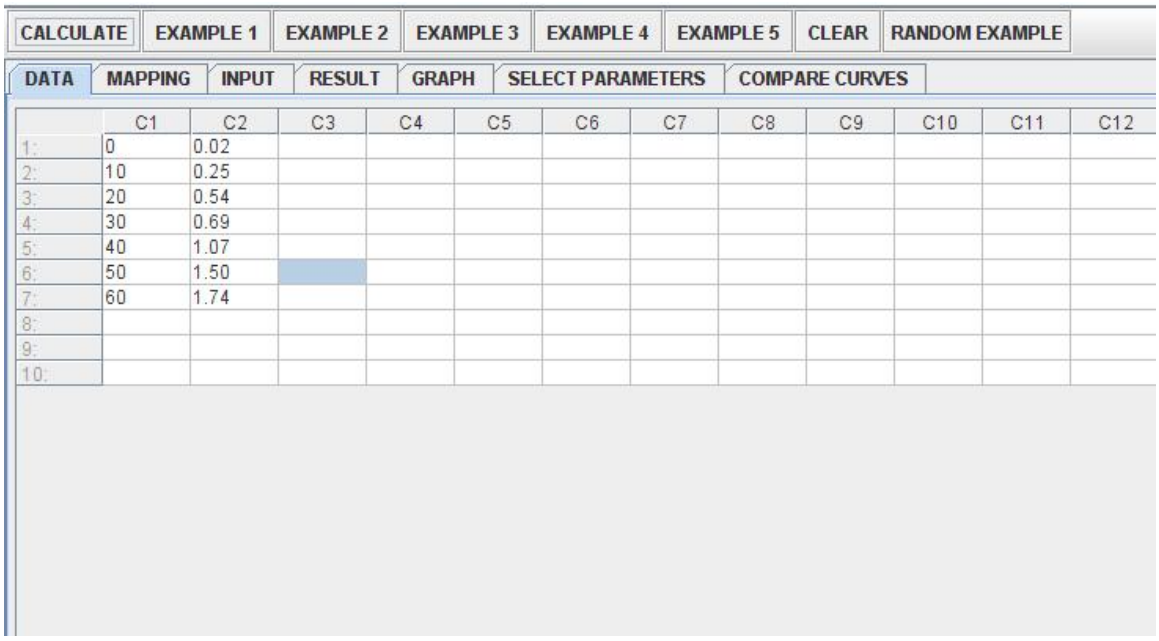
Stats 13.1, Fall 2008, Homework 7 Solutions

