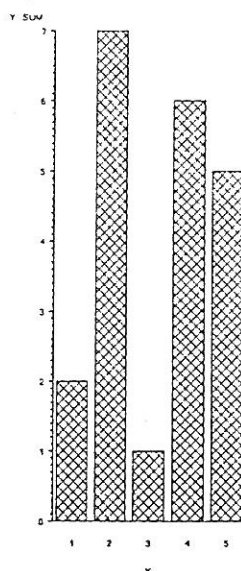


SAS-kursus

MILJØMINISTERIET

DANMARKS MILJØUNDERSØGELSER
Afd. for Flora- & Faunaøkologi



Forelæsningsnoter
af
Mogens Ring Petersen

Formål

Disse noter viser, hvordan man kommer i gang med SAS, hvordan man indlæser data, og hvordan man får noget ud igen. Idé til udformning af programmer og fremgangsmåde er delvis hentet fra *EDB for biologer, Forår 1991 - bind 2* af Jørgen Gomme, Henrik Nielsen, Peter S. Jørgensen og Mogens Ring Petersen.

Jeg har bestræbt mig på at vise nogle eksempler, der (for de første 6 programmers vedkommende) ikke vil skræmme nybegynderen væk, men hvor der også kan være små fif til den, der er kommet længere. Forklaringen er med vilje gjort så kort som mulig.

For en mere udtømmende dokumentation henvises til SAS-manualer og diverse SAS-bøger.

Side 3-14 er de grundlæggende. Derefter har jeg vist nogle eksempler på SAS-anvendelse og har puttet en masse detaljer ind i programmet (SASKURS7.SAS). Her kan man finde simple grafiske afbildninger og simpel statistik. Listingfilen til programmet følger lige bagefter. Senere er der eksempler på SASGRAPH og endelig til sidst har jeg vist, hvordan man tester sine data statistisk med nogle af de almindeligste tests.

Indholdsfortegnelse

	Side
1) Start af SAS og Display Manager	3
2) Indlæsning af tekstdata	6
3) Sammensætning af datasæt	13
4) Eksempel på SAS-anvendelse	15
5) SASGRAPH	21
6) Statistisk anvendelse af SAS	27

Man starter SAS ved at skrive SAS
, og derpå kommer følgende tre vinduer frem:

```

OUTPUT
Command ==>

LOG
Command ==>

NOTE: Copyright(c) 1985,86,87 SAS Institute Inc., Cary, NC 27512-8000, U.S.A.
NOTE: SAS (r) Proprietary Software Release 6.04
      Licensed to MILJOEMINISTERIET, Site 21321001.

NOTE: AUTOEXEC processing completed.

PROGRAM EDITOR
Command ==>

00001
00002
00003
00004
    
```

Man kan så skrive sit program i programvinduet. Når man er færdig, sendes det afsted vha af ordren **SUBMIT**.

```

OUTPUT
Command ==>

LOG
Command ==>

NOTE: Copyright(c) 1985,86,87 SAS Institute Inc., Cary, NC 27512-8000, U.S.A.
NOTE: SAS (r) Proprietary Software Release 6.04
      Licensed to MILJOEMINISTERIET, Site 21321001.

NOTE: AUTOEXEC processing completed.

PROGRAM EDITOR
Command ==> submit
00001
00002 *-----;
00003 *  SASKURS2.SAS                                KALØ 18.1.1993 ;
00004 *-----;
    
```

Herved kommer der resultater i OUTPUT-vinduet og meddelelser om programforløbet i LOG-vinduet:

```

OUTPUT
Command ==>

  7  Thyra      70446    2666      166      .      145      110
  8  Tommy     260453    1803      187     780      .      .
  9  Viggo      71046     231      182     802     115      60
 10  Freddy    220952     891      170     615     130      70
 11  Yrsa     150248     446      179     685     120      80

LOG
Command ==>

NOTE: The data set PERMA.HELBRED has 11 observations and
      7 variables.
NOTE: The DATA statement used 2.00 seconds.
      41  run;
NOTE: The PROCEDURE PRINT used 1.00 seconds.

PROGRAM EDITOR
Command ==>

00001
00002
00003
00004
    
```

```
PROGRAM EDITOR
Command ==> file 'mitprog.sas'
```

```
00001
00002  Sådan gemmer man et program under navnet mitprog.sas.
00003  Næste gang behøver man kun at skrive 'file'.
00004
```

ZOOM

```
PROGRAM EDITOR
Command ==> submit (eller <F10> med "standardopsætning")
```

```
00001
00002  Sådan udfører man programmet.
00003
```

ZOOM

```
PROGRAM EDITOR
Command ==> clear
```

```
00001
00002  Sådan clearer man et vindue.
00003
```

ZOOM

```
PROGRAM EDITOR
Command ==> recall (eller <F9> med "standardopsætning")
```

```
00001
00002  Sådan kalder man programmet tilbage.
00003
```

ZOOM

```
PROGRAM EDITOR
Command ==> inc 'oldprog.sas'
```

```
00001
00002  Sådan henter man et gammelt program.
00003
```

ZOOM

Her er nogle nyttige liniekommandoer, når man skal redigere i sit program:

```
PROGRAM EDITOR
Command ==>
```

```
00001  <-- Her i talkolonnerne kan bl.a. skrives følgende ordrer:
00002
00003  I      (indsætter 1 linje)
00004  I9    (      "      9 linjer)
00005
00006  R      (gentager linjen)
00007  R8    (      "      "      8 gange)
00008
00009  RR3    (gentager blokken, i det her tilfælde linje 9-10 3 gange)
00010  RR
00011
00012  C      (kopierer linjen)
00013
00014  CC     (kopierer blokken, i det her tilfælde linje 14-16)
00015
00016  CC
00017
00018  M      (flytter linjen)
00019
00020  MM     (flytter blokken)
00021  MM
00022
00023  B      ("Before", bruges ved C, M og I)
00024
00025  A      ("After",      "      )
00026
00027  >      (indrykker linjen 1 position)
00028
00029  >>     (indrykker blokken 1 position)
00030  >>
00031
00032  >6     (indrykker linjen 6 positioner)
00033
00034  >>6    (indrykker blokken 6 positioner)
00035  >>
00036
```

ZOOM

Det er praktisk at lægge de ordrer, man bruger tit (som fx SUBMIT) ind i funktionstasterne <F1>-<F12> m.fl..

Her er en fortegnelse over, hvordan du kan sætte dine "keys" op. Du skriver KEYS i programvinduet (ud for pilen) og går ud i vinduet og skriver den ordre, du vil have tasten til at udføre.

Du gemmer det hele ved at skrive END udfor Command ==>. Derved bliver filen F:\FF\ (dit user-id)\SASUSER\PROFILE.SCT (, som altså ligger på dit eget drev), opdateret med de nye ændringer.

