GeoWing Academy Case Study: RGB Drones for Ag Veg Health?

Can RGB drones & Machine Learning be used to analyse crop health?

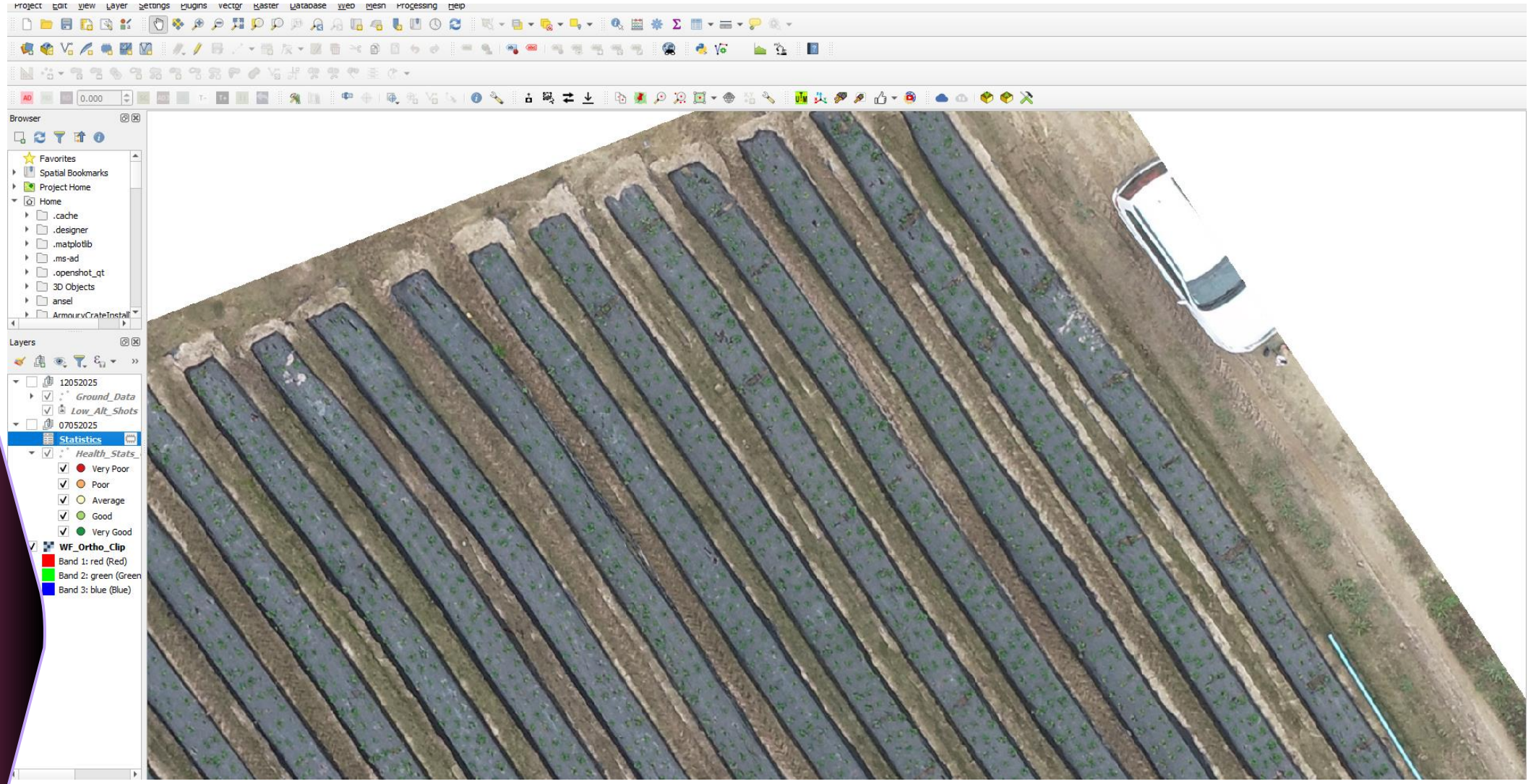
OpenDroneMap

GEOWING ACADEMY

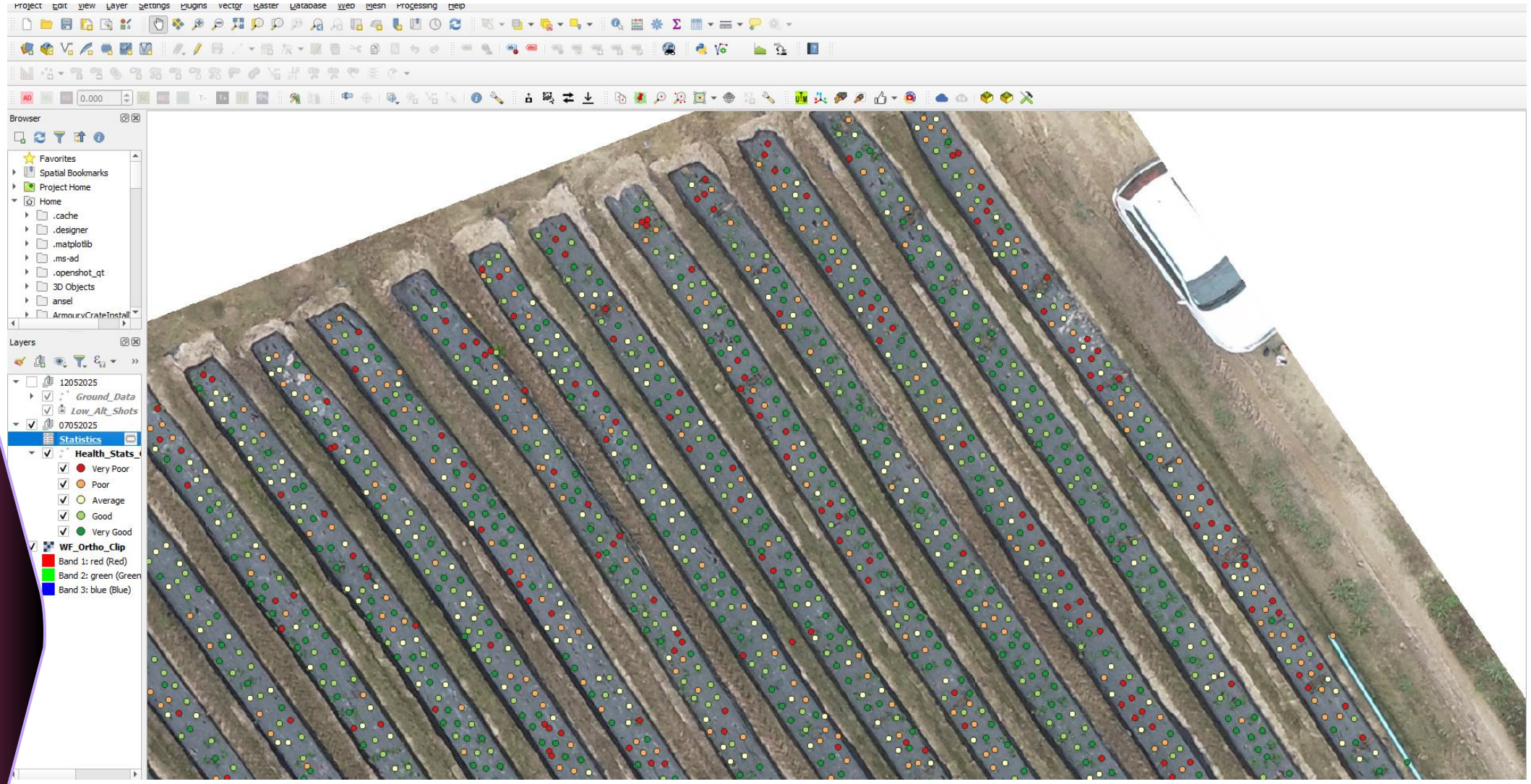
QGIS

GEOWING ACADEMY

