LESSON QUESTIONS: Bar charts

FOCUS QUESTION: How can I show proportions and relative sizes of different data groups?

Contents

- EXAMPLE 1: Load the data about New York contagious diseases
- EXAMPLE 2: Calculate individual and overall monthly totals (combining columns)
- EXAMPLE 3: Create a bar chart of the measles monthly case totals (bar(Y))
- EXAMPLE 4: Label the bars explicitly (set(gca, ...))
- EXAMPLE 5: Create a bar chart of the measles yearly case totals (bar(X,Y))
- EXAMPLE 6: Create a side-by-side bar chart of monthly totals
- EXAMPLE 7: Create a stacked bar chart of monthly totals
- EXAMPLE 8: Create horizontal stacked bar chart using a different color scheme (colormap)

EXAMPLE 1: Load the data about New York contagious diseases

```
load NYCDiseases.mat;
```

EXAMPLE 2: Calculate individual and overall monthly totals (combining columns)

```
measlesByMonth = sum(measles, 1);

mumpsByMonth = sum(mumps, 1);

CPByMonth = sum(chickenPox, 1);

byMonth = [measlesByMonth', mumpsByMonth', CPByMonth'];
```

<table>
<thead>
<tr>
<th>Questions</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the size of sum(measles)?</td>
<td>Since <code>measles</code> is an array with 41 rows and 12 columns, the result of <code>sum(measles)</code> is an array with 1 row of 12 columns. (Remember that <code>sum</code> without a second argument calculates the sums of the individual columns for 2D arrays.)</td>
</tr>
<tr>
<td>What does [a, b, c] mean?</td>
<td>The square brackets ([ ]) indicate a new array formed by placing the arrays a, b, and c side-by-side.</td>
</tr>
<tr>
<td>What does the prime (') mean?</td>
<td>The prime (') means to take the transpose of the array. The transpose operator flips the array along the main diagonal to produce a new array. (The rows of the original array are the columns)</td>
</tr>
</tbody>
</table>
### Questions | Answers
--- | ---
What is the size of `measlesByMonth'`? | Since `measlesByMonth` is a single row with 12 columns, the size of `measlesByMonth'` is 1 column of 12 rows (i.e., 1 x 12).
Why does `byMonth` have 12 rows and 3 columns? | MATLAB forms the array by placing side-by-side three arrays, each consisting of a single column of 12 rows.
Why was the transpose operator necessary? | Without the transpose operator, the row vectors `measlesByMonth`, `mumpsByMonth` and `CPByMonth` would be placed side-by-side to form a single row vector of 36 elements. Plotting `byMonth` using a bar graph results in the 12 bars of the measles monthly totals, followed by the 12 bars of the mumps monthly totals, and ending with 12 bars of the chicken pox monthly totals. In order to have a bar graph where the bars representing January for the three diseases are grouped together, the array must have the January values as the first row.

**EXAMPLE 3: Create a bar chart of the measles monthly case totals (bar(Y))**

```matlab
figure

bar(measlesByMonth./1000)

xlabel('Month')

ylabel('Cases (in thousands)')

title('Measles by month in NYC: 1931-1971')
```
Questions | Answers
--- | ---
How many bars does this graph have? | The graph has 12 bars, representing the 12 months.
Why rescale the data using units of thousands of cases rather than just cases? | Tick mark labels of large magnitude are difficult to read and comprehend. MATLAB tries to alleviate this problem by presenting the tick mark labels in scientific notation with the exponent displayed above the axis. However, the exponent is easy to miss. The axis rescaling makes the graph easier for most viewers to understand.

**EXAMPLE 4: Label the bars explicitly (set(gca, ...))**

```matlab
mylabels = {'J', 'F', 'M', 'A', 'M', 'J', 'J', 'A', 'S', 'O', 'N', 'D'};

figure

bar(measlesByMonth./1000)

xlabel('Month')

ylabel('Cases (in thousands)')

title('Measles by month in NYC: 1931-1971')
```
set(gca,'XTickLabelMode', 'manual', ...
'
'XTickLabel', mylabels)

