SAS Analyst Tutorial

First time users of SAS Analyst should read **SAS Analyst: General Guidelines** to acquaint themselves with the Analyst environment, terminology and commands.

The following tutorial provides instructions on how to use Analyst to perform basic data manipulations and analyses (using the height/weight data set).

I. Getting Started

1. Create the temporary data set:

 Using either Internet Explorer or Netscape 4.73 or higher, go to <u>http://www.umanitoba.ca/centres/mchpe/teaching/sasmanual/db_htwt.html</u>. Copy the SAS program and then paste it in the Program Editor of SAS.

NOTE: In order to view and analyze the entire dataset (N=18), delete the following lines from the program:

/* View the values of the 1st 10 observations */ proc print data=htwt(obs=10); /* Begin a PROC step */ run; /* End the PROC step */

NOTE: Users of Netscape 4.73 or higher: Delete the spaces at the start of each line of raw data so that they are aligned against the left margin. This ensures that the actual columns of data match column numbers identified in the INPUT line above.

• Submit the program: Click on the SUBMIT icon

×

at the top of the screen

2. Launch Analyst:

From the menu at the top of the screen, select **SOLUTIONS** \rightarrow **ANALYSIS** \rightarrow **ANALYST**

V SAS						_ & ×
Eile Edit View Tools Solutions Window Help						
- Analygis	•	3D ⊻isua	l Analysis			
Results Development and Programmin	1g 🕨	Analyst				
Beporting	•	Design of	f Experiments	3		
Accessories	•	Geograp	hic Informatio	n S <u>y</u> stem		ervations:
ASSIST		<u>G</u> uided D	lata Analysis			exes:
Desktop		Interactive	e Data Analy	sis		ervation Le
EIS / OLAP Application Builder	r	l <u>m</u> vestme	nt Analysis			ressed:
Data Set Type:	_	Market R	esearch			ted:
Laber.		Eroject M	anagement			
	En	Quality Im	provement.			
		Queueing	Simulation			
Data Set Page Siz Number of Data Se	e: t. Pager	Time Ser	ies <u>F</u> orecast	ing System		
First Data Page:	e roge	Time Ser	ies Viewer			
Max Obs per Page: Obs in First Data	Page :	18				
Number of Data Se	t Repair	rs: 0		-		
File Name: Belease Created:		C:\TE 8.000	MP\SAS Ter ONO	porary F	iles_	_TD222\htwt
Host Created:		HIN_N	T			
0	lphabet	ic List o	f Variable	s and At	tr ibut	tes
		Variable	Туре	Len	Pos	
	3	ace	Num	8	0	
	4	height	Nun	8	8	
	2	name	Char	10	24	
	5	weight	Nun	ė,	16	
						<u> 1</u>
B Results GEncloser Output - (Untitled)	(Untitled)	X	Program Ed	itor - htwt		
Start the Analyst application			C:WIN	NT\Profiles	Puth-	

3. Access the height/weight data set in Analyst:

From the menu at the top of the screen, select FILE \rightarrow OPEN BY SAS NAME \rightarrow WORK \rightarrow HTWT \rightarrow OK.

Open by SAS name			×
Library MAPS V7 C:\P SASHELP V7 ('C SASUSER V7 C:\W WORK V7 C:\T	Member HTTNT	OK Cancel Options Help	
Description:	<u> </u>		

* Note: The above window may appear slightly differently for users of SAS 8.

****Note:** At the start of every new session, the temporary data set must be re-created following the above steps.

The data set will look like:

Analyst: (nr Analyst: (nr New Project	ew project) ew project) lysis	Htwt	Graphs 3 Grause) (Browse) Aubrey Ron	Stenistics	Window	Height	weight	2 • •
Analyst: (ne New Project	ew project) _{Vsis}	- <u>Htwt</u>	(Browse)	N CL 1	820 (B) 820 (B) 41	Fau 🛋 ⊵ 👌	weight	2 • • •
Analyst: (no New Project	ew project) Iysis	- Htut	(Browse) name Aubrey Ron	tex M	age 41	height 74	weight	
New Project	ysis	- <u>Htut</u>	(Browse) name Aubrey Ron	sex M	age	height	weight	
New Project	lysis		Aubrey Ron	sex M	age 41	height 74	weight	
E Pitet Anal	ysis	1	Aubrey	M	41	74	MONUTE A TO	-
Hwi		2	Ron	14		/4	1/0	
199, 1994		3		M	42	68	166	
			Carl	M	32	70	155	
		4	Antonio	M	39	72	167	
		5	Deborah	F	30	66	124	
		6	Jacqueline	F	33	66	115	
		7	Helen	F	26	64	121	
		8	David	M	30	71	158	
		9	Jame:	M	53	72	175	
		10	Michael	M	32	69	143	
		11	Ruth	F	47	69	139	
		12	Joel	м	34	72	163	
		13	Donna	F	23	62	98	
		14	Roger	м	36	76	160	
		15	Yao	м		70	145	
		16	Elizabeth	F	31	67	135	
		17	Tin	м	29	71	176	
			Susan	F	28	65	131	

II. View the data: Analyst Procedures

Example 1: Listing all variables and all values

You can display the values for any of the variables and for any number of observations in the SAS data set.

1. From the menu at the top of the screen, select **REPORTS** → **LIST DATA**. The following dialogue box appears:

List Data: Htwt				×
Candidate Box				
-	Print		Id	
C name] [UK
C sex				Cance 1
height				Reset
weight				Save Options
				Help
Reaove				
		_		1
		Options	Titles	Variables

- 2. The **Candidate Box** lists all the variables in the data set. Select the variables you want to view and then click on **PRINT**. Note: if you accidentally select the wrong variable, just highlight it in the **Print Box** and click on **REMOVE**.
- 3. Click on **OPTIONS** ensure that the following options are checked: *Column names*, *Single*, and *observation number*.
- Click on TITLES these comprise the top three lines at the top of the output. Type in the titles one on each line: View The Data: Example 1 Variable List
- 5. Click **OK** when finished.

NB: The values of the variables identified in the **ID BOX** (ID variables) are used instead of observation numbers to identify the observations in the listing.

- leave blank for this example.

🐺 SAS													_ 8 ×
<u>File E</u> o	dit <u>V</u> iew	<u>T</u> ool	s <u>D</u> ata	<u>R</u> epo	ts <u>G</u> raµ	ohs	<u>S</u> tatistic	∷s <u>W</u> indow	<u>H</u> elp				
				-	🗅 🖻		a	19 90 (m)	🛍 🔝	🙅 😿 🎫	p=05 📈 🍭		
, Besults		×I	BB ALL			-0							
	l:	=1	ar Anaiy	/st: (ne	w projec	cŋ.							
📑 🖓 🖓	ults	- 11	New	Proiect			<u>Htwt (</u>	Browse)					
	ontents	- 11	⊥					name	sex	age	height	weight	-
	rint	- 11	[-] [_] H	twt Analy	SIS		1	Aubrey	M	41	74	170	
		- 11	-	Htwt			2	Ron	M	42	68	166	
		- 11		- List Detr			3	Carl	м	32	70	155	
		- 11			, 		4	Antonio	M	39	72	167	
		- 11		- 🔡 Listir	ng		5	Deborah	F	30	66	124	
		- 11		- 🗐 Code	<u> </u>		6	Jacqueline	F	33	66	115	
		- 11						Helen	F	26	54	121	
		- 11	— Ŀ 🌽 н	twt Analy	sis		8	David	M	30	71	158	
		- 11	- 1	Htwt			- 9	James	M	53	72	1/5	
		- 11		Di int Date			11	Duth	E E	32	63	143	
		- 11			3		12		Г М	47	72	153	
		- 11		- 🛒 Code	,		12	Doppa	F	22	62	99	
		- 11		List Date	•		14	Boger	M	36	75	160	
		- 11					15	Yao	M		70	145	
		- 11		- 🔡 Listir	ng		16	Flizabeth	F	31	67	135	
		- 11		- 🖃 Code	,		17	Tim	M	29	71	176	
		- 11		<u> </u>			18	Susan	F	28	65	131	
		- 11]					
		- 11											
		- 11											
		- 11				Ŀ	14						
			<u> </u>			-							
- Result	s		🖪 Outpu	ut - (Untit	led)	Ē	Log - (Un	ititled)	Pri	ogram Editor	'- htw 💱 Ar	nalyst: (nev	w pr
											T\Profiles\Rutl	h	

* To view the actual SAS code used to generate the output, double-click on **CODE** in the procedure tree next to the data listing,

Each new procedure performed is added to the tree. You can view any one at any time by right-clicking on one and choosing **VIEW** or **OPEN**.

