

Customizing SAS® Studio in SAS® Viya® – The Next Step

Danny Modlin Sr. Analytical Training Consultant SAS - Cary NC

Developing a Custom Step for the KCLUS Procedure

1. Starting from SAS Drive, use the applications menu to proceed to **Develop Code and Flows**. From the menu bar, click **New** and select **Custom Step**.

DATA Pane

- 2. The Custom Step area opens into the Designer with Page 1 appearing. On the properties panel, set the ID of the page to **data** and the label to **DATA**.
- In the Control Library, scroll down to the Data area and drag Input Table to the DATA page. Set ID to inputdata and Label to Select table for analysis. Leave the check box next to Required selected.
- 4. Add two Column Selector items to the page. Set the following properties:

	Column Selector 1	Column Selector 2	
ID	catvars	intvars	
Label	Select categorical inputs	Select continuous inputs	
Link to input table	(inputdata)	(inputdata)	
Column Type	All types	Numeric	

OPTIONS Pane

- 5. In the Control Library, click **Add Page**. With this new page selected, set **ID** to **options** and **Label** to **OPTIONS** in the properties panel.
- 6. Add two Check Box items to the page. Set the following properties:

	Check Box 1	Check Box 2
ID	imputeint imputecat	
Label	Replace interval Replace nom	
	missing values missing value	
	with mean.	with mode.

- 7. Add a radio button group to the page. Set ID to radiocluster and Label to Number of clusters:. Click Add Many to add several items to the radio list. In the box that appears, enter Specify number of clusters and Calculate number of clusters, where each option appears on its own line. Click OK.
- 8. If the original default radio button is still present in the list, click the box in front of that option (in properties) and click the trash button.
- 9. Click the **Map Values** button to expand the options within the radio list. A new column named **Value** appears. On each line, change **Value** to **specify** and **calculate** respectively.
- 10. Set the default radio button to **Specify number of clusters**.
- 11. Add a text or numeric field to the page. Set **ID** to **numcluster** and **Label** to **Maximum number of clusters**. Change **Type** to **Numeric**. Set the default value to 4, the minimum value to 2, and the maximum value to 6.

12. Expand the **Dependencies** area. Under **Visibility**, type ["\$radiocluster","=","specify"].

OUTPUT Pane

- 13. In the Control Library, click **Add Page**. With this new page selected, set **ID** to **output** and **Label** to **OUTPUT**.
- 14. Add an output table to the page. Set **ID** to **outputdata** and **Label** to **Select location of output table**. Select the box for **Required**.
- 15. Go to the PROGRAM CODE area. Copy and paste the provided code into the Program area within the Custom Step Build area. You can also copy and paste the program code from the provided TXT file.
- 16. Save the step using the name **KCLUSExample**.

PROGRAM CODE Area

The following program code should be copied into the Program area of the custom step. It is important that the IDs that were created within the Designer are exact as they are referenced in the program as macro variables.

