

BioMedStat

Version 1.0

User Guide

© 2004-2005 Peterson Software Lab

Baylor College of Medicine

Table of Contents

INPUT FILE FORMAT

Tab-delimited Text Files with Variable Names

Tab-delimited Text Files with Variable and Record Names

EXAMPLE 1 - OPENING A TAB-DELIMITED TEXT FILE

EXAMPLE 2 – SUMMARY STATISTICS OF SEVERAL VARIABLES

EXAMPLE 3 – RECODE A CONTINUOUS VARIABLE INTO QUARTILES, AND THEN RECODE INTO 4 NEW INDICATOR VARIABLES

EXAMPLE 4 – PARAMETRIC 2-SAMPLE PAIRED T-TEST

EXAMPLE 5 – NON-PARAMETRIC 2-SAMPLE PAIRED WILCOXON SIGNED RANK TEST

EXAMPLE 6 – PARAMETRIC INDEPENDENT 2-SAMPLE T-TEST

EXAMPLE 7 – NON-PARAMETRIC INDEPENDENT 2-SAMPLE MANN-WHITNEY U TEST

EXAMPLE 8 – PARAMETRIC INDEPENDENT K-SAMPLE ANALYSIS OF VARIANCE

EXAMPLE 9 – NON-PARAMETRIC INDEPENDENT K-SAMPLE KRUSKAL-WALLIS TEST

EXAMPLE 10 – PEARSON PRODUCT MOMENT AND SPEARMAN RANK CORRELATION

EXAMPLE 11 – COVARIANCE ANALYSIS

EXAMPLE 12 – CHI-SQUARE 2-WAY CONTINGENCY TABLE ANALYSIS

EXAMPLE 13 – LINEAR REGRESSION

EXAMPLE 13 – LOGISTIC REGRESSION (UNCONDITIONAL)

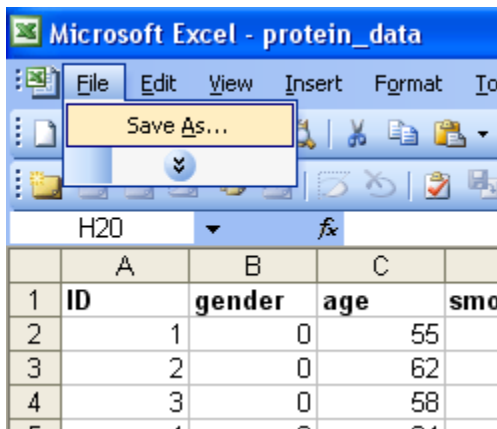
Input File Format

Tab-delimited Text Files with Variable Names

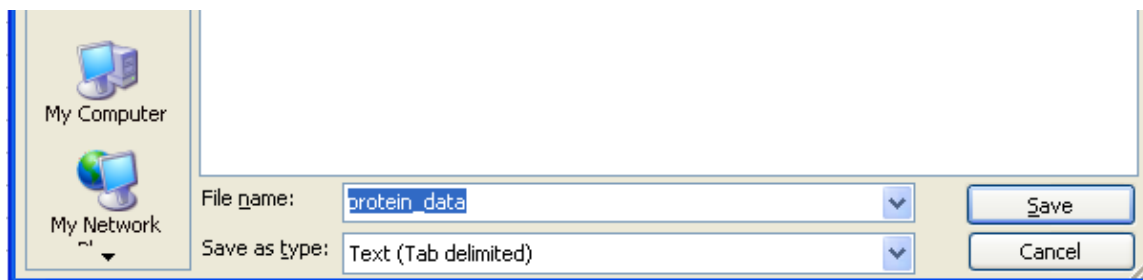
Tab-delimited text files are comprised of data separated (i.e., delimited) with a character tab. Tab-delimited text files used for data analysis typically have variable names in the first row of the file. The simplest way to generate a tab-delimited text file is by specifying “tab-delimited text” as a file save option in Microsoft Excel. The example below shows the setup necessary for saving data and variables names to a tab-delimited text file in Excel.

	A	B	C	D	E	F	G
1	ID	gender	age	smoke	protein A	protein B	
2	1	0	55	0	0.56	1.48	
3	2	0	62	0	1.10	0.22	
4	3	0	58	0	0.97	1.08	
5	4	0	64	1	0.90	0.41	
6	5	0	47	0	0.30	0.72	
7	6	1	59	1	1.01	3.35	
8	7	1	38	0	4.50	2.48	
9	8	1	75	0	2.29	1.57	
10	9	1	56	1	0.82	6.78	
11	10	1	40	1	0.26	1.61	
12							
13							

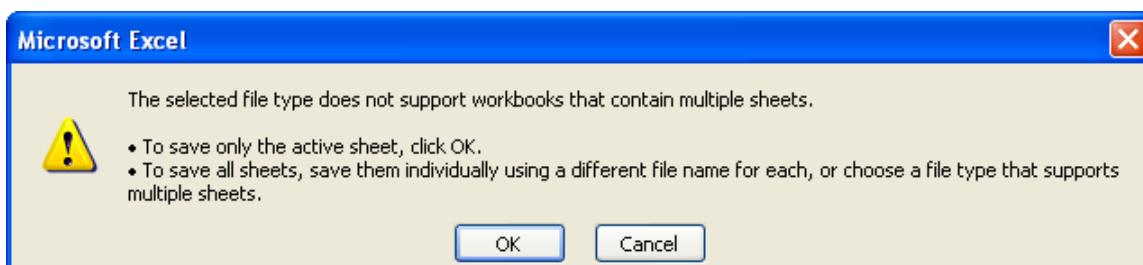
To save the data listed above, select File, then Save As, then select Tab delimited (txt), as in:



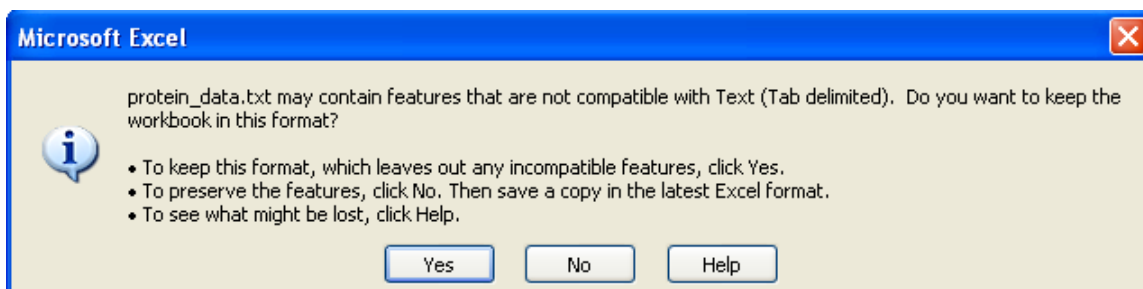
Next, in the Save As window, specify the file name as "protein_data":



You will notice two popup windows, the first of which states that multiple sheets cannot be saved in the file, so click OK:



And the second which states that you can lose special features of Excel when saving data into a tab-delimited text file. For example, bold fonts and colors cannot be saved with the data, thus, click Yes:



Next, open the "protein_data.txt" file just saved, and the following will appear:

protein_data - Notepad

ID	gender	age	smoke	protein A	protein B
1	0	55	0	0.56	1.48
2	0	62	0	1.10	0.22
3	0	58	0	0.97	1.08
4	0	64	1	0.90	0.41
5	0	47	0	0.30	0.72
6	1	59	1	1.01	3.35
7	1	38	0	4.50	2.48
8	1	75	0	2.29	1.57
9	1	56	1	0.82	6.78
10	1	40	1	0.26	1.61

Tab-delimited Text Files with Variable and Record Names

The following illustrates a file setup with both variable and record names. For this example, the variable names are the sample (patient) identifiers and the record names in the last column are the genes or proteins. This setup is commonly used for DNA microarray data, however, BioMedStat was designed for clinical statistical analysis for clinical data in which records represent patients and columns represent variables or measurements made on the experimental units (patients). (If you want to analyze microarray data, then use the [ChipST2C](#) program). The example is shown below:


	A	B	C	D	E	F	G	H	I	J	K
1	samp1	samp2	samp3	samp4	samp5	samp6	Gene				
2	1.95	6.34	0.08	0.49	1.14	0.41	calmodulin 2 (phospho				
3	0.79	0.33	1.12	3.11	0.78	2.22	Human 90-kDa heat-shock protein gene; cDNA; complete cds_3				
4	1.89	0.85	0.51	0.47	1.12	1.15	phosphofructokinase; liver_36950				
5	3.30	0.44	1.57	2.52	0.83	0.27	glutamate dehydrogenase 1_39285				
6	0.77	0.60	0.65	1.19	0.74	0.29	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (hydroxyme				
7	0.32	0.36	0.24	0.54	0.23	0.60	accessory proteins BAP31/BAP29_37196				
8	2.77	0.44	0.55	2.81	0.32	7.07	glycogen synthase kinase 3 beta_37904				
9	0.82	3.35	0.25	1.14	0.58	1.50	Human cyclin G1 interacting protein (1500GX1) mRNA; complete				
10	0.61	2.14	3.02	1.44	0.12	2.97	methionine aminopeptidase; eIF-2-associated p67_39093				
11	3.29	0.41	3.21	5.17	0.32	0.91	superoxide dismutase 1; soluble (amyotrophic lateral sclerosis 1				
12											
13											

To continue, after the data above were saved as a tab-delimited text file and opened, the following format will be observed:

protein_sample_data - Notepad

samp1	samp2	samp3	samp4	samp5	samp6	Gene
1.95	6.34	0.08	0.49	1.14	0.41	calmodulin 2 (phosphory
0.79	0.33	1.12	3.11	0.78	2.22	Human 90-kDa heat-shock
1.89	0.85	0.51	0.47	1.12	1.15	phosphofructokinase; 11
3.30	0.44	1.57	2.52	0.83	0.27	glutamate dehydrogenase
0.77	0.60	0.65	1.19	0.74	0.29	3-hydroxymethyl-3-methy
0.32	0.36	0.24	0.54	0.23	0.60	accessory proteins BAP3
2.77	0.44	0.55	2.81	0.32	7.07	glycogen synthase kinas
0.82	3.35	0.25	1.14	0.58	1.50	Human cyclin G1 interac
0.61	2.14	3.02	1.44	0.12	2.97	methionine aminopeptida
3.29	0.41	3.21	5.17	0.32	0.91	superoxide dismutase 1;

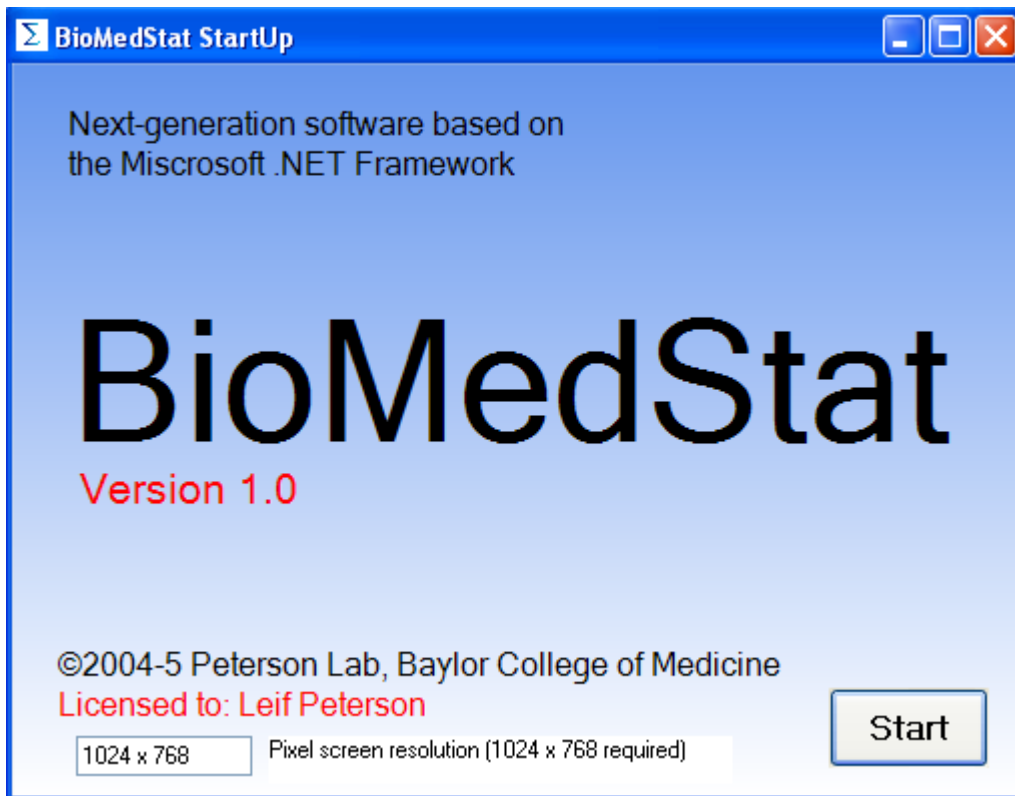
Example 1 - Opening a Tab-Delimited Text File

When BioMedStat is installed, the icon  is placed on the Desktop with a shortcut link to the BioMedStat program. The desktop shortcut appearance is as follows:

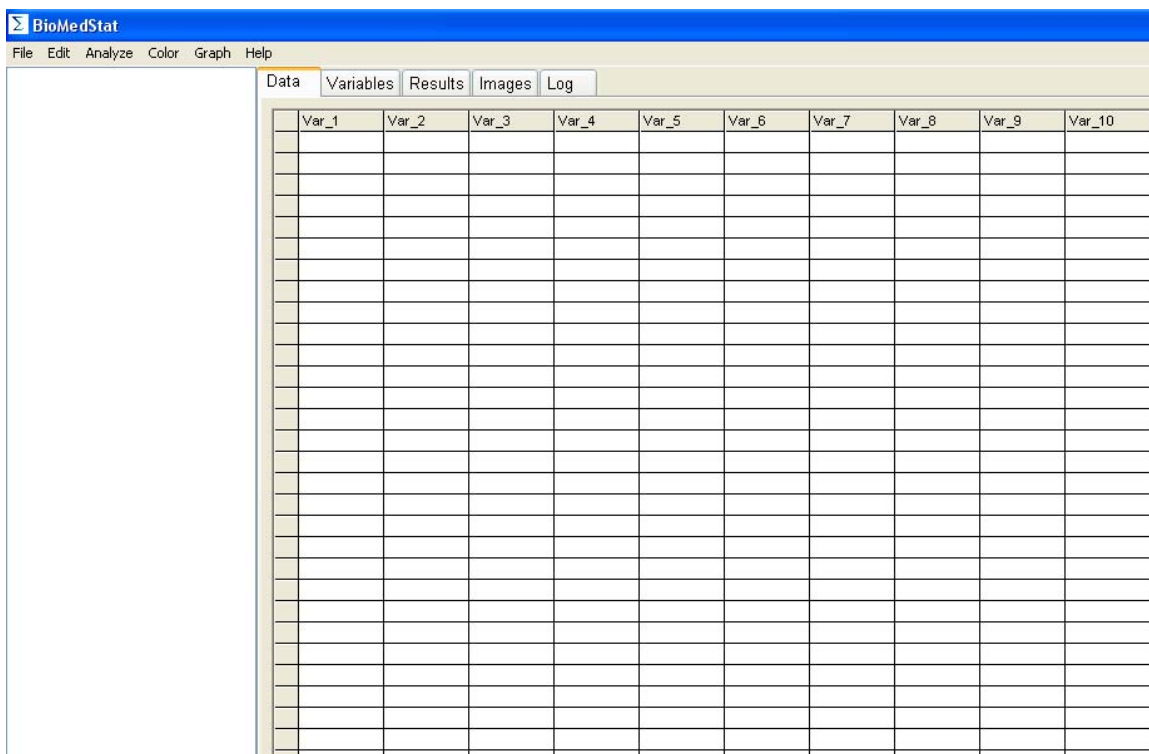


(Note that in some cases, BioMedStat may be installed without a desktop shortcut requiring program startup by on the Start, then Programs, then BioMedStat.)

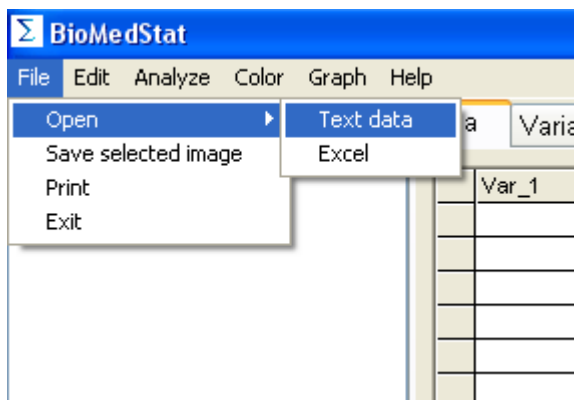
To start BioMedStat, double-click on the desktop icon. You will then see a splash screen:



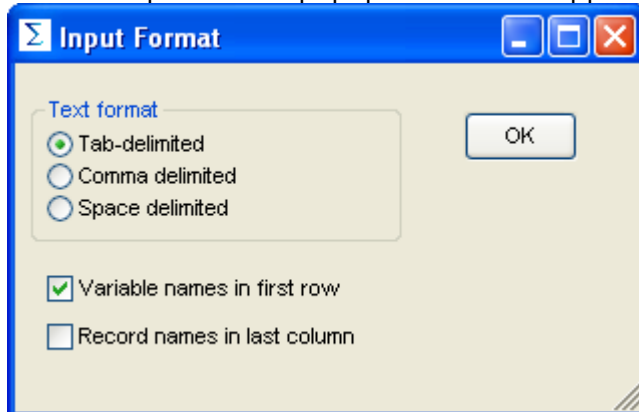
Next click on the Start button, and the BioMedStat application will then be visible, as shown below:



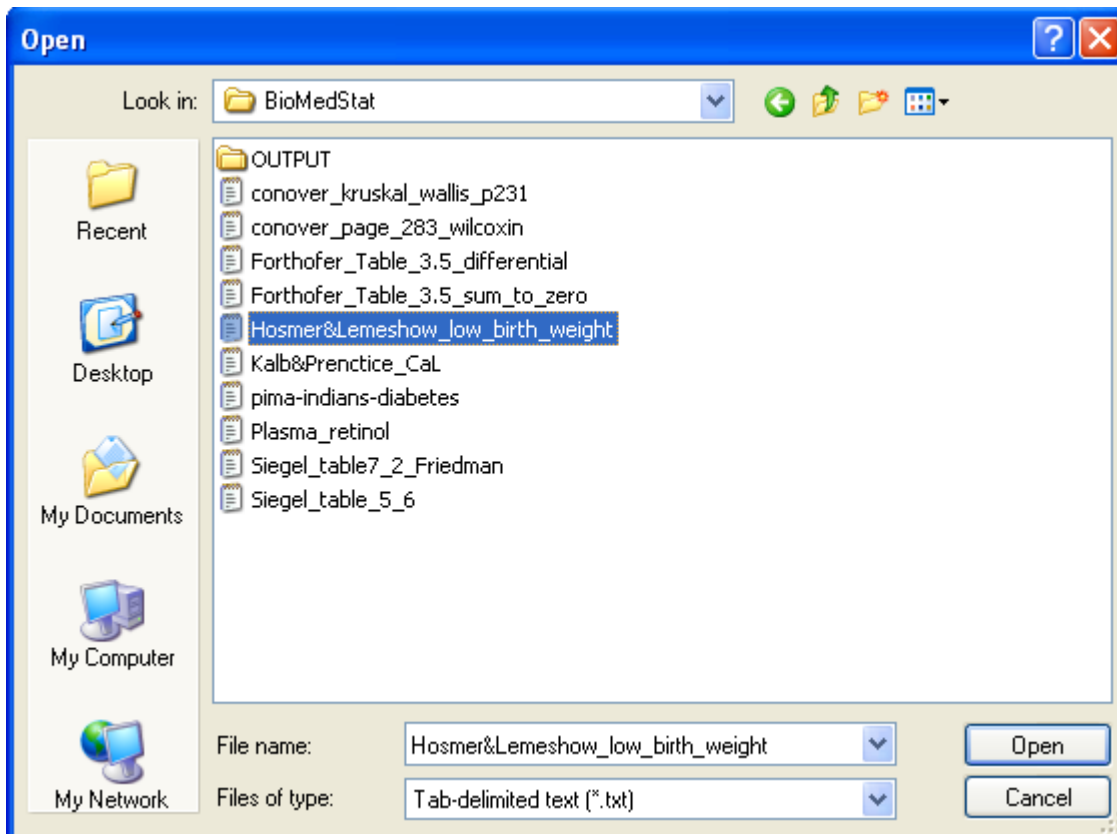
To open a file, select File, Open, Text data, as shown below



and the Input format popup window will appear:



