AP Stats #1 | 1.1, 2, 3, 14

1.1 Pulse rate data

a) class members (male/female)
b) two variables: rate, gender
c) beats per minute (rate)
male or female (gender)
d) rate is quantitative
gender is categorical

1.2 Region is categorical
All others are quantitative

1.3

Categorical	Quantitative
gender	age
race	blood pressure
smoker	calcium level

1.14 a) employees of the company (names)
b) categorical: gender, race, job type
c) quantitative: age, salary
units years $ per year

AP Stats 1.9, 10, 12

2

1.9

10	139
11	5
12	69
13	77
14	08
15	244
16	55
17	8
18	
19	
20	0

200 seems to be an outlier.

The center seems to be between 137 and 140 (if we exclude the outlier)

The spread is from 101 to 178 if we exclude the outlier.

17|8 means 178

1.10

1	0
1	5
2	00233 44
2	59
3	000112 22
3	

The distribution is skewed left (toward the smaller values)

The center is 25.

The are no outliers.

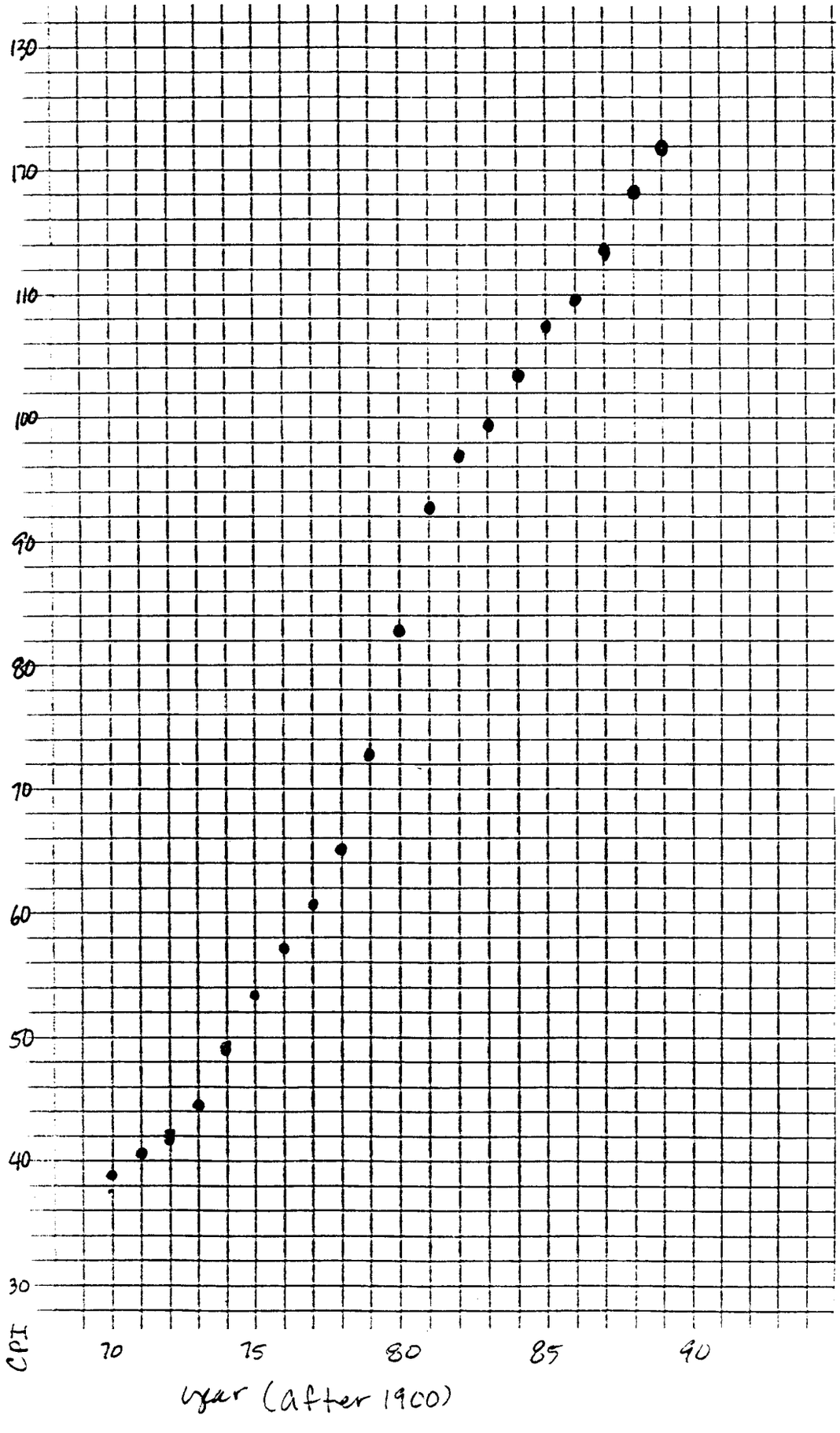
1|5 means 15

1.12

a) see graph b) try this on the calculator

c) The overall trend is up (increasing). The CPI never declined during the years given.

d) The fastest increase was between 1978 and 1982 (the slope is steepest). The slowest increase was from 1970 to 1974


130
120
110
100
90
80
70
60
50
40
30
CPI
70
75
80
85
90
Year (after 1900)

AP Stats | 1.25 - 1.30 |

1.25 a) $\frac{154+109+137+\cdots+148}{18} = \frac{2539}{18} = \boxed{141.058 = \bar{x}}$

b) $\bar{x} = \frac{2339}{17} = 137.588$ This is closer to the "center" found in 1.9. The outlier makes the mean higher than it "should" be.

1.26 if $\bar{x} = 1.2$ million and $n = 25$
The annual payroll is $25(1.2 \text{ mil}) = \boxed{\$30 \text{ million}}$.
