**Flash Data Collection and Analysis: A First Day Activity**

Joanna S. Hunter

Radford University

**Abstract**

This first day activity for an introductory class is designed to introduce students to the discipline of sociology, particularly its research methods. Students engage in “flash data collection” by quickly filling out a card about a classmate and collecting data on their attributes, then they use that data to draw inferences about their subjects on a variety of behavioral outcomes. This activity introduces concepts such as variables, attributes, correlation/causation, and the role of empirical data collection and analysis in the discipline of sociology.

**Learning Goals**

After completing this activity, students will have been introduced to the concepts of variables, attributes, correlation and causation, and the role of empirical data collection and analysis in the discipline of sociology.

**Goals Assessment**

Students will demonstrate their familiarity with the concepts listed above by using them appropriately in Phase 3 of the exercise.

**Procedure**

Prior to the first day of the semester, I create a spreadsheet of the student roster with the student names and two numbers that depended on the enrollment of the class. The roster is divided into two groups, and each student’s name is associated with two numbers: one in the lower 50% of the course enrollment and one in the upper 50% of the course enrollment. For example, in a class of 90 students, each student will have a number between 1 and 45 and one number between 46 and 90. The procedure described below is based on a class of 90.

I then create index cards for each student enrolled in the class. Each student’s name is listed at the top of the card, and each card has the two numbers listed in two different colors for ease of explanation. At the beginning of class on the first day, I distribute adhesive name tags to the students as I distribute syllabi and other documents and ask students to write their first and last name on their nametag and attach it to the front of their bodies. I distribute an attendance sheet that lists all names of enrolled students, to be sure that each student present has a notecard with their name on it – if students on the list are not present, I remove that card from the stack, and if a student who is not on my roster is present, I quickly create a new card for them. I go about the business of the first day of class, going over the syllabus and introducing the material and the course website. Then, I introduce the activity.

I begin by shuffling the notecards and distributing them at random to the students. I tell students that they should have a card with a name on it that is not familiar to them. If, by chance, they get their own card, or the card of a friend, I ask them to switch with a neighbor. I display an example card and call their attention to the two numbers printed on the card in two different colors (red and green in this example). I then display the example respondent card pictured below and included in the accompanying slide deck.

|  |  |
| --- | --- |
|  |  |

I explain that the exercise includes several phases, during which each student will have the opportunity to act as a *responder* and as a *collector* and that when we have completed Phases One and Two, that they will be holding a card that contains data from a student in the class on the ten variables listed on the example respondent card and that a different student will have data about them. I specify that students should not rely on their own perception of their respondents’ attributes, but should take care to ask each question. I find that variables are self-explanatory and familiar to students as conceptual ideas, and specify that “undecided” is an acceptable response for College Major, and that “Group Affiliation” can be defined however the respondent wishes to do so – Greek organization, affinity club, political party – or could even be “none.”

For Phase One, I draw their attention to one of the two numbers printed on the card they are holding (here, using two colors – red and green). During Phase One, the red number on the card is relevant. If the red number on the card they are holding is between 1 and 45, they are to be data *collectors* in Phase One; they are to find the person whose name is listed on their card and survey them. If the red number on the card they are holding is between 46 and 90, they are to be *subjects* in Phase One; they are to stay in their seats and wait for a classmate to find them and survey them. When they are finished collecting data from their respondents, Phase One *collectors* return to their seats to signal that they have completed their task.



Once all students have returned to their seats, Phase Two begins. During Phase Two, the green number is relevant: if the green number is between 1 and 45 (in a class of 90 students), they are to find the person whose name is listed on their card and survey them. If the green number on the card they are holding is between 46 and 90, they are to be *subjects* in Phase Two; they are to stay in their seats and wait for a classmate to find them and survey them. When they are finished collecting data from their respondents, Phase Two *collectors* return to their seats to signal that they have completed their task.



Once the students have completed Phases One and Two of the activity and everyone is seated, we begin Phase Three. In Phase Three, I display the slide displayed below and included in the accompanying slide deck.



I assure students that their respondents will not see their responses (I collect and discard the notecards at the end of class). After giving the students a moment to record their responses on the backs of their cards, we discuss the activity. I begin by explaining the relationship between variables and attributes (as shown on the slide below). We discuss the levels of measurement and how some variables (like age) could be transformed from one level of measurement to another (from an interval variable to an ordinal variable).



I then discuss the concepts of independent and dependent variables and provide a simple example. As this activity is intended to be a quick introduction to these concepts, I assure students that we will study these concepts in much more depth in the weeks to come. I lead a discussion about how we can be sure that a certain variable is indeed causal, or if it is just correlational. We also discuss the usefulness of certain variables and I ask students to name those variables that they felt were of little to no use. Students often indicate that shoe size and favorite color seem irrelevant to the inferences they are attempting to draw. I point out the research between height (which is highly correlated with shoe size) and income, and the observation that a disproportionate number of people say their favorite color is blue. This leads to a discussion of what data we need to answer certain questions, and that sometimes we have data that is not appropriate to our research questions.



I ask students if they used information they didn’t formally collect as responses to the variables to draw their inferences, even if subconsciously, or if there were any inferences they logically drew that made them feel uncomfortable. How might their respondent’s general friendliness or seeming competence influence their judgement about their capacity to earn an A in the course? How did they feel about using someone’s race to infer their future fertility plans or their group membership to infer their use of alcohol? This leads into another important first day conversation, about sociology and how it deals with controversial and uncomfortable topics. I find that linking this necessary conversation about respect and intellectual conversations to a concrete example (this activity) is more meaningful than discussing them in the abstract

**Usage Notes**

This activity is designed to take about twenty minutes of class time on the first day of the term. I have used it in classes ranging in size from thirty-five to ninety students, but it could be adapted for other sizes. The attached file includes slides and a sample spreadsheet assuming a class of ninety students using a random name generator.