STAT107 Data Science Discovery

Project: Mosaic Extra Office Hour

Man Fung (Heman) Leung

Spring, 2022

University of Illinois at Urbana-Champaign

Collaboration

ILLINOIS

- Please work in a group of 2–4 students
  - collaboration is important in data science!
  - meet new friends and discuss :)
  - let us know if you have any questions
- Attendance form
  - you can come up if you do not want to use this form
  - submit before you leave the lab

Comment: lab_simulation

- Check email for score decomposition
- Do not be afraid of expressing your opinion! Thoughtful answers are good answers even if they may not be correct in the first place
- 3.1: -0.5 if you use uniform for any hero other than Mantis
- 3.4: exact probability can be found but they are not the numbers in 3.2/3.3. You need to use transformation of random variables, which is tedious even if you know how. I treat this question as opinion-based so it is fine if your elaboration is not theoretically correct
- 4.3: -0.5 if you simply write Monte Carlo method because it is usually an alternative name but not an application of simulation. Originally I should take off points from answers without elaboration as well but I did not do so

Extra credit: lab_simulation

ILLINOIS

- 1 point for displaying correct animation at the end
- 1 point for using appropriate sequence
  $n = 400, 800, \ldots, 10000$
  - exactly the same sequence is not required but it should be arithmetic
- 1 point for using the correct condition to separate points inside/outside circle
- 1 point for any additional things that you try/follow the hints, e.g.,
  - changing the `markersize`
  - using the points stored in `df4` instead of resimulation
- A few of you get full score :D Congratulations if you work it out on your own

- Overview
  - similar to a k-nearest neighbors algorithm
  - use your own tiles (2 extra points last semester)
- Ideas for extra credit in Section 9
  - vectorization using `np.mean` (sections 4 and 6)
  - subsampling (section 7)
  - other distance measure, e.g., sum of absolute difference (section 7)
  - filter the best tile, e.g., brighter/darker (section 7)
  - increase k in k-nearest neighbors (section 7)
- Feel free to:
  - ask us questions
  - leave whenever you submit the attendance form