

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

- 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.

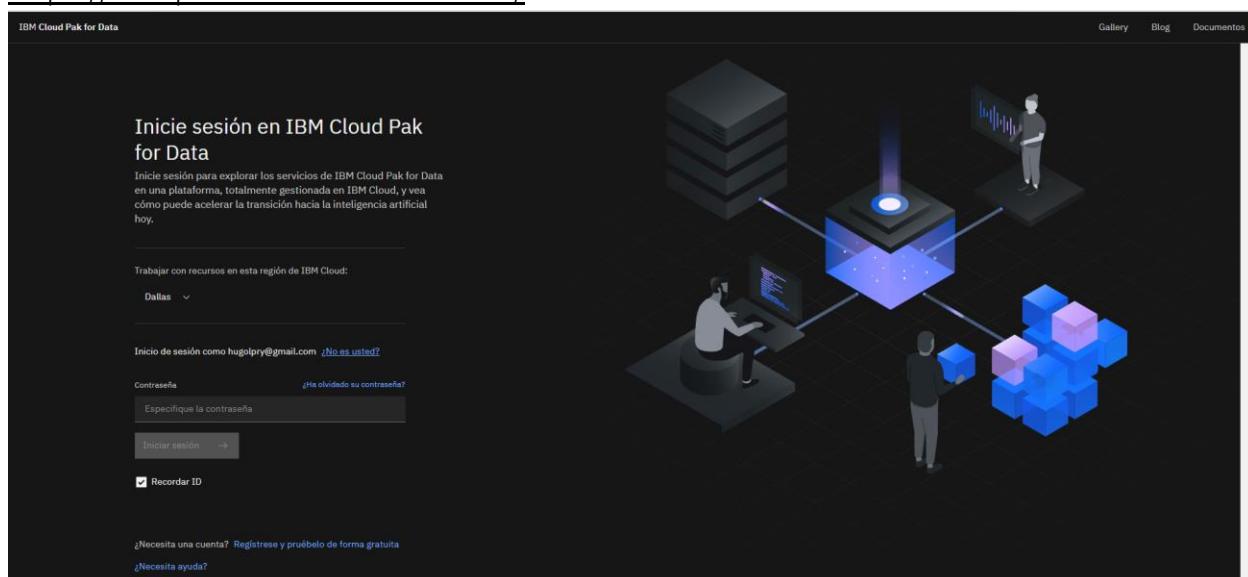
- b) Debe cargar sus datos en la plataforma (archivos del punto a)

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:

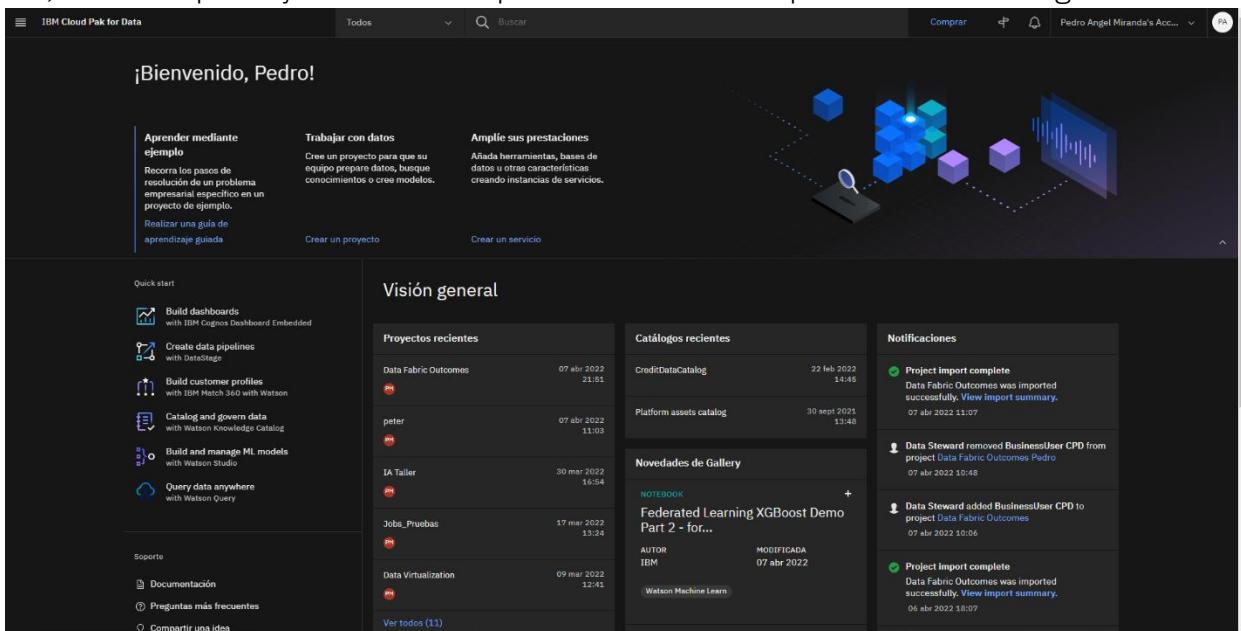
<https://dataplatform.cloud.ibm.com/>



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 'IBM Cloud Pak for Data', a search bar, and a user profile 'Pedro Angel Miranda's Acc...'. Below the navigation bar is a large banner with the text '¡Bienvenido, Pedro!' and a decorative graphic of data cubes and a magnifying glass. The main content area is divided into several sections:

- Aprender mediante ejemplo:** Recorra los pasos de resolución de un problema empresarial específico en un proyecto de ejemplo.
- Trabajar con datos:** Cree un proyecto para que su equipo prepare datos, busque conocimientos o cree modelos.
- Amplíe sus prestaciones:** Añada herramientas, bases de datos u otras características creando instancias de servicios.
- Quick start:** A list of quick start options:
 - Build dashboards with IBM Cognos Dashboard Embedded
 - Create data pipelines with DataStage
 - Build customer profiles with IBM Match 360 with Watson
 - Catalog and govern data with Watson Knowledge Catalog
 - Build and manage ML models with Watson Studio
 - Query data anywhere with Watson Query
- Visión general:** A summary section with three tables:
 - Proyectos recientes:**

| | |
|----------------------|-------------------|
| Data Fabric Outcomes | 07 abr 2022 21:51 |
| peter | 07 abr 2022 11:05 |
| IA Taller | 30 mar 2022 16:54 |
| Jobs_Puebas | 17 mar 2022 13:24 |
| Data Virtualization | 09 mar 2022 12:42 |
| Ver todos (11) | |
 - Catálogos recientes:**

| | |
|-------------------------|--------------------|
| CreditDataCatalog | 22 feb 2022 14:46 |
| Platform assets catalog | 30 sept 2021 13:48 |
 - Notificaciones:**
 - Project import complete: Data Fabric Outcomes was imported successfully. View import summary. 07 abr 2022 11:07
 - Data Steward removed BusinessUser CPD from project Data Fabric Outcomes Pedro 07 abr 2022 10:48
 - Data Steward added BusinessUser CPD to project Data Fabric Outcomes 07 abr 2022 10:06
 - Project import complete: Data Fabric Outcomes was imported successfully. View import summary. 06 abr 2022 18:07

Nota: Vaya al menú de “Hamburguesa” (esquina superior izquierda), seleccione “Servicios” y luego “Instancias de servicio” (color naranja)

The screenshot shows the IBM Cloud Pak for Data dashboard. On the left, there's a sidebar with various menu items like 'Inicio', 'Conexiones de plataforma', 'Proyectos', 'Ver todos los proyectos', 'Trabajo', 'Catálogo', 'Biblioteca', 'Bibliotecas', and 'Servicios'. The 'Servicios' item is highlighted with an orange box. Below it, under 'Catalogo de servicios', the 'Instancias de servicio' option is also highlighted with an orange box. The main content area displays a 'Visión general' (General View) with sections for 'Proyectos' (Projects), 'Catálogo' (Catalog), and 'Notificaciones' (Notifications). The 'Notificaciones' section lists several deployment and import status messages.

A continuación, tome nota de los servicios que fueron habilitados de manera automática, y siga al siguiente paso. En caso de tener ya habilitado el servicio que se le pide, avance al siguiente servicio.

This screenshot shows the 'Instancias de servicio' (Services Instances) page. At the top, there's a message about updating a service plan. Below that, there are filters for 'Filter by:' (Grupos