

23 Y 24 DE JUNIO

# HACKATHON IA 2022

## HACK THE GAP

CERRANDO LA BRECHA DE STEM\*

→ CONVOCATORIA

\*SCIENCE, TECHNOLOGY, ENGINEERING AND MATH  
\*CIENCIA, TECNOLOGÍA, INGENIERÍA Y MATEMÁTICAS

# Análisis de Riesgo Crediticio

## Introducción:

- a) Le proporcionaremos dos archivos “csv” que serán sus fuentes de datos:

[https://github.com/peterwtf/Taller\\_RiesgoCrediticio](https://github.com/peterwtf/Taller_RiesgoCrediticio)

- Riesgo Crediticio - Histórico. Contiene información histórica de clientes que han solicitado créditos y su comportamiento.
- Riesgo Crediticio – Solicitantes. Contiene información de personas que están solicitando un nuevo crédito y que decidiremos si se los otorgamos o no.

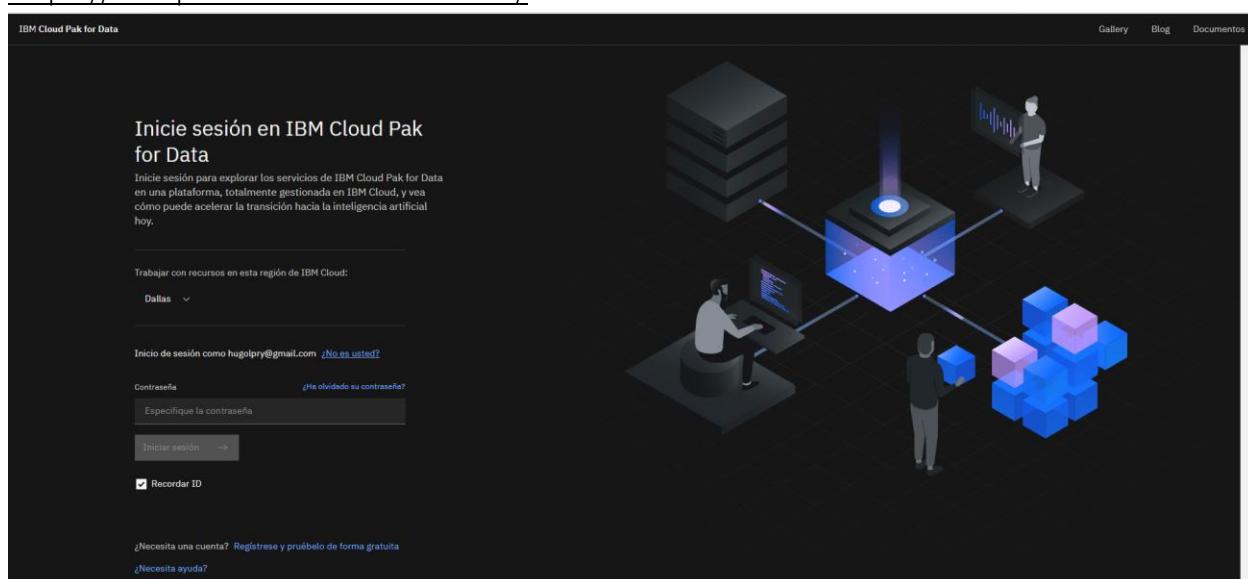
- b) Debe cargar sus datos en la plataforma (archivos del punto a)

## Parte 1

### 1. Entrar a la plataforma de Cloud Pak for Data

Entrar a Cloud Pak for Data as a Service en la siguiente liga:

<https://dataplatform.cloud.ibm.com/>



1.a) Si no tiene cuenta vaya a la sección “¿Necesita una cuenta?” (parte baja izquierda de la pantalla) y cree su cuenta siguiendo las instrucciones, al definir su cuenta elija región “Dallas”.

1.b) Si ya tiene cuenta, proporcione su usuario/password y entre a la plataforma.

1.c) Una vez que haya entrado a la plataforma verá una pantalla como la siguiente:

The screenshot shows the IBM Cloud Pak for Data platform interface. At the top, there's a navigation bar with links for 'Todos', 'Buscar', 'Comprar', and a user account section. The main header says '¡Bienvenido, Pedro!'. Below the header, there are three main sections: 'Aprender mediante ejemplo' (with a link to 'Recorrido de pasos de resolución de problemas empresariales específicos en un proyecto de ejemplo.'), 'Trabajar con datos' (with a link to 'Crear un proyecto para que su equipo准备 datos, busque conocimientos o cree modelos.'), and 'Amplie sus prestaciones' (with a link to 'Añada herramientas, bases de datos u otras características creando instancias de servicios.'). To the right of these sections is a large, stylized graphic of a smartphone displaying data cubes and a waveform. On the left side, there's a 'Quick start' sidebar with links to various services like 'Build dashboards', 'Create data pipelines', 'Build customer profiles', etc. The central area features a 'Visión general' (General View) dashboard with three main cards: 'Proyectos recientes' (Recent Projects) showing items like 'Data Fabric Outcomes', 'peter', 'IA Taller', 'Jobs\_Puebas', and 'Data Virtualization'; 'Catálogos recientes' (Recent Catalogs) showing 'CreditDataCatalog' and 'Platform assets catalog'; and 'Notificaciones' (Notifications) showing recent activity like 'Project import complete' and 'Data Steward removed BusinessUser CPD from project Data Fabric Outcomes Pedro'. At the bottom left, there's a 'Soporte' (Support) section with links to 'Documentación', 'Preguntas más frecuentes', and 'Compartir una idea'.

## 2. Aprovisionar los servicios que utilizaremos para el ejercicio.

2.a) Ir al menú de “Hamburguesa” (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”

The screenshot shows the IBM Cloud Pak for Data interface. On the left, there's a navigation sidebar with various options like Inicio, Datos, Proyectos, Catálogos, Gobernabilidad, Despliegues, Servicios, and Catalogo de servicios. The 'Catalogo de servicios' option is highlighted with a red box. The main area has a dark background with a central graphic of data cubes and a magnifying glass. It includes sections for 'Trabajar con datos' (Create a project for your team to prepare data, search for knowledge or create models) and 'Amplie sus prestaciones' (Add tools, data bases and other characteristics by creating instances of services). Below these are buttons for 'Crear un proyecto' and 'Crear un servicio'. A 'Visión general' section displays 'Proyectos recientes' (Data Fabric Outcomes, 07 abr 2022), 'Catálogos recientes' (CreditDataCatalog, 22 feb 2022), and a 'Notificaciones' section with a green checkmark and the message 'Project import complete'.

2.b) Utilizando la opción de búsqueda (lupa) encuentre el servicio “cloud object storage” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).

This screenshot shows the 'Catalogo de servicios' page with a search bar containing 'Cloud object storage'. A red box highlights the search term. Below the search bar, a category 'Almacenamiento' is selected. A green box highlights the 'Cloud Object Storage' service card. The card shows the service name, category, a brief description ('Almacenamiento de datos habilitado para desarrolladores para aplicaciones en la nube integradas con servicios d...'), and a 'Lite • Gratuito' status.

2.c) Una vez dentro del servicio de Cloud Object Storage, seleccione la opción “IBM cloud”, “Lite” (color rojo) y elija la opción de crear (color verde).

The screenshot shows the IBM Cloud interface for creating a Cloud Object Storage service. The 'Lite' plan is selected and highlighted with a red box. The 'Crear' (Create) button is highlighted with a green box.

2.d) Si el servicio de Cloud Object Storage fue creado correctamente deberá ver una pantalla como la siguiente:

The screenshot shows the IBM Cloud Pak for Data interface displaying a list of service instances. A single instance named "CloudObjectStorage" is listed in the table.

Nombre	Grupo	Ubicación	Producto	Plan	Estado
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Activas

2.e) Vaya de nuevo al menú de “Hamburguesa” nuevamente (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”

The screenshot shows the IBM Cloud Pak for Data interface. On the left, there's a navigation sidebar with various options like 'Inicio', 'Datos', 'Proyectos', 'Catálogos', 'Gobernabilidad', 'Despliegues', 'Servicios', and 'Catalogo de servicios'. The 'Catalogo de servicios' option is highlighted with a red box. The main area displays a dashboard with sections for 'Trabajar con datos' (Working with data) and 'Amplie sus prestaciones' (Expand your offerings). Below these are buttons for 'Crear un proyecto' (Create a project) and 'Crear un servicio' (Create a service). A large graphic on the right shows a computer monitor displaying a chart, with several 3D cubes of different colors (blue, green, purple) floating around it, representing data and AI components.

2.f) Utilizando la opción de búsqueda (lupa) encuentre el servicio “Watson Studio” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).

The screenshot shows the 'Catalogo de servicios' (Services Catalog) search results. In the search bar at the top left, the text 'Watson Studio' is entered and highlighted with a red box. Below the search bar, a list of categories is shown: 'IA / Aprendizaje automático', 'bases de datos', 'Integración', 'Analíticas', and 'Almacenamiento'. Under 'IA / Aprendizaje automático', there is a single result for 'Watson Studio', which is also highlighted with a green box. The result card for 'Watson Studio' includes the text 'Integre la IA y el aprendizaje automático en su negocio. Cree modelos personalizados utilizando su...', a 'Lite • Gratuito' button, and a small icon of a person with a gear.

2.g) Una vez dentro del servicio de Watson Studio, seleccionar una región “Dallas” además la opción “Lite” (color rojo) y elija la opción de crear (color verde).

Watson Studio

Autor: IBM • Fecha de última actualización: Mar 31, 2022 • Documentos

Crear Acerca de

Seleccionar una región

Seleccionar una región

Dallas

Plan de precios

Los precios mostrados no incluyen impuestos. Los precios mensuales que se muestran son para el país o región: United States

Plan	Características	Precios
Lite	<ul style="list-style-type: none"> <li>1 usuario autorizado</li> <li>Límite mensual de 50 unidades de capacidad por hora</li> <li>Entorno = Número de unidades de capacidad necesarias por hora           <ul style="list-style-type: none"> <li>• 1 vCPU + 4 GB RAM = 0,5</li> <li>• 2 vCPU + 8 GB RAM = 1</li> <li>• 4 vCPU + 16 GB RAM = 2</li> <li>• Optimización de decisiones = Entorno + 5</li> </ul> </li> </ul> <p>El plan Lite para Watson Studio le ofrece todo lo que necesita para convertirse en un mejor científico de datos o experto en dominios en un entorno de colaboración.</p> <p>Los servicios del plan Lite se suprime tras 30 días de inactividad.</p>	Gratis
Professional	<ul style="list-style-type: none"> <li>Unlimited collaborators</li> <li>Unlimited elastic compute environments</li> <li>Entorno = # of capacity units required per hour           <ul style="list-style-type: none"> <li>• 1 vCPU + 4 GB RAM = 0,5</li> <li>• 2 vCPU + 8 GB RAM = 1</li> </ul> </li> </ul>	\$0.98 USD/Capacity Unit-Hour

Crear Ver términos

2.h) Si el servicio de Watson Studio fue creado correctamente deberá ver una pantalla como la siguiente:

Instancias de servicio

Para actualizar un plan de servicio, en primer lugar [actualice](#) su cuenta de IBM Cloud. A continuación, seleccione [Actualizar servicio](#) o [Gestionar en IBM Cloud](#) desde el menú en la fila del servicio.

Filter by: Grupos de recursos Ninguna Producto Plan de servicio

Buscar instancias de servicio

Nombre	Grupo	Ubicación	Producto	Plan	Estado
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Activas
Watson Studio - lab	Default	Dallas	Watson Studio	Lite	Activas

Añadir servicio +

2.i) Vaya una vez más al menú de “Hamburguesa” (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”

The screenshot shows the IBM Cloud Pak for Data dashboard. On the left, there's a sidebar with a 'Filtrar navegación' dropdown and several menu items: Inicio, Datos, Proyectos, Catálogos, Gobernabilidad, Despliegues, Servicios, and Catalogo de servicios (which is highlighted with a red box). Below this are sections for Instancias de servicio, Gallery, Administración, and Soporte. The main area has a header 'Buenos días, Pedro!' and sections for 'Trabajar con datos' and 'Amplíe sus prestaciones'. There are buttons for 'Crear un proyecto' and 'Crear un servicio'. A central 'Visión general' section displays 'Proyectos recientes' (Data Fabric Outcomes, 07 abr 2022), 'Catálogos recientes' (CreditDataCatalog, 22 feb 2022), and a 'Notificaciones' card with a green checkmark and the text 'Project import complete'. The background features a 3D cube visualization.

2.j) Utilizando la opción de búsqueda (lupa) encuentre el servicio “Cognos Dashboard Embedded” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).

The screenshot shows the 'Catalogo de servicios' search results. A search bar at the top contains the text 'Cognos Dashboard E...'. The results list 'Analíticas' under 'Categoría'. The first result is 'IBM Cognos Dashboard Embedded' under 'Analíticas', which is highlighted with a green box. The description below it reads: 'Ponga los datos en acción directamente desde su aplicación con este servicio de visualización potente y fácil de usar! Lite • Gratuito'. Other categories listed include IA / Aprendizaje automático, bases de datos, Integración, Análiticas, and Almacenamiento.

2.k) Una vez dentro del servicio de Cognos Dashboard Embedded, seleccionar una región “Dallas” además la opción “Lite” (color rojo) y elija la opción de crear (color verde).

**Catálogo de servicios /**

**IBM Cognos Dashboard Embedded**

Autor: IBM • Fecha de última actualización: Dec 10, 2021 • Documentos

**Crear** Acerca de

Seleccionar una región

Seleccionar una región

Dallas

Plan de precios

Los precios mostrados no incluyen impuestos. Los precios mensuales que se muestran son para el país o región: United States

Plan	Características	Precios
Lite	50 sesiones al mes  Una sesión es un periodo de 60 minutos en el que los usuarios finales pueden realizar interacciones ilimitadas con un panel de control incluido.  Los servicios del plan Lite se suprimen tras 30 días de inactividad.	Gratis
Pago por uso	Tras 50 sesiones Conexión directa con los datos subyacentes Incluye paneles de control allí donde estén los usuarios sin que se pierda la interactividad Creación inteligente de visualizaciones Exploración interactiva de los datos por medio de filtros y rutas de navegación	\$0.05 USD/Session

**Resumen**

**IBM Cognos Dashboard Embedded**  
Región: Dallas  
Plan: Lite  
Nombre de servicio: IBM Cognos Dashboard Embedded-ju  
Grupo de recursos: Default

**Crear**

Ver términos

2.l) Si el servicio de Cognos Dashboard Embedded fue creado correctamente deberá ver una pantalla como la siguiente:

**Instancias de servicio**

Para actualizar un plan de servicio, en primer lugar [actualice](#) su cuenta de IBM Cloud. A continuación, seleccione [Actualizar servicio](#) o [Gestionar en IBM Cloud](#) desde el menú en la fila del servicio.

Filter by: Grupos de recursos ▾ Ubicaciones ▾ Ninguna ▾ Producto ▾ Plan de servicio ▾

Buscar instancias de servicio

Nombre	Grupo	Ubicación	Producto	Plan	Estado
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Activas
IBM Cognos Dashboard Embedded-db	Default	Dallas	IBM Cognos Dashboard Embedded	Lite	Activas
Watson Studio - lab	Default	Dallas	Watson Studio	Lite	Activas

**Añadir servicio**

2.m) Vaya una vez más al menú de “Hamburguesa” (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”

The screenshot shows the IBM Cloud Pak for Data dashboard. On the left, there's a sidebar with a hamburger menu icon. Below it, under 'Servicios', the 'Catalogo de servicios' option is highlighted with a red box. The main area has a dark background with a central graphic of data cubes and a monitor. The 'Visión general' section includes tabs for 'Proyectos recientes' (Data Fabric Outcomes, 07 abr 2022), 'Catálogos recientes' (CreditDataCatalog, 22 feb 2022), and 'Notificaciones' (Project import complete).

2.n) Utilizando la opción de búsqueda (lupa) encuentre el servicio “DataStage” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).

The screenshot shows the 'Catalogo de servicios' page with a search bar containing 'DataStage'. A red box highlights the search bar. Below the search bar, there are two sections: 'bases de datos' and 'Integración'. In the 'Integración' section, a green box highlights the 'DataStage' service card. The card shows the service icon, name, category (Analytics > bases de datos > Integración), a brief description, and 'Lite • Gratuito' status.

2.ñ) Una vez dentro del servicio de DataStage, seleccionar una región “Dallas” además la opción “Lite” (color rojo) y elija la opción de crear (color verde).

IBM Cloud Pak for Data

Catálogo de servicios /

## DataStage

Autor: IBM • Fecha de última actualización: Mar 1, 2022 • Documentos • Documentos de API

**Crear** Acerca de

Seleccionar una región:

Dallas

Plan de precios

Los precios mostrados no incluyen impuestos. Los precios mensuales que se muestran son para el país o región: United States

Plan	Características	Precios
Estándar	Plan estándar de IBM DataStage Entorno = Número de unidades de capacidad necesarias por hora • 1 vCPU + 4 GB RAM = 2 CUH • 2 vCPU + 8 GB RAM = 4 CUH • 4 vCPU + 16 GB RAM = 8 CUH • 8 vCPU + 32 GB RAM = 16 CUH • 16 vCPU + 64 GB RAM = 32 CUH	\$1.75 USD/Capacity Unit-Hour
Small Enterprise Bundle	Fully managed, serverless data integration with readily scalable DataStage PX engine with 5000 CUH included • Bulk purchase with same functionality as standard plan	\$7900.00 USD/Instance \$1.75 USD/Capacity Unit-Hour
Medium Enterprise Bundle	Fully managed, serverless data integration with readily scalable DataStage PX engine with 10000 CUH included • Bulk purchase with same functionality as standard plan	\$14000.00 USD/Instance \$1.75 USD/Capacity Unit-Hour
Large Enterprise Bundle	Fully managed, serverless data integration with readily scalable DataStage PX engine with 25000 CUH included • Bulk purchase with same functionality as standard plan	\$30600.00 USD/Instance \$1.75 USD/Capacity Unit-Hour
Lite	El Plan Lite de IBM DataStage permite a los usuarios acceder a las prestaciones de DataStage para un máximo de 15 CUH. Plan Lite de IBM DataStage Los servicios del plan Lite se suprinen tras 30 días de inactividad.	Granulado

**Crear**

Ver términos

2. o) Si el servicio de DataStage fue creado correctamente deberá ver una pantalla como la siguiente:

IBM Cloud Pak for Data

Todos | Buscar | Comprar | Pedro Angel Miranda's Acc... | PA

## Instancias de servicio

Para actualizar un plan de servicio, en primer lugar [actualice](#) su cuenta de IBM Cloud. A continuación, seleccione [Actualizar servicio](#) o [Gestionar en IBM Cloud](#) desde el menú en la fila del servicio.

