SDLC Strengthening Data Literacy across the Curriculum

Investigating Income Inequality in the U.S.

Module Overview and Sample Lessons

The *Investigating Income Inequality in the U.S.* module focuses on describing, comparing, and making sense of quantitative variables. Students deepen their understanding of this content by investigating questions such as: How have incomes for higher- and lower-income individuals in the U.S. changed over time? How much income inequality exists between males and females in the U.S.? Does education explain the wage gap between males and females?

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 4, titled *How have middle-income earners in the U.S. been doing over time?*; 3) the teacher guide for Lesson 6, titled *How much income inequality exists between males and females in the U.S.?*; and 4) the team data investigation.

Education Development Center, Inc.
Waltham, Massachusetts

California Polytechnic State University
San Luis Obispo, California

The Concord Consortium
Concord, Massachusetts

National Science Foundation
This material is based upon work supported by the National Science Foundation under Grant #1813956. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Investigating Income Inequality in the U.S.

Lesson Overview and Learning Objectives

SDLC Strengthening Data Literacy across the Curriculum

Investigating Income Inequality in the U.S. Module

Lesson Overview and Learning Objectives

1. **What is income inequality?**
   The activities in this opening lesson of the module are designed to engage students in the topic of income inequality and to introduce them to the Data Investigation Cycle.

   Learning Objectives
   Students will be able to:
   - Pose questions about income inequality and identify which ones can be answered with data.
   - Identify the four steps of the Data Investigation Cycle.

2. **How do we learn about people's incomes in the U.S.?**
   The activities in this lesson are designed to engage students in exploration of the data source of the module: The U.S. Census Bureau's American Community Survey. Students will discuss strengths and weaknesses of the data and connect tables and graphs of data from the survey, displayed using CODAP.

   Learning Objectives
   Students will be able to:
   - Give a definition of income as measured by the ACS and examples of other ACS variables (attributes).
   - Describe some strengths and weaknesses of an ACS data sample.
   - Describe how the variables (attributes) in a data set relate to the ACS questionnaire.
   - Describe how data in a data table relate to data in a graph.

3. **What was the average income in the U.S. in 2017, and how accurately can we estimate it?**
   In this lesson, students will use sample data from the ACS to estimate the population mean. CODAP will calculate the mean for different samples, and students will consider how increasing the sample size reduces the sample-to-sample variation. The goal is to convince students that a sample of just 1,000 provides a reasonable estimate of the mean of the entire population.

   Learning Objectives
   Students will be able to:
   - Graph and describe what the *distribution* of incomes looked like in the U.S. in 2017.
   - Estimate the average annual income earned from wages for someone in the U.S. in 2017.
   - Explain what sample size can provide a reasonable estimate of the average annual income from wages in the population.

4. **How have middle-income earners in the U.S. been doing over time?**
   In this lesson, students will make some predictions about how average income and the distribution of income has changed in the last 50+ years. Then, they will use CODAP to examine distributions at three points in time, exploring the differences between two measures of center: mean and median. Students will examine how the location of the two measures of center in relation to one another changes based on the skewness of the distribution.

   Learning Objectives
   Students will be able to:
   - Describe the difference between the mean and the median of a distribution.
   - Distinguish between typical income and average income.

- Explain how the shape of a distribution influences the location of the mean and median in relation to one another.
- Explain why the median is less influenced by extreme values than the mean.
  Describe how mean and median incomes in the U.S. have changed over time, and provide implications for this pattern.

5. **How have higher- and lower-income earners in the U.S. been doing over time?**
   In this lesson, students will examine comparison points at different percentiles to develop the practice of examining the features of an entire distribution. Students will then examine how U.S. incomes at different percentiles have changed over time. This lesson prompts students to think about how they might measure income inequality in the U.S. and to assess its scale using tools that students have reviewed in earlier lessons (such as the mean, median, and percentiles of the income distribution).

   Learning Objectives
   Students will be able to:
   - Identify the incomes at different percentiles of the income distribution.
   - Describe how higher- and lower-income earners have been doing over time.
   - Quantify the variability within the U.S. income distribution using the interquartile range (IQR).
   - Describe how measures of center, variability, and shape in the U.S. income distribution have changed over time, and what these changes suggest about income inequality in past decades and today.

6. **How much income inequality exists between males and females in the U.S.?**
   This lesson focuses directly on the theme of income inequality between groups and asks students to investigate: Is there income inequality between males and females in the U.S. To explore this question, students will examine the incomes of males and females in a sample of the U.S. population from 2017 to see how they may differ. They will also begin to speculate about what might explain a gap in incomes by sex.

   Learning Objectives
   Students will be able to:
   - Compare the mean and median incomes for a sample of males and females in the U.S.
   - Compare the variability in incomes for a sample of males and females in the U.S.
   - Estimate the difference in typical incomes between males and females.
   - Summarize a conclusion in response to a question, supported with clear and accurate evidence, including graphical displays.

