

## BAB IV

### DESAIN MODEL REGRESI LINIER GANDA

Teknik regresi ganda yang telah kita pelajari bisa sangat berguna, tapi juga bisa sangat membahayakan jika tidak tepat dalam penafsiran dan penggunaannya. Sebelum memecahkan masalah yang besar dengan metode regresi linier ganda, wajar bila kita merencanakan kegiatan sejauh mungkin, mengkhususkan suatu masalah dalam kegiatan dan serta menetapkan pilihan-pilihan sebagai tindak lanjut kegiatan yang sedang berjalan atau dilaksanakan. Perencanaan ini akan menjadi pokok yang penting dalam bab ini. Pertama-tama kita akan membicarakan 3 ciri-ciri utama dalam model matematika yang kerap kali digunakan oleh para peneliti yakni :

1. Model fungsional
2. Model kontrol
3. Model prediksi

#### 4.1. Model Fungsional

Jika hubungan fungsional yang benar antara sebuah respons dan variabel-variabel prediktor dalam sebuah masalah diketahui, maka peneliti dalam keadaan yang baik bisa mengontrol, mengerti dan meramal respons. Meskipun begitu, sangat sedikit sekali keadaan seperti itu dalam praktek, sehingga model-model seperti itu bisa ditentukan. Bahkan dalam situasi seperti itu, persamaan fungsional biasanya

sangat rumit, sulit ditafsirkan dan digunakan, dan biasanya dalam bentuk non linear. Contohnya, banyak proses kimia ditampilkan dengan sistim persamaan differensial yang mengarah ke model-model non linear. Dalam situasi seperti itu prosedur regresi linier tidak bisa diaplikasikan atau model-model linier yang lain bisa digunakan hanya sebagai pendekatan model-model yang sebenarnya dalam prosedur estimasi.

#### 4.2. Model Kontrol

Bahkan seandainya model fungsional diketahui lengkap, model fungsional tidak selalu layak untuk mengontrol sebuah variabel respons. Contohnya, dalam permasalahan banyaknya uap air yang digunakan dalam sebuah tanaman, salah satu variabel terpenting adalah temperatur, dan variabel ini tidak terkontrol, tapi variabel lain seperti proses tekanan dan proses lain variabelnya terkontrol. Sangat jelaslah bahwa model yang berisi variabel-variabel terkontrol adalah penting untuk mengontrol sebuah respons.

Sebuah model kontrol yang berguna kadang-kadang bisa dibuat dengan teknik regresi ganda, tetapi menggunakannya harus hati-hati. Jika sebuah rancangan percobaan menggunakan variabel-variabel terkontrol adalah mudah didapat, maka pengaruh variabel-variabel ini pada respon bisa didapat dari sebuah aplikasi sederhana dalam regresi ganda. Meskipun begitu, ada beberapa situasi dimana rancangan percobaan tidak

