Exam 1

STAT 251 Section 03

Student Name: _____ Last Four of Vandal Number: ___ ___ ___ ___

Test Version: **A**

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

1. (2pts) Multiple choice: A boxplot is a visual representation of the five-number summary of a quantitative variable. Which of the following provides the correct five numbers?

(a) minimum, mean, range, median, maximum

(b) minimum, Q1, median Q3, maximum

(c) minimum, Q1, mean, Q3, maximum

(d) minimum, IQR, mean, standard deviation, maximum

2.) (2pts) True or False: The selection of the number of bins k can significantly influence the shape of a histogram

Answer:_____

3.) (2pts) Multiple choice: A plot which shows distinct values of a variable on the x-axis and the frequency or relative frequency on the y-axis is called a:

(a) Dot plot

(b) Pie chart

(c) Box and Whisker Plot

(d) Histogram

4.) (2pts) Multiple choice: A preliminary exploration and summary of the data

(a) descriptive statistics

(b) sampling design

(c) interquartile range

(d) box and whisker plot

5.) (2pts) True or False: Inferential statistics can be applied to both samples and populations

Answer:_____

6.) (2pts) Write in the letter(s) of the words or phrases that best completes the following sentence: "The mean is the _____ of a distribution while the median is the _____ of a distribution."

(a) middle value

(b) frequency

(c) measure of spread

(d) center of gravity

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

7.) (2pts) Write in the letter(s) of the words or phrases that best complete the following sentence: "The variance is a _____ of a distribution. It is _____ to outliers"

(a) measure of spread

(b) susceptible

(c) measure of location

(d) resistant

8.) (6pts total) The provided data originates from a survey conducted among a representative sample of 20 college students in Georgia. This comprehensive study, administered by the state, aims to explore various factors influencing academic performance concerning student behavior. As part of the survey, participants were asked to self-report the daily time they dedicate to studying in whole hours. Complete the frequency table below by providing the missing information, and address the following questions (a)-(c):

Study Time (Hrs)	Frequency	Relative frequency	Cumulative RF
3	9	0.45	0.45
6		0.3	
9	3		0.9
12	1		
15	1	0.05	1

(a) (2pts) What kind of variable is “Study Time (Hrs)"?

(b) (2pts) Using the frequency table above, Confirm that the mean amount of time a student spends studying in this sample is 5.85 hours

(c) (2pts) What proportion of students study 6 hours a day or less?

9.) (4pts total) consider the following sample of 3 observations of a quantitative variable X

$$X = \{-4.4, -0.2, 1.5\}$$

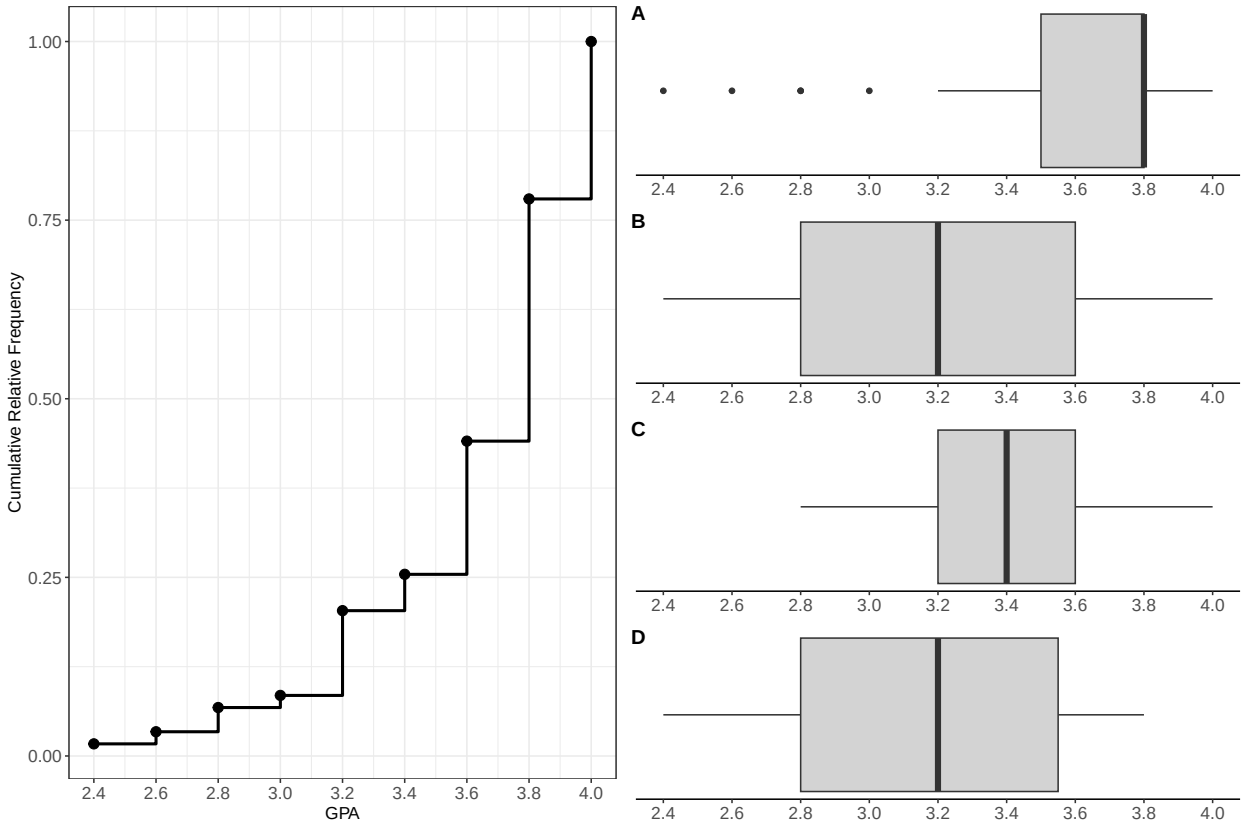
Answer parts (a) - (c)

