VAMPIRES AND STAR-CROSSED LOVERS: SECONDARY TEACHERS' REASONING ABOUT THE CONNECTIONS BETWEEN MULTIVARIATE DATA AND VISUALIZATION

CHELSEY LEGACY
University of Minnesota
legac006@umn.edu

ANDREW ZIEFFLER
University of Minnesota
zief0002@umn.edu

V. N. VIMAL RAO
University of Illinois
raovnv@illinois.edu

ROBERT DELMAS
University of Minnesota
delma001@umn.edu

ABSTRACT

As ideas from data science become more prevalent in secondary curricula, it is important to understand secondary teachers' content knowledge and reasoning about complex data structures and modern visualizations. The purpose of this case study is to explore how secondary teachers make sense of mappings between data and visualizations, especially depictions of multivariate relationships. The participants were 14 in-service secondary teachers who were video recorded as they worked through three sets of activities. In these activities, participants created a visualization (network graph) from multivariate data, encoded raw data for several attributes from visualizations depicting multivariate relationships, and structured data into a tidy format. With minimal instruction, participants were able to create visualizations when given data representing multivariate relationships. They were also able to structure non-tidy data into a tidy format with some scaffolding and discussion. Notably, creating data tables from visualizations, especially relational tables, seemed more challenging for them. These results provide insight into secondary teachers' reasoning about connections between multivariate data and visualization.

Keywords: *Statistics education research; multivariate thinking; data science education; teacher training*

1. INTRODUCTION

The prevalence of and reliance on data for decision making has made data literacy an important outcome for active citizenship. Because data visualization is often the medium through which information from data is communicated, it is important that students build proficiency in reading and interpreting information from visualization, especially multivariate data visualizations (Bargagliotti et al., 2021; Franklin et al., 2015; GAISE College Report ASA Revision Committee, 2016; IDSSP Curriculum Team, 2019; ProCivicStat Partners, 2018; Ryan et al., 2019).

Building this proficiency requires that students have multiple experiences with different types of data (e.g., numeric, categorical, text) and the tools and methods used to extract information from that data. At the K–12 level in the United States, however, multivariate exposure has historically been rare, with curricula primarily focused on univariate and bivariate visualizations (e.g., College Board, 2021; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). While some tertiary courses have begun to include multivariate visualization (e.g., Bargagliotti

et al., 2020; Çetinkaya-Rundel & Ellison, 2020; Stander & Dalla Valle, 2017), this is not typical (Ridgway, Nicholson, & McCusker, 2007; Schield, 2004).

Though there is research on the development of students' and teachers' reasoning and the challenges they face when interpreting univariate and bivariate data visualizations (e.g., Pfannkuch, 2007; Shaughnessy & Noll, 2006; Zieffler & Garfield, 2009), the research on students' and teachers' understanding of multivariate visualizations is still nascent. This work often has focused on data visualization literacy (e.g., Engebretsen, 2020; Gould, 2017; Prodromou & Dunne, 2017; Ridgway, 2016). Given the prominence of multivariate visualization and reasoning in several curricular recommendations, it is important to understand how students and teachers develop reasoning around multivariate visualization.

2. BACKGROUND

While the curricular recommendations promote giving students experiences with more complex graphs and multivariate thinking earlier in their educational tenure, several challenges will need to be addressed. For example, in-service teachers, the primary source for introducing concepts to secondary students, often have little experience or training related to working with data structures and complex visualizations that are useful for understanding the relationships within multivariate data (Burgess, 2002). Because these experiences may ultimately be important for helping students develop multivariate reasoning, teachers need opportunities to work with complex data structures that underlie multivariate visualization.

2.1. DATA STRUCTURES UNDERLYING MULTIVARIATE VISUALIZATIONS

Creating visualizations that are multivariate in nature brings with it particular challenges. Computer programs used to create these data visualizations often require the raw data to be in a specific format. For example, most require data tables with standard row-and-column, case-by-attribute organization. However, given the complex nature of some visualizations, the cases and attributes might be difficult to determine (e.g., repeated measures data). Determining cases and attributes requires students to be able to identify the unit of observation (e.g., a person, a person at a medical checkup, a person at a medical checkup at a particular time point). As Kaplan (2018) points out, this is not always obvious and needs to be addressed repeatedly in the curriculum.

One way to incorporate instruction on data structure into the curriculum is to promote reasoning about *tidy data* (Wickham, 2014). Tidy data is a standard for describing the underlying semantics (meaning) of a set of data. At its core, it is based on three principles:

- **Tidy Data Principle #1:** Each variable forms a column, where a variable contains all values (quantitative or categorical) that measure the same underlying attribute across observations.
- **Tidy Data Principle #2:** Each observation forms a row, where an observation contains all values measured on the same unit (like a person, a day, or a school) across attributes.
- **Tidy Data Principle #3:** Each type of observational unit forms a distinct table.

Consider the following example: A teacher collects data from students at multiple time points to examine their learning over time. In the teacher's gradebook (Figure 1, left panel), the data may be recorded with students as cases (rows) and each exam as a separate attribute (columns). While this structure may be useful for quickly scanning how each student is progressing, it may not be ideal for computation, for example, creating a graph that shows a student's exam scores over time. As Murrell (2013, p. 31) states, "[t]he problem is that we inevitably end up wanting to do more with the data, which means working with the data using software, which means explaining the format of the data to the software, which in turn means that we end up wishing that the data were formatted for consumption by a computer, not human eyeballs."

Student	Exam_01	Exam_02	Exam_03
Nilla	5	8	8
Iggy	2	3	5
Sadie	1	4	2
Hank	8	5	7

Student	Exam	Score
Nilla	Exam_01	5
Nilla	Exam_02	8
Nilla	Exam_03	8
Iggy	Exam_01	2
Iggy	Exam_02	3
Iggy	Exam_03	5
Sadie	Exam_01	1
Sadie	Exam_02	4
Sadie	Exam_03	2
Hank	Exam_01	8
Hank	Exam_02	5
Hank	Exam_03	7

Figure 1. Two different formats for organizing a teacher's gradebook data

In the tidy data paradigm (Figure 1, right panel), the gradebook data would be restructured so that all students' exam scores are in a single column. This is consistent with Tidy Data Principle #1, which states that a variable (column) contains all values that measure the same underlying attribute, in this case, all the students' exam scores. The tidy data would also include another column with data about the corresponding exam. Including the exam number as data in a column fixes one of the common problems Wickham identifies, namely that "[c]olumn headers are values, not variable names" (p. 5).

While the tidy data standard can facilitate easier computation, we believe it also has the advantage of making students think more critically about the data and their intentions for analysis. For example, to formally analyze students' learning over time, the case (row) in the data table must be a student at a particular time point. By understanding that the purpose of analysis changes how one chooses to structure data, students may develop a more flexible understanding of data structure. Wickham (2014) also suggests that because the same raw data can be structured in multiple ways, tidy data also provides a way to "describe why the two tables represent the same data" (p. 3).

The third tidy data principle alludes to the fact that sometimes there should be multiple data tables to describe the data fully. For example, educational data might include one data table to record student attributes and a separate data table to describe school attributes. Both tables would also require an attribute that links the two tables. This principle requires an understanding of relational data, a topic that, in our experience, most students rarely encounter in undergraduate coursework, let alone at the secondary or primary levels.

Research on students' understanding of data structure has suggested that ideas about the complex and flexible structure of data may prove challenging for students (e.g., Cobb & Moore, 1997; Pfannkuch et al., 2016). Younger students struggled to reason about case-by-attribute structures and associated data displays (Hancock et al., 1992; Lehrer & Schauble, 2004), while older students were generally able to reason about case-by-attribute structures with multiple variables but struggled to reason about hierarchical or nested data (Lehrer & Schauble, 2000).

In a study that investigated extracting data from visualizations and creating data structures, Konold et al. (2017) presented participants with a depiction of traffic and asked them to produce an organized record of the data. Nearly all middle school and high school students and most of the adults in the study who produced tables (as opposed to narratives) produced either one or a series of flat tables in a case-by-attribute structure, albeit while sometimes violating Tidy Principles #1 and #2. Approximately one-quarter of the adults in the study employed nesting within a single table to denote relationships between cases. Both nesting and serialization of tables encode information hierarchically and efficiently record

information relative to a single flat table. However, nested and hierarchical tables are typically not appropriate data storage formats that can be easily read by a computer and often violate Tidy Principle #3.

