Name : _____ AP Stat Activity Ch 5 --- Sampling and Bias

➢ Part I: Sampling Techniques

Problem: A farmer has just cleared a field for corn that can be divided into 100 smaller plots. The land has his house on one side of the field and a river running down the opposite side from his house.
The farmer isn't sure whether harvesting the entire field is worth the expense. So he decides to harvest 10 plots and use this information to estimate the total yield. Based on this information, he will decide whether to harvest the remaining plots.

Assume the farmer does not have access to any data yet concerning this field. What are some criteria that should impact his decision as to which sampling choice would be in the **best interest of his business?**

_____

_____

A year has passed and the crop of corn is up and ready to harvest. Below is a grid showing the yield per plot. **REMEMBER, you would never have such information. You would only know this for the plots you planted. But for this exercise, this information is provided.**

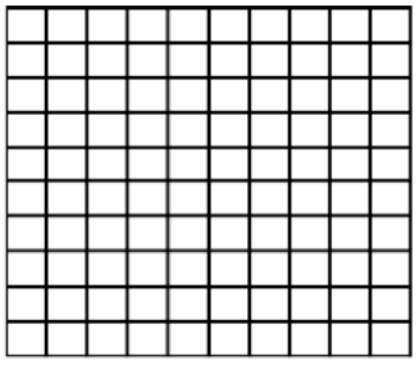
6	17	20	38	47	55	69	76	82	97
7	14	23	34	43	56	63	75	81	92
2	14	28	30	50	50	62	80	85	96
9	15	27	34	43	51	65	72	88	91
4	15	28	32	44	50	64	76	82	97
5	16	27	31	48	59	69	72	86	99
5	18	28	34	50	60	62	75	90	90
8	15	20	38	40	54	62	77	88	93
7	17	29	39	44	53	61	77	80	90
7	19	22	33	49	53	67	76	86	97

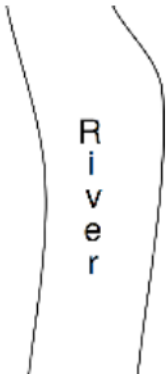
❖ Method 1: Convenience Sample

The farmer begins by choosing 10 plots that are **most convenient** for him to harvest without any other considerations. What do you think would determine this selection? _____

_____

Mark these plots on the grid below. Use the yield from these plots to calculate total yield, mean yield, and estimate of the entire field's yield. Complete the table with summary statistics.





Sampling Method	Convenience Sample	SRS	Stratified by Column	Stratified by Row
Mean				
Standard Deviation				
Minimum				
Q1				
Median				
Q3				
Maximum				
Outliers				

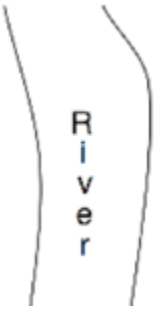
Total Yield: sum of the yield of your chosen plots = _____
Mean Yield per Plot: divide your total yield by 10 (Why?) = _____
Estimate of Field's Yield: multiply your mean yield by 100 (Why?) = _____

❖ Method 2: Simple Random Sample (SRS)

Begin at **line 120** on a random digit table and choose 10 plots at random to harvest. Mark them on the grid below with an X. Use the yield from these plots to calculate total yield, mean yield, and estimate of field's yield. Go back and complete the table of summary statistics. **[Pay attention to <u># of digits</u> you must read to represent each plot.]**

Is repetition of digits allowed? Why or why not? _____

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100



Total Yield: _____ Mean Yield per Plot: _____ Estimate of Field's Yield: _____

❖ Method 3: Stratified by Column

Think of the field as grouped in 10 vertical columns (the strata).

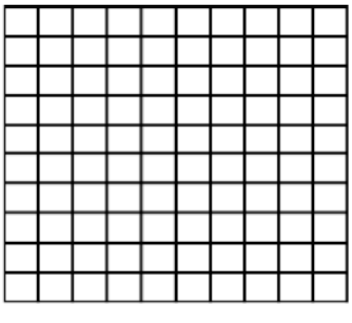
What is the common characteristic in the strata that groups the field this way? _____

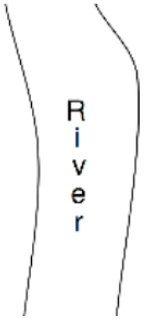
_____

How should you number your grid ***differently*** than we did for Method 2?

Is repetition of digits allowed? Why or why not? _____

_____





Begin at **line 105** on a random digit table and randomly choose one plot from each vertical strata and mark these plots on the grid with an X. Use the yield from these plots to calculate total yield, mean yield, and estimate of field's yield. **Go back and complete the table of summary statistics.**

Total Yield: _____ Mean Yield per Plot: _____ Estimate of Field's Yield: _____

❖ **Method 4: Stratified by Row**

Think of the field as grouped in 10 horizontal rows (the strata).

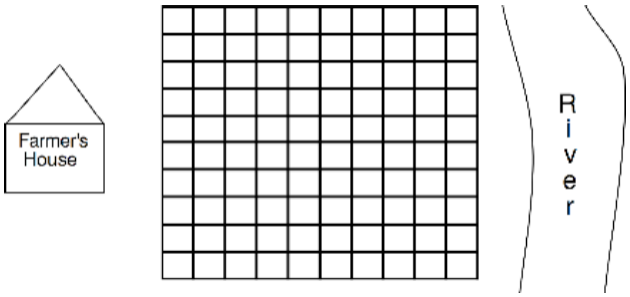
What is the common characteristic in the strata that groups the field this way? _____

_____

How should you number your grid ***differently*** than we did for Method 2 & Method 3?

Is repetition of digits allowed? Why or why not? _____

_____



Begin at **line 160** on a random digit table and randomly choose one plot from each horizontal strata and mark these plots on the grid with an X. Use the yield from these plots to calculate total yield, mean yield, and estimate of field's yield. **Go back and complete the table of summary statistics**.

Total Yield: _____ **Mean Yield per Plot: _____** **Estimate of Field's Yield: _____**

- **Part II: Analysis of Data**

The goal for choosing a good sampling method is one that will **best predict** the mean of the population distribution with as **small variation** as possible. Using the summary statistics, construct boxplots for each method.

SRS

STRATCOL

STRATROW

Based on **our** results, which probability sampling method would you recommend to the farmer and why? Relate your answer to the actual yield of the farmer's field. Explain using statistics.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____