**Multiple Choice [32 pts] -- choose the ONE BEST answer for each question by writing the corresponding letter in the blank to the left of the question.**

A\_\_\_\_\_\_

E\_\_\_\_\_\_

D\_\_\_\_\_\_

A\_\_\_\_\_\_

B\_\_\_\_\_\_

D\_\_\_\_\_\_

A\_\_\_\_\_\_

B\_\_\_\_\_\_

A\_\_\_\_\_\_

C\_\_\_\_\_\_

A\_\_\_\_\_\_

B\_\_\_\_\_\_

E\_\_\_\_\_\_

D\_\_\_\_\_\_

C\_\_\_\_\_\_

E\_\_\_\_\_\_

1. The heights of three waves were 1.7. 0.2, and 3.6 feet, respectively. This difference between individuals is an example of what type of variability?

**A.** Natural **B.** Extreme **C.** Process **D.** Variable **E.** Sampling

1. Two samples of 80 Oak trees each resulted in 31% and 47% of trees exhibiting wilt (a type of disease), respectively. This difference among statistics is an example of what type of variability?

**A.** Natural **B.** Extreme **C.** Process **D.** Variable **E.** Sampling

1. What is the summary of all possible individuals of interest, whether actually examined or not, called?

**A.** Sample **B.** Statistic **C.** Gang **D.** Parameter **E.** Population

1. What is the group of individuals actually examined in a statistical study called?

**A.** Sample **B.** Statistic **C.** Gang **D.** Parameter **E.** Population

1. What type of variable is level of difficulty rating for ski hills (from easiest to most difficult) – “green circle”, “blue square”, “black diamond”, “black diamond”?

**A.** Nominal **B.** Ordinal **C.** Response **D.** Continuous **E.** Discrete

1. What type of variable is the depth of base of snow for a ski hill (in meters)?

**A.** Nominal **B.** Ordinal **C.** Response **D.** Continuous **E.** Discrete

1. Which graph would best be used to examine the distribution of preferences for computer operating system (e.g., Windows, MacOS, Linux)?

**A.** Bar Chart **B.** Dot Plot **C.** Histogram **D.** Scatterplot **E.** Stemplot

1. What is the symbol used to represent the sample standard deviation?

**A.** x **B.** s **C.**  **D.**  **E.** Q3

1. The mean is \_\_\_\_\_\_\_\_ the median for an extremely left-skewed distribution.

**A.** less than **B.** equal to **C.** greater than **D.** five times **E.** a sibling of

1. On any normal distribution, what proportion of the individuals are within +2 of ?

**A.** 0.680 **B.** 0.900 **C.** 0.950 **D.** 0.997 **E.** 1.000

1. On a N(18,9) distribution, what proportion of the individuals are negative?

**A.** 0.025 **B.** 0.16 **C.** 0.50 **D.** 0.84 **E.** 0.975

1. What type of normal distribution question is “What percentage of dogs can jump higher than 1 meter?”

**A.** forward, left-of **B.** forward, right-of **C.** forward, between **D.** reverse, left-of **E.** reverse, right-of

1. What type of normal distribution question is “What is the length of duration such that 20% of light bulbs last longer?”

**A.** forward, left-of **B.** forward, right-of **C.** forward, between **D.** reverse, left-of **E.** reverse, right-of

1. What is the name of the variable that we are interested in predicting or explaining?

**A.** Continuous **B.** Discrete **C.** Explanatory **D.** Response **E.** Sampling

1. What type of study is it if every individual has the same chance of being selected for the sample?

**A.** Convenience **B.** Inference **C.** Simple Random **D.** Regression **E.** Voluntary Response

1. What type of study is it if the researcher sends a survey to the entire population, but only some of the individuals return it?

**A.** Convenience **B.** Inference **C.** Simple Random **D.** Regression **E.** Voluntary Response

**Answer the following two questions with your final answer clearly identified (e.g., circled). You must show all of your work to receive full credit (i.e., just providing the final answer will not receive full, if any, credit).**

1. Compute the mean **[4 pts]** and standard deviation **[8 pts]** for the following data: 67, 23, 43, 12, 34, and 13.

x diffs diffs.sq

1 67 35 1225

2 23 -9 81

3 43 11 121

4 12 -20 400

5 34 2 4

6 13 -19 361

sum 192 0 2192

Mean = x-bar = 192 / 6 = 32

Variance = s^2 = 2192 / 5 = 438.4 Std. Dev = s = sqrt(438.4) = **20.938**

18. Compute the median **[4 pts]** and IQR **[6 pts]** for these data: 45, 23, 65, 34, 89, 90, 34, 23, 32, 52, 57, 51, 87, and 7.