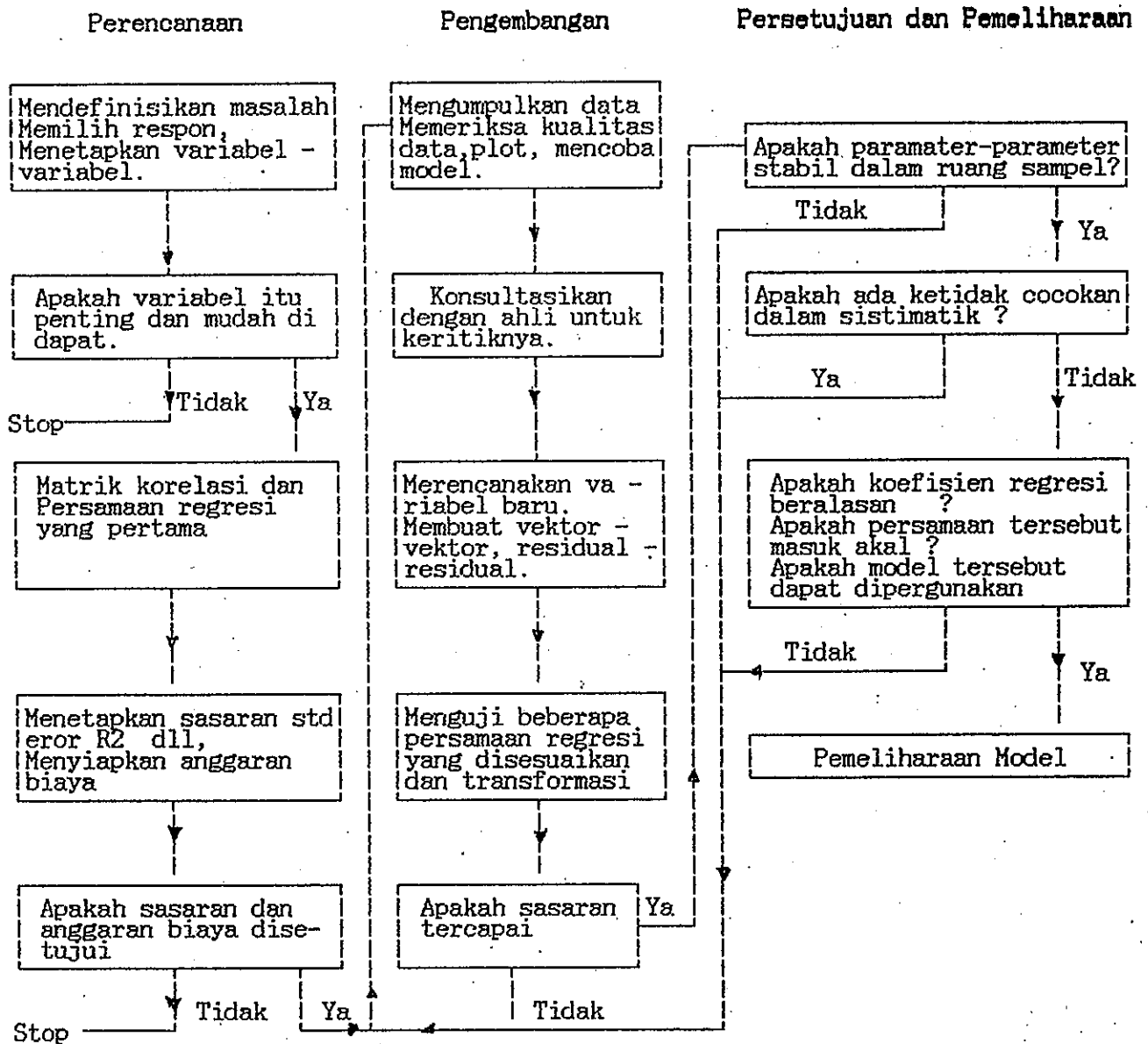
mudah dikerjakan. Contohnya, sebuah percobaan yang dilakukan dalam sebuah super market mungkin bisa dirancang dan ditangani dengan baik, tapi ada faktor-faktor yang tak terkontrol akan membuat beberapa pengaruh perhitungan matematika dalam variabel terkontrol menjadi sangat membingungkan, sehingga rancangan percobaan menjadi tak berguna. Situasi seperti ini mengarah praktis untuk menggunakan model prediksi.

#### 4.3. Model Prediksi

Jika model fungsional sangat kompleks dan kemampuan memperoleh estimasi variabel kontrol pengaruhnya terbatas, model yang kerap kali bisa digunakan adalah model prediksi linier. Gagasan ini dalam beberapa hal yang wajar mungkin tak realistik, tapi setidaknya bisa memberi ciri-ciri utama kelakuan respon yang dipelajarinya. Model prediksi ini sangat berguna dan pada situasi yang pasti bisa menuntun kepada pengertian yang nyata kedalam suatu proses atau masalah. Teknik regresi ganda mempunyai sumbangan yang besar dalam menetapkan model prediksi yang kadang-kadang berisi data-data saling berkorelasi dengan sangat tinggi yang disebut multi kolinearitas, dan dapat menunjukkan dengan tepat variabel-variabel penting yang akan dimasukkan dalam model. Tetapi menggunakannya harus hati-hati, karena mudah salah menggunakan dan mengartikan. Sebuah

perencanaan yang terorganisasi untuk memecahkan masalah yang disetujui untuk sebuah pendekatan regresi ganda adalah tepat dan perlu. Pada bab ini dimaksudkan hanya sebagai usulan atau rancangan, dan skema rancangan ini akan mempermudah dan memperbaiki keadaan suatu masalah.

