

AP Statistics Summer Assignment – 2018-2019

Mr. Krar
Marjory Stoneman Douglas High School
Email: frank.krar@browardschools.com
Follow me on Twitter @FrankKrar

May 2018

To all students enrolled in AP Statistics and their parents/guardians,

Welcome to Advanced Placement Statistics. This course is like no other mathematics course offered at MSD in that the emphasis is placed on your ability to **think, reason, explain,** and **support** your answers as opposed to simply performing computations. Don't like word problems? We do *nothing but* word problems!

Students – seniors in particular – sometimes sign up to take AP Statistics thinking they are somehow “escaping” so-called “difficult” math (such as Pre-Calculus or Calculus) and will have it easy taking this class. Let me assure you, AP Statistics is **not** an easy class. No Advanced Placement class is easy. Like in any class, you can expect to spend time completing assignments (reading, watching videos, completing problems, projects, etc.) outside of class, as well as in class. However, AP Statistics is special; it is a course that combines both mathematical and verbal skills. On the AP exam, you will be asked to write descriptive paragraphs and concluding sentences. You will have to explain the reasoning behind the methods you use and the conclusions you draw.

Throughout the course of the year, AP Statistics will expose you to four themes: 1) Exploring Data; 2) Sampling and Experimentation; 3) Anticipating Patterns and Probability; and 4) Statistical Inference. Since statistics helps us describe and understand the world around us, the knowledge you gain in this class will be useful as you move into your future careers.

How can you prepare for this class?

- 1) Complete the attached summer packet by the first day of class in August. It will be graded **for correctness** and will count as your **first quiz grade!** I will accept it 1 day late for a 10-point deduction (out of 100). If you turn it in on the third day of class, I will accept it for a 25-point deduction. After that, I will not accept the assignment and you will receive a zero.
- 2) Having your own graphing calculator is required and you will need to bring it to class **every day**. I HIGHLY recommend getting the TI-84 Plus CE, but any Texas Instruments calculator in the TI-84 family will work fine. I don't recommend using the TI-83, as there are functions it cannot do that will limit its usefulness in certain topics. Our textbook includes “TI-Tips” throughout the book; they show you how to use the TI-84 calculators specifically.
- 3) Lastly, you will need to be an active participant in the course. This means you **MUST** be willing to work with your fellow classmates and me often during the year. If you are the type of student that does not want to work, would rather sit and do nothing during class time, does not enjoy the mental challenge of a good question, or are looking for a class where you can get an “easy A” without much effort, then this is probably not the course for you.

If you've read this far and are still interested in the course (i.e., I haven't scared you off yet), then I bid you welcome. Attached is your summer assignment. Be sure to read the directions **thoroughly**. I look forward to working with and teaching each and every one of you over the course of the coming school year.

Enjoy your summer and see you at the start of school!

AP Statistics Summer Assignment – 2018-2019

Welcome to AP Statistics, future statisticians! The purpose of this assignment is to make you familiar with the concepts of data analysis and to be able to hit the ground running on the first day of school.

The summer assignment is composed of two parts:

- 1) Reading and Vocabulary:** You will use a free online statistical tutoring site that will give you information on variables and data displays. While reviewing the information on the site, you will be completing a vocabulary list (see pages 3 – 6). Follow the steps below:
- Go to www.stattrek.com
 - Click on “AP Statistics” then on “AP Tutorial”
 - On the left side of the screen is a list of general topics. Under each general topic is a list of subtopics. You will explore the following subtopics to complete the vocabulary list. **I will expect you to be familiar with ALL of these terms on the first day of school.**

EXPLORING DATA	
The Basics	Variables Population vs. Sample Central Tendency Variability Position
Charts and Graphs	Patterns in Data Dotplots Histograms Stemplots Boxplots Scatterplots Comparing Data Sets

- 2) Practice Problems:** After reading all the material above you should be able to complete the questions in the remaining pages of this packet. You should do so in the spaces provided.

Remember, a graphing calculator is required for this course. I recommend getting your calculator before beginning this summer assignment. Again, I recommend the TI-84 Plus CE. Using your calculator to help you complete this assignment – even if you don’t know how to use the “fancy” functions – will go a long way to making you more comfortable with some of its statistical features.

