The Evolution of Textbook Illustrations in Probability and Statistics: A Comparison

Berke Aygulhan[1]
Supervisor(s): Martin Skrodzki[1]**, Mrinal Dhume**[1]
[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 21, 2026

Name of the student: Berke Aygulhan
Final project course: CSE3000 Research Project
Thesis committee: Martin Skrodzki, Mrinal Dhume, Ranga Rao Venkatesha Prasad

The Evolution of Textbook Illustrations in Probability and Statistics: A Comparison

Berke Aygulhan

Delft University of Technology

Abstract. Illustrations play a major role in helping students understand probability and statistics, yet there are limited studies researching how textbook illustrations evolve alongside technology. This paper compares illustrations from two editions of the same probability and statistics textbook published 34 years apart. Because the author and target audience stayed the same, we were able to isolate how illustration practices changed. We analyzed 103 illustrations across three chapters using thematic analysis. We aimed to answer two questions: how did the use of illustrations change between two editions, and how do illustration types differ depending on the given topic? We found that, over time, illustrations shifted away from showing step-by-step procedures and instead focused on final results. The newer edition uses color to highlight answers, includes output of computer software, and places more detailed images closer to the text. However, when looking across the different topics, we found that specific topics use specific types of illustrations.

Keywords: Probability and Statistics · Textbook Illustrations · Thematic Analysis.

1 Introduction

Pictures in a mathematics textbook are not just for decoration. A good image can make an abstract idea easy to see [3]. In probability and statistics (P&S), where students must understand the concepts and patterns of data, pictures show what words simply cannot [3]. Therefore, the way textbook images are designed affects how students learn. However, limited number of studies have investigated how these illustrations change over time [7], and current existing research looks at geometry.

This could be an ideal opportunity to research this subject because the practices in the field have changed quite a lot [4]. In the past, students calculated numbers and drew graphs by hand, while today this work is typically done by software, and students are expected to read the results. The textbooks might also follow this. One approach to study this trend is to compare P&S textbooks from different eras. To address this gap, we compare the figures in one probability and statistics textbook by one author across two editions published 34 years apart. We aim to answer two research questions:

- **Research Question 1:** How did the type, amount, and purpose of illustrations change between the two editions for the same topic?
- **Research Question 2:** What patterns emerge when comparing the use of illustrations across topics for the same book?

For the first research question, there are two main themes. First, illustrations produced by software focus more on showing final results rather than explaining how to draw the graph by hand. The new edition uses screenshots from software and uses color to highlight results inside plots. Second, the modern illustrations contain more features and sit closer to the text that explains them, even though their captions are much shorter. For the second research question, we found that the author consistently uses specific types of illustrations for specific topics, and this did not change between 1982 and 2016. One explanation for this stability is simply that the same author produced both editions. The pairing of topic and figure type may be a personal preference, and that preference is one of the things our same-author design holds constant. So we cannot tell from this study whether the topic-figure pairing reflects something about the subject matter or just the consistent habits of one author. However what can be said is that across a 34-year span in which the figures changed substantially in other respects, this particular pairing stayed stable.

2 Background and Related Work

This section reviews relevant prior work and identifies the gap that the present study addresses. The visual form of a figure means something by itself [3]. Carter examined whether diagrams are more useful than symbols in mathematics. By analyzing several examples of propositions and their proofs, she found that both diagrams and symbols can lead to new ideas. However, she identified two features that only apply to diagrams: they can show a relationship instead of just stating it, and they can use space in two dimensions. This is important for our study because it means that changing a figure's two-dimensional layout also changes what it tells the reader, justifying our comparison of how each edition's figure looks.

The look of textbook diagrams changes over time [7]. Dimmel and Herbst built a method for studying textbook diagrams and used it on about 2,300 figures from 22 geometry textbooks from the 20th century. Their work showed that the visual features of diagrams carry meaning and that these features change over time. This is the closest study to ours, and one of the motivations we made this study. However, their framework was built specifically for geometry, which relies on features like labeled points and angle markings. These do not match the plots, abstract diagrams, and shaded regions used in probability and statistics. That's the reason our study used their work as motivation rather than copying their process.

Slough and McTigue's GAP framework checks how well textbook graphics support learning [11]. It looks at form and function. It looks at whether a graphic helps the reader build a mental model, and whether the text is placed close to the figure. We borrow two of its parts, caption quality and spatial contiguity. Bustamante-Valdés et al studied statistical graphs in primary school textbooks and labeled them by graph type, reading level, complexity, and task type [2]. Their graph-type code fits our study, but their reading-level and task-type codes do not because they studied student exercises in primary school.

