IQ Scores and Making a Histogram

You have probably heard that the distribution of IQ scores follows a bell shaped pattern. Here are some actual IQ scores from 60 5th grade students. The students were chosen at random from one school.

145	139	126	122	125	130	96	110	118	118
101	142	134	124	112	109	134	113	81	113
123	94	100	136	109	131	117	110	127	124
106	124	115	133	116	102	127	117	109	137
117	90	103	114	139	101	122	105	97	89
102	108	110	128	114	112	114	102	82	101

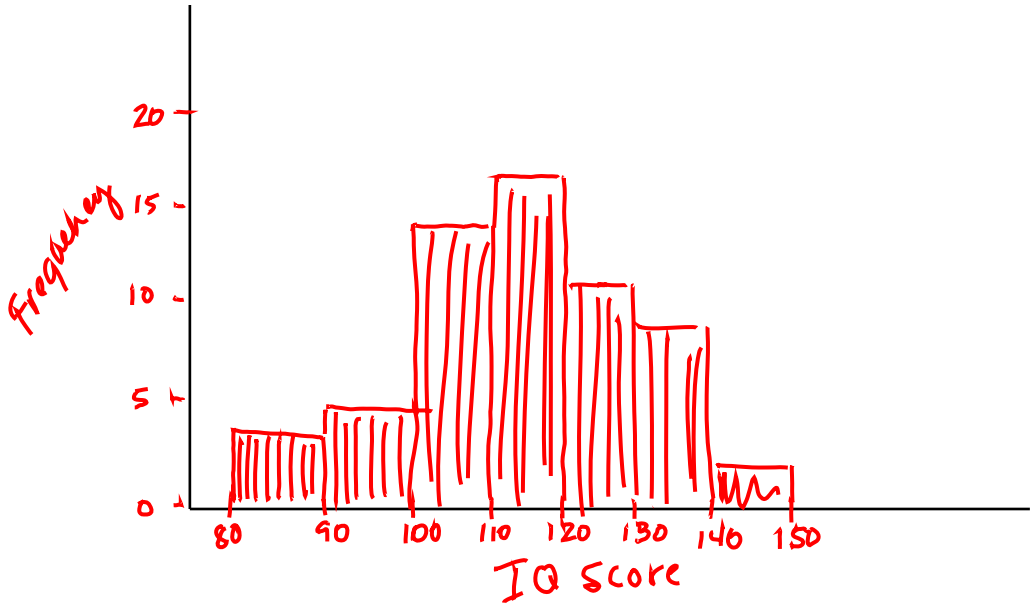
Divide the range of data into classes of equal width. Low 81 and high 145
80-89
90-99
100-109
110-119
120-129
130-139
140-149

Be sure to specify the class widths so that each individual falls into exactly one class.

Create a frequency table for all of the classes.

Class	Count	Class	Count
80-89	3	120-129	11
90-99	4	130-139	9
100-109	14	140-149	2
110-119	17		

Draw your histogram and label the axes.



Tips for Histograms

Be sure to choose class widths that are equal width.
There is no one right choice for the number of class. Five is a good minimum.

Histogram Versus Bar Graphs

Histogram shows: distribution of counts or percents of a single quantitative variable

Bar Graph shows: categorical variable on the x axis, shows comparison between variables, not a distribution

Examining Distributions

G AP

S HAPE

O UTLIER

C ENTER

S PREAD

Relative Frequency and Cumulative Frequency

Ogive: relative cumulative frequency graph

How to construct an ogive—

1. Decide on class intervals and make a frequency table, but add three new columns: Relative Frequency, Cumulative Frequency and Relative Cumulative Frequency.

Relative Frequency = count/total

Cumulative Frequency = total count so far

Relative cumulative frequency = cumulative frequency/total

2. Label and scale your axes for your graph. Age at inauguration is the x and cumulative relative frequency is the y.

3. Plot a point corresponding to the relative cumulative frequency in each class interval at the LEFT endpoint of the NEXT class interval. For example, for the 40 to 44 interval, plot a point at 4.7% above the age value of 45. Begin the ogive at 0% above the left endpoint of the first class interval.

