

Unit 1 - Summarizing Data
Week 1 - Practice Problems

Solutions

#1.

- a. Qualitative - ordinal
- b. Qualitative - nominal
- c. Quantitative - ratio
- d. Qualitative - nominal
- e. Quantitative - ratio
- f. Quantitative - interval

#2a. By hand, here is the stem and leaf diagram I constructed. Other groupings for the stem are okay.

| Stem | Leaf |
|------|---------------------|
| 0 | 1 1 1 1 |
| 0 | 2 3 3 3 3 3 3 |
| 0 | 4 4 4 4 4 5 5 5 5 5 |
| 0 | 7 7 7 7 7 7 7 |
| 0 | 8 8 8 8 8 8 |
| 1 | 0 0 0 1 1 |
| 1 | 2 2 3 3 |
| 1 | |
| 1 | 7 7 |

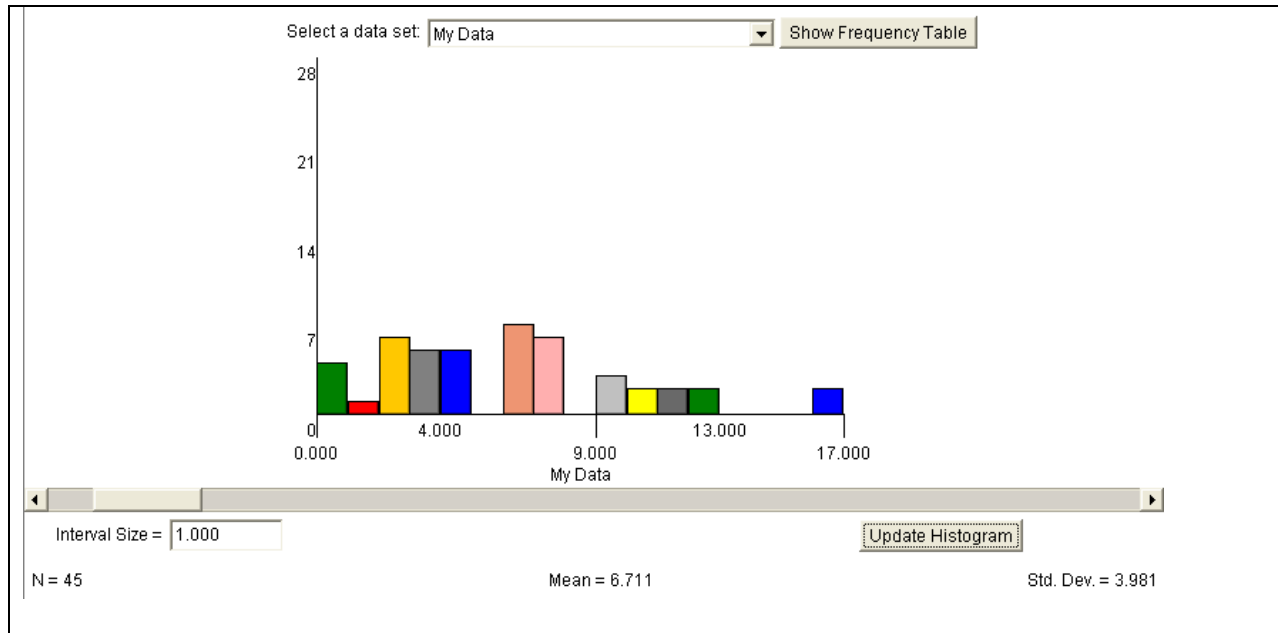
#2b. By hand, this is what I produced. Other class intervals are okay.

| Class Interval | Frequency | Relative Frequency | Cumulative Frequency | Cumulative Rel. Frequency |
|----------------|-----------|--------------------|----------------------|---------------------------|
| 0-1 | 4 | .0889 | 4 | .0889 |
| 2-3 | 7 | .1556 | 11 | .2444 |
| 4-5 | 10 | .2222 | 21 | .4667 |
| 6-7 | 7 | .1556 | 28 | .6222 |
| 8-9 | 6 | .1333 | 34 | .7556 |
| 10-11 | 5 | .1111 | 39 | .8667 |
| 12-13 | 4 | .0889 | 43 | .9556 |
| 14-15 | 0 | 0 | 43 | .9556 |
| 16-17 | 2 | .0444 | 45 | 1.0000 |
| TOTAL | 45 | 1.0000 | | |