KEYS <DMKEYS>	
Command ==>	
Key	Description
F1	help
F2	keys
F3	log;zoom on
F4	output;zoom on
F5	next
F6	pgm;zoom on
F7	zoom
F8	subtop
F9	pgm;recall;zoom on
F10	zoom off;submit

Et andet vindue er Options-vinduet, som du ligeledes kan rette i og gemme med <F10>. Jeg har fx sat PAGESIZE op til 500, så jeg slipper for utidige sideskift.

OUTPUT		OPTIONS	
Command ==> Options		Command ==>	
		Option	Value
		CENTER	ON
		CHARCODE	OFF
		DATE	OFF
		DEVICE	VGA
		ERRORS	20
		FIRSTOBS	1
		FORMCHAR	---- + ----+
		KANJI	OFF
		LINESIZE	64
		MISSING	.
		MPRINT	ON
		MTRACE	ON
		NEWS	
		NOTES	ON
		NUMBER	ON
		OBS	MAX
		OVP	ON
		PAGENO	1
		PAGESIZE	500
		PROBSIG	0
		PROFILE	OFF
		REMOTE	
		REPLACE	ON
		SOURCE	ON
		SOURCE2	OFF
		STIMER	ON
		SYMBOLGEN	OFF
		LAST	_NULL_

Dette er en liste over nogle personer, som er blevet undersøgt helbredsmæssigt.

Kolonnerne er som følger:

- 1) navn
- 2) fødselsdag
- 3) sidste fire cifre i personnummer
- 4) højde i cm
- 5) vægt i cg
- 6) systole (den højeste værdi, man får ved blodtryksmåling)
- 7) diastole (den laveste værdi, man får ved blodtryksmåling)

Her er filen SASKURS2.TXT :

Erna	080551	2602	175	670	.	.
Else	020954	4998	184	695	135	90
Hans	210552	2713	174	765	120	80
Jesper	270345	1113	180	762	.	.
Børge	060149	0621	176	750	135	90
Thyra	040147	0258	169	680	105	65
Thyra	070446	2666	166	.	145	110
Tommy	260453	1803	187	780	.	.
Viggo	071046	0231	182	802	115	60
Freddy	220952	0891	170	615	130	70
Yrsa	150248	0446	179	685	120	80

På den næste side ser du et program (SASKURS1.SAS), der læser denne tekst ind vha CARDS og i sidste program (SASKURS2.SAS) læses fra ekstern fil og der laves samtidigt et permanent SAS-datasæt ud af teksten.

Der er ikke noget i vejen for, at man også kan lave et permanent datasæt vha CARDS-indlæsningen og man kan også her læse i frit format. De to programmer er blot lavet med så mange variationer, så du ser nogle flere eksempler på, hvad man kan gøre.

```

*-----;
*   SASKURS1.SAS                                KALØ 18.1.1993   ;
*-----;
*   Indlæser nedenstående tekstdata vha Cards og kolonnevis ind- ;
*   læsning (Column Input).                                     ;
*-----;

Data temp1; /* Etledet navn giver et temporært datasæt */

Input navn $ 1-10 foeddag 12-17 persnr 20-23 hoejde 26-28
       vaegt 30-32 systole 34-36 diastole 38-40;

/* Man skal skrive $, hvis teksten skal ind i SASdata-sættet som
tekst, ellers vil SAS (default) prøve at omsætte variabel-inputtet
til binær talkode */

/* Tip: ved at skrive COLS ude i kommandofeltet (tallene til venstre
i Display Manager) får du en lineal, så du kan se i hvilken kolonne
de enkelte tal befinder sig */

/* i dette program vil vi skrive tekstdata direkte ind i programmet */
/* her kunne man også have brugt ordet 'Datalines' */
Cards;
Else          020954  4998  184 695 135  90
Erna          080551  2602  175 670
Hans          210552  2713  174 765 120  80
Jesper       270345  1113  180 762
Børge        060149  0621  176 750 135  90
Thyra        040147  0258  169 680 105  65
Thyra        070446  2666  166   . 145 110
Tommy        260453  1803  187 780
Viggo        071046  0231  182 802 115  60
Freddy       220952  0891  170 615 130  70
Yrsa         150248  0446  179 685 120  80
;

Proc print; /* her får man bare de samme data skrevet ud igen */

run; /* for at få det sidste trin med - ellers stopper den ved
      Proc print */

```

```

*-----;
*   LISTINGFIL:                                     ;
*-----;
OBS   NAVN   FOEDDAG  PERSNR  HOEJDE  VAEGT  SYSTOLE  DIASTOLE ;
1    Erna    80551    2602    175     670    .         .
2    Else    20954    4998    184     695    135        90
3    Hans    210552   2713    174     765    120        80
4    Jesper  270345   1113    180     762    .         .
5    Børge   60149     621    176     750    135        90
6    Thyra   40147     258    169     680    105        65
7    Thyra   70446    2666    166     .     145       110
8    Tommy   260453   1803    187     780    .         .
9    Viggo   71046     231    182     802    115        60
10   Freddy  220952   0891    170     615    130        70
11   Yrsa    150248   0446    179     685    120        80
*/

```



```

*-----;
*   SASKURS2.SAS                                KALØ 18.1.1993 ;
*-----;
*   Indlæser nedenstående textdata vha ekstern fil og bruger ind- ;
*   læsning i frit format (List Input). ;
*-----;
*   Betingelser: ;
*   Manglende værdier erstattes med punktum. ;
*   Tekst til hver variabel må ikke være over 8 tegn bred - ellers ;
*   skal man bruge length (se et senere eksempel). ;
*-----;

Libname perma 'F:\ff\ffmrp'; /* her skal datasættet sendes ud -
                               filen kommer til at hedde
                               'helbred.ssd' */

Data temp2; /* etledet navn giver temporært datasæt */
*-----;
*   Normalt kan man allerede her angive et toledet navn og der- ;
*   ved få et permanent datasæt, men her på Kalø er der proble- ;
*   mer med at køre det på den måde. Man kan så løse problemet ;
*   ved først at lave et temporært datasæt, og så senere lave ;
*   et permanent ud af dette (jfr. 'set temp2' nedenfor). ;
*-----;

Infile 'F:\ff\ffmrp\saskurs2.txt';
*-----;
*   Herfra læses teksten. ;
*   Hvis man kun vil have læst fra linje 17 til og med 27 skrives ;
*   i stedet: ;
*   Infile 'F:\ff\ffmrp\saskurs2.txt' firstobs=17 obs=27;
*-----;

Length navn $ 10;
Input navn $ foeddag persnr hoejde vaegt systole diastole;

Data perma.helbred2; /* toledet navn giver permanent datasæt */
    set temp2;

Proc print;

run;

```

```

*-----;
*   LISTINGFIL: ;
*-----;
OBS   NAVN   FOEDDAG  PERSNR  HOEJDE  VAEGT  SYSTOLE  DIASTOLE ;
1     Erna   80551    2602    175     670     .         .
2     Else   20954    4998    184     695    135       90
3     Hans   210552   2713    174     765    120       80
4     Jesper 270345   1113    180     762     .         .
5     Børge  60149    621    176     750    135       90
6     Thyra  40147    258    169     680    105       65
7     Thyra  70446   2666    166     .     145      110
8     Tommy  260453   1803    187     780     .         .
9     Viggo  71046    231    182     802    115       60
10    Freddy 220952   891    170     615    130       70
11    Yrsa   150248   446    179     685    120       80
*/

```

Her er filen SASKURS3.TXT :

Dette er en ASCII-fil. I denne tekstfil består fx tallet 155 af tre ASCII-tegn.

