ini, akan kita gunakan metode MCA LSD.
- Klik OK.



Multiple Comparisons

Data

LSD

(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Tembakau 1.00	Tembakau 2.00	-12.32222*	1.97615	.000	-16.4008	-8.2437
	Tembakau 3.00	-18.35000*	2.04012	.000	-22.5606	-14.1394
2.00	1.00	12.32222*	1.97615	.000	8.2437	16.4008
	3.00	-6.02778*	2.08988	.008	-10.3411	-1.7145
3.00	1.00	18.35000*	2.04012	.000	14.1394	22.5606
	2.00	6.02778*	2.08988	.008	1.7145	10.3411

*. The mean difference is significant at the 0.05 level.

- Uji Hipotesis
 - H_0 : tidak ada perbedaan yang signifikan
 - H_1 : ada perbedaan yang signifikan
- Tingkat Signifikansi
 - $\alpha=0,05$
- Statistik Uji

Modul Analisis Data Statistik

Perbandingan Jenis Tembakau	P-value
1 dengan 2	0,000
1 dengan 3	0,000
2 dengan 3	0,008

- Daerah Kritik

H_0 ditolak jika

$P\text{-value} < \alpha$

Perbandingan Jenis Tembakau	P-value	Daerah Kritik	Kesimpulan	Mean Differences	Perbandingan
1 VS 2	0,000	H_0 ditolak	Ada perbedaan signifikan	-12,32222	1<2
1 VS 3	0,000	H_0 ditolak	Ada perbedaan signifikan	-18,35000	1<3
2 VS 3	0,008	H_0 ditolak	Ada perbedaan signifikan	-6,02778	2<3

- Kesimpulan

Dari tabel di atas dapat kita simpulkan bahwa $3 > 2 > 1$. Oleh karena itu, dapat disimpulkan bahwa jenis tembakau yang paling tinggi kadar nikotinnya adalah jenis tembakau 3.

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 atau fungsional. Untuk menetapkan kedua variabel mempunyai hubungan kausal atau tidak, maka harus didasarkan pada teori atau konsep-konsep tentang dua variabel tersebut.

Analisis regresi merupakan suatu analisis antara dua variabel yaitu variabel independent(predictor) dan variabel dependen(respon) dimana predictor diasumsikan mempengaruhi respon secara linear, sehingga variabel respon dapat diduga dari variabel predictor. Misalnya jika kita ingin mengetahui hubungan antara biaya promosi dengan hasil penjualan suatu produk, maka kita dapat menduga hasil penjualan melalui analisis regresi jika biaya promosi telah diterapkan.

REGRESI LINEAR SEDERHANA

Regresi sederhana didasarkan pada hubungan fungsional ataupun kausal satu variabel independen (X) dengan satu variabel dependen (Y).

Persamaan umum regresi linear sederhana adalah: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i=1,2,\dots,n$

Dimana:

Y_i = Subyek dalam variabel dependen yang diprediksikan

β_0 = Harga Y_i bila $X=0$ (harga konstan)

β_1 = Angka arah atau koefisien regresi yang menunjukkan angka peningkatan ataupun penurunan variabel dependen yang didasarkan pada variabel independen

X_i = Subyek pada variabel independen yang mempunyai nilai tertentu

β_0 (konstan/intersep) dan β_1 (gradient/slope) adalah parameter yang tidak diketahui nilainya dan akan diestimasi dengan statistik b_0 dan b_1 . Estimator b_0 dan b_1 dicari menggunakan metode Least Square Error (LSE) untuk meminimalkan jumlah kuadrat error.

Harga b_1 merupakan fungsi dari koefisien korelasi. Bila koefisien bernilai tinggi, maka harga b_1 juga besar, sebaliknya bila koefisien bernilai rendah maka harga b_1 juga rendah (kecil). Selain itu bila koefisien korelasi negatif maka harga b_1 juga negatif, dan sebaliknya bila koefisien korelasi positif maka harga b_1 juga positif.

Asumsi-asumsi Regresi

- Data variabel dependen (respon) berasal dari populasi berdistribusi normal.
- Linearitas (hubungan antara variabel independen dengan variabel dependen harus linear).

ANALISIS KORELASI

Korelasi digunakan untuk mencari hubungan dan membuktikan hipotesis hubungan dua variabel bila data kedua variabel berbentuk interval atau ratio, dan sumber data dari dua variabel atau lebih adalah sama. Berikut ini dikemukakan rumus yang paling sederhana yang dapat digunakan untuk menghitung koefisien korelasi, yaitu rumus 1.0. Koefisien korelasi untuk populasi diberi symbol rho (ρ) dan untuk sampel diberi simbol r , sedang untuk korelasi ganda diberi simbol R .

Rumus 1.0

$$r_{xy} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\{n \sum x_i^2 - (\sum x_i)^2\} \{n \sum y_i^2 - (\sum y_i)^2\}}}$$

Untuk dapat memberikan penafsiran terhadap koefisien korelasi yang ditemukan besar atau kecil, maka dapat berpedoman pada ketentuan yang tertera pada tabel di bawah ini.

Interpretasi terhadap koefisien korelasi

Interval Koefisien	Tingkat Hubungan
0,00 – 0,199	Sangat rendah
0,20 – 0,399	Rendah
0,40 – 0,599	Sedang
0,60 – 0,799	Kuat
0,80 – 1,00	Sangat kuat

Dalam analisis korelasi terdapat suatu angka yang disebut dengan koefisien determinasi, yang besarnya adalah kuadrat dari koefisien korelasi (r^2). Koefisien ini disebut penentu, karena varian yang terjadi pada variabel dependen dapat dijelaskan melalui varian yang terjadi pada variabel dependen.

Studi Kasus :

Berikut ini merupakan data penelitian tentang massa otot seseorang dan umur dalam suatu populasi masyarakat. Apakah banyaknya umur berpengaruh terhadap massa otot seseorang? Ceklah asumsi-asumsi yang diperlukan, tulis model regresi yang anda susun, kemudian lakukan analisis regresi untuk menemukan model terbaik, dan berikan interpretasinya.

PENYELESAIAN

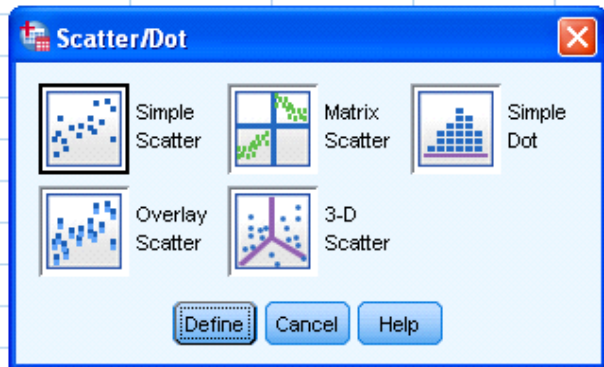
- Masukkan data ke dalam SPSS dengan format sebagai berikut :

	massa	umur
1	82	71
2	91	64
3	100	43
4	68	67
5	87	56
6	73	73
7	78	67
8	80	56
9	65	76
10	84	65
11	116	45
12	76	58
13	97	45
14	100	53
15	105	49
16	77	78
17		

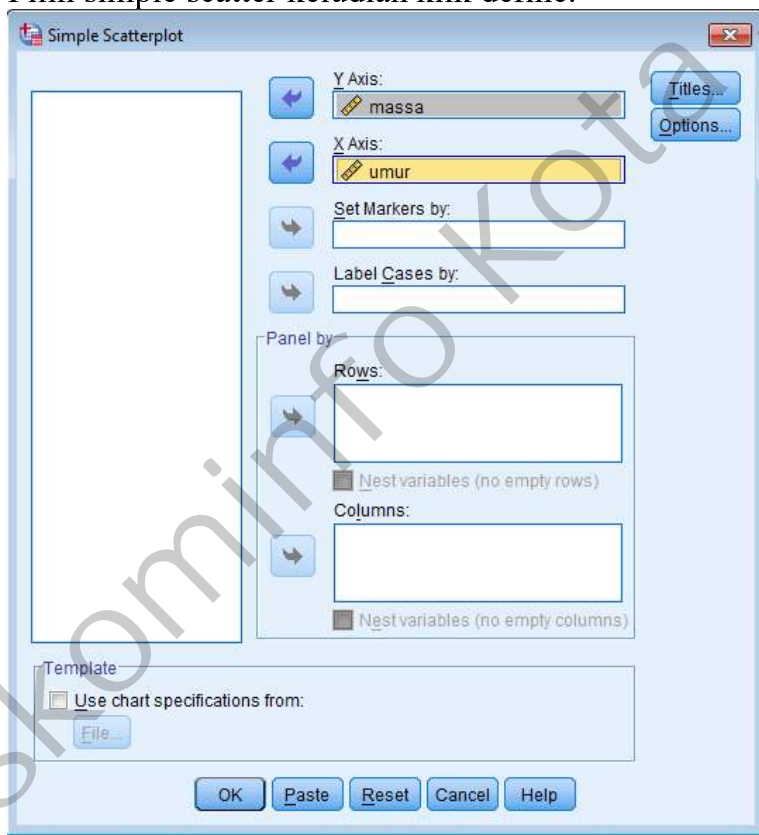
2. UJI ASUMSI

- **LINIERITAS DATA** (Scatter antara X dan Y)

Klik graphs → legacy dialogs → scatter/dot dan akan muncul kotak dialog seperti gambar dibawah ini :

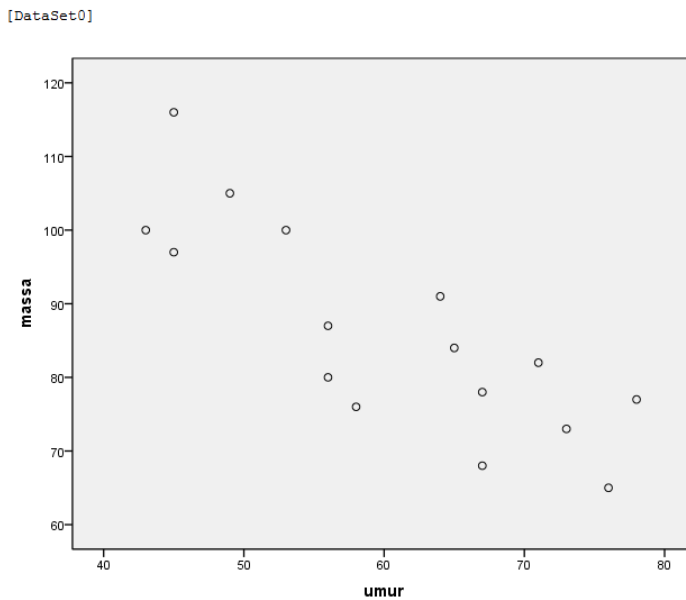


Pilih simple scatter kemudian klik define.