PROMPT UI

The following UI code is generated as items are added to the Designer. This code can be copied into the Prompt UI area of the custom step. After they are copied, the corresponding items appear in the Designer.

```
{
     "showPageContentOnly": true,
     "pages": [
           {
                "id": "data",
                "type": "page",
                "label": "DATA",
                "children": [
                      {
                           "id": "inputdata",
                           "type": "inputtable",
                           "label": "Select table for analysis",
                           "required": true,
                           "placeholder": "",
                           "visible": ""
                      },
                           "id": "catvars",
                           "type": "columnselector",
                           "label": "Select categorical inputs",
                           "order": false,
                           "columntype": "a",
                           "max": null,
                           "min": null,
                           "visible": "",
                           "table": "inputdata"
                      },
```

```
"id": "intvars",
                           "type": "columnselector",
                           "label": "Select continuous inputs",
                           "order": false,
                           "columntype": "n",
                           "max": null,
                           "min": null,
                           "visible": "",
                           "table": "inputdata"
                     }
                ]
           },
                "id": "options",
                "type": "page",
                "label": "OPTIONS",
                "children": [
                     {
                           "id": "imputeint",
                           "type": "checkbox",
                           "label": "Replace interval missing values
with mean.",
                           "visible": ""
                     },
                           "id": "imputecat",
                           "type": "checkbox",
                           "label": "Replace nominal missing values
with mode.",
                           "visible": ""
                     },
                           "id": "radiocluster",
                           "type": "radiogroup",
                           "label": "Number of clusters:",
                           "items": [
                                {
                                      "value": "specify",
                                      "label": "Specify number of
clusters"
                                },
                                      "value": "calculate",
                                      "label": "Calculate number of
clusters"
                                }
                           "visible": ""
```

```
},
                      {
                           "id": "numcluster",
                           "type": "numberfield",
                           "label": "Maximum number of clusters",
                           "placeholder": "",
                           "required": false,
                           "max": 6,
                           "min": 2,
                           "excludemin": false,
                           "excludemax": false,
                           "visible": [
                                 "$radiocluster",
                                 "=",
                                 "specify"
                           1
                      }
                ]
           },
                "id": "output",
                "type": "page",
                "label": "OUTPUT",
                "children": [
                      {
                           "id": "outputdata",
                           "type": "outputtable",
                           "label": "Select location of output
table.",
                           "required": true,
                           "placeholder": "",
                           "visible": ""
                      }
                ]
           }
     ],
     "values": {
           "inputdata": {
                "library": "",
                "table": ""
           },
           "catvars": [],
           "intvars": [],
           "imputeint": false,
           "imputecat": false,
           "radiocluster": {
                "value": "specify",
                "label": "Specify number of clusters"
           },
```

Developing a Custom Step for the REGSELECT Procedure

1. Starting from SAS Drive, use the applications menu to proceed to **Develop Code and Flows**. From the menu bar, click **New** and select **Custom Step**.

DATA Pane

- 2. The Custom Step area opens into the Designer with Page 1 appearing. On the properties panel, set the ID of the page to **data** and the label to **DATA**.
- In the Control Library, scroll down to the Data area and drag Input Table to the DATA page. Set ID to dataset and Label to Select data set to analyze:. Leave the check box next to Required selected.
- 4. Under the common area of the Control Library, drag a **Text** item to the bottom of the page. Set **ID** to **labelPartition** and **Enter text** to **Partition Data**:
- 5. Drag a second **Text** item to the bottom of the page. Set **ID** to **partitiontext** and **Enter text** to **In** addition to training data, the data set contains:
- 6. Add three text or numeric fields to the page. Set the following properties:

	Numeric Field 1	Numeric Field 2	Numeric Field 3
ID	ValidProp	TestProp	seed
Label	bel Proportion of Proportion of test Inpu		Input a random generating seed:
	validation cases:	cases:	
Туре	Numeric	Numeric	Numeric
Default Value	0	0	8675309
Minimum Value	0	0	0 (Exclude min in range)
Maximum Value	1	1	

- 7. Drag a **Text** item to the page. Set **ID** to **labelRoles** and **Enter text** to **Roles**:.
- 8. Add three Column Selector items to the page. Set the following properties:

	Column Selector 1	Column Selector 2	Column Selector 3
ID	target	intvars	catvars
Label	Interval Target:	Interval Inputs:	Nominal Inputs:
Link to input table	(dataset)	(dataset)	(dataset)
Column Type	Numeric	Numeric	All types
Min Columns	1		
Max Columns	1		

MODELOPTIONS Pane

- 9. In the Control Library, click **Add Page**. With this new page selected, set **ID** to **model** and **Label** to **MODELOPTIONS** in the properties panel.
- 10. Add four Check Box items to the page. Set the following properties:

	Check Box 1	Check Box 2	Check Box 3	Check Box 4
ID	clbeta	type3	stbeta	chkvif
Label	Display confidence limits	Display the Type 3/Joint tests of	Display standardized	Display variance inflation factors
	confidence infins	3/301111 16515 01	Stariuaruizeu	IIIIIalion lactors
	for parameters	effects	beta estimates	for the estimates

