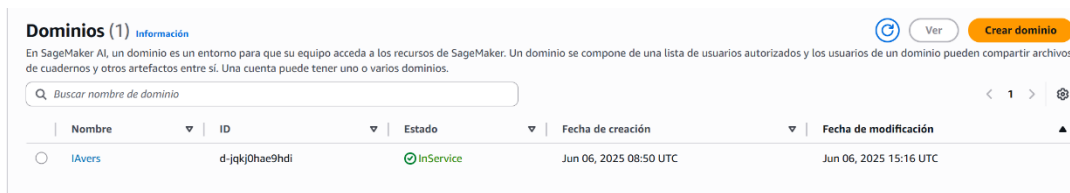


# 1. Configuración del entorno

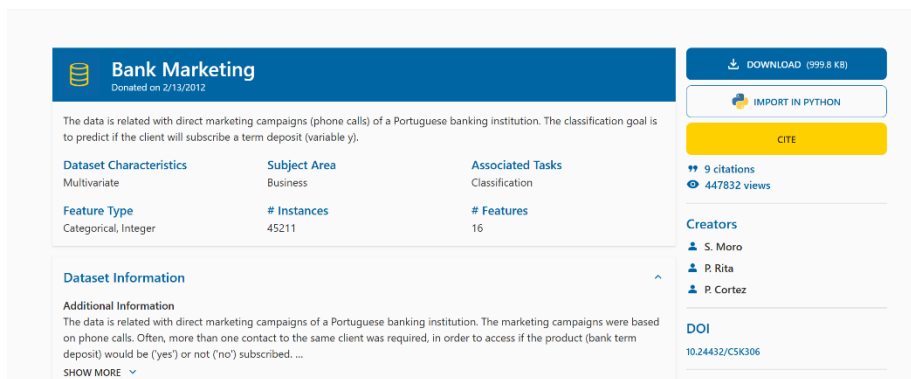
## 1. Se crea el dominio con nombre “IAvers”



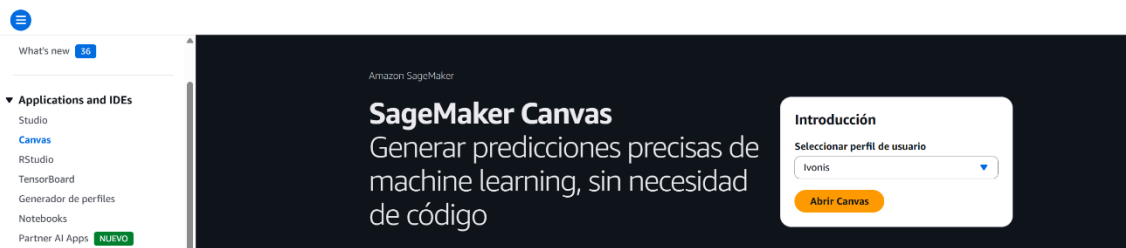
## 2. Se crea un perfil de usuario “Ivonis”



