

Answer to exercise BOC

Question 1

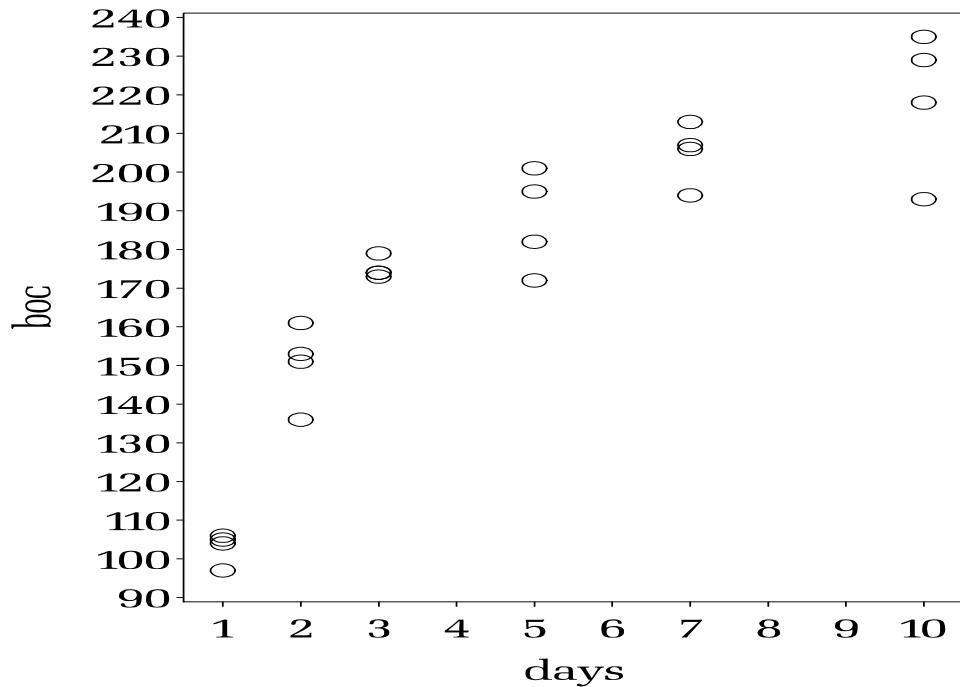
The reading of data may be performed as:

```
data sasuser.boc;
input days boc;
datalines;
 1      105
 1      97
 1      104
 .
 .
 .
 .
 10     193
 10     235
 10     229
;
run;
```

and a scatter plot with a reasonable appearance can be obtained from the code

```
proc gplot data=sasuser.boc;
  plot boc*days
    / haxis=axis1 vaxis=axis2 frame;
axis1
  value=(H=3)
  offset=(3,3)
  minor=NONE
  label=(H=3);
axis2
  value=(H=3)
  minor=NONE
  label=(A=90 R=0 H=3);
symbol1 v=circle i=none c=BLACK l=1 h=3 w=2;
run;
```

resulting in the plot below



The above plot shows that `boc` as a function of `days` is certainly **not linear**.

The biologists claim that the relation between `boc` and `days` can be described by a relation of the form

$$\text{boc} = \gamma \exp(-\beta/\text{days})$$

This relation is obviously nonlinear, but may be transformed to linearity as we shall see below (question 3).

Question 2

As time passes (when `days` $\rightarrow \infty$), the term $\exp(-\beta/\text{days})$ gets close to 1, so `boc` $\rightarrow \gamma$. The parameter γ may therefore be interpreted as the maximal oxygen loss for this volume of waste water.

The parameter β is inversely related to the velocity of the oxygen loss. If β is large, the process is slow in reaching its maximum. More precisely, the time needed for the process to increase to half of its maximum is $\frac{\beta}{\log(2)}$ (this result may be obtained by setting $\text{boc}=\frac{\gamma}{2}$)

Question 3

If we transform the above equation with the natural logarithm, we get

$$\log(\text{boc}) = \log(\gamma) - \beta/\text{days}$$

With

$$\begin{aligned} y &= \log \text{boc} = \log(\text{boc}) \\ x &= \text{invdays} = 1/\text{days} \text{ and} \\ \alpha &= \log(\gamma) \end{aligned}$$

