

# SAS BASICS

Technology Short Courses: Fall 2008  
(October, 2008)  
Kentaka Aruga

# Object of the course

---

- Sub-windows in SAS
- Basics of managing data files
- Basic commands in SAS

# Introduction: What is SAS?

---

- What is SAS?
  - Originally an acronym for Statistical Analysis System
  - Provided by SAS institute since the 1970s
  - A software used for statistical analysis, graphing, and presenting data

# Introduction: DATA Step

---

- Two distinct categories
  - DATA step
  - PROC step
- DATA Step
  - Provides data management
  - Use
    - Reading data
    - Data transformation
    - Creating or removing variables

# Introduction: PROC Step

---

- PROC Step
  - Performs a wide variety of analysis on data those are retrieved and transformed from the DATA Step
  - Examples
    - PROC MEANS, CONTENTS, SORT, FREQ, PRINT, PLOT etc.

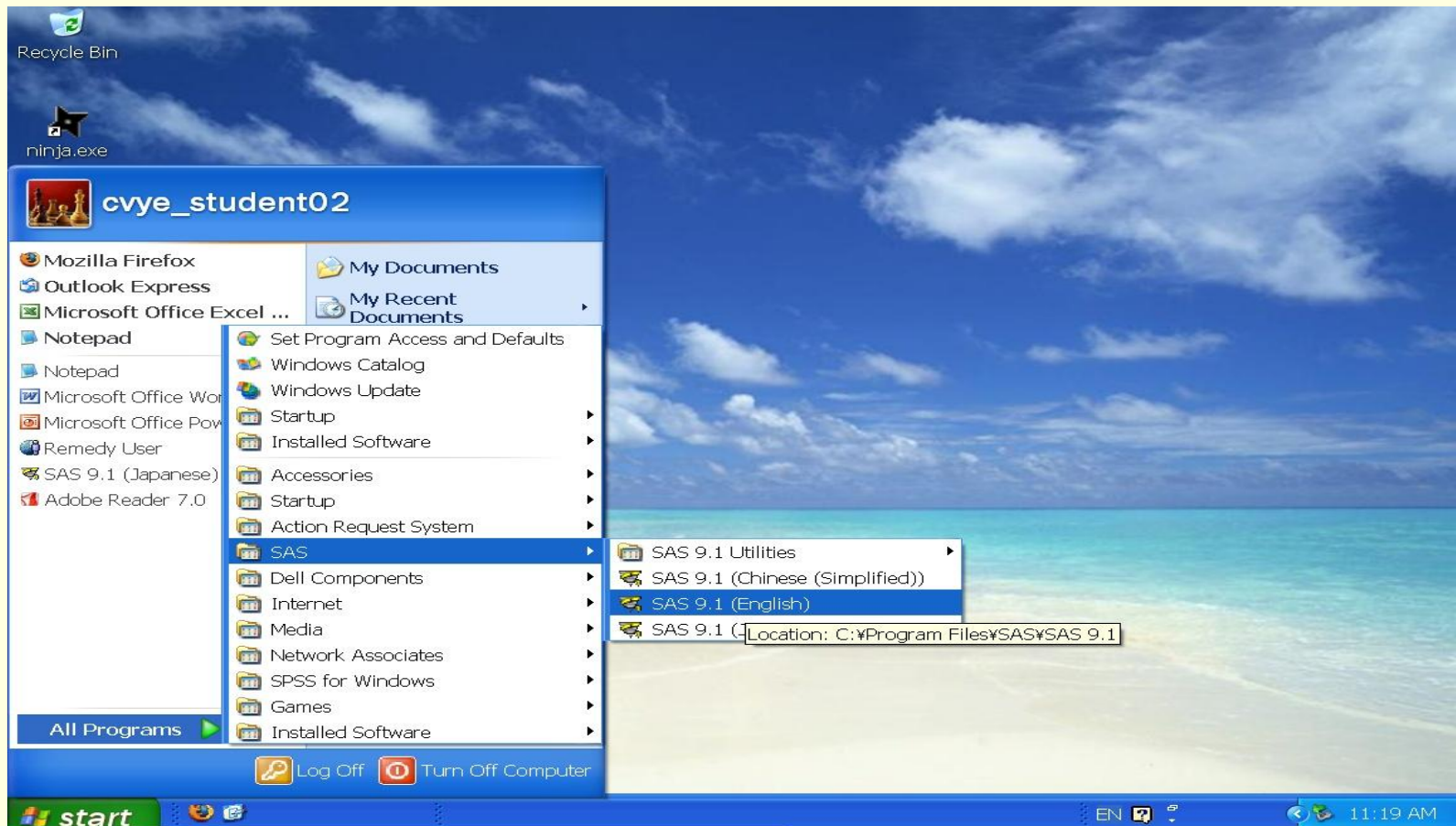


# **Section 1**

## Learning About the Sub-windows

# Opening SAS

- Start → All Programs → SAS → SAS 9.1



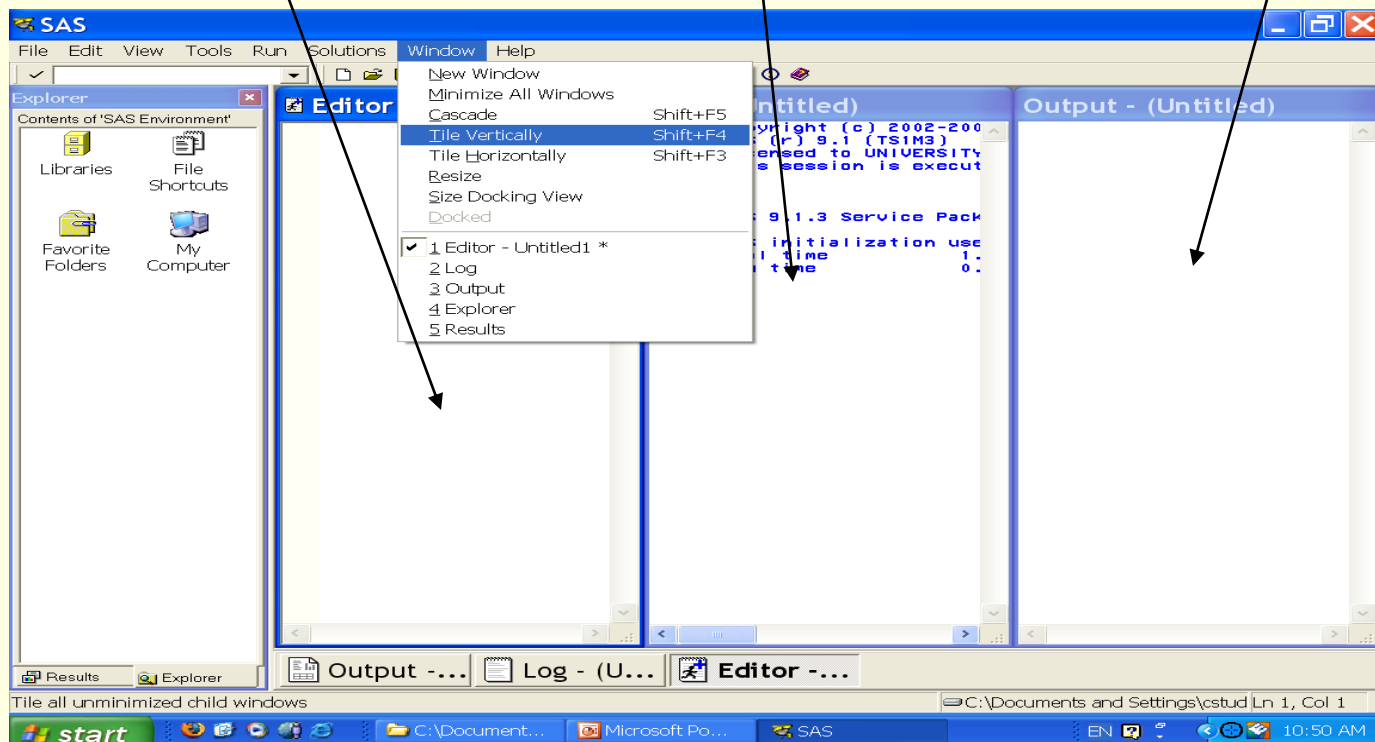
# Three main windows: Program editor

On the top bar click 'Window' and then click 'Tile Vertically.' You will be able to see three sub-windows.

1. Program editor

2. Log window

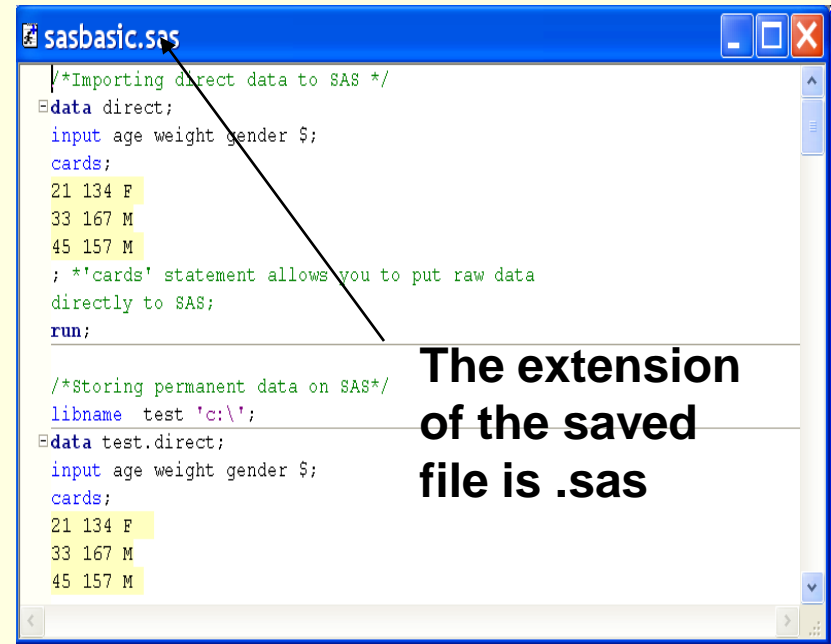
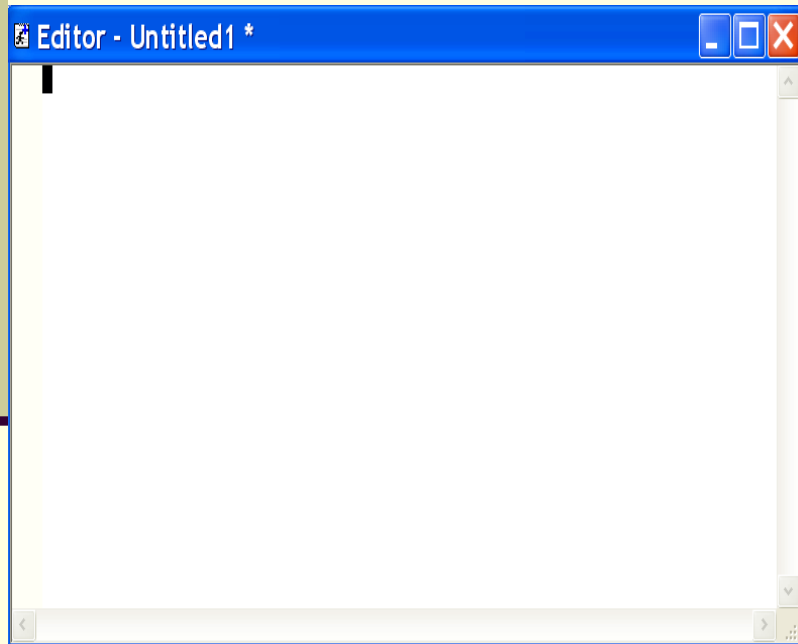
3. Output window





# Three main windows: Program editor

- Program editor
  - Entering and editing SAS command lines

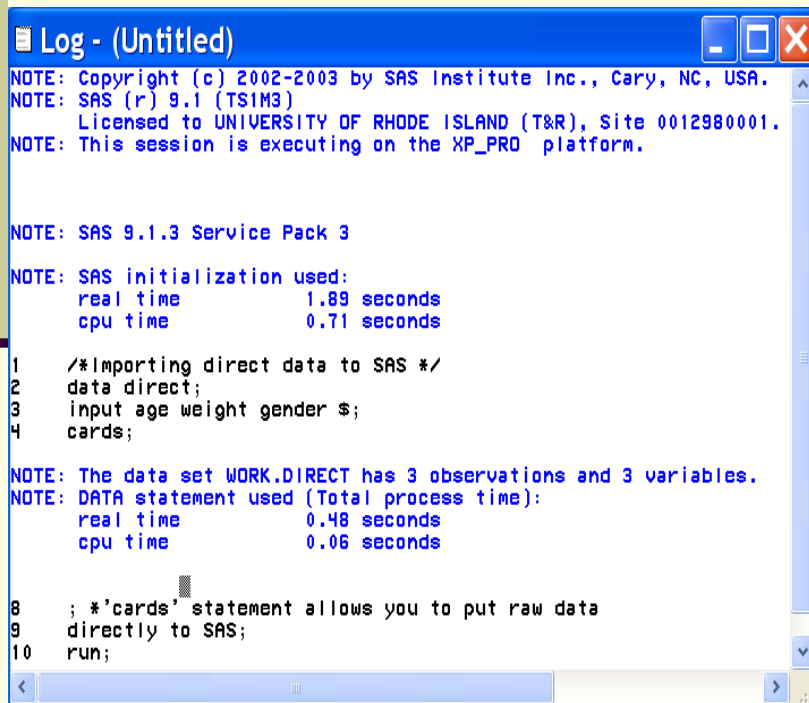


# Three main windows: Log window

- Log window

- This window keeps track of your command runs, and lists SAS notes and error messages (shown in red)

## Commands written correctly



```
Log - (Untitled)
NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) 9.1 (TS1M3)
      Licensed to UNIVERSITY OF RHODE ISLAND (T&R), Site 0012980001.
NOTE: This session is executing on the XP_PRO platform.

NOTE: SAS 9.1.3 Service Pack 3

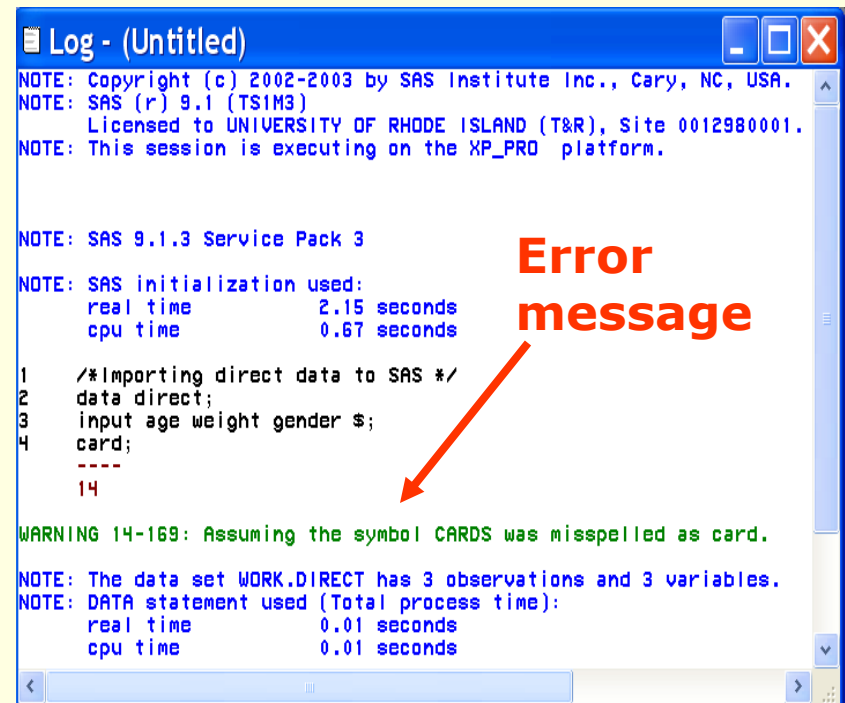
NOTE: SAS initialization used:
      real time      1.89 seconds
      cpu time       0.71 seconds

1  /*Importing direct data to SAS */
2  data direct;
3  input age weight gender $;
4  cards;

NOTE: The data set WORK.DIRECT has 3 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time      0.48 seconds
      cpu time       0.06 seconds

8  ; *'cards' statement allows you to put raw data
9  directly to SAS;
10 run;
```

