

Copyright Notice

These slides are distributed under the Creative Commons License.

[DeepLearning.AI](#) makes these slides available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite [DeepLearning.AI](#) as the source of the slides.

For the rest of the details of the license, see <https://creativecommons.org/licenses/by-sa/2.0/legalcode>



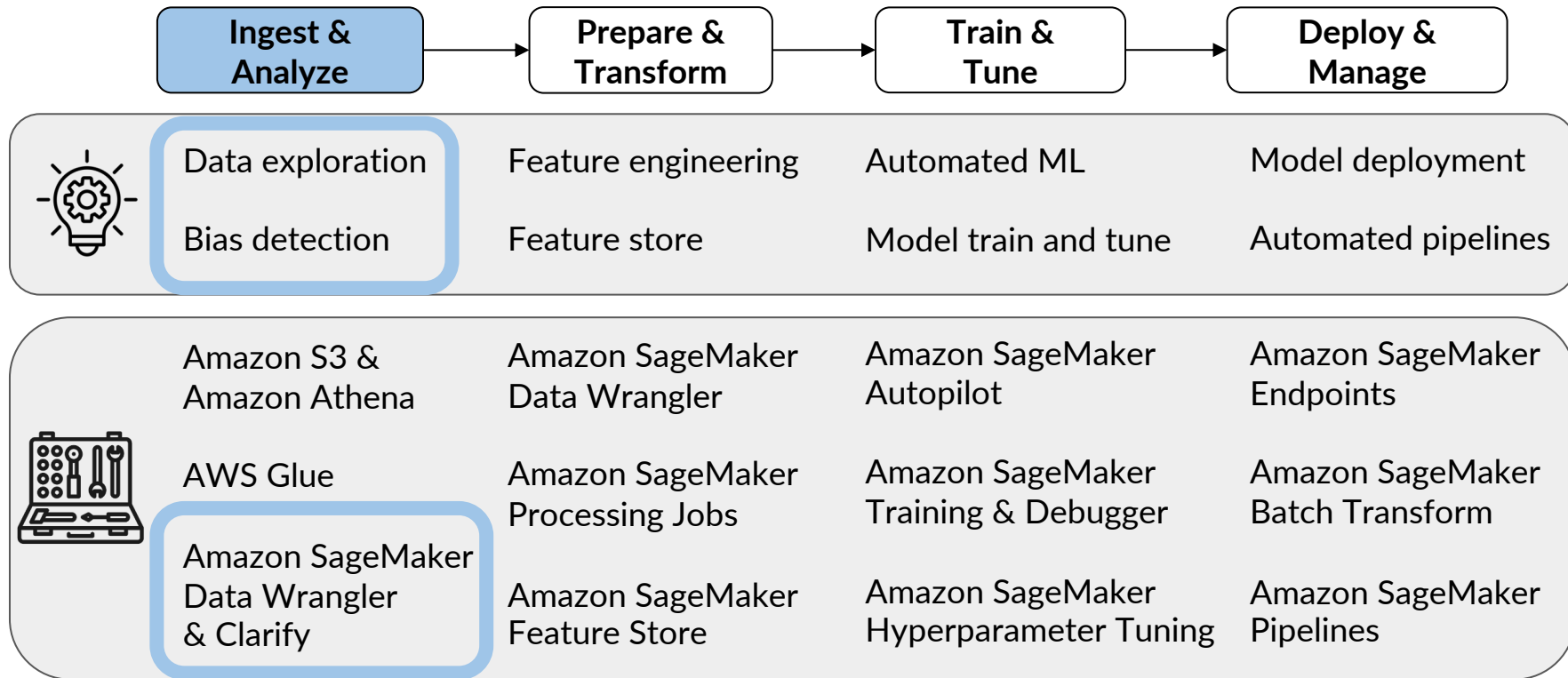
DeepLearning.AI



Practical Data Science

Statistical Bias and Feature Importance

Machine Learning Workflow



Statistical Bias

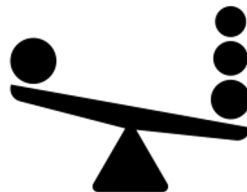


Statistical Bias

- Training data does not comprehensively represent the problem space
- Some elements of a dataset are more heavily weighted or represented



