

AP STATISTICS

Civic Knowledge & Action: Voter Registration

TEACHER RESOURCE

Overview

According to data from the U.S. Census Bureau, only 46% of the population of voting-age citizens between the ages of 18 and 29 were registered to vote and voted in the 2016 U.S. presidential election. Every two years since 1964, the Census Bureau conducts the Voting and Registration Supplement to the Current Population Survey (CPS), often the week immediately after a national election. These survey data inform the most reliable estimates of the demographic characteristics of the American voting population.

In this lesson, students will develop data analysis skills by exploring CPS data across various demographic categories. Students will also be introduced to content and skills associated with collecting data. In an optional extension project, students will create a final product that spotlights the associations among variables within voting and registration data and provides an opportunity to call others to action to increase voting engagement within their community.

ACTIVITY: GUIDED EXPLORATION (90 MINUTES)

Part 1: Describing Voting Data Represented in a Table (20 minutes) 2.A

- Students will describe categorical voting data represented in a table of frequency counts and relative frequencies.

Part 2: Identifying Information from Voting Data and Constructing Tables to Compare Two Variables (25 minutes) 1.A 1.B 2.B 2.C

- Students will identify questions that might be answered by voting data.
- Students will identify key and relevant information from voting data.
- Students will construct contingency tables to represent two categorical variables pertaining to voting data with frequency counts and relative frequencies.

Part 3: Constructing Graphs to Describe Associations in Voting Data (15 minutes) 2.B 2.D

- Students will construct graphical representations of voting data and use them to determine if two categorical variables pertaining to voting data are associated.

Part 4: Describing Voting Data Collection Methods and Identifying Appropriate Generalizations (25 minutes) 1.C 4.A

- Students will identify types of studies and appropriate generalizations that can be made for sample surveys that collect voting data.

Debrief and Check for Understanding (5 minutes)

SUGGESTED TIMING

The guided exploration activity in this lesson can be completed in one 90-minute class or two 45-minute classes. If desired, the lesson can be extended with the optional investigative mini-project, which is designed to be completed in an additional 90-minute class or two 45-minute classes. If Personal Progress Check (PPC) data from Units 1 and 2 related to Skill Category 2: Data Analysis indicate that students need additional practice on these skills, then the guided exploration activity should be extended to provide extra practice.

AP Exam Connections

The multiple-choice questions on the AP Exam will assess students' abilities to select methods for collecting data and to describe patterns, trends, associations, and relationships in data. One of the free-response questions on the AP Exam will primarily focus on collecting data (Skill 1.C), while another will primarily focus on exploring data (Skills 2.A, 2.B, 2.C, and/or 2.D).

OPTIONAL EXTENSION: INVESTIGATIVE MINI-PROJECT (90 MINUTES)**Part A: Identifying a Question Based on Voting Data (25 minutes)** 1.A 1.B

- Students will identify a question to be answered about a possible relationship between two categorical variables pertaining to voting data.
- Students will identify sources of voting data and relevant information.

Part B: Describing Patterns and Associations in Voting Data (20 minutes)

2.B 2.C 2.D

- Students will compute summary statistics and construct graphical representations for two categorical variables pertaining to voting data.
- Students will use the summary statistics and graphical representations to determine if two categorical variables pertaining to voting data are associated.

Part C: Identifying Generalizations and Drawing Conclusions about Voting Data (20 minutes) 1.C 4.A

- Students will identify appropriate generalizations that can be made based on the method used to gather voting data.
- Students will draw an appropriate conclusion that addresses the question from Part A about a possible relationship between two categorical variables pertaining to voting data.

Part D: Student Presentations (25 minutes)

- Students will present the conclusions from their investigations and invoke a call to action to increase voting engagement within their community.

Connections to the AP Statistics Course Framework**UNIT 1: EXPLORING ONE-VARIABLE DATA**

- **Topic 1.3:** Representing a Categorical Variable with Tables
 - **Learning Objective UNC-1.B:** Describe categorical data represented in frequency or relative tables. 2.A
- **Topic 1.4:** Representing a Categorical Variable with Graphs
 - **Learning Objective UNC-1.C:** Represent categorical data graphically. 2.B

UNIT 2: EXPLORING TWO-VARIABLE DATA

- **Topic 2.1:** Introducing Statistics: Are Variables Related?
 - **Learning Objective VAR-1.D:** Identify questions to be answered about possible relationships in data. 1.A
- **Topic 2.2:** Representing Two Categorical Variables
 - **Learning Objective UNC-1.P:** Compare numerical and graphical representations for two categorical variables. 2.D
- **Topic 2.3:** Statistics for Two Categorical Variables
 - **Learning Objective UNC-1.Q:** Calculate statistics for two categorical variables. 2.C
 - **Learning Objective UNC-1.R:** Compare statistics for two categorical variables. 2.D

Interdisciplinary Opportunities

Additional lessons on displaying and analyzing voting data are available in the following AP courses:

- AP Computer Science Principles
- AP English Language and Composition
- AP Seminar
- AP Spanish Language and Culture
- AP U.S. Government and Politics

If students at your school are enrolled in one or more of these courses, you may wish to collaborate with your colleagues on the best positioning of this lesson within the school year, overlapping content and skills, and additional interdisciplinary connections that can be reinforced between courses. You may also wish to conduct the investigative mini-project in direct collaboration with another AP teacher. For example, AP Statistics students could present their investigative mini-project to an audience of AP U.S. Government and Politics students when they study Topic 5.2: Voter Turnout and/or Disciplinary Practice 3: Data Analysis.

UNIT 3: COLLECTING DATA

- **Topic 3.2:** Introduction to Planning a Study
 - **Learning Objective DAT-2.A:** Identify the type of study. **1.C**
 - **Learning Objective DAT-2.B:** Identify appropriate generalizations and determinations based on observational studies. **4.A**

Formative Assessment with AP Classroom Topic Questions and Personal Progress Checks

- In advance of this lesson, assign the **Topic Questions** for Topics 1.3, 1.4, 2.2, and 2.3. Review the results to identify and address any gaps in students' prior knowledge as you plan your instruction for this lesson.
- After this lesson, assign the **Topic Questions** for Topic 3.2. Review the results to identify and address any student misunderstandings as instruction continues in Unit 3.
- At the conclusion of instruction on Unit 3, assign **Personal Progress Check 3**, which contains two free-response questions that primarily focus on collecting data and exploring data. Review the results to identify and address any student misunderstandings.

Before You Start

This lesson revisits instruction on Skill Category 2: Data Analysis and assumes that students have already worked with the content and skills in Unit 1 and Topics 2.2 and 2.3 in Unit 2. If this lesson will introduce Topics 2.2 and 2.3 to students for the first time, you may wish to forgo the optional investigative mini-project in order to provide additional scaffolding and opportunities for student practice throughout the guided exploration activity.

This lesson introduces Topic 3.2 to students for the first time and serves as a launch for the additional instruction on the content and skills in Unit 3.

Virtual Teaching Tip

Some video conferencing platforms include a "breakout room" feature that can be used to create multiple video meeting spaces within a single video call. If enabled, use this feature to place students, randomly or manually, into collaborative groups throughout this activity. Simulate both the whole-class and collaborative group dynamic by moving students between the main video call and their breakout room throughout the activity.