## Commands with error



```
Log - (Untitled)
NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) 9.1 (TS1M3)
      Licensed to UNIVERSITY OF RHODE ISLAND (T&R), Site 0012980001.
NOTE: This session is executing on the XP_PRO platform.

NOTE: SAS 9.1.3 Service Pack 3

NOTE: SAS initialization used:
      real time      2.15 seconds
      cpu time       0.67 seconds

1  /*Importing direct data to SAS */
2  data direct;
3  input age weight gender $;
4  card;
   ----
   14

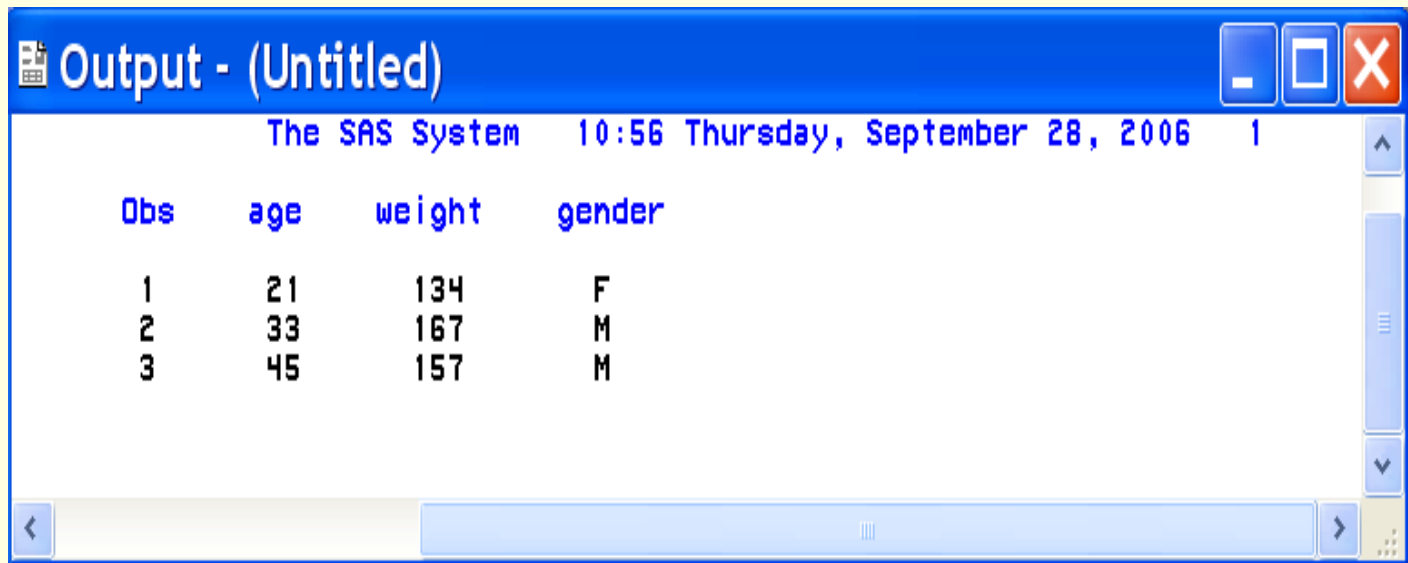
WARNING 14-169: Assuming the symbol CARDS was misspelled as card.

NOTE: The data set WORK.DIRECT has 3 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time      0.01 seconds
      cpu time       0.01 seconds
```

**Error message**

# Three main windows: Output window

- Output window
  - Shows the results of SAS procedures
  - The extension of the saved file is “.lst”



The screenshot shows the SAS Output window titled "Output - (Untitled)". The window displays the following data:

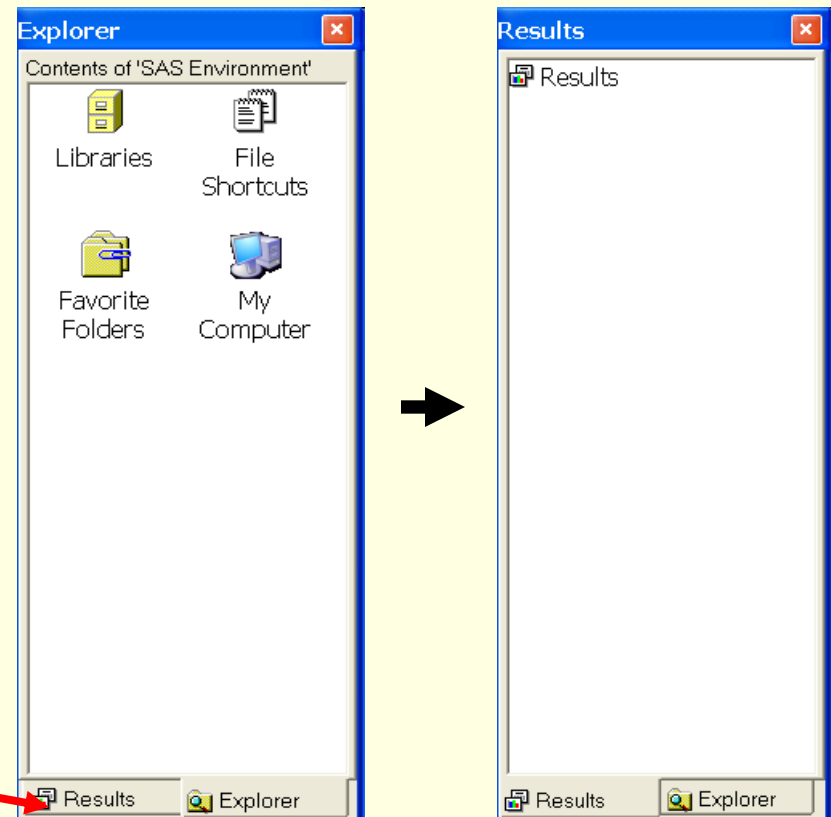
Obs	age	weight	gender
1	21	134	F
2	33	167	M
3	45	157	M

The window also shows the SAS System status bar at the top: "The SAS System 10:56 Thursday, September 28, 2006 1".

# 'Explorer' and 'Results' window

The 'Explorer' and 'Results' *Windows* will appear on the left side of your screen.

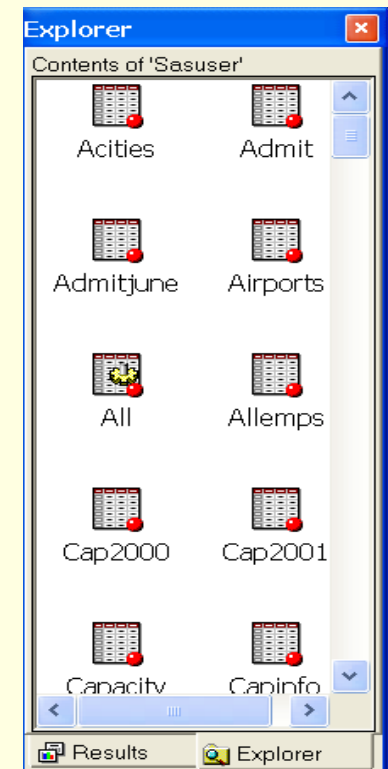
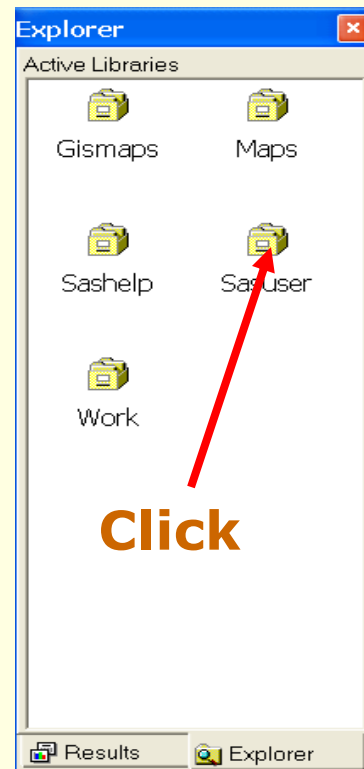
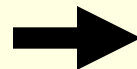
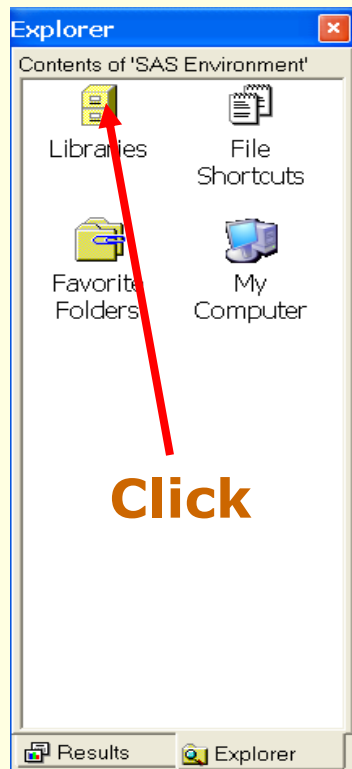
- Explorer window
  - This window is used to explore various default libraries that contain a number of sample SAS data sets
- Results window
  - Organizes the information contained in the Output Window in a hierarchical fashion.



**Click**

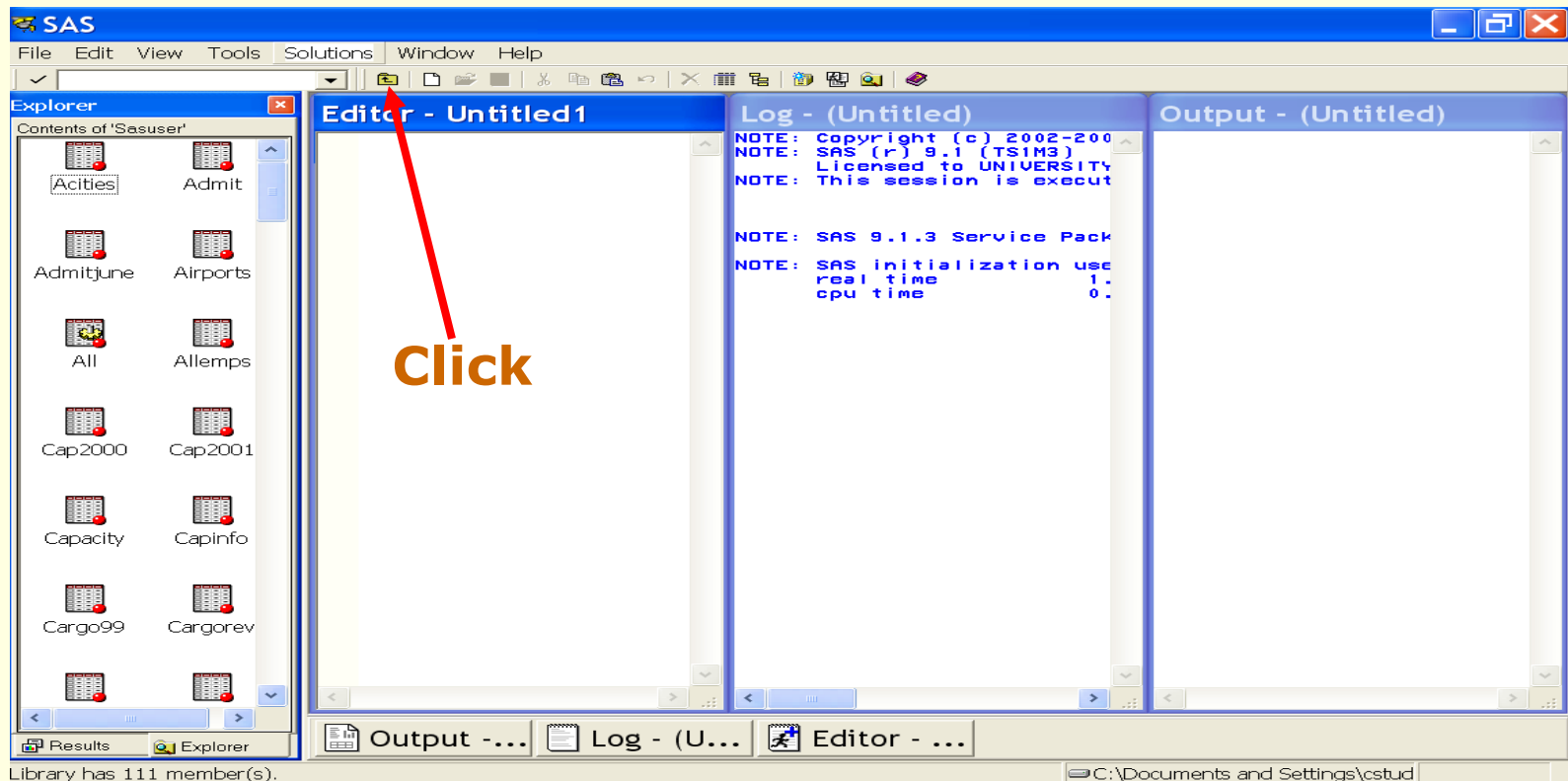
# 'Explorer' window

- Click 'Libraries' icon in the Explorer window. Then you will see several subfolders. You can find the raw SAS data in these subfolders.



# 'Explorer' window (Cont'd)

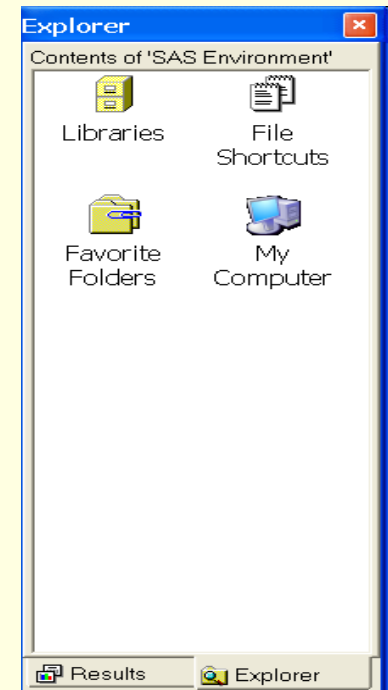
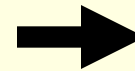
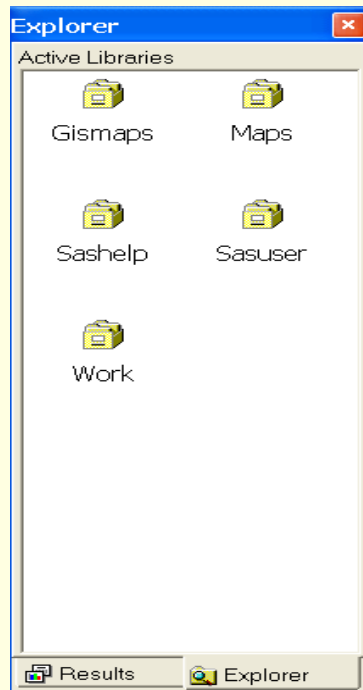
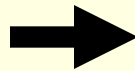
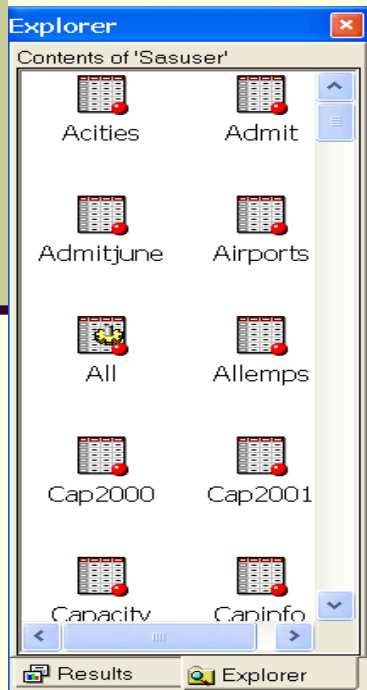
- To move backward from one folder to another in the Explorer Window, simply click the left most icon on the toolbar that looks like a folder.



