

**AP Statistics Summer Assignment**

**Name:** \_\_\_\_\_

1. The graph below displays the scores of 32 students on a recent exam. Scores on this exam ranged from 64 to 95 points.



a. What does the graph tell you about how the grades are distributed?

b. Explain how you would estimate the mean and the median? Use your method to estimate the mean and the median.

c. Based on (a) and (b), which summary statistics, the mean or the median should the teacher use to report the overall performance of the students on the exam? Justify your answer.

**Each question is worth 10 points**

2. The Better Business Council of a large city has concluded that students in the city's schools are not learning enough about economics to function in the modern world. These findings were based on test results from a random sample of 20 twelfth-grade students who completed a 46 question multiple-choice test on the basic economics concepts. The data set below shows the number of questions that each of the 20 students in the sample answered correctly.

12 16 18 17 18 33 41 44 38 35 19 36 19 13 43 8 16 10 9 13

a. Display these data in a stemplot and a boxplot.

**Each question is worth 10 points**

b. Explain what stands out in the stemplot that is not visible in the boxplot.

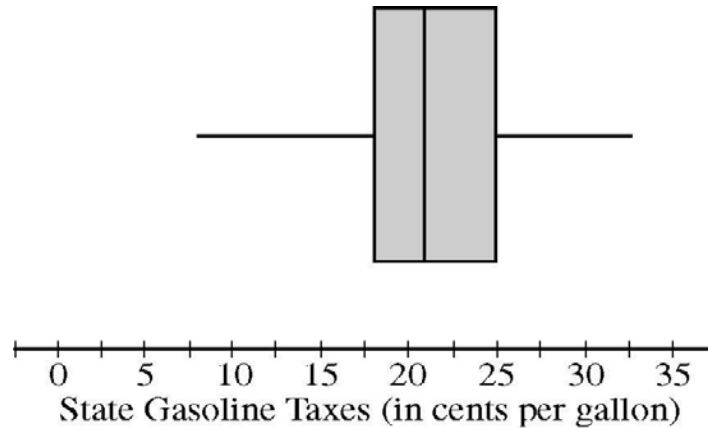
c. Explain what stands out in the boxplot that is not in the stemplot.

d. What does the graph tell you about the distribution of the correctly answered questions for the 20 students?

e. Why would it be misleading to report only a measure of center for this score distribution?

Each question is worth 10 points

3. As gasoline prices have increased in recent years, many drivers have expressed concern about the taxes they pay on gasoline for their cars. In the United States, gasoline taxes are imposed by both the federal government and by individual states. The boxplot below shows the distribution of the state gasoline taxes, in cents per gallon, for all 50 states on January 1, 2006.



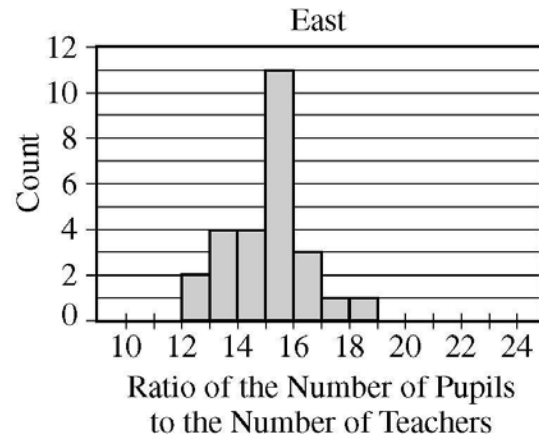
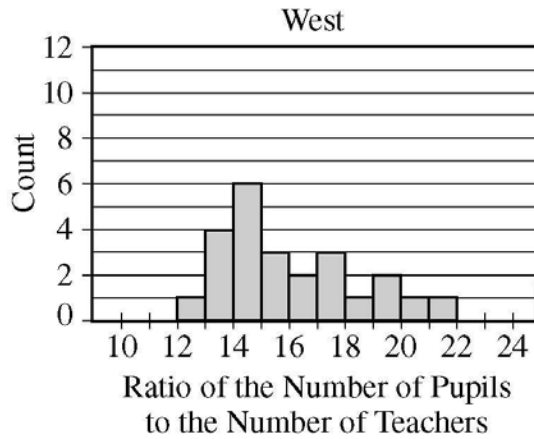
a. Based on the boxplot, what are the approximate values of the median and the interquartile range of the distribution of state gasoline taxes, in cents per gallon? Mark and label the boxplot to indicate how you found the approximated values.

