

# STATISTICS

## INFORMED DECISIONS USING DATA

Fifth Edition

### STATISTICS

INFORMED DECISIONS USING DATA 5e

Michael Sullivan III



## Chapter 2

### Organizing and Summarizing Data

## 2.1 Organizing Qualitative Data

### Learning Objectives

1. Organize Qualitative Data in Tables
2. Construct Bar Graphs
3. Construct Pie Charts

## 2.1 Organizing Qualitative Data

### Introduction

When data is collected from a survey or designed experiment, they must be organized into a manageable form. Data that is not organized is referred to as **raw data**.

#### Ways to Organize Data

- Tables
- Graphs
- Numerical Summaries (Chapter 3)

## 2.1 Organizing Qualitative Data

### 2.1.1 Organize Qualitative Data in Tables (1 of 8)

A **frequency distribution** lists each category of data and the number of occurrences for each category of data.

#### EXAMPLE Organizing Qualitative Data into a Frequency Distribution

The data on the next slide represent the color of M&Ms in a bag of plain M&Ms.

Construct a frequency distribution of the color of plain M&Ms.

## 2.1 Organizing Qualitative Data

### 2.1.1 Organize Qualitative Data in Tables (3 of 8)

#### EXAMPLE Organizing Qualitative Data into a Frequency Distribution

brown, brown, yellow, red, red, red, brown, orange, blue, green,  
blue, brown, yellow, yellow, brown, red, red, brown, brown, brown,  
green, blue, green, orange, orange, yellow, yellow, yellow, red,  
brown, red, brown, orange, green, red, brown, yellow, orange, red,  
green, yellow, yellow, brown, yellow, orange

Color	Tally	Frequency
Brown		12
Yellow		10
Red		
Orange		
Blue		
Green		

## 2.1 Organizing Qualitative Data

### 2.1.1 Organize Qualitative Data in Tables (4 of 8)

#### Frequency table

Color	Tally	Frequency
Brown		12
Yellow		10
Red		9
Orange		6
Blue		3
Green		5

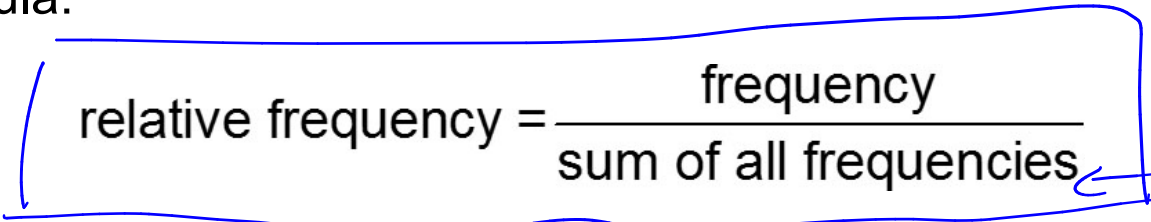


## 2.1 Organizing Qualitative Data



### 2.1.1 Organize Qualitative Data in Tables (5 of 8)

The **relative frequency** is the proportion (or percent) of observations within a category and is found using the formula:


$$\text{relative frequency} = \frac{\text{frequency}}{\text{sum of all frequencies} \leftarrow \text{total}}$$

A **relative frequency distribution** lists each category of data with the relative frequency.

## 2.1 Organizing Qualitative Data

### 2.1.1 Organize Qualitative Data in Tables (6 of 8)

#### EXAMPLE Organizing Qualitative Data into a Relative Frequency Distribution

Use the frequency distribution obtained in the prior example to construct a relative frequency distribution of the color of plain M&Ms.

Color	Tally	Frequency	Relative Frequency
Brown		12	$\frac{12}{40} = 0.3$
Yellow		10	$\frac{10}{40} = 0.25$
Red		9	$\frac{9}{40} = 0.225$
Orange		6	$\frac{6}{40} = 0.15$
Blue		3	$\frac{3}{40} = 0.075$

Copyright © 2017, 2013, 2010 Pearson Education, Inc. All Rights Reserved

Pearson

total

$$12 + 10 + 9 + 6 + 3 = 40$$

add these:  
1.00



## 2.1 Organizing Qualitative Data

### 2.1.2 Construct Bar Graphs (1 of 13)

A **bar graph** is constructed by labeling each category of data on either the horizontal or vertical axis and the frequency or relative frequency of the category on the other axis.