SELECTION Pane

- 11. In the Control Library, click **Add Page**. With this new page selected, set **ID** to **selection** and **Label** to **SELECTION**.
- 12. Add a Text item to the page. Set ID to modelselectlabel and Enter text to Model Selection:.
- 13. Add a Check Box item to the page. Set **ID** to **checkselect** and **Label** to **Process Using Model Selection**.
- 14. Add a radio button group to the page. Set ID to radioSelect and leave Label blank. Click Add Many to add several items to the radio list. In the box that appears, enter Forward Selection, Backwards Elimination, and Stepwise Selection, where each option appears on its own line. Click OK.
- 15. Click the **Map Values** button to expand the options within the radio list. A new column named **Value** appears. On each line, change **Value** to **forward**, **backward**, and **stepwise** respectively.
- 16. Set the default radio button to **Stepwise Selection**.
- 17. Expand the **Dependencies** area. Under **Visibility**, type **"\$checkselect"**.
- 18. Add three drop-down lists to the page. Set the following properties:

	Drop-Down List 1	Drop-Down List 2	Drop-Down List 3
ID	select	stop	choose
Label	Add/Remove Inputs based upon:	Stop adding/removing inputs based upon:	Select best model using:
Determine where values come from	Static List	Static List	Static List

- 19. For the select drop-down List, click **Add Many**. In the box that appears, enter **Adjusted R-sq**, **AIC**, **AICC**, **Mallows' C(p)**, **R-sq**, **SBC**, and **Significance**, where each option appears on its own line. Click **OK**.
- 20. Click the **Map Values** button to expand the options within the radio list. A new column named **Value** appears. On each line, change the **Value** to **ADJRSQ**, **AIC**, **AICC**, **CP**, **R-sq**, **SBC**, and **Significance** respectively.
- 21. Set the **Default** item to **SBC**.

- 22. Expand the **Dependencies** area. Under **Visibility**, type "\$checkselect".
- 23. For the stop drop-down list, click **Add Many**. In the box that appears, enter **Adjusted R-sq**, **AIC**, **AICC**, **Mallows' C(p)**, **SBC**, **Significance**, and **None**, where each option appears on its own line. Click **OK**.
- 24. Click the **Map Values** button to expand the options within the radio list. A new column named **Value** appears. On each line, change **Value** to **ADJRSQ**, **AIC**, **AICC**, **CP**, **SBC**, **Significance**, and **None** respectively.
- 25. Set the **Default** item to **SBC**.
- 26. Expand the **Dependencies** area. Under **Visibility**, type "\$checkselect".
- 27. For the choose drop-down list, click **Add Many**. In the box that appears, enter **Adjusted R-sq**, **AIC**, **AICC**, **Mallows' C(p)**, and **SBC**, where each option appears on its own line. Click **OK**.
- 28. Click the **Map Values** button to expand the options within the radio list. A new column named **Value** appears. On each line, change **Value** to **ADJRSQ**, **AIC**, **AICC**, **CP**, and **SBC** respectively.
- 29. Set the **Default** item to **SBC**.
- 30. Expand the **Dependencies** area. Under **Visibility**, type "\$checkselect".
- 31. Add two text or numeric fields to the page. Set the following properties:

	Numeric Field 1	Numeric Field 2
ID	sle	sls
Label	Significance level to enter the model:	Significance level to remain in model:
Туре	Numeric	Numeric
Default Value	0.05	0.05
Minimum Value	0.01	0.01
Maximum Value	0.99	0.99

32. For both numeric fields, expand **Dependencies**. Under **Visibility**, type "\$checkselect".

OUTPUT Pane

- 33. In the Control Library, click **Add Page**. With this new page selected, set **ID** to **output** and **Label** to **OUTPUT**.
- 34. Add a check box to the page. Set **ID** to **checkoutput** and **Label** to **Include output statistics**.
- 35. Add an output table to the page. Set **ID** to **outputtable** and **Label** to **Select the CAS data table to contain the desired statistics:**. Select the box for **Required**. Expand **Dependencies**. Under **Visibility**, type "\$checkoutput".
- 36. Add a check box to the page. Set **ID** to **checkcopyvars** and **Label** to **Include the model variables in the output table**. Expand **Dependencies**. Under **Visibility**, type **"\$checkoutput"**.
- 37. Add a Text object to the page. Set **ID** to **text1** and **Enter text** to **Additional statistics to save in output table:**. Expand **Dependencies**. Under **Visibility**, type "\$checkoutput".

