**Multiple Choice [16 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 \_\_

\_\_ E \_\_

\_\_ B \_\_

\_\_ B \_\_

\_\_ D \_\_

\_\_ C \_\_

\_\_ B \_\_

\_\_ C \_\_

\_\_ A \_\_

\_\_ A \_\_

\_\_ A \_\_

\_\_ E \_\_

\_\_ D \_\_

\_\_ E \_\_

\_\_ C \_\_

1. The temperature on two days was 3oF and -17oF, respectively. This difference between individuals is an example of what type of variability?

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

1. Three samples of Northland students resulted in average preferred January temperatures of 20oF, 24oF, and 22oF, 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 set 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 summary of 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 the feeling of coldness – “bitter cold”, “cold”, “warm”, “hot”, “too hot”?

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

1. What type of variable is the daily high temperature (oF)?

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

1. Which graph would best be used to examine the distribution of the amount of money that townships in Wisconsin have spent on snow removal this year?

**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 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(6,3) 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 days was the temperature below zero?”

**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 temperature such that 20% of days were warmer?”

**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 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

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

**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 **[2 pts]** and standard deviation **[4 pts]** for the following data: 17, 24, 15, 6, 9, and 13.

x diffs diffs.sq

1 17 3 9

2 24 10 100

3 15 1 1

4 6 -8 64

5 9 -5 25

6 13 -1 1

sum 84 0 200

Mean = x-bar = 84 / 6 = **14**

Variance = s^2 = 200 / 5 = 40

Std. Dev = s = sqrt(40) = **6.324555**

18. Compute the median **[2 pts]** and IQR **[4 pts]** for the following data: 63, 22, 27, 29, 21, 22, 36, 38, 49, 38, 50, 53, 41, 54, 43, 46, 20, 60, 62, and 64.

**Median (=42)** is the average of values in positions 10 and 11.

20 21 22 22 27 29 36 38 38 [41 43] 46 49 50 53 54 60 62 63 64

**Q1 (=28)** is average of values in positions 5 and 6 of the lower half.

20 21 22 22 [27 29] 36 38 38 41

**Q3 (=53.5)** is average of values in positions 5 and 6 of the upper half.

43 46 49 50 [53 54] 60 62 63 64

**Use distrib() in 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 distrib() are on the next page.**

19. **[8 pts]** Suppose that it is known that the distribution of commute times for staff of Northland College is normally distributed with a mean of 9 minutes and a standard deviation of 2.5 minutes. Use this information to answer the questions below *to one decimal place*.

1. What percentage of the staff have a commute to campus longer than 13 minutes?

|  |
| --- |
| > ( distrib(13,mean=9,sd=2.5,lower.tail=FALSE) )  [1] 0.05479929 🡺 **5.5%** |

1. What is the commute time for the staff with the 10% longest commutes to campus?

> ( distrib(0.10,mean=9,sd=2.5,type="q",lower.tail=FALSE) )

[1] 12.20388 🡺 **12.2 minutes**

1. What percentage of the staff commute to campus in between 5 and 10 minutes?

> ( ab <- distrib(10,mean=9,sd=2.5) )

[1] 0.6554217

> ( a <- distrib(5,mean=9,sd=2.5) )

[1] 0.05479929

> ab-a

[1] 0.6006224 🡺 **60.1%**

1. What is the IQR for time to commute to campus?

> ( distrib(0.25,mean=9,sd=2.5,type="q") )

[1] 7.313776

> ( distrib(0.25,mean=9,sd=2.5,type="q",lower.tail=FALSE) )

[1] 10.68622 🡺 **Between 7.3 and 10.7 minutes**

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

20. **[6 pts]** An SCD (Sustainable Community Development) student was interested in determining the mean amount of

**I** -- \_**A Wisconsin community in 2012**\_

**V** -- \_**Amount spent on “green amenities**\_

**Po** -- \_**ALL Wisconsin communities in 2012**\_

**Pa** -- \_**Mean amount spent on “green amenities” by ALL Wisconsin communities in 2012**\_

**Sa** -- \_**34 Wisconsin communities examined**\_

**St** -- \_**Mean amount spent on “green amenities” by the 34 Wisconsin communities sampled in 2012**\_

1. **[7 pts]** Students designed a simple experiment to determine if the duration of a D cell battery differed significantly

a. What is the response variable? \_**Time to drain the battery**\_

b. What is/are the factor(s)? \_**Battery Type, Temperatures**\_

c. What is/are the number of levels? \_**4 and 3**\_

d. What is the number of treatments? \_**12**\_

e. What is the number of replicates per treatment? \_**4**\_

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

21. **[5 pts]** A Northland student examined the basal area (cm) of Hemlock at a site in Iron County. A histogram and

**Right-skewed, no outliers, median of 1.57 cm, and IQR of 0.63-2.71 cm**

22. **[2 pts]** The Strategic Research Initiative polled 802 Wisconsin residents last Fall and asked them “From what you

***Lots of leeway here* … More support than oppose, but those that oppose more strongly oppose than those that support.**

**Short (Paragraph) Answers [9 pts] -- Answer the following questions with complete sentences.**

1. Thoroughly describe what the two major goals of statistics are **AND** why each is important.

**Summarize large quantitities of data … so as to better identify the “story” the data are telling,**

**Make inferences … sample to population because we can’t see the entire population.**

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

**Natural … variability among individuals**

**Sampling … variability among statistics from samples**

**Example … make sure that sampling is not from two populatios.**

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

**Control – isolate the effect on the response variable.**

**Randomization – remove potential bias when selecting individuals from the popn or allocating to tx**

**Replication – removes potential bias of using too few individuals.**

**Multiple Choice [16 pts] -- choose the ONE BEST answer for each question by writing the corresponding letter in the blank to the left of the question.**

\_\_ E \_\_

\_\_ A \_\_

\_\_ B \_\_

\_\_ E \_\_

\_\_ D \_\_

\_\_ B \_\_

\_\_ C \_\_

\_\_ D \_\_

\_\_ A \_\_

\_\_ C \_\_

\_\_ A \_\_

\_\_ E \_\_

\_\_ A \_\_

\_\_ D \_\_

\_\_ C \_\_

\_\_ E \_\_

1. Three samples of Northland students resulted in average preferred January temperatures of 20oF, 24oF, and 22oF, respectively. This difference among statistics is an example of what type of variability?

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

1. The temperature on two days was 3oF and -17oF, respectively. This difference between individuals 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 group of individuals actually examined in a statistical study called?

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

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

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

1. What type of variable is the daily high temperature (oF)?

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

1. What type of variable is the feeling of coldness – “bitter cold”, “cold”, “warm”, “hot”, “too hot”?

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

1. Which graph would best be used to examine the distribution of the amount of money that townships in Wisconsin have spent on snow removal this year?

**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 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(6,3) 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 is the temperature such that 20% of days were warmer?”

**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 days was the temperature below zero?”

**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