

## Populations and Sampling

Populations and sampling are critical issues when doing survey research. A population refers to all of the units in a set. If we're interested in the population of the United States, for example, every single person (each unit) in the United States is part of the set. If we are interested in how pop songs from 1968 differ from pop songs of the 2018, then all pop songs from both years are the population (of course, we would need to define what counts as a pop song).

The larger a population becomes, the more likely social scientists are to use a sample—or a subset—of that population to survey. Probability, or random, samples contain essentially the same variations that exist in the population. What does that mean? If the population you're interested in is comprised of 50 percent women and 50 percent men, we would expect a random sample of that population to closely match those percentages. Unless the sample is chosen randomly, it will not be *representative* of that population. Generally, we want to be able to generalize from a sample to the whole population; therefore, we use random samples. Knowing the population and how that population was (or was not) sampled helps us understand to whom we can generalize the findings from the survey.

Now that you have had a chance to think about how researchers collect data using populations and sampling, let's explore the ARDA to see different ways these are used.

Go the [www.theARDA.com](http://www.theARDA.com) and click on the tab for the Data Archive. Look at the structure of how the data are broken into categories to get a sense of what types of populations might be used. The archive is broken into several categories of data, including International Surveys and Data, U.S. Church Membership Data, U.S. Surveys, and Other Data. Click on the link for the Baylor Religion Survey and scroll to Wave III from 2010. Click on the "Summary" tab. There is a variety of information available from the summary, including the number of cases and variables, Data Collection and Collection Procedures, (summary pages often include a section for Sampling Procedures, as well).

As you read through the summary what does it tell you about the population targeted for the survey and how that population was sampled? According to the summary the target population was "a national sample" of adults age 18 and older in the continental United States. These adults were sampled using random digit telephone sampling (random digit dialing).



Find five additional data files in the ARDA's Data Archive. List the name of the data set, the target population, sampling procedure, and how the data were collected.

1. Data File: \_\_\_\_\_

2. Population: \_\_\_\_\_

3. Was random sampling used for this data set? (circle one)

Yes    No

4. How was the data collected?

\_\_\_\_\_  
\_\_\_\_\_

5. Who can be generalized to from this data set?

\_\_\_\_\_  
\_\_\_\_\_

6. Data File: \_\_\_\_\_

7. Population: \_\_\_\_\_

8. Was random sampling used for this data set? (circle one)

Yes    No

9. How was the data collected?

\_\_\_\_\_  
\_\_\_\_\_

10. Who can be generalized to from this data set?

\_\_\_\_\_  
\_\_\_\_\_

11. Data File: \_\_\_\_\_

12. Population: \_\_\_\_\_

13. Was random sampling used for this data set? (circle one)

Yes    No

14. How was the data collected?

---

---

15. Who can be generalized to from this data set?

---

---

16. Data File: \_\_\_\_\_

17. Population: \_\_\_\_\_

18. Was random sampling used for this data set? (circle one)

Yes    No

19. How was the data collected?

---

---

20. Who can be generalized to from this data set?

---

---

21. Data File: \_\_\_\_\_

22. Population: \_\_\_\_\_

23. Was random sampling used for this data set? (circle one)

Yes    No

24. How was the data collected?

\_\_\_\_\_  
\_\_\_\_\_

25. Who can be generalized to from this data set?

\_\_\_\_\_  
\_\_\_\_\_