7. **Does education explain the wage gap between males and females?**
   This lesson continues to explore income inequality among males and females. The data analysis has revealed that typical (median) incomes for males are higher than for females in the U.S. What could help explain this gap? Could differences in education levels between males and females play a role?

   The key idea of this lesson is to consider whether introducing a third variable into the analysis can explain or even change the original association between two variables. Students may have encountered this idea previously in terms of "confounding variables," which are additional variables related to the variables of interest in the study.

Learning Objectives

Students will be able to:

- Describe how a third "confounding" variable may affect the comparison between two variables.
- Analyze the association of two variables across different categories of a third variable.

8. Team Data Investigations

For the team data investigation, students will choose and examine a third variable other than education to see whether the female/male wage gap looks different once we consider this variable. They will also complete all four steps of the Data Investigation Cycle.

Questions for further investigation of income inequality between males and females:

- Could wage disparity between males and females be related **to the number of hours or weeks** that females work as compared to males? (i.e. Are females more likely to be part-time workers?)
- What does the income gap between males and females look like among people of a particular **racial/ethnic background**?
- What does the income gap between males and females look like among people in **by regions or division of the U.S.**?
- What does the income gap between males and females look like among people who have **the same or similar occupation or industry?**
- What does the income gap between males and females look like among people in **a particular age group**?

Investigating Immigration to the U.S.

Lesson 4 Teacher Guide

Investigating Income Inequality in the U.S.
Lesson 4 Teacher Guide

How have middle-income earners in the U.S. been doing over time?

Lesson Overview
In this lesson, students will make some predictions about how average income and the distribution of income has changed in the last 50+ years. Then, they will use CODAP to examine distributions at three points in time, exploring the differences between two measures of center: mean and median. Students will examine how the location of the two measures of center in relation to one another changes based on the skewness of the distribution.

Learning Objectives
Students will be able to:

- Describe the difference between the mean and the median of a distribution;
- Distinguish between typical income and average income;
- Explain how the shape of a distribution influences the location of the mean and median in relation to one another;
- Explain why the median is less influenced by extreme values than the mean; and
- Describe how mean and median incomes in the U.S. have changed over time, and provide implications for this pattern.

Lesson Introduction and Opening Discussion (5 minutes)
How have the middle-income earners in the U.S. been doing over time?

Use **slide** to share lesson objectives.
Consider reviewing and making these points for your students:

- In the last lesson we explored the distribution of income-wages of *employed* individuals in 2017.
- We talked briefly about the behavior of the distribution, including why it makes sense that it is skewed to the right.
- We saw that a sample of 1,000 employed people should give us a reasonable estimate of the population mean income.
- In this lesson, you will investigate how the distribution of incomes has changed over time. We will further explore what measures of center tell us about a distribution, and in particular, how income inequality has changed over time.

Activity: Investigate the Lesson Question (50 minutes)

Step 1: Ask Questions (and Make Predictions) (5 minutes)
The first step in the investigation cycle is to pose a question. In this lesson, the question has already been framed for students, so in this step, students will make some guesses about the answer. This is a good way to dig into the exploration.

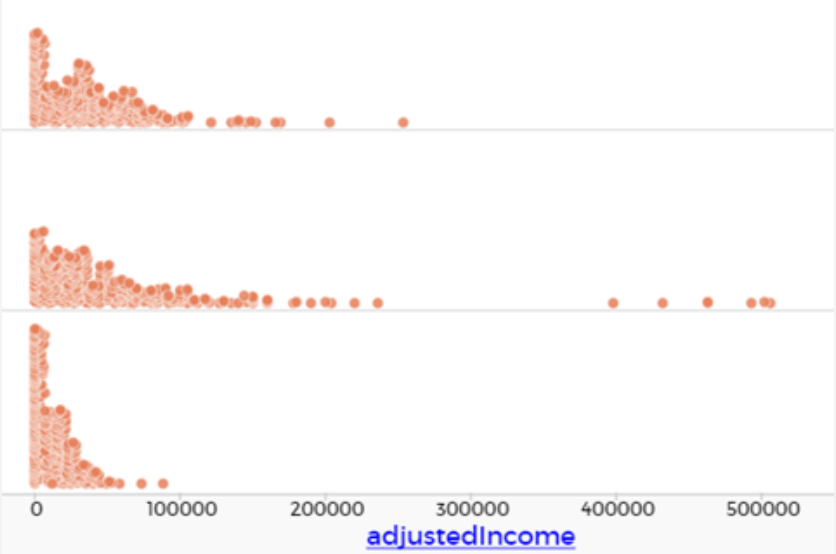
Make sure students record answers for the prediction questions but emphasize that guessing is fine. They will have a chance to revisit their predictions later in the lesson.

Work with your partner to make some predictions. Record your answers in your shared Google doc.

Question 1. PREDICT: How do you think the average income has changed in the U.S. over the last 40 years? (Underline or highlight your answer.)

INCREASED DECREASED STAYED THE SAME

Question 2. PREDICT: Below are three graphs, one of 1940 income-wages that have been adjusted to 2017 dollars, one of 1980 income-wages adjusted to 2017 dollars, and one of 2017 income-wages. Identify which graph you think belongs to which year and explain your reasoning. Record the year of your choice, 1940, 1980 or 2017, in the text box next to each graph.