Class	Frequency	Relative Frequency	Cumulative Frequency	Relative Cumulative Frequency
40-44	2	2/43 4.7 %	2	2/43 4.7%
45-49	6	6/43 14.0%	8	8/43 18.6%
50-54	13	13/43 30.2%	21	21/43 48.8%
55-59	12	12/43 27.9%	33	33/43 76.7%
60-64	7	7/43 16.3%	40	40/43 93.0%
65-69	3	3/43 7.0%	43	43/43 100%

Time plot:

Plots each observation against the time at which it was observed. Time is always on the horizontal axis and measured variable is on the vertical axis.

Average price for a gallon of gasoline.

09/15/1981 1.471
10/15/1981 1.470
11/15/1981 1.470
12/15/1981 1.468
01/15/1982 1.466
02/15/1982 1.448
03/15/1982 1.408
04/15/1982 1.351
05/15/1982 1.355
06/15/1982 1.418
07/15/1982 1.443
08/15/1982 1.439
09/15/1982 1.429
10/15/1982 1.421
11/15/1982 1.412
12/15/1982 1.394
01/15/1983 1.376
02/15/1983 1.338
03/15/1983 1.308
04/15/1983 1.360
05/15/1983 1.397
06/15/1983 1.411
07/15/1983 1.421
08/15/1983 1.419
09/15/1983 1.410
10/15/1983 1.395
11/15/1983 1.384
12/15/1983 1.376
01/15/1984 1.369
02/15/1984 1.361
03/15/1984 1.362
04/15/1984 1.375
05/15/1984 1.380
06/15/1984 1.377
07/15/1984 1.370
08/15/1984 1.355
09/15/1984 1.360
10/15/1984 1.365
11/15/1984 1.364
12/15/1984 1.354
01/15/1985 1.304
02/15/1985 1.290
03/15/1985 1.310
04/15/1985 1.340
05/15/1985 1.360
06/15/1985 1.371
07/15/1985 1.367
08/15/1985 1.359
09/15/1985 1.349
10/15/1985 1.342
11/15/1985 1.339
12/15/1985 1.344
01/15/1986 1.336
02/15/1986 1.282
03/15/1986 1.160
04/15/1986 1.061
05/15/1986 1.075
06/15/1986 1.100
07/15/1986 1.045
08/15/1986 0.999
09/15/1986 1.010
10/15/1986 0.987
11/15/1986 0.980
12/15/1986 0.984
01/15/1987 1.007
02/15/1987 1.047
03/15/1987 1.052
04/15/1987 1.073
05/15/1987 1.079
06/15/1987 1.098
07/15/1987 1.115
08/15/1987 1.139
09/15/1987 1.136
10/15/1987 1.128
11/15/1987 1.125
12/15/1987 1.119
01/15/1988 1.095
02/15/1988 1.082
03/15/1988 1.074
04/15/1988 1.088
05/15/1988 1.105
06/15/1988 1.111
07/15/1988 1.123
08/15/1988 1.138
09/15/1988 1.130
10/15/1988 1.119
11/15/1988 1.116
12/15/1988 1.101
01/15/1989 1.091
02/15/1989 1.100
03/15/1989 1.115
04/15/1989 1.221
05/15/1989 1.278
06/15/1989 1.278
07/15/1989 1.264
08/15/1989 1.233
09/15/1989 1.213
10/15/1989 1.209
11/15/1989 1.187
12/15/1989 1.170
01/15/1990 1.230
02/15/1990 1.227
03/15/1990 1.218
04/15/1990 1.233
05/15/1990 1.248
06/15/1990 1.271
07/15/1990 1.272
08/15/1990 1.369
09/15/1990 1.467
10/15/1990 1.554
11/15/1990 1.559
12/15/1990 1.537
01/15/1991 1.431
02/15/1991 1.321
03/15/1991 1.264
04/15/1991 1.281
05/15/1991 1.331
06/15/1991 1.338
07/15/1991 1.313
08/15/1991 1.318
09/15/1991 1.324
10/15/1991 1.307
11/15/1991 1.318
12/15/1991 1.309
01/15/1992 1.267
02/15/1992 1.248
03/15/1992 1.250
04/15/1992 1.268
05/15/1992 1.317
06/15/1992 1.359
07/15/1992 1.362
08/15/1992 1.348
09/15/1992 1.346
10/15/1992 1.345
11/15/1992 1.351
12/15/1992 1.330
01/15/1993 1.313
02/15/1993 1.301
03/15/1993 1.294
04/15/1993 1.304
05/15/1993 1.319
06/15/1993 1.321
07/15/1993 1.305
08/15/1993 1.294
09/15/1993 1.282
10/15/1993 1.323
11/15/1993 1.305
12/15/1993 1.268
01/15/1994 1.240
02/15/1994 1.245
03/15/1994 1.243
04/15/1994 1.260