Masukkan variabel 'berat_badan' pada **Y axis** dan variable 'umur' pada **X axis** lalu klik OK.

• Output



Akan muncul hasilnya pada lembar output, untuk mendapatkan garis linear kita harus double klik pada graph di atas sehingga muncul kotak dialog seperti di bawah ini.

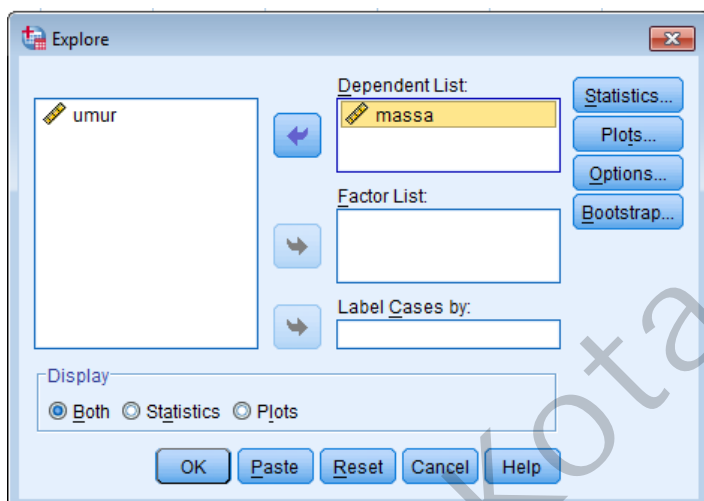
Klik icon yang dilingkari garis merah, kemudian pada kotak dialog properties pilih linier kemudian klik apply lalu klik close.

Interpretasi :

Karena terdapat garis lurus dan terdapat titik – titik disekitar garis tersebut sehingga disimpulkan terdapat hubungan linier (negatif) antara variable dependen (massa otot) dan variable independen (umur). Semakin besar nilai X (umur) maka nilai Y (massa otot) akan semakin kecil.

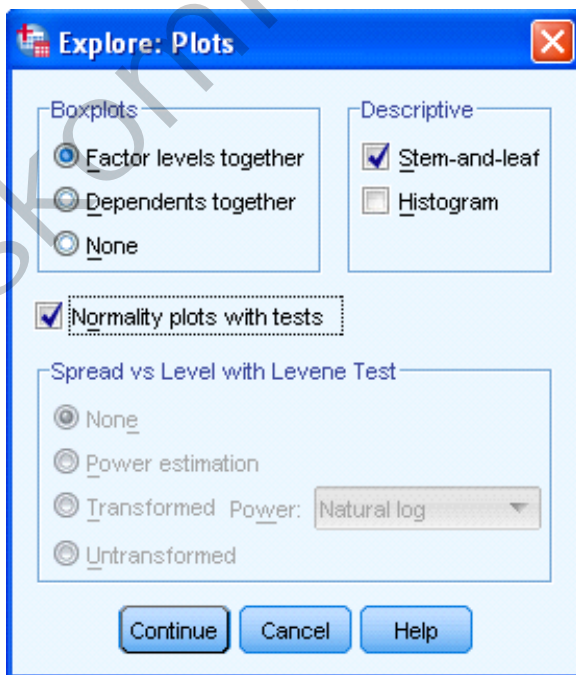
- **UJI NORMALITAS VARIABLE DEPENDEN**

Klik analyze > descriptive statistics > explore



Masukkan massa (response) pada kotak dependent list.

Kemudian klik tab plots dan akan muncul kotak dialog seperti gambar dibawah ini.



Check list pada normality plots with tests. Kemudian klik Continue. Pada kotak dialog explore klik OK.

Kemudian didapatkan output sebagai berikut :

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
massa	.124	16	.200 [*]	.964	16	.730

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Uji normalitas variable dependen (massa)

1. Hipotesis

H₀: Data berdistribusi normal

H₁: Data berdistribusi tidak normal

2. Tingkat signifikansi = $\alpha = 0,05$

3. Uji Statistik

P-value

Didapat dari SPSS P-value = 0,730

(kita lihat pada Sig. Shapiro-Wilk karena data kita ≤ 50)

4. Daerah kritik

H₀ ditolak jika P_value < α

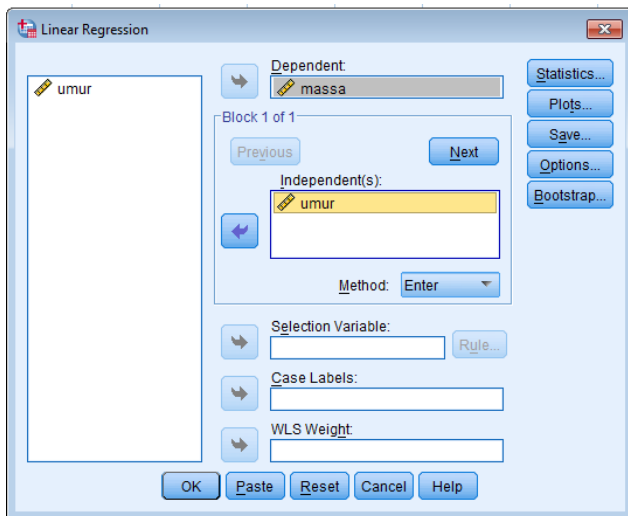
5. Kesimpulan

Karena P-value (0,730) > (0,05) maka H₀ tidak ditolak, jadi variable dependen (massa otot) berdistribusi normal.

Untuk kasus real dan apabila data tidak berdistribusi normal, maka dapat dilakukan transformasi data atau dapat pula dilakukan penambahan data.

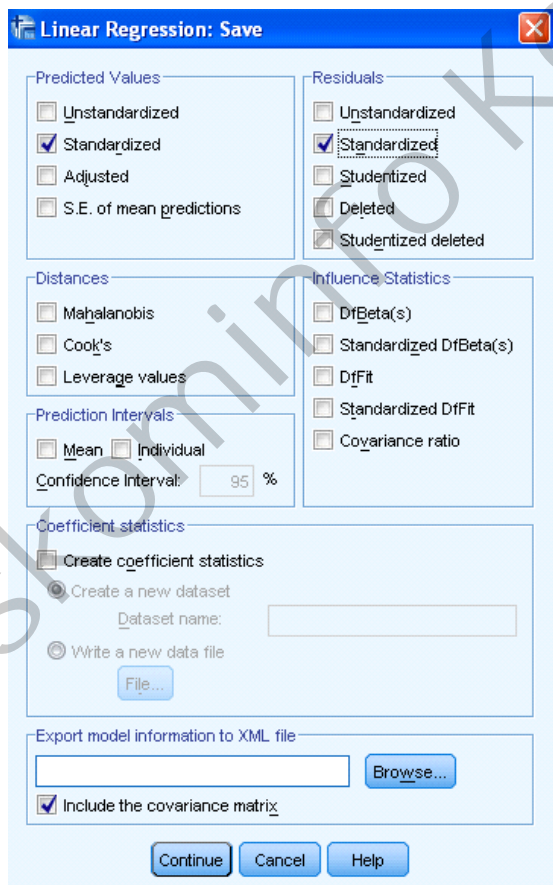
3. ANALISIS REGRESI LINIER SEDERHANA

Klik analyze > regression > linear



Masukkan massa pada kotak **dependent** dan umur pada kotak **independent**.

Pada tab **save** check list standardized pada predicted values dan residuals, seperti terlihat di bawah ini.



• **OUTPUT Model Summary**

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.824 ^a	.678	.655	8.352

a. Predictors: (Constant), umur
 b. Dependent Variable: massa

Interpretasi :

R = 0,824 menunjukkan derajat hubungan linear (negatif) yang cukup erat antara variabel respon Y(massa otot) dan variabel prediktor X(umur).

R Square = 0,678 menyatakan 67,8% variansi dalam variabel dependen (Y) dapat diterangkan oleh variabel independen (X).

Adjusted R Square tidak digunakan dalam analisis regresi linear sederhana.

St. Error = 8,352, ukuran besarnya variansi model regresi.

Persamaan regresi/ model regresi

$$y = \beta_0 + \beta_1 x$$

• **UJI OVERALL (Kelayakan Model)**

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2057.905	1	2057.905	29.503	.000 ^a
	Residual	976.532	14	69.752		
	Total	3034.438	15			

a. Predictors: (Constant), umur
 b. Dependent Variable: massa

1. Hipotesis

$H_0 : \beta_i = 0$ (model regresi tidak layak) ($i=0,1$)

$H_1 : \beta_i \neq 0$ (model regresi layak)

2. Tingkat signifikansi

$$\alpha = 0,05$$

3. Uji Statistik

P-value

Didapat dari spss P-value = 0,000

4. Daerah kritik

H_0 ditolak jika $P\text{-value} < 0,05$

5. Kesimpulan

Karena $P\text{-value} (0,000) < (0,05)$ maka H_0 ditolak, jadi model regresi layak digunakan.

