**GitGood Project Outline**

**Group:** GitGood

**Members:** Sarah Tucker, Nikita Konstantinov, Liam Joiner, Pallavi Gupta, Kai Ellis

**Dataset:** <https://datadryad.org/resource/doi:10.5061/dryad.120f446?fbclid=IwAR1rRF4M7UMN-z9dkWZH9MOh7JplgY8T6UTcUyGABU7_qgEd8kpycD_vKXw>

**Dataset description:**

The dataset contains Operational Taxonomic Unit (OTU) sequences from arthropods used in a study by [Creedy et al. (2019)](https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.4839) in a FASTA file format, as well as a .csv file reporting the number of reads of each OTU in each sequenced library. The occurrence of these OTU’s were counted in samples gathered by placing 1 m^2 metal trays in the canopy and collecting organisms that fell into them after being fogged. In total, 13 tray samples were collected from an individual *Liquidambar styracaflua* tree canopy.

**Data Files:**

Map.csv - a table showing the number of reads of each OTU in a well

Otu.fa - fasta file of all the OTUs from the study

Summary\_metadata.csv - summary table of the sample data for each well  
MBTrial\_tray-stem\_distances.csv - sample location in reference to base of tree

Community\_metabarcoding\_papers.csv - additional papers of interest that may be used to write the final report  
<https://github.com/sarahttucker/BIO432GitGoodProject>

<https://datadryad.org/resource/doi:10.5061/dryad.120f446/1>

**Questions to address:**

1. Is there a relationship between species composition and size class?
   1. SNP data by size class of arthropods - create phylogeny
   2. regression depicting relationship between SNPs and size class
2. What has a stronger effect on arthropod communities: the presence/absence of certain taxa, or sampling location?

**Analysis workflow/division of work:**

1. Data Clean up: (Everyone **😔**)

* General clean up - reformat data

↓

1. Is there a relationship between species composition and size class? (sarah, nikita, Liam)

* 2.1 Compare OTU data by creating a neighbour joining tree (label ends based on size - bray curtis, get rid of reads below 5)/scatterplot using size class

↓

3. What has a stronger effect on arthropod communities: the presence/absence of certain taxa, or sampling location? (Pallavi, Kai)

* 3.1 NMDS plot comparing effect of presence absence of taxa
* 3.2 NMDS plot comparing effect of sampling location

↓

4. Write up an .rmd report (Kai, Liam, Pallavi)

* Everyone - make figure captions for own work, etc)
* Nikita and Sarah edit and make sure .html file is working and has no errors