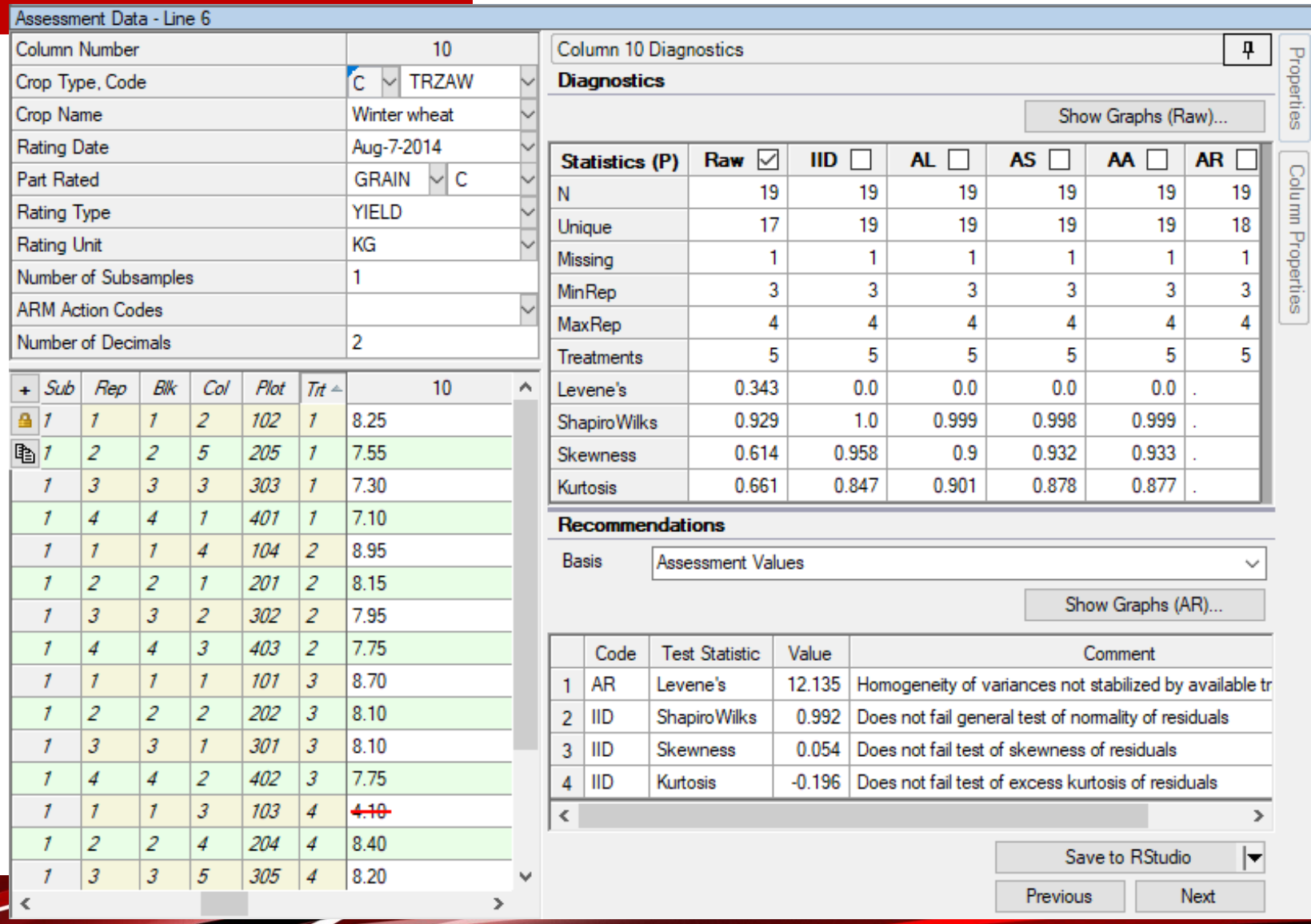
Column Diagnostics & Non-parametric statistics

Column Diagnostics

- New tool for reviewing data
- Statistical tests for assumptions of AOV, run on residuals
- ARM recommends actions from results
- View diagnostic plots of data and residuals
- Find outliers



Column Diagnostics

Diagnostics

Show Graphs (Raw)...

