Nama Mhs : .....
NIM : .....
Hari/Jam : .....
No. Komp : .....

Penuntun/Modul Praktikum

STATISTIKA
(PSA102)

Materi:

SPSS

(Statistical Product & Service Solution)

TIM PENYUSUN

Oleh:

TIM Pengasuh Mata Kuliah Statistika (PSA102)
Program Studi Agribisnis, Jurusan Sosial Ekonomi Pertanian,
Fakultas Pertanian, Universitas Syiah Kuala
&
Laboratorium Statistik dan Sosial, Fakultas Pertanian, Universitas Syiah Kuala

Fakultas Pertanian
Universitas Syiah Kuala
2021

KATA PENGANTAR

Panduan praktikum ini ditujukan sebagai pelengkap kuliah Statistika pada Fakultas Pertanian Universitas Syiah Kuala dan untuk menambah pemahaman mahasiswa terhadap materi perkuliahan. Panduan praktikum ini menggunakan program statistik SPSS for Windows version 20.

Buku panduan ini terdiri dari tujuh bab yang menjelaskan pengenalan software SPSS for Windows sebagai sarana untuk mempermudah perhitungan dan materi-materi kuliah perkuliahan Statistika yang telah disusun berdasarkan keeratan hubungan antar materinya.

Pemilihan software SPSS for Windows didasarkan pada kemudahan dalam pengoperasiannya, tampilan grafis yang menarik, kemudahan dalam memahami output hasil perhitungan dan membuat kesimpulan serta tidak memerlukan spesifikasi komputer yang tinggi.

Tim penyusun mengucapkan terima kasih kepada para dosen yang telah memberikan dukungan untuk menyelesaikan buku panduan ini. Tim juga dengan senang hati menerima segala bentuk saran dan kritik demi perbaikan panduan praktikum ini di masa yang akan datang.

Semoga bermanfaat.

Tim Penyusun

DAFTAR ISI

KATA PENGANTAR	.....	i
DAFTAR ISI	.....	ii
BAB I PENDAHULUAN DAN PENGENALAN SPSS	.....	1
BAB II DISTRIBUSI FREKUENSI DAN PENGELOMPOKAN DATA	.....	14
BAB III STATISTIKA DESKRIPTIF	.....	20
BAB IV TRANSFORMASI VARIABEL	.....	27
BAB V TABEL	.....	31
BAB VI GRAFIK	.....	36
BAB VII PENGUJIAN HIPOTESIS	.....	45
DAFTAR PUSTAKA	.....	55

BAB I

Pendahuluan dan Pengenalan SPSS

A. Pendahuluan

Aplikasi ilmu statistika dapat dibagi dalam dua bagian:

1. Statistik Deskriptif

   Statistik Deskriptif berusaha menjelaskan atau menggambarkan berbagai karakteristik data seperti berapa rata-ratanya, seberapa jauh data-data bervariasi dari rata-ratanya, berapa median data, dan sebagainya.

2. Statistik Induktif (Inferensi)

   Statistik Induktif membuat berbagai inferensi terhadap sekumpulan data yang berasal dari suatu sampel. Tindakan inferensi tersebut seperti melakukan perkiraan besaran populasi, uji hipotesis, peramalan, dan sebagainya.

Skala Pengukuran

Berdasarkan tingkat pengukurannya (level of measurement), data statistik dapat dibedakan dalam empat jenis:

A). Data Kualitatif, disebut juga data yang buka berupa angka. Pada data ini tidak bisa dilakukan operasi matematika. Data kualitatif dapat dibagi dua:

1. **Nominal.** Skala nominal hanya bisa membedakan benda atau peristiwa yang satu dengan yang lainnya berdasarkan nama (predikat). Skala pengukuran nominal digunakan untuk mengklasifikasi obyek, individual atau kelompok dalam bentuk kategori dan memberikan angka pada tiap-tiap kategori.

   Pemberian angka atau simbol pada skala nomial tidak memiliki maksud kuantitatif hanya menunjukkan ada atau tidak adanya atribut atau karakteristik pada objek yang diukur. Misalnya, jenis kelamin diberi kode 1 untuk laki-laki dan kode 2 untuk perempuan. Angka ini hanya berfungsi sebagai label kategori, tanpa memiliki nilai instrinsik dan tidak memiliki arti apa pun. Kita tidak bisa mengatakan perempuan dua kali dari laki-laki. Kita bisa saja mengkode laki-laki menjadi 2 dan perempuan dengan kode 1, atau bilangan apapun asal kodenya berbeda antara laki-laki dan perempuan. Misalnya lagi untuk agama, kita bisa mengkode 1 = Islam, 2 = Kristen, 3 = Hindu, 4 = Budha dan seterusnya.

Kita bisa menukar angka-angka tersebut, selama suatu karakteristik memiliki angka yang berbeda dengan karakteristik lainnya. Karena tidak memiliki nilai instrinsik, maka angka-angka (kode-kode) yang kita berikan tersebut tidak memiliki sifat sebagaimana bilangan pada umumnya. Oleh karenanya, pada variabel dengan skala nominal tidak dapat diterapkan operasi matematika standar (aritmatik) seperti pengurangan, penjumlahan, perkalian, dan lainnya. Peralatan statistik yang sesuai dengan skala nominal adalah peralatan statistik yang berbasiskan (berdasarkan) jumlah dan proporsi seperti modus, distribusi frekuensi, Chi Square dan beberapa peralatan statistik non-parametrik lainnya.

2. **Ordinal.** Skala Ordinal sering disebut dengan skala peringkat. Hal ini karena dalam skala ordinal, lambang-lambang bilangan hasil pengukuran selain menunjukkan pembedaan juga menunjukkan urutan atau tingkatan obyek yang diukur menurut karakteristik tertentu. Misalnya tingkat kepuasan seseorang terhadap produk. Bisa kita beri angka dengan 5 = sangat puas, 4 = puas, 3 = kurang puas, 2 = tidak puas dan 1= sangat tidak puas. Atau misalnya dalam suatu lomba, pemenangnya diberi peringkat 1,2,3 dstnya.

   Dalam skala ordinal, tidak seperti skala nominal, ketika kita ingin mengganti angka-angkanya, harus dilakukan secara berurut dari besar ke kecil atau dari kecil ke besar. Jadi, tidak boleh dibuat 1 = sangat puas, 2 = tidak puas, 3 = puas dan seterusnya.

   Selain itu, yang perlu diperhatikan dari karakteristik skala ordinal adalah meskipun nilainya sudah memiliki batas yang jelas tetapi belum memiliki jarak (selisih). Kita tidak tahu berapa jarak kepuasan dari tidak puas ke kurang puas. Dengan kata lain, walaupun sangat puas kita beri angka 5 dan sangat tidak puas kita beri angka 1, kita tidak bisa mengatakan bahwa kepuasan yang sangat puas lima kali lebih tinggi dibandingkan yang sangat tidak puas.

   Sebagaimana halnya pada skala nominal, pada skala ordinal kita juga tidak dapat menerapkan operasi matematika standar (aritmatik) seperti pengurangan, penjumlahan, perkalian, dan lainnya. Peralatan statistik yang sesuai dengan skala ordinal juga adalah peralatan statistik yang berbasiskan (berdasarkan) jumlah

dan proporsi seperti modus, distribusi frekuensi, Chi Square dan beberapa peralatan statistik non-parametrik lainnya.

B). Data Kuantitatif, disebut juga data yang berupa angka dalam arti sebenarnya. Sehingga bisa dilakukan operasi matematika. Terdiri dari dua jenis data:

3. **Interval.** Skala interval mempunyai karakteristik seperti yang dimiliki oleh skala nominal dan ordinal dengan ditambah karakteristik lain, yaitu berupa adanya interval yang tetap. Dengan demikian, skala interval sudah memiliki nilai intrinsik, sudah memiliki jarak, tetapi jarak tersebut belum merupakan kelipatan. Pengertian "jarak belum merupakan kelipatan" ini kadang-kadang diartikan bahwa skala interval tidak memiliki nilai nol mutlak. Misalnya pada pengukuran suhu. Kalau ada tiga daerah dengan suhu daerah A = 10^oC, daerah B = 15^oC dan daerah C = 20^oC. Kita bisa mengatakan bahwa selisih suhu daerah B, 5^oC lebih panas dibandingkan daerah A, dan selisih suhu daerah C dengan daerah B adalah 5^oC. (Ini menunjukkan pengukuran interval sudah memiliki jarak yang tetap). Tetapi, kita tidak bisa mengatakan bahwa suhu daerah C dua kali lebih panas dibandingkan daerah A (artinya tidak bisa jadi kelipatan). Karena dengan pengukuran yang lain, misalnya dengan Fahrenheit, di daerah A suhunya adalah 50^oF, di daerah B = 59^oF dan daerah C = 68^oF. Artinya, dengan pengukuran Fahrenheit, daerah C tidak dua kali lebih panas dibandingkan daerah A, dan ini terjadi karena dalam derajat Fahrenheit titik nolnya pada 32, sedangkan dalam derajat Celcius titik nolnya pada 0.

   Skala interval ini sudah benar-benar angka dan, kita sudah dapat menerapkan semua operasi matematika serta peralatan statistik kecuali yang berdasarkan pada rasio seperti koefisien variasi.

4. **Skala rasio.** Skala rasio adalah skala data dengan kualitas paling tinggi. Pada skala rasio, terdapat semua karakteristik skala nominal,ordinal dan skala interval ditambah dengan sifat adanya nilai nol yang bersifat mutlak. Nilai nol mutlak ini artinya adalah nilai dasar yang tidak bisa diubah meskipun menggunakan skala yang lain. Oleh karenanya, pada skala ratio, pengukuran sudah mempunyai nilai perbandingan/rasio. Pengukuran-pengukuran dalam skala rasio yang sering digunakan adalah pengukuran tinggi dan berat. Misalnya berat benda A adalah

30 kg, sedangkan benda B adalah 60 kg. Maka dapat dikatakan bahwa benda B dua kali lebih berat dibandingkan benda A.

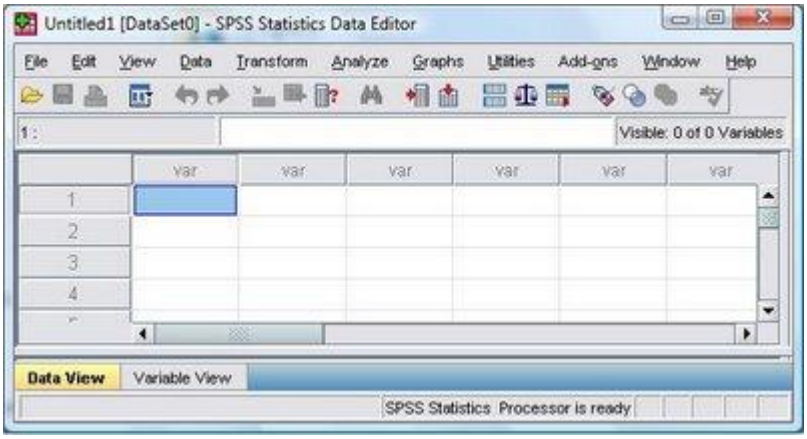
Jenis-jenis data yang telah dijelaskan di atas harus dipahami dengan baik karena penerapan dalam statistik akan berbeda untuk jenis data yang berbeda.

B. Pengenalan SPSS

SPSS (Statistical Package for the Social Science) merupakan software statistik yang pada awalnya digunakan untuk riset dibidang sosial dan melayani berbagai jenis user. SPSS merupakan paket program statistik yang paling populer dan paling banyak digunakan di seluruh dunia. Hal inilah yang yang membuat kepanjangan SPSS saat ini adalah Statistical Product and Service Solution. Dengan SPSS semua kebutuhan pengolahan data dapat diselesaikan dengan mudah dan cepat. Kemampuan yang dapat diperoleh dari SPSS meliputi pemrosesan segala bentuk file data, modifikasi data, membuat tabulasi berbentuk distribusi frekuensi, analisis statistik deskriptif, analisis statistik lanjut yang sederhana maupun kompleks, pembuatan grafik dan sebagainya.

Memasukkan Data pada SPSS

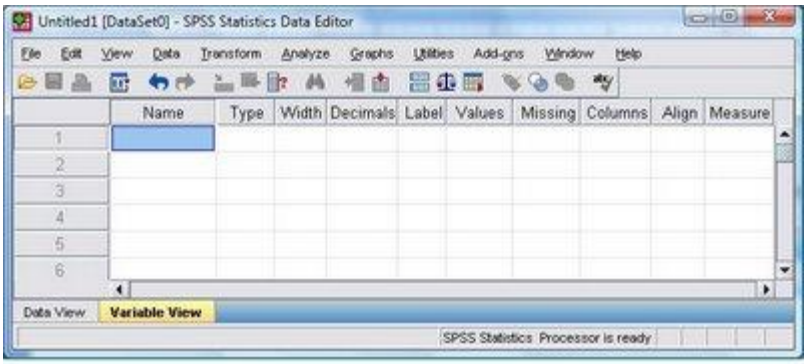
Untuk menginput data, buka Program SPSS melalui **Start → Programs → IBM SPSS Statistics → IBM SPSS Statiscs**. Pertama kali akan muncul tampilan sebagai berikut:



Gambar 1.1 Tampilan layar Data View

Tampilan tersebut adalah tampilan Data Editor dalam SPSS yang mempunyai fungsi utama untuk mendefinisikan, menginput, mengedit dan menampilkan data. Sebelum menginput data, definisikan terlebih dahulu data (variabel) yang akan diinput. Perhatikan di sudut kiri bawah dari tampilan data editor di atas, terdapat

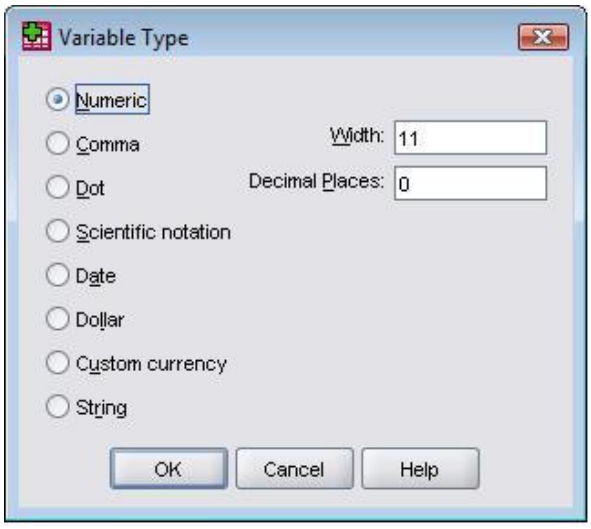
menu Data View dan menu Variable View. Untuk mendefinisikan data (variabel), klik Variable View, maka akan muncul tampilan berikut:



Gambar 1.2 Tampilan layar Variable View

Setiap baris dalam tampilan diatas digunakan untuk mendefinisikan satu variabel (data) yang akan diinput. Ada beberapa informasi yang perlu dimasukkan, yaitu:

1. **Name.** Isikan nama variabel. Persyaratan dalam pemberian nama variabel adalah:
   a. Nama variabel tidak boleh duplikasi dengan nama variabel lainnya.
   b. Nama variabel paling panjang hanya 64 karakter dan harus diawali oleh huruf atau @, #, $. Karakter berikutnya boleh kombinasi huruf, @, #, $ atau angka. Nama variabel yang diawali dengan tanda $ menunjukkan bahwa variabel tersebut adalah suatu variabel sistem.
   c. Variabel tidak boleh mengandung spasi dan kata-kata kunci perintah SPSS yaitu ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, dan WITH.
2. **Type.** Definisikan tipe variabel. Ketika anda mengklik sel di bawah type, akan muncul titik tiga (…). Klik titik tiga tersebut, maka akan muncul tampilan berikut:



Gambar 1.3 Tampilan kotak dialog Variable Type

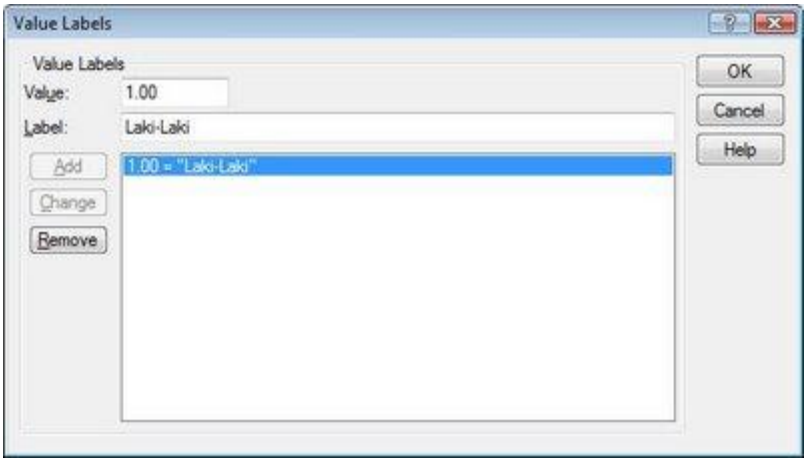
Dari tampilan diatas, terdapat beberapa pilihan tipe variabel sebagai berikut:

a. Numeric. Variabel yang berbentuk angka
b. Comma. Variabel numerik dengan tampilan koma untuk setiap 3 angka (memisahkan ribuan), dan titik untuk memisahkan desimal. Misalnya, jika data yang diinput adalah 2567932, maka akan ditampilkan dalam SPSS sebagai 2,567,932.00
c. Dot. Variabel numerik dengan tampilan titik untuk setiap 3 angka (memisahkan ribuan), dan koma untuk memisahkan desimal. Misalnya, jika data yang diinput adalah 2567932, maka akan ditampilkan dalam SPSS sebagai 2.567.932,00
d. Scientific notation. Variabel numerik dengan tampilan scientific. Misalnya, jika data yang diinput adalah 2567932, maka akan ditampilkan dalam SPSS sebagai 2.57E+006
e. Date. Variabel numerik dengan nilai yang ditampilkan dalam format tanggal atau waktu. Jika anda pilih tipe data ini, akan muncul tampilan pilihan format tanggal atau waktu.
f. Dollar. Variabel numerik dengan tampilan tanda $.
g. Custom currency. Variabel numerik yang ditampilkan dalam format uang yang anda inginkan (custom currency) misalnya dalam bentuk Rp. Penggunaan pilihan format ini harus didefinisikan terlebih dahulu dalam menu Options pada Currency tab.
h. String. Variabel yang tidak berbentuk numerik (angka) dan karenanya tidak digunakan dalam perhitungan. Jenis ini juga dikenal sebagai variabel alphanumeric

Setelah memilih jenis variabel, lanjutkan dengan mengisi Width, yaitu jumlah karakter (angka/huruf) maksimum dari data yang akan diinput. Setelah itu, tentukan jumlah decimal yang ingin ditampilkan. Selanjutnya klik OK.

