Investigating Immigration to the U.S.
Module Overview and Sample Lessons

The Investigating Immigration to the U.S. module focuses on describing, comparing, and making sense of categorical variables. Students investigate questions such as: Are there more immigrants in the U.S. today than in previous years? Where have immigrants to the U.S. come from, now and in the past? Are immigrants as likely as the U.S. born to be participating in the labor force, after adjusting for education?

This module was developed for 12th grade non-AP mathematics and statistics courses and contains seven lessons and one final team data investigation. The module is designed for two to three weeks of instruction.

This sample document contains 1) an overview of the module lessons and learning objectives, 2) the teacher guide for Lesson 5, titled Where have most immigrants been coming from?, and 3) the team data investigation.
Investigating Immigration to the U.S.

Lesson Overview and Learning Objectives
Investigating Immigration to the U.S. Module
Lesson Overview and Learning Objectives

1. **Who are immigrants in the U.S.?**
   This lesson introduces the overarching themes of the module: Who has been immigrating to the U.S., and how have immigrants shaped U.S. society over time? The lesson also introduces some myths about immigrants that the module seeks to dispel through data analysis. This lesson also explores some infographics that show the contributions of immigrants in various sectors of the U.S. economy.

   **Learning Objectives**
   Students will be able to:
   - Identify questions about U.S. immigrants that can and cannot be answered with data.
   - Summarize key information about immigrants that is collected by the U.S. Census Bureau on the American Community Survey (ACS) questionnaire.
   - Explain how the 4-step Data Investigation Cycle can support exploration of questions and help break down myths, misconceptions, and stereotypes about immigrants to the U.S.

2. **What percentage of the U.S. population were immigrants in 2017?**
   In this lesson, students begin to examine the following myth: *The U.S. is being overrun with immigrants like never before.* They do so by examining data from the ACS to determine the percentage of the U.S. population who were immigrants in 2017. The lesson discusses the concepts of sample proportions and percentages; how they are used to analyze categorical variables/attributes; and the approximate margin of error around sample estimates.

   **Learning Objectives**
   Students will be able to:
   - Use a sample proportion or percentage to estimate the population proportion or percentage.
   - Explain why sample proportions will vary from sample to sample.
   - Explain what the approximate *margin of error* means and why it is important to include this calculation when reporting findings.
   - Use 2017 data to comment on whether or not the U.S. is overrun with immigrants.

   **Focus CODAP Skills:**
   Students will learn to:
   - Create a binary attribute from a quantitative attribute.
   - Take screenshots and create links to share graphs.

3. **Are there more immigrants in the U.S. in 2017 than in previous years?**
   This lesson continues to use data to explore the following myth: *The U.S. is being overrun with immigrants.* Students will investigate how levels of immigration to the U.S. have changed over time by examining the percentage of immigrants in the U.S. population in 1920, 1970, and 2017.

   **Learning Objectives**
   Students will be able to:
   - Describe the distribution of a categorical attribute using percentages and proportions.
   - Compare two different percentages and write a conclusion about whether or not the difference is larger than the approximate margin of error.
   - Interpret a timeplot using background information.
Focus CODAP Skills:
Students will learn to:
- Recode a quantitative attribute into a categorical attribute.
- Use CODAP to calculate sample percentages.

4. Where have immigrants to the U.S. come from, now and in the past?
The goal of this lesson is to have students examine the changing world origins of immigrants over time, and to reflect on causes behind these changes. In this investigation, students will gain experience generating and comparing percentages of categorical data from different years.

Learning Objectives
Students will be able to:
- Create and interpret a segmented bar chart that displays conditional percentages.
- Produce conditional proportions and percentages to compare categorical attributes.
- Summarize the relationship between two categorical variables using graphs and numerical summaries.

Focus CODAP Skills:
Students will learn to:
- Use the Fuse Dots feature to compare categorical data at two points in time.

5. Where have immigrants settled in the U.S.?
The goal of this lesson is to have students examine variation in immigrants’ geographic settlement patterns. Through this investigation, students will gain further experience generating and comparing percentages of categorical data.

Learning Objectives
Students will be able to:
- Read and interpret percentages from bar graphs.
- Attend to what is the “part” and what is the “whole” when calculating a percentage.
- Use data to analyze where U.S. immigrants were most likely to live in 2017, and how the settlement patterns of U.S. immigrants compared to those of U.S.-born individuals in 2017.

