EXPLORATIONS

Chapter 5

Describing Categorical Data

Topic 13 covers categorizing data using tables and bar charts.

Topic 13—Tables and Bar Charts

Example: Tall buildings in North American cities saved in folder **BLDTALL**. (Set current folder to **BLDTALL** from the MODE screen.)

Note: *Although heights are numerical values, they can be grouped into three height categories.*

Cities in North America with over 20 tall buildings (500 feet or taller)

Height (in ft)	New York City	Chicago	Houston	Los Angeles	Totals
500-599	88	46	11	9	154
600-699	43	13	6	4	66
≥ 700	26	13	10	8	57
Total	157	72	27	21	277

(Source: Reprinted with permission from the World Almanac and Book of Facts 2000. © 2000 World Almanac Education Group, Inc. All rights reserved.)

From the two-way frequency table, you can find *joint frequencies*, such as the 46 buildings that are jointly in Chicago and in the 500 to 599 feet tall height group. You can find the *marginal frequencies*, such as the total of 72 tall buildings in Chicago or the total of 154 buildings that are between 500 to 599 feet tall combining all four cities. There are four marginal frequencies or totals in the lower margin and three in the right margin. There are 12 (or 4 x 3) joint frequencies in the table.

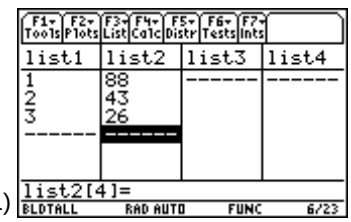
Bar Charts

Example: Consider the New York City data repeated below.

Height (in ft)	500-599	600-699	≥700	Total
Buildings in NYC	88	43	26	157

1. Clear the lists in the **BLDTALL** folder by highlighting the names of the lists to be cleared and pressing [CLEAR].
2. Press [MODE] and change Exact/Approx to **Auto**.
3. Enter the values **1**, **2**, and **3** in **list1** and **88**, **43**, and **26** in **list2** (screen 1).

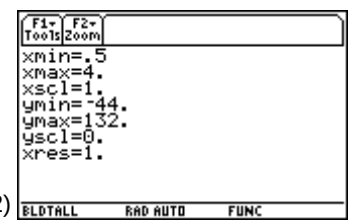
(1)



4. Set up the window using ◆ [WINDOW] with the following entries:
   - **xmin = .5**
   - **xmax = 4**
   - **xscl = 1**
   - **ymin = -44**
   - **ymax = 132**
   - **yscl = 0**
   - **xres = 1**

   (See screen 2.)

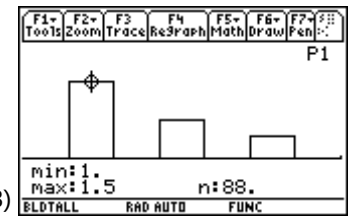
(2)



5. From the Plot Setup screen, set up and define **Plot 1** as a histogram with X List: **list1**, Hist. Bucket Width: **0.5**, and Freq: **list2**. (Similar to Topic 3, screen 21.)

6. Press ♦ [GRAPH], and then press F3 **Trace** for the bar chart (screen 3).

(3)



Note: *You could also plot the relative frequencies as in Topic 3, screen 28.*

Drawing Comparative Bar Charts

Example: Compare the New York City building data and the Chicago building data.

1. From the Stats/List Editor, enter the values **1.4**, **2.4**, and **3.4** in **list3** and **46**, **13**, and **13** in **list4** (screen 4).

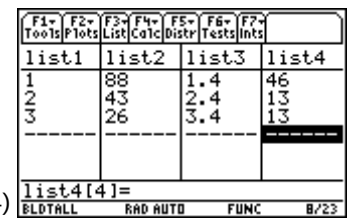
2. Change only one value in the **Plot 1** setup from above with Hist. Bucket Width: **0.2**.

3. Set up and define **Plot 2** as a histogram with X List: **list3**, Hist. Bucket Width: **0.2**, and Freq: **list4**.

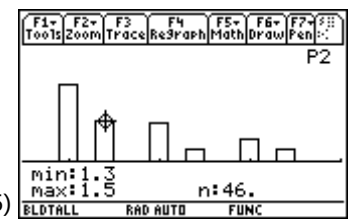
4. With both **Plot 1** and **Plot 2** selected, press ♦ [GRAPH], F3 **Trace**, and the arrow keys (screen 5).

   New York City has more buildings in each height category, with twice as many in the first and third categories. There are three times as many buildings in New York City in the middle, or 600 to 699 feet category, as compared to Chicago.

