СТАТИСТИЧЕСКИ СОФТУЕР

Упражнение 3: Анализ на данни със SAS. Статистически изводи.

3.1. Използване на IF и ELSE IF условие

```
data conditional;
 length Gender $ 1
 Quiz $ 2;
 input Age Gender Midterm Quiz FinalExam;
 if Age lt 20 and not missing(age) then AgeGroup = 1;
 else if Age ge 20 and Age lt 40 then AgeGroup = 2;
 else if Age ge 40 and Age lt 60 then AgeGroup = 3;
 else if Age ge 60 then AgeGroup = 4;
 datalines;
 21 M 80 B- 82
 . F 90 A 93
 35 M 87 B+ 85
48 F . . 76
 59 F 95 A+ 97
 15 M 88 . 93
 67 F 97 A 91
 . M 62 F 67
 35 F 77 C- 77
 49 M 59 C 81
;
title "Listing of CONDITIONAL";
proc print data=conditional noobs;
run;
```

Syntax

IF *expression* **THEN** *statement*;
<**ELSE** *statement*;>

Listing of CONDITIONAL

Gender	Quiz	Age	Midterm	FinalExam	AgeGroup
M	B-	21	80	82	2
F	A	.	90	93	.
M	B+	35	87	85	2
F		48	.	76	3
F	A+	59	95	97	3
M		15	88	93	1
F	A	67	97	91	4
M	F	.	62	67	.
F	C-	35	77	77	2
M	C	49	59	81	3

Логически оператори:

Logical Comparison	Mnemonic	Symbol
Equal to	EQ	=
Not equal to	NE	^= or ~= or ¬= *
Less than	LT	<
Less than or equal to	LE	<=
Greater than	GT	>
Greater than or equal to	GE	>=
Equal to one in a list	IN	

СТАТИСТИЧЕСКИ СОФТУЕР

Упражнение 3: Анализ на данни със SAS. Статистически изводи.

3.2. Използване на специфична форма на IF условие

```
data females;
 length Gender $ 1
 Quiz $ 2;
 input Age Gender Midterm Quiz FinalExam;
 if Gender eq 'F';
 datalines;
 21 M 80 B- 82
 . F 90 A 93
 35 M 87 B+ 85
 48 F . . 76
 59 F 95 A+ 97
 15 M 88 . 93
 67 F 97 A 91
 . M 62 F 67
 35 F 77 C- 77
 49 M 59 C 81
 ;
 title "Listing of FEMALES";
 proc print data=Females noobs;
 run;
```

Listing of FEMALES

Gender	Quiz	Age	Midterm	FinalExam
F	A	.	90	93
F		48	.	76
F	A+	59	95	97
F	A	67	97	91
F	C-	35	77	77

СТАТИСТИЧЕСКИ СОФТУЕР

Упражнение 3: Анализ на данни със SAS. Статистически изводи.

3.3.Използване на IF и OR условие

- **data females;**
- **length Gender $ 1**
- **Quiz $ 2;**
- **input Age Gender Midterm Quiz FinalExam;**
- **if Quiz = 'A+' or Quiz = 'A' or Quiz = 'B+' or Quiz = 'B'**
- **then QuizRange = 1;**
- **else if Quiz = 'B-' or Quiz = 'C+' or Quiz = 'C'**
- **then QuizRange = 2;**
- **else if not missing(Quiz) then QuizRange = 3;**
- **datalines;**

21 M 80 B- 82
. F 90 A 93
35 M 87 B+ 85
48 F . . 76
59 F 95 A+ 97
15 M 88 . 93
67 F 97 A 91
. M 62 F 67
35 F 77 C- 77
49 M 59 C 81
;
title "Listing of FEMALES";
proc print data=Females noobs;
run;

Listing of FEMALES

Gender	Quiz	Age	Midterm	FinalExam	QuizRange
M	B-	21	80	82	2
F	A	.	90	93	1
M	B+	35	87	85	1
F		48	.	76	.
F	A+	59	95	97	1
M		15	88	93	.
F	A	67	97	91	1
M	F	.	62	67	3
F	C-	35	77	77	3
M	C	49	59	81	2

СТАТИСТИЧЕСКИ СОФТУЕР

Упражнение 3: Анализ на данни със SAS. Статистически изводи.