2.2. AESTHETIC MAPPINGS: LINKING DATA STRUCTURE AND VISUALIZATION

Wilkinson (2005) introduced a formal grammar that could be used for creating graphs from data called the *Grammar of Graphics*. Within this framework, when data are used to create a visualization, the grammar requires that we define how features in the data are mapped to perceptual aesthetics in the plot (e.g., position, size, color)—aesthetic mappings. An aesthetic mapping is thus the link between attributes in a dataset and their representations depicted visually in a graph.

While explicit instruction and understanding of the entire grammar is likely unnecessary for most students, the idea of aesthetic mappings could serve as a foundation for learning and reasoning about how data visualizations are created and the data structures they represent. For example, students could use substantive questions to identify the attributes from the data they need to consider for a given research question or investigation. Then, they try mapping different attributes in a dataset to aesthetics to see how different relationships between attributes are highlighted (called working from data-to-viz). Conversely, by working from viz-to-data (considering the aesthetics in each plot and how they are mapped to data attributes), students could garner insight into how the raw data need to be structured to create the plot. Providing students with knowledge of aesthetic mappings may provide a generalizable framework to make sense of data and graphs. This is especially important as classroom instruction is unlikely to expose students to all forms of data visualizations and data structures. Instead, students need to be prepared to reason about these extramural cases on their own.

2.3. STUDY RATIONALE AND PURPOSE

The goal of this study is to begin to understand how in-service teachers' knowledge and reasoning about data structures and corresponding plots develop as they complete a sequence of intentionally designed activities. In particular, we ask to what extent can these teachers:

- use multivariate data to create a visualization that allows them to make sense of the potential multivariate relationships?
- reason from a data visualization depicting multivariate relationships to the raw data used to create the visualization?
- produce tidy data from a data visualization depicting multivariate relationships?

3. METHODS

We used Groth's (2013) theoretical framing of Statistical Knowledge for Teaching (SKT) as a guide to situate and interpret the results of this study. Based on Ball et al.'s (2008) model of Mathematical Knowledge for Teaching, SKT categorizes knowledge for teaching as subject matter or pedagogical content knowledge and includes hypotheses about the relationships between constructs within this framework. The focus of the professional development activities used in this study was to build teachers' subject matter knowledge by exposing them to content outside the current curriculum (specifically, what Groth refers to as "horizon knowledge").

Interpreting results within this framework also helps us identify where teachers are making conceptual advancements in their reasoning and knowledge—which Simon (2006) refers to as key developmental understandings (KDUs). This, in turn, gives us insight into how to refine the materials used in the Professional Development (PD).

To answer the research questions, we developed a set of activities to build teachers' horizon knowledge around multivariate data and visualization. Because of the exploratory nature of the work, we employed a case study to investigate the development of this horizon knowledge in in-service teachers. This methodology is appropriate given our research goal of describing the extent to which teachers reason about this content as they interact with the PD materials. While not generalizable, the insights gained from this case study can further research by helping to identify and describe KDUs

pertaining to multivariate reasoning. These insights could also inform the development of curricular material for teacher PD of specific KDUs.

3.1. PARTICIPANTS AND PROCEDURE

The participants were 14 in-service secondary teachers with a broad range of teaching experience (10–30 years) currently teaching statistics as part of the University of Minnesota College in the Schools (CIS) program. At the time of the study, they all taught utilizing the CATALST curriculum in their secondary schools (see Zieffler & Huberty (2015) for more about the CIS program; see Justice et al. (2020) for more about the CATALST curriculum). Nearly all have taught since the 2015–2016 academic year. Additionally, a couple of the teachers had upwards of 15 years of experience teaching statistics at the secondary level. Despite their experience, none of the participants had an advanced degree in statistics, and the formal statistical preparation of the participants prior to the CIS program was limited to one or two undergraduate courses. This lack of formal coursework in statistics is consistent with secondary mathematics teachers more broadly in the United States (Franklin et al., 2015).

Study participants were video recorded and observed as they worked through multiple activities during three PD sessions. The first two PD sessions were conducted remotely, over Zoom, and the third was conducted in person. Video recordings, observer notes, and artifacts from the PD sessions were analyzed for empirical traces of participants' understanding and reasoning. The study obtained approval from the University of Minnesota IRB: STUDY00010778.

3.2. PD ACTIVITIES

A set of PD activities was designed to build teachers' horizon knowledge around multivariate data and visualization. The tasks in each activity were developed to promote learning outcomes that were directly connected to the research questions (e.g., use multivariate data to create a visualization that allows them to make sense of the potential multivariate relationships). These tasks were written and sequenced to build on their prior knowledge and experiences, as well as to elicit participants' reasoning. Participants worked collaboratively on these activities in keeping with the culture established in this professional cohort. Figure 2 offers a timeline of the three PD sessions and provides an outline of activities in each of the PD sessions. All PD activities can be found at: https://laser-umn.github.io/posts/d2g.html.

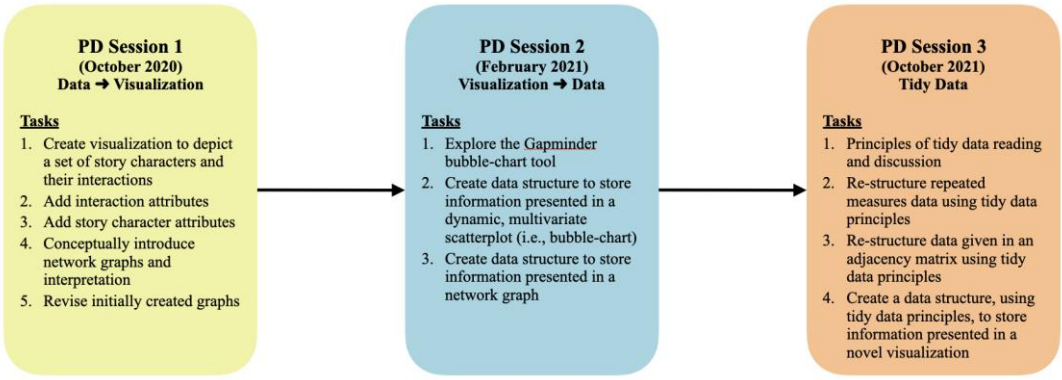


Figure 2. Timeline showing the chronology of the tasks within and across the PD sessions

Activity 1 The tasks in the initial PD session (referred to as Activity 1) had participants create a network graph (a data visualization not introduced in the CATALST curriculum nor typical introductory statistics curricula) from multiple data tables by encoding different aspects of a multivariate relationship. Participants were asked to use the visualizations they created to respond to a series of contextual questions to understand how they reason about information and relationships embedded in the plot. We did this via a structured sequence of five tasks based on character relationships in the

Twilight movies. Figure 3 depicts the tasks and learning outcomes for Activity 1. It also includes the participants' prior knowledge and experiences we assumed when designing the tasks in this activity.

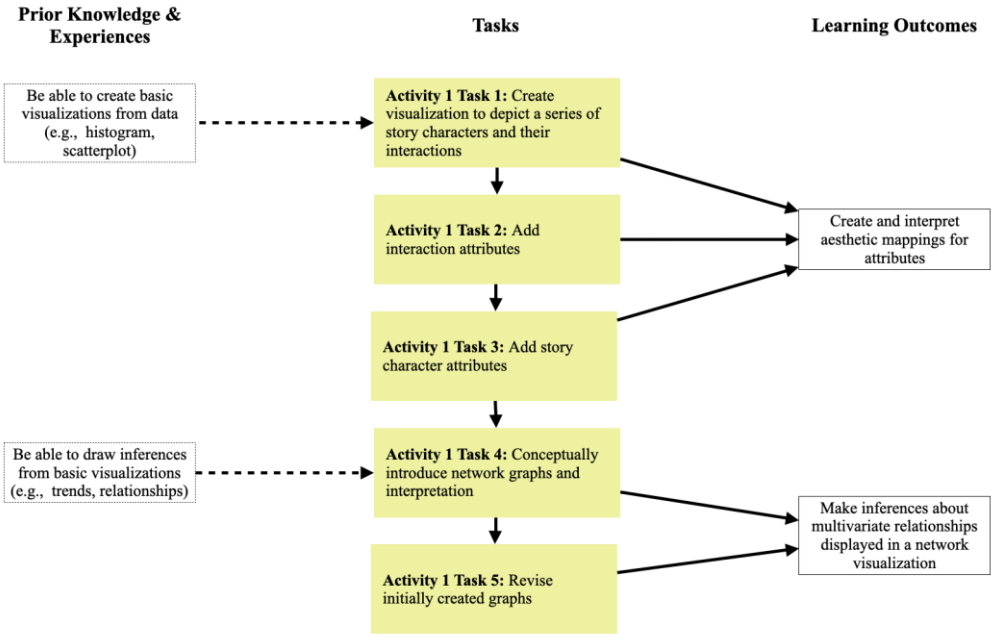


Figure 3. Diagram of the participants' assumed prior knowledge/experiences along with the tasks and learning outcomes for Activity 1. Dashed lines indicate the hypothesized role of prior experience. Solid lines indicate the designed contribution of each task to future tasks and learning outcomes.

