Name: ____Kameron Mace_____

# **Random Sampling Activity:** Sampling Segregation

*Most American cities have economic segregation. This means that households in the same neighborhood have similar incomes, but incomes across neighborhoods tend to be quite different. Multiple reports have found that San Antonio is one of the most economically segregated cities in the United States. The following activity investigates the challenges of sampling incomes in San Antonio.*

**Judgement Sample:** Select 5 homes that you think are *representative* of the whole population of San Antonio. Label the home numbers on the top lines and their incomes on the bottom lines. Calculate and write down the mean of these incomes in the space provided.

Home Number: 1 34 59 65 30

Income: 25k 56.5 139 20 58.5 Mean: 59.8

Pros: Easy to implement

Con: Biased

**Simple Random Sample (SRS):** Use a random number generator to select 5 unique homes numbered between 1-100. Label the home numbers, their incomes, and mean of their incomes.

Home Number: _23_ _4_ _77__ _1_ _2__

Income: 30.5 56 76 25 26.5 Mean: 42.8

Pros: Unbiased, moderately accurate

Cons: would have to drive all over

**Cluster Random Sample:** Use a random number generator to find 1 home numbered between 1 and 100. Select that home as well as the four homes located closest to it on the map.

Home Number: _37__ __36_ 38___ _39__ _30__

Income: 44.5 42.5 41 43 58.5 Mean: 45.9

Pros: unbiased

Cons: high variability



Lesson created by Skew The Script (skewthescript.org) and made possible through support from The University of Texas San Antonio School of Data Science

**Stratified Random Sample:** Find the page with the title "Strata by Race-Ethnicity", in which the homes are divided by predominant race. Use a random number generator to select 1 home within each of the 5 strata. For example, use RandInt(67,71,1) to select a Native American home, since the range for that strata is 67-71.

Home Number: _58__ _62__ _42__ _71__ _76__

Income: ___68__ __26__ __92.5_ __25___ __56____ Mean: _53.5____

Pros: Unbiased, accurate, low variability Cons: Hardest to implement , need to have lots of data

**Systematic Random Sample:** Select a random number between 1 and 20. This is the first home. Add 20 to the first home number to select the second home. Add 20 more to get the third home. Repeat until you have 5 homes.

Home Number: 13___ __33_ _53__ __73_ __93_

Income: __35.5__ _5.5____ ___47___ __110____ ___186___ Mean: __86.7_____

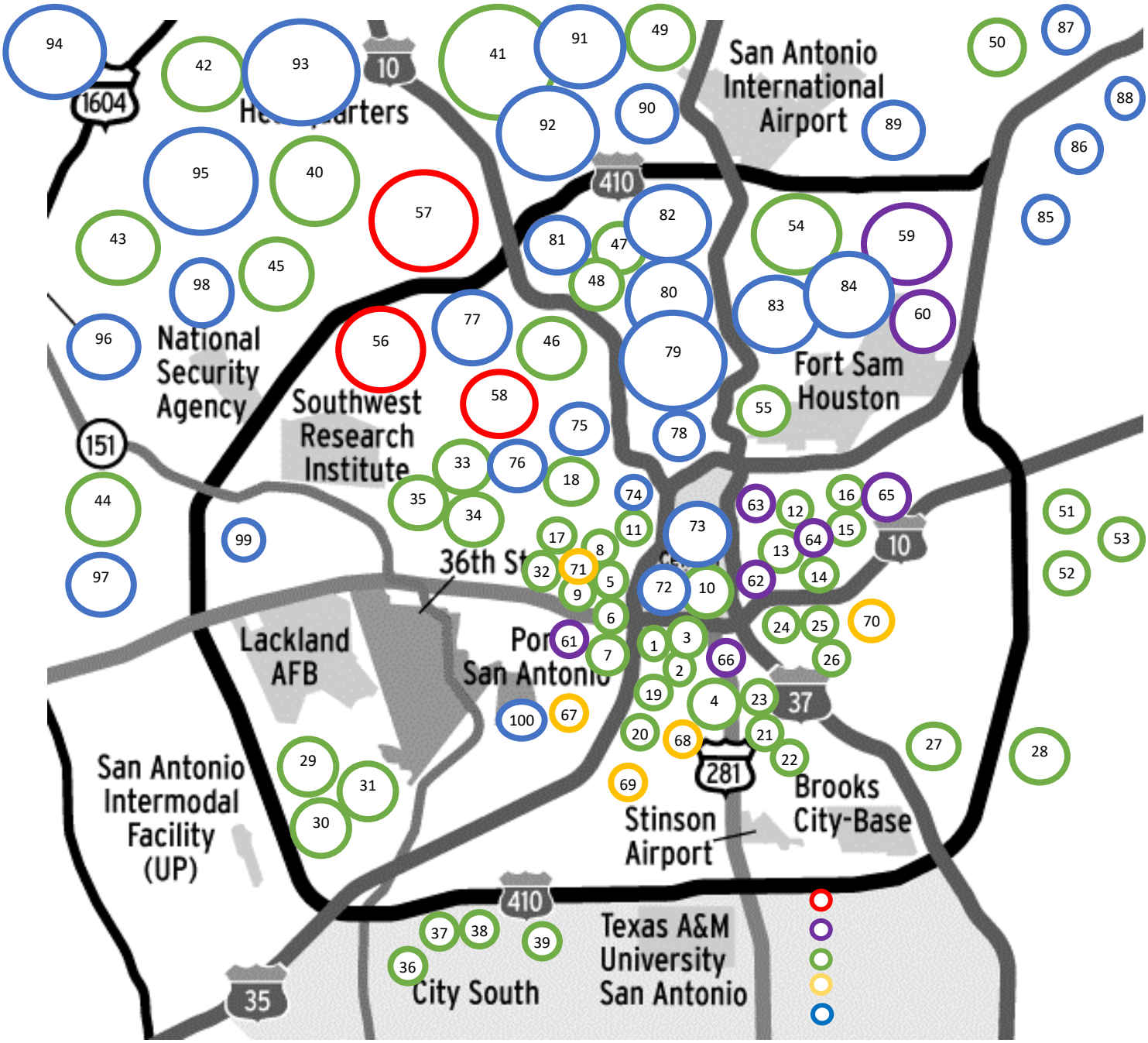
Pros: Unbiased Cons: moderate variability

TRUE MEAN: 61,000



Lesson created by Skew The Script (skewthescript.org) and made possible through support from The University of Texas San Antonio School of Data Science

UTSA School of Data Science




**Key (Predominant Race in Home):**

- Asian
- Black
- Hispanic
- Native American
- White



Lesson created by Skew The Script (skewthescript.org) and made possible through support from The University of Texas San Antonio School of Data Science

UTSA School of Data Science

| Home | Income |
|---|---|
| 1 | $25,000 |
| 2 | $26,500 |
| 3 | $42,000 |
| 4 | $56,000 |
| 5 | $23,000 |
| 6 | $24,500 |
| 7 | $44,500 |
| 8 | $25,500 |
| 9 | $25,500 |
| 10 | $54,000 |
| 11 | $26,000 |
| 12 | $25,000 |
| 13 | $35,500 |
| 14 | $25,000 |
| 15 | $27,500 |

| Home | Income |
|---|---|
| 16 | $23,500 |
| 17 | $23,500 |
| 18 | $53,000 |
| 19 | $36,000 |
| 20 | $28,500 |
| 21 | $27,500 |
| 22 | $28,000 |
| 23 | $30,500 |
| 24 | $27,000 |
| 25 | $31,500 |
| 26 | $29,500 |
| 27 | $45,500 |
| 28 | $61,000 |
| 29 | $61,000 |
| 30 | $58,500 |

| Home | Income |
|---|---|
| 31 | $64,000 |
| 32 | $35,000 |
| 33 | $55,000 |
| 34 | $56,500 |
| 35 | $57,000 |
| 36 | $42,500 |
| 37 | $44,500 |
| 38 | $41,000 |
| 39 | $43,000 |
| 40 | $109,000 |
| 41 | $214,500 |
| 42 | $92,500 |
| 43 | $95,000 |
| 44 | $68,500 |
| 45 | $76,000 |

| Home | Income |
|---|---|
| 46 | $65,000 |
| 47 | $62,500 |
| 48 | $78,000 |
| 49 | $85,500 |
| 50 | $61,500 |
| 51 | $48,000 |
| 52 | $44,500 |
| 53 | $47,000 |
| 54 | $108,500 |
| 55 | $44,500 |
| 56 | $102,000 |
| 57 | $139,000 |
| 58 | $68,000 |
| 59 | $107,500 |
| 60 | $85,000 |

