Problem Statement

The price of a property depends on a lot of factors, ranging from the type of neighborhood the property is located in to its structural features and its aesthetics. In this assessment, you are required to use features that describe various aspects of a property and predict the sales price for the property.

The data comprises 79 explanatory features describing a subset of properties for sale in the city of Ames, Iowa. The type of features are varied, ranging from numbers like lot area to categorical types like type of property for sale. The target variable is *SalePrice*, i.e. the price at which the property might be sold, in US dollars. For each *Id* in the test set, you must predict the value of the *SalePrice* variable.

It is not mandatory to use all the 79 features to make the prediction. In fact, you can even creatively combine multiple features to come up with derived features that could be relevant.

The data files are arranged in the following way:

- **data_description.txt**: This file contains explanations for each of the 79 features that describe a property.
    - A briefer version of this description is also provided in this file in the last section.
- **train.csv:** The training dataset which contains values for all the 79 features as well as the target variable *SalePrice.*
- **test.csv:** The test dataset which contains values for the 79 features but not the target variable *SalePrice.* The main task of this assignment is to predict the sale price for the properties in the test dataset.

What are we interested in seeing?

- What features you choose and your reasoning for choosing them.
- What algorithm you choose to perform the regression.
- What loss function you choose for this regression task to train the algorithm.

What to submit and how?

You can submit your work by uploading **the code and results** to a public github repository and sharing the link to it. We expect you to provide the predictions result in a csv file in the following format:

Id	SalePrice
1461	170000.5
1462	190000
1463	230000.5

We also expect you to write the code in python and use different libraries for data manipulation, visualization and modeling.
It doesn't matter where you write the code although we recommend you to code it in a Jupyter notebook and perform some exploratory data analysis before you attempt to solve the regression task.

Data Description

A detailed description for each of the fields will be provided as part of the email along with this document.

- SalePrice - the property's sale price in dollars. This is the target variable that you're trying to predict.
- MSSubClass: The building class
- MSZoning: The general zoning classification
- LotFrontage: Linear feet of street connected to property
- LotArea: Lot size in square feet
- Street: Type of road access
- Alley: Type of alley access
- LotShape: General shape of property
- LandContour: Flatness of the property
- Utilities: Type of utilities available
- LotConfig: Lot configuration
- LandSlope: Slope of property
- Neighborhood: Physical locations within Ames city limits
- Condition1: Proximity to main road or railroad
- Condition2: Proximity to main road or railroad (if a second is present)
- BldgType: Type of dwelling
- HouseStyle: Style of dwelling
- OverallQual: Overall material and finish quality
- OverallCond: Overall condition rating
- YearBuilt: Original construction date
- YearRemodAdd: Remodel date
- RoofStyle: Type of roof
- RoofMatl: Roof material
- Exterior1st: Exterior covering on house

- Exterior2nd: Exterior covering on house (if more than one material)
- MasVnrType: Masonry veneer type
- MasVnrArea: Masonry veneer area in square feet
- ExterQual: Exterior material quality
- ExterCond: Present condition of the material on the exterior
- Foundation: Type of foundation
- BsmtQual: Height of the basement
- BsmtCond: General condition of the basement
- BsmtExposure: Walkout or garden level basement walls
- BsmtFinType1: Quality of basement finished area
- BsmtFinSF1: Type 1 finished square feet
- BsmtFinType2: Quality of second finished area (if present)
- BsmtFinSF2: Type 2 finished square feet
- BsmtUnfSF: Unfinished square feet of basement area
- TotalBsmtSF: Total square feet of basement area
- Heating: Type of heating
- HeatingQC: Heating quality and condition
- CentralAir: Central air conditioning
- Electrical: Electrical system
- 1stFlrSF: First Floor square feet
- 2ndFlrSF: Second floor square feet
- LowQualFinSF: Low quality finished square feet (all floors)
- GrLivArea: Above grade (ground) living area square feet
- BsmtFullBath: Basement full bathrooms
- BsmtHalfBath: Basement half bathrooms
- FullBath: Full bathrooms above grade
- HalfBath: Half baths above grade
- Bedroom: Number of bedrooms above basement level
- Kitchen: Number of kitchens
- KitchenQual: Kitchen quality
- TotRmsAbvGrd: Total rooms above grade (does not include bathrooms)
- Functional: Home functionality rating
- Fireplaces: Number of fireplaces
- FireplaceQu: Fireplace quality
- GarageType: Garage location
- GarageYrBlt: Year garage was built
- GarageFinish: Interior finish of the garage
- GarageCars: Size of garage in car capacity
- GarageArea: Size of garage in square feet
- GarageQual: Garage quality
- GarageCond: Garage condition
- PavedDrive: Paved driveway
- WoodDeckSF: Wood deck area in square feet
- OpenPorchSF: Open porch area in square feet
- EnclosedPorch: Enclosed porch area in square feet
- 3SsnPorch: Three season porch area in square feet

- ScreenPorch: Screen porch area in square feet
- PoolArea: Pool area in square feet
- PoolQC: Pool quality
- Fence: Fence quality
- MiscFeature: Miscellaneous feature not covered in other categories
- MiscVal: $Value of miscellaneous feature
- MoSold: Month Sold
- YrSold: Year Sold
- SaleType: Type of sale
- SaleCondition: Condition of sale