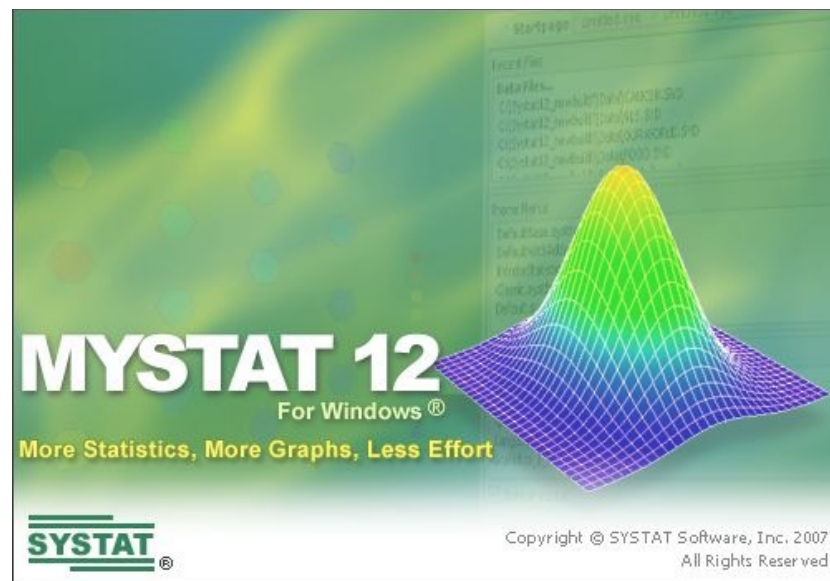


MODUL 1

ANALISIS DATA MULTIVARIAT MENGUNAKAN MYSTAT 12



Oleh:
Bertho Tantular

DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PADJADJARAN
2022

PENGANTAR MYSTAT 12

Saat ini analisis data statistika merupakan hal yang penting dalam dunia pendidikan terutama yang berkaitan dengan penelitian atau riset. Untuk melakukan suatu penelitian seorang peneliti membutuhkan data yang akan mendukung pada analisis dan pengambilan kesimpulan penelitiannya. Analisis statistika merupakan salah satu cara dalam mengolah data sehingga dapat diperoleh kesimpulan atas penelitian yang dilakukan seorang peneliti, terutama penelitian yang melibatkan data kuantitatif.

Beragam aplikasi software komputer telah dibuat untuk memudahkan seseorang untuk mengolah data dan melakukan analisis statistika. Software yang pada saat ini banyak digunakan oleh banyak kalangan adalah SPSS karena sifatnya yang *user friendly* atau mudah digunakan. Selain itu software yang juga cukup banyak peminatnya adalah Minitab, Statistica, S-Plus dan SAS meskipun lebih banyak digunakan oleh kalangan statistikawan.

Kendala terbesar dari software-software tersebut adalah lisensi yang harus dibayar kepada perusahaan pemegang lisensi software tersebut masih terlalu mahal, terutama bagi kalangan mahasiswa. Belum lagi apabila software-software tersebut memberikan jangka waktu lisensinya yang akan semakin memberatkan.

Beberapa pemegang lisensi software telah membuat software yang bisa digunakan secara gratis oleh kalangan mahasiswa atau untuk kepentingan pendidikan diantaranya R, Openstat dan MYSTAT 12.

Software R merupakan software statistika yang sangat "*powerfull*" dan merupakan software statistika yang perkembangannya paling cepat, sehingga banyak peneliti yang menggunakan software ini bahkan ikut berkontribusi terhadap software ini karena sifatnya yang *open source*. Akan tetapi software R bukan software yang *user friendly* karena membutuhkan pengetahuan tentang *syntax* dan bahasa pemrogramannya.

MYSTAT 12 merupakan software yang *user friendly* dan dapat dipergunakan secara gratis oleh kalangan mahasiswa. Software ini merupakan *student version* dari software SYSTAT. Selain *user friendly* software ini juga mempunyai tahap-tahap analisis yang mirip dengan SPSS yang saat ini (mungkin) paling banyak digunakan dikalangan mahasiswa di Indonesia. Meskipun demikian software MYSTAT 12 memiliki kemampuan analisis yang terbatas karena memang di-*setting* khusus untuk kalangan mahasiswa.

DAFTAR ISI

PENGANTAR MYSTAT 12

MODUL 1 MANAJEMEN DATA DALAM MYSTAT12

1. Pendahuluan
2. Membuat File Data
3. Penggunaan Data Editor
4. Penyimpanan Data
5. Pendefinisian Data
6. Transformasi Data

MODUL 2 ANALISIS REGRESI DAN KORELASI

1. Pendahuluan
2. Analisis Regresi Linier
3. Korelasi

MODUL 5 ANALISIS FAKTOR

1. Pendahuluan
2. Analisis Faktor Dalam MYSTAT 12

DAFTAR PUSTAKA



BAB 1

MANAJEMEN DATA DALAM MYSTAT 12

I. Pendahuluan

Sebelum informasi yang ada dapat dianalisis dengan suatu program terlebih dahulu data/informasi tersebut harus didefinisikan ke dalam bentuk-bentuk variabel, yang kemudian dimasukkan (*entry*) dalam suatu file data.

Demikian pula halnya apabila dikehendaki bahwa analisis data dilakukan dengan mempergunakan software MYSTAT 12, dalam sistem ini pun tidak terlepas dari keberadaan data yang sesuai dengan format MYSTAT 12 yang dipergunakan.

Dalam modul ini akan dibahas beberapa hal yang berkaitan dengan pendataan, antara lain:

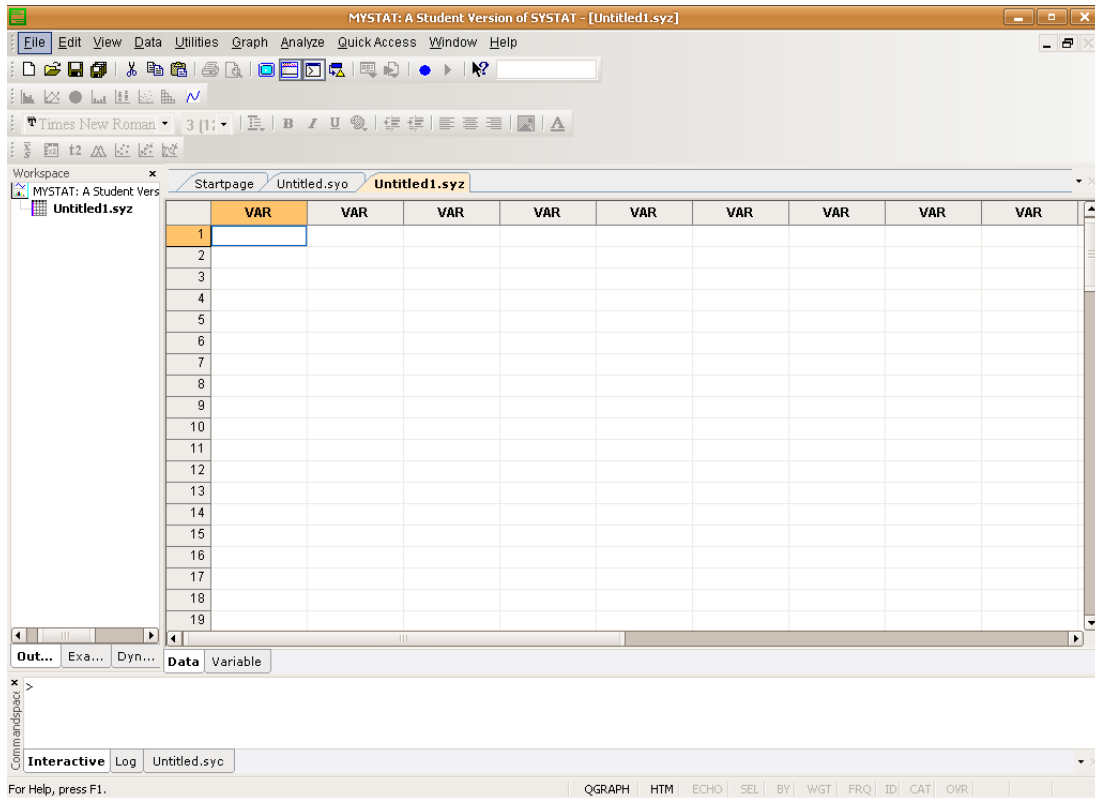
1. Membuat file data
2. Penyimpanan data dan Pemanggilan data
3. Penggunaan Data Editor
4. Pendefinisian data
5. Transformasi data

II. Membuat File Data

Untuk membuat file data dalam MYSTAT 12 bisa dalam bentuk format MYSTAT 12 atau dapat dilakukan dengan menggunakan program lain di luar MYSTAT 12 sendiri (Spreadsheet files atau ASCII text files).

1. Struktur dasar program MYSTAT 12

MYSTAT 12 merupakan program statistika yang *user friendly* yang berarti mudah dalam penggunaannya. Oleh sebab itu dalam MYSTAT 12 disediakan menu-menu untuk manajemen data, analisis statistika dan grafik. Dalam program MYSTAT 12 terdiri dari beberapa sub jendela seperti terlihat dalam Gambar dibawah. Sub jendela pertama berada disebelah kiri terdiri dari tiga buah Tab yaitu **Output Organizer**, **Example** dan **Dynamic Explorer**. Sub jendela kedua merupakan yang terbesar berada di sebelah kiri yang terdiri dari tiga buah Tab yaitu **Startpage** (informasi halaman awal MYSTAT 12), **Output** (berekstensi *.syo), dan **Data** (berekstensi *.syz). Sub jendela ketiga adalah **Command space** yaitu tempat mengisi syntax (command) juga untuk melihat log bila kita tidak menggunakan menu.



Untuk data, terdiri dari dua subtab yang berada disebelah kiri bawah, yaitu

A. Sheet data yang dapat dijelaskan sebagai berikut

- a. Tiap baris adalah sebuah observasi
- b. Tiap kolom adalah sebuah variabel
- c. Data dapat langsung di-input di dalam sel.