• **UJI PARSIAL (Uji Kelayakan Koefisien)**

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	148.183	11.603		12.771	.000
	umur	-1.027	.189	-.824	-5.432	.000

a. Dependent Variable: massa

Uji parsial untuk koefisien X

1. Hipotesis

$H_0 : \beta_1 = 0$ (koefisien X tidak signifikan)

$H_1 : \beta_1 \neq 0$ (koefisien X signifikan)

2. Tingkat signifikansi

$\alpha = 0,05$

3. Uji Statistik

P-value

Didapat dari SPSS $P\text{-value} = 0,000$

4. Daerah kritis

H_0 ditolak jika $P\text{-value} < 0,05$

5. Kesimpulan

Karena $P\text{-value} (0,000) < (0,05)$ maka H_0 ditolak, jadi koefisien X signifikan.

Uji parsial untuk konstan

1. Hipotesis

$H_0 : \beta_0 = 0$ (konstan tidak signifikan)

$H_1 : \beta_0 \neq 0$ (konstan signifikan)

2. Tingkat signifikansi

$\alpha = 0,05$

3. Uji Statistik

P-value

Didapat dari SPSS p-value = 0,000

4. Daerah kritik

H_0 ditolak jika P-value < 0,05

5. Kesimpulan

Karena P-value (0,003) < (0,05) maka H_0 ditolak, jadi konstan signifikan.

Interpretasi :

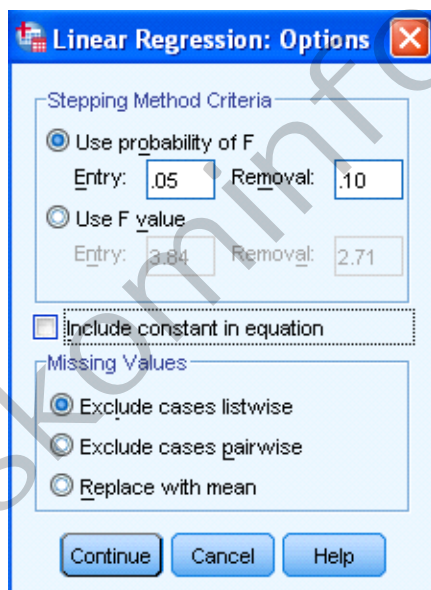
Dari tabel coefisien didapat β_0 yaitu 3,025 dan β_1 yaitu 0,507. Maka model regresi yang didapat yaitu $Y = 148,183 - 1,027 X$ atau **massa otot = 148,183 - 1,027 umur**

• **Tambahan :**

Jika ada variable konstan yang tidak signifikan maka dilakukan uji regresi lagi dengan mengeluarkan konstan.

Kita melakukan lagi analisis regresi linier sederhana.

Klik analyze > regression > linear namun pada tahap ini klik tab options lalu uncheck list pada include constant in equation. Klik continue lalu klik OK.



ANALISIS DATA KATEGORIK

TABEL KATEGORIK 2 x 2

1. Pengantar

Data kualitatif atau sering disebut data kategorikal adalah salah satu skala pengukuran yg terdiri atas sekumpulan kategori. Skala data yang dikategorikan data kualitatif adalah skala ordinal dan skala nominal.

Contoh data dengan skala ordinal adalah tingkat kepuasan konsumen (sangat tidak puas, tidak puas, biasa, puas, sangat puas), predikat kelulusan (memuaskan, sangat memuaskan, dan cumlaude), dan pendidikan terakhir (SD, SMP, SMA, Diploma, Sarjana, Magister, Doktor). Sedangkan contoh data dengan skala nominal adalah jenis kelamin (pria, wanita), jenis kendaraan angkutan barang (truk, pesawat terbang, kereta, kapal laut).

Data kategorik termasuk data kelas, tingkat, sering dimanfaatkan frekuensinya. Dalam inferensi sederhana juga telah dikenal analisis data kategorik dalam uji proporsi baik satu populasi maupun dua populasi. Pada topik kali ini kita akan mempelajari analisis data kategorik berupa dua data kategorik yang diklasifikasikan silang dalam bentuk tabel silang 2 x 2 atau ukuran yang lebih besar.

Dipunyai n observasi yang diklasifikasikan silang dalam dua variabel kategorik dan dihasilkan empat sell observasi seperti dalam desain di bawah ini :

Variabel A	Variabel B			
		B	Bc	Jumlah
	A	a	b	n1=a+b
	Ac	c	d	n2=c+d
	Jumlah	m1=a+c	m2=b+d	n=n1+n2

Ada dua macam uji hipotesis dalam analisis data kategorik tabulasi silang :

1. Uji Homogenitas (dua sampel)
2. Uji Independensi (satu sampel)

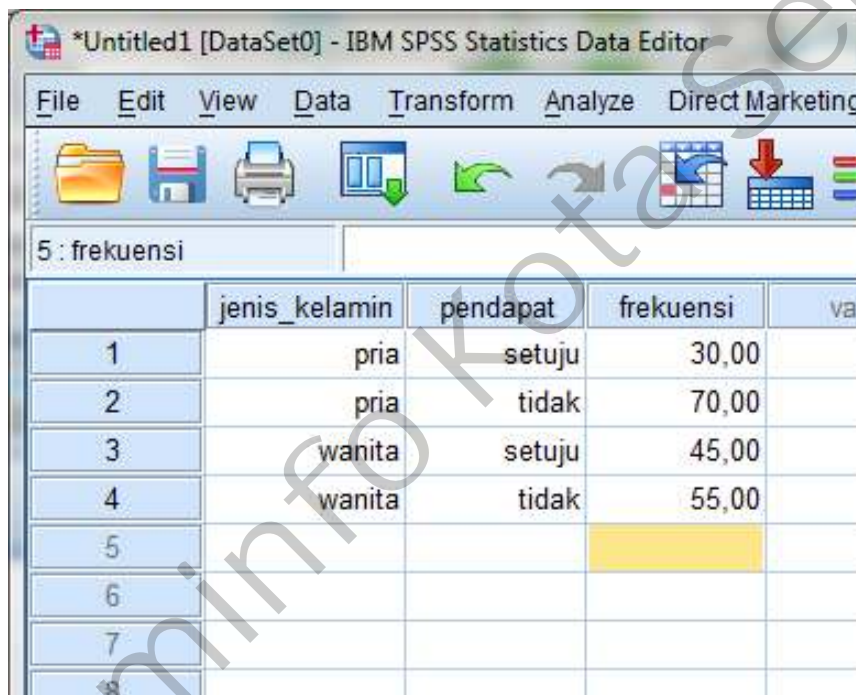
Contoh

Dua sampel random masing-masing terdiri dari 100 pria dan 100 wanita. Diajukan pertanyaan, "Apakah anda setuju wanita mempunyai hak dan kewajiban yang sama dengan pria?". Diperoleh tabulasi silang sebagai berikut:

Jenis Kelamin	Pendapat			
		Setuju	Tidak	Total
	Pria	30	70	100
	Wanita	45	55	100
	Total	75	125	200