38. Add four check boxes to the page. Set the following properties:

	Check Box 1	Check Box 2	Check Box 3	Check Box 4
ID	checkpred	checkres	checkconfidence	checkinfluence
Label	Predictions	Residuals	Confidence Limits Influence Statistics (COOKSD, COVRATIO,	
			Mean)	so on)

- 39. For each of these check boxes, expand **Dependencies**. Under **Visibility**, type "\$checkoutput".
- 40. Go to the PROGRAM CODE area. Copy and paste the provided code into the Program area within the Custom Step Build area. You can also copy and paste the program code from the provided TXT file.
- 41. Save the step using the name **REGSELECTExample**.

PROGRAM CODE Area

The following program code should be copied into the Program area of the custom step. It is important that the IDs that were created within the Designer are exact because they are referenced in the program as macro variables.

```
****************
*Creates the lists of categorical and interval inputs;
%let catvarslist=;
%let intvarslist=;
%macro predictors;
%do i=1 %to &catvars count;
     %let catvarslist = &catvarslist &&catvars &i. name;
%end;
%do j=1 %to &intvars count;
     %let intvarslist = &intvarslist &&intvars &j. name;
%end;
%mend;
%predictors;
%let predictorlist = &catvarslist &intvarslist;
*****************
*Creates the options line for the model statement;
%let modeloptions=;
%macro model;
%if (&clbeta) %then %do; %let modeloptions= &modeloptions CLB;
%end;
```

```
%if (&type3) %then %do; %let modeloptions= &modeloptions SS3; %end;
%if (&stbeta) %then %do; %let modeloptions= &modeloptions STB;
%end;
%if (&chkvif) %then %do; %let modeloptions= &modeloptions VIF;
%end;
%mend;
%model;
****************
*Creates the options for the selections statement after the method;
%let selectionoptions=;
%macro selections;
%if (&checkselect) %then %do;
%let selectionoptions = selection method= &radioSelect (select=
&select stop=&stop choose=&choose ;
%if (&radioSelect=forward) %then %do; %let selectionoptions =
&selectionoptions sle=&sle); %end;
%if (&radioSelect=backward) %then %do; %let selectionoptions =
&selectionoptions sls=&sls); %end;
%if (&radioSelect=stepwise) %then %do; %let selectionoptions =
&selectionoptions sls=&sls sle=&sle); %end;
                                      %end;
%mend;
%selections;
*****************
*Creates the options for the OUTPUT statement;
%let outputoptions=;
%macro outputopt;
%if (&checkoutput) %then %do;
%let outputoptions= &outputoptions OUTPUT out=&outputtable ;
```

```
%if (&checkcopyvars) %then %do; %let outputoptions = &outputoptions
COPYVARS= (&predictorlist) ; %end;
%if (&checkpred) %then %do; %let outputoptions = &outputoptions
PRED ; %end;
%if (&checkres) %then %do; %let outputoptions = &outputoptions
RESID ; %end;
%if (&checkconfidence) %then %do; %let outputoptions =
&outputoptions LCL=LCL LCLM=LCLM UCL=UCL UCLM=UCLM ; %end;
%if (&checkinfluence) %then %do; %let outputoptions =
&outputoptions COOKSD=COOKSD COVRATIO=COVRATIO DFFIT=DFFIT H=H
LIKEDIST=LIKEDIST PRESS=PRESS RSTUDENT=RSTUDENT STUDENT=STUDENT ;
%end;
                                            %end;
%mend;
%outputopt;
******************
*Core code that takes advantage of the preceding code generations;
proc regselect data=&dataset;
class &catvarslist;
model &target = &predictorlist / &modeloptions;
partition fraction(test=&TestProp validate=&ValidProp seed=&seed);
&selectionoptions;
&outputoptions;
run; quit;
******************
```

