PAGE 44

Data Viz and Machine Learning on SAS Viya

Created an automatic chart showing the Product Cost of Sale by Facility Continent

Product Cost of Sale by Facility Continent

Facility Continent

Product Cost of Sale (millions)

1
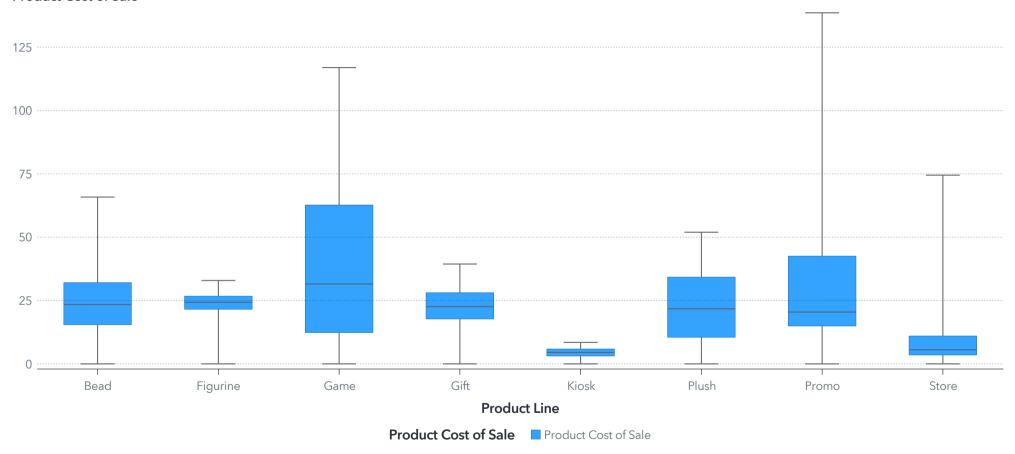
box plot showing the Product Cost of Sale by Product Line

Product Cost of Sale by Product Line

Product Cost of Sale

Product Cost of Sale ■ Product Cost of Sale

Unit Capacity by Transaction Month
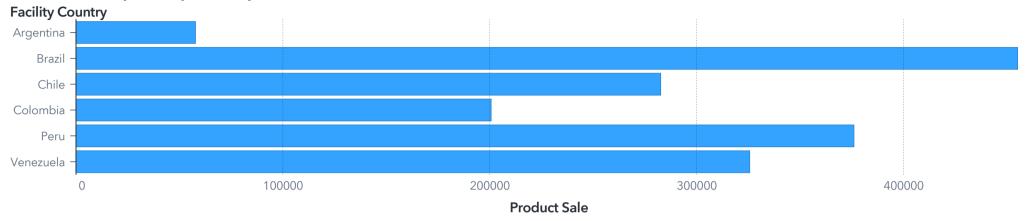
Product Sale by Facility Country

Facility Country

France	
Norway	Mexico
Brazil	
Sweden	Italy
United States	
United Kingdom	Germany
Spain	Canada

Chile
Denmark
Colombia
Venezuela
Peru

23M
17K

Product Sale

product sales for each country in South America using three types of charts

Product Sale by Facility Country

Facility Country

Product Sale

Product Sale by Facility Country

Product Sale

1.7M

Facility Country

■ Brazil ■ Peru ■ Venezuela ■ Chile ■ Colombia ■ Argentina

5
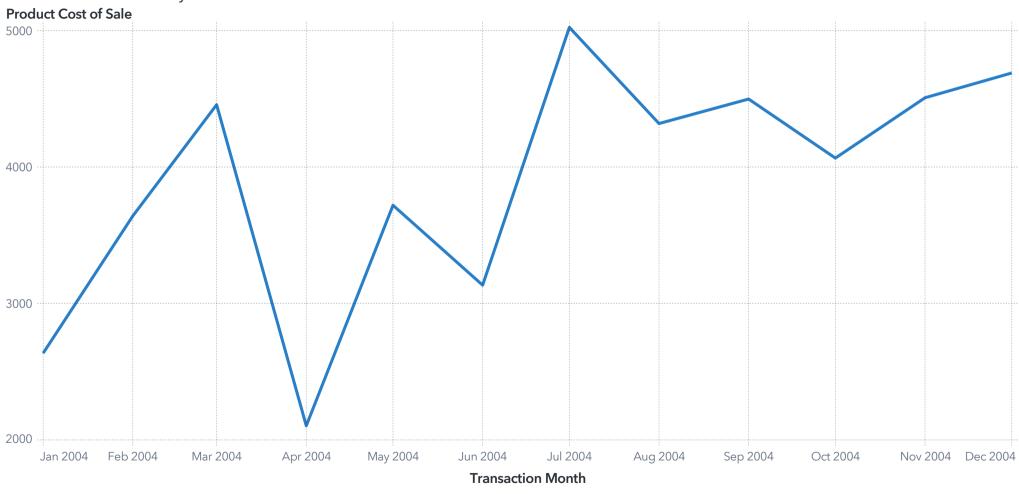
two months in 2004 when the cost of sale was lowest in Argentina

Product Cost of Sale by Transaction Month

6

Global Performance Overview

Product Cost of Sale by Transaction Month

Facility Efficiency

Product Sale by Facility Continent

Global Performance Overview

Insights:

Product Sales by Facility Continent: This will show which continents are driving the most product sales, allowing you to identify top-performing regions.
Key Insight: North America has the highest sales, while South America lags behind.
Product Cost of Sale Trend by Month: This will help in understanding cost fluctuations over time.

Regional Performance Overview

Product Cost of Sale by Facility Region

Facility Age

Product Material Cost by Facility Region

Regional Performance Overview

Insights:

Product Sales by Facility Region: A treemap to illustrate sales distribution across various regions.

Key Insight: The North-East region outperforms other regions, accounting for 25% of total sales.

Product Material Cost by Region: Box plot to analyze material costs across regions.

Financial Data Overview

Product Cost of Sale by Product Line

Product Sale by Product Make

Financial Data Overview

Insights:

Product Cost of Sale by Product Line: Bar chart to compare costs across different product lines.

Key Insight: Product Line A has the highest cost, indicating potential inefficiencies or higher quality materials.

Detailed Financial Metrics: Table to provide a detailed view of financial metrics like cost, price, and sales.

Marketing Data

Customer Satisfaction by Sales Rep

Sales Rep Customer Base by Facility Continent, Frequency

Marketing Data Overview

Insights:

Customer Satisfaction by Sales Rep: Bar chart to show satisfaction levels per sales rep.
Key Insight: Sales Rep John Doe has the highest satisfaction score.
Sales Rep Customer Base by Facility Continent: Heat map to illustrate the customer base distribution.

Summary

Detailed Points and Insights for Each Chart:

Global Performance Overview
Bar Chart: Breakdown of sales by continent shows North America leading with 40% of total sales.
Line Chart: Monthly cost trends indicate peaks in July and December, suggesting seasonal impacts.
Gauge: Global efficiency at 85%, short of the 90% target, indicating room for improvement.

