

## 4.2: The Statistical Process

Statistical Inference can be thought of as a process that can be used for testing claims and making estimates.

### Steps of a Statistical Process

- Step 1 (Problem):** Ask a question that can be answered with sample data.
- Step 2 (Plan):** Determine what information is needed.
- Step 3 (Data):** Collect sample data that is representative of the population.
- Step 4 (Analysis):** Summarize, interpret and analyze the sample data.
- Step 5 (Conclusion):** State the results and conclusion of the study.

In Step 3, we introduce the concept of a representative sample. Let's define it here.

### Definition: Representative sample

A **representative sample** has characteristics, behaviors and attitudes similar to the population from which the sample is selected.

### Definition: Biased sample

A sample that is not representative is a biased sample.

Representative samples are necessary to make valid claims about the population. We will explore methods of obtaining representative samples in a later section.

### Example: Online dating trends



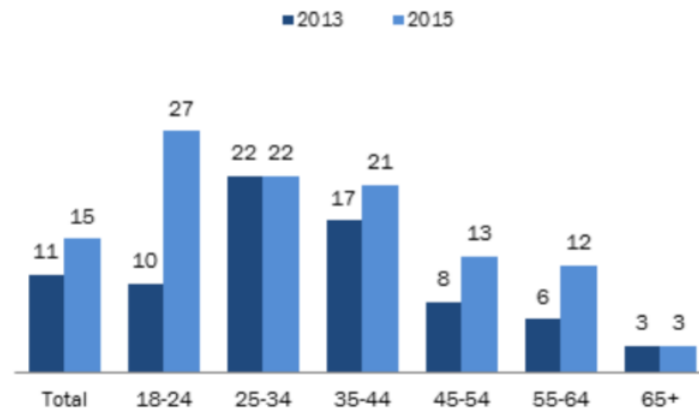
In 2015, the Pew Research Center was investigating trends in online dating; this culminated in a study published in February, 2016.<sup>41</sup> Pew Research wanted to investigate a belief that American's use of online dating website and mobile applications had increased from an earlier study done in 2013, especially among younger adults.

A survey was conducted among a national sample of 2,001 adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia. Fully 701 respondents were interviewed on a landline telephone, and 1,300 were interviewed on a cell phone, including 749 who had no landline telephone. Calls were made using random digit dialing. In addition to questions about online dating, researchers collected demographic data as well (age, gender, ethnicity, etc).

The survey found that in 2015, 15% of American adults have used online dating sites and mobile apps, compared to 11% in 2013. However, for young adults aged 18-24, the increase was dramatic: from 10% in 2013 to 27% in 2015. All age groups are summarized in the graph.

## Use of online dating sites or mobile apps by young adults has nearly tripled since 2013

% in each age group who have ever used an online dating site and/or mobile dating app



Source: Survey conducted June 10-July 12, 2015.

PEW RESEARCH CENTER

Let's first identify the population and the sample in this study.

The **population** is all American adults living in all 50 states and the District of Columbia. The **sample** is the 2,001 adults surveyed.

In this example we can investigate how Pew Research Center followed the Steps of a Statistical Process in performing this analysis.

1: Ask a question that can be answered with sample data.	Has there been an increase in American's use of online dating in the last two years? Are these rates affected by age?
2: Determine what information is needed.	The percentage of adults who are using online dating service. The age of each individual.
3: Collect sample data that is representative of the population.	Since the researchers surveyed both land lines and cell phones using a random dialer, the sample should be representative of the population.
4: Summarize, interpret and analyze the sample data.	15% of American Adults have used online dating sites and mobile apps, compared to 11% in 2013. For young adults aged 18-24, the increase was dramatic: from 10% in 2013 to 27% in 2015. Other age groups are displayed in the graph.
5: State the results and conclusion of the study.	Adults are using online dating sites and mobile dating apps at increasing rates, especially younger adults.

This page titled [4.2: The Statistical Process](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Maurice A. Geraghty](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.