Dette er en liste over flere personer, som er blevet undersøgt helbredsmæssigt.

Kolonnerne er som følger:

- 1) navn
- 2) fødselsdag
- 3) sidste fire cifre i personnummer
- 4) højde i cm
- 5) vægt i cg
- 6) systole (den højeste værdi, man får ved blodtryksmåling)
- 7) diastole (den laveste værdi, man får ved blodtryksmåling)

```
-----  
Børge      150248  0449  172 800 155 110  
Olfert     030748  0909  164 524 170 120  
Hugo       010101  0000  
Carl-Ejnar 080645  1117  191 905 150 100  
-----
```

På næste side ser du et program, der læser denne tekst ind og laver et permanent SAS-datasæt ud af teksten.

```

*-----;
*  SASKURS3.SAS                                KALØ 18.1.1993 ;
*-----;
*  Indlæser nedenstående textdata vha ekstern fil og bruger ind- ;
*  læsning i frit format (List Input). ;
*-----;

Libname perma 'F:\ff\ffmrp';

Data temp3;
  Infile 'F:\ff\ffmrp\saskurs3.txt' firstobs=18 obs=21 missover;
  /* MISSEVER bevirker, at den ikke læser videre på en ny linje,
     hvis de sidste variable mangler i en observation */

  Length navn $ 10;
  Input navn foeddag persnr hoejde vaegt systole diastole;

Data perma.helbred3; /* toledet navn giver permanent datasæt */
  set temp3;

Proc print;

run;

```

```

*-----;
*  LISTINGFIL: ;
*-----; /*
  OBS NAVN          FOEDDAG PERSNR HOEJDE VAEGT SYSTOLE DIASTOLE

    1  Børge          150248   449    172   800    155    110
    2  Olfert          30748   909    164   524    170    120
    3  Hugo            10101     0     .     .     .     .
    4  Carl-Ejnar      80645  1117    191   905    150    100
*/

```

Her er filen SASKURS4.TXT :

Dette er en ASCII-fil. F.eks.
består '1969' af ASCII-
tegnene for 1, 9, 6 og 9.

Dette er en liste over personer i undersøgelsesmaterialet,
som er jægere. Tallet efter personnummeret angivet det år,
den enkelte fik sit jagttegn.

```
-----  
Børge      060149-0621  1969  
Hans       210552-2713  1974  
Thyra      040147-0258  1969  
Erna       080551-2602  1975  
Thyra      070446-2666  1966  
Børge      150248-0449  1972  
Tommy      260453-1803  1973  
Carl-Ejnar 080645-1117  
Jesper     270345-1113  1980  
Else       020954-4998  1984  
-----
```

På næste side ser du et program, der læser denne tekst ind og
laver et permanent SAS-datasæt ud af teksten.

```
*-----;
* SASKURS4.SAS                                KALØ 18.1.1993 ;
*-----;
* Indlæser nedenstående textdata vha ekstern fil og bruger ind- ;
* læsning i forskellige formater. ;
*-----;
```

Libname perma 'F:\ff\ffmrp';

Data temp;

Infile 'F:\ff\ffmrp\saskurs4.txt'

firstobs=11 obs=20 missover lrecl=200;

/* MISSEVER bevirker, at den ikke læser videre på en ny linje,
hvis de sidste variable mangler i en observation */

/* LRECL=200 bevirker, at SAS kan læse i tekstfilen ud til
kolonne 200. Uden denne option vil SAS kun kunne læse ud
til (default-værdien) kolonne 132. I dette tilfælde er
denne option ikke nødvendig */

Length navn \$ 10;

Input navn foeddag 12-17 persnr 19-22 jaeger;

Data perma.helbred4; /* toledet navn giver permanent datasæt */
set temp;

Proc print;

run;

```
*-----;
* LISTINGFIL: ;
*-----; /*
```

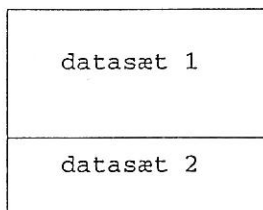
OBS	NAVN	FOEDDAG	PERSNR	JAEGER
1	Børge	60149	621	1969
2	Hans	210552	2713	1974
3	Thyra	40147	258	1969
4	Erna	80551	2602	1975
5	Thyra	70446	2666	1966
6	Børge	150248	449	1972
7	Tommy	260453	1803	1973
8	Carl-Ejnar	80645	1117	.
9	Jesper	270345	1113	1980
10	Else	20954	4998	1984

```
*/
```

```

*-----;
*   SASKURS5.SAS                                KALØ 18.1.1993 ;
*-----;
*   Konkatenerer to datasæt, dvs. sætter dem sammen på denne måde: ;
*   ;
*   ;
*   ;
*   ;
*   ;
*   ;
*   ;
*   ;
*   ;
*   ;
*   ;
*-----;

```



```

Libname perma 'F:\ff\ffmrp';

```

```

Data temp1;
  set perma.helbred2;

```

```

Data temp2;
  set perma.helbred3;

```

```

Data temp3;
  set temp1 temp2;

```

```

Data perma.konkat;
  set temp3;

```

```

Proc print;

```

```

run;

```

```

*-----;
*   LISTINGFIL:                                ;
*-----; /*

```

OBS	NAVN	FOEDDAG	PERSNR	HOEJDE	VAEGT	SYSTOLE	DIASTOLE
1	Erna	80551	2602	175	670	.	.
2	Else	20954	4998	184	695	135	90
3	Hans	210552	2713	174	765	120	80
4	Jesper	270345	1113	180	762	.	.
5	Børge	60149	621	176	750	135	90
6	Thyra	40147	258	169	680	105	65
7	Thyra	70446	2666	166	.	145	110
8	Tommy	260453	1803	187	780	.	.
9	Viggo	71046	231	182	802	115	60
10	Freddy	220952	891	170	615	130	70
11	Yrsa	150248	446	179	685	120	80
12	Børge	150248	449	172	800	155	110
13	Olfert	30748	909	164	524	170	120
14	Hugo	10101	0	.	.	.	.
15	Carl-Ejnar	80645	1117	191	905	150	100

```

*/

