CHAPTER 1
Exploring Data

Introduction

Data Analysis:
Making Sense of Data

The Practice of Statistics, 5th Edition
Starnes, Tabor, Yates, Moore



Bedford Freeman Worth Publishers

Data Analysis: Making Sense of Data

Learning Objectives

After this section, you should be able to:

- ✓ IDENTIFY the individuals and variables in a set of data
- ✓ CLASSIFY variables as categorical or quantitative

Data Analysis

Statistics is the science of data.

Data Analysis is the process of *organizing*, *displaying*, *summarizing*, and *asking questions* about data.

Individuals

✓ objects described by a set of data

Variable

✓ any characteristic of an individual

Categorical Variable

✓ places an individual into one of several groups or categories.

Quantitative Variable

✓ takes numerical values for which it makes sense to find an average.

Data Analysis

A **variable** generally takes on many different values.

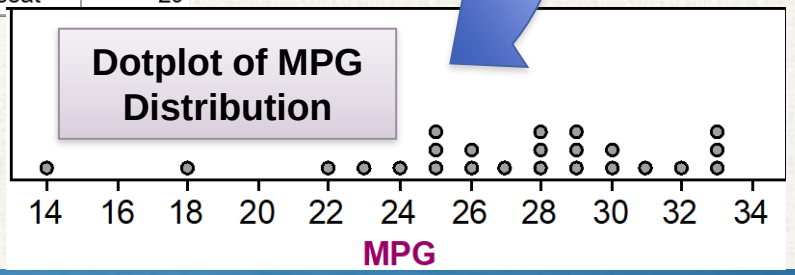
- We are interested in how often a variable takes on each value.

Distribution

✓ tells us what values a variable takes and how often it takes those values.

MODEL	MPG	MODEL	MPG	MODEL	MPG
Acura RL	22	Dodge Avenger	30	Mercedes-Benz E350	24
Audi A6 Quattro	23	Hyundai Elantra	33	Mercury Milan	29
Bentley Arnage	14	Jaguar XF	25	Mitsubishi Galant	27
BMW 5281	28	Kia Optima	32	Nissan Maxima	26
Buick Lacrosse	28	Lexus GS 350	26	Rolls Royce Phantom	18
Cadillac CTS	25	Lincolon MKZ	28	Saturn Aura	33
Chevrolet Malibu	33	Mazda 6	29	Toyota Camry	31
Chrysler Sebring	30	Mercedes-Benz E350	24	Volkswagen Passat	29

Variable of Interest: MPG

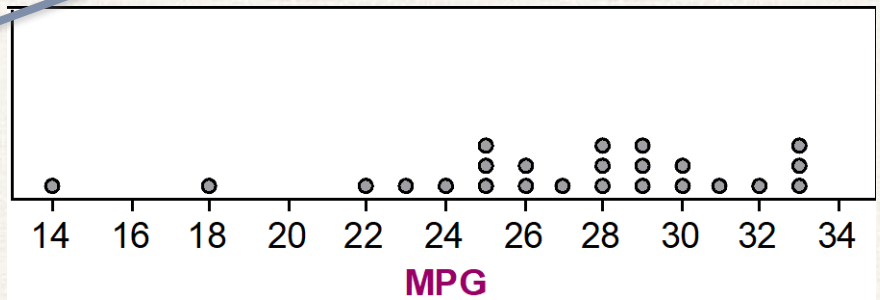


How to Explore Data

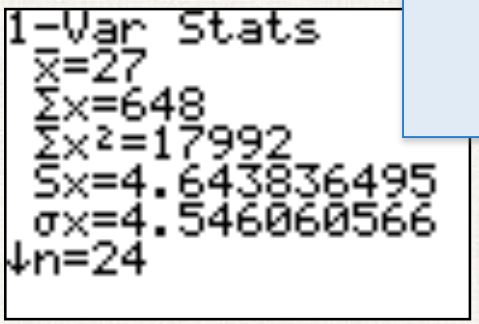
Examine each variable by itself.
Then study relationships among the variables.

MODEL	MPG	MODEL	MPG	MODEL	MPG
Acura RL	22	Dodge Avenger	30	Mercedes-Benz E350	24
Audi A6 Quattro	23	Hyundai Elantra	33	Mercury Milan	29
Bentley Arnage	14	Jaguar XF	25	Mitsubishi Galant	27
BMW 5281	28	Kia Optima	32	Nissan Maxima	26
Buick Lacrosse	28	Lexus GS 350	26	Rolls Royce Phantom	18
Cadillac CTS	25	Lincolon MKZ	28	Saturn Aura	33
Chevrolet Malibu	33	Mazda 6	29	Toyota Camry	31
Chrysler Sebring	30	Mercedes-Benz E350	24	Volkswagen Passat	29

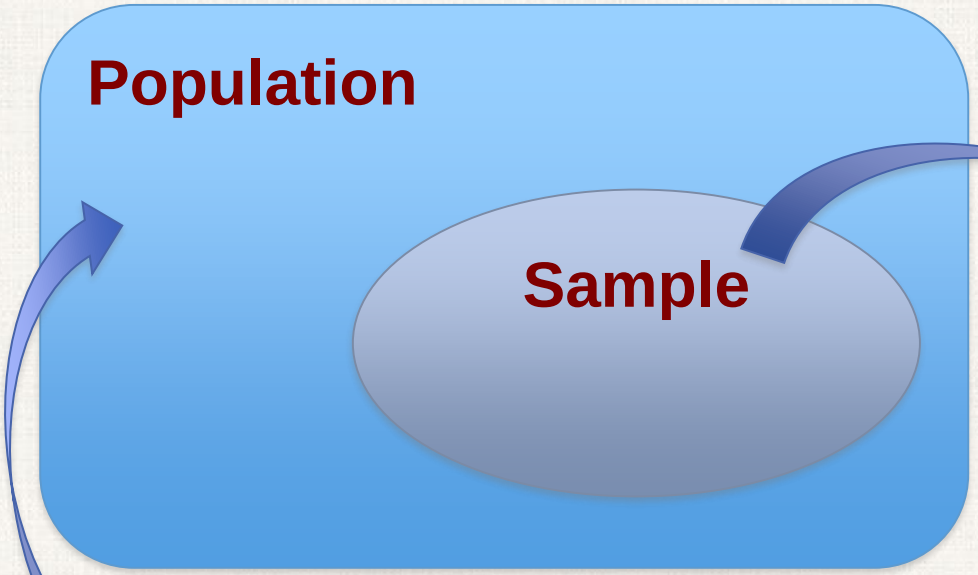
Start with a graph or graphs



Add numerical summaries



From Data Analysis to Inference



Collect data from a representative **Sample**...

Perform **Data Analysis**, keeping probability in mind…

Make an **Inference** about the **Population**.

Data Analysis: Making Sense of Data

Section Summary

In this section, we learned that…

- ✓ A dataset contains information on **individuals**.
- ✓ For each individual, data give values for one or more **variables**.
- ✓ Variables can be **categorical** or **quantitative**.
- ✓ The **distribution** of a variable describes what values it takes and how often it takes them.
- ✓ Inference is the process of making a conclusion about a population based on a sample set of data.