Exercise

In a study, researchers treated 570 people who smoke with either nicotine gum or a nicotine patch. Among those treated with nicotine gum, 191 continued to smoke and the other 59 stopped smoking. Among those treated with nicotine patch, 263 continued to smoke and the other 57 stopped smoking (based on data from the Center for Disease Control and Prevention). Construct the relative frequency distribution.

Solution

Obtain the relative frequencies by dividing the given frequencies by the total of 570.

Those whose smoking		***Relative Frequency***
Continued after the gum	$\frac{191}{570} = 0.335$	33.5 %
Stopped after the gum	$\frac{59}{570} \approx 0.104$	10.4 %
Continued after the patch	$\frac{263}{570} \approx 0.461$	46.1 %
Stopped after the patch	$\frac{57}{570} \approx 0.1$	10.0 %
		100.0 %

Exercise

Heights of statistics students were obtained by the author as part of a study conducted for class. The last digits of those heights are listed below. Construct a frequency distribution with 10 classes. Based on the distribution, do the heights appear to be reported or actually measured? What do you know about the accuracy of the results?

0 0 0 0 0 0 0 0 0 1 1 2 3 3 3 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 6 6 8 8 8 9

Solution

Digit	***Frequency***
0	9
1	2
2	1
3	3
4	1
5	15
6	2
7	0
8	3
9	1
	37

The heights are not actually measured because the data have been round off and showed unusually high numbers of 0's and 5's.

Exercise

Listed below are amounts of strontium-90 (in millibecquerels) in a simple random sample of baby teeth obtained from Pennsylvania residents born after 1979, Construct a frequency distribution with eight classes. Begin with a lower class limit of 110, and use a class width of 0. Cite a reason why such data are important.

155	142	149	130	151	163	151	142	156	133	138	161	128	144
172	137	151	166	147	163	145	116	136	158	114	165	169	145
150	150	150	158	151	145	152	140	170	129	188	156		

Solution

Lowest number: 114 Highest number: 188

level	***Frequency***
110 – 119	2
120 – 129	2
130 – 139	5
140 – 149	9
150 – 159	13
160 – 169	6
170 – 179	2
180 – 189	1
	40

Such data are important because they can be helpful in detecting potentially dangerous situations and in making recommendations for future action

Exercise

Refer to the data below and use the 40 voltage measurements from the generator. Construct a frequency distribution with seven classes. Begin with a lower class limit of 123.9 volts, and use a class width of 0.20 volt. Using a very loose interpretation of the relevant criteria, does the result appear to have a normal distribution?

124.8	124.3	125.2	124.5	125.1	124.8	125.1	125.0	124.8	124.7	124.5	125.2	124.4	124.7
124.9	124.5	124.8	124.8	124.5	124.6	125.0	124.7	124.9	124.9	124.7	124.2	124.7	124.8
124.4	124.6	124.4	124.0	124.7	124.4	124.6	124.6	124.6	124.8	124.3	124.0		

Solution

Voltage (volts)	*Frequency*
123.9 – 124.0	2
124.1 – 124.2	1
124.3 – 124.4	6
124.5 – 124.6	9
124.7 – 124.8	13
124.9 – 125.0	5
125.1 – 125.2	4

Yes;
The voltages do appear to follow a normal distribution.
There are many values near the center of the distribution, and the frequencies diminish toward either end.

Exercise

As part of the Garbage Project at the University of Arizona, the discarded garbage for 62 households was analyzed. Refers to the 62 weights from table below and construct a frequency distribution. Begin with a lower class of 1.00 lb., and use a class width of 4.00 lb. Do the weights of discarded paper appear to have a normal distribution?

2.41	11.08	9.45	5.88	7.57	12.43	12.32	8.26	9.55	6.05	20.12	12.45
8.82	13.61	7.72	10.58	8.72	6.98	6.16	5.87	6.96	14.33	7.98	8.78
6.83	13.31	9.64	11.03	11.42	3.27	8.08	12.29	16.08	6.67	10.99	20.58
6.38	17.65	13.11	12.56	13.05	12.73	3.26	9.92	11.36	9.83	1.65	3.45
15.09	16.39	10	9.09	2.8	6.33	8.96	3.69	6.44	9.19	9.46	2.61
5.86	9.41										

Solution

Weight (lbs)	*Frequency*
1.00 – 4.99	8
5.00 – 8.99	21
9.00 – 12.99	22
13.00 – 16.99	8
17.00 – 20.99	3
	62

Yes,
The weights appear to have a distribution that is approximately normal.