B. Sheet *Variables* terdiri atas

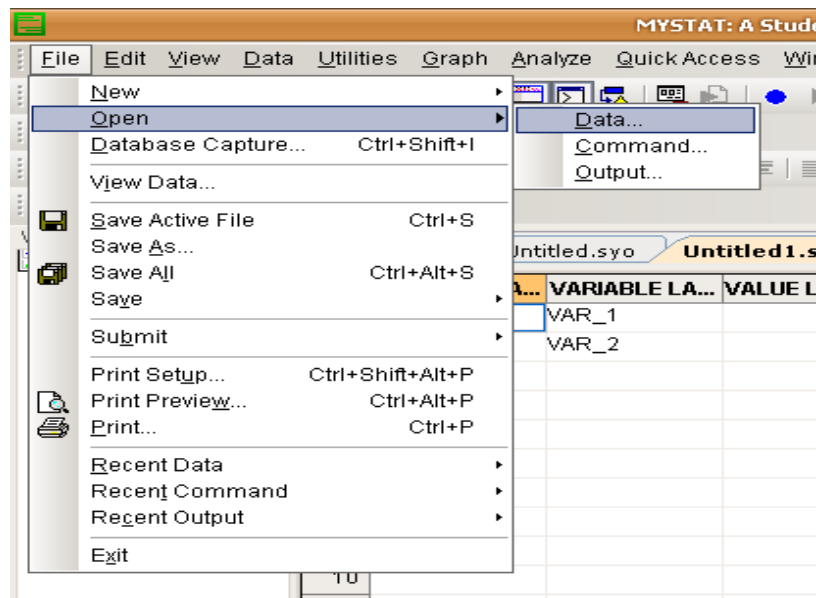
- a. *Variabel Name*
- b. *Variable Label*
- c. *Value Labels* (untuk data kategori)
- d. *Type* (Tipe variabel)
- e. dan beberapa pengaturan jumlah karakter, desimal dan kategori

2. Membuka dan *Import data*

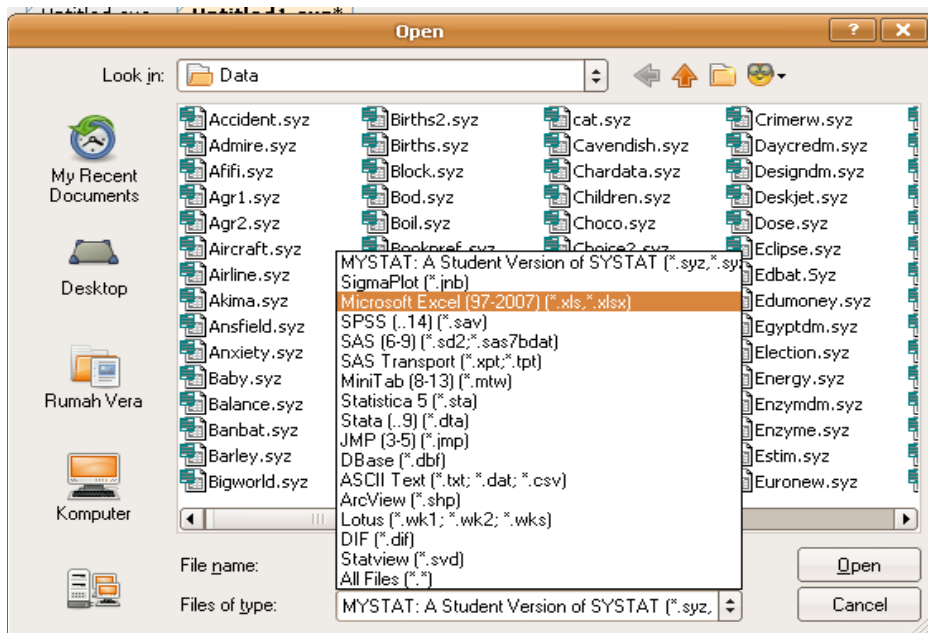
Selain input langsung didalam sheet MYSTAT 12 juga dapat membuka file data berformat MYSTAT 12 atau SYSTAT (*.syz) spreadsheet (Lotus, Excel), paket program statistik lain (SPSS, SAS, Minitab, Statistica, Stata) dan file data text (ASCII dan lain-lain) dengan cara sebagai berikut:



- a. Pilih menu *File* sub menu *Open* dan *Data...*



- b. Pilih *file of type* yang sesuai dengan ekstensioannya (misalnya *Microsoft Excel*)

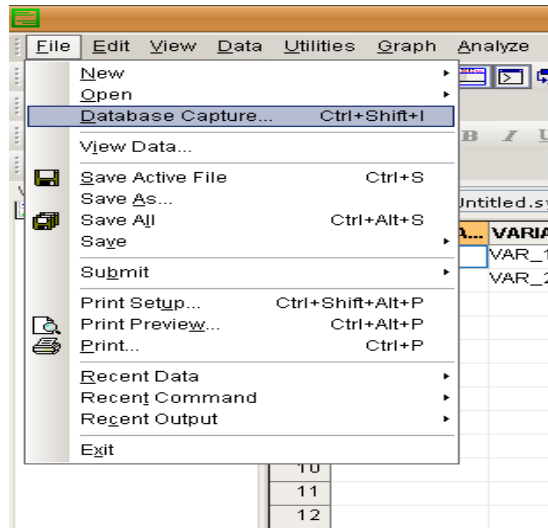


- b. Pilih file sesuai dengan nama file yang kita inginkan

2. Import data File Database Capture

Untuk membaca beberapa sumber data ODBC atau SQL dalam data editor:

- a. Pilihlah sub menu *Database Capture* dari menu *File* dan pilihlah *New Query*.



- b. Pilih sumber database yang akan diimport dan kemudian klik *Next*
- c. Pilihlah *file* yang akan diambil, kemudian ambil *field-field* yang akan di dijadikan data file MYSTAT 12.

III. Penggunaan Data Editor

A. Entri Data Numerik

1. Pilihlah sebuah sel dalam data editor
2. Masukkan nilai data, nilai-nya di display dalam sel editor di bagian atas data editor
3. Tekan Enter atau pilih sel yang lain untuk mencatat nilainya.

B. Entri Data Non-Numerik

1. Klik dua kali pada nama variabel yang terletak di atas kolom atau klik sembarang dalam kolom variabel dan dari menu pilihlah

Data

Variable Properties...

2. Pilih tipe dalam dialog box *variable type*

Untuk tipe *String* nama variabel (*Variable name*) harus diakhiri dengan \$

Untuk *categorical* bisa ber tipe *sting* maupun *numerik*

3. Klik *OK*



C. Modifikasi nilai data

Untuk menghapus nilai lama dan memasukkan nilai baru:

1. Klik sel. Nilai sel di display dalam sel editor
2. Masukkan nilai baru, itu akan merubah nilai lama dalam sel editor
3. Klik *Enter* (*move* ke sel lain) untuk mencatat nilai baru

D. Menyisipkan (Insert)

Cases diantara Cases

1. Pilih angka pada judul (nama) *case* atau baris sebelum atau sesudah posisi dimana yang kita inginkan disisipkan *case* baru.
2. Dari menu pilihlah(atau klik kanan pada nomor atau judul *case*)

Edit

Data > Insert Case(s)...

3. Isikan berapa banyak *case* yang akan disisipkan dan pilih *before* (jika akan disisipkan sebelum *case* yang dipilih) atau *after* (jika sesudahnya). Klik OK.

Variabel diantara Variabel

1. Pilih judul variabel atau kolom sebelum atau sesudah posisi dimana yang kita inginkan disisipkan variabel baru.
2. Dari menu pilihlah (atau klik kanan pada judul *variable*)

Edit

Data > Insert Variable(s)...

3. Isikan berapa banyak variabel yang akan disisipkan dan pilih *before* (jika akan disisipkan sebelum *case* yang dipilih) atau *after* (jika sesudahnya). Klik OK.

E. Memindahkan atau Menyalin *variables* dan *cases* dalam data editor

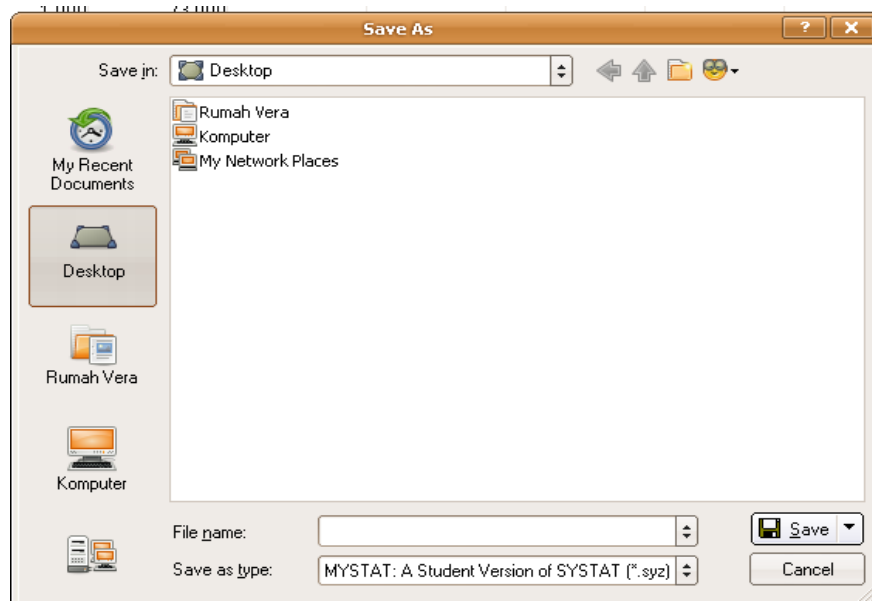
Data editor dalam MYSTAT 12 berfungsi seperti *spreadsheet* sehingga kita dapat memotong (*cut*) , menyalin (*copy*) dan memindahkannya (*paste*). Caranya dengan memblok satu atau beberapa *case* kemudian klik kanan untuk *cut*, atau *copy* dan pindahkan pada *case* yang akan ditempatkan dan klik kanan kemudian pilih *paste*. Hal yang sama dapat dilakukan untuk variabel dan sel.



IV. Penyimpanan Data

Apabila data telah dimasukkan ke dalam data editor, data tersebut perlu disimpan agar dapat dipergunakan lagi kemudian. Adapun perintah yang dipergunakan untuk menyimpan data adalah:

1. Pilih menu *File* sub menu *Save > Data* (atau tekan *Ctrl+D*)
2. Ketik nama file sesuai dengan yang diinginkan, letakkan pada *directory* (folder) yang telah ditentukan dan tentukan tipe filenya (secara default berekstensi *.syz) kemudian tekan *Save*.