(4)



(5)



Note: *This chart can be extended to include more cities.*

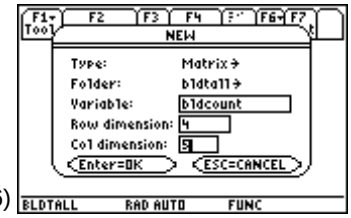
Relative Frequencies Using Matrices

Matrices are good ways of storing and calculating data in two-way tables. Create a matrix named **bldcount**:

1. Press APPS, **6:Data/Matrix Editor**, **3:New** for the NEW screen.

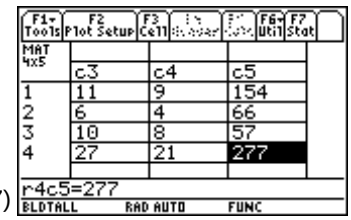
2. Set Type: **Matrix**, Folder: **bldtall**, Variable: **bldcount**, Row dimension: **4** (three height categories and a total), and Col dimension: **5** (four cities and a total). See screen 6.

(6)



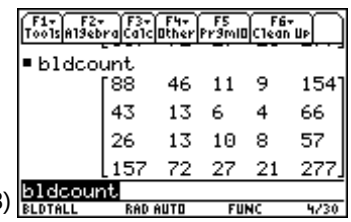
3. Press ENTER ENTER for the Stats/List Editor. Enter the tall building data in this topic's example by row; after each value is entered press ENTER (screen 7).

(7)



4. From the Home screen, enter **bldcount** and press ENTER (screen 8).

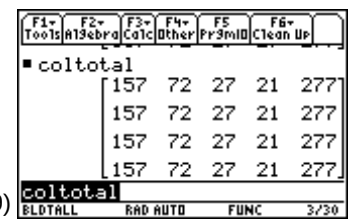
(8)



Note: *This matrix fits on the Home screen. See Topic 9, screen 7 on the procedure for moving around a larger matrix on the Home screen (scrolling up or down, left or right).*

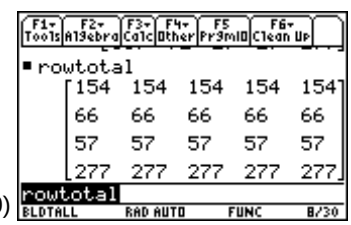
5. Repeat steps 1 through 4 to create a 4 by 5 **coltotal** matrix, with each row representing the lower marginal totals (screen 9).

(9)



6. Create a 4 by 5 **rowtotal** matrix, with each column representing the right margin totals (screen 10).

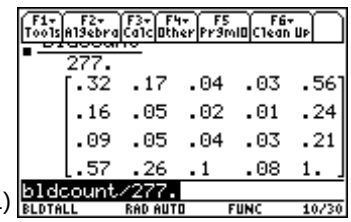
(10)



Relative Frequencies (Joint and Marginal)

1. Press MODE and change the third value, **Display Digits**, to **G:FLOAT 2**. This is done so the screen will contain all the values without scrolling. For a larger matrix or to show more decimals, you can scroll as explained in Topic 9, screen 7.
2. From the Home screen, enter **bldcount** ÷ **277.0**, and then press ENTER (screen 11).

(11)



From screen 11, the *joint relative frequency* of about 17% (46) of all the buildings being considered (277) are in Chicago and are between 500 and 599 feet tall.

The *marginal relative frequency* of about 26% (72) of all the buildings being considered (277) are in Chicago, and about 56% (154) are between 500 and 599 feet tall.

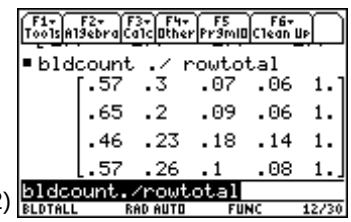
Notice that New York City has over half (57%) of all the tall buildings being considered.

Conditional Relative Frequencies

1. Enter **bldcount** . ÷ **rowtotal** and press ◆ ENTER (screen 12).

   Given the condition of the tallest buildings being considered (57 that are ≥ 700 ft.), 46% are in New York City, 23% in Chicago, 18% in Houston, and 14% in Los Angeles for a total of 100% = 1.

(12)



Note: . ÷ *divides each element in the first matrix by the corresponding element in the second matrix. In this case, each value in a row is divided by that row total.*

.46 + .23 + .18 + .14 = 1.01 *because of rounding to two decimals.*

2. Enter **bldcount** . ÷ **coltotal** and press ◆ ENTER (screen 13).

(13)