http://www.stat.ucla.edu/~dinov/courses_students.dir/08/Fall/STAT13.1.dir/

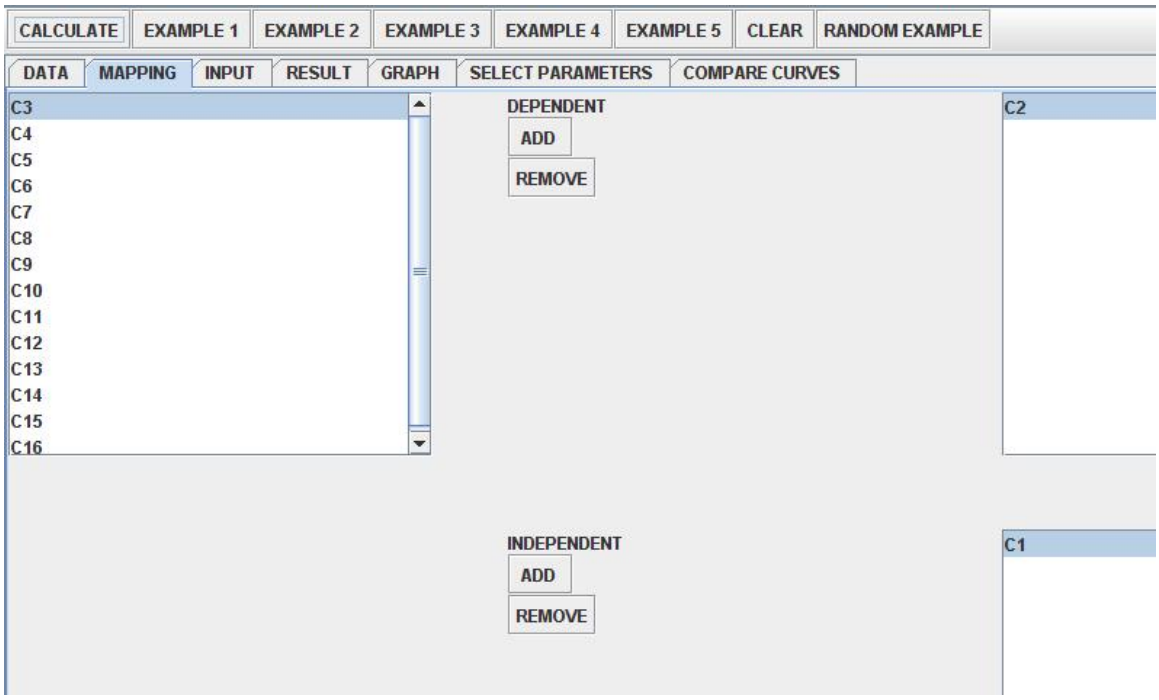
1. Problem 1

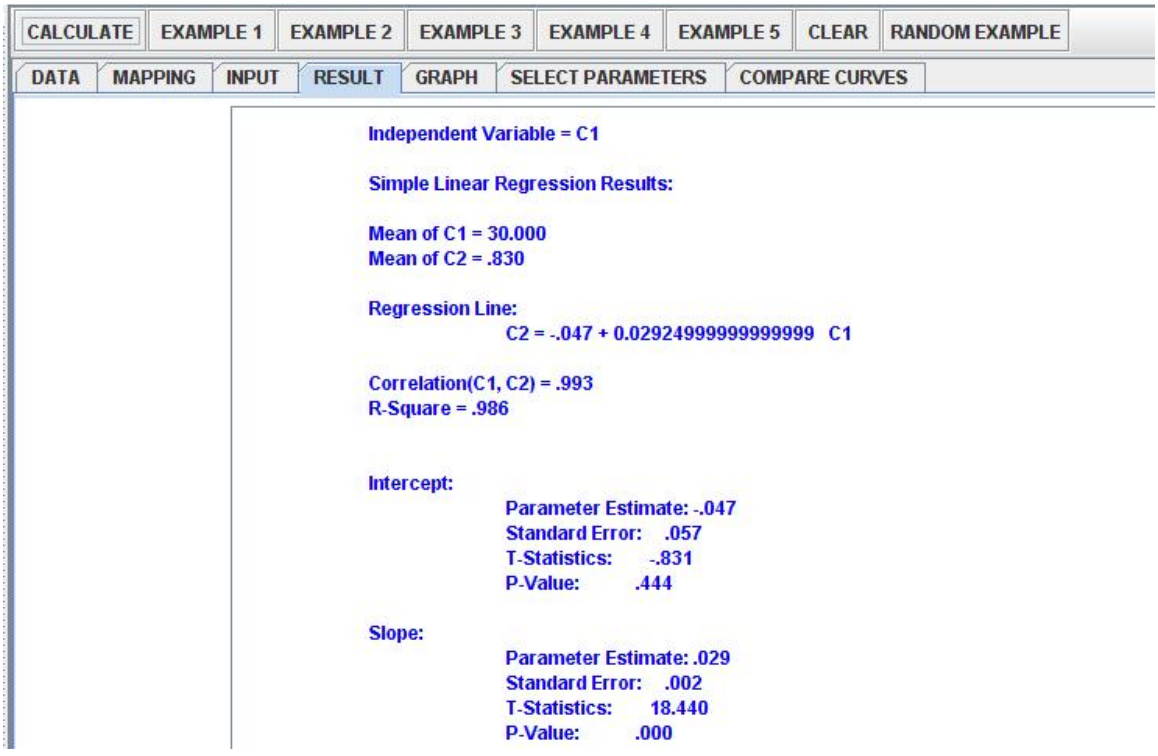
(part 1) Using the equation $\hat{b} = \frac{\sum_{i=1}^{N}(x_i y_i) - N\overline{xy}}{\sum_{i=1}^{N}(x_i)^2 - N\overline{x}^2}$, your estimated rate of incorporation is $\frac{256.2 - 7(30)(0.83)}{9100 - 7(30)^2} = 0.02925$

