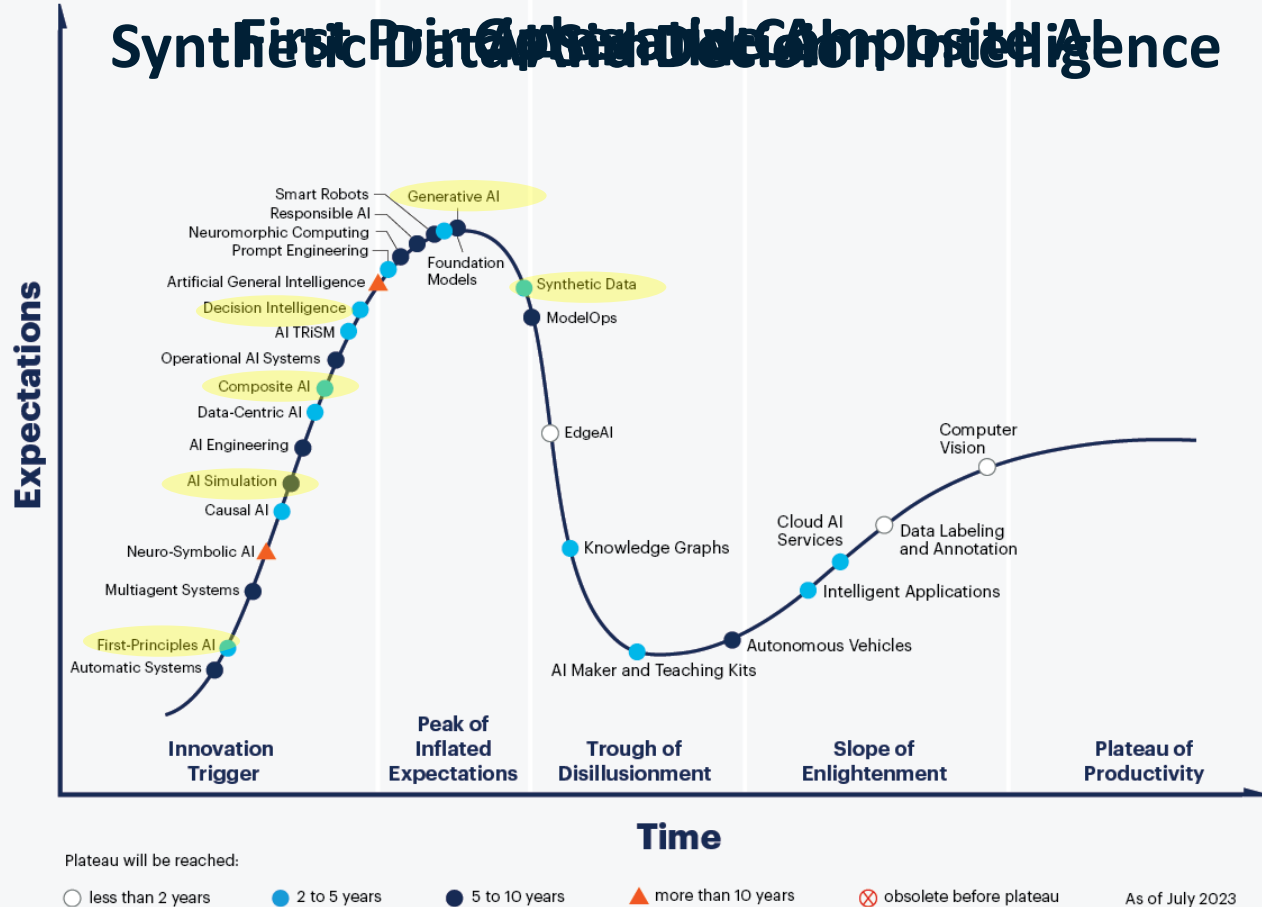


A hand is shown holding a glowing, translucent sphere. The background is dark blue with a network of glowing blue lines and dots, suggesting a digital or data environment. The overall aesthetic is futuristic and technological.

Artificial Intelligence and Analytics Trends

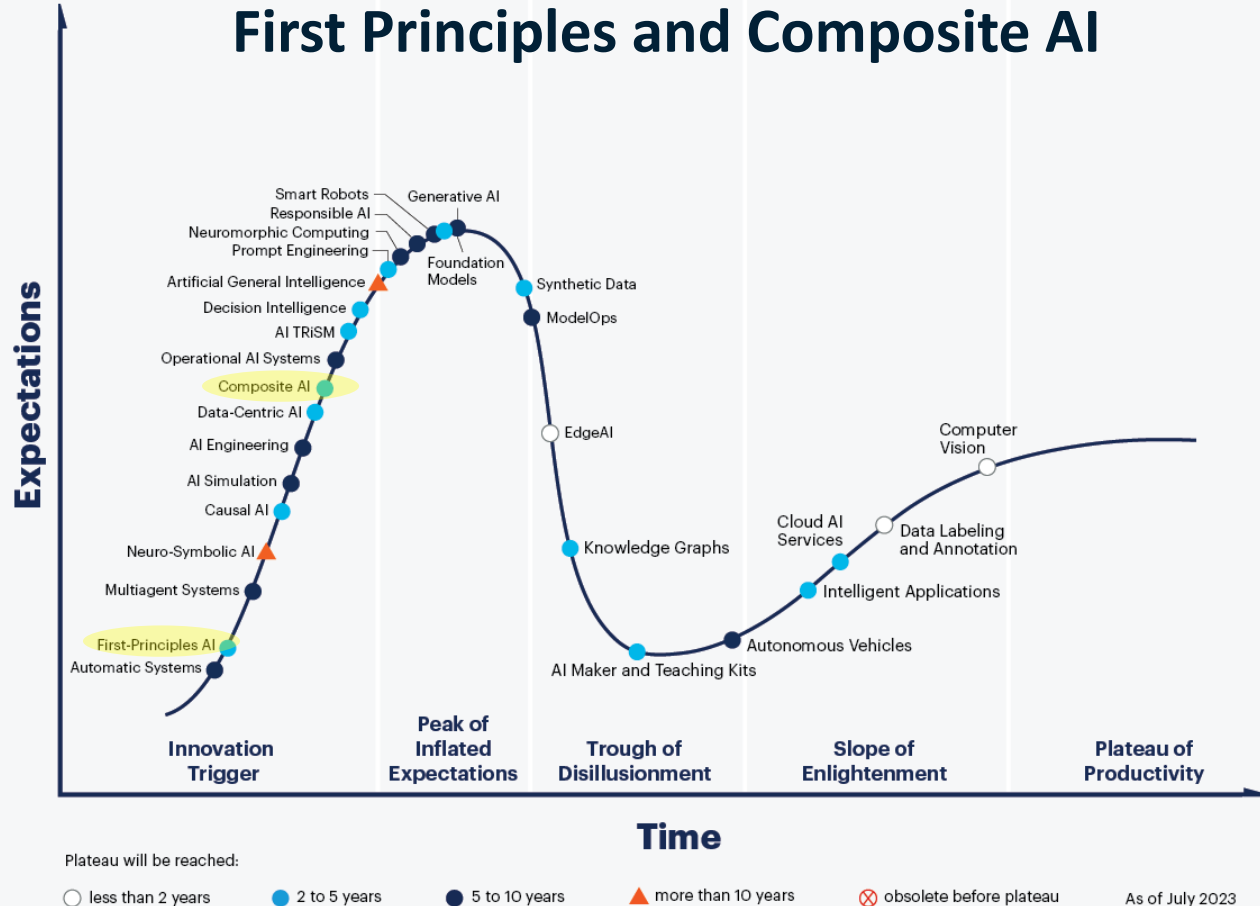
John Gottula
Director of Crop Science
SAS

Hype Cycle for Artificial Intelligence, 2023



Hype Cycle for Artificial Intelligence, 2023

First Principles and Composite AI



Crop Composition Data for EU Regulatory Submission

REGULATIONS

COMMISSION IMPLEMENTING REGULATION (EU) No 503/2013
of 3 April 2013

on applications for authorisation of genetically modified food and feed in accordance with Regulation (EC) No 1829/2003 of the European Parliament and of the Council and amending Commission Regulations (EC) No 641/2004 and (EC) No 1981/2006



