Paper SAS4609-2020

Smarter and Faster with SAS® Visual Analytics

Rick Styll, SAS Institute Inc., Cary, NC

ABSTRACT

SAS® Visual Analytics is the smartest business intelligence tool available. Automated Explanation, the new name for automated analysis, has been rewritten and redesigned to give you smarter and clearer insights, more interactivity, and easier-to-read explanations. In seconds, you can get the analytical story for the business intelligence hidden in your data that would take you hours to do manually. On top of that, you can automatically see suggested visualizations and identify related measures. For more advanced analytical visualizations, like decision trees, you get human-friendly natural language descriptions, drawing out insights that are easy to digest.

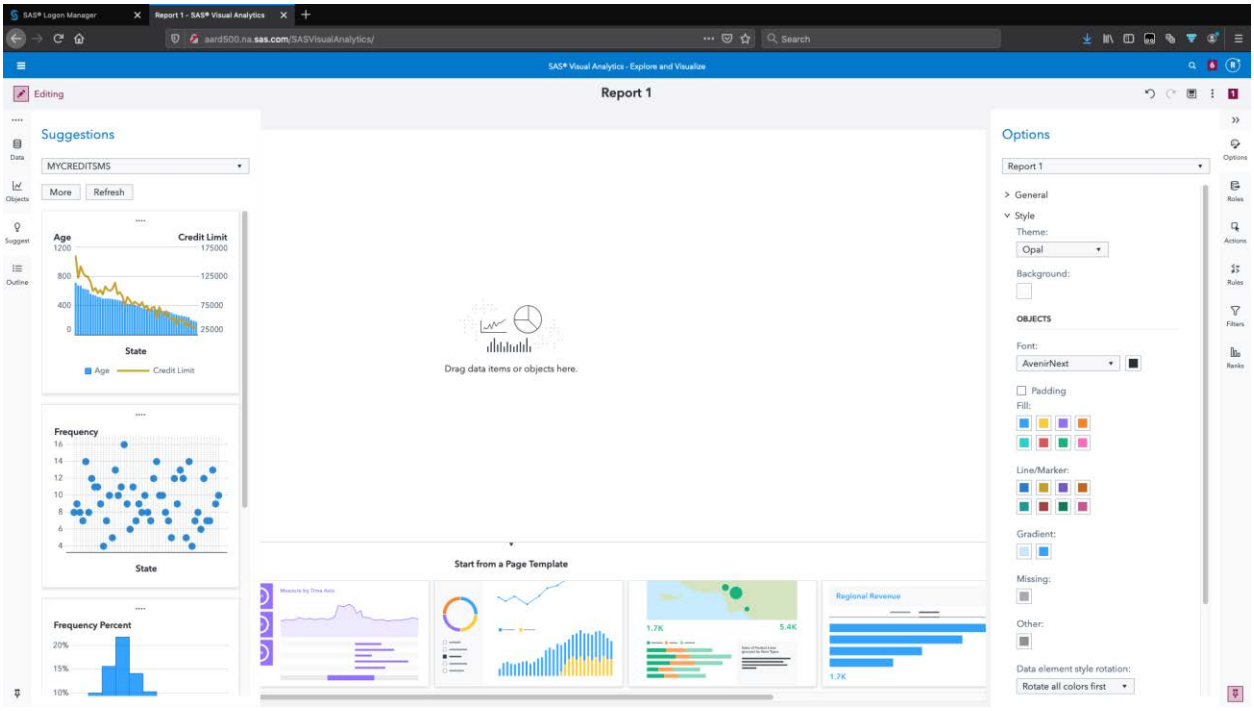
INTRODUCTION

SAS® Visual Analytics is growing smarter. With two releases of SAS Visual Analytics in 2019, SAS delivered new and enhanced AI-driven functionality so that you can do more, faster. Automated Analysis, introduced in SAS Visual Analytics 8.3 in September 2018, is now Automated Explanation, and it has been improved based on feedback from early users. In August 2019, SAS Visual Analytics 8.4 added suggested visualizations to the left pane to give you several smart visualizations without you first having to select data items. Strong correlation relationships between your measure data items are automatically highlighted subtly right in your left data pane. Automated Prediction was new for SAS Visual Analytics 8.5 in November 2019. You get a new predicted outcome for one data item you are interested in whenever you tweak input values for one or more data items that SAS Visual Analytics automatically displays as strong predictor data items. If your site is licensed for SAS® Data Preparation, the SAS Data Studio interface included in SAS Visual Analytics includes additional suggestions that automatically recommend data quality and other transformations to your data. You will learn more about each of these automated features. With every release, SAS is weaving subtle yet powerful artificial intelligence throughout your entire experience.

SUGGESTED VISUALIZATIONS

Suggested Visualizations provide a quick and fun way to explore and visualize your data. SAS Visual Analytics has been suggesting charts to you since it launched in 2012. We called it Auto-Charting. You selected some columns from your data and when you dropped them on the canvas, you got a suggested chart.

With SAS Visual Analytics 8.4, Suggested Visualizations takes this to the next level. Now you don't have to select columns. Just open a dataset and open the Suggestions Pane to get suggested charts based on all your data. You can refresh the suggestions to get a whole new set of charts with the refresh button. You can click more to get more suggested charts added to the existing suggestions. When you find something you like, just drag and drop it into your report. You can find more details by referring to the product documentation (SAS Institute Inc., 2019a). You will see an example of Suggested Visualization in Display 1 on the following page. Suggested Visualizations doesn't build your report for you, but it makes it effortless to quickly explore a new dataset or learn more about a dataset that you're familiar with.



