**Restructure Node** and **Derive Node** in SPSS Modeler

Building a regression model with gender discrimination data

The dataset **gender discrimination.xlsx** has a few categorical variables, which needs to be transformed into dummy variables.

**Gender** has values Male, Female. We need one dummy variable with 1 for Male, 0 for female

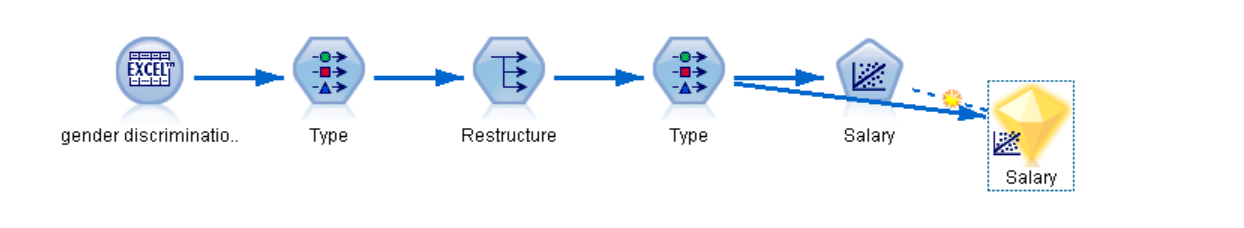
**HasDoctorate** has Yes, or No. We need only one dummy variable with 1 for Hasdoctorate, 0 if the instructor has No Doctorate.

**Rank** has 4 possible values Professor, Associate Professor, Assistant Professor, Lecturer

So we need 3 dummy variable

These additional columns can be added with the help of the **Restructure** Node in SPSS Modeler.

Note that the Type node is added twice. Once before restructure, and again after restructure.



**Derive Node** in SPSS Modeler:

Building and running a Binomial Logistic Regression Model with Breast cancer data.

SPSS modeler complains that **Malignant** is too long a name for a Flag Variable.

So our plan is to derive a new column (Field) which has T or F. T for Malignant, F for Benign. We will call the new field **cancerous.**

Since we need to derive a new **Target** column, add Derive node before Type node, derive a new column, and then set Target with the Type node.