05/15/1994 1.274
06/15/1994 1.300
07/15/1994 1.327
08/15/1994 1.367
09/15/1994 1.364
10/15/1994 1.345
11/15/1994 1.354
12/15/1994 1.337
01/15/1995 1.324
02/15/1995 1.316
03/15/1995 1.306
04/15/1995 1.325
05/15/1995 1.383
06/15/1995 1.411
07/15/1995 1.384
08/15/1995 1.352
09/15/1995 1.332
10/15/1995 1.315
11/15/1995 1.292
12/15/1995 1.290
01/15/1996 1.317
02/15/1996 1.311
03/15/1996 1.348
04/15/1996 1.431
05/15/1996 1.507
06/15/1996 1.481
07/15/1996 1.453
08/15/1996 1.421
09/15/1996 1.417
10/15/1996 1.408
11/15/1996 1.428
12/15/1996 1.438
01/15/1997 1.441
02/15/1997 1.434
03/15/1997 1.415
04/15/1997 1.413
05/15/1997 1.409
06/15/1997 1.411
07/15/1997 1.388
08/15/1997 1.433
09/15/1997 1.458
10/15/1997 1.426
11/15/1997 1.397
12/15/1997 1.363
01/15/1998 1.319
02/15/1998 1.271
03/15/1998 1.229
04/15/1998 1.237
05/15/1998 1.275
06/15/1998 1.279
07/15/1998 1.268
08/15/1998 1.244
09/15/1998 1.230
10/15/1998 1.236
11/15/1998 1.225
12/15/1998 1.187
01/15/1999 1.171
02/15/1999 1.155
03/15/1999 1.186
04/15/1999 1.367
05/15/1999 1.370
06/15/1999 1.339
07/15/1999 1.378
08/15/1999 1.441
09/15/1999 1.468
10/15/1999 1.464
11/15/1999 1.454
12/15/1999 1.486
01/15/2000 1.486
02/15/2000 1.551
03/15/2000 1.723
04/15/2000 1.698
05/15/2000 1.682
06/15/2000 1.786
07/15/2000 1.773
08/15/2000 1.689
09/15/2000 1.764
10/15/2000 1.744
11/15/2000 1.738
12/15/2000 1.679
01/15/2001 1.657
02/15/2001 1.671
03/15/2001 1.638
04/15/2001 1.748
05/15/2001 1.934
06/15/2001 1.881
07/15/2001 1.695
08/15/2001 1.636
09/15/2001 1.726
10/15/2001 1.560
11/15/2001 1.427
12/15/2001 1.312
01/15/2002 1.323
02/15/2002 1.330
03/15/2002 1.450
04/15/2002 1.622
05/15/2002 1.625
06/15/2002 1.606
07/15/2002 1.607
08/15/2002 1.620
09/15/2002 1.619
10/15/2002 1.643
11/15/2002 1.643
12/15/2002 1.589
01/15/2003 1.666
02/15/2003 1.828
03/15/2003 1.924
04/15/2003 1.846
05/15/2003 1.729
06/15/2003 1.700
07/15/2003 1.710
08/15/2003 1.808
09/15/2003 1.911
10/15/2003 1.789
11/15/2003 1.724
12/15/2003 1.686
01/15/2004 1.779
02/15/2004 1.858
03/15/2004 1.949
04/15/2004 2.012
05/15/2004 2.186
06/15/2004 2.225
07/15/2004 2.130
08/15/2004 2.091
09/15/2004 2.082
10/15/2004 2.215
11/15/2004 2.203
12/15/2004 2.080
01/15/2005 2.017
02/15/2005 2.105
03/15/2005 2.251
04/15/2005 2.468
05/15/2005 2.403
06/15/2005 2.365
07/15/2005 2.502
08/15/2005 2.701
09/15/2005 3.130
10/15/2005 3.001
11/15/2005 2.560
12/15/2005 2.393
01/15/2006 2.521
02/15/2006 2.519
03/15/2006 2.603
04/15/2006 2.967
05/15/2006 3.169
06/15/2006 3.139
07/15/2006 3.219
08/15/2006 3.207
09/15/2006 2.819
10/15/2006 2.493
11/15/2006 2.459
12/15/2006 2.550
01/15/2007 2.501
02/15/2007 2.509
03/15/2007 2.818
04/15/2007 3.093
05/15/2007 3.348
06/15/2007 3.281
07/15/2007 3.200
08/15/2007 3.018
09/15/2007 3.021
10/15/2007 3.037
11/15/2007 3.307
12/15/2007 3.264
01/15/2008 3.291
02/15/2008 3.272
03/15/2008 3.502
04/15/2008 3.690
05/15/2008 4.003
06/15/2008 4.319
07/15/2008 4.350
08/15/2008 4.045
09/15/2008 3.940
10/15/2008 3.432
11/15/2008 2.433
12/15/2008 1.951
01/15/2009 2.036
02/15/2009 2.182
03/15/2009 2.197
04/15/2009 2.309