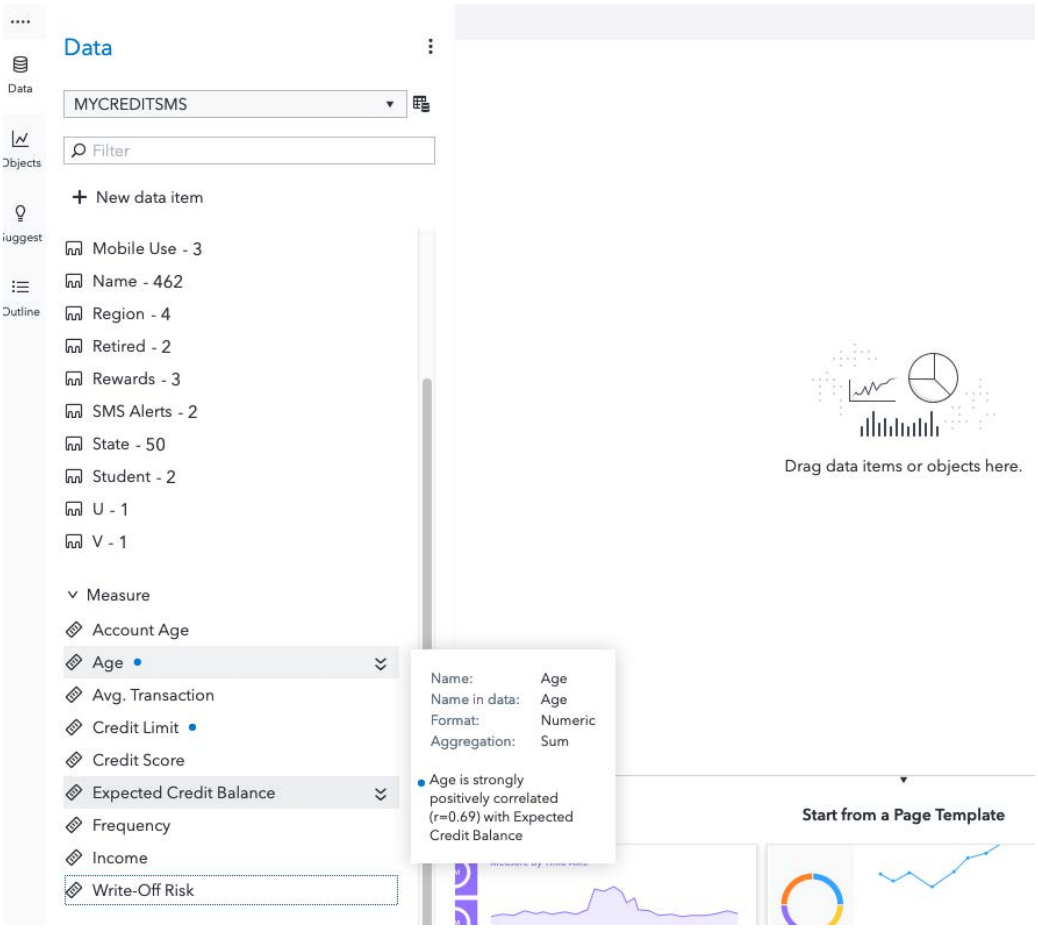
Display 1. Display of Suggested Visualizations From the Left Pane

Keep the following in mind when using Suggested Visualizations.

- The Suggestions Pane uses a series of heuristics to make chart recommendations.
- Geographical maps and hierarchies are currently not supported in suggestions.

RELATED MEASURES

One of the best examples of how SAS is weaving intelligence through SAS Visual Analytics is a feature that we call "Related Measures." The Related Measures feature automatically detects potential relationships in your measure data items and notifies you in a subtle way. As you work with data in the data pane, SAS Visual Analytics automatically discovers insights from your data. When you select a measure, SAS Visual Analytics determines correlations across all your measures, and indicates the relationship with a small blue indicator, as shown in Display 2 below. You can hold your pointer over the blue indicator to see the relationship with your original column selection.



Display 2. Age is Highlighted as a Strong Correlation with Expected Credit Balance