Example 2: Illustrating Viewing Options

1. Change mode from Browse to Edit: From the top of the screen, select EDIT \rightarrow MODE \rightarrow EDIT.

This allows you to manipulate the data and analyze them in various ways.

le i	Edit ⊻iew	Tools	Data	Beports	Graphs	Statistics	Window	Help		
1	inpert Calur		•	-	🗅 🚅 🖬 🖉	5 D I	a 🗗	🐃 🔺 🔄 j	X 🖸 🛤 🐱	۲
	Add Bows									
^	Duslicate				(m.).					
1	Delete			Htut	(Browse)					
Ы					name	1090	age	height	weight	
	condition on			1	Aubrey	M	41	/4	1/0	
	Mode		•	 Browse 	3	14 H	42	50	100	
				Edit		C -	36	70	155	
						F	30	F6	124	
			- L	Shared	Edit	F	33	55	115	
				7	Helen	F	26	64	121	
				8	David	м	30	71	158	
				9	James	M	53	72	175	
				10	Michael	M	32	69	143	
				11	Buth	F	47	69	139	
				12	Joel	м	34	72	163	
				13	Donna	F	23	62	98	
				14	Roger	м	36	75	160	
				15	Yao	м		70	145	
				16	Elizabeth	F	31	67	135	
				17	Im	M	29	/1	176	
				18	Susan	P	28	65	131	

2. Add labels to the variable names:

More descriptive names (called labels) can be added for each variable. In this example, you will add labels to four variables as follows:

Variable Name	Label
Name	Name of Student
Weight	Weight in Pounds
Sex	Gender of Student
Age	Age of Student

• Right-click on the desired column and select **PROPERTIES**. In the dialogue window that pops up, type in the appropriate label. Click **OK**. Do this for each of the four variables.

SAS			- 8 3
Eile ⊻iew <u>T</u> ools <u>S</u> olutions <u>W</u> in	dow <u>H</u> elp		
~	. D 🛎 🖬 🚳 🗅 🞼 🖞	8 🖨 🐃 🔺 🖄 🕱 🗄	🖸 🚧 🔟 😔
Column Properties			X
Colunn		_	
Nane: nane		OK	1 70
Label: Name of Student	1	Cancel	66
		Pacet	
Type: JC	Uppercase all values	- Neset	- 6
Length: 10		Help	hs
Fornat: \$10.	i ⇒ I		21
Informatic Elle			25
Informat: joio.	-		43
			29
	12 000 P	34 72 22 62	163
	14 Boost M	36 75	160
	15 Yao M	. 70	145
1	16 Elizabeth F	31 67	135
	17 Tim M	29 71	176
	18 Susan F	28 65	131
, کے	1		لتى.
🗋 Output - (Untitled) 📃 Log - (L	Intified) 🚺 Program Edi	tor-htwt 🍑 Column I	Properties

3. Work with only the first ten observations:

• Select rows 11-18: Click on the button that identifies row 11. Hold down the shift key and click on the button that identifies row 18. From the top of the screen, select Edit → DELETE.

4. Print a list of the data values for sex, age and weight.

You can print your data in a listing report. You can also specify the variables to be included in the report and some details about the report format.

- From the menu at the top of the screen, select **REPORTS** → **LIST DATA**. In the dialogue window that pops up, select *sex, age*, and *weight* in the **Candidate Box** and click on **PRINT**.
- 5. Identify records by name rather than by observation number:
 - Select *name* in the Candidate Box and click on ID.
- 6. Sum the values for the *weight* variable:
 - Click on OPTIONS → SUM tab → WEIGHT → OK.

Elle View Tools Solutions	Window Help	. 8 ×
_ √ [🔄 🗅 🚅 🖬 🖨 🗅 12 📴 🖻 🐃 🔺 🗠 X 🖾 🚧 🖉	
List Data: Options		
General Sum	Sum OK	ļ
	Help n2	
Resove		
	Y	
1	15 Yao M . 70 145	
	16 Elizabeth F 31 67 135	
	18 Susan F 28 65 131	
	2	-
2 X	Г Ш .	
🕒 Output - (Untitled)	og - (Untitled) 🕼 Program Editor - htwt 🎜 List Data: Options	
	C:\WINNT\Profiles\Puth-	

- 7. Display sex with values of male and female rather than with M and F:
 - From the menu at the top of the screen select DATA → TRANSFORM → RECODE VALUES
 - In the dialogue window that pops up, specify the variable to be recoded by clicking in the **COLUMN TO BE RECODED** line and selecting **SEX** from the list of variables. Since **sex** is a character variable, select this option under **COLUMN TYPE**.

Column to recode: sex veight New column name: : weight height age : : sex 170 height age : sex 165 sex 165 iname : iname :	weight 170		V				macon. ri	values init	
age sex name 155 Cance I Help 167 115 115 7 Helen F 26 64 121 8 David M 30 71 158 9 James M 53 72 175 10 Michael M 32 63 143	100			cter	@ Chara	jht ht	sex (weigh , heigh	o recode: nn name: nn type:	column to lev column lev column
7 Helen F 26 64 121 8 David M 30 71 158 9 James M 53 72 175 10 Michael M 32 69 143	155 167 124 115		He1p	ince l	C4	8	age sex name		
B David N 30 71 158 9 James M 53 72 175 10 Michael M 32 69 143	64 121	64	26	F	Helen	7		Code	- <u>B-</u> 100
10 Michael M 32 69 113	71 150	71	50	PE LA	Lawer	- 0			
11 D.A. F 47 00 110	69 143	69	32	M	Michael	10			
4/ 53 L 9	69 139	69	47	F	Buth	11			
12 Joel M 34 72 163	72 163	72	34	M	Joel	12			
13 Donna F 23 62 98	62 98	62	23	F	Donna	13			
14 Roger M 36 75 160	75 160	75	36	M	Roger	14			
15 Yao M	70 145	70		м	Yao	15			
16 Elizabeth F 31 67 135	67 135	67	31	F	Elizabeth	16			
17 Tin M 29 71 176	71 176	71	29	м	Tin	17			
18 Sutan F 28 65 131	65 131	65	28	F	Susan	18			

• In the new dialogue box that opens, type in *female* and *male* beside F and M, respectively. Click **OK**. This will add a new column in the data listing entitled Sex_recoded.

🐺 SAS	
<u>Eile ⊻iew Tools S</u> olu	tions <u>W</u> indow <u>H</u> elp
1	- D 🚅 🖬 🖽 D 🕼 🔐 B 🖘 🔺 🗠 X 🖸 🚧 🖉 🔗
Bacada Valuas: Htt	
Precode Values, rive	
Enter a new column	value to correspond to each original
nissing value will	be used.
Original Value	New Value (Character)
F	Female Cancel
M	Male Beset
	нетр
T CONTRACTOR OF CONTRACTOR	
_	
	18 Susan F 28 65 131
	2
<u>x</u>	
Dutput - (Untitled)	E Log - (Untitled) Recode Values:
	C WARNING Doctory
	C. (William I Promission

Note:

If you want to change the name of the new column, right-click on new column and select **PROPERTIES.** In the dialogue window that pops up, replace the column name with the new name of your choice (e.g. *sex2*). Click **OK**

If you want to change the order of the columns, right-click on any column title and select **MOVE** from the pull-down menu. In the dialogue window that pops up, highlight the variable you wish to move and click on the up or down arrows as desired. As you click on the arrow, the variable will move up or down the list, as appropriate. Click **OK** when finished re-ordering the columns.

Example 3: Sorting the Data

These examples show you how to sort the data and then create a listing of the values of specified variables.

- 1. Sort the data:
 - Right-click on the column to be sorted (in this case, *name*) and select **SORT** from the pull-down menu.
 - If it is not selected already, highlight *name* in the **Candidate Box** and click on **SORT BY**. Click on the **ASCENDING/DESCENDING** button to choose the sort order. Click **OK** when finished.

The data in the data list will now be listed in alphabetical order of name.

🐺 SA	s										_ # ×
Eile	<u>E</u> dit ⊻iew	Tools	<u>D</u> ata	Beports	Graphs	Statistic	s <u>W</u> indow	Help			
1				•) 🖻 🗖	AS D	ାର ଜଣ୍ଡ ସ	5 X 10	X 10 per	🗵 📣	
		_	_				** UP 14		A 🗆		
1 A	nalyst: (new	project))								
1.0	New Project		-	Htut	(Edit)						
1	new Project				name	189	age	height	weight	1842	14
	Htwt Analysis			1	Antonio	м	39	72	167	Male	
	Hwt			2	Aubrey	M	41	74	170	Male	
				3	Carl	M	32	70	155	Male	
	- J _ List Data			-4	David	M	30	71	158	Male	
	- 🖾 Listing			5	Deborah	F	30	66	124	Female	
	- Phone			6	Donna	F	23	62	98	Female	
	E Code			7	Elizabeth	F	31	67	135	Female	
1	E Pecode Va	alues		8	Helen	F	26	64	121	Female	
	- Pl Code			9	Jacqueline	F	33	66	115	Female	
				10	James	M	53	72	175	Male	
					Joel	M	34	72	163	Male	
				12	Michael	M	32	63	143	Male	
				13	Roger	M	36	75	160	Male	
				14	Hon	M	42	68	166	Male	
				15	Ruth	F	47	69	139	Female	
				16	Susan	P	28	65	131	Female	
				1/	Im	M	29	/1	1/6	Male	
				18	Yao	м		/0	145	Male	
											-
			1.1								-
1				10							2
_											
<u> </u>)utput - (Untitle	d)	E Log	- (Untified		Progra	m Editor - hh	ML. 🤡 Ana	lyst: (new p	ro	
								30	(WINNT\Pro	files\Ruth-	

• To sort by two variables, in the data listing, highlight both columns, right-click on them and select **SORT**. Follow the above instructions to sort them.