b. List the possible gasoline taxes for each state that would produce the boxplot above. Explain why the data that you listed would coincide with the box plot.

c. The federal tax imposed on gasoline was 18.4 cents per gallon at the time the state taxes were in effect. The federal tax was added the state gasoline tax for each state to create a new distribution of combined gasoline taxes. What are the approximate values in cents per gallon, of the mean, median, standard deviation and the interquartile range of the new distribution of combined gasoline taxes? Justify your answer.

Each question is worth 10 points

4. Records are kept by each state in the United States on the number of pupils enrolled in public schools and the number of teachers employed by public schools for each school year. From these records, the ratio of the number of pupils to the number of teachers (P-T ratio) can be calculated for each state. The histograms below show the P-T ratio for every state during the 2001–2002 school year. The histogram on the left displays the ratios for the 24 states that are west of the Mississippi River, and the histogram on the right displays the ratios for the 26 states that are east of the Mississippi River.



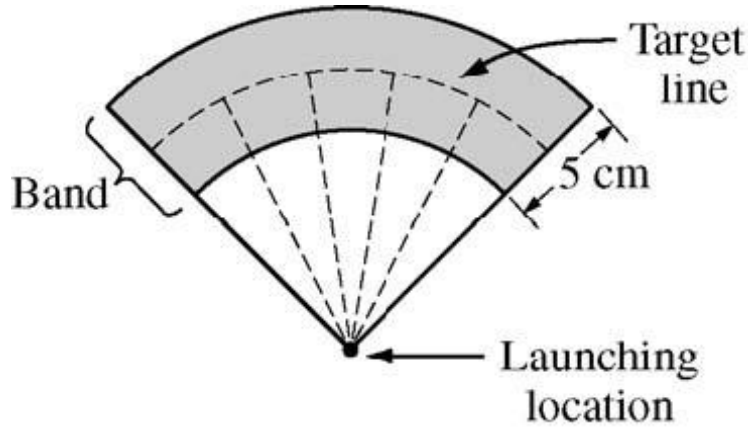
a. Describe how you would use the histograms to estimate the median P-T ratio for each group (west and east) of states. Then use this procedure to estimate the median of the west group and the median of the east group.

b. Write a few sentences comparing the distributions of P-T ratios for states in the two groups (west and east) during the 2001–2002 school year. In clued in your discuss the differences or similarities in the medians, the ranges and whether the shape of the distributions.

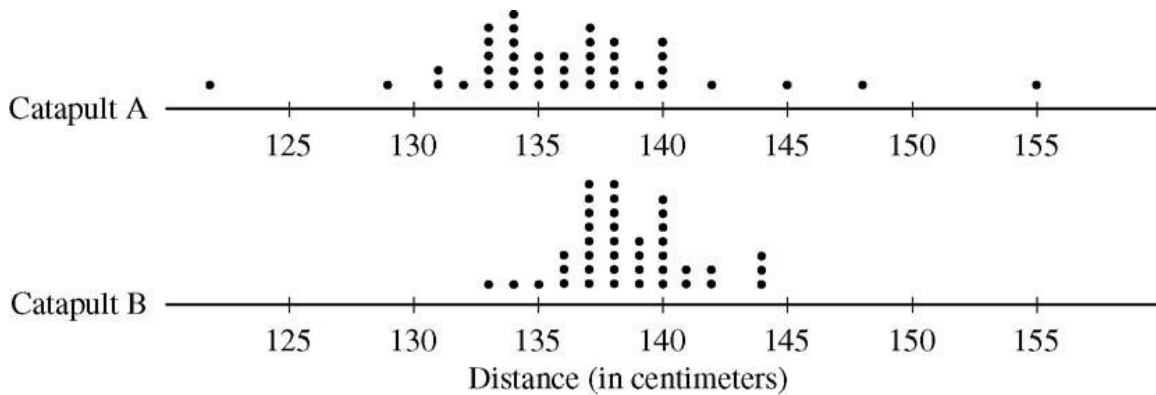
c. Using your answers in parts (a) and (b), Explain which distribution would have the larger mean.