Focus CODAP Skills:
Students will learn to:
- Choose between displays of row or column percentages to make appropriate comparisons among categorical attributes.

6. Are immigrants as likely as the U.S. born to be employed or looking for a job?
In this lesson, students explore the following myth: Immigrants don't pay taxes and are a drain on society and the economy. To do so, students investigate the differences in labor force participation between U.S.-born individuals and immigrants. Through this investigation, students will gain further experience generating and comparing percentages of categorical data where there are more than two possible categories. They will also gain experience with data cleaning and subsetting.

Learning Objectives
Students will be able to:
- Determine percentages based on what is the “part” and what is the “whole.”
- Interpret percentages, and use them to compare and contrast groups.
- Describe how, if at all, U.S. immigrants differ from U.S. born individuals in 2017 with regard to whether they are employed, looking for jobs, or unemployed.
Focus CODAP Skills: Students will learn to:
● Identify attributes of individual cases in a graph by selecting and finding the cases in the data table.

7. **Are immigrants as likely as the U.S. born to be participating in the labor force, after adjusting for education?**
   This lesson builds directly on Lesson 6 and continues to address the following myth: *Immigrants don’t pay taxes and are a drain on society and the economy.* The lesson aims to develop students’ understanding that the association between two variables may change after accounting for a third variable. Specifically, students investigate what happens to labor force participation rates of U.S.-born individuals and immigrants when adjusting for education levels. Students will also gain further experience generating and comparing percentages from categorical data using attributes with more than two possible categories. They will also gain experience with data cleaning and subsetting.

**Learning objectives**
Students will be able to:
● Determine percentages based on what is the “part” and what is the “whole.”
● Read and interpret percentages from bar graphs.
● Describe what may happen to the association between two attributes when adjusting (or controlling) for a third attribute.
● Describe whether U.S. immigrants are as likely as the U.S. born to participate in the labor force, after adjusting (or controlling) for whether or not people have a college degree.

Focus CODAP Skills: Students will learn to:
● Create multiple graphs (such as in a side-by-side layout) to examine the relationship between two attributes when adjusting (or controlling) for a third attribute.

8. **Team Data Investigations**
Students will work in teams to choose and investigate a new question related to immigrants in the U.S. They will apply their understanding of the 4-step Data Investigation Cycle as well as statistical concepts to address their chosen question. Specifically, they will identify a question to answer, assemble a data set using ACS and/or decennial census data, use graphs and tables to analyze the data, and justify conclusions based on the data. After investigating patterns between two attributes, students will examine how the relationship between two attributes may change when adjusting for a third attribute. Students will be encouraged to share their work and findings with their peers.

Questions for further investigation of immigrants in the U.S.:
1) **What types of occupations are immigrants most likely to hold compared to U.S.-born individuals?**
2) **How does the typical wage of immigrants compare to the typical wage of U.S.-born individuals?**
3) **Are immigrants in 2017 less likely than immigrants in 1980 to speak English well?**

Possible third attributes:
● Sex, race/ethnicity, education, age, marital status, U.S. region, a different year (an option for question 1 only), birthplace (an option for question 3 only).
Investigating Immigration to the U.S.

Lesson 5 Teacher Guide
Lesson Overview
The goal of this lesson is to have students examine the changing world origins of immigrants over time, and to reflect on causes behind these changes. In this investigation, students will gain experience generating and comparing percentages of categorical data from different years.

Learning Objectives
*Students will build the following skills and concepts:*
- Create a graph and calculate numbers to compare percentages across groups.
- Use this information to analyze the association between two categorical variables.

Focus Immigration Claim: In this lesson you will examine data to address the claim: *Most immigrants are Mexican.*

Focus CODAP Skills
*Students will build the following skills:*
- Compute row percentages and column percentages.

Lesson Introduction and Opening Discussion
Use slides to share the lesson overview and objectives. Remind students that the investigation in the previous lesson used ACS data to examine the percentage of the population who were immigrants over time. Today we are going to look at where immigrants have come from now and in the past.

