Lab #5.

Dummy Variables Ex. (Thanksgiving Turkeys)

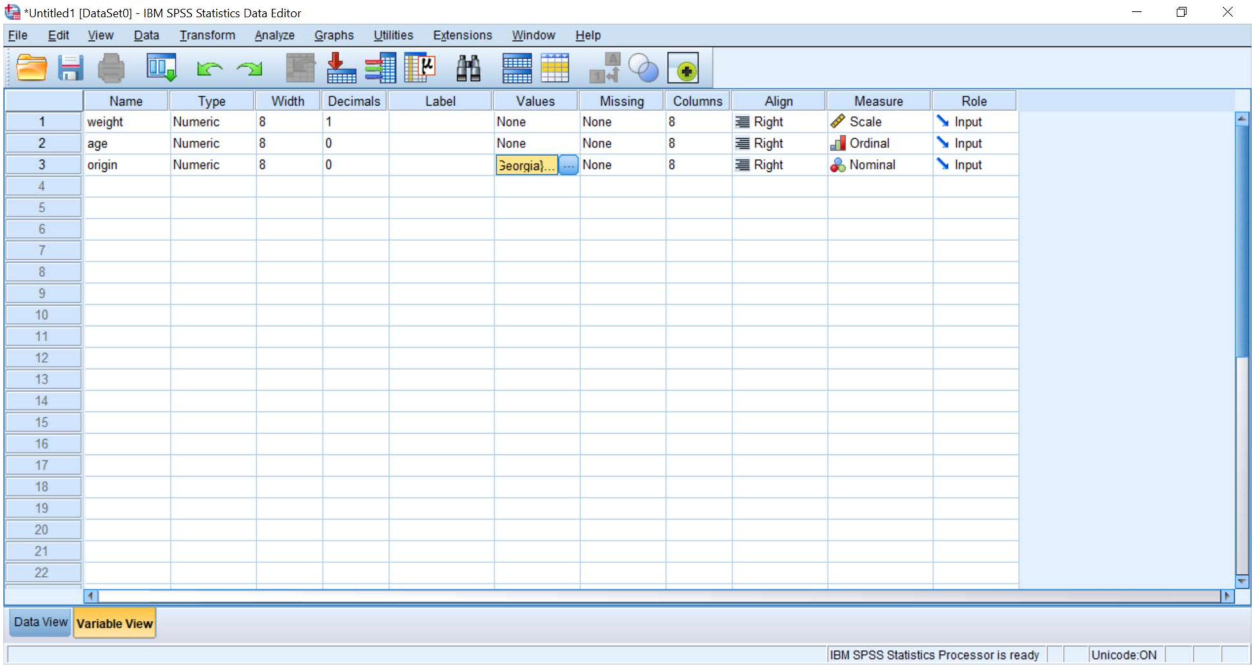
Obs	Ages(X_1)	Weights(Y)	Origin
1	28	13.3	G
2	20	8.9	G
3	32	15.1	G
4	22	10.4	G
5	29	13.1	V
6	27	12.4	V
7	28	13.2	V
8	26	11.8	V
9	21	11.5	W
10	27	14.2	W
11	29	15.4	W
12	23	13.1	W
13	25	13.8	W

y = turkey weight (in pounds)
$\boldsymbol{x_1}$ = ages (in weeks)
n = 13 thirteen Thanksgiving turkeys.
Origin = 1) Four of these turkeys were reared in Georgia(G),
2) Four in Virginia(V)
3) Five in Wisconsin(W).

