Most conventional statistical analyses of brain-image data, and most statistical data-mining approaches are based on general linear models (GLMs), which in turn requires that data are multivariate Gaussian. In this talk I describe a complementary approach for categorical (i.e., discrete) variables. This data-mining framework, which is Bayesian, makes very few assumptions about the data, is not mass-univariate, does not suffer from the multiple-comparison problem, degrades gracefully with increasing noise and with missing data, requires no user input, and has clear semantics. This approach can detect multivariate nonlinear associations among variables, whether they are image variables—such as fMR activation or abnormal morphometry at a voxel—or clinical variables, such as experimental-group membership (e.g., treated/placebo). Our software takes as input a series of registered, segmented brain images, along with text-format clinical data, and searches for voxels whose states are associated with that of a clinical variable. The algorithm groups voxels that have similar associations with the clinical variable into a cluster, which is not necessarily spatially contiguous. The resulting clusters, or regions of interest, jointly classify new subjects with respect to a clinical variable, with relatively high accuracy, despite being based on categorical, rather than continuous variables. I will illustrate the application of this approach using data from different MR-based research projects.