

Instructions

Due Week 10 · Friday, October 30

Prof. Jack Reilly

Fall 2026

Complete the following problems and upload your answers as a PDF file to **Blackboard**. Upload your .do file as well.

Questions

1. In 2022, approximately 17% of all land on the globe had been preserved and protected.¹ Not all countries protect an equal number of lands, however. The dataset `protectedlands2022.dta` gives the distribution of all protected lands across all countries in the world in the year 2022.
 - a. A researcher samples 10 countries. What is the probability that these sampled countries protect more than 30% of their lands? (*Hint: calculate the population mean and standard deviation from Stata, and use these to manually calculate a z-score.*)
 - b. If the researcher sampled 50 countries, rather than 10, what would the probability be that these sampled countries protect more than 20% of their lands?
2. The `sample` command in Stata draws random samples from data in memory, with the number after `sample` referring to the size of the sample to draw. (So, for instance, the command `sample 10` would draw a 10% sample of the dataset.)
 - a. Draw a histogram of the distribution of protected lands.
 - b. Draw a 20% sample of countries from the dataset. What is the mean and standard deviation of this sample? Draw a histogram of the distribution. How close does this distribution look to the population distribution?
 - c. Draw a 20% sample of countries 30 more times, and record the mean each time. Draw a histogram of these sample means.² What is the shape of this distribution? What is its standard deviation?

¹Source: <https://ourworldindata.org/grapher/terrestrial-protected-areas>

²Note: this will require you to manually record what those sample means are, and then create a new dataset within Stata.