The rate according to the SOCR regression analysis (C1 (the x values) is independent and C2 (the y values) is dependent) is very close to our estimated rate. See figures below.

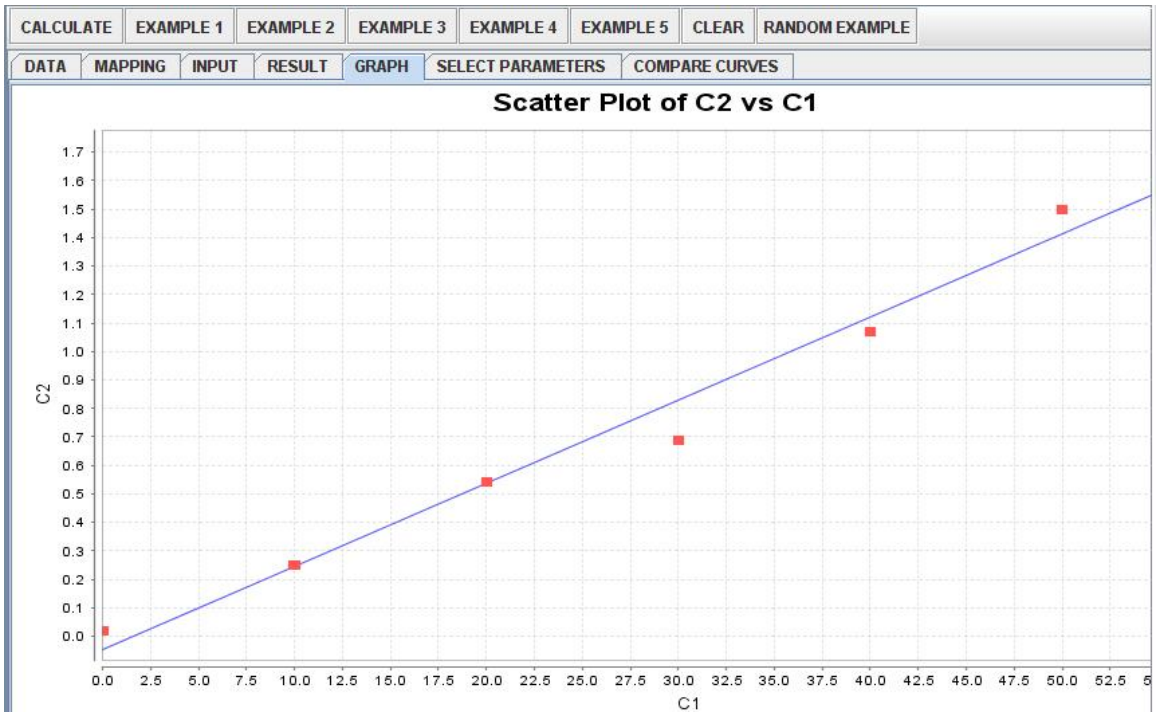


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1:	0	0.02										
2:	10	0.25										
3:	20	0.54										
4:	30	0.69										
5:	40	1.07										
6:	50	1.50										
7:	60	1.74										
8:												
9:												
10:												


CALCULATE
EXAMPLE 1
EXAMPLE 2
EXAMPLE 3
EXAMPLE 4
EXAMPLE 5
CLEAR
RANDOM EXAMPLE
DATA
MAPPING
INPUT
RESULT
GRAPH
SELECT PARAMETERS
COMPARE CURVES
C3
C4
C5
C6
C7
C8
C9
C10
C11
C12
C13
C14
C15
C16
DEPENDENT
ADD
REMOVE
C2
INDEPENDENT
ADD
REMOVE
C1