Regional Performance Overview
Treemap: Sales heavily concentrated in the North-East region.
Box Plot: The South-East region benefits from lower material costs.
Gauge: The average age of facilities is 12 years, with Midwest facilities averaging 15 years.

Financial Data Overview
Bar Chart: Product Line A's higher costs might require a review of its production process.
Table: Detailed metrics show Product X's high cost is justified by its leading sales figures.
Pie Chart: Dominance of Product Make 1 suggests focusing marketing efforts here could be beneficial.

Marketing Data Overview
Bar Chart: Customer satisfaction high for John Doe, indicating successful sales strategies.
Heat Map: North America's strong customer base suggests it's a key market.
Scatter Plot: Western region's high sales rep ratings highlight successful regional management.
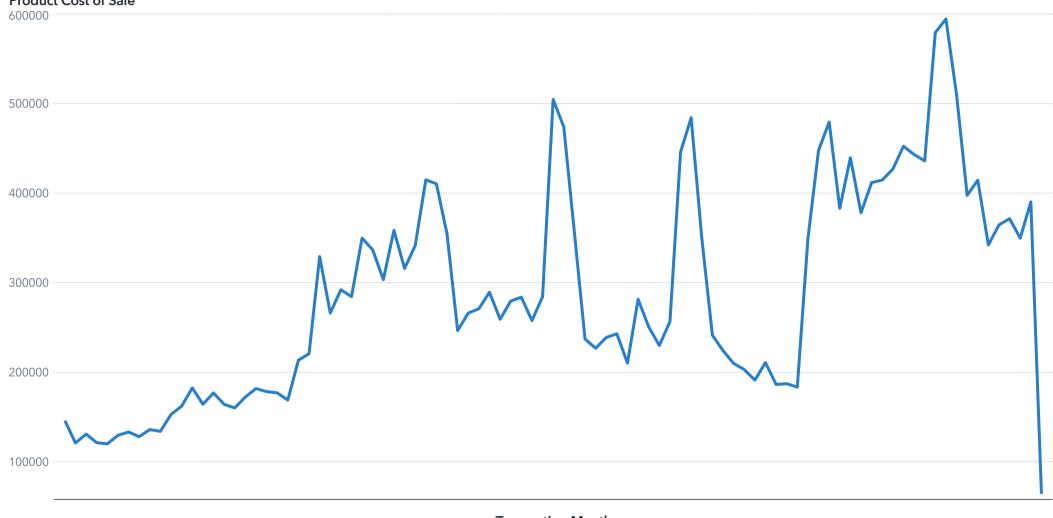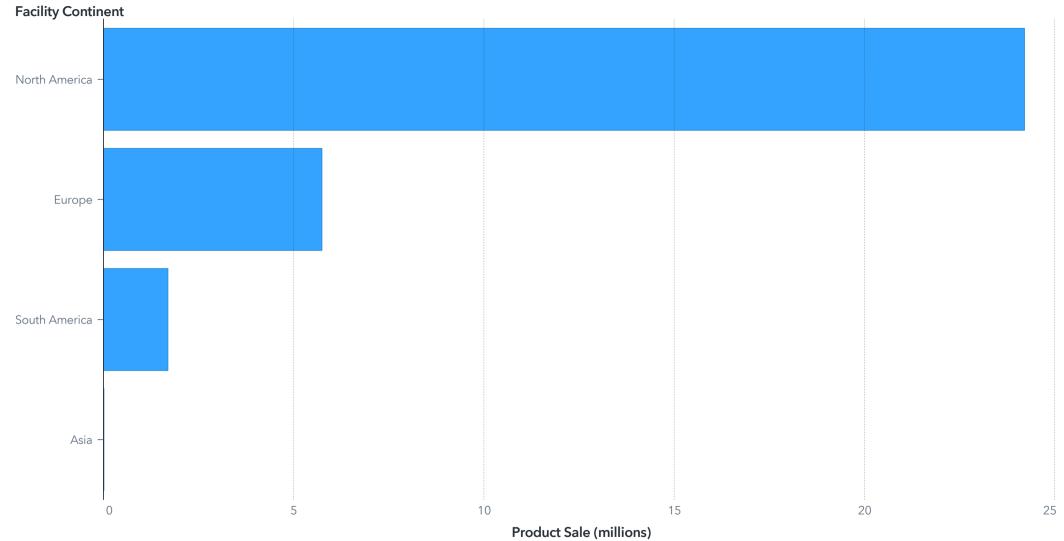
Product Cost of Sale by Transaction Month

Product Cost of Sale

Transaction Month

Product Sale by Facility Continent

Facility Continent

Product Sale (millions)

13

Global Performance Overview

Insights:

Product Sales by Facility Continent: This will show which continents are driving the most product sales, allowing you to identify top-performing regions.

Key Insight: North America has the highest sales, while South America lags behind.
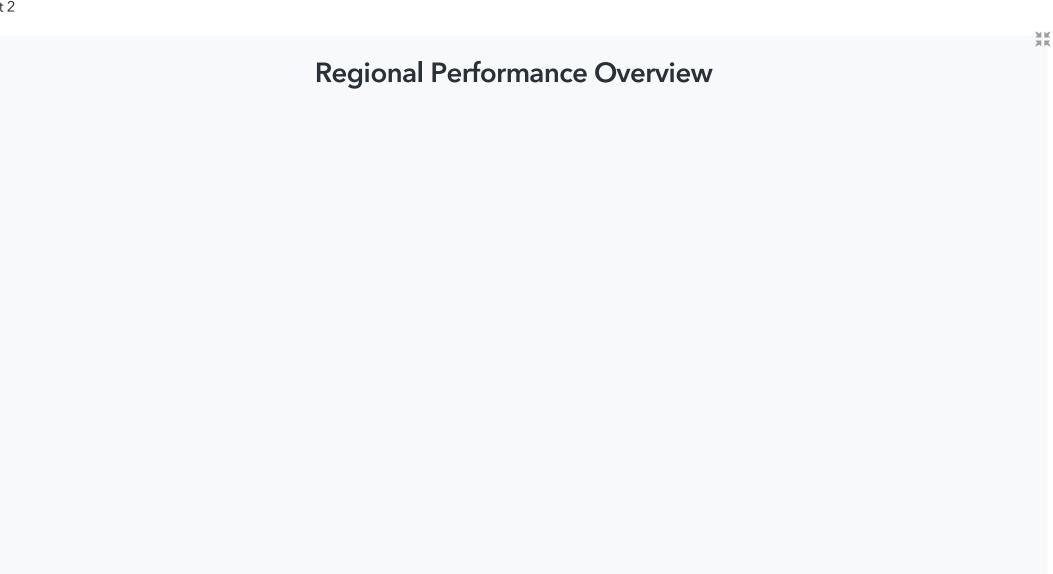
Product Cost of Sale Trend by Month: This will help in understanding cost fluctuations over time.