How statistics are taught has changed because of computers [9]. Moore argued that technology strongly affects both what is taught and how it is taught in statistics. This idea became official in the GAISE College Report, which tells teachers to use real data and analyze data with technology instead of by hand [4]. Therefore, as statistics moved from hand calculation to software, the textbooks start showing software and its results. The way textbooks are produced has also changed. Advances in printing technology altered textbook layouts, making four-color printing the norm by the 1980s and driving down production costs [10]. At the same time, the statistics field shifted from constructing plots by hand to relying on software for graphs [9]. These overlapping changes help explain why a P&S textbooks' figures might change a lot over the years.

These four ideas point in the same direction, but leave a gap. The look of a figure carries meaning [3], textbook figures change over time [7], multiple systems for studying figures existing [11,2], and statistics has moved from hand calculation to software while the technology for making books changed at the same time [4,9]. What is missing is a study that connects them. The research on how textbook figures change over time has focused mostly on geometry, where figures are points, lines, and angles. Much less attention has been paid to probability and statistics, where the figures are plots, software output, and shaded regions. Therefore, it is still unclear how the figures in a P&S textbook changed as the technology around them changed.

This study seeks to address this gap. Our study compares the figures in one probability and statistics textbook by one author across two editions which are published 34 years apart. Due to the author, the book and its main content staying the same, changes in the figures can be attributed more confidently to the change in technology and methodology. Using an analysis built from the frameworks above, we ask not only whether the figures changed. We also ask in what specific ways, and whether the type of figure used for each topic changed along with them.

3 Methodology

To understand how certain properties of the illustrations changed between the two editions, we used thematic analysis. Thematic analysis is a qualitative method for identifying and interpreting recurring patterns across a dataset [1]. This is suitable for our study because the aim is to find patterns across a set of figures. It lets us to start with features for each illustration and then group these features into higher-level themes. This way of moving from individual features to general themes allows us to see what changes happened rather than confirming a change we predetermined.

To classify the illustrations, iterative open coding was carried out in ATLAS.ti. Open coding is an exploratory process in which the researcher continuously questions what their observations might represent in order to generate new ideas and guide further data collection [12]. We used thematic analysis with a codebook built from existing frameworks and refined iteratively. This approach suits our study, where the goal is to characterize how illustrations are used rather than to test a predefined theory.

To study how technology changed illustrations, we need textbooks where the technology was the main differentiator. To reduce other differentiators, we chose a single textbook by a single author and compared two of its editions: Devore's 1st edition [5] and the 9th edition [6]. Holding the author, book, and audience constant lets us attribute differences in the figures more confidently to the change in production technology rather than to a different writer. From this book, we then narrowed the analysis to three chapters.

Moreover, three subtopics were selected for analysis using purposive sampling, in which cases are chosen for their value in informing the research question [13]. The selection criterion was the chapters with the largest increase in number of figures, maximizing the figure data available for comparison. The selected chapters were Descriptive Statistics, Joint Probability, and Linear Regression, as shown in Table 1. Other than having the highest increase, these topics contain various different types of illustration ranging from function plots to conceptual diagrams, thereby ensuring a diverse visual corpus as well.

Chapter Title	**1st Ed.**	**9th Ed.**	**# Fig.**
Overview and Descriptive Statistics	12	23	11
Probability	11	14	3
Discrete Random Variables and Probability Distributions	9	8	-1
Continuous Random Variables and Probability Distributions	28	38	10
Joint Probability Distributions and Random Samples	12	23	11
Point Estimation	7	7	0
Tests of Hypotheses Based on a Single Sample	12	10	-2
Simple Linear Regression and Correlation	11	22	11

Table 1. Figure Count Comparison for All Common Chapters

Our codebook was built using an abductive coding process, meaning we moved back and forth between established theories and our own data to refine our categories[8]. First, we used deductive coding to apply existing frameworks to our figures. From Slough and McTigue's GAP framework [11], we utilized "caption quality" and "spatial contiguity," and from Carter's taxonomy [3], we utilized the "text–illustration relationship." Second, we used inductive coding to capture unique features in the P&S figures that fell outside of those frameworks. For example, we created a new code to identify if a figure was a direct software output. Combining these two directions allowed us to best explain the data.

