

Statistics is the science of data, and is used to evaluate claims.

Ex: I make 80% of free throws I shoot.

Chapter 1: Picturing Distributions with Graphs

Def: An individual is an object described by data.

Ex: Person, city, animal, company.

Def: A variable is a characteristic of an individual.

Ex: Age, population, species, profit.

Ex: We randomly select 4 people in the US and ask them to report their age and gender. We also ask them what state they're living in.

| State | Age | Reported Gender |
|------------|-----|-----------------|
| Kentucky | 61 | Female |
| Florida | 27 | Female |
| Wisconsin | 27 | Male |
| California | 33 | Female |

4 individuals and 3 variables measured for each individual

catagorical (pointing to State)
quantitative (pointing to Age)
catagorical (pointing to Reported Gender)

Def: A variable is quantitative if it takes numerical values and arithmetic

makes sense.

Def: A variable is categorical if it is not quantitative.

Now we ask for zip codes

| State | Age | Reported Gender | Zip |
|------------|-----|-----------------|-------|
| Kentucky | 61 | Female | 91375 |
| Florida | 27 | Female | 93402 |
| Wisconsin | 27 | Male | 97403 |
| California | 33 | Female | 49102 |

↑
categorical!

Ex: A study classifies bison in Yellowstone as young or adult. State the

individuals, variables, and the type of variable.

Bison, age, categorical

Def: The distribution of a variable is the information of both its possible values and how often they occur.

Ex:

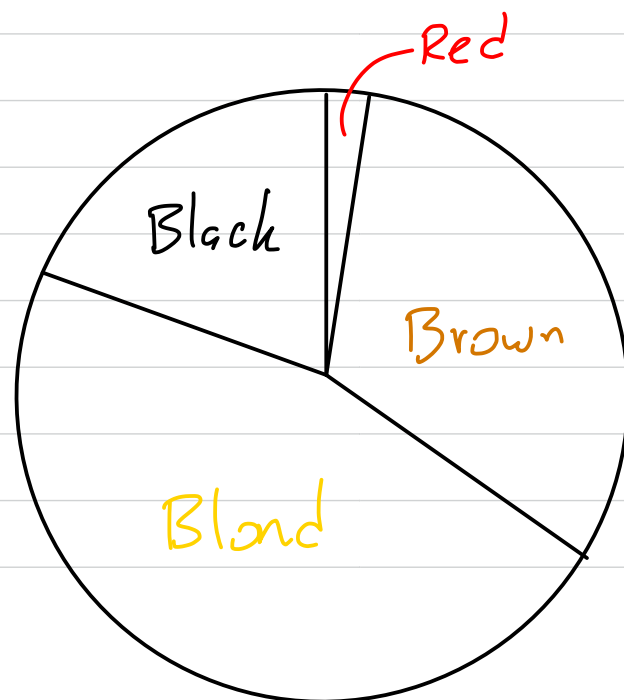
| Student ID | Hair color |
|------------|------------|
| 003 | Red |
| 005 | Brown |
| 035 | Brown |
| 089 | Black |

← not a distribution

| Hair color | % of students w/ this color |
|------------|-----------------------------|
| Red | 2 % |
| Brown | 35 % |
| Blond | 43 % |
| Black | 20 % |