CALCULATE
EXAMPLE 1
EXAMPLE 2
EXAMPLE 3
EXAMPLE 4
EXAMPLE 5
CLEAR
RANDOM EXAMPLE
DATA
MAPPING
INPUT
RESULT
GRAPH
SELECT PARAMETERS
COMPARE CURVES
Independent Variable = C1
Simple Linear Regression Results:
Mean of C1 = 30.000
Mean of C2 = .830
Regression Line:
C2 = -.047 + 0.02924999999999999 C1
Correlation(C1, C2) = .993
R-Square = .986
Intercept:
Parameter Estimate: -.047
Standard Error: .057
T-Statistics: -.831
P-Value: .444
Slope:
Parameter Estimate: .029
Standard Error: .002
T-Statistics: 18.440
P-Value: .000

(part 2) The regression line is a straight line that shows an estimation of the y value at any given x value. See the plot below:



(part 3) Residual = (y value) - {(intercept) + (slope * x value)}. For example, for index 3, the residual will equal 0.54 - {-0.047+(0.02925*20)} = 0.002. Calculate the residuals for each of the 7 indexes and calculate the standard deviation using the formula $\sqrt{\frac{\sum_{i=1}^{7}(residual_i - avgresidual)^2}{7-1}}$. The residuals (from index 1 to 7) are 0.067, 0.0045, 0.002, -0.1405, -0.053, 0.0845, and 0.032. The avgresidual (or average residual) is -0.0005. Therefore the standard deviation of the residuals is 0.0766.