Key Insight: Costs peak during certain months, possibly due to seasonal demand or supply chain issues.

Global Facility Efficiency: Gauge to show overall efficiency of facilities worldwide.

Key Insight: Average facility efficiency is 85%, with a target of 90%.

Regional Performance Overview

Product Cost of Sale by Facility Region

Facility Region
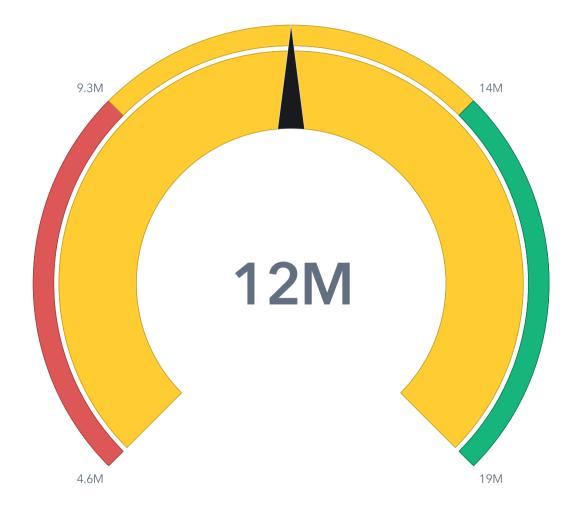
- Nebraska
- New York
- Washington
- Colorado
- Arizona
- Lazio
- Florida
- Oslo
- Montana
- Indiana
- Rio de Janeiro
- Tennessee
- Minas Gerais
- Manchester
- City of London
- South Dakota
- Wyoming
- Utah
- Iowa
- Sao Paulo
- Virginia
- Georgia
- Baja California
- Jalisco
- Alabama
- Arkansas
- Missouri
- Nevada
- New Mexico
- Comunitat Valenciana
- Michigan
- Louisiana
- Bogota, D.C.
- Distrito Federal
- Mississippi
- Kentucky
- Pennsylvania
- Minnesota
- British Columbia
- Madrid (Comunidad de)
- Cataluna
- District of Columbia
- Maryland
- West Virginia
- Oklahoma
- South Carolina
- Illinois
- California
- Oregon
- Kansas
- Wisconsin
- Södermanland
- Ohio
- North Dakota
- Lima
- North Carolina
- Idaho
- Ontario
- Berlin
- Hamburg
- Hovedstaden
- Santiago
- Ile-de-France
- Distrito Capital
- New Jersey
- Texas

1.8M
14K

Product Cost of Sale

16

Facility Age

9.3M

14M

12M

4.6M

19M

Product Material Cost by Facility Region

Product Material Cost

Facility Region
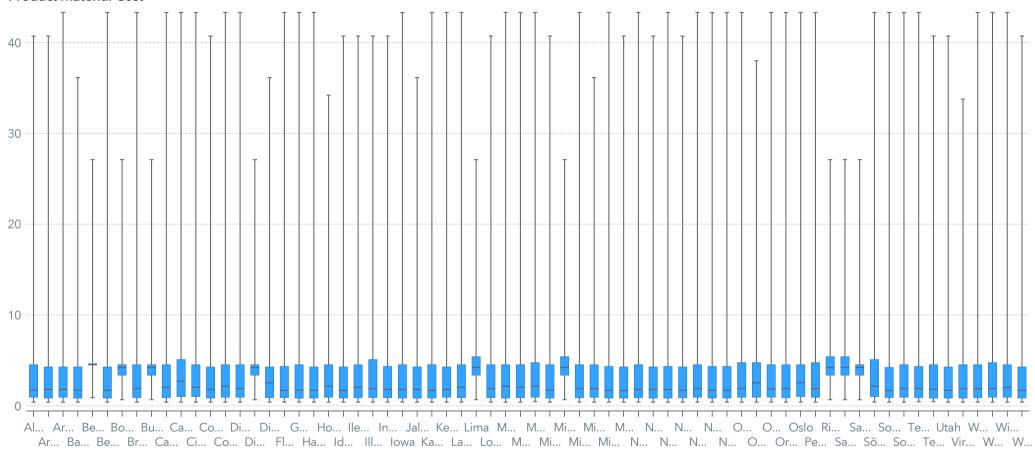
Product Material Cost ■ Product Material Cost
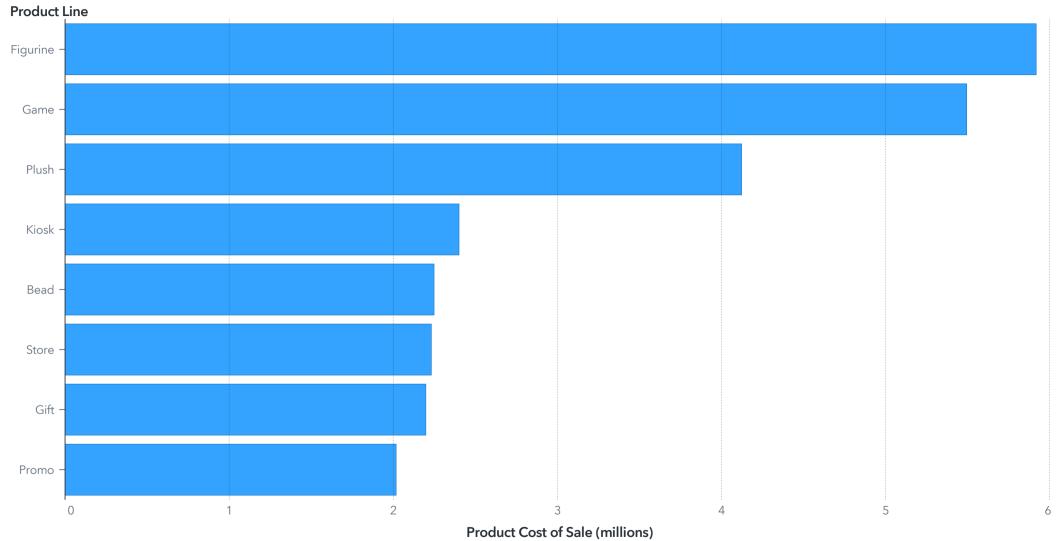
18

Regional Performance Overview

Insights:

Product Sales by Facility Region: A treemap to illustrate sales distribution across various regions.
Key Insight: The North-East region outperforms other regions, accounting for 25% of total sales.
Product Material Cost by Region: Box plot to analyze material costs across regions.
Key Insight: Material costs vary significantly, with the South-East region having the lowest median cost.
Average Facility Age by Region: Gauge to show the average age of facilities.
Key Insight: Older facilities in the Midwest may need upgrades.