Make a time plot of the average price of a gallon of gasoline using the month in which you were born from 1981 (if possible) to 2009 (if possible).

Assignment: pgs. 55-58 1.7, 1.9, 1.12 pgs. 64-66 1.13, 1.15, 1.16

Measuring Center

Mean: $\bar{x}$ x bar $\bar{x} = \frac{\sum x_i}{n}$

Median: midpoint, middle #, 50th %tile

1. order from least to greatest

2. odd number the median is middle number

3. even number the median is average of the two middle numbers

Measuring Spread

Quartiles:

Q1: 1st quartile, 25th %tile, median of the lower half of data

Q3: 3rd quartile, 75th %tile, median of the upper half of data

Five Number Summary and Box Plots

Minimum Q1 Median Q3 Maximum

Box Plot: a visual representation of the 5 number summary

IQR: Inter Quartile Range (Q3-Q1) Measure of Spread

Outlier Rule: Q3 + 1.5 * IQR and Q1 – 1.5*IQR Any value that lies outside those numbers are outliers

Assignment: pgs. 74-75 1.27,28,29,30 pgs. 82-83 1.33, 34

Measuring Spread

Variance s^2 : average of the squares of the deviations of the observations from their mean.

$$s^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{n-1}$$

$$s^2 = \frac{1}{n-1}\sum (x_i-\bar{x})^2$$

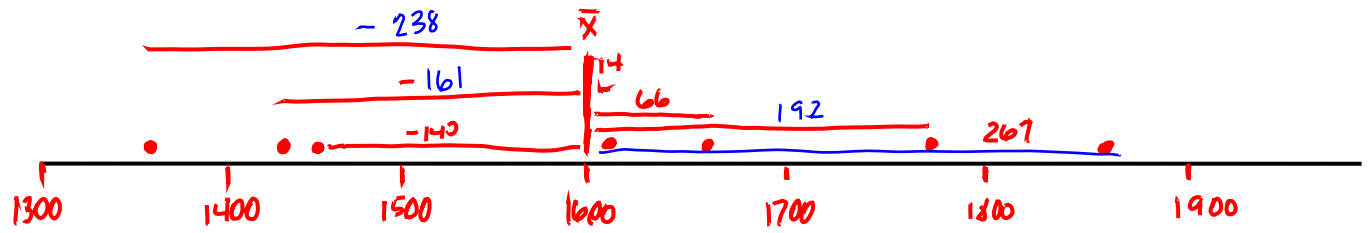
Standard Deviation: the square root of the variance.

