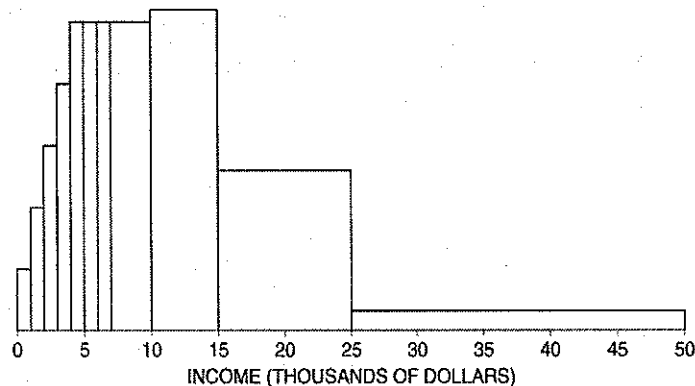


Chapter 3: The Histogram

A histogram represents percentages by AREA, not by height

Total area = 100%

Figure 1. A histogram. This graph shows the distribution of families by income in the U.S. in 1973.



Source: Current Population Survey.

Don't plot the percentages!

Figure 3. Don't plot the percents.

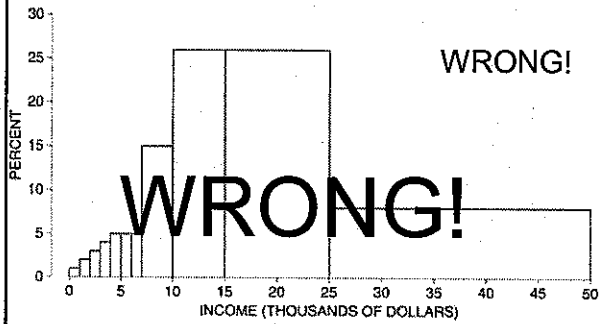
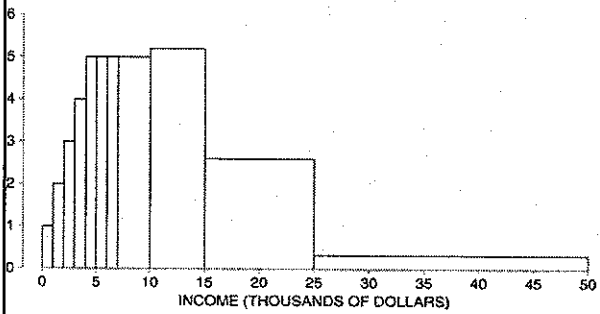


Figure 2. The histogram from figure 1, with a vertical scale supplied.



Income	%
\$0 - \$1,000	1
\$1,000 - \$2,000	2
\$2,000 - \$3,000	3
\$3,000 - \$4,000	4
\$4,000 - \$5,000	5
\$5,000 - \$6,000	5
\$6,000 - \$7,000	5
\$7,000 - \$10,000	15
\$10,000 - \$15,000	26
\$15,000 - \$25,000	26
\$25,000 - \$50,000	8
\$50,000 and over	1

HOW TO MAKE A HISTOGRAM:

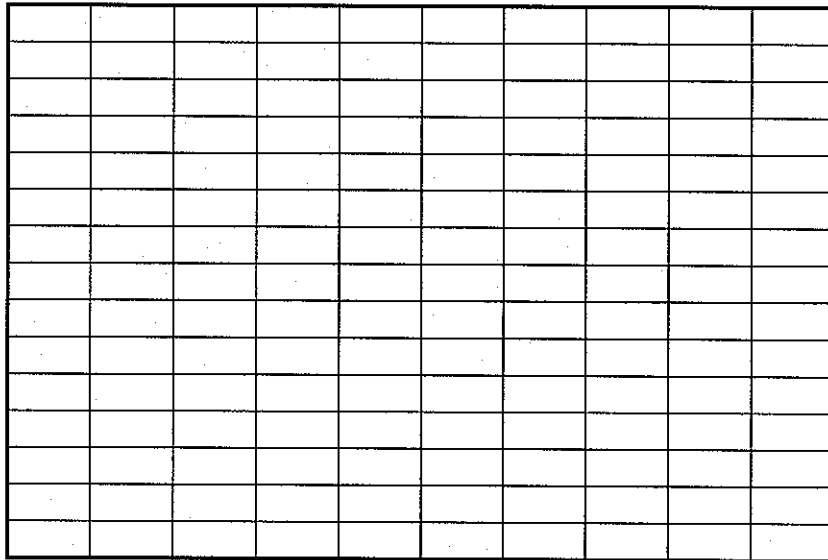
1. Determine the class intervals. Decide on the endpoint convention.
2. Determine frequency, relative frequency, and percentage for each interval.
3. Height of the block = $\frac{\%}{\text{Width of the interval}}$
4. Draw