Exercise

a) Refer to the data below for the FICO credit rating scores. Construct a frequency distribution beginning with a lower class limit of 400, and use a class width of 50. Does the result appear to have a normal distribution? Why or why not?

708	713	781	809	797	793	711	681	768	611	698	729	829
836	768	532	657	559	741	792	701	753	745	681	594	744
598	693	743	444	502	739	755	835	714	517	787	706	752
714	497	636	637	797	568	714	618	830	579	818	722	783
751	731	850	591	802	756	689	789	654	617	849	604	630
628	692	779	756	782	760	503	784	798	611	709	661	579
591	834	694	795	660	651	696	638	697	732	796	753	782
635	795	519	682	824	603	709	777	664				

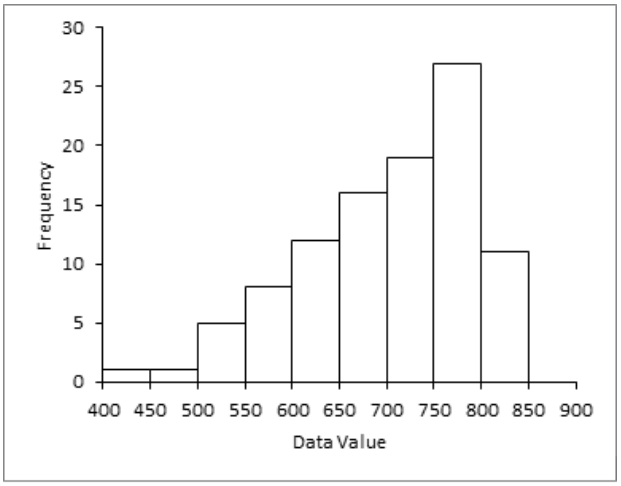
b) Use the table to construct a histogram. Does the result appear to be normal distribution? Why or why not?

Solution

a)

FICO score	***Frequency***
400 – 449	1
450 – 499	1
500 – 549	5
550 – 599	8
600 – 649	12
650 – 699	16
700 – 749	19
750 – 799	27
800 – 849	10
850 – 899	1
	100

b)



The data appear to have a distribution that is not normal. While there is a maximum value with progressively smaller frequencies on either side of the maximum, the distribution is definitely not symmetric.

Exercise

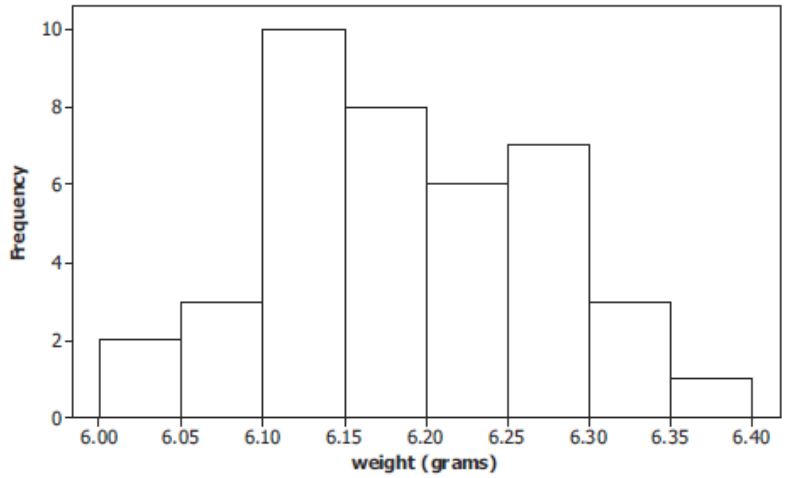
Refer to the data in the table below. Construct a frequency distribution. Begin with lower class limit of 6.0000 g, and use a class width of 0.0500 g.