Statistics (P)	Raw ☑	IID ☐	AL ☐	AS ☐	AA ☐	AR ☐
N	19	19	19	19	19	19
Unique	17	19	19	19	19	18
Missing	1	1	1	1	1	1
Treatments	5	5	5	5	5	5
Levene's	0.343	0.0	0.0	0.0	0.0	.
ShapiroWilks	0.929	1.0	0.999	0.998	0.999	.
Skewness	0.614	0.958	0.9	0.932	0.933	.
Kurtosis	0.661	0.847	0.901	0.878	0.877	.

Statistical tests for assumptions of AOV analysis

- Levene's: homogeneity of variance
- Shapiro-Wilk's: general test for normality
- Skewness/Kurtosis: tests for normality

Tests performed on **residuals**, not raw data

Significant P value -> test fails, so assumption is not met

Column Diagnostics

Recommends actions for failed tests of AOV assumptions

- Transformations
- Non-parametric analysis

Recommendations

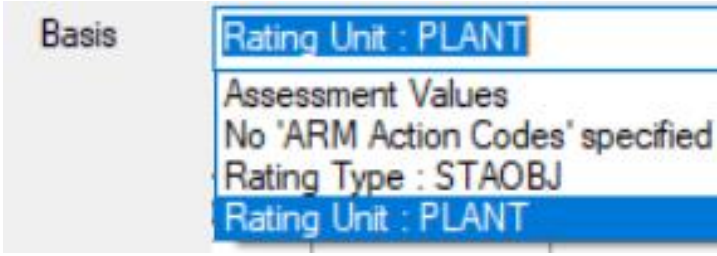
Basis: Rating Unit : PLANT

Show Graphs (AS)...

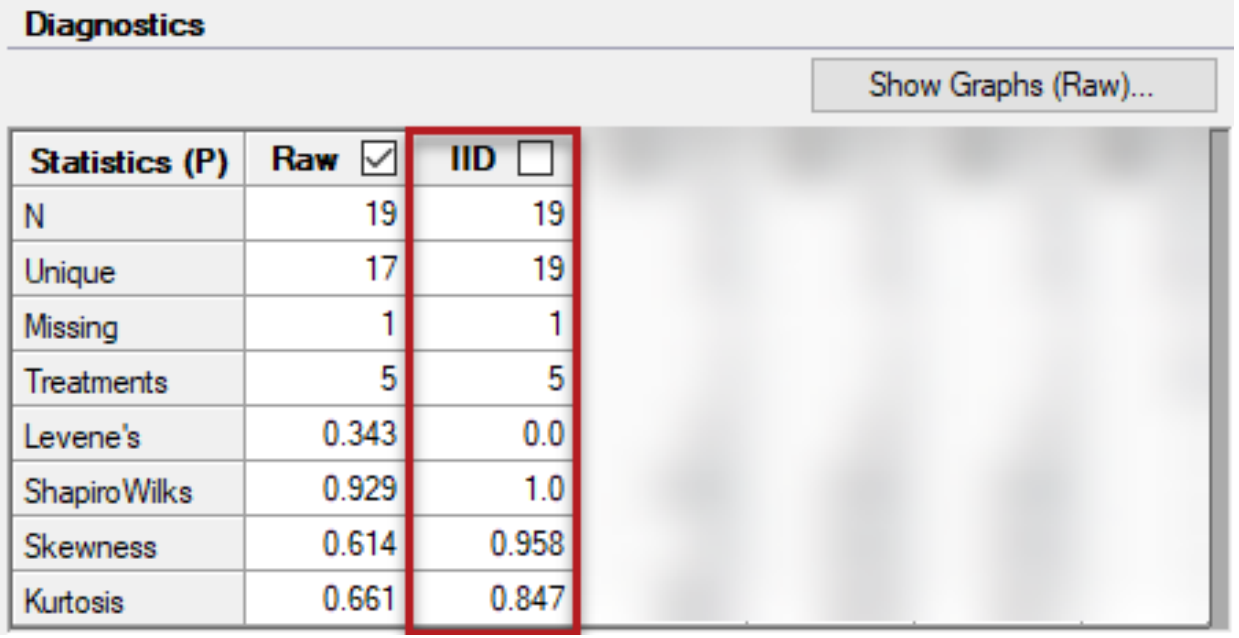
	Code	Test Statistic	Value	Comment
1	AS	Levene's	3.026	Transform to stabilize variance
2	AR	ShapiroWilks	0.873	Available transformations do not improve normality
3	IID	Skewness	0.143	Does not fail test of skewness of residuals
4	AR	Kurtosis	7.126	Available transformations do not correct kurtosis of residuals