```

datasæt 1	d a t a 2
-----------	-----------------------

```
Data temp1;
    set perma.konkat;

Data temp2;
    set perma.helbred4;

Proc sort data=temp1;
    by navn foeddag persnr;

Proc sort data=temp2;
    by navn foeddag persnr;

Data temp3;
    merge temp1 temp2;
    by navn foeddag persnr;

Data perma.matmerg;
    set temp3;

Proc print;

run;
```

```
*-----;
* LISTINGFIL:;
*-----:/*
```

		F	P	H	S	D		
		O			Y	I		
		E	E	O	S	A	J	
	N	D	R	V	T	S	A	
O	A	D	S	E	O	O	E	
B	V	A	N	D	G	L	G	
S	N	G	R	E	T	E	R	
1	Børge	60149	621	176	750	135	90	1969
2	Børge	150248	449	172	800	155	110	1972
3	Carl-Ejnar	80645	1117	191	905	150	100	.
4	Else	20954	4998	184	695	135	90	1984
5	Erna	80551	2602	175	670	.	.	1975
6	Freddy	220952	891	170	615	130	70	.
7	Hans	210552	2713	174	765	120	80	1974
8	Hugo	10101	0	.	.	.	.	.
9	Jesper	270345	1113	180	762	.	.	1980
10	Olfert	30748	909	164	524	170	120	.
11	Thyra	40147	258	169	680	105	65	1969
12	Thyra	70446	2666	166	.	145	110	1966
13	Tommy	260453	1803	187	780	.	.	1973
14	Viggo	71046	231	182	802	115	60	.
15	Yrsa	150248	446	179	685	120	80	.

* /

```
proc chart;
  hbar alder;
  by koen;
  title2 '---> proc chart <--';
  title2 '---> hbar alder <--';
  title3 '---> by koen <--';

proc plot;
  plot hoejde*vaegt;
  title2 '---> proc plot <--';
  title3 '---> plot hoejde*vaegt <--';

proc plot;
  plot hoejde*vaegt=koen;
  title2 '---> proc plot <--';
  title3 '---> plot hoejde*vaegt=koen <--';

run;
```

```
*-----;
* LISTINGFIL:                               ;
*-----; /*
```

----- KOEN=K -----

NAVN	FØDT	ALDER	HØJDE	VÆGT I KG
Else	020954	38	184	69.5
Erna	080551	41	175	67.0
Thyra	040147	45	169	68.0
Thyra	070446	46	166	.
Yrsa	150248	44	179	68.5

----- KOEN=M -----

NAVN	FØDT	ALDER	HØJDE	VÆGT I KG
Børge	060149	43	176	75.0
Børge	150248	44	172	80.0
Carl-Ejnar	080645	47	191	90.5
Freddy	220952	40	170	61.5
Hans	210552	40	174	76.5
Jesper	270345	47	180	76.2
Olfert	030748	44	164	52.4
Tommy	260453	39	187	78.0
Viggo	071046	46	182	80.2

Kommentar:

Bemærk, at Hugo ikke er kommet med pga programtrinnet
IF ALDER > 50 THEN DELETE;

CORRELATION ANALYSIS

2 'VAR' Variables: HOEJDE VAEGT

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
HOEJDE	14	176.4	7.9383	2469.0	164.0	191.0
VAEGT	13	72.5615	9.6012	943.3	52.4000	90.5000

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0
/ Number of Observations

	HOEJDE	VAEGT
HOEJDE	1.00000 0.0 14	0.74894 0.0032 13
VAEGT	0.74894 0.0032 13	1.00000 0.0 13

UNIVARIATE PROCEDURE

Variable=HOEJDE

Moments

N	14	Sum Wgts	14
Mean	176.3571	Sum	2469
Std Dev	7.938292	Variance	63.01648
Skewness	0.230822	Kurtosis	-0.65947
USS	436245	CSS	819.2143
CV	4.501259	Std Mean	2.121598
T:Mean=0	83.12468	Prob> T	0.0001
Sgn Rank	52.5	Prob> S	0.0001
Num ^= 0	14		

UNIVARIATE PROCEDURE

Variable=HOEJDE

Quantiles(Def=5)

100% Max	191	99%	191
75% Q3	182	95%	191
50% Med	175.5	90%	187
25% Q1	170	10%	166
0% Min	164	5%	164
		1%	164
Range	27		
Q3-Q1	12		
Mode	164		

```

*-----;
*   SASKURS7.SAS 1                                KALØ 18.1.1993   ;
*-----;
*   Eksempel på SAS-anvendelse                               ;
*-----;

options linesize=68 pagesize=70;
libname perma 'f:\ff\ffmrp';

data temp;
  set perma.matmerg;
  vaegt=vaegt/10;
  if mod(persnr,2)=1 then koen='M';
  else koen='K';
  alder = 92 - mod(foeddag,100);
  if alder > 50 then list; /* bliver skrevet til LOG-filen */
  if alder > 50 then delete;
  jage_ald = mod(jaeger,100) - alder;
  BT_mean = (systole + diastole) / 2;
  drop systole diastole; /* smider variablene SYSTOLE
                        og DIASTOLE væk */

proc sort;
  by koen;

proc print split = '//';
  id navn;
  by koen;
  var foeddag alder hoejde vaegt;
  title2;
  title3;
  label foeddag = 'FØDT';
  label hoejde = 'HØJDE';
  label vaegt = 'VÆGT /I KG';
  format alder 3.0 vaegt 4.1 hoejde 4. foeddag Z6.0;

proc corr;
  var hoejde vaegt;

proc univariate;
  var hoejde;

options linesize=68 pagesize=25;
*** pagesize=25 giver netop en side, der fylder hele skærmen ***;

proc chart;
  hbar koen;
  title2 '---> proc chart <---';
  title3 '---> hbar koen <---';

proc chart;
  vbar koen / type = percent;
  title2 '---> proc chart <---';
  title3 '---> vbar koen / type = percent <---';

proc chart;
  vbar koen / type = mean sumvar = alder;
  title2 '---> proc chart <---';
  title3 '---> vbar koen / type = mean sumvar = alder <---';

proc chart;
  hbar alder;
  title2 '---> proc chart <---';
  title3 '---> hbar alder <---';

```

¹ Dette program er en omskrivning af programmer anvendt til øvelse 8 i bogen EDB for biologer, forår 1991, bind 2 af Jørgen Gomme, Henrik Nielsen, Peter S. Jørgensen og Mogens Ring Petersen.