6.2771	6.2371	6.1501	6.0002	6.1275	6.2151
6.2866	6.0760	6.1426	6.3415	6.1309	6.2412
6.1442	6.1073	6.1181	6.1352	6.2821	6.2647
6.2908	6.1661	6.2674	6.2718	6.1949	6.2465
6.3172	6.1487	6.0829	6.1423	6.1970	6.2441
6.3669	6.0775	6.1095	6.1787	6.2130	6.1947
6.1940	6.0257	6.1719	6.3278		

Solution

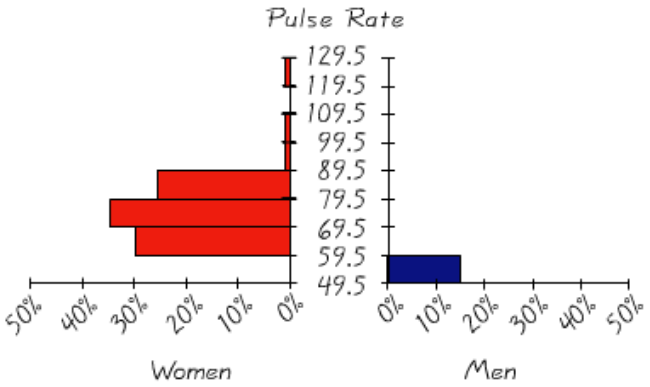
Weight (g)	***Frequency***
6.0000 – 6.04999	2
6.0500 – 6.0999	3
6.1000 – 6.1499	10
6.1500 – 6.1999	8
6.2000 – 6.2499	6
6.2500 – 6.2999	7
6.3000 – 6.3499	3
6.3500 – 6.3999	1
	40

b)



Exercise

When using histograms to compare two data sets, it is sometimes difficult to make comparisons by looking back and forth between the two histograms. A back-to-back relative frequencies histogram uses a format that makes the comparison much easier. Instead of frequencies, we should use relative frequencies (percentages or proportions) so that the comparisons are not distorted by different sample sizes. Complete the back-to-back relative frequency histograms shown below by using the data below. Then use the result to compare the two data sets.

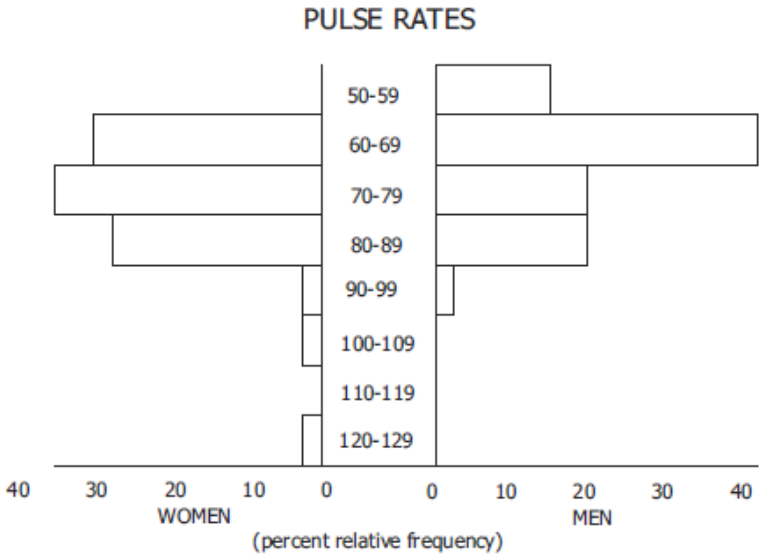


Pulse Rates (*beats per minute*) of Females and Males

Females																			
76	72	88	60	72	68	80	64	68	68	80	76	68	72	93	72	68	72	64	80
64	80	76	76	76	80	104	88	60	76	72	72	88	80	60	72	88	88	124	64
Males																			
68	64	88	72	64	72	60	88	76	60	96	72	56	64	60	64	84	76	84	88
72	56	68	64	60	68	60	60	56	84	72	84	88	56	64	56	56	60	64	72