If you knew the median you would not be able to calculate the total salary.

1.27 Ruth med. = 46 Maris med. = 24.5

1.28 $M = 138.5$ $\bar{x} = 141.058$ The median is smaller than the mean because the outlier makes the distribution skewed to the right. The outlier "pulls" the mean toward itself.

1.29 $\bar{x} = \frac{480000}{8} = 60000$ $M = 22000$

Seven of the 8 employees (everyone but the owner) earned less than the mean. A recruiter could say that the average salary at this company is \$60,000. That would be true, but no one but the owner actually makes that much money.

1.30 \$490000 is the median
\$1,160,000 is the mean

The distribution would be skewed to the right since a few players get really high salaries. In that situation the mean would be "pulled" to the right and would be larger than the median.

AP Stats #3 | 1.5, 6, 7, 8

1.5 a) The shape is almost symmetric
b) The center seems to be between 1 and 20 (≈ 15
c) The smallest is between -60% and -70%.
The largest is between 100% and 110%.
d) Add the bars to the left of zero. (11+5+3+1+1+1+1)
The total is approximately 23%

1.6 a) The distribution is skewed to the right. The center is about 3 (31 are less than 3, 11 equal to 3, 23 more than 3). The spread is 0 to 10 with no outliers.
b) 15 out of 65 years the number of days equals zero. $15/65 \approx .23$ so approximately **23%**

1.7 Lightning: Center - around noon (between 11:30 - 12:30)
Spread - 7 to 17 (about 6:30 am to 17:30 or 5:30 pm)
Shakespear: Center - 4 (45 less than 4, 32 more, 24 equal to 4)
Spread - 1 to 12

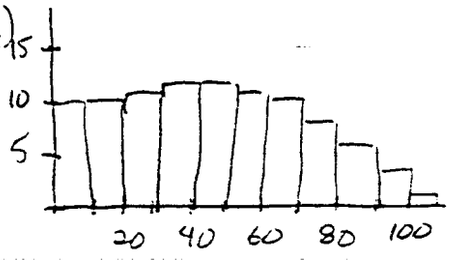
1.8 a)

