

# Manajemen Data Relasional (STK261)

#### PENYIAPAN DATA DI SAS

Dr. Agus Mohamad Soleh

Departemen Statistika Fakultas Matematika dan IPA Ganjil 2021



#### Pernyataan pada Tahapan Data

- ■Pernyataan DATA
  untuk menyatakan nama gugus data SAS yang akan dibuat
- ■Pernyataan INPUT untuk menyatakan nama-nama peubah yang ada dalam data
- ■Pernyataan INFILE
  untuk menyatakan nama gugus data sebagai data input eksternal yang akan dibuat gugus data SAS
- Pernyataan CARDS (DATALINES) untuk menyatakan data input *in-stream* yang akan dibuat gugus data SAS
- ■Pernyataan LIBNAME
  untuk menyatakan lokasi (folder) gugus data SAS sebagai input atau
  output yang disimpan secara permanen



#### Pernyataan DATA

(nama gugus data SAS)

#### Bentuk umum:

### DATA namagugusdataSAS;

bisa menyisipkan lebih dari 1 gugus data

- Hasil dari tahapan DATA:
   Sementara ('WORK.namagugusdataSAS') atau permanen
- Setiap PROC menggunakan gugusdata terbaru
- Contoh:

DATA; nama dari gugus data -> sementara selesai membuat data ini maka proc yang terakhir akan langsung mengeksekuis data ini. proc nya jg bersifat sementara

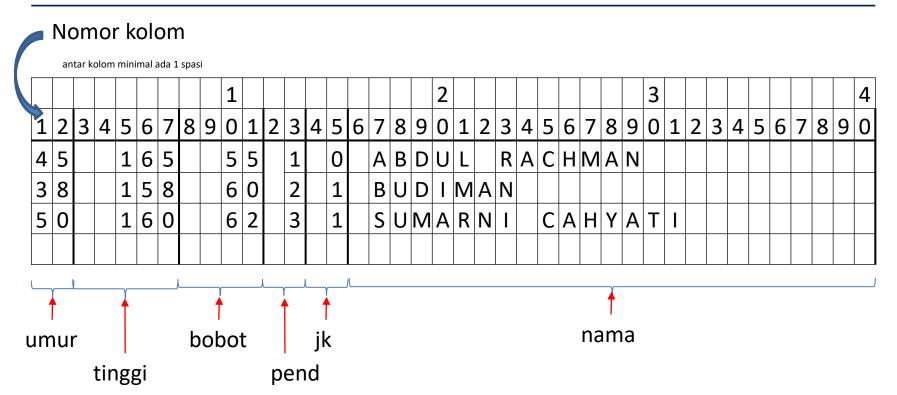
DATA statlab;

PROC PRINT;

PROC PRINT DATA=statlab;



#### Contoh DATA





(nama-nama peubah)

#### Bentuk umum:

untuk menyiapkan varibel maksimum 32 karakter tipe/format : karakter & numerik

karakter & numerik punya banyak tipe lagi

### INPUT namapeubah tipe lokasi;

contoh untuk tanggal (numerik) bisa pake

■INPUT /ist (tanpa posisi kolom) scanning: jika dipisahkan oleh minimal 1 spasi, maka input akan langsung memisahkannya berdasarkan vatiabel yang ada (tiap ada spasi)

Contoh:

untuk pendidikan dan jenis kelamin bisa diberikan simbol \$, karena karakter. karena mreka tidak bisa

INPUT umur tinggi bobot pend\$ jk\$ nama\$;

rachman dan cahyani tidak akan dimasukkan ke data 3 3 4 5 6 3 6 8 9 A|B|D|U4 5 6 0 H M A N 3 8 6 B|U|D|I | M | A | NAHYA 6



(nama-nama peubah)

#### Bentuk umum:

#### INPUT namapeubah tipe lokasi;

#### ■INPUT kolom

Contoh:

INPUT nama \$ 17-40 umur 1-2 tinggi 4-7 bobot 9-11 pend 13 jk 15;