2. List Data:

- Generate a listing report for sex and age and identify each observation by name: From the menu at the top of the screen, select REPORTS → LIST DATA
- In the dialogue window that pops up, select *sex* and *age* from the **Candidate Box** and click on **PRINT**.
- Then select *name* from the **Candidate Box** and click on **ID**.

- Add titles to the report: Click on the TITLES button. Type in the following titles, one per line: View the Data: Example 3 Data List for Sex and Age Sorted by Name Click OK when finished.
- The output should look like:

View the Data: Example Data List for sex and ag	3 ge 10 Sorted by I	4 54 Tuesday, August 22, 2000 NAME
name	sex	age
Antoni o	М	39
Aubrey	М	41
Carl	М	32
Davi d	Μ	30
Deborah	F	30
Donna	F	23
El i zabeth	F	31
Hel en	F	26
Jacquel i ne	F	33
James	М	53
Joel	М	34
Mi chael	М	32
Roger	М	36
Ron	М	42
Ruth	F	47
Susan	F	28
Tim	М	29
Yao	Μ	

* Note: This does not need to be saved

III. Explore the Data: Statistics for Numeric Data

Example 1: Calculating Means

* If starting a brand new session, access the height/weight data set in Analyst:

 From the menu at the top of the screen, select FILE → OPEN BY SAS NAME → WORK → HTWT → OK.

1. Generate descriptive statistics:

• From the menu at the top of the screen, select STATISTICS → DESCRIPTIVE → SUMMARY STATISTICS

🐺 SA	s												. đ >
Eile	Edit	⊻iew	Tool	s <u>D</u> ata	<u>Reports</u>	Graphs	Statist	ics Window	Help				
1					•) 🚅 🗆	Det	criptive	•	Summary S	Statistics		
A Property of		_					Tat	le Analysis		Distribution	s		le les
ALCON A	8		-	🐮 Analy	rst: (new p	project)	Hva	onthesis Tests		Correlation	8		. 🗆 ×
PP F	Results		- 11	20 Name	Deplact	-	ANI	ηνα		Frequency	Counts		
Đ-1	🖥 Con	tents: Th	he i	218 HOW	Project.		Dev.	version	• T		neight	Height	É.
			- 11	📄 🖻 🎾 Ht	wt Analysis		Det	gression		41	74	170	
			- 11	15	Hiwit		Mu	tivanate		42	68	166	
			- 11				Sur	vivel	•	32	70	155	
			- 11				Sar	nple Size	•	39	72	167	
			- 11				Inde	3X		30	66	124	
			- 11				Б	Jacqueine	F	33	66	115	
			- 11				7	Helen	F	26	64	121	
			- 11				8	David	м	30	71	158	
			- 11				9	James	м	53	72	175	
			- 11				10	Michael	м	32	69	143	
			- 11				11	Ruth	F	47	69	139	
			- 11				12	Joel	м	34	72	163	
			- 11				13	Donna	F	23	62	98	
			- 11				14	Roger	м	36	75	160	
			- 11				15	Yao	м		70	145	
			- 11				16	Elizabeth	F	31	67	135	
			- 11				17	Tin	м	29	71	176	
			- 11				10	Susan	F	28	65	131	
				×			u.						ل لح
E Re	esuits	🔍 Explo	<u>ы</u>	🖸 Outpu	rt - (Unfitled) [6) Log - (U	Initiled)	🕅 Pro	ogram Editor -	htw 👉 An	alyst (new	/ pr
Comp	ute sur	mmary s	tetistic	08							Profiles\Ruth		

• From the **Candidate Box**, select the variables for which you wish to generate statistics (*age, weight*, and *height*) and click on **ANALYSIS**

🐺 SAS								. 8 ×
Eile ⊻iew Tools §	Solutions Window Help							
1		1.45 D	12 Se (-	5 56 1 6	10 V 10 mil 1	× 📣		
) *		а ша щ	+z 2∏ti :	1	C A D ** 1	~ ~		
Hesuits 💌	Summary Statistics:	Htwt				×		
Results	ſ							
🗄 🐻 Contents: The :			Analusis	-		F	meight	Ĩ.
			indiyara		OK		170	
	Cnane				Coursel	-	166	
	age				Lancel		155	
	height				Reset		167	
	weight				Save Dotion	s	124	
			Class .		auto operation		115	
				_	Help		121	
		7		1			158	
							175	
	Renove	Qtat is	tion	Plote	Dutout		143	
		Paura	Dete 1	r i ti ti ann	Unsighter		163	
		29/6	uata I	Itles	Variables		98	
							160	
		15	Yao	м		70	145	
		16	Elizabeth	F	31	67	135	
		17	Tin	м	29	71	176	
		18	Susan	F	28	65	131	
	× ×	10						<u>ا</u> ال
4 F								
🛱 Results 🖳 Explorer	Output - (Untitled)	Log - (Ur	nitled)	A Pro	ogram Editor - htw	🛹 Su	immary Sta	tisti
					C:\WINNT\Profi	iles\Rut	in l	

• Click on the **STATISTICS** button to select the statistics you want generated.

The basic statistics most often chosen are: *mean, standard deviation, minimum, maximum, and number of observations.* Click **OK** when finished.

SAS								_ 8
ile ⊻iew <u>T</u> ools <u>S</u> olutions <u>W</u> i	ndaw <u>H</u>	jelp						
~	• C) 🖙 🖬 é	12 Da 12	k 🖻 🐐	🔺 🖄 🛣		š 😣	
ummary Statistics: Htwt								
Summary Statistics: Statistic	9					×		1
í literatura de la companya de la co							sex2	
							Male	
Print statistics for anal	ysis va	riables_			OK	- 1 2	Male	- 11
Standard deviation	Numbe Numbe	r of obse r of niss	sing valu	es	Cance I		Male	- 8
Standard error	Skewn	688			Basat	- 1	Fenale	- 11
I Minimun	Stude	nt's t			neset		Fenale	- 11
Maximun Passa	Proba	bility of	ft		Help		Fenale	
□ Sun	Corre	cted sum	of squar	88			Fenale	
Med i an	Uncor	rected su	un of squ	ares			Fenale	- 11
							Male	- 11
						1	Male	- 11
							Male	- 11
					75	160	Male	- 11
	45	10.4		17	68	166	Male	- 11
🔁 🍃 List Data	15	Huth	F	47	69	139	Fenale	- 11
EPIL inting	15	Susan	P	28	50	131	Female	- 11
En closely	18	Yan	M	23	70	145	Male	- 11
L BU Code		1100	14		10	145	Halo	-
🖻 🍃 Summary Statistics								
ITA Summary Statistics of little								
TTT Opening A Digeorge of Lifest								
Code								
- 🖻 Code 🗾								1
Dutput - (Untitled)	Untified)		Listing		🎺 Summa	ry Statist	i	1

- 2. Add titles to the output:
 - Click on the **TITLES** button and type in the titles, one per line: View the data: Example 1 Calculating Means: Age, Weight & Height

Click **OK** when finished.

3. Click OK.

The output should look like:

		Data Evenuela 1						
	View the	e Data: Example 1	10 5			4		
	Data Lis	t for sex and age	10: 5	4 Tuesday	, August	22, 2000		
		0						
		So	rted by NA	ME				
				_				
		name	sex	č	ige			
		Antoni o	М		39			
		Aubrev	М		41			
		Carl	М		32			
		Davi d	Μ		30			
		Deborah	F		30			
		Donna	F		23			
		Elizabeth	F		31			
		Hel en	F		26			
		Jacquel i ne	F		33			
		James	Μ		53			
		Joel	Μ		34			
		Mi chael	Μ		32			
		Roger	Μ		36			
		Ron	Μ		42			
		Ruth	F		47			
		Susan	F		28			
		Tim	Μ		29			
		Yao	Μ					
		View th	e Data: Ex	ample 1				5
		Calculating Mea	ns: Age, W	eight & ⊦	lei ght			
		_	-	-	10: 54	Tuesday,	August 22,	2000
		The M	EANS Proce	dure				
Vari abl e	Ν	Mean	Std D	ev	Mi ni mum		Maximum	
ſſſſſſ	ffffffff.	ffffffffffffffffffff	fffffffff	ffffffff	fffffffff.	ffffffff	ſſſſſſ	
age	17	34. 4705882	7.76303	46 2	3. 0000000	53.	000000	
hei ght	18	69.0555556	3. 52256	96 6	2.0000000	75.	000000	
wei ght	18	146. 7222222	22. 54095	76 9	8. 0000000	176.	000000	
fffffff	<i>ffffffff</i> f	ſſſſſſſſſſſſſſſ	ffffffff	fffffffff	ffffffff	ffffffff	ffffff	

Example 2: Calculating Means: CLASS Option

This example demonstrates how to generate descriptive statistics with other options.

