Title: Funga Ecological Data Scientist Role-related Assignment

Instructions: We are hoping the new ecological data scientist at Funga will help our team expand our current computational capacity specifically in the analysis of forest data and development of models for forest data. The following assignment is meant to be open ended and assess the ability of the candidate to communicate their analyses quickly to their team members. This assignment will also give us the chance to view your coding and communication style. Communication is important as we scale the company and work asynchronously. Please format your response to the following questions in any way you see fit. Any coding or markup language is allowed. AI is also allowed but please specify if it is used. This is meant to take no longer than two hours. Please explain the steps you took to arrive at your answers, show your work allowing your analysis to be reproducible, and include next steps you might take in the future with your analysis and/or with the development of your analysis pipeline.

Datasets: There are two attached datasets. *tree_level_lob_dat.csv* contains FIA tree-level measurements (diameter at breast height in inches and tree height in feet) at 200 sites as well as the year they were measured and forest stand age. *site_level_lob_covariates.csv* contains covariate information for each site where "mat", "t_seas", "t_range", "t_wet_q", "t_warm_q", "map", "p_wet_q" "p_warm_q" are climate variables from WorldClim (https://www.worldclim.org/) and "sand", "silt", "clay", "bdod", "cec", "cfvo", "nitrogen", "phh2o", and "bdticm" are soil variables from Soil Grid 2.0 (https://soilgrids.org/)

Deadline: Return to ann@funga.earth by 5pm ET Friday, March 28th, 2023

Instructions:

1. Use the attached datasets to model biomass.
2. Predict biomass at stand age 25 across the sites and hypothesize about what is driving the variability in the predictions. Then, pretend a particular stakeholder is interested specifically in the effect of soil pH on future productivity and provide an assessment of this relationship.
3. Assume climate change will decrease soil pH across the dataset. Provide a recommendation for sites that will be productive planting locations under this scenario.