**Step 1: Ask Questions (and Make Predictions)**
*Slide.* The first step in the investigation cycle is to pose a question. In this lesson, the question has already been framed for students. Here is the question for today’s investigation:
- **What were the percentages of U.S. immigrants from different world regions in 2017? How about in 1920?**

In this step, students will make some guesses about the answer. This is a good way to dig into the exploration. Remind students that in this lesson they will continue to use data from the ACS and recall that Question 7 asks “Where is this person from?”
Recall that the ACS asks where an individual was born. The responses to “Outside of the United States” are categorized into (188!) different world regions. We have broken this down into 9 world regions in the data portal:

- U.S. state, territory, or outlying region
- Canada, Australia, New Zealand, or Pacific Islands
- Mexico, Central America, South America, or the Caribbean
- Northern or Western Europe
- Southern Europe, Central/Eastern Europe, or Russia
- East, Southeast, or South Asia
- Middle East or Southwest Asia
- Africa
- Unknown
When students have completed their guesses, take each question at a time and invite students to share their responses in the chat or in discussion. One important idea to help students understand the question: We are not asking where did most people who came in 2017 come from. We are asking, if you think of all of the immigrants who are here in the U.S. in 2017 which regions will have the most immigrant representation?

CODAP Activity

In this lesson, students are provided a dataset of randomly selected immigrants who responded to questionnaires from the U.S. Census Bureau. (They will not create their own dataset from the Data Portal as they did in previous lessons. Make sure students understand that by using the same dataset, everyone in the class will be analyzing the same sample of people.) The dataset includes several attributes. Check in on understanding of “attributes” by briefly hearing from students what attributes they might expect to be included in the dataset. Hopefully, they will mention an attribute that tells where the person immigrated from. Ask students to look for such an attribute when they examine the dataset with their team. Don’t take a lot of time in discussion here.
because it's best to talk about this after they've looked at the dataset. Have students complete through Question 4 and then confirm their responses to questions 3 and 4.

**Step 2: Assemble Data Using CODAP and Step 3: Analyze the Data**

**Step 2: Assemble Data Using CODAP**

For this lesson, we will all use a random sample of immigrants who responded to the ACS survey in 2017 and the decennial census in 1920. Open the data file in CODAP: [Link](#).

Describe your sample:

**Question 3.** How many immigrants are in this sample?  
>> 1092

In this data file, we have immigrants who responded in 2017 and 1920.
  * Determine the number of immigrants that were surveyed in 1920 or in 2017 by creating a graph and dragging the Year attribute to the horizontal axis.  
  * Use the Ruler to show Count.

If you get stuck and need guidance, see the [CODAP Reference Guide](#) for more instructions.

**Question 4.** Examine who is in your sample.
   A. How many people in this sample were surveyed in 1920?  >> 505  
   B. How many people in this sample were surveyed in 2017?  >> 587

*Stop here and be prepared to share your responses.*

**Step 3: Analyze the Data, Questions**

**Slides.** In this step students will analyze the data. Show slide and orient students to the purple box: *Make sense of the data by creating and analyzing graphs. What story is the data telling? Give your graphs titles.*

Have students continue to Step 3, Questions 5 and 6. You can have students work independently and share answers or complete this together as a class and have students give input. Then, guide them through Questions 7-9 and provide support to make sense of percentages.
Step 3: Analyze Data

To evaluate the claim that most immigrants are Mexican, we can compare the percentage of immigrants who were born in “Mexico, Central America, South America, or the Caribbean” to the percentage of immigrants from other world regions. We can also see whether this percentage was much different between 1920 and 2017.

- Drag the Birthplace_recode attribute to the y axis.

First examine the distribution of responses in 2017.

**Question 5.** In 2017, how many immigrants in the sample were born in Mexico, Central America, South America, or the Carribean? >> **294**

Was this the most common birth place in 2017? **>> There were more people with this birth place in 2017 than any other region. (It’s important to note, though, that this region includes several other countries than just Mexico. We don’t have a breakdown of this data by country, so it’s not clear how many are from Mexico and how that compares to other countries in this region and other regions in the world. Since 50% of all immigrants are from this region that includes several countries, it’s safe to assume that most (more than half) immigrants are not from Mexico.)**

Now consider the distribution of responses in 1920.

**Question 6.** Was Mexico/Central America/South America/Carribean still the most common region? **>> No**

If not, what was the most common region? **>> Southern Europe, Central/Easter Europe, or Russia**

Because we do not have the same number of immigrants in this sample each year, it is helpful to convert the “number of immigrants” to percentages in order to compare the distributions for the two years. Let’s see what CODAP does when we ask for percentages.

Use the Ruler to check the Percent box. Note that the “Row” button is selected.

**Question 7.** What percentages are reported for the Mexico, Central America, South America, Carribean category in each year?

