Generative AI

Die Position von SAS und die ethischen Herausforderungen einer neuen Ära der künstlichen Intelligenz

Anette Almer, Advisory Solutions Architect, SAS
Tamara Fischer, Principal Data Scientist, SAS

"Generative AI (GAI) is a category of AI that can **create new content**, including video, audio, images and text." *SAS*

"A generative model can take what it has learned from the examples it's been shown and **create something entirely new** based on that information. Hence the word "generative!"" *Google*

"Generative artificial intelligence [...] is artificial intelligence **capable of generating** text, images, or other media, using generative models. Generative AI models learn the patterns and structure of their input training data and then generate new data that has similar characteristics." *Wikipedia*

"Despite its capabilities, GPT-4 has similar limitations as earlier GPT models. Most importantly, it still is not fully reliable (it "hallucinates" facts and makes reasoning errors). Great care should be taken when using language model outputs, particularly in high-stakes contexts, with the exact protocol (such as human review, grounding with additional context, or avoiding high-stakes uses altogether) matching the needs of a specific use-case."

https://openai.com/research/gpt-4

Principles for Responsible Usage of AI

SAS Data Ethics Practice

Human-centricity

Putting people at the forefront

Transparency

Understanding the reasons behind development

Robustness

Awareness of limitations and risks

Privacy and security

Keeping everything safe

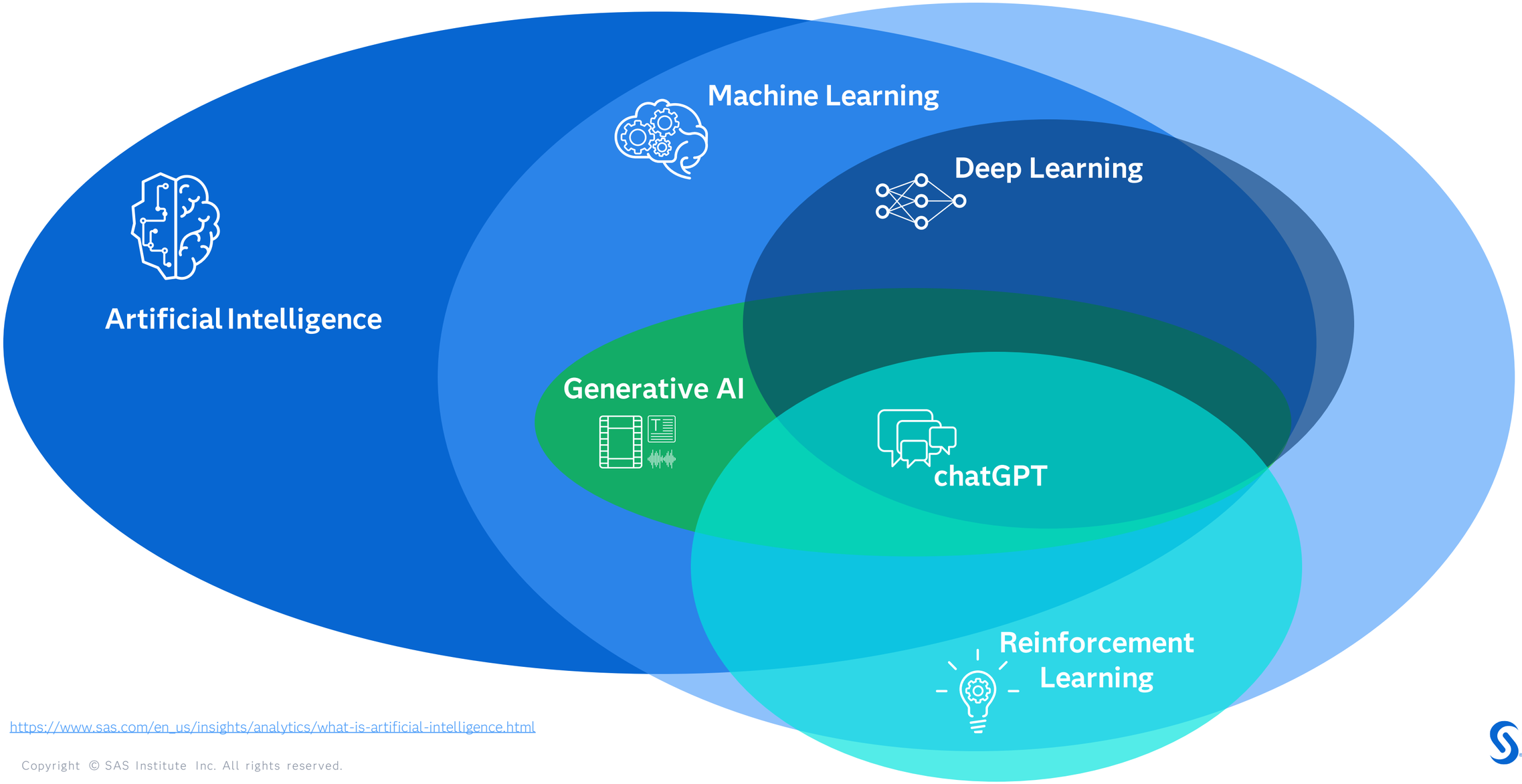
Inclusivity

Recognizing diverse needs and perspectives

Accountability

Prioritizing receiving feedback

Relationship Between AI, ML, DL, RL and GenAI

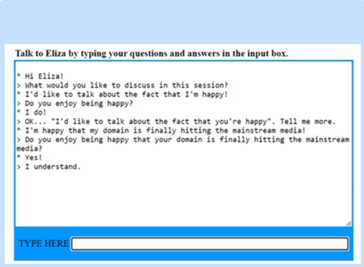


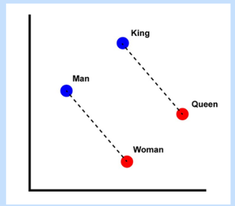
https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html

A Brief History of Modern ChatBots

1966

ELIZA Pattern Matching

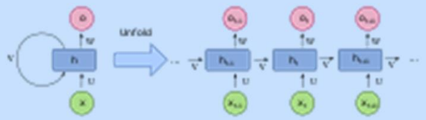




https://en.wikipedia.org/wiki/Word2vec#/media/File:Word_vector_illustration.jpg

word2vec

Recurrent Neural Network

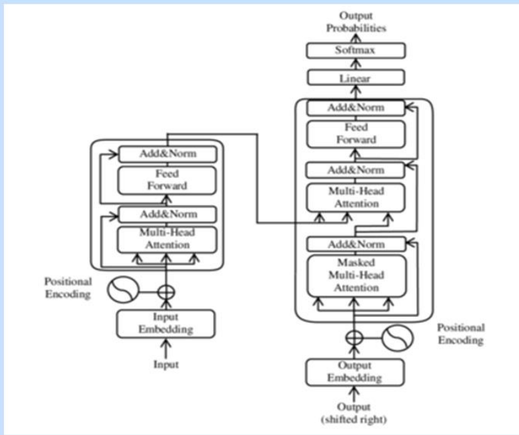


https://en.wikipedia.org/wiki/Recurrent_neural_network

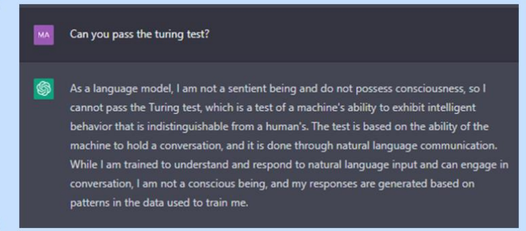
https://en.wikipedia.org/wiki/Recurrent_neural_network

Long Short-Term Memory Network

Transformers & Large Language Models



https://en.wikipedia.org/wiki/Transformer_(machine_learning_model)



GPT-3.5 or chatGPT

2022

https://blogs.sas.com/content/subconsciousmusings/2023/02/13/curious-about-chatgpt/

How chatGPT Works...

https://arxiv.org/abs/2203.02155

Demo

David Weik, Sr Pre-Sales Solutions Architect

SAS Portal

Conversational AI Life Cycle

Conversational AI Life Cycle

Enter your analysis question here...