Suggestions based on assumed distribution of data

- Determined from: Rating Type, Rating Unit, range of data values



Column Diagnostics



Statistics (P)	Raw ☑	IID ☐
N	19	19
Unique	17	19
Missing	1	1
Treatments	5	5
Levene's	0.343	0.0
ShapiroWilks	0.929	1.0
Skewness	0.614	0.958
Kurtosis	0.661	0.847

IID - new ARM Action Code

- ***I**dentically and **I**ndependently **D**istributed*
- Signifies that column meets all assumptions of AOV

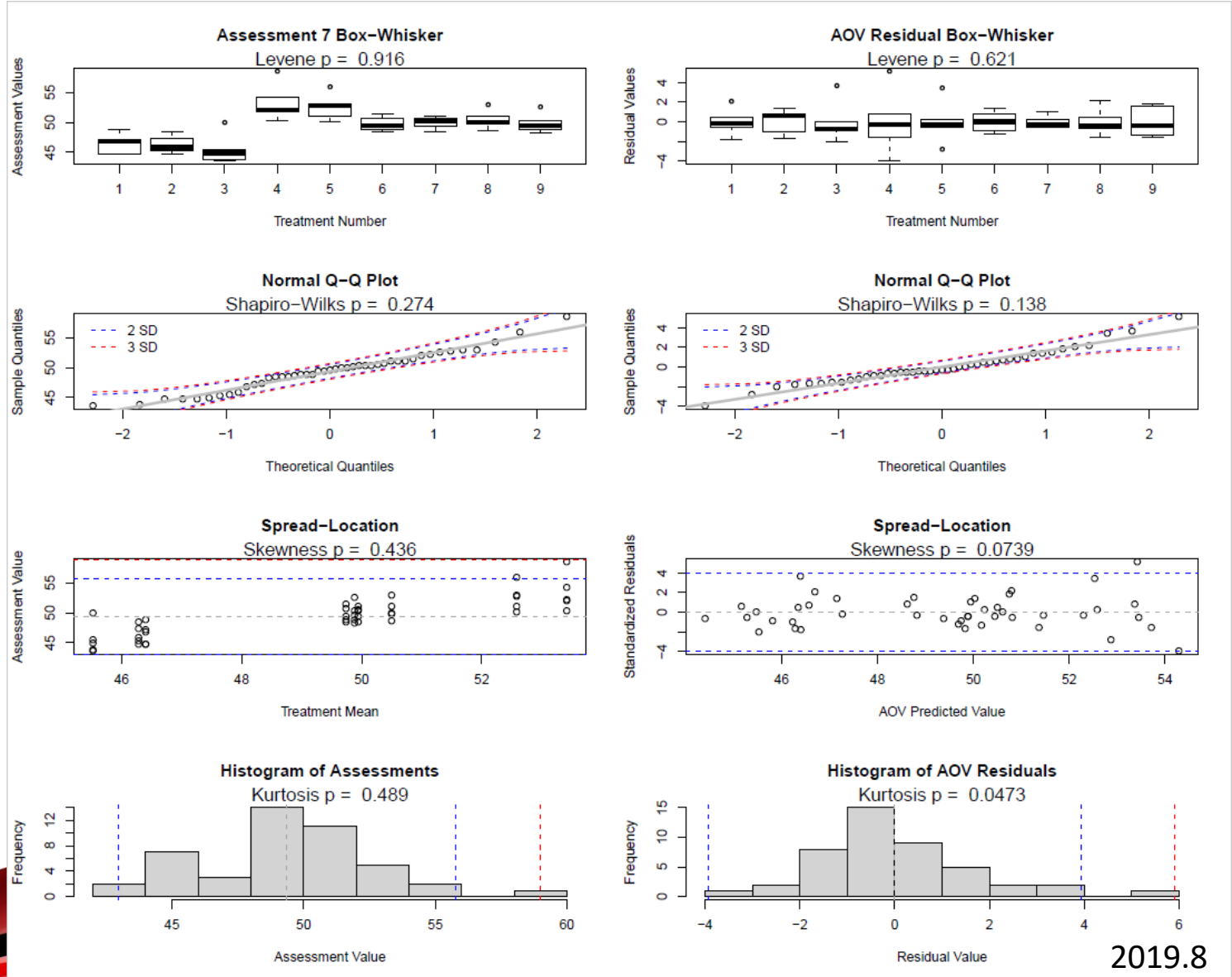
A way to mark column as reviewed, but without any corrections required

