Assumptions:

- The number of bins and the number of balls are the same (n)

Input parameters:

- n: number of bins and balls
- d: number of bins selected each time to do load balancing (is given, d=2, d=4)
- nRuns: number of runs of the simulation (each with a different random seed)

Output metrics:

- Occupancy for algorithm (1), (2), (3)
- Confidence interval

Main data structures:

- 3 dictionaries for the empirical results (1 for each algorithm) then aggregated in a dataframe
- nRuns dataframe with the results for each value of and and for each algorithm

Main algorithms:

- Random dropping (1):
  1. Select a random bin
  2. Increment the number of balls in the selected bin
- Random Load Balancing d=2 (2):
  1. Select d random bins
  2. Select the least occupied one of the d random bins
  3. Increment the number of balls in the selected bin
- Random Load Balancing d=4 (3) -> same as above

- SIMULATOR:
  1. For nRuns random seed do:
     a. Compute max_occupancy for each of the 3 algorithms
     b. Aggregate the results
     c. Append to the list of dataframe the results
  2. Compute mean and std deviation of all dataframes (obtaining a value for every n)
  3. Compute the confidence interval for every n
  4. Plot

Extension:

- Computing and plotting the confidence interval

RESULTS:

Below the results obtained with 10 runs, showing the 95% confidence interval