

**MODUL PRAKTIKUM
PRAKTEK STATISTIK**



**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN ILMU KOMPUTER
EL RAHMA
YOGYAKARTA
2011**

MODUL I
PENGENALAN SPSS DAN TABULASI DATA

A. MAKSUD DAN TUJUAN

1. MAKSUD
Mengenalkan SPSS for Windows sebagai salah satu alat untuk mengolah serta menganalisis data statistik
2. TUJUAN
 - a. Agar mahasiswa dapat melakukan operasi dasar SPSS for Windows.
 - b. Agar mahasiswa mampu memasukkan data secara langsung dan menyuntingnya.

B. TEORI SINGKAT

SPSS merupakan singkatan dari *Statistical Product and Service Solution* salah satu program aplikasi komputer untuk menganalisis data statistik. Dilihat dari namanya, SPSS banyak membantu memecahkan berbagai permasalahan ilmu - ilmu sosial, khususnya dalam analisis statistik. Namun demikian fleksibilitas yang dimilikinya menyebabkan berbagai problem analisis data diluar ilmu sosial juga dapat diatasi dengan baik. Keunggulan SPSS for Windows antara lain :

- a. Mudah dalam pengoperasiannya (user friendly), antara lain penekanan modus pengoperasian pada menu full-down dengan *dialog box interface*.
- b. Unggul dalam menampilkan grafik hasil analisis sekaligus kemudahan - kemudahan dalam penyuntingan grafik.
- c. Menyediakan *syntax window* sebagai sarana penyusunan perintah program, untuk mengatasi masalah seperti pembobotan case dengan bobot sesuai yang dikendaki, instruksi- instruksi dengan menu full-down sangat tidak efisien bila dibandingkan dengan penyusunan perintah dengan modus interaktif atau modus batch.
- d. Menyediakan fasilitas analisis yang lebih lengkap dan menyeluruh.
- e. Bagi penyusun program SPSS for Window jauh lebih praktis, karena sudah tidak memerlukan teks editor lainnya.
- f. *Help Window* pada SPSS for Windows, menyediakan *Glossary* yang akan membantu dalam mendefinisikan berbagai istilah statistik
- g. Tersedia *SPSS Syntax* yang akan mendefinisikan berbagai istilah khususnya perintah, subperintah dan kata kunci yang digunakan SPSS For Windows.

Kelemahan SPSS for Windows

- a. Pada menu full-down, diantaranya pada uji beda nyata terkecil (*Least-significant difference*) pada prosedur One-way ANOVA hanya menyediakan tingkat signifikan sebesar 0,05 saja, dimana bisa ditentukan tingkat signifikan yang lain bila bekerja dalam modus pemrograman.
- b. Meskipun menu Help sangat membantu, namun disarankan untuk tidak terlalu tergantung pada bantuan informasi tersebut. karena adanya sedikit kesalahan yang cukup mengganggu pada beberapa informasi tersebut, seperti Help untuk prosedur *Runs Test* (Non Parametrik) yang memuat informasi salah dalam hal pembagian kategori Cut Point, juga informasi yang semestinya tidak perlu ada pada help untuk prosedur *chi-Square*.

1. Mempersiapkan data

Untuk bisa dimengerti oleh Proessor pada SPSS for Windows, data tersebut harus mempunyai format, jenis, dan aturan - aturan tertentu. Hal ini mengingat bahwa selain data yang dibuat melalui SPSS data Editor yang telah siap pakai dalam SPSS for Windowsini, paket program ini juga dapat mengakses data yang disusun dengan menggunakan text editor di luar SPSS for Windows, seperti WordStart (Non document), Side Kick, Notepad, Norton Comander dan text editor yang menggunakan kode ASCII lainnya.

2. Struktur Data

Struktur data pada SPSS terdiri dari mahasiswa baris dan n kolom, tiap baris dinamakan case (kasus), tiap kolom data mempunyai heading yang dinamakan variabel. Terdapat 8 type variabel yang dikenal oleh SPSS yaitu: **Numeric, Comma, Dot, Scientific notation, Date, Dollar, Custom Currency, dan String**. Selain type date dan string type - type data tersebut pada hakekatnya sama dengan type numeric yang membedakan hanya format tampilannya saja. Variabel bertype date berbasis numeric yang didefinisikan untuk format penanggalan dan atau waktu (jam), sehingga variabel ini bisa dianalisis sebagaimana variabel numerik. Jadi prinsipnya hanya ada 2 type variabel pada SPSS yaitu Numeric dan String.

| | | | |
|----------|---|----------------------|----------|
| Arina | 2 | Jl. Adi Sucipto 23 | 14.07.99 |
| Camellia | 2 | Jl. Agus Salim 67 | 19.01.77 |
| Heri | 1 | Jl. Aji Spinggir 50 | 28.11.78 |
| Alfina | 2 | Jl. Merdeka Barat 54 | 05.07.75 |
| Yudi | 1 | Jl. Timoho 90 | 23.12.77 |

Tabel diatas menunjukkan listing data yang terdiri dari 4 variabel dan 5 case. Misal variabel - variabel tersebut diberi nama NAMA, SEX, ALAMAT dan TGL.LAHIR, maka NAMA dan ALAMAT adalah variabel yang bertype string SEX bertype Numeric dan TGL.LAHIR bertype date.

Aturan pemberian nama variabel

- Maksimum 8 karakter
- Tidak boleh ada spasi
- Karakter pertama harus berupa huruf atau karakter @
- Karakter terakhir tidak boleh berupa titik
- Hindari istilah - istilah yang biasa digunakan SPSS, seperti ALL, AND, BY, EQ, GE, GT, LE, NE, NOT, OR, TO, WITH
- Huruf besar dan huruf kecil dianggap sama

Interaksi antara tiap variabel dan tiap case adalah value. Pada contoh diatas value dari variabel NAMA pada xase pertama 'ARINA' dan value dari variabel SEX pada case kedua adalah 2 yang dalam hal ini berarti perempuan.

3. Format Data

File data yang dibuat tidak melalui SPSS Data Editor mempunyai 2 jenis format, yaitu *fixed* (tertentu) dan *freefield* (bebas), sedangkan file data yang dibuat melalui SPSS Data Editor sudah mempunyai format baku yang berorientasi pada

jenis fixed. Format fixed berindikasi bahwa setiap variabel menempati posisi pada lokasi (kolom) tertentu contoh pada tabel diatas. Sedangkan pada format freefield penempatan variabel tidak perlu pada lokasi (kolom) tertentu, contoh pada tabel dibawah ini :

```
'Arina' 2 'Jl. Adi Sucipto 23' 14.07.99
'Camellia' 2 'Jl. Agus Salim 67' 19.01.77
'Hery' 1 'Jl. Aji Spinggir 50' 28.07.78
'Alfina' 2 ' Jl. Merdeka Barat 54' 05.07.75
'Yudi' 1 'Jl. Timoho 90' 23.12.77
```

Terlihat bahwa format freefield sangat tidak rapi jika dibandingkan dengan format fixed, akan tetapi untuk menanganan data dalam jumlah besar format freefield merupakan tip untuk penghematan memori. Hal yang perlu diperhatikan dalam format freefield ialah urutan variabel pada tiap case harus tetap. Value antar variabel pada tiap case harus dipisahkan dengan paling sedikit satu spasi atau satu tanda koma (,).

Untuk menyatakan value suatu variabel string yang mempunyai lebih dari satu kata sebagai satu kesatuan, seperti value 'Jl. Adi Sucipto', maka value tersebut harus diapit dengan tanda apostrof (') atau tanda petik (").

4. Missing value

Missing value digunakan untuk mendeklasraqsikan data yang hilang / tidak lengkap, hal ini diperhatikan karena data yang hilang akan sangat berpengaruh pada hasil pengolahan maupun analisis dari keseluruhan data. Ada 2 jenis missing value yang dikenal oleh SPSS, yaitu :

- User missing value ; mempunyai keentuan yang ditetapkan oleh pemakai, ditentukan harga 0 sebagai missing value.
- System missing value; ditentukan secara otomatis oleh SPSS bila dijumpai harga yang ilegal, seperti didapatinya karakter alphabetic pada variabel berjenis numerik atau perhitngan yang menghasilkan nilai tak terdefinisi pada perintah transformasi data seperti pembagian dengan 0.

5. Mengaktifkan SPSS

Untuk mengaktifkan SPSS dapat dijalankan melalui tombol Star dengan mengklik SPSS, sehingga akan muncul logo SPSS yang diikuti dengan munculnya dua buah window dibelakangnya yaitu Output Windows dan Window Data Editor yang siap dioperasikan, sebagaimana gambar berikut ini :

6. Membuat file Data

Pada saat SPSS diaktifkan secara otomatis SPSS Data Editor akan terbuka dengan nama New Data. Bila pada Data Editor telah berisi data, dapat dibuat file data baru dari menu utama dengan :

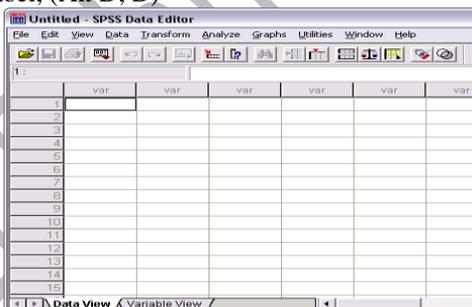
Klik File; New; Data (Alt-F; N; D)

7. Mendefinisikan Variabel

Ada 2 cara untuk mendefinisikan variabel

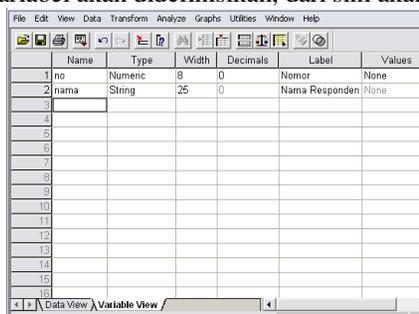
Cara I :

- Pilih sembarang sel pada kolom dimana variabel akan didefinisikan
- Klik Data; Define Variabel; (Alt D; D)



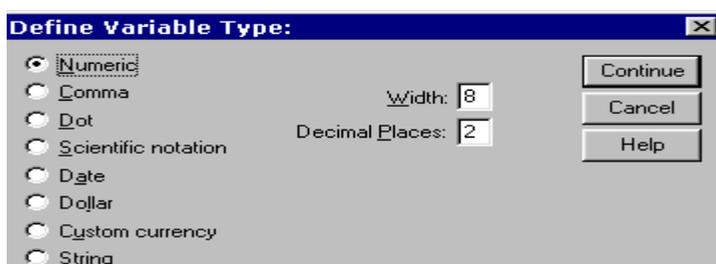
Cara II

Double klik pada judul kolom dimana variabel akan didefinisikan, dari sini akan diperoleh tampilan sebagai berikut :



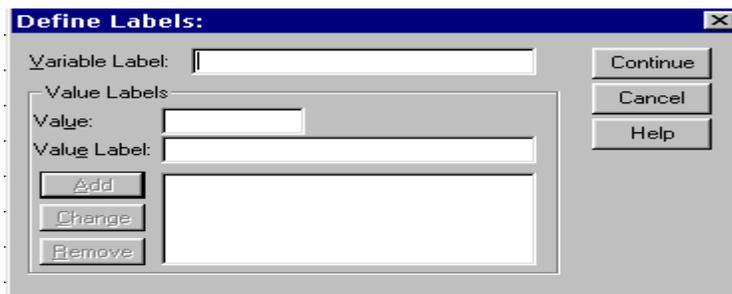
8. Tombol Type

Untuk menentukan tipe variabel, lebar variabel (field) dan jumlah angka desimal. Apabila kita klik pada tombol type akan muncul tampilan sebagai berikut :



9. Tombol Label

Untuk menentukan variabel dan harga dari label tersebut (bila diperlukan), apabila diklik tombol Labels akan muncul tampilan sebagai berikut :



10. Tombol Missing Value

Untuk menentukan harga - harga dari suatu variabel yang akan dideklarasikan sebagai missing value.

11. Tombol Column Format

Untuk menentukan lebar kolom variabel dan untuk menentukan jenis perataannya (kiri, tengah, kanan), penentuan lebar kolom dapat dilakukan dengan drag pada pembatas judul kolom.

12. Menyimpan File Data

Data pada SPSS Data Editor dapat disimpan dengan langkah :

- a. Klik File; Save Data (Alt-F; S)
- b. Beri nama file pada kotak dialog Save Data As. (dalam pemberian nama file data tak perlu dituliskan ekstensinya, SPSS akan memberi ekstensi *.sav secara otomatis)
- c. Tekan OK

C. PRAKTEK

Terdapat data nilai Statistik dari 10 Mahasiswa sebagai berikut :

| No | NIM | NAMA | JK | TGS | UTS | UAS |
|----|-------|--------------------|----|-----|-----|-----|
| 1 | 12001 | Erwin Hadi | 1 | 70 | 50 | 60 |
| 2 | 12002 | Ratnachoerniawatie | 2 | 80 | 52 | 70 |
| 3 | 12003 | Tri Karuniea Ratna | 2 | 75 | 60 | 50 |
| 4 | 12004 | Bayu Anggoro | 1 | 63 | 72 | 60 |
| 5 | 12005 | Bayu El Reyzal | 1 | 85 | 70 | 75 |
| 6 | 12006 | Krisna Fahriesta | 2 | 80 | 80 | 82 |
| 7 | 12007 | Izmaiel Fahmie | 1 | 72 | 72 | 80 |
| 8 | 12008 | Elvin Rizqyanie | 2 | 75 | 75 | 90 |
| 9 | 12009 | Helvy Triana Rosa | 2 | 80 | 80 | 60 |
| 10 | 12010 | Fariq Dian Wisesa | 1 | 70 | 70 | 70 |

Masukkan data tersebut diatas dengan menggunakan SPSS melalui langkah - langkah sebagai berikut :

1. Aktifkan SPSS dari menu utama Windows melalui tombol Start dengan mengklik SPSS.
2. Masukkan data diatas pada Window Data Editor dengan cara :
 - a. Double klik pada judul kolom pertama, sehingga akan muncul kotak dialog Define Variabel.
 - b. Isi variabel name dengan 'No'
 - c. Klik tombol Type, sehingga muncul kotak dialog Define Variable Type, pilih type numeric, dengan With = 4, Decimal = 0, klik tombol Continue, akan muncul kotak Define Variable lagi, klik OK kemudian pindah ke judul kolom berikutnya.
 - d. Ulangi langkah a, b, c, sehingga secara keseluruhan seperti terlihat pada tabel berikut ini

| Variable Name | Type | Width | Dec | Label |
|---------------|---------|-------|-----|--------------------------|
| No | NUMERIC | 4 | 0 | |
| NIM | STRING | 6 | | |
| NAMA | STRING | 15 | | |
| JK | STRING | 3 | | 1=Laki-laki, 2=Perempuan |
| TGS | NUMERIC | 4 | 0 | |
| UTS | NUMERIC | 4 | 0 | |
| UAS | NUMERIC | 4 | 0 | |

- e. Isikan data pada setiap sel seperti mengisi data pada excel
- f. Simpan file data dengan nama Modul 1.
3. Tampilkan file data tersebut dengan klik Menu Analyze; Report; Case Summarize, sehingga muncul kotak dialog List-Case. Pindahkan variabel - variabel yang akan ditampilkan pada kotak variable(s) dengan mengklik tanda panah berurutan, klik OK sehingga muncul tampilan pada Window Output.
4. Simpan output tersebut dengan nama output 1

**MODUL II
MENYUNTING DATA DAN TRANSFORMASI DATA**

A. MAKSUD DAN TUJUAN

1. MAKSUD

- Melakukan operasi - operasi terhadap file data
- Mengenalkan beberapa perintah transformasi data

2. TUJUAN

- Agar mahasiswa bisa menghapus, menyisipkan, mengkopi, mentranspose, mengurutkan data, memecah file dan memilih kasus.
- Agar mahasiswa mampu mentransformasikan (modifikasi atau merubah) data dengan perintah COMPUTE, COUNT, RECODE.

B. TEORI SINGKAT

1. Menghapus data

- a. Menghapus isi sel
 - Pilih sel yang akan dihapus isinya dengan dengan cara klik dengan mouse
 - Tekan tombol delete pada keyboard, untuk menghapus isi sejumlah sel sekaligus, pilih sejumlah sel tersebut dengan drag.
- b. Menghapus isi satu kolom
 - Klik judul kolom (nama variabel) yang akan dihapus isi -isi selnya.
 - Tekan tombol delete pada keyboard, untuk menghapus sejumlah kolom, pilih sejumlah kolom tersebut dengan drag pada bagian judul (nama variabel).
- c. Menghapus isi satu baris
 - Klik nomor case yang akan dihapus
 - Tekan tombol delete pada keyboard, untuk menghapus sejumlah case sekaligus, pilihlah sejumlah case tersebut dengan drag pada bagian nomor case.

