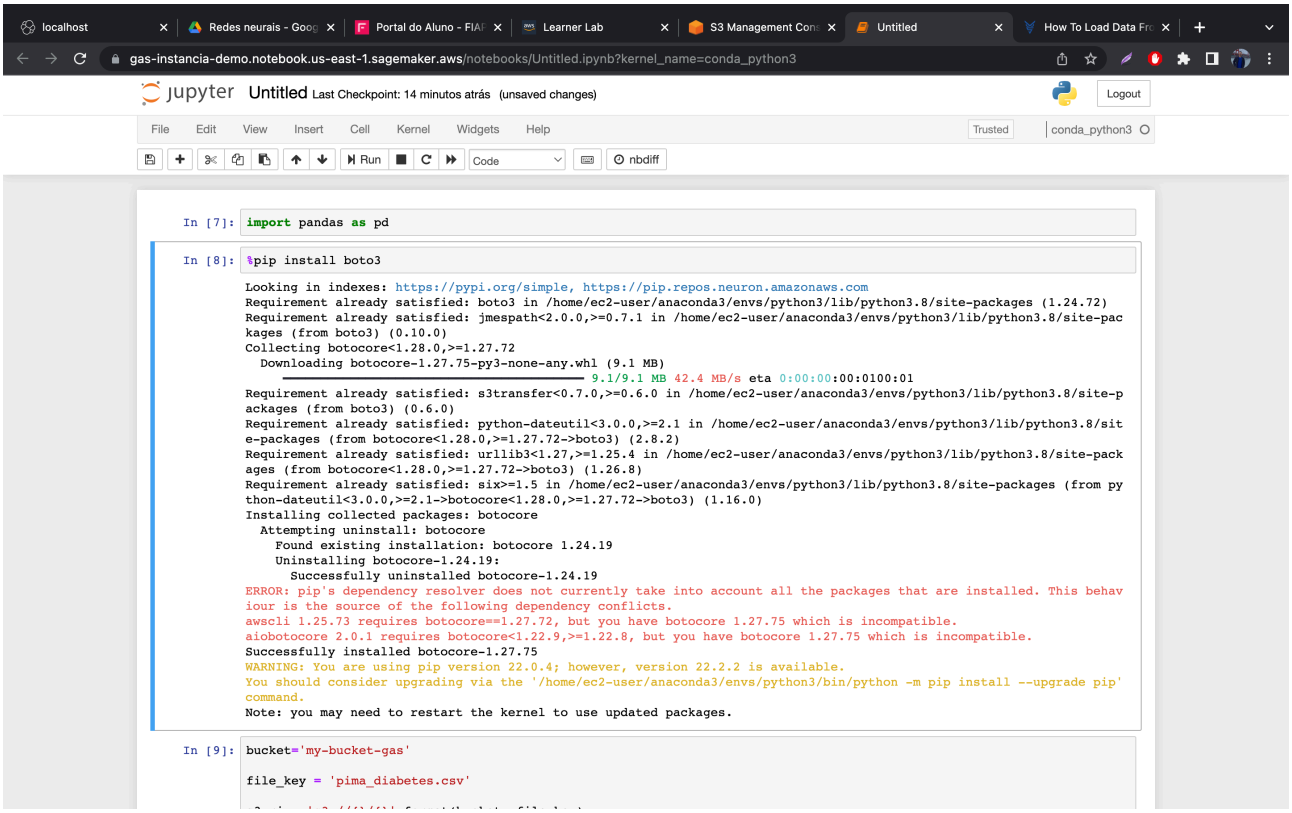
Plataformas e Serviços Cognitivos - Fernando Timóteo Fernandes

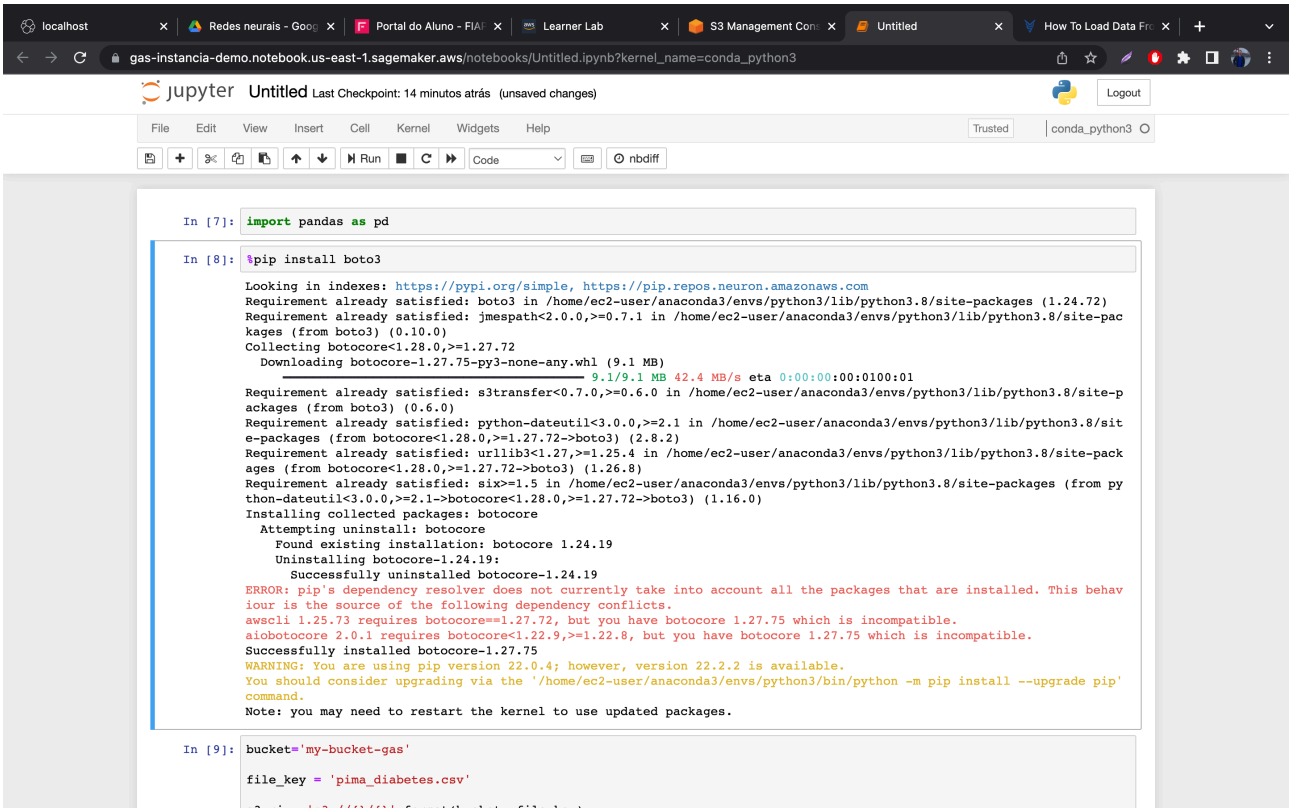
Integrantes do Sprint 2 - Grupo FeelingWhat:

Henrico Nardelli Bela - RM 95985
Guilherme Antonio Silva - RM 95044
Felype Nunes De Souza - RM 96232
Rafael Pereira da Silva - RM 94341

