Lesson Title: Factory Bias Analysis

Suggested Time for Lesson: 85 minutes (Part 1 - 25 minutes, Part 2- 50 minutes, Part 3 - 10 minutes)

Lesson involves:
- Collaborative In-class work
- Individual homework

Overview of Lesson and Learning Objectives:
This lesson involves students discussing and actively engaging in real-world applications of data types to better develop their mathematical understanding of qualitative and quantitative data and their graphical representations.
After this lesson, students should be able to:
- Differentiate between and identify qualitative and quantitative data.
- Identify the best representation for various data types.
- Read, interpret, and use bar graphs.
- Understand ratios and proportions.
- Recognize misrepresentation of data.
- Read, create, and interpret scatter plots.

How Lesson Addresses MIP Pillars:
Active Learning
The lesson encourages active learning by scaffolding students through a real life scenario to interpret and create various representations of qualitative and quantitative data. The students can work individually, however, groups are encouraged. The activity will guide towards open discussion where students can communicate with others to answer the open ended questions. The use of technological approaches such as Google Sheets or Excel Sheets will encourage the students to actively engage with previously discussed mathematical concepts.

Meaningful Applications
The lesson provides students the opportunity to work with meaningful applications as they make and justify claims made by others. This requires not only engaging with a real life scenario, but using their mathematical prowess to extract meaning from the scatter plots they are provided and to see how some might try to misrepresent the scenario when using graphs that might be construed as being deliberately misleading.

Academic Success Skills

The lesson allows for productive student engagement as students build on their life experiences and integrate mathematics to make critical decisions that are discussed with their peers. This should establish students as mathematical competent and part of a community of learners who are able to interpret data that is provided to them in a way that should instill confidence in their abilities.

Prerequisite Ideas/Skills for Lesson:
- Recognize and interpret bar graphs
- Understand ratio and proportions
- Classify data types
- Identify data misrepresentations
- Determine best representation methods
- Basic understanding of Google Sheets or Excel

Materials Needed to Carry Out Lesson:
- Factory Bias Slide Show
- Device to access Google Docs, Google Sheets or Excel

Instructional Plan:

The activities are outlined on the Factory Bias Slide Show. The instructor should read this in advance. Even though a slide show is provided, the idea is not to lecture. Instead, the instructor should use the slide show as a visual guide to free time up from writing all the information on the board so that there is more time to circulate about the classroom as students work through the questions in small groups. It is best to establish the small groups right after the initial scenario is presented to allow them to discuss it, as well participate in each embedded task.

Below is an outline of explanations for each question posed in the slide deck. Note that the final problems are intended as individual homework, but it could be assigned as either individual or collaborative in-class work.

The outline below is followed by instructions on how to make scatter plots using Google Sheets.

Part 1: Sam's Concern (Slides 2-13)

To improve student engagement, ask students to first work individually to answer the questions. Then they should discuss and compare their answers in small groups. Furthermore, ask one representative from each group to share their answers with the class.

Question 1 Explanation:

a) The horizontal axis represents the gender of managers. The vertical axis represents the number of managerial positions.
b) The horizontal axis (representing gender) is qualitative data, while the vertical axis (representing the number of managerial positions) is quantitative data.
c) Sam's graph makes it look like females are being favored for managerial positions.
d) Sam's graph is not misrepresented because the vertical axis starts at 12. To create a properly represented graph, we should start the vertical axis at 0.

Question 2 Explanation:

a) Both graphs are from the same data set, but Sam's vertical axis starts at 12 while the union's starts at 0. Only the union's graph is properly represented.
b) There are slightly more female managers than male managers when looking at the union's graph. By manipulating the vertical axis, Sam's graph makes it look like there are significantly more female managers. We need to investigate further to see if Sam's suspicions are correct. Students' answers will vary.
c) We should use a bar graph here because we are comparing two categories of qualitative data (male vs. female). A histogram or line chart is inappropriate here because we are working with qualitative data, not quantitative data.

Question 3 Explanation: Here we want students to consider the male to female ratio of all employees, not just managers. We should take this a step further and investigate the ratio of all applicants to avoid hiring bias.

Question 4 Explanation: $13/30 = 0.43333$ or 43% of managers are male and $17/30 = 0.56667$ or 57% of managers are female.

Question 5 Explanation: Since 40% of the applicants were male, the factory is promoting a nice balance of men and women. Students' conclusion should mention the percentages of male and female applicants as well as the demographics of the region.

Part 2: Sally's Concern (Slides 14-19)

To improve student engagement, ask students to first work individually to answer the questions. Then they should discuss and compare their answers in small groups. Furthermore, ask one representative from each group to share their answers with the class.

