Winter 2014, Exam1 **NAME: _____ _____/60**

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.

_____ 1. Three samples of Northland students resulted in average preferred January temperatures of 20°F, 24°F, and 22°F, respectively. This difference among statistics is an example of what type of variability?
A. Natural **B.** Extreme **C.** Process **D.** Variable **E.** Sampling

_____ 2. The temperature on two days was 3°F and -17°F, respectively. This difference between individuals is an example of what type of variability?
A. Natural **B.** Extreme **C.** Process **D.** Variable **E.** Sampling

_____ 3. 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

_____ 4. What is all possible individuals of interest, whether actually examined or not, called?
A. Sample **B.** Statistic **C.** Gang **D.** Parameter **E.** Population

_____ 5. What is the symbol used to represent the population mean?
A. $\bar{x}$ **B.** s **C.** μ **D.** σ **E.** Q3

_____ 6. What type of variable is the daily high temperature (°F)?
A. Nominal **B.** Ordinal **C.** Response **D.** Continuous **E.** Discrete

_____ 7. 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

_____ 8. Which graph would best be used to examine the distribution of responses to "what is your favorite month of the year?
A. Bar Chart **B.** Dot Plot **C.** Histogram **D.** Scatterplot **E.** Stemplot

_____ 9. 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

_____ 10. 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

_____ 11. Which measures should be used if the distribution is symmetric with no outliers?
A. $\bar{x}$ & s **B.** $\bar{x}$ & IQR **C.** $\bar{x}$ & range **D.** Median & s **E.** Median & IQR

_____ 12. On any normal distribution, what proportion of the individuals are within $\pm 2\sigma$ of μ?
A. 0.680 **B.** 0.900 **C.** 0.950 **D.** 0.997 **E.** 1.000

_____ 13. On any normal distribution, what proportion of individuals are between Q1 and Q3?
A. 0.500 **B.** 0.680 **C.** 0.900 **D.** 0.950 **E.** 0.997

_____ 14. On a N(6,3) distribution, what proportion of the individuals are positive?
A. 0.025 **B.** 0.16 **C.** 0.50 **D.** 0.84 **E.** 0.975

_____ 15. 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

_____ 16. 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

Answer the following two questions on a separate sheet of paper with the question number clearly labeled and 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).

17. Compute the mean **[2 pts]** and standard deviation **[4 pts]** for the following data: 14, 21, 12, 3, 6, and 10.

18. Compute the median **[2 pts]** and IQR **[4 pts]** for the following data: 53, 12, 17, 19, 11, 12, 26, 28, 39, 28, 40, 43, 31, 44, 33, 36, 10, 50, and 54.

Use `distrib()` in RStudio to produce the result(s) needed to answer the next question. On a separate sheet of paper, write your answer in a complete sentence with the code used to produce the result below your sentence.

19. **[10 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*.

    A. What percentage of people have a commute to campus longer than 12 minutes?

    B. What is the commute time for the staff with the 15% longest commutes to campus?

    C. What are the most common 80% of commute times to campus?

    D. What percentage of people commute to campus in between 5 and 10 minutes?

    E. What is the seventh decile for time to commute to campus?

```
library(NCStats)
distrib(x,mean=##,sd=##,lower.tail=XXXXX,type="X")
```

where **x** is replaced with the value of the quantitative variable or the area

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

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

lower.tail=XXXXX has XXXXX replaced with TRUE (default) for a "left-of" and FALSE for a "right-of" calculation

type="X" has X replaced with p (default) for a forward and q for a reverse question

---

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 money that Wisconsin communities spent on so-called "green amenities" in 2012. To examine this question the student obtained a sample of 34 communities and, from their published 2012 budgets, determined the amount that each spent on "green amenities." Use this information to identify the **I**ndividual, **V**ariable, **Po**pulation, **Pa**rameter, **Sa**mple, and **St**atistic.

I -- _____

V -- _____

Po -- _____

Pa -- _____

Sa -- _____

St -- _____

Complete a thorough univariate EDA appropriate to the type of variable in each of the following two questions. Your answer should be written with complete sentences on a separate sheet of paper.

21. **[5 pts]** A Northland student examined the basal area (cm) of Hemlock at a site in Iron County. A histogram and descriptive statistics for his sample is presented in Figure 1 and Table 1, respectively.

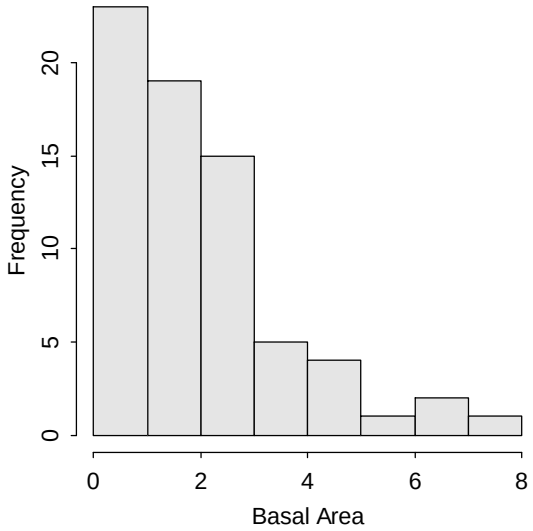

Figure 1. Histogram of Hemlock basal area.

Table 1. Descriptive statistics of Hemlock basal area.

mean	1.98
sd	1.67
min	0.08
Q1	0.63
median	1.57
Q3	2.71
max	7.94

22. **[2 pts]** The Strategic Research Initiative polled 802 Wisconsin residents last Fall and asked them "From what you know about the Affordable Care Act, also known as Obamacare, would you say that you strongly support, somewhat support, somewhat oppose, or strongly oppose this policy?" Table 2 contains the percentages of respondents by their level of support.

Table 2. Percentage of respondents by level of support for the Affordable Care Act.

Stronly Support	Somwhat Support	Somewhat Oppose	Strongly Oppose	Not Sure
22%	32%	12%	32%	2%

Short (Paragraph) Answers -- Answer <u>THREE</u> of the following questions with complete sentences on a separate sheet of paper. <u>Circle the questions below</u> that you have chosen to answer and make sure to clearly label your answers on the separate sheet. Each question is worth 3 points.

23. Thoroughly describe what the two major goals of statistics are **AND** why each is important.
24. Define natural and sampling variability. Provide a thoughtful example that depicts each type of variability.
25. **COMPLETELY** describe the underlying philosophical differences in how the mean and median measure center.
26. Describe **HOW** and **WHY** you would decide to use either the mean and standard deviation or the median and IQR to measure center and dispersion in a univariate EDA for quantitative data.
27. Describe two major principles or realities that lead to the importance of statistics in everyday life and scientific research.