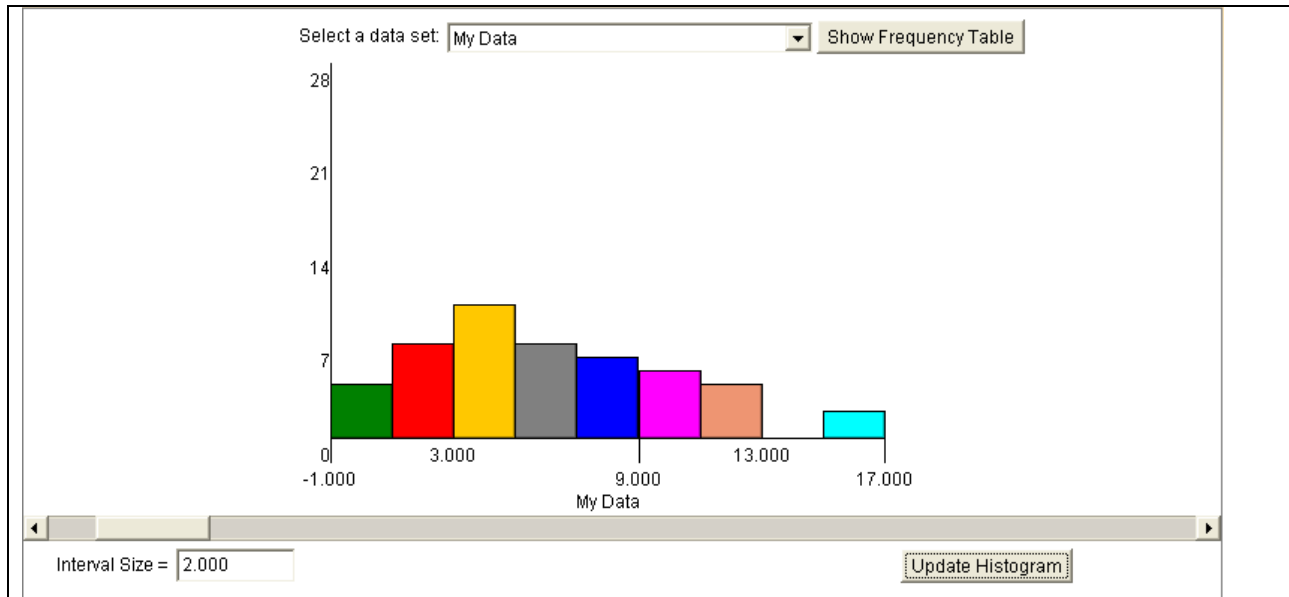
#2c. I did not construct a histogram by hand, I used the Shodor applet that can be found under

<http://www.shodor.org/interactivate/activities/Histogram/>

At the dialogue box “select a data set”, I scrolled down to choose **My Data**. Next, I scrolled down until I found a data entry box. I entered my data, selected **interval size = 1.00** and then clicked on **Update Histogram**. Here is what I got.



If you like, you can play with different choices of interval size. For example, interval size=2.00 yields the following.

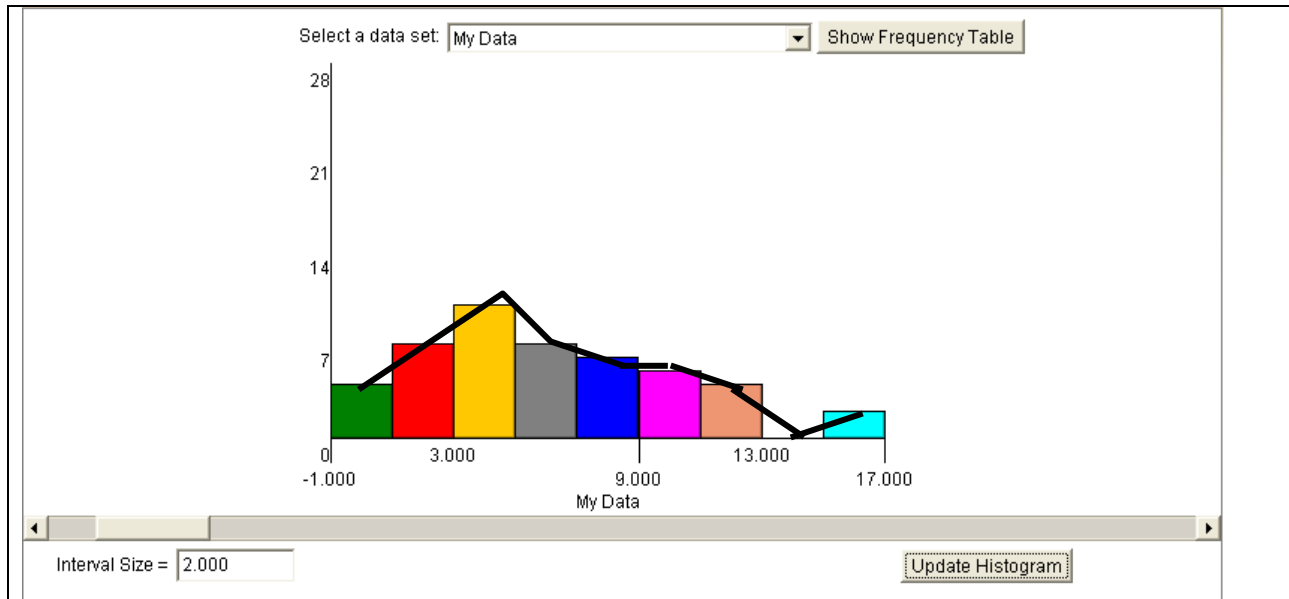


If you would like to try generating a histogram in SAS or Stata or Minitab visit the Summarizing Data topic page in the course resource website. Go to

