Winter 2020

Playfair's Introduction of Bar and Pie Charts to Represent Data

Diana White
University of Colorado Denver

River Bond

Joshua Eastes

Negar Janani

Playfair's Introduction of Bar and Pie Charts to Represent Data

Diana White, River Bond, Joshua Eastes, Negar Janani

February 12, 2020

1 Introduction

William Playfair (1759 – 1823), a Scottish engineer and political economist, was instrumental in the use of graphics to visualize quantitative data. His accomplishments in this area include inventing numerous different types of diagrams, including the bar graph, line graph, and pie chart.

It is easy for us today to think that these are all trivial things, as we've seen them since our childhood and are accustomed to them. However, someone had to think of them first!

While there were some prior attempts to visualize quantitative data, William Playfair was the first to make extensive efforts in this direction. His efforts were met with mixed acceptance, as many viewed data in tabular format as more precise. As such, he spent some time trying to convince the reader of the utility of his approach, as shown in the following quote from his book *Statistical Breviary* [Playfair, 1801].

> As knowledge increases amongst mankind, and transactions multiply, it becomes more and more desirable to *abbreviate* and *facilitate* the modes of conveying information from one person to another, and from one individual to the many.
>
> Algebra has abbreviated arithmetical calculations; logarithmic tables have shortened and simplified questions in geometry. The studies of history, genealogy, and chronology have been much improved by copper-plate charts. Is it now thirteen years since I first thought of applying lines to subjects of Finance.

Some of his representations were quite complex, as we'll see toward the end of the project. For now, though, let's get started with the bar graph and pie chart.

2 Bar Graph

Playfair is credited with using the first bar graph, which is a means of representing data using vertical or horizontal bars whose height or length represents the frequency of the data. The bar graph below,

taken from his 1786 book *Commercial and Political Atlas*, is actually something that today we would call a compound bar graph, as it compares data classified with two or more categories.[1]

Take a couple of minutes to look it over, then complete the tasks below it.

