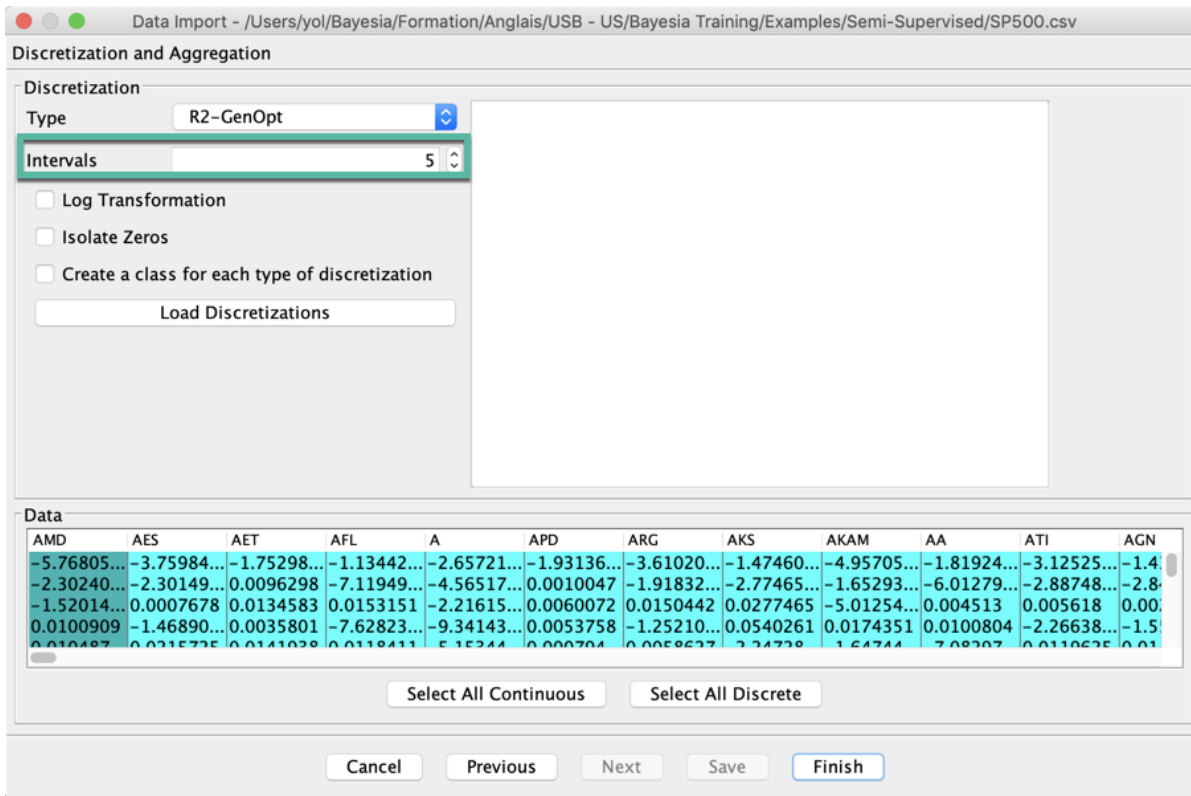


Discretization Wizard (8.0)

Context

Data Import | Discretization and Aggregation Wizard





New Feature: Intervals

As of version 8.0, the number of bins proposed by default in the discretization wizard is automatically calculated using the number of observations. This number is between 3 and 7.



We chose 3 instead of 2 as the minimum value to be able to capture non monotonic relationships. Indeed, with two bins, all relationships are represented linearly.

Except for 3 when there are few observations, the choice is rather conservative. This means that BayesiaLab should be able to discover two-parent structures when the number of observations exceeds a few hundred.

 The number of bins proposed in **Learning | Discretization** and **Node Editor | Curve | Generate a Discretization** tools does not use this new heuristic. It is either your previous choice, or the one defined in  **Window | Preferences | Discretization**