$$s = \sqrt{\frac{1}{n-1}\sum (x_i-\bar{x})^2}$$

Here are seven metabolic rates from men who took part in a study concerning dieting.

1792	1666	1362	1614	1460	1867	1439

Calculate the mean: $\bar{x} = 1600$ $s = 189.24$



Properties of the standard deviation:

1. s measures spread about the mean and should only be used when the mean is the measure of the center.
2. s = 0 only when there is no spread or variability. This only happens when all observations are the same value.
3. s, like the mean, is not resistant. A few outliers can make s very large.

Choosing a Summary:

The five number summary is usually best for describing a skewed distribution. Use mean and standard deviation only when the distribution is reasonably symmetric and free of outliers.

Investment	Mean Return	Standard Deviation
Common Stocks	13.2%	17.6%
Treasury Bills	5.0%	2.9%

Assignment: p.89-90 1.39 to 1.44

ANDERSON COUNTY SCHOOLS CERTIFIED SALARY SCHEDULE 2009-2010

Years Experience	RANK III 187 Days	RANK II 187 Days	RANK I 187 Days	Doctorate 187 Days	RANK IV 187 Days
0	34,869	38,458	42,048	45,413	34,869
1	35,459	39,151	42,843	46,145	
2	36,049	39,843	43,638	46,877	
3	36,638	40,535	44,432	47,609	RANK V
4	37,228	41,227	45,227	48,341	187 Days
5	37,818	41,920	46,022	49,074	26,243
6	38,408	42,612	46,817	49,806	
7	38,996	43,304	47,611	50,538	
8	39,586	43,996	48,406	51,271	
9	40,176	44,689	49,201	52,003	
10	40,766	45,381	49,996	52,735	
11	41,355	46,073	50,790	53,517	
12	41,945	46,765	51,585	54,785	
13	42,535	47,458	52,380	55,203	
14	43,125	48,150	53,175	55,664	
15	43,714	48,842	53,970	56,396	
16	44,304	49,534	54,765	57,128	
17	44,894	50,227	55,560	57,861	
18	45,484	50,919	56,355	58,593	
19	46,073	51,611	57,149	59,325	
20	46,663	52,303	57,944	60,058	
21	47,253	52,996	58,739	60,790	
22	47,843	53,688	59,534	61,522	
23	48,432	54,380	60,328	62,254	
24	49,021	55,072	61,123	62,986	
25	49,611	55,765	61,918	63,718	
26	50,201	56,457	62,713	64,451	
27	50,790	57,149	63,507	65,184	
28	51,380	57,841	64,302	65,916	
29	51,970	58,534	65,097	66,648	
30	52,560	59,226	65,892	67,380	

Using the Rank III column, find the mean, standard deviation and 5 number summary for the salary distribution.

Suppose each teacher in the Rank III column receives a raise of $3000. How will this change the shape, center and spread of the distribution? Give the mean, standard deviation and 5 number summary for the new distribution. Compare to the original.

Suppose that instead of a $3000 raise, the teachers in the Rank III column were given a 5% raise. How will this change the shape, center and spread of the distribution? Give the mean, standard deviation and 5 number summary for the new distribution. Compare to the original.

On the axes below, draw three box plots. One for each of the distributions given above. Make sure you use the same scale for each.

How do the box plots compare?

Linear transformations:

- Multiplying each observation by a positive number b multiplies measures of center and measures of spread by b.
- Adding the same number a to each observation adds a to measures of center and to quartiles, but does not change measures of spread.

Assignment: p. 97-99 1.47 to 1.50