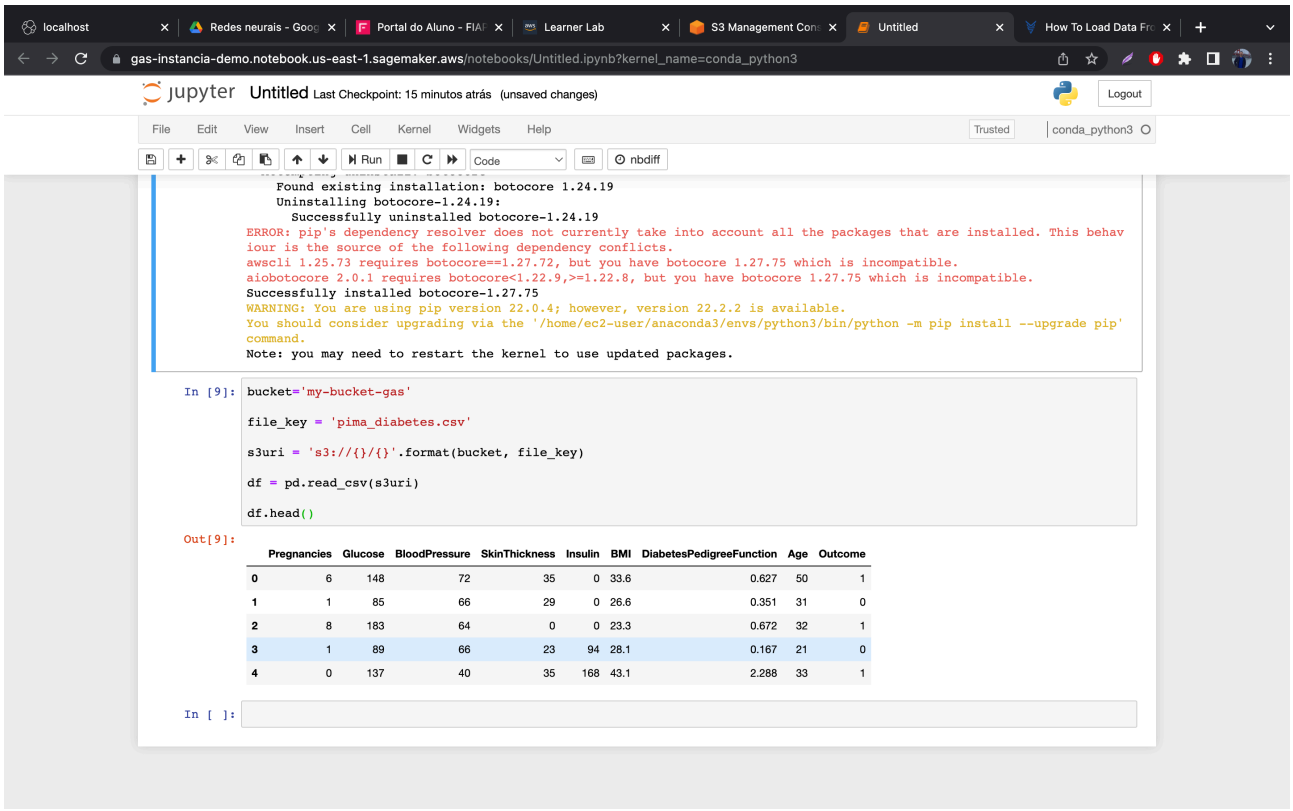
1) Crie uma instância no Jupyter Notebook:



2) Carregue o arquivo CSV (pima_diabetes.csv da pasta challenge4)



3)Exiba as informações básicas usando o pandas (ex: describe, shape, info)



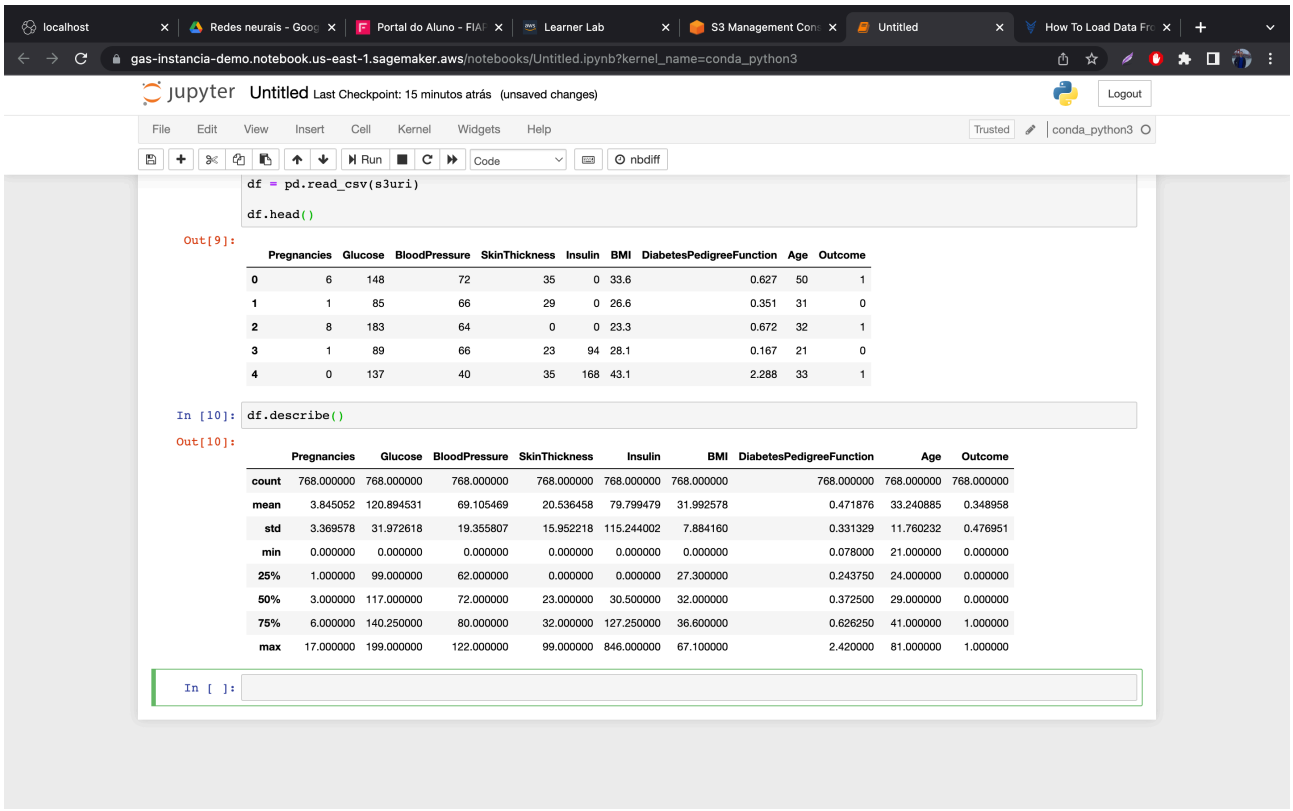
```
  Found existing installation: botocore 1.24.19
  Uninstalling botocore-1.24.19:
    Successfully uninstalled botocore-1.24.19
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behav
iour is the source of the following dependency conflicts.
awscli 1.25.73 requires botocore==1.27.72, but you have botocore 1.27.75 which is incompatible.
aiobotocore 2.0.1 requires botocore<1.22.9,>=1.22.8, but you have botocore 1.27.75 which is incompatible.
Successfully installed botocore-1.27.75
WARNING: You are using pip version 22.0.4; however, version 22.2.2 is available.
You should consider upgrading via the '/home/ec2-user/anaconda3/envs/python3/bin/python -m pip install --upgrade pip'
command.
Note: you may need to restart the kernel to use updated packages.
```