Rectangles of equal width are drawn for each category.

The height of each rectangle represents the category's frequency or relative frequency.

**KEY: in a bar graph, the bars DO NOT TOUCH each other!**

## 2.1 Organizing Qualitative Data

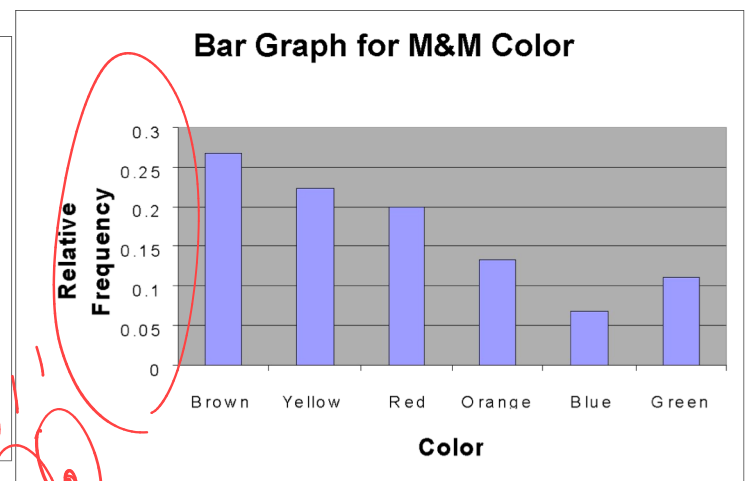
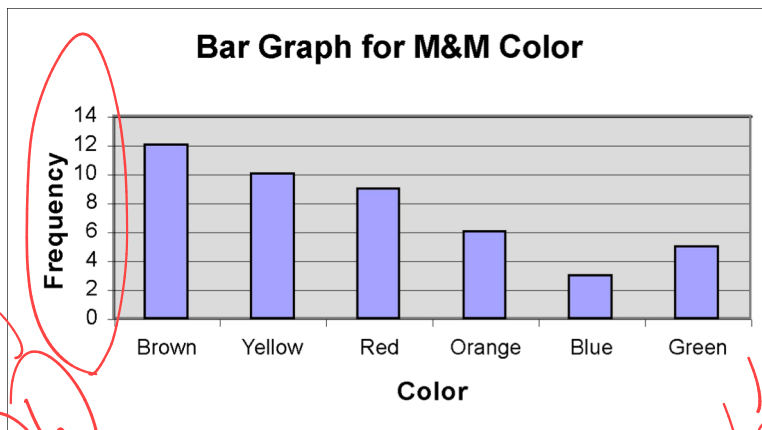
### 2.1.2 Construct Bar Graphs (2 of 13)

#### EXAMPLE Constructing a Frequency and Relative Frequency Bar Graph

Use the M&M data to construct

- a) a frequency bar graph and
- b) a relative frequency bar graph.

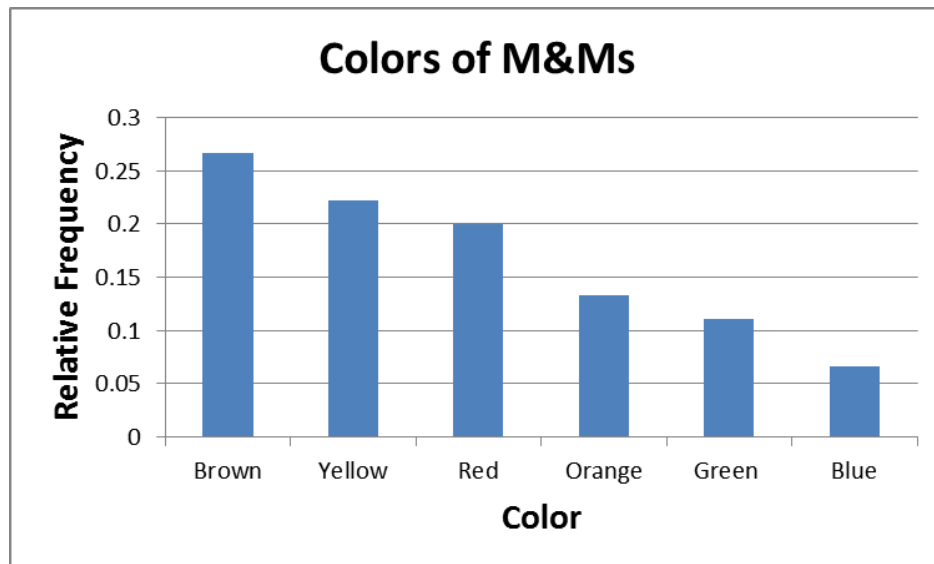
Color	Tally	Frequency	Relative Frequency
Brown		12	$\frac{12}{45} \approx 0.2667$
Yellow		10	0.2222
Red		9	0.2
Orange		6	0.1333
Blue		3	0.0667