Year:

Year:

Year:

Stop here for instructions from your teacher.

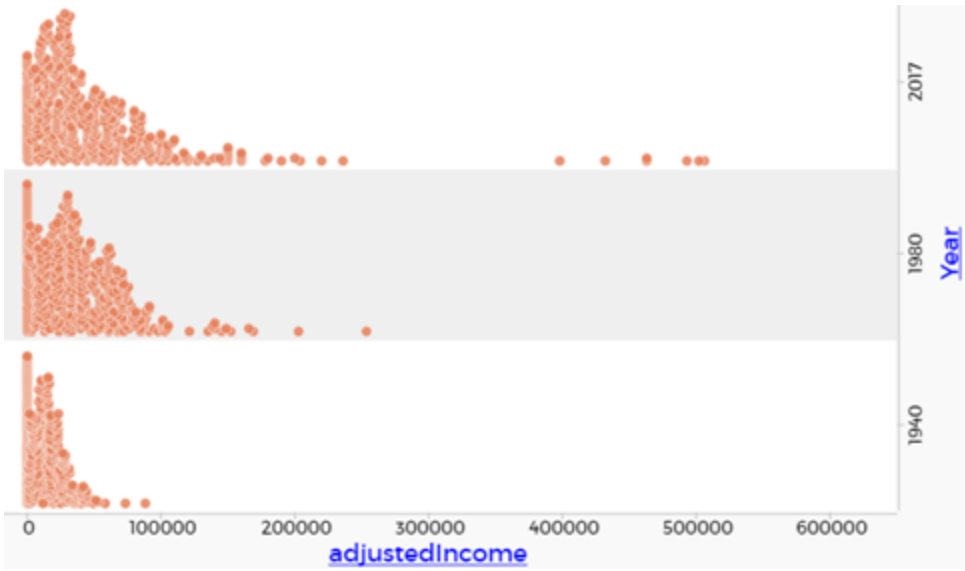
When students have completed their predictions, assemble the class to follow along with the next step of the lesson.

Step 2: Assemble Data Using CODAP (10 minutes)

A teacher-guided activity. Students will follow along, working with their partners in CODAP. Students will record their answers in their shared Google doc.

Learn about adjusted income.

Students open the CODAP link.



CODAP shows a graph of the three years, in chronological order. Have students look back at their predictions and explanations in Step 1 to see how they compare. Then invite students to share their ideas and wonderings.

Discussion Questions

- The graphs have been labeled. Do they match your predictions?
- What do you notice about these three graphs? (They may notice that they are all skewed right but the 1940 values are pushed to the left more than in the other years; they may also notice that the horizontal axis is labeled "adjusted income" rather than "income wages."
- To compare incomes in two different years, we need to adjust incomes to reflect dollar values from the same year because the value of a dollar has changed over time. In 1940, $1 was equivalent in purchasing power to $17.51 in 2017. (CPI Inflation Calculator)
- Invite students to find the column titled **CPI2017**. The variable CP12017 provides the multiplier that is available from the Bureau of Labor Statistics to adjust for inflation and to convert dollar figures to constant 2017 dollars. Dollar figures from different years are comparable once they have been converted to constant dollars from a particular year. To calculate an adjusted income, compute:

  Income wages x CPI2017 = income value in 2017 dollars.

- Show students an example. Take the second person in the data set.

  This individual was 45 when she submitted her data to the Census Bureau. She was employed and lived in New York. Her 1940 income wages were $2800. If she worked the same exact role in 2017, would you expect her wages to be higher or lower than $2800?

  We can use the calculator in the top left tool bar in CODAP to calculate her income in 2017 dollars.

  2800 x 17.652 = $49,425.60

  In 2017 dollars, her salary would be $49,425.60, as shown in the adjusted income column in the CODAP data table.

- Conclude that when the wages are adjusted to the same year's value of the dollar, we can compare the graphs to study changes. Have students complete Question 4 with their partners.

Follow along with the teacher demonstration. Work with your partner in your shared Google doc.

- Open this link.
  Notice how the year labels shown on the graphs compare to what you predicted. Follow along with your teacher's CODAP exploration.

Question 3. Click on the **Ruler** icon and select **Mean** to display the means for each of the three years. Record the means for each year in the table below.

Year	1940	1980	2017
Mean	*13,125*	*37,862*	*50,168*

Question 4. How did the average adjusted income-wage change from 1940 to 1980 to 2017? Is this what you expected?

Pause here for some class discussion of your data set.

Step 3: Analyze the Data (30 min)

A. Exploring the movable line and the median. (10 min)
Students follow along as you demonstrate the movable line feature of CODAP.

- Discuss and complete the table together to record the % of individuals above the mean in each year.
- Model how to use the movable line to find the midpoint where there are 50% on each side.
- Then as a wrap-up, connect this with the definition of median.

Follow along with the teacher demonstration. Work with your partner to record answers.