Analyze

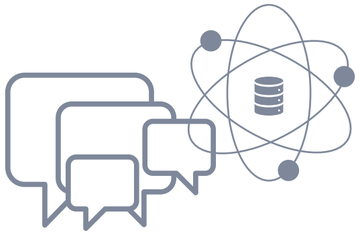
Activate full automation

Review Every Step

Step	API-Endpoint	Request	Result
Init Conversational AI Life Cycle	self	Waiting for user request	-

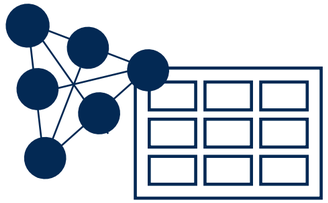
SAS Considers Three Categories to be Generative in Nature

Large Language Models



A natural language processing neural network-based model with millions or billions of parameters & trained on a massive corpus of real-world data that tracks & learns relationships in data to respond (in the form of novel language, summarization, classification, & question answering) to prompts

Synthetic Data Generation



On demand, self-service, or automated data generated by algorithms or rules, vs gathered in the real-world, to meet conditions lacking in real data

Synthetic data reproduces the same statistical properties, probability, & characteristics of the real-world dataset from which the synthetic data is trained

Considered to be a privacy preserving technique

Digital Twin

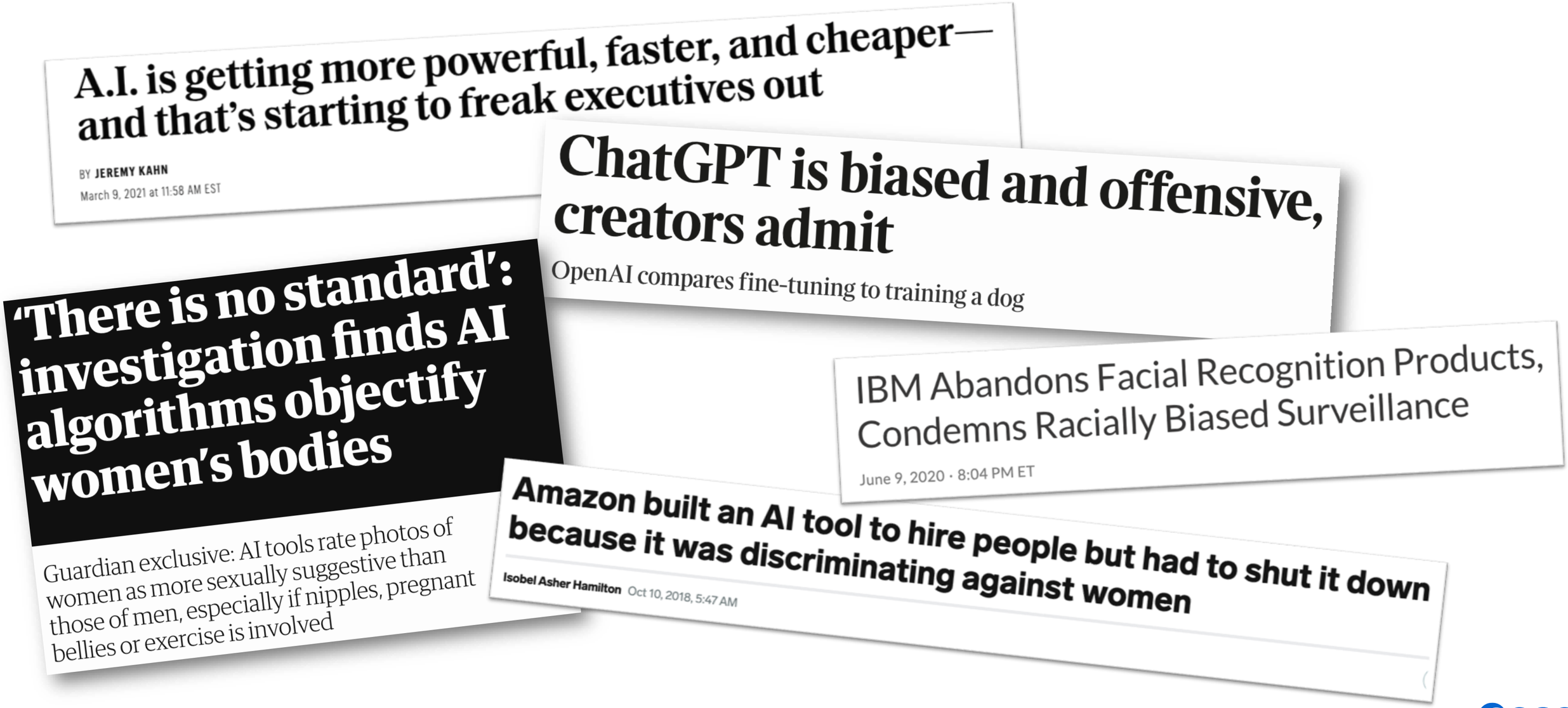
A digital, animate, dynamic ecosystem – comprised of an interconnected network of software, generative & non-generative models, & (historical, real-time, & synthetic) data – that both mirrors & synchronizes with a physical system

Digital twins simulate "what-if" scenarios & stress test systems in the digital world to prescribe actions that optimize the physical world – to improve the lives of individuals, populations, cities, organizations, the environment, systems, products, & more

Large Language Model Considerations

1. Expensive to build
2. Expensive to run
3. May not utilize recent data
4. May use prompts as training data
5. May present IP, reputational, security risks
6. May be biased
7. May not be accurate
8. May not be open
9. Currently, not explainable
10. Has a large carbon footprint