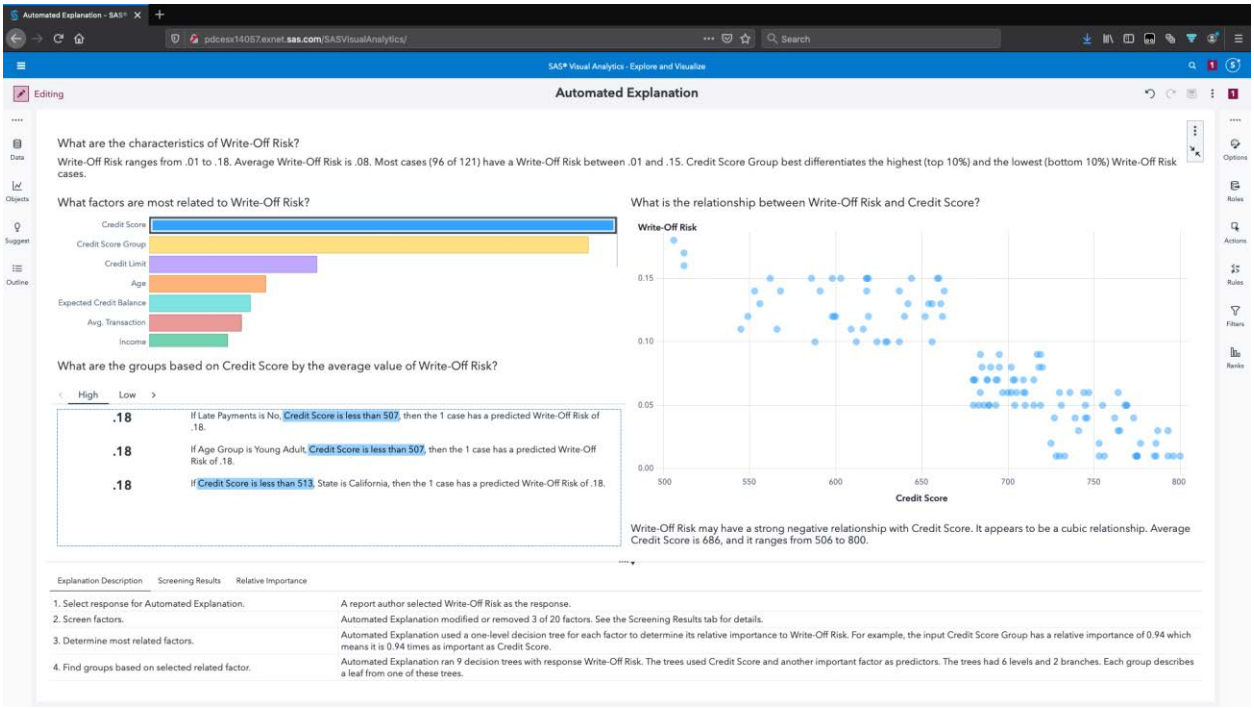
You should be aware of the following characteristics for Related Measures.

- Whether an indicator appears, depends on the size of the data. Very small datasets (<56 rows) have progressively higher correlation thresholds. This reduces the likelihood of highlighting correlations due to random chance (SAS Institute Inc., 2019b).
- Correlation thresholds are set at .6 for data with more than 56 observations. Lower thresholds would increase the likelihood of catching random chance correlations.
- Correlation thresholds are not user adjustable.
- Blue dots are not displayed in the Data pane if there are more than 1000 measures in the data source (SAS Institute Inc., 2019b).

AUTOMATED EXPLANATION

THE BASICS

With Automated Explanation you can get hours of analysis done in seconds. Automated Explanation is the fastest way to find insights in your data. You simply right-click the column you're interested in and click explain. Behind the scenes, SAS Visual Analytics screens your data, runs various algorithms based on the kind of data you're working with, and presents the results in an easy to understand analytical story, with natural language. You can quickly learn about the characteristics of the variable you're interested in, see how other columns help explain your data, and see groups within your data with common characteristics. Display 3 shows an example of Automated Explanation for Write-Off Risk.



Display 3. Automated Explanation Example for Write-Off Risk

- Automated Explanation is only supported in English. Support for other languages is planned.
- Automated Explanation can explain categorical and measure columns.

THE CHARACTERISTICS OF YOUR TARGET VARIABLE

Automated explanation helps you learn about the characteristics of the column you're interested in. The first section of the Automated Explanation results describes the characteristics of your data. For a measure, this includes the range, the average, the middle 80% of your data, the column that best differentiates the top and bottom 10% of your data, and outliers if there are any. For a categorical column, you'll learn what proportion of your data belongs to that category and how it ranks compared to other categories in that column.

To find what best differentiates the top 10% and bottom 10% of the data, SAS Visual Analytics runs a decision tree on the tails of your data. The most informative split in this tree is the column that best differentiates the tails of the data.

MOST RELEVANT FACTORS

Automated Explanation finds the factors that are most related to your column of interest. The screened factors that are most related to the column that is being explained are displayed in the bar chart. The most related factors are at the top, the least at the bottom. You can select one or more of these factors to get additional information about the relationship and see groups based on the factors you selected. If a factor doesn't make sense, you can select it (or multi-select multiple factors) and remove them from the explanation.

- You can right-click a related factor and have that factor explained to you in a new Automated Explanation object.
- These results are not from a decision tree, and will not be the same as a relative

importance plots in a decision tree object with the same variables.

- You can remove factors from the analysis
- You can adjust the number of related factors shown in the object in the options pane.