ACTIVITY: GUIDED EXPLORATION

Part 1 Describing Voting Data Represented in a Table

TIME:
20 minutes

- Place students into collaborative groups of 2-3.
- Distribute **Student Handout 1: Describing Voting Data Represented in a Table**.
- Use the *Notice and Wonder* instructional strategy to allow students to engage with the table at the top of the handout. Allow three minutes of independent work for students to list two things that they "notice" and two things that they "wonder" about the data in the table or the way that the data are displayed. Ask a few students to share what they "notice" or "wonder" before proceeding.

Instructional Strategy:
Notice and Wonder

This instructional strategy allows students to access the relevant information being conveyed by the table without worrying about having to answer something "correctly." Students can share what they "notice" or "wonder" about the table in small groups or with the whole class.

- Direct students to the remaining questions on the handout. Use the *Think-Pair-Share* (or *Wait, Turn, and Talk*) instructional strategy to allow students three minutes to process their responses individually, two minutes to share with a partner, and three minutes for a whole-class debrief.
- Guide students towards the understanding of two issues with the table:
 - The group *Voted* (column 4) is a subset of the group *Registered* (column 2).
 - The table does not include information on the individuals who are not registered to vote.

Build the rationale for why we will construct new contingency tables by highlighting student responses that call out these issues.

- Use the **Student Handout 1: Teacher Key** to drive the discussion and assist with any questions.

Before Moving On

In Part 1, students should be able to describe what frequency counts and relative frequencies are from the provided table. Students should also be able to explain the rationale for building new contingency tables to describe the relationship between *voter registration and turnout* across *age group* and *sex*.

Part 2 Identifying Information from Voting Data and Constructing Tables to Compare Two Variables

TIME:

25 minutes

- Ask students to name an eligibility requirement for registering to vote in the United States that they recall from their history or government class. The requirements that will apply to this lesson are:
 - Be 18 years old on or before Election Day (CPS data only lists ages for the month of November 2016, not the specific date of the election)
 - Be a native or naturalized United States citizen

Guide students towards the understanding that these requirements frame the group of individuals that we are interested in examining. Note that the term “population” will be introduced later in the lesson.
- Distribute **Student Handout 2: Identifying Information from Voting Data and Constructing Tables to Compare Two Variables**.
- Direct students to the question 1 on the handout that uses the *Sentence Starters* strategy. Allow students 1–2 minutes to formalize the questions that will be investigated through the analysis of these data. Ask a student to share their two questions with the whole class. Emphasize the importance of defining “in the 2016 U.S. presidential election” since voter registration and turnout may differ by election type and/or year.
- Distribute **Student Handout A: Voting Data Tables**.

Instructional Strategy: Think-Pair-Share (or Wait, Turn, and Talk)

This instructional strategy allows students uninterrupted time to process their initial ideas and then revise those ideas with a partner in preparation for sharing them with a larger group.

Common Student Misconception

Students may incorrectly interpret the percentages in the table by conditioning on the incorrect variable, such as stating “17.4% of people in the 18 to 29 age group were registered to vote in 2016” or “Only 46.8% of males voted in the 2016 election.” Redirect this misconception by encouraging students to sum the percentages to help identify on which variable the percentage is being conditioned.

Instructional Strategy: Sentence Starters

This instructional strategy helps students practice communication skills by providing a starting point and modeling how to structure questions about the relationship between two categorical variables. Modify the question template to align with the level of scaffolding that your students need.

- Display the second table on this handout. Ask students to call out what they notice about the new data. Answer questions that may arise about the presentation of the categories of the variable “nativity and citizenship status.”
- Release students to work in their collaborative groups on questions 2 and 3 for 8–10 minutes. Use the **Student Handout 2: Teacher Key** while circulating around the room to address student misunderstandings or incorrect computations.

Differentiated Instruction

Guide students who are struggling with the task by probing student thinking instead of providing answers. A sample of common student questions (S) and teacher responses (T) are included below:

- S: How do I compute the number of people who registered but did not vote?
 - T: What relevant information does the first table on Student Handout A provide?
 - T: What was the difference we discussed between the *Registered* and *Voted* categories?
- S: Where do I get the information to fill in the *Not Registered* column of the tables?
 - T: What relevant information does the second table on Student Handout A provide?
 - T: What were the requirements we discussed to be eligible to register to vote?
 - T: How can we use those requirements to compute the total number of people eligible to register to vote?
 - T: How can we use the frequency counts that we now know in the table to compute the number of people eligible to register to vote that did not register?
- When groups have begun to finish, use the *Discussion Groups* strategy to provide students an opportunity to turn to a neighboring group and compare their answers. If groups disagree on any frequency counts they will need to have a discussion to arrive at a consensus for the correct value.
- Ask students to critique to the following claim: “The 18-to-29 age group had about the same number of individuals not registered as the 45-to-64 age group.” Guide students towards the understanding that while the claim is factually correct (since “number” implies a count and 20,071 is roughly equivalent to 20,200), we cannot compare the distributions of *voter registration and turnout across age group* or *sex* unless we compare relative frequencies within each category.
- Release students to work in their collaborative groups on questions 2 and 3 for 5–6 minutes.
- As before, when groups have begun to finish have them compare their answers with a different neighboring group and come to a consensus on any values in disagreement.
- In order to build the rationale for the next part of the activity, ask students to identify what type of representation would help them compare the distributions represented in each row of the tables.

Before Moving On

In Part 2, students should be able to identify the two questions about relationship between *voter registration and turnout*, and *age group* or *sex*. Students should be able to use the provided data tables to identify relevant information from voting data and calculate frequency counts, marginal relative frequencies, and conditional relative frequencies. Students should be able to use the statistics to construct four contingency tables that represent these categorical variables.

Virtual Teaching Tip

A spreadsheet file (.xls) is provided with this lesson that includes pre-populated formulas for students to use to construct contingency tables. Students can interact with this file through spreadsheet software, such as opening it in Microsoft Excel or importing it into Google Sheets.

Instructional Strategy: Discussion Groups

This instructional strategy aids understanding through the sharing of ideas used to identify the information needed to construct the contingency tables.

Review and Reinforce

The total of 100% for each age group and sex category are provided in the contingency tables to guide students towards the understanding of which variable is being conditioned. Reinforce this concept by asking students why the value in those cells is 100% as you circulate around the room.

Part 3 Constructing Graphs to Describe Associations in Voting Data

TIME:

15 minutes

- Ensure students remain in their collaborative groups of 2-3. Distribute **Student Handout 3: Constructing Graphs to Describe Associations in Voting Data**.
- Use the *Look for a Pattern* instructional strategy to allow students 6–8 minutes to complete the segmented bar graph template that is provided for them in the first two questions on the handout. Circulate around the room to ensure that students are labeling axes appropriately and including a legend that defines different categories of *voter registration and turnout*. Use the **Student Handout 3: Teacher Key** to address student misunderstandings or incorrect graphs.
- Conduct a three-minute class discussion by asking students to share the trends or associations they notice in the graphical representations. Some guiding questions to initiate discussion include:
 - Does there appear to be a difference in the distribution of the *voter registration and turnout* variable between the 18-to-29 age group and the 65 and older age group? Between the 45-to-64 age group and the 65 and older age group? Between the 30-to-44 age group and 45-to-64 age group?
 - How can you identify that a difference exists between each age group?
 - Does there appear to be a difference in the distribution of the *voter registration and turnout* variable between males and females?
 - Why can we say that the distribution of the *voter registration and turnout* variable for males and females is relatively the same when the relative frequencies within each category are not exactly equal?
- Allow students four minutes to complete the remaining questions on the handout in their collaborative groups. Emphasize the use of evidence from the segmented bar graph (e.g., comparing specific relative frequencies, comparing heights of bars, etc.) to justify their answers.