In a histogram, the areas of the blocks represent percentages.

Example 1: Drawing a histogram

Exam Score	Number of students	Percentage of students	Width of interval	Height
0-50	25			
50-70	50			
70-80	75			
80-90	75			
90-100	25			

Example 1: Drawing a histogram



Using a histogram

1. What percentage of students scored 80 or above?
2. What percentage scored at least 60 but less than 70?
3. What percentage scored at least 25 but less than 55?

Labeling the Axes

The y-axis label should be "Percent per" followed by the UNITS OF MEASUREMENT of the x axis.

x is height in inches

y label is "Percent per inch"

x is weight in kilograms

y label is "Percent per kilogram"

x is income in thousands

y label is "Percent per thousand"

x is number of packs of cigarettes

y label is "Percent per pack"

NLVM: Histogram Manipulative

Variables

A variable is a characteristic of a person or thing which is of interest in a study. For example, in a health study we might be interested in people's blood pressure, weight, height, cholesterol level, etc.

Examples:

- Age
- Height
- Weight
- Income
- Family size
- Occupation
- Race
- Religion

Types of Variables

A **quantitative** variable is measured by a number.

Examples:

- Age
- Height
- Weight
- Income
- Family size

A **qualitative** variable is measured by a category.

Examples:

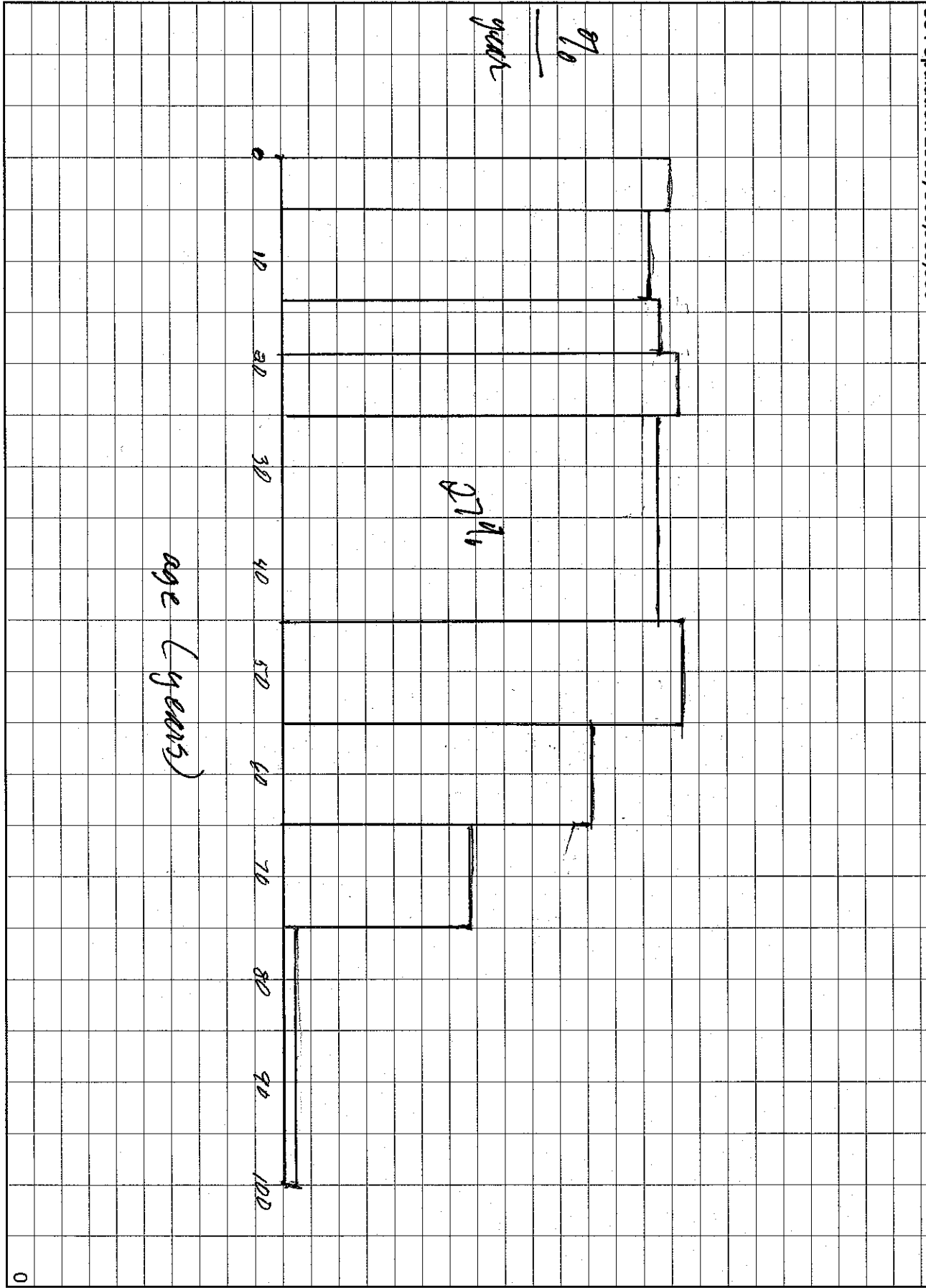
- Occupation
- Race
- Religion

B

Population of the United States for the Year 2009					306,983,000
Age Group	Total	%	Width	Height	
Under 5	21,300,000	6.94	5	1.39	
5 to 13	36,487,000	11.89	9	1.32	
14 to 17	16,761,000	5.46	4	1.36	
18 to 24	30,412,000	9.91	7	1.42	
25 to 34	41,567,000	13.54	10	1.35	
35 to 44	41,530,000	13.53	10	1.35	
45 to 54	44,592,000	14.53	10	1.45	
55 to 64	34,763,000	11.32	10	1.13	
65 to 74	20,792,000	6.77	10	0.68	
75 and over	18,779,000	6.12	25	0.24	

$$\frac{21,300,000}{306,983,000} \times 100\% = 6.94\%$$

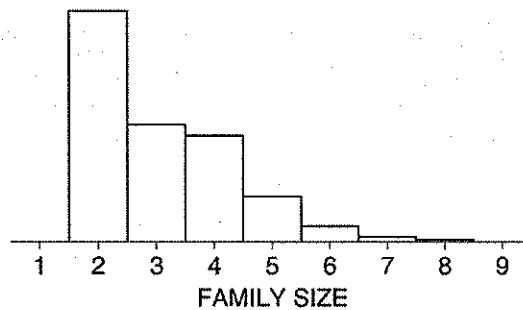
US Population 2009, 306,983,000



Histograms are done differently for discrete variables!

Usually, each class interval represents one number and the number is placed in the middle of the bar.

Figure 6. Histogram showing distribution of families by size in 2005. With a discrete variable, the class intervals are centered at the possible values.



Source: March 2005 Current Population Survey; CD-ROM supplied by the Bureau of the Census.