INTERACTING WITH RELATED FACTORS

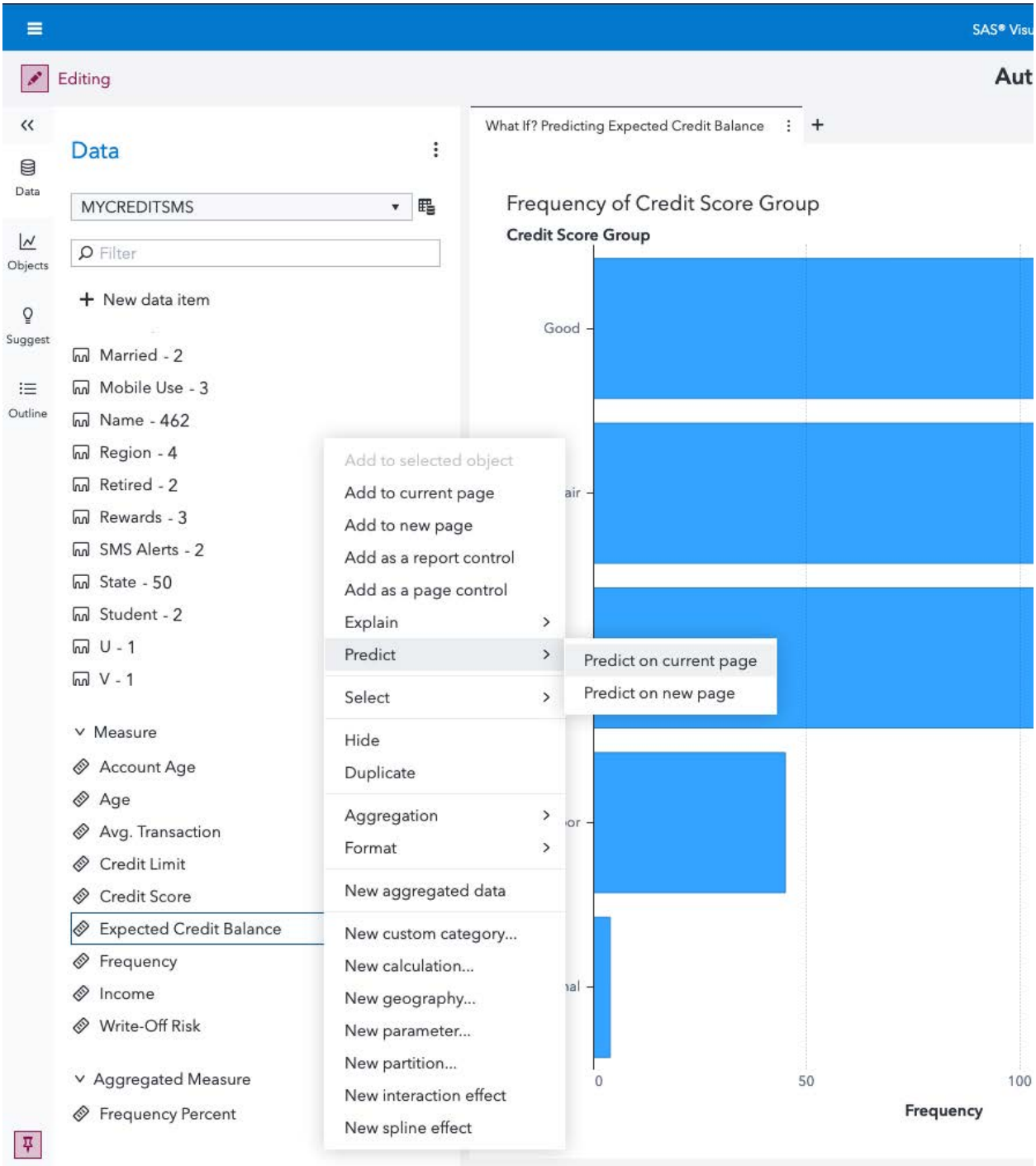
You can select a related factor to identify groups based on that factor and to learn exactly how that factor relates to your column of interest. You can select factors in the Related Factors chart and Automated Explanation generates new insights based on your selected factor and the variable that you want explained. Automated Explanation identifies high and low groups based on your selections. You can read about the characteristics of these groups or even derive the group as a column in your data that identifies whether each row belongs to that group. A visualization on the right lets you see the exact relationship between the factor you've selected and the variable you're interested in. Natural language explanation provides basic characteristics of the selected factor and a description of the relationship.

- You can adjust the number of groups Automated Explanation identifies in the options pane.
- High and low groups are identified using a data transform action that uses either decision trees or regression trees based on the input data.

For more details about Automated Explanation, you can refer to the product documentation (SAS Institute Inc., 2019c). In addition, a SAS Global Forum paper dedicated to Automated Explanation in 2019, when the feature was introduced as Automated Analysis, covers this topic in more detail (Styll, 2019).

AUTOMATED PREDICTION

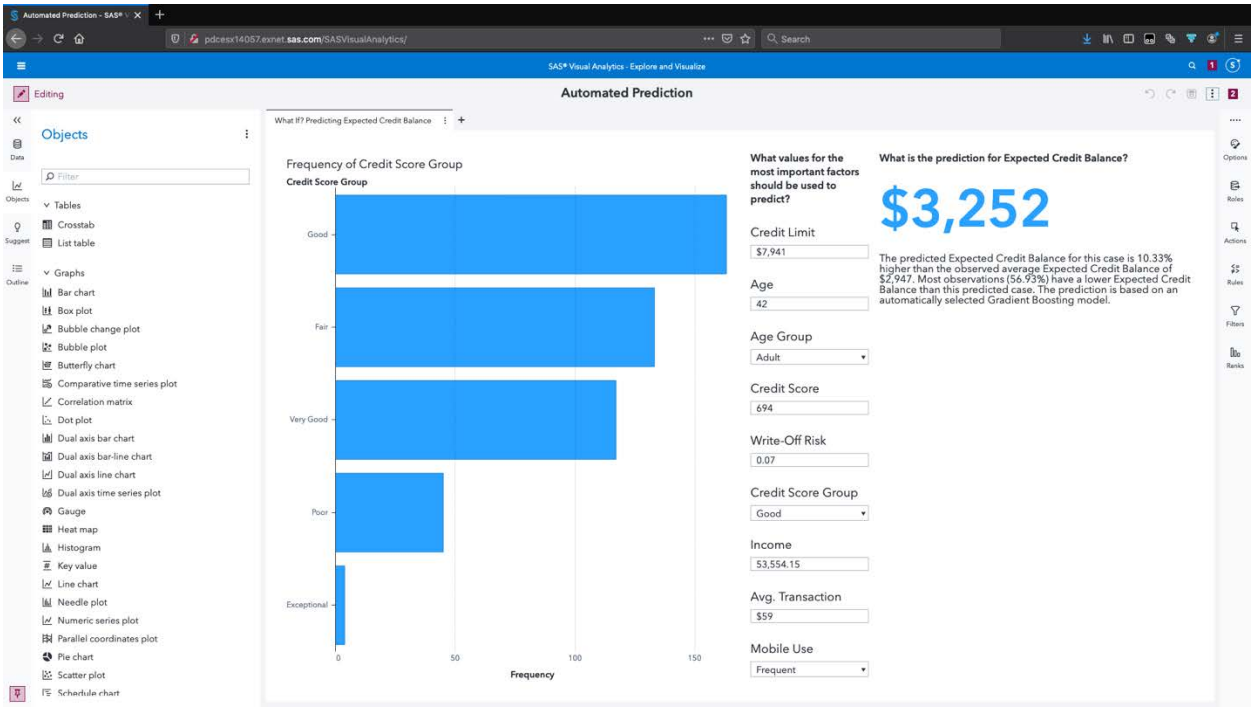
With a single click, you can have SAS Visual Analytics build multiple machine learning models for your data based on the data item that represents the response or outcome that you're interested in. SAS Visual Analytics then automatically chooses the best model to help you explore your data. In the following example, we'll select the Expected Credit Balance data item, as shown in Display 4. When you right-click Expected Credit Balance, you will see Predict in the menu with a choice to predict on the current page or a new page.