V. Pendefinisian Data

A. Definisi *Variables*

Definisikan data yang ada dalam data editor dengan cara memberikan keterangan tambahan pada setiap variabel agar lebih memberikan informasi. pada outputnya, selain itu juga memberikan keterangan terhadap nilai-nilai yang diberikan dalam suatu variabel, pada output yang dihasilkannya. Perintah untuk definisi data ada pada subtab *Variables* seperti telah dijelaskan sebelumnya. Berikut penjelasan tentang bagian-bagian dalam subtab *Variables*.

1. *VARIABLE NAME*

Adalah nama untuk setiap variabel. Secara umum nama variabel harus memenuhi:

- a. Diawali dengan huruf



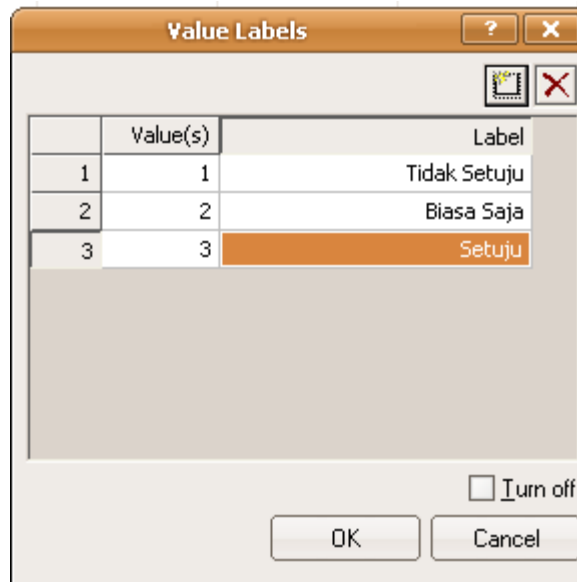
- b. boleh diikuti dengan huruf dan angka
- c. tidak boleh mengandung spasi dan karakter lain

2. VARIABLES LABELS

Berisi keterangan tentang Variabelnya apabila tidak cukup dengan nama variabel.

3. VALUE LABELS

Digunakan untuk data kategori yang telah dikoding (tipe numerik). Misalnya Variabel mempunyai tiga kategori yaitu; *Tidak Setuju*, *Biasa Saja* dan *Setuju*. Data tersebut diberi koding masing-masing 1, 2 dan 3. Cara membuat *Value Labels* adalah klik dua kali pada cell dalam baris variabel yang akan dibuat *value labels*-nya maka akan terlihat seperti gambar berikut



Untuk menambah label klik tombol sebelah kiri yang ada di pojok kanan atas. Dan untuk menghapusnya klik ebelah kanannya. kemudian isi masing-masing nilai datanya dalam tabel disebelah kiri dan labelnya disebelah kanan.

4. CATEGORICAL

Berisi keterangan tentang data berupa kategori atau bukan. Klik dua kali pada sel CATEGORY pada baris variabel yang diperhatikan. Beri nilai YES jika berupa kategori dan NO jika bukan.

5. CHARACTERS

Isi bagian ini untuk menentukan banyak karakter dalam sel yang akan ditampilkan



6. DECIMALS

Adalah banyaknya angka dibelakang koma. Isi bagian ini sesuai yang diinginkan.

7. DISPLAYS

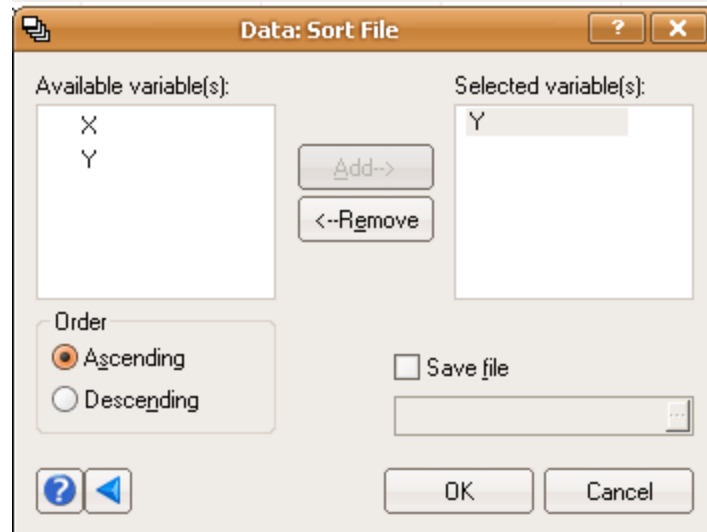
Tampilan sel dalam data editor ada tiga pilihan yaitu: *Normal* (untuk tampilan biasa), *Exponential Notation* (jika tampilan angka dalam eksponen) dan *Date Time* (untuk tampilan waktu dan tanggal).

8. DATE/TIME FORMAT

Akan berfungsi apabila DISPLAYS yang dipilih adalah Date Time. Pilih format tanggal atau waktu yang diperlukan.

B. Mengurutkan Data (*SORT FILE...*)

Untuk mengurutkan data dari kecil ke besar (*Ascending*) atau pun dari besar ke kecil (*Descending*) dapat menggunakan menu **Data** dan sub menu **Sort File...** seperti terlihat pada gambar berikut:



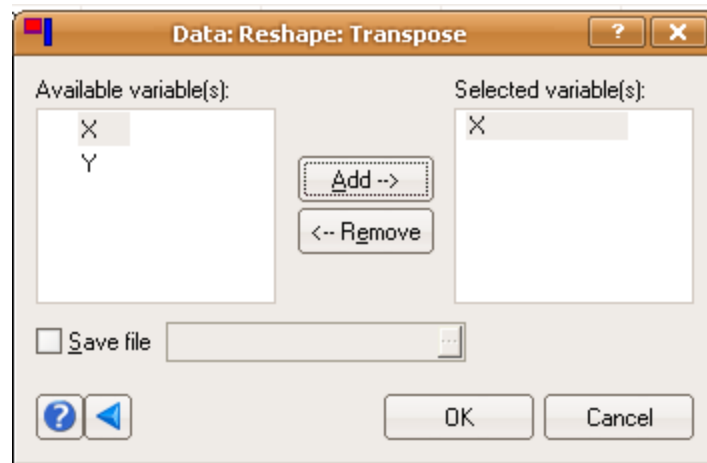
Pilih variabel yang akan diurutkan dengan memilih variabelnya kemudian menekan tombol Add--> dan pilih metode pengurutannya pada box **Order** kemudian klik **OK**.

C. Transpose

Fungsi transpose ini digunakan apabila kita ingin merubah dari observasi (cases) menjadi variabel atau sebaliknya dari variabel menjadi cases. Perintah yang dilakukan untuk mengoperasikan fungsi transpose adalah:



- a. Pilih menu *Data* sub menu *Reshape > Transpose*
- b. Pilih variabel yang akan di-transpose,



kemudian tekan *OK*.

D. Menggabungkan Variabel (*STACK*)

Fungsi *Stack* ini digunakan apabila kita ingin menggabungkan variabel-variabel kedalam satu variabel.

E. Menggabungkan cases atau variabel dari file lain (*MERGE FILES*)

Fungsi ini digunakan apabila kita ingin menggabungkan observasi (*add cases*) atau variabel (*add variable*) dari file data yang lain ke dalam data editor yang sedang aktif.

VI. Transformasi Data

Seringkali data yang sudah ada tidak dapat dipergunakan sebagai mana mestinya, sehingga perlu dilakukan modifikasi-modifikasi, misal data pada suatu file tidak akan diproses semuanya, namun hanya data laki-laki saja, ataupun yang diproses adalah data logaritmanya bukan data asli, ataupun bentuk-bentuk lainnya, yang memaksa peneliti melakukan modifikasi atau transformasi. Berikut adalah beberapa cara untuk memodifikasi ataupun transformasi data.

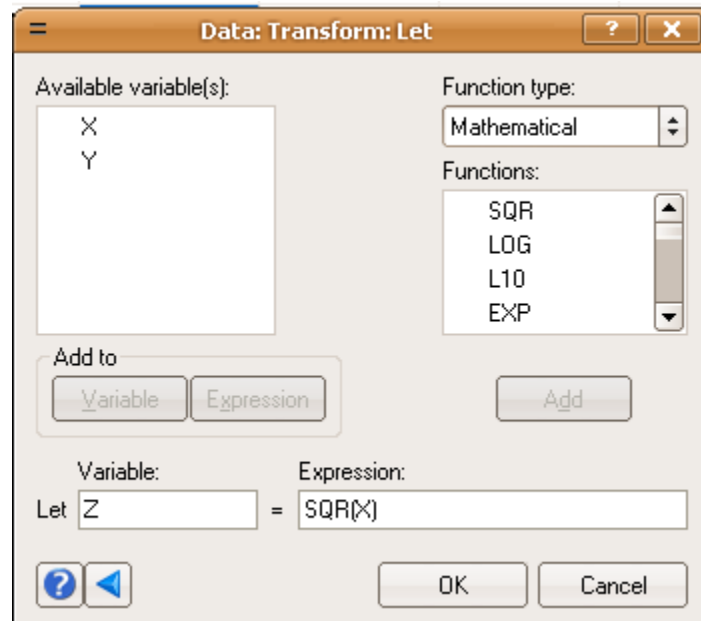
A. Menu Transform

> *Let...*

Dipergunakan bila kita ingin menghitung variabel baru dengan menggunakan variabel yang sudah ada. Langkah kerja menggunakan perintah *compute* ini adalah:



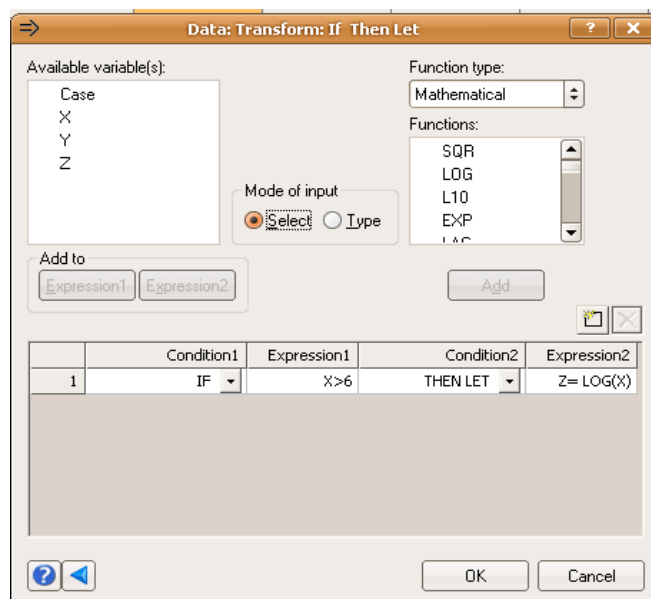
- a. Pilih menu *Data* kemudian sub menu *Transform* > *Let...* terlihat jendela seperti pada gambar dibawah ini



- b. Isi variabel hasil pada kotak *Variable* dan Fungsi matematik atau statistik (atau yang lainnya) sesuai dengan yang diinginkan kemudian pilih kode fungsi-fungsinya seperti (SQR, LOG, L10, EXP dsb), atau dapat juga diisi dengan operasi matematika. (+, -, *, /).