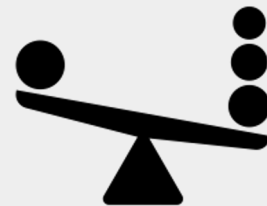
Fraud Detection



Biased models

- Imbalances in product review dataset

Statistical Bias Causes



Statistical Bias – Causes

Statistical Bias – Causes



Activity Bias

Social Media Content

Statistical Bias – Causes



Activity Bias

Social Media Content



Societal Bias

Human Generated Content

Statistical Bias – Causes



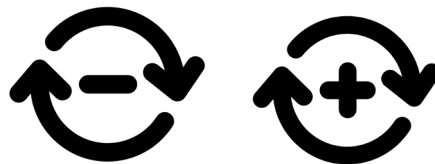
Activity Bias

Social Media Content



Societal Bias

Human Generated Content



Selection Bias

Feedback loop

Statistical Bias – Causes



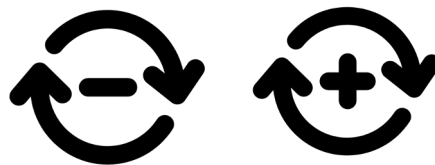
Activity Bias

Social Media Content



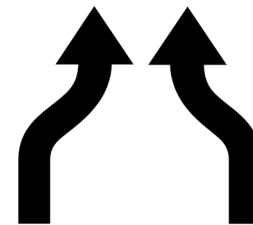
Societal Bias

Human Generated Content



Selection Bias

Feedback Loop



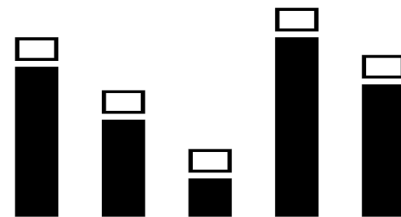
- *Covariant Drift*
- *Prior probability Drift*
- *Concept Drift*

Measuring Statistical Bias



Measure Imbalance in Data - Metrics

Class Imbalance (CI)



- Measures the imbalance in the number of members between different facet values.
- Does a ***product_category*** has disproportionately more reviews than others?

Measure Imbalance in Data - Metrics

Difference in Proportions of Labels (DPL)



- Measures the imbalance of positive outcomes between different facet values.
- Does a ***product_category*** has disproportionately higher ratings than others

Detecting Statistical Bias

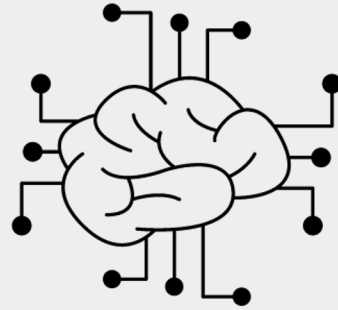
- Amazon SageMaker Data Wrangler
- Amazon SageMaker Clarify



AWS Toolkit

Detect Statistical Bias

Amazon SageMaker Data Wrangler



Detect Statistical Bias - Amazon SageMaker Data Wrangler



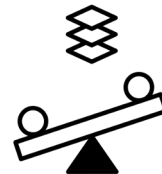
Source



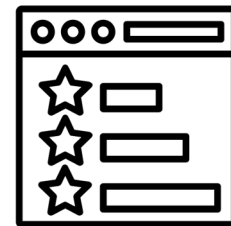
Visualization



Transform



Statistical
Bias Report



Feature
Importance

Detect Statistical Bias - Amazon SageMaker Data Wrangler



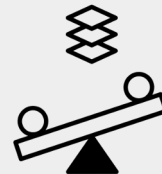
Source



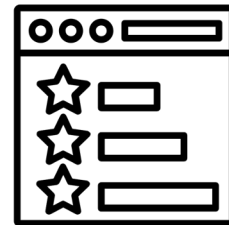
Visualization



Transform



Statistical
Bias Report



Feature
Importance

Browser tabs: Isengard, Amazon SageMaker, JupyterLab

Address bar: <https://d-gpf8j6ix5mch.studio.us-west-2.sagemaker.aws/jupyter/default/lab?>