Display 4. Automated Prediction From the Data Item Right-Click Context Menu

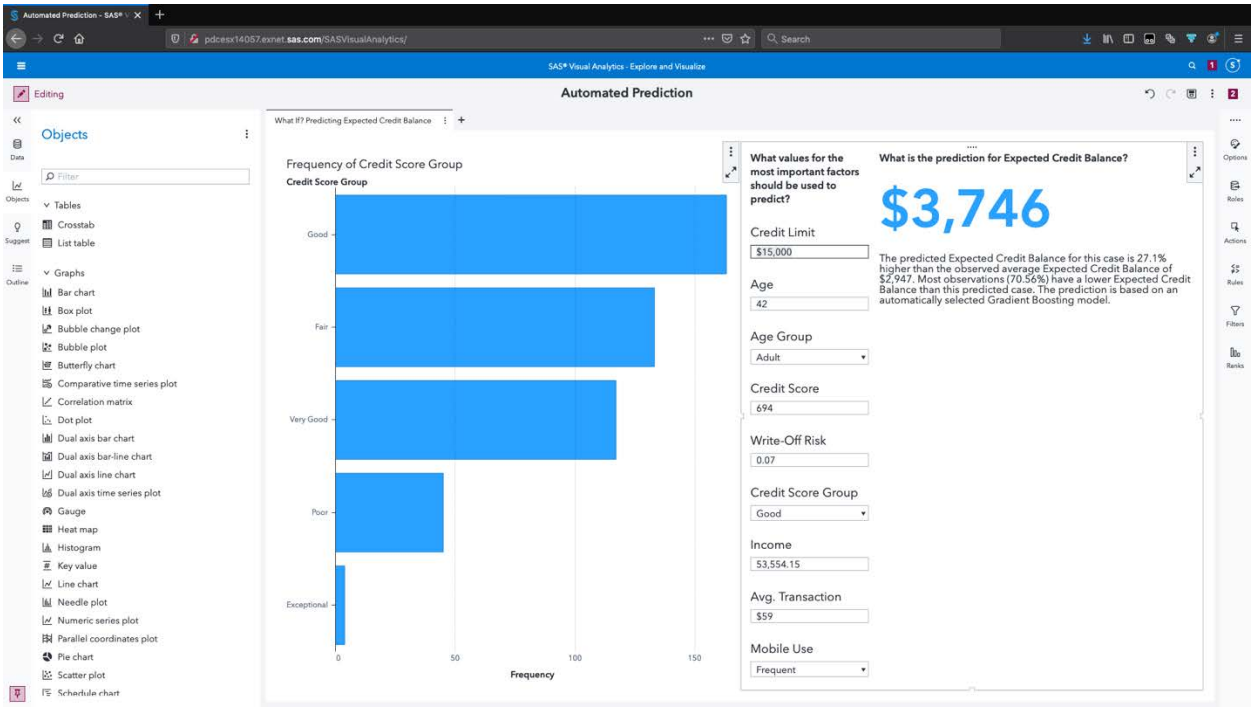
The automated prediction object runs multiple models on a response data item that you select in the data pane. Most of the other data items are automatically added to each model as underlying factors for model building. SAS Visual Analytics automatically determines the champion model. Then the predicted value for the response data item of interest along with the underlying factor data items are displayed within the report content. The initial model prediction is calculated by using the median values for measure data item factors or the most common value (mode) for category data item factors as inputs to each model. Consider, "What If?" At this point, you can adjust one or multiple values of the data item factors now displayed as form controls in the report in order to get a new prediction.

If your response data item of interest is a category, SAS Visual Analytics will automatically run a logistic regression (binary response only), gradient boosting, and decision tree models. The champion model is chosen based on the highest accuracy, which is 100 – misclassification rate. If your response data item is a measure, then a linear regression, gradient boosting, and decision tree models are run. The champion model is determined based on the lowest average squared error (ASE). The automated prediction does not update when page or report level filter controls change, nor does it update in response to other user actions on the report such as filter relationships between objects (SAS Institute Inc., 2019d).



