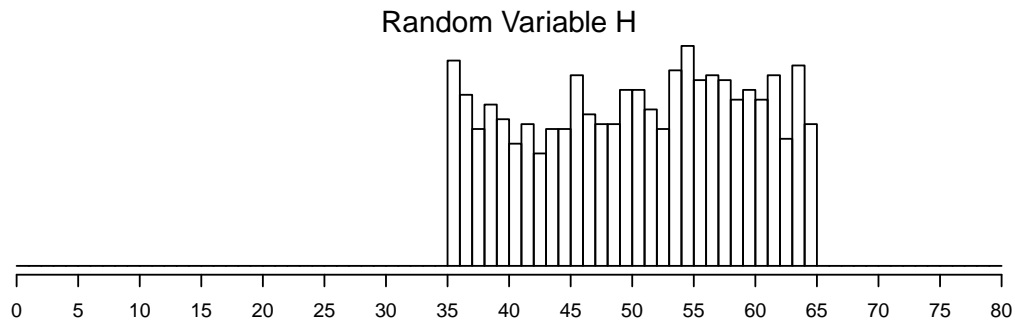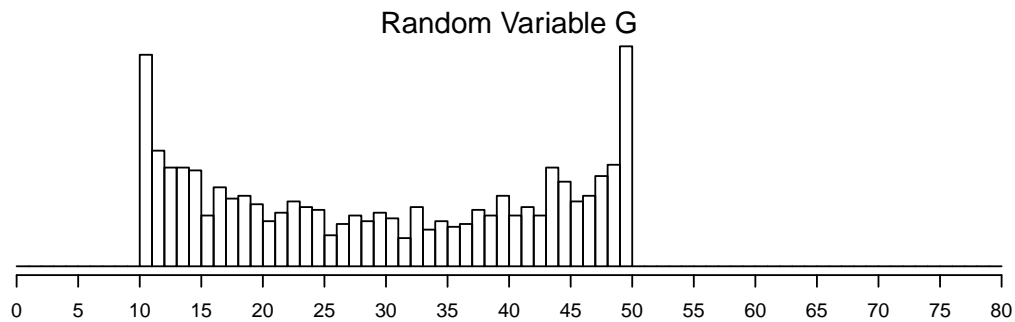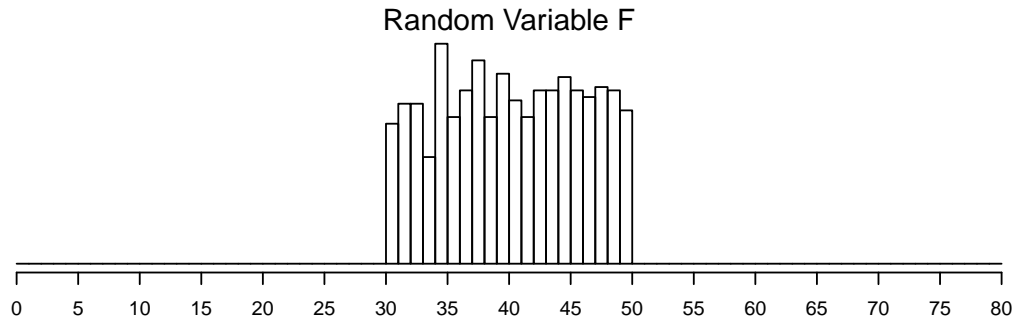1. **Problem**

   Three random variables (F, G, and H) were measured 1000 times each. The resulting histograms show the three distributions.

Random Variable F

Random Variable G

Random Variable H

   (a) Which distribution has the highest mean? (F, G, or H)

   (b) Which distribution has the lowest mean? (F, G, or H)

   (c) Which distribution has the largest standard deviation? (F, G, or H)

   (d) Which distribution has the smallest standard deviation? (F, G, or H)

2. **Problem**
From a very large population, a small sample of measurements was taken.

54	57	34	72	47	60

Please calculate the Average Absolute Deviation (also called the Average Distance from Mean: ADM) using the following formula:

$$\text{AAD} = \frac{\sum |x - \bar{x}|}{n}$$

3. **Problem**
   A continuous random variable was measured 300 times. The resulting frequency distribution is shown below.

class	frequency
10–11	10
11–12	39
12–13	28
13–14	26
14–15	28
15–16	29
16–17	32
17–18	31
18–19	29
19–20	34
20–21	14

   (a) Describe the overall shape of the distribution. (symmetric mound, skew left, skew right, uniform, or bimodal)

   (b) Estimate the range of the distribution (range = max-min).