Age	1950	2075
0-9	19.4%	11.2%
10-19	14.4	11.5
20-29	15.9	11.8
30-39	15.1	12.3
40-49	12.8	12.2
50-59	10.3	12.1
60-69	9.3	11.1
70-79	3.6	8.8
80-89	1.1	6.1
90-99	0.1	2.5
100-109	0.0	0.5

b)

Children under 10 represent the single largest group. About 1 out of 5 (20%) Americans was under 10 in 1950. There is a slight dip in the 10-19 age group. The percentages decline gradually after 20-29.

1.8 c)



The projections show a much higher proportion of persons in the higher age brackets. There is a gradual rise up to age 40-49 followed by a decline.

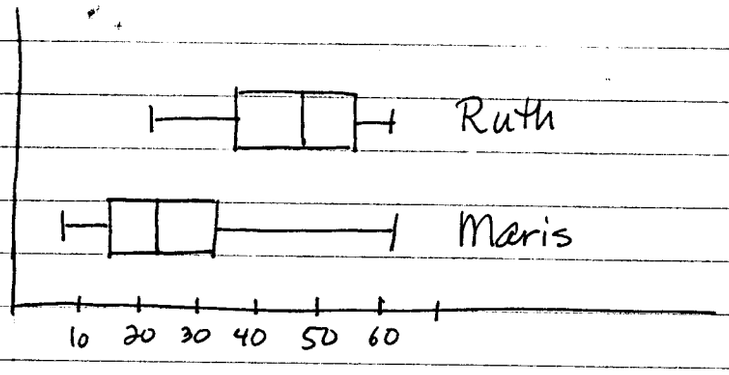
AP Stats 1.31–34, 39, 41

1.31 The IQR is a resistant measure of spread.
For example: {1, 2, 3, 4, 5, 6} $\bar{x} = 3.5$ $M = 3.5$
$Q_1 = 2$ $Q_3 = 5$
$IQR = 5 - 2 = 3$

If the data were changed so that the numbers were {1, 2, 3, 4, 5, 100}
$\bar{x} = 25.17$ $M = 3.5$ $Q_1 = 2$ $Q_3 = 5$ $IQR = 5 - 2 = 3$

The mean is greatly affected. The median and IQR are not changed.

1.32 Ruth 22 35 46 54 60
Maris 8 14 24.5 33 61

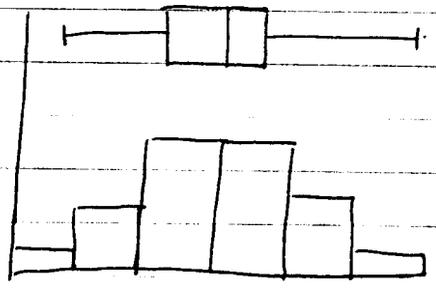


Babe Ruth was a much better hitter, overall, than Roger Maris. Maris' 61 homerun season was unusual for him.

1.33 a) It seems that M should be approximately the same as $\bar{x}$ since there is no particular skewness.
b) 42 51 55 58 69 $\bar{x} = 54.833$

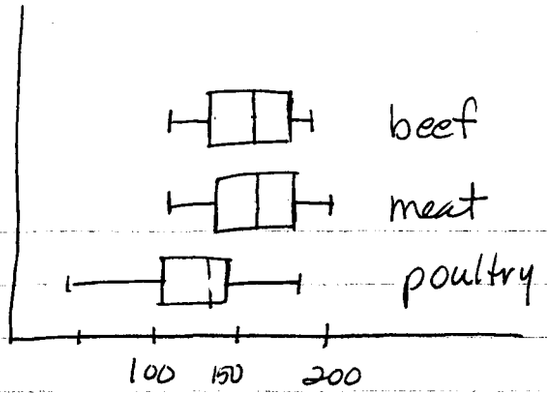
c) $IQR = 58 - 51 = 7$ $1.5(IQR) = 10.5$ $58 + 10.5 = 68.5$

d)



69 (Ronald Reagan) is an outlier. Harrison was 68 but $1.5(IQR) + 58 = 10.5 + 58$ gives 68.5 so he fits just within the guideline.