Product Cost of Sale by Product Line

Product Line

Figurine

Game

Plush

Kiosk

Bead

Store

Gift

Promo

0   1   2   3   4   5   6

Product Cost of Sale (millions)

20
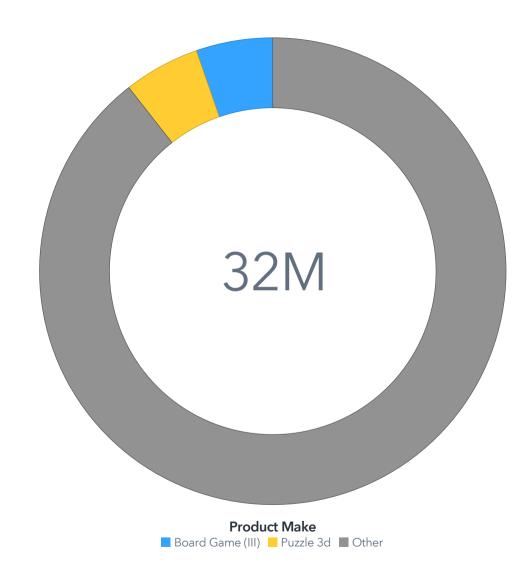
Product Sale by Product Make

Product Sale

Product Make
■ Board Game (III) ■ Puzzle 3d ■ Other

32M

List table - Product Cost of Sale 1

Product Cost of Sale ▲	Product Material Cost	Product Price (target)	Product Sale
26632640	5888709	35767243	31668376

Financial Data Overview
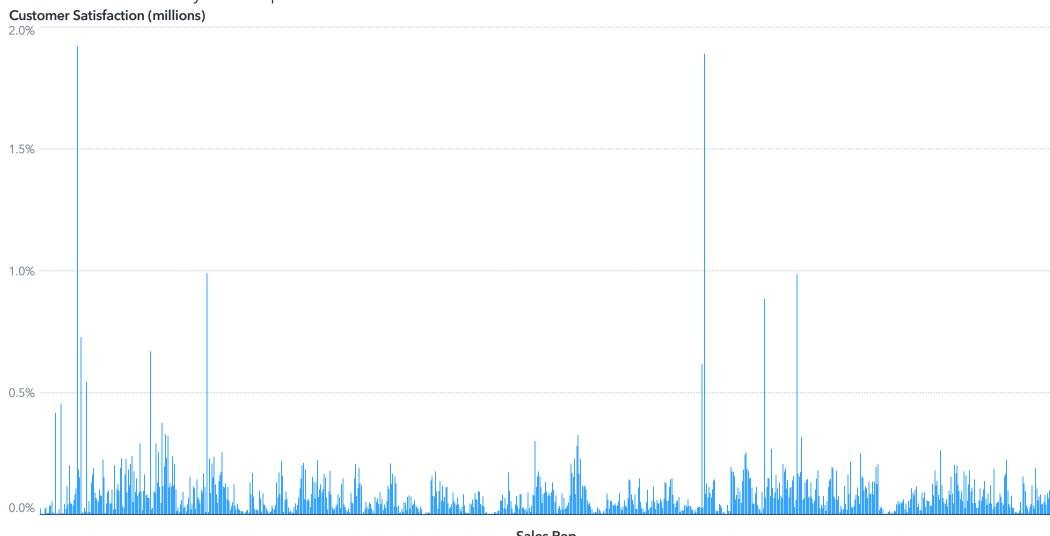
Insights:
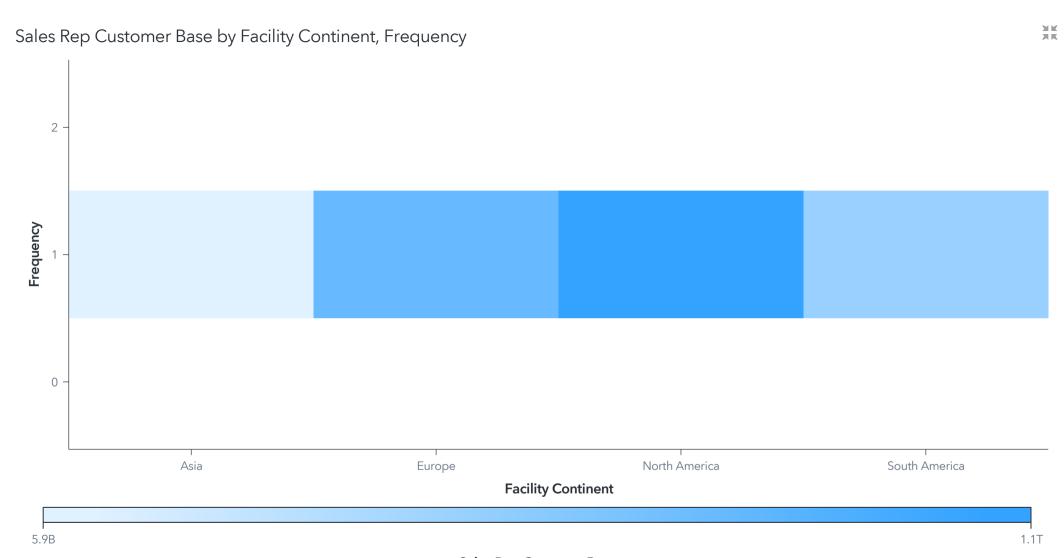
Product Cost of Sale by Product Line: Bar chart to compare costs across different product lines.
Key Insight: Product Line A has the highest cost, indicating potential inefficiencies or higher quality materials.
Detailed Financial Metrics: Table to provide a detailed view of financial metrics like cost, price, and sales.
Key Insight: Product X has a high cost but also generates the highest sales.
Product Sale Distribution by Product Make: Pie chart to show the proportion of sales by product make.
Key Insight: Product Make 1 dominates the market with 60% of sales.

Customer Satisfaction by Sales Rep

Customer Satisfaction (millions)

Sales Rep

Sales Rep Customer Base by Facility Continent, Frequency

25

Marketing Data Overview

Insights:

Customer Satisfaction by Sales Rep: Bar chart to show satisfaction levels per sales rep.
Key Insight: Sales Rep John Doe has the highest satisfaction score.
Sales Rep Customer Base by Facility Continent: Heat map to illustrate the customer base distribution.
Key Insight: The majority of customers are in North America, with the lowest in Australia.
Sales Rep Rating by Facility Region: Scatter plot to show ratings distribution.
Key Insight: High ratings are clustered in the Western region.
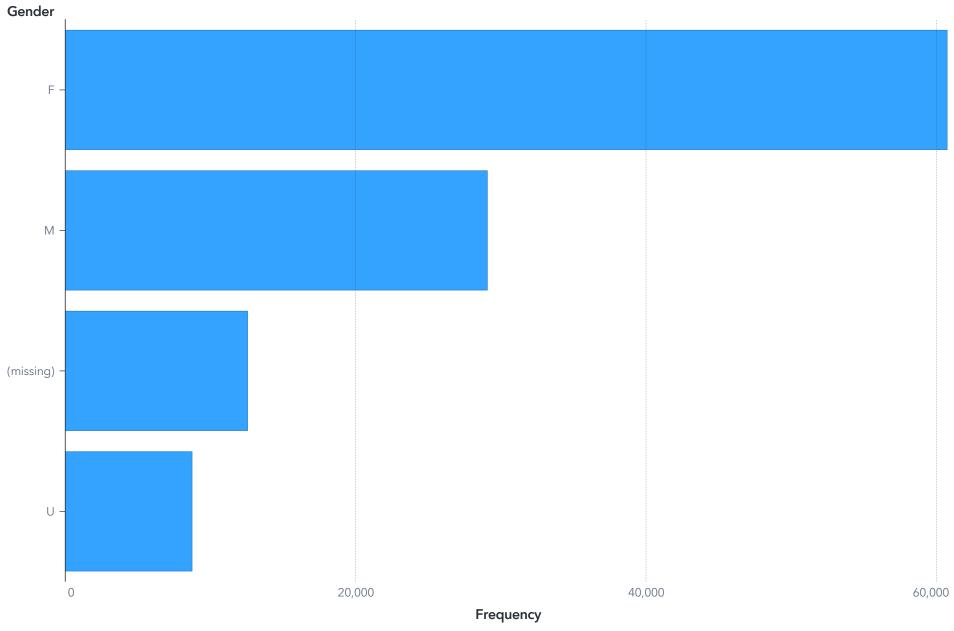
PAGE 72

