

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

English ▾

SageMaker Immersion Day > Lab 8. Build ML Model with No Code Using Sagemaker Canvas

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

- [Overview](#)
- [Prerequisite](#)
- [Import and manage data](#)
- [Build a model](#)
- [Evaluate your model](#)
- [Make predictions](#)
- [Optional - Standard build model](#)

Overview

Amazon SageMaker Canvas is a new visual, no code capability that allows business analysts to build ML models and generate accurate predictions without writing code or requiring ML expertise. Its intuitive user interface lets you browse and access disparate data sources in the cloud or on-premises, combine datasets with the click of a button, train accurate models, and then generate new predictions once new data is available.

SageMaker Canvas leverages the same technology as [Amazon SageMaker](#) to automatically clean and combine your data, create hundreds of models under the hood, select the best performing one, and generate new individual or batch predictions. It supports multiple problem types such as binary classification, multi-class classification, numerical regression, and time series forecasting. These problem types let you address business-critical use cases, such as fraud detection, churn reduction, and inventory optimization, without writing a single line of code.

Prerequisite

First, we need to launch the SageMaker Canvas application. For that, navigate to SageMaker service and then in the Canvas screen, select "Open Canvas":

- Lab3c. Bring your own Container
- ▼ Lab 4. Autopilot, Debugger and Model Monitor
 - Autopilot
 - Debugger
 - Model Monitor
- ▼ Lab 5. Bias and Explainability
 - Bias and Explainability-Tabular Data
- ▼ Lab 6. SageMaker Pipelines
 - Option 1: Train Pipeline (SageMaker Pipelines)
 - Option 2: Batch Inference Pipeline (SageMaker Pipelines)
 - Option 3: SageMaker Projects
- Lab 7. Real Time ML inference on Streaming Data
- Lab 8. Build ML Model with No Code Using Sagemaker Canvas**
- ▶ Lab 9. Amazon SageMaker JumpStart
- ▶ Lab 10. ML Governance Tools for Amazon SageMaker
- ▶ Lab 11. SageMaker Notebook Instances

Q sagemaker Canvas X

Search results for 'sagemaker'

Services (2)

Features (4)

Blogs (614)

Documentation (58,491)

Amazon SageMaker

Build, Train, and Deploy Machine Learning Models

Amazon SageMaker

SageMaker Canvas

Generate accurate machine learning predictions — no code required

Get Started
Select user profile
sagemakeruser
Open Canvas

It will open the Canvas app in new tab:

▼ Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

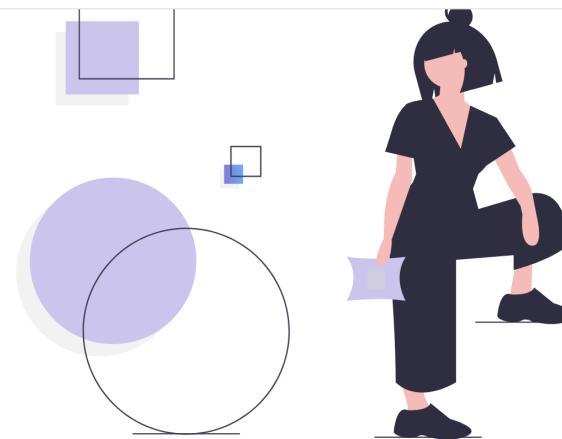


Models



Search models

+ New model



You haven't created any models yet.

Create a new model.

If you click on the bottom left arrow (>) button you see Models and Datasets menu item on the left side.

We already setup SageMaker domain for you as part of this workshop. Now, you need to give yourself the ability to upload local files in SageMaker Canvas. SageMaker Canvas automatically creates a S3 bucket with name that uses the pattern: sagemaker-Region-your-account-id.

If you'd like to have the ability to upload files from your local machine to SageMaker Canvas, you attach a CORS policy to it. To attach a CORS policy, use the following procedure.

1. Go to AWS console and navigate S3 service by searching S3 in search box

The screenshot shows the AWS search interface. A search bar at the top contains the text 's3'. Below the search bar, a sidebar on the left lists various services and documentation categories. On the right, the main area displays search results for 'S3', with a prominent 'Services' section featuring the S3 icon and the text 'Scalable Storage in the Cloud'. Below this, there's a 'Top features' section with links to 'Buckets', 'Access points', and 'Batch Operations'.

2. Choose the bucket with the name that uses the pattern: sagemaker- AWS-Region - AWS-account-id``.

The screenshot shows the 'Amazon S3 > Buckets' page. On the left, a sidebar provides navigation links for buckets, access points, object lambda, multi-region access points, batch operations, and IAM access analyzer. The main content area displays an 'Account snapshot' summary and a table of existing buckets. The table has columns for Name, AWS Region, Access, and Creation date. Two buckets are listed: one named 'sagemaker-us-east-1' which is highlighted with an orange background, and another unnamed bucket in the same row. The table also includes buttons for 'Copy ARN', 'Empty', 'Delete', and 'Create bucket'.

Name	AWS Region	Access	Creation date
sagemaker-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	March 24, 2023, 10:00:54 (UTC+00:00)
	US East (N. Virginia) us-east-1	Objects can be public	March 24, 2023, 17:05:35 (UTC+00:00)

3. Choose **Permissions**.

4. Navigate to **Cross-origins resource sharing (CORS)**.

5. Choose **Edit**.

6. Add the following CORS policy:

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

```
[  
  {  
    "AllowedHeaders": [  
      "*"  
    ],  
    "AllowedMethods": [  
      "POST"  
    ],  
    "AllowedOrigins": [  
      "*"  
    ],  
    "ExposeHeaders": []  
  }  
]
```

▼ Content preferences

Language

- Lab3c. Bring your own Container
- ▼ Lab 4. Autopilot, Debugger and Model Monitor
 - Autopilot
 - Debugger
 - Model Monitor
- ▼ Lab 5. Bias and Explainability
 - Bias and Explainability- Tabular Data
- ▼ Lab 6. SageMaker Pipelines
 - Option 1: Train Pipeline (SageMaker Pipelines)
 - Option 2: Batch Inference Pipeline (SageMaker Pipelines)
 - Option 3: SageMaker Projects
- Lab 7. Real Time ML inference on Streaming Data
- Lab 8. Build ML Model with No Code Using Sagemaker Canvas**
- Lab 9. Amazon SageMaker JumpStart
- Lab 10. ML Governance Tools for Amazon SageMaker
- Lab 11. SageMaker Notebook Instances

Cross-origin resource sharing (CORS)

The CORS configuration, written in JSON, defines a way for client web requests to access resources located in a different domain. [Learn more](#) 

1	[
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	}
14]

```
"AllowedHeaders": [
    "*"
],
"AllowedMethods": [
    "POST"
],
"AllowedOrigins": [
    "*"
],
"ExposeHeaders": []
```

7. choose save changes.

Your SageMaker Canvas setup is complete now.

Import and manage data

In this workshop, you will create a model that solve a binary classification problem. Model predicts whether a lead will convert into a sale or not based on the different marketing activities/metrics captured for that lead by the marketing campaigns. Download the [web](#)

marketing data [CSV](#)! and lead data [CSV](#)! files to your local desktop/laptop. You will upload these 2 files later in the Sagemaker Canvas for training the model. Let's have a look into the data. open both csv files in your laptop.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Lead data csv file provides data about the lead and each lead has a unique ProspectID associated with it. It provides information including job role (JobRole), lead profile (LeadProfile) , whether they used marketing promotion or not (UsedPromo), region (Region), unique Id (prospectID) and whether they converted into a sales or not (Converted) etc. This "converted" field is our target feature for model prediction.

LeadData

LeadSource	JobRole	ContactNotes	LeadProfile	UsedPromo	DoNotReachOut	Region	Converted	ProspectID
Digital Ad	Engineering	ContentPresentation	Large	NoPromo	FALSE	NAMER	0	5f974944-eee6-414f-8bd3-7c22697404ca
Digital Ad	HR	NeedsCaseStudy	Large	NoPromo	FALSE	LATAM	0	5e84c78f-6438-4d91-aa96-b492f7e91029
Event	Sales	FeatureRequest	Medium	NoPromo	TRUE	EMEA	0	38fab8b5-fdbf-4171-b7e2-2a99355a425e
Organic	Management	ContentPresentation	Large	NoPromo	FALSE	LATAM	0	cb9a6f52-a6b5-4be7-8743-9032a0d73884
Referral	Marketing	ContentPresentation	Large	UsedPromo	TRUE	EMEA	1	43afcad1-5292-4bd4-9df8-1bfa688e301e
Digital Ad	Sales	ContentPresentation	Small	UsedPromo	FALSE	APAC	1	8062ca5f-295c-4769-a76b-ad98720c4432

WebMarketingData csv file provides data on what all different marketing activities / matrices were performed by the lead under different campaigns run by marketing team. It provides data including the last campaign activity performed by the lead (LastCampaignActivity), number of page views per visit (PageViewsPerVisit), total time spend on website (TotalTimeOnWebsite), whether the lead attended the marketing event or not (AttendMarketingEvent), whether the lead viewed the advertisement or not (ViewedAdvertisement) etc.

WebMarketingData

ProspectID	LastCampaignActivity	PageViewsPerVisit	TotalTimeOnWebsite	TotalWebVisits	AttendedMarketingEvent	OrganicSearch	ViewedAdvertisement
5f974944-eee6-414f-8bd3-7c22697404ca	Event RSVP	6.342099362	102	7	DidNot	Y	Y
5e84c78f-6438-4d91-aa96-b492f7e91029	DownloadedMaterial	6.393732165	125	7	DidNot	Y	N
38fab8b5-fdbf-4171-b7e2-2a99355a425e	Event RSVP	5.477249709	126	10	DidNot	Y	N
cb9a6f52-a6b5-4be7-8743-9032a0d73884	DownloadedMaterial	7.073818254	138	14	DidNot	N	Y
43afcad1-5292-4bd4-9df8-1bfa688e301e	Open Email	5.099326251	92	17	DidNot	Y	Y
8062ca5f-295c-4769-a76b-ad98720c4432	Web Visit	8.732570244	155	10	DidNot	Y	Y

Please Note: In both files the field, "ProspectID" joins each lead in LeadData csv file to WebMarketingData csv file.

Now you import both of these files in Sagemaker Canvas to train the model.

Navigate to "Datasets" menu item on the left side and click on "import" button on the top right side:

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Amazon SageMaker
Canvas

Models

Datasets

Datasets

Dataset

Join data

Import

Name	Source	Columns	Rows	Cells
canvas-sample-product-descriptions.csv	S3	5	120	600
canvas-sample-housing.csv	S3	10	1,000	10,000
canvas-sample-shipping-logs.csv	S3	12	1,000	12,000
canvas-sample-maintenance.csv	S3	9	1,000	9,000
canvas-sample-sales-forecasting.csv	S3	5	1,000	5,000
canvas-sample-diabetic-readmission.csv	S3	16	1,000	16,000
canvas-sample-loans-part-1.csv	S3	19	1,000	19,000
canvas-sample-loans-part-2.csv	S3	5	1,000	5,000

Please note, at the time of publishing this workshop, you can import data from local files, Amazon S3, Amazon Redshift and Snowflake into SageMaker canvas. You can import data from multiple sources into a single dataset.

Select the "Upload" button and upload both csv files you downloaded earlier.



akadam ▾

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Upload

S3

Add connection

Upload files to import

Drag a CSV file here

or

Select files from your computer

Cancel

Import data

Import

Upload

S3

Add connection

Upload files to import

Drag a CSV file here

or

Select files from your computer

2 datasets ready to import

Delete all

LeadData.csv



WebMarketingData.csv



Preview

Cancel

Import data

Preview both files in the canvas also before importing them. Click on "Import data" button to import both files.

[Privacy policy](#) [Terms of use](#)

Name	Source	Columns	Rows	Created	Status
WebMarketingData.csv	Local	8	10,000	02/16/2022 1:30 PM	Ready
LeadData.csv	Local	9	10,000	02/16/2022 1:30 PM	Ready

Now we join both datasets which create a new dataset for model training . Select the button “Join data” in the top right corner and drag and drop both of the datasets in the Canvas UI:

We will use “ProspectID” field to join both datasets .

Join Datasets

Datasets + Import

Search...

