

Test Question (45 Minutes)

Write a Python script to analyze taxonomic data and produce summary statistics and visualizations. The script must be containerized using Docker.

Problem:

You are provided with a [dataset\(taxonomic_data.csv\)](#) representing species counts grouped by taxonomic classifications. Your task is to:

1. Save results in a `csv` file after calculating:
 - Total species count for each taxonomic phylum.
 - Average species count per species within each phylum.
2. Generate a bar chart showing the total species count for each phylum.

Deliverables:

Github repo with instructions to run python script in readme

- The container should:
 - Read the input CSV file mounted from the host system.
 - Save the output files to the same location.

Note:

- If any part of the problem seems unclear, make reasonable assumptions, clearly document them as comments in your code.
- Your code should follow the best coding standards.
- Ensure your script gracefully handles missing or invalid data.