* If starting a brand new session, access the height/weight data set in Analyst:

From the menu at the top of the screen, select FILE \rightarrow OPEN BY SAS NAME \rightarrow WORK \rightarrow HTWT \rightarrow OK.

1. Generate descriptive statistics:

- From the menu at the top of the screen, select STATISTICS → DESCRIPTIVE → SUMMARY STATISTICS
- If there are any variables listed in the **ANALYSIS** box, remove them (highlight them and click the **REMOVE** button)

- The analysis in this example applies only to the *age* variable, so it should be selected as the only **ANALYSIS** variable.
- The analysis will be separated by values of **sex** so this should be the chosen as the **CLASS** variable.

🐺 SAS						- 8 :
Eile ⊻iew Io	ols <u>S</u> olutions <u>W</u> inde	aw <u>H</u> elp				
~		10 🚔 🖬 🚳	0 19 0 8 (a	1 🐜 🔺 🔄 🕱 🛛	🗉 🕶 🛃 🔗	1
Bes to I						
	Summary Statis	lics: Hiwt			×	
Presults						
E 🐻 Contents		final;	2818			the
	Corre			ак	74	170
	height	age		Cance 1	68	166
	weight			Barrat	70	155
				heset	72	167
				Save Options	R 66	124
		Ela	8.8	Halo	66	115
		sex		neip	64	121
		I I	2		71	158
					72	175
	Renove	Phatinting	Plata	Dutmut	69	143
			Plots	ourepute	69	139
		Save Data	Titles	Variables	72	163
					62	98
					75	160
		15	Yao N	l .	70	145
		16	Elizabeth F	31	67	135
		17	Tim M	29	71	176
		18	Susan F	28	65	131
		-1				
1 1 1		E				
🚰 Res. 🔍 E	🖸 Output - (Untitled)	🖾 Log - (Untilled)	Program	Editor	mary Statist	🥩 Summary Sta
				C:(WIN	INT\Profiles\P	uths

• To select the statistics you want to generate click on the STATISTICS button.

The basic statistics are: Mean, Minimum, Maximum, Number of observations, Number of missing values and sum. Click **OK** when finished.

• Specify the number of decimal places: click on the **OUTPUT** button. Type **1** for number of decimals. Click **OK**.

V SAS							_ # ×
Eile ⊻iew <u>T</u> oo	ils <u>S</u> olutions <u>W</u> ind	law <u>H</u> elp					
×		- 🗅 🖨 🖬	a D. 19 []	8 (a) 🎭	🔺 🔄 🛣	🔛 priš 🛃	۰
Results 💌	Summary Statistic	sc. Hited					
- Deculte	Summary S	tatistics: Output			×		
Contents	i i i i i i i i i i i i i i i i i i i						
Coments							weight
					nik I	24	170
	Dutput a	pearance				68	166
	E sald as		_	Car	ncel	F 12	155
	Field wi	ath:		Be	set	H 12	167
	Number o	f decinals:	1			8 60	124
	C Pr in	t variable la		H	elp	66	115
		c variable la	010				121
						71	138
						12	1/5
						63	143
		Saure Da	ata Titl	ee U.	e lables	- 63	1/2
		0070 10		00 75	0.000.000	12	163
						52	36
		15	V.			20	145
		10	Tao Electroth	F		70	140
		17	E szabelni Tim	P N	31	21	130
		10	1 m	F	29	/1	176
		10	Susan	r	20	60	131
		A					
<u> </u>							
📅 Res. 🔍 E	Dutput - (Untitled	📋 Log - (Untifie	ed) 🕅 Pro	igram Edi	tor 🖸 Su	mmery Statist.	Summary Sta
Save and close wir	ndow				DC:W	INNT/Profiles	Ruths

2. Add titles to the output report: click on the **TITLES** button. Type in the titles, one per line:

Example 2: Generating Means Use of CLASS option Classed by Gender

Click OK

3. Click OK.

Output should look like:

SAS -	- [Sumr Edit	nary Stat Help	istics of	htwt]			
~					谷道 古田橋・	ः 🗊 🖫 🚉 🗴	0 😣
Γ				Example 2 Use of Class The Mi Analysis	: Generating Means CLASS option sed by Gender EANS Procedure & Variable : age	10:52 Thursday	2, September 7, 2000
	sex	N Obs	N	Nean	Std Dev	Niniewe	Nax inum
L	F	7	7	31.1	7.7	23.0	47.0
	н	11	10	36.8	7.3	29.0	53.0
							D
🖸 Outp	ut - (Unt	itled)	🖹 Log -	(Untitled)	ditor - Untitled1 🥼	Analyst Test_proje	ct 🔝 Summary Statist
Start	23 Pe	gasus Ma	ii [Wicrosoft Word - S.	. Tahoo! - Netscape	8 Jas SAS - [Sum	mar 02168 10:58

Example 3: Generating Means: BY option

In this example, the data will first be sorted by sex.

1. In order to edit the data set, the edit mode must be changed from BROWSE to EDIT. From the menu at the top of the screen, select EDIT → MODE → EDIT.

Note that the current mode appears in parentheses beside the name of the data set.

١

🐺 SAS -	[Analyst (n	ew proje	ect)]														_ 8 ×
💱 <u>E</u> le	Edit View	Tools	Data	Bepo	rts (Grapt	is <u>Stati</u>	stics	. <u>₩</u> in	dow	r <u>H</u> e	lp					_ & ×
~	Insert <u>C</u> olu	imnis	•	• I [) 🦻		a 🗅	12 1	la a	5	- 1 X	23	ΧE	1-16	<u>k</u> (9	
Results	Add Bows	3		-10	-	Htut	(Edit.	1		_	_		_				1
🔐 Resi	Duglisste				1 1			ane	20	x	age	1	he	idht		ecidht	
🗄 🕞 C	Delete				111	1	Aubre	,	м			41		74	4	170	
	Beneme					2	Ron		м			42		68	3	166	
	1-35jitesiites					3	Col.,		м			32		70	0	155	
	Mode		•	Brow	vse			0	м			39		77	2	167	
1.1.1		🖻 🗋 Sur	tinary :	- Edit				sh	F			30		68	6	124	
				_				eline	F			33		68	6	115	
				Shar	red E	fit			F			26		6	1	121	
			_			8	David		м			- 30		7	1	158	
						9	Janeo		м			53		77	2	175	
						10	Micha	el	м			32		68	9	143	
						11	Ruth		F			47		68	9	139	
						12	Joel		м			34		77	2	163	
						13	Donne		F			23		6	2	98	
						14	Roger		м			36		75	5	160	
						15	Yao		м					70)	145	
						16	Elizab	sh 🛛	F			31		67	7	135	
						17	Tim		M			29		7	1	175	
						18	Susan		F			28		62	5	131	
	1				, -	1											ل نې
🗗 Res. [<u>ع</u> د 🗅	Output - (Untitled	ŋ	Ē	.og - (Untitled)		ľ] Pri	ogram	Edito	r - ht	wt	a An	alyst: (nev	v pro
Access d	ata in edit mod	le e									1	⇒ C:(WIN	NT/Pro	ofiles¥	Ruths	

2. To sort analyses by sex, right click on the top of the sex column and choose **SORT** from the popdown menu. Ensure that it will be sorted in ascending order. Click **OK**:

File View To	ools Solutions Window	Help			_ 6
) 🖻 🖬 🚳 🗅 🕼 🕅	i i i i i i i i i i i i i i i i i i i	🖂 🖻 🔣 🔗	
esults Results E Contents	C name age height weight	(A) sex		OK Cancel Reset Help	X 170 166 155 167 124 115 121 158 175 143 128
	Beaove	Ascend/Desce	nd		163
	Renove	Ascend/Desce		70	163 98 160
	Renove	Ascend/Desce	M .	70 67	163 98 160 145
	Renove	Ascend/Desce	M . F 31 M 29	70 67 71	163 163 160 145 135 176
		15 Yao 16 Eksabeth 17 Tim 18 Susan	M . F 31 M 29 F 28	70 67 71 65	163 98 160 145 135 176 131
		Ascend/Desce	M . F 31 M 29 F 28	20 67 71 65	163 98 160 145 135 176 131
Res. Q. E.	Benove	Ascend/Desce	M . F 31 M 29 F 28	20 67 21 65	163 98 160 145 135 176 131

In the list of data, all females should now appear at the beginning.