3.4. Използване на WHEN условието

```
data conditional;
 length Gender $ 1
 Quiz $ 2;
 input Age Gender Midterm Quiz FinalExam;
 select;
 when (missing(Age)) AgeGroup = .;
 when (Age lt 20) AgeGroup = 1;
 when (Age lt 40) AgeGroup = 2;
 when (Age lt 60) AgeGroup = 3;
 when (Age ge 60) Agegroup = 4;
 otherwise;
 end;
 datalines;
```

```
21 M 80 B- 82
 . F 90 A 93
 35 M 87 B+ 85
 48 F . . 76
 59 F 95 A+ 97
 15 M 88 . 93
 67 F 97 A 91
 . M 62 F 67
 35 F 77 C- 77
 49 M 59 C 81
 ;
 title "Listing of FEMALES";
 proc print data= conditional
noobs;
 run;
```

Syntax

SELECT <(*select-expression*)>;
  WHEN-1 (*when-expression-1* <..., *when-expression-n*>) *statement*;
  <... **WHEN**-*n* (*when-expression-1* <..., *when-expression-n*>) *statement*;>
    <**OTHERWISE** *statement*;>

END;

```
data believe_it_or_not;
 input X;
 if X = 3 or X = 4 then Match = 'Yes';
 else Match = 'No';
 datalines;
 3
 7
 .
 ;
 title "Listing of BELIEVE_IT_OR_NOT";
 proc print data=believe_it_or_not noobs;
 run;
```

Listing of BELIEVE_IT_OR_NOT

X	Match
3	Yes
7	No
.	No

- **data** females;
- set conditional;
- where Gender eq 'F';
- **run**;

	Gender	Quiz	Age	Midterm	FinalExam	AgeGroup
1	F	A	.	90	93	.
2	F		48	.	76	3
3	F	A+	59	95	97	3
4	F	A	67	97	91	4
5	F	C-	35	77	77	2

Използване на оператора : WHERE

Operator	Description	Example
IS MISSING	Matches a missing value	where Subj is missing
IS NULL	Equivalent to IS MISSING	where Subj is null
BETWEEN AND	An inclusive range	where age between 20 and 40
CONTAINS	Matches a substring	where Name contains Mac
LIKE	Matching with wildcards	where Name like R_n%
=*	Phonetic matching	where Name =* Nick

- **Използване на оператора : WHERE**

Expression	Matches
where Gender is null	A missing character value
where Age is null	A missing numeric value
where Age is missing	A missing numeric value
where Age between 20 and 40	All values between 20 and 40, including 20 and 40
where Name contains mac	macon immaculate
where Name like R_n%	Ron Ronald Run Running
where Name =* Nick	Nick Nack Nikki

```
data revenue;
 retain Total 0;
 input Day : $3.
 Revenue : dollar6.;
 if not missing(Revenue) then Total = Total + Revenue;
 format Revenue Total dollar8.;
 datalines;
 Mon $1,000
 Tue $1,500
 Wed .
 Thu $2,000
 Fri $3,000
 ;
```

	Total	Day	Revenue
1	$1,000	Mon	$1,000
2	$2,500	Tue	$1,500
3	$2,500	Wed	.
4	$4,500	Thu	$2,000
5	$7,500	Fri	$3,000

```
data revenue;
input Day : $3.
Revenue : dollar6.;
Total + Revenue;
format Revenue Total dollar8.;
datalines;
Mon $1,000
Tue $1,500
Wed .
Thu $2,000
Fri $3,000
;
```

	Day	Revenue	Total
1	Mon	$1,000	$1,000
2	Tue	$1,500	$2,500
3	Wed	.	$2,500
4	Thu	$2,000	$4,500
5	Fri	$3,000	$7,500

```
data test;
input x;
if missing(x) then MissCounter + 1;
datalines;
2
.
7
.
;
```

	x	MissCounter
1	2	0
2	.	1
3	7	1
4	.	2

```
data compound;
 Interest = .0375;
 Total = 100;
 Year + 1;
 Total + Interest*Total;
 output;
 Year + 1;
 Total + Interest*Total;
 output;
 Year + 1;
 Total + Interest*Total;
 output;
 format Total dollar10.2;
 run;
 title "Listing of COMPOUND";
 proc print data=compound noobs;
 run;
```

```
data compound;
Interest = .0375;
Total = 100;
do Year = 1 to 3;
Total + Interest*Total;
output;
end;
format Total dollar10.2;
run;
```

(iterative DO loop)

	Interest	Total	Year
1	0.0375	$103.75	1
2	0.0375	$107.64	2
3	0.0375	$111.68	3

Listing of COMPOUND

Interest	Total	Year
0.0375	$103.75	1
0.0375	$107.64	2
0.0375	$111.68	3

```
data table;
 do n = 1 to 10;
 Square = n*n;
 SquareRoot = sqrt(n);
 output;
 end;
 run;
 title "Table of Squares and Square Roots";
 proc print data=table noobs;
 run;
```

Table of Squares and Square Roots

n	Square	SquareRoot
1	1	1.00000
2	4	1.41421
3	9	1.73205
4	16	2.00000
5	25	2.23607
6	36	2.44949
7	49	2.64575
8	64	2.82843
9	81	3.00000
10	100	3.16228

```
data equation;
 do X = -10 to 10 by .01;
Y = X-8;
 output;
 end;
 run;
 symbol value=none interpol=sm width=2;
 title "Plot ofY versus X";
 proc gplot data=equation;
 plotY * X;
 run;
```