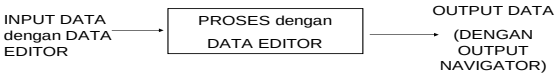
SEKILAS TENTANG SPSS

Salah satu program yang dapat digunakan untuk mengolah data melalui komputer adalah SPSS. Program ini telah dipakai pada berbagai macam industri untuk menyelesaikan berbagai macam permasalahan seperti, riset perilaku konsumen, peramalan bisnis dsb. Selain itu SPSS juga telah banyak dipergunakan di dunia pendidikan untuk membantu para akademisi dan mahasiswa dalam melakukan penelitian ilmiah.

1. KOMPUTER

2. STATISTIK

3 SPSS



Adapun cara memproses data melalui SPSS adalah sebagai berikut:

1. Data yang akan diproses dimasukkan dalam menu DATA EDITOR yang secara otomatis muncul dilayar
2. Data yang telah diinput kemudian diproses pada menu DATA EDITOR
3. Hasil pengolahan data muncul di layar (window) lain yang disebut output navigator. Hasilnya berupa:
    - Teks atau tulisan
    - Tabel
    - Bagan atau Grafik

Sebelum menentukan metoda pengolahan data yang sesuai dengan tujuan penelitian, maka para peneliti juga harus memahami karakteristik dari data yang ada sehingga metoda statistik yang digunakan tepat.

Statistik Parametrik

Metoda statistik Parametrik akan tepat digunakan pada kondisi :

- Tipe Data Interval/Rasio
- Data berdistribusi normal

- Jumlah data >30

Catatan : apabila jumlah data < 30 akan tetapi distribusi populasi normal, maka dapat menggunakan t test

Statistik Non Parametrik

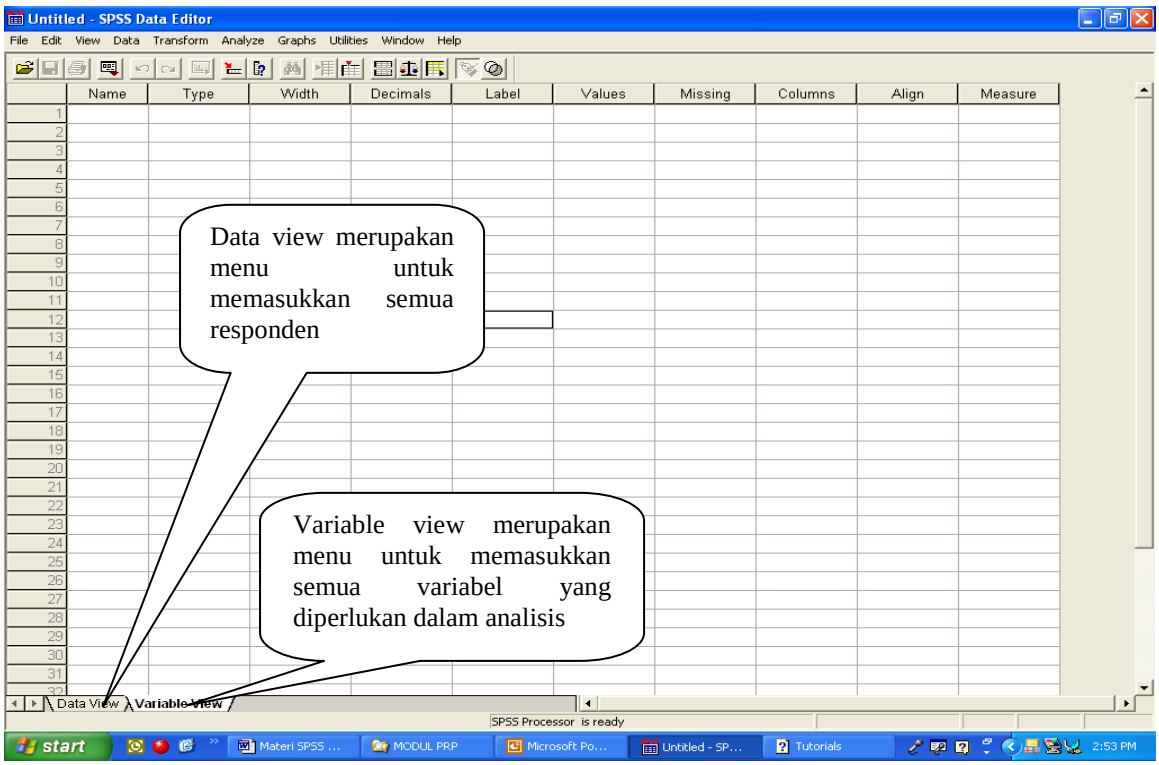
Metoda statistik Non Parametrik akan tepat digunakan pada kondisi :

- Tipe Data Nominal/Ordinal
- Data tidak berdistribusi normal
- Jumlah data < 30

PERSIAPAN DATA

A. SPSS DATA EDITOR

Data editor menyediakan suatu metode yang mirip seperti lembaran buku kerja untuk membuat dan mengedit data dalam file. Data editor terbuka secara otomatis saat Anda membuka program SPSS.



Data editor menyediakan dua menu untuk melihat data Anda:

1. Data view : menunjukkan nilai dari data aktual atau *defined value labels*
2. Variable view: memperlihatkan informasi mengenai semua variabel termasuk *name, type, width, decimals, label, values, missing, column, align, measure*.

Dalam kedua menu tersebut, Anda dapat menambah, mengubah dan menghapus data yang berada dalam file SPSS.

B. MENU VARIABLE VIEW

Dalam variabel view terdapat 10 kolom yang terdiri dari:

1. Name

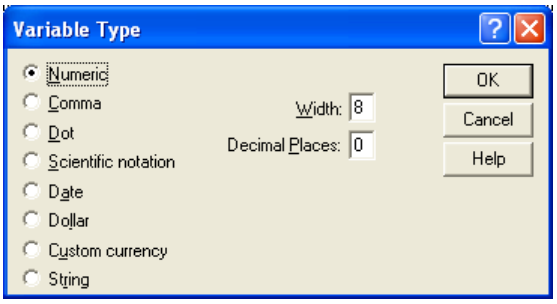
Bagian ini untuk mengisi nama variabel yang terdapat dalam penelitian. SPSS selalu menuliskan nama variabel dalam huruf kecil dengan jumlah karakter maksimal 8 huruf.

2. Type

Kolom ini digunakan untuk menentukan tipe data dari variabel yang akan dimasukkan ke dalam program SPSS.

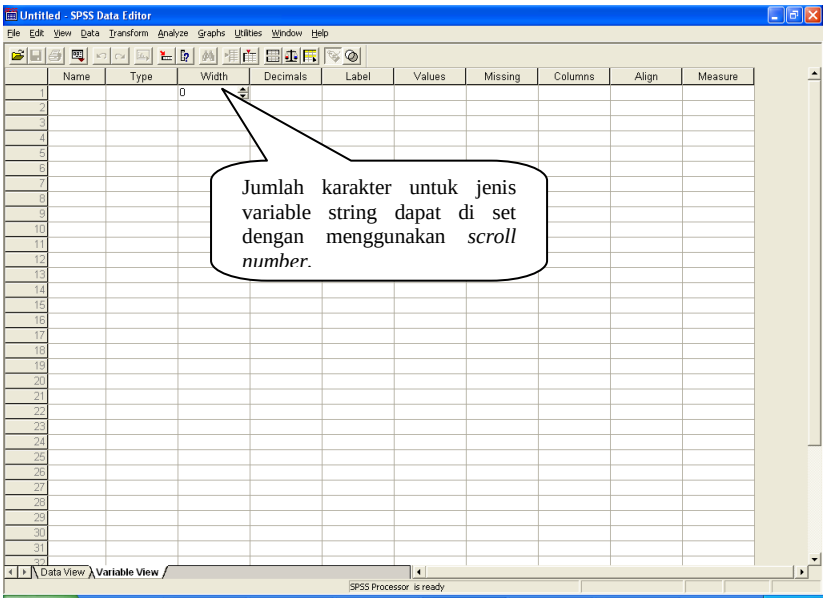
Pada bagian ini Anda memiliki 8 pilihan jenis variabel. Namun jenis variabel yang paling sering digunakan adalah numeric, date dan string,

Numeric digunakan untuk data yang berupa angka. Date digunakan untuk data tanggal, biasanya dipergunakan dalam analisis *time series.* Sedangkan *string* digunakan untuk data yang berupa karakter.



3. Width

Pilihan ini menyediakan masukan antara 1 hingga 255 karakter untuk jenis variabel *string*. Pengisian angka WIDTH dapat dilakukan dengan mengetik jumlah karakter kyang diinginkan atau dengan menggunakan *scroll number* yang terletak di sebelah kanan kolom WIDTH.



4. Decimals

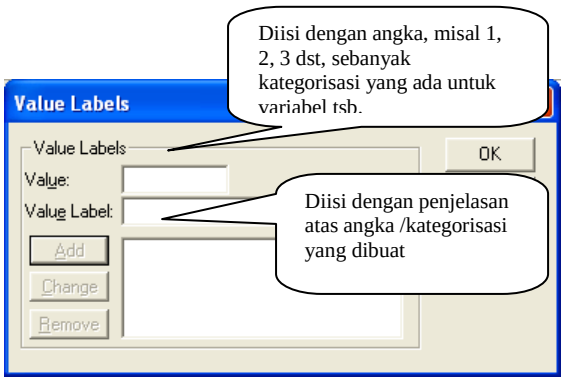
Desimal akan muncul untuk jenis variable numeric.

5. Label

Label adalah keterangan yang lebih lengkap untuk nama variabel. Jika sebelumnya Anda hanya dibatasi untuk memberikan nama variabel sejumlah 8 karakter pada kolom **name**, maka pada bagian ini Anda dapat menuliskan nama variabel selengkap mungkin. Nama pada label ini juga yang akan muncul pada output SPSS.

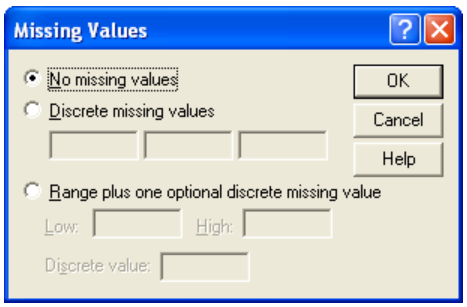
6. Value Labels

Kolom value diisi untuk variabel yang memiliki kategorisasi tertentu yang mewakili data non numeric (misal 1= pria, 2 =wanita). Hal ini dapat lebih memudahkan pengisian data tanpa perlu melakukan pengetikan yang berulang-ulang. Jika anda meng-klik di bagian kanan kolom tersebut, maka akan muncul tampilan berikut ini:



7. Missing

Kolom ini dipergunakan untuk menjelaskan mengenai perlakuan terhadap data yang hilang.



8. Column

Kolom ini memberikan informasi mengenai lebar kolom yang diperlukan untuk memasukkan data.

9. Align

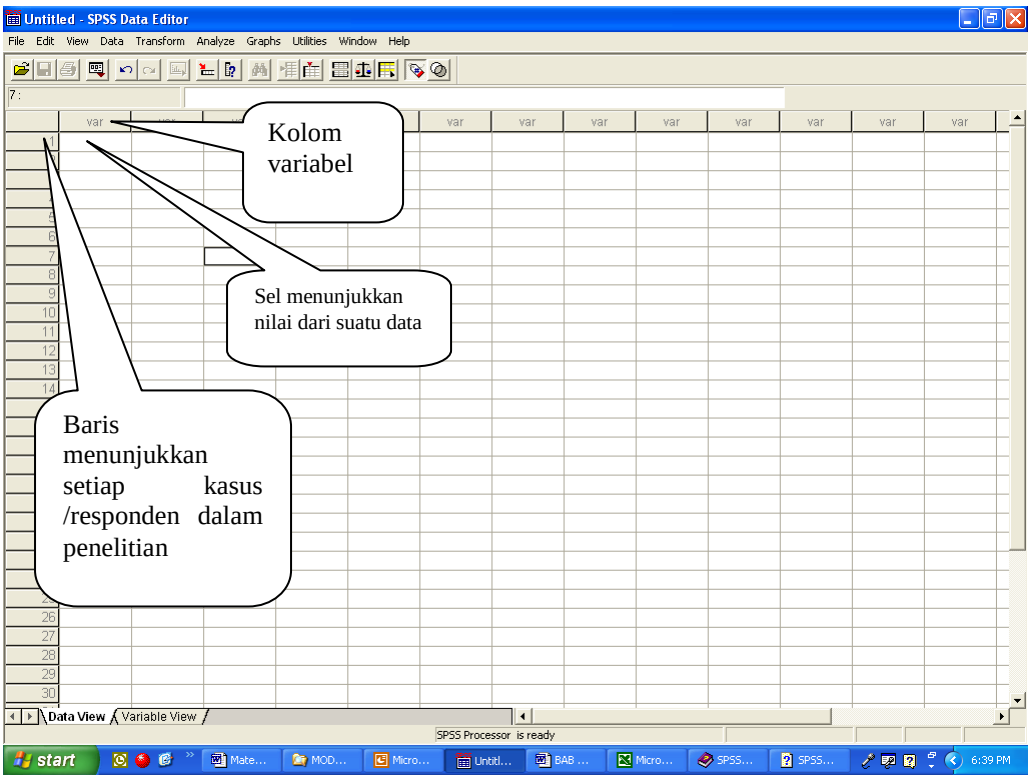
Kolom ini menunjukkan mengenai letak atau posisi data, apakah ingin diletakkan di sebelah kiri, kanan, atau tengah sel.

10. Measure

Kolom ini merupakan kolom yang sangat penting dalam SPSS karena pemilihan jenis pengukuran dalam kolom ini akan menentukan jenis analisis yang dipilih.

C. MENU DATA VIEW

Sebagian besar fitur dalam Data View hampir serupa dengan fitur dalam lembar kerja (spreadsheet). Namun, dalam SPSS kita tetap menemukan beberapa macam aplikasi yang berbeda:



1. **Baris**

Baris dalam menu ini menunjukkan jumlah kasus yang ada. Setiap baris dalam data view mewakili setiap kasus atau hasil observasi yang ada. Setiap responden dalam penelitian mewakili satu kasus.

2. Kolom

Kolom dalam menu Data View mewakili variabel atau karakteristik yang diukur dalam penelitian. Setiap item dalam kuesioner merupakan variabel.

3. Sel

Sel dalam dalam data view berisi nilai dari setiap variabel untuk setiap kasus. Sel merupakan perpotongan dari kasus dan variabel. Sel hanya berisi nilai dari data dan tidak berisi rumus-rumus seperti halnya dalam lembar kerja yang terdapat dalam program Excel.

Contoh cara memasukkan data ke SPSS.

Seorang peneliti yang melakukan penelitian mengenai

Nomor Responden	Nama	Usia	Jenis Kelamin	Pekerjaan	Pendapatan (dalam Juta)	Tingkat Pendidikan
1	Diah	38	Wanita	Pegawai Swasta	1	SMU/Sederajat
2	Andi	27	Pria	Wiraswasta	5	S1
3	Febi	37	Wanita	PNS	2	S1
4	Anita	39	Wanita	PNS	2	S1
5	Rina	43	Wanita	Pegawai Swasta	6	S2
6	Ahmad	49	Pria	Pegawai Swasta	6	S2
7	Rani	40	Wanita	Wiraswasta	6	SMU/Sederajat
8	Heri	42	Pria	PNS	2.5	S3
9	Iwan	29	Pria	Pegawai Swasta	1.2	Diploma
10	Dedi	26	Pria	PNS	1.5	Diploma

Dari data tersebut kita melihat bahwa terdapat kurang lebih 7 variabel untuk dimasukkan ke dalam program SPSS, yakni : Nomor responden, nama, usia, jenis kelamin, pekerjaan, pendapatan (dalam juta rupiah) dan tingkat pendidikan.

Buka menu **SPSS data editor** pada bagian **variable view** lalu masukkan data diatas dengan langkah sebagai berikut:

1. Nomor Responden

- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik nomor (bagian ini hanya dapat diisi oleh maksimal 8 karakter, tanpa mengizinkan adanya spasi)

- **Type.** Tipe data untuk nomor responden adalah **string** (non-numeric). Meskipun nomor responden seolah-olah terlihat dalam bentuk angka (numeric), namun sebetulnya angka tersebut sejenis dengan nomor-nomor lainnya yang hanya menunjukkan identitas, seperti halnya nomor KTP, NPM, dsb, yang tidak akan diolah atau dianalisis.
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi
- **Label.** Untuk keseragaman ketik Nomor responden
- **Values.** Karena nomor responden bukan merupakan data kuantitatif dan tidak memiliki kategorisasi, maka kita dapat mengabaikan pilihan ini.
- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.
- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.
- **Measure.** Untuk data kualitatif atau string, SPSS hanya menyediakan 2 pilihan yakni nominal atau ordinal. Untuk variabel ini kita memilih **nominal**. karena nomor responden tidak menunjukkan adanya perbedaan jenjang pada responden.

2. Nama

- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik nama
- **Type.** Tipe data untuk nama responden adalah **string** (non-numeric).
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi
- **Label.** Untuk keseragaman ketik Nama respoden
- **Values.** Karena nama responden bukan merupakan data kuantitatif dan tidak memiliki kategorisasi, maka kita dapat mengabaikan pilihan ini.
- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.
- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.
- **Measure.** Untuk data kualitatif atau string, SPSS hanya menyediakan 2 pilihan yakni nominal atau ordinal. Untuk variabel ini kita memilih **nominal**.