SAS - [Analyst: (new proje	ct)]		-				_ 8 ×
<u>∛ E</u> ile <u>E</u> dit ⊻iew <u>T</u> ools	Data Beports	s <u>G</u> raphs	Statistics	Window	w <u>H</u> elp		X
~	· D	൙ 🖬 🍏	D. 12 2	8 (a) 4	i 🔺 ⊵ 🕱	🔛 PH 🛃 (
	al brut ((u+)					1
New Project			[] ·		halahé	unista	
E phitwt Analysis		Deborah	Local a	20	nagni CC	wegni 124	1
Land Land	2	Lacquellee	E	22	23	115	
- See Introd	2	Lalao	E .	26	60	121	
Table Analysis		Puth	E .	47	+0 03	129	
		Doppa	F	22	63	90	
	6	Fizabeth	F		67	135	
	7	Susan	F	28	65	131	
	8	Autoreu	M	41	74	120	
	9	Bon	M	42	68	186	
	10	Carl	M	32	70	155	
	11	Antonio	M	39	72	167	
	12	David	M	30	71	158	
	13	James	M	53	72	175	
	14	Michael	M	32	69	143	
	15	Joel	M	34	72	163	
	16	Boger	M	36	75	160	
	17	Yao	M		70	145	
	18	Tim	M	29	71	176	
							_
	_						
1	<u>ار ک</u>						<u>ل</u> ے.
		1			_		
🖸 Output - (Untitled)	.og - (Untified)	2 A 12	nalyst (n	ew pro			

To separate output by sex:

- 3. From the menu at the top of the screen, select STATISTICS → DESCRIPTIVE → SUMMARY STATISTICS. Clear the ANALYSIS and CLASS boxes by highlighting the variable that appears in them and clicking on the REMOVE button
- 4. Select *age* as the **ANALYSIS** variable.
- 5. Click on the **VARIABLES** button. Highlight the *sex* variable and click on the **BY GROUP** button. Click **OK** when finished.

														_ 6
Eile ⊻iew I(slac	Solutions	Window	/ <u>H</u> el	р									
~			•		2 🖬	- 65	D 12	ឿង (;	a 🐐	🔺 🔄	ΧE	pris 🛃 (9	
les la 🕷		Summan	Statistic		rinhla	-	-		_					
El Davilla	1	Summary	ordusiit	.s. va	nabie	3								_
Presuits														
Comence					BJ	1 Gro	pup	ι.		4e ight			w	
		C name] [sex			٦Г					<u> </u>	
		height						-			7	Can	cel	
		- weight							-			Rea	set	
									- Fr	equenc	y	He	10	
								-			-			
								_						
		Here	0.46											
	⊩				Save I	Data	T	tles	Va	riable	2	21	150	
	⊩				Save I	Data	Ti	tles	Va	riable	8	71	158	
					Save I	Data	Ti	tles	Va	riabla	2	71 72 69	158 175 143	
				1	Save I	Data	Joel	tles	Va M	riable 3	2	71 72 69 72	158 175 143 163	
				1	Save	Data 15	Joel Roger	tles	. Uq м м	r kable 3 3	8 4 6	71 72 69 72 75	158 175 143 163 160	
					Save	Data 15	Joel Roger Yao	tles		r kabla 3 3	z 4 6	71 72 69 72 75 70	158 175 143 163 160 145	
				-	Save 1	Data 15	Joel Roger Yao Tim	tles		r kabla 3 3	z 4 6 9	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
				2	Save	Data 15 16 17	Joel Roger Yao Tim	tles	M M M M	r kable 3 3	4 6 9	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
				2	5arve 1	Data 15	Joel Roger Yao Tim	tles	M M	r kab 1 a 3 3 2	4 6	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
				_	Save 1	Data 15	Joel Roger Yao Tim	tlez	- U а М	r kabilar 3 3 2	4 6	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
				-	Save	Data 15 16 17	Joel Roger Yao Tim	tles	- Uia М. М. М.	r kabilar 3 3 2	4 6	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
						Data 15 16 17	Joel Roger Yao Tim	tles	M M	r kabilar 3 3 2	z 4 9	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
4 						Data 15	Joel Roger Yao Tim	tles	0 0 M M M	r kabilar 3 3 2	4 6 9	71 72 69 72 75 70 71	158 175 143 163 160 145 176	
• ــــــــــــــــــــــــــــــــــــ) Output - (L	Intified)		1 1 1 1	0ata 16	Joel Roger Yao Tim	tles		r iab la 3 3 2 gram Ed	2 4 6 9	71 72 75 70 71	158 175 143 163 160 145 176 mmary Stat	isti

- 6. To change the title of the output report, click on the **TITLES** button. Up to three titles can be added one in each line:
 - Calculating Means: Example 3
 Sorted by Gender

Output will look like:

	AS - [Summary Sta Eile Edit <u>H</u> elp	tistics of Htwt)			5 B B *) 🗊 🔁 💁 🗶 🛈 d		. 8 > . 8 >
			Pr	oc Means: E Use of BY Sorted by	xample 3 option Gender	09:54 Tuesday, Augus	t 8, 2000	5 *
			1 Ana	he MEANS Pr lysis Varia	ocedure ble : age			
	Mean	Std Dev	N	N Miss	Minimum	Max invin	Sun	- 1
	31.1	7.7	7	0	23.0	47.0	218.0	- 8
	Nean	Std Dev	Ana	nlysis Varia N Miss	ble : age Minimum	Nextmun	Sun	-
	36.8	7.3	10	1	29.0	53.0	368.0	
	Output - (Untitt	🖹 Log - (Unifiled)		Program Edito.	Summerv	Stati	Summarv	*
÷.						2 · · · · · · · · · · · · ·		

7. To generate more statistics on the *age* variable, go to the menu at the top of the screen, select **STATISTICS** → **DESCRIPTIVE** → **SUMMARY STATISTICS**. Click on the **STATISTICS** button in this window. The typical statistics chosen are:

Mean, Standard deviation, Standard error, Variance, Minimum, Maximum, Range, Sum, Median, Number of observations, Skewness, Kurtosis, Student's t, Coefficient of variation, Corrected sum of squares, Uncorrected sum of squares

IV. Explore the Data: Creating tables

Calculating frequencies produces summary data in tabular form. Each of the following five examples show various options available with this procedure.

Example 1: Frequency Tables

- 1. To generate frequency counts for each variable, from the menu at the top of the screen, select STATISTICS → FREQUENCY COUNTS.
- 2. In Box 1, highlight all variables and then click on the **FREQUENCIES** button.

* To generate counts for only specific variables, highlight only those variables and then click on the FREQUENCIES button.

J SAS								. 8 ×
Eile View Tools So	olutions <u>W</u> indow <u>H</u> elp							
1	JORE	1.45 D	12 20		. Is 🐨 🖂 🕬	k 🔌		
) • I		а фу щ	+2 LL			<u> </u>		
Results 💌	Frequency Counts: Hte	rt 👘			2	C		
📅 Results								1
🗉 👼 Contents: The :		Erea	encie	. 1			uniald 1	- 1
🗷 👸 Print: The SAS	C		, criere		OK	66	124	
E 🐻 Contents:	C sex				Cance 1	66	115	
	age height				Reset	64	121	
	weight				Course Contribution	52	98	
					save uptions	57	135	
					Help	65	131	
				1		74	170	
						68	166	
	Promotion					70	155	
	L163m(0.0-8.	Inpu	it	Plots	Tables	72	167	
		Titl		Uariables		71	158	
				var rabitea	·]	72	175	
						69	143	
		15	Joel	м	34	72	163	
		16	Roger	м	36	75	160	
		17	Yao	м		70	145	
		18	Tin	м	29	71	176	
	L. 2		_					<u>ل</u> ے .
Besuits St Explorer	Output - (Untitled)	Log - (Ur	ntitled)	E F	Program Editor - htw.	. 🐼 Fre	quency Co	ount
					C(WINNT/Pro	files\Ruth-		

3. Click on the **TABLES** button to choose the counts you want displayed in the tables. Ensure that the first option is checked. Click **OK**.

🐺 SA	s												_ # ×
Eile	⊻iew	Tools	Solutions	Window	Help								
<hr/>				•	D 🦻	🖬 🚳 🖻	42 🗄	6 (B) 7	1	🖄 🛣 🖻 P	• 🗵 🧶		
Results		l l	Ere	quency C	ounts: T	ables						×	
B B B	esults	nts: The											F.
E-	🖗 Print	The SAS	i i c	nt	iee and	d nercent		aith a	and a loss	tiuno	OK		
•	Conte	ents:	l à	Frequen	cles and	d cunulat	ive fi	requen	cies	LIVES	Cance	1	
			i i	Frequen	ies on	ly ly	ages				Reset		
											Help		
								_			58	168	1
				Beauve	1						70	155	
				1101-10-10		(np	ut	Plo	te	Tables	72	167	
						Titl	1642	Variat	blez		71	158	
											72	175	
											69	143	
						15	Joel		м	34	72	163	
						16	Roger		м	36	75	160	
						17	Yao		м		70	145	
						18	Tin		м	29	71	176	
4				nut - (1 loit	en) [<u></u>	witled)		a Pm	arem Editor - I	tw Free	wency (لتر teuri
Me He	saute 1	Exploie		par-touro		E togs (0	(nued)		01110	gran Euror-r	and a rec	facincy c	ounc
									-	⇒C:\WINNT\	Profiles\Puth-		

4. To add a title to the table, click on the **TITLES** button. In the first line, type in: *Example 1: Frequency Table.* Click **OK** when finished.