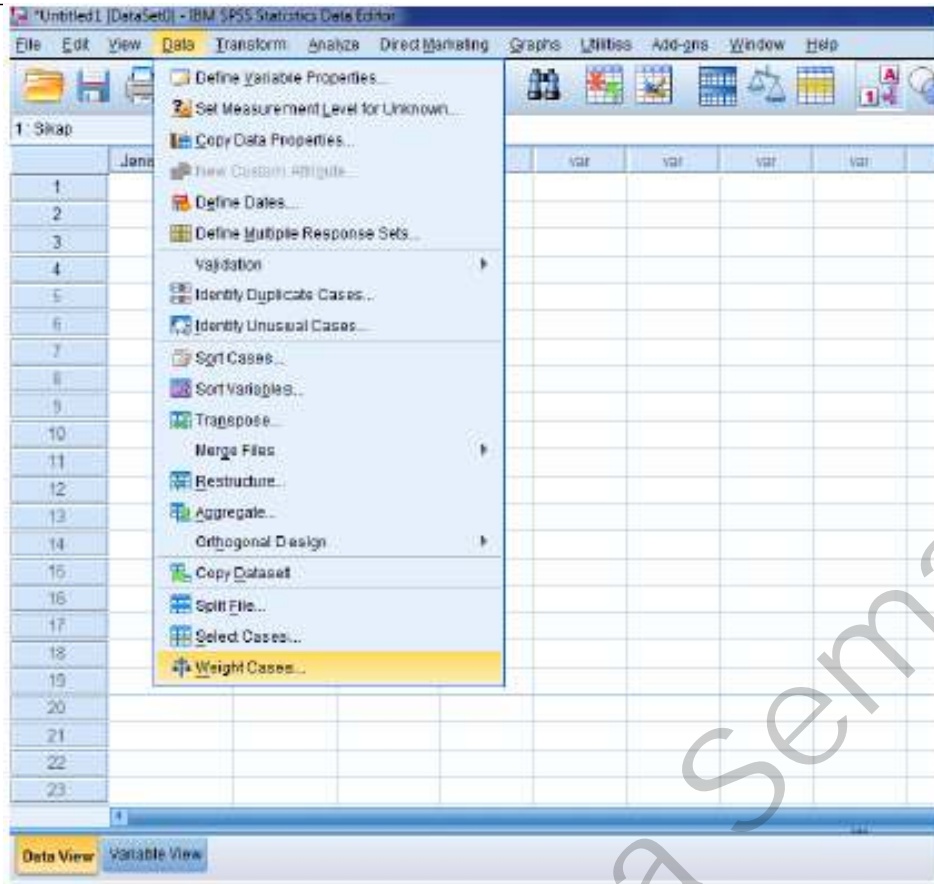
Uji Homogenitas dengan SPSS

1. Masukkan data dengan format seperti berikut ini :

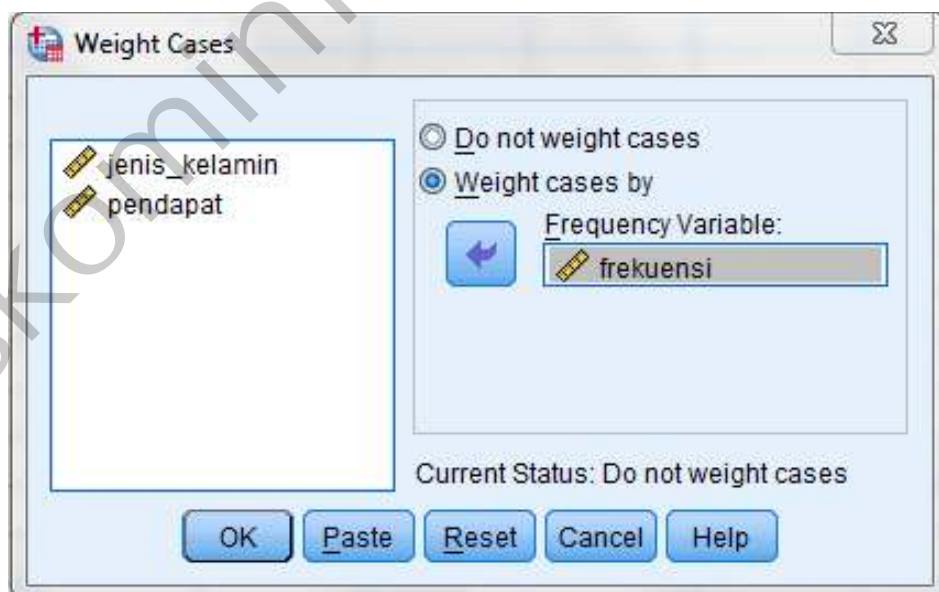


2. Karena dalam analisis tabulasi silang yang dipakai variabel baris dan variabel kolom, dalam hal ini variabel jenis kelamin dan variabel sikap, maka kedua variable tersebut perlu diberi bobot dengan frekuensi data yang ada. Caranya adalah klik menu **data** → **Weight Cases**.

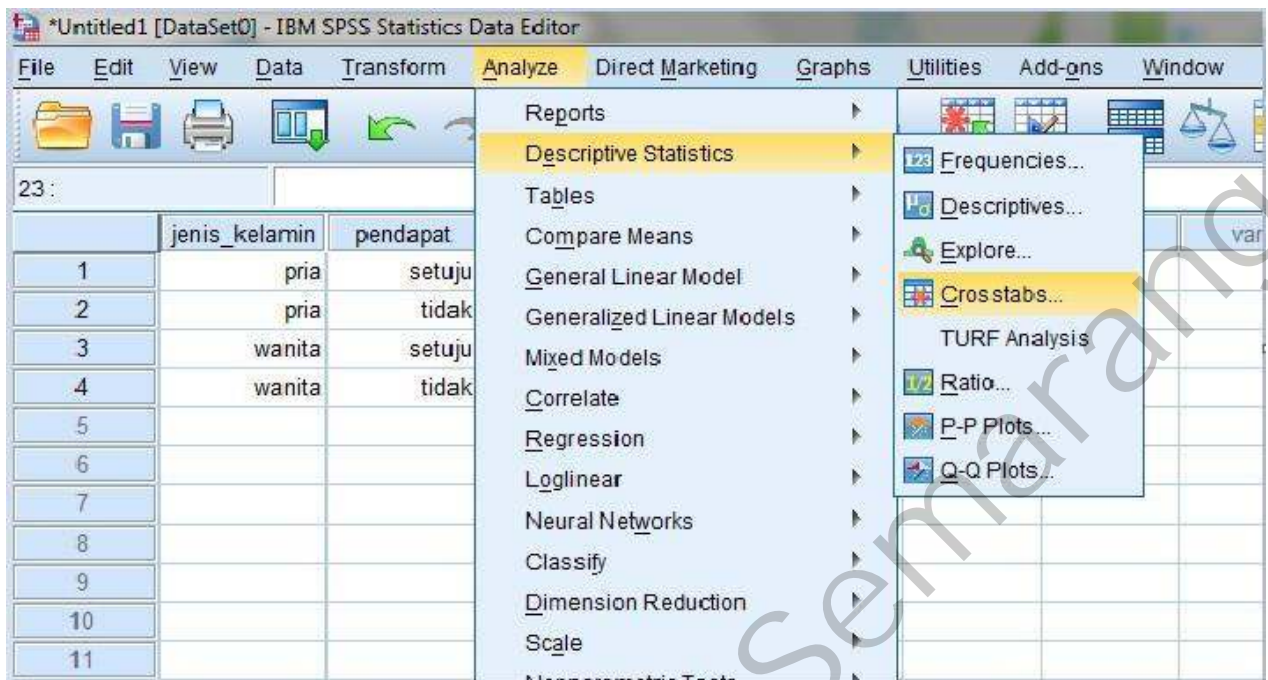
Modul Analisis Data Statistik



3. Pada kotak **Weight Cases** klik pilihan **Weight Cases by**, kemudian kotak **Frequency Variable** diisi dengan variabel **frekuensi**. Maksudnya adalah variable kategorik jenis kelamin dan sikap dibobot dengan frekuensi nya. Selanjutnya klik **OK**. Setelah di klik **OK**, sebagai tanda bahwa variabel sudah di bobot adalah akan muncul tulisan **Weight On** di pojok kiri bawah.



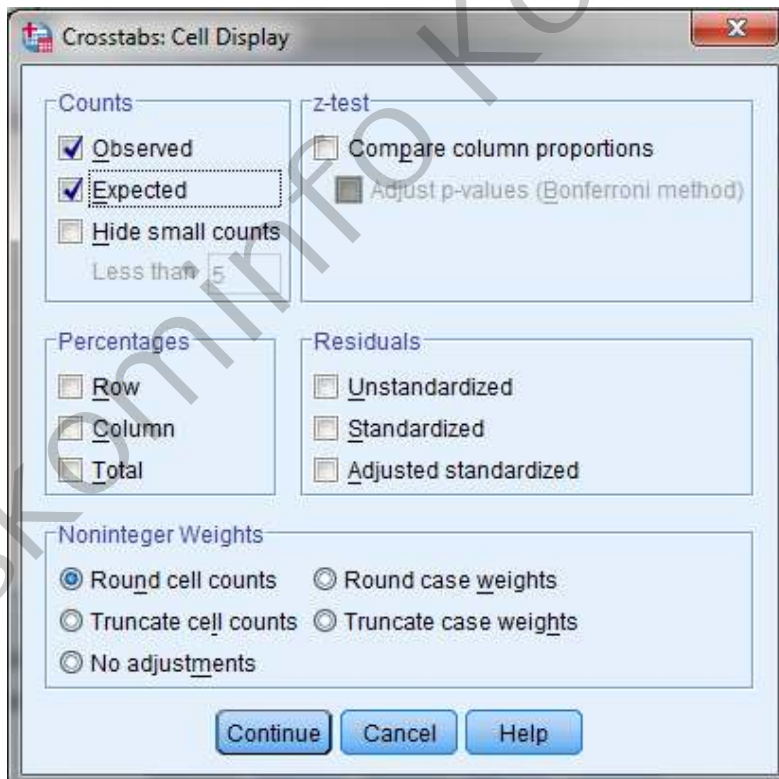
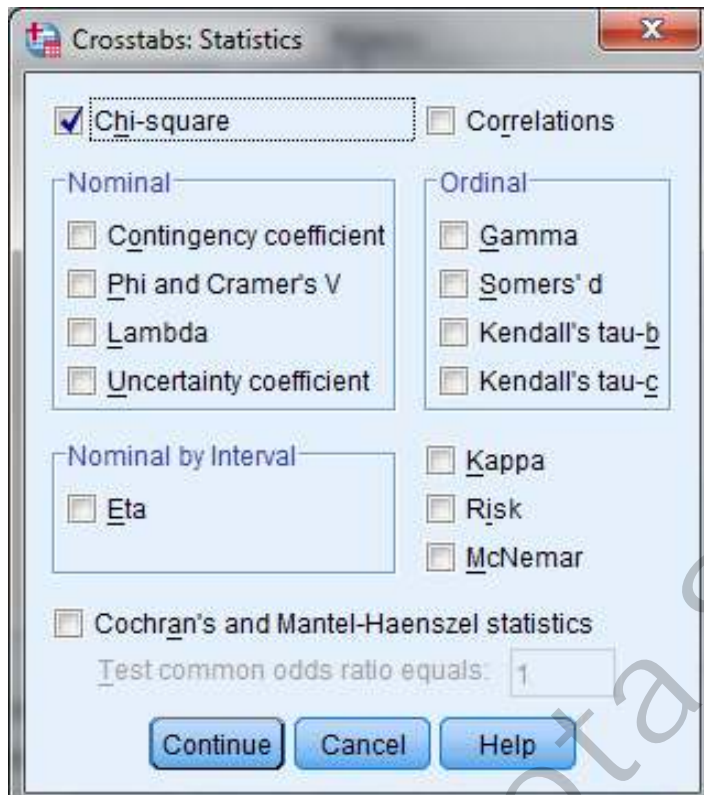
4. Data sudah siap dilakukan analisis tabulasi silang. Klik menu **Analyze** → **Descriptive Statistics** → **Crosstabs** seperti dibawah ini :



5. Masukkan variabel jenis kelamin ke kotak Row(s), variabel sikap ke kotak Column(s), seperti gambar di bawah ini :



6. Klik kotak **Statistics**, dan pilihlah kotak **Chi Square**. Klik lah kotak **Cells**, dan pilihlah kotak **Observed**, **Expected**. Kemudian klik **OK** pada menu utama Crosstabs.