> *If...Then Let...*

Dipergunakan seperti *Let...* tetapi dengan kondisi tertentu yang didefinisikan sebelumnya. Misalnya variabel z bernilai 0 jika x bernilai 1 seperti terlihat pada gambar berikut.(Kondisi dapat lebih dari satu)

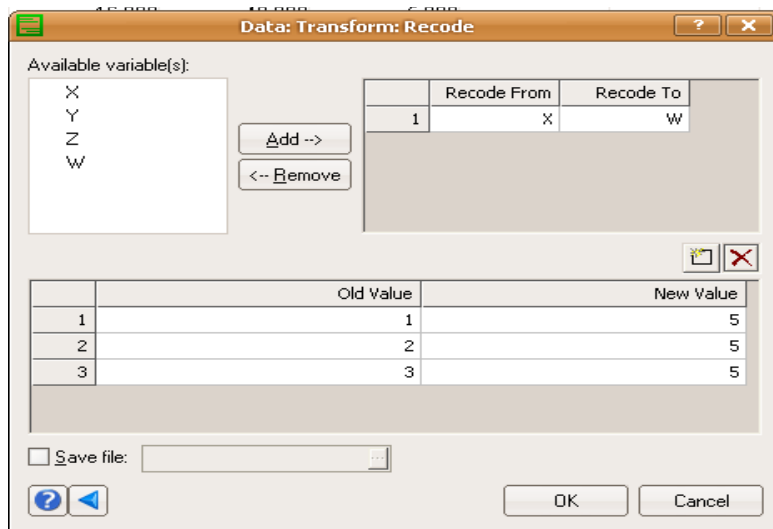




> **Recode..**

Perintah ini digunakan untuk membuat perubahan-perubahan khusus dalam nilai-nilai variabel. Langkah kerja menggunakan perintah *recode* ini adalah:

- a. Dalam *Recode*, kita bisa melakukan *recode* dalam variabel yang sama atau dalam variabel yang berbeda.
- b. Setelah memilih variabel yang akan di *recode*, klik tombol *Insert Row* (yang ada di kanan atas tabel)
- c. Sesuaikan *old value*-nya dan isilah *new values* yang sesuai.
- d. Klik *OK* apabila sudah selesai melakukan *recode*.

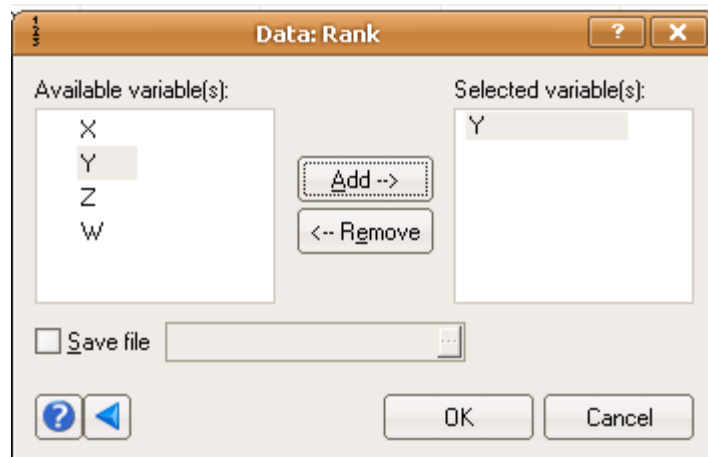


B. Membuat RANK CASES

Fungsi ini digunakan apabila akan me-ranking cases dari suatu variabel tertentu. Fungsi digunakan apabila ada hubungannya dengan statistika non parametrik.

Langkah kerja menggunakan perintah ini adalah:

- a. Pilih menu *Data* sub menu *Rank...*





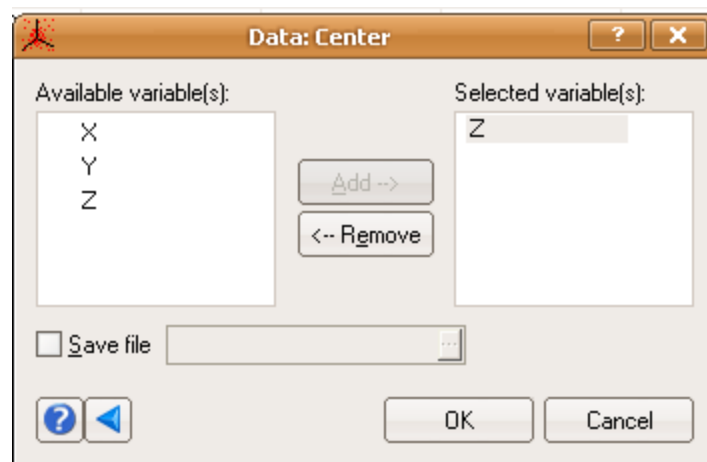
- b. Dalam sub menu ini, pilihlah variabel yang akan di-ranking
- c. Klik *OK*

Secara otomatis variabel terpilih akan berubah nilainya menjadi nilai ranking.

C. Membuat Data Baku

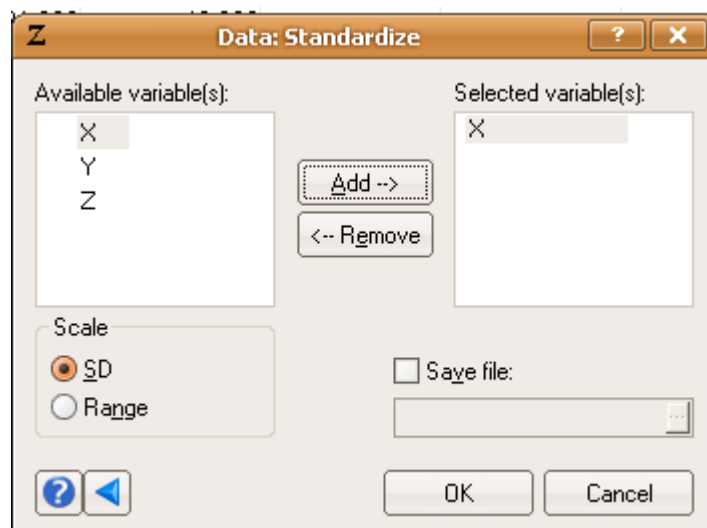
> *Center*

Fungsi ini digunakan apabila akan memusatkan data dari suatu variabel terhadap rata-ratanya.



> *Standardize*

Fungsi yang digunakan untuk membuat data baku.





BAB 2

ANALISIS REGRESI DAN KORELASI

I. Pendahuluan

Analisis Regresi adalah suatu analisis yang menyatakan hubungan fungsional antara sebuah variabel tak bebas dengan satu atau lebih variabel bebas. Regresi banyak digunakan di berbagai bidang ilmu. Sedangkan korelasi adalah digunakan suatu ukuran yang menyatakan keeratan hubungan antara dua variabel.

II. Analisis Regresi Linier dalam MYSTAT

Hubungan linier antara sebuah variabel tak bebas dengan satu variabel bebas dapat dinyatakan dalam bentuk suatu persamaan seperti berikut

$$Y = a + bX + e$$

Y adalah variabel tak bebas

X adalah variabel bebas

e adalah galat atau kekeliruan

a adalah koefisien intersep

b adalah koefisien arah

Dalam pekerjaan praktis dalam regresi yang ingin diketahui adalah nilai dari a dan b dengan beragam asumsi yang mendasarinya. Asumsi utama dalam regresi adalah bahwa e harus berdistribusi Normal. Sehingga pada awal pekerjaan kita harus menguji asumsi ini terlebih dahulu. Untuk itu perhatikan contoh berikut ini

Suatu penelitian lingkungan bertujuan untuk mengetahui tingkat pencemaran yang berasal dari mobil. Dalam hal ini diperkirakan bahwa tingkat emisi hidrokarbon (HC) dari mobil tergantung dari jaraknya. Dengan demikian, mobil yang masih baru lebih sedikit mengeluarkan HC daripada mobil tua. Untuk itu sebanyak 10 mobil merek tertentu dipilih secara acak, kemudian diperiksa berapa jarak tempuh (dalam ribuan kilometer) dari mobil tersebut dan diukur tingkat emisi HC-nya (dalam ppm). Hasilnya adalah sebagai berikut:

Jarak (x)	31	38	48	52	63	67	75	84	89	99
Emisi (y)	553	590	608	650	700	680	834	752	845	960

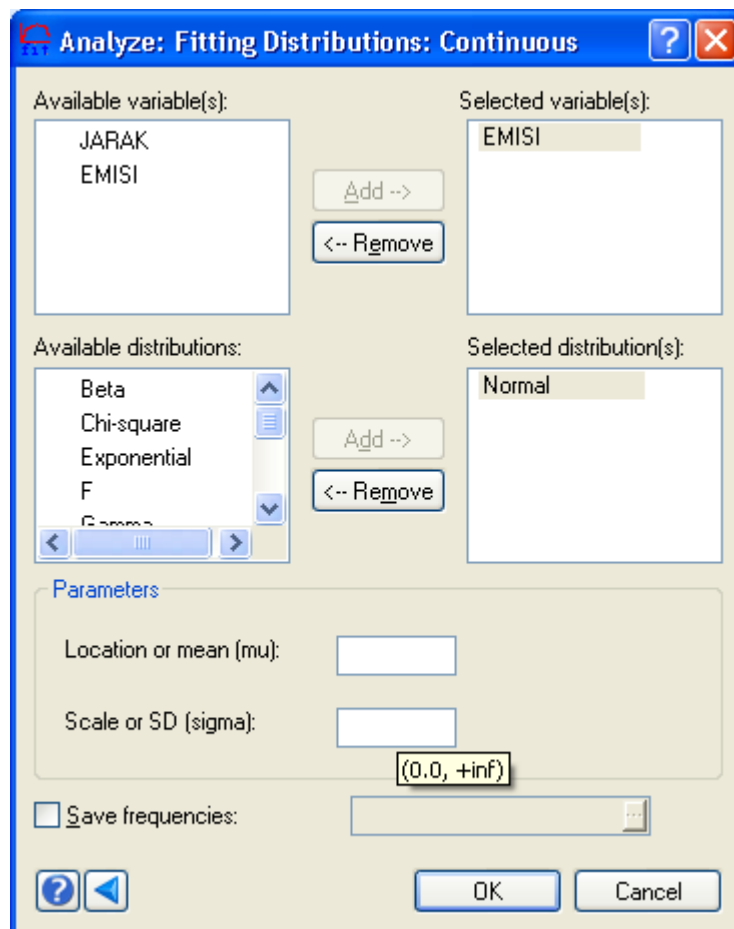
Pada taraf $\alpha = 0.05$ tentukan persamaan regresinya dan apakah Jarak tempuh mempengaruhi emisi HC?