Statistical considerations for GMOs safety

SCIENTIFIC OPINION

Scientific Opinion on

Statistical considerations for the safety evaluation of GMOs¹

EFSA Panel on Genetically Modified Organisms (GMO)^{2,3}

European Food Safety Authority (EFSA), Parma, Italy



EFSA Journal 2011; 9(5):2150

SCIENTIFIC OPINION

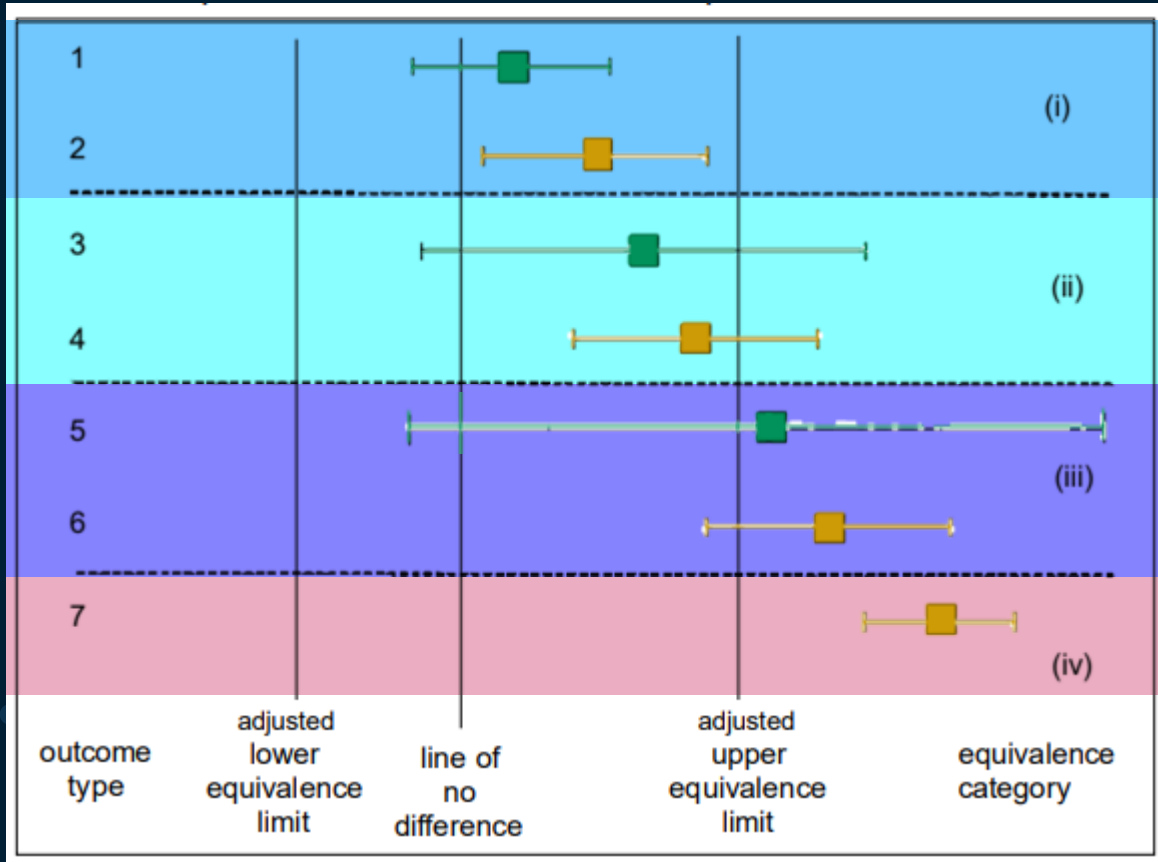
Guidance for risk assessment of food and feed from
genetically modified plants¹

EFSA Panel on Genetically Modified Organisms (GMO)^{2,3}

European Food Safety Authority (EFSA), Parma, Italy

A	B	C	D	E	F	G	H	I	L
analyte Catego	analyte	Unit of Me	genoty	Background	genoty	notype	rep	site	y
Amino Acids	Alanine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	5.65
Amino Acids	Alanine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	5.29
Amino Acids	Alanine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	5.14
Amino Acids	Alanine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	5.71
Amino Acids	Arginine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	3.17
Amino Acids	Arginine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	2.88
Amino Acids	Arginine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	3.02
Amino Acids	Arginine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	3.25
Amino Acids	Aspartic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	4.67
Amino Acids	Aspartic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	4.62
Amino Acids	Aspartic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	4.11
Amino Acids	Aspartic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	4.53
Amino Acids	Cystine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	1.55
Amino Acids	Cystine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	1.53
Amino Acids	Cystine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	1.43
Amino Acids	Cystine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	1.49
Amino Acids	Glutamic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	14.29
Amino Acids	Glutamic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	13.49
Amino Acids	Glutamic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	13.04
Amino Acids	Glutamic Acid	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	13.70
Amino Acids	Glycine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	2.72
Amino Acids	Glycine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	2.57
Amino Acids	Glycine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	2.50
Amino Acids	Glycine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	2.53
Amino Acids	Histidine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	2.17
Amino Acids	Histidine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	2.07
Amino Acids	Histidine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	1.96
Amino Acids	Histidine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	2.20
Amino Acids	Isoleucine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	2.58
Amino Acids	Isoleucine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	2.47
Amino Acids	Isoleucine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	2.29
Amino Acids	Isoleucine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	2.14
Amino Acids	Leucine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	9.61
Amino Acids	Leucine	mg/g DW	Comp	omp Background	Non-GM	Comp	2	Atlantic, IA	9.08
Amino Acids	Leucine	mg/g DW	Comp	omp Background	Non-GM	Comp	3	Atlantic, IA	8.78
Amino Acids	Leucine	mg/g DW	Comp	omp Background	Non-GM	Comp	4	Atlantic, IA	9.56
Amino Acids	Lysine	mg/g DW	Comp	omp Background	Non-GM	Comp	1	Atlantic, IA	2.30

EFSA Difference and Equivalence



Compared to comparator:

No difference

Difference

Compared to references:

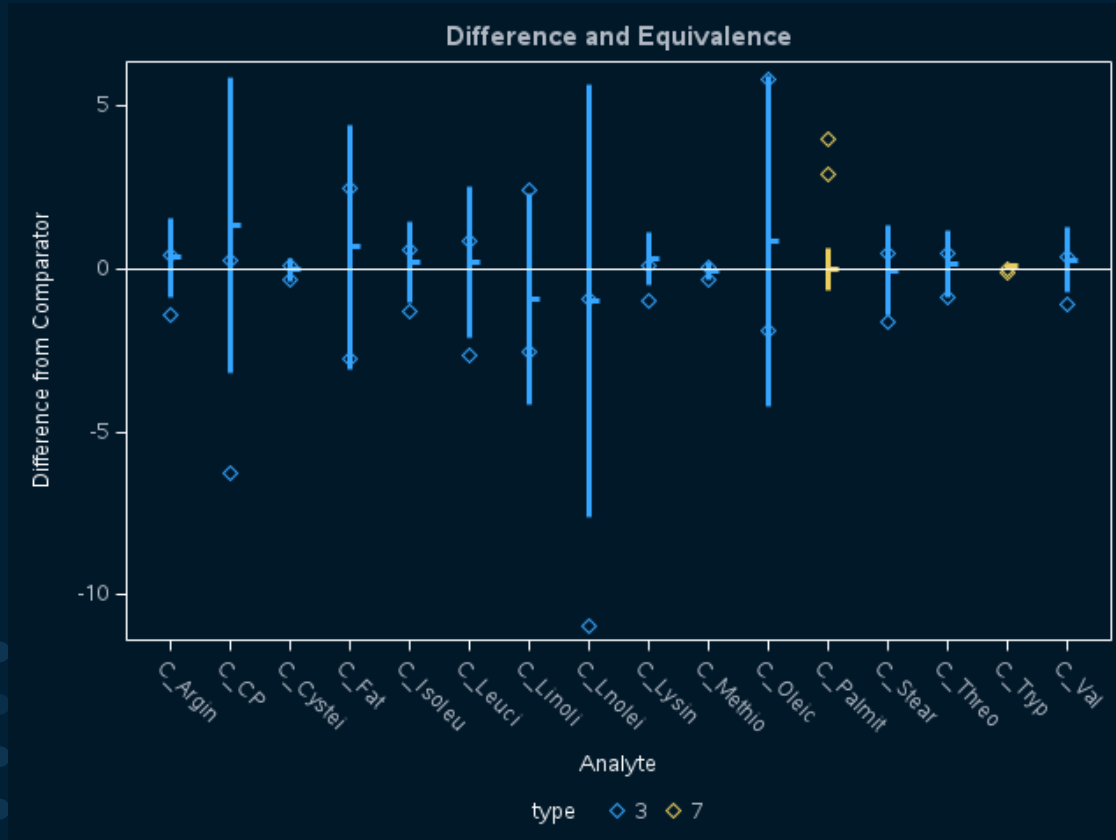
Equivalent

Likely Equivalent

Less Likely Equivalent

Not Equivalent

Common flaws to implement EFSA's approach



- Comparator's genetic composition not equivalent to GMO
- Reference varieties do not capture adequate range of diversity
- Reference varieties different from Comparator

Composite AI – “Self-Service” Analytics

Run Cancel Copy to My Tasks + Code to Flow

Oct 9, 2023, 1:37:22 PM

Data Model Options Output Information

Code Log Results Output Data (1) Edit Code

1 /*
2 *
3 * Task code generated by SAS® Studio 6.0
4 *
5 * Generated on '10/9/23, 1:38 PM'
6 * Generated on SAS version 'V.04.00M0P091823'
7 * Generated on browser 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML,
8 * Generated on web client 'https://agtech2-viya4.eastus2.cloudapp.azure.com/SASStudio/main?lo
9 */
10
11 ods noproctitle;
12 ods graphics / imagemap=on;
13
14 proc sort data=WORK.COMPSTUDY1 out=Work._tempsorted_
15 by Analyte;
16 run;
17
18 proc mixed data=Work._tempsorted_ method=reml alpha=0.1;
19 class genotype rep site Background genotypegroup;
20 model y=genotypegroup / ddfm=kenrog;
21 random rep site rep*site genotype*Background /;
22 weight InverseCv;
23 by Analyte;
24 run;
25
26 proc delete data=Work._tempsorted_
27 run;

WORK.COMPSTUDY1

Filter: (none)

Roles

Dependent variable: *

y

Explanatory Variables

Classification variables:

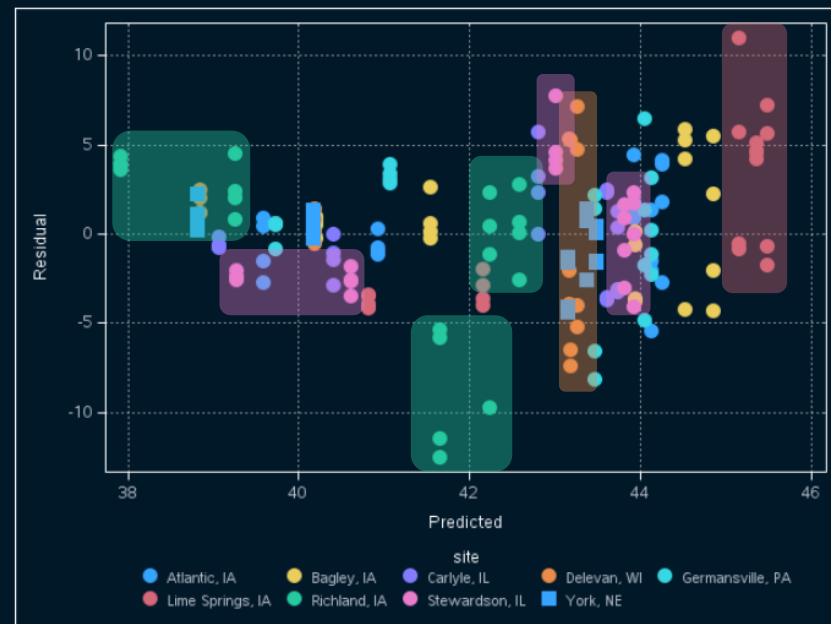
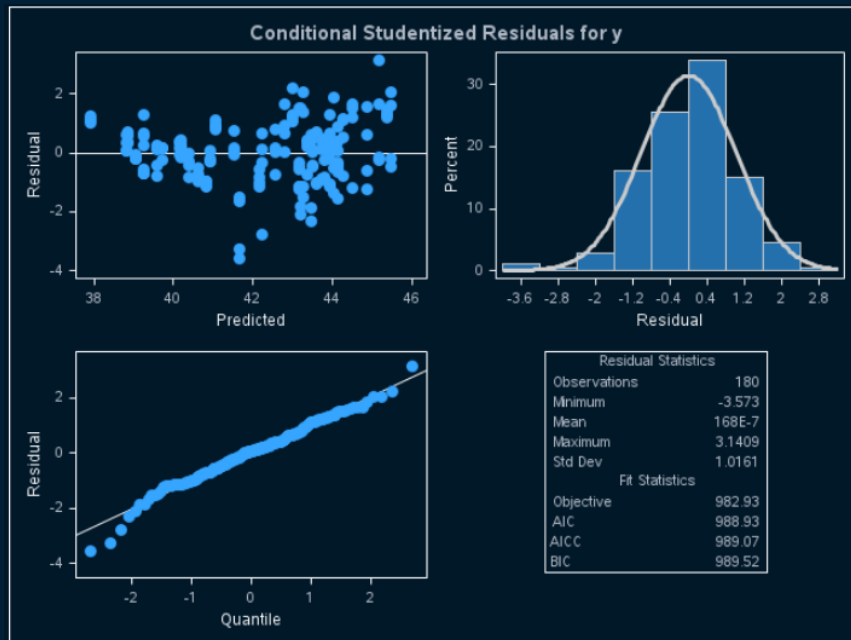
- genotype
- rep
- site
- Background
- genotypegroup

> Treatment of Missing Values

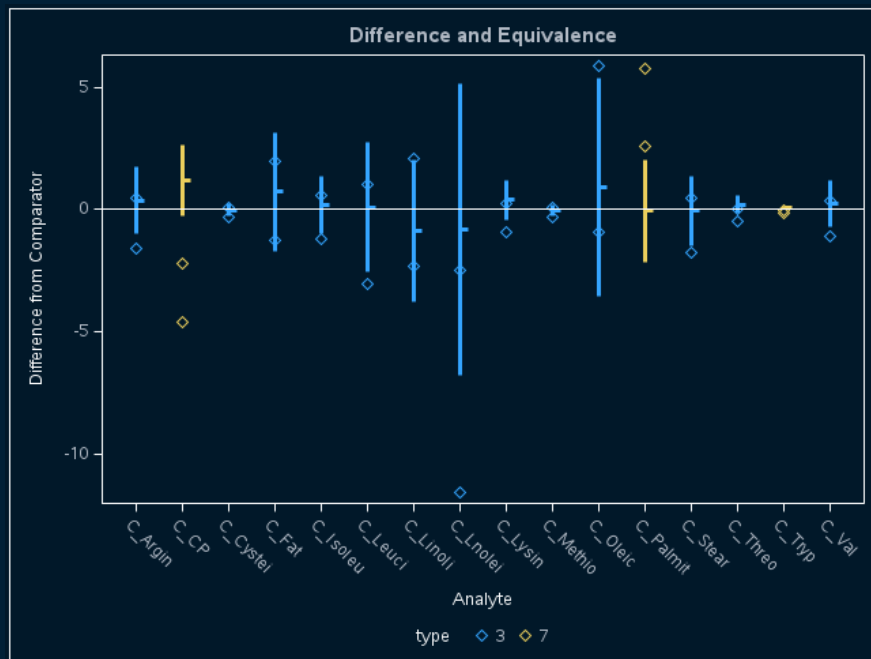
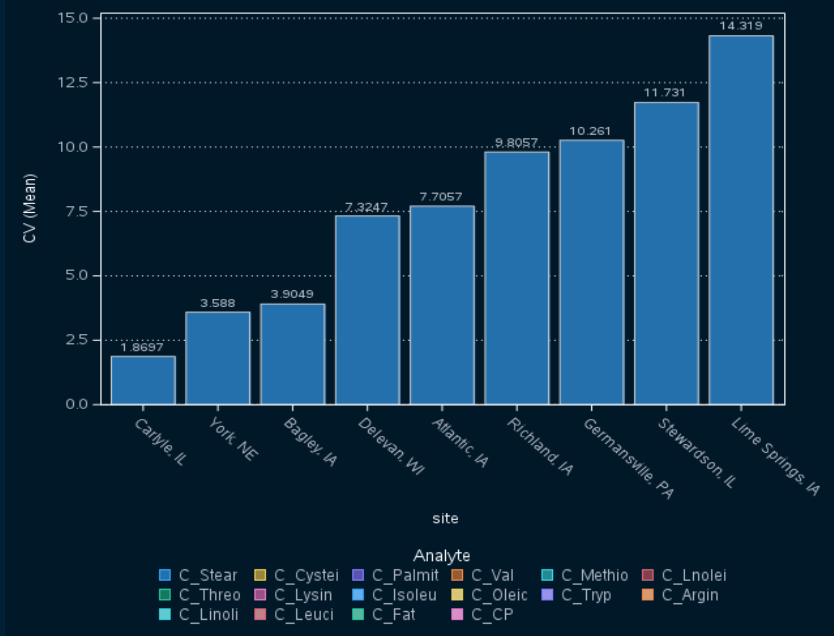
Continuous variables:

Add numeric variables

First Principles AI – Thinking Through the Question



Quantifying and Weighting Errors



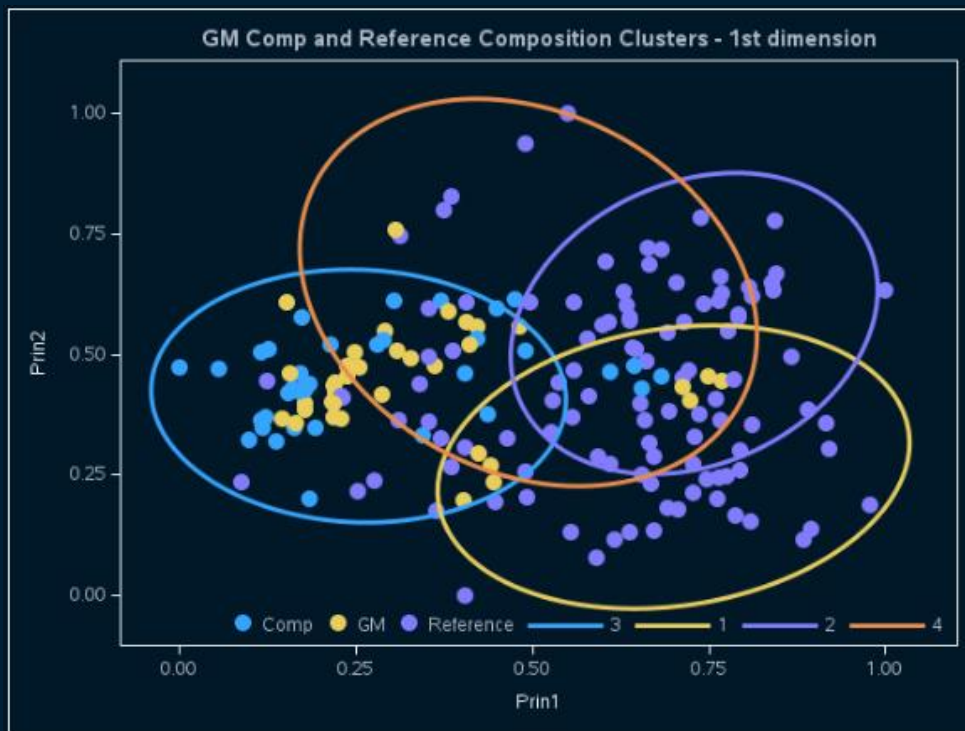
Analyzing Composition Components All at Once?

