1. Precipitations = [53.71,34.35,43.63,46.33,44.04, 38.94, 55.12, 39.05, 60.18, 39.37, 53.05, 52.87,36.11,37.21, 38.67, 46.36, 37.16, 33.44, 25.59, 53.21]
   1. mean = ∑(precipitations) / length(precipitations) =

= 868.39 / 20 = **43.42**

* 1. median = 39.37, 53.05, or mean([39.37, 53.05]) = **41.5**
  2. μ = population mean = 43.42

N = population size = 20

σ = standard deviation = sqrt(∑(Xi – μ)2 / N-1)

= sqrt(∑[105.89, 82.25, 0.04, 8.47, 0.38, 20.06, 136.90, 19.09, 280.91, 16.39, 92.74, 89.31, 53.42, 38.55, 22.55, 8.64, 39.18, 99.59, 317.89, 95.85]/ 20)

= sqrt(1528.18 / 19) = **8.97**

* 1. variance = σ2 = 8.972 = **80.43**
  2. range = max(precipitation) – min(precipitation)

= 60.18 – 25.59 = **34.59**

1. The best dispersion measure to compare income inequalities between counties over the years is the **Coefficient of Variance (CV),** since it allows direct comparison across different datasets and removes the influence of magnitude. CV is calculated as the sample standard deviation over the sample mean.
2. The Binomial Distribution could be used to estimate the probability of the lower Brahmaputra River flooding at least once in the next 3 out of 10 rainy seasons.

p = 0.4

N = 10

P(3) = (10! 0.43 (1-0.4)10-3) / (3! (10 – 3)!) = (10! 0.43 0.67) / (3! 7!) = 0.2149

1. To calculate the likelihood of 5 or more days of rainfall this May we use the Poisson Distribution. We could either calculate the sum of the probabilities that there are 5 – 31 days of rainfall, or we could calculate the difference between 1.0 and the sum of probabilities that it rains 0 – 4 days.

μ = 3.2

P(0) = (e-μμ0) / 0! = 0.04

P(1) = (e-μμ1) / 1! = 0.13

P(2) = (e-μμ2) / 2! = 0.20

P(3) = (e-μμ3) / 3! = 0.22

P(4) = (e-μμ4) / 4! = 0.17

P(0) + P(1) + P(3) + P(4) = 0.78

1. – 0.7808 = **0.22**

To verify the claim that P(5-31) = 1.0 – P(0-4), I wrote a python script to calculate the sum of probabilities from 5 to 31 and got 0.22.

1. = 26.55”
   1. Probability of more than 31” of annual snowfall:

Zi = (31 - / s = (31 – 26.55) / 5.37 **=** 0.8286

Using Normal Table 0.83 -> 0.2967 -> 0.5 – 0.2967 = **0.2033**

* 1. The annual snowfall likely to be exceeded with a probability of 0.9 is:

(X – 26.55) / 5.36 = z score corresponding to (0.9 – 0.5 = 0.4) = -1.28

X = **19.67”**

* 1. The probability of snowfall between 20”-30”:

z1 = (20 – 26.55) / 5.37 = -1.22 -> using the Normal Table -1.22 -> 0.3888

z2 = (30 – 26.55) / 5.37 = 0.64 -> using the Normal Table 0.64 -> 0.2389

z1 + z2 = 0.3888 + 0.2389 = **0.6277**

1. If a researcher is studying soil chemistry along Bayfront Beach and wants to oversample tidal wetland area, which makes up about 15% of the total study area, then she is advocating for a (disproportionate) stratified point sampling approach. She thinks the human impact on the soil is most severe in this area and wants to control the effect of this influence by separating this area into its own “strata” and disproportionally oversampling from it. If she wanted to collect 100 samples in total, then she would collect **more than 15 from tidal wetland area, and less than 85 from the rest of the study area**. One possible example would be to collect 30 samples from the tidal wetland area and 70 samples from the rest beach area being studied.
2. A physical geographer is interested in observing the pH values at the Holiday Farm after the wildfire in 2020.

n = 50, = 7.8, s = 1.2

* 1. 90% Confidence Interval:

± z (s/sqrt(n))

z = (0.9 / 2 = 0.45) -> 1.65 (from table)

7.8 ± 1.65 (1.2 / sqrt(50))

[7.8 - 0.28, 7.8 + 0.28] = **[7.52, 8.08]**

* 1. 95% Confidence Interval:

± z (s/sqrt(n))

z = (0.95 / 2 = 0.475) -> 1.96 (from table)

7.8 ± 1.96 (1.2 / sqrt(50))

[7.8 - 0.33, 7.8 + 0.33] = **[7.47, 8.13]**

1. A preliminary sampling of 50 households has been conducted, and the calculated standard deviation is 65 liters.

S = 65, E = 10

* 1. Samples needed to be 95% confident that estimates are within 10L of the true population mean:

n = ((1.96 \* 65) / 10)2 = 162.30 = **163 samples**

* 1. Samples needed to be 90% confident that estimates are within 10L of the true population mean:

n = ((1.65 \* 65) / 10)2 = 115.02 = **116 samples**