Comece com ML hoje

CHARLES BAUER

Links úteis

Diagrama de uso do Azure ML Studio:

https://docs.microsoft.com/pt-br/azure/machine-learning/machine-learning-studio-overview-diagram

Livro com passo a passo para começar a usar o Azure ML Studio:

https://blogs.msdn.microsoft.com/microsoft_press/2015/04/15/free-ebook-microsoft-azure-essentials-azure-machine-learning/

Que tipo de Perguntas ML Responde?

O que houve? Quem são? Quantos são? (Análise Descritiva)

- Quem são meus clientes?
- Quantos tipos de clientes temos?

Por que tal coisa está acontecendo? (Análise Diagnóstica)

- Por que estamos perdendo clientes todos os meses?
- Por que o número de ligações está aumentando?

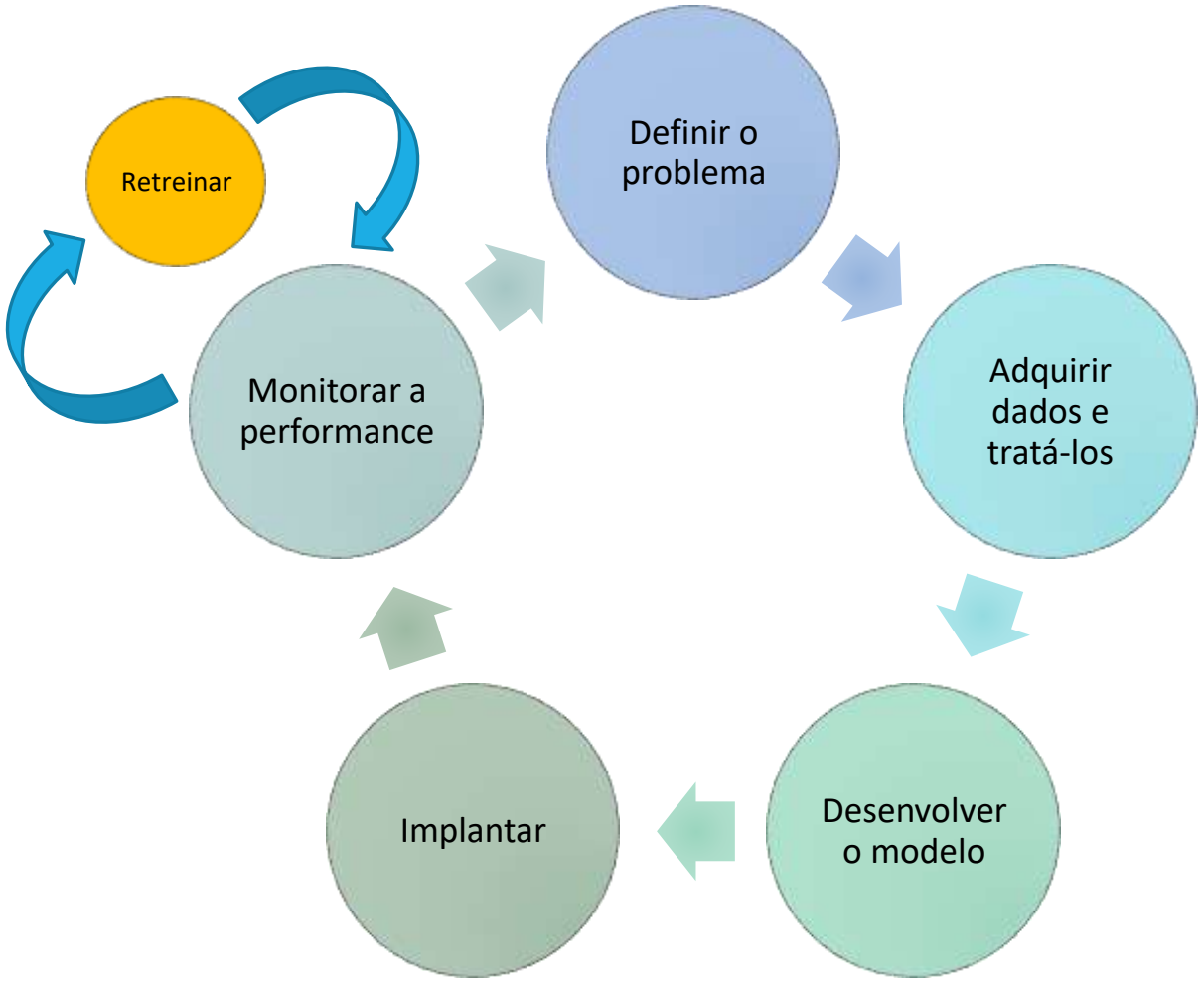
O que vai acontecer? (Análise Preditiva)

- Quais clientes vão comprar nosso produto Premium?
- Quais clientes vão nos deixar?

O que podemos fazer no futuro? (Análise Prescritiva)

- Que serviços devemos oferecer para tais clientes?
- Quantas pessoas devemos colocar para atender as ligações no natal?

Como fazemos isso?



Que técnicas utilizamos?

Classificação de dados
- Ex.: Classificar uma oportunidade dentro das áreas da empresa com base em anotações do CRM;
- Dizer se o cliente vai cancelar ou não um contrato.

Clusterização
- Ex.: Agrupar clientes por similaridade;
- Agrupar documentos semelhantes.

Regressão
- Ex.: Dizer quanto devemos cobrar por um projeto;
- Dizer em que temperatura provavelmente o sistema irá parar.

Análise de Conteúdo
- Ex.: Dizer quais comentários no Twitter são negativos;
- Identificar rostos em fotos.

Recomendação
- Ex.: Dizer que produtos ou serviços devemos oferecer para determinado cliente;
- Identificar que produtos que normalmente são vendidos em conjunto.

Com utilizar tudo isso no dia-a-dia?

API´s e bibliotecas para integração em seus sistemas atuais ou novos;

Consultas via Tabelas do Excel;

Entrega de relatórios pontuais.

Exemplo prático

Prever a renda anual de uma pessoa

Problema:

- Com base em dados históricos prever a renda anual de uma determinada pessoa.

Obter e Tratar dados:

- Dados do senso Demográfico Americano de 1994-1995;
- Obter dados do senso disponibilizados pela UCI (https://archive.ics.uci.edu/ml/datasets/Adult);
- Remover colunas que não nos ajudam na análise;
- Remover linhas que contém dados faltando pois também atrapalham o algoritmo escolhido;

Desenvolver o modelo:

- Escolhemos um algoritmo binário para termos apenas duas respostas (<=50k ou >50k);

Implantamos o modelo:

- Utilização via Excel;
- Utilização via programa em Python;

Monitoramos a performance/assertividade do modelo:

- Utilizamos gráficos e técnicas de avaliação do modelo.

Onde faremos isso?

Precisamos de poder computacional:

- Para treinar o modelo;
- Ter escalabilidade;

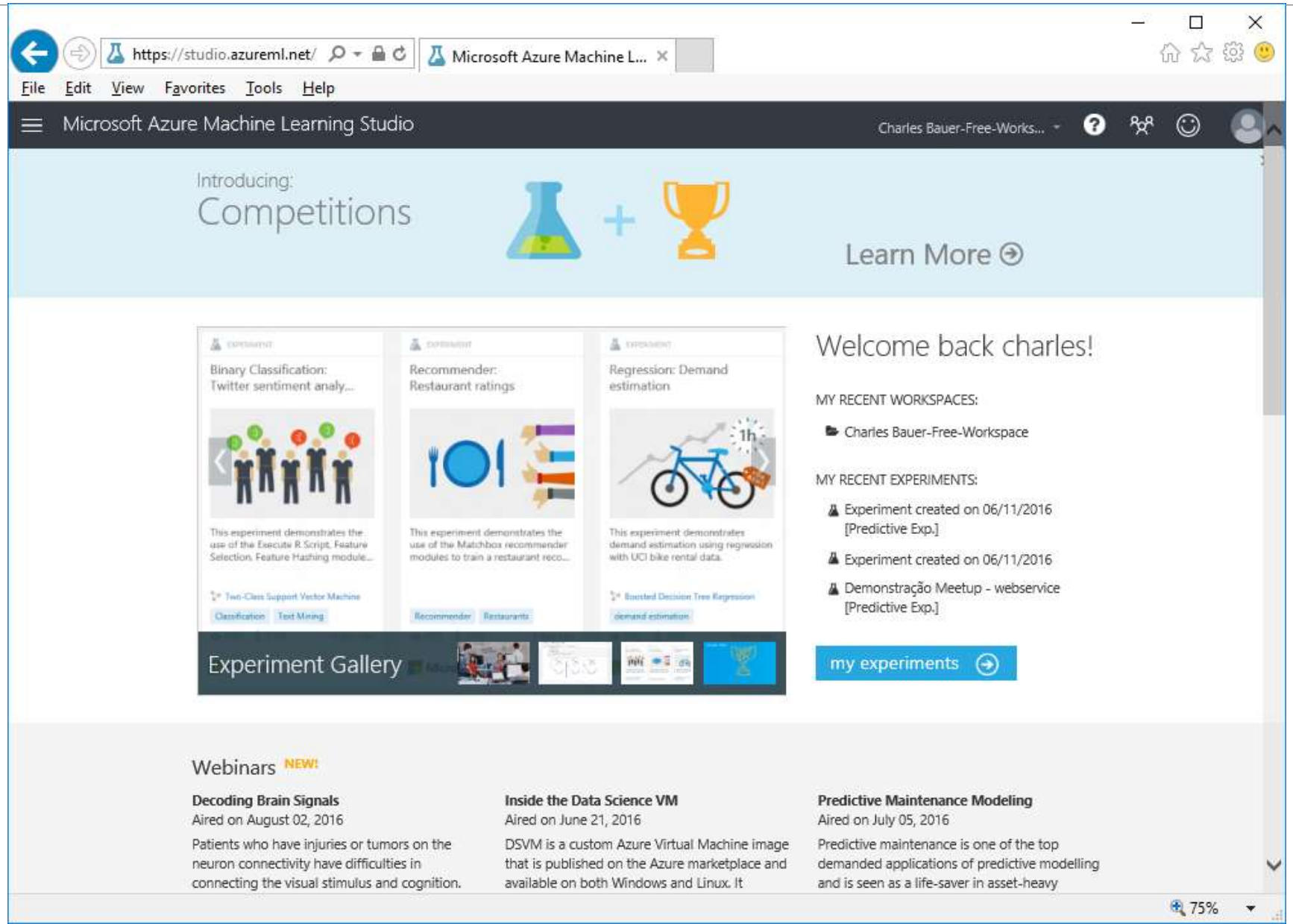
Queremos começar hoje:

- Não queremos ficar instalando pacotes, compilando, parametrizando, lendo sobre algoritmos...

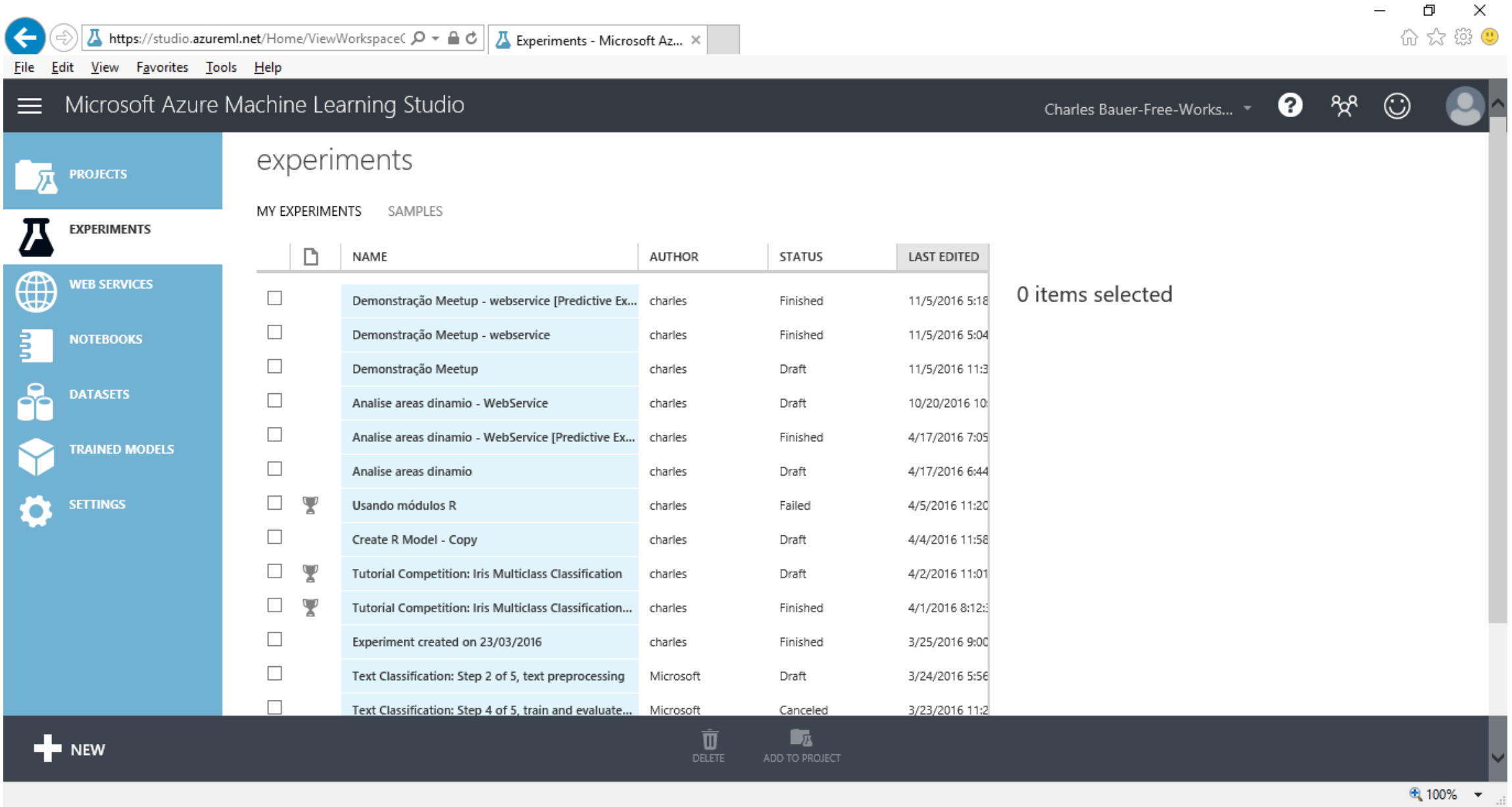
Escolha:

- Cloud Computing: Microsoft Azure.
- Opção Gratuita.