<table>
<thead>
<tr>
<th></th>
<th>2017</th>
<th>1920</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico, Central America, South America, Carribean</td>
<td>&gt;&gt; 92%</td>
<td>&gt;&gt; 8.1%</td>
</tr>
</tbody>
</table>

What do these percentages mean? To interpret proportions and percentages it is important to identify the numbers that form the “part” and the “whole.” For these “row percentages” (which, apart from rounding, sum to 100%), the “whole” is the total number of individuals who said they were born in Mexico, Central America, South America, or the Carribean. This total number becomes our denominator.
**Question 8.** What is the total number of responses for this region, combining the two years together? 

>> 320

The “part” is how many responded from each year. This number becomes the numerator. In this sample,

\[
\frac{294}{294+26} \times 100\% = 92\% \\
\text{part} \quad \text{whole}
\]

So 92% of those born in Mexico, Central America, South America, or the Carribean responded to the survey in 2017 and 8% responded in 1920.

**Question 9.** Now consider the Southern, Central, or Eastern Europe row. Determine the percentages that responded in each year (show your work) and complete the sentence.

>>

\[
\text{12}_\% \text{ of immigrants born in Southern, Central, or Eastern Europe responded in } 2017 \text{ and } 88\% \text{ responded in } 1920.
\]

Next, help students understand the terminology “association.” Association between attributes corresponds to differences between groups. Both analyses (row and column percentages) are summarizing the same association, but the interpretation is probably easier for students if we condition on “year” rather than on “region.” In this next section, they will also explore column percentages to condition on region.
Because these percentages (92%/8% vs. 12%/88%) are so different, we say there is an association between birthplace and year. But it might be more clear to describe the association another way.

Use CODAP to change the row percentages to “column percentages.”
- Use the Ruler to select the Column percent rather than the Row percent.

**Question 10.** What two percentages are reported for the Mexico, Central America, South America, Carribean category?

<table>
<thead>
<tr>
<th></th>
<th>2017</th>
<th>1920</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico, Central America, South America, Carribean</td>
<td>&gt;&gt;50%</td>
<td>&gt;&gt;5.1%</td>
</tr>
</tbody>
</table>

**Question 11.** Explore these values to see how to interpret them:

Do these values sum to 100%? >> No

What is the “whole” (denominator) for the values in 2017? (Hint: What numbers do sum to 100%?)

The whole is all of the individuals in the sample who responded in 2017. Or all of the individuals who responded in 1920. Each of the years is the whole.

Show how the percentage for 2017 was calculated: \( \frac{294}{587} \times 100\% \)

**Question 12.** Complete the sentences below to interpret these values.

Of the \(587\) immigrants in the 2017 sample, \(50\) % were born in Mexico, Central America, South America, or the Carribean.

Of the \(505\) immigrants in the 1920 sample, \(5.1\) % were born in Mexico, Central America, South America, or the Carribean.

**Question 13.** Would you say these two percentages from Question 12 are similar to each other or quite different?

>> These are quite different. The 2017 percentage is \(\frac{1}{2}\) of the sample. In 1920, the individuals from this region were much fewer -- less than 10%.

Next, have students work on their own to explore the data and birthplace regions further by completing Question 14. Hear their ideas before moving to the lesson Wrap-Up.
Question 14. Identify another region where the percentage of immigrants born in that region changed substantially between 1920 and 2017. Record the region and the percentages below.

<table>
<thead>
<tr>
<th>Region</th>
<th>Percentage of U.S. immigrants from the region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2017</td>
</tr>
<tr>
<td></td>
<td>1920</td>
</tr>
</tbody>
</table>

Definition: Because the distributions of birthplaces are so different between 1920 and 2017, we say there is an association between birthplace and year. In other words, we have evidence that the pattern of birthplaces of immigrants has changed between 1920 and 2017.

Class Discussion

Slides. When students have finished Question 14, facilitate a whole-class discussion. Possible questions are listed below. You can record results in the slide or on board.

Which world regions were most U.S. immigrants from in 2017? What were the percentages of immigrants that year who were born in these world regions?  
*Note that immigrants may move among multiple countries before coming to the U.S. In this analysis, we considered immigrants’ reported birthplace as where they are from.*

Which world regions were most U.S. immigrants from in 1920? What were the percentages of immigrants that year who were born in these world regions?

*Consider recording students’ responses to both questions above for the class to examine together.*

What changes do you notice in immigrants’ world origins or birthplaces between 1920 and 2017?

*Have students compare the percentage of immigrants from a specific world region in 1920 and 2017. Consider having students calculate the percentage point change vs. the percent change.*

What are you wondering about after analyzing these data?  
How could you learn more to explore these questions?