									1										2										3										4
1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0
4	5			1	6	5			5	5		1		0		Α	В	D	U	L		R	Α	С	Н	M	Α	N											
3	8			1	5	8			6	0		2		1		В	U	D	I	M	Α	N																	
5	0			1	6	0			6	2		3		1		S	U	M	Α	R	N			C	Α	Н	Υ	Α	Т	I									
																																			_				



(nama-nama peubah)

#### Bentuk umum:

#### INPUT namapeubah tipe lokasi;

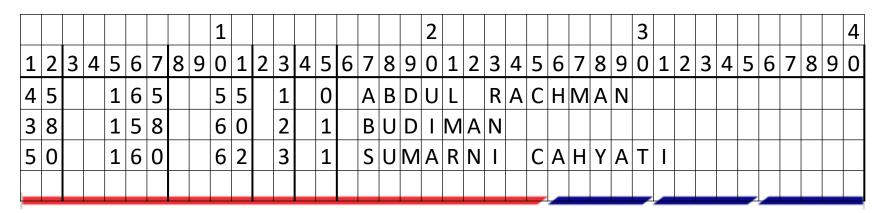
■INPUT campuran (list dan kolom)

Contoh:

selama ada lokasinya, bisa ditempatkan di mana saja (variabelnya)

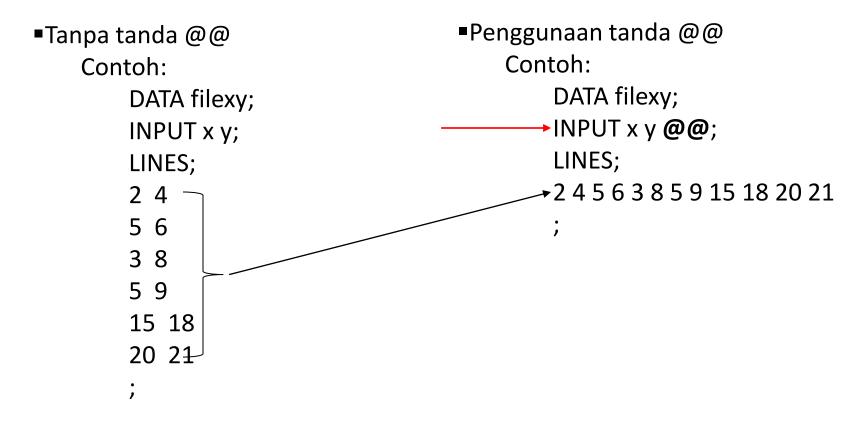
karena tinggi & bobot tidak ada angkanya, maka ia akan ditempatkan sesuai urutan penamaan variabel (kolom 3 & 4)

INPUT nama \$ 17-40 umur 1-2 tinggi bobot pend 13 jk 15;





(nama-nama peubah)





#### Pernyataan INFILE

(nama gugus data sebagai input eksternal)

#### Bentuk umum:

INFILE namafile\_dan\_lokasinya;

umumnya jika sudah menggunakan infil, maka tidak ada lagi pernyataan cards, dataline, lines dll.

```
/* data eksternal */
/* 'stkdat1.txt' */

10 4 12

9 9 15

INFILE "D:\stk\stkdat1.txt";

RUN;

/* data SAS sementara*/
/* data eksternal

*/
DATA stk371;
INFILE "D:\stk\stkdat1.txt";
RUN;
```



### Pernyataan DATALINE (CARDS)

(data input in-stream)

#### Bentuk umum:

#### DATALINES;

```
Contoh: /* data in-stream */
DATA stk261;
INPUT x y z;
DATALINES;
10 4 12
9 9 15
13 8 14
;
RUN;
```



#### Contoh Program

```
/* data SAS sementara*/
/* data SAS sementara*/
                                       /* data eksternal
/* data in-stream
                                       DATA sta261;
DATA sta261;
                                       INFILE "D:\stk\stkdat1.txt";
INPUT x y z;
                                       INPUT x y z;
DATALINES;
                                       RUN;
10 4 12
 9 9 15
13 8 14
                                       /* data eksternal */
                                       /* 'stkdat1.txt'
RUN;
                                       10 4 12
                                        9 9 15
          nama gugus data SAS pada
                                       13 8 14
          Explorer: WORK.sta261
```