Before Moving On

In Part 3, students should be able to construct segmented bar graphs to represent the relationship between *voter registration and turnout* and *age group* or *sex* and use those graphs to determine if there is an association between the variables.

Part 4 Describing Voting Data Collection Methods and Determining Appropriate Generalizations

TIME:

25 minutes

- Distribute **Student Handout 4: Describing Voting Data Collection Methods and Identifying Appropriate Generalizations**.

Virtual Teaching Tip

A spreadsheet file (.xls) is provided with this lesson that includes pre-populated formulas for students to use to construct graphical representations. Students can interact with this file through spreadsheet software, such as opening it in Microsoft Excel or importing it into Google Sheets.

Instructional Strategy: Look for a Pattern

This instructional strategy helps students to identify patterns that may be used to make predictions. Students observe the information from the contingency tables they constructed and create a visual representation to determine if a trend exists in the voting data.

Differentiated Instruction

The template that is provided to students for constructing the segmented bar graphs on Student Handout 3 can be modified or removed depending on the level of scaffolding that your students need. Students who have already mastered constructing segmented bar graphs by hand can be challenged to instead create their graphs from scratch (without using the provided template) in Microsoft Excel or Google Sheets.

- Direct students to the first four questions on the handout. Use the *Think-Pair-Share* (or *Wait, Turn, and Talk*) instructional strategy to allow students 5–6 minutes to process their responses individually, 2 minutes to share with a partner, and 3 minutes for a whole-class debrief.
- As students share their responses during the whole-class debrief, reinforce the concept that samples must be representative of the population in order to make appropriate generalizations about that population.
- Direct students to the excerpt from the U.S. Census Bureau report. Before having students read the excerpt, ask them to read questions 5 and 6. Use the *Marking the Text* instructional strategy to allow students 5 minutes to read the excerpt and highlight, underline, and/or annotate the text to focus on key information that may be relevant to answering questions 5 and 6.
- Ask students to discuss their initial thoughts on questions 5 and 6 with their collaborative groups and form a group consensus before recording their responses on the handout. Conduct a brief whole-class discussion before moving on. Use the **Student Handout 4: Teacher Key** to drive the discussion and address student misunderstandings.
- Allow students 1–2 minutes to independently apply their understandings from questions 5 and 6 and the class discussion in answering question 7. Ask one student to share their responses with the whole class.
- Direct students to the last question on the handout. Read the claim aloud before directing them to independently respond to both parts of the question.
- Ask for 1–2 students to share the reason they listed in part (a) of the question. Reinforce the concept that it is not possible to determine causal relationships between variables using data collected in an observational study.
- With the remaining time in the class, ask students to share the modification they wrote in part (b) of the question. Use the *Peer Critique* strategy to have students comment on whether they feel the claim is an appropriate generalization or provide a reason why they feel it is not.

Before Moving On

In Part 4, students should be able to identify types of studies and appropriate generalizations that can be made for sample surveys that collect voting data.



Debrief and Check for Understanding

TIME:

5 minutes

- Conduct a whole-class debrief in which students discuss one or more of the following:
 - What statistics and graphical representations can we use to represent two categorical variables?
 - What did the statistics that we calculated and the graphical representations that we constructed allow us to identify about the voting data we examined?
 - How does the way we collect data influence what we can and cannot say about a population?

Instructional Strategy: Marking the Text

This instructional strategy helps students identify important information in the text and make notes about the interpretation of tasks required and concepts to apply to reach a solution.

Differentiated Instruction

When completing question 7 on Student Handout 4, some students may point out that the earlier parts of the activity displayed CPS data at such a scale that it appeared to represent the population of voting-age citizens and not a sample. Point them to the language that the U.S. Census Bureau uses in a note at the bottom of the table on Student Handout 1: “Numbers are based on weighted reports of voting behavior derived from a survey sample.” While the data was scaled to represent the population, it was in fact collected from a sample.

Instructional Strategy: Peer Critique

This instructional strategy helps establish criteria for a model response based on content and communication.

- What surprised you about the relationship between variables in the voting data we examined?

Before Moving On

In this lesson, students should be able to calculate statistics and construct graphical representations for two categorical variables, compare distributions to determine if variables are associated, and identify how the data collection methods used influence the generalizations that we can make about a population.

**Teacher Notes**

OPTIONAL EXTENSION: INVESTIGATIVE MINI-PROJECT

Part A Identifying a Question Based on Voting Data

TIME:

25 minutes

- Before beginning this project:
 - Ensure students have access to an internet-connected computer that can open spreadsheet (.xls or .csv) files.
 - Determine whether students will choose their partner for the project or will be assigned a partner.
- Distribute **Student Handout 5: Investigative Mini-Project**.
- Ask a student to read the overview of the project aloud. Review the expectations for Part A: Identifying a Question Based on Voting Data. Emphasize the requirements of the non-trivial question that students will identify.
- Allow students to work with their partner to complete the tasks in Part A of the handout.
- Circulate around the room to listen in on the conversations in which students are engaging. Redirect students who may be considering a question that is trivial or too complex. Ensure students are identifying a question that compares two categorical variables.

Before Moving On

In Part A, students should have identified a question about a possible relationship between two categorical variables to be answered using publicly available state or national election data that pertains to voter registration and/or voter turnout. Students should have identified the source of voting data and relevant information.

Differentiated Instruction

Students may need to oscillate between Task 1 (develop a non-trivial question) and Task 2 (identify one or more sources) as data availability may inform the question they can ask.

Review and Reinforce

Encourage students to refer back to the beginning of Student Handout 2 as they formalize the wording of their project question.

Part B Describing Patterns and Associations in Voting Data

TIME:

20 minutes

- Review the expectations for Part B: Describing Patterns and Associations in Voting Data. Emphasize the requirements of the data analysis component of the project.
- Allow students to work with their partner to complete the tasks in Part B of the handout.
- Circulate around the room to listen to students' conversations. Redirect students who may be considering inappropriate numerical or graphical representations of their project data. Probe student thinking by asking them questions such as:
 - What are the categories of the two variables you selected?

Review and Reinforce

Encourage students to refer back to Student Handout 2 and/or Student Handout 3 as they consider the numerical and/or graphical representations needed to analyze their project data.

Virtual Teaching Tip

If appropriate, students can modify the spreadsheet file that was provided in the guided exploration activity to apply to their project data or create their own spreadsheet file.

- What statistics should be calculated in order to compare these two categorical variables?
- What graphical representations should be constructed in order to compare these two categorical variables?
- Does there appear to be an association between the two categorical variables? What evidence from your graphical representations can you provide?

Before Moving On

In Part B, students should have computed summary statistics and constructed graphical representations, and used them to determine whether there is an association between the two variables in their project question.

Part C Identifying Generalizations and Drawing Conclusions about Voting Data

TIME:

20 minutes

- Review expectations for Part C: Identifying Generalizations and Drawing Conclusions about Voting Data. Emphasize the requirements of the conclusion component of the project.
- Allow students to work with their partner to complete the tasks in Part C of the handout.
- Circulate around the room to listen to students' conversations. Redirect students who may be considering an inappropriate generalization or determination from their project data. Probe student thinking by asking them questions such as:
 - What type of study was conducted to collect the data you used?
 - Is your data from a sample that is representative of the population? How do you know?
 - What generalization can you make about the population?
 - Is there a causal relationship between the two variables in your project question?
- As students transition to creating their infographic handout or digital poster, refer them to the provided template as an example of how they can structure and organize their information.