Display 5. Automated Prediction Initial Display

In the example above in Display 5, the prediction object is added to the current page beside an existing bar chart. A predicted value for Expected Credit Balance of $3,252 is displayed in a large font near the top write. To the left, an input control for each underlying factor data item is presented with default values based on the rules above. Note that the default value for Credit Limit is $7,941. When you change Credit Limit to $15,000, as shown in Display 6, the predicted Expected Credit Balance changes to $3,746.

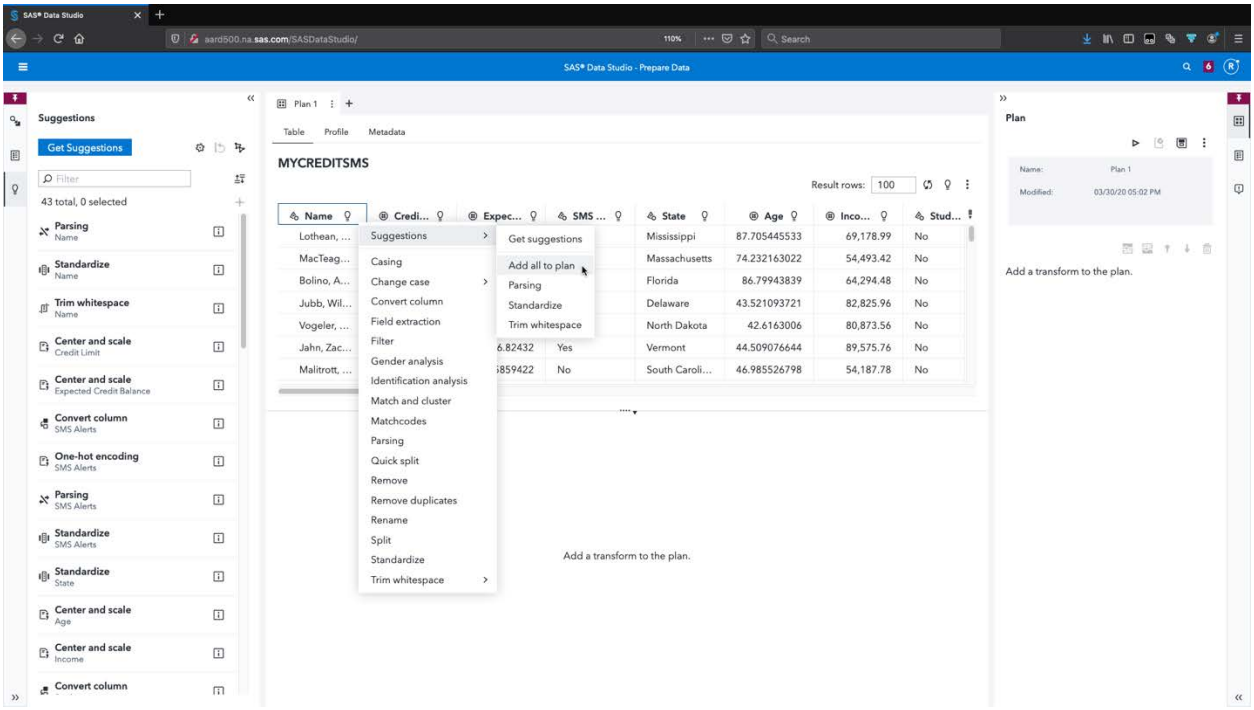


Display 6. Predicted Expected Credit Balance After Adjusting Credit Limit

DATA PREPARATION SUGGESTIONS

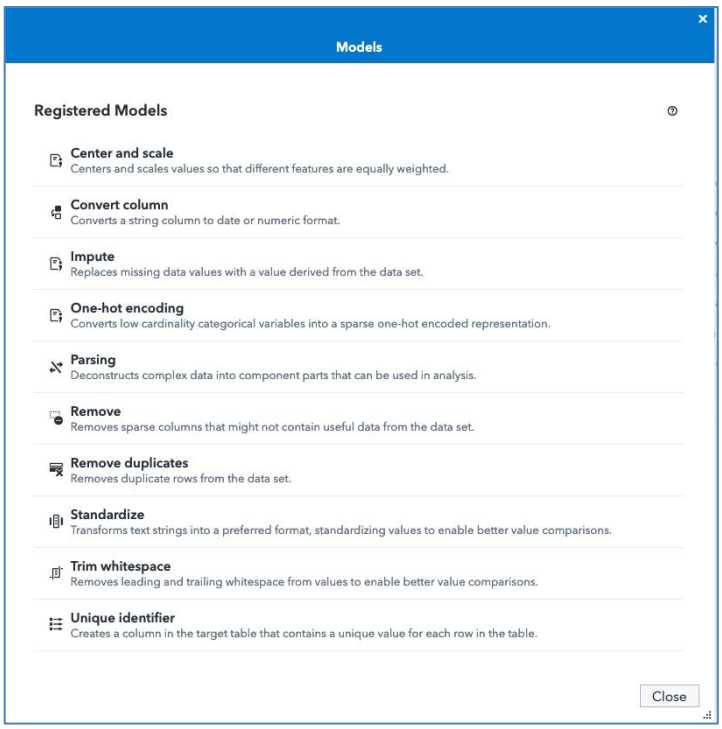
You can get data quality and transformation suggestions while preparing your data for analysis so that you can work faster and spend more time analyzing your data. If your site is licensed for SAS Data Preparation, and you are using SAS Data Studio through SAS Visual Analytics, then additional options are included. These options automatically suggest data quality and other transformations when getting data ready for analysis and reporting. Justifications are provided for each suggestion to help you better understand the recommendation and make the best decision for your analysis. Suggestions can automatically generate code-based analytic transformations like one-hot encoding, center and scale, and impute that are customized for your data – no manual coding needed.

