

An Examination of Tableau as a Supplement to Excel to Enhance Data Literacy Skills

Mark P. Sena
sena@xavier.edu

Thilini Ariyachandra
ariyachandrat@xavier.edu

Business Analytics and Information Systems Department
Xavier University
Cincinnati, OH USA

Abstract

Data literacy has become a much sought after skill by organizations as the importance of data driven decision making continues to rise in importance. This study explores the teaching of Tableau as a supplement to Excel to enhance the data literacy skills of students. Twenty eight students taking an introductory course in Business Analytics and Information Systems completed a series of five tasks in both Excel and Tableau and answered a two question survey on their perceptions of the relative ease of use and ability to produce correct and visually pleasing results for Tableau versus Excel. An assessment of the results showed that Tableau submissions produced higher scores in terms of correctness and more professionally formatted visualizations. Student survey responses also supported a significant preference for Tableau over Excel in terms of both ease of use and ability to produce correct results.

Keywords: Tableau, Excel, Data Literacy, Analytics, Data Visualization.

1. INTRODUCTION

Recent disruptive changes in the global scape have led organizations to realize the value of data driven decision making. Information is now a necessary ingredient to become nimble and thrive in disruptive business environments. In over 90 percent of organizations, information is explicitly recognized as a critical enterprise asset in corporate strategies (Rollings, Duncan and Logan 2022). Yet its potential remains untapped as an Accenture survey of over 9000 employees revealed that 75 percent are underprepared for the growing use of data in their roles (Vohra 2020). Only 21 percent of survey respondents were confident in their data literacy skills. According to Gartner, one of the two main road blocks to the creation of a data driven business environment is poor data literacy within an organization (Boobier 2021). As a result of these

gaps in data literacy, business leaders become blind to potential business opportunities as well as potential threats and weaknesses (Goasduff 2020).

Today, data literacy is seen as a critical skill for those entering the workforce. A survey of 1200 global C-level executives and 6000 employees indicates that data literacy will be the most in-demand skill by 2030 (QLIK Technologies, Inc., 2022). Over 85 percent of executive respondents believed that data literacy will be vital to the future akin to the ability to use a computer today. A recent Forrester survey further confirm evidence of the growing demand for these skills as survey respondents indicated a 40 percent rise in the heavy use of data by employees in the last four years (Blackborow 2022). The demand for data literacy skills by employers suggests a shift in the hiring requirements as well as a shift in the

existing role requirements in organizations.

Consequently, institutions of higher education need to incorporate data literacy skills into its core curriculum to meet the needs of potential employers. Heeding the call for greater data literacy skills in future employees, recently academic institutions have begun using data visualization course work as a means to infuse data literacy into the curriculum (Batt, Grealis, Harmon, and Tomolonis 2020; Schuff 2018; Usova and Laws 2021). However, it has not been a fluid integration (Hunt 2005; Philip, Olivares-Pasillas and Rocha 2016; Wolff, Gooch, Montaner, Rashid, and Kortuem 2016). The most common IT application used to teach data literacy skills is Microsoft Excel (Slayter and Higging 2018).

The introductory IS course found in the typical undergraduate business curriculum generally contains a sizable share of data skills content taught to students using Microsoft Excel (Modaresnezhad and Schell 2019). More recently, an introduction to data visualization applications, such as Tableau and Power BI, have begun to emerge in the content presented in introductory IS courses. A tool that is used for the sole purpose of visualizing data such as Tableau offers an easy means of supplementing the core data skills students learn through MS Excel. This paper explores the teaching of Tableau as a supplement to Excel to enhance the data literacy skills of students. In so doing, the presents options for instructors hoping to incorporate content to bolster the data literacy skills currently being taught in Introductory IS courses.

To do so, the paper first introduces the concepts of data literacy and its importance. Next, it describes the methodology used to explore the usefulness of Tableau as a supplement to MS Excel skills taught in the introductory IS course. This is followed by a detailed discussion of the study findings. Finally, the conclusion of the study is presented along with limitations and future research options.