Before Moving On

In Part C, students should have identified appropriate generalizations that they can make based on the method used to gather the data. Students should have drawn an appropriate conclusion that addresses their project question and constructed an infographic handout or digital poster to communicate their conclusions.

Review and Reinforce

Encourage students to refer back to Student Handout 4 as they consider the appropriateness of generalizations and determinations they wish to make given the method used to gather the data.

Virtual Teaching Tip

Provide a digital copy of the infographic handout template that students can manipulate or restructure to best communicate their conclusions. At the teacher's discretion, students may also choose to use another medium (e.g., slide deck, webpage) to construct their final product.

Part D Student Presentations

TIME:

25 minutes

- Consider the most appropriate method to organize the student presentations for your classroom and students.
- Review with students any additional expectations you have for presenters and audience members. Depending on class size and available time, you may wish to give students a few additional minutes to prepare for their presentation with their partner.
- Conduct the student presentations and, if time allows, allow for the audience to ask questions of the presenters.
- Consider seeking approval from your school's administration to publicly display your students' final products or identify ways in which the students' final products may be incorporated into a larger conversation about voter registration and turnout within the school's community.

Before Moving On

In Part D, students should have presented the conclusions from their investigation and invoked a call to action to increase voting engagement within their community.

**Teacher Notes**

Interdisciplinary Opportunities

Recall that you may wish to conduct Part D in direct collaboration with another AP teacher. For example, AP Statistics students could present their investigative mini-project to an audience of AP U.S. Government and Politics students when they study Topic 5.2: Voter Turnout and/or Disciplinary Practice 3: Data Analysis.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

Civic Knowledge & Action: Voter Registration

STUDENT HANDOUT A

Voting Data Tables

Registration and Voter Characteristics: 2016

(Numbers in thousands)

Characteristic	Registered	Percent	Voted	Percent
Total	157,596	100.0	137,537	100.0
Age				
18 to 29	27,405	17.4	21,620	15.7
30 to 44	36,129	22.9	30,933	22.5
45 to 64	57,394	36.4	51,668	37.6
65 and older	36,667	23.3	33,314	24.2
Sex				
Male	73,761	46.8	63,801	46.4
Female	83,835	53.2	73,735	53.6

Note: Numbers are based on weighted reports of voting behavior derived from a survey sample.

Source: U.S. Census Bureau, Current Population Survey, November 2016.

<https://www.census.gov/content/dam/Census/library/publications/2018/demo/P20-582.pdf>

NATIVITY AND CITIZENSHIP STATUS BY SEX AND AGE: 2016

(Numbers in thousands. Civilian noninstitutionalized population.¹)

Sex and age	Total		Nativity and citizenship status							
			Native		Foreign born ²					
					Total		Naturalized		Not a citizen	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Both sexes	318,868	100.0	275,798	86.5	43,070	13.5	20,086	6.3	22,984	7.2
0 to 14 years	60,965	100.0	59,241	97.2	1,724	2.8	393	0.6	1,332	2.2
15 to 17 years	13,097	100.0	12,161	92.9	936	7.1	254	1.9	681	5.2
18 to 24 years	29,791	100.0	26,708	89.7	3,083	10.3	892	3.0	2,191	7.4
25 to 29 years	22,434	100.0	18,835	84.0	3,599	16.0	1,041	4.6	2,558	11.4
30 to 34 years	21,329	100.0	17,182	80.6	4,147	19.4	1,220	5.7	2,926	13.7
35 to 39 years	20,387	100.0	15,796	77.5	4,591	22.5	1,648	8.1	2,943	14.4
40 to 44 years	19,618	100.0	15,061	76.8	4,557	23.2	1,915	9.8	2,642	13.5
45 to 49 years	20,679	100.0	16,320	78.9	4,359	21.1	2,156	10.4	2,203	10.7
50 to 54 years	21,887	100.0	17,917	81.9	3,970	18.1	2,215	10.1	1,755	8.0
55 to 59 years	21,835	100.0	18,559	85.0	3,276	15.0	2,046	9.4	1,230	5.6
60 to 64 years	19,300	100.0	16,602	86.0	2,698	14.0	1,779	9.2	919	4.8
65 to 69 years	16,517	100.0	14,279	86.5	2,238	13.5	1,574	9.5	664	4.0
70 to 74 years	11,425	100.0	9,997	87.5	1,428	12.5	1,072	9.4	355	3.1
75 to 79 years	8,421	100.0	7,341	87.2	1,080	12.8	820	9.7	260	3.1
80 to 84 years	5,637	100.0	4,940	87.6	697	12.4	522	9.3	175	3.1
85 years and over	5,547	100.0	4,859	87.6	688	12.4	538	9.7	150	2.7

Sex and age	Total		Nativity and citizenship status							
			Native		Foreign born ²					
	Total				Naturalized		Not a citizen			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Male	156,205	100.0	135,234	86.6	20,971	13.4	9,309	6.0	11,662	7.5
0 to 14 years	31,114	100.0	30,263	97.3	851	2.7	185	0.6	666	2.1
15 to 17 years	6,624	100.0	6,142	92.7	482	7.3	117	1.8	365	5.5
18 to 24 years	15,095	100.0	13,486	89.3	1,609	10.7	464	3.1	1,145	7.6
25 to 29 years	11,287	100.0	9,464	83.8	1,823	16.2	500	4.4	1,323	11.7
30 to 34 years	10,558	100.0	8,480	80.3	2,078	19.7	593	5.6	1,485	14.1
35 to 39 years	10,056	100.0	7,743	77.0	2,313	23.0	742	7.4	1,571	15.6
40 to 44 years	9,644	100.0	7,426	77.0	2,218	23.0	887	9.2	1,330	13.8
45 to 49 years	10,127	100.0	7,990	78.9	2,137	21.1	996	9.8	1,141	11.3
50 to 54 years	10,712	100.0	8,784	82.0	1,928	18.0	1,034	9.7	894	8.3
55 to 59 years	10,626	100.0	9,019	84.9	1,606	15.1	993	9.3	614	5.8
60 to 64 years	9,153	100.0	7,878	86.1	1,274	13.9	825	9.0	449	4.9
65 to 69 years	7,775	100.0	6,708	86.3	1,067	13.7	750	9.6	317	4.1
70 to 74 years	5,186	100.0	4,594	88.6	592	11.4	463	8.9	128	2.5
75 to 79 years	3,854	100.0	3,384	87.8	470	12.2	360	9.4	110	2.9
80 to 84 years	2,400	100.0	2,152	89.7	248	10.3	186	7.7	62	2.6
85 years and over	1,995	100.0	1,720	86.2	275	13.8	214	10.7	61	3.1
Female	162,664	100.0	140,564	86.4	22,100	13.6	10,777	6.6	11,322	7.0
0 to 14 years	29,851	100.0	28,978	97.1	874	2.9	208	0.7	666	2.2
15 to 17 years	6,473	100.0	6,020	93.0	454	7.0	137	2.1	316	4.9
18 to 24 years	14,696	100.0	13,222	90.0	1,474	10.0	428	2.9	1,046	7.1
25 to 29 years	11,146	100.0	9,371	84.1	1,776	15.9	541	4.9	1,235	11.1
30 to 34 years	10,771	100.0	8,702	80.8	2,069	19.2	628	5.8	1,441	13.4
35 to 39 years	10,331	100.0	8,053	77.9	2,278	22.1	906	8.8	1,372	13.3
40 to 44 years	9,974	100.0	7,635	76.5	2,339	23.5	1,028	10.3	1,311	13.1
45 to 49 years	10,552	100.0	8,330	78.9	2,222	21.1	1,160	11.0	1,062	10.1
50 to 54 years	11,174	100.0	9,132	81.7	2,042	18.3	1,181	10.6	861	7.7
55 to 59 years	11,209	100.0	9,539	85.1	1,670	14.9	1,053	9.4	617	5.5
60 to 64 years	10,148	100.0	8,724	86.0	1,423	14.0	954	9.4	470	4.6
65 to 69 years	8,742	100.0	7,572	86.6	1,170	13.4	824	9.4	346	4.0
70 to 74 years	6,239	100.0	5,403	86.6	836	13.4	609	9.8	227	3.6
75 to 79 years	4,567	100.0	3,957	86.7	610	13.3	460	10.1	150	3.3
80 to 84 years	3,237	100.0	2,787	86.1	450	13.9	337	10.4	113	3.5
85 years and over	3,552	100.0	3,140	88.4	413	11.6	324	9.1	89	2.5

¹Plus armed forces living off post or with their families on post.

