

For HW 5 - PART 3, the Data Structures I plan on using to store my data will be an Array and a Heap. Below I talk about and explain the justification and operations that I will be using for this final project:

## 1. Array (NumPy)

- **Justification:**

- NumPy arrays are highly efficient for numerical computations and manipulations.
- Ideal for tasks like normalization, calculations, and feature transformations.

- **Operations:**

- Compute correlations between features (ex. PSS\_score and Neuroticism).
- Normalize numerical values such as screen\_on\_time.
- Performs calculations on derived features like sleep efficiency.

## 2. Heap (Heapq)

- **Justification:**

- Heaps work well for retrieving the smallest or largest elements in a dataset, making them ideal for tasks like ranking participants by stress levels or screen time.
- Python's heapq module provides an easy-to-use implementation of heaps.

- **Operations:**

- Finding Extremes: Use heaps to efficiently find participants with the highest PSS\_score or longest screen\_on\_time.

- Dynamic Tracking: Maintain a priority queue to dynamically track the top  $n$  participants with the most sleep efficiency.
- Sorting: Perform heap-based sorting for specific subsets of the data (ex. stress scores for a particular group).