Creation Date: Saturday, August 24, 2024, 01:34:15 PM
Author: mohammed-danish.mustafa@edu.dsti.institute

Frequency of Gender

Gender

Frequency
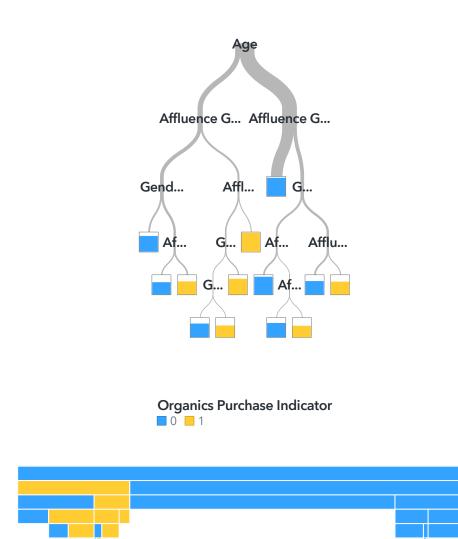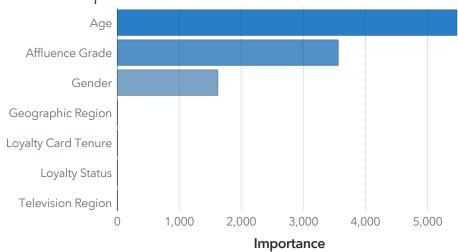
Decision Tree of Organics Purchase Indicator

Event: **1**   Fit: **KS (Youden) 0.4707**   Observations: **111K of 111K**

Tree

Organics Purchase Indicator
■ 0 ■ 1
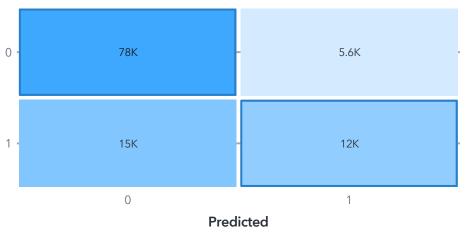
Variable Importance

Confusion Matrix

Observed

	0	1
0	78K	5.6K
1	15K	12K
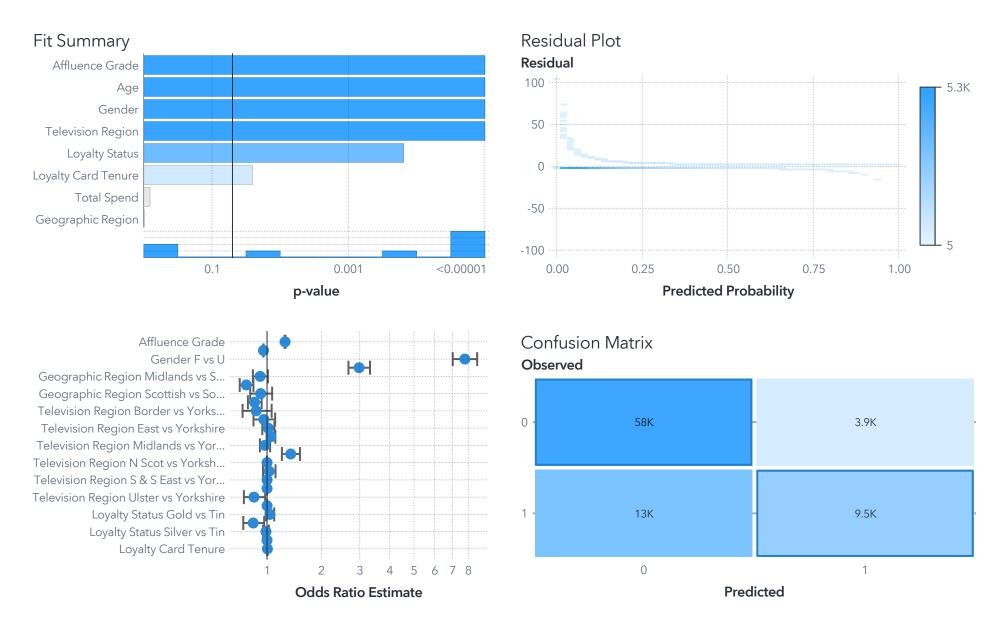
Predicted
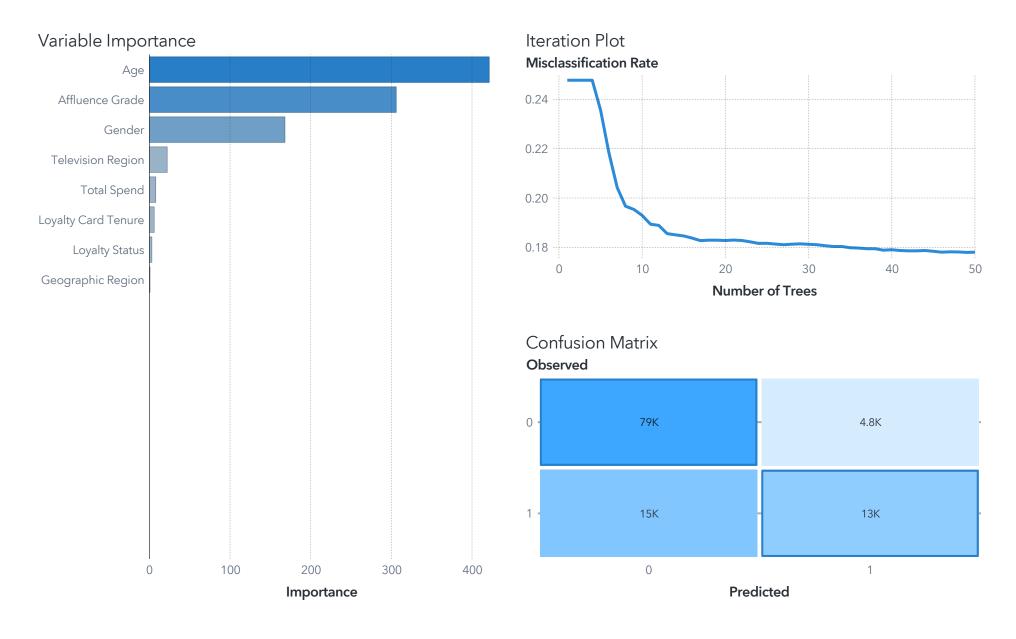
Logistic Regression of Organics Purchase Indicator