A. Pengujian Normalitas

Untuk menguji normalitas pada kasus diatas menggunakan MYSTAT 12 adalah sebagai berikut

1. Input data terlebih dahulu
2. Pilih *Analyze* kemudian *Fitting Distribution* dan pilih *Continues...*
3. Pada jendela Distribusi pilih data yang akan diuji dalam hal ini adalah **EMISI**
4. Pilih distribusinya **Normal** kemudian klik OK.



Maka akan keluar output berikut ini

▼ [Fitting Continuous Distribution](#)

Variable Name : EMISI
Distribution : Normal

Estimated Parameter(s)

Location or Mean(mu) : 717.200
Scale or SD(sigma) : 123.045



Estimation of Parameter(s): Maximum Likelihood Method

Test Results

WARNING Chi-square test results may not be good for this sample size.

Kolmogorov-Smirnov Test Statistic : 0.156
Lilliefors Probability : 0.862

Shapiro-Wilk Test Statistic : 0.950
p-value : 0.663

Dari output terlihat hasil statistik Kolmogorov-Smirnov sebesar 0.156 dengan p-value sebesar 0.862. yang artinya data mengikuti distribusi normal.

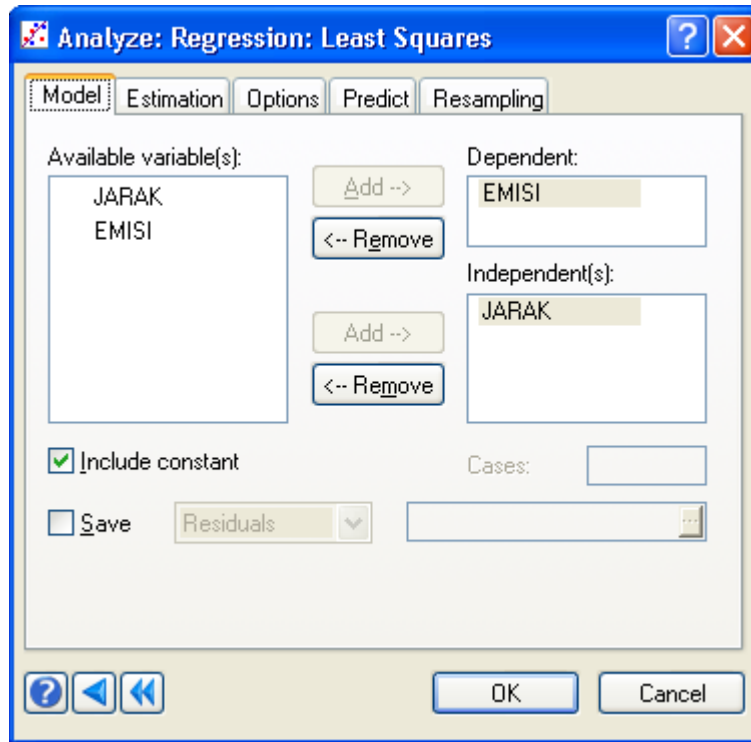
B. Plot Data

Sebelum melakukan analisis regresi sebaiknya data plot terlebih dahulu untuk melihat apakah ada pola linier yang terbentuk. Plot data digunakan sebagai langkah awal dalam menentukan regresi yang akan digunakan.

C. Menentukan Nilai a dan b

Dalam menentukan nilai a dan b (atau penaksir dari koefisien regresi) dapat menggunakan berbagai metode salah satunya adalah Metode Kuadrat Terkecil (*Least Square Method*). Berikut langkah-langkah dalam MYSTAT 12

1. Pilih *Analyze* kemudian Pilih *Regression* dan pilih *Least Square...*
2. Isi *Dependent* dengan EMISI dan *Independent* dengan JARAK
3. Klik Tab *Options* untuk menguji Normalitas cek *Kolmogorov-Smirnov*
4. Klik OK



Output yang dihasilkan adalah sebagai berikut

▼ OLS Regression

Dependent Variable	EMISI
N	10
Multiple R	0.948
Squared Multiple R	0.899
Adjusted Squared Multiple R	0.886
Standard Error of Estimate	43.814

Regression Coefficients $B = (X'X)^{-1}X'Y$

Effect	Coefficient	Standard Error	Std. Coefficient	Tolerance	p-value
CONSTANT	363.669	44.222	0.000	8.224	0.000
JARAK	5.473	0.650	0.948	1.000	8.418

Analysis of Variance

Source	SS	df	Mean Squares	F-ratio	p-value
Regression	136,042.607	1	136,042.607	70.869	0.000
Residual	15,356.993	8	1,919.624		



Test for Normality

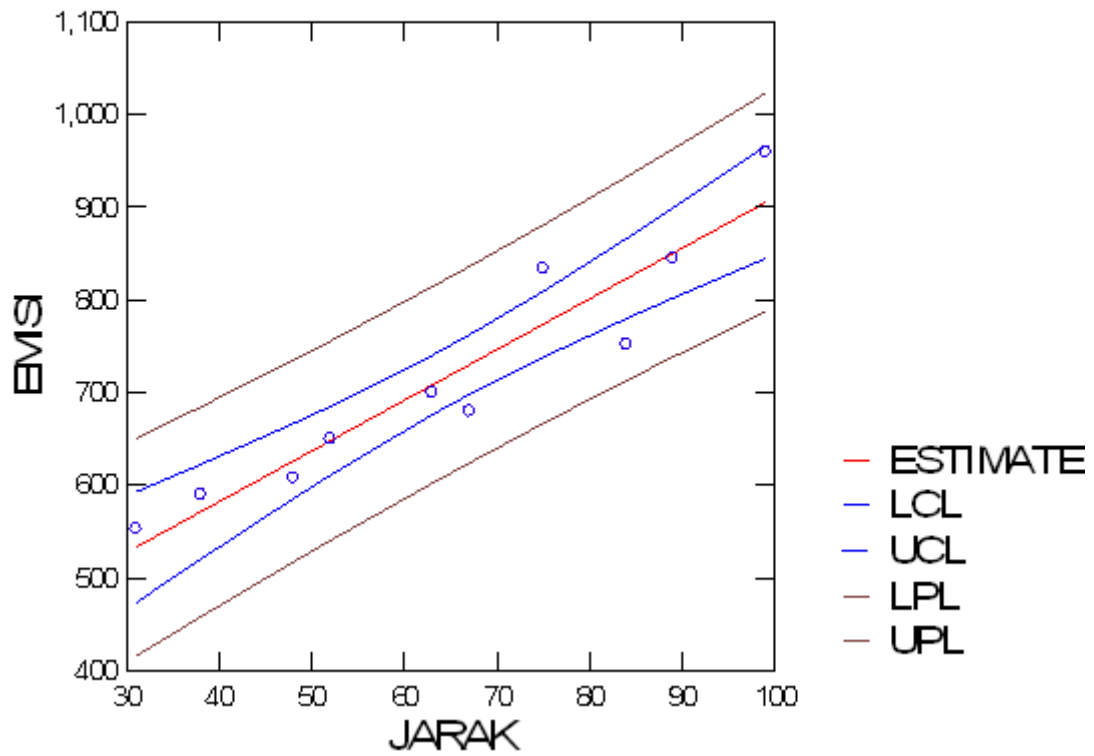
	Test Statistic	p-value
K-S Test (Lilliefors)	0.128	1.000

Durbin-Watson D Statistic	2.665
First Order Autocorrelation	-0.442

Information Criteria

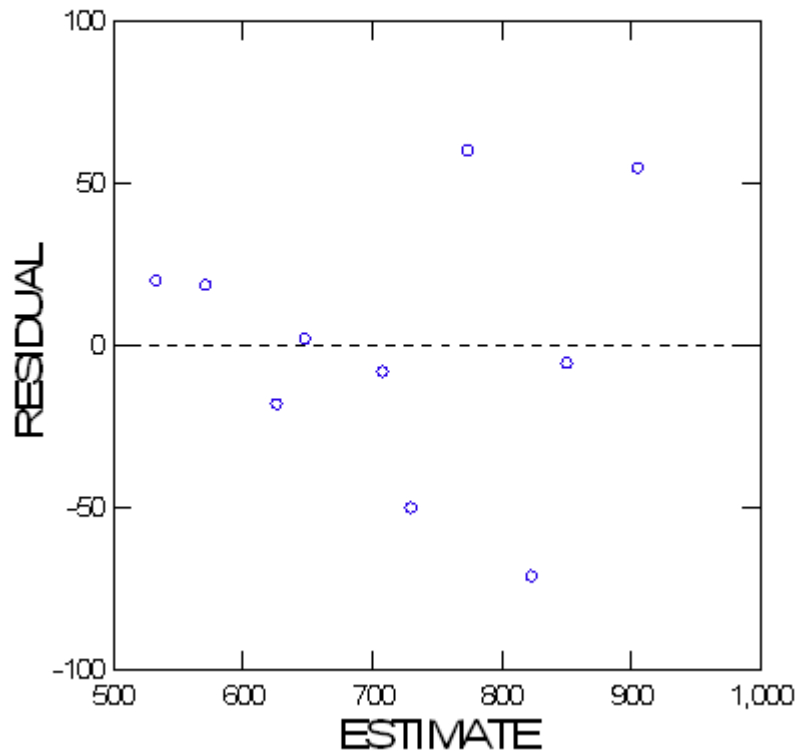
AIC	107.746
AIC (Corrected)	111.746
Schwarz's BIC	108.654

Confidence Interval and Prediction Interval





Plot of Residuals vs Predicted Values



Dari *output* terlihat pada tabel yang kedua bahwa nilai penaksir a adalah 363,669 dan nilai penaksir b adalah 5.437. Pengujian koefisien regresi secara keseluruhan terlihat pada tabel yang ketiga *Analysis of Variance* diperoleh hasil F sebesar 70.869 dengan p -value sebesar 0.000. Hasil ini sangat signifikan artinya terdapat perbedaan diantara koefisien regresi. Hal ini juga dibuktikan pada uji parsial (terlihat pada uji t dalam tabel kedua). Pada tabel pertama terlihat bahwa nilai koefisien determinasi (*square multiple R*) sebesar 0.899 yang berarti sebesar 89.9% emisi gas buang dapat dijelaskan oleh jarak tempuh dengan standard error sebesar 43.814.