The initial task provided participants with a single flat data file containing information about the interactions between movie characters from the Twilight movie series. Participants were asked to devise a way to visualize the interactions between movie characters with no specific requirements for how this should be achieved. In the second and third tasks of the activity, participants were asked to modify their previous visualization (or create a new visualization) to incorporate additional characteristics for the interactions between characters and additional characteristics for the characters themselves.

In the fourth task, participants were provided a network graph representing the interactions between players in Shakespeare's Romeo & Juliet. Participants were asked to identify what the nodes and edges represent and to identify node and edge attributes and their aesthetic mappings. This task was meant to serve as a scaffold for participants' knowledge of network graphs, to prompt reflection about the visualizations they created in earlier tasks, and to serve as an exemplar for subsequent activities. Participants were then asked (in Task 5) to specifically create a network graph of the interactions between Twilight movie characters using all the data presented in earlier tasks. Finally, each participant was asked to complete Task 5 individually, outside the PD. Participants were asked to interpret the network graph their group created in Task 4 by responding to contextual questions related to character relationships.

Activity 2 In the second PD, participants extracted the raw data for several attributes from two visualizations depicting multivariate relationships—a bubble plot and a network graph. For each of these visualizations, they were asked to identify and organize the underlying data. Tasks 1 and 2 utilized Gapminder's World data bubble plot, and Task 3 returned to the Romeo & Juliet network graph presented in Activity 1. Figure 4 depicts the tasks and learning outcomes for Activity 2. It also includes the participants' prior knowledge and experiences we assumed when designing the tasks in this activity.

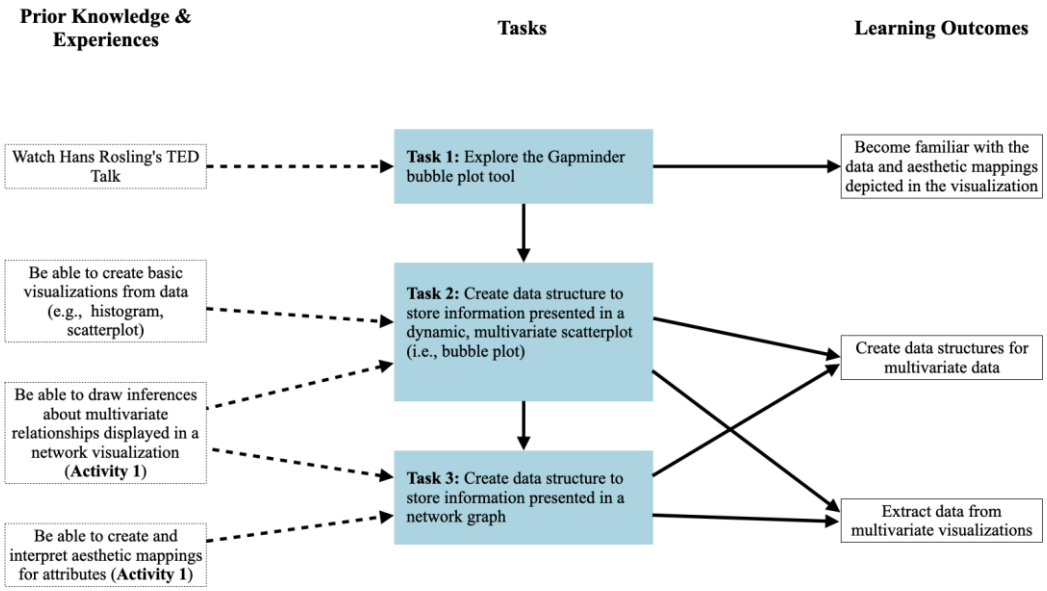


Figure 4. Diagram of the participants' assumed prior knowledge/experiences along with the tasks and learning outcomes for Activity 2. Dashed lines indicate the hypothesized role of prior experience. Solid lines indicate the designed contribution of each task to future tasks and learning outcomes.

In Task 1, participants acquainted themselves with the Gapminder World data bubble plot using the default chart featuring countries' life expectancy, per capita GDP, region, and population. They spent time toggling the interactive elements of the plot, highlighting the cases to learn more about them, and using the animation to watch how the variables interacted over time. They were asked a series of questions prompting them to describe relationships between two variables and then explore how those variables changed over time.

Participants were then asked to consider the aesthetic mappings used in the bubble plot. After identifying the attributes from these mappings, the participants were asked to create a data table that they thought could be used to create such a plot. In this table, they were asked to record three years' worth of data for 11 different countries. The final item in this activity asked participants to write instructions (pseudocode) for creating a similar bubble plot using the data recorded in their table. The purpose was to investigate how the participants determined which attributes were needed to recreate the plot, how they mapped those to aesthetic features in the plot, and how they structured the raw data in their data tables.

After completing Task 2, the participants were shuffled into different groups and asked to share their data tables and instructions for creating the bubble plot. After discussing this, they engaged in Task 3. Within this activity, the participants again had to record the raw data in a data table(s) needed to recreate a multivariate visualization, the Romeo & Juliet network graph, and provide a set of instructions or pseudocode that someone else could use to create the graph using the data in their tables.

Activity 3 In the third PD, participants were introduced to the tidy data principles (through an assigned reading) and then asked to tidy the data structures they created in the second PD session. In the culminating activity of this PD, participants were given the traffic visualization used in Konold et al. (2017) and asked to produce a tidy data structure of the underlying data. Figure 5 depicts the tasks and learning outcomes for Activity 3. It also includes the participants' prior knowledge and experiences we assumed when designing the tasks in this activity.

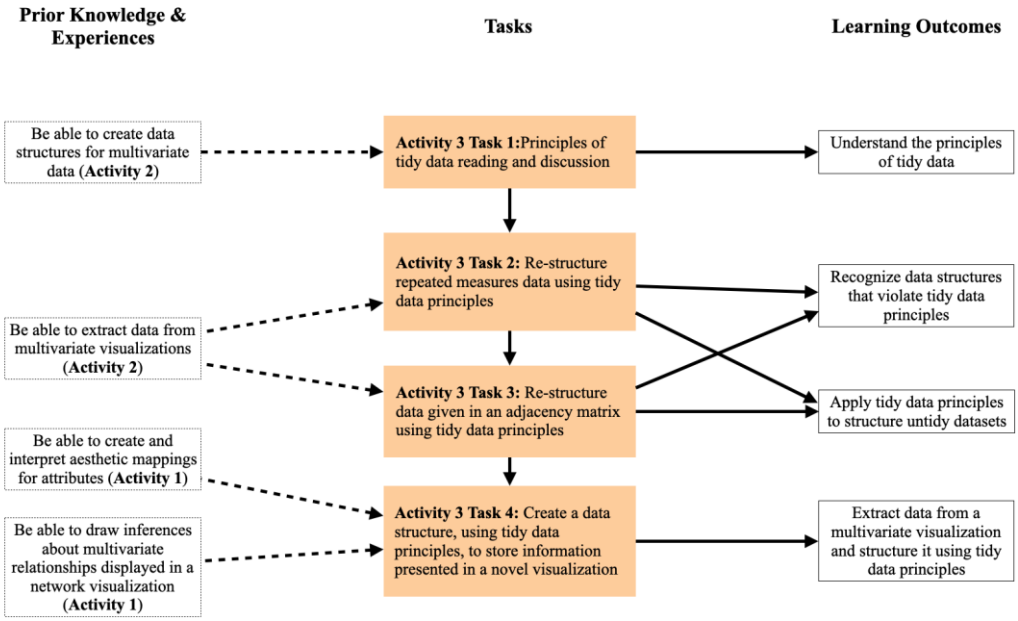


Figure 5. Diagram of the participants' assumed prior knowledge/experiences along with the tasks and learning outcomes for Activity 3. Dashed lines indicate the hypothesized role of prior experience. Solid lines indicate the designed contribution of each task to future tasks and learning outcomes.

