Jacob Bachtarie 1985185
Nikolas Velazquez 1819671
Tam Nguyen 1538439
Braddock Mitchell Parks 2151920

Helios : Progress Report

For task 1, we, as a group, have divided the different subtasks amongst ourselves; Most of us have been spending our time on understanding the dataset and how to develop the kernel density function methods. We decided to use Python for our project, some of our members have not used python before for data science; They are learning and getting used to using the different libraries and navigating Jupyter Notebooks. We are most likely going to use K-Nearest-Neighbor as one of our methods and are actively discussing the different possible algorithms we could implement to use for our second method. We are considering the benefits and costs of different methods/approaches, e.g. Kernel Density Estimation.

We have made progress into the different models and are in the process of creating visual representation to present alongside the different models. Utilizing the seaborn and matplotlib libraries to create the visual representations. We have discussed the method we want to approach for the hotspot discovery algorithm and how we plan to get d1 and d2. We want to approach whether an area is a hotspot by looking at the different percentiles of our intensity estimation and finding what we believe would classify as high or significantly higher than the average.

For our group, our main setbacks have been not recognizing how to divide tasks into subtasks efficiently and understanding how to approach tasks. Time is the biggest constraint in completing this project. Thus, the limited time available gives a significant challenge that needs to be carefully managed to ensure we have a good analysis by the end of the project.