MATLAB Tutorials

16.62x Experimental Projects

Violeta Ivanova, Ph.D.
Educational Technology Consultant
MIT Academic Computing

violeta@mit.edu

IST Information Services and Technology

MIT

This Tutorial

- Class materials

  web.mit.edu/acmath/matlab/16.62x
- Topics
  - MATLAB Basics Review
  - Data Analysis
  - Statistics Toolbox
  - Curve Fitting Toolbox

Other References

- Mathematical Tools at MIT

  web.mit.edu/ist/topics/math
- Course16 Tutorials
  - Unified MATLAB: web.mit.edu/acmath/matlab/unified
  - 16.06 &16.07 MATLAB & Simulink: web.mit.edu/acmath/matlab/course16

MATLAB Basics Review

Toolboxes & Help

Matrices & Vectors

Built-In Functions

Graphics

Help in MATLAB

- Command line help

```
>> help <command>
          e.g. help regress
>> lookfor <keyword>
          e.g. lookfor regression
```

- Help Browser
  - Help->Help MATLAB

MATLAB Help Browser

- MATLAB
  - + Mathematics
  - + Data Analysis
  - + Programming
  - + Graphics
- Curve Fitting Toolbox
- Statistics Toolbox
  - + Linear Models
  - + Hypothesis Tests
  - + Statistical Plots

Vectors

- Row vector

```
>> R1 = [1 6 3 8 5]
>> R2 = [1 : 5]
>> R3 = [-pi : pi/3 : pi]
```

- Column vector

```
>> C1 = [1; 2; 3; 4; 5]
>> C2 = R2'
```

Matrices

- Creating a matrix

```
>> A = [1 2.5 5 0; 1 1.3 pi 4]
>> A = [R1; R2]
```

- Accessing elements

```
>> A(1,1)
>> A(1:2, 2:4)
>> A(:,2)
```

Matrix Operations

- Operators + and –

>> X = [x_1 x_2]; Y = [y_1 y_2]; A = X+Y

A =

x_1+y_1 x_2+y_2

- Operators *, /, and ^

>> Ainv = A^-1 Matrix math is default!

- Operators .*, ./, and .^

>> Z = [z_1 z_2]; B = [Z.^2 Z.^0]

B =

z_1^2 z_2^2 1 1

Built-In Functions

- Matrices & vectors

```
>> [n, m]= size(A)
>> n = length(X)
>> M1 = ones(n, m)
>> M0 = zeros(n, m)
>> En = eye(n)
>> N1 = diag(En)
```

- And many others …

```
>> y = exp(sin(x)+cos(t))
```

Graphics

- 2D linear plots: `plot`

```
>> plot (t, z, ‘r-’)
```

Colors: `b, r, g, y, m, c, k, w`

Markers: `o, *, ., +, x, d`

Line styles: `-, --, -., :`

- Annotating graphs

```
>> legend (‘z = f(t)’)
>> title (‘Position vs. Time’)
>> xlabel (‘Time’)
>> ylabel (‘Position’)
```

Multiple Plots

- Multiple datasets on a plot

```
>> p1 = plot(xcurve, ycurve)
>> hold on
>> p2 = plot(Xpoints, Ypoints, ‘ro’)
>> hold off
```

- Subplots on a figure

```
>> s1 = subplot(1, 2, 1)
>> p1 = plot(time, velocity)
>> s2 = subplot(1, 2, 2)
>> p2 = plot(time, acceleration)
```

MATLAB Data Analysis

Preparing Data

Basic Fitting

Correlation

Data Input / Output

- Import Wizard for data import

  `File->Import Data …`

- File input with `load`

  ```
  B = load('datain.txt')
  ```

- File output with `save`

  ```
  save('dataout', 'A', '-ascii')
  ```

Missing Data

- Removing missing data
  - Removing NaN elements from vectors

```
>> x = x(~isnan(x))
```

  - Removing rows with NaN from matrices

```
>> X(any(isnan(X),2),:) = []
```

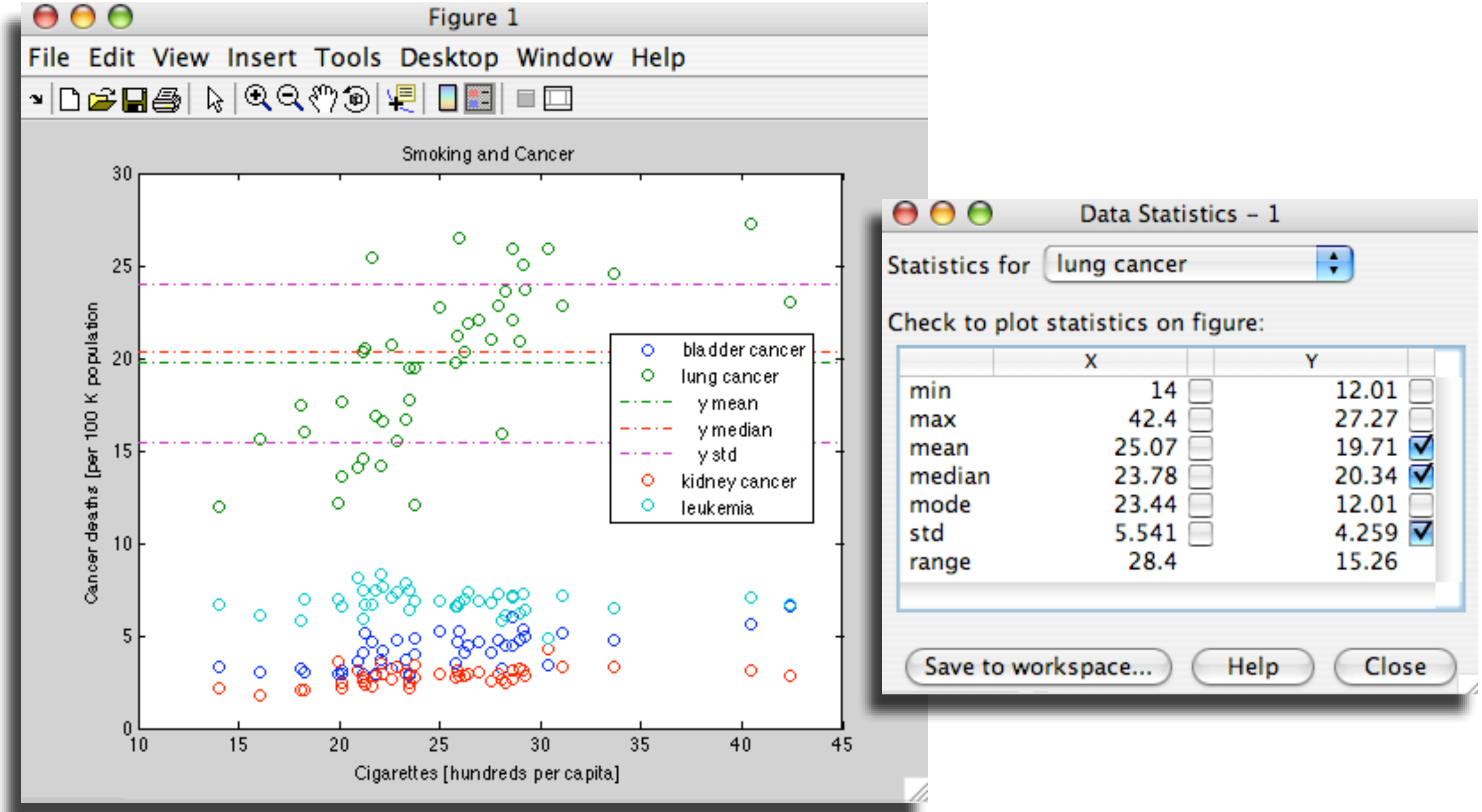
- Interpolating missing data

```
YI = interp1(X, Y, XI, ‘method’)
```

Methods: ‘spline’, ‘nearest’, ‘linear’, …

Data Statistics

- Figure window: `Tools->Data Statistics`



Statistics for: lung cancer

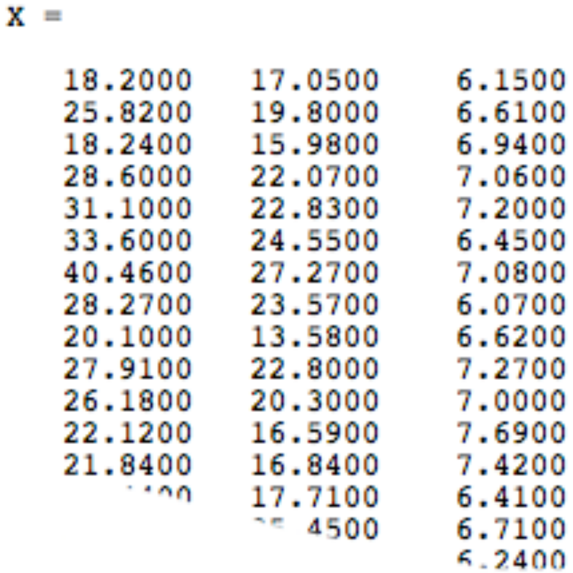
Check to plot statistics on figure:

	X		Y	
min	14	□	12.01	□
max	42.4	□	27.27	□
mean	25.07	□	19.71	☑
median	23.78	□	20.34	☑
mode	23.44	□	12.01	□
std	5.541	□	4.259	☑
range	28.4		15.26	

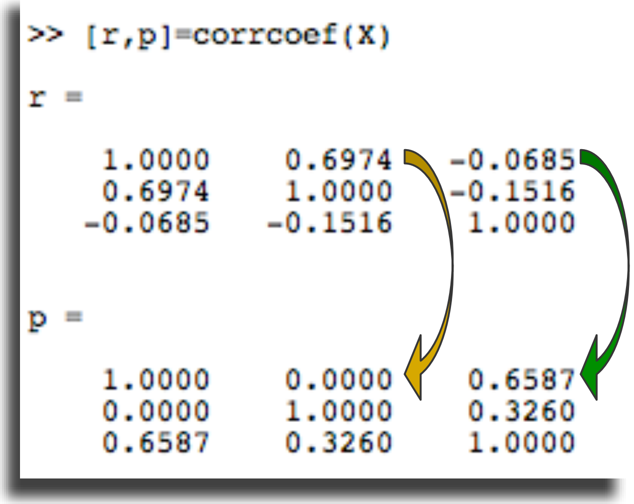
Correlation

- Correlation coefficient & confidence interval

```
>> [R, P, Rlo, Rup, alpha] = corrcoef(X);
>> [i, j] = find(P < 0.05);
```



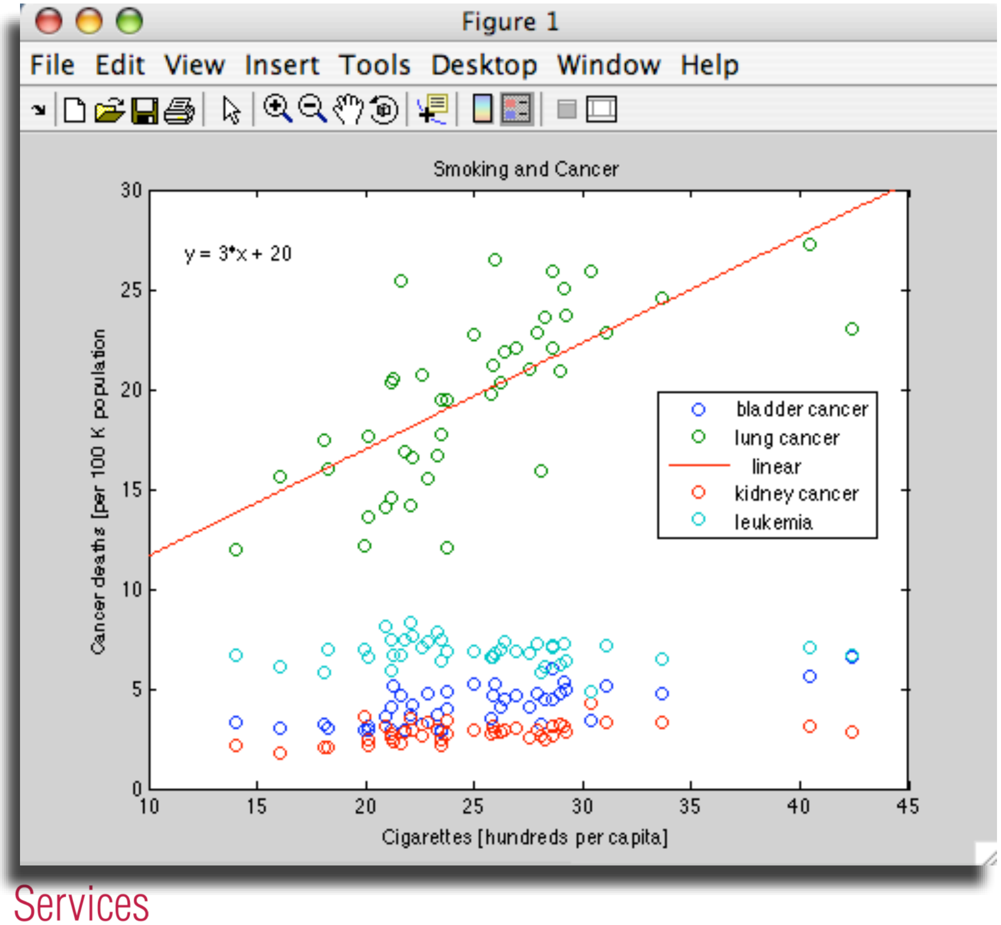
```
X =

   18.2000   17.0500    6.1500
   25.8200   19.8000    6.6100
   18.2400   15.9800    6.9400
   28.6000   22.0700    7.0600
   31.1000   22.8300    7.2000
   33.6000   24.5500    6.4500
   40.4600   27.2700    7.0800
   28.2700   23.5700    6.0700
   20.1000   13.5800    6.6200
   27.9100   22.8000    7.2700
   26.1800   20.3000    7.0000
   22.1200   16.5900    7.6900
   21.8400   16.8400    7.4200
             17.7100    6.4100
                4500    6.7100
                        6.2400
```



```
>> [r,p]=corrcoef(X)

r =

    1.0000    0.6974   -0.0685
    0.6974    1.0000   -0.1516
   -0.0685   -0.1516    1.0000


p =

    1.0000    0.0000    0.6587
    0.0000    1.0000    0.3260
    0.6587    0.3260    1.0000
```

Basic Fitting

- Figure window: `Tools->Basic Fitting …`



Polynomials

- Evaluating polynomials

$$y = p_1 x^n + p_2 x^{n-1} \ldots + p_n x + p_{n+1}$$

```
>> p = [p1 p2 … ]

>> t = [-3 : 0.1 : 3]

>> z = polyval(p, t)
```

- Fitting a polynomial

```
>> X = [x1 x2 … xn]; Y = [y1 y2 … yn]

>> Pm = polyfit(X, Y, m)
```

Statistics Toolbox

Probability Distributions

Descriptive Statistics

Linear & Nonlinear Models

Hypothesis Tests

Statistical Plots

Descriptive Statistics

- Central tendency

```
>> m   = mean(X)
>> gm  = geomean(X)
>> med = median(X)
>> mod = mode(X)
```

- Dispersion

```
>> s = std(X)
>> v = var(X)
```

Probability Distributions

- Probability density functions

```
>> Y = exppdf(X, mu)
>> Y = normpdf(X, mu, sigma)
```

- Cumulative density functions

```
>> Y = expcdf(X, mu)
>> Y = normcdf(X, mu, sigma)
```

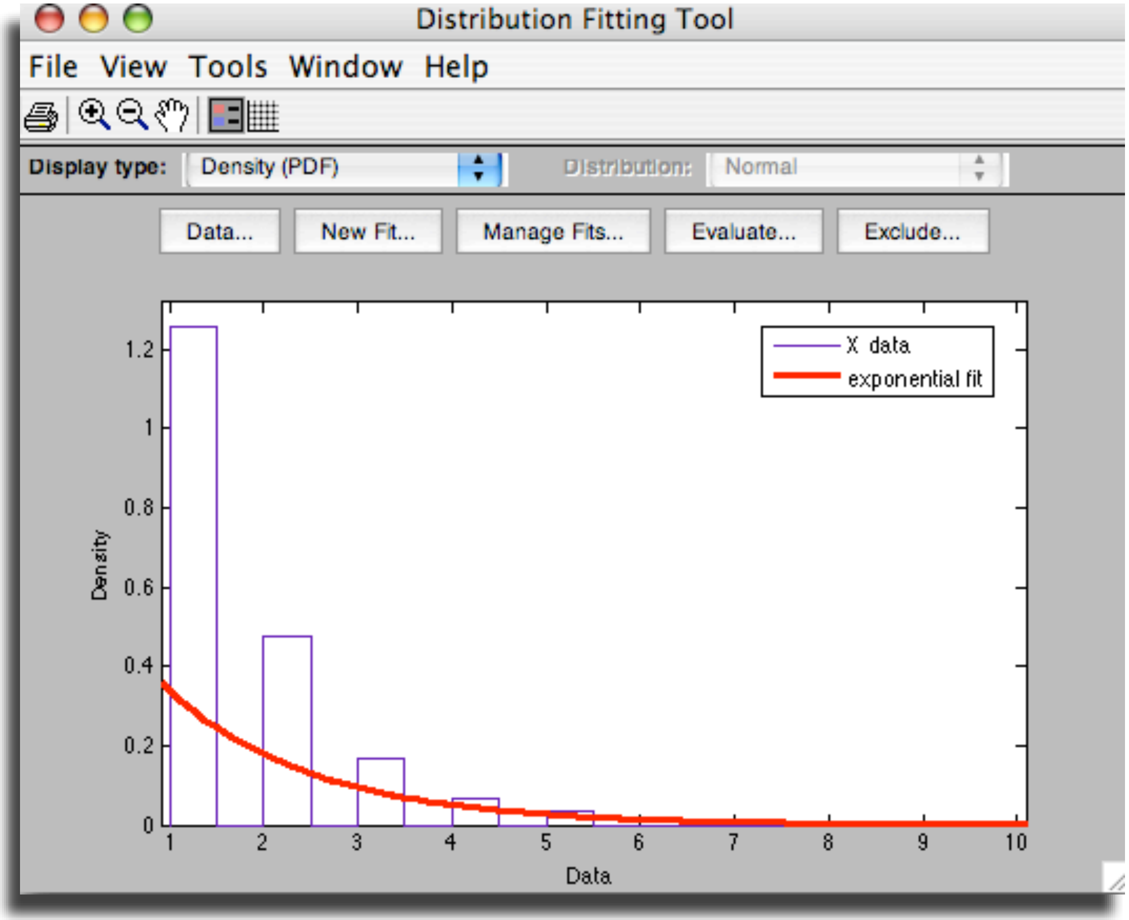
- Parameter estimation

```
>> m = expfit(data)
>> [m, s] = normfit(data)
```

Distribution Fitting Tool

- Start from command line window

  `>> dfittool`



Linear Models

- Definition:

$$y = X\beta + \varepsilon$$

y: *n x 1* vector of observations

X: *n x p* matrix of predictors

β: *p x 1* vector of parameters

ε: *n x 1* vector of random disturbances

Linear Regression

- Multiple linear regression

```
>> [B, Bint, R, Rint, stats] = regress(y, X)
```

  `B`: vector of regression coefficients

  `Bint`: matrix of 95% confidence intervals for B

  `R`: vector of residuals

  `Rint`: intervals for diagnosing outliners

  `stats`: vector containing R^2 statistic etc.

- Residuals plot

```
>> rcoplot(R, Rint)
```

Polynomial Fitting Tool

```
>> polytool(X, Y)
```

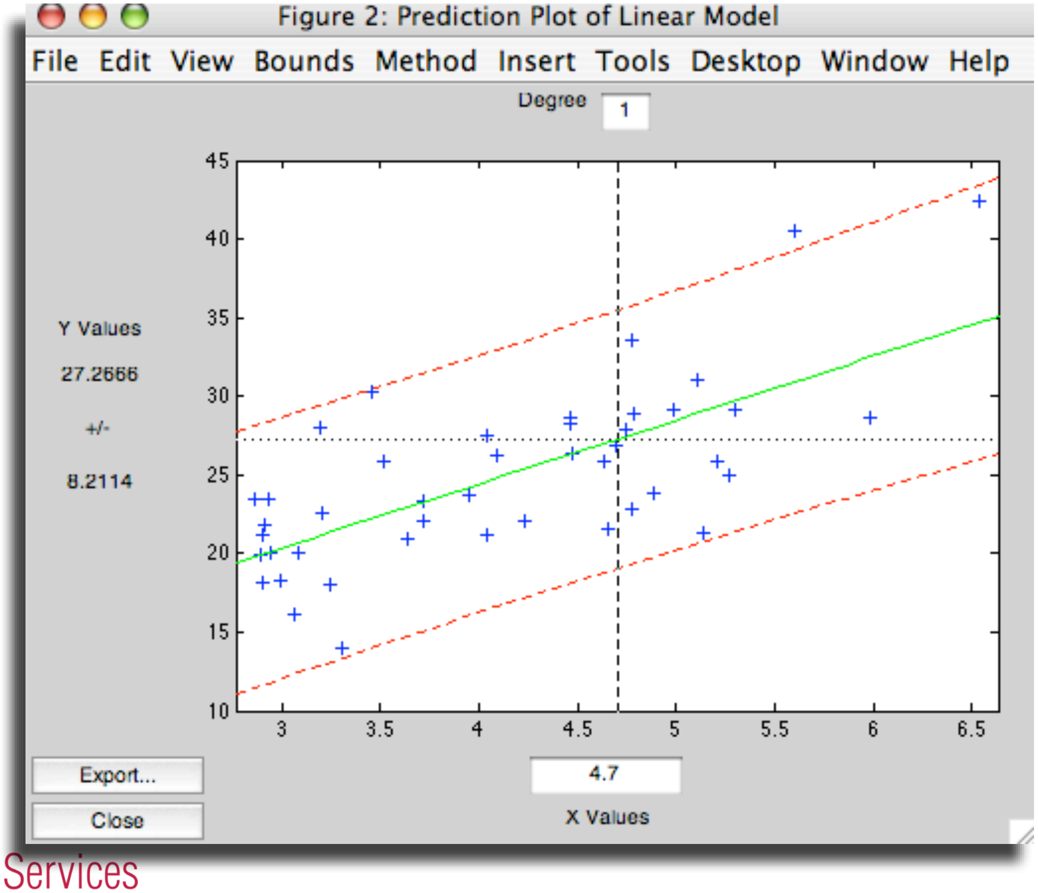
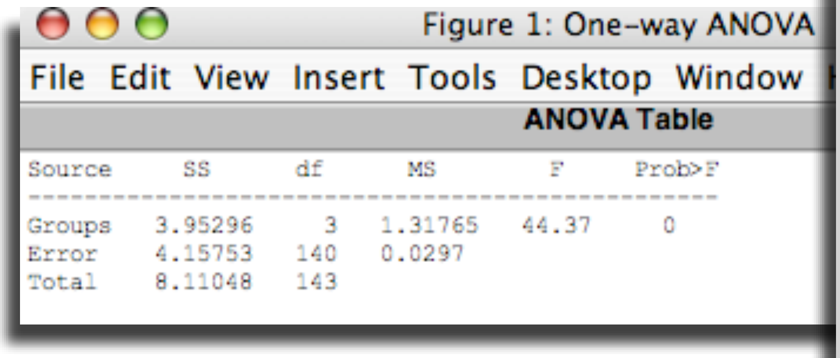