UNIVARIATE PROCEDURE

Variable=HOEJDE

Extremes

Lowest	Obs	Highest	Obs
164(	12)	180(	11)
166(	4)	182(	14)
169(	3)	184(	1)
170(	9)	187(	13)
172(	7)	191(	8)

```
--> proc chart <--
--> hbar koen <--
```

FREQUENCY OF KOEN

KOEN		FREQ	CUM FREQ	PERCENT	CUM PERCENT
K	*****	5	5	35.71	35.71
M	*****	9	14	64.29	100.00

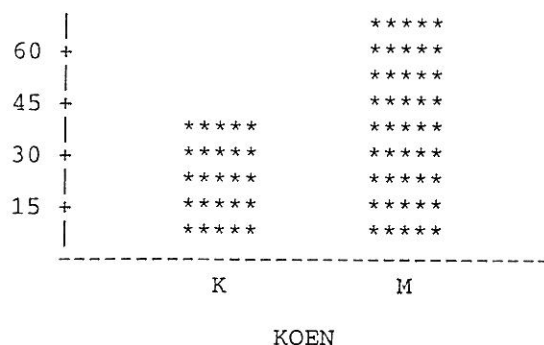
-----+-----+-----+-----+-----
2 4 6 8

FREQUENCY

```
--> proc chart <--
--> vbar koen / type = percent <--
```

PERCENTAGE OF KOEN

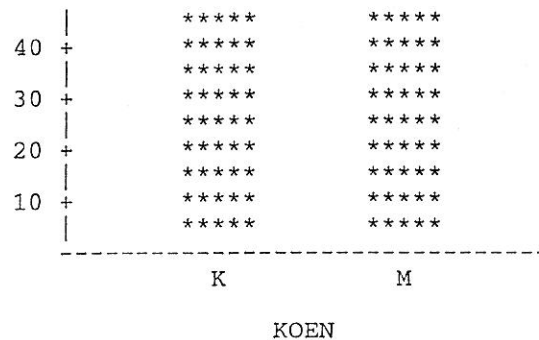
PERCENTAGE



```
--> proc chart <--
--> vbar koen / type = mean sumvar = alder <--
```

MEAN OF ALDER BY KOEN

ALDER MEAN



```
--> proc chart <--
--> hbar alder <--
```

FREQUENCY OF ALDER

ALDER MIDPOINT		FREQ	CUM FREQ	PERCENT	CUM PERCENT
39	*****	2	2	14.29	14.29
41	*****	3	5	21.43	35.71
43	*****	1	6	7.14	42.86
45	*****	4	10	28.57	71.43
47	*****	4	14	28.57	100.00
	-----+-----+-----+-----+				
	1 2 3 4				

FREQUENCY

```
--> hbar alder <--
--> by koen <--
```

----- KOEN=K -----

FREQUENCY OF ALDER

ALDER MIDPOINT		FREQ	CUM FREQ	PERCENT	CUM PERCENT
38.75	*****	1	1	20.00	20.00
41.25	*****	1	2	20.00	40.00
43.75	*****	1	3	20.00	60.00
46.25	*****	2	5	40.00	100.00
	-----+-----+-----+				
	1 2				

FREQUENCY

SASGRAPH - sådan får du grafik ud:

1) Direkte til printeren: (standardmetoden)

Gå i office ind i undermenu **U** og videre til **N** netprintervalg (eller skriv blot **printer** i DOS. Vælg derefter den printer, du vil bruge ved at taste fx **4** for LJ2_546K, når meddelelsen

VÆLG printer - tast nummer:

kommer på skærmen. Forbind den derpå til LPT1:, ved at skrive tallet **1**, ved næste spørgsmål

vælg printerport

Herpå bliver printet automatisk sendt til printeren, når du inde i programmet har skrevet: (ovenstående printer-eksempel)

```
goptions cback=none colours=(black)
          noprompt device=HPLJS2;
```

2) Printet bliver sendt til fil:

I SASKURS9.SAS er det vist, hvordan man kan sende printet til en fil, som derefter kan kopieres **binært** til printeren. **HUSK!!!** at angive option **/B** efter **COPY** (altså **COPY /B**) **!!!**, da der ellers vil komme en 1-2 cm tyk bunke af printerpapir ud med nogle få kruseduller på hver side. Det kan nogle gange være fordelagtigt at anvende denne metode, hvis man er ude for, at figuren bliver "klippet over".

3) Hente filen ind i Word Perfect:

I Word Perfect: 1) <Alt-F9> 2) Figur ... osv.

Alle figurer i dette skrift er lavet på denne måde.

- - -

I næste program, kan du godt skrive hele blokken ind i dit program og så fjerne de enkelte stjerner '*' ud for GOPTIONS, når du får brug for dem - og ellers vil SAS bare opfatte sætningen (dvs. det, der står ud til ';') som en kommentar. Du behøver ikke at sætte en stjerne ud for den option, der angiver skærmtypen. Når du bruger en af de andre options til en af printerne, kommer der nemlig ikke noget på skærmen. Du kan dog komme ud for ikke at have hukommelse nok, og så skal du sætte en stjerne ud for skærm-devicen.

Et helt simpelt SAS-program med SASGRAPH ser sådan ud (næste side):

```
*-----;
* SASKURS8.SAS KALØ 18.1.1993 ;
*-----;
* Eksempel på SASGRAPH (simpelt eksempel)
*-----;
```

```
*---- NULSTILLING -----;
options reset=all;
```

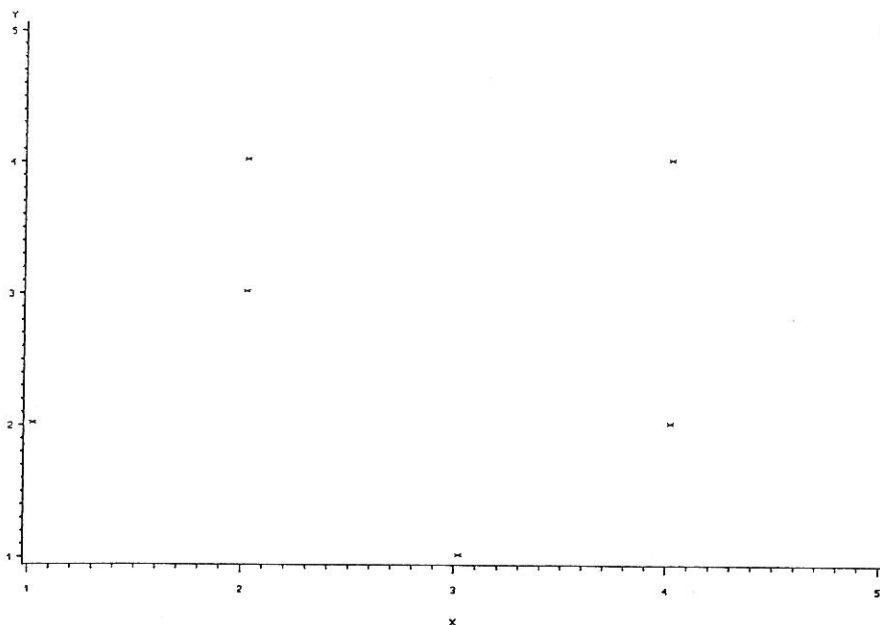
```
*----- SKÆRM -----;
options device=ps2ega;
```

```
data temp;
input x y;
```

```
cards; Alle figurer på side 22-24 er lavet med disse data:
```