3. **Label.** Label adalah keterangan mengenai variable dan dapat dibuat sampai 256 karakter. Selain itu, label dapat menggunakan spasi maupun karakter-karakter yang tadinya tidak dapat digunakan pada nama variabel.
4. **Values.** Values ini secara khusus berguna jika data yang kita gunakan merupakan kode numerik (dalam bentuk angka) yang mewakili kategori non-numerik,

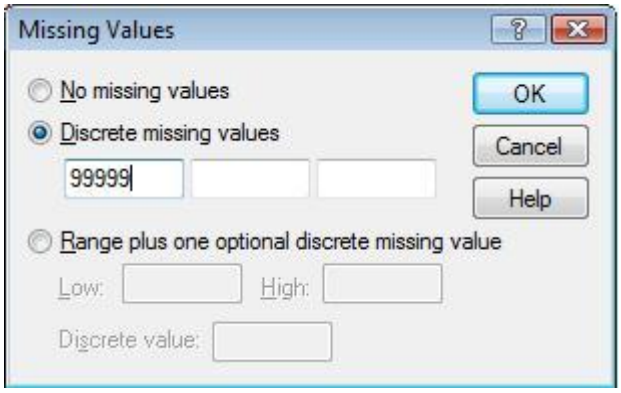
misalnya kode 1 untuk laki-laki dan kode 2 untuk perempuan. Untuk menginput values dari masing-masing variabel, klik sel di bawah Values, akan muncul titik tiga (…). Klik titik tiga tersebut, maka akan muncul tampilan berikut:



Gambar 1.4 Tampilan kotak dialog Value Labels

Isikan kode pada kotak Values dan labelnya pada kotak Label. Misalnya, kode 1 untuk Laki-laki. Setelah itu, klik Add. Lanjutkan untuk kode-kode berikutnya, setelah itu klik OK.

5. **Missing.** Menentukan nilai "missing" ini berguna jika dalam pertanyaan survai, ada responden yang tidak memberikan/menolak memberikan jawaban, sehingga tidak tersedia data untuk diinput. Misalnya, jika ada responden yang menolak memberikan jawaban mengenai pendapatannya. Jika kita menginput jawaban responden tersebut dengan angka 0, maka dalam pengolahannya, SPSS akan memasukkan dalam perhitungan (sehingga akan berpengaruh terhadap rata-rata keseluruhan maupun terhadap distribusi frekuensi). Tetapi jika kita mendefinisikan suatu angka untuk menyatakan nilai missing tersebut, maka SPSS akan mengeluarkan dari perhitungan. Untuk mendefinisikan nilai "missing" dari masing-masing variabel, klik sel di bawah "Missing", akan muncul titik tiga (…). Klik titik tiga tersebut, maka akan muncul tampilan berikut:



Gambar 1.5 Tampilan kotak dialog Missing Value

Kita bisa mendefinisikan tiga deretan angka yang berbeda untuk nilai missing masing-masing variabel. Kita juga memberikan range nilai untuk mendefinisikan nilai missing tersebut. Dalam contoh tampilan diatas, misalnya didefinisikan hanya satu deretan angka untuk nilai missing yaitu 99999. Dengan demikian, jika terdapat data yang kosong (atau tidak terisi) dari variabel, maka inputkan angka 99999. Setelah mendefinisikan nilai missing, klik OK.

6. **Coloumn.** Menentukan lebar kolom dari worksheet SPSS untuk input data. Lebar kolom ini ditentukan minimal sama dengan "Width" variabel yang telah ditentukan sebelumnya.
7. **Align.** Perataan dari tampilan input data. Jika diklik sel dibawah align, akan muncul tiga pilihan yaitu left (rata kiri), right (rata kanan) dan center (rata tengah).
8. **Measure.** Skala pengukuran dari masing-masing variabel. Ada tiga pilihan yaitu Scale, Ordinal dan Nominal. Scale kita pilih jika skala pengukuran kita adalah skala interval atau ratio.

Setelah selesai dengan tahap terakhir pendefinisian variabel ini, klik kembali menu Data View (yang ada disudut kiri bawah). Dengan cara ini, kita akan masuk ke worksheet SPSS dan siap untuk menginput data.

Sebagai latihan, input data hasil penelitian terdapat 18 responden penelitian sebagai berikut:

No	Responden	Jenis Kelamin	Umur	Pendidikan	Pendapatan	Konsumsi
1	Susan	Perempuan	25	SLTP	5000	3750
2	Susanto	Laki-laki	20	SLTP	2500	1875
3	Marni	Perempuan	23	SLTA	-	-
4	Semar	Laki-laki	30	S1	2000	1500
5	Tini	Perempuan	35	D3	1500	1400
6	Tono	Laki-laki	40	SLTA	1300	1200
7	Fadli	Laki-laki	43	SLTP	2000	1500
8	Fadliani	Perempuan	42	S1	2200	1650
9	Purwanti	Perempuan	30	S1	1700	1275
10	Purwana	Laki-laki	25	D3	1100	1000
11	Purwani	Perempuan	40	SLTA	3000	2250
12	Purwano	Laki-laki	25	D3	3200	2400
13	Mutia	Perempuan	38	SLTA	1800	1350
14	Tiara	Perempuan	45	SLTA	-	-
15	Joni	Laki-laki	55	SLTP	1400	1300
16	Jono	Laki-laki	50	SLTA	1600	1400
17	Kris Dayanti	Perempuan	35	D3	1900	1300
18	Kris Jhon	Laki-laki	57	D3	1700	1300

Pada contoh data latihan di atas, kita punya enam variabel (data) yang akan diinput yaitu nama responden, jenis kelamin, umur, pendidikan, penghasilan dan konsumsi. Mari kita definisikan masing-masing variabel sebagai berikut:

Variabel pertama:

Nama Variabel : Responden

Type : String (karena variabel ini tidak berbentuk numerik)

Width : 18 (untuk data kita ini, jumlah karakter terbanyak 18)

Decimal : 0 (untuk data type string, desimal akan otomatis 0)

Label : Nama Responden

Values : None (untuk data type string, values akan otomatis none)

Missing : None (untuk data type string, missing akan otomatis none)

Column : 18 (ukuran kolom ini sesuaikan dengan jumlah karakter dari nama variabel atau maksimum karakter dari data pada variabel tersebut, mana yang paling banyak)

Align : Left (untuk data string sebaiknya dibuat rata kiri)

Measure : Nominal (untuk data string, pilih saja measure nominal)

Variabel kedua:

Nama Variabel : Sex

Type : Numeric

Width : 2 (sebenarnya input data yang akan kita masukkan nanti hanya berupa kode 1 dan 2, atau hanya terdiri dari 1 karakter, tetapi width nya sebaiknya kita lebihkan 1 karakter)

Decimal : 0 (karena tidak memerlukan angka dibelakang koma)

Label : Jenis Kelamin Responden

Values : 1 = laki-laki, 2 = perempuan

Missing : None (karena informasi mengenai jenis kelamin tersedia pada semua responden)

Column : 4

Align : Rigth (untuk data numerik sebaiknya dibuat rata kanan)

Measure : Nominal (angka untuk pengkodean jenis kelamin ini, adalah termasuk data skala nominal)

Variabel ketiga:

Nama Variabel : Umur

Type : Numeric

Width : 3

Decimal : 0 (karena tidak memerlukan angka dibelakang koma)

Label : Umur Responden

Values : None (tidak ada pengkodean numerik untuk variabel ini)

Missing : None (karena informasi mengenai umur tersedia pada semua responden)

Column : 5

Align : Rigth (untuk data numerik sebaiknya dibuat rata kanan)

Measure : Scale (karena umur merupakan data berskala ratio)

Variabel keempat:

Nama Variabel : Pendidikan

Type : Numeric

Width : 2 (karena pendidikan akan diinput dengan kode 1 – 5)

Decimal : 0 (karena tidak memerlukan angka dibelakang koma)

Label : Pendidikan Responden

Values : 1 = SD, 2= SLTP, 3= SLTA, 4= D3, 5= S1

Missing : None (karena informasi mengenai pendidikan tersedia pada semua responden)

Column : 8

Align : Rigth (untuk data numerik sebaiknya dibuat rata kanan)

Measure : Ordinal (karena pendidikan merupakan data berskala ordinal)

Variabel kelima:

Nama Variabel : Pendapatan

Type : Numeric

Width : 4

Decimal : 0 (karena tidak memerlukan angka dibelakang koma)

Label : Pendapatan Responden (dalam ribuan Rp)

Values : None (tidak ada pengkodean numerik untuk variabel ini)

Missing : terdapat responden yang tidak memiliki informasi mengenai pendapatan. Untuk itu, sebagai latihan kita berikan kode 9999 untuk responden yang tidak kita dapatkan informasi pendapatannya tersebut

Column : 9

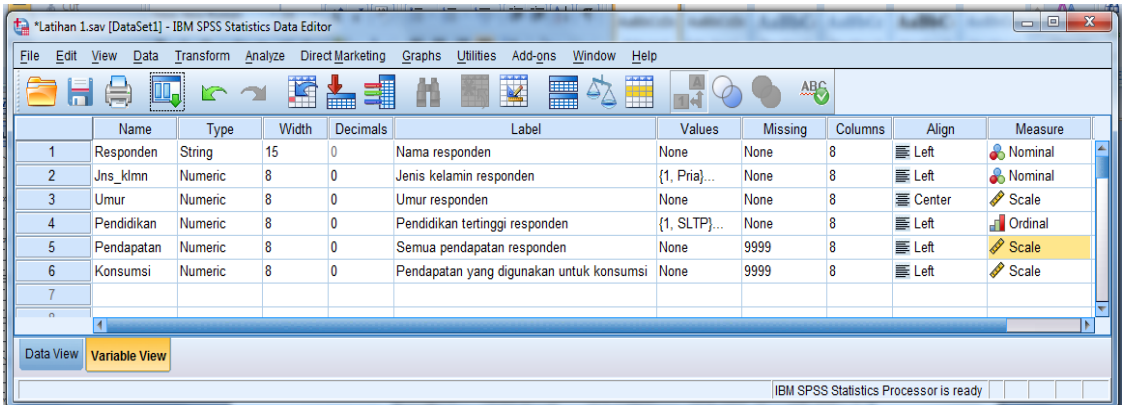
Align : Rigth (untuk data numerik sebaiknya dibuat rata kanan)

Measure : Scale (karena pendapatan merupakan data berskala ratio)

Variabel keenam:

Semua isian data sama dengan variable ke lima kecuali nama variable.

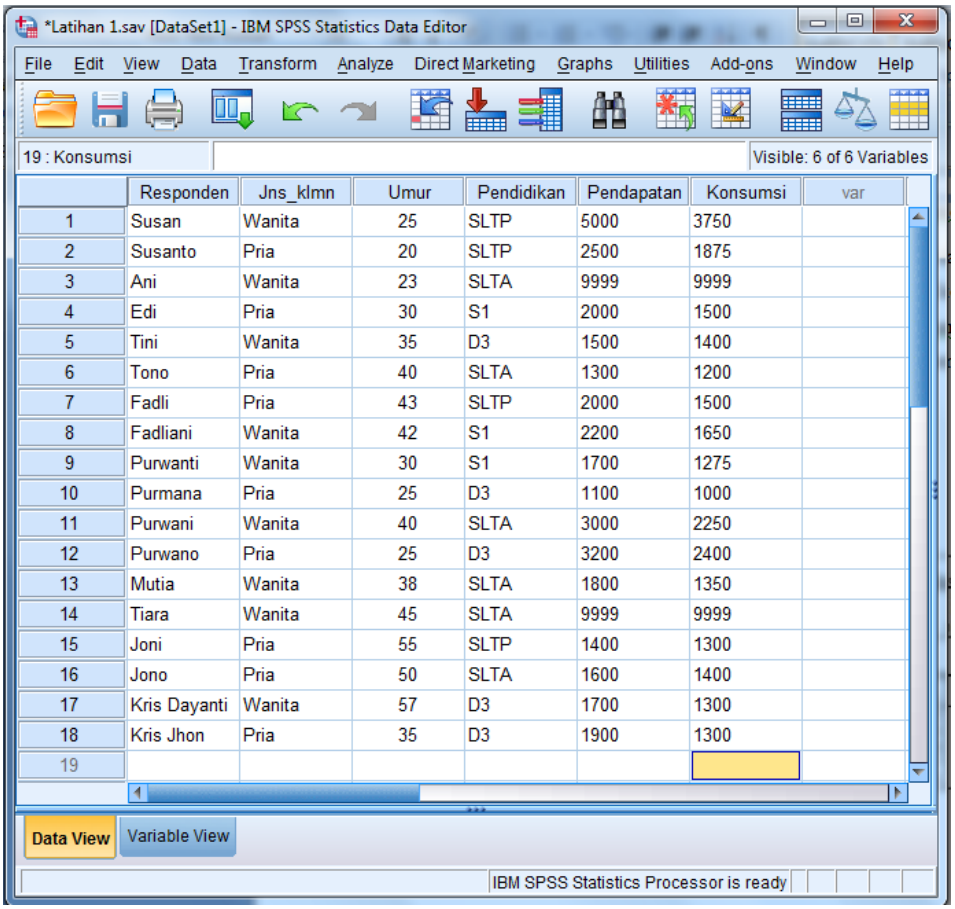
Hasil dari tahapan-tahapan yang kita lakukan akan memberikan tampilan kira-kira sebagai berikut:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Responden	String	15	0	Nama responden	None	None	8	Left	Nominal
2	Jns_klmn	Numeric	8	0	Jenis kelamin responden	{1, Pria}...	None	8	Left	Nominal
3	Umur	Numeric	8	0	Umur responden	None	None	8	Center	Scale
4	Pendidikan	Numeric	8	0	Pendidikan tertinggi responden	{1, SLTP}...	None	8	Left	Ordinal
5	Pendapatan	Numeric	8	0	Semua pendapatan responden	None	9999	8	Left	Scale
6	Konsumsi	Numeric	8	0	Pendapatan yang digunakan untuk konsumsi	None	9999	8	Left	Scale
7										

Gambar 1.6 Tampilan Variable View yang telah didefenisikan

Setelah itu, klik menu Data View yang ada disudut kiri bawah, dan data siap diinput kemudian simpan data tersebut dengan nama **Latihan distribusi frekuensi**. Berdasarkan data latihan sebelumnya, tampilan input data adalah sebagai berikut:



	Responden	Jns_klmn	Umur	Pendidikan	Pendapatan	Konsumsi	var
1	Susan	Wanita	25	SLTP	5000	3750	
2	Susanto	Pria	20	SLTP	2500	1875	
3	Ani	Wanita	23	SLTA	9999	9999	
4	Edi	Pria	30	S1	2000	1500	
5	Tini	Wanita	35	D3	1500	1400	
6	Tono	Pria	40	SLTA	1300	1200	
7	Fadli	Pria	43	SLTP	2000	1500	
8	Fadliani	Wanita	42	S1	2200	1650	
9	Purwanti	Wanita	30	S1	1700	1275	
10	Purmana	Pria	25	D3	1100	1000	
11	Purwani	Wanita	40	SLTA	3000	2250	
12	Purwano	Pria	25	D3	3200	2400	
13	Mutia	Wanita	38	SLTA	1800	1350	
14	Tiara	Wanita	45	SLTA	9999	9999	
15	Joni	Pria	55	SLTP	1400	1300	
16	Jono	Pria	50	SLTA	1600	1400	
17	Kris Dayanti	Wanita	57	D3	1700	1300	
18	Kris Jhon	Pria	35	D3	1900	1300	
19							

Gambar 1.7 Tampilan Data View yang telah didefenisikan

Latihan

1. Diberikan data mengenai jenis kelamin (sex), daerah, tingkat pendidikan dan pendapatan dari 60 orang responden penelitian. Simpan data dengan nama **Latihan tabel dan grafik**.
   Jenis kelamin dikode sebagai berikut:
   1 = laki-laki
   2 = perempuan
   Daerah dikode sebagai berikut:
   1 = Kota
   2 = Desa
   Tingkat pendidikan dikategorikan sebagai:
   1 = responden yang berpendidikan SLTP kebawah
   2 = responden yang berpendidikan SLTA
   3 = responden yang berpendidikan D3 ke atas.