- Composition endpoints are structurally related
- Composition endpoints have correlated errors
- Independent Analysis of each results in a Multiple Testing Problem

Distill each endpoint into few information rich vectors (e.g. Principal Components)

Generate Clusters among observations

Test GM for 'cluster belonging' and within-cluster variance



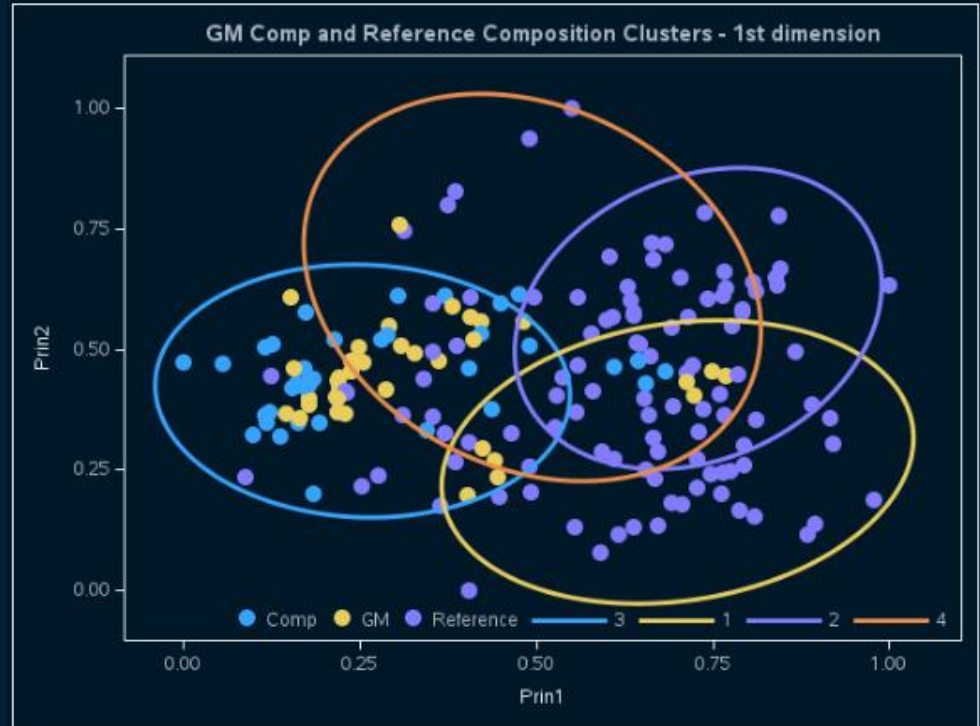
Cluster-based Difference Testing

Difference: Does the GM cluster with its comparator? (χ^2)

Frequency Row Pct	Table of genotypegroup by CLUSTER					
	genotypegroup	CLUSTER(Cluster)				Total
		1	2	3	4	
Comp	0	4	26	6	36	
	0.00	11.11	72.22	16.67		
GM	3	1	29	3	36	
	8.33	2.78	80.56	8.33		
Total	3	5	55	9	72	

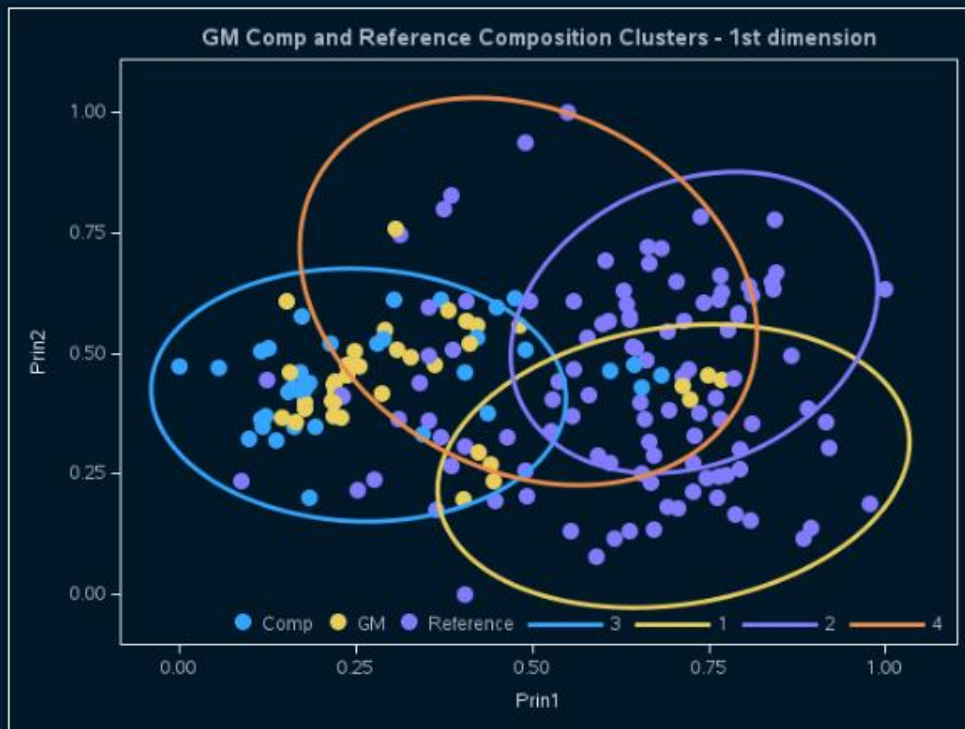
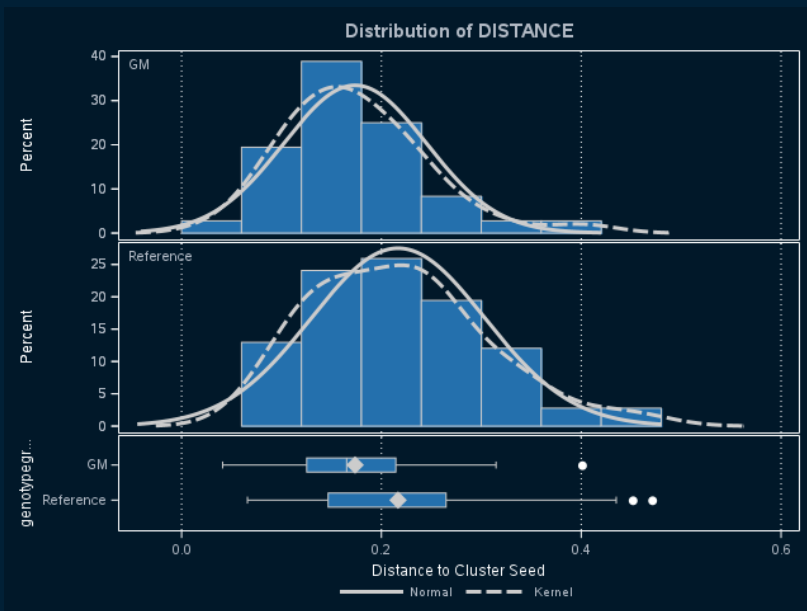
Statistics for Table of genotypegroup by CLUSTER

Statistic	DF	Value	Prob
Chi-Square	3	5.9636	0.1134
Likelihood Ratio Chi-Square	3	7.2694	0.0638
Mantel-Haenszel Chi-Square	1	1.3683	0.2421
Phi Coefficient		0.2878	
Contingency Coefficient		0.2766	
Cramer's V		0.2878	



Cluster-based Equivalence Testing

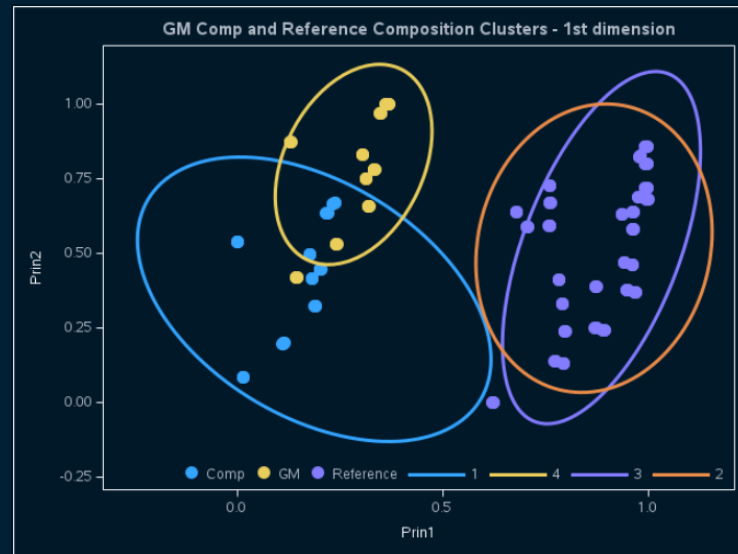
How do GM and Reference compare in their cluster centrality (t-test)



