



# Lecture 2

## **Displaying Categorical Data with Table**

By the end of this section, you should be able to:

- ✓ Make and interpret frequency table, relative frequency table and two-way table
- ✓ Calculate marginal and joint relative frequencies from a two-way table
- ✓ Calculate conditional relative frequencies from a two-way table

# How to measure categorical variables?

The diagram illustrates the relationship between **Variables** and **Individuals**. Variables are categorized and mapped to individuals to calculate Frequency (counts) and Relative Frequency (percentage/proportion).

	Gender (M/F)	Age	Weight (lbs.)	Height (in.)	Smoking (1=No, 2=Yes)	Race
Patient #1	M	59	175	69	1	White
Patient #2	F	67	140	62	2	Black
Patient #3	F	73	155	59	1	Asian

**Frequency (counts)**  
**Relative Frequency (percentage/proportion)**

**DEFINITION** Frequency table, Relative frequency table

A **frequency table** shows the number of individuals having each value.

A **relative frequency table** shows the proportion or percent of individuals having each value.

Some people use the terms frequency distribution table and relative frequency distribution table instead.

**PROBLEM:** Census At School is an international project that collects data about primary and secondary school students using surveys. Hundreds of thousands of students from Australia, Canada, Ireland, Japan, New Zealand, South Africa, South Korea, the United Kingdom, and the United States have taken part in the project. Data from the surveys are available online. We used the site’s “Random Data Selector” to choose 10 Canadian students who completed the survey in a recent year. The table displays the data.

Province	Gender	Number of languages spoken	Handedness	Height (cm)	Wrist circumference (mm)	Preferred communication
Saskatchewan	Male	1	Right	175.0	180	In person
Ontario	Female	1	Right	162.5	160	In person
Alberta	Male	1	Right	178.0	174	Facebook
Ontario	Male	2	Right	169.0	160	Cell phone
Ontario	Female	2	Right	166.0	65	In person
Nunavut	Male	1	Right	168.5	160	Text messaging
Ontario	Female	1	Right	166.0	165	Cell phone
Ontario	Male	4	Left	157.5	147	Text messaging
Ontario	Female	2	Right	150.5	187	Text messaging
Ontario	Female	1	Right	171.0	180	Text messaging

**Note that the frequencies and relative frequencies listed in these tables are not data !!!!!!!**

Preferred communication	Preferred Communication	Frequency
In person	In Person	3
In person		
Facebook	Facebook	1
Cell phone		
In person	Cell Phone	2
Text messaging		
Cell phone		
Text messaging	Text Messaging	4
Text messaging		
Text messaging		

**Note that the frequencies and relative frequencies listed in these tables are not data !!!!!!!**

Preferred Communication	Frequency	Relative Frequency Proportion	Relative Frequency Percent
In Person	3	0.3	30%
Facebook	1	0.1	10%
Cell Phone	2	0.2	20%
Text Messaging	4	0.4	40%

# TWO-WAY TABLE

A ~~two-way table~~ describes two categorical variables organizing counts according to a row variable and a column variable.

**A two-way table is sometimes called a contingency table.**

Young adults by gender and chance of getting rich			
	Female	Male	Total
Almost no chance	96	98	<b>194</b>
Some chance, but probably not	426	286	<b>712</b>
A 50-50 chance	696	720	<b>1416</b>
A good chance	663	758	<b>1421</b>
Almost certain	486	597	<b>1083</b>
<b>Total</b>	<b>2367</b>	<b>2459</b>	<b>4826</b>

What are the variables described by this two-way table?



The **marginal distribution** of one of **responses** in a two-way table of counts is the distribution of values of **responses** among all individuals described by the table.

## Marginal Distributions

The **marginal distribution** of one of **the categorical variables** in a two-way table of counts is the distribution of values of **that variable** among all individuals described by the table.

Young adults by gender and chance of getting rich		
		Total
Almost no chance		194
Some chance, but probably not		712
A 50-50 chance		1416
A good chance		1421
Almost certain		1083
Total		4826

# Marginal Distributions

The **marginal distribution** of one of the categorical variables in a two-way table of counts is the distribution of values of that variable among all individuals described by the table.