2. DATA LITERACY

While a relatively recent addition to the growing body of literacies (e.g., IT literacy and statistical literacy), the definition of data literacy is evolving as the extant universe of data and its characteristics keep growing. Similar to other types of literacy, a simple definition for data literacy is the ability to understand and use data (Frank, Walker, Attard and Tygel 2016). D'Ignazio and Bhargava (2015), express data literacy in terms of the ability to carry out four

specific actions related to data; specifically, reading data, working with data, analyzing data and arguing with data. They describe each of these actions as follows:

- 1) Reading data, which means understanding what data is and the aspects of the world it represents.
- 2) Working with data involves including creating, acquiring, cleaning, and managing it.
- 3) Analyzing data involves filtering, sorting, aggregating, comparing, and performing other analytic operations on it.
- 4) Arguing with data involves using data to support a larger narrative that is intended to communicate some message or story to a particular audience." (2 pg, D'Ignazio and Bhargava 2015).

Within the context of an organization, data literacy has been defined as the ability of employees to understand and work with data to the appropriate degree (Brown 2021). Tableau provides a more comprehensive description to data literacy. They define data literacy as "the ability to explore, understand and communicate with data using critical thinking skills to use, interpret and make decisions; then, convey its value and significance to others" (Tableau 2022). While assumed in the Tableau definition, recent data literacy definitions make specific reference to data visualization. For example, Shreiner and Dykes (2020) state it as the ability to comprehend, analyze, and interpret data and data visualizations. Furthermore, an individual can achieve data literacy by gaining verbal, numerical and graphical literacy (Brown 2022). Graphical literacy, the understanding of how to decode and interpret visual representations of meaning or data visualization (Roberts and Brugar 2017) is a key means by which individuals engage in data literacy today.

The specific inclusion of graphical literacy as part of data literacy suggests that the introduction and use of data visualization applications such as Tableau in an introductory IS course can enhance the content currently being taught in the course. Next, the methodology section next describes the process used to empirically examine the value of a data visualization tool to enhance for data literacy in an introductory IS course.

3. METHODOLOGY

In order to explore the relative effectiveness of teaching Tableau as a compliment to Excel, two sections enrolled in an introductory undergraduate course in Business Analytics and Information Systems were asked to complete an

ungraded test with identical questions using both Excel and Tableau. Students were also asked to complete a two question survey asking their opinion on the ease of use and effectiveness of using the Excel vs Tableau for the test questions. The purpose of the assessment was to collect data for this study while providing students with an opportunity to test their skills in Tableau while relating them to skills previously learned in Excel.

The course is a 200-level class that is required of all business students and is comprised of mostly sophomores and juniors at a private university with approximately 6000 students located in the Midwest of the United States. All students take the course as a requirement of their major. Students are comprised of various business majors along with selected other majors (particularly health and sport management) throughout the university. Prior to the test, students had completed approximately seven weeks of Excel training and approximately two weeks of Tableau training, with the Tableau training coming immediately prior to the test. A major objective of the course is to acquire skills in Excel that are useful across the various majors of enrolled students. Most students have limited exposure to Excel prior to enrollment. After learning the basics and intermediate Excel skills, students are given approximately two weeks of training on data analysis skills (such as tables, pivot tables, and correlation functions). This is followed by the Tableau training that covers similar content resulting in an approximately equal amount of time spent on each tool for the purposes of learning data analysis.

The test asked students to download an Excel file that is available on Kaggle (Kaggle, 2022). The file is a widely used training dataset with fields that include the gas mileage per gallon of various automobile models by year and number of engine cylinders. The questions included the following:

1. Create a chart that shows the average mpg by year grouped in 5 year intervals.
2. Create a chart that shows each car's mpg and cylinders in a scatter plot. Label the model.
3. Create a chart that shows the average mpg by Japanese automakers vs others. These include Datsun, Honda, Toyota, Suburu, and Mazda.
4. Show a list of the top ten cars with the highest mpg.
5. Show a list of models that have at least 6 cylinders and at least 30 mpg.