<http://www-unix.oit.umass.edu/~biep540w/webpages/summarizing.htm>

#2d. I did not do this by hand, either. A frequency polygon plot is similar to a histogram. The first step is to choose class intervals. Next, note for each class interval the frequency (or relative frequency of data values in the interval). Plotted on the x-axis is the midpoint of the interval. Plotted on the y-axis is the frequency (or relative frequency).

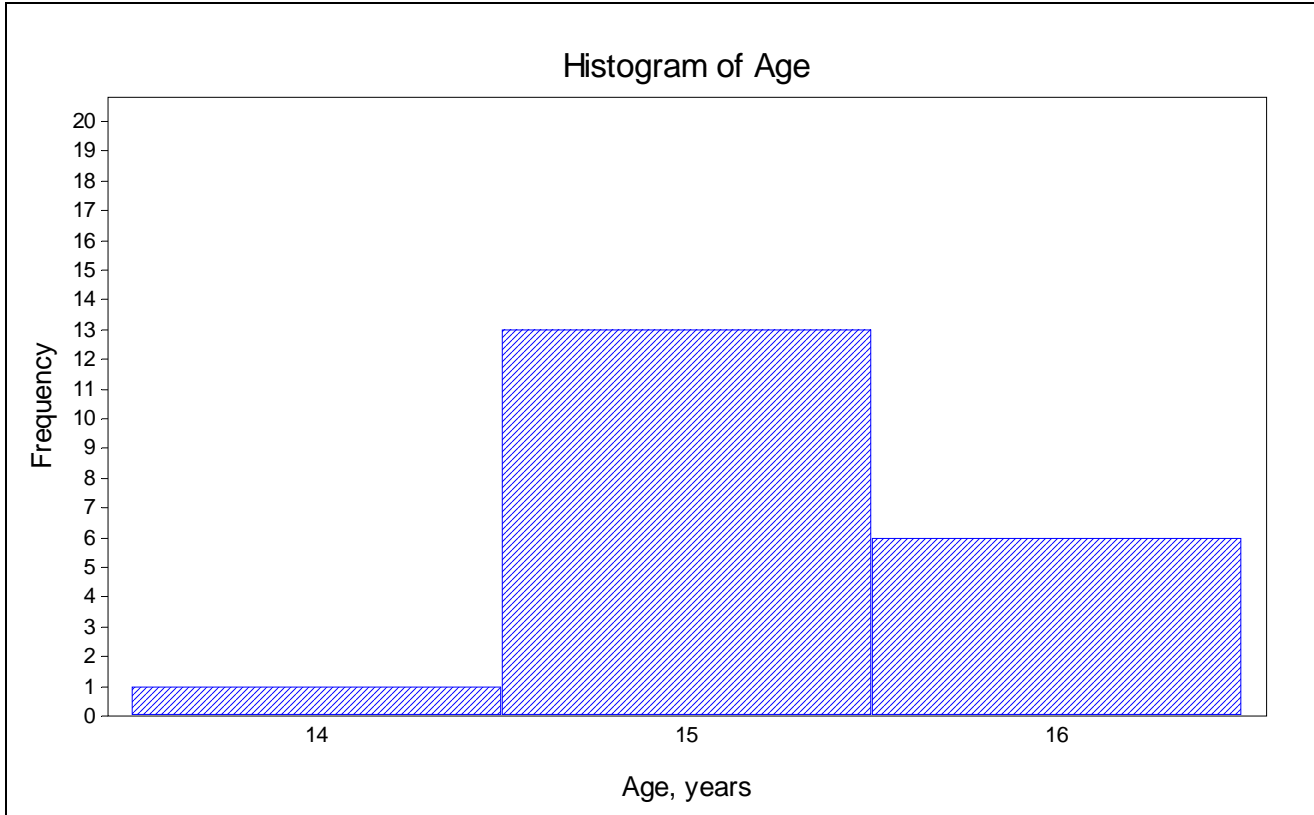
Thus, a frequency polygon can be appreciated as an overlay of the histogram (and therefore communicating the same summarization as the histogram). With a little bit of artistic license (MS word doesn't allow lots of precision), the frequency polygon is the graphed line below in bold black.