   (c) What percent of the measurements are less than 12?

   (d) What percent of the measurements are greater than 11?

   (e) What percent of the measurements are between 11 and 12?

   (f) What percent of the measurements are within 1.5 of 13.5? In other words, what percent of measurements satisfy $|x - 13.5| \leq 1.5$?

   (g) Of the measurements less than 12, what percent are greater than 11?

   (h) Estimate the value of the 25.67th percentile. In other words, determine a value such that 25.67% of the measurements are less than or equal to it.

4. **Problem**
   A continuous random variable was measured 200 times. The resulting histogram is shown below.

Histogram of data
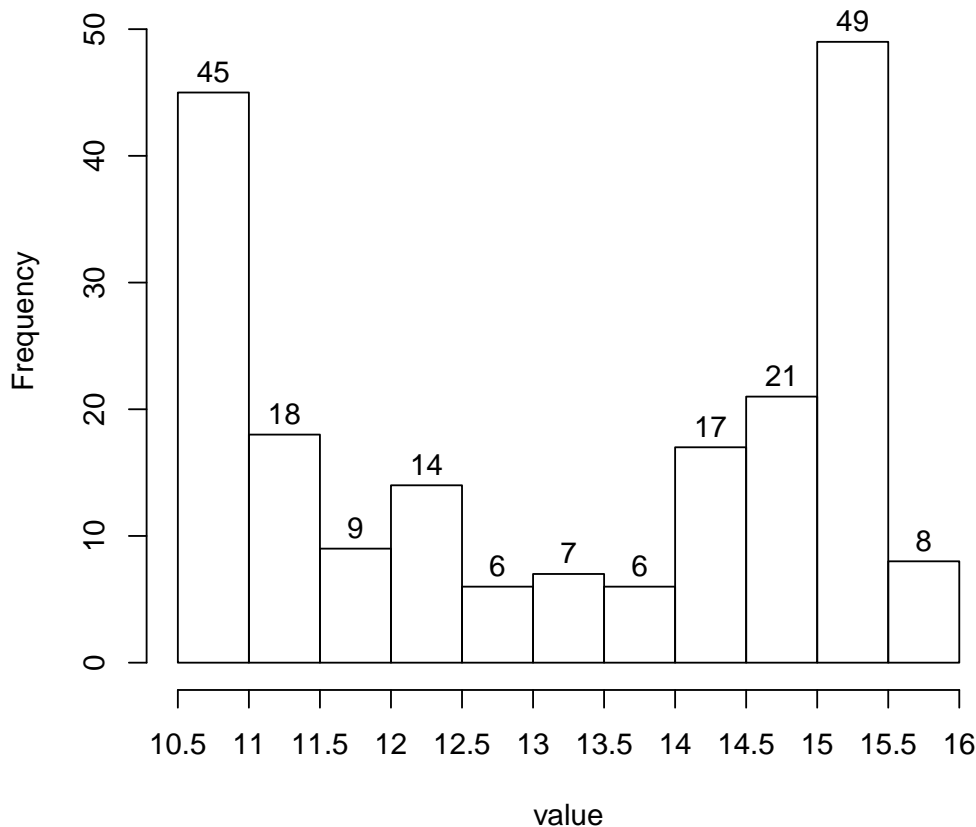
(a) Describe the overall shape of the distribution. (symmetric mound, skew left, skew right, uniform, or bimodal)

(b) Estimate the range of the distribution (range = max-min).

(c) What percent of the measurements are less than 13?

(d) What percent of the measurements are less than 12?

(e) What percent of the measurements are between 12 and 13?

(f) What percent of the measurements are within 0.25 from 13.25? In other words, what percent of measurements satisfy $|x - 13.25| \leq 0.25$?

(g) Of the measurements less than 13, what percent are less than 12?