- Use the **Ruler** to click the **Movable Value** button and select **Add**.
- Use the **Ruler** to select the **Count** and **Percent** boxes.
- For the year that you are assigned, drag the movable line to match the mean line (as close as you can).

Question 5. What percentage of the incomes are larger than the mean? Record the percentages in the table below.

Year	1940	1980	2017
% above the mean	*47%*	*41%*	*33%*

Tell students it's okay if they couldn't get the lines exactly on top of the mean line. It is easier to get close with a larger graph, so make sure they did enlarge their graph window.

In 2017, the mean has been pulled pretty far to the right of the main clump of the distribution. So what would you say is a *middle* income in 2017? Discuss how if you wanted to know a *typical* income, this mean value might not be as informative as a value that was in the middle, with 50% of incomes on each side. Let's see what that value would be for 2017. It's possible that a student may recognize this as the median (and that's great). Postpone the vocabulary, though, until they have found the value. This will allow students to focus on the concept without worrying about its name.

Determine the middle income for 2017.

- Drag the movable line to the left so that you find the income-wage with 50% of the values to the left and 50% of the values to the right (as close as you can).

Question 6. Report the income value that you find that has approximately 50% of the distribution to its left and approximately 50% of the distribution to its right.

Definition: Another way to characterize the center of a distribution is the *median*. The *median income* is the income that has 50% of the incomes smaller than it and 50% of the incomes larger than it. Because the median is always in the middle, it is often referred to as a *typical* value of a distribution.

Question 7. Check the box on the ruler to get the medians for all three years. Record the medians in the table with the means below and find the differences.

Year	1940	1980	2017
Mean	*13,125*	*37,862*	*50,168*
Median	*10,591*	*32,127*	*32,000*
Difference (Mean-Median)	*2,534*	*5,735*	*18,167*

Question 8. What do you notice about how the difference between the sizes of the mean and median have changed over time?

Please pause here to share your findings with the class.

If you wish, you can have students fill in a class chart of the data or display a student's completed data table for reference during the discussion. Hear students' ideas about Questions 8 and 9.

B. Investigate the influence of outliers on the mean and the median. (15 minutes)

Materials: 11 cards printed double-sided and cut (these cards are based on the individual ACS respondents from Lesson 2.)

- Explain to students that you have 11 cards, each of which has a short description of an individual that includes an income-wage. Read a couple of examples before randomly dealing out the cards to 11 students.
- Ask the class to calculate the median. Students should realize they need to put themselves in order and display the income on their card.
- Record the response ($50,000).
- Ask the class to determine the mean income for these 11 values ($62,636). You can appoint two students to calculate and have them check each other's work. Then get another volunteer to take the mean card and position themselves appropriately. Use a card of a different color to represent the mean.
- Now read the grey card and give it to Upsilon, the individual with the highest income.
    - How will the median change? (It doesn't.)
    - How will the mean change? Have them recalculate. ($82, 545)

The class discusses how Upsilon's raise impacts the mean (make that mean person move) and the median (median person doesn't move). This activity is intended to illustrate how the mean is sensitive to outliers, and gets pulled in their direction, making it a less representative of the typical person in the group when the data have outliers. The median is resistant to outliers.

Alpha	30,000		Alpha	30,000	
Beta	59,000		Beta	59,000	
Delta	83,000		Delta	83,000	
Epsilon	20,000		Epsilon	20,000	
Zeta	130,000		Zeta	130,000	
Eta	35,000		Eta	35,000	
Pi	36,000		Pi	36,000	

Upsilon	131,000		Upsilon	350,000	
Xi	65,000		Xi	65,000	
Tau	50,000		Tau	50,000	
Phi	50,000		Phi	50,000	
	62,636	**MEAN**		82,545	**MEAN**
Epsilon	20,000		Epsilon	20,000	
Alpha	30,000		Alpha	30,000	
Eta	35,000		Eta	35,000	
Pi	36,000		Pi	36,000	
Tau	50,000		Tau	50,000	
Phi	**50,000**	**MEDIAN**	**Phi**	**50,000**	**MEDIAN**
Beta	350,000		Beta	350,000	
Xi	65,000		Xi	65,000	
Delta	83,000		Delta	83,000	
Zeta	130,000		Zeta	130,000	
Upsilon	131,000		Upsilon	131,000	

Have students work in pairs or teams to complete questions 9 and 10.

After the card activity, work with your partner to complete Questions 9 and 10.

Question 9. Explain why, for the 2017 income distribution, the median is smaller than the average (mean) income.

