Chapter 2: Frequency Distributions

"If I had only one day left to live, I would live it in my statistics class;
it would seem so much longer."
~ Anonymous

Learning Objectives

Upon completion of this chapter, students should know

- How to create ranked frequency distributions.
- How to create simple frequency distributions.
- How to generate grouped frequency distributions.
- How to compute and differentiate between simple frequency, cumulative frequency, relative frequency, cumulative relative frequency, and cumulative percent.

Key Terms

Terms are listed in the order they appear in the chapter.

Raw data are the scores or numbers that have been collected but not organized or summarized.

Ranked distribution is a distribution of numbers in which scores are arranged in order (ranked), with the highest number at the top and the lowest number at the bottom of a list.

Simple frequency distributions are created by listing all possible score values in a distribution and then indicating the frequency.

Frequency (*f*) is the number of times a score occurs. The sum of the frequencies in a distribution is equal to the total number of scores in the distribution. $\sum f = n$

In a **grouped frequency distribution**, the raw data are combined into equal-sized groups called class intervals. The rule of thumb is to create between 10 and 20 class intervals and adjust the size of the intervals accordingly.

$$\text{number of intervals} = \frac{\text{range}}{i} \qquad i = \frac{\text{range}}{\text{number of intervals}}$$

Class intervals are equal-sized groups of raw data used to summarize data in a grouped frequency distribution.

Apparent limits are the limits of a class interval in the same units as the original data.

Range is the full extent of scores from the highest to the lowest in the distribution.

Real limits are the true extensions of the lower apparent limit minus 0.5 unit and the upper apparent limit plus 0.5 unit.

Midpoint is the average (center) of a class interval. $\text{Midpoint} = \dfrac{\text{lower limit} + \text{upper limit}}{2}$

Cumulative frequency (cum *f*) is the total number of scores that fall below the upper real limit of an interval.

Relative frequency (*rel f*) is the proportion of scores from the distribution that fall within the real limits of an interval. $rel\ f = \dfrac{f}{n}$

Cumulative percent (*cum %*) is also known as percentile. It is the percentage of scores that fall below the exact upper limit of the interval. *cum %* = *cum rel f* · 100

Cumulative relative frequency (cum *rel f*) is the total proportion of scores that lie below the real upper limit of the interval. $cum\ rel\ f = \dfrac{cf}{n}$

Lecture and Demonstration Ideas

The purpose of descriptive statistics is to organize and summarize raw data in a meaningful manner. Researchers usually end up with long lists of data and students quickly understand that large amounts of data are virtually meaningless without organization and summarization.

Wechsler IQ Scores. The IQ scores collected in Willerman, Schultz, Rutledge, & Bigler's (1991) brain size and intelligence study can be used to demonstrate how to generate a grouped frequency distribution. In this study, 40 introductory psychology students took four subsets (Vocabulary, Similarities, Block Design, and Picture Completion) of the Wechsler (1981) Adult Intelligence Scale-Revised. Their scores are shown in Handout 2-A. A completed grouped frequency distribution for this dataset using an interval size of 10 is shown below.

range = 144 - 77 = 77 $\quad i = 10 \quad \dfrac{77}{10} = 7.7 \sim 8$

Demonstration Solution *i* = 10

Real Limits	Apparent Limits	*f*	Midpoint	Cumulative Frequency	Relative Frequency	Cumulative Relative Frequency	Cumulative Percent
139.5-149.5	140-149	6	144.5	40	.150	1.00	100
129.5-139.5	130-139	14	134.5	34	.350	.85	85
119.5-129.5	120-129	0	124.5	20	0	.50	50
109.5-119.5	110-119	0	114.5	20	0	.50	50
99.5-109.5	100-109	4	104.5	20	.100	.50	50
89.5-99.5	90-99	6	94.5	16	.150	.40	40
79.5-89.5	80-89	9	84.5	10	.225	.25	25
69.5-79.5	70-79	1	74.5	1	.025	.025	2.5

Census 2000. Use the portion of the Census 2000 - California Population shown in Table 1(Transparency 2-1) as an example of a simple frequency distribution. Although large numbers are used, you may want to use this data to demonstrate how grouped frequency distributions are generated. If so, Table 2 shows the complete grouped frequency distribution for this dataset using an interval size of 4.

Dataset Note: Since the "under 1 year" interval is not shown in months, the "under 1 year " and the "1-year" age groups were combined to simplify the dataset.

Census 2000: California Population = 33,871,648

California Population 21 yrs and younger = 11,195,956 (33.05% of the population)

Table 1.

Census 2000: California Population by Each Single Age (both sexes combined)

Under 1 year	483,143	**11 years**	521,774
1 year	486,587	**12 years**	511,339
2 years	489,336	**13 years**	494,910
3 years	504,490	**14 years**	494,250
4 years	523,425	**15 years**	488,690
5 years	531,405	**16 years**	483,269
6 years	536,382	**17 years**	494,187
7 years	544,099	**18 years**	490,003
8 years	553,828	**19 years**	494,739
9 years	560,166	**20 years**	490,829
10 years	548,549	**21 years**	470,556

Demonstration Solution. $i = 4$

Table 2.