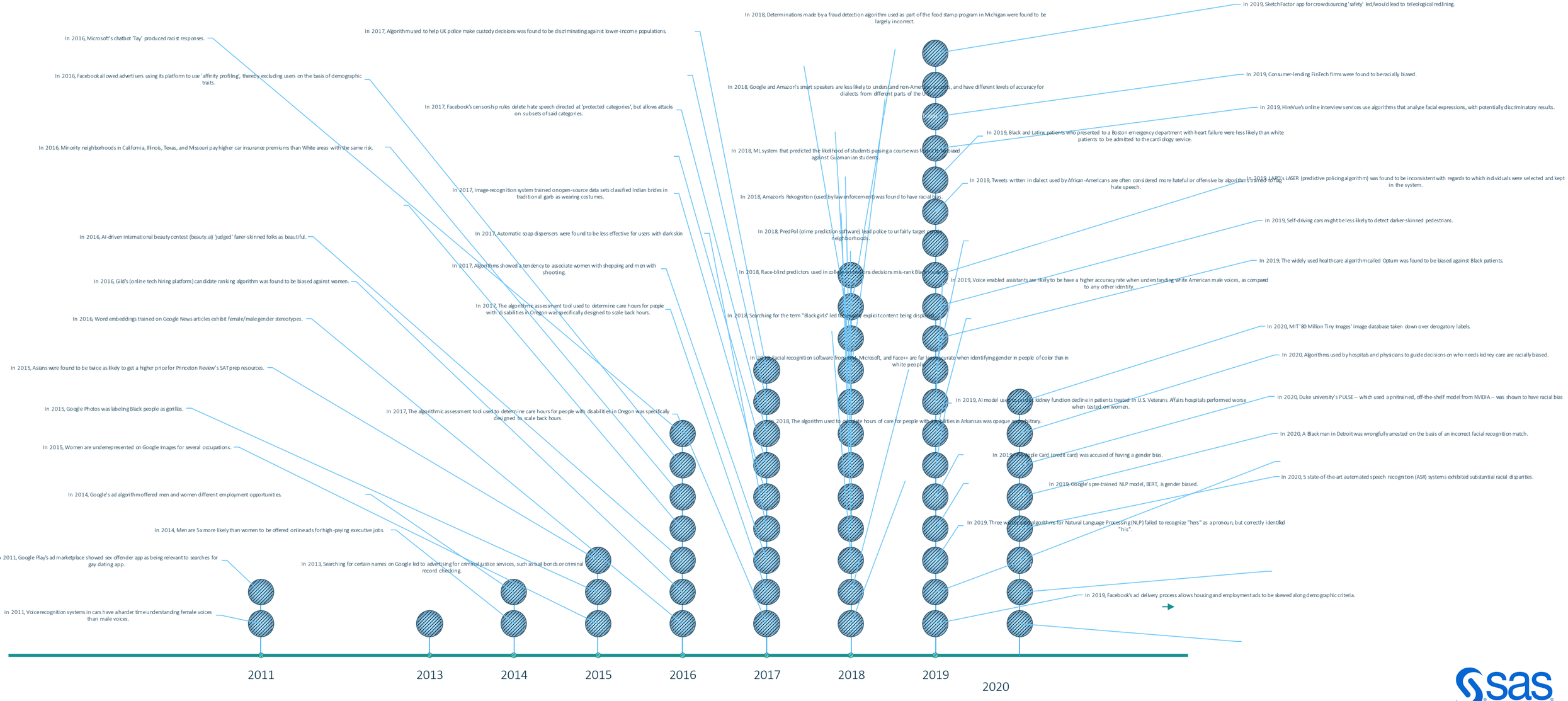
Artificial Intelligence Harm in the News



Unintended Harm…

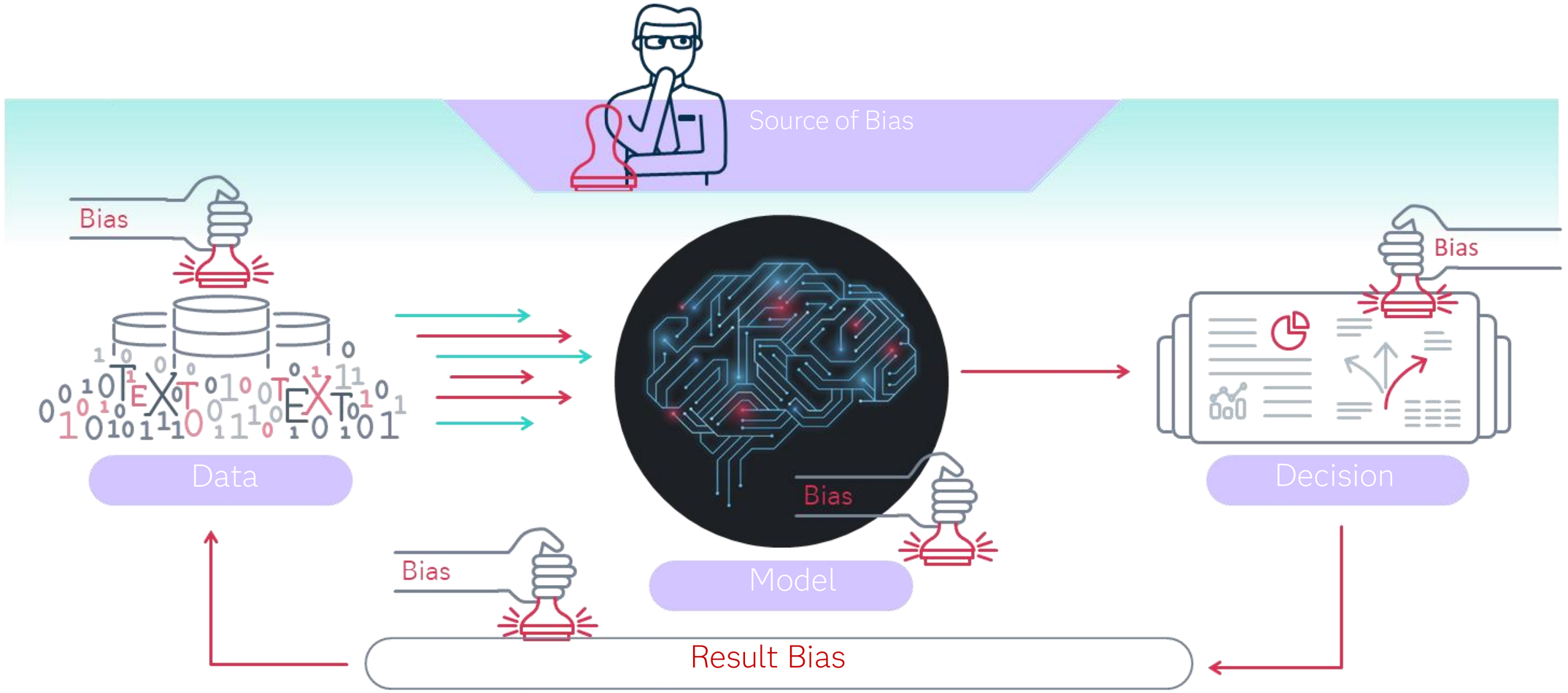
BERKELEY HAAS CENTER FOR EQUITY, GENDER & LEADERSHIP Examples of Bias in Artificial Intelligence

...Inceases Dramatically



Bias Risks

Data Bias | Model Bias | Decision Bias



Activation

Oversight
AI Governance, Strategy and Enforcement

Compliance
Performance and Risk Management

Operations
Operating Procedures with Infrastructure

Culture
Systems, Norms and Practices

Technology

Data Management
- Data Quality
- Variable Metadata
- Data Preparation
- Data Asset Catalog

Explanation
- Natural Language Explanation
- Explainable ML
- Counterfactual Explanation
- Surrogate Model Interpretation
- Causal Inference

Detection
- Bias Detection
- Fairness Assessment

Privacy & Security
- Privacy Preservation
- Model Security
- Autonomy Preservation
- Consent & Control

Mitigation
- Bias Mitigation
- Bias Prevention
- Synthetic Data Generation

ModelOps
- Model Cards
- Decisioning
- Lifecycle Management
- Metric Monitoring
- Model Robustness

