Multiple Linear Regression in Python - Automatic Backward Elimination

Section 5, Lecture 48

Hi guys,

if you are also interested in some automatic implementations of Backward Elimination in Python, please find two of them below:

Backward Elimination with p-values only:

```
import statsmodels.formula.api as sm
def backwardElimination(x, sl):
    numVars = len(x[0])
    for i in range(0, numVars):
        regressor_OLS = sm.OLS(y, x).fit()
        maxVar = max(regressor_OLS.pvalues).astype(float)
        if maxVar > sl:
            for j in range(0, numVars - i):
                if (regressor_OLS.pvalues[j].astype(float) == maxVar):
                    x = np.delete(x, j, 1)
    regressor_OLS.summary()
    return x

SL = 0.05
X_opt = X[:, [0, 1, 2, 3, 4, 5]]
X_Modeled = backwardElimination(X_opt, SL)
```

Backward Elimination with p-values and Adjusted R Squared: