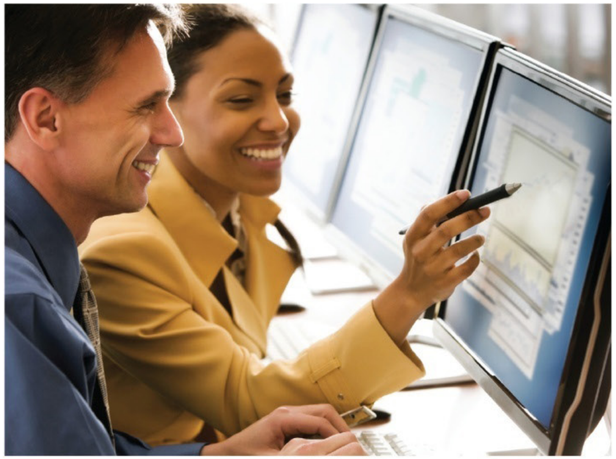
Analyzing Data

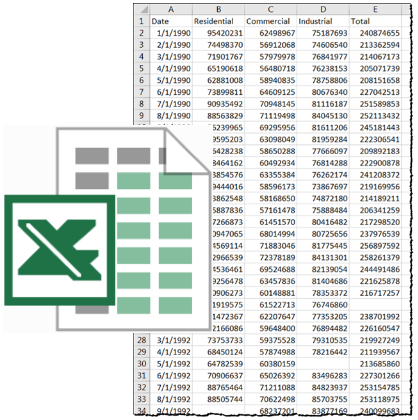
MATLAB® Fundamentals for Aerospace Applications



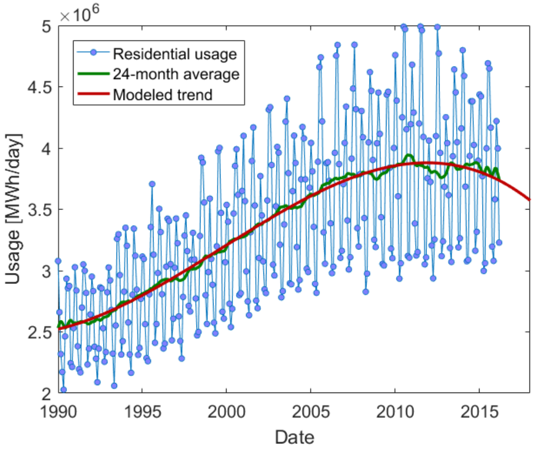
MathWorks | Training Services

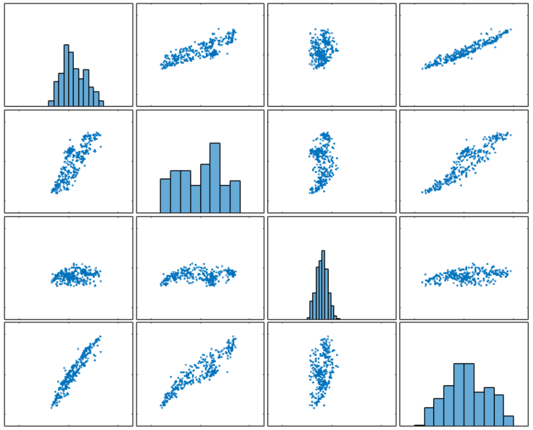
Outline

- Importing data from file
- Normalizing data
- Dealing with missing data
- Polynomial fitting
- Creating customized visualizations

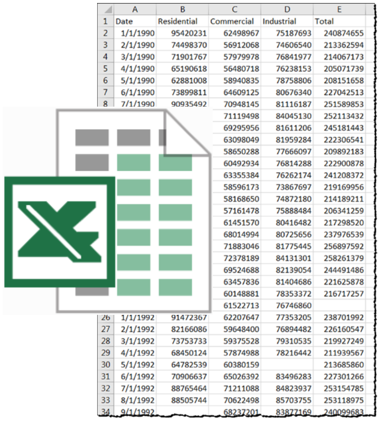


	A	B	C	D	E
1	Date	Residential	Commercial	Industrial	Total
2	1/1/1990	95420231	62498967	75187693	240874655
3	2/1/1990	74498370	56912068	74606540	213362594
4	3/1/1990	71901767	57979978	76841977	214067173
5	4/1/1990	65190618	56480718	76238153	205071739
6	5/1/1990	62881008	58940835	78758806	208151658
7	6/1/1990	73899811	64609125	80676340	227042513
8	7/1/1990	90935492	70948145	81116187	251589853
9	8/1/1990	88563829	71119498	84045130	252113432
		6239965	69295956	81611206	245181443
		9595203	63098049	81959284	222306541
		6428238	58650288	77666097	209892183
		8464162	60492934	76814288	222900878
		3854576	63355384	76262174	241208372
		9444016	58596173	73867697	219169956
		3862548	58168650	74872180	214189211
		5887836	57161478	75888484	206341259
		7266873	61451570	80416482	217298520
		0947065	68014994	80725656	237976539
		4569114	71883046	81775445	256897592
		2966539	72378189	84131301	258261379
		4536461	69524688	82139054	244491486
		9256478	63457836	81404686	221625878
		0906273	60148881	78353372	216717257
		1919575	61522713	76746860	
		1472367	62207647	77353205	238701992
		2166086	59648400	76894482	226160547
28	3/1/1992	73753733	59375528	79310535	219927249
29	4/1/1992	68450124	57874988	78216442	211939567
30	5/1/1992	64782539	60380159		213685860
31	6/1/1992	70906637	65026392	83496283	227301266
32	7/1/1992	88765464	71211088	84823937	253154785
33	8/1/1992	88505744	70622498	85703755	253118975
34	9/1/1992		68237201	83877169	240099683



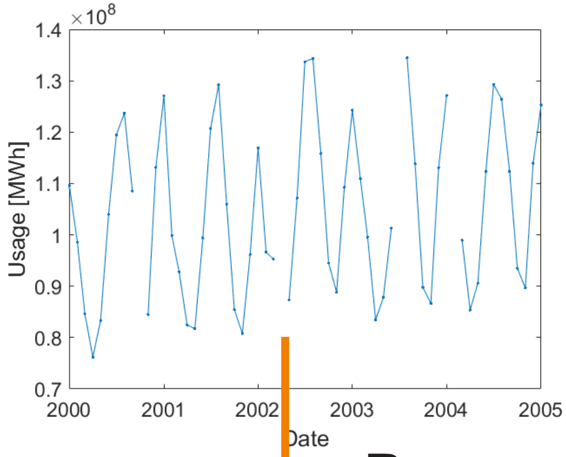


Course Example: Modeling Electricity Consumption

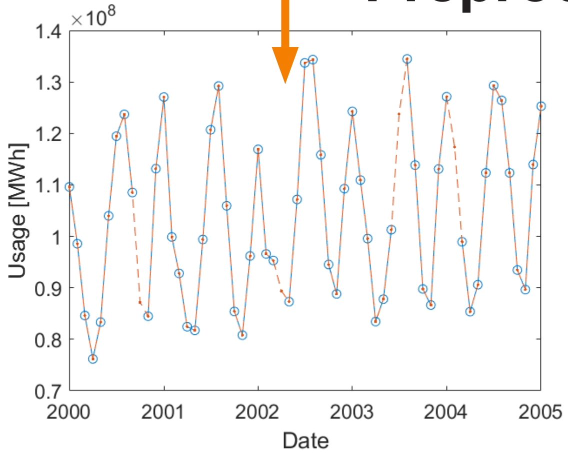


	A	B	C	D	E
1	Date	Residential	Commercial	Industrial	Total
2	1/1/1990	95420231	62498967	75187693	240874655
3	2/1/1990	74498370	56912068	74606540	213362594
4	3/1/1990	71901767	57979978	76841977	214067173
5	4/1/1990	65190618	56480718	76238153	205071739
6	5/1/1990	62881008	58940835	78758806	208151658
7	6/1/1990	73899811	64609125	80676340	227042513
8	7/1/1990	90935492	70948145	81116187	251589853
			71119498	84045130	252113432
			69295956	81611206	245181443
			63098049	81959284	222306541
			58650288	77666097	209892183
			60492934	76814288	222900878
			63355384	76262174	241208372
			58596173	73867697	219169956
			58168650	74872180	214189211
			57161478	75888484	206341259
			61451570	80416482	217298520
			68014994	80725656	237976539
			71883046	81775445	256897592
			72378189	84131301	258261379
			69524688	82139054	244491486
			63457836	81404686	221625878
			60148881	78353372	216717257
			61522713	76746860	
26	1/1/1992	91472367	62207647	77353205	238701992
27	2/1/1992	82166086	59648400	76894482	226160547
28	3/1/1992	73753733	59375528	79310535	219927249
29	4/1/1992	68450124	57874988	78216442	211939567
30	5/1/1992	64782539	60380159		213685860
31	6/1/1992	70906637	65026392	83496283	227301266
32	7/1/1992	88765464	71211088	84823937	253154785
33	8/1/1992	88505744	70622498	85703755	253118975
34	9/1/1992		68237201	83877169	240099683

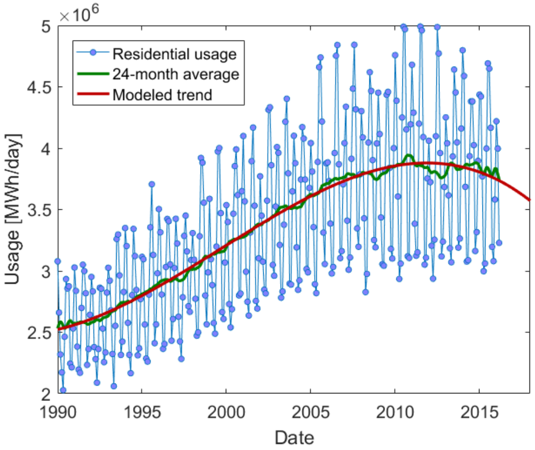
Import



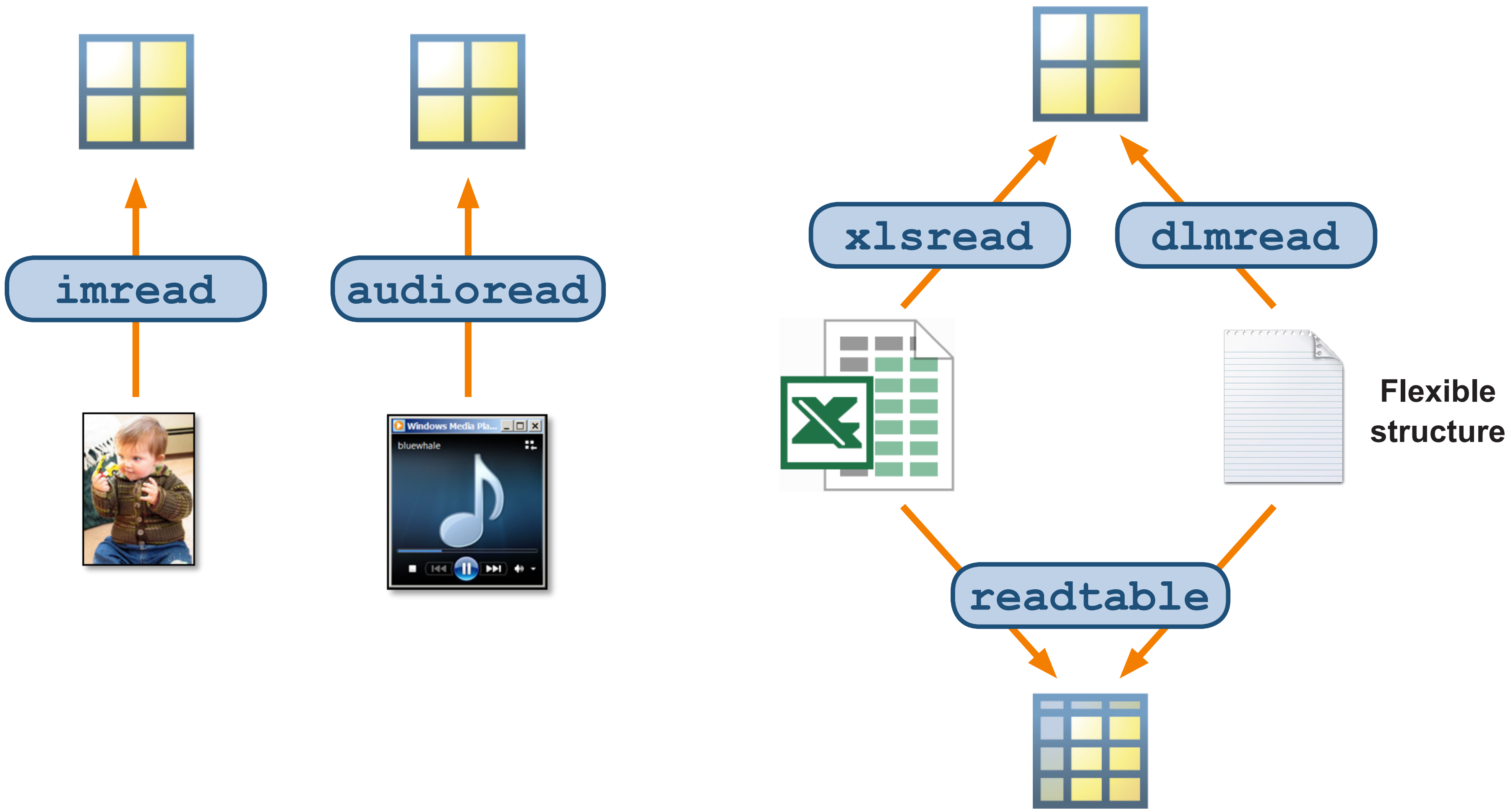
Preprocess



Model



Importing Data Programmatically



imread

audioread

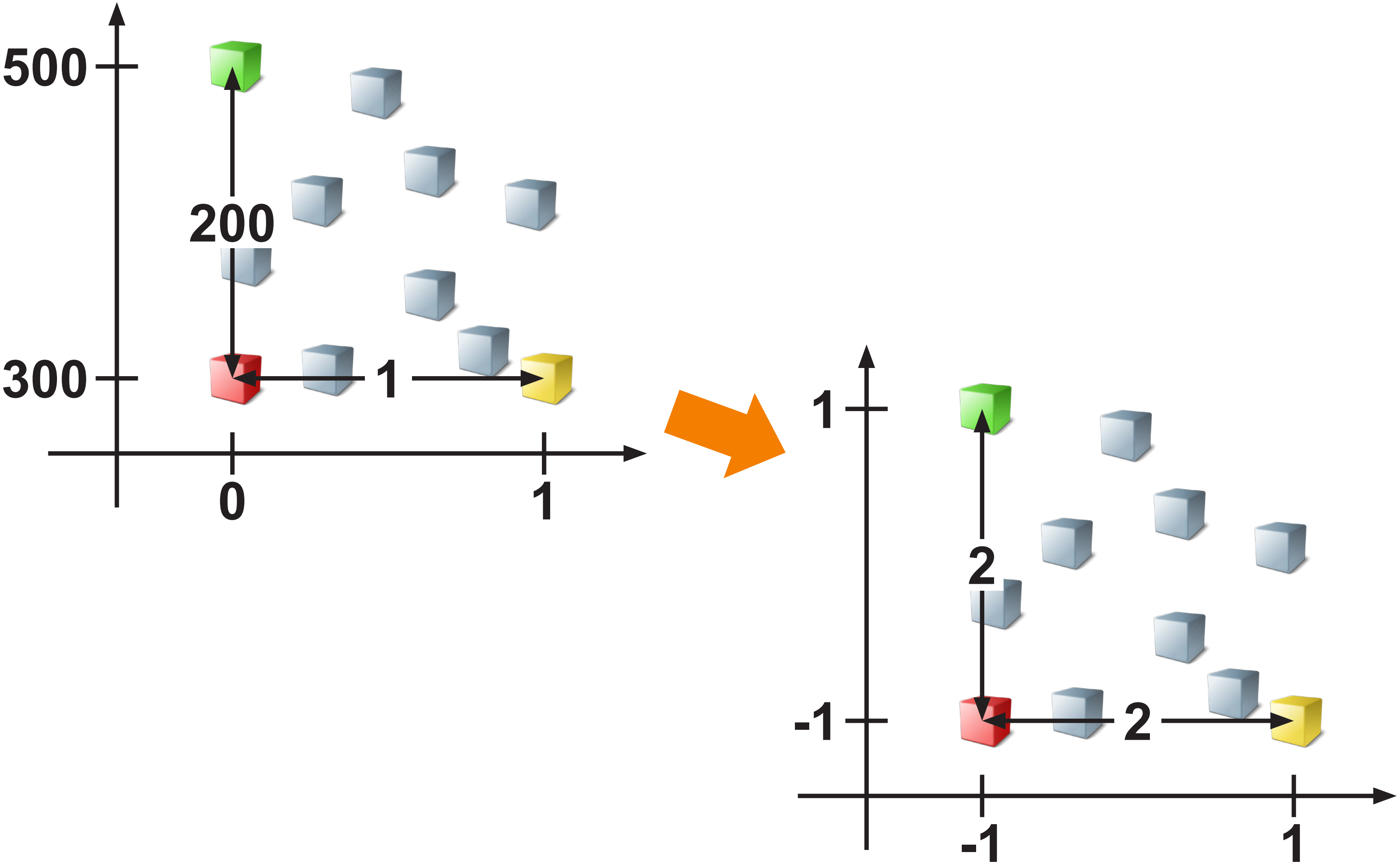
xlsread

dlmread

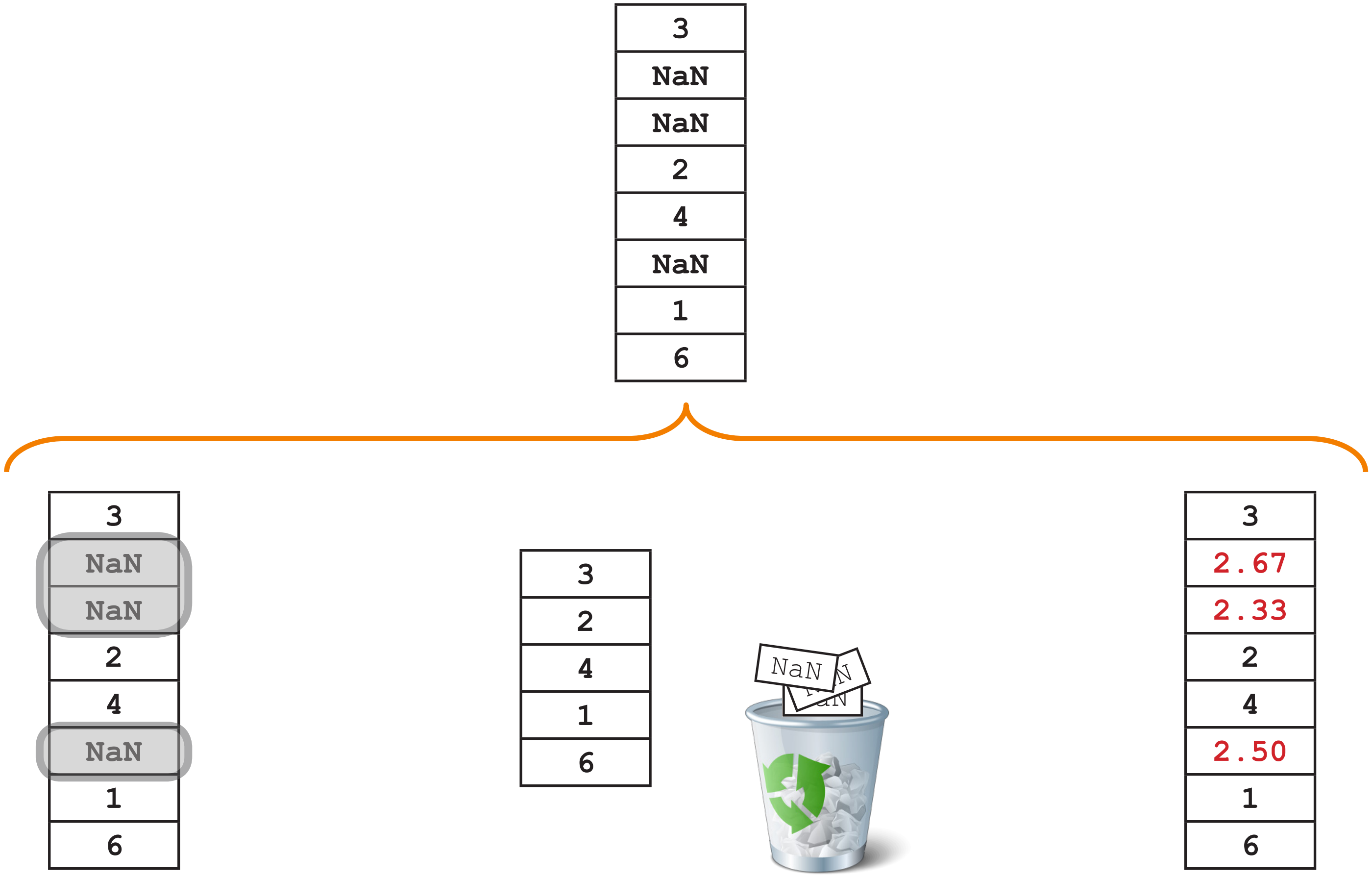
readtable

Flexible structure

Normalizing Data



Dealing with Missing Data



3
NaN
NaN
2
4
NaN
1
6

3
NaN
NaN
2
4
NaN
1
6

Ignore

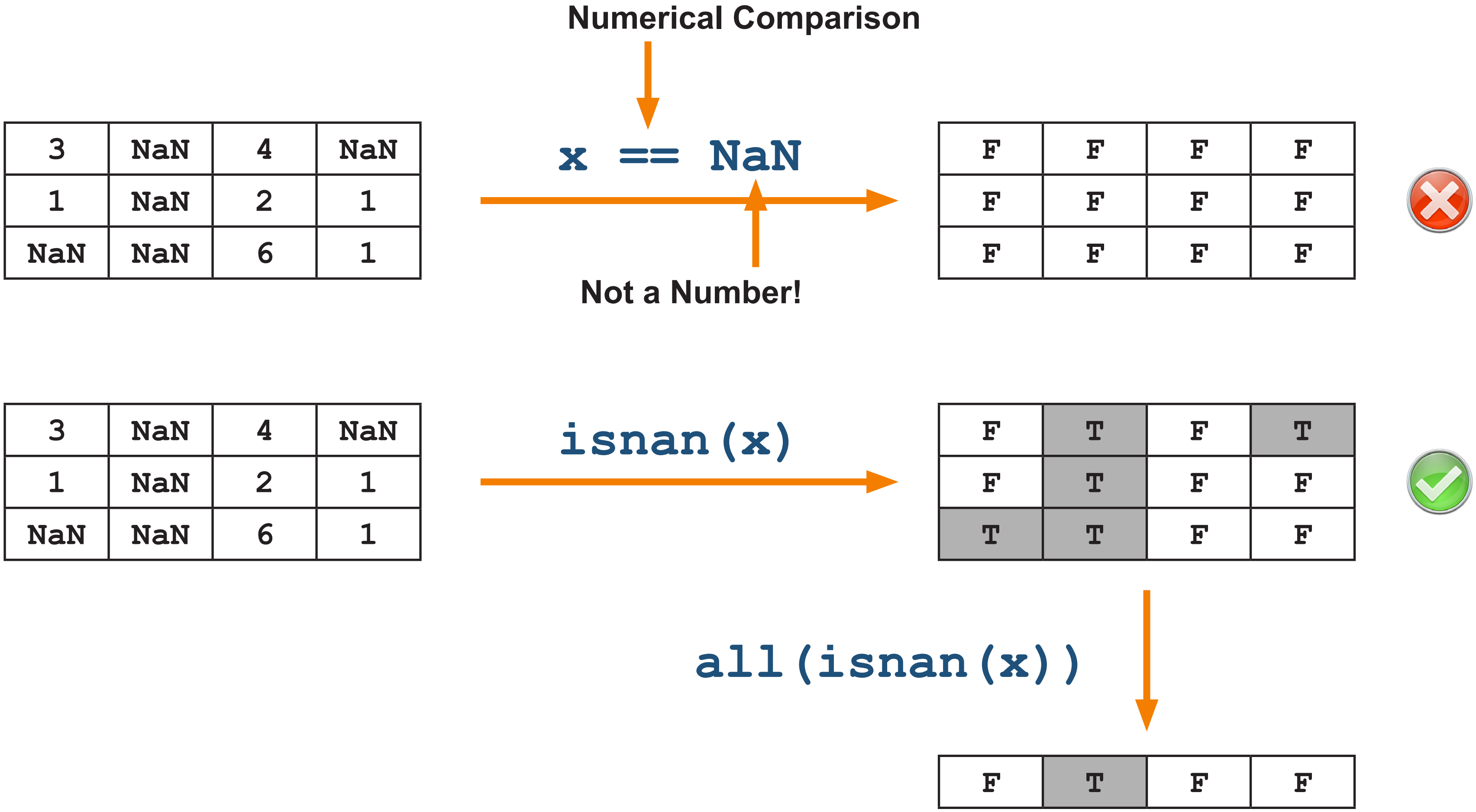
3
2
4
1
6

Delete

3
2.67
2.33
2
4
2.50
1
6

Replace

Locating Missing Values



Numerical Comparison

3	NaN	4	NaN
1	NaN	2	1
NaN	NaN	6	1

```
x == NaN
```

Not a Number!

F	F	F	F
F	F	F	F
F	F	F	F

✗

3	NaN	4	NaN
1	NaN	2	1
NaN	NaN	6	1

```
isnan(x)
```

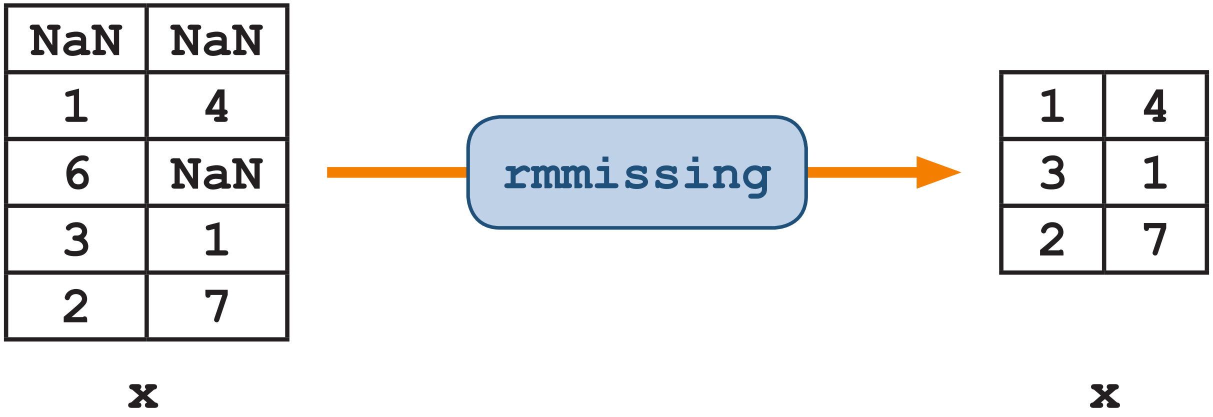
F	T	F	T
F	T	F	F
T	T	F	F

✓

```
all(isnan(x))
```

F	T	F	F

Removing Missing Values



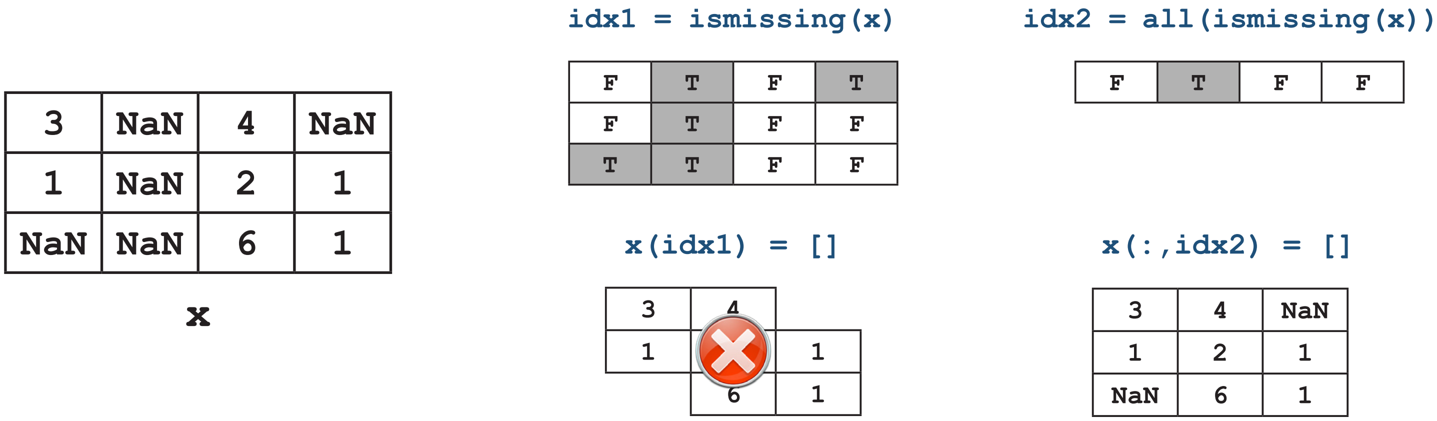
NaN	NaN
1	4
6	NaN
3	1
2	7

x

`rmmissing`

1	4
3	1
2	7

x



3	NaN	4	NaN
1	NaN	2	1
NaN	NaN	6	1

x

```
idx1 = ismissing(x)
```

F	T	F	T
F	T	F	F
T	T	F	F

```
x(idx1) = []
```

3	4	
1		1
	6	1

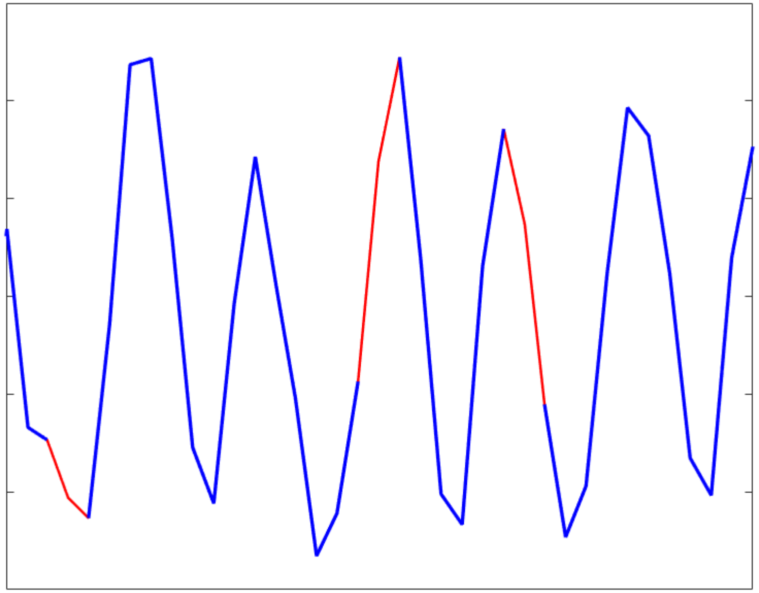
```
idx2 = all(ismissing(x))
```

F	T	F	F

```
x(:,idx2) = []
```

3	4	NaN
1	2	1
NaN	6	1

Replacing Missing Values



3
NaN
NaN
2
4
NaN
1
6

x

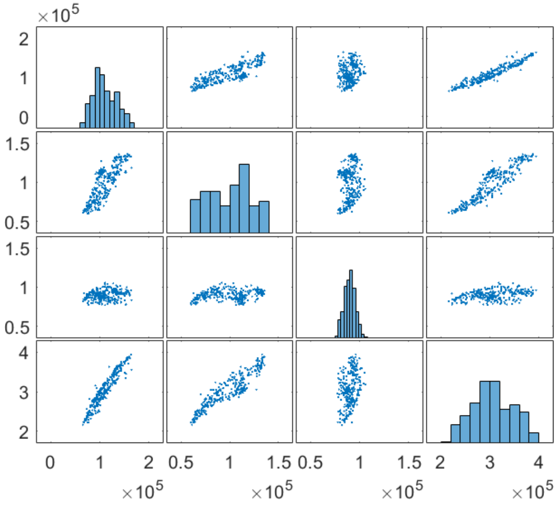
```
y = fillmissing(x,'linear')
```

3
2.67
2.33
2
4
2.50
1
6

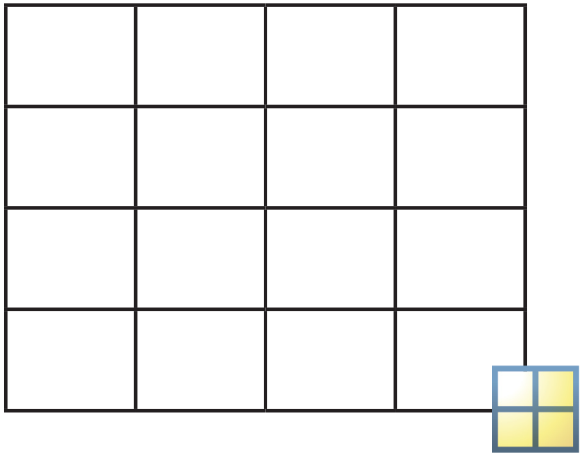
y

Linear Correlation

`plotmatrix`

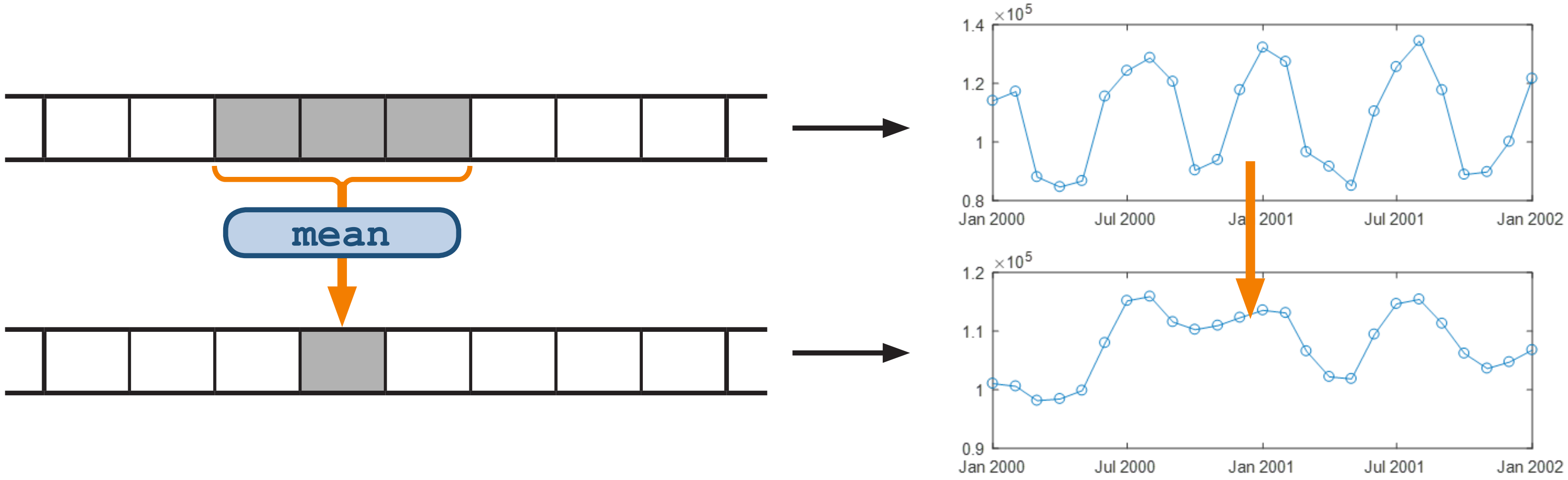


`corrcoef`

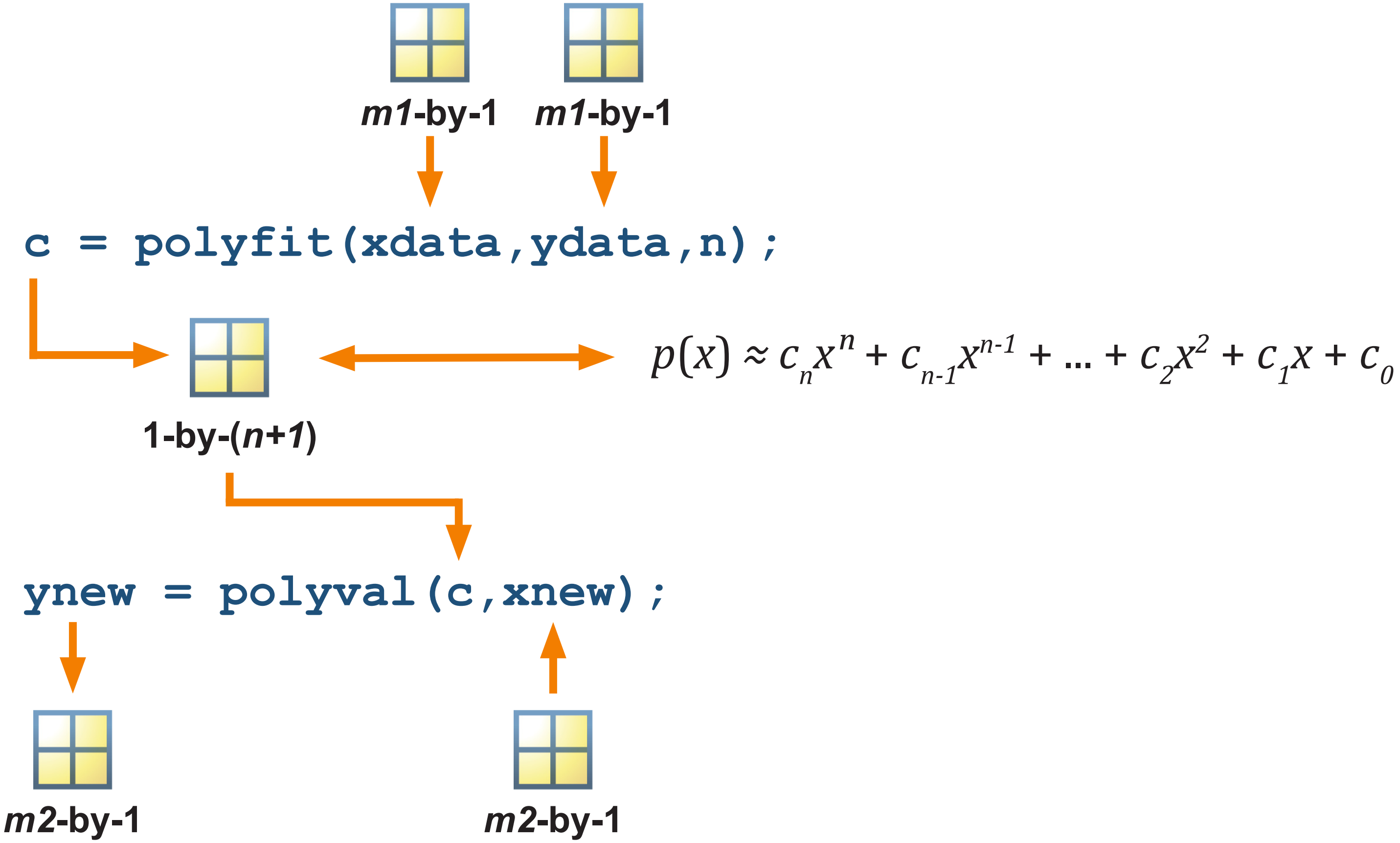


Moving Window Operations

```
s = movmean(x,3)
```



Fitting a Polynomial

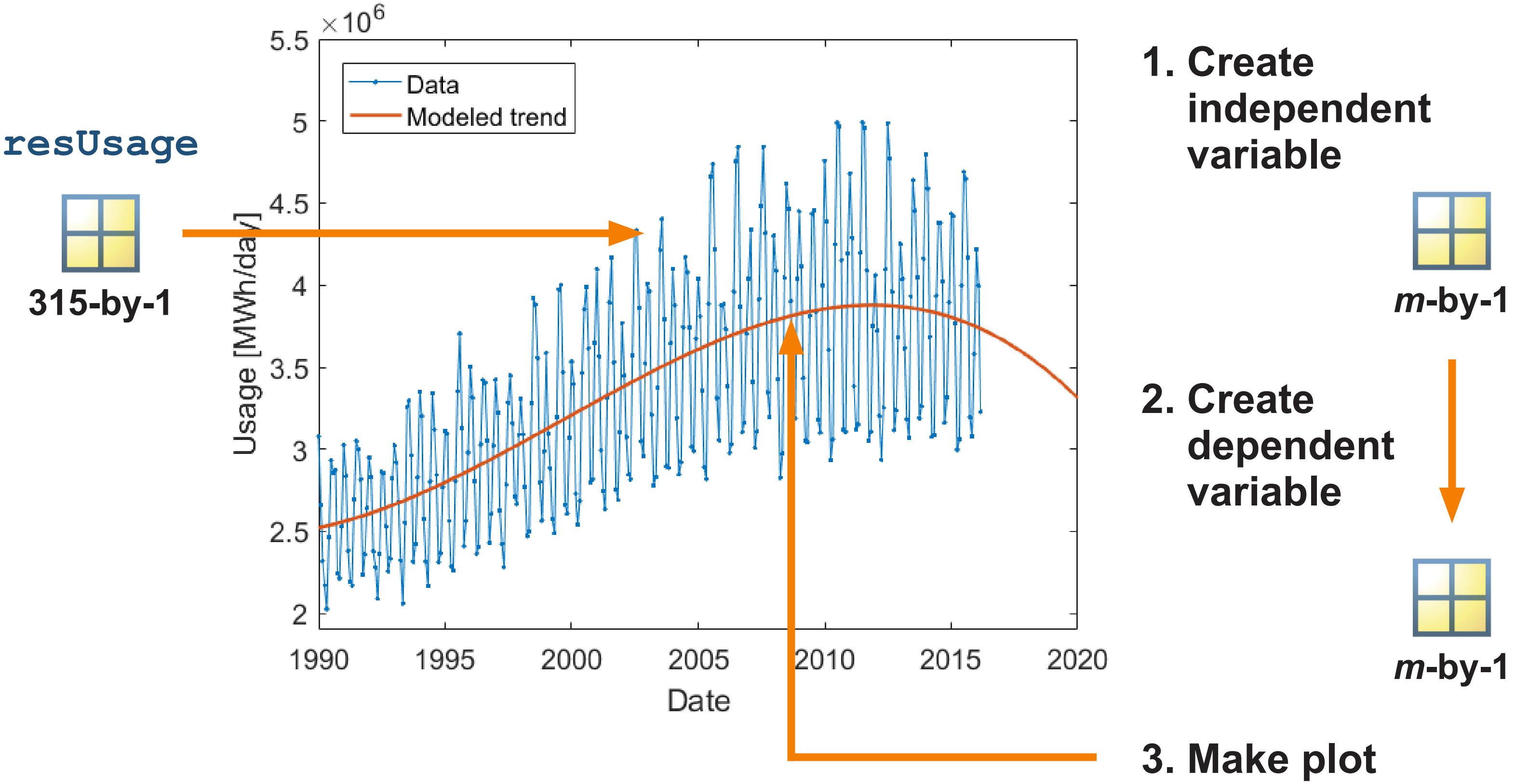


```
c = polyfit(xdata,ydata,n);
```

$$p(x) \approx c_n x^n + c_{n-1} x^{n-1} + \ldots + c_2 x^2 + c_1 x + c_0$$

```
ynew = polyval(c,xnew);
```

Adding a Theoretical Curve



`resUsage`

315-by-1

1. Create independent variable

m-by-1

2. Create dependent variable

m-by-1

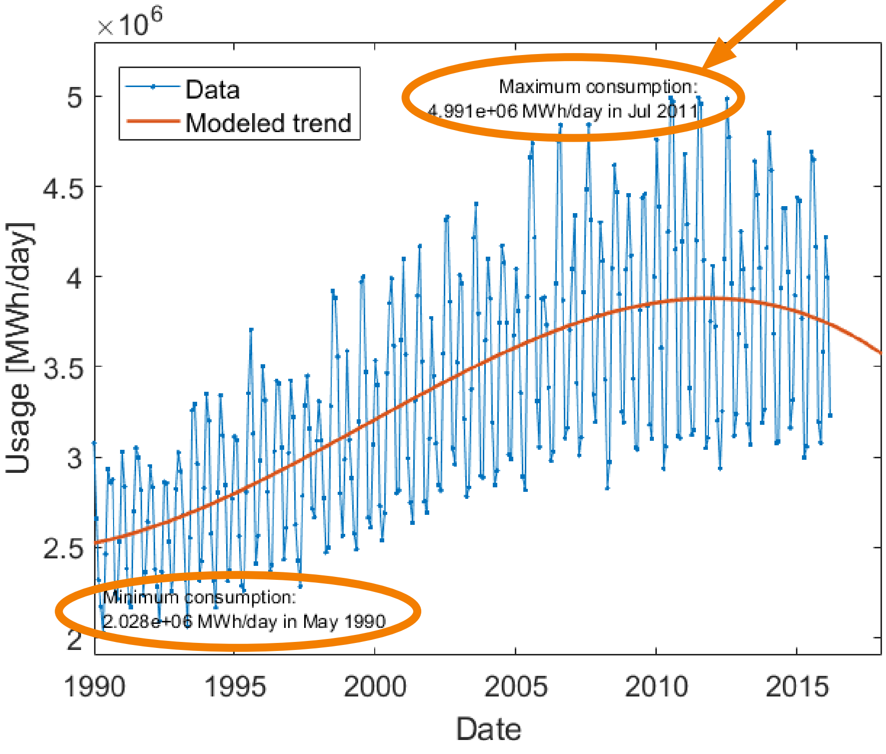
3. Make plot

Adding Annotations

Adding calculated information

`text(x,y, ch )`

`num2str`



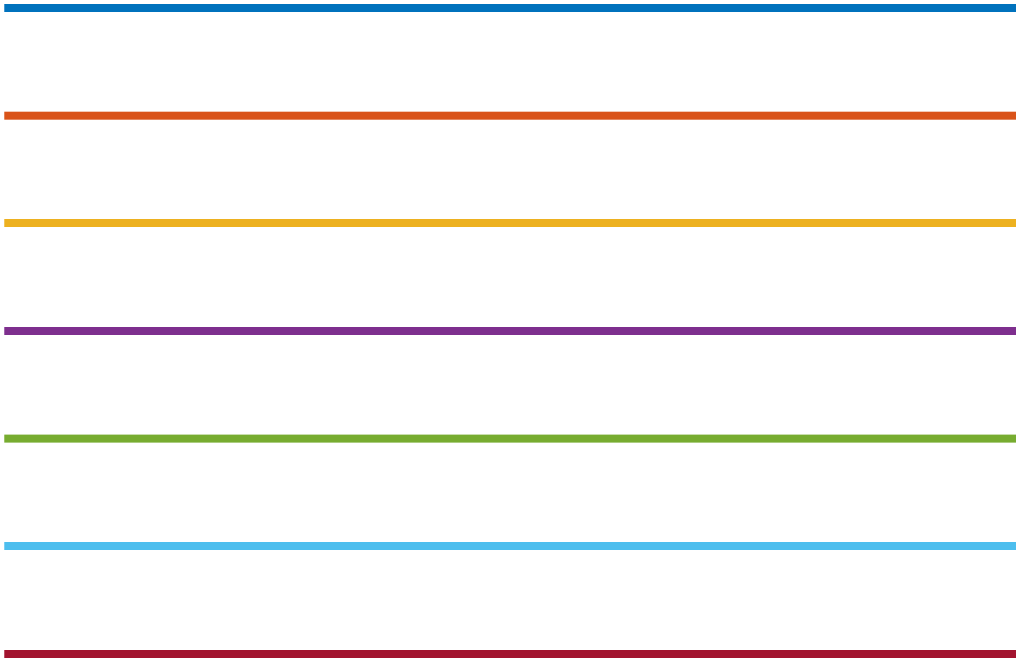
Multiple lines

{ } ch ch

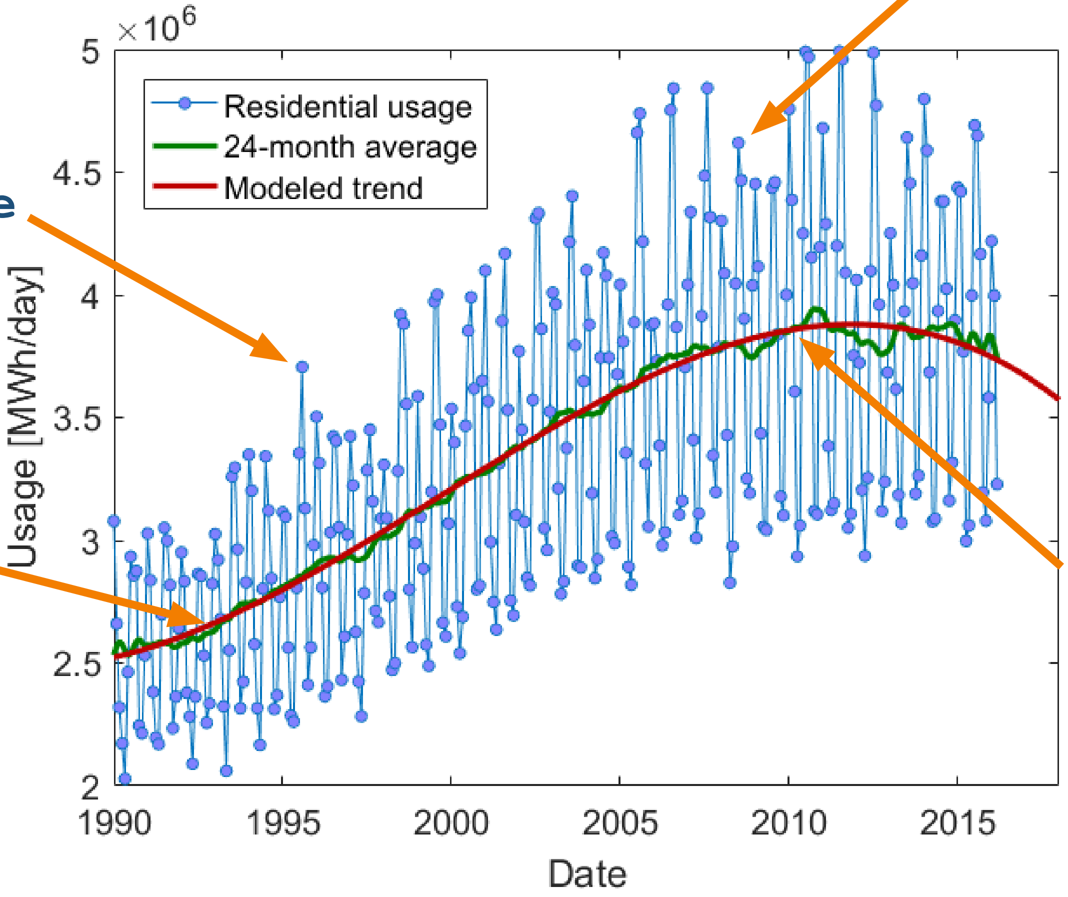
Specifying Color

```
[ R      G      B   ]

[0.00   0.45   0.74]
[0.85   0.33   0.10]
[0.93   0.69   0.13]
[0.49   0.18   0.56]
[0.47   0.67   0.19]
[0.30   0.75   0.93]
[0.64   0.08   0.18]
```



Customizing Plots



`MarkerFaceColor`

`MarkerSize`

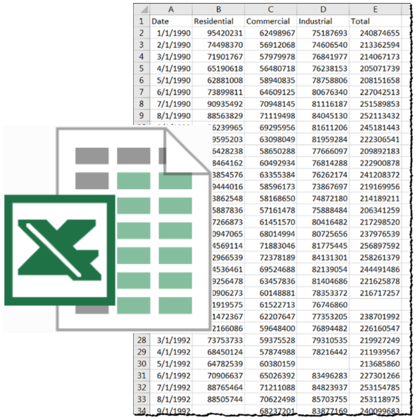
`Color`

`LineWidth`

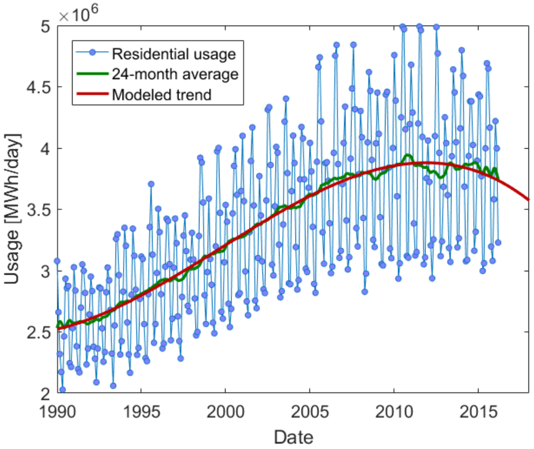
```
plot(x,y,'PropertyName',Value)
```

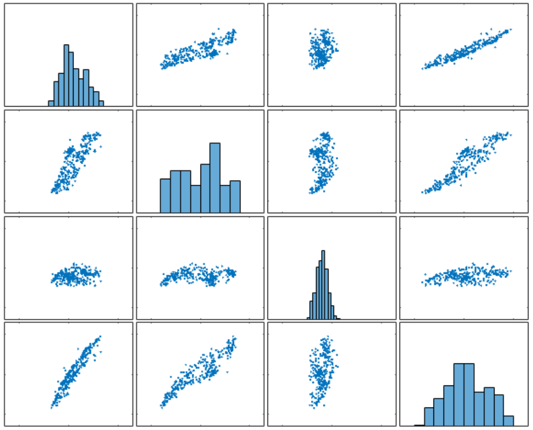
Summary

- Importing data from file
- Normalizing data
- Dealing with missing data
- Polynomial fitting
- Creating customized visualizations



	A	B	C	D	E
1	Date	Residential	Commercial	Industrial	Total
2	1/1/1990	95420231	62498967	75187693	240874655
3	2/1/1990	74498370	56912068	74606540	213362594
4	3/1/1990	71901767	57979978	76841977	214067173
5	4/1/1990	65190618	56480718	76238153	205071739
6	5/1/1990	62881008	58940835	78758806	208151658
7	6/1/1990	73899811	64609125	80676340	227042513
8	7/1/1990	90935492	70948145	81116187	251589853
9	8/1/1990	88563829	71119498	84045130	252113432
		6239965	69295956	81611206	245181443
		9595203	63098049	81959284	222306541
		6428238	58650288	77666097	209892183
		8464162	60492934	76814288	222900878
		3854576	63355384	76262174	241208372
		9444016	58596173	73867697	219169956
		3862548	58168650	74872180	214189211
		5887836	57161478	75888484	206341259
		7266873	61451570	80416482	217298520
		0947065	68014994	80725656	237976539
		4569114	71883046	81775445	256897592
		2966539	72378189	84131301	258261379
		4536461	69524688	82139054	244491486
		9256478	63457836	81404686	221625878
		0906273	60148881	78353372	216717257
		1919575	61522713	76746860	
		1472367	62207647	77353205	238701992
		2166086	59648400	76894482	226160547
28	3/1/1992	73753733	59375528	79310535	219927249
29	4/1/1992	68450124	57874988	78216442	211939567
30	5/1/1992	64782539	60380159		213685860
31	6/1/1992	70906637	65026392	83496283	227301266
32	7/1/1992	88765464	71211088	84823937	253154785
33	8/1/1992	88505744	70622498	85703755	253118975
34	9/1/1992		68237201	83877169	240099683





Test Your Knowledge

1. Which of the following makes a plot with a thick line?

   A.
   ```
   p = plot(x,y);
   LineWidth(p,4)
   ```

   B.
   ```
   plot(x,y,'LineWidth'=4)
   ```

   C.
   ```
   p = plot(x,y);
   p(LineWidth) = 4;
   ```

   D.
   ```
   plot(x,y,'LineWidth',4)
   ```

Test Your Knowledge

2. Given 1-by-50 vectors `x` and `y`, what is the result of the command `z = polyfit(x,y,3)`?

   A. A 1-by-3 vector of points interpolating `y` as a function of `x`

   B. A 1-by-4 vector representing the coefficients of a cubic polynomial fitted to `y` as a function of `x`

   C. A 1-by-50 vector of the values of a cubic polynomial fitted to `y` as a function of `x`

   D. An error message