(h) Estimate the value of the 61th percentile. In other words, determine a value such that 61% of the measurements are less than or equal to it.
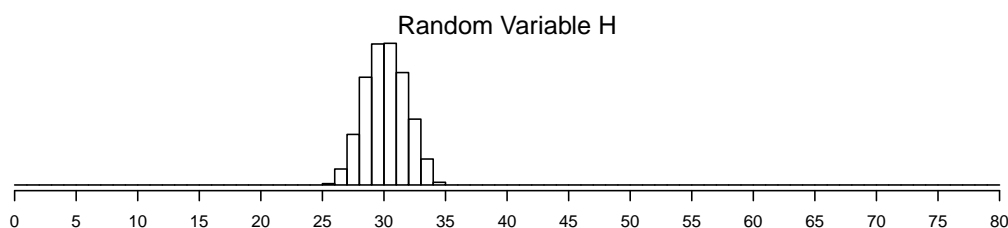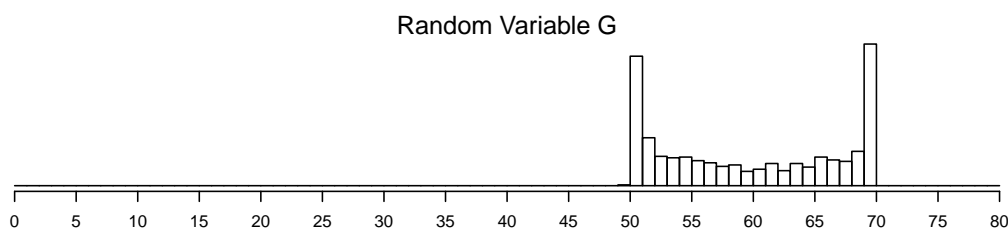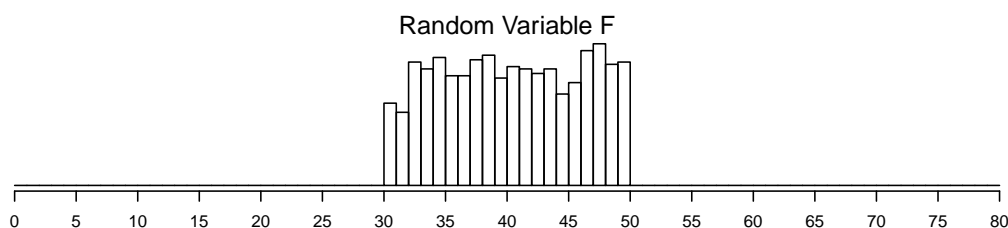
5. **Problem**
   We can estimate the mean of **symmetric** distributions.

$$\bar{x} \approx \frac{\max(x) + \min(x)}{2}$$

We can **roughly** estimate the standard deviation of certain distributions.

Shape	SD estimate
bell	range/6
uniform	range/4
bimodal	range/2

Three random variables (F, G, and H) were measured 1000 times each. The resulting histograms show the three distributions.

Random Variable F

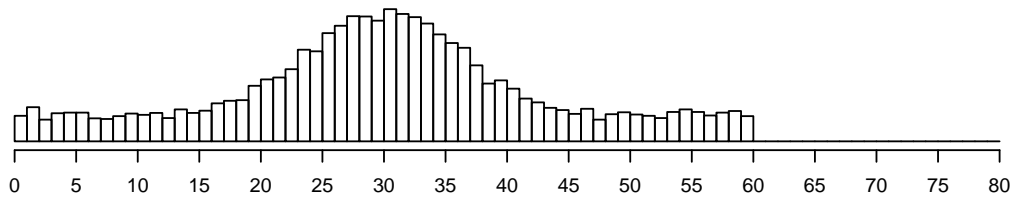Random Variable G

Random Variable H

(a) Estimate the mean of F.

(b) Estimate the mean of G.

(c) Estimate the mean of H.

(d) Estimate the standard deviation of F.

(e) Estimate the standard deviation of G.

(f) Estimate the standard deviation of H.

6. **Problem**

   Three random variables (F, G, and H) were measured 1000 times each. The resulting histograms show the three distributions.

   Random Variable F
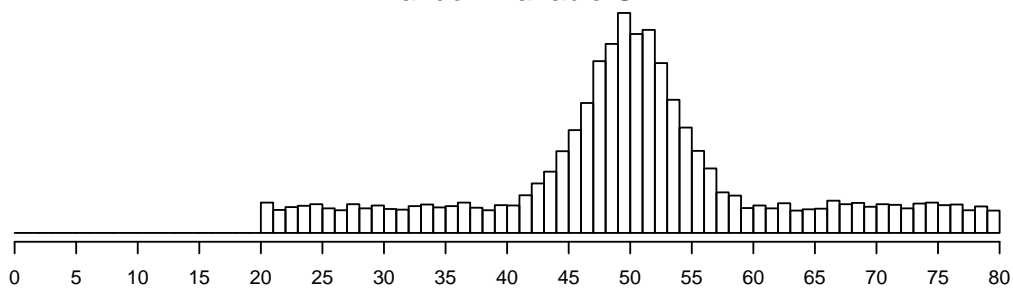
   Random Variable G

   Random Variable H
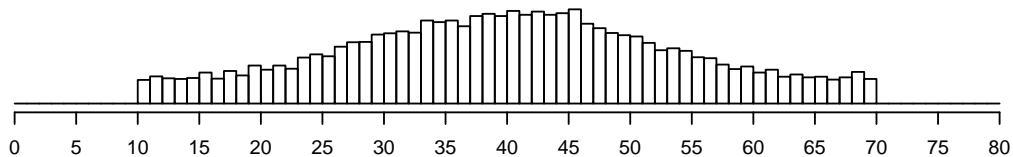
   (a) Which distribution has the highest mean? (F, G, or H)