2. Mengcopi Data

- a. Mengcopi isi sel
 - Pilih sel yang akan dicopi isinya
 - Tekan tombol Ctrl C
 - Pindahkan penunjuk sel ke sel yang dituju
 - Tekan Ctrl V
- b. Mengcopi isi satu kolom
 - Klik judul kolom (nama variabel) yang akan dicopi isi - isi selnya
 - Tekan Ctrl C
 - Pindahkan penunjuk sel ke baris pertama yang dituju
 - Tekan Ctrl V

Hasil dari instruksi diatas adalah akan mengcopi kolom sekaligus dengan format variabelnya (tipe variabel, lebar kolom value label, dsb), dan pasti tidak mengubah nama variabel. Bila dikehendaki tidak ada perubahan format variabel pada kolom yang dituju yang dilakukukan adalah :

- Klik judul kolom (nama variabel) yang akan dicopi isi - isi selnya
- Tekan Ctrl C
- Pindahkan penunjuk sel ke baris pertama kolom yang dituju
- Tekan Ctrl V

Untuk mengcopi isi sejumlah kolom sekaligus pilihlah sejumlah kolom tersebut dengan drag pada bagian judul kolom.

c. Mengcopi case

- Klik nomor case yang akan di copi
- Tekan Ctrl C
- Klik nomor case yang dituju atau pindahkan penunjuk sel ke kolom pertama baris yang dituju.
- Tekan Ctrl V

Untuk mengcopi sejumlah case sekaligus pilihlah sejumlah case tersebut dengan drag pada bagian nomor case

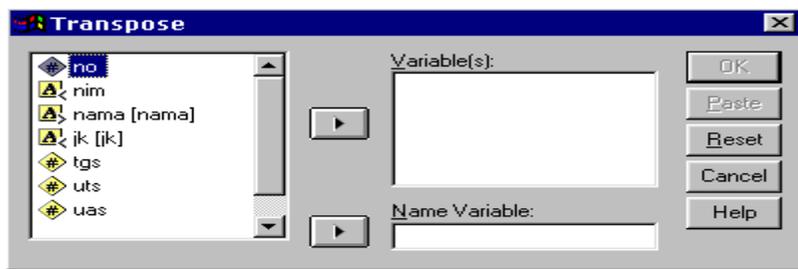
3. Menyisipkan Data

- a. Menyisipkan kolom
 - Pindahkan penunjuk sel pada kolom yang akan disisipi
 - Klik Data ; Insert Variable (Alt D;V)
- b. Menyisipkan baris
 - Pindahkan penunjuk sel pada baris yang akan disisipi
 - Klik Data; Insert Case (Alt D; I)

4. Transpose Data

Transpose data adalah memindahkan kolom data menjadi baris data dan sebaliknya baris data kolom data. Perintah trasnpose ini akan menghasilkan file data baru. Langkah - langkah berikut ini untuk mentranspose file data dengan nama file *Modul 1.sav*

- a. Buka file modul 1.sav
- b. Klik Data; Transpose (Alt D; N)



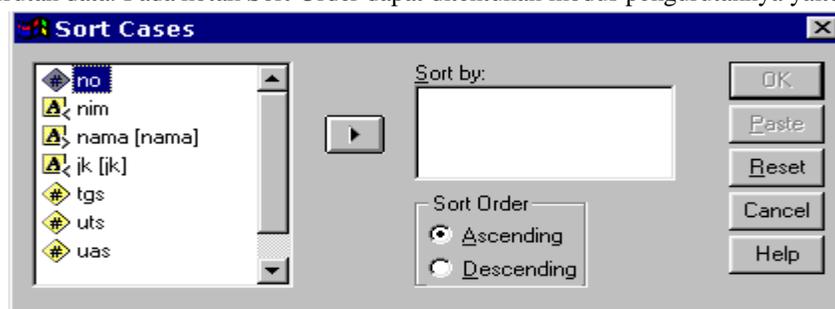
- c. Pindahkan satu atau beberapa variabel ke kotak *variable(s)*.
- d. Tekan OK.

Maka variabel - variabel yang akan dipindahkan akan menjadi case (baris) pada file data baru dan variabel - variabel yang tidak dipindahkan akan dihilangkan.

5. Pengurutan

Pengurutan case pada file data didasarkan pada satu atau beberapa variabel dapat dilakukan dengan menggunakan perintah Sort Cases pada menu Data, dari perintah ini akan didapat kotak dialog sebagai berikut :

Pindahkan satu atau beberapa variabel ke kotak *sort by*, maka variabel - variabel tersebut akan dijadikan dasar pengurutan data. Pada kotak Sort Order dapat ditentukan modus pengurutannya yaitu *Ascending* atau *Descending*.



6. Split File

Perintah Split File digunakan untuk memecah file data ke dalam sub group - subgroup sehingga nantinya dalam analisis maupun dalam pembuatan chart didasarkan pada tiap - tiap subgroup tersebut (analisis) terpisah.

Untuk mendapatkan analisis terpisah terhadap masing - masing group dari beberapa group case, pilihlah Repeat Analysis for each group, dan pindahkan satu atau beberapa variabel (maks 8 variabel) ke kotak *Group Base On*. Variabel - variabel yang dipindahkan tersebut adalah variabel - variabel yang akan dipisah - pisah casenya ke dalam beberapa group. Ada 2 ketentuan yang perlu diperhatikan mengenai jumlah variabel yang anda pindahkan ke kota *Group Based On*:

- Bila hanya satu variabel yang akan dipindahkan maka nantinya dalam analisis dikelompokkan berdasarkan masing - masing value dari variabel tersebut.
- Bila lebih dari satu variabel yang akan dipindahkan maka nantinya dalam analisis dikelompokkan berdasarkan kombinasi value - value dari variabel - variabel yang dipindahkan tersebut..

Untuk Restore (menghilangkan pengaruh perintah Split File yang telah dijalankan), pilih tombol Option Analyze all case atau membuka kembali file data baru.

Dua buah tombol option di bawah kotak Group Based On dapat diaplikasikan hanya bila terlebih dahulu ditentukan pilihan pada tombol option Repeat analysis for each group, karena SPSS telah melakukan perintah pengurutan secara otomatis berdasarkan variabel group maka kemudian ditentukan pilihan file is already sorted. Variabel group yang digunakan sebagai dasar pengelompokkan dapat dilihat pada current status dibagian bawah kotak dialog Split File.

7. Select Case

Select case dapat digunakan untuk menentukan case - case yang akan diikutsertakan dalam analisis data selanjutnya. Terdapat 5 pilihan untuk menentukan dasar penyeleksian case :

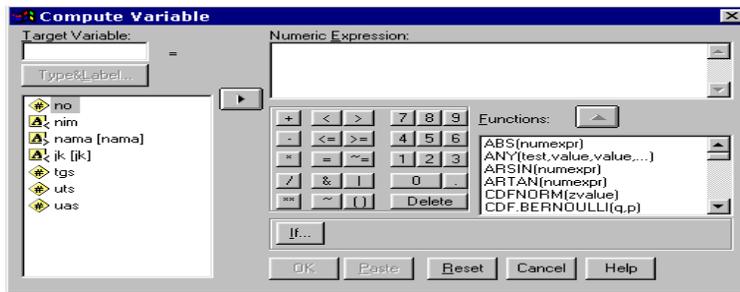
- a. All Case, Seluruh case pada file kerja akan diikutsertakan dalam analisis data selanjutnya, pilihan ini merupakan default.
- b. If Condition is Satisfied ; Digunakan untuk memilih case - case yang memenuhi kondisi logika tertentu, tekan tombol if... untuk menentukan kondisi logika tersebut.
- c. Random Sample of Cases; Digunakan bila hendak menggunakan case - case random untuk pemrosesan data (menentukan sample data random dari populasi). Tekan tombol sample... untuk menentukan ukuran sample data tersebut.

Ada 2 pilihan kotak dialog dibawah ini :

- a. Aproximately -% of all case. Untuk menentukan persentase case yang akan diproses dari seluruh case pada file data, masukkan integer (bilangan bulat positif) antara 1 hingga 99 pada kotak yang tersedia.
- b. Exactly case from the firstcase. Untuk menentukan proporsi sampel yang akan digunakan, masukan integer pada 2 buah kotak yang tersedia, dimana isi pada kotak yang pertama harus lebih kecil dari isi pada kotak yang kedua, dan isi kotak yang kedua seharusnya tidak lebih besar dari jumlah cacah case pada file data. Pilihan ini akan menghasilkan sebuah variabel baru yang oleh SPSS diberi nama Filter_\$.
- c. Based on time or case range. Digunakan bila hendak memproses case -case yang berada dalam interval tertentu, tekan tombol Range.. untuk menentukan batas interval tersebut. Masukan integer pada dua buah kotak yang tersedia, yakni kotak first case dan kotak last case yang masing - masing untuk nomor awal case dan nomor akhir case yang akan dianalisis.
- d. Use Filter Variable. Digunakan bila hendak menggunakan value - value dari variabel - variabel numerik yang tersedia pada kotak listing variabel untuk mengontrol case filter.

8. Perintah COMPUTE

Perintah Compute digunakan untuk melakukan perhitungan terhadap value - value dari variabel - variabel yang sudah ada maupun untuk variabel baru pada file data, perhitungan ini berlaku untuk semua case atau untuk case - case yang memenuhi kondisi tertentu. Apabila kita jalankan perintah Compute dari menu Trasform akan terlihat tampilan sebagai berikut :



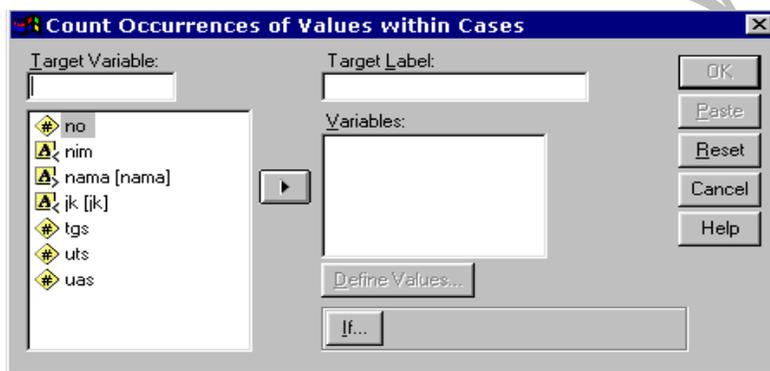
Keterangan :

- a. Kotak Target Variable ; Digunakan untuk mengisikan nama variabel target, nama variabel target bisa berupa nama variabel yang sudah ada pada file data
- b. Kotak Numeric Expression; Digunakan untuk menyusun ekspresi yang akan digunakan untuk transformasi variabel target. Untuk mengisikan ekspresi tersebut bisa menerapkan operasi dan angka yang ada pada calculator pad, bisa dengan memindahkan nama variabel dari kotak listing variabel atau memindahkan fungsi - fungsi yang diperlukan dari kotak listing fungsi atau bisa pula dengan mengetikan langsung pada kotak Numeric expression.

Tombol If...; Digunakan untuk menentukan kondisi logika bila dalam transformasi variabel memenuhi kondisi tertentu, yang perlu diperhatikan disini adalah bahwa tombol Include if case satisfied condition harus dipilih sebelum menyusun logika tersebut, defaultnya adalah pilihan Include all case.

9. Perintah COUNT

Perintah count digunakan untuk menghitung jumlah cacah value dan seluruh variabel yang di daftar yang memenuhi syarat value yang telah didefinisikan pada tiap case, apabila kita jalankan perintah COUNT dari menu transform maka akan terlihat tampilan sebagai berikut :



- a. Kotak Target Variable; Fungsi kegunaannya sama dengan pada perintah COMPUTE
- b. Kotak Target Label; Pada kotak ini diisikan dengan label dari variable target
- c. Kotak Variables; digunakan untuk mendaftarkan satu atau lebih variabel dari kotak listing variabel, dengan syarat variable - variable tersebut harus mempunyai tipe yang sama.
- d. Tombol Define Value; Digunakan untuk mendefinisikan syarat value yang harus dipenuhi, sebelum mendefinisikan syarat value harus dipilih dahulu salah satu dari 6 spesifikasi syarat berikut :

- Value; Digunakan untuk menghitung cacah value yang berharga sama dengan harga yang didefinisikan
- System Missing; Digunakan untuk menghitung cacah value yang berharga system missing value
- System or user Missing; Digunakan untuk menghitung cacah value yang berharga missing
- Range dengan Interval; Digunakan untuk menghitung cacah value yang berharga lebih besar atau sama dengan awal interval dan lebih kecil atau sama dengan akhir interval
- Range lebih kecil atau sama dengan interval; Digunakan untuk menghitung cacah value yang berharga lebih besar atau sama dengan awal interval dan lebih kecil atau sama dengan akhir interval.
- Range dengan interval; Digunakan untuk menghitung cacah value yang berharga lebih besar atau sama dengan awal interval dan lebih kecil atau sama dengan akhir interval.

Setelah mengisi syarat value, maka syarat value dapat dipindahkan ke kotak value to count dengan menekan tombol Add, jika menghendaki syarat lebih dari satu dapat diulang kembali.

Untuk mengedit syarat value yang telah dipindahkan ke kotak value to count maka pilih syarat value tersebut dan edit pada kotak value selanjutnya tekan tombol *Change*.

Untuk menghapusnya pilih syarat value tersebut kemudian tekan tombol *Remove*

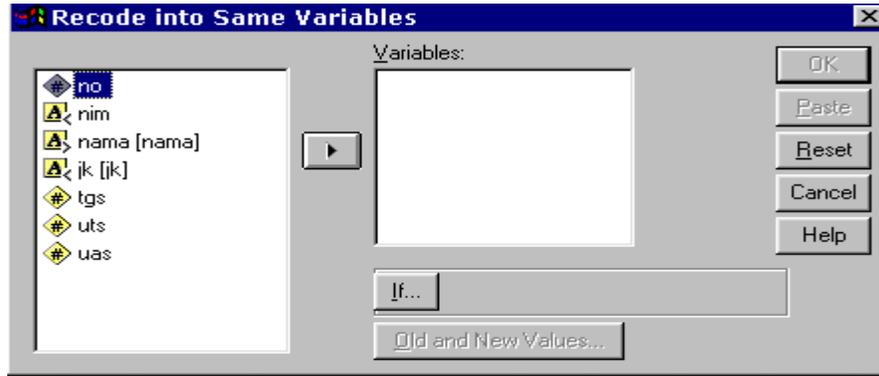
10. Perintah RECODE

Perintah Recode digunakan untuk memodifikasi atau menggantikan value dari variabel yang didaftarkan menjadi value dengan nilai baru. Apabila perintah Recode dijalankan dari menu transform maka akan muncul pilihan :

- Into same variables yakni recode untuk variable yang sudah ada
- Into Different Variables yakni recode untuk membentuk variabel baru

- a. Memodifikasi value variable dengan Recode

Untuk memodifikasi value variabel maka dipilih perintah *Into Same Variables* sehingga akan muncul kotak dialog sbb:

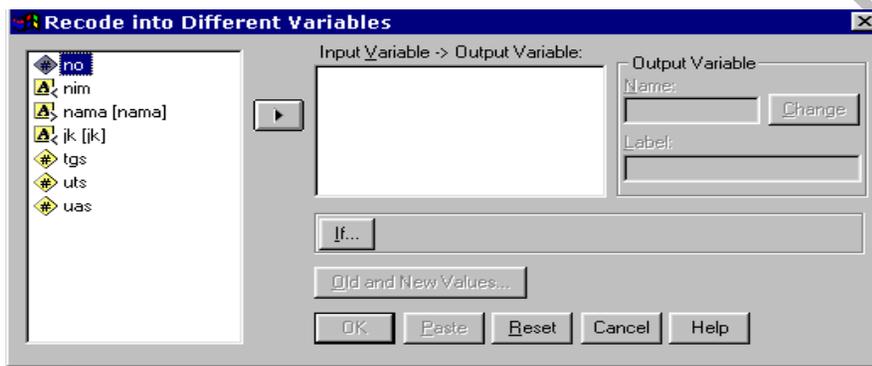


Tombol Old and New Value....

Tombol ini digunakan untuk menentukan value yang akan diganti dan value pengganti dari variabel yang didaftar, setelah mengisi kotak old value dan new value maka klik tombol Add, ulangi jika ingin menyusun ketentuan yang lain.

- b. Membentuk variabel baru dengan Recode

Untuk membentuk variabel baru maka dipilih perintah Into Different Variables sehingga akan muncul kotak dialog Sbb:



C. PRAKTEK

1. Menyisipkan, transpose data, mengcopi data, dan mengurutkan data
 - a. Buka file data (Latihanku.sav)
 - b. Sisipkan variabel JUR diantara variabel NIM dan NAMA, isilah dengan Jurusan MI, SI, TK, TI, KA dst
 - c. Copilah case no 1-3, letakan di baris 12
 - d. Hapus variabel tugas, lihat hasilnya dan simpan dengan nama MODUL 2
 - e. Buka kembali file MODUL 1 lalu pilih data, transpose dan masukan tgs, uts dan uas pada kotak variabel(s) dengan cara mengklik tanda panah
 - f. Pindahkan nama ke kotak Name variabel dengan cara meng-klik tanda panah lalu klik OK, lihat hasilnya dan simpan dengan nama TRANS
 - g. Buka kembali MODUL 1, lakukan perintah Sort secara urut naik berdasarkan nama, simpan dengan nama URUT 1. Urutkan juga berdasarkan NIM secara urut turun, simpan dengan nama URUT 2.
2. Perintah Compute
 - a. Buka file data anda
 - b. Klik menu Transform, Compute sehingga keluar kotak dialog compute Variabels
 - c. Isi kotak target Variable dengan nama Rerata
 - d. Isi kotak Numerik Expression denga $(tgs + uts + uas) / 3$ dan klik OK
 - e. Simpan dengan nama Rerata.

| | no | nim | nama | jk | tgs | uts | uas | rerata |
|----|----|-------|--------|----|-----|-----|-----|--------|
| 1 | 1 | 55233 | Doni | 1 | 70 | 50 | 60 | 60,00 |
| 2 | 2 | 55004 | Kiki | 2 | 80 | 52 | 70 | 67,33 |
| 3 | 3 | 55223 | Ruli | 2 | 75 | 60 | 50 | 61,67 |
| 4 | 4 | 56780 | Dimas | 1 | 63 | 72 | 60 | 65,00 |
| 5 | 5 | 52490 | Riko | 1 | 85 | 70 | 75 | 76,67 |
| 6 | 6 | 53256 | Rika | 2 | 80 | 80 | 82 | 80,67 |
| 7 | 7 | 54200 | Tomas | 1 | 71 | 72 | 80 | 74,33 |
| 8 | 8 | 52550 | Arum | 2 | 60 | 75 | 90 | 75,00 |
| 9 | 9 | 54443 | Mayang | 2 | 68 | 80 | 60 | 69,33 |
| 10 | 10 | 52439 | Puguh | 1 | 90 | 70 | 70 | 76,67 |

3. Perintah Count
 - a. Buka file data anda
 - b. Klik menu Transform, Count sehingga muncul kotak dialog Count Occurences of values within cases
 - c. Isi kotak target Variables dengan nama Cacah1
 - d. Isi kotak Numeric Variables dengan memindahkan variable yang telah tersimpan ke kotak Numeric Variables dengan cara mengklik tanda panah.