To use suggestions in SAS Data Studio, you first select the lightbulb icon on the left top open the Suggestions pane, as show in Display 7 below. Clicking Get Suggestions initiates an automated analysis of your tables columns that produces a list of suggested transforms that will appear in the Suggestions pane. Depending on your settings for Suggestions, you might see a suggested column transforms, such as standardize and parse and you might see table suggestions, such as remove duplicate rows. You can add one or more of the suggested transforms to your plan using the plus icon near the top. Alternatively, you can add one or multiple suggestions to your plan for a specific column in your data with a right-click on any column that has a lightbulb icon.



Display 7. Automated Data Preparation Suggestions Based on Machine Learning

The following list of models shown in Display 8 represents the full set of suggestions available in SAS Data Studio.

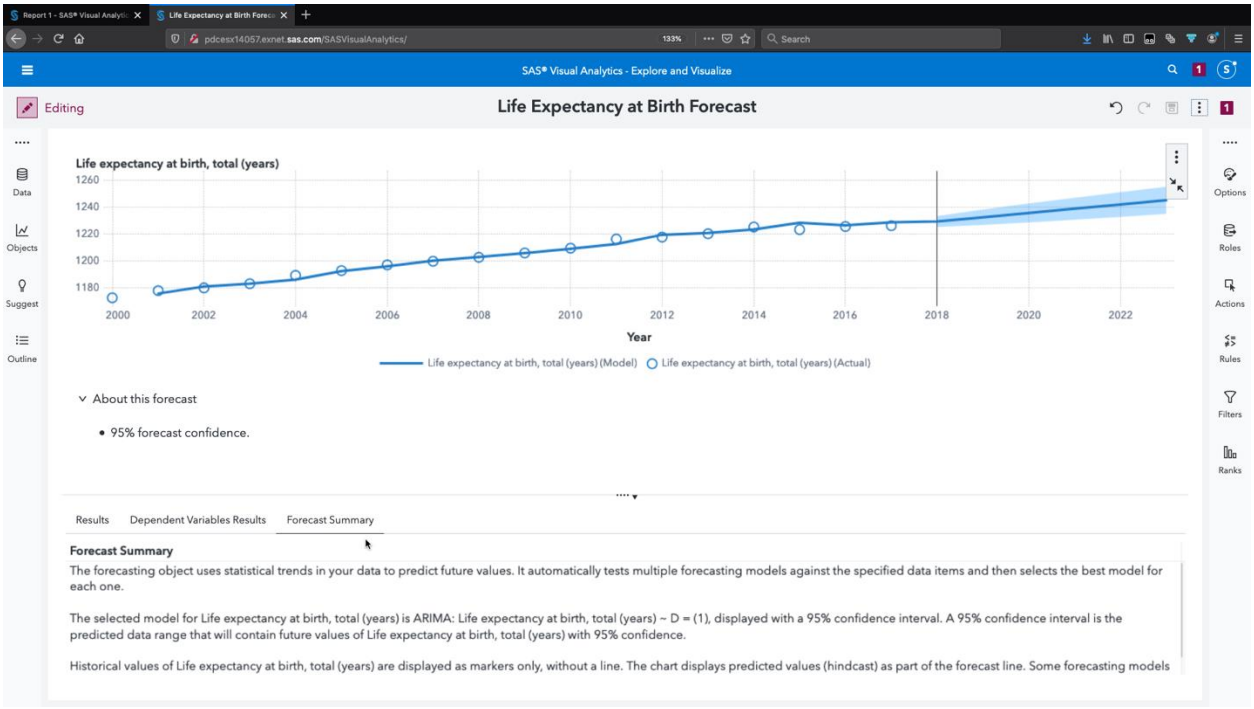


Display 8. List of Built-In Models for Data Preparation Suggestions

For more information about each of the steps above and more, please read Mary Kathryn Queen's recent blog post (Queen, 2020). You can also refer to the SAS Data Studio 2.5 product documentation for details (SAS Institute Inc., 2019e).

DYNAMIC EXPLANATION TEXT

Are you a data scientist? AI is for you too. Interpretability supports everyone whether they're prepping data, building reports, visualizing data, or building analytical models and pipelines. Dynamic explanation text on Forecast, Decision Tree, Text Topics, Network Analysis, Path Analysis, Automated Explanation, and Automated Prediction helps you to better understand the results. Even data preparation suggestions are clearly explained so that you can better understand what areas in your data could be improved and how you could improve them. Display 9 below shows a text-based summary of forecasted Life Expectancy at Birth just below the timeseries chart.



Display 9. Natural Language Explanation of a Life Expectancy at Birth Forecast

CONCLUSION

SAS is weaving intelligence into SAS Visual Analytics. Real artificial intelligence doesn't stand alone. Real artificial intelligence blends with the way you work. It makes everything you use smarter. Suggested visualizations provide a quick and fun way to explore and visualize your data. The Related Measures feature automatically detects potential relationships in your measure variables and then notifies you in a subtle way. Automated Explanation helps you to easily and quickly learn about the characteristics of the outcome that you're interested in. You can get suggestions on your data prep so that you can work faster. Dynamic explanation text on advanced visualizations helps you to better understand the results. SAS has woven some great new AI capabilities into SAS Visual Analytics.