First Principles and Composite AI

Applied to EFSA Difference and Equivalence Testing

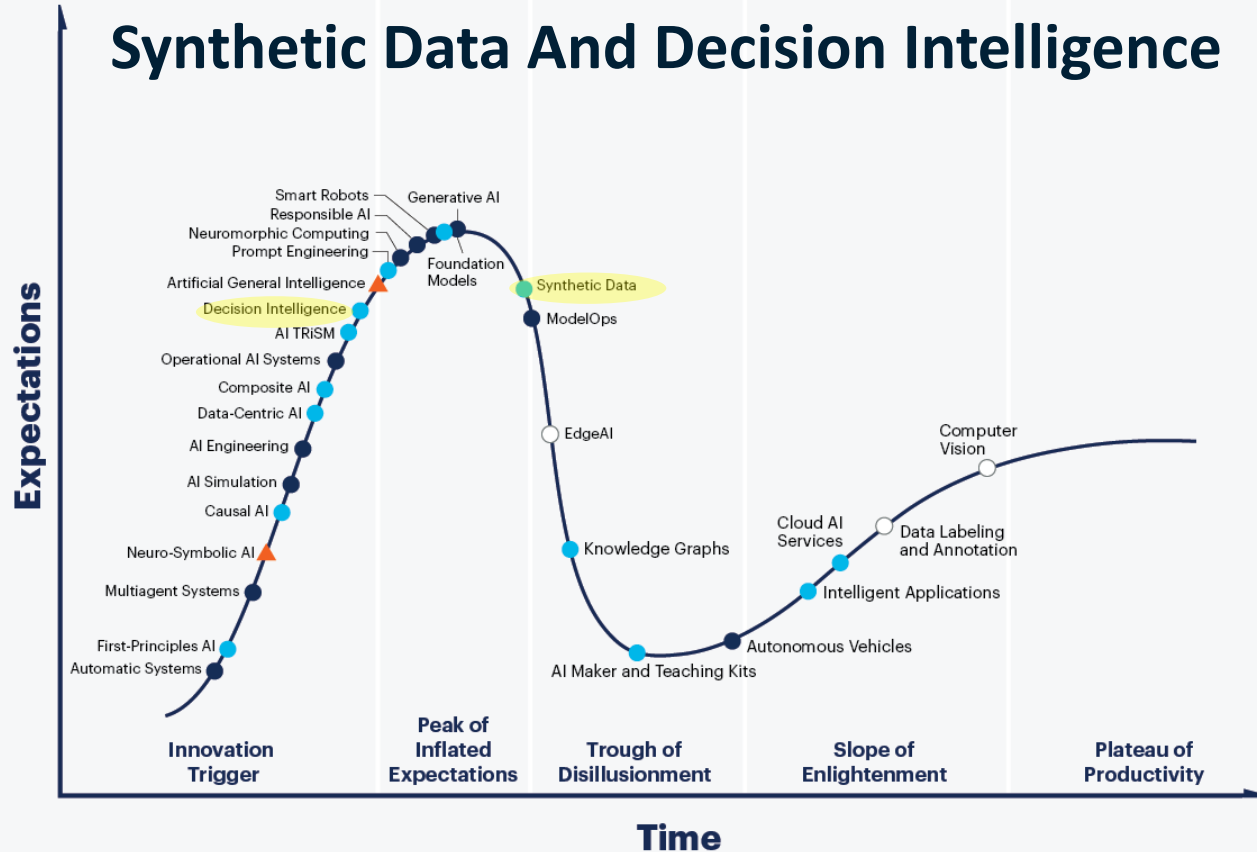
- Self-service data analysis crucial for front-line business and scientific decisions
- Data generators possess invaluable understanding of data (and error) generation process
- Site-Analyte inverse variance weighting leads to less equivalence (and probably more differences)
- Clustering and cluster statistics can test wholistic GMO compositional difference and equivalence



Clustering following best linear unbiased estimates of each site*analyte combo

Hype Cycle for Artificial Intelligence, 2023

Synthetic Data And Decision Intelligence



Plateau will be reached:

○ less than 2 years

● 2 to 5 years

● 5 to 10 years

▲ more than 10 years

⊗ obsolete before plateau

As of July 2023

USDA Germplasm Resources

If your reference varieties aren't working, change them!

USDA-ARS Germplasm Resources Information Network (GRIN)

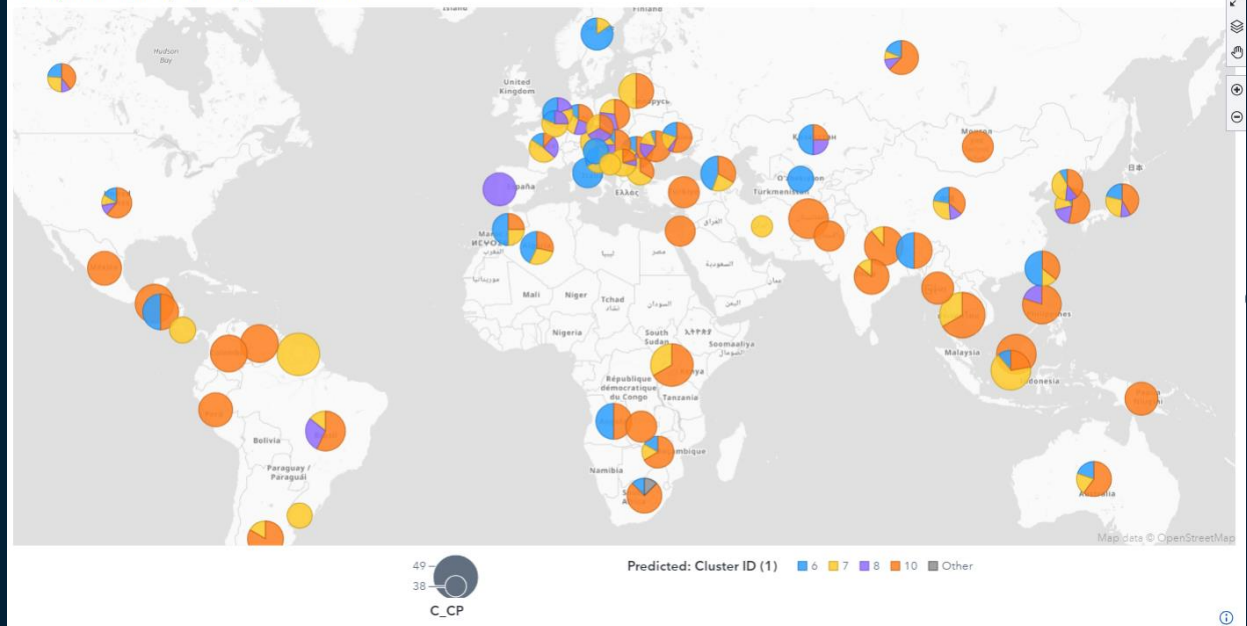
Collections ▾ Crop Germplasm Committees ▾ N

GRIN-Global:
Empowering
and enabling
crop diversity

Search over **600,000** active
accessions

Search

Germplasm Diversity Sized by Crude Protein



Germplasm Phenotypes → case for synthetic data

Before

A	C	G	H	I	J	K	L	M	N	O
D_Access	D_Orig	C_Argin	C_Cystei	C_Isoleu	C_Leuci	C_Lysin	C_Palmit	C_Threo	C_Tryp	C_Val
FC 31557	United States						15.2			
FC 31579	Illinois, Unit	7.48	1.48	4.6	7.84	6.59	10.4	3.67	1.16	5.15
FC 32141	South Dako	7.5	1.5	4.6	8	6.5	11.5	3.7	1.4	5
PI 189926	United States						15			
PI 248398	Illinois, United States									
PI 248402	Illinois, United States						10.7			
PI 355067	Hawaii, United States						11.2			
PI 355068	Hawaii, United States						11.5			
PI 404154	United State	7.57	1.45	4.55	7.87	6.44	11.3	3.58	1.19	5.18
PI 417576	United States						13			
PI 417579	United State	6.9	2	5.5	8.3	6.1	12.05	3.3	1.2	4.9
PI 417582	United State	7.4	1.7	5.1	7.9	6.6	11.9	3.6	1.5	5.2
PI 438492	United States						12.8			
PI 438493	United States						12.3			
PI 438494	United State	7.4	1.5	4.4	7.5	6.2	11.45	3.6	1.3	5
PI 438500	United State	7.3	1.4	5.4	8.1	6.7	10.95	3.5	1.2	4.6
PI 506417	Illinois, Unit	7.6	1.5	5.1	8	7	10.5	3.8	1.2	5
PI 508083	Minnesota,	7.5	1.6	5	8.1	6.8	10.25	3.9	1.3	5.2
PI 508084	Minnesota,	7.3	1.5	4.5	7.6	6.3	11.55	3.7	1.3	5
PI 508266	North Carol	7.6	1.6	4.9	7.7	6.5	12.34	3.6	1.3	5.1
PI 508268	Virginia, Un	7.7	1.7	4.9	8.2	6.6	11.9	3.8	1.4	5.3
PI 508269	Virginia, United States						10.5			