Lastly, let me again remind you that this summer assignment is **due on the first day of school**. It will be graded **for correctness** and will count as your **first quiz grade!**

If you have any questions or need clarification, please email me at frank.krar@browardschools.com, or send me a message on Twitter @FrankKrar. I will get back to you as soon as I can.

Name _____

PART 1 – VOCABULARY: Define each of the following terms from the information on the stattrek website. Words in **BLUE on the website can be clicked on for more information. When asked to provide an example or a sketch of the word, please provide a *unique example NOT given on the website*.**

1. Categorical Variable

Example:

2. Quantitative Variable

Example:

3. Discrete Variable

4. Continuous Variable

5. Univariate Data

6. Bivariate Data

7. Population

Example:

8. Sample

Example:

9. Median

10. Mean

Formula:

11. Outlier

12. Parameter

13. Statistic

14. Range

15. Standardized Score (z-score)

Formula

16. Center

17. Spread

18. Variance

Formula:

19. Standard Deviation

Formula:

20. Symmetry

Sketch:

21. Unimodal

Sketch:

22. Bimodal

Sketch:

23. Skewness

Sketch Skewed Left:

Sketch Skewed Right:

24. Uniform

Sketch:

25. Gaps

Sketch:

26. Outliers

Sketch:

27. Dot Plots

28. Bar Chart

29. Histogram

30. Difference between a Bar Chart and a Histogram

31. Stemplots

32. Boxplots

33. Quartiles

34. Range

35. Interquartile Range

36. Four Ways to Describe Data Sets

37. Types of Graphs that can be used for Comparing Data

PART 2 – PRACTICE PROBLEMS

A.) CATEGORICAL or QUANTITATIVE - Determine if the variables listed below are *quantitative (Q)* or *categorical (C)*.

1. Time it takes to get to school
2. Number of people under the age of 18 living in a household
3. Hair color
4. Temperature of a cup of coffee
5. Teacher salaries
6. Gender
7. Whether someone is a smoker or not
8. Height
9. Amount of oil spilled
10. Age of Academy Award winners
11. Type of depression medication
12. Jellybean flavors
13. Social Security Number
14. Type of meat
15. Number of pairs of shoes owned

B.) STATISTICS – WHAT IS THAT?

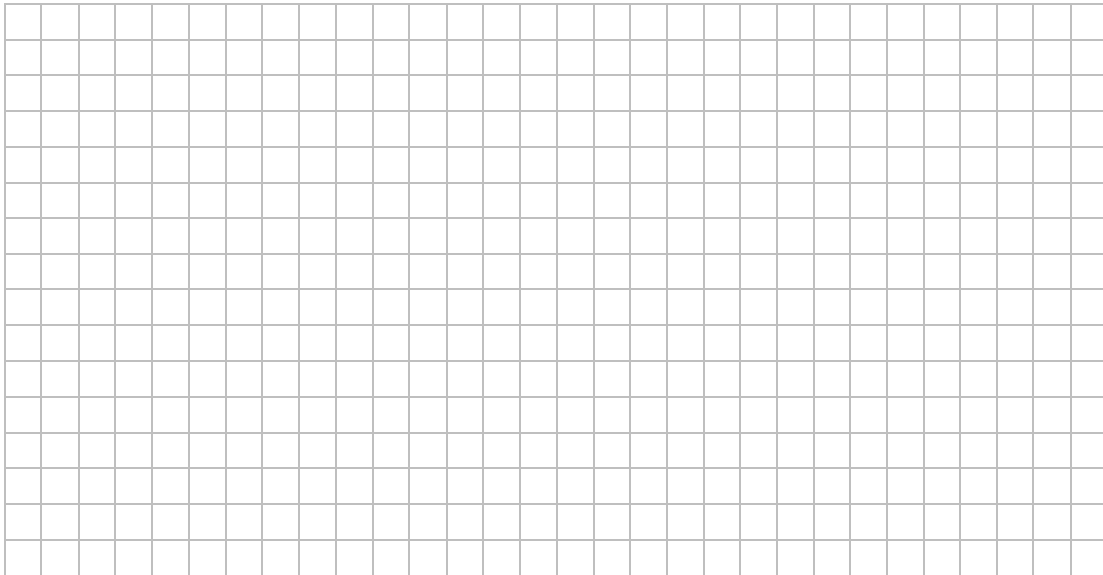
A statistic is a number calculated from data. Quantitative data have many different statistics that can be calculated. Determine the given statistics from the data below on the number of home runs Mark McGuire hit in each season from 1982 – 2001.

