1. Open the WHAS500 data set in the software program of your choice

Obs	ID	AGE	GENDER	HR	SYSBP	DIASBP	BMI	CVD	AFB	SHO	CHF	AV3	MIORD	MITYPE	YEAR	LOS	DSTAT	LENFOL	FSTAT	time_yrs
1	1	83	0	89	152	78	25.5405	1	1	0	0	0	1	0	1	5	0	2178	0	5.96304
2	2	49	0	84	120	60	24.0240	1	0	0	0	0	0	1	1	5	0	2172	0	5.94661
3	4	70	0	65	123	76	26.6319	1	0	0	1	0	0	1	1	10	0	297	1	0.81314
4	5	70	0	63	135	85	24.4125	1	0	0	0	0	0	1	1	6	0	2131	0	5.83436
5	6	70	0	76	83	54	23.2424	1	0	0	0	1	0	0	1	1	1	1	1	0.00274

It is always a good idea to peek at the first few rows of a dataset to orient yourself at the start.

a. Calculate a Cox regression model for systolic blood pressure (sysbp) by itself

The PHREG Procedure

Model Information	
Data Set	WORK.TIME_RECODE
Dependent Variable	time_yrs
Censoring Variable	FSTAT
Censoring Value(s)	0
Ties Handling	BRESLOW

Number of Observations Read	500
Number of Observations Used	500

Summary of the Number of Event and Censored Values			
Total	**Event**	**Censored**	**Percent Censored**
500	215	285	57.00

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Without Covariates**	**With Covariates**
-2 LOG L	2455.158	2450.998
AIC	2455.158	2452.998
SBC	2455.158	2456.368

The p-value is less than 0.05 and the hazard ratio is less than 1. There is evidence of a statistically significant decline in mortality as sysbp increases.

a. Calculate a Cox regression model for systolic blood pressure (sysbp) by itself

The PHREG Procedure

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	4.1606	1	0.0414
Score	4.0922	1	0.0431
Wald	4.0902	1	0.0431

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter Estimate**	**Standard Error**	**Chi-Square**	**Pr > ChiSq**	**Hazard Ratio**
SYSBP	1	-0.00450	0.00223	4.0902	0.0431	0.996

The p-value is less than 0.05 and the hazard ratio is less than 1. There is evidence of a statistically significant decline in mortality as sysbp increases.

and then adjusted for gender and age.

The PHREG Procedure

Model Information	
Data Set	WORK.TIME_RECODE
Dependent Variable	time_yrs
Censoring Variable	FSTAT
Censoring Value(s)	0
Ties Handling	BRESLOW

Number of Observations Read	500
Number of Observations Used	500

Summary of the Number of Event and Censored Values			
Total	**Event**	**Censored**	**Percent Censored**
500	215	285	57.00

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Without Covariates**	**With Covariates**
-2 LOG L	2455.158	2309.238
AIC	2455.158	2315.238
SBC	2455.158	2325.350

The inclusion of gender and age does not appear to have much effect on the hazard ratio for sysbp.