Event: **1**   Fit: **KS (Youden) 0.4561**   Observations: **85K of 111K**

Fit Summary

Residual Plot

Residual

Confusion Matrix

Observed
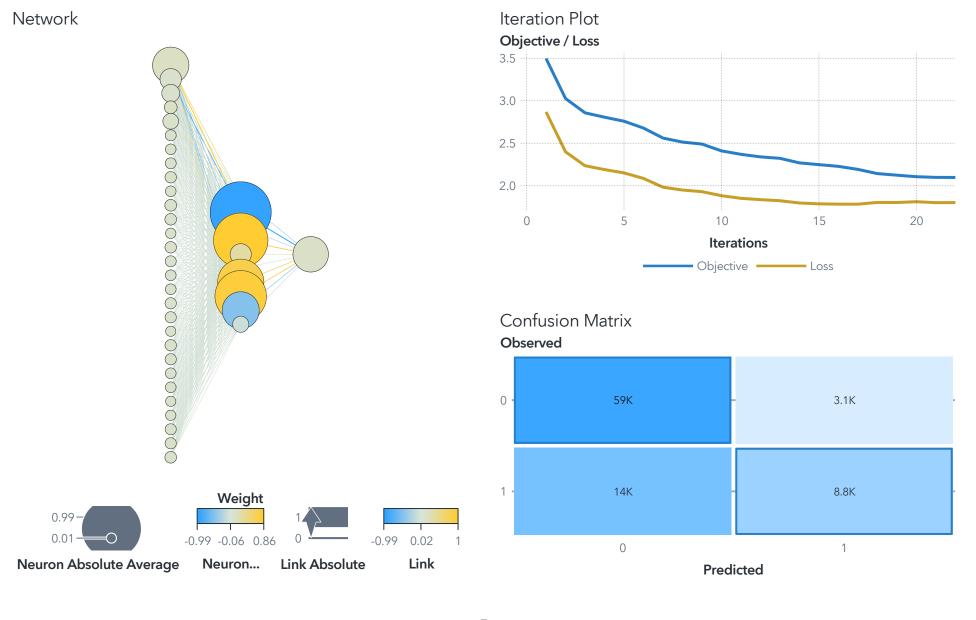
Gradient Boosting of Organics Purchase Indicator

Event: **1**   Fit: **KS (Youden) 0.5119**   Observations: **111K of 111K**

Variable Importance

Iteration Plot
Misclassification Rate

Confusion Matrix
Observed

Neural Network of Organics Purchase Indicator

Event: **1** Fit: **KS (Youden) 0.4571** Observations: **85K of 111K**

Network

Neuron Absolute Average

0.99
0.01

Weight

Neuron...

-0.99 -0.06 0.86

Link Absolute

1
0

Link

-0.99 0.02 1

Iteration Plot

Objective / Loss

Iterations

Objective Loss

Confusion Matrix

Observed

	Predicted 0	Predicted 1
0	59K	3.1K
1	14K	8.8K

Predicted
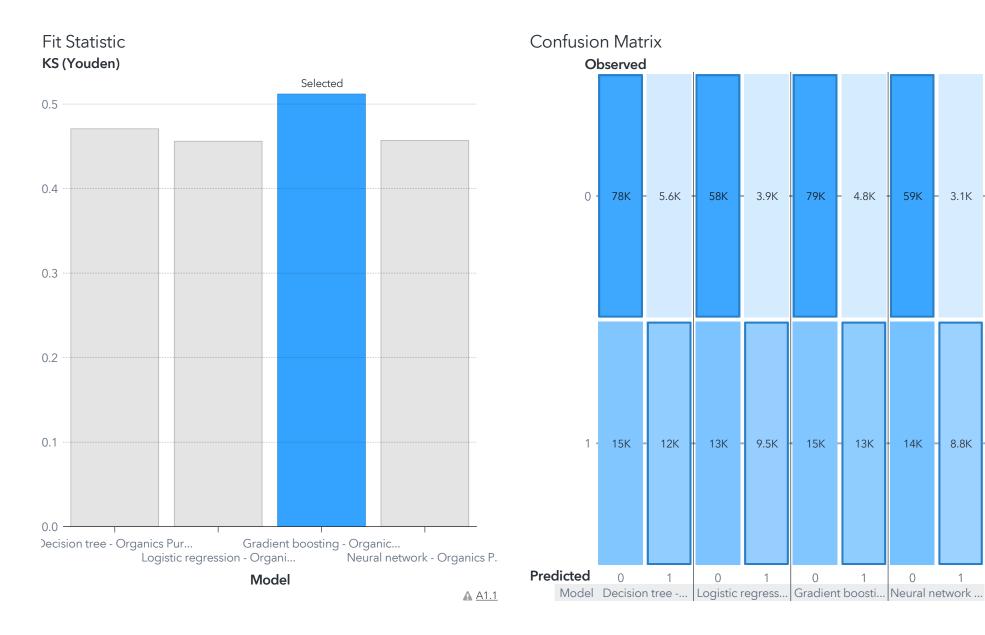
Model Comparison of Organics Purchase Indicator

Event: **1**

Fit Statistic

KS (Youden)

Selected

0.5	
0.4	
0.3	
0.2	
0.1	
0.0	

Decision tree - Organics Pur...
Logistic regression - Organi...
Gradient boosting - Organic...
Neural network - Organics P...

Model

⚠ A1.1

Confusion Matrix

Observed

	Decision tree		Logistic regress...		Gradient boosti...		Neural network ...	
0	78K	5.6K	58K	3.9K	79K	4.8K	59K	3.1K
1	15K	12K	13K	9.5K	15K	13K	14K	8.8K
Predicted	0	1	0	1	0	1	0	1
Model	Decision tree -...		Logistic regress...		Gradient boosti...		Neural network ...	

