

Name: _____

Student Number: _____

Exam 1A

**Introductory Statistics
Resource Economics 211**

Please read the following rules and sign below to indicate that you do understand our rules. Thanks!

- ❖ Please have nothing but a pencil/pen and your calculator on your desk. All book bags, backpacks need to be stored at the front of the class.
- ❖ You cannot share calculators.
- ❖ You cannot talk with those around you. Raise your hand, if you have any questions or problems.
- ❖ No cell phones. If you are seen with a cell phone, we will take your exam.
- ❖ Please check your Exam; there should be **8 pages** and three parts:
 - A – information for us.
 - B – 15 MC questions.
 - C – Problems, **questions 1-5**. You must **provide exact formulas where required**.
- ❖ We will provide you with a formula sheet and a bubble sheet. Only the bubble sheet answers count for the MC questions.
- ❖ You will need a #2 pencil for the bubble sheet.
- ❖ Neatness counts! If we cannot read your answers, then we cannot give you credit.
- ❖ Please keep the parts stapled together! To be safe, please put your name on all pages!
- ❖ Relax, you'll think better. Stress is bad. Good Luck.

Okay, Dan. I've read and agree to these rules. _____

Part A. No points here, just necessary information. Declare your weights. You can revise these weights at Exam 2 and again at the Final. Answer the following questions on the ScanTron Bubble Answer sheet with a number 2 pencil.

1. What is the color of your exam?
 - a. Yellow
 - b. Blue

2. What weight do you wish to place on Exam 1?
 - a. 10%
 - b. 20%
 - c. 30%

3. What weight do you wish to place on Exam 2?
 - a. 10%
 - b. 20%
 - c. 30%

4. What weight do you wish to place on the Final Exam?
 - a. 20%
 - b. 30%
 - c. 40%

5. What weight do you wish to place on OWL quizzes?
 - a. 10%
 - b. 20%

Name: _____

Student Number: _____

Part B: 2 points each. Answer the following multiple choice questions on the ScanTron Bubble Answer sheet.

6. In the section of the Gallup article The Interview Itself, their telephone methods are discussed. Gallup typically tries several times to contact adult household members in their telephone sample. What possible problem are they trying to avoid?
- Missing data.
 - Selection bias.
 - Non-response bias.
 - Measurement error.
 - None of the above
7. The universe or "totality of items or things" under consideration is called
- a parameter.
 - a statistic.
 - a population.
 - a sample.
8. The chancellor of a major university was concerned about alcohol abuse on her campus and wanted to find out the portion of students at her university who visited campus bars every weekend. Her advisor took a random sample of 250 students and computed the proportion or percentage of the students in the sample who visited campus bars every weekend. The proportion or percentage of the students in the sample who visited campus bars every weekend is an example of a
- Continuous random variable.
 - Categorical random variable.
 - Sample statistic.
 - Discrete random variable.
 - Population parameter.
9. Those methods involving the collection, presentation, and characterization of a set of data in order to properly describe the various features of that set of data are called:
- Descriptive statistics.
 - Statistical inference.
 - The scientific method.
 - Sampling.
 - None of the above
10. The number of observations that fall into one of the distinct classes for the distribution divided by the total number of observations in the data set is called a
- relative frequency.
 - descriptive statistic.
 - frequency.
 - class boundary.
 - None of the above

Name: _____

Student Number: _____

11. If the age distribution of customers at a major retail chain is thought to be bell shaped with a mean equal to 43 years and a standard deviation equal to 7 years, the percentage of customers between the ages of 29 and 50 years is:
 - a. At least 75.
 - b. Approximately 95.
 - c. Approximately 68.
 - d. Approximately 48.
 - e. Approximately 82.

12. Classes should be set up so that they do not overlap; each observation should fit in one and only one class. Classes with this property are called
 - a. frequency classes.
 - b. equal width classes.
 - c. mutually exclusive classes.
 - d. all inclusive classes.
 - e. None of the above

13. Which of the following data collection methods will most likely generate the largest non-response?
 - a. Telephone surveys.
 - b. Personal interviews.
 - c. Direct observation.
 - d. Mail surveys.
 - e. None of the above.