   (b) Which distribution has the lowest mean? (F, G, or H)

   (c) Which distribution has the largest standard deviation? (F, G, or H)

   (d) Which distribution has the smallest standard deviation? (F, G, or H)

7. **Problem**
   Please make a frequency table and a dot plot from the following (unsorted) data.

15	16	18	20	19
15	15	15	14	19
18	14	18	17	20
19	16	14	13	18

8. **Problem**

A continuous random variable X was measured 27 times. The sorted measurements are shown below.

40.78	46.55	48.21	54.29	54.95	55.8	56.19	61.61	92.92	94.55
95.24	95.47	97.8	100.1	102.7	103.3	107.6	108.8	115.6	119.5
119.9	120.4	122.3	124.2	127.5	129.5	131.8			

The total of the measurements is 2527.56.

(a) Determine the percentile rank of the measurement 46.55. In other words, determine what percent of data are less than or equal to 46.55.

(b) Determine the measurement corresponding to a percentile rank of 0.778. In other words, determine x such that 77.8% of the data are less than or equal to x.

(c) Determine the mean of the measurements.

(d) Determine the median of the measurements.

9. **Problem**

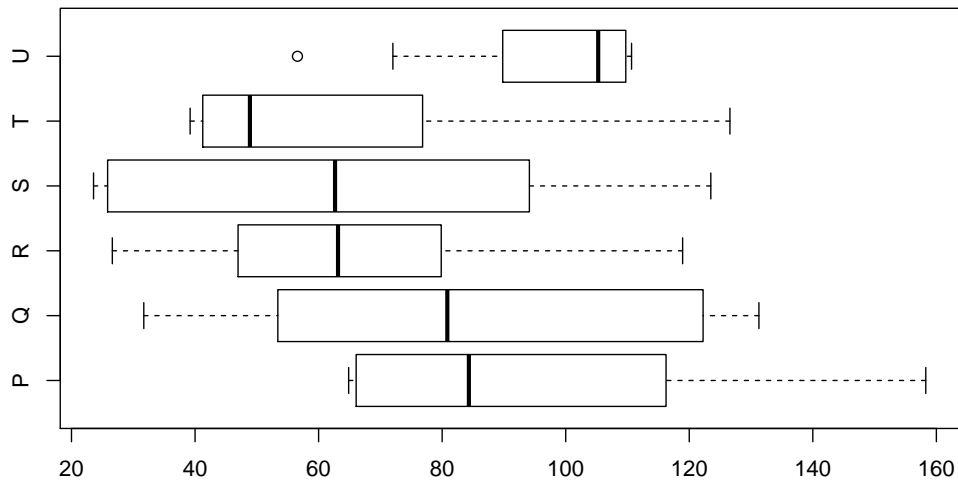From a very large population, a small sample of measurements was taken.

| 107 | 107 | 109 | 107 |

Please calculate the (Bessel corrected) sample standard deviation using the following formula:

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

10. **Problem**
    Six random variables were each measured 25 times. The resulting boxplots are shown.

    (a) Which variable produced the largest measurment?

    (b) Which variable produced the smallest measurment?

    (c) Which distribution has the largest median?

    (d) Which distribution has the smallest median?

    (e) Which distribution has the largest 25th percentile?

    (f) Which distribution has the smallest 25th percentile?

    (g) Which distribution has the largest 75th percentile?

    (h) Which distribution has the smallest 75th percentile?

    (i) Which distribution has the largest IQR?

    (j) Which distribution has the smallest IQR?