Each question is worth 10 points

5. Two parents have each built a toy catapult for use in a game at an elementary school fair. To play the game, students will attempt to launch Ping-Pong balls from the catapults so that the balls land within a 5-centimeter band. A target line will be drawn through the middle of the band, as shown in the figure below. All points on the target line are equidistant from the launching location.



If a ball lands within the shaded band, the student will win a prize. The parents have constructed the two catapults according to slightly different plans. They want to test these catapults before building additional ones. Under identical conditions, the parents launch 40 Ping-Pong balls from each catapult and measure the distance that the ball travels before landing. Distances to the nearest centimeter are graphed in the dotplots below.



**Each question is worth 10 points**

a. Comment on any similarities and any differences in the two distributions of distances traveled by balls launched from catapult A and catapult B.

(b) If the parents want to maximize the probability of having the Ping-Pong balls land within the band, which one of the two catapults, A or B, would be better to use than the other? Justify your choice.

(c) Using the catapult that you chose in part (b), how many centimeters from the target line should this catapult be placed? Explain why you chose this distance.

**Each question is worth 10 points**

6. Select at random 10 senators from the list below. (See list of US Senators at the end of the questions.)

i. Explain the method you used to select the senators that you chose randomly.

ii. Record the years of service for the senators that you selected in the table below.

<b>ID#</b>										
<b>Years of service</b>										

iii. Construct a dotplot for the years of service of the 10 senators that you selected.

iv. Use the distribution of the dotplot along with the mean, the median, the standard deviation and the interquartile range to draw conclusions about the number of years of service of US Senators in 1994.



Each question is worth 10 points

US Senators (1994)					
ID#	Name	Sex	Party	State	Years
01	Akaka	M	Dem	Hawaii	4
02	Baucus	M	Dem	Montana	16
03	Bennett	M	Rep	Utah	1
04	Biden	M	Dem	Delaware	21
05	Bingaman	M	Dem	New Mexico	11
06	Bond	M	Rep	Missouri	7
07	Boren	M	Dem	Oklahoma	15
08	Boxer	F	Dem	California	1
09	Bradley	M	Dem	New Jersey	15
10	Breaux	M	Dem	Louisiana	7
11	Brown	M	Rep	Colorado	3
12	Bryan	M	Dem	Nevada	5
13	Bumpers	M	Dem	Arkansas	19
14	Burns	M	Rep	Montana	5
15	Byrd	M	Dem	West Virginia	35
16	Campbell	M	Dem	Colorado	1
17	Chafee	M	Rep	Rhode Island	18
18	Coats	M	Rep	Indiana	5
19	Cochran	M	Rep	Mississippi	16
16	Cohen	M	Rep	Maine	15
21	Conrad	M	Dem	North Dakota	7
22	Coverdell	M	Rep	Georgia	1
23	Craig	M	Rep	Idaho	3
24	D'Amato	M	Rep	New York	13
23	Danforth	M	Rep	Missouri	18
26	Daschle	M	Dem	South Dakota	7
27	DeConcini	M	Dem	Arizona	17
28	Dodd	M	Dem	Connecticut	13
29	Dole	M	Rep	Kansas	25
30	Domenici	M	Rep	New Mexico	21
31	Dorgan	M	Dem	North Dakota	1
32	Durenburger	M	Rep	Minnesota	16
33	Exon	M	Dem	Nebraska	15
34	Faircloth	M	Rep	North Carolina	1
35	Feingold	M	Dem	Wisconsin	1

