ST 305: Exam 1

By handing in this completed exam, I state that I have neither given nor received assistance from another person during the exam period. I have used no resources other than the exam itself and the basic mathematical functions of a calculator (ie, no notes, electronic communication, notes stored in calculator memory, etc.) Using your calculator for values from probability distributions like the normal or t is OK; however, if you are doing a calculation from a normal distribution show your work all the way to the point of calculating z-scores. I have not copied from another person's paper. I understand that the penalty if I am found guilty of any such cheating will include failure of the course and a report to the NCSU Office of Student Conduct. **I understand that I must show all work/calculations, even if they seem trivial, to get credit for my answers.**

Name: KEY

ID#: _____

$\bar{x} = \frac{1}{n}\sum x_i$ $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ $Z = \frac{X - \mu}{\sigma}$		

Definitions. (5 points each) Clearly define each of the following terms.

1. Variable:

   A characteristic of an individual

2. Distribution:

   The possible values of a variable, and the frequency of each value.

3. Resistant:

   A resistant measure is (relatively) unaffected by skewness or outliers

Multiple Choice. (5 points each) Select the one best answer.

4. The 5-number summary is most appropriate for use with

A (a.) skewed distributions
   b. small data sets
   c. data sets containing no outliers

5. Stemplots are most appropriate for use with

A (a.) small data sets
   b. categorical variables
   c. data sets with two or more variables per individual

6. The standard normal distribution
   a. has mean equal to 1

B (b.) has standard deviation equal to 1
   c. can have any possible value for the mean and standard deviation

7. When a distribution is skewed to the right, the mean will tend to be
   a. approximately equal to the median

C b. less than the median
   (c.) greater than the median

8. The 68-95-99.7 rule applies to
   a. outliers

B (b.) all normal distributions
   c. all symmetric distributions

For the remaining problems, SHOW YOUR WORK. Numerical answers with no supporting work or explanation will receive zero credit, even if the calculations are trivial.

9. The table below includes the percentage of votes cast for the Democrat candidate for President in the past two elections in 10 southeastern states:

State	% Democrat 2012	% Democrat 2016
Alabama	38.4	34.6
Florida	50.0	47.8
Georgia	45.4	45.6
Kentucky	37.8	32.7
Mississippi	43.5	39.7
North Carolina	48.4	46.7
South Carolina	44.0	40.8
Tennessee	39.0	34.9
Virginia	50.8	49.9
West Virginia	35.5	26.5

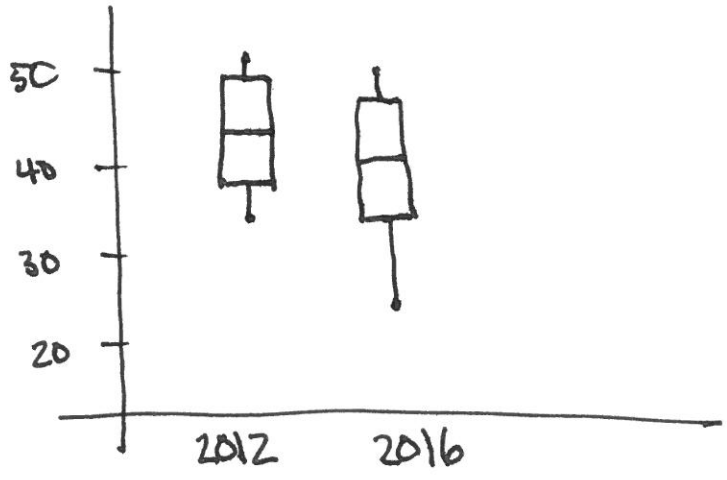
a) Compute the 5-number summaries for each election year, and display them with boxplots on a single graph. (10 points)

2012: 35.5 37.8 38.4 39.0 43.5 44.0 45.4 48.4 50.0 50.8

min = 35.5, Q_1 = 38.4, M = 43.75, Q_3 = 48.4, max = 50.8

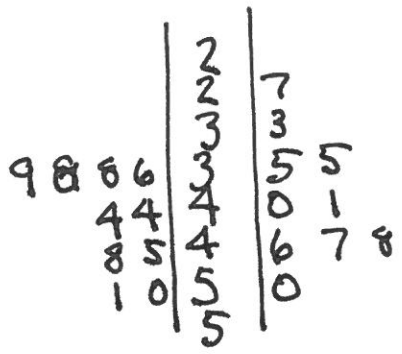
2016: 26.5 32.7 34.6 34.9 39.7 40.8 45.6 46.7 47.8 49.9

min = 26.5, Q_1 = 34.6, M = 40.25, Q_3 = 46.7, max = 49.9



b) Display the data using a back-to-back stemplot. Think carefully about how to process the values to best display these distributions. (10 points)

2012 2016



c) What were the **variables** in the data set above? (5 points)

% Dem 2012 ; % Dem 2016

d) What were the **individuals** in the data set? (5 points)

states

e) Are there outliers in either of the data sets? Justify your answer with a calculation. (10 points)

2012 : IQR = 48.4 − 38.4 = 10 1.5 × IQR = 15

$Q_1 - 1.5 \times IQR = 38.4 - 15 = \underline{23.4}$

$Q_3 + 1.5 \times IQR = 48.4 + 15 = \underline{63.4}$

No values outside this range ⇒ no outliers

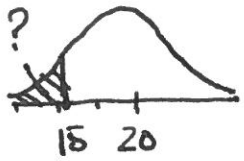
2016 : IQR = 46.7 − 34.6 = 12.1 1.5 × IQR = 18.15

34.6 − 18.15 = 16.45

46.7 + 18.15 = 64.85 ⇒ no outliers

10. The time it takes to complete a patient appointment in a veterinary office follows a normal distribution with mean 20 minutes and standard deviation 3 minutes.

a. What percentage of appointments last less than 15 minutes? (5 points)

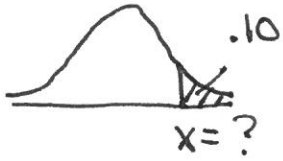


Let X = appointment length $X \sim N(20, 3)$

$$z = \frac{15-20}{3} = -1.67$$

$$L(-1.67) = .0475 \Rightarrow \boxed{4.75\%}$$

b. The office owner would like to identify the longest 10 percent of appointments. What is the cutoff value for being in the 10% of appointments with the longest times? (5 points)



$R(z) = .10 \Rightarrow z = 1.28$

$$z = \frac{x-\mu}{\sigma} \Rightarrow 1.28 = \frac{x-20}{3}$$

solve for x: $x = \boxed{23.84 \text{ minutes}}$

10. We discussed two different quantities for describing the center of a distribution.

a. What were the two quantities? (just name them) (5 points)

mean, median

b. We discussed some of the good and bad properties for these quantities. For each of the two quantities, name 1 good property and 1 bad property. (5 points)

	Good	Bad
mean	Fast to compute	Not resistant to outliers
median	Slow to compute	Resistant

swap