Data-to-graphs and back: Secondary teachers' reasoning about aesthetic mappings in data visualizations

Rao, V.N.V., Legacy, C., Brown, J.M., Zieffler, A., & delMas, R.

Introduction

As ideas from data science become more prevalent in secondary curricula, it is important to understand secondary teachers' reasoning about data attributes and their corresponding visual aesthetics. The purpose of this multiple case study is to begin to explore how these teachers make sense of aesthetic mappings between data and graphs depicting multivariate relationships.

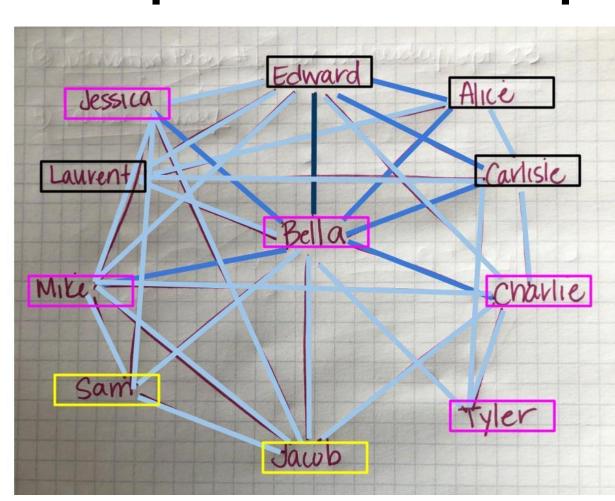
Study 1: Data to Graphs

In our first study we asked secondary teachers to create a graph from a series of data tables. Teachers' creations were nearly identical in aesthetic characteristics. Below is an example of one group's creation. The data provided represented information about interactions between characters from the movie *Twilight*.

Data Provided

	Frequency of interactions:			· S	
Character:	A little	Some	A lot		
Alice	Carlisle, Laurent	Bella, Edward			
Bella	Jacob, Laurent, Sam, Tyler	Alice, Carlisle, Charlie, Jessica, Mike	Edward	8	
Carlisle	Alice, Charlie, Laurent, Tyler	Bella, Edward		Species:	Character:
Charlie	Carlisle, Edward, Jacob, Mike, Tyler	Bella		Human	Bella, Charlie, Jessica, Mike, Tyler
Edward	Charlie, Jessica, Laurent, Mike	Alice, Carlisle	Bella	Vampire	Alice, Carlisle, Edward, Laurent
Jacob	Bella, Charlie, Jessica, Mike, Sam			Werewolf	Jacob, Sam
Jessica	Edward, Jacob, Mike, Sam	Bella			*
Laurent	Alice, Bella, Carlisle, Edward				
Mike	Charlie, Edward, Jacob, Jessica, Sam	Bella			
Sam	Bella, Jacob, Jessica, Mike			1	
Tyler	Bella, Carlisle, Charlie			1	

Example Teachers' Graph



Given multivariate data, teachers were able to create effective visualizations.

However, given a visualization, they had difficulty producing the multivariate data structure.

Conclusions & Future Work

In the first study, teachers were able to easily create visualizations to depict given data and intuitively incorporated many easy-to-interpret aesthetic elements (e.g., color). However, they faced more challenges when asked to create data tables from a given graph. Two moments of impasse led to a variety of decisions and paths of reasoning: (1) whether or not characters should repeat across rows, and (2) how to add character attributes to an adjacency matrix.

Teachers' decisions led most of our participants away from a case-by-variable form required by many computer programs to create visualizations. This indicated that the teachers' would require more scaffolding to reason about complex data structures related to innovative visualizations.

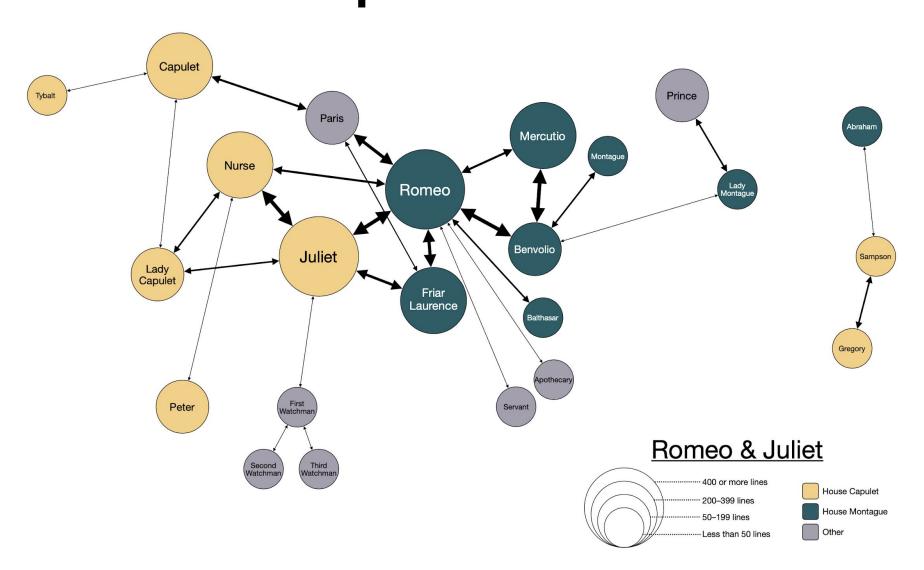
Given how important computing is for modern statistics courses, additional training and research is required to ensure teachers possess the pedagogical content knowledge to teach data science skills, such as formatting data for computer software.



Study 2: Graphs to Data

In our second study we asked the same participants to create data tables from a graph. As opposed to the first study, teachers' creations greatly differed in both data structure and their underlying reasoning. The graph depicted interactions between characters from the play *Romeo and Juliet*.

Graph Provided



Repeated Attributes Information

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

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

Some teachers put each interaction on its own line within a single data table, and repeated character information across rows.

Visual Aesthetics for a Data Table

	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	х	4	0	3	Many lines together
Romeo	4	2	4	x	3	2	Some lines together
Paris	1	0	0	3	х	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		

Some teachers avoided repeating a character across rows by creating an adjacency matrix, and chose to add character attributes to the matrix by giving the table aesthetic characteristics such as size and color.

Multiple Data Tables

		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 interactio	n			

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

Some teachers appended character attributes as extra columns in their adjacency table, and with gentle prompting separated the information.