- e. Tekan tombol define value, pilih range isi kotak lowest trough dengan nilai 60; tekan Add; tekan Continue
- f. Tekan tombol IF; tekan Include if Case Satisfies Condition; tulis $jk=1$ pada kotak yang tersedia; tekan Continue; tekan OK.
- g. Simpan dengan nama file Cacah1

| | no | nim | nama | jk | tgs | uts | uas | rerata | cacah1 |
|----|----|-------|--------|----|-----|-----|-----|--------|--------|
| 1 | 1 | 55233 | Doni | 1 | 70 | 50 | 60 | 60,00 | 2,00 |
| 2 | 2 | 55004 | Kiki | 2 | 80 | 52 | 70 | 67,33 | . |
| 3 | 3 | 55223 | Ruli | 2 | 75 | 60 | 50 | 61,67 | . |
| 4 | 4 | 56780 | Dimas | 1 | 63 | 72 | 60 | 65,00 | 3,00 |
| 5 | 5 | 52490 | Riko | 1 | 85 | 70 | 75 | 76,67 | 3,00 |
| 6 | 6 | 53256 | Rika | 2 | 80 | 80 | 82 | 80,67 | . |
| 7 | 7 | 54200 | Tomas | 1 | 71 | 72 | 80 | 74,33 | 3,00 |
| 8 | 8 | 52550 | Arum | 2 | 60 | 75 | 90 | 75,00 | . |
| 9 | 9 | 54443 | Mayang | 2 | 68 | 80 | 60 | 69,33 | . |
| 10 | 10 | 52439 | Puguh | 1 | 90 | 70 | 70 | 76,67 | 3,00 |

4. Perintah RECODE

- a. Buka file data anda
- b. Klik menu Transform; Recode; Into Defference Variables, sehingga muncul kotak dialog Recode Into Deference Variable
- c. Isi kotak Numeric variable dengan variabel Rerata
- d. Isi kotak output dengan Nilai; klik Change
- e. Klik tombol Old &New Values, muncul kotak dialog
- f. Klik Output variables Are Strings
- g. Klik Range; isi kotak Higest Trought dengan nilai 80; klik value, ketik A; klik Add.
- h. Klik Range; isi kotak Trought dengan nilai 65 -79; klik value; ketik B, klik Add.
- i. Klik Range; isi kotak Trought dengan nilai 50 - 64; klik value; ketik C, klik Add.
- j. Klik Range; isi kotak Trought dengan nilai 40 - 49; klik value; ketik D, klik Add.
- k. Klik Range; isi kotak Lowest Trought dengan nilai 40; klik value; ketik E, klik Add
- l. Klik Continue; OK
- m. Simpan dengan nama file Nilai

| | no | nim | nama | jk | tgs | uts | uas | rerata | cacah | nilai |
|----|----|-------|--------|----|-----|-----|-----|--------|-------|-------|
| 1 | 1 | 55233 | Doni | 1 | 70 | 50 | 60 | 60,00 | 2,00 | C |
| 2 | 2 | 55004 | Kiki | 2 | 80 | 52 | 70 | 67,33 | . | B |
| 3 | 3 | 55223 | Ruli | 2 | 75 | 60 | 50 | 61,67 | . | C |
| 4 | 4 | 56780 | Dimas | 1 | 63 | 72 | 60 | 65,00 | 3,00 | B |
| 5 | 5 | 52490 | Riko | 1 | 85 | 70 | 75 | 76,67 | 3,00 | B |
| 6 | 6 | 53256 | Rika | 2 | 80 | 80 | 82 | 80,67 | . | A |
| 7 | 7 | 54200 | Tomas | 1 | 71 | 72 | 80 | 74,33 | 3,00 | B |
| 8 | 8 | 52550 | Arum | 2 | 60 | 75 | 90 | 75,00 | . | B |
| 9 | 9 | 54443 | Mayang | 2 | 68 | 80 | 60 | 69,33 | . | B |
| 10 | 10 | 52439 | Puguh | 1 | 90 | 70 | 70 | 76,67 | 3,00 | B |

MODUL III
 PROSEDUR FREQUENCIES, DESCRIPTIVES DAN CROSSTABS

A. MAKSUD DAN TUJUAN

1. MAKSUD

- Menggunakan prosedur Frequencies dan Descriptives dalam menu Analyze, untuk menyajikan dan menganalisis data yang sudah disimpan dalam file data.
- Mengenalkan prosedur Crosstabs sebagai salah satu metode analisis data statistik

2. TUJUAN

- a. Mengetahui perintah - perintah untuk analisis data dalam menu Analyze
- b. Trampil menggunakan perintah - perintah dalam prosedur Frequencies dan Descriptives
- c. Agar mahasiswa dapat menyajikan data dalam tabulasi silang dengan prosedur Crosstabs

B. TEORI SINGKAT

Setelah data diperoleh, biasanya kemudian dianalisis dengan terlebih dahulu menghitung nilai statistiknya, misal rata - rata, median, modus, simpangan baku, dan kemiringan. Dalam SPSS mempunyai fasilitas untuk menghitung nilai - nilai statistik dan analisisnya, yang tersedia dalam menu analyze.

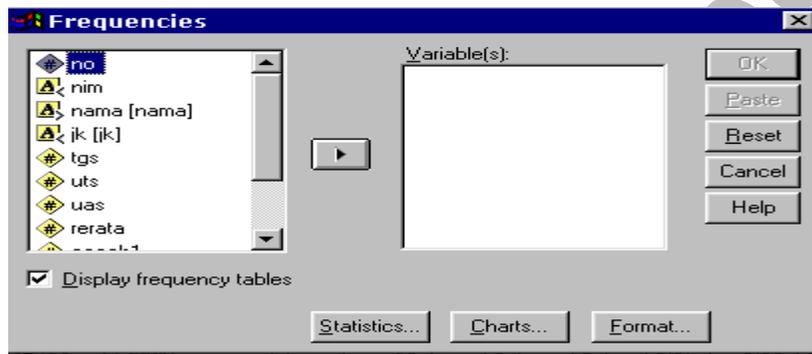
Dalam submenu Descriptives statistic tersebut diantaranya terdapat prosedur Frequencies, Descriptives dan Crosstabs.

1. PROSEDUR FREQUENCIES

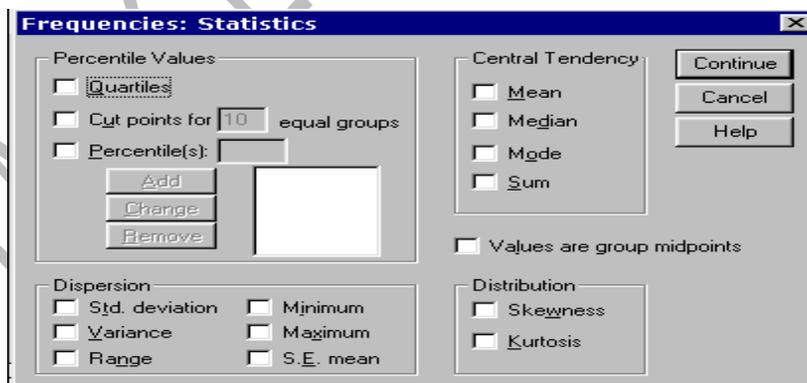
Prosedur frequencies digunakan untuk membuat tabel frekuensi, yaitu berisi cacah harga semua case dari variabel - variabel dalam daftar, cacah persentase semua case, cacah persentase valid (tanpa missing value), dan persentase kumulatif.

Bila dipilih Frequencies dari menu Descriptive statistic akan keluar tampilan seperti berikut :

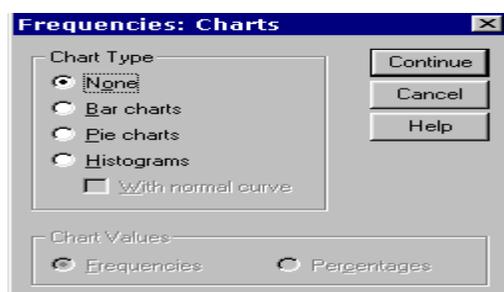
Spesifikasi minimum dari prosedur ini adalah memindahkan sebuah variabel ke kotak *Variable(s)*. check box pada display frequency bila ditandai akan dihasilkan tabel frekuensi, selain itu terdapat tiga tombol lain yaitu *statistides*, *charts* dan *format*



- a. tombol statistics; Tombol ini digunakan untuk menampilkan deskripsi univariat dari variabel numerik yang ada di daftar, antara lain ukuran tendensi sentral (Mean, Median, Modus, Jumlah), ukuran letak (Kuartil, Desil, Persentil), ukuran Dispersi (Nilai maksimum, nilai minimum, Range, Varians, Standard Deviasi), Kemiringan (Skewness) dan Kurtosis. Statistik - statistik yang diinginkan dapat diminta dengan menandai check box yang tersedia, bila tombol tersebut diklik maka keluar kotak dialog sebagai berikut :



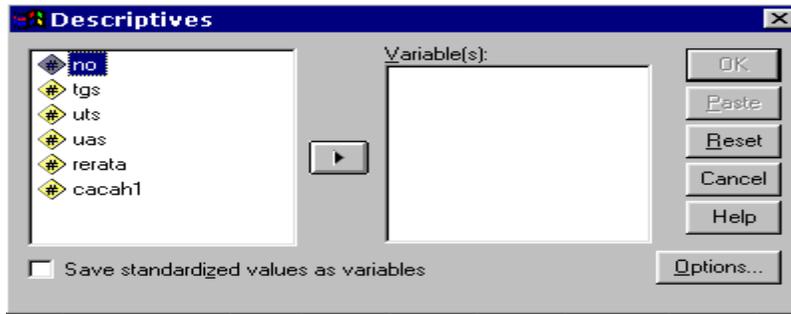
- b. Tombol chart; Tombol ini digunakan untuk menampilkan chart bar atau histogram pada chart carousel dari tabel frekuensi yang dihasilkan, bila tombol ini di klik maka akan muncul kotak dialog sebagai berikut :



Tombol Format; Untuk menentukan format tabel frekuensi, yaitu format halaman dan format pengurutan.

2. PROSEDUR DESCRIPTIVES

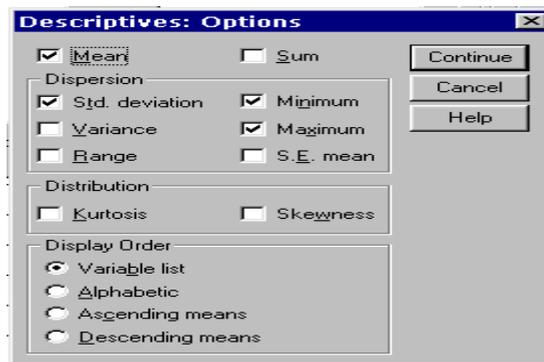
Prosedur ini juga digunakan untuk menampilkan deskripsi statistik univariat dari variabel numerik pada daftar (sebagaimana yang juga bisa dilakukan oleh prosedur Frequencies). Bedanya adalah prosedur ini tidak menampilkan susunan value terurut dalam tabel frekuensi. Adapun kotak dialog yang muncul jika prosedur Descriptives dipilih adalah seperti gambar berikut :



Spesifikasi minimum dari prosedur ini adalah memindahkan sebuah variabel ke kotak Variable(s). dibandingkan dengan prosedur frekuensi, prosedur deskriptif ini sangat efisien untuk menampilkan berbagai ukuran statistik univariat dari banyak variabel sekaligus.

Tombol option

Tombol ini menampilkan mean, sum, ukuran penyebaran (dispersi), distribusi (kemiringan dan kurtosis) dan tipe pengurutan, seperti gambar berikut :



3. PROSEDUR CROSSTABS

Prosedur Crosstabs atau Xtabs digunakan untuk menampilkan tabulasi silang (tabel kontingensi) yang menunjukkan distribusi bersama, deskripsi statistik bivariatnya dan berbagai pengujian dari 2 variabel atau lebih, khususnya variabel berbentuk kategori.

Perintah ini juga digunakan untuk mengetahui apakah ada hubungan diantara 2 variabel

Ada 3 macam kotak untuk menampung variabel - variabel yang berperan dalam pengelompokkan tabel yakni :

1. Kotak Rows; Untuk mendaftarkan variabel - variabel yang casenya akan ditampilkan secara baris (sebagai heading baris tabel)
2. Kotak Columns; Untuk mendaftarkan variabel - variabel yang casenya akan ditampilkan secara kolom (sebagai heading kolom tabel)
3. Kotak Layers; Untuk mendaftarkan variabel - variabel kontrol, yakni variabel - variabel yang akan mengelompokkan tabel per case variabel - variabel tersebut

Apabila check box suppress tables ditandai maka pada output akan menampilkan statistik yang dipilih dan tidak menampilkan tabulasi silang.

Kotak Suppress Table....

Apabila kotak ini ditandai maka pada output hanya menampilkan statistik yang dipilih, sedang tabulasi silangnya tidak ditampilkan.

Tombol Statistics....

Tombol ini digunakan untuk menampilkan statistik dua variabel dari variabel baris dan variabel kolom

Tombol Cell....

Tombol ini digunakan untuk menentukan isi sel. Defaultnya isi sel berupa cacah case pada sel tersebut, selain default tersebut dapat pula ditambahkan :

1. Expected; Banyaknya case yang diharapkan pada sel bila variabel baris dan variabel kolom independen
2. Rows; Persentase baris
3. Column; Persentase kolom
4. Total; persentase total
5. Unstandardized residual; residual yang tak terstandar
6. Standardized residual; residual yang terstandar, residual yang dibagi akar kuadrat dari expected
7. Adjusted standardized residual; residual dibagi estimasi standar errornya

Tombol format...

Tombol ini digunakan untuk menentukan format tabel yang dihasilkan. Pada kotak labels digunakan untuk menentukan pemunculan label dan value variabel, pada kotak row order digunakan untuk menentukan modus pengurutan.

Sedangkan check box Boxes Araund cells dan Index of Tables masing - masing untuk menentukan ada atau tidaknya garis pada sel tabel dan untuk menentuukan penampilan indeks.

C. PRAKTEK

1. Untuk perintah Frequencies
 - a. Buka file data anda
 - b. Klik menu Analyze, Frequencies sehingga tampil kotak dialog frequencies
 - c. Isi kotak Variable dengan TGS; klik tombol statistics sehingga muncul kotak dialog frequencies statistics
 - d. Klik Quartiles; Percentiles isi dengan nilai 75; klik mean, median, mode, sum; klik Standard Deviation, Variances, Minimum, Maximum
 - e. Klik Continue; OK
 - f. Simpan output dengan nama file Frekuensi
2. Untuk perintah Descriptif
 - a. Buka file data anda
 - b. Klik menu Analyze; Descriptives Statistics; Descriptives, sehingga muncul kotak dialog Descriptives
 - c. Isi kotak variable dengan TGS; klik tombol option sehingga muncul kotak dialog Descriptives
 - d. Klik mean, sum, standard deviation, Variances, minimum, maximum
 - e. Klik Continue; OK
 - f. Simpan output dengan nama file Deskriptif.
3. Untuk perintah Crosstabs
 - a. Buka file data anda
 - b. Dari menu Transform pilih compute, pada kotak target variable tuliskan mean, dan pada kotak Numeric Expression masukan Tgs, Uts dan uas untuk menghitung rata - rata dari ketiga nilai tersebut.
 - c. Gunakan perintah Recode untuk membuat variabel baru dengan nama variabel NILAI, dari variabel ini dengan ketentuan :
 - " Kurang " untuk nilai < 60
 - " Cukup " untuk nilai 60 - 80
 - " Baik " untuk nilai > 80
 - d. Dari menu Analyze; Descriptive Statistics; klik Crosstabs sehingga tampil kotak Crosstabs
 - e. Pindahkan Variabel JK ke kotak Row(s) dengan mengklik tanda panah
 - f. Pindahkan variabel NILAI ke kotak Column(s) dengan mengklik tanda panah
 - g. Tekan tombol Cells sehingga muncul kotak dialog Crosstabs, Cell Display, tandai Check Box Expected, row, column, total, standardize
 - h. Klik Continue; OK
 - i. Simpan dengan nama file Crosstabs

MODUL IV
PENGUJIAN RATA-RATA SATU DAN DUA SAMPEL

A. MAKSUD DAN TUJUAN

1. MAKSUD

- Menguji data dengan pengujian rata-rata satu sampel
- Menguji data dengan pengujian rata-rata dua sampel

2. TUJUAN

- Mahasiswa mampu menguji data dengan pengujian rata-rata satu sampel
- Mahasiswa mampu menguji data dengan pengujian rata-rata dua sampel

B. TEORI SINGKAT

Hipotesis dapat diartikan sebagai dugaan mengenai suatu hal, atau hipotesis merupakan jawaban sementara suatu masalah, atau juga hipotesis dapat diartikan sebagai kesimpulan sementara tentang hubungan suatu variabel dengan satu atau lebih variabel yang lain. Namun menurut Prof. Dr. S. Nasution definisi hipotesis adalah pernyataan tentatif yang merupakan dugaan mengenai apa saja yang sedang kita amati dalam usaha untuk memahaminya.