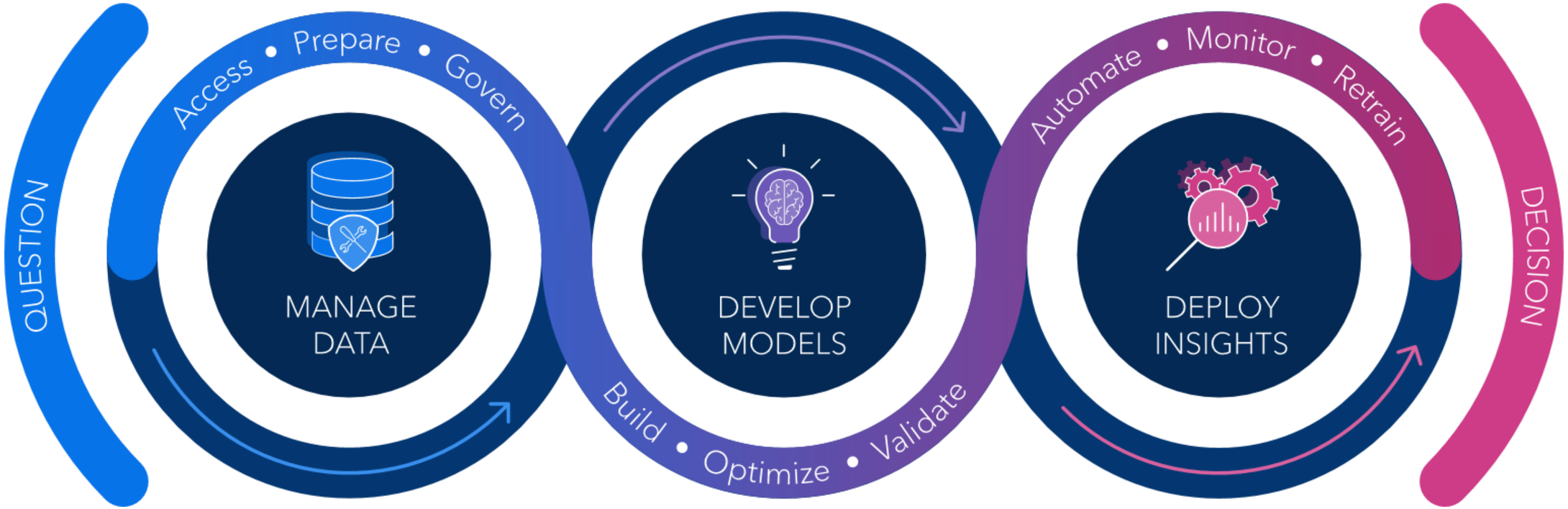
Solutions (industry & domain)

Accessible and Intuitive UI

Logging & Auditability

Interoperability

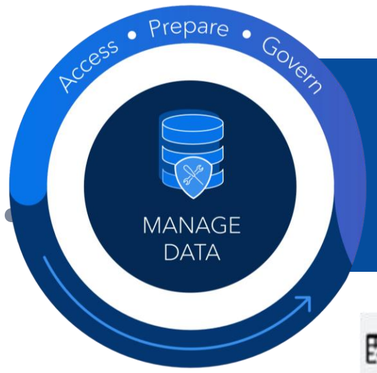
ANALYTICS LIFECYCLE



ROBUST DATA GOVERNANCE

TRANSPARENT AI

SECURE MODEL OPS PROCESSES



MANAGE DATA

DATA QUALITY

WATER_CLUSTER
Samples

Columns	Rows	Size
21	46.7 K	9.6 MB

Completeness: 95%

Status: Approved

Actions

Overview | Column Analysis | Sample Data

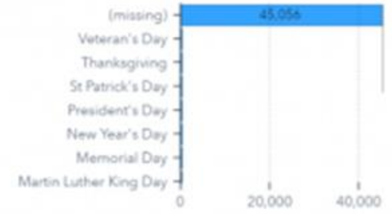
Date analyzed: 27 Jan 2021 4:25 pm

Descriptive Measures | Metadata Measures | Data Quality Measures

#	Name	Completeness	Uniqueness	Most Common Value	Least Common Value	Pattern Count	Semantic
9	Zip	100%	0%	77094	77020 (2 more)	--	POSTAL C
10	Lat	100%	0%	-95.408872 (1 more)	-95.71878 (19+ more)	--	LATITUDE
11	Long	100%	0%	29.785747 (1 more)	29.651167 (19+ more)	--	LONGITU
12	Property_type	100%	0%	0	0	--	GENERIC
13	Meter_Location	100%	0%	external	internal	1	
14	Cli	100%	0%	HSTNTXJA	HSTNTXHO	1	
15	DMA	100%	0%	2	1	--	
16	Weekday	100%	0%	4 (1 more)	1	--	
17	Weekend	100%	0%	0	1	--	
18	Daily_W_C_M3	100%	11%	0	0.023 (19+ more)	--	
19	Week	100%	0%	51 (19+ more)	52	--	DATE
20	US Holiday	4%	0%	Veteran's Day (11 m...	Independence Day	13	
21	CLUSTER	100%	0%	4	1	--	

Column Properties

Frequency distribution:



Column name:	US Holiday
Label:	--
Type:	char
Actual type:	String
Format:	$CHAR
Length:	27
Minimum length:	8
Maximum length:	27
Primary key candidate:	No
Logical type:	Nominal
Semantic type:	--
Information privacy:	--

MANAGE DATA

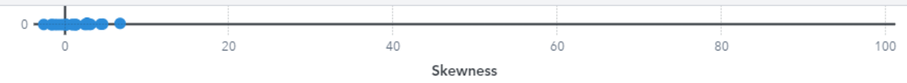
DATA EXPLORATION

TS_NC > "Data Exploration" Results

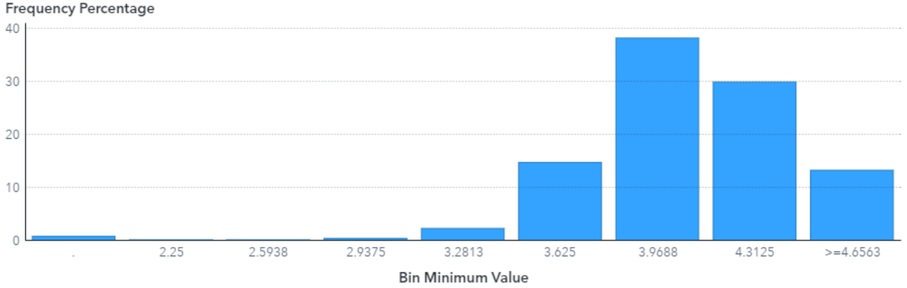
Close

Summary Output Data

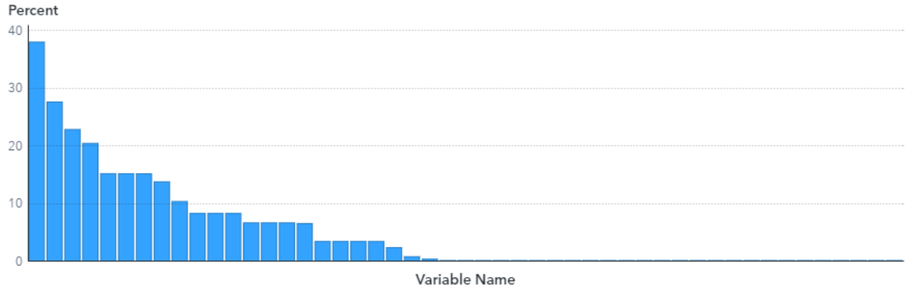
CLASS_SIZE	7	4,630	327.2290	148.1479	2.7501	76.2903	0.4527	623.5:



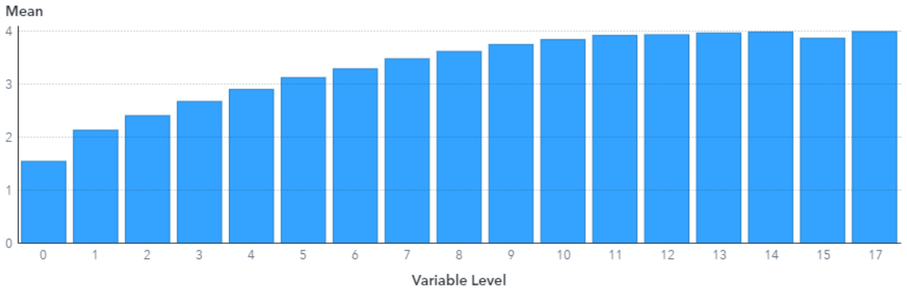
Interval Variable Distributions HS_GPA



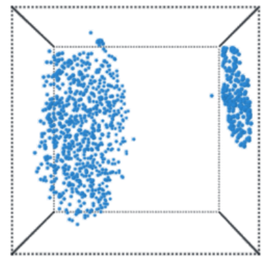
Missing Values



Target Variable Means A



t-SNE Projection of Interval Inputs



MANAGE DATA

INFORMATION PRIVACY

SAS® Information Catalog - Discover Information Assets

Catalog Home › Search Results › AUSCUST

AUSCUST
COMP_GEL

Columns 16 | Rows 5 | Size 128 KB

Completeness: 100%

Overview | Column Analysis | Sample Data

Content Overview

Information Privacy: Sensitive

Time Period Covered: (none found)

Area Covered: SE, VIC, VA, NSW, MELBOURNE

Analyzed using discovery locale:
Australia (English)

hat it tracks or measures.

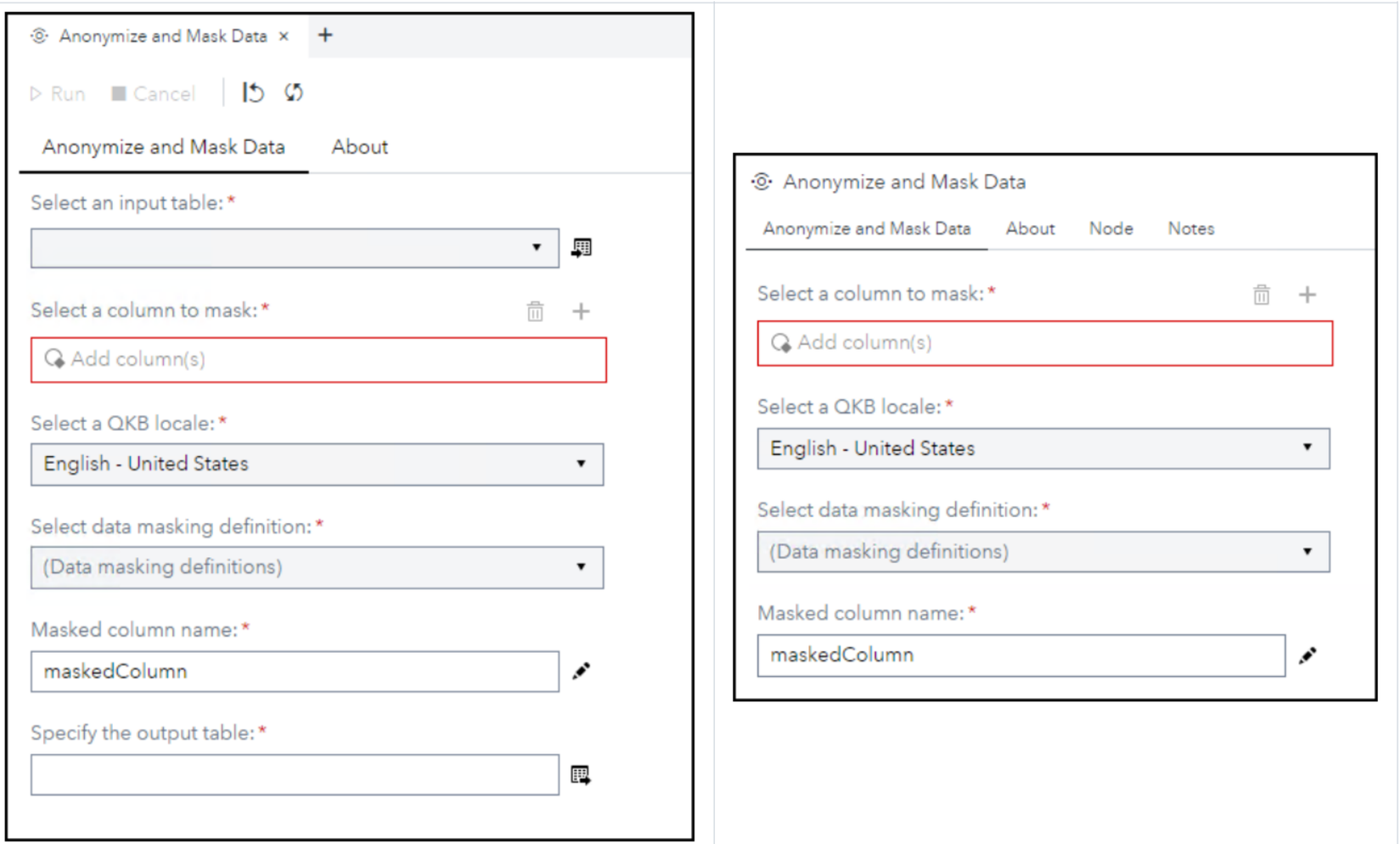
Sensitive:
- MARITAL_STATUS

Private:
- INDIVIDUAL
- AGE
- BIRTH DATE
- GENDER
- E-MAIL
- PHONE

Candidate:
- ADDRESS
- CITY
- POSTAL CODE
- STATE/PROVINCE

MANAGE DATA

DATA MASKING



MANAGE DATA

DATA SUPPRESSION

SAS® Visual Analytics - Explore and Visualize

Report 1

Create Calculation

Name:

Count

Type:

Data suppression

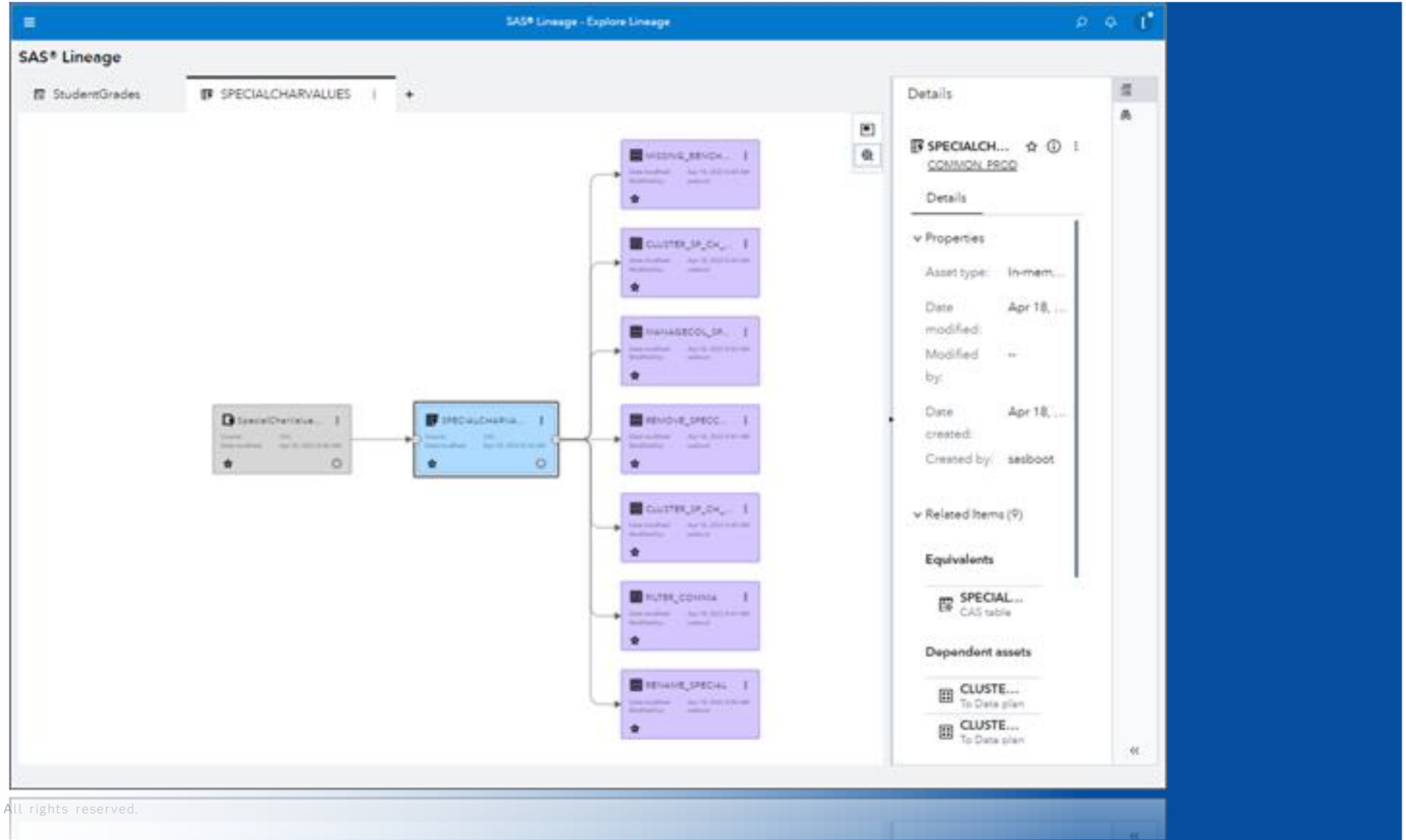
Suppress data if count less than:

5

OK Cancel

MANAGE DATA

DATA LINEAGE



DEVELOP MODELS

NATURAL LANGUAGE INSIGHTS

What are the characteristics of Customer Satisfaction?

Customer Satisfaction ranges from 1 to 5. Average Customer Satisfaction is 3.9. Most cases (951 of 1.2K) have a Customer Satisfaction between 2 and 5. Warehouse (Ship To) best differentiates the highest (top 10%) and the lowest (bottom 10%) Customer Satisfaction cases. The three most related factors are Warehouse (Ship To), Customer State, and Customer Loan-to-Value (LTV).

What factors are most related to Customer Satisfaction?

What are the groups based on Customer State by the average value of Customer Satisfaction?

< High | Low >

2.3	If Customer State is GA, KS, MD, NC, or TX, Social Media Score is 3, then the 11 cases have a predicted Customer Satisfaction of 2.3.
2.4	If Online Reviews is 1 or 2, Customer State is KS, then the 11 cases have a predicted Customer Satisfaction of 2.4.
2.5	If Online Reviews is 1 or 2, Customer State is NC, then the 14 cases have a predicted Customer Satisfaction of 2.5.

What is the relationship between Customer Satisfaction and Customer State?

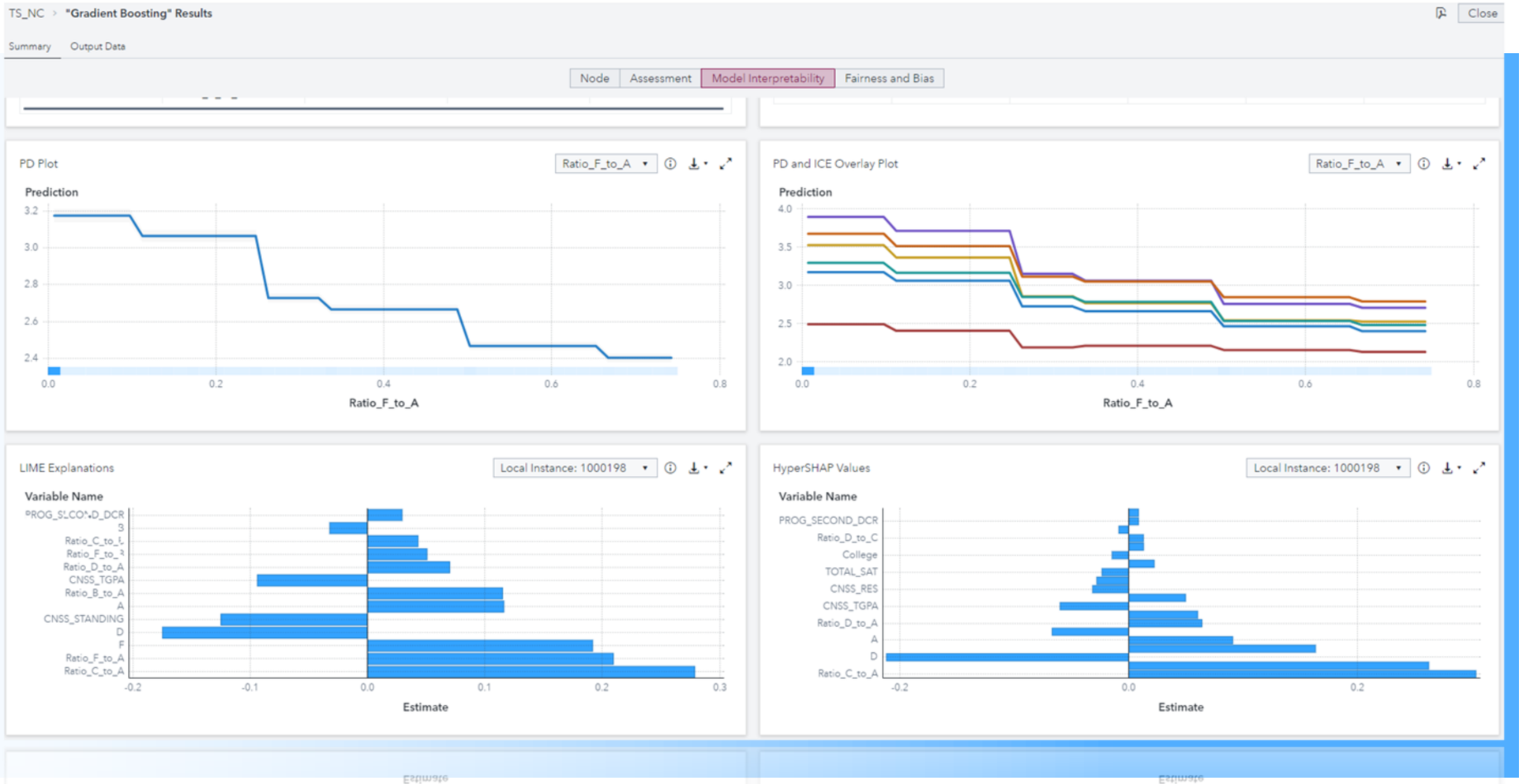
When Customer State is UT, CO, NJ, OR, NY, WA, CA, AZ or IL, the average of Customer Satisfaction is a high value. When Customer State is TX, KS, GA, NC or MD, the average of Customer Satisfaction is a low value. The most common Customer State value is NC.