PROMPT UI

The following UI code is generated as items are added to the Designer. This code can be copied into the Prompt UI area of the custom step. After they are copied, the corresponding items will appear in the Designer.

```
{
     "showPageContentOnly": true,
     "pages": [
           {
                "id": "data",
                "type": "page",
                "label": "DATA",
                "children": [
                     {
                           "id": "dataset",
                           "type": "inputtable",
                           "label": "Select data set to analyze:",
                           "required": true,
                           "placeholder": "",
                           "visible": ""
                     },
                           "id": "labelPartition",
                           "type": "text",
                           "text": "Partition Data:",
                           "visible": ""
                     },
                      {
                           "id": "partitiontext",
                           "type": "text",
                           "text": "In addition to training data,
the data set contains:",
                           "visible": ""
                     },
                           "id": "ValidProp",
                           "type": "numberfield",
                           "label": "Proportion of validation
cases:",
                           "placeholder": "",
                           "required": false,
                           "max": 1,
                           "min": 0,
                           "excludemin": false,
                           "excludemax": false,
                           "visible": ""
                     },
```

```
"id": "TestProp",
                           "type": "numberfield",
                           "label": "Proportion of test cases:",
                           "placeholder": "",
                           "required": false,
                           "max": 1,
                           "min": 0,
                           "excludemin": false,
                           "excludemax": false,
                           "visible": ""
                     },
                           "id": "seed",
                           "type": "numberfield",
                           "label": "Input a random generating
seed:",
                           "placeholder": "",
                           "required": false,
                           "max": null,
                           "min": 0,
                           "visible": "",
                           "excludemin": true
                     },
                           "id": "labelRoles",
                           "type": "text",
                           "text": "Roles:",
                           "visible": ""
                     },
                      {
                           "id": "target",
                           "type": "columnselector",
                           "label": "Interval target:",
                           "order": false,
                           "columntype": "n",
                           "max": 1,
                           "min": 1,
                           "visible": "",
                           "table": "dataset"
                     },
                           "id": "intvars",
                           "type": "columnselector",
                           "label": "Interval Inputs:",
                           "order": false,
                           "columntype": "n",
                           "max": null,
                           "min": null,
                           "visible": "",
```

```
"table": "dataset"
                      },
                           "id": "catvars",
                           "type": "columnselector",
                           "label": "Nominal Inputs:",
                           "order": false,
                           "columntype": "a",
                           "max": null,
                           "min": null,
                           "visible": "",
                           "table": "dataset"
                      }
                1
           },
                "id": "model",
                "type": "page",
                "label": "MODELOPTIONS",
                "children": [
                      {
                           "id": "clbeta",
                           "type": "checkbox",
                           "label": "Display confidence limits for
parameters",
                           "visible": ""
                      },
                           "id": "type3",
                           "type": "checkbox",
                           "label": "Display the Type 3/Joint tests
of effects",
                           "visible": ""
                      },
                           "id": "stbeta",
                           "type": "checkbox",
                           "label": "Display standardized beta
estimates",
                           "visible": ""
                      },
                      {
                           "id": "chkvif",
                           "type": "checkbox",
                           "label": "Display variance inflation
factors for the estimates",
                           "visible": ""
                      }
```

```
},
                "id": "selection",
                "type": "page",
                "label": "SELECTION",
                "children": [
                     {
                           "id": "modelselectLabel",
                           "type": "text",
                           "text": "Model Selection:",
                           "visible": ""
                     },
                      {
                           "id": "checkselect",
                           "type": "checkbox",
                           "label": "Process Using Model Selection",
                           "visible": ""
                     },
                           "id": "radioSelect",
                           "type": "radiogroup",
                           "label": "",
                           "items": [
                                {
                                      "value": "forward",
                                      "label": "Forward Selection"
                                },
                                 {
                                      "value": "backward",
                                      "label": "Backwards
Elimination"
                                },
                                      "value": "stepwise",
                                      "label": "Stepwise Selection"
                                }
                           "visible": "$checkselect"
                     },
                           "id": "select",
                           "type": "dropdown",
                           "label": "Add/Remove Inputs based upon:",
                           "items": [
                                 {
                                      "value": "ADJRSQ",
                                      "label": "Adjusted R-sq"
                                },
```

```
"value": "AIC",
                                      "label": "AIC"
                                 },