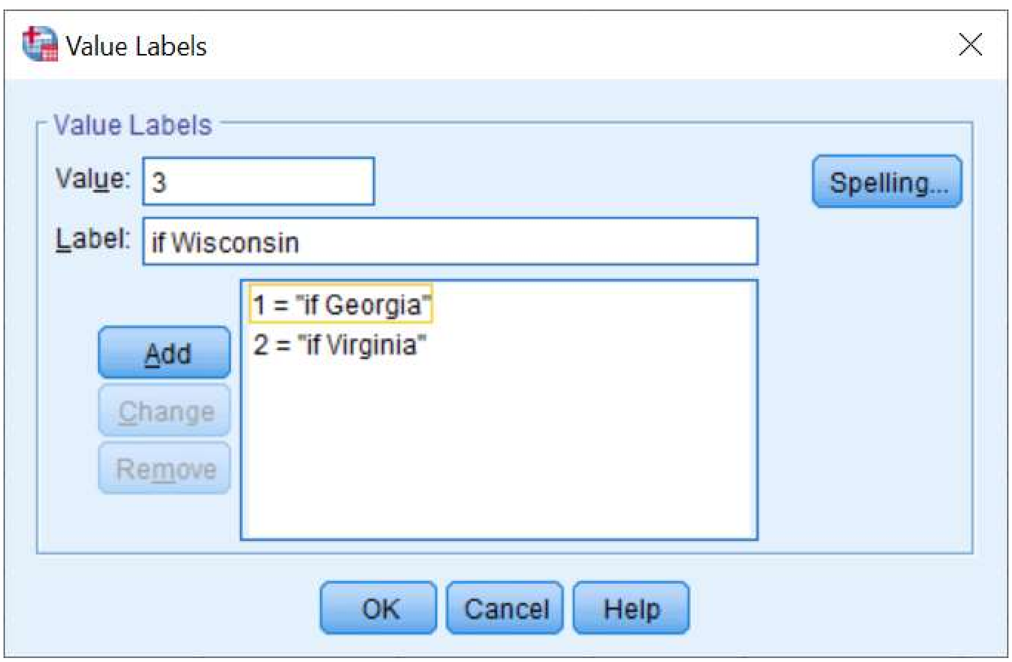
Let us define 2 dummy variables (because categorical variable "origin" has 3 categories: G, V and W).

Let $x_2 = \begin{cases} 1; & \text{if G} \\ 0; & \text{ow} \end{cases}$ and $x_3 = \begin{cases} 1; & \text{if V} \\ 0; & \text{ow} \end{cases}$

We enter the y, x1 and origin variables in to SPSS:



With categorical variable “origin” having the values defined as:

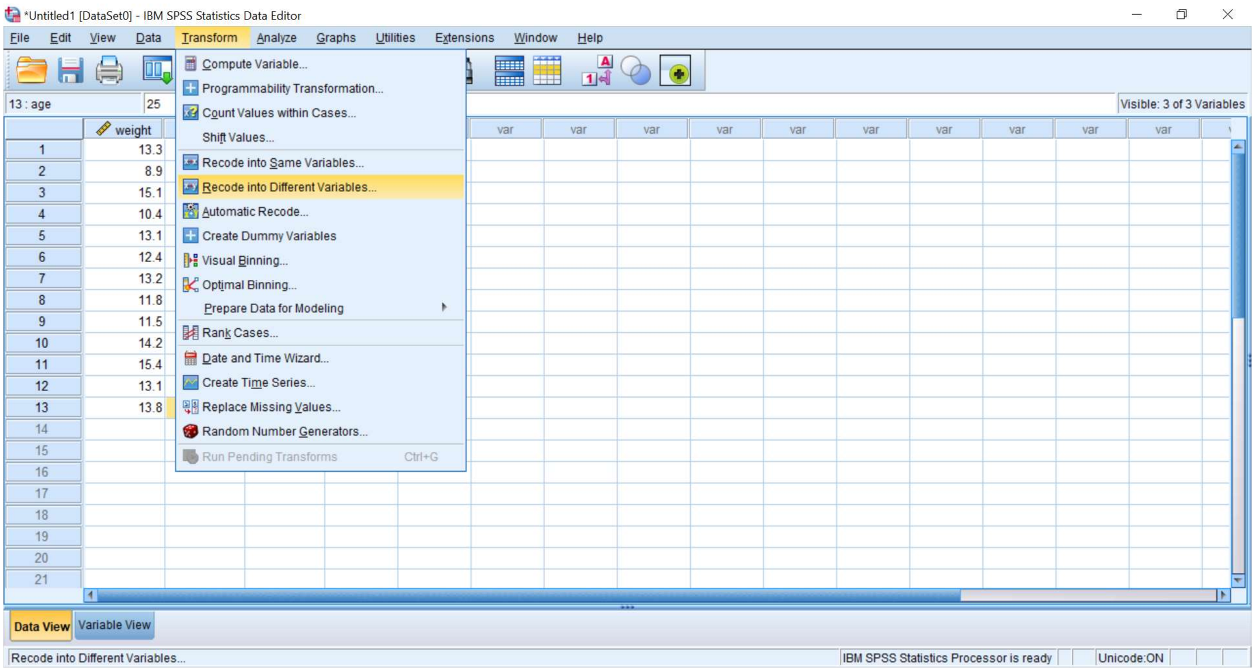


i.e “value” 1 , “label” if Georgia → “Add” → “OK”,

“value” 2 , “label” if Virginia → “Add” → “OK”,

“value” 3 , “label” if Wisconsin → “Add” → “OK”.13

Next , we need to define our dummy variables x_2 (as 1 if G and 0 otherwise) and x_3 (as 1 if V and 0 otherwise).