For the first half hour, the participants worked in small groups to discuss the assigned tidy data reading. This discussion focused on identifying the importance of tidy data and describing its key features. Then, for Task 2, the participants were given a non-tidy dataset created by one group in Activity 2 Task 2. They were asked to tidy the data based on the three tidy principles. The participants again worked in small groups to restructure the data table into a tidy data table on the whiteboards around the classroom. A whole group discussion followed once all groups were finished.

Next, the participants were shown a data table created by a group in Activity 2 Task 3 and again asked to tidy the data. They continued to work in small groups and displayed their answers on whiteboards around the classroom. Another whole group discussion of the nuanced difficulties of tidying this dataset followed.

Finally, in Task 4, the participants were given a copy of the traffic visualization from Konold et al. (2017). They continued to work in small groups to determine how to record the data from the visualization into a tidy data table.

3.3. ANALYSIS

We approach this research from an interpretivist paradigm, believing that teachers' knowledge is socially constructed through their interactions (Guba & Lincoln, 1994). Our primary goal within this research was to gain insight into how these teachers construct knowledge to make sense of the relationships between multivariate data and visualization as they progress through a series of scaffolded activities. To do this, we analyzed the teachers' discussions and artifacts produced (e.g., visualizations) as they completed the PD activities in small groups of 3–4 teachers. A secondary goal was to identify potential Key Developmental Understandings (KDUs) that would be pivotal in developing teachers' reasoning about connections between multivariate data and visualizations.

Data collected included (1) video recordings of each group of teachers and their shared screen in Zoom as they worked on activities, (2) observer notes taken by a member of the research team for each group of teachers, and (3) the teachers' answers to the activity prompts and their work artifacts.

The data were analyzed using an inductive reflective thematic analysis. In this variant of thematic analysis, the codes and themes that are used to identify and interpret patterns in data are fully generated by the content and data itself, without the initial influence of an outside theory (e.g., Braun & Clarke,

2006; 2019). This type of analysis is a non-sequential iterative process composed of generating, coding, reviewing, and refining thematic elements in the data (Saldaña, 2016).

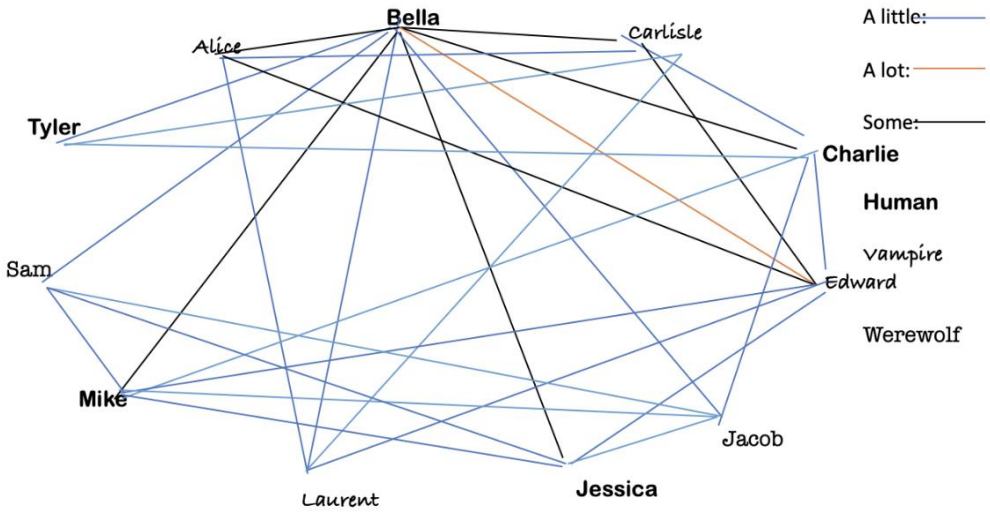
To begin the analytic process, each co-author observed and took notes on a group of teachers as they completed the activities. At the conclusion of each PD session, all co-authors met to discuss and share evidence related to teachers' reasoning. These observations and the authors' discussion after PD sessions served to familiarize all authors with the data we had collected. To begin the coding step in our analysis, we reviewed the groups' work artifacts and observer notes, which helped us identify themes related to the teachers' construction of knowledge and reasoning throughout these activities. The first and third co-authors then watched all video recordings to review the empirical evidence in light of these elements in an effort to review and refine them. Then, all authors came together to ensure there was consensus about the identified themes and corroborating evidence. To answer our research questions, we provide a structured narrative summary of participants' reasoning based on the thematic elements identified in our analysis.

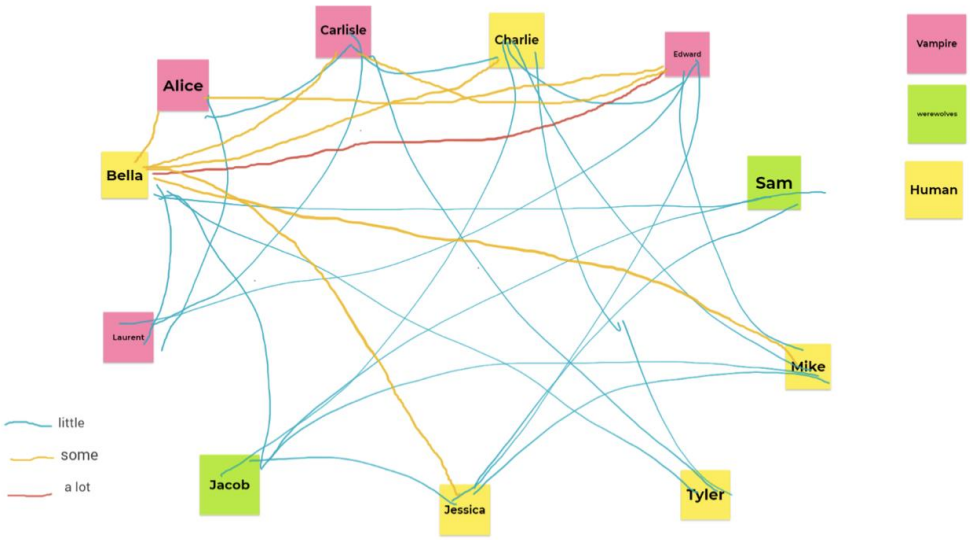
4. RESULTS

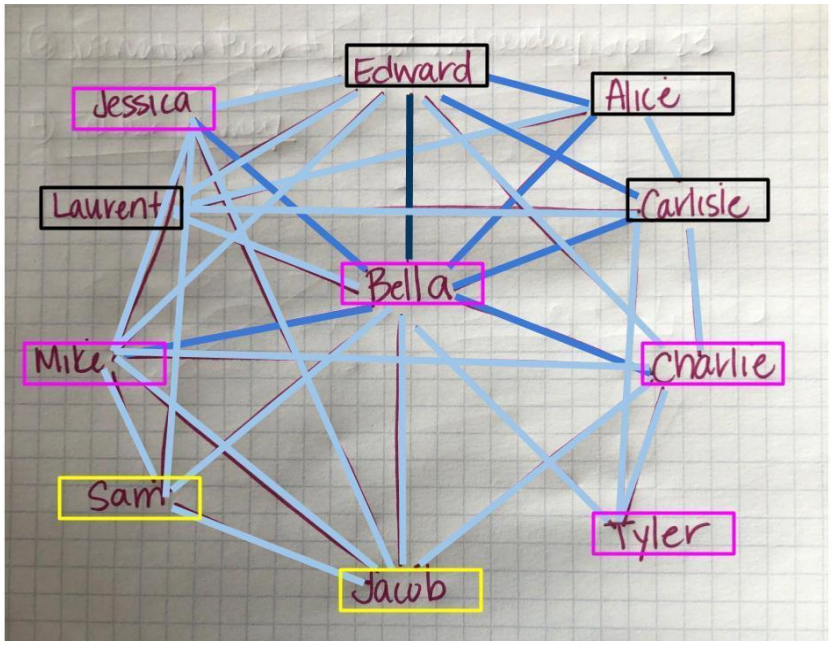
This section presents the results for each research question with examples, figures, and quotes from the participants.

4.1. TO WHAT EXTENT CAN THESE TEACHERS USE MULTIVARIATE DATA TO CREATE A VISUALIZATION THAT ALLOWS THEM TO MAKE SENSE OF RELATIONSHIPS WITHIN THE DATA?