In [9]:
```
bucket='my-bucket-gas'

file_key = 'pima_diabetes.csv'

s3uri = 's3://{}/{}'.format(bucket, file_key)

df = pd.read_csv(s3uri)

df.head()
```

Out[9]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

In [ ]:



```
df = pd.read_csv(s3uri)

df.head()
```

Out[9]:

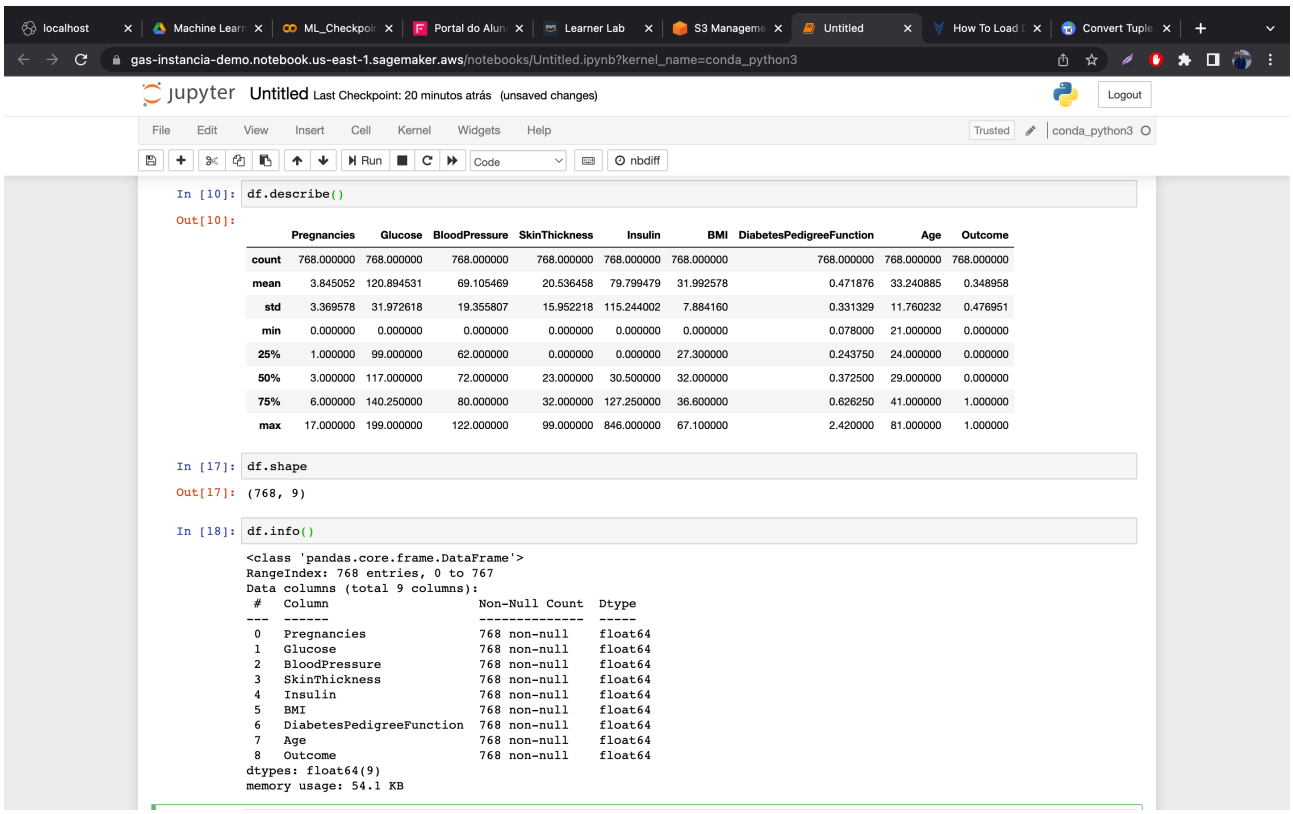
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

In [10]:
```
df.describe()
```

Out[10]:

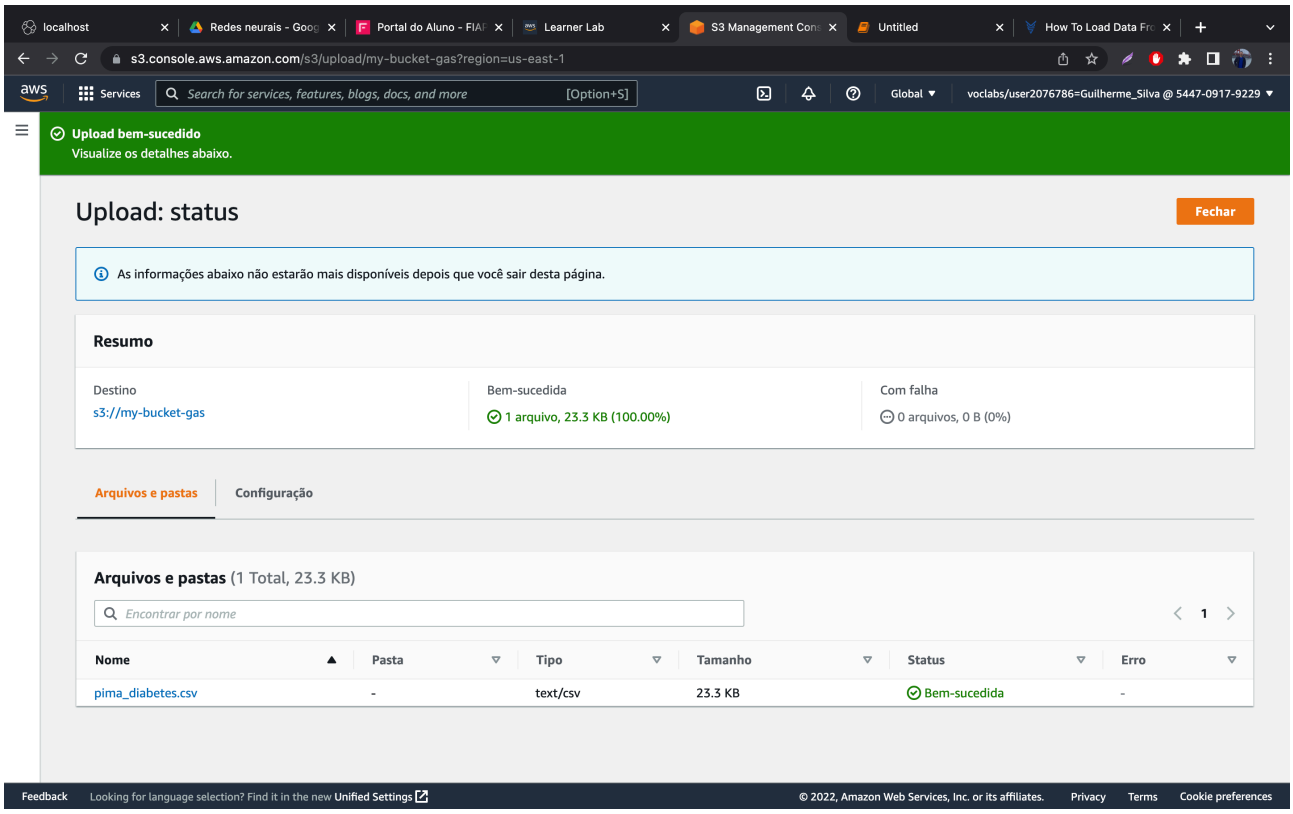
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.799479	31.992578	0.471876	33.240885	0.348958
std	3.369578	31.972618	19.355807	15.952218	115.244002	7.884160	0.331329	11.760232	0.476951
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.078000	21.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000	0.243750	24.000000	0.000000
50%	3.000000	117.000000	72.000000	23.000000	30.500000	32.000000	0.372500	29.000000	0.000000
75%	6.000000	140.250000	80.000000	32.000000	127.250000	36.600000	0.626250	41.000000	1.000000
max	17.000000	199.000000	122.000000	99.000000	846.000000	67.100000	2.420000	81.000000	1.000000