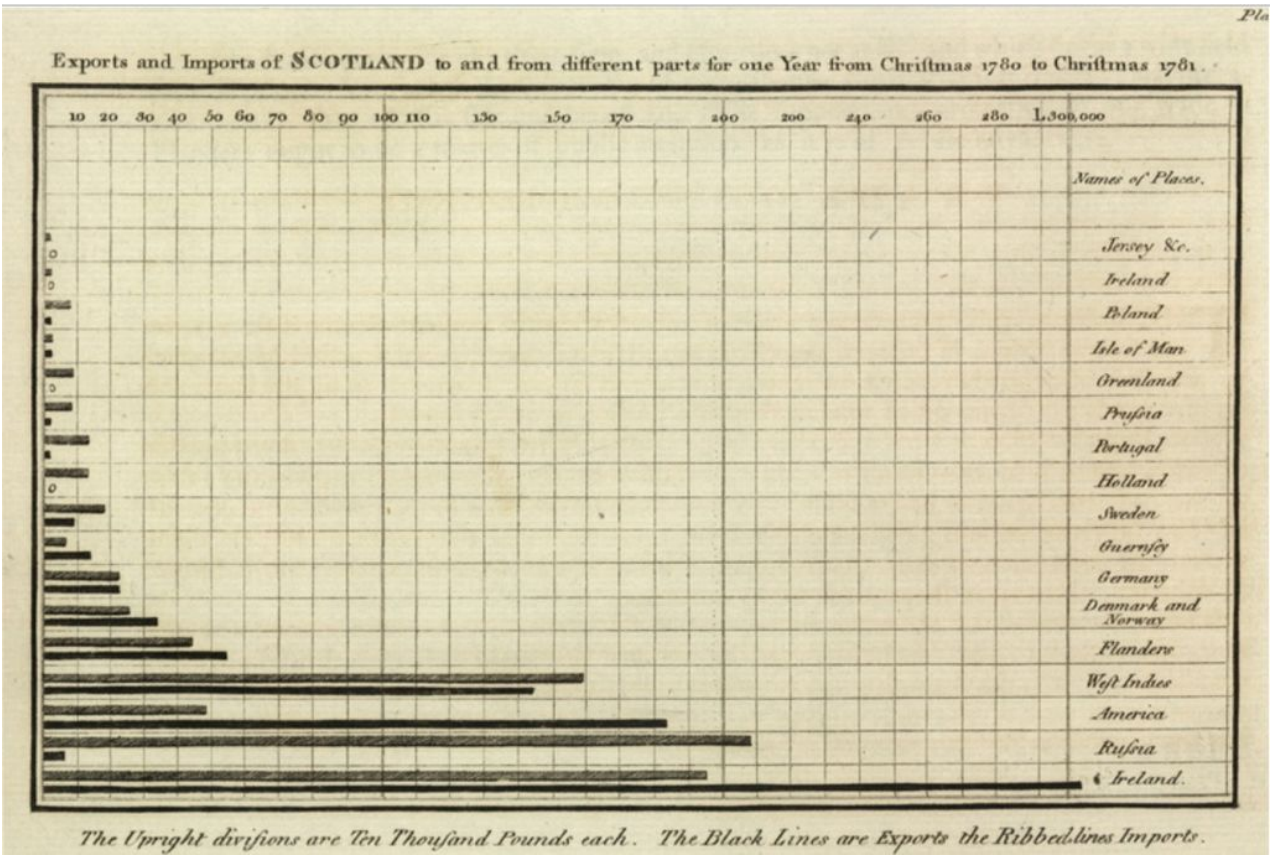


Figure 1: Chart of the Exports and Imports of Scotland to and from different parts for one year, from Christmas 1780 to Christmas 1781.

Task 1

(a) What data does this graph summarize?

(b) What does the x-axis represent? Is the scale of the x-axis constant?

(c) What does the y-axis represent?

Task 2

(a) Why is this a compound bar graph?

(b) What are the "two or more categories" referenced above in the initial description of a bar graph?

[1]The bar chart can be found online: `https://commons.wikimedia.org/wiki/William_Playfair#/media/File:1786_Playfair_-_Exports_and_Imports_of_Scotland_to_and_from_different_parts_for_one_Year_from_Christmas_1780_to_Christmas_1781.jpg`.

Task 3

(a) What are some examples of countries with which Scotland imported more than it exported?

(b) Were there any countries:

(i) with no imports from Scotland?

(ii) with no exports to Scotland?

(iii) where the imports and exports appear to be about the same?

(c) Which country had the biggest difference in imports/exports?

Task 4 What in this data or about this representation surprises you?

Task 5 How would you use this data to determine whether Scotland imported or exported more in total during this time frame?

Task 6 How does this representation of a compound bar graph compare or contrast with how we typically display compound bar graphs today?

3 Pie Chart

Playfair is also credited with using the first pie charts. In general, pie charts are useful for showing part-whole relationships. They are circles whose sections show the percentage of frequencies in each category of the distribution.

The following early pie chart is from Playfair's translation of Denis Francois Donnant's (1805) *Statistical Account of the United States of America*. Playfair has labeled the segments and also stated the actual values of the land areas in square miles on each segment. [2]

[2] This chart can be found online:
`https://blog.usejournal.com/why-humans-love-pie-charts-9cd346000bdc`.