Perencanaan dibawah ini meliputi pengembangan sebuah model matematika untuk tujuan prediksi. Rancangan ini secara umum cukup untuk digunakan dalam menyusun baik model fungsional maupun model kontrol. Rancangan dibagi dalam tiga tahap yaitu perencanaan pengembangan dan pemeliharaan. Sebuah skema diagram rancangan diperlihatkan dalam gb.41, dan akan dibicarakan secara lebih mendetail dibawah ini.



Gb. 4.1. Rangkaian Prosedur menyusun model

(Draper dan Smith, 1981, hal. 414)

#### 4.3.1. Proses Perencanaan Membuat Model

Mendefinisikan masalah, memilih respons dan menetapkan variabel-variabel Prediktor.

Pernyataan tertentu dalam suatu masalah adalah tahap penting dalam prosedur pemecahan masalah. Pernyataan tersebut adalah perlu karena insinyur, ilmuawan dan orang-orang bisnis memerlukan dalam pengkhususan masalahnya. Misalnya, seperti pernyataan "Mengapa konsumen membeli produk kami ? atau Mengapa pada tahap 5 operasi tidak berjalan baik hari ini ?. Hal ini adalah menarik, tapi itu tidak cukup untuk beberapa pernyataan yang terbaik. Pendefinisian masalah harus jelas dan ringkas, baik variabel respon maupun variabel prediktor, keduanya harus diidentifikasi secara jelas. Pada tahap permulaan perencanaan ini, para peneliti tidak boleh menahan diri dalam mencatat setiap variabel prediktor dan variabel respon yang mungkin yang dia anggap mempunyai andil atau pengaruh dalam masalahnya. Catatan ini akan menjadi besar, tapi menurut kesepakatan yang masuk akal akan berangsur-angsur mengurangi banyaknya variabel-variabel prediktor yang masuk dalam modelnya menjadi beberapa variabel yang layak. Petunjuk penting yang harus diingat dalam pemilihan-pemilihan variabel-variabel tidak boleh meninggalkan beberapa prosedur

statistika, termasuk prosedur regresi ganda yang akan dibicarakan pada bab VI. Akhirnya kita sampai pada kenyataan masalah yang khusus dengan sebuah respon yang tertentu atau respon-respon yang harus diselidiki dalam hubungannya dengan sekumpulan variabel-variabel prediktor yang potensial.

Apakah variabel-variabel pokok yang diusulkan itu penting untuk suatu masalah yang dibicarakan dan apakah variabel-variabel itu mudah didapat ?

Selanjutnya, daftar variabel-variabel yang diperoleh dari suatu masalah tertentu yang dibicarakan harus diuji secara hati-hati. Kebanyakan variabel-variabel prediktor ini mungkin tak terukur, untuk contohnya, penurunan temperatur dalam suatu proses dapat dianggap sebagai variabel yang penting tapi tak bisa diukur pada saat itu. Maka dibutuhkan variabel pengganti yang terukur dan masih ada hubungannya dengan penurunan temperatur tersebut, atau dibutuhkan peralatan baru yang bisa mengukur penurunan temperatur itu setiap saat atau waktu alternatif terakhir ini akan membutuhkan biaya yang besar dan para peneliti harus mempunyai keputusan yang mana dari dua alternatif itu yang terbaik. Pada tahap ini pelaksanaan penaksiran semua variabel-variabel harus dikerjakan dalam tahap perencanaan, sebelum data keseluruhan

dikumpulkan.

Pertanyaan selanjutnya yang perlu dijawab adalah " Bisakah kita mendapat himpunan yang lengkap pada saat yang sama ?, Apakah sekumpulan data tersebut komplit ?". Ada banyak situasi dimana sampel-sampel mudah dikerjakan pada saat yang sama tetapi dibutuhkan pemeriksaan pengukuran setelah itu ? Berdasarkan beban kerja, akan ada pengunduran beberapa minggu sebelum data itu diambil dari laboratorium, haruskah analisi menunggu ? haruskah kita membuang gagasan mengumpulkan pasangan data tersebut ?.