70	52	22	49	3	32	58	39
39	65	42	29	9	32	9	33

Mean	
Minimum	
Q1	
Median	
Q3	
Maximum	
Range	
IQR	

C.) ACCIDENTAL DEATHS - In 1997 there were 92,353 deaths from accidents in the United States. Among these were 42,340 deaths from motor vehicle accidents, 11,858 from falls, 10,163 from poisoning, 4051 from drowning, and 3601 from fires. The rest were listed as “other causes.”

1. Find the percent of accidental deaths from each of these causes, rounded to the nearest percent.
2. What percent of accidental deaths were from “other causes?”
3. NEATLY create a well-labeled **bar graph** of the distribution of causes of accidental deaths. Be sure to include an “other causes” bar.

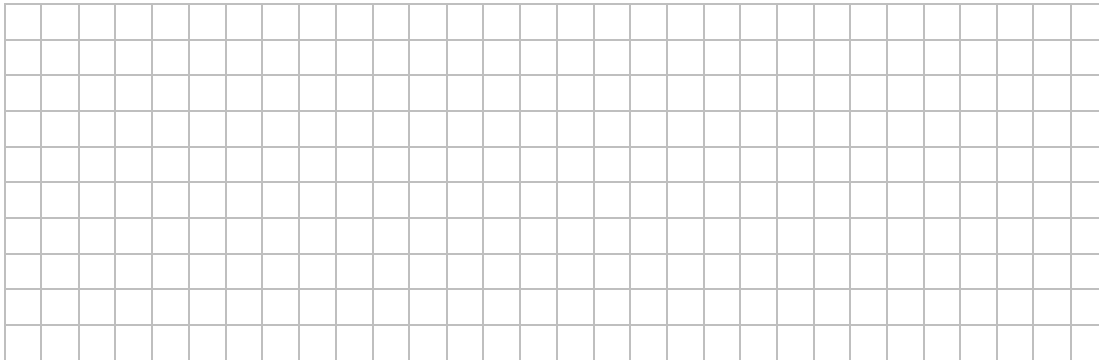


4. A pie chart is another graphical display used to show all the categories in a categorical variable relative to each other. Create a pie chart for the accidental death percentages. You may try using software or Internet source to make one and paste in the space below (*Microsoft Excel works well*).

D.) HURRICANE'S A'COMIN' – The data below give the number of Atlantic hurricanes that made landfall each year from 1944 through 2000 as reported by *Science* magazine.

3	2	1	4	3	7	2	3	3	2	5	2	2	4
2	2	6	0	2	5	1	3	1	0	3	2	1	0
1	2	3	2	1	2	2	2	3	1	1	1	3	0
1	3	2	1	2	1	1	0	5	6	1	3	5	3

1. Make a horizontal dotplot to display these data. Make sure you include appropriate labels, title, and scale. The graph paper below should help ensure you space your markings (you may use X's or dots) consistently.



E.) SHOPPING SPREE! – A marketing consultant observed 50 consecutive shoppers at a supermarket. One variable of interest was how much each shopper spent in the store. Here are the data – rounded to the nearest dollar – arranged in increasing order.

3	9	9	11	13	14	15	16	17	17
18	18	19	20	20	20	21	22	23	24
25	25	26	26	28	28	28	28	32	35
36	39	39	41	43	44	45	45	47	49
50	53	55	59	61	70	83	86	86	93

1. Make a stemplot using tens of dollars as the stem and dollars as the leaves. Make sure you include appropriate labels, title, and especially a key.



