**Variable**
A type of information or property about the thing being observed

size, color, weight, age, gender, temperature, …

**Categorical Data**
Data which falls into categories

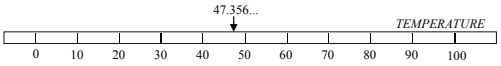
red, male, Republican, rich, Yankee fan,...

**Numerical Data**
Data values which are the result of measurements

temperature, income, weight, GPA, cost

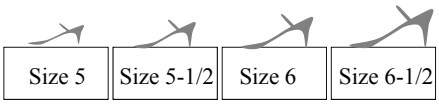
**(Numerical) Continuous Data**
Data which can take on any value between two numbers

height, temperature, annual salary, blood pressure, ...




**(Numerical) Discrete Data**
Data which can take on only certain values

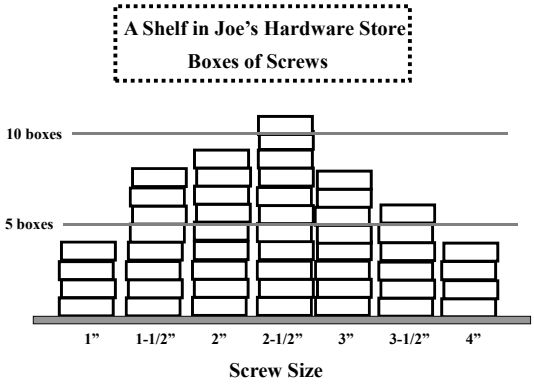
shoe size, number of children in a family, ...








**Distribution**
A presentation of data along with the number of times each data value occurs.

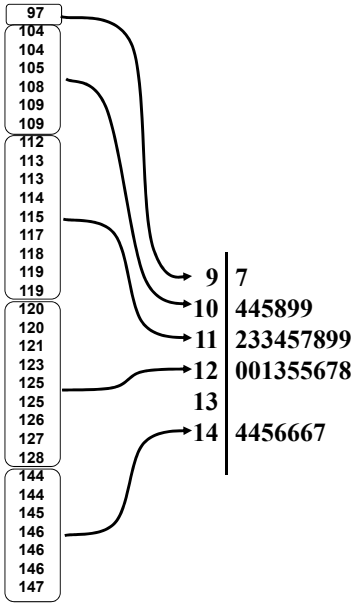







**Stem and Leaf Display**
A form of data presentation which uses the data to visually display how the data are distributed.

**Blood Pressures of a Sample of Men sorted Low to High)**








**Four Levels of Measurement**

**Nominal (Qualitative Data)**
Names or categories
Alabama, Georgia, California...

**Ordinal (Qualitative or Quantitative)**
A category which has an inherent order
Small, Medium, Large, Extra Large, Jumbo

**Interval (Quantitative)**
A category in which only the value of the difference between the numbers has a meaning.
Temperature Measures (100°F is 50F° warmer than 50°F, but it is not "twice as warm" because 0°F is arbitrary.
Dress Sizes: A 12 is not twice as large as a 6.

**Ratio (Quantitative)**
A category in which both the interval and the ratio have meaning.
Lengths, heights, weights, etc. (A 16 oz. hamburger weighs 8 ozs. more than an 8 oz. hamburger, and is also twice as heavy as an 8 oz hamburger.

# Students Enrolled in Various College Programs

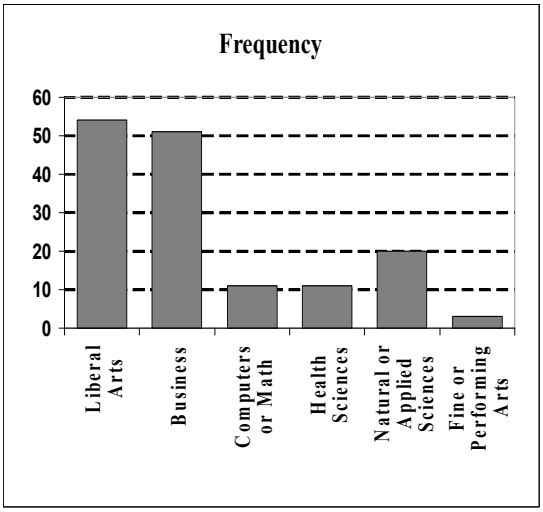
**Frequency Distribution Table for Categorical Data**

| Classes | Frequency |
|---|---|
| Liberal Arts | 54 |
| Business | 51 |
| Computers or Math | 11 |
| Health Sciences | 11 |
| Natural or Applied Sciences | 20 |
| Fine or Performing Arts | 3 |