1.39



There seems to be little difference between beef and meat hot dogs. Poultry dogs are generally lower in calories. The median of poultry dogs is less than the lower quartiles of the other two. The poultry lower quartile is less than the minimum of the other types.

1.41

48	8
49	
50	7
51	0
52	6 7 9 9
53	0 4 4 6 9
54	2 4 6 7
55	0 3 5 7 8
56	1 2 3 5 8
57	5 9
58	5

48|8 means 4.88

5 number summary

4.88 5.295 5.46 5.615 5.85

According to the 1.5 IQR rule, there are no outliers.

$Q_1 - 1.5(IQR) = 5.295 - 1.5(0.32) = 4.815$

The $\bar{x}$ and s are reasonable in this setting.

$\bar{x} = 5.4479$ $s = 0.22095$

The mean ($\bar{x}$) is our best estimate of the earth's density.

APStats 1.42-46, 48, 49

1.42 Since there are two definite outliers (Alaska and Florida), the five-number summary is preferrable.

4.2 11.35 12.65 13.7 18.3

$\bar{x} = 12.544$ $s = 2.121$

1.43

0	1111111112224
0	5668
1	0
1	5
2	1133
2	
3	03
3	
4	0
4	
5	
5	9
6	2

The distribution is clearly skewed to the right. The two or three are outliers.

The 5 number summary is

109 158 635 2300 6200

1.44 $M = 8000$ $\bar{x} = 69000$ The distribution must be skewed sharply to the right. The must have been some very large awards.

1.45 a) $\bar{x}$ Incomes are likely to be skewed, but you still need the mean to estimate the total amount of taxes.

b) M A "typical" income would be the median because it is not affected by the extreme high or low values.

1.46 You have the median (actually a median ... it could be any of several values between the high and low speeds).

1.48 a) We expect heart attacks to occur randomly so each day is equally likely. The graph shows some variation, but, in general, it supports our thinking.

b) Mon.–Thurs. are similar, but Fri. is much higher and Sat. and Sun. are lower. Patients do have some choice about when they go home so they may choose to leave on Fri. to spend the weekend at home. Also, hospitals may choose to discharge as many patients as possible so they need less staffing on the weekend.