#### Pernyataan LIBNAME

(lokasi (folder) gugus data SAS)

#### Bentuk umum:

#### LIBNAME lokasi;

setiap nama var & lib, tidak boleh menggunakan spasi

```
Contoh:
                      /* data SAS permanen*/
                        LIBNAME nlib "D:\stk\";
                                                          nama library pada Explorer
menyimpan data secara permanen di folder stk
                        DATA nlib.sta261
                                                            nama gugus data SAS pada Explorer:
                        INPUT x y z;
                                                            nlib.sta261
                        DATALINES;
                        10 4 12
                           9 15
                        13 8 14
                                            nama gugus data SAS pada folder D:\stk\:
                                            sta261.sas7bdat
                        RUN;
```



### **Contoh Program**

```
/* data SAS permanen*/
/* data SAS permanen*/
                                     /* data eksternal
/* data in-stream
                                      LIBNAME nlib "D:\stk\";
LIBNAME nlib "D:\stk\";
                                      DATA nlib.sta261;
DATA nlib.sta261;
                                      INFILE "D:\stk\stkdat1.txt";
INPUT x y z;
                                      INPUT x y z;
DATALINES;
                                      RUN;
10 4 12
 9 9 15
13 8 14
                                      /* data eksternal */
                                      /* 'stkdat1.txt'
RUN;
                                      10 4 12
           sta261.sas7bdat
                                       9 9 15
                                      13 8 14
```



#### PROC CONTENTS DATA=gugusdataSAS;

```
DATA skor;
INPUT x y z nama $;
LINES;
23 46 5 Budi
32 75 2 Norma
40 55 8 Andre
;
PROC CONTENTS DATA=skor;
RUN;
```



#### Output:

		banyaknya data	
Data Set Name	WORK.SKOR	Observations	3
Member Type	DATA	Variables	4

posisi file/otak file	
Filename	C:\Users\HP\AppData\Local\Temp\SAS Temporary Files\_TD7688_DESKTOP-E80JVJ4_\skor.sas7bdat

Alphabetic List of Variables and Attributes										
#	Variable	Туре	Len							
4	nama	Char	8							
1	x	Num	8							
2	У	Num	8							
3	z	Num	8							

defaultnya 8

#### PROC CONTENTS DATA=gugusdataSAS;

lib name hanya perlu sekali saja. tapi jika lokasinya berbeda, maka lib namenya juga harus berbeda

```
LIBNAME nlib "D:\education\statcomp\";
PROC CONTENTS DATA=nlib.sta261;
RUN;
```

```
PROC CONTENTS

DATA="D:\education\statcomp\sta261";
RUN;
```



#### Output:

	The CONTENTS F	Procedure						
Data Set Name	NLIB.STA261	Observations	3					
Member Type	DATA	Variables	3					
Fngine	v9 (versi sas 9)	Indeves	n					
Filename D:\education\statcomp\sta261.sas7bdat								

Alphabetic List of Variables and Attributes										
#	Variable	Туре	Len							
1	x	Num	8							
2	у	Num	8							
3	z	Num	8							



### Periksa Sintaks Progam SAS

dibaca secara baris, sampai bertemu;

```
1. data bobot;
                                     tim harusnya karakter, tapi tidak error. hanya akan warning dan nilainya adalah missing
   input nomor 1-4 nama $ 6-24 tim bobotsblm bobotssdh;
                                        tidak ada;
3. turun = bobotsblm - bobotssdh
                                           maka akan error
4. lines:
5. 1023 Dodi Suradi
                                merah
                                           189 165
6. 1049 Amelia Serrano
                                kuning
                                           145 124
7. 1219 Alan Nance
                                merah
                                           210 192
8. 1246 Ravi Sinha
                                           194 177
                                kuning
9. 1078 Asley Makatano
                                merah
                                            127 118
10. :;
11. proc print nobs;
12. var nama bobotblm bobotssdh turun;
13. proc sort data boobt by nomor;
                       typo
14. run;
```