Amazon SageMaker Studio Launcher

OPEN TABS: Launcher

Get started

Explore one-click solutions, models, and tutorials

SageMaker JumpStart

[Solution: Detect malicious users and transactions](#) →

[Solution: Demand forecasting](#) →

[Go to SageMaker JumpStart](#) →

Build models automatically

SageMaker Autopilot

[Video: Get started with Autopilot](#) →

[Blog: Getting started with Autopilot](#) →

[New autopilot experiment](#) →

Run open-source models with one click

SageMaker JumpStart

[Model: Popular image classification based on ResNet](#) →

[Model: State-of-the-art BERT text processing](#) →

[Explore models](#) →

ML tasks and components

New feature group

Create a new feature group in the feature store to logically group and manage features. [View feature store](#)

New data flow

Prepare and visualize your data with SageMaker Data Wrangler. [View data flows](#)

New project

Organize ML components and automate MLOps with built-in or custom project templates. [View projects](#)

New Autopilot experiment

Create prediction models from your data and start making predictions in a few clicks. [View current experiments](#)

Notebooks and compute resources

Select a SageMaker image: Data Science ▼

Notebook

Python 3

Console

Python 3

Image terminal

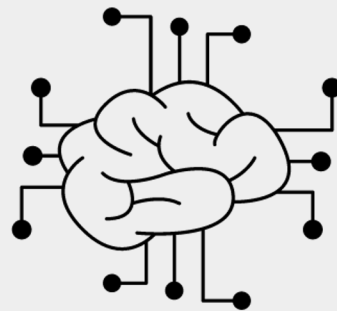
Image Terminal

Launcher

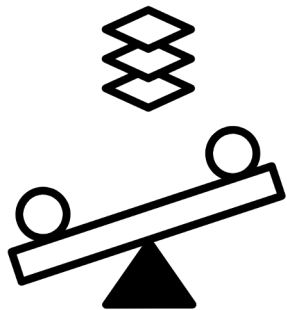
<https://d-gpf8j6ix5mch.studio.us-west-2.sagemaker.aws/jupyter/default/lab?#>

Detect Statistical Bias

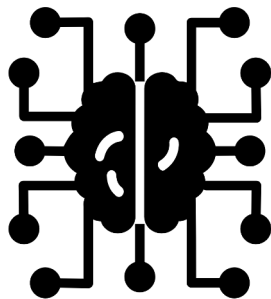
Amazon SageMaker Clarify



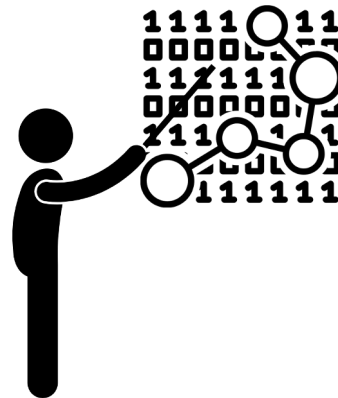
Detect Statistical Bias - Amazon SageMaker Clarify



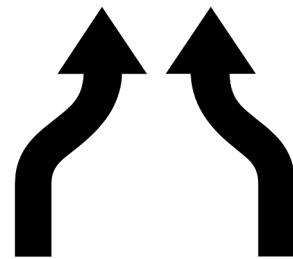
Statistical Bias
Report



Model Bias
Report

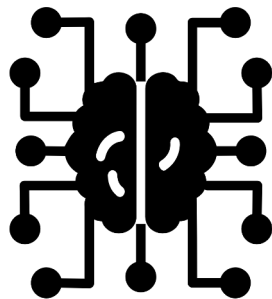
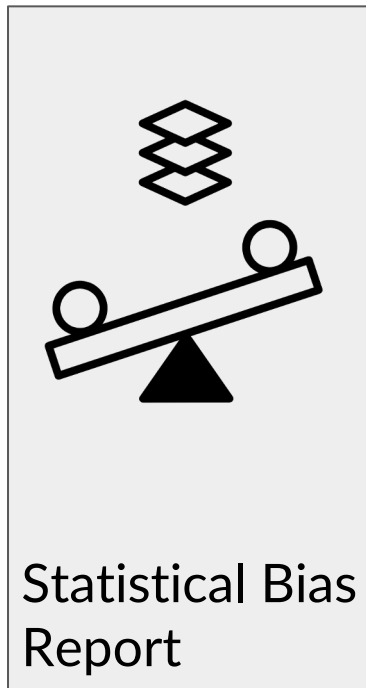