Pendapatan kelompokkan atas tiga yaitu:
1 = pendapatan rendah (kurang dari Rp. 1.000.000)
2 = pendapatan menengah (Rp. 1.000.000 – 3.000.000)
3 = pendapatan tinggi (diatas Rp 3.000.000)

2. Diberikan data mengenai investasi dan tabungan domestik Indonesia (dalam trilyun Rp) selama periode tahun 1996 – 2013. Simpan data dengan nama **Latihan grafik garis**.

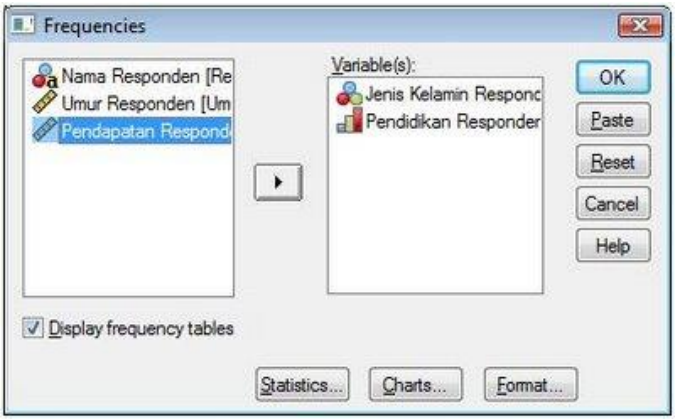
Tahun	Investasi	Tabungan
1996	58,9	53,7
1997	68,0	61,1
1998	75,4	69,7
1999	93,4	86,8
2000	115,8	108,3
2001	142,3	130,8
2002	157,8	138,8
2003	178,3	163,8
2004	242,8	283,8
2005	205,6	283,7
2006	220,1	249,2
2007	375,1	450,2
2008	380,7	441,0
2009	405,3	465,6
2010	515,3	518,8
2011	657,6	738,7
2012	805,4	895,0
2013	983,9	1084,3

BAB II

Distribusi Frekuensi dan Pengelompokan Data

A. Distribusi Frekuensi

Dari file Latihan distribusi frekuensi, dapat dibuat distribusi frekuensi untuk jenis kelamin dan pendidikan, sedangkan untuk distribusi frekuensi umur dan pendapatan, sebaiknya data dikelompokkan dulu. Untuk menampilkan distribusi frekuensi pada SPSS adalah dengan cara klik menu **Analyze → Descriptive Statistics → Frequencies**. Akan muncul tampilan berikut:



Gambar 2.1 Tampilan Frequencies.

Masukkan (pindahkan) variabel jenis kelamin dari kotak sebelah kiri ke kotak sebelah kanan dengan cara klik variabel jenis kelamin di kotak kiri, kemudian klik panah yang menuju ke sebelah kanan. Lakukan hal yang sama untuk variabel pendidikan. Kemudian klik OK. Output dari distribusi frekuensi diberikan sebagai berikut:

Statistics

		Jenis Kelamin Responden	Pendidikan Responden
N	Valid	18	18
	Missing	0	0

Jenis Kelamin Responden

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Laki-Laki	9	50.0	50.0	50.0
	Perempuan	9	50.0	50.0	100.0
	Total	18	100.0	100.0	

Pendidikan Responden

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	SLTP	4	22.2	22.2	22.2
	SLTA	6	33.3	33.3	55.6
	D3	5	27.8	27.8	83.3
	S1	3	16.7	16.7	100.0
	Total	18	100.0	100.0	

Tabel pertama memberikan keterangan mengenai variabel yang diolah, yaitu jumlah observasinya dan jumlah observasi missing. Dari tabel tersebut terlihat bahwa jumlah observasi sebanyak 18 dan tidak ada observasi missing.

Tabel kedua dan ketiga masing-masingnya memberikan distribusi frekuensi untuk jenis kelamin responden dan pendidikan responden. Tabel distribusi frekuensi menampilkan lima kolom sebagai berikut:

- Kolom pertama: kategori yang difrekuensikan (sebagai contoh pada jenis kelamin adalah pria dan wanita)
- Kolom kedua: frekuensi masing-masing kategori
- Kolom ketiga: persentase frekuensi masing-masing kategori (persentase dihitung dari total observasi termasuk observasi missing)
- Kolom keempat : persentase frekuensi masing-masing kategori tetapi persentase dihitung dengan mengeluarkan observasi missing. (Catatan: berhubung tidak ada observasi missing, baik untuk jenis kelamin maupun pendidikan, maka kolom 3 dan 4 menjadi sama).
- Kolom kelima: Cumulative Percent yaitu persentase kumulatif yang dihitung dari valid percent. Sebagai contoh pada tabel frekuensi pendidikan. Baris pertama adalah 22,2 persen. Pada baris kedua adalah 55,6 persen yang dihitung dari 22,2 + 33,3 (catatan: perbedaan perhitungan karena pembulatan).

B. Pengelompokkan Data

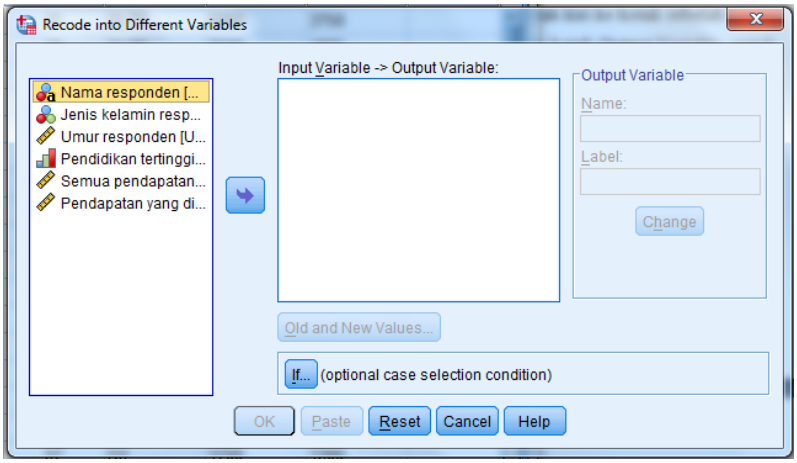
Selain menampilkan ukuran statistik deskriptif dari kumpulan data, kita juga dapat menampilkan distribusi frekuensi dari data tersebut. Jika data sudah dalam bentuk kategori (misalnya pendidikan SD, SLTP,SLTA dst) atau data sudah dikategorikan dalam kelompok-kelompok interval tertentu misalnya pendapatan rendah (< 1.000.000), menengah (1.000.000 – 2.000.000), tinggi (>2.000.000), maka kita dapat secara langsung membuat distribusi frekuensinya. Tetapi jika data belum terkelompok dalam kategori-kategori tertentu, tabel distribusi frekuensinya akan sangat panjang mengikuti keragaman dari nilai-nilai data tersebut. Misalnya jika dibentuk tabel distribusi frekuensi dari data umur, maka akan terbentuk tabel distribusi yang tidak ringkas, sehingga kita akan sulit menarik kesimpulan dari data tersebut, seperti terlihat dibawah ini:

Umur Responden

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20	1	5.6	5.6	5.6
	23	1	5.6	5.6	11.1
	25	3	16.7	16.7	27.8
	30	2	11.1	11.1	38.9
	35	2	11.1	11.1	50.0
	38	1	5.6	5.6	55.6
	40	2	11.1	11.1	66.7
	42	1	5.6	5.6	72.2
	43	1	5.6	5.6	77.8
	45	1	5.6	5.6	83.3
	50	1	5.6	5.6	88.9
	55	1	5.6	5.6	94.4
	57	1	5.6	5.6	100.0
	Total	18	100.0	100.0	

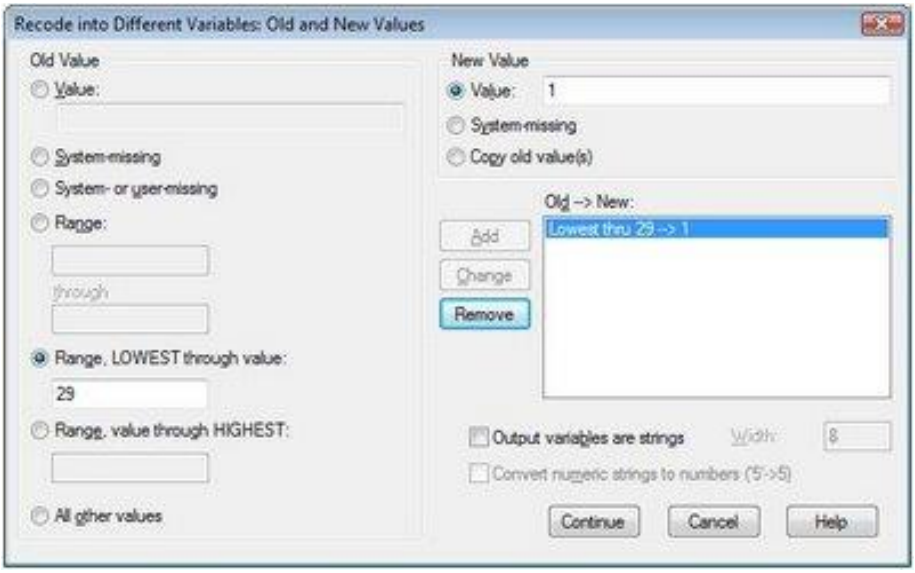
Gambar 2.3 Tampilan output distribusi frekuensi variable umur

Agar tabel distribusi frekuensi menjadi lebih ringkas sehingga mudah diinterpretasikan, data umur sebaiknya dikelompokkan terlebih dahulu dengan cara klik menu **Transform**. Ada dua pilihan dalam mengelompokkan data yaitu **Recode into Same Variables** (kode pengelompokkan akan menindih data asli) dan **Recode into Different Variables** (kode pengelompokan akan dibuat pada variabel yang baru). Pilih Recode into Different Variables, maka akan muncul tampilan berikut:



Gambar 2.4 Tampilan kotak dialog Recode into Different Variables

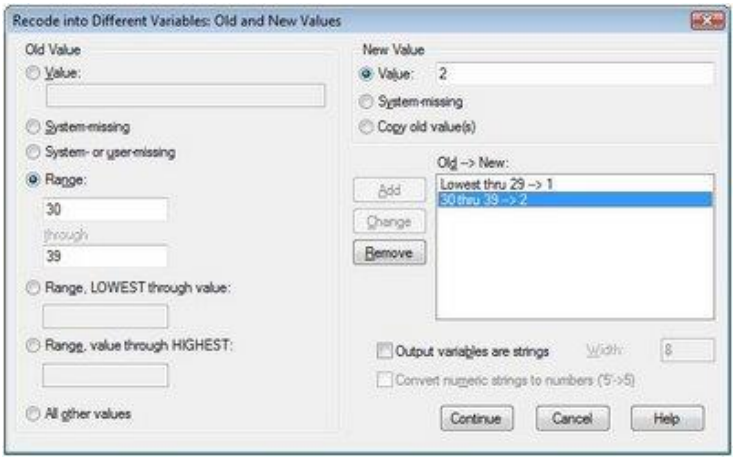
Pindahkan variable umur ke kotak Input Variables → Output Variables. Kemudian pada kotak Output Variable, untuk Name tuliskan Kel.Umur dan untuk Label tuliskan Kelompok Umur, kemudian klik **Change → Old and New Values**, akan muncul tampilan berikut:



Gambar 2.5 Tampilan Recode into Different Old and New Values

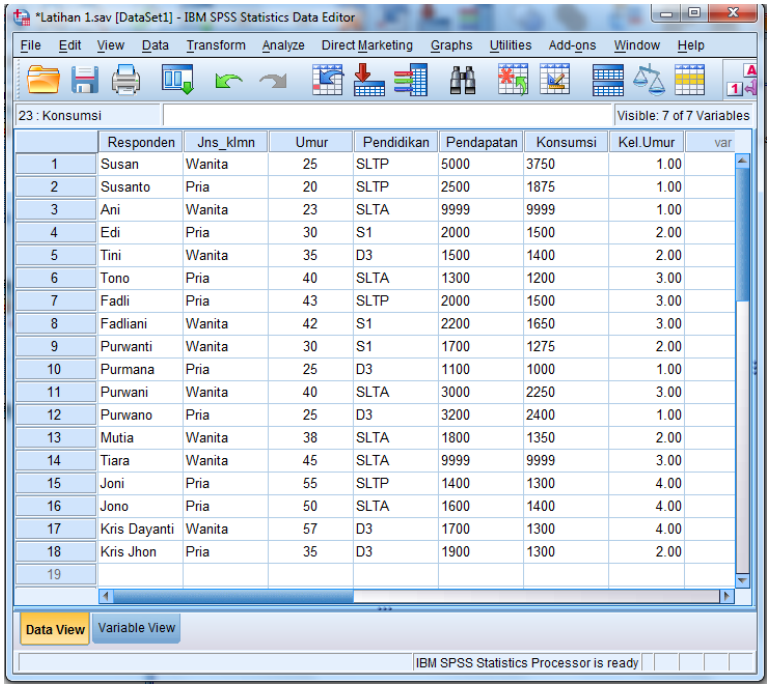
Misalnya variable umur akan dikelompokkan menjadi ≤ 29, 30 – 39, 40 – 49, dan ≥ 50. Untuk mengelompokkan umur ≤ 29, pada bagian **Old Value**, klik **Range, LOWEST through value**, kemudian isikan pada kotak dibawahnya angka 29. Pada bagian **New Value**, pada kotak **Value** isikan angka 1, lalu klik **Add** (lihat tampilan diatas). Untuk mengelompokkan umur 30 – 39, pada bagian **Old Value**, klik **Range**, kemudian isikan pada kotak dibawahnya angka 30 dan kotak dibawah **through**

angka 39. Selanjutnya pada bagian **New Value**, pada kotak **Value** isikan angka 2, kemudian klik **Add** (lihat tampilan berikut ini)



Gambar 2.6 Tampilan Recode into Different Old and New Values

Dengan cara yang sama lakukan untuk kelompok umur 40 – 49 dan pada New Value beri kode 3. Selanjutnya untuk kelompok umur ≥ 50, pada bagian **Old Value**, klik **Range, value through HIGHEST**, isikan pada kotak dibawahnya angka 50. Selanjutnya pada bagian **New Value**, pada kotak **Value** isikan angka 4, lalu klik **Add**. Setelah selesai memberikan kode untuk pengelompokan umur ini, kemudian klik **Continue** dan **OK**. Hasilnya pada sheet Data View akan ada tambahan variabel baru yaitu kelompok umur sebagai berikut:



	Responden	Jns_klmn	Umur	Pendidikan	Pendapatan	Konsumsi	Kel.Umur	var
1	Susan	Wanita	25	SLTP	5000	3750	1.00	
2	Susanto	Pria	20	SLTP	2500	1875	1.00	
3	Ani	Wanita	23	SLTA	9999	9999	1.00	
4	Edi	Pria	30	S1	2000	1500	2.00	
5	Tini	Wanita	35	D3	1500	1400	2.00	
6	Tono	Pria	40	SLTA	1300	1200	3.00	
7	Fadli	Pria	43	SLTP	2000	1500	3.00	
8	Fadliani	Wanita	42	S1	2200	1650	3.00	
9	Purwanti	Wanita	30	S1	1700	1275	2.00	
10	Purmana	Pria	25	D3	1100	1000	1.00	
11	Purwani	Wanita	40	SLTA	3000	2250	3.00	
12	Purwano	Pria	25	D3	3200	2400	1.00	
13	Mutia	Wanita	38	SLTA	1800	1350	2.00	
14	Tiara	Wanita	45	SLTA	9999	9999	3.00	
15	Joni	Pria	55	SLTP	1400	1300	4.00	
16	Jono	Pria	50	SLTA	1600	1400	4.00	
17	Kris Dayanti	Wanita	57	D3	1700	1300	4.00	
18	Kris Jhon	Pria	35	D3	1900	1300	2.00	
19								

Berikan keterangan untuk masing-masing kode pengelompokan umur tersebut pada Value Label dengan kode 1 (≤ 29), kode 2 (30-39), kode 3 (40 – 49), kode 4 (≥ 50). Lalu bentuklah distribusi frekuensi untuk kelompok umur tersebut sehingga output adalah sebagai berikut:

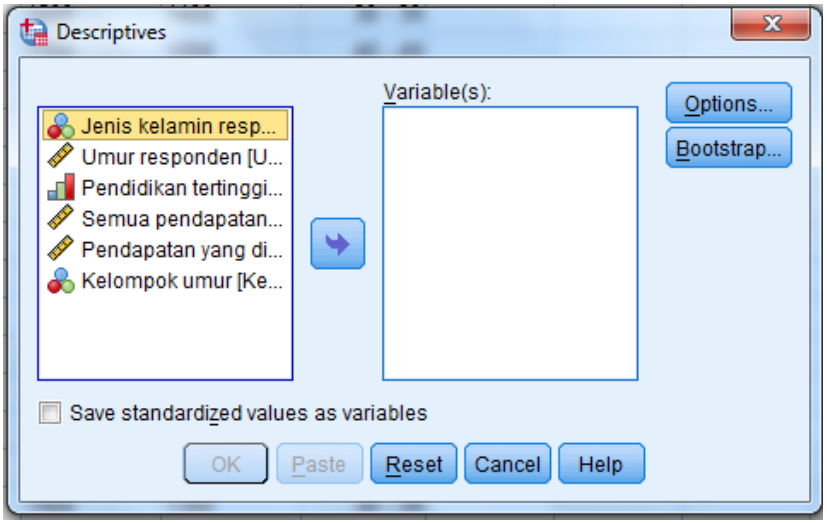
Kelompok Umur

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	5	27.8	27.8	27.8
	2.00	5	27.8	27.8	55.6
	3.00	5	27.8	27.8	83.3
	4.00	3	16.7	16.7	100.0
	Total	18	100.0	100.0	

BAB III
Statistik Deskriptif

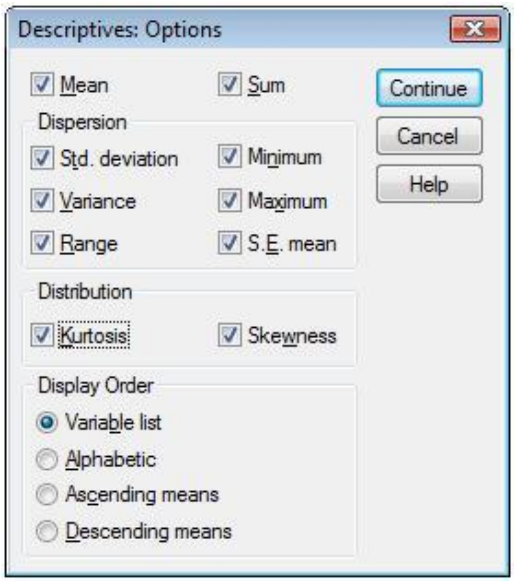
Ukuran-ukuran statistik deskriptif dalam pengolahan data bertujuan untuk mendapatkan gambaran ringkas dari sekumpulan data, sehingga dapat disimpulkan keadaan data secara mudah dan cepat. Selain itu, melalui ukuran-ukuran statistik deskriptif ini, dapat ditentukan jenis pengolahan statistik lebih lanjut yang sesuai dengan karakteristik data tersebut.

Pada file Latihan distribusi frekuensi terdapat data umur dan pendapatan dari 18 responden penelitian yang telah diinput pada SPSS. Perhatikan pada responden ketiga dan keempat belas, pendapatannya tertulis 9999. Angka tersebut bukanlah pendapatan dari responden, tetapi adalah kode untuk "missing" data (data yang tidak tersedia). Selanjutnya untuk mendapatkan ukuran-ukuran statistik deskriptif, klik **Analyze → Descriptive Statistics → Descriptives**. Akan muncul tampilan berikut:



Gambar 3.1 Kotak dialog Descriptives

Pindahkan variabel umur dan pendapatan (yang tadinya ada dikotak sebelah kiri) ke kotak sebelah kanan, dengan cara klik variabel yang bersangkutan, kemudian klik panah yang menuju ke arah kanan. Selanjutnya, klik **Options**, akan muncul tampilan berikut:



Gambar 3.2 Kotak dialog pilihan Descriptives

Terdapat berbagai pilihan ukuran numerik statistik deskriptif. Sebagai latihan, klik semua pilihan tersebut. Selain itu, terdapat juga pilihan **Display Order** (urutan tampilan output).

- **Variable list**, output akan ditampilkan dengan urutan sesuai dengan urutan variabel yang diinput (dalam contoh ini, tampilan outputnya umur kemudian pendapatan).
- **Alphabetic**, output akan ditampilkan berdasarkan urutan abjad awal dari nama variabel (dalam hal ini pendapatan kemudian umur).
- **Ascending means**, urutan tampilan output dimulai dari variabel dengan rata-rata terkecil.
- **Descending means**, urutan tampilan output dimulai dari variabel dengan rata-rata terbesar.

Dalam contoh diambil pilihan **Variable list**.

Setelah mengambil pilihan-pilihan yang diinginkan, klik **Continue → OK**. Akan muncul output statistik deskriptif sebagai berikut:

	N	Range	Minimum	Maximum	Sum	Mean	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Umur responden	18	37	20	57	658	36.56	2.592
Pendapatan responden	16	3900	1100	5000	33900	2118.75	239.830
Valid N (listwise)	16						

	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Umur responden	10.999	120.967	.305	.536	-.765	1.038
Pendapatan responden	959.318	920291.667	2.017	.564	4.878	1.091
Valid N (listwise)						

Keterangan:

- Kolom pertama dari output menunjukkan variabel yang diolah.
- Kolom kedua adalah jumlah observasi. Perhatikan untuk umur responden, jumlah observasi adalah 18, untuk pendapatan responden adalah 16. Mengapa? Karena dua observasi sesuai dengan contoh latihan adalah data missing. SPSS dalam hal ini hanya akan mengolah data yang valid dengan mengeluarkan data missing.
- Kolom ketiga adalah range (jarak). Range merupakan pengukuran yang paling sederhana untuk dispersi (penyebaran) data. Rumus untuk range adalah nilai maksimum dikurangkan nilai minimum. Dalam contoh, misalnya range untuk umur adalah 37, karena nilai maksimum 57 dan nilai minimum 20.
- Kolom keempat adalah nilai minimum (terendah) dari data
- Kolom kelima adalah nilai maksimum (tertinggi) dari data
- Kolom keenam adalah jumlah (sum) dari keseluruhan data.
- Kolom ketujuh adalah nilai rata-rata, yaitu jumlah dibagi dengan banyaknya observasi. Dalam kasus umur = 658/18 = 36.56
- Kolom kedelapan adalah standar error dari rata-rata (Standard error of Mean). Ini adalah pengukuran untuk mengukur seberapa jauh nilai rata-rata bervariasi dari

satu sampel ke sampel lainnya yang diambil dari distribusi yang sama. Cara menghitung standard error dari rata-rata (misalnya untuk umur) adalah:

$$SE = \frac{S}{\sqrt{n}} = \frac{10,999}{\sqrt{18}} = 2,592$$

Keterangan

SE : standar error dari rata-rata

S : adalah standar deviasi (lihat kolom 9)

n : adalah jumlah observasi

Apa perbedaan standard error (of mean) dengan standar deviasi (kolom kesembilan)? Kalau standard deviasi adalah suatu indeks yang menggambarkan sebaran data terhadap rata-ratanya, maka standard error (of mean) adalah indeks yang menggambarkan sebaran rata-rata sampel terhadap rata-rata dari rata-rata keseluruhan kemungkinan sampel (rata-rata populasi). Pengukuran ini berguna, terutama untuk menjawab pertanyaan "seberapa baik rata-rata yang kita dapatkan dari data sampel dapat mengestimasi rata-rata populasi?"

- Kolom kesembilan adalah standar deviasi, yang dihitung dengan rumus:

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

Sebagai contoh perhitungan untuk standard deviasi umur sebagai berikut:

X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
25	-11.56	133.53
20	-16.56	274.09
23	-13.56	183.75
30	-6.56	42.98
35	-1.56	2.42
40	3.44	11.86
43	6.44	41.53
42	5.44	29.64
30	-6.56	42.98
25	-11.56	133.53
40	3.44	11.86
25	-11.56	133.53
38	1.44	2.09
45	8.44	71.31
55	18.44	340.20
50	13.44	180.75
57	20.44	417.98
35	-1.56	2.42
Jumlah		2056.44

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1}\sum(Xi - \bar{X})^2} = \sqrt{\frac{1}{18-1}(2056.44)} = 10.999$$

- Kolom kesepuluh adalah varians dari data. Secara matematis, varians dan standar deviasi saling terkait, dimana standar deviasi adalah akar varians, atau varians adalah kuadrat dari standar deviasi. Dengan demikian untuk varians umur adalah $10.999^2 = 120.967$
- Kolom kesebelas adalah skewness data. Skewness merupakan alat ukur dalam menelusuri distribusi data yang diperbandingkan dengan distribusi normal. Skewness merupakan pengukuran tingkat ketidaksimetrisan (kecondongan) sebaran data di sekitar rata-ratanya. Distribusi normal merupakan distribusi yang simetris dan nilai skewness adalah 0. Skewness yang bernilai positif menunjukkan ujung dari kecondongan menjulur ke arah nilai positif (ekor kurva sebelah kanan lebih panjang). Skewness yang bernilai negatif menunjukkan ujung dari kecondongan menjulur ke arah nilai negatif (ekor kurva sebelah kiri lebih panjang). Rumus skewness adalah sebagai berikut:

$$Skewness = \frac{n}{(n-1)(n-2)}\sum\left(\frac{Xi - \bar{X}}{S}\right)^3$$

Sebagai contoh, perhitungan skewness untuk data umur adalah sebagai berikut:

Xi	$\left(\frac{Xi-X}{S}\right)^3$
25	-1.16
20	-3.41
23	-1.87
30	-0.21
35	0.00
40	0.03
43	0.20
42	0.12
30	-0.21
25	-1.16
40	0.03
25	-1.16
38	0.00
45	0.45
55	4.72
50	1.83
57	6.42
35	0.00
Jumlah	4.61

$$Skewness = \frac{18}{(18-1)(18-2)}(4.61) = 0{,}305$$

- Kolom keduabelas adalah standar error dari skewness. Untuk menghitung standar error dari skewness ini (sebagai contoh umur) adalah sebagai berikut:

$$SE\ \ skewness = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} = \sqrt{\frac{6(18)(18-1)}{(18-2)(18+1)(18+3)}} = 0.536$$

- Kolom ketiga belas adalah Kurtosis. Sebagaimana skewness, kurtosis juga merupakan alat ukur dalam menelusuri distribusi data yang diperbandingkan dengan distribusi normal. Kurtosis menggambarkan keruncingan (peakedness) atau kerataan (flatness) suatu distibusi data dibandingkan dengan distribusi normal. Pada distribusi normal, nilai kurtosis sama dengan 0. Nilai kurtosis yang positif menunjukkan distribusi yang relatif runcing, sedangkan nilai kurtosis yang negatif menunjukkan distribusi yang relatif rata.

  Rumus kurtosis adalah:

$$Kurtosis = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{Xi - \bar{X}}{S} \right)^4 \right\} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

Contoh perhitungan untuk data umur sebagai berikut:

(X)	$\left(\frac{Xi-\bar{X}}{S}\right)^4$
25	1.22
20	5.13
23	2.31
30	0.13
35	0.00
40	0.01
43	0.12
42	0.06
30	0.13
25	1.22
40	0.01
25	1.22
38	0.00
45	0.35
55	7.91
50	2.23
57	11.94
35	0.00
Jumlah	33.98

Sehingga kurtosisnya adalah:

$$Kurtosis = \left\{ \frac{18(19)}{(17)(16)(15)} (33.98) \right\} - \frac{3(17)^2}{(16)(15)} = -0{,}765$$

- Kolom keempat belas adalah standar error dari kurtosis, yang dihitung dengan rumus berikut:

$$SE\ kurtosis = \sqrt{\frac{4(n^2-1)(SEs^2)}{(N-3)(N+5)}}$$

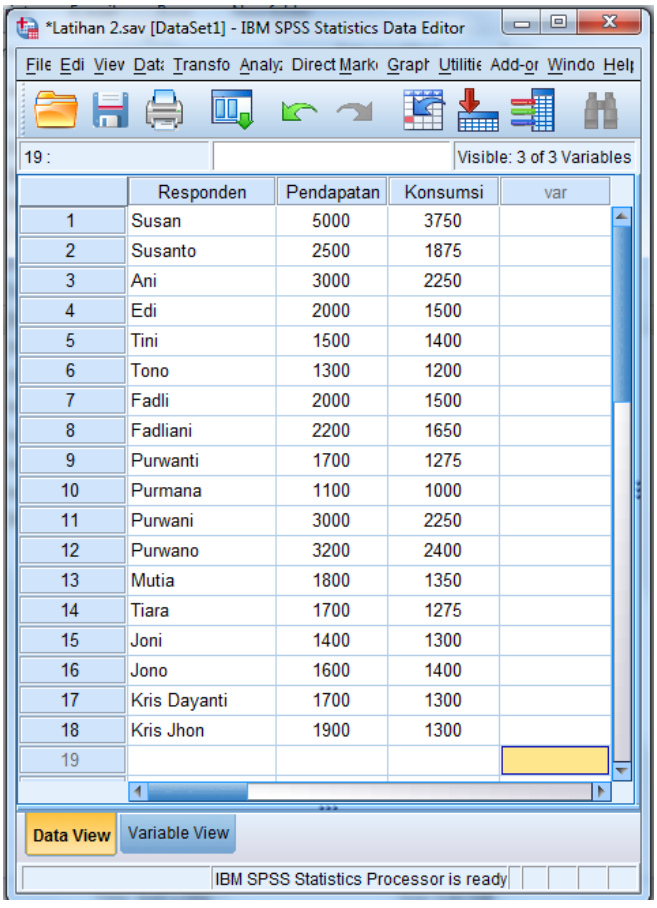
Dimana Ses adalah Standar error dari skewness yang telah kita hitung sebelumnya. Dengan demikian, standar error kurtosis untuk kasus umur adalah:

$$SE\ kurtosis = \sqrt{\frac{4(18^2-1)(0.536^2)}{(18-3)(18+5)}} = 1.038$$

BAB IV

Transformasi Variabel

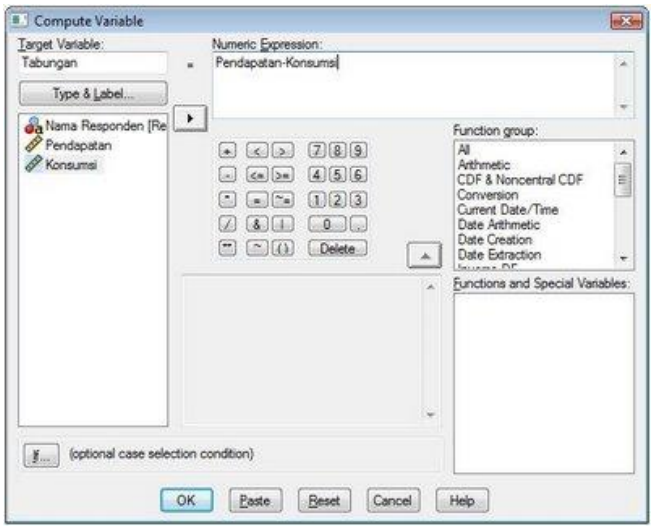
Salah satu keunggulan SPSS sebagai salah satu software statistik adalah fasilitas transformasi variabelnya yang dapat dilakukan secara mudah seperti program-program spreadsheet lainnya (misalnya Excel). Sebagai latihan sederhana, gunakan file Latihan distribusi frekuensi yang sudah diinput pada SPSS sebagai berikut (isikan data ketiga dan keempat belas seperti yang terlihat pada tampilan dibawah):