# Students Enrolled in Various College Programs (Displayed as a Bar Graph)








# Steps in Creating a Stem and Leaf Plot

**Chapter 2 - Organizing and Presenting Data**

**Cholesterol Levels (Mg Cholest/100ml)**

**Stem-and-Leaf Display**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 263 | 258 | 240 | 233 | 225 | 222 | 199 | 282 | 232 |
| 283 | 200 | 212 | 225 | 235 | 240 | 258 | 263 | 259 |
| 241 | 237 | 226 | 213 | 269 | 199 | 253 | 201 | 238 |
| 242 | 259 | 233 | 238 | 229 | 215 | 202 | 319 | 239 |
| 243 | 248 | 245 | 219 | 276 | 246 | | | |



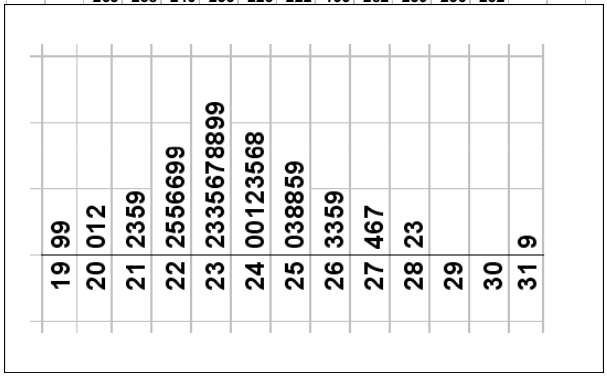


# Steps in Creating a Stem and Leaf Plot

**Cholesterol Levels (Mg Cholest/100ml)**

**Stem-and-Leaf Display**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 263 | 258 | 240 | 233 | 225 | 222 | 199 | 282 | 239 | 236 | 232 | |



# Steps in Creating a Stem and Leaf Plot

**Cholesterol Levels (Mg Cholest/100ml)**

**Stem-and-Leaf Display**

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 263 | 258 | 240 | 233 | 225 | 222 | 199 | 282 | 239 | 236 | 232 |
| 283 | 200 | 212 | 225 | 235 | 240 | 258 | 263 | 274 | 250 | 259 |
| 241 | 237 | 226 | 213 | 269 | 199 | 253 | 201 | 265 | 226 | 238 |
| 242 | 259 | 233 | 238 | 229 | 215 | 202 | 319 | 277 | 229 | 239 |
| 243 | 248 | 245 | 219 | 276 | 246 | | | | | |

**Sort the Data in Descending Order**

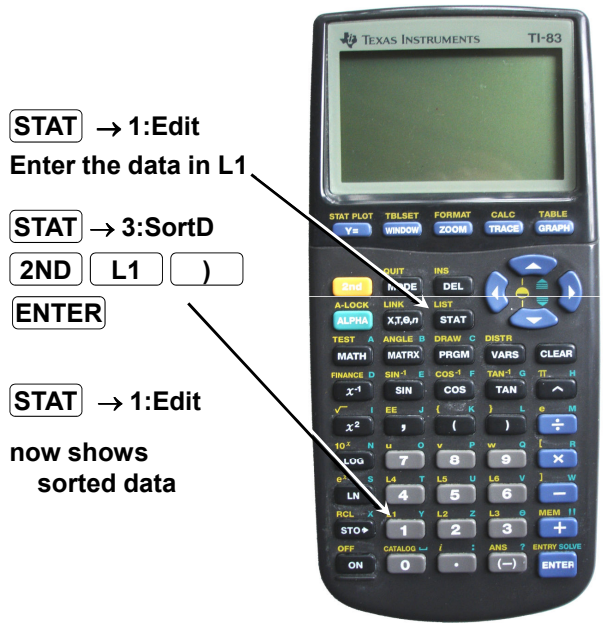
| | |
|---|---|
| 19 | 99 |
| 20 | 012 |
| 21 | 2359 |
| 22 | 2556699 |
| 23 | 2335678899 |
| 24 | 00123568 |
| 25 | 038859 |
| 26 | 3359 |
| 27 | 467 |
| 28 | 23 |
| 29 | |
| 30 | |
| 31 | 9 |