# 'Explorer' window (Cont'd)

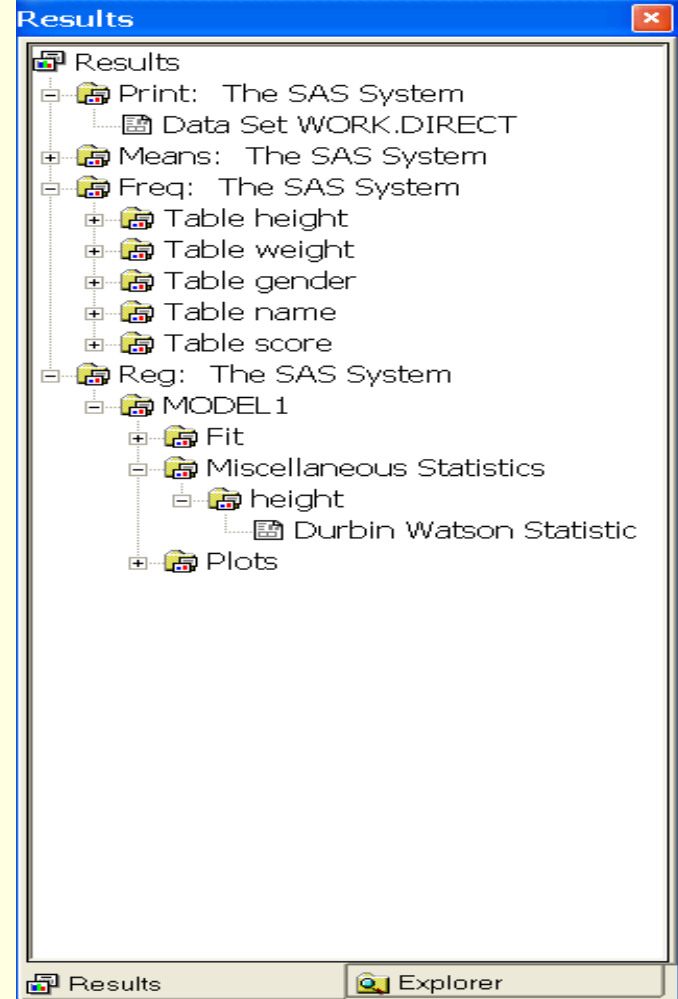
- To move backward from one folder to another in the Explorer Window, simply click the left most icon on the toolbar that looks like a folder.

Click



# 'Results' window

- Results window
  - This window allows you to view all the results of procedures you have executed in the program editor.
  - Use the expansion icons (+ or - icons) next to the folder to open or hide its contents.





# Points to Remember in SAS program

---

- All SAS statements begin with a keyword and end with a semicolon (;)
- Except for within the data section, SAS is not sensitive to spacing between words: the amount of space you put between words does not matter.
- Comments are entered in a SAS program using either the following formats:
  - /\* comments \*/ (used for large comment blocks)
  - \* comments ; (used for single line comments)

## **Section 2**

Basics of managing data files:

DATA step, LIBNAME, PROC export,  
and PROC import, and data  
transformation

# Practice Round: Getting data

---

- Go to  
[http://www.uri.edu/its/instructional\\_on-line\\_materials/sasbasics.html](http://www.uri.edu/its/instructional_on-line_materials/sasbasics.html)
- Download the SAS program from  
<http://www.uri.edu/its/research/basics.txt>
- Download two data files from  
<http://www.uri.edu/its/research/scores.txt>  
<http://www.uri.edu/its/research/scores2.txt>
- After opening these files, select 'Save As' under File. Save these as C:\ basics.txt, C:\ scores.txt, and C:\ scores2.txt.

# Importing direct data

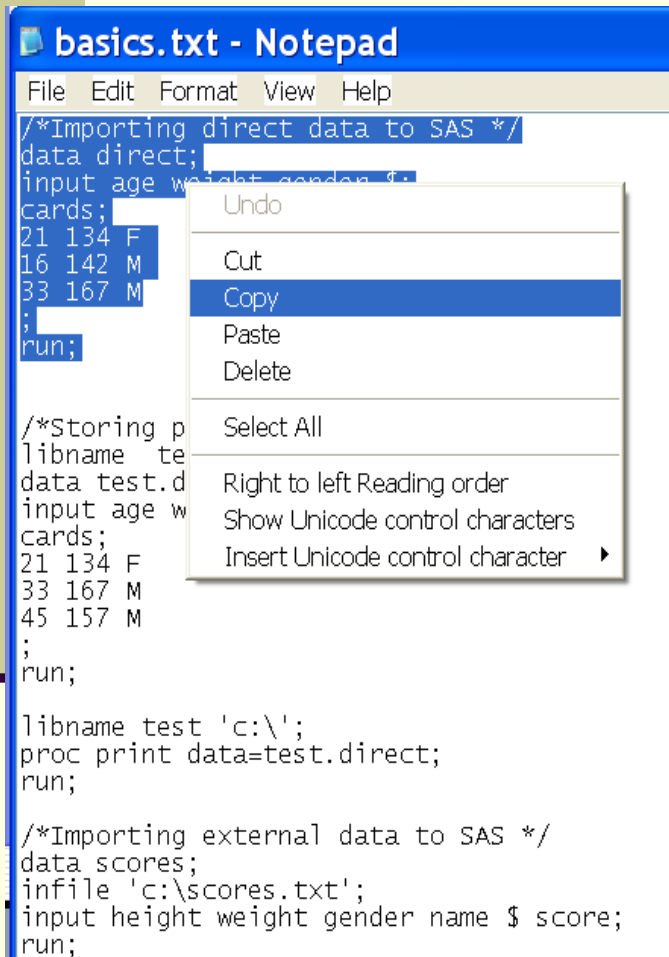
---

- Open basics.txt with 'MS Word' or 'Notepad.'
- Drag lines shown below in the file and copy and paste it to the 'editor' window in SAS.

```
data direct;  
input age weight gender $;  
cards;  
21 134 F  
33 167 M  
45 157 M  
;  
run;
```

- 'cards' statement allows you to put raw data directly to SAS

You can copy and paste also with your key board.  
Copy is Ctrl-C and paste is Ctrl-V.

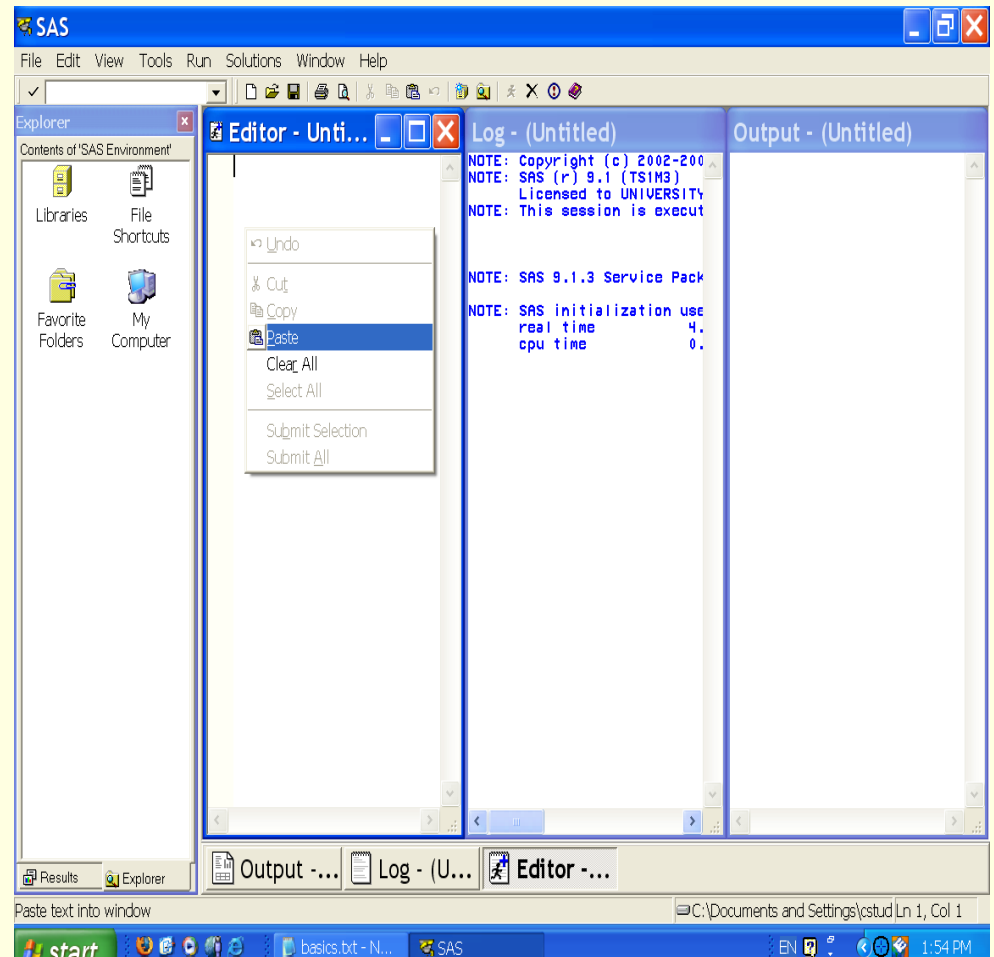


```
File Edit Format View Help
/*Importing direct data to SAS */
data direct;
input age weight gender $;
cards;
21 134 F
16 142 M
33 167 M
;
run;

/*Storing p
libname test 'c:\';
data test.direct;
input age w
cards;
21 134 F
33 167 M
45 157 M
;
run;

libname test 'c:\';
proc print data=test.direct;
run;

/*Importing external data to SAS */
data scores;
infile 'c:\scores.txt';
input height weight gender name $ score;
run;
```



SAS

File Edit View Tools Run Solutions Window Help

Explorer Contents of 'SAS Environment'

Libraries File Shortcuts

Favorite Folders My Computer

Editor - Untitled

Log - (Untitled)

Output - (Untitled)

NOTE: Copyright (c) 2002-200...  
NOTE: SAS (r) 9.1 (TS1M3)  
Licensed to UNIVERSITY  
NOTE: This session is execut

NOTE: SAS 9.1.3 Service Pack

NOTE: SAS initialization use  
real time 4.  
cpu time 0.

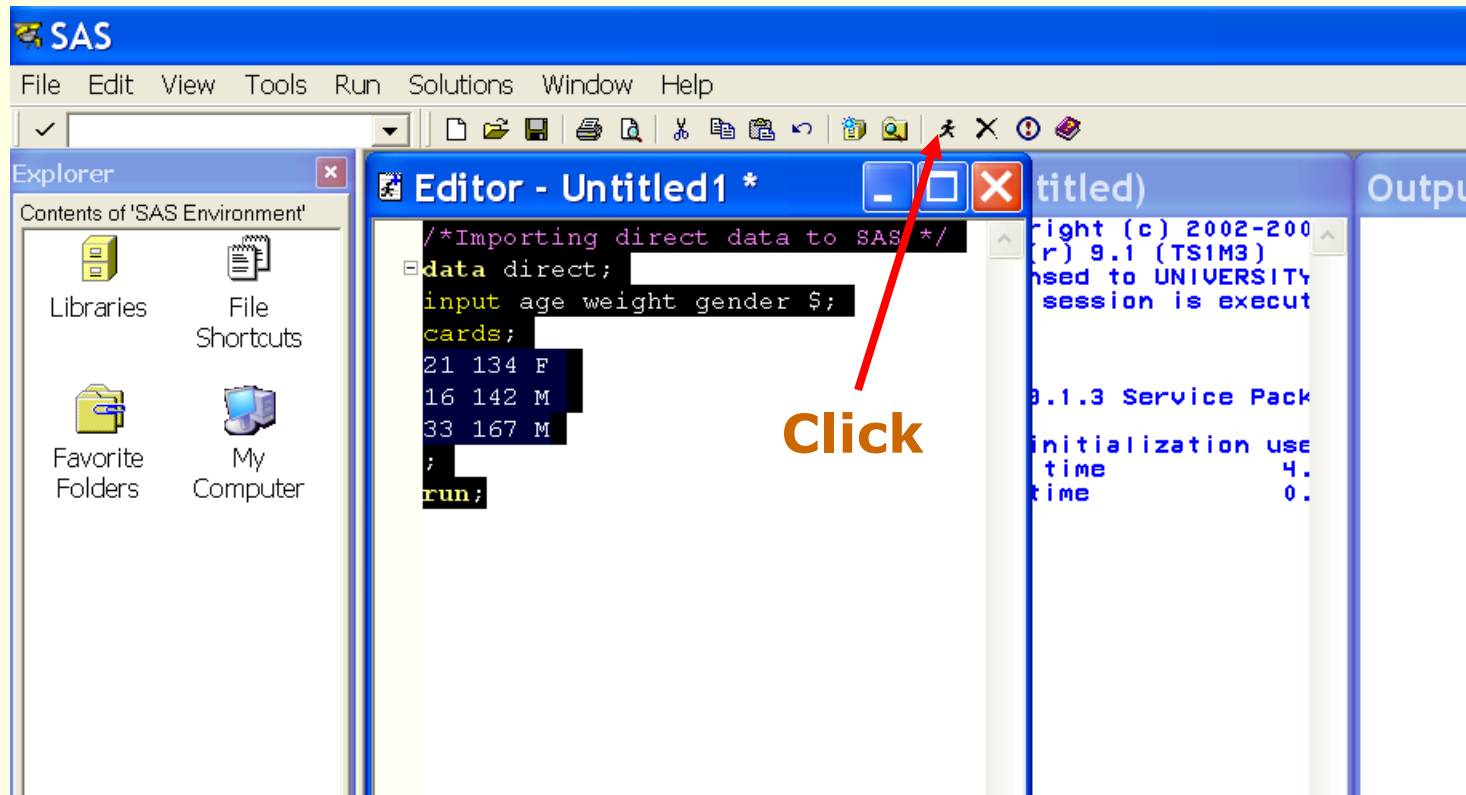
Results Explorer Output -... Log - (U... Editor -...

