1. 20 meters are roughly the side width of the South Hall that I live in, and for example, small to medium-sized buildings, cars, trees, small pond, and tennis/basketball court would fit in a 20 \* 20 meters square. A bigger building, parking lot, sectioned land, and roads would partially fit a 20 \* 20 m area. Green pixels would represent grassland, trees, fields on campus, while gray pixels represent buildings, and blue pixels are exclusively representing the tennis courts.

2. The algal booms in Oneida lake are mostly located in the south-eastern regions. The southern and eastern bank of the lake has more human constructions such as buildings, while on the eastern end there are also two streams connected to the lake.

3. There are 2297430 (1005 \* 2286) pixels in the raster image. The higher resolution provides a more details description of actual objects on the earth and makes identifying them easier (e.g. human vs. natural). Moreover, the HD image are more colorful, because the lower resolution image has to pact multiple pixels into one, hurting the color palettes.

4.



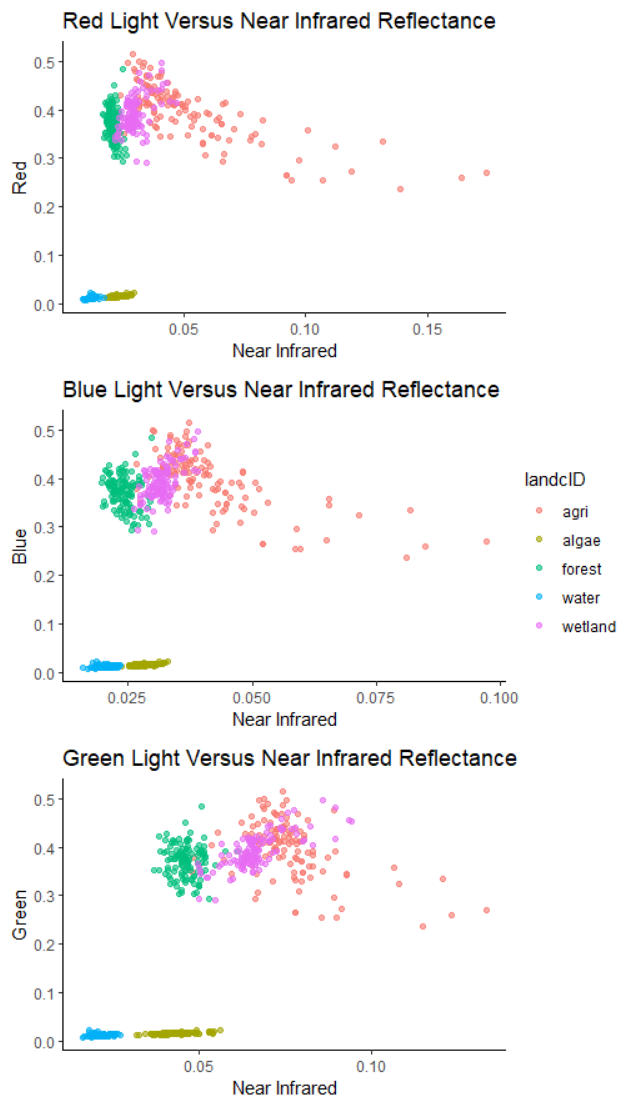
Above images are blue + red + green; infrared; and infrared + blue + green.

The last image actually provides a better contrast in terms of the lake situation. A lot of nearly negligible green observations from the previous graphs are better highlighted, and this graph with infrared in addition to blue and green image also prevented the confusion of the small isles in the lake being mistakenly considered as algae.

5. Overall the lands are green, while lighter colors indicating buildings, roads, and clouds. On the other hand, inside the lake, the waters are shown in a gradient between bright yekkiw to pink. Nearing the banks, the south eastern sides are remarkably more pink than other edges of waters, and for the most part the water has a yellow tone, while gradually deeper colorings on the south and east corners.