Filter by: Grupos de recursos ▾ Ubicaciones ▾ Ninguna ▾ Producto ▾ Plan de servicio ▾

Buscar instancias de servicio

Nombre	Grupo	Ubicación	Producto	Plan	Estado	Añadir servicio
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Activas	+ Añadir servicio
DataStage-34	Default	Dallas	DataStage	Lite	Activas	+ Añadir servicio
IBM Cognos Dashboard Embedded-db	Default	Dallas	IBM Cognos Dashboard Embedded	Lite	Activas	+ Añadir servicio
Watson Studio - lab	Default	Dallas	Watson Studio	Lite	Activas	+ Añadir servicio

2.p) Vaya una vez más al menú de “Hamburguesa” (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”

The screenshot shows the IBM Cloud Pak for Data interface. On the left, there is a navigation sidebar with various options like 'Inicio', 'Datos', 'Proyectos', 'Catálogos', 'Gobernabilidad', 'Despliegues', 'Servicios', and 'Catálogo de servicios'. The 'Catálogo de servicios' option is highlighted with a red box. The main area displays a dashboard with sections for 'Trabajar con datos' and 'Amplíe sus prestaciones'. Below this is a 'Visión general' section with cards for 'Proyectos recientes' (Data Fabric Outcomes, 07 abr 2022), 'Catálogos recientes' (CreditDataCatalog, 22 feb 2022), and 'Notificaciones' (Project import complete). A large graphic on the right depicts a computer monitor displaying a chart, with several 3D cubes of different colors (blue, green, purple) floating around it, representing data and machine learning components.

2.q) Utilizando la opción de búsqueda (lupa) encuentre el servicio “Machine Learning” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).

The screenshot shows the 'Catálogo de servicios' search results. In the search bar at the top left, the text 'Machine Learning' is entered. Below the search bar, a list of categories is shown: 'Categoría' followed by 'IA / Aprendizaje automático', 'bases de datos', 'Integración', 'Analíticas', and 'Almacenamiento'. To the right of this list, a search result for 'Machine Learning' is displayed in a box with a green border. The result includes the service name 'Machine Learning', its category 'IA / Aprendizaje automático', a brief description 'Despliegue, gestione e integre modelos de aprendizaje automático en sus aplicaciones y servicios en tan solo u...', and a status 'Lite • Gratuito'. The entire search results area is set against a light gray background.

2.r) Una vez dentro del servicio de Machine Learning, seleccionar una región “Dallas” además la opción “Lite” (color rojo) y elija la opción de crear (color verde).

The screenshot shows the 'Machine Learning' service page in the IBM Cloud Pak for Data interface. The 'Region' dropdown is set to 'Dallas'. A red box highlights the 'Lite' plan option in the 'Características' section. A green box highlights the 'Crear' button at the bottom right.

Plan	Características	Precios
Lite	Instancia de servicio 20 horas de unidades de capacidad (CUH) incluidas: Tipos de ejecutores: - 1 vCPU y 4 GB RAM = 0,5 unidades de capacidad necesarias por hora - 2 vCPU y 8 GB RAM = 1 unidad de capacidad necesaria por hora - 4 vCPU y 16 GB RAM = 2 unidades de capacidad necesarias por hora - 8 vCPU y 32 GB RAM = 4 unidades de capacidad necesarias por hora - 16 vCPU y 64 GB RAM = 8 unidades de capacidad necesarias por hora - 32 vCPU y 128 GB RAM = 16 unidades de capacidad necesarias por hora IA automática con datos Unidos • 1 Controlador: 2 vCPU y 8 GB de RAM + 1 Ejecutor: 2 vCPU y 8 GB de RAM = 10 unidades de capacidad necesarias por hora, más 5 unidades de capacidad por hora para cada ejecutor adicional. Una vez que se han unido los datos, CUH se consume a una velocidad de 20 unidades de capacidad por hora Optimización de decisiones • 2 vCPU y 8 GB RAM = 30 unidades de capacidad necesaria por hora • 4 vCPU y 16 GB RAM = 40 unidades de capacidad necesaria por hora • 8 vCPU y 32 GB RAM = 50 unidades de capacidad necesaria por hora • 16 vCPU y 64 GB RAM = 60 unidades de capacidad necesaria por hora Máximo 2 trabajos por lotes paralelos de Optimización de decisiones por despliegue Número predefinido de trabajos de despliegue retenidos por espacio de despliegue: 100	Gratis

La instancia de plan básico del servicio IBM Watson Machine Learning proporciona 20 horas de unidades de capacidad al mes durante las cuales los modelos se pueden entrenar, evaluar, desplegar y puntuar.

**Crear**

2. s) Si el servicio de Machine Learning fue creado correctamente deberá ver una pantalla como la siguiente:

The screenshot shows the 'Instancias de servicio' (Service Instances) page. It lists several active services: CloudObjectStorage, DataStage-34, IBM Cognos Dashboard Embedded-db, WatsonMachineLearning, and Watson Studio - lab, all in the 'Activas' state.

Nombre	Grupo	Ubicación	Producto	Plan	Estado
CloudObjectStorage	Default	Global	Cloud Object Storage	Lite	Activas
DataStage-34	Default	Dallas	DataStage	Lite	Activas
IBM Cognos Dashboard Embedded-db	Default	Dallas	IBM Cognos Dashboard Embedded	Lite	Activas
WatsonMachineLearning	Default	Dallas	Machine Learning	Lite	Activas
Watson Studio - lab	Default	Dallas	Watson Studio	Lite	Activas

“Felicitaciones ha habilitado todos los servicios que estaremos utilizando” para crear esta demostración”.

### 3. Crear proyecto

Ahora vamos a crear nuestro proyecto de Riesgo Crediticio, vaya al menú de “Hamburguesa” y seleccione la opción “Proyectos” y “Ver Todos los Proyectos” (color rojo).

The screenshot shows the IBM Cloud Pak for Data interface. The top navigation bar includes the IBM logo, a search bar, and a 'Todos' dropdown. The left sidebar has a 'Filtrar navegación' button. The main content area displays a 'nido, Pedro!' message. The 'Proyectos' section is expanded, showing options like 'Ver todos los proyectos' (highlighted with a red box), 'Trabajos', 'Catálogos', 'Gobernabilidad', 'Despliegues', 'Servicios', 'Gallery', 'Administración', and 'Soporte'. To the right, there are sections for 'ediante', 'Trabajar con datos', and 'Amplíe sus prestaciones'. A 'Visión general' and 'Proyectos recientes' section are also visible.

3.a) Si usted seleccionó la opción correcta verá una pantalla como ésta y ahí debe seleccionar la opción “Nuevo Proyecto” (color rojo) para crear su primer proyecto.

The screenshot shows the 'Proyectos' page. The top navigation bar includes the IBM logo, a search bar, and a 'Todos' dropdown. The main content area shows a table with columns for 'Nombre', 'Última modificación', 'Su rol', and 'Colaboradores'. A 'Nuevo proyecto' button is located in the top right corner of the table header, which is highlighted with a red box.

3.b) Ahora de clic en la opción “Crear un proyecto vacío” (color rojo).

3.c) A continuación defina el “Nombre” y “Descripción” del proyecto como “Riesgo Crediticio” (color rojo), el “Almacenamiento” le debe aparecer auto-llenado con el nombre del servicio de Cloud Object Storage que creó anteriormente, si todo aparece como se ve en la siguiente pantalla entonces oprima el botón de “Crear” (color amarillo, esquina inferior derecha):

3.d) Ahora debe ver la pantalla principal de su proyecto “Riesgo Crediticio” y debe estar ubicado en la pestaña de “Visión general”, por favor seleccione la pestaña “Activos” (color rojo).

3.e) Ahora descargue desde Github hacia su PC los dos archivos que han sido proporcionados para la demo como parte del curso (“18-1 - Riesgo Crediticio - Histórico” y “18-1 - Riesgo Crediticio - Solicitantes” (tome nota del directorio donde queden guardados).

A continuación, regrese a CP4DaaS seleccione la opción “Suelte aquí archivos de datos o examine en busca de datos que cargar” (a la derecha de su pantalla, color rojo) y seleccione o arrastre los archivos al proyecto.

Si todo lo hizo bien, una vez que haya cargado los dos archivos deben aparecer ambos en la sección “Activos de Datos” (color verde)

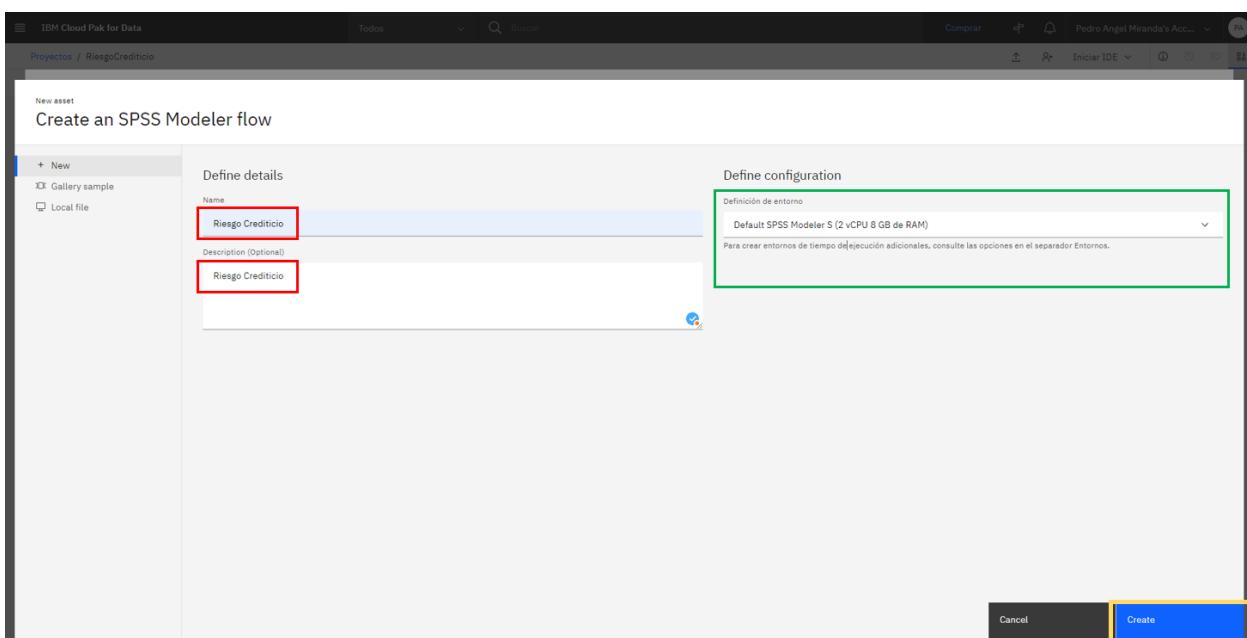
“Felicitaciones ya creo su proyecto y descargó las dos fuentes de datos que va a utilizar para su modelo”.

#### 4. Crear modelo analítico

Ahora vamos a crear un modelo analítico utilizando una herramienta visual (SPSS Modeler) que se encuentra dentro de Watson Studio, oprima el botón azul “Nuevo Activo” (a su derecha en la parte superior de la pantalla, marcado con color rojo)

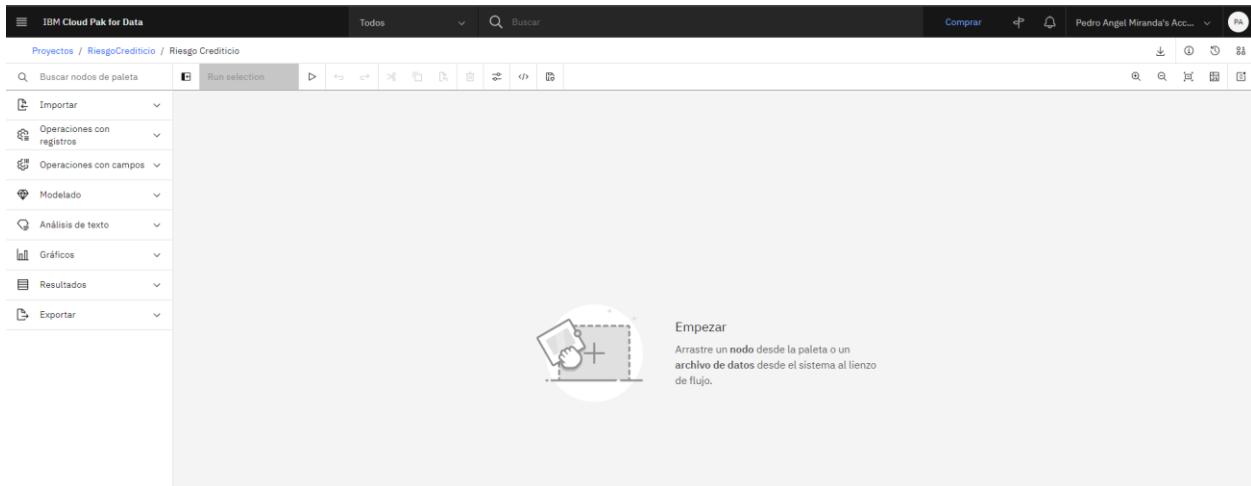
y a continuación la opción “SPSS Modeler” (color verde).

4.a) Nombre su modelo como “Riesgo Crediticio” (color rojo) y en la descripción “Riesgo Crediticio” (color rojo), deje el valor que por default le aparece en “Definición de entorno” (2vCPU 8GB RAM) (color verde) y oprima la opción de “Crear” (color amarillo, en la esquina inferior derecha), una vez que oprima Crear espere a que se generé en el entorno de ejecución del SPSS Modeler, puede tardar unos pocos minutos, sea paciente.

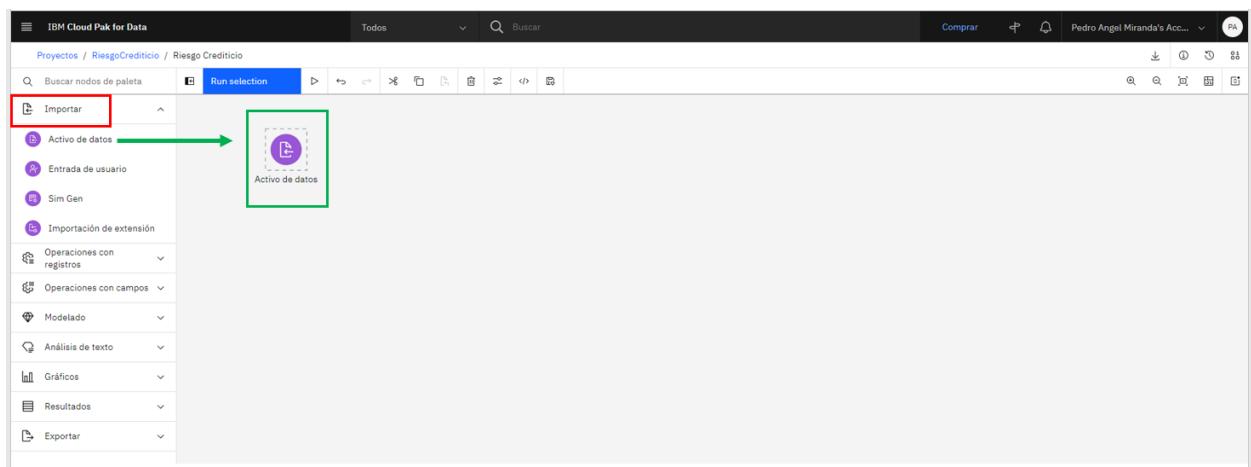


Nota: Para fines prácticos del Hackathon siempre utilizar Default SPSS Modeler (2vCPU 8GB RAM)

4.b) Si se ejecutó bien el paso anterior le aparecerá la siguiente pantalla con un lienzo en blanco.

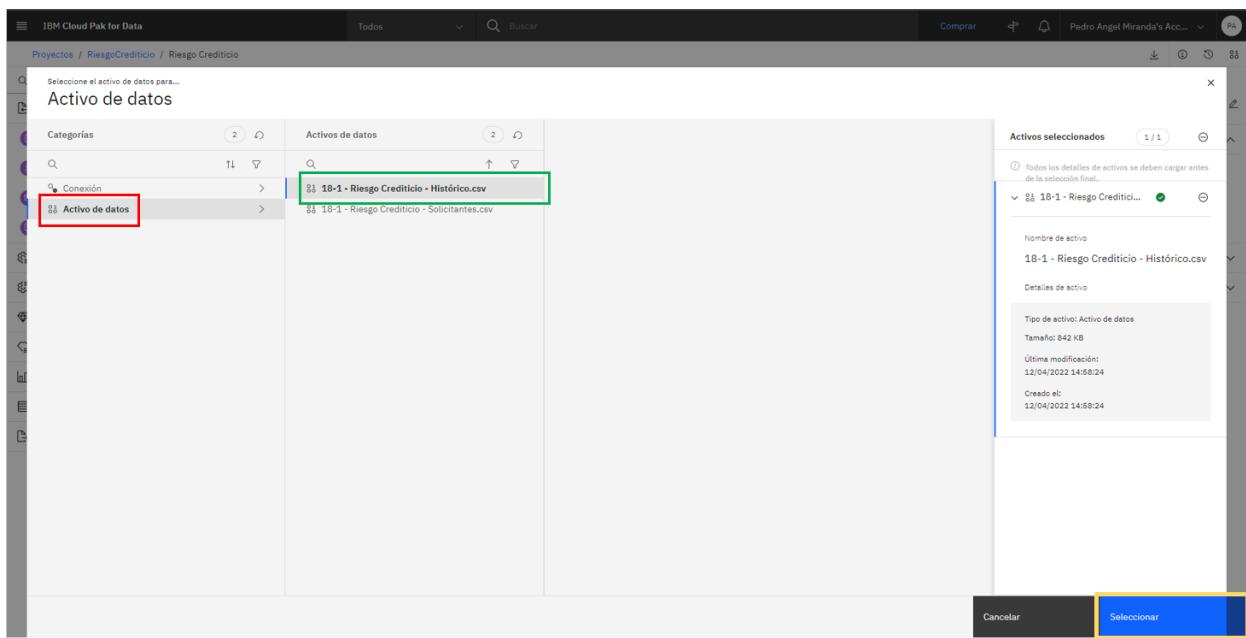


4.c) A continuación vamos a **crear el flujo de entrenamiento del modelo**, para eso debe ir a la opción de “Importar” (color rojo) y arrastrar hacia el lienzo el nodo “Activo de Datos” (flecha verde) y quedará como sigue:



4.d) Haga clic derecho sobre el nodo “Activo de Datos” y dé clic en “Abrir”, a continuación, le aparecerá la siguiente pantalla y debe dar clic en “Activos de Datos” (color rojo), aparecerán a la derecha las dos fuentes de datos que tenemos en nuestro proyecto, seleccione la fuente de datos “18 - 1 - Riesgo Crediticio - Histórico” (color verde) y finalmente de clic en “Seleccionar” (color amarillo, esquina inferior derecha).