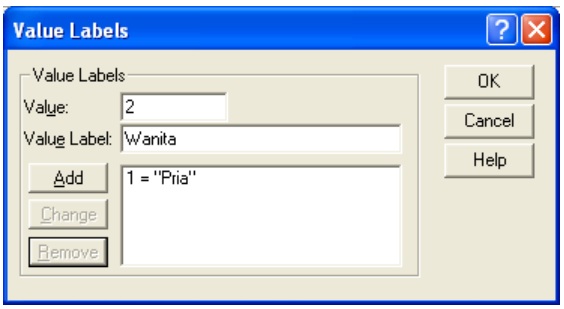
3. Usia

- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik usia

- **Type.** Tipe data untuk usia responden adalah **numeric**.
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi
- **Label.** Untuk keseragaman ketik Usia respoden
- **Values.** Usia responden merupakan data kuantitatif yang tidak perlu kita kategorisasi, maka pilihan tersebut dapat diabaikan.
- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.
- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.
- **Measure.** Untuk data kuantitatif atau numeric, SPSS menyediakan 3 pilihan yakni nominal, ordinal dan scale. Untuk variabel ini kita memilih **scale** karena datanya bersifat rasio. Yang mana SPSS menggabungkan data interval dan rasio ke dalam satu jenis yakni scale.

4. Jenis Kelamin

- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik jenis kelamin
- **Type.** Tipe data untuk jenis kelamin responden adalah **string** (non-numeric). Namun untuk mempermudah pengisian data, kita dapat membuat kategorisasi pria dan wanita ke dalam angka. Untuk itu kita pilih numeric.
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi
- **Label.** Untuk keseragaman ketik jenis kelamin respoden
- **Values.** Karena jenis kelamin responden akan kita kategorisasi, maka kita dapat mengklik pilihan ini, lalu akan muncul tampilan berikut ini
  Untuk kesepakatan, beri kategorisasi 1= Pria dan 2= wanita

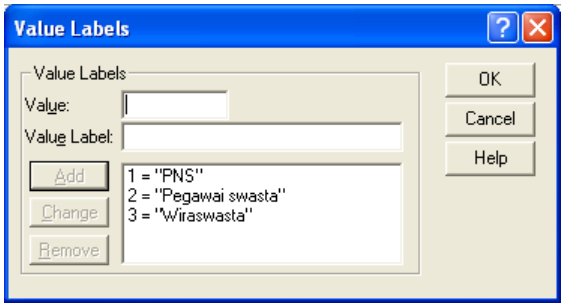


- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.
- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.

- **Measure.** Untuk data string dengan value, SPSS akan menyediakan 3 pilihan yakni nominal, ordinal dan scale. Untuk variabel ini kita memilih **nominal** karena angka yang kita berikan hanya dimaksudkan untuk kategorisasi.

5. Pekerjaan

- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik kerja.
- **Type.** Tipe data untuk pekerjaaan responden adalah **numeric** karena kita akan melakukan kategorisasi terhadap pekerjaan responden.
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi
- **Label.** Untuk keseragaman ketik pekerjaan responden.
- **Values.** Meskipun pekerjaan responden bukan merupakan data kuantitatif namun ia memiliki kategorisasi, maka kita akan melakukan langkah yang sama seperti yang dilakukan pada variabel sebelumnya. Untuk kesepakatan kita akan menggolongkan pekerjaan responden ke dalam tiga golongan.
  1= PNS, 2= Pegawai swasta, 3= Wiraswasta



- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.
- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.
- **Measure.** Untuk data kualitatif yang dikategorisasi menjadi numerik, SPSS menyediakan 3 pilihan yakni nominal, ordinal dan scale. Untuk variabel ini kita memilih **nominal**.

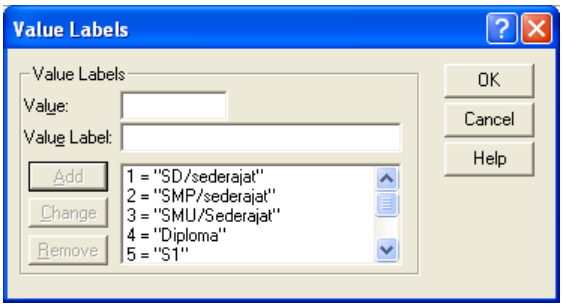
6. Pendapatan (dalam juta rupiah)

- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik income.
- **Type.** Tipe data untuk income adalah numeric.
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi.

- **Label.** Untuk keseragaman ketik Pendapatan respoden.
- **Values.** Pendapatan responden merupakan data kuantitatif yang tidak memiliki kategorisasi, maka kita dapat mengabaikan pilihan ini.
- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.
- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.
- **Measure.** Untuk data kuantitatif atau numeric, SPSS menyediakan 3 pilihan yakni nominal, ordinal dan scale. Untuk variabel ini kita memilih scale karena datanya bersifat **rasio**. Yang mana SPSS menggabungkan data interval dan rasio ke dalam satu jenis yakni scale.

7. Tingkat pendidikan

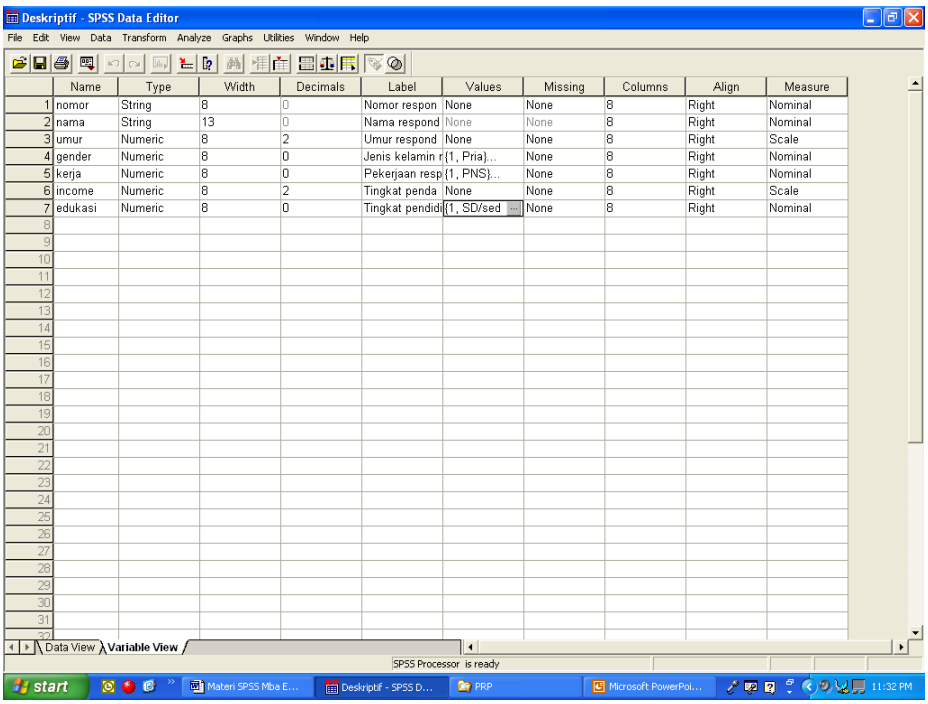
- **Name.** Sesuai kasus letakkan pointer di bawah kolom Name, *klik ganda* pada sel tersebut lalu ketik edukasi.
- **Type.** Tipe data untuk pendidikan responden adalah **numeric** karena kita akan melakukan kategorisasi terhadap pendidikan responden.
- **Width.** Biarkan saja sesuai dengan default SPSS.
- **Decimals.** Tidak usah diisi
- **Label.** Untuk keseragaman ketik pendidikan responden.
- **Values.** Meskipun pendidikan responden bukan merupakan data kuantitatif namun ia memiliki kategorisasi, maka kita akan melakukan langkah yang sama seperti yang dilakukan pada variabel sebelumnya. Untuk kesepakatan kita akan menggolongkan pendidikan responden ke dalam tujuh golongan 1= SD/sederajat, 2= SMP/sederajat, 3= SMU/sederajat, 4=Diploma, 5=S1, 6=S2, 7=S3.



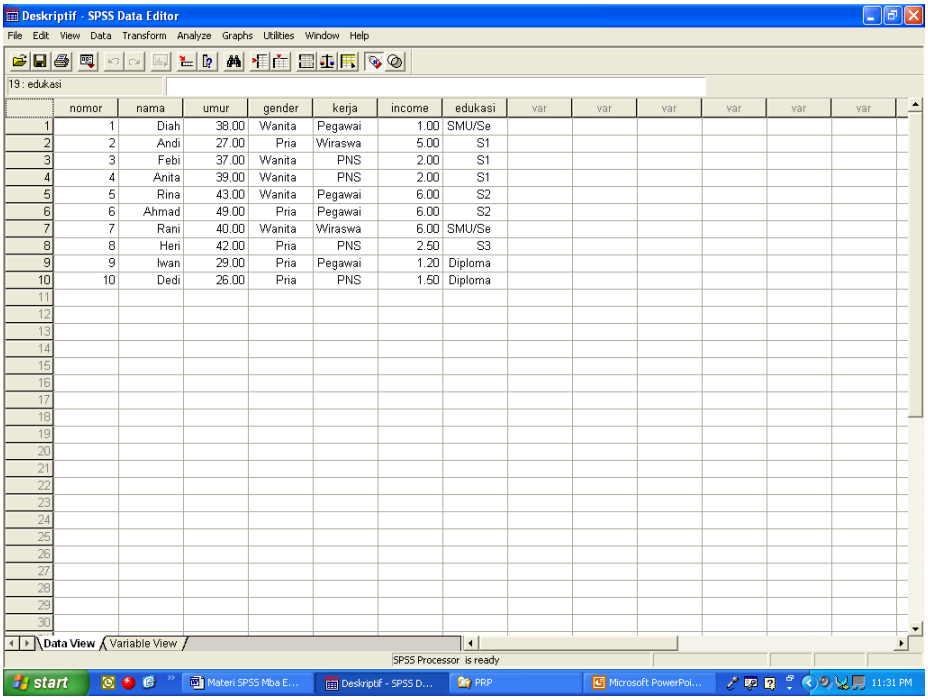
- **Missings.** Dianggap tidak ada data yang hilang.
- **Column.** Abaikan pilihan ini sehingga kita akan menggunakan default yang ada.

- **Align.** Abaikan pilihan ini sehingga akan menggunakan default dari SPSS yakni kanan.
- **Measure.** Untuk data kualitatif atau string yang dikategorisasi kedalam angka, SPSS menyediakan 3 pilihan yakni nominal, ordinal dan rasio. Untuk variabel ini kita memilih **nomina**l.

Setelah Anda memasukkan semua data tersebut, maka pada menu variable view akan tampak tampilan berikut ini:



Setelah selesai, masukkan data untuk masing-masing responden secara berurutan. Sehingga tampak tampilan berikut ini



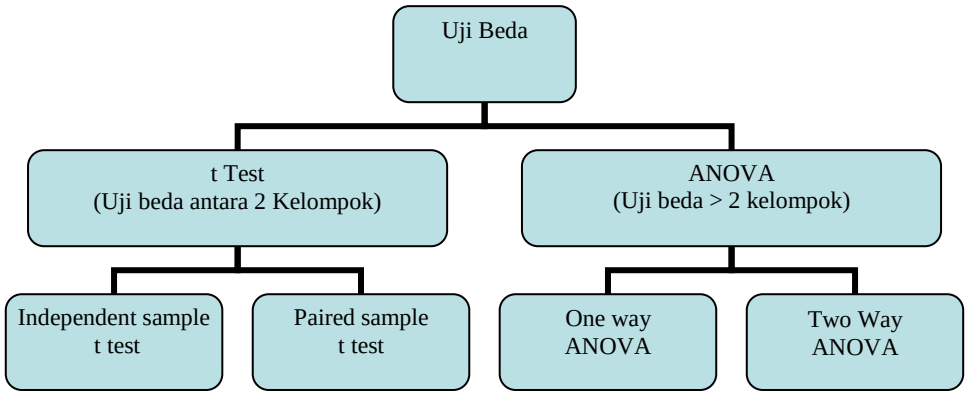
Setelah semua data dimasukkan, simpanlah file tersebut dengan memilih menu utama **FILE**, lalu pilih **Save As...** kemudian simpan file tersebut sebagai **DESKRIPTIF UNTUK KESERAGAMAN.**

PENGOLAHAN DATA

UJI BEDA : t Test dan ANOVA

Uji beda pada dasarnya dapat dilakukan melalui berbagai teknik. Dua diantara teknik analisis untuk uji beda adalah t test dan ANOVA. Uji beda pada prinsipnya dapat dilakukan pada minimal dua kelompok yang mendapatkan perlakuan yang berbeda.

Adapun perbedaan t test dilakukan jika uji beda hanya terdiri dari 2 kelompok saja (1 sampel atau 2 sampel). Sedangakan ANOVA dilakukan jika kita ingin membandingkan perbedaan antara berbagai kelompok yang jumlahnya lebih besar dari 2.



A. INDEPENDENT SAMPEL t-TEST

Pada uji beda ini, maka peneliti ingin mengetahui apakah dua kelompok yang ada berasal dari populasi yang sama atau tidak. Untuk mengetahui hal tersebut biasanya seorang peneliti memberikan perlakuan yang berbeda pada dua kelompok yang ada.

Perhatikan kata 'independen' atau 'bebas', yang berarti tidak ada hubungan antara dua sampel yang akan di uji. Hal ini berbeda dengan uji berpasangan (Paired sample t-Test), dimana satu kasus diobservasi lebih dari sekali. Dalam uji independen satu kasus hanya di data satu kali.

Kasus :

Contohnya adalah Seorang marketer ingin mengetahui bagaimana kondisi penjualan produk penghisap debu di dua wilayah pemasarannya, yakni Jakarta Pusat dan Jakarta Selatan selama 2 tahun (8 kuartal) berturut-turut. Hasil penjualan dari masing-masing daerah selama 8 kuartal tersebut adalah sebagai berikut:

Kuartal	Daerah Penjualan	Penjualan (dalam unit)
1	Jakarta Selatan	120
2	Jakarta Selatan	133
3	Jakarta Selatan	124
4	Jakarta Selatan	168
5	Jakarta Selatan	137
6	Jakarta Selatan	145
7	Jakarta Selatan	267
8	Jakarta Selatan	300
1	Jakarta Pusat	245
2	Jakarta Pusat	432
3	Jakarta Pusat	222
4	Jakarta Pusat	543
5	Jakarta Pusat	221
6	Jakarta Pusat	156
7	Jakarta Pusat	425
8	Jakarta Pusat	189

Keterangan:

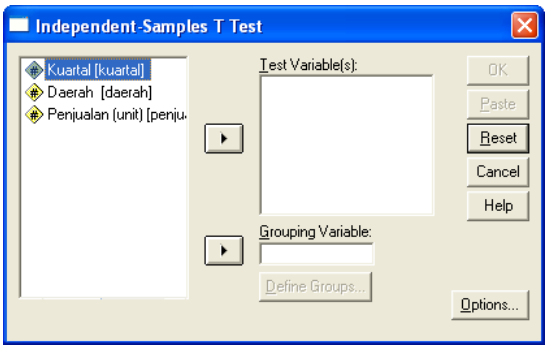
Kasus di atas terdiri atas dua sampel yang bebas satu dengan yang lain, yaitu sampel daerah Jakarta Pusat dan Jakarta Selatan.

Disini populasi diketahui berdistribusi normal dan karena sampel sedikit, dipakai uji t untuk dua sampel.

Pengolahan data dengan SPSS

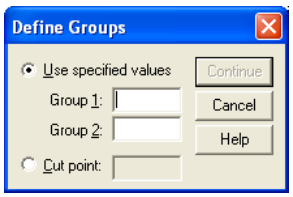
Langkah-langkah :

- buka lembar **t test 1**
- dari menu utama SPSS pilih menu **Analyze**, kemudian pilih submenu **Compare-Means.** Dari serangkaian pilihan test, sesuai kasus pilih **independent-samples t test**
- klik mouse pada pilihan tersebut, maka akan tampak di layar :



Pengisian :

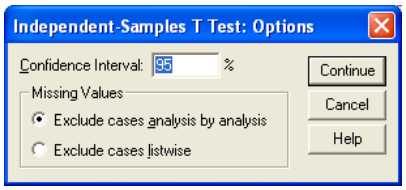
- Test variable(s) atau variabel yang akan di uji. Oleh karena di sini akan diuji data penjualan, maka masukkan variabel **penjualan**.
- Grouping variabel atau variabel group. Oleh karena variabel pengelompokkan ada pada variabel daerah maka masukkan variabel **daerah**
- Pengisian Group : klikmouse pada **Define Group,** tampak di layar :



- untuk group 1, isi dengan 1 yang berarti grup 1 berisi tanda 1 atau 'Jakarta Selatan'
- untuk group 2, isi dengan 2, yang berarti 'Jakarta Pusat'

  setelah pengisian selesai, tekan **Continue** untuk melanjutkan ke menu sebelumnya.

- Untuk kolom **Option** atau pilihan yang lain, dengan mengklik mouse akan tampak dilayar :



Pengisian :

- untuk Confidence Interval atau tingkat kepercayaan, biarkan pada angka default 95%
- untuk Missing Values ata data yang hilang. Oleh karena dalam kasus semua pasangan data komplit (tidak ada yang kosong), maka abaikan saja bagian ini.
- Tekan **Continue** jika pengisian di anggap selesai.

- Tekan OK untuk mengakhiri pengisian prosedur analisis. Terlihat SPSS melakukan pekerjaan analisis dan terlihat output SPSS.

Output SPPS dan Analisis

Berikut output dari independent t test :

Group Statistics

	Daerah	N	Mean	Std. Deviation	Std. Error Mean
Penjualan (unit)	Jakarta Selatan	8	174.2500	69.5573	24.5922
	Jakarta Pusat	8	304.1250	141.6075	50.0658

Pada bagian pertama terlihat ringkasan statistik dari kedua sampel. Untuk penjualan di daerah Jakarta Selatan (tanda 1) mempunyai penjualan rata-rata sebesar 174, 25 unit (dibulatkan menjadi 174 unit), yang jauh di bawah penjualan rata-rata Jakarta Pusat yaitu sebesar 304 unit penghisap debu.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Penjualan (unit)	Equal variances assumed	8.054	.013	-2.328	14	.035	-129.8750	55.7796	-249.5103	-10.2397
	Equal variances not assumed			-2.328	10.192	.042	-129.8750	55.7796	-253.8430	-5.9070

Pertama analisis menggunakan F test, untuk menguji apakah ada kesamaan varians pada data penjualan di Jakarta Selatan dan Jakarta Pusat.