Pengujian asumsi terlihat pada tabel dan grafik berikutnya.

1. Tabel untuk menguji normalitas
2. Tabel untuk menguji autokorelasi
3. Grafik nilai prediksi dan interval konfidensi untuk melihat linieritas
4. Grafik nilai prediksi dan nilai residu untuk melihat heteroskedastisitas

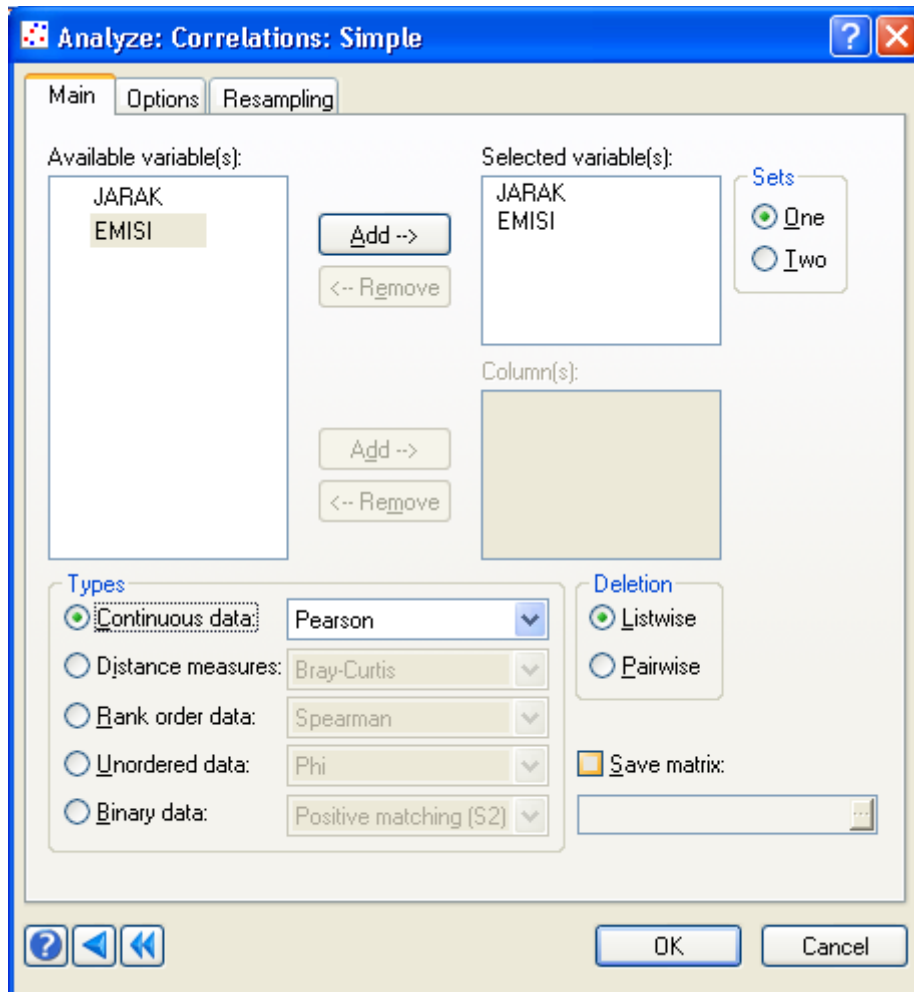
III Korelasi dalam MYSTAT

Korelasi digunakan untuk menyatakan kuat tidaknya hubungan antar dua variabel. Selain itu dari korelasi juga dapat dilihat bentuk hubungannya saling berbanding lurus (positif) atau berbanding terbalik (negatif). Nilai korelasi sangat ditentukan oleh skala ukur dari masing-masing variabel. Misalnya untuk variabel yang berskala



minimal interval (data kontinu) keduanya dapat digunakan korelasi Pearson, untuk variabel berskala ordinal (*rank order*) digunakan korelasi Spearman atau Tau Kendall dan sebagainya. Dalam MYSTAT 12 fasilitas untuk korelasi cukup lengkap untuk berbagai tipe data.

Korelasi bisa dilihat pada menu *Analyze* lalu pilih *Correlation* kemudian pilih *Simple*. Masukkan variabel-variabel yang akan dikorelasikan kedalam kotak *Selected Variables*. Pilih opsi tipe untuk tipe korelasi yang akan digunakan misalnya untuk *Continuous data* adalah *Pearson*. Kemudian Klik OK.



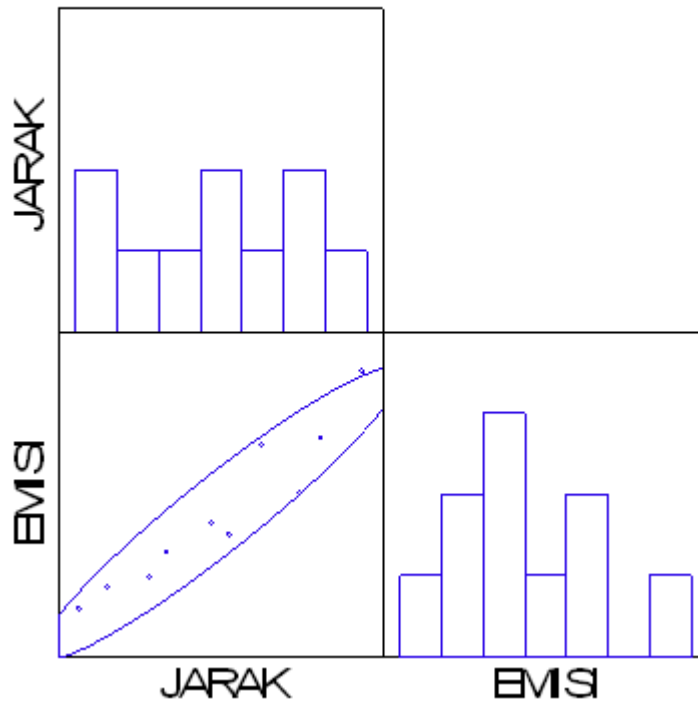
Maka akan keluar output seperti dibawah ini

▼Correlation: Pearson

Number of Observations: 10

Pearson Correlation Matrix		
	JARAK	EMISI
JARAK	1.000	
EMISI	0.948	1.000

Scatter Plot Matrix



Dari output terlihat bahwa nilai korelasi untuk Emisi dan Jarak tempuh sebesar 0.948 yang berarti hubungan kedua variabel tersebut positif dan sangat kuat. Output kedua adalah grafik untuk kedua variabel tersebut.



BAB 3

ANALISIS FAKTOR

I. Pendahuluan

Analisis Faktor adalah suatu cara menjelaskan suatu set variabel berdasarkan dimensi yang lebih umum. Pada dasarnya analisis faktor bertujuan untuk memudahkan interpretasi melalui struktur pola hubungan atau untuk mereduksi variabel. Hal ini dilakukan dengan cara mengidentifikasi struktur yang terdapat dalam set variabel yang terobservasi.

Secara umum ada tiga kegunaan utama dari Analisis Faktor yaitu:

1. Eksplorasi (disebut **Eksploratory factor analysis (EFA)**) yaitu membentuk variabel baru yang diperoleh melalui reduksi variabel
2. Konfirmasi (disebut **Confirmatory Factor Analysis (CFA)**) yaitu menguji struktur variabel yang dihipotesiskan berdasarkan banyaknya faktor yang signifikan dan besarnya faktor loading
3. Alat Pengukur (**model measurment**) yaitu pembentukan indeks-indeks yang akan digunakan sebagai pengamatan baru dalam analisa selanjutnya

Misalkan terdapat satu set variabel dengan banyak variabel adalah p dan ternyata antar variabel tersebut mempunyai tingkat korelasi yang tinggi. Dimungkinkan ada satu atau lebih variabel yang tidak terobservasi (disebut sebagai variabel laten) yang merupakan penyebab p variabel di atas. Keberadaan variabel yang tak terobservasi yang mampu menjelaskan variabel yang teramati merupakan pembahasan dalam Analisis Faktor. Dengan demikian pada intinya Analisis Faktor menjelaskan hubungan struktur kovarians dari variabel yang teramati dengan variabel yang tidak teramati.

Analisis Faktor dapat dirumuskan dalam suatu model persamaan linier. Misalkan vektor acak X dengan p komponen memiliki rata-rata μ dan matriks *covariance* Σ . Maka dapat dibentuk model persamaan faktornya adalah

$$\begin{aligned} X_1 - \mu_1 &= l_{11} F_1 + l_{12} F_2 + \dots + l_{1m} F_m + \varepsilon_1 \\ X_2 - \mu_2 &= l_{21} F_1 + l_{22} F_2 + \dots + l_{2m} F_m + \varepsilon_2 \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ X_p - \mu_p &= l_{p1} F_1 + l_{p2} F_2 + \dots + l_{pm} F_m + \varepsilon_p \end{aligned}$$



dengan:

μ_i = rata-rata variabel asal ke i

ε_i = spesifik faktor ke i

F_j = Common faktor ke j .

l_{ij} disebut *loading* dari peubah asal ke i pada faktor ke j .