The output should look like this:

Example 1: Frequency Table			6			
			1	4:04 Wednesday,	August 9	, 2000
	The	e FREQ Proce	dure			
			Cumul ati ve	Cumul ati ve		
name	Frequency	Percent	Frequency	Percent		
fffffffffff	ſſſſſſſſſſſ		ſſſſſſſſſſſ	ſſſſſſſſſſſſ		
Antoni o	1	5.56	1	5.56		
Aubrey	1	5.56	2	11. 11		
Carl	1	5.56	3	16.67		
Davi d	1	5.56	4	22.22		
Deborah	1	5.56	5	27.78		
Donna	1	5.56	6	33.33		
Elizabeth	1	5.56	7	38.89		
Hel en	1	5.56	8	44.44		
Jacquel i ne	1	5.56	9	50.00		
James	1	5.56	10	55.56		
Joel	1	5.56	11	61.11		
Mi chael	1	5.56	12	66.67		

28 Analyst / rs

Roger	1	5.56	13	72. 22		
Ron	1	5.56	14	77.78		
Ruth	1	5.56	15	83.33		
Susan	1	5.56	16	88.89		
Tim	1	5.56	17	94.44		
Yao	1	5.56	18	100.00		
-	_		Cumulative	Cumul ati ve		
sex H	-requency	Percent	Frequency	Percent		
	, , , ,		<i>, , ,</i>			
F	/	38.89	10	38.89		
IVI	11	01.11	10	100.00		
	Exam	ple 1: Frequ	ency Tabl e			7
				14:04 Wednesday,	, August 9, 20	000
	-					
		THE FREU Pro	cedul e			
			Cumulative	Cumul ati ve		
age F	requency	Percent	Frequency	Percent		
	ſ <i>ſſſſſſſſſ</i>	ſ <i>ſſſſſſſſſſ</i>	ſſſſſ	ſſſſſſſſſſſ		
23	1	5.88	1	5.88		
26	1	5.88	2	11.76		
28	1	5.88	3	17.65		
29	1	5.88	4	23.53		
30	2	11.76	6	35.29		
31	1	5.88	7	41.18		
32	2	11.76	9	52.94		
33	1	5.88	10	58.82		
34	1	5.88	11	64.71		
36	1	5.88	12	70.59		
39	1	5.88	13	76.47		
41	1	5.88	14	82.35		
42	1	5.88	15	88.24		
47	1	5.88	16	94.12		
53	1	5.88	17	100.00		
	Fra	equency Miss	ina = 1			
	110	equency wirss	· ···9 – ·			
			Cumulative	Cumulative		
hei aht	Frequency	Percent	Frequency	Percent		
ffffffff	ffffffffffff	fffffffffff	fffffffffffff	fffffffffff		
62	1	5. 56	1	5. 56		
64	1	5.56	2	11. 11		
65	1	5.56	- 3	16. 67		
66	2	11. 11	5	27.78		
67	- 1	5.56	6	33. 33		
68	1	5.56	7	38.89		
69	2	11. 11	9	50.00		
70	2	11. 11	11	61. 11		
71	2	11. 11	13	72.22		
72	3	16.67	16	88.89		
74	1	5.56	17	94.44		
75	1	5.56	18	100.00		

	Exam	ple 1: Freque	ency Table			8
			5	14:04 Wednesday,	August 9	, 2000
				5	0	
		The FREQ Prod	cedure			
			Cumul ati ve	Cumulative		
wei ght	Frequency	Percent	Frequency	Percent		
ffffff	ſſſſſſſſſſſſ		ſſſſſſſſſſſſ	ſſſſſſſſſſſ		
98	1	5.56	1	5.56		
115	1	5.56	2	11. 11		
121	1	5.56	3	16. 67		
124	1	5.56	4	22. 22		
131	1	5.56	5	27.78		
135	1	5.56	6	33. 33		
139	1	5.56	7	38.89		
143	1	5.56	8	44.44		
145	1	5.56	9	50.00		
155	1	5.56	10	55.56		
158	1	5.56	11	61.11		
160	1	5.56	12	66. 67		
163	1	5.56	13	72.22		
166	1	5.56	14	77.78		
167	1	5.56	15	83.33		
170	1	5.56	16	88.89		
175	1	5.56	17	94.44		
176	1	5.56	18	100.00		

Example 2: Creating 2-way tables

This example produces a two-way table (cross-tab) from a subset of the data.

- 1. Generate a cross-tab (*sex* BY *age*) only for records where *age* is between 0 and 29. From the menu at the top of the screen, select DATA → FILTER→ SUBSET DATA.
- In the dialogue window that pops up, select the *age* variable from the AVAILABLE COLUMNS. It automatically appears in the WHERE box of the window. In this box you will write a statement identifying the conditions that determine which records will be included (i.e. you specify: 0 ≤ *age* ≥ 29).
- 3. When you select *age*, a list of possible OPERATORS also appears. These operators help set the conditions for the *age* variable: EQ stands for Equal To

NE stands for Not Equal To GT stands for Greater Than and so on...

Because one record in the data set is missing the *age* variable, you must first specify that this record is not to be included. Select **OTHER Operators** to extend the list.

▼SAS Eile View Tools Solutions Window Help ▼ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■ ■	- 8 × •
Available Columns CONSTANT enter value> age height eight EO NE GT LT GE LE * j OTHER Operators	Undo Help Reset Cancel
Output - (Untitled)	sVRuths

4. From this extended list, select 'IS NOT MISSING'.

* **NOTE:** Use the **UNDO** button at the side of the window to delete/replace any incorrect selections you make.

- 5. In order to specify that *age* must be less/equal to 29, you must make a compound statement. Click on the **OPERATORS** button and select **AND**.
- 6. Select *AGE* from the **AVAILABLE COLUMNS**. Select **LE** (*Less or Equal to*) from the list of **OPERATORS**.

🐺 SAS	- [Subs	et: WO	RKTMP_									₫ ×
🐴 Ele	⊻iew	Tools	Solutions	Window	Help						-	l∉ ×
✓ [1 🖉 🖬	1 26 L & D	1 🛍 KA 🛛 🚺) 🔁 迫	* 🕚	۲		
An op	erator	nust f	ollow an	operand	_							
200	ISTANT	columns enter	ualue)		-	a					ov.	1
nam	8	circor	varue/		-	Operators				_	UK	1
age											Undo	1
vei	ght											
											He1p	1
										В	eset	1
										_		
										C	ancel	
age	ls Not	Missin	g AND age	• LE 29					1			
🖸 Out	put - (Un	itled)	E Log-	(Untified)	l l	Editor - Untitled1	📑 S	ubset: W	ORK			
							=	C:/WINN	T\Profiles\	Ruths		

7. To specify the maximum *age*, select '<LOOKUP distinct values>' from the AVAILABLE COLUMNS. This will open another window that lists all the values for *age* in the data set. Select 29. Click OK.

The resulting data listing should be as follows:

File	AS	Viau	Tools	Date	Re	norte	Get	mhe	Stal	lietine	- \A	lindo	ar k	Halo							_ 6	×
Lie	Ear	TIGM	Tonis	Dava	De	al I r	200		210		- <u>-</u>	te (=		Teih	les.	v			le.			
<u> </u>	<u> </u>				_	1) L			9	LQ, I	2 1	12 =			×	A		P-11	<u> </u>	۵.	 	_
27	Analyst	: (new	project)																			×
	New Pro	sject		-	ų.	htut	(Ed	it)	Mhen	e ag	e la	s No	t M	1881	ng (NND	age	LE	29		 	
ΠĨ	Chiana	Amahanin						nar	18	SEX		aga		hei	ght		Mei	ight _			÷	
	- TRWG	Anarysis				7	He	len		F			26		1	64		12	21			
	- 18 h	twt				13	Do	nna		F			23		-	62		5	98			
	_				1.	17	T.	n		м			29			71		17	76			
						18	Su	nan		F			28		1	65		13	31			
																					_	
<u> </u>						_																
	🗈 Output - (Untitled) 🔄 Log - (Untitled) 📝 Program Editor - (Unt 😥 Analyst: (new pro																					
																:\W	INN	∏/Pro	files	VPuth-		

 Now it's time to generate the two-way table. From the menu at the top of the screen, select STATISTICS → TABLE ANALYSIS. The window that pops up allows you to specify the layout of the table:

🐺 SAS			_ # ×
<u>F</u> ile ⊻iew	Tools Solutions	Window Help	
✓		🔄 🗅 🖙 🖬 🚳 🗅 19 📴 🖨 🐃 🔺 🗠 X 🖸 🚧 🔌	
Analyst (n	C name C sex age height weight	is: htwt Row Colunn OK Cancel Reset Save Options Help I I Input Select Tab Statistics Tables Titles Variables	
⊡ Output-	(Untitled)	og - (Untified) Program Editor - (Unt Program Editor - (Unt Program Editor - (Unt)	لتر

- 9. In this example, you want to create a table that plots *sex* in rows and and *age* in columns. Highlight *sex* and then click on **ROW**. Highlight *age* and click on **COLUMN**.
- 10. To choose the chi-square test, click on the **STATISTICS** button and check it off. Click **OK**.