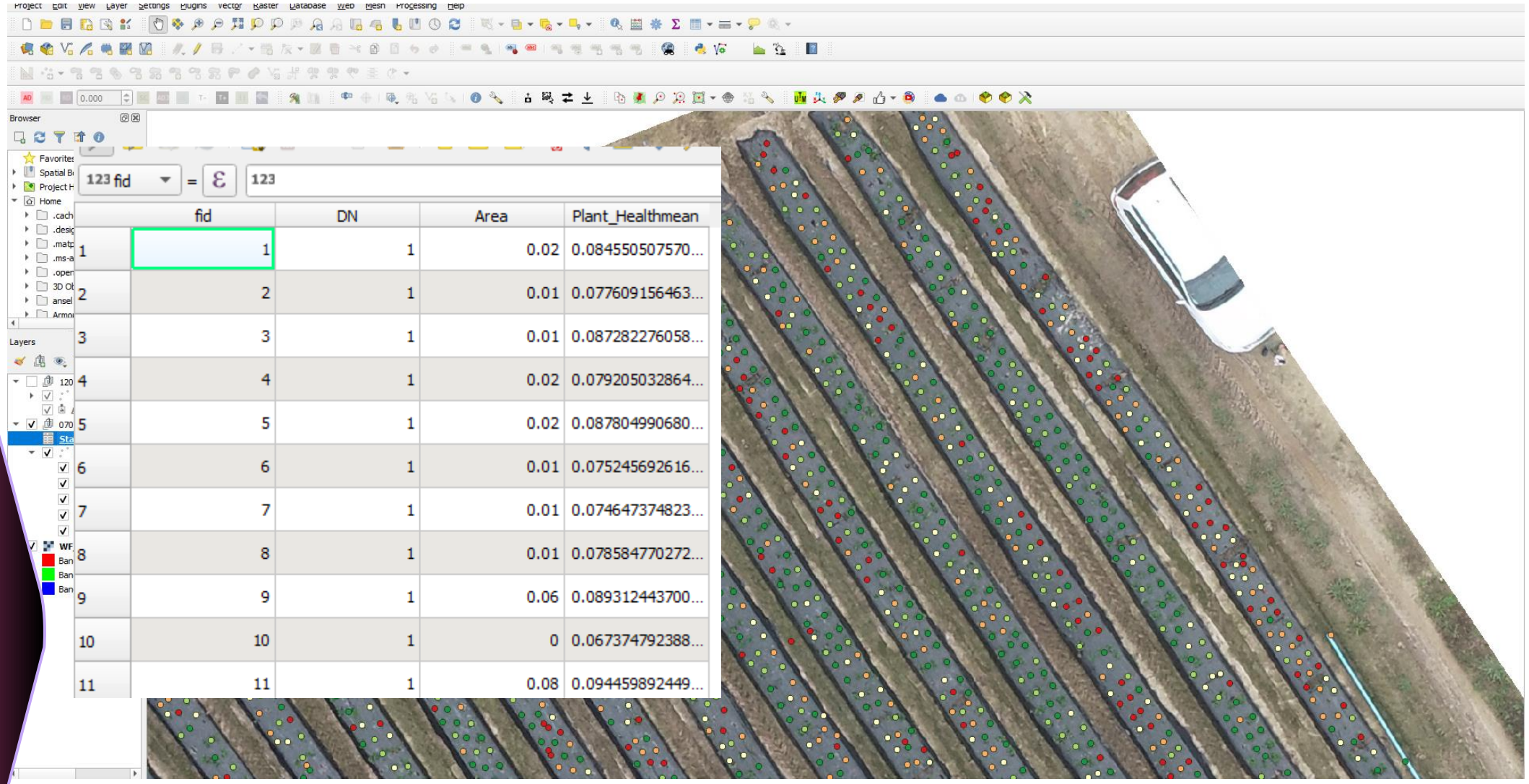
Integrated Machine Learning



Integrated Machine Learning

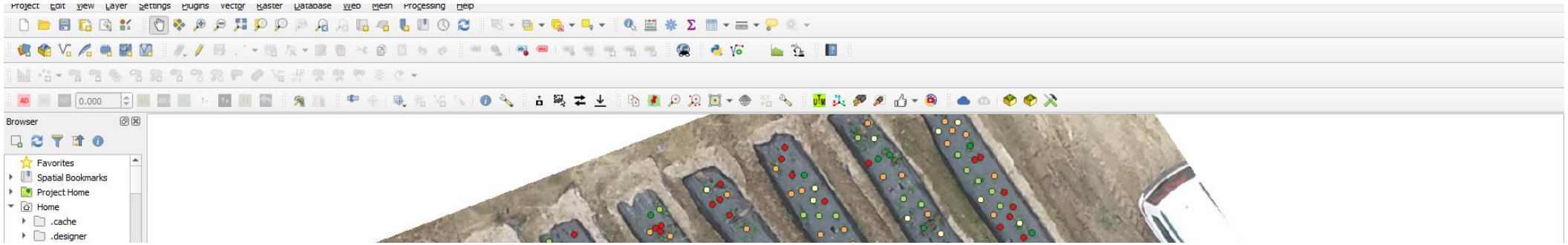


Individual Plant Health Stats

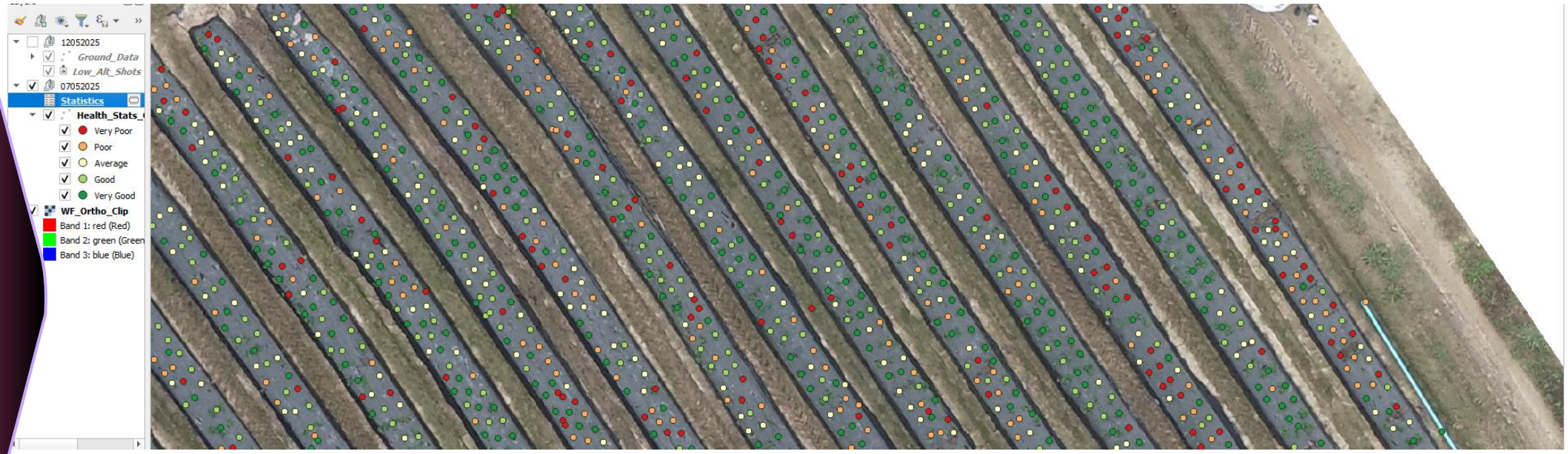


	fid	DN	Area	Plant_Healthmean
1	1	1	0.02	0.084550507570...
2	2	1	0.01	0.077609156463...
3	3	1	0.01	0.087282276058...
4	4	1	0.02	0.079205032864...
5	5	1	0.02	0.087804990680...
6	6	1	0.01	0.075245692616...
7	7	1	0.01	0.074647374823...
8	8	1	0.01	0.078584770272...
9	9	1	0.06	0.089312443700...
10	10	1	0	0.067374792388...
11	11	1	0.08	0.094459892449...