All but one of the groups created a network graph for Activity 1 (Tasks 1-3), despite many not having been introduced to this type of visualization in their formal education and given no instruction to do so in the activity. The one group that did not create a network graph created an adjacency matrix instead. All the groups' final visualizations are shown below.







	Alice	Bella	Carlisle	Charlie	Edward	Jacob	Jessica	Laurent	Mike	Sam	Tyler
Alice		2	1	0	2	0	0	1	0	0	0
Bella	2		2	2	3	1	2	1	2	1	1
Carlisle	1	2		1	2	0	0	1	0	0	1
Charlie	0	2	1		1	1	0	0	1	0	1
Edward	2	3	2	1		0	1	1	1	0	0
Jacob	0	1	0	1	0		1	0	1	1	0
Jessica	0	2	0	0	1	1		0	1	1	0
Laurent	1	1	1	0	1	0	0		0	0	0
Mike	0	2	0	1	1	1	1	0		1	0
Sam	0	1	0	0	0	1	1	0	1		0
Tyler	0	1	1	1	0	0	0	0	0	0	

	human / human-human interaction
	vampire / vampire-vampire interaction
	werewolf / werewolf-werewolf interaction
	human-werewolf interaction
	human-vampire interaction
	vampire-werewolf interaction
0	NO interaction
1	A LITTLE interaction
2	SOME interaction
3	A LOT interaction

Figure 6. Group visualizations created for Activity 1

The groups that created network graphs began by creating all the graph nodes that represented characters and then added edges to identify interactions between characters. In positioning the graph nodes, participants who were familiar with the Twilight series drew on their contextual knowledge and placed Bella near the center of their network graph. For example, one participant noted, *"I [thought] 'Oh, I'll start with Alice' and then I realized right away that was going to be a disaster, so then I used the context and I said, 'Well, Bella, Edward, and Jacob are going to have the most'."* Those participants who were not familiar with the Twilight series seemed to use the data to make decisions about node placement, as exemplified by one participant who stated, *"[be]cause Bella interacts with everyone it might make sense visually to have her in the center."*

There was little variation in how the groups represented the interactions between characters. Each group drew a line connecting the nodes of the interacting characters (see Figure 6). No discussion indicated this choice was intentional, so we speculate this is likely because of the proximity of the initial node placement rather than a feature they wanted to display.

Interestingly, the group that created the adjacency matrix (see the bottom panel of Figure 6) first considered a network graph to visualize the data. They abandoned this after deciding that such a plot would not allow them to calculate numerical summaries of the data (which we did not ask them to do in the activity). As stated by one participant:

> I started drawing [a network graph] on my tablet and then I abandoned it and switched to google sheets … [since] I knew it was going to be a big mess…and I was like, I bet we're going to have to do some sort of calculation or summarization.

Regardless of the visualization they produced, the participants were able to modify their initial visualizations by mapping additional attributes to visual aesthetics in their plots. When making these modifications, most groups used color to represent these attributes. While every group employed color to visualize at least one aesthetic, one group used font type to encode character species, and another group used numeric values to encode the frequency of interactions between characters. Even the groups that adopted color discussed alternatives. For example, when adding character species to their visualization, one group briefly considered changing the shape of the node, suggesting that they could *"insert a little picture of a human, or vampire, or werewolf—create a symbol for each one."* Similar alternatives were discussed when groups added the frequency of interactions between characters. One group considered varying the thickness of the edges, another considered annotating the visualization with text to indicate the frequencies, and the group creating the adjacency matrix mapped the frequency of interactions to numerical values (1–3) and included them in the cells of their matrix.

Participants also seemed to make thoughtful choices about the specific encodings used to visualize character species and frequency of interactions, with many relating the encodings to the problem context. For example, the group that chose font type to encode character species wanted to select fonts that invoked the essence of each species; *"vampire, let's see, like a scrolly font … like a creepy one … or something gothic."* Another group used similar reasoning when adopting their color palette, saying, *"shouldn't the vampires be red for the blood?"*

This reasoned choice of encoding was also employed when considering the frequency of character interactions. Two groups specifically discussed how to use color to encode the ordinal nature of these data. One group employed a color palette that encoded interactions with higher frequency to rainbow colors with a longer wavelength, *"[we decided to] go with the ROY G BIV idea, and the more reddish it gets the more interactions they have."* Another group also briefly considered the rainbow palette before deciding to use various shades of blue, saying:

> I think 'light-to-dark' does a better job of graduating—the more intense the color, the more the interaction. And I think we could pick shades in between if we ended up with five levels, instead of just picking rainbow colors that don't show that as much.

4.2. CONSIDERATIONS AFTER NETWORK GRAPH INTRODUCTION

After being introduced to network visualizations more formally in Task 4, participants were able to connect and integrate some of what they had seen in the Romeo & Juliet graph into the visualizations they created in Task 5. However, this seemed somewhat dependent on the contextual similarities of the two activities. For example, after seeing how the Romeo & Juliet network graph mapped the frequency of character interactions to edge thickness, most groups suggested that they could have done the same in their Twilight graph. On the other hand, after seeing how size was used to map a characteristic of the nodes (number of lines) in the Romeo & Juliet graph, many participants were not able to consider other characteristics aside from number of lines that could be mapped to node size. One participant stated, *"[w]e don't have that information though on our characters [from Twilight]."*

Those who tried to employ node size considered creating a new variable to measure the total number of individuals a character interacts with (degree of the node). Upon reflection, one participant stated that *"if I would have had more context to follow I think I would have made different circles for the people bigger or smaller depending on the number of interactions, so Bella would have been a big huge circle, right, 'cuz she's connected with everyone."* This statement seems to suggest that participants are aware that multiple attributes can be mapped to a single node by varying multiple aesthetics, in this case, both color (for species) and size (for number of characters interacted with) of the node.

After being formally exposed to a network graph, the participants also reconsidered the validity of the adjacency matrix as a visualization of the data. Even though the adjacency matrix does depict the multivariate relationships in the data, the participants seemed to gravitate toward the network graph when revisiting their initial visualization from Tasks 1-3.

4.3. INTERPRETATION OF NETWORK GRAPHS

Despite initially creating a network graph to depict the Twilight characters and interactions (Activity 1 Tasks 1–3), some groups questioned the visualization's utility, making comments such as "that's a lot of lines," "I have no idea what my web really says," and "[e]verything is muddled." Much of this confusion subsided after they were formally introduced to network graphs. At this point, participants were generally able to identify and interpret key features of the graph, including the identification of each aesthetic mapping. Moreover, nearly all participants were able to make sense of the multivariate relationships depicted in their initial visualizations to draw inferences about the characters and their relationships. For example, one participant noted that "other than Romeo and Juliet the interactions usually stay within the same house." However, this understanding was not complete among all participants.

Participants also considered the inferences they could make from the arrangement and positioning of nodes, although no data or instructions related to edge length or node position were presented to participants. For example, in Activity 1 Task 4, their reasoning seemed to be contextual and related to the location of the nodes for Romeo and Juliet, with most participants recognizing that "[t]he main characters are front and center."

After acknowledging that the node positioning for Romeo and Juliet may be relevant to the interpretation of the graph, participants made different inferences about the position of the other nodes. Some thought that the "position of the nodes in two-dimensional space represent a higher quantity of, or stronger relationships between those characters." Relatedly, another participant noted, "the ones that interact a lot are really close together too … so it's the distance between them". Others weren't as sure, saying, "Romeo and Juliet are in the center for a reason, but I don't know if the other placements mean anything," or "[w]e notice that Romeo and Juliet are central but there is no evidence that the rest of the placements have meaning."

While in some network graphs the node positioning and edge length are meaningful, they were not in the Romeo & Juliet graph. We suspect that some of the participants' belief that node positioning and edge length were meaningful stemmed from their creation of the initial graph in Task 1. In creating that visualization, many of the groups made explicit decisions about these attributes (especially node positioning). This decision-making may have led them to believe these attributes were also intentional in the Romeo & Juliet graph. Interestingly, no participant recognized that other attributes presented to them could be used to evaluate many of their hypotheses about these attributes. For example, they had

been asked to map the frequency of interaction to edge thickness, which many thought was also represented in edge length.

4.4. TO WHAT EXTENT CAN THESE TEACHERS REASON FROM A DATA VISUALIZATION DEPICTING MULTIVARIATE RELATIONSHIPS TO THE RAW DATA USED TO CREATE THE VISUALIZATION?