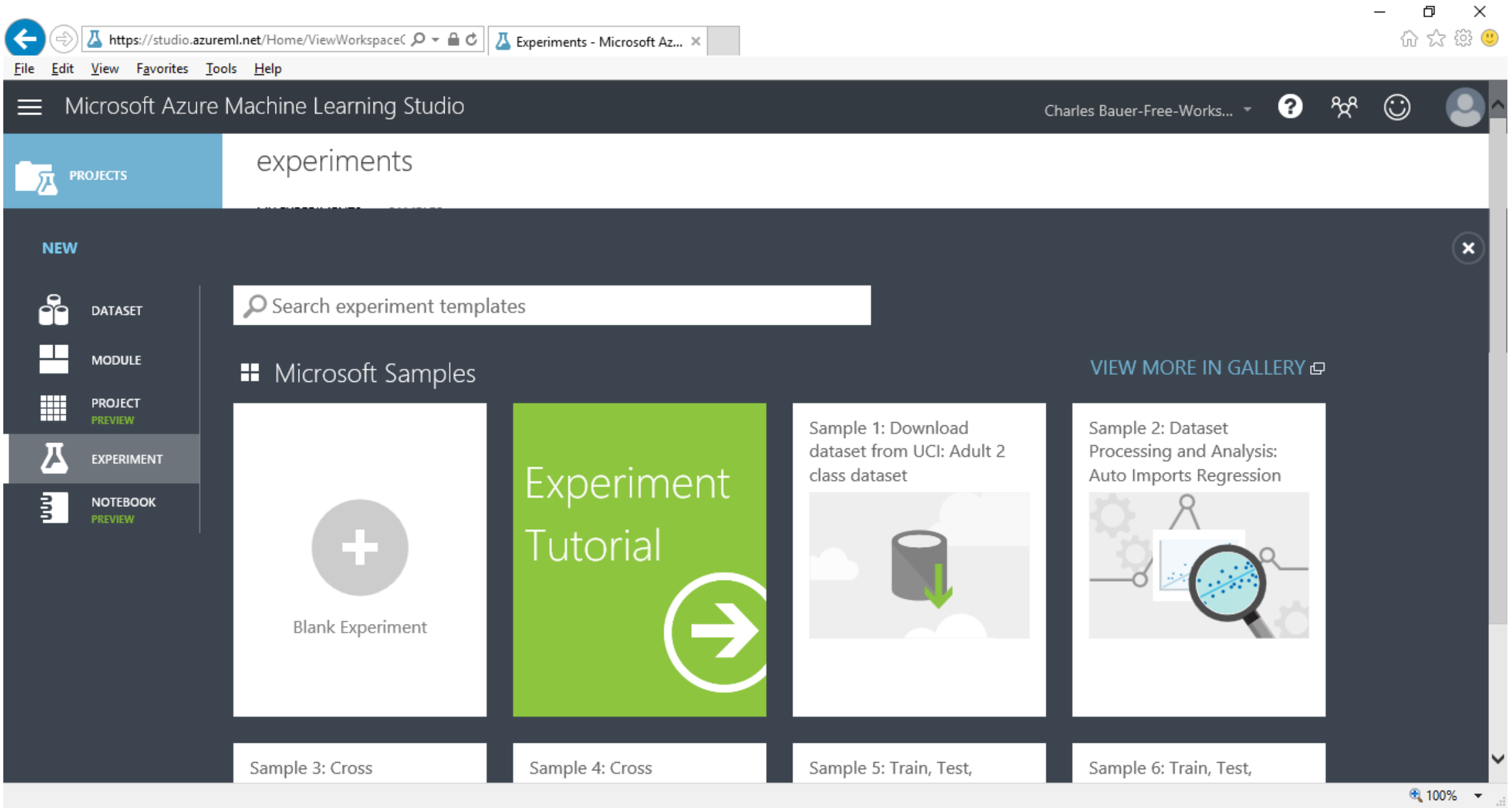
ML Studio - https://studio.azureml.net



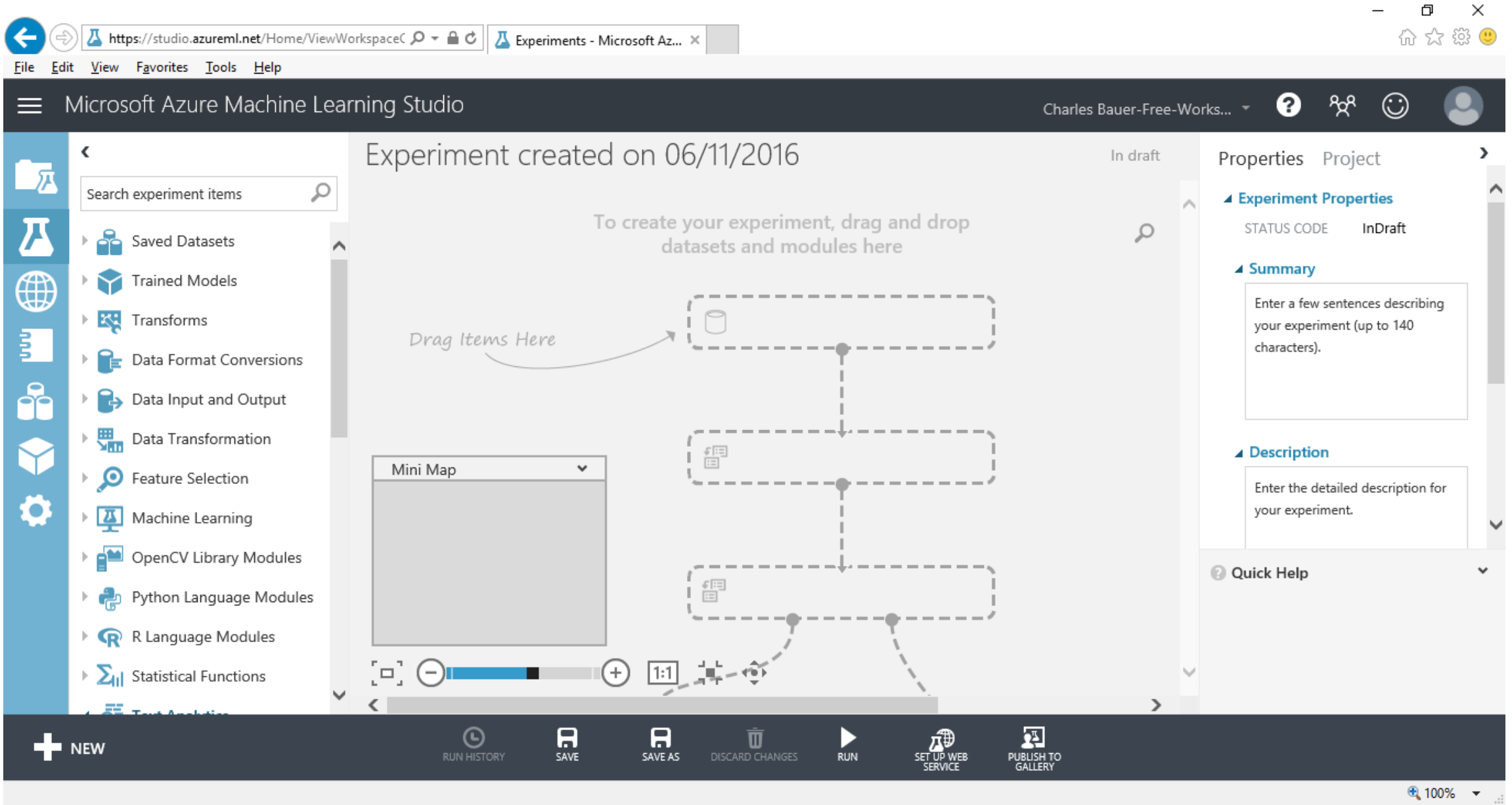
ML Studio


https://studio.azureml.net/Home/ViewWorkspaceC
Experiments - Microsoft Az...
File Edit View Favorites Tools Help
Microsoft Azure Machine Learning Studio
Charles Bauer-Free-Works...
PROJECTS
EXPERIMENTS
WEB SERVICES
NOTEBOOKS
DATASETS
TRAINED MODELS
SETTINGS
experiments
MY EXPERIMENTS
SAMPLES
NAME
AUTHOR
STATUS
LAST EDITED
Demonstração Meetup - webservice [Predictive Ex... charles Finished 11/5/2016 5:18
Demonstração Meetup - webservice charles Finished 11/5/2016 5:04
Demonstração Meetup charles Draft 11/5/2016 11:3
Analise areas dinamio - WebService charles Draft 10/20/2016 10
Analise areas dinamio - WebService [Predictive Ex... charles Finished 4/17/2016 7:05
Analise areas dinamio charles Draft 4/17/2016 6:44
Usando módulos R charles Failed 4/5/2016 11:20
Create R Model - Copy charles Draft 4/4/2016 11:58
Tutorial Competition: Iris Multiclass Classification charles Draft 4/2/2016 11:01
Tutorial Competition: Iris Multiclass Classification... charles Finished 4/1/2016 8:12:
Experiment created on 23/03/2016 charles Finished 3/25/2016 9:00
Text Classification: Step 2 of 5, text preprocessing Microsoft Draft 3/24/2016 5:56
Text Classification: Step 4 of 5, train and evaluate... Microsoft Canceled 3/23/2016 11:2
0 items selected
NEW
DELETE
ADD TO PROJECT
100%