	Responden	Pendapatan	Konsumsi	var
1	Susan	5000	3750	
2	Susanto	2500	1875	
3	Ani	3000	2250	
4	Edi	2000	1500	
5	Tini	1500	1400	
6	Tono	1300	1200	
7	Fadli	2000	1500	
8	Fadliani	2200	1650	
9	Purwanti	1700	1275	
10	Purmana	1100	1000	
11	Purwani	3000	2250	
12	Purwano	3200	2400	
13	Mutia	1800	1350	
14	Tiara	1700	1275	
15	Joni	1400	1300	
16	Jono	1600	1400	
17	Kris Dayanti	1700	1300	
18	Kris Jhon	1900	1300	
19				

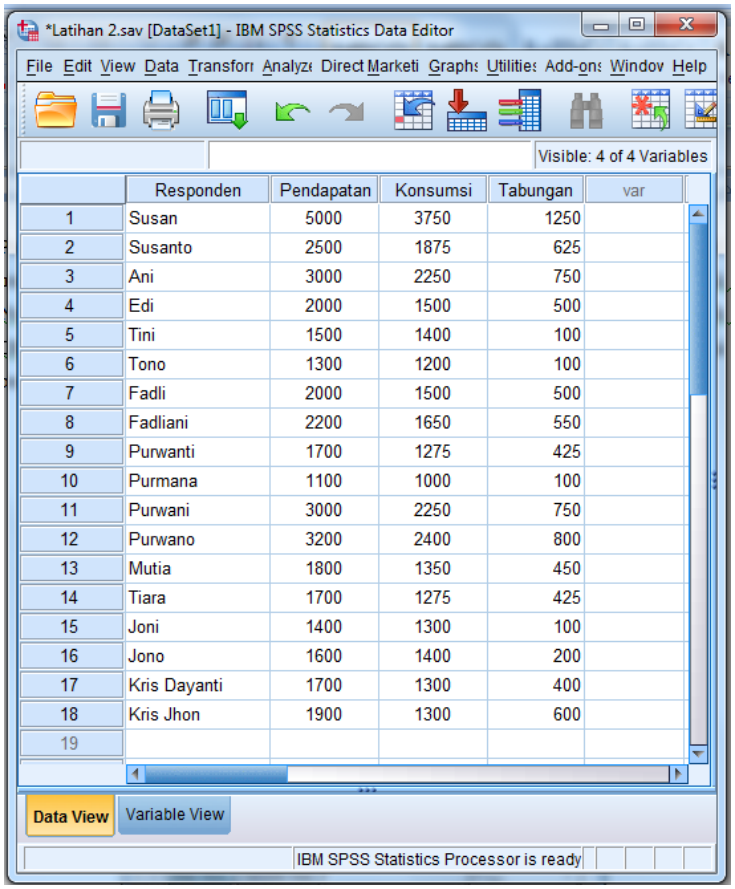
Gambar 4.1 Data transformasi variabel

Berdasarkan data tersebut, ingin dihitung tabungan masing-masing responden (yang berasal dari pendapatan dikurangi konsumsi), yang akan digunakan untuk pengolahan data lebih lanjut. Untuk melakukan itu, klik **Transform → Compute Variable**. Akan muncul tampilan berikut:



Gambar 4.2 Tampilan compute variable

Pada kotak Target Variable, masukkan nama variabel untuk menampung data hasil transformasi variabel tersebut. Misalnya dalam contoh diberi nama Tabungan. Pada kotak Numeric Expression, isikan perintah transformasinya, yaitu **Pendapatan – Konsumsi** lalu klik OK. Pada worksheet SPSS akan muncul variabel baru dengan nama Tabungan sebagai berikut:



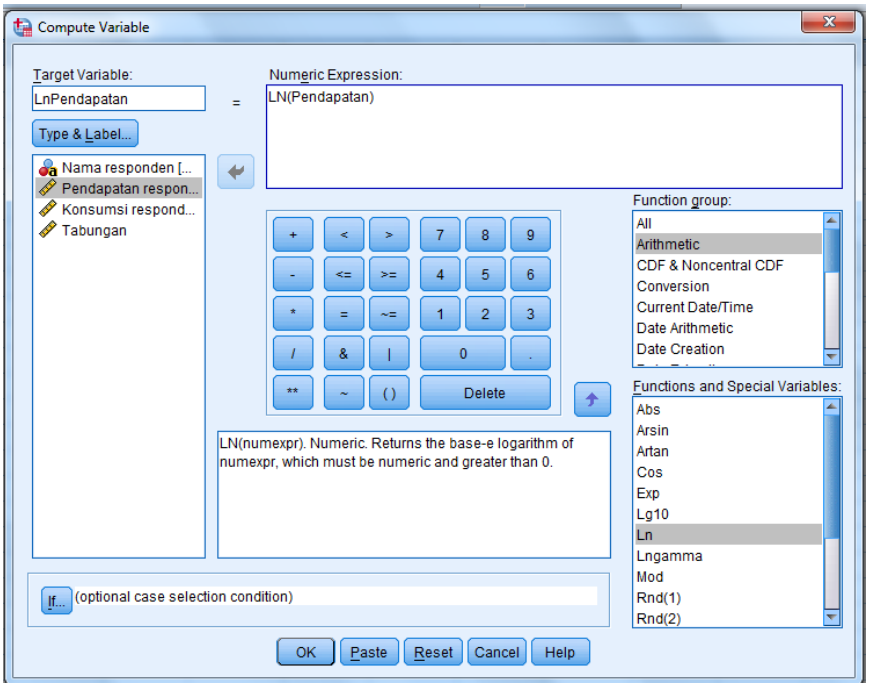
	Responden	Pendapatan	Konsumsi	Tabungan	var
1	Susan	5000	3750	1250	
2	Susanto	2500	1875	625	
3	Ani	3000	2250	750	
4	Edi	2000	1500	500	
5	Tini	1500	1400	100	
6	Tono	1300	1200	100	
7	Fadli	2000	1500	500	
8	Fadliani	2200	1650	550	
9	Purwanti	1700	1275	425	
10	Purmana	1100	1000	100	
11	Purwani	3000	2250	750	
12	Purwano	3200	2400	800	
13	Mutia	1800	1350	450	
14	Tiara	1700	1275	425	
15	Joni	1400	1300	100	
16	Jono	1600	1400	200	
17	Kris Dayanti	1700	1300	400	
18	Kris Jhon	1900	1300	600	
19					

Pada prinsipnya, dalam perintah transformasi ini ada dua cara yang dapat dilakukan:

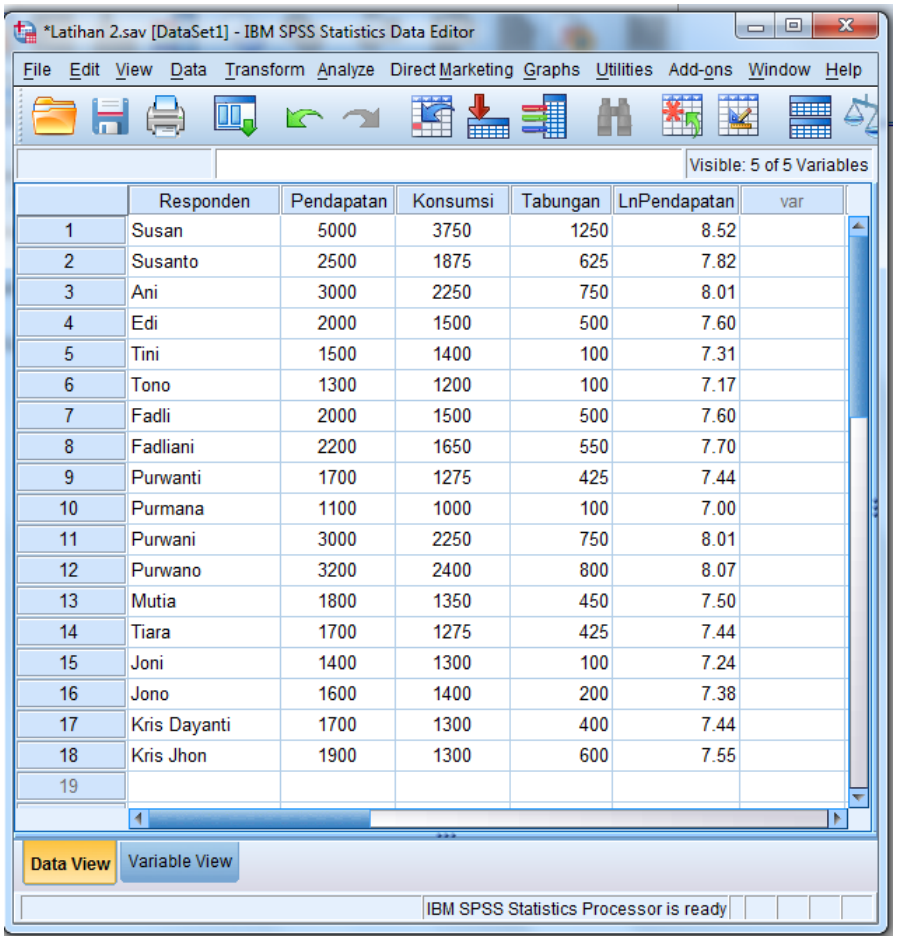
a. Membuat rumus sendiri, seperti contoh diatas.

b. Menggunakan fungsi yang sudah disediakan SPSS dalam kotak Function Group.

SPSS menyediakan berbagai fungsi baik fungsi aritmatik, statistik, fungsi waktu dan lainnya. Salah satu contoh memanfaatkan fungsi ini, misalnya ingin menghitung logaritma natural (Ln) variabel pendapatan. Fungsi tersebut disediakan oleh SPSS dalam Fungsi Arithmethic. Klik **Arithmethic**, maka akan muncul di kotak **Function and Special Variables**, fungsi-fungsi yang tersedia pada kelompok arithmetic seperti tampilan dibawah ini.



Isikan pada kotak **Target Variable**, nama variabel sebagai penampung transformasi tersebut. Misalnya **LnPendapatan**. Kemudian klik **Ln**, dan klik tanda panah yang mengarah ke atas pada gambar diatas. Maka pada kotak **Numeric Expression** akan muncul tulisan **LN(?)**. Selanjutnya klik variabel Pendapatan (variabel asal yang akan ditransformasikan), dan klik panah yang mengarah kekanan dari gambar diatas. Prosedur ini akan mengganti tanda tanya diatas menjadi Pendapatan, sehingga tulisan pada kotak Numeric Expression menjadi **LN(Pendapatan)**. Setelah itu klik **OK**, maka akan akan keluar output dari logaritma natural dari pendapatan yang berada pada variabel baru yang bernama LnPendapatan seperti tampilan berikut:



	Responden	Pendapatan	Konsumsi	Tabungan	LnPendapatan	var
1	Susan	5000	3750	1250	8.52	
2	Susanto	2500	1875	625	7.82	
3	Ani	3000	2250	750	8.01	
4	Edi	2000	1500	500	7.60	
5	Tini	1500	1400	100	7.31	
6	Tono	1300	1200	100	7.17	
7	Fadli	2000	1500	500	7.60	
8	Fadliani	2200	1650	550	7.70	
9	Purwanti	1700	1275	425	7.44	
10	Purmana	1100	1000	100	7.00	
11	Purwani	3000	2250	750	8.01	
12	Purwano	3200	2400	800	8.07	
13	Mutia	1800	1350	450	7.50	
14	Tiara	1700	1275	425	7.44	
15	Joni	1400	1300	100	7.24	
16	Jono	1600	1400	200	7.38	
17	Kris Dayanti	1700	1300	400	7.44	
18	Kris Jhon	1900	1300	600	7.55	
19						

Latihan.

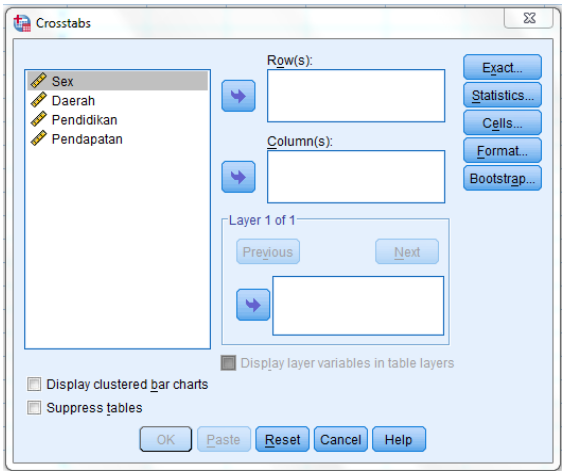
Dari contoh data latihan di atas, gunakan fungsi aritmatik, statistik, fungsi waktu dan lainnya.

BAB V
Tabel

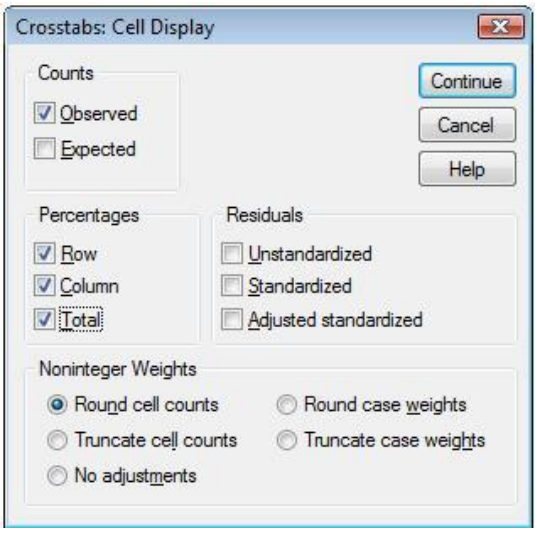
Salah satu cara untuk menggambarkan keterkaitan antar variabel secara sederhana adalah dengan membentuk tabel silang (crosstabs) antar variabel tersebut. Tabel silang adalah tabel distribusi frekuensi yang menghubungkan dua atau lebih variabel.

A. Tabel Silang Dua Variabel

Dalam contoh pembuatan table silang dua variable akan digunakan variable tingkat pendidikan dan pendapatan responden pada file **Latihan Tabel dan grafik** yang ada pada latihan pertemuan 1. Untuk membuat table ini, klik **Analyze → Descriptive Statistics → Crosstabs**. Akan muncul tampilan berikut:



Masukkan variabel **Pendapatan** ke kotak **Row(s)** dan variabel **Pendidikan** ke kotak **Column(s)**. Selanjutnya klik **Cells**, akan muncul tampilan berikut:



Klik **Observed, Row, Column, Total**. Pilihan Observed bertujuan untuk menampilkan frekuensi data sebenarnya, pilihan Row bertujuan untuk menampilkan persentase baris, column untuk persentase kolom dan total untuk persentase total dalam tabel silang. Tentunya, pilihan-pilihan persentase ini dalam prakteknya tidak akan digunakan semuanya (karena akan memperumit pembacaan tabel). Tapi untuk sekedar latihan, silakan klik saja semua pilihan persentase tersebut. Selanjutnya, klik **Continue → OK**. Akan muncul output tabel silang sebagai berikut:

Pendapatan * Pendidikan Crosstabulation

			Pendidikan			Total
			=< SLTP	SLTA	>= D3	
Pendapatan	Rendah	Count	16	7	2	25
		% within Pendapatan	64.0%	28.0%	8.0%	100.0%
		% within Pendidikan	69.6%	33.3%	12.5%	41.7%
		% of Total	26.7%	11.7%	3.3%	41.7%
	Menengah	Count	3	11	5	19
		% within Pendapatan	15.8%	57.9%	26.3%	100.0%
		% within Pendidikan	13.0%	52.4%	31.3%	31.7%
		% of Total	5.0%	18.3%	8.3%	31.7%
	Tinggi	Count	4	3	9	16
		% within Pendapatan	25.0%	18.8%	56.3%	100.0%
		% within Pendidikan	17.4%	14.3%	56.3%	26.7%
		% of Total	6.7%	5.0%	15.0%	26.7%
Total		Count	23	21	16	60
		% within Pendapatan	38.3%	35.0%	26.7%	100.0%
		% within Pendidikan	100.0%	100.0%	100.0%	100.0%
		% of Total	38.3%	35.0%	26.7%	100.0%

Dari output SPSS, Count adalah frekuensi dari data yang diamati (observed). Cara membacanya:

- Angka 16 pada baris pertama kolom pertama dari tabel diatas berarti bahwa terdapat 16 responden dengan pendidikan =< SLTP yang berpendapatan rendah.