De esta manera habrá asociado una fuente de datos al nodo “Activo de Datos”. A continuación, le aparecerá una segunda pantalla del lado derecho donde debe seleccionar la opción “Guardar” y habrá terminado este paso.

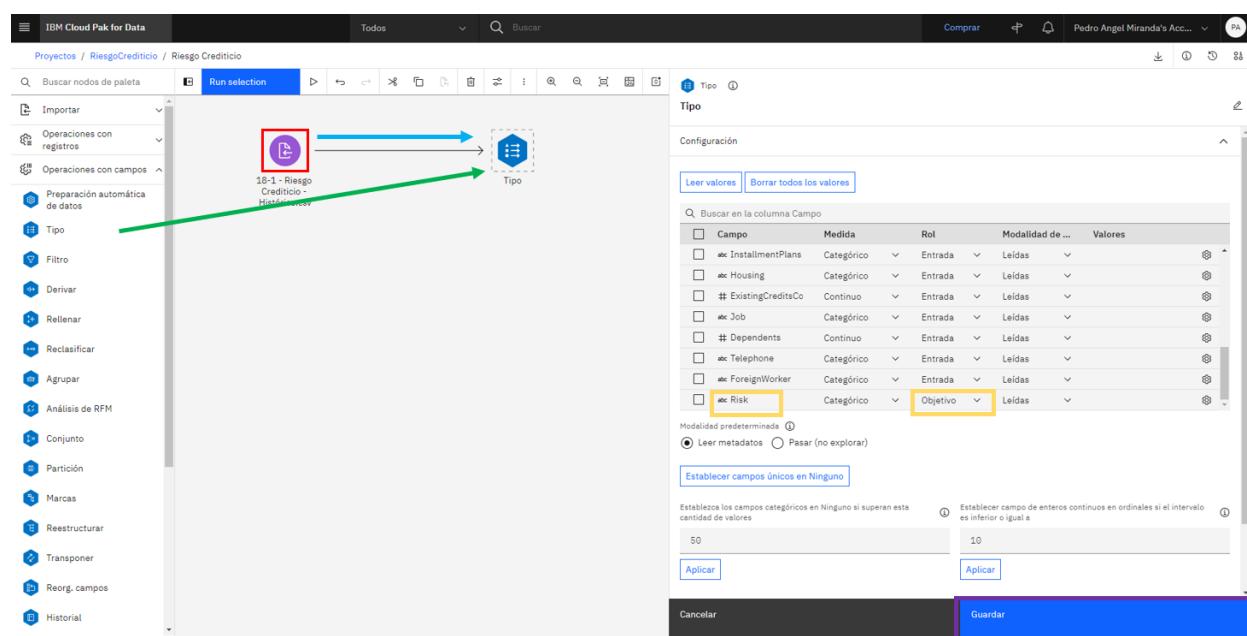


4.e) Si hizo todo bien regresará al lienzo donde tendrá el ícono de “Activo de datos”, pero habrá sido renombrado con el nombre de la fuente de datos que le acaba de asociar (color rojo).

Ahora debe ir a las opciones de su lienzo que están a la izquierda y seleccionar la opción **“Operaciones con campos”** y arrastrar el ícono “Tipo” al lienzo (flecha verde) y conectar el nodo “Activo de Datos” con el nodo “Tipo” (color azul).

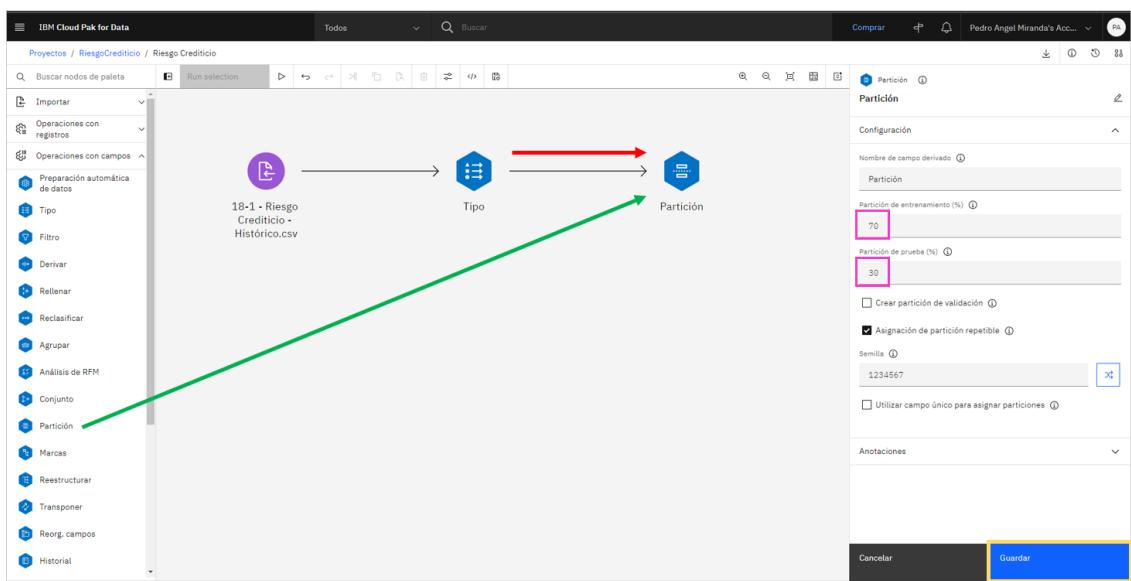
Ahora abra el ícono de “Tipo” (botón derecho y opción “Abrir”), aparecerá una nueva pantalla a un lado donde estarán todos los campos de la fuente de datos, busque el que se llama “Risk” (color amarillo, 1er columna) cambie el “rol” de ese campo a “Objetivo” (3<sup>a</sup> columna, color amarillo).

Así habrá definido el campo “Risk” como la variable objetivo en su modelo. A continuación de clic en guardar (color morado, esquina inferior derecha).



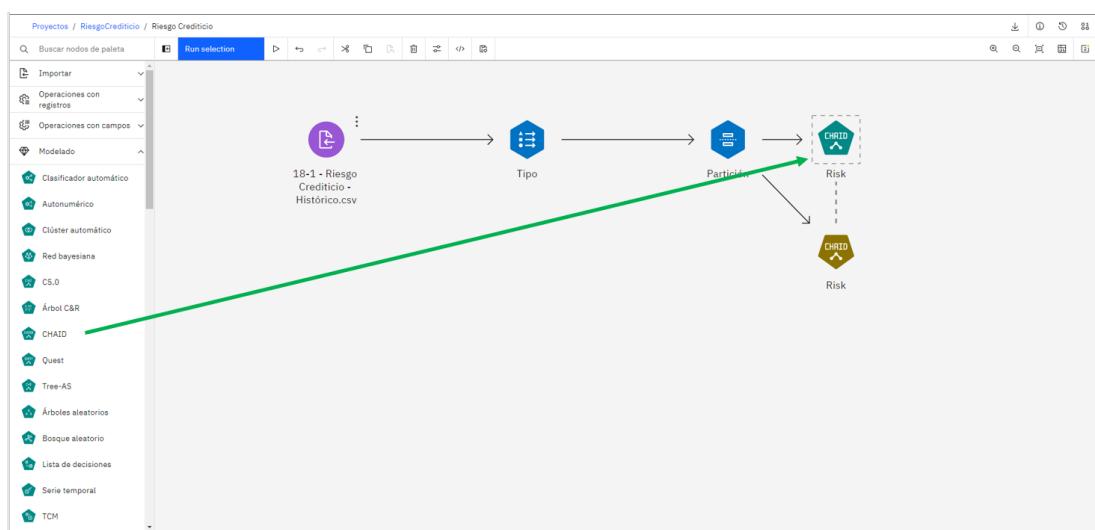
4.f) A continuación en “Operaciones con Campos” seleccione el nodo “Partición” y arrástrelo al lienzo (flecha verde), conecte el nodo “Tipo” al nodo “Partición” (flecha roja).

Ahora “Abra” el nodo “Partición” (botón derecho sobre el ícono y seleccionar “Abrir”), se abrirá una pantalla adicional a la derecha, ponga 70 en el campo “Partición de entrenamiento (%)” (color rosa) y 30 en “Partición de prueba (%)” (color rosa), finalmente dé clic en el botón “Guardar” (color amarillo).



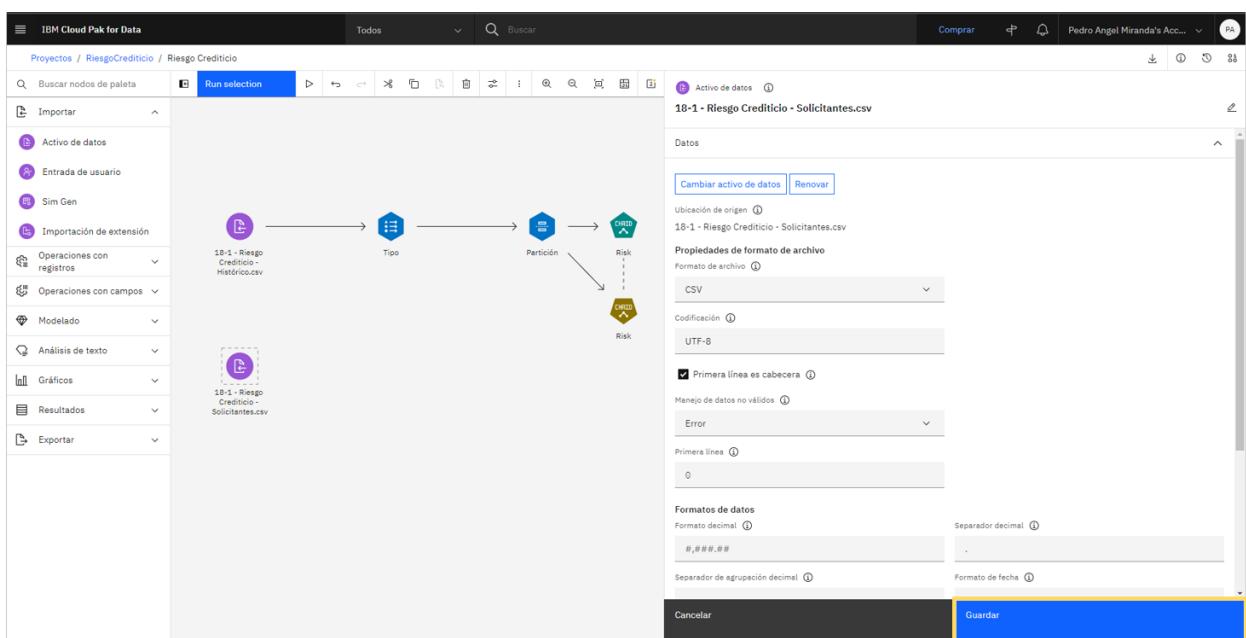
4.g) Ahora vaya a las opciones de la derecha y busque la sección de “Modelado” y arrastrar el ícono “CHAID” al lienzo (flecha Verde), conectar el nodo “Partición” con el nodo “CHAID”

Ahora de click derecho el nodo “CHAID” y seleccionar la opción “Ejecutar” que como resultado creará un nodo “CHAID” dorado (**modelo entrenado**), si se abre una ventana de mensajes en la parte superior derecha del lienzo, sólo ciérrela dando clic en la “x”.



Como siguiente paso debe arrastrar al lienzo un nuevo nodo de “Activo de Datos” (lo encuentra en la sección “Importar”), debe abrir ese nodo y asociarle la segunda fuente de datos que se le proporcionó (Opción “Activos de datos” y fuente de datos “18-1 - Riesgo Crediticio - Solicitantes”, dar clic al botón “Seleccionar” y en la siguiente pantalla dar clic al botón “Guardar” [color amarillo]).

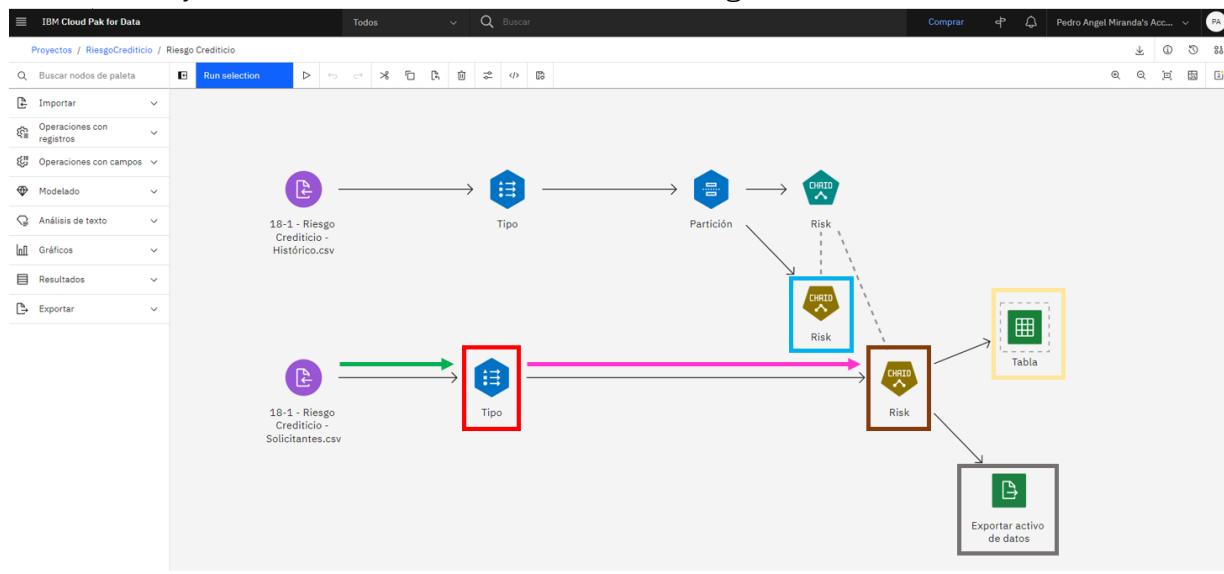
Con éste nuevo ícono de “Activo de Datos” estamos iniciando la creación del flujo de predicción.



4.h) Ahora vamos a **crear el flujo completo de predicción** como sigue:

- Arrastre un nodo “Tipo” (“Operaciones con campos”/”Tipo”, color rojo)
- conecte el nodo “Activo de Datos” con el nuevo nodo “Tipo” (color verde)
- dé clic sobre el nodo “CHAID dorado” (color azul) para tenerlo seleccionado y a continuación usa las teclas Ctrl-C y Ctrl-V para crear una copia de ese nodo
- conecte el nodo “Tipo” (color rojo) con la copia del nodo (Flecha rosa) “CHAID” dorado (color café)
- a continuación, arrastre un nodo de “Tabla” (“Resultados”/”Tabla”) al lienzo (color amarillo) y un nodo de “Exportar Activo de Datos” (“Exportar”/”Exportar Activo de Datos”) al lienzo (color gris)
- abra el nodo “Exportar Activo de Datos” (color gris) y cambie el “Nombre de archivo” de “output.csv” a “Prediccion”, de click en guardar
- a continuación, conecte la copia del nodo CHAID dorado (color café) con el nodo de “Tabla” (color amarillo) y al nodo de “Exportar Activo de Datos” (color gris).
- De click derecho en el nodo “Tabla” (color amarillo) y seleccione la opción “Ejecutar”
- De click derecho en el nodo “Exportar Activo de Datos” (color gris) y de la opción “Ejecutar”.

Al final el flujo de su modelo se deberá ver como sigue:



4.i) Obtenga una vista previa del nodo (color azul) “Tabla” (color amarillo, punto anterior) y del nodo “Exportar Activo de Datos” (color gris, punto anterior) y confirme que en ambos aparezcan los campos: “\$R-Risk” y “\$RC-Risk” (color rojo) que se generan como resultado de correr el modelo de predicción.