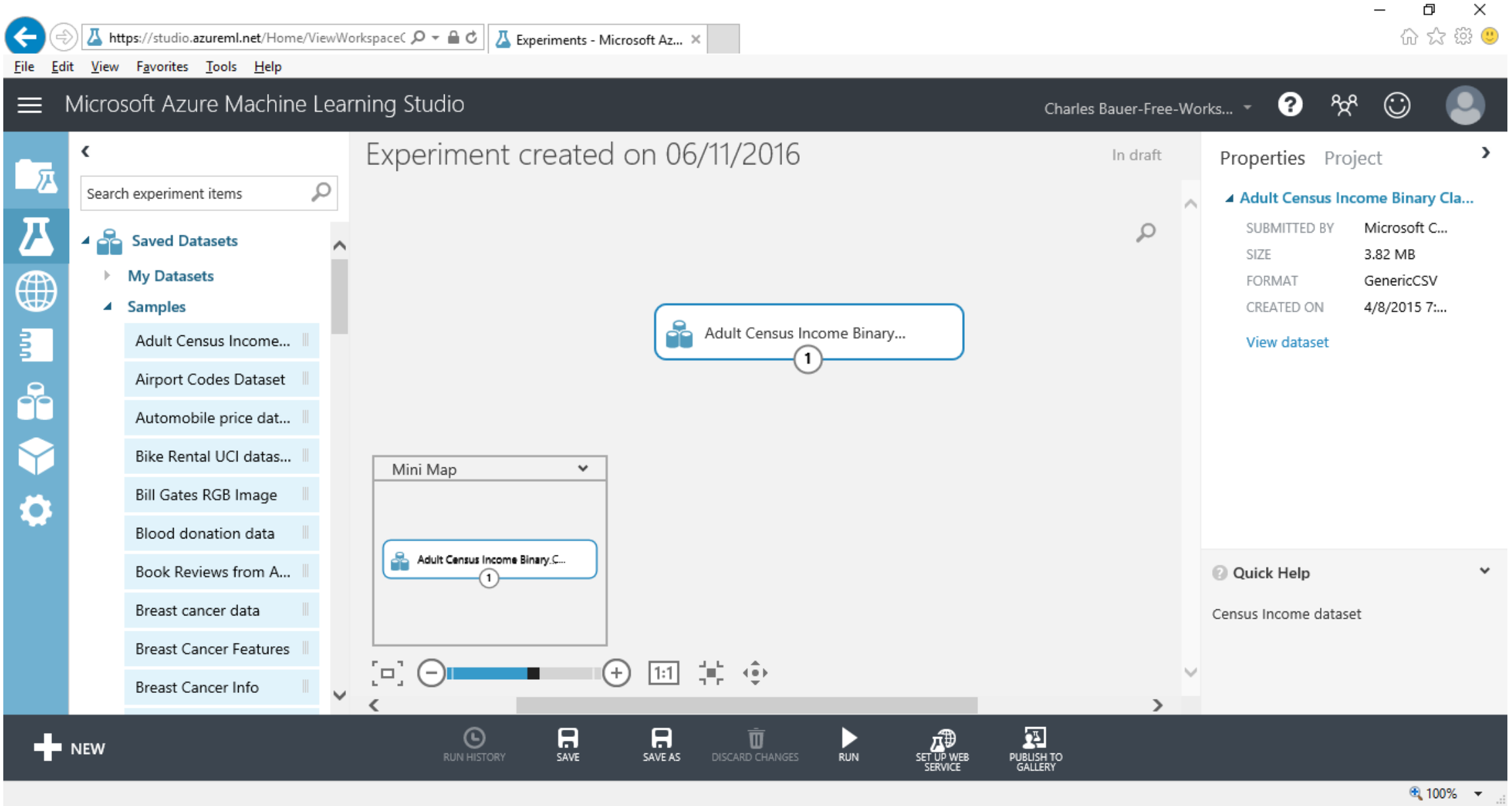
ML Studio: Novo experimento



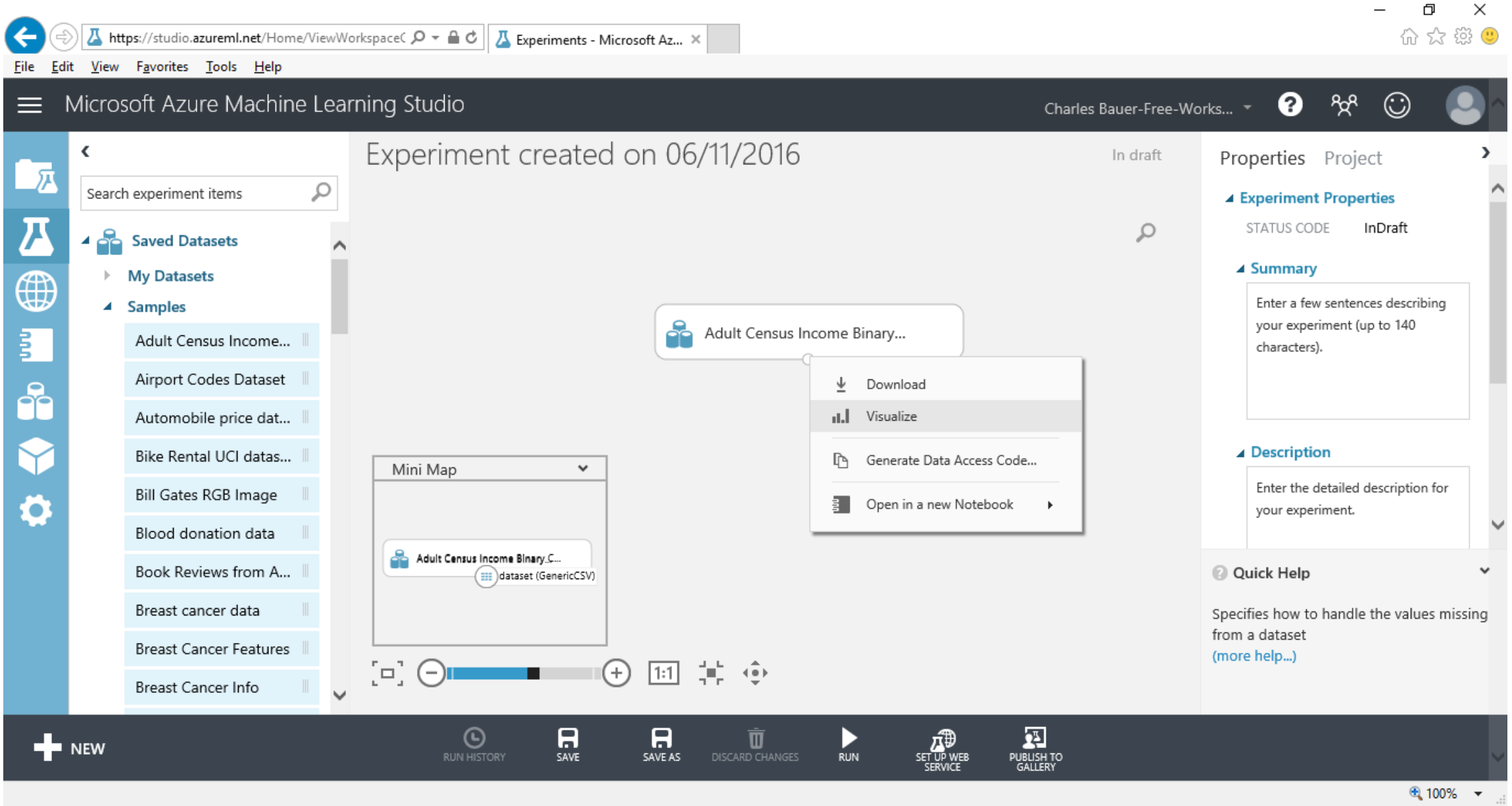
ML Studio: A Interface



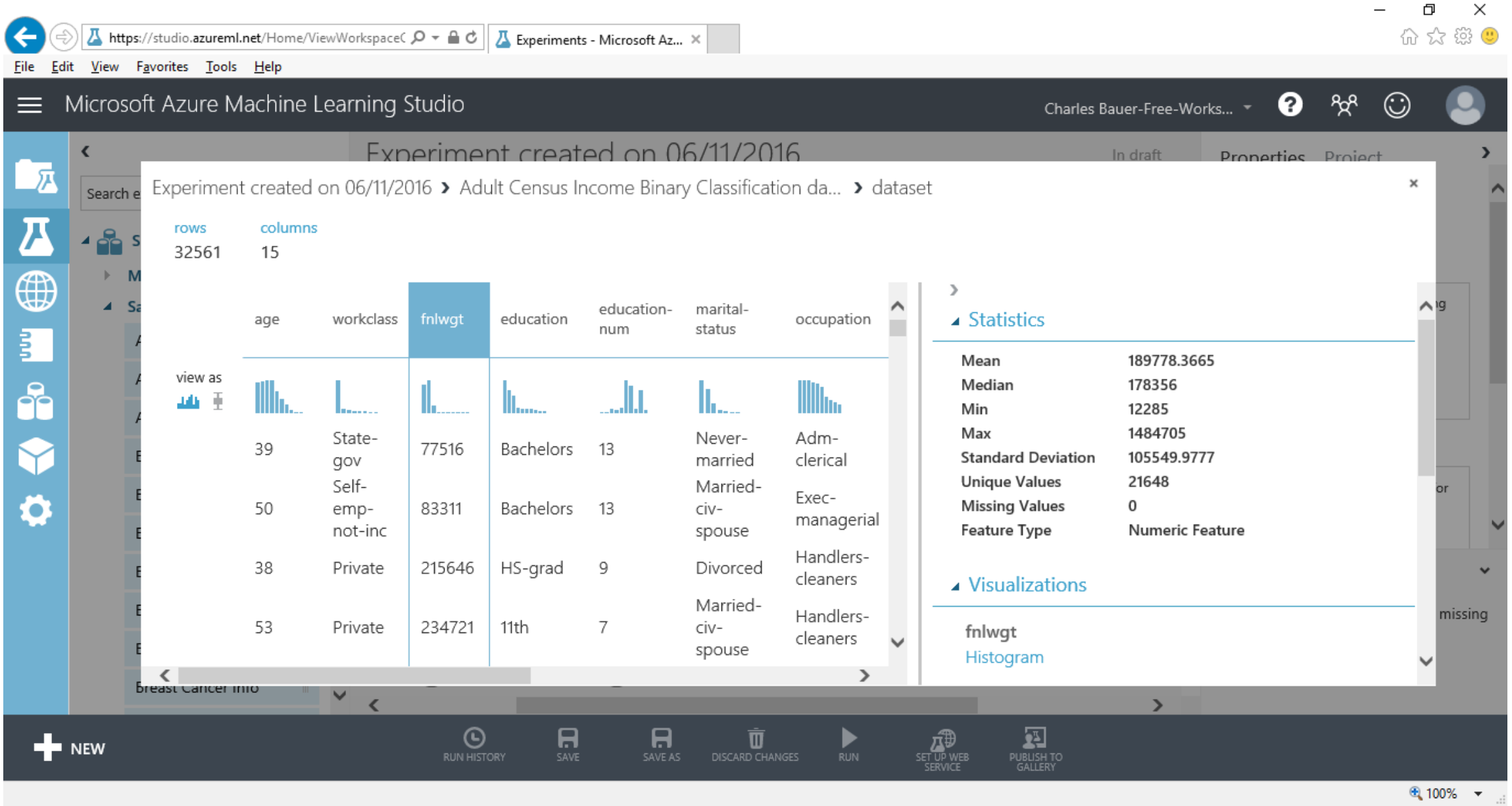
ML Studio: Obtendo os dados



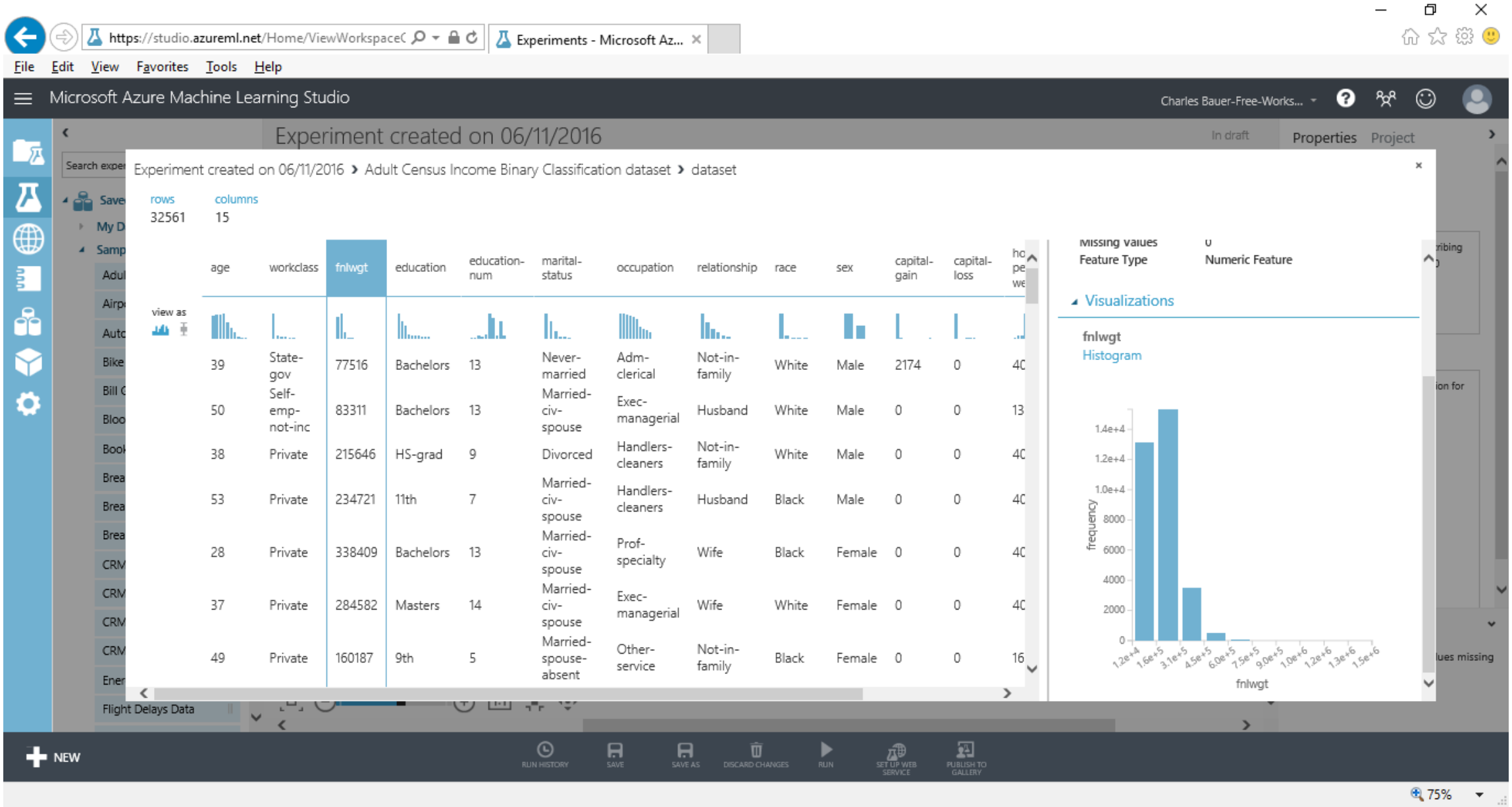
ML Studio: Analisando os dados



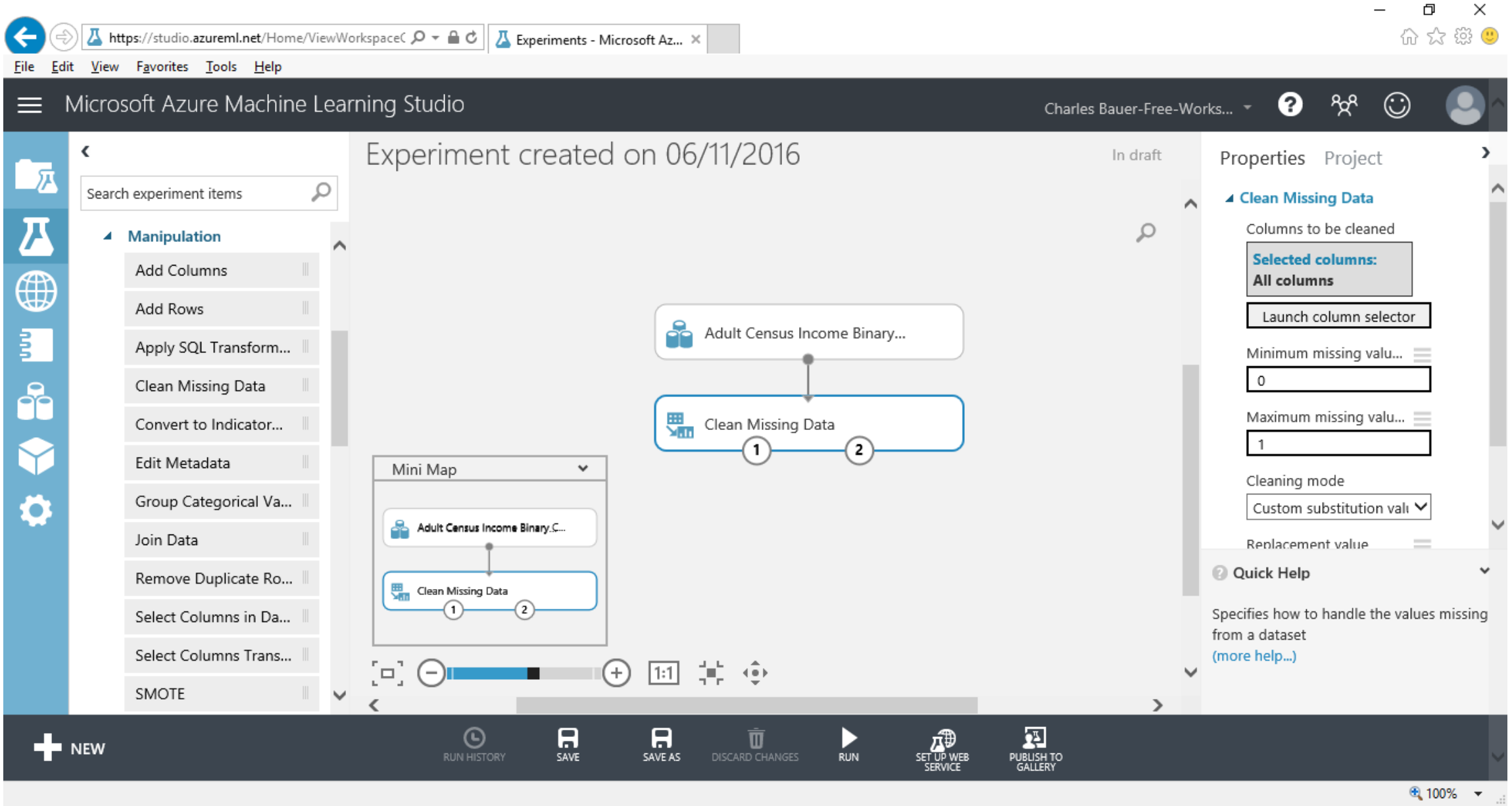
ML Studio: Analisando os dados



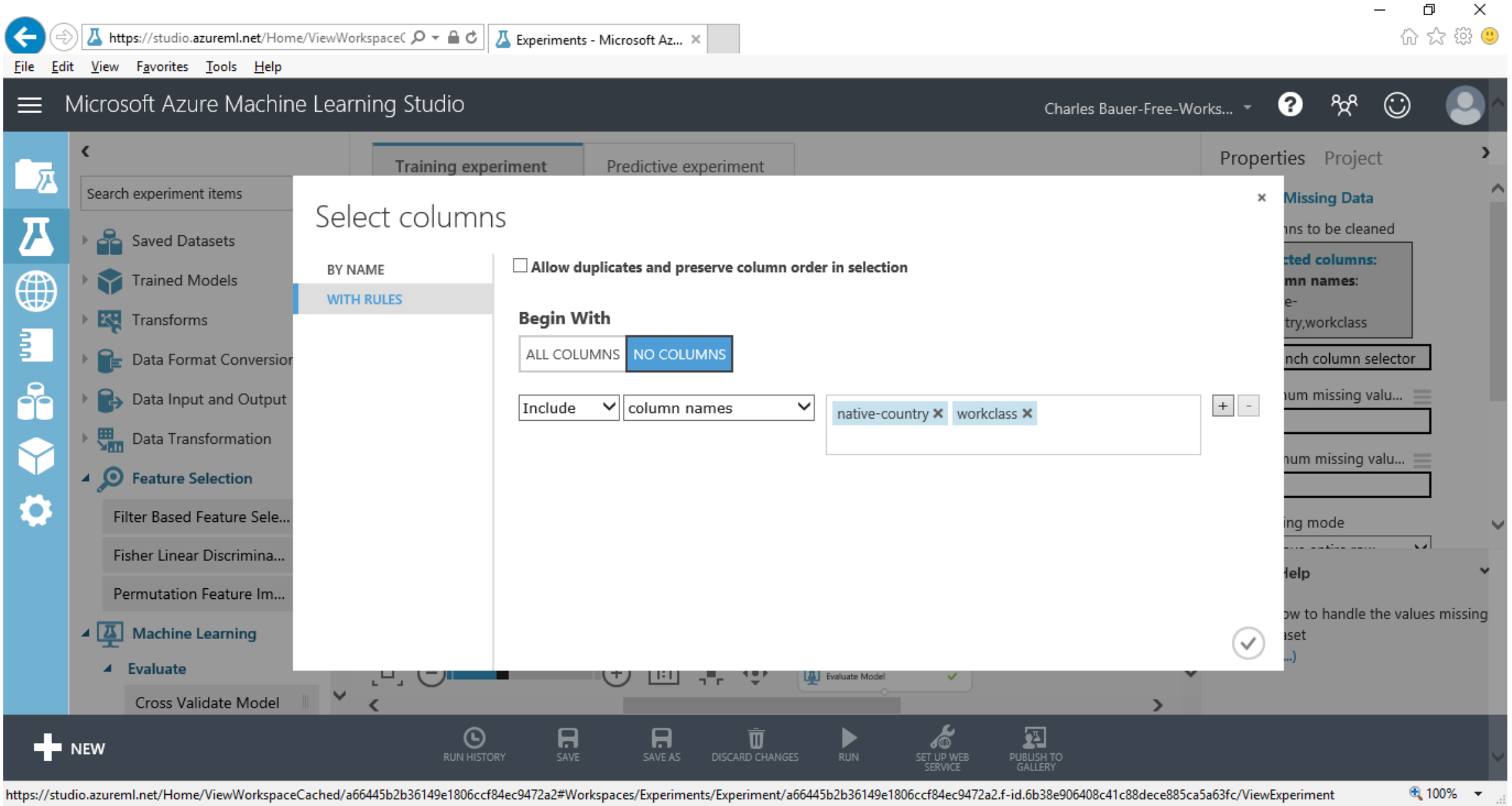