Atau dalam bentuk matriks menjadi

$$\begin{array}{l} (\underline{X} - \underline{\mu}) \\ (p \times 1) \end{array} = \begin{array}{l} \underline{L} \quad \underline{F} + \underline{\varepsilon} \\ (p \times m) \quad (m \times 1) \quad (p \times 1) \end{array}$$

Asumsi yang digunakan dalam analisis faktor adalah:

- a. Data berasal dari populasi yang berdistribusi normal univariat
- b. Dalam set data terdapat multikolinieritas dapat diuji dengan menggunakan Uji Bartlett.
- c. $E(\underline{F}) = \underline{0}$, $E(\underline{\varepsilon}) = \underline{0}$
- d. $\text{Cov}(\underline{F}) = E(\underline{F}\underline{F}') = \underline{I}$
- e. $\text{Cov}(\underline{\varepsilon}) = E(\underline{\varepsilon}\underline{\varepsilon}') = \underline{\Psi} = \text{diag}(\Psi_1, \dots, \Psi_p)$
- f. \underline{F} dan $\underline{\varepsilon}$ saling bebas,
- g. $\text{Cov}(\underline{\varepsilon}, \underline{F}) = E(\underline{\varepsilon}\underline{F}') = \underline{0}$

Koefisien l_{ij} disebut loading dari variabel asal ke i pada faktor ke j , maka matriks \underline{L} adalah matriks *factor loading* dan F_1, F_2, \dots, F_m , $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ adalah tidak terobservasi.

Berikut adalah hal-hal yang perlu dilakukan dalam analisis faktor

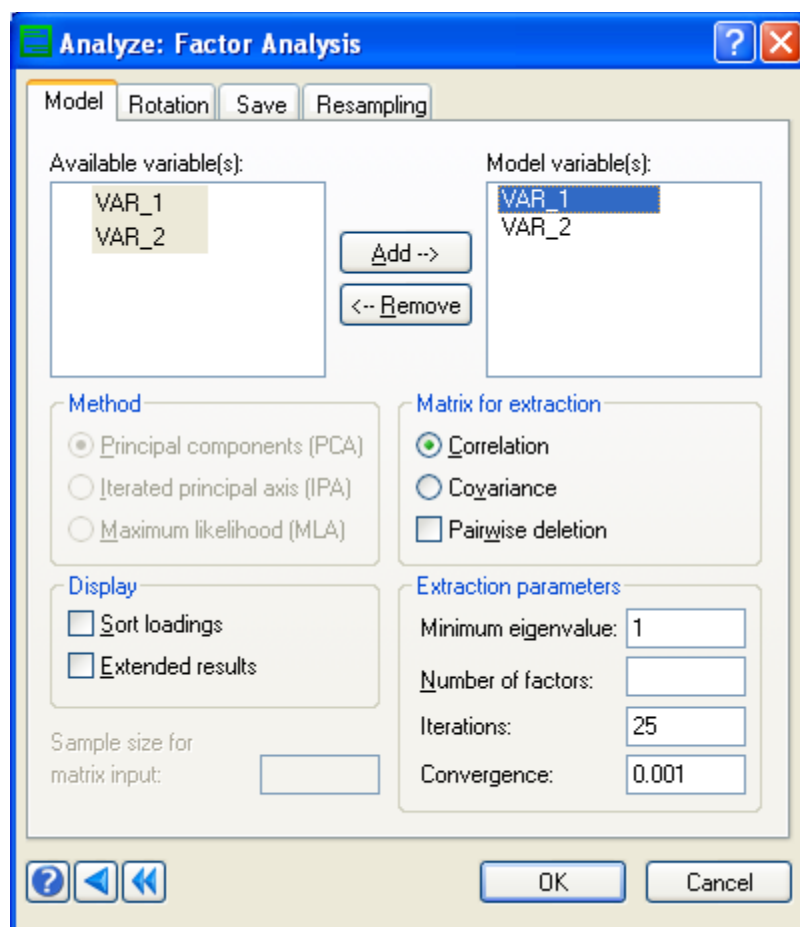
- a. Mengidentifikasi struktur
- b. Menentukan jumlah faktor (*scree plot, eigen values, proporsi varians*)
- c. Menduga parameter (*factor loading* dan *sistematik varians*)
 - Metode Komponen Utama
 - Metode Kemungkinan Maksimum
 - Metode Kuadrat Terkecil
- d. Rotasi faktor (**ortogonal**: *varimax, quartimax, equimax*; **oblique**: *oblimax, quartimin, oblimin*)
- e. Interpretasi faktor (*eigen values, explained varianses, factor scores, koefisien faktor*)



II. Analisis Faktor dalam MYSTAT

Analisis faktor dapat menggunakan metode komponen utama (*principal component*) dan metode kemungkinan maksimum (*maximum likelihood*). Tetapi dalam software Mystal hanya disediakan analisis faktor menggunakan metode komponen utama. Dalam software Mystal menyediakan opsi untuk rotasi faktor, sort data, plot faktor, dan *save* nilai *factor loading*. Selain itu juga dapat di *save* nilai *factor scores* dan koefisiennya.

Dalam MYSTAT analisis faktor dapat diakses pada menu **Analyze** sub menu **Factor Analysis...**



Hal-hal yang perlu diperhatikan adalah

Model Variables adalah variabel-variabel yang akan digunakan

Display digunakan untuk menampilkan Faktor Loading dan hasil-hasil semua faktor yang mungkin

Matrix Extractions digunakan untuk memilih matriks yang akan diekstraksi dalam analisis faktor, umumnya yang digunakan adalah matriks korelasi.



Extraction parameters adalah parameter-parameter yang digunakan dalam ekstraksi faktor

Minimum eigen values adalah nilai eigen terkecil yang dapat digunakan

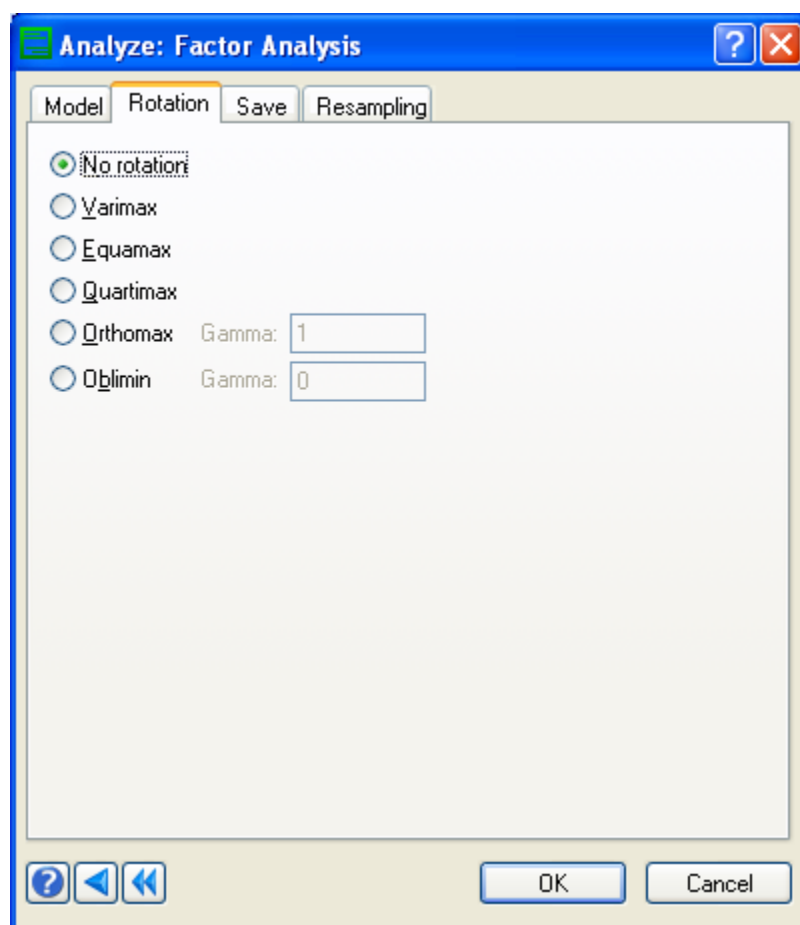
Numbers of factors adalah banyak faktor yang akan dibuat

Iterations dan **Convergence** tidak digunakan dalam metode komponen utama

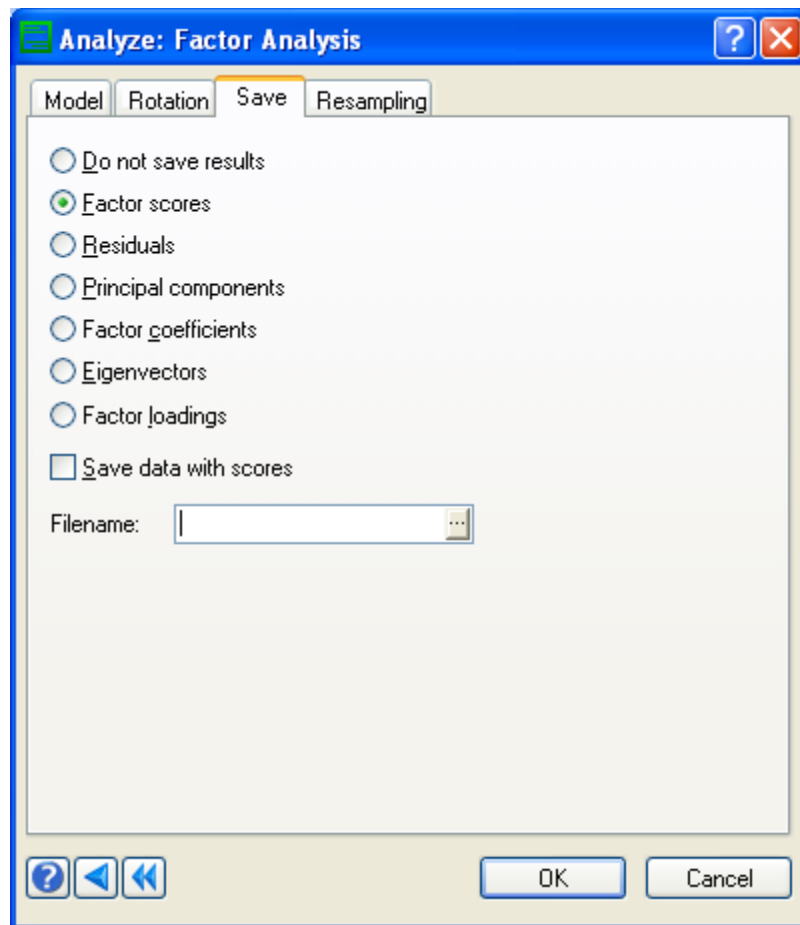
Sample Size for Matrix Input digunakan pada saat data yang diinputkan merupakan matriks korelasi atau matriks kovarians