In the prior section on data literacy, we summarized four major skills of data literacy: reading data, working with data, analyzing data

and arguing with data (D'Ignazio and Bhargava (2015)). In this study, our research questions focus primarily on the relative ability of students to use Tableau and Excel to perform the third set of skills. Because the sample data file was provided to students and did not require any manipulation or cleansing, the first two categories of data literacy were not measured in this study. Due to time constraints, students were not asked to write a narrative explaining their findings. Specifically, our study focuses on the following research questions:

Research Question 1: To what extent are students able to use Excel vs Tableau for filtering, sorting, aggregating, comparing, and performing other analytic operations on a set of five problems?

Research Question 2: To what extent are students able to use Excel vs Tableau to communicate the results of their analysis of the five problems to a particular audience in a clear and professionally formatted presentation?

Each test question was graded on five-point scale for correctness with 5 representing completely correct and 1 representing completely incorrect or not attempted. The questions were also graded for clarity and professionalism of appearance with 5 representing very clear and professional in appearance and 1 representing very unclear or unprofessional in formatting or appearance. The second set of evaluation is intended to measure the extent to which Tableau allows for increased ease in formatting and creating visually pleasing visualizations.

In one course section, students completed the test using Excel first while the other section completed the test using Tableau before Excel. Students took the test in proctored classroom and were required to work individually but were allowed "open book" access to the Internet, prior training materials, etc,. However, the test was timed and limited to fifty minutes. Twenty eight students successfully completed both sections of the test and were included in this analysis. There were a few students who failed to submit both sections or had submissions errors (did not attach the correct file) and were excluded from the dataset.

After completing the tasks in both Excel and Tableau, students were required to upload the files then were asked to complete the following two survey questions:

- Please select the response which best fits your opinion as to which tool was easier to

use:

1. Tableau was much easier to use
 2. Tableau was somewhat easier to use
 3. Tableau and Excel were equally easy to use
 4. Excel was somewhat easier to use
 5. Excel was much easier to use
- Please select the response which best fits your opinion as to which tool produced the better results (correct answer, visually pleasing, etc):
 1. Tableau produced much better results
 2. Tableau produced somewhat better results
 3. Tableau and Excel produced equal results
 4. Excel produced somewhat better results
 5. Excel produced much better results.

4. FINDINGS

1. *Create a chart that shows the average mpg by year grouped in 5 year intervals.*

In order to complete this task in Excel, students would insert a pivot table then select the Year field as a row and the MPG field as the Sum Value field. The Value Field setting for MPG would need to be changed from the default of Sum to Average. The Year field would need to be grouped in fixed intervals of five years. In Tableau (after loading the Excel data source), students would need to create a "Bin" using the Year field to allow for five year sets of groups. Afterwards, students would drag the two fields, Year Bin and MPG onto the row and column sections of the screen then would need to adjust the Measure setting of MPG to Average.

Question: Avg MPG by Yr	
Average of Excel Correct	4.68
Average of Tableau Correct	4.86
P(T<=t) two-tail	0.31
Average of Excel Clarity	4.04
Average of Tableau Clarity	4.39
P(T<=t) two-tail	0.12

Table 1: Question 1 Results

As shown in Table 1, students generally performed well in both Excel and Tableau for this question. The moderately lower scores for Excel clarity are due to the inherent pivot table formatting where numbers need to be formatted, headings need to be replaced from generic ("Row Label") and chart headings need to be added. In Tableau, fields created using Bins default to

showing just the first entry in a grouped range (such as 1970 instead of 1970-1974) but otherwise Tableau creates visually appealing charts more easily than Excel pivot tables.

2. *Create a chart that shows each car's mpg and cylinders in a scatterplot. Label the model.*

In order to create the scatterplot in Excel, students simply needed to select the ranges for both MPG and Cylinders and select the chart option to insert a Scatterplot. Creating labels for the car model would require advanced skills to allow for models to show only when hovered (otherwise the labels would make the scatterplot difficult to read). In Tableau, students would drag to two fields onto the Row and Column but would need to unselect Aggregate Measures from the ribbon to allow for the plotting of individual records instead of summary values (the default would plot just one point). Students would also need to drag the Model field on to the Label option in the Marks section of the screen.