• Fungsi

- Untuk menguji kebenaran suatu teori
- Memberikan gagasan baru untuk mengembangkan suatu teori.
- Memperluas pengetahuan peneliti mengenai suatu gejala yang sedang dipelajari.

• Pengujian hipotesis

Hipotesis yang baik selalu memenuhi dua pernyataan, yaitu :

- Menggambarkan hubungan antar variabel.
- Dapat memberikan petunjuk bagaimana pengujian hubungan tersebut.

Oleh karena itu hipotesis perlu dirumuskan terlebih dahulu sebelum dilakukan pengumpulan data. Hipotesis ini disebut **Hipotesis Alternatif** (H_a) atau **Hipotesis kerja** (H_k) atau H_1 . Hipotesis kerja atau H_1 merupakan kesimpulan sementara dan hubungan antar variabel yang sudah dipelajari dari teori-teori yang berhubungan dengan masalah tersebut. Untuk pengujian H_1 perlu ada pembandingan yaitu **Hipotesis Nol** (H_0). H_0 disebut juga sebagai Hipotesis Statistik, karena digunakan sebagai dasar pengujian.

Langkah atau prosedur untuk menentukan apakah menerima atau menolak Hipotesis Statistik (H_0) disebut Pengujian Hipotesis. Oleh karena itu dalam pengujian Hipotesis, penarikan kesimpulan mengenai populasi didasarkan pada informasi sampel bukan populasi itu sendiri, maka kesimpulannya dapat saja keliru. Dalam Pengujian Hipotesis terdapat dua kekeliruan atau galat, yaitu :

| Kesimpulan | Keadaan sebenarnya H_0 | |
|--------------|----------------------------|----------------------------|
| | H_0 benar | H_0 salah |
| Terima H_0 | tepat | Galat jenis II (β) |
| Tolak H_0 | galat jenis I (α) | tepat |

Penarikan kesimpulan dinyatakan tepat apabila kita menerima H_0 , karena memang H_0 benar, atau menolak H_0 , karena memang H_0 salah. Apabila kita menyimpulkan menolak H_0 padahal H_0 benar, maka kita telah melakukan kekeliruan yang disebut kekeliruan atau galat jenis I (α). Begitu pula sebaliknya jika kita menyimpulkan untuk menerima H_0 padahal H_0 salah, maka kita telah melakukan kekeliruan yang disebut kekeliruan atau galat jenis II (β).

Jika nilai α diperkecil, maka akan menjadi β besar. Nilai α biasanya ditetapkan sebesar 0,05 atau 0,01. Jika $\alpha = 0,05$, artinya 5 dari setiap 100 kesimpulan kita akan menolak H_0 , yang seharusnya diterima. Harga $(1 - \beta)$ disebut Kuasa Uji atau Kekuatan Uji.

Teknik dalam pengujian hipotesis dilakukan berdasarkan :

- Pengujian Satu Pihak

$H_0 : \alpha = \alpha_0$

$H_1 : \alpha > \alpha_0$

$H_1 : \alpha < \alpha_0$

- Pengujian Dua Pihak

$H_0 : \alpha = \alpha_0$

$H_1 : \alpha \neq \alpha_0$

1. Pengujian Rata-rata Satu Sampel

Pengujian rata-rata satu sampel dimaksudkan untuk menguji nilai tengah atau rata-rata populasi μ sama dengan nilai tertentu μ_0 , lawan hipotesis alternatifnya bahwa nilai tengah atau rata-rata populasi μ tidak sama dengan μ_0 . Jadi kita akan menguji :

$H_0 : \alpha = \alpha_0$ lawan $H_1 : \alpha \neq \alpha_0$

H_0 merupakan hipotesa awal.

2. Pengujian Rata-rata Dua Sampel

Untuk pengujian rata-rata dua sampel terdapat 2 jenis data :

- Dua Sampel Berpasangan.
Artinya kedua sampel bersifat mutually exclusive (saling asing) dan banyaknya pengamatan (ulangan) sama pada masing-masing sampel.
- Sampel Bebas / Independen.
Pada pengujian rata-rata dua sampel berpasangan, banyaknya nilai pengamatan harus sama ($n_1=n_2$), sedangkan pada dua sampel yang bebas banyaknya pengamatan tidak harus sama.

C. PRAKTEK

1. Pengujian Rata-rata Satu Sampel

Seorang mahasiswa melakukan penelitian mengenai galon susu murni yang rata-rata isinya 10 liter. Telah diambil sampel secara acak dari 10 botol yang telah diukur isinya, dengan hasil sebagai berikut : 10,2 ; 9,7 ; 10,1 ; 10,3 ; 10,1 ; 9,8 ; 9,9 ; 10,4 ; 10,3 ; 9,8. Dengan $\alpha = 0,01$

Analisa secara manual :

1. Hipotesis $H_0 : \alpha = 10$ lawan $H_1 : \alpha \neq 10$
2. Uji statistik t (karena α tidak diketahui atau $n < 30$).
3. $\alpha = 0.01$
4. Wilayah kritik : $t < t_{\alpha/2(n-1)}$ atau $t > t_{\alpha/2(n-1)}$.
5. Perhitungan, dari data : rata-rata $\bar{x} = 10.06$ dan simpangan baku sampel $s = 0.2459$.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = 0,772$$

Karena $t = 0,772$ terletak diantara -3,250 dan 3,250 disimpulkan untuk menerima H_0 , artinya pernyataan bahwa rata-rata isi galon susu murni 10 liter dapat diterima.

Analisa menggunakan SPSS :

1. Masukkan data diatas pada Data View, namun sebelumnya kita harus menentukan nama dan tipe datanya pada Variable View.
2. klik Menu Analyze Compare Means One Sample T-Test.
3. Sehingga menghasilkan hasil analisa sebagai berikut :

T-Test

[DataSet0]

One-Sample Statistics

| | N | Mean | Std. Deviation | Std. Error Mean |
|-----------|----|--------|----------------|-----------------|
| ISI GALON | 10 | 10.060 | .2459 | .0777 |

One-Sample Test

| | Test Value = 10 | | | | | |
|-----------|-----------------|----|-----------------|-----------------|---|-------|
| | t | df | Sig. (2-tailed) | Mean Difference | 99% Confidence Interval of the Difference | |
| | | | | | Lower | Upper |
| ISI GALON | .772 | 9 | .460 | .0600 | -.193 | .313 |

Keterangan hasil analisa :

- Std error = Standar Error
- T = nilai hitung
- Df = derajat kebebasan
- Sig (2-tailed)= probabilitas ($\alpha/2$)
- Mean difference = perbandingan rata-rata
- Ho diterima apabila $\text{sig} > (\alpha/2)$, Ho ditolak apabila $\text{sig} < (\alpha/2)$,

2. Pengujian Rata-rata Dua Sampel

| Mahasiswa aktif | Mahasiswa Tidak Aktif |
|-----------------|-----------------------|
| 7 | 4 |
| 8 | 3 |
| 5 | 5 |
| 3 | 4 |
| 5 | 2 |

Langkah penyelesaiannya adalah sebagai berikut :

1. Masukkan data diatas pada Data View, namun sebelumnya kita harus nenentukan nama dan tipe datanya pada Variable View.
2. Klik Analyze >> Compare Means >> Paired Samples T-Test
Maka akan muncul tampilan sebagai berikut :

Paired Samples Statistics

| Pair | | Mean | N | Std. Deviation | Std. Error Mean |
|------|-----------------------|------|---|----------------|-----------------|
| 1 | Mahasiswa Aktif | 5,60 | 5 | 1,949 | ,872 |
| | Mahasiswa Tidak Aktif | 3,60 | 5 | 1,140 | ,510 |

Paired Samples Correlations

| | | N | Correlation | Sig. |
|--------|---|---|-------------|------|
| Pair 1 | Mahasiswa Aktif & Mahasiswa Tidak Aktif | 5 | -,202 | ,744 |

Paired Samples Test

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|--------|---|--------------------|----------------|-----------------|---|-------|-------|----|-----------------|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | |
| | | | | | Lower | Upper | | | |
| Pair 1 | Mahasiswa Aktif - Mahasiswa Tidak Aktif | 2,000 | 2,449 | 1,095 | -1,041 | 5,041 | 1,826 | 4 | ,142 |

Keimpulan H_0 diterima karena $p\text{-value } \alpha/2 > 0,05$

- misal kita akan menguji sampel pada taraf nyata $\alpha = 0.05$ bahwa masa putar roda pada sepeda 1 berbeda dengan sepeda 2. Data masa putar roda (menit) kedua sepeda tersebut adalah :

| | Masa Putar (menit) | | | | | | |
|----------|--------------------|----|----|----|----|----|----|
| | sepeda 1 | 55 | 58 | 55 | 52 | 59 | |
| sepeda 2 | 65 | 59 | 60 | 64 | 62 | 65 | 58 |

Langkah penyelesaiannya adalah sebagai berikut :

- Masukkan data diatas pada Data View, namun sebelumnya kita harus menentukan nama dan tipe datanya pada Variable View.
- Klik Analyze >> Compare Means >> Independent Samples T-Test

Tugas :

1. Pengujian Rata-rata Satu Sampel

Seorang pengusaha berpendapat bahwa rata-rata penjualan perhari karyawan-karyawannya adalah sebesar Rp. 1.020,00 dengan alternatif tidak sama dengan itu. Untuk maksud pengujian pendapatnya, pengusaha tersebut melakukan wawancara terhadap 20 orang karyawannya yang dipilih secara acak. Dengan menggunakan $\alpha = 0,05$. ujliah pendapat tersebut dan berikan analisa anda. Hasil wawancaranya adalah sebagai berikut.

| Nama | Penjualan (Rp.) |
|-------|-----------------|
| aan | 1000 |
| andi | 980 |
| beril | 880 |
| bona | 970 |
| cici | 850 |
| dimas | 750 |
| erik | 770 |
| gogon | 920 |
| Hari | 870 |
| heru | 900 |

| Nama | Penjualan (Rp.) |
|------|-----------------|
| ila | 930 |
| osin | 1080 |
| mima | 1200 |
| neni | 1040 |
| sila | 1040 |
| Siqi | 850 |
| Tata | 950 |
| Tita | 1100 |
| Wina | 1110 |
| zula | 990 |

Tuliskan hasil analisisnya dibawah ini, dan apakah H_0 diterima?

2. Pengujian Rata-rata Dua Sampel

Seorang guru berpendapat bahwa tidak ada perbedaan nilai rata-rata murid kelas A dan murid kelas B, namun dengan alternatif ada perbedaan. Untuk menguji pendapat tersebut, kemudian dilakukan penelitian berdasarkan penarikan sampel secara acak dimana ada 8 murid kelas A dan 6 murid kelas B. Ternyata hasil penelitian nilai siswa adalah sebagai berikut :

Kelas A : 7,5 ; 8,5 ; 7 ; 7,3 ; 8 ; 7,7 ; 8,4 ; 8,5

Kelas B : 7 ; 6,7 ; 7,3 ; 7,5 ; 6,6

Dengan menggunakan $\alpha = 5\%$, uji pendapat tersebut.

MODUL V
ANALISA DATA KATEGORIK

A. MAKSUD DAN TUJUAN

1. MAKSUD
Menganalisa data dengan analisa data kategorik
2. TUJUAN
Mahasiswa mampu menganalisa data dengan analisa data kategorik

B. TEORI SINGKAT

Dalam bab ini kita akan membahas analisis data kategorik (analysis of categorical data) menggunakan uji khi kuadrat (chi-square test), yaitu uji khi kuadrat untuk satu variabel kategorik (one categorical variable) yang disebut uji kebaikan suai khi kuadrat (chi-square goodness-of-fit test) dan uji khi kuadrat untuk dua variabel kategorik (two categorical variable) yang disebut uji khi kuadrat untuk kebebasan (chi square test for independence).

- Dalam pengujian hipotesis, kriteria untuk menolak atau tidak menolak H_0 berdasarkan p-value adalah sebagai berikut.
 - Jika P-value $< \alpha$, maka H_0 ditolak
 - Jika P-value $> \alpha$, maka H_0 tidak dapat ditolak.
- dalam program SPSS digunakan istilah **significance** (yang disingkat **Sig**) untuk P-value; dengan kata lain **P-value = Sig**.

uji kebaikan suai khi kuadrat (χ^2)

digunakan untuk menguji apakah frekuensi data yang diamati dari suatu variable kategorik sesuai dengan frekuensi harapan (expected frequencies).

Hipotesis untuk uji kebaikan suai khi kuadrat selalu berbentuk uji hipotesis dua sisi (two-sided atau two-tailed test) dengan hipotesis:

$$H_0 : \pi_i = \pi_{i0}$$

$$H_1 : \text{tidak semua } \pi_i = \pi_{i0}$$

Dimana :

- .i = 1, 2, 3 ... k
- k = banyaknya kategori
- π_{i0} = probalitas atau proporsi atau frekuensi acuan.

Rumus chi-square test

Rumus untuk uji kebaikan suai (goodness-of-fit test) adalah sebagai berikut :

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$$

Dengan derajat kebebasan =(k-1).

Dimana :

- O_i = adalah frekuensi data yang diamati (observed frequencies).
- e_i = adalah frekuensi harapan (expected frequencies)
- k = banyaknya kategori.

Rumus untuk uji khi-kuadrat untuk kebebasan (chi-square test for independence)

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$e_{ij} = \frac{n_i \cdot n_j}{n_{..}}$$

Derajat kebebasan =(r-1)(c-1)

Dimana :

- o_{ij} : adalah frekuensi data yang di observasi pada baris ke- i kolom ke- i
- e_{ij} : adalah frekuensi harapan pada baris ke- i kolom ke- i
- n_i : jumlah frekuensi pada baris baris ke- i
- n_j : jumlah frekuensi pada baris kolom ke- i
- $n_{..}$: jumlah total frekuensi
- r : adalah jumlah baris (row)
- c : adalah jumlah kolom (column)

C. PRAKTEK

1. Uji kebaikan suai dengan frekuensi harapan sama.

Sebagai contoh uji kebaikan suai khi kuadrat untuk model dengan frekuensi harapan sama (equal expected frequencies) akan digunakan data sebagai berikut :

Sebuah perusahaan pasta gigi ingin memasarkan pasta gigi dengan rasa : strawberry, vanilla, coklat, jeruk, dan nanas. Perusahaan tersebut ingin mengetahui apakah konsumen memiliki preferensi tertentu terhadap kelima rasa pasta gigi tersebut. Perusahaan melakukan suatu survey dengan membagikan kemasan kecil dari setiap rasa pasta gigi kepada 200 konsumen. Data preferensi dari 200 konsumen terhadap kelima rasa pasta gigi tersebut adalah sebagai berikut :

| | |
|-----------------|---|
| Rasa pasta gigi | Frekuensi konsumen yang memilih rasa tersebut |
| Strawberry | 32 |
| Vanilla | 30 |
| Coklat | 28 |
| Jeruk | 58 |
| Nanas | 52 |
| total | 200 |

Apakah ada preferensi tertentu konsumen terhadap kelima rasa pasta gigi tersebut ? (gunakan $\alpha = 0,01$).

Penyelesaiannya :

Ada 5 kategori untuk variable kategorik rasa pasta gigi, $k = 5$.

Frekuensi harapan = $\frac{200}{5} = 40$.

Bentuk hipotesis :

H_0 : preferensi terhadap kelima rasa pasta gigi sama.

H_1 : preferensi terhadap kelima rasa pasta gigi tidak sama.

Atau

$H_0 : \pi_{\text{strawberry}} = \pi_{\text{vanila}} = \pi_{\text{coklat}} = \pi_{\text{jeruk}} = \pi_{\text{nanas}} = 40$

$H_1 : \text{tidak semua } \pi_{\text{strawberry}} = \pi_{\text{vanila}} = \pi_{\text{coklat}} = \pi_{\text{jeruk}} = \pi_{\text{nanas}} \text{ sama dengan } 40.$

Prosedur dengan SPSS

- a. pada lembar **variable view** kita definisikan variabel **rasa pasta gigi** dengan nama **rasa** (dengan data value '1= strawberry'; '2=vanila'; '3=coklat'; '4=jeruk'; '5=nanas') dan variabel **frekuensi** dengan nama **frekuensi**.

| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure |
|---|-----------|---------|-------|----------|-----------------|-----------------|---------|---------|-------|---------|
| 1 | rasa | Numeric | 8 | 0 | rasa pasta gigi | {1, strawberry} | None | 8 | Right | Scale |
| 2 | frekuensi | Numeric | 8 | 0 | frekuensi | None | None | 8 | Right | Scale |
| 3 | | | | | | | | | | |
| 4 | | | | | | | | | | |

- b. pada lembar **data view**, masukkan data **rasa** dan **frekuensi** :

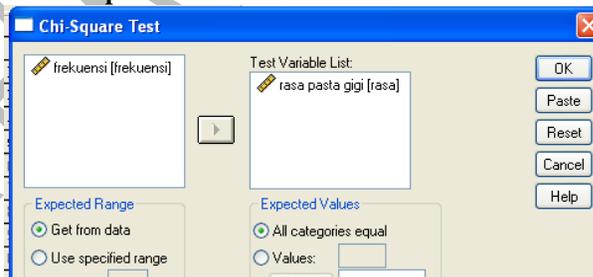
| | rasa | frekuensi |
|---|------------|-----------|
| 1 | strawberry | 32 |
| 2 | vanila | 30 |
| 3 | coklat | 28 |
| 4 | jeruk | 58 |
| 5 | nanas | 52 |
| 6 | | |
| 7 | | |

- c. klik **data** lalu pilih **weight cases**.