Hipotesis

Hipotesis untuk kasus ini adalah :

Ho : Kedua varians populasi adalah identik (varians populasi penjualan di Jakarta Selatan dan Jakarta Pusat adalah sama)

Hi : kedua varians populasi tidak identik (varians populasi penjualan di Jakarta Selatan dan Jakarta Pusat adalah berbeda)

Pengambilan Keputusan

Dasar pengambilan keputusan :

Jika probabilitas > 0,05, Ho di tolak

Jika probabilitas < 0,05, Hi di terima

Keputusan :

Terlihat bahwa F hitung untuk tinggi badan dengan Equal variance assumed (diasumsi kedua varians sama atau menggunakan pooled variance t test) adalah 8.054 dengan probabilita 0,013. oleh karena probabilita < 0,05 maka Ho ditolak, atau kedua varians benar-benar berbeda.

Perbedaan yang nyata dari kedua varians membuat penggunaan varians untuk membandingkan rata-rata populasi dengan t test, sebaiknya menggunakan dasar **Equal variance not assumed** (diasumsi kedua varians tidak sama).

Selanjutnya dilakukan analisis dengan memakai t test untuk asumsi varians tidak sama.

Hipotesis

Hipotesis untuk kasus ini :

Ho : kedua rata-rata populasi adalah identik (rata-rata populasi penjualan di Jakarta Selatan dan Jakarta Pusat adalah sama)

Hi : kedua rata-rata populasi adalah tidak identik (rata-rata populasi penjualan di Jakarta Selatan dan Jakarta Pusat adalah sama)

Keputusan

Terlihat bahwa t hitung untuk penjualan dengan Equal varianced not assumed (diasumsi kedua varians tidak sama atau menggunakan separate varians test) adalah -2,328 dengan probabilita 0,042. Oleh karena probabilita < 0,05, maka Ho di tolak atau kedua rata-rata (mean) populasi penjualan di Jakarta Selatan dan Jakarta Pusat adalah berbeda berbeda, dalam artian Jakarta Pusat mempunyai rata-rata penjualan yang lebih dari Jakarta Selatan.

B. PAIRED SAMPLE T-TEST

Paired Sampel yang berpasangan diartikan sebagai sebuah sampel dengan subyek yang sama namun mengalami dua perlakuan atau pengukuran yang berbeda, seperti Subjek X akan mendapat perlakuan I kemudian perlakuan II.

Kasus

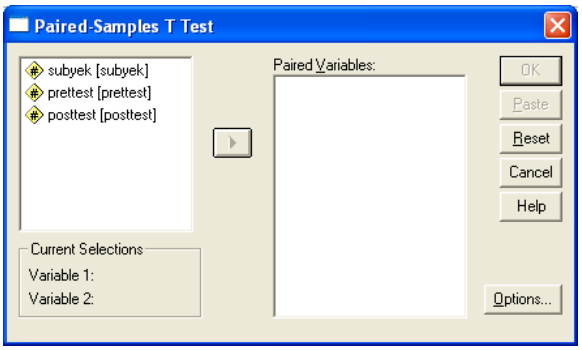
Sebuah Universitas berusaha meningkatkan kemampuan staf pengajar mereka dalam berbahasa Inggris dengan memberikan pelatihan selama 1 semester. Untuk mengetahui keberhasilan program pelatihan tersebut, pihak Universitas mengadakan *pretest* dan *posttest* pada beberapa staf pengajar yang mengikuti program tersebut. Berikut ini data perolehan nilai mereka dalam sebuah skala penilaian 50-150.

Subyek	Pretest	Posttest
1	60	107
2	85	111
3	90	117
4	110	125
5	115	122

Dari data diatas, kita melihat bahwa seorang subyek mengalami dua perlakuan yang sama, untuk itu kita menggunakan analisis paired sample t-test.

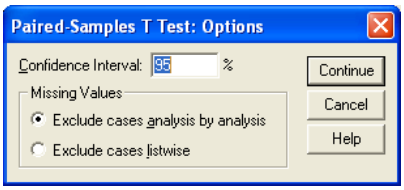
Pengolahan Data dengan SPSS

- Buka file **paired t test**
- Dari menu utama SPSS pilih menu **Analyze,** kemudian pilih submenu **Compare-Means.** Dari serangkaian pilihan test, sesuai kasus pilih **Paired-Samples T Test.**
  Tampak di layar :



- **Paired Variable(s)** atau Variabel yang akan diuji. Karena di sini yang akan diuji adalah data **pretest** dan **posttest ,** klik mouse pada variabel **pretest** kemudian klik mouse sekali lagi pada variabel **posttest.** Maka terlihat pada kolom Current Selection dibawah, terdapat keterangan untuk variabel 1 dan 2. Kemudian klik mouse pada tanda > (yang sebelah atas), maka Paired Variables terlihat tanda **pretest...posttest.**
  Keterangan : variabel pretest dan posttest harus dipilih secara bersamaan. Jika tidak, SPSS tidak bisa menginput dalam kolom Paired Variables, dengan tidak aktifnya tanda →

- Untuk kolom **option** atau pilihan yang lain, dengan mengklik mouse, tampak dilayar:



Untuk **Confidence Interval** atau tingkat kepercayaan. Sebagai default, SPSS menggunakan tingkat kepercayaan 95% atau tingkat signifikansi 100%-95%=5%

Untuk **Missing Values** atau data yang hilang. Oleh karena dalam kasus semua pasangan data lengkap (tidak ada yang kosong), maka abaikan saja bagian ini (tetap pada default dari SPSS yaitu Exclude cases analysis by analysis).

Tekan **Continue** jika pengisian selesai, sekarang SPSS akan kembali ke kotak dialog utama uji t paired.

- Kemudian tekan OK untuk mengakhiri pengisian prosedur analisis. Terlihat SPSS melakukan pekerjaan analisis dan terlihat output SPSS

Interpretasi Output SPSS

Berikut output dari uji_t_paired :

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	prettest	92.0000	5	21.9659	9.8234
	posttest	116.4000	5	7.4699	3.3407

Pada bagian pertama terlihat ringkasan statistik dari kedua sampel. Untuk kemampuan rata-rata para staf pengajar dalam berbahasa inggris sebelum mengikuti pelatihan adalah sebesar 92. Sedangkan setelah mengikuti pelatihan adalah sebesar 116.4.

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	prettest & posttest	5	.946	.015

Bagian kedua output adalah hasil korelasi antara kedua variabel, yang menghasilkan angka 0,946 dengan nilai probabilitas sebesar 0,015. Hal ini menyatakan bahwa korelasi antara nilai pretest dan posttest sangat nyata.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	prettest - posttest	-24.4000	15.0930	6.7498	-43.1405	-5.6595	-3.615	4	.022

Hipotesis

Hipotesis untuk kasus ini adalah :

Ho : Kedua rata-rata populasi adalah identik (rata-rata nilai pretest dan posttest tidak berbeda secara nyata)

Hi : Kedua rata-rata populasi adalah tidak identik (rata-rata nilai pretest dan posttest adalah memang berbeda secara nyata)

Pengambilan Keputusan

Dasar pengambilan keputusan berdasarkan nilai probabilitas:

jika probabilita > 0,05 ; maka Ho ditolak

jika probabilita < 0,05 ; maka Ho ditolak

Keputusan:

Terlihat bahwa t hitung adalah -3.615 dengan probabilitas 0,022 oleh karena probabilitas < 0,05 maka Ho ditolak atau kedua rata-rata populasi adalah tidak identik (rata-rata nilai pretest dan posttest berbeda secara nyata)

C. ANOVA 1 ARAH (ONE-WAY ANOVA)

Tujuan utama dari ANOVA adalah untuk membandingkan *mean* dari tiga kelompok atau lebih, untuk memberikan informasi apakah perbedaan yang teramati (*observed differences*) antar kelompok tersebut terjadi karena kebetulan *(chance)* atau karena suatu pengaruh tertentu yang bersifat sistematis *(systematic effect)*.

Analisis Varians (ANOVA) mensyaratkan adanya Variabel Dependen (DV) yang memiliki skala interval atau rasio dan satu atau lebih Variabel Independen (IV) yang seluruhnya bersifat kategori atau yang merupakan kombinasi dari variabel bersifat kategorik dengan variabel berskala interval atau rasio.

ANOVA berusaha membandingkan variabilitas skor yang terjadi dalam suatu kelompok (within group, yakni variabilitas yang disebabkan oleh *sampling error* itu sendiri) dengan variabilitas yang terjadi antar kelompok (between group, yakni variabilitas yang disebabkan karena efek dari suatu perlakuan/*treatment* dan variabilitas yang disebabkan karena *sampling error*).

Post Hoc Test

Konsep ANOVA sesungguhnya hampir sama dengan t-test, hanya saja ANOVA biasanya membedakan lebih dari dua kelompok. Pada ANOVA kita biasanya akan melakukan apa yang disebut Post hoc test. Hal ini dilakukan untuk membandingkan satu kelompok dengan kelompok lainnya satau per satu dan tidak secara bersamaan. Proses ini pada dasarnya sama

dengan uji beda t-test. Adapun metode-metode dalam post hoc test tergolong banyak sekali jumlahnya. Penulis hanya akan membahas dua metode saja di sini.

1. Scheffe's Test

Metode ini paling banyak digunakan untuk melakukan post hoc comparison. Metode tersebut memungkinkan peneliti untuk membandingkan mean secara berpasangan dengan kombinasi yang sangat kompleks.

Dalam setiap post hoc comparison, distribusi t dapat kita gunakan untuk menentukan apakah perbedaaan mean yang terjadi antara dua kelompok sampel terjadi secara sistematis atau kebetulan semata.

Untuk setiap perbandingan, maka hipotesis nol dan hipotesis alternative yang kita buat ditulis untuk menunjukkan kontras yang terjadi antar populasi (C) : H_o:C=0 dan H_1:C$\neq$ 0.

2. Tukey's HSD Test

Tukey's HSD (Honestly Significant Difference) test dirancang untuk dapat melakukan perbandingan mean antar kelompok pada semua tingkat signifikansi tes. Tes ini jauh lebih kuat dibandingkan Scheffe's test. , namun tidak dapat digunakan untuk menguji perbandingan yang bersifat kompleks.

Kasus

Seorang pengusaha ingin mengatahui apakah terdapat perbedaan penjualan pada tiga strategi distribusi yang dipilihnya (intensif, selektif dan eksklusif). Sebagai seorang konsultan, lakukanlah analisis terhadap data penjualan (dalam milyar rupiah) yang dimiliki oleh pengusaha tersebut!

Intensif	Selektif	Eksklusif
5	1	3
6	2	3
8	3	4

Kasus diatas memiliki satu variabel independen berjenis kategorik, yakni strategi distribusi. Variabel dependen-nya berskala metric yakni penjualan.

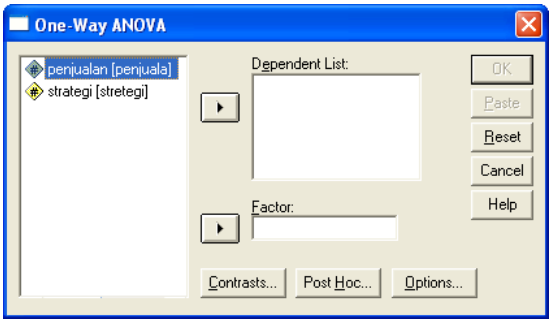
Pengolahan data dengan SPSS

Buka lembar kerja / file **one way anova** sesuai kasus diatas, atau jika sudah terbuka, ikuti prosedur berikut ini:

- Dari menu utama SPSS, pilih menu **Analyze**, kemudian pilih submenu **Compare-Means**.

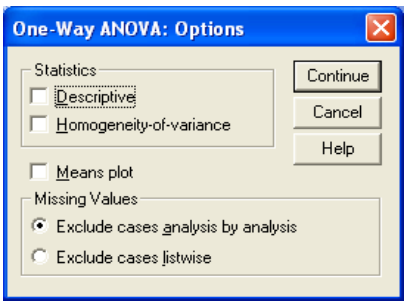
  Dari serangkaian pilihan test, sesuai kasus pilih One-Way Anova…

  Klik pilihan tersebut, maka tampak dilayar:



Pengisian:

- **Dependent List** atau variabel dependen yang akan diuji. Oleh karena yang akan diuji tingkat penjualan, maka klik variabel tingkat penjualan, kemudian klik tanda> (yang sebelah atas), maka variabel tingkat penjualan berpindah ke Dependent List.
- **Factor atau grup**.Oleh karena itu, variabel pengelompokkan ada pada variabel strategi promosi, maka klik strategi promosi, kemudian klik tanda >, maka variabel jenis strategi promosi akan berpindah ke Factor. Untuk kolom Option atau pilihan lain dengan mengkliknya maka tampak di layar:



Pengisian:

Untuk Statistics atau perhitungan statistic yang akan dilakukan, untuk keseragaman akan dipilih Descriptive dan Homogeinity of Variance. Untuk itu, klik kedua pilihan tersebut.

Interpretasi Output SPSS:

Descriptives

penjualan

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
intensif	3	6.3333	1.5275	.8819	2.5388	10.1279	5.00	8.00
selektif	3	2.0000	1.0000	.5774	-.4841	4.4841	1.00	3.00
eksklusif	3	3.3333	.5774	.3333	1.8991	4.7676	3.00	4.00
Total	9	3.8889	2.1473	.7158	2.2383	5.5395	1.00	8.00

Dapat diketahui bahwa rata-rata penjualan pada melalui strategi intensif adalah sebesar 6, 3 milyar.

Penjualan terendah untuk strategi intensif adalah sebesar 5 milyar, dan tertinggi adalah 8 milyar

Penjualan dengan strategi lainnya dapat diinterpretasikan dengan cara yang sama.

Test of Homogeneity of Variances

penjualan

Levene Statistic	df1	df2	Sig.
1.217	2	6	.360

Analisis Homogeneity of Variances ini bertujuan untuk menguji berlaku tidaknya asumsi untuk ANOVA, yaitu apakah penjualan dengan menggunakan ketiga strategi promosi yang ada mempunyai varians yang sama:

Hipotesis pada bagian ini adalah:

Ho: Ketiga varians populasi adalah identik

Hi: Kedua varians populasi tidak identik

Pengambilan keputusan:

jika probabilita > 0,05 ; maka Ho diterima

jika probabilita < 0,05 ; maka Ho ditolak

Keputusan:

dari informasi diatas, Sig 0,360 >dari 0,05. oleh karena itu Ho diterima. Dengan kata lain varians identik.

ANOVA

penjualan

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	29.556	2	14.778	12.091	.008
Within Groups	7.333	6	1.222		
Total	36.889	8			

Bagian ini menguji apakah ketiga populasi diatas memiliki mean yang sama:

Hipotesis pada bagian ini adalah:

Ho: Ketiga mean populasi adalah identik

Hi: Kedua mean populasi tidak identik

Pengambilan keputusan

Berdasar nilai signifikansi:

Jika Sig >0,05 maka Ho diterima

Jika Sig<0,05 maka Ho ditolak

Signifikansi analisis ini adalah sebesar 0,008 berarti <0,05. Hal ini menunjukkan bahwa rata-rata penjualan dengan menggunakan tiga strategi promosi tersebut berbeda secara signifikan

Multiple Comparisons

Dependent Variable: penjualan

	(I) promosi	(J) promosi	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	intensif	selektif	4.3333*	.9027	.007	1.5637	7.1030
		eksklusif	3.0000*	.9027	.037	.2303	5.7697
	selektif	intensif	-4.3333*	.9027	.007	-7.1030	-1.5637
		eksklusif	-1.3333	.9027	.365	-4.1030	1.4363
	eksklusif	intensif	-3.0000*	.9027	.037	-5.7697	-.2303
		selektif	1.3333	.9027	.365	-1.4363	4.1030
Scheffe	intensif	selektif	4.3333*	.9027	.009	1.4382	7.2284
		eksklusif	3.0000*	.9027	.044	.1049	5.8951
	selektif	intensif	-4.3333*	.9027	.009	-7.2284	-1.4382
		eksklusif	-1.3333	.9027	.394	-4.2284	1.5618
	eksklusif	intensif	-3.0000*	.9027	.044	-5.8951	-.1049
		selektif	1.3333	.9027	.394	-1.5618	4.2284

*. The mean difference is significant at the .05 level.

Setelah kita mengetahui bahwa terdapat perbedaan yang signifikan pada ketiga daerah, maka kita akan berusaha untuk mengetahui perbedaan antara satu strategi dengan strategi lainnya.

Scheffe Test.

Catatan: Tanda * pada kolom mean differences menunjukkan bahwa perbedaan yang ada tergolong signifikan.

- Perbandingan antara strategi intensif dengan strategi selektif tergolong signifikan, begitu pun perbandingan antara strategi intensif dengan strategi ekslusif tergolong signifikan.

  Interpretasi yang sama di lakukan terhadap strategi lainnya.

D. ANOVA 2 ARAH (TWO WAY ANOVA)

Uji Anova dua arah disebut juga Faktorial Anova, yang bermanfaat untuk menguji beberapa hipotesis mengenai perbedaan mean dalam desain faktorial (variabel independen $\geq$ 2, sehingga terdapat $\geq$ 4 kelompok dalam desain tersebut)

Persyaratan utama dari desain ini adalah :

1. Terdapat dua variabel independent, yang mana masing-masing variabel independent tersebut memiliki dua atau lebih level/kelompok. Dua variabel independent tersebut diasumsikan saling bersilangan satu sama lain.
2. Level atau kelompok dalam masing-masing variabel independen dapat berbeda baik secara kuantitatif maupun kualitatif.
3. Subyek hanya boleh muncul dalam satu sel desain faktorial sekali saja.

Tujuan dari Anova dua arah ini adalah untuk membandingkan nilai mean dari empat atau lebih kelompok dalam suatu desain faktorial, sehingga dapt diketahui:

1. Apakah perbedaan mean yang ada, terjadi karena kebetulan, ataukah merupakan pengaruh dari faktor pertama (*main effect* untuk faktor A).
2. Apakah perbedaan mean yang ada, terjadi karena kebetulan, ataukah merupakan pengaruh dari faktor kedua (*main effect* untuk faktor B).
3. Apakah perbedaan mean yang ada, terjadi karena kebetulan, ataukah merupakan pengaruh dari interaksi antara faktor pertama dengan faktor kedua (*interaction effect*).