Since tab-delimited is the default text file format it won't need to be specified, however, check the "Variable names in first row" option and click OK. In the Open popup window, you should see the text files that were installed in the c:\Program Files\BioMedStat\BioMedStat\ directory, as shown below:



Select the Hosmer & Lemeshow low birth weight file, and then click on Open (or double-click the filename). When the file is read into BioMedStat, the status text field and progressbar on the log tab will indicate the amount of data read and copied into the data viewing spreadsheet. When completed, the data viewing tab will show the data that were read in. **(Note: the spreadsheet in the Data tab is only used for viewing data and not editing, i.e., changing, copying, pasting, etc.)**

The screenshot shows the BioMedStat software window. The 'Data' tab is selected, displaying a table with 189 rows (observations) and 11 columns (variables). The variables are: id, low, age, lwt, race, smoke, pti, ht, ui, ftv, and bwt. The data is organized into a grid with alternating light and dark rows for readability. The 'id' column ranges from 1 to 189. The 'low' column contains 0s and 1s. The 'age' column ranges from 17 to 33. The 'lwt' column ranges from 90 to 182. The 'race' column contains 1, 2, and 3. The 'smoke' column contains 0 and 1. The 'pti' column contains 0 and 1. The 'ht' column ranges from 58 to 72. The 'ui' column contains 0 and 1. The 'ftv' column ranges from 0 to 3. The 'bwt' column ranges from 2523 to 2977.

	id	low	age	lwt	race	smoke	pti	ht	ui	ftv	bwt
1	85	0	19	182	2	0	0	0	1	0	2523
2	86	0	33	155	3	0	0	0	0	3	2551
3	87	0	20	105	1	1	0	0	0	1	2557
4	88	0	21	108	1	1	0	0	1	2	2594
5	89	0	18	107	1	1	0	0	1	0	2600
6	91	0	21	124	3	0	0	0	0	0	2622
7	92	0	22	118	1	0	0	0	0	1	2637
8	93	0	17	103	3	0	0	0	0	1	2637
9	94	0	29	123	1	1	0	0	0	1	2663
10	95	0	26	113	1	1	0	0	0	0	2665
11	96	0	19	95	3	0	0	0	0	0	2722
12	97	0	19	150	3	0	0	0	0	1	2733
13	98	0	22	95	3	0	0	1	0	0	2750
14	99	0	30	107	3	0	1	0	1	2	2750
15	100	0	18	100	1	1	0	0	0	0	2769
16	101	0	18	100	1	1	0	0	0	0	2769
17	102	0	15	98	2	0	0	0	0	0	2778
18	103	0	25	118	1	1	0	0	0	3	2782
19	104	0	20	120	3	0	0	0	1	0	2807
20	105	0	28	120	1	1	0	0	0	1	2821
21	106	0	32	121	3	0	0	0	0	2	2835
22	107	0	31	100	1	0	0	0	1	3	2835
23	108	0	36	202	1	0	0	0	0	1	2836
24	109	0	28	120	3	0	0	0	0	0	2863
25	111	0	25	120	3	0	0	0	1	2	2877
26	112	0	28	167	1	0	0	0	0	0	2877
27	113	0	17	122	1	1	0	0	0	0	2906
28	114	0	29	150	1	0	0	0	0	2	2920
29	115	0	26	168	2	1	0	0	0	0	2920
30	116	0	17	113	2	0	0	0	0	1	2920
31	117	0	17	113	2	0	0	0	0	1	2920
32	118	0	24	90	1	1	1	0	0	1	2948
33	119	0	35	121	2	1	1	0	0	1	2948
34	120	0	25	155	1	0	0	0	0	1	2977
35	121	0	25	125	2	0	0	0	0	0	2977
36	123	0	29	140	1	1	0	0	0	2	2977
37	124	0	19	138	1	1	0	0	0	2	2977

Example 2 – Summary statistics of several variables

Data used for this example were published in the book Hosmer, D.W. and Lemeshow, S. *Applied Logistic Regression*, New York, Wiley (1989) and are available from the University of Massachusetts (Amherst) Statistical Software Information Internet resources at URL <http://people.umass.edu/statdata/statdata/>. These data are copyrighted by John Wiley & Sons Inc. and must be acknowledged and used accordingly. Data were collected at Baystate Medical Center, Springfield, Massachusetts during 1986.

The data set is comprised of 189 observations for 11 variables (risk factors) associated with giving birth to a low birth weight baby (weighing less than 2500 grams). Data were collected on 189 women, 59 of which had low birth weight babies and 130 of which had normal birth weight babies. Four variables which were thought to be of importance were age, weight of the subject at her last menstrual period, race, and the number of physician visits during the first trimester of pregnancy.

Columns	Variable	Abbreviation
2-4	Identification Code	ID
10	Low Birth Weight (0 = Birth Weight \geq 2500g, 1 = Birth Weight $<$ 2500g)	LOW
17-18	Age of the Mother in Years	AGE
23-25	Weight in Pounds at the Last Menstrual Period	LWT

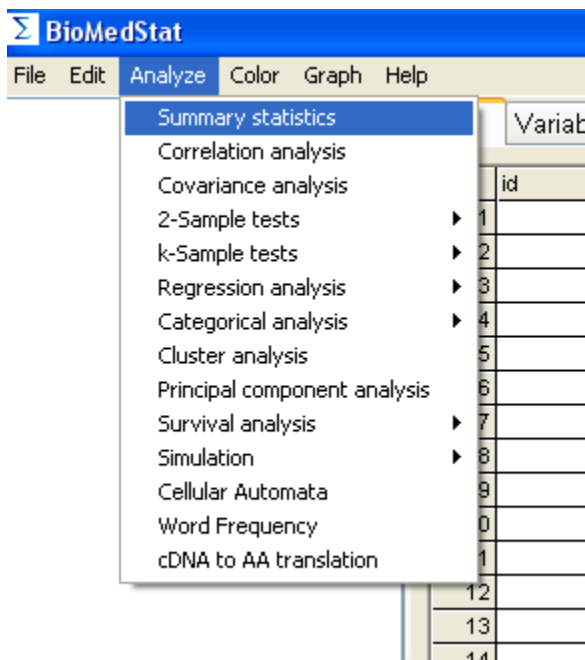
32	Race (1 = White, 2 = Black, 3 = Other)	RACE
40	Smoking Status During Pregnancy (1 = Yes, 0 = No)	SMOKE
48	History of Premature Labor (0 = None 1 = One, etc.)	PTL
55	History of Hypertension (1 = Yes, 0 = No)	HT
61	Presence of Uterine Irritability (1 = Yes, 0 = No)	UI
67	Number of Physician Visits During the First Trimester (0 = None, 1 = One, 2 = Two, etc.)	FTV
<hr/> 73-76 Birth Weight in Grams		BWT <hr/>

Background information for the data:

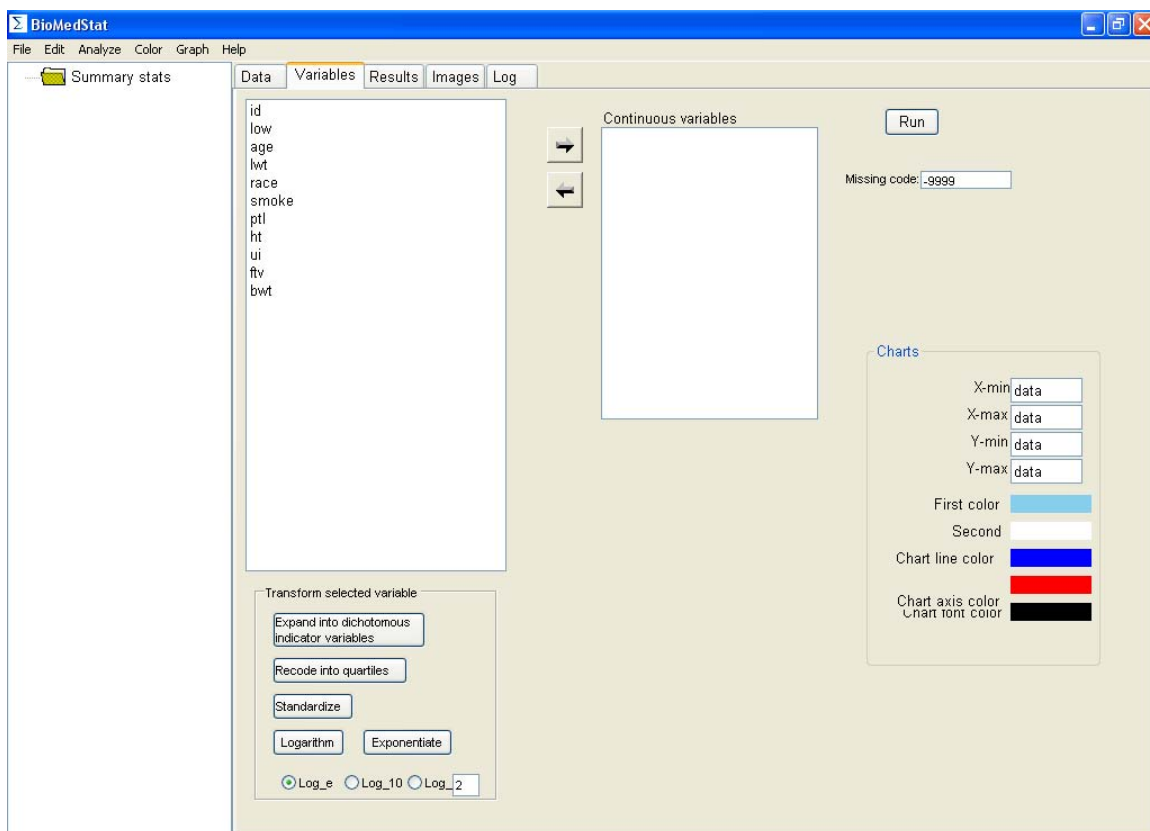
- Low birth weight is an outcome that has been of concern to physicians for years.
- Infant mortality rates and birth defect rates are very high for low birth weight babies
- A woman's behavior during pregnancy (including diet, smoking habits, and receiving prenatal care) can greatly alter the chances of carrying the baby to term and, consequently, of delivering a baby of normal birth weight

The risk factors described above have been shown to be associated with low birth weight in the obstetrical literature. The goal of the Hosmer & Lemeshow study was to ascertain if these variables were important in the population being served by the medical center where the data were collected.

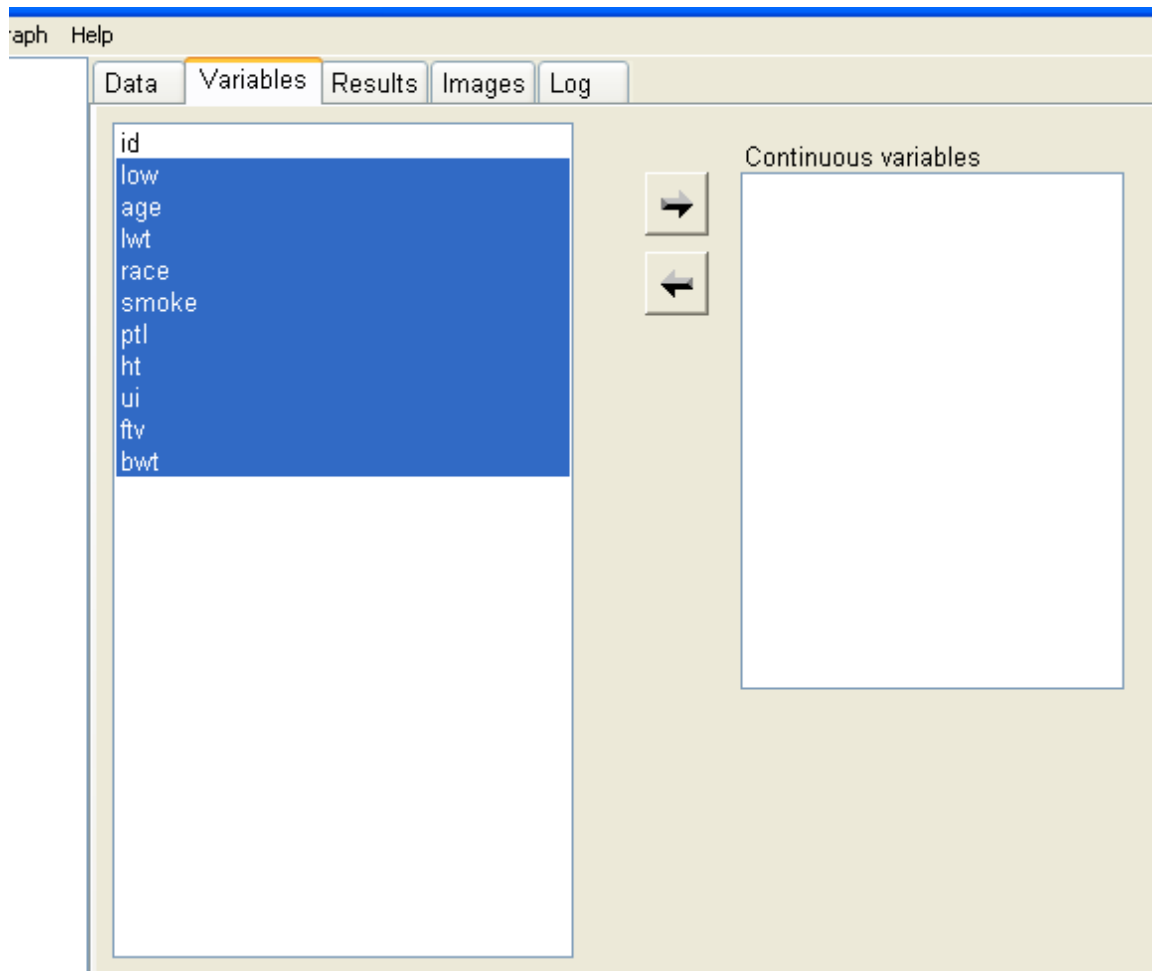
Open the tab-delimited text file for the Hosmer & Lemeshow low birth weight data (see Example 1). Select the Summary statistics command of the Analyze pull-down menu, shown as




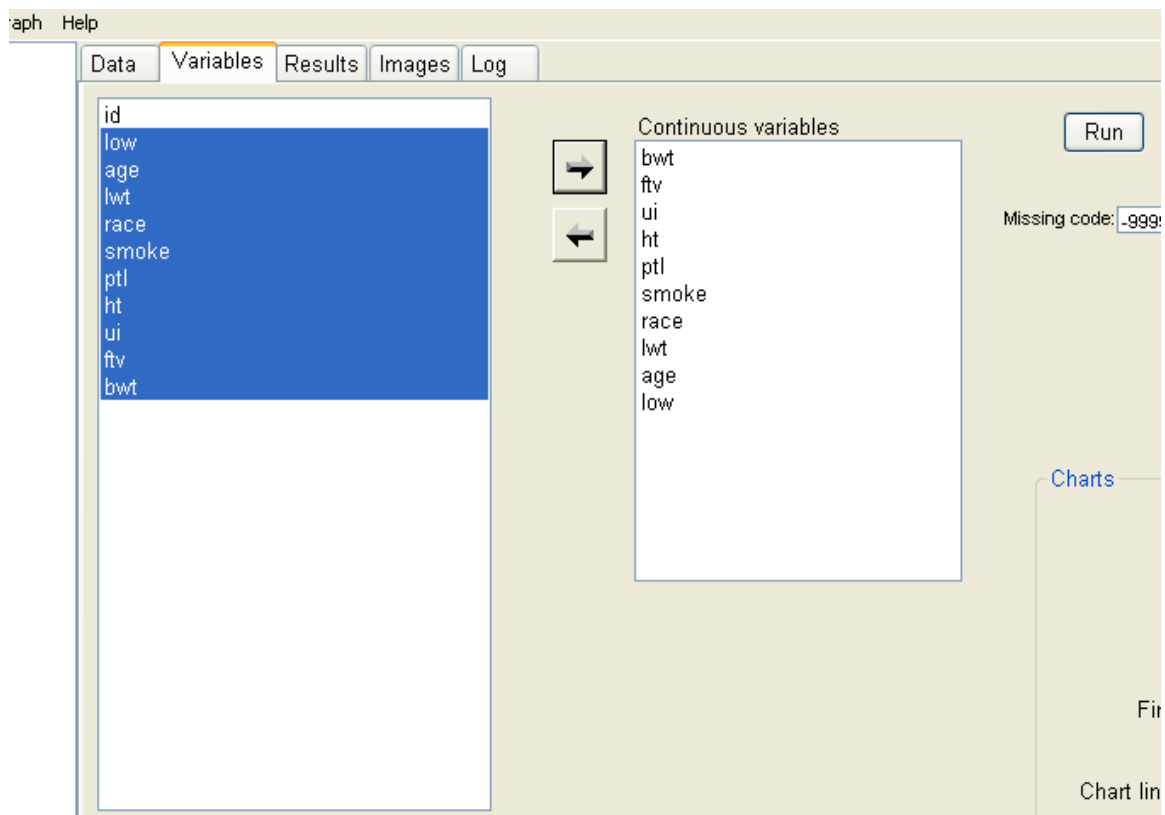
Notice that the Variables tab opens, shown as follows:



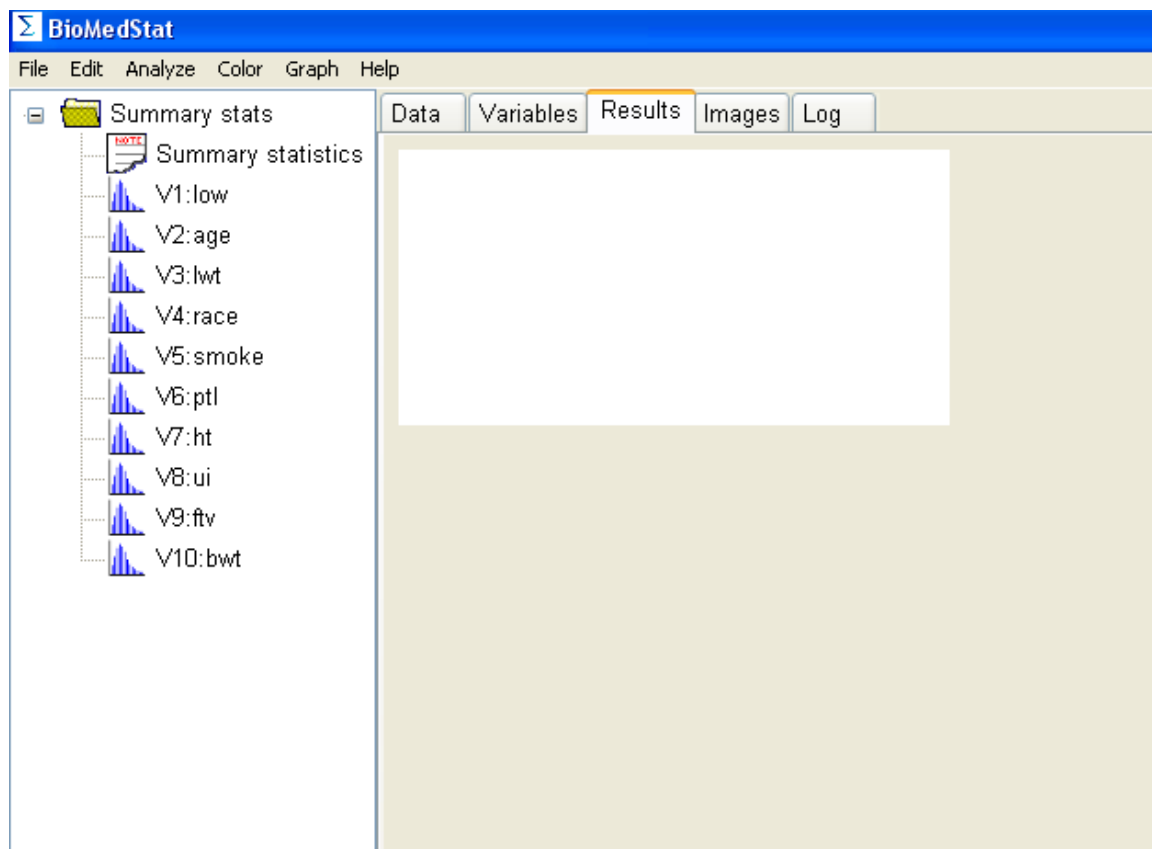
Select all but the first variable (ID number) in the following manner:




Then click on the  button to add the variables to the list of selected variables:



and finally, click on the  button. The Treeview on the left will show a number of icons in the form:



To view the summary statistics for each variable, click on the  icon, and the text output containing summary statistics for the specified variables will become visible:

BioMedStat

File Edit Analyze Color Graph Help


Summary stats Data Variables Results Images Log

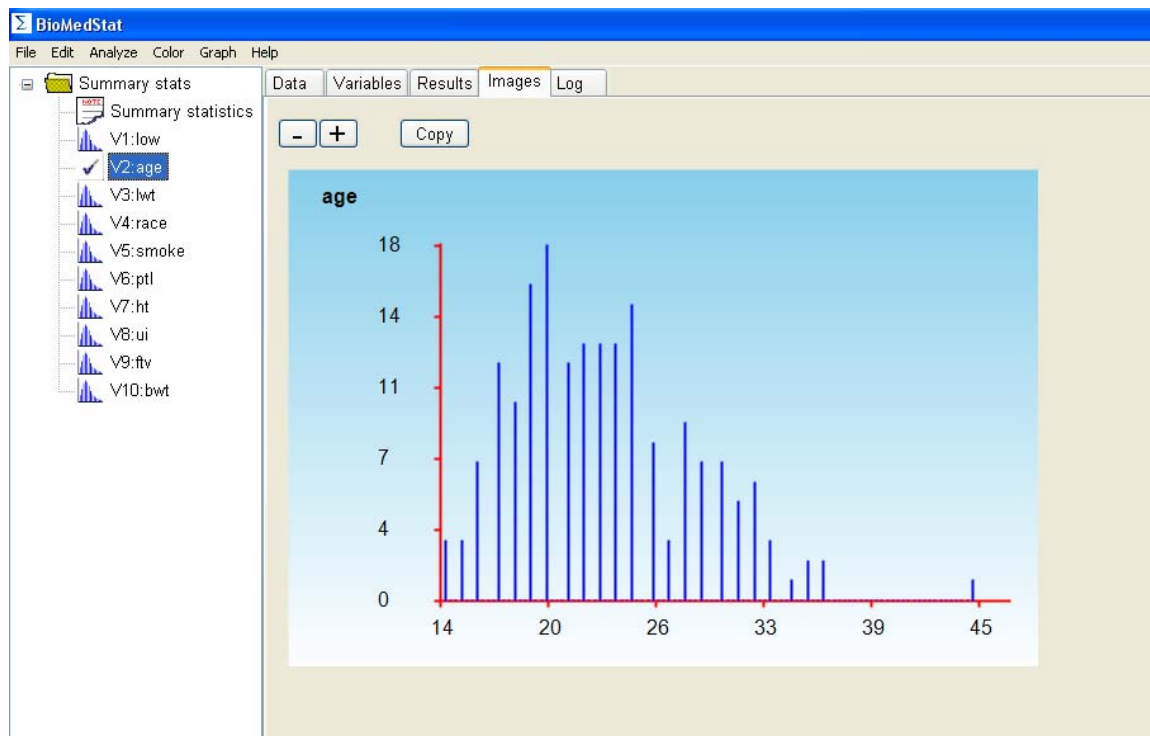
Summary statistics

V1:low
V2:age
V3:lwt
V4:race
V5:smoke
V6:ptl
V7:ht
V8:ui
V9:ftv
V10:bwt

Summary Statistics

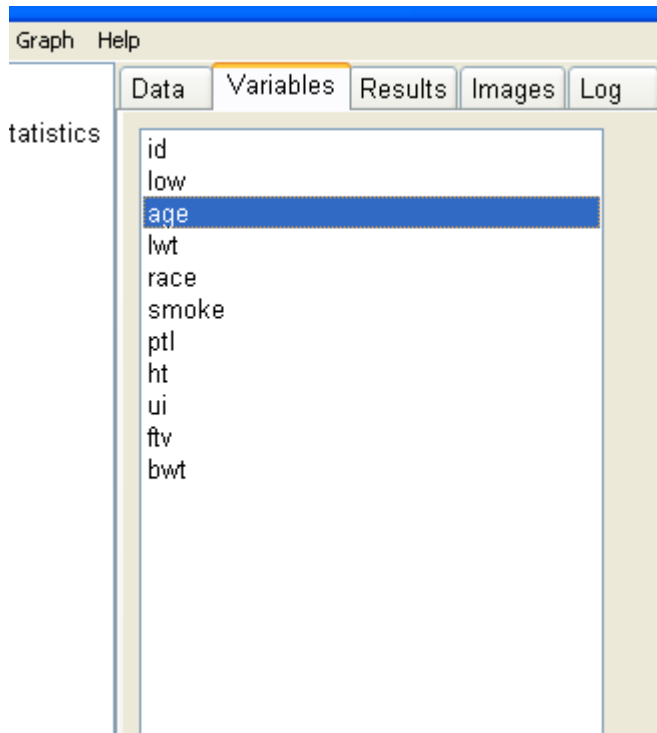
Variable	N	Average	Std. Dev.	Min	Max	Range	Median	Q1	Q2	Q3	Skewness
low	189	0.312	0.465	0.000	1.000	1.000	0.000	0.000	0.000	1.000	0.8
age	189	23.238	5.299	14.000	45.000	31.000	22.000	19.000	22.000	26.000	0.7
lwt	189	129.815	30.579	80.000	250.000	170.000	121.000	110.000	121.000	140.000	1.3
race	189	1.847	0.918	1.000	3.000	2.000	1.000	1.000	1.000	3.000	0.3
smoke	189	0.392	0.489	0.000	1.000	1.000	0.000	0.000	0.000	1.000	0.4
ptl	189	0.196	0.493	0.000	3.000	3.000	0.000	0.000	0.000	0.000	2.7
ht	189	0.063	0.244	0.000	1.000	1.000	0.000	0.000	0.000	0.000	3.5
ui	189	0.148	0.356	0.000	1.000	1.000	0.000	0.000	0.000	0.000	1.9
ftv	189	0.794	1.059	0.000	6.000	6.000	0.000	0.000	0.000	1.000	1.5

To observe the frequency histogram of age, click on the  icon and the following histogram will appear:

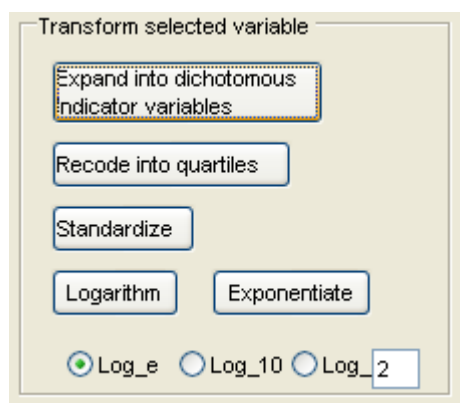


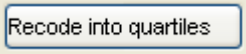
Example 3 – Recode a continuous variable into quartiles, and then recode into 4 new indicator variables

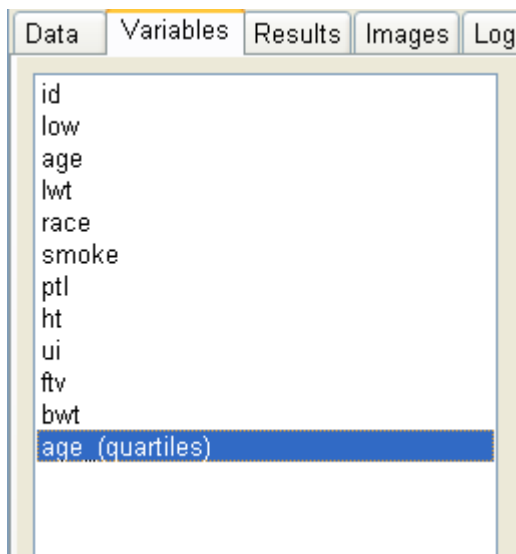
Open the Hosmer & Lemeshow low birth weight data (see Example 1), select the Variables tab and then select (highlight) the “age” variable as:

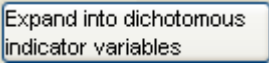


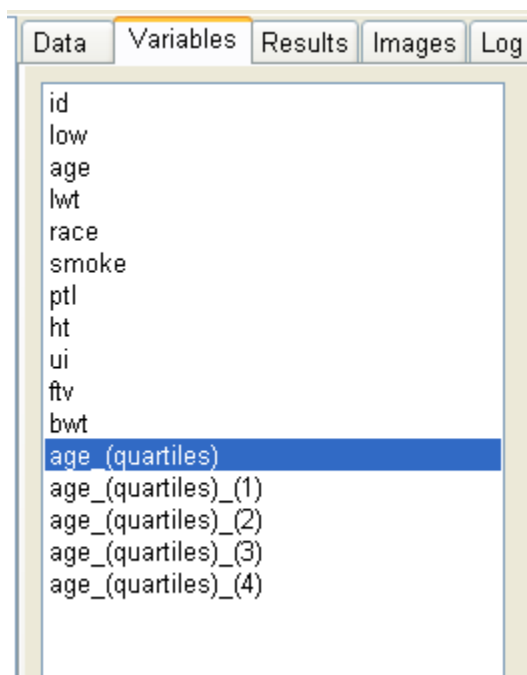
Next, at the bottom of the Variables tab are several buttons for transforming the values of variables into new values:



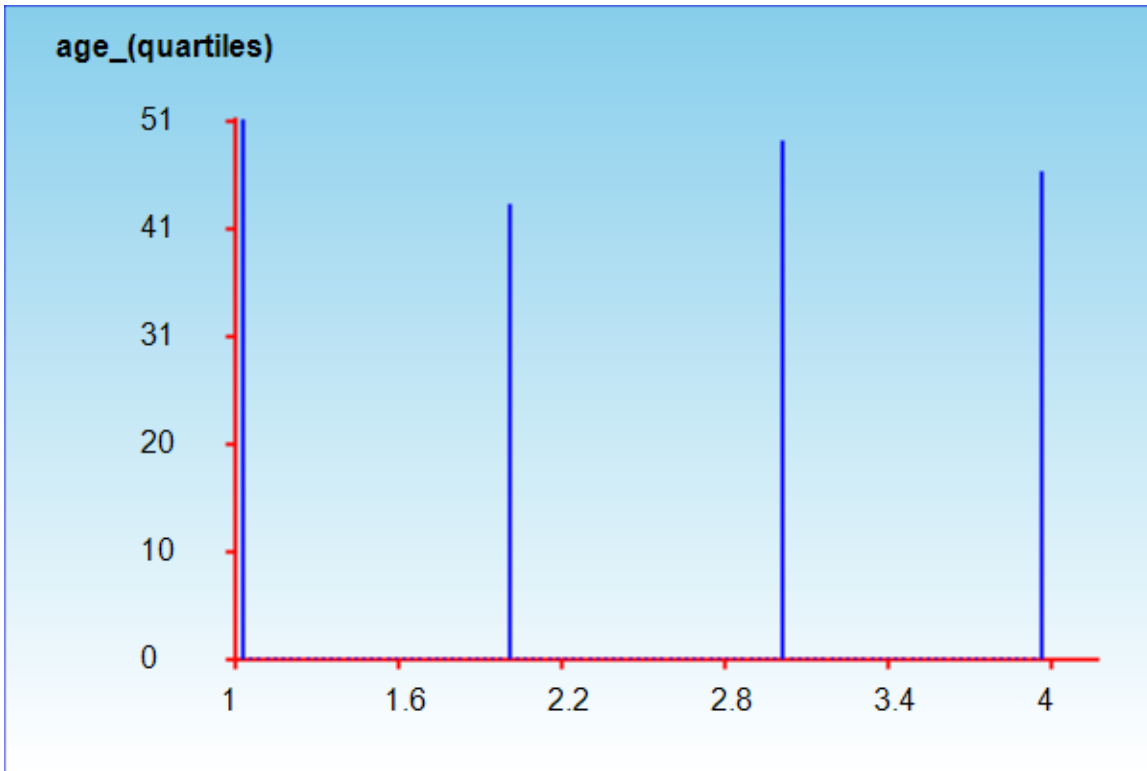
Now that the “age” variable is selected in the variable list, click on the  button, and a new (ordinally ranked) **categorical** variable called “age_(quartiles)” will be generated and shown in the variable list. At run-time, the cutpoints for quartiles are computed (to see their values, run the summary statistics option on the age variable) and the resulting values for the new variable are 1, 2, 3, 4, representing the age quartile that each record (patients) is assigned to. Select the new variable (“age_(quartiles)”) in the fashion



and then click on the  button. Upon completion of the transformation, there will be four new **indicator** or “dummy” variables added to the list, appearing as:



The four indicator variables have values of 0 or 1 depending on whether or not a patient’s age falls within the given quartile represented by the variable. If you run summary statistics on the age_(quartiles) variable, the following histogram will be generated:



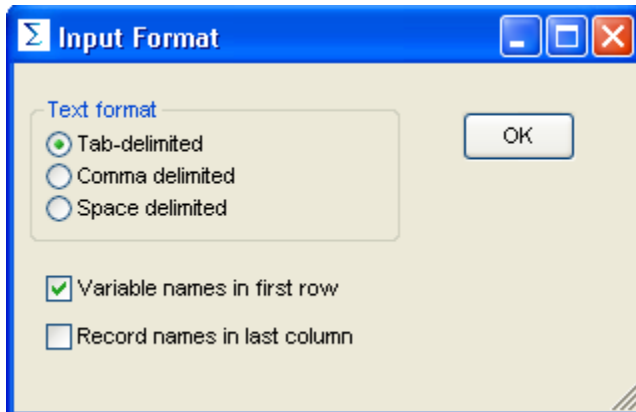
which clearly shows that the variable takes on values of 1, 2, 3, or 4 at various frequencies of occurrence.

Example 4 – Parametric 2-sample paired T-test

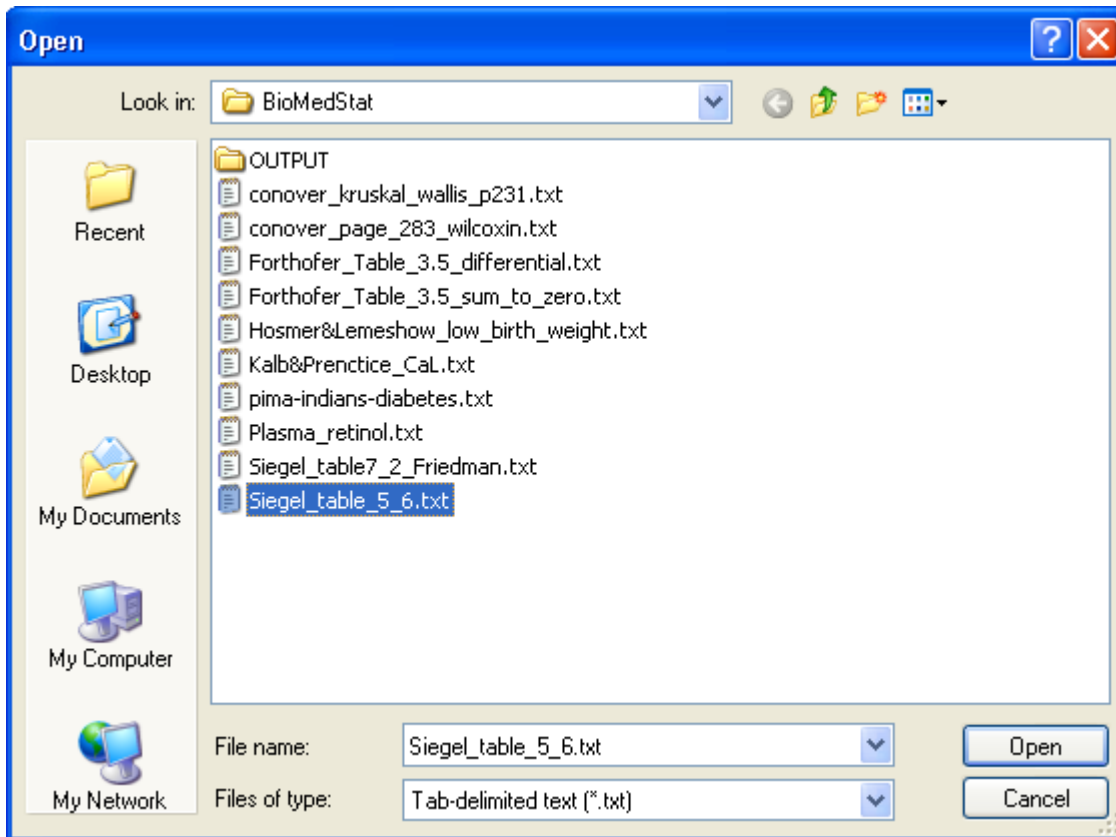
Let us consider the example listed on page 77 of Siegel (Siegel, S., *Non-parametric Statistics for the Behavioral Sciences*, New York, McGraw-Hill, 1956) involving matched data on social perceptiveness scores from identical twin pairs. The hypothesized experiment for this example tests whether or not there is a difference between social perceptiveness scores between twins after one of the twins attends a single term of nursery school and the other stays home. The data are listed below:

<i>Identical twin pair</i>	<i>Social perceptiveness score of twin in nursery school</i>	<i>Social perceptiveness score of twin at home</i>
<i>a</i>	82	63
<i>b</i>	69	42
<i>c</i>	73	74
<i>d</i>	43	37
<i>e</i>	58	51
<i>f</i>	56	43
<i>g</i>	76	80
<i>h</i>	65	82

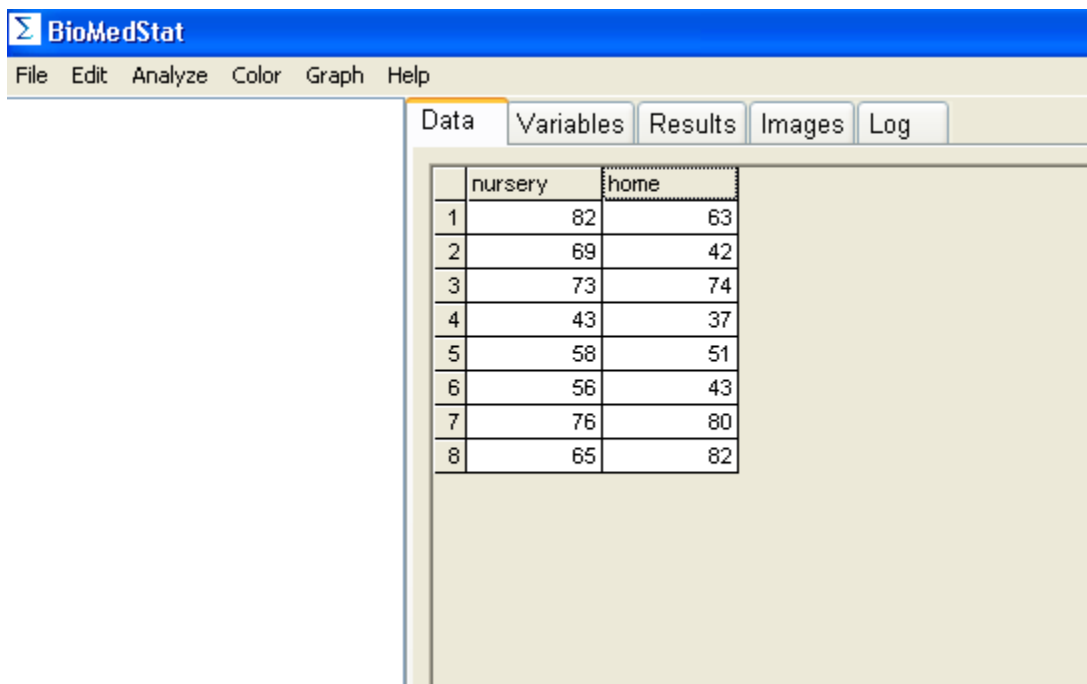
Let's use a "paired T-test" to determine if the average social perceptiveness scores are the same between the twin pairs. Start BioMedStat, and first specify the Open command of the File pull-down menu, then Text, and then specify variable names in first row as shown below:



Next, select the file Siegel_table_5_6.txt, shown as:

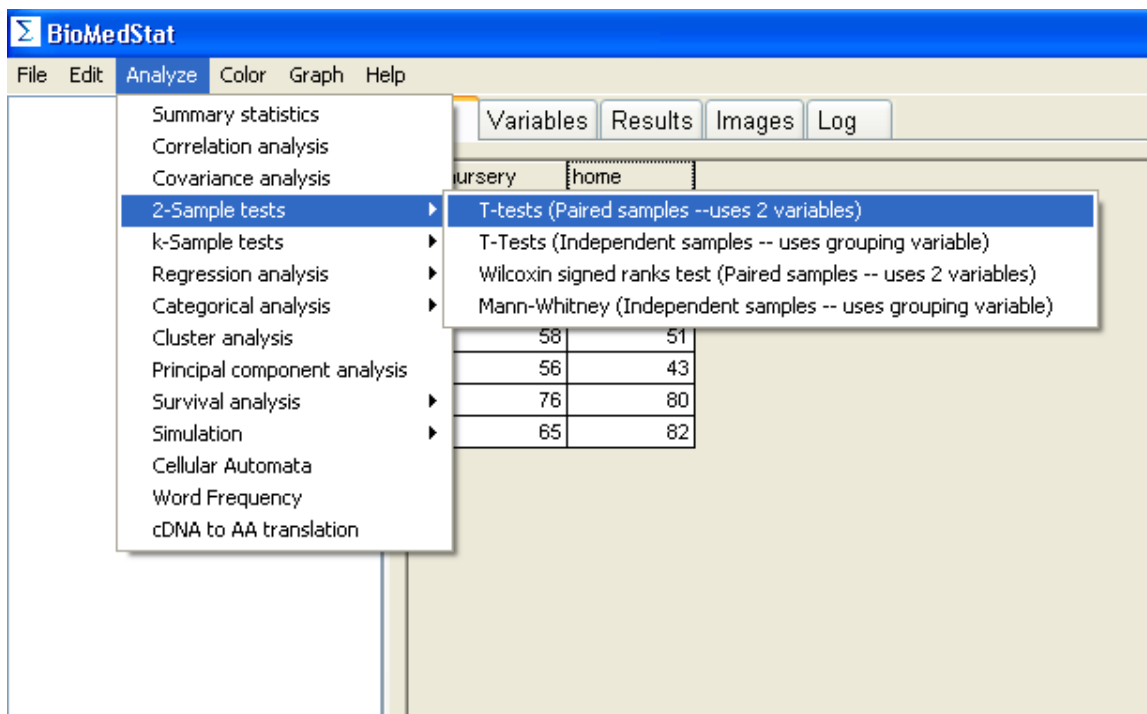


and you will see the data in the viewing spreadsheet as follows:

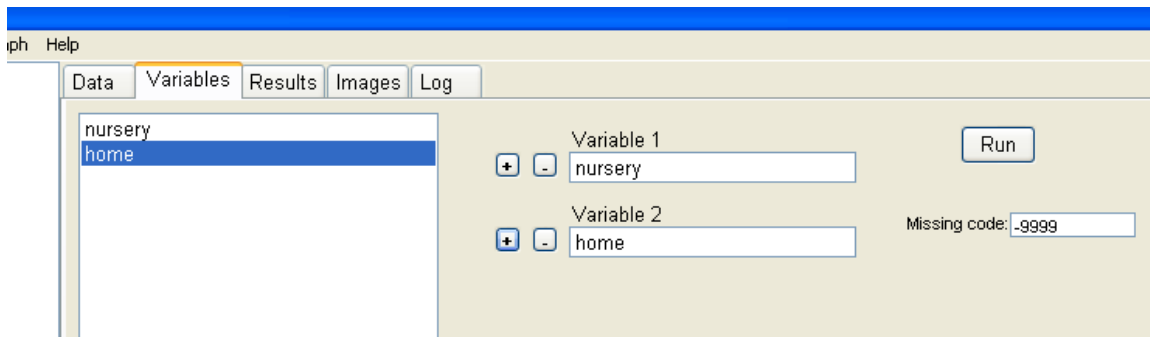


	nursery	home
1	82	63
2	69	42
3	73	74
4	43	37
5	58	51
6	56	43
7	76	80
8	65	82

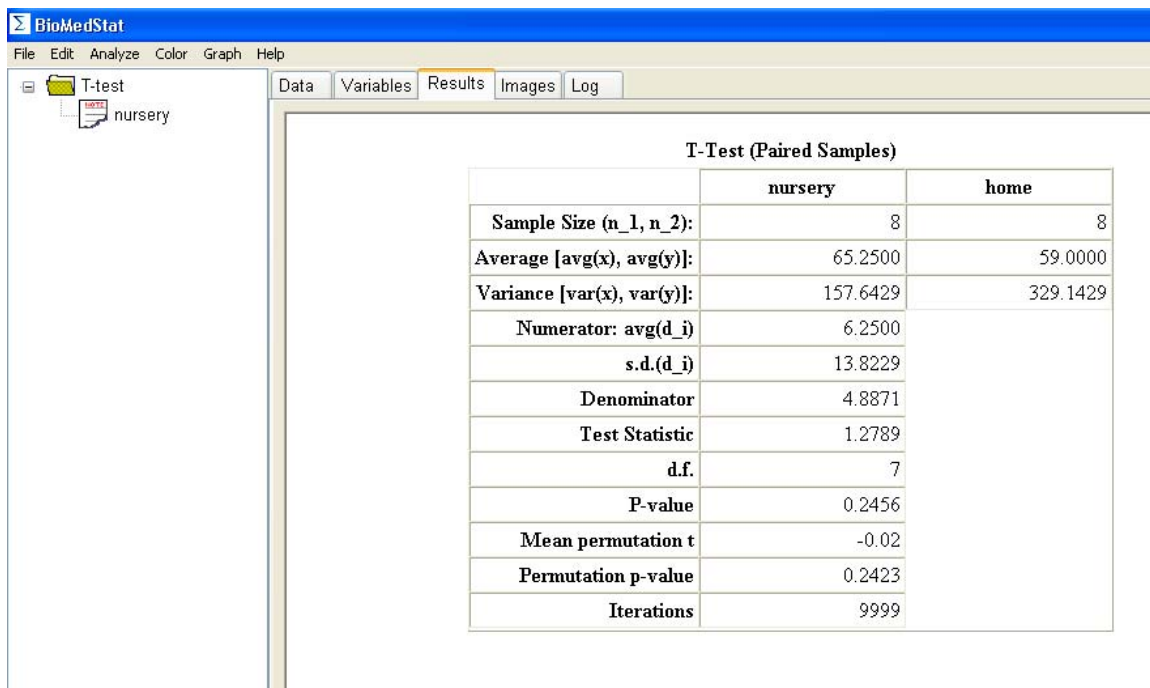
To begin the analysis, select the “T-tests (paired samples)” command of the 2-sample command of the Analyze pull-down menu as follows:



Specify the nursery variable as Variable 1, and home variable as Variable 2 (shown below) and then click on Run.



And the resulting icons and table (after clicking “nursery” icon) will appear as:



The test statistic of 1.2789 for a two-tailed test is compared with a tabled critical value of $t_{0.05; 7} = 2.365$, so the decision rule is to accept the null hypothesis that there is no difference between the average social perceptiveness score among matched pairs of twins. The relevant mathematical formulae are given as:

T-Test (Paired Samples)		
	nursery	home
Sample Size (n_1, n_2):	8	8
Average [avg(x), avg(y)]:	65.2500	59.0000
Variance [var(x), var(y)]:	157.6429	329.1429
Numerator: avg(d_i)	6.2500	\bar{d}
s.d.(d_i)	13.8229	$\sigma_{\bar{d}}$
Denominator	4.8871	$\sigma_{\bar{d}} / \sqrt{n} = 13.8229 / \sqrt{8}$

Test Statistic*	1.2789	$\frac{\bar{d}}{\sigma_{\bar{d}} / \sqrt{n}} = \frac{4.8871}{1.2789}$ $t_{0.05,7} = 2.365$
d.f.	7	
P-value	0.2456	
Mean permutation t	-0.02	
Permutation p-value	0.2423	
Iterations	9999	
<i>*Daniel, W.W. Paired Comparisons (Sect. 6.4). In: Biostatistics: A Foundation for Analysis in the Health Sciences (5th Edition). New York, Wiley (1991).</i>		

While the SPSS (12) benchmark results are as follows:

Paired Samples Statistics

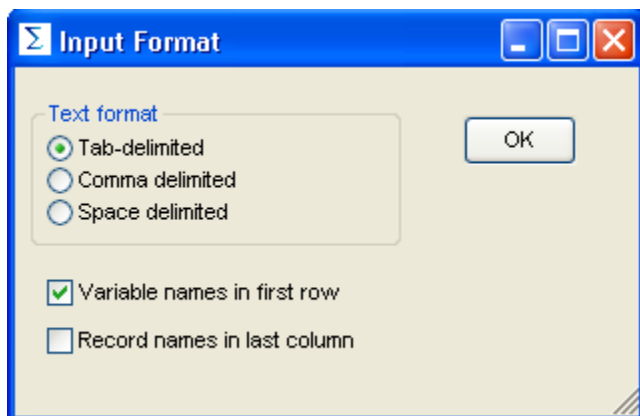
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	nursery	65.250	8	12.5556	4.4391
	home	59.000	8	18.1423	6.4143

Paired Samples Test

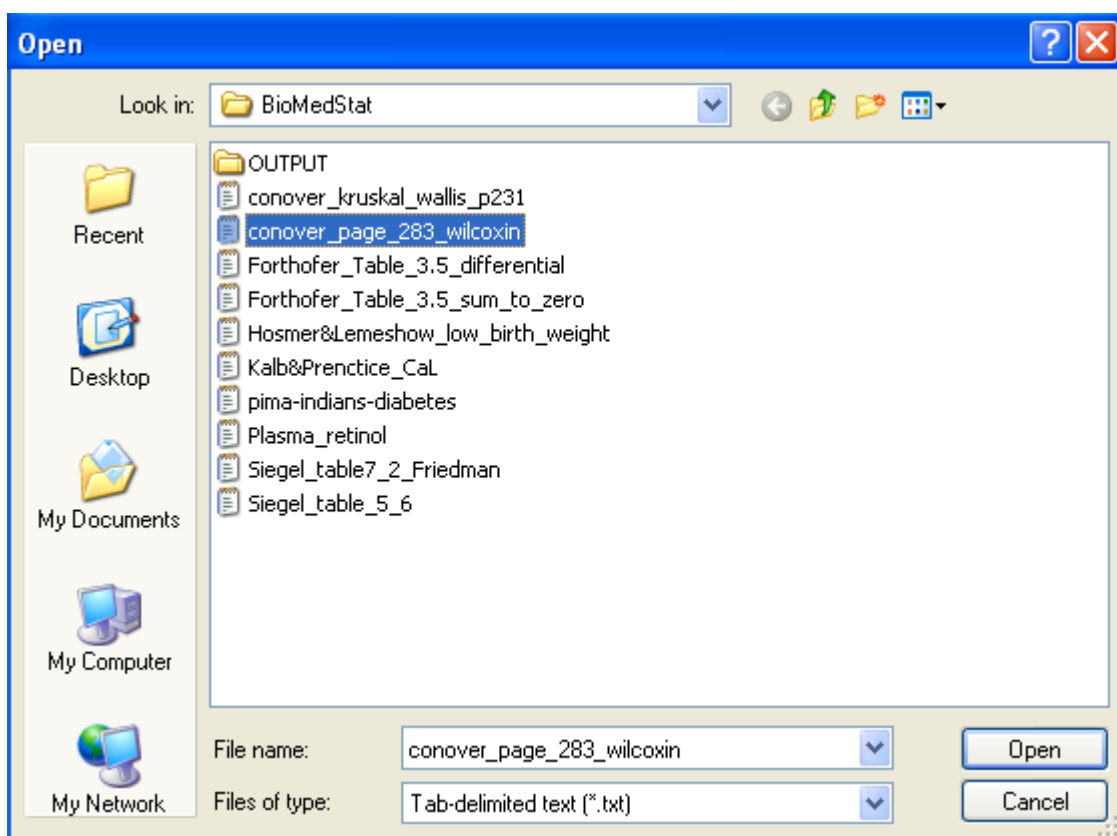
		Paired Differences						
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower	Upper	t	df
Pair 1	nursery - home	6.2500	13.8229	4.8871	-5.3062	17.8062	1.279	7
								Sig. (2-tailed)
								.242

Example 5 – Non-parametric 2-sample paired Wilcoxon signed rank test

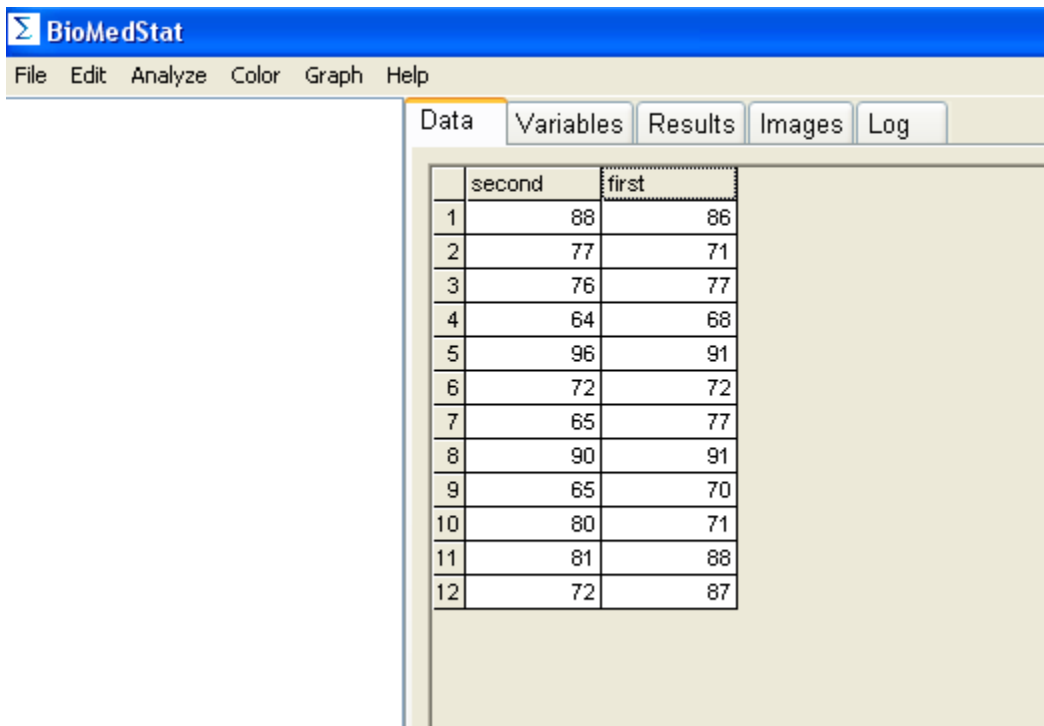
This example will also use an aggressiveness score among 12 sets of identical twins, given on page 283 of Conover, W.J. *The Wilcoxon Signed Rank Test. In: Practical Non-Parametric Statistics (2nd Edition). New York, Wiley (1980).* Start BioMedStat, and first specify the Open command of the File pull-down menu, then Text, and then specify variable names in first row as shown below:



Next, select the file "conover_page_283_Wilcoxon.txt, shown as:



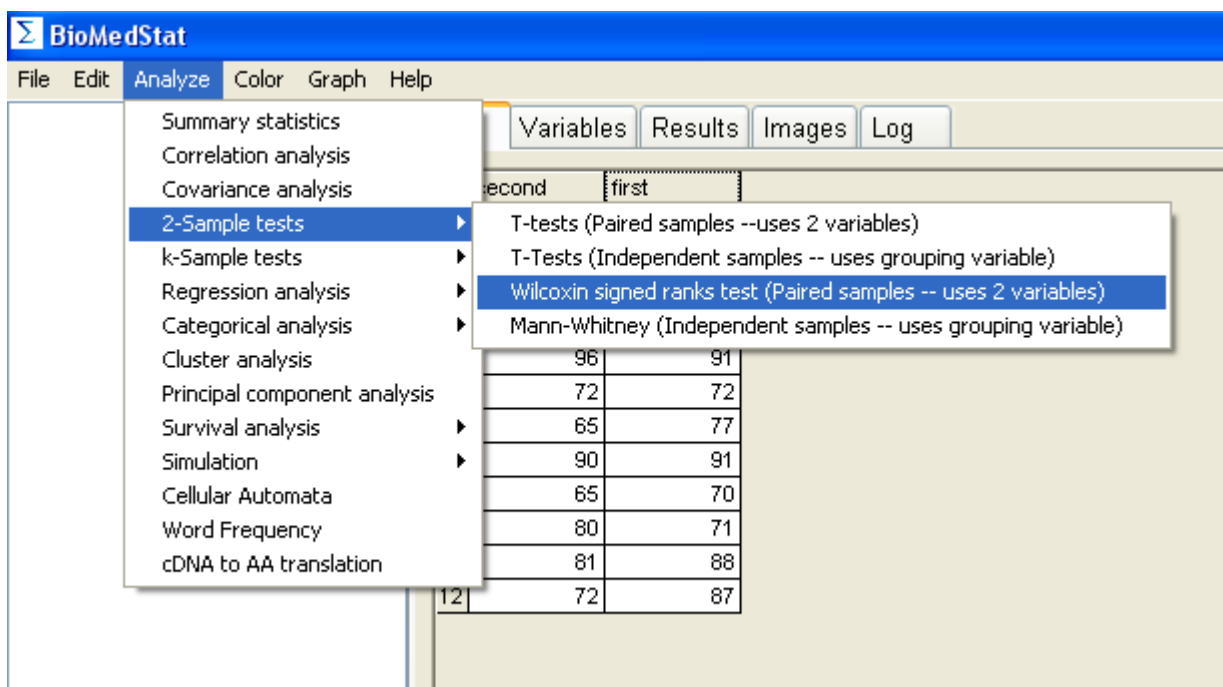
and you will see the data in the viewing spreadsheet as follows:



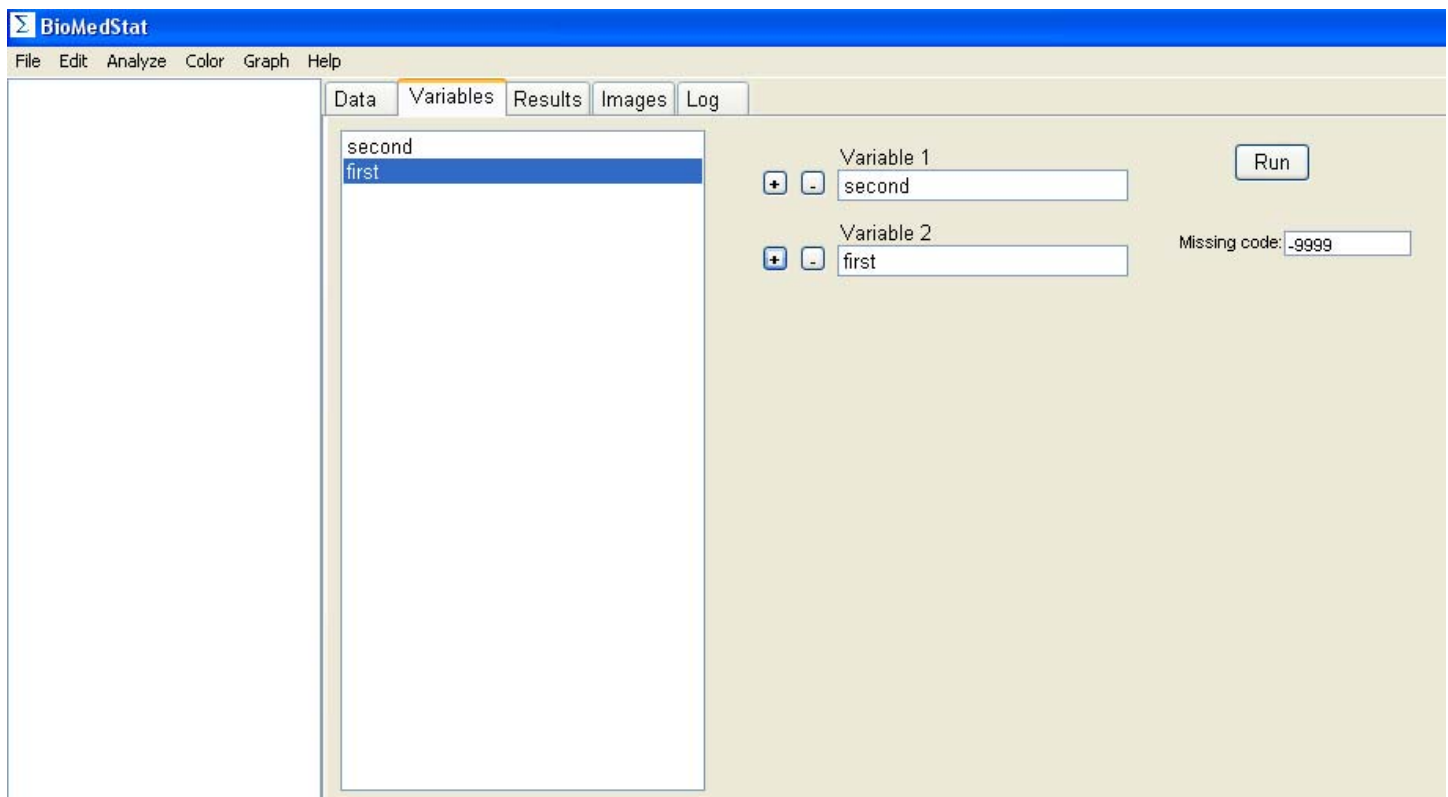
The screenshot shows the BioMedStat software interface. The 'Data' tab is selected, displaying a table with two columns: 'second' and 'first'. The table contains 12 rows of data.

	second	first
1	88	86
2	77	71
3	76	77
4	64	68
5	96	91
6	72	72
7	65	77
8	90	91
9	65	70
10	80	71
11	81	88
12	72	87

To begin the analysis, select the “Wilcoxon signed ranks test” command of the 2-sample command of the Analyze pull-down menu as follows:



Specify the second variable as Variable 1, and first variable as Variable 2 (shown below) and then click on Run.