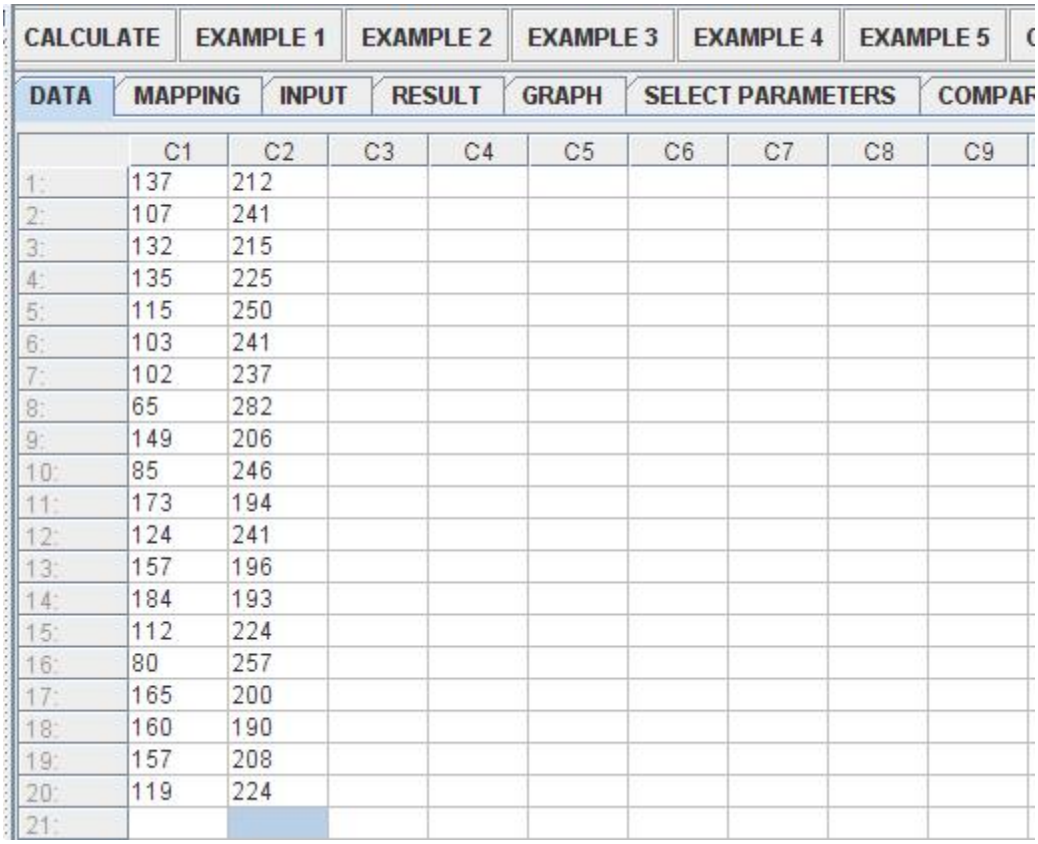
2. Problem 2

(part one) There are three steps involved in using the SOCR Regression Analysis:

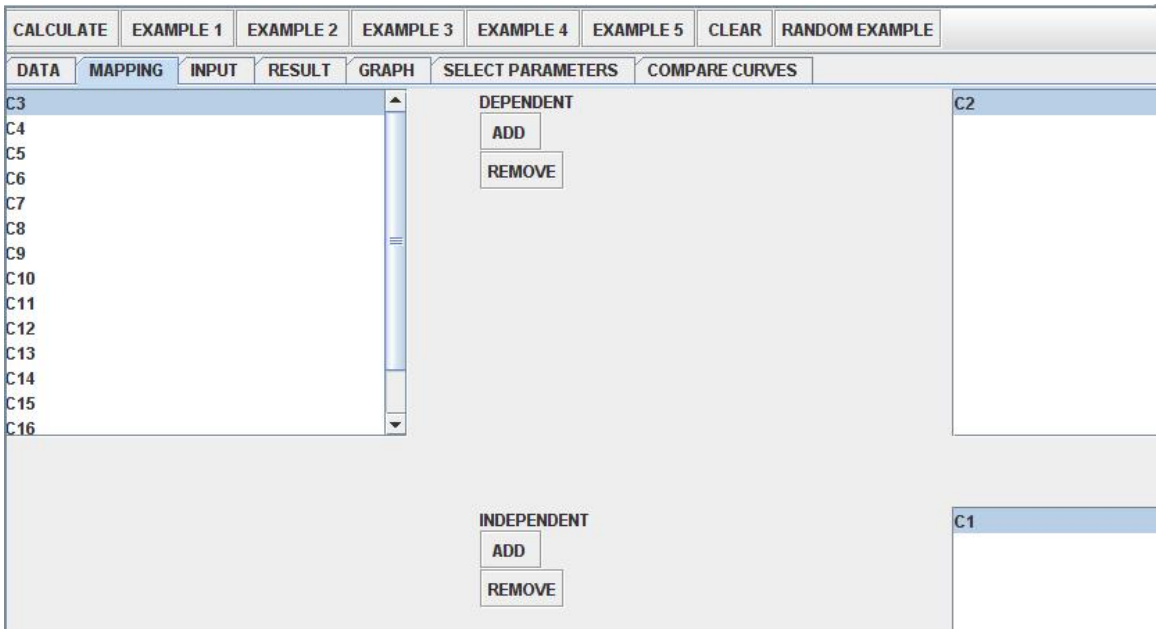
Step 1: Enter the data like the first figure.