Solution

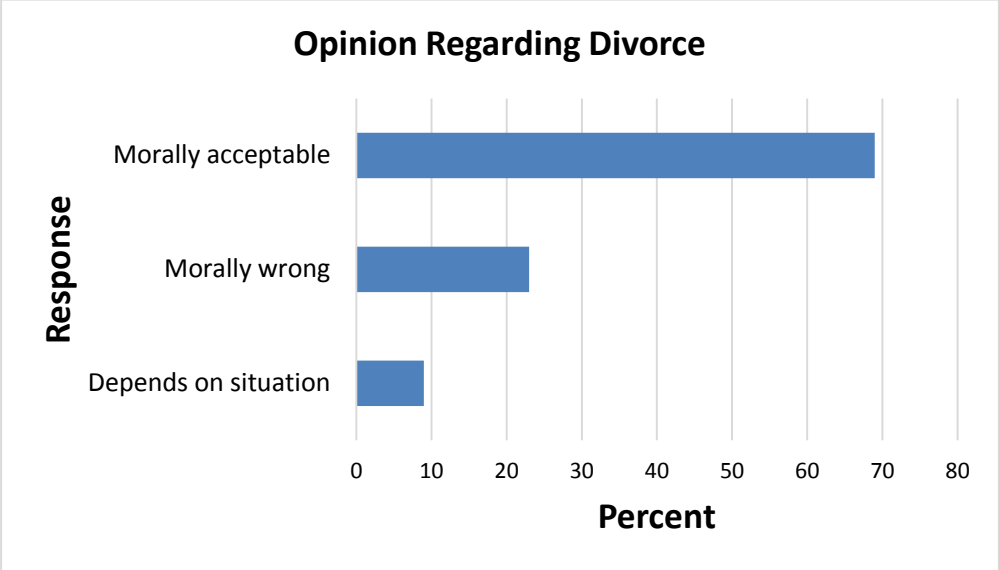
The back-to-back relative frequency histograms are given below. The pulse rates of the males tend to be lower than those of the females.



Exercise

The following graph represents the results of a survey, by Gallup in May 2010, in which a random sample of adult Americans was asked, "Please tell me whether you personally believe that in general divorce is morally accepted or morally wrong."

Response	***Percent***
Depends on situation	9
Morally wrong	23
Morally acceptable	69



a) What percent of the respondents believe divorce is morally acceptable?
b) If there were 240 million adult Americans, how many believe that divorce is morally wrong?
c) If Gallup claimed that the results of the survey indicate that 8% of adult Americans believe that divorce is acceptable in certain situations, would you say this statement is descriptive or inferential? Why?

Solution

a) 69% of the respondents believe divorce is morally acceptable
b) 23% believe divorce is morally wrong. So, 240 million * 0.23 = 55.2 million adult Americans believe divorce is morally wrong.
c) This statement is inferential, since it is a generalization based on the observed data.

Exercise

In a national survey conducted by the Centers for Disease Control to determine health-risk behaviors among college students, college students were asked, "How often do you wear a seat belt when driving a car?" The frequencies were as follow:

Response	***Frequency***
I do not drive a car	249
Never	118
Rarely	249
Sometimes	345
Most of the time	716
Always	3093

a) Construct a relative frequency distribution
b) What percentage of respondents answered "Always"?
c) What percentage of respondents answered "Never" or "Rarely"?
d) Construct a frequency bar graph.
e) Construct a relative frequency bar graph.
f) Construct a pie chart
g) Suppose that a representative from the Centers for Disease Control says, "2.5% of the college students in this survey responded that they never wear a seat belt." Is this a descriptive or inferential statement?