And the resulting icons and table (after clicking “Wilcoxon” icon) will appear as:

The screenshot shows the BioMedStat software interface with the 'Results' tab selected. On the left, a tree view shows the 'Wilcoxin test' folder expanded, revealing 'Calculations' and 'Wilcoxin test' (which is checked). On the right, a table titled 'Wilcoxin Signed Rank Test' displays the results for the comparison 'second vs. first'.

Wilcoxin Signed Rank Test	
	second vs. first
Sample Size	12
Sum(R _i)	-17.00
Sqrt[Sum(R _i ²)]	505.00
Z-score	-0.76
P-value	0.4899
P-value(Permutation)	0.1190
Avg. perm. Z	-0.5135
Iterations	1000

The test statistic of -0.76 for a two-tailed test is compared with a tabled critical value of -1.96, so the decision rule is to accept the null hypothesis that the two distributions are alike. The relevant mathematical formulae are given as:

Wilcoxon Signed Rank Test		
	Second vs. first	
Sample Size	12	n
Sum(R _i)	-17.00	$\sum_{i=1}^n R_i$
Sqrt[Sum(R _i ²)]	505.00	$\sum_{i=1}^n R_i^2$
Z-score*	-0.76	$\frac{\sum_{i=1}^n R_i}{\sqrt{\sum_{i=1}^n R_i^2}}$
P-value	0.4899	
P-value(Permutation)	0.1190	
Avg. perm. Z	-0.5135	
Iterations	1000	
*Conover, W.J. <i>The Wilcoxon Signed Rank Test. In: Practical Non-Parametric Statistics (2nd Edition). New York, Wiley (1980)</i>		

The SPSS (12) benchmarked results for the Wilcoxon test are:

Ranks

		N	Mean Rank	Sum of Ranks
first - second	Negative Ranks	4(a)	6.13	24.50
	Positive Ranks	7(b)	5.93	41.50
	Ties	1(c)		
	Total	12		

a first < second

b first > second

c first = second

Test Statistics(b)

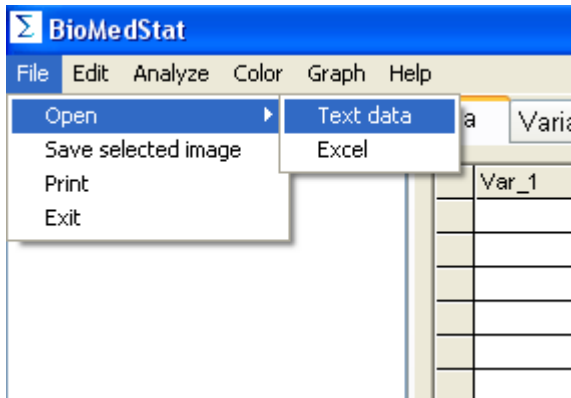
	first - second
Z	-.756(a)
Asymp. Sig. (2-tailed)	.449

a Based on negative ranks.

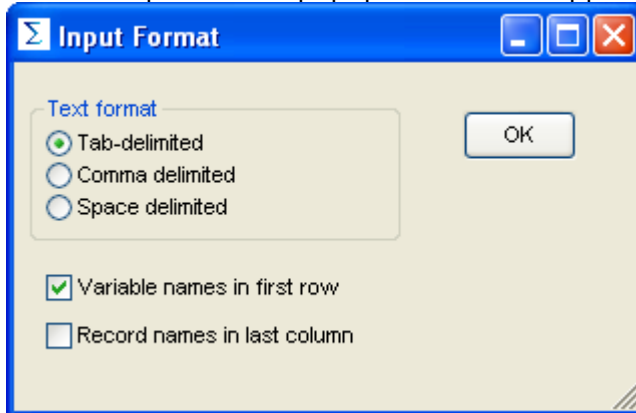
b Wilcoxon Signed Ranks Test

Example 6 – Parametric independent 2-sample t-test

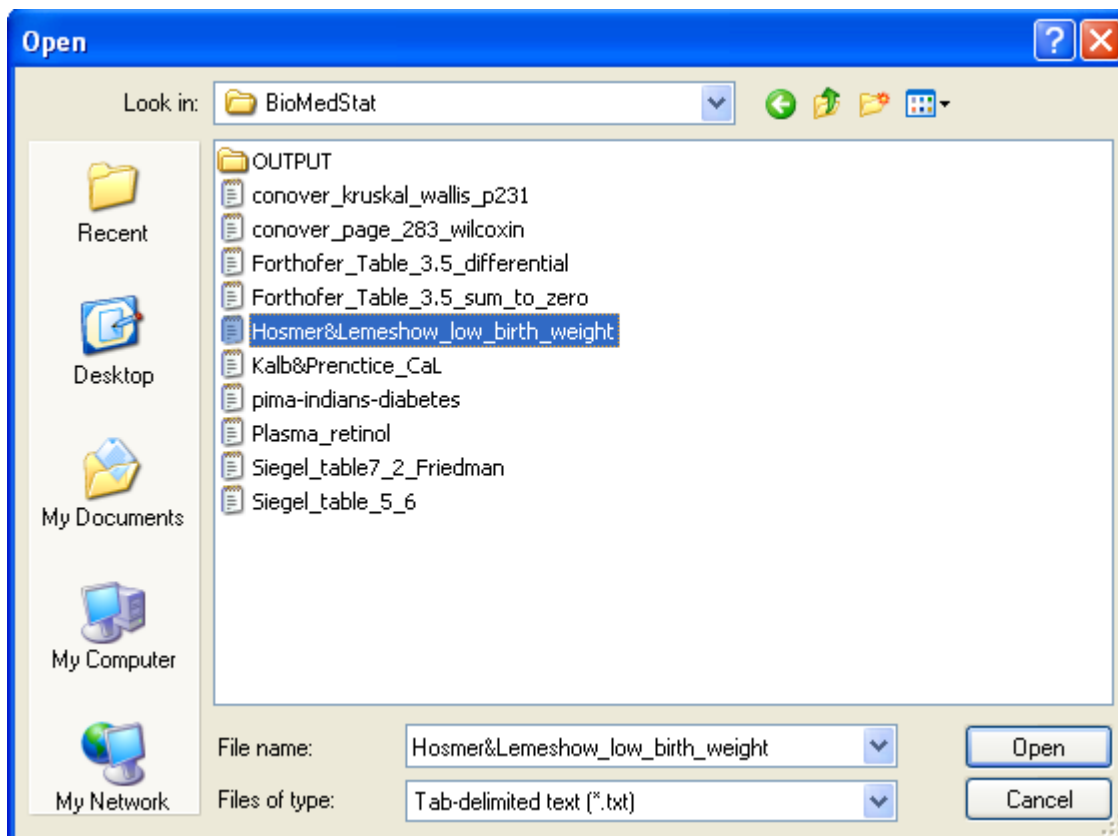
This example uses the Hosmer and Lemeshow low birth weight data described in detail in Example 2. To begin, select File, Open, Text data, as shown below:



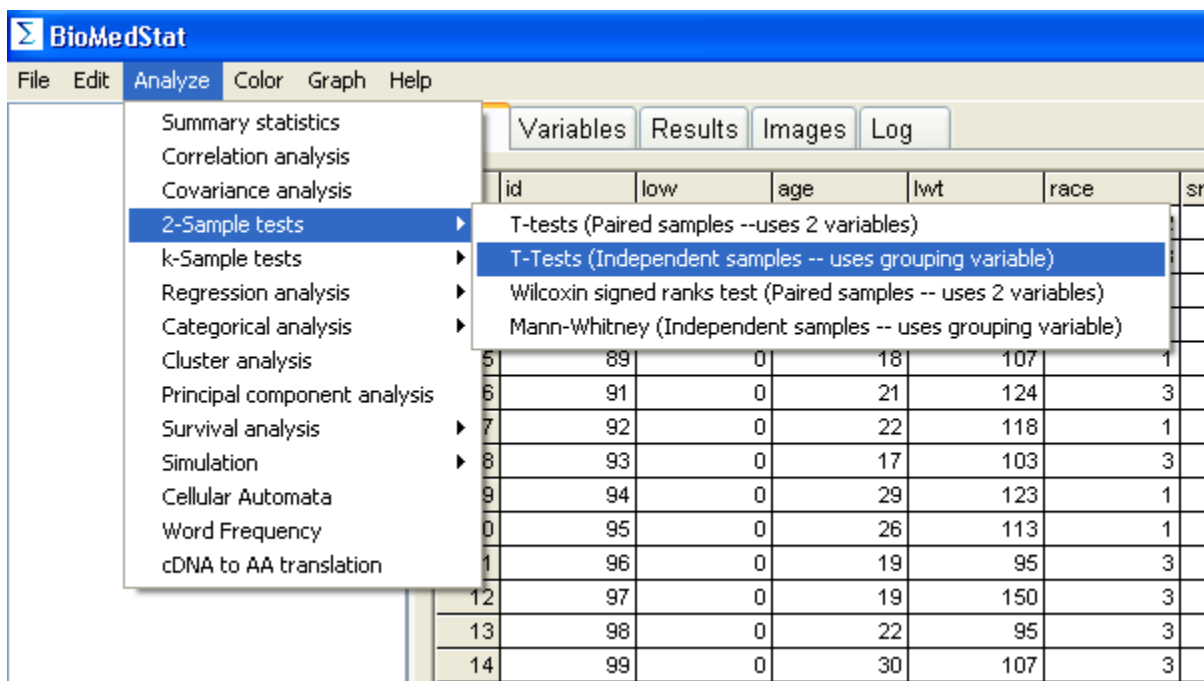
and the Input format popup window will appear:



check the “Variable names in first row” option and click OK. Open the Hosmer and Lemeshow low birth weight data file as shown below:

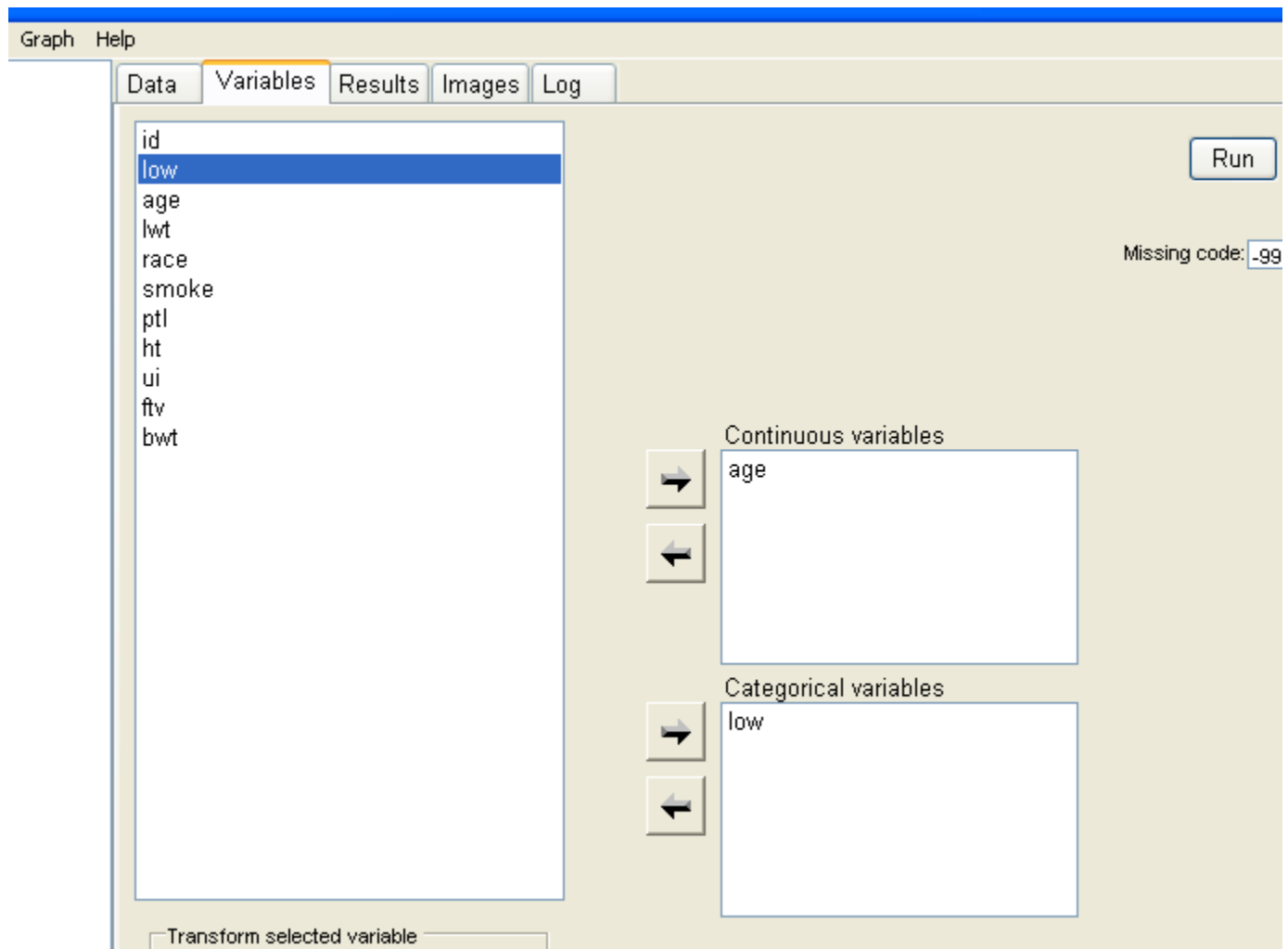



In the Analyze pull-down menu select 2-sample tests, then T-tests (independent samples) shown as:



Recall that in the low birth weight data set, the variable “low” takes on a value of 0 when birth weight \geq 2500g, and 1 when birth weight $<$ 2500g. Using “low” as the grouping variable for the two groups,

let's test if there is a significant age difference between women giving birth to a normal weight baby (low=0) or low weight baby (low=1). Next, select "age" as the continuous variable, and "low" as the categorical variable, shown as:



Then, click Run and click on the  age icon and you will observe the following tables containing results:

T-Test (Independent samples)		
Test var:	Age	Equal variances assumed
Grouping var:	Low	
	1	2
Sample Size (n_1, n_2):	130	59
Average [avg(x), avg(y)]:	23.6615	22.3051
Variance [var(x), var(y)]:	31.1869	20.3536
Equal variance test:	F-ratio: 0.2916(1,187)	Prob: 0.4102

	$W = \frac{(n-k) \sum_{i=1}^k n_i (Z_{i.} - Z_{..})^2}{(k-1) \sum_{i=1}^k \sum_{j=1}^{n_i} (Z_{ij} - Z_{i.})^2}$ $\sim F(k-1, n-k)$	
Numerator: avg(x)-avg(y)	1.3565	$\bar{X}_1 - \bar{X}_2$
Pooled common variance	27.8268	$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$
Denominator	0.8281	$s_p^2 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
Test Statistic	1.6381	$\frac{\bar{X}_1 - \bar{X}_2}{s_p^2 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$
d.f.	187	$n_1 + n_2 - 2$
P-value	0.1072	
Mean permutation t	0.00	
Permutation p-value	0.1028	
Iterations	9999	

T-Test (Independent samples)		
Test var:	age	Unequal variances assumed
Grouping var:	low	
	1	2
Sample Size (n_1, n_2):	130	59
Average [avg(x), avg(y)]:	23.6615	22.3051
Variance [var(x), var(y)]:	31.1869	20.3536
Numerator: avg(x)-avg(y)	1.3565	$\bar{X}_1 - \bar{X}_2$
Denominator	0.7648	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Test Statistic	1.7737	$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
d.f. (Dixon & Massey correction)	137	$df' = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{(s_1^2/n_1)^2}{n_1} + \frac{(s_2^2/n_2)^2}{n_2}}$
P-value	0.0809	
Mean permutation t	0.04	

Permutation p-value	0.0766
Iterations	9999
<i>*Daniel, W.W. The t Distribution (Sect. 5.6). In: Biostatistics: A Foundation for Analysis in the Health Sciences. New York, Wiley (1974).</i>	

The first table contains results for Levene's variance ratio test for equal variances, as well as the t-test results assuming equal variances. The equal variance test has a resulting p-value of 0.42, so we accept the null hypothesis that the variances are equal. The test statistic for the t-test (assuming equal variances) is 1.638, for which the p-value is 0.11, so for these data there is not a significant difference in age between mothers giving birth to low and normal weight babies.

The SPSS(12) benchmarked results are:

Group Statistics

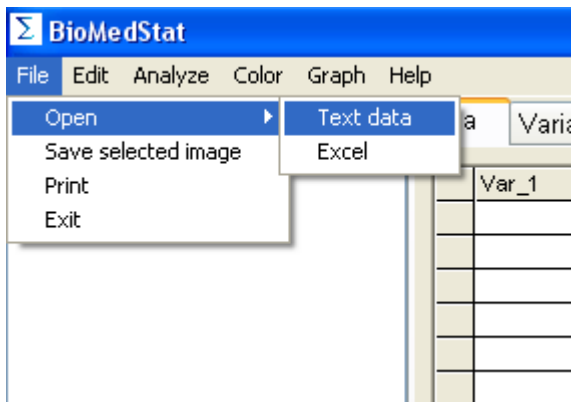
	low	N	Mean	Std. Deviation	Std. Error Mean
age	0	130	23.662	5.5845	.4898
	1	59	22.305	4.5115	.5873

Independent Samples Test

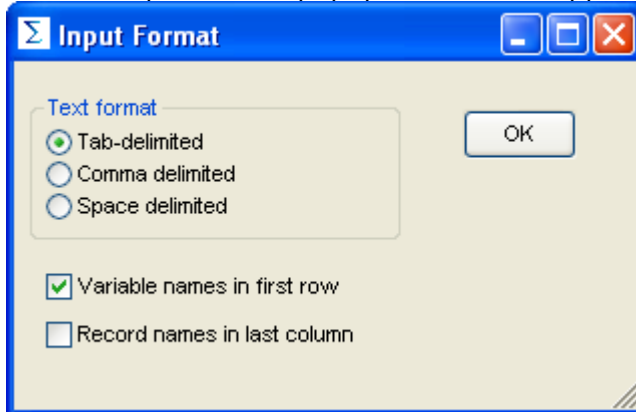
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower Upper
age	Equal variances assumed	3.429	.066	1.638	187	.103	1.3565	.8281	-.2771 2.9901
	Equal variances not assumed			1.774	136.941	.078	1.3565	.7648	-.1558 2.8688

Example 7 – Non-parametric independent 2-sample Mann-Whitney U Test

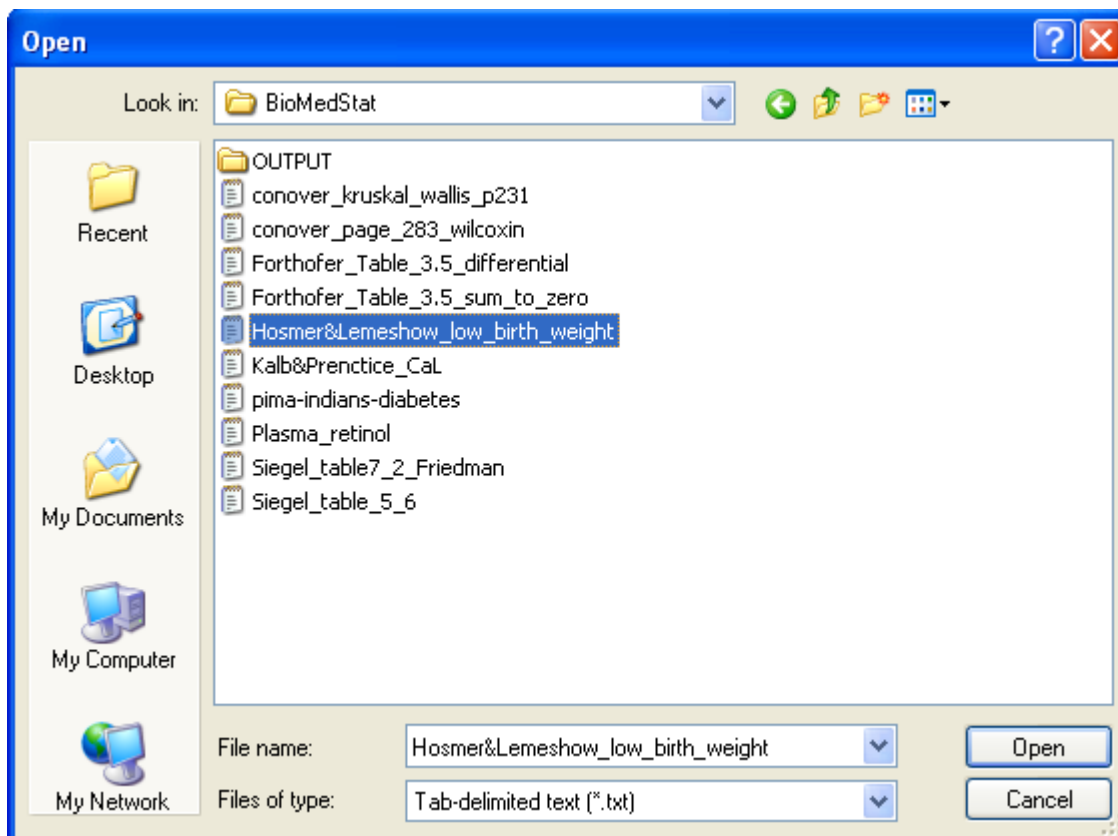
This example uses the Hosmer and Lemeshow low birth weight data described in detail in Example 2. To begin, select File, Open, Text data, as shown below:



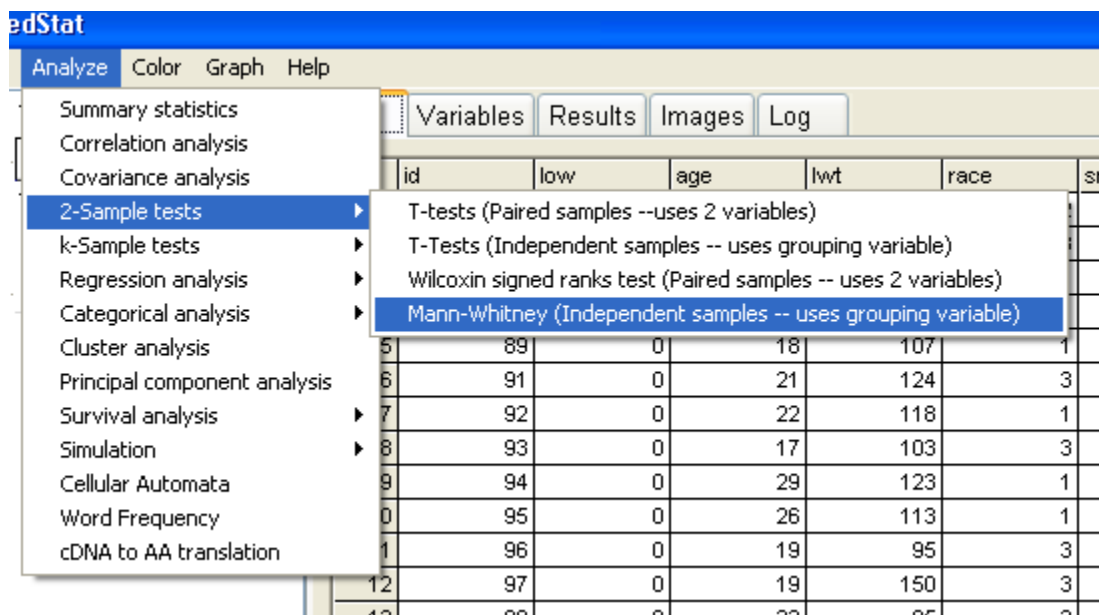
and the Input format popup window will appear:



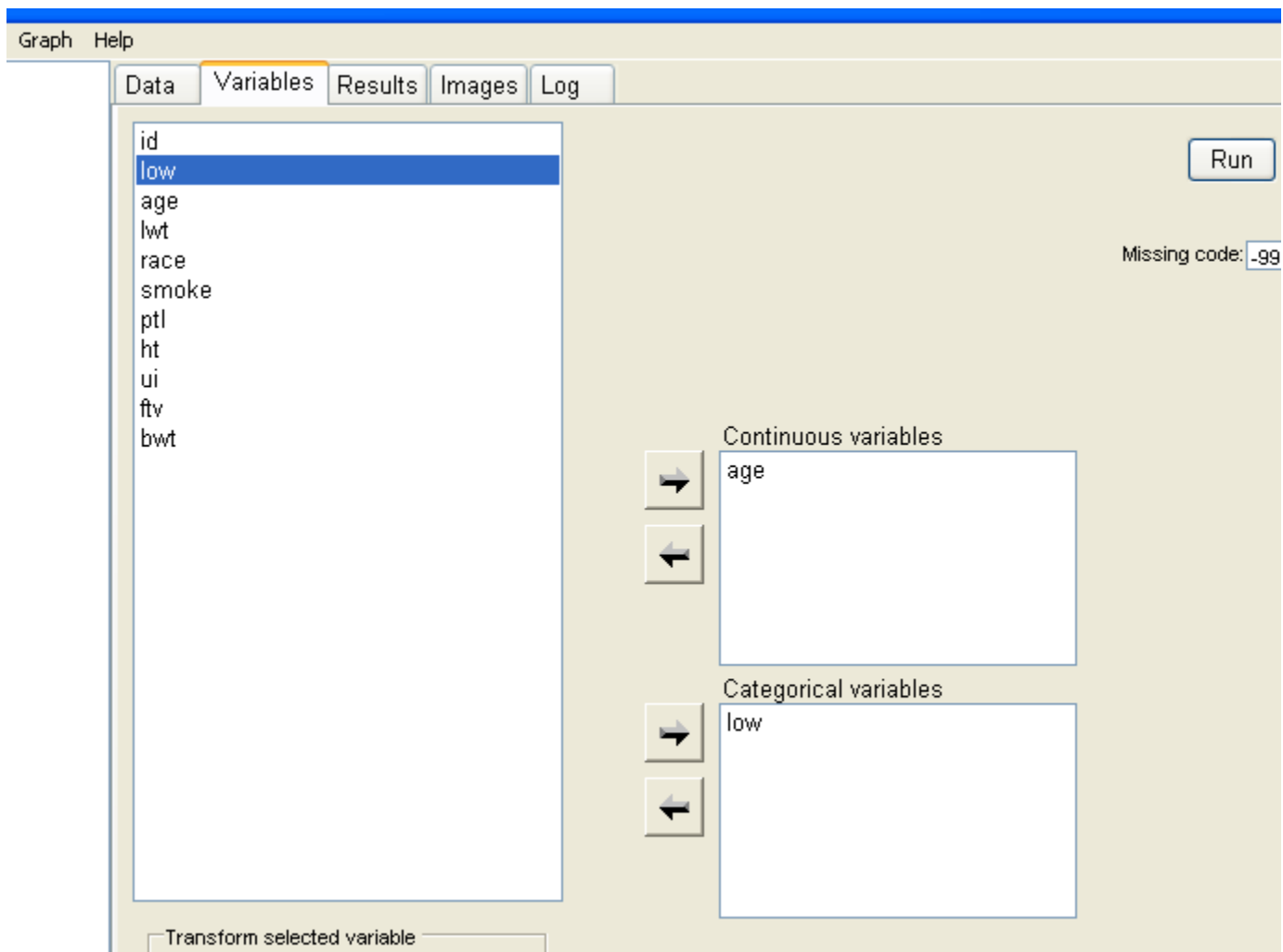
check the “Variable names in first row” option and click OK. Open the Hosmer and Lemeshow low birth weight data file as shown below:




In the Analyze pull-down menu select 2-sample tests, then T-tests (independent samples) shown as:



Next, select "age" as the continuous variable, and "low" as the categorical variable, shown as:



Then, click Run and click on the  age icon and you will observe the following tables containing results:

Mann-Whitney Test		
Test var:	age	
Grouping var:	low	
	1	2
Sample Size	130	59
R	12753.00	5202.00
U1,U2	3432.00	4238.00
U based on Min(U1,U2)	3432.00	
Mean U	3835.00	
s.d.(U)	348.48	
Z-score	-1.16	
P-value	0.3014	
P-value(Permutation)	0.2467	

Avg. perm. Z	-0.7980	
Iterations	9999	

Benchmark results from SPSS(12) are:

Ranks

	low	N	Mean Rank	Sum of Ranks
age	0	130	98.10	12753.00
	1	59	88.17	5202.00
	Total	189		

Test Statistics(a)

	age
Mann-Whitney U	3432.000
Wilcoxon W	5202.000
Z	-1.159
Asymp. Sig. (2-tailed)	.247

a. Grouping Variable: low

Example 8 – Parametric independent k-sample Analysis of Variance

This example focuses on a one-way analysis of variance (ANOVA) to determine statistically significant differences between age among three racial groups in the Hosmer and Lemeshow low birth weight data described in Example 2. In the low birth weight study, the categorical variable “race” was coded to represent 1-white, 2-black, and 3-other. In the following procedures, we will use BioMedStat to test for a significant difference in age among the three groups using an independent k-sample F-test as part of the ANOVA.

To start, open the Hosmer and Lemeshow data set as described in Example 2. Under the “k-sample” command of the Analyze pull-down menu specify “F-test (independent samples),” shown as follows:

edStat

Analyze Color Graph Help

Variables Results Images Log

Summary statistics
Correlation analysis
Covariance analysis
2-Sample tests
k-Sample tests
Regression analysis
Categorical analysis
Cluster analysis
Principal component analysis
Survival analysis
Simulation
Cellular Automata
Word Frequency
cDNA to AA translation

F-test (Paired Samples -- uses k variables)
F-test (Independent Samples - uses grouping variable)
Friedman test (Paired - uses k variables)
Kruskal-Wallis test (Independent Samples -- uses grouping variable)

	id	low	age	lwt	race	smoke
1	85	0	19	182	2	
6	91	0	21	124	3	
7	92	0	22	118	1	
8	93	0	17	103	3	
9	94	0	29	123	1	
10	95	0	26	113	1	
11	96	0	19	95	3	
12	97	0	19	150	3	
13	98	0	22	95	3	

Select age as the continuous variable and race as the categorical variable, and then click on Run.

elp

Data Variables Results Images Log

id
low
age
lwt
race
smoke
ptl
ht
ui
ftv
bwt

Continuous variables
age

Categorical variables
race

Transform selected variable

The results of the ANOVA re shown as follows:

Analysis of Variance			
Test var:		age	
Grouping var:		race	
Group:	Sample size	Mean	s.d.
1	96	24.292	5.655
2	26	21.538	5.109
3	67	22.388	4.536
ANOVA Table		SS	MSE
Treatment		230.08	115.04
Error		5048.21	27.14
P-value		0.0158	
P-value(Permutation)		0.0810	
Avg. perm. F-ratio		2.0286	
Iterations		1000	

SPSS(ANOVA) results for the same ANOVA are:

Descriptives

age

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	96	24.292	5.6548	.5771	23.146	25.437	14	45
2	26	21.538	5.1087	1.0019	19.475	23.602	15	35
3	67	22.388	4.5359	.5541	21.282	23.494	14	33
Total	189	23.238	5.2987	.3854	22.478	23.998	14	45

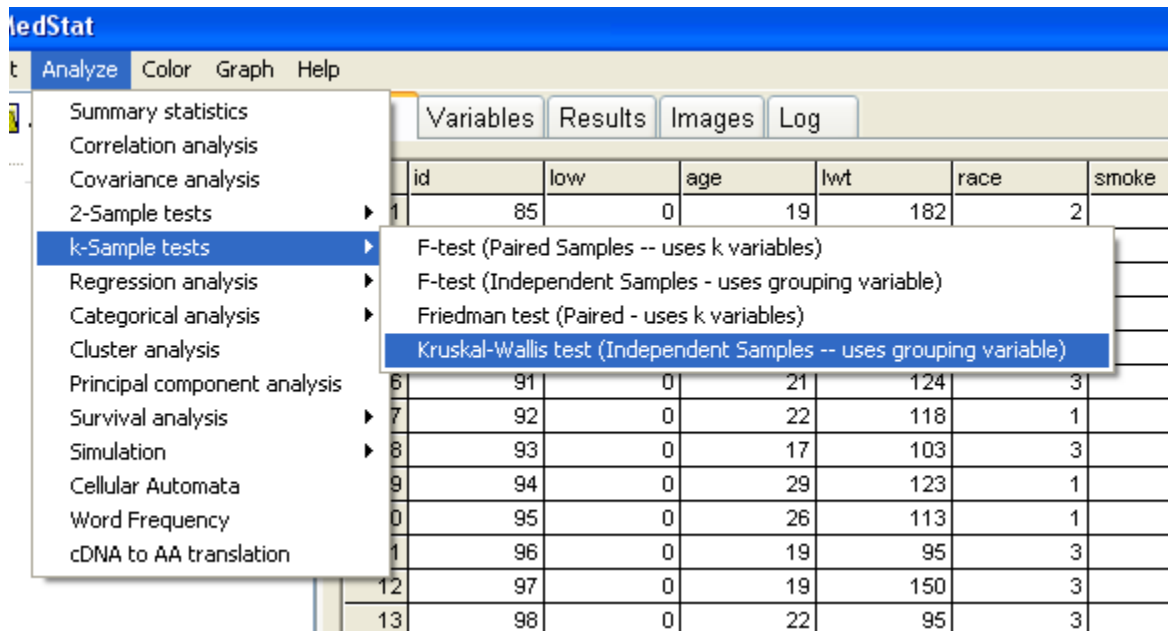
ANOVA

age

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	230.080	2	115.040	4.239	.016
Within Groups	5048.205	186	27.141		
Total	5278.286	188			

Example 9 – Non-parametric independent k-sample Kruskal-Wallis Test

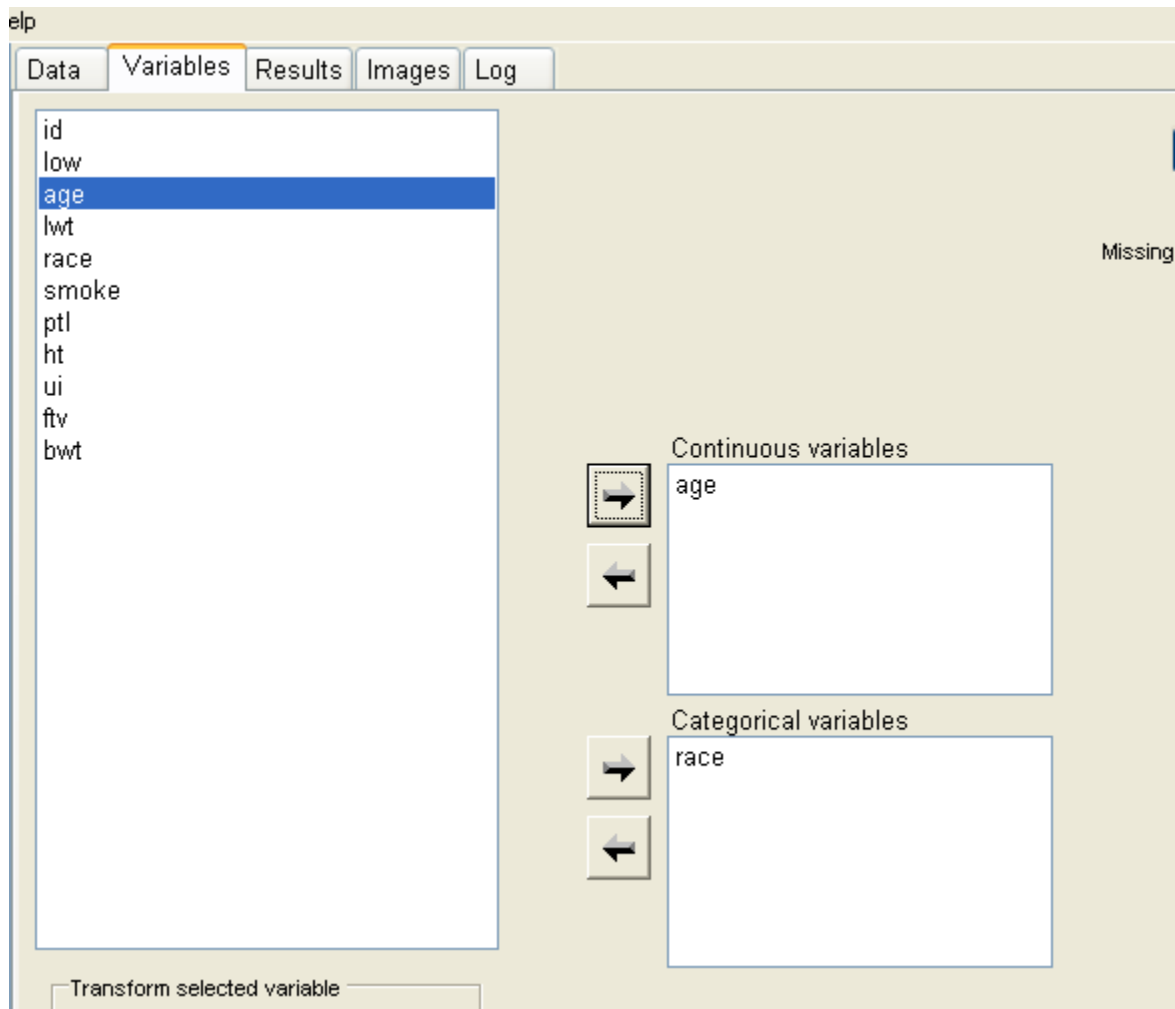
Open the Hosmer and Lemeshow data (see Example 2), and specify a Kruskal-Wallis test:



The screenshot shows the MedStat software interface. The 'Analyze' menu is open, and the 'k-Sample tests' option is selected. A sub-menu is displayed, showing the 'Kruskal-Wallis test (Independent Samples -- uses grouping variable)' as the chosen test. The background shows a data table with columns: id, low, age, lwt, race, and smoke.

	id	low	age	lwt	race	smoke
1	85	0	19	182	2	
6	91	0	21	124	3	
7	92	0	22	118	1	
8	93	0	17	103	3	
9	94	0	29	123	1	
10	95	0	26	113	1	
11	96	0	19	95	3	
12	97	0	19	150	3	
13	98	0	22	95	3	

Next, specify age as the continuous variable and race as the categorical variable as



And then click on Run. The results of the Kruskal-Wallis test are shown below:

Kruskal-Wallis Test			
Test var:	age		
Grouping var:	race		
Group:	Sample size, n(j)	Sum of ranks, R(j)	Average rank, R(j)/n(j)
1	96	10055.50	104.74
2	26	1968.00	75.69
3	67	5931.50	88.53
Chi-square (d.f.)	7.22 (2)		
P-value	0.0270		
P-value(Permutation)	0.0190		
Avg. perm. Chi-square	1.9026		
Iterations	1000		

Benchmark results for SPSS(12) are:

Ranks

	race	N	Mean Rank
age	1	96	104.74
	2	26	75.69
	3	67	88.53
	Total	189	

Test Statistics(a,b)

	age
Chi-Square	7.251
df	2
Asymp. Sig.	.027

a Kruskal Wallis Test

b Grouping Variable: race

Example 10 – Pearson product moment and Spearman rank correlation

Using the Hosmer and Lemeshow data again, below are the BioMedStat results for correlation analysis.