²The foreign born includes anyone who was not a U.S. citizen or U.S. national at birth. Conversely, native includes anyone who was a U.S. citizen at birth.

Note: Details may not sum to totals because of rounding.

Source: U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplement, 2016.

https://www2.census.gov/programs-surveys/demo/tables/age-and-sex/2016/age-sex-composition/2016gender_table4.xls

Describing Voting Data Represented in a Table

In September 2018, the U.S. Census Bureau released a report entitled "Characteristics of Voters in the Presidential Election of 2016." The following table is a subset of a larger table from the report.

Registration and Voter Characteristics: 2016

(Numbers in thousands)

Characteristic	Registered	Percent	Voted	Percent
Total	157,596	100.0	137,537	100.0
Age				
18 to 29	27,405	17.4	21,620	15.7
30 to 44	36,129	22.9	30,933	22.5
45 to 64	57,394	36.4	51,668	37.6
65 and older	36,667	23.3	33,314	24.2
Sex				
Male	73,761	46.8	63,801	46.4
Female	83,835	53.2	73,735	53.6

Note: Numbers are based on weighted reports of voting behavior derived from a survey sample.

Source: U.S. Census Bureau, Current Population Survey, November 2016.

<https://www.census.gov/content/dam/Census/library/publications/2018/demo/P20-582.pdf>

- Examine the table above and the data it represents. List two things that you "notice" and two things that you "wonder" about these data and/or the way in which they are displayed.

Notice	Wonder
■	■
■	■

- Interpret in context of the value 17.4 from the table.

3. A student examining these data believes that there were 27,405 thousand people age 18–29 who **registered but did not vote** in the 2016 U.S. presidential election. Explain the error in the student's claim and how the way in which the data are displayed might have contributed to the student's misconception.

4. What additional data are needed in order to determine the percentage of each age group (or sex) who registered for and voted in the 2016 U.S. presidential election?

Gathering Voting Data and Constructing Tables to Compare Two Variables

While each U.S. citizen (native or naturalized) who is 18 years or older has the constitutional right to vote, each individual must register to vote within their home state. Each individual in the population can thus be classified by *voter registration and turnout* as: not registered for, registered for but did not vote, or registered for and voted in the 2016 U.S. presidential election.

In this exercise, we wish to examine the role that the variables *age group* and *sex* have on *voter registration and turnout* in the 2016 U.S. presidential election.

1. Identify the two questions that we wish to answer about possible relationships in these data.

- Is there an association between _____ and _____
in _____?
- Is there an association between _____ and _____
in _____?

Use the data available in **Student Handout A: Voting Data Tables** to complete the following tasks.

2. Construct a contingency table to compare *voter registration and turnout* across *age group*.

Age Group	Not Registered (thousands)	Registered, Did Not Vote (thousands)	Registered, Voted (thousands)	Total
18 to 29				
30 to 44				
45 to 64				
65 and older				
Total				

3. Construct a contingency table to compare *voter registration* and *turnout* across *sex*.

Sex	Not Registered (thousands)	Registered, Did Not Vote (thousands)	Registered, Voted (thousands)	Total
Male				
Female				
Total				

As the total counts in each category of the *age group* and *sex* variables vary, relative frequencies must be used to compare the distribution of *voter registration* and *turnout* across the categories.

4. Construct a new contingency table to compare the conditional relative frequencies of *voter registration* and *turnout* across each category of *age group*.

Age Group	Not Registered (percent)	Registered, Did Not Vote (percent)	Registered, Voted (percent)	Total
18 to 29				100%
30 to 44				100%
45 to 64				100%
65 and older				100%
Total				100%

5. Construct a new contingency table to compare the conditional relative frequencies of *voter registration* and *turnout* across each category of *sex*.

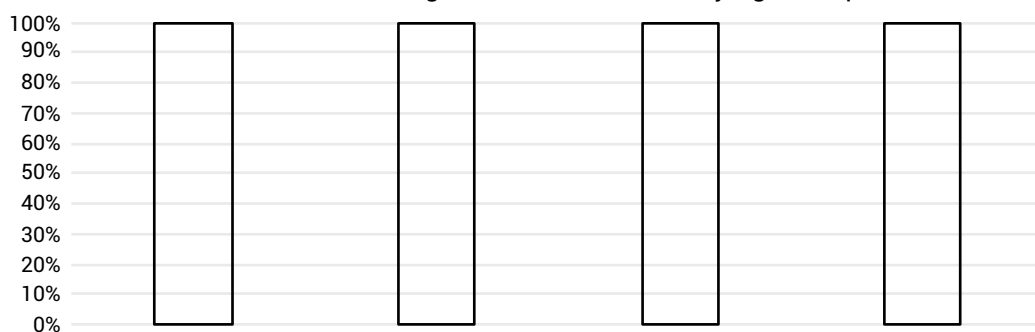
Sex	Not Registered (percent)	Registered, Did Not Vote (percent)	Registered, Voted (percent)	Total
Male				100%
Female				100%
Total				100%

Constructing Graphs to Describe Associations in Voting Data

Use the tables constructed in **Student Handout 2: Gathering Voting Data and Constructing Tables to Compare Two Variables** to complete the following tasks.

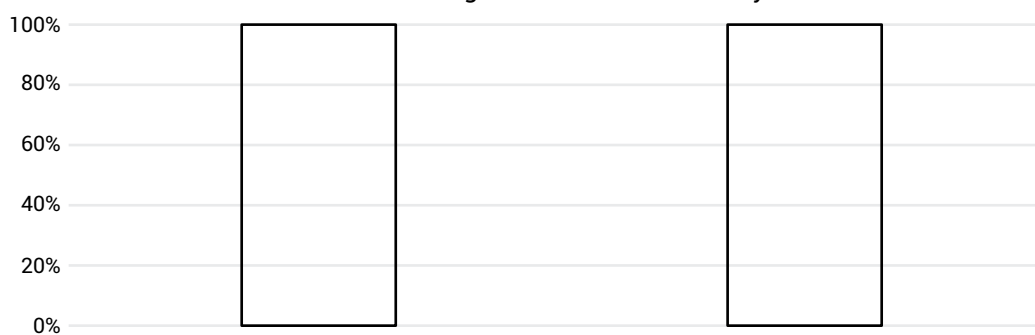
1. Construct a segmented bar graph to compare *voter registration* and *turnout* across *age group*.