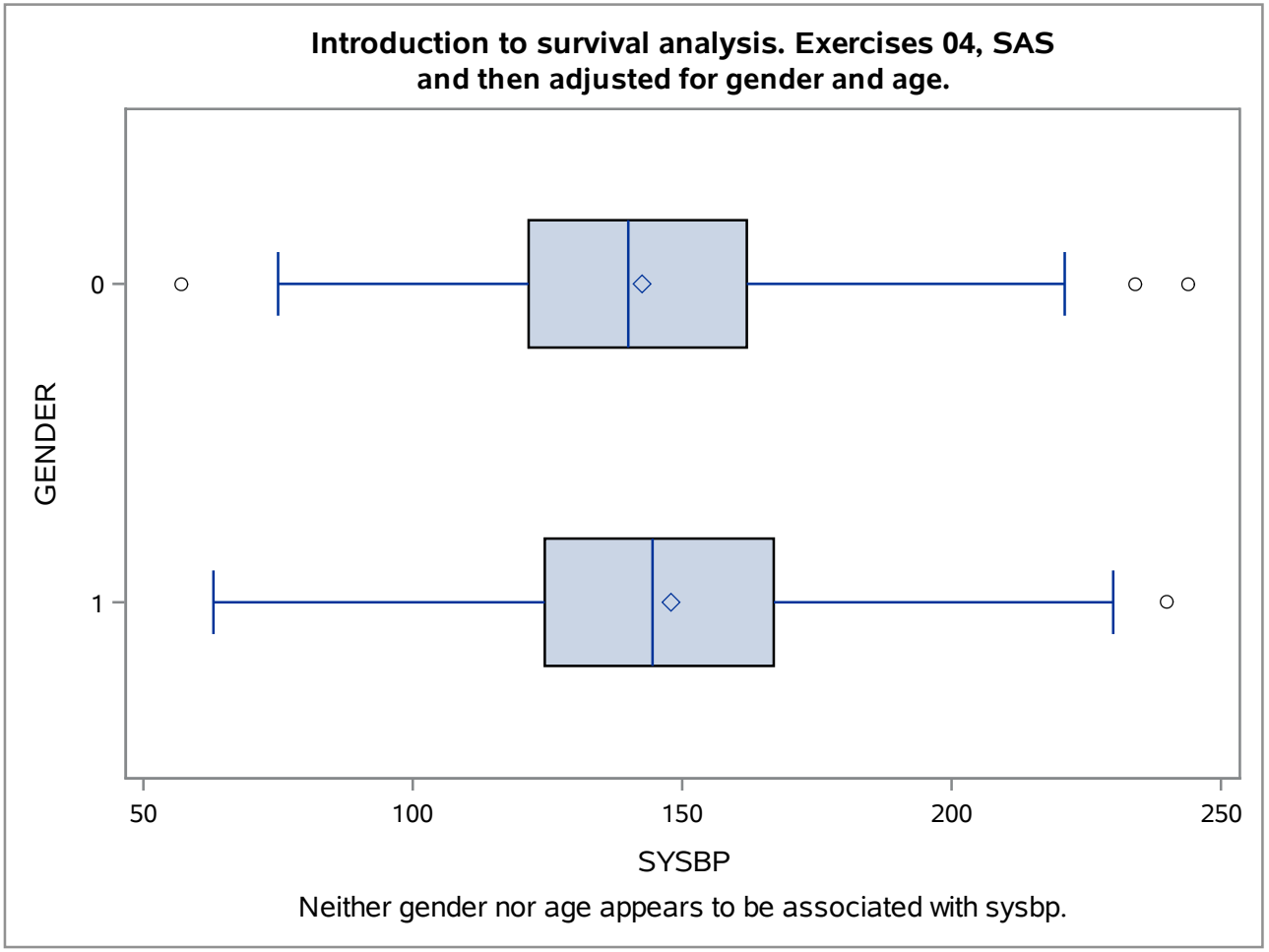
and then adjusted for gender and age.