Each illustration was coded along nine categories, which can be seen in Table 2 with their definitions. The coding scheme was developed iteratively: after each edition was completed, codes were reviewed, merged, or split as needed, and their definitions sharpened. Decision trees were constructed for some categories to ensure consistency during coding. We created simple rules for categories that could be confusing. For example, we only coded a figure as 'software-output' if the book explicitly named the software used such as Minitab, Python, or R. This was regardless of how the computer generated image looked. Simpler categories did not need these kinds of rules. For example, coding for "color" just meant checking if the color blue was present.

Code	**What it records**
Illustration Type	histogram, boxplot and so on.
Communicative Function	showing data, explaining a concept and so on.
Text–illustration Relationship	How the picture connects to the writing around it.
Caption Quality	Whether the caption just names the picture or actually explains it.
Spatial Contiguity	Whether the picture is on the exact same page as the text talking about it.
Text Reference	Whether the main text mentions the picture (like saying "See Figure 1").
Features	Details added to the picture (like titles, labels, or legends).
Color	Whether the picture uses color or is just black and white.
Software Output	Whether the picture is a direct output from a software.

Table 2. Illustration Codes for the Codebook

We coded every figure coded nine categories. The coded data served as our supporting evidence. When we were coding the book, we identified specific shifts between the two editions. We checked these theories against our coded data. Patterns that recurred across chapters and were supported by the coded evidence were retained as themes. The themes that were not supported by the codebook were discarded.

4 Results

4.1 The figures were remade as digital illustrations

The most visible difference between the two editions is color. How the color is used matters more than the fact that it is used in this case, because color is a known device for directing a reader's attention to specific parts of a figure [7]. In our codebook, a figure was categorized as colored if any part of it had blue, and grayscale if not. All 35 figures in the 1st edition are grayscale. In the 9th edition, 56 of the 68 figures have color. The 9th edition colors the part that carries the result.

Figures 1 and 2 are the same figure in both editions. Both figures are the error distribution in panel (a) and the regression line in panel (b). In the 1st edition, the whole figure is grayscale, so the result line sits at the same visual level as the axes. This makes it hard to distinguish the line from the rest of the figure. In the 9th edition, the figure shows the exact same thing. However, this time the line and curves are blue while the axes and labels stay grayscale. Both figures convey the same content, but the 9th edition's colored figure elevates the result.

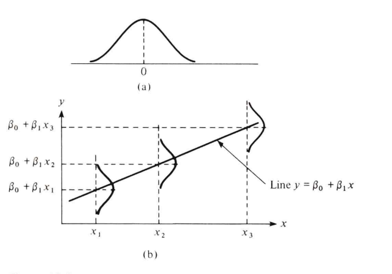

Fig. 1. Error distribution and regression line (1st Edition, grayscale).

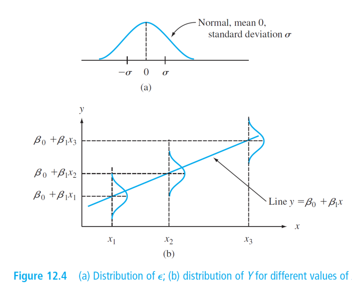

Fig. 2. Error distribution and regression line (9th Edition, Blue color highlight).

One can also argue that the extra color is simply a consequence of color printing becoming cheap, and the 9th edition just follows this general trend like the rest of the modern textbooks [10]. Cheaper color printing is part of the reason. If color were only used just because it was cheap, we would expect it to fill the whole figure. Instead, the figures show the opposite. In the 9th edition, it was used selectively. Only certain parts, like the result line, have the blue color. That is hard to explain if color is just a cheap decoration. Cheap printing explains why color was available, but it does not explain why it was used selectively. It is more possible if the author is using color on purpose to point the reader straight to the answer.

The second change in the 9th edition is a new type of figure that shifts the reader from doing a calculation to reading its result. In our codebook, we categorized a figure as "software output" if the book states it was produced by Minitab, Python, or R. These are programs that compute and display results the reader would otherwise work out by hand. According to our codebook, software output appears 18 times in the 9th edition and never in the 1st edition.

The clearest way to see the shift is to look at how both editions teach the same regression topic. The 1st edition uses idea figures, scatterplots of points around a line in Figure 3, and the error distribution in Figure 4. They show what regression is and set up a calculation that the reader then can do by hand. The 9th edition shows a finished Minitab output instead. It gives the regression equation, the R-sq value of 79.1%, and a full analysis-of-variance table. The reader computes none of it. The book even adds its own blue arrows and labels on top, pointing to the parts the reader should focus on. The 1st edition never needed labels like these, because it had no "software output" to explain. So for the same topic, the 1st edition shows the idea and lets the reader do the math, while the 9th edition shows the finished result and teaches the reader how to read it.

