Data Analysis: Measures of Spread; Empirical Rule Name: Key

1. The times in minutes for swimmers during a race are shown in the following table. What is the approximate mean race time? What is the approximate standard deviation? Round to the nearest hundredth. Pop.

Swimmer Number	Time (minutes)	Swimmer Number	Time (minutes)
1	2.10	7	2.08
2	2.30	8	2.19
3	2.57	9	2.35
4	2.10	10	2.20
5	2.22	11	2.12
6	2.47	12	2.26

mean: 2.25 standard deviation: .15

2. The following are the [10] sample means of weighted boxes in a delivery truck. What is the approximate standard deviation of the sample means of weighted boxes? Round your answer to the nearest hundredth.
11.4, 9.8, 10.0, 8.2, 10.5, 9.8, 9.9, 11.1, 11.4, 10.2 sample

.91

3. Determine the mean and standard deviation of the following data:
Round [correctly] to the nearest hundreth.

Height (in inches)	Frequency
56	1
58	4
60	6
62	11
64	8
66	5
68	2

Sample

mean: 62.38 standard deviation: 2.90

4. Jeremy wants to perform an experiment to analyze the weights of players on his soccer team. He assigns a number to each of the players and takes a sample of ten players on his list using a random number generator.. Biased or unbiased?

Unbiased SRS.

5. Devin wants to perform an experiment to analyze the heights of apple trees in an orchard. He assigns each apple tree in the orchard a unique number and picks the first ten trees. Biased or unbiased?

biased Convenience

6. James, Maggie, and Simone are collecting data to perform an experiment to analyze the heights of students in their school. There are a total of 465 students in the school. James chooses a random sample of 90 students and records their heights. Maggie chooses a random sample of 75 students and records their heights. Simone randomly chooses a sample of 60 students and records their heights. Whose sample would best represent the heights of all students in the school? Why?

name: None of Them

why? They are not independent

If independent Then James Larger sample better representation

7. Given a mean of 65 and a standard deviation or 5, build a normal distribution and answer the following questions.

a. What percent of the data lies below 65? 50%

b. What percent of the data lies above 65? 50%

c. What percent of the data lies wit in one standard deviation of the mean? 68%

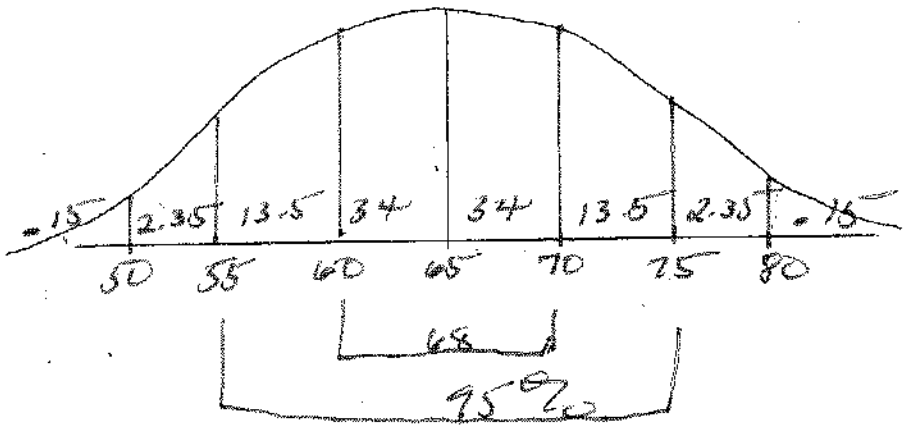
d. What percent of the data lies within 2 standard deviations of the mean? 95%

e. What percent of the data lies between 50 and 70? 83.85%

f. What percent of the data lies below 50? .15%

g. What percent of the data lies above 75? 2.5%

h. What percent of the data lies between 60 and 80? 83.85%



Worksheet on Normal Distribution # 7 Name: Key

For each question, construct a normal distribution curve and label the horizontal axis. Then answer each question.

1. The mean life of a tire is 30,000 km. The standard deviation is 2000 km.
   a) 68.3% of all tires will have a life between 28000 km and 32000 km.
   b) 95.5% of all tires will have a life between 26000 km and 34000 km.
   c) What percent of the tires will have a life that exceeds 26,000 km? 1 − .0062 = .9938
   d) If a company purchased 2000 tires, how many tires would you expect to last more than 28 000 km? 1 − .1587 = .8413 × 2000 = 1682.6 = 1683

2. The shelf life of a particular dairy product is normally distributed with a mean of 12 days and a standard deviation of 3 days.
   a) About what percent of the products last between 9 and 15 days? 68%
   b) About what percent of the products last between 12 and 15 days? 34%
   c) About what percent of the products last 6 days or less? 2.5%
   d) About what percent of the products last 15 or more days? 16%

3. A line up for tickets to a local concert had an average (mean) waiting time of 20 minutes with a standard deviation of 4 minutes.
   a) What percentage of the people in line waited for more than 28 minutes? 2.5%
   b) If 2000 ticket buyers were in line, how many of them would expect to wait for less than 16 minutes?

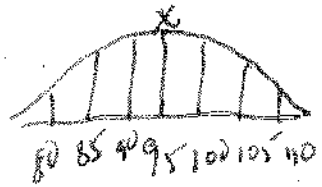
.16 × 2000 = 320 people

5. In an Oreo factory, the mean mass of a cookie is given as 40 g. For quality control, the standard deviation is 2 g.
   a) If 10,000 cookies were produced, how many cookies are within 2 g of the mean? .68 × 10000 = 6800 cookies
   b) Cookies are rejected if they weigh more than 44 g or less than 36 g. How many cookies would you expect to be rejected in a sample of 10,000 cookies?

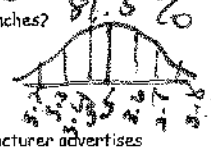
.025 + .025 = .05 × 10,000 = 500 cookies

6. The speeds of cars on the highway have a mean of 95 km/h with a standard deviation of 5 km/h.
   a) What percentage of cars averaged less than 85 km/h? 2.5%
   b) If a police car stopped cars that were going more than 105 km/h, how many cars would they stop if there were 8000 cars on the highway?

.025 × 8000 = 200 cars



7. The Floppy Disk Company makes 3.5 inch floppy disks for disk drives that are 3.7 inches wide. The size of a manufactures disk is normally distributed with a standard deviation of 0.1 inches. The company manufactures 1000 disks every hour.
   a) What % of the disks would you expect to be greater than 3.5 inches? 50%
   b) In one hour, how many disks would you expect to be between 3.4 inches and 3.7 inches? .815 × 1000 = 815
   c) About how many disks will be unable to fit in the disk drive (3.7 inch won't fit)? .025 × 1000 = 25



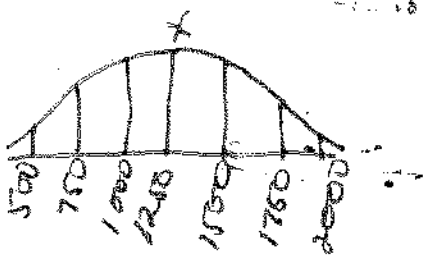
8. The mean life of a battery is 50 hours with a standard deviation of 6 hours. The manufacturer advertises that they will replace all batteries that last less than 38 hours. If 50,000 batteries were produced, how many would they expect to replace? .025 × 50,000 = 1250

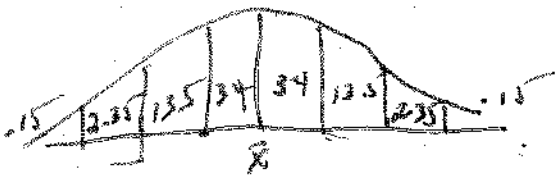
9. A bottle of fruit punch contains at least 473 ml. The machine that fills the bottles is set so that the mean volume is 477 ml. The volumes in the bottles are normally distributed.
   a) What percent of the bottles are underfilled if the standard deviation is 2 ml? 2.5%
   b) What percent of the bottles are underfilled if the standard deviation is 4 ml? 16%

$\frac{473-477}{2} = -2$ $\frac{473-477}{4} = -1$

10. A grading scale is set up for 1000 students' test scores. It is assumed that the scores are normally distributed with a mean score of 75 and a standard deviation of 15
   a) How many students will have scores between 45 and 75? P(45 < x < 75) = P(−2 < z < 0) = .5 − .0228 = .4772 × 1000 = 477
   b) P(x > 60) = P(z > −1) = 1 − .1587 = .8413 × 1000 = 841

11. The monthly income of 5,000 workers at the Microsoft plant are distributed normally. Suppose the mean monthly income is $1,250 and the standard deviation is $250.
   a) How many workers earn more than $1500 per month? .16 × 5000 = 800
   b) How many workers earn less than $750 per month? .025 × 5000 = 125
   c) What percentage of the workers earn between $750 and $1500 per month? 81.5%
   d) What percentage of the workers earn less than $1750 per month? 97.5%





Key

WORKSHEET: (Normal Distribution and Z scores)

The following are a set of practice problems.

Conversion of Variables:

1. Each year thousands of high-school students take either the SAT or ACT, standardized tests used in the college admission process. Combined SAT Math and Verbal scores go a s high as 1600, while the maximum ACT composite is 36. Different scales are used so the comparisons are different. SAT= 40 X ACT +150
   An admissions officer reported the following statistics about the ACT scores of 2355 students who applied to her college one year. Find the summaries of equivalent SAT spores.

ACT		SAT	
Lowest score=19		Lowest score=	910
Mean=27		Mean=	1230
Standard deviation=3		Standard deviation=	270
3rd Quartile= 30		3rd Quartile=	1350
Median= 28		Median=	1270
IQR= 6		IQR=	340

2. A high school senior uses the Internet to get information on February temperatures in the town where he'll be going to college. The website gives information in degree Celsius.
   Conversion formula: °F = (9/5)°C + 32
   Determine the Fahrenheit equivalents for the summary information below:

Max Temp= 11°C		Max Temp=	51.8°F
Range = 33°C		Range=	91.4°F
Mean= 1°C		Mean=	33.8°F
Standard Deviation = 7°C		SD =	44.6°F
Median = 2°C		Median =	35.6°F
IQR= 16°C		IQR =	60.8°F

$z = \frac{x - \mu}{\sigma}$

Z-Scores and standardization:

3. Nicole's score on the Stats midterm was 80 points. The class average was 75and the SD was 5 points. What was her z-score?

   $z = \frac{80 - 75}{5}$

   Ans: $z = 1$
4. Cars currently sold in the US have an average of 135 horsepower with an SD of 40 horsepower. What is the z-score for a car with 195 horsepower?

   $z = \frac{195 - 135}{40}$

   Ans: $z = 1.5$
5. The average score on a Stats midterm was 75 points with an SD of 5 points. If Gregor's z-score is -2, how many points did he score?

   $-2 = \frac{x - 75}{5}$ $-10 = x - 75$ $65 = x$

   Ans: $x = 65$
6. People with z-scores greater than 2.5 on an IQ test are considered as geniuses. If IQ tests have a score of a mean of 100 and SD of 16 points. What is the cut off score for a genius to prove himself as one?

   $2.5 = \frac{x - 100}{16}$

   Ans: $x = 140$
7. A town's January high temperatures average 36°F with an SD of 10°F, while in July the mean high is 74°F and SD is 8°F. In which month is it more unusual to have a day with a high temp of 55°F? Why?

   $z = \frac{55 - 36}{10} = 1.9$ $z = \frac{55 - 74}{8} = -2.375$

   Ans: July you are farther away from the avg Temp
8. Anna scored 83 on both her French and Spanish final. Megan scored 77 on French and 95 on her Spanish final. Overall student scores on the French exam had a mean of 81 and SD of 5, while Spanish exam mean was 74 and SD was 15.
   a) To qualify for honors a major must maintain at least an 85 average for all language courses taken. Which student qualifies?

      Ans: _____
   b) Which student's overall score is better?

      Ans: _____
   c) Would there be a difference if all the Spanish exam marks were increased by 5 marks? Ans a) and b) again?

      Ans: _____

Anna

$Fr = \frac{83 - 81}{5} = .4$

$Sp = \frac{83 - 74}{15} = .6$

Megan

$Fr = \frac{77 - 81}{5} = -.8$

$Sp = \frac{95 - 74}{15} = 1.4$

Normal Models:

9. What percent of a standard normal model is found in each region? Draw a picture.

a) $z > 1.5$	Ans:	$1 - .9332 = .0668$ 6.68%
b) $z < 2.25$	Ans:	$.9878$ 98.78%
c) $-1 < z < 1.15$	Ans:	$.7162$ 71.62%
d) $\lvert z \rvert > 0.5$	Ans:	$.6170$ 61.7%