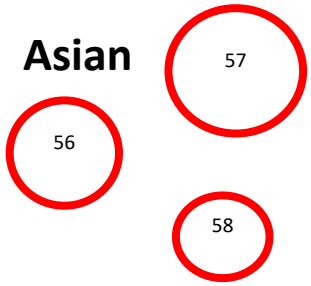
| Home | Income |
|---|---|
| 61 | $66,500 |
| 62 | $26,000 |
| 63 | $24,500 |
| 64 | $22,000 |
| 65 | $20,000 |
| 66 | $28,500 |
| 67 | $23,000 |
| 68 | $20,000 |
| 69 | $27,500 |
| 70 | $36,000 |
| 71 | $25,000 |
| 72 | $57,000 |
| 73 | $110,000 |
| 74 | $28,000 |
| 75 | $58,000 |

| Home | Income |
|---|---|
| 76 | $56,000 |
| 77 | $76,000 |
| 78 | $44,500 |
| 79 | $212,500 |
| 80 | $119,000 |
| 81 | $67,000 |
| 82 | $124,000 |
| 83 | $109,000 |
| 84 | $111,500 |
| 85 | $48,000 |
| 86 | $43,500 |
| 87 | $47,000 |
| 88 | $39,000 |
| 89 | $58,500 |
| 90 | $62,000 |

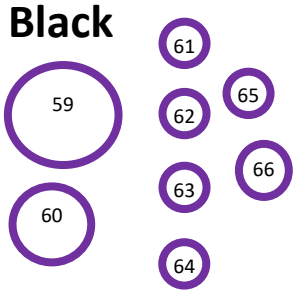
| Home | Income |
|---|---|
| 91 | $108,000 |
| 92 | $157,000 |
| 93 | $186,000 |
| 94 | $166,000 |
| 95 | $174,000 |
| 96 | $80,000 |
| 97 | $42,000 |
| 98 | $48,500 |
| 99 | $28,000 |
| 100 | $27,000 |

## Strata: By Race-Ethnicity

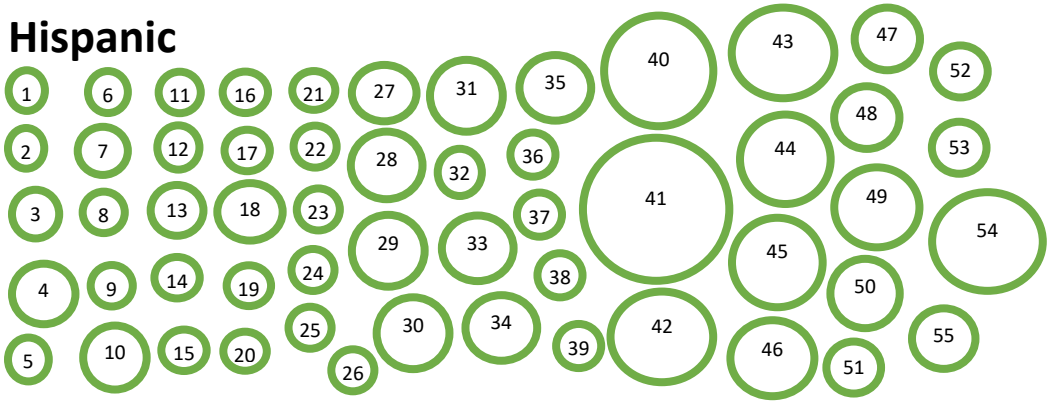
### Asian




### Black




### Hispanic

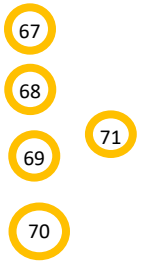




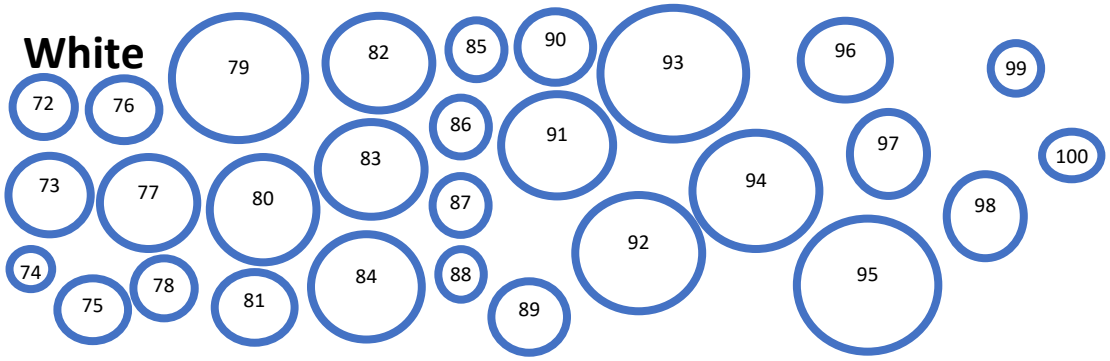

Lesson created by Skew The Script (skewthescript.org) and made possible through support from The University of Texas San Antonio School of Data Science

