STATISTICS 100 EXAM 1 Version 1 Key (Exam)

February 10th, 2016
Spring 2016

PRINT NAME _____ _____
(Last name) (First name)

NETID: _____

CIRCLE SECTION: **L1 (Laska MWF 12pm)** **L2 (Laska Tues/Thurs 11am)**

Write answers in appropriate blanks. When no blanks are provided **CIRCLE** your answers.
SHOW WORK when requested, otherwise no credit.
If you need scrap paper, raise your hand and ask for some!

Make sure you have all 7 pages including the normal table (15 problems).

DO NOT WRITE BELOW THIS LINE

The numbers written in each blank below indicate how many points you missed on each page. The numbers printed to the right of each blank indicate how many points each page is worth.

Page 1 _____13

Page 2 _____10

Page 3 _____30

Page 4 _____22

Page 5 _____ 13

Page 6 _____12

Total Score _____

***Karle Laska's Sections: There is NO class Thursday or Friday! Have a great Valentine's Day weekend!**

Scores will be posted in Compass early Friday morning ☺

Floor

Statistics 100 Exam 1 Exam Page 1 February 10th, 2016

Question 1 *(8 points total)*

A recent study was done to test the effectiveness of acupuncture in the treatment of tension-type headaches. The subjects were 270 adult volunteers who reported having had tension headaches for at least eight days a month in the previous three months.

The subjects were ***randomly*** divided into 3 groups. One received 8 weeks of a traditional form of acupuncture, one received 8 weeks of a fake acupuncture (superficial needling at non-acupuncture points) and a third group were told they were on a waiting list and received no treatment for the 8 weeks.

The subjects didn't know if they were receiving the traditional or superficial acupuncture. All subjects kept headache diaries for 8 weeks. The evaluators recorded the difference in the number of headaches before and after treatment for each subject not knowing which subjects were in which group.

The average number of headaches after treatment decreased by 7 in both the real and fake acupuncture groups, but only decreased by 1.5 in the waitlist group.

a) *(2 points)* Which of the following best describes this study? ***Choose one:***

i) It's an observational study

ii) It's a randomized controlled experiment with a placebo and "blind" evaluators.

iii) It's a non-randomized controlled double-blind experiment.

b) *(2 points)* Which of the following statements is best? ***Choose one:***

i) This study is very strong evidence that traditional acupuncture works better than a placebo in the treatment of tension-type headaches.

ii) This study only shows an *association* between acupuncture and reduced headaches. It does not prove or disprove that acupuncture *caused* fewer headaches since there are likely to be other differences between those who received the real and those who received the fake acupuncture that could confound the results.

iii) This study is strong evidence that acupuncture (both fake and real) has a strong placebo effect since the real acupuncture worked no better than the fake acupuncture but both worked better than nothing.

c) *(2 points)* Which of the following are likely to confound the results of this study? ***Choose one:***

i) Pain Tolerance- People who choose acupuncture may tolerate pain better and thus report fewer headaches.

ii) Alternative Medicine- People who choose acupuncture are more likely to be taking alternative therapies such as herbal cures and massage which could help alleviate headaches.

iii) Tension -- People who seek acupuncture are more likely to have tension headaches than those who don't.

iv) All of the above are likely confounders.

v) None of the above are likely confounders.

d) *(2 points)* Say I wanted to do a similar randomized controlled experiment at U of I, but I could only get 24 volunteers. I want to make sure my treatment and control group have the same amount of males and females in each group. I'm worried that when I randomly divide my subjects into a treatment and control group, the small sample size could be a problem and my groups may be unequal in terms of gender. What's the best method I can use to avoid this?

i) I could hand pick the groups to have the same amount of males and females.

ii) At the end of the experiment, I could "stratify" based on gender. Only compare the males in the treatment group to the males in the control group and do the same with the females.

iii) There is nothing that the researcher could do here to help.

iv) The researcher could "block" the subjects based on gender first, then randomly assign half of the males to treatment and half to control. They would then do the same thing with the females.

Question 2 *(5 points total)*

Two language instruction schools both gave the same standardized Spanish and Chinese pass/fail exams to their graduating classes. The table below gives the results for the past 5 years.

	School A			School B		
	# Pass	# Fail	% Pass	# Pass	#Fail	% Pass
Spanish	900	100	90%	850	150	85%
Chinese	600	400	60%	30	70	30%
Total	1500	500	75%	880	220	80%

a) *(1 point)* Which school has a higher passing rate in Spanish? i) School A ii) School B iii) unknown

b) *(1 point)* Which school has a higher passing rate in Chinese? i) School A ii) School B iii) unknown

c) *(1 point)* Which group has the higher overall passing rate? i) School A ii) School B iii) unknown

d) *(2 points)* Which conclusion is best supported based only on the data in the table?

i) School A has better instruction in both Spanish and Chinese.

ii) School B has better instruction in both Spanish and Chinese.

iii) The relative quality of the instruction at the two schools depends on which language is being taught.

Questions 3 *(4 points total)*

Suppose I wanted to determine whether or not offering optional homework problems to supplement the required homework would help students learn. I randomly split the class into 2 groups—the treatment group was offered optional problems to supplement the required homework while the control group was just given the required problems.

At the end of the semester I compared the exam scores of the 2 groups and saw no significant difference. **Both groups had an average of exam score of 80% and a SD of 20%.**

But only half the students in the treatment group actually did the supplemental problems. Those who chose to do the optional homework did much better than those who chose not to do it. The table below gives the data:

	Treatment group Exam Average	Control Group Exam Average
Chose to do extra homework	90%	Not applicable
Chose not to do extra homework	70%	Not applicable
Total	80%	80%

a) *(2 points)* Which 2 percentages are most appropriate to compare to determine whether or not the extra problems helped?
Choose one: **i)** 80% vs. 80% **ii)** 90% vs. 80% **iii)** 80% vs. 70% **iv)** 90% vs. 70%

b) *(2 points)* Which conclusion is most appropriate to draw from the data? *Choose one:*

i) Giving students the option of doing supplemental problems helps them learn.

ii) Giving students the option of doing supplemental problems doesn't seem to make any difference in learning. Students who chose to do them are probably more serious students who would have done well without the supplemental problems.

iii) Giving students the option of doing supplemental problems impairs learning.

iv) Those who choose to do the supplemental problems are clearly helped by them and those who choose not to do the supplemental problems are clearly hurt by them.

Questions 4-8 *(6 points total)*

A recent study shows that children who play violent video games as kids are more likely to get into trouble with the law as adults. The researchers found that children who played violent video games were about 5 times more likely to get into trouble with the law as adults, which was statistically significant.

4. (1 point) Based only on the information above, this study is an example of… ***Choose one.***

a) An observational study
b) A randomized controlled experiment without a placebo
c) A randomized controlled double-blind experiment
d) A non-randomized controlled experiment

5. (2 points) Based only on the information above, which statement is best? ***Choose one.***

a) This study only shows that playing violent video games as a kid is **associated with** getting into trouble with the law; it doesn't show whether or not there's a causal relationship.
b) This study is strong evidence that violent video games **cause** people to get into trouble with the law as adults.
c) This study shows that there is no relationship between playing violent video games and committing crimes as an adult.
d) This study shows that playing violent video games as a kid has harmful effects and all parents can prevent their children from getting into trouble with the law by banning violent video games during their childhood.

Below are either confounders that mix up the study, causal links that explain the conclusion, or neither. ***Circle which is which.***

6. (1 point) Brain Development- Video games, even violent ones, force kids to use critical thinking skills to solve problems.
a) Confounder ***b)*** Causal Link ***c)*** Neither

7. (1 point) Uninvolved Parents- Kids who play violent video games probably don't have parents who are very active in their life. If the parents aren't watching their kids, they're more likely to play violent video games. They're also more likely to start hanging out with the wrong crowd when their older, and potentially partake in risky behavior that gets them in trouble with the law.
a) Confounder ***b)*** Causal Link ***c)*** Neither

8. (1 point) Exposure to fighting, death, and guns- Violent video games often include fighting, killing, and guns. When young kids see this while playing video games, it makes them think that this is a normal part of life, hence more likely to partake in these types of activities when they're older and get in trouble with the law.
a) Confounder ***b)*** Causal Link ***c)*** Neither

Question 9 *(20 points total)*

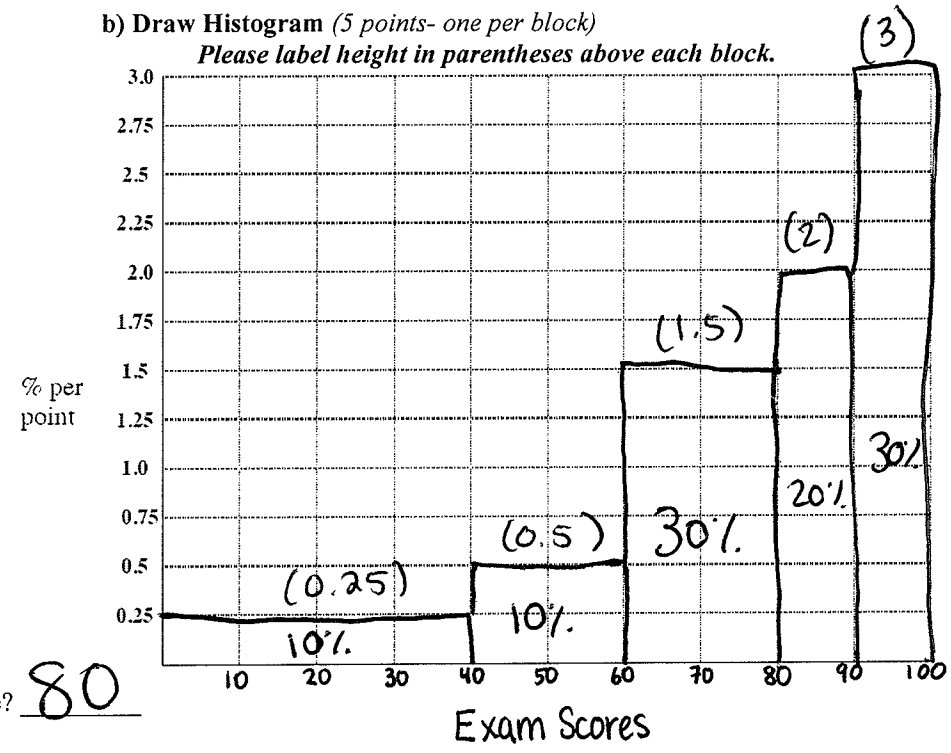
A distribution table for the scores on a Stat 100 exam is shown below on the left. Fill in the blanks in the height column then draw the histogram in the graph below:

a) Fill in table *(5 points)*

Points	%	Height (%/ point)
0-40	10	0.25
40-60	10	0.5
60-80	30	1.5
80-90	20	2
90-100	30	3

b) Draw Histogram *(5 points- one per block)*

Please label height in parentheses above each block.



c) *(2 points)* What is the median exam score? 80

d) *(2 points)* The median is > average.

Fill in blank above with > (greater than), < (less than), or =.

e) *(1 point)* What percentage of the people scored below 90? 70 % ***Fill in the blank with a number.***

f) *(2 points)* Would it be appropriate to use normal approximation with this data to figure percentages within different intervals?

Choose one:

i) Yes, if we knew the average and the SD we'd lose little accuracy in using the normal curve.

ii) Yes, normal approximations are always accurate in approximating percentages.

iii) Yes, because even though the histogram is far from normal, it will become normal after the data is converted to z-scores.

iv) No, this histogram is not close to normal, so the normal approximations would not be close to the real percentages.

g) *(3 points)* One student's exam was re-graded. His original score was changed from a 7 to a 77. A new average, median and SD were computed. How did the new average, median and SD compare to the old ones?

Fill in the blanks below with > (greater than), < (less than) or =.

i) The new average was > than the old one.

ii) The new median was = than the old one.

iii) The new SD was = than the old one.

Question 10 pertains to the following list of 5 numbers: *(10 points total)* 6, -3, 2, 3, 2

a) *(2 points)* The average is 2 and the median is 2. -3, 2, 2, 3, 6

b) *(5 points)* The deviations from the average are: 4, -5, 0, 1, 0

c) *(1 point)* The sum of the deviations from the average must always = 0

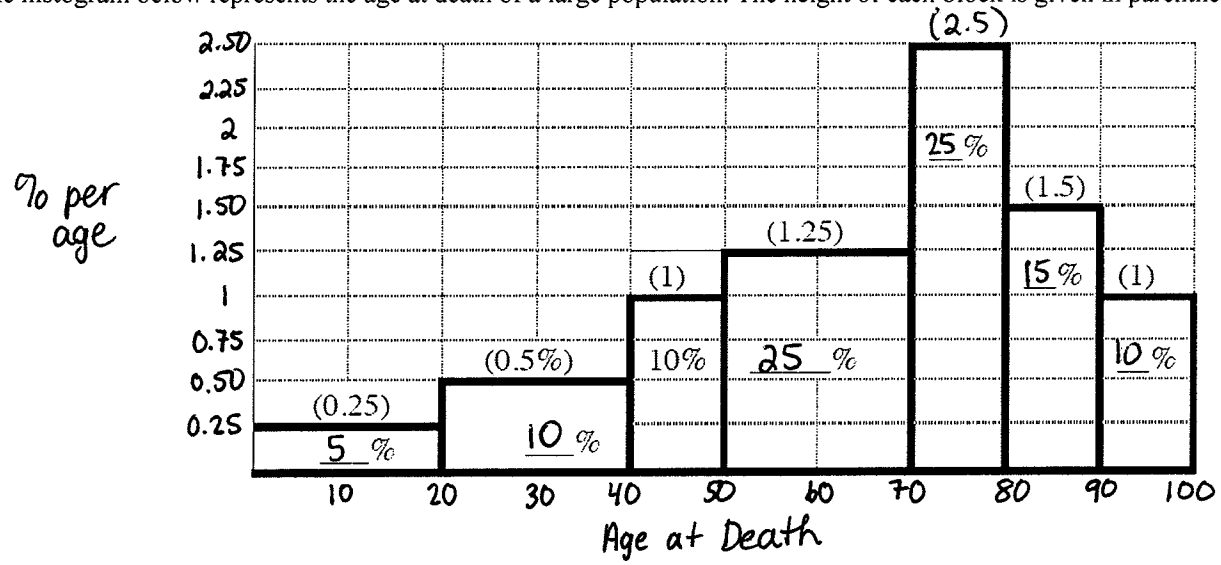
d) *(2 points)* Compute the SD. ***Show work using the deviations you got in part (b) above.***

Round answer to two decimal places. Circle answer.

$$\sqrt{\frac{16+25+0+1+0}{5}} = 2.90 \text{ or } 2.89$$

Question 11 pertains to the histogram below: *(16 points total)*

The histogram below represents the age at death of a large population. The height of each block is given in parentheses.



a) *(7 points)* **What percent died in each of the 7 intervals?** Fill in the blanks below with the appropriate areas.

0-20 5 **20-40** 10 **40-50** 10 **50-70** 25 **70-80** 25 **80-90** 15 **90-100** 10

b) *(1 point)* What is the median age of death? 70 years ***(Fill in the blank with a number)***

c) *(1 point)* The average is < the median. ***Fill in the blank with > (greater than), < (less than), or =.***

d) *(1 point)* What % of the population died at 75 years? (Assume an equal distribution throughout the interval.) 2.5 %

e) *(3 points)* If 5 years were added to everyone's life span how would it affect the average, median and SD of the histogram above?

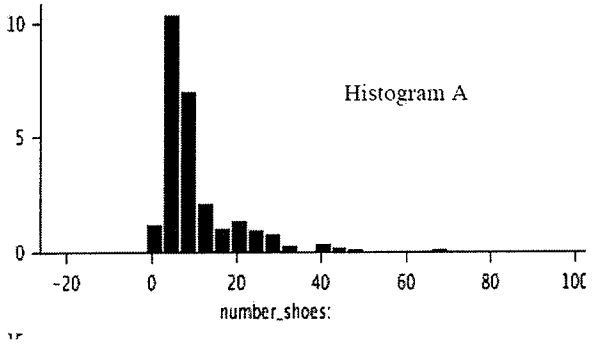
i) The average would … ***Choose one:*** **(a)** increase **b)** decrease **c)** stay the same **d)** Not enough info given
ii) The median would ***Choose one:*** **(a)** increase **b)** decrease **c)** stay the same **d)** Not enough info given
iii) The SD would … ***Choose one:*** **a)** increase **b)** decrease **(c)** stay the same **d)** Not enough info given

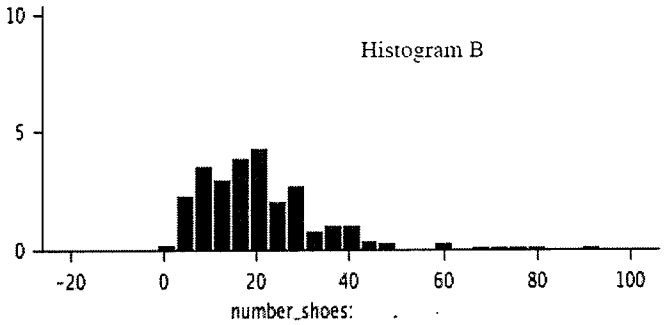
f) *(3 points)* If 5 years were *only* added to those in the 0-20 interval, how would it affect the average, median and SD?

i) The average would.. ***Choose one:*** **(a)** increase **b)** decrease **c)** stay the same **d)** Not enough info given
ii) The median would.. ***Choose one:*** **a)** increase **b)** decrease **(c)** stay the same **d)** Not enough info given
iii) The SD would … ***Choose one:*** **a)** increase **(b)** decrease **c)** stay the same **d)** Not enough info given

Question 12 pertains to the 2 histograms below which depict the male and female answers from this class to the survey question: "How many pairs of shoes do you own?" *(6 points total)*

a) *(2 points)* Which histogram represents the *female* answers? ***Circle one:*** **i)** Histogram A **(ii)** Histogram B





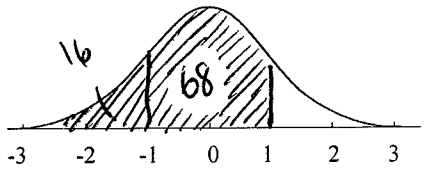
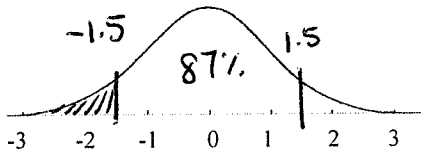
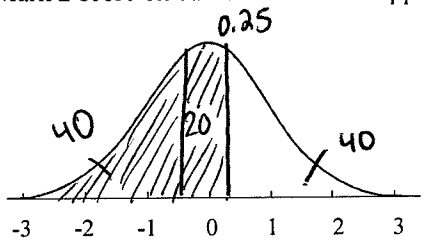
b) *(4 points)* The following 4 numbers (in no particular order) are the averages & medians of the 2 histograms: **7, 19, 10.152, & 20.168.** ***Fill in each of the 4 blanks below with the 4 given numbers.***

i) The median of Histogram A is 7 and the median of histogram B is 19

ii) The average of Histogram A is 10.152 and the average of Histogram B is 20.168

For Questions 13 and 15 use the normal table. You may "round" z scores and areas to fit the closest line on the normal table and you may round middle areas on the table to the nearest whole number. Hint for questions 13 and 15: Make sure your answers make sense!

Question 13 *(10 points total)*
Suppose math SAT scores are normally distributed with an ***average=520 and a SD=80***. In the table below you're given either a person's SAT score, z-score, or percentile and you have to fill in the missing blanks.