Paste text into window C:\Documents and Settings\cstud Ln 1, Col 1

start basics.bt - N... SAS EN 1:54 PM

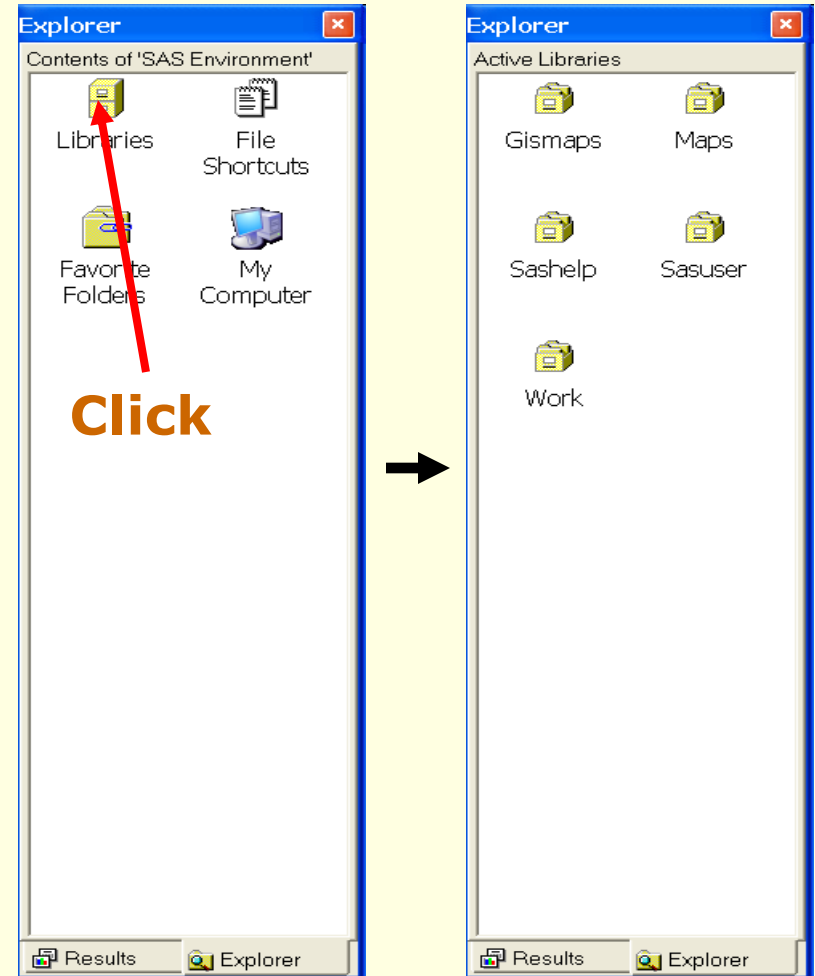
# Importing direct data: Executing the commands

- To execute the commands, highlight it and click the 'submit' icon or select 'submit' under the Run menu.



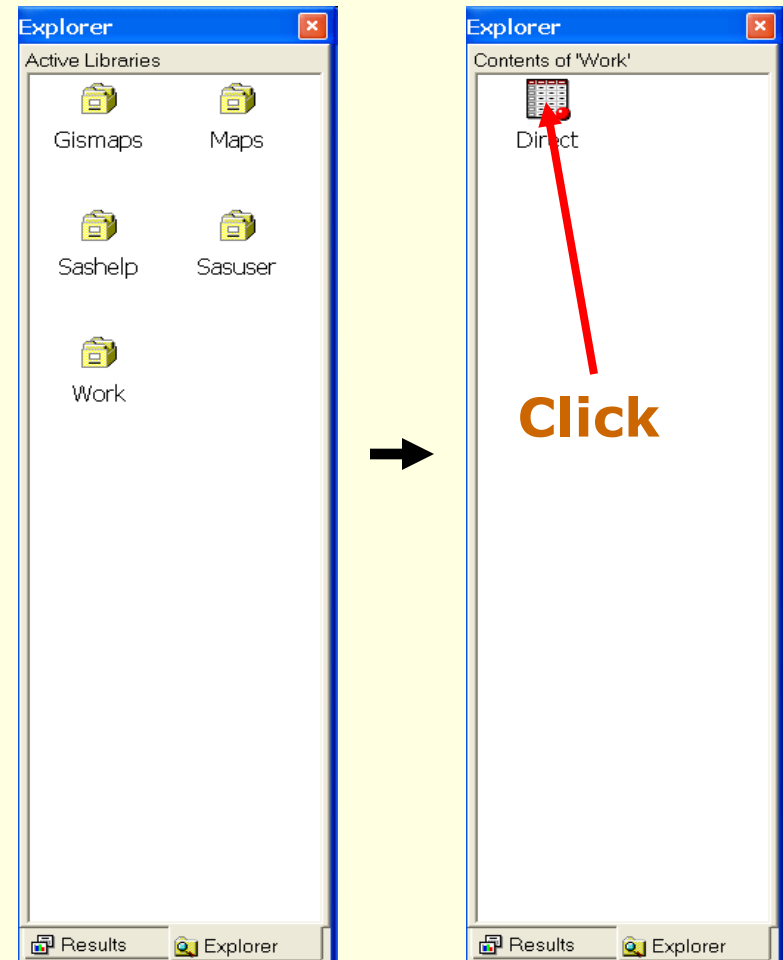
# Data command

- data direct;
  - Allows SAS to create a temporary SAS data file.
  - In this example the file was named 'direct' but you can have your own name by renaming 'direct.'
  - In the 'Explorer' window click Libraries.



# Data command: How to see your data in the SAS library

- Now click and go into the 'Work' library.
  - You should see the 'direct' file you have just created in the library.
- Finally click the 'direct' file in the work library.
  - You should be able to see the 'veiwtable window'







# VIEWTABLE: Work.Direct

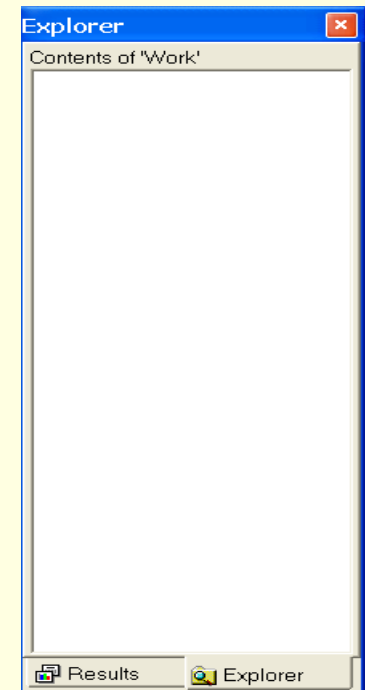


	age	weight	gender
1	21	134	F
2	16	142	M
3	33	167	M

# 'Work' library

---

- The data in the 'Work' library is not stored permanently in SAS. The work folder store files only **temporarily**. Once you exit the SAS program the file will be erased from the folder.
  - End SAS session.
  - Open SAS again and look in the Work library. **NO DATA FILES!**



# LIBNAME statement

- To store the data permanently, you need to create and reference a library  
⇒ Use LIBNAME statement
- Drag the lines shown below from the file 'basics.txt' and copy and paste it to the 'editor' window in SAS.

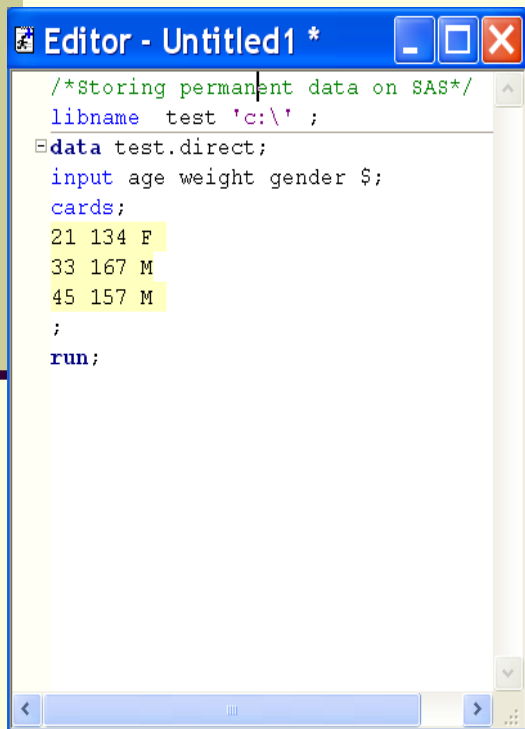
```
libname test 'C:\';  
data test.direct;  
input age weight gender $;  
cards;  
21 134 F  
33 167 M  
45 157 M  
;  
run;
```

Name of the  
library

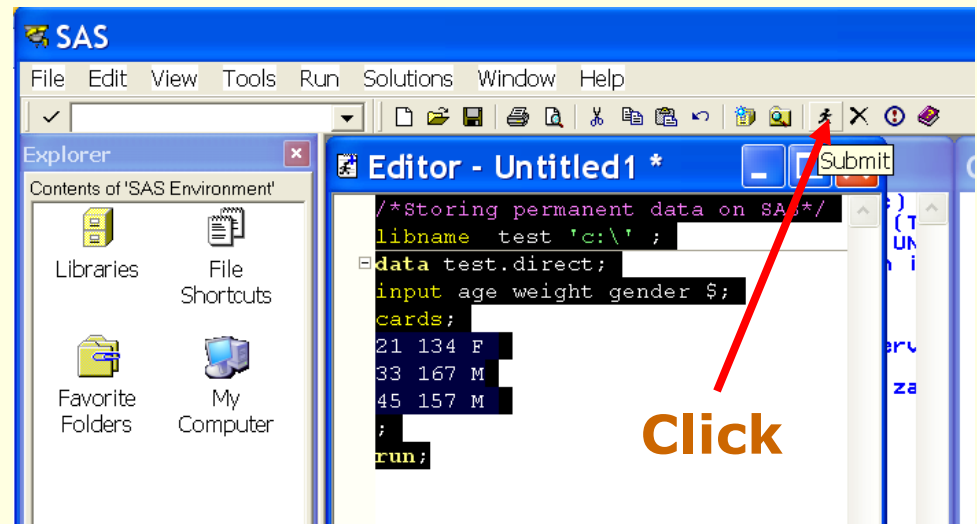
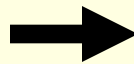
Name of the file

# LIBNAME statement (cont'd)

- After pasting the commands to the 'editor' window of SAS, highlight the commands and then click submit.



```
Editor - Untitled1 *  
/*Storing permanent data on SAS*/  
libname test 'c:\';  
data test.direct;  
input age weight gender $;  
cards;  
21 134 F  
33 167 M  
45 157 M  
;  
run;
```



SAS

File Edit View Tools Run Solutions Window Help

Submit

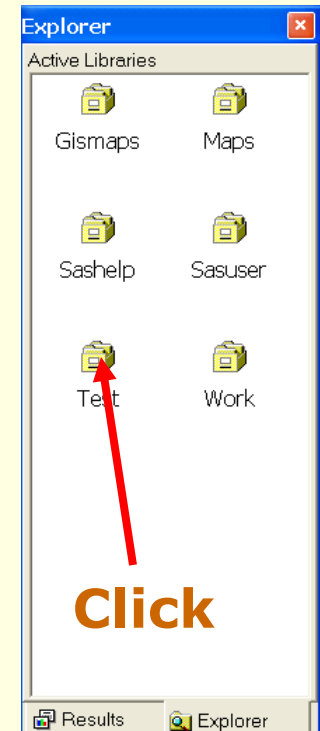
```
Editor - Untitled1 *  
/*Storing permanent data on SAS*/  
libname test 'c:\';  
data test.direct;  
input age weight gender $;  
cards;  
21 134 F  
33 167 M  
45 157 M  
;  
run;
```

Click

# LIBNAME statement (cont'd)

- The command submitted has created a new library named 'test' on SAS, and saved data file 'direct' in this library and in the 'C:\' folder of your computer.

- In the 'Explorer' window click Libraries. Then go into the 'test' library.
- Click the 'test library'



# LIBNAME statement (cont'd)

- You will now see the 'Direct' file in the 'test' library.
- To view the 'Direct' data file click 'direct'
- You will also find the same file in the 'C:\' folder of your computer.



VIEWTABLE: Test.Direct

	age	weight	gender
1	21	134	F
2	33	167	M
3	45	157	M

# LIBNAME statement (cont'd)

---

- Once you stored your data file into your C:/ drive with the LIBNAME statement, you can refer to the file without importing the raw data again.
  - Example:
    - Close SAS session, re-open it.
    - Then copy and paste the following commands from 'basics.txt' to the 'editor' window in SAS.

```
libname test 'C:\';  
proc print data=test.direct;  
run;
```

- Click the submit icon to execute the command

basics.txt - Note... [min] [max] [close]

File Edit Format View Help

```
/*Storing permanent data on SAS*  
libname test 'c:\';  
data test.direct;  
input age weight gender $;  
cards;  
21 134 F  
33 167 M  
45 157 M  
;  
run;  
  
libname test 'c:\';  
proc print  
run;
```

/\*Importin  
data score  
infile 'c:  
input heig  
run;  
data score  
infile 'c:  
input heig  
name \$ 8-1  
run;  
  
/\*To put o  
than one line\*/  
data linecontrol;

Undo  
Cut  
Copy  
Paste  
Delete  
Select All  
Right to left Reading order  
Show Unicode control character  
Insert Unicode control characte

Editor - Untitled1 \* [min] [max] [close]

```
libname test 'c:\';  
proc print data=test.direct;  
run;
```

SAS [min] [max] [close]

File Edit View Tools Run Solutions Window Help

Explorer  
Contents of 'SAS Environment'

- Libraries
- File Shortcuts
- Favorite Folders
- My Computer

Editor - Untitled1 \* [min] [max] [close]

```
libname test 'c:\';  
proc print data=test.direct;  
run;
```

Output - (Untitl... [min] [max] [close]

Click

Results Explorer Output -... Log - (U... Editor -...

C:\Documents and Settings\Ln 3, Col 5



# You will see the same data as before!

The screenshot displays the SAS software interface. The main window is titled "SAS" and contains a menu bar (File, Edit, View, Tools, Solutions, Window, Help) and a toolbar. The interface is divided into several panes:

- Results**: A pane on the left showing a tree view with "Results" and "Print: The SAS System".
- Editor - Untitled**: A central pane containing SAS code:

```
libname test  
proc print  
run;
```
- Output - (Untitled)**: A pane on the right displaying the output of the code. It shows the text "The SAS System" followed by the date and time "09:19 Friday," and a table of data:

Obs	age	weight	gender
1	21	134	F
2	33	167	M
3	45	157	M

At the bottom of the window, there is a taskbar with icons for "Output ...", "Log - (U...", and "Editor - ...". The status bar at the very bottom shows "NOTE: At top." on the left and "C:\Documents and Settir" on the right.