In Activity 2 Task 2, most groups began creating a data table of information for several countries over three different years. They recorded the data for a single year in a case-by-attribute table, with each country in a different row and the attributes (Country, Life Expectancy, Income, Region, and Population) as separate columns. As the groups recognized that there were multiple years of data for each country, there was quite a bit of discussion about how to incorporate multiple years of data into the table.

All the groups initially considered recording the data in multiple tables. However, three of the four groups ultimately decided to organize the data into a single table because they felt it would be a more efficient way to record the information, with one participant noting, "There has to be a better way than making three tables." When probed by one of the researchers inquiring why they chose one table over several, another participant stated:

> To me they are conceptually the same. We could have a data table for each country, year, we just happen to have it in one because we are so efficient. But I don't think there is any difference. As far as having the data all in one place I don't think there's any advantage to having more than one [table].

Another participant further explained, "This is a database, not the data needed to create an individual graph."

Despite adopting a single table to record the data, all three groups created a different table. Two groups opted to include multiple columns per year to record the data (e.g., wide format). One of these groups included the year information in the variable name (Figure 7, panel A), while another used a header denoting year spanned over the other variables (see Figure 7, panel B). A third group included the variable Year as a separate column and incorporated the data by adding multiple rows for each country (e.g., long format). In this organization, each row represents the data for a country each year.

A

	1907 Income	1982 Income	2019 Income	1907 Pop	1982 Pop	2019 Pop	1907 life	1982 life	2019 life
Algeria	1890	11400	14000	5200000	7970000	7970000	29.5	65.1	78.1
Andorra	N/A	29200	N/A	N/A	39100	N/A	N/A	78.3	N/A
Bolivia	1660	4030	7150				34.1	56.7	73.3

B

	A	B	C	D	E	F	G	H	I	J	K
1			1907			1982			2019		
2		World Region	Population	Income	Life Expectancy	Population	Income	Life Expectancy	Population	Income	Life Expectancy
3	Algeria										
4	Andorra										
5	Bolivia										

Figure 7. Cropped example data table from two different groups. (A) Table with year included in the variable name. (B) Table with year spanned across each variable

The fourth group created one table per country, putting each country's data in a different tab. This group also organized the data in each table so that each row represented the data for a country in a given year. As they considered how to record multiple years of data for each country, one participant wondered, "[h]ow are we going to collect subcolumns… we almost need a three-dimensional table? Or do we need … separate worksheets for each [country]?" This discussion led to the group's decision to divide the information into different tabs in their Google Sheet.

4.5. CREATING A DATA TABLE BASED ON A NETWORK GRAPH

Teachers were also asked to create a data table to record the data portrayed in Activity 2 Task 3. Though the focus on data extraction was consistent with the aim of the previous task, we note that organizing the data into a table for this visualization seemed more challenging for participants than organizing the Gapminder data. The participants seemed to be able to identify the variables at the character and interaction level. However, many groups were again focused on storing the data in a single table. Some participants noted, *"I'd prefer to have only one sheet,"* or *"I think that's the easiest way to put all the data in one spot."* At the suggestion of creating a tab for each character, one participant explained, *"I don't like [when] they are in separate tabs. I've got to look down [at the tab name] to see who they are talking to."*

Two of the groups decided to organize the interaction-level data into rows, repeating character-level data in these rows (see Figure 8).

A

Player	House	Number of Lines	Interaction	Line Thickness
Tibault	Capulet	< 50	Capulet	1
Mercutio	Montague	200-399	Romeo	2
Mercutio	Montague	200-399	Benvolio	4
Juliet	Capulet	400+	Nurse	4
Juliet	Capulet	400+	Friar Laurence	3
Juliet	Capulet	400+	First Watchman	1
Juliet	Capulet	400+	Lady Capulet	2
Juliet	Capulet	400+	Romeo	4
First Watchman	Other	< 50	Second Watchman	1
First Watchman	Other	< 50	Third Watchman	1
First Watchman	Other	< 50	Juliet	1
Lady Montague	Montague	< 50	Benvolio	1
Lady Montague	Montague	< 50	Prince	2

B

Character	Number of Lines	House	Interaction
Romeo	400 +	Montague	Juliet - a lot
Romeo	400 +	Montague	Paris - a lot
Romeo	400 +	Montague	Servant - very little
Juliet	400 +	Capulet	Nurse - a lot
Juliet	400 +	Capulet	Lady Capulet - some

Figure 8. Example tables of interaction data from two different groups.

The third group wanted to include the interaction- and character-level data in a single table but did not want to repeat data across rows. Their solution was to create an adjacency matrix (see Figure 9) to represent the interactions. They added character-level attributes to the matrix by using aesthetic characteristics such as font size and color.

	Fryer Laurence	Nurse	Juliet	Romeo	Paris			
Fryer Laurence	x	0	3	4	1			Interaction Key
Nurse	0	x	4	2	0		4	A lot of lines together
Juliet	3	4	x	4	0		3	Many lines together
Romeo	4	2	4	x	3		2	Some lines together
Paris	1	0	0	3	x		1	Very little lines together
							0	Less than 30 lines together
		House						
		Montegeau		Font Size 18	400+ Lines			
		Capulet		Font Size 14	200-399 Lines			
		Other		Font Size 10	50-199 Lines			
				Font 6	Less Than 50			

Figure 9. Adjacency matrix that depicts interactions between characters for the Romeo & Juliet network graph in a single table

The fourth group also created an adjacency matrix to organize the interaction-level data. They initially added columns to this matrix depicting the character-level data for the character in each row of the adjacency matrix. After being prompted by a researcher to explain why they had done this, they noted that there was a difference in the type of data that was depicted in the adjacency matrix side of the table. Upon further consideration and discussion, they decided it would be clearer if they put the character-level data into a separate table (see Figure 10).

		Table 1	Interactions		
	Romeo	Juliet	Mercutio	Capulet	Friar Laurance
Romeo	0	1	1	0	1
Juliet	1	0	0	0	1
Mercutio	1	0	0	0	0
Capulet	0	0	0	0	0
Friar Laurance	1	1	0	0	0
	1 = an interaction of at least 30 lines				
	0 = no interaction				

	Table 2	
	House	Number of Lines
Romeo	Montage	400+
Juliet	Capulet	400+
Mercutio	Montage	200-399
Capulet	Capulet	200-399
Friar Laurance	Montage	200-399

Figure 10. Adjacency matrices with interactions and character-level data in separate table

4.6. TO WHAT EXTENT CAN THESE TEACHERS PRODUCE TIDY DATA FROM A DATA VISUALIZATION DEPICTING MULTIVARIATE RELATIONSHIPS?

When presented with untidy data from Activity 2 Task 2, all groups were able to transform it into a tidy data table. Some groups discussed the relevance of column order but ultimately concluded it did not matter, and all groups produced a data table with an organization similar to that seen in Figure

11. (Note: Because the groups did not record complete data for each observation, we are making an inference that the data would be tidy.)

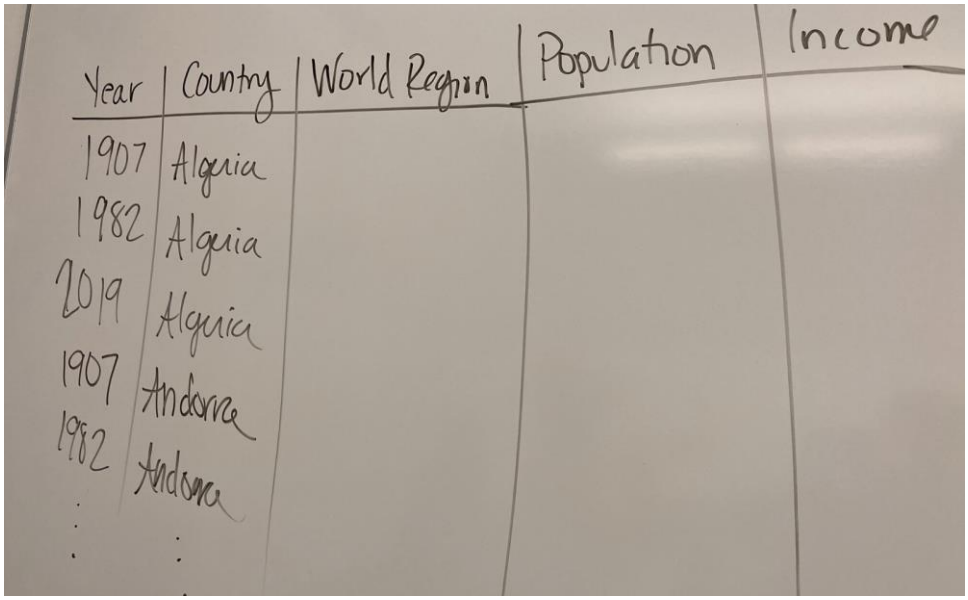


Figure 11. Example of a tidy data table for the Gapminder Bubble data.