"Software output" also comes with a cost that recoloring does not. A recolored figure does the same job as before, just more clearly. On the other hand, a "software output" shows a finished result without showing how it was reached. The contrast with the 1st edition and the 9th edition is what makes this more apparent. In our codebook, we categorized figures as "demonstrate-technique" if the figure's job is to show how to carry out a procedure step by step. Both editions have the same 6 figures for "demonstrate-technique". But the more interesting finding is that the "software output" figures were added around them instead of replacing them. Therefore, the addition of new figures did not affect the number of "demonstrate-technique" figures, but it complemented them. Hence, the older figures more often showed the reader how an answer was reached, while the "software output" figures show the answer and leave the method inside the program.

Though, the two changes do not always move together. Only about half of the "software output" figures are colored at all. This is far below the 82% colored rate for figures of the 9th edition. The 9th edition seems to have deliberately left many of them in grayscale, possibly to match what a student sees on screen when running the program. "Software output" is therefore partly outside the color trend, not an example of it. One further exception sharpens the point: of

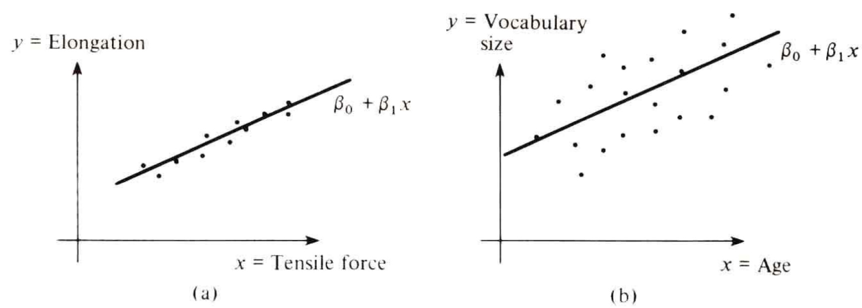

Figure 12.4 Typical sample for (a) small σ^2, (b) large σ^2

Fig. 3. Scatterplot of points around a line (1st Edition).

```
The regression equation is
cet num = 75.2 − 0.209 iod val

                            β̂0  β̂1
Predictor          Coef      SE Coef         T         P
Constant         75.212        2.984     25.21     0.000
iod val        −0.20939      0.03109     −6.73     0.000
                                100r²
s  =  2.56450   R-sq  =  79.1%    R-sq(adj)  =  77.3%

Analysis  of  Variance             SSE

SOURCE         DF         SS         MS          F         P
Regression      1     298.25     298.25      45.35     0.000
Error          12      78.92       6.58
Total          13     377.17  ←──────────────  SST
```

Figure 12.12 Minitab output for the regression of Examples 12.4 and 12.9

Fig. 4. Finished Minitab regression output with annotations (9th Edition).

the 18 software outputs, only one is coded as illustrating a concept rather than presenting a result in Figure 5, and even that one falls back on a histogram to carry the idea. None of the "software output" figure is an abstract diagram; they are all conventional displaying figures, such as histograms, tables, and scatterplots. Therefore, even when a software output teaches a concept, it does so through a traditional display. Taken together, "the figures were remade as digital" turns out to mean two different things: old figures recolored to highlight results, and a genuinely new figure type that moves the reader from performing calculations to reading them.

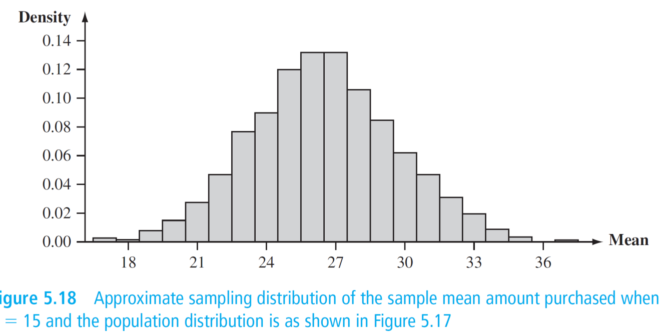

Fig. 5. Minitab sampling-distribution output illustrating a concept. (9th edition)

4.2 The figures became more elaborate and more tightly integrated

Color is not the only thing that changed between the two editions. The 9th edition is also more organized about how it presents the figures. The figures carry more detail, combine a wider variety of figures, and sit closer to the text that explains them.