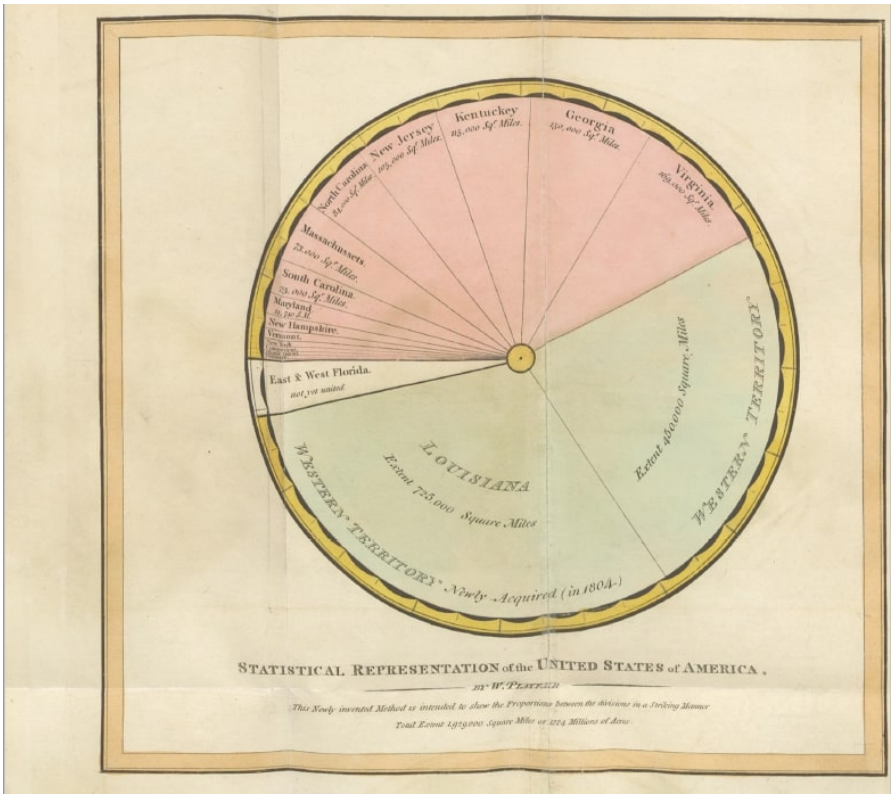


Figure 2:The Statistical Account of the United States of America in 1805.

Task 7 What does each slice of this pie chart represent? What do the colors represent?

Task 8

(a) According to this graph, which state/territory is the largest?

(b) What are some examples of states/territories:

(i) that appear to be about the same size?

(ii) where one appears to be about twice as large as the other?

(iii) that together compose approximately one-fourth of the total area of the country?

Task 9 Were there states/territories in 1805 that are different in relative size than they are today? That is, are there any pairs of states where the larger one in 1805 is now the smallest of the two? Or vice-versa?

Pie charts are still commonly used today, though some experts advice strongly against them. Consider what the statistician Edward Tufte says about them in his book *Visual Display of Quantitative Information*, see Tufte [2001].

Tables are preferable to graphics for many small data sets. A table is nearly always better than a dumb pie chart; the only thing worse than a pie chart is several of them, for then the viewer is asked to compare quantities located in spatial disarray both within and between pies – Given their low data-density and failure to order numbers along a visual dimension, pie charts should never be used.

Task 10

(a) What do you think of this argument?

(b) Pie charts are common in the media today. Look around and find some. Can you find any examples from the media where pie charts are used, but where you think another means of portraying the information would be better?

In today's computer era, it has become more common to use three-dimensional versions of pie charts. While these look attractive, they can also be very deceptive because of the way that the introduction of the third dimension distorts proportions.

Consider the following graph[3]

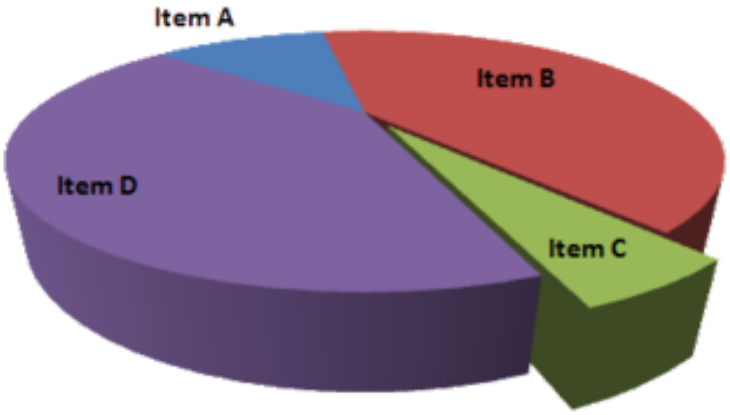


Figure 3: Three-Dimensional Misleading Pie Chart

Task 11 Compare the four items in the chart. Which is largest? smallest?

Now look at the following two-dimensional pie chart of the same data:

[3]This chart can be found online:
`https://en.wikipedia.org/wiki/Misleading_graph`.

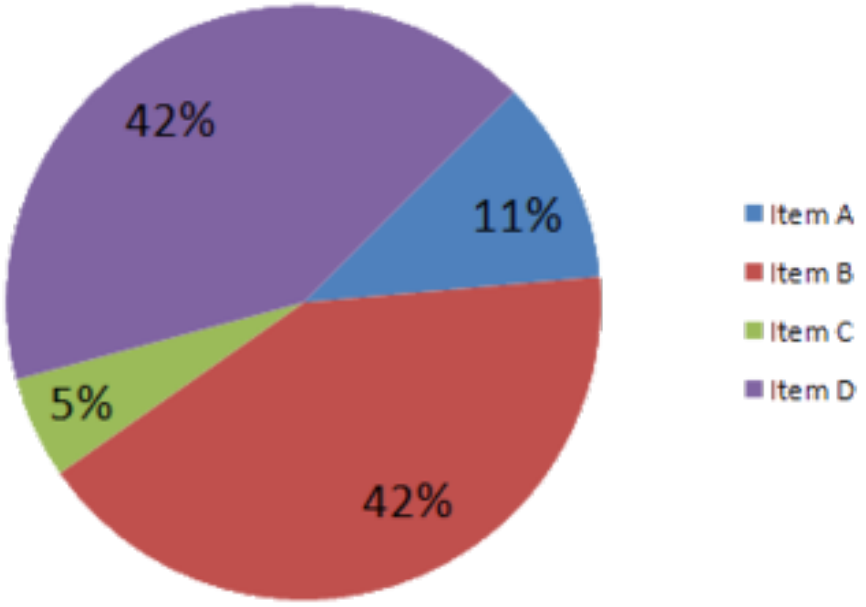


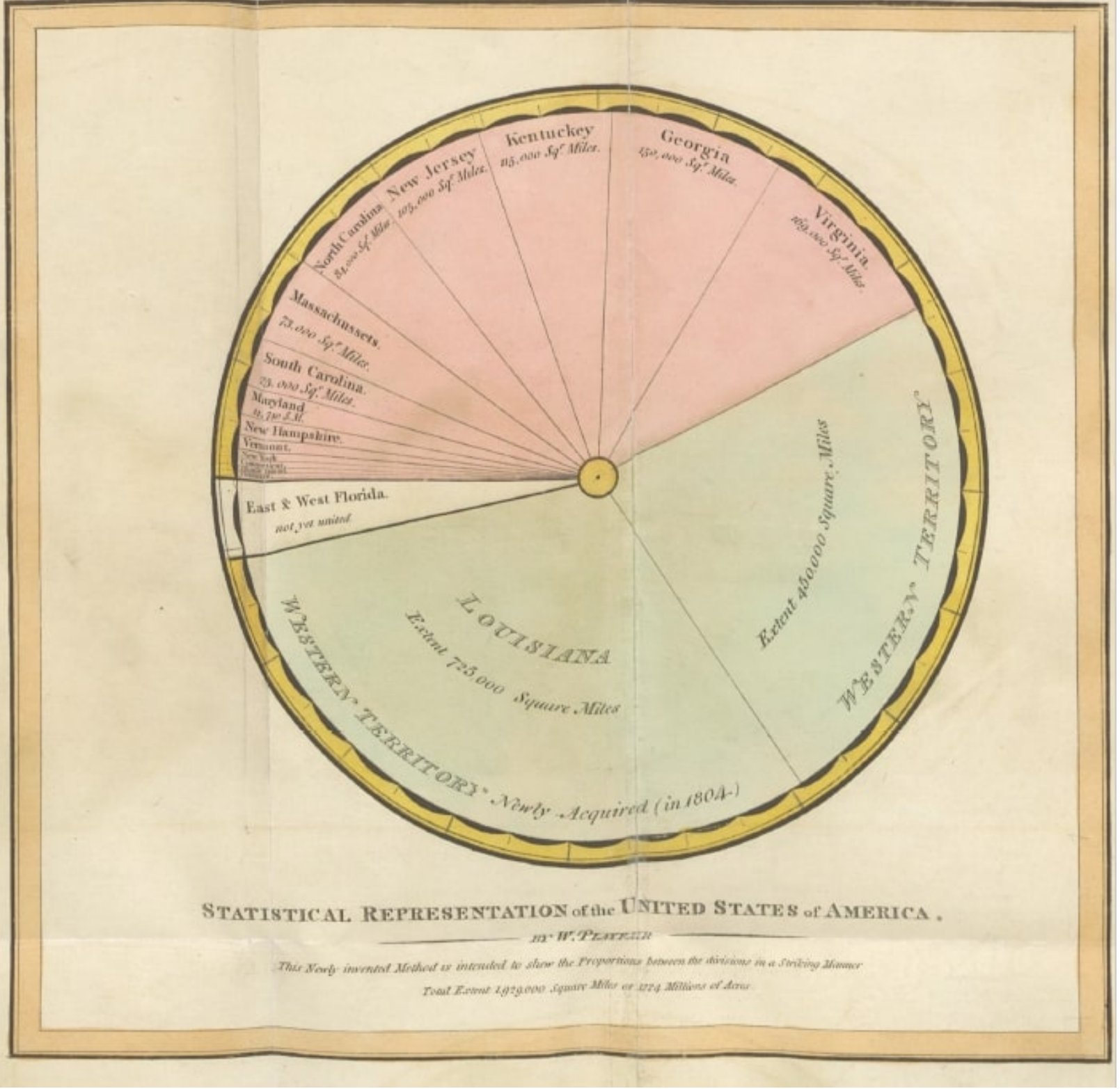
Figure 4: Two-dimensional Pie Chart

Task 12 (a) Based on this 2-dimensional chart, which item is the smallest? Largest?

(b) Compare your answers to (a) with your answers to the prior task. What can you conclude?

4 Final Section

Visualization of data has become an integral part of our society, especially with the more modern trend toward data science. William Playfair catalyzed this transition in the way we communicate data, breaking from the traditions of his time to focus on visual representations of data as opposed to tables. A variety of other representations of his are available in his various books, found in the references. His impact continues today.



A Statistical Account of the United States of America in 1805.

References

William Playfair. Export import to and from all north america from 1700 to 1800 (from 3e edition, 1801), 1786a. URL https://commons.wikimedia.org/wiki/File:1786_Playfair_-_5_Export_\%26_Import_to_and_from_all_North_America_from_1700_to_1800_(from_3e_edition,_1801).jpg. [Source: The Commercial and Political Atlas, 1786 (3th ed. edition 1801)].

William Playfair. Exports and imports of scotland to and from different parts for one year from christmas 1780 to christmas 1781, 1786b. URL https://commons.wikimedia.org/wiki/File:1786_Playfair_-_Exports_and_Imports_of_Scotland_to_and_from_different_parts_for_one_Year_from_Christmas_1780_to_Christmas_1781.jpg. [Source: The Commercial and Political Atlas, 1786 (3th ed. edition 1801)].

William Playfair. *The Commercial and Political Atlas*. T. Button, Little Queen Street, 1786c.

William Playfair. *The Statistical Breviary: Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe ... to Which Is Added, a Similar Exhibition of the Ruling Powers of Hindoostan*. J. Wallis, 1801.

William Playfair. Statistical representation of the united states of america (1805), 1805. URL https://blog.usejournal.com/why-humans-love-pie-charts-9cd346000bdc.

William Playfair. In Howard Wainer and Ian Spence, editors, *The Commercial and Political Atlas and Statistical Breviary*. Cambridge University Press, 2005.

Smallman12q. Two-dimensional pie chart, 2012. URL https://upload.wikimedia.org/wikipedia/commons/8/87/Sample_Pie_Chart.png.

Ian Spence and Howard Wainer. *Introduction to Playfair's Commercial and Political Atlas and Statistical Breviary*. Cambridge University Press, 01 2006. ISBN 9780521855549.

J. Stillwell. *Mathematics and Its History (Undergraduate Texts in Mathematics) 3e*. Springer, 2010.

Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press USA, Cheshire, CT, 2001.

Notes to Instructors

PSP Content: Topics and Goals

This Primary Source Project (PSP) aims to use original displays of William Playfair as a means to deepen student understanding of basic ways to interpret visual displays of data.

It is intended for use in an introductory statistics or data science course at the undergraduate level. However, it could also be used in courses for pre-service teachers, mathematics for liberal arts courses, professional development courses/workshops for teachers, or in history of mathematics courses. It is also potentially suitable for use at the middle- or high-school level.

Student Prerequisites

Prerequisites for this project are minimal to non-existent, as students likely already have basic familiarity with seeing and interpreting data from their K12 education and from the media.

PSP Design, and Task Commentary

The PSP has an introductory section followed by sections on the pie chart and bar graph. The section on the pie chart contains a task in which students are asked to read and comment on a quote from Edward Tufte arguing against ever using pie charts. The students are also presented with a misleading 3-dimensional pie chart to bring out pitfalls for them to be aware of in visual representations of data. Interestingly, the authors tried to recreate the pie chart themselves using Microsoft Excel and Google Sheet but the software seemed to auto-adjust the orientation in a way that did clearly represent the data.

The following questions inspired the task development throughout (i.e., what do we want the students to understand):

1. What is the data being shown?
2. How is the data being shown? Is it effectively shown, why (or why not)?
3. What"story" does the data tell (i.e. what can we infer from the data)? How is the story built from multiple figures? Does anything about the story surprise you?

Suggestions for Classroom Implementation

As with most PSPs, we encourage instructors to assign reading and/or homework tasks before and after class, and to incorporate active learning into class time. Students could discuss their responses in small groups or present to the class, with the instructor facilitating a discussion on what they are learning.

Given the brief nature of the PSP, the students could be expected to read some or all of the sections beforehand and have some answers and opinions to share with a group. Although group work is not strictly necessary, we find that it does significantly enhance the learning experience. The advance reading of the material is an important aspect of the PSP. Some of the figures in this PSP would be best seen on a large screen or overhead projector rather than a laptop or tablet. We strongly suggest you have the students show up with some preliminary answers that you can ask them to revise, or not revise. However, each student should have reasons for their revisions or lack thereof.

LaTeX code of this entire PSP is available from the author by request to facilitate preparation of advanced preparation / reading guides or 'in-class worksheets' based on tasks included in the project. The PSP itself can also be modified by instructors as desired to better suit their goals for the course.

Sample Implementation Schedules

For a lower-level course with a standard amount of work outside of class (2-3 hours per hour in class), this PSP could be implemented in approximately one class period.

An instructor could probably multiply this by 1.5 if the majority of the PSP work is done in-class, with minimal work outside of class, and with students working in pairs or groups during class.

The actual amount of time spent on each section and on this PSP naturally depends on the instructor's goals and on how the PSP is actually implemented with students. For example, it could take 1.5 days with a math for liberal arts class, or less than a class period in an upper-level mathematics class for pre-service secondary teachers.

Online Implementation

The authors of this PSP have implemented the extended, full-length version of this PSP in a 3-credit, lower-level, online introductory statistics course taught during an 8-week summer session. We used Canvas as our course management software, dividing the students into groups of approximately 3-4 using the "Discussion" feature of the software. The students were asked to answer some of the questions from the PSP and then read the responses and communicate with their peers about the similarities and differences in their individual answers. The students were then tasked with re-writing their answers while incorporating what they'd learned from the discussion with their peers and submitting a final write-up. This PSP served as the class project for the summer. As such, it was interspersed with other coursework and implemented over an approximately 2-week period toward the end of the first third of the course.

We split this PSP into two sections for an online class where the first section covered tasks 1 through 12, which are the contents of this PSP.

We asked the following questions for the online discussion:

Read the answers from the other members of your group. You should have 2-3 other members in your group. Respond specifically to each classmate with at least the following information:

1. Are there any responses of theirs that differed from yours? Helped you see something in a different or more clear way? That you think are particularly well explained?

2. Are there any changes you'll make to your final submission based on your classmates' answers? (Note that eventually you'll submit one well-written set of responses to these and additional tasks.)

3. Do you see anything in your classmates' answers that you disagree with or think may be incorrect? If so, be specific about what and why you think differently. (Recall that you are NOT graded on accuracy of these answers yet, so this is only helpful feedback to your other group members. It's also a chance for you to think deeply and assess their answers, which is a higher cognitive skill than just answering yourself. Please be kind and professional, as this is a learning environment

and perfection is not expected, but also be specific and direct.)

4. Is there anything else you'd like to tell them about their responses that you haven't already? If so, include it as well.

This PSP has several tasks that have no single answer and/or no right answer. The tasks with *no one answer* are: 3(a), 15 and 18. The tasks with *no right answer* are: 3(b)(iii), 8(b)(i), 10(a), and 10(b).

Connections to other Primary Source Projects

This PSP connects loosely with all of the PSPs involving statistics. There are two other PSPs by the same authors that could be used to complement this one - M31:Playfair's Introduction of Bar and Pie Charts to Represent Data and M33: Playfair's Novel Displays of Data. In addition, there is *Seeing and Understanding Data* by Beverly Wood and Charlotte Bloch as well as *Regression to the Mean* and *Quantifying Certainty: the p-value*, both of which are authored by Dominic Klyve.

Additional Historical Notes

Printing illustrations was quite challenging at the time of Playfair.

Recommendations for Further Reading

There are several books and resources in the references. The first author found the following book chapter to provide particularly useful historical context:

Introduction to Playfair's Commercial and Political Atlas and Statistical Breviary, January 2006, by Ian Spence and Howard Wainer, Cambridge University Press.

The Statistical Breviary: Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe ... to Which Is Added, a Similar Exhibition of the Ruling Powers of Hindoostan, 1801, by William Playfair, J. Wallis.

Acknowledgments

The development of this student project has been partially supported by the TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) Program with funding from the National Science Foundation's Improving Undergraduate STEM Education Program under Grant Nos. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily represent the views of the National Science Foundation.