In [ ]:

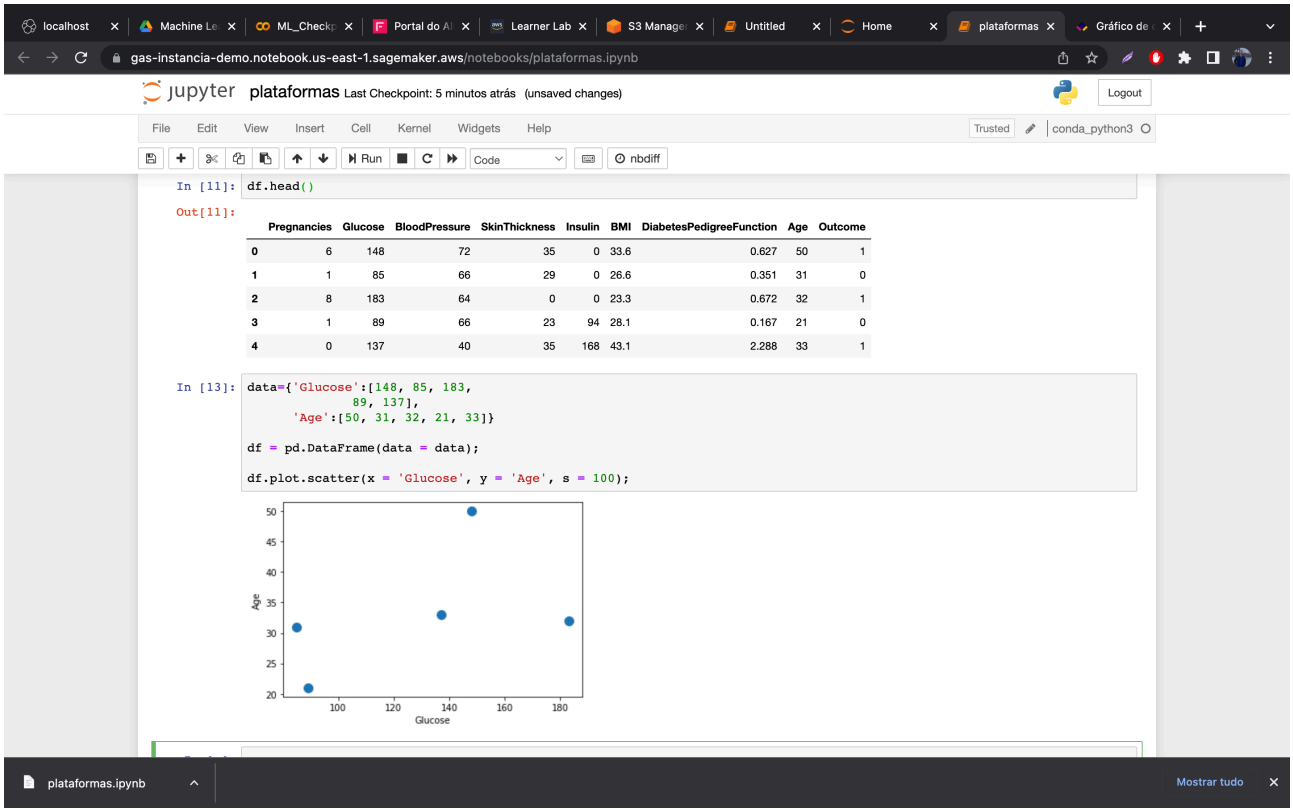


	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.799479	31.992578	0.471876	33.240885	0.348958
std	3.369578	31.972618	19.355807	15.952218	115.244002	7.884160	0.331329	11.760232	0.476951
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.078000	21.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000	0.243750	24.000000	0.000000
50%	3.000000	117.000000	72.000000	23.000000	30.500000	32.000000	0.372500	29.000000	0.000000
75%	6.000000	140.250000	80.000000	32.000000	127.250000	36.600000	0.626250	41.000000	1.000000
max	17.000000	199.000000	122.000000	99.000000	846.000000	67.100000	2.420000	81.000000	1.000000