UTSA School of Data Science

## Native American

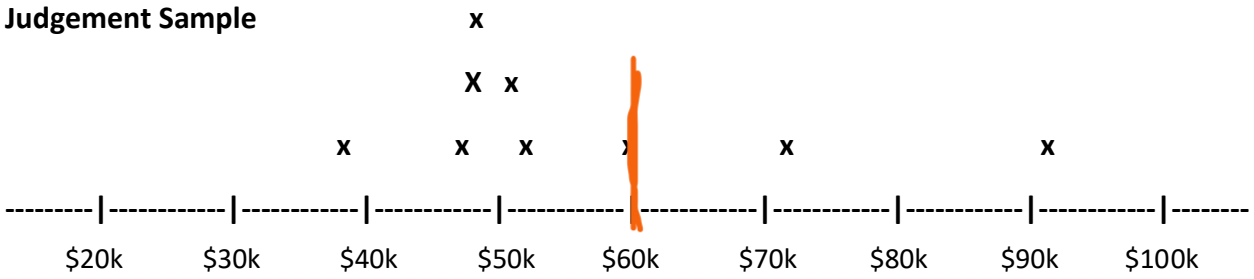



## White

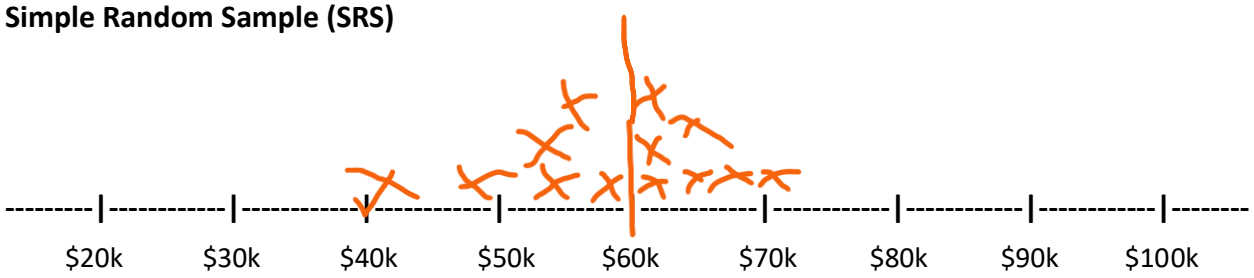



### Sampling Distributions for Each Sample Type

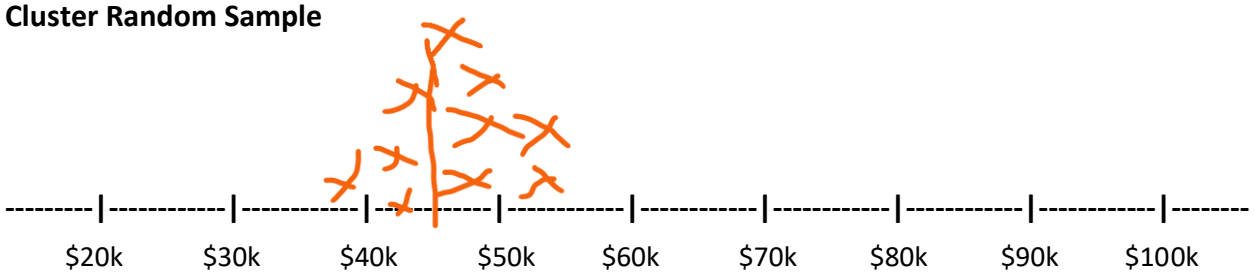
**Judgement Sample**




**Simple Random Sample (SRS)**




**Cluster Random Sample**

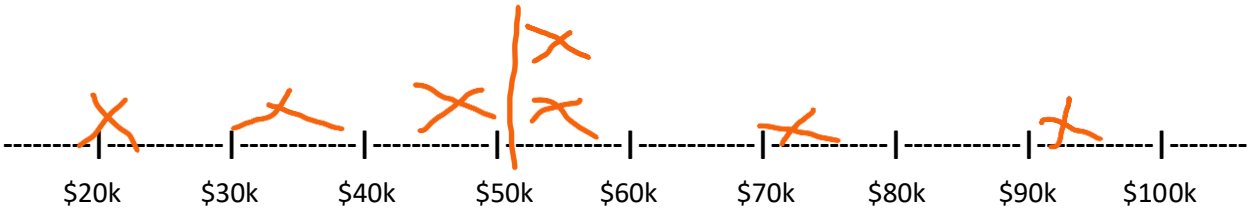





Lesson created by Skew The Script (skewthescript.org) and made possible through support from The University of Texas San Antonio School of Data Science

School of Data Science

**Stratified Random Sample**




**Systematic Random Sample**