SAT Score	Z score	Percentile
Murphy has SAT= 600 $Z = \frac{600-520}{80} = 1$	**Z=** 1 *(1 point)* **Show work:**	***Murphy is in the*** 84 percentile *(Fill in the blank-1 point)* Mark z-score on curve & shade the appropriate percentile. *(1 point)* ***Round middle area to nearest whole number.***  $\frac{100-68}{2} = 16$ $68+16 = 84$
Kyla SAT score = 400 *(Fill in the blank- 1 point)* **Show work:** value = 520 + (-1.5)(80) = 400 **Do NOT round answer.**	**Z= -1.5**	***Kyla is in the*** 6.5 ***percentile.*** *(Fill in the blank-1 point)* Mark z-score on curve & shade the appropriate percentile. *(1 point)* ***Round middle area to nearest whole number.***  $\frac{100-87}{2} = 6.5$
Gibson SAT score= 540 *(Fill in the blank- 1 point)* **Show work:** value = 520 + (0.25)(80) **Do NOT round answer.**	**Z=** 0.25 *(1 point)*	***Gibson is in the*** **60th percentile** What middle area should you look up on the normal table to find the correct Z score? 20 % ***(Fill in blank- 1 point)*** *If the middle area is between two lines on the table, use the closest line.* Mark z-score on curve & shade the appropriate percentile. *(1 point)* 

Question 14: Describe the following variables. *(3 points total)*

a) Eye Color	**i)** Quantitative Continuous	**ii)** Quantitative Discrete	**iii)** Qualitative
b) Income	**i)** Quantitative Continuous	**ii)** Quantitative Discrete	**iii)** Qualitative
c) Number of Friends You Have	**i)** Quantitative Continuous	**ii)** Quantitative Discrete	**iii)** Qualitative

Statistics 100 Exam 1 February 10th, 2016

Question 15 *(12 points total)*

According to Survey data, **female** Stat 100 students have an ***average weight of 133 pounds and a SD of 23 pounds*** and the histogram of their weights is close to the normal curve.

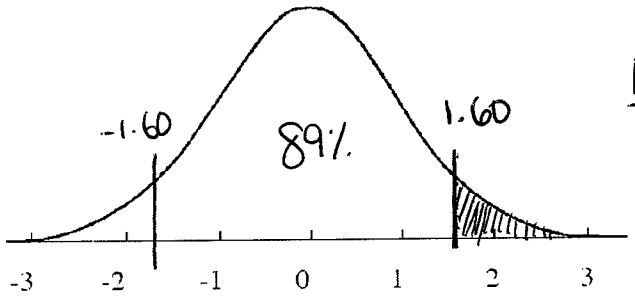
a) *(2 points)* Approximately **95%** of female students are between 87 lbs. and 179 lbs.
(Fill in the blanks with weights, NOT z scores) $z=-2$ $z=2$

b) What percent of the females in the class weigh more than 170 pounds?

i) *(1 point)* First convert **170 pounds** to a **z**-score. ***Show work and write your answer in the blank provided.***

z-score = 1.60 $z = \frac{170-133}{23} = 1.6$

ii) *(2 points)* Now use the normal table to determine what % of the females in the class weigh ***more than 170 lbs.*** ***Mark the z score accurately on the curve below and shade the appropriate region.***



$\frac{100-89}{2} = 5.5$

5.5 % weigh more than 170

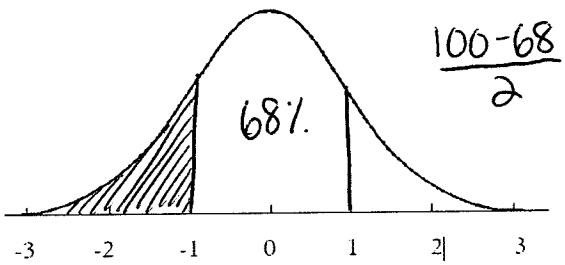
Fill in the blank above and round your answer to the nearest tenth. (one decimal place.)

c) What **percent** of the females weigh less than 110 pounds?

i) *(1 point)* First convert 110 pounds to a z-score. ***Show work and write your answer in the blank provided.***

z-score = -1 $z = \frac{110-133}{23} = -1$

ii) *(2 points)* Now use the normal table to determine what % of the females in the class weigh less than 110 pounds. ***Mark the z-score accurately on the curve below and shade the appropriate region.***

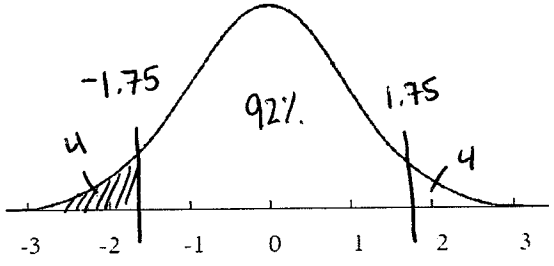


$\frac{100-68}{2} = 16\%$

16 % weigh less than 110

Fill in the blank above and round your answer to the nearest whole number.

iii) What weight corresponds to the **4th** percentile? Mark the **4th** percentile on the curve below and use the normal table to find the correct Z score. Fill in the blanks in the box.