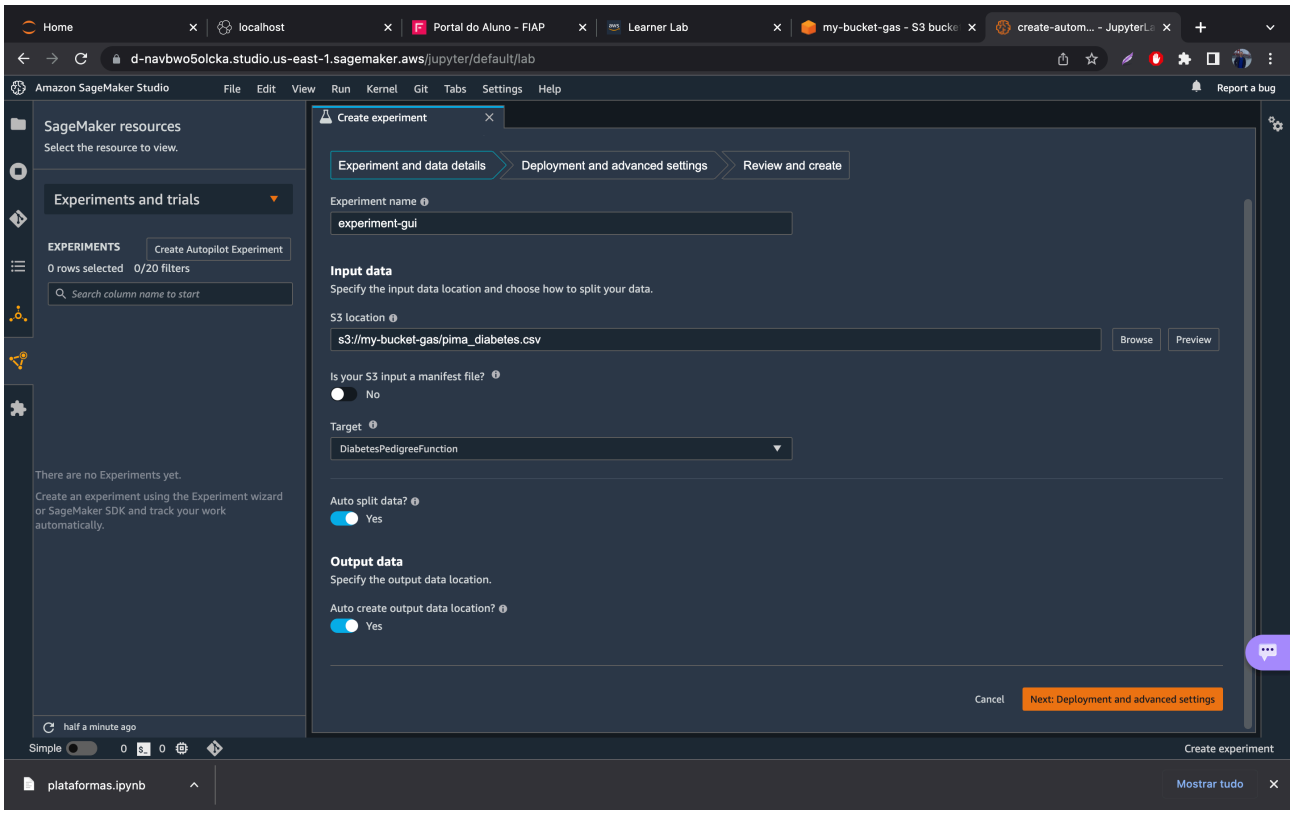
4) Arquivo adicionado ao bucket no s3:

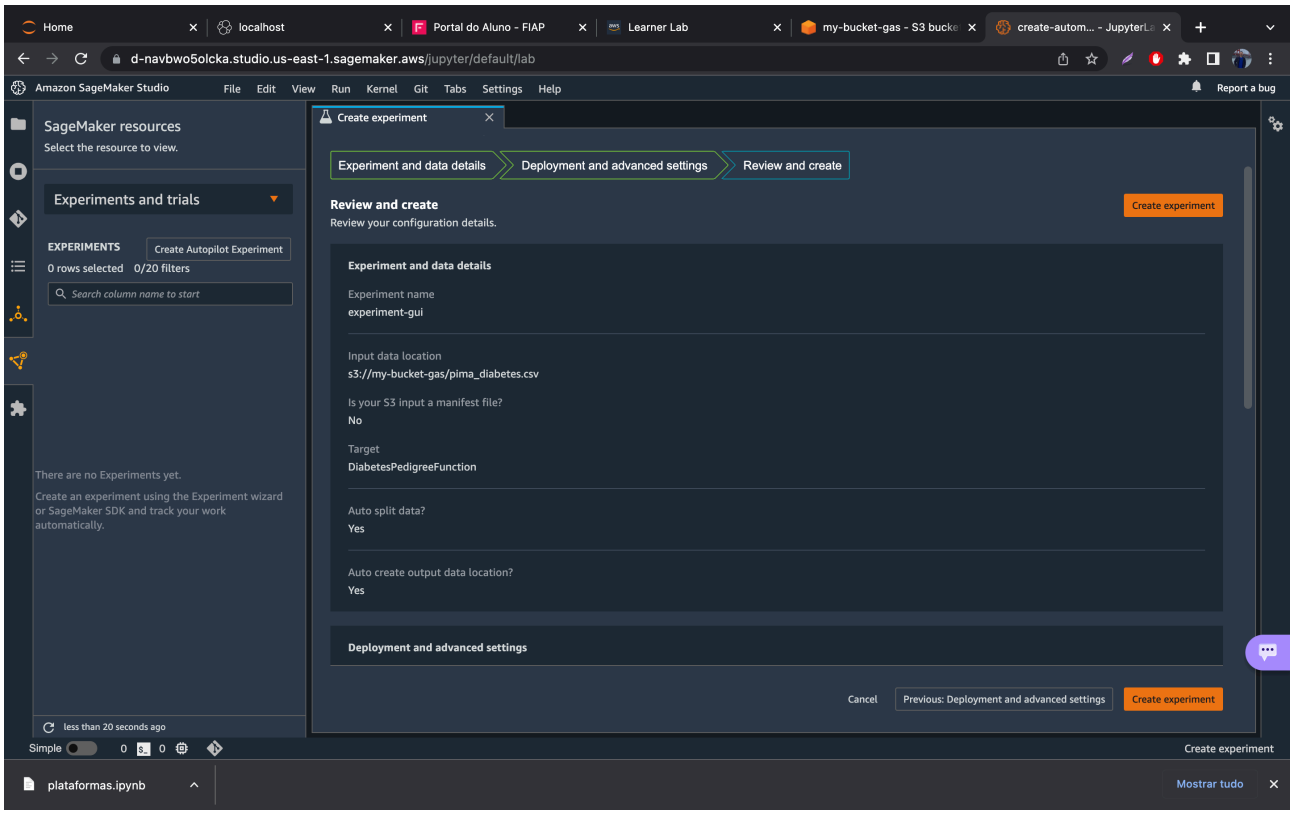


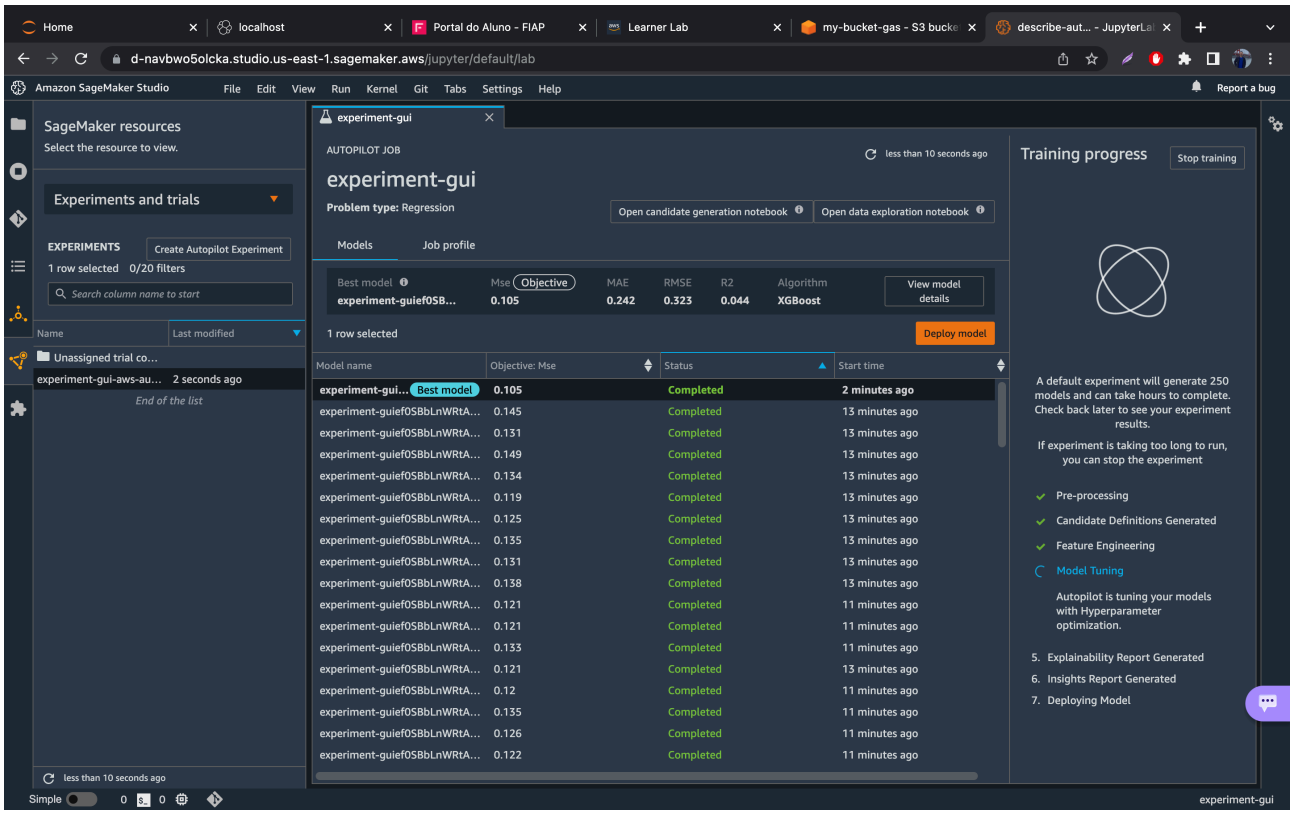
5) No **AWS Learner Lab**, crie uma instância do Jupyter Lab, adicione um notebook (visualizar-dados.ipynb) e gere um gráfico de dispersão entre 2 variáveis à sua escolha.