Solution

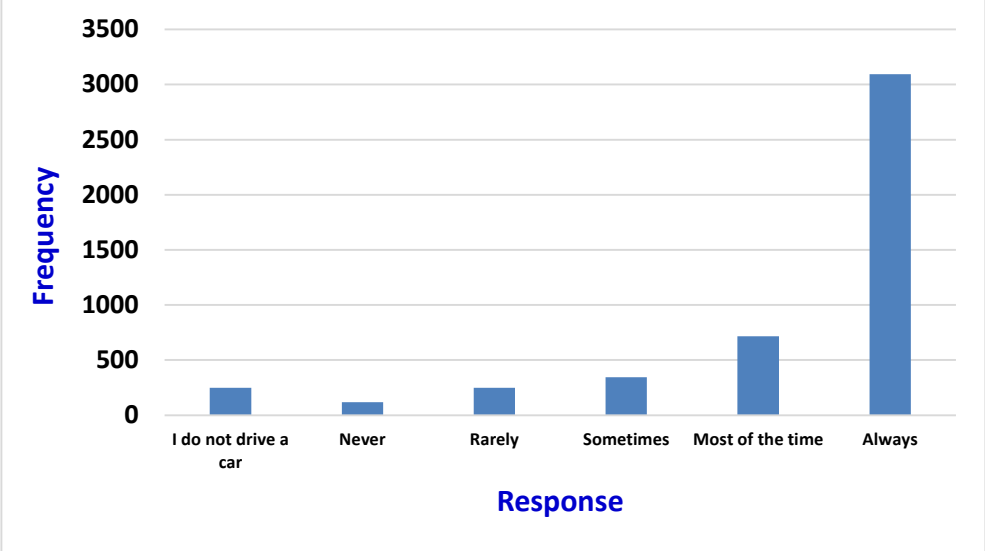
a)

Response	***Frequency***	***Relative Frequency***
I do not drive a car	249	$\frac{249}{4770} \approx 0.0522$
Never	118	$\frac{118}{4770} \approx 0.0247$
Rarely	249	$\frac{249}{4770} \approx 0.0522$
Sometimes	345	$\frac{345}{4770} \approx 0.0723$
Most of the time	716	$\frac{716}{4770} \approx 0.1501$
Always	3093	$\frac{3093}{4770} \approx 0.6484$
	4770	