As shown in Table 2, students performed much better using Tableau for this task than using Excel. Many students assumed that the task should be completed using a pivot table (providing average MPG by cylinder), which does not produce the required result.

Question: MPG vs Cylinders Scatterplot	
Average of Excel Correct	2.46
Average of Tableau Correct	4.32
P(T<=t) two-tail	<.001
Average of Excel Clarity	2.36
Average of Tableau Clarity	4.29
P(T<=t) two-tail	<.001

Table 2: Question 2 Results

3. *Create a chart that shows the average mpg by Japanese automakers vs others. These include Datsun, Honda, Toyota, Subaru, and Mazda*

Similar to question 1, this task required students to create a pivot table in Excel with the Auto Maker field in the row and MPG (with a setting of Avg) as the Sum Value field. Students then needed to manually group the Auto Maker range and format the label of the two groups. The challenge with this grouping is that the required records were not in adjacent rows so students would need to hold the Control key (or Command key in Mac) to select multiple entries. In Tableau, students would begin by creating a

group field from the existing Auto Maker field. However, the grouping is conducted in a pop-up window that allows for easy selection of group entries and labels. The remaining task would simply be to drag the grouped Auto Maker field and MPG onto the screen in the row/column and again change MPG measure from sum to average.

As shown in Table 3, students struggled with the task of selecting the required groupings in Excel resulting in much lower scores for both correctness and clarity.

Question: Avg MPG for Japanese automakers vs. other automakers	
Average of Excel Correct	3.21
Average of Tableau Correct	4.46
P(T<=t) two-tail	<.001
Average of Excel Clarity	2.75
Average of Tableau Clarity	4.75
P(T<=t) two-tail	<.001

Table 3: Question 3 Results

4. Show a list of the top ten cars with the highest mpg.

In Excel, the preferred method of completing this task would be to insert a table then apply a top ten filter to the MPG field. Student could also simply create a copy of the raw data and apply a sort and manually select the first ten rows. Students often failed to include a heading for the audience. In Tableau, students would drag the Model and MPG (with measure of Avg) onto the row and column, then drag the Model into the Filter section of the screen and select the Top Ten option (which defaults to top ten largest MPG average).

Question: Top Ten Models with Highest Average MPG	
Average of Excel Correct	4.25
Average of Tableau Correct	4.25
P(T<=t) two-tail	1.00
Average of Excel Clarity	3.71
Average of Tableau Clarity	4.57
P(T<=t) two-tail	.013

Table 4: Question 4 Results

As shown in Table 4, students performed equally well in getting the correct solution (with the sort option counted as an alternative correct result) but with less clarity as compared to the Tableau version. One modest formatting advantage in Tableau is that worksheet names are

automatically used as chart headings so the audience can more clearly understand the result.

5. Show a list of models that have at least 6 cylinders and at least 30 mpg.

This was another task that could be completed in various ways. In Excel, ideally students would insert a table and apply numeric filters to limit the entries that met the requirements. Students could also apply a two level sort and manually select the records. Some students created a new column using either a nested IF or IF(AND) function. As in question 4, students often failed to label or include a heading indicating the purpose of the analysis. In Tableau, there were also multiple ways of achieving the result. The simplest approach was to apply two filters by dragging the MPG and Cylinder fields into the Filter area of the screen then applying the minimum filter options. Some students chose to create a new "Calculated" field where students enter a formula into a popup window. In this option, students would also need to drag the calculated field into the filter and limit the entries to "True".

As shown in Table 5, students struggled with this question in both tools, particularly in the clarity of the presentation. It is important to note that because the test was timed, some students may have had limited time to complete the last question.