7. Maka akan di dapatkan beberapa output, yaitu :

jenis_kelamin * pendapat Crosstabulation

			pendapat		Total
			setuju	tidak	
jenis_kelamin	pria	Count	30	70	100
		Expected Count	37,5	62,5	100,0
	wanita	Count	45	55	100
		Expected Count	37,5	62,5	100,0
Total		Count	75	125	200
		Expected Count	75,0	125,0	200,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4,800 ^a	1	,028		
Continuity Correction ^b	4,181	1	,041		
Likelihood Ratio	4,825	1	,028		
Fisher's Exact Test				,041	,020
Linear-by-Linear Association	4,776	1	,029		
N of Valid Cases	200				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 37,50.

b. Computed only for a 2x2 table

Pada output **jenis_kelamin * pendapat Crosstabulation** pada bagian Expected Count. Jika expected count < 5. Maka lihat output **Chi-Square Tests** pada bagian Fisher's Exact Test, dan lihat pada kolom Exact Sig. (2 – sided) dan lihat lah hasilnya. Maka itulah p-value nya.

Sedangkan bila expected count > 5. Maka lihat output **Chi-Square Tests** pada bagian Pearson Chi-Square, dan lihat pada kolom Asymp. Sig. (2 – sided) dan lihat lah hasilnya. Maka itulah p-value nya.

Pada kasus ini, expected count nya > 5, maka p-value nya adalah 0,028. Karena p- value (0,028) < α (0,05). Maka H0 ditolak, jadi dapat disimpulkan bahwa proporsi antara kedua populasi terhadap pendapat tentang kesamaan hak dan kewajiban antara wanita dengan pria adalah tidak homogen (tidak sama).

2. Uji Independensi

Pada uji Independensi, sampel berasal dari **satu** populasi. Jadi ditentukan terlebih dahulu besarnya ukuran sampel yaitu **n**, sedangkan **n₁** dan **n₂** diperoleh dari observasi sebagai bagian dari sampel. Tujuan dari uji Independensi adalah untuk mengetahui hubungan dependen variabel **A** dan variabel **B**.

Contoh

Pada suatu hari Qia mahasiswa Statistika UGM ingin meneliti apakah orang yang mengonsumsi sambal akan menyebabkan sakit perut. Pada saat makan siang, Qia menanyakan mahasiswa yang sedang makan siang di Kantin Kluster Sains. Didapatkan data sebagai berikut.

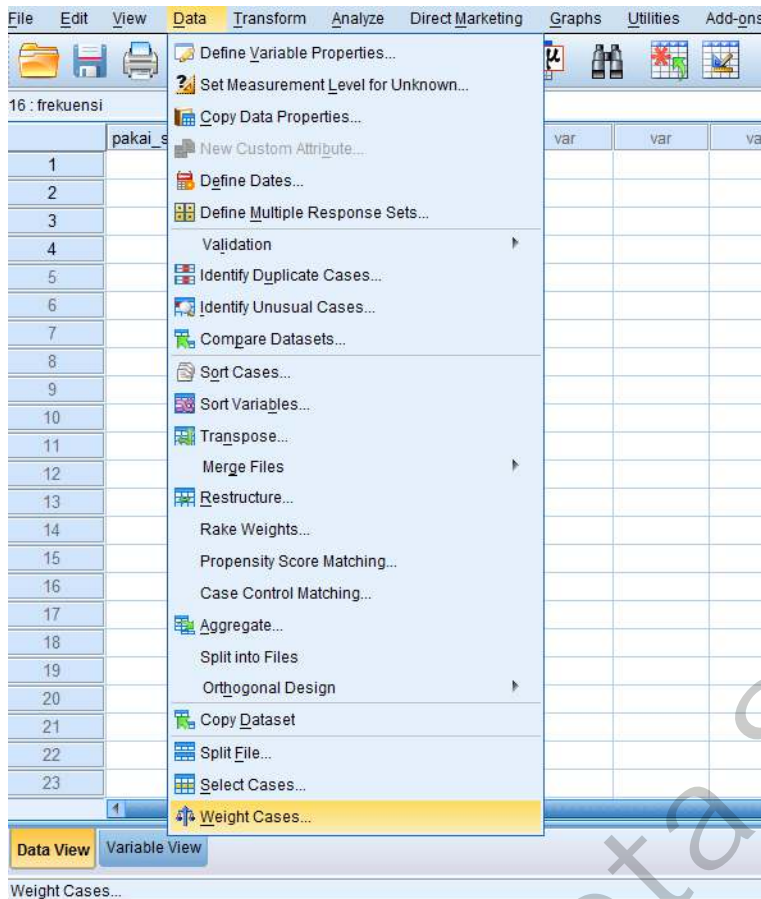
Memakai sambal	Sakit perut		Total
	Ya	Tidak	
Pakai	109	116	225
Tidak pakai	4	34	38
Total	113	150	263

Uji Independensi dengan SPSS

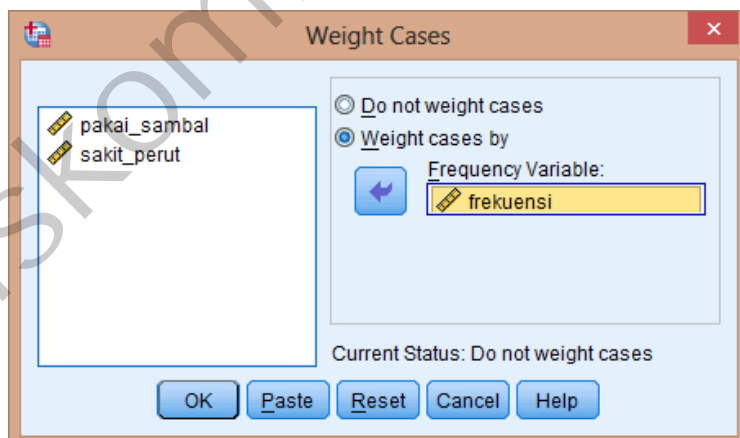
1. Masukkan data dengan format seperti berikut ini :

	pakai_sam...	sakit_perut	frekuensi	var	var
1	1,00	1,00	109,00		
2	1,00	2,00	116,00		
3	2,00	1,00	4,00		
4	2,00	2,00	34,00		
5					
6					

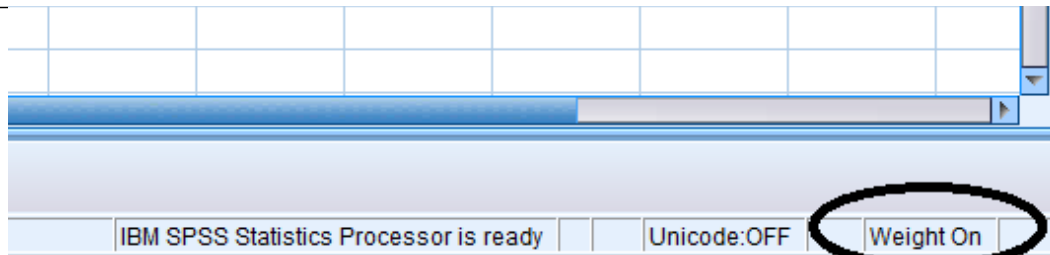
2. Karena dalam analisis tabulasi silang yang dipakai variabel baris dan variable kolom, dalam hal ini variabel **memakai sambal** dan **sakit perut**, maka kedua variabel tersebut perlu diberi bobot dengan frekuensi data yang ada. Caranya adalah klik menu **Data** → **Weight Cases**.



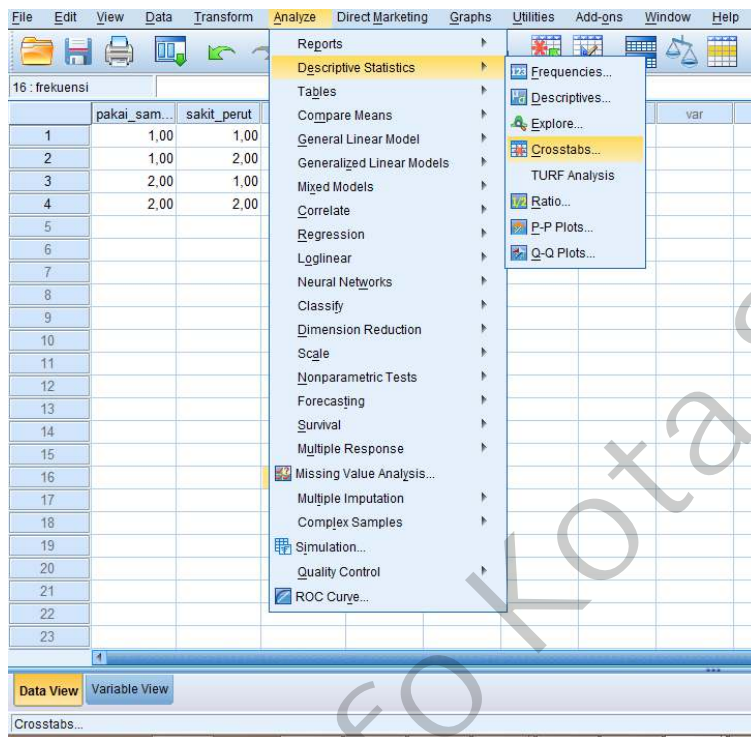
3. Pada kotak **Weight Cases** klik pilihan **Weight Cases by**, kemudian kotak **Frequency Variable** diisi dengan variabel **Frekuensi**. Maksudnya adalah variabel kategorik **memakai sambal** dan **sakit perut** dibobot dengan frekuensi nya. Selanjutnya klik **OK**. Setelah di klik **OK**, sebagai tanda bahwa variabel sudah di bobot adalah akan muncul tulisan **Weight On** di pojok kanan bawah.