distribution

Pie chart



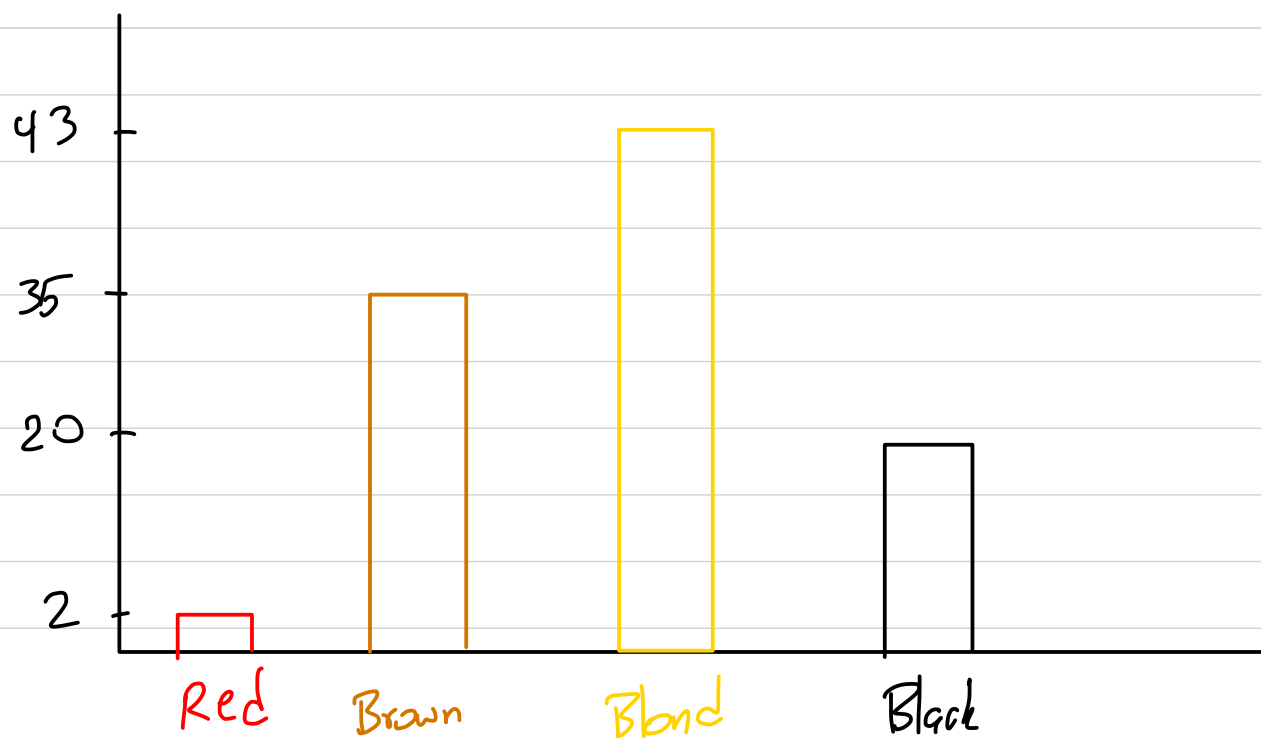
Comment: Only use pie charts when the values the variable can take are mutually exclusive — i.e. every individual has at most one value. Hair color is mutually exclusive since you can have at most one. A survey asking which types of soda you'd had in the past month would not be mutually exclusive since you could have had more than one type.

| | % POP |
|------------|-------|
| Sprite | 30 % |
| Dr. Pepper | 25 % |

this doesn't reflect the people who have had both

| Hair color | % of students w/ this color |
|------------|-----------------------------|
| Red | 2 % |
| Brown | 35 % |
| Blond | 43 % |
| Black | 20 % |

Bar graph:



| <u>Ex</u> | Music source | % of 12-24 year olds who have used it |
|-----------|--------------|---------------------------------------|
| | Radio | 72 |
| | YouTube | 77 |
| | iTunes | 47 |

Don't use a pie chart, because the different music sources aren't mutually exclusive!

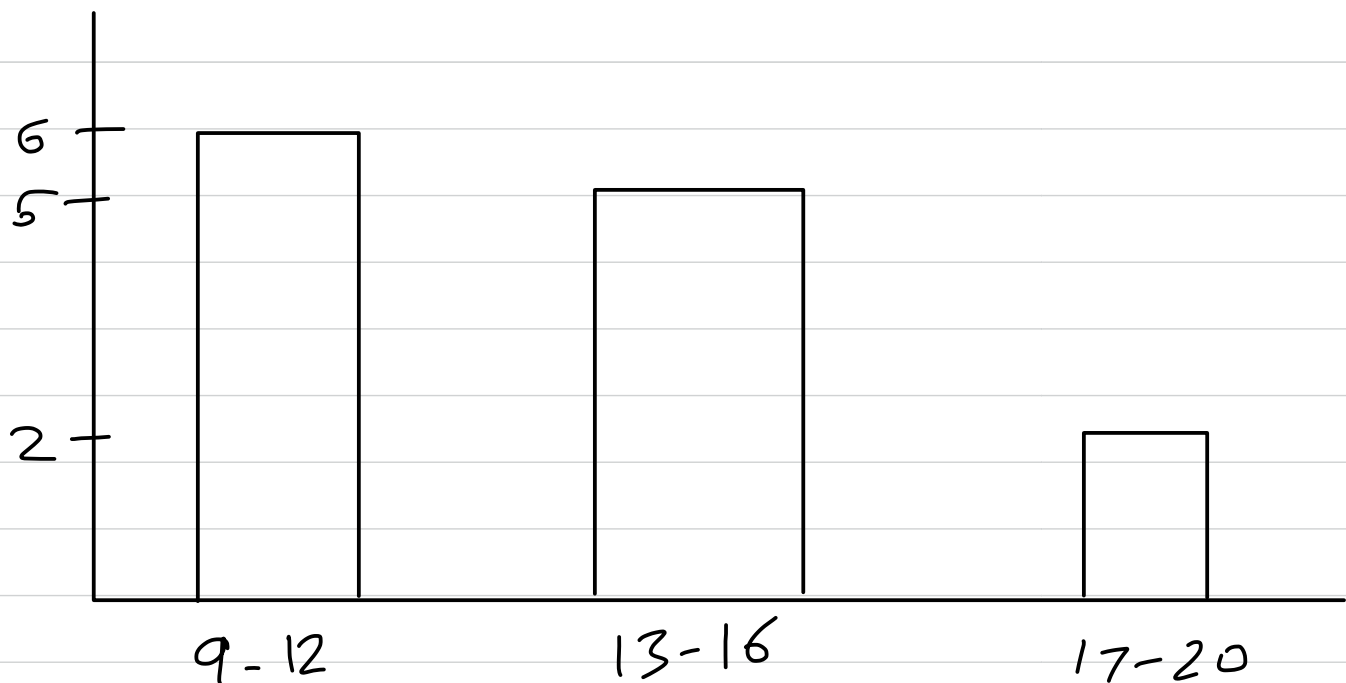
Histograms : when given a sample of individuals, you can make a histogram by dividing the data into ranges (called classes) and counting the number of individuals in each class. Then we make a bar graph of the result. This