Appendix

A1.1   Fit Statistic

Warnings: Number of observations for all models do not match.
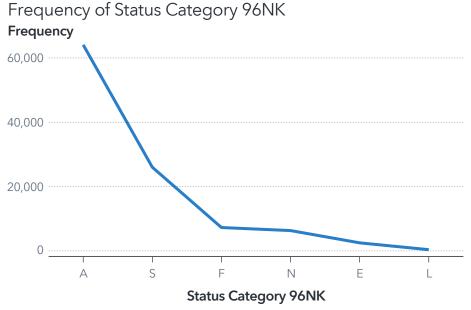
PAGE 72.1
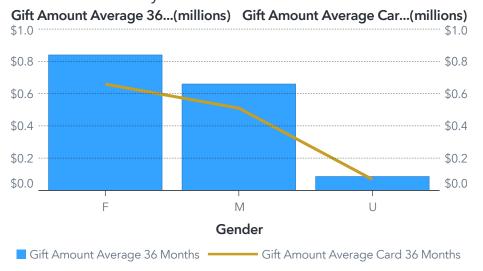
Creation Date: Saturday, August 24, 2024, 01:35:01 PM
Author: mohammed-danish.mustafa@edu.dsti.institute

Frequency of Target Gift Flag

Target Gift Flag

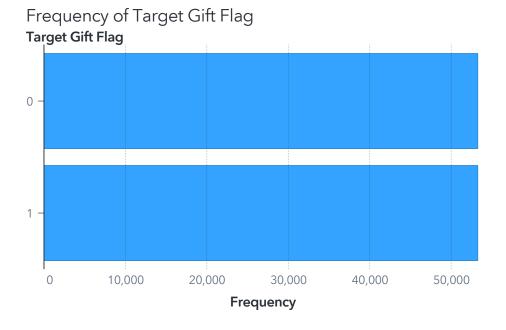**Frequency**

Frequency of Status Category 96NK

Frequency

Status Category 96NK

Gift Amount Average 36 Months, Gift Amount Average Card 36 Months by Gender

Gift Amount Average 36...(millions)   **Gift Amount Average Car...(millions)**

Gender

■ Gift Amount Average 36 Months   —— Gift Amount Average Card 36 Months
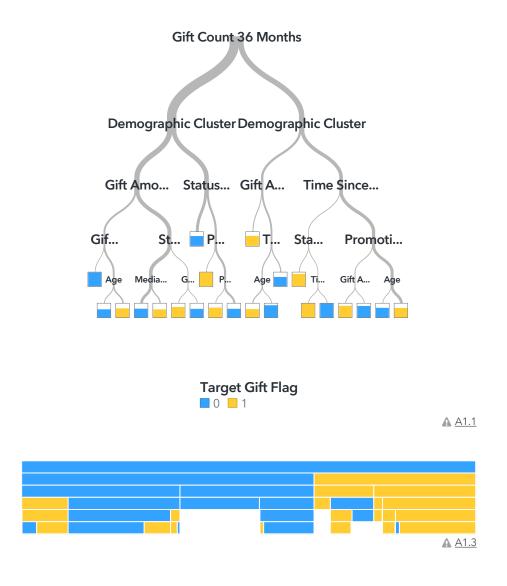
Target Gift Amount, Target Gift Amount with Zero by Target Gift Flag

Target Gift Amount (millions)   **Target Gift Amount with Zero (millions)**

Target Gift Flag

■ Target Gift Amount   ■ Target Gift Amount with Zero
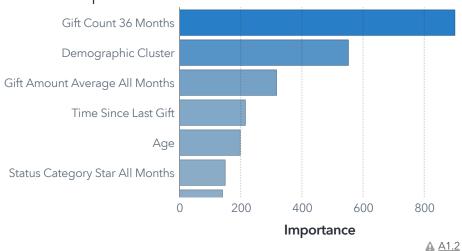
Decision Tree of Target Gift Flag

Event: **1**   Fit: **KS (Youden) 0.1993**   Observations: **107K of 107K**

Tree

Gift Count 36 Months

Demographic Cluster   Demographic Cluster

Gift Amo...   Status...   Gift A...   Time Since...

Gif...   St...   P...   T...   Sta...   Promoti...

Age   Media...   G...   P...   Age   Ti...   Gift A...   Age

Target Gift Flag
■ 0   ■ 1

⚠ A1.1

⚠ A1.3

Variable Importance

Gift Count 36 Months
Demographic Cluster
Gift Amount Average All Months
Time Since Last Gift
Age
Status Category Star All Months

0   200   400   600   800

Importance

⚠ A1.2

Confusion Matrix

Observed

	Predicted 0	Predicted 1
0	35K	19K
1	24K	29K
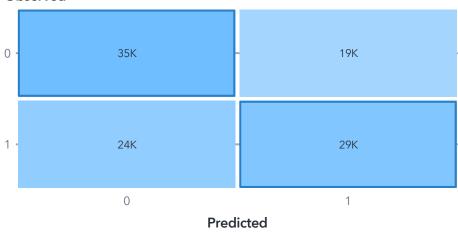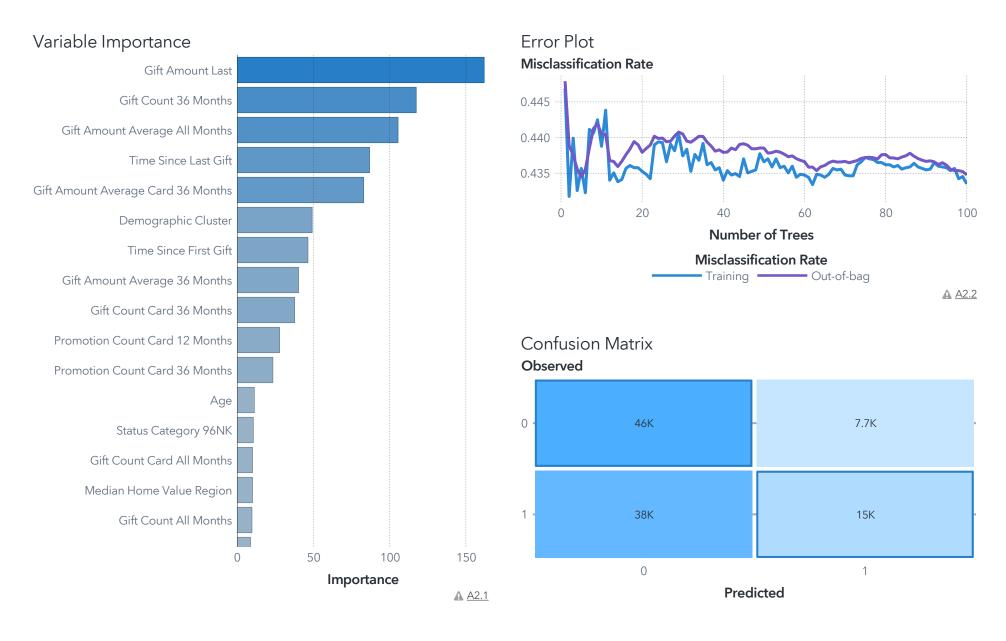
Predicted