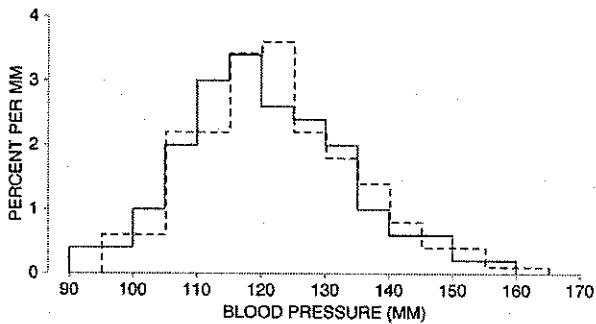
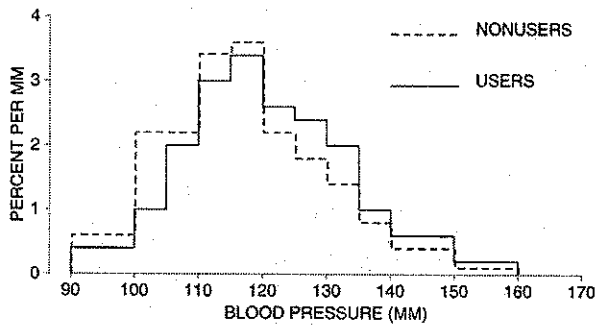
Bar Chart

Table 2. Systolic blood pressure by age and pill use, for women in the Contraceptive Drug Study, excluding those who were pregnant or taking hormonal medication other than the pill. Class intervals include the left endpoint, but not the right. - means negligible. Table entries are in percent; columns may not add to 100 due to rounding.

<i>Blood pressure (millimeters)</i>	<i>Age 17-24</i>		<i>Age 25-34</i>		<i>Age 35-44</i>		<i>Age 45-58</i>	
	<i>Non- users</i>	<i>Users</i>	<i>Non- users</i>	<i>Users</i>	<i>Non- users</i>	<i>Users</i>	<i>Non- users</i>	<i>Users</i>
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
under 90	-	1	1	-	1	1	1	-
90-95	1	-	1	-	2	1	1	1
95-100	3	1	5	4	5	4	4	2
100-105	10	6	11	5	9	5	6	4
105-110	11	9	11	10	11	7	7	7
110-115	15	12	17	15	15	12	11	10
115-120	20	16	18	17	16	14	12	9
120-125	13	14	11	13	9	11	9	8
125-130	10	14	9	12	10	11	11	11
130-135	8	12	7	10	8	10	10	9
135-140	4	6	4	5	5	7	8	8
140-145	3	4	2	4	4	6	7	9
145-150	2	2	2	2	2	5	7	9
150-155	-	1	1	1	1	3	2	4
155-160	-	-	-	1	1	1	1	3
160 and over	-	-	-	-	1	2	2	5
Total percent	100	98	100	99	100	100	99	99
Total number	1,206	1,024	3,040	1,747	3,494	1,028	2,172	437

Comparing blood pressure for pill users and non-users

Figure 7. The effect of the pill. The top panel shows histograms for the systolic blood pressures of the 1,747 users and the 3,040 non-users age 25-34 in the Contraceptive Drug Study. The bottom panel shows the histogram for the non-users shifted to the right by 5 mm.



The crosstabs for pill use and blood pressure are easier to interpret when we do a histogram

Example: intelligence of rats

Figure 8. Tryon's experiment. Distribution of intelligence in the original population.

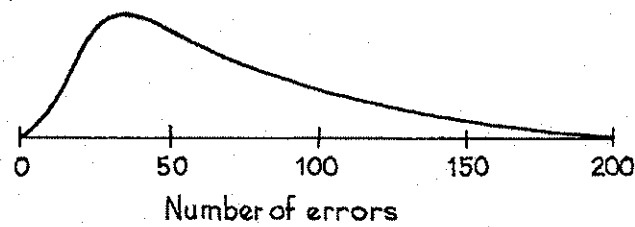
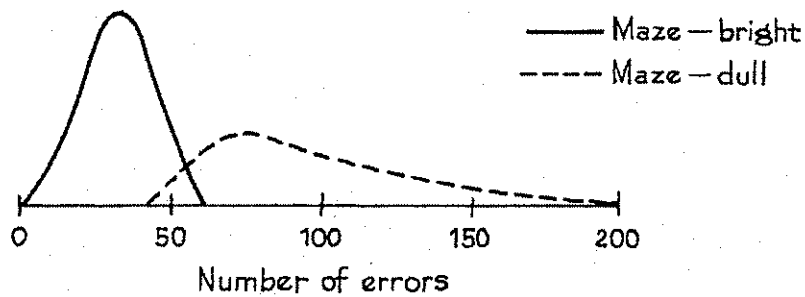




































Figure 9. Tryon's experiment. After seven generations of selective breeding, there is a clear separation into "maze-bright" and "maze-dull" strains.



