

## Kepler - Bug #2044

### R actor for means and error by group

03/11/2005 03:37 PM - Matt Jones

<b>Status:</b>	Resolved	<b>Start date:</b>	03/11/2005
<b>Priority:</b>	Immediate	<b>Due date:</b>	
<b>Assignee:</b>	Dan Higgins	<b>% Done:</b>	0%
<b>Category:</b>	actors	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	1.0.0alpha6	<b>Spent time:</b>	0.00 hour
<b>Bugzilla-Id:</b>	2044		
<b>Description</b>			
<p>This bug is to create an actor that wraps an R script (using the generic R actor described in bug <a href="#">#1342</a>) for calculating the mean, standard deviation, and standard error of a numeric variable that is passed into the script. Optionally the stats should be calculated using a grouping variable (by group), with one mean/std/se for each group.</p> <p>There are two inputs:</p> <ol style="list-style-type: none"><li>1) dataValue (type double)</li><li>2) group (type string)</li></ol> <p>The actor will accumulate all data passed in as a series of {dataValue, group} tuples and pass this along to R to calculate the mean, stdev, and stderr by group. If group is not present then the tuples will be {dataValue}. When the R script is finished, a matrix of {group, mean, stdev, stderr} will be passed back to the R actor and then emitted on the output port. So for N tokens that come in on the input port the actor will output on 1 token on the output port after all input data has been received (this is a grouping operation).</p>			
<b>Related issues:</b>			
Blocked by Kepler - Bug #2043: create suite of common statistical actors using R		<b>Resolved</b>	<b>03/11/2005</b>

### History

#### #1 - 06/27/2005 01:21 PM - Dan Higgins

An RExpression actor with this functionality has been created and is present in the Kepler CVS distribution as an example workflow.

#### #2 - 03/27/2013 02:18 PM - Redmine Admin

Original Bugzilla ID was 2044