roughly approximates the distribution.

Ex: We get a set of ages:

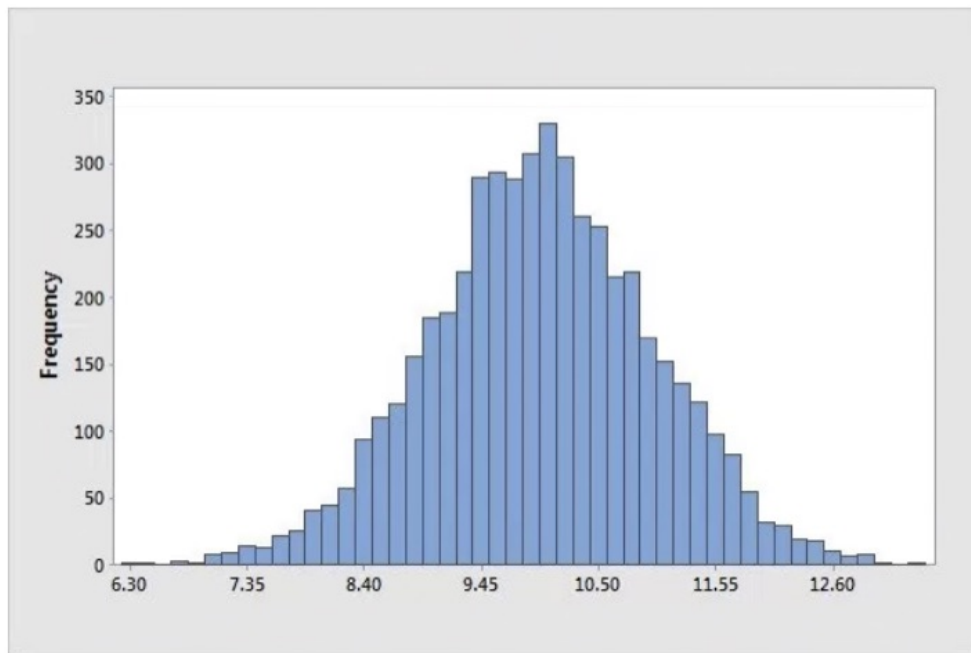
9 10 10 11 12 12 14 15 15
16 16 18 20

Classes: 9-12, 13-16, 17-20
~~~~~  
6 5 2

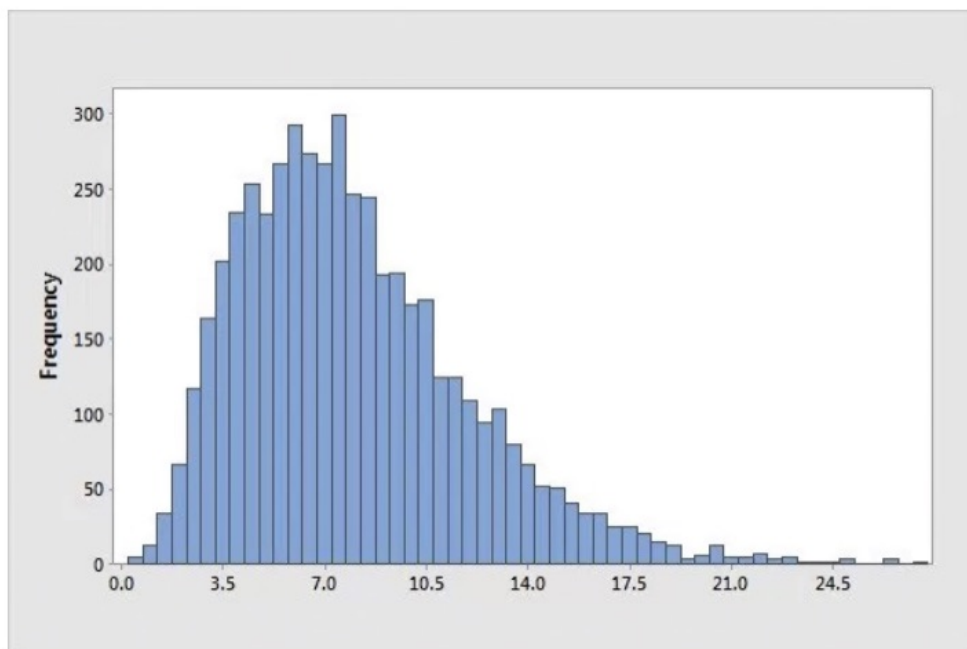


A **symmetric** distribution.

Ex: Heights of young women, Lengths of bird bills

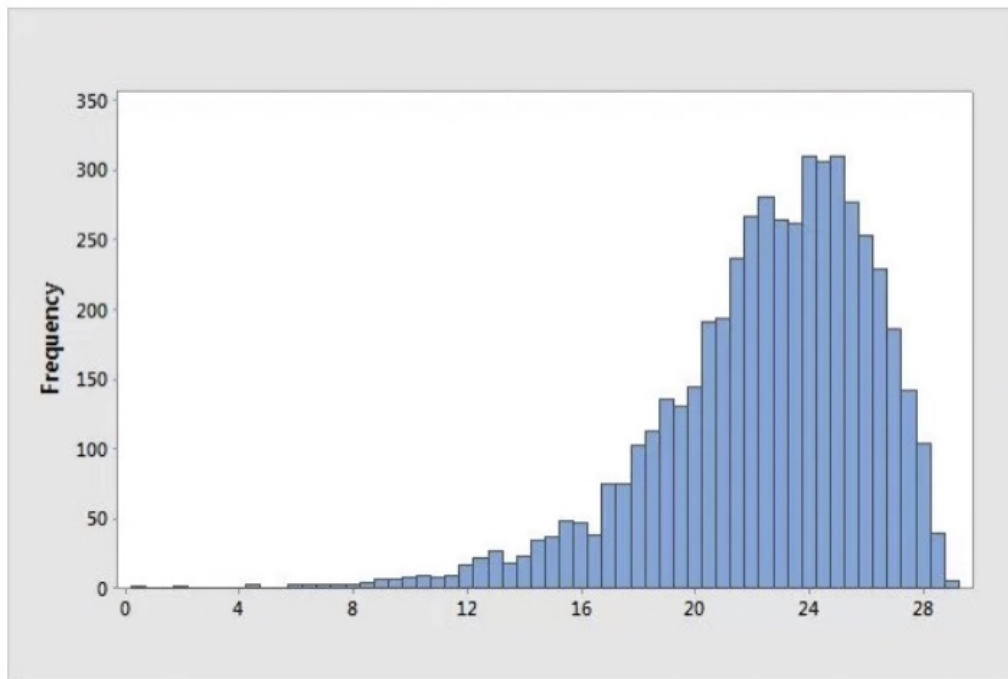


Right-skewed

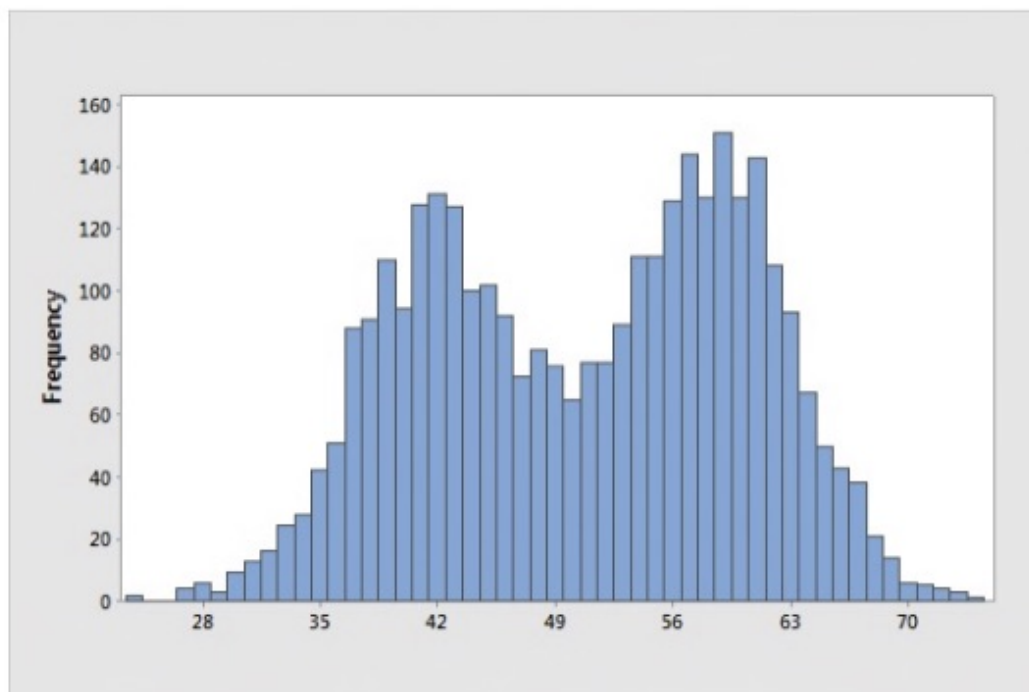


Ex: incomes