Grouped Frequency Distribution of Census 2000: California Population Data (partial)

Real Limits	Apparent Limits	Frequency	Cumulative Frequency	Relative Frequency	Cumulative Relative Frequency	Cumulative Percent Frequency
19.5 – 23.5	20 – 23	961,385	11,195,956	.086	1.000	100
15.5 – 19.5	16 – 19	1,962,198	10,234,571	.175	.914	91.4
11.5 – 15.5	12 – 15	1,989,189	8,272,373	.178	.739	73.9
7.5 – 11.5	8 – 11	2,184,317	6,283,184	.195	.561	56.1
3.5 – 7.5	4 – 7	2,135,311	4,098,867	.191	.366	36.6
0 – 3.5	0 - 3	1,963,556	1,963,556	.175	.175	17.5

Active-Learning Activities

Activity 1: Census 2000

Have students practice generating a grouped frequency distribution using the full dataset, Census 2000: California Population By Single Year, shown in Handout 2-B (see Dataset Note below). This generates questions and discussion about interval size selection and the number of class intervals. Take this opportunity to discuss the problems when too few or too many class intervals are used. Assign different interval sizes to students and compare these distributions in terms of information loss or gain.

Additional Assignments

The Census 2000: United States Population By Each Single Year of Age By State dataset can be downloaded and distributed to the class (see below). Divide the class into small groups and ask each group to select and generate a relative frequency distribution for one state. Consider using the same interval size for all the distributions so these distributions can be used to create relative frequency polygons in the next chapter.

> ***Census 2000*:** This dataset lists the total number of people (both sexes are combined) for each single age, ranging from 1 to 110 years, living at census. Data are listed separately for each state. Instructor's Area: Filename for this dataset is Census 2000. Format: Excel

References

The Administration on Aging (2001). Census 2000: Single year of age by state. Accessed 9/01/01. http://www.aoa.dhhs.gov/research.html

Willerman, L., Schultz, R., Rutledge, J. N., & Bigler, E. (1991). In vivo brain size and intelligence. *Intelligence, 15,* 223-228.

Handout 2-A

Wechsler IQ Scores			
144	135	103	89
141	135	103	89
141	133	101	88
140	133	100	85
140	133	99	83
140	133	97	83
139	133	96	83
139	132	92	81
138	132	91	80
137	130	90	77

WORKSHEET

Real Limits	Apparent Limits	*f*	Midpoint	Cumulative Frequency	Relative Frequency	Cumulative Relative Frequency	Cumulative Percent

Handout 2-B.

The State of California					
Total both Sexes, all ages	33,871,648	This data set lists the total number of people (both sexes are combined) for each single age, ranging from 1 to 110 years, living in California at census.			
Under 1 year	483,143	**41 years**	528,663	**82 years**	99,417
1 year	486,587	**42 years**	538,628	**83 years**	88,555
2 years	489,336	**43 years**	519,315	**84 years**	80,787
3 years	504,490	**44 years**	508,057	**85 years**	74,108
4 years	523,425	**45 years**	513,038	**86 years**	64,654
5 years	531,405	**46 years**	475,896	**87 years**	55,775
6 years	536,382	**47 years**	465,367	**88 years**	46,702
7 years	544,099	**48 years**	443,306	**89 years**	39,656
8 years	553,828	**49 years**	434,185	**90 years**	32,445
9 years	560,166	**50 years**	444,820	**91 years**	26,411
10 years	548,549	**51 years**	410,344	**92 years**	21,662
11 years	521,774	**52 years**	415,046	**93 years**	16,851
12 years	511,339	**53 years**	404,876	**94 years**	13,019
13 years	494,910	**54 years**	324,757	**95 years**	9,682
14 years	494,250	**55 years**	324,412	**96 years**	7,355
15 years	488,690	**56 years**	309,786	**97 years**	5,362
16 years	483,269	**57 years**	306,166	**98 years**	3,758
17 years	494,187	**58 years**	271,187	**99 years**	2,876
18 years	490,003	**59 years**	255,701	**100 to 104 years**	4,856
19 years	494,739	**60 years**	252,240	**105 to 109 years**	356
20 years	490,829	**61 years**	236,405	**110 years and over**	129
21 years	470,556	**62 years**	229,913		
22 years	468,991	**63 years**	216,443		
23 years	472,881	**64 years**	211,840		
24 years	478,031	**65 years**	212,489		
25 years	496,124	**66 years**	193,991		
26 years	479,629	**67 years**	194,757		
27 years	500,830	**68 years**	189,485		
28 years	515,683	**69 years**	193,813		
29 years	551,275	**70 years**	193,874		
30 years	575,175	**71 years**	180,024		
31 years	530,404	**72 years**	180,578		
32 years	526,149	**73 years**	175,122		
33 years	519,912	**74 years**	173,690		
34 years	533,881	**75 years**	172,110		
35 years	573,400	**76 years**	163,420		
36 years	562,657	**77 years**	154,760		
37 years	560,032	**78 years**	150,567		
38 years	558,614	**79 years**	138,490		
39 years	560,040	**80 years**	125,225		
40 years	575,935	**81 years**	108,847		

Data Source:
Data Set: Census 2000 Summary File 1 (SF 1) 100-Percent Data
The U.S. Department of Health and Human Services
The Administration on Aging
Census 2000: Single Year of Age by State
http://www.aoa.dhhs.gov/research.html

Transparency 2-1.

Census 2000: California Population by Each Single Age (both sexes combined)			
Under 1 year	483,143	**11 years**	521,774
1 year	486,587	**12 years**	511,339
2 years	489,336	**13 years**	494,910
3 years	504,490	**14 years**	494,250
4 years	523,425	**15 years**	488,690
5 years	531,405	**16 years**	483,269
6 years	536,382	**17 years**	494,187
7 years	544,099	**18 years**	490,003
8 years	553,828	**19 years**	494,739
9 years	560,166	**20 years**	490,829
10 years	548,549	**21 years**	470,556