The 9th edition combines a wider variety of figures into single illustrations. In our codebook, we call a figure a "composite" if it combines different types of figures under one. Both editions use "composite" figures, but they became more common, rising from 6% to 12% of figures. This is a change of its own, not a consequence of "software output". Out of the eight ninth-edition composites, four contain a "software output", and four do not. In our codebook, we also categorized what types of figures each "composite" figure contains. In the 9th edition, "composite" figures also draw on a wider range of figure types. While the 1st edition combines three different kinds of figures, the 9th edition combines seven.

The clearest way to see the change is to put a "composite" from each edition side by side. Figure 6, from the 1st edition, has two panels: a frequency-distribution table beside a histogram of the same data. Figure 7, from the 9th edition, has five panels: four histograms showing how the sampling distribution approaches normality as the sample size grows from 5 to 30, plus a normal probability plot. Both do the same basic job. Both place related views together, so the reader can compare them in one place instead of across several pages. But what gets compared grew. The 1st edition "composite" sets two representations of a single dataset against each other while the 9th edition one lays out a whole sequence of distributions at once. So "composite" figures did not just become more frequent. They also came to support richer comparisons.

This is what "more organized" means for the figures. Putting more panels under one number lets the reader stay on the topic without flipping between pages. And because the panels sit together, the reader can compare them directly, which would be harder if the figures were separated around the edition. An increase in the variety of figures also makes it easier for the reader to see different ideas. The 9th edition does not just use composites more often; it also combines a wider variety of figures in them.

The figures carry more features of their own in the 9th edition. In our codebook, a "feature" is something added to a figure to help read it, like axis labels, titles, gridlines, or legends. Titles, gridlines, and legends were absent in the 1st edition, so these are basically new. Adding these features lets a figure explain more of itself. The change is real but modest. The 9th edition adds new kinds of features rather than overfilling every figure with them. Still, a figure with a title or a legend can say more about itself, so a reader can tell what a figure is about without searching through the text. This makes the figure more self-contained.

Change in captions between the two editions points in the opposite way than the previous changes. In the 1st edition, 6 figures have no caption at all. In the 9th edition, every figure has one. Not only were captions fixed, but the captions were also changed. Captions that add extra information fell from 49% to 35%, while captions that only name the figure rose from 34% to 65%. Thus, captions became more common, but they say less on their own. The figures become less self-contained.

The interesting part is that the extra information from the captions did not disappear, but moved. In the 1st edition, a caption often carried an explanation the reader needed. In the 9th edition, that explanation is usually in the body text instead. At the same time, the figures moved closer to that text. Figure being on the same page as the text referring to them rose from 59% to 80%, and text explicitly referencing the figures rose from 87% to 99%. In the 1st edition, a figure was often set beside or apart from the text that explained it. This was the opposite in the 9th. In the end the figure and its explanation, once divided between a caption and a paragraph that might be elsewhere, now sit together.

The 9th edition figures pull in two directions at once. The figure itself got more detailed with more features, so in a way, it explains more of itself. On the other side, its caption got less detailed, and the figure now depends more on the text that explains it. Because the explanation did not disappear, it did not get harder to understand. It just moved from the caption into two new places: the figure's own features and the paragraph beside it. When read next to its

Table 1.2 *Frequency Distribution and Histogram for Motorcycle Data*

Frequency distribution			Histogram
Manufacturer	Frequency	Relative frequency	
1. Honda	41	.34	
2. Yamaha	27	.23	
3. Kawasaki	20	.17	
4. Suzuki	18	.15	
5. Harley-Davidson	3	.03	
6. Other	11	.09	
	120	1.01	

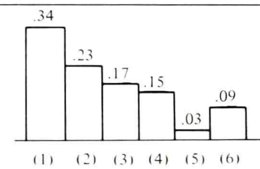

Fig. 6. 1st Edition composite: frequency distribution and histogram (2 panels).