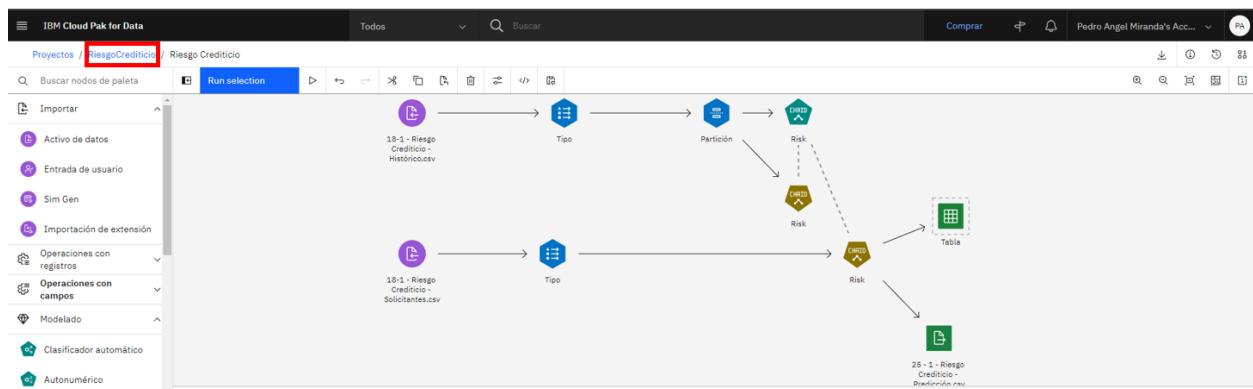
The screenshot shows the IBM Cloud Pak for Data interface. On the left, there's a sidebar with various options like Importar, Operaciones con registros, Modelado, and Gráficos. The main area displays a flow diagram with nodes: 'Run selection' (blue), 'Importar' (grey), 'Operaciones con registros' (grey), 'Operaciones con campos' (grey), 'Modelado' (grey), 'Análisis de texto' (grey), 'Gráficos' (grey), 'Resultados' (grey), and 'Exportar' (grey). A specific node labeled '18-1 - Riesgo Credítico - Solicitantes.csv' is highlighted. Arrows connect these nodes to a central 'Risk' node, which then connects to a 'CHAID' node and a 'Tabla' (Table) node. A context menu is open over the 'Tabla' node, with 'Vista previa' (Preview) selected. The preview table shows data with columns: 'holdersOnLoan', 'CurrentResidenceDuration', 'OwnsProperty', 'Age', 'InstallmentPlans', 'Housing', 'ExistingCreditsCount', 'Job', 'Dependents', 'Tel', 'Nuevo comentario', 'Preview SQL pushback', 'Risk', '\$Risk', and '\$RC-Risk'. The '\$Risk' and '\$RC-Risk' columns are highlighted with red boxes. The 'Risk' column also has a red box around it.

holdersOnLoan	CurrentResidenceDuration	OwnsProperty	Age	InstallmentPlans	Housing	ExistingCreditsCount	Job	Dependents	Tel	Nuevo comentario	Preview SQL pushback	Risk	\$Risk	\$RC-Risk
ne	3	savings_insurance	32	none	own	1	skilled	1	no			0.869		
ne	2	savings_insurance	37	stores	own	2	skilled	1	no	Ejecutar		0.896		
ne	3	real_estate	28	none	own	2	skilled	1	yes	no	No Risk	0.878		
ne	2	savings_insurance	32	none	own	1	skilled	1	none	yes	No Risk	0.640		
ne	3	unknown	57	none	own	2	skilled	1	none	yes	Risk	0.716		
-applicant	5	unknown	57	none	free	2	skilled	2	yes	yes	Risk	0.950		

¡Felicitaciones su modelo está listo!

## 5. Creación de un Tablero a partir de los datos que tenemos.

Regrese a su “Proyecto” (tip en la esquina superior izquierda dice: “Proyectos / Riesgo Crediticio / Riesgo Crediticio”, de clic en el link “Riesgo Crediticio” en color azul, el que viene después de la diagonal de “Proyectos”(color rojo).



Vaya a la parte de “Gestionar” (color verde), despues el la sección “Servicios e integraciones” y finalmente en asociar un servicio (color amarillo)

Palomea “IBM Cognos Dashboard Embedded” y seleccionar Asociar servicio (color rojo)

1 item selected						
Nombre	Tipo	Plan	Ubicación	Estado	Grupo	
<input type="checkbox"/> Watson OpenScale-op	Watson OpenScale	Lite	Dallas	No asociado	Default	
<input type="checkbox"/> Data Replication-kd	Data Replication	Lite	Global	No asociado	Default	
<input checked="" type="checkbox"/> IBM Cognos Dashboard Embedded-db	IBM Cognos Dashboard Embedded	Lite	Dallas	No asociado	Default	
<input type="checkbox"/> WatsonMachineLearning	Machine Learning	Lite	Dallas	No asociado	Default	

Si todo sale bien vera una paloma verde en la parte de “Estado” y si servicio estara asociado al proyecto

Asociar servicio

El servicio IBM Cognos Dashboard Embedded-db se ha asociado correctamente al proyecto RiesgoCrediticio.

Nombre	Tipo	Plan	Ubicación	Estado	Grupo
<input type="checkbox"/> Watson OpenScale-op	Watson OpenScale	Lite	Dallas	No asociado	Default
<input type="checkbox"/> Data Replication-kd	Data Replication	Lite	Global	No asociado	Default
<input checked="" type="checkbox"/> IBM Cognos Dashboard Embedded-db	IBM Cognos Dashboard Embedded	Lite	Dallas	Asociado	Default
<input type="checkbox"/> WatsonMachineLearning	Machine Learning	Lite	Dallas	No asociado	Default

5.a) Regrese a la parte de “Activos” dé clic al botón azul “Nuevo activo” (color rojo)

IBM Cloud Pak for Data

Proyectos / RiesgoCrediticio

Visión general Activos Trabajos Gestionar

2 assets

Todos los activos

Nombre Última modificación

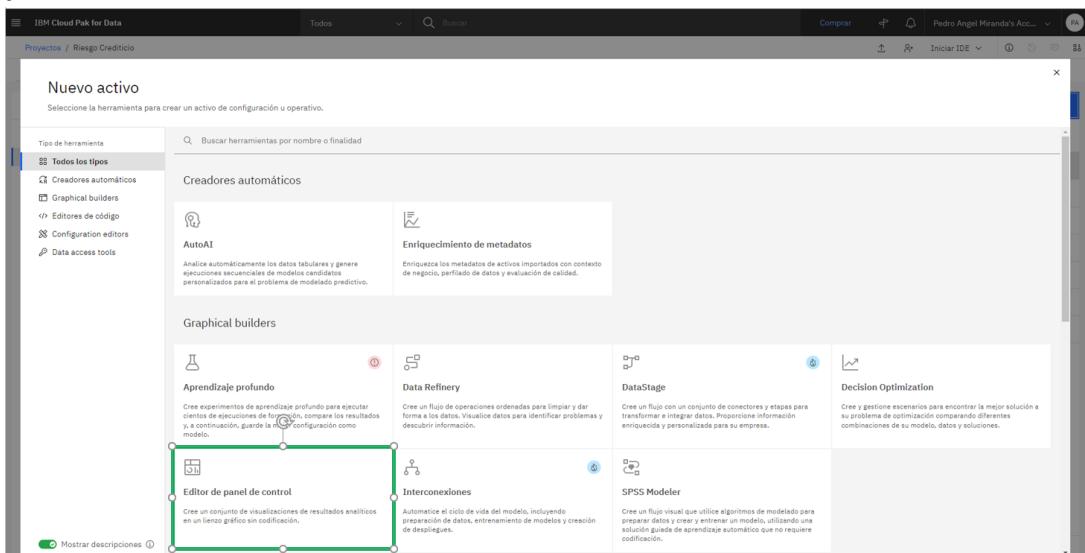
18-1 - Riesgo Crediticio - Solicitantes.csv Ahora Pedro Angel Miranda (Usted)  
18-1 - Riesgo Crediticio - Histórico.csv Ahora Pedro Angel Miranda (Usted)

Añadir activo +

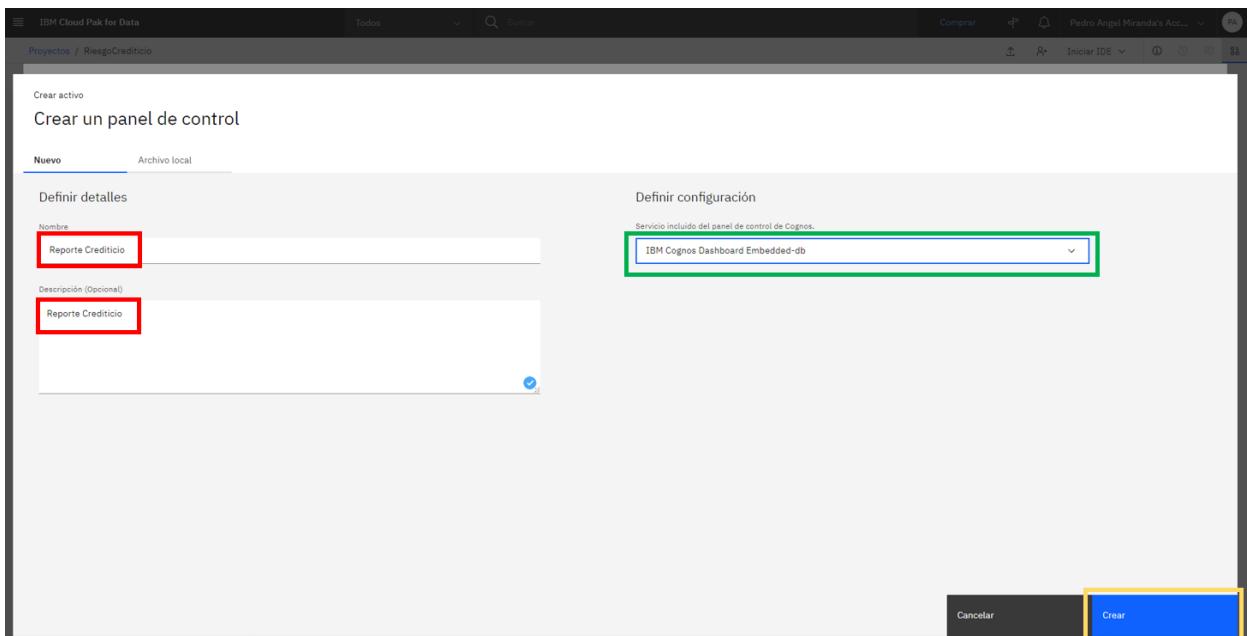
Nuevo activo +

Suelte aquí archivos de datos o examine en busca de datos que cargar

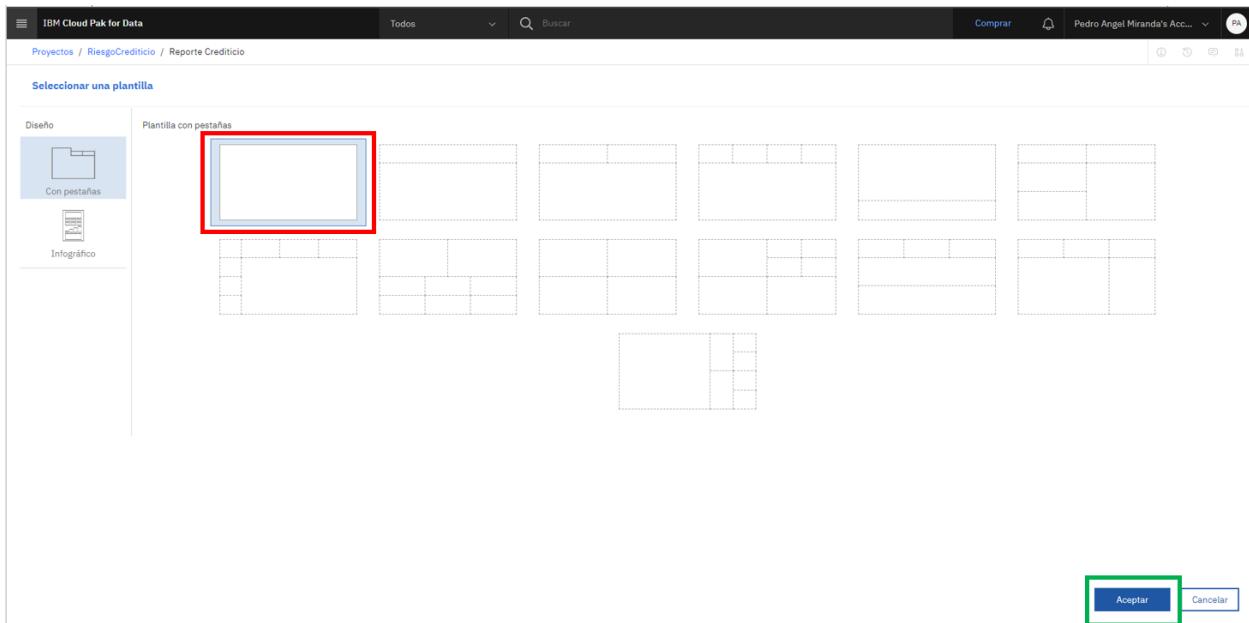
y seleccione “Editor Panel de Control” (color verde)



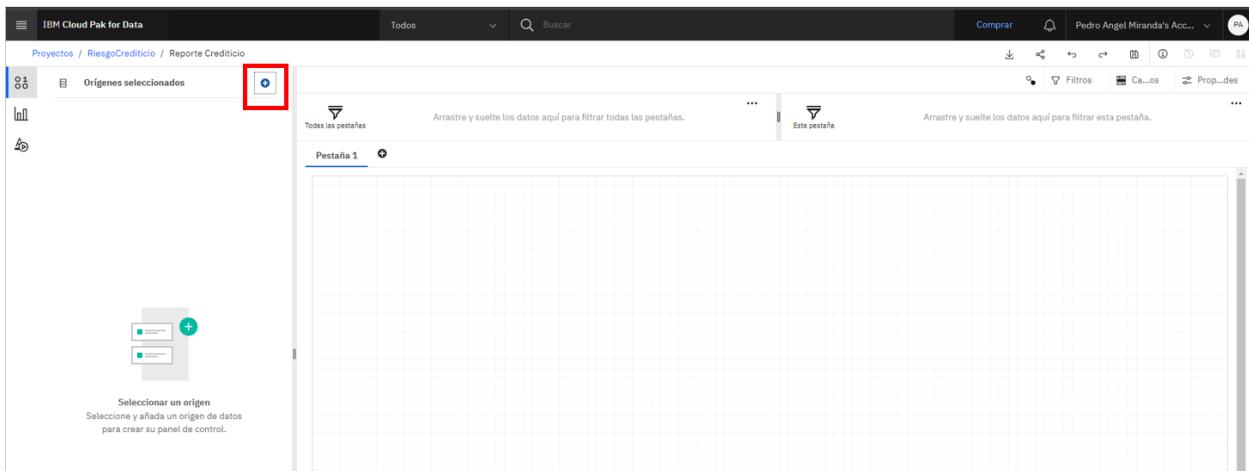
5.b) Ponga en Nombre y Descripción “Reporte Crediticio” (color rojo), confirme que ya tiene una instancia de Cognos asociada al proyecto (color verde) y de clic en “Crear” (color amarillo).



5.c) Seleccione la “Plantilla con pestañas” en blanco (color rojo) y de clic en “Aceptar” (color verde).



5.d) Dé clic en el símbolo de “+” (color rojo) en la sección de “Orígenes de Datos”.



5.e) Seleccione “Activos de datos” (color rojo), luego seleccione el archivo “25 - 1 - Riesgo Crediticio - Predicción.csv” (color verde) y de clic en “Seleccionar” (color amarillo, esquina inferior derecha)

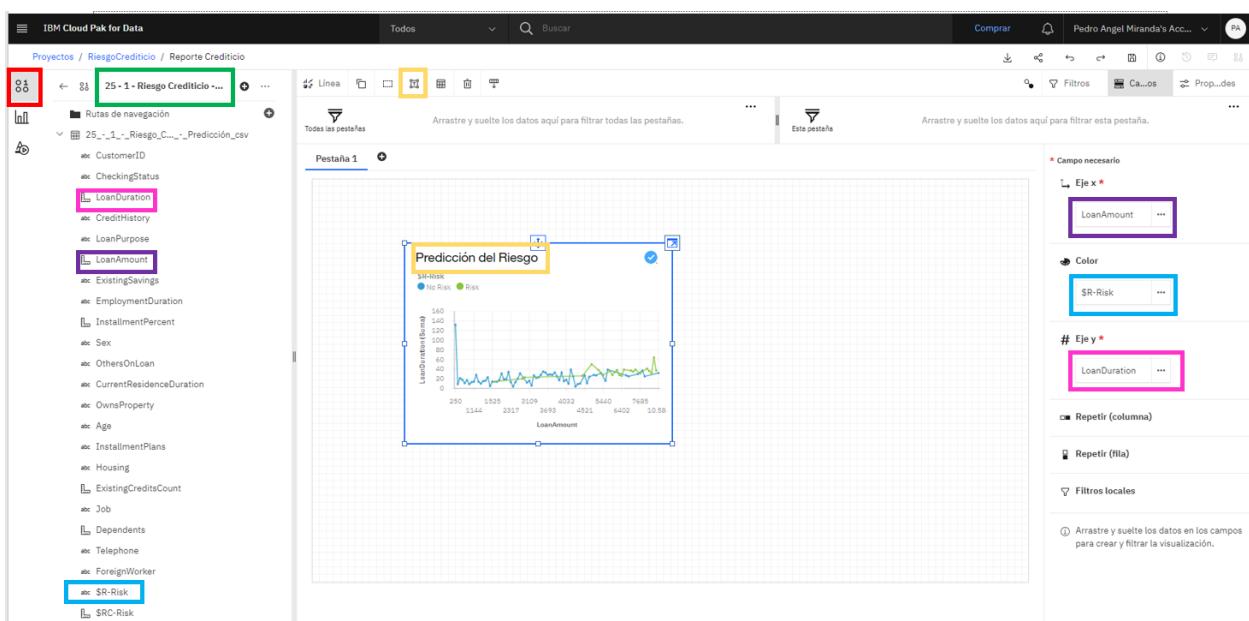
Seleccionar origen de datos

5.f) Seleccione los tipos de gráficas (color rojo), arrastre la “gráfica de línea” al lienzo (flecha verde).

5.g) Seleccione las fuentes de datos (color rojo) y el archivo “25 - 1 - Riesgo Crediticio - Predicción” (color verde) y:

- arrastre “LoanAmount” de los campos de la derecha y póngalo en el Eje x (color morado)
- arrastre LoanDuration y póngalo en el Eje y (color rosa)
- arrastre “\$R-Risk” y póngalo en “Color” (color azul)
- selección la opción “T” (color amarillo) en las opciones superiores del lienzo y ponga el título “Predicción del Riesgo” a su gráfica.

Si todo lo hizo bien tendrá la gráfica que aparece en la pantalla.



5.h) Seleccione una gráfica de “Barras apiladas” (tip: seleccione el ícono visualizaciones [color rojo] y es la segunda arriba a su izquierda) arrástrela a su lienzo a la derecha de la gráfica que ya creó.