Seringkali ada yang mengajukan pertanyaan, mengapa peneliti tidak melakukan suatu riset yang berbeda untuk masing-masing variabel. Hal tersebut tidak dilakukan dengan alasan untuk:

1. Menghemat waktu dan sumber daya
2. Memperoleh informasi yang lebih lengkap dari desain faktorial, karena informasi ini tidak dapat diprediksi dari pengaruh faktor 1 dan faktor 2 secara independent.
3. Estimasi dari varians error akan lebih akurat dalam desain faktorial dibandingkan dengan desain Anova satu arah.

Dalam anova satu arah, total variabilitas antar skor dibagi ke dalam 2 sumber variabilitas yang bersifat independent. Yaitu variabilitas dalam kelompok (within-group variability) dan variabilitas antar kelompok atau (between group variability).

Total variability=variability between group + variability within group (error).

Sedangkan dalam anova dua arah, sumber variabilitasnya dibagi ke dalam 4 sumber

Total variabilitas = variabilitas karena faktor A+ variabilitas karena faktor B+ variabilitas karena interaksi A&B+variabilitas dalam kelompok(error).

Kasus :

Seorang produsen ingin mengetahui apakah perbedaan kandungan alkohol dalam produk parfum yang diluncurkannya baru-baru ini memiliki pengaruh terhadap tingkat pembelian. Parfum itu sendiri baru diluncurkan pada satu cabang tokonya di Jawa dan Bali. Berikut ini data tingkat pembelian (dalam juta) produk dari konsumen parfum tersebut selama bulan April dan Mei:

Daerah		Tingkat Kandungan Alkohol		
		5% Mid winter	10% Delice	(15%)Together
Bali	April	100	100	700
	Mei	200	200	800
Jawa	April	800	300	300
	Mei	900	400	400

Pemasukan data ke SPSS

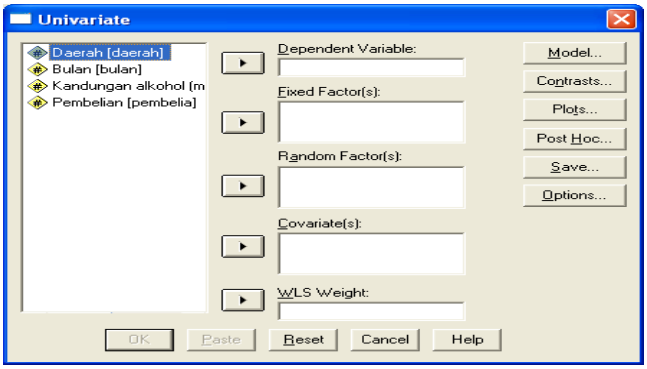
Dari menu utama File, Anova 2 arah

Pengolahan data dengan SPSS

Langkah-langkah :

- Buka file **two way anova**
- Dari menu utama SPSS pilih menu **Analyze**, kemudian pilih submenu **General Linear Model**. Dari serangkaian pilihan test, sesuai kasus pilih **Univariate**...

  Maka tampak di layar:



Pengisian:

- **Dependent** atau variabel dependen yang akan diuji. Yang akan diuji adalah tingkat pembelian konsumen, klik variabel tingkat pembelian, kemudian klik tanda >maka variabel pembelian akan berpindah ke dependen.
- **Factors atau faktor /grup**, masukkan variabel daerah dan kandungan alkohol (merek parfum) Catatan: Perhatikan isi fixed factor selalu berupa data nominal seperti 1,2 dan seterusnya.
- Klik **Model**, pilih full factorial, kemudian continue
- Klik **Post hoc**, kemudian klik daerah dan merek lalu klik tanda ">" ke bagian Post hoc test for. Kemudian klik **Equal variances assumed**, lalu pilih **Scheffe** dan **Tukey test**. Kemudian continue.
  Abaikan pilihan lain. Lalu klik OK.

Interpretasi Output SPSS

Between-Subjects Factors

		Value Label	N
Daerah	1.00	Bali	6
	2.00	Jawa	6
Kandungan alkohol (merk)	1.00	Mid Winter	4
	2.00	Delice	4
	3.00	Together	4

Pada bagian ini terlihat ringkasan dari data yang diproses. Dari tabel terlihat bahwa untuk masing-masing daerah, yakni Jawa dan Bali, terdapat 6 kasus yang diproses. Sedangkan berdasarkan kandungan alcohol (merk), ada 4 kasus yang diproses untuk masing-masing merek.

Tests of Between-Subjects Effects

Dependent Variable: Pembelian

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	896666.667[a]	5	179333.333	35.867	.000
Intercept	2253333.333	1	2253333.333	450.667	.000
DAERAH	83333.333	1	83333.333	16.667	.006
MERK	206666.667	2	103333.333	20.667	.002
DAERAH * MERK	606666.667	2	303333.333	60.667	.000
Error	30000.000	6	5000.000		
Total	3180000.000	12			
Corrected Total	926666.667	11			

a. R Squared = .968 (Adjusted R Squared = .941)

Sesuai dengan tujuan dari desain faktorial, maka ada 3 perbedaan mean yang akan kita uji:

1. **Apakah perbedaan mean yang ada, terjadi karena kebetulan, ataukah merupakan pengaruh dari interaksi antara faktor pertama dengan faktor kedua (*interaction effect*).**

Efek interaksi harus dikedepankan dalam interpretasi Anova dua arah, sebab nilai pembelian parfum oleh konsumen mungkin lebih disebabkan karena kombinasi dari efek faktor pertama (daerah) dan faktor kedua (kandungan alcohol/merk)

Pengambilan Keputusan

Hipotesis pada bagian ini adalah:

Ho: Tidak ada interaksi antara daerah dengan kandungan alcohol (merek parfum)

Hi: Ada interaksi antara daerah dengan kandungan alcohol (merek parfum)

Pengambilan Keputusan:

jika probabilita > 0,05 ; maka Ho diterima

jika probabilita < 0,05 ; maka Ho ditolak

Keputusan:

Terlihat bahwa F nilai signifikansi adalah sebesar 0,000 yang berarti <0,005, maka Ho ditolak. Berarti memang ada interaksi antara daerah dengan kandungan alcohol (merek parfum).

2. **Apakah perbedaan mean yang ada, terjadi karena kebetulan, ataukah merupakan pengaruh dari faktor pertama (*main effect* untuk faktor daerah).**

Hipotesis

Ho: Mean kedua populasi (Bali dan Jawa) adalah identik

Hi: Mean kedua populasi (Bali dan Jawa) tidak identik

Pengambilan Keputusan

jika probabilita > 0,05 ; maka Ho diterima

jika probabilita < 0,05 ; maka Ho ditolak

Keputusan:

Terlihat nilai signifikansi adalah sebesar 0,006 yang berarti <0,005, maka Ho ditolak. Berarti mean pembelian parfum di kedua daerah tersebut memang berbeda secara nyata.

3. **Apakah perbedaan mean yang ada, terjadi karena kebetulan, ataukah merupakan pengaruh dari faktor kedua (*main effect* untuk faktor kandungan alkohpol / merk parfum).**

Hipotesis:

Ho: Mean ketiga populasi (Mid winter, Delice,dan Together) adalah identik

Hi: Mean ketiga populasi (Mid winter, Delice,dan Together) tidak identik

Pengambilan Keputusan:

jika probabilita > 0,05 ; maka Ho diterima

jika probabilita < 0,05 ; maka Ho ditolak

Keputusan:

Terlihat bahwa F nilai signifikansi adalah sebesar 0,002 yang berarti <0,005, maka Ho ditolak. Berarti mean pembelian parfum ketiga merk tersebut memang berbeda secara nyata.

Multiple Comparisons

Dependent Variable: Pembelian

	(I) Kandungan alkohol (merk)	(J) Kandungan alkohol (merk)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Mid Winter	Delice	250.00*	50.00	.006	96.59	403.41
		Together	-50.00	50.00	.603	-203.41	103.41
	Delice	Mid Winter	-250.00*	50.00	.006	-403.41	-96.59
		Together	-300.00*	50.00	.002	-453.41	-146.59
	Together	Mid Winter	50.00	50.00	.603	-103.41	203.41
		Delice	300.00*	50.00	.002	146.59	453.41
Scheffe	Mid Winter	Delice	250.00*	50.00	.007	89.64	410.36
		Together	-50.00	50.00	.630	-210.36	110.36
	Delice	Mid Winter	-250.00*	50.00	.007	-410.36	-89.64
		Together	-300.00*	50.00	.003	-460.36	-139.64
	Together	Mid Winter	50.00	50.00	.630	-110.36	210.36
		Delice	300.00*	50.00	.003	139.64	460.36

Based on observed means.

*. The mean difference is significant at the .05 level.

Setelah kita mengetahui bahwa terdapat perbedaan yang signifikan baik dari variabel daerah maupun variabel kandungan alkohol (merk), maka kita akan berusaha untuk mengetahui perbedaan antara satu merek dengan merek lainnya satu persatu. Kita tidak mungkin melakukan post hoc comparison untuk variabel daerah karena hanya terdiri dari 2 kelompok (Post hoc mensyaratkan minimal ada 3 kelompok untuk dibandingkan).

Baik Tukey maupun Scheffe Test menghasilkan nilai yang tidak jauh berbeda.

Catatan: Tanda * pada kolom mean differences menunjukkan bahwa perbedaan yang ada tergolong signifikan.

- Perbandingan perbedaan mean antara Mid Winter dengan Delice menunjukkan perbedaan yang sangat signifikan. Namun perbandingan antara Mid Winter dengan Together tidak menunjukkan adanya hubungan yang signifikan.
- Perbandingan perbedaan mean antara Delice dengan Mid Winter dan Together cukup signifikan.
- Perbandingan antara Together dengan Midwinter tergolong tidak signifikan, namun perbandingan antara Together dengan Delice tergolong signifikan.

Pembelian

	Kandungan alkohol (merk)	N	Subset 1	Subset 2
Tukey HSD[a,b]	Delice	4	250.00	
	Mid Winter	4		500.00
	Together	4		550.00
	Sig.		1.000	.603
Scheffe[a,b]	Delice	4	250.00	
	Mid Winter	4		500.00
	Together	4		550.00
	Sig.		1.000	.630

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 5000.000.

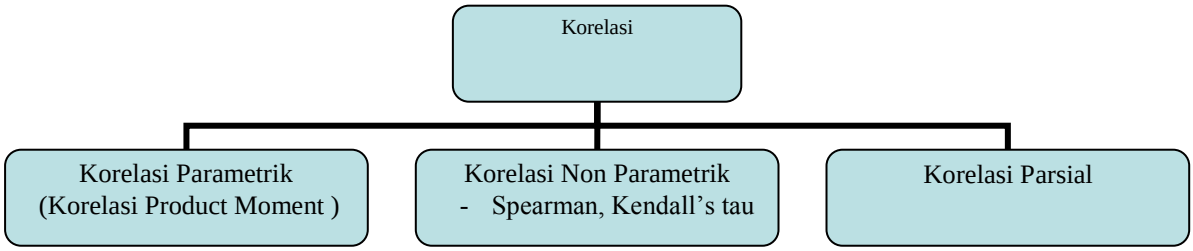
a. Uses Harmonic Mean Sample Size = 4.000.

b. Alpha = .05.

Pada bagian ini, yang akan kita cari adalah grup/subset mana saja yang mempunyai perbedaan mean yang tidak signifikan.

Terlihat bahwa sampel terbagi ke dalam 2 subset, yang menunjukkan bahwa ada 2 merk yang memiliki mean pembelian yang tidak terlalu berbeda secara nyata. Dua merek yang tidak berbeda pembeliannya adalah Mid Winter dan Together.Mean dari pembelian parfum Delice berbeda secara signifikan dengan parfum merk Mid Winter dan Together. Namun dari informasi diatas, terlihat bahwa rata-rata pembelian dari parfum Delice tergolong lebih rendah, jauh dibawah Mid Winter dan Together.

UJI KORELASI



A. KORELASI PARAMETRIK

Korelasi product moment (*r)* adalah suatu indeks statistik yang paling sering digunakan untuk untuk mengukur kekuatan hubungan antara dua variabel metrik. Misalnya saja X dan Y. Dengan kata lain, indeks tersebut dapat menunjukkan pada kita seberapa besar hubungan antara variasi yang terjadi pada variabel X dengan variasi yang terjadi pada variabel Y. Perlu kita catat bahwa dalam korelasi kita belum menentukan dengan pasti variabel independent dan dependen-nya, seperti yang kita lakukan dalam analisis regresi.

Kasus

Seorang peneliti ingin mengetahui apakah rasa humor seseorang berhubungan dengan sifat keterbukaan seseorang (extraversion). Dengan menggunakan *extraversion scale* dan *humor scale*, diperolehlah data seperti dibawah ini.

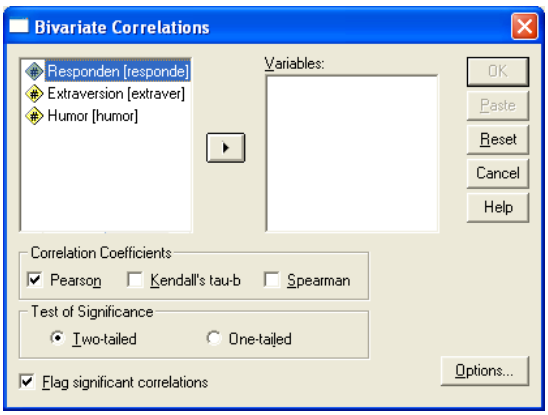
Responden	Extraversion	Humor
1	1	4
2	1	8
3	2	10
4	3	14
5	4	15
6	5	16
7	6	17
8	7	18
9	8	18

Pengolahan data dengan SPSS

Langkah-langkah :

- Buka file **Korelasi Parametrik**
- Dari menu utama SPSS pilih menu **Analyze**, kemudian pilih submenu **Correlate**. Lalu klik **Bivariate**

  Tampak di layar:



Pengisian:

- Variable atau variable yang akan dikorelasikan. Karena variabel yang akan kita korelasikan adalah extraversion dengan humor, maka klik-lah extraversion lalu tekan tanda '>', maka variabel promosi akan berpindah ke variable.
- Demikian pula untuk variable humor.
- Untuk kolom **Correlation Coefficient** atau alat hitung koefisien korelasi, maka klik **Pearson** karena kedua variabel memiliki skala metric.
- Untuk kolom **Test of Significance**, karena yang akan diuji dua sisi, maka pilih two tailed.
- Untuk **Flag Significant correlations** atau berkenaan dengan tanda untuk tingkat signifikansi 5% dan 10% akan ditampilkan pada output atau tidak. Untuk keseragaman pilihan tersebut dipakai hingga nanti pada output ada tanda * untuk 5% dan atau tanda ** untuk 1%.
- Tekan **continue** jika sudah selesai. Kemudian tekan **OK** untuk mengakhiri prosedur korelasi melalui SPSS. Kemudian akan kita peroleh output

Interpretasi Output SPSS :

Correlations

		Extraversion	Humor
Extraversion	Pearson Correlation	1	.915**
	Sig. (2-tailed)	.	.001
	N	9	9
Humor	Pearson Correlation	.915**	1
	Sig. (2-tailed)	.001	.
	N	9	9

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

Dilihat dari tanda ** maka korelasi antara Extraversion dengan Humor seseorang tergolong signifikan, yakni sebesar 0,915. Berarti ada hubungan yang positif antara kedua variabel. Sebagai informasi, Angka korelasi berkisar antara -1 hingga 1. Dimana 0 menunjukkan tidak adanya korelasi. Sedangkan tanda negative positf menunjukkan arah hubungan antar varaiabel.

B. Korelasi Non Parametrik

Prosedur ini dapat dimanfaatkan untuk mengukur hubungan antar variabel non metrik. Perlu kita ingat bahwa variabel dengan skala non metrik tidak memiliki sifat seperti data berskala interval atau rasio.

Kita dapat menggunakan korelasi Spearman rho atau korelasi Kendall's Tau. Spearman rho dapat kita gunakan terutama saat kita memiliki jumlah kategori yang tergolong besar. Sedangkan Kendall's tau lebih disarankan untuk digunakan pada data yang tergolong besar namun hanya dikelompokkan dalam kategori yang lebih sedikit.

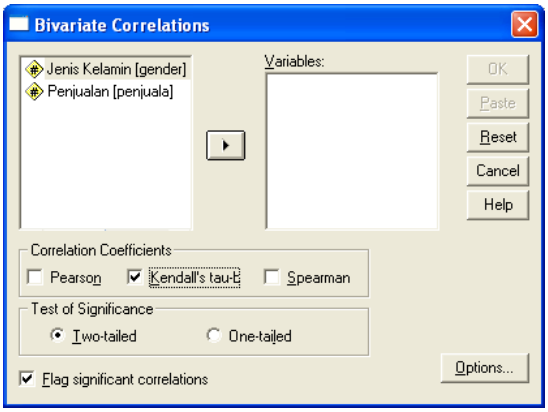
Kasus :

Seorang produsen ingin mengetahui apakah ada hubungan antara jenis kelamin tenaga penjualannya dengan tingkat penjualan (dalam juta) yang dihasilkannya setiap bulan.

Gender	Penjualan
Pria	15
Wanita	14
Pria	12
Wanita	14
Pria	15
Pria	17
Wanita	13
Wanita	10

Langkah-langkah :

- Buka file **korelasi non parametrik**
- Dari menu utama SPSS pilih menu **Analyze**, kemudian pilih submenu **Correlate**. LALU Klik **Bivariate**



Pengisian:

- Variable atau variable yang akan dikorelasikan. Karena variabel yang akan kita korelasikan adalah jenis kelamin dengan penjualan, maka klik-lah jenis kelamin lalu tekan tanda '>', maka variabel jenis kelamin akan berpindah ke variable. Demikian pula untuk variable penjualan.
- Untuk kolom **Correlation Coefficient** atau alat hitung koefisien korelasi, maka klik **Kendall's** atau Spearman rho karena salah satu variabel memiliki skala non metric.
- Untuk kolom **Test of Significance**, karena yang akan diuji dua sisi, maka pilih **two tailed**.
- Untuk **Flag Significant correlations** atau berkenaan dengan tanda untuk tingkat signifikansi 5% dan 10% akan ditampilkan pada output atau tidak. Untuk keseragaman pilihan tersebut dipakai hingga nanti pada output ada tanda * untuk 5% dan atau tanda ** untuk 1%.