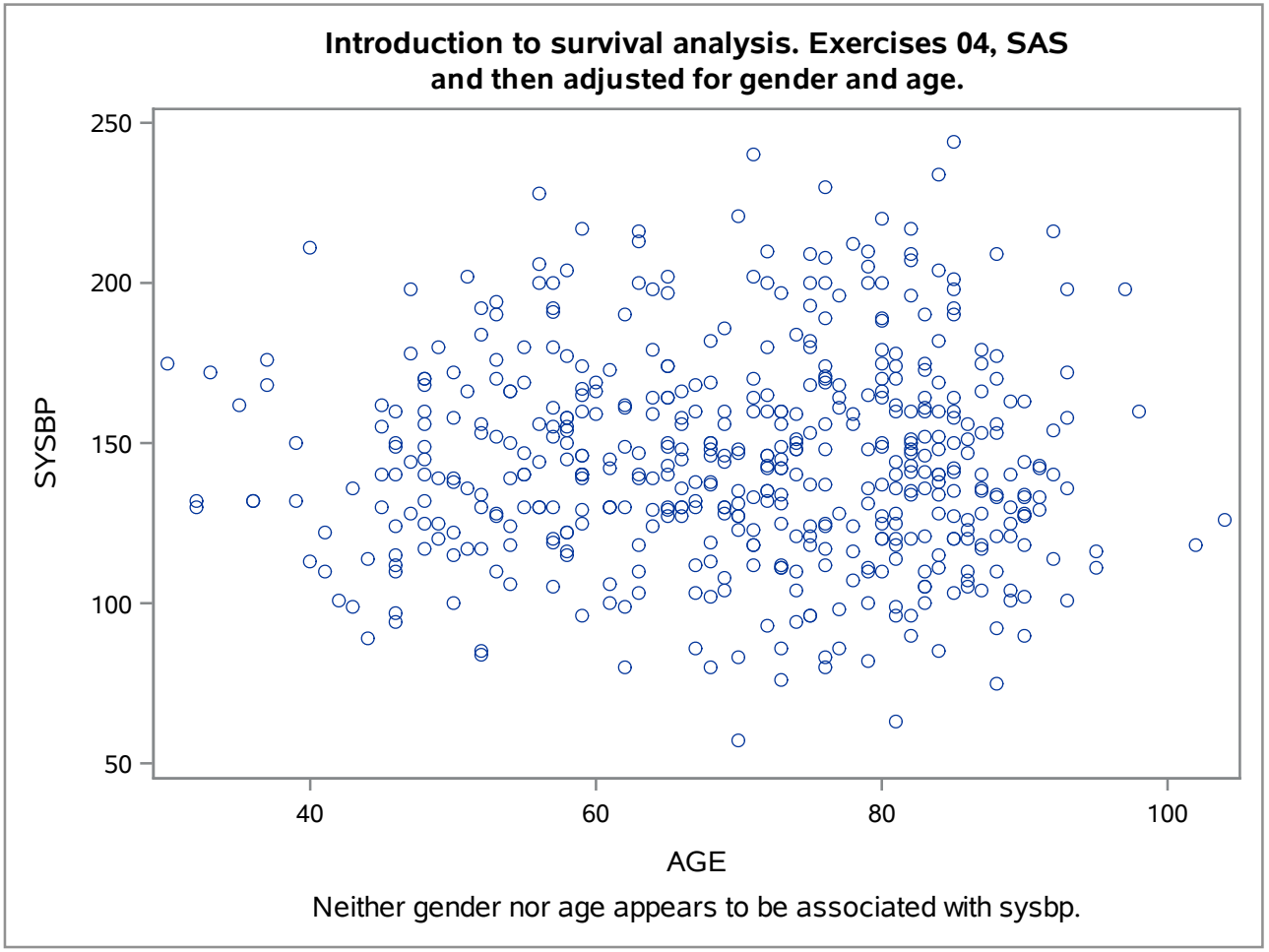
The PHREG Procedure

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	145.9202	3	<.0001
Score	131.4801	3	<.0001
Wald	124.1651	3	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter Estimate**	**Standard Error**	**Chi-Square**	**Pr > ChiSq**	**Hazard Ratio**
SYSBP	1	-0.00426	0.00218	3.8241	0.0505	0.996
GENDER	1	-0.05337	0.14080	0.1437	0.7047	0.948
AGE	1	0.06646	0.00618	115.8405	<.0001	1.069

The inclusion of gender and age does not appear to have much effect on the hazard ratio for sysbp.


Introduction to survival analysis. Exercises 04, SAS
and then adjusted for gender and age.
GENDER
0
1
50
100
150
200
250
SYSBP
Neither gender nor age appears to be associated with sysbp.


Introduction to survival analysis. Exercises 04, SAS
and then adjusted for gender and age.
SYSBP
250
200
150
100
50
40
60
80
100
AGE
Neither gender nor age appears to be associated with sysbp.

Calculate the unadjusted survival curves for patients with systolic blood pressures of 120, 140, and 160.

The PHREG Procedure

Model Information	
Data Set	WORK.TIME_RECODE
Dependent Variable	time_yrs
Censoring Variable	FSTAT
Censoring Value(s)	0
Ties Handling	BRESLOW

Number of Observations Read	500
Number of Observations Used	500

Summary of the Number of Event and Censored Values			
Total	**Event**	**Censored**	**Percent Censored**
500	215	285	57.00

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Without Covariates**	**With Covariates**
-2 LOG L	2455.158	2450.998
AIC	2455.158	2452.998
SBC	2455.158	2456.368

The unadjusted comparison shows a small decrease in risk of death as sysbp increases.

Calculate the unadjusted survival curves for patients with systolic blood pressures of 120, 140, and 160.

