

In Python the feature selection method is called the Recursive Feature Elimination

```
In [ ]: import pandas as pd
        from sklearn.linear_model import LinearRegression
        from sklearn.feature_selection import RFE, RFECV #This is the Recursive Feature Elimination
```

```
In [ ]: data=pd.read_csv("D:\\Workshops\\W 15 - Data Science Masterclass\\Data\\Boston.CSV")
        data.head()
```

```
In [ ]: df_x=data.iloc[:,12]
        df_y=data.iloc[:,12]
```

Creating a Recursive Feature Elimination object

```
In [ ]: rfe = RFE(estimator=LinearRegression(), n_features_to_select=5)
```

Training the RFE object with independent and response data ¶

```
In [ ]: elm_fit=rfe.fit(df_x, df_y)
```

```
In [ ]: elm_fit
```

Number of features selected

```
In [ ]: elm_fit.n_features_
```

Which columns have been selected

```
In [ ]: elm_fit.support_
```

```
In [ ]: df_x.columns
```

```
In [ ]: selected=df_x.columns[elm_fit.support_]
        selected #This gives the selected columns in the independent data
```

Selection ranking is given as

```
In [ ]: elm_fit.ranking_
```

Selected data frame

```
In [ ]: df_x[selected]
```

Now these selected features can be used for the model

```
In [ ]: x=df_x[selected].values
        y=df_y.values
```

```
In [ ]: x
```

```
In [ ]: y
```

This x matrix can be obtained using following way as well

```
In [ ]: x_sel= rfe.fit_transform(df_x,df_y)
x_sel
```

RFE Cross Validation (If the number of features is not decided)

```
In [ ]: rfe=RFECV(estimator=LinearRegression(),min_features_to_select=1,cv=10)
```

```
In [ ]: elm_fit=rfe.fit(df_x,df_y)
```

```
In [ ]: elm_fit.support_
```

```
In [ ]: df_x.columns[elm_fit.support_]
```