Explanation Description | Screening Results | Relative Importance | Anomalies

1. Select response for Automated Explanation	A report author selected Customer Satisfaction as the response.
2. Screen factors	Automated Explanation modified or removed 6 of 12 factors. See the Screening Results tab for details.
3. Determine most related factors	Automated Explanation used a one-level decision tree for each factor to determine its relative importance to Customer Satisfaction. For example, the input Customer State has a relative importance of 0.75 which means it is 0.75 times as important as Warehouse (Ship To).
4. Find groups based on selected related factor	Automated Explanation ran 8 decision trees with response Customer Satisfaction. The trees used Customer State and another important factor as predictors. The trees had 6 levels and 2 branches. Each group describes a leaf from one of these trees.

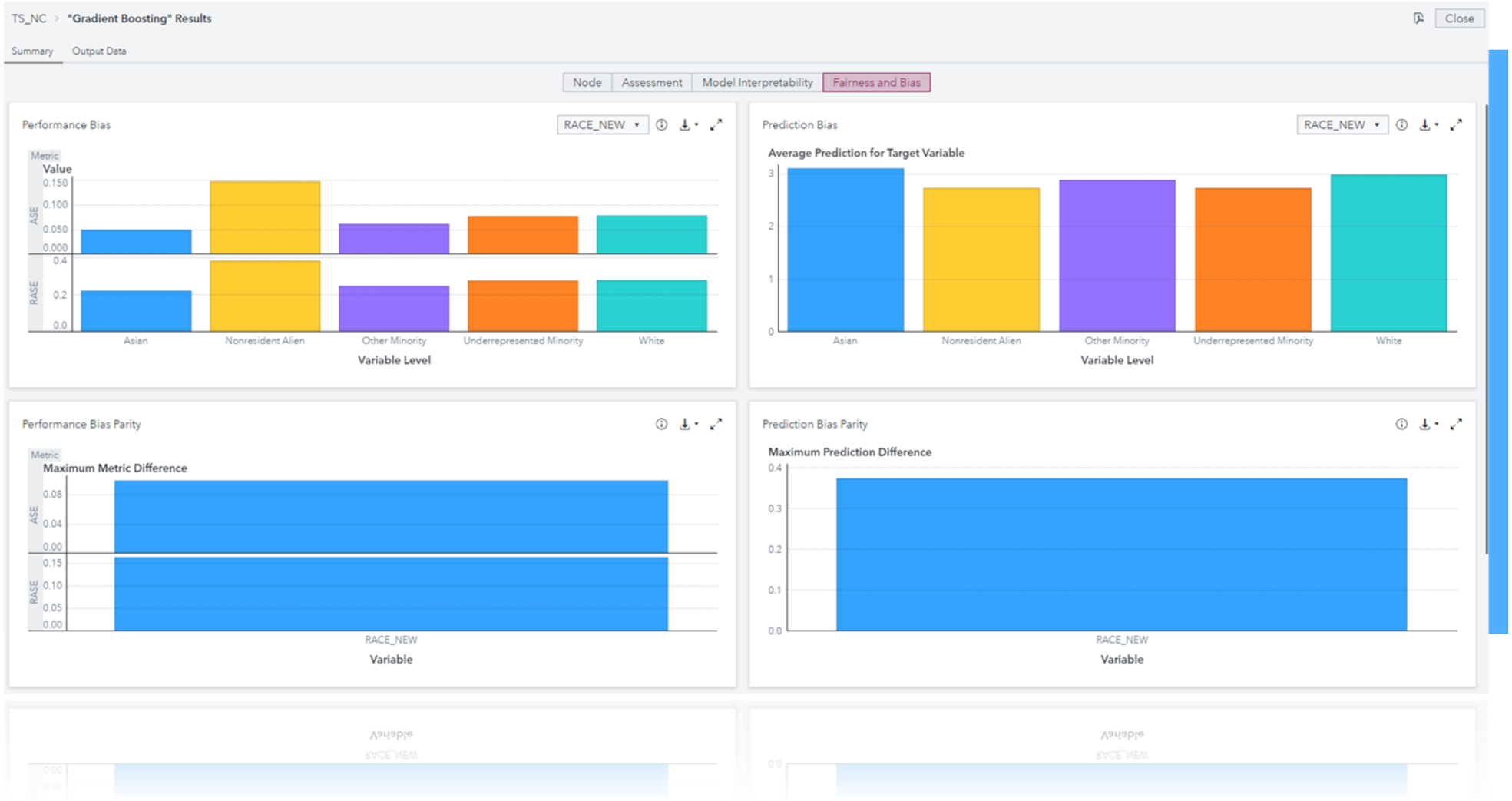
DEVELOP MODELS

MODEL INTERPRETABILITY



DEVELOP MODELS

FAIRNESS ASSESSMENT & BIAS MITIGATION



DEPLOY INSIGHTS

MODEL GOVERANCE

SAS® Model Manager - Manage Models

Home
Models
Projects
Deployments
Tasks

Projects

Hide tiles

388

Total Number of Projects

266
122

122 Projects with Published Models

Projects per Model Function

388 Total

Analytical
Classification
Clustering
Computer vision - Image classification
Computer vision - Object detection
Not specified
Prediction
Single-series forecasting
Text analytics