Each question is worth 10 points

US Senators (1994)					
ID#	Name	Sex	Party	State	Years
36	Feinstein	F	Dem	California	1
37	Ford	M	Dem	Kentucky	20
38	Glenn	M	Dem	Ohio	20
39	Gorton	M	Rep	Washington	13
40	Graham	M	Dem	Florida	7
41	Gramm	M	Rep	Texas	9
42	Grassley	M	Rep	Iowa	13
43	Gregg	M	Rep	New Hampshire	1
44	Harkin	M	Dem	Iowa	9
45	Hatch	M	Rep	Utah	17
46	Hatfield	M	Rep	Oregon	27
47	Heflin	M	Dem	Alabama	15
48	Helms	M	Rep	North Carolina	21
49	Hollings	M	Dem	South Carolina	28
50	Hutchinson	F	Rep	Texas	1
51	Inouye	M	Dem	Hawaii	31
52	Jeffords	M	Rep	Vermont	5
53	Johnston	M	Dem	Louisiana	22
54	Kassebaum	F	Rep	Kansas	16
55	Kempthorne	M	Rep	Idaho	1
56	Kennedy	M	Dem	Massachusetts	32
57	Kerry, J	M	Dem	Massachusetts	9
58	Kerry, R	M	Dem	Nebraska	5
59	Kohl	M	Dem	Wisconsin	5
60	Lautenberg	M	Dem	New Jersey	12
61	Leahy	M	Dem	Vermont	19
62	Levin	M	Dem	Michigan	15
63	Lieberman	M	Dem	Connecticut	5
64	Lott	M	Rep	Mississippi	5
65	Lugar	M	Rep	Indiana	17
66	Mack	M	Rep	Florida	5
67	Matthew	M	Dem	Tennessee	1
68	McCain	M	Rep	Arizona	7
69	McConnell	M	Rep	Kentucky	9
70	Metzenbaum	M	Dem	Ohio	18

Each question is worth 10 points

US Senators (1994)					
ID#	Name	Sex	Party	State	Years
71	Mikulski	F	Dem	Maryland	7
72	Mitchell	M	Dem	Maine	14
73	Moseley-Baum	F	Dem	Illinois	1
74	Moynihan	M	Dem	New York	17
75	Murkowski	M	Rep	Alaska	13
76	Murray	F	Dem	Washington	1
77	Nickles	M	Rep	Oklahoma	13
78	Nunn	M	Dem	Georgia	22
79	Packwood	M	Rep	Oregon	25
80	Pell	M	Dem	Rhode Island	33
81	Pressler	M	Rep	South Dakota	15
82	Pryor	M	Dem	Arkansas	15
83	Reid	M	Dem	Nevada	7
84	Riegle	M	Dem	Michigan	18
85	Robb	M	Dem	Virginia	5
86	Rockefeller	M	Dem	West Virginia	9
87	Roth	M	Rep	Delaware	23
88	Sarbanes	M	Dem	Maryland	17
89	Sasser	M	Dem	Tennessee	17
90	Shelby	M	Dem	Alabama	7
91	Simon	M	Dem	Illinois	9
92	Simpson	M	Rep	Wyoming	15
93	Smith	M	Rep	New Hampshire	3
94	Specter	M	Rep	Pennsylvania	13
95	Stevens	M	Rep	Alaska	26
96	Thurmond	M	Rep	South Carolina	38
97	Wallop	M	Rep	Wyoming	17
98	Warner	M	Rep	Virginia	15
99	Wellstone	M	Dem	Minnesota	16
100	Wofford	M	Dem	Pennsylvania	3

### **Course Description**

In AP Statistics, we will follow four conceptual themes: Exploring Data, Planning a Study, Anticipating Patterns, and Statistical Inferences. This involves deriving methods to collect, summarize, analyze and interpret numerical information that relates to the real world. Probability is applied along with these statistical methods to observe consistent behavior in data and to develop rules such as the Central Limit Theorem. This allows us to calculate Confidence Intervals to approximate results from data and to perform Hypothesis Testing and Chi-Square Testing to check the validity of results from data. These statistical tools are used to make inferences and draw conclusions about data.

My goal as your teacher is to provide a clear understanding of how statistics works in the real world and how it affects our thinking and our ability to solve problems.

### **Essential Questions:**

1. How do we use numbers to provide information and to make decisions?
2. How do we use numbers to evaluate decisions that affect us?

### **Course Requirements**

- Two three ring binder (one for notes and one for homework)
- Pens, Pencils, a Ruler and a Stapler
- TI-84+ graphing calculator

You are expected to come to class prepared. Failure to do so will affect your grade. All assignments should be neatly stapled before they are submitted.

