EXPLORATIONS

Activity 4

Describing Categorical Data

Topic 17 looks at the distribution of the population of the United States by the categorical variable race for one age group. Percentages are calculated, and visual comparison is made with a bar chart.

Topic 18 extends the discussion to a two-way table of race by age group. Age by itself is a quantitative variable, but ages are grouped into different categories (youngest, older, ...oldest) to have a manageable table. Percents and bar charts are compared.

Topic 17—Bar Charts

📖 *For this topic, set the mode for two decimal places. (Refer to "Setting Modes" under Do This First.)*

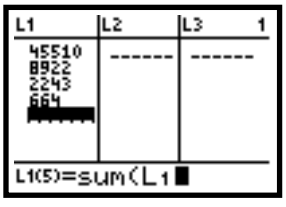
Population data, in thousands, for those under 15 years old follows.

Race:	**White**	**Black**	**Asian Pacific Islander**	**Am. Indian Eskimo, Aleut**
(thousands):	45510	8922	2243	664

(Source: Population by Age and Race 1994, Statistical Abstract of US 1995.)

1. Put the data in the spreadsheet in **L_1, as shown in** screen 1.
2. With the fifth row highlighted, press 2nd [LIST] **<MATH> 5:sum(**L_1, as shown in the bottom line of screen 1.

(1)



Activity 4, Describing Categorical Data (cont.)

3. Press [ENTER] for the first column of screen 2 with the sum of the first four rows showing in the fifth row.

4. Highlight **L2** as shown in screen 2, and type **L1** [÷] **L1** [(] **5** [)] [×] **100** in the bottom line.

   Press [ENTER] for the second column in screen 3, which lists the percent of those under 15 years old by race (**79.37**, **15.56**, **3.91**, and **1.16** for a total of **100** percent).

5. Enter **1**, **2**, **3**, **4**, and **9** in **L3**, as shown in screen 3.

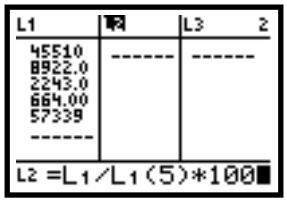
To get a visual comparison of the difference in population for each race, you will construct a bar chart. This is easy to do by hand and not difficult with the TI-83. The **Histogram** plot of the TI-83 is used, so you might want to look at Topic 2.

Screen 4 sets up for a **Histogram** with the **Xlist** of **L3**, which contains the values 1, 2, 3, 4, and 9 for reasons explained in the next paragraph. The **Ylist** is **L2**, which contains the percent of each race.

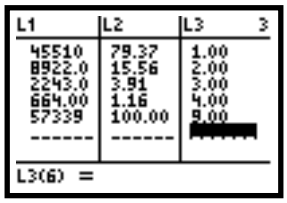
Note that the **WINDOW** in screen 5 has **Xmax = 4.5** and **Xscl = 0.5**. The first bar starts with the first $x = 1$ and has a width of 0.5, so there is a gap between bars with the second bar starting at $x = 2$. The fifth value in **L2** has the total of 100 percent and $x = 9$ in **L3**, which is beyond **Xmax = 4.5**. Thus, a bar for the total will not be on the screen.

The bar chart in screen 6 clearly shows the large majority of whites in the population of those under 15 years old. Bar charts are often given in order of largest magnitude to smallest. The plot compares percents, but it has the exact proportions and shape of one that compares population counts. The percents add up to 100 percent. If population counts were used, they would add up to the total population.

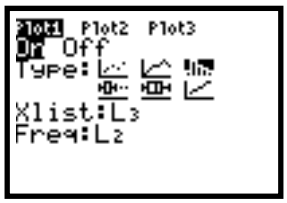
(2)



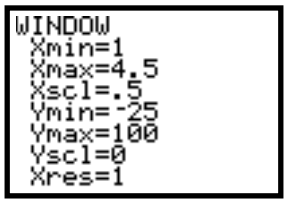
(3)



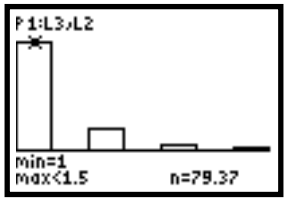
(4)



(5)



(6)



Topic 18—Two-Way Table

A more complete table of the data used in Topic 17 of the U.S. population of age by race, including more age groups, is given below. The population counts are in thousands.