Pearson Correlation Matrix

	low	age	lwt	race	smoke	ptl	ht	Ui	ftv
age	-0.12								
lwt	-0.17*	0.18*							
race	0.14	-0.17*	-0.17*						
smoke	0.16*	-0.04	-0.04	-0.34**					
ptl	0.20**	0.07	-0.14	0.01	0.19**				
ht	0.15*	-0.02	0.24**	0.02	0.01	-0.02			
ui	0.17*	-0.08	-0.15*	0.05	0.06	0.23**	-0.11		
ftv	-0.06	0.22**	0.14	-0.10	-0.03	-0.04	-0.07	-0.06	
bwt	-0.79**	0.09	0.19**	-0.20**	-0.19**	-0.15*	-0.15*	-0.28**	0.06

Significant, negative:

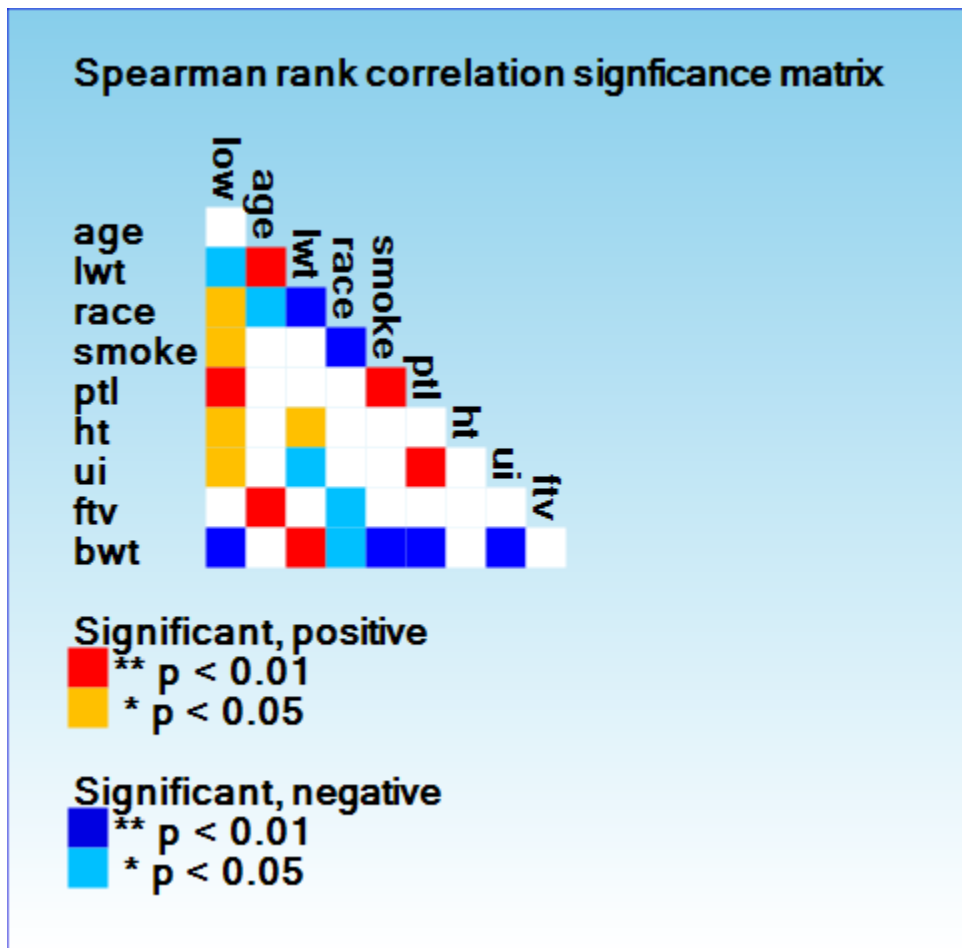
** p < 0.01

* p < 0.05

Significant, positive:

****** $p < 0.01$

***** $p < 0.05$



SPSS(12) benchmarking results for Pearson correlation are:

Correlations

		low	age	lwt	race	smoke	ptl	ht	ui
low	Pearson	1	-.119	-.170(*)	.138	.161(*)	.196(**)	.152(*)	.169(*)
	Correlation								
	Sig. (2-tailed)		.103	.020	.059	.027	.007	.036	.020
age	N	189	189	189	189	189	189	189	189
	Pearson	-.119	1	.180(*)	-.173(*)	-.044	.072	-.016	-.075
	Correlation								
lwt	Sig. (2-tailed)	.103		.013	.017	.545	.328	.829	.304
	N	189	189	189	189	189	189	189	189
	Pearson	-.170(*)	.180(*)	1	-.165(*)	-.044	-.140	.236(**)	-.153(*)
race	Correlation								
	Sig. (2-tailed)	.020	.013		.023	.546	.055	.001	.036
	N	189	189	189	189	189	189	189	189
smoke	Pearson	.138	-.173(*)	-.165(*)	1	-.339(**)	.008	.020	.054
	Correlation								

smoke	Sig. (2-tailed)	.059	.017	.023		.000	.914	.785	.464
	N	189	189	189	189	189	189	189	189
	Pearson Correlation	.161(*)	-.044	-.044	-.339(**)	1	.188(**)	.013	.062
ptl	Sig. (2-tailed)	.027	.545	.546	.000		.010	.855	.395
	N	189	189	189	189	189	189	189	189
	Pearson Correlation	.196(**)	.072	-.140	.008	.188(**)	1	-.015	.228(**)
ht	Sig. (2-tailed)	.007	.328	.055	.914	.010		.833	.002
	N	189	189	189	189	189	189	189	189
	Pearson Correlation	.152(*)	-.016	.236(**)	.020	.013	-.015	1	-.109
ui	Sig. (2-tailed)	.036	.829	.001	.785	.855	.833		.137
	N	189	189	189	189	189	189	189	189
	Pearson Correlation	.169(*)	-.075	-.153(*)	.054	.062	.228(**)	-.109	1
ftv	Sig. (2-tailed)	.020	.304	.036	.464	.395	.002	.137	
	N	189	189	189	189	189	189	189	189
	Pearson Correlation	-.063	.215(**)	.141	-.098	-.028	-.044	-.072	-.060
bwt	Sig. (2-tailed)	.389	.003	.054	.178	.702	.544	.322	.416
	N	189	189	189	189	189	189	189	189
	Pearson Correlation	-.785(**)	.090	.186(*)	-.196(**)	-.189(**)	-.155(*)	-.146(*)	-.283(**)
	Sig. (2-tailed)	.000	.219	.010	.007	.009	.034	.045	.000
	N	189	189	189	189	189	189	189	189

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

The Spearman rank correlation results from BioMedStat are:

Spearman Correlation Matrix									
	low	age	Lwt	race	smoke	ptl	ht	ui	ftv
age	-0.08								
lwt	-0.18*	0.19**							
race	0.14*	-0.15*	-0.19**						
smoke	0.16*	-0.05	-0.09	-0.34**					
ptl	0.26**	0.12	-0.11	0.04	0.19**				
ht	0.15*	-0.01	0.17*	0.02	0.01	0.00			
ui	0.17*	-0.08	-0.18*	0.05	0.06	0.19**	-0.11		
ftv	-0.09	0.23**	0.09	-0.15*	-0.08	-0.01	-0.08	-0.06	
bwt	-0.80**	0.06	0.25**	-0.18*	-0.20**	-0.20**	-0.11	-0.27**	0.07

Significant, negative:

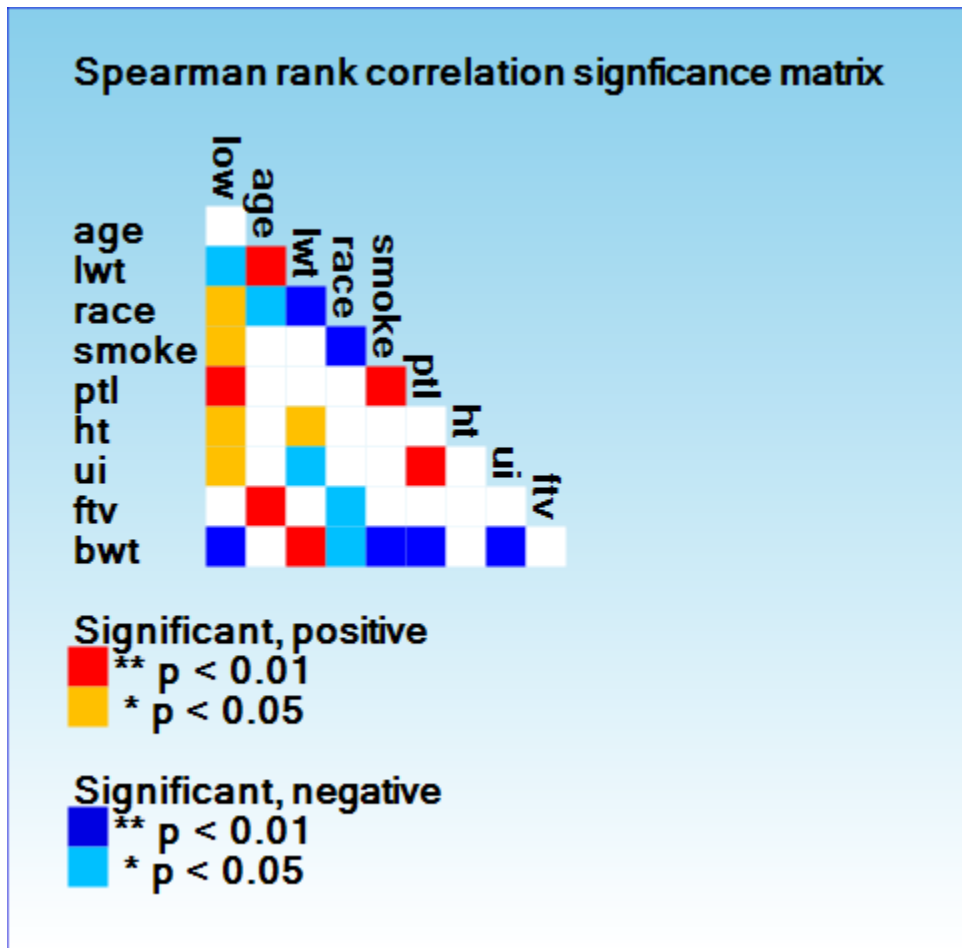
** p < 0.01

* p < 0.05

Significant, positive:

**** p < 0.01**

*** p < 0.05**



while the SPSS(12) benchmarking results for Spearman rank correlation are:

Correlations

Spearman		low	age	lwt	race	smoke	ptl	ht	ui
low	Correlation	1.000	-.085	-.182(*)	.142	.161(*)	.261(**)	.152(*)	.169(*)
	Coefficient								
	Sig. (2-tailed)		.248	.012	.051	.027	.000	.036	.020
age	N	189	189	189	189	189	189	189	189
	Correlation	-.085	1.000	.186(*)	-.152(*)	-.048	.119	-.005	-.075
	Coefficient								
lwt	Sig. (2-tailed)	.248	.010	.037	.515	.104	.944	.302	
	N	189	189	189	189	189	189	189	189
	Correlation	-.182(*)	.186(*)	1.000	-.189(**)	-.085	-.111	.170(*)	-.175(*)
race	Coefficient								
	Sig. (2-tailed)	.012	.010	.009	.243	.129	.019	.016	
	N	189	189	189	189	189	189	189	189
smoke	Correlation	.142	-.152(*)	-.189(**)	1.000	-.338(**)	.036	.024	.051
	Coefficient								

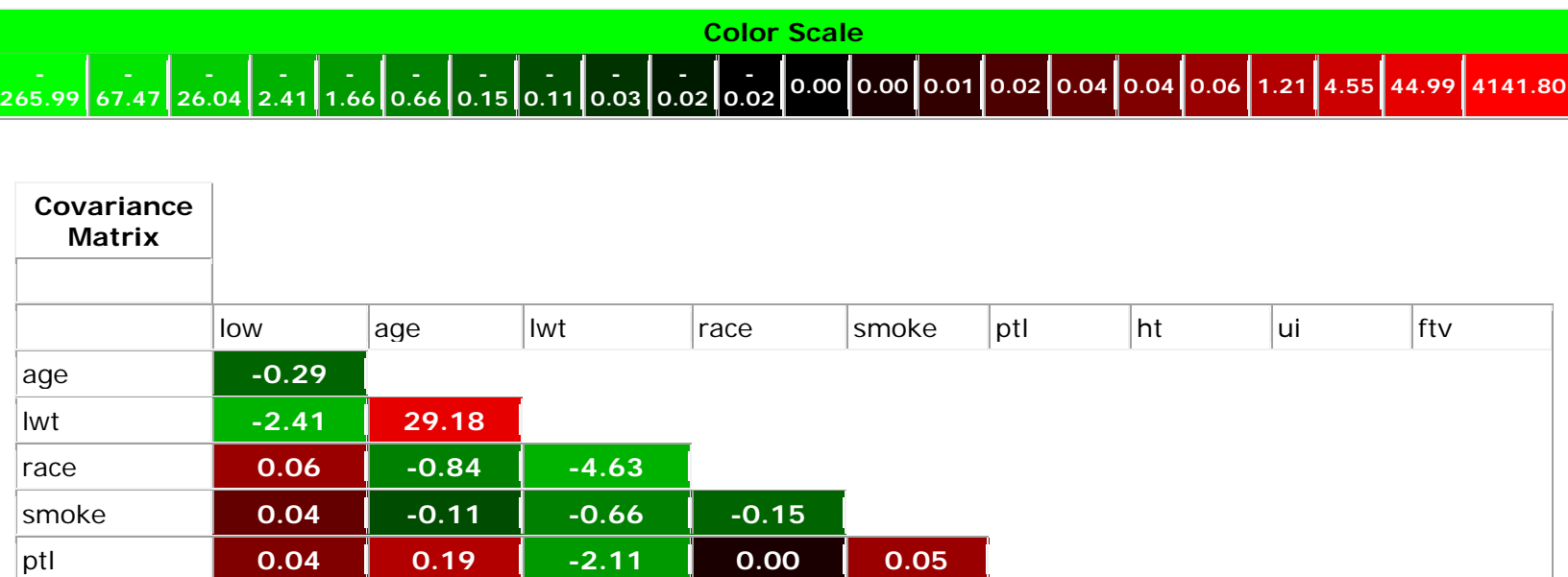
	Coefficient								
	Sig. (2-tailed)	.051	.037	.009	.	.000	.623	.742	.482
	N	189	189	189	189	189	189	189	189
smoke	Correlation								
	Coefficient	.161(*)	-.048	-.085	-.338(**)	1.000	.187(**)	.013	.062
	Sig. (2-tailed)	.027	.515	.243	.000	.	.010	.855	.395
	N	189	189	189	189	189	189	189	189
ptl	Correlation								
	Coefficient	.261(**)	.119	-.111	.036	.187(**)	1.000	.002	.194(**)
	Sig. (2-tailed)	.000	.104	.129	.623	.010	.	.980	.008
	N	189	189	189	189	189	189	189	189
ht	Correlation								
	Coefficient	.152(*)	-.005	.170(*)	.024	.013	.002	1.000	-.109
	Sig. (2-tailed)	.036	.944	.019	.742	.855	.980	.	.137
	N	189	189	189	189	189	189	189	189
ui	Correlation								
	Coefficient	.169(*)	-.075	-.175(*)	.051	.062	.194(**)	-.109	1.000
	Sig. (2-tailed)	.020	.302	.016	.482	.395	.008	.137	.
	N	189	189	189	189	189	189	189	189
ftv	Correlation								
	Coefficient	-.086	.234(**)	.089	-.155(*)	-.083	-.014	-.080	-.056
	Sig. (2-tailed)	.239	.001	.223	.034	.254	.845	.275	.442
	N	189	189	189	189	189	189	189	189
bwt	Correlation								
	Coefficient	-.803(**)	.061	.248(**)	-.184(*)	-.196(**)	-.204(**)	-.115	-.266(**)
	Sig. (2-tailed)	.000	.404	.001	.011	.007	.005	.117	.000
	N	189	189	189	189	189	189	189	189

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Example 11 – Covariance analysis

Using the Hosmer and Lemeshow data again, below are the BioMedStat results for covariance analysis.



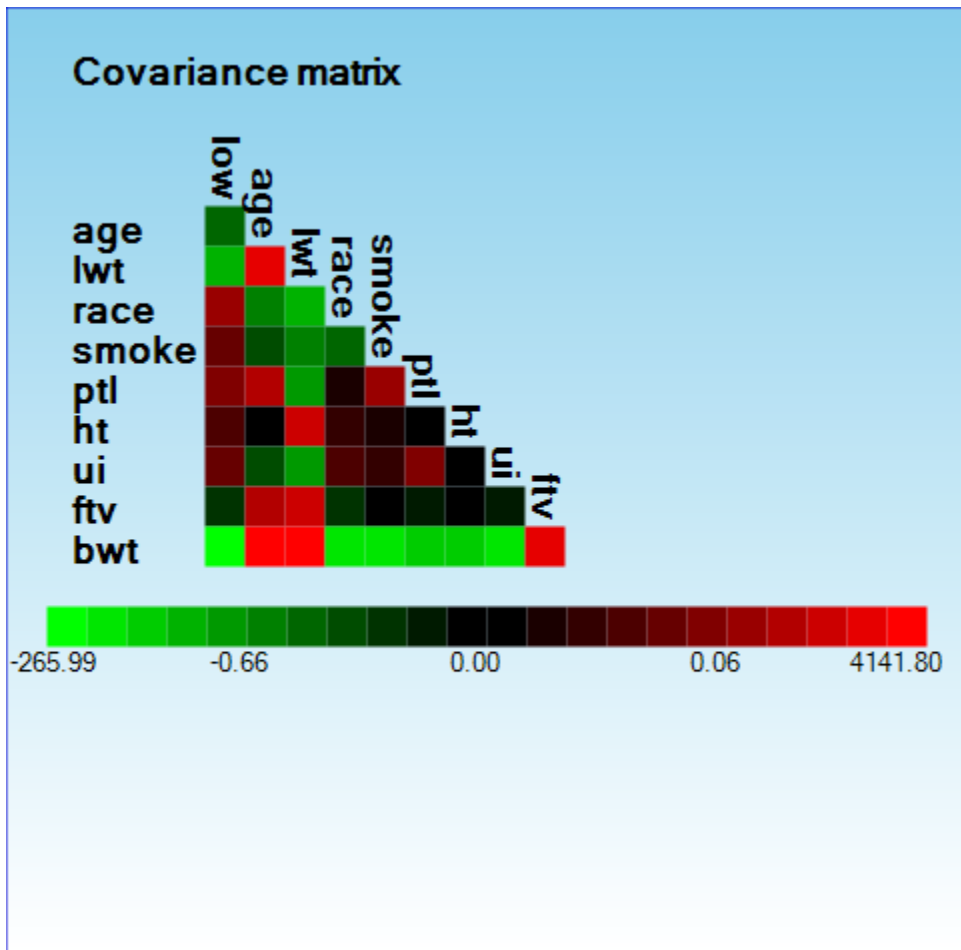
ht	0.02	-0.02	1.77	0.00	0.00	0.00				
ui	0.03	-0.14	-1.66	0.02	0.01	0.04	-0.01			
ftv	-0.03	1.21	4.55	-0.10	-0.01	-0.02	-0.02	-0.02		
bwt	-265.99	347.14	4141.80	-131.36	-67.47	-55.65	-26.04	-73.61	44.99	

The SPSS(12) covariance matrix is shown below:

Inter-Item Covariance Matrix

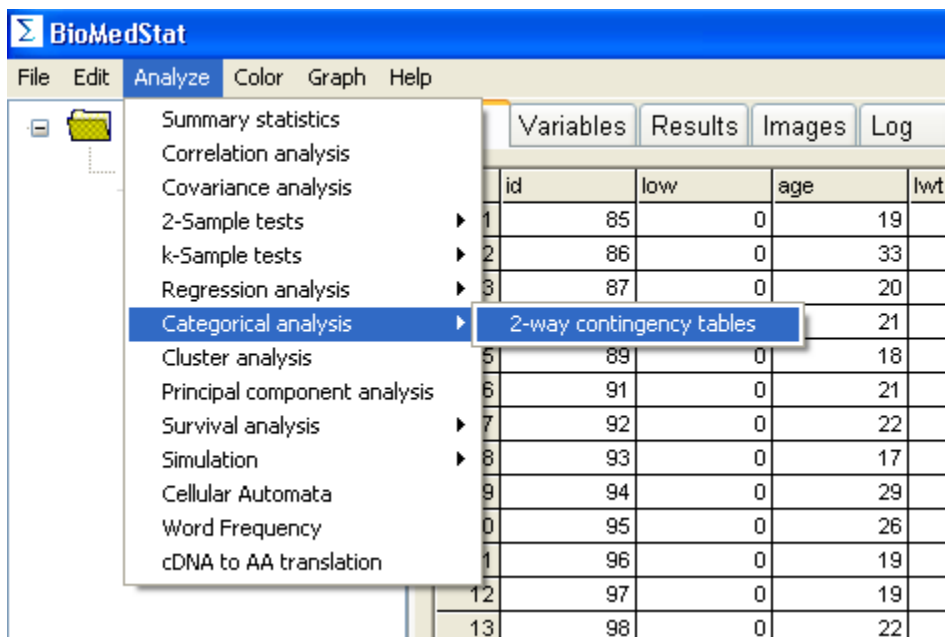
	low	age	lwt	race	smoke	ptl	ht	ui	ftv	bwt
low	.216	-.293	-2.410	.059	.037	.045	.017	.028	-.031	-265.993
age	-.293	28.076	29.177	-.841	-.115	.187	-.021	-.142	1.209	347.141
lwt	-2.410	29.177	935.099	-4.635	-.661	-2.112	1.767	-1.664	4.552	4141.798
race	.059	-.841	-4.635	.843	-.152	.004	.004	.018	-.096	-131.356
smoke	.037	-.115	-.661	-.152	.240	.045	.002	.011	-.015	-67.471
ptl	.045	.187	-2.112	.004	.045	.243	-.002	.040	-.023	-55.650
ht	.017	-.021	1.767	.004	.002	-.002	.060	-.009	-.019	-26.037
ui	.028	-.142	-1.664	.018	.011	.040	-.009	.127	-.022	-73.608
ftv	-.031	1.209	4.552	-.096	-.015	-.023	-.019	-.022	1.122	44.992
bwt	-265.993	347.141	4141.798	-131.356	-67.471	-55.650	-26.037	-73.608	44.992	531473.682

The covariance matrix is calculated and used in the analysis.