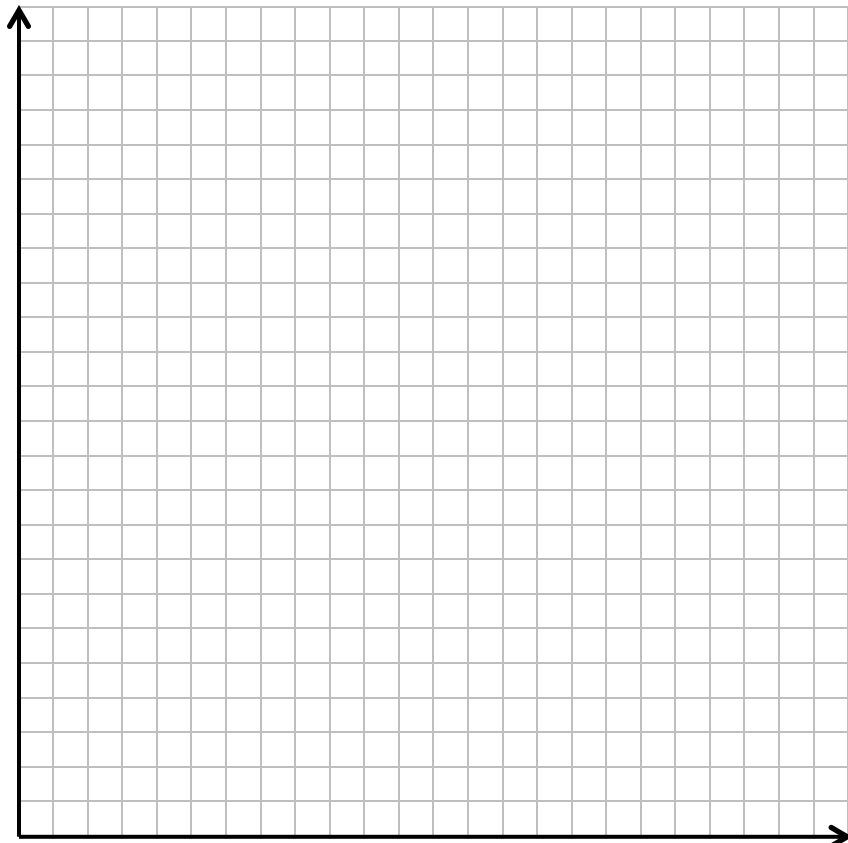
F.) WHERE DO OLDER FOLKS LIVE? – The table gives the percentage of residents 65 or older in each state.

State	Percent	State	Percent	State	Percent
Alabama	13.1	Louisiana	11.5	Ohio	13.4
Alaska	5.5	Maine	14.1	Oklahoma	13.4
Arizona	13.2	Maryland	11.5	Oregon	13.2
Arkansas	14.3	Massachusetts	14.0	Pennsylvania	15.9
California	11.1	Michigan	12.5	Rhode Island	15.6
Colorado	10.1	Minnesota	12.3	South Carolina	12.2
Connecticut	14.3	Mississippi	12.2	South Dakota	14.3
Delaware	13.0	Missouri	13.7	Tennessee	12.5
Florida	18.3	Montana	13.3	Texas	10.1
Georgia	9.9	Nebraska	13.8	Utah	8.8
Hawaii	13.3	Nevada	11.5	Vermont	12.3
Idaho	11.3	New Hampshire	12.0	Virginia	11.3
Illinois	12.4	New Jersey	13.6	Washington	11.5
Indiana	12.5	New Mexico	11.4	West Virginia	15.2
Iowa	15.1	New York	13.3	Wisconsin	13.2
Kansas	13.5	North Carolina	12.5	Wyoming	11.5
Kentucky	12.5	North Dakota	14.4		

Histograms are a way to display groups of quantitative data into bins (the bars). These bins have the same width and scale and are touching because the number line is continuous. To make a histogram you must first decide on an appropriate bin width and count how many observations are in each bin. The bins for percentage of residents aged 65 or older have been started for you below.

1. Finish the chart of bin widths and then create a histogram using those bins on the grid below. Make sure you include appropriate labels, title, and scales.

Bin Widths	Frequency
4 to <6	1
6 to <8	
8 to <10	



G.) SSHA SCORES – Here are the scores on the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women:

154	109	137	115	152	140	154	178	101
103	126	126	137	165	165	129	200	148

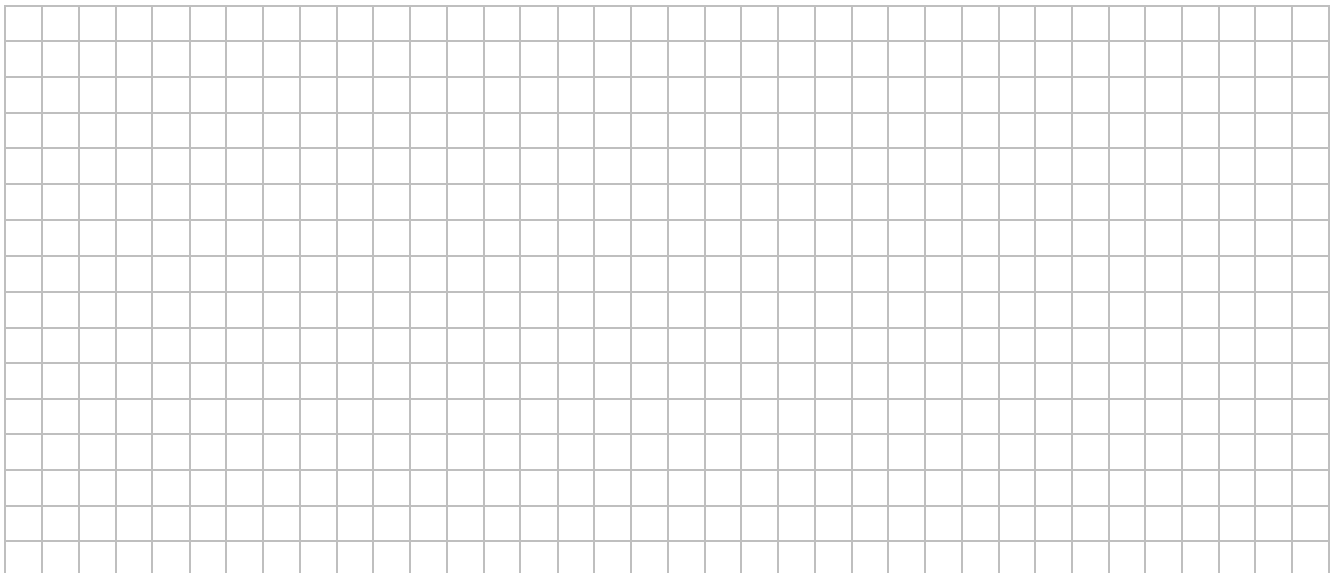
and for 20 first-year college men:

108	140	114	91	180	115	126	92	169	146
109	132	75	88	113	151	70	115	187	104

1. Put the data values in order for each gender. Compute the following numerical summaries for each gender.

Women		Men	
Mean		Mean	
Minimum		Minimum	
Q1		Q1	
Median		Median	
Q3		Q3	
Maximum		Maximum	
Range		Range	
IQR		IQR	

2. Using the Minimum, Q1, Median, Q3, and Maximum from each gender, make parallel boxplots to compare the distributions.



H.) To answer the following, refer to the readings on www.stattrek.com titled "Survey Sampling Methods".

The 7 types of sampling designs are:

- A. Voluntary response B. Convenience C. Simple Random Sampling (SRS) D. Stratified
E. Cluster F. Multistage G. Systematic

1. The Florida division of Weight Watchers (WW) is doing research to determine how many people on the Weight Watchers diet cheat at least once per week. They decide that anonymous surveys will give them an accurate representation but do not have time to get responses from ALL the Florida Weight Watchers participants.

Read the scenarios below and determine which of the 7 sampling methods best describes it.

- _____ I. Randomly select 10 members from each of the WW centers in Florida.
_____ II. Use an alphabetical listing of all Florida WW members. Randomly choose a starting person on the list. Then select every 20th person thereafter.
_____ III. Randomly select 2 or 3 branches of the Florida division and survey every member of that center.
_____ IV. The Florida regional office is in Orlando, so they survey members at the WW center in Orlando.
_____ V. Send out a survey to every member of the Florida division. Place drop boxes in each WW center. Anyone who returns the survey will be in the sample.
_____ VI. From a numbered list of all Florida WW members, use a computer to randomly select 100 numbers and survey all members with those corresponding numbers.

2. What is the population of interest in the Weight Watchers situation?

I.) ALGEBRA!!!! – The prerequisite for AP Statistics is Algebra II. You will not find very much equation solving in this course, but some quick review of Algebra I and Algebra II content will be helpful.

Here is a formula that is used frequently in AP Statistics: $z = \frac{x - \bar{x}}{s}$. Use your algebra skills...