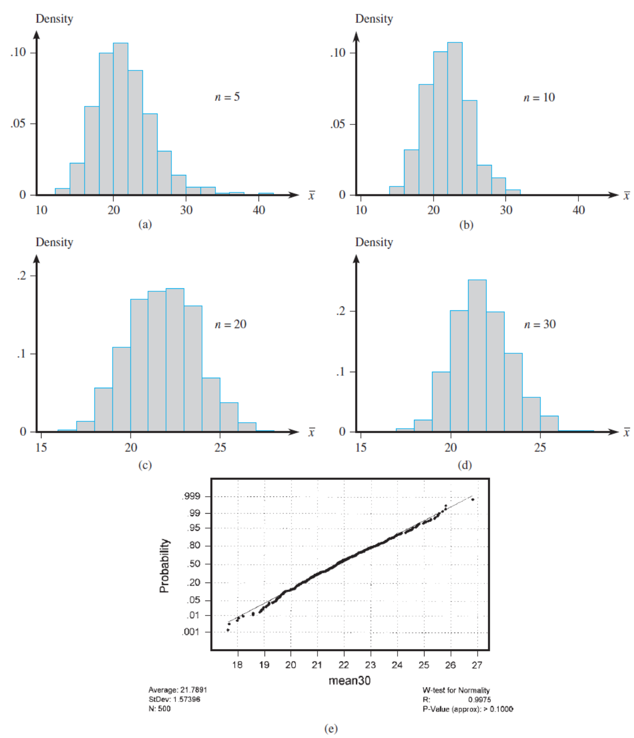

Figure 5.14 Results of the simulation experiment of Example 5.24: (a) $\bar{x}$ histogram for $n = 5$; (b) $\bar{x}$ histogram for $n = 10$; (c) $\bar{x}$ histogram for $n = 20$; (d) $\bar{x}$ histogram for $n = 30$; (e) normal probability plot for $n = 30$ (from Minitab)

Fig. 7. 9th Edition composite: sequence of normal probability plots (5 panels).

paragraph, a 9th-edition figure is as clear as a 1st-edition one, so nothing was lost. What changed was the figure. A figure with a basic caption needs that paragraph next to it in a way that a figure with a full explanatory caption does not. Consequently, the 9th edition figures did not become impossible to read alone. They just stand alone less than the 1st edition figures did.

4.3 Topics determines illustration types

The second research question asks about topics rather than editions. The finding is that the topic leans on one kind of figure for each topic, and the kind fits what the topic asks the reader to do. Descriptive statistics is full of data, so it uses figures that specifically display data like histograms and boxplots. Joint probability is more abstract, so it uses diagrams that build an idea. Regression is about the relationship between two variables, so it uses figures that display a relationship, such as scatterplots.

Each topic has a different leading function Table 3: descriptive statistics leans on figures that display a distribution, joint probability on figures that illustrate a concept, and regression on figures that display a summary. The types of figures sort even more cleanly. Dotplots, stem-and-leaf plots, and boxplots appear only in descriptive statistics. Scatterplots make up 52% of regression figures, and the only other topic they appear in is joint probability. In our codebook, we coded figures as "conceptual diagram" if it explains an abstract concept.Conceptual diagrams make up 41% of joint probability figures.

Function	**Descriptive statistics (n=36)**	**Joint probability (n=34)**	**Regression (n=33)**
Displays a distribution	47%	15%	15%
Illustrates a concept	36%	76%	36%
Displays a summary	11%	3%	30%
Compares groups	11%	24%	24%
Compares representations	11%	3%	3%
Demonstrates a technique	3%	24%	9%

Table 3. Figure function by topic — % of figures coded with each topic, both editions combined. Figures may carry more than one function code, so columns exceed 100%.

Three figures, one per topic, show why the pairing makes sense because each does a different job. A boxplot in descriptive statistics just shows data that the text already gave. It just makes the shape of the data easier to see. It builds nothing new. A conceptual diagram in joint probability does the opposite. It draws the region you integrate over, which the text has a hard time visualizing, so the picture has to carry the idea. A scatterplot in regression shows how one variable moves with another, which is the whole point of the topic.

This pairing is what did not change between the editions. The same topics utilize the same kinds of figures in both 1982 and 2016: descriptive statistics uses "data displays", probability uses "conceptual diagrams", and regression uses "scatterplots". One explanation for this stability is simply that the same author made both editions, so the pairing may be a personal habit that did not change. This study cannot separate that from the topics themselves. What we can say is that, across 34 years in which the figures changed a lot in other ways, such as color, software output, captions, and integration, this one pairing stayed similar.

The changes between editions reached the topics unevenly in the 9th edition. "Software output" makes up about a third of the figures in descriptive statistics and regression, but less than a fifth in probability. color shows a different story: it appears in 75% of descriptive statistics figures, but 91% of probability figures and 86% of regression figures. This points to something counter-intuitive. Probability relies on figures the most to carry its abstract concepts, yet it is the least touched by software output. The reason comes down to the type of figure each topic needs. Probability figures are "conceptual diagrams" where color simply shades the region to integrate, rather than displaying a computed result. In contrast, descriptive statistics and regression produce exactly the kind of output that software generates. So the modernization reached the topics whose figures software could make, and left mostly untouched the one topic whose figures still have to be drawn to show an idea.