Interpretasi Output SPSS:

Correlations

			Jenis Kelamin	Penjualan
Kendall's tau_b	Jenis Kelamin	Correlation Coefficient	1.000	-.490
		Sig. (2-tailed)	.	.144
		N	8	8
	Penjualan	Correlation Coefficient	-.490	1.000
		Sig. (2-tailed)	.144	.
		N	8	8
Spearman's rho	Jenis Kelamin	Correlation Coefficient	1.000	-.552
		Sig. (2-tailed)	.	.156
		N	8	8
	Penjualan	Correlation Coefficient	-.552	1.000
		Sig. (2-tailed)	.156	.
		N	8	8

Penafsiran angka korelasi antara metric dan non metric pada dasarnya sama dengan korelasi Pearson. Dimana korelasi dibawah 0,5 menunjukkan korelasi yang lemah, dan korelasi diatas 0,5 menunjukkan korelasi yang kuat.

Dari data diatas, terlihat bahwa korelasi antara jenis kelamin ternyata sangat rendah yakni hanya sebesar 0,144

Hipotesis:

Ho: tidak ada korelasi antara jenis kelamin dengan tingkat penjualan

Ha: ada korelasi antara jenis kelamin dengan tingkat penjualan

Pengambilan keputusan:

Jika probabilita > 0,05 ; maka Ho diterima

Jika probabilita < 0,05 ; maka Ho ditolak

Keputusan:

Karena probabilita korelasi antara jenis kelamin dengan penjualan adalah sebesar 0.156 atau 0,144 yang nilainya > 0,05 maka Ho diterima.Berarti tidak ada korelasi antara jenis kelamin dengan tingkat penjualan.

C. KORELASI PARSIAL

Korelasi parsial ini berusaha untuk mengukur kekuatan antara dua variabel yang ada, namun setelah kita melakukan kontrol terhadap variabel lain yang dapat mempengaruhi hubungan kedua variabel yang sedang kita ukur tersebut.

Kasus :

Seorang produsen ingin mengetahui apakah ada hubungan antara besarnya budget promosi yang dikeluarkan perusahaan dengan tingkat penjualan produknya. Dengan mempertimbangkan variabel ketiga yang dapat mempengaruhi keduanya yaitu jumlah SPG Berikut ini data budget promosi, penjualan dan jumlah SPG perusahaan tersebut selama 7 tahun belakangan ini

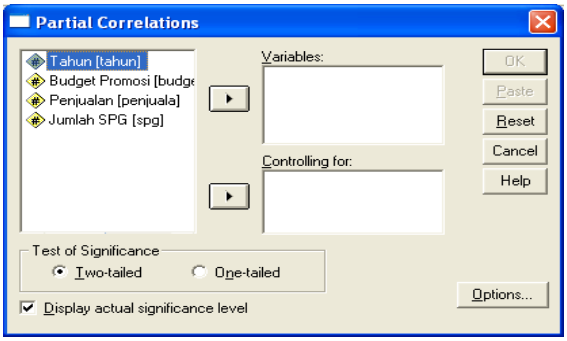
Tahun	Budget Promosi	Penjualan	Jumlah
	(dalam ratusan juta)	(dalam ratusan juta)	SPG
1996	8	10	8
1997	7	8	6
1998	3	2	4
1999	5	6	7
2000	7	9	7
2001	2	2	4
2002	4	5	4

Pengolahan Data dengan SPSS

Langkah-langkah :

- Buka file **Korelasi Parsial**
- Dari menu utama SPSS pilih menu Analyze, kemudian pilih submenu Correlate. LALU Klik partial…

  Tampak di layar:



Pengisian:

- **Variable** atau variable yang akan dikorelasikan. Karena variabel yang akan kita korelasikan adalah budget promosi dengan penjualan, maka klik-lah budget promosi lalu tekan tanda '>', maka variabel budget promosi akan berpindah ke variable.

  Demikian pula untuk variable penjualan.

  Controlling for atau variabel pengontrol pilih SPG
- Untuk kolom **Test of Significance**, karena yang akan diuji dua sisi, maka pilih two tailed.
- Untuk **Flag Significant correlations** atau berkenaan dengan tanda untuk tingkat signifikansi 5% dan 10% akan ditampilkan pada output atau tidak. Untuk keseragaman pilihan tersebut dipakai hingga nanti pada output ada tanda * untuk 5% dan atau tanda ** untuk 10%.
- Tekan **continue** jika sudah selesai. Kemudian tekan OK untuk mengakhiri prosedur korelasi melalui SPSS.

Interpretasi Output SPSS

- - - P A R T I A L C O R R E L A T I O N C O E F F I C I E N T S - - -

Controlling for.. SPG

	PENJUALA	BUDGET
PENJUALA	1.0000	.9264
	(0)	(4)
	P= .	P= .008
BUDGET	.9264	1.0000
	(4)	(0)
	P= .008	P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

Signifikansi hasil korelasi:

Dari output terlihat bahwa besaran korelasi parsial adalah 0.008, yang < dari 0,05. Arti angka Korelasi:

Coba lakukan terlebih dahulu antara penjualan dengan budget yang menghasilkan nilai korelasi yang tidak memperhitungkan variabel pengontrol yakni SPG, sebesar 0.983. Setelah dilakukan korelasi parsial, terlihat penurunan nilai korelasi dari 0,983 menjadi 0,9264. Sedangkan tanda korelasi masih positif. Ini berarti SPG tidak terlalu mempengaruhi tingkat penjualan dari perusahaan tersebut. Oleh karena itu perusahaan perlu meningkatkan efektifitas SPG-nya.

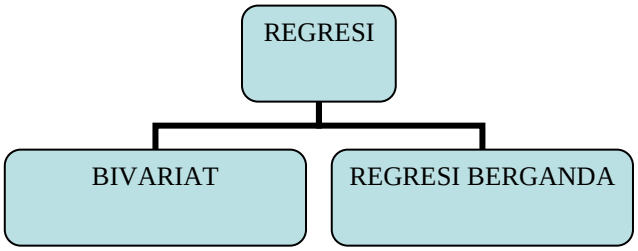
ANALSIS MULTIVARIATE

Analisis Multivariate terkait dengan metoda statistik yang menganalisa secara serentak berbagai pengukuran pada setiap objek penelitian yang terdiri dari jumlah variabel yang lebih dari 2 variabel. Berbagai metoda dapat digunakan pada analisis multivariate, diantaranya multiple regression, factor analysis, discriminant analysis, dan cluster analysis.

A. REGRESSI

Regresi adalah suatu prosedur statistik untuk menganalisa hubungan asosiatif antara suatu variabel dependen dengan variabel independent dalam bentuk metrik.

Regresi secara umum dapat di bagi dua yakni:



Meskipun variabel independent yang ada dapat menjelaskan variasi yang terjadi dalam variabel independent, namun hal tersebut tidak secara pasti menunjukkan hubungan kausalitas.Penggunaan istilah variabel dependen (*criterion variables*) dan variabel independent (*predictor*)

A. REGRESI BIVARIAT (BIVARIATE REGRESSION)

Regresi bivariat merupakan suatu prosedur untuk menurunkan suatu hubungan matematis, dalam bentuk persamaan, antara suatu variabel dependen berbentuk metrik dengan sebuah variabel independent yang juga berbentuk metric.

Rumus umum dari persamaan regresi bivariat:

$Y = \beta_o + \beta_1 X$

X= Variabel dependen

Y = Variabel independent atau predictor

β_o = Intercept dari garis

β_1 =Slope dari garis

Model diatas menunjukkan suatu hubungan yang bersifat deterministik dari Y, yang secara total hanya dijelaskan oleh X. Nilai dari Y akan dapat kita ketahui jika nilai B_o dan B_1 diketahui

Estimasi Parameter

Pada sebagian besar kasus, B_o dan B_1 tidak dapat kita ketahui ,sehingga kita perlu melakukan estimasi dari sampel yang diobservasi, dengan menggunakan persamaan:

$\hat{Y}_i = a + b_{xi}$

dimana:

$\hat{Y}_i$ merupakan estimasi dari Y, a dan b merupakan estimator dari β_o dan β_1. Konstanta b biasanya mengacu koefisien regresi yang tidak terstandardisasi. Angka tersebut merupakan slope dari garis regresi dan menunjukkan ekspektasi perubahan pada Y saat X berubah sebesar 1 unit.

Catatan: Perlu diingat bahwa dalam regresi, kita pasti melihat korelasi, namun korelasi belum tentu regresi. Sebab dalam korelasi kita sudah menentukan variabel dependen dan independennya. Dalam korelasi, kita tidak membuat persamaan matematis, sedangkan dalam regresi kita mencoba membuat model persamaan regresi

Kasus :

Seorang produsen ingin membuat sebuah model regresi dengan variabel independen besarnya budget promosi yang dikeluarkan perusahaan, dan tingkat penjualan produk sebagai variabel dependennya. Berikut ini data budget promosi dan penjualan perusahaan tersebut selama 7 tahun belakangan ini

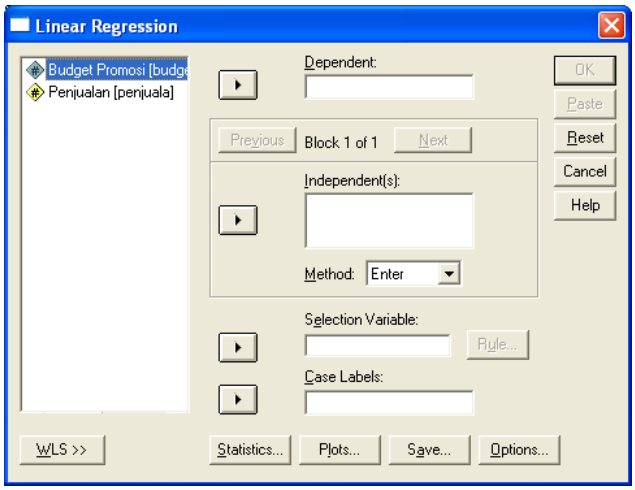
Tahun	Budget Promosi (dalam ratusan juta)	Penjualan (dalam ratusan juta)
1996	8	10
1997	7	8
1998	3	2
1999	5	6
2000	7	9
2001	2	2
2002	4	5

Pengolahan Data dengan SPSS

Langkah-langkah :

- Buka file **regresi bivariate**
- Dari menu utama SPSS pilih menu **Analyze**, kemudian pilih submenu Regression. LALU Klik linear…

Tampak di layar:



Pengisian:

- **Variable dependent** atau variabel terikat adalah penjualan, maka klik-lah penjualan lalu tekan tanda '>', maka variabel penjualan akan berpindah ke dependent.
- **Independent** atau variabel bebas. Pilih variabel budget promosi.
- **Method** atau cara memasukkan /seleksi variabel. Metode ini bermacam-macam seperti Stepwise, remove,backward, dan forward. Karena persamaan regresi ini masih bivariate, maka pilih saja default.

  Pilih kolom **Statistics** dengan mengklik mouse pada pilihan tersebut:

  Pilihan ini berkenaan dengan perhitungan statistic regresi yang akan digunakan. Perhatikan default yang ada di SPSS adalah Estimates dan Model Fit.

  Regression coefficient atau perlakuan koefisien regresi, pilih default atau estimates

  Klik mouse pada pilihan descriptives

  Residuals dikosongkan
- Klik continue
- Abaikan pilihan lain, lalu klik OK.

Interpretasi Output SPSS:

Descriptive Statistics

	Mean	Std. Deviation	N
Penjualan	6.0000	3.2146	7
Budget Promosi	5.1429	2.2678	7

Rata-rata penjualan dengan selama 7 tahun belakangan ini adalah sebesar 600 juta rupiah. Sedangkan mean budget promosi adalah sebesar 514,29 juta.

Correlations

		Penjualan	Budget Promosi
Pearson Correlation	Penjualan	1.000	.983
	Budget Promosi	.983	1.000
Sig. (1-tailed)	Penjualan	.	.000
	Budget Promosi	.000	.
N	Penjualan	7	7
	Budget Promosi	7	7

Korelasi antara budget promosi dengan penjulalan ternyata cukup signifikan (0,000 yang < dari 0,05) Jadi korelasi antara kedua variabel tersebut cukup kuat yaitu sebesar 0,983

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Budget Promosi[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Penjualan

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.983[a]	.966	.960	.6448

a. Predictors: (Constant), Budget Promosi

Standard error for the estimate adalah sebesar 0,6448, bandingkan dengan standar deviasi variabel penjualan pada tabel output pertama yang menunjukkan angka 3,2146. Karena Standard error for the estimates lebih kecil dari standar deviasi, maka model regresi ini lebih bagus untuk bertindak sebagai predictor penjualan dibandingkan rata-rata penjualan itu sendiri.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	59.921	1	59.921	144.131	.000[a]
	Residual	2.079	5	.416		
	Total	62.000	6			

a. Predictors: (Constant), Budget Promosi

b. Dependent Variable: Penjualan

Dari tabel Anova, terlihat bahwa F hitung yang kita peroleh adalah sebesar 144,131 dengan signifikansi sebesar 0,000 yang < dari 0,005. Berarti model regresi dapat dipakai untuk memprediksi variabel penjualan.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.167	.645		-1.809	.130
	Budget Promosi	1.394	.116	.983	12.005	.000

a. Dependent Variable: Penjualan

Tabel ini menunjukkan persamaan regresi
Y=-1,167+1,3946X

Dimana :
Y=Penjualan
X= Budget Promosi

Konstanta sebesar -1.167 menyatakan bahwa jika tidak ada budget promosi, maka penjualan tidak akan terjadi, karena nilainya – (minus)
Koefisien regresi sebesar 1,3964 menunjukkan bahwa setiap kita menaikkan budget promosi sebesar 100 juta rupiah (karena satuan budget dalam ratusan juta rupiah), maka penjualan akan meningkat sebesar 139,4 ratus juta rupiah.
Untuk menguji signifikansi konstanta kita dapat melihat hasil dari uji t atau dari nilai signifikansi-nya. Untuk memudahkan maka kita melihat nilai signifikansi-nya saja.

Pengambilan Keputusan:
jika probabilita > 0,05 ; maka Ho diterima
jika probabilita < 0,05 ; maka Ho ditolak

Terlihat bahwa probabilita dari promosi adalah sebesar 0,000 jauh dibawah 0,05. Ini berarti bahwa budget promosi memang berpengaruh secara signifikan terhadap penjualan.

B. REGRESI BERGANDA

Regresi berganda (*multiple regression)* adalah suatu teknik statistik yang secara simultan mengembangkan hubungan matematis antara dua atau lebih variabel independent dan sebuah variabel dependen.

Rumus umum dari persamaan regresi berganda:

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + e$$

yang diestimasi melalui persamaan berikut

$$\hat{Y}_i = a + b_{xi}$$

Seperti pada regresi bivariat (*bivariate regression)* maka koefisien menunjukkan intercept, namun nilai b pada regresi berganda (*multiple regression*) ini merupakan koefisien regresi yang bersifat parsial. Semua asumsi yang ada pada regresi bivariat (*bivariate regression)* juga diterapkan pada regresi berganda (*multiple regression*)

Kasus

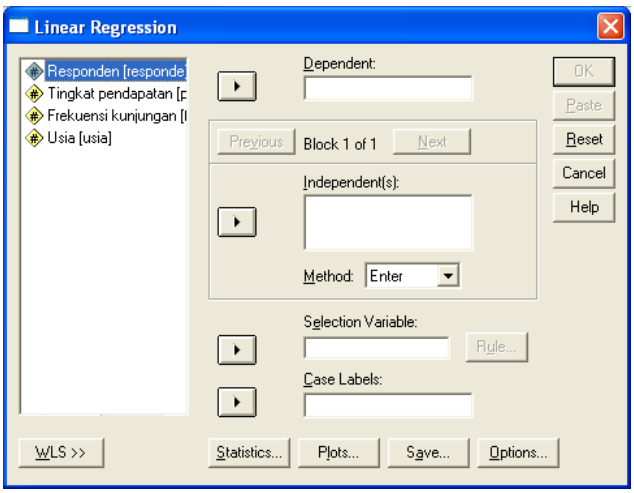
Untuk mengetahui pengaruh usia dan tingkat pendapatan terhadap frekuensi kunjungan ke mal, seorang marketer melakukan sebuah riset kecil. Berikut ini data dari masing-masing responden dari riset tersebut.

Responden	Tingkat pendapatan (dalam Juta)	Frekuensi kunjungan ke mal/bulan	Usia
1	4	6	21
2	2	3	34
3	3	2	39
4	2	1	33
5	1.5	2	24
6	7	2	27
7	10	4	29
8	2	6	29
9	2	4	37
10	3	3	22
11	5	8	56
12	1	3	58
13	6	3	26
14	2.7	2	45
15	4	4	40

Pengolahan Data dengan SPSS :

- Buka file regresi ganda
- Dari menu utama SPSS pilih menu Analyze, kemudian pilih submenu Regression. LALU Klik linear…

  Tampak di layar:



Pengisian:

- **Variable dependent** atau variabel terikat adalah frekuensi kunjungan ke mal, maka klik-lah frekuensi kunjungan ke mal lalu tekan tanda '>', maka variabel frekuensi kunjungan ke mal akan berpindah ke dependent.
- **Independent** atau variabel bebas. Pilih variabel usia dan variabel tingkat pendapatan
- **Method** atau cara memasukkan /seleksi variabel. Metode ini bermacam-macam seperti Stepwise, remove,backward, dan forward. Pilih saja enter yang berarti kita akan memasukkan kedua variabel independent secara bersamaan.
- Pilih kolom **Statistics** dengan mengklik mouse pada pilihan tersebut:

  Pilihan ini berkenaan dengan perhitungan statistic regresi yang akan digunakan. Perhatikan default yang ada di SPSS adalah Estimates dan Model Fit.