1. (a) H
   (b) G
   (c) G
   (d) F

2. We fill out the table column by column.

x	$x - \bar{x}$	$\lvert x - \bar{x} \rvert$
54	0	0
57	3	3
34	-20	20
72	18	18
47	-7	7
60	6	6
========	========	========
$\sum x = 324$		$\sum \lvert x - \bar{x} \rvert = 54$
$\bar{x} = 54$		

We are ready for the formula.

$$s = \frac{\sum \lvert x - \bar{x} \rvert}{n}$$
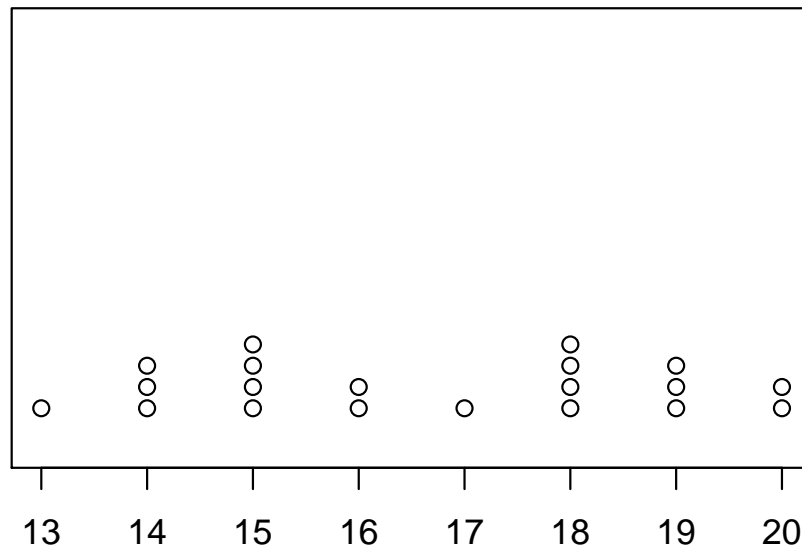
$$= \frac{54}{6}$$

$$= \boxed{9}$$

3. (a) uniform
   (b) 11
   (c) 16.33%
   (d) 96.67%
   (e) 13%
   (f) 27.33%
   (g) 79.59%
   (h) 13

4. (a) bimodal
   (b) 5.5
   (c) 46%
   (d) 36%
   (e) 10%

   (f) 3.5%

   (g) 78.26%

   (h) 14.5

5. (a) 40

   (b) 60

   (c) 30

   (d) 5

   (e) 10

   (f) 1.6666667

6. (a) G

   (b) F

   (c) H

   (d) G

7. Make a frequency table.

value	frequency
13	1
14	3
15	4
16	2
17	1
18	4
19	3
20	2

Make the dot plot.

8. Let x represent a measurement of interest. Let i represent that measurement's index. Let ℓ represent that measurement's percentile. Let n represent the sample size (number of measurements). In general,

$$\ell = \frac{i}{n}$$

(a) We are given $x = 46.55$. This means $i = 2$. We know $n = 27$. Determine the percentile ℓ.

$$\ell = \frac{2}{27}$$

$$\ell = 0.0741$$

So, the percentile rank is $\boxed{0.0741}$, or 7.41th percentile.

(b) We are given $\ell = 0.778$. We can use algebra to solve for i.

$$\ell = \frac{i}{n}$$

Multiply both sides by n.

$$n \cdot (\ell) = n \cdot \left(\frac{i}{n}\right)$$

Simplify both sides.

$$n\ell = i$$

To make me happy, switch the sides.

$$i = n\ell$$

Now, we can evaluate i.

$$i = (27)(0.778)$$

$$i = 21$$

Determine the x associated with $i = 21$.

$$x = \boxed{119.9}$$

(c) The mean: $\bar{x} = \frac{2527.56}{27} = \boxed{93.613}$

(d) If n is odd, then median is $x_{i=\frac{n+1}{2}}$, the value of x when $i = \frac{n+1}{2}$. Otherwise, if n is even, the median is mean of $x_{i=\frac{n}{2}}$ and $x_{i=\frac{n}{2}+1}$. In this case, $n = 27$ and so n is odd.

$$\text{median} = x_{(27+1)/2,} = x_{14}$$

So, median = $\boxed{100.1}$.

9. We fill out the table column by column.

x	$x - \bar{x}$	$(x - \bar{x})^2$
107	-0.5	0.25
107	-0.5	0.25
109	1.5	2.25
107	-0.5	0.25
========	========	========
$\sum x = 430$		$\sum(x - \bar{x})^2 = 3$
$\bar{x} = 107.5$		

We are ready for the formula.

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

$$= \sqrt{\frac{3}{4 - 1}}$$

$$= \sqrt{1}$$

$$= \boxed{1}$$

10. (a) P

(b) S

(c) U

(d) T

(e) U

(f) S

(g) Q

(h) T

(i) Q

(j) U