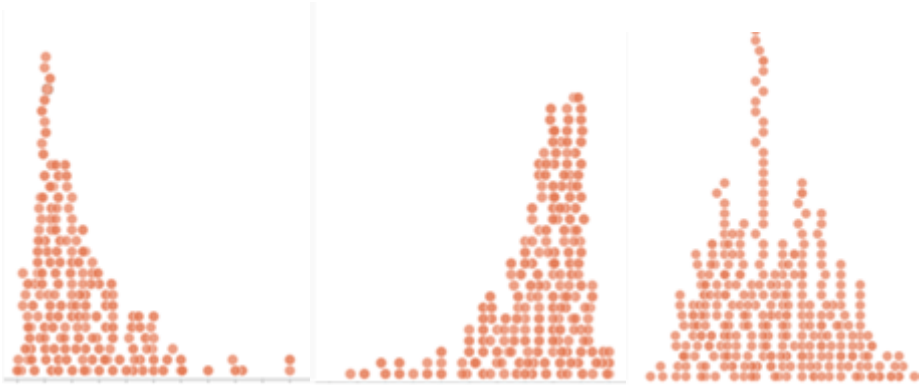
Question 10: Describe what the shape of a distribution would look like if the median is larger than the mean. Use the draw feature to sketch it if you want.

Please stop here to share your ideas with the class.

Discuss how, through the averaging of every single income-wage value, the mean was pulled to the right by the longer right tail and the high outliers.

Class Discussion

The shape of the distribution often predicts the relationship between the mean and median. Show students the slide with the following graphs and ask them to predict which will be larger, the mean or the median.

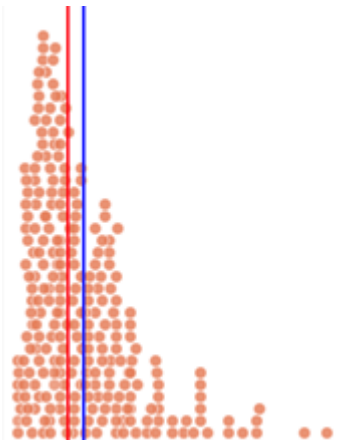


Skewed to the right | Skewed to the left | Symmetric

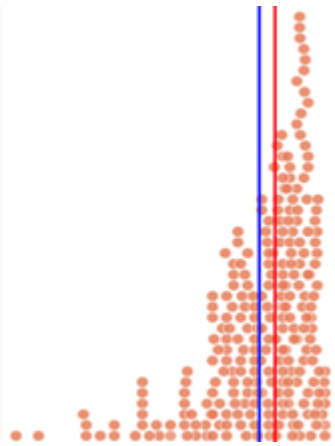
Talk about what variables (attributes) might be represented in these graphs.

- The left graph could be income.
- The graph on the right could be age.
- The graph in the middle could be age at retirement.

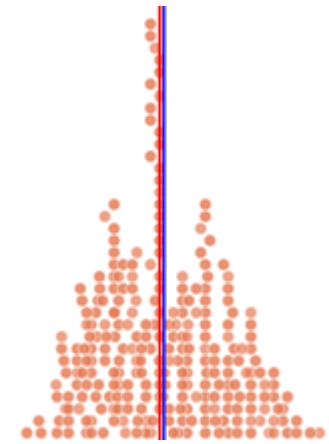
You can then have students try to picture the location of the mean and median in each graph. They may imagine that the mean and median will be quite similar when the distribution is symmetric and the mean will be pulled out toward the tail when there is skewness.



Mean > Median



Median > Mean



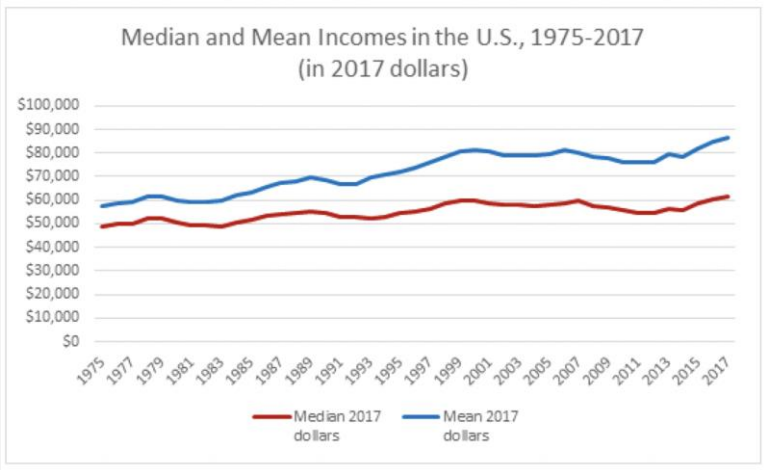
Mean = Median (approx.)

Key Idea: An advantage to using the median as the measure of center is it will *always* be in the middle of the observations; whereas the mean will be pulled in the direction of the skewness of a distribution or towards extreme observations. (Though the impact of a few extreme observations won't be as noticeable when the sample size is large.)

The median income will often be reported, especially for skewed data, because the median may represent a more typical income. Mean represents an average income.

C. Interpreting a graph of mean and median over time. (5 minutes)
Display the slide of Figure 1 and refer students to the graph in their Google doc if they prefer to look at it there.

Figure 1



Think about these discussion questions:

- What is represented by the values on the horizontal axis?
- What is represented by the values on the vertical axis?
- How would you define "2017 dollars"?
- What does this graph tell you about how the income distribution has changed over time?