Homicide Rates

Rank	Countries	Amount ▼	
# 1	<u>South Africa:</u>	1.19538 per 1,000 people	
# 2	<u>Seychelles:</u>	0.788294 per 1,000 people	
# 3	<u>Australia:</u>	0.777999 per 1,000 people	
# 4	<u>Montserrat:</u>	0.749384 per 1,000 people	
# 5	<u>Canada:</u>	0.733089 per 1,000 people	
# 6	<u>Jamaica:</u>	0.476608 per 1,000 people	
# 7	<u>Zimbabwe:</u>	0.457775 per 1,000 people	
→ # 8	<u>Dominica:</u>	0.34768 per 1,000 people	
→ # 9	<u>United States:</u>	0.301318 per 1,000 people	
# 10	<u>Iceland:</u>	0.246009 per 1,000 people	
# 11	<u>Papua New Guinea:</u>	0.233544 per 1,000 people	
# 12	<u>New Zealand:</u>	0.213383 per 1,000 people	
# 13	<u>United Kingdom:</u>	0.142172 per 1,000 people	
# 14	<u>Spain:</u>	0.140403 per 1,000 people	
# 15	<u>France:</u>	0.139442 per 1,000 people	
# 16	<u>Korea, South:</u>	0.12621 per 1,000 people	
# 17	<u>Mexico:</u>	0.122981 per 1,000 people	
# 18	<u>Norway:</u>	0.120836 per 1,000 people	
# 19	<u>Costa Rica:</u>	0.118277 per 1,000 people	
# 20	<u>Venezuela:</u>	0.115507 per 1,000 people	
# 21	<u>Finland:</u>	0.110856 per 1,000 people	
# 22	<u>Netherlands:</u>	0.100445 per 1,000 people	
# 23	<u>Denmark:</u>	0.0914948 per 1,000 people	
# 24	<u>Germany:</u>	0.0909731 per 1,000 people	
# 25	<u>Bulgaria:</u>	0.0795973 per 1,000 people	
# 26	<u>Chile:</u>	0.0782179 per 1,000 people	
# 27	<u>Thailand:</u>	0.0626305 per 1,000 people	
# 28	<u>Kyrgyzstan:</u>	0.0623785 per 1,000 people	
# 29	<u>Poland:</u>	0.062218 per 1,000 people	
# 30	<u>Sri Lanka:</u>	0.0599053 per 1,000 people	
# 31	<u>Hungary:</u>	0.0588588 per 1,000 people	
# 32	<u>Estonia:</u>	0.0547637 per 1,000 people	
# 33	<u>Ireland:</u>	0.0542829 per 1,000 people	
# 34	<u>Switzerland:</u>	0.0539458 per 1,000 people	