W SAS	_ # ×
Eile View Tools Solutions Window Help	
	2 🗄 🖹 🕷 🗶 🖄 🏹 🛄 🚧 🖉
Table Analysis: Statistics	
Statistics Chi-square statistics Heasures of association Mantel-Haenszel statistics Heasures of agreement	0K 29
.Exact test	Reset 31
Exact test for (r x c) table	Help ionz
Note: Computational time may be prohibitive for certain data. The option must be set each time it is used.	
□ Print statistics only (no tables) □ Include missing values in calculations	t los Tes
Output - (Untitled)	Editor - (Unt
	C:\WINNT\Profiles\Puth-

11. To specify the frequency and percentage values you want in the table, click on the **TABLES** tab. Check off the desired values.

T SAS		_ 8 ×
E SAS		_ @ ×
, Eile View Tools Solutions Win	idaw <u>H</u> elp	
	🖸 🗅 😂 🖬 🚳 🖬 💱 📴 🗎	Tu 🔺 ⊵ 🕱 🛄 👐 🖉 🧶
Table Analysis: Tables		X
Frequencies Per Deserved Expected Deviation	centages OK Cell Row Cancel Column Reset Help	29 0K 21 58 75 31 e Options HeTp
Prenaryva	Input Select Tebles Title	Tab Statistics vs Variables
	1	تى
Cutput - (Untitled)	Untified) 🔀 Program Editor - (U	Jnt 🌌 Table Analysis:
Cancel selections and close	,	C(WINNT)Profiles)Ruth

- 11. To add titles to the table, click on the TITLES button. Up to three titles can be added one in each line:
 1 Calculating Frequencies: Example 2
 2 Two-way Table: Sex X Age
 3 Subsetting ages 0 to 29
 Click OK when finished.
- 12. Click **OK** to run analysis.

Output should look like the following:

Calculating Frequencies: Example 2 3 Two-Way Table: Sex X Age 10:09 Thursday, August 10, 2000 Subsetting ages 0 to 29 Where age Is Not Missing AND age LE 29 $\,$ The FREQ Procedure Table of sex by age sex age Frequency, Percent , Row Pct , Col Pct , 28, 29, Total 23, 26, 1, F 1, 1, 0, 3 , 25.00 , 25.00 , 25.00 , 0.00 , 75.00 , , 33.33 , 33.33 , 33.33 , 0.00, , 100.00 , 100.00 , 100.00 , 0.00, 0, 0, М 0, 1, 1 , 0.00 , 0.00 , $0.\ 00\ , \quad 25.\ 00\ , \quad 25.\ 00$, 0.00, 0.00, 0.00 , 100.00 , , 0.00 , 0.00, 0.00 , 100.00 , Total 1 1 4 1 1 100.00 25.00 25.00 25.00 25.00

Statistics for Table of sex by age

Statistic	DF	Val ue	Prob
ffffffffffffffffffffffffffffffffffff	fffffff	ſſſſſſſſſ	fffffff
Chi-Square	3	4.0000	0. 2615
Likelihood Ratio Chi-Square	3	4. 4987	0. 2124
Mantel-Haenszel Chi-Square	1	1. 1905	0. 2752
Phi Coefficient		1.0000	
Contingency Coefficient		0. 7071	
Cramer's V		1.0000	

WARNING: 100% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 4

Example 3: Three-Way Table on a Subset of data

This example creates a three-way table using three variables on a subset of the data. The first variable (*name*) represents the control variable for which separate output (cross-tabs of the other two variables - *height*, *weight*) is created for each of its values.

1. Filter the data and create a subset based on age, keeping only ages 0 to 27:

- From the menu at the top of the screen, select Data → Filter → Subset Data.
- In the dialogue window that pops up, select the *age* variable from the **AVAILABLE COLUMNS**.
- Following the instructions in Example 3, 'write' a statement identifying the conditions to determine which records will be included. From the list of possible **OPERATORS** that also appears beside the **WHERE** box, select **GE** (which stands for 'greater than or equal to).
- To specify the minimum *age* to be included, select **<CONSTANT enter value>** from the **AVAILABLE COLUMNS**. Type in 20 in the window that pops up and click **OK**.
- To finish the statement of conditions, select *age* from AVAILABLE COLUMNS. Select AND, and then LT (less than) from the list of OPERATORS.
- To specify the maximum *age*, select **<CONSTANT enter value>** and type in 28. Click **OK**.
- Click **OK** when finished
- The data list will now show only records for Donna and Helen since they are the only ones whose ages are between 0 and 27.

VORKT	MP_0]			_ 8 ×
Ele View Icols Solut	tions Window Help			X
_ ✓		公園 水田嶋・	이 🗊 봰 🛄 🗶 🕐	
Available Columns <pre></pre> <pre></pre> <pre></pre> <pre></pre> <pre> <pre>Available Columns </pre> <pre></pre> <pre></pre> <pre> <pre> <pre> <pre> </pre> <pre> <pre> <pre> </pre> <pre> </pre> <pre> </pre> <pre> <pre> </pre> </pre> <pre> </pre> </pre> <pre> </pre> </pre> <pre> </pre> <pre> </pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <pre> <</pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre>	>	Operators		OK Undo Help Reset Cancel
Output - (Untitled)	Log - (Untitled)	Program Editor - htwt	Subset: WORK	
			C:\WINNT\Profiles\P	iuth-s

- 2. Create tables:
- From the menu at the top of the page, select **STATISTICS** → **TABLE ANALYSIS**
- In the dialogue window that pops up, you can specify the row, column and strata variables. Select the desired variable and click on the appropriate varible type: *height* is the ROW variable, *weight* is the COLUMN variable, and *name* is the STRATA variable (also determines the name of the table).

▼ SAS	_ # ×
Eile View Tools Solutions Window Help	
) 🗸 📄 🗁 🖬 🖨 🗋 42 📴 🖻 🔺 🗠 X 🖾 🚧 🖉	
Analyst: (new project)	
Table Analysis: Htwt	;
Row Column C sex age height 0K Benove Strata Cell Counts Renove Reset	l
Tables Titles Variables	
Et De Table Analysis	
- Table Analysis of Htwt	
	1
🖸 Output - (Untitled) 📔 Log - (Untitled) 📝 Program Editor - htwt 🌌 Table Analysis:	
C:\WINNT\Profiles\Puth-	

- Given the selected variables, two tables are possible. To select the preferred table, click on the **SELECT TAB** button and then highlight the three way table option: **NAME*HEIGHT*WEIGHT**. Click **OK**.
- To specify the information to be included in the table, click on the **TABLES** button. For this example, under **FREQUENCIES**, check off **observed**, and under **PERCENTAGES**, check off **cell**, **row**, and **column**. Click **OK**.

₩ SAS	_ # ×
Eile View Tools Solutions Window Help	
	۲
Analyst: (new project)	
Table Analysis: Tables	i ا
C se: Frequencies Percentages OK	
Expected Row Cancel	
Reset	
Help	
Print missing value frequencies	
Insut Palent Tab Statistics	
Tables Titles Usriables	
I CALL TRAVE AVIRATORS OF TRAT	
- PhCode	
- Em Lacie Analysis of Hwit	
L Code J,	<u>اتر</u>
🗈 Output - (Untitled) 📄 Log - (Untitled) 🚺 Program Editor - htwl	
C/WINIT/Profiles	sVRuth-

• The output should look like:

Where age GE 20 AND age LT 28 The FREQ Procedure Table 1 of height by weight Controlling for name=Donna hei ght weight Frequency, Percent , Row Pct , Col Pct , 98, 121, Total fffffffffffffffffffffffff 62, 1, 0, 1 , 100.00 , 0.00 , 100.00 , 100.00 , 0.00 , , 100.00 , *ffffffffffffffffffffffffffffffff* 64, 0, 0, 0 , 0.00 , 0.00 , 0.00 · , . , , 0.00 , . , fffffffffffffffffffffffff 1 Total 0 1 0.00 100.00 100.00 Table 2 of height by weight Controlling for name=Helen hei ght weight Frequency, Percent , Row Pct , Col Pct , 98, 121, Total fffffffffffffffffffffffff 62, 0, 0, 0 0.00 , 0.00 , 0.00 , , fffffffffffffffffffffffff 0, 1, 64 , 1 0.00 , 100.00 , 100.00 , 0.00 , 100.00 , , . , 100.00 , *fffffffffffffffffffffffffffffff* Total 0 1 1 0.00 100.00 100.00

IV. Data Manipulation: Creating New Variables

A. Create new variables that incorporate computations made on existing variables.

If continuing from the previous example, take off the filter: from the menu at the top of the screen, select DATA \rightarrow FILTER \rightarrow NONE.

Ensure that the edit mode is set to EDIT: From the menu at the top of the screen, select EDIT \rightarrow MODE \rightarrow EDIT.

Each of the following computations must be made separately in the **Compute window**. Go to the menu at the top of the screen and select **DATA** \rightarrow **Transform** \rightarrow **Compute**.



