VISUALIZATION RUBRIC RUBRIC

	STRONG	MEDIUM	WEAK
COVERS THE RIGHT CONTENT	The content of the rubric represents the best thinking in the field of data visualization about what it means to create meaningful and truthful graphics.	Much of the content of the rubric represent the best thinking in data visualization, but there are a few places that are questionable.	It is difficult to tell what the rubric is intended to assess and the content is far removed from the best thinking in data visualization.
	The content of the rubric aligns with the objectives of the course and covers issues of credibility, ethics, aesthetics, and accessibility.	Some features don't align with the objectives of the course and omit features related to credibility, ethics, aesthetics, or accessibility.	The rubric does not alight with the objectives of the course and does not cover features related to credibility, ethics, aesthetics, or accessibility.
CRITERIA ARE WELL ORGANIZED	The rubric is divided into easily understandable criteria. The number of criteria reflects the complexity of assessing data visualizations.	The number of criteria needs to be adjusted a little—either a single criterion should be made into two criteria, or two criteria should be combined.	The rubric is a long list of everything; there is no organization. The rubric looks like a brainstormed list.
	The details that are used to describe a criterion go together; you can see how they are facets of the same criterion.	Some details that are used to describe a criterion are in the wrong criterion, but most are placed correctly.	The rubric seems mixed up. Descriptors that go together don't seem to be placed together; things that are different are put together.
	The relative emphasis on various features of performance is right—things that are more important are stressed more; things that are less important are stressed less.	The emphasis on some criteria or descriptors is either too small or too great; others are okay.	The rubric is out of balance—features of more importance are emphasized the same as features of less importance.
	The criteria are independent. Each important feature appears in only one place in the rubric.	Some features are included in more than one criterion.	Features are so redundant that the criteria don't really cover different things.
NUMBER OF LEVELS	The number of levels of quality used in the rating scale makes sense. There are enough levels to accurately rate the visualization, but not so many levels that it is impossible to distinguish among them.	More levels would be helpful to make finer distinctions in the assessment of the visualization.	The number of levels is not appropriate—there are either so many levels that it is impossible to reliably distinguish between them, or too few to make important distinctions.
LEVELS DEFINED WELL	Each score point is defined with indicators or descriptions.	Only the top level is defined; the other levels are not defined.	No levels are defined; the rubric is little more than a list of categories to rate followed by a rating scale.
	There is enough descriptive detail to allow you to match the visualization's performance to the right score.	There is some attempt to define terms and include descriptors, but some key ideas are fuzzy in meaning.	Wording of levels (if present) is vague or confusing.
	Two independent users raters assign the same rating most of the time.	It is uncertain if two independent testers, even with practice, could assign the same rating most of the time.	It is unlikely that two independent raters could assign the same rating.
	Wording is descriptive, not evaluative.	Wording is mostly descriptive, but there are a few instances of evaluative labels.	Wording is evaluative and not descriptive (e.g. the visualization is "interesting" or "above average")
PARALLEL LEVELS	The levels of the rubric are parallel in content—if a feature is discussed in one level, it is discussed in all levels. If the levels are not parallel, there is a good explanation why.	The levels are mostly parallel in content, but there are some places where there is a feature at one level that is not present at the other levels.	Levels are not parallel in content and there is no explanation of why.

Adapted from Bonnie B. Mullinix, "Rubric for Assessing Rubrics" (December 2003), accessed September 18, 2017, http://www.asu.edu/courses/asu101/ asuonline/temp/rubric_%20for_rubrics.pdf and Judith A. Arter and Jan Chappuis, *Creating & Recognizing Quality Rubrics* (Boston: Pearson, 2006).