Name: **Answer Key** Per: __________ Date: __________________
Serafino · ProbStatPrecalc M T W R F

Show all work for each answer. If you used your calculator, write in what your typed.

T1–R Statistics

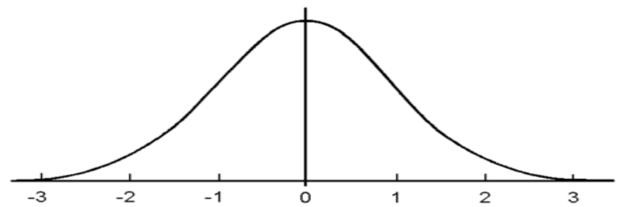
Test 1 Review

1. The weights of Roma tomatoes are normally distributed with a mean of 74 grams with a standard deviation of 3.3 grams.

   a. Label the bell curve with the appropriate weights at each standard deviation value.

   b. What are the lower and upper boundaries within 3 standard deviations (99.7%)?

   64.1 and 83.9



64.1 67.4 70.7 74 77.3 80.6 83.9

   c. What percent of tomatoes weigh…

   i. More than 80 g? ii. Less than 75 g? iii. Between 68 and 78 g?\

   NormCdf(80, 999, 74, 3.3) = 3.45%

   NormCdf(−999, 75, 74, 3.3) = 61.91%

   NormCdf(68, 78, 74, 3.3) = 85.27%

   d. What are the weights of…

   i. The lightest 15%? Ii. The heaviest 30%? Iii. The central 25%?

   InvNorm(.15, 74, 3.3) 70.58 g and less

   NormCdf(0.7, 74, 3.3) 75.73 g and more

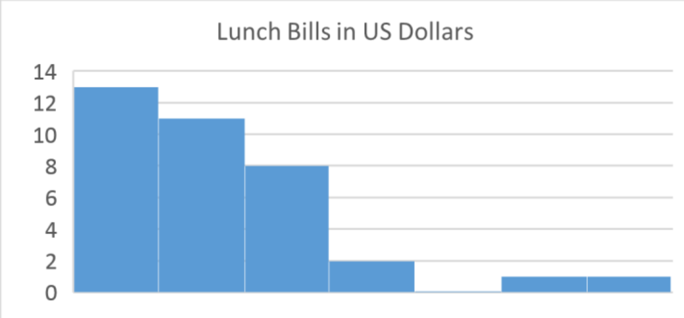
   NormCdf for .375 and 0.625 between 72.945 and 75.052 g

2. A college professor gives test grades as z-scores. Ben got a z–score of 2.20. Jerry got a z–score of 0. Explain what each of their scores mean about the performance of these students. *(i.e. Think: what information does this scoring methodology give the students? How did they do, relative to the rest of the class?)*

   Ben's z–score of 0 means his test score was exactly the mean of the class; he did better than 50% of the class.
   Jerry's z–score of 2.2 means he performed 2.2 standard deviations above the mean; he did better than 98.6%
   NormalCdf(−999, 2.2, 0, 1)

3. A 37–gram mouse and a 1260–lb moose are arguing about their body weights. Each one thinks they are heavier, relative to their respective groups. The mouse's colony has a mean weight of 30 grams and a standard deviation of 3.4 grams. The moose's herd has a mean weight of 910 pounds and a standard deviation of 185 pounds. Who is right? Support your answer with calculations and what they mean.

   The mouse: z = (37–30)/3.4 = 2.0588 The moose (1260 – 910) / 185 = 1.8919
   The mouse is heavier, relative to its population (98^{th} percentile of its colony). The moose is only in the 97^{th} percentile)



Milwaukee is shown below.

63 74 76 80 72 50 40 28

58.5, 73, 80 Range: 57

IQR: 38 Outliers: 35 – (38 x 1.5) = –22 73 + (38 x 1.5) = 130 **None**

5. A new lunch place is opening in town, and they want RHS students to be a main lunchtime demographic. The manager wants to find out how much the average student–generated lunch bill is in various restaurants around town. A random sample of 36 local lunch orders was collected.

2	3	4	4	5	6	7	7	7
8	9	9	10	11	12	12	13	14
14	15	16	17	18	19	21	22	23
23	25	26	27	28	29	30	49	63

a. Represent your data in the following ways (by hand):

i. Stem–and–leaf plot:

Stem	Leaf
0	2, 3, 4, 4, 5, 6, 7, 7, 7, 8, 9, 9
1	0, 1, 2, 2, 3, 4, 4, 5, 6, 7, 8, 9
2	1, 2, 3, 3, 5, 6, 7, 8, 9
3	0
4	9
6	3

ii. Histogram:

0 9.5 19.5 29.5 39.5 49.5 59.5 69.5

b. Are there any outliers? Should that data be kept in or not? What might explain them?

Q1 = 7.5, Q3 = 23; IQR = 15.5 7.5 – 1.5(15.5) = –15.75 23 + 1.5(15.5) = 46.25

There are two outliers; 49 and 63. Answers may vary. They could be because those bills were for a large group.

c. A few months later, the luncheonette is in business. They introduce a "Student Lunch Combo Special" at a discounted price, which they advertise on Instagram. The manager wanted to know the effectiveness of that ad, so after every lunch order was placed, they noted whether or not the Combo was ordered, and customer was asked, "Did you see the ad for the Combo on Instagram?"

i. Is this a good way to gauge whether or not the ad is effective? Explain why or why not.