A continuación, vaya al mismo archivo de datos (color verde) con el que ya está trabajando y:

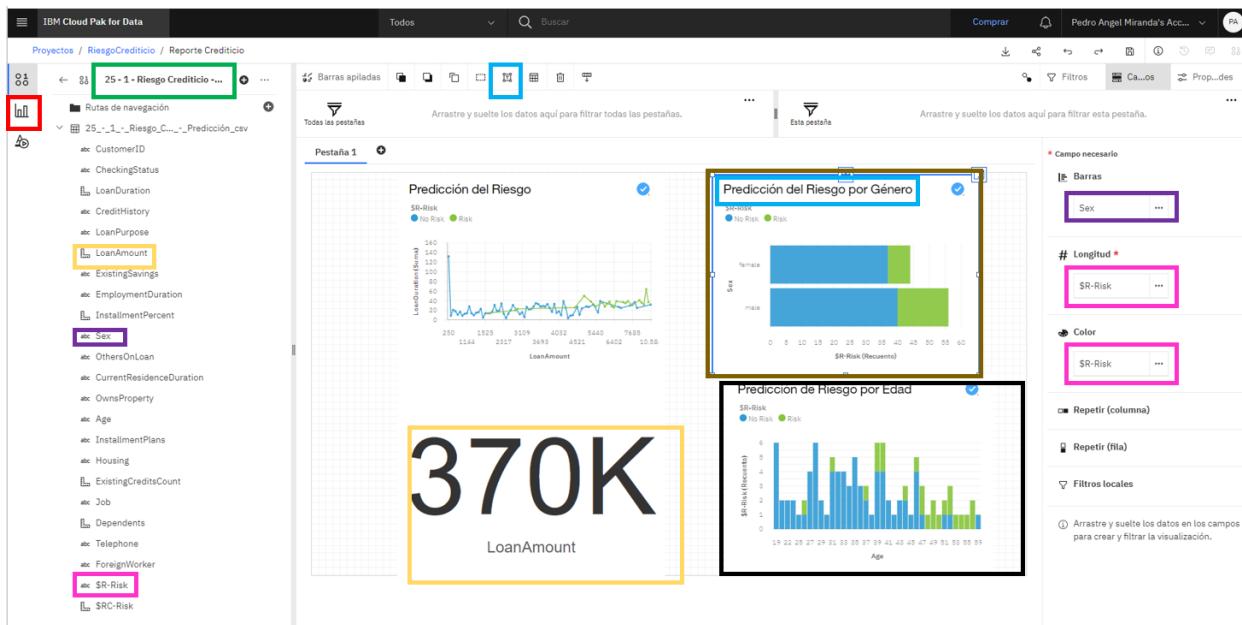
- arrastre el campo “Sex” que está en el lado izquierdo hacia “Barras” en las opciones de la gráfica del lado derecho (color morado)
- arrastre el campo “\$R-Risk” a “Longitud” y “Color” (color rosa)
- finalmente póngale el título “Predicción del Riesgo por Género” (color azul)

Si lo hizo todo bien verá la gráfica de barras que aparece abajo (color café).

Ahora:

- arrastre el indicador de “LoanAmount” (color amarillo) bajo la primera gráfica que hizo
- finalmente cree una nueva gráfica de columnas apiladas poniendo el campo “Age” en “Barras”
- y el campo “\$R-Risk” en “Longitud” y “Color” con el título “Predicción de Riesgo por Edad”

Si la hizo bien será igual a la que parece abajo (color negro).



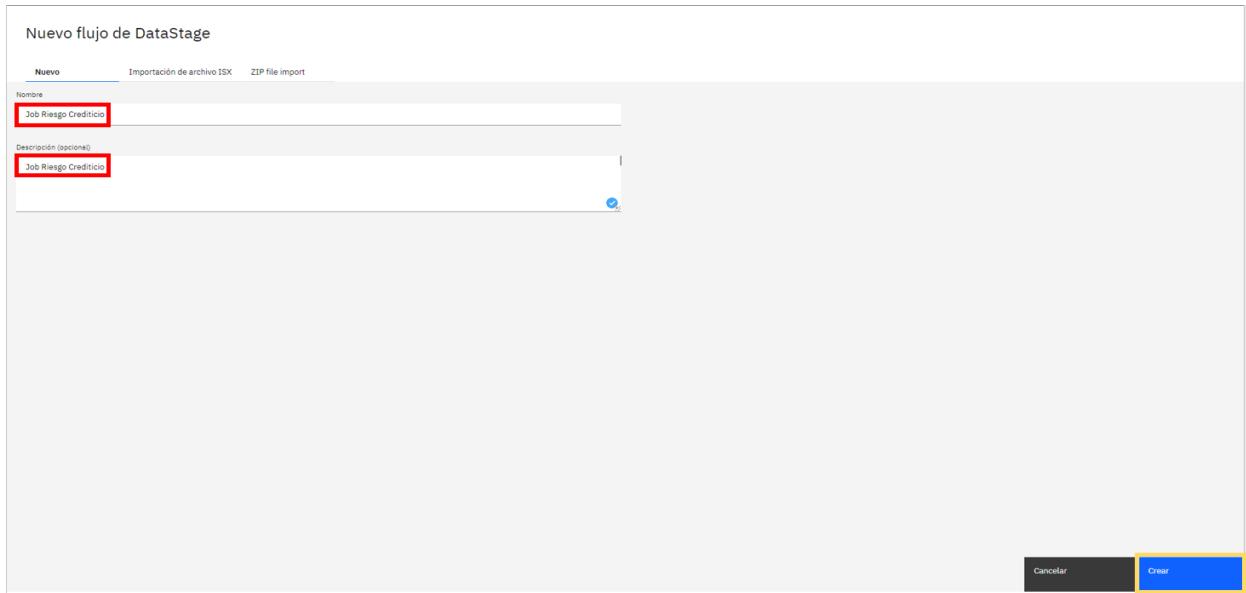
## Parte 2

6.a) Regrese a la parte de “Activos” dé clic al botón azul “Nuevo activo” (color rojo).

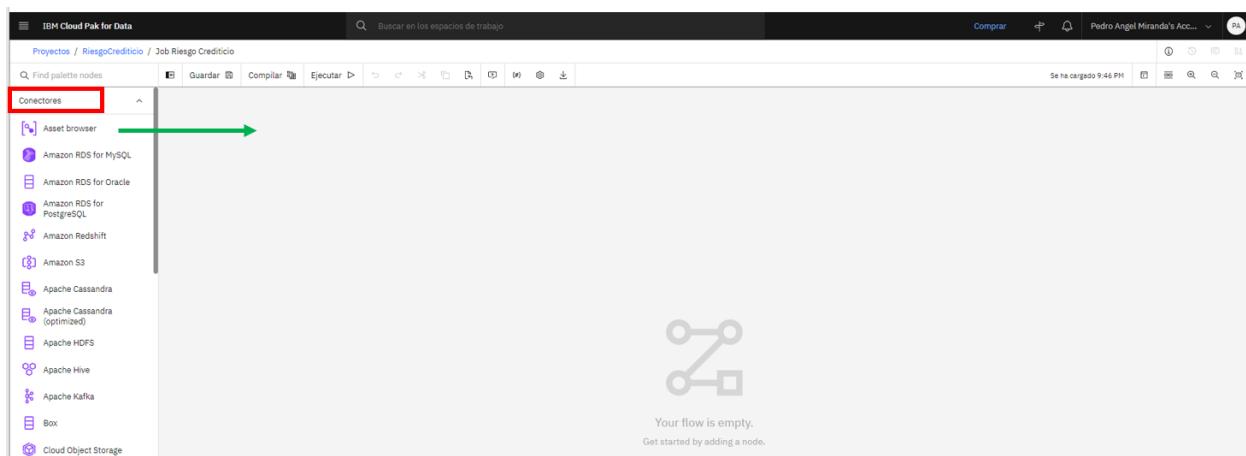
y seleccione “DataStage” (color verde).

Nombre	Última modificación
1B-1 - Riesgo Crediticio - Solicitantes.csv	Ahora Pedro Angel Miranda (Listed)
1B-1 - Riesgo Crediticio - Histórico.csv	Ahora Pedro Angel Miranda (Listed)

Nombre su flujo de DataStage y la descripción como “Job Riesgo Crediticio” (color rojo) y oprima la opción de “Crear” (color amarillo).



6.b) A continuación primero vamos a **crear los componentes (Stages) de nuestro flujo de DataStage**, para eso debe ir a la opción de “Conectores” (color rojo) y arrastrar hacia el lienzo el nodo “Activo de Datos” (flecha verde)

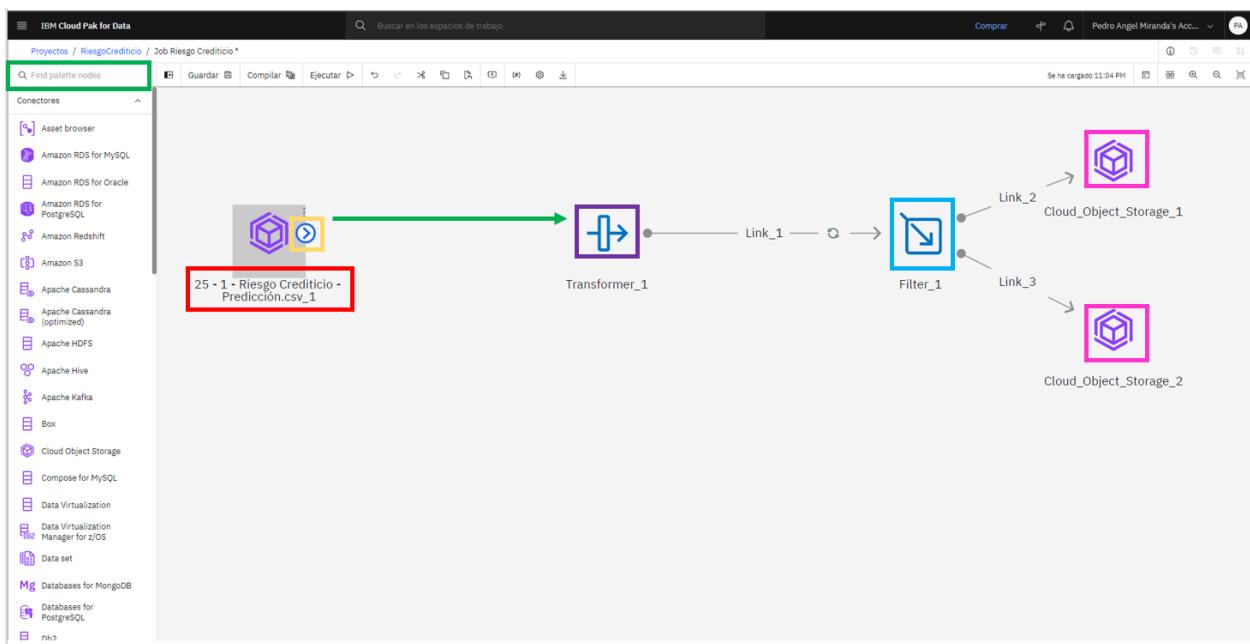


A continuación, le aparecerá la siguiente pantalla, y debe dar clic en “Activos de Datos” (color rojo), aparecerán a la derecha las dos fuentes de datos que tenemos en nuestro proyecto, seleccione la fuente de datos “25 - 1 - Riesgo Crediticio - Predicción.csv” (color rojo) y finalmente de clic en “Añadir” (color amarillo, esquina inferior derecha).

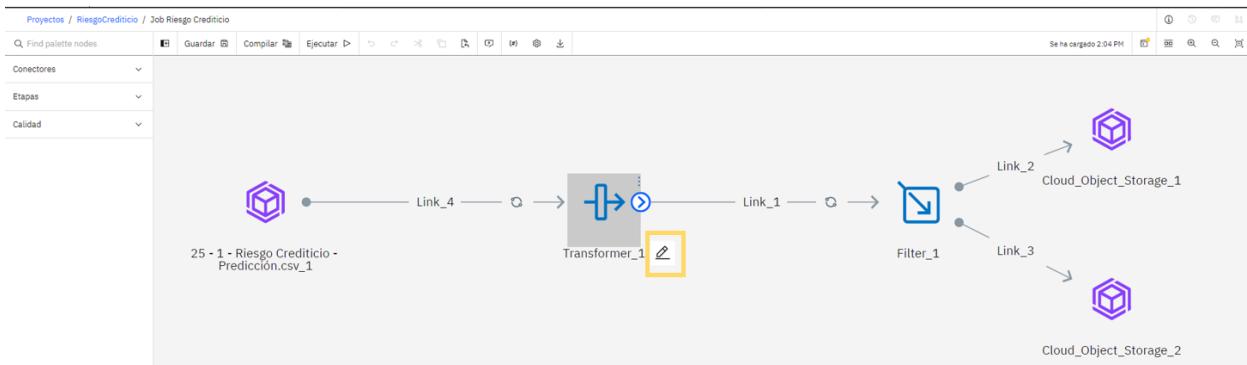
Si los pasos anteriores se realizaron correctamente obtendrá la siguiente pantalla con el nombre del activo de datos.

6.c) Ahora tendremos que poner los componentes(stages) de nuestro flujo, vaya a “Find palette nodes” (color verde) y escriba “Transformer”, arrástrelo y póngalo en el lienzo, haga lo mismo para “Filter”, y “Cloud object Storage”.

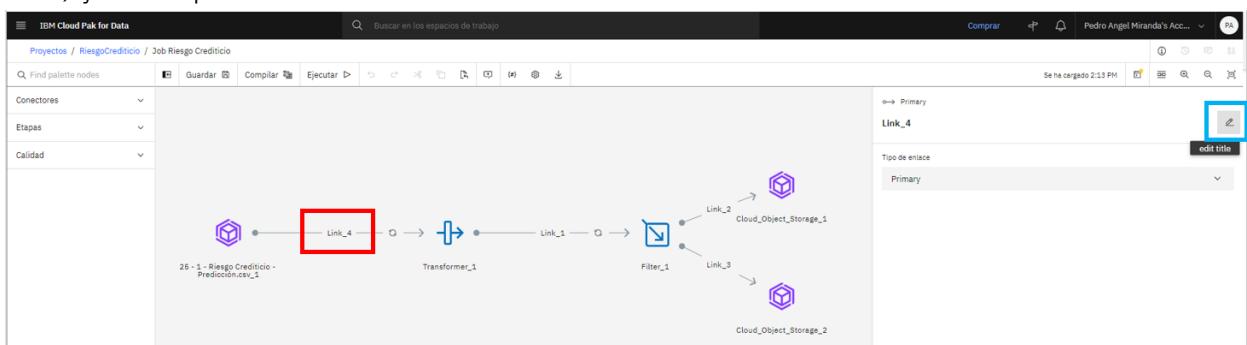
Una vez que tenga todos sus componentes enlácelos como se muestra a continuación. De lado derecho de cada componente tendrá un icono “>” (color amarillo) el cual tendrá que arrastrar hasta el siguiente componente(flecha verde), primero una el Cloud Object Storage “25 - 1 - Riesgo Crediticio - Predicción.csv” (color rojo) con el “Transformer” (color azul), el “Transformer” con el “Filter” (color azul) y finalmente el “Filter” con ambos Cloud Object Storage (color rosa).



6.d) Nombraremos cada componente, posíónese sobre un componente y vera que aparecerá un lápiz (color amarillo) cambie los nombres “Transformer\_1” por “Status\_de\_Credito”, “Filter\_1” por “Filtro”, “Cloud\_Object\_Storage\_1” por Aprobados y finalmente “Cloud\_Object\_Storage\_2” por “No\_aprobados”



Para nombrar los enlaces de doble clic en cualquier enlace en este caso “Link\_4” (color rojo) y de lado derecho aparecerá una ventana en donde se encuentra un lápiz (color azul) y ahora podrá renombrarlos.



Nómbrellos como se muestra a continuación.