we may write this equation as

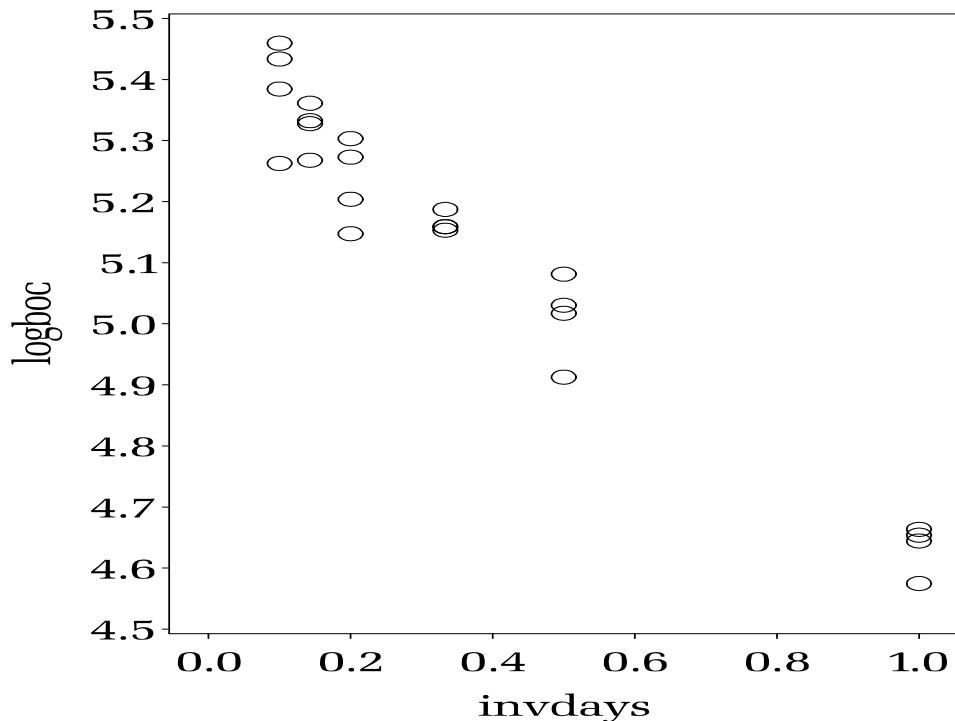
$$y = \alpha - \beta x$$

i.e. a **linear relation**, only with a minus sign on the slope.
We write

```
data boc;
set sasuser.boc;

logboc=log(boc);
invdays=1/days;
run;
```

and make a scatter plot of these new variables



This plot gives a totally different basis for doing a linear regression. The relation looks straight and the variation seems to be reasonably constant over time.

Question 4.

The analysis may be performed using `reg` in SAS:
based on the analysis

```
proc reg data=boc;
  model logboc=invdays / clb;
run;
```

with the result

Dependent Variable: logboc

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1.48266	1.48266	434.02	<.0001
Error	22	0.07515	0.00342		
Corrected Total	23	1.55781			
Root MSE	0.05845	R-Square	0.9518		
Dependent Mean	5.12480	Adj R-Sq	0.9496		
Coeff Var	1.14048				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	5.43125	0.01894	286.76	<.0001
invdays	1	-0.80781	0.03878	-20.83	<.0001

Parameter Estimates

Variable	DF	95% Confidence Limits
Intercept	1	5.39197 - 5.47053
invdays	1	-0.88823 - 0.72740

The linear regression model gives us the estimates:

$$\begin{aligned} \text{intercept} : \hat{\alpha} &= \log(\hat{\gamma}) = 5.431(0.019) \\ \text{slope} : \hat{\beta} &= -0.808(0.039) \end{aligned}$$

Since $\alpha = \log(\gamma)$, we therefore have:

$$\log(\hat{\gamma}) = 5.431(0.019)$$

with corresponding 95% confidence limits (5.392, 5.471)

By noting that $\text{boc}(\infty) = \gamma = \exp(\alpha)$,

we may transform the above estimate of α with confidence interval directly back to an estimate with confidence interval for $\text{boc}(\infty)$ simply by taking exponentials of the estimate as well as the end points of the confidence interval.

In this way, we find the estimate of $\text{boc}(\infty)$ to be $\exp(5.431) = 228.38$ with the 95% confidence interval
 $(\exp(5.392), \exp(5.471)) = (219.6, 237.7)$

Question 5

We may perform a formal **test for linearity** by checking whether an ANOVA-type model (allowing every single day to have its own mean value) fits much better, i.e. if we can make the model reduction from the ANOVA-model

```
proc glm data=boc;
  class days;
  model logboc=days;
run;
```

to the linear regression model.

This may be done by formally including both effects in the same model and then test the significance of the **class** variable **days**:

```
proc glm data=boc;
  class days;
  model logboc=invdays days / solution;
run;
```

The test below shows that linearity is a **reasonable description**, since the extra flexibility allowing each day to have its own mean with no relation to the means of the other days, does not give us a significantly better description ($P=0.49$).

