**Family data analysis 08/09/2022**

**Git repository link-** [**https://github.com/Renuka-3/Family\_microbiome**](https://github.com/Renuka-3/Family_microbiome)

**Completed tasks**

1. Phyloseq: Create phyloseq object using metadata, OTU table and taxa table
2. Beta diversity analysis: PCoA and PERMANOVA analysis was carried out with confounding factors such as age, gender, geographical location, diet and family.
3. Hierarchial Clustering: Hierarchial Clustering carried out using Ward2 method.
4. Alpha diversity analysis: alpha diversity estimated using Shannon diversity index for all co-factors.
5. [CST analysis](https://github.com/Renuka-3/Family_microbiome/blob/main/CST.md): Community state types analysis performed on geographical location.

**ToDo**

1. Random forest analysis for family predictions
2. PERMANOVA analysis by adjusting P values for multiple testing

**Needs to be discuss**

1. Data cleaning/rarefaction- Singleton removal (Currently we have 211 singleton sequences out of 1009)
2. Analysis pipeline- Phyloseq or TSE/mia?
3. If TSE then what about tree??
4. Tracing family bacterial signatures?