

Description: This document describes the data included in a single csv file that represents all of the data gathered in Season 10 by the Gantry. The document includes column descriptions, general field and season information and clarification on things that may give rise to potential confusion.

Summary: This CSV file contains data from RGB, thermal, PS2, and 3D sensors that was collected in 2019/2020 over that year's lettuce season (Season 10). Ideally, each sensor would scan the field on the same day, however, that is not feasible based on the logistics of getting the best data from each sensor. The data itself reflects the sparse, staggered nature of Field Scanalyzer scans. At the start of the season, RGB was the only sensor being used to collect data and this remained until the plant's were large enough to provide meaningful data. Due to the nature of the sensors, the data is at the individual plant scale except for PS2. To cut down on scan time, PS2 captures were centered above a single plot, therefore, these data are at the plot level.

CSV Column Description:

Column Name	Description
date	Collection date.
plant_name	Unique identifier for an individual plant. This name consists of "genotype + _ + unique number assigned". The plant_name does not inform the reader of its location in the field.
plot	The agricultural plot, i.e where it is located in the field. This value consists of the range and column values for that particular plot of interest (Range + Plot).
year	Year of growing season.
experiment	Small description of the particular experiment that includes the crop, season and particular instance.
field	Location of the field. When discussing the Gantry field, there is a North and a South field. Different experiments are often grown in each field simultaneously, requiring a method to distinguish the two.
treatment	Value is associated with the irrigation treatment that was implemented. "treat1" = well-watered, "treat2" = moderately water-limited, and "treat3" = Severely water

	limited.
rep	Replication number of a plot/plant. This means repeating individual treatment plots within the field research area
range	The range in the field represents the row that the particular plot is in.
column	The column in the field represents the column that the plot is in.
genotype	The name of that particular genotype/accession/line.
bounding_area_m2	Calculated using the RGB sensor, this describes the size of the bounding box that encompasses that particular plant. The bounding area is a direct measurement of plant size in square meters.
double_lettuce	Indicates if there were two plants in the same location (this issue arose due to poor thinning). 0 = single plants 1 = more than 1 plant. You can choose to remove 1+ plants during data cleaning.
lon	Longitudinal coordinate of the individual plant
lat	Latitudinal coordinate of the individual plant
nw_lat	Latitudinal coordinate of the northwest point of the bounding box
se_lat	Latitudinal coordinate of the southwest point of the bounding box
se_lon	Longitudinal coordinate of the southwest point of the bounding box
roi_temp	Temperature of the region of interest
quartile_1	Represents the lowest 25% of the temperature values extracted from the individual plant's canopy
mean	The mean of all pixel values representing canopy temperature of the individual plant
median	The median of all pixel values representing canopy temperature of the individual plant
quartile_3	Represents the upper 25% of the temperature values extracted from the individual plant's canopy
variance	The spread of the pixel values representing canopy temperature of the individual plant

std_dev	The standard deviation of the pixel values representing canopy temperature of the individual plant
F0	Minimal level of fluorescence
FM	Maximal level of fluorescence
FV	Maximal variable fluorescence (FM - F0)
FV/FM	Maximum quantum efficiency of PSII photochemistry if all capable reaction centers were open
min_x	Minimum value in the x axis of all of the points in the plant 3D point cloud
min_y	Minimum value in the y axis of all of the points in the plant 3D point cloud
min_z	Minimum value in the z axis of all of the points in the plant 3D point cloud
max_x	Maximum value in the x axis of all of the points in the plant 3D point cloud
max_y	Maximum value in the y axis of all of the points in the plant 3D point cloud
max_z	Maximum value in the z axis of all of the points in the plant 3D point cloud
num_points_0	Total number of points in the 3D point cloud of the individual plant (x-axis)
num_points_1	Total number of points in the 3D point cloud of the individual plant (y axis)
num_points_2	Total number of points in the 3D point cloud of the individual plant (z axis)
hull_volume	Convex hull volume enclosing an individual plant
oriented_bounding_box	Bounding box aligned with the x,y,z axis
axis_aligned_bounding_box	Bounding box with slight rotation, not aligned with the coordinate system, but fits plant geometry more tightly
persistence_entropy_0 (TDA)	Measure of the entropy of the points in a persistence diagram
persistence_entropy_1 (TDA)	Measure of the entropy of the points in a persistence diagram

persistence_entropy_2 (TDA)	Measure of the entropy of the points in a persistence diagram
Number_points_0 (TDA)	Number of points in x axis
Number_points_1 (TDA)	Number of points in y axis
Number_points_2 (TDA)	Number of points in z axis
Amplitude_landscape_0 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using landscape distance (distance between persistence landscapes)
Amplitude_landscape_1 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using landscape distance (distance between persistence landscapes)
Amplitude_landscape_2 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using landscape distance (distance between persistence landscapes)
Amplitude_bottleneck_0 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using bottleneck notions of distance
Amplitude_bottleneck_1 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using bottleneck notions of distance
Amplitude_bottleneck_2 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using bottleneck notions of distance
Amplitude_wasserstein_0 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using wasserstein notions of distance
Amplitude_wasserstein_1 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using wasserstein notions of distance
Amplitude_wasserstein_2 (TDA)	Amplitude of a subdiagram as its distance from the diagonal diagram using wasserstein notions of distance
Amplitude_betti_0 (TDA)	Amplitude of a subdiagram as its distance between Betti curves.
Amplitude_betti_1 (TDA)	Amplitude of a subdiagram as its distance between Betti curves.
Amplitude_betti_2 (TDA)	Amplitude of a subdiagram as its distance between Betti curves.
Amplitude_silhouette_0 (TDA)	Amplitude of a subdiagram as its distance between silhouettes
Amplitude_silhouette_1 (TDA)	Amplitude of a subdiagram as its distance between

	silhouettes
amplitude_silhouette_2 (TDA)	Amplitude of a subdiagram as its distance between silhouettes
amplitude_heat_0 (TDA)	Amplitude of a subdiagram as its distance between Gaussian-smoothed diagrams
amplitude_heat_1 (TDA)	Amplitude of a subdiagram as its distance between Gaussian-smoothed diagrams
amplitude_heat_2 (TDA)	Amplitude of a subdiagram as its distance between Gaussian-smoothed diagrams
amplitude_persistence_image_0 (TDA)	Amplitude of a subdiagram as its distance between Gaussian-smoothed diagrams represented on birth-persistence axes
amplitude_persistence_image_1 (TDA)	Amplitude of a subdiagram as its distance between Gaussian-smoothed diagrams represented on birth-persistence axes
amplitude_persistence_image_2 (TDA)	Amplitude of a subdiagram as its distance between Gaussian-smoothed diagrams represented on birth-persistence axes

Known discrepancies:

- Some PS2 data rows will not have a plant_name associated with them. This is due to the fact that some PS2 scans did not overlap with other sensors, which contained individual plant phenotypes. In instances where PS2 did overlap, each individual plant within a plot was given the same PS2 values (F0, FV, FM, FV/FM).