

Instructions:

1. ALL DUE TIMES ARE IN EST
2. Upload a Python file to this BB Assignment.
3. All answers must be in your own words, and copy-and-paste answers will receive no credit.
4. You must submit a **.py file** that contains the results of your code.
5. You are limited to 2 submissions
6. **You must be in the lab session to get the lab's credit.**
7. If you cannot submit the lab by the end of the day, you can still submit it by the end of the next day with a 25% penalty. Any lab submissions after two days of the lab class will be given a zero

Download the attached python file and TSV file (eukaryotes.tsv) from the Blackboard assignment. After loading the data into the `euk` DataFrame, investigate the structure and contents of the DataFrame. Then use Pandas to solve the following:

- A. Write the code to create a new DataFrame, `euk_float`, which only includes rows where the column "Size (Mb)" is less than 4,000 and drops any rows with missing values.
- B. Write a line of code to filter the `euk` DataFrame to include only observations where the ratio of "Number of proteins" to "Number of genes" is at least 1.1. Store the result in a new DataFrame called `euk_filtered`.
- C. How many fungal species have genomes bigger than 100Mb? What are their names?