#3a.

| Age | Frequency | Relative Frequency | Cumulative Frequency | Cumulative Rel. Frequency |
|-------|-----------|--------------------|----------------------|---------------------------|
| 14 | 1 | .05 | 1 | .05 |
| 15 | 13 | .65 | 14 | .70 |
| 16 | 6 | .30 | 20 | 1.00 |
| TOTAL | 20 | 1.00 | | |

#3b.



#3c. Males tend to be taller than females

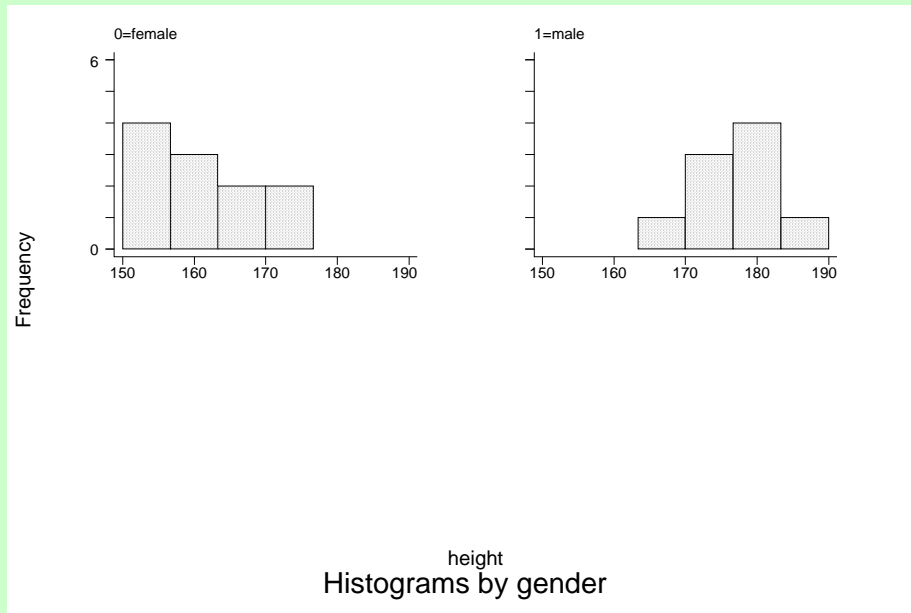
| Females | Stem | Males |
|---------|------|-------|
| 1 3 4 | 15 | |
| 6 9 | 15 | |
| 0 3 | 16 | |
| 6 7 | 16 | 7 |
| 0 1 | 17 | 4 3 3 |
| | 17 | 8 7 |
| | 18 | 3 3 |
| | 18 | 5 |

#3d.

| Class Interval | <u>FEMALES</u> | | <u>MALES</u> | |
|----------------|----------------|-----------|--------------|------------|
| | Freq. | Re. Freq. | Freq. | Rel. Freq. |
| 150-159 | 5 | .45 | 0 | 0 |
| 160-169 | 4 | .36 | 1 | .11 |
| 170-179 | 2 | .18 | 5 | .56 |
| 180-189 | 0 | 0 | 3 | .33 |

I then used STATA version 7 to produce the histogram. You can use whatever tool you like (by hand, SAS, SPSS, Minitab)

```
. graph height, by(gender) bin(6) freq ytick(0,1,2,3,4,5,6) xlabel
```



#4a.

$$\begin{aligned}(X_1 + X_2 + X_3 + X_4)^2 &= \left[\sum_{i=1}^4 X_i \right]^2 \\ &= (3 + 1 + 4 + 6)^2 \\ &= 14^2 \\ &= 196.\end{aligned}$$

#4b.

$$\begin{aligned}X_1^2 + X_2^2 + X_3^2 + X_4^2 &= \sum_{i=1}^4 X_i^2 \\ &= 3^2 + 1^2 + 4^2 + 6^2 \\ &= 9 + 1 + 16 + 36 \\ &= 62.\end{aligned}$$

#4c.

$$\begin{aligned}\sum_{i=1}^4 (X_i - 1)^2 &= (3-1)^2 + (1-1)^2 + (4-1)^2 + (6-1)^2 \\ &= 2^2 + 0^2 + 3^2 + 5^2 \\ &= 4 + 0 + 9 + 25 \\ &= 38.\end{aligned}$$

Note:

$$\begin{aligned}\sum_{i=1}^4 (X_i - 1)^2 &= \sum_{i=1}^4 [X_i^2 - 2X_i + 1] \\ &= \sum_{i=1}^4 X_i^2 - 2 \sum_{i=1}^4 X_i + 1 \sum_{i=1}^4 1 \\ &= 62 - (2)(14) + (1)(4) \\ &= 38.\end{aligned}$$

#4d.

$$\begin{aligned}\sum_{i=1}^4 3X_i &= 3 \sum_{i=1}^4 X_i \\ &= 3(14) \\ &= 42\end{aligned}$$