- d. pilih option **weight cases by**. Masukkan variabel **frekuensi**. Lalu klik **ok**.

- e. lalu klik **analyze** pilih **nonparametric Test, Chi-Square**. Pindahkan variabel **rasa pasta gigi [rasa]** ke dalam box test variable(s). klik **option**. Pilih **descriptive** lalu klik **continue**.



- f. kemudian klik **ok** maka akan di dapat hasil analisis chi-square goodness-of-fit test sebagai berikut :

Descriptive Statistics

| | N | Mean | Std. Deviation | Minimum | Maximum |
|-----------------|-----|------|----------------|---------|---------|
| rasa pasta gigi | 200 | 3.34 | 1.419 | 1 | 5 |

Chi-Square Test

Frequencies

rasa pasta gigi

| | Observed N | Expected N | Residual |
|------------|------------|------------|----------|
| strawberry | 32 | 40.0 | -8.0 |
| vanila | 30 | 40.0 | -10.0 |
| coklat | 28 | 40.0 | -12.0 |
| jeruk | 58 | 40.0 | 18.0 |
| nanas | 52 | 40.0 | 12.0 |
| Total | 200 | | |

| | |
|---------------|-----------------|
| | rasa pasta gigi |
| Chi-Square(a) | 19.400 |
| df | 4 |
| Asymp. Sig. | .001 |

a 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 40.0.

interpretasi hasil

untuk chi-square goodness-of-fit test, SPSS memberikan hasil $\chi^2 = 19.4$ dengan derajat kebebasan = $k-1 = 5-1 = 4$ dan P-value = 0.001. karena P-value=0.001 lebih kecil dari $\alpha=0.01$, maka $H_0 : \pi_{strawberry} = \pi_{vanila} = \pi_{coklat} = \pi_{jeruk} = \pi_{nanas} = 40$ di tolak.

Kesimpulan preferensi konsumen terhadap kelima rasa pasta gigi tidak sama.

2. Uji kebaikan suai dengan frekuensi harapan tak sama.

Contoh kasus untuk model ini adalah sebagai berikut :

Pabrik permen coklat kacang M&M menyatakan dalam setiap kantong permen coklat kacang kemasan 500 gram terdapat 30% permen warna coklat, 20% permen warna hijau, 20% warna merah, 20% warna kuning, dan 10% warna biru. Seorang naka membeli 1 kemasan permen tersebut dan di dalamnya terdapat 188 permen dengan rincian warna sebagai berikut ; 67 warna coklat, 24 warna hijau, 51 warna merah, 22 warna kuning, dan 24 warna biru. Gunakan taraf signifikan $\alpha=0.01$, untuk menguji apakah distribusi warna permen sesuai dengan pernyataan pabrik.

Penyelesaiannya:

Bentuk hipotesis :

$H_0 : \pi_{coklat} = 56,4; \pi_{hijau} = 37,6; \pi_{merah} = 37,6; \pi_{kuning} = 37,6; \pi_{biru} = 18,8.$

$H_1 : \text{Distribusi warna permen coklat kacang tidak sesuai dengan } \pi_{coklat} = 56,4; \pi_{hijau} = 37,6; \pi_{merah} = 37,6; \pi_{kuning} = 37,6; \pi_{biru} = 18,8.$

Dimana :

$\pi_{coklat} = 30\% \times 188 = 56,4.$

$\pi_{kuning} = 20\% \times 188 = 37,6.$

$\pi_{hijau} = 20\% \times 188 = 37,6.$

$\pi_{biru} = 10\% \times 188 = 18,8.$

$\pi_{merah} = 20\% \times 188 = 37,6.$

Prosedur dengan SPSS

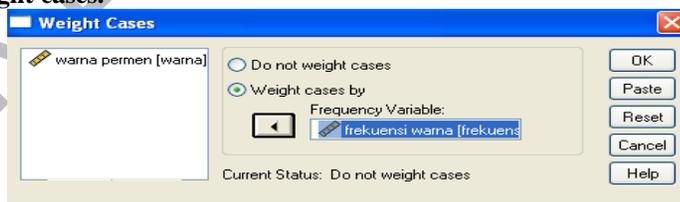
- a. pada lembar **variable view** kita definisikan variabel **warna permen coklat** dengan nama **warna** (dengan data value '1=coklat' ; '2=hijau' ; '3=merah' ; '4=kuning' ; '5=biru') dan variabel **frekuensi** dengan nama **frekuensi**.

| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure |
|---|-----------|---------|-------|----------|-----------------|---------------|---------|---------|-------|---------|
| 1 | warna | Numeric | 8 | 0 | warna permen | {1, coklat}.. | None | 8 | Right | Scale |
| 2 | frekuensi | Numeric | 8 | 0 | frekuensi warna | None | None | 8 | Right | Scale |
| 3 | | | | | | | | | | |
| 4 | | | | | | | | | | |

- b. pada lembar **data view**, masukkan data **rasa** dan **frekuensi** :

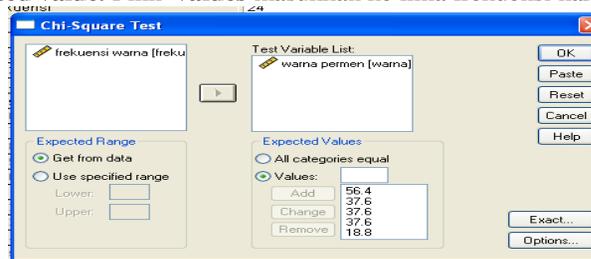
| | warna | frekuensi |
|---|--------|-----------|
| 1 | coklat | 67 |
| 2 | hijau | 24 |
| 3 | merah | 51 |
| 4 | kuning | 22 |
| 5 | biru | 24 |
| 6 | | |

- c. klik **data** lalu pilih **weight cases**.



- d. pilih option **weight cases by**. Masukkan variabel **frekuensi**. Lalu klik **ok**.

- e. lalu klik **analyze** pilih **nonparametric Test, Chi-Square**. Pindahkan variabel **warna permen [warna]** ke dalam box test variable(s). pada **Expected value**. Pilih **values** masukkan ke lima frekuensi harapan lalu klik **add**.



- f. klik menu **option**. Pilih **descriptive** lalu klik **continue**. Lalu klik **ok**.

- g. maka akan di dapat hasil analisis chi-square goodness-of-fit test sebagai berikut :

Descriptive Statistics

| | N | Mean | Std. Deviation | Minimum | Maximum |
|--------------|-----|------|----------------|---------|---------|
| warna permen | 188 | 2.53 | 1.404 | 1 | 5 |

Chi-Square Test
Frequencies

warna permen

| | Observed N | Expected N | Residual |
|--------|------------|------------|----------|
| coklat | 67 | 56.4 | 10.6 |
| hijau | 24 | 37.6 | -13.6 |
| merah | 51 | 37.6 | 13.4 |
| kuning | 22 | 37.6 | -15.6 |
| biru | 24 | 18.8 | 5.2 |
| Total | 188 | | |

Test Statistics

| | warna permen |
|---------------|--------------|
| Chi-Square(a) | 19.598 |
| df | 4 |
| Asymp. Sig. | .001 |

a 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 18.8.

interpretasi hasil

karena P-value=0.001 lebih kecil dari $\alpha=0.01$, maka $H_0 : \pi_{coklat}=56,4; \pi_{hijau} = 37,6; \pi_{merah} =37,6; \pi_{kuning} =37,6; \pi_{biru} = 18,8$. **Ditolak**

Kesimpulan distribusi warna permen coklat kacang **tidak** sesuai dengan pernyataan pabrik.

3. Uji khi kuadrat (χ^2) untuk kebebasan

Uji khi kuadrat (χ^2) untuk kebebasan (chi-square (χ^2) test for independence) atau disebut juga **contingency-table analysis** digunakan untuk menguji apakah dua variabel kategorik bersifat independen atau dependen.

Sebagai contoh kasusnya sebagai berikut :

Seorang penegak hukum mengklasifikasi tindakan kriminal menjadi dua tipe: ‘dengan kekerasan’ atau ‘tanpa kekerasan’. Suatu investigasi dilakukan untuk mempelajari tipe tindak kriminal tergantung pada usia pelaku tindak kriminal tersebut. Suatu sampel acak dari 100 pelaku tindak kriminal diambil dari data kepolisian. Data ini di tabulasi-silangkan sebagai berikut :

| Tipe tindak kriminal | Kelompok usia (dalam tahun) | | |
|----------------------|------------------------------|-------|-----------|
| | Dibawah 25 | 25-49 | 50 keatas |
| Dengan kekerasan | 15 | 30 | 10 |
| Tanpa kekerasan | 5 | 30 | 10 |

Apakah data tersebut menunjukkan bahwa tipe tindak kriminal tergantung pada usia pelaku? gunakan $\alpha =0.05$.

penyelesaiannya

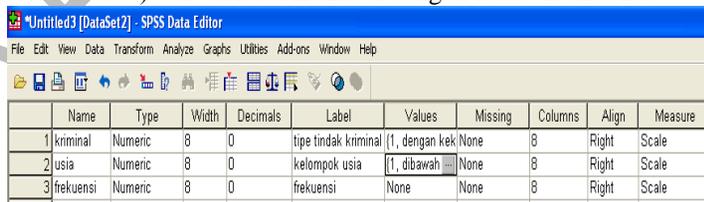
hipotesis:

H_0 : tipe tindak kriminal **tidak** tergantung pada usia pelaku.

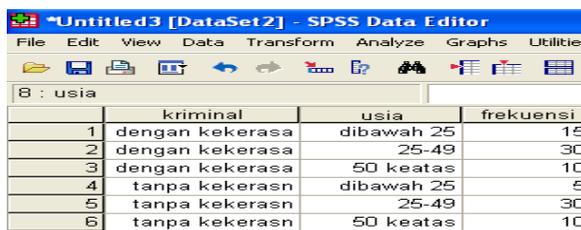
H_1 : tipe tindak kriminal tergantung pada usia pelaku.

Prosedure dengan SPSS

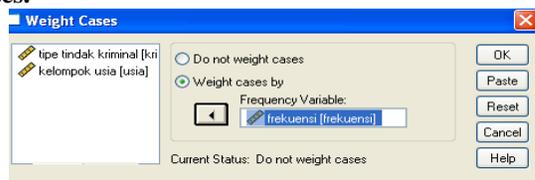
- a. pada lembar **variable view** kita definisikan variabel **tipe tindak kriminal** dengan nama **kriminal** (dengan data value ‘1=dengan kekerasan’ ; ‘2=tanpa kekerasan’), variabel **kelompok usia** dengan nama **usia** (dengan data value ‘1= dibawah 25’; ‘2=25-49’; ‘3=50 keatas’) dan variabel **frekuensi** dengan nama **frekuensi**.



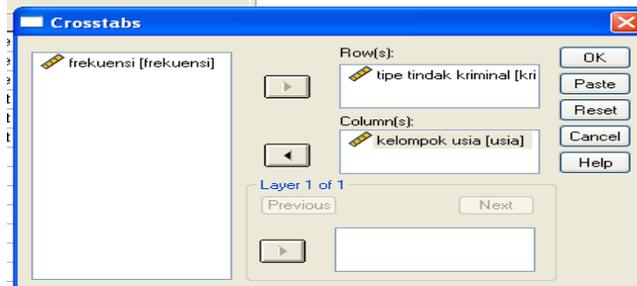
- b. kemudian pada lembar **data view** kita masukkan data sebagai berikut :
 - untuk tipe **tindak kriminal** ‘1 = dengan kekerasan’ dan **usia** ‘1 = dibawah 25’ dengan frekuensi data **15**.
 - untuk tipe **tindak kriminal** ‘1 = dengan kekerasan’ dan **usia** ‘2 = 25-49’ dengan frekuensi data **30**.
 - dan seterusnya...



- c. klik **data** lalu pilih **weight cases**.



- d. kemudian klik **analyze**, **descriptive statistics**, dan pilih **crosstabs**.
- e. pindahkan variabel **tipe tindak kriminal [kriminal]** ke box **row(s)** dan variabel **kelompok usia[usia]** ke dalam box **column(s)**.



- f. klik **cells**, pada pilihan counts pilih **observed** dan **expected**, lalu klik continue.
- g. kemudian klik **statistics**, lalu continue



- h. kemudian klik **ok**, maka akan didapat hasil sebagai berikut :

Case Processing Summary

| | Cases | | | | | |
|--------------------------------------|-------|---------|---------|---------|-------|---------|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| tipe tindak kriminal * kelompok usia | 100 | 100.0% | 0 | .0% | 100 | 100.0% |

tipe tindak kriminal * kelompok usia Crosstabulation

| | | | kelompok usia | | | Total |
|----------------------|------------------|----------------|---------------|-------|-----------|-------|
| | | | dibawah 25 | 25-49 | 50 keatas | |
| tipe tindak kriminal | dengan kekerasan | Count | 15 | 30 | 10 | 55 |
| | | Expected Count | 11.0 | 33.0 | 11.0 | 55.0 |
| | tanpa kekerasn | Count | 5 | 30 | 10 | 45 |
| | | Expected Count | 9.0 | 27.0 | 9.0 | 45.0 |
| Total | | Count | 20 | 60 | 20 | 100 |
| | | Expected Count | 20.0 | 60.0 | 20.0 | 100.0 |

Chi-Square Tests

| | Value | df | Asymp. Sig. (2-sided) |
|------------------------------|----------|----|-----------------------|
| Pearson Chi-Square | 4.040(a) | 2 | .133 |
| Likelihood Ratio | 4.231 | 2 | .121 |
| Linear-by-Linear Association | 2.500 | 1 | .114 |
| N of Valid Cases | 100 | | |

a 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.00.

interpretasi hasil

untuk chi-square test for independence, SPSS memberikan hasil $\chi^2 = 4,040$ dengan derajat kebebasan $= (r-1)(c-1) = (2-1)(3-1) = 2$ dan $P\text{-value} = 0,133$. Karena $P\text{-value} = 0,133$ lebih besar dari $\alpha = 0,05$, maka H_0 : tipe tindak kriminal **tidak** tergantung pada kelompok usia pelaku **tidak** dapat ditolak. Kesimpulan tipe tindak kriminal **tidak** tergantung pada kelompok usia pelaku.

MODUL VI
ANALISA RAGAM SATU ARAH

A. MAKSUD DAN TUJUAN

1. MAKSUD
Menganalisa data dengan analisa ragam satu arah
2. TUJUAN
Mahasiswa mampu menganalisa data dengan analisa ragam satu arah

B. TEORI SINGKAT

Analisa ragam satu arah (oneway ANOVA) digunakan untuk membandingkan mean lebih dari satu.. Bentuk hipotesisnya adalah sebagai berikut :

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

$$H_1 : \text{minimal ada dua mean popuasi yang tidak sama.}$$

C. PRAKTEK

Seorang manajer yang melakukan supervisi terhadap 3 production line tertarik akan kinerja ketiga production line tersebut. Selama 6 minggu, manajer itu mengumpulkan data jumlah produk yang cacat per 1000 unit yang produksi. Dua dari production line itu harus tutup selama 2 minggu karena harus perbaikan peralatan. Data yang di dapat manajer adalah sebagai berikut :

| Production line 1 | Production line 2 | Production line 3 |
|-------------------|-------------------|-------------------|
| 4.1 | 2.5 | 2.6 |
| 3.9 | 2.4 | 2.2 |
| 4.1 | 3.0 | 2.2 |
| 3.5 | 1.5 | 2.5 |
| | | 1.2 |
| | | 1.2 |

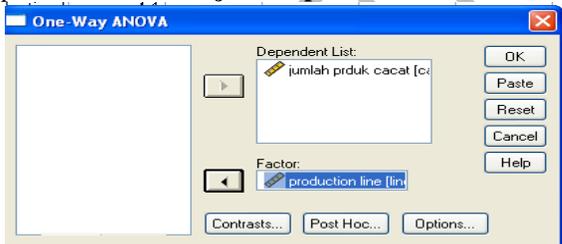
Apakah ketiga production line tersebut menghasilkan produk dengan kualitas yang sama ? gunakan $\alpha = 0,01$.

Penyelesaiannya

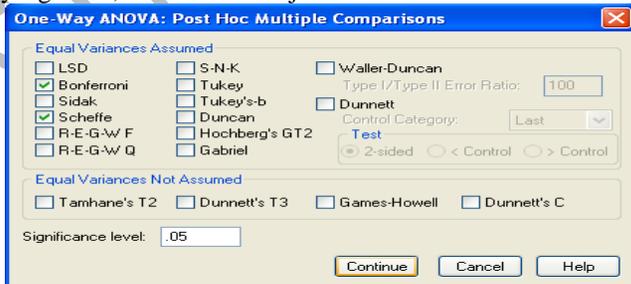
1. pada lembar **variable view** kita definisikan variabel **production line** dengan nama variabel **line** dan di beri label **production line** seta value '1 = production line 1'; '2 = production line 2'; '3 =production line 3'. Untuk variabel **jumlah produk yang cacat** gunakan nama **cacat** dan diberi label **jumlah produk cacat**.
2. kemudian pada lembar **data view** kita masukkan data **line** dan **cacat** sebagai berikut

| line | cacat |
|------|-------|
| 1 | 4.1 |
| 2 | 3.9 |
| 3 | 4.1 |
| 3 | 3.5 |
| 2 | 2.5 |
| 2 | 2.4 |
| 1 | 3.0 |
| 2 | 1.5 |
| 3 | 2.6 |
| 2 | 2.2 |
| 1 | 2.2 |
| 1 | 2.5 |
| 2 | 1.2 |
| 3 | 1.2 |

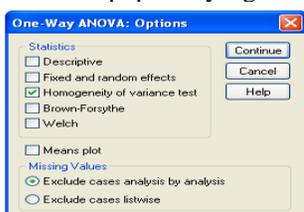
3. kemudian klik **analyze, compare means**. Lalu pilih **one-way ANOVA**.
4. pindahkan variabel **jumlah produk cacat [cacat]** ke dependen list. Dan variabel **production line** ke factor.



