

## How Much Do Fans Love Justin Timberlake? Day 2

In the next city, Justin Timberlake's concert promoter again wants to find out how much fans enjoy his concerts. He will ask fans, "From 1 to 100, where 100 is the most, how much did you enjoy the concert?" Again, he wants to take a sample of 10 fans. He also would like to try out a couple of new methods for sampling.

### 1. Method #1:

Take a simple random sample (SRS) of 10 fans.

Simple Random Sample

$RNG(1, 50) \Rightarrow 10$   
different numbers for 10 different seats.



		X		X					
					X	X			X
X						X			
									X
			X					X	

### 2. Method #2:

Yesterday we decided the row would have a big impact their enjoyment. The promoter decides to sample **entire columns** (sample every fan in the selected columns).

Cluster Sample

a. Why would sampling all the fans in a column give a good estimate?

you would get a person in each row which affects enjoyment.



X			X						
X			X						
X			X						
X			X						
X			X						

b. How many columns will the promoter need to select to get a sample of 10 fans? Randomly choose the columns and mark the fans that will be sampled.

2,  $RNG(1, 10) \Rightarrow 1, 3$

### 3. Method #3:

Justin's manager thinks it is important to sample fans that have different views of the stage. He wants to sample every 7<sup>th</sup> fan.

Systematic Random Sample

a. First, we need to figure out the starting fan. Randomly select a fan and mark with an X.

$RNG(1, 50) \Rightarrow 43$



			X						
X							X		
				X					
X	X	X							X
						X			

b. Begin marking every 7<sup>th</sup> seat until you get a sample of 10 seats (start back at the beginning if you need to).

### 4. Which method do you think is best? Why?

Systematic, you can't get a group of all the same row or column.

5. Now, it's time for the actual data. For each of your samples on the previous page, calculate the average enjoyment. Add your average to the dotplots on the board.

Sample #1: 
$$\frac{+ + + + +}{10} = \bar{x}$$



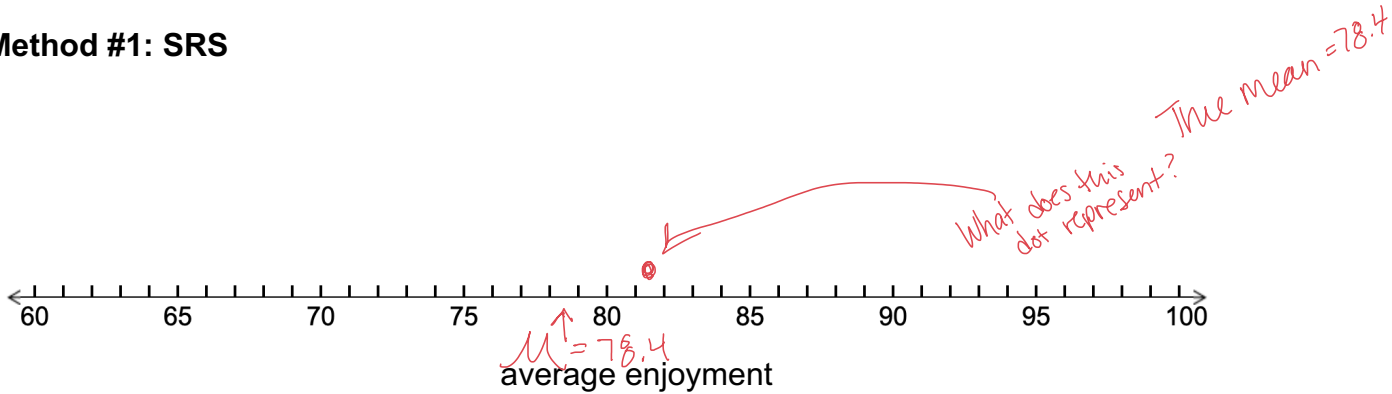
Sample #2: 
$$\frac{+ + + + +}{10} = \bar{x}$$

Sample #3: 
$$\frac{+ + + + +}{10} = \bar{x}$$

92	89	90	88	95	10	98	93	95	84
82	86	90	88	86	91	90	89	85	83
80	74	80	67	81	82	76	77	74	65
72	68	74	73	70	69	72	70	68	67
69	67	68	68	64	66	63	63	70	68

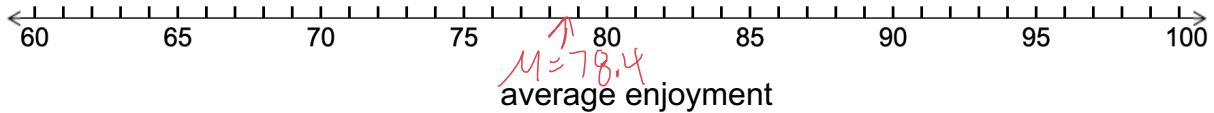
**Method #1: SRS**

*Unbiased*



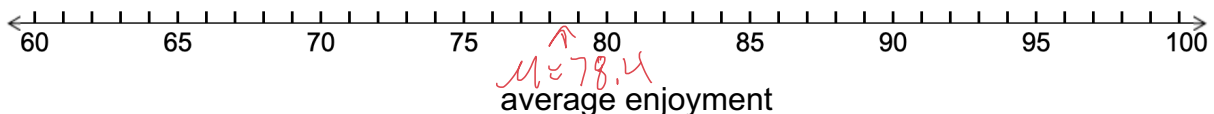
**Method #2: Cluster Sample**

*Unbiased  
lower variability*



**Method #3: Systematic Random Sample**

*Unbiased*



## More Random Sampling Methods Day 2

<p><b>Important Ideas:</b> Cluster sample: Split population into groups based on location. Use an SRS to choose groups. Sample <u>all</u> individuals in those groups. Cluster <math>\rightarrow</math> heterogeneous groups Strata <math>\rightarrow</math> homogeneous groups</p>	<p><b>Systematic Random Sample:</b> Sample where you randomly choose a starting position then sample with an equal interval between individuals.</p>	<p><b>Benefits of types of samples:</b> SRS: Unbiased Stratified: Unbiased &amp; low variability Cluster: Easy to take sample Systematic: Don't need to label everyone.</p>
---	--	---

### Check Your Understanding:

A large hospital would like to survey their patients on their level of satisfaction with their hospital room. The hospital has 10 floors, each with 15 rooms (total of 150 rooms). The hospital staff would like to take a sample of 30 rooms.

1. Describe how to select a stratified random sample of 30 rooms.

Each floor is labeled with rooms 1 to 15. Start with first floor. RNG(1,10) for 3 unique numbers. Select the corresponding rooms. Repeat for all floors.

2. Describe how to select a cluster sample of 30 rooms.

Each floor is labeled 1 to 10. RNG(1,10) for 2 unique numbers. Select corresponding floors and sample all rooms on these floors.

3. Describe how to select a systematic random sample of 30 rooms.

Label every room from 1 to 150. RNG(1,150) for 1 number. That's the room to start. Select every 5<sup>th</sup> room from there.

4. Explain a benefit of using each of the three types of sampling methods in this context.

Stratified: All floors are represented in the sample - Reduced variability.

Cluster: It's the easiest and fastest.

Systematic: Don't have to label every room. All floors are represented.