ML Studio: Analisando os dados



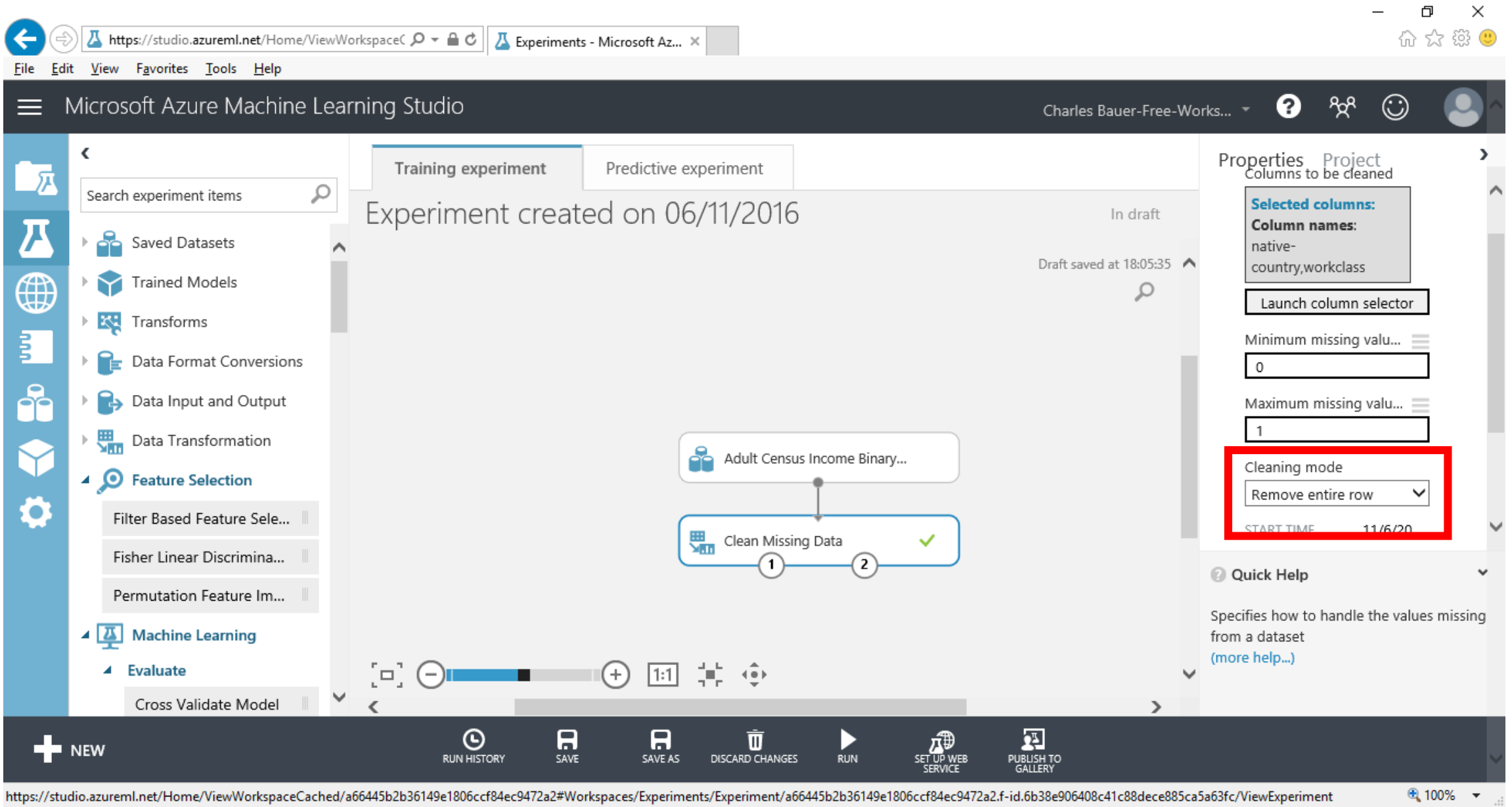
ML Studio: Tratando os dados



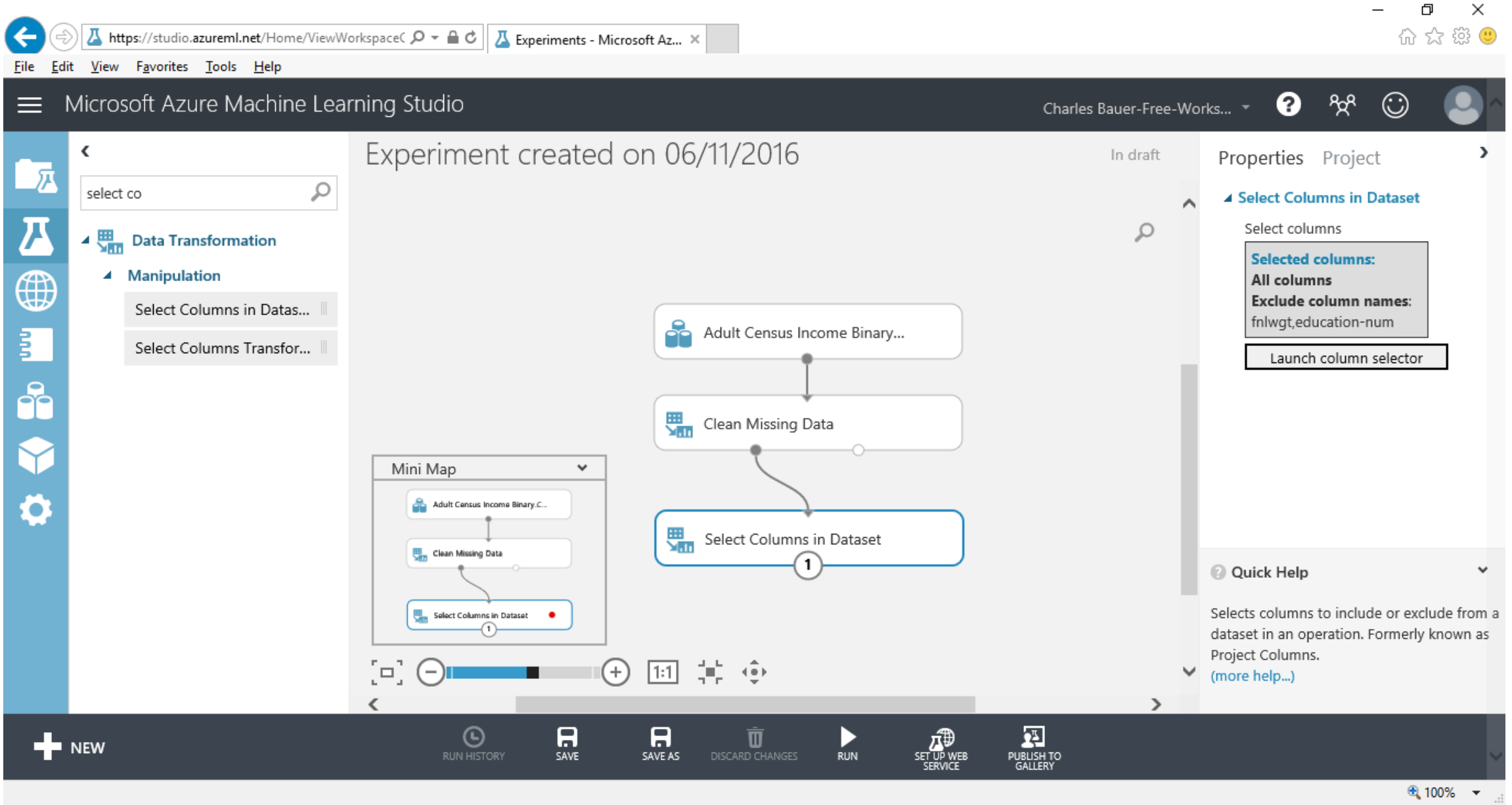
ML Studio: Tratando os dados



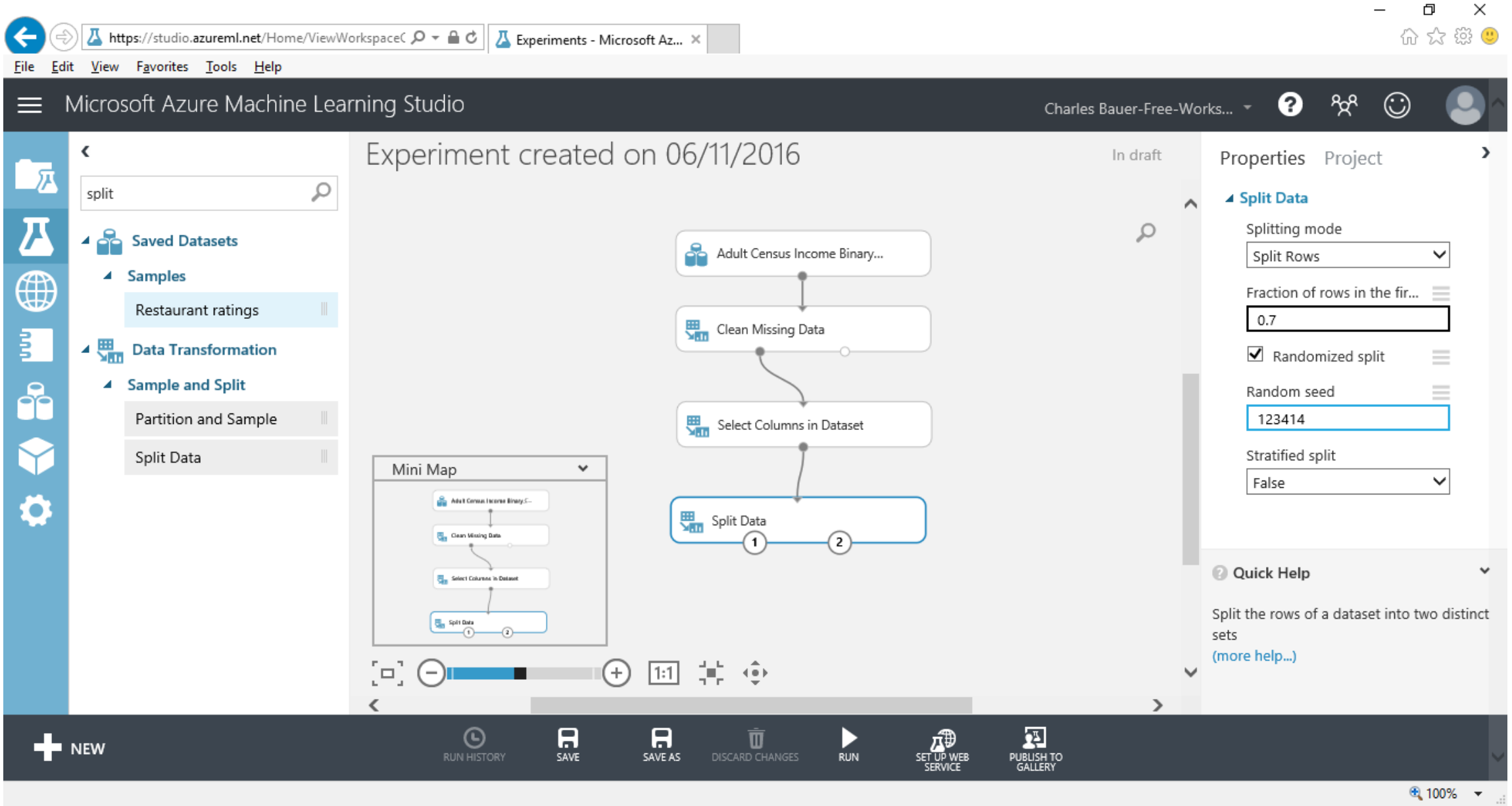
ML Studio: Tratando os dados: remover linhas



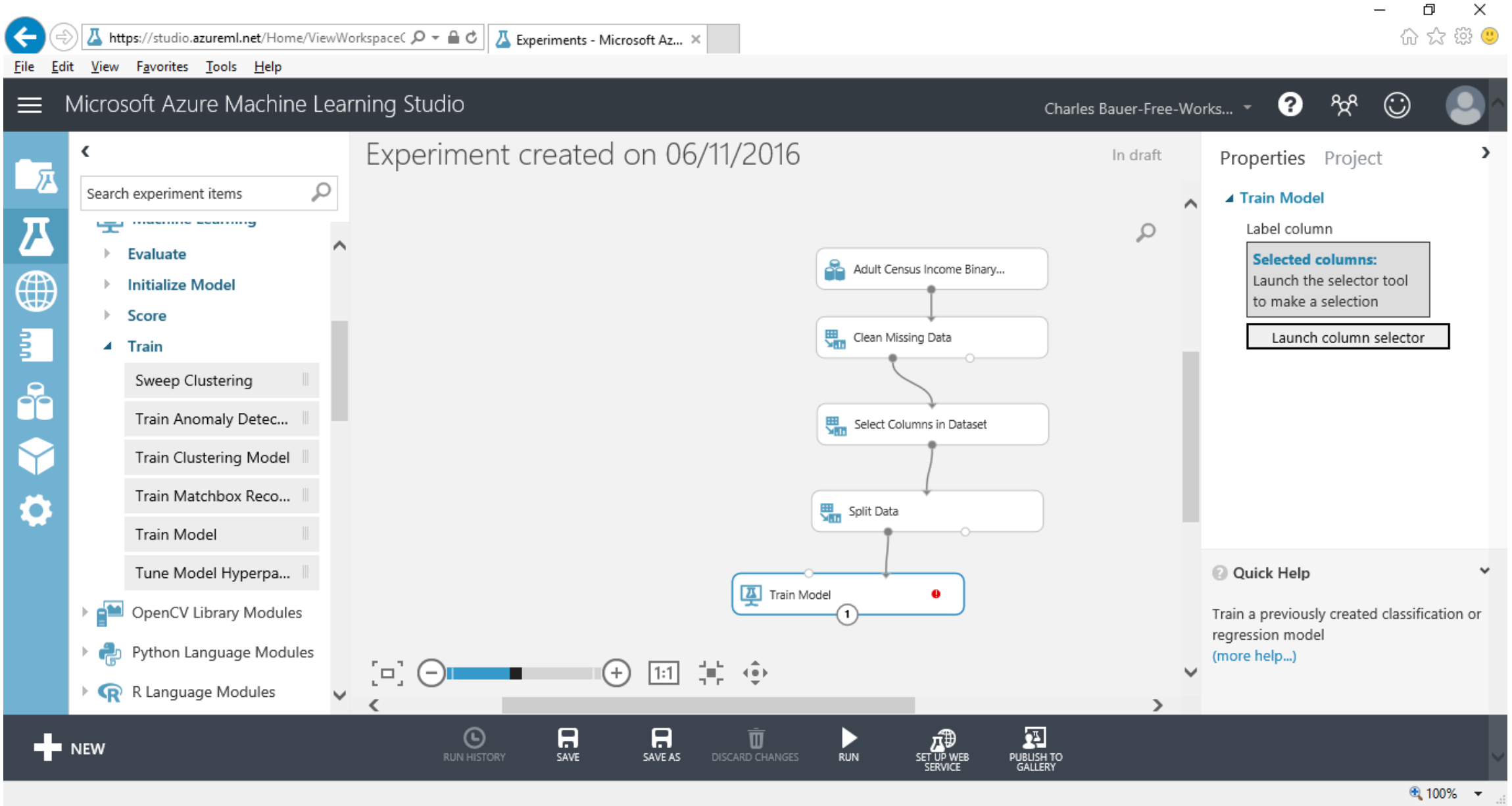
ML Studio: Tratando os dados: selecionando colunas