The participants were also able to tidy an untidy data table based on Activity 2 Task 3. Again, this seemed to be more challenging, and the groups took more time to complete this task. Most groups started with a data table similar to that in Figure 12A. Again, their discussions focused on how to incorporate all interaction and character-level data into a single table and eliminate redundant information. A considerable group discussion in which they were reminded of the principles of tidy data prompted them to consider using separate tables for each observational unit. After this group discussion, all the groups split the data into two tables: one for character-level data and another for interaction-level data (see Figure 12B). They also included a key column (e.g., ID) in each table to connect information across the tables.

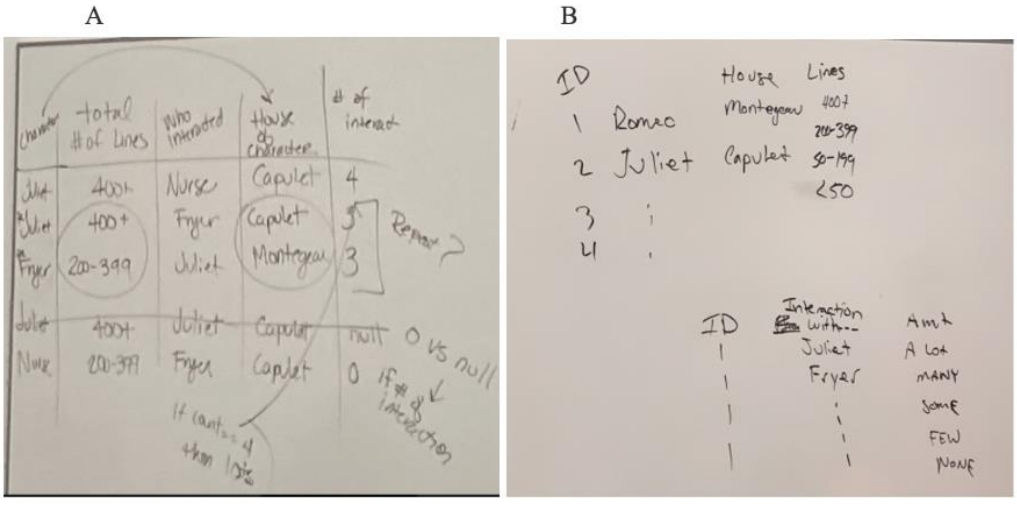


Figure 12. Examples of data tables for the Romeo & Juliet data. (A) Initial attempt at a tidy data table incorporating all interaction and character-level data. (B) Tidy data table with separate linked tables for character-level and interaction-level data.

When creating tidy data for the traffic data visualization, the participants, primed with the tidy principles from the previous two tasks, quickly determined that they needed multiple tables and discussed the observational units for each table. All groups determined they needed separate tables for road segment data and car data.

Interestingly, many groups included redundant or extraneous information in their tables. For example, the data tables created by one group included a summary of which vehicles were on each segment of the road in their road table, which is information already captured (albeit not summarized) in the car table. Another group aimed to eliminate redundant row information by creating three tables: (1) road segment, (2) vehicle speed and distance, and (3) vehicle type and direction. In their attempt to avoid redundant information within a single table, they created multiple tables for the same observational unit (vehicle), which violates Tidy Principle #3. They also introduced redundant information to their Segment column when they created the Road Segment ID column.

5. DISCUSSION

This set of activities designed for the PD sessions exposed in-service secondary statistics teachers to multivariate data visualizations consistent with current guidelines and recommendations for teaching statistics. As the participants engaged in the activities, they gained experience creating visualizations from multivariate data and considering the data underlying such a visualization. In doing so, they also made decisions about how the data should be structured into a format that computer programs typically require to create these data visualizations. This explicit exposure and practice are critical to developing teachers' horizon knowledge that can support their instruction of multivariate visualizations and data structures.

5.1. SUMMARY

We saw evidence from Activity 1 that the teachers were able to produce network graphs from data provided in a table and intuitively reason about the relationships depicted by the nodes and edges in their visualization. Yet this reasoning seemed less intuitive when they did not produce the visualization. For example, when presented with the Romeo and Juliet network (Activity 1, Task 4), many participants confused the frequency of interaction between two characters, an edge attribute, with the number of other characters a particular character interacted with, a node attribute. This points to a potential KDU related to interpreting network graphs and making sense of multivariate relationships:

1. The need to understand and differentiate node and edge attributes, especially in graphs they did not create.

When interpreting network graphs, the teachers were prone to identifying aesthetic variation even when such differences were coincidental and not related to any information provided in any legend. For example, many participants commented on the length of an edge despite this having no interpretational value in Activity 1, Task 4. In this same activity, teachers interpreted the node positioning as meaningful when, in fact, the node position did not serve as an aesthetic mapping in this activity. This points to another potential KDU related to interpreting network graphs and making sense of multivariate relationships:

2. The need to understand which visual elements in the graph are meaningful (i.e., map to variables in the data) and which elements are not (i.e., artifacts of creating the network).

When working on the activity to create data tables from graphs, the teachers did not demonstrate a preference to organize the data into tables that are efficiently structured for computation. The teachers instead focused on the efficiency of both table creation and human extraction of information from their tables (i.e., seeing all the data in one place). This is consistent with the findings of Lehrer and Schauble (2000) and Konold et al. (2017), which state that adults may struggle to reason about hierarchical or nested data and naturally prefer flat tables rather than a series of relational tables. One group went as far as encoding aesthetic features into their table to avoid creating multiple tables. While this might be

a function of human efficiency, it might also point to a lack of experience with complex data formats. This points to a potential KDU:

3. The need to understand the difference between data structures efficient for computation and those for human consumption.

When working to produce tidy datasets, the teachers easily created a table from the bubble plot. However, they needed additional scaffolding and reminders about the Tidy Data Principles when creating tables from the network graph. This is likely because the data structure for the bubble plot had a single observational unit, which could be represented in a single data table. In contrast, the network graph had two observational units, requiring the creation of two linked data tables. The teachers also had difficulty reasoning about and creating data tables with multiple observational units in Activity 3 Task 4, including redundant information across their tables. This activity made it clear that even though they were aware of the Tidy Data Principle, which states that each observational unit forms a distinct table, this was not enough for the teachers to apply it consistently. This points to a potential KDU related to creating tidy data tables:

4. The need to identify observational units in a visualization and understand which visual aesthetics correspond to each observational unit.

5.2. IMPLICATIONS

Given the prevalence of computing and focus on data science in the discipline of statistics, it is important that introductory statistics instructors have the horizon knowledge related to data structures and visualizations. We need additional research to determine what professional development would enhance teachers' reasoning. Moreover, the field of data science and computing is changing rapidly, so we need to understand not only what teachers currently understand but their propensity to reason and adapt as they encounter new complex visualizations or data structures.

One opportunity for additional research is to investigate these KDUs further. For example, it could be important to understand if these KDUs are generalizable beyond this group of teachers. Additionally, the KDUs could be used to better inform the learning content and scaffolding in the activities, which could then be studied to evaluate their efficacy for developing participants' reasoning about the connections between data and visualizations. This could include proposing a learning trajectory and investigating the scope and sequencing of activities needed to support this development. It would also be important to study whether these tasks and activities are appropriate for students who likely have less experience and knowledge than the teachers who participated in this study.

There are also some potential implications for teaching, including that while they were able to correctly answer almost all questions, demonstrating a high level of basic fluency despite little to no explicit training and instruction besides the scaffolding activities, there were points of confusion that posed difficulties to teachers, indicating that they need more experience and training before teaching these concepts in the classroom. Additionally, teachers may need more hands-on experience working with data analysis software to fully appreciate the need for tidy data structures. For a fuller discussion of these implications, see Rao et al. (2023).

