**DATA Set: Indicate one: Apartments**

Directions: Fill in the answer to the questions posed below. Attach the specified printouts where indicated and upload the project in Canvas on or before **the deadline**. **Please give specific interpretations as generic answers will not receive full credit.**

1. Unemployment is a big concern due to people being laid off due to the Coronavirus pandemic. It is estimated that 20% of all households have been directly affected. Suppose we randomly sample 25 households and determine the number of households that have been directly affected.
2. Identify the type of distribution that can be used to analyze this data. Make sure to identify any parameters that would be needed in this identification (ex. N, p, µ, σ).
3. What is the probability that at least 6 of these 25 sampled households has been directly affected? Show work to get full credit.
4. What is the probability that at least 4 but less than 9 of the sampled household has been directly affected? Show work to get full credit.
5. Provide a printout in which you find the mean and standard deviation for the 75 **sizes/mileages** of your data set. Also, list your size/mileage for your 10th and your 40th observation below. For all of question 2, treat these values as the mean and standard deviation of the population when answering questions.

Variable N Mean SD

Size 75 1316.6 614.03

Size/Mileage for Observation #10: \_\_\_\_\_\_\_\_\_\_\_

Size/Mileage for Observation #40: \_\_\_\_\_\_\_\_\_\_\_

* 1. For this problem, assume that the distribution for all sizes/mileages in the population are normally distributed (may not really be true). Find the proportion of all sizes/mileages that would fall between the sizes/mileages of your 10th and 40th observations. **To receive full credit, you should include pictures of the normal curve (labeled with both x and z-values) with the pertinent probabilities shaded in the picture.**



X



z

* 1. **Assuming the distribution of the sizes/mileages are normal, use the mean and standard deviation you found above to locate the 97.5th percentile in the size/mileage distribution. To receive full credit, you show as much work as possible.**



* 1. **Suppose you take a future sample of 100 sizes/mileages from your population. Describe the sampling distribution for the x-bar value that you would find. (hint: there should be three parts to this answer)**
	2. Use the sampling distribution to find the probability that the sample mean from that future sample would exceed the size/mileage for Observation #40. **To receive full credit, you should include pictures of the normal curve (labeled with both x-bar and z-values) with the pertinent probabilities shaded in the picture.**

$$\overbar{x}$$





z

1. Use two of the techniques discussed in class to assess the normality of the Price distribution from your data set. You can choose the two techniques you use. Just make sure you show us enough information (including printouts) to check your work.

Stem and Leaf Plot of size

Leaf Digit Unit = 100 Minimum 587.00

0 5 represents 500 Median 1123.0

 Maximum 3168.0

Depth Stem Leaves

 3 0 555

 17 0 66666677777777

 28 0 88899999999

 (14) 1 00000111111111

 33 1 222222333

 24 1 44455

 19 1 677

 16 1 8889

 12 2 01

 10 2 223

 7 2 4555

 3 2 67

 1 2

 1 3 1

Descriptive Statistics

Variable N Mean SD 1st Quarti 3rd Quarti

rent 75 1675.9 523.90 1300.0 1995.0