# Forms of INPUT statement

---

## ■ Example 1

### ■ input age weight gender \$;

- This statement allows SAS to read the variables used for the raw data.
- In this example three variables (age, weight, and gender) were put into SAS
- SAS initially only reads numeric variable so in order to read character values you need to use modifiers:
  - The variable 'gender' is a character variable. You need to use '\$'
    - \$: enables SAS to read character values with default size of eight characters with no embedded blanks
    - &: enables SAS to read character values with embedded blanks

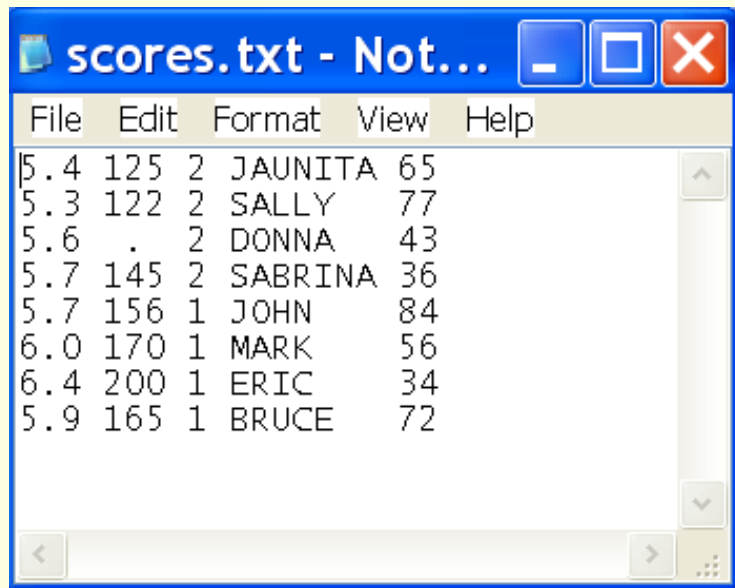
# INPUT statement: Example 2

---

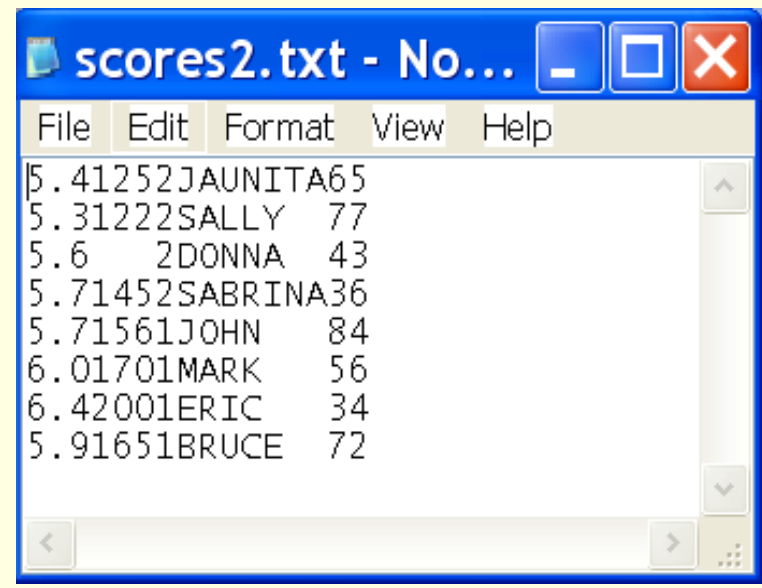
- input height **1-3** weight **4-6** gender **7** name \$ **8-14** score **15-16**;
  - If the data contain the followings you need to set up a column input mode to specify the column positions of the pointer
    - Standard character and numeric data
    - Values entered in fixed column positions
    - Character values longer than eight characters
    - Character values with embedded blanks

# Importing external data

- Open scores.txt, and scores2.txt from c:/ drive and compare.



```
5.4 125 2 JAUNITA 65
5.3 122 2 SALLY 77
5.6 . 2 DONNA 43
5.7 145 2 SABRINA 36
5.7 156 1 JOHN 84
6.0 170 1 MARK 56
6.4 200 1 ERIC 34
5.9 165 1 BRUCE 72
```



```
5.41252JAUNITA65
5.31222SALLY 77
5.6 2DONNA 43
5.71452SABRINA36
5.71561JOHN 84
6.01701MARK 56
6.42001ERIC 34
5.91651BRUCE 72
```

# Importing external data (Cont'd)

---

- Open basics.txt with 'MS Word' or 'Notepad.'
- Drag the lines shown below on the file, copy and paste it to 'editor' window in SAS, and execute the commands.

```
data scores;  
infile 'C:\scores.txt';  
input height weight gender name $ score;  
run;  
data scores2;  
infile 'C:\scores2.txt';  
input height 1-3 weight 4-6 gender 7  
name $ 8-14 score 15-16;  
run;
```

# Importing external data (Cont'd)

- Go to the 'Explorer' window, click the work folder, and open 'scores' and 'scores2.' You will see exactly the same file.



	height	weight	gender	name	score
1	5.4	125	2	JAUNITA	65
2	5.3	122	2	SALLY	77
3	5.6	.	2	DONNA	43
4	5.7	145	2	SABRINA	36
5	5.7	156	1	JOHN	84
6	6	170	1	MARK	56
7	6.4	200	1	ERIC	34
8	5.9	165	1	BRUCE	72

	height	weight	gender	name	score
1	5.4	125	2	JAUNITA	65
2	5.3	122	2	SALLY	77
3	5.6	.	2	DONNA	43
4	5.7	145	2	SABRINA	36
5	5.7	156	1	JOHN	84
6	6	170	1	MARK	56
7	6.4	200	1	ERIC	34
8	5.9	165	1	BRUCE	72

# INPUT statement: Example 3

- How to put observations in more than one line
  - #n: moves the pointer to record n.

- Example

```
data linecontrol;
```

```
input #1 name $ height weight #2 country & $24.
```

```
#3 score1 score2;
```

```
cards;
```

```
Ken 5.9 158
```

```
Great Britain
```

```
44 36
```

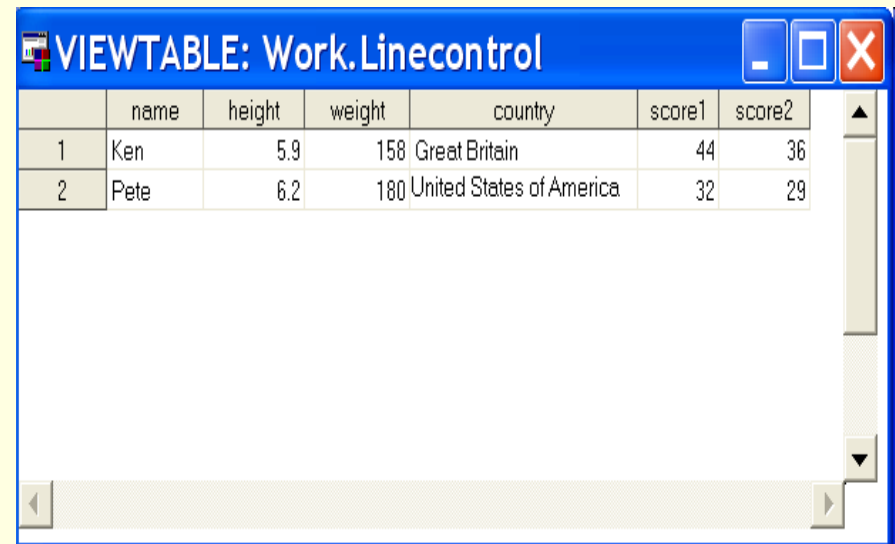
```
Pete 6.2 180
```

```
United States of America
```

```
32 29
```

```
;
```

```
run;
```

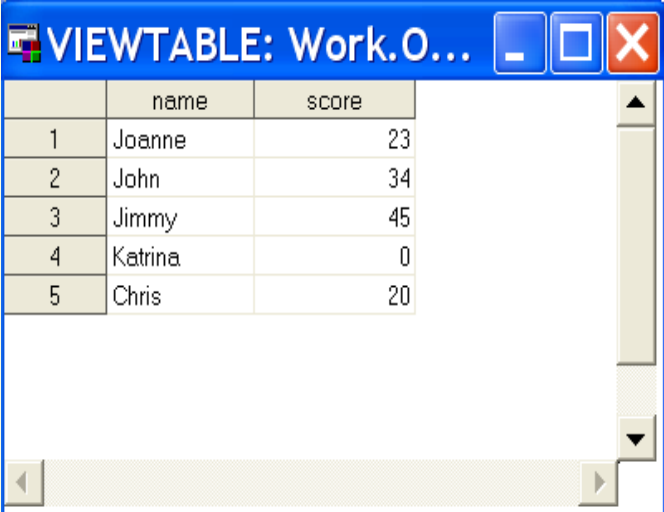


	name	height	weight	country	score1	score2
1	Ken	5.9	158	Great Britain	44	36
2	Pete	6.2	180	United States of America	32	29

# INPUT statement: Example 4

- How to put several observations in one line
  - @@: Used when each input line contains values for several observations
  - Example

```
data oneline;  
input name $ score @@;  
cards;  
Joanne 23 John 34 Jimmy 45  
Katrina 0 Chris 20  
;  
run;
```



VIEWTABLE: Work.O...

	name	score
1	Joanne	23
2	John	34
3	Jimmy	45
4	Katrina	0
5	Chris	20



# Exporting & Importing MS Excel data 1

---

**/\*Exporting data to MS Excel data\*/**

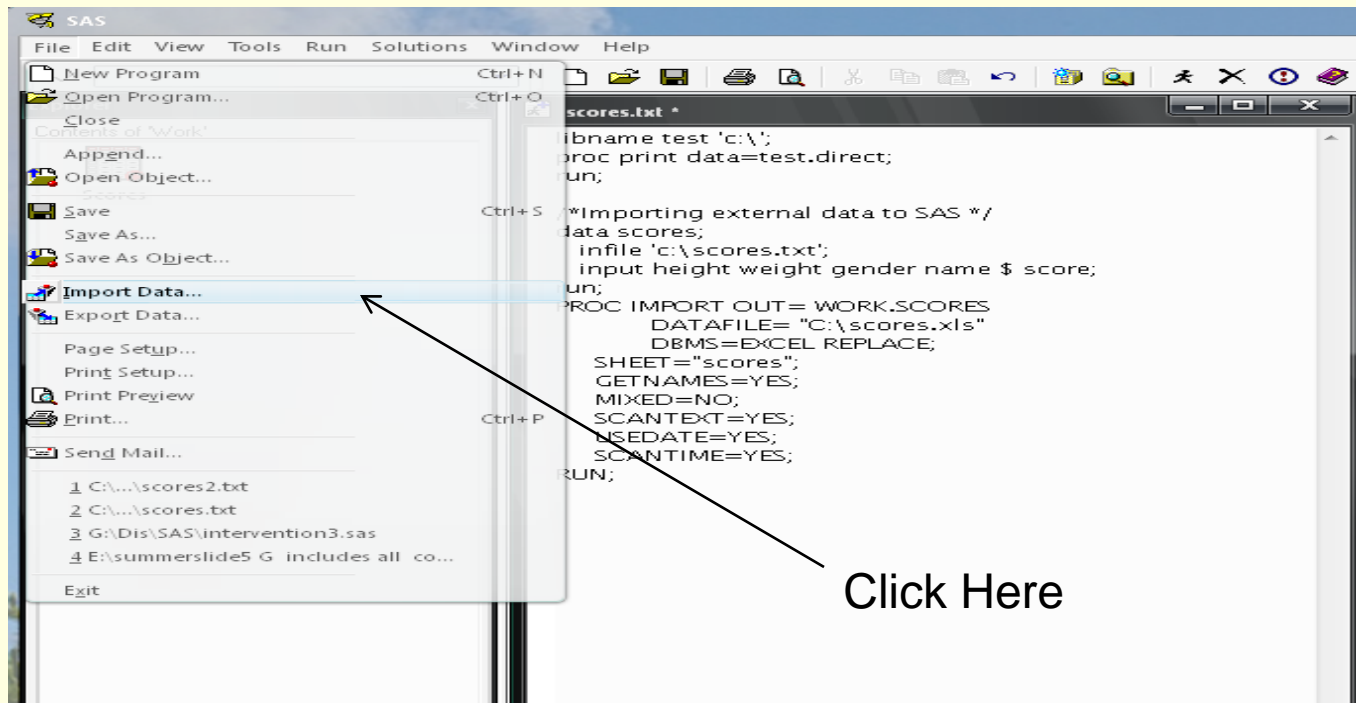
```
proc export data=scores  
    outfile="C:\scores.xls"  
    dbms=excel2000 replace;  
    sheet="scores";  
run;
```

**/\*Importing data from MS Excel\*/**

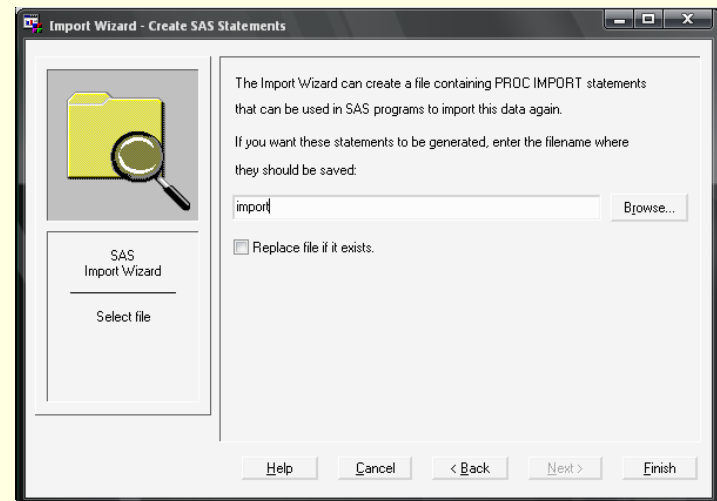
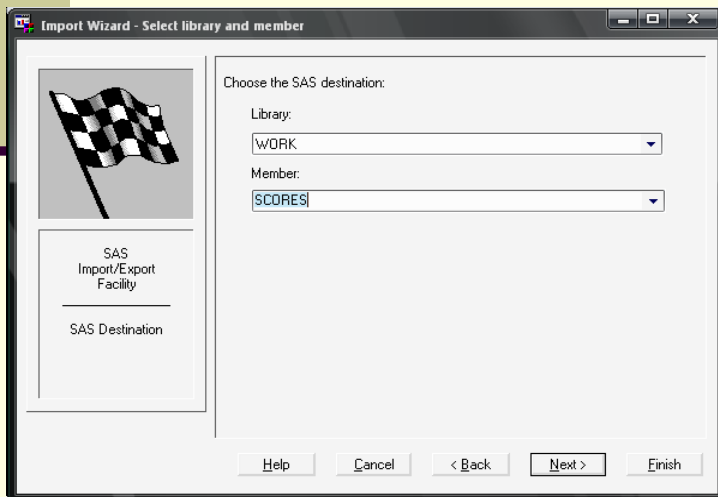
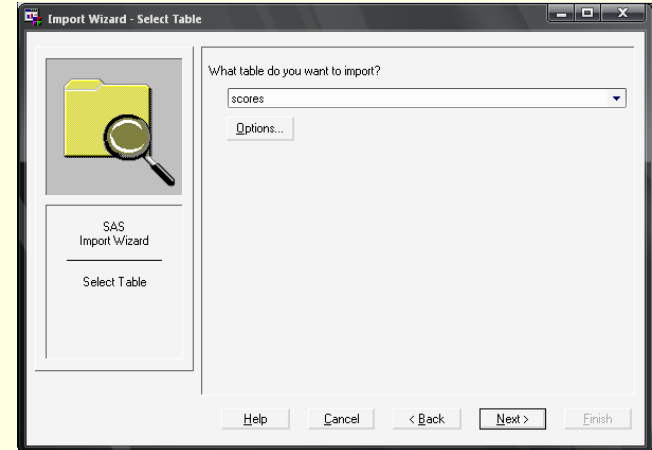
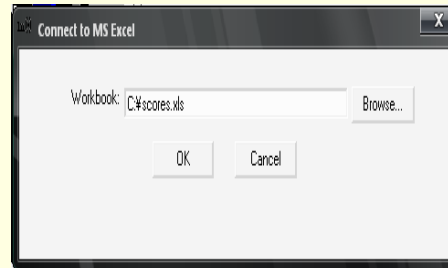
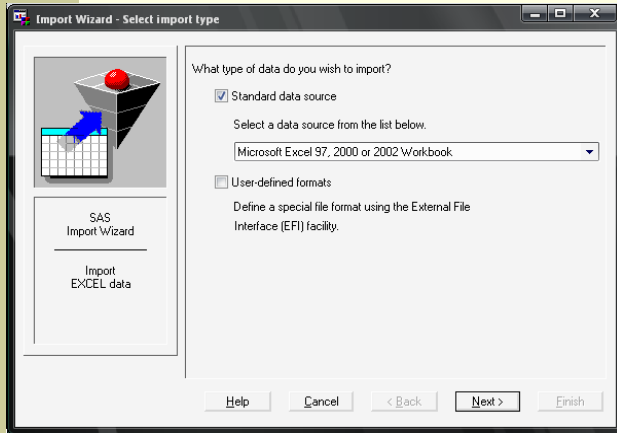
```
proc import out=impscores  
    datafile="C:\scores.xls"  
    dbms=excel2000 replace;  
    sheet="scores";  
    getnames=yes;  
    mixed=yes;  
  
run;
```