WebMarketingData.csv

LeadData.csv

Drag and drop datasets to join

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Join Datasets

X

- Lab3c. Bring your own Container
- ▼ Lab 4. Autopilot, Debugger and Model Monitor
 - Autopilot
 - Debugger
 - Model Monitor
- ▼ Lab 5. Bias and Explainability
 - Bias and Explainability- Tabular Data
- ▼ Lab 6. SageMaker Pipelines
 - Option 1: Train Pipeline (SageMaker Pipelines)
 - Option 2: Batch Inference Pipeline (SageMaker Pipelines)
 - Option 3: SageMaker Projects
- Lab 7. Real Time ML inference on Streaming Data
- Lab 8. Build ML Model with No Code Using Sagemaker Canvas**
- ▶ Lab 9. Amazon SageMaker JumpStart
- ▶ Lab 10. ML Governance Tools for Amazon SageMaker
- ▶ Lab 11. SageMaker Notebook Instances

Datasets [+ Import](#) Autosaved 2/16/22 at 1:42:11 PM

Search...

WebMarketingData.csv
LeadData.csv

Join preview Search columns Show dropped columns

<input checked="" type="checkbox"/>	Column name	Data type	Missing values	Invalid values	Distinct values

Close Import data

Select the intersection of both the datasets and select "Outer" join option as shown below. Sagemaker Canvas automatically identifies "ProspectID" as the joined key in both of the dataset and select for you.

Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Join Datasets

Datasets + Import Autosaved 2/16/22 at 1:44:00 PM

Search... WebMarketingData.csv LeadData.csv

Join preview

Column name	Data type	Missing
ProspectID	String	0

Join

Inner Left Right Outer

WebMarketingData.csv LeadData.csv

ProspectID = ProspectID

+ Add Key Save & close Close Import data

```
graph LR; A[WebMarketingData.csv] --> J(( )); J --> B[LeadData.csv]
```

Click “Save & Close” button. Before importing the data, you preview the joined dataset in case you want as shown below where it shows statistical data in numerical and graphical forms; for each column in the new dataset. Click on the double arrows to expand the join preview section:

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Join Datasets



Join preview



Search columns

Show dropped columns



<input checked="" type="checkbox"/>	Column name	Data type	Missing values	Invalid values	Distinct values	Mean/Mode
<input checked="" type="checkbox"/>	ProspectID	Text	0	--	100	00a7d348-3bae-48ae-9cf2-ee55f717af
<input checked="" type="checkbox"/>	LastCampaignActivity	Text	0	--	4	Open Email
<input checked="" type="checkbox"/>	PageViewsPerVisit	Numeric	0	--	100	10.02773543
<input checked="" type="checkbox"/>	TotalTimeOnWebsite	Numeric	0	--	69	102
<input checked="" type="checkbox"/>	TotalWebVisits	Numeric	0	--	13	10
<input checked="" type="checkbox"/>	AttendedMarketingEvent	Binary	0	--	2	DidNot
<input checked="" type="checkbox"/>	OrganicSearch	Binary	0	--	2	Y
<input checked="" type="checkbox"/>	ViewedAdvertisement	Binary	0	--	2	Y
<input checked="" type="checkbox"/>	LeadSource	Text	0	--	5	Digital Ad
<input checked="" type="checkbox"/>	JobRole	Text	0	--	7	HR
<input checked="" type="checkbox"/>	ContactNotes	Text	0	--	4	NeedsTraining
<input checked="" type="checkbox"/>	LeadProfile	Text	0	--	3	Small
<input checked="" type="checkbox"/>	UsedPromo	Binary	0	--	2	UsedPromo
<input checked="" type="checkbox"/>	DoNotReachOut	Binary	0	--	2	TRUE
<input checked="" type="checkbox"/>	Region	Text	0	--	4	LATAM
<input checked="" type="checkbox"/>	Converted	Binary	0	--	2	1

Close

Import data

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

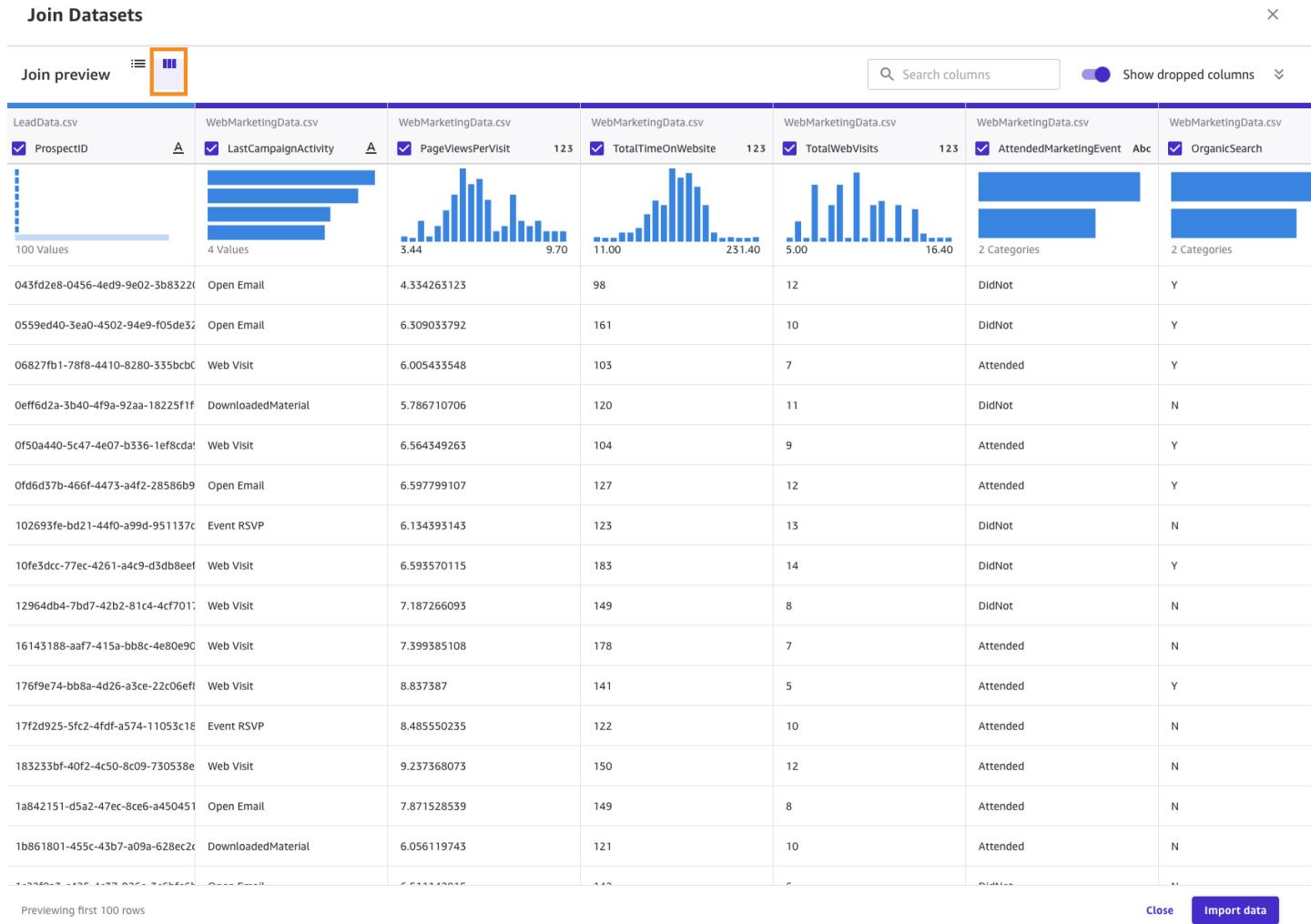
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



Click on “import data” button and provide import dataset name as “joined_campaign_data”. Click on import data button again to save the new dataset. it takes few seconds to process and new dataset comes in ready status as shown below:

Amazon SageMaker Canvas

Models

Datasets

Datasets

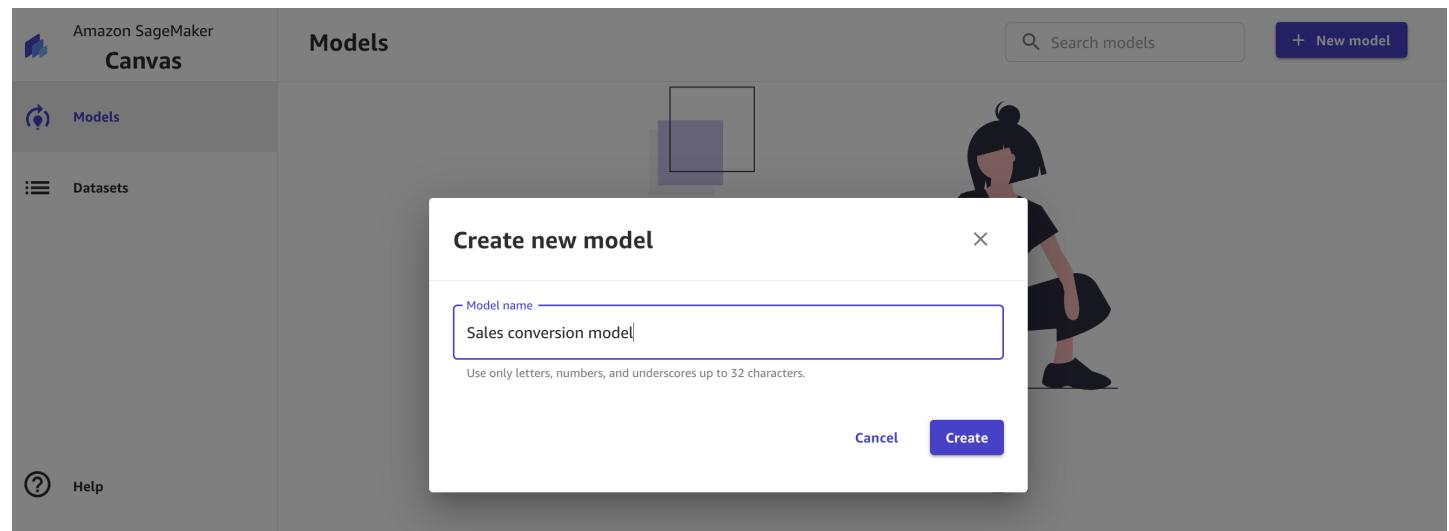
Dataset Name Source Columns Rows Cells Created Status

joined_campaign_data	Joined	16	10,000	160,000	03/24/2023 5:24 PM	Ready
WebMarketingData.csv	Local	8	10,000	80,000	03/24/2023 5:15 PM	Ready
LeadData.csv	Local	9	10,000	90,000	03/24/2023 5:15 PM	Ready

We use this new dataset to train our model.

Build a Model

Navigate to “Models” menu item in left pane and click on “New Model” button and give model name as “Sales conversion model” and click “Create”:



Select the ‘joined-campaign_data’ as dataset and click on select dataset button.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Amazon SageMaker Canvas

Models Datasets Help Account

Sales conversion model

V1 • Draft Add version Share :

Select Build Analyze Predict

Select dataset

You can import a tabular dataset or choose one that has already been imported. Your dataset must contain at least one input column and a target column.

Search datasets in Canvas

All Joined

Name	Source	Columns	Rows	Created	Status
joined_campaign_data	Joined	16	10,000	02/16/2022 1:58 PM	Ready
WebMarketingData.csv	Local	8	10,000	02/16/2022 1:30 PM	Ready
LeadData.csv	Local	9	10,000	02/16/2022 1:30 PM	Ready

Close Select dataset

Select the Target column as "Converted".

As soon as you select your dataset, you're presented with an overview that outlines the data types, missing values, mismatched values, unique values, and the mean or mode values of the respective columns.

After you import the data and select a target column, you're given a choice to validate your data. Data Validation is a new feature in SageMaker Canvas to proactively check for potential data quality issues. Click on "Validate data".

Sales conversion model

Select Build Analyze Predict

① Validate your data
It might take several minutes, depending on the dataset size.

Select a column to predict
Choose the target column. The model that you build predicts values for the column that you select.

Target column: Converted

Value distribution: 1 (blue bar) 0 (blue bar)

Model type
SageMaker Canvas automatically recommends the appropriate model type for your analysis.

2 category prediction
Your model classifies Converted into two categories.