                                 {
                                      "value": "AICC",
                                      "label": "AICC"
                                 },
                                 {
                                      "value": "CP",
                                      "label": "Mallows' C(p)"
                                 },
                                 {
                                      "value": "R-sq",
                                      "label": "R-sq"
                                 },
                                      "value": "SBC",
                                      "label": "SBC"
                                 },
                                      "value": "Significance",
                                      "label": "Significance"
                                 }
                           ],
                           "required": false,
                           "placeholder": "",
                           "visible": "$checkselect"
                      },
                           "id": "stop",
                           "type": "dropdown",
                           "label": "Stop adding/removing inputs
based upon:",
                           "items": [
                                 {
                                      "value": "ADJRSQ",
                                      "label": "Adjusted R-sq"
                                 },
                                 {
                                      "value": "AIC",
                                      "label": "AIC"
                                 },
                                      "value": "AICC",
                                      "label": "AICC"
                                 },
                                 {
                                      "value": "CP",
                                      "label": "Mallows' C(p)"
```

```
},
           {
                "value": "SBC",
                "label": "SBC"
           },
                "value": "Significance",
                "label": "Significance"
           },
           {
                "value": "None",
                "label": "None"
           }
     ],
     "required": false,
     "placeholder": "",
     "visible": "$checkselect"
},
{
     "id": "choose",
     "type": "dropdown",
     "label": "Select best model using:",
     "items": [
           {
                "value": "ADJRSQ",
                "label": "Adjusted R-sq"
           },
           {
                "value": "AIC",
                "label": "AIC"
           },
                "value": "AICC",
                "label": "AICC"
           },
                "value": "CP",
                "label": "Mallows' C(p)"
           },
                "value": "SBC",
                "label": "SBC"
           }
     ],
     "required": false,
     "placeholder": "",
     "visible": "$checkselect"
},
```

```
"id": "sle",
                           "type": "numberfield",
                           "label": "Significance level to enter the
model:",
                           "placeholder": "",
                           "required": false,
                           "max": 0.99,
                           "min": 0.01,
                           "visible": "$checkselect"
                     },
                           "id": "sls",
                           "type": "numberfield",
                           "label": "Significance level to remain in
model:",
                           "placeholder": "",
                           "required": false,
                           "max": 0.99,
                           "min": 0.01,
                           "visible": "$checkselect"
                     }
                ]
          },
                "id": "output",
                "type": "page",
                "label": "OUTPUT",
                "children": [
                     {
                           "id": "checkoutput",
                           "type": "checkbox",
                           "label": "Include output statistics",
                           "visible": ""
                     },
                           "id": "outputtable",
                           "type": "outputtable",
                           "label": "Select the CAS data table to
contain the desired statistics:",
                           "required": true,
                           "placeholder": "",
                           "visible": "$checkoutput"
                     },
                           "id": "checkcopyvars",
                           "type": "checkbox",
                           "label": "Include the model variables in
the output table.",
```

```
"visible": "$checkoutput"
                      },
                      {
                           "id": "text1",
                           "type": "text",
                           "text": "Additional statistics to save in
output table:",
                           "visible": "$checkoutput"
                      },
                      {
                           "id": "checkpred",
                           "type": "checkbox",
                           "label": "Predicitions",
                           "visible": "$checkoutput"
                      },
                      {
                           "id": "checkres",
                           "type": "checkbox",
                           "label": "Residuals",
                           "visible": "$checkoutput"
                      },
                           "id": "checkconfidence",
                           "type": "checkbox",
                           "label": "Confidence Limits (Individual
and Mean)",
                           "visible": "$checkoutput"
                      },
                           "id": "checkinfluence",
                           "type": "checkbox",
                           "label": "Influence Statistics (COOKSD,
COVRATIO, etc)",
                           "visible": "$checkoutput"
                      }
                1
           }
     ],
     "values": {
           "dataset": {
                "library": "",
                "table": ""
           "ValidProp": 0,
           "TestProp": 0,
           "seed": 8675309,
           "target": [],
           "intvars": [],
           "catvars": [],
```

```
"clbeta": false,
           "type3": false,
           "stbeta": false,
           "chkvif": false,
           "checkselect": false,
           "radioSelect": {
                "value": "stepwise",
                "label": "Stepwise Selection"
           },
           "select": {
                "value": "SBC",
                "label": "SBC"
           },
           "stop": {
                "value": "SBC",
                "label": "SBC"
           },
           "choose": {
                "value": "SBC",
                "label": "SBC"
           },
           "sle": 0.05,
           "sls": 0.05,
           "checkoutput": false,
           "outputtable": {
                "library": "",
                "table": ""
           },
           "checkcopyvars": false,
           "checkpred": false,
           "checkres": false,
           "checkconfidence": false,
           "checkinfluence": false
     }
}
```