### **Criteria for Success**

- i. **Homework**  
Homework serves two main purposes. It allows you to practice what was taught in class and to determine whether you understand the material. You will be selected to present homework assignments in class. I will collect and grade your homework book and your notebook. You will also get pop quizzes based on the homework. Failure to do the homework will affect your grade. Homework questions will be reviewed in class. If you do not understand a homework question you are required to send me an email with questions about the assignment. You can also see me in the morning for extra help.
- ii. **Homework of the Week (HOW)**  
You will be given a weekly homework assignment (**HOW**) in addition to your regular homework assignments. The questions on the HOWs will be taken from past AP Statistics exams.

**iii. Test/Project**

Tests and Quizzes are given on a regular basis. Review sessions for test will be scheduled before or after school. There will also be pop quizzes based on the homework. Projects are assigned to provide deeper insights into the topics that are covered and to develop problem-solving skills.

**iv. Notes**

Good note-taking is essential to success. In addition to class notes, you will be given reading assignments from the textbook. These reading assignments will be collected and graded.

**v. Participation/Presentation**

Participation and presentations in class is also critical to success. You are expected to share your ideas and pay attention when others are speaking. This is an integral part of class discussions and group work. You are also expected to present problems on the board. Your grades for presentation and participation reflect your contribution to the learning process.

**Grading Policy**

Test/Quiz/Project	60 pts
Homework	15 pts
Notes	5 pts
Participation (Includes working in pairs and in groups)	15 pts
Presentation	<u>5 pts</u>
	100 pts

**vi. Reflection**

Your thoughts about what we do in class allow us to create and maintain an environment that is conducive to learning.

Please sign below confirming that you understand and are in agreement with everything that is stated above.

**Student Name (Print):**\_\_\_\_\_ **Signature:**\_\_\_\_\_

**Parent Name (Print):**\_\_\_\_\_ **Signature:**\_\_\_\_\_

## **COURSE OUTLINE**

**Primary Textbook:** Peck, Roxy , Chris Olsen, Jay Devore. *Introduction to Statistics and Data Analysis*, 3<sup>rd</sup> Edition. California: Brooks/Cole, 2001

**Supplementary Textbook:** Sheaffer, Richard, Mrudulla Gnanadesikan, Ann Watkins, Jeffrey Witmer. *Activity Based Statistics*. New York: Springer-Verlag 1996

Rossman, Allan, J. Barr Von Oehsen. *Workshop Statistics*. New York: Springer-Verlag 1997

Roxy Peck and Daren Starnes. *Making Sense of Statistical Studies*. Virginia: American Statistical Association, 2009

Daniel S. Yates, David S. Moore and Daren S. Starnes. *The Practice of Statistics*, 3<sup>rd</sup> Edition. New York: W. H. Freeman Co., 2008

David E. Bock, Paul F. Velleman, Richard D. De Veaux. *Stats: Modeling the World AP Edition 3<sup>rd</sup> Edition*. Boston: Addison-Wesley, 2004

**Review Book:** Anne Carrol, Ruth Carver, Susan Peters, Janice Ricks. *AP Statistics for Stats: Modeling the World 3<sup>rd</sup> Edition*. Boston: Addison-Wesley, 2010

## **Theme I: Numerical Data**

### **I. Planning a Survey Study for Numerical Data (1 weeks)**

- i. Numerical Data: Discrete and Continuous
- ii. Sample and Population
- iii. Good sampling methods (SRS, stratified sampling etc)
- iv. Bad sampling methods and biases (voluntary sampling, etc)
- v. Types of Data: Numerical and Categorical

### **II. Exploring Numerical Data (2 weeks)**

- i. Numerical data: Discrete and Continuous
- ii. Representing and Displaying Numerical data
- iii. Measure of Central Tendency
- iv. Measure of Variability
- v. Normal Distribution

### **III. Anticipating Patterns in Numerical Data (1 weeks)**

- i. Mean and Standard Deviation: Statistic and Parameter
- ii. Sampling Distribution (The German Tank Project)
- iii. Sampling Distribution of Means