Change type

Quick build

Preview model

joined_campaign_data
Full dataset: 10.0k rows

Column name	Data type	Missing	Mismatched	Unique	Mean / Mode	Correlation to target
ViewedAdvertisement	Binary	0.00% (0)	0.00% (0)	2	N	-0.016
UsedPromo	Binary	0.00% (0)	0.00% (0)	2	NoPromo	0.528
TotalWebVisits	Numeric	0.00% (0)	0.00% (0)	19	10	0.131
TotalTimeOnWebsite	Numeric	0.00% (0)	0.00% (0)	240	125	0.189
Region	Categorical	0.00% (0)	0.00% (0)	4	APAC	N/A
ProspectID	Text	0.00% (0)	0.00% (0)	10,000	0009de69-00a3-463a...	N/A
PageViewsPerVisit	Numeric	0.00% (0)	0.00% (0)	10,000	1.93	0.172
OrganicSearch	Binary	0.00% (0)	0.00% (0)	2	N	0.049
LeadSource	Categorical	0.00% (0)	0.00% (0)	5	Cold Outbound	N/A
LeadProfile	Categorical	0.00% (0)	0.00% (0)	3	Large	N/A
LastCampaignActivity	Categorical	0.00% (0)	0.00% (0)	4	DownloadedMaterial	N/A
JobRole	Categorical	0.00% (0)	0.00% (0)	7	HR	N/A
DoNotReachOut	Binary	0.00% (0)	0.00% (0)	2	False	0.046
Converted	Target	0.00% (0)	0.00% (0)	2	1	--
ContactNotes	Categorical	0.00% (0)	0.00% (0)	4	ContentPresentation	N/A

Total columns: 16 Total rows: 10,000 Total cells: 160,000 Show dropped columns

If you choose to validate your data, Canvas analyzes your data for numerous conditions including:

- Too many unique labels in your target column – for the category prediction model type
- Too many unique labels in your target column for the number of rows in your data – for the category prediction model type
- Wrong model type for your data – the model type doesn't fit the data you're predicting in the Target column
- Too many invalid rows – missing values in your target column
- All feature columns are text columns – they will be dropped for standard builds

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

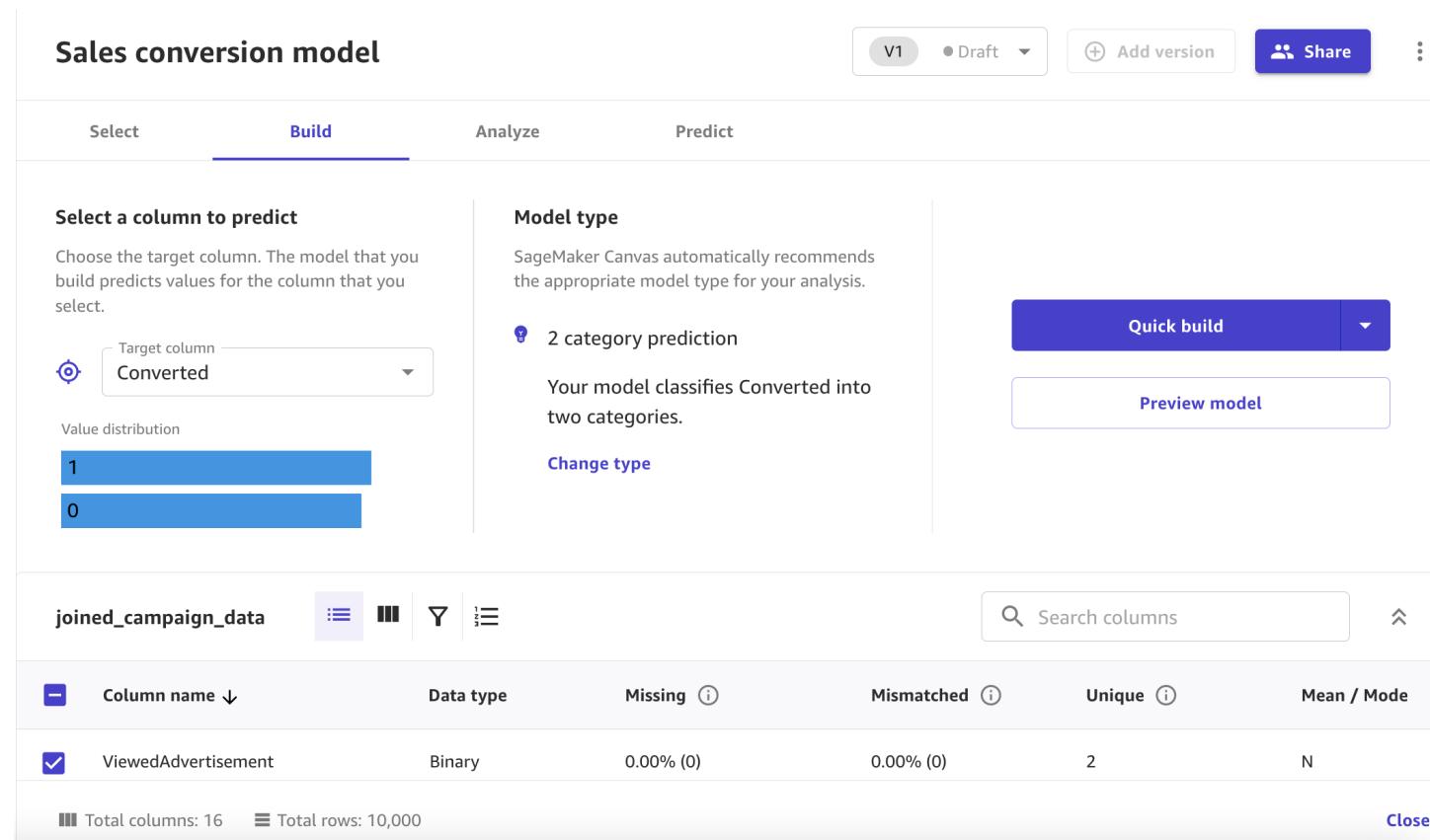
Language

- Too few columns – too few columns in your data
- No complete rows – all of the rows in your data contain missing values

After a few second, data validation should be completed, and for this dataset no issues should have been found.



Expand the joined_campaign_data preview window by clicking on double arrow button and remove the select mark against the ProspectID field as it doesn't have any impact on 'converted' target feature and hence not required to train the model.



Column name	Data type	Missing	Mismatched	Unique	Mean / Mode
ViewedAdvertisement	Binary	0.00% (0)	0.00% (0)	2	N

Target column: Converted

2 category prediction

Change type

Quick build

Preview model

joined_campaign_data
Full dataset: 10.0k rows

Column name	Data type	Missing	Mismatched	Unique	Mean / Mode	Correlation to target
ViewedAdvertisement	Binary	0.00% (0)	0.00% (0)	2		-0.016
UsedPromo	Binary	0.00% (0)	0.00% (0)	2	NoPromo	0.528
TotalWebVisits	Numeric	0.00% (0)	0.00% (0)	19	10	0.131
TotalTimeOnWebsite	Numeric	0.00% (0)	0.00% (0)	240	125	0.189
Region	Categorical	0.00% (0)	0.00% (0)	4	APAC	N/A
ProspectID	Text	0.00% (0)	0.00% (0)	10,000	0009de69-00a3-463a...	N/A
PageViewsPerVisit	Numeric	0.00% (0)	0.00% (0)	10,000	1.93	0.172
OrganicSearch	Binary	0.00% (0)	0.00% (0)	2		0.049
LeadSource	Categorical	0.00% (0)	0.00% (0)	5	Cold Uutbound	N/A
LeadProfile	Categorical	0.00% (0)	0.00% (0)	3	Large	N/A
LastCampaignActivity	Categorical	0.00% (0)	0.00% (0)	4	DownloadedMaterial	N/A
JobRole	Categorical	0.00% (0)	0.00% (0)	7	HR	N/A
DoNotReachOut	Binary	0.00% (0)	0.00% (0)	2	False	0.046
Converted	Target	0.00% (0)	0.00% (0)	2	1	--
ContactNotes	Categorical	0.00% (0)	0.00% (0)	4	ContentPresentation	N/A
AttendedMarketingEvent	Binary	0.00% (0)	0.00% (0)	2	Attended	-0.19

Click on "model recipe" button (as highlighted in the picture below) which demonstrates that ProspectID feature is dropped now:

Target column: Converted

2 category prediction

Change type

Quick build

Preview model

joined_campaign_data
Full dataset: 10.0k rows

Model recipe

Drop column ProspectID

Column name	Data type	Missing	Mismatched	Unique	Mean / Mode	Correla
ViewedAdvertisement	Binary	0.00% (0)	0.00% (0)	2		-0.016

Before going ahead to build a model, SageMaker Canvas provides you functionality of previewing your model before building it. SageMaker Canvas gives you the ability to get insights from your data before you build a model by choosing Preview model. For example, you can see how the data in each column is distributed. For models built using categorical data, you choose Preview model to generate an Estimated accuracy prediction of how well the model can analyze your data. SageMaker Canvas uses a subset of your data to build a model quickly to check if your data is ready to generate an accurate prediction. Using this sample model, you can understand the current model accuracy and the relative impact of each column on predictions.

Click on the "Preview model" button and wait for couple of minutes till it generates estimated accuracy and column impact of each column in this model. Amazon SageMaker Canvas automatically handles missing values in your dataset while it builds the model. It infers the missing values by using adjacent values that are present in the dataset.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

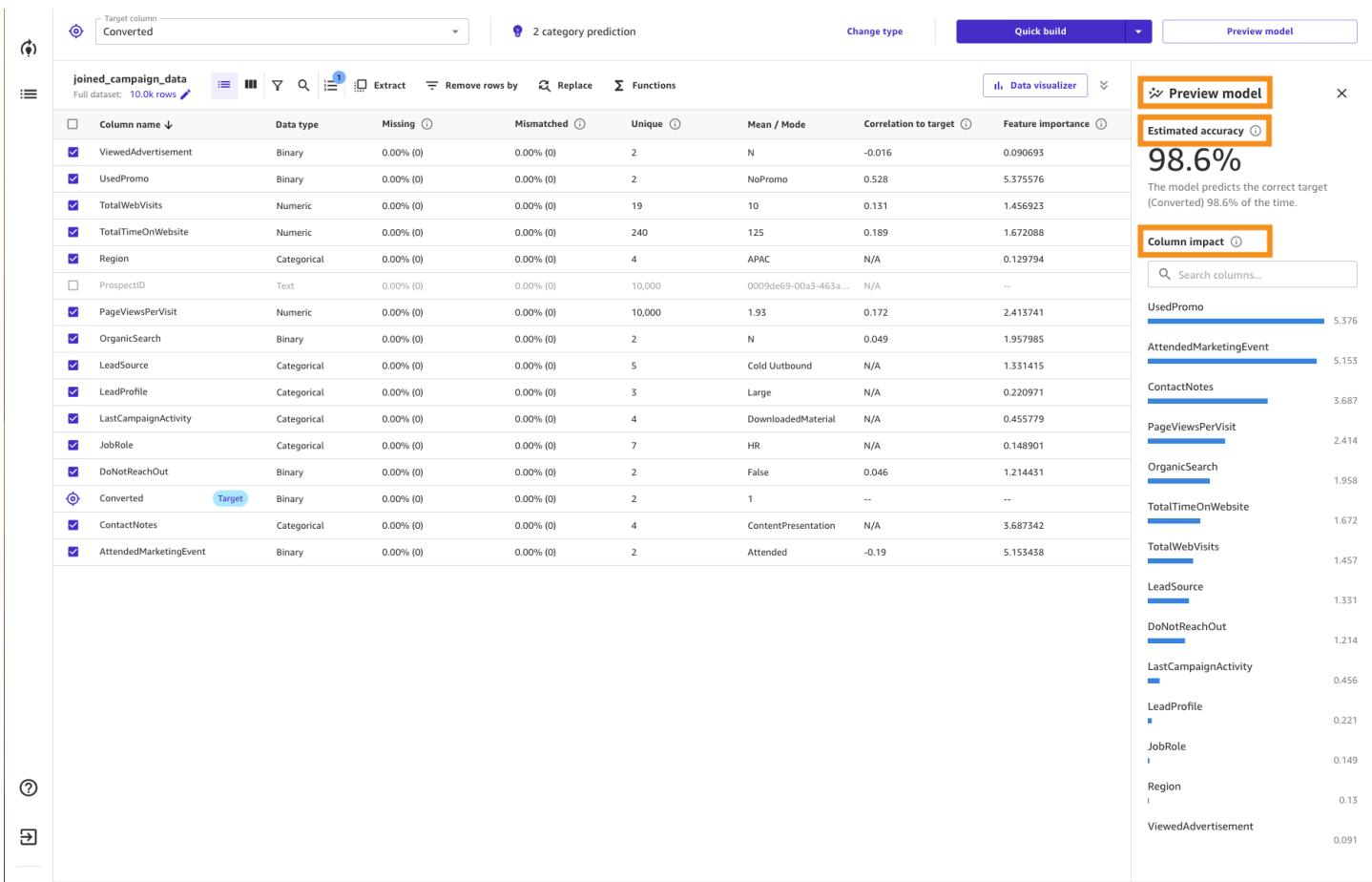
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



As we can see here that the accuracy is quite good (your model accuracy and column impact order may differ from above as it depends what sample dataset was picked by preview mode functionality).

Now select the "Grid View" (as highlighted in orange colour below) to see how the data in each column is distributed.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

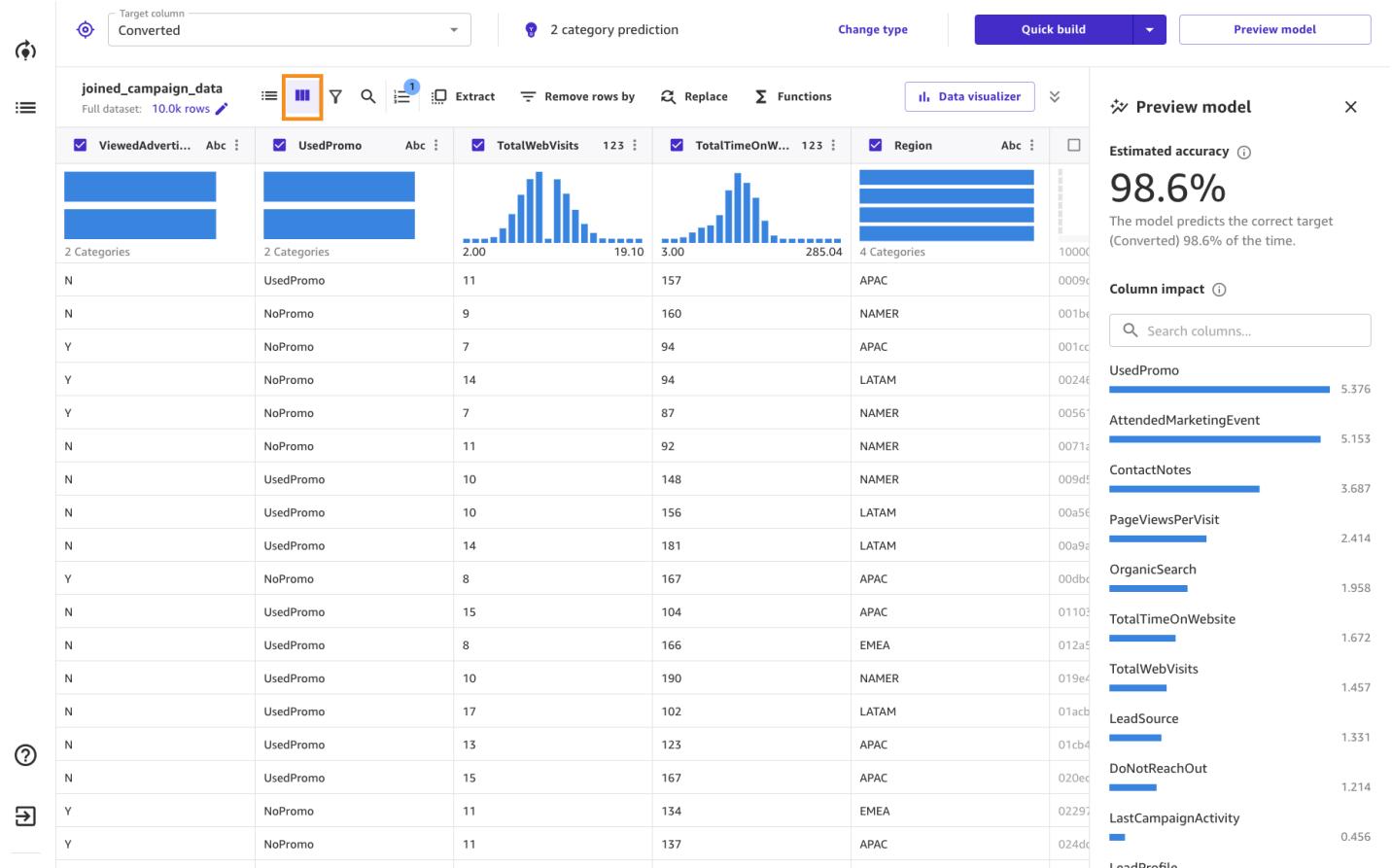
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



Now click on the “Data visualizer” button. This feature in Canvas allows you to see if there are relationships between elements. You can explore and visualize your data, to help you gain advanced insights into your data before building your ML models. You can visualize using scatter plots, bar charts, and box plots, which can help you understand your data and discover the relationships between features that could affect the model accuracy.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

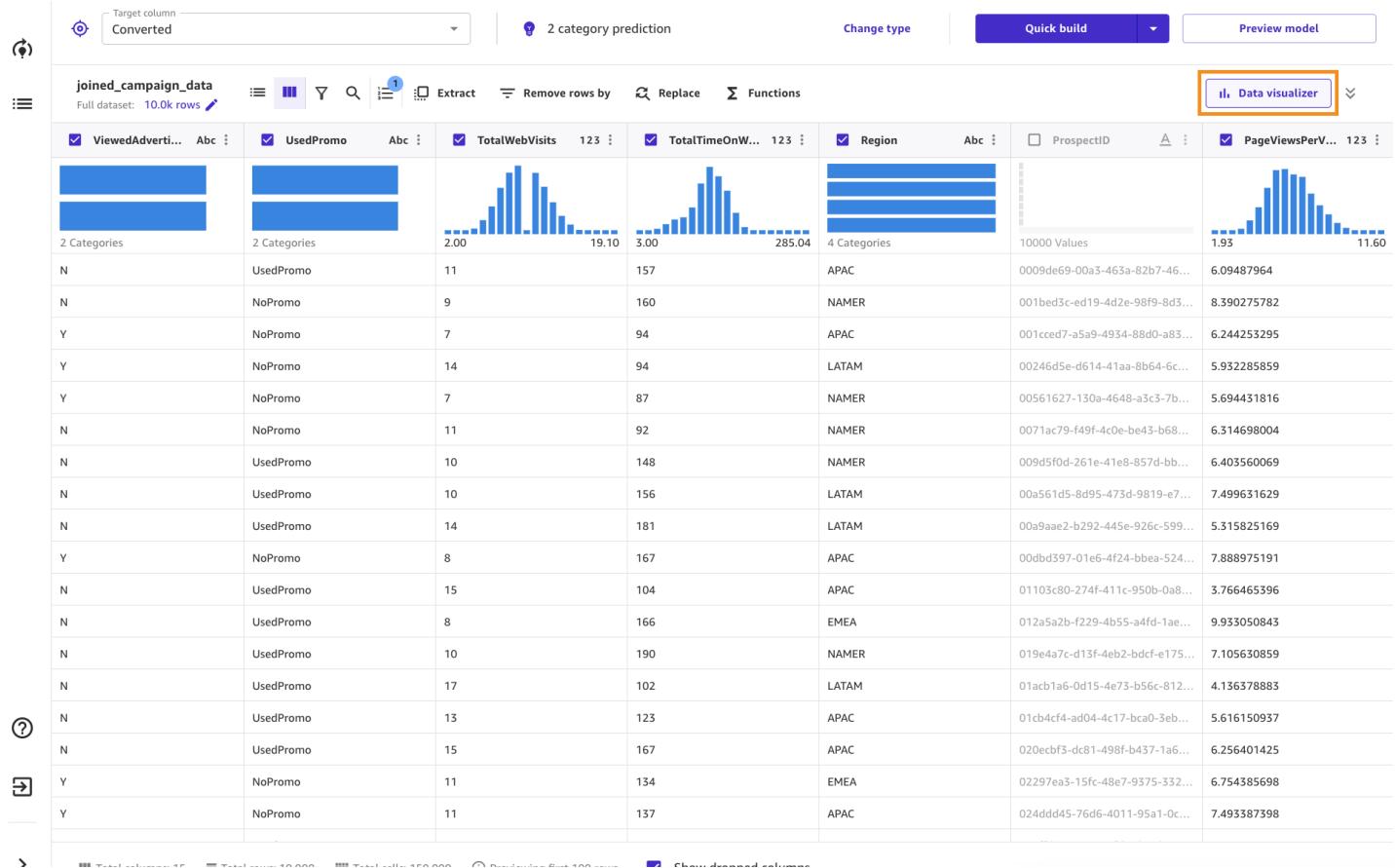
► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language



You may want to change the sample size based on your dataset to get a better perspective. To do this in the Data Visualizer choose number of rows next to Visualization sample and then use the slider to select your desired sample size. Once you have selected 350 rows, click "Update".

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

The screenshot shows the Amazon SageMaker Canvas interface. At the top, there's a header with 'Target column: Converted', '2 category prediction', 'Change type', 'Quick build', and 'Preview model'. Below the header, a table lists 'joined_campaign_data' with '100 rows' selected. The interface has tabs for 'Visualizations' and 'Analytics', with 'Visualizations' currently active. Under 'Visualizations', there are icons for 'Scatter plot', 'Bar chart', and 'Box plot'. A 'Columns' section lists various columns like 'LeadSource', 'ContactID', 'UsedP...', 'Region', 'Prospect...', 'PageViews...', 'TotalTi...', 'TotalWe...', and 'Attendee...'. A 'Random sample' dialog box is open in the foreground, showing a slider set to '350 rows'. The background shows a faint visualization of a scatter plot.

Under visualization you will see "Scatter plot" is selected by default. A scatter plot shows the relationship between two quantitative variables measured for the same individuals. In our case, it's important to understand the relationship between values to check for any correlation. For this we are going to select the column "PageViewsPerVisit" and drag it to the Y axis, and select the column "TotalTimeOnWebsite" and drag it to the X axis.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

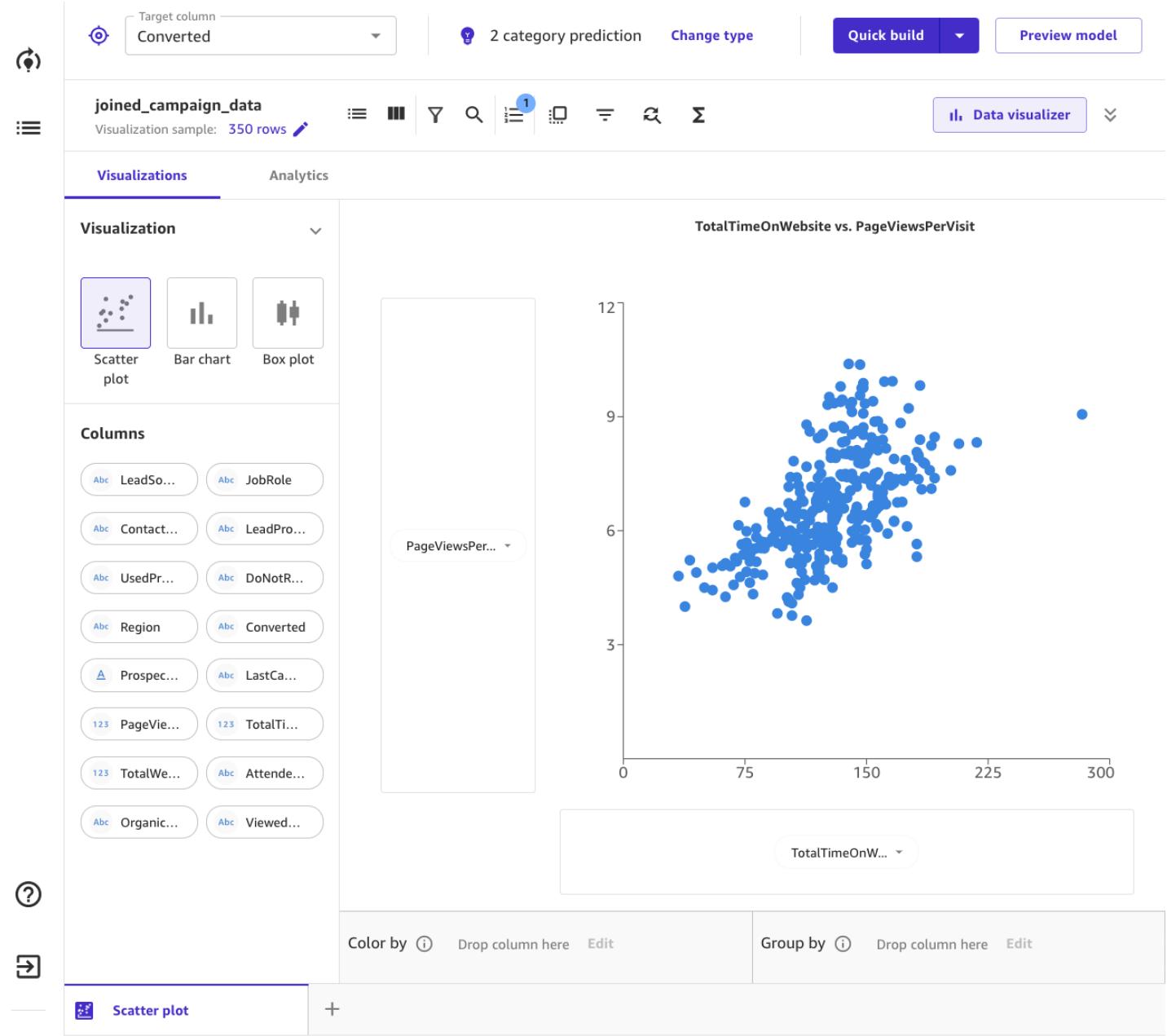
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