# Importing MS Excel data 2-1

- Click file under the main tab and open “Import Data”



# Importing MS Excel data 2-2



# Data Transformation

---

- How to transform data in SAS

```
data trans;
```

```
set scores;
```

```
* 'Set' command allows reusing created SAS  
data;
```

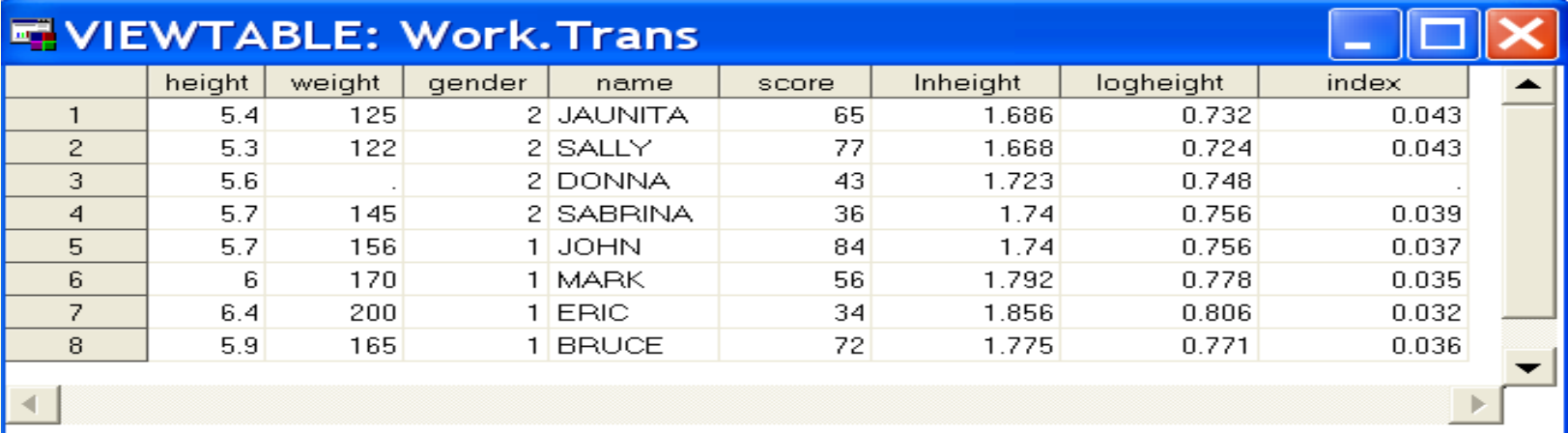
```
lnheight=log(height);
```

```
logheight=log10(height);
```

```
index=height/weight;
```

```
run;
```

# Data Transformation (Cont'd)



	height	weight	gender	name	score	lnheight	logheight	index
1	5.4	125	2	JAUNITA	65	1.686	0.732	0.043
2	5.3	122	2	SALLY	77	1.668	0.724	0.043
3	5.6	.	2	DONNA	43	1.723	0.748	.
4	5.7	145	2	SABRINA	36	1.74	0.756	0.039
5	5.7	156	1	JOHN	84	1.74	0.756	0.037
6	6	170	1	MARK	56	1.792	0.778	0.035
7	6.4	200	1	ERIC	34	1.856	0.806	0.032
8	5.9	165	1	BRUCE	72	1.775	0.771	0.036

## ■ Note

- **LOG(x)** : the natural logarithm of  $x$
- **LOG10(x)** : the log base ten of  $x$
- **LOG2(x)** : the log base two of  $x$

# Arithmetic and Comparison Operators

## Arithmetic Operators

Symbol	Definition	Example
**	exponentiation	a**3
*	multiplication	2*y
/	division	var/5
+	addition	num+3
-	subtraction	sale-discount

## Comparison Operators

Symbol	Definition	Example
=	equal to	a = 3
^= or NE	not equal to	a ne 3
≠ or NE	not equal to	
~= or NE	not equal to	
> or GT	greater than	num > 5
< or LT	less than	num < 8
>= or GE	greater than or equal to	sales >= 300
<= or LE	less than or equal to	sales <= 100

# Data Modification: If / then Statements

---

- How to delete certain observations from data
  - Example: The following command deletes observations having weight more than 160

```
data modify;  
set trans; *'Set' command allows reusing created  
SAS data;  
if weight > 160 then delete;  
run;
```
- Open the created data file 'modify' in the 'Work' folder of your library and compare that from the data file 'trans.'

You can see that observations for 'Mark,' 'Eric,' and 'Bruce,' have been deleted in 'modify.'

	height	weight	gender	name	score	lnheight	logheight	index
1	5.4	125	2	JAUNITA	65	1.686	0.732	0.043
2	5.3	122	2	SALLY	77	1.668	0.724	0.043
3	5.6	.	2	DONNA	43	1.723	0.748	.
4	5.7	145	2	SABRINA	36	1.74	0.756	0.039
5	5.7	156	1	JOHN	84	1.74	0.756	0.037
6	6	170	1	MARK	56	1.792	0.778	0.035
7	6.4	200	1	ERIC	34	1.856	0.806	0.032
8	5.9	165	1	BRUCE	72	1.775	0.771	0.036

	height	weight	gender	name	score	lnheight	logheight	index
1	5.4	125	2	JAUNITA	65	1.686	0.732	0.043
2	5.3	122	2	SALLY	77	1.668	0.724	0.043
3	5.6	.	2	DONNA	43	1.723	0.748	.
4	5.7	145	2	SABRINA	36	1.74	0.756	0.039
5	5.7	156	1	JOHN	84	1.74	0.756	0.037



# Section 3

Basic commands in SAS:

PROC step

# Proc Steps: proc print

- Use: to see the SAS data file in the output window  
proc print data=scores;  
run;

Obs	height	weight	gender	name	score
1	5.4	125	2	JAUNITA	65
2	5.3	122	2	SALLY	77
3	5.6	143	2	DONNA	43
4	5.7	145	2	SABRINA	36
5	5.7	156	1	JOHN	84
6	6.0	170	1	MARK	55
7	6.4	200	1	ERIC	34
8	6.9	165	1	BRUCE	72

# Proc Steps: proc contents

- Use: to see the contents of SAS data file  
`proc contents data=scores;`  
`run;`

The screenshot shows the SAS Output window for the PROC CONTENTS procedure. The output is as follows:

```
The SAS System 11:33 Friday, September 29, 2006
The CONTENTS Procedure

Data Set Name      WORK.Scores      Observations      8
Member Type       DATA           Variables         5
Engine            US              Indexes           0
Created           Friday, September 29, 2006 11:34:12 AM      Observation Length 4
Last Modified     Friday, September 29, 2006 11:34:12 AM      Deleted Observations 0
Protection
Data Set Type
Label
Data Representation WINDOWS_32
Encoding          wlatin1 - Western (Windows)
Compressed
Sorted            NI

Engine/Host Dependent Information

Data Set Page Size      4096
Number of Data Set Pages 1
First Data Page        1
Max Obs per Page       101
Obs in First Data Page 8
Number of Data Set Repairs 0
File Name               C:\DOCUME~1\cstudent\LOCALS~1\Temp\SAS
                       Temporary Files\_TD3688\scores.sas7bdat
Release Created        9.0101M3
Host Created           XP_PRD

Alphabetic List of Variables and Attributes

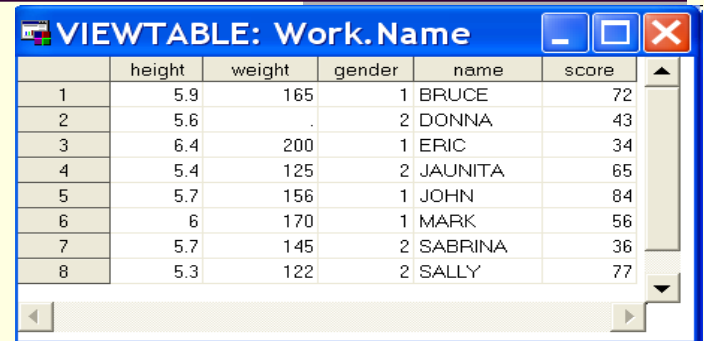
#   Variable   Type   Len
3   gender     Num    8
1   height     Num    8
4   name       Char   8
5   score      Num    8
2   weight     Num    8
```

# Proc Steps: proc sort

■ Use: to sort SAS data file  
**proc sort data=scores out=name**  
**by name;** \*Sorts the data by name  
in alphabetical orders  
**run;**

**proc sort data=scores out=height**  
**by height;** \*Sorts the data by height  
in ascending orders;  
**Run;**

**proc sort data=scores out=height2**  
**by descending height;**  
\*Sorts the data by height  
in descending orders;  
**run;**



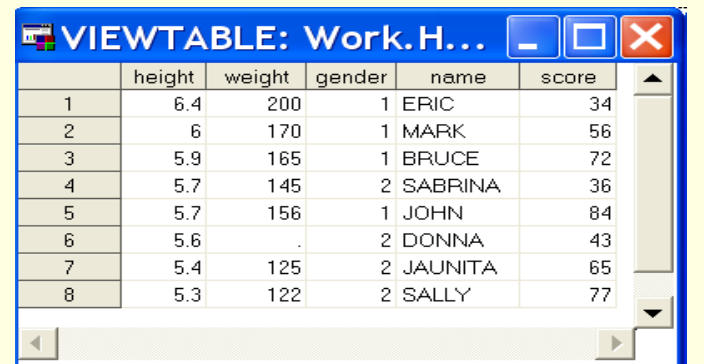
VIEWTABLE: Work.Name

	height	weight	gender	name	score
1	5.9	165	1	BRUCE	72
2	5.6	.	2	DONNA	43
3	6.4	200	1	ERIC	34
4	5.4	125	2	JAUNITA	65
5	5.7	156	1	JOHN	84
6	6	170	1	MARK	56
7	5.7	145	2	SABRINA	36
8	5.3	122	2	SALLY	77



VIEWTABLE: Work.Height

	height	weight	gender	name	score
1	5.3	122	2	SALLY	77
2	5.4	125	2	JAUNITA	65
3	5.6	.	2	DONNA	43
4	5.7	145	2	SABRINA	36
5	5.7	156	1	JOHN	84
6	5.9	165	1	BRUCE	72
7	6	170	1	MARK	56
8	6.4	200	1	ERIC	34



VIEWTABLE: Work.H...

	height	weight	gender	name	score
1	6.4	200	1	ERIC	34
2	6	170	1	MARK	56
3	5.9	165	1	BRUCE	72
4	5.7	145	2	SABRINA	36
5	5.7	156	1	JOHN	84
6	5.6	.	2	DONNA	43
7	5.4	125	2	JAUNITA	65
8	5.3	122	2	SALLY	77

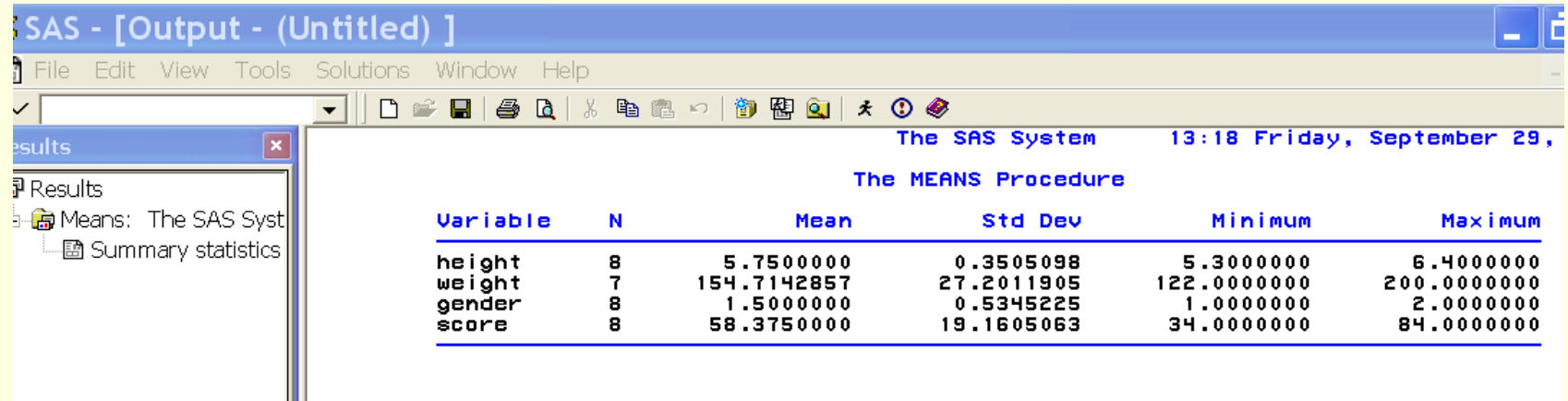
# Proc Steps: proc means

- Use: to see basic simple statistics of data

```
proc means data=scores;
```

```
run;
```

\*This provides the number of obs, mean, std, min, and max of all numeric variables;

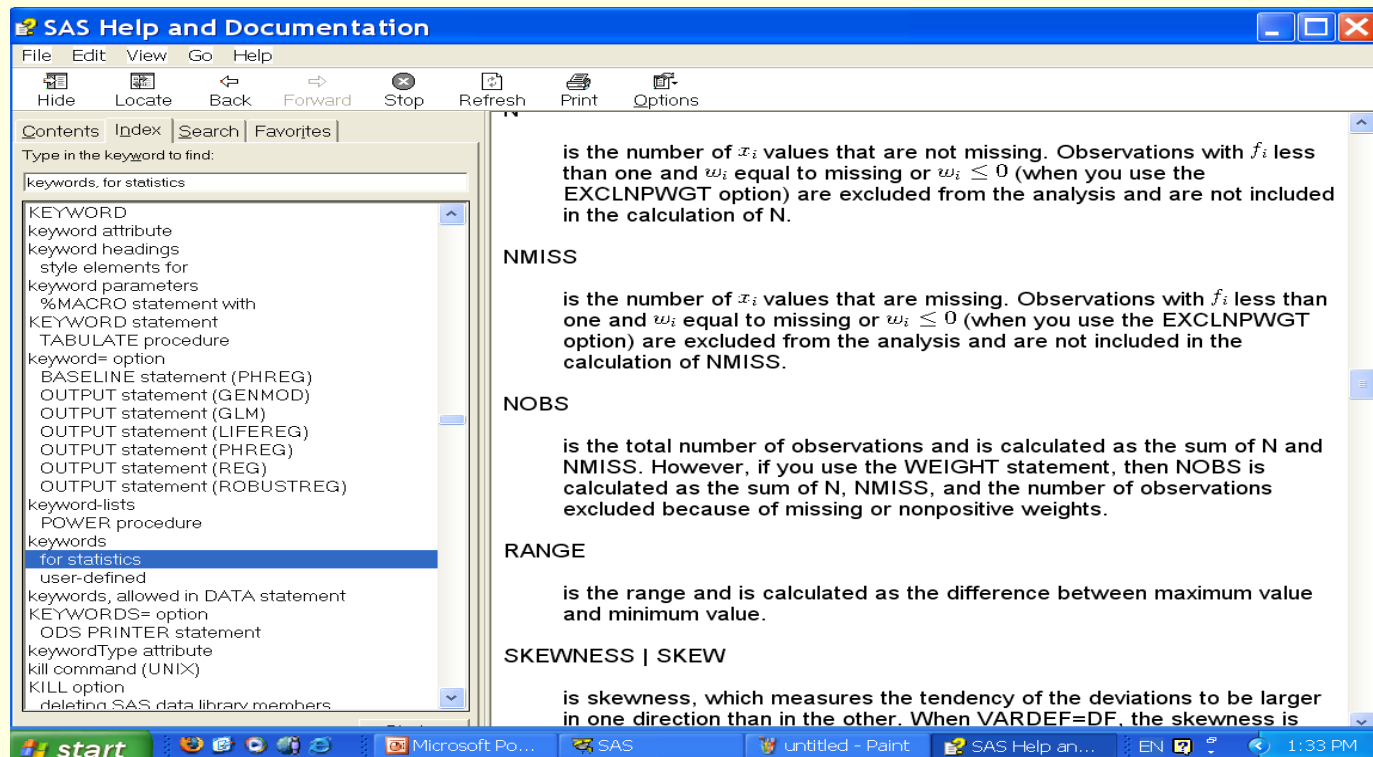


The screenshot shows the SAS Output window titled "SAS - [Output - (Untitled)]". The window displays the results of a PROC MEANS procedure. The output is titled "The SAS System" and "The MEANS Procedure". The results are presented in a table with the following columns: Variable, N, Mean, Std Dev, Minimum, and Maximum. The variables listed are height, weight, gender, and score.