Explainability

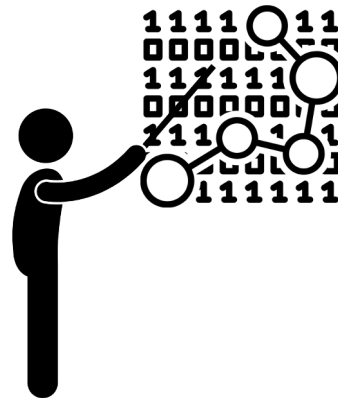


Drift

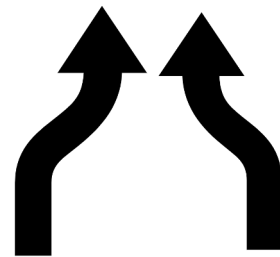
Detect Statistical Bias - Amazon SageMaker Clarify



Model Bias Report



Explainability



Drift

Detect Statistical Bias - Amazon SageMaker Clarify

```
from sagemaker import clarify
```

```
clarify_processor = clarify.SageMakerClarifyProcessor(  
    role=role,  
    instance_count=1,  
    instance_type='ml.c5.2xlarge',  
    sagemaker_session=sess)
```

**Distributed
cluster size**

**Type of each
instance**

```
bias_report_output_path = << Define S3 path >>
```

**S3 location to
store bias report**

Detect Statistical Bias - Amazon SageMaker Clarify

```
bias_data_config = clarify.DataConfig(  
    s3_data_input_path=...,  
    s3_output_path=...,  
    label='sentiment',  
    headers=df_balanced.columns.to_list(),  
    dataset_type='text/csv')
```



Data Configuration

Detect Statistical Bias - Amazon SageMaker Clarify

```
bias_config = clarify.BiasConfig(  
    label_values_or_threshold=[...],  
    facet_name='product_category')
```

Bias Configuration

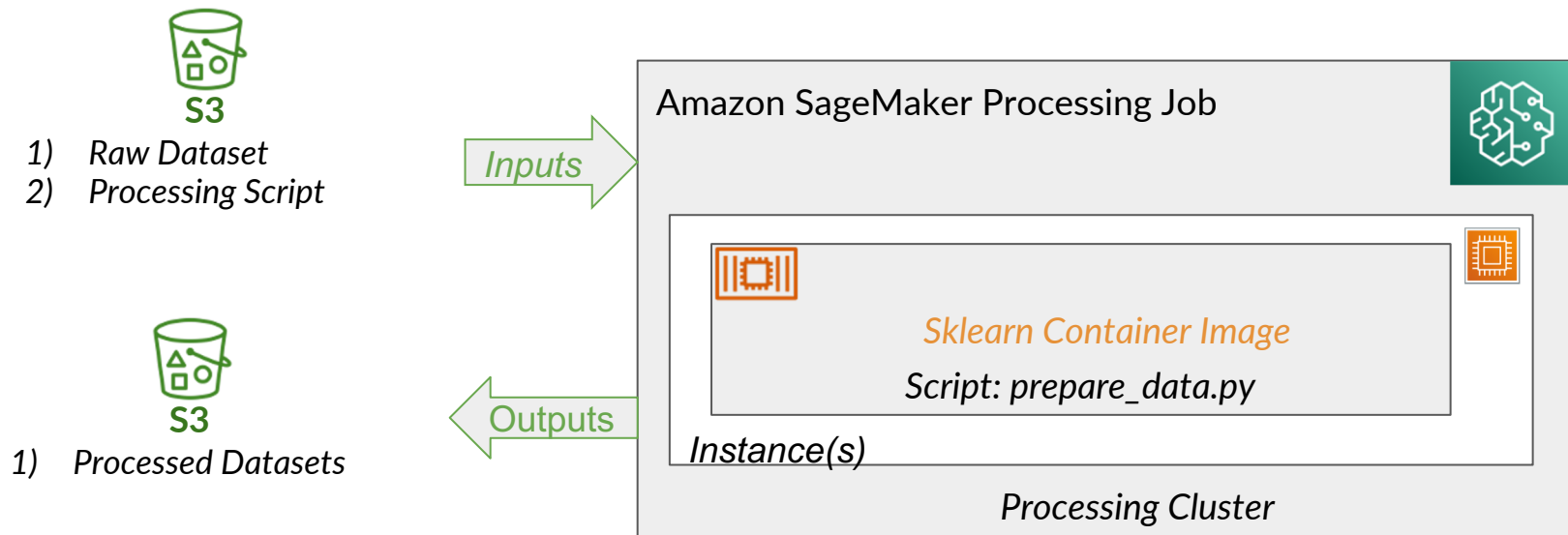
Detect Statistical Bias - Amazon SageMaker Clarify

```
clarify_processor.run_pre_training_bias(  
    data_config=...,  
    data_bias_config=...,  
    methods=["CI", "DPL", ...],  
    wait=<<False/True>>,  
    logs=<<False/True>>)
```