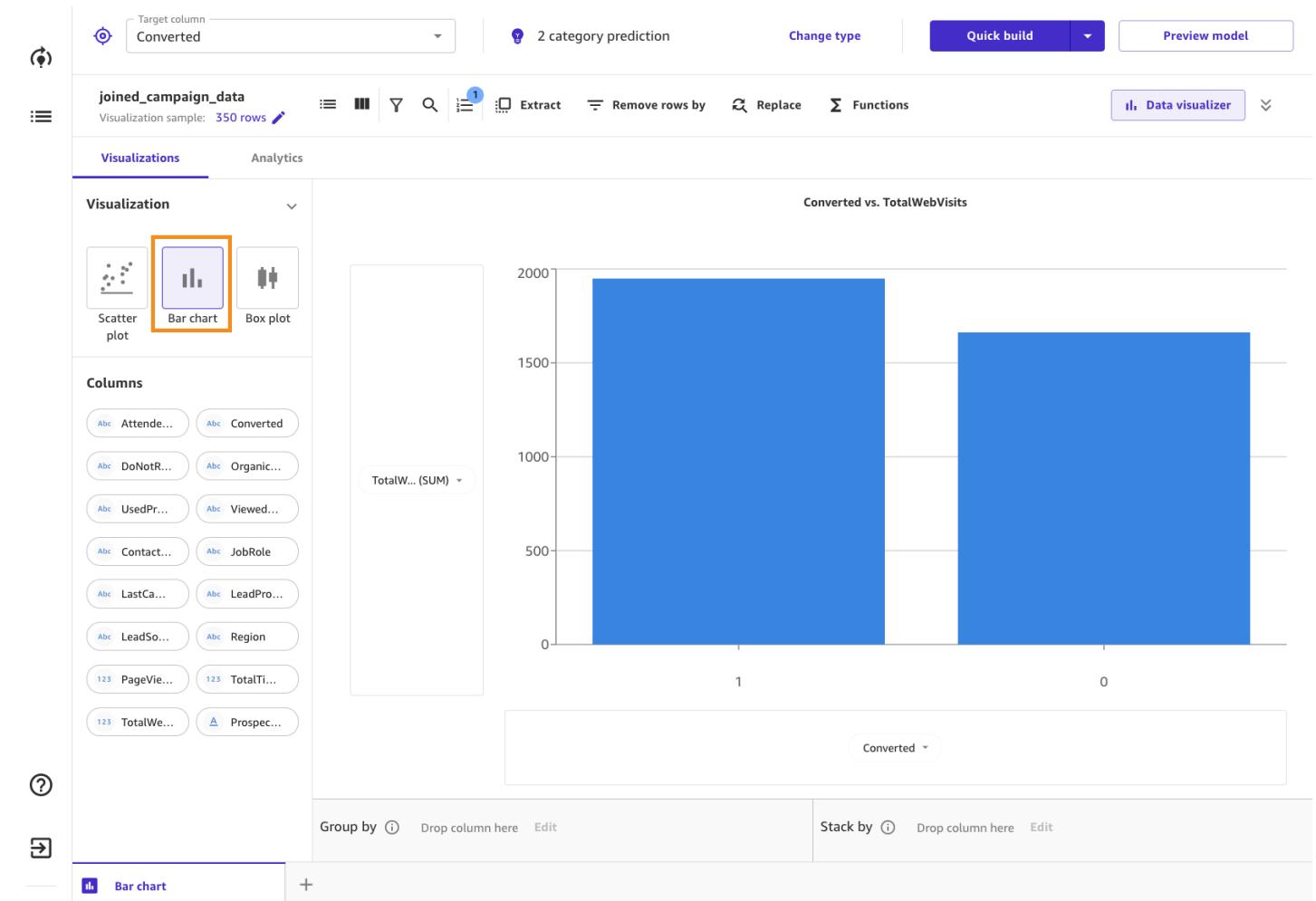
► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



Now let's look at data balance and variation. We will create a bar chart to see the how the number of web visits are distributed across our target column Converted or not. Under visualization select "Bar chart". Select column TotalWebVisits and drag it to the Y axis, and select column "Converted" and drag it to the X axis.



It looks like behaviour for web visits didn't have in our target column.

Now let's look at Box plots. Box plots are useful because they show differences in behaviour of data by class (churn or not). Because we're going to predict churn (target column), let's create a box plot of some features against our target column to infer descriptive statistics on the dataset such as mean, max, min, median, and outliers.

Select "Box Plots" and drag and drop "TotalTimeOnWebsite" and "Converted" to the y-axis and x-axis, respectively.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

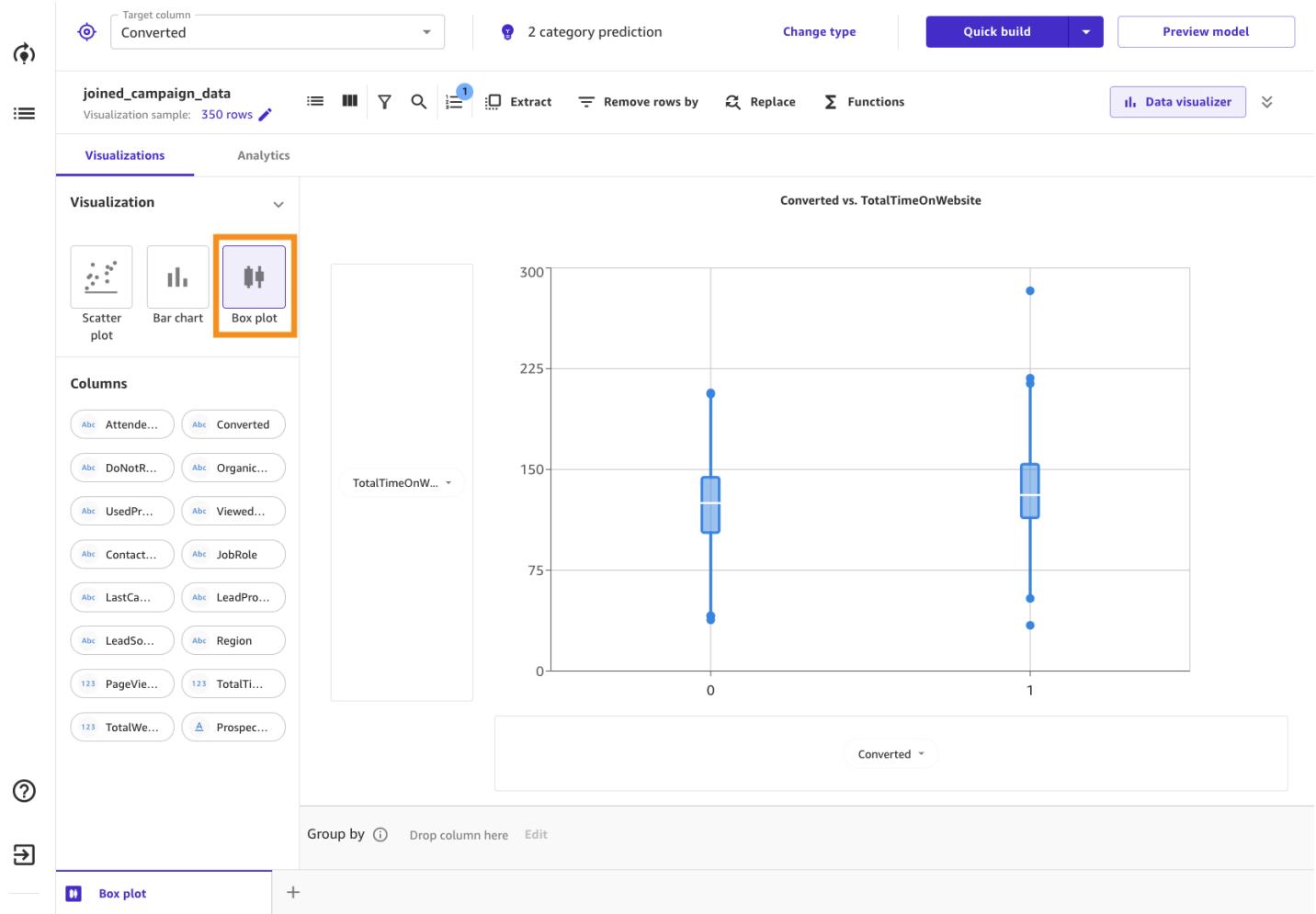
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



From our observations, we can determine that the dataset is fairly balanced. We want the data to be evenly distributed across true and false values so that the model isn't biased towards one value.

Now we are ready to build the model.

As we can see here that the accuracy is quite good (your model accuracy and column impact order may differ from above as it depends what sample dataset was picked by preview mode functionality) and this suggests that our dataset is ready to generate the model. Click on the arrow of "Quick build" button (as highlighted in the picture below in red colour) and you will see 2 options to build the model. First one is the **standard build** and second one is **quick build**. As name suggest "quick build" option provides speed over accuracy as building model takes around 2 to 15 minutes but you can't share quick build models. With standard build, it supersedes accuracy over speed and building model usually takes 2 to 4 hours. Standard build runs experiments using different combinations of hyperparameters and generates many models in the backend (using Sagemaker auto pilot functionality) and then picks the best model. You can share the model using standard build option.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

The screenshot shows the AWS SageMaker Canvas interface. At the top, there's a target column dropdown set to 'Converted' and a prediction type dropdown set to '2 category prediction'. Below this is a data preview for 'joined_campaign_data' with 10.0k rows. A 'Change type' button is highlighted with a blue border. To its right are 'Quick build' and 'Standard build' buttons. A tooltip for 'Quick build' is displayed, stating: 'Choose speed over accuracy. Building usually takes 2-15 minutes. You can't share quick build models.' Another tooltip for 'Standard build' is also visible, stating: 'Choose accuracy over speed. Building usually takes between 2-4 hours.' On the far right, there's a 'Data visualizer' button.

For this workshop first we build model using "quick build" and will predict the conversion of sample data using this model. Then we will generate another model using "standard build" option. As standard build takes around 1.5 hours to 2 hours to complete hence this portion of the workshop will be optional here.

Click on the "Quick build" button and it starts building the model. It takes around 5-10 minutes to complete:

The screenshot shows the 'Sales conversion model' view in SageMaker Canvas. The tab bar has 'Select', 'Build', 'Analyze' (which is selected), and 'Predict'. Below the tabs is a 'Model overview' section with the message: 'Your model is being created. Quick build usually takes 2-15 minutes. You can now leave this view.' To the right, there's a progress bar with a person icon working at a laptop. Below the progress bar, it says 'Generating column impact'. At the bottom, there are status indicators for 'joined_campaign_data' (Total columns: 16, Total rows: 10000), 'Converted' (2 category prediction), and a 'Close' button.