Class Level Information

Class	Levels	Values
-------	--------	--------

days	6	1 2 3 5 7 10
------	---	--------------

Number of observations in data set = 24

Dependent Variable: logboc

Source	DF	Sum of Squares		Mean Square		F Value	Pr > F
Model	5	1.4949614		0.2989923		85.63	0.0001
Error	18	0.0628520		0.0034918			
Corrected Total	23	1.5578134					

R-Square	C.V.	Root MSE	logboc Mean
0.959654	1.153046	0.0591	5.1248

Source	DF	Type I SS	Mean Square	F Value	Pr > F
invdays	1	1.4826592	1.4826592	424.61	0.0001
days	4	0.0123022	0.0030756	0.88	0.4949

Source	DF	Type III SS	Mean Square	F Value	Pr > F
invdays	0	0.0000000	.	.	.
days	4	0.0123022	0.0030756	0.88	0.4949

However, this test is not very powerful, since it does not use the *ordering* of the days. A much more powerful test of linearity will

be to test against the alternative of a second order polynomium (a parabola):

```

data boc;
set boc;

invdays2=invdays**2;
run;

proc reg;
model logboc=invdays invdays2;
run;

```

with the output

```

The REG Procedure
Dependent Variable: logboc

Number of Observations Used           24

Analysis of Variance

Source          DF      Sum of Squares      Mean Square      F Value      Pr > F
Model            2        1.48645       0.74322       218.71      <.0001
Error           21        0.07136       0.00340
Corrected Total  23        1.55781

Root MSE         0.05829      R-Square       0.9542
Dependent Mean   5.12480      Adj R-Sq       0.9498
Coeff Var        1.13750

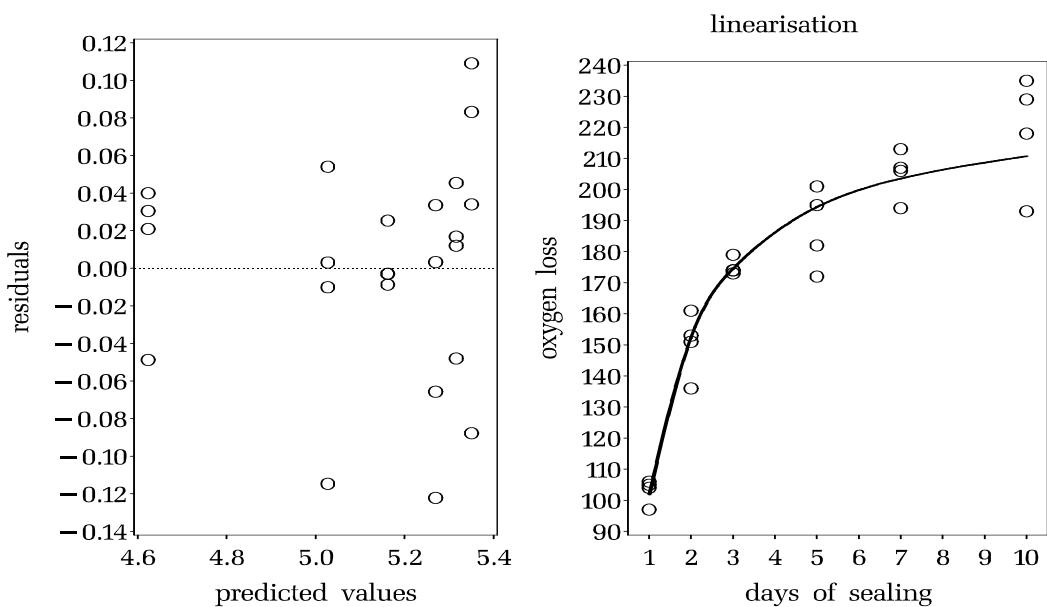
Parameter Estimates

Variable        DF      Parameter Estimate      Standard Error      t Value      Pr > |t|
Intercept       1        5.46146       0.03427       159.36      <.0001
invdays         1       -0.99135       0.17803       -5.57      <.0001
invdays2        1        0.16525       0.15646       1.06       0.3029

```

Since this test also fails to reject linearity, and since the relation looks really straight to the eye, we conclude that the linear description is adequate.

Below, you can see a residual plot and the actual nonlinear fit:



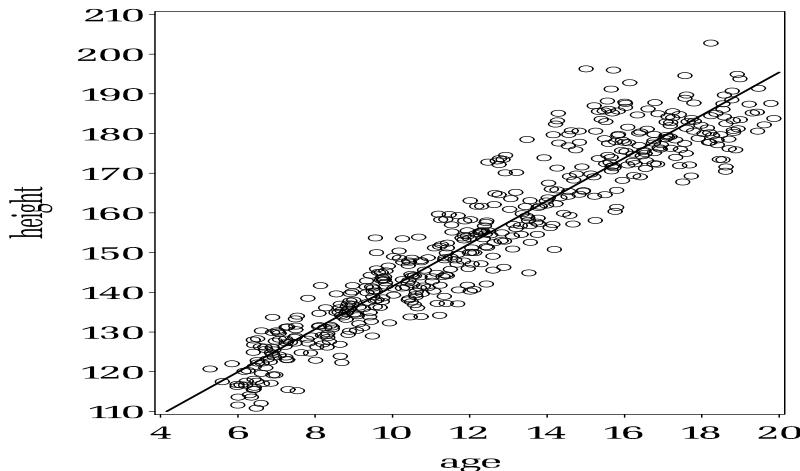
Answer to exercise 'height vs. age' (Juul)

