Statistics 100 Exam 1                                      February 11th, 2015

**Questions 1-4 pertains to the following study:** *(8 points total)*
A study on the influence of alcohol on memory was reported in the Jan, 2003 issue of *Psychological Science*. In the study, 148 undergraduates were **randomly** split into two groups, half were told they were getting a vodka tonic ( an alcoholic beverage) and the other half were told they were getting only tonic water (a non-alcoholic beverage). **In reality, all were getting plain tonic water with limes.**

The research was carried out in a room equipped with bartenders, vodka bottles, tonic bottles, and glasses, to give the feel of a bar-like setting; flat tonic water was poured from sealed vodka bottles that appeared to be brand new and all glasses were rimmed with limes.

Once the students consumed their drinks, they watched a sequence of slides depicting a crime. The subjects then read a summary of the crime riddled with misleading information. Neither the subjects nor those who evaluated them knew who was in which group.

**Results: Those who were told they were drinking vodka had significantly poorer memory powers, were more suggestible, and were less reliable eye-witnesses than those who were told they were drinking only tonic water.**

1. *(2 points)* Which of the following statements best describes this study?
   *Circle one:*
   a)    This was a randomized controlled experiment without a placebo.
   b)    This was an observational study with controls.
   c)    This was a randomized controlled double-blind experiment.
   d)    This was a non-randomized controlled experiment with a placebo.

2. *(2 points)* Based only on the results of this study, which of the following statements is most appropriate?
   *Circle one:*
   a)    This study is strong evidence that drinking alcohol impairs memory.
   b)    This study is strong evidence that simply *thinking* one has been drinking alcohol impairs memory.
   c)    This study only shows that there is an *association* between thinking one has been drinking alcohol and impaired memory. It does not show a causal relationship. A likely confounder is the social effect of a bar setting.
   d)    This study has cause and effect reversed. It's more likely that poor memory led to drinking alcohol than vice versa. After all, those who *remembered* past drunken mishaps were more likely to abstain from drinking the vodka.

3. *(2 points)* Which of the following could confound (mix up) the results?
   *Circle one:*

   a)    Responsibility- people who abstain from drinking are likely to be more responsible in general and would therefore take the eyewitness test more seriously.
   b)    Sex-men are more likely to drink in a bar and less likely to pay attention to a test in a bar.
   c)    Both of the above are likely confounders.
   d)    None of the above are likely confounders.

4. *(2 points)* In this study, the treatment group believes that they've been given alcohol when in fact they have not. When people believe they've been given the treatment, this affects their response. This phenomenon is known as?
   *Circle one:*
   a)    The placebo effect.
   b)    Double blind
   c)    Evaluator bias
   d)    Confounder

Statistics 100 Exam 1                                                                    February 11th, 2015

**Questions 5-9**   *(6 points total)*
A study was done to see if healthy school lunches could prevent childhood obesity. The subjects were all 1st graders in the same school district. 200 kids were randomly assigned to treatment and 800 to control. In the treatment group, the children ate in a special cafeteria and were given only fresh, mostly organic food, much of which was grown in their own school gardens. The control group ate the regular (highly processed) food in the regular cafeteria

5.   *(1 point)* Based only on the information above, which of the following best describes this study?
   **a) Randomized controlled experiment**      *b)* Observational study      *c)* Non-Randomized controlled experiment

**The program lasted through 5th grade and the children were periodically weighed. Throughout the study there were no significant differences in the average and the SD of the children's weights in the 2 groups.**
Even though everyone in the treatment group was encouraged to eat the healthy food only half of them did. The other half usually brought lunches from home. Many kids in the control group also brought lunches from home. Here's the data:

|  | Treatment Group | | | Control Group | | |
|---|---|---|---|---|---|---|
|  | # Kids | # Obese | % Obese | # Kids | # Obese | % Obese |
| Regularly ate school lunches (adherers) | 100 | 5 | 5% | 500 | 115 | 23% |
| Regularly brought lunch from home (non-adherers) | 100 | 35 | 35% | 300 | 45 | 15% |
| **Total** | 200 | 40 | 20% | 800 | 160 | 20% |

6.   *(2 points)* To evaluate whether or not the program helped reduce obesity, which 2 percents in the table should be compared?
   *a)*   Adherers in the treatment group (5%) vs. Adherers in the control group (23%)
   *b)*   Adherers in the treatment group (5%) vs. Non-adherers in the treatment group (35%)
   **c)**   Everyone in treatment (20%) vs. everyone in control (20%)