(a) (2pts) given that $\bar{x} \approx 1$, **show how to compute** the sample variance s^2

(b) (2pts) Assuming the variance of X above $s^2 \approx 9.22$, what is the standard deviation s?

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

10.) (4pts total) The plot on the left shows the cumulative distribution of college grade point average from the Georgia Student Survey. Circle the letter of the boxplot that correctly depicts the cumulative distribution.



Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

11.) (4 pts total) A survey of college students in Georgia asked the following questions:

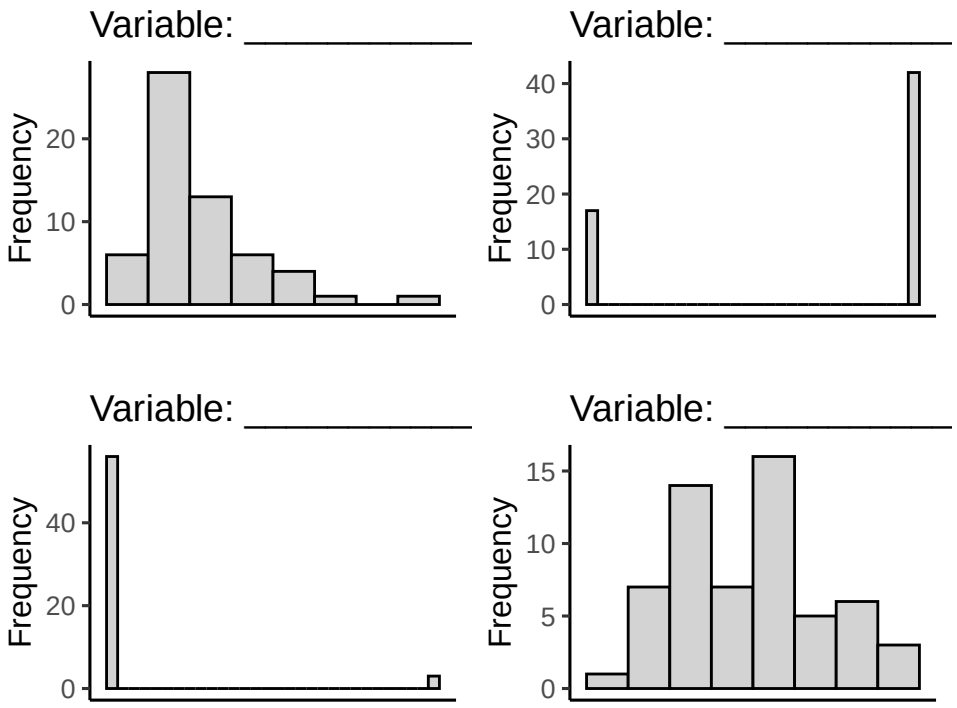
A.) Are you male or female? (recorded as male = 0, female = 1)

B.) What is you height in inches?