Question 1

Fitting a straight line to height for males in the age range 5-20 and making the corresponding illustration is performed by writing:

```
proc reg data=juul;
  where age ge 5 and age le 20 and sex='male';
  model height=age;
run;

proc gplot data=juul;
where age ge 5 and age le 20 and sex='male';
  plot height*age
    / haxis=axis1 vaxis=axis2 frame;
axis1 value=(H=3) minor=NONE label=(H=3);
axis2 value=(H=3) minor=NONE label=(A=90 R=0 H=3);
symbol1 v=circle i=rl c=black h=2 l=1 w=2;
run;
```

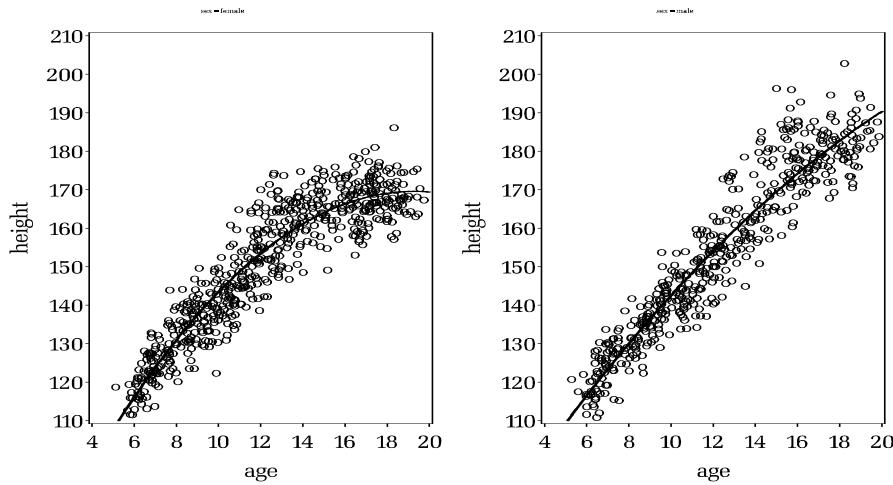


It looks as if the variation increases somewhat with age, but for ease of interpretation, we shall here abstain from performing any transformation. More importantly, it looks as if linearity is not entirely adequate, at least not extending beyond the age of 16. This is indeed not very surprising.

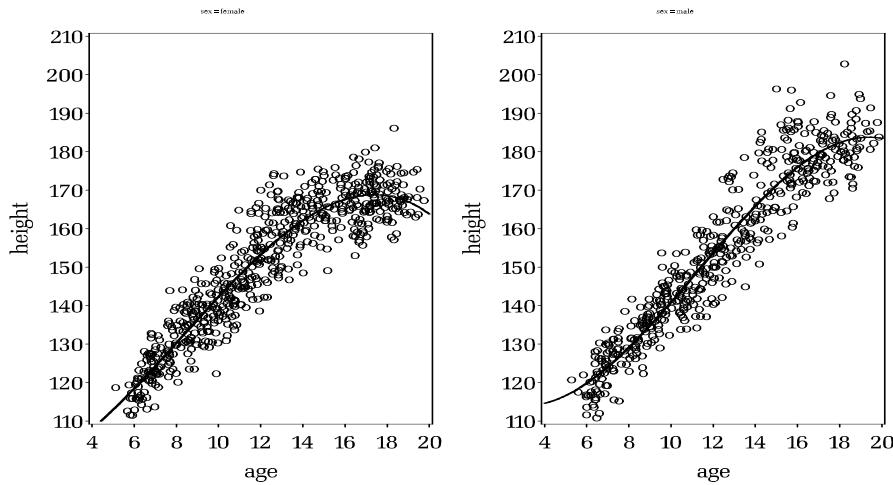
Question 2

Graphical fits of parabolas and cubic functions may be performed simply by using `i=rq` or `i=rc` in the `symbol`-statement in `gplot`. We get

Quadratic fits:



Cubic fits:



In order to fit the models for males and females simultaneously, we may use two different parametrisations. First, we choose the estimation friendly version:

```

proc glm data=juul; where age ge 5 and age le 20;
  class sex;
  model height=sex sex*age sex*age2 sex*age3 / noint solution;
run;