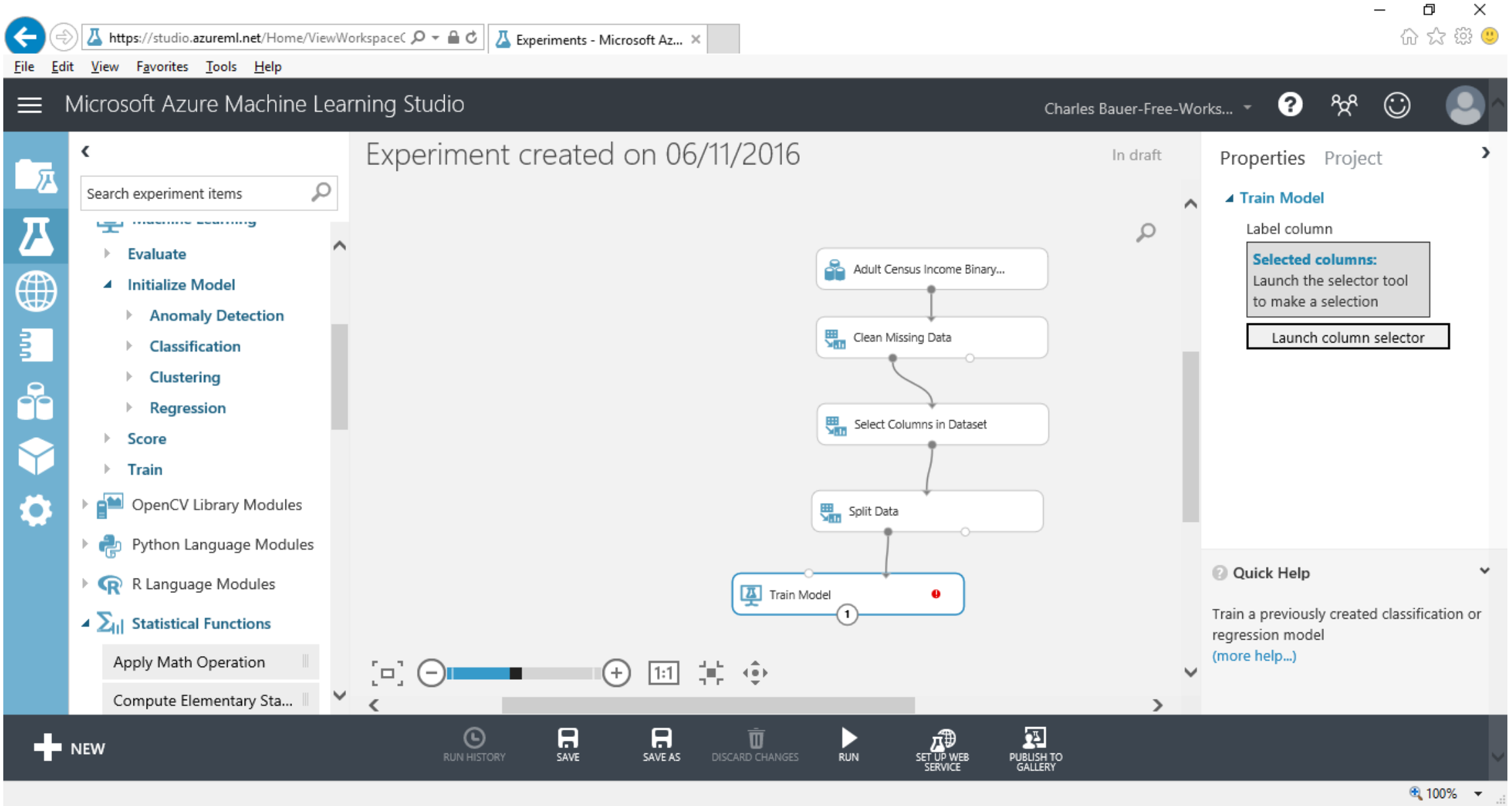
ML Studio: Separando dados de aprendizado e teste



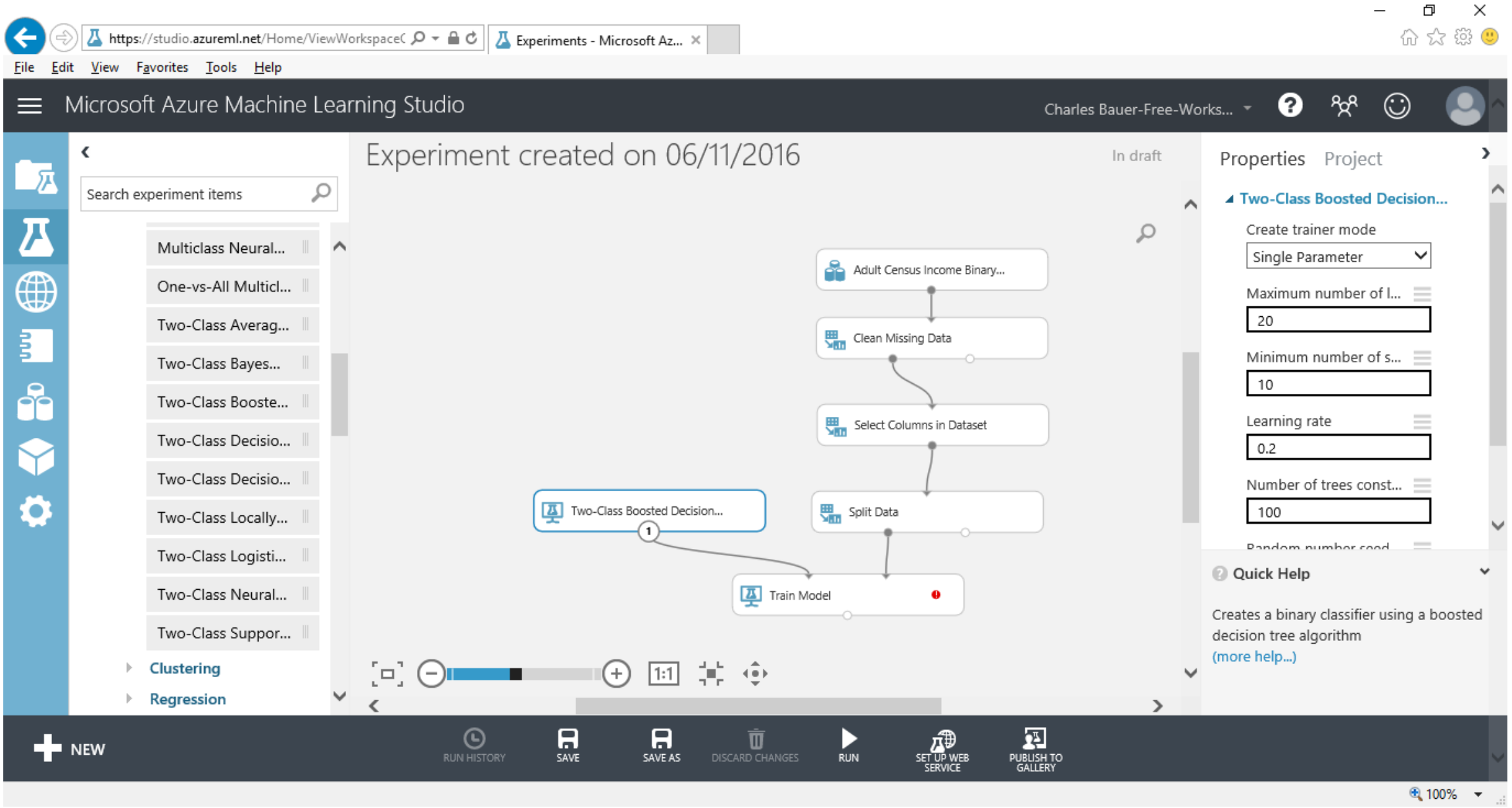
ML Studio: Escolhendo o algoritmo



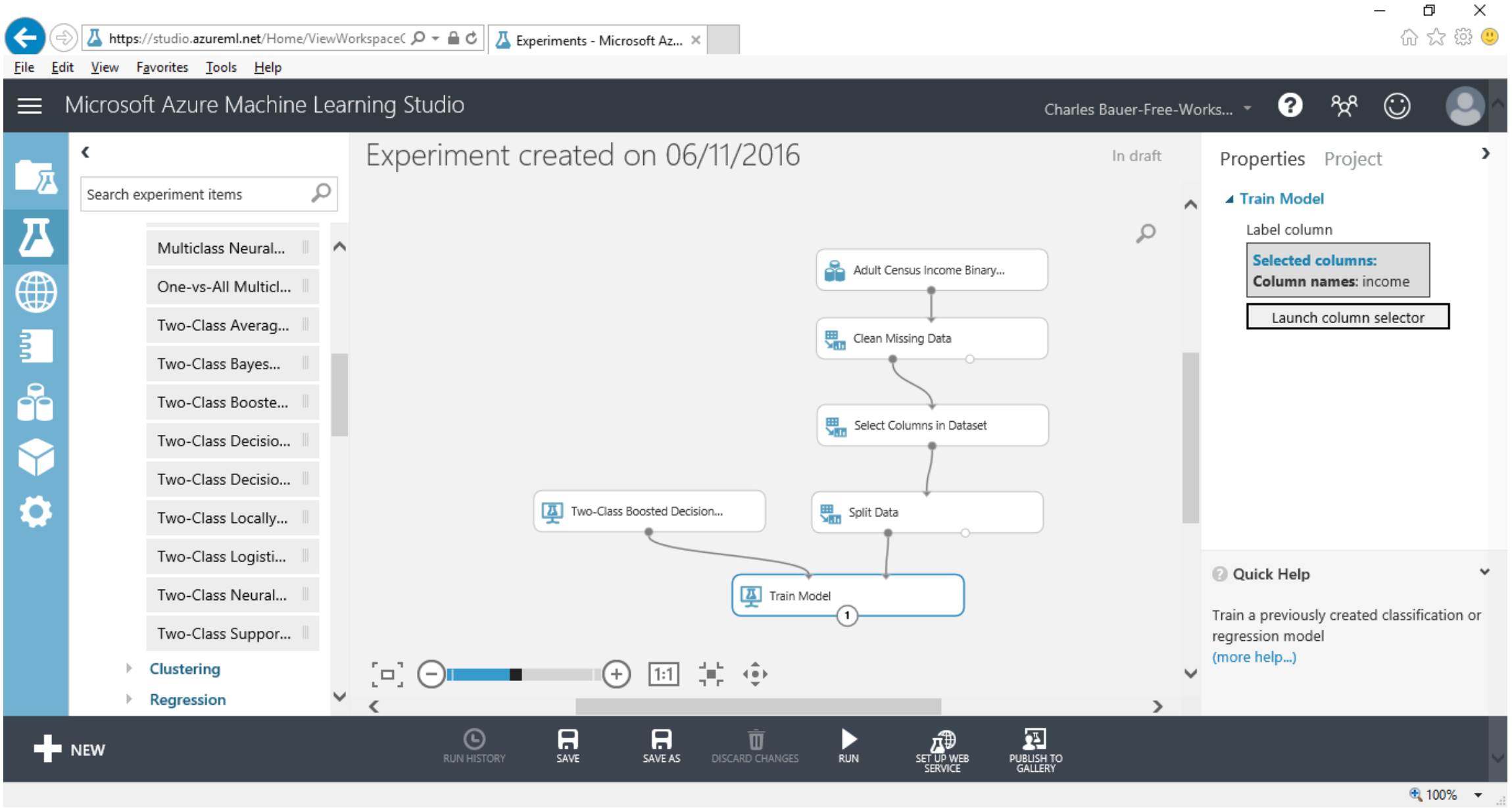
ML Studio: Escolhendo o algoritmo



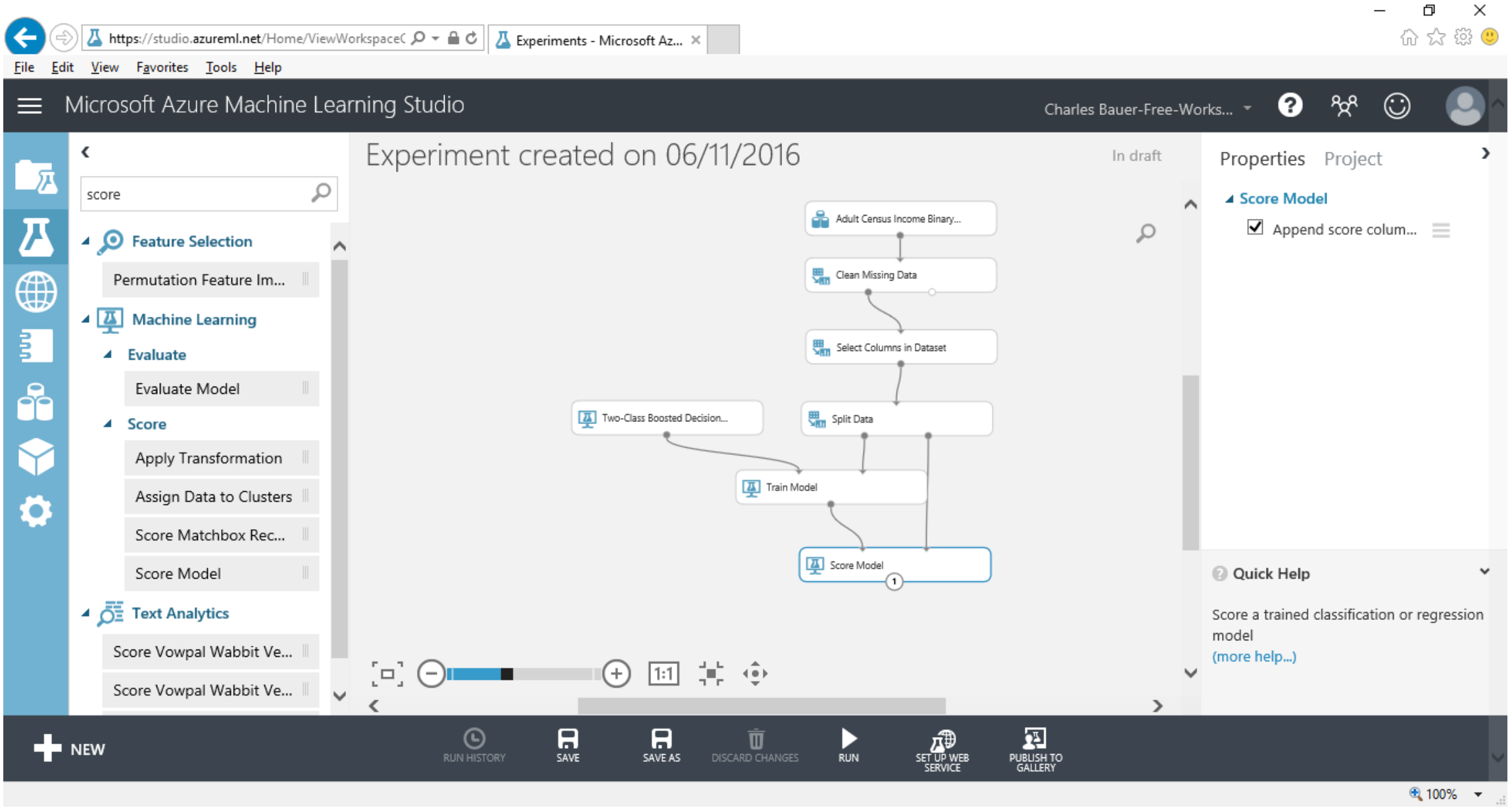
ML Studio: Escolhendo o algoritmo



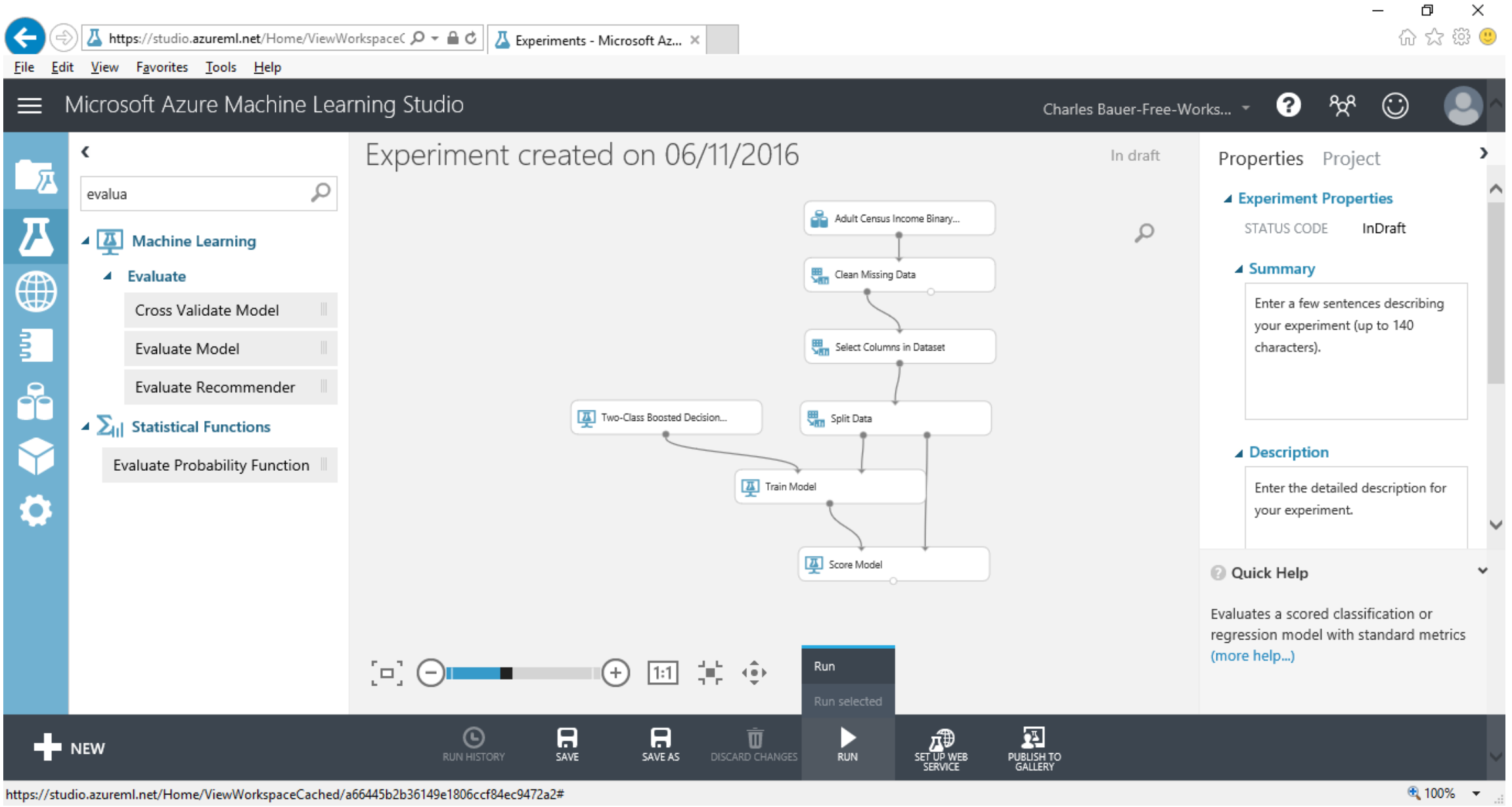
ML Studio: Escolhendo qual coluna queremos prever



