# Analytics Products Mathematics Standards

Our mathematics standards’ purpose is to ensure that the tools we deliver

1. are correct to the limits of machine precision
2. embody scientifically acceptable (generally accepted in the scientific literature, or proven optimal in a peer-reviewed journal article) state-of-the-art modeling techniques
3. expose commonly used model features
4. generally encourage end users to apply mathematical best practices.

Equally important, enforcing our math standards during the design process ensures our development activities have a predictable duration.

In what follows, reviewing the scientific literature means developing a comprehensive understanding of the state of the art about a family of algorithms (such as linear models or centered clustering models), such as one would obtain by reading a current systematic-review article in a peer-reviewed journal, or a beginning-graduate textbook on the subject. Where such reviews exist, they suffice.

(Continued)

## General Standards

|  |  |
| --- | --- |
| **ID** | **Description** |
| M1 | Each time we create a new tool or substantially enhance an existing tool, the primary developer will review relevant literature, compile a list of reviewed sources, save the list on Quip, and share the list with the math reviewers at least one business day before the math review. The secondary developer will read the material identified by the primary developer. The remainder of the team will review the same material to their individual satisfaction. |
| M2 | Decisions about which model features to expose to naïve and sophisticated users, and how, will reflect a review of online sources including bulletin boards, competitors’ technical documentation, as well as professional and research publications and customer input. |
| M3 | Each tool’s design document will identify the published sources of the mathematical techniques implemented in the tool. These sources must be among the publications available for the math review. |
| M4 | When a model commonly uses a variable transformation, or when applying the transformation before using the model is a best practice, a tool implementing the model will incorporate the transformation as an option that executes by default. |
| M5 | Every tool will implement all applicable, widely used model-quality metrics, whether these are general purpose, category specific, or model specific. |
| M6 | Every tool will include an embedded variable-selection method where it exists. Otherwise, if a wrapper variable-selection method is widely used, the tool will include it. |
| M7 | Every tool will include an embedded null-replacement method where it exists. If several embedded methods exist, the math review process should determine which to implement. |
| M8 | When a model can be implemented with any of several commonly used fitness functions, a tool implementing the model will support all of them. |
| M9 | When a model can be implemented with any of several commonly used fitting algorithms, the choice of fitting algorithm should reflect correctness, robustness, performance, and scalability concerns; and the literature review should determine how the alternatives compare along these four dimensions. |

## Transformation Standards

## Model-Quality Standards

## Variable-Selection Standards

## Null-Replacement Standards

## Fitness-Function Standards

## Fitting-Algorithm Standards

## Numerical Standards