$.9749$ $.1587$

$\pm .5 \times .5$

$.3085 \times 2$

$x_1 + x_2 = 155$

$\sigma = \sqrt{5^2 + 15^2} = 15.81$

$\frac{85 - 81}{5} = .8$

$\frac{85 - 74}{15} = .73$

Megan = 170 Anna = 166

French $\mu = 81$ $\sigma = 5$

Spanish $\mu = 74$ $\sigma = 15$

$z = \frac{77 - 81}{5} = -.8$ $z = \frac{83 - 81}{5} = .4$

$z = \frac{95 - 74}{15} = 1.4$ $z = \frac{83 - 74}{15} = .6$

10. In a standard Normal Model what values of Z cuts off the region described. Draw a picture.

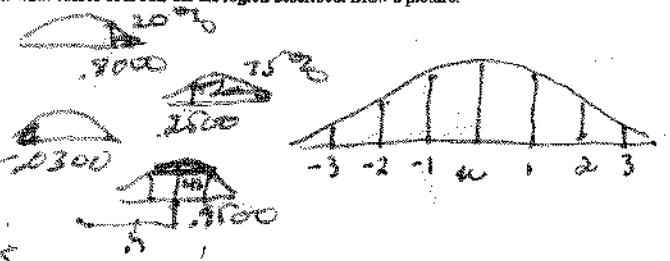
a) The highest 20%
Ans: .84

b) The highest 75%
Ans: -.67

c) The lowest 3%
Ans: -1.88

d) The middle 90%
Ans: ±1.645

e) The middle 50%
Ans: ±.67



11. Based on the IQ ~ N(100,16) scores, what percent of people's IQs would you expect to be:

a) Over 80? Ans: 89.44%
b) Under 90? Ans: 26.43%
c) Between 112 and 132? Ans: 20.38%
d) Cut off value of highest 5%? Ans: 126.32
e) Cut off value of lowest 30%? Ans: 91.168
f) Cut off value of middle 80%? Ans: 79.52 to 120.48
g) What is the IQR of the IQs? Ans: 21.44
h) What IQ represents the 98th percentile? Ans: 132.8

11. a. $z = \frac{80-100}{16} = -1.25$

b. $z = \frac{90-100}{16} = -.625$

c. $z = \frac{112-100}{16} = .75$ $\quad z = \frac{132-100}{16} = 2$

.9773 − .7734 = 20.39

D $1.645 = \frac{x-100}{16}$ = 126.32

E $-.52 = \frac{x-100}{16}$ = 91.68

F $1.28 = \frac{x-100}{16}$ $\quad -1.28 = \frac{x-100}{16}$

120.48 $\quad$ 79.52

g $.67 = \frac{x-100}{16}$ $\quad 110.72$ $\quad 89.28$

h $2.05 = \frac{x-100}{16}$

12. Find the missing parameter?

a) μ=88, 2% below 50; σ= 18.54
b) μ=0.64, 12% below 0.7; σ= .05 Not good
c) σ=0.5, 90% above 10; μ= 10.64
d) σ=15.6, 10% below 11.2; μ= 31.168

d. $-1.28 = \frac{11.2-\mu}{15.6}$ = 31.168

Miscellaneous:

13. While only 5% of babies have learned to walk by the age of 10 months, 75 % are walking by 13 months of age. If the age at which babies develop the ability to walk can be described by a Normal model find mean and standard deviation?

$.670 = \frac{13-\bar{x}}{.67}$ $\quad -1.645$ $\quad \sigma = \frac{10-\bar{x}}{-1.645}$ $\quad -.67 = \frac{x-100}{16}$

$z = .67$

± 1.645

$\frac{13-\bar{x}}{.67} = \frac{10-\bar{x}}{-1.645}$

$-1.645(13-\bar{x}) = .67(10-\bar{x})$

$-21.385 + 1.645\bar{x} = 6.7 - .67\bar{x}$

$2.315\bar{x} = 28.085$

$\bar{x} = 12.13$

$\sigma = \frac{13-12.13}{.67}$ $\quad \sigma = 1.3$

14. Wild life biologists believe that the weights of adult trout can be described by a normal model. They collected data from fishermen, finding that 22% of the trout caught were thrown back because they were below 2-lbs minimum and only 6 % weighed over 5 lbs. What are the parameters? (mean and SD)

$\sigma = \frac{2-\bar{x}}{.77}$ $\quad \sigma = \frac{5-\bar{x}}{1.555}$

$.77(5-\bar{x}) = 1.555(2-\bar{x})$

$-3.85 + .77\bar{x} = 3.11 - 1.555\bar{x}$

$2.325\bar{x} = 6.96$

$\bar{x} = 2.994$

$\sigma = \frac{2-2.994}{.77}$ $\quad \sigma = 1.29$

15. Based on a long term investigation, researchers have suggested that the acidity (pH) of rainfall in the Shenandoah mountains can be described by the normal model N(4.9,0.6). The lower the pH, the more acidic the rain. What is the pH level for the most acidic 20% of all storms and that of the least acidic 5%? $\bar{x} = 4.9$ $\quad \sigma_x = .6$

20% $\quad .84 = \frac{x-4.9}{.6}$ $\quad x > 5.404$

5% $\quad -1.645 = \frac{x-4.9}{.6}$ $\quad x < 3.913$

16. Avoiding an accident when driving can depend on reaction time. That time measured from the moment the driver first sees the danger until he or she gets his foot on the brake pedal is thought to follow a Normal model with mean 1.5 seconds and SD of 0.18 seconds.
What is the reaction time of the slowest one third (1/3) of all drivers. Give the 68-95-99.7 Rule.

$\bar{x} = 1.5$ $\quad \sigma = .18$

.333

$-.43 = \frac{x-1.5}{.18}$

$x < 1.4226$