5 Discussion

As one can see from the results section, the book changed a lot in terms of presentation, but it did not change the kind of figure each topic uses. In both 1982 and 2016, descriptive statistics uses data displays, probability uses conceptual diagrams, and regression uses scatterplots. What changed is how the figures look, how they are structured on the page, and how much they rely on the surrounding text. The figures were modernized. This shift is most possibly tied in to how production technology has changed in the 34 years between the two editions. This change made a different kind of presentation plausible.

In 1982, producing figures was harder and more limited. The figures were simpler. They were grayscale with limited features. color printing was expensive, which is consistent with the 1st edition having no color at all. Editing the structure of the book was also more difficult with the technology of the time. The advancements in technology most likely lowered these limits. Color printing became cheaper [10]. Color went from absent to standard in the book. Digital typesetting most likely made it easier for authors to edit a book, arrange it on a computer, and keep the figures' references consistent. This is supported by the fact that nearly every figure in the 9th edition had a caption and an explicit reference. Cheaper production and digital tools also made it more convenient to add features that 1st edition rarely include such as titles, grid lines, and legends. Thus the technology set what was plausible for the authors. But it does not, on its own, explain how the book transformed. color is the clearest case of this. Despite the availability, yet the 9th edition does not use it for every figure. It leaves 12 out of 68 figures in grayscale with only some recoloring where it was needed. That selective use is an authorial choice. Technology made color accessible; the author decided where it was used.

A similar split appears with software output, one of the biggest additions of the 9th edition. Software output is not only a product of new production technology. It also reflects how the field of probability and statistics has changed. Statistics is now practiced through software, and statisticians are expected to use it [9,4]. The 9th edition follows this trend. It uses software output regularly, and presents computed results for the reader to read rather than showing how to calculate them by hand. Here as well, technology made the figure easier to implement, but the choice to showing the finished results and teaching the reader how to read them is a pedagogical one.

Our findings align with earlier work on textbook figures. Dimmel and Herbst looked at geometry diagrams across many textbooks from the 20th century and found that the visual features of diagrams change over time [7]. We found a similar result for P&S. The figures in a textbook are not fixed, but change as the way books are produced and how statistics change. The setups of the two studies differ. Dimmel and Herbst used many books by different authors, so differences between authors could affect the figures. We did the opposite by using the same book by the same author. This reduces the effect of the author, but it does not remove the effect since an author's style can change across 34 years. The kinds of change in our results also differ. One of the biggest changes we found is that "software output" is becoming a new type of figure, which does not appear in their study. The two studies support each other. Their study shows that textbook figures change over time, and ours shows what that change looks like for the P&S field in relation to Devore's textbook.

5.1 Limitations

This study only compares certain chapters from the selected editions. Furthermore, the study only uses 1 book with 2 different versions. The findings are only true for the selected chapters of the selected books. Thus, this study cannot generalize findings; it can only compare them with the more general ones from other studies. Another limitation is that there is only one coder, so the codebook is generated with the bias of the researcher. Some codes might be mislabeled, or there might not be enough code variety due to the researcher's lack of creativity. This study also only uses books from 1982 and 2016, so the found themes are limited to that time frame.

5.2 Future Work

The results point to several specific future improvements. The clearest one comes from the "software output" finding. We seek to show that software output presents results instead of showing how to reach them, but we cannot say whether this helps or hurts learning. A study could test whether students learn regression better from an annotated "software output", or from a worked hand-calculation.

Because we only have the 1st and 9th editions, we can see that color and software output appeared, but not when. Coding the editions in between would show when these changes entered the book, and whether they arrived slowly or all at once.

Our most surprising finding was that probability is the least digitized chapter, even though it relies on figures the most. We could only see this in one book. Checking whether the same topic-by-topic pattern holds in other authors' P&S textbooks would show whether it is a property of the subject or just of this author.

Finally, the single-coder limitation could be addressed directly. A second coder applying our codebook, with an inter-rater reliability score, would test how much the codes depend on one person's judgment. New codes could also be added for the things our scheme could not measure, like where color falls in a figure or how many panels a composite has.

6 Conclusion

This study compared the illustrations in two editions of an engineering statistics textbook published 34 years apart. We analyzed 103 figures across three chapters based on nine different categories. We designed this study to separate changes caused by technological change from differences caused by the topic difference. Because the author and the content of the book stayed the same, the changes between the 1982 and 2016 editions can be attributed more confidently to updates in how books are made rather than to the author. Meanwhile, comparing the three chapters showed us how illustration practices depended on the specific topic being taught.