199 225 236 243 263
199 225 237 245 263
200 226 238 246 265
201 226 238 248 269
202 229 239 250 274
212 229 239 253 276
213 232 240 258 277
215 233 240 258 282
219 233 241 259 283
222 235 242 259 319





# Using the Calculator to Help Create a Stem and Leaf Plot








# Using a Back-to-Back Stem-and-Leaf Plot to Compare Groups of Data

**Cholesterol Levels of Men**

**Regular Diet vs. Low Fat Diet**

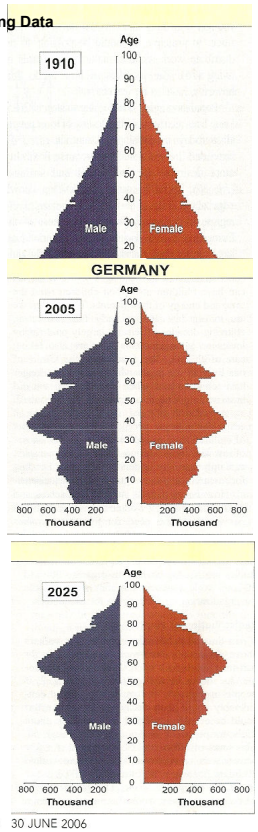
**Back-to-Back Stem-and-Leaf Display**

| Regular Diet | | Low Fat Diet |
|---|---|---|
| | 16 | 158 |
| | 17 | 122567 |
| | 18 | 0144899 |
| 99 | 19 | 003355789 |
| 210 | 20 | 000255999 |
| 9532 | 21 | 01114789 |
| 9966552 | 22 | 008 |
| 988765332 | 23 | 01 |
| 86532100 | 24 | |
| 998830 | 25 | |
| 533 | 26 | |
| 764 | 27 | |
| 32 | 28 | |
| | 29 | |
| | 30 | |
| 9 | 31 | |





Stem and Leaf plots for German population in 1910, 2005 and 2025. The data for 1910 resemble a pyramid with many children and few elderly people, but by 2005 there was a bulge of adults around age 40. This bulge will rise to age 60 in 2025.