The PHREG Procedure

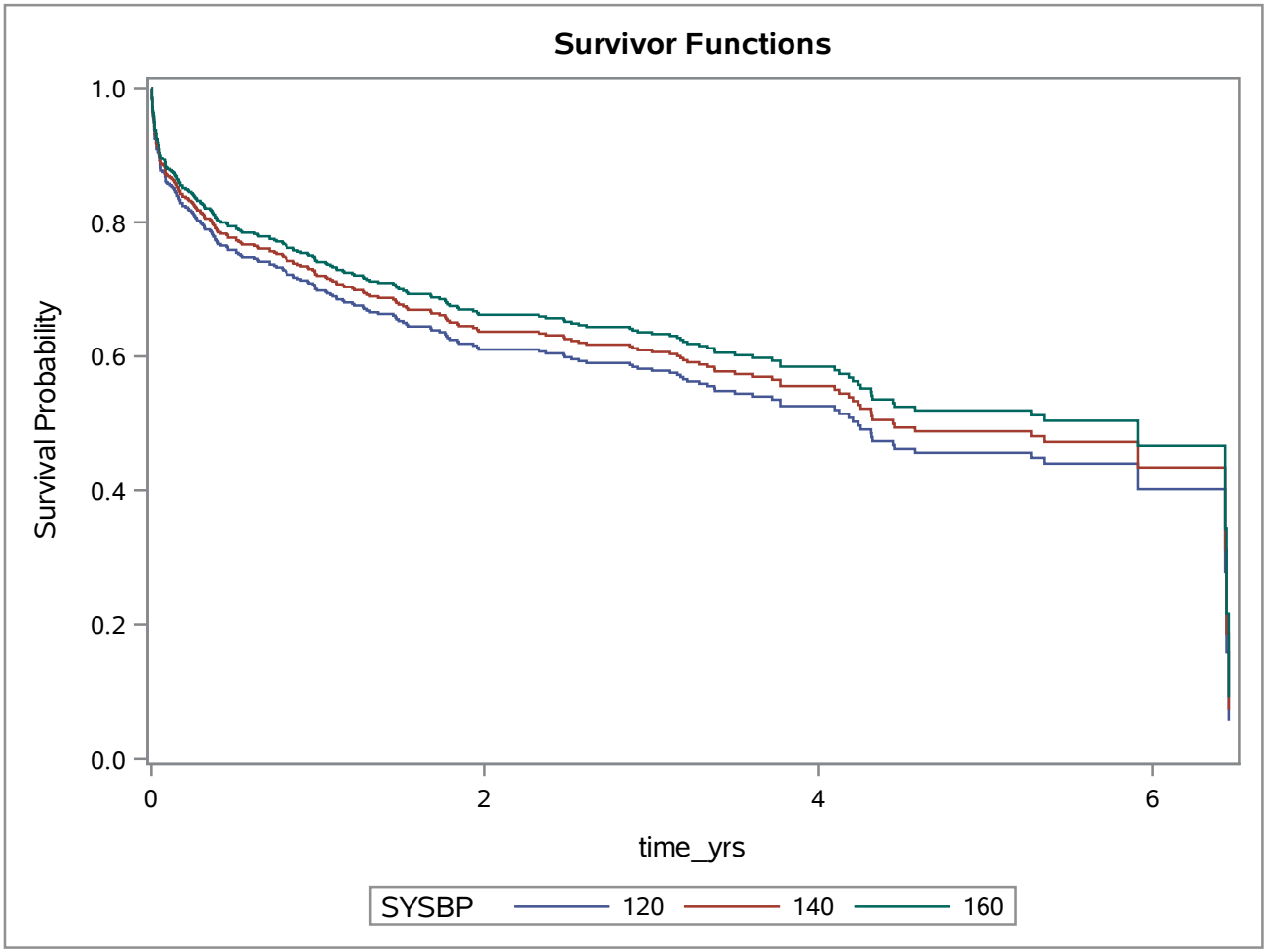
Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	4.1606	1	0.0414
Score	4.0922	1	0.0431
Wald	4.0902	1	0.0431

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter Estimate**	**Standard Error**	**Chi-Square**	**Pr > ChiSq**	**Hazard Ratio**
SYSBP	1	-0.00450	0.00223	4.0902	0.0431	0.996

The unadjusted comparison shows a small decrease in risk of death as sysbp increases.

Calculate the unadjusted survival curves for patients with systolic blood pressures of 120, 140, and 160.

The PHREG Procedure



The unadjusted comparison shows a small decrease in risk of death as sysbp increases.

Then recalculate these survival curves with age set to the overall average age, and to a population that is 30% female. Interpret your results.

The PHREG Procedure

Model Information	
Data Set	WORK.TIME_RECODE
Dependent Variable	time_yrs
Censoring Variable	FSTAT
Censoring Value(s)	0
Ties Handling	BRESLOW

Number of Observations Read	500
Number of Observations Used	500

Summary of the Number of Event and Censored Values			
Total	**Event**	**Censored**	**Percent Censored**
500	215	285	57.00

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Without Covariates**	**With Covariates**
-2 LOG L	2455.158	2309.238
AIC	2455.158	2315.238
SBC	2455.158	2325.350

The results are largely unchanged after adjustment.

Then recalculate these survival curves with age set to the overall average age, and to a population that is 30% female. Interpret your results.