Based on the above information *circle* whether the following statements are *true or false.*
7.   *(1 point)* There is no evidence that the program either caused or prevented obesity.  **a) TRUE**      b) FALSE

8.   *(1 point)* There is strong evidence that the program prevented obesity for those who ate the school lunch but led to obesity in those who brought their lunch from home.                 a) TRUE          **b) FALSE**

9.   *(1 point)* You cannot compare obesity *rates* between the treatment and control groups because there were only 200 kids in treatment while there were 800 in control.                 a) TRUE          **b) FALSE**

**Questions 10-13 pertain to the following situation:** *(5 points total)*
While in college, Karle Laska & her fiancé Steve decided to take some courses online. The courses are divided into 2 groups (A & B) that cover the same subjects, but at a different level of difficulty. Group A courses are very difficult & include courses such as Japanese, Physics, & Calculus with Several Variables. Group B courses are easier & include courses such as Japanese Movies, Physical Education, & Calculators with Several Fingers. Here's a chart of how they did:

|  | Karle | | Steve | |
|---|---|---|---|---|
|  | # Courses | GPA | # Courses | GPA |
| Group A | 1 | 1.0 | 9 | 3.0 |
| Group B | 9 | 3.6 | 1 | 3.9 |
| Total | 10 | 3.34 | 10 | 3.09 |

10. *(1 point)* Who had the *higher* GPA for Group A courses?   *a)* Karle       **b) Steve**       *c)* Cannot be determined

11. *(1 point)* Who had the *higher* GPA for Group B courses?   *a)* Karle       **b) Steve**       *c)* Cannot be determined

12. *(1 point)* Who had the *higher* overall GPA?              **a) Karle**       *b)* Steve       *c)* Cannot be determined

13. *(2 points)* Which conclusion is best? *Choose one.*
   *a)*   Here it's best to compare total GPAs and clearly Karle is the better student.  She has a better overall GPA (3.34) than Steve who has an overall GPA of 3.09.
   **b)**   If you look at Group A and B courses separately, it is obvious that Steve is the better student.  His overall GPA is lower because the results are confounded by the difficulty level of the courses.
   *c)*   Here we cannot determine any conclusion because we don't know if Karle and Steve actually did their homework for each class.
   *d)*   There must have been a mistake when calculating the total GPAs for Karle and Steve since Karle's total GPA is better than Steve's.

Statistics 100 Exam 1                                                      February 11ᵗʰ, 2015

**Questions 14** *(2 points total)*
A recent study was done at a hospital and found that people who wear sunscreen regularly were **more likely** to get skin cancer than those who do not. This goes against what doctors had originally thought. They determined that it is very likely that a confounder is present and mixing up our results. Which of the following could confound the results of this study?
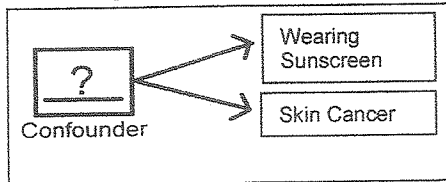


a)  Chemicals in sunscreen- When applying sunscreen to your skin, you're exposing yourself to certain harmful chemicals that may cause skin cancer.
b)  Genetics- Cancer is known to run in some families, making it more likely for you to get skin cancer whether you wear sunscreen or not.
c)  Health problems- Studies have shown that too much sun exposure is bad for your health.
d)  People with fair skin- People with fair or lighter skin easily get sunburned and therefore wear more sunscreen to prevent this. Also, people with fair skin have a higher risk of getting skin cancer.

**Questions 15-20** *(12 points total)*
A Mayo clinic study of 8,236 men & women found that people who had 1 or 2 alcoholic drinks a day were about half as likely to be obese as those who never drink. The researchers concluded that drinking a moderate amount alcohol may help prevent obesity.

15. *(2 points)* Which of the following statements best describes this study? ***Choose one.***
    a)  This was a randomized controlled experiment without a placebo.
    b)  This was an observational study.
    c)  This was a randomized controlled double-blind experiment.
    d)  This was a non-randomized controlled experiment with a placebo.

16. *(2 points)* Based on the results of this study alone, which of the following statements is best? ***Choose one.***
    a)  Never drinking alcohol causes more obesity than moderate drinking.
    b)  Never drinking alcohol is *associated* with and *may* cause more obesity than moderate drinking.
    c)  Never drinking alcohol is *associated* with but clearly does *not* cause more obesity than moderate drinking.
    d)  People who wish to lose weight should drink one or two alcoholic drinks each day.