## 2.1 Organizing Qualitative Data

### 2.1.2 Construct Bar Graphs (7 of 13)

A **Pareto chart** is a bar graph where the bars are drawn in **decreasing order** of frequency or relative frequency.



## 2.1 Organizing Qualitative Data

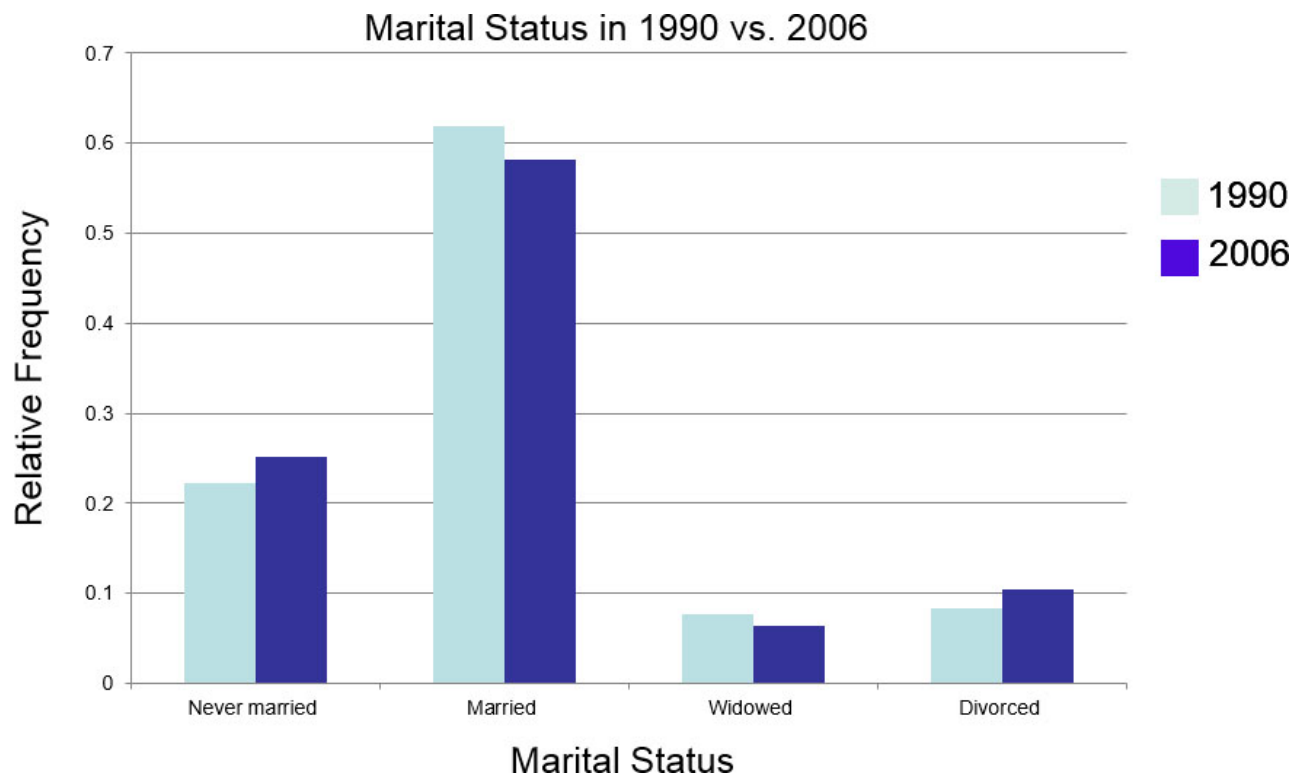
### 2.1.2 Construct Bar Graphs (9 of 13)

#### Side-by-Side Bar Graphs

Suppose we want to know whether more people are finishing college today than in 1990. We could draw a **side-by-side bar graph** to compare the data for the two different years. Data sets should **be compared by using relative frequencies**, because different sample or population sizes make comparisons using frequencies difficult or misleading.