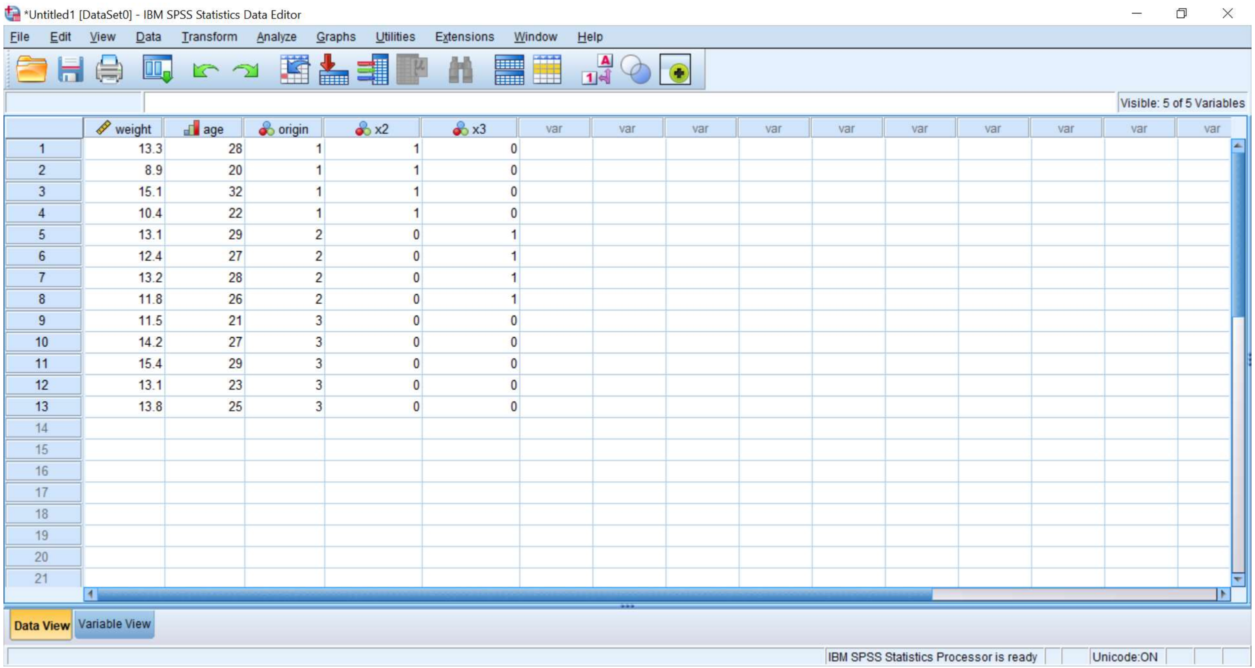
→ “Transform” → “Recode into different variables” → select “origin” under the <u>input variable</u> and type “x2” under <u>output variable</u> → click on “old and new values” → and under old values type ‘1’, then under new value type ‘1’ → click on “add”. Repeat with: under old values type ‘2’, then under new value type ‘0’ → click on “add”. Repeat with: under old values type ‘3’, then under new value type ‘0’ → click on “add” → click on “continue” → “change” → OK.

(**<u>NOTE;</u>** here we defined x_2 dummy variable that is 1 if origin is Georgia and 0 if origin is Virginia or Wisconsin)

Repeat, the same process, by defining a dummy variable x_3 such that x_3=1 if origin is Virginia and x_3=0 if origin is Georgia or Wisconsin.

i.e. → “Transform” → “Recode into different variables” → select “origin” under the <u>input variable</u> and type “x3” under <u>output variable</u> → click on “old and new values” → and under old values type ‘1’, then under new value type ‘0’ → click on “add”. Repeat with: under old values type ‘2’, then under new value type ‘1’ → click on “add”. Repeat with: under old values type ‘3’, then under new value type ‘0’ → click on “add” → click on “continue” → “change” → OK.

