```
% Load dataset. Note: this example uses MATLAB's table data structure to
% store the content of the .csv file (table is available since
% MATLAB R2013b)
data = readtable('../data/regression.csv');

% Convert the Sex column from strings to MATLAB's categorical data type.
data.Sex = categorical(data.Sex);

% Plot the data and colour the points by gender. The third argument (empty
% square brackets, i.e. []) tells MATLAB to use its default.
figure;
scatter(data.Age, data.OI, [], data.Sex, 'filled');
xlabel('Age');
ylabel('OI');

% Clean up the data by removing rows with negative age
data = data(data.Age > 0, :);

% Log transform OI for data fitting
data.logOI = log(data.OI);

% Fit a multiple regression model according to the formula:
%   log(OI) ~ Age + Sex
% Note that by not terminating the line with ; we get the ANOVA summary
model = fitlm(data, 'logOI ~ Age + Sex')

% Standard residual plot. Other residuals are available: Raw, Pearson,
% Standardized, Studentized
figure;
qqplot(model.Residuals.Raw);

% Fitted values vs residual plot
figure;
scatter(model.Fitted, model.Residuals.Raw, [], data.Sex, 'filled');
xlabel('Fitted values');
ylabel('Residual');

% Age vs residual plot coloured by gender
figure;
scatter(data.Age, model.Residuals.Raw, [], data.Sex, 'filled');
xlabel('Age');
ylabel('Residual');

% Boxplot of residuals separated by gender
figure;
boxplot(model.Residuals.Raw, data.Sex)
xlabel('Sex');
ylabel('Residual');
```