The PHREG Procedure

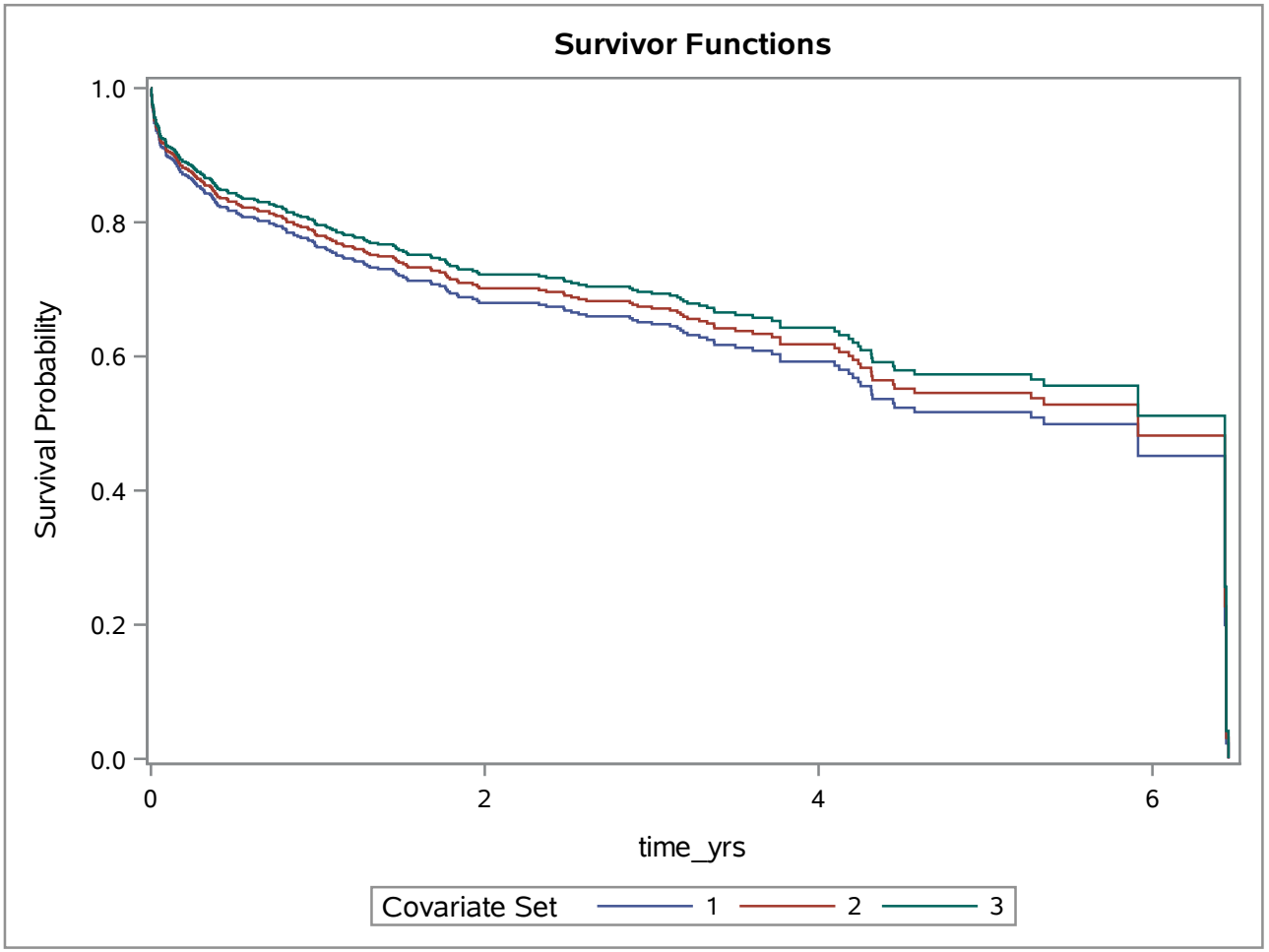
Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	145.9202	3	<.0001
Score	131.4801	3	<.0001
Wald	124.1651	3	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter Estimate**	**Standard Error**	**Chi-Square**	**Pr > ChiSq**	**Hazard Ratio**
SYSBP	1	-0.00426	0.00218	3.8241	0.0505	0.996
AGE	1	0.06646	0.00618	115.8405	<.0001	1.069
GENDER	1	-0.05337	0.14080	0.1437	0.7047	0.948

The results are largely unchanged after adjustment.

Then recalculate these survival curves with age set to the overall average age, and to a population that is 30% female. Interpret your results.

