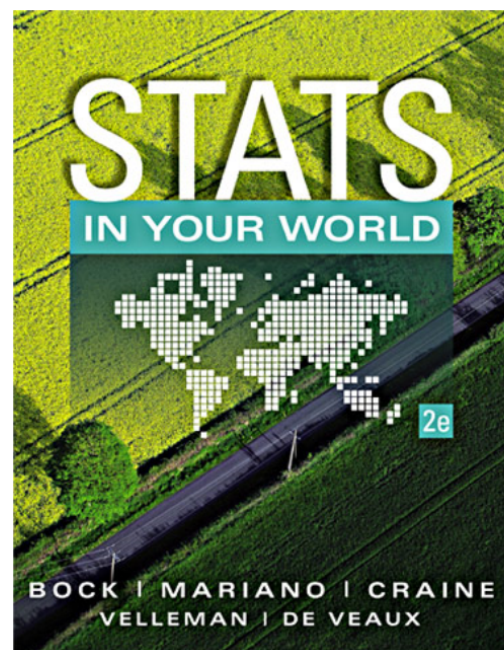


Word Problem Percent Homework:

5. 37 questions
6. 24 exhibits
7. 33.3%
8. 50 students
9. 17 owners
10. 22 ounces

Chapter 2

Stories Categorical
Data Tell



To Do Statistics Right...

Think, Show, Tell (from Chapter 1)

Think - think carefully about what the question is asking and what techniques to use to answer the question.

Show - with calculations and graphs

Tell - explain your results so someone can understand what you've learned through your analysis.

The Four C's for writing conclusions:
(Part for the Tell portion)

Make sure your conclusions are:

- ✓ Clear
- ✓ Concise
- ✓ Complete
- ✓ in Context

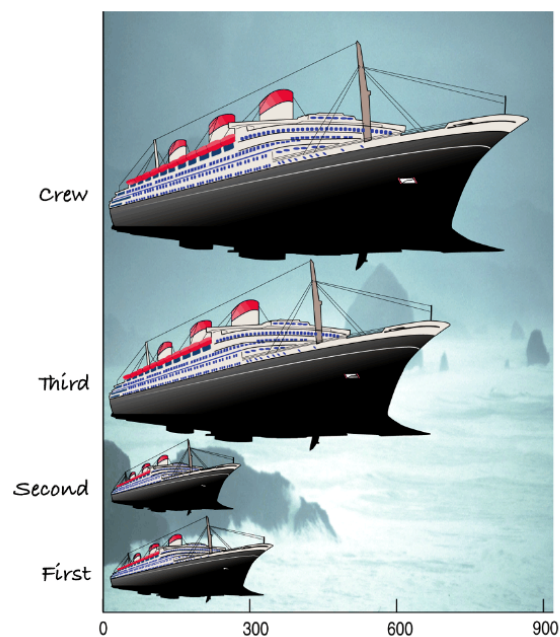
The Three Rules of Data Analysis

The three rules of data analysis won't be difficult to remember:

1. **Make a picture**—things may be revealed that are not obvious in the raw data. These will be things to *think* about.
2. **Make a picture**—important features of and patterns in the data will *show* up. You may also see things that you did not expect: extraordinary data or unexpected patterns.
3. **Make a picture**—the best way to *tell* others about your data is with a well-chosen picture.

What's Wrong With This Picture?

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



Frequency Tables: Making Piles

- We can “pile” the data by counting the number of data values in each category of interest.
- We can organize these **counts** into a **frequency table**, which records the totals and the category names.

Class	Count
First	325
Second	285
Third	706
Crew	885
2201	

Frequency Tables: Making Piles (cont.)

- A **relative frequency table** is similar, but gives the percentages (instead of counts) for each category. (percent)

Class	%	
First	14.77	(325/2201)
Second	12.95	(285/2201)
Third	32.08	(706/2201)
Crew	40.21	(885/2201)

Frequency Tables: Making Piles (cont.)