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

17. *(2 points)* Genetics- It's likely that certain genes influence whether or not you are likely to gain weight and become obese.
    *a)* Confounder            *b)* Causal Link            *c)* Neither

18. *(2 points)* Metabolism- Consuming alcohol may raise your metabolism, which would lead to weight loss.
    *a)* Confounder            *b)* Causal Link            *c)* Neither

19. *(2 points)* Income- Wealthy people tend to have a lifestyle that includes a glass or two of wine with dinner and they also tend to have a lifestyle that makes it easier to stay fit and not become obese.
    *a)* Confounder            *b)* Causal Link            *c)* Neither

20. *(2 points)* Suppose my research team thinks that a possible confounder for the study done above about alcohol and obesity is age. In other words, people who drink moderately may be younger than people who don't. Younger people are also more likely to live active healthy lifestyles, meaning that it is unlikely that they will be obese. To confirm that age is a confounder, my research team should…
    **Choose one:**
    a)  Break the subjects up into different categories based on the response (obesity) and see what type of people are in each group and how much alcohol they drink.
    b)  Stop the study here and determine that it is inconclusive because age is a confounder that is messing up our results.
    c)  Break the subjects into subgroups along the confounding factor (age) where the age is the same in each group. If the difference in obesity goes away and the scales are balanced for each treatment and control group, you've found the confounder.

Statistics 100 Exam 1                                           February 11$^{th}$, 2015

**Question numbers 21-26 refer to the histogram below:** *(8 points total)*

The histogram below gives the ages of the audience attending a recent concert at Krannert, except the block for 40-50 years is missing. The heights of the blocks are labeled in parenthesis. No one in the group was younger than 10 or older than 50. (Assume an even distribution throughout each interval) Fill in the percentage for each block on the histogram below and draw the missing block accurately in the histogram below.
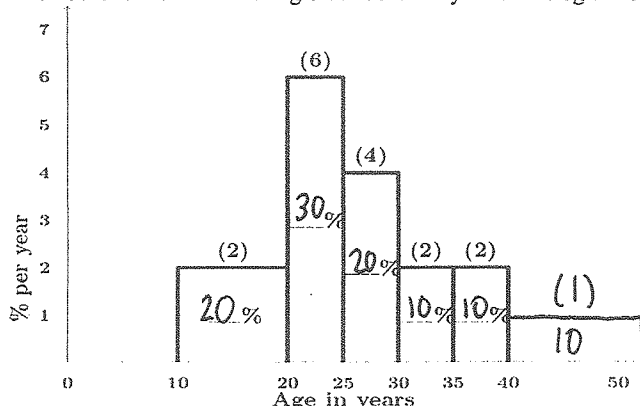


*21. (1 point)* The percent of audience between ages 40-50 years is

   a) 2%        b) 30%        **c) 10%**        d) 20%        e) none of the given choices

*22. (1 point)* The percent of audience between ages 10-20 years is _____ the percent of the audience between 30-40 years.

   a) less than        b) more than        **c) same as**        d) none of the given choices

*23. (1 point)* The median is closest to

   a) 5        b) 10        **c) 25**        d) 50        e) none of the given choices

*24. (2 points)* The area of the entire histogram (including the missing block) is closest to...

   a) 50%        b) 60%        c) 80%        d) 90%        **e) 100%**

*25. (1 point)* The 80th percentile is closest to...

   a) 40        b) 30        **c) 35**        d) 25        e) none of the given choices

*26. (2 points)* The percentage of audience that were exactly 15 years old is

   a) 1%        **b) 2%**        c) 6%        d) 20%        e) none of the given choices

**Questions 27-30 pertain to the new histogram drawn three years later:** *(4 points total)*
I decided to check back 3 years later and actually found the same people watching! This means that everyone's age is increased by 3 years.

