Venture Hive is a South Florida company incubator—it provides other businesses, usually startups, with the guidance and tools they need to grow and be successful.

To incubate efficiently, Venture Hive collects large amounts of data about regional companies; however, it needs a system that mines and sifts the data to retrieve useful information.

Frequent mining leads to the discovery of associations among items in large relational data sets. Venture Hive could benefit from finding these association patterns in its databases, especially in order to assess the replicability of practices that are correlated with desirable results.

Currently, Venture Hive does not have a data mining system in place; the company’s incorporation of the Regional Miner into its operational structure will mark its incursion into new territory. The idea is, on the one hand, for Venture Hive to identify salient correlations between specific business attributes in order to make informed recommendations to its clients; on the other hand, to sort businesses and entrepreneurs into groups, or “clusters”, based on shared features, with the purpose of tailoring its incubation approach according to the cluster to which a client belongs.

Having a human sort vast amounts of data in search of these associations and clusters can be very costly, even unfeasible; having a program perform this task is much more manageable and effective.

Venture Hive Regional Miner shall:
- Allow the user to select filtered data and view its details.
- Display which algorithms require data adjustments in order to be executed.
- Allow the user to perform data cleaning routines (e.g. replace missing values) to the selected data set for a specific algorithm.
- Allow the user to perform data transformation to provide better results (e.g. discretize and normalize) given a data set and specific algorithm.
- Allow the user to run multiple algorithms, regardless of type (i.e. association and clustering), in a single instruction.
- Allow the user to save a run configuration (i.e. data set, data adjustments and algorithms) for future execution.

The Regional Miner can help Venture Hive incubate its clients in two main ways.

First, it picks out correlations or associations between specific attributes; this allows Venture Hive to make recommendations to its clients based on what sorts of business practices are correlated with good results (e.g., number of founders being positively correlated with annual revenue).

Second, it assembles groups of businesses or entrepreneurs based on key features that they have in common, no matter how apparent the features may be to a human observer. Different incubation methods will suit different companies, yet companies in the same cluster will likely benefit from the same method. Venture Hive can “profile” a client according to the cluster to which it belongs in order to reduce the range of appropriate incubation methods.