5.3. LIMITATIONS

This study was conducted in the Fall of 2020 and 2021 and the Spring of 2021 amid the COVID-19 pandemic. Given the immense strain high school teachers were under at that time, and how the study data was collected virtually, we acknowledge many limitations to this study. The PD was conducted via Zoom, which was not a familiar environment to all teachers as many used Google Meet more frequently. They were limited in their choice of tools for designing their graphs and recording their tables of information. There is no way to know how much the technology impeded their choices for how they responded to the prompts given to them. It was clear on occasion that they had ideas for creating graphs and tables that were difficult, time-consuming, or impossible to implement virtually.

Although we conclude that they preferred one table to many tables, this could have been a function of trying to complete the task within the virtual environment.

Another limitation was that the teachers had limited resources to complete each task during each PD. For example, they only used applications such as Jamboard or Google Sheets because they were easy to share over Zoom. These tools may have limited their choices in aesthetic mapping, potentially explaining their preference for color as it was the easiest to add to their visualization. Because of the time limitations of the PD, the teachers may not have had enough time to fully engage in some of the activities. For example, in Activity 1 Task 5, most groups were too short on time to do more than briefly describe a few aesthetic mappings for the network graph they would create.

6. CONCLUSION

With the increased prevalence of complex modern data visualizations depicting multivariate relationships, horizon knowledge for secondary teachers is important. This research was able to identify four potential KDUs related to supporting the development of teachers' reasoning about multivariate data visualization and the data structures commonly used to create them: (1) understand and differentiate node and edge attributes (in network graphs), (2) understand which visual elements are meaningful and which are not, (3) understand the difference between data structures for computation and those for consumption by humans, and (4) identify observational units in a visualization and identify their corresponding aesthetics. Further research on teachers' and students' understanding of these KDUs and their implementation in curricular materials can promote teachers' horizon knowledge, ultimately supporting students' learning about multivariate data and visualizations.

REFERENCES

Ball, D., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Bargagliotti, A., Arnold, P., & Franklin, C. (2021). GAISE II: Bringing data into classrooms. *Mathematics Teacher: Learning and Teaching PK-12*, *114*(6), 424–435. https://doi.org/10.5951/MTLT.2020.0343

Bargagliotti, A., Binder, W., Blakesley, L., Eusufzai, Z., Fitzpatrick, B., Ford, M., Huchting, K., Larson, S., Miric, N., Rovetti, R., Seal, K., & Zachariah, T. (2020). Undergraduate learning outcomes for achieving data acumen. *Journal of Statistics Education*, *28*(2), 197–211. https://doi.org/10.1080/10691898.2020.1776653

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Burgess, T. (2002). Investigating the "data sense" of preservice teachers. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics. Cape Town, South Africa*. International Statistical Institute and International Association for Statistics Education. https://iase-web.org/documents/papers/icots6/6e4_burg.pdf?1402524962

Çetinkaya-Rundel, M., & Ellison, V. (2020). A fresh look at introductory data science. *Journal of Statistics Education*, 1–11. https://doi.org/10.1080/10691898.2020.1804497

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801–823. https://doi.org/10.1080/00029890.1997.11990723

College Board. (2021). *AP Statistics Course Overview 2021*. https://apcentral.collegeboard.org/courses/ap-statistics/course

Engebretsen, M. (2020). From decoding a graph to processing a multimodal message: Interacting with data visualisation in the news media. *Nordicom Review*, *41*(1), 33–50. https://doi.org/10.2478/nor-2020-0004

Franklin, C., Bargagliotti, A., Case, C., Kader, G., Scheaffer, R., & Spangler, D. (2015). *Statistical Education of Teachers* (p. 88). American Statistical Association. http://www.amstat.org/asa/files/pdfs/EDU-SET.pdf

GAISE College Report ASA Revision Committee. (2016). *Guidelines for Assessment and Instruction in Statistics Education College Report 2016* (p. 143). http://www.amstat.org/education/gaise

Gould, R. (2017). Data literacy is statistics literacy. *Statistics Education Research Journal*, *16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209

Groth, R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, *15*(2), 121–145. https://doi.org/10.1080/10986065.2013.770718

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, *2*(163-194), 105.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist*, *27*(3), 337–364. https://doi.org/10.1207/s15326985ep2703_5

IDSSP Curriculum Team. (2019). *Curriculum Frameworks for Introductory Data Science*. http://www.idssp.org/files/IDSSP_Data_Science_Curriculum_Frameworks_for_Schools_Edition_1.0.pdf

Justice, N., Le, L., Sabbag, A., Fry, E., Ziegler, L., & Garfield, J. (2020). The CATALST Curriculum: A story of change. *Journal of Statistics Education*, *28*(2), 175–186. https://doi.org/10.1080/10691898.2020.1787115

Kaplan, D. (2018). Teaching stats for data science. *The American Statistician*, *72*(1), 89–96. https://doi.org/10.1080/00031305.2017.1398107

Konold, C., Finzer, W., & Kreetong, K. (2017). Modeling as a core component of structuring data. *Statistics Education Research Journal*, *16*(2), 191–212. https://doi.org/10.52041/serj.v16i2.190

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning*, *2*(1–2), 51–74. https://doi.org/10.1207/S15327833MTL0202_3

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, *41*(3), 635–679. https://doi.org/10.3102/00028312041003635

Murrell, P. (2013). Data intended for human consumption, not machine consumption. In *Bad Data Handbook* (pp. 31–51). O'Reilly Media, Inc.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards Mathematics*.

Pfannkuch, M. (2007). Year 11 students' informal inferential reasoning: A case study about the interpretation of box plots. *International Electronic Journal of Mathematics Education*, *2*(2), 149–167. https://doi.org/10.29333/iejme/181

Pfannkuch, M., Budgett, S., Fewster, R., Fitch, M., Pattenwise, S., Wild, C., & Ziedins, I. (2016). Probability modeling and thinking: What can we learn from practice? *Statistics Education Research Journal*, *15*(2), 11–37. https://doi.org/10.52041/serj.v15i2.238

ProCivicStat Partners. (2018). Engaging Civic Statistics: A call for action and recommendations. ProCivicStat. https://iase-web.org/islp/pcs/documents/ProCivicStat_Report.pdf

Prodromou, T., & Dunne, T. (2017). Data visualisation and statistics education in the future. In *Data Visualization and Statistical Literacy for Open and Big Data* (pp. 1–28). IGI Global. https://doi.org/10.4018/978-1-5225-2512-7.ch001

Rao, V.N.V., Legacy, C., Zieffler, A., & delMas, R. (2023). Designing a sequence of activities to build reasoning about data and visualization. *Teaching Statistics*, *45*(S1), S80–S92. https://doi.org/10.1111/test.12341

Ridgway, J. (2016). Implications of the data revolution for statistics education. *International Statistical Review*, *84*(3), 528–549. https://doi.org/10.1111/insr.12110

Ridgway, J., Nicholson, J., & McCusker, S. (2007). Reasoning with multivariate evidence. *International Electronic Journal of Mathematics Education*, *2*(3), 245–269. https://doi.org/10.29333/iejme/212

Ryan, L., Silver, D., Laramee, R. S., Ebert, D., & Rhyne, T.-M. (2019). Teaching data visualization as a skill. *IEEE Computer Graphics and Applications*, *39*(2), 95–103. https://doi.org/10.1109/MCG.2018.2889526

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3E [Third edition]). SAGE.

Schield, M. (2004). *Statistical Literacy Curriculum Design*. 54–74. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.144.8102&rep=rep1&type=pdf

Shaughnessy, M., & Noll, J. (2006). School mathematics students' reasoning about variability in scatterplots. In S. Alatorre, J. L. Cortina, M. Sáiz, & A. Méndez (Eds.), *Proceedings of the 28th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2).

Simon, M. A. (2006). Key developmental understandings in mathematics: A direction for investigating and establishing learning goals. *Mathematical thinking and learning*, *8*(4), 359–371.

Stander, J., & Dalla Valle, L. (2017). On enthusing students about big data and social media visualization and analysis using R, RStudio, and RMarkdown. *Journal of Statistics Education*, *25*(2), 60–67. https://doi.org/10.1080/10691898.2017.1322474

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, *59*(10). https://doi.org/10.18637/jss.v059.i10

Wilkinson, L. (2005). *The grammar of graphics* (2nd ed.). Springer.

Zieffler, A., & Huberty, M. D. (2015). A catalyst for change in the high school math curriculum. *CHANCE*, *28*(3), 44–49. https://doi.org/10.1080/09332480.2015.1099365

Zieffler, A. S., & Garfield, J. B. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, *8*(1), 7–31.

CHELSEY LEGACY
56 E River Road
Minneapolis, MN 55455