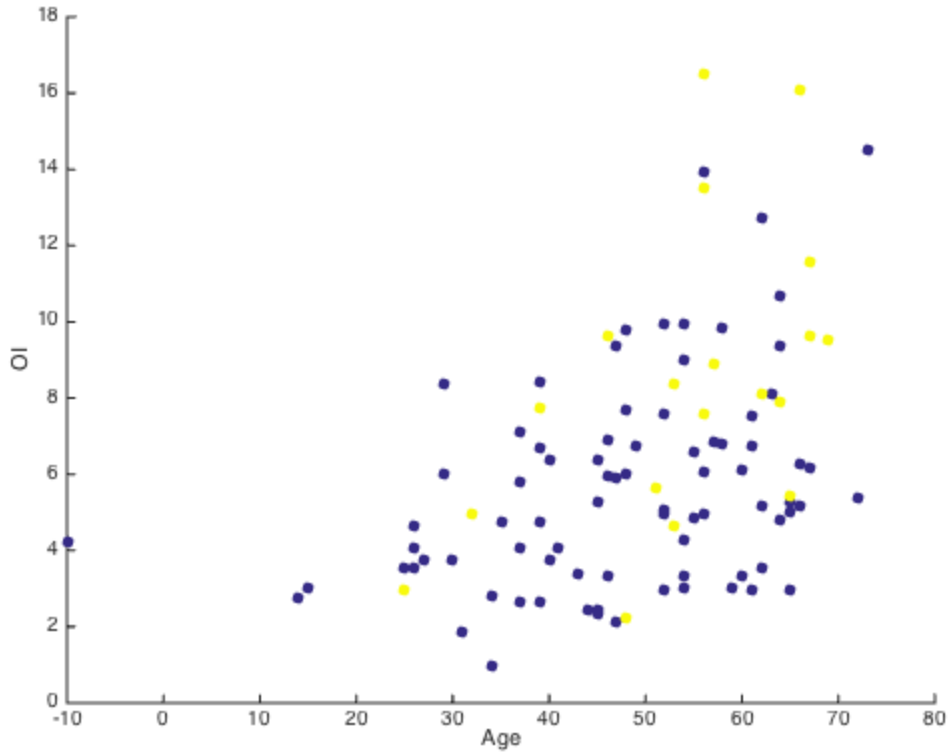
```
model =
```

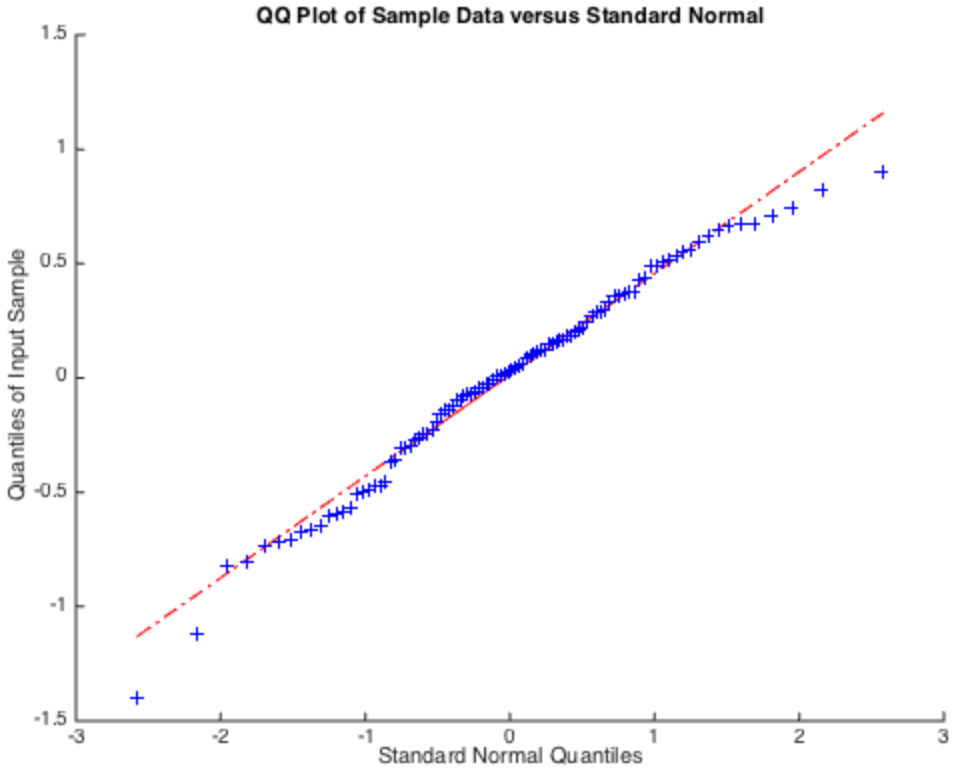
Linear regression model:
logOI ~ 1 + Age + Sex

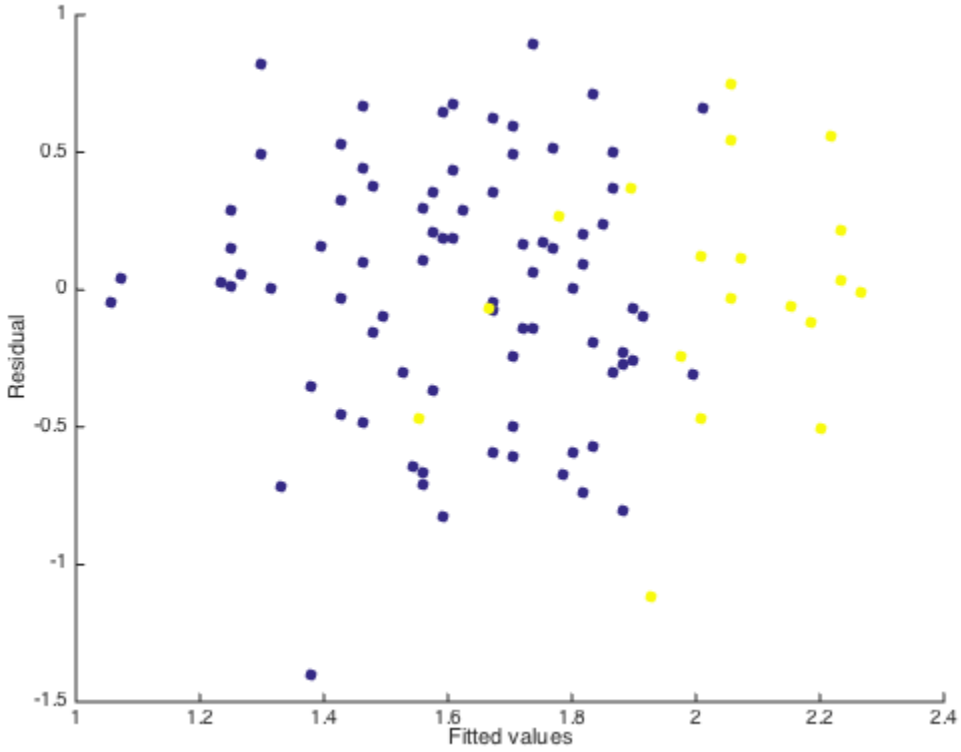
Estimated Coefficients:

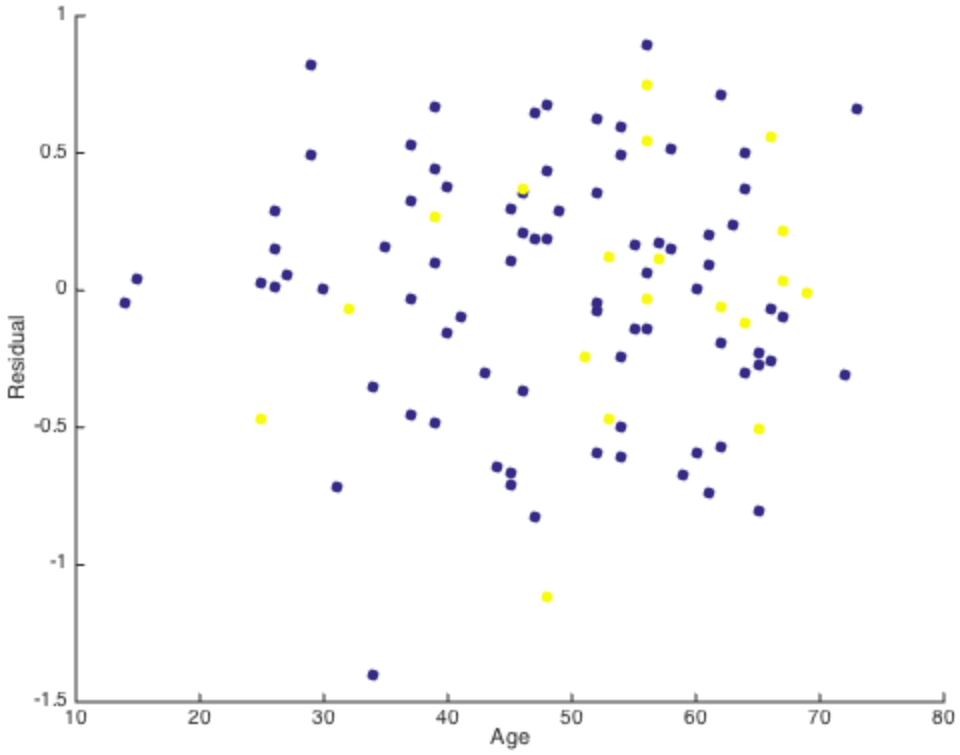
	Estimate	SE	tStat	pValue
(Intercept)	0.8292	0.17771	4.6661	9.8341e-06
Age	0.016208	0.0035215	4.6027	1.2637e-05
Sex_Male	0.3189	0.11568	2.7567	0.0069774

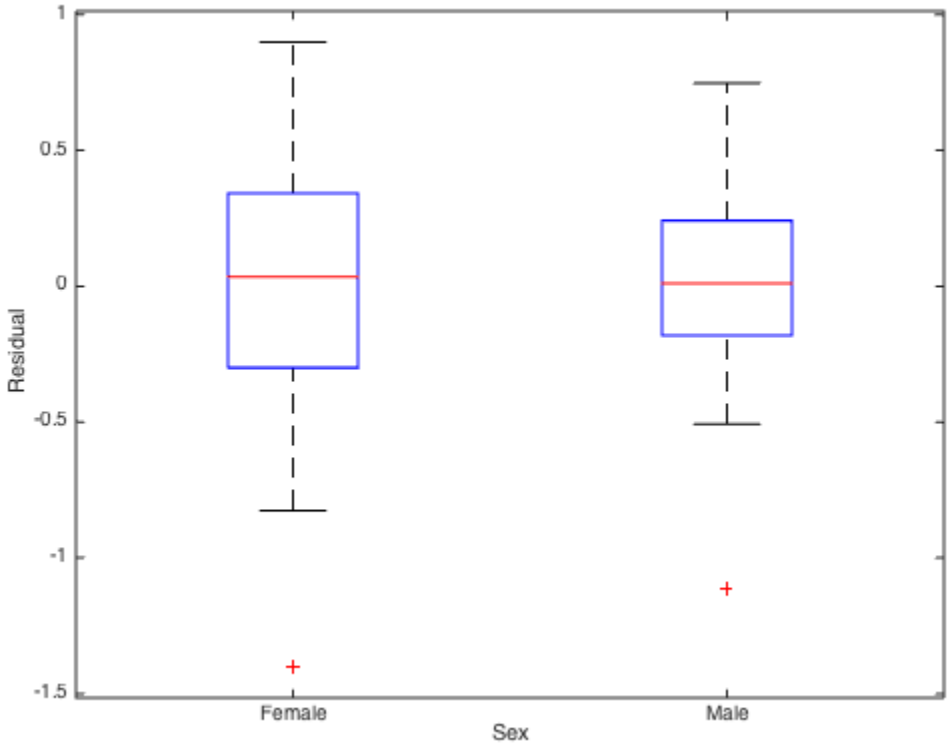
Number of observations: 100, Error degrees of freedom: 97
Root Mean Squared Error: 0.455
R-squared: 0.262, Adjusted R-Squared 0.247
F-statistic vs. constant model: 17.2, p-value = 3.96e-07




QQ Plot of Sample Data versus Standard Normal
Quantiles of Input Sample
Standard Normal Quantiles
1.5
1
0.5
0
-0.5
-1
-1.5
-3
-2
-1
0
1
2
3


Residual
Fitted values
1
0.5
0
-0.5
-1
-1.5
1
1.2
1.4
1.6
1.8
2
2.2
2.4


Residual
Age
1
0.5
0
-0.5
-1
-1.5
10
20
30
40
50
60
70
80


Residual
Sex
Female
Male
1
0.5
0
-0.5
-1
-1.5

Published with MATLAB® R2014b