14. Which of the following is most likely to be a parameter as opposed to a statistic?
 - a. The average height of people randomly selected from a database.
 - b. The average score of the first five students completing an assignment.
 - c. The proportion of trucks stopped yesterday and cited for bad brakes.
 - d. The proportion of females registered to vote in a county.

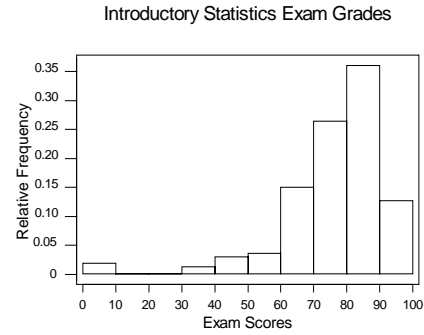
15. In conducting a personal interview, what problem can result if the interviewer is allowed to arbitrarily decide who should be interviewed?
 - a. Missing data.
 - b. Bias.
 - c. Non-responses.
 - d. None of the above

16. Suppose two researchers each drew samples from the same population of equal size. What can we expect from their analyses?
 - a. This is a poor question. You can't draw more than one sample from a population.
 - b. Their sample means will be equivalent, only the individual data values in their samples will be different.
 - c. Their sample means are estimates of the population mean; they will likely be different.
 - d. The sample statistics they present will likely be identical since both samples were drawn from the same population.
 - e. None of the Above

Name: _____

Student Number: _____

Use the histogram at the right to answer each of the following questions.



17. Approximately what *proportion* of students earned a grade of 70 or higher on the exam?

- a. 0.75
- b. 0.50
- c. 0.15
- d. 50
- e. 75

18. Approximately what *percentage* of students earned a grade of less than 60?

- a. 0.05
- b. 0.10
- c. 0.15
- d. 25
- e. 10

The stem and Leaf diagram at the right shows the final exam grades of 40 students in an advanced biology course. Use this diagram to answer each of the following two questions.

19. How many students earned a grade of 80 or greater on the final exam?

- a. 4
- b. 17
- c. 8
- d. 21
- e. 30

Stem-and-leaf of FinalExam N =40
Leaf Unit = 1.0

```
2 0 00
2 1
5 2 068
7 3 37
10 4 128
19 5 125666899
(4) 6 4579
17 7 000112778
8 8 1455677
1 9 9
```

20. A passing grade was 55. How many students earned a *passing grade or better* the final exam?

- a. 7
- b. 19
- c. 21
- d. 28
- e. 45

Name: _____

Student Number: _____

Part C: Please answer all questions in this section. Point totals are provided in the left margin.

Show all work: No formulas, no credit. No calculations, no credit.

1. The Cheetah is the fastest land mammal and has been clocked at speeds in excess of 70 mph. The data at the right gives the speeds, in miles per hour (mph), over a ¼ mile for 35 cheetahs. Use these data to complete the following questions.

- (2) a. The data at the right are conveniently organized for you. What has been done to the data and what would you call this column of data?

These data have been sorted in ascending order – they constitute an **array**.

- (14) b. Complete a **grouped data table** for these data. Use **52 mph as the first limit** and a **class width of 2 mph**. The table below has sufficient space for you to complete this task. Be sure to include and properly label the three most important columns. **Where necessary, round to 3 decimal places.**

Cheetah Speed (mph)	Frequency	Relative Frequency
52 < 54	2	0.057
54 < 56	4	0.114
56 < 58	7	0.200
58 < 60	7	0.200
60 < 62	8	0.229
62 < 64	3	0.086
64 < 66	1	0.029
66 < 68	1	0.029
68 < 70	1	0.029
70 < 72	1	0.029
72 < 74	0	0.000
	35	1.000