### Periksa Sintaks Progam SAS

```
data bobot;
input nomor 1-4 nama $ 6-24 tim bobotsblm bobotssdh;
turun = bobotsblm - bobotssdh
lines:
1023 Dodi Suradi merah 189 165
1049 Amelia Serrano
                       kuning 145 124
                       merah 210 192
1219 Alan Nance
1246 Ravi Sinha
                       kuning 194 177
1078 Asley Makatano merah 127 118
<u>:;</u>
proc print nobs;
var nama bobotblm bobotssdh turun;
proc sort data boobt by nomor;
run;
```



proc means

#### Penggunaan Paket SAS

```
data bobot;
input nomor 1-4 nama $ 6-24 tim $ bobotsblm bobotssdh;
turun = bobotsblm - bobotssdh;
lines;
1023 Dodi Suradi
                         merah
                                  189 165
1049 Amelia Serrano
                         kuning
                                  145 124
                         merah
                                  210 192
1219 Alan Nance
1246 Ravi Sinha
                                  194 177
                         kuning
                         merah
                                  127 118
1078 Asley Makatano
proc print noobs;
var nama bobotsblm bobotssdh turun;
proc sort data=bobot; by nomor;
run;
       proc dalam sas:
```

harus ada s nya, tapi jika tidak, maka sas akan tetap menjalankannya dengan program yang sama



### Pernyataan pada Tahapan Data

■Pernyataan SET

menggunakan gugus data yang sudah ada, dibuat ke gugus data yang baru. mirip merge tapi bukan merge

set tidak hanya sebagai gugus data, tapi juga bisa sebagai variabel

untuk menyatakan nama gugus data SAS sebagai sumber data

Pernyataan DROP

untuk menyatakan nama-nama peubah yang **tidak akan dimasukkan** ke dalam suatu gugus data SAS

misal DATA A; & DATA B; -> DATA AB;

Pernyataan KEEP

untuk menyatakan nama-nama peubah yang **akan dimasukkan** ke dalam suatu gugus data SAS

Pernyataan IF

untuk menyatakan kondisi tertentu dalam suatu program SAS

Pernyataan DO-loop

untuk menyatakan proses iteratif dalam suatu program SAS



#### Pernyataan SET

#### Bentuk umum:

#### SET namagugusdataSAS;

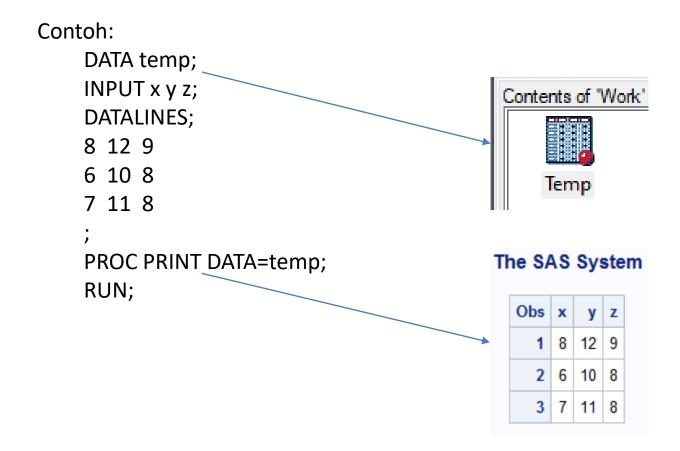
- Tidak ada pernyataan INPUT bersamaan dengan pernyataan SET
- Untuk membuat subset data dari suatu gugusdataSAS

```
DATA temp;
INFILE 'D:/xyzdata'; menyimpan gugus data temp
INPUT x y z;
DATA jumlah; gugus data jumlah ada 4 var:
x, y, z, jum

SET temp;
jum = x+y+z;
PROC PRINT DATA=jumlah;
```