The PHREG Procedure



The results are largely unchanged after adjustment.

b. Calculate cubic spline model for systolic blood pressure with four degrees of freedom.

The PHREG Procedure

Model Information	
Data Set	WORK.SYSBP_RECODE
Dependent Variable	time_yrs
Censoring Variable	FSTAT
Censoring Value(s)	0
Ties Handling	BRESLOW

Number of Observations Read	500
Number of Observations Used	500

The spline is statistically significant.

b. Calculate cubic spline model for systolic blood pressure with four degrees of freedom.

The PHREG Procedure

Knots for Spline Effect sysbp_spline5	
Knot Number	sysbp_c
1	-56.53333
2	-25.36667
3	5.80000
4	36.96667
5	68.13333

The spline is statistically significant.

b. Calculate cubic spline model for systolic blood pressure with four degrees of freedom.

The PHREG Procedure

Basis Details for Spline Effect sysbp_spline5		
Column	Power	Break Knot
1	0	
2	1	
3	3	-56.53333
4	3	-25.36667
5	3	5.80000

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
500	215	285	57.00

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	2455.158	2442.424
AIC	2455.158	2450.424
SBC	2455.158	2463.907

The spline is statistically significant.

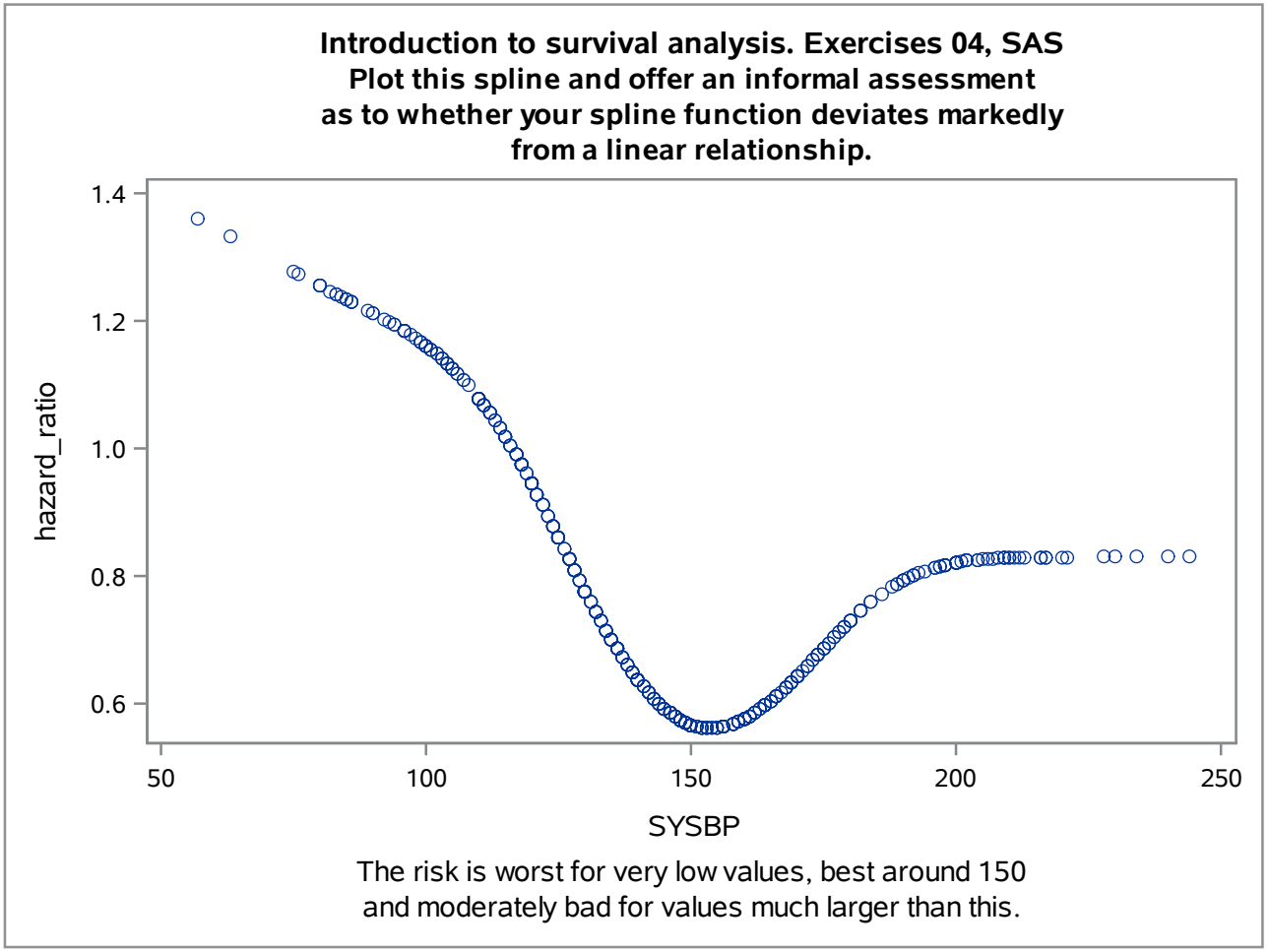
b. Calculate cubic spline model for systolic blood pressure with four degrees of freedom.