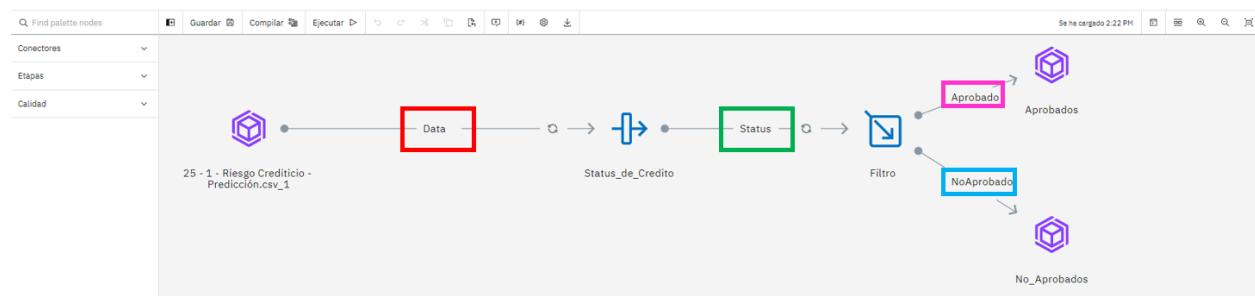
25 - 1 - Riesgo Crediticio - Predicción.csv – Status\_de\_Credito = Data (color rojo)

Status\_de\_Credito-Filtro= Status (color verde)

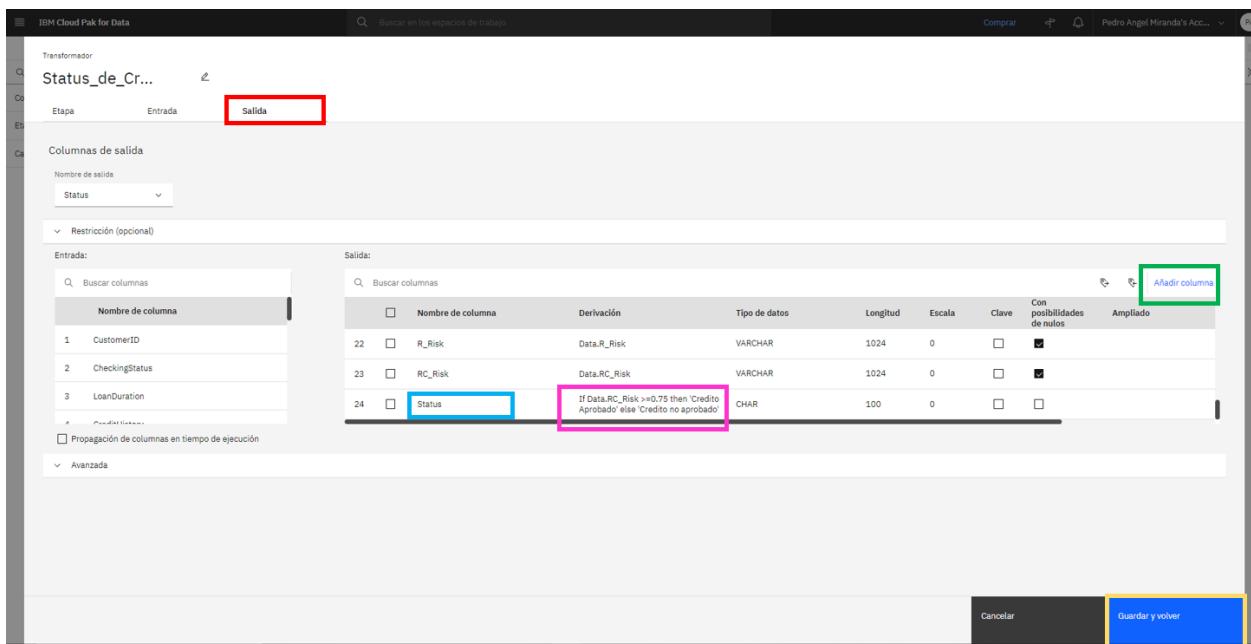
Filtro-Aprobados=Aprobado (color rosa)

Filtro-No\_Aprobados=NoAprobado (color azul)

De manera que le quede como se muestre a continuación:



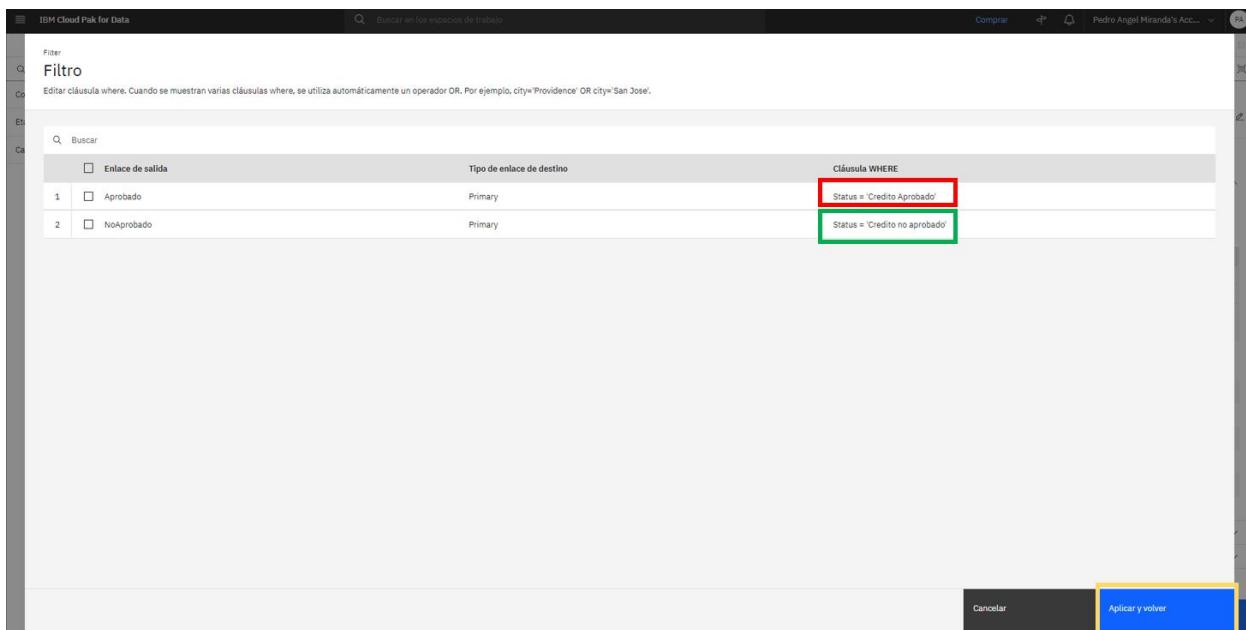
6.e) Ahora vamos a configurar nuestro Transformer llamado “Status\_de\_Credito”, de doble click sobre el componente, var a la parte de “Salida” (color rojo), de click en “Añadir columna” (color verde), vaya hasta la parte de debajo y vera que tiene una nueva columna la cual vamos a editar con el lápiz que se encuentra en el cuadro (color azul) y la nombraremos como “Status”, así mismo en “Derivación” (color rosa) pondremos **If Data.RC\_Risk >=0.75 then 'Credito Aprobado' else 'Credito no aprobado'**, hacemos click en Guardar y volver (color amarillo) para finalizar damos click en “Save”



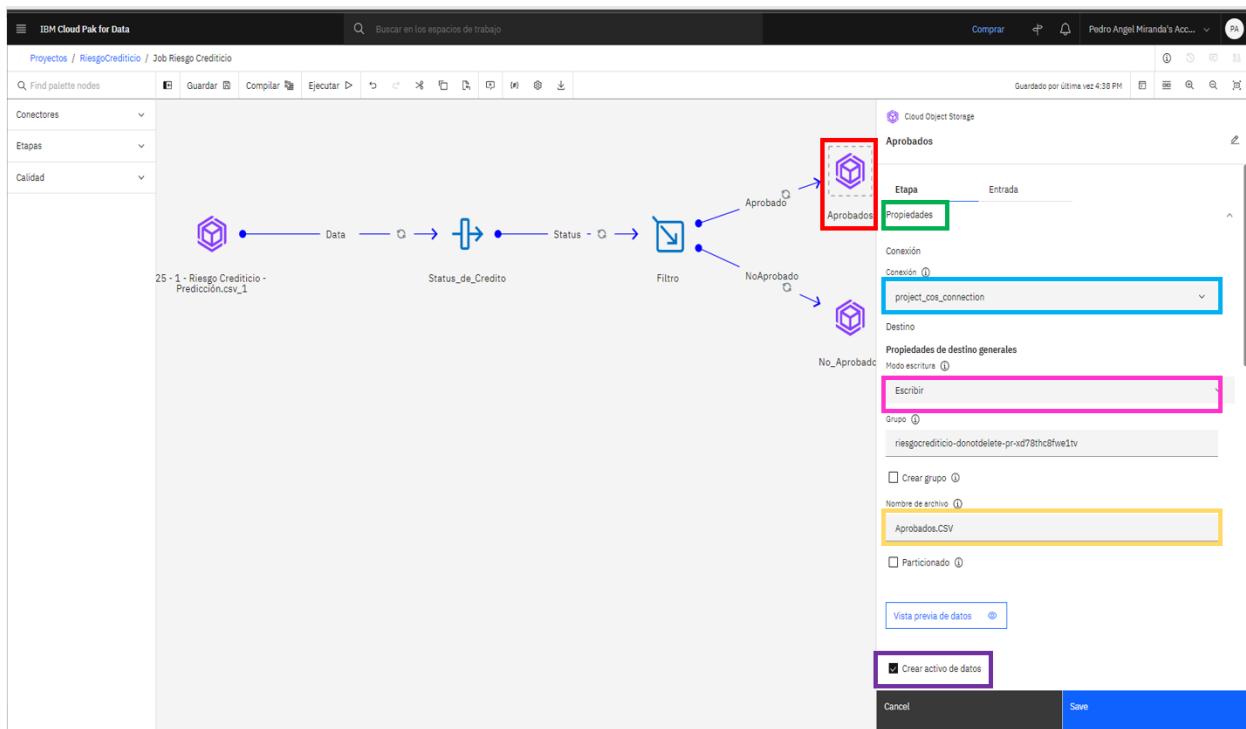
Nombre de columna	Derivación	Tipo de datos	Longitud	Escala	Clave	Con posibilidades de nulos	Ampliado
R_Risk	Data.R_Risk	VARCHAR	1024	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
RC_Risk	Data.RC_Risk	VARCHAR	1024	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Status	If Data.RC_Risk >=0.75 then 'Credito Aprobado' else 'Credito no aprobado'	CHAR	100	0	<input type="checkbox"/>	<input type="checkbox"/>	

6.f) Vaya al componente Filter llamado "Filtro" de doble clic, expanda la pestaña llamada "Propiedades", en la parte de "predicados" de clic en "Edit" y en Cláusula WHERE coloque

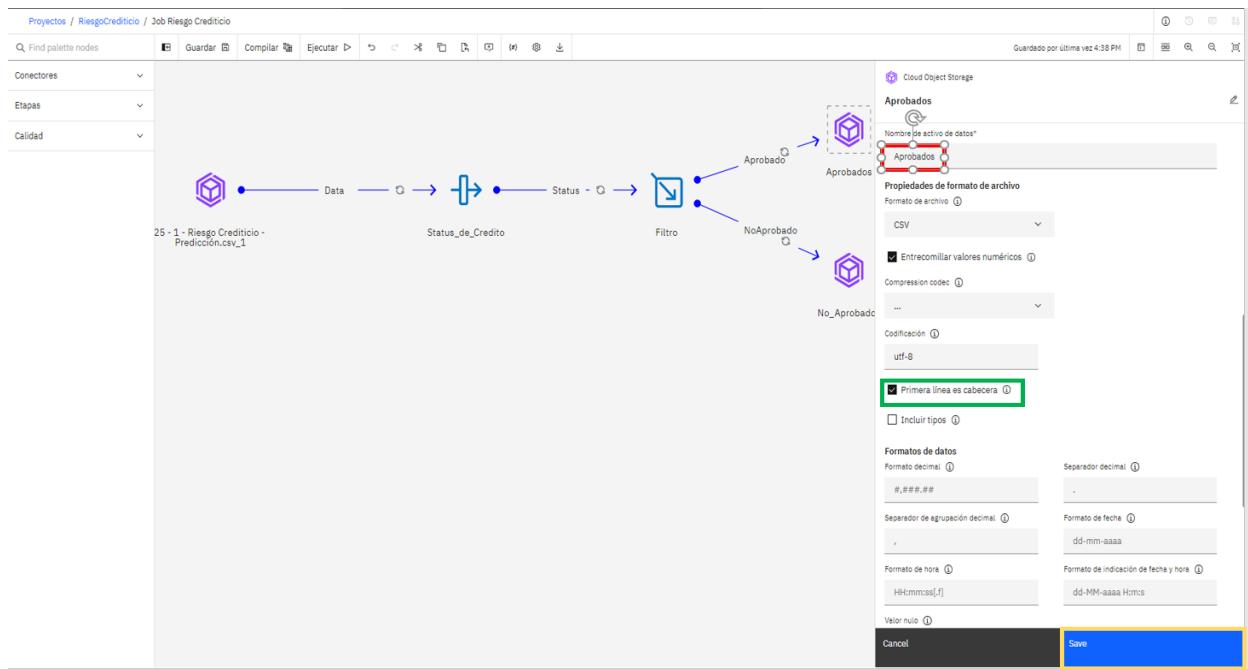
Para Aprobados **Status = 'Credito Aprobado'** (color rojo) y para NoAprobado **Status = 'Credito no aprobado'** (color verde) finalmente de click en Aplicar y volver (color amarillo)



6.g) vaya al componente Cloud Object Storage llamado “Aprobado” (color rojo) de doble click y expanda la pestaña “Propiedades”(color verde) en la parte de “Conexión” seleccione la única que existe(color azul), en “Modo de escritura” ponemos “Escribir”(color rosa), en “Nombre del archivo” ponemos **Aprobados.CSV** (color amarillo), palomeamos el cuadro que dice “Crear activos de datos”(color morado).

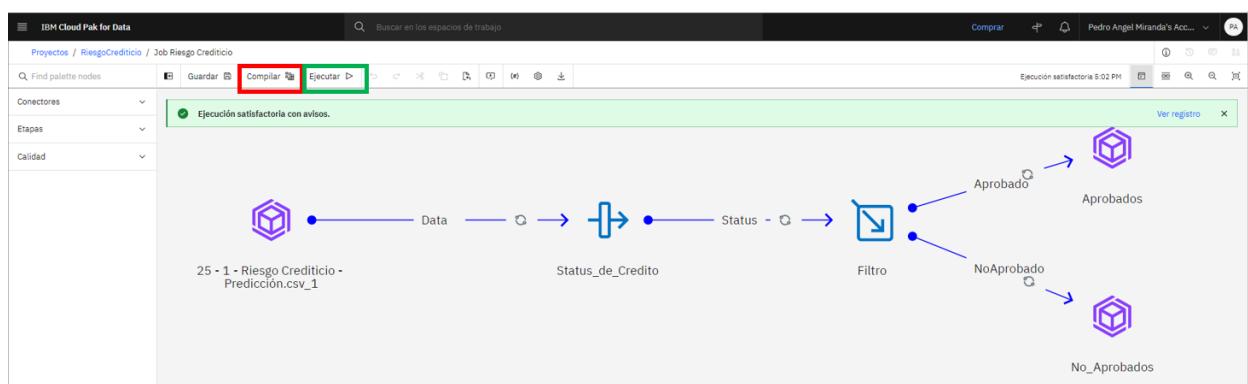


En “Nombre de activo de datos” ponemos **Aprobados** (color rojo), palomeamos el cuadro que dice “Primera línea es cabecera” (color verde) y finalmente hacemos click en “Save” (color amarillo)



Repita el paso 6.g) pero ahora con componente Cloud Object Storage llamado “No\_Aprobado”, cambiando solo las partes de “Nombre del archivo” a NoAprobados.CSV y “Nombre de activo de datos” a NoAprobados.

6.h) Su secuencia de DataStage esta lista para correr. Vaya y haga click en “Compilar” (color rojo), si no tiene errores le mostrara un mensaje en color verde que dice “La compilación ha sido satisfactoria”, ahora haga click en “Ejecutar” (color verde) este paso tardara un par de minutos cuando termine le mostrara un mensaje de “Ejecución satisfactoria”.



Para validar que tenemos datos en las salidas colóquese en el componente Cloud Object Storage “Aprobados” de click izquierdo y seleccionar “Vista previa de los datos”, podrá darse cuenta que tenemos la nueva columna con puros Status “Credito Aprobado”.

	residenceDuration	OwnsProperty	Age	InstallmentPlans	Housing	ExistingCreditsCount	Job	Dependents	Telephone	ForeignWorker	R_Risk	RC_Risk	Status
1	savings_insurance	32	none	own	1	skilled	1	none	yes	No Risk	0.8698852	Credito Aprobado	
2	real_estate	28	none	own	2	skilled	1	yes	no	No Risk	0.878049	Credito Aprobado	
3	savings_insurance	22	none	own	1	management_self-employed	1	none	yes	No Risk	0.8698852	Credito Aprobado	
4	real_estate	19	none	rent	1	management_self-employed	1	none	yes	No Risk	0.993902	Credito Aprobado	
5	car_other	52	none	own	1	skilled	2	yes	yes	No Risk	0.793893	Credito Aprobado	
6	real_estate	26	none	own	1	skilled	1	none	yes	No Risk	0.988372	Credito Aprobado	
7	real_estate	28	none	own	1	skilled	1	none	yes	No Risk	0.988372	Credito Aprobado	
8	savings_insurance	24	none	own	2	skilled	2	yes	yes	No Risk	0.83871	Credito Aprobado	
9	real_estate	19	stores	rent	1	skilled	1	none	yes	No Risk	0.835294	Credito Aprobado	
10	real_estate	19	stores	rent	1	skilled	1	none	yes	No Risk	0.878049	Credito Aprobado	
11	car_other	38	none	own	1	skilled	2	yes	yes	No Risk	0.860606	Credito Aprobado	
12	real_estate	31	none	own	1	skilled	1	none	yes	No Risk	0.983607	Credito Aprobado	
13	savings_insurance	35	none	own	1	skilled	1	yes	yes	No Risk	0.793893	Credito Aprobado	
14	real_estate	27	none	own	1	skilled	1	none	yes	No Risk	0.958763	Credito Aprobado	

Para finalizar regrese a su proyecto (color rojo) y vera que ahora tiene dos nuevos sets de datos en su proyecto (color verde)

IBM Cloud Pak for Data

Proyectos RiesgoCredito

Activos

Visión general Trabajos Gestión

Buscar activos

Añadir activo Nuevo activo

Nombre	Última modificación
NoAprobados CSV	Hace 35 minutos Pedro Angel Miranda (United)
Aprobados CSV	Hace 35 minutos Pedro Angel Miranda (United)
25-1 - Riesgo Crediticio - Predicción.csv	Hace 1 día Pedro Angel Miranda (United)
10-1 - Riesgo Crediticio - Solicitantes.csv	Hace 3 días Pedro Angel Miranda (United)
18-1 - Riesgo Crediticio - Histórico.csv	Hace 3 días Pedro Angel Miranda (United)

Tipos de activos

- Acceso a datos
- Activo de datos
- Flujos
- Visualizaciones
- Código fuente

11 assets

✓ Todos los activos

83 patos

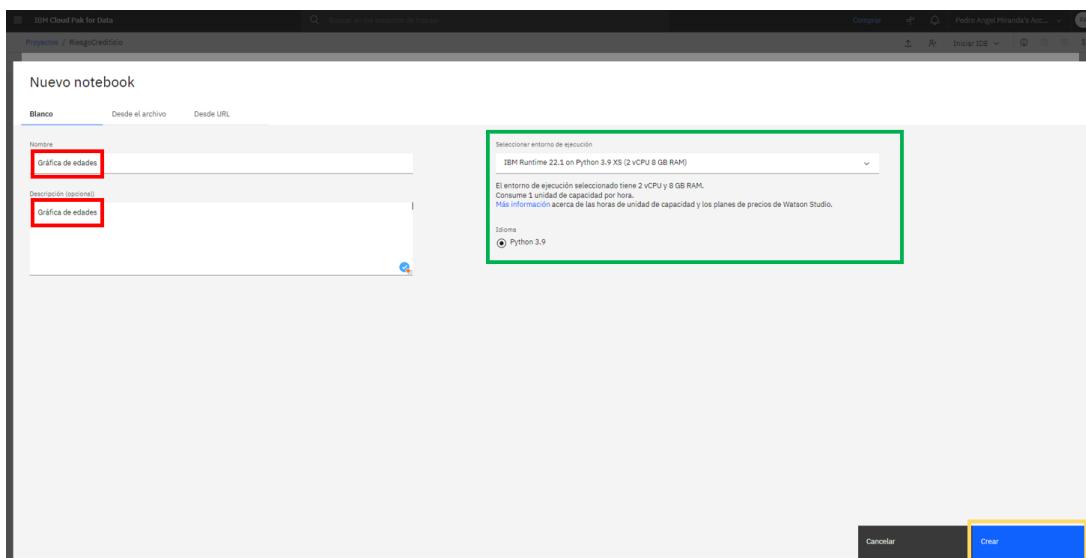
Felicitaciones ha terminado su flujo de DataStage

## 11 . Crear un Notebook de Jupyter

Regrese a la parte de “Activos” dé clic al botón azul “Nuevo activo” (color rojo).

y seleccione “Editor de cuaderno de Jupyter” (color verde).

