**CAPSTONE PROJECT IDEAS**

**One:**

**House Prices**

**Problem:** This project would aim at predicting the final price of a residential home in Ames, Iowa based on different essential parameters that influences price negotiations other than the number of bedrooms or a white-picket fence.

**Data:** The dataset contains 2930 observations and 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, with information from the Ames Assessor’s Office used in computing assessed values for individual residential properties sold in Ames, IA from 2006 to 2010. Below is the link to the dataset:

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data>

**Two:**

**Corporación Favorita Grocery Sales Forecasting**

**Problem:** This project would aim at building a model that forecasts product sales more accurately in order to be able to please customers by having just enough of the right products at the right time. This model will help solve the problem of  [Corporación Favorita](http://www.corporacionfavorita.com/), a large Ecuadorian based grocery retailer inaccurate purchasing and sales forecasting.

**Data:** The dataset contains both the training and the test set with the training data containing dates, store and item information, whether that item was being promoted, as well as the unit sales while the test data also contains the dates, store and item information that are to be predicted, along with the on-promotion information. Below is the link to the dataset:

<https://www.kaggle.com/c/favorita-grocery-sales-forecasting/data>

**Three:**

**Bank Marketing**

**Problem:** The goal of this classification is to predict if a client will subscribe a term deposit of a Portuguese banking institution. This project would aim at predicting if a client will say yes or no to a bank term deposit subscription.

**Data:** This is a multivariate dataset with 45211 instances with 17 attributes. The data is related with direct marketing campaigns of a Portuguese banking institution and these campaigns were based on phone calls. Ideally, it was required for a client to be contacted more than one time, in order to access if the product “bank term deposit” would be a “yes” or not a “no” subscription. Below is the link to the dataset:

<http://archive.ics.uci.edu/ml/datasets/Bank+Marketing>