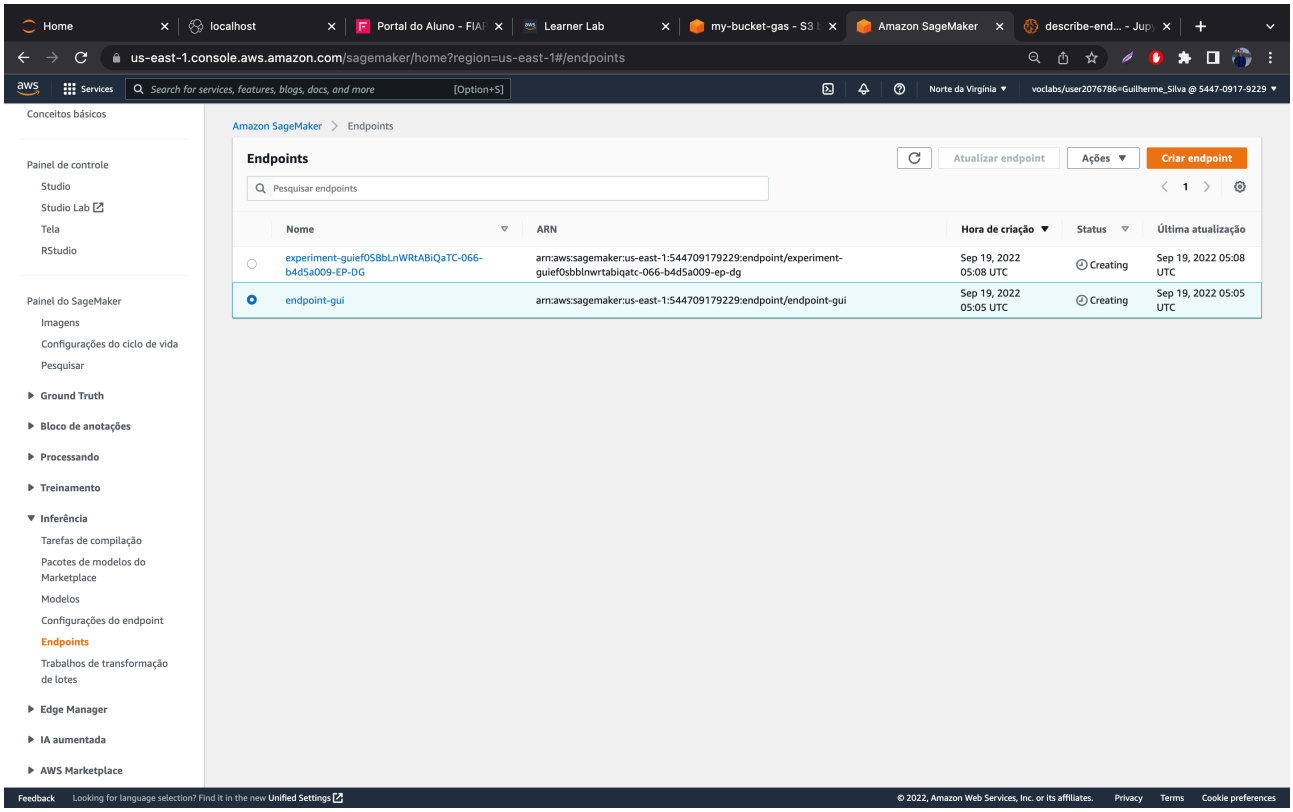


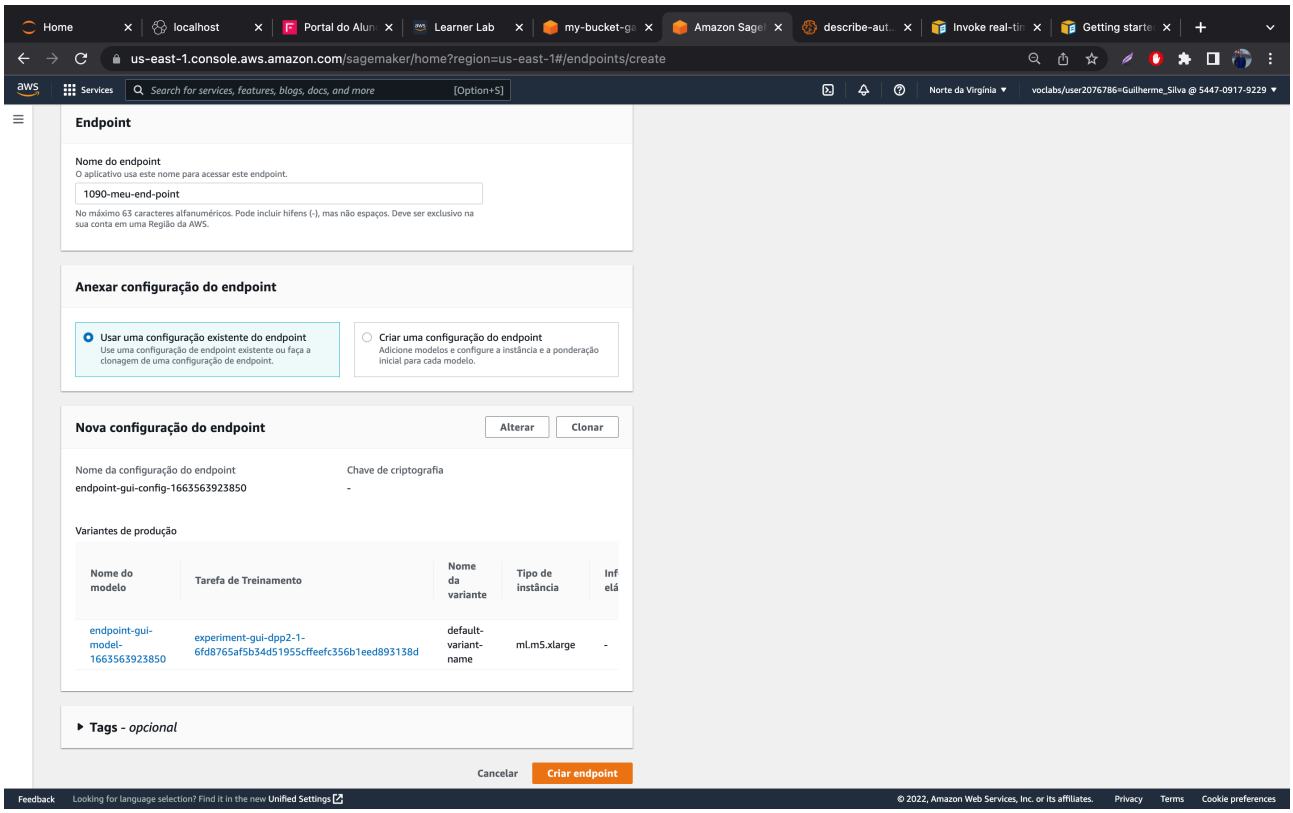
6) No **AWS Learner Lab**, utilize o auto pilot do sagemaker para criar um experimento para predição de risco de desenvolvimento de diabetes (use o dataset da pasta Challenge4 - pima_diabetes.csv). Faça o deploy e simule pelo menos 2 novas predições para o endpoint.

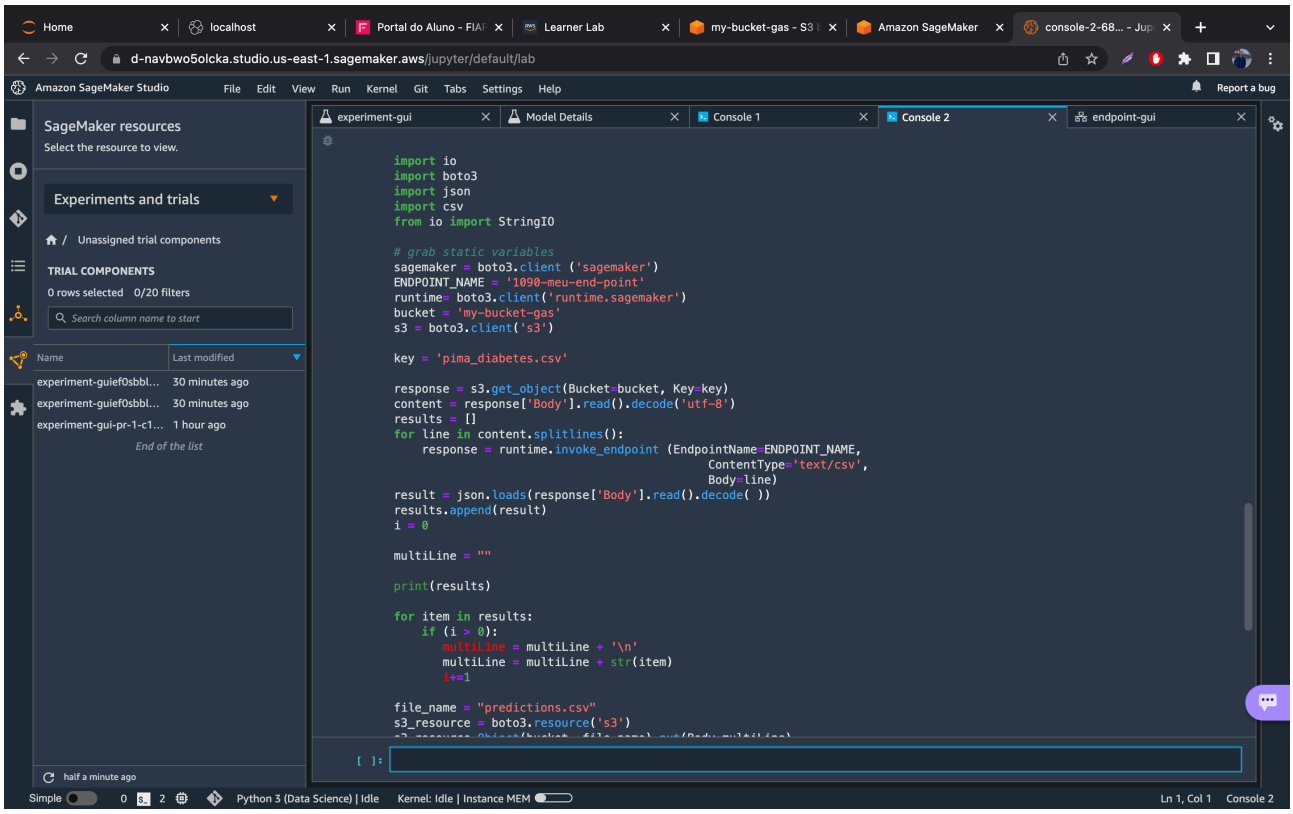



Home
localhost
Portal do Aluno - FIAP
Learner Lab
my-bucket-gas - S3 bucke
create-autom... - JupyterLa
d-navbwo5olcka.studio.us-east-1.sagemaker.aws/jupyter/default/lab
Amazon SageMaker Studio
File Edit View Run Kernel Git Tabs Settings Help
Report a bug
SageMaker resources
Select the resource to view.
Experiments and trials
EXPERIMENTS
Create Autopilot Experiment
0 rows selected 0/20 filters
Search column name to start
There are no Experiments yet.
Create an experiment using the Experiment wizard or SageMaker SDK and track your work automatically.
less than 20 seconds ago
Create experiment
Experiment and data details
Deployment and advanced settings
Review and create
Review and create
Review your configuration details.
Create experiment
Experiment and data details
Experiment name
experiment-gui
Input data location
s3://my-bucket-gas/pima_diabetes.csv
Is your S3 input a manifest file?
No
Target
DiabetesPedigreeFunction
Auto split data?
Yes
Auto create output data location?
Yes
Deployment and advanced settings
Cancel
Previous: Deployment and advanced settings
Create experiment
Simple
0 0
Create experiment
plataformas.ipynb
Mostrar tudo