Step 4: Summarize Conclusions (Exit Task) (5 minutes)

Have students complete individually. If time allows, you could have them discuss first with a partner or small group.

Question 1. Reflect on your predictions in Step 1, Questions 1 and 2 and on what you have learned from your own data and from the Figure 1 graph that we discussed as a class. Summarize what you have learned by writing a response to the lesson question below.

How have middle-income earners in the U.S. been doing over time?

Investigating Income Inequality in the U.S.

Lesson 6 Teacher Guide

Investigating Income Inequality in the U.S.
Lesson 6 Teacher Guide

How much income inequality exists between males and females in the U.S.?

Lesson Overview
This lesson focuses directly on the theme of income inequality between groups and asks students to investigate: Is there income inequality between males and females in the U.S. To explore this question, students will examine the incomes of males and females in a sample of the U.S. population from 2017 to see how they may differ. They will also begin to speculate about what might explain a gap in incomes by sex.

Learning Objectives
Students will be able to:

- Compare the mean and median incomes for a sample of males and females in the U.S.;
- Compare the variability in incomes for a sample of males and females in the U.S.;
- Estimate the difference in typical incomes between males and females; and
- Summarize a conclusion in response to a question, supported with clear and accurate evidence, including graphical displays.

Lesson Introduction and Opening Discussion (5 minutes)

Use **Slides** to share lesson overview and objectives.

Activity: Investigate the Lesson Question (40 minutes)

Step 1: Ask Questions (and Make Predictions) (7-8 minutes)

The first step in the investigation cycle is to pose a question. In this lesson, the question has already been framed for students, so in this step, students will share some initial ideas. Remind students to explain what information helped inform their answers.

Work with your partner to make some predictions about the incomes of males and females. Record your answers in your shared Google doc for questions #1-3.

Question 1. Who do you think tends to make more money in the U.S., males or females? How much more on average? Provide your best guess, and explain what information has informed your answer.

Question 2. Make a guess about the variability (or spread) of the incomes of males and females. Choose (highlight or underline) from the statements below:

- ❑ **Males have more variability (or spread)** in their incomes than females.
- ❑ **Females have more variability (or spread)** in their incomes than males.
- ❑ The **variability (or spread) in incomes is about the same** for males and female.

Question 3. Briefly explain the reasoning you used to make your guess.

Stop here for instructions from your teacher.

When students have completed their predictions, you can have them share briefly with another team.

Step 2: Assemble Data Using CODAP (5 minutes)

This lesson uses the sample of employed individuals of all workers in the U.S. in 2017.

Have students complete the Step 2 question. Let them know to continue on to Step 3: Analyze the data.

- Open the data set.

Question 4. Write 2-3 sentences to describe the data set including the number of males and females in the data set and overview of what variables (attributes) are included.

Step 3: Analyze the Data (20 minutes)

Have students continue on to Step 3.

A. Compare the centers.

Step 3: Analyze the Data

To describe and compare the incomes of males and females, start with graphs that separate the incomes of males and females, but compare them on the same scale.

- Create a graph with the attribute **Income_wages** on the horizontal axis and **Sex** on the vertical axis.

Question 5. Describe the shapes of the income distributions for males and females. Be sure to note the ways they are similar and different.

Next, we want to use a measure of center to compare incomes for males and females in the U.S. in 2017. Use the **Ruler** to check the **Mean** and **Median** boxes. This will add the lines to each graph.

Compare the centers.

Question 6. Report the means for the means and females. Compute the difference between the mean income for males and females and write a sentence to summarize the comparison.

Male mean: $60,008 Female mean: $41,466 Difference: $18,542

Question 7. Report the medians for the means and females. Compute the difference between the median income for males and females and write a sentence to summarize the comparison.

Male median: $38,000 Female median: $30,000 Difference: $8,000

Question 8. Another way to compare these values is by computing the ratio.

Calculate the ratio of female mean income to male mean income and complete the sentence below:

For every dollar that a male makes, on average, a female makes $ ___0.69_____.

Hint: Express this ratio as $\frac{mean\ female\ income}{mean\ male\ income}$ a decimal to determine what a female earns compared to one dollar earned by a male.

Question 9. Calculate the ratio of female median income to male median income and complete the sentence below:

For every dollar that a male makes (typically), a female makes $ __0.78___.

Hint: Express this ratio $\frac{median\ female\ income}{median\ male\ income}$ as a decimal to determine what a female earns compared to one dollar earned by a male.

Questions for Discussion:

- What did you learn about the mean incomes of females as compared to males?
- What did you learn about the median incomes of females as compared to males?
- How did you calculate to determine how much money a female earns for each dollar a male makes?
- What does it say to you that there is more disparity when you compare the mean incomes?

B. Compare the variability.

Next, you will investigate the variability of male and female incomes by determining the Interquartile Range (IQR) for each group. Have students complete questions 10 and 11.

Compare the variability.

- Use the **Ruler** to select the **Boxplot**. This will add a boxplot to each graph. If you hover over the ends of the boxes, you will find the 75th percentile (or "upper quartile" which CODAP labels "Q3") and the lower 25th percentile ("lower quartile", "Q1").