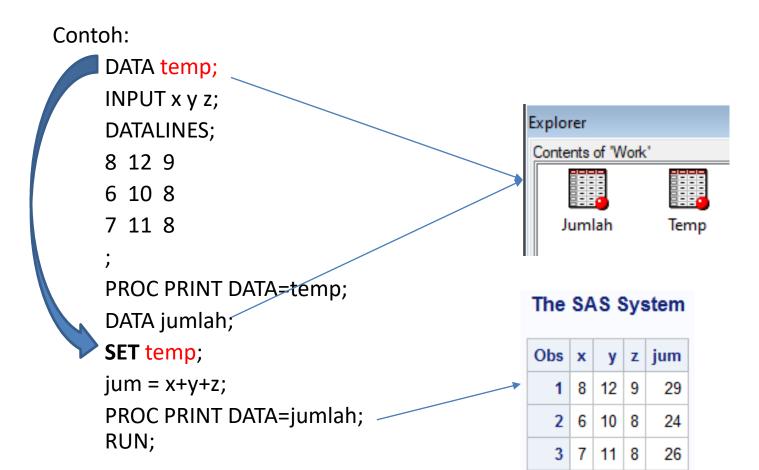


### Pernyataan SET





### Pernyataan SET





#### Pernyataan DROP dan KEEP

#### Bentuk umum:

tidak boleh diletakkan dalam suatu tahapan data (yang sama)

```
DROP peubah ...;
KEEP peubah ...;
```

#### Contoh:

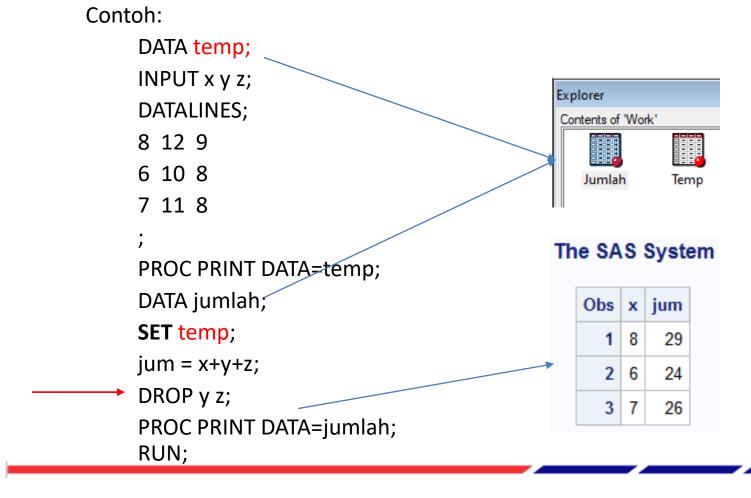
```
DATA temp;
INFILE 'D:/xyzdata';
INPUT x y z;
DATA jumlah;
SET temp;
jum = x+y+z;
DROP x y z;
KEEP jum;
PROC PRINT DATA=jumlah;
```