C.) Do you routinely smoke cigarettes? (recorded as no = 0, yes = 1)

D.) How many hours do you spend studying?

The figure below shows histograms of the student responses to each of these questions in scrambled order and without scale markings. Match each variable to its corresponding histogram by indicating the appropriate letter in the space provided on each plot.



12.) (2 pts) Consider the following five number summary of the sugar content in milligrams in a single serving of 20 different U.S cereal brands

Minimum	Q1	Median	Q3	Maximum
0	4	9.5	13	18

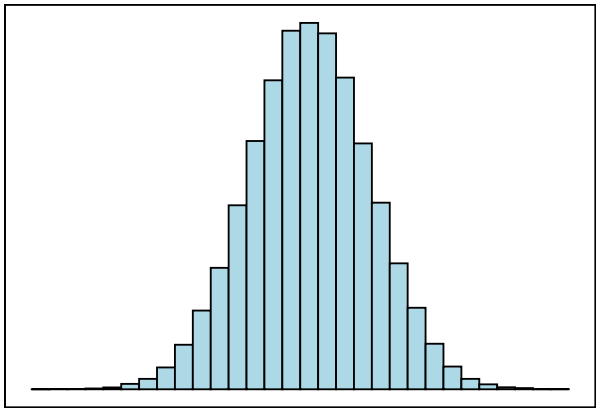
Utilizing the $1.5 \times IQR$ rule and the table above, at what sugar content level does a cereal brand become classified as an outlier?

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

13.) (6pts total) Describe the shape of the following distributions and for each distribution identify if the mean will be greater than, less than, or equal to the median (use symbols $<$, $>$, $=$).