Selain itu disediakan opsi-opsi pada tab-tab berikut

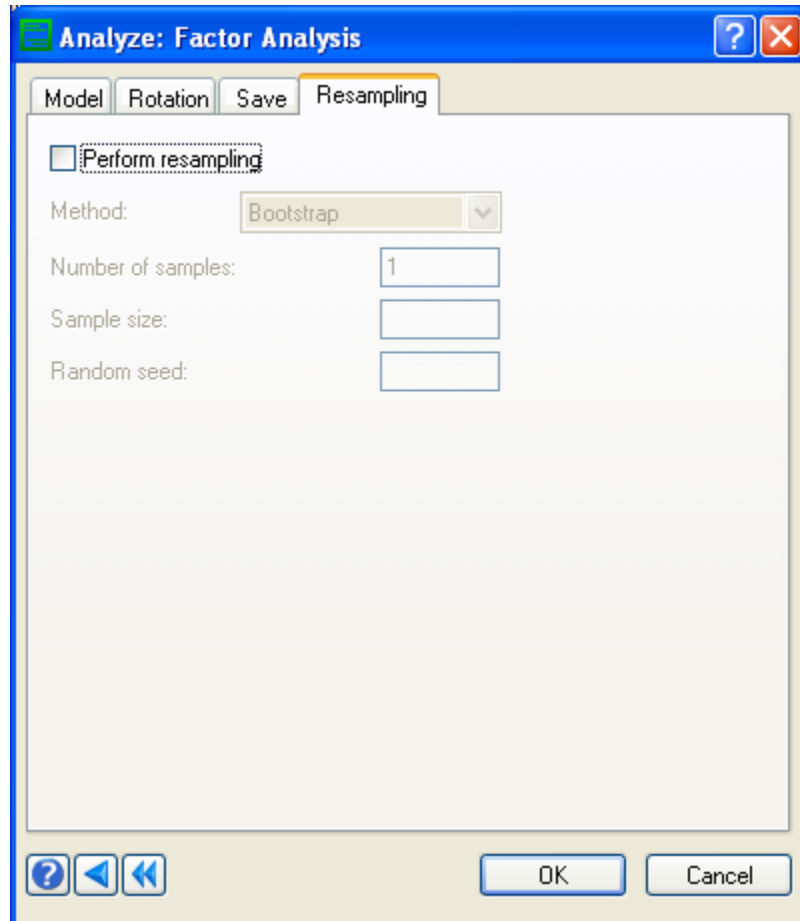
Rotation adalah metode rotasi yang digunakan, umumnya menggunakan varimax yang merupakan rotasi ortogonal.



Save adalah opsi untuk menyimpan hasil-hasil analisis faktor pada file. Hasil yang dapat disimpan diantaranya: factor scores, nilai residual, nilai komponen utama, koefisien faktor, vektor eigen dan Factor loadings.



Resampling digunakan pada saat data yang digunakan tidak mencukupi



Selain menggunakan menu Software Mystat juga menyediakan syntax atau **command** untuk analisis. Hal ini diperlukan pada saat ada beberapa metode yang tidak disediakan dalam menu. Penjelasan berikut mengenai penggunaan **command** dalam MYSTAT

FACTOR

MODEL *list_variabel*

SAVE *nama_file* / SCORES DATA LOAD COEF VECTORS PC RESID

ESTIMATE / METHOD = PCA

LISTWISE or PAIRWISE N=*n* CORR or COVA

NUMBER=*n* EIGEN=*n opsi_lainnya*

ROTATE = VARIMAX

Input data untuk analisis faktor dalam Mystat dapat berupa rawdata, matriks korelasi atau matriks kovarians. Apabila input data berupa matriks korelasi atau matriks kovarians maka nilai *n* harus diketahui besarnya.



Data yang digunakan untuk contoh penggunaan MYSTAT adalah data Youth.syz yang ada dalam software MYSTAT. Untuk analisis faktor menggunakan menu dapat diikuti petunjuk pada bagian sebelumnya. Sedangkan bila menggunakan command dapat ditulis baris-baris berikut pada command line yang ada di bagian bawah program MYSTAT

```
>FACTOR
>USE youth.syz
>MODEL HEIGHT .. WIDTH
>ESTIMATE / METHOD=PCA N=305 SORT ROTATE=VARIMAX
```

Kedua metode tersebut (melalui menu ataupun melalui command) akan menghasilkan output yang sama yaitu sebagai berikut

Session Start: Wednesday, February 10th, 2010.

▼ Factor Analysis

Latent Roots (Eigenvalues)

1	2	3	4	5	6	7	8
4.673	1.771	0.481	0.421	0.233	0.187	0.137	0.096

Component Loadings

	1	2
HEIGHT	0.859	0.372
ARM_SPAN	0.842	0.441
LOWERLEG	0.840	0.395
FOREARM	0.813	0.459
WEIGHT	0.758	-0.525
BITRO	0.674	-0.533
WIDTH	0.671	-0.418
GIRTH	0.617	-580

Variance Explained by Components

1	2
4.673	1.771



Percent of Total Variance Explained

1	2
58.411	22.137

Rotated Loading Matrix (VARIMAX, Gamma = 1.000000)

	1	2
ARM_SPAN	0.930	0.195
FOREARM	0.919	0.164
HEIGHT	0.900	0.260
LOWERLEG	0.899	0.229
WEIGHT	0.251	0.887
BITRO	0.181	0.840
GIRTH	0.107	0.840
WIDTH	0.251	0.750

"Variance" Explained by Rotated Components

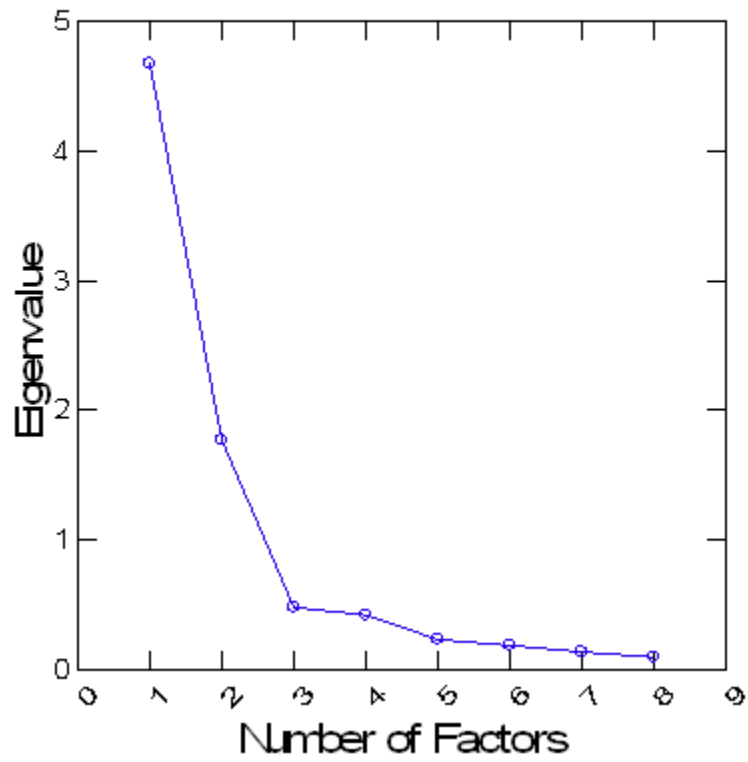
1	2
3.497	2.947

Percent of Total Variance Explained

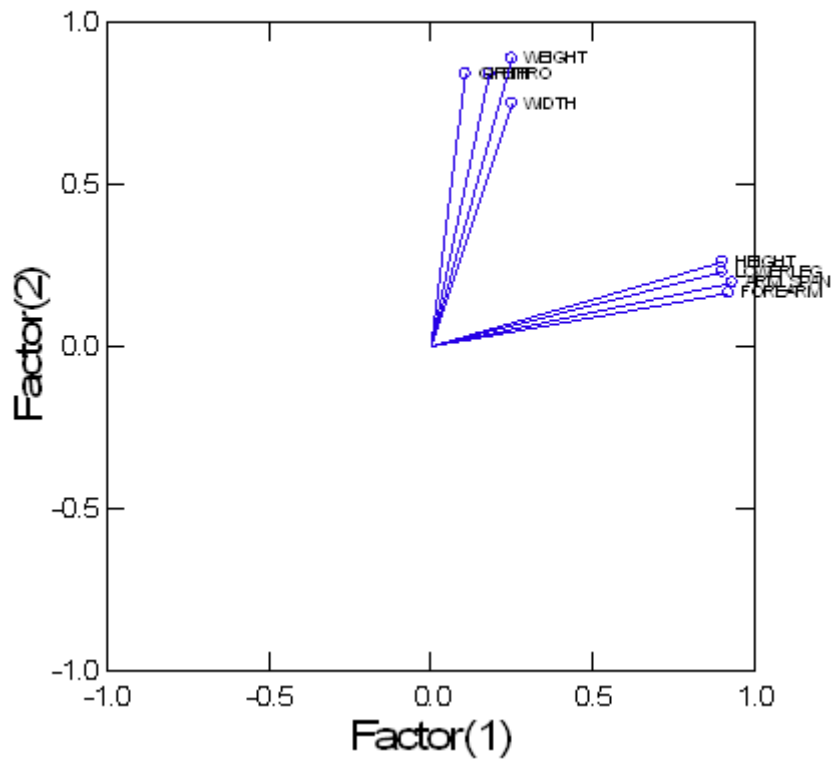
1	2
43.717	36.832



Scree Plot



Factor Loadings Plot





DAFTAR PUSTAKA

1. Johnson, Richard A, Wichern, Dean W. (2002) *Applied Multivariate Statistical Analysis*. Pearson Education International, New Jersey.
2. Manual Book: SYSTAT 11 Statistics. 2004. SYSTAT Software, Inc.