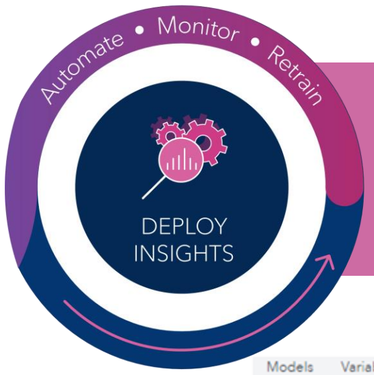
Published Models per Destination

249 Total

SAS-Cloud-Analytic-Service
maslocal

Search name | Advanced search | New Project

Name	Published Status	Operational Status	Model Function	Date Created	Date Modified	Latest Version
TS_deeplearning	●	Prototype		Feb 19, 2021 11:04 AM	Jan 28, 2022 07:58 PM	Version 1
ESP_astore	●	Prototype	Classification	Feb 23, 2021 11:54 AM	Jan 28, 2022 07:58 PM	Version 1
GELTEST_LogisticRegression		Prototype	Classification	Mar 10, 2021 01:24 PM	Jan 28, 2022 07:58 PM	Version 4
Dave		Prototype	Classification	Mar 10, 2021 02:02 PM	Jan 28, 2022 07:58 PM	Version 1
Credit Card Default	●	Prototype	Classification	Mar 15, 2021 04:29 AM	Jan 28, 2022 07:58 PM	Version 1
CV_image_classificatio		Prototype	Computer vision - Image classification	Mar 18, 2021 10:07 AM	Jan 28, 2022 07:58 PM	Version 1
sasjst_Model		Prototype	Prediction	Apr 19, 2021 09:51 PM	Jan 28, 2022 07:58 PM	Version 1
Fashion Influencer Prediction		Prototype	Classification	Apr 20, 2021 02:51 PM	May 13, 2022 04:34 PM	Version 1
RETAIL_Customers_Influencer		Prototype	Classification	Apr 23, 2021 03:38 PM	Jan 28, 2022 07:58 PM	Version 1
Identify Influencers		Prototype	Classification	Apr 29, 2021 03:51 PM	Jan 28, 2022 07:58 PM	Version 1
Predicting Diabetes (ICU)		Prototype	Classification	May 3, 2021 09:53 PM	Jan 28, 2022 07:58 PM	Version 1
Drug Report Anxiety and Depression - Pip…		Prototype	Text analytics	May 4, 2021 02:06 PM	Jan 28, 2022 07:58 PM	Version 1
Drug Report Anxiety and Depression - Pip…	●	Prototype	Text analytics	May 5, 2021 01:34 PM	Jan 28, 2022 07:58 PM	Version 1
open source		Prototype	Classification	May 5, 2021 04:49 PM	Jan 28, 2022 07:58 PM	Version 1
Identify Influencers_1		Prototype	Classification	May 6, 2021 08:27 AM	Jan 28, 2022 07:58 PM	Version 1
HMEQ_snlidl	●	Prototype	Classification	May 6, 2021 10:02 AM	Jan 28, 2022 07:58 PM	Version 1
HMEQ_OS		Prototype	Classification	May 7, 2021 09:18 AM	Jan 28, 2022 07:58 PM	Version 1
python		Prototype	Classification	May 7, 2021 05:02 PM	Jan 28, 2022 07:58 PM	Version 1



DEPLOY INSIGHTS

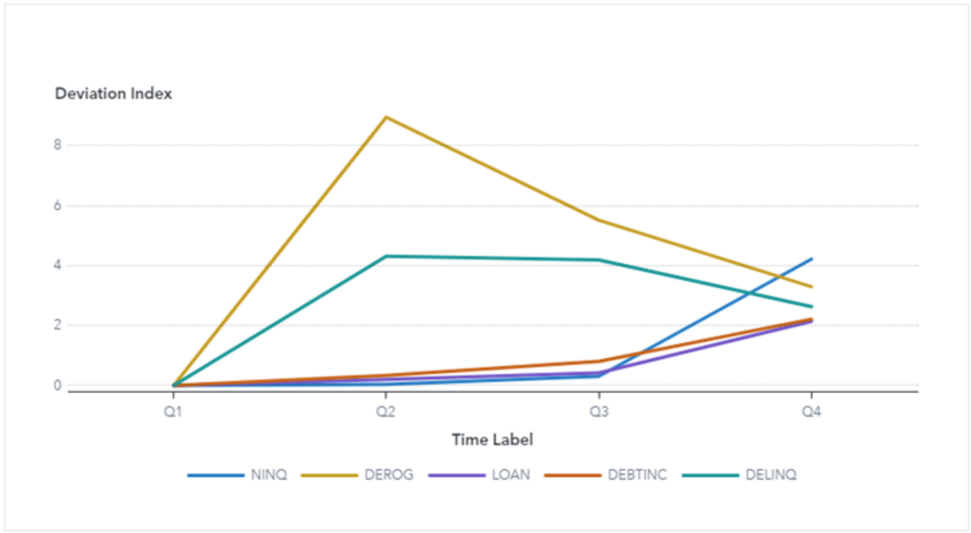
MODEL MONITORING

Models Variables Properties Files Scoring Performance Workflow History

hmeq_python

Run ▾ Edit

Summary Python_xgboost (3.0) SAS Random Forest (Open-source pipeline) (1.0)

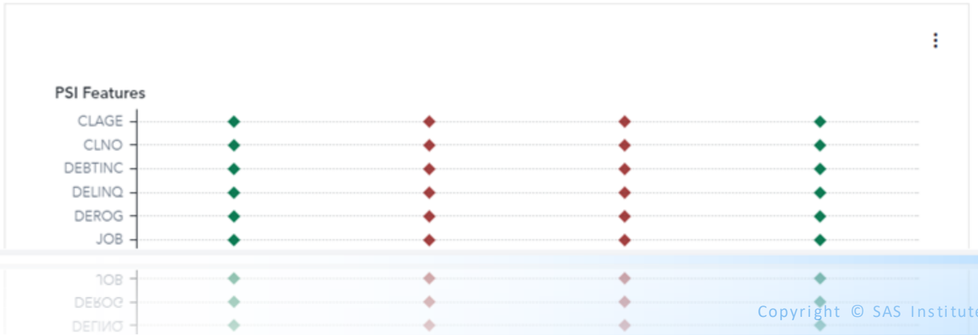


Input Variable Characteristic

This chart shows the shifts in the distribution of variable values in the input data over time. These shifts can point to significant changes in customer behavior that are due to new technology, competition, marketing promotions, new laws, or other influences.

To find shifts, the distributions of the variables in the training data are compared to the current data.

If large enough shifts occur in the distribution of variable values over time, the original model might not be the best predictive or classification tool to use with the current data.



PSI Out-of-Bounds Indicators for Input Variables

This chart shows the out-of-bounds indicators of the population stability index (PSI) for input variables.

Input variables that are marked as out-of-bounds have substantial shifts in their distribution. You should scrutinize input variables that exhibit an increasing trend of being out-of-bounds.

DEPLOY INSIGHTS

DECISION ACCOUNTABILITY

Loan Decision (1.0)

Publish

Decision Flow | Decision | Properties | Variables | Scoring | Versions | History

Search name

All types

Start

Home_Equity_Variable_Init...

Initial_Loan_Check_Simplif...

DeclineFlag = True

Yes

No

Deny Loan (1.0)

Credit_Class = 'PRIME'

Yes

No

Gradient Boosting (Adv Te...

Forest (Adv Template)

P_BAD1 > 0.85

P_BAD1 > .5

Yes

No

Yes

No

Deny Loan (1.0)

Approve Loan (1.0)

Deny Loan (1.0)

Approve Moderate Interest...

End

“Our $1 billion investment in industry solutions includes the integration of trustworthy generative AI capabilities that are accurate, explainable and defendable.”

Brian Harris - SAS Executive Vice President and Chief Technology Officer

SAS & Generative AI

"SAS and Microsoft are jointly developing a generative AI integration that combines the scale of Microsoft Azure OpenAI with SAS' orchestration of enterprise tasks and existing analytics used by enterprises to make operational decisions. The generative AI integration will be available in a private preview in Q4 2023."

SAS Explore 2023

Live Webinar:

SAS' Generative AI Strategy for Enterprise Transformation

Oct. 12 • 10 a.m. ET | 4 p.m. CET • Cost: Complimentary

https://www.sas.com/en_us/webinars/generative-ai-strategy.html

Vielen Dank

Anette.Almer@sas.com
Tamara.Fischer@sas.com