SPEED

52.4
52.6
54.0
54.8
54.9
55.5
56.0
56.5
56.6
57.5
57.6
57.8
57.8
58.1
58.3
58.5
58.7
58.8
59.6
59.7
60.1
60.3
60.4
60.6
60.7
60.8
60.9
61.6
62.3
62.6
63.4
64.0
66.2
68.2
71.3

Important features:

- ❖ Good column labels on the table.
- ❖ Correct classes including the lower limits and upper limits. The upper limits are not included in the classes for these continuous data.
- ❖ Correct counts to determine the frequencies.
- ❖ Correct relative frequencies – simply divide each frequency by the total number of observations (35) to get the relative frequencies.
- ❖ Column totals. These help you check your work. Frequencies sum to n and the relative frequencies sum to 1.000 or something very close.

Name: _____

Student Number: _____

(11) c. Using the same cheetah data, determine a **5-Number Summary**. (The data are presented again in the right margin.) Place the 5-Number Summary in the boxes below and label each box. **Show formulas used and all work required** to determine the 5-Number Summary.

Min.	Q1	Q2 (Median)	Q3	Max.
52.4	56.6	58.8	60.9	71.3

Quartile Locations:

$$Q1 @ \frac{(n+1)}{4} @ \frac{(35+1)}{4} @ 9$$

$$Q2 @ \frac{2(n+1)}{4} @ \frac{2(35+1)}{4} @ 18$$

$$Q3 @ \frac{3(n+1)}{4} @ \frac{3(35+1)}{4} @ 27$$

These are the locations. Go to these observations to find the values for the quartiles:

$$Q1 = x_9 = 56.6$$

$$Q2 = x_{18} = 58.8$$

$$Q3 = x_{27} = 60.9$$

- (2) d. The mean speed for these 35 cheetahs was 59.4 mph. What would you conclude about the **shape of the distribution**?

Given that the mean is greater than the median, we would expect that the distribution is skewed to the right, or there are outlying or extreme values. But, there is not too much of a difference, so we might also expect the distribution is fairly symmetrical.

- (6) e. Are there any **potential outliers** among these data? Show the formulas and determine the lower and upper limits. **What do you conclude?**

$$\text{Lower Limit} = Q1 - 1.5 (IQR) = 56.6 - 1.5 (4.3) = 50.15$$

$$\text{Upper Limit} = Q3 + 1.5 (IQR) = 60.9 + 1.5 (4.3) = 67.35$$

Because the final two values in the data set exceed the upper limit, we would label those two values (68.2 and 71.3) as outlying observations. We may want to delete them from the data set before doing any further analyses.

SPEED

52.4
52.6
54.0
54.8
54.9
55.5
56.0
56.5
56.6
57.5
57.6
57.8
57.8
58.1
58.3
58.5
58.7
58.8
59.6
59.7
60.1
60.3
60.4
60.6
60.7
60.8
60.9
61.6
62.3
62.6
63.4
64.0
66.2
68.2
71.3