Question 6 Explanation:

a) The horizontal axis represents the age of workers in years. The vertical axis represents the yearly salary in dollars.
b) Both axes represent quantitative data. No qualitative data is discussed here.

Question 7 Explanation: According to Sally's data, there might be a legitimate concern. Sally's scatterplot shows that males of all ages do earn more than females. Being older doesn't imply that an employee has worked at the factory longer. We should investigate tenure at the factory rather than age.

To improve student engagement on submissions, ask students to bring a device to class. Give students time to work on creating the scatter plot during class. This will allow you an opportunity to help groups with the technology. This can be given as an in class group work or take home individual assignment.

1. Create and analyze the scatter plot:
    a)

Work experience vs salary

FEMALE SALARY   MALE SALARY

You can find a completed example with sorted data here.
Click this link to find directions to create a scatter plot.

b) Notice that the trendlines almost overlap each other.
c) Sally should not go to the union. Our new scatter plot shows that most employees who have worked at the factory the same amount of time make about the same salary. There are 3 employees that have worked at the factory 22 years: one male earning $42,000 and two females earning $43,500 and $46,500. In this case the females actually earn more than males.

Chart should include a title, labels for the vertical and horizontal axes, and units on the vertical axis ($).

d) A scatterplot is the best method because we can use it to compare two related *discrete quantitative variables* to determine their dependency on each other.

Work experience (in years) & salary ($) are two quantitative variables and, male & female are two distinct qualitative categories discussed in the scenario. When looking

at the problem, the main concept is to understand the relationship between the work experience and the salary. Here salary depends on the work experience. So, work experience is the experimental (independent) variable and salary is the response (dependent) variable. The best representation method to see the relationship between experimental and response variables of related discrete quantitative variables is the scatter plot. Here we use two overlapping scatter plots, distinguished by color to compare the data related to male (purple color) and female (green color). Goal is to see the dependency of the salary with the work experience and compare the trend of male and female data. So, scatter plot is still the best method to represent the data in this scenario.

Note: A bar graph will not work since we are not using qualitative data. A line chart will not work since our data is not continuous.

Part 3: Simon's Concern (Slides 24-25)

Since students do not get much opportunity to choose appropriate tools for analyzing these real world situations, we have created an optional assignment as Part 3. They are led through what graph to consider using a second set of data with different variables. In this part, students take on the responsibility to direct the investigation on their own.

If instructors choose to assign Part 3, they should ask students to bring a device to class. Give students time to work on creating the pie chart during class. This will allow you an opportunity to help with the technology. This can be given as an in class individual work or take home individual assignment.

What you need to submit:
2. Analyze the productivity data:
    a) Shifts represent a qualitative data set while number of cases per shift represent a quantitative discrete data set. However, the scenario described here considers a representation of a qualitative data set.

    b) The correct graph is a pie chart. The shifts represent a part of a day. Since each input represents part of a whole, a pie chart is best.

c)

Productivity in Cases per Shift

Night
20.9%
200

Morning
27.2%
260

Afternoon
18.3%
175
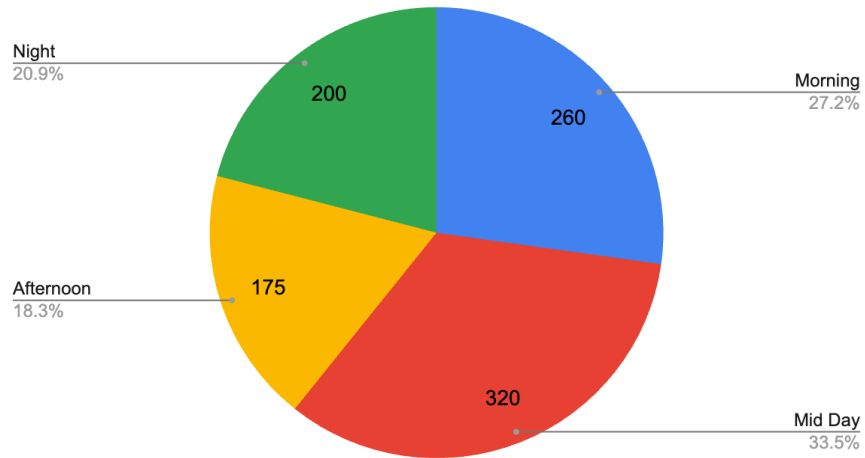
Mid Day
33.5%
320

d) The graveyard shift is not more productive than the other shifts. The mid day shift is the most productive.

3. Students should summarize the conclusions from Sam, Sally and Simon's concerns from the union's perspective. They address a more generalized concern. (This open ended question answers will vary.)