Variable	N	Mean	Std Dev	Minimum	Maximum
height	8	5.7500000	0.3505098	5.3000000	6.4000000
weight	7	154.7142857	27.2011905	122.0000000	200.0000000
gender	8	1.5000000	0.5345225	1.0000000	2.0000000
score	8	58.3750000	19.1605063	34.0000000	84.0000000

# Proc means: How to see other simple statistics

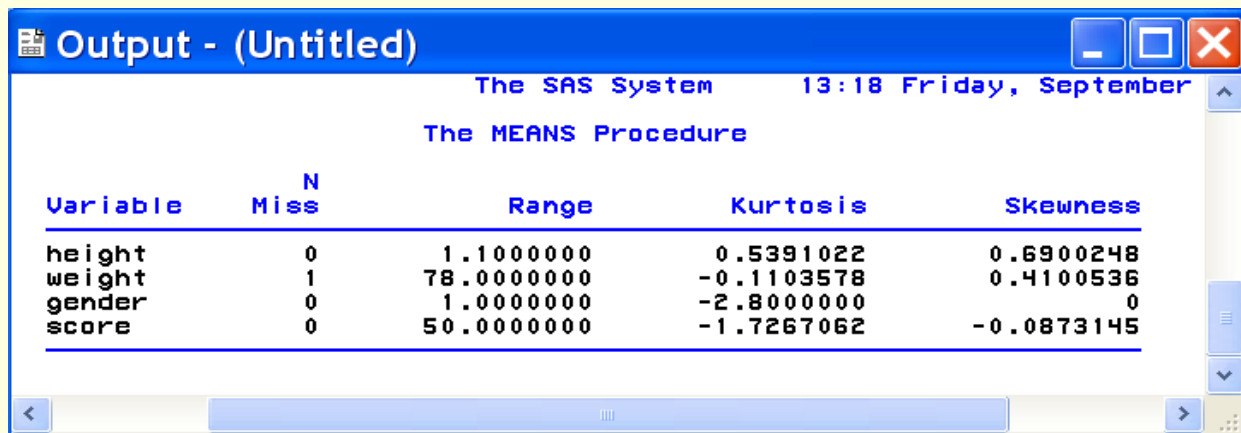
- To find out the commands for other simple statistics click the help icon and then click index. Then type in 'keywords' in the search box and enter. Finally, click 'for statistics';



# Proc Steps: How to see other simple statistics (Cont'd)

- Insert the commands for the simple statistics you want to calculate with SAS before the command 'data="file name"':

```
proc means nmiss range kurt skew data=scores;  
run;
```



The screenshot shows a window titled "Output - (Untitled)" from "The SAS System" on "Friday, September 13:18". The window displays the output of "The MEANS Procedure". The output is a table with the following columns: Variable, N Miss, Range, Kurtosis, and Skewness. The data rows are: height (N Miss: 0, Range: 1.1000000, Kurtosis: 0.5391022, Skewness: 0.6900248), weight (N Miss: 1, Range: 78.0000000, Kurtosis: -0.1103578, Skewness: 0.4100536), gender (N Miss: 0, Range: 1.0000000, Kurtosis: -2.8000000, Skewness: 0), and score (N Miss: 0, Range: 50.0000000, Kurtosis: -1.7267062, Skewness: -0.0873145).

Variable	N Miss	Range	Kurtosis	Skewness
height	0	1.1000000	0.5391022	0.6900248
weight	1	78.0000000	-0.1103578	0.4100536
gender	0	1.0000000	-2.8000000	0
score	0	50.0000000	-1.7267062	-0.0873145

# Proc Steps: proc freq

---

## ■ proc freq

- Use: to analyze frequency of the variables and to create frequency tables for variables

```
proc freq data=scores;
```

```
run; *shows one-way frequencies;
```

```
proc freq data=scores;
```

```
tables gender*weight;
```

```
run; *creates cross-tabulation table;
```



Output - (Untitled) The SAS System 13:18 Friday, September

The FREQ Procedure

height	Frequency	Percent	Cumulative Frequency	Cumulative Percent
5.3	1	12.50	1	12.50
5.4	1	12.50	2	25.00
5.6	1	12.50	3	37.50
5.7	2	25.00	5	62.50
5.9	1	12.50	6	75.00
6	1	12.50	7	87.50
6.4	1	12.50	8	100.00

weight	Frequency	Percent	Cumulative Frequency	Cumulative Percent
122	1	14.29	1	14.29
125	1	14.29	2	28.57
145	1	14.29	3	42.86
156	1	14.29	4	57.14
165	1	14.29	5	71.43
170	1	14.29	6	85.71
200	1	14.29	7	100.00

Frequency Missing = 1

gender	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	4	50.00	4	50.00
2	4	50.00	8	100.00

name	Frequency	Percent	Cumulative Frequency	Cumulative Percent
------	-----------	---------	----------------------	--------------------

Output - (Untitled) The SAS System 14:11 Monday, October 2, 20

The FREQ Procedure

Table of gender by weight

gender	weight							Total
	122	125	145	156	165	170	200	
1	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 14.29 25.00 100.00	1 14.29 25.00 100.00	1 14.29 25.00 100.00	1 14.29 25.00 100.00	4 57.14
2	1 14.29 33.33 100.00	1 14.29 33.33 100.00	1 14.29 33.33 100.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	3 42.86
Total	1 14.29	1 14.29	1 14.29	1 14.29	1 14.29	1 14.29	1 14.29	7 100.00

Frequency Missing = 1

# Proc Steps: proc reg

---

## ■ proc reg

- One of a general-purpose procedures for regression analysis in SAS

```
proc reg data=scores;  
model height=weight / dw alpha=0.01 clb ;  
plot height*weight / cframe=ligr conf pred ;  
run;
```

$$\text{height} = \alpha + \beta \text{weight} + \varepsilon$$

Results

- Reg: The SAS System
  - MODEL1
    - Fit
      - height
        - Number of Observations
        - Analysis of Variance
        - Fit Statistics
        - Parameter Estimates
      - Miscellaneous Statistics
        - height
          - Durbin Watson Stat
      - Plots

Click

Click

Results Explorer

Output - (Untitled)

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	99% Confidence Limits	
Intercept	1	3.66848	0.12483	29.39	<.0001	3.16516	4.17180
weight	1	0.01359	0.00079634	17.07	<.0001	0.01038	0.01680

Output - (Untitled)

The SAS System 13:18 Friday

The REG Procedure  
Model: MODEL1  
Dependent Variable: height

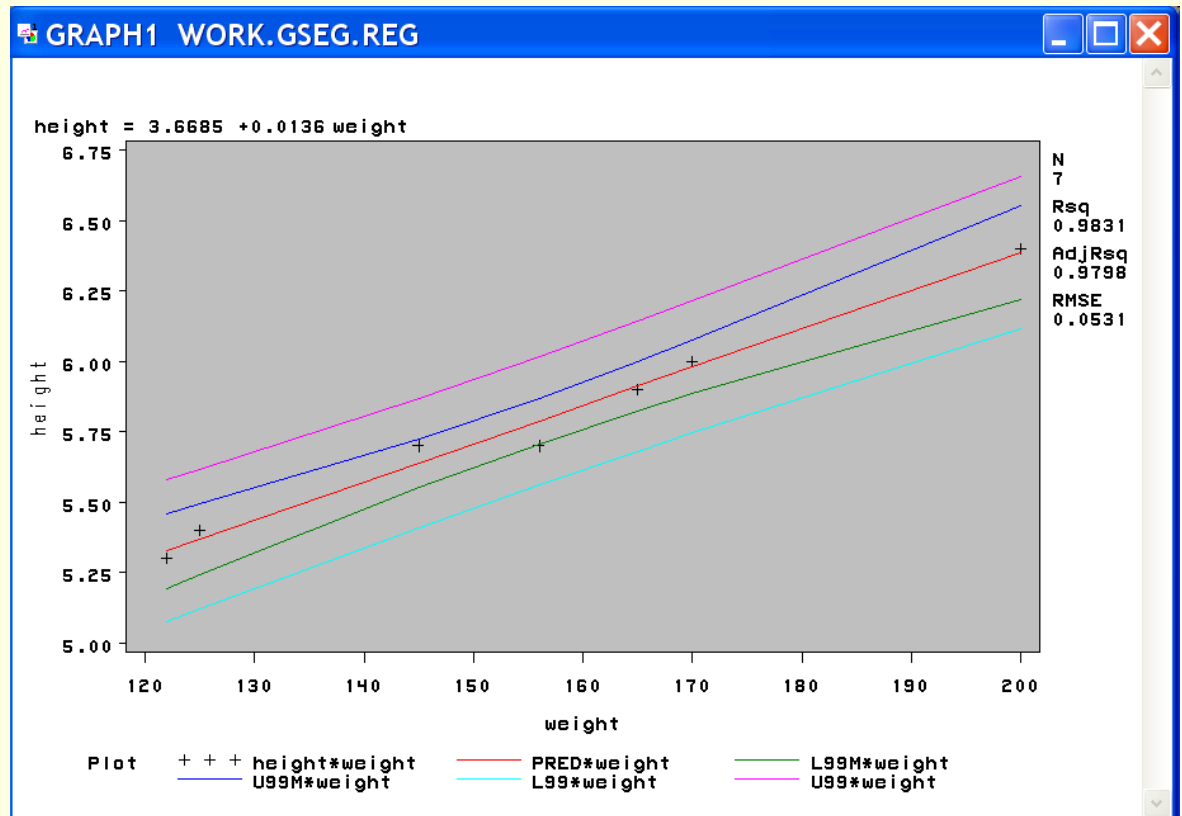
Durbin-Watson D 3.281  
Number of Observations 7  
1st Order Autocorrelation -0.682

Results

- Reg: The SAS System
  - MODEL1
    - Fit
      - height
        - Number of Obs
        - Analysis of Var
        - Fit Statistics
        - Parameter Esti
      - Miscellaneous Statisti
        - height
          - Durbin Watson
      - Plots
        - height\*weight

Click

Results Explorer



# Proc Steps: proc gplot

---

- proc gplot

- Use: to plot the values of two or more variables on a set of coordinate axes

```
proc gplot data=scores;
```

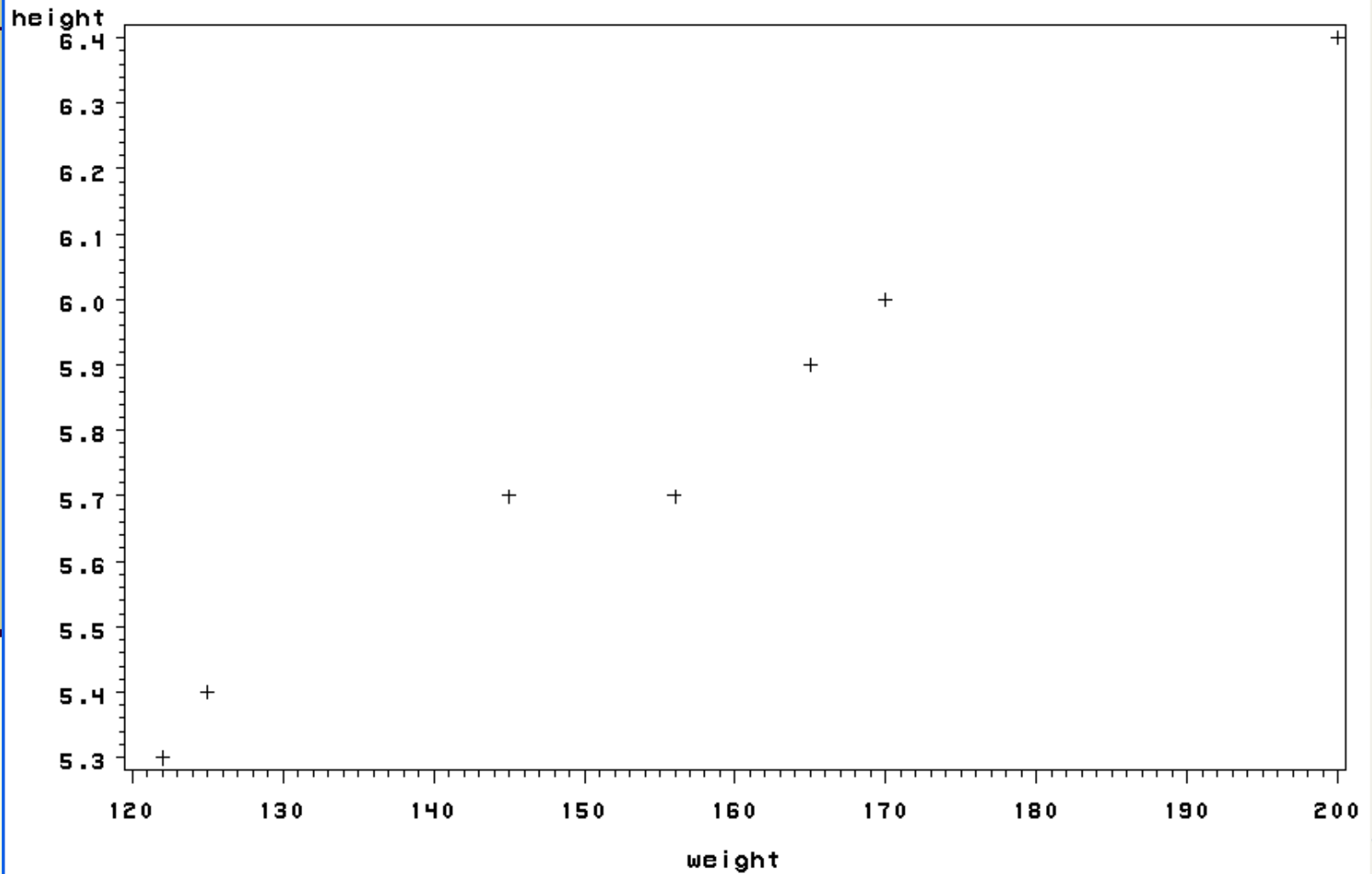
```
plot height*weight;
```

```
*height=vertical axis,
```

```
weight=horizontal axis;
```

```
run;
```