2016 Voter Registration and Turnout by Age Group



2. Construct a segmented bar graph to compare *voter registration* and *turnout* across *sex*.

2016 Voter Registration and Turnout by Sex



3. Does there appear to be an association between *voter registration* and *turnout* and *age group*? Use evidence from the segmented bar graph to justify your answer.

4. Does there appear to be an association between voter registration and turnout and sex? Use evidence from the segmented bar graph to justify your answer.

Describing Voting Data Collection Methods and Identifying Appropriate Generalizations

1. Each row of the table below provides an example of a population and a sample from that population. Read each example and use them to write a definition, in your own words, for the two terms at the top of the table.

Population:	Sample:
All Ohio residents	All registered voters in Ohio
All individuals in the U.S. population	All individuals in the U.S. who are eligible to register to vote
All individuals in the U.S. who are eligible to register to vote	All native U.S. citizens
All individuals who were registered to vote in the 2016 U.S. presidential election	All individuals who voted in the 2016 U.S. presidential election

2. Do you think each sample above is *representative* of the corresponding population? Explain why or why not.

3. When using data from a sample to make generalizations about a population, why is it essential for the sample to be *representative* of the population?

4. Read the following example of an observational study and an experiment. Use the example to write a definition, in your own words, for the two terms at the top of the table.

Observational study:	Experiment:
Survey each voter in a sample of voters on who they plan to vote for.	Show a political ad to half of the voters in a sample and then survey each voter in the sample on who they plan to vote for.

The data used throughout this lesson were obtained from the U.S. Census Bureau Current Population Survey (CPS) in November 2016. Read the following excerpt from the U.S. Census Bureau report about these data.

The CPS is a monthly labor force survey in which interviews are conducted in approximately 56,000 households across the country... and every November following national elections, the CPS includes supplemental questions about voting and registration. Either self-responses or proxy responses, where a single respondent can provide answers for themselves and all additional eligible household members, are allowed. The voting questions are asked of all U.S. citizens who are 18 years or older, and these eligibility criteria are determined using questions about age and citizenship status asked during the core survey ... Analysis of the Voting and Registration Supplement can vary based on the selected population universe. Typical population universes used by researchers include the total voting-age population (anyone aged 18 and older), the citizen voting-age population (anyone aged 18 and older who is also a citizen), and the registered voter population.

The data in this report are based on responses to the CPS November Voting and Registration Supplements, which survey the civilian noninstitutionalized population in the United States. Voting estimates from the CPS and other sample surveys have historically differed from those based on administrative records, such as the official reports from each state disseminated collectively by the Clerk of the U.S. House of Representatives and the Federal Elections Commission. In general, voting rates from sample surveys like the CPS tend to be higher than official results (Bauman and Julian, 2010; DeBell, et al., 2015; McDonald, 2015; Tourangeau, 2010). Potential explanations for these differences include question misreporting, problems with memory or knowledge of others' behavior, and methodological issues related to question wording, method of survey administration, and survey nonresponse bias. Despite these observed differences between CPS estimates and official tallies, the CPS data remain the most comprehensive data source available for examining the social and demographic composition of American voters in federal elections, particularly when examining broad historical results.

Source: U.S. Census Bureau, Characteristics of Voters in the Presidential Election of 2016, September 2018
<https://www.census.gov/content/dam/Census/library/publications/2018/demo/P20-582.pdf>

5. Were the data from the CPS November Voting and Registration Supplements collected through an observational study or an experiment? Explain your answer.

6. Is the sample from which these CPS data were collected *representative* of the population of voting-age citizens? Explain your answer.

7. Based on the excerpt from the report, do you believe that it is appropriate to generalize the associations you identified between *voter registration and turnout* and *age group* and *sex* to the population of voting-age citizens?

8. A student examining these data makes the following claim:

"A person's older age causes them to be more participatory in voting."

- a. Provide a reason why this student's claim is not valid.

- b. How could the student's claim be modified to represent an appropriate generalization of the data?

Investigative Mini-Project

OVERVIEW

Voting is considered to be one of the most fundamental civic duties in the U.S. However, data from the U.S. Census Bureau show that only 46% of the population of voting-age citizens between the ages of 18 and 29 were registered to vote and voted in the 2016 U.S. presidential election. Analyzing data from state and national elections can identify trends among subgroups of the population of voting-age citizens that can be utilized to drive change and calls to action.

For this task, you and a partner will investigate a question pertaining to voter registration and turnout data that is of interest to you. You will identify relevant data, calculate appropriate statistics and construct graphical representations to analyze the data, describe any patterns or associations that are in the data, and identify appropriate generalizations in the communication of your conclusions.

PART A: IDENTIFYING A QUESTION BASED ON VOTING DATA

1. Develop a non-trivial question that will provide the basis of your investigation. The question must:

- be answerable using state or national election data that are publicly available,
- pertain in some way to voter registration and/or voter turnout, and
- invoke the comparison of the distributions of two categorical variables and an examination of whether the two categorical variables are associated.

With your partner, make a list of potential *categorical* variables you are interested in examining.

-
-
-
-
-
-

As you examine your list of potential variables and consider the data that are publicly available for those variables, choose the two categorical variables that you are interested in comparing for this project.

Categorical Variable 1	Categorical Variable 2

Identify the question to be answered about the possible relationship between the two variables that you selected above.

Project Question

2. Identify one or more sources that you will use for data on both categorical variables represented in your project question. The data set(s) do not need to contain raw data and can instead provide counts or relative frequencies represented in tables. Only use publicly available, permanent data sets from credible sources that support a meaningful investigation.

Examples of relevant sources from the U.S. Census Bureau include data tables on:

- **Age & Sex**
- **Educational Attainment**
- **Hispanic Origin**
- **Income**
- **Poverty**
- **Race**
- **Voting & Registration**

With your partner, make a list of potential sources you could use for relevant data for both categorical variables represented in your project question.

<ul style="list-style-type: none"> ■ ■ ■ ■
--

PART B: DESCRIBING PATTERNS AND ASSOCIATIONS IN VOTING DATA

Apply appropriate data analysis techniques to answer your question. This should include:

- constructing contingency tables to represent counts and/or calculating relative frequencies,
- constructing graphical representations, and
- examining the data for an association between the two variables.

PART C: IDENTIFYING GENERALIZATIONS AND DRAWING CONCLUSIONS ABOUT VOTING DATA

1. Use the numerical and graphical representations you constructed to identify appropriate generalizations and draw conclusions from the investigation. Your conclusion should:
 - provide an answer to the question you identified,
 - identify the extent to which you can generalize your conclusions to a broader population, and
 - describe whether you can establish a causal relationship between the two variables.
2. Create an infographic handout or digital poster to communicate your conclusions in a clear and articulate manner to the broader audience of the students in your school. You may use the provided template as a guide in structuring your infographic handout or digital poster.

PART D: STUDENT PRESENTATIONS

Present your infographic handout or digital poster to an audience of your peers. Use your conclusions to leave your audience with a lasting takeaway or invoke a call to action to increase voter registration and turnout within your community. If you are presenting outside of your statistics class, ensure your use of course content is communicated appropriately in order to effectively reach your audience.

Project Question

Data Sources & Graphical Representation(s)

Conclusions & Call to Action

THIS PAGE IS INTENTIONALLY LEFT BLANK.

Civic Knowledge & Action: Voter Registration

STUDENT HANDOUT 1: TEACHER KEY

Describing Voting Data Represented in a Table

In September 2018, the U.S. Census Bureau released a report entitled "Characteristics of Voters in the Presidential Election of 2016." The following table is a subset of a larger table from the report.