Plant Health Stats

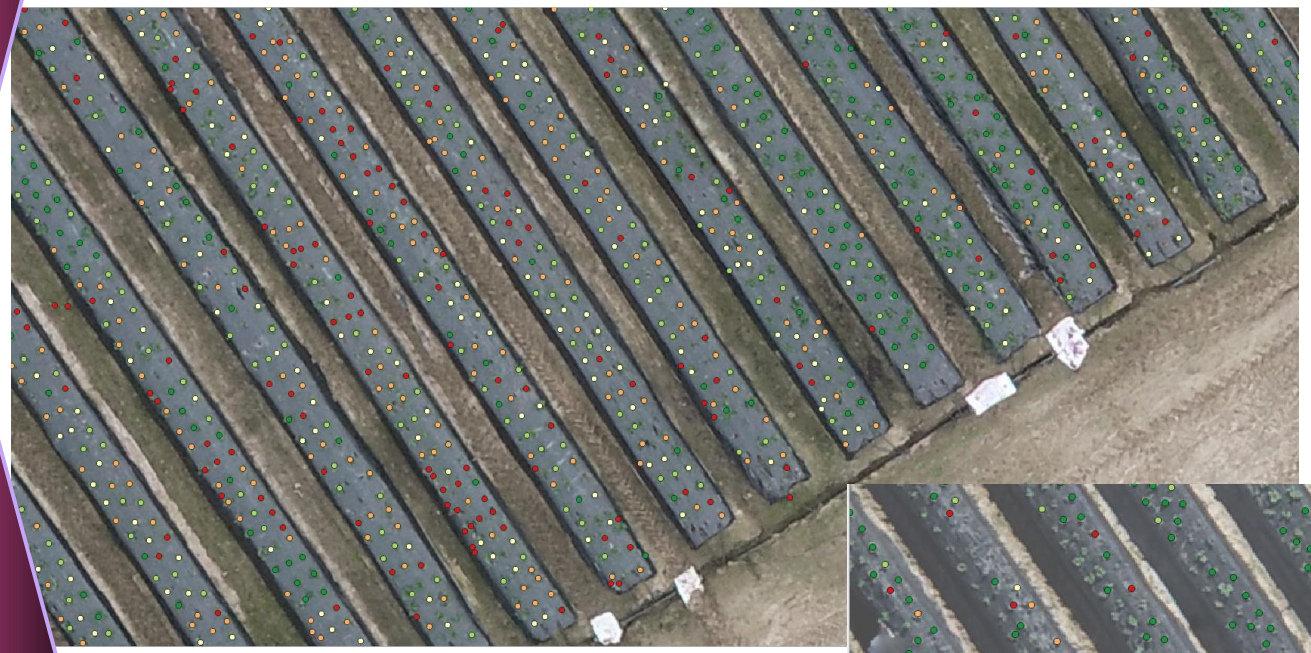


	count	unique	empty	filled	min	max	range	sum	mean	median	stddev	cv	minority	majority	q1	q3	iqr
1	32713	32713	0	32713	0.026823071254...	0.179422425275...	0.152599354020...	2863.502187927...	0.087534074769...	0.087515951693...	0.015891015427...	0.181540908149...	0.026823071254...	0.026823071254...	0.076627939582...	0.098562637721...	0.021934698138...



Cloud vs Sun Data Difference when using RGB

Cloudy (note consistent detections, uniform light)



Sunny (note lost detections, scattered light)

Shadow Confused with Grren when using RGB

Shadow Confused with Green when using RGB

*Health Index = MPRI

Cloud vs Sun Statistic Difference when using RGB

Statistic	Meaning	Sunny	Cloudy	Difference / Interpretation (incl. plant health & shadow/M.L. impact)
Count	Number of observations	30,689	33,229	Cloudy has more data points — slightly broader coverage. On cloudy days, detections are more consistent, improving overall statistical reliability.
Min	Minimum value in dataset	0.0228	0.026	Both minimums likely represent shadows or bare soil. On sunny days, deep shadows result in very low values, reducing ML detection accuracy in dark zones.
Max	Maximum value	1	0.2346	Sunny has an extreme high — possibly from sunlit plant tops or overexposure. Cloudy max is more balanced, improving index reliability.
Range	Max - Min	0.9772	0.2086	Sunny's large range shows high variation due to sunlight/shadow contrast, reducing ML consistency. Cloudy range is narrower, aiding stable model training.
Sum	Total of all values	5817.36	2892.5	Sunny has a higher total reflectance, but this is skewed by sunlight intensity and shadows. Cloudy sum reflects more balanced plant health data.
Mean	Average value	0.1896	0.087	Sunny mean appears healthier but includes false positives in bright areas and missed detections in shadow. Cloudy mean is more reliable for mapping.
Median	Middle value when sorted	0.169	0.0873	Sunny shows higher central values but is impacted by uneven lighting. Cloudy median offers more consistent input to ML classification.