```

which gives

```

The GLM Procedure
Dependent Variable: height

Sum of
Source          DF      Squares      Mean Square   F Value   Pr > F
Model           8       26516958.98    3314619.87   82370.4   <.0001
Error          1122      45149.75        40.24
Uncorrected Total 1130      26562108.73

R-Square      Coeff Var      Root MSE      height Mean
0.894745      4.171330      6.343533      152.0746

Source          DF      Type I SS      Mean Square   F Value   Pr > F
sex              2       26136381.25    13068190.63   324753   <.0001
age*sex          2       367006.25     183503.12    4560.17   <.0001
age2*sex         2       11576.94      5788.47     143.85   <.0001
age3*sex         2       1994.55      997.27     24.78   <.0001

Source          DF      Type III SS      Mean Square   F Value   Pr > F
sex              2       10227.77296    5113.88648   127.08   <.0001
age*sex          2       202.69405     101.34703    2.52    0.0810
age2*sex         2       1315.69653    657.84826    16.35   <.0001
age3*sex         2       1994.54539    997.27270   24.78   <.0001

Parameter        Estimate      Standard
Parameter        Estimate      Error      t Value   Pr > |t|
sex female      95.5763245    9.28773561   10.29   <.0001
sex male        122.7018171   10.07682834   12.18   <.0001
age*sex female   1.0495601    2.47354806   0.42    0.6714
age*sex male    -5.8700641    2.66352795   -2.20   0.0277
age2*sex female  0.6060279    0.20792957   2.91    0.0036
age2*sex male   1.0928101    0.22214015   4.92    <.0001
age3*sex female -0.0243940   0.00555157   -4.39   <.0001
age3*sex male   -0.0323690   0.00588451   -5.50   <.0001

```

or with the test friendly parametrisation:

```

proc glm data=juul; where age ge 5 and age le 20;
  class sex;
  model height=age age2 age3 sex sex*age sex*age2 sex*age3 / solution;
run;

```

which gives

The GLM Procedure
Dependent Variable: height

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	383805.4938	54829.3563	1362.54	<.0001
Error	1122	45149.7473	40.2404		
Corrected Total	1129	428955.2411			

R-Square	Coeff Var	Root MSE	height Mean
0.894745	4.171330	6.343533	152.0746

Source	DF	Type I SS	Mean Square	F Value	Pr > F
age	1	360150.8657	360150.8657	8949.98	<.0001
age2	1	9798.6726	9798.6726	243.50	<.0001
age3	1	1583.1409	1583.1409	39.34	<.0001
sex	1	3611.3187	3611.3187	89.74	<.0001
age*sex	1	6850.1039	6850.1039	170.23	<.0001
age2*sex	1	1772.2879	1772.2879	44.04	<.0001
age3*sex	1	39.1042	39.1042	0.97	0.3245

Source	DF	Type III SS	Mean Square	F Value	Pr > F
age	1	70.770426	70.770426	1.76	0.1851
age2	1	1254.425022	1254.425022	31.17	<.0001
age3	1	1981.075696	1981.075696	49.23	<.0001
sex	1	157.656447	157.656447	3.92	0.0480
age*sex	1	145.824982	145.824982	3.62	0.0572
age2*sex	1	102.993585	102.993585	2.56	0.1099
age3*sex	1	39.104156	39.104156	0.97	0.3245

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	122.7018171 B	10.07682834	12.18	<.0001
age	-5.8700641 B	2.66352794	-2.20	0.0277
age2	1.0928101 B	0.22214015	4.92	<.0001
age3	-0.0323690 B	0.00588451	-5.50	<.0001
sex female	-27.1254926 B	13.70417827	-1.98	0.0480
sex male	0.0000000 B	.	.	.

age*sex	female	6.9196242 B	3.63494444	1.90	0.0572
age*sex	male	0.0000000 B	.	.	.
age2*sex	female	-0.4867822 B	0.30427119	-1.60	0.1099
age2*sex	male	0.0000000 B	.	.	.
age3*sex	female	0.0079749 B	0.00808995	0.99	0.3245
age3*sex	male	0.0000000 B	.	.	.

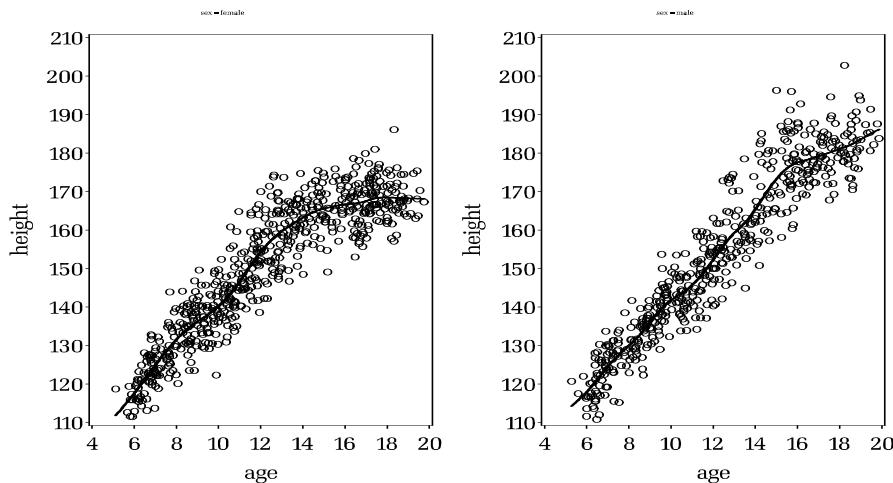
NOTE: The $X'X$ matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

We see that the two genders do not deviate in the coefficient of the third order term ($P=0.32$), but (seen from Type I) they do differ in the coefficient of the second order term ($F=44.04$, $P < 0.0001$).