Years:	White	Black	Asian Pacific Islander	Am. Indian Eskimo, Aleut
0 - 14	45510	8922	2243	664
15 - 29	44350	8020	2192	558
30 - 44	53045	7878	2400	512
45 - 59	35040	4190	1295	282
60 - 74	25405	2584	665	140
≥75	13120	1080	193	53

If you are only interested in a particular row or column of data, then put that data in **L1** and proceed as shown in Topic 17. If you wish to make many comparisons, then proceed as follows.

1. Set the mode to two decimals as shown in Do This First at the beginning of this handbook.
2. Put the four columns of data (of six rows each) in lists **L1** to **L4** (screen 7), and then calculate the column sums as follows.
   a. From the home screen, press [2nd] [LIST] **<MATH> 5:sum(L1** [STO▶] L1 [(] 7 [)]. This takes the sum of the values in **L1** (six values) and stores the sum in the seventh row of **L1**.
   b. Use [2nd] [ENTRY] and edit the list used to calculate the sum for each list, as shown in screen 7.
3. Use **L1** [+] **L2** [+] **L3** [+] **L4** [STO▶] **L5** [ENTER] to store the sum of each row (age group) in **L5**. (See screen 8.)

   The complete table of data, including the row and column totals, is shown in screens 9 and 10 and stored in lists **L1** to **L5**. You will also want to save this data in a matrix (see the next section).

(7)
```
sum(L1→L1(7)
          216470.00
sum(L2→L2(7)
           32674.00
sum(L3→L3(7)
            8988.00
sum(L4→L4(7)
```

(8)
```
                2209.00
L1+L2+L3+L4→L5
{57339.00 55120…
```

Note: *If your data source gave totals, you could have entered them. Because of rounding, they may have been a bit different from the totals obtained here.*

(9)

L1	L2	L3 1
45510	8922.0	2243.0
44350	8020.0	2192.0
53045	7878.0	2400.0
35040	4190.0	1295.0
25405	2584.0	665.00
13120	1080.0	193.00
216470	32674	8988.0

L1(1)=45510

(10)

L4	L5 4
57339	57339
558.00	55120
512.00	63835
282.00	40807
140.00	28794
53.00	14446
2209.0	260341

54

Saving Data in Matrix [A]

To save the original count data, including totals, so you can make calculations that change the values in the spreadsheet, proceed as follows.

1. Press [2nd] [LIST] **<OPS> 0:List▸matr(** **L1** [,] **L2** [,] **L3** [,] **L4** [,] **L5** [,] **[A]**, where **[A]** is pasted by pressing [MATRX] **<NAMES> 1:[A]** (see screen 11).

2. Press [ENTER] for **Done**, and then press [MATRX] **<NAME> 1:[A]** to paste **[A]** to the home screen, as shown in the last line of screen 11.

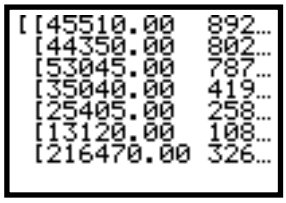
   After you paste **[A]** and press [ENTER], part of matrix **[A]** is revealed, as shown in screen 12.

3. Press the cursor control keys to reveal the rest of matrix **[A]**.

   The first column of matrix **[A]** is list **L1**, the second column, **L2**, and so on.

(11)

(12)



Calculating Column Percents

To calculate column percents, proceed as follows.

1. Take a list , and divide it by the seventh value (the total).

2. Multiply by 100.

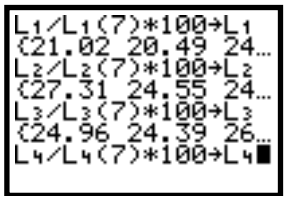
3. Store the resulting percents to the same list.

   This procedure is done for **L1** to **L5** in screens 13 and 14. The results are revealed in the spreadsheet in screens 15 and 16.

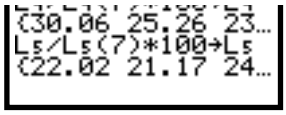
4. Edit list **L6**, and a newly created list named **LC7** to contain values to draw comparative bar graphs.