*27. (1 point)* The total area (in percentage) of the new histogram would increase by 3.        a) True   **b) False**

*28. (1 point)* The average and median both increases by 3, but SD does not change.        **a) True**   b) False

*29. (1 point)* The average, median, and SD all increase by 3.        a) True   **b) False**

*30. (1 point)* The percentage of people in the audience who are now exactly 26 years old is 6%.        **a) True**   b) False

**Question 31-34 pertain to a list of numbers that has *a Standard Deviation = 0:*** *(4 points total)*

*31. (1 point)* The average of the list must be 0.        a) True   **b) False**        c) Not enough information to decide

*32. (1 point)* Every number on the list must be 0.        a) True   **b) False**        c) Not enough information to decide

*33. (1 point)* Every number on the list must be the same.   **a) True**   b) False        c) Not enough information to decide

*34. (1 point)* The median of the list must be 0.        a) True   **b) False**        c) Not enough information to decide

## February Page 5

**Questions 35-41 pertain to drawing a histogram for the distribution of exam scores given in the table below.**
*(10 points+ 5 points for histogram=15 points total)*
Complete the table for the height column. _Draw the histogram and write the height in parenthesis over each block._
Then answer the questions that follow.

| Exam Score | % of students | Height (% per point) |
|------------|---------------|----------------------|
| 0-60 | 15 | 0.25 |
| 60-70 | 10 | 1 |
| 70-80 | 25 | 2.5 |
| 80-90 | 30 | 3 |
| 90-100 | 20 | 2 |
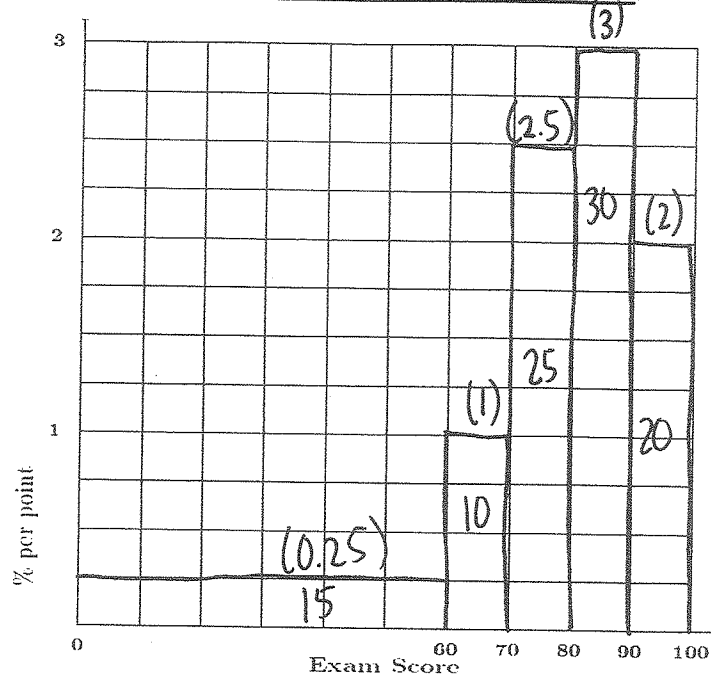


**35.** *(1 point)* The median score is:
a) 90   **(b) 80**   c) 70   d) 60   e) 50

**36.** *(1 point)* The histogram for the exam scores has a _____.
a) Long right hand tail
**b) Long left hand tail**
c) Is symmetric
d) Cannot be determined

**37.** *(1 point)* The average is _____ the median.
**a) less than**    b) greater than
c) equal to    d) cannot be determined

**38.** *(2 points)* If 10 points are added to everyone's score (the maximum score is now 110) then, the SD would...
a) increase    b) decrease    **c) stay the same**    d) cannot be determined

**39.** *(2 points)* If 10 points are added only to the score of those who failed (i.e. the 0-60 interval) then the SD would...
a) increase    **b) decrease**    c) stay the same    d) cannot be determined

**40.** *(1 point)* If Timothy is at the 25th percentile for scores in the class, then Timothy's score on the exam is closest to...
**a) 70**    b) 80    c) 50    d) 65    e) cannot be determined

**41.** *(2 points)* Would it be appropriate to use a normal approximation for this data?
**a) No, the histogram is far from normal, so we should not be using a normal approximation.**
b) Yes, because converting to z-scores will change the shape and make the histogram normal.
c) No, this histogram is already normal so there is no need to approximate.
d) Yes, because we can get the average and SD from the data and then use the normal approximation

**Question 42 (a) – (f) pertain to the following:** *(13 points total)*
Consider the following list of 4 numbers: 0, 2, -2, 8    $-2, 0, 2, 8$

a) *(2 points)* The average is $2$ .

b) *(2 points)* The median is $1$ .

c) *(4 points)* The deviations are $-4, -2, 0, 6$ .

d) *(1 point)* The sum of the deviations should *always* = $0$ (Check that the 5 deviations above sum to what they should)

e) *(2 points)* The standard deviation is $3.74$ . (Round answer to 2 decimal places.)
Show work. You may start by using the deviations you got in (c) above.
$$\frac{16+4+0+36}{4} = \frac{56}{4} = 14 \qquad \sqrt{14} = 3.74$$

f) *(2 points)* Suppose you multiplied all the numbers on the list above by -10, the SD would ....
**i) be multiplied by 10.**    ii) be multiplied by -10.    iii) decrease by 10.    iv) increase by 10.    v) stay the same.