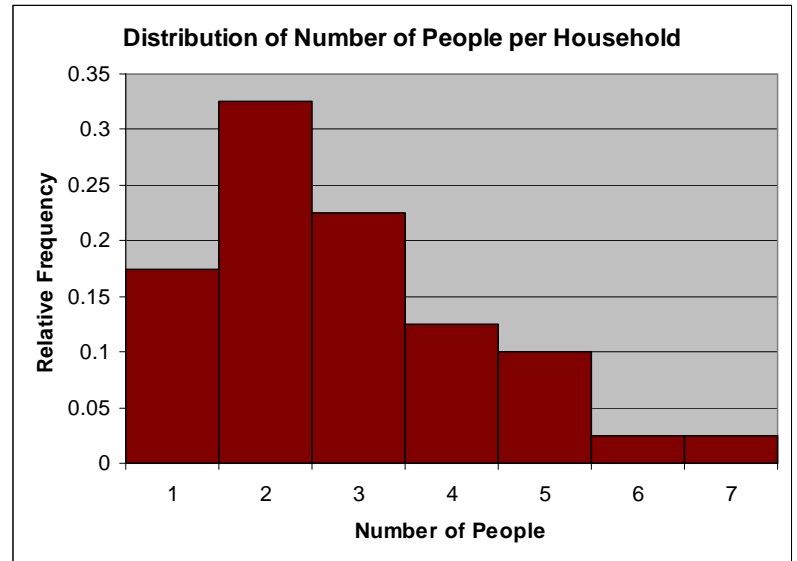
Name: _____

Student Number: _____

2. Construct a graph to show the following distribution. Axes are provided. Be sure to **completely label each graph**, use the space wisely and **label the graph completely**.

- (8) The following table gives the distribution for number of people per household for a sample of 40 US households. Construct a graph to show the *relative frequency distribution*.

<i>Number of People</i>	<i>Frequency</i>	<i>Relative Frequency</i>
1	7	0.175
2	13	0.325
3	9	0.225
4	5	0.125
5	4	0.100
6	1	0.025
7	1	0.025



I did my "histogram" in Excel. Importantly, the axes are properly labeled, it is a **relative frequency histogram**, and it has a title! Because these are discrete data, with just a few classes (single valued classes) a bar chart is also acceptable.

- (4) 3. The following regression was estimated relating the prices of used Honda Accords to the car's mileage:
 $y = 16,400 - 0.0315x$, where y is the price in \$ and x is the number of miles on the odometer.
a. Predict the value of a used 2000 Accord with 92,000 miles. (Round to the nearest \$.)

Use the regression equation and just plug in the mileage for the car in question:

$$y = 16,400 - 0.0315x = 16,400 - (0.0315) \cdot (92,000) = \$13,502$$

- b. Suppose you can purchase this car for \$12,500. Would you be willing to pay this amount? **Explain.**

This car looks to be a pretty good deal. Our regression suggests it's worth \$13,502 and you can get it for \$12,500. Of course, you'd want to see the car first!

- (3) 4. The variable X is known to have a population mean of 4,500, but we have no information about the shape of the distribution. At least what percentage of the data will fall within 2 standard deviations of the population mean μ ? Report your answer as a percentage to one decimal place.

Use Chebychev's rule for this. $\left(1 - \frac{1}{k^2}\right) \cdot 100\% = \left(1 - \frac{1}{2^2}\right) \cdot 100\% = \left(1 - \frac{1}{4}\right) \cdot 100\% = 75\%$

Name: _____

Student Number: _____

5. Below are IQ scores for a random sample of 10 students. Use these data to answer the following questions.

IQ Score (x)	$(x - \bar{x})$	$(x - \bar{x})^2$	x^2
91	-23.4	547.56	8281
97	-17.4	302.76	9409
101	-13.4	179.56	10201
102	-12.4	153.76	10404
110	-4.4	19.36	12100
116	1.6	2.56	13456
122	7.6	57.76	14884
126	11.6	134.56	15876
137	22.6	510.76	18769
142	27.6	761.76	20164
1144	0.00	2670.40	133544

- (4) a. Compute the *mean* for these data. Be sure to **show the formula used** and **round to 1 decimal place**.

$$\bar{x} = \frac{\sum x}{n} = \frac{1144}{10} = 114.4$$

- (10) b. Complete the additional columns needed in the table to show *deviations* for all observations and to calculate the *standard deviation*. Label your columns and include the formula for the standard deviation. (**Round to two decimal places.**)

The formula for a deviation is shown as the column heading in the table. All deviations are show as well as all squared deviations.

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} = \sqrt{\frac{2670.40}{9}} = 17.23$$

- (6) c. Now, verify your answer in part b by using the **short-cut formula to compute the standard deviation**. Use the final column of the table and create the additional data needed to use the short-cut formula.

The additional required column is for the values x^2 . They are shown in the last column of the table.

$$s = \sqrt{\frac{\sum x^2 - (\sum x)^2/n}{n-1}} = \sqrt{\frac{133,544 - (1144)^2/10}{9}} = 17.23$$