What middle area should you look up on the normal table to find the correct Z score? 92 % (Fill in blank)

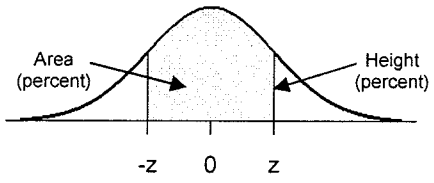
(If the middle area is between 2 lines on the table, use the line that is closest to find Z.)

Z= -1.75 Weight = 92.75 lbs (Don't round)

Show work (how you got the girl's weight from the Z score.)

value = 133 + (-1.75)(23)

STANDARD NORMAL TABLE



Standard Units

z	*Height*	*Area*		*z*	*Height*	*Area*		*z*	*Height*	*Area*
0.00	39.89	0.00		1.50	12.95	86.64		3.00	0.443	99.730
0.05	39.84	3.99		1.55	12.00	87.89		3.05	0.381	99.771
0.10	39.70	7.97		1.60	11.09	89.04		3.10	0.327	99.806
0.15	39.45	11.92		1.65	10.23	90.11		3.15	0.279	99.837
0.20	39.10	15.85		1.70	9.40	91.09		3.20	0.238	99.863
0.25	38.67	19.74		1.75	8.63	91.99		3.25	0.203	99.885
0.30	38.14	23.58		1.80	7.90	92.81		3.30	0.172	99.903
0.35	37.52	27.37		1.85	7.21	93.57		3.35	0.146	99.919
0.40	36.83	31.08		1.90	6.56	94.26		3.40	0.123	99.933
0.45	36.05	34.73		1.95	5.96	94.88		3.45	0.104	99.944
0.50	35.21	38.29		2.00	5.40	95.45		3.50	0.087	99.953
0.55	34.29	41.77		2.05	4.88	95.96		3.55	0.073	99.961
0.60	33.32	45.15		2.10	4.40	96.43		3.60	0.061	99.968
0.65	32.30	48.43		2.15	3.96	96.84		3.65	0.051	99.974
0.70	31.23	51.61		2.20	3.55	97.22		3.70	0.042	99.978
0.75	30.11	54.67		2.25	3.17	97.56		3.75	0.035	99.982
0.80	28.97	57.63		2.30	2.83	97.86		3.80	0.029	99.986
0.85	27.80	60.47		2.35	2.52	98.12		3.85	0.024	99.988
0.90	26.61	63.19		2.40	2.24	98.36		3.90	0.020	99.990
0.95	25.41	65.79		2.45	1.98	98.57		3.95	0.016	99.992
1.00	24.20	68.27		2.50	1.75	98.76		4.00	0.013	99.9937
1.05	22.99	70.63		2.55	1.54	98.92		4.05	0.011	99.9949
1.10	21.79	72.87		2.60	1.36	99.07		4.10	0.009	99.9959
1.15	20.59	74.99		2.65	1.19	99.20		4.15	0.007	99.9967
1.20	19.42	76.99		2.70	1.04	99.31		4.20	0.006	99.9973
1.25	18.26	78.87		2.75	0.91	99.40		4.25	0.005	99.9979
1.30	17.14	80.64		2.80	0.79	99.49		4.30	0.004	99.9983
1.35	16.04	82.30		2.85	0.69	99.56		4.35	0.003	99.9986
1.40	14.97	83.85		2.90	0.60	99.63		4.40	0.002	99.9989
1.45	13.94	85.29		2.95	0.51	99.68		4.45	0.002	99.9991