Angka 7 pada kolom kedua baris pertama berarti bahwa terdapat 7 responden dengan pendidikan SLTA yang berpendapatan rendah.

- % within Pendapatan adalah persentase baris dari tabel silang ini. Misalnya, angka 64.0% (baris kedua kolom pertama) dari tabel diatas adalah berasal dari (16/25) x 100%. Cara membacanya adalah dari total responden berpendapatan rendah (sebanyak 25 orang), 64,0 persen diantaranya adalah mereka yang berpendidikan SLTP ke bawah.
- % within Pendidikan adalah persentase kolom dari tabel silang ini. Misalnya angka 69.6% (baris ketiga kolom pertama) dari tabel diatas adalah berasl dari (16/23) x 100%. Cara membacanya adalah dari total responden yang berpendidikan SLTP ke bawah (sebanyak 23 orang), 69,6 persen diantaranya berpendapatan rendah.
- % within total adalah persentase total dari tabel silang ini. Misalnya angka 26.7% (baris keempat kolom pertama) adalah berasal dari (16/60) x 100%. Cara membacanya adalah dari total responden (sebanyak 60 orang), 26,7 persen diantaranya memiliki pendidikan SLTP kebawah dengan pendapatan rendah.

Sekali lagi, sebagai catatan, dalam prakteknya kita tidak perlu menggunakan semua jenis persentase ini. Silakan pilih sesuai dengan kebutuhan analisis, agar tampilan tabel silang tidak ruwet seperti diatas.

B. Tabel Silang Tiga Variabel

Tabel silang pada dasarnya tidak hanya dapat dibentuk antar dua variabel, tetapi juga dapat dibentuk untuk melihat keterkaitan lebih dari dua variabel. Sebagai latihan, misalnya diberikan data mengenai jenis kelamin (sex), tingkat pendidikan dan pendapatan dari 60 orang responden penelitian (file Latihan table dan grafik).

Untuk membuat table ini, klik **Analyze → Descriptive Statistics → Crosstabs**. Masukkan variabel Pendapatan ke kotak Row(s), variabel Pendidikan ke kotak Column(s) serta variabel Sex ke kotak Layer 1of 1. (Catatan: ini hanya salah satu contoh cara mengkombinasikan penempatan Row,column dan layernya. Sebagai latihan silakan diganti-ganti dan lihat bentuk outputnya). Selanjutnya klik Cells. Untuk kepentingan latihan saat ini, silakan klik saja Observed dan Column. Pilihan Observed bertujuan untuk menampilkan frekuensi data sebenarnya, pilihan Row bertujuan untuk menampilkan persentase baris, column untuk persentase kolom dan

total untuk persentase total dalam tabel silang. Selanjutnya, klik Continue dan klik OK. Akan muncul output tabel silang sebagai berikut:

Pendapatan * Pendidikan * Sex Crosstabulation

Sex				Pendidikan =< SLTP	Pendidikan SLTA	Pendidikan >= D3	Total
Laki-Laki	Pendapatan	Rendah	Count	7	5	0	12
			% within Pendidikan	77.8%	35.7%	.0%	40.0%
		Menengah	Count	1	8	3	12
			% within Pendidikan	11.1%	57.1%	42.9%	40.0%
		Tinggi	Count	1	1	4	6
			% within Pendidikan	11.1%	7.1%	57.1%	20.0%
	Total		Count	9	14	7	30
			% within Pendidikan	100.0%	100.0%	100.0%	100.0%
Perempuan	Pendapatan	Rendah	Count	9	2	2	13
			% within Pendidikan	64.3%	28.6%	22.2%	43.3%
		Menengah	Count	2	3	2	7
			% within Pendidikan	14.3%	42.9%	22.2%	23.3%
		Tinggi	Count	3	2	5	10
			% within Pendidikan	21.4%	28.6%	55.6%	33.3%
	Total		Count	14	7	9	30
			% within Pendidikan	100.0%	100.0%	100.0%	100.0%

Cara membaca table.

Count adalah frekuensi dari data yang diamati (observed), misalnya angka 7 pada baris pertama kolom pertama dari tabel diatas berarti bahwa terdapat 7 responden laki-laki dengan pendidikan =< SLTP yang berpendapatan rendah.

% within Pendidikan adalah persentase kolom. Angka 77.8% (baris kedua kolom pertama) didapat dari (7/9) x 100% adalah total responden laki-laki yang berpendidikan SLTP ke bawah (sebanyak 9 orang), 77,8 persen diantaranya berpendapatan rendah.

C. Tabel Silang Empat Variabel

Sebagai latihan, misalnya kita punya data mengenai jenis kelamin (sex), daerah, tingkat pendidikan dan pendapatan dari 60 orang responden penelitian.

Untuk membuat table ini, klik **Analyze → Descriptive Statistics → Crosstabs**. Masukkan variabel Pendapatan ke kotak Row(s), Variabel Pendidikan ke kotak Column(s) serta Variabel Sex ke kotak Layer 1of 1. (Catatan: ini hanya salah satu contoh cara mengkombinasikan penempatan Row,column dan layer nya. Sebagai latihan silakan diganti-ganti dan lihat bentuk outputnya)

Selanjutnya klik Next, masukkan variabel Daerah ke kotak Layer 2 of 2. (Catatan: Begitu seterusnya untuk tabel silang lebih dari empat variabel)

Selanjutnya klik Cells, klik Observed dan Column. Selanjutnya, klik Continue dan klik OK. Akan muncul output tabel silang sebagai berikut:

Pendapatan * Pendidikan * Sex * Daerah Crosstabulation

Daerah	Sex				Pendidikan			Total
					=< SLTP	SLTA	>= D3	
Kota	Laki-Laki	Pendapatan	Rendah	Count	5	5	0	10
				% within Pendidikan	71.4%	41.7%	.0%	38.5%
			Menengah	Count	1	6	3	10
				% within Pendidikan	14.3%	50.0%	42.9%	38.5%
			Tinggi	Count	1	1	4	6
				% within Pendidikan	14.3%	8.3%	57.1%	23.1%
		Total		Count	7	12	7	26
				% within Pendidikan	100.0%	100.0%	100.0%	100.0%
	Perempuan	Pendapatan	Rendah	Count	2		1	3
				% within Pendidikan	100.0%		25.0%	50.0%
			Menengah	Count	0		1	1
				% within Pendidikan	.0%		25.0%	16.7%
			Tinggi	Count	0		2	2
				% within Pendidikan	.0%		50.0%	33.3%
		Total		Count	2		4	6
				% within Pendidikan	100.0%		100.0%	100.0%
Desa	Laki-Laki	Pendapatan	Rendah	Count	2	0		2
				% within Pendidikan	100.0%	.0%		50.0%
			Menengah	Count	0	2		2
				% within Pendidikan	.0%	100.0%		50.0%
		Total		Count	2	2		4
				% within Pendidikan	100.0%	100.0%		100.0%
	Perempuan	Pendapatan	Rendah	Count	7	2	1	10
				% within Pendidikan	58.3%	28.6%	20.0%	41.7%
			Menengah	Count	2	3	1	6
				% within Pendidikan	16.7%	42.9%	20.0%	25.0%
			Tinggi	Count	3	2	3	8
				% within Pendidikan	25.0%	28.6%	60.0%	33.3%
		Total		Count	12	7	5	24
				% within Pendidikan	100.0%	100.0%	100.0%	100.0%

Cara membaca tabel.

Count adalah frekuensi dari data yang diamati (observed). Angka 5 pada baris pertama kolom pertama berarti terdapat 5 responden laki-laki di kota dengan pendidikan =< SLTP yang berpendapatan rendah.

% within Pendidikan adalah persentase kolom. Angka 71.4% (baris kedua kolom pertama) didapat dari (5/7) x 100% adalah dari total responden laki-laki di kota yang berpendidikan SLTP ke bawah (sebanyak 7 orang), 71,4 persen diantaranya berpendapatan rendah.

Perhatikan bahwa dari tabel kita diatas terdapat banyak sel yang kosong. Hal ini dikarenakan jumlah observasi yang relatif sedikit dibandingkan jumlah sel yang terbentuk dari tabel silang ini. Oleh karenanya, jumlah observasi harus menjadi pertimbangan ketika kita ingin membentuk tabel silang dengan jumlah variabel atau kategori yang banyak.

BAB VI

Grafik

A. Batang Sederhana

Sebagai latihan misalnya kita punya data tingkat pendidikan dan pendapatan dari 30 orang responden penelitian. Tingkat pendidikan misalnya dikategorikan sebagai:

1 = responden yang berpendidikan SLTP kebawah

2 = responden yang berpendidikan SLTA

3 = responden yang berpendidikan D3

4 = responden yang berpendidikkan S1.

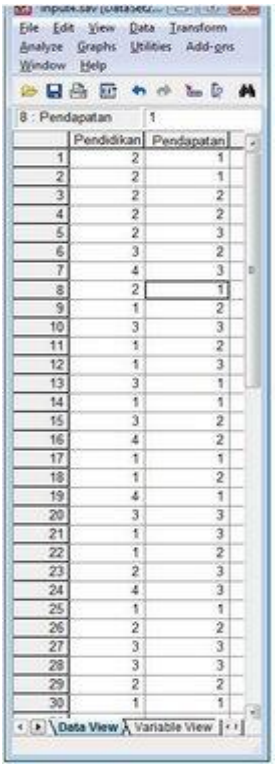
Pendapatan juga kita kelompokkan atas tiga yaitu:

1 = pendapatan rendah (kurang dari Rp. 1.000.000)

2 = pendapatan menengah (Rp. 1.000.000 – 3.000.000)

3 = pendapatan tinggi (diatas Rp 3.000.000)

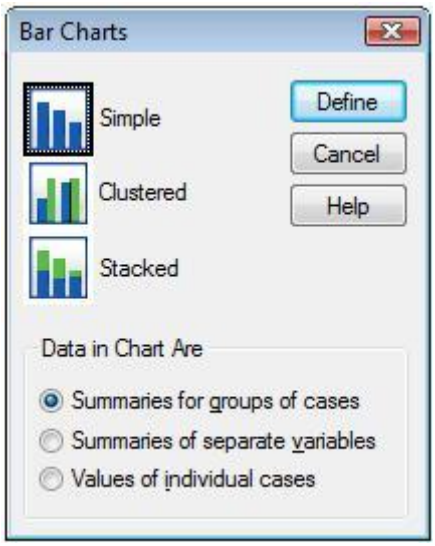
Data tingkat pendidikan dan pendapatan responden tersebut diberikan di bawah ini.



Terdapat tiga jalur (cara) dalam pembuatan grafik pada SPSS. Pada tulisan ini kita akan membahas cara yang paling sederhana saja, sebagai berikut:

Klik Graphs > Legacy Dialogs. Akan terdapat beberapa pilihan grafik yang tersedia yaitu: Bar (grafik batang), 3-D (grafik batang tiga dimensi), Line (grafik garis), Pie (grafik lingkaran), High-Low, Boxplot, Error Bar, Population Pyramid, Scatter/Dot (sebaran/titik), dan Histogram. Masing-masing grafik memiliki karakteristik-karakteristik tertentu yang sesuai dalam penggambaran data. Grafik yang berbasis batang, umumnya digunakan untuk menggambarkan perbandingan antar variabel/kategori. Grafik yang berbasis garis, umumnya (lebih sesuai) untuk menggambarkan perkembangan data. Grafik yang berbasis lingkaran, umumnya untuk menggambarkan data yang bersifat proporsi. Grafik yang berbasis titik umumnya untuk menggambarkan pencaran/sebaran data.

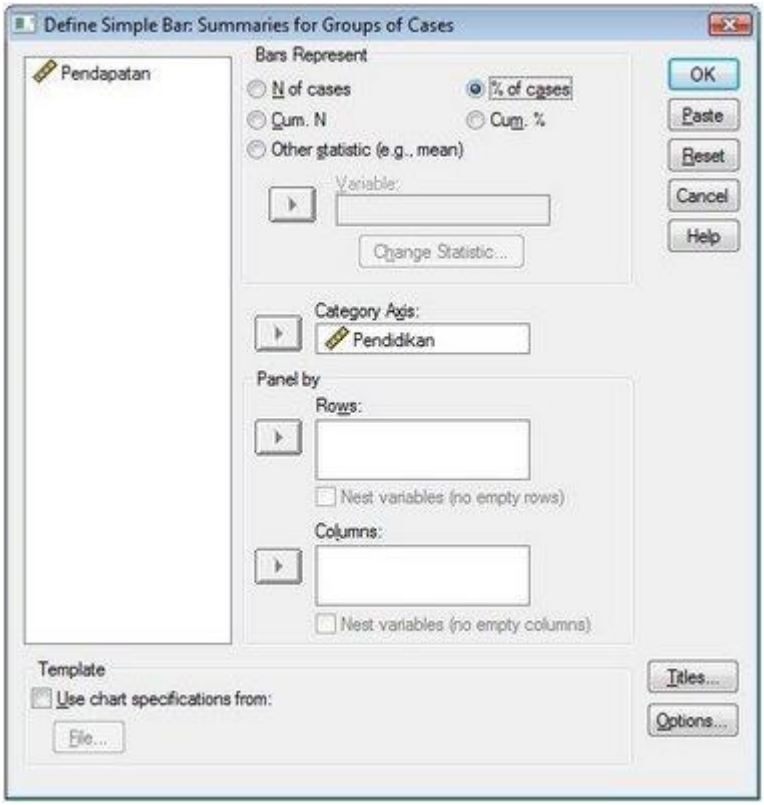
Dalam konteks data latihan kita, pada tulisan ini kita akan membahas terlebih dahulu mengenai grafik batang. Untuk itu klik Bar, akan muncul tampilan berikut:



Terdapat tiga pilihan grafik batang, yaitu Simple, Clustered dan Stacked. Pilihan Simple digunakan untuk menggambarkan grafik dari variabel tunggal. Pilihan Clustered dan Stacked digunakan untuk menggambarkan grafik dari variabel tunggal tetapi dikelompokkan berdasarkan kategori dari variabel lainnya. Pengelompokan pada tipe grafik Clustered dilakukan secara horizontal, sedangkan pada tipe Stacked secara vertikal. Kemudian terdapat pilihan tampilan data untuk

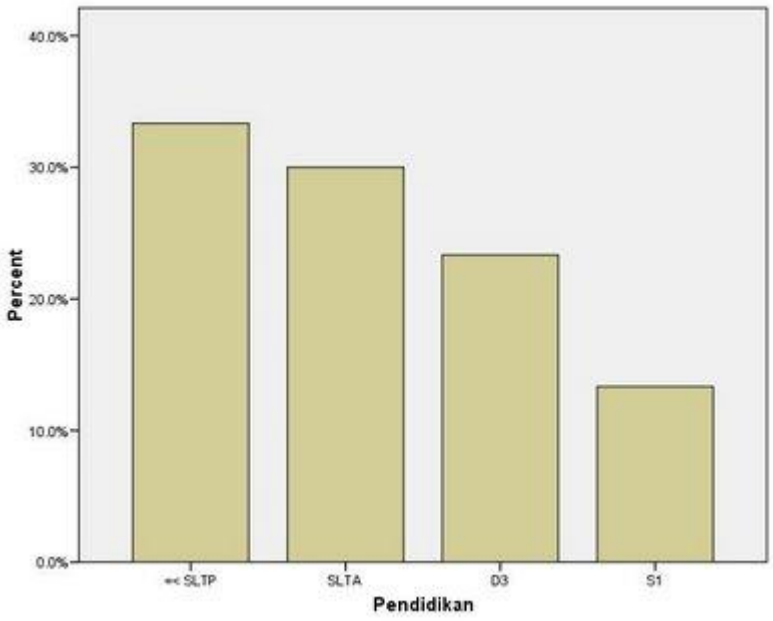
grafik (Data in Chart Are), yaitu diringkas berdasarkan kategori (Summaries for groups of cases), diringkas berdasarkan pemisahan variabel (Summaries of separate variables) atau menampilkan data individual. (Catatan: anda bisa mencoba-coba pilihan tersebut untuk memahami maknanya).

