

# Tabulating and Depicting Data

James H. Steiger

## Key Concepts for Today

- *Shape* of the distribution of a data set is often important for understanding key aspects of the data-generation process.

In this section, we discuss:

- Several methods for displaying distributions
- Frequency tables, and several terms associated with them

## Creating a Frequency Distribution Table

---

- # Suppose I just gave a statistics exam to a class of 50 people, and got the following list of grades:

79, 55, 79, 56, 83, 74, 77, 46, 84, 68,  
77, 84, 80, 62, 75, 64, 63, 78, 80, 88,  
76, 75, 80, 88, 71, 75, 85, 75, 73, 66,  
79, 46, 61, 59, 80, 81, 76, 80, 84, 62,  
61, 72, 77, 73, 82, 77, 89, 68, 78, 73

$N = 50$

---

## Creating a Frequency Distribution Table

---

- # A “grouped frequency distribution display” breaks the range of the data into about 10 equal intervals, and counts the number of scores in each interval.
  - # For example, with the present data, the scores range from 46 to 89, so we have 10 intervals 4 units wide. We then tabulate the data like this (usually with a computer statistics package).
-

## A Grouped Frequency Distribution

▣ Here is an example

<i>Interval</i>	<i>Tally</i>	<i>Frequency</i>
85-89		4
80-84		11
75-79		15
70-74		6
65-69		3
60-64		6
55-59		3
50-54		0
45-49		2
40-44		0
		<i>N</i> = 50

## A Stem-Leaf Diagram

▣ This preserves the original data, and also shows distributional shape.

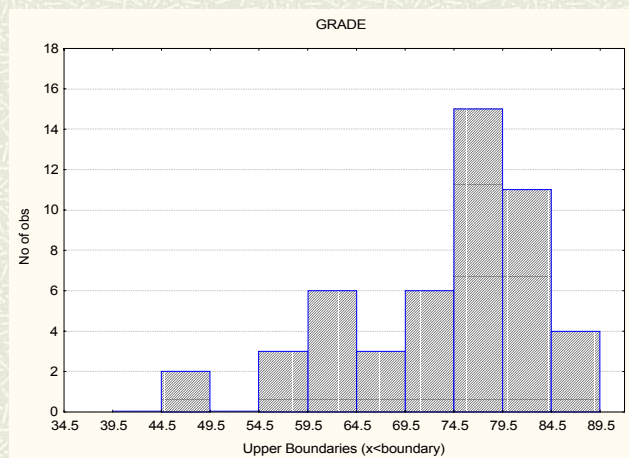
8*	8859
8.	34400001042
7*	997758655596778
7.	413233
6*	868
6.	243121
5*	569
5.	
4*	66
4.	

# Advantages of the Stem-Leaf Diagram

- *Advantages over the tally method?*
  - Original numbers preserved.
  - Summary statistics easily calculated.
  - Easy re-alignment of intervals.
  - Easier recovery from errors.

# A Frequency Histogram

Frequency Histogram



## A Frequency Distribution Table

#

<i>Nominal Interval</i>	<i>Real Interval</i>	<i>Freq.</i>	<i>Relative Freq.</i>	<i>Cum. Freq.</i>	<i>Cum. Rel. Freq.</i>
85–89	84.5–89.5	4	.08	50	1.00
80–84	79.5–84.5	11	.22	46	.92
75–79	74.5–79.5	15	.30	35	.70
70–74	69.5–74.5	6	.12	20	.40
65–69	64.5–69.5	3	.06	14	.28
60–64	59.5–64.5	6	.12	11	.22
55–59	54.5–59.5	3	.06	5	.10
50–54	49.5–54.5	0	.00	2	.04
45–49	44.5–49.5	2	.04	2	.04
40–44	39.5–44.5	0	.00	0	.00
		$N = 50$			

## Quantiles

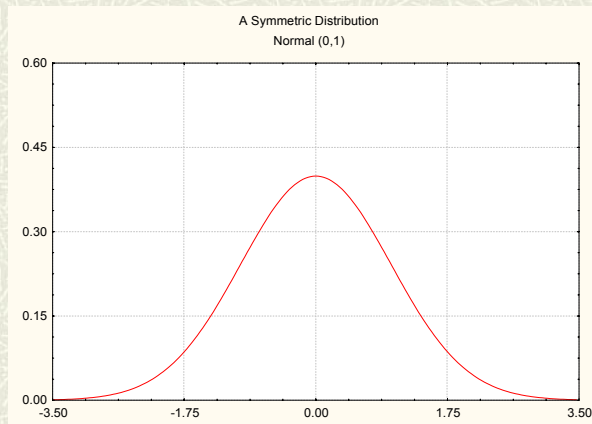
# **Quantiles** – The  $n$ th quantile in a distribution is that point at or below which  $n$  quantities fall.

# **Examples.**

- The 95<sup>th</sup> *percentile*. A point in the distribution at or below which 95 percent of the cases fall.
- The 3<sup>rd</sup> *decile*. A point in the distribution at or below which 3 tenths of the cases fall.
- The 2<sup>nd</sup> *quartile*. A point in the distribution at or below which 2 quarters of the cases fall.
- The third decile is equal to the 30<sup>th</sup> percentile.
- The second quartile is equal to the 50<sup>th</sup> percentile.

# Distributional Shape

## # A Symmetric Distribution



# Distributional Shape

## # A skewed distribution