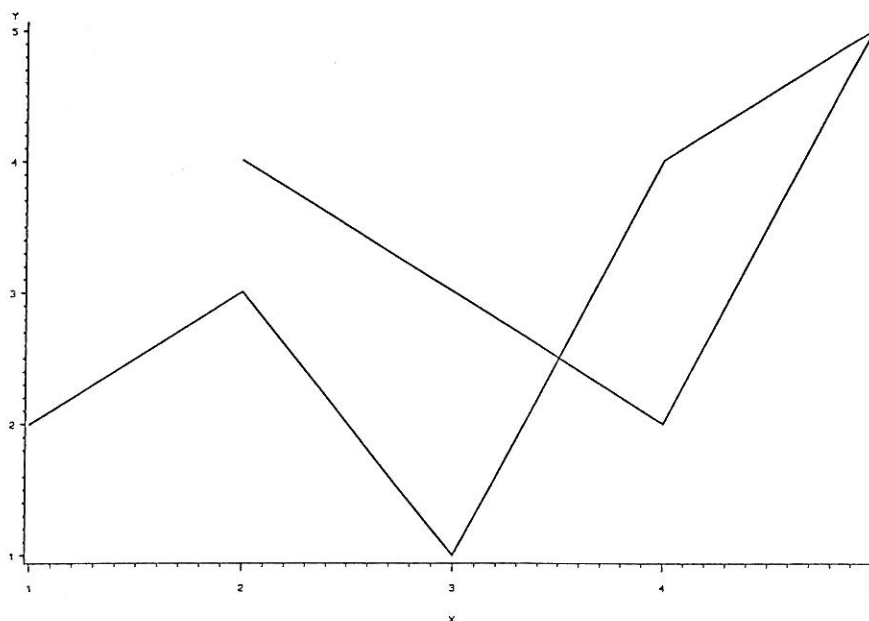
```
1 2
2 3
3 1
4 4
5 5
4 2
2 4
;
```

```
proc gplot;
plot y * x = 1;
symbol1 V=X;
```

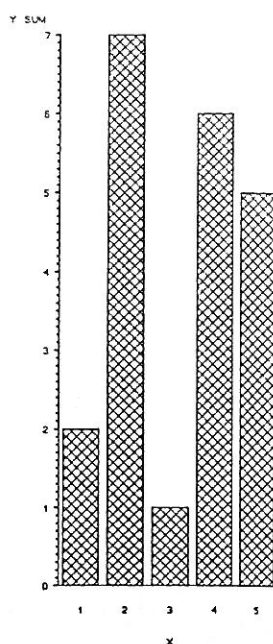


/* Du kan godt lave tallene større, sætte overskrifter på osv. - se nærmere i SASGRAPH-manualen. En lille del af faciliteterne er vist i SASKURS9.SAS */

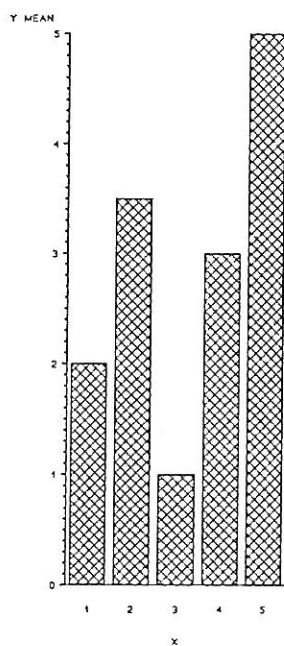
```
proc gplot;
  plot y * x = 1;
  symbol1 I=j; /* 'j' står for 'join', dvs. linje mellem hver punkt */
```



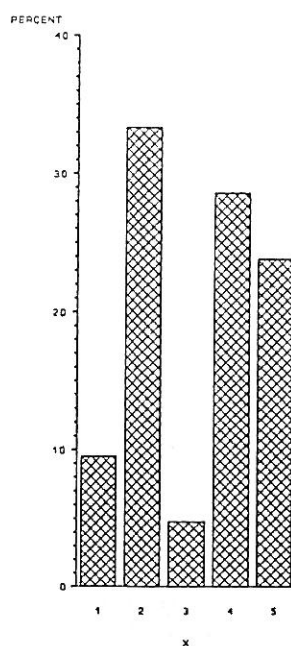
```
proc gchart;
  vbar x / sumvar=y discrete; /* lægger y-værdierne sammen */
  /* DISCRETE, ellers laver SAS en inddeling,
  , hvor der godt kan være decimaler i
  (sammenlign med ALDER i figur, side 2) */
```




```
proc gchart;
  vbar x / sumvar=y type=mean discrete; /* gennemsnit */
```



```
proc gchart;
  vbar x / freq=y type=percent discrete; /* procent */
run;
```



```

*-----;
*   SASKURS9.SAS                                KALØ 18.1.1993   ;
*-----;
*   Eksempel på SASGRAPH (flere options) ;
*-----;

*----- NULSTILLING -----;
  goptions reset=all;

*----- SKÆRM -----;
  goptions device=ps2ega;

*----- HP-PLOTTER -----;
* goptions gaccess='sasgastd>plotfil.hp'
  noprompt device=HP7550A gsfmodes=replace ;

*----- WORD-PERFECT -----;
  goptions gaccess='sasgastd>plotfil.hp'
  cback=none colours=(black)
  noprompt device=HP7550A gsfmodes=replace ;

*----- HP-PRINTER II -----;
* goptions gaccess='sasgastd>filnavn'
  vsize=11 cm hsize=17 cm
  cback=none colours=(black)
  noprompt device=HPLJS2 gsfmodes=replace;
*-----*
!   NB!!!                                     !
!   SKAL SKRIVES UD MED 'COPY /B filnavn LPT1:' !
*-----*

*----- HP-PRINTER III -----;
* goptions gaccess='sasgastd>filnavn'
  vsize=11 cm hsize=17 cm
  cback=none colours=(black)
  noprompt device=HPLJ5P3 gsfmodes=replace;
*-----*
!   NB!!!                                     !
!   SKAL SKRIVES UD MED 'COPY /B filnavn LPT1:' !
*-----*

*----- HP-PRINTER 4 -----;
* (som for HP-PRINTER III);

data temp;
  input x @@;
  do I=1 to 20;
    input y @@;
    output;
  end;

cards;
1 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3
2 2 3 4 2 3 4 2 3 4 2 3 4 2 3 4 2 3 4 2 3
3 2 3 4 5 2 3 4 5 2 3 4 5 2 3 4 5 2 3 4 5
4 2 3 4 5 6 2 3 4 5 6 2 3 4 5 6 2 3 4 5 6
5 2 3 4 5 6 7 2 3 4 5 6 7 2 3 4 5 6 7 2 3
6 2 3 4 5 6 7 8 2 3 4 5 6 7 8 2 3 4 5 6 7
7 2 3 4 5 6 7 8 9 2 3 4 5 6 7 8 9 2 3 4 5
;