Sebagai latihan awal, kita akan membuat grafik untuk variabel jenis kelamin. Kita pilih jenis grafik Simple (klik) dan tampilan data adalah Summaries for groups of cases. Kemudian klik Define, akan muncul tampilan berikut:



Tentukan terlebih dahulu ukuran yang akan ditampilkan oleh batang dari grafik kita (Bar Represent). Ada beberapa pilihan yaitu ukuran frekuensi absolut (N of cases), kumulatif frekuensi (Cum.N), persentase frekuensi (% of cases), kumulatif persentase frekuensi (Cum.%), atau ukuran statistik lainnya (Other Statistics). Anda juga bisa mencoba-coba berbagai pilihan ini untuk melihat perbedaan output grafiknya.

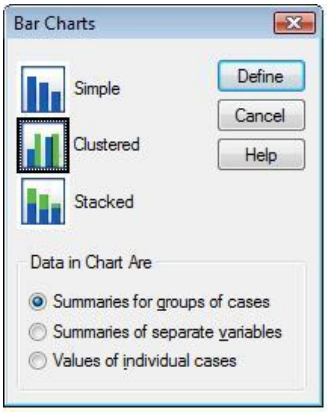
Untuk latihan ini, kita pilih % of cases. Selanjutnya masukkan variabel Pendidikan ke dalam kotak Category Axis, dan kemudian klik OK. Akan muncul output grafik sebagai berikut:



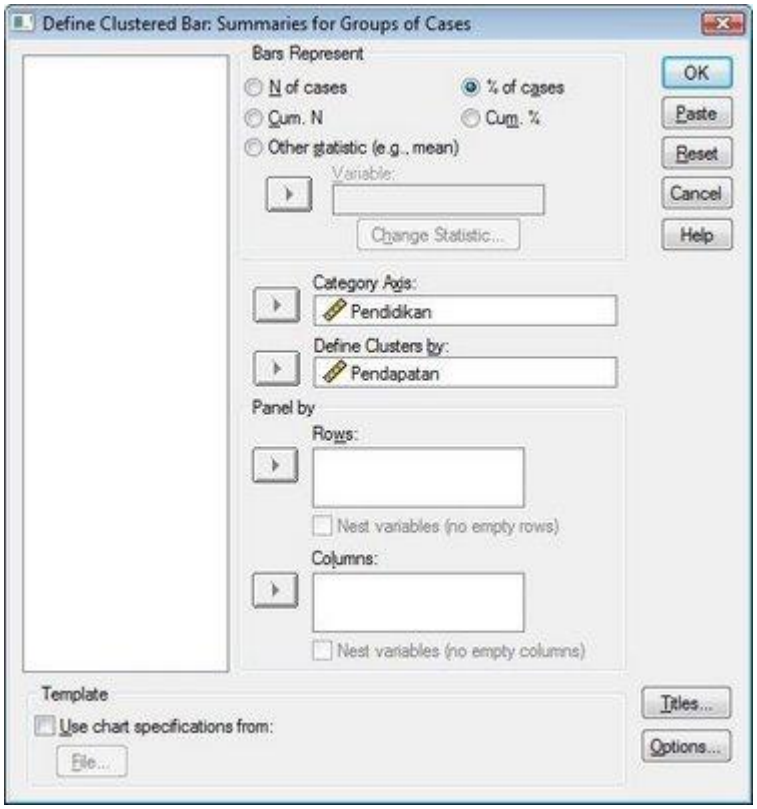
B. Grafik Batang Clustered dan Stacked

Kali ini akan dibahas mengenai grafik batang clustered dan stacked. Data yang digunakan untuk latihan adalah data tingkat pendidikan dan pendapatan yang telah diberikan sebelumnya.

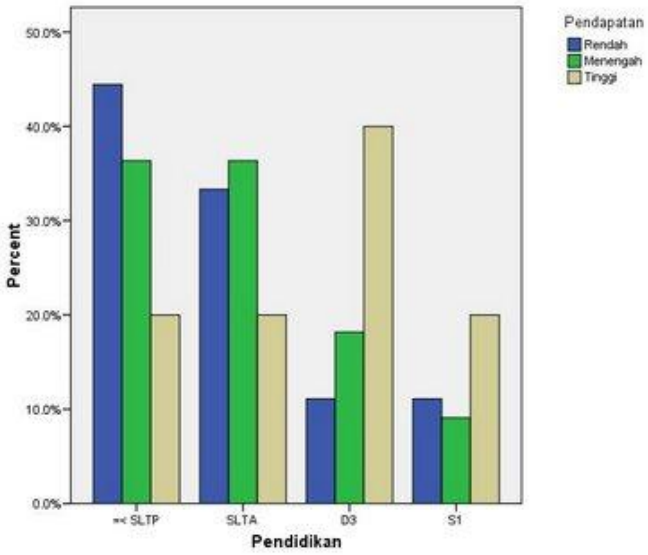
Setelah menginput data tersebut, klik Graphs > Legacy Dialogs > Bar. Akan muncul tampilan berikut:



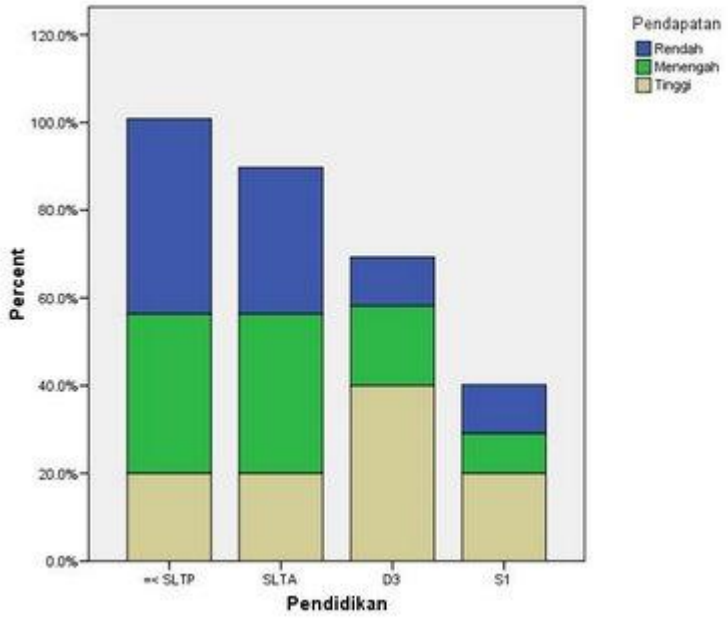
Sebagai latihan, kita akan membuat grafik pendidikan yang dikelompokkan berdasarkan tingkat pendapatan. Klik Clustered, klik Summaries for groups of cases, kemudian klik Define, akan muncul tampilan berikut:



Pada Bar Represent klik % of cases (tentunya anda bisa mencoba pilihan lainnya). Masukkan variabel Pendidikan pada kotak Category Axis dan variabel Pendapatan pada kotak Define Cluster by. Kemudian klik OK. Akan muncul output grafik sebagai berikut:



Ini adalah grafik batang clustered. Apa perbedaannya dengan grafik batang stacked? Jika anda ambil pilihan stacked pada proses sebelumnya, output grafik akan menjadi sebagai berikut:

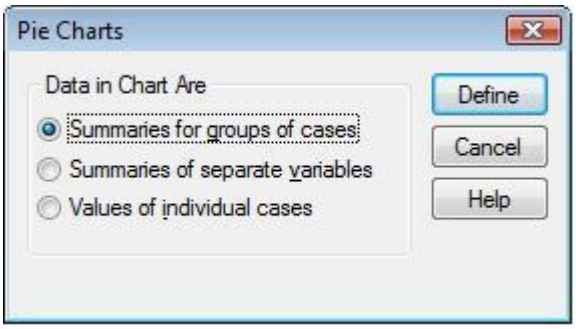


Terlihat bahwa perbedaan antara grafik batang clustered dengan stacked adalah jika pada clustered pengelompokkan pendapatan dilakukan secara horizontal (kesamping), pada stacked dilakukan secara bertumpuk vertikal.

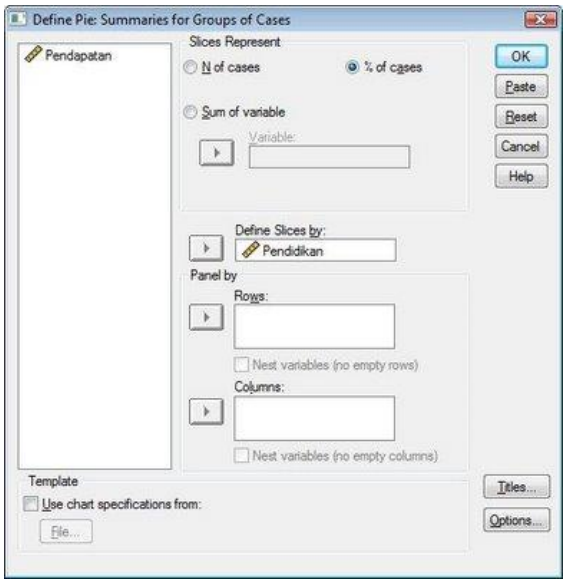
C. Grafik Lingkaran

Salah satu jenis grafik yang sering digunakan terutama untuk menggambarkan proporsi dari kategori data adalah grafik lingkaran (pie). Sebagai latihan, data yang kita gunakan sama dengan data sebelumnya (yaitu data tingkat pendidikan dan pendapatan dari 30 orang responden penelitian).

Setelah menginput data tersebut, klik Graphs > Legacy Dialogs > Pie. Akan muncul tampilan berikut:

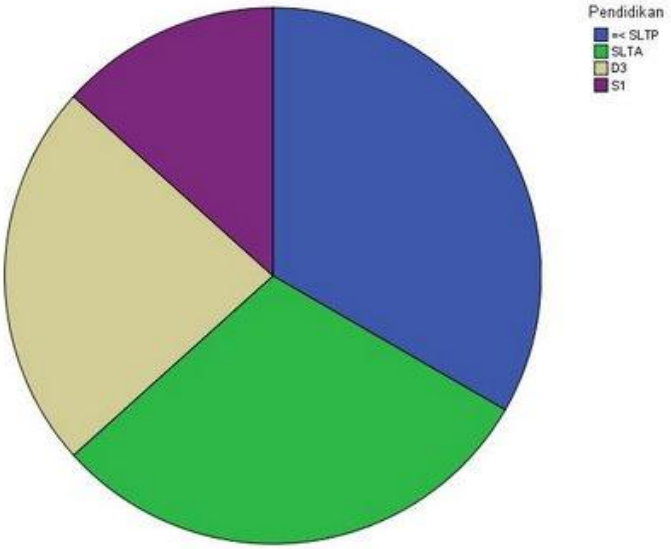


Dari tampilan tersebut, pilih Summaries for group of cases (lihat catatan pada tulisan sebelumnya untuk penggunaan pilihan yang lainnya). Kemudian klik Define. Akan muncul tampilan berikut:

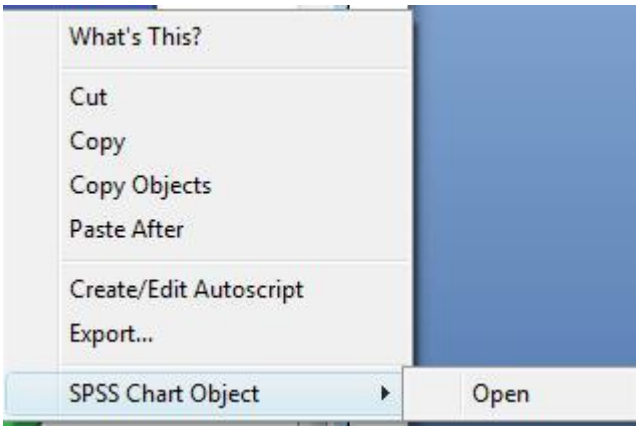


Misalnya kita ingin menggambarkan grafik lingkaran untuk pendidikan. Pada Slices Represent, kita diminta untuk memilih apakah irisan dari grafik lingkaran akan menggambarkan jumlah kasus (N of cases), persentase atau proporsi masing-masing

kategori (% of cases) atau jumlah dari variabel (Sum of variable). Untuk latihan kita pilih saja % of cases. Selanjutnya pada kotak Define Slices by: masukan variabel Pendidikan. Kemudian klik OK, maka akan muncul output grafik lingkaran sebagai berikut:

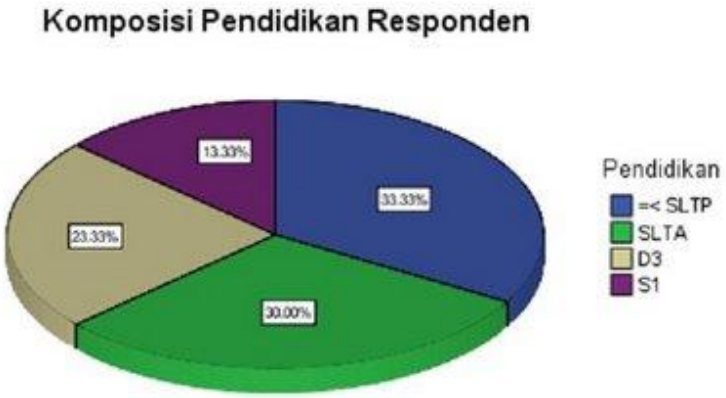


Perhatikan, grafik lingkaran yang dihasilkan masih sangat sederhana dengan elemen grafik yang sangat terbatas. Untuk menambah elemen grafik agar lebih komunikatif dalam penggambaran data, buka Chart Editor dengan cara klik kanan pada bidang grafik kemudian klik SPSS Chart Object dan klik Open (lihat tampilan dibawah ini)



Akan muncul tampilan Chart Editor. Melalui menu-menu yang ada pada Chart Editor tersebut, kita bisa mengatur ukuran dari lingkaran, membuat tampilan jadi tiga dimensi, memberikan label pada masing-masing irisan dari lingkaran, memberi judul, mengganti warna dan lainnya. Silakan dicoba-coba (tidak dibahas dalam postingan ini karena akan terlalu panjang).

Salah satu contoh setelah pengeditan (diberi judul, label irisan dan dijadikan tiga dimensi) diberikan sebagai berikut:



BAB VII
PENGUJIAN HIPOTESIS

Pengujian hipotesis statistik adalah prosedur yang memungkinkan keputusan dapat dibuat, yaitu keputusan untuk menolak atau atau tidak menolak hipotesis yang sedang dipersoalkan. Hipotesis yang akan diuji diberi symbol Ho (hipotesis nol) dan langsung disertai Ha (hipotesis alternatif). Ha akan secara otomatis diterima, apabila Ho ditolak.

A. Uji Hipotesis Rata-Rata Sampel Tunggal (One Sample t Test)

Uji rata-rata satu sampel atau sering di kenal sebagai uji one sample t test berguna untuk mengetahui perbedaan nilai rata-rata populasi yang digunakan sebagai pembading dengan rata-rata sebuah sampel. Dari hasil uji ini akan diketahui apakah rata-rata populasi yang digunakan sebagai pembanding berbeda nyata secara signifikan dengan rata-rata sebuah sampel, jika ada perbedaan rata-rata manakah yang lebih tinggi.

Rumusan hipotesis

- $H_o : \mu = \mu_o$, untuk menguji apakah rata-rata sample μ sama dengan rata-rata μ_o yang diberikan.
- $H_a : \mu > \mu_o$, untuk menguji apakah rata-rata sample μ lebih dari rata-rata μ_o yang diberikan.
- $H_a : \mu < \mu_o$, untuk menguji apakah rata-rata sample μ kurang dari rata-rata μ_o yang diberikan.

Kasus untuk sampel kecil

Sebuah asrama putri menyatakan bahwa menerima rata-rata 6 orang mahasiswa selama 2 tahun terakhir. Uji pernyataan tersebut pada taraf α = 0,05 jika diketahui data penerimaan mahasiswi selama 24 bulan bulan terakhir adalah sebagai berikut (nama variable **kost**):

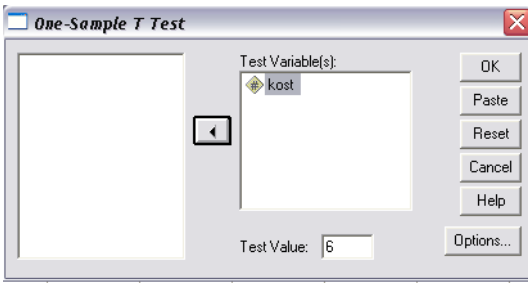
7	5	7	6	5	3	8	4
8	8	3	4	7	7	5	6
6	4	5	8	4	6	7	5

Penyelesaian:

Kasus diatas terdiri atas satu sample yang akan dipakai dengan nilai populasi hipotesis 6 orang. Di sini populasi dianggap berdistribusi normal dan karena sample < 30, maka dipakai uji t.