ini sebenarnya salah, seharusnya salah satunya saja

terlihat bahwa KEEP lebih efisien

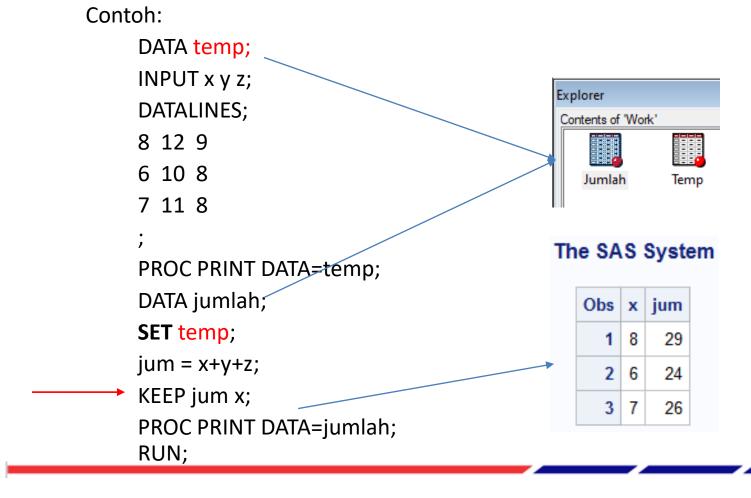


### Pernyataan DROP dan KEEP





### Pernyataan DROP dan KEEP





```
IF expression; jika benar, akan disimpan
IF expression THEN DELETE;
```

jika benar, akan dihapus

```
IF Sex='f';
IF Sex='m' THEN DELETE;

* Choose only comedies;

DATA comedy;

INFILE 'c:\MyRawData\Shakespeare.dat';

INPUT Title $ 1-26 Year Type $;

IF Type='comedy';

PROC PRINT DATA=comedy;

RUN;
```

```
DATA CLASS;
    INPUT jk$ nama...;

LINES;

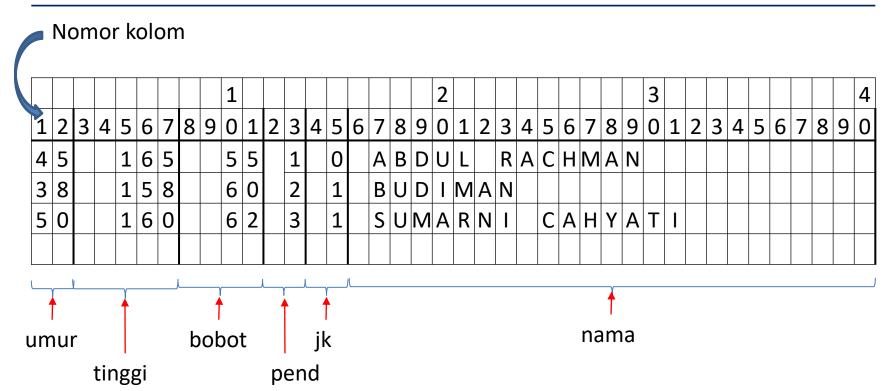
DATA PRIA;
    SET CLASS;
    IF jk = 'M';
    IF jk = 'F' THEN DELETE;

DATA PRIA WANITA;
    SET CLASS;
    IF jk = '1' THEN OUTPUT PRIA;
    IF jk = '2' THEN OUTPUT WANITA;

data akan disimpan pada gugus data WANITA
```



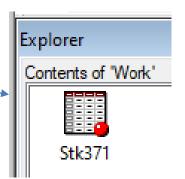
#### Contoh DATA





Contoh: tanpa pernyataan IF

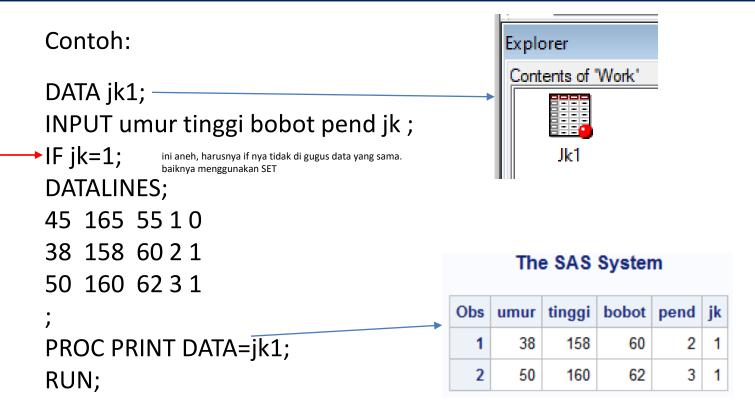
```
DATA stk371;
INPUT umur tinggi bobot pend jk;
DATALINES;
45 165 55 1 0
38 158 60 2 1
50 160 62 3 1
;
PROC PRINT DATA=stk371;
RUN;
```



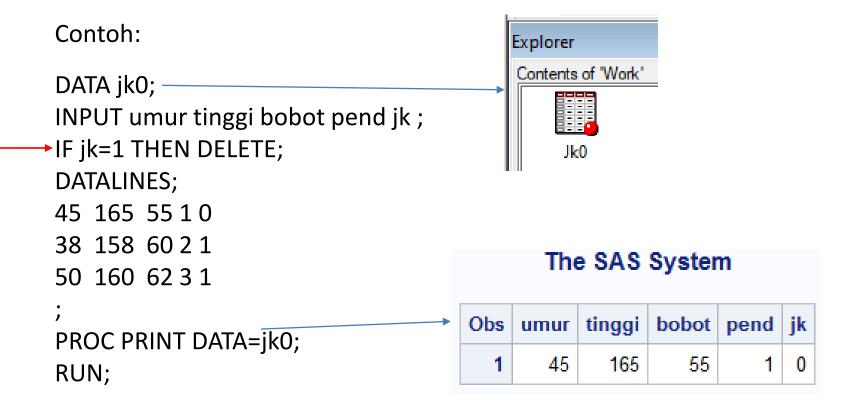
			-		
Obs	umur	tinggi	bobot	pend	jk
1	45	165	55	1	0
2	38	158	60	2	1
3	50	160	62	3	1

