Questions

1. Why is this discriminant analysis of three different file formats of paleoclimatology data based off of 2 lakes in the Russian Federation important to perform?
2. Why is **Age** the chosen common column variable among the three different file formats?
3. Do you think **Depth** could have been a chosen common column variable?
4. Do you think **PANN** could have been a chosen common column variable?
5. Do you think **Picea** could have been a chosen common column variable?
6. Why were no columns removed from any dataset?
7. Is renaming the chosen common column variable important to structuring the data into a database?
8. Shouldn't the outliers be removed from the chosen common column variable?
9. Why were a txt, JSON, and xls file formats chosen to perform this discriminant analysis?
10. Do you think a database is efficient when studying the paleoclimatology in the Russian Federation?