Langkah-langkah:

1. Klik menu **Analyze → Compare-Means → One Sample T test…**
   Maka tampak dilayar:



2. Masukkan variable **kost** ke kotak **Test Variable(s)**.
3. Untuk nilai yang akan diuji, isi angka 6 di kotak **Test Value**.
4. Klik tombol **Options…** untuk mengganti tingkat kepercayaan (Confidence Interval), gunakan tingkat kepercayaan 95%. Kemudian **OK.**

Akan muncul output:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
KOST	24	5.75	1.595	.326

One-Sample Test

	Test Value = 6					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
KOST	-.768	23	.450	-.25	-.92	.42

Penjelasan:

Hipotesis:

H_o : rata-rata penerimaan mahasiswa = 6

H_a : rata-rata penerimaan mahasiswa ≠ 6

Daerah penolakan:

Uji 2 arah: Tolak Ho bila $|t| > t_{\alpha/2,n-1}$ atau $P_{value} < \alpha$

Kesimpulan:

Dari keluaran di atas diperoleh nilai P = 0,45, sedangkan taraf nyata yang diuji adalah 0,05. Karena P = 0,45 > α = 0,05, maka terima H_o. Jadi dapat disimpulkan bahwa asrama putri menerima rata-rata 6 orang mahasiswi selama 2 tahun terakhir.

Kesimpulan yang sama juga akan diperoleh jika menggunakan statistik uji t. H_o ditolak bila $t > t_{\alpha/2,n-1}$ atau jika t bernilai negative bila $t < t_{\alpha/2,n-1}$. Untuk kasus ini, α = 0,05 dan n = 24 diperoleh nilai $t_{0.025,23}$ = 2,069. Karena t = -0,768 > $t_{0.025,23}$ = -2,069; maka terima H_o

Kasus untuk sampel besar

Seorang dosen mengatakan berat badan rata-rata mahsiswa di universitas 150 pon. Seorang mahasiswi ingin menguji kebenaran kata-kata dosennya itu dengan tingkat kepercayaan 95%. Dia mengambil sample acak 40 mahasiswa dengan berat sebagai berikut:

128	138	135	164	165	150	144	132	157	144
125	149	145	152	140	154	156	153	119	148
136	163	147	176	147	135	142	150	145	173
135	142	138	126	140	161	146	168	198	146

Penyelesaian:

Pada SPSS pengujian sample besar tetap menggunakan **One Sample T test…** sehingga tahapan pengerjaan untuk contoh ini digunakan langkah-langkah di atas.

Akan muncul output:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
BERAT	40	147.80	15.274	2.415

One-Sample Test

	Test Value = 150					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
BERAT	-.911	39	.368	-2.20	-7.08	2.68

Penjelasan:

Ho : Rata-rata berat badan mahasiswa = 150 pon

Ha : Rata-rata berat badan mahasiswa ≠ 150 pon

Daerah penolakan: Uji 2 arah, tolak Ho apabila $|z| > z_{\alpha/2}$ atau $P_{value} < \alpha$.

Kesimpulan:

Dari keluaran di atas diperoleh nilai P = 0,368, sedangkan taraf nyata α yang diuji adalah 0,05. Karena P > α maka terima Ho. Jadi pernyataan dosen yang menyatakan berat badan rata-rata mahasiswa 150 pon adalah benar.

Kesimpulan yang sama juga akan diperoleh jika menggunakan statistik uji z. Ho ditolak apabila $z > z_{\alpha/2}$ atau jika z bernilai negative $z < -z_{\alpha/2}$. Untuk kasus ini α = 0,05 maka diperoleh $z_{0,025}$ = -1,96. Karena $z > -z_{0,025}$ (-0,91 > -1,96) maka terima Ho.

Latihan.

1. Suatu daerah akan mengembangkan salah satu kecamatan yang terpencil. Salah satu alternative pengembangannya adalah dengan cara mengembangkan peternakan sapi perah oleh masyarakat daerah itu. Menurut perhitungan yang telah dilakukan sebelumnya, peternakan sapi perah hanya akan menguntungkan apabila produksi rata-ratanya mencapai 16 liter per hari. Kepala-kepala desa setempat mengatakan bahwa rata-rata produksi susu setiap ekor sapi per hari mencapai 16 liter. Untuk melaksanakan program pengembangan peternakan sapi perah itu tentu saja harus diuji apakah produksi susu rata-ratanya benar-benar 16 liter, dengan kata lainapakah pernyataan para kepala desa itu benar. Oleh karena itu diteliti sampel sebanyak 6 ekor sapi dengan kualitas sesuai dengan program

yang dimiliki oleh penduduk di daerah itu. Ternyata produksi susunya sebagai berikut: 14, 15, 19, 12, 17, dan 18 liter. Uji kebenaran hipotesis dengan $\alpha = 5\%$!

2. Suatu perusahaan penangkar benih menawarka suatu KUD untuk memasarkan benih yang dihasilkannya kepada para petani. Benih ini dijamin memiliki rata-rata daya kecambah 80% dengan ragam 9%. Karena menyangkut produktivitas padi sawah di arealnya, maka manajer KUD ini menguji kebenaran jaminan tersebut dengan melakukan percobaan dengan menanam benih masing-masing 1.000 butir benih yang diambil secara acak pada 8 petak sawah. Dua minggu setelah tanam, diperoleh persen daya kecambah pada masing-masing petak 85, 86, 87, 84, 91, dan 75%, $\alpha = 5\%$. Menurut saudara, maukah manajer tersebut memasarkan benih ini?

B. Uji Hipotesis Beda Rata-rata Dua Sampel Berpasangan (Paired Sample t Test)

Dua sampel berpasangan artinya sampel dengan subjek yang sama namun mengalami dua perlakuan atau pengukuran yang berbeda. Uji perbedaan rata-rata dua sampel berpasangan atau uji paired sample t test digunakan untuk menguji ada tidaknya perbedaan mean untuk dua sampel bebas (independen) yang berpasangan. Adapun yang dimaksud berpasangan adalah data pada sampel kedua merupakan perubahan atau perbedaan dari data sampel pertama atau dengan kata lain sebuah sampel dengan subjek sama mengalami dua perlakuan.

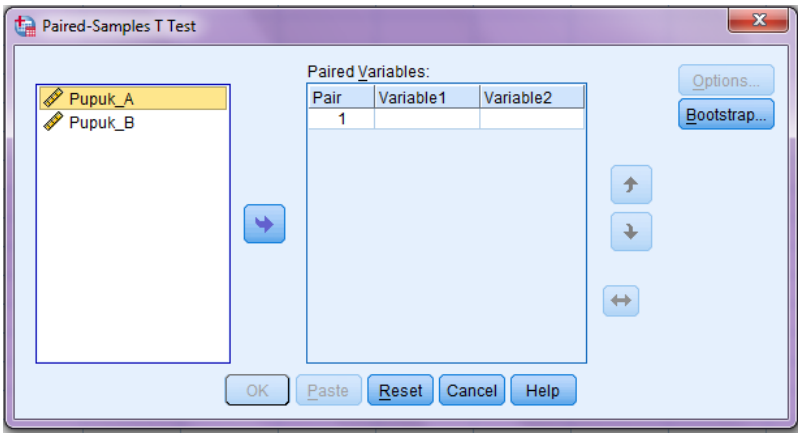
Contoh kasus.

Pengujian produktivitas padi (kwintal) yang diberi dua jenis pupuk.

Plot	Pupuk A	Pupuk B
1	7	8
2	6	6
3	5	7
4	6	8
5	5	6
6	4	6
7	4	7
8	6	7
9	6	8
10	7	7
11	6	6
12	5	7

Langkah-langkah pengujian:

1. Masukkan data seperti data di atas sesuai dengan variabelnya.
2. Klik menu **Analyze – Compare Means – Paired-Samples T Test**.
3. Klik Pupuk_A dan drag ke kotak Paired Variables pada kolom Variable 1 dan Pupuk_B ke Variable 2.



4. Untuk **Options**, gunakan tingkat kepercayaan 95% atau tingkat signifikansi 5%, klik **Continue.**
5. Untuk mengakhiri klik **OK**, maka akan ditampilkan outputnya sebagai berikut:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pupuk_A	5.58	12	.996	.288
	Pupuk_B	6.92	12	.793	.229

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pupuk_A & Pupuk_B	12	.412	.183

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pupuk_A - Pupuk_B	-1.333	.985	.284	-1.959	-.708	-4.690	11	.001

Penjelasan output:

- Tabel pertama menjelaskan deskriptif dari hasil panen padi untuk masing-masing pupuk. Rata-rata panen padi yang diberi pupuk A 5,58 kwintal dengan standar deviasi 0,996 kwintal dan standard error 0,228 kwintal. Sedang padi yang diberi pupuk jenis B memiliki rata-rata lebih besar sedangkan standar deviasi dan standar error lebih kecil yaitu 0.793 dan 0,229 kwintal.
- Tabel kedua menyajikan korelasi antara kedua variabel, yang menghasilkan angka 0,412 dengan nilai probabilitas (sig.) 0,183. Hal ini menyatakan bahwa korelasi antara sebelum diet dan sesudah diet mempunyai hubunganyang sedang bahkan cenderung lemah, karena nilai probabilitas >0,05.
- Tabel Ketiga (Paired Samples Test)

  Hipotesis :

  *Ho = Kedua rata-rata populasi sama (rata-rata hasil panen padi yang diberi pupuk A dan pupuk B adalah sama atau tidak berbeda secara nyata)

  *H1 = Kedua rata-rata populasi tidak sama (rata-rata hasil panen padi yang diberi pupuk A dan pupuk B adalah tidak sama atau berbeda secara nyata)

  Berdasarkan perbandingan nilai probabilitas (Sig.)

  *Jika probabilitas > 0,05, maka Ho diterima

  *Jika probabilitas < 0,05, maka Ho ditolak

  Keputusan :

  Terlihat bahwa nilai probabilitas 0,001. Oleh karena probabilitas 0,001 < 0,05, maka Ho ditolak, yang berarti rata-rata kedua populasi tidak sama atau berbeda nyata. Dalam output juga disertakan berbedaan mean sebesar 1,333 kwintal yaitu selisih rata-rata hasil panen kedua padi.

Sebagai latihan bagi para mahasiswa, lakukan pengambilan keputusan menggunakan t hitung dan t table.

Latihan.

Ujilah apakah terdapat perbedaan penghasilan para petambak ikan pada musim penghujan dan kemarau. Diambil sampel sebanyak 5 petambak yang memiliki luas kolam yang sama. Penghasilan dalam ton adalah sebagai berikut:

No	1	2	3	4	5
Musim Hujan	15	14	11	21	10
Musim Kemarau	13	10	9	25	9

C. Uji Hipotesis Beda Rata-Rata Dua Sampel Saling Bebas (Independent Sample t Test)

Uji Independent Sample T Test digunakan untuk membandingkan rata-rata dari dua group yang tidak berhubungan satu dengan yang lain, apakah kedua group tersebut mempuyai rata-rata yang sama ataukah tidak secara signifikan. Data kuantitatif dengan asumsi data berdistribusi normal dan jumlah data sedikit yakni di dibawah 30.

Contoh kasus.

Ada anggapan bahwa ada perbedaan IP antara mahasiswa Prodi A dengan Prodi B. Dari sampel acak mahasiswa yang dipilih, diperoleh data IP sebagai berikut:

Prodi A	2,11	3,15	2,75	3,10	2,95	2,95	3,00	2,50	2,79	2,50
Prodi B	3,05	2,70	2,90	2,67	3,15	2,03	2,65	2,37		

Langkah-langkah pengujian:

1. Pada kolom Name, ketik IP pada baris pertama dan Prodi pada baris kedua.
2. Pada kolom Label, ketik Indeks Prestasi untuk baris pertama dan Prodi untuk baris kedua.
3. Pada baris kedua, pada kolom Values, klik mouse pada kotak kecil di kanan sel. Pada kotak isikan Value ketik 1, pada kotak isian Value label, ketik Prodi A, Klik tombol Add, selanjutnya isi kembali untuk value, ketik 2 dan pada Velue label ketik Prodi B, klik kembali tombol Add, kerena sudah selesai maka klik **OK**.
4. Klik Data View, pada SPSS Data Editor dan masukkan datanya seperti data di atas sesuai dengan variabelnya.
5. Setelah itu, klik menu **Analyze – Compare Means – Independent – Samples T test.**
6. Muncul kotak dialog baru, pada kotak tersebut klik variable Indeks Prestasi, masukkan ke kotak Test Variables.
7. Pada Grouping Variable, klik Define Groups ketik 1 pada Group 1 dan ketik 2 pada Group 2, kemudian klik **Continue.**
8. Untuk **Option**, gunakan tingkat kepercayaan 95% atau tingkat signifikansi 5%, klik **Continue**.
9. Untuk mengakhiri perintah Klik **OK**. Maka akan muncul output SPSS

Group Statistics

	Prodi	N	Mean	Std. Deviation	Std. Error Mean
IP	Prodi A	10	2.7300	.36624	.11582
	Prodi B	8	2.6900	.36359	.12855

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
IP	Equal variances assumed	.311	.585	.231	16	.820	.04000	.17318	-.32712	.40712
	Equal variances not assumed			.231	15.191	.820	.04000	.17303	-.32839	.40839

Penjelasan output.

- Output Bagian Pertama (Group Statistics)

  Pada bagian pertama ini menyajikan deskripsi variabel yang dianalisis, yang meliputi rata-rata (mean) Indek Prestasi Prodi A = 2,7800 dengan standar deviasi 0,3253 dan rata-rata Indek Prestasi Prodi B = 2,6900 dengan Standar deviasi 0,3636.

- Output Bagian Kedua (Independent Sample Test)

  Analisis Uji F

  Hipotesis :

  H0 = Kedua varians populasi adalah sama (homogen)

  H1 = Kedua varians populasi adalah tidak sama (tidak homogen)

  Pengambilan Keputusan :

  • Jika nilai probabilitas > 0,05, maka H0 diterima

  • Jika nilai probabilitas < 0,05, maka H0 ditolak

Keputusan :

Terlihat bahwa Fhitung untuk Indek Prestasi adalah 0,006 dengan probabilitas 0,937. Oleh karena nilai probabilitas > 0,05, maka H0 diterima atau kedua varians populasi adalah sama (homogen)

Analisis Uji t

Hipotesis :

Ho = rata-rata Indek Prestasi antara Prodi A dan Prodi B adalah sama

H1 = rata-rata Indek Prestasi antara Prodi A dan Prodi B adalah tidak sama

Pengambilan keputusan dalan analisis Uji t dapat dilakukan dengan dua cara yakni berdasarkan perbandingan antara thitung dengan t tabel, dan berdasarkan perbandingan nilai probabilitas atau nilai signifikansi.

• Jika probabilitas > 0,05, maka H0 diterima

• Jika probabilitas < 0,05, maka H0 ditolak

Keputusan :

Terlihat bahwa thitung adalah dengan probabilitas 0,587. Oleh karena probabilitas 0,587 > 0,05, maka Ho diterima, maka dapat disimpulkan bahwa rata-rata indek prestasi antara Prodi A dan Prodi B adalah sama.

Latihan

Terdapat dua tanaman utama yang diusahakan oleh petani, yaitu kopi dan sayuran. Seorang peneliti tertarik untuk mengamati perbedaan pendapatan kedua jenis petani ini. Hasil survey terhadap masing-masing 10 petani kopi dan 10 petani sayuran yang terpilih secara acak ada pada table ini.

Petani	1	2	3	4	5	6	7	8	9	10
Kopi	2,5	2,7	3,7	7,5	7	6	5,2	5,5	8,1	9,1
Sayuran	1,5	3,1	2,7	1,5	1,4	1,7	6,1	3,5	1,9	2,7

Dengan $\alpha = 5\%$, ujilah perbedaan pendapatan kedua kelompok petani tersebut.

Daftar Pustaka

Sumber buku.

Kemas Ali hanafiah. 2006. *Dasar-Dasar Statistika*. Jakarta: PT Raja Grafindo Persada

Pangestu Subagyo. 2004. *Statistika Terapan*. Yogyakarta: BPFE-Yogyakarta.

Sugiyono. 2010. *Statistika untuk Penelitian*. Bandung: Alfabeta.

Sumber Internet

http://tip-trik-spss.blogspot.com

http://junaidichaniago.wordpress.com/2010/02/10/sekilas-mengenai-spss/

http://konsisten.com

http://statistikceria.com