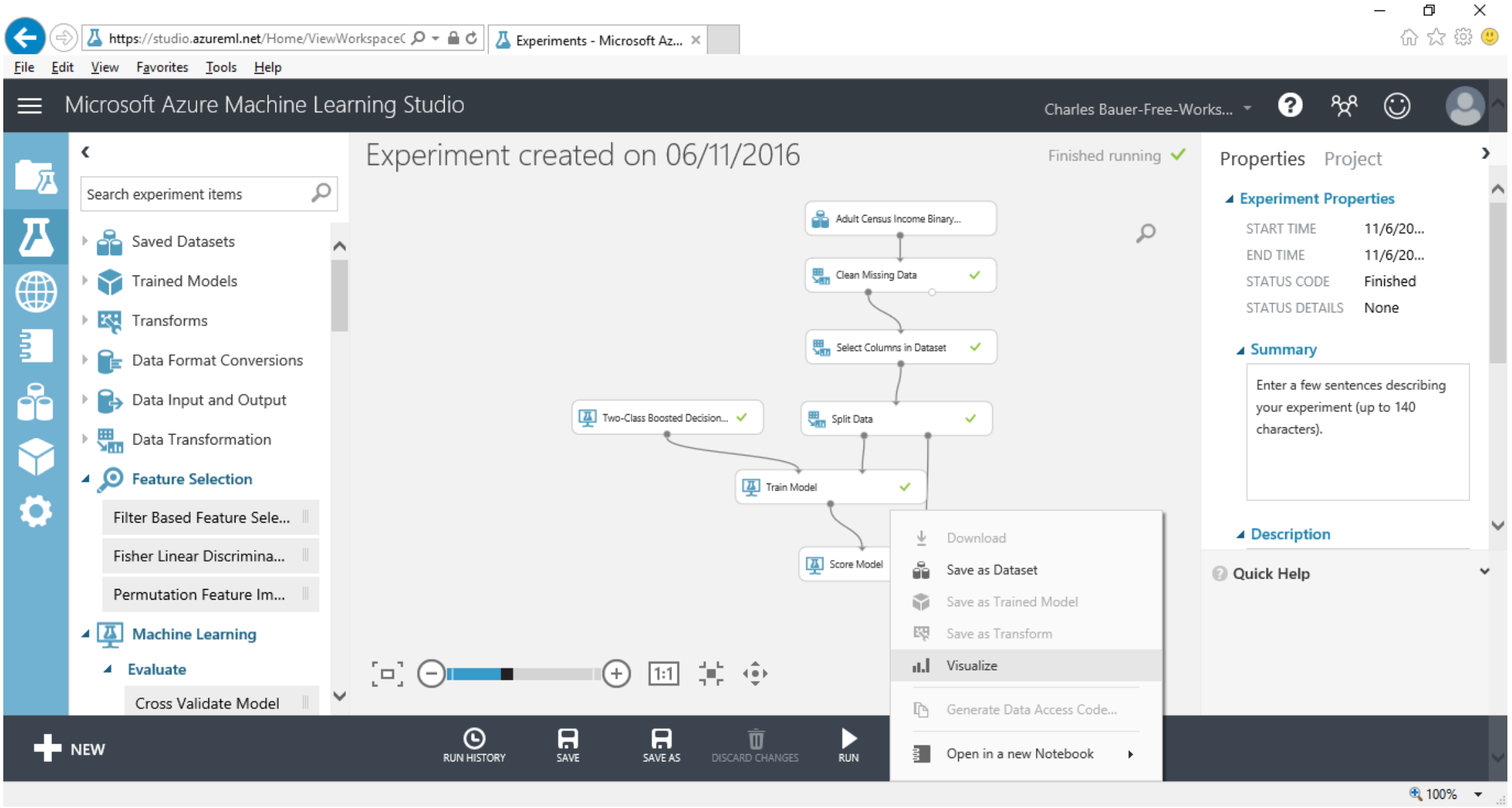
ML Studio: Pontuando o modelo



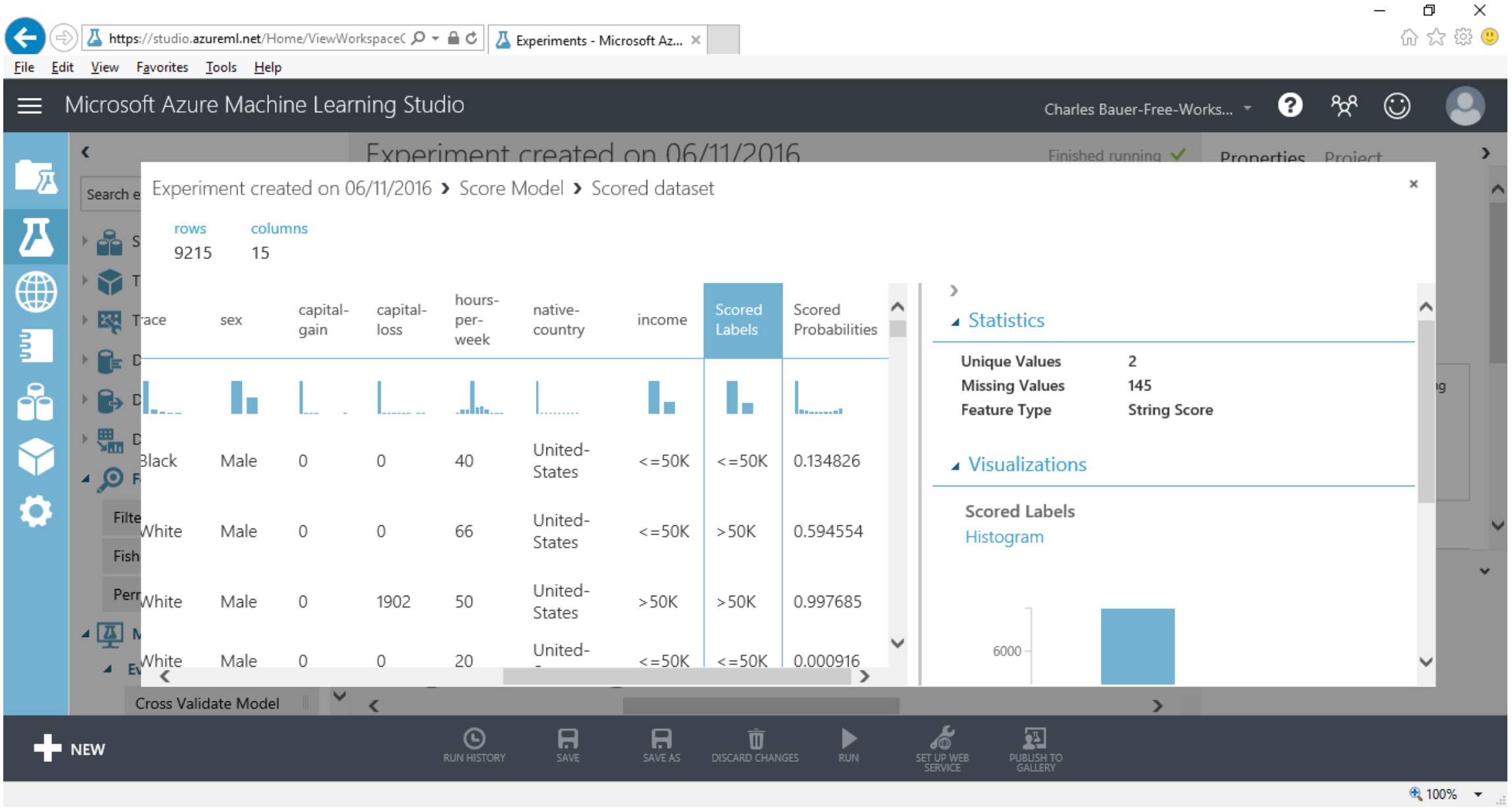
ML Studio: Pontuando o modelo



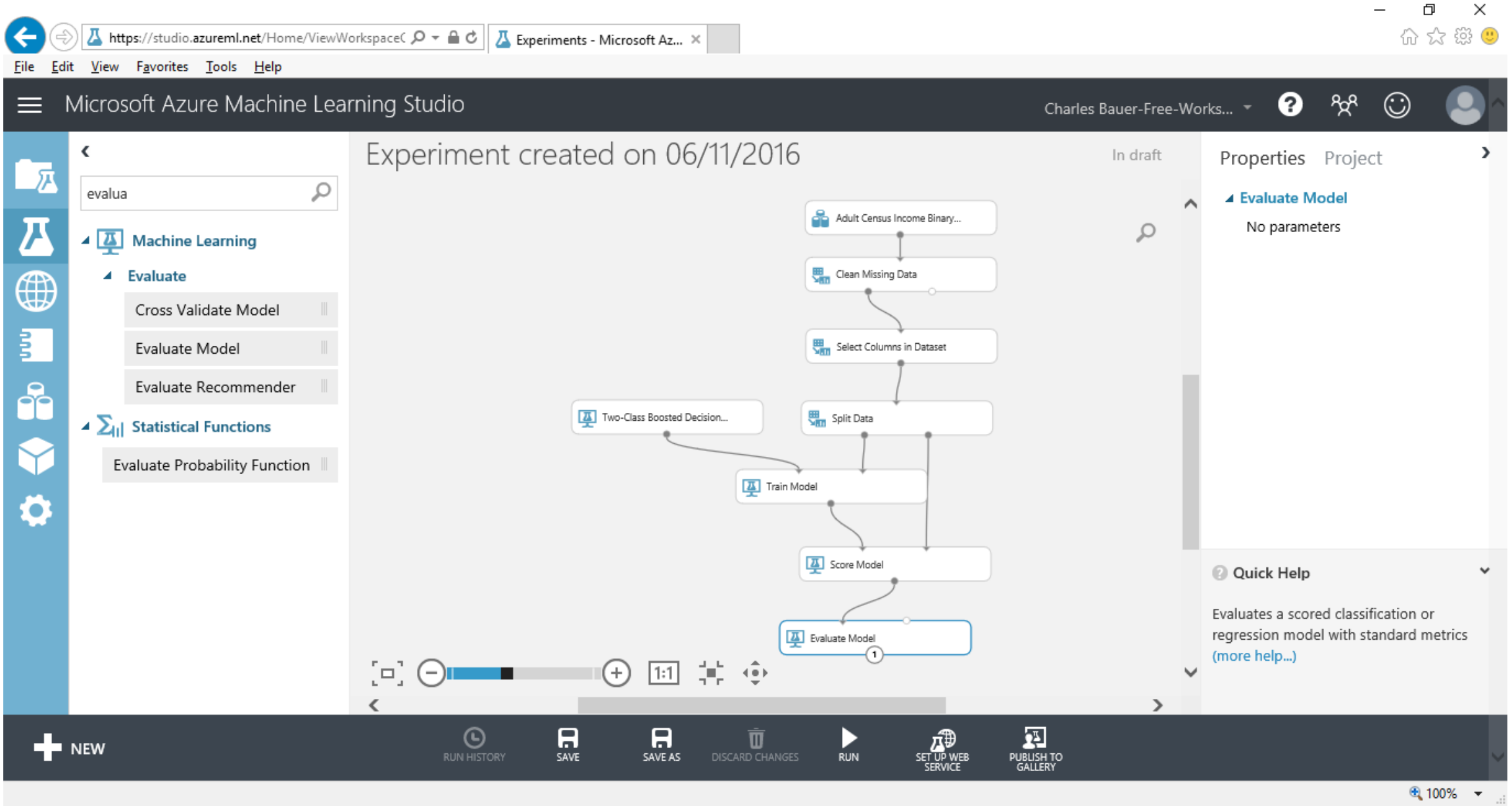
ML Studio: Pontuando o modelo



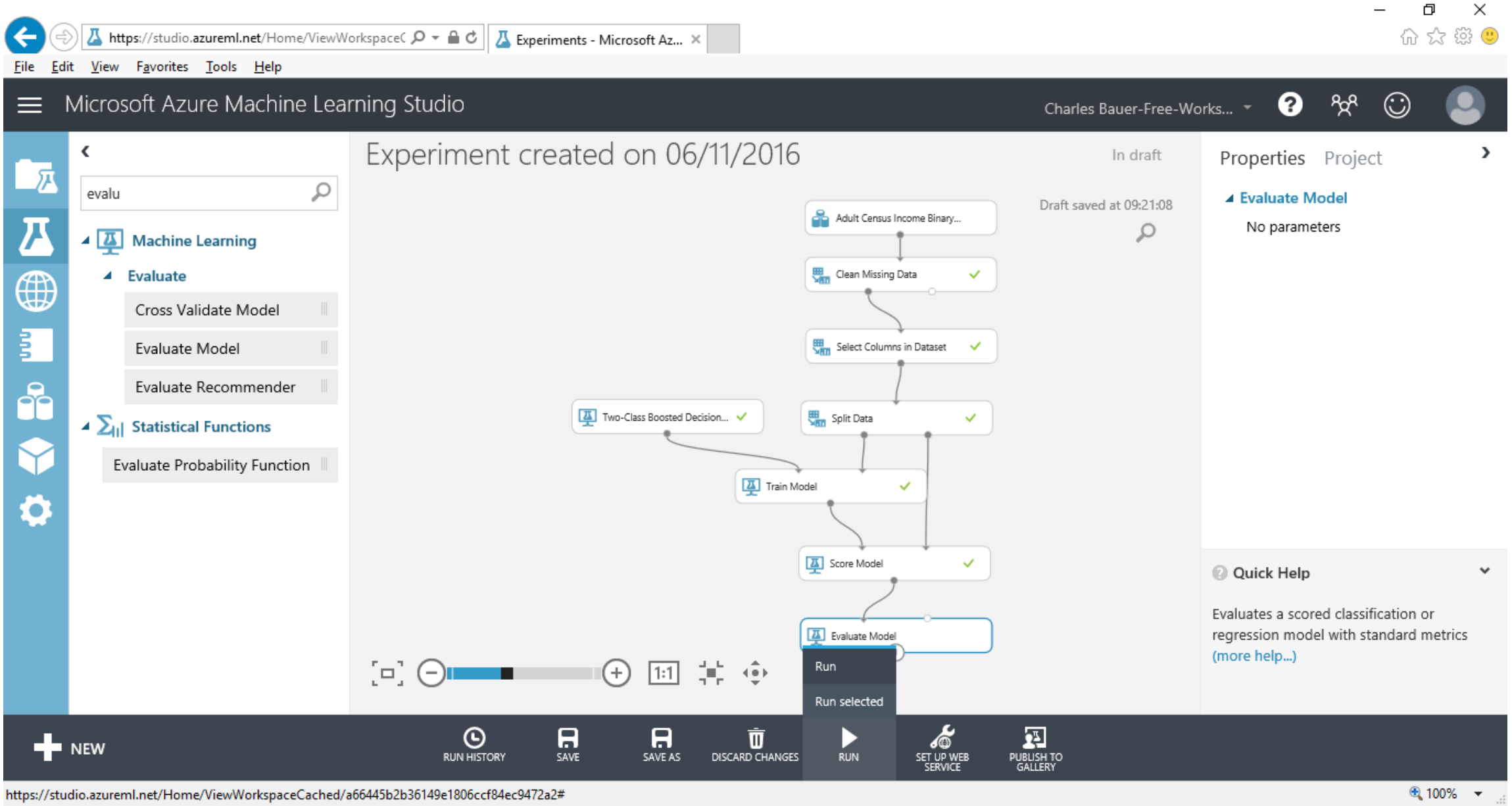
ML Studio: Pontuando o modelo



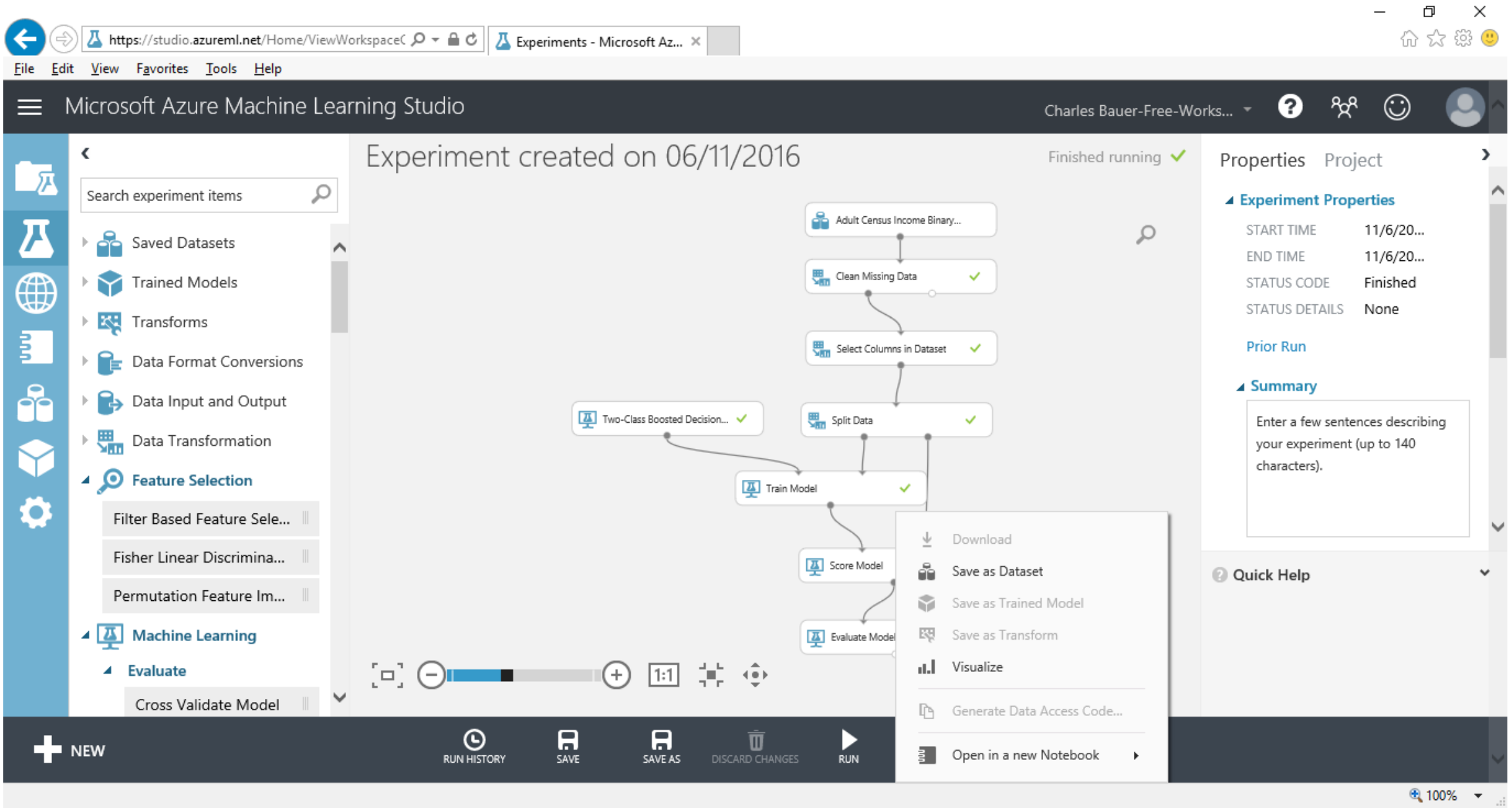
ML Studio: Pontuando o modelo



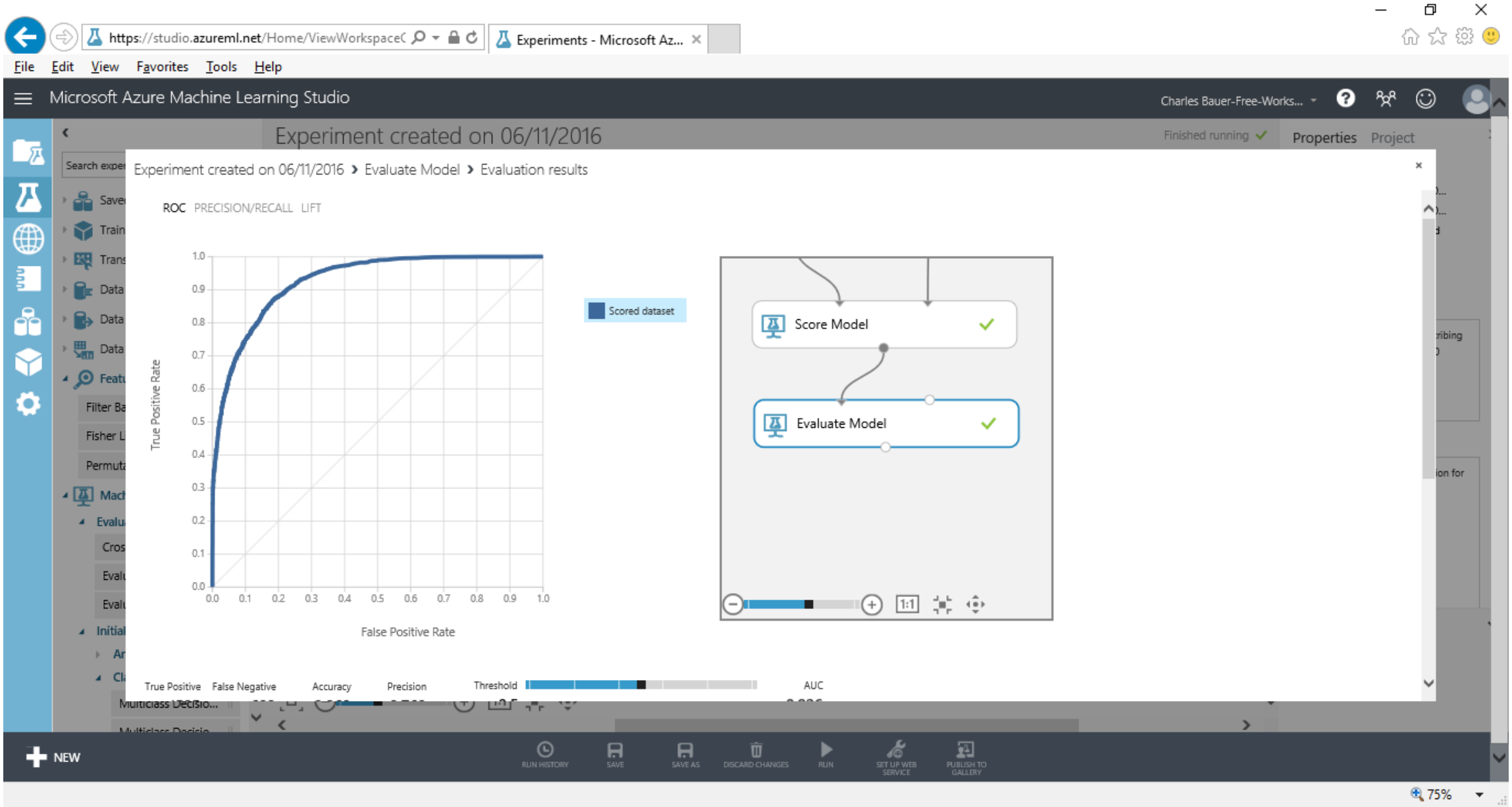
ML Studio: Pontuando o modelo



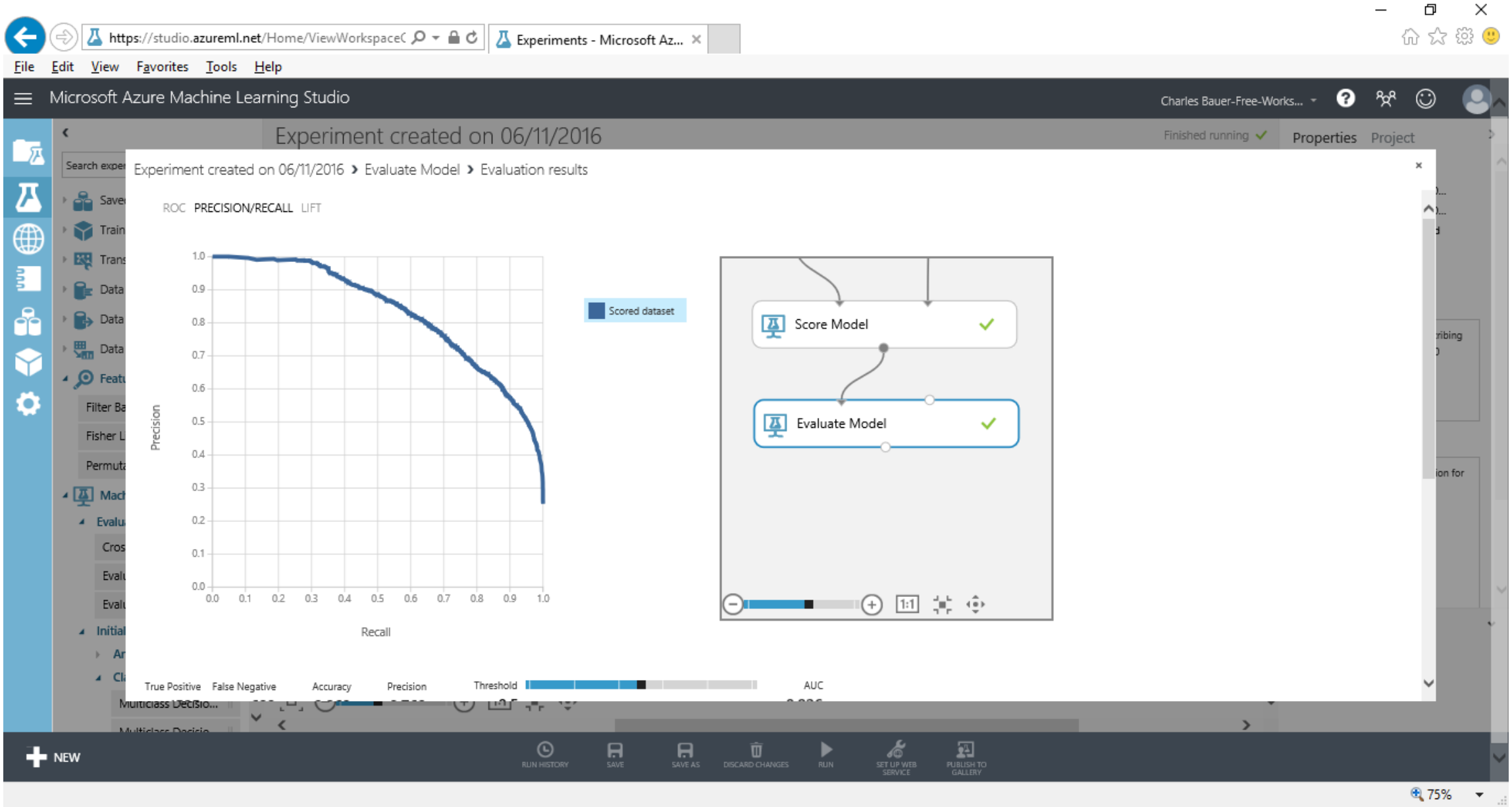
ML Studio: Pontuando o modelo



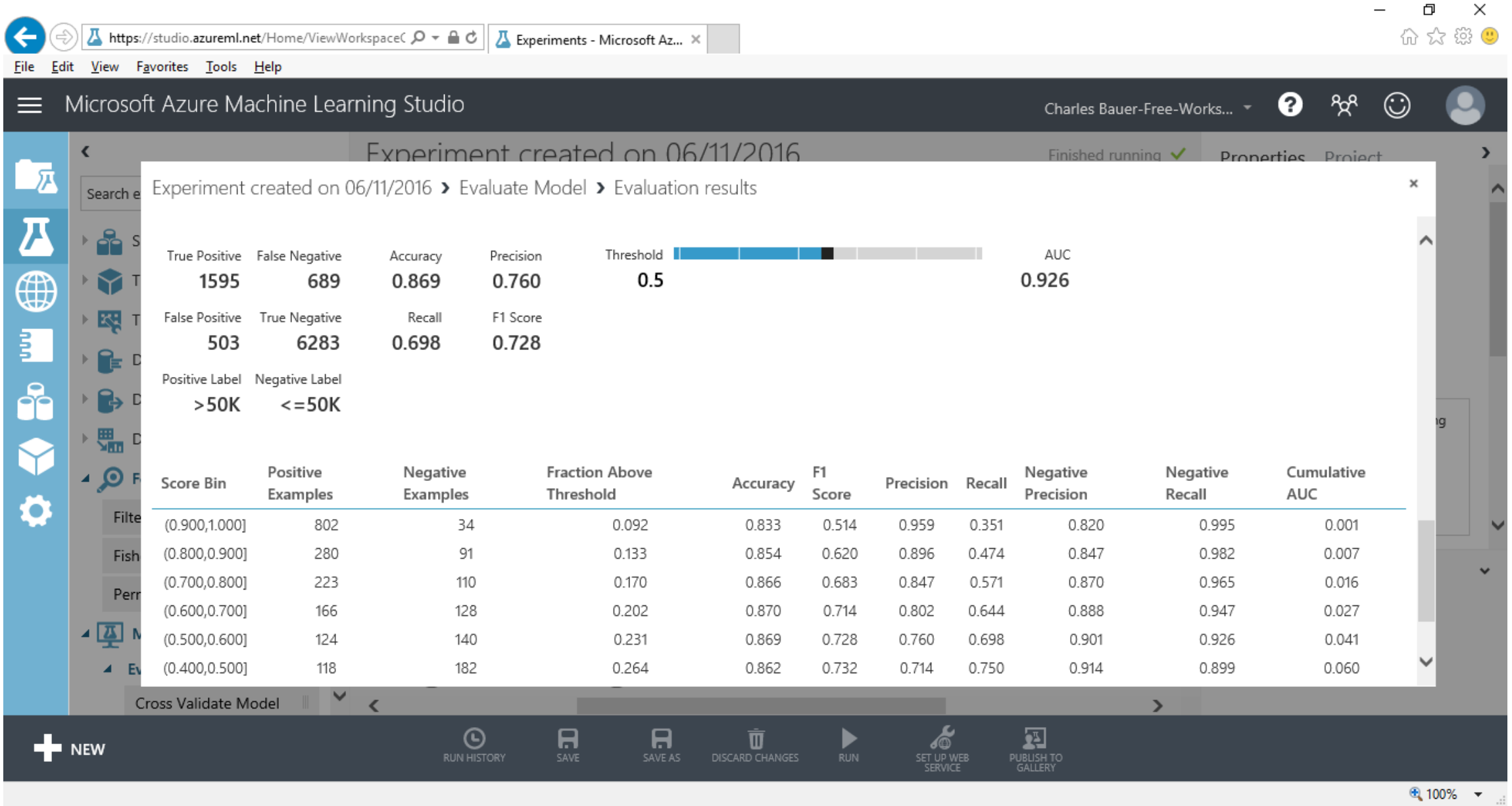