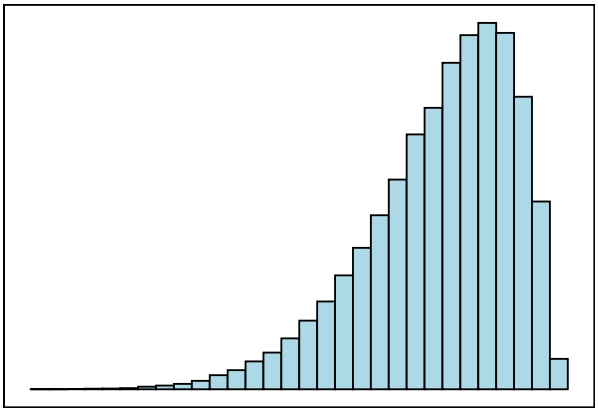
Shape: _____

mean _____ median



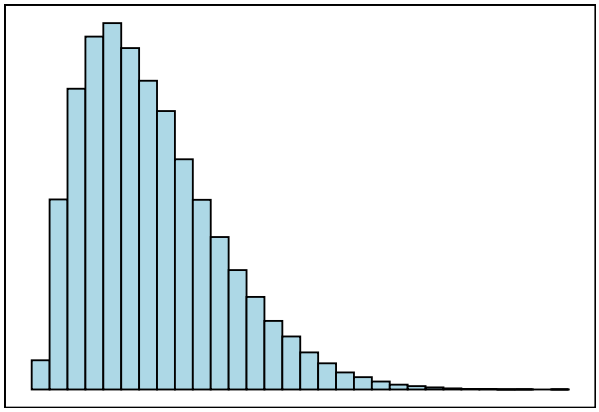
Shape: _____

mean _____ median



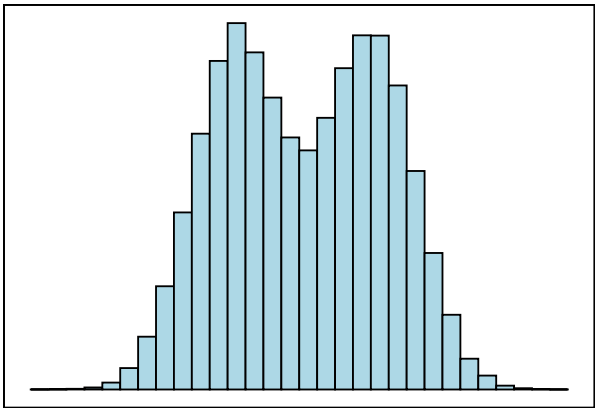
Shape: _____

mean _____ median



Shape: _____

mean _____ median



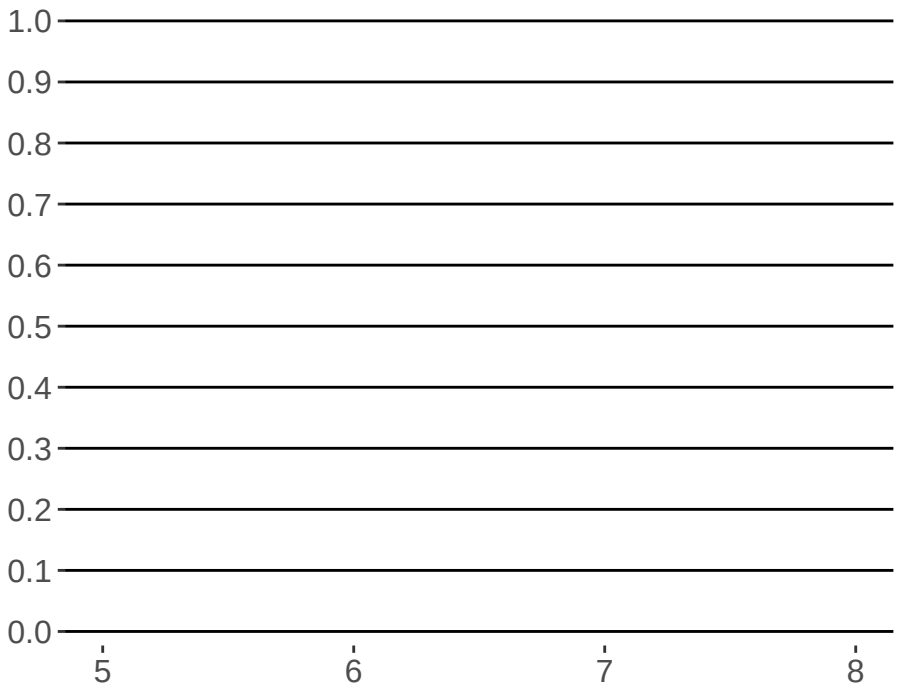
Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

14.) (6 points total) The table below gives the distribution of a quantitative variable X.

Table 1:

X	Frequency(X)	Relative Frequency(x)	Cumulative Relative Fequency
5	8	0.40	0.40
6	5	0.25	0.65
7	2	0.10	0.75
8	5	0.25	1.00

Plot the cumulative distribution (3pts) and use this plot to find the $25th$, $50th$, and $75th$ percentiles of X (1pt each)