- **Regression coefficient** atau perlakuan koefisien regresi, pilih default atau estimates
- Klik mouse pada pilihan descriptives
- Residuals dikosongkan
- Klik continue
- Abaikan pilihan lain, lalu klik OK.

Interpretasi Output SPSS:

Descriptive Statistics

	Mean	Std. Deviation	N
Frekuensi kunjungan	3.5333	1.8848	15
Tingkat pendapatan	3.6800	2.4408	15
Usia	34.6667	11.4059	15

Mean frekuensi kunjungan ke mal adalah 3,5 kali dalam sebulan. Mean tingkat pendapatan adalah sebesar 3,6 juta. Sedangkan mean usia adalah sebesar 34, 6 tahun.

Correlations

		Frekuensi kunjungan	Tingkat pendapatan	Usia
Pearson Correlation	Frekuensi kunjungan	1.000	.176	.192
	Tingkat pendapatan	.176	1.000	-.231
	Usia	.192	-.231	1.000
Sig. (1-tailed)	Frekuensi kunjungan	.	.265	.247
	Tingkat pendapatan	.265	.	.204
	Usia	.247	.204	.
N	Frekuensi kunjungan	15	15	15
	Tingkat pendapatan	15	15	15
	Usia	15	15	15

Korelasi antara frekuensi kunjungan ke mal dengan tingkat pendapatan ternyata tidak cukup signifikan (0,265 yang > dari 0,05) Jadi korelasi antara kedua variabel tersebut tidak cukup kuat yaitu sebesar 0,265. Begitu pula hasil korelasi dari frekuensi kunjungan ke mal dengan usia.

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Usia, Tingkat pendapatan[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Frekuensi kunjungan

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.297[a]	.088	-.064	1.9440

a. Predictors: (Constant), Usia, Tingkat pendapatan

Model persamaan regresi yang diperoleh ternyata tidak cukup baik untuk menjelaskan frekuensi kunjungan ke mal, karena nilai R swuare hanya sebesar 0.088 atau hanya menjelaskan 8 persen dari frekunsi kunjungan ke mal.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.385	2	2.192	.580	.575[a]
	Residual	45.349	12	3.779		
	Total	49.733	14			

a. Predictors: (Constant), Usia, Tingkat pendapatan

b. Dependent Variable: Frekuensi kunjungan

Dari tabel Anova, terlihat bahwa F hitung yang kita peroleh adalah sebesar 0.580 dengan signifikansi sebesar 0,575 yang > dari 0,005. Berarti model regresi tidak dapat dipakai untuk memprediksi variabel frekuensi kunjungan ke mal.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.464	2.034		.720	.485
	Tingkat pendapatan	.180	.219	.233	.823	.427
	Usia	4.057E-02	.047	.245	.866	.403

a. Dependent Variable: Frekuensi kunjungan

Dari tabel di atas persamaan regresinya adalah Y=1,464+0,180X1+0,041X2.Artinya tanpa ada variabel tingkat pendapatan dan usia frekuensi kunjungan ke mal sebesar 1,464. Perubahan Rp 1 juta dalam tingkat pendapatan akan menaikkan frekuensi kunjungan ke mal sebesar 0,180 dan kenaikkan dalam usia akan menaikkan frekuensi kunjungan ke mal sebesar 0,041.

Probabilitas tingkat pendapatan dan usia 0,427 dan 0,403 atau di atas 0,05 , maka terima Ho atau tingkat pendapatan dan usia tidak berpengaruh secara signifikan dengan frekuensi kunjungan ke mall.

C. ANALISIS FAKTOR

Analisis faktor pada prinsipnya digunakan untuk mereduksi data, yaitu proses untuk meringkas sejumlah variabel menjadi lebih sedikit dan menamakannya sebagai faktor. Jadi, dapat saja lebih dari 10 atribut yang mempengaruhi sikap konsumen, setelah dilakukan analisis faktor, sebenarnya 10 atribut tersebut dapat di ringkas menjadi 3 faktor utama saja.

Secara garis besar, tahapan pada analisis faktor :

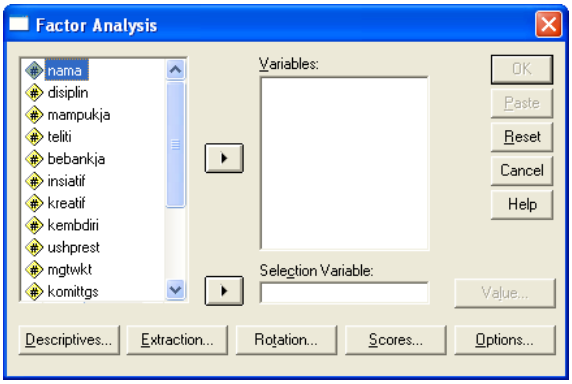
1. Memilih variabel yang layak dimasukkan dalam analisis faktor. Oleh karena analisis faktor berupaya mengelompokkan sejumlah variabel, maka seharusnya ada korelasi yang cukup kuat di antara variabel sehingga akan terjadi pengelompokkan. Jika sebuah variabel atau lebih berkorelasi lemah dengan variabel lainnya, maka variabel tersebut akan dikeluarkan dari analisis faktor. Dengan melihat contoh di atas, dari 15 variabel, mungkin saja dalam seleksi ada satu atau lebih variabel yang gugur. Alat seperti MSA atau Barlett's dapat digunakan untuk keperluan ini.
2. Setelah jumlah variabel terpilih, maka dilakukan 'ekstraksi' variabel tersebut hingga menjadi satu atau beberapa faktor. Beberapa metode pencarian faktor yang populer adalah Principal Component dan Maximum Likehood.
3. Faktor yang terbentuk, pada banyak kasus, kurang menggambarkan perbedaan diantara faktor-faktor yang ada. untuk itu, jika isi faktor masih diragukan, dapat dilakukan Proses Rotasi untuk memperjelas apakah faktor terbentuk sudah secara signifikansi berbeda dengan faktor lain.
4. Setelah faktor benar -benar sudah terbentuk, maka proses dilanjutkan dengan menamakan faktor yang ada, seperti pada contoh di atas. Kemudian beberapa langkah akhir juga perlu dilakukan, yaitu validasi hasil faktor.

Seperti telah disebut di depan, analisis faktor meliputi beberapa tahapan analisis, yaitu memilih variabel yang dapat dianalisis,kemudian menganalisis variabel terpilih, serta validasi hasil analisis.

a. Memilih Variabel

Langkah :

- buka software SPSS
- buka file **faktor analisis**
- dari menu utama SPSS, buka menu **Analyze,** submenu **Data Reduction,** kemudian pilih **factor.**
- Tampak dilayar :

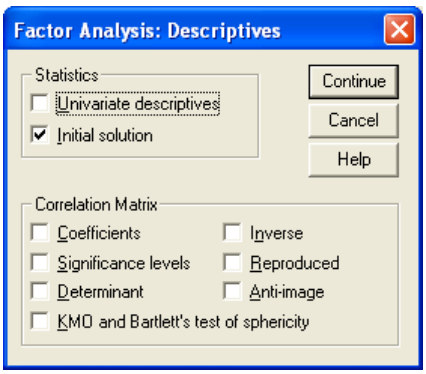


Pengisian :

- **Variables** atau variabel apa saja yang akan di proses. Oleh karena akan diuji semua variabel, masukkan SEMUA VARIABEL. Untuk itu, masukkan variabel **disiplin, kemampuan kerja, dst** ke kotak VARIABLES di sebelah kanan

Langkah praktis untuk memilih semua variabel, pilih variabel teratas, tekan tombol shift, lalu sambil menekan tombol Shift, pilih variabel terbawah. Otomatis semua variabel tersorot dan tinggal klik tanda > untuk memasukkan semua variabel tersebut.

- Buka icon **Descriptives,** hingga tampak di layar :



Pengisian :

- Untuk CORRELATION MATRIX , pilih (aktifkan) **KMO and Bartlett'stest of sphericity** dan **anti image.** Abaikan pilihan yang lain
- Untuk STATISTICS, biarkan pilihan pada default INITIAL SOLUTION

Tekan continue untuk kembali ke kotak dialog utama. Abaikan bagian yang lain dari kotak dialog Factor. Kemudian tekan OK untuk proses pengujian variabel.

OUTPUT

Simpan dalam file **factor analisis 1.**

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.873
Bartlett's Test of Sphericity	Approx. Chi-Square	984.164
	df	78
	Sig.	.000

- Pada tabel pertama KMO and Bartlett's test, terlihat angka KMO Measure of Sampling Adequacy (MSA) adalah 0,873 Oleh karena angka MSA diatas 0,5 maka kumpulan variabel tersebut dapat diproses lebih lanjut. Selanjutnya tiap variabel dianalisis untuk mengetahui mana yang dapat diproses lebih lanjut dan mana yang harus dikeluarkan. Kesimpulan yang sama dapat dilihat pula pada angka KMO and Bartlett's test (yang ditampakkan dengan angka Chi-Square) sebesar 984.164 dengan signifikansi 0,000.

Anti-image Matrices

		disiplin	mampukja	teliti	bebankja	insiatif	kreatif	kembdiri	ushprest	mgtwkt	komittgs	komitorg	comm	kjsama
Anti-image Covariance	disiplin	.733	.034	-.026	-.035	.026	-.032	-.025	-.019	-.122	-.081	-.097	-.032	-.018
	mampukja	.034	.567	-.102	-.024	-.113	-.062	.097	-.098	-.020	-.026	.018	-.013	-.091
	teliti	-.026	-.102	.673	.024	.049	-.074	-.028	.020	-.066	-.163	.008	-.125	.055
	bebankja	-.035	-.024	.024	.617	.005	-.027	-.057	-.068	-.111	-.094	-.054	-.060	.011
	insiatif	.026	-.113	.049	.005	.413	-.198	-.061	-.036	-.040	.010	-.095	-.005	-.018
	kreatif	-.032	-.062	-.074	-.027	-.198	.427	-.088	-.060	.018	.029	.028	-.021	.031
	kembdiri	-.025	.097	-.028	-.057	-.061	-.088	.548	-.182	-.023	-.034	.101	.017	-.005
	ushprest	-.019	-.098	.020	-.068	-.036	-.060	-.182	.417	-.046	.000	-.084	.073	-.091
	mgtwkt	-.122	-.020	-.066	-.111	-.040	.018	-.023	-.046	.608	-.089	-.073	.011	.033
	komittgs	-.081	-.026	-.163	-.094	.010	.029	-.034	.000	-.089	.512	-.107	-.039	-.112
	komitorg	-.097	.018	.008	-.054	-.095	.028	.101	-.084	-.073	-.107	.662	-.084	.057
	comm	-.032	-.013	-.125	-.060	-.005	-.021	.017	.073	.011	-.039	-.084	.616	-.258
	kjsama	-.018	-.091	.055	.011	-.018	.031	-.005	-.091	.033	-.112	.057	-.258	.594
Anti-image Correlation	disiplin	.925[a]	.053	-.036	-.052	.047	-.057	-.039	-.035	-.183	-.133	-.139	-.048	-.028
	mampukja	.053	.894[a]	-.165	-.041	-.233	-.125	.174	-.201	-.034	-.049	.029	-.022	-.157
	teliti	-.036	-.165	.857[a]	.038	.094	-.137	-.047	.038	-.103	-.279	.012	-.194	.087
	bebankja	-.052	-.041	.038	.937[a]	.009	-.053	-.097	-.135	-.181	-.167	-.085	-.097	.018
	insiatif	.047	-.233	.094	.009	.855[a]	-.472	-.128	-.086	-.081	.022	-.182	-.011	-.036
	kreatif	-.057	-.125	-.137	-.053	-.472	.858[a]	-.182	-.143	.035	.062	.052	-.040	.062
	kembdiri	-.039	.174	-.047	-.097	-.128	-.182	.845[a]	-.381	-.040	-.065	.168	.030	-.009
	ushprest	-.035	-.201	.038	-.135	-.086	-.143	-.381	.877[a]	-.091	.000	-.160	.143	-.184
	mgtwkt	-.183	-.034	-.103	-.181	-.081	.035	-.040	-.091	.923[a]	-.160	-.116	.018	.055
	komittgs	-.133	-.049	-.279	-.167	.022	.062	-.065	.000	-.160	.890[a]	-.184	-.069	-.204
	komitorg	-.139	.029	.012	-.085	-.182	.052	.168	-.160	-.116	-.184	.870[a]	-.131	.091
	comm	-.048	-.022	-.194	-.097	-.011	-.040	.030	.143	.018	-.069	-.131	.802[a]	-.427
	kjsama	-.028	-.157	.087	.018	-.036	.062	-.009	-.184	.055	-.204	.091	-.427	.805[a]

a. Measures of Sampling Adequacy(MSA)

- Pada tabel kedua (Anti Image Matrices), khususnya pada bagian bawah (Anti Image Correlation), terlihat sejumlah angka yang membentuk diagonal, yang bertanda 'a' menandakan besaran MSA sebuah variabel. Seperti layout yang mempunyai MSA variabel. Seperti variabel disiplin yang mempunyai MSA 0,925 kemudian variabel kemampuan kerja dengan MSA sebesar 0,894 dan seterusnya.

Pedoman :

- Apakah ada angka MSA yang di bawah 0,5? Ternyata tidak ada untuk itu kita dapat meneruskan analisis ke tabel berikutnya

Communalities

	Initial	Extraction
disiplin	1.000	.531
mampukja	1.000	.567
teliti	1.000	.399
bebankja	1.000	.503
insiatif	1.000	.709
kreatif	1.000	.715
kembdiri	1.000	.570
ushprest	1.000	.693
mgtwkt	1.000	.606
komittgs	1.000	.648
komitorg	1.000	.472
comm	1.000	.706
kjsama	1.000	.668

Extraction Method: Principal Component Analysis.

Untuk variabel DISIPLIN, angka adalah 0,531 Hal ini berarti sekitar 53,1% varians dari variabel Disiplin dapat dijelaskan oleh faktor yang nanti terbentuk (jika dilihat pada tabel terakhir, yaitu komponent matrix, ada 2 component yang berarti ada 3 faktor terbentuk)

Untuk variabel KEMAMPUAN KERJA, angka adalah 0,567. Hal ini berarti sekitar 56,7% varians dari variabel Kemampuan kerja dapat dijelaskan oleh faktor yang nanti terbentuk. Demikian seterusnya, dengan ketentuan bahwa semakin kecil communalities sebuah variabel berarti semakin lemah hubungannya dengan faktor terbentuk.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.238	40.293	40.293	5.238	40.293	40.293
2	1.463	11.253	51.546	1.463	11.253	51.546
3	1.085	8.345	59.891	1.085	8.345	59.891
4	.799	6.148	66.039			
5	.782	6.015	72.054			
6	.655	5.041	77.094			
7	.593	4.563	81.658			
8	.540	4.151	85.809			
9	.499	3.836	89.645			
10	.420	3.233	92.878			
11	.358	2.756	95.634			
12	.300	2.311	97.945			
13	.267	2.055	100.000			

Nilai eigen value masih di atas 1

Extraction Method: Principal Component Analysis.

Dari tabel di atas terlihat bahwa hanya tiga faktor yang terbentuk, karena dengan satu faktor, angka eigenvalues di atas 1 dengan dua ataupun tiga factor, angka eigenvalues juga masih diatas 1, namun untuk 4 faktor angka eigen values sudah dibawah 1.

Component Matrix[a]

	Component		
	1	2	3
disiplin	.534	.264	-.420
mampukja	.663	-.126	.334
teliti	.550	.303	.065
bebankja	.671	.042	-.226
insiatif	.714	-.427	.127
kreatif	.697	-.461	.128
kembdiri	.606	-.444	-.072
ushprest	.756	-.347	-.022
mgtwkt	.657	.112	-.403
komittgs	.695	.395	-.095
komitorg	.575	.221	-.304
comm	.516	.517	.416
kjsama	.562	.293	.517

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

Setelah diketahui bawa tiga faktor adalah jumlah yang paling optimal, maka tabel ini menunjukkan distribusi ketigabelas variabel tersebut pada tiga faktor yang ada. Sedangkan angka yang ada pada tabel tersebut adalah faktor loadings atau besar korelasi antara suatu variabel dengan faktor 1, faktor 2 dan faktor 3.

Seperti pada variabel Disiplin, korelasi antar variable Disiplin dengan faktor 1 adalah 0,534 (cukup kuat) sedang korelasi variabel Disiplin dengan faktor 2 adalah 0,264 (lemah) dan korelasi dengan faktor 3 adalah sebesar -0,420 yang lebih kecil dari

korelasi dengan faktor 1. Dengan demikian dapat dikatakan variabel Disiplin dapat dimasukkan sebagai komponen faktor 1.

Secara ringkas dapat kita katakan bahwa pengelompokkan faktor tersebut adalah sebagai berikut

Faktor 1	1, 2, 3, 4, 5, 6, 7,8,9,10,11,13
Faktor 2	12
Faktor 3	?

Jika kita menggunakan output diatas, maka pngelompokkan faktor menjadi kurang jelas, untuk itu, kita perlu melakukan rotasi atas faktor yang ada.