5. untuk menghitung **post Hoc Multiple Comparison** dengan asumsi ketiga sample **production line** memiliki ragam (variance) yang sama, klik tombol berjudul **Post Hoc**. Pilih **Bonferroni** dan **Scheffe** lalu **continue**.



6. kemudian klik **option** pilih **homogeneity of variance test** untuk menguji asumsi apakah ketiga sampel **production line** berasal dari populasi yang mempunyai ragam (variance) sama. Lalu klik **continue**.



7. kemudian klik **ok**, maka akan di dapat hasil:

Test of Homogeneity of Variances

| Levene Statistic | df1 | df2 | Sig. |
|------------------|-----|-----|------|
| 1.613 | 2 | 11 | .243 |

ANOVA

jumlah produk cacat

| | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|----|-------------|--------|------|
| Between Groups | 9.274 | 2 | 4.637 | 15.098 | .001 |
| Within Groups | 3.378 | 11 | .307 | | |
| Total | 12.652 | 13 | | | |

Post Hoc Tests

| Multiple Comparisons | | | | | | | |
|--|---------------------|---------------------|-----------------------|------------|------|-------------------------|-------------|
| Dependent Variable: jumlah prduk cacat | | | | | | | |
| | (I) production line | (J) production line | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
| | | | | | | Lower Bound | Upper Bound |
| Scheffe | production line 1 | production line 2 | 1.5500* | .3919 | .008 | .444 | 2.656 |
| | | production line 3 | 1.9167* | .3577 | .001 | .907 | 2.926 |
| | production line 2 | production line 1 | -1.5500* | .3919 | .008 | -2.656 | -.444 |
| | | production line 3 | .3667 | .3577 | .605 | -.643 | 1.376 |
| | production line 3 | production line 1 | -1.9167* | .3577 | .001 | -2.926 | -.907 |
| | | production line 2 | -.3667 | .3577 | .605 | -1.376 | .643 |
| Bonferroni | production line 1 | production line 2 | 1.5500* | .3919 | .007 | .445 | 2.655 |
| | | production line 3 | 1.9167* | .3577 | .001 | .908 | 2.925 |
| | production line 2 | production line 1 | -1.5500* | .3919 | .007 | -2.655 | -.445 |
| | | production line 3 | .3667 | .3577 | .982 | -.642 | 1.375 |
| | production line 3 | production line 1 | -1.9167* | .3577 | .001 | -2.925 | -.908 |
| | | production line 2 | -.3667 | .3577 | .982 | -1.375 | .642 |

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

Homogeneous Subsets

jumlah produk cacat

| production line | N | Subset for alpha = .05 | |
|-------------------|---|------------------------|-------|
| | | 1 | 2 |
| Scheffe(a, b) | | | |
| production line 3 | 6 | 1.983 | |
| production line 2 | 4 | 2.350 | |
| production line 1 | 4 | | 3.900 |
| Sig. | | .624 | 1.000 |

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 4.500.

b The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Interpretasi hasil

- Dari test of homogeneity of variance di dapat nilai P-value = 0,243 yang lebih besar dari $\alpha = 0,01$. Sehingga $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ tidak dapat ditolak. Kesimpulan ketiga sampel production line berasal dari populasi yang memiliki ragam sama.
- Dari tabel ANOVA di dapat nilai stistik F= 15,098 dengan derajat kebebasan K-1= 3 - 2 dan n-k=14-3 = 11 dan P-value = 0,0001. Karena P-value lebih kecil dari $\alpha = 0,01$, maka $H_0 : \mu_1 = \mu_2 = \mu_3$ ditolak. Kesimpulan ketiga production line menghasilkan produk dengan kualitas yang berbeda.
- Dari tabel homogeneous subsets terlihat bahwa subset 1 beranggotakan production line 3 dan production line 2. ini berarti production line 3 dan production line 2 tidak berbeda.

Tugas :

seseorang ingin mengetahui apakah ada perbedaan terhadap pertambahan berat badan 3 kelompok bayi berusia 2 bulan apabila di perlakukan dengan 3 cara pemberian ASI dan diperoleh datanya sebagai berikut :

| BB1 | BB2 | BB3 |
|-----|-----|-----|
| 350 | 300 | 342 |
| 375 | 300 | 310 |
| 380 | 275 | 290 |
| 370 | 290 | 370 |
| 410 | 310 | 340 |
| 400 | 330 | 330 |
| 500 | 350 | 350 |
| 475 | 290 | 400 |
| 385 | 310 | 450 |
| 420 | 305 | 325 |

Ket:

BB1 = berat badan bayi berumur 2 bulan yang hanya diberi ASI saja

BB2 = berat badan bayi berumur 2 bulan yang Tidak diberi ASI saja

BB3 = berat badan bayi berumur 2 bulan yang hanya diberi ASI dan susu bayi

Gunakan $\alpha = 0,05$.

MODUL VII
ANALISA RAGAM DUA ARAH

A. MAKSUD DAN TUJUAN

1. MAKSUD
Menganalisa data dengan analisa ragam dua arah
2. TUJUAN
Mahasiswa mampu menganalisa data dengan analisa ragam dua arah

B. TEORI SINGKAT

Analisa ragam dua arah (Twoway Analysis of Variance) digunakan untuk membandingkan mean lebih dari dua sampel yang diklasifikasikan menjadi dua factor atau dua klasifikasi. Ada tiga bentuk hipotesis analisa ragam dua arah :

Untuk factor pertama (A) :

$$H_0 : \mu'_1 = \mu'_2 = \mu'_3 = \dots = \mu'_r$$

$$H_1 : \text{minimal ada dua mean populasi yang tidak sama.}$$

Untuk factor kedua (B) :

$$H_0 : \mu''_1 = \mu''_2 = \mu''_3 = \dots = \mu''_c$$

$$H_1 : \text{minimal ada dua mean populasi yang tidak sama.}$$

Untuk interkasi anatar kedua factor A dan factor B (bila ada) ;

$$H_0 : (\mu_{AB})_1 = (\mu_{AB})_2 = (\mu_{AB})_3 = \dots = (\mu_{AB})_{r \times c}$$

$$H_1 : \text{minimal ada dua mean interkasi populasi yang tidak sama.}$$

Ket :

- μ'_1 = mean populasi baris ke-1 dari factor pertama.
- μ'_2 = mean populasi baris ke-2 dari factor pertama.
- μ'_3 = mean populasi baris ke-3 dari factor pertama.
- μ'_r = mean populasi baris ke-r dari factor pertama.
- μ''_1 = mean populasi baris ke-1 dari factor kedua.
- Dan seterusnya.

C. PRAKTEK

Data pada tabel di bawah ini adalah data hasil kuis yang didapat oleh 5 mahasiswa untuk 4 mata kuliah, yaitu : matematika, statistika, bahasa inggris, dan bahasa Indonesia. Gunakan $\alpha = 0,05$ untuk melakukan.

- a. Uji hipotesis bahwa kelima mahasiswa mempunyai kemampuan yang berbeda.
- b. Uji hipotesis bahwa keempat mata kuliah tersebut memiliki tingkat kesulitan yang berbeda.
- c. Uji hipotesis bahwa tidak ada interaksi antara mahasiswa dan mata kuliah.

| Nama mahasiswa | Nilai kuis | | | |
|----------------|------------|------------|----------------|------------------|
| | Matematika | statistika | Bahasa inggris | Bahasa indonesia |
| andy | 75 | 70 | 78 | 77 |
| | 74 | 72 | 79 | 79 |
| | 72 | 73 | 76 | 80 |
| diah | 79 | 85 | 82 | 80 |
| | 77 | 87 | 80 | 83 |
| | 78 | 88 | 81 | 81 |
| hendra | 70 | 74 | 81 | 76 |
| | 71 | 76 | 84 | 75 |
| | 72 | 77 | 83 | 73 |
| sinta | 55 | 54 | 53 | 54 |
| | 54 | 53 | 50 | 50 |
| | 53 | 56 | 52 | 53 |
| yanti | 80 | 84 | 82 | 83 |
| | 82 | 83 | 81 | 82 |
| | 84 | 81 | 85 | 80 |

Penyelesaiannya.

1. pada **variable view** kita definisikan :
 - a. variabel **mahasiswa** dengan nama **mahasiswa** dan labelnya mahasiswa serta value label ' 1 = andy';'2 = diah';'3 = hendra';'4 = sinta';'5 = yanti'. Skala variabel mahasiswa adalah **nominal**.
 - b. variabel **matakuliah** dengan nama **mtkuliah** dan labelnya mata kuliah serta value label '1 = matematika';'2 = statistika';'3 = bahasa inggris';'4 = bahasa Indonesia'. Skala variabel matakuliah adalah **nominal**.
 - c. variabel **nilai kuis** dengan nama nilai dan diberi variabel label **nilai kuis**.
2. pada lembar **data view** kita masukkan data di atas.

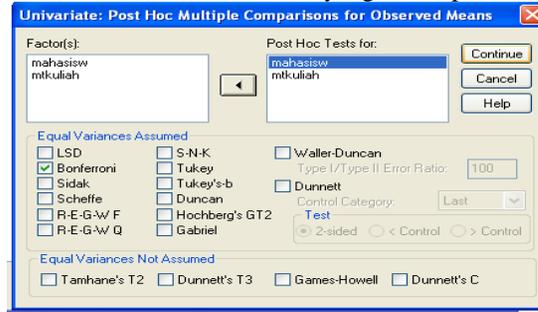
| 12 : | | | | |
|------|----------|------------|-------|----|
| | mahasisw | mtkuliah | nilai | vs |
| 1 | Andy | Matematik | 75 | |
| 2 | Andy | Statistika | 70 | |
| 3 | Andy | Bahasa Ing | 78 | |
| 4 | Andy | Bahasa Ind | 77 | |
| 5 | Andy | Matematik | 74 | |
| 6 | Andy | Statistika | 72 | |
| 7 | Andy | Bahasa Ing | 79 | |
| 8 | Andy | Bahasa Ind | 79 | |
| 9 | Andy | Matematik | 72 | |
| 10 | Andy | Statistika | 73 | |
| 11 | Andy | Bahasa Ing | 76 | |
| 12 | Andy | Bahasa Ind | 80 | |
| 13 | Diah | Matematik | 79 | |
| 14 | Diah | Statistika | 85 | |
| 15 | Diah | Bahasa Ing | 82 | |
| 16 | Diah | Bahasa Ind | 80 | |
| 17 | Diah | Matematik | 77 | |

3. kemudian klik **analyze** dan klik **general linear model** lalu pilih **univariate**.

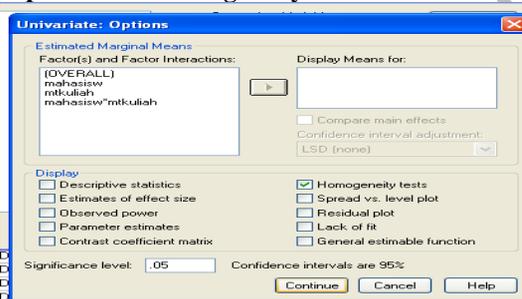
4. pindahkan variabel **nilai kuis [nilai]**. Kedependen variable dan **variabel mahasiswa [mahasiswa]** serta variabel **matakuliah[mtkuliah]** ke Fixed factor(s).



5. untuk menghitung Post Hoc Multiple Comparison terhadap variabel **mahasiswa** dan variabel **mata kuliah**. Klik **post Hoc**. Pundahkan factor mahasiswa dan mtkuliah ke post hoc test for dan kemudian pilih **Bonferroni** biladiasumsikan sama dan **Games-Howell** bila di asumsikan variance tidak sama. Disini kita memilih **Bonferroni** karena belum tahu asumsi mana yang bisa dipenuhi.



6. kemudian klik **continue**. Lalu klik **option**. Pilih **Homogeneity test**.



7. kemudian klik **continue** lalu klik **ok**. Maka akan didapat hasil sebagai berikut :

**Univariate Analysis of Variance
Between-Subjects Factors**

| | | Value Label | N |
|------------|---|------------------|----|
| Mahasiswa | 1 | Andy | 12 |
| | 2 | Diah | 12 |
| | 3 | Hendra | 12 |
| | 4 | Sinta | 12 |
| | 5 | Yanti | 12 |
| Matakuliah | 1 | Matematika | 15 |
| | 2 | Statistika | 15 |
| | 3 | Bahasa Inggris | 15 |
| | 4 | Bahasa Indonesia | 15 |

Levene's Test of Equality of Error Variances(a)
Dependent Variable: Nilai Kuis

| F | df1 | df2 | Sig. |
|------|-----|-----|------|
| .370 | 19 | 40 | .989 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept+mahasiswa+mtkuliah+mahasiswa * mtkuliah

Tests of Between-Subjects Effects

Dependent Variable: Nilai Kuis

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|----------------------|-------------------------|----|-------------|------------|------|
| Corrected Model | 7300.600(a) | 19 | 384.242 | 163.507 | .000 |
| Intercept | 325901.400 | 1 | 325901.400 | 138681.447 | .000 |
| mahasiswa | 6854.267 | 4 | 1713.567 | 729.177 | .000 |
| mtkuliah | 92.600 | 3 | 30.867 | 13.135 | .000 |
| mahasiswa * mtkuliah | 353.733 | 12 | 29.478 | 12.544 | .000 |
| Error | 94.000 | 40 | 2.350 | | |

| | | | | | |
|-----------------|------------|----|--|--|--|
| Total | 333296.000 | 60 | | | |
| Corrected Total | 7394.600 | 59 | | | |

a R Squared = .987 (Adjusted R Squared = .981)

post hoc test mahasiswa

Multiple Comparisons

Dependent Variable: Nilai Kuis
Bonferroni

| (I) Mahasiswa | (J) Mahasiswa | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---------------|---------------|-----------------------|------------|-------|-------------------------|-------------|
| | | | | | Lower Bound | Upper Bound |
| Andy | Diah | -6.33(*) | .626 | .000 | -8.19 | -4.47 |
| | Hendra | -.58 | .626 | 1.000 | -2.44 | 1.28 |
| | Sinta | 22.33(*) | .626 | .000 | 20.47 | 24.19 |
| | Yanti | -6.83(*) | .626 | .000 | -8.69 | -4.97 |
| Diah | Andy | 6.33(*) | .626 | .000 | 4.47 | 8.19 |
| | Hendra | 5.75(*) | .626 | .000 | 3.89 | 7.61 |
| | Sinta | 28.67(*) | .626 | .000 | 26.81 | 30.53 |
| | Yanti | -.50 | .626 | 1.000 | -2.36 | 1.36 |
| Hendra | Andy | .58 | .626 | 1.000 | -1.28 | 2.44 |
| | Diah | -5.75(*) | .626 | .000 | -7.61 | -3.89 |
| | Sinta | 22.92(*) | .626 | .000 | 21.06 | 24.78 |
| | Yanti | -6.25(*) | .626 | .000 | -8.11 | -4.39 |
| Sinta | Andy | -22.33(*) | .626 | .000 | -24.19 | -20.47 |
| | Diah | -28.67(*) | .626 | .000 | -30.53 | -26.81 |
| | Hendra | -22.92(*) | .626 | .000 | -24.78 | -21.06 |
| | Yanti | -29.17(*) | .626 | .000 | -31.03 | -27.31 |
| Yanti | Andy | 6.83(*) | .626 | .000 | 4.97 | 8.69 |
| | Diah | .50 | .626 | 1.000 | -1.36 | 2.36 |
| | Hendra | 6.25(*) | .626 | .000 | 4.39 | 8.11 |
| | Sinta | 29.17(*) | .626 | .000 | 27.31 | 31.03 |

Based on observed means.

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

Matakuliah

| Multiple Comparisons | | | | | | |
|--|------------------|-----------------------|------------|-------|-------------------------|-------------|
| Dependent Variable: Nilai Kuis Bonferroni | | | | | | |
| (I) Matakuliah | (J) Matakuliah | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
| | | | | | Lower Bound | Upper Bound |
| Matematika | Statistika | -2.47* | .560 | .000 | -4.02 | -.91 |
| | Bahasa Inggris | -3.40* | .560 | .000 | -4.95 | -1.85 |
| | Bahasa Indonesia | -2.00* | .560 | .006 | -3.55 | -.45 |
| Statistika | Matematika | 2.47* | .560 | .000 | .91 | 4.02 |
| | Bahasa Inggris | -.93 | .560 | .620 | -2.49 | .62 |
| | Bahasa Indonesia | .47 | .560 | 1.000 | -1.09 | 2.02 |
| Bahasa Inggris | Matematika | 3.40* | .560 | .000 | 1.85 | 4.95 |
| | Statistika | .93 | .560 | .620 | -.62 | 2.49 |
| | Bahasa Indonesia | 1.40 | .560 | .099 | -.15 | 2.95 |
| Bahasa Indonesia | Matematika | 2.00* | .560 | .006 | .45 | 3.55 |
| | Statistika | -.47 | .560 | 1.000 | -2.02 | 1.09 |
| | Bahasa Inggris | -1.40 | .560 | .099 | -2.95 | .15 |

Based on observed means.

