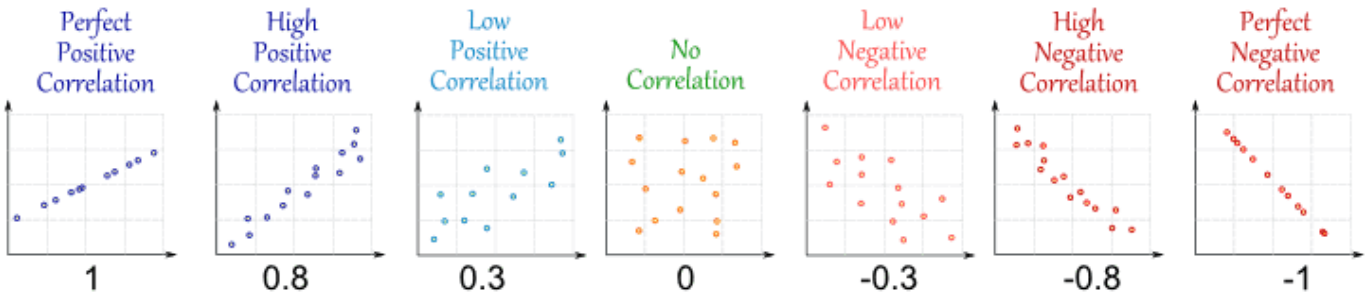
# Weka

## Working with Weka

### Regression for Prepredictiion

```
1.  Explorer
2.  Preprocess
3.  Open File (csv or arff): HHRealstate.csv (to be downloaded from
github)
    note: * If you open .csv you can save as .arff
          * To view .csv files, change file type to .csv
4.  Choos filter -> filters -> unsupervised -> attributes -> Normalize ->
Apply
5.  Classify -> Choose
6.  Classifier -> functions -> LinearRegression (explore the other
functions)
7.  Select Test Option
8.  Start
```

**Output: Correlation and Errors**



(source: http://www.mathsisfun.com/data/correlation.html)

Mean absolute error is:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{\theta}_i - \theta_i|$$

Root mean square error is:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta_i\right)^2}$$

Relative absolute error:

$$\mathrm{RAE} = \frac{\sum_{i=1}^{N}|\hat{\theta}_i - \theta_i|}{\sum_{i=1}^{N}|\overline{\theta} - \theta_i|}$$

where $\overline{\theta}$ is a mean value of $\theta$.

Root relative squared error:

$$\mathrm{RRSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta_i\right)^2}{\sum_{i=1}^{N}\left(\overline{\theta} - \theta_i\right)^2}}$$

**Save/Load Trained model**

```
1. Right Click on the function used for training and Select Save Model.
2. Exit Weka
3. Create copy of HHrealstate.csv file and call it realstat.csv, this file
represent a file that need to be predicted.
4.  In the file, remove the last colomn(unit price and replace it with ?)
```

```
5. save the file.
6. Open Weka, and load HHRealstate.csv OR ANY OTHER FILE, we do this to
activate the ability to Right Click on the used function.
7. Classify
8. Right Click the used function
9. Load Model -> Select your saved model
10. from Test Options , Select : Supplied Test Set
11. select the file you will use for prediction (realstate.csv)
12. Make sure that the Class field is selected.
13. Select: Close
14. Select: More Options -> Output prediction -> choose : Plane Text
15. Right Click on the function.LinearRegression -> "Re-evalute model on
current test set"
```

## Classification

```
*   We use weather data and iris data.
```