Stddev (Standard Deviation)	Spread of values around mean	0.1129	0.0166	Sunny has high spread due to shadow/light extremes, which hurts ML performance. Cloudy data's low spread improves detection reliability.
CV (Coefficient of Variation)	Relative variability (stddev/mean)	0.5955	0.1909	Sunny data is less stable — inconsistent detections due to shadows. Cloudy CV is low, offering better model training and generalization.
Q1 (1st Quartile)	25% of values are below this	0.1104	0.0761	Shadows lower Q1 in Sunny data, misrepresenting health. Cloudy Q1 is more accurate for weak vegetation or soil detection.
Q3 (3rd Quartile)	75% of values are below this	0.2347	0.0984	Sunny Q3 shows higher reflectance from sunlit plants, but shadows reduce full-plant detection. Cloudy offers more consistent classification of healthy plants.
IQR (Interquartile Range)	Q3 - Q1: middle 50% spread	0.1243	0.0223	Sunny's wider IQR shows inconsistencies due to light variation. Cloudy IQR reflects more stable and reliable crop health analysis across the field.

Cloud vs Sun Statistic Difference Summary

- **Cloudy days** produce **lower values overall**, but lighting is even, resulting in **higher consistency, fewer shadows, and better machine learning performance** for plant detection and health analysis.
- **Sunny days** offer a wider data range and **highlight plant vigour**, but **shadows reduce the accuracy** of machine learning models in detecting individual plants — leading to **missed or misclassified plants**, especially in rows or under canopy.
- For **statistical reliability and ML classification**, **cloudy conditions are generally more dependable**, even if the raw values are lower.
- How to improve results
  - **Fly on over cast days** to ensure high detection accuracy for individual plant referencing
  - Give the training data **more inputs** to figure out the contrast between "green" and "shadow"
  - Use **Multispectral** Sensor

GeoWing Academy Uses Open Source Software, Provides all UAV Training Data & Material Including Additional Course Related Training Documentation

**You do not need a drone

OpenDroneMap

GEOWING ACADEMY

QGIS