Pertanyaan macam ini harus dipertimbangkan secara hati-hati. Setelah pemeriksaan secara lengkap pada semua variabel-variabel yang telah diambil dikerjakan, penaksiran ulang adalah layak dilakukan dalam menjawab permasalahan penting.

**Apakah permasalahan mampu terpecahkan ?**

Prosedur pemilihan diatas mungkin dapat membatasi banyaknya variabel prediktor dari pertimbangan-pertimbangan, sekaligus mengurangi kesempatan dalam memecahkan masalah pada umumnya. Umumnya hal ini tidak akan terjadi. Namun demikian, salah satu dari tiga keputusan kemungkinan harus diambil pada tahap perencanaan ini.



1. Seluruh kegiatan harus dihentikan
2. Seluruh kegiatan harus didefinisikan ulang dengan menambah pengetahuan pada permasalahan ini.
3. Seluruh kegiatan mungkin bisa dilaksanakan maka perencanaan selanjutnya bisa dilanjutkan.

#### **Matrik korelasi dan membuat persamaan regresi ?**

Jika kegiatan tersebut bisa dilanjutkan, perencana kegiatan selanjutnya harus mengambil beberapa ciri-ciri dari data untuk dianalisa. Sehingga perlu membuat jadwal waktu untuk penyelesaian kegiatan tersebut. Yang berisi misalnya karyawan, biaya dan lain-lain. Jika mungkin sekumpulan data sampel harus diperoleh, ringkasan statistik, matrik korelasi dan matrik korelasi invers harus dihitung.

Selanjutnya menghitung korelasi antara variabel-variabel dengan respon. Untuk setiap respon, harus ada satu atau lebih variabel prediktor yang mempunyai korelasi yang tinggi. Jika tidak, pertimbangkan keadaan tersebut, apakah variabel-variabel prediktor yang penting tidak dimasukan ?

Mungkin pengaruh variabel prediktor terhadap respon sangat kecil. Yang harus diingat bahwa korelasi antara variabel-variabel prediktor terhadap respon

harus besar, memang mencapai peramalan yang baik adalah sulit, memperpendek dan mencoba pada tahap awal perencanaan ini memungkinkan peneliti membuat estimasi yang bijaksana dalam waktu dan biaya yang dibutuhkan dan juga untuk menaksir kesempatan menyusun model prediksi yang baik.

### Menetapkan Sasaran dan Menyiapkan Anggaran Biaya

Pada tahap perencanaan ini, para peneliti harus menetapkan sasaran untuk kegiatan, memperkirakan jadwal waktu untuk menjawab permasalahan, menyiapkan karyawan dan keperluan komputer. Pilihan waktu dan karyawan harus disetujui dan kegiatan secara keseluruhan diajukan untuk disetujui pula. Dibawah ini diperlihatkan contoh usulan yang akan diajukan untuk disetujui.

#### Bentuk Kegiatan

JUDUL KEGIATAN : Persamaan estimasi untuk banyaknya konsumsi air pada tanaman A.

PELAKSANA KEGIATAN : UNTUNG HARTONO

JENIS ANALISA YANG DIGUNAKAN : Regresi linier ganda

SASARAN KEGIATAN :

1. Persamaan akhir mempunyai variasi 80% ( $R^2 > 0,8$ )

2. Standar kesalahan estimasi harus tidak lebih dari 5 % dalam rata-rata banyaknya air yang digunakan.
3. Banyaknya variabel prediktor adalah ..... (5)
4. Semua koefisien estimasi dalam model akhir adalah dengan tingkat signifikan statistik : 0,05.
5. Tidak ada pola-pola yang menyimpang dalam residual.

