

**Learning Target:**

- Explain why random sampling is advantageous when conducting a survey.

**LEARNING STRATEGIES:** Close Reading, Questioning the Text, Role Play, Summarizing, Paraphrasing, Debriefing, Discussion Groups

Jorge is a member of the student government at a large school with over 2500 students. The student government would like to recommend that students with part-time jobs be permitted to get a class credit in business. Knowing that Jorge is a good statistics student, the student government asked him to estimate the proportion of students at the school who have part-time jobs.

1. What difficulties might Jorge encounter if he tries to ask every student about having a part-time job?

Sometimes you may want to know some characteristic of a large population, such as the median income of households in your state or the proportion of students at a large school who have part-time jobs. Since it is often difficult or impossible to survey everyone in the population, you may wish to **survey** a **sample** of the population and infer conclusions from the sample about the population.

Jorge considers different methods for obtaining a sample.

2. Jorge is thinking about posting the question, “Do you have a part-time job?” on Facebook and collecting **responses** to his post. He knows that not everyone will reply, but he thinks he’ll still get a large number of responses. Explain why, even if a large number of people replied (even as much as half of the student body), Jorge would be unwise to suppose that the proportion of people who posted that they have a part-time job is the same as the proportion of all students who have a part-time job.

**My Notes****MATH TERMS**

A **survey** is a study in which subjects are asked a question or series of questions.

An answer provided by a subject to a survey question is called a **response**.

**MATH TERMS**

A **sample** is part of a population of interest. Data are collected from the individuals in the sample.

## My Notes

3. Jorge is on the football team at his school and is thinking of asking everyone on the football team if they have a part-time job. Why might this give him a poor estimate of the actual proportion of students at his school with part-time jobs?
  
4. Jorge is considering standing beside an exit of the school one day after the last class is over and asking every student who passes by if he or she has a part-time job. How might this method produce an inaccurate estimate of the actual proportion of students at his school with part-time jobs?

## MATH TERMS

A sample shows **bias** if the composition of the sample favors certain outcomes.

## MATH TERMS

A **simple random sample (SRS)** is a sample in which all members of a population have the same probability of being chosen for the sample.

Sampling can give very good results even if only a small sample of the population is surveyed, but it is critical that the sample be *representative* of the population with respect to the survey question. If the design of a sample favors one outcome over another, the sample is said to be **biased**. Each of Jorge's sampling methods described in Items 2, 3, and 4 display bias, and your responses indicate how this bias was manifested in the results.

How can you be sure that a sample is representative of the population? Many methods of sampling people could produce samples of people that would tend to favor one type of survey response over another.

One way to avoid favoring some types of response over others is to sample people at random, with every person being equally likely to be chosen. Such a sample is called a **simple random sample**, abbreviated **SRS**. A simple random sample is impartial because it does not favor anyone over anyone else. When a simple random sampling process is used to select members from a population, then everyone is as likely to be included in the sample as everyone else, and one person's inclusion in the sample has no effect on anyone else's inclusion in the sample.

5. There was bias in each of the sampling methods described in Items 2, 3, and 4 of this activity. Describe how a simple random sample would have avoided such bias.

6. Jorge has access to a full roster of all 2500 students at his school. One way to get a simple random sample of students would be for him to write the names of all 2500 students on index cards, put the cards into a large cardboard box and mix them up thoroughly, and then to draw out the desired number of names at random. What difficulties might Jorge encounter in his attempt to take a simple random sample in this way?

Another way to get a simple random sample is to number the list of students from 1 to 2500, and then use technology to randomly generate integers between 1 and 2500 until you have the desired sample size. For example, on TI-84 calculators, the following command generates a random integer between 1 and 2500:

`randInt(1,2500)`

Use the command to generate random integers that are matched up with the numbered list (ignoring repeated numbers) until you have identified all those names chosen to be in your sample.

7. Use your graphing calculator to choose 20 random integers between 1 and 100. Write the calculator syntax and your 20 random integers.

My Notes

TECHNOLOGY TIP

To find the *randInt*( function on a TI-84 calculator, press the **MATH** button, scroll to *PRB*, and then choose *randInt*(.

**My Notes**

Another method for generating random numbers from 1 to 2500 involves using a random digits table. Since the largest number in this range has four digits, you need to represent all numbers from 1 to 2500 as four-digit numbers. For example, 23 would be represented as 0023, and 798 would be represented as 0798. Then choose a line of the table at random and begin inspecting clusters of four digits. When a four-digit number matches one on Jorge's list, that name is selected as part of the sample. If a number is not on the list, then it is disregarded, as are repeated occurrences of the same number.

Random digits								
Line								
101	19223	95034	05756	28713	96409	12531	42544	82853
102	73676	47150	99400	01927	27754	42648	82425	36290
103	45467	71709	77558	00095	32863	29485	82226	90056
104	52711	38889	93074	60227	40011	85848	48767	52573
105	95592	94007	69971	91481	60779	53791	17297	59335
106	68417	35013	15529	72765	85089	57067	50211	47487
107	82739	57890	20807	47511	81676	55300	94383	14893
108	60940	72024	17868	24943	61790	90656	87964	18883
109	36009	19365	15412	39638	85453	46816	83485	41979
110	38448	48789	18338	24697	39364	42006	76688	08708
111	81486	69487	60513	09297	00412	71238	27649	39950
112	59636	88804	04634	71197	19352	73089	84898	45785
113	62568	70206	40325	03699	71080	22553	11486	11776
114	45149	32992	75730	66280	03819	56202	02938	70915
115	61041	77684	94322	24709	73698	14526	31893	32592
116	14459	26056	31424	80371	65103	62253	50490	61181
117	38167	98532	62183	70632	23417	26185	41448	75532
118	73190	32533	04470	29669	84407	90785	65956	86382
119	95857	07118	87664	92099	58806	66979	98624	84826
120	35476	55972	39421	65850	04266	35435	43742	11937
121	71487	09984	29077	14863	61683	47052	62224	51025
122	13873	81598	95052	90908	73592	75186	87136	95761
123	54580	81507	27102	56027	55892	33063	41842	81868
124	71035	09001	43367	49497	72719	96758	27611	91596
125	96746	12149	37823	71868	18442	35119	62103	39244
126	96927	19931	36089	74192	77567	88741	48409	41903
127	43909	99477	25330	64359	40085	16925	85117	36071
128	15689	14227	06565	14374	13352	49367	81982	87209
129	36759	58984	68288	22913	18638	54303	00795	08727
130	69051	64817	87174	09517	84534	06489	87201	97245

- Beginning at line 122 on the random digit table, identify the first five numbers that would correspond to names on Jorge's list. Compare this method to using the random integer generator on the graphing calculator.
- Suppose that Jorge uses the random number generator on his graphing calculator to choose an SRS of 100 students at his school. He then surveys these students to determine whether they have part-time jobs. He notices that two of the 100 students in his sample are friends who both have part-time jobs working at the local auto garage. Jorge is worried about the over-inclusion of people with part-time jobs in his sample. Should he be concerned?