**Pre training bias
job**

Amazon SageMaker Processing

Execute preprocessing, post processing, model evaluation



Detect Statistical Bias - Amazon SageMaker Clarify

```
clarify_processor.run_pre_training_bias(  
    data_config=...,  
    data_bias_config=...,  
    methods=["CI", "DPL", ...],  
    wait=<<False/True>>,  
    logs=<<False/True>>)
```

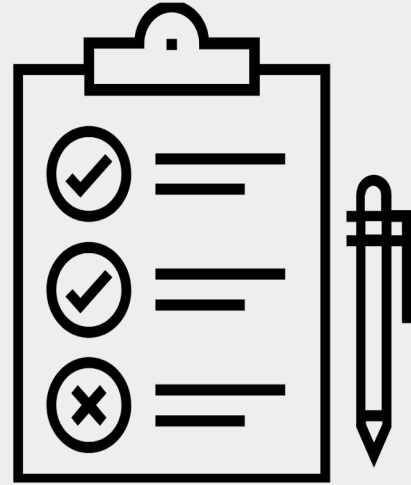
Result ?

Detecting Statistical Bias - Two Approaches

- Amazon SageMaker Data Wrangler
- Amazon SageMaker Clarify

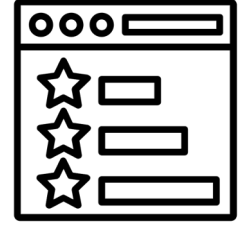
Feature Importance

SHAP



Feature Importance

- Explains the features that make up the training data using a score (importance).
- How useful or valuable the feature is relative to other features.
- Predict the sentiment for a product → Which features play a role?



Feature Importance

- **Open Source Framework - SHapley Additive exPlanations**

Feature Importance

- **Open Source Framework - SHAP**
 - Shapley values based on game theory.

Feature Importance

- **Open Source Framework - SHAP**
 - Shapley values based on game theory.
 - Explain predictions of a ML model
 - Each feature value of training data instance is a player in a game
 - ML prediction is the payout

Feature Importance

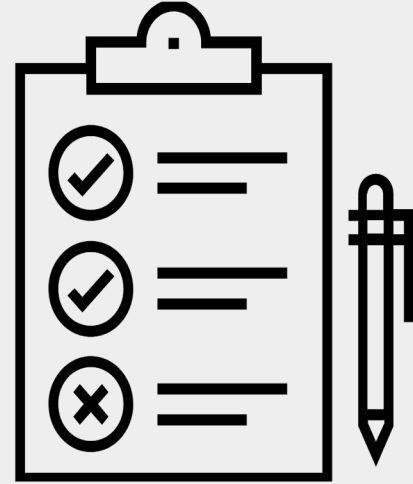
- **Open Source Framework - SHAP**
 - Shapley values based on game theory.
 - Explain predictions of a ML model
 - Each feature value of training data instance is a player in a game
 - ML prediction is the payout
 - Local vs global explanations

Feature Importance

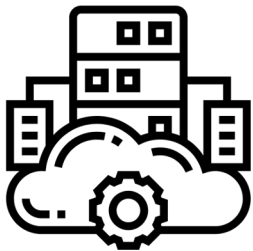
- **Open Source Framework - SHAP**
 - Shapley values based on game theory.
 - Explain predictions of a ML model
 - Each feature value of training data instance is a player in a game
 - ML prediction is the payout
 - Local vs global explanations
 - SHAP can guarantee consistency and local accuracy.