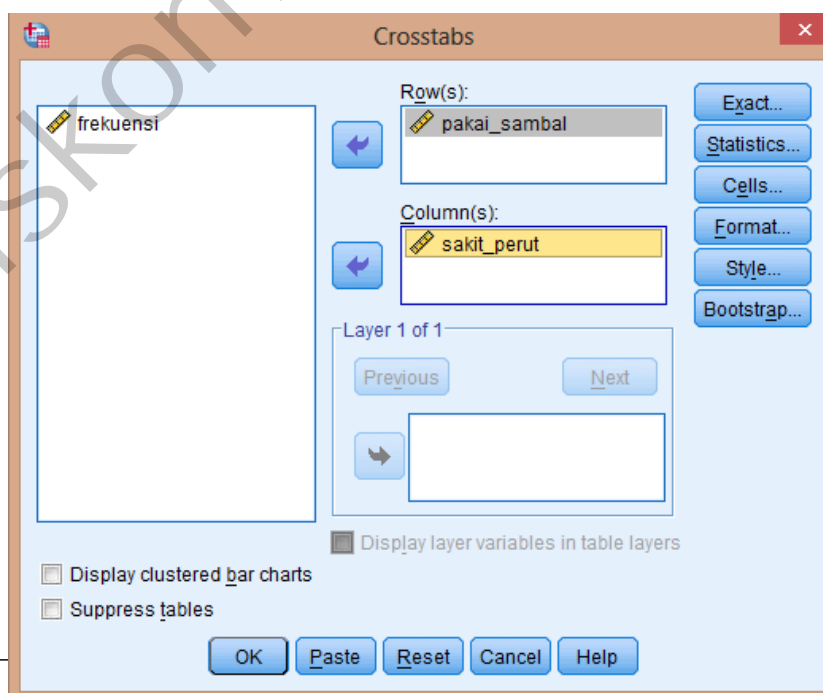
Modul Analisis Data Statistik



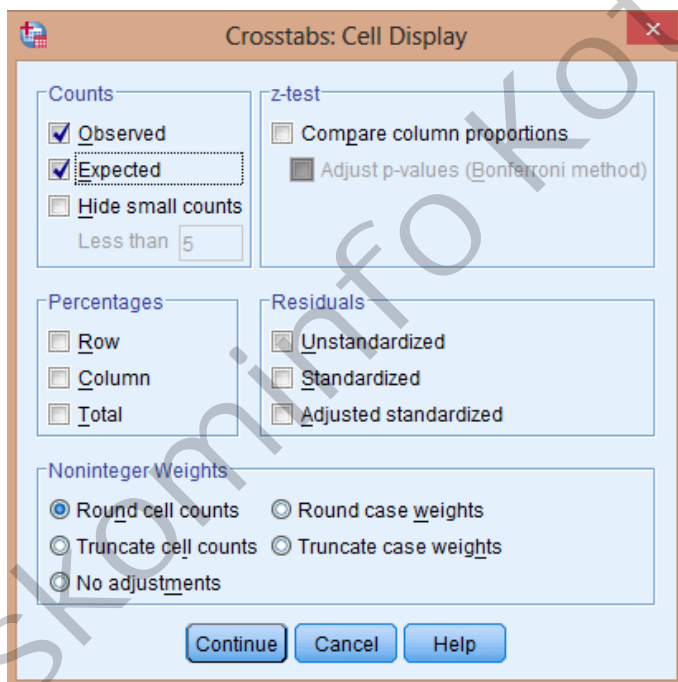
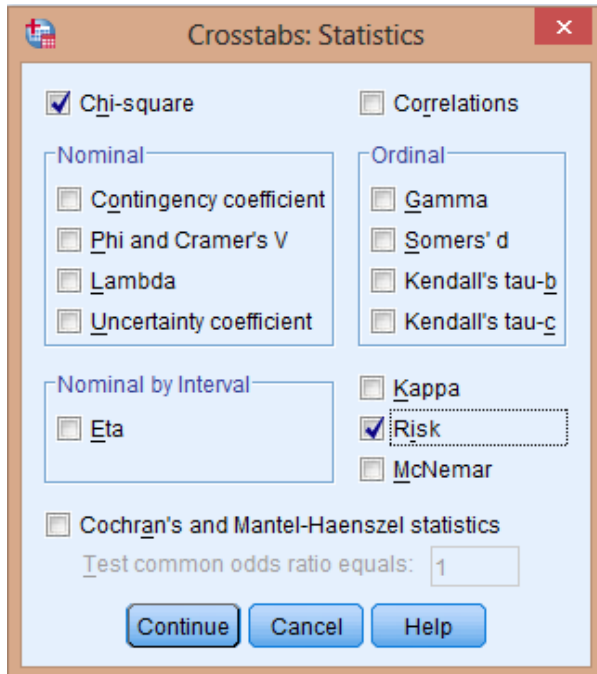
4. Data sudah siap dilakukan analisis tabulasi silang. Klik menu **Analyze** → **Descriptive Statistics** → **Crosstabs** seperti di bawah ini :



5. Masukkan variabel **pakai_sambal** ke kotak **Row(s)**, variabel **sakit_perut** ke kotak **Column(s)**, seperti gambar dibawah ini :



6. Klik kotak **Statistics**, dan pilihlah kotak **Chi Square** ->**Continue** . Kemudian, klik lah kotak **Cells**, dan pilihlah kotak **Observed**, **Expected** lalu setelah itu **Continue**. Kemudian klik **OK** pada menu utama **Crosstabs**.



7. Maka akan di dapatkan beberapa output, yaitu :

pakai_sambal * sakit_perut Crosstabulation

			sakit_perut		Total
			Ya	Tidak	
pakai_sambal	Ya	Count	109	116	225
		Expected Count	96,7	128,3	225,0
	Tidak	Count	4	34	38
		Expected Count	16,3	21,7	38,0
Total	Count	113	150	263	
	Expected Count	113,0	150,0	263,0	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	19,074 ^a	1	,000		
Continuity Correction ^b	17,558	1	,000		
Likelihood Ratio	22,101	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	19,002	1	,000		
N of Valid Cases	263				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 16,33.

b. Computed only for a 2x2 table

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for pakai_sambal (Ya / Tidak)	7,987	2,744	23,251
For cohort sakit_perut = Ya	4,602	1,804	11,743
For cohort sakit_perut = Tidak	,576	,488	,681
N of Valid Cases	263		

Pada output **pakai_sambal * sakit_perut Crosstabulation** pada bagian Expected Count. Jika expected count < 5. Maka lihat output **Chi-Square Tests** pada bagian Fisher's Exact Test, dan lihat pada kolom **Exact Sig. (2 – sided)** dan lihat lah hasilnya. Maka itulah **p-value nya**.

Sedangkan bila expected count > 5. Maka lihat output **Chi-Square Tests** pada bagian Pearson Chi-Square, dan lihat pada kolom **Asymp. Sig. (2 – sided)** dan lihat lah hasilnya.

Maka itulah p-value nya.

Pada kasus ini, expected count nya > 5 , maka p-value nya adalah **0,000**. Karena **p-value** (0,000) $< \alpha$ (0,05). Maka H_0 ditolak, jadi dapat disimpulkan bahwa variabel **pemakaian sambal** dependen terhadap variabel **sakit perut**.

Karena diperoleh kesimpulan bahwa **pemakaian sambal** berhubungan dengan **sakit perut**, maka dapat dihitung pula keeratan hubungan antara keduanya menggunakan Odds Ratio pada tabel **Risk Estimate**. Diperoleh nilai OR adalah **7,987** yang artinya seseorang yang memakan **memakai sambal** memiliki resiko untuk terserang **sakit perut** 7,987 x lebih besar daripada seseorang yang tidak **memakan memakai sambal**.

MODUL VII

UJI VALIDITAS DAN RELIABILITAS

Salah satu instrumen yang sering dipakai untuk pengambilan data primer dalam penelitian ilmiah adalah kuesioner, yang bertujuan untuk mengetahui pendapat seseorang mengenai suatu hal. Sebuah kuesioner dapat disusun dengan pertanyaan yang bersifat terbuka (contoh: Bagaimanakah pendapat anda mengenai kinerja Lembaga Survei Indonesia?) atau pertanyaan tertutup (dalam sebulan terakhir ini anda mengalami sakit sebanyak : 1. tidak pernah, 2. satu kali, 3. dua kali, 4. tiga kali, 5. lebih dari tiga kali).

Dalam penelitian **kualitatif** yang menggunakan instrumen kuesioner sebagai salah satu alat ukur, ada dua syarat penting yang harus dipenuhi yaitu keharusan sebuah kuesioner untuk Valid dan Reliabel. Suatu kuesioner dikatakan Valid jika pertanyaan pada suatu kuesioner mampu untuk mengungkapkan sesuatu yang akan diukur oleh kuesioner tersebut. Sedangkan suatu kuesioner dikatakan reliabel (handal) jika jawaban seseorang terhadap pertanyaan adalah konsisten atau stabil dari waktu ke waktu.

Pengukuran reliabilitas pada dasarnya dapat dilakukan dengan dua cara :

- Repeated Measure atau mengukur ulang. Di sini seseorang akan disodori pertanyaan yang sama pada waktu berbeda, dan kemudian dilihat apakah dia tetap konsisten dengan jawabannya atau tidak.
- One short atau sekali saja. Di sini pengukuran hanya sekali dan kemudian hasilnya dibandingkan dengan hasil pertanyaan lain.