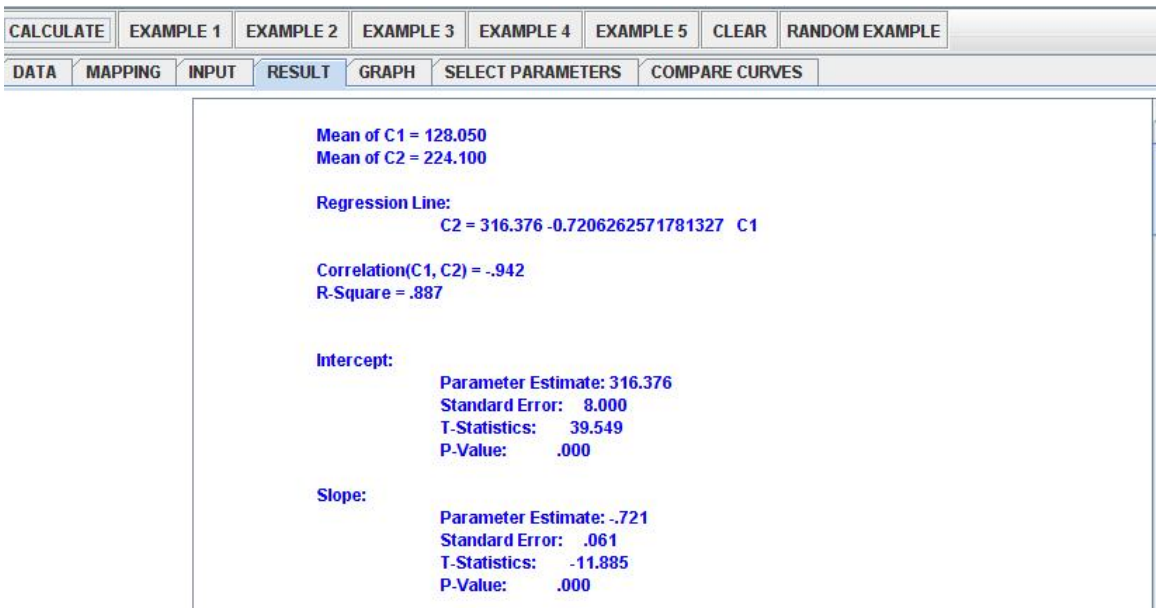
Step 2: Map the columns so that the C1 (x) is independent and C2 (y) is dependent.

Step 3: Click on CALCULATE and the results will be shown in the RESULTS tab.