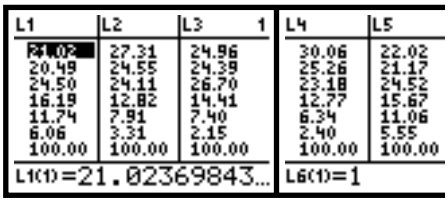
   The results are shown in screen 17.

(13)



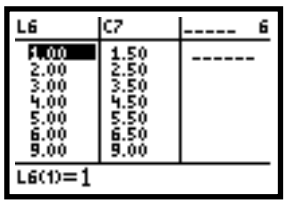
(14)





(15) (16)

(17)

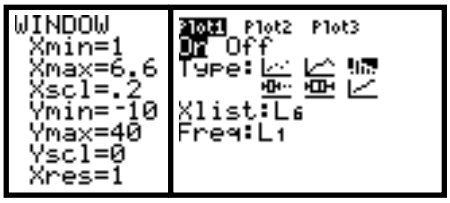


Drawing Comparative Bar Charts

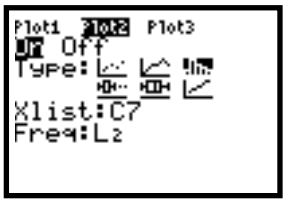
Before starting here, it is a good idea to first read Topic 17 about single bar charts.

1. To get a visual comparison of the age distributions for two groups, white and black, set up the **WINDOW** and **Plot1** and **Plot2** for two **Histograms**, as shown in screens 18, 19, and 20.

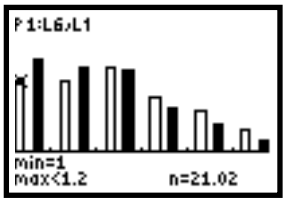
2. Press [TRACE] for the bar charts, as shown in screen 21.



(18) (19)



(20)



(21)

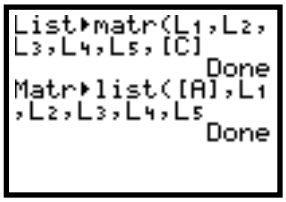
As you can see from the charts and from the previous table of column percents, a larger proportion of blacks are in the two youngest age groups than the proportion of whites in the same age groups. (The shading can be added by hand or in **Dot** mode with six commands: [2nd] [DRAW] **7:Shade(0 , L2 (1) (1.43 < X) (X < 1.57))**, and then **7:Shade(0 , L2 (2) (2.43 < X) (X< 2.57))**, and so on; where **<** is pasted from [2nd] [TEST] **5:<**.)

When making comparisons, do not confuse *percents* with *counts*. The fact that a larger percentage of blacks are in the youngest groups than the percentage of whites in the same groups does not mean that there are more blacks than whites in these groups.

Before you can continue making the comparative bar charts (described below), you must switch some data about as follows.

3. Save the column percents table in matrix **[C]** by pressing [2nd] [LIST] **<OPS> 0:List▸matr(L1 , L2 , L3 , L4 , L5 , [C]** [ENTER] for **Done**. (See screen 22.)

4. Return the original data, with totals, in matrix **[A]** to the spreadsheet by pressing [2nd] [LIST] **<OPS> A:Matr▸list([A] , L1 , L2 , L3 , L4 , L5** [ENTER] for **Done**. (See screen 22.)

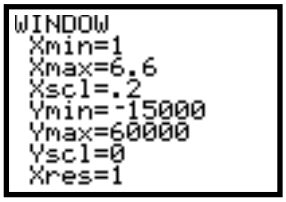


(22)

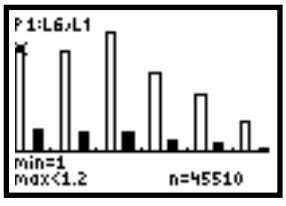
To continue making the comparative bar charts:

5. Compare the counts for each age group with the original count data as shown in column **L1** and **L2**. The plots are set up the same as the previous bar charts of percents, but **Ymin** and **Ymax** have been adjusted to include all the counts (see screen 23).

6. Press [TRACE], and add shading for the bar charts shown in screen 24.



(23)



(24)

You can see that there are many more whites than blacks in all age groups.

Activity 4, Describing Categorical Data (cont.)

