One approach to using graphical models to define high-dimensional joint probability distributions is to model dependence between two variables by adding an edge between them in the graph. An alternative approach to this problem is to assume that the observed variables have some correlation because they arise from a hidden common cause. Models such as these with hidden variables are known as latent variable models, also called LVMs. These models are harder to fit than models with no latent variables, but they can have advantages. For example, there is the fact that LVMs have fewer parameters than models that directly represent correlation in the visible space. Additionally, the hidden variables in an LVM can serve as a bottleneck, which computes a compressed representation of the data, and this forms the basis of unsupervised learning. In general there are L latent variables and D visible variables, usually where $D >> L$. If we have $L > 1$, then there are many latent factors contributing to each observation, so there is a many-to-many mapping. If $L = 1$, there is only a single latent variable, and in this case z_i is usually discrete, so we have a one-to-many mapping. There can also be a many-to-one mapping representing different competing factors or causes for each observed variable. These types of models form the basis of probabilistic matrix factorization. Lastly, there can also be a one-to-one mapping, which can be represented as $z_i \rightarrow x_i$. By allowing z_i or x_i to be vector-valued, this representation can subsume all others. Depending on the form of the likelihood $p(x_i|z_i)$ and the prior $p(z_i)$ we can generate a variety of different models.