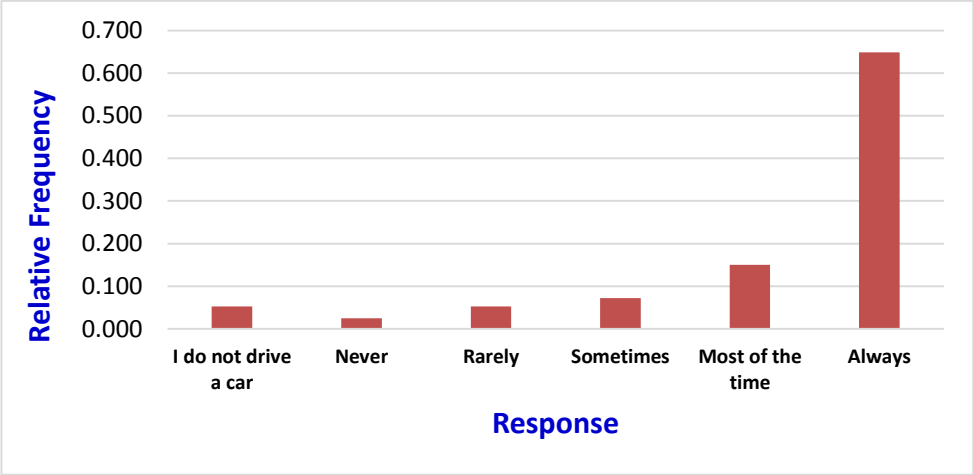
b) $0.6484 = 64.84\%$

c) $0.0247 + 0.0522 = 0.0769$ 7.69%

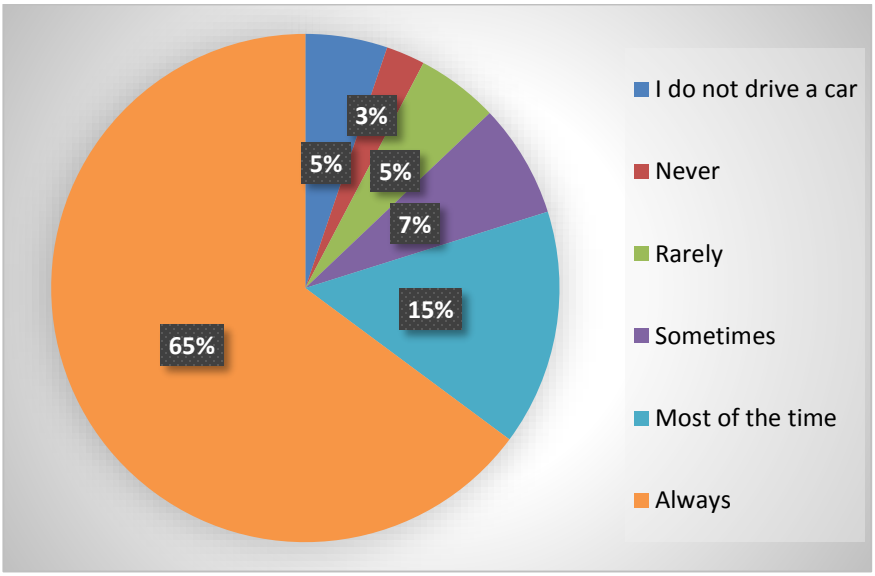
d) Frequency bar graph



e) Relative frequency bar graph



f) How often do you wear a seat belt when driving a car?



g) The statement is descriptive because it is describing the particular sample.

Exercise

A phlebotomist draws the blood of a random sample of 50 patients and determines their blood types as shown.

O	*B*	*AB*	*O*	*AB*	*O*	*O*	*O*	*O*	*O*
O	*O*	*B*	*O*	*O*	*A*	*A*	*B*	*O*	*A*
A	*B*	*A*	*A*	*A*	*A*	*O*	*A*	*O*	*O*
A	*A*	*B*	*A*	*B*	*O*	*AB*	*A*	*A*	*A*
O	*O*	*AB*	*O*	*A*	*A*	*A*	*O*	*O*	*O*

a) Construct a frequency distribution
b) Construct a relative frequency distribution
c) According to the data, which blood type is most common?
d) According to the data, which blood type is least common?
e) Use the results of the sample to conjecture the percentage of the population that has type *O* blood. Is this an example of descriptive or inferential statistics?
f) Contact a local hospital and ask them the percentage of the population that us blood type *O*. Why might the results differ?
g) Draw a frequency bar graph
h) Draw a relative frequency bar graph
i) Draw a pie chart

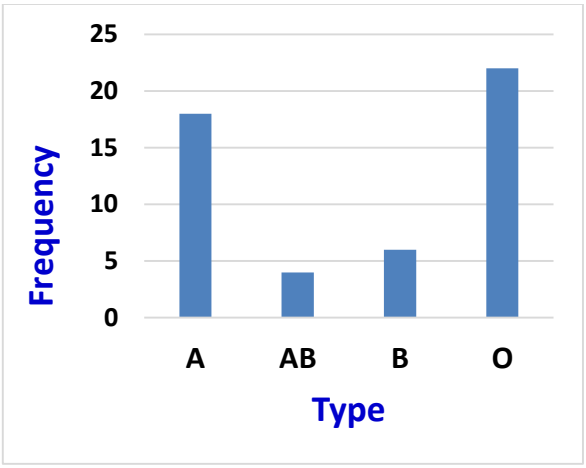
Solution

a)
b)

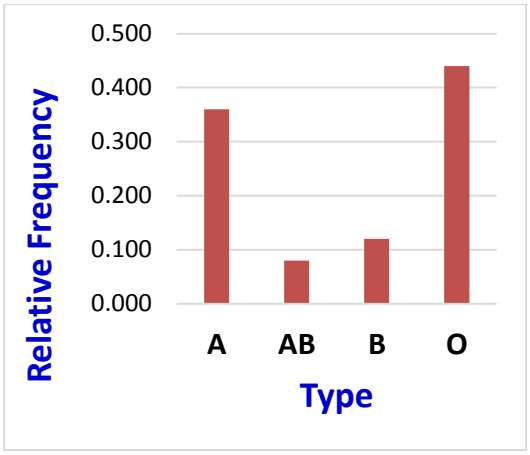
Blood Type	***Frequency***	***Relative Frequency***	***Percentage***
A	18	0.360	36%
AB	4	0.080	8%
B	6	0.120	12%
O	22	0.440	44%

c) Type *O* is the most common
d) Type *AB* is the least common
e) We estimate that 44% of the population has type *O* blood. This is considered inferential statistics because a conclusion about the population is being drawn based on sample data.
f) Answers will vary; in 2008 the Red Cross reported that 45% of the population had type *O* blood (either + or –). Results will differ because of sampling variability.

g)



h)



i) Blood Types