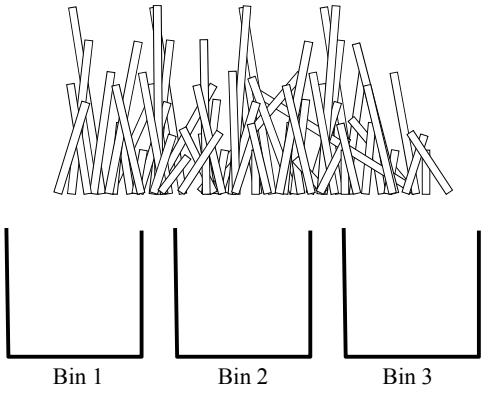

www.sciencemag.org SCIENCE VOL 312 30 JUNE 2006

## Class Definitions

**You go to work in a lumber yard.**

**On your first day, you find piles of wood of many lengths and need to separate them into 3 bins.**

**Longest piece = 20 feet  Shortest piece = 3 feet**



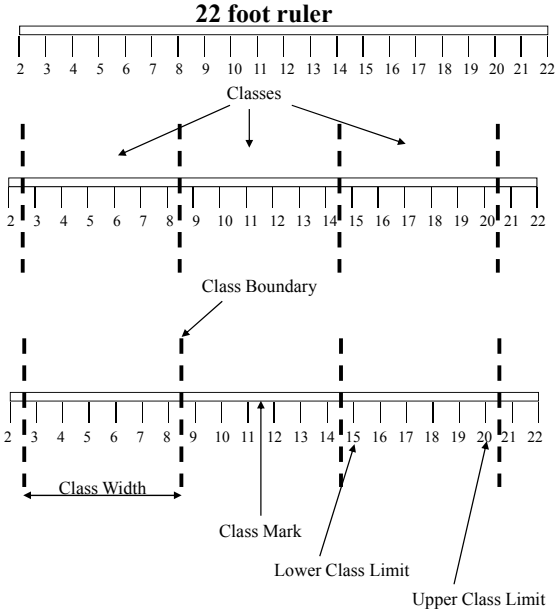

***Which lengths should go into each bin?***





## Class Definitions

**Problem: sort lumber lengths between 3 feet and 20 feet into three equal length categories (classes).**








**From DeSanto Pg 43 Section 2.4**

**Table: Study Hours per Week Reported by Students**

**Frequency Distribution Table for Numerical Data**

| Stdy Hrs | Freq | Stdy Hrs | Freq | Stdy Hrs | Freq | Stdy Hrs | Freq |
|---|---|---|---|---|---|---|---|
| 6 | 1 | 14 | 2 | 22 | 3 | 30 | 4 |
| 7 | 3 | 15 | 9 | 23 | 5 | 31 | 3 |
| 8 | 7 | 16 | 6 | 24 | 9 | 32 | 1 |
| 9 | 7 | 17 | 8 | 25 | 4 | 33 | 3 |
| 10 | 10 | 18 | 6 | 26 | 4 | 34 | 0 |
| 11 | 6 | 19 | 5 | 27 | 7 | 35 | 0 |
| 12 | 9 | 20 | 12 | 28 | 7 | 36 | 1 |
| 13 | 0 | 21 | 5 | 29 | 3 | | |





**Constructing a Frequency Distribution Table**

**Frequency Distribution Table for Numerical Data**

| Stdy Hrs | Freq | Stdy Hrs | Freq | Stdy Hrs | Freq | Stdy Hrs | Freq |
|---|---|---|---|---|---|---|---|
| 6 | 1 | 14 | 2 | 22 | 3 | 30 | 4 |
| 7 | 3 | 15 | 9 | 23 | 5 | 31 | 3 |
| 8 | 7 | 16 | 6 | 24 | 9 | 32 | 1 |
| 9 | 7 | 17 | 8 | 25 | 4 | 33 | 3 |
| 10 | 10 | 18 | 6 | 26 | 4 | 34 | 0 |
| 11 | 6 | 19 | 5 | 27 | 7 | 35 | 0 |
| 12 | 9 | 20 | 12 | 28 | 7 | 36 | 1 |
| 13 | 0 | 21 | 5 | 29 | 3 | | |

**Constructing a Frequency Distribution Table**

**Frequency Distribution Table for Numerical Data**

| Stdy Hrs | Freq | Stdy Hrs | Freq | Stdy Hrs | Freq | Stdy Hrs | Freq |
|---|---|---|---|---|---|---|---|
| 6 | 1 | 14 | 2 | 22 | 3 | 30 | 4 |
| 7 | 3 | 15 | 9 | 23 | 5 | 31 | 3 |
| 8 | 7 | 16 | 6 | 24 | 9 | 32 | 1 |
| 9 | 7 | 17 | 8 | 25 | 4 | 33 | 3 |
| 10 | 10 | 18 | 6 | 26 | 4 | 34 | 0 |
| 11 | 6 | 19 | 5 | 27 | 7 | 35 | 0 |
| 12 | 9 | 20 | 12 | 28 | 7 | 36 | 1 |
| 13 | 0 | 21 | 5 | 29 | 3 | | |

*Reconstruct above table for 7 categories*

$$\text{approximate class width} = \frac{\text{largest value - smallest value}}{\text{number of classes}}$$

$$= \frac{36-6}{7} = \frac{30}{7} = 4.29 \rightarrow 5$$

*make the class width = 5*

**"New" Frequency Distribution Table for Numerical Data**

| STDY HRS | FREQ |
|---|---|
| 6-10 | 28 |
| 11-15 | 26 |
| 16-20 | 37 |
| 21-25 | 26 |
| 26-30 | 25 |
| 31-35 | 7 |
| 36-40 | 1 |





**Class Boundaries and Class Marks**

**"New" Frequency Distribution Table for Numerical Data**

| STDY HRS | FREQ |
|---|---|
| 6-10 | 28 |
| 11-15 | 26 |
| 16-20 | 37 |
| 21-25 | 26 |
| 26-30 | 25 |
| 31-35 | 7 |
| 36-40 | 1 |

$$\text{Class Boundary} = \frac{\text{Upper Class Limit of one class + Lower Class Limit of next class}}{2}$$