ML Studio: Pontuando o modelo



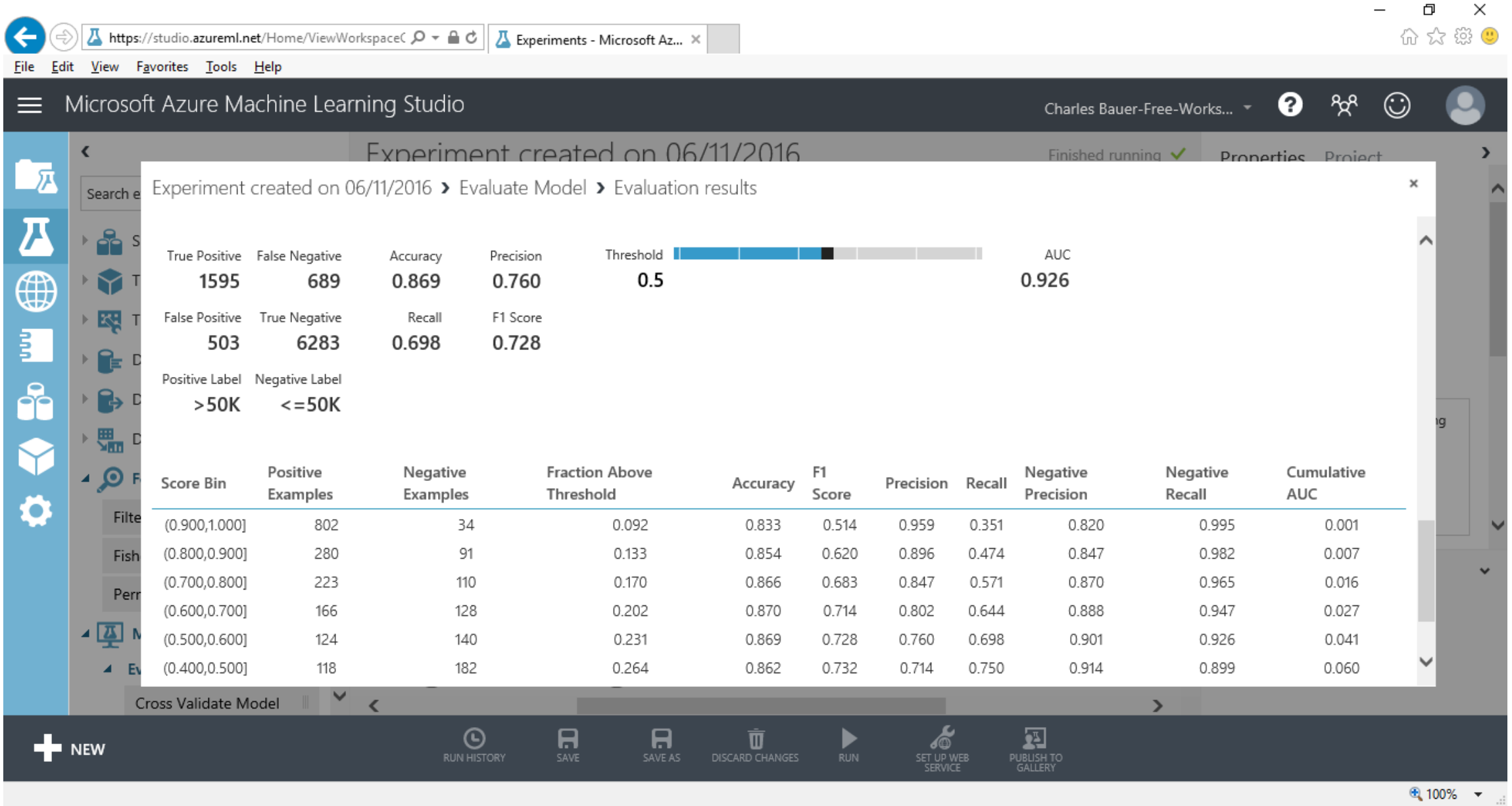
ML Studio: Pontuando o modelo



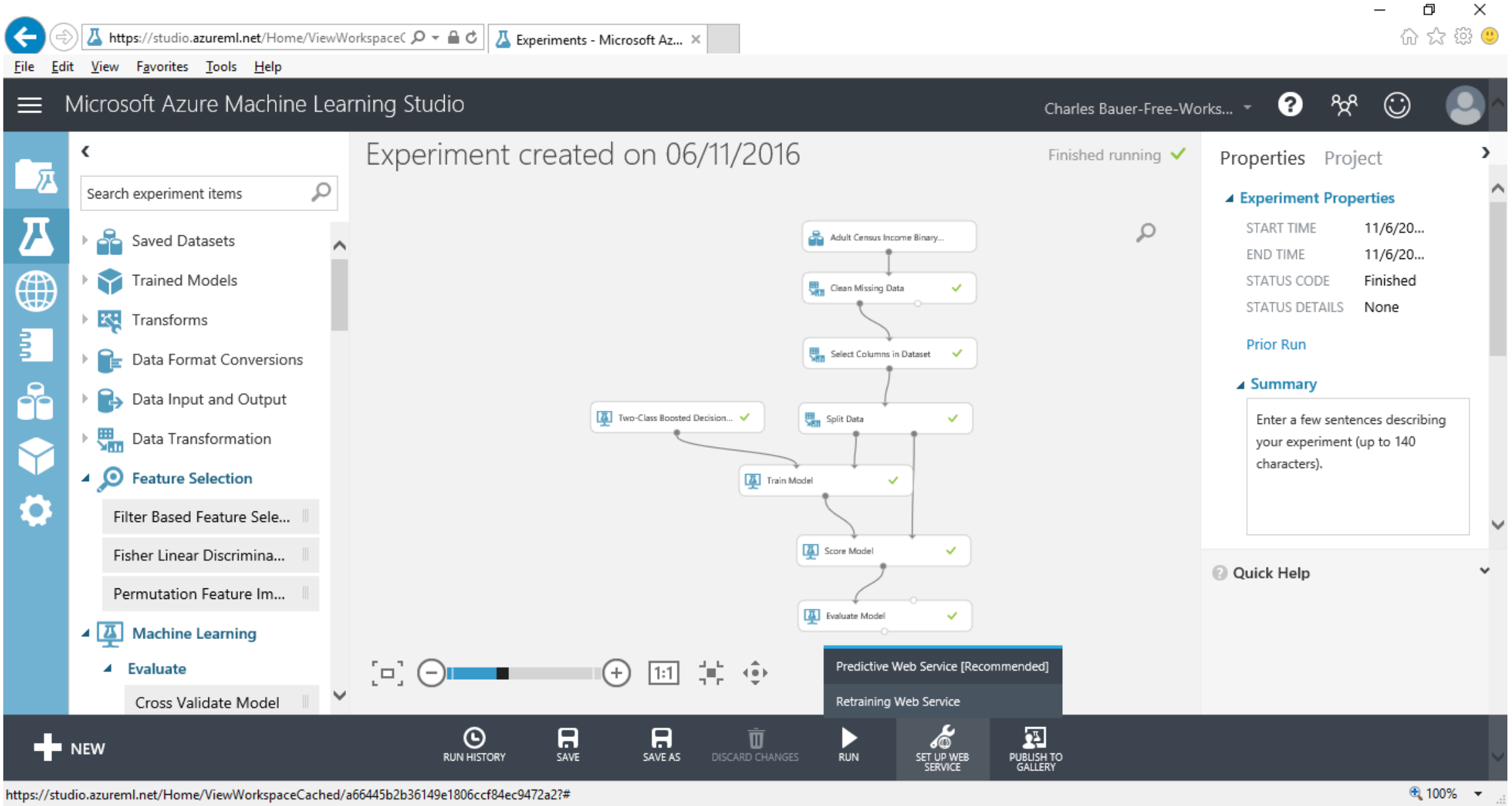
ML Studio: Pontuando o modelo



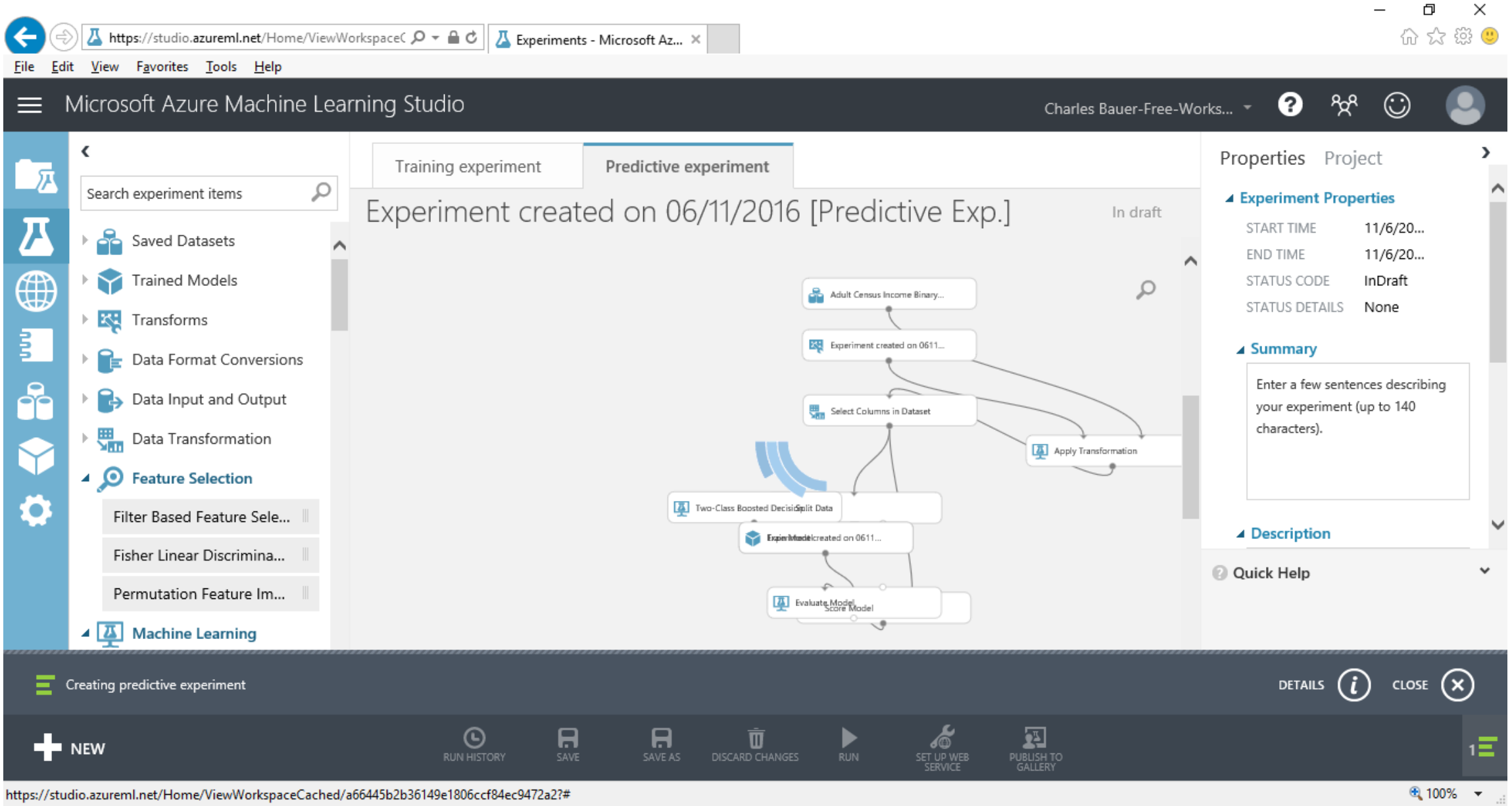
ML Studio: Pontuando o modelo



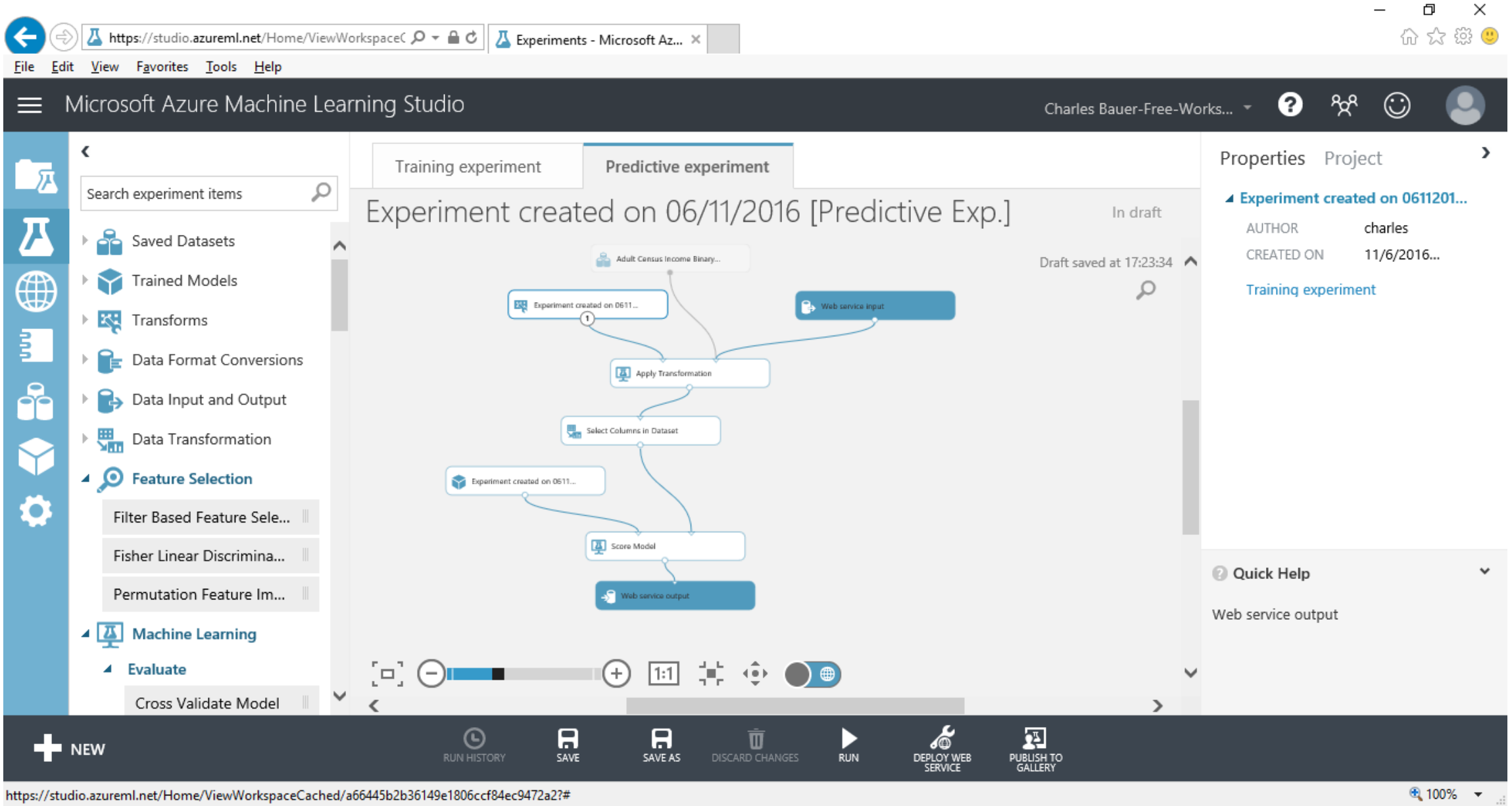
ML Studio: Preparando o Webservice



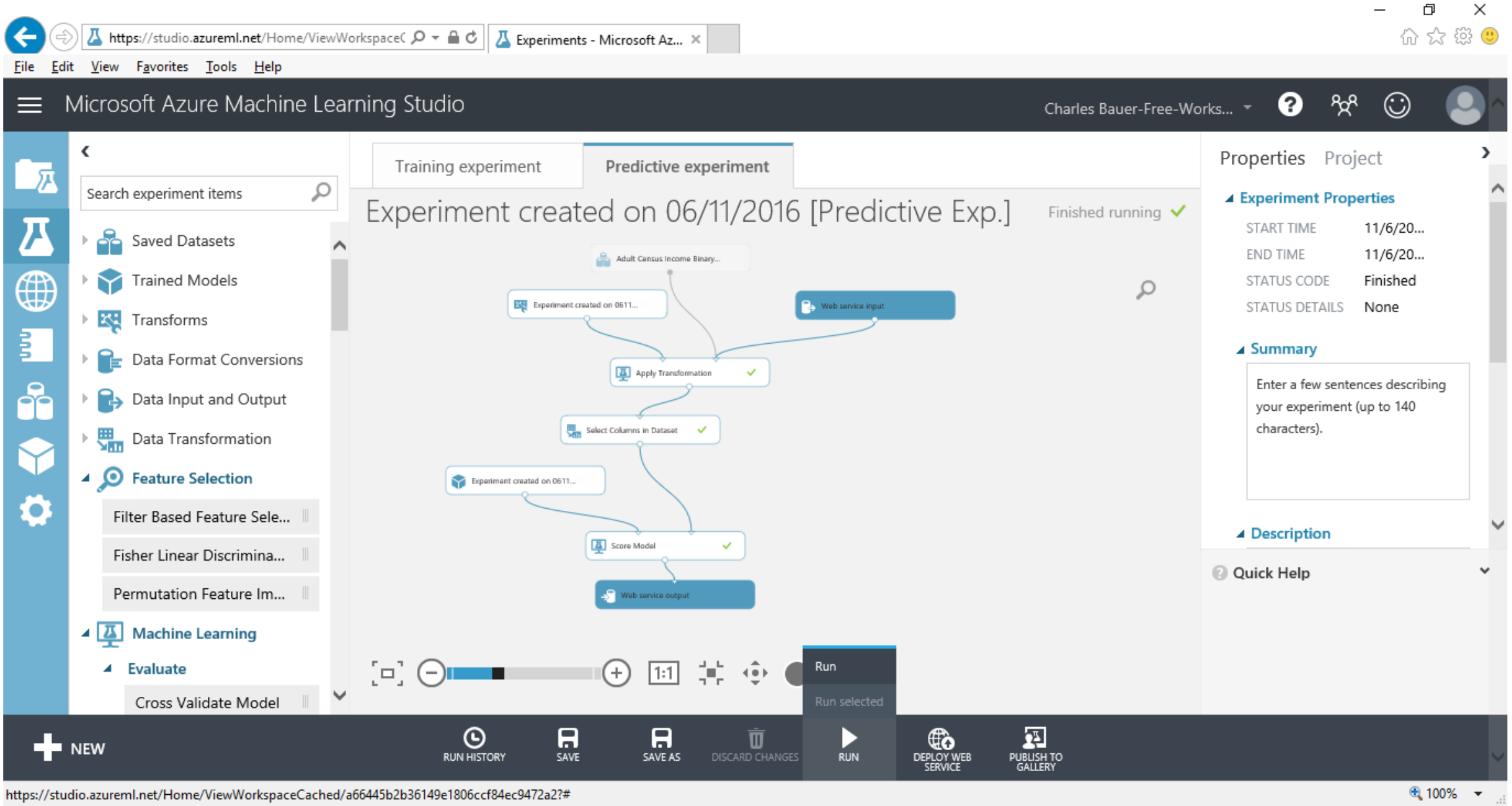
ML Studio: Preparando o Webservice



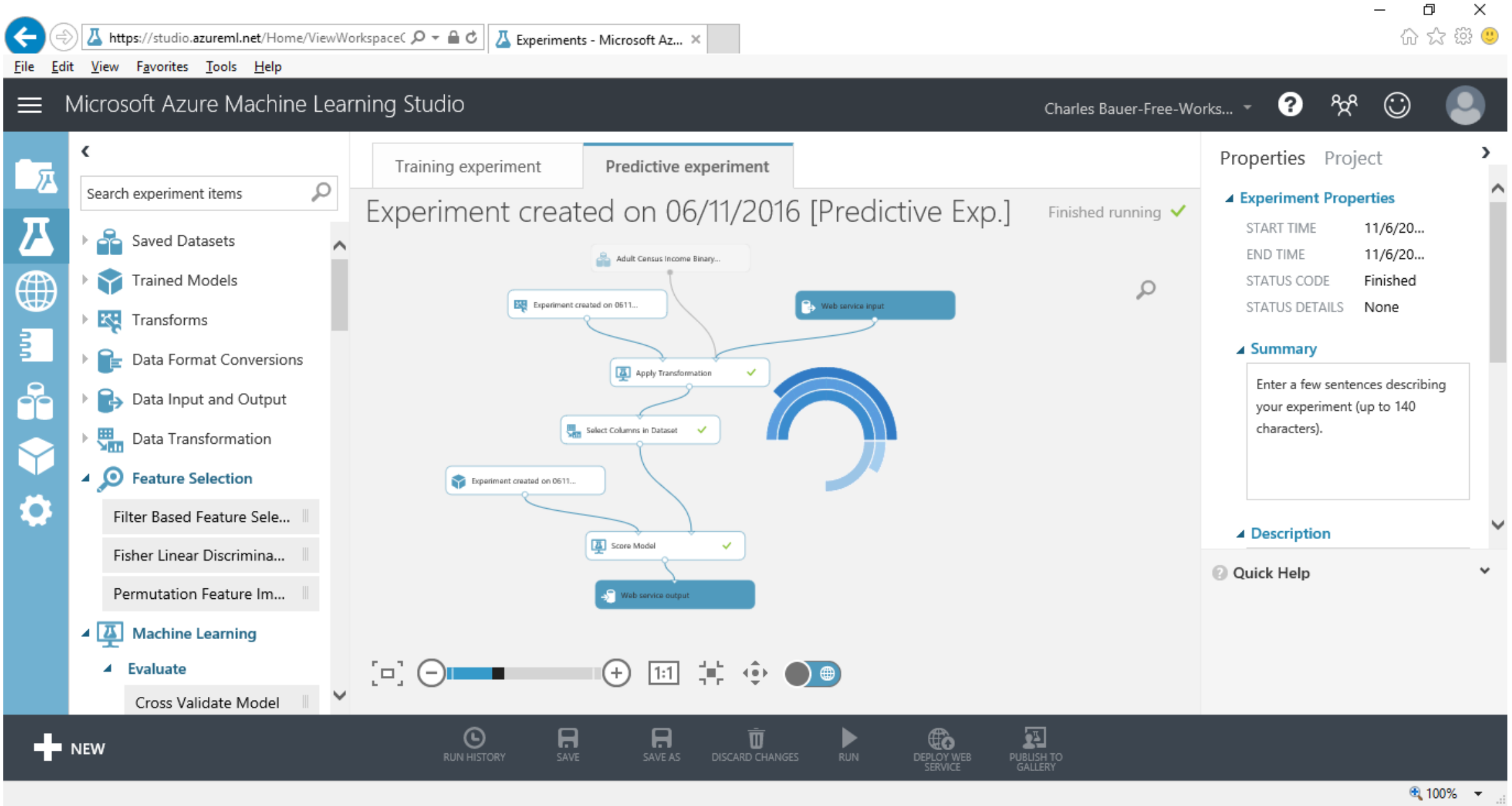
ML Studio: Preparando o Webservice



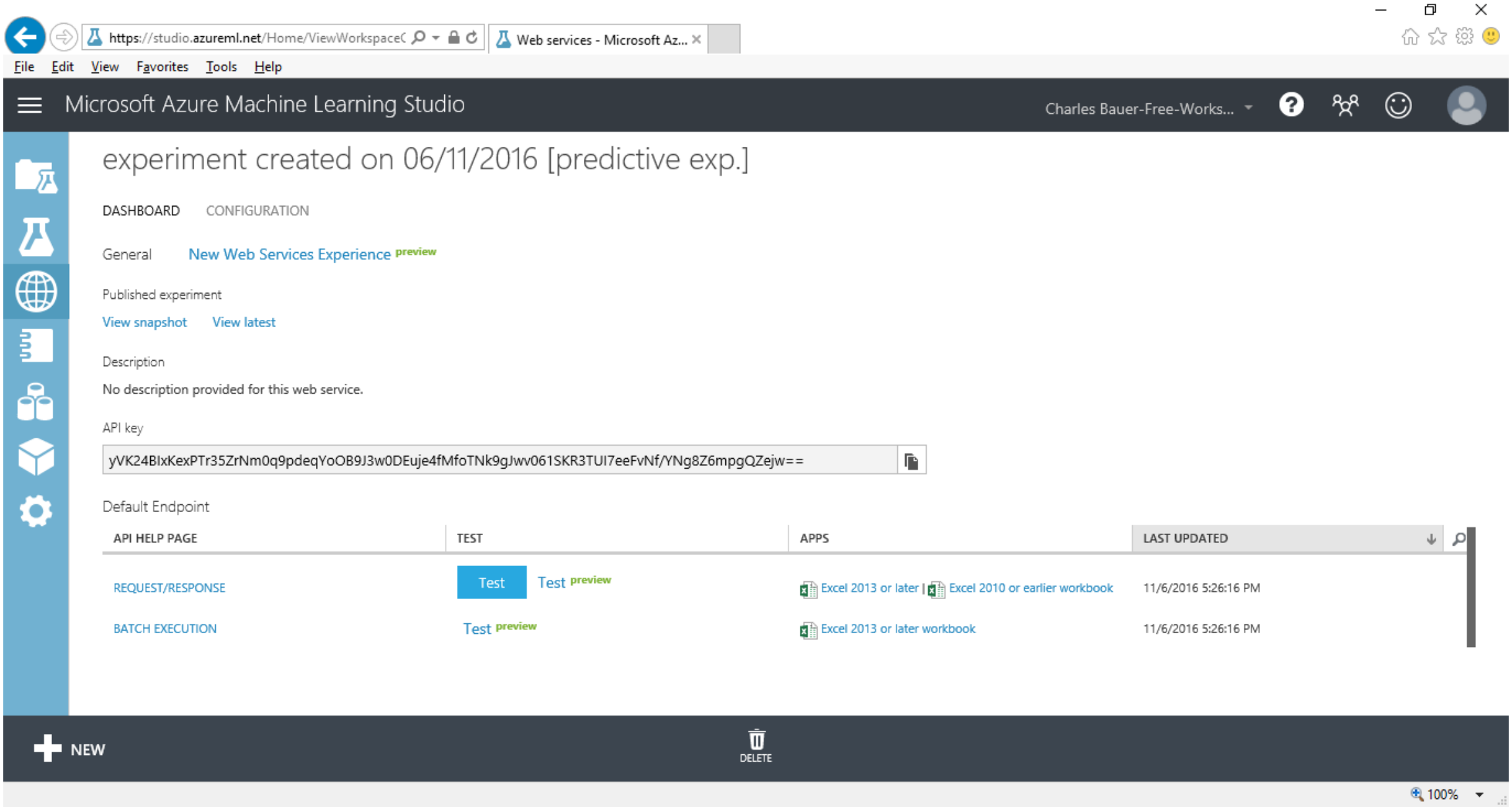
ML Studio: Preparando o Webservice para Publicação