Lesson Wrap-Up

Step 4: Summarize Conclusions

Slide. Ask students to complete Step 4 of the Data Investigation Cycle. Summarize key learning by revisiting the lesson learning objectives with students. Let students know that in the Lesson Wrap Up they will summarize their conclusion to the claim the class has been investigating: *Most Immigrants are Mexican.*
Step 4: Summarize Your Conclusions

Use the link that your teacher provides to complete the Lesson 5 Google Form that contains the 3 questions below. If you prefer, you can type your answers here and then copy and paste them into the Google Form.

**Wrap-Up Question 1.** Recall the claim that Most immigrants are Mexican. How would you respond to this claim? Use data to support your argument and discuss if there is any additional data you would like to have to further explore the claim.

**Wrap-Up Question 2.** How has the distribution of birth places changed between 1920 and 2017? (Hint: Focus on describing which regions have had bigger changes.)

**Wrap-Up Question 3.** Reflect or re-read sections of the Immigration Timeline article you read in the last lesson. Cite at least one reason that might help to explain the difference in the distribution of birth places between 1920 and 2017 from this article.

**Make sure that you have submitted your responses in your Lesson 5 Google Form at this link.**
Investigating Immigration to the U.S.

Team Data Investigation
Team Data Investigation: Investigating Immigration to the U.S.

Overview
Work as part of a team to apply your understanding of the Data Investigation Cycle and statistical concepts by completing the cycle for a new question using ACS and decennial census data. You will identify a question, assemble a data set, analyze the data, and draw conclusions. You will share your work and findings with your peers.

Step 1: Ask a Question

Review the suggested list of questions below. Choose a question from Part A that interests you. Then choose a third attribute from Part B to add to your analysis.

Part A. Choose one of the bolded questions below to investigate:
- Question 1. What types of occupations are immigrants most likely to hold compared to U.S.-born individuals?
- Question 2. How does the typical income of immigrants compare to the typical income of U.S.-born individuals?
- Question 3. Are immigrants in 2017 less likely than immigrants in 1980 to speak English well?

Part B. Choose a third attribute to extend your analysis. In particular, you will explore how your findings change when you adjust (or control) for this third attribute.
- Sex
- Race/ethnicity
- Education
- Age
- Marital status
- U.S. region
- Birthplace (Option for Question 3 only)

1. List your question below.
2. List the third attribute you will investigate.

3. Make some predictions about what you expect to see in the data.

4. How do you expect the results will differ when you consider the third attribute?

**Step 2: Assemble Data**

If you chose **Question 1 or 2** in Part A above, you will use this dataset. 
If you chose **Question 3** in Part A above, you will use this dataset.

5. What is the sample size?

6. List the attributes (variables) included in the dataset

7. List the attributes you plan to use in your analysis. Reference the data code book and include a description of the attribute.

8. Will you clean the data at all? Record notes here on what you do so that you can describe your actions in your conclusion summary.

9. Which attributes might you want to recode? Why? Record notes here on what you do so that you can describe your actions in your conclusion summary.

**Step 3: Analyze Data**

10. Make graphs to address your first data investigation question (Part A).
    - Display appropriate counts and percentages on the graph or create a table to display the counts and/or percentages.
    - Give each of your graphs a title.
    - Describe any data-cleaning (setting aside or hiding) that you did and why.

11. Insert screenshots of your graphs and paste a link to CODAP below.
12. What story do you think is told by the data? What patterns do you see?

Next, you will create and analyze graph(s) that incorporate the third variable you have selected.

Step 1: Ask a Question

13. Pose your question(s) involving the third variable here:

14. What do you expect to see? Make some predictions about how the graphs will look or what relationship you expect you might find?

Step 2: Assemble Data

15. Will you recode the third variable? If so, how?

Step 3: Analyze Data

- Create new graphs to investigate your question with the third attribute. Make sure you keep your prior graphs and update your CODAP link to keep all of your graphs.
- Display appropriate counts and percentages on the graphs or create a table to display the counts and/or percentages.
- Give your graphs titles.
- Read Guidance for the Third Attribute to get more information on how to work with your third attribute.

16. Describe your new graphs and insert screenshots of them. Update your CODAP link.

Step 4: Summarize Conclusions

17. What story do you think is told by the graphs? What could help explain the relationships you see among variables or other patterns in the data?

18. Pose at least one new question for future investigation.