- 1.) If $z = 2.5$, $x = 102$, and $\bar{x} = 100$, what is s ? Show your work.

- 2.) If $z = -3.35$, $x = 60$, and $s = 4$, what is \bar{x} ? Show your work.

J.) LINEAR FUNCTIONS & SCATTERPLOTS – It is expected that you have a thorough understanding of linear functions and scatterplots.

1.) The USDA reported that in 1990 each person in the United States consumed an average of 133 pounds of natural sweeteners. They also claimed this amount has decreased by about 0.6 pounds each year.

a.) If 1990 could be considered “year 0”, which of the above numbers represents the slope and which represents the y -intercept?

b.) What is the equation of the line of best fit using the slope and y -intercept above?

c.) Predict the average consumption of sweeteners per person for the year 2005.

2.) The following equation can be used to predict the average height of boys anywhere between birth and 15 years old: $\hat{y} = 2.79x + 25.64$, where x is the age (in years) and \hat{y} is the predicted height (in inches).

a.) What does the slope represent in this problem? Interpret it in the context of this situation.

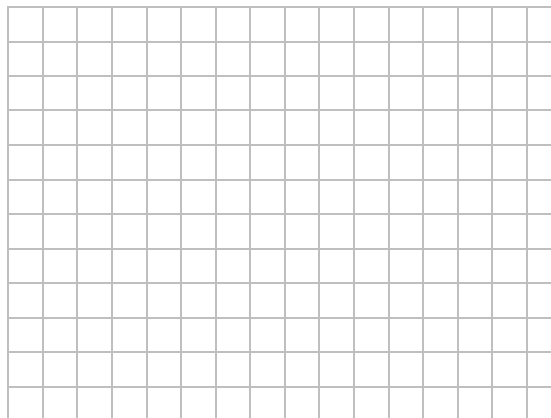
b.) What does the y -intercept represent in this problem? Interpret it in context.

3.) Angie wonders if people of similar heights tend to date each other. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches).

Women	66	64	66	65	70	65
Men	72	68	70	68	74	69

a.) Construct a scatterplot of the data. Include labels and scales.

b.) Describe the association between the heights of the women and the men they date.



K.) PROBABILITY – You are expected to have a basic understanding of simple probability. If you find these problems less than intuitive, there numerous sites available online that provide basic probability explanations. I recommend YouTube – search “basic probability rules”.

1.) A special lottery is to be held to select the student who will live in the only deluxe room in a dormitory. There are 100 seniors, 150 juniors, and 200 sophomores who applied. Each senior’s name is placed in the lottery 3 times; each junior’s name 2 times; and each sophomore’s name 1 time. What is the probability that a senior’s name will be chosen?

- A.) $\frac{1}{8}$ B.) $\frac{2}{9}$ C.) $\frac{2}{7}$ D.) $\frac{3}{8}$ E.) $\frac{1}{2}$

2.) Which of the following has a probability closest to 0.5?

- A.) The sun will rise tomorrow.
B.) It will rain tomorrow.
C.) You will see a dog with only three legs when you leave the room.
D.) A fair die will come up with a score of 6 four times in a row.
E.) There will be a plane crash somewhere in the world within the next five minutes.

3.) If a coin is tossed twice, what is the probability of getting heads on the first toss, and tails on the second toss?
(HINT: Think about the possible outcomes when you toss a coin twice)

- A.) $\frac{1}{6}$
B.) $\frac{1}{3}$
C.) $\frac{1}{4}$
D.) $\frac{1}{2}$
E.) 1

4.) If a coin is tossed twice, what in the probability that it will land either heads both times OR tails both times?

- A.) $\frac{1}{8}$
B.) $\frac{1}{6}$
C.) $\frac{1}{4}$
D.) $\frac{1}{2}$
E.) 1

5.) Calculate the following probabilities and arrange them in order from **least to greatest**.

- I. The probability that a fair die will produce an even number. _____
- II. A random digit from 1 to 9 (inclusive) is chosen, with all digits being equally likely. The probability that when it’s squared the answer will contain the digit 1. _____
- III. The probability that a letter chosen from the alphabet will be a vowel (not counting “y”). _____
- IV. A random number from 1 to 20 (inclusive) is chosen. The probability that it’s square root will not be an integer. _____

Order: _____, _____, _____, _____