$$\text{Class Mark} = \frac{\text{upper class limit + lower class limit}}{2}$$

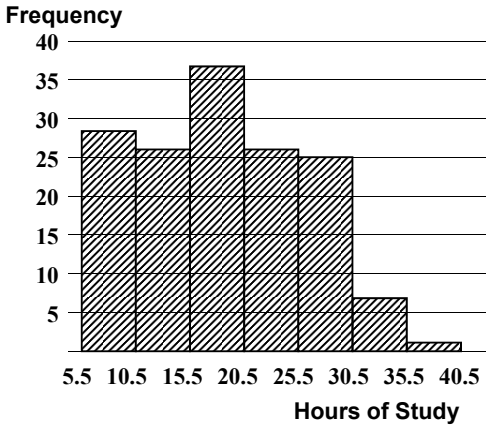
| STDY HRS | CLASS BOUNDARIES | CLASS MARK | FREQ |
|---|---|---|---|
| 6-10 | 5.5 - 10.5 | 8 | 28 |
| 11-15 | 10.5 - 15.5 | 13 | 26 |
| 16-20 | 15.5 - 20.5 | 18 | 37 |
| 21-25 | 20.5 - 25.5 | 23 | 26 |
| 26-30 | 25.5 -30.5 | 28 | 25 |
| 31-35 | 30.5 - 35.5 | 33 | 7 |
| 36-40 | 35.5 - 40.5 | 38 | 1 |



**2.5: Bar Graphs, Histogram, Frequency Polygon and Ogive**

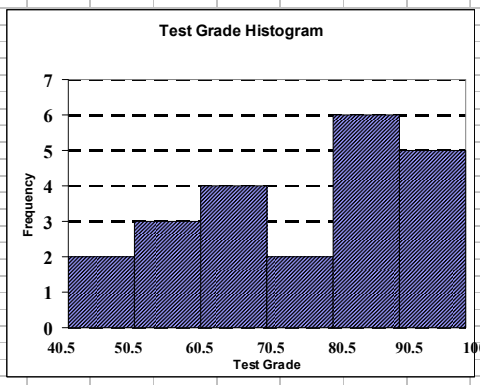
## A HISTOGRAM OF STUDENT STUDY HOURS

| Study Hrs | Class Boundaries | Class Mark | Freq |
|---|---|---|---|
| 6 – 10 | 5.5 -10.5 | 8 | 28 |
| 11 - 15 | 10.5 – 15.5 | 13 | 26 |
| 16 – 20 | 15.5 – 20.5 | 18 | 37 |
| 21 – 25 | 20.5 – 25.5 | 23 | 26 |
| 26 – 31 | 25.5 – 30.5 | 28 | 25 |
| 31 – 35 | 30.5 – 35.5 | 33 | 7 |
| 36 – 40 | 35.5 – 40.5 | 38 | 1 |






**Chapter 2: Organizing and Presenting Data**




a) How many test grades are represented in the histogram?
b) How many test grades are greater than 70?
c) How many test grades are between 61 and 80 (inclusive)?
d) How many test grades are less than 81?
e) Which class has a frequency of 4?
f) Which classes have the lowest frequency?

**Construct a Histogram (Pg 75 / Prob 20)**

**Ages of members visiting a health club on a Friday night**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 19 | 31 | 52 | 34 | 84 | 63 | 52 | 37 | 24 | 29 |
| 33 | 46 | 19 | 32 | 41 | 49 | 26 | 32 | 46 | 44 |
| 28 | 76 | 49 | 34 | 73 | 63 | 56 | 35 | 74 | 66 |
| 59 | 39 | 61 | 50 | 37 | 29 | 30 | 51 | 54 | 41 |

**Decide on the number of classes in the table (...say 5)**

**Find the largest and smallest data values (... 19 and 84)**

**Calculate an initial estimate of class width ( ... (84-19)/5 = 13**

| CLASS | | |
|---|---|---|
| 19 | - | 31 |
| 32 | - | 44 |
| 45 | - | 57 |
| 58 | - | 70 |
| 71 | - | 83 |

**...which leaves out the value 84**

**So what do we do now?**

**Increase the class width to 14 ...or...**

**Increase the number of classes to 6?**

**Let's make the number of classes = 6**

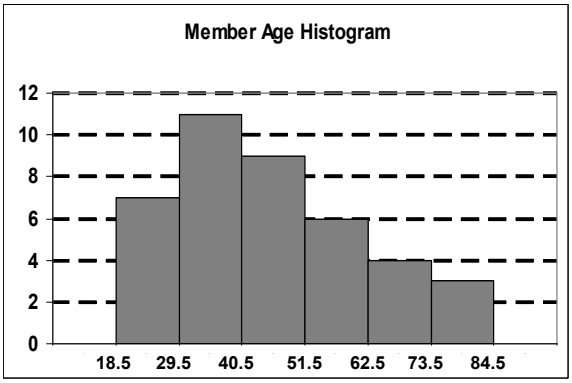
**Class Width = (84 -19)/6 = 10.8 =11**

| CLASS | FREQ |
|---|---|
| 19 - 29 | 7 |
| 30 – 40 | 11 |
| 41 – 51 | 9 |
| 52 – 62 | 6 |
| 63 – 73 | 4 |
| 74 - 84 | 3 |