## 2.1 Organizing Qualitative Data

### 2.1.2 Construct Bar Graphs (11 of 13)

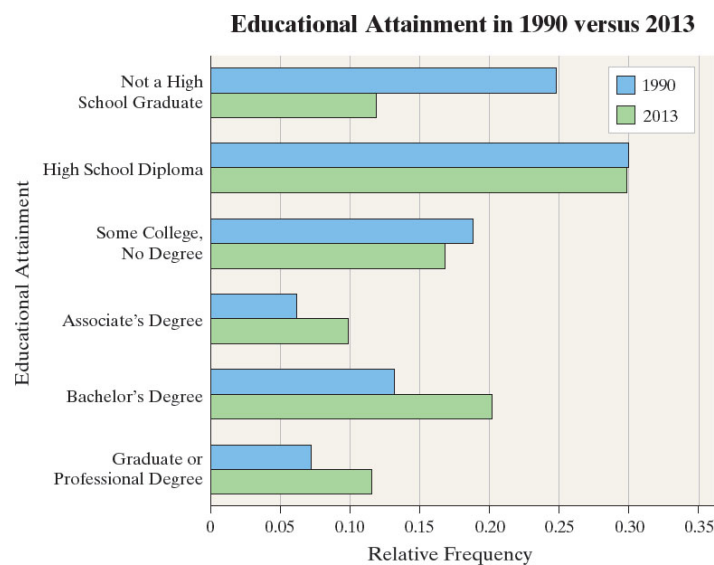


# 2.1 Organizing Qualitative Data

## 2.1.2 Construct Bar Graphs (12 of 13)

### Horizontal Bars

Bar graphs may also be drawn with horizontal bars. Horizontal bars are preferable when category names are lengthy.

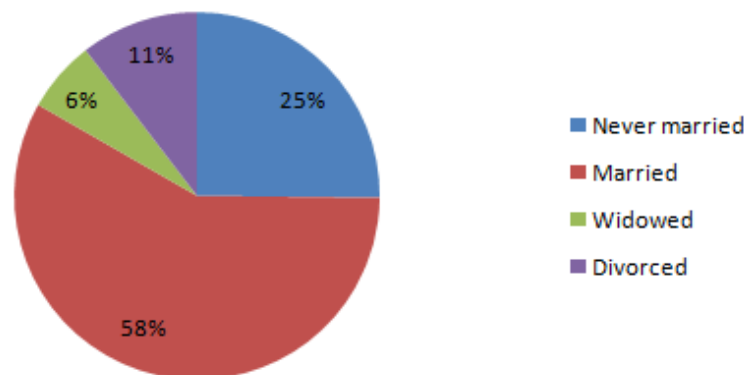


## 2.1 Organizing Qualitative Data

### 2.1.3 Construct Pie Charts (1 of 3)

A **pie chart** is a circle divided into sectors. Each sector represents a category of data. The area of each sector is proportional to the frequency of the category.

**Marital Status, 2006**



out of 100%

## 2.3 Additional Displays of Quantitative Data

### 2.3.2 Construct Frequency and Relative Frequency

#### Ogives (1 of 3)

An **ogive** (read as “oh jive”) is a graph that represents the cumulative frequency or cumulative relative frequency for the class. It is constructed by plotting points whose **x-coordinates are the upper class limits** and whose **y-coordinates are the cumulative frequencies or cumulative relative frequencies** of the class.

Then line segments are drawn connecting consecutive points. An additional line segment is drawn connecting the first point to the horizontal axis at a location representing the upper limit of the class that would precede the first class (if it existed).