Question: List of Models with 30+ MPG and 6+ Cylinders	
Average of Excel Correct	2.68
Average of Tableau Correct	3.36
P(T<=t) two-tail	0.07
Average of Excel Clarity	2.18
Average of Tableau Clarity	2.86
P(T<=t) two-tail	0.04

Table 5: Question 5 Results

Table 6 shows that, overall students produced more correct results using Tableau to complete the tasks and had substantially better clarity in terms of formatting, labels and professional looking visualizations. Students who completed the questions in Tableau prior to Excel actually had higher overall means for correctness in both Tableau and Excel. This may have been due to time constraints because the Excel solutions were less structured and the recent training in Tableau may have allowed those who completed Tableau first to have more time remaining. It may also be

a function of small sample sizes given the relatively small number (28) of students who completed both sections of the test.

	Excel	Tableau	Excel	Tableau
	Correct Solution		Clarity	
Overall Avg (n=28)	3.46	4.25 *	3.01	4.17
Excel First (n=18)	3.34	4.12	3.03	4.17
Tableau First (n=10)	3.66	4.48	2.96	4.18

Table 6: Overall Averages and Order of Tasks

* P(T<=t) two-tail < .001 for all comparisons

After completing the ungraded test, students were asked their opinion as to which tool was produced the better result and which tool was easier to use. As shown in Figure 1, the majority of students found Tableau produced much better results and was much easier to use than Excel for the specific questions in the test.

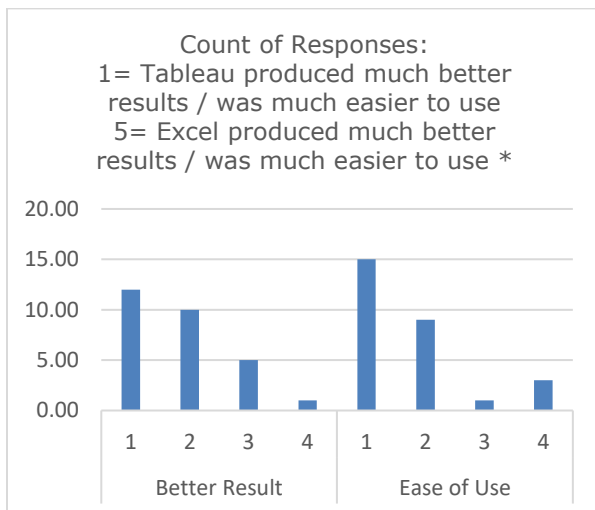


Figure 1: Student Opinions: Tableau vs Excel

* Note: no students selected "5" as a response

It is important to note that students were preparing for a graded skills test in Tableau and the recent training on Tableau could have resulted in some halo effect in the opinions. The questions in the test were appropriate for Tableau analysis. Clearly Tableau is not a replacement for all Excel features as Tableau is not designed for the wide spectrum of Excel uses. For example, students could not easily use Tableau to create a budget or compute payments on a mortgage, etc. Thus, the findings should be limited to data literacy as they pertain to tasks that can be

accomplished easily in Tableau.

5. CONCLUSIONS

As organizations strive to increase data literacy skills among employees, universities play an important role in providing students with a foundation. Software such as Tableau (along with similar tools such as Microsoft's Power BI) can play an important role in expanding the capability of students to manage, cleanse, interpret, and present data in a visual manner that students find intuitive and engaging. Similarly, Excel, which has long been the key application in introducing students to data literacy skills remains an important component in business analytics education.

This study provides evidence in support of using Tableau for both research questions examined in the study. The results support the enhanced ability of using Tableau for filtering, sorting, aggregating, comparing, and performing other analytic operations and to better communicate the results of their analyses in a clear and professionally formatted presentation.

Clearly, the implications of this study are limited by several factors. Our sample size of only 28 students on a small set of five problems provides only preliminary evidence. With problems that are unstructured and result in students providing various solutions, grading on a five point scale as we did in this study is inherently subjective. Our conclusions are also influenced by the specific questions chosen, the prior experience of the students and the manner in which the test was administered. Future studies could focus on establishing a more robust set of questions with more specific grading rubrics and a larger sample size. Particularly for the assessment of clear and professionally formatted presentation, a more descriptive set of criteria is needed to add rigor to the understanding of the importance of this aspect of data literacy.