Figure 2: Prediction Plot of Linear Model
File Edit View Bounds Method Insert Tools Desktop Window Help
Degree
1
45
40
35
30
25
20
15
10
3
3.5
4
4.5
5
5.5
6
6.5
Y Values
27.2666
+/-
8.2114
Export...
Close
4.7
X Values

Analysis of Variance (ANOVA)

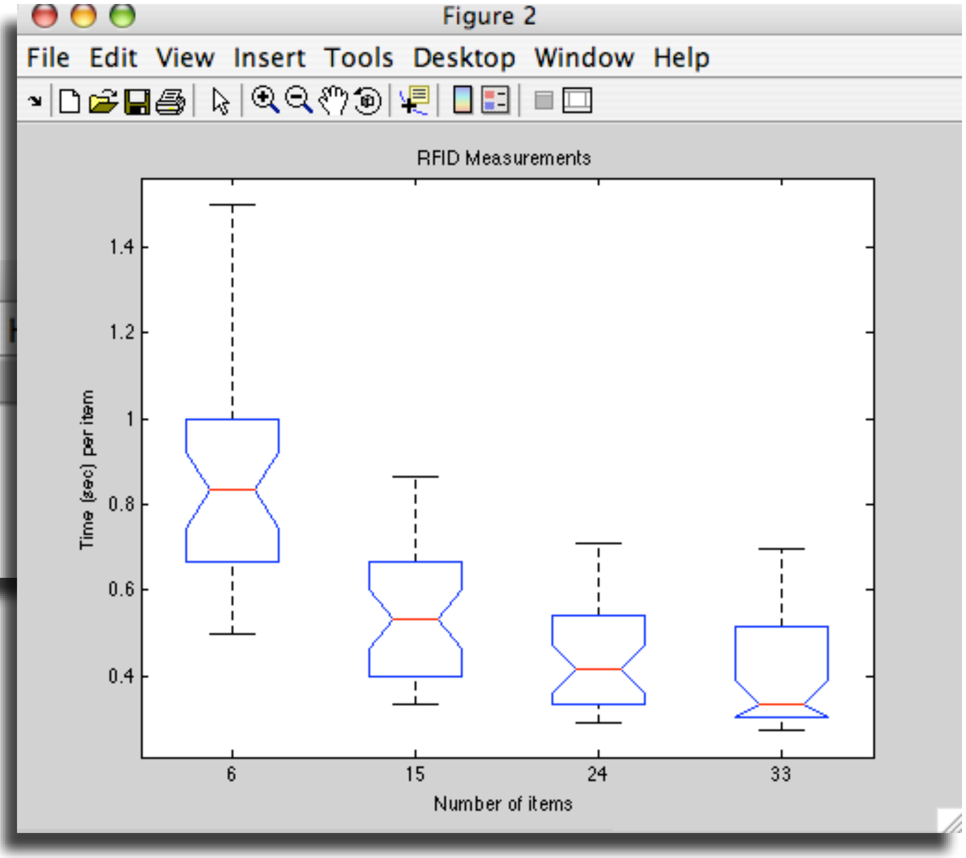
- One-way ANOVA

```
>> anova1(X,group)
```



ANOVA Table

```
Source      SS       df      MS        F     Prob>F
---------------------------------------------------
Groups    3.95296     3   1.31765   44.37     0
Error     4.15753   140   0.0297
Total     8.11048   143
```



Multiple Comparisons

```
>> [p, tbl, stats] = anova1(X,group)
>> [c, m] = multcompare(stats)
```

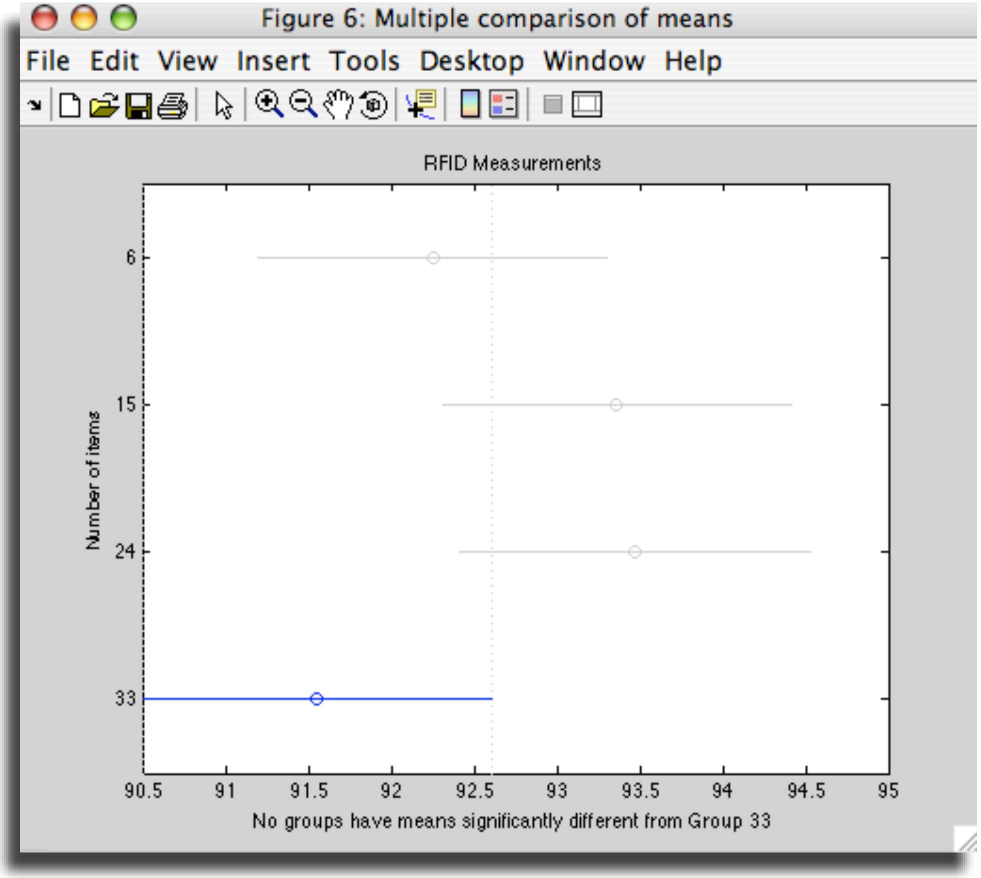


Figure 6: Multiple comparison of means
File Edit View Insert Tools Desktop Window Help
RFID Measurements
Number of items
6
15
24
33
90.5
91
91.5
92
92.5
93
93.5
94
94.5
95