The PHREG Procedure

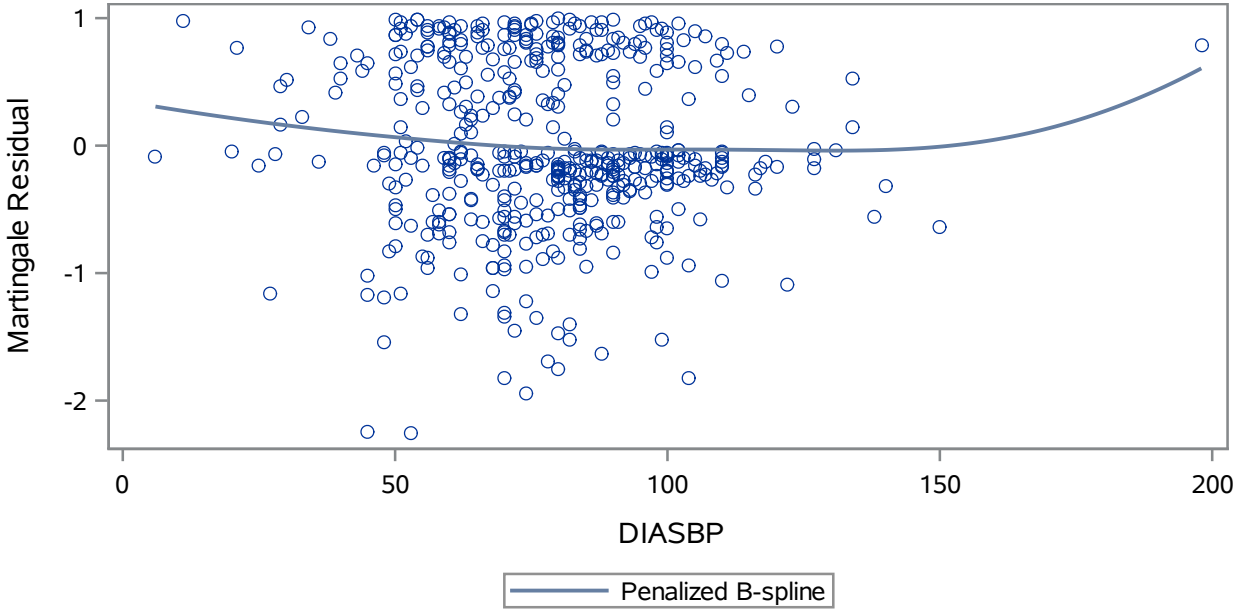
Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	12.7340	4	0.0127
Score	13.3401	4	0.0097
Wald	13.0454	4	0.0111

Analysis of Maximum Likelihood Estimates								
Parameter		**DF**	**Parameter Estimate**	**Standard Error**	**Chi-Square**	**Pr > ChiSq**	**Hazard Ratio**	**Label**
sysbp_spline5	**1**	0	0	.	.	.	.	sysbp_spline5 1
sysbp_spline5	**2**	1	-0.00351	0.01080	0.1056	0.7452	.	sysbp_spline5 2
sysbp_spline5	**3**	1	-0.0005558	0.0004879	1.2978	0.2546	.	sysbp_spline5 3
sysbp_spline5	**4**	1	0.00164	0.00114	2.0538	0.1518	.	sysbp_spline5 4
sysbp_spline5	**5**	1	-0.00156	0.00107	2.1553	0.1421	.	sysbp_spline5 5

