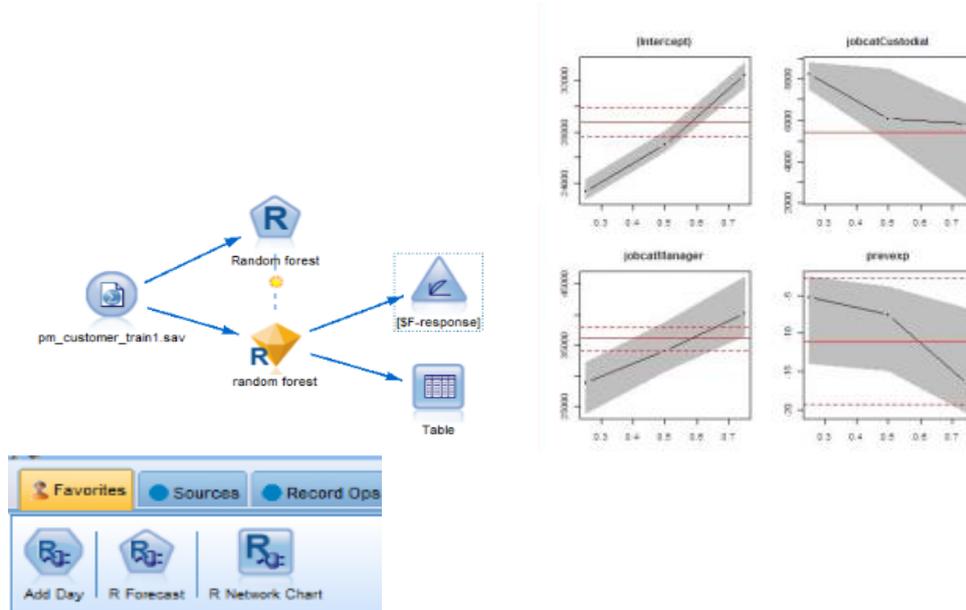


IBM SPSS Modeler en open source zoals R

- Native — first-class citizen
- Robust — data preparation, build/scoring and output
- Scale — score in-database and in-Hadoop with no data movement



Custom Dialog Builder

The Custom Dialog Builder interface includes a 'Tools' sidebar with various controls like Source List, Target List, Check Box, Combo Box, List Box, Text control, Number control, Static Text, Item Group, Radio Group, Check Box Group, File Browser, and Sub-dialog Button. The main canvas shows a dialog design with 'Variables' (Variable1, Variable2, Variable3), 'Target List' (Variable), 'Predictors List' (Variable), and an 'Automatic Learning' checkbox.

The 'Syntax Template' dialog box contains the following text:

Enter your syntax template below. The syntax template is used to generate syntax for your dialog. Enter replaceable values using the corresponding control identifier enclosed within '%' symbols.

For example, one line of a syntax template may contain:
`/VARIABLES %%var_list%%`

Tip: You can select from a list of all available control identifiers to insert by positioning the text caret where you would like to insert an identifier and then pressing the <CTRL><SPACEBAR> keys together. The list contains the control identifiers followed by the elements of command syntax available for the current context.

The list of identifiers includes:

```

1 # model building
2   attach(modelerData)
3   model<-lm(%%tgt_var%% ~ %%pred_var%%)
4 # model
5
6 -----
7
8 # scoring
9   attach(modelerData)
10  output<-predict(model)
11 # output
12

```

Buttons: OK, Cancel, Help

Target Variable List Properties

Property	Value
Identifier	pred_var
Title	Predictors List
ToolTip	
Target list type	Multiple item list
Mnemonic Key	
Required for execution	True

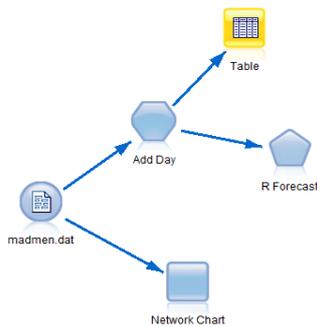
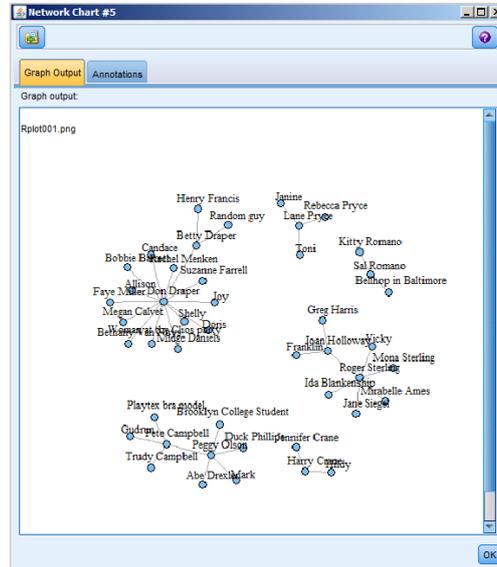
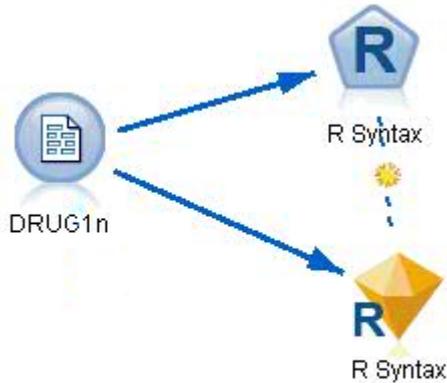
Predictive Extensions Demonstration

The diagram shows a workflow: Data → Auto Data Prep → Type → Select → Filler. From the Filler node, arrows point to Table, SQL, and Google Heatmaps (circled in red). The interface also shows a 'Streams' panel on the right with a tree view of extensions like Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment.