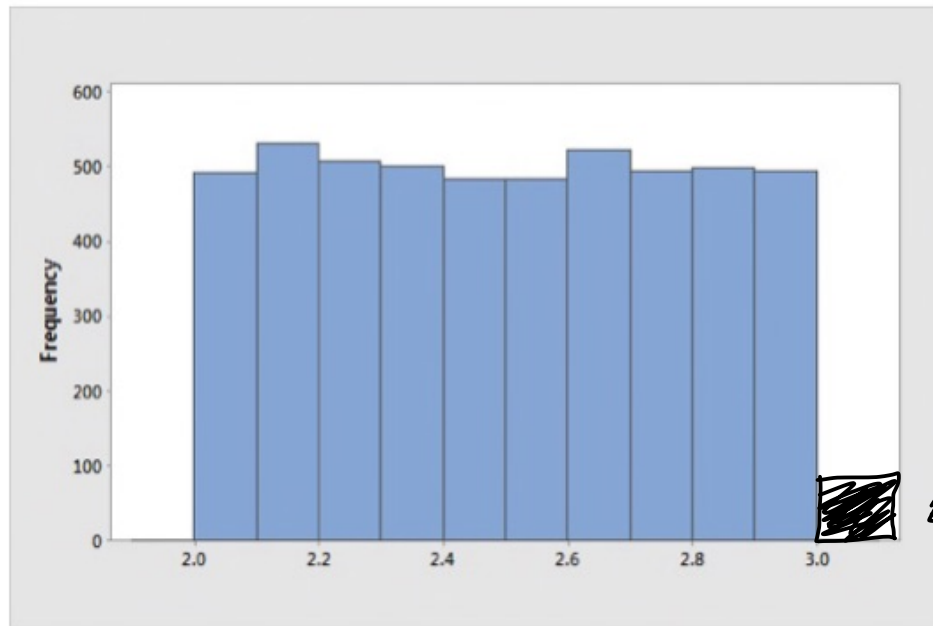
A **left-skewed** distribution.  
Ex: Grades on an easy test



A **bimodal** distribution.  
Ex: Exam scores when one group studied and another didn't



An **approximately uniform** distribution.  
Ex: Rolling a die



outlier

Def: The center of the distribution is the mean or median. The variability is roughly how spread out the distribution is. Outliers are individuals who don't fit the pattern.

Def: Given a set of quantitative data, we can form a stem-and-leaf plot: take all of the numbers and split them into the last digit and all the other digits. Then write the second piece (i.e. the prefix) and all the final digits with that prefix.

Ex: 9 10 10 11 12 12 14 15 15  
16 16 18 20

|   |  |                       |
|---|--|-----------------------|
| 0 |  | 9                     |
| 1 |  | 0 0 1 2 2 4 5 5 6 6 8 |
| 2 |  | 0                     |