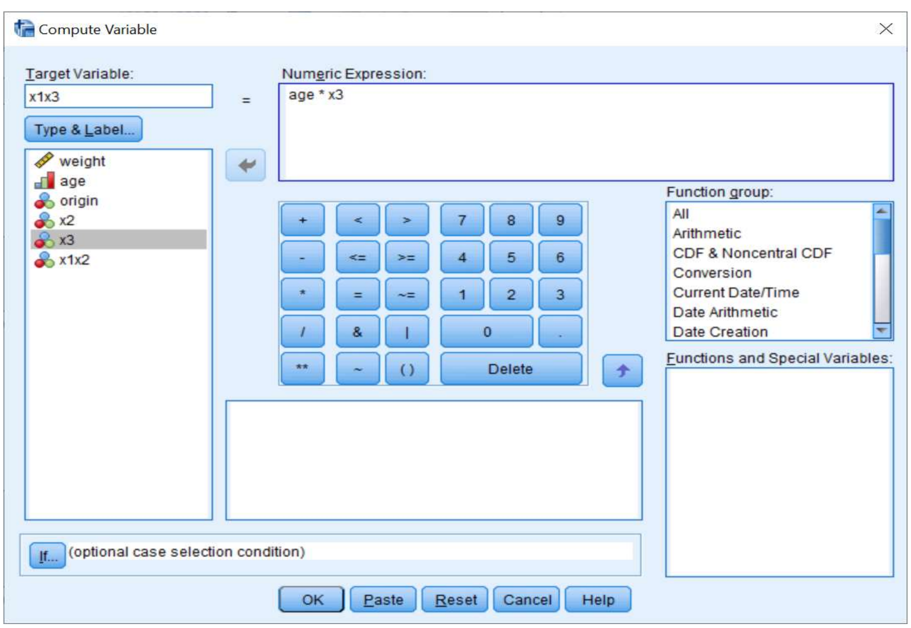
We get following variables:



Our <u>full model</u> is: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \varepsilon$

<u>To define the interaction terms:</u>

Go to “Transform” → “Compute Variable” → type ‘x1x2’ under <u>target variable</u> and then under <u>numeric expression</u> type ‘origin*x2’ → OK. Repeat with defining ‘x1x3’ as ‘origin*x3’ → OK.



To run the full model:

Go to "Analyze" → "Regression" → "Linear" → put 'weight' under dependent variable and 'age', x2, x3, x1x2 and x1x3 under independent variables → OK

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	38.711	5	7.742	76.744	.000[b]
	Residual	.706	7	.101		
	Total	39.417	12			

a. Dependent Variable: weight

b. Predictors: (Constant), x1x3, age, x2, x1x2, x3

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.475	1.264		1.959	.091
	age	.445	.050	.871	8.861	.000
	x2	-3.454	1.531	-.916	-2.257	.059
	x3	-2.775	4.109	-.736	-.675	.521
	x1x2	.061	.060	.423	1.013	.345
	x1x3	.025	.151	.182	.166	.873

a. Dependent Variable: weight

- Test whether the 3 lines have the same y-intercepts, i.e.

$H_0 : \beta_2 = \beta_3 = 0$

$H_a :$ at least one of β's $\neq 0$

The reduced model is: $y = \beta_0 + \beta_1 x_1 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \varepsilon$

To run the reduced model:

Go to "Analyze" → "Regression" → "Linear" → put 'weight' under dependent variable and 'age', x1x2 and x1x3 under independent variables → OK

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	38.196	3	12.732	93.838	.000[b]
	Residual	1.221	9	.136		
	Total	39.417	12			

a. Dependent Variable: weight

b. Predictors: (Constant), x1x3, age, x1x2

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.107	.814		.132	.898
	age	.539	.033	1.054	16.423	.000
	x1x2	-.074	.010	-.510	-7.555	.000
	x1x3	-.083	.010	-.608	-8.574	.000

a. Dependent Variable: weight

Test-statistic: $$F_{part} = \frac{\left[SSR_f - SSR_r\right] / \left[df_{SSR_f} - df_{SSR_r}\right]}{MSE_f} = \frac{(38.711-38.196)/(5-3)}{0.101} =$$

$$= F_{drop} = \frac{\left[SSE_r - SSE_f\right] / \left[df_{SSE_r} - df_{SSE_f}\right]}{MSE_f} = \frac{(1.221-0.706)/(9-7)}{0.101}$$

R.R.: reject Ho if $F_{part}\ (or\ F_{drop}) > F_{\alpha;(2,7)}$