- Catalogue of extensions made available through community. Use what's there, add your own
- Take advantage of tutorials, usage examples, expertise, events, other downloads and more

Extend the Use of R

- Build and score R models through Modeler GUI
- Scale R execution by leveraging database vendor provided R engines
- Use R processes and generate output



R is growing in popularity.

Results for the most recent Rexer Analytics Annual Data Mining Survey had R usage at approximately 70% of respondents. So if you haven't come across it at an opportunity yet, you are either very new, or not trying hard enough.

Though its popularity it is not the only tool they are using, and in most cases, not their primary tool either!

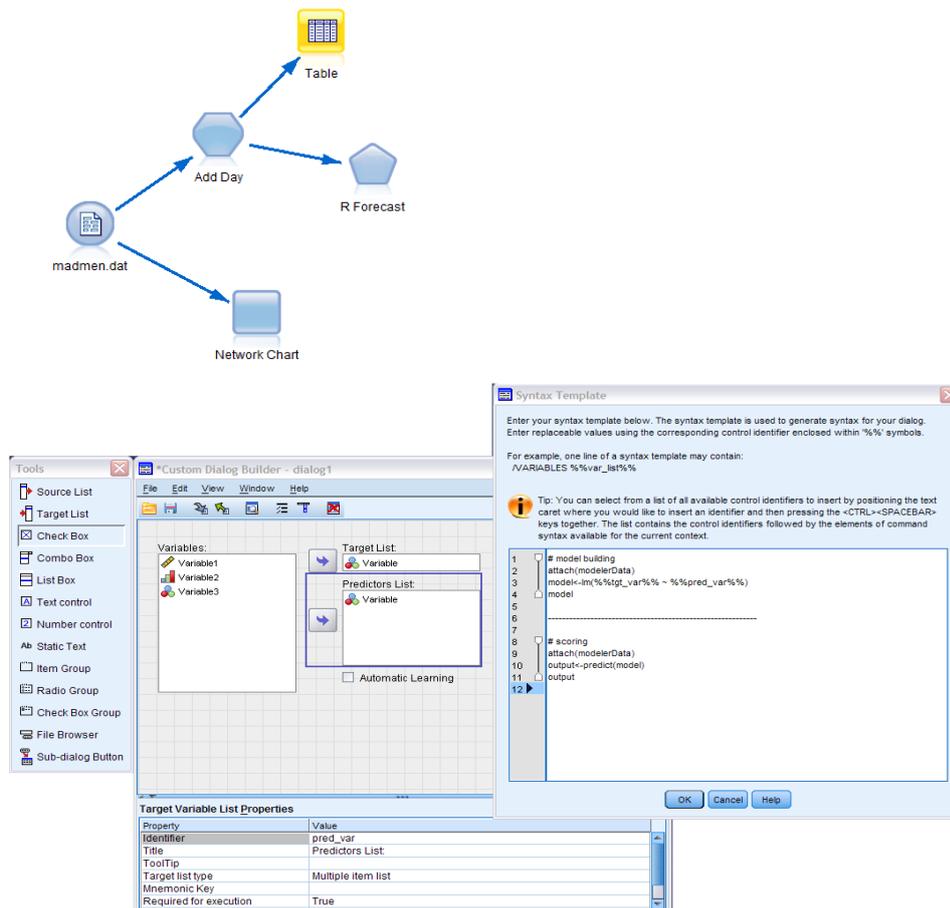
For those that aren't familiar with R, it is an open source programming language used for analytics. As its programming based, it doesn't have the sexy GUI of Modeler, but what it does have is a large dedicated user base that is constantly creating and publishing R procedures or packages. Version 16 of Modeler allows users to build and score models through the GUI, which means that the potential number of algorithms in Modeler just got a whole lot larger.

And with the use of vendor provided R extensions, can allow R to be scaled inside the database which will improve performance.

Finally if there are specific R outputs that are required, they can be generated also from Modeler.

Extend the Use of R

- Custom Dialog Builder for R
- Provides the ability to create new Modeler Algorithm nodes and dialogs that run R processes
- Makes R usable for non-programmers



Programming R is not going to appeal to everyone. They may not have the skill, they may prefer to use a GUI, or they may have used it that much at university or college that they just don't like it.

So the custom dialog builder will be right up their alley.

Similar to the capability that was added in Statistics many versions ago, you can now create a dialog box and use it for R.

This means that not only are we appealing those users that want to use R in addition to Modeler, Modeler becomes appealing to those that could or need to use R but can't.

This means that when you encounter R in an account, rather than viewing it as your competition and trying to displace it - embrace it.

We are a leader in deploying predictive analytics - adding R just strengthens the breadth of analytics that we can use - we make it very easy to use R and very scalable. On top of that, our other features enable you to do more with R - i.e. ease of data preparation - leverage ensembles, etc. And we have 40+ years of proven technology. The argument that we compete with R can really be used with a lot of what we do - you could say we compete with Netezza, etc - but our focus is analytics and allowing customers to build and deploy the best analytics - this allows them to pick and choose between our algorithms, netezza, R, etc and when you add Analytic Server to the mix, you can scale R to big data... to add the most business value.