Macam validitas umumnya digolongkan ke dalam tiga kategori besar, yaitu validitas isi (content validity), validitas berdasarkan kriteria (criterion-related validity) dan validitas konstruk. Pada penelitian ini akan dibahas hal menyangkut validitas untuk menguji apakah pertanyaan-pertanyaan itu telah mengukur aspek yang sama. Untuk itu dipergunakanlah validitas konstruk. Uji validitas dilakukan dengan mengukur korelasi antara variabel/item dengan skor total variabel. Cara mengukur validitas konstruk yaitu dengan mencari korelasi antara masing- masing pertanyaan dengan skor total menggunakan rumus teknik korelasi product moment, sebagai berikut:

dimana r : koefisien korelasi product moment

X : skor tiap pertanyaan/ item

Y : skor total

N : jumlah responden

Setelah semua korelasi untuk setiap pertanyaan dengan skor total diperoleh, nilai-nilai tersebut dibandingkan dengan nilai kritik. Selanjutnya, jika nilai koefisien korelasi product moment dari suatu pertanyaan tersebut berada diatas nilai tabel kritik, maka pertanyaan tersebut signifikan.

Contoh :

Untuk mengetahui bagaimana perilaku konsumen terhadap roti produksi DM, perusahaan mengukur dengan pendapat konsumen terhadap kualitas roti (faktor_1) dan harga roti (faktor_2). Untuk itu perusahaan menyusun kuesioner yang dibagikan kepada konsumen yang berisi :

Faktor 1 (kualitas roti) terdiri dari 10 butir pertanyaan misalnya :

Butir 1 : Apakah Anda setuju dengan rasa roti DM yang enak dan lembut?

Faktor 2 (harga roti) terdiri dari 7 butir pertanyaan.

Semua jawaban berupa pilihan dengan skala likert lima skala :

5 = sangat setuju 4 = setuju 3 = ragu-ragu 2 = tidak setuju 1 = sangat tidak setuju Namun sebelum diedar secara resmi kuesioner tersebut harus diuji validitas dan reliabilitasnya dengan menyebarkan kepada 30 responden. Berikut hasilnya :

Data_faktor_1

Nama variabel	Tipe	Label	keterangan
Nama	String	-	Nama responden
Butir1	Numerik	-	Jawaban responden
Butir2	Numerik	-	Jawaban responden
-----	-----	-----	-----
Butir10	Numerik	-	Jawaban responden

Data_faktor_2

Nama variabel	Tipe	Label	keterangan
Nama	String	-	Nama responden
Butir11	Numerik	-	Jawaban responden
Butir12	Numerik	-	Jawaban responden
-----	-----	-----	-----
Butir17	Numerik	-	Jawaban responden

dalam tampilan variabel view sebagai berikut :

	Name	Type	Width	Decimals	Label	Values	Missing
1	nama	String	8	0		None	None
2	butir1	Numeric	8	0		None	None
3	butir2	Numeric	8	0		None	None
4	butir3	Numeric	8	0		None	None
5	butir4	Numeric	8	0		None	None
6	butir5	Numeric	8	0		None	None
7	butir6	Numeric	8	0		None	None
8	butir7	Numeric	8	0		None	None
9	butir8	Numeric	8	0		None	None
10	butir9	Numeric	8	0		None	None
11	butir10	Numeric	8	0		None	None
12	butir11	Numeric	8	0		None	None
13	butir12	Numeric	8	0		None	None
14	butir13	Numeric	8	0		None	None
15	butir14	Numeric	8	0		None	None
16	butir15	Numeric	8	0		None	None
17	butir16	Numeric	8	0		None	None
18	butir17	Numeric	8	0		None	None

Isian Data :

Nama	Butir 1	Butir 2	Butir 3	Butir 4	Butir 5	Butir 6	Butir 7	Butir 8
ali	3	4	3	3	2	4	4	3
amir	4	3	3	3	2	3	3	3
amin	2	3	3	4	2	3	4	4
anna asoi	3	5	3	3	1	3	3	2
badu	4	3	4	5	2	4	3	4
budi bari	2	4	3	4	2	2	4	3
basi	4	4	4	5	2	4	4	3
cecep	2	5	3	4	1	5	5	5
cicil cuko	4	5	3	3	1	1	5	5
dede dani	4	3	4	5	4	1	4	4
didik euis	3	4	4	3	2	2	2	3
emil	3	2	4	2	5	2	1	2
farah	2	3	4	1	3	4	4	3
faisal	3	5	3	3	2	3	3	4
guna	2	4	3	5	3	2	2	5
gandi	4	5	5	3	1	4	4	3
heru	3	5	5	4	2	3	3	5
husain	2	4	4	4	2	4	4	4
imron	3	3	3	4	1	2	4	5
jujuk kiki	4	4	4	3	2	4	3	5

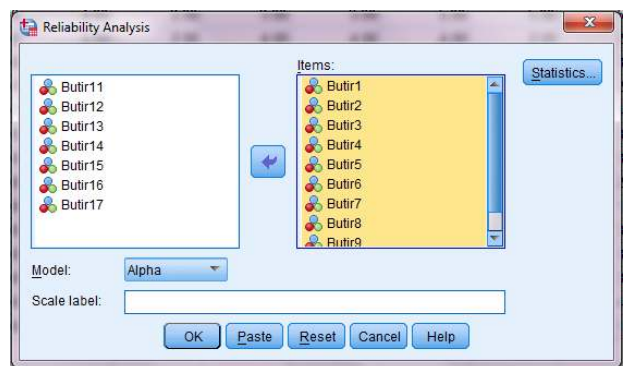
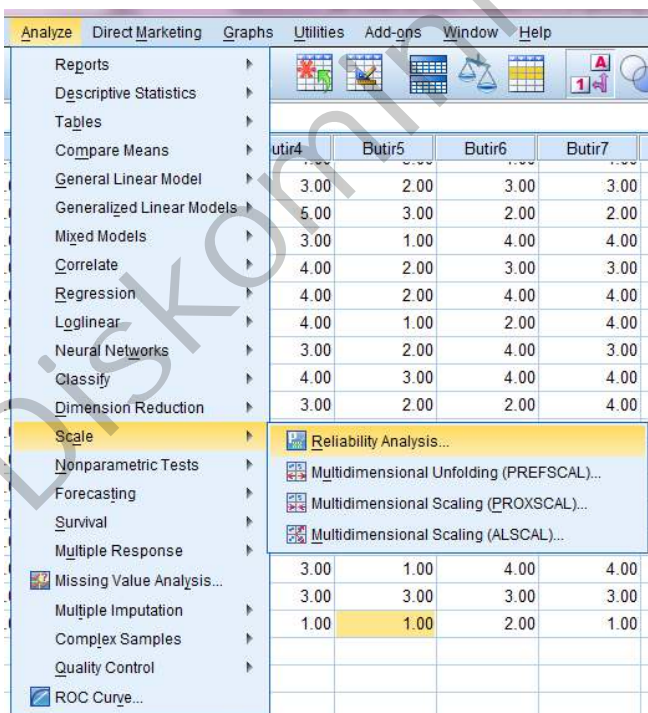
Butir9	Butir10	Butir11	Butir12	Butir13	Butir14	Butir15	Butir16	Butir17
4	3	3	4	3	3	2	4	4
3	3	4	3	3	3	2	3	3
2	2	2	3	3	4	2	3	4
2	2	3	5	3	3	1	3	3
4	4	4	3	4	5	2	4	3
3	3	2	4	3	4	2	2	4
5	2	4	4	4	5	2	4	4
2	2	4	5	3	3	1	1	5
5	2	4	3	4	5	4	1	4
4	4	3	4	4	3	2	2	2
2	2	3	2	4	2	5	2	1
3	3	2	3	4	1	3	4	4
2	2	3	5	3	3	2	3	3
3	3	2	4	3	5	3	2	2
4	4	4	5	5	3	1	4	4
5	5	3	5	5	4	2	3	3
2	2	2	4	4	4	2	4	4
3	3	3	3	3	4	1	2	4
2	2	4	4	4	3	2	4	3
4	4	2	3	3	4	3	4	4
2	3	3	4	5	3	2	2	4
3	3	4	3	3	2	2	2	2

sehingga diperoleh tampilan pada variabel view sebagai berikut :

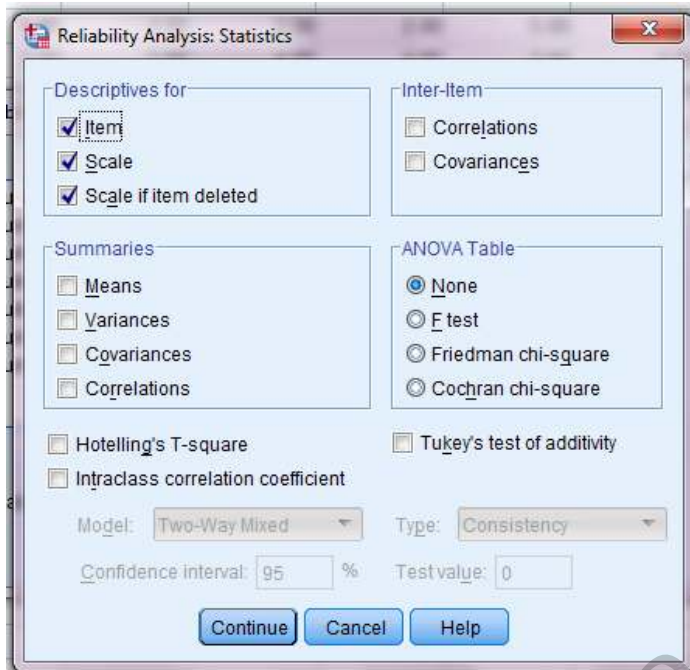
	nama	butir1	butir2	butir3	butir4	butir5	butir6	butir7	butir8	butir9
1	ali	3	4	3	3	2	4	4	3	4
2	amir	4	3	3	3	2	3	3	3	3
3	amin	2	3	3	4	2	3	4	4	2
4	anna	3	5	3	3	1	3	3	2	2
5	asoi	4	3	4	5	2	4	3	4	4
6	badu	2	4	3	4	2	2	4	3	3
7	budi	4	4	4	5	2	4	4	3	5
8	bari	2	5	3	4	1	5	5	5	4
9	basi	4	5	3	3	1	1	5	5	2
10	cecep	4	3	4	5	4	1	4	4	5
11	cicil	3	4	4	3	2	2	2	3	4
12	cuko	3	2	4	2	5	2	1	2	2
13	dede	2	3	4	1	3	4	4	3	3
14	dani	3	5	3	3	2	3	3	4	2
15	didik	2	4	3	5	3	2	2	5	3
16	euis	4	5	5	3	1	4	4	3	4
17	emil	3	5	5	4	2	3	3	5	5
18	farah	2	4	4	4	2	4	4	4	2
19	faisal	3	3	3	4	1	2	4	5	3
20	guna	4	4	4	3	2	4	3	5	2
21	gandi	2	3	3	4	3	4	4	4	4
22	hari	3	4	5	3	2	2	4	4	4