Statistics 100 Exam 1 ........................................ February 11$^{th}$, 2015
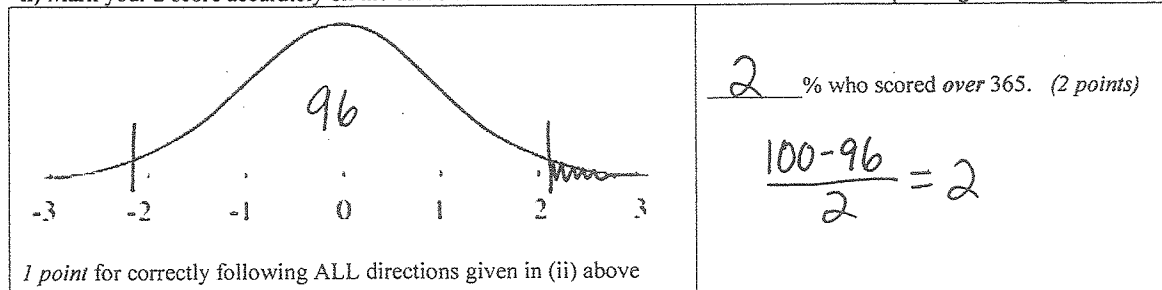
**Question 43** *(10 points total)*
At Hogwarts School of Witchcraft and Wizardry, students in their 5$^{th}$ year must take OWL (Ordinary Wizarding Level) exams. The OWL scores of the witches & wizards happened to follow the normal curve with <u>**Average = 300**</u> & <u>**SD=31**</u>.

**a)** Hermione Granger scored an outstanding score of 365 on her OWL exams! What percentage of the witches and wizards at Hogwarts scored better than Hermione? In other words, what percentage of people scored over 365?

**i)** First, convert 365 to a Z-score. (Remember: average =300 and SD =31) **Show work for full credit and round your z-score to 2 decimal places.** *(2 points)*

Z-score = __2.10__     $Z = \frac{365-300}{31} = 2.10$

**ii)** Mark your z-score accurately on the curve & mark the middle area. Shade the area corresponding to scoring over 365.



96

-3  -2  -1  0  1  2  3

*1 point* for correctly following ALL directions given in (ii) above

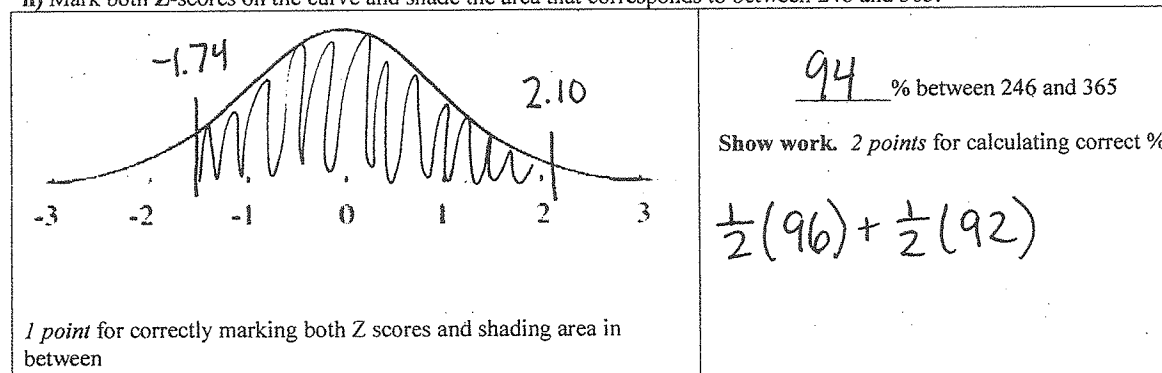__2__ % who scored *over* 365. *(2 points)*

$\frac{100-96}{2} = 2$

**b)** Ron Weasley scored a 246 on his OWLs. What percentage of students scored in between Ron and Hermione? In other words, what percent of students scored between 246 and 365?