- 1. New Variable Name
- 2. Variable List
- a) Variable Select Buttonb) Function Select Button
- 4. Statement Box

- 5. Mathematical Operators
- 6. Attributes
- 7. Functions Categories
- 8. Function Sub-categories

- 1. Add '2' to height for each record.
 - Select *'height'* from the Variable List (#2) and click on the VARIABLE SELECT button (#3a). This variable now appears in the Statement Box (#4).
 - Click on the PLUS SIGN from the selection of Mathematical Operators (#5).
 - Type '2' in the **Statement Box**
 - To verify that the expression is valid, click on the **VERIFY** button just below the **Statement Box**.
 - Type in the New Variable Name (#1): ht2p. Click OK when finished.
- 2. Subtract '1' from age for each record.
 - Select *age* from the Variable List and click on the VARIABLE SELECT button.
 - Click on the **MINUS** sign from the selection of **Math Operators**.
 - By default the column name is *Comp1*. Replace it with **agem1**
 - Click **OK** when finished.
- 3. Convert *inches* to *feet* for *height* (i.e. divide value for height by 12) and limit to one decimal place for each record.
 - Select 'height' from the Variable List and click on the VARIABLE SELECT button
 - Click on the **DIVIDE** sign from the selection of math operators. Type in 12 on the same line.
 - Type in the new column name: htfeet.
 - To limit the number of decimal places, click on the **ATTRIBUTES** button.

V SAS	×
Eile View Tools Solutions Window Help	
Column Attributes	
Colunn	
Name: htfeet OK	
Label: Cancel	
Beset	
Type: • Numeric Character	
Length: 8 Help	
Fornat: best8.	
Informat: 8.	
17 Tin M 29 71 176	
18 Susan F 28 65 131	
تے ہے۔ تے ا	
	-
🖸 Output - (Untitled) 📄 Log - (Untitled) 🖉 Program Editor - htwL.	
C:\WINNT\Profiles\Puth.	

The **COLUMN ATTRIBUTES WINDOW** enables you to specify the attributes for your computed column.

- The name of the column should be htfeet.
- Specify whether the column type is Numeric or Character. In this example, it is numeric.
- The variable length in the **Length** field can remain at the default of **8**.
- The arrows next to **Format** and **Informat** allow you to select the format and informat for the column. They should be consistent with each other and with the variable length. A format controls how values are displayed, while an informat controls how values are read by the SAS System.

Scroll down the list of **FORMATNAMES** until you come to: **w.d. standard numeric width, decimal,** and select it. This option allows you to set the number of decimal places to **1**. Click **OK** when finished.

<u>Note</u>: the **WIDTH** field specifies the maximum number of places before the decimal places. For this example, leave the default of 8.

📲 5A5
Eile View Tools Solutions Window Help
X
Format Names
s370fribu IBM 370 unsigned interger binary
s370fpdu IBM 370 packed decinal
s370fpib IBM 370 positive integer binary s370frb IBM 370 real binary
s370fzd IBM 370 zoned decimal Valid Width
s370fzds IBM 370 zoned decimal Bange: 1-32
s370fzdt IBM 370 zoned decimal Width: 8
tine time values - Decimal: II
tod time-of-day
Vid a candard home re vider de man
weekdate date values
weekdate date values weekdatx date values weekday day of week
weekdate date values weekdatx date values weekday day of week
weekdate date values weekdaty day of week
weekdate date values weekdatx date values weekday day of week
weekdate date values weekday day of week 17 Tin M 29 71 176
weekdate date values weekday day of week 17 Tin M 29 71 176 18 Susan F 28 65 131
weekdate date values weekdat date values weekday day of week 17 Tim M 29 71 176 18 Sutan F 28 65 131
weekdate date values weekdat date values weekday day of week 17 Tim M 29 71 176 18 Susan F 28 65 131
weekdate date values weekday day of week 17 Tim M 29 71 176 18 Susan F 28 65 131
weekdate date values weekday day of week
weekdate date values weekday day of week 17 Tin M 29 71 176 18 Susan F 28 65 131
weekdate date values weekdatx date values weekday day of week OK OK OK Cancel Image: Concel Image:

• Follow the same procedure to change **INFORMAT** to **1** decimal place.

4. Convert pounds to ounces for *weight* for each record.

- Select *weight* from the VARIABLE LIST and click on the VARIABLE SELECT button.
- Click on the **multiplication** sign from the selection of math operators. Type in **16** on the same line.
- Type in the new column name: **wtounce**

• Click on OK.

B. Create new variables that truncate the values of existing variables.

Ensure that the edit mode has been changed from BROWSE to EDIT: From the menu at the top of the screen, select EDIT \rightarrow MODE \rightarrow EDIT.

Each of the following computations must be made separately in the **Compute window**. Go to the menu at the top of the screen and select **DATA** \rightarrow **Transform** \rightarrow **Compute**.

Refer to the figure in the previous example to help identify the various fields and boxes mentioned below.

1. Truncate *name* to display only the first two values.

- Change column name to **name1**.
- Because *name* is a character variable, change the Functions Categories (Box #7) to Character.
- Click on the **ATTRIBUTES** button and change the **VARIABLE TYPE** to **character**. Click **OK**.
- Scroll down the FUNCTION SUB-CATEGORIES and select SUBSTR.
- The following expression appears in the **STATEMENT BOX**: substr(string,n<,n>)
- Delete *string* and replace with *name* by highlighting *name* in the VARIABLE LIST and clicking on the FUNCTION SELECT BUTTON.
- The first *n* represents the starting point for the truncation. Replace it with 1 and delete the '<' beside it.
- The second *n* represents the number of values you want to be displayed. Replace it with 2 and delete the '>' beside it.
- Click OK.

2. Truncate name to display the first 3 values.

Follow the instructions from Example 1 except:

- Change column name to **name2**
- Replace the *n*'s in the substring expression with **1** and **3**, respectively.
- 3. Truncate *name* to display the first 3 values, starting at the second value.

Follow the instructions from Example 1 except:

- Change the column name to **name3**
- Replace the *n*'s in the substring expression with 2 and 3, respectively.

C. Creating and grouping formats to create new variables

Character Variables

Character variables can be grouped in one of two ways - by recoding the values or by integrating SAS code with the point-and-click approach of Analyst.

This example will show you how to recode the *name* variable both ways. 'Elizabeth', 'David', and 'James' will be grouped as one new variable.

1. Recode Values

- From the menu at the top of the screen, select DATA → TRANSFORM → RECODE VALUES
- Click in the line for COLUMN TO RECODE for a list of variables. Select 'name'
- Select CHARACTER for NEW COLUMN TYPE
- Click OK
- In the dialogue window that pops up, enter '1' as the new value for 'Elizabeth', 'David', and 'James' and 'o' for the other names. Hit the **TAB** key twice to move to the next line.

Recode Values: Htw	
inter a new column column value. If missing value will	value to correspond to each original you do not enter a new value, a be used.
Original Value	New Value (Character)
Antonio	
Autorey Devi	Heset
Jan	Help
Advid .	
Aecoran	
Anna III.	
aizabeth	
lacqueine	
Aurrados Incel	
loei Kahaal	
elcriment	<u>د</u>
	18 Susan F 28 65 131
	and the second sec

• A new column entitled 'name_recoded' now appears in the data list.

NOTE: This method is simple and straightforward for a small data set or for binary (dichotomous) variables, but not very efficient for large data sets or for non-dichotomous variables with several value options.

2. PUT statement

SAS uses formats to read, display, and print data stored in a SAS data set.

The PUT function associates a format with an existing variable to create a new variable. It can be used with either numeric or character variables. The PUT statement can also be used to convert numeric variables into character variables.

• Create a format for the *name* variable: In the PROGRAM EDITOR of (base) SAS, type in the following code: proc format;

Value \$namef 'Elizabeth', 'David', 'James' = '1'

Run; The '\$' denotes character values.



• Submit the program by clicking on the SUBMIT icon at the top of the screen

This creates a 'look up' table that converts or labels variables, and can be referred to for further analyses in the same SAS session.

Other = '0':

- In ANALYST, from the menu at the top of the screen select DATA → TRANSFORM → COMPUTE.
- In the dialogue window that pops up, any valid SAS code can be entered into the statement box. Type in the following statement: put (name,\$namef.)

This statement instructs SAS to create a new variable using the format created above.

♥SAS Eile ⊻iew Iools <u>S</u> ok	utions Window Help
Compute: Htwt name2 - put (name,\$namef.) Var iables C name C sex age he ight we ight C name_recode	Becall Image: Second se
1	
Dutput - (Untitled)	🖹 Log - (Untified) 🚺 Program Editor - htwt 🍑 Compute: Htwt
	C:\WINNT\Profiles\Puth.

- To ensure that Analyst knows this is a character variable, click on the **ATTRIBUTES** button and check off this option in the dialogue window that pops up. Click **OK**.
- Change the name of the new variable by replacing COMP1 with **name2.** Click **OK**.

This generates a new column with values of either 0 or 1 instead of actual names.

Numeric Variables

<u>Note:</u> Numeric variables can also be converted to character variables to save storage space or if there are leading zeros you wish to keep.