From the estimation friendly parametrisation, we may also note that the linear term is not significant for girls ($P=0.67$), but this is rather hard to interpret.....

Question 3

Automatic fit with e.g. `i=sm40` gives the result



which pretty much resembles the parabolic fits.

Question 4.

We now consider cutpoints at the ages 10, 12, 13 and 15 and define variables

denoting number of years above these respective thresholds.

```

data juul;
set juul;

/* definition af new intercept, at the age 5 */
age5=age-5;
extra_age10=max(age-10,0);
extra_age12=max(age-12,0);
extra_age13=max(age-13,0);
extra_age15=max(age-15,0);
run;

```

We now use these new variables to define a linear spline, and at the same time we test linearity using a **contrast**-statement

```

proc glm data=juul; where age ge 5 and age le 20; by sex;
  classes tanner;
  model lheight=age5 extra_age15 extra_age13
        extra_age12 extra_age10 / solution;
  contrast 'all' extra_age10 1,
            extra_age12 1,
            extra_age13 1,
            extra_age15 1;
run;

```

From this we get

```

sex=male

The GLM Procedure

Number of Observations Read      506
Number of Observations Used     502

Dependent Variable: height

          Sum of
Source      DF      Squares      Mean Square      F Value      Pr > F
Model           5   215410.0237    43082.0047    1004.27    <.0001
Error         496   21277.7593      42.8987
Corrected Total 501   236687.7829

```

R-Square	Coeff Var	Root MSE	height Mean		
0.910102	4.254027	6.549711	153.9649		
Source					
	DF	Type I SS	Mean Square	F Value	Pr > F
age5	1	212089.1768	212089.1768	4943.95	<.0001
extra_age15	1	2642.2952	2642.2952	61.59	<.0001
extra_age13	1	605.1047	605.1047	14.11	0.0002
extra_age12	1	20.4972	20.4972	0.48	0.4897
extra_age10	1	52.9498	52.9498	1.23	0.2671
Source					
	DF	Type III SS	Mean Square	F Value	Pr > F
age5	1	11288.78957	11288.78957	263.15	<.0001
extra_age15	1	1054.91319	1054.91319	24.59	<.0001
extra_age13	1	0.03990	0.03990	0.00	0.9757
extra_age12	1	69.40845	69.40845	1.62	0.2040
extra_age10	1	52.94981	52.94981	1.23	0.2671
Contrast					
all	4	3320.846895	830.211724	19.35	<.0001
				Standard	
Parameter	Estimate	Error	t Value	Pr > t	
Intercept	112.4674646	1.25870755	89.35	<.0001	
age5	5.8573756	0.36107839	16.22	<.0001	
extra_age15	-5.3782734	1.08456759	-4.96	<.0001	
extra_age13	0.0728337	2.38830149	0.03	0.9757	
extra_age12	2.8737997	2.25929236	1.27	0.2040	
extra_age10	-1.0928339	0.98365782	-1.11	0.2671	

Question 5

We find the first deviation from the straight line at the age of 13 (based on the Type I tests, P=0.0002).

Question 6

For girls, we get similarly

sex=female	Sum of	Source	DF	Squares	Mean Square	F Value	Pr > F
The GLM Procedure							
Number of Observations Read	636						
Number of Observations Used	628						
Dependent Variable: height							

Model	5	166447.1678	33289.4336	916.50	<.0001
Error	622	22592.5271	36.3224		
Corrected Total	627	189039.6950			
R-Square	Coeff Var	Root MSE	height Mean		
0.880488	4.002832	6.026806	150.5635		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
age5	1	154917.0709	154917.0709	4265.06	<.0001
extra_age15	1	10750.4962	10750.4962	295.97	<.0001
extra_age13	1	652.0044	652.0044	17.95	<.0001
extra_age12	1	35.4492	35.4492	0.98	0.3236
extra_age10	1	92.1472	92.1472	2.54	0.1117
Source	DF	Type III SS	Mean Square	F Value	Pr > F
age5	1	11743.51638	11743.51638	323.31	<.0001
extra_age15	1	250.71123	250.71123	6.90	0.0088
extra_age13	1	79.06736	79.06736	2.18	0.1406
extra_age12	1	7.95188	7.95188	0.22	0.6400
extra_age10	1	92.14715	92.14715	2.54	0.1117
Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
all	4	11530.09691	2882.52423	79.36	<.0001
Parameter	Estimate	Standard Error	t Value	Pr > t	
Intercept	113.9969325	1.03724473	109.90	<.0001	
age5	5.3601137	0.29810011	17.98	<.0001	
extra_age15	-2.3762747	0.90447583	-2.63	0.0088	
extra_age13	-2.7555641	1.86766551	-1.48	0.1406	
extra_age12	-0.8536511	1.82445325	-0.47	0.6400	
extra_age10	1.3088112	0.82171887	1.59	0.1117	

and again, we see the first deviation at the age of 13 ($F=17.95, P < 0.0001$).

