**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.**

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

1. The number of plants in three plots was 44, 33, and 26, 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 Northland students found that 31% and 47%, respectively, were from Wisconsin. 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 the individuals actually examined?

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

1. What is all individuals, whether actually examined or not?

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

1. What type of variable is type of “smart phone” (e.g., iPhone, Android, Samsung Galaxy)?

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

1. What type of variable is the number of permanent docks on a lake?

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

1. Which graph would best be used to examine the distribution of people’s favorite breakfast restaurant in Ashland (e.g., Blue Wave, Breakwater, Buddies)?

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

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

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

1. The mean is \_\_\_\_\_\_\_\_ the median for an extremely right-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 +1 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 positive?

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

1. What type of normal distribution question is “What is the depth of ice such that 20% of lakes have more ice?”

**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 percentage of ink toner cartridges can print more than 600 pages?”

**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 individuals choose to be (or not to be) part of the sample?

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

1. What type of study is it if the researcher surveys respondents at location where many respondents are likely to be (thus, making it easy to get a large sample)?

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

**Answer questions 17 and 18 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: 13, 17, 87, 57, and 21.

18. Compute the median **[4 pts]** and IQR **[6 pts]** for these data: 28, 19, 32, 21, 1, 11, 27, 28, 13, 28, 30, 32, and 7.

**Answer questions 19 and 20 in the space provided. Please be as specific as possible.**

19. **[12 pts]** The Northland College Student Association is interested in determine the percentage of current (Winter 2017) Northland students that support a resolution requesting that the Board of Trustees divest Northland’s investment portfolio of companies with socially or environmentally questionable business practices. To estimate this percentage they asked a sample of 90 students whether they supported the resolution or not. Use this information to identify the **I**ndividual, **V**ariable, **Po**pulation, **Pa**rameter, **Sa**mple, and **St**atistic.

20. **[14 pts]** A student wanted to see how far she could hit two brands of softballs with three types of bats. The three bats were all 34 inches long, weighed 34 ounces, and looked similar. Four new balls of each brand were placed into a pitching machine in random order. The student picked one brand of bat at random to hit each of the eight pitched balls (assume that each ball was hit the first time it was pitched). The student recorded the distance (in feet) that each ball was hit. This process was repeated for each bat. Use this information to answer the following questions.

a. What is the response variable? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

b. What is/are the factor(s)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

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

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

**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 below.**

**library(NCStats)**

**distrib(#,mean=##,sd=##,lower.tail=FALSE,type=”q”)**

where **#** is replaced with the value of the quantitative variable or the area (i.e., the percentage as a proportion).

**mean=##** has ## replaced by the value of the population mean

**sd=##** has ## replaced by the value of the population standard deviation

**lower.tail=FALSE** is included for a “right-of” calculation

**type=”q”** is included for a reverse calculation

21. **[16 pts]** A company produces packages of grape tomatoes that are advertised as weighing 16 ounces. If a package weighs less than 15.8 ounces or more than 16.5 ounces it is rejected for sale. Suppose that it is known that the distribution of package weights prior to possible rejection is normal with a mean of 16.1 ounces and a standard deviation of 0.3 ounces. Use this information to answer the questions below *to one decimal place*.

1. What percentage of packages are rejected for being too heavy?
2. What percentage of packages are acceptable for sale?
3. What should the weight be if the company wanted to reject only 5% of packages for being too heavy?
4. What is the IQR for weight of packages?

**Complete a thorough univariate EDA appropriate to the type of variable in each of question 22 and 23.**

22. **[10 pts]** A student summarized the ages of United States Presidents on their inauguration day in Figure 1 and Table 1.

|  |  |
| --- | --- |
| **Figure 1.** Histogram of President’s ages. | **Table 1.** Descriptive statistics of Presidents ages.  Min. 42.9  1st Qu. 51.1  Median 55.2  3rd Qu. 58.9  Max. 70.6  Mean 55.5  St. Dev. 6.6 |

23. **[4 pts]** A recent Quinnipiac University poll asked a sample of 899 voters “How concerned are you about climate change?” The percentage responses are shown in Table 2.

**Table 2.** Percentage of respondents by answer to the climate change concern question.

Very Somewhat Not so Not Concerned

Concerned Concerned Concerned at all

45.2% 29.3% 14.4% 13.1%

**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 each is important.

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