Registration and Voter Characteristics: 2016

(Numbers in thousands)

Characteristic	Registered	Percent	Voted	Percent
Total	157,596	100.0	137,537	100.0
Age				
18 to 29	27,405	17.4	21,620	15.7
30 to 44	36,129	22.9	30,933	22.5
45 to 64	57,394	36.4	51,668	37.6
65 and older	36,667	23.3	33,314	24.2
Sex				
Male	73,761	46.8	63,801	46.4
Female	83,835	53.2	73,735	53.6

Note: Numbers are based on weighted reports of voting behavior derived from a survey sample.

Source: U.S. Census Bureau, Current Population Survey, November 2016.

<https://www.census.gov/content/dam/Census/library/publications/2018/demo/P20-582.pdf>

- Examine the table above and the data it represents. List two things that you "notice" and two things that you "wonder" about these data and/or the way in which they are displayed.

Notice	Wonder
<p>Responses will vary.</p> <p>Sample responses are listed below:</p> <ul style="list-style-type: none"> Each count is represented in thousands of people. The 45 to 64 age group is the greatest percentage of Registered and of Voted. 	<p>Responses will vary.</p> <p>Sample responses are listed below:</p> <ul style="list-style-type: none"> Why are the percentages conditioned on Voted and Registered instead of Age and Sex? How many individuals within each category were not registered to vote in 2016?

- Interpret in context of the value 17.4 from the table.

The value 17.4 represents that 17.4% of individuals who were registered to vote in the 2016 U.S. presidential election were in the 18 to 29 age group.

3. A student examining these data believes that there were 27,405 thousand people age 18–29 who **registered but did not vote** in the 2016 U.S. presidential election. Explain the error in the student's claim and how the way in which the data are displayed might have contributed to the student's misconception.

The value 27,405 represents that there were 27,405 thousand (or 27,405,000) individuals in the 18 to 29 age group who were registered to vote in 2016. The student did not recognize that this value includes both registered individuals who voted in the 2016 election and registered individuals who did not vote in the 2016 election. Since 27,405 is from the "Registered" category and not from the "Voted" category, the student may have misunderstood "Registered" to indicate individuals who were registered but did not vote.

4. What additional data are needed in order to determine the percentage of each age group (or sex) who registered for and voted in the 2016 U.S. presidential election?

We would need the total number of people within each category of age group (or sex) who are eligible to register to vote in the 2016 U.S. presidential election (i.e., the number of U.S. citizens, either native or naturalized, who are 18 years or older before the election in 2016).

Identifying Information from Voting Data and Constructing Tables to Compare Two Variables

While each U.S. citizen (native or naturalized) who is 18 years or older has the constitutional right to vote, each individual must register to vote within their home state. Each individual in the population can thus be classified by *voter registration and turnout* as: not registered for, registered for but did not vote, or registered for and voted in the 2016 U.S. presidential election.

In this exercise, we wish to examine the role that the variables *age group* and *sex* have on *voter registration and turnout* in the 2016 U.S. presidential election.

1. Identify the two questions that we wish to answer about possible relationships in these data.

- Is there an association between voter registration and turnout and age group in the 2016 U.S. presidential election?
- Is there an association between voter registration and turnout and sex in the 2016 U.S. presidential election?

Use the data available in **Student Handout A: Voting Data Tables** to complete the following tasks.

2. Construct a contingency table to compare *voter registration and turnout* across *age group*.

Age Group	Not Registered (thousands)	Registered, Did Not Vote (thousands)	Registered, Voted (thousands)	Total
18 to 29	20,071	5,785	21,620	47,476
30 to 44	16,693	5,196	30,933	52,822
45 to 64	20,200	5,726	51,668	77,594
65 and older	9,275	3,353	33,314	45,942
Total	66,239	20,060	137,535	223,834

3. Construct a contingency table to compare *voter registration* and *turnout* across sex.

Sex	Not Registered (thousands)	Registered, Did Not Vote (thousands)	Registered, Voted (thousands)	Total
Male	34,074	9,960	63,801	107,835
Female	32,165	10,100	73,735	116,000
Total	66,239	20,060	137,536	223,835

As the total counts in each category of the *age group* and sex variables vary, relative frequencies must be used to compare the distribution of *voter registration* and *turnout* across the categories.

4. Construct a new contingency table to compare the conditional relative frequencies of *voter registration* and *turnout* across each category of *age group*.

Age Group	Not Registered (percent)	Registered, Did Not Vote (percent)	Registered, Voted (percent)	Total
18 to 29	42%	12%	46%	100%
30 to 44	32%	10%	59%	100%
45 to 64	26%	7%	67%	100%
65 and older	20%	7%	73%	100%
Total	30%	9%	61%	100%

5. Construct a new contingency table to compare the conditional relative frequencies of *voter registration* and *turnout* across each category of sex.

Sex	Not Registered (percent)	Registered, Did Not Vote (percent)	Registered, Voted (percent)	Total
Male	32%	9%	59%	100%
Female	28%	9%	64%	100%
Total	30%	9%	61%	100%

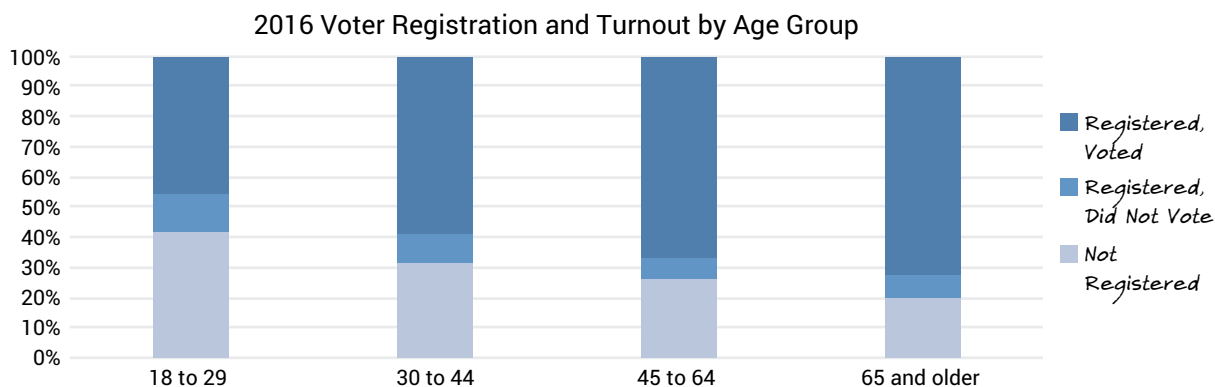
Civic Knowledge & Action: Voter Registration

STUDENT HANDOUT 3: TEACHER KEY

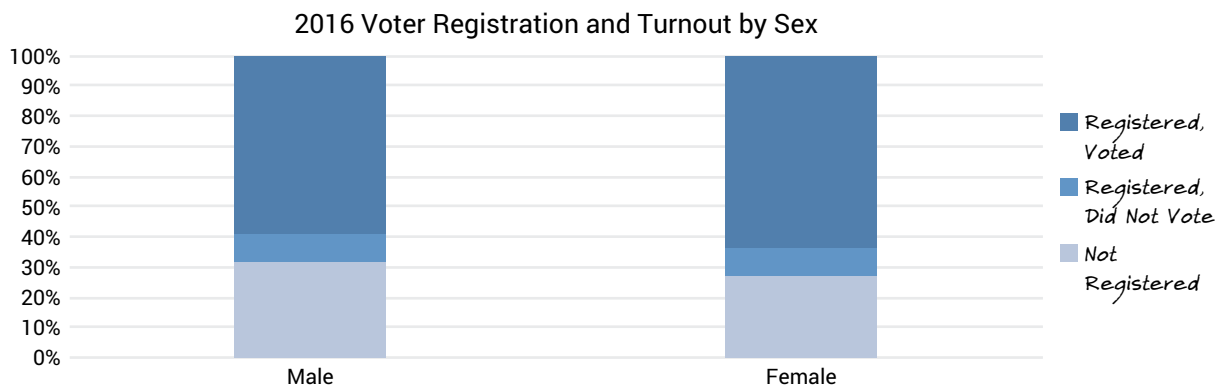
Constructing Graphs to Describe Associations in Voting Data

Use the tables constructed in **Student Handout 2: Gathering Voting Data and Constructing Tables to Compare Two Variables** to complete the following tasks.