Question 7

A model for the two genders simultaneously is defined as

```
proc glm data=juul; where age ge 5 and age le 20;
  class sex;
  model height=age5 extra_age10 extra_age12 extra_age13 extra_age15
    sex
    sex*extra_age15 sex*extra_age13
    sex*extra_age12 sex*extra_age10 sex*age5 / solution;
run;
```

and yields

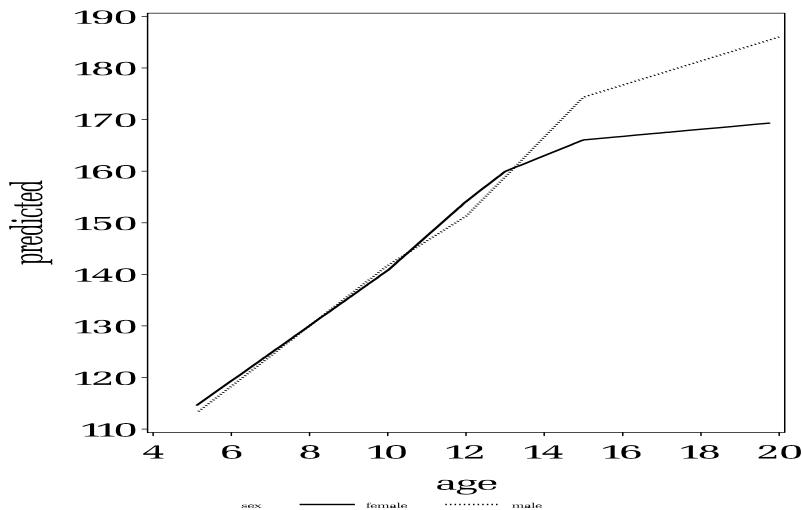
The GLM Procedure						
Dependent Variable: height						
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	
Model	11	385084.9547	35007.7232	892.14	<.0001	
Error	1118	43870.2864	39.2400			
Corrected Total	1129	428955.2411				
R-Square	Coeff Var	Root MSE	height Mean			
0.897728	4.119150	6.264182	152.0746			
Source	DF	Type I SS	Mean Square	F Value	Pr > F	
age5	1	360150.8657	360150.8657	9178.16	<.0001	
extra_age10	1	4713.9147	4713.9147	120.13	<.0001	
extra_age12	1	4298.0950	4298.0950	109.53	<.0001	
extra_age13	1	2213.3050	2213.3050	56.40	<.0001	
extra_age15	1	1120.8001	1120.8001	28.56	<.0001	
sex	1	3541.6014	3541.6014	90.25	<.0001	
extra_age15*sex	1	6894.9301	6894.9301	175.71	<.0001	
extra_age13*sex	1	1973.9843	1973.9843	50.31	<.0001	
extra_age12*sex	1	16.5181	16.5181	0.42	0.5166	
extra_age10*sex	1	115.8652	115.8652	2.95	0.0860	
age5*sex	1	45.0751	45.0751	1.15	0.2841	
Source	DF	Type III SS	Mean Square	F Value	Pr > F	
age5	1	22938.10276	22938.10276	584.56	<.0001	
extra_age10	1	1.13371	1.13371	0.03	0.8651	
extra_age12	1	19.37534	19.37534	0.49	0.4824	
extra_age13	1	31.42844	31.42844	0.80	0.3710	
extra_age15	1	1204.03894	1204.03894	30.68	<.0001	
sex	1	35.14931	35.14931	0.90	0.3441	
extra_age15*sex	1	180.44654	180.44654	4.60	0.0322	
extra_age13*sex	1	34.93412	34.93412	0.89	0.3456	
extra_age12*sex	1	65.96394	65.96394	1.68	0.1951	
extra_age10*sex	1	140.18557	140.18557	3.57	0.0590	
age5*sex	1	45.07508	45.07508	1.15	0.2841	
Parameter	Estimate	Standard Error	t Value	Pr > t		
Intercept	112.4674646 B	1.20383527	93.42	<.0001		
age5	5.8573756 B	0.34533748	16.96	<.0001		
extra_age10	-1.0928339 B	0.94077610	-1.16	0.2456		
extra_age12	2.8737997 B	2.16080044	1.33	0.1838		
extra_age13	0.0728337 B	2.28418553	0.03	0.9746		
extra_age15	-5.3782734 B	1.03728680	-5.18	<.0001		
sex	female	1.5294679 B	1.61601834	0.95	0.3441	
sex	male	0.0000000 B	.	.	.	

extra_age15*sex female	3.0019987	B	1.39991144	2.14	0.0322
extra_age15*sex male	0.0000000	B	.	.	.
extra_age13*sex female	-2.8283978	B	2.99764307	-0.94	0.3456
extra_age13*sex male	0.0000000	B	.	.	.
extra_age12*sex female	-3.7274509	B	2.87490153	-1.30	0.1951
extra_age12*sex male	0.0000000	B	.	.	.
extra_age10*sex female	2.4016451	B	1.27063702	1.89	0.0590
extra_age10*sex male	0.0000000	B	.	.	.
age5*sex female	-0.4972619	B	0.46396076	-1.07	0.2841
age5*sex male	0.0000000	B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

We see some rather small signs of the genders differ already from the age of 10, although this does not reach significance. From the age 13 and onwards, they clearly differ, since the boys grow more quickly than the girls.