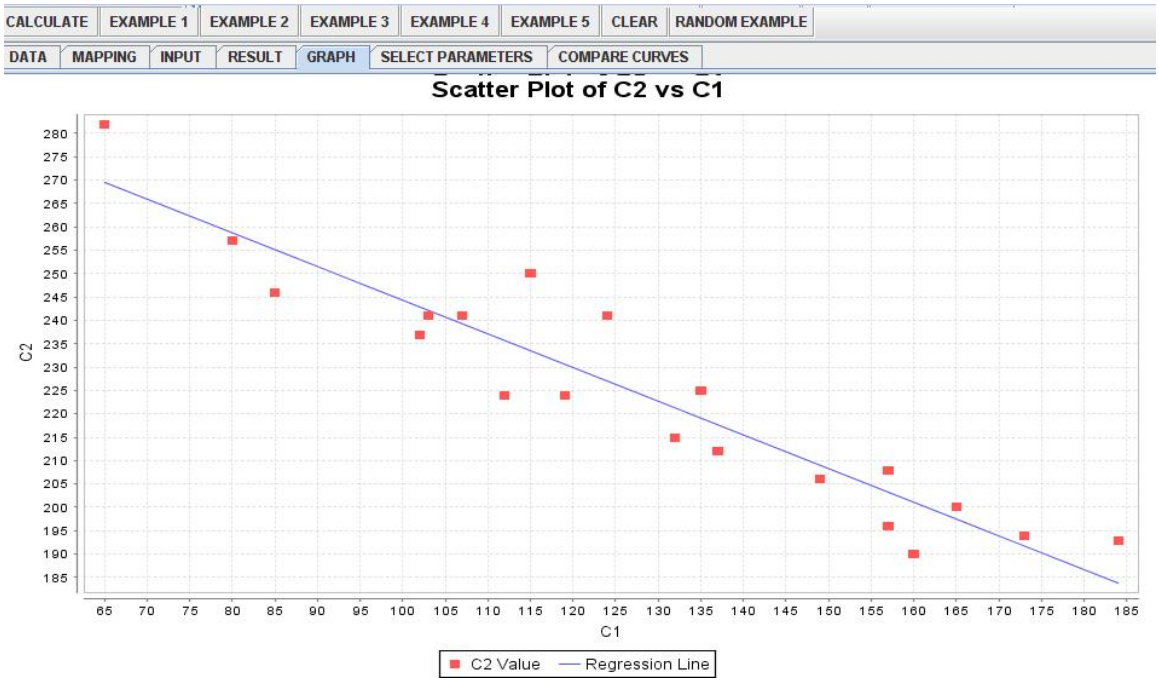


	C1	C2	C3	C4	C5	C6	C7	C8	C9
1:	137	212							
2:	107	241							
3:	132	215							
4:	135	225							
5:	115	250							
6:	103	241							
7:	102	237							
8:	65	282							
9:	149	206							
10:	85	246							
11:	173	194							
12:	124	241							
13:	157	196							
14:	184	193							
15:	112	224							
16:	80	257							
17:	165	200							
18:	160	190							
19:	157	208							
20:	119	224							
21:									





(part 2) The plot of the data with the line:



(part 3) The slope is -0.721 indicating that the cob weight decreases by 72 units as the plant density increases by 100 units. The intercept is 316.376 indicating that the cob weight is estimated to be 316 units if there are no plants in the plot (x=0). The correlation is -0.942 meaning that the cob weight highly depends on the plant density.

3. **Problem 3**

(effect on sodium) Testing H_0: sodium levels are the same for all three meat groups.
Step 1: Input the data.

CALCULATE | EXAMPLE 1 | EXAMPLE 2 | EXAMPLE 3 | EXAMPLE 4 | EX

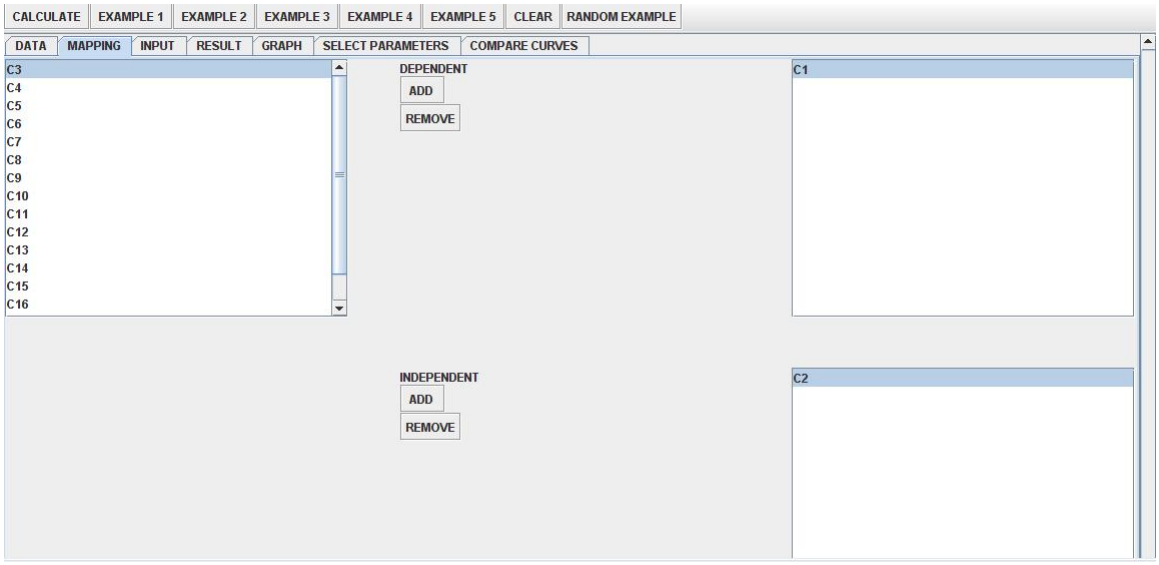
DATA | MAPPING | INPUT | RESULT | GRAPH | SELECT PARAMETERS