After

D_Access	D_Orig	C_Argin	C_Cystei	C_Isoleu	C_Leuci	C_Lysin	C_Palmit	C_Threo	C_Tryp	C_Val
FC 31557	United Sta	7.43	1.35	4.7	7.94	6.42	15.2	3.59	1.28	4.89
FC 31579	Illinois, Un	7.48	1.48	4.6	7.84	6.59	10.4	3.67	1.16	5.15
FC 32141	South Dak	7.5	1.5	4.6	8	6.5	11.5	3.7	1.4	5
PI 189926	United Sta	7.69	1.55	4.77	8.04	6.9	15	3.61	1.35	4.92
PI 248398	Illinois, Un	7.15	1.43	4.35	7.84	5.96	12.57842	3.57	1.27	4.88
PI 248402	Illinois, Un	7.35	1.59	4.58	7.88	6.8	10.7	3.89	1.32	4.97
PI 355067	Hawaii, Ur	7.62	1.41	4.45	7.55	6.23	11.2	3.39	1.13	4.62
PI 355068	Hawaii, Ur	7.48	1.44	4.42	7.45	6.17	11.5	3.38	1.21	4.85
PI 404154	United Sta	7.57	1.45	4.55	7.87	6.44	11.3	3.58	1.19	5.18
PI 417576	United Sta	7.53	1.49	4.97	8.28	6.27	13	3.46	1.33	4.98
PI 417579	United Sta	6.9	2	5.5	8.3	6.1	12.05	3.3	1.2	4.9
PI 417582	United Sta	7.4	1.7	5.1	7.9	6.6	11.9	3.6	1.5	5.2
PI 438492	United Sta	7.46	1.46	4.77	8.08	6.11	12.8	3.52	1.32	4.99
PI 438493	United Sta	7.5	1.55	4.75	7.94	6.44	12.3	3.67	1.28	5.07
PI 438494	United Sta	7.4	1.5	4.4	7.5	6.2	11.45	3.6	1.3	5
PI 438500	United Sta	7.3	1.4	5.4	8.1	6.7	10.95	3.5	1.2	4.6
PI 506417	Illinois, Un	7.6	1.5	5.1	8	7	10.5	3.8	1.2	5
PI 508083	Minnesota	7.5	1.6	5	8.1	6.8	10.25	3.9	1.3	5.2
PI 508084	Minnesota	7.3	1.5	4.5	7.6	6.3	11.55	3.7	1.3	5
PI 508266	North Carc	7.6	1.6	4.9	7.7	6.5	12.34	3.6	1.3	5.1
PI 508268	Virginia, U	7.7	1.7	4.9	8.2	6.6	11.9	3.8	1.4	5.3
PI 508269	Virginia, U	7.63	1.45	5.41	8.29	6.31	10.5	3.73	1.37	4.98

Imputation can backtrack the most likely values in sparse data sets

...



Correlations

How does it work?

Discover relationships among variables

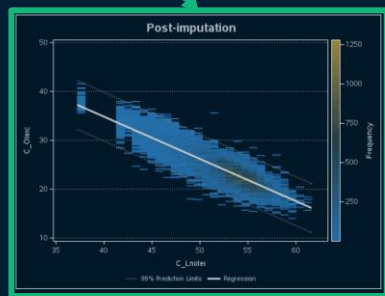
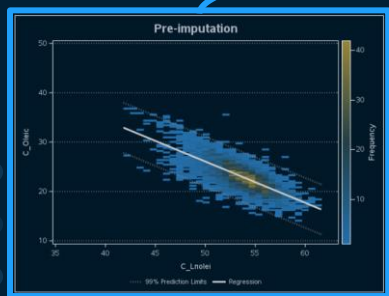


Evaluating Synthetic Data Validity

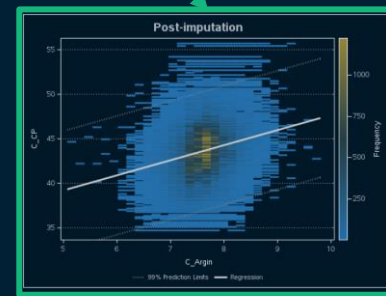
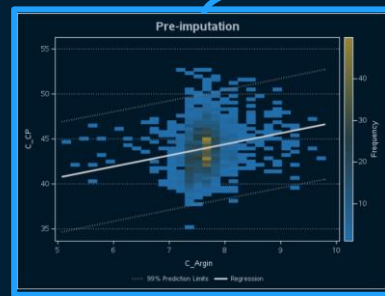
Is there sufficient mutual information?

Group	C_Argin	C_CP	C_Cystei	C_Fat	C_Isoleu	C_Leuci	C_Linoli	C_Lnolei	C_Lysin	C_Methio	C_Oleic	C_Palmit	C_Steat	C_Threo	C_Tryp	C_Val	Freq
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	2066
2	.	X	.	X	.	.	X	X	.	X	X	X	X	.	.	.	521
3	.	X	.	X	.	.	X	X	.	X	316
4	.	X	.	X	.	.	X	X	.	.	X	X	X	.	.	.	2013
5	.	X	.	X	.	.	X	X	18
6	.	X	.	X	4
7	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	91

Are the output relationships reasonable?