[illegible]

# Marginal Distributions

**Frequency Table**

Almost no chance	<b>194</b>
Some chance, but probably not	<b>712</b>
A 50-50 chance	<b>1416</b>
A good chance	<b>1421</b>
Almost certain	<b>1083</b>
<b>Total</b>	<b>4826</b>

**Relative Frequency Table**

Almost no chance	<b>0.040</b>
Some chance, but probably not	<b>0.148</b>
A 50-50 chance	<b>0.293</b>
A good chance	<b>0.394</b>
Almost certain	<b>0.224</b>
<b>Total</b>	<b>1</b>

These percents or proportions are known as **marginal relative frequencies** because they are calculated using values in the margins of the two-way table.

	optimistic	pessimistic
Male	190	10
Female	190	1810

**Note:** Percentages are often more informative than counts, especially when comparing groups of different sizes.

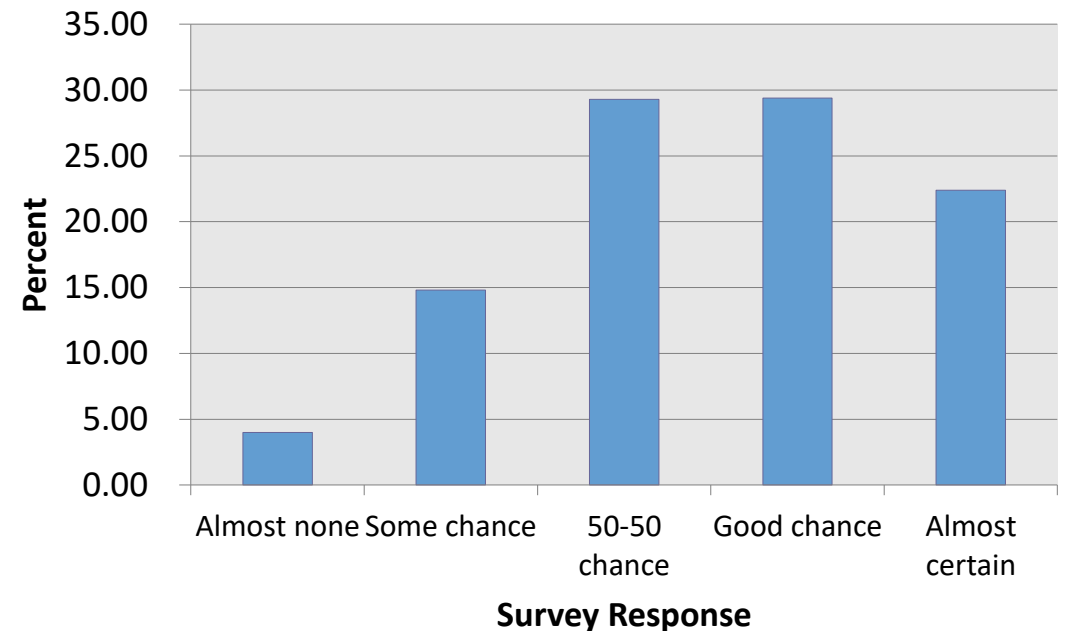
## How to examine a marginal distribution:

- 1) Use the **Examine the marginal distribution of chance of getting rich.**
- 2) Make a graph to display the marginal distribution.

**Relative Frequency Table**

Almost no chance	<b>0.040</b>
Some chance, but probably not	<b>0.148</b>
A 50-50 chance	<b>0.293</b>
A good chance	<b>0.294</b>
Almost certain	<b>0.224</b>
<b>Total</b>	<b>1</b>

**Chance of being wealthy by age 30**




# Conditional Distributions

## Young adults by gender and chance of getting rich

	Female	Male	Total
Almost no chance	96	98	<b>194</b>
Some chance, but probably not	426	286	<b>712</b>
A 50-50 chance	696	720	<b>1416</b>
A good chance	663	758	<b>1421</b>
Almost certain	486	597	<b>1083</b>
<b>Total</b>	<b>2367</b>	<b>2459</b>	<b>4826</b>

-Condition = 'given'

-What percent of people  
“almost certain”,  
given they are female?  
Given they are male?



Response	Female
Almost no chance	$96/2367 = 4.1\%$
Some chance	$426/2367 = 18.0\%$
A 50-50 chance	$696/2367 = 29.4\%$
A good chance	$663/2367 = 28.0\%$
Almost certain	$486/2367 = 20.5\%$

Young adults by gender and chance of getting rich	
	Female
Almost no chance	96
Some chance, but probably not	426
A 50-50 chance	696
A good chance	663
Almost certain	486
<b>Total</b>	<b>2367</b>

**Conditional  
relative  
frequency**



# Homework