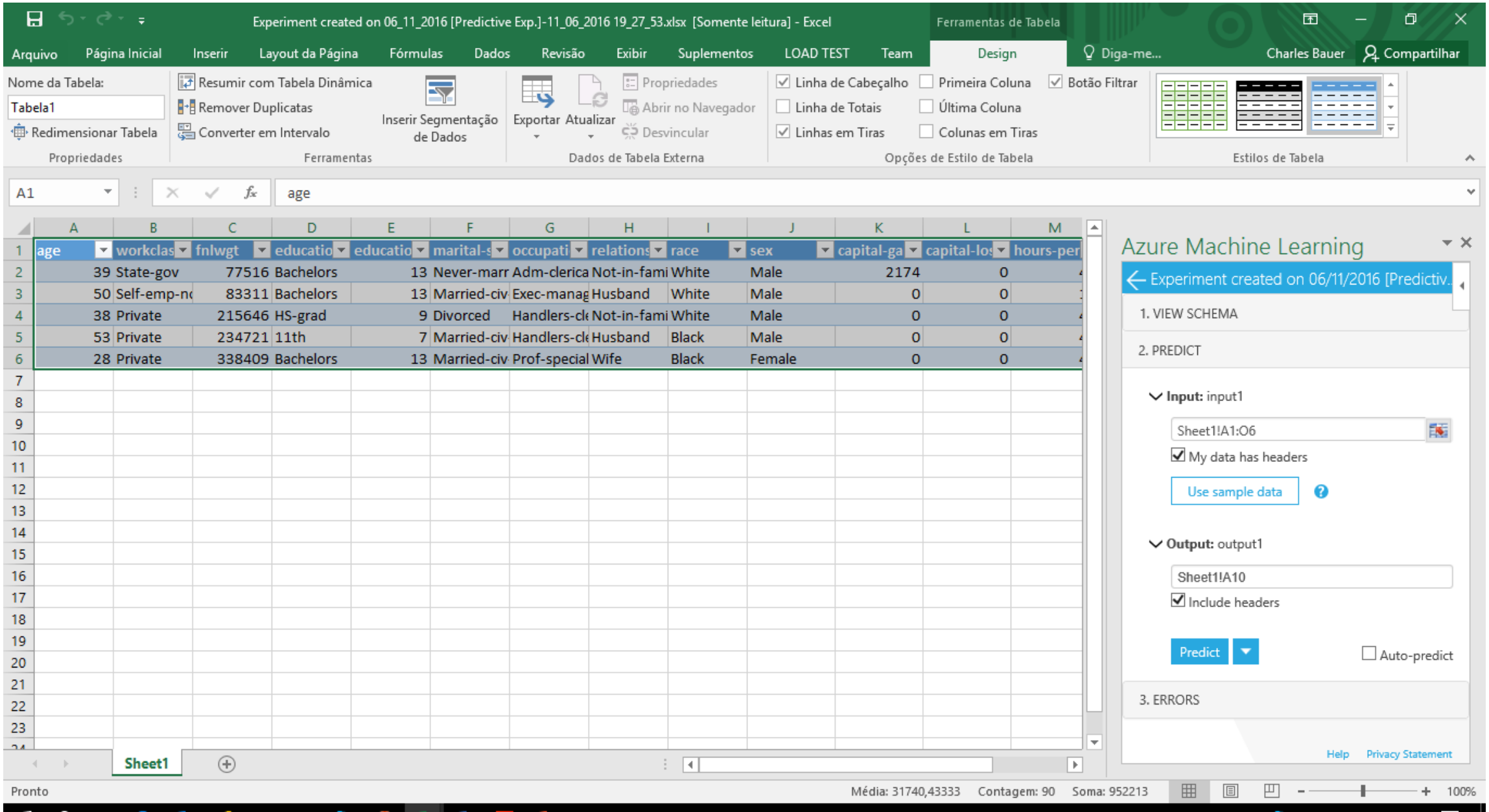
ML Studio: Publicando o WebService



ML Studio: Informações do WebService

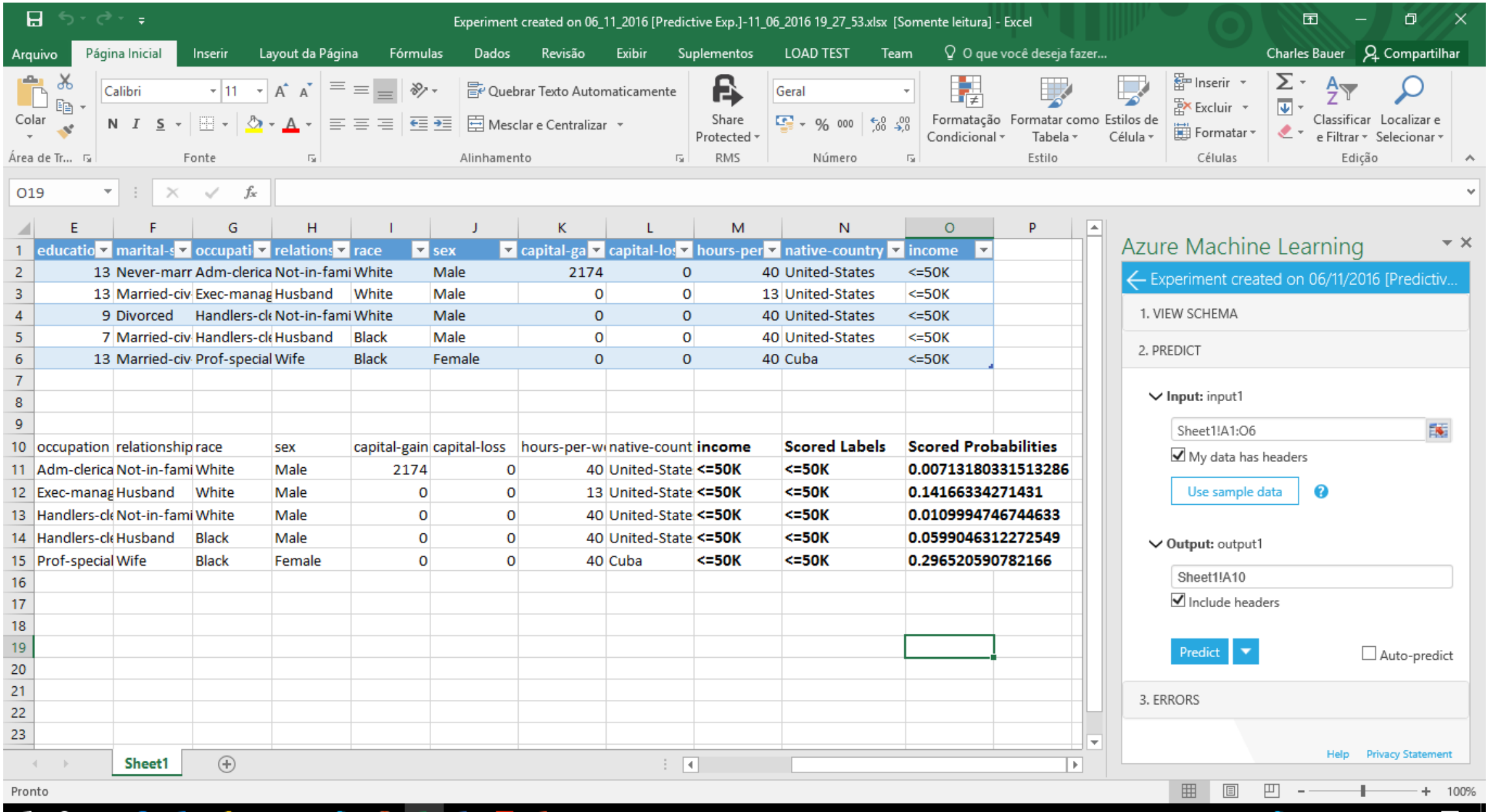


ML Studio: Submetendo dados dos Excel



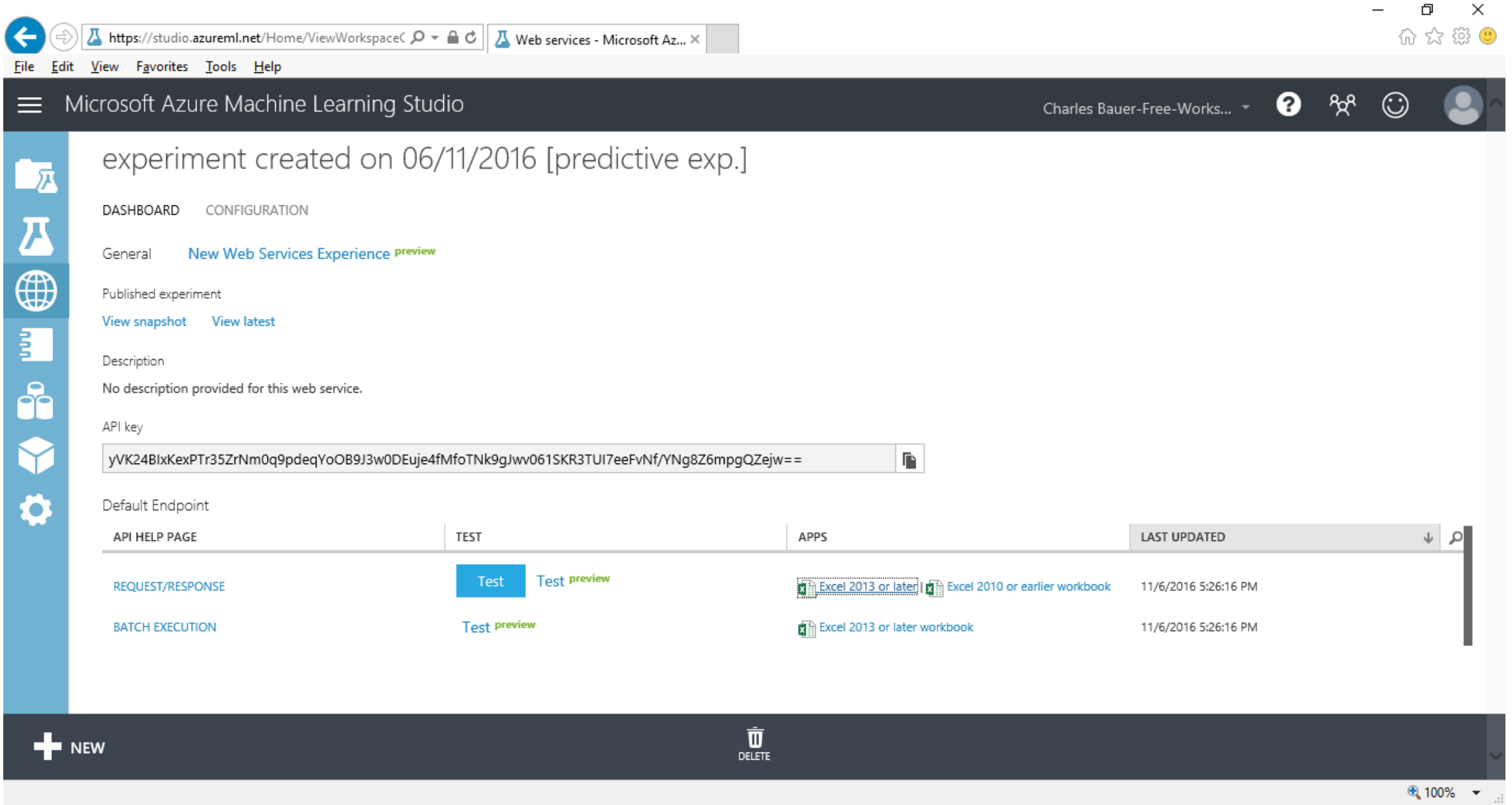
age	workclas	fnlwgt	educatio	educatio	marital-s	occupati	relations	race	sex	capital-ga	capital-los	hours-per
39	State-gov	77516	Bachelors	13	Never-marr	Adm-clerica	Not-in-fami	White	Male	2174	0	
50	Self-emp-nc	83311	Bachelors	13	Married-civ	Exec-manag	Husband	White	Male	0	0	
38	Private	215646	HS-grad	9	Divorced	Handlers-cl	Not-in-fami	White	Male	0	0	
53	Private	234721	11th	7	Married-civ	Handlers-cl	Husband	Black	Male	0	0	
28	Private	338409	Bachelors	13	Married-civ	Prof-special	Wife	Black	Female	0	0	

ML Studio: Submetendo dados dos Excel

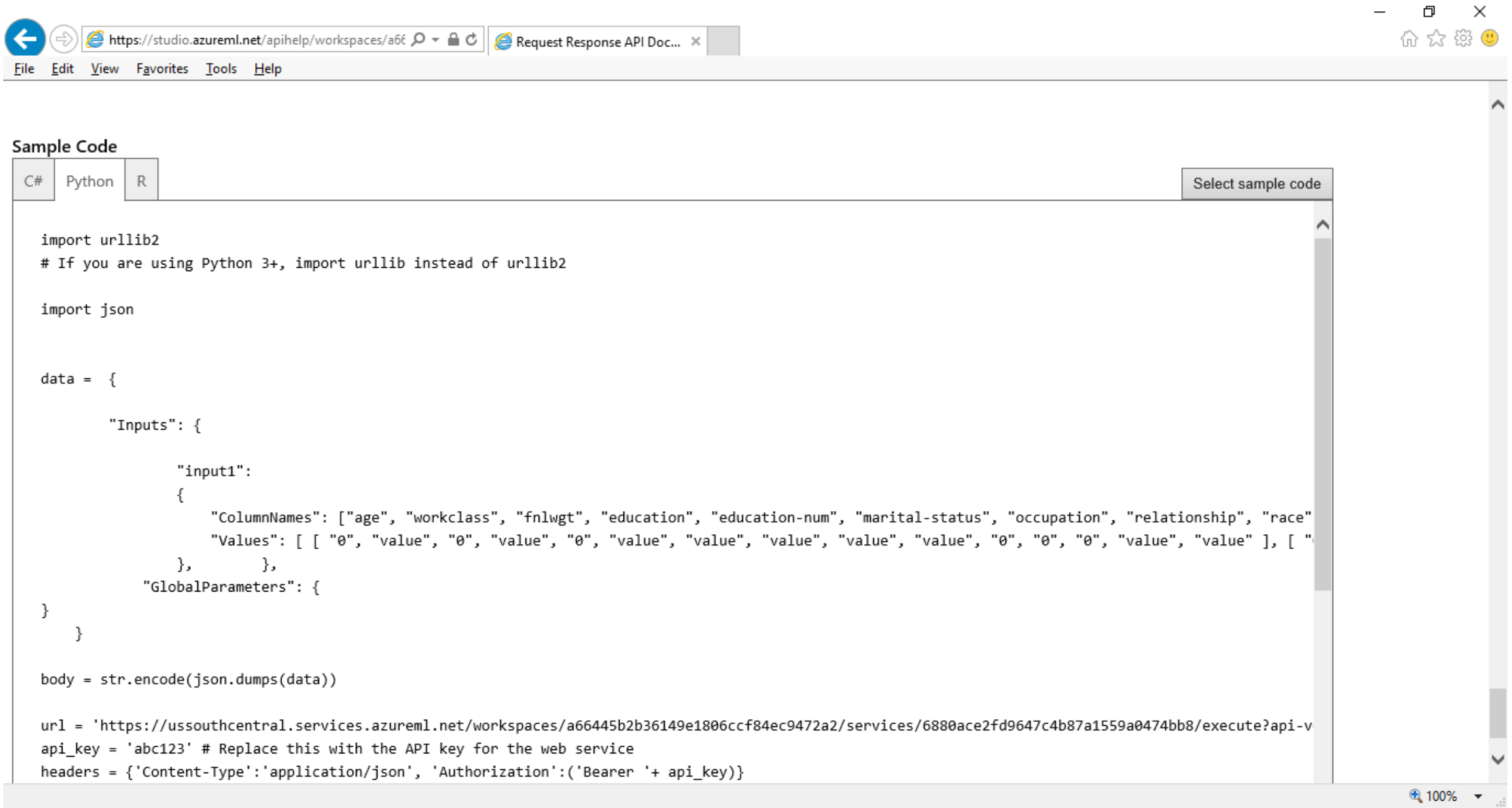


	E	F	G	H	I	J	K	L	M	N	O	P
1	educatio	marital-s	occupati	relations	race	sex	capital-ga	capital-los	hours-per	native-country	income	
2	13	Never-marr	Adm-clerica	Not-in-fami	White	Male	2174	0	40	United-States	<=50K	
3	13	Married-civ	Exec-manag	Husband	White	Male	0	0	13	United-States	<=50K	
4	9	Divorced	Handlers-cl	Not-in-fami	White	Male	0	0	40	United-States	<=50K	
5	7	Married-civ	Handlers-cl	Husband	Black	Male	0	0	40	United-States	<=50K	
6	13	Married-civ	Prof-special	Wife	Black	Female	0	0	40	Cuba	<=50K	
7												
8												
9												
10	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-w	native-count	**income**	**Scored Labels**	**Scored Probabilities**	
11	Adm-clerica	Not-in-fami	White	Male	2174	0	40	United-State	**<=50K**	**<=50K**	**0.00713180331513286**	
12	Exec-manag	Husband	White	Male	0	0	13	United-State	**<=50K**	**<=50K**	**0.14166334271431**	
13	Handlers-cl	Not-in-fami	White	Male	0	0	40	United-State	**<=50K**	**<=50K**	**0.0109994746744633**	
14	Handlers-cl	Husband	Black	Male	0	0	40	United-State	**<=50K**	**<=50K**	**0.0599046312272549**	
15	Prof-special	Wife	Black	Female	0	0	40	Cuba	**<=50K**	**<=50K**	**0.296520590782166**	
16												
17												
18												
19												
20												
21												
22												
23												

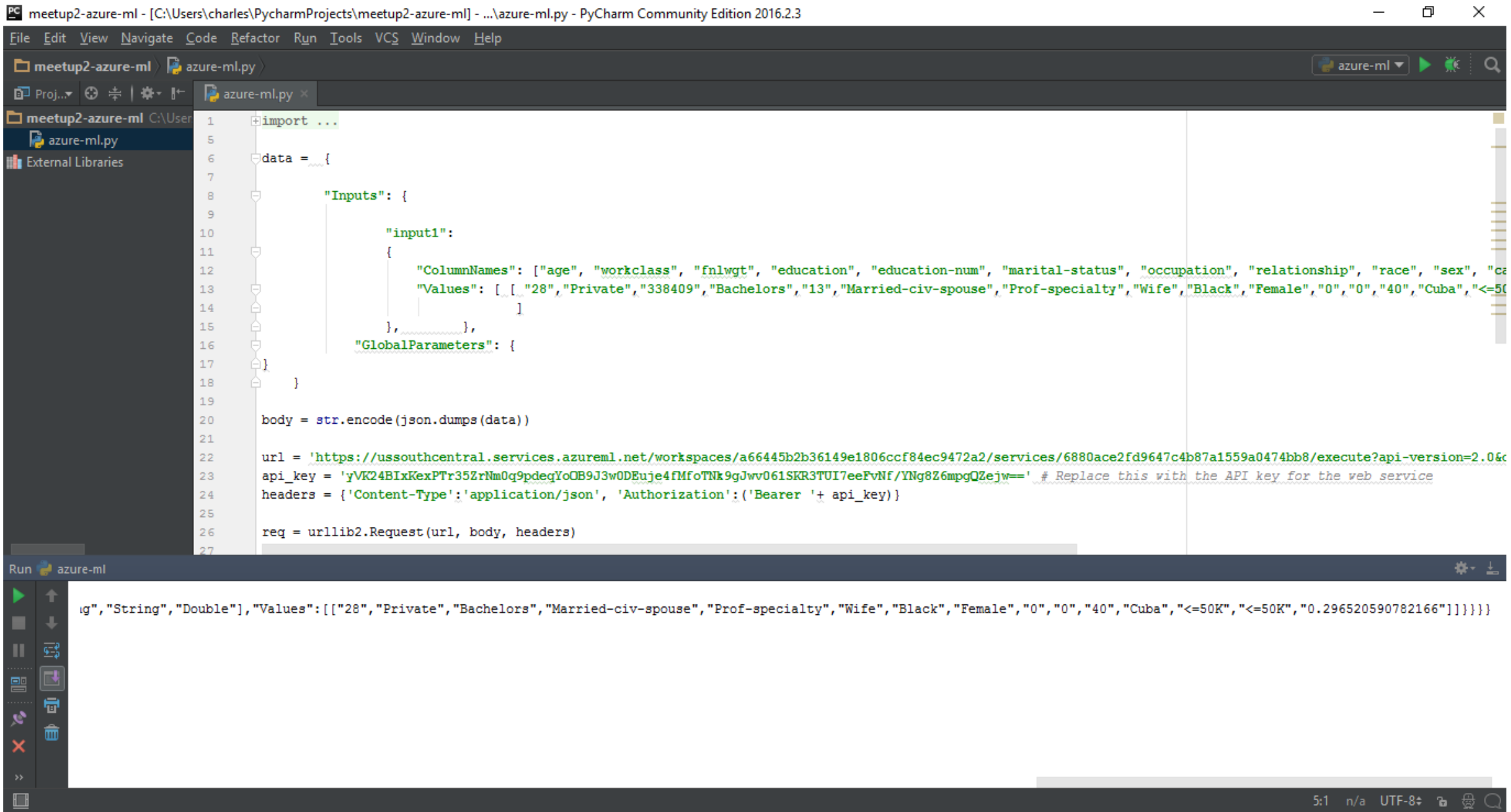
ML Studio: Usando o modelo em apps Python



ML Studio: Usando o modelo em apps Python



ML Studio: Usando o modelo em apps Python



Obrigado!

CHARLESBAUER@GMAIL.COM

+55 47-99125-6666