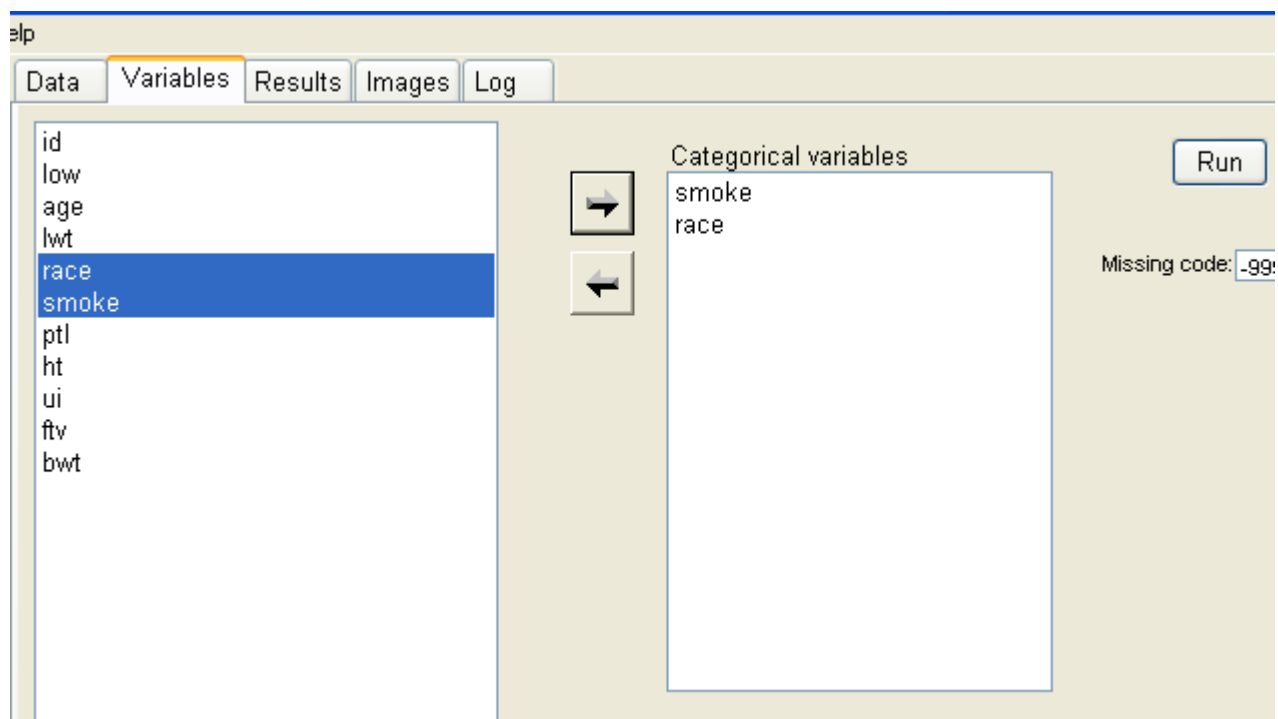


Example 12 – Chi-square 2-way contingency table analysis

Open the Hosmer and Lemeshow data again (Example 2). Specify a 2-way contingency table analysis.



Next select race and smoke as the categorical variables, and click on Run:



The result of the run is as follows:

Chi-Square test				
		Smoke		
		1	2	Total

race	1	44	52	96
	2	16	10	26
	3	55	12	67
	Total	115	74	189
	Chi-square (d.f.)	21.78(2)		
	Tabled chi-square(0.95)	6.4516		
	P-value	0.0000		
	Iterations	1000		
	Mult. prob (observed)	8.97795524039906E-06		
	Randomization P-value	0.0000		

SPSS(12) results are as follows:

race * smoke Crosstabulation

Count

		smoke		Total
		0	1	
race	1	44	52	96
	2	16	10	26
	3	55	12	67
Total		115	74	189

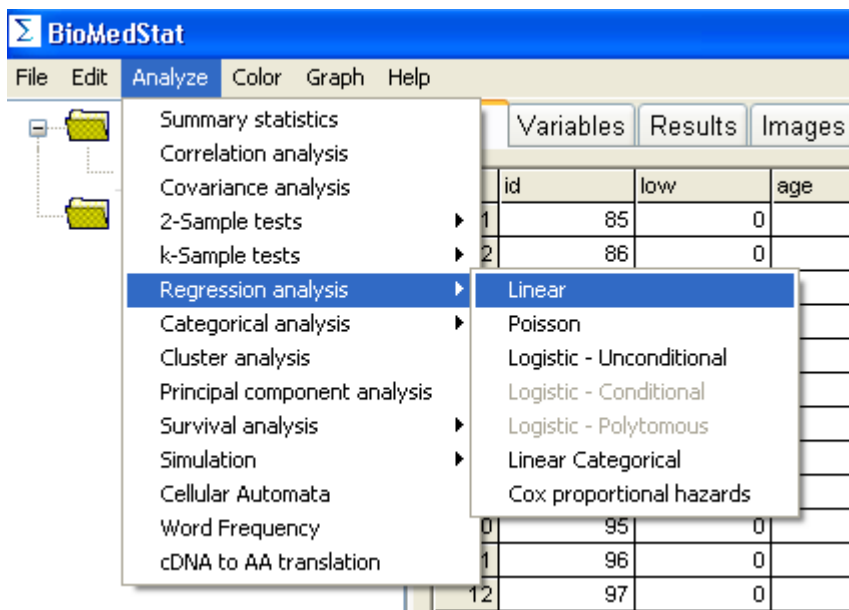
Chi-Square Tests

	Value	df	Asymp. Sig. (2- sided)
Pearson Chi-Square	21.779(a)	2	.000
Likelihood Ratio	22.997	2	.000
Linear-by-Linear Association	21.609	1	.000
N of Valid Cases	189		

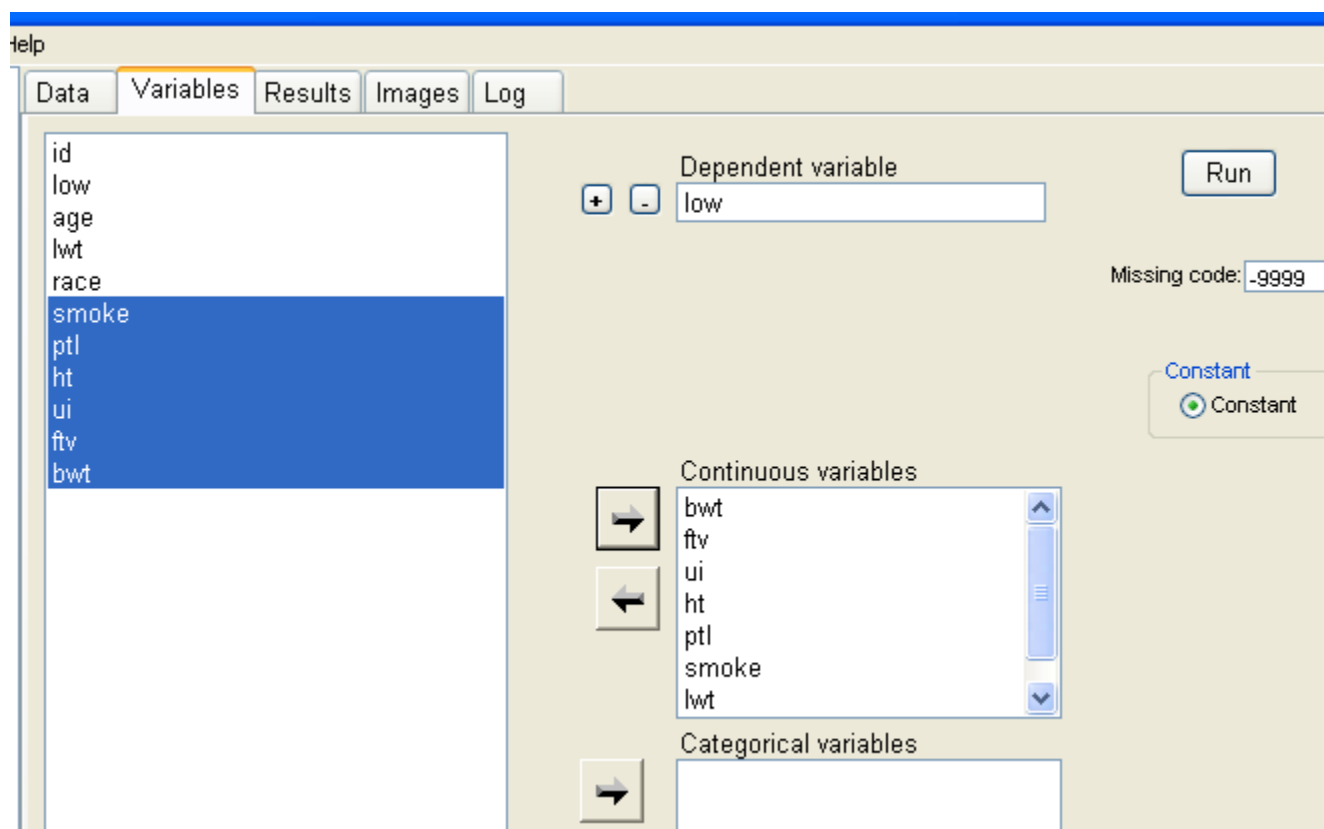
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.18.

Example 13 – Linear regression

Open the Hosmer and Lemeshow data again (Example 2). Specify a linear regression as shown below:



Next, specify low as the dependent variable and all other variables except for race as the independent variables (shown below), and click on Run:



The result of the run is as follows:

Linear Regression

Dependent variable: low

Regression Coefficients					
Variable	Coef.	s.e.	Wald	Prob.	Score
Const	1.9292	0.1647	137.2897	0.0000	0.0000
age	-0.0050	0.0042	1.4596	0.2144	0.0000
lwt	-0.0003	0.0007	0.1946	0.6591	0.0000
smoke	-0.0027	0.0443	0.0038	0.9508	0.0000
ptl	0.0897	0.0451	3.9572	0.0461	0.0000
ht	0.0685	0.0917	0.5587	0.4548	0.0000
ui	-0.1002	0.0637	2.4790	0.1125	0.0000
ftv	0.0000	0.0205	0.0000	0.9993	0.0000
bwt	-0.0005	0.0000	244.4939	0.0000	0.0000

The SPSS(12) result of the run are as follows:

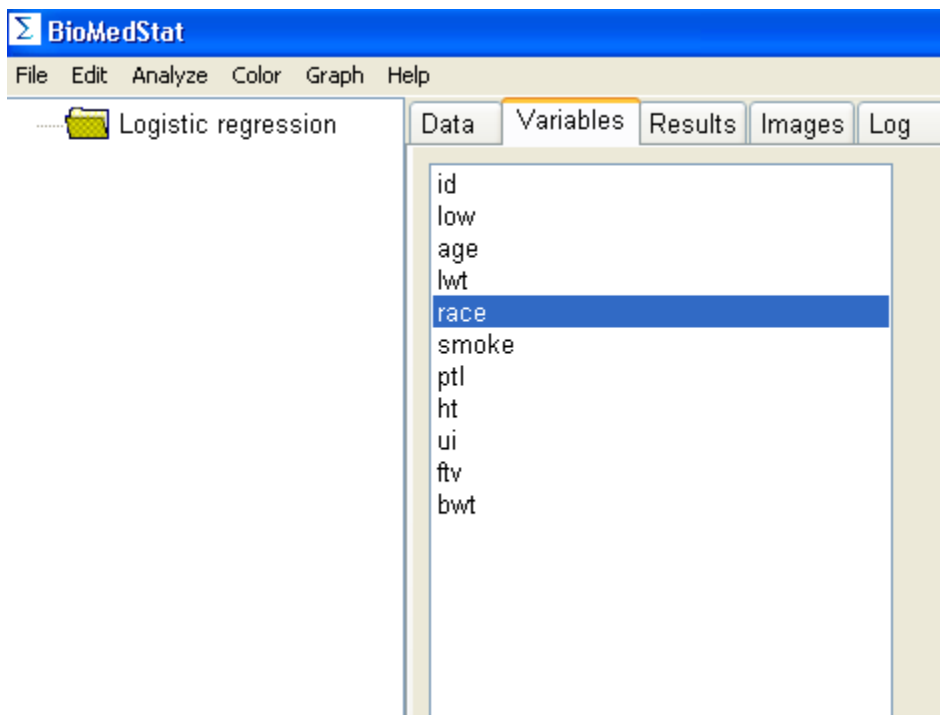
Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.927	.149		12.948	.000
	age	-.005	.004	-.057	-1.208	.229
	lwt	.000	.001	-.022	-.441	.660
	smoke	-.003	.044	-.003	-.062	.951
	ptl	.090	.045	.095	1.989	.048
	ht	.069	.092	.036	.747	.456
	ui	-.100	.064	-.077	-1.574	.117
	ftv	-1.779E-05	.021	.000	-.001	.999
	bwt	.000	.000	-.778	-15.636	.000

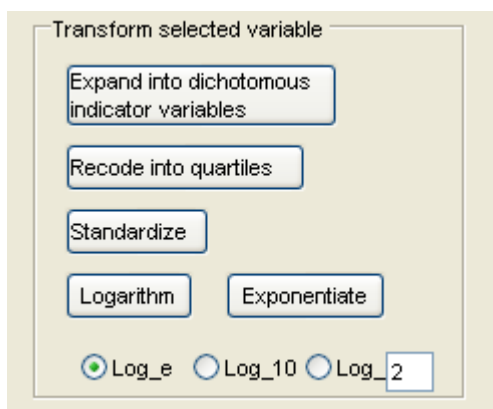
a. Dependent Variable: low

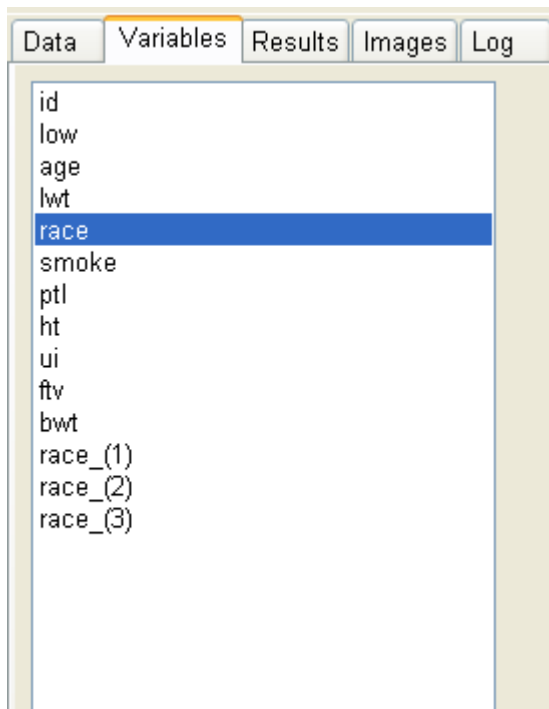
Example 13 – Logistic regression (unconditional)

Open the Hosmer and Lemeshow data (Example 2), and specify Regression→ Unconditional logistic. Highlight the race variable as shown below:

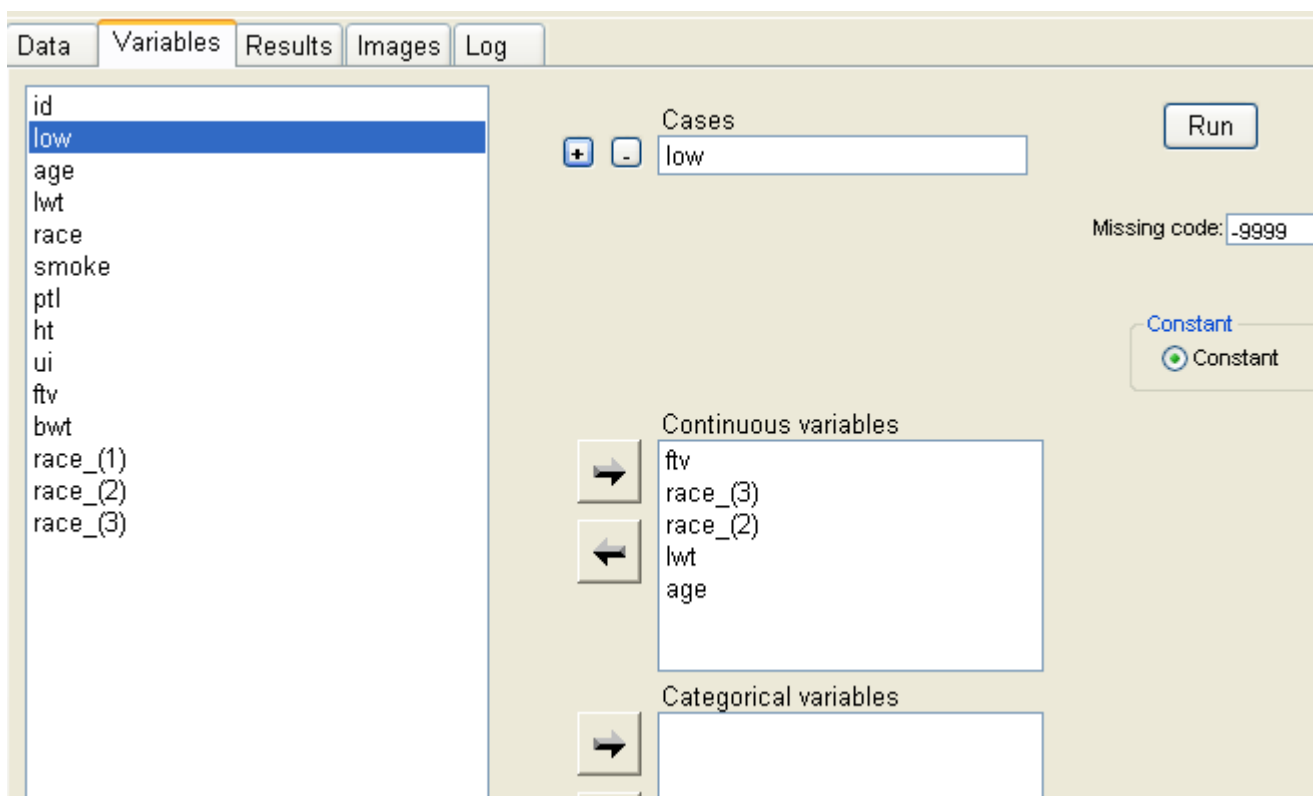


Next, click on the “Expand into dichotomous indicator variables” button as shown below:





Now add low as the “case” variable and age, lwt, race_(2), race_(3), and ftv as independent variables, as shown below:



Click on Run. The results are shown below:

Unconditional Logistic Regression

Dependent variable: low

Regression Coefficients					
Variable	Coef.	s.e.	Wald	Prob.	Score
Const	1.2954	1.0714	1.4617	0.2141	0.0000
age	-0.0238	0.0337	0.4988	0.4800	0.0000
lwt	-0.0142	0.0065	4.7430	0.0292	0.0000
race_(2)	1.0039	0.4979	4.0660	0.0433	0.0000
race_(3)	0.4331	0.3622	1.4296	0.2186	0.0000
ftv	-0.0493	0.1672	0.0869	0.7681	0.0000

Odds Ratios			
Variable	OR	Lower 95% CI	Upper 95% CI
Const	3.652331566	0.447238627	29.826417184
age	0.976458553	0.913992390	1.043193923
lwt	0.985856363	0.973298538	0.998576214
race_(2)	2.728897888	1.028492574	7.240580891
race_(3)	1.542043413	0.758143585	3.136474321
ftv	0.951887601	0.685849579	1.321120596

Contrasts

--

Const	age	lwt	race_(2)	race_(3)	ftv	RR	Lower 95% CI	Upper 95% CI	Chi-square	Prob
	+	+				0.96	0.90	1.03	0.00	1.0000
	+		+			2.66	1.00	7.13	4.06	0.1312
	+			+		1.51	0.45	5.08	1.43	0.4897
	+				+	0.93	0.26	3.27	0.00	1.0000
		+	+			2.69	1.02	7.13	4.07	0.1309
		+		+		1.52	0.46	5.08	1.43	0.4894
		+			+	0.94	0.27	3.27	0.00	1.0000
			+	+		4.21	1.15	15.36	4.41	0.1102
			+		+	2.60	0.68	9.87	4.07	0.1309

				+	+	1.47	0.66	3.25	1.40	0.4958
	+	+	+			2.63	0.74	9.31	4.06	0.2660
	+	+		+		1.48	0.35	6.34	1.43	0.6991
	+	+			+	0.92	0.21	4.06	0.00	1.0000
	+		+	+		4.11	0.57	29.65	4.21	0.2492
	+		+		+	2.54	0.34	18.78	4.06	0.2660
	+			+	+	1.43	0.17	12.37	1.40	0.7052
		+	+	+		4.15	0.33	52.30	4.41	0.2285
		+	+		+	2.56	0.20	32.96	4.07	0.2655
		+		+	+	1.45	0.10	21.03	1.40	0.7048
			+	+	+	4.01	0.20	81.46	4.41	0.2290

SPSS(12) results are:

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1(a)						
age	-.024	.034	.499	1	.480	.976
lwt	-.014	.007	4.743	1	.029	.986
race			4.411	2	.110	
race(1)	1.004	.498	4.066	1	.044	2.729
race(2)	.433	.362	1.430	1	.232	1.542
ftv	-.049	.167	.087	1	.768	.952
Constant	1.295	1.071	1.462	1	.227	3.652

a Variable(s) entered on step 1: age, lwt, race, ftv.

The same regression coefficients and standard errors are listed in Table 2.2 on page 30 of Hosmer and Lemeshow, *Applied Logistic Regression*, New York, Wiley, 1989.