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

Interpretasi hasil

- Dari tabel **levens test of equality of error variance** yang menguji hipotesis
 H_0 : variance diasumsikan sama
 H_1 : variance diasumsikan tidak sama
 Didapat hasil P-value = 0,989 yang lebih besar dari $\alpha = 0,05$ sehingga H_0 : variance diasumsikan sama **tidak** dapat ditolak. Kesimpulan variance sama.
- Dari tabel ANOVA di dapat nilai statistik untuk main effect sebagai berikut :
 - Faktor **mahasiswa** : nilai uji F = 729,177 dengan derajat kebebasan r -1 = 5 -1 = 4 dan rc(n-1) = 5x4(3-1) = 40 serta P-value=0,000. Karena P-value lebih kecil dari $\alpha=0,05$, maka $H_0 : \mu_{andy} = \mu_{diah} = \mu_{hendra} = \mu_{sinta} = \mu_{yanti}$ **ditolak**. Kesimpulan kelima mahasiswa mempunyai kemampuan yang berbeda.
 - Faktor **matakuliah** nilai uji F = 13,135 dengan derajat kebebasan c -1 = 4-1 = 3 dan rc(n-1) = 5x4(3-1) = 40 serta P-value=0,000. Karena P-value lebih kecil dari $\alpha=0,05$, maka $H_0 : \mu_{matematika} = \mu_{statistika} = \mu_{bahasainggris} = \mu_{bahasa indonesia}$ **ditolak**. Kesimpulan keempat mata kuliah mempunyai tingkat kesulitan yang berbeda.
 - Faktor **interaksi** : nilai uji F=12,544 dengan derajat kebebasan (r-1)(c-1) =

$(5-1)(4-1) = 12$ dan $rc(n-1) = 5 \times 4(3-1) = 40$ serta $P\text{-value} = 0,000$. Karena $P\text{-value}$ lebih kecil dari $\alpha = 0,05$, maka $H_0 : (\mu_{mhs-mtkliah})_1 = (\mu_{mhs-mtkliah})_2 = \dots = (\mu_{mhs-mtkliah})_{20}$ **ditolak**. Kesimpulan ada interaksi diantara kedua faktor.

Tugas :

Suatu perusahaan roti menduga bahwa penataan roti pada etalase yang dipajang mempengaruhi penjualan roti. Penataan roti pada etalase meliputi tinggi (A : A_1, A_2, A_3) dan lebar (B_1, B_2, B_3). Apabila tingkat penjualan di ukur dari banyaknya roti yang terjual dan asumsikan $\alpha = 0,05$ diperoleh data sebagai berikut :

| Faktor A | Faktor B | |
|----------------|------------------|---------------------|
| | B_1 (regular) | B_2 (non regular) |
| A_1 (dasar) | 47 | 46 |
| | 43 | 40 |
| A_2 (tengah) | 62 | 67 |
| | 68 | 71 |
| A_3 (tengah) | 41 | 42 |
| | 39 | 46 |

STMIK El Rahma Yogyakarta

MODUL VIII
ALALISIS REGRESI SEDERHANA

A. MAKSUD DAN TUJUAN

1. MAKSUD
Menganalisa data dengan analisa regresi sederhana
2. TUJUAN
Mahasiswa mampu menganalisa data dengan analisa regresi sederhana

B. TEORI SINGKAT

Analisis regresi sederhana digunakan untuk memprediksi nilai suatu variabel dependen y berdasarkan nilai variabel independen x. analisis regresi juga dapat digunakan untuk melihat pengaruh variabel independen x terhadap variabel dependen y. variabel independen x sering disebut sebagai variabel prediktor, sedangkan variabel dependen y sering disebut sebagai variabel respons.

C. PRAKTEK

Sebuah perusahaan jeans memperkirakan bahwa iklan di televisi akan meningkatkan permintaan produk jeans perusahaan tersebut. Bagian marketing perusahaan tersebut membuat model persamaan regresi untuk memprediksi permintaan produk berdasarkan biaya iklan yang pernah dianggarkan dan digunakan selama 19 tahun terakhir seperti tercantum dalam tabel berikut :

| Jumlah permintaan jeans (dlm ribuan) | Biaya iklan (dlm puluhan juta) |
|--------------------------------------|---------------------------------|
| 94 | 0.473 |
| 96 | 0.753 |
| 95 | 0.929 |
| 95 | 0.939 |
| 94 | 0.832 |
| 95 | 0.983 |
| 94 | 1.049 |
| 104 | 1.178 |
| 104 | 1.176 |
| 106 | 1.292 |
| 108 | 1.403 |
| 110 | 1.499 |
| 113 | 1.529 |
| 113 | 1.599 |
| 118 | 1.749 |
| 115 | 1.746 |
| 121 | 1.897 |
| 127 | 2.040 |
| 131 | 2.231 |

Tentukan persamaan regresi untuk data diatas. Apakah regresi yang didapat signifikan. Gunakan $\alpha = 0,05$.

Penyelesaiannya

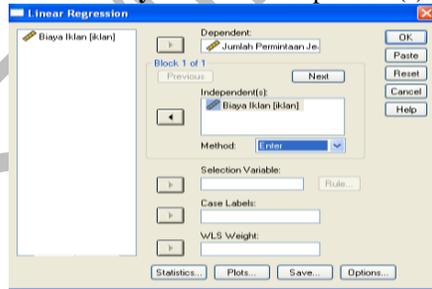
Bentuk hipotesis untuk menguji koefisien b_0 dan b_1

$$H_0 : \beta_0 = 0, H_0 : \beta_0 \neq 0 \text{ Dan } H_1 : \beta_1 = 0, H_0 : \beta_1 \neq 0$$

Dimana β_0 dan β_1 adalah parameter dari model regresi $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$

Prosedure dengan SPSS

- a. pada lembar **variable view** kita definisikan variabel **jumlah permintaan jeans** dengan nama variabel **jeans** dan labelnya **jumlah permintaan jeans**. Untuk variabel **biaya iklan** gunakan nama **iklan** dan labelnya **biaya iklan**.
- b. pada lembar **data view** masukkan data diatas.
- c. kemudian klik **analyze, regression**. Lalu pilih **linear**. Pindahkan variabel **jumlah permintaan jeans** ke dependent dan variabel **biaya iklan** ke independent(s).



- d. pastikan anda memilih **method : enter**. Kemudian klik **ok.**, maka akan di dapat hasil sebagai berikut :
- e.

Model Summary

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|---------|----------|-------------------|----------------------------|
| 1 | .969(a) | .938 | .935 | 3.019 |

a Predictors: (Constant), Biaya Iklan

ANOVA(b)

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|-------|------------|----------------|----|-------------|---------|---------|
| 1 | Regression | 2363.055 | 1 | 2363.055 | 259.266 | .000(a) |
| | Residual | 154.945 | 17 | 9.114 | | |
| | Total | 2518.000 | 18 | | | |

- a Predictors: (Constant), Biaya Iklan
- b Dependent Variable: Jumlah Permintaan Jeans

Coefficients(a)

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|-------|-------------|-----------------------------|------------|---------------------------|--------|------------|
| | | B | Std. Error | Beta | B | Std. Error |
| 1 | (Constant) | 74.673 | 2.124 | | 35.161 | .000 |
| | Biaya Iklan | 24.280 | 1.508 | .969 | 16.102 | .000 |

- a Dependent Variable: Jumlah Permintaan Jeans

interpretasi hasil

- dari R^2 (R square) dari tabel Model Summary menunjukkan bahwa 93,8% dari variance “ jumlah permintaan jeans” dapat di jelaskan oleh perubahan dalam variabel “ biaya iklan “
- tabel ANOVA diatas mengindikasikan bahwa regresi secara statistik sangat signifikan dengan nilai $F = 259.266$ untuk derajat kebebasan $k = 1$ dan $n-k-1 = 19 - 1 - 1 = 17$ dan $P\text{-value} = 0,000$ yang jauh lebih kecil dari $\alpha = 0,05$.
- Uji F menguji secara serentak hipotesis $H_0 : \beta_1 = \beta_2 = \beta_3 = \dots \beta_k = 0$ terhadap $H_1 : \text{tidak semua } \beta_i, i = 1, 2, \dots, k \text{ sama dengan nol}$. Tetapi karena pada regresi sederhana hanya ada satu β_1 , maka kita hanya menguji $H_0 : \beta_0 = 0$ terhadap $H_1 : \beta_1 \neq 0$. Dari tabel ANOVA jelas sekali terlihat bahwa H_0 ditolak karena $P\text{-value} = 0,000$ lebih kecil dari $\alpha = 0,05$
- Persamaan garis regresi menggunakan metode kuadrat terkecil (least square method) yang didapat adalah :
 $\hat{y} = 74,673 + 24,280 x$
 Dimana \hat{y} = jumlah permintaan jeans dan x = biaya iklan.
- Untuk menguji signifikan masing-masing koefisien regresi digunakan uji statistik t . untuk menguji $\beta_1 : H_0 : \beta_1 = 0$ terhadap $H_1 : \beta_1 \neq 0$. Di dapat nilai $t = 16,102$ dengan derajat kebebasan $n - 2 = 19 - 2 = 17$ dan $P\text{-value} = 0.000$. hal ini merupakan bukti kuat penolakan $H_0 : \beta_1 = 0$, karena $P\text{-value} = 0,000$ lebih kecil dari $\alpha = 0,05$

Tugas :

sebuah penelitian tentang hubungan antara tinggi badan dengan berat badan pada siswa sebuah sekolah. Diambil secara acak 15 siswa pada sekolah tersebut dan diperoleh data sebagai berikut :

| nama | Tinggi badan (x) | Berat badan (y) |
|---------|------------------|-----------------|
| Animah | 120 | 38.4 |
| Haryadi | 126 | 41.6 |
| Maya | 135 | 46.2 |
| Ina | 135 | 49.8 |
| Dewi | 143 | 55.9 |
| Yayuk | 150 | 61.2 |
| Masihah | 150 | 59.8 |
| Mafaza | 155 | 66.5 |
| Uniana | 155 | 63.4 |
| Maruija | 155 | 65.8 |
| Rendy | 160 | 67.5 |
| Fafa | 162 | 68.7 |
| Rangga | 162 | 81.8 |
| Jaka | 170 | 75.8 |
| wahana | 172 | 78.6 |

Tentukan persamaan regresi untuk data diatas. Apakah regresi yang didapat signifikan. Gunakan $\alpha = 0,05$.

MODUL IX
ANALISA REGRESI LINEAR BERGANDA

A. MAKSUD DAN TUJUAN

1. MAKSUD
Menganalisa data dengan analisa regresi regresi linier berganda
2. TUJUAN
Mahasiswa mampu menganalisa data dengan analisa regresi linier berganda

B. TEORI SINGKAT

Analisa regresi linear berganda adalah pengembangan dari analisa regresi linear sederhana dimana terdapat lebih dari satu variabel independen x. analisa ini digunakan untuk melihat sejumlah variabel independen x_1, x_2, \dots, x_k terhadap variabel dependen y berdasarkan nilai variabel-variabel independen x_1, x_2, \dots, x_k .

C. PRAKTEK

Suatu perusahaan memiliki data usia, income sales person, dan pengalaman kerja sebagai sales. Perusahaan itu ingin membuat model regresi berganda untuk memprediksi income berdasarkan usia dan pengalaman kerja. Data dapat dilihat pada tabel dibawah ini ;

| Usia (x_1) | Pengalaman kerja (x_2) | Income (y) |
|----------------|----------------------------|------------|
| 31 | 4 | 35400 |
| 3 | 4 | 41200 |
| 38 | 5 | 45000 |
| 39 | 2 | 40300 |
| 30 | 0 | 22000 |
| 28 | 3 | 28000 |
| 20 | 0 | 13000 |
| 23 | 1 | 22000 |
| 25 | 2 | 26000 |
| 28 | 4 | 27000 |
| 29 | 5 | 30000 |

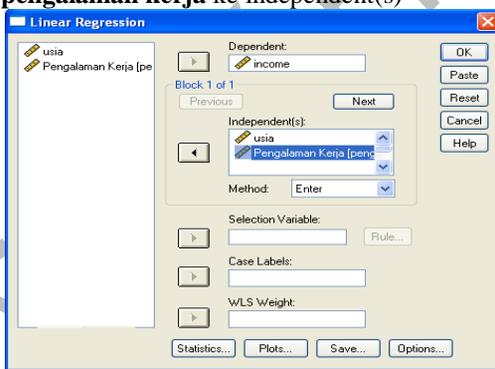
Tentukan koefisiensi dari persamaan regresi berganda dan tentukan apakah koefisiensi yang diperoleh signifikan. Lakukan pula estimasi untuk seorang sales yang berusia 40 tahun dengan pengalaman kerja 4 tahun. Gunakan $\alpha = 0,05$.

Penyelesaiannya

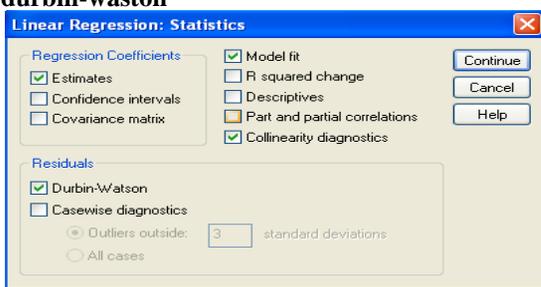
1. pada lembar **variable view** kita definisikan variabel **usia** dengan nama **usia**. Variabel **pengalaman kerja** dengan nama **pengalaman**, dan variabel **income** dengan nama **income**. Untuk variabel **pengalaman kerja** di beri label **pengalaman kerja**.
2. pada lembar **data view** kita masukkan data diatas.

| | usia | pengalam | income | var |
|----|------|----------|--------|-----|
| 1 | 31 | 4 | 35400 | |
| 2 | 37 | 4 | 41200 | |
| 3 | 38 | 5 | 45000 | |
| 4 | 39 | 2 | 40300 | |
| 5 | 30 | 0 | 22000 | |
| 6 | 28 | 3 | 28000 | |
| 7 | 20 | 0 | 13000 | |
| 8 | 23 | 1 | 22000 | |
| 9 | 25 | 2 | 26000 | |
| 10 | 28 | 4 | 27000 | |
| 11 | 29 | 5 | 30000 | |

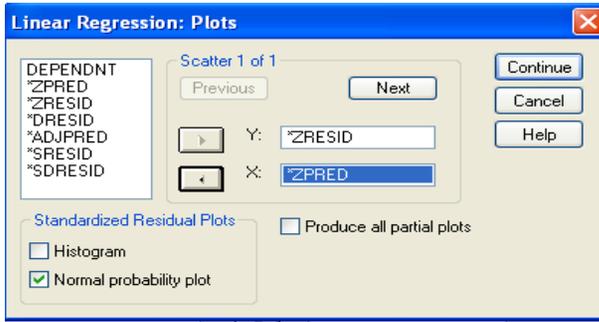
3. kemudian klik **analyze, regression**, lalu pilih **linear**. Pindahkan variabel **income** ke dependent dan variabel **usia** serta **pengalaman kerja** ke independent(s)



4. pastikan **method : enter** telah terpilih, lalu klik **statistics** dan pilih **estimates, model fit, collinearity diagnostics** dan **durbin-waston**



5. kemudian klik **continue** , lalu klik **plots**. Pilih **normal probability plot**. Kemudian pindahkan standardized residual ***BZRESID** ke dalam kotak **Y** dan standardized predicted value ***ZPRED** ke dalam kotak **X**.