A scatter plot of the predictions looks like



If we reduce the model to include only changes in the slope from the age 13 and onwards, we get

```
The GLM Procedure

Class Level Information

Class      Levels      Values
sex          2    female male

Number of Observations Read      1142
```

Number of Observations Used 1130

Dependent Variable: height

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	384843.4137	64140.5690	1632.89	<.0001
Error	1123	44111.8273	39.2803		
Corrected Total	1129	428955.2411			

R-Square	Coeff Var	Root MSE	height Mean
0.897164	4.121269	6.267403	152.0746

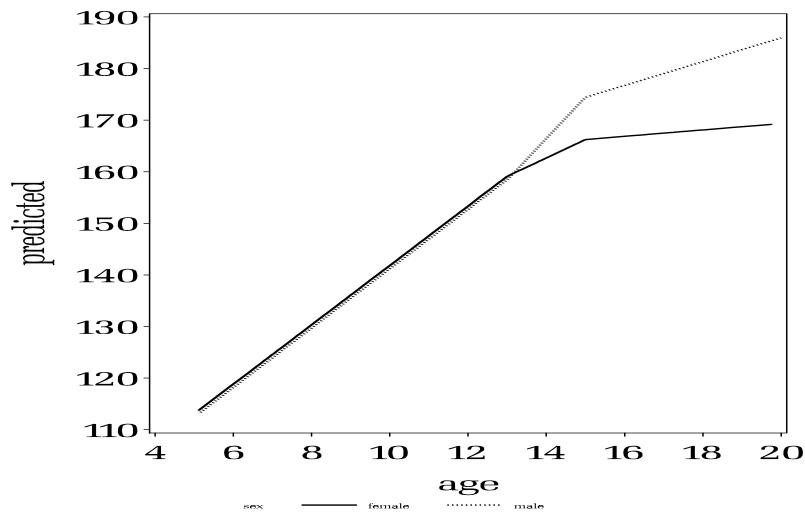
Source	DF	Type I SS	Mean Square	F Value	Pr > F
age5	1	360150.8657	360150.8657	9168.73	<.0001
extra_age13	1	10188.6587	10188.6587	259.38	<.0001
extra_age15	1	2093.4652	2093.4652	53.30	<.0001
sex	1	3514.6923	3514.6923	89.48	<.0001
extra_age15*sex	1	6923.1946	6923.1946	176.25	<.0001
extra_age13*sex	1	1972.5372	1972.5372	50.22	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
age5	1	108563.1404	108563.1404	2763.80	<.0001
extra_age13	1	0.1119	0.1119	0.00	0.9574
extra_age15	1	2234.0449	2234.0449	56.87	<.0001
sex	1	72.4196	72.4196	1.84	0.1748
extra_age15*sex	1	269.7065	269.7065	6.87	0.0089
extra_age13*sex	1	1972.5372	1972.5372	50.22	<.0001

Parameter	Estimate	Standard			
		Error	t Value	Pr > t	
Intercept	112.4445342 B	0.62833302	178.96	<.0001	
age5	5.7517869	0.10940804	52.57	<.0001	
extra_age13	2.2311796 B	0.55472003	4.02	<.0001	
extra_age15	-5.6783749 B	0.78831730	-7.20	<.0001	
sex	female	0.6486417 B	0.47771016	1.36	0.1748
sex	male	0.0000000 B	.	.	.
extra_age15*sex	female	2.7299440 B	1.04182720	2.62	0.0089
extra_age15*sex	male	0.0000000 B	.	.	.
extra_age13*sex	female	-4.4145862 B	0.62296701	-7.09	<.0001
extra_age13*sex	male	0.0000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

with a new plot of predicted values:



From the output as well as from the figures, we note that between the ages 13 and 15, boys grow an extra 4.4 cm (s.e.=0.62cm) a year on average, compared to the girls. Using the estimation friendly version of this model gives us

Parameter		Estimate	Standard Error	t Value	Pr > t
age5		5.7517869	0.10940804	52.57	<.0001
sex	female	113.0931759	0.61030017	185.31	<.0001
sex	male	112.4445342	0.62833302	178.96	<.0001
extra_age15*sex	female	-2.9484310	0.75863194	-3.89	0.0001
extra_age15*sex	male	-5.6783749	0.78831730	-7.20	<.0001
extra_age13*sex	female	-2.1834066	0.53562336	-4.08	<.0001
extra_age13*sex	male	2.2311796	0.55472003	4.02	<.0001