11.a) Nombre su Notebook y la descripción como “Gráfica de edades” (color rojo) en la sección Seleccionar entorno de ejecución utilice “IBM Runtime 22.1 on Python 3.9 XS (2vCPU y 8GB RAM)”, Idioma “Python 3.9”(color verde) y oprima la opción de “Crear” (color amarillo).



Se comenzará a crear una instancia de entorno.

11.b) De click en cell (color rojo), vaya a la opción Cell Type y seleccione Markdown y ponga la siguiente descripción (color verde):

**# Gráfica de Edades**

Vamos a visualizar la información de un activo y graficaremos una columna



11.c) Vaya a la sección Insert (color rojo) y agregue dos celdas en “Insert Cell Below”:

- Colóquese en la primera celda después vaya a Buscar y añadir datos (color verde), seleccione “18-1 - Riesgo ... - Histórico.csv” despliegue Insertar en código y escoja pandas DataFrame(color rosa).
- Colóquese en la segunda celda y ponga el siguiente código (color amarillo):

```
import matplotlib.pyplot as plt
import numpy as np
plt.bar((np.arange(0, len(df_data_1.Age), 1)[0:5]), (df_data_1.Age[0:5]))
plt.show()
```

The screenshot shows a Jupyter-style notebook interface. At the top, there's a toolbar with 'File', 'Edit', 'View', 'Insert' (which is highlighted in red), 'Cell', 'Kernel', and 'Help'. The status bar indicates the notebook is saved, using Python 3.9, and shows the user's account.

**Code Cells:**

- In [1]:** Contains Python code for reading a CSV file from Object Storage and displaying its first 5 rows. It also includes a call to plt.show() at the end.
- In [2]:** Contains the same bar chart code as the first cell, enclosed in a yellow box.

**Datos Panel:**

- Archivos:** Shows a placeholder for dragging files and a dropdown menu with "18-1 - Riesgo ... - Histórico.csv". This dropdown has "pandas DataFrame" selected, highlighted in pink.
- Conexiones:** Shows dropdown menus for "Aprobados", "Insertar en código", and "NoAprobados".

11.d) Finalmente corra cada celda (color rojo), podrá ver los primero cinco registros de nuestro activo de datos además de ver en la gráfica las edades de esos registros.

The screenshot shows a Jupyter Notebook interface with the following details:

- Title Bar:** IBM Cloud Pak for Data / RiesgoCredito / Gráfica de edades
- Toolbar:** File, Edit, View, Insert, Cell, Kernel, Help, Run (highlighted in red)
- Code Cell [18]:**

```
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
client = ibm.object_storage.ObjectStorageClient(
    endpoint='https://object-storage-01.cloud-object-storage.appdomain.cloud',
    ibm_api_key_id='1rnn...7dbX0V4PCKCFQG_f5e5e040CCKUpjyfRgWid',
    ibm_auth_endpoint='https://iam.cloud.ibm.com/oidc/token',
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.us.cloud-object-storage.appdomain.cloud'
)

body = client.get_object(Bucket='riesgo_credito-donotdelete-pr-xd78thc8fweiyv',Key='18-1 - Riesgo Credito - Histórico.csv')['Body']

# and read it as a pandas DataFrame object, or pandas accepts body as file-like object
if not hasattr(body, '__iter__'): body = iter([body])
df_data_1 = pd.read_csv(body)
df_data_1.head()
```
- Data Preview:** Out[18] displays the first 5 rows of the dataset.

	CustomerID	CheckingStatus	LoanDuration	CreditHistory	LoanPurpose	LoanAmount	ExistingSavings	EmploymentDuration	InstallmentPercent	Sex	OwnsProperty	Age	InstallmentPlans	Housing	ExistingCreditsCount	Job	Dependents	Telephone	ForeignWorker	Risk
0	4e561ef3-a927-4107-aa12-16a5772e0d3	less_0	6	at_credit_paid_back	car_used	250	less_100	1_M-4	2	male	... savings_insurance	26	stores	rent	1	skilled	1	none	yes	Risk
1	20815-4053-409-4cda-8	less_0	14	all_credit_paid_back	appliances	1431	less_100	unemployed	1	female	... car_other	25	stores	own	1	skilled	1	none	yes	Risk
2	c13240f7-7b76-4265-89b6-ea1185654f3	greater_200	5	credit_paid_to_date	car_used	260	less_100	4_M-7	3	male	... savings_insurance	42	none	rent	1	skilled	1	none	yes	No Risk
3	c19905c4-2b0a-4f66-8550-13a2a090908	less_0	11	all_credit_paid_back	furniture	2016	less_100	less_1	1	female	... car_other	34	bank	own	1	unemployed	1	none	yes	No Risk
4	c8b7ae96-7354-4ef5-835a-4cb61008ca1	less_0	4	credits_paid_to_date	car_new	250	less_100	unemployed	1	female	... real_estate	29	none	rent	1	skilled	1	none	yes	No Risk

5 rows × 22 columns

- Code Cell [19]:**

```
In [19]: import matplotlib.pyplot as plt
import numpy as np
plt.bar(np.arange(0, len(df_data_1.Age),1 )[0:5],(df_data_1.Age[0:5]))
plt.show()
```
- Figure:** A histogram showing the distribution of ages for the first five rows of data. The x-axis represents age groups (0, 1, 2, 3, 4) and the y-axis represents frequency (0 to 40+).

## 12. AutoAI y Despliegue de modelos

Regrese al proyecto y vaya a la parte de “Gestionar” (color verde), despues el la seccion “Servicios e integraciones” y finalmente en asociar un servicio (color amarillo)

The screenshot shows the 'Gestionar' tab selected in the top navigation bar. Below it, the 'Servicios e integraciones' section is active. A yellow box highlights the 'Asociar servicio' button at the top right of the service list area.

### 12.a) Palomea “WatsonMachineLearning” y seleccionar Asociar servicio (color rojo)

The screenshot shows the 'Asociar servicio' dialog box. A red box highlights the 'Asociar servicio' button at the top right. The table lists several services, with 'WatsonMachineLearning' selected and its status changed to 'Asociado'.

Nombre	Tipo	Plan	Ubicación	Estado	Grupo
Watson OpenScale-op	Watson OpenScale	Lite	Dallas	No asociado	Default
Data Replication-kd	Data Replication	Lite	Global	No asociado	Default
IBM Cognos Dashboard Embedded-db	IBM Cognos Dashboard Embedded	Lite	Dallas	Asociado	Default
<input checked="" type="checkbox"/> WatsonMachineLearning	Machine Learning	Lite	Dallas	No asociado	Default

Si todo sale bien vera una paloma verde en la parte de “Estado” y si servicio estara asociado al proyecto

The screenshot shows the 'Asociar servicio' dialog box with a green message at the top stating 'El servicio WatsonMachineLearning se ha asociado correctamente al proyecto RiesgoCrediticio.' A red box highlights the row for 'WatsonMachineLearning' in the table, which now shows a green status icon in the 'Estado' column.

Nombre	Tipo	Plan	Ubicación	Estado	Grupo
Watson OpenScale-op	Watson OpenScale	Lite	Dallas	No asociado	Default
Data Replication-kd	Data Replication	Lite	Global	No asociado	Default
IBM Cognos Dashboard Embedded-db	IBM Cognos Dashboard Embedded	Lite	Dallas	Asociado	Default
<input checked="" type="checkbox"/> WatsonMachineLearning	Machine Learning	Lite	Dallas	Asociado	Default

12.b) Regrese a la parte de “Activos” dé clic al botón azul “Nuevo activo” (color rojo).

y seleccione “Auto AI” (color verde).

12.c) Nombre su experimento de AutoAI y la descripción como “Modelo Riesgo Crediticio” (color rojo) y oprima la opción de “Crear” (color amarillo).

New asset  
Nuevo experimento de AutoAI

+ Nuevo  
ID: Ejemplo de Gallery

Definir detalles

Nombre: **Modelo Riesgo Crediticio**

Descripción: **Modelo Riesgo Crediticio**

Definir configuración

Instancia de servicio de Watson Machine Learning: WatsonMachineLearning

Definición de entorno: 8 vCPU y 32 GB de RAM

Etiquetas (opcional): Comience escribiendo para añadir etiquetas

Crear

Se mostrará la siguiente pantalla, de click en Seleccionar desde proyecto.

Configurar experimento AutoAI

Modelo Riesgo Crediticio

Añadir orígenes de datos

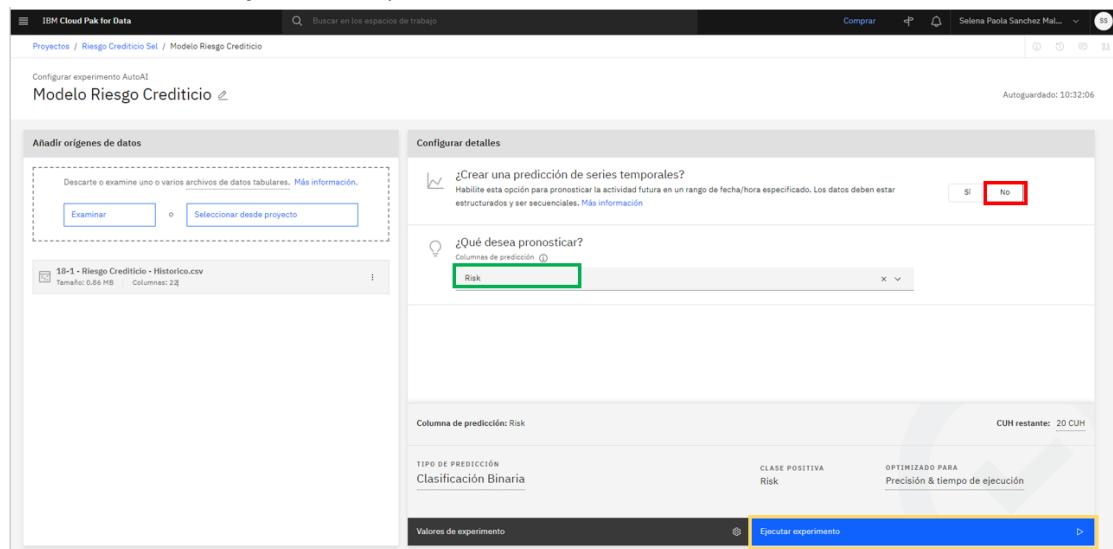
Añadir orígenes de datos  
Drop or browse for one or more tabular data files. [Más información.](#)

Examinar o Seleccionar desde proyecto

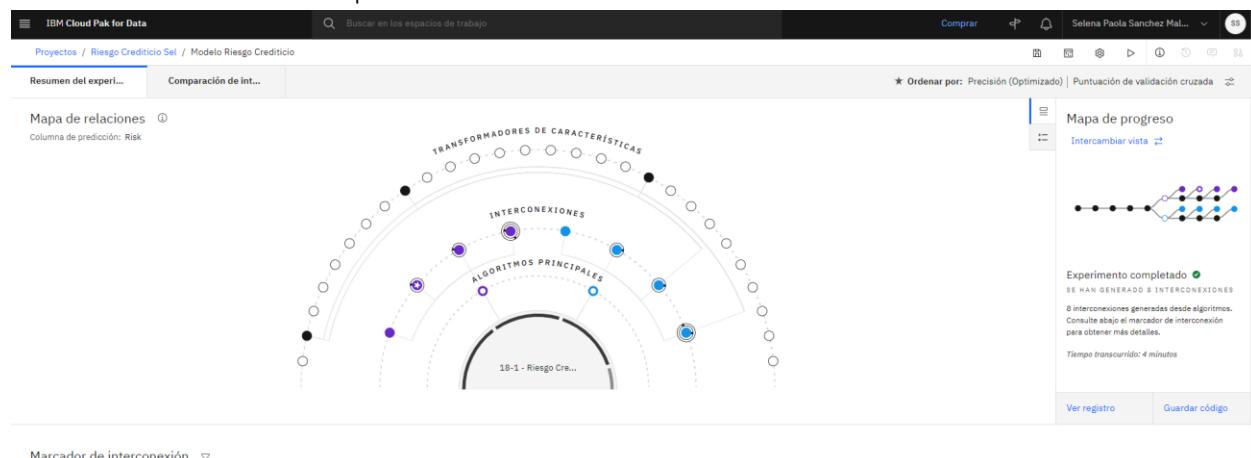
12.d) A continuación, le aparecerá la siguiente pantalla, y debe dar clic en “Activos de Datos” (color rojo), aparecerán a la derecha las fuentes de datos que tenemos en nuestro proyecto, seleccione la fuente de datos “18-1 - Riesgo Crediticio – Historico.csv” (color rojo) y finalmente de clic en “Seleccionar archivo” (color amarillo, esquina inferior derecha).

The screenshot shows the 'Activos de datos de RiesgoCrediticio' (Assets of RiesgoCrediticio) screen in the IBM Cloud Pak for Data interface. On the left, there's a sidebar with 'Categorías' and a 'Conexión' section. Below it, a red box highlights the 'Activo de datos' section. In the center, there's a grid titled 'Activos de datos' with several rows. One row, '18-1 - Riesgo Crediticio - Histórico.csv', has a checked checkbox and is also highlighted with a red box. To the right, a panel titled 'Activos seleccionados' shows a single item: '18-1 - Riesgo Crediticio - Histórico.csv'. At the bottom right of the panel, a blue button labeled 'Seleccionar activo' is highlighted with a yellow box.

12.e) Le preguntara si desea crear una predicción de seria temporales, seleccionar **no** (color rojo). En la parte de Que desea pronosticar seleccionar la opción **Risk** (color verde) ahora dar click en Ejecutar experimento (color amarillo)



Durante el proceso se irán enlistando los pipelines generados con los resultados obtenidos con la aplicación del algoritmo seleccionado por AutoAI y estos estarán ordenados en base a su precisión.



Marcador de interconexión ▾

Clasificación	Nombre	Algoritmo	Precisión (Optimizado) Validación Cruzada	Mejoras	Tiempo de creación
★ 1	2 interconexión	Clasificador XGB	0.803	HPO-1	00:00:13
2	3 interconexión	Clasificador XGB	0.803	HPO-1 FE	00:00:46
3	4 interconexión	Clasificador XGB	0.803	HPO-1 FE HPO-2	00:00:33

A continuación podrá ver como se listaron los mejores modelos y el más apto para predecir la variable es el que tiene un símbolo de estrella.

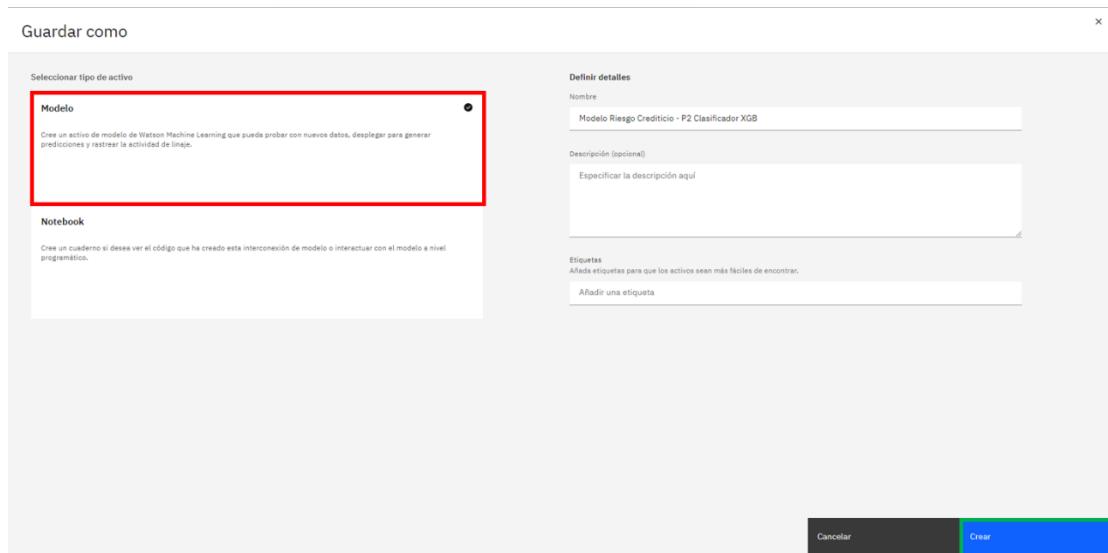
Clasificación	Nombre	Algoritmo	Precisión (Optimizado) Validación Cruzada	Mejoras	Tiempo de creación
★ 1	2 Interconexión	Clasificador XGB	0.803	HPO-1	00:00:13
2	3 Interconexión	Clasificador XGB	0.803	HPO-1 FE	00:00:46
3	4 Interconexión	Clasificador XGB	0.803	HPO-1 FE (HPO-2)	00:00:33
4	1 Interconexión	Clasificador XGB	0.798	Ninguno	00:00:01
5	7 Interconexión	Clasificador de árboles adicionales	0.788	HPO-1 FE	00:00:27
6	8 Interconexión	Clasificador de árboles adicionales	0.788	HPO-1 FE (HPO-2)	00:00:01
7	5 Interconexión	Clasificador de árboles adicionales	0.781	Ninguno	00:00:01
8	6 Interconexión	Clasificador de árboles adicionales	0.781	HPO-1	00:00:01

Una vez concluido el proceso, podemos guardar el modelo de Machine Learning generado o bien crear un Jupyter notebook para recrear el análisis que llevo a cabo AutoAI en Python.