Calculating the Percent of Each Cell

From the last value in **L5** in screen 10, you can see that the grand total of the population is **260,341** (thousand).

In the first age group (0 to 14 year olds), there are 45,510 (thousand) whites and 8,922 (thousand) blacks. This is about 17.5 percent of the whole population for whites in this age group, 100 ∗ (45510/260341) = **45510 ∗ (100/260341) = 17.48**, and about 3.4 percent for blacks in this age group, **8922 ∗ (100/260341) = 3.43**. These calculations are shown in screen 25.

(25)
```
45510*(100/26034
1)
                17.48
8922*(100/260341
)
                3.43
```

All the cells of the table (stored in **[A]**) can be multiplied by **100** and divided by **260,341**, as shown in screen 26. Note that the **17.48** percent and **3.43** percent you calculated above are in the first row of the resulting matrix (screen 27).

(26)
```
[A]*(100/260341)
```

(27)
```
[[17.48 3.43  .…
 [17.04 3.08  .…
 [20.38 3.03  .…
 [13.46 1.61  .…
 [9.76  .99   .…
 [5.04  .41   .…
 [83.15 12.55 3…
```

Calculating Row Percents

The example data for Topic 17 is really the first row of the two-way table. To calculate the percents with the data and totals in list **L1** to **L5**, proceed as follows.

1. Enter **L1** [×] **100** [÷] **L5** [STO▶] **L1** [ENTER] for the first two lines in screen 28.
2. Press [2nd] [ENTRY] to recall the last entry, press [◀] to move the flashing cursor over **L1,** and then change **L1** to **L2**.
3. Press [▲] to jump to the first **L1**, change this to **L2**, and then press [ENTER] for the third and fourth lines of screen 28.
4. Repeat the above steps for **L3**, **L4**, and **L5**, which are made up of the totals from the other rows, and thus all have values of 100 percent. (See the fourth line of screen 29.)

   You can observe the row percentages in **L1** to **L5** under [STAT] **1:Edit**.

(28)
```
L1*100/L5→L1
{79.37 80.46 83…
L2*100/L5→L2
{15.56 14.55 12…
L3*100/L5→L3
{3.91 3.98 3.76…
```

5. Store the results in matrix **[B]** as follows.

   a. Press 2nd [LIST] **<OPS> 0:List▸matr(L1 , L2 , L3 , L4 , L5 , [B]**, where **[B]** is pasted by pressing MATRX **<NAMES> 2:[B]** (see screen 29).

   b. Press ENTER for **Done**, and then press MATRX **<NAMES>2:[B]** to paste **[B]** to the home screen, as shown in the last line of screen 29.

   After you paste **[B]** and press ENTER, part of matrix [B] is revealed, as shown in screen 30.

6. Press the cursor control keys to reveal the rest of matrix **[B]**.

   The values in the first row (**79.37**, **15.56**, **3.91**, and **1.16** percent) are the same as shown in Topic 17.

(29)

```
L4*100/L5→L4
{1.16 1.01 .80 …
L5*100/L5→L5
{100.00 100.00 …
List▸matr(L1,L2,
L3,L4,L5,[B]
                Done
[B]
```

(30)

```
[[79.37 15.56 3…
 [80.46 14.55 3…
 [83.10 12.34 3…
 [85.87 10.27 3…
 [88.23 8.97  2…
 [90.82 7.48  1…
 [83.15 12.55 3…
```

Making Bar Charts for Rows

If you want the rows of percents to be set up in lists to construct bar charts with the TI-83, one solution is to store the transpose of matrix **[B]** in matrix **[E]** with the original rows of **[B]** now columns in **[E]**.

1. First, paste **[B]** to the home screen, and then press MATRX **<MATH> 2:ᵀ** STO▸ **[E]** for the first line in screen 31.

2. Press ENTER for matrix **[E]**, part of which is shown in screen 31.

   The first row of **[B]** represents the youngest age groups' percents now located in the first column of **[E]**.

3. Transfer these lists from **[E]** to the spreadsheet, and plot bar charts as shown previously in this topic.

(31)

```
[B]ᵀ→[E]
[[79.37  80.46 …
 [15.56  14.55 …
 [3.91   3.98  …
 [1.16   1.01  …
 [100.00 100.00…
```