Feature Importance

Amazon SageMaker Data Wrangler



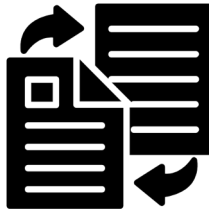
Feature Importance - Amazon SageMaker Data Wrangler



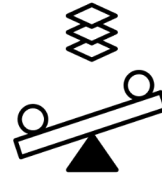
Source



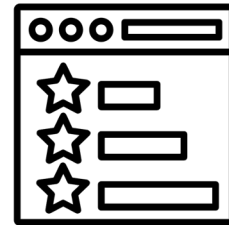
Visualization



Transform

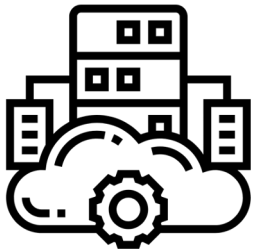


Bias
Report



Feature
Importance

Feature Importance - Amazon SageMaker Data Wrangler



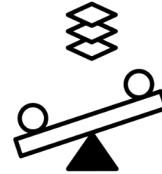
Source



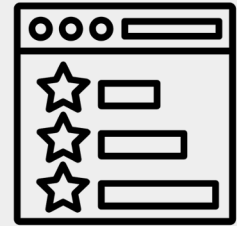
Visualization



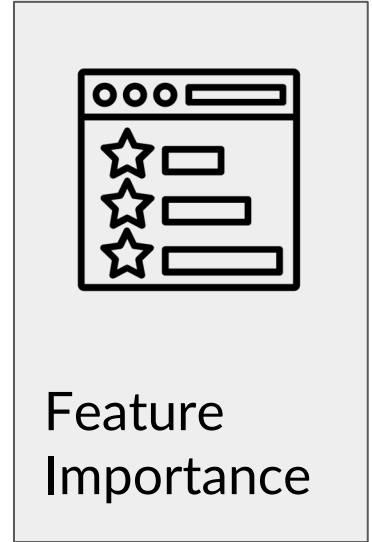
Transform



Bias
Report



Feature
Importance



Amazon SageMaker Studio interface showing the "untitled.flow" notebook. The interface includes a sidebar with file explorer, a top menu bar, and a main workspace. The workspace displays the "Create Analysis" wizard, which is currently in the "Configure" tab. The wizard shows the imported dataset "amazon_reviews_us_giftcards_software_videogames_balanced.csv" and the analysis name "AmazonReviewDataset_F1". The "Label" dropdown menu is open, showing a list of features, with "star_rating" selected. Below the wizard, a data table is displayed, showing columns: marketplace, customer_id, review_id, product_id, product_parent, product_title, and product_category. The table contains 6 rows of data.

Imported datasets / Transform: amazon_reviews_us_giftcards_software_videogames_balanced.csv / AmazonReviewDataset_F1

Create Analysis

Create an analysis of your data. [Learn more.](#)

Quick Model: AmazonReviewDataset_F1

No Preview available

Use Configure for built-in analyses

Use Code to create a custom analysis

Data table

marketplace	customer_id	review_id	product_id	product_parent	product_title	product_category
US	43630866	R3K6VPY2NE6262	B004LLIKVU	473048287	Amazon.com eGift Cards	Gift Card
US	52721479	R2V5BSVBPO5N94	B00CT77B7M	473048287	Amazon.com eGift Cards	Gift Card
US	25341626	RC62YV8GHYO67	B004LLIKVU	473048287	Amazon.com eGift Cards	Gift Card
US	42898788	R35VSTQDN6CPDG	B004LLIKVU	473048287	Amazon.com eGift Cards	Gift Card
US	2943280	R3RA8FDVJD5KB5	B00H58MH44	81025991	Amazon eGift Card - Ho...	Gift Card
US	3929348	R19X9YCLW1614Z	B002VBWIP6	232803743	Xbox Live Subscription	Digital_Video...

Analysis type: Quick Model

A limit of 100,000 rows is used for this analysis.

Analysis name: AmazonReviewDataset_F1

Optional

Label: Select...

marketplace

customer_id

review_id

product_id

product_parent

product_title

product_category

star_rating

helpful_votes

total_votes

Cancel Preview Create