Biaya	Total rupiah
Karyawan	400.000
2 orang /bln	
computer 6 jam	60.000
	<hr/>
Total	460.000

#### JADWAL WAKTU

- |  |          |
|--|----------|
| 1. Pengumpulan data, pembuatan modal, dan memeriksa data | 2 minggu |
| 2. Modifikasi dan pengembangan model                     | 4 minggu |
| 3. Ringkasan dan melengkapi kekurangan                   | 2 minggu |

**Apakah Sasaran dan Anggaran Biaya dapat Diterima ?**

Jika usulan proyek disetujui, kegiatan selanjutnya adalah pada tahap pengembangan, jika tidak kegiatan kegiatan direvisi menjadi sederhana atau distop.

#### 4.3.2. Pengembangan Model Matematika

**Mengumpulkan data, Memeriksa kualitas data, Plot dan Mencoba Model.**

Memeriksa proses pengumpulan data harus hati-hati, jangan mencoba menyusun sebuah model dengan data yang jelek. Dalam penelitian sosial kerap kali dijumpai data-data semacam itu, misalnya seorang laki-laki tingginya 14 inch, wanita dengan berat 1000 pound, pelajar dengan " tidak mempunyai paru-paru" dan lain-lain. Didalam bagian pengolahan data suatu industri pertambangan, kerap kali dijumpai data-data yang jelek pula, misalnya ada 10.000 pound material didalam sebuah kapasitas 100 pound barel, dan banyak contoh kesalahan yang hampir mirip lainnya. Semua penelitian dan perencanaan didunia tidak akan bisa menghilangkan masalah-masalah jenis ini.

Hal ini adalah kesalahan manusia, bukan kesalahan komputer, meskipun komputer kerap kali melakukan kesalahan.

Salah satu dari semua kegiatan ini yang harus dikerjakan adalah proses penyusunan model dapat dimulai, menggambarkan data, kecocokan model, pemeriksaan residual dan lain-lain.

**Rundingkan dengan Ahli untuk sarannya.**

Sebagai kemajuan pekerjaannya, persamaan regresi potensial akan diajukan ke ahli dalam bidangnya, untuk persetujuan dan saran mereka.

**Menetapkan variabel-variabel Baru dan Menyusun Vektor-vektor yang menerangkan Residual-residual.**

Jika para ahli mengusulkan variabel-variabel baru untuk kemungkinan bisa digunakan, data ini harus diperoleh dan diuji dalam kaitannya dengan harga variabel respons dan residual-residual dari persamaan regresi yang cocok.

**Menguji Beberapa Persamaan Regresi Yang Disesuaikan**

Variabel prediktor yang terpilih sekarang bisa dicoba dalam persamaan regresi, variabel transformasi mungkin juga bisa digunakan. Sebagai misalnya sebuah plot residual tertentu mungkin menunjukkan bahwa bentuk logaritma variabel prediktor yang akan digunakan lebih baik daripada variabel prediktor itu sendiri. Banyak transformasi pada awalnya ditetapkan oleh kebiasaan ilmuwan dengan proses yang diamatinya. Pendekatan transformasi pada variabel asli yang dimasukan

dalam persamaan regresi jarang berlangsung dengan baik, bahkan pemakaian variabel transformasi kemungkinan sia-sia saja.

#### **Apakah Sasaran Tercapai ?**

Akhirnya, kadang-kadang setelah banyak percobaan-percobaan, peneliti telah sampai pada apa yang dia anggap persamaan terbaiknya dan akan menguji persamaan tersebut dalam sasaran-sasaran yang jelas yang ditetapkan dalam tahap perencanaan ini. Jika tidak memenuhi standar ini, keputusan harus dibuat seperti misalnya :

Apakah proses tersebut harus dihentikan ? atau apakah siklus dalam pengembangan tersebut harus diulangi ? Mungkin banyak biaya yang akan dibutuhkan, atau mungkin perubahan sasaran adalah tepat. Hal ini adalah pilihan lain dalam suatu proses.

#### **4.3.3. Persetujuan dan Pemeliharaan Model Matematika**

Setelah sebuah persamaan atau model menemui sekumpulan sasaran dalam tahap perencanaan dan diterima sebagai model prediksi yang berguna, merupakan suatu keharusan untuk menetapkan prosedur untuk pemeliharaan dan validasi model.