The results of the study provide some initial evidence that could form the foundation for future studies in the application of data visualization tools for developing data literacy skills. The analysis of the results are also insightful to software makers to show examples of the challenges that users face in solving the sample questions in the test. Additional academic and pedagogical research should focus on the benefits and success factors of using data visualization software to enhance the data literacy of students and employees.

6. REFERENCES

- Batt, S., Grealis, T., Harmon, O., & Tomolonis, P. (2020). Learning Tableau: A data visualization tool. *The Journal of Economic Education*, 51(3-4), 317-328.
- Blackborow, J., (2022). Building Data Literacy. Forrester Consulting. https://www.tableau.com/sites/default/files/2022-03/Forrester_Building_Data_Literacy_Tableau_Mar2022.pdf
- Brown, S. (2021, February 9). How to build data literacy in your company. MIT Sloan. Retrieved July 12, 2022, from <https://mitsloan.mit.edu/ideas-made-to-matter/how-to-build-data-literacy-your-company>
- Boobier, T. (2021, July 7). Perhaps Execs don't need data literacy after all, just good judgement. Customer Insight Leader. Retrieved July 12, 2022, from <https://www.customerinsightleader.com/others/execs-dont-need-data-literacy/>
- Frank, M., Walker, J., Attard, J., & Tygel, A. (2016). Data Literacy-What is it and how can we make it happen?. *The Journal of Community Informatics*, 12(3).
- Goasduff, L. (2020, July 9). Avoid 5 Pitfalls When Building Data and Analytics Teams. Gartner. Retrieved July 12, 2022, from <https://www.gartner.com/smarterwithgartner/avoid-5-pitfalls-when-building-data-and-analytics-teams>
- Hunt, K. (2005). The challenges of integrating data literacy into the curriculum in an undergraduate institution. *IASSIST Quarterly*, 28(2-3), 12-12.
- Kaggle (2022). Auto-mpg dataset. Retrieved July 14, 2022 from <https://www.kaggle.com/datasets/uciml/automp-g-dataset>
- Modaresnezhad, M., & Schell, G. (2019). The soul of the introductory information systems course. *Information Systems Education Journal*, 17(5), 39.
- Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, 34(4), 361-388.
- QLIK Technologies, Inc. (2022, March 22). Data Literacy to be Most In-Demand Skill by 2030 as AI Transforms Global Workplaces. GlobeNewswire News Room. Retrieved July 12, 2022, from <https://www.globenewswire.com/en/news-release/2022/03/22/2407543/0/en/Data-Literacy-to-be-Most-In-Demand-Skill-by-2030-as-AI-Transforms-Global-Workplaces.html>
- Roberts, K. L., & Brugar, K. A. (2017). The view from here: Emergence of graphical literacy. *Reading Psychology*, 38(8), 733-777.
- Rollings, M., Duncan, A., & Logan, V. (2019). 10 Ways CDOs Can Succeed in Forging a Data-Driven Organization. Gartner (G00378249).
- Schuff, D. (2018). Data science for all: a university-wide course in data literacy. In *Analytics and data science* (pp. 281-297). Springer, Cham.
- Shreiner, T. L., & Dykes, B. M. (2021). Visualizing the teaching of data visualizations in social studies: A study of teachers' data literacy practices, beliefs, and knowledge. *Theory & Research in Social Education*, 49(2), 262-306.
- Slayter, E., & Higgins, L. M. (2018). Hands-On Learning: A Problem-Based Approach to Teaching Microsoft Excel. *College Teaching*, 66(1), 31-33.
- Tableau. (n.d.). Data Literacy. Retrieved July 13, 2022, from <https://www.tableau.com/why-tableau/data-literacy>
- Usova, T., & Laws, R. (2021). Teaching a one-credit course on data literacy and data visualisation. *Journal of Information Literacy*, 15(1).
- Vohra, S. (2020, January 16). The human impact of data literacy. Accenture. <https://www.accenture.com/us-en/insights/technology/human-impact-data-literacy>.
- Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12(3)