The survey may give some information but it poses a few problems: • The survey isn't random, they're only asking the people who are already in the restaurant… and only asking the person at the register. • They're not asking whether or not the ad affected their lunch choice. • People can lie or forget or are unaware of whether something motivates a choice. • The customer may not follow the restaurant on Instagram at all, etc.

Data from the "lunch combo survey" is recorded in the **Frequency table.**

ii. Fill in and label the marginal distributions

	Bought Combo	Did Not Buy Combo	Totals
Saw the ad	32	33	65
Did not see the ad	17	18	35
Totals	49	51	100

iii. Make a Relative Frequency Table

	Combo	No Combo	Totals
Ad	32/100 = 32%	33/100 = 33%	65/100 = 65%
No Ad	17/100 = 17%	18/100 = 18%	35/100 = 35%
Totals	49/100 = 49%	51/100 = 51%	100/100 = 100%

iv. Conditional Relative Frequency Tables:

	Combo	No Combo
Ad	32/49= 65.3%	33/51 = 64.7%
No Ad	17/49 = 34.6%	18/51 = 35.3%
Totals	49/100 = 49%	51/100 = 51%

	Combo	No Combo	Totals
Ad	32/65 = 49.2%	33/65 = 50.8%	65/100 = 65%
No Ad	17/35 = 48.6%	18/35 = 51.4%	35/100 = 35%

i. What do you think? Is the ad effective? Explain why or why not and support with calculations.

Answers may vary… but use data from the Relative Frequency & Conditional Frequency Tables. Sample answers: Yes, because 65% of the customers saw the ad 65% of the people who bought the combo saw the ad. No, because 65% of the people who didn't buy the combo also saw the ad. And amongst people who saw the ad, about half bought it and half didn't.

6. Use the distribution below to answer the following questions:

a. Is the data normal, uniform, skewed left or right? Skewed left

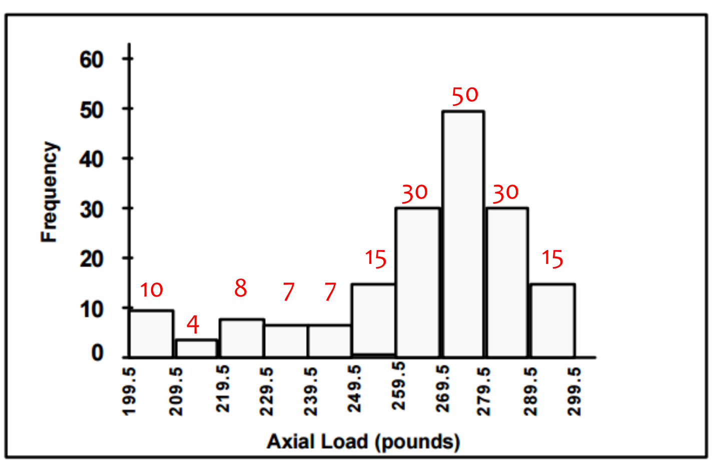
b. Which is greater? The mean or the median? The median is greater

7. What is the mean for the histogram?

10(204.5) + 4(214.5) + 8(224.5) + 7(234.5) + 7(244.5) + 15(254.5) + 30(264.5) + 50(274.5) + 30(284.5) + 15 (294.5) = 46,412

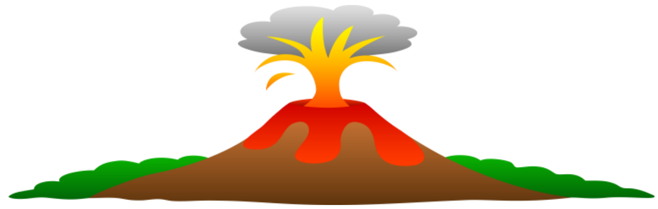
10+4+8+7+7+15+30+50+30+15 = 176

46,412 / 176 = **263.7045**



8. Several years ago, a volcano erupted. Lava is still oozing down its mountainside. The following ordered pairs gives one scientists' prediction of height of the lava, in y–feet, x -years from this year:

(2, 253), (3, 241), (5, 215), (6, 205), (9, 168)



a. What is the LSR model?

h = –12.13333t + 277.0667

b. What is the correlation coefficient? r = –.99

c. What is the average rate of change for the lava? Describe in words.

The lava loses altitude at an average of 12.1333 feet per year

d. What is the altitude of the lava this year?

This year is 0 years from now, so h(0) = 277.067

e. Predict the height of the lava in 2030.

The year 2030 will be 14 years from now, so h(14) = 107.2 ft

f. In what year will the lava finally hit the ocean?

0 – 012.1333t + 277.0667; t = 22.835 Approximately 2038

g. What is the residual for x = 6? What does that mean, in real life?

h(6) = 204.267. The data point is 205, which is .0733 above the LSR line. The residual is 0.733, meaning the scientist's prediction is 0.733 feet above the LSR prediction.

h. The volcano from which the lava erupted is 1,000 feet high. In what year did the volcano erupt?

1000 – 012.1333t + 277.0667 t = –59.58 In the year 1956 or 1957

9. Use the information to crease the LSR line:

$\bar{x} = 121.67$, $\bar{y} = 167.5$, $s_x = 107.165$, $s_y = 165.39$, $r = .967$

First, you use find Slope: $m = \frac{S_y}{S_x} r$ $\frac{165.39}{107.165}(.967) = 1.4924$

Then, solve the familiar y = mx + b equation, but use the means for x and y instead of the coordinates.

$\bar{y} = m\bar{x} + b$ 167.5 = (1.4929)(121.67) + b b = –14.1411 $\hat{y} = 1.4924x - 14.1411$