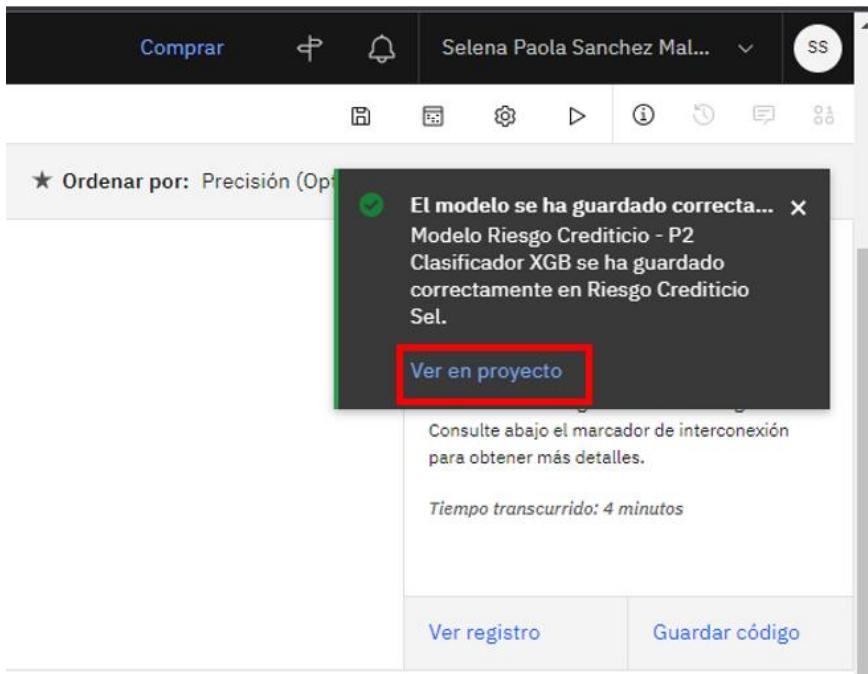
12.f) Seleccione el pipeline con la mejor clasificación (con la estrella) y haga click en el botón “Guardar como”(color rojo)

Clasificación	Nombre	Algoritmo	Precisión (Optimizado) Validación Cruzada	Mejoras	Tiempo de creación	
★ 1	2 interconexión	Clasificador XGB	0.803	HPO-1	00:00:13	<b>Guardar como</b>

12.g) Selecciona “Modelo” (color rojo) y haz click en el botón “Crear”(color verde). (Opcional: puede repetir el paso seleccionando “Notebook” para crear el Jupyter notebook en Python y revisarlo en el proyecto en la sección de “Notebook”



12.h) Haz click en “Ver en proyecto” (color rojo)



12.i) Ahora vamos a desplegar nuestro modelo para que pueda ser consumido por otras aplicaciones y puedan aprovechar el uso del modelo.

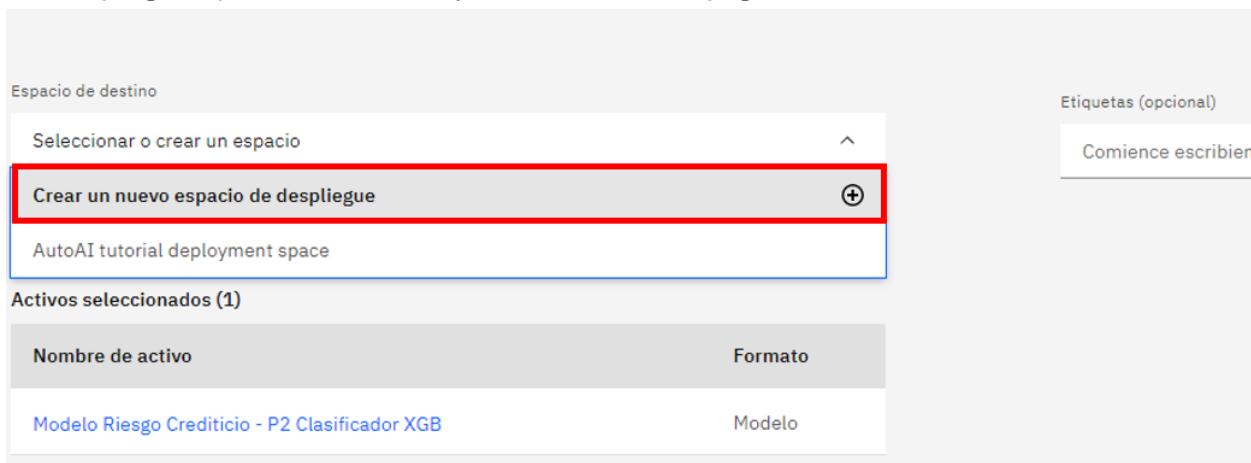
Haz click en “Promocionar al espacio de despliegue” (color rojo)

The screenshot shows the "Modelo Riesgo Crediticio - P2 Clasificador XGB" page. On the right side, there is a "Promocionar al espacio de despliegue" button, which is highlighted with a blue rectangular box. To the right of the button, there is a detailed view of the model's deployment space, including its ID, creation date, type, and software specifications.

12.j) Haz click en “Seleccionar o crear un espacio” y luego selecciona “Crear nuevo espacio de despliegue”

### Promocionar al espacio

Utilice un espacio de despliegue para organizar recursos de soporte como, por ejemplo, datos de entrada y entornos; desplegar modelos o funciones para generar predicciones o soluciones; y ver o editar detalles de despliegue.



Espacio de destino

Seleccionar o crear un espacio

**Crear un nuevo espacio de despliegue**

AutoAI tutorial deployment space

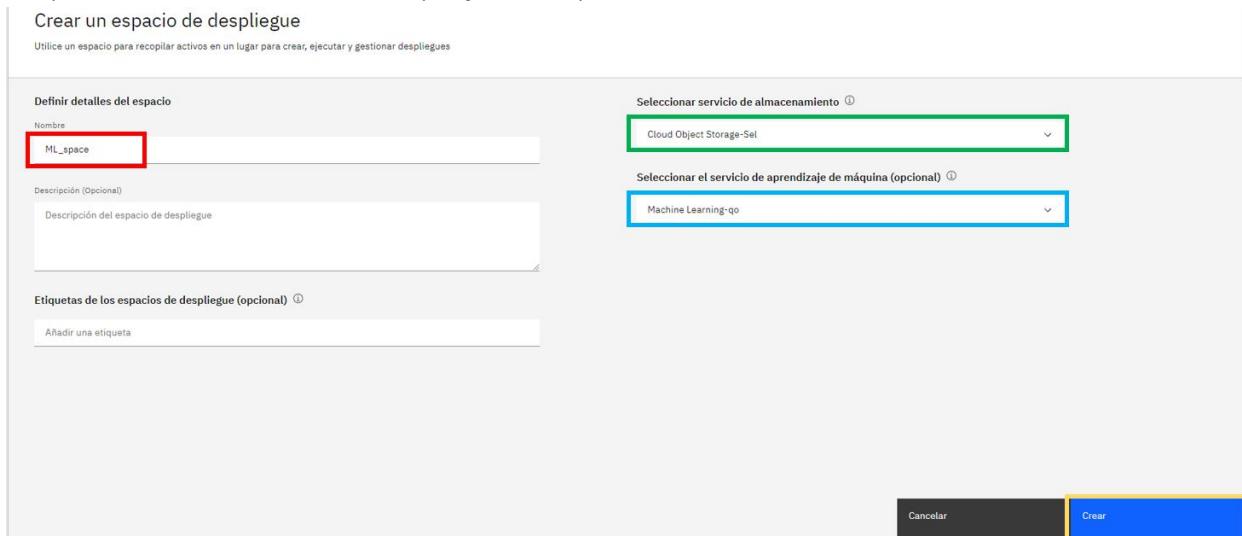
Activos seleccionados (1)

Nombre de activo	Formato
Modelo Riesgo Crediticio - P2 Clasificador XGB	Modelo

Etiquetas (opcional)

Comience escribiendo...

12.k) En Nombre introduzca “ML\_space” (color rojo) y asegúrese que en seleccionados los servicios Almacenamiento (color verde) y Machine Learning(color azul) (que fueron creados al inicio del workshop). Haz click en el botón “Crear” (color amarillo). Tomará un par de minutos crear el deployment space



Crear un espacio de despliegue

Utilice un espacio para recopilar activos en un lugar para crear, ejecutar y gestionar despliegues

Definir detalles del espacio

Nombre  
**ML\_space**

Descripción (Opcional)

Descripción del espacio de despliegue

Seleccionar servicio de almacenamiento ⓘ

Cloud Object Storage-Sel

Seleccionar el servicio de aprendizaje de máquina (opcional) ⓘ

Machine Learning-qo

Etiquetas de los espacios de despliegue (opcional) ⓘ

Añadir una etiqueta

Cancelar **Crear**

Al finalizar la creación del Espacio de despliegue, haz click en el botón “Cerrar”

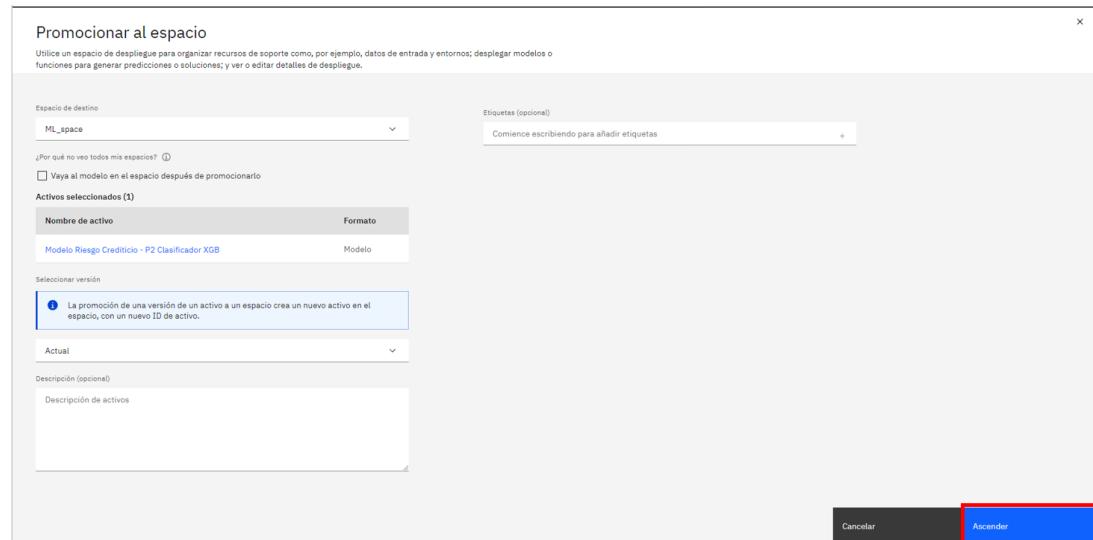
El espacio está listo

Cierre esta notificación para reanudar su trabajo. Pulse Deployments en el panel de navegación para ver y acceder al nuevo espacio.

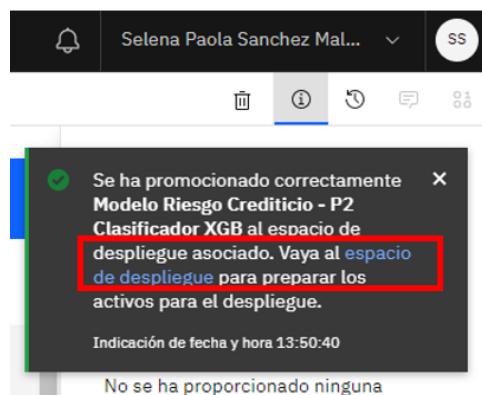
✓ Paso 1 de 1. Creación del espacio de despliegue.

**Cerrar**

12.l) Haz click en el botón "Ascender" (color rojo)



12.m) Una vez promovido nuestro modelo, haz click en el link “espacio de despliegue” (color rojo)



12.n) Posiciona el cursor sobre el modelo y haz click en el ícono  (color rojo)

Nombre	Tipo	Especificación de software	Etiquetas	Última modificación	
 Modelo Riesgo Crediticio - P2 Clasificador XGB	wml-hybrid_0.1	hybrid_0.1		28 de abr. de 2022 13:50	 

12.ñ) Nuestro modelo puede ser consumido de forma En línea o por lotes. Seleccione “En línea”(color rojo) e introduzca en nombre “Modelo\_Riesgo\_Crediticio” (color verde). Haz click en el botón “Crear”(color amarillo)

Crear un despliegue

Activo asociado  
Modelo Riesgo Crediticio - P2 Clasificador XGB

Tipo de despliegue

En línea  
Ejecute modelo sobre datos en tiempo real, puesto que los datos los recibirá un servicio web.

Por lotes  
Ejecute modelo con relación a los datos como un proceso por lotes.

Nombre

Nombre del servicio 

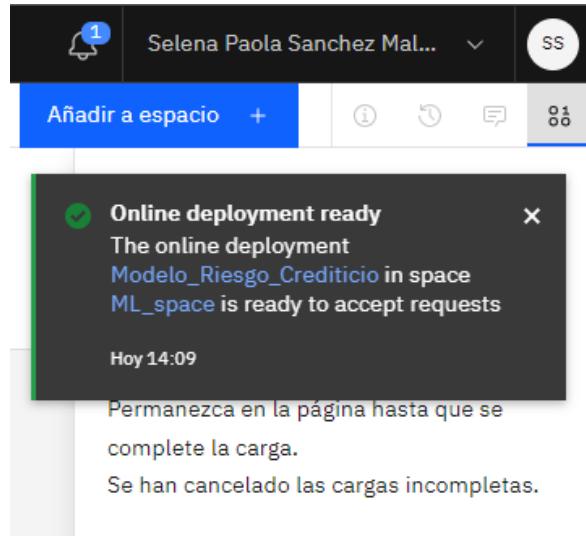
Nombre de servicio de despliegue

Descripción

Etiquetas  
Añadir etiquetas para que sea más fácil encontrar los activos.

Aparecerá un mensaje de que fue exitoso el despliegue del modelo en unos segundos.



12.o Vaya a la sección “Despliegues” en la barra de menú.

Podemos observar que el modelo fue **Desplegado** y procederemos a probarlo.  
Haga click sobre el nombre “**Modelo\_Riesgo\_Crediticio**”

Nombre	Tipo	Estado	Activo	Etiquetas	Última modificación
Modelo_Riesgo_Crediticio	En linea	Desplegado	Modelo Riesgo Crediticio - P2 Clasificador XGB		28 de abr. de 2022 14:09

Al desplegar el modelo se generan automáticamente el URL Endpoint (color rojo) con el cual puede ser llamado el Modelo por otras aplicaciones, así mismo se muestran ejemplos de Code Snippets (color verde) en los diferentes lenguajes de programación, navegue por estas opciones.

12.p) Al finalizar, de click en la sección de “Prueba” (color rosa)

```
# NOTE: you must set $API_KEY below using information retrieved from your IBM Cloud account.
curl --insecure -X POST -H "Content-Type: application/x-www-form-urlencoded" --header "Accept: application/json"
--data-uzlenicode "grant_type=password;username=$grant_type;password=$password"
--data-uzlenicode "apikey=$API_KEY" "https://iam.cloud.ibm.com/identity/token"
# The above CURL request will return an auth token that you will use as $IAM_TOKEN in the scoring request below
# TODO: manually define any pass values to be scored below
curl -X POST -H "Content-Type: application/json" --header "Accept: application/json" --header "Authorization: Bearer $IAM_TOKEN" -d '{"input_data": [{"fields": ["$ARRAY_OF_INPUT_FIELDS"], "values": ["$ARRAY_OF_VALUES_TO_BE_SCORED", "$ANOTHER_ARRAY_OF_VALUES_TO_BE_SCORED"]}], "url": "https://us-south.ml.cloud.ibm.com/v4/deployments/f577be6e-61c9-487f-a7ca-2a4ce59c1012/predictions?version=2022-04-28"}'
```

12.q) Para simular la evaluación en el otorgamiento de un crédito, introduzca algunos valores ficticios como **Loanduration**, **Age**, **LoanAmount**, **CreditHistory**, etc. (entre más datos, mejor) Con esta información, el modelo evaluará y arrojará el porcentaje de probabilidad de que podamos pagar el crédito. Haz click en “Añadir a la lista” (color rojo). Añada tantos casos quiera evaluar

### Modelo\_Riesgo\_Crediticio Desplegado En Línea

Referencia de API Prueba Detalles del despliegue

Especificar datos de entrada

CustomerID	4e561ef3-a927-4107-aa12-18ac5772ec63
CheckingStatus	less_0
LoanDuration	6
CreditHistory	all_credits_paid_back
LoanPurpose	car_used
LoanAmount	250
ExistingSavings	less_100

**Añadir a la lista** +

Por ejemplo:

Customer ID: 4e561ef3-a927-4107-aa12-18ac5772ec63  
CheckingStatus: less\_0  
LoanDuration: 6  
CreditHistory: all\_credit\_paid\_back  
LoanPurpose: car\_used  
LoanAmount: 250  
ExistingSavings: less\_100

12.r) Cuando haya completado su lista de casos, haga click en el botón “Predecir” para evaluar los casos con el modelo

**Lista de entrada (1)**

```
[ 4e561ef3-a927-4107-aa12-18ac5772ec63, less_0,
  6, all_credits_paid_back, car_used, 250, less_100, 1_to
  _4, 2, male, none, 2, savings_insurance, 28, stores, ren
  t, 1, skilled, 1, none, yes ]
```

**Predecir (1)**

Con base a la información de cada caso, el modelo arrojará la probabilidad de otorgar o no el crédito solicitado

**Resultado**

```
0 {
1   "predictions": [
2     {
3       "fields": [
4         "prediction",
5         "probability"
6       ],
7       "values": [
8         [
9           "No Risk",
10          [
11            0.9602444767951965,
12            0.03975552320480347
13          ]
14        ]
15      ]
16    }
17  ]
18 }
```

En este ejemplo, podemos ver que nos arroja “No Risk” lo que nos dice que no hay un riesgo de falta de pago, y con la predicción del 0.93, el crédito es aprobado.