### **IV. Making Inferences About Numerical Data (2 weeks)**

- i. Confidence Interval for one Mean
- ii. Hypothesis Test for one Mean

## **Theme II: Categorical Data**

### **I. Planning a Study for Categorical Data (1 week)**

- i. Designing Questions
- ii. Reviewing Biases
- iii. Applying Stratified Sampling

### **II. Exploring Categorical Data (1 week)**

- i. Representing and Displaying Categorical Data
- ii. Proportion: Statistic and Parameter

### **III. Anticipating Pattern for Categorical Data (2 weeks)**

- i. Probability: Basic Rules
- ii. Independence and Mutually Exclusive Events
- iii. Simulation
- iv. Probability Distribution
- v. Binomial and Geometric Probability Distribution
- vi. Combining Distributions

### **IV. Making Inferences about Categorical Data (2 weeks)**

- i. Normal Approximation of Binomial Probability
- ii. Sampling Distribution of Proportions
- iii. Confidence Interval for One Proportion
- iv. Hypothesis Test for One Proportion
- v. Chi Square Test for Goodness Fit

### **Theme III: Bivariate Numerical Data**

#### **I. Planning an Observational Study for Association (Bivariate Numerical Data) ( 1 week)**

- i. Defining Observational Studies for Bivariate Numerical Data
- ii. Lurking Variables

#### **II. Exploring Bivariate Numerical Data (2 weeks)**

- i. Displaying Bivariate Numerical Data
- ii. Correlation Coefficient and The Regressions Line
- v. Residuals and Residual Plots
- vi. Outliers and Influential Data
- vii. Coefficient of determination and Standard Error
- viii. Linearizing Regression

#### **III. Anticipating Patterns for Bivariate Numerical Data (1 week)**

- i. Standard deviation of Slope
- ii. Sampling Distribution of Slopes

#### **IV. Making Inferences About Bivariate Numerical Data ( 1 week)**

- i. Confidence Interval of Slopes
- ii. Sampling Distribution of Slopes

### **Theme IV: Bivariate Categorical Data**

#### **Planning an Observational Study (Bivariate Categorical Data) (1 week)**

- i. Defining Observational Study for Bivariate Categorical Data
- ii. Lurking Variable

#### **II. Exploring Bivariate Categorical Data (1 week)**

Displaying Bivariate Categorical Data

#### **III. Anticipating Patterns for Bivariate Categorical Data (1 week)**

- i. Independence
- ii. Review Basic Rule of Probability

#### **IV. Making Inferences (1 week)**

Chi Square Test for Association or Independence



## **Theme V: Comparative Observational Studies and Experiments for Numerical Data**

### **Planning a Comparative Observational Study and an Experiment for Numerical Data (2 weeks)**

- i. Comparative Observation Study and Experiment
- ii. Treatment, Control Groups, Placebo and Placebo Effect
- iii. Randomization, Replication, Control and Blocking
- iv. Complete Randomized Design
- v. Blocked and Matched Pairs Designs

### **II. Exploring two or more Numerical Data sets (1 week)**

- i. Comparing Numerical Data sets Graphically
- ii. Comparing Numerical Data sets with numbers

### **III. Anticipating Patterns for Numerical Data Sets (1 week)**

- i. Review Combining Distributions
- ii. Sampling Distribution for Difference of two means

### **IV. Making Inferences about Numerical Data Sets (1 weeks)**

- i. Confidence Interval for two means
- ii. Hypothesis Test for two means

## **Theme VI: Comparative Studies and Experiments for Categorical Data**

### **I. Planning a Comparative Observation Study and an Experiment for Categorical Data (1 week)**

Review the Four Principle of Experimental Design (Randomization, Replication, Control and Blocking)

### **II. Exploring two or more Categorical Data Sets (1 week)**

- i. Comparing Categorical Data sets Graphically
- ii. Comparing Categorical Data sets Numerically

### **III. Anticipating Patterns for Categorical Data Sets (1 week)**

Review Independence

### **IV. Making Inferences about Categorical Data Sets (1 weeks)**

- i. Confidence Interval for two Proportions
- ii. Hypothesis Test for two Proportions
- iii. Chi Square Test for Homogeneity