Forest of Target Gift Flag

Event: **1**   Fit: **KS (Youden) 0.1565**   Observations: **107K of 107K**

Variable Importance

A2.1

Error Plot

Misclassification Rate

Number of Trees

Misclassification Rate
— Training   — Out-of-bag

A2.2

Confusion Matrix

Observed

Predicted

Gradient Boosting of Target Gift Flag
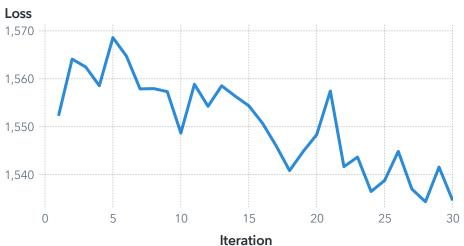
Event: **1**   Fit: **KS (Youden) 0.3452**   Observations: **107K of 107K**

Variable Importance

A3.1

Iteration Plot
Misclassification Rate

Number of Trees

A3.2

Confusion Matrix
Observed

Predicted
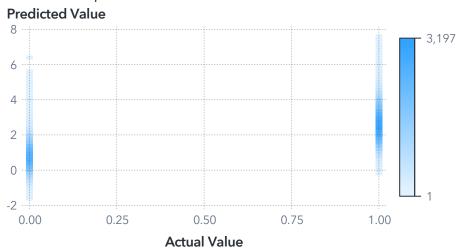
Factorization Machine of Status Category Star All Months

Fit: **ASE 2.8** Observations: **65K of 107K**
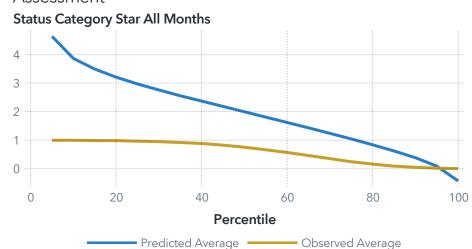
Iteration Plot

Loss

Iteration

Scored Response

Predicted Value

Actual Value

Top Recommendations

Target Gift Flag	Rank
1	1
0	2

Assessment

Status Category Star All Months

Percentile

——— Predicted Average ——— Observed Average
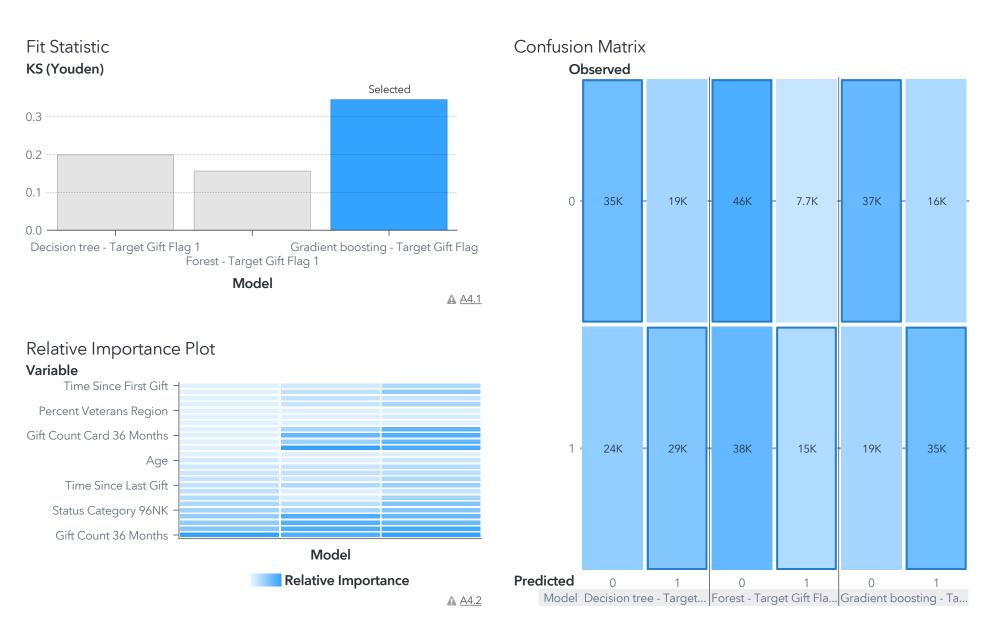
Model Comparison of Target Gift Flag

Event: **1**

Fit Statistic
KS (Youden)

⚠ A4.1

Relative Importance Plot
Variable

- Time Since First Gift
- Percent Veterans Region
- Gift Count Card 36 Months
- Age
- Time Since Last Gift
- Status Category 96NK
- Gift Count 36 Months

Model

Relative Importance

⚠ A4.2

Confusion Matrix
Observed

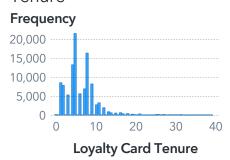|	Decision tree - Target...		Forest - Target Gift Fla...		Gradient boosting - Ta...	
Predicted	0	1	0	1	0	1
0	35K	19K	46K	7.7K	37K	16K
1	24K	29K	38K	15K	19K	35K

Appendix

A1.1 Tree

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A1.2 Variable Importance

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A1.3 Icicle

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A2.1 Variable Importance

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A2.2 Error Plot

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A3.1 Variable Importance

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A3.2 Iteration Plot

Warnings: One or more predictors have been removed from the analysis due to their cardinalities are high.

A4.1 Fit Statistic

Warnings: Decision tree - Target Gift Flag 1: One or more predictors have been removed from the analysis due to their cardinalities are high.

A4.2 Relative Importance Plot

Warnings: Decision tree - Target Gift Flag 1: One or more predictors have been removed from the analysis due to their cardinalities are high.

REPORT 3

Creation Date: Saturday, August 24, 2024, 01:32:49 PM
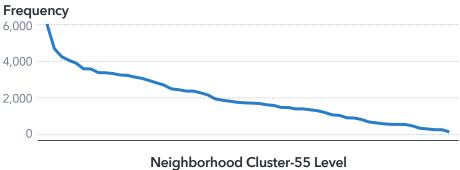Author: mohammed-danish.mustafa@edu.dsti.institute

Frequency of Loyalty Card Tenure

Frequency

Loyalty Card Tenure

Total Spend

491M

Frequency Percent of Gender

Gender

| F | M | U |
| | | (missing) |

55%
7.9%

Frequency Percent

Organics Purchase Count by Loyalty Status

Organics Purchase Count

Gold   Platinum   Silver   Tin

Loyalty Status

Frequency of Customer Loyalty ID

A0... A0...   A0000001120 A0...
A0000003328   A0...
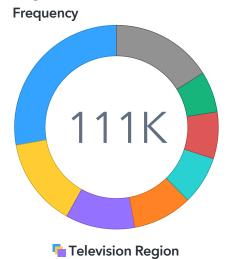A0000004529   A0000002313
A0000009814   A0000010219
A0000005886   A0000007420
A0000002771   A0000003131
A0000000620   A0...
A0000010006   A0000000868

T Frequency

⚠ A1.1

Frequency of Television Region

Frequency

111K

Television Region

Frequency of Neighborhood Cluster-55 Level

Frequency

Neighborhood Cluster-55 Level

Appendix

A1.1   Frequency of Customer Loyalty ID

Warnings: Not all terms are shown.
                 Only 100 rows of the data appear.