### Apakah Parameter Stabil dalam Ruang Sampel ?

Kita bisa membedakan dua macam perbedaan sekumpulan data yaitu data longitudinal dan data cross seksional dan dapat digunakan ketika menguji kestabilan parameter.

**DATA LONGITUDINAL.** Jika model yang telah dibuat menggunakan pengamatan yang diambil melalui sebuah rentang waktu yang panjang, maka seseorang dapat menguji kestabilan koefisien estimasi dengan mencocokkan model pada waktu rentang yang pendek dan menentukan berturut-turut pola estimasi sekumpulan parameter-parameter. Sebagai contohnya, jika data diambil pada waktu bulanan untuk empat tahun berturutan, maka model yang terpilih bisa dicocokkan untuk data masing-masing dalam empat tahun terpisah, sehingga mendapat empat kumpulan koefisien regresi. Jika estimasi koefisien-koefisien tidak stabil, yaitu parameter parameter menunjukkan pola yang berubah-ubah, maka menggunakan model yang cocok yang didapat dari semua data untuk tujuan prediksi menjadi tidak bijaksana.

**DATA CROSS SEKSIONAL.** Jika data bisa dianggap sebagai kumpulan informasi pada "titik" yang sama didalam ukuran waktu, seperti misalnya didalam

kelompok kerja yang sama atau didalam sekumpulan barang-barang mentah yang sama, maka metode untuk menguji kestabilan parameter-parameter adalah pertama-tama membagi sekumpulan data yang lengkap kedalam subset dengan kriteria yang masuk akal dan kemudian menggunakan salah satu subset dari subset dari data untuk membuat model prediksi dan bagian yang lain untuk menguji kebenaran model.

#### **Apakah Ada Ketidakcocokan Sistimatis ?**

Bahkan jika sebuah model yang cocok adalah diterima dengan parameter-parameter yang sangat stabil. Kemungkinan variabel-variabel prediktor telah dihilangkan. Residual akan selalu diuji dalam semua cara yang mungkin untuk dilihat, jika ada beberapa pola yang dapat dilihat menunjukkan variabel tersebut dihilangkan.

#### **Apakah Koefisien Regresi Pantas ?**

Ini mungkin kelihatan sebagai sebuah pertanyaan yang aneh, tapi harus diingat bahwa model tersebut akan digunakan oleh beberapa orang yang tidak menyadari akan kenyataan bahwa koefisien regresi kuadrat terkecil adalah menyesuaikan untuk variabel lain dalam persamaan regresi. Jadi mereka mungkin mencoba untuk meramal



respons dengan hanya perubahan dalam satu variabel, dan menggunakan koefisien itu untuk memutuskan banyaknya pengaruh perubahan respons. Jika variabel-variabel prediktornya variabel orthogonal, maka ini mungkin mengakibatkan kesalahan yang kecil. Tetapi, jika variabel-variabel prediktor merupakan korelasi yang tinggi, maka kepercayaan terhadap koefisien-koefisien secara individual sangat tidak bijaksana.

Apakah persamaan tersebut Masuk Akal ?

Apakah persamaan tersebut dapat Digunakan ?

Model akhir mungkin berisi sekumpulan variabel-variabel yang bisa dipergunakan untuk meramal respon, tetapi mungkin tidak bisa untuk menilai pengaruh setiap variabel prediktor secara individual.

#### **Pemeliharaan Model.**

Jika semua kriteria sebelumnya dan pilihan-pilihan telah memuaskan, maka perlu membuat sebuah prosedur untuk pemeliharaan model. Kondisi fisik menuntut perubahan dan perlu juga membuat suatu keadaan yang menentukan apabila terdapat penyimpangan-penyimpangan pengamatan yang nyata dari harga prediksi menunjukkan bahwa model mulai

menunjukkan tanda-tanda menjadi usang. Maka perlu analisis membuat sebuah diagram control untuk penyimpangan-penyimpangan tersebut. Prosedur pengendalian mutu harus mampu memberikan kecocokan pada sebuah model secara periodik dengan analisa statistik.