## 3. Se descarga el Dataset Bank Marketing.

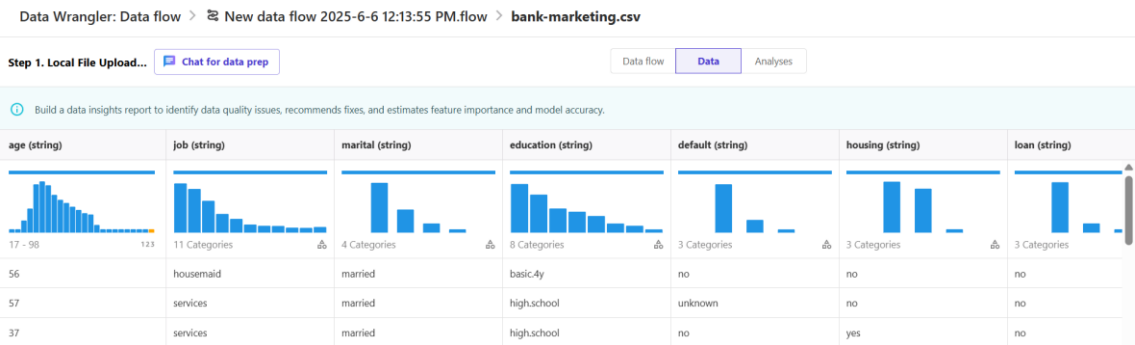


## 4. Se procede a entrar a SageMaker Canvas con el perfil que se ha creado.



## 2. Preparación del Dataset

5. Se carga el dataset en Data Wrangler como CSV tabular. En este caso mediante Local upload, aunque bien podría ser usando Amazon S3. A continuación se mostrará una previsualización del dataset.



6. Se genera un reporte rápido con el preset “Data quality and insights report: Data Basics” planteando un problema de clasificación para predecir “y” como “sí” o “no”.

### Data quality and insights report: Data Basics

Target column	Type	Dataset
y	Classification	bank-marketing.csv

7. A continuación se quitan algunas columnas que no aportarán precisión al modelo, como identificadores o campos irrelevantes que podrían introducir ruido.

### 3. Drop column

Move, drop, duplicate or rename columns in the dataset. [Learn more.](#)

Transform \* ⓘ

Drop column X ▼

Columns to drop \*

emp.var.rate X  
cons.price.idx X  
cons.conf.idx X  
nr.employed X pdays X  
+ 3 items selected

### 3. Entrenamiento del modelo

8. Se crea un nuevo modelo con nombre “Act9” con clasificación binaria con métrica objetivo F1 como objetivo al tratarse de un problema con clases desbalanceadas. Parámetros ajustados: máximo de horario de 30’ y método Standard, que hace varias pruebas de entrenamiento y prioriza la precisión.

The screenshot shows the 'Build' tab of the Data Science Canvas interface. At the top, there's a navigation bar with 'Select', 'Build', 'Analyze', 'Predict', and 'Deploy'. Below it, a blue banner indicates 'Preprocessing your dataset. This can take a few minutes. You can navigate away from this page. This won't interrupt the process. Learn more.' The main area is divided into two sections: 'Select a column to predict' and 'Model type'. In the 'Select a column to predict' section, the target column 'y' is selected. In the 'Model type' section, '2 category prediction' is chosen. Below these sections, a table displays the dataset statistics for 'Dataset 20250606\_104641'.

Column name	Data type	Feature type	Missing	Mismatched	Unique	Mode
y	Text	Binary	0.00% (0)	0.00% (0)	2	no
previous	Numeric	-	0.00% (0)	0.00% (0)	1	0
previous	Text	Categorical	0.00% (0)	0.00% (0)	1	remediated
plays	Numeric	-	0.00% (0)	0.00% (0)	1	999
reemployed	Numeric	Binary	0.00% (0)	0.00% (0)	2	1228.1
month	Text	Categorical	0.00% (0)	0.00% (0)	4	may
marital	Text	Categorical	0.00% (0)	0.00% (0)	4	married
loan	Text	Categorical	0.00% (0)	0.00% (0)	5	no
job	Text	Categorical	0.00% (0)	0.00% (0)	12	blue-collar
housing	Text	Categorical	0.00% (0)	0.00% (0)	3	no

At the bottom, summary statistics are shown: Total columns: 21, Total rows: 43,180, Total cells: 864,948. A 'Show dropped columns' checkbox is also present.