No groups have means significantly different from Group 33

More Built-In Functions

- Two-way ANOVA

```
>> [P, tbl, stats] = anova2(X, reps)
```

- Statistical plots

```
>> boxplot(X, group)
```

- Other hypothesis tests

```
>> H = ttest(X)
>> H = lillietest(X)
```

Exercise 1: Data Analysis

- RFID and Barcode Scanning Tests
  - Script m-file: `dataanalysis.m`

Follow instructions in the m-file …

Questions?

Curve Fitting Toolbox

Curve Fitting Tool

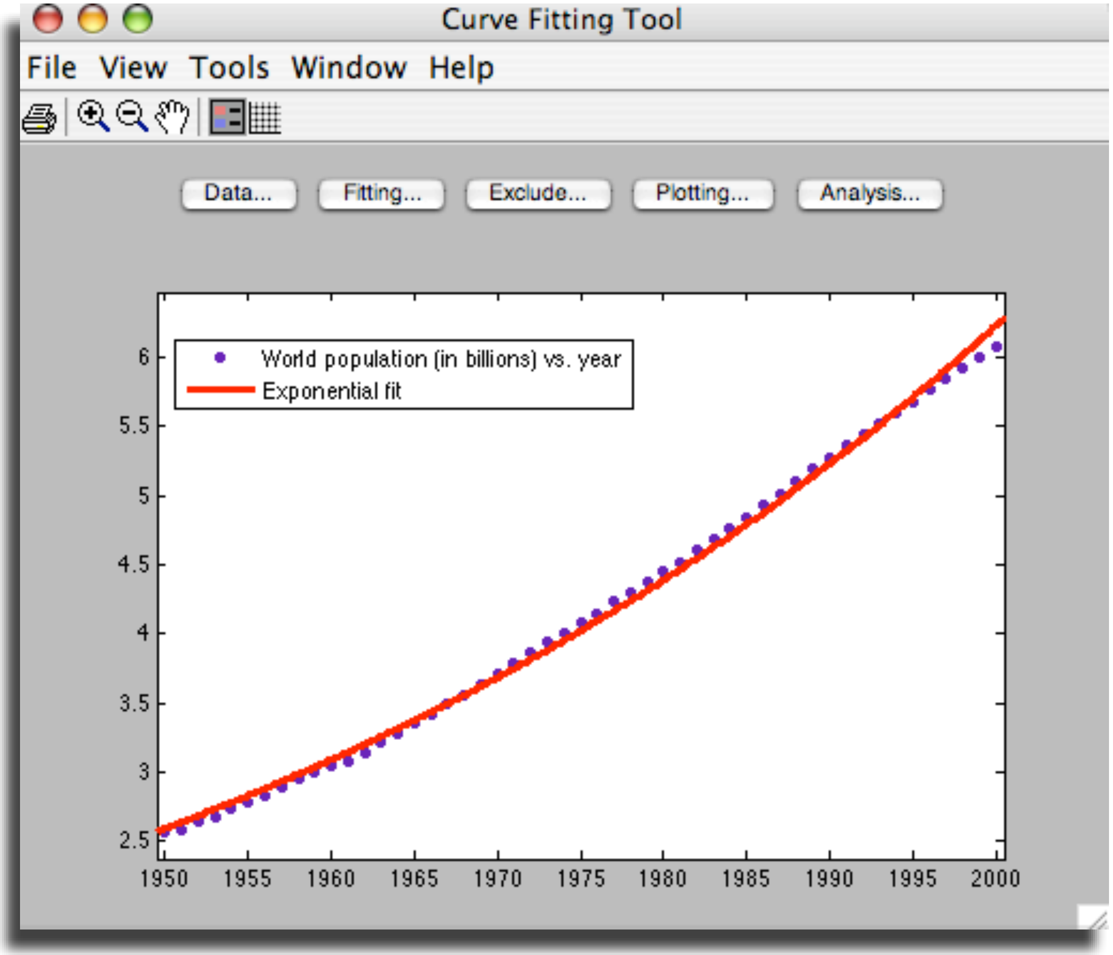
Goodness of Fit

Analyzing a Fit

Fourier Series Fit

Curve Fitting Tool

```
>> cftool
```



Goodness of Fit Statistics

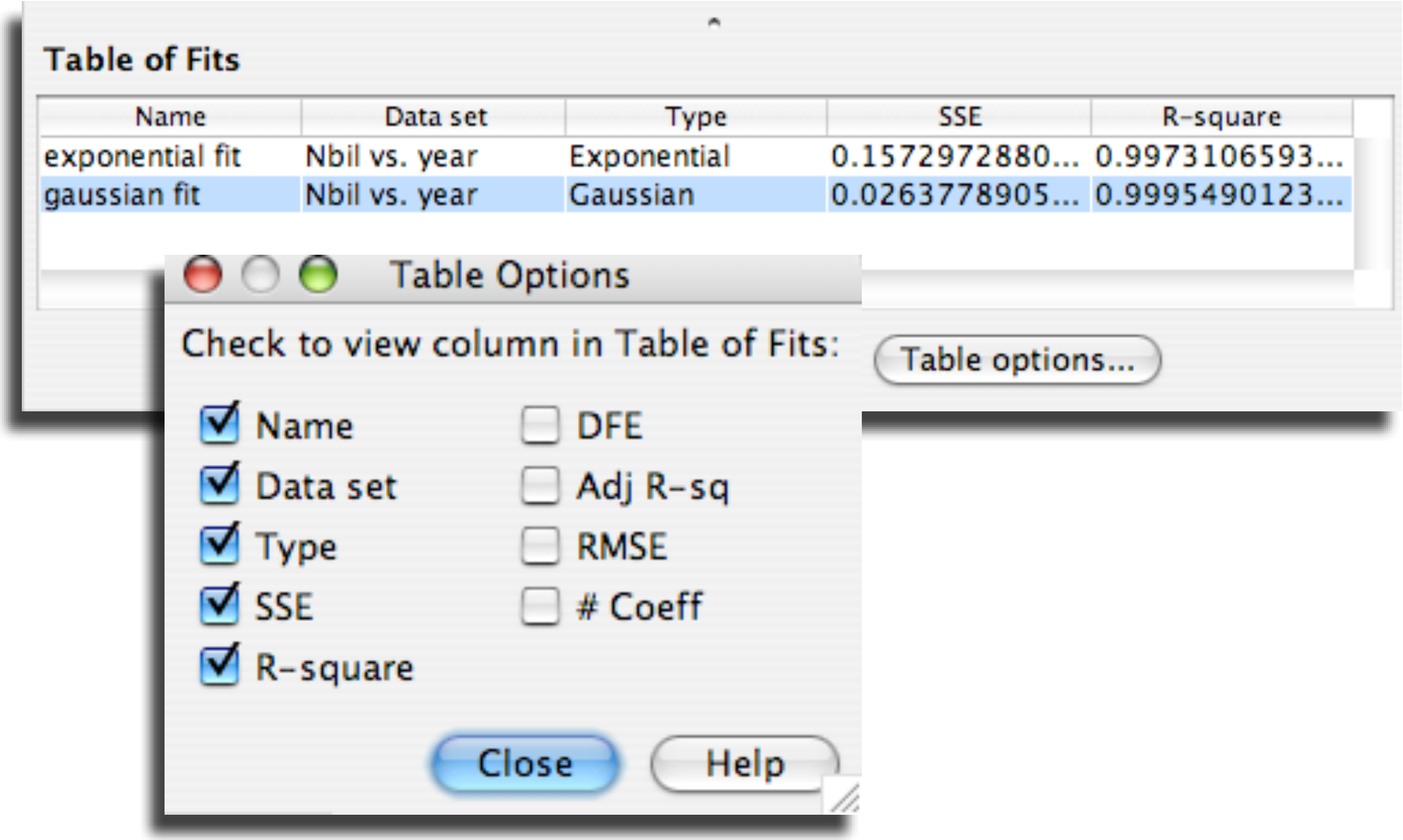


Table of Fits

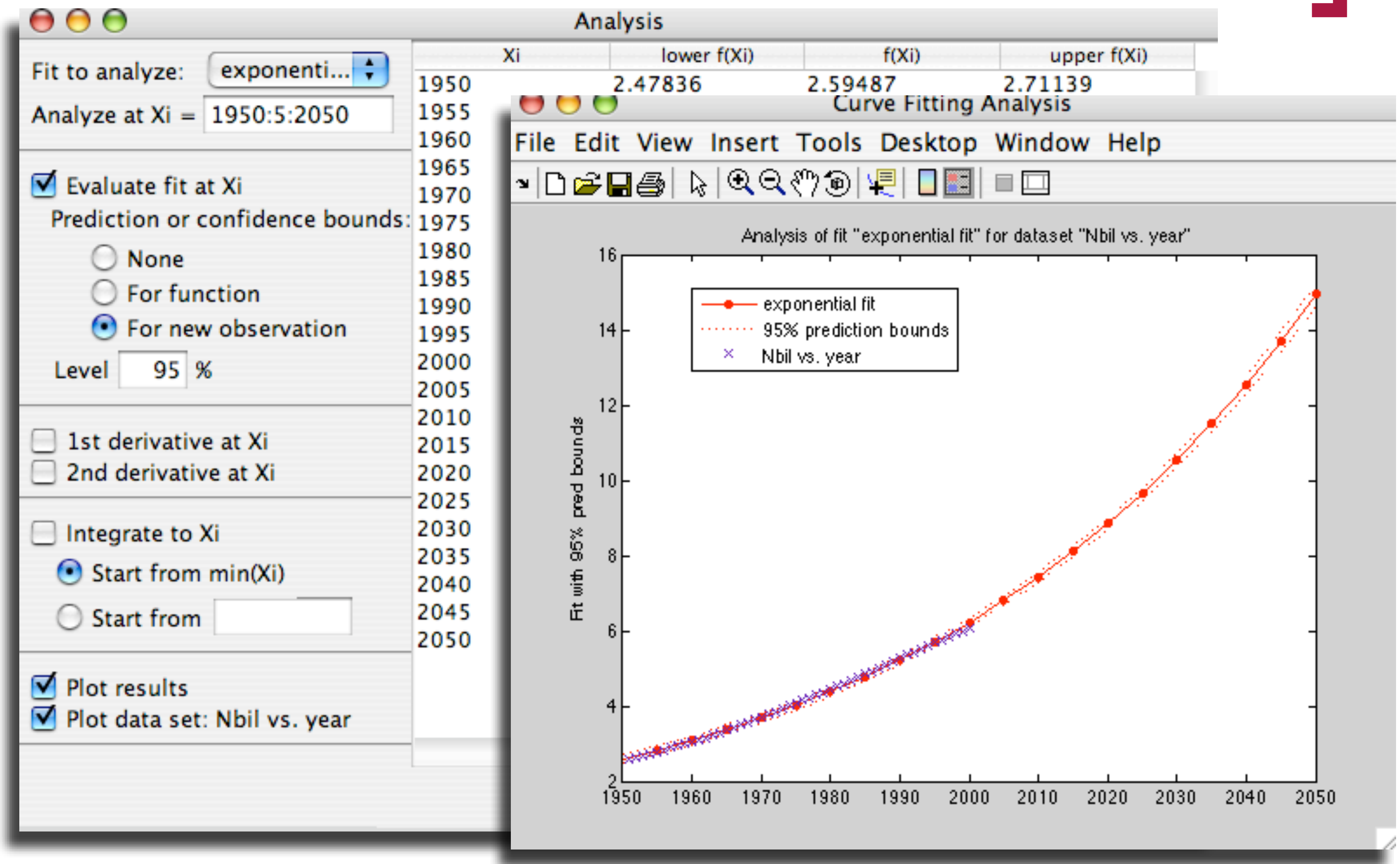
Name	Data set	Type	SSE	R-square
exponential fit	Nbil vs. year	Exponential	0.1572972880...	0.9973106593...
gaussian fit	Nbil vs. year	Gaussian	0.0263778905...	0.9995490123...