Median (**=48**) is the average of values in positions 7 and 8.

7 23 23 32 34 34 [45 51] 52 57 65 87 89 90

Q1 (**=32**) is the value in position 4 of the lower half.

7 23 23 [32] 34 34 45

Q3 (**=65**) is the value in position 4 of the upper half.

51 52 57 [65] 87 89 90

**Use RStudio to produce the result(s) needed to answer the next question. Circle your answer and include the R code used to produce your result. Hints about use of distrib() are on the next page.**

19. **[16 pts]** Suppose that it is known that the distribution of long “crosse”s (a stick used to handle a lacrosse ball) by all midfielders on women’s Division III lacrosse teams is normally distributed with a mean of 58 inches and a standard deviation of 3.2 inches. Use this information to answer the questions below *to one decimal place*.

1. What percentage of long “crosse”s are longer than 64 inches?

( distrib(64,mean=58,sd=3.2,lower.tail=FALSE) ) 0.03039636 🡪 **3.0%**

1. What is the length such that 20% of all long “crosse”s are shorter?

( distrib(0.2,mean=58,sd=3.2,type="q") ) 55.30681 🡪 **55.3**

1. What percentage of long “crosse”s are between 60 and 65 inches?

distrib(65,mean=58,sd=3.2)-distrib(60,mean=58,sd=3.2) 0.2516325 🡪 **25.2%**

1. What is the IQR for length of long “crosse”s?

( distrib(0.75,mean=58,sd=3.2,type="q") ) 60.15837 🡪 **60.2**

( distrib(0.25,mean=58,sd=3.2,type="q") ) 55.84163 🡪 **55.8**

1. What percentage of long “crosse”s are between than 55 and 59 inches?

distrib(59,mean=58,sd=3.2)-distrib(55,mean=58,sd=3.2) 0.448419 🡪 **44.8%**

1. What is the length such that 10% of all long “crosse”s are longer?

( distrib(0.1,mean=58,sd=3.2,type="q",lower.tail=FALSE) ) 62.10097 🡪 **62.1 in**

1. What percentage of long “crosse”s are longer than 48 inches?

( distrib(48,mean=58,sd=3.2,lower.tail=FALSE) ) 0.999111 🡪 **99.9%**

1. What is the IQR for length of long “crosse”s?

( distrib(0.75,mean=58,sd=3.2,type="q") ) 60.15837 🡪 **60.2**

( distrib(0.25,mean=58,sd=3.2,type="q") ) 55.84163 🡪 **55.8**

**Answer the following question in the space provided. Please be as specific as possible.**

20. **[12 pts]** The National Collegiate Athletic Association (NCAA) is interested in determining the mean team grade point

**I** -- \_ collegiate women’s athletic teams\_

**V** -- \_ team gpa for collegiate women’s athletic teams\_

**Po** -- \_ALL collegiate women’s athletic teams\_

**Pa** -- \_mean team gpa of ALL collegiate women’s athletic teams\_

**Sa** -- \_ 72 collegiate women’s athletic teams actually examined \_

**St** -- \_ mean team gpa of 72 collegiate women’s athletic teams actually examined \_

21. **[14 pts]** Maret and Collins (1996) set up an experiment to test the effects of food ration (low, medium, and high

a. What is the response variable? \_growth of larval salamanders\_

b. What is/are the factor(s)? \_food ration and density of competitor\_

c. What is/are the number of levels? \_3 & 3\_

d. What is the number of treatments? \_9\_

e. What is the number of replicates per treatment? \_5\_

**Complete a thorough univariate EDA appropriate to the type of variable in each of the following two questions.**

22. **[10 pts]** The ambient air temperature (oC) between 15-Nov and 31-Mar was measured in limestone caves in **Bimodal, outlier at 13.10, mean of 7.09, sd of 2.11**

23. **[4 pts]** The General Social Survey (GSS) received responses to the question “What is your religious preference?”

**Majority were Protestants, very few Jewish**

**Short (Paragraph) Answers [15 pts] -- Answer the following questions with complete sentences.**

24. Define natural and sampling variability. Provide a thoughtful example that depicts each type of variability.

25. Describe the major principles of experimental design and why is each important.

26. Describe at least three situations (be specific) where observational studies are valuable (even though strict cause-and-effect statements cannot be made).