	C1	C2	C3	C4	C5	C6
1:	495	A				
2:	477	A				
3:	425	A				
4:	322	A				
5:	482	A				
6:	587	A				
7:	370	A				
8:	322	A				
9:	479	A				
10:	375	A				
11:	330	A				
12:	300	A				
13:	386	A				
14:	401	A				
15:	645	A				
16:	440	A				
17:	317	A				
18:	319	A				
19:	298	A				
20:	253	A				
21:	458	B				
22:	506	B				
23:	473	B				
24:	545	B				
25:	496	B				
26:	360	B				
27:	387	B				
28:	386	B				
29:	507	B				

On the next page there is more data…

30:	393	B
31:	405	B
32:	372	B
33:	144	B
34:	511	B
35:	405	B
36:	428	B
37:	339	B
38:	430	C
39:	375	C
40:	396	C
41:	383	C
42:	387	C
43:	542	C
44:	359	C
45:	357	C
46:	528	C
47:	513	C
48:	426	C
49:	513	C
50:	358	C
51:	581	C
52:	588	C
53:	522	C
54:	545	C

Step 2: Map the columns so that the column with sodium (in this case C1) is dependent.



Step 3: Hit CALCULATE and check out the results from the RESULTS tab:

The P-value is greater than 0.05, so we will not reject the null hypothesis that the sodium levels of the three meat groups are the same.

(effect on calories) Testing H_0: the calories are the same for all three meat groups.
Step 1: Input the data:

CALCULATE | EXAMPLE 1 | EXAMPLE 2 | EXAMPLE 3 | EXAMPLE 4 | EX

DATA | MAPPING | INPUT | RESULT | GRAPH | SELECT PARAMETERS

	C1	C2	C3	C4	C5	C6
1:	186	A				
2:	181	A				
3:	176	A				
4:	149	A				
5:	184	A				
6:	190	A				
7:	158	A				
8:	139	A				
9:	175	A				
10:	148	A				
11:	152	A				
12:	111	A				
13:	141	A				
14:	153	A				
15:	190	A				
16:	157	A				
17:	131	A				
18:	149	A				
19:	135	A				
20:	132	A				
21:	173	B				
22:	191	B				
23:	182	B				
24:	190	B				
25:	172	B				
26:	147	B				
27:	146	B				
28:	139	B				
29:	175	B				

More data is on the next page…

DATA | MAPPING | INPUT | RESULT | GRAPH | SELECT PARAMETER

	C1	C2	C3	C4	C5	C6
30:	136	B				
31:	179	B				
32:	153	B				
33:	107	B				
34:	195	B				
35:	135	B				
36:	140	B				
37:	138	B				
38:	129	C				
39:	132	C				
40:	102	C				
41:	106	C				
42:	94	C				
43:	102	C				
44:	87	C				
45:	99	C				
46:	107	C				
47:	113	C				
48:	135	C				
49:	142	C				
50:	86	C				
51:	143	C				
52:	152	C				
53:	146	C				
54:	144	C				

Step 2: Map the columns so that the column with calories (in this case C1) is dependent. (see the figure from the analysis of sodium)

Step 3: Hit CALCULATE and check out the results from the RESULTS tab on the next page.

The p-value is very small, so the null hypothesis will be rejected. Also note the huge F-value which is another reason to reject the null hypothesis.

4. Problem 4

Step 1: Input the data so that all CPI values are from the same year, which is just one way of doing it.

Step 2: Map the columns so that the CPI values are dependent.

Step 3: Click CALCULATE and see the results from the RESULTS tab. Looking at the next page, you can see that the null hypothesis that all CPI item had the same effect on the CPI value because 1) the F-value is big and 2) the p-value is tiny.

Sample Size = 48

Independent Variable = C2
Dependent Variable = C1

Results of One-Way Analysis of Variance:

Model:
Degrees of Freedom = 3
Residual Sum of Squares = 211590.919
Mean Square Error = 70530.306

Error:
Degrees of Freedom = 44
Residual Sum of Squares = 668.348
Mean Square Error = 15.190

Corrected Total:
Degrees of Freedom = 47
Residual Sum of Squares = 212259.267

F-Value = 4643.289
P-Value = < 1E-15

R-Square = .997