Home
localhost
Portal do Aluno - FIAP
Learner Lab
my-bucket-gas - S3 bucke
describe-aut... - JupyterLa
d-navbwo5olcka.studio.us-east-1.sagemaker.aws/jupyter/default/lab
Amazon SageMaker Studio
File Edit View Run Kernel Git Tabs Settings Help
Report a bug
SageMaker resources
Select the resource to view.
Experiments and trials
EXPERIMENTS
Create Autopilot Experiment
1 row selected 0/20 filters
Search column name to start
Name
Last modified
Unassigned trial co...
experiment-gui-aws-au... 2 seconds ago
End of the list
less than 10 seconds ago
experiment-gui
AUTOPILOT JOB
experiment-gui
Problem type: Regression
less than 10 seconds ago
Open candidate generation notebook
Open data exploration notebook
Models
Job profile
Best model
Mse Objective
MAE
RMSE
R2
Algorithm
experiment-guief0SB...
0.105
0.242
0.323
0.044
XGBoost
View model details
1 row selected
Deploy model
Model name
Objective: Mse
Status
Start time
experiment-gui... Best model 0.105 Completed 2 minutes ago
experiment-guief0SBbLnWRtA... 0.145 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.131 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.149 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.134 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.119 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.125 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.135 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.131 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.138 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.121 Completed 11 minutes ago
experiment-guief0SBbLnWRtA... 0.121 Completed 11 minutes ago
experiment-guief0SBbLnWRtA... 0.133 Completed 11 minutes ago
experiment-guief0SBbLnWRtA... 0.121 Completed 13 minutes ago
experiment-guief0SBbLnWRtA... 0.12 Completed 11 minutes ago
experiment-guief0SBbLnWRtA... 0.135 Completed 11 minutes ago
experiment-guief0SBbLnWRtA... 0.126 Completed 11 minutes ago
experiment-guief0SBbLnWRtA... 0.122 Completed 11 minutes ago
Training progress
Stop training
A default experiment will generate 250 models and can take hours to complete. Check back later to see your experiment results.
If experiment is taking too long to run, you can stop the experiment
Pre-processing
Candidate Definitions Generated
Feature Engineering
Model Tuning
Autopilot is tuning your models with Hyperparameter optimization.
5. Explainability Report Generated
6. Insights Report Generated
7. Deploying Model
less than 10 seconds ago
Simple
0 0
experiment-gui


Home
localhost
Portal do Aluno - FIAP
Learner Lab
my-bucket-gas - S3
Amazon SageMaker
describe-end... - Jupyter
us-east-1.console.aws.amazon.com/sagemaker/home?region=us-east-1#/endpoints
Services
Search for services, features, blogs, docs, and more
[Option+S]
Norte da Virgínia
voclabs/user2076786=Guilherme_Silva @ 5447-0917-9229
Conceitos básicos
Painel de controle
Studio
Studio Lab
Tela
RStudio
Painel do SageMaker
Imagens
Configurações do ciclo de vida
Pesquisar
Ground Truth
Bloco de anotações
Processando
Treinamento
Inferência
Tarefas de compilação
Pacotes de modelos do Marketplace
Modelos
Configurações do endpoint
Endpoints
Trabalhos de transformação de lotes
Edge Manager
IA aumentada
AWS Marketplace
Amazon SageMaker > Endpoints
Endpoints
Atualizar endpoint
Ações
Criar endpoint
Pesquisar endpoints
Nome
ARN
Hora de criação
Status
Última atualização
experiment-guief0SBbLnWRtABiQaTC-066-b4d5a009-EP-DG
arn:aws:sagemaker:us-east-1:544709179229:endpoint/experiment-guief0sbblnwrtabiqatc-066-b4d5a009-ep-dg
Sep 19, 2022 05:08 UTC
Creating
Sep 19, 2022 05:08 UTC
endpoint-gui
arn:aws:sagemaker:us-east-1:544709179229:endpoint/endpoint-gui
Sep 19, 2022 05:05 UTC
Creating
Sep 19, 2022 05:05 UTC
Feedback
Looking for language selection? Find it in the new Unified Settings
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Home
localhost
Portal do Alun
Learner Lab
my-bucket-ga
Amazon Sage
describe-aut
Invoke real-tin
Getting starte
us-east-1.console.aws.amazon.com/sagemaker/home?region=us-east-1#/endpoints/create
Services
Search for services, features, blogs, docs, and more
[Option+S]
Norte da Virgínia
voclabs/user2076786=Guilherme_Silva @ 5447-0917-9229
Endpoint
Nome do endpoint
O aplicativo usa este nome para acessar este endpoint.
1090-meu-end-point
No máximo 63 caracteres alfanuméricos. Pode incluir hífens (-), mas não espaços. Deve ser exclusivo na sua conta em uma Região da AWS.
Anexar configuração do endpoint
Usar uma configuração existente do endpoint
Use uma configuração de endpoint existente ou faça a clonagem de uma configuração de endpoint.
Criar uma configuração do endpoint
Adicione modelos e configure a instância e a ponderação inicial para cada modelo.
Nova configuração do endpoint
Alterar
Clonar
Nome da configuração do endpoint
endpoint-gui-config-1663563923850
Chave de criptografia
-
Variantes de produção
Nome do modelo
Tarefa de Treinamento
Nome da variante
Tipo de instância
endpoint-gui-model-1663563923850
experiment-gui-dpp2-1-6fd8765af5b34d51955cffeefc356b1eed893138d
default-variant-name
ml.m5.xlarge
-
Tags - opcional
Cancelar
Criar endpoint
Feedback
Looking for language selection? Find it in the new Unified Settings
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Home
localhost
Portal do Aluno - FIAP
Learner Lab
my-bucket-gas - S3
Amazon SageMaker
console-2-68... - Jupyter
d-navbwo5olcka.studio.us-east-1.sagemaker.aws/jupyter/default/lab
Amazon SageMaker Studio
File Edit View Run Kernel Git Tabs Settings Help
Report a bug
SageMaker resources
Select the resource to view.
Experiments and trials
/ Unassigned trial components
TRIAL COMPONENTS
0 rows selected 0/20 filters
Search column name to start
Name
Last modified
experiment-guief0sbbl... 30 minutes ago
experiment-guief0sbbl... 30 minutes ago
experiment-gui-pr-1-c1... 1 hour ago
End of the list
half a minute ago
experiment-gui
Model Details
Console 1
Console 2
endpoint-gui
import io
import boto3
import json
import csv
from io import StringIO
# grab static variables
sagemaker = boto3.client ('sagemaker')
ENDPOINT_NAME = '1090-meu-end-point'
runtime= boto3.client('runtime.sagemaker')
bucket = 'my-bucket-gas'
s3 = boto3.client('s3')
key = 'pima_diabetes.csv'
response = s3.get_object(Bucket=bucket, Key=key)
content = response['Body'].read().decode('utf-8')
results = []
for line in content.splitlines():
response = runtime.invoke_endpoint (EndpointName=ENDPOINT_NAME,
ContentType='text/csv',
Body=line)
result = json.loads(response['Body'].read().decode( ))
results.append(result)
i = 0
multiLine = ""
print(results)
for item in results:
if (i > 0):
multiLine = multiLine + '\n'
multiLine = multiLine + str(item)
i+=1
file_name = "predictions.csv"
s3_resource = boto3.resource('s3')
[ ]:
Simple
Python 3 (Data Science) | Idle
Kernel: Idle | Instance MEM
Ln 1, Col 1
Console 2