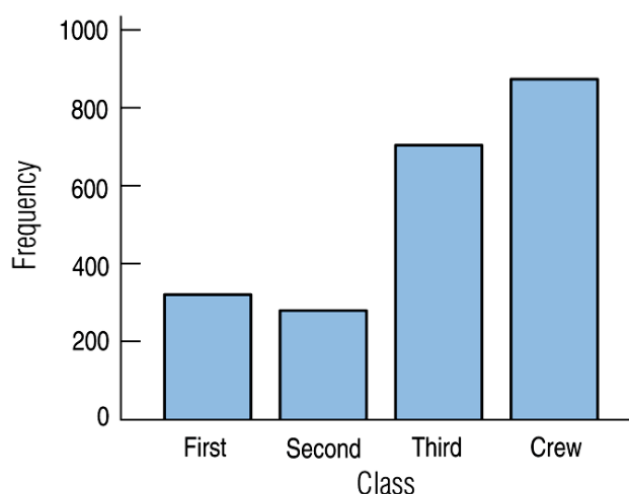
- Both types of tables show how cases are distributed across the categories.
 - They describe the **distribution** of a categorical variable because they name the possible categories and tell how frequently each occurs.
 - Percentages are easier to understand and interpret.
- ★ Pro-tip: **relative frequency** is just a fancy name for percent.

The Area Principle

- The ship display makes it look like most of the people on the *Titanic* were crew members, with a few passengers along for the ride.
- When we look at each ship, we see the *area* taken up by the ship, instead of the *length* of the ship.
- The ship display violates the **area principle**:
 - The area occupied by a part of the graph should correspond to the size of the value it represents.

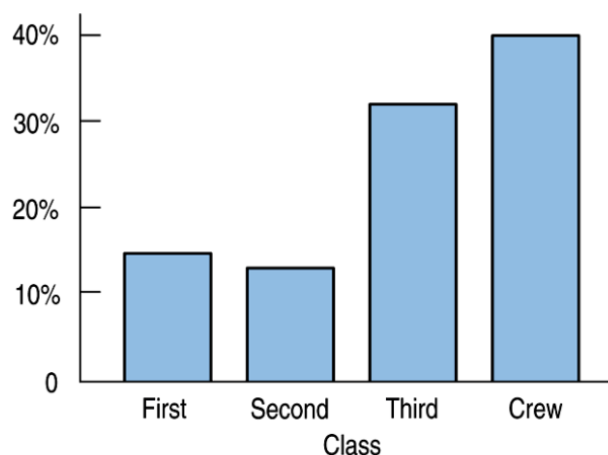
Bar Charts

- A **bar chart** displays the distribution of a categorical variable, showing the counts for each category next to each other for easy comparison.
- A bar chart stays true to the area principle.
- The bars are usually spaced apart to make the graph more readable.



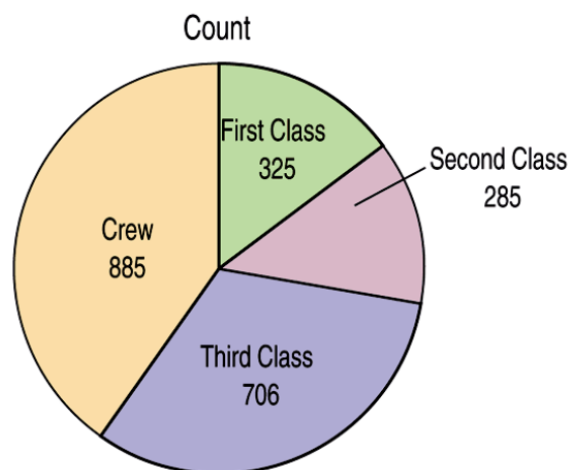
Bar Charts (cont.)

- A **relative frequency bar chart** displays the relative *proportion* of counts for each category. (**percent**)
- A relative frequency bar chart also stays true to the area principle.
- Replacing counts with percentages in the ship data:
- Doesn't look much different, does it?



Pie Charts

- When you are interested in parts of the whole, a **pie chart** might be your display of choice.
- Pie charts show the whole group of cases as a circle.
- They slice the circle into pieces whose size are proportional to the fraction of the whole in each category.



Homework

Read Pg. 16-21

Reading Guide #1, 2, 3(1,2), 4

Packet pg. 13-14

Pg. 20
in text
↓

↑
back of
worksheet