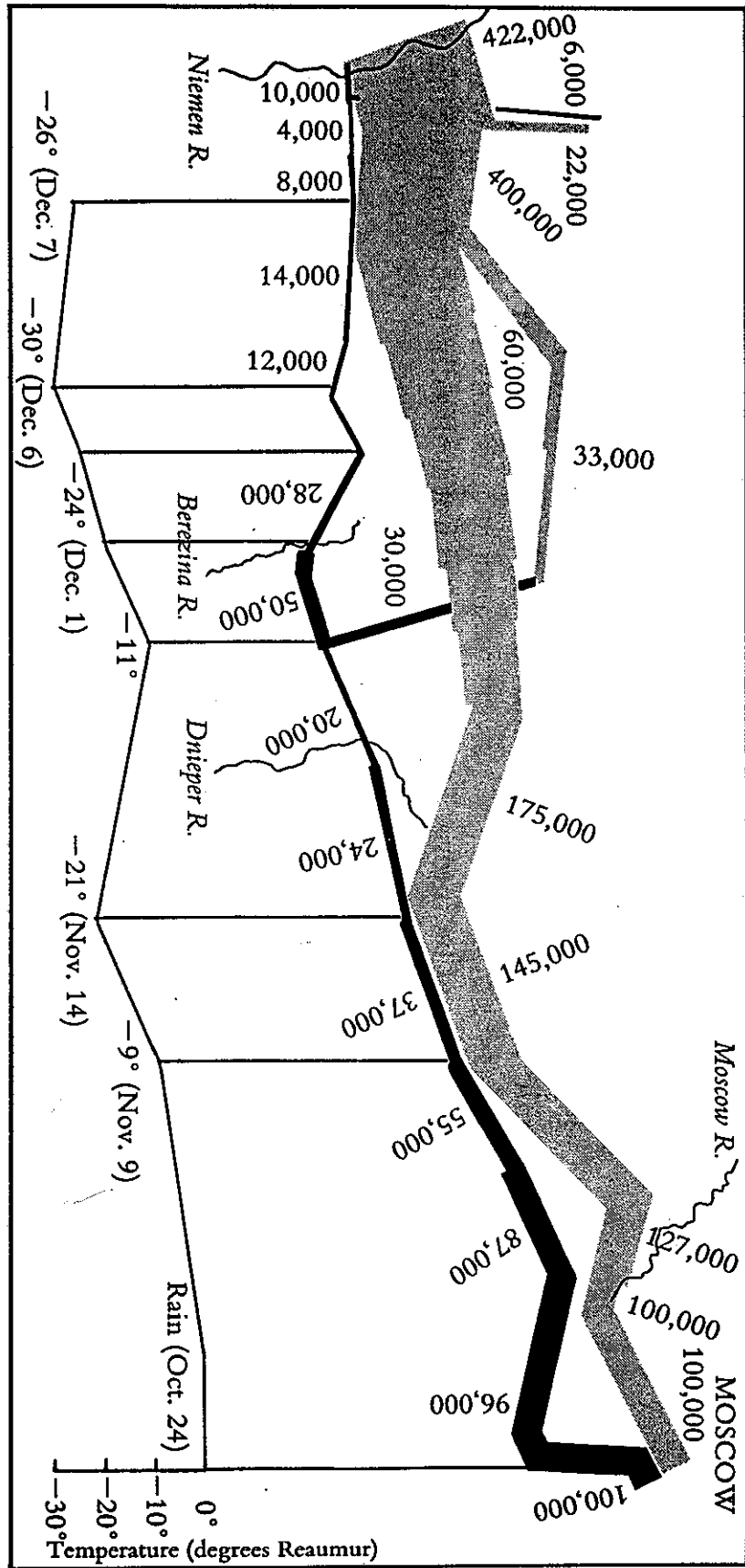


Figure 6.1 Redrawing of Charles Minard's 1861 graph of Napoleon's Russian campaign. (In the Reaumur temperature scale, water boils at 80° R and freezes at 0° R.)