AOV analysis is then run on the **residuals** (previously were run on raw data)

Column Diagnostics

View diagnostic plots:

- Data vs. residuals
- AOV residuals vs. transformed residuals



Column Diagnostics

Search for outliers

- Search for outliers based on residuals
- Calculates a studentized residual (accounts for number of observations)

- *(coming soon)* Exclude all calculated outliers in the column

Recommendations

	ActionCode	Criteria	Value	Comment
1	EX	Outliers	NA	Exclude outliers to reduce skewness

Outliers

	plot	treatment	replicate	column	assessment31	StdRes
1	102	9	1	2	1	-4.5

Non-parametric statistics

- Relies on ranks to analyze data, instead of means and st. dev.
- New ARM Action Code: **AR**
  (automatic rank transformation)
  - Rank the data points
    (Kruskal-Wallis or Friedman's test)
  - Perform mean comparison test
    on rank means (LSD)
- Analysis included with other AOV columns

Pest Type			W Weed	W Weed
Pest Name			Oxeye	Oxeye
Rating Date			Sep-4-2016	Sep-4-2016
Rating Type			STAOBJ	STAOBJ
Rating Unit			PLANT	PLANT
Number of Subsamples			1	1
ARM Action Codes				AR
Trt No.	Treatment Name	Appl Code	34	35
1	Burn	A	0.0 c	3.8 b
2	Burn	A	0.1 c	4.7 b
3	Burn	A	0.0 c	4.5 b
4	Mow	A	0.0 c	3.8 b
5	Mow	A	0.0 c	3.8 b
6	Mow	A	0.0 c	3.8 b
7	Spray	A	0.0 c	3.8 b
8	Spray	A	1.7 a	8.9 a
9	Spray	A	0.8 b	7.9 a
LSD P=.05			1.66	1.13
Levene's F			2.794	2.40
Levene's Prob(F)			0.016*	0.035*
Friedman's X2			.	34.526
P(Friedman's X2)			.	0.00
Skewness			2.1418*	0.4492
Kurtosis			3.6057*	4.7736

Reported values for columns 35 are rank means and not assessment means

Why non-parametric???

- ANOVA has assumptions that must be met
  - Otherwise results may not be valid
- Non-parametric analysis does not have these restrictions
- Use non-parametric when:
  - Data cannot be corrected to fit assumptions of AOV
  - Data is not real-valued, like counts or index scales