The SAS System











#### Bentuk umum:

```
DO var1 = var2 TO var3 [BY pertambahan];
:
:
END;
```

#### Contoh:

```
DATA contoh1;

INPUT pon @@;

kilo = pon/2.2;

LINES;

OUTPUT;

jika ada OUTPUT maka setiap baris (loop) akan disimpan nilainya ke gugus data contoh2

END;
```

dalam sas ada juga While. tapi nanti



```
Contoh:

DATA normal1;

DO z = -4 TO 4.05 BY 0.2;

density = (1.0/SQRT(2*3.1415926536))*EXP(-z**2/2);

OUTPUT;

END;

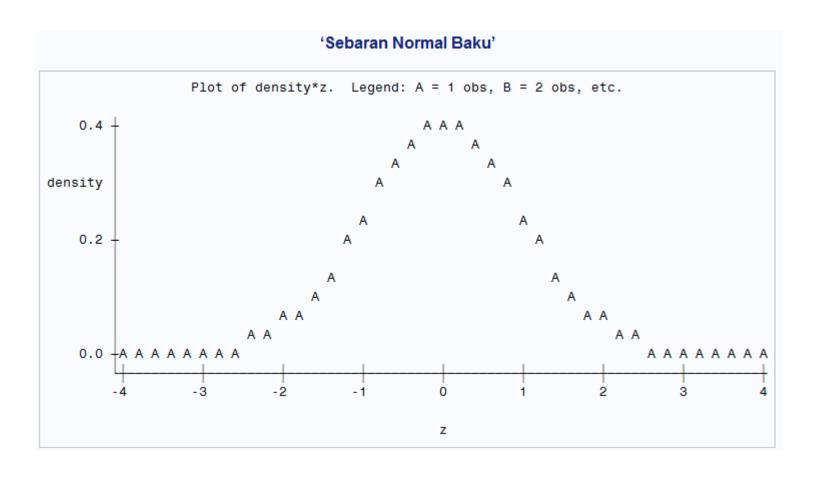
PROC PLOT DATA=normal1;

PLOT density*z / HAXIS=-4 TO 4 BY 1;

TITLE 'Sebaran Normal Baku';

RUN;
```





```
DATA normal2;

DO i = 1 TO 40; kalo tanpa pernyataan by, maka akan naik sebanyak 1 satuan

x = 50 + 10*NORMAL(0);

OUTPUT;

END;

PROC CHART DATA=normal2;

VBAR x / MIDPOINTS=20 TO 80 BY 5;

TITLE 'Bangkitan data normal dengan rataan 50 dan simpangan 10';

RUN;
```

```
DATA seragam;
DO i = 1 TO 100;
y = UNIFORM(0);
OUTPUT;
END;
PROC CHART DATA=seragam;
VBAR y / MIDPOINTS=0.05 TO 0.95 BY 0.1;
TITLE 'Bangkitan data seragam (uniform) dengan selang nilai (0,1)';
RUN;
```



```
DATA normal3;
n = 40;
DO i = 1 \text{ TO } 500;
   DO j = 1 TO n;
         x = 50 + 10*NORMAL(0);
         jum = jum + x;
   END;
   Rataan = jum / n;
   OUTPUT;
END;
PROC PRINT DATA=normal3;
TITLE 'Bangkitan data 500 kali masing2 dengan n=40';
PROC CHART DATA=normal3;
VBAR rataan / MIDPOINTS=40 TO 60 BY 2;
RUN;
```



## Terima Kasih



Inspiring Innovation with Integrity in Agriculture, Ocean and Biosciences for a Sustainable World