Question 10. Find these values for the males and females. Use these values to compute the interquartile range (Q3 – Q1).

Males Q3: $70,000 Males Q1: $20,000 Difference: $50,000

Females Q3: $52,000 Females Q1: $12,000 Difference: $40,000

Question 11. According to the interquartile range, which group (males or females) appears to have more variability in their income-wages? Males have a higher IQR and more variability in their income-wages.

Pause here for class discussion about the chart.

Have students share their calculations in Question 10 and response to Question 11. Then direct students to complete Step 4: Summarize Conclusions

Step 4: Summarize Conclusions (7-8 minutes)

Summarize what you have learned about how the income distributions compare for males and females. In particular, use information from your answers in Step 3 to answer the following questions.

Question 12. Who tends to make more money in the U.S.: males or females? How much more? (Be sure to say whether you are comparing the means or medians and describe your calculations.)

Question 13. Use evidence (including graphs and calculations from Step 3) and reasoning to summarize and explain your conclusions.

Question 14. Describe how your findings compare to your initial guess in Step 1.

Lesson Wrap-Up (10 minutes)

Summarize key learning by revisiting the lesson learning objectives with students. Focus on a couple of the lesson objectives and invite students to share their ideas and examples.

Final Activity: Exit Task (5 minutes)

Have students complete the exit task individually.

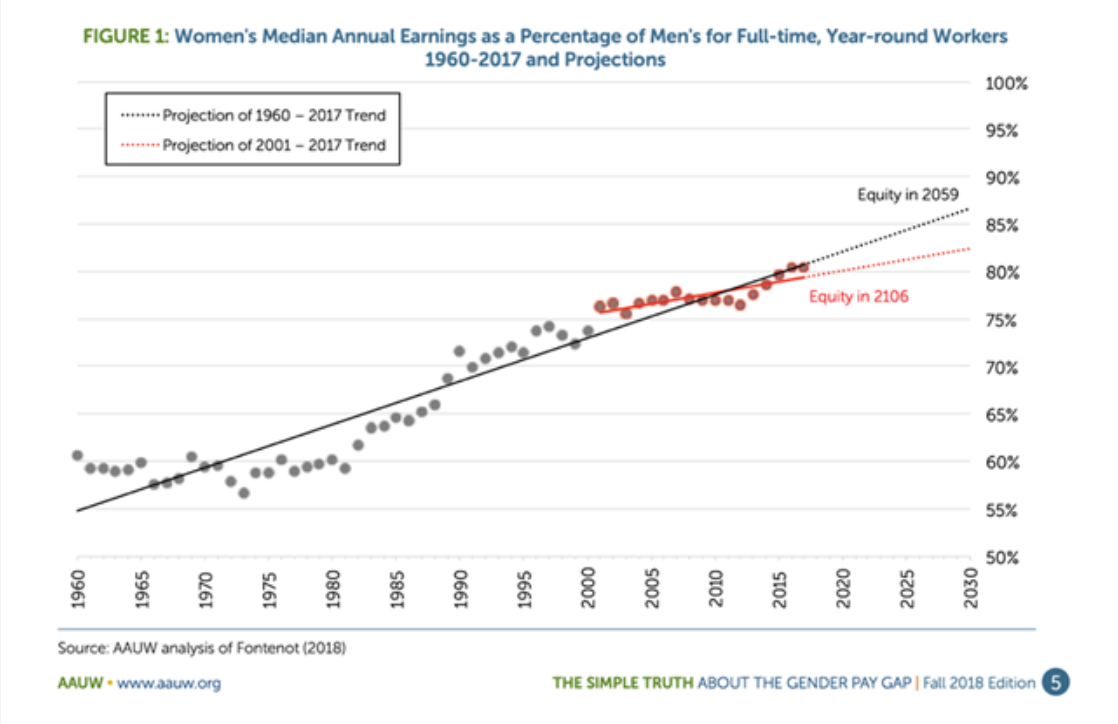
Exit Task

A final part of Step 4 of the data investigation cycle is to consider what new questions have arisen as a result of your data analysis. Here is a new question for you to answer.

Question 1. What might explain differences in income-wages and differences in income variability between males and females? Suggest at least 1-2 reasons for each and describe how you think they might impact income.

Income wages:
Income variability:

Extension Question. Examine the graph of Earnings and Projections shown below. What do you notice? What do you wonder?



