Student name:

MATH 120: Elementary Statistics
Section 1

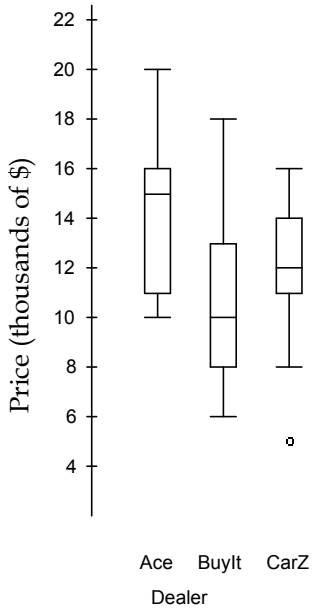
Test 1
February 21, 2019

Instructions:

- This is a regular "closed-book" test, and is to be taken without the use of notes, books, or other reference materials. Collaboration or group work is not permitted.
- Cell-phone usage of any kind is prohibited for the entire duration of the test. This also applies to any restroom breaks taken during the test.
- The time limit for taking this test is 80 minutes from the scheduled start time. Please turn in your test promptly when time is called to avoid late penalties.
- This test adds up to 50 points.

---

1. [4 pts] The boxplots show prices of used cars, in thousands of dollars, advertised for sale at three different car dealers. Write a few sentences comparing prices at these three dealerships. Be sure to include the context and units in your discussion.



2. [4 pts] In their October 2003 issue, *Consumer Reports* evaluated the price and performance of 23 models of cell phones. Computer output gives these summaries for the prices:

Min	Q1	Median	Q3	Max	Mean	SD
15	30	50	110	200	71.75	52.08

(a) Were any of the prices outliers? Justify your claim.

(b) What statistics would best summarize the center and spread of these data?

3. [4 pts] According to the U.S. Bureau of Labor Statistics, in 2009 the mean hourly wage for Chief Executives was \$93 and the median was \$78.28. Is this wage distribution likely to be symmetric, skewed left, or skewed right? Show a plausible sketch (or histogram profile) for the distribution.

4. [4 pts] The table below compares the daily weather forecast in a city with the actual weather for 1 year

		Forecast		Total
		rain	**no rain**	
Actual	**rain**	102	31	133
	no rain	46	186	232
	Total	148	217	365

(a) For what percent of days was the forecast correct?
(b) On what percent of days did it actually rain?

5. [4 pts.] In a psychology experiment, researchers studied the association between social media usage and academic performance among college students. Social media usage was measured in average number of hours per week, and academic performance was measured by GPA. The association was roughly linear, with a correlation of $r = -0.2$. The researchers decided to construct a linear regression model to predict GPA from social media usage.
(a) Identify the explanatory variable and the response variable, including their units.
(b) Suppose a student's social media usage is 2 standard deviations higher than the mean usage. What does this model predict about the student's GPA?

6. [10 pts] Now that cigarette smoking has been clearly tied to lung cancer, researchers are exploring possible links to other diseases. An article in the *American Journal of Public health* gives data on smoking rates and coronary heart disease (CHD) in 21 countries. The mean cigarette consumption in these countries was 2148 cigarettes per adult per year, with a standard deviation of 809 cigarettes per adult per year. The mean CHD rate was 144.9 deaths per 100,000 citizens, with standard deviation of 66.5 deaths per 100,000. The association between cigarette consumption and CHD rates was found to be approximately linear, positive, and with correlation $r = 0.73$.

(a) Construct a linear regression model to predict CHD death rates from cigarette consumption rate.

(b) Interpret its slope in this particular application context.

(c) Find R^2 and interpret its meaning.

7. [10 pts] A company's manufacturing process uses 500 gallons of water at a time, which is later treated to remove a chemical pollutant, before the water is released into a nearby lake. Local laws allow a maximum of 80 ppm (parts per million) of the chemical in the treated water. The output from the company's treatment process follows a normal model, with mean pollutant concentration of 75 ppm and standard deviation 4.2 ppm.
(a) Sketch and label the normal model, showing key numerical values and percentages of the distribution of data.
(b) In what percent of the cases does the treated water exceed the legal maximum limit of 80 ppm?
(c) Management wants to ensure that the 80 ppm limit is only exceeded 1% of the time. What standard deviation must the treatment process have in order to achieve this goal?

8. [10 pts] The highest temperature ever recorded (in °F) in 32 different states in the U.S. is shown below, in ascending order:

98	102	104	106	109	110	112	114
114	115	115	116	116	118	118	118
120	120	120	120	121	121	122	122
123	124	125	126	126	128	129	134

(a) Make a suitably scaled, neat, histogram for these data. Show detailed steps.
(b) Compute suitable summary statistics for the center and spread of the distribution. Feel free to use your calculator short-cut keys for this part.
(c) If we were to convert all temperatures to °C, what would the new values of your center and spread statistics be? [FYI: $°C = \frac{5}{9}(°F - 32)$]