REFERENCES

SAS Institute Inc. 2019. *SAS Visual Analytics 8.4: About the Suggestions Pane.* Cary, NC: SAS Institute Inc. Available https://go.documentation.sas.com/?cdcId=vacdc&cdcVersion=8.4&docsetId=vaobj&docsetTarget=p1rw5h4dxwv661n19ag9tvlmxi46.htm

SAS Institute Inc. 2019. *SAS Visual Analytics 8.4: View Related Measures in the Data Pane.* Cary, NC: SAS Institute Inc. Available https://go.documentation.sas.com/?cdcId=vacdc&cdcVersion=8.4&docsetId=vareportdata&docsetTarget=p0x9lknf2s6m9nn1xybuk6ereudi.htm#p1ljvjsdx0vdn1n17ohprqa61lki

SAS Institute Inc. 2019. *SAS Visual Analytics 8.4: Working with Automated Explanation Objects.* Cary, NC: SAS Institute Inc. Available https://go.documentation.sas.com/?cdcId=vacdc&cdcVersion=8.4&docsetId=vaobj&docsetTarget=n0ewwfd6udhv7qn1nropd18lof9j.htm

Styll, Rick. 2019. "AI is Coming for Your BI: Automated Analysis in SAS Visual Analytics." *Proceedings of the SAS Global 2019*, Cary, NC: SAS Institute Inc. Available https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3526-2019.pdf

SAS Institute Inc. 2019. *SAS Visual Analytics 8.5: Working with Automated Prediction Objects.* Cary, NC: SAS Institute Inc. Available https://go.documentation.sas.com/?cdcId=vacdc&cdcVersion=8.5&docsetId=vaobj&docsetTarget=p0sytm7v7ft0c6n1x2a7k5i8sa9b.htm

SAS Institute Inc. 2019. *SAS Data Studio 2.5: Working with Suggestions.* Cary, NC: SAS Institute Inc. Available https://go.documentation.sas.com/?docsetId=datastudioadv&docsetTarget=n07ph5njtwqsw6n1c23u70nii37j.htm&docsetVersion=2.5

Queen, Mary Kathrin. February 2020. "How to use Suggestions in SAS Data Studio." The SAS Data Science Blog. Available https://blogs.sas.com/content/subconsciousmusings/2020/02/24/how-to-use-suggestions-in-sas-data-studio. Last modified February 24, 2020. Accessed on March 13, 2020.

RECOMMENDED READING

- Queen, Mary Kathrin. "How to use Suggestions in SAS Data Studio" Available https://blogs.sas.com/content/subconsciousmusings/2020/02/24/how-to-use-suggestions-in-sas-data-studio. Last modified February 24, 2020. Accessed on March 13, 2020.
- Deigaard, Jeppe. "Juletip #13 The Future of Business Intelligence is Here" Available https://communities.sas.com/t5/SAS-Nordic-Users-Group/Juletip-13-The-Future-of-Business-Intelligence-is-here/gpm-p/611524# Accessed on March 13, 2020.
- Brown, Anna. "Automated Analysis What's the deal with AI-powered automated analysis in SAS Visual Analytics?" Available https://communities.sas.com/t5/SAS-Visual-Analytics/What-s-the-deal-with-AI-powered-automated-analysis-in-SAS-Visual/td-p/539716. Last modified March 1, 2019. Accessed on March 13, 2019.
- Styll, Rick. "AI is Coming for Your BI: Automated Analysis in SAS Visual Analytics" Available https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3526-2019.pdf Accessed on March 13, 2019.
- Carey, Melanie. "How SAS Visual Analytics' automated analysis takes customer care to the next level – Part 1" Available https://blogs.sas.com/content/sgf/2019/01/02/how-sas-visual-analytics-automated-analysis-takes-customer-care-to-the-next-level-part-1. Last modified January 2, 2019. Accessed on March 13, 2020.
- Carey, Melanie. "How SAS Visual Analytics' automated analysis takes customer care to the next level – Part 2" Available https://blogs.sas.com/content/sgf/2019/01/07/how-sas-visual-analytics-automated-analysis-takes-customer-care-to-the-next-level. Last modified January 7, 2019. Accessed on March 13, 2020.
- Carey, Melanie. "How SAS Visual Analytics' automated analysis takes customer care to the next level – Part 3" Available https://blogs.sas.com/content/sgf/2019/01/14/how-sas-visual-analytics-automated-analysis-takes-customer-care-to-the-next-level-2. Last modified January 14, 2019. Accessed on March 13, 2020.
- Assa, Atrin. "Automated Analysis with SAS Visual Analytics" Available https://www.youtube.com/watch?v=Jl-zZ7LhXfE. Available December 19, 2018. Accessed on March 13, 2020.
- Carey, Melanie. "Using the Automated Analysis Feature in SAS Visual Analytics in SAS Viya" Available https://video.sas.com/detail/video/5978210714001/using-the-automated-analysis-feature-in-sas®-visual-analytics-in-sas®-viya®. Available December 11, 2018. Accessed March 13, 2020.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Rick Styll
SAS Institute Inc.
rick.styll@sas.com