Once model build is complete you can see estimated accuracy and column impact list (with their SHAP values) as below:

SHAP is a game theoretic framework inspired by **shapley values** that provides local explanations for any model. SHAP has gained popularity in recent years, probably due to its strong theoretical basis. The SHAP package contains several algorithms that, when given a sample and model, derive the SHAP value for each of the model's input features. The SHAP value of a feature represents its contribution to the model's prediction.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

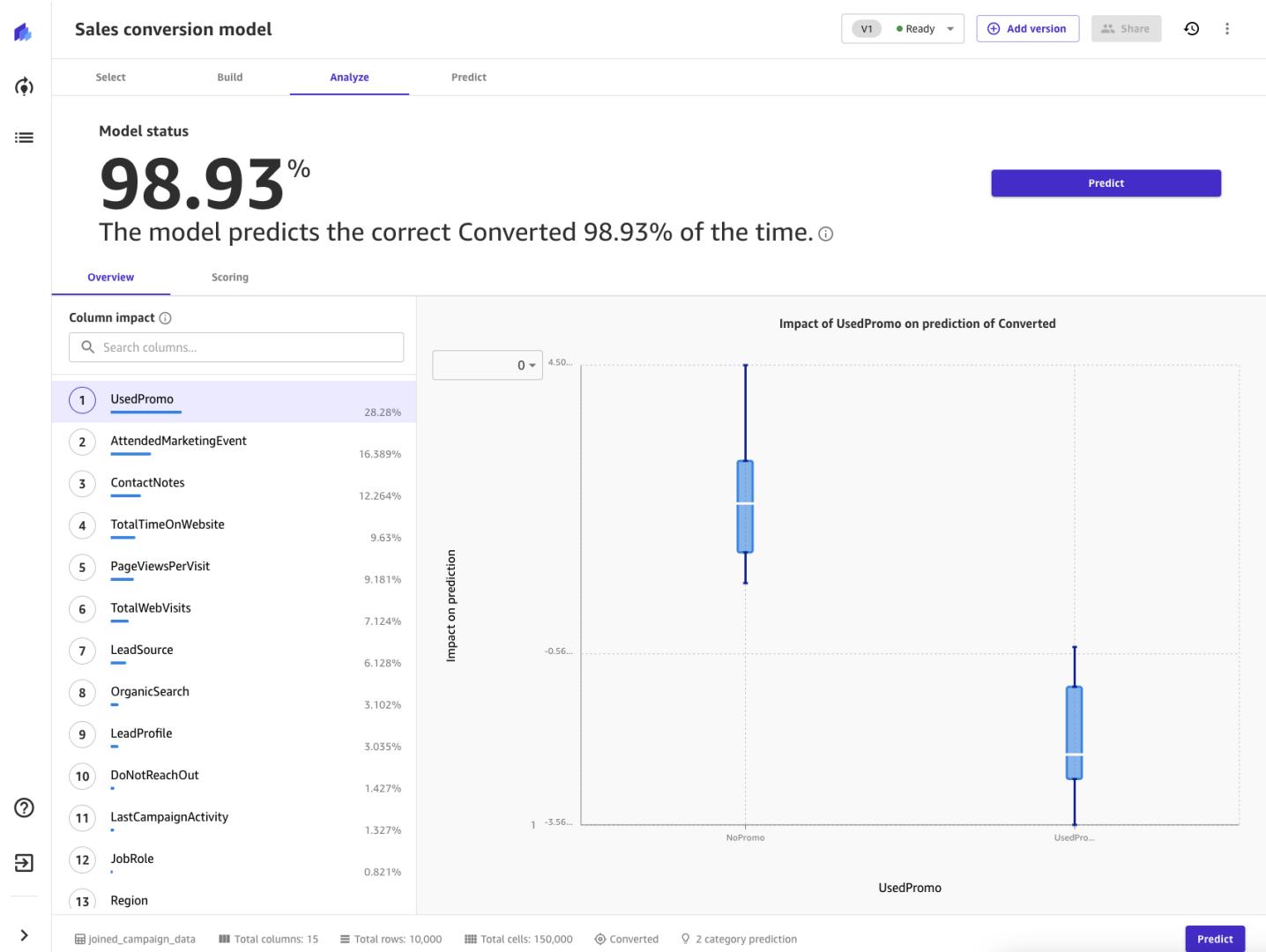
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



Evaluate the model

You evaluate how well your model performed on your data before you start using it to make predictions by using the following:

- **Column impact** is a percentage score that indicates how much weight a column has in making predictions in relation to the other columns. For a column impact of 25%, SageMaker Canvas weighs the prediction as 25% for the column and 75% for the other columns.
- **Scoring** is a section that shows visualizations and figures that you can use to get more insights into your model's performance beyond the overall accuracy metric. For a categorical prediction, you can see the predicted values in contrast to the actual values.

- **Advanced metrics** is a section that contains information that you can use for a deeper understanding of your model's performance.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker
JumpStart

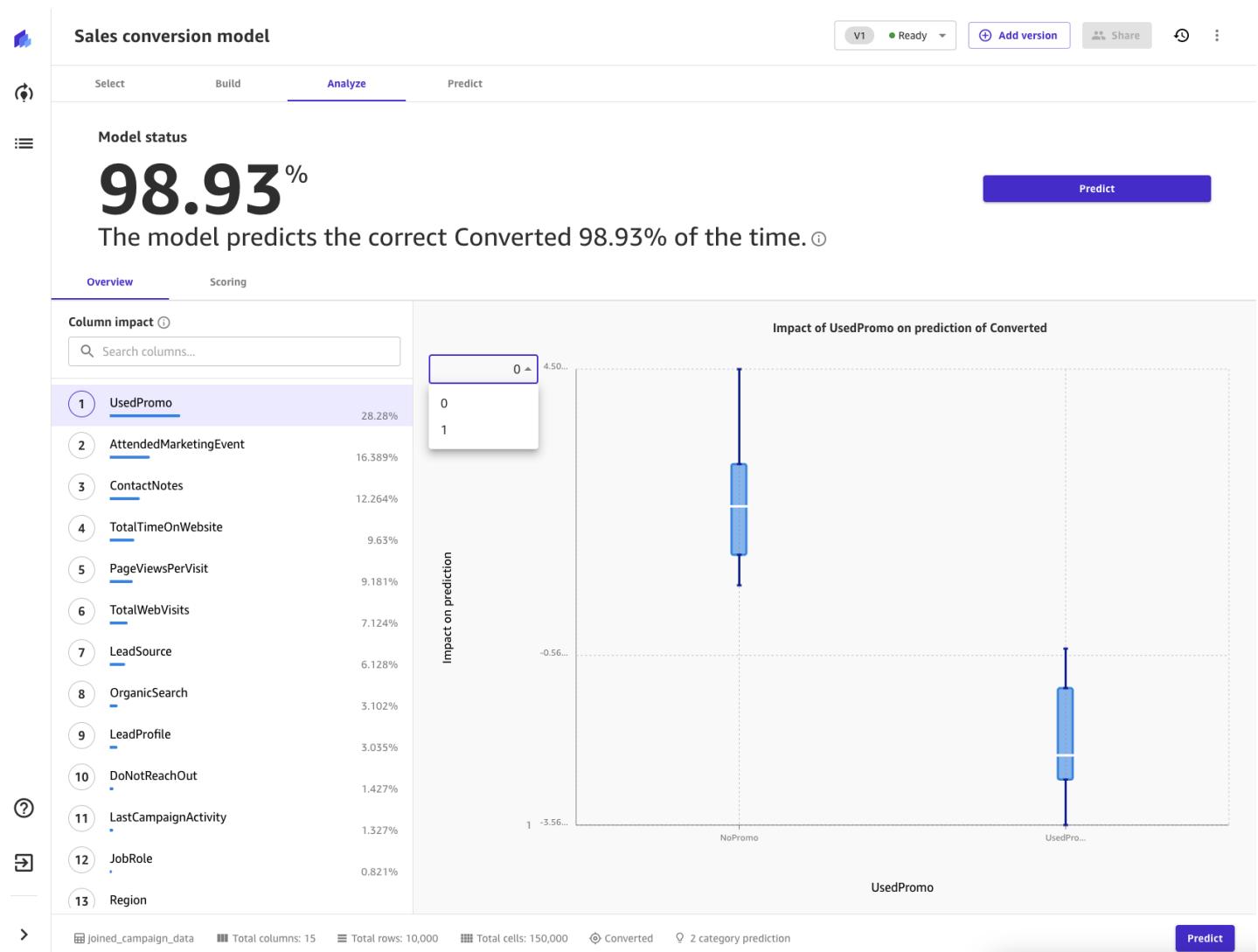
► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Scroll down through the "Column impact" order list that depicts which feature has high impact on the inference of target feature using this model. you can click on any column t and change the target value (0 or 1) to see the column impact:



Click on the "Scoring" tab (as highlighted in red colour bellow). Here you see on the right hand side, model accuracy insights are provided to gain more insights of the model.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

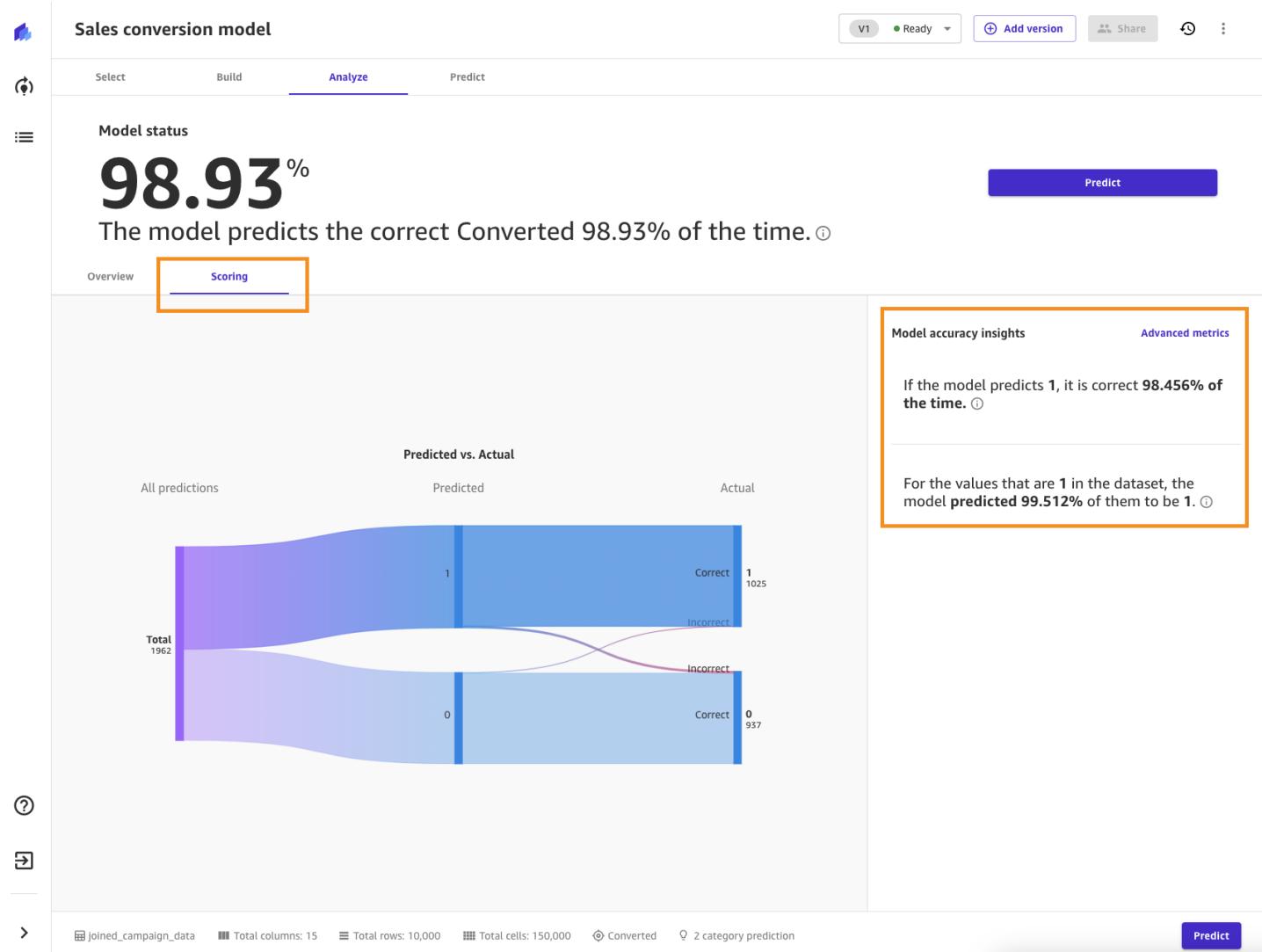
► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language



Above, you see that actual "No Conversion" records (0 value) were 979 only but model predicts 984 records as "No Conversion" and 1021 actual "Yes Conversion" records (1 value) were there but model predicted 1016 as "Yes Conversion". Based on these outcomes, SageMaker canvas generated Model accuracy insights.