1. Construct a segmented bar graph to compare *voter registration* and *turnout* across *age group*.



2. Construct a segmented bar graph to compare *voter registration* and *turnout* across *sex*.



3. Does there appear to be an association between *voter registration* and *turnout* and *age group*? Use evidence from the segmented bar graph to justify your answer.

Yes, there appears to be an association between voter registration and turnout and age group because the distribution of voter registration and turnout is different for the 18 to 29 group, with 42% not registered, 12% registered but did not vote, and 46% registered and voted, compared to that of the other age groups. While the distribution of voter registration and turnout differs from one age group to another, the most significant difference in the distribution appears between the 18 to 29 and 65 and older age groups.

4. Does there appear to be an association between *voter registration* and *turnout* and *sex*? Use evidence from the segmented bar graph to justify your answer.

No, there does not appear to be an association between voter registration and turnout and sex because the distribution of voter registration and turnout is relatively the same for males, with 32% not registered, 9% registered but did not vote, and 59% registered and voted, compared to that of females, with 28% not registered, 9% registered but did not vote, and 64% registered and voted. This is best visualized in the graph as the bars for both sexes appear about the same.

Civic Knowledge & Action: Voter Registration

STUDENT HANDOUT 4: TEACHER KEY

Describing Voting Data Collection Methods and Identifying Appropriate Generalizations

1. Each row of the table below provides an example of a population and a sample from that population. Read each example and use them to write a definition, in your own words, for the two terms at the top of the table.

Population: <i>All items or subjects of interest</i> <i>Responses will vary.</i>	Sample: <i>A subset of the population</i> <i>Responses will vary.</i>
All Ohio residents	All registered voters in Ohio
All individuals in the U.S. population	All individuals in the U.S. who are eligible to register to vote
All individuals in the U.S. who are eligible to register to vote	All native U.S. citizens
All individuals who were registered to vote in the 2016 U.S. presidential election	All individuals who voted in the 2016 U.S. presidential election

2. Do you think each sample above is *representative* of the corresponding population? Explain why or why not.

Responses will vary.

Sample response: No, each sample is not likely to be representative of the corresponding population since there are subgroups of each population not represented within each sample. For example, the sample of all registered voters in Ohio excludes residents of Ohio who are not registered to vote.

3. When using data from a sample to make generalizations about a population, why is it essential for the sample to be *representative* of the population?

Responses will vary.

Sample response: If using data from a sample to make generalizations about a population, the sample should represent the various subgroups of the population as best as possible. For example, residents of Ohio who are not registered to vote may have differing views, opinions, or perspectives than Ohio residents who are registered. By excluding them in the sample, we may make inappropriate generalizations about the population of all Ohio residents.

4. Read the following example of an observational study and an experiment. Use the example to write a definition, in your own words, for the two terms at the top of the table.

Observational study: <i>A study in which data from participants or subjects are observed without the imposition of different conditions (treatments) Responses will vary.</i>	Experiment: <i>A study in which different conditions (treatments) are assigned to experimental units (participants or subjects) Responses will vary.</i>
Survey each voter in a sample of voters on who they plan to vote for.	Show a political ad to half of the voters in a sample and then survey each voter in the sample on who they plan to vote for.

The data used throughout this lesson were obtained from the U.S. Census Bureau Current Population Survey (CPS) in November 2016. Read the following excerpt from the U.S. Census Bureau report about these data.

The CPS is a monthly labor force survey in which interviews are conducted in approximately 56,000 households across the country... and every November following national elections, the CPS includes supplemental questions about voting and registration. Either self-responses or proxy responses, where a single respondent can provide answers for themselves and all additional eligible household members, are allowed. The voting questions are asked of all U.S. citizens who are 18 years or older, and these eligibility criteria are determined using questions about age and citizenship status asked during the core survey ... Analysis of the Voting and Registration Supplement can vary based on the selected population universe. Typical population universes used by researchers include the total voting-age population (anyone aged 18 and older), the citizen voting-age population (anyone aged 18 and older who is also a citizen), and the registered voter population.

The data in this report are based on responses to the CPS November Voting and Registration Supplements, which survey the civilian noninstitutionalized population in the United States. Voting estimates from the CPS and other sample surveys have historically differed from those based on administrative records, such as the official reports from each state disseminated collectively by the Clerk of the U.S. House of Representatives and the Federal Elections Commission. In general, voting rates from sample surveys like the CPS tend to be higher than official results (Bauman and Julian, 2010; DeBell, et al., 2015; McDonald, 2015; Tourangeau, 2010). Potential explanations for these differences include question misreporting, problems with memory or knowledge of others' behavior, and methodological issues related to question wording, method of survey administration, and survey nonresponse bias. Despite these observed differences between CPS estimates and official tallies, the CPS data remain the most comprehensive data source available for examining the social and demographic composition of American voters in federal elections, particularly when examining broad historical results.

Source: U.S. Census Bureau, Characteristics of Voters in the Presidential Election of 2016, September 2018
<https://www.census.gov/content/dam/Census/library/publications/2018/demo/P20-582.pdf>

5. Were the data from the CPS November Voting and Registration Supplements collected through an observational study or an experiment? Explain your answer.

Since a sample survey was conducted in an attempt to learn about the population from which the sample was taken, the data was collected through an observational study. There was no imposition of different conditions (treatments) on the participants in order to measure a change in response to the survey questions.

6. Is the sample from which these CPS data were collected *representative* of the population of voting-age citizens? Explain your answer.

Yes. While the excerpt mentions that there were differences in actual voting rates from the data in this sample survey, it also states that the data from the CPS remains "the most comprehensive data source available for examining the social and demographic composition of American voters" and "typical population universes" include the "citizen voting-age population."

7. Based on the excerpt from the report, do you believe that it is appropriate to generalize the associations you identified between voter registration and turnout and age group and sex to the population of voting-age citizens?

Yes. Since the data used to compare voter registration and turnout with age group and sex were from the CPS and the CPS data were representative of the population of voting-age citizens, then the associations can be generalized to the population of voting-age citizens.

8. A student examining these data makes the following claim:

"A person's older age causes them to be more participatory in voting."

- a. Provide a reason why this student's claim is not valid.

There are many variables, other than a person's age, that could contribute to the differences in the distribution of voter registration and turnout. While there appears to be an association between age group and voter registration and turnout, that does not imply that there is a causal relationship between the two variables.

- b. How could the student's claim be modified to represent an appropriate generalization of the data?

Responses will vary.
Sample response: The student's claim could be modified to say, "Individuals in older age groups tend to be more participatory in voting than those in younger age groups."