6. kemudian klik **continue** , lalu klik **ok** maka akan di dapat hasil sebagai berikut :

Variables Entered/Removed(b)

| Model | Variables Entered | Variables Removed | Method |
|-------|---------------------------|-------------------|--------|
| 1 | Pengalaman Kerja, usia(a) | . | Enter |

- a All requested variables entered.
- b Dependent Variable: income

Model Summary(b)

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
|-------|---------|----------|-------------------|----------------------------|---------------|
| 1 | .970(a) | .941 | .927 | 2615.354 | 1.497 |

- a Predictors: (Constant), Pengalaman Kerja, usia
- b Dependent Variable: income

ANOVA(b)

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|-------|------------|----------------|----|---------------|--------|---------|
| 1 | Regression | 876968463.994 | 2 | 438484231.997 | 64.105 | .000(a) |
| | Residual | 54720626.915 | 8 | 6840078.364 | | |
| | Total | 931689090.909 | 10 | | | |

- a Predictors: (Constant), Pengalaman Kerja, usia
- b Dependent Variable: income

Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
|-------|------------------|-----------------------------|------------|---------------------------|--------|------|-------------------------|-------|
| | | B | Std. Error | Beta | | | Tolerance | VIF |
| 1 | (Constant) | -10360.5 | 4218.164 | | -2.456 | .040 | | |
| | usia | 1201.098 | 157.935 | .765 | 7.605 | .000 | .726 | 1.377 |
| | Pengalaman Kerja | 1663.516 | 524.950 | .319 | 3.169 | .013 | .726 | 1.377 |

- a. Dependent Variable: income

Collinearity Diagnostics^a

| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | | |
|-------|-----------|------------|-----------------|----------------------|------|------------------|
| | | | | (Constant) | usia | Pengalaman Kerja |
| 1 | 1 | 2.801 | 1.000 | .00 | .00 | .02 |
| | 2 | .183 | 3.910 | .05 | .01 | .79 |
| | 3 | .016 | 13.443 | .95 | .98 | .19 |

- a. Dependent Variable: income

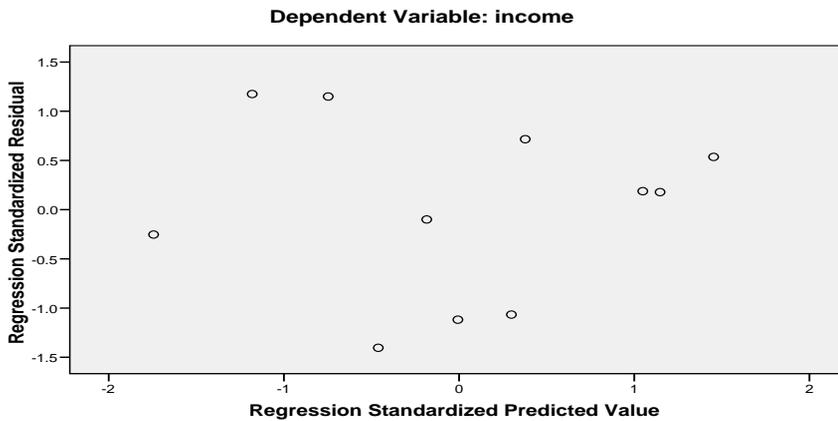
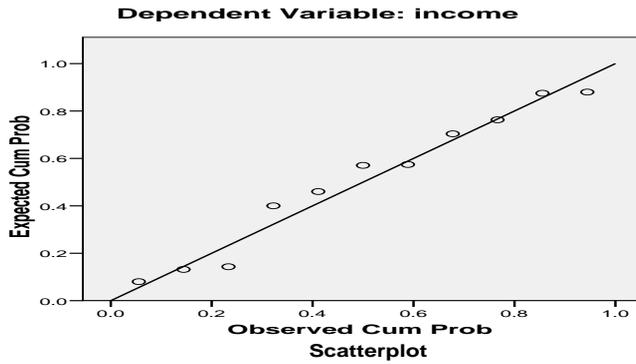
Residuals Statistics(a)

| | Minimum | Maximum | Mean | Std. Deviation | N |
|----------------------|-----------|----------|----------|----------------|----|
| Predicted Value | 13661.45 | 43598.79 | 29990.91 | 9364.659 | 11 |
| Residual | -3672.429 | 3071.740 | .000 | 2339.244 | 11 |
| Std. Predicted Value | -1.744 | 1.453 | .000 | 1.000 | 11 |
| Std. Residual | -1.404 | 1.175 | .000 | .894 | 11 |

- a Dependent Variable: income

chart

Normal P-P Plot of Regression Standardized Residual



Interpretasi hasil

- Nilai R^2 (*R square*) dari tabel model summary menunjukkan bahwa 94.1% dari variance “ **income** “ dapat di jelaskan oleh perubahan dalam variabel “ **usia**” dan “ **pengalaman kerja** “
- Nilai uji statistik durbin-watson = 1,497. Jadi dapat diasumsikan tidak terjadi autocorrelation.
- Tabel ANOVA mengidentifikasi bahwa regresi berganda secara statistic sangat signifikan dengan uji statistic $F=64,105$ dan derajat kebebasan $k = 2$ dan $n-k-1 = 11 - 2 - 1 = 8$, $P\text{-value} = 0.000$ lebih kecil dari $\alpha = 0,05$.
- Uji F menguji hipotesis $H_0 : \beta_1 = \beta_2 = 0$ terhadap $H_1 : \beta_1$ dan β_2 tidak sama dengan nol
- Dari $P\text{-value} = 0,000$ yang lebih kecil dari $\alpha = 0,05$., terlihat bahwa $H_0 : \beta_1 = \beta_2 = 0$ ditolak secara signifikan. Ini berarti koefisien β_1 dan β_2 tidak semuanya bernilai nol.
- Untuk menguji apakah masing-masing koefisien regresi signifikan, digunakan uji- t dengan hasil sebagai berikut :
 - a. Variabel **usia** $H_0 : \beta_1$ terhadap $H_1 : \beta_1 \neq 0$
 Hasil uji- $t : t = 7,605$ dengan derajat kebebasan $n-k = 11-2-1= 8$, dan $P\text{-value} = 0.000$ yang lebih kecil dari $\alpha = 0,05$.. hal ini merupakan bukti kuat penolakan $H_0 : \beta_1 = 0$.
 - b. Variabel **pengalaman kerja** : $H_0 : \beta_2 = 0$. Terhadap $H_1 : \beta_2 \neq 0$
 Hasil uji- $t . t = 3,169$ dengan derajat kebebasan $n-k = 11-2-1= 8$, dan $P\text{-value} = 0.013$ yang lebih kecil dari $\alpha = 0,05$.. hal ini merupakan bukti kuat penolakan $H_0 : \beta_2 = 0$.
 Sehingga dapat disimpulkan bahwa koefisien regresi tidak ada yang bernilai nol.
- Persamaan regresi berganda yang diperoleh dengan menggunakan metode kuadrat terkecil kriteria adalah

$$\hat{y} = -10360,5 + 1201,098x_1 + 1663,516 x_2$$
 dimana : \hat{y} = income, x_1 = usia, x_2 = pengalaman kerja.
- Dari tabel coefficients juga terlihat bahwa nilai VIF = 1,377 sehingga masih dapat dianggap tidak terjadi multicollinearity (atau tepatnya hanya low collinearity).
- Dari normal probability plot juga terlihat bahwa titik-titik data berbentuk pola linear sehingga konsisten dengan distribusi normal.

Scatterplot antara *ZRESID dan *ZPRED tidak membentuk pola tertentu, sehingga bias dianggap residual mempunyai variance konstan. (homoscedasticity)

MODUL X
ANALISA REGRESI LOGISTIK

A. MAKSUD DAN TUJUAN

1. MAKSUD
Menganalisa data dengan analisa regresi logistik
2. TUJUAN
Mahasiswa mampu menganalisa data dengan analisa regresi logistik

B. TEORI SINGKAT

Analisa regresi logistik digunakan untuk melihat pengaruh sejumlah variabel independen $x_1, x_2 \dots x_k$ terhadap variabel dependen y yang berupa variabel kategorik (binominal, multi nominal, atau ordinal) atau juga untuk memprediksi nilai suatu variabel dependen y (yang berupa variabel kategorik) berdasarkan nilai variabel-variabel independen $x_1, x_2 \dots x_k$

C. PRAKTEK

1. Regresi logistik biner

Adalah regresi logistik dimana variabel dependennya berupa variabel dikotomi atau variabel biner, misalnya sukses-gagal, ya-tidak, benar-salah dll.

Contoh kasusnya adalah sebagai berikut :

Seorang peneliti mengamati pengaruh merokok dan berat badan terhadap detak jantung saat responden beristirahat. Detak jantung responden di kategorikan menjadi **1= tinggi** dan **0=rendah**, variabel merokok dikategorikan menjadi **1=merokok** dan **0= tak merokok**. Berikut data ini datanya:

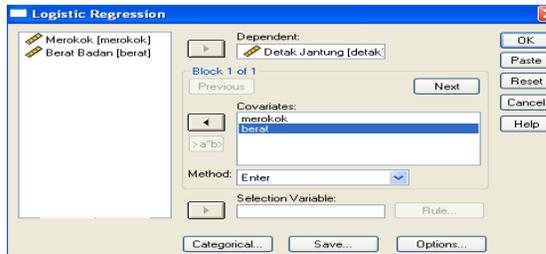
| Detak jantung 1= tinggi, 0= rendah | Merokok 1= merokok, 0= tak merokok | Berat badan kg |
|---------------------------------------|---------------------------------------|-------------------|
| 1 | 1 | 68 |
| 0 | 1 | 71 |
| 0 | 0 | 57 |
| 0 | 0 | 86 |
| 0 | 0 | 70 |
| 0 | 1 | 77 |
| 0 | 0 | 70 |
| 0 | 0 | 98 |
| 0 | 0 | 59 |
| 1 | 1 | 59 |
| 0 | 0 | 70 |
| 0 | 0 | 74 |
| 0 | 0 | 57 |
| 1 | 1 | 61 |
| 0 | 0 | 57 |
| 1 | 0 | 54 |
| 0 | 0 | 55 |
| 0 | 0 | 52 |
| 0 | 0 | 46 |
| 0 | 0 | 75 |
| 0 | 0 | 68 |
| 0 | 0 | 50 |
| 1 | 0 | 53 |
| 0 | 1 | 79 |
| 1 | 0 | 43 |
| 1 | 1 | 57 |
| 0 | 0 | 60 |
| 0 | 1 | 80 |
| 1 | 0 | 68 |
| 0 | 0 | 90 |

Tentukan persamaan regresi logistik biner dan tentukan apakah koefisien regresi yang diperoleh signifikan. Gunakan $\alpha = 0,05$

Penyelesaiannya

$n = 30$, $y =$ detak jantung, $x_1 =$ merokok, $x_2 =$ berat badan.

- a. pada lembar **variable view** kita definisikan :
 - untuk variabel **detak jantung** beri nama **detak** dan labelnya **detak jantung**, sedang value labelnya '1 = tinggi'; '2 = rendah'.
 - untuk variabel **merokok** beri nama **merokok** dan labelnya **merokok**, sedang value labelnya '1 = merokok'; '2 = tak merokok'.
 - untuk variabel **berat badan** beri nama **berat** dan labelnya **berat badan**,
- b. pada **data view** kita masukkan data di atas.
- c. kemudian klik **analyze, regression**. Pilih **binary logistic**. Pindahkan variabel **detak jantung [detak]** ke dependent dan variabel **merokok [merokok]** dan **berat badan[berat]** ke covariates.



d. pastikan **method : enter** telah terpilih. Kemudian klik **ok**, maka akan didapat hasil sebagai berikut :

Logistic Regression
Case Processing Summary

| Unweighted Cases(a) | | N | Percent |
|---------------------|----------------------|----|---------|
| Selected Cases | Included in Analysis | 30 | 100.0 |
| | Missing Cases | 0 | .0 |
| | Total | 30 | 100.0 |
| Unselected Cases | | 0 | .0 |
| Total | | 30 | 100.0 |

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

| Original Value | Internal Value |
|----------------|----------------|
| Rendah | 0 |
| Tinggi | 1 |

Block 0: Beginning Block

Classification Table(a,b)

| Observed | | | Predicted | | |
|--------------------|---------------|--------|---------------|--------|--------------------|
| | | | Detak Jantung | | Percentage Correct |
| | | | Rendah | Tinggi | Rendah |
| Step 0 | Detak Jantung | Rendah | 22 | 0 | 100.0 |
| | | Tinggi | 8 | 0 | .0 |
| Overall Percentage | | | | | 73.3 |

a. Constant is included in the model. b. The cut value is .500

Variables in the Equation

| | B | S.E. | Wald | df | Sig. | Exp(B) |
|-----------------|--------|------|-------|----|------|--------|
| Step 0 Constant | -1.012 | .413 | 6.004 | 1 | .014 | .364 |

Variables not in the Equation

| Step | Variables | Score | df | Sig. |
|------|--------------------|-------|----|------|
| 0 | merokok | 3.037 | 1 | .081 |
| | berat | 3.768 | 1 | .052 |
| | Overall Statistics | 8.143 | 2 | .017 |

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

| Step | Chi-square | df | Sig. |
|--------|------------|----|------|
| Step 1 | 11.617 | 2 | .003 |
| Block | 11.617 | 2 | .003 |
| Model | 11.617 | 2 | .003 |

Model Summary

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|------|-------------------|----------------------|---------------------|
| 1 | 23.178(a) | .321 | .468 |

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Classification Table^a

| Observed | | | Predicted | | |
|--------------------|---------------|--------|---------------|--------|--------------------|
| | | | Detak Jantung | | Percentage Correct |
| | | | Rendah | Tinggi | |
| Step 1 | Detak Jantung | Rendah | 21 | 1 | 95.5 |
| | | Tinggi | 3 | 5 | 62.5 |
| Overall Percentage | | | | | 86.7 |

a. The cut value is .500

Variables in the Equation

| Step | B | S.E. | Wald | df | Sig. | Exp(B) | |
|------|----------|-------|-------|-------|------|--------|----------|
| 1 | merokok | 3.425 | 1.560 | 4.821 | 1 | .028 | 30.712 |
| | berat | -.166 | .073 | 5.080 | 1 | .024 | .847 |
| | Constant | 8.016 | 4.018 | 3.980 | 1 | .046 | 3027.815 |

a. Variable(s) entered on step 1: merokok, berat.

Interpretasi hasil

- tabel **dependen variable encoding** menunjukkan variabel detak jantung diberi kode 1= tinggi dan 0= rendah.
- Output block : 0 beginning block
 1. Classification table menunjukkan tabel 2x2 dengan kolom berupa predicted values dari variabel dependen dan baris berupa niali data aktual yang diamati. Untuk model yang sempurna, semua cases akan terletak pada diagonal tabel dan overall percentage akan bernilai 100%. Jika model regresi logistic mempunyai variance sama, maka nilai persen (%) pad akedua baris hampir sama. Overall percentage yang memprediksi model dengan benar mempunyai nilai cukup baik sebesar $\frac{22}{30} \times 100\% = 73,3 \%$.
 2. Tabel **variables in the equation** yang hanya berisi **constant** memberikan nilai $b_0 = -1,102$ atau $\exp(-1,102) = e^{-1,102} = 0,364$. Karena responden yang mempunyai detak jantung tinggi ada 8 dan yang mempunyai detak jantung rendah ada 22, maka $\text{odd ratio} = \frac{8}{22} = 0,364$.
 3. Uji **wald** pada tabel **variables in the equation** digunakan untuk menguji apakah masing-masing koefisien regresi logistik signifikan. Uji **wald** sama dengan kuadrat dari rasio koefisien regresi logistic B dan standar error S.E . dalam contoh ini uji wald

$$= \left[\frac{B}{S.E} \right]^2 = \left[\frac{-1,012}{0,413} \right]^2 = 6,004. P\text{-value} = 0,014 \text{ lebih kecil dari } \alpha = 0,05. \text{ Maka kesimpilannya } \mathbf{constant} \text{ dari model regresi logistic ini signifikan.}$$
- Pada output block 1 : method enter
 1. Tabel **omnibus test of model coefficients** memberikan nilai chi-square goodness-of-fit test sebesar 11,617 dengan derajat kebebasan =2 $P\text{-value}=0,003$ lebih kecil dari $\alpha = 0,05$.. sehingga hasil uji ini sngat signifikan, chi-square goodness-of-fit test disini digunakan untuk menguji hipotesis :

$$H_0 : \text{memasukkan variabel independen ke dalam model tidak akan menambah kemampuan predeksi model regresi logistik}$$
 2. Tabel **model summary** memberikan nilai statistic **-2 loglikelihood = 23,178** . semakin kecil nilai **-2 loglikelihood** semakin baik.
 3. Koefisien **cox & snall R square** pada tabel **model summary** dapat diinterpretasikan sama seperti koefisien determinasi R^2 pada regresi berganda. Tetapi karena nilai **cox & snall R square** biasanya lebih kecil dari 1 maka sukar untuk di interpretasikan dan jangan digunakan
 4. Koefisien **nagelkerke R square** pada tabel **model summary** merupakan modifikasi dari koefisiensi **cox & snall R square** agar nilai maksimumnya bias mencapai satu dan mempunyai kisaran nilai antara 0 dan 1, sama seperti koefisien determinasi R^2 pada regresi linear nerganda. Nilai koefisien **nagelkerke R square** umumnya lebih besar dari koefisien **cox & snall R square** tapi cenderung lebih kecil dibandingkan dengan nilai koefisien R^2 pada regrei linear berganda. Dalam contoh ini koefisien **nagelkerke R square = 0,468**.
 5. Hasil perhitungan koefisien dari model regresi logistik biner ini terlihat pada tabel **variables in the equation** sebagai berikut :

$$\ln \left(\frac{\pi}{1-\pi} \right) = 8,016 - 0,166 \text{ berat} + 3,425 \text{ merokok}$$
 Atau

$$\frac{\pi}{1-\pi} = \exp (8,016 - 0,166 \text{ berat} + 3,425 \text{ merokok})$$
 6. kolom **Exp(B)** merupakan **odds ratio** yang diprediksi oleh model :
 - a. untuk koefisien variabel **merokok** :

$$\exp (3,425) = e^{3,425} = 30,712$$
 - b. untuk koefisien variabel **berat** :

$$\exp (-0,166) = e^{-0,166} = 0,847$$
 - c. Untuk **constant** : $\exp 8,016) = e^{8,016} = 3027,815$
 7. **uji wald** manguji masing-masing koefisien regresi logistik :
 - a. untuk koefisien variabel **merokok**:

$$= \left(\frac{B}{S.E} \right)^2 = \left(\frac{3,425}{1,560} \right)^2 = 4,821. P\text{-value} = 0,028 \text{ lebih kecil dari } \alpha = 0,05, \text{ maka koefisien regresi untuk variabel } \mathbf{merokok} \text{ signifikan.}$$
 - b. untuk koefisien variabel **berat** :

$$\left(\frac{-0,166}{0,073} \right)^2 = 0,024 P\text{-value} = 0,024 \text{ lebih kecil dari } \alpha = 0,05, \text{ maka koefisien regresi untuk variabel } \mathbf{berat} \text{ signifikan.}$$
 - c. untuk **constant** :

$$\left(\frac{8,016}{4,018} \right)^2 = 3,980. P\text{-value} = 0,046 \text{ lebih kecil dari } \alpha = 0,05, \text{ maka koefisien regresi untuk variabel } \mathbf{constant} \text{ signifikan.}$$
- 2. **Regresi logistik multinomial**

Adalah regresi logistic dimana variabel dependennya berupa variabel kategorik yang terdiri lebih dari dua nilai, seperti : merah, biru, kuning, hitam atau islem, Kristen, hindu, budha dll
- 3. **Regresi logistik ordinal**

Adalah regresi logistic dimana variabel dependennya berupa variabel dengan skala ordinal seperti : sangat setuju, setuju, netral, tidak setuju, sangat tidak setuju, atau halus, sedang, kasar. dll