ANALISIS FAKTOR DENGAN ROTATION

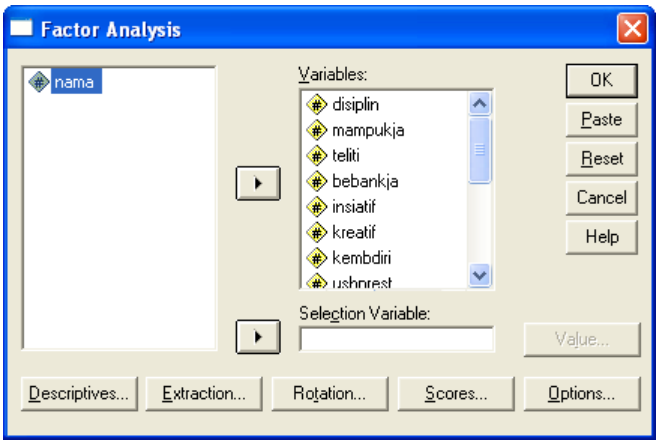
Analisis faktor di atas diperluas dengan proses rotasi

Langkah :

a. Buka file **faktor analisis** (atau jika sudah terbuka, lakukan prosedur selanjutnya)

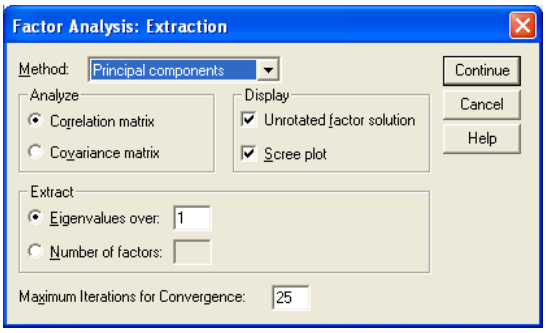
b. Dari menu utama SPSS pilih menu **Analyze**, lalu submenu **Reduction,** kemudian pilih **Factor.....**

c. Tampak di layar :



Pengisian :

- masukkan variabel **semua variable kecuali nama** ke dalam kotak variables. Ke enam variabel ini akan dilakukan proses faktorisasi.
- klik icon **EXTRACTION**. Tampak di layar



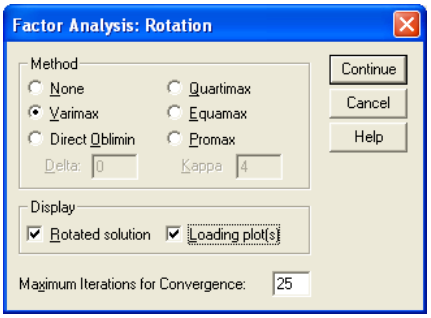
pengisian :

- untuk METHOD, pilih **Principal Component**
- untuk DISPLAY, selain default **UnRotated Solution** aktifkan juga pilihan **Scree Plot**

Abaikan bagian lain dan tekan continue untuk kembali ke kotak dialog utama

Abaikan icon lain dan tekan OK untuk proses data.

- Klik icon Rotation. Tampak dilayar :



Pengisian :

- untuk METHOD, pilih **Varimax**
- untuk DISPLAY, selain default **Rotated Solution**, aktifkan juga pilihan **Loadings Plot**

Abaikan bagian lain dan tekan continu untuk kembali ke kotak dialog utama

Abaikan icon lain dan tekan OK untuk proses data.

OUTPUT

Secara umum, output yang dikeluarkan oleh SPSS adalah sama dengan proses sebelumnya. Output yang berbeda dan cukup penting adalah tabel Rotated component matrix.

Rotated Component Matrix[a]

	Component		
	1	2	3
disiplin	.095	.716	.096
mampukja	.580	.113	.467
teliti	.149	.402	.463
bebankja	.378	.575	.172
insiatif	.804	.170	.180
kreatif	.819	.145	.154
kembdiri	.716	.238	-.018
ushprest	.751	.331	.139
mgtwkt	.290	.718	.080
komittgs	.153	.637	.468
komitorg	.171	.643	.172
comm	.021	.220	.811
kjsama	.231	.090	.779

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Dari table diatas, maka kita dapat melihat pengelompokkan faktor sebagai berikut:

Faktor 1	2,5,6,7,8
Faktor 2	1,4,9,,10,11,
Faktor 3	3,12,13

Sekarang kita sudah dapat melihat adanya pengelompokkan yang lebih jelas.

MENAMAKAN FAKTOR

Setelah didapat tiga faktor, langkah berikut adalah memberi nama pada ketiga faktor tersebut. Tentu saja penamaan faktor ini bergantung pada nama-nama variabel yang menjadi satu kelompok, pada intepretasi masing-masing analisis dan aspek lainnya. Sehingga sebenarnya pemberian nama bersifat subjektif serta tidak ada ketentuan yang pasti mengenai pemberian nama tersebut.

D. ANALISIS CLUSTER

Cluster Analysis adalah teknik yang digunakan untuk mengklasifikasikan objek kedalam kelompok yang relatif homogen yang disebut cluster. Dimana Objek dalam tiap cluster cenderung memiliki kemiripan satu dengan lainnya. Cluster Analysis pada prinsipnya digunakan untuk mereduksi data, yaitu meringkas sejumlah variabel menjadi lebih sedikit dan menamakannya sebagai cluster. Cluster analysis dapat digunakan untuk riset segmentasi pasar, dimana konsumen dikelompokkan ke dalam suatu cluster berdasarkan variabel-variabel tertentu.

Analisis Cluster dapat dibagi menjadi dua jenis, yaitu Hierarchical Cluster dan K-Means cluster. Pengelompokan secara hierarki biasanya digunakan untuk jumlah sampel yang relatif

sedikit. Sedangkan untuk data yang banyak dapat digunakan K-Means Cluster (lebih populer digunakan).

Pada modul ini, penulis hanya akan membahas mengenai Hierarchical cluster.

A. HIERARCHICAL CLUSTER

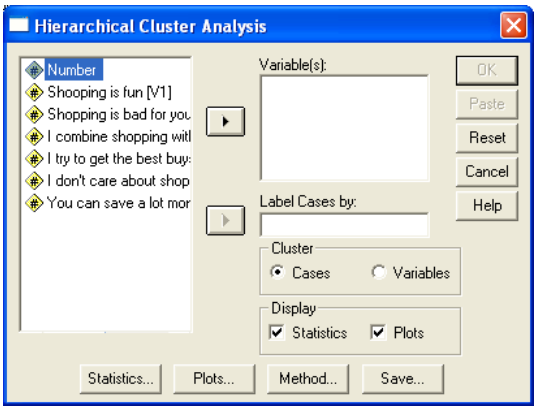
Cluster hierarki lebih berupaya mengelompokkan cases (individu/responden) berdasar kemiripan yang ada pada persepsi mereka, untuk jumlah responden tertentu yang jumlahnya sedikit. Sehingga boleh dikatakan bahwa K-Means cluster lebih efektif untuk membuat cluster pada case yang banyak. Cluster Hierarki akan melakukan proses membandingkan setiap pasang case, sehingga akan menyulitkan bila jumlah case-nya banyak.

Kasus

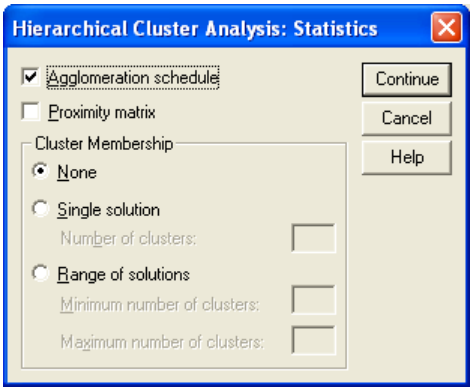
Seorang pemilik minimarket mengadakan riset kecil kepada 20 orang pelanggan setia yang berbelanja di tokonya. Ia ingin mengetahui pengelompokkan pelanggan tersebut berdasarkan sikapnya dalam berbelanja. Untuk menjawab pertanyaan tersebut, maka dilakukan suatu cluster analysis atas data yang diperolehnnya melalui penelitian tersebut.

Langkah-langkahnya:

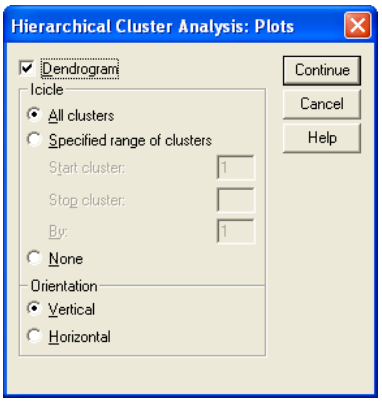
- buka file **cluster**
- dari menu utama SPSS, pilih menu **analyze**, lalu submenu **classify**, kemudian pilih **hierarchical cluster**



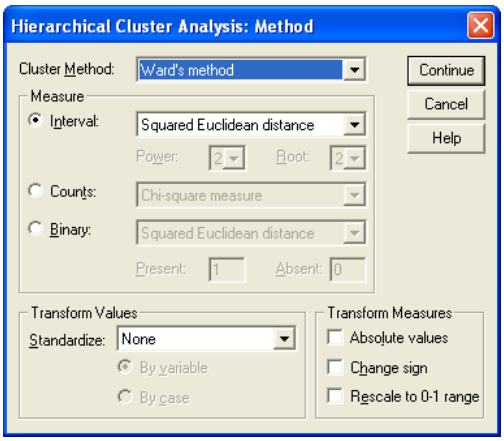
- pada kotak dialog yang muncul, masukkan keenam variabel tersebut yang ada.
- klik tombol "statistics" , lalu pada kotak dialog yang timbul aktifkan pilihan range of solution dan isi from dengan 2 dan through dengan 4 yang berarti kemungkinan akan dibuat 2 hingga 4 cluster, klik tombol "continue" untuk kembali ke kotak dialog utama.



- Klik tombol "plots", lalu pada kotak dialog yang timbul aktifkan kotak dendogram untuk memperlihatkan terjadinya proses cluster secara grafis,. Sedangkan untuk bagian "icicle" pilih none yang berarti tidak ada icicle yang diperlihatkan dalam output. Lalu klik tombol "continue" untuk kembali ke kotak dialog utama.



- Klik tombol "methods", lalu pada kotak dialog yang muncul pilihlah ward's method pada pilihan cluster method. Hal ini berarti kita akan menggunakan metode ward's. Lalu klik tombol "continue" untuk kembali ke kotak dialog utama.



- Klik tombol "OK" untuk proses pengolahan data

OUTPUT

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
20	95.2	1	4.8	21	100.0

a. Squared Euclidean Distance used

b. Ward Linkage

Agglomeration Schedule

	Cluster Combined			Stage Cluster First Appears		
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
1	14	16	1.000	0	0	6
2	6	7	2.000	0	0	7
3	2	13	3.500	0	0	15
4	5	11	5.000	0	0	11
5	3	8	6.500	0	0	16
6	10	14	8.167	0	1	9
7	6	12	10.500	2	0	10
8	9	20	13.000	0	0	11
9	4	10	15.583	0	6	12
10	1	6	18.500	0	7	13
11	5	9	23.000	4	8	15
12	4	19	27.750	9	0	17
13	1	17	33.100	10	0	14
14	1	15	41.333	13	0	16
15	2	5	51.833	3	11	18
16	1	3	64.500	14	5	19
17	4	18	79.667	12	0	18
18	2	4	172.667	15	17	19
19	1	2	328.600	16	18	0

Pada stage 1 responden 14 dan 16 yang paling mirip maka keduanya masuk ke dalam satu kelompok. Next stage 6 berarti proses dilanjutkan ke proses 6 dimana responden ke 10 masuk pada kelompok yang sudah terbentuk yaitu 14 dan 16. Next stage 9 berarti proses selanjutnya di stage 9 dimana responden 4 masuk ke dalam kelompok yang sudah terbentuk yaitu 10 dan 14. Next stage 12 berarti proses selanjutnya di stage 12 dimana responden 19 masuk ke kelompok yang sudah terbentuk yaitu 4 dan 10. Proses selanjutnya di stage 17 dimana responden 18 masuk pada kelompok yang sudah terbentuk sebelumnya. Proses selanjutnya di stage 18 dimana responden 2 masuk ke kelompok yang sudah terbentuk sebelumnya. Proses selanjutnya di stage 19 dimana responden 1 masuk ke dalam kelompok yang sudah terbentuk sebelumnya.

Proses ini terus dilanjutkan pada cluster berikutnya

Cluster Membership

Case	4 Clusters	3 Clusters	2 Clusters
1	1	1	1
2	2	2	2
3	1	1	1
4	3	3	2
5	2	2	2
6	1	1	1
7	1	1	1
8	1	1	1
9	2	2	2
10	3	3	2
11	2	2	2
12	1	1	1
13	2	2	2
14	3	3	2
15	1	1	1
16	3	3	2
17	1	1	1
18	4	3	2
19	3	3	2
20	2	2	2

Seperti pada proses input cluster, akan dibuat 2 sampai 4 cluster. Tabel cluster membership secara praktis memberi informasi anggota yang ada pada tiap cluster jika dibuat 2,3, dan 4 cluster.

Jika dibuat 4 cluster, maka terlihat anggota cluster 1 adalah responden 1,3,6,7,8,12,15,17. Demikian seterusnya, jika akan dibuat 3 cluster dan 4 cluster, maka setiap grosir akan dimasukkan

```
* * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S
* * * * * *


 Dendrogram using Average Linkage (Between Groups)

                         Rescaled Distance Cluster Combine

   C A S E    0         5        10        15        20        25
  Label  Num  +---------+---------+---------+---------+---------+

         14   ⇩⇘
         16   ⇩⇧⇩⇘
         10   ⇩⇙ ▭⇩⇘
          4   ⇩⇩⇩⇙ ▭⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇘
         19   ⇩⇩⇩⇩⇩⇙          ▭⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇘
         18   ⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇙              ⇔
          2   ⇩×⇩⇩⇩⇩⇩⇩⇩⇘
▭⇩⇩⇩⇩⇩⇩⇩⇩⇩⇘
         13   ⇩⇙      ⇔                       ⇔       ⇔
          5   ⇩×⇩⇘     ▭⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇙
⇔
         11   ⇩⇙ ▭⇩⇘    ⇔                              ⇔
          9   ⇩⇩⇩⇙ ▭⇩⇩⇩⇙                               ⇔
         20   ⇩⇩⇩⇩⇩⇙                                   ⇔
          3   ⇩×⇩⇩⇩⇩⇩⇩⇩⇩⇩⇘
⇔
          8   ⇩⇙        ⇔                              ⇔
          6   ⇩×⇩⇘       ▭⇩⇘                           ⇔
          7   ⇩⇙ ⇔       ⇔ ⇔                           ⇔
         12   ⇩⇩⇩✓⇩⇩⇩⇘  ⇔
▭⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇙
          1   ⇩⇩⇩⇙   ▭⇩⇩⇩⇙ ⇔
         17   ⇩⇩⇩⇩⇩⇩⇩⇙     ⇔
         15   ⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇩⇙
```

Dendogram

Hasil proses agglomerasi di atas dapat juga ditampilkan dengan sebuah dendogram. Terlihat sesuai dengan proses agglomerasi, responden 14 bergabung terlebih dengan responden 16 menjadi 1 cluster terlebih dahulu, kemudian responden 10 bergabung dengan cluster tadi. Demikian seterusnya sampai terbentuk sebuah cluster besar, dimana semua case (responden) sudah termasuk di dalamnya.

E. ANALISIS DISKRIMINAN

Teknik untuk menganalisis data yang memiliki variabel dependen dalam bentuk *kategori* dan variabel independen dalam bentuk *metric*

Tujuan Discriminant Analysis

- Menganalisis apakah terdapat perbedaan yang cukup signifikan antar kelompok dalam hal variabel independen
- Penentuan variabel mana yang memberikan kontribusi terbesar terhadap perbedaan yang terjadi antar kelompok
- Klasifikasi setiap kasus ke dalam satu kelompok berdasarkan nilai dari prediktor
- Evaluasi terhadap akurasi klasifikasi

A. TEKNIK DISCRIMINANT ANALYSIS

Berdasarkan jumlah kategori dalam dependent variabel, maka analisis diskriminan dibagi menjadi:

- *Two-Group Discriminant Analysis*
  -DV terdiri dari 2 kategori
  -Hanya dapat menurunkan 1 fungsi diskriminan
- *Multiple Discriminant Analysis*
  -DV terdiri dari >2 kategori
  ->1 fungsi diskriminan

B. MODEL DISCRIMINANT ANALYSIS

Bentuk Kombinasi linier:

D=bo+bX+b2X2+b3X3+…+bkXk

D= Skor Diskriminan
b= Koefisien diskriminasi atau bobot
X= Predictor atau Variabel Independen

Kasus

Seorang pengelola resort ingin mengetahui karakter-karakter dari keluarga yang berlibur dengan mereka yang tidak berlibur ke pantai selama dua tahun terakhir ini.
Berikut ini data responden yang dimiliki oleh pengelola resort tersebut.

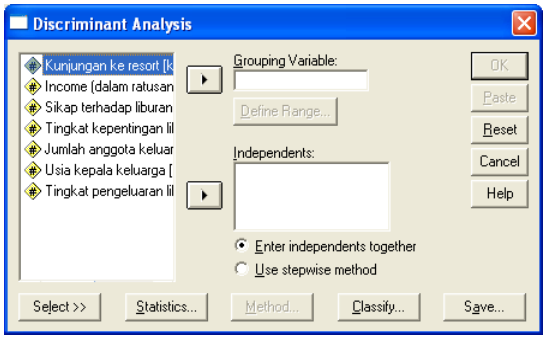
No	Kunjungan ke resort	Income (dalam ratusan ribu)	Sikap thd liburan	Tingkat kepentingan liburan	Jumlah anggota keluarga	Usia kepala keluarga	Tingkat pengeluaran liburan
1	Tidak	50.2	5	8	3	43	M (2)
2	Tidak	70.3	6	7	4	61	H (3)
3	Tidak	62.9	7	5	6	52	H (3)
4	Tidak	48.5	7	5	5	36	L (1)
5	Tidak	52.7	6	6	4	55	H (3)
6	Tidak	75	8	7	5	68	H (3)
7	Tidak	46.2	5	3	3	62	M (2)
8	Tidak	57	2	4	6	51	M (2)
9	Tidak	64.1	7	5	4	57	H (3)
10	Tidak	68.1	7	6	5	45	H (3)
11	Tidak	73.4	6	7	5	44	H (3)
12	Tidak	71.9	5	8	4	64	H (3)
13	Tidak	56.2	1	8	6	54	M (2)
14	Tidak	49.3	4	2	3	56	H (3)
15	Tidak	62	5	6	2	58	H (3)
16	Ya	32.1	5	4	3	58	L (1)
17	Ya	36.2	4	3	2	55	L (1)
18	Ya	43.2	2	5	2	57	M (2)
19	Ya	50.4	5	2	4	37	M (2)
20	Ya	44.1	6	6	3	42	M (2)
21	Ya	38.3	6	6	2	45	L (1)
22	Ya	55	1	2	2	57	M (2)
23	Ya	46.1	3	5	3	51	L (1)
24	Ya	35	6	4	5	64	L (1)
25	Ya	37.3	2	7	4	54	L (1)
26	Ya	41.8	5	1	3	56	M (2)
27	Ya	57	8	3	2	36	M (2)
28	Ya	33.4	6	8	2	50	L (1)
29	Ya	37.5	3	2	3	48	L (1)
30	Ya	41.3	3	3	2	42	L (1)

Pengolahan data dengan SPSS

Langkah-langkah :

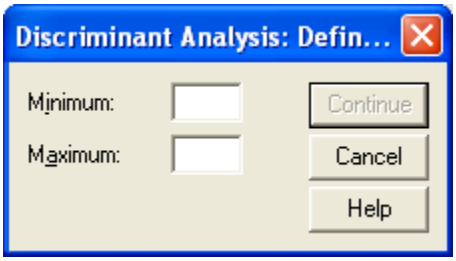
- Buka file **analisis diskriminan**
- Dari menu utama SPSS pilih menu Analyze, kemudian pilih submenu Classify kemudian pilih Discriminant ...