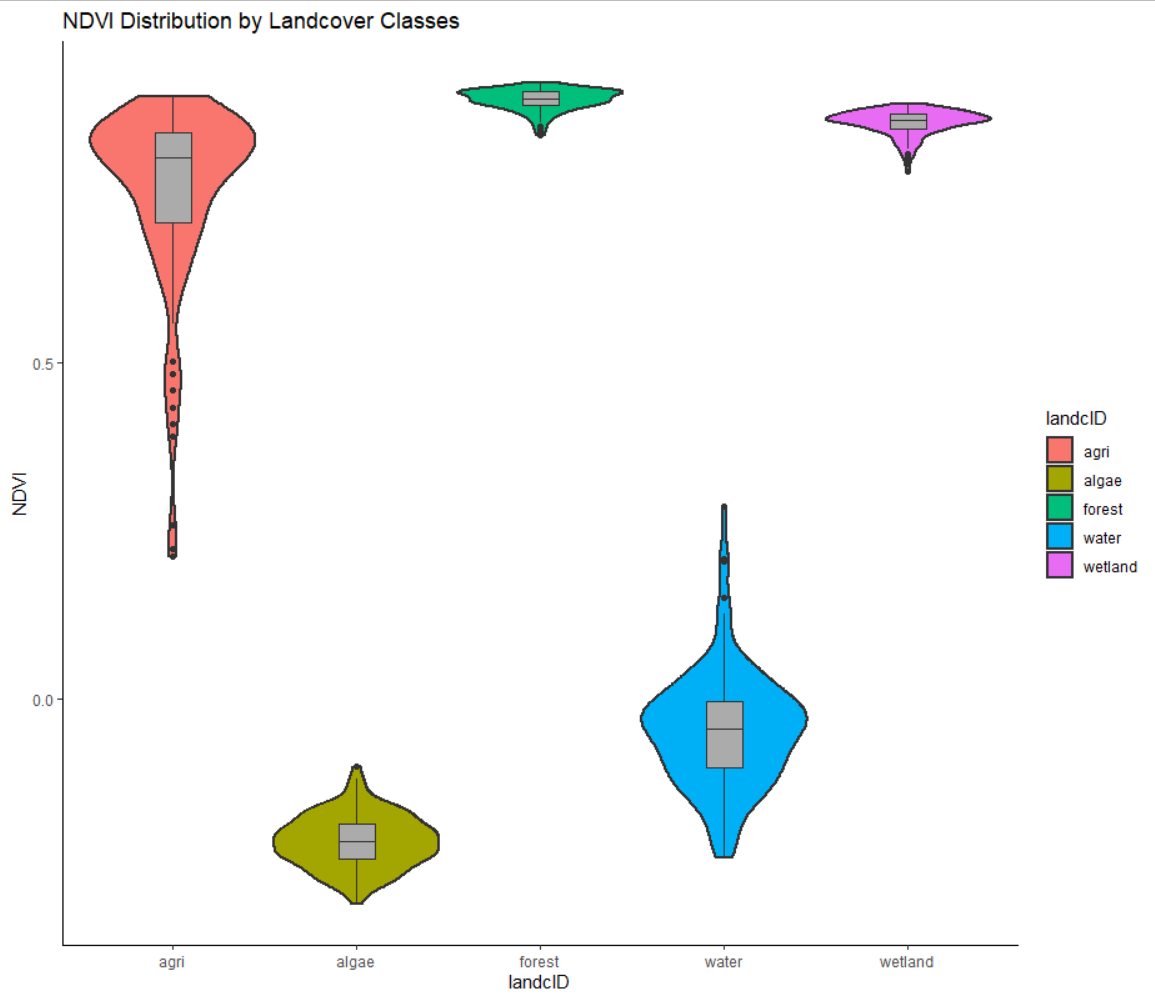
6. “Rep” replicates the values in the given vector, which in this case contains all the labels of each land type ID. “Each” means repeat each elements in the given vector such many times, while “times” would repeat the entire vector. It would be wrong to use “times” because x and y valuesare presented by groups, but this argument would make the ID labels --- algae, water, agri, forest, wetland, algae, water, agri, forest, wetland… What we need is “each” that would provide --- algae, … ,algae, water, … , water, agri, … , agri, forest,…, forest, wetland,…, wetland.

7.



From the above graphs, I would say that the landcover classes of algae and wetland are a bit hard to distinguish suing differences in reflectance alone.

8.



9. The NDVI for agircultural fields, forests, and wetlands does not ***seems*** to differ ***much*** as the graph shown above – most of their distributions are overlapping on the higher values; however, it could be due to the fact that the graph is not detailed enough to showcase the real difference up top. Therefore, an one-sided ANOVA test can be introduced:

> landc.aov <- aov(data = agri\_fore\_wetl, NDVI ~ landcID)

> summary(landc.aov)

Df Sum Sq Mean Sq F value Pr(>F)

landcID 2 1.230 0.6148 92.82 <2e-16 \*\*\*

Residuals 357 2.365 0.0066

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

>

> TukeyHSD(landc.aov)

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = NDVI ~ landcID, data = agri\_fore\_wetl)

$landcID

diff lwr upr p adj

forest-agri 0.13755053 0.11282137 0.162279687 0.0000000

wetland-agri 0.10314015 0.07841099 0.127869308 0.0000000

wetland-forest -0.03441038 -0.05913954 -0.009681222 0.0033195

And it is easy to tell that in general all three groups do not have the same variances, and pair-wise comparison also shows that each type of landclass significantly differs from others. Both procedure yielded extremely low p-values.

10. <https://github.com/guozhaosengzs/ENVDS/blob/master/activity8/activity8_script.R>