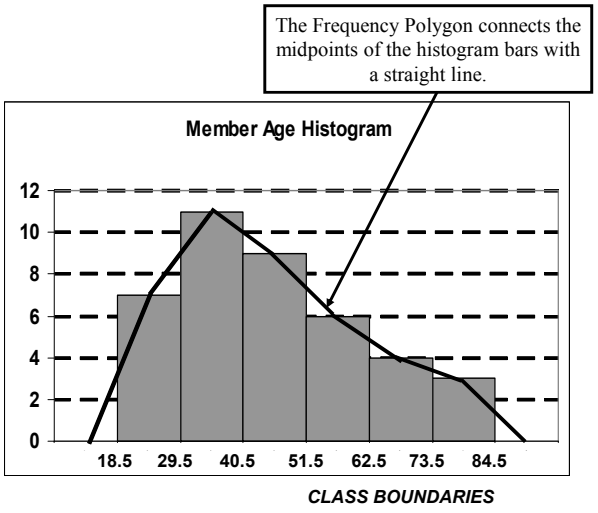
## Member Age Histogram
## Recalculated Using 6 Classes








## Member Age Histogram
## Recalculated Using 6 Classes



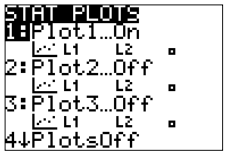





## Using the Calculator to Create a Histogram

**STAT** → **1:Edit**

**Enter the data in L1**



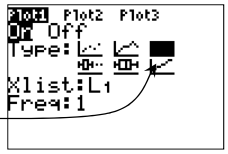

**2ND** → **STAT PLOT**

**Select 1: Plot 1**

**Place cursor on "ON"**

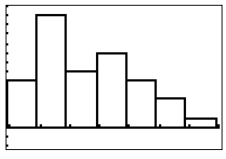
**then ENTER and**

**select histogram icon**

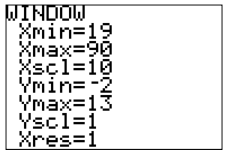



**ZOOM** → **9: ZoomStat**

**(The calculator has selected a scale to fit the histogram in its window)**



**To set a scale of your own choosing, select** **WINDOW**




**Xmin = (lower boundary desired)**

**Xmax= (larger than your largest data value)**

**Xscl=(Set to desired class width)**

**Ymin=(Set to 0)**

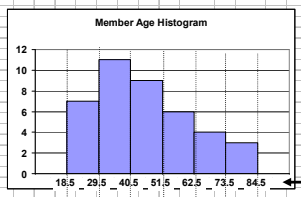
**Ymax=(set high enough to display height of largest box)**

**Yscl=(sets y-axis marks)**

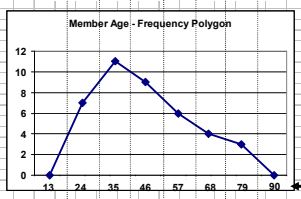
**Xres=(set to 1 for best resolution)**

**Comparison: Histogram, Frequency Polygon and Ogive**

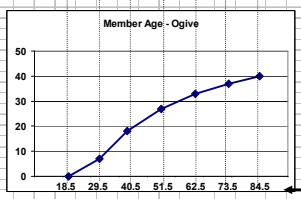
| LOWER CLASS BNDRY | CLASS | | UPPER CLASS BNDRY | CLASS MARK | FREQ |
|---|---|---|---|---|---|
| 18.5 | 19 | 29 | 29.5 | 24 | 7 |
| 29.5 | 30 | 40 | 40.5 | 35 | 11 |
| 40.5 | 41 | 51 | 51.5 | 46 | 9 |
| 51.5 | 52 | 62 | 62.5 | 57 | 6 |
| 62.5 | 63 | 73 | 73.5 | 68 | 4 |
| 73.5 | 74 | 84 | 84.5 | 79 | 3 |




*Class Boundaries*




*Class Marks*




*Class Boundaries*