We found that newer figures focus more on showing the final result rather than explaining step-by-step. The 9th edition uses color to highlight results, includes software output, and has more detailed illustrations that are placed closer to the text. However, when looking at different topics, we found that specific topics always require specific types of illustrations. This underlying rule stayed the same across both editions.

Ultimately, these findings show that modern redesign changed almost everything about how the textbook's illustrations look, feel, and organization. However, the redesign changed very little about the underlying content of the illustrations. The modernization was mostly visually driven by digitalization and a shift in the statistics field from hand calculations to computer software. Because this study only looked at one textbook and was coded by one researcher, future research will be needed to see if these same patterns hold true for other authors, subjects, and books.

7 Responsible Research

The study is coded, designed, and conducted by one person only. Consequently, subjective categories rely on this person's personal judgment and bias. It may have been influenced by our early expectation that the newer book would use more color and software. To minimize bias and ensure consistency, we used decision trees. However, because a second researcher did not cross-check the coding, readers should treat these findings as one person's findings rather than a generalization for the whole field. Another independent researcher can try to reproduce our results using our codebook but they might have a different result if they assign some codes to different illustrations.

References

1. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative Research in Psychology **3**(2), 77–101 (Jan 2006). https://doi.org/10.1191/1478088706qp063oa, https://www.tandfonline.com/doi/full/10.1191/1478088706qp063oa
2. Bustamante-Valdés, M., Díaz-Levicoy, D., Alarcón-Bustamante, E.: Analysis of Formative and Evaluative Activities on Statistical Graphs in Textbooks for Chilean Rural Multigrade Education. European Journal of Investigation in Health, Psychology and Education **14**(5), 1396–1412 (May 2024). https://doi.org/10.3390/ejihpe14050092, https://www.mdpi.com/2254-9625/14/5/92
3. Carter, J.: Exploring the fruitfulness of diagrams in mathematics. Synthese **196**(10), 4011–4032 (Oct 2019). https://doi.org/10.1007/s11229-017-1635-1, http://link.springer.com/10.1007/s11229-017-1635-1
4. Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Roswell, G.H., Velleman, P., Witmer, J., Wood, B.: Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016. Tech. rep., American Statistical Association (Jul 2016), https://commons.erau.edu/publication/1083
5. Devore, J.L.: Probability and statistics for engineering and the sciences. Brooks/Cole Publ, Monterey, Calif (1982)
6. Devore, J.L.: Probability and statistics for engineering and the sciences. Cengage Learning, Boston, MA, ninth edition edn. (2016)
7. Dimmel, J.K., Herbst, P.G.: The Semiotic Structure of Geometry Diagrams: How Textbook Diagrams Convey Meaning. Journal for Research in Mathematics Education **46**(2), 147–195 (Mar 2015). https://doi.org/10.5951/jresematheduc.46.2.0147, https://pubs.nctm.org/view/journals/jrme/46/2/article-p147.xml
8. Lester, J.N. (ed.): Approaches to qualitative research: key concepts in qualitative methods. Key concepts in qualitative methods, Routledge, Abingdon, Oxon New York, NY (2026)
9. Moore, D.S.: New Pedagogy and New Content: The Case of Statistics. International Statistical Review **65**(2), 123–137 (Aug 1997). https://doi.org/10.1111/j.1751-5823.1997.tb00390.x, https://onlinelibrary.wiley.com/doi/10.1111/j.1751-5823.1997.tb00390.x

10. Mulcahy, P., Samuels, S.J.: Three Hundred Years of Illustrations in American Textbooks. In: Houghton, H.A., Willows, D.M. (eds.) The Psychology of Illustration, pp. 1–52. Springer US, New York, NY (1987). https://doi.org/10.1007/978-1-4612-4706-7_1, http://link.springer.com/10.1007/978-1-4612-4706-7_1
11. Slough, S.W., McTigue, E.: Development of the Graphical Analysis Protocol (GAP) for Eliciting the Graphical Demands of Science Textbooks. In: Khine, M.S. (ed.) Critical Analysis of Science Textbooks, pp. 17–30. Springer Netherlands, Dordrecht (2013). https://doi.org/10.1007/978-94-007-4168-3_2, https://link.springer.com/10.1007/978-94-007-4168-3_2
12. Tavory, I., Timmermans, S.: Abductive analysis: theorizing qualitative research. The Univ. of Chicago Press, Chicago (2014)
13. Tenny, S., Brannan, J., Brannan, G.: Qualitative Study. In: StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL) (Sep 2022), https://www.ncbi.nlm.nih.gov/books/NBK470395/