**i)** Convert 246 to a Z-score. (You've already converted 365 to a Z score above) **Show work for full credit.** *(2 points)*

Z-score for 246= __-1.74__     $Z = \frac{246-300}{31} = -1.74$

**ii)** Mark both Z-scores on the curve and shade the area that corresponds to between 246 and 365.



-1.74          2.10

-3  -2  -1  0  1  2  3

*1 point* for correctly marking both Z scores and shading area in between

__94__ % between 246 and 365

**Show work.** *2 points* for calculating correct %
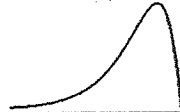
$\frac{1}{2}(96) + \frac{1}{2}(92)$

**Question 44-45 pertain to the following:** *(2 points total)*
Below are rough sketches of 2 histograms. One depicts scores on an easy exam where most students did well. One depicts scores on a hard exam where most students did poorly. The horizontal axis ranges from 0% to 100%.

Histogram I                    Histogram II



0        100            0        100

**44.** *(1 point)* Which histogram depicts the easy exam?     **a)** Histogram I     (**b)** Histogram II

**45.** *(1 point)* Which histogram has the average greater than the median?
(**a)** Histogram I          **b)** Histogram II          **c)** Both histograms          **d)** Neither histogram

February Page 7
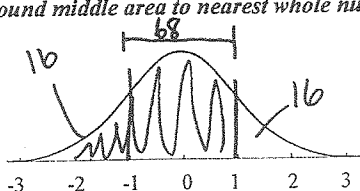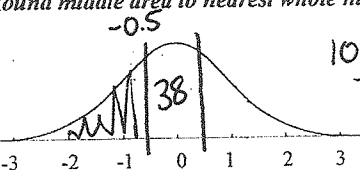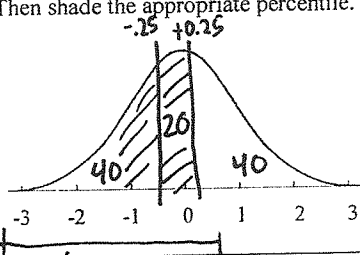
February 11th, 2015

**Question 46** *(11 points total)*
Last year, STAT 100 students were asked the question: How many valentines did you get on Valentine's day? The histogram of the number of valentines students received followed a normal curve with an *average = 520 and a SD = 80.* (Yes, STAT 100 students are popular on campus!) In the table below you're given either a student's valentine count, Z score, or percentile and you have to fill in the missing blanks.
**For all these problems, please round the areas given in the normal table to the nearest whole number.**

| Valentine Count | Z score | Percentile |
|---|---|---|
| Eleanor got 600 valentines. | Z= __1__ *(1 point)* Show work: $$Z = \frac{600-520}{80} = 1$$ | *Eleanor is in the* __84th__ *percentile (Fill in the blank-1 point)* <br> Mark z-score on curve & divide the curve into a middle area and 2 tails. Then shade the appropriate percentile. *(1 point)* *Round middle area to nearest whole number.* <br> 16   68   16   68+16=84 <br> -3 -2 -1 0 1 2 3 |
| Elliot got __480__ valentines. *(Fill in the blank- 1 point)* Show work: $$Value = 520 + (-0.5)(80$$ Do NOT round answer. | Z= -0.5 | *Elliot is in the* __31st__ *percentile. (Fill in the blank-1 point)* <br> Mark z-score on curve & divide the curve into a middle area and 2 tails. Then shade the appropriate percentile. *(1 point)* *Round middle area to nearest whole number.* <br> -0.5  38   $$\frac{100-38}{2} = \frac{62}{2} = 31$$ <br> -3 -2 -1 0 1 2 3 |
| Harmony got __540__ valentines. *(Fill in the blank- 1 point)* Show work: $$Value = 520 + (0.25)(80)$$ Do NOT round answer. | Z= __0.25__ *(1 point)* | *Harmony is in the 60th percentile* <br> What middle area should you look up on the normal table to find the correct Z score? __20__ % *(Fill in blank- 1 point)* <br> *If the middle area is between two lines on the table, use the closest line.* <br> Mark z-score on curve & divide the curve into a middle area and 2 tails. Then shade the appropriate percentile. <br> -.25 +0.25  20  40   40 <br> -3 -2 -1 0 1 2 3 <br> 60 |
| William got __500__ valentines. *(Fill in the blank- 1 point)* Show work: $$Value = 520 + (-0.25)(80)$$ | Z= __-0.25__ *(1 point)* | **William is in the 40th percentile.** <br> No work is necessary. Just use the normal curve from above for the 40th percentile to get the corresponding z-score. |