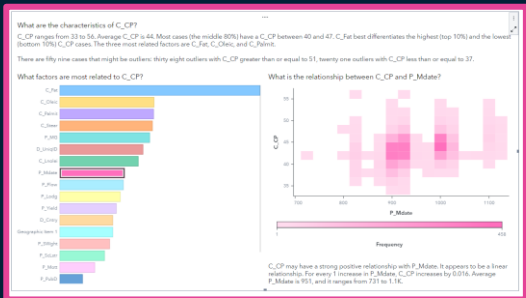
Linoleic vs Oleic



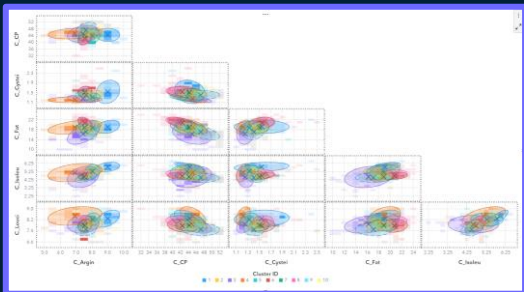
Arginine vs Crude Protein

Decision Intelligence

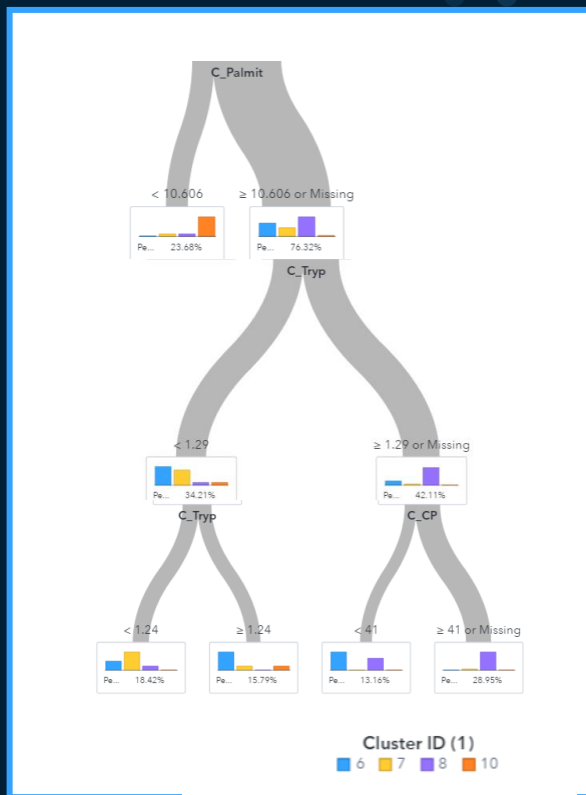
Identify important variables



Assess scenarios



Undertake classification



With human-in-the-loop

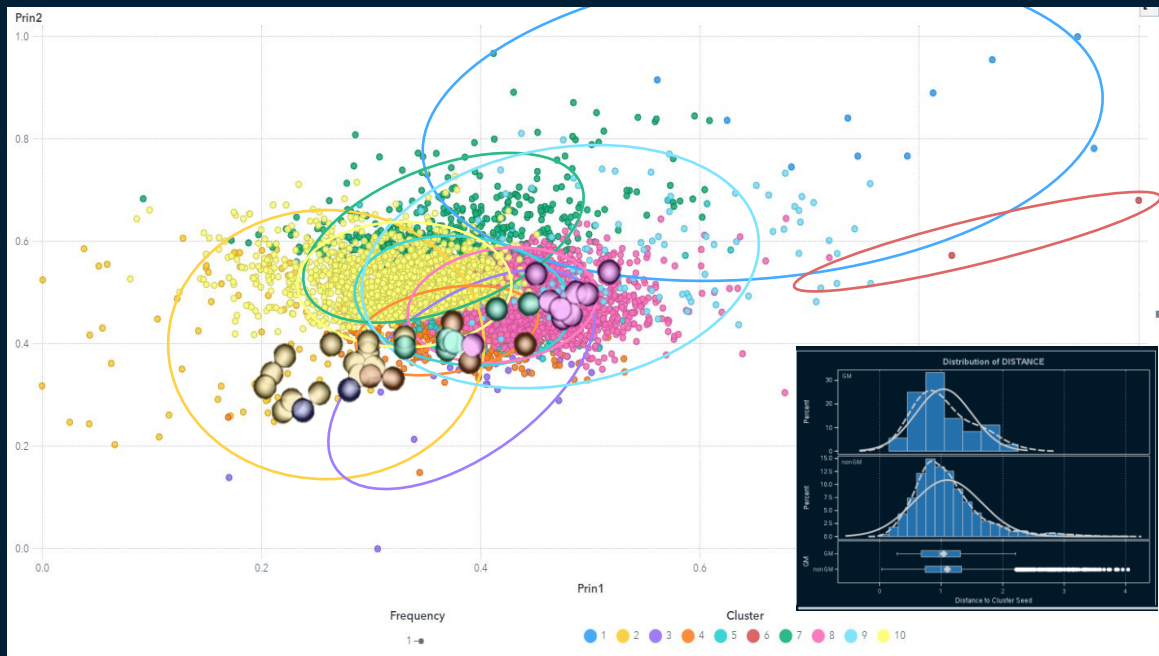
Take the decision

Cluster ID (1)	D_Orig	High_Yield	MG3	D_Access	C_Tryp	C_CP	C_Palmit	P_Mdate	P_SWght
6	Illinois, United States	High_Yield	MG3	PI 548628	1.29	43	12.15	1002	19
6	Iowa, United States	High_Yield	MG3	PI 548543	1.27	43	10.9	927	20
6	Ohio, United States	High_Yield	MG3	PI 584470	1.24	44	12	925	23
6	Nebraska, United States	High_Yield	MG3	PI 595754	1.30	41	10.65	928	16
6	United States	High_Yield	MG3	PI 404154	1.19	44	11.3	923	16
7	Illinois, United States	High_Yield	MG3	PI 597385	1.20	42	11.8	919	16
8	Illinois, United States	High_Yield	MG3	PI 518675	1.30	42	11.85	912	15
8	Ohio, United States	High_Yield	MG3	PI 567902	1.40	42	11.1	929	16
8	Nebraska, United States	High_Yield	MG3	PI 610670	1.37	41	10.65	924	17
8	Nebraska, United States	High_Yield	MG3	PI 595753	1.40	41	10.2	929	17
8	Illinois, United States	High_Yield	MG3	PI 593257	1.40	42	11.3	922	15
8	Illinois, United States	High_Yield	MG3	PI 548556	1.23	44	10.866666667	1004	17
8	Illinois, United States	High_Yield	MG3	PI 548362	1.30	42	11.65	929	16
10	Iowa, United States	High_Yield	MG3	PI 548522	1.31	45	10.6	929	19
10	Indiana, United States	High_Yield	MG3	PI 548585	1.23	42	10.5	921	16
10	Ohio, United States	High_Yield	MG3	PI 593463	1.34	43	8.8	929	18

Synthetic Data and Decision Intelligence

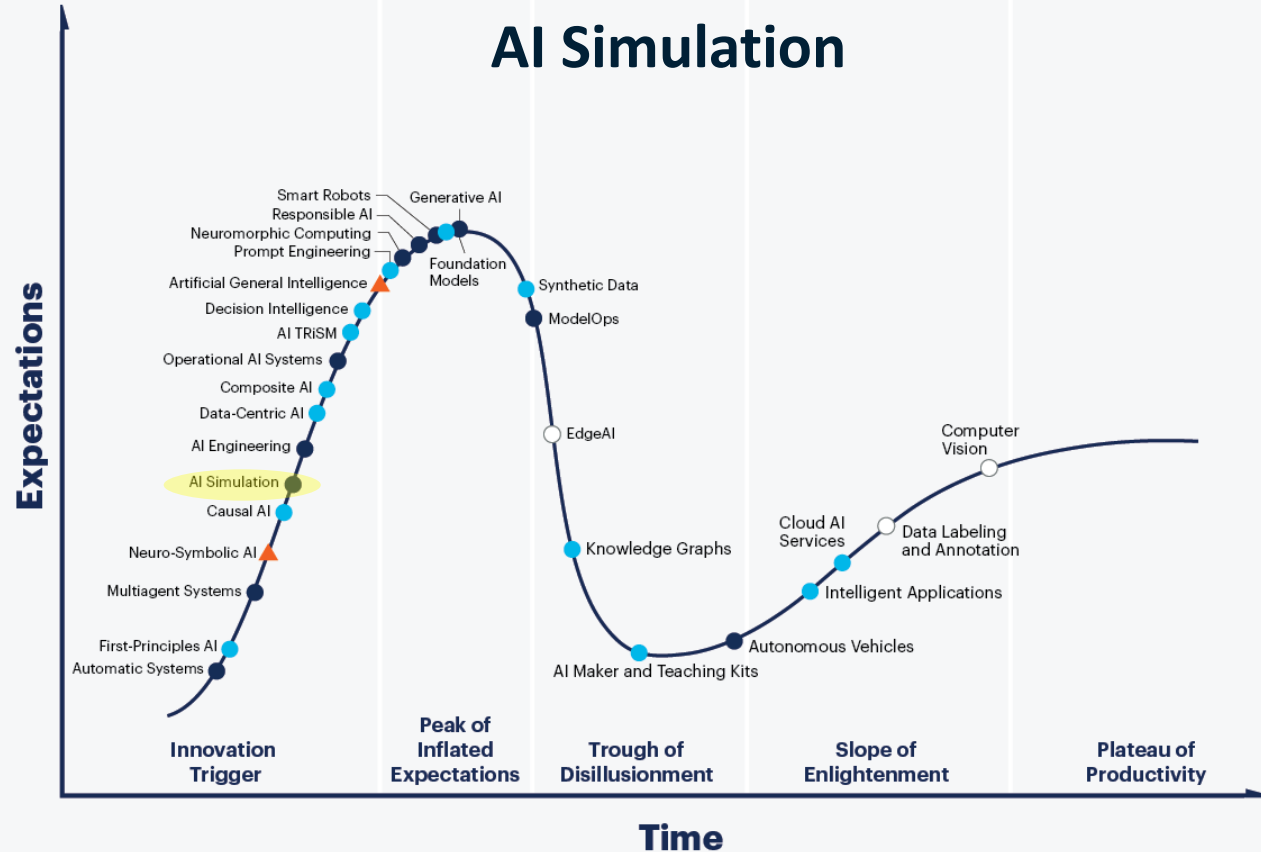
Applied to Reference Varieties

GM distribution across axes of soybean compositional variability



- Synthetic data is an enabler to create models where data is sparse or unknown
 - Imputation is one form of synthetic data
- Decision intelligence reduces intensive data sets into actionable classifications
- USDA Germplasm a useful source of reference information

Hype Cycle for Artificial Intelligence, 2023



Plateau will be reached:

○ less than 2 years

● 2 to 5 years

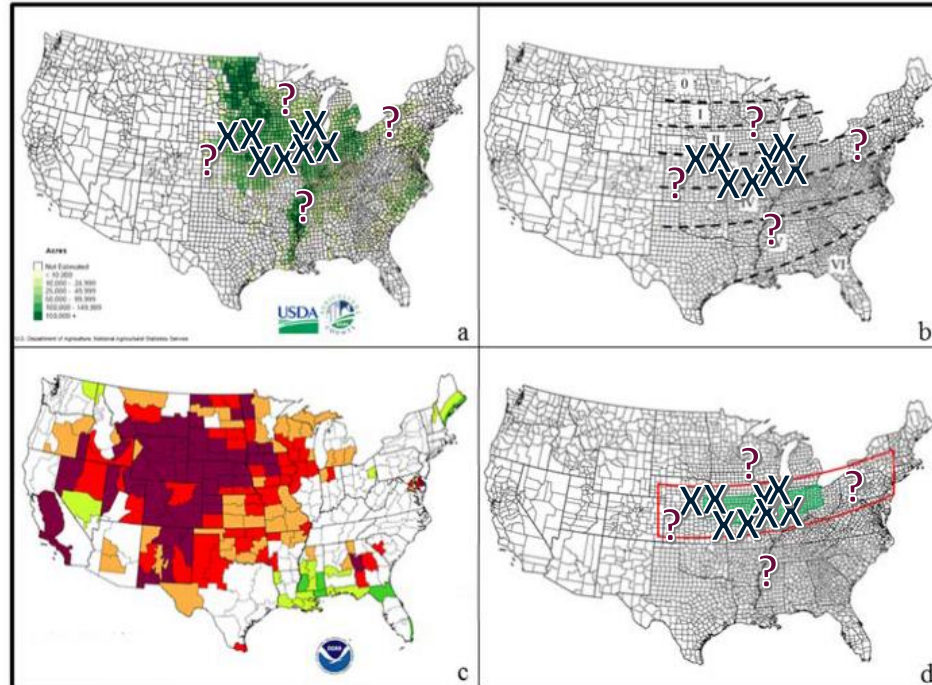
● 5 to 10 years

▲ more than 10 years

⊗ obsolete before plateau

As of July 2023

AI Simulation



Data Simulation

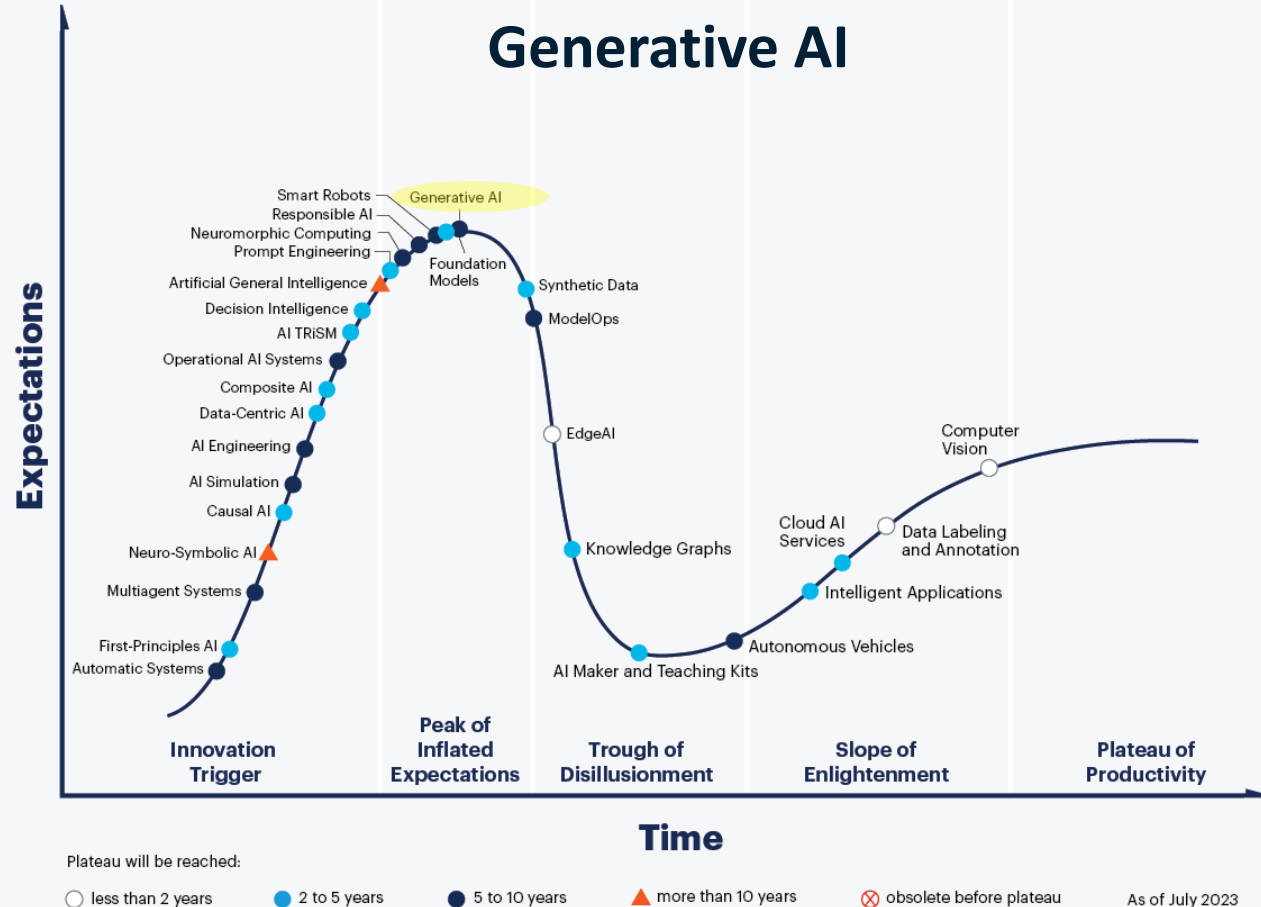
Applied to Compositional Extrapolation

- Algorithms associated with statistics and machine learning can both provide estimates and predictions
- Partitioning into training and holdout data sets ensures generalizability
- Continuous covariate regressors provide a basis for extrapolation
- Anything can be modeled... do you have the data to do it?



Model to predict reference varieties in a sea of germplasm based on composition

Hype Cycle for Artificial Intelligence, 2023



Generative AI Application

What is the best way to analyze GMO composition data to assess safety?:

Analyzing GMO (Genetically Modified Organism) composition data to assess safety typically involves a multi-faceted approach that combines various methods and considerations. Here's a general guideline for analyzing GMO composition data:

Define the Safety Parameters: Identify the specific safety parameters you want to assess, such as allergenicity, toxicity, and nutritional equivalence.

Collect Comprehensive Data: Gather data on the composition of the GMO and its non-GMO counterpart, including proteins, carbohydrates, fats, vitamins, minerals, and other relevant components.

Statistical Analysis: Use statistical analysis to compare the GMO and non-GMO data sets. Common statistical methods include t-tests, ANOVA, and multivariate analysis techniques to detect any significant differences.

Evaluate Known Allergens and Toxins: Assess whether the GMO contains known allergens or toxins and compare their levels to established safety thresholds.

Use Bioinformatics: Employ bioinformatics tools to analyze the DNA and protein sequences of the GMO for potential allergenicity or toxicity. Compare these sequences to established databases of known allergens and toxins.

Generative AI Application

Tell me more about statistics and machine learning for nutritional equivalence

Statistics and machine learning techniques play a vital role in assessing nutritional equivalence between genetically modified organisms (GMOs) and their non-GMO counterparts. Here's how they can be applied:

Data Collection and Preprocessing:

Collect data on the nutritional composition of GMOs and non-GMOs, including information on proteins, carbohydrates, fats, vitamins, minerals, and other relevant nutrients.

Preprocess the data by handling missing values, outliers, and normalizing variables to ensure that it's suitable for analysis.

Descriptive Statistics:

Start with basic descriptive statistics to summarize and visualize the data. This can include measures like mean, median, standard deviation, and histograms to understand the distribution of nutritional components.

Hypothesis Testing:

Use statistical hypothesis testing techniques to determine if there are significant differences in nutrient levels between GMOs and non-GMOs. For instance, you can perform t-tests or ANOVA to assess differences in means.

Machine Learning Classification:

Employ machine learning classification algorithms to predict whether a sample is GMO or non-GMO based on its nutritional composition. Feature selection and engineering are crucial in this step.

Common algorithms include logistic regression, decision trees, random forests, and support vector machines.

Generative AI Application

Tell me about a robust data analysis process to address EFSA GMO crop composition requirements for import registration. Apply the rules of European Commission Implementing Regulations 503/2013, EFSA Guidance for risk assessment of food and feed from genetically modified plants, EFSA Statistical consideration for the safety evaluation of GMOs, EFSA Guidance on the agronomic and phenotypic characterisation of genetically modified plants, and other relevant European sources. Apply additional well-suited statistical techniques to the problem to support a more rigorous safety evaluation within the data presented.

Addressing EFSA GMO crop composition requirements for import registration involves a systematic and robust data analysis process that complies with European Commission regulations and EFSA guidelines. Here's a step-by-step approach:

Step 1: Data Collection

1.1. Gather comprehensive data on the genetically modified (GM) crop and its non-GM counterpart(s), including information on genetic modifications, cultivation practices, and environmental conditions.

Step 2: Data Preprocessing

2.1. Clean the data to remove any outliers or errors.

2.2. Normalize or transform the data as necessary to ensure comparability.

2.3. Separate the data into appropriate categories or groups based on crop type, location, and other relevant factors.

Step 3: Statistical Analysis

3.1. Descriptive Statistics:

Calculate summary statistics (mean, median, standard deviation, etc.) for each compositional parameter in GM and non-GM crops.

Visualize data using histograms, box plots, and scatter plots to identify trends and differences.

Generative AI Application

Also, describe an audacious and popular modeling technique that uses available data and would suitably test GMO compositional difference and equivalence.

An audacious and popular modeling technique that can be applied to test GMO compositional differences and equivalence is Machine Learning (ML), specifically Random Forest:

Data Preparation: Utilize the preprocessed data as input features and label GM and non-GM crops. Ensure the data is properly balanced and split into training and testing sets.

Random Forest Modeling:

Train a Random Forest classifier to distinguish between GM and non-GM crops based on their compositional data.

Random Forest can handle high-dimensional data, capture complex interactions, and provide feature importance rankings.

Feature Importance Analysis: Use the Random Forest model to identify which compositional parameters contribute the most to the classification decision. This helps pinpoint the most critical differences between GM and non-GM crops.

Validation and Interpretation:

Validate the model's performance using appropriate metrics (accuracy, precision, recall, etc.) on the testing dataset.

Interpret the model's results to assess the significance and biological relevance of compositional differences.

Equivalence Testing:

Integrate the model's predictions with equivalence testing techniques to determine if the observed differences are within acceptable limits.

Common AI excellence threads across organizations



Metadata

- Common definitions
- Data Dictionary



Data Ownership

- Go-to-person
- Usage guidelines



Data trust

- Consistency across sources
- Ease of validation



Data Access

- Ease of use
- Tools



Reporting / Analytics

- Single access point
- Growing need



Data Integration

- Unfragmented data
- Data ready for use



Architecture

- Standards
- Scalable solution(s)



Culture

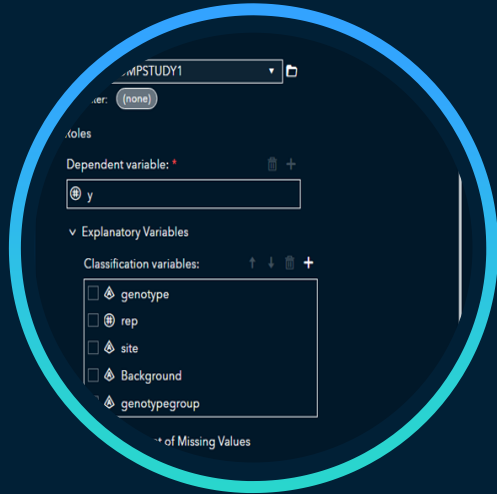
- Break division silos
- Create transparency



Data Regulation

- Data provenance
- Firewalls

How SAS supports 21st century regulatory science teams



Facile – for all study personnel



Traceable and integrated



Auditable from data to report

Overall Summary

Trends in AI and Analytics

- Composite and First Principles AI enables better [statistical] inference
- Synthetic data and decision intelligence facilitate faster and more robust decision frameworks
- When done well, AI simulation supports extrapolation to ‘edge scenarios’.
- Generative language models can summarize lots of information quickly and can spark ideation
- SAS enables regulatory scientists to adopt and expand multimodal analytics

Questions