Representing and Interpreting Data From Playfair

Diana White
River Bond
Joshua Eastes
Negar Janani

Follow this and additional works at: https://digitalcommons.ursinus.edu/triumphs_statistics

Part of the Curriculum and Instruction Commons, Educational Methods Commons, Higher Education Commons, Science and Mathematics Education Commons, and the Statistics and Probability Commons

Click here to let us know how access to this document benefits you.

Recommended Citation
White, Diana; Bond, River; Eastes, Joshua; and Janani, Negar, "Representing and Interpreting Data From Playfair" (2019). Statistics and Probability. 4.
https://digitalcommons.ursinus.edu/triumphs_statistics/4

This Course Materials is brought to you for free and open access by the Transforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) at Digital Commons @ Ursinus College. It has been accepted for inclusion in Statistics and Probability by an authorized administrator of Digital Commons @ Ursinus College. For more information, please contact aprock@ursinus.edu.
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.

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 for two or more categories.¹

Take a couple of minutes to look it over², 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?

---

¹The 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](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).

²Note that for added clarity and ease of reading, larger versions of all of the figures in this PSP are at the end of the PSP.
(b) What are the “two or more categories” referenced above in the initial description of a bar graph?

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 Domnant’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.

---

3 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 chart.

![Three-Dimensional Misleading Pie Chart](https://en.wikipedia.org/wiki/Misleading_graph)

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:

---

4This 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 Time Series - Exports and Imports

Playfair created the “Time Series” graph to visually represent measurements over specified time intervals. For instance, the figures in this section provide a window into various financial measurements between the years 1700 and 1800. Note how the intervals are marked.

The following is an example of a Time Series graph that depicts imports and exports of England to and from North America.\(^5\)

\(^5\)This chart can be found online on https://commons.wikimedia.org/wiki/File:1786_Playfair_-_5_Export_\&_Import_to_and_from_all_North_America_from_1700_to_1800_(from_3e_edition,_1801).jpg.
Figure 5: English exports and imports to and from North America (1700 to 1800).

**Task 13**

(a) What does the x-axis represent? How about the y-axis?

(b) What do the shaded areas represent?

**Task 14**

(a) During which years does the difference between imports and export appear to be the largest?

(b) During which years does the balance appear to favor North America?

(c) During which years does the balance appear to favor England?

(d) When does the balance of imports and exports appear to be equal? How did you determine this?

The following chart depicts similar information to Figure 5 but during a more narrow time interval.⁶

---

⁶You can find this graph online at [https://commons.wikimedia.org/wiki/File:1786_Playfair_-_Chart_of_import_and_exports_of_England_to_and_from_all_North_America_from_the_year_1770_to_1782.jpg](https://commons.wikimedia.org/wiki/File:1786_Playfair_-_Chart_of_import_and_exports_of_England_to_and_from_all_North_America_from_the_year_1770_to_1782.jpg).
Playfair’s work is sometimes critiqued for accuracy, and Figures 5 and 6 give us one example of this. Each represent the same data, just during different time intervals. Yet if you look closely at the various high and low points (“local maxima and local minima”, for those who have had calculus) you’ll see that there are some inconsistent data.

**Task 15** Find at least two ways in which the specific data in Figures 5 and 6 are inconsistent. That is, give the year and associated volume of imports/exports for each figure at two instances where they clearly disagree.

**Task 16** (Bonus, take a guess, or try to look it up - the authors have some ideas, but we’re not sure ourselves.) Why do you think the balance of imports was in favor of North America from late 1774 to mid 1775?
5 Compound Time Series Graphs

Take a few minutes to look over the following figure, then complete the tasks below it.

Figure 7: Linear Chronology, Exhibiting the Revenues, Expenditure, Debt, Price of Stocks and Bread, from 1770 to 1824

---

**Task 17**

What are some advantages and disadvantages about how the data is represented?

**Task 18**

Compare the various lines across the time intervals. What do you notice? For example, as one graph increases/decreases, do other graphs seem to increase/decrease? What is the term for these potential relationships?

**Task 19**

(a) How might war affect the price of bread?

(b) Does a time of peace seemingly have the opposite effect on the price of bread? Explain.

(c) What dangers do we risk if we infer more than what the graph depicts?

(d) What is the difference between correlation and causation?

(e) Give an example of two things (not necessarily from Figure 7) that are highly correlated, but where there is no causation between them.

(f) Look up “spurious correlations” and post at least two time series that show spurious correlations. For your amusement, we found one showing how between 2000 and 2009 the per capita cheese consumption and the number of people who died by becoming tangled in their bed sheets has a 95% correlation. Also, divorces in the state of Maine and per capita consumption of margarine were allegedly 99% correlated between 2000 and 2009.