$25th =$ _____ $50th =$ _____ $75th =$ _____

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

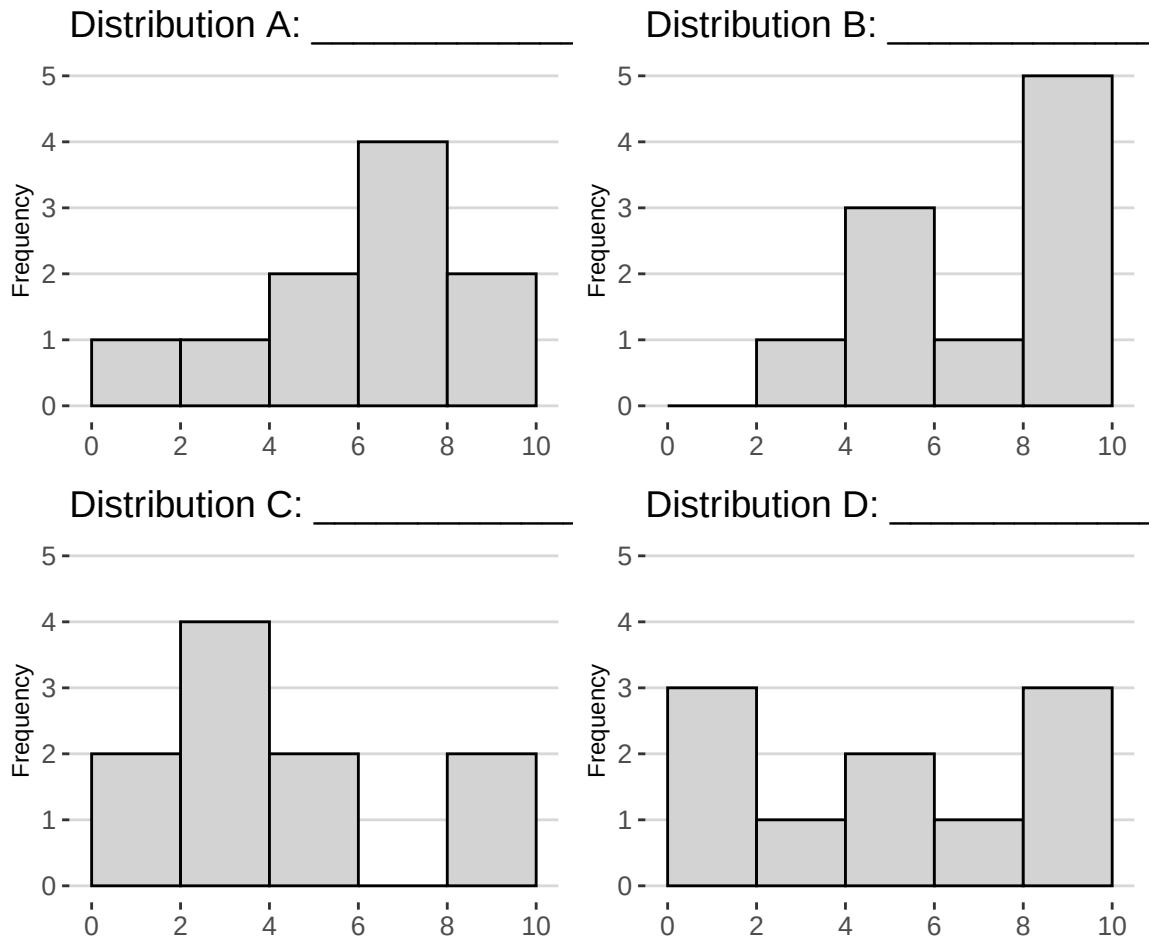
(**15**) (4pts total) Consider the following four sets of observations of a quantitative variable x. For your convenience the observations have been sorted in increasing order. Match datasets $1-4$ with the correct histogram (labeled $A-D$)

$$\text{Dataset } 1 = \{0.1, 1.1, 2.6, 2.7, 3.4, 3.4, 4.1, 4.4, 8.8, 9.6\}$$

$$\text{Dataset } 2 = \{1.1, 3.8, 5.3, 6.0, 6.2, 6.9, 7.9, 7.9, 8.1, 8.7\}$$

$$\text{Dataset } 3 = \{0.1, 0.3, 1.2, 2.4, 4.4, 4.5, 8.0, 8.9, 9.3, 9.3\}$$

$$\text{Dataset } 4 = \{3.4, 4.5, 5.4, 5.6, 7.0, 8.5, 8.9, 9.2, 9.7, 9.7\}$$



Extra Credit: Any points earned on extra credit problems will be applied to the total score of this examination not to exceed the total of **50** points possible.

(bonus) (2pts) Why is the sample variance divided by $n-1$ instead of n like the sample mean? explain you answer (a mathematical demonstration can also help)

(bonus) (2pts) Why is it advisable to complement a boxplot with another type of plot, such as a dot plot or histogram, when performing a descriptive analysis of a variable? Explain the rationale behind not relying solely on a boxplot for a comprehensive understanding of the data.

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!

END OF EXAM

1 Formulas

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\bar{x} = \frac{1}{n}\sum_{x} xF(x)$$

$$\bar{x} = \sum_{x} xRF(X)$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\text{range}(x) = \min(x) - \max(x)$$

$$IQR = Q3 - Q1$$

$$x < Q1 - 1.5 \times (Q3 - Q1)$$

$$x > Q3 - 1.5 \times (Q3 - Q1)$$

Instructions: Carefully read each question. This exam is worth a total of 50 points. For some problems, you may need to use your calculator. You should write any/all computations in the space next to the multiple choice answers for each problem. Good luck!