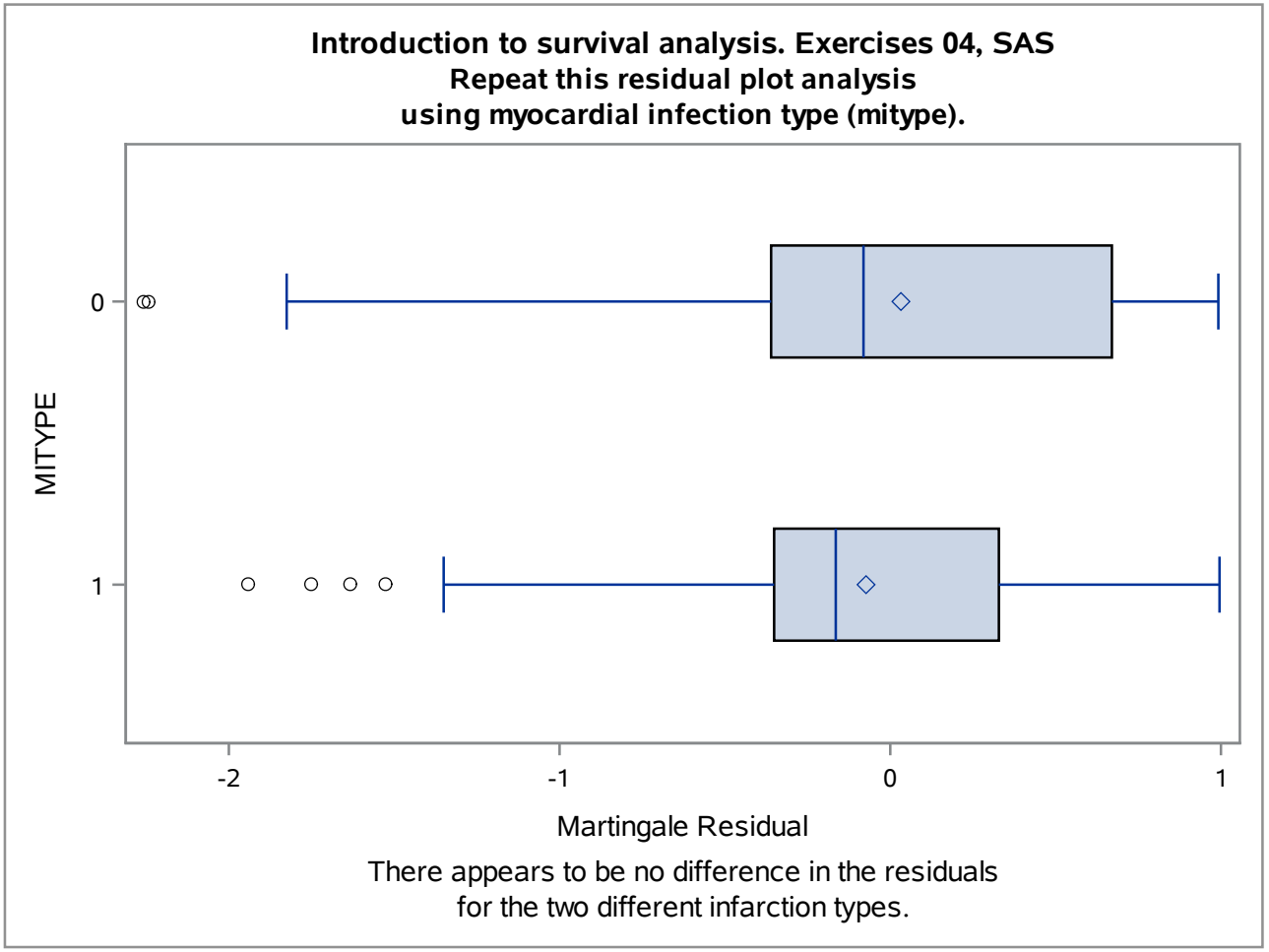
The spline is statistically significant.


Introduction to survival analysis. Exercises 04, SAS
Plot this spline and offer an informal assessment
as to whether your spline function deviates markedly
from a linear relationship.
hazard_ratio
1.4
1.2
1.0
0.8
0.6
50
100
150
200
250
SYSBP
The risk is worst for very low values, best around 150
and moderately bad for values much larger than this.

Introduction to survival analysis. Exercises 04, SAS
c. Calculate the Martingale residuals from your Cox model with a linear term for systolic blood pressure and for age and a term for gender. Plot these residuals versus diastolic blood pressure.



There may be an effect for diastolic blood pressure similar to what you saw for the spline model for systolic blood pressure.


Introduction to survival analysis. Exercises 04, SAS
Repeat this residual plot analysis
using myocardial infection type (mitype).
MITYPE
0
1
-2
-1
0
1
Martingale Residual
There appears to be no difference in the residuals
for the two different infarction types.