```

```
proc gplot;
plot y * x = 1;

symbol1 V=point C=red I=std1T;
title1 F=cartog H=3.0 C=green 'MMNMNNN ' C=yellow F=swiss ' DMU '
      F=cartog H=3.0 C=green ' NMNMNNN';
title2 draw=(10 pct, 90 pct, 90 pct, 90 pct);
title3 F=script H=2.5 C=red A=10 'Der bliver vist een'
      ' gang standardafvigelse' A=-135 ' -->';
footnote F=simplex H=2.0 C=white 'SASGRAHP-eksempel';

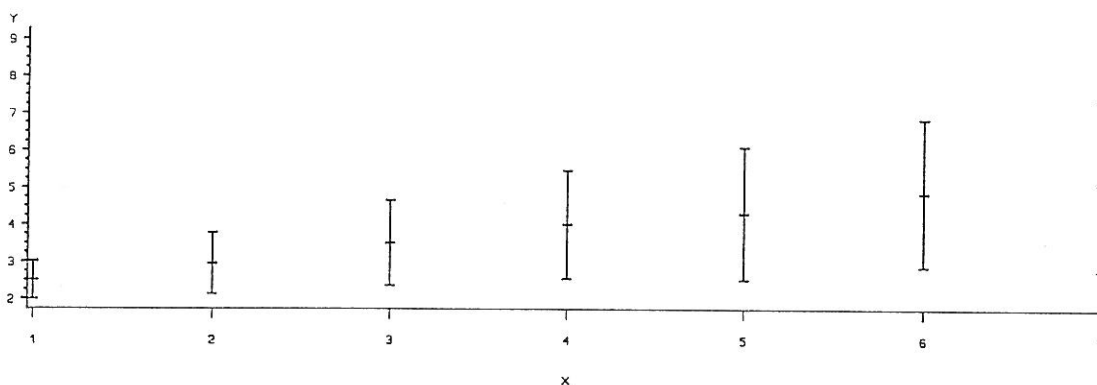
run;
```

Og her er så resultatet: (dog uden farver)

Denne figur viser, hvordan man kan lege med SASGRAPH, hvis man er træt af computerspil. Resultatfilen er hentet ind i et grafikfelt i Word-Perfect.



Der bliver vist en gang standardafvigelse



SASGRAHP-eksempel

```

*-----;
*   SASKURS10.SAS                                KALØ 18.1.1993 ;
*-----;
*   Eksempel på statistisk anvendelse - Chi2-test
*-----;

```

```
/*
```

Man har lavet nogle forsøg med forskellige hagltyper med
30 blyhaglspatroner og 19 jernhaglspatroner med følgende resultat:

	Dræbt	Ikke dræbt
Bly5	19	11
Jern4	13	6

Vi vil nu undersøge, om der er forskel på de to hagltyper vha en chi2-test.

```
*/
```

```
options ps=500;
```

```
data temp1;
```

```
  input Bly5 Jern4 antal ;
```

```
cards;
```

```
1 1 19
```

```
1 2 11
```

```
2 1 13
```

```
2 2 6
```

```
;
```

```
proc freq formchar='|---||| |';
```

```
  tables bly5*jern4 / expected deviation cellchi2 chisq
                        norow nocol nopercnt;
```

```
weight antal;
```

* WEIGHT, denne option bruges fordi vi allerede har de
sammenregnede tal. Herved får vi selve tallene (19,
11, 13 og 6) med - hvis den undlades, får hver rude
bare værdien 1 (, da det er frekvensen af tal i hver
rude);

```

*-----;
*   LISTINGFIL:
*-----;/*

```

TABLE OF BLY5 BY JERN4

BLY5	JERN4		
Frequency			
Expected			
Deviation			
Cell Chi-Square	1	2	Total
1	19	11	30
	19.592	10.408	
	-0.592	0.5918	
	0.0179	0.0337	
2	13	6	19
	12.408	6.5918	
	0.5918	-0.592	
	0.0282	0.0531	
Total	32	17	49

STATISTICS FOR TABLE OF BLY5 BY JERN4

Statistic	DF	Value	Prob
Chi-Square	1	0.133	0.715
Likelihood Ratio Chi-Square	1	0.134	0.715
Continuity Adj. Chi-Square	1	0.003	0.955
Mantel-Haenszel Chi-Square	1	0.130	0.718
Fisher's Exact Test (Left)			0.480
(Right)			0.748
(2-Tail)			0.767
Phi Coefficient		-0.052	
Contingency Coefficient		0.052	
Cramer's V		-0.052	

Sample Size = 49

*/

```

*-----;
* SASKURS11.SAS KALØ 18.1.1993 ;
*-----;
* Eksempel på statistisk anvendelse - Variansanalyse ;
*-----;

```

* Vi vil så undersøge, om der er forskel på vægtene på nogle dyr fra 3 forskellige områder B, L og K;

Data;

```

input omraade $ @@;
do I=1 to 20;
  input vaegt @@;
  output;
end;

```

* Data bliver nødt til at stå på 2 linjer, da man ikke kan læse ud over kolonne 80 i Cards;

Cards;

```

B 16.0 15.5 20.5 16.5 19.0 18.5 14.5 17.0 16.0 17.5
   18.5 17.0 17.0 15.5 16.5 16.5 18.5 15.0 15.5 .
L 15.5 15.0 18.5 16.5 16.0 16.5 17.0 18.5 17.0 16.0
   15.5 20.5 20.5 16.5 18.5 14.5 . . . .
K 14.9 16.5 11.7 13.0 12.0 11.2 14.3 13.0 13.0 15.0
   12.6 11.2 10.2 11.0 10.2 11.0 15.0 12.2 13.2 15.2
;

```

proc glm;

```

class omraade;
model vaegt=omraade / solution; /* SOLUTION: parameter estimator */
means omraade / Tukey; /* multiple sammenligninger,
                        Tukey's Studentized Range (HSD) */

```

```

*-----;
* LISTINGFIL: ;
*-----; /*

```

General Linear Models Procedure
Class Level Information

Class	Levels	Values
OMRAADE	3	B K L

Number of observations in data set = 60

NOTE: Due to missing values, only 55 observations can be used in this analysis.

