Bemus Landscape - UCSD Capstone Project Proposal

**Background:**

The company’s current methodology for estimating new contract pricing is a combination of multipliers or production rates applied to numerical parameters obtained via satellite measurements. These production rates are based on historical field testing and represent averages across the company’s branches. This estimating methodology makes no adjustments based on other parameters, historical data, or parameter interactions and has not been updated for many years.

In addition to the above, non-contract work represents approximately half of the company’s total revenue and more than half of its gross profit. There have been no past attempts to predict non-contract (ancillary) revenue that may come with a future contract client. Conventional wisdom suggests a benchmark penetration rate of ancillary services relative to contract dollars, but that figure has not been re-evaluated in some time and is not more sophisticated than assuming a percentage of the contract. In recent years and for reasons yet unexplored, that conventional wisdom is not always holding true.

The goal of this capstone project is to accomplish the following:

* Create a predictive model for total hours (i.e. price agnostic of labor cost) needed for a future contract client using historical data.
* Create predictive models for ancillary revenue across three separate non-contract divisions assuming Bemus has secured the base contract.
* Combine the above predictive models into a new ensemble model that predicts a client’s potential two-year value.

If successful, the combination of models will become the new pricing model for each future client. Knowing the total two-year value of a client will allow the business to price base contracts more competitively to increase close percentage and maximum revenue across divisions that typically don’t consult with each other.

**Key Contacts:**

Robin Reese - Director of Operations (Bemus)

[robin.reese@bemus.com](mailto:robin.reese@bemus.com)

Corin Bemus - CEO (Bemus)

[corin.bemus@bemus.com](mailto:corin.bemus@bemus.com)

Colin Bemus - Chief Operating Officer (Sperber)

[cbemus@sperbercompanies.com](mailto:cbemus@sperbercompanies.com)

**Data:**

An overview of the data and its size is below. There are relatively few tables to work with though the tables are very wide relative to their height. Not more than 20% of the available columns are likely relevant to this project. Significant cleaning and aggregation of the data will be required to put the raw data into a usable format. One complicating factor is this data lives in two different instances of our ERP platform (pre- and post-acquisition of the company). In addition to this users will find that naming conventions and divisions have changed over time to reflect accounting needs. Your SME will guide you through how to manage this.

General dimensions of the data:

* Properties ~1100 rows x 117 columns
* Issues ~28,000 rows x 25 columns
* Opportunities ~246,000 rows x 185 columns
* Work Tickets ~450,000 rows x 124 columns