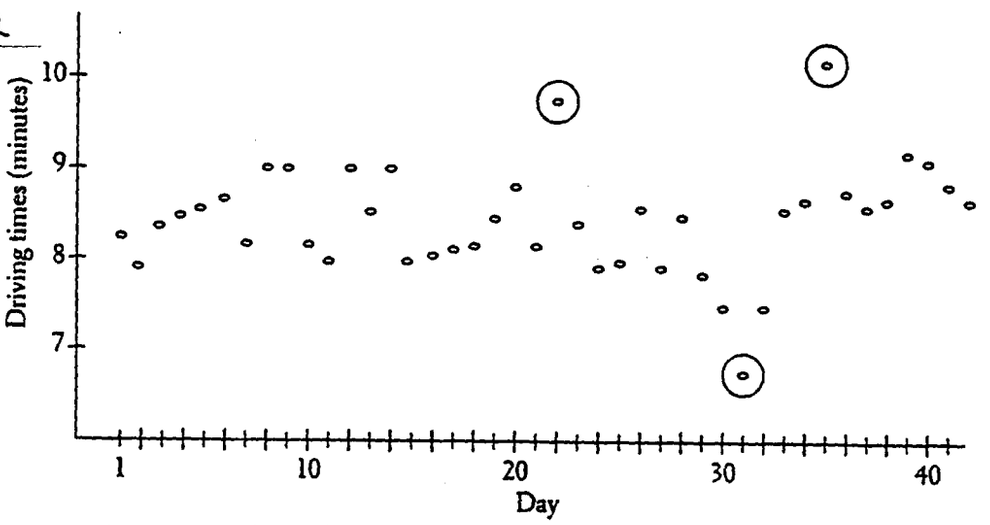
1.49

6	8
7	44
7	88888999
8	01122333 444
8	55566777788 8
9	000012
9	8
10	2

a) Symmetric – no outliers

b) Outliers show on second graph.

c) $\bar{x} = 8.3628$ $s = 0.4645$



1.54 a)

	min	Q_1	M	Q_3	max
	-34.04	-2.95	3.47	8.45	58.68

b) best 58.68% $1000(.5868) = 586.8$
$1000 + 586.80 = \underline{\$1586.80}$

worst -34.04% $1000(-.3404) = -340.4$
$1000 - 340.40 = \underline{\$659.60}$

c) $IQR = 8.45 - (-2.95) = 11.401358$
$1.5(IQR) = 17.102037$
$Q_1 - 17.1 = -2.95 - 17.1 = \boxed{-20.05}$

$Q_3 + 17.1 = 8.45 + 17.1 = \boxed{25.55}$

It seems that SPLUS uses the 1.5 IQR rule.

d) The distribution is fairly symmetric with a single peak in the middle (5 to 9). There are four outliers below -20.05 and five above 25.55. The mean, 3.064, is slightly less than the median, 3.4691. This suggests a slight skew to the left.

55 a) After the first two years the median is above zero every year except 1984. There doesn't seem to be a trend in the typical monthly return.

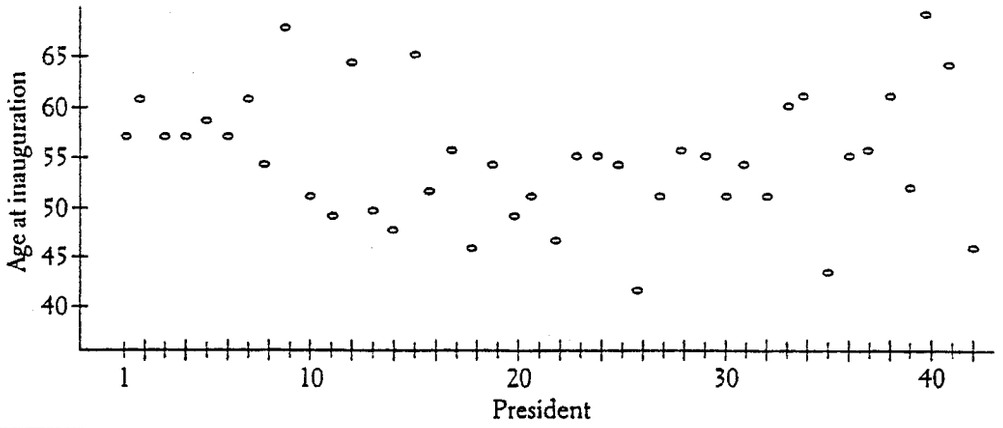
b) The spread is smaller in the later years (except 1987). than it was earlier.

c) We can see four of the high outliers 58.7 in 1973, 57.9 in 1975, 42 or 41 in 1974, and 32 in 1979. The other one (42 or 41) must have been in '73, '74 or '75.

1.55

c) The low outliers -34.04 in 1973, -31.25 must be also in 1973, -27.06 in 1987 (or -26.6). The other one must have been included in 1973 or 1987. This supports the conclusion that the outliers occurred in the earlier years. The low in 1987 is a deviation from the pattern of less spread in later years.

1.56



There is no apparent trend to support either position. There seems to be a greater spread in ages in recent years.