Now click on the "Advanced metrics" link (link is there on the right hand side, under Model accuracy insights section). It opens below screen:

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

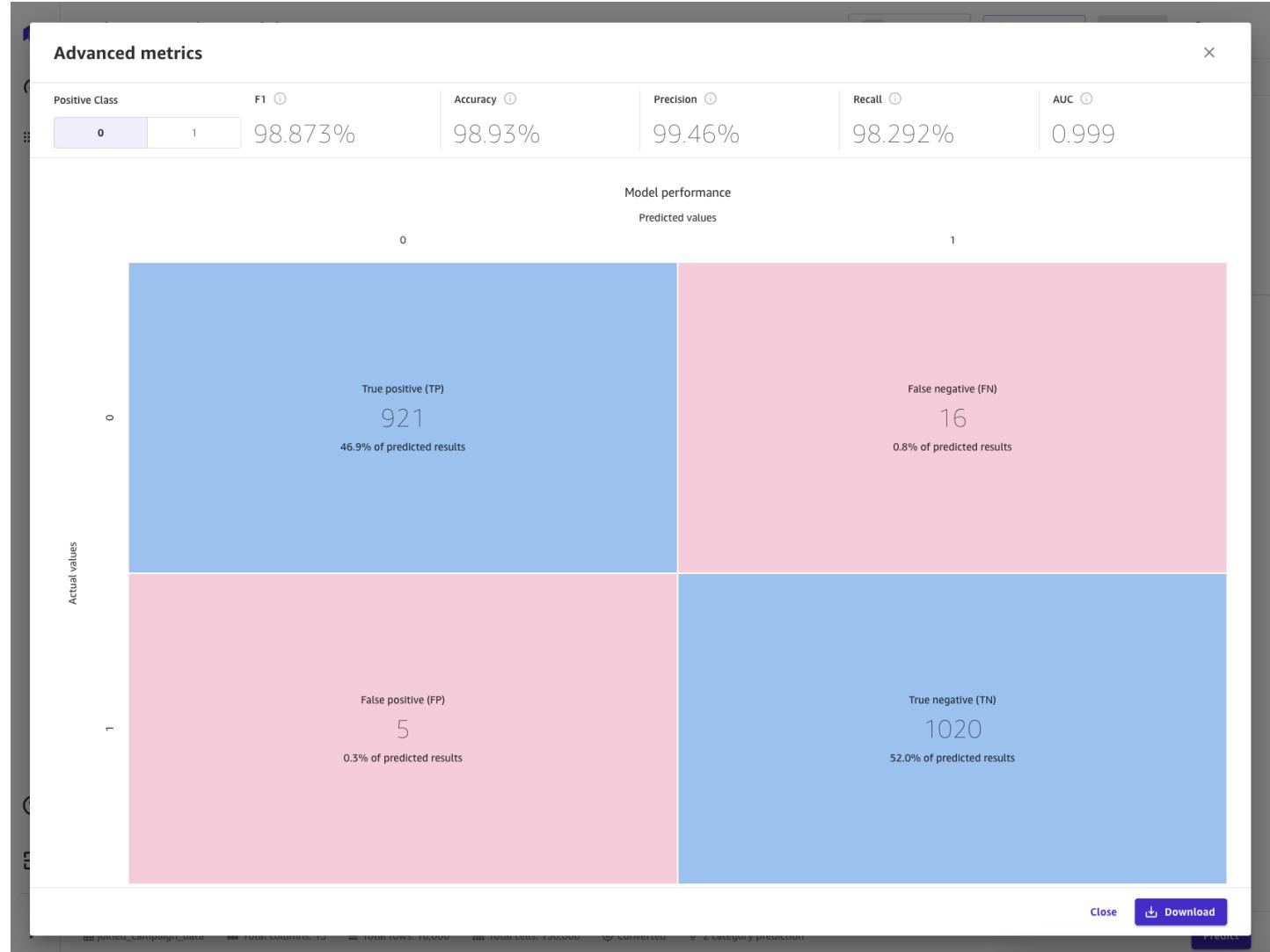
► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language



SageMaker Canvas uses confusion matrices to help you visualise when a model makes predictions correctly. As this is a binary classification problem hence advance metrics like F1 score, Accuracy, Precision, recall and AUC are shown above. Click on the close button to close above screen.

Make Prediction

You can do a single prediction or batch prediction using Sagemaker Canvas. We do a single prediction for this workshop. click on the "Predict" button and then select "Single Prediction" tab. Sagemaker canvas automatically generates a sample data and predicts the target feature (converted) for this data. As you can see here, it predicted "No" as converted target feature for this sample data.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

Sales conversion model

Select Build Analyze Predict

Predict target values

Batch prediction Single prediction

Modify values to predict Converted in real time.

Filter columns

Column	Feature importance	Value
LeadSource	24.415%	Cold Uutbound
PageViewsPerVisit	19.459%	1.931872658
UsedPromo	18.801%	NoPromo
AttendedMarketingEvent	11.24%	Attended
OrganicSearch	8.248%	N
DoNotReachOut	4.63%	False
TotalTimeOnWebsite	4.586%	125
LastCampaignActivity	3.601%	DownloadedMat...
LeadProfile	2.326%	Large
ContactNotes	1.076%	ContentPresent...
TotalWebVisits	0.794%	10
JobRole	0.538%	HR
ViewedAdvertisement	0.191%	N
Region	0.096%	APAC
ProspectID	0%	0009de69-00a3-463a-82b

Converted Prediction

No

Average prediction

Yes 1.627% ⓘ

No 98.373% ⓘ

Copy

Download

You can play around with this sample data and update the prediction after changing the data. like I changed "Usedpromo = UsedPromo" and "LastCampaignActivity = Web Visit" and hit the update button to predict the target feature again. This time it predicted "Yes" as converted target feature for this changed sample data.

▼ Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Sales conversion model

Select Build Analyze Predict

Predict target values

Batch prediction Single prediction

Modify values to predict Converted in real time.

Filter columns

Column	Feature importance	Value
LeadSource	24.415%	Cold Uutbound
PageViewsPerVisit	19.459%	1.931872658
UsedPromo	18.801%	UsedPromo
AttendedMarketingEvent	11.24%	Attended
OrganicSearch	8.248%	N
DoNotReachOut	4.63%	False
TotalTimeOnWebsite	4.586%	125
LastCampaignActivity	3.601%	DownloadedMat...
LeadProfile	2.326%	Large
ContactNotes	1.076%	ContentPresent...
TotalWebVisits	0.794%	10
JobRole	0.538%	HR
ViewedAdvertisement	0.191%	N
Region	0.096%	APAC
ProspectID	0%	0009de69-00a3-463a-82b

Converted Prediction

You made 1 change to the values. Update for a new prediction and feature importance metrics.

Update

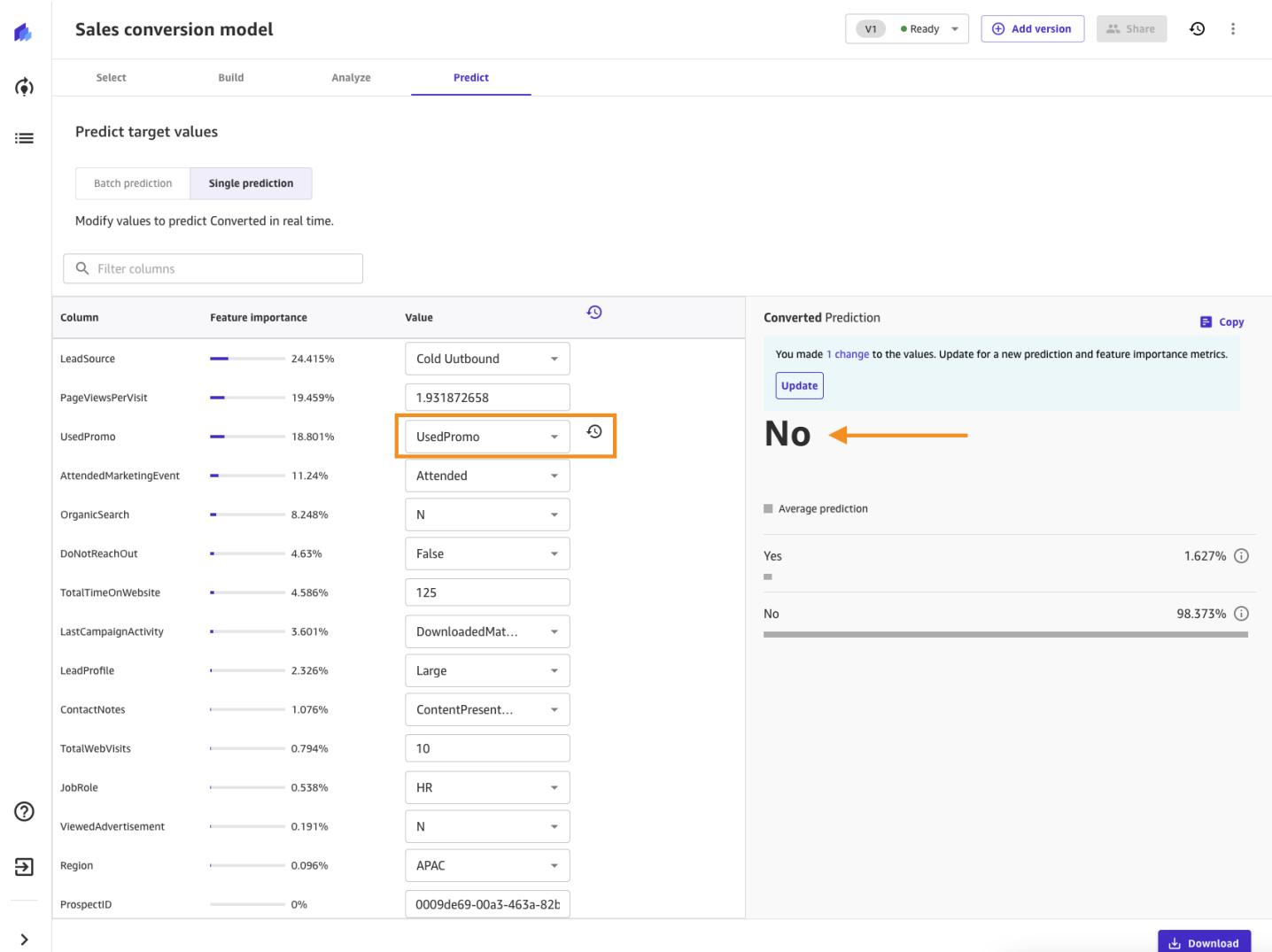
No

Average prediction

Yes 1.627% ⓘ

No 98.373% ⓘ

Download



Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Sales conversion model

Select Build Analyze Predict

Predict target values

Batch prediction Single prediction

Modify values to predict Converted in real time.

Filter columns

Column	Feature importance	Value
LeadSource	7.818%	Cold Uutbound
PageViewsPerVisit	35.715%	1.931872658
UsedPromo	34.216%	UsedPromo
AttendedMarketingEvent	3.325%	Attended
OrganicSearch	2.869%	N
DoNotReachOut	2.985%	False
TotalTimeOnWebsite	9.017%	125
LastCampaignActivity	0.446%	DownloadedMat...
LeadProfile	0.801%	Large
ContactNotes	2.234%	ContentPresent...
TotalWebVisits	0.285%	10
JobRole	0.144%	HR
ViewedAdvertisement	0.002%	N
Region	0.142%	APAC
ProspectID	0%	0009de69-00a3-463a-82b

Converted Prediction

Yes

New prediction Average prediction

Yes 73.215% ⓘ

No 26.785% ⓘ

Copy Download

OPTIONAL: Build another version of the Model using Standard build option

SageMaker Canvas supports building multiple version of a model. Now we create a new version of the above model but this time we train it using the "Standard build" option.