### 4. Evaluación del modelo.

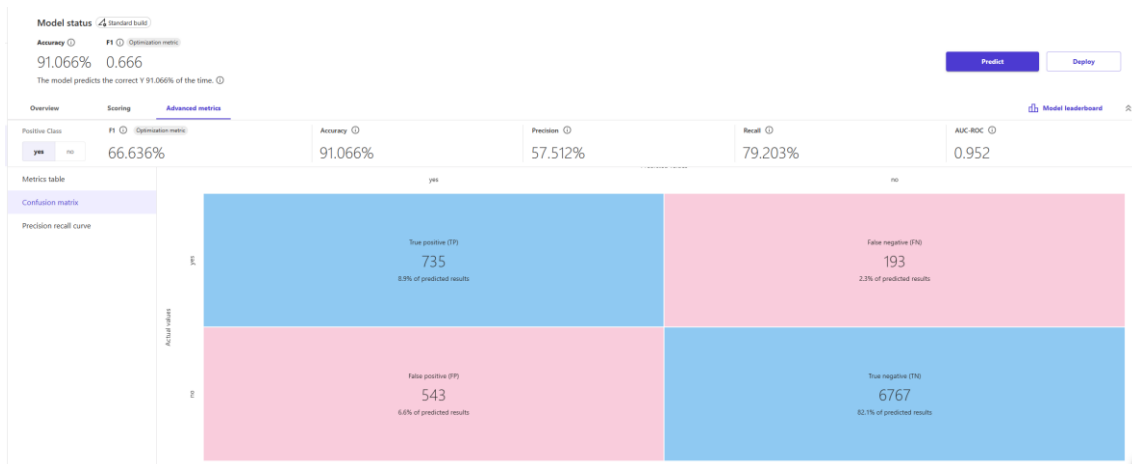
9. En la sección “Overview” se puede ver que el modelo predice correctamente aproximadamente el 91% de las veces, lo cual está bien.

The screenshot shows the 'Analyze' tab of the Data Science Canvas interface. The top navigation bar includes 'Select', 'Build', 'Analyze', 'Predict', and 'Deploy'. The 'Model status' section shows 'Standard built' and 'Accuracy 91.066%' with an 'F1 0.666' score. Below this, it states 'The model predicts the correct Y 91.066% of the time.' The 'Overview' section is active, displaying a 'Column impact' table and a scatter plot titled 'Impact of duration on prediction of y'.

Column	Impact
1 duration	47.094%
2 euribor3m	14.000%
3 reemployed	7.754%
4 emp.var.rate	6.993%
5 month	5.154%
6 cons.price.idx	2.811%
7 campaign	2.771%
8 contact	1.943%
9 cons.conf.idx	1.802%

The scatter plot shows the relationship between 'duration' (x-axis) and 'impact on prediction' (y-axis). The data points are blue dots, showing a clear downward trend as duration increases, indicating that longer durations have a lower impact on the prediction.

10. En la sección “Advanced metrics” se encuentran otras métricas como la matriz de confusión para continuar evaluando.



## 5. Generación de inferencias

11. Se selecciona la base de datos sobre la cual queremos generar la predicción, y a continuación se puede elegir entre “Batch” o “Single” prediction.

The screenshot shows the 'Predict' tab of a machine learning interface. The 'Predict target values' section has a 'Single prediction' button selected. Below this, there is a search bar for 'Filter columns'. A table lists the input features and their values:

Column	Value
age	35
job	blue-collar
marital	married
education	high.school
default	no
housing	no
loan	no
contact	telephone

On the right, a dropdown menu is open, showing options: 'Add to Model Registry', 'Send to QuickSight', 'Deploy', 'View Notebook', 'Rename model', and 'Delete model'. Below the table, the predicted value is 'No' with a probability of 99.961%. The 'Download prediction' button is at the bottom right.

## 6. Despliegue

12. Con esto se finaliza el ciclo completo de despliegue de un modelo en SageMaker Canvas, desde la carga de datos hasta la generación de predicciones, permitiendo obtener valor en producción sin escribir una sola línea de código.

### Create Deployment ×

Deploy your model to a SageMaker endpoint so that you can make predictions from outside of the Canvas application, test and monitor your model to proactively detect issues such as model drift.

**Selected model version**

Act9

V1 ✓ Ready Created: 06-06-2025-2:05 PM

**Deployment type**

Real-time ⓘ

**Deployment name**

Deployment name  
Deplyment1

**Instance type** ⓘ [Learn about pricing](#) ↗

ml.m5.12xlarge Recommended

**Instance count** ⓘ

2

Cancel Deploy 💬