  Tampak di layar:



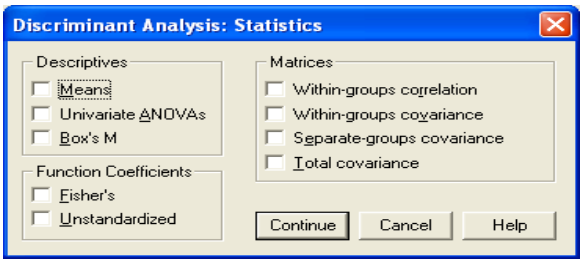
Pengisian:

- Masukkan variabel pengeluaran liburan ke bagian GROUPING VARIABEL
- Kemudian buka define range…, hingga tampak di layar:



Masukkan angka 1 pada angka minimum dan 3 pada angka maksimum.

- Tekan: Continue untuk kembali ke kotak dialog utama.
- Masukkan variabel kunjungan, income, sikap terhadap liburan, tingkat kepentingan liburan, jumlah angota keluarga, dan usia kepala keluarga ke dalam kotak INDEPENDENT.
- Klik Icon STATISTICS . Tampak di layar:



Pada bagian DESCRIPTIVES, aktifkan pilihan Means dan Univariate ANOVA's

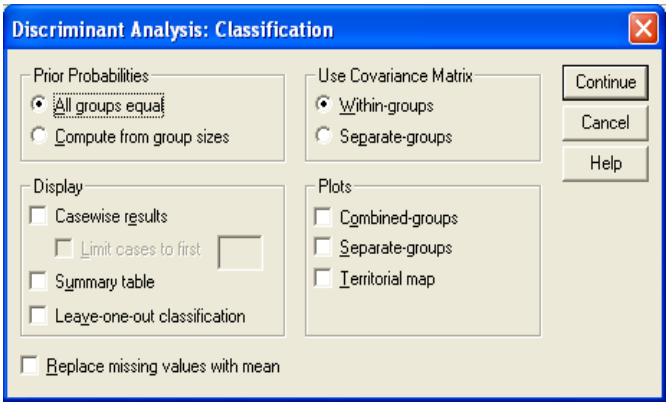
Pada bagian FUNCTION COEFFICIENT, aktifkan pilihan Unstandardized.

- Abaikan bagian lain dan tekan Continue untuk kembali ke kotak dialog utama.
- Kemudian klik mouse pada pilihan Use Stepwise method (yang terletak di bagian bawah), maka secara otomatis icon METHOD akan terbuka aktif.

  Pengisian:

  Pada bagian METHOD, pilih Mahalanobis distance

  Pada bagian CRITERIA, pilih Use probability of F, namun jangan mengubah isi yang sudah ada (default).
- Abaikan bagian lain, kemudian tekan continue untuk kembali ke kotak dialog utama.
- Klik Icon CLASSIFY. Tampak di layar:



Pengisian:

- Pada bagian DISPLAY, aktifkan pilihan Casewise results.
- Masih pada bagian DISPLAY, aktifkan pilihan Leave-oneout-classification.

- Abaikan bagian lain, lalu tekan Continue untuk kembali ke kotak dialog utama.
- Tekan OK untuk proses data.

OUTPUT

Interpretasi Output SPSS:

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		30	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		30	100.0

Tabel ini menyatakan bahwa responden (jumlah kasus atau baris SPSS) semuanya valid (sah untuk diproses). Oleh karena itu, tidak ada satu pun data yang hilang (*missing*).

Group Statistics

Tingkat pengeluaran liburan		Mean	Std. Deviation	Valid N (listwise) Unweighted	Weighted
Low	Kunjungan ke resort	1.9000	.3162	10	10.000
	Income (dalam ratusan ribu)	38.5700	5.2972	10	10.000
	Sikap terhadap liburan	4.5000	1.7159	10	10.000
	Tingkat kepentingan liburan	4.7000	1.8886	10	10.000
	Jumlah anggota keluarga	3.1000	1.1972	10	10.000
	Usia kepala keluarga	50.3000	8.0973	10	10.000
Medium	Kunjungan ke resort	1.6000	.5164	10	10.000
	Income (dalam ratusan ribu)	50.1100	6.0023	10	10.000
	Sikap terhadap liburan	4.0000	2.3570	10	10.000
	Tingkat kepentingan liburan	4.2000	2.4855	10	10.000
	Jumlah anggota keluarga	3.4000	1.5055	10	10.000
	Usia kepala keluarga	49.5000	9.2526	10	10.000
High	Kunjungan ke resort	1.0000	.0000	10	10.000
	Income (dalam ratusan ribu)	64.9700	8.6143	10	10.000
	Sikap terhadap liburan	6.1000	1.1972	10	10.000
	Tingkat kepentingan liburan	5.9000	1.6633	10	10.000
	Jumlah anggota keluarga	4.2000	1.1353	10	10.000
	Usia kepala keluarga	56.0000	7.6012	10	10.000
Total	Kunjungan ke resort	1.5000	.5085	30	30.000
	Income (dalam ratusan ribu)	51.2167	12.7952	30	30.000
	Sikap terhadap liburan	4.8667	1.9780	30	30.000
	Tingkat kepentingan liburan	4.9333	2.0998	30	30.000
	Jumlah anggota keluarga	3.5667	1.3309	30	30.000
	Usia kepala keluarga	51.9333	8.5740	30	30.000

Dari data diatas, terlihat bahwa dalam masing-masing kelompok pengeluaran untuk liburan terdiri dari 10 orang kelompok.

Dari kolom income, dapat kita ketahui bahwa mean dari pendapatan keluarga yang jarang berlibur adalah sebesar 38.57 (dalam ratusan ribu)

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Kunjungan ke resort	.440	17.182	2	27	.000
Income (dalam ratusan ribu)	.262	37.997	2	27	.000
Sikap terhadap liburan	.788	3.634	2	27	.040
Tingkat kepentingan liburan	.881	1.830	2	27	.180
Jumlah anggota keluarga	.874	1.944	2	27	.163
Usia kepala keluarga	.882	1.804	2	27	.184

Jika Sig.> 0,05 berarti tidak ada perbedaan antara tiga kelompok

Jika Sig.<0,05 berarti ada perbedaan antara tiga kelompok

Contoh analisis:

Variabel KUNJUNGAN KE RESORT, nilai signifikansi adalah sebesar 0,000 <0,05. Berarti ada perbedaan antara ketiga kelompok dalam hal kunjungan ke resort.

Variabel INCOME, nilai signifikansi adalah sebesar 0,000 <0,05. Berarti ada perbedaan antara ketiga kelompok dalam hal kunjungan ke resort.

Variabel TINGKAT KEPENTINGAN LIBURAN,nilai signifikansi adalah sebesar 0,180 >0,05. Berarti tidak ada perbedaan antara ketiga kelompok dalam hal tingkat kepentingan liburan.

Variables Entered/Removed[a,b,c,d]

Step	Entered	Min. D Squared					
		Statistic	Between Groups	Exact F			
				Statistic	df1	df2	Sig.
1	Income (dalam ratusan ribu)	2.889	Low and Medium	14.444	1	27.000	7.483E-04

At each step, the variable that maximizes the Mahalanobis distance between the two closest groups is entered.

a. Maximum number of steps is 12.

b. Maximum significance of F to enter is .05.

c. Minimum significance of F to remove is .10.

d. F level, tolerance, or VIN insufficient for further computation.

Variables in the Analysis

Step		Tolerance	Sig. of F to Remove
1	Income (dalam ratusan ribu)	1.000	.000

Tabel ini menunjukkan variabel mana yang akan dimasukkan ke dalam persamaan Diskriminan. Ternyata dari 6 variabel yang ada, hanya 1 variabel saja yang dimasukkan ke dalam persamaan yakni Income.

Variables Not in the Analysis

Step		Tolerance	Min. Tolerance	Sig. of F to Enter	Min. D Squared	Between Groups
0	Kunjungan ke resort	1.000	1.000	.000	.736	Low and Medium
	Income (dalam ratusan ribu)	1.000	1.000	.000	2.889	Low and Medium
	Sikap terhadap liburan	1.000	1.000	.040	.076	Low and Medium
	Tingkat kepentingan liburan	1.000	1.000	.180	.060	Low and Medium
	Jumlah anggota keluarga	1.000	1.000	.163	.054	Low and Medium
	Usia kepala keluarga	1.000	1.000	.184	.009	Low and Medium
1	Kunjungan ke resort	.911	.911	.204	3.025	Low and Medium
	Sikap terhadap liburan	.997	.997	.196	3.020	Low and Medium
	Tingkat kepentingan liburan	.906	.906	.345	3.537	Low and Medium
	Jumlah anggota keluarga	.855	.855	.705	3.089	Low and Medium
	Usia kepala keluarga	.956	.956	.146	2.960	Low and Medium

Melalui Tabel ini, kita ingin menguji variabel mana yang tidak dimasukkan dalam analisis. Namun untuk itu, pada langkah pertama, semua variabel dimasukkan terlebih dahulu ke dalam persamaan. Pada langkah ke dua, tenyata kelima variabel yang lain yakni (kunjungan ke resort, sikap terhadap liburan, tingkat kepentingan liburan, jumlah anggota keluarga, dan usia kepala keluarga) langsung dikeluarkan dari persamaan Diskriminan.

Wilks' Lambda

Step	Number of Variables	Lambda	df1	df2	df3	Exact F			
						Statistic	df1	df2	Sig.
1	1	.262	1	2	27	37.997	2	27.000	1.414E-08

Pairwise Group Comparisons[a]

Step	Tingkat pengeluaran		Low	Medium	High
1	Low	F		14.444	75.595
		Sig.		.001	.000
	Medium	F	14.444		23.951
		Sig.	.001		.000
	High	F	75.595	23.951	
		Sig.	.000	.000	

a. 1, 27 degrees of freedom for step 1.

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.815[a]	100.0	100.0	.859

a. First 1 canonical discriminant functions were used in the analysis.

Melalui tabel Eigenvalues, kita dapat melihat bahwa hanya ada satu faktor yang dapat menjelskan varians dari variabel tingkat pengeluaran liburan, yakni INCOME.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.262	36.148	2	.000

Tabel Eigenvalues dan Wilk's Lambda pada dasarnya dapat saling melengkapi. Pada tabel ini terlihat bahwa nilai Chi-square dari persamaan Diskriminan ini adalah sebesar 36.148 dengan signifikansi sebesar 0.000. Berarti ada perbedaan yang sangat nyata antara ketiga kelompok dalam hal tingkat pengeluaran liburan. Hal tersebut ternyata dapat dijelaskan secara signifikan melalui Income saja.

andardized Canonical Discriminant Function Coefficient

	Function
	1
Income (dalam ratusan ribu)	1.000

Structure Matrix

	Function
	1
Income (dalam ratusan ribu)	1.000
Jumlah anggota keluarga[a]	.380
Tingkat kepentingan liburan[a]	.307
Kunjungan ke resort[a]	-.298
Usia kepala keluarga[a]	-.209
Sikap terhadap liburan[a]	.051

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within functio

a. This variable not used in the analysis.

Pada tabel, ke-enam variabel dimasukkan ke dalam fungsi persamaan Diskriminan. Namun lihat tanda'a' pada setiap variabel. Tanda tersebut menunjukkan bahwa variabel tersebut akhirnya tidak digunakan dalam analisis.

anonical Discriminant Function Coefficient

	Function
	1
Income (dalam ratusan ribu)	.147
(Constant)	-7.543

Unstandardized coefficients

Functions at Group Centroids

Tingkat pengeluaran liburan	Function
	1
Low	-1.863
Medium	-.163
High	2.026

Unstandardized canonical discriminant functions evaluated at group means

Tabel ini menunjukkan bahwa variabel tingkat pengeluaran liburan baik untuk kelompok rendah, medium, dan tinggi dapat dijelaskan hanya melalui INCOME.

Tanda negative pada kelompk rendah dan medium, menunjukkan bahwa kelompok tersebut memiliki Income yang rendah. Sedangkan tanda positif pada kelompok tinggi, menunjukkan bahwa mereka memiliki income yang tinggi.

Classification Processing Summary

Processed		30
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		30

Prior Probabilities for Groups

Tingkat pengeluaran liburan	Prior	Cases Used in Analysis	
		Unweighted	Weighted
Low	.333	10	10.000
Medium	.333	10	10.000
High	.333	10	10.000
Total	1.000	30	30.000

Classification Results[b,c]

		Tingkat pengeluaran liburan	Predicted Group Membership			Total
			Low	Medium	High	
Original	Count	Low	8	2	0	10
		Medium	3	7	0	10
		High	0	2	8	10
	%	Low	80.0	20.0	.0	100.0
		Medium	30.0	70.0	.0	100.0
		High	.0	20.0	80.0	100.0
Cross-validated[a]	Count	Low	8	2	0	10
		Medium	3	7	0	10
		High	0	2	8	10
	%	Low	80.0	20.0	.0	100.0
		Medium	30.0	70.0	.0	100.0
		High	.0	20.0	80.0	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 76.7% of original grouped cases correctly classified.

c. 76.7% of cross-validated grouped cases correctly classified.

Tabel ini menunjukkan bahwa ketepatan prediksi dari model persamaan Diskriminan ini adalah sebesar 76,7%. Angka yang diperoleh setelah dilakukan validasi ulang juga tetap sama. Hal ini berarti bahwa pengelompokkan yang kita lakukan sudah tepat.

LATIHAN

1. Seorang eksekutif pemasaran berusaha untuk meningkatkan penjualan produknya dengan melakukan inovasi produk dan melakukan pengujian sikap konsumen di beberapa daerah dengan penilaian :
   5=sangat setuju 4=setuju 3=cukup setuju 2=tidak setuju
   1= sangat tidak setuju

   Hasil pengujian adalah sebagai berikut :

Kota	Sikap
Medan	Sangat setuju
Medan	Setuju
Medan	Tidak setuju
Medan	Setuju
Medan	Sangat tidak setuju
Medan	Sangat setuju
Medan	Cukup setuju
Depok	Sangat tidak setuju
Depok	Setuju
Depok	Sangat setuju
Depok	Setuju
Depok	Tidak setuju
Depok	Tidak setuju
Depok	Sangat setuju
Depok	Cukup setuju

   Selanjutnya dia ingin mengetahui apakah ada perbedaan sikap antara konsumen kota Medan dengan Depok dan dia meminta bantuan saudara, apa yang dapat saudara simpulkan dari tabel di atas untuk menjawab keingintahuan eksekutif tersebut ?

2. Menghadapi kondisi persaingan yang semakin ramai, seorang praktisi pemasaran menerapkan strategi pengurangan harga (discount) terhadap produknya. Setelah 1 bulan, dia ingin mengetahui apakah ada perbedaan penjualan sebagai akibat kebijakan discount tersebut pada 7 toko yang dimilikinya di Jakarta. Dari tabel berikut apa yang dapat saudara simpulkan ?

Toko	Sales sebelum disc	Sales setelah disc
1	1200.00	1230.00
2	1350.00	1450.00
3	1200.00	1250.00
4	1100.00	1140.00
5	1200.00	1200.00
6	1300.00	1300.00
7	1290.00	1300.00

3. Untuk memperbaiki kinerja bagian penjualan, perusahaan mendidik mereka dengan memberikan keahlian penjualan. Karyawan bagian penjualan dikelompokkan menjadi 3 kelompok (masing-masing 10 orang) dan dididik di 3 tempat berbeda, yaitu di JRP, PKL, dan USH. Setelah pendidikan, hasil kerja mereka di evaluasi sebagai berikut : Apakah ada pengaruh tempat pendidikan dengan kinerja penjualan mereka ?

Minggu	Penjualan		
	Kel-JRP	Kel-PKL	Kel-USH
1	52	55	49
2	51	45	48
3	50	53	52
4	55	55	54
5	56	53	52
6	53	50	51
7	48	49	54
8	51	48	48

4. Kondisi penjualan PT. MANDIRI MAKMUR selama 10 tahun terakhir adalah sebagai berikut :

Tahun	Penjualan	Biaya Promosi	Jml. Karyawan	Jml. Toko
1990	1500.00	300.00	50.00	35.00
1991	1450.00	320.00	48.00	34.00
1992	1545.00	320.00	53.00	35.00
1993	1498.00	310.00	55.00	35.00
1994	1600.00	350.00	55.00	40.00
1995	1300.00	250.00	51.00	40.00
1996	1469.00	300.00	55.00	30.00
1997	1690.00	360.00	58.00	45.00
1998	1480.00	320.00	60.00	35.00
1999	1500.00	300.00	59.00	40.00

Dapatkan dikatakan bahwa penjualan ditentukan oleh promosi, jumlah karyawan, dan jumlah toko ? seberapa besar pengaruhnya terhadap penjualan ?