Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

AP STATISTICS – MS. KLIMCZUK

**Chapter 2 Questions (From Notes)**

1. Observe the frequency table below of Titanic passengers.



 Now make a relative frequency table of the Titanic passengers below:

1. You might think that a good way to show the *Titanic* data is with this display:



1. What is wrong with this?
2. This shape display violates the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
3. a) How can we fix the last graphical display?

 b) Create a new graphical display that does not violate the area principle.

1. Make a relative frequency bar chart of the Titanic Data below:
2. When you are interested in parts of a whole, what type of display might you choose? Draw this graphical display below:
3. a) What type of table allows us to look at two categorical variables together?

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 Example: We can examine the class of ticket and whether a person survived the *Titanic*:



 b) The margins of the table, both on the right and on the bottom, give totals and the frequency distributions for each of the variables. These are called:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 c) What is the marginal distribution of survival?

 d) Each cell of the table gives the count for a combination of values of the two values. For example,

 what does the second cell in the crew column tells us?

1. A **conditional distribution** shows the distribution of one variable for just the individuals who satisfy some condition on another variable.

The following is the conditional distribution of ticket *Class*, conditional on having survived:



 Now, create the conditional distribution of ticket *Class*, conditional on having perished:

1. Now you can make two pie charts for the two distributions.



1. Is the distribution of *Class* for the survivors different from that of the non-survivors? Explain.
2. Do you think that Class and Survival are associated? Explain.
3. So is Class and Survival independent of each other? Why or why not?

**Independence**: The variables are considered independent when the distribution of one variable in a contingency table is the same for all categories of the other variable.

1. a) What other type of graphical display uses parts of a whole?



How is this type of display similar to a pie chart?

How is it different from a pie chart?