1812

Fatality Rate
All Highways 1968-1982

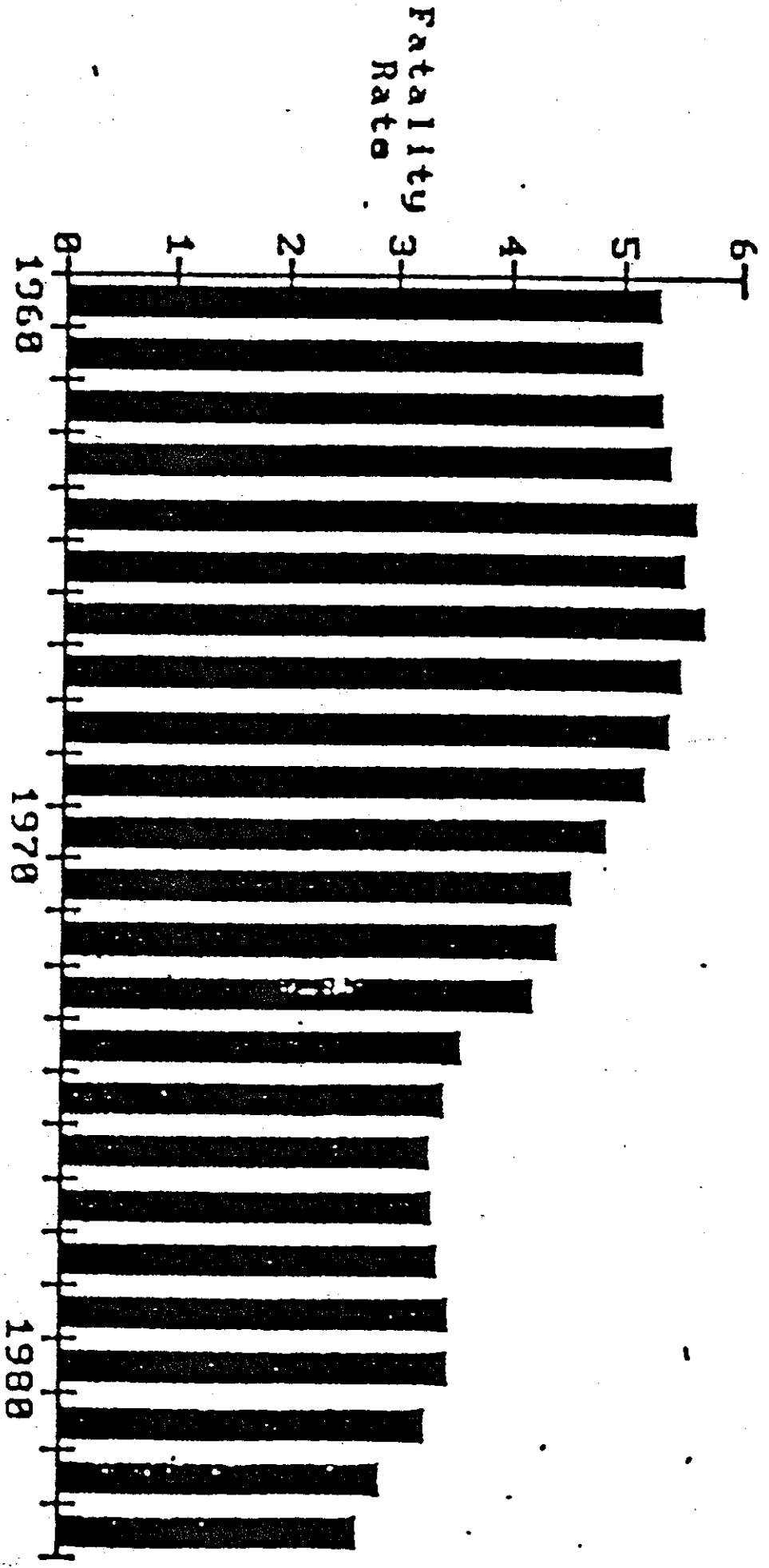
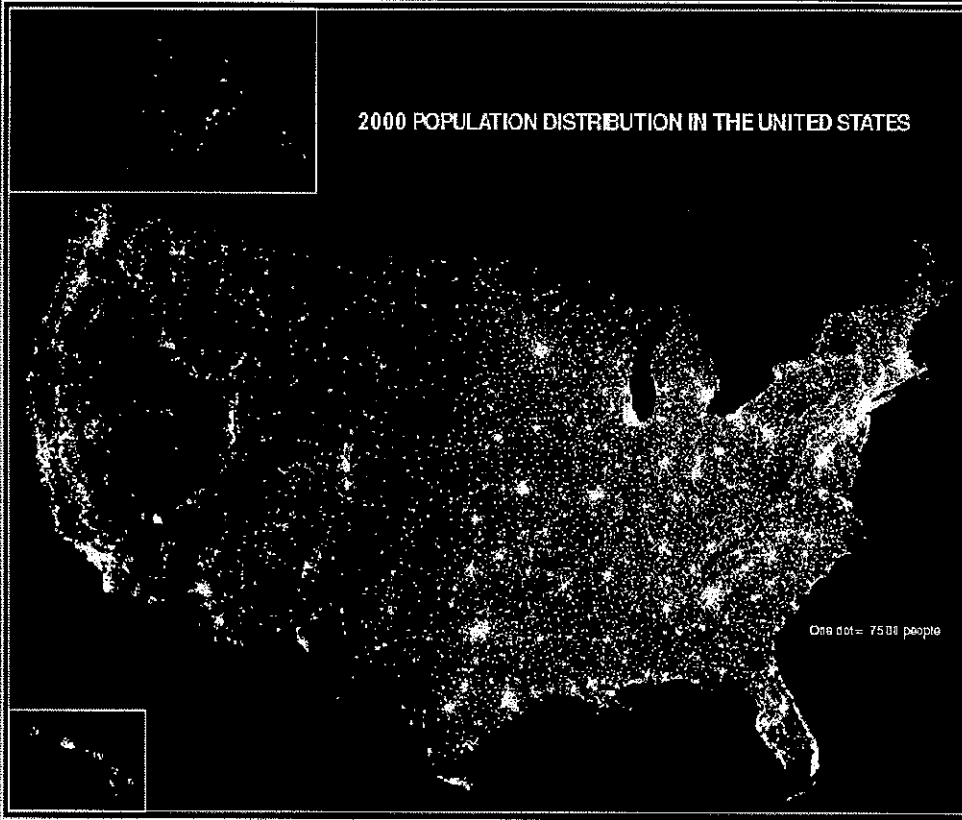