Langkah Uji Validitas :

- masukkan data faktor 1 pada SPSS seperti pada tampilan variabel view diatas.
- klik menu Analyze >> scale >> reliability analysis, seperti pada tampilan berikut :
- masukkan semua variable butir faktor 1 ke kotak item yang ada di sebelah kanan.
- pada bagian Model, biarkan pilihan pada alpha



- klik menu statistics, pada bagian descriptive for check list ketiganya (item, scale, scale if item deleted).



- abaikan bagian lain dan klik tombol continue untuk kembali ke dialog utama.
- Klik ok.
- Sehingga diperoleh output sebagai berikut :

Reliability Statistics

Cronbach's Alpha	N of Items
.724	10

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Butir1	27.7000	30.631	-.161	.766
Butir2	27.2667	24.478	.462	.690
Butir3	27.4667	24.671	.523	.684
Butir4	27.6667	21.885	.612	.659
Butir5	28.6667	31.885	-.270	.793
Butir6	28.0000	25.586	.301	.715

Modul Analisis Data Statistik

Butir7	27.6667	23.126	.524	.677
Butir8	27.4000	22.041	.610	.660
Butir9	27.9000	21.128	.717	.639
Butir10	28.0667	23.720	.555	.675

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
30.8667	29.844	5.46294	10

Interpretasi Output

Bagian pertama berisi data singkat dari ke sepuluh butir, yang mencakup mean dan standar deviasi untuk kesepuluh butir tersebut, yaitu 30.8667 dan 5.4629 yang sebenarnya tidak berarti karena data tersebut menggunakan skala ordinal.

Bagian kedua adalah hasil dari proses validitas. Langkah **uji validitas** adalah sebagai berikut:

Uji Validitas :

1. Hipotesis

H_0 : Skor butir berkorelasi positif dengan skor faktor

H_1 : Skor butir tidak berkorelasi positif dengan skor faktor.

2. Tingkat Signifikansi

Dari tabel r, untuk $df = \text{jumlah kasus} - 2$, atau dalam kasus ini $df = 30 - 2 = 28$, dengan nilai $\alpha = 5\%$ terdapat angka 0,239.

Disini uji dilakukan satu arah karena hipotesis menunjukkan arah tertentu, yaitu positif. Untuk df yang tidak ada pada tabel, seperti angka 31 dan 44 dan sebagainya bisa dilakukan interpolasi (misal 44 bisa interpolasi antara 40 dan 60).

3. Statistik Uji

Nilai r untuk tiap item dapat dilihat pada kolom **CORRECTED ITEM – TOTAL CORRELATION**.

4. Daerah Kritik

H_0 tidak ditolak jika :

r hasil positif, serta r hasil $>$ r tabel, sehingga butir tersebut dikatakan Valid.

5. Kesimpulan

Dalam kasus faktor 1 tersebut dapat dituliskan uji validitasnya sebagai berikut :

Uji Validitas butir 1

1. Hipotesis

H_0 : Skor butir 1 berkorelasi positif dengan skor factor

H_1 : Skor butir 1 tidak berkorelasi positif dengan skor faktor.

2. Tingkat Signifikansi

$\alpha = 0,05$, $df = 28$, r tabel = 0,239

3. Statistik Uji

r hasil = - 0,161

4. Daerah Kritik

H_0 tidak ditolak jika :

r hasil positif, serta r hasil > r tabel, sehingga butir tersebut dikatakan Valid.

5. Kesimpulan

Karena r hasil (-0,161) tidak positif maka H_0 ditolak yang berarti bahwa skor butir 1 tidak berkorelasi positif dengan skor faktor sehingga butir 1 dikatakan tidak Valid.

Uji Validitas butir 2

1. Hipotesis

H_0 : Skor butir 2 berkorelasi positif dengan skor faktor

H_1 : Skor butir 2 tidak berkorelasi positif dengan skor faktor.

2. Tingkat Signifikansi

$\alpha = 5\%$, $df = 28$, r tabel = 0,239

3. Statistik Uji

r hasil = 0,462

4. Daerah Kritik

H_0 tidak ditolak jika : r hasil positif, serta r hasil > r tabel, sehingga butir tersebut dikatakan Valid.

5. Kesimpulan

Karena r hasil (0,462) > r tabel (0,239) maka H_0 tidak ditolak yang berarti bahwa skor butir 2 berkorelasi positif dengan skor faktor sehingga butir 2 dikatakan Valid.

dan seterusnya hingga Uji Validitas butir 10.

Kemudian, setelah dilakukan uji validitas diperoleh tabel sebagai berikut :

No	Variabel	r hasil	r tabel	keterangan
1	Butir 1	-0,161	0,239	Tidak valid
2	Butir 2	0,462	0,239	Valid
3	Butir 3	0,523	0,239	Valid
4	Butir 4	0,612	0,239	Valid
5	Butir 5	-0,270	0,239	Tidak valid
6	Butir 6	0,301	0,239	Valid
7	Butir 7	0,524	0,239	Valid
8	Butir 8	0,610	0,239	Valid
9	Butir 9	0,717	0,239	Valid
10	Butir 10	0,555	0,239	Valid

Dari tabel tersebut terlihat bahwa terdapat 2 variabel yang tidak valid yaitu variabel 1 dan 5, maka butir tersebut dikeluarkan, dan proses analisis seperti diatas diulang untuk butir yang valid saja. Setelah dilakukan pengeluaran variabel butir 1 dan variabel butir 5 akan diperoleh output sebagai berikut :

Reliability Statistics

Cronbach's Alpha	N of Items
.836	8

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Butir2	21.9000	26.300	.554	.819
Butir3	22.1000	27.955	.449	.831
Butir4	22.3000	24.355	.617	.810
Butir6	22.6333	27.551	.375	.841
Butir7	22.3000	24.700	.628	.809
Butir8	22.0333	24.033	.665	.803
Butir9	22.5333	24.189	.655	.805
Butir10	22.7000	26.148	.577	.816

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
25.5000	32.741	5.72201	8

Lalu dilakukan uji hipotesis validitas seperti langkah diatas, sehingga dapat dibuat tabel sebagai berikut :

No	Variabel	r hasil	r tabel	keterangan
1	Butir 2	0,554	0,239	Valid
2	Butir 3	0,449	0,239	Valid
3	Butir 4	0,617	0,239	Valid
4	Butir 6	0,375	0,239	Valid
5	Butir 7	0,628	0,239	Valid
6	Butir 8	0,665	0,239	Valid
7	Butir 9	0,655	0,239	Valid
8	Butir 10	0,577	0,239	Valid

Maka telah diperoleh semua variabel pada faktor 1 valid.

Jika semua butir telah valid, kemudian dilakukan **uji reliabilitas**.

Langkah Uji Reliabilitas :

1. Hipotesis

H_0 : Skor butir berkorelasi positif dengan komposit faktornya

H_1 : Skor butir tidak berkorelasi positif dengan komposit faktornya

2. Statistik

Uji

Cronbach'

s alpha

3. Daerah Kritik

H_0 tidak ditolak jika :

Nilai dari Cronbach's alpha > 0.60 , sehingga butir tersebut dikatakan Reliabel.

4. Kesimpulan

Dalam kasus tersebut diperoleh sebagai berikut :

Uji Reliabilitas faktor 1 :

1. Hipotesis

H_0 : Faktor 1 reliabel

H_1 : Faktor 1 tidak reliabel

2. Statistik Uji

Cronbach's alpha = 0,836

3. Daerah Kritik

H_0 tidak ditolak jika :

Nilai Cronbach's alpha > 0.60 , sehingga butir tersebut dikatakan Reliabel.

4. Kesimpulan

Karena Cronbach's Alpha (0.836) positif dan > 0.60 maka H_0 tidak ditolak yang berarti bahwa faktor 1 reliabel.

Keterangan :

Dengan demikian, setelah melewati 1 putaran, kedelapan butir pada data faktor 1 adalah valid dan reliabel.

DAFTAR PUSTAKA

Junaidi, Junaidi. (2015). *Statistik Deskriptif dengan Microsoft Office Excel*. 10.13140/RG.2.1.1297.4248.

Mustari, Andhie S dan Daud Eliezar. (2019). *Modul Analyse / Analisis*. Jakarta: Pusat Pendidikan dan Pelatihan Badan Pusat Statistik.

Diskominfo Kabupaten Blitar. (2017). *Modul Statistik Dasar dan Metode Perhitungan Statistik Sektoral*. Pemerintah Kabupaten Blitar.



PEMERINTAH KOTA SEMARANG