GRAPH1 WORK.GSEG.GPLOT



# Using advanced options in SAS

---

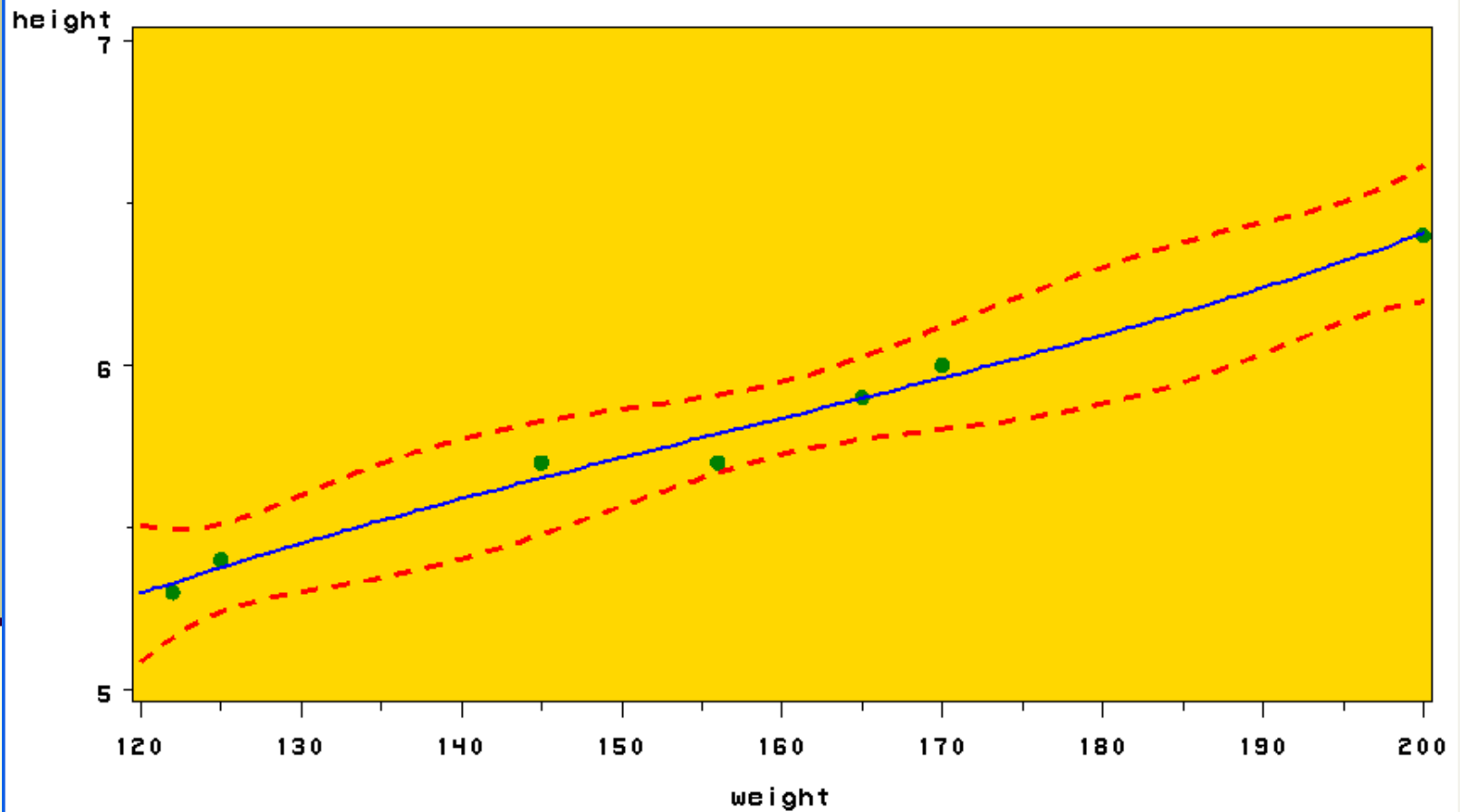
```
proc gplot data=scores;
plot height*weight
/skipmiss haxis=120 to 200 by 10 hminor=1
vaxis=5.0 to 7.0 by 1.0 vminor=1
Regeqn cframe=gold; *Options for the plot statement;

title font=arial c=blue box=3 bcolor=yellow
'Study of Height vs Weight'; *Putting a title
for your graph;

symbol i=rcclm95 value=dot height=1
cv=green ci=blue co=red width=2;
*Setting the colors and size for the plot symbol
and lines. i= can be also expressed as interpol=;
run;
```



Study of Height vs Weight



Regression Equation:  
 $height = -0.219354 + 0.090498 * weight - 0.000498 * weight^2 + 1.054E-6 * weight^3$



# Useful supports

- In the tool bar click the help menu or the help icon



**SAS Help and Documentation**

File Edit View Go Help

Hide Locate Back Forward Stop Refresh Print Options

Contents Index Search Favorites

Type in the keyword to find:

- ' (quotation mark)
- as literal character
- variable indicator
- ~ (tilde) format modifier
- ~ (tilde) format modifier, definition
- \_ (underscore), in SAS names
- ! (exclamation points), concatenation operator
- !DDE\_FLUSH string
- !SASROOT directory
- renaming
- utilities directory
- # (pound sign), variables as plot point labels
- #BYLINE option, text string specifications
- #BYVAL option, text string specifications
- #BYVAR option, text string specifications
- #n, column-pointer control
- #n, line-pointer control
- DATA step execution and skipping input variables
- \$ (dollar sign)
- defining character variables in input data
- in variable names
- \$ASCIIw. format
- \$ASCIIw. informat
- \$BINARYw. format
- \$BINARYw. informat

Display

## Using the Enhanced Editor

- [Enhanced Editor Features](#)
- [Using the Enhanced Editor Window](#)
- [Scrolling and Line Number Commands](#)
- [Creating Your Own Keywords](#)
- [Associating File Extensions with File Types](#)
- [Setting Enhanced Editor Options](#)
- [Using Keyboard Shortcuts to Customize the Enhanced Editor](#)
- [Enabling and Disabling the Enhanced Editor](#)

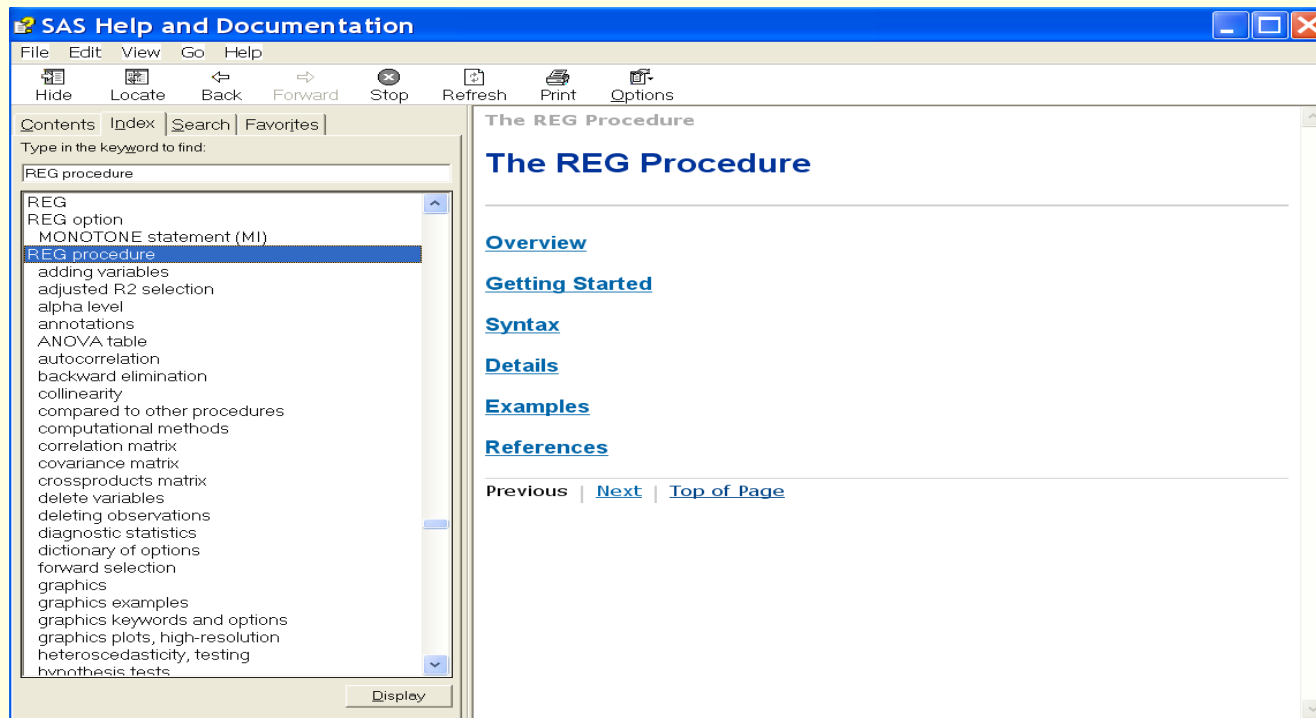
### Enhanced Editor Features

While retaining some familiar Program Editor features, the Enhanced Editor enables you to

- use color-coding to identify SAS and SCL program elements as well as HTML and XML document elements. Color-coding settings can be saved in a color scheme.
- create and format your own keywords.

# Useful supports: using the Help in SAS

- Example: click index and type 'reg.' Then double click 'REG procedure'



# Useful supports: other useful sites

---

- Online SAS manuals

<http://www.uri.edu/sasdoc>

This will automatically link you to

<http://support.sas.com/documentation/onlinedoc/sas9doc.html>

- Statbookstore: useful site for finding program examples

<http://www.geocities.com/statbookstore/>

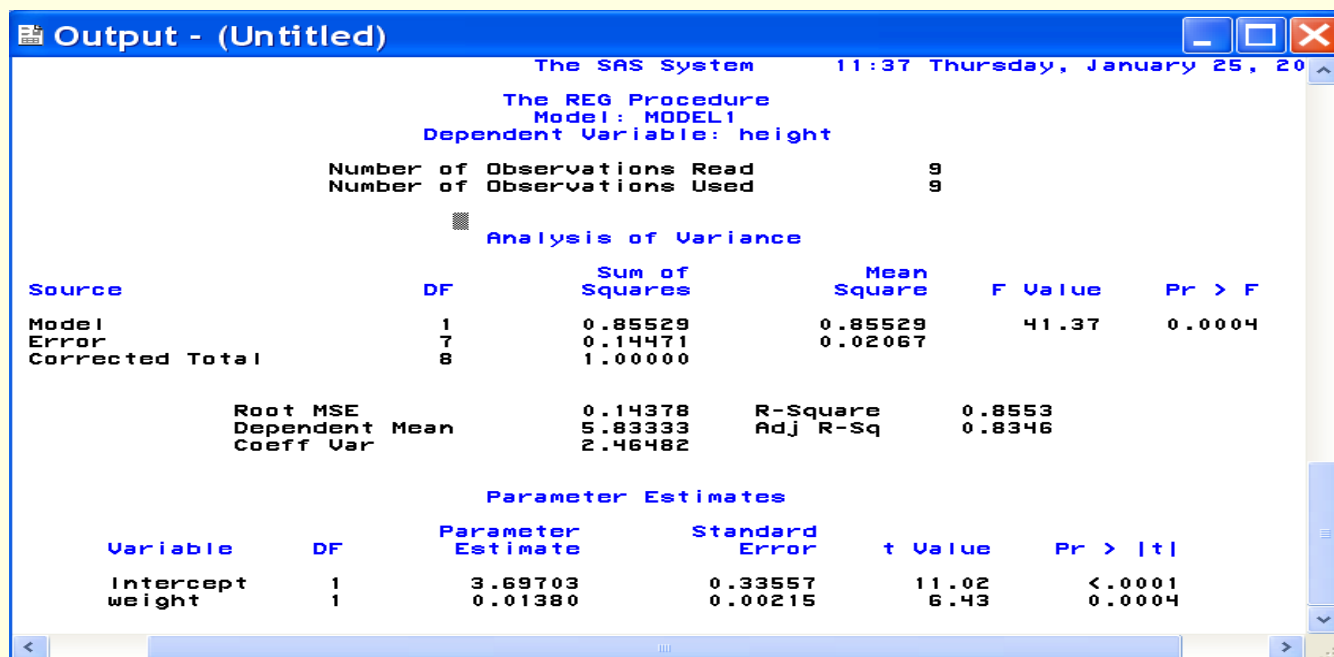
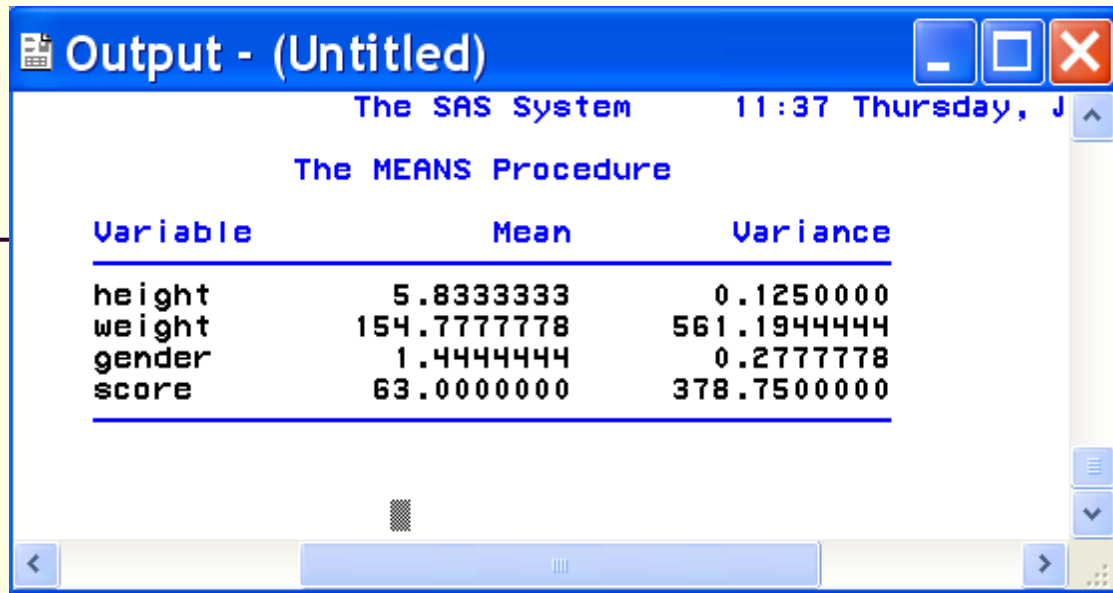
# Exercise

- Import the following data and use the libname statement to save the data to your 'c:/' drive of the computer.
- Use SAS to determine the mean and variance of 'height' and 'score' of the data.
- Determine the intercept (b1) and the coefficient (b2) of the model,  
$$\text{height} = b1 + b2 * \text{weight} + e$$
using the data.

height	weight	gender	name	score
5.4	125	2	JAUNITA	65
5.3	122	2	SALLY	77
5.7	145	2	SABRINA	36
5.9	150	2	KATE	55
5.7	156	1	JOHN	84
6	170	1	MARK	56
6.4	200	1	ERIC	34
5.9	165	1	BRUCE	72
6.2	160	1	TOM	88

# Solution

```
libname test 'c:\';  
data test.scores3;  
input height weight gender name $ score;  
cards;  
5.4 125 2 JAUNITA 65  
5.3 122 2 SALLY 77  
5.7 145 2 SABRINA 36  
5.9 150 2 KATE 55  
5.7 156 1 JOHN 84  
6 170 1 MARK 56  
6.4 200 1 ERIC 34  
5.9 165 1 BRUCE 72  
6.2 160 1 TOM 88  
;  
run;  
proc means mean var data=test.scores3;  
run;  
proc reg data=test.scores3;  
model height = weight;  
run;
```





For further Questions:  
[kentaka@mail.uri.edu](mailto:kentaka@mail.uri.edu)