Click on the "Add version" button and it will create a V2 (in draft status) of this model:

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

The screenshot shows the AWS SageMaker Canvas interface. At the top, there's a header bar with 'Import bookmarks...' and 'Bookmarks Toolbar' options. Below the header, there are four tabs: 'Select', 'Build', 'Analyze', and 'Predict', with 'Predict' being the active tab. The main area is titled 'Sales conversion model' and contains a sub-section 'Predict target values'. At the bottom right of the main area, there's a dropdown menu showing two versions: 'V1' (Ready, Created Feb 16, 2022 3:47 PM) and 'V2' (Draft, Created Feb 16, 2022 7:03 PM). A checkmark is next to 'V1'. To the right of the dropdown, there are 'Add version', 'Share', and more options buttons.

select the draft V2 version in the dropdown and below screen will appear. Click on the "Go to Build" button:

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Sales conversion model

V2 Draft Add version Share

Select Build Analyze Predict

Prediction

Build a model before you run predictions

Go to Build

Below screen appears. Now, select "Go to select" button and select "joint_Campaign_Dataset" option and click on select dataset button to select the dataset,:;

The screenshot shows the Amazon SageMaker Canvas interface. At the top, there are buttons for 'V2', 'Draft', 'Add version', 'Share', and more. Below that, tabs for 'Select', 'Build', 'Analyze', and 'Predict' are present, with 'Predict' being the active tab. On the right side, there's a graphic of a person thinking with a speech bubble containing blue dots. Below the graphic, the word 'Prediction' is written. Further down, a message says 'Build a model before you run predictions' with a 'Go to Build' button highlighted by an orange box. To the left of the main area, a sidebar contains a list of lab sections: 'Lab3c. Bring your own Container', 'Lab 4. Autopilot, Debugger and Model Monitor', 'Lab 5. Bias and Explainability', 'Lab 6. SageMaker Pipelines', 'Lab 7. Real Time ML inference on Streaming Data', 'Lab 8. Build ML Model with No Code Using Sagemaker Canvas' (which is bolded), 'Lab 9. Amazon SageMaker JumpStart', 'Lab 10. ML Governance Tools for Amazon SageMaker', and 'Lab 11. SageMaker Notebook Instances'. At the very bottom, there's a 'Content preferences' section with a 'Language' dropdown.

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Import bookmarks... Bookmarks Toolbar

Sales conversion model

V2 • Draft Add version Share ...

Select Build Analyze Predict

Build

Select dataset to review and build.

Go to Select Close

Now build model screen appear (like shown below) and de-select the ProspectID feature as shown below (as we don't want to train our model with this feature):

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language

Sales conversion model

Select Build Analyze Predict

Validate your data
It might take several minutes, depending on the dataset size.

Validate data

Select a column to predict
Choose the target column. The model that you build predicts values for the column that you select.

Target column: Converted

Value distribution

Model type
SageMaker Canvas automatically recommends the appropriate model type for your analysis.

2 category prediction
Your model classifies Converted into two categories.

Quick build

Preview model

joined_campaign_data
Full dataset: 10.0k rows

Column name	Data type	Missing	Mismatched	Unique	Mean / Mode	Correla
ViewedAdvertisement	Binary	0.00% (0)	0.00% (0)	2	N	-0.016
UsedPromo	Binary	0.00% (0)	0.00% (0)	2	NoPromo	0.528
TotalWebVisits	Numeric	0.00% (0)	0.00% (0)	19	10	0.131
TotalTimeOnWebsite	Numeric	0.00% (0)	0.00% (0)	240	125	0.189
Region	Categorical	0.00% (0)	0.00% (0)	4	APAC	N/A
ProspectID	Text	0.00% (0)	0.00% (0)	10,000	0009de69-00a3-463a...	N/A
PageViewsPerVisit	Numeric	0.00% (0)	0.00% (0)	10,000	1.93	0.172
OrganicSearch	Binary	0.00% (0)	0.00% (0)	2	N	0.049
LeadSource	Categorical	0.00% (0)	0.00% (0)	5	Cold Uutbound	N/A
LeadProfile	Categorical	0.00% (0)	0.00% (0)	3	Large	N/A
LastCampaignActivity	Categorical	0.00% (0)	0.00% (0)	4	DownloadedMaterial	N/A
JobRole	Categorical	0.00% (0)	0.00% (0)	7	HR	N/A
DoNotReachOut	Binary	0.00% (0)	0.00% (0)	2	False	0.046
Converted	Target	0.00% (0)	0.00% (0)	2	1	--
ContactNotes	Categorical	0.00% (0)	0.00% (0)	4	ContentPresentation	N/A

Model recipe

Drop column ProspectID

Total columns: 15 Total rows: 10,000 Total cells: 150,000 Show dropped columns

Now click on the dropdown of "Quick build" button and select "Standard build" option. Now click on "Standard build" button:

Sales conversion model

Select Build Analyze Predict

Validate your data
It might take several minutes, depending on the dataset size.

Select a column to predict
Choose the target column. The model that you build predicts values for the column that you select.

Target column: Converted

Value distribution: 1 (blue bar) 0 (blue bar)

Model type
SageMaker Canvas automatically recommends the appropriate model type for your analysis.

2 category prediction
Your model classifies Converted into two categories.

Quick build
Choose speed over accuracy. Building usually takes between 2-4 hours.

Standard build
Choose accuracy over speed. Building usually takes between 2-15 minutes. You can't share quick build models.

Data visualizer

joined_campaign_data
Full dataset: 10.0k rows

Column name	Data type	Missing	Mismatched	Unique	Mean / Mode	Correla
ViewedAdvertisement	Binary	0.00% (0)	0.00% (0)	2	N	-0.016
UsedPromo	Binary	0.00% (0)	0.00% (0)	2	NoPromo	0.528
TotalWebVisits	Numeric	0.00% (0)	0.00% (0)	19	10	0.131
TotalTimeOnWebsite	Numeric	0.00% (0)	0.00% (0)	240	125	0.189
Region	Categorical	0.00% (0)	0.00% (0)	4	APAC	N/A
ProspectID	Text	0.00% (0)	0.00% (0)	10,000	0009de69-00a3-463a...	N/A
PageViewsPerVisit	Numeric	0.00% (0)	0.00% (0)	10,000	1.93	0.172
OrganicSearch	Binary	0.00% (0)	0.00% (0)	2	N	0.049
LeadSource	Categorical	0.00% (0)	0.00% (0)	5	Cold Uutbound	N/A
LeadProfile	Categorical	0.00% (0)	0.00% (0)	3	Large	N/A
LastCampaignActivity	Categorical	0.00% (0)	0.00% (0)	4	DownloadedMaterial	N/A
JobRole	Categorical	0.00% (0)	0.00% (0)	7	HR	N/A
DoNotReachOut	Binary	0.00% (0)	0.00% (0)	2	False	0.046
Converted	Target	0.00% (0)	0.00% (0)	2	1	--
ContactNotes	Categorical	0.00% (0)	0.00% (0)	4	ContentPresentation	N/A

Total columns: 15 Total rows: 10,000 Total cells: 150,000 Show dropped columns

Model recipe

Drop column ProspectID

Now wait for model build completion as it will take around 1 hour and 45 minutes or less:

Lab3c. Bring your own Container

Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

Lab 5. Bias and Explainability

Bias and Explainability- Tabular Data

Lab 6. SageMaker Pipelines

Option 1: Train Pipeline (SageMaker Pipelines)

Option 2: Batch Inference Pipeline (SageMaker Pipelines)

Option 3: SageMaker Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

Lab 9. Amazon SageMaker JumpStart

Lab 10. ML Governance Tools for Amazon SageMaker

Lab 11. SageMaker Notebook Instances

Content preferences

Language

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference
on Streaming Data

**Lab 8. Build ML Model with
No Code Using Sagemaker
Canvas**

► Lab 9. Amazon SageMaker
JumpStart

► Lab 10. ML Governance Tools
for Amazon SageMaker

► Lab 11. SageMaker Notebook
Instances

▼ Content preferences

Language

Sales conversion model

Select Build Analyze Predict

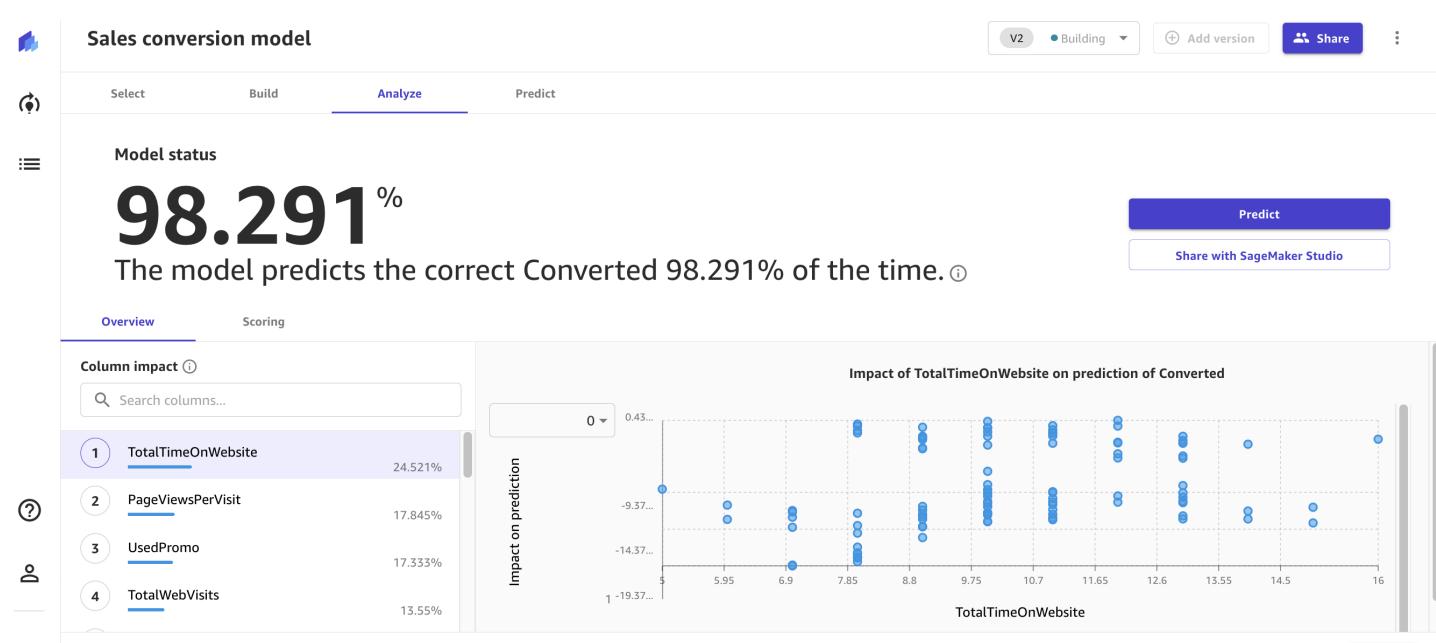
Model overview

Your model is being created. Standard build usually takes between 2–4 hours. You can now leave this view.

Time elapsed: 12 sec Expected build time: 1 hr 45 min Build type: Standard build Detailed progress: Starting training job



Once the model build is complete, you will see accuracy and column impact list as below:



Evaluate this model same as the previous model; by navigating to "Scoring" and "Advance Metrics" sections. you can also now see that version of this model (V2) is showing in "Ready" status. You can use this model to do either batch prediction or single prediction.

[Previous](#)[Next](#)

Lab3c. Bring your own Container

▼ Lab 4. Autopilot, Debugger and Model Monitor

Autopilot

Debugger

Model Monitor

▼ Lab 5. Bias and Explainability

Bias and Explainability-
Tabular Data

▼ Lab 6. SageMaker Pipelines

Option 1: Train Pipeline
(SageMaker Pipelines)

Option 2: Batch Inference
Pipeline (SageMaker
Pipelines)

Option 3: SageMaker
Projects

Lab 7. Real Time ML inference on Streaming Data

Lab 8. Build ML Model with No Code Using Sagemaker Canvas

► Lab 9. Amazon SageMaker JumpStart

► Lab 10. ML Governance Tools for Amazon SageMaker

► Lab 11. SageMaker Notebook Instances

▼ Content preferences

Language