**Springboard-Data Science Career Program**

**Capstone 2 – House price prediction**

**Problem**: It’s challenging to a home buyer to describe their dream house, and difficult to know the factors driving the price of the house i.e whether it is number of bedrooms or neighborhood etc.

Goal isto predict the sales price for each house using 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa.

**Data:**

Data Source is House price dataset from Kaggle competition

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

1. Train.csv, contains several features including target variable SalePrice of house. Use this dataset to train the machine learning regression model.
2. Test.csv, contains all 80 features in train.csv file except the target variable. So, predict the SalePrice of the houses using model trained on train.csv dataset.

**Approach:**

Since target variable is continuous, supervised machine learning approach i.e Linear regression will be explored to find what features drive predicting the price of the house.

**Deliverables:**

The project deliverables will include Jupyter notebooks containing machine learning methods used and code to support analysis, a final report detailing methods, conclusions and recommendations.