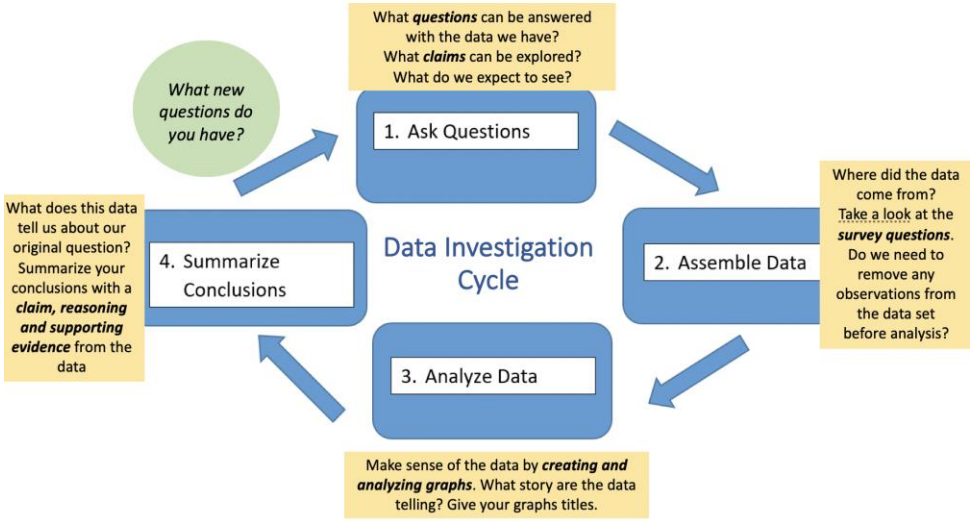
Investigating Income Inequality in the U.S.

Team Data Investigation

Team Data Investigation: Investigating Income Inequality in the U.S.

Overview

Apply your understanding of the 4-Step Data Investigation Cycle and statistical concepts by implementing the process for a new question using 2017 ACS income data. You will choose 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

1. Using the USS codebook, the ACS questionnaire, and the questions listed below, identify a third attribute that you think is related to income-wages and perhaps to differences in income between sexes. You will work together with your partner or team to explore:
    - The distribution of the third attribute
    - How this attribute is related to income-wages
    - Whether or not the female/male wage gap looks different once we consider this attribute

    Sample Questions for further investigation of income inequality between males and females:
    - Could wage disparity between males and females be related **to the number of hours or weeks** that females work as compared to males? (i.e. Are females more likely to be part-time workers?)
    - What does the income gap between males and females look like among people of a particular **racial/ethnic background**?
    - What does the income gap between males and females look like among people in **by different regions of the U.S.**?
    - What does the income gap between males and females look like among people in **a particular age group**?
    - What does the income gap between males and females look like when you look only at those males and females **based on whether or not they have children**?
    - What does the income gap between males and females look like when you look only at those males and females **based on whether or not they have ever been married**?

2. Why did you choose this third attribute? Why do you find it of interest? What are you curious about?
3. Hypothesize about the third attribute. Make a prediction about whether the income- wages gap between males and females is any different when you consider this third variable.

Step 2: Assemble Data

1. Use the **USS Data Portal** to create a data set of 3000 cases that you will use to investigate your question. (In the data portal you can get 1000 cases at a time. Select **Keep existing data** to get 1000 cases two more times.) List the variables (attributes) that you will include in your data set.
2. Create a graph of this attribute and describe the nature of this new variable (e.g., what are the categories, what do the categories mean?) **Note:** You can grab a category to move it/reorder the categories in the graph.
3. Do you want to remove any categories? If so, do so and document your steps. Be sure to save this graph and give it a descriptive title (**Graph1: title**).
4. Before you begin your analysis, be sure to include only those individuals in your data set who are employed. Document your steps so you can report them in your presentation. Remember to **Set Aside** cases in the data window, rather than hiding the cases in the graph.
5. Write a brief description of your data assembly process, including how you chose your attributes and cleaned your data set. Also discuss any limitations to your data set.

Step 3: Analyze Data

1. Using CODAP, produce a graph exploring the relationship between income-wages and your new attribute
2. Save this as **Graph2-[Title]**. Compare the distributions of income-wages (*Hint*: You should compare shapes, centers, and spreads aka variability) across the categories of your new attribute with appropriate numerical support.
3. Using CODAP, produce a graph saved as **Graph3-[Title]** that explores the male/female income-wage distributions (sex on vertical and income-wages on horizontal).* Now add the new attribute (e.g., along the top or right side) and the means, medians, and box plots. Compare the male/female wage age within each category of the new attribute.
4. What story do you think is told by the graphs?
5. What could help explain the relationships you see among variables or other patterns in the data? How might you explore possible explanations?

Step 4: Summarize Conclusions

- Summarize what you learned about how the new attribute is related to income-wages.
- Summarize what you learned about how the male/female wage gap changes across the categories of the new attribute. Would you say the male/female wage gap increases or decreases or stays the same across these categories?
- Pose at least one new question for future investigation.
- Create a short slideshow or report of your investigation to present in a small group. Your presentation should include the slides described below. You are welcome to include additional slides as needed.
- Prepare a short presentation using your slides that you will present to another team.

For Graph 2, include on your slide summary sentences of what the graph reveals about the distribution of income-wage for each category of the third attribute. Include statements about the center, shape, and variability of each distribution.

For Graph 3, include a description of what the gender gap in income-wages looks like for each category of the third attribute. Cite some numbers from your analysis to indicate how much the male-female gap in income-wages changes across categories of the third attribute.