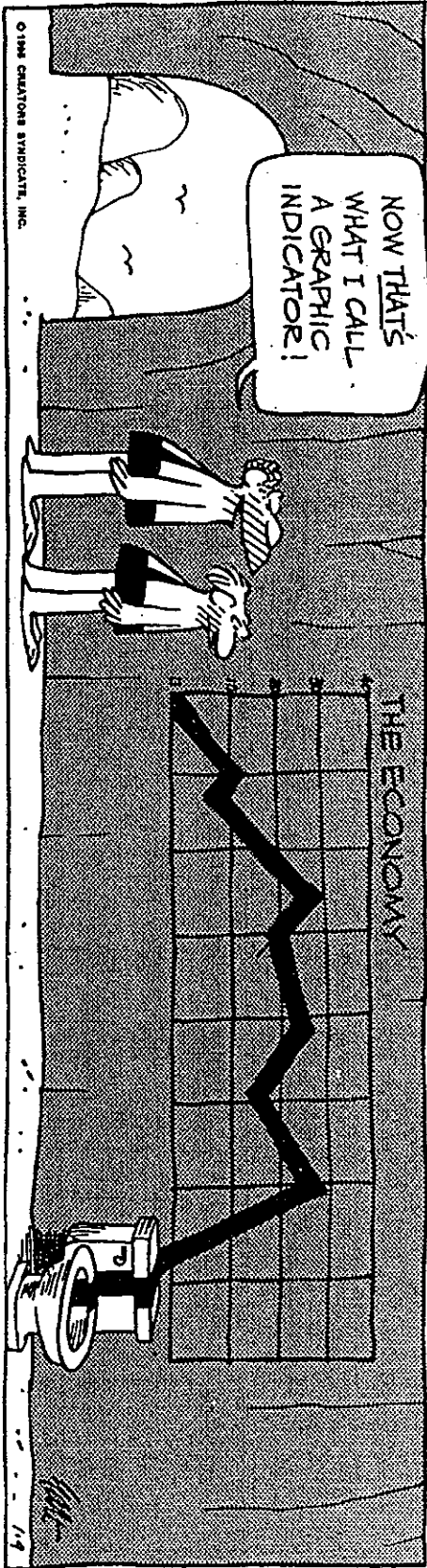
Figure 1

2000 POPULATION DISTRIBUTION IN THE UNITED STATES



Prepared by Geography Division, U.S. Department of Commerce, Economic and Statistical Administration, U.S. Census Bureau

B.C.



© 1986 CREATORS SYNDICATE, INC.