Table Options

Check to view column in Table of Fits:

☑ Name ☐ DFE
☑ Data set ☐ Adj R-sq
☑ Type ☐ RMSE
☑ SSE ☐ # Coeff
☑ R-square

Analyzing a Fit



Analysis

Fit to analyze: exponenti...

Analyze at Xi = 1950:5:2050

☑ Evaluate fit at Xi

Prediction or confidence bounds:

○ None

○ For function

● For new observation

Level 95 %

□ 1st derivative at Xi

□ 2nd derivative at Xi

□ Integrate to Xi

● Start from min(Xi)

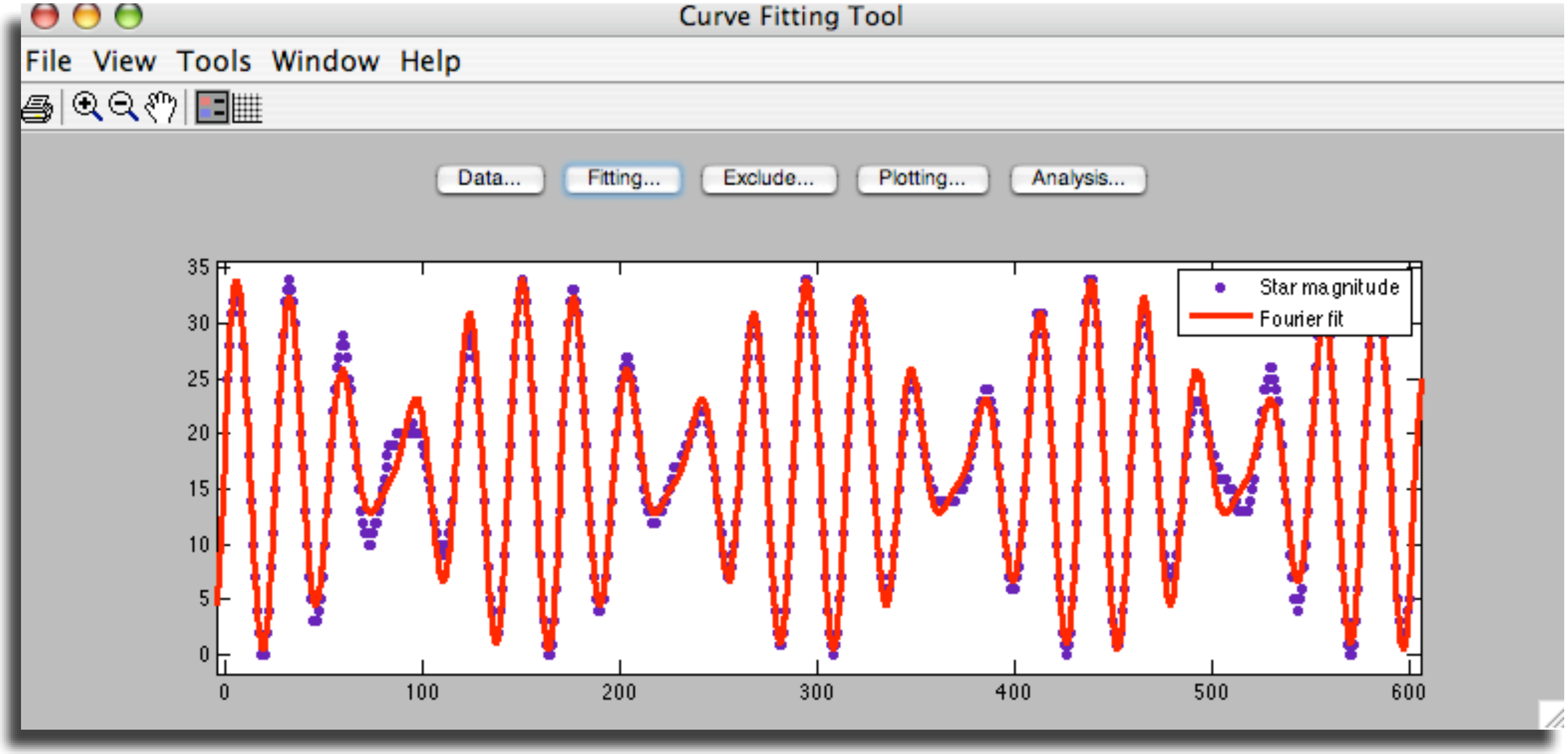
○ Start from

☑ Plot results

☑ Plot data set: Nbil vs. year

Xi	lower f(Xi)	f(Xi)	upper f(Xi)
1950	2.47836	2.59487	2.71139
1955			
1960			
1965			
1970			
1975			
1980			
1985			
1990			
1995			
2000			
2005			
2010			
2015			
2020			
2025			
2030			
2035			
2040			
2045			
2050			

Fourier Series Fit



Exercise 2: Regression

- Linear regression & other line fitting
  - Script m-file: `regression.m`

Follow instructions in the m-file …

Questions?