General Linear Models Procedure

Dependent Variable: VAEGT

Source	DF	Sum of Squares	F Value	Pr > F
Model	2	218.00124222	36.58	0.0001
Error	52	154.95584868		
Corrected Total	54	372.95709091		
	R-Square	C.V.	VAEGT Mean	
	0.584521	11.17113	15.45272727	

Source	DF	Type I SS	F Value	Pr > F
OMRAADE	2	218.00124222	36.58	0.0001
Source	DF	Type III SS	F Value	Pr > F
OMRAADE	2	218.00124222	36.58	0.0001

Parameter	Estimate	T for H0: Parameter=0	Pr > T	Std Error of Estimate
INTERCEPT	17.03125000 B	39.46	0.0001	0.43156113
OMRAADE B	-0.13651316 B	-0.23	0.8166	0.58573280
K	-4.21125000 B	-7.27	0.0001	0.57900002
L	0.00000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular and a generalized inverse was used to solve the normal equations. Estimates followed by the letter 'B' are biased, and are not unique estimators of the parameters.

General Linear Models Procedure

Tukey's Studentized Range (HSD) Test for variable: VAEGT

NOTE: This test controls the type I experimentwise error rate.

Alpha= 0.05 Confidence= 0.95 df= 52 MSE= 2.97992
Critical Value of Studentized Range= 3.412

Comparisons significant at the 0.05 level
are indicated by '***'.

OMRAADE Comparison		Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
L	- B	-1.277	0.137	1.550	
L	- K	2.814	4.211	5.608	***
B	- L	-1.550	-0.137	1.277	
B	- K	2.740	4.075	5.409	***
K	- L	-5.608	-4.211	-2.814	***
K	- B	-5.409	-4.075	-2.740	***

*/

```

*-----;
* SASKURS12.SAS KALØ 18.1.1993 ;
*-----;
* Eksempel på statistisk anvendelse - Lineær regression ;
*-----;

```

```
libname perma 'F:\FF\FMRP';
```

```
data temp;
  set perma.matmerg; /* vores datasæt fra side 13 */
  vaegt=vaegt/10;
```

```
proc reg;
  model vaegt=hoejde;
  output out=nydata R=resid P=yhat L95=nedre U95=underste;
  /* L95 og U95 bruger vi ikke i dette eksempel, men blot taget
     med for at vise, at disse kan man også få ud */
```

```

*-----;
* LISTINGFIL: ;
*-----; /*

```

```
Model: MODEL1
Dependent Variable: VAEGT
```

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	620.47878	620.47878	14.052	0.0032
Error	11	485.71199	44.15564		
C Total	12	1106.19077			

Root MSE	6.64497	R-square	0.5609
Dep Mean	72.56154	Adj R-sq	0.5210
C.V.	9.15770		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0
INTERCEP	1	-93.788402	44.41467282	-2.112
HOEJDE	1	0.939014	0.25049650	3.749

Variable	DF	Prob > T
INTERCEP	1	0.0584
HOEJDE	1	0.0032

```
*/
```

```

*---- NULSTILLING ----;
  goptions reset=all;

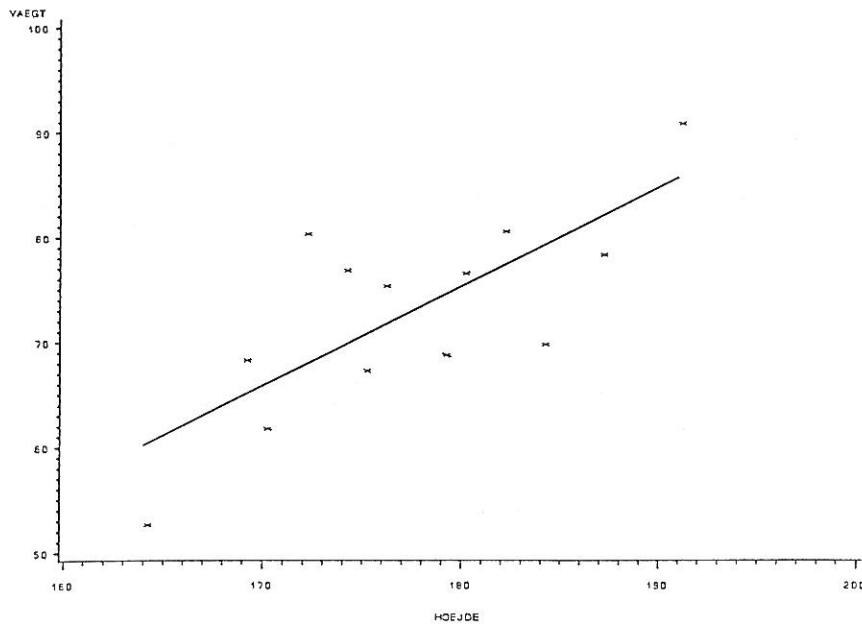
```

```

*----- SKÆRM -----;
  goptions device=ps2ega;

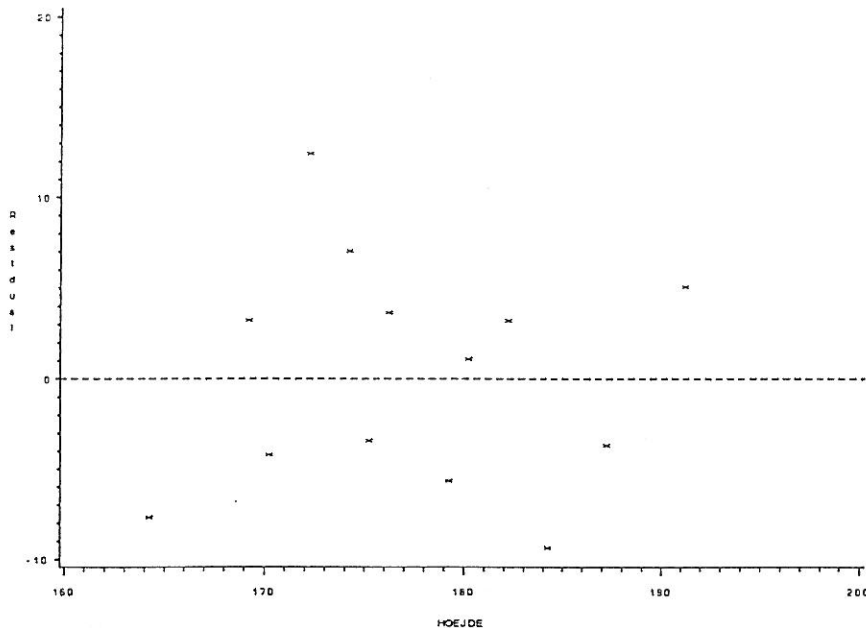
```

```
proc gplot;
  plot vaegt*hoejde=1 yhat*hoejde=2 / overlay;
  symbol1 V=X I=none c=red;
  symbol2 V= I=j c=green;
```



/* Modelkontrol. Hvis linjen er rigtig må afvigelserne (RESID) fordele sig jævnt på hver side af 0 */

```
proc gplot;
  plot resid*hoejde= 1 / vref=0 LV=2;
  symbol1 V=X; /* VREF giver vandret linje (Y=0),
               LV=2 (linjetype, nr. 2 er stiplet) */
run;
```



God fornøjelse!