<table>
<thead>
<tr>
<th>Questions</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is gca?</td>
<td>The gca identifier is a reference or handle to the current axes. Use gca to get or set graphical properties of the current axes in the current figure. You can save this handle in a variable and refer to this axes later, when you have created other axes and this axes is not the current one.</td>
</tr>
<tr>
<td>Why do I need to set the mode for the x-axis tick labels to be manual?</td>
<td>The 'manual' setting of the 'XTickLabelMode' property specifies that MATLAB should not try to automatically set the tick mark labels. Instead, MATLAB should use the labels that you have given using the 'XTickLabel' property of the current axis.</td>
</tr>
<tr>
<td>What tick labels does MATLAB automatically choose for bar graphs?</td>
<td>MATLAB uses the bar positions 1, 2, ... if you don't specify your own tick mark labels.</td>
</tr>
<tr>
<td>Why are the tick mark labels enclosed with single quotes?</td>
<td>The enclosing single quotes specify that the labels are strings (i.e., messages not a variables).</td>
</tr>
<tr>
<td>Why was the list of tick mark labels enclosed with curly braces rather</td>
<td>Use square brackets ([ ] ) to create an ordinary array, which must have elements of the same</td>
</tr>
<tr>
<td>Questions</td>
<td>Answers</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>than square brackets?</td>
<td>size and type. Use curly braces ({ }) to create a cell array, which may have elements of different sizes and types. In this example the array elements are string, which usually differ in length.</td>
</tr>
</tbody>
</table>

**EXAMPLE 5: Create a bar chart of the measles yearly case totals (bar(X,Y))**

```matlab
figure

bar(years, sum(measles, 2)./1000)

xlabel('Year')

ylabel('Cases (in thousands)')

title('Measles by year in NYC: 1931-1971')
```

**EXAMPLE 6: Create a side-by-side bar chart of monthly totals**

```matlab
figure
```
Questions | Answers
--- | ---
What does each bar in this graph represent? | An individual bar represents the total number of cases for a particular disease in a particular month.
How are the bars grouped? | MATLAB groups by row number. That is, all of the bars for the first row are displayed at position 1, etc. Thus, the three bars at position 1 correspond to the totals for the three diseases in January.
What is the order of the bars within each group? | MATLAB determines the order by the column position. Thus, the bars in group 1 correspond to the January totals for measles, mumps, and chicken pox, respectively.

EXAMPLE 7: Create a stacked bar chart of monthly totals
Questions | Answers
---|---
What does each bar in this graph represent? | The total height of each bar corresponds to the total of a row of `byMonth` (in thousands). For example, bar 1 is the total number of cases of all three diseases for January.
What does each band in this graph represent? | Each element in `byMonth` corresponds to one band in this graph. The bands in bar 1 correspond to the 3 elements in the first row of `byMonth`, that is, the totals of each of the three diseases for January.
How is the order of the bands determined within each bar?

MATLAB shows the bands in the column order of elements, with the value in column 1 shown on the bottom. Thus, the band at the bottom of bar 1 corresponds to the total number of cases of measles for January.

How do the groupings in Examples 6 and 7 differ?

Each group plots the same elements, but the 'stack' option draws the bands on top of each other in a single bar, while the other version draws the bands as side-by-side bars.

Why use stacked bar charts?

A stacked bar chart allows you to better estimate and compare the overall totals of the groups. You can comprehend a larger number of groups with the stacked bars. The chart of Example 6 is rather crowded.

EXAMPLE 8: Create horizontal stacked bar chart using a different color scheme (colormap)

Create a new cell in which you type and execute:

```matlab
figure
colormap spring
barh(byMonth./1000, 'stack')
xlabel('Cases (in thousands)')
ylabel('Month')
title('Childhood diseases by month in NYC: 1931-1971')
legend('Measles', 'Mumps', 'Chicken pox')
```

You should see a Figure Window containing a labeled horizontal stacked bar chart that uses a spring color scheme:
<table>
<thead>
<tr>
<th>Questions</th>
<th>Answers</th>
</tr>
</thead>
</table>
| How does the horizontal bar chart in this example differ from the vertical bar chart of Example 7? | The two types of charts show the same information, but MATLAB draws these charts in a different orientations.  

How do the orientations of vertical and horizontal bar charts differ?  
For the vertical bar chart, MATLAB draws the bars in an up and down orientation and the displays groups from left to right. For the horizontal bar chart, MATLAB draws the bars in a side ways orientation and displays the groups from bottom to top. Many details can be changed by setting properties.  

What is a color map?  
A color map is a way of assigning colors based on integers. For example, there is a color 1, color 2, etc.  

What determines how colors are assigned?  
MATLAB keeps a table of colors for each figure designated by the Colormap property. You can change this property for a figure by using the colormap command or by directly editing the figure.  

What is spring?  
MATLAB has several standard tables of colors, and spring designates one of those standard tables.  

These questions were written by Kay A. Robbins of the University of Texas at San Antonio and last modified on 01-Sep-2013. Please contact krobbins@cs.utsa.edu with comments or suggestions.