---

This graph can be found online on https://commons.wikimedia.org/wiki/File:Linear_Chronology,_Exhibiting_the_Revenues,_Expenditure,_Debt,_Price_of_Stocks_&_Bread,_from_1770_to_1824,_1824.jpg
Just as we did with pie charts, we now turn our attention to some ways in which time series graphs can be misleading.

**Task 20** Use the internet to find two ways in which time series graphs can be used to deceive viewers. Put an example of each of them here, explaining each in your own words.

## 6 More Complicated Visualizations

Playfair also provided numerous and more complicated displays of visual information.

The figures in this section present information about the value of wheat and the wage of labor from 1565 to 1821. The value of wheat provided in both shillings and the daily wages of labor.

The following chart shows the price of the quarter of wheat\(^8\) and the wage of labor by the week from 1565 to 1821.

---

![Chart Showing at One View with the Price of the Quarter of Wheat, and the Wages of Labor by the Week, from 1565 to 1821.]

---

**Task 21** Can you find any relation between the price of the wheat and the wage of labor during different time periods?

**Task 22** Which of these variables was more stable over time and which ones fluctuated?

---

\(^8\) This refers to a quarter of a bushel of wheat and is an imperial unit of measure.
Task 23: Compare the price of the wheat at the beginning of each governor’s term with the end of their term? How about the wages of laborers at the beginning and end of each governor’s term?

Task 24: What was the effect of changing governors on the price of the wheat and on the weekly wage of laborers?

The following graph shows the value of the wheat in shillings and in a day’s wage for a laborer, which can provide meaningful statistical information for the study of economic and financial fluctuations from the 16th to the 19th century.

Figure 12: Chart Showing the Value of the Quarter of the Wheat in Shilling and in days wages of a Good Mechanic from 1565 to 1821.

Task 25: How has the price of the wheat changed over the time period?

Task 26: How has the daily wage of the laborers changed from 16th century to the 19th century?

Task 27: In which century do we see more fluctuation in daily wages? How about the price of wheat?

Task 28: (a) How did the purchasing power of the laborers change over time?
   (b) Compare the purchasing power of the laborers in any time period.
   (c) During which times do you think laborers faced more financial problems? Explain your reasoning.

Task 29: Where can we see a period of recession reflected in this graph?

Task 30: How does the graph reflect any historical changes?
7 Wrap Up

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. His impact continues today. A variety of other representations of his are available in his various books, found in the references. We encourage you to take a look at some of them!
Chart of the Exports and Imports of Scotland to and from different parts for one Year from Christmas 1780 to Christmas 1781.
A Statistical Account of the United States of America in 1805.
English Exports and Imports to and from North America (1700 to 1800).
English Imports and Exports to and from North America (1770 to 1782).
Linear Chronology, Exhibiting the Revenues, Expenditures, Debt, Price of Stocks and Bread, from 1770 to 1824.
Chart Showing at One View with the Price of the Quarter of Wheat, and the Wages of labor by the Week, from 1565 to 1821
Chart Showing the Value of the Quarter of the Wheat in Schilling and in Days Wages of a Good Mechanic from 1565 to 1821.
References


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. Parts of it, in particular the last section, could also be used to introduce more difficult displays of data after a standard textbook treatment.

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 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 is structured by topic, starting with some basic introductory material, followed by sections on the pie chart, bar graph, time series, and more complicated compound time series. In several of the sections, we have “bonus” tasks that focus on either concerns or common misrepresentations of data related to that section. For example, 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 each 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.

The students should be expected to read 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. Many 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 a standard week of semester class time (three 50-minute periods or two 75-minute periods). A possible implementation plan for 50-minute class periods is as follows: For three classes, 50-minute classes: Day One: Sections 1,2 & 3; Day Two: Sections 4 & 5; Day Three: Sections 6. For two classes, 75 minute classes: Day One: Sections 1,2 & 3; Day Two: Sections 4,5 & 6

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 number of class periods 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 weeks with a math for liberal arts class, one week with an introductory statistics course, and as little as half a week with an upper-level history of mathematics course.

**Online Implementation**

The authors of this PSP have implemented it 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 and the second section covered tasks 13 through 21. We did not cover section 6 (tasks 22 - 31) for this implementation.

We used the following prompts and 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), 10(b), 16, 17, 19, 20(a)-(g) and 21.

Possible Modifications of the PSP
This PSP is also available as three mini-PSPs, each of which is designed to be approximately one class period in length:


Connections to other Primary Source Projects
This PSP connects loosely with all of the PSPs involving statistics, in particular with 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:


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.

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (https://creativecommons.org/licenses/by-sa/4.0/legalcode). It allows re-distribution and re-use of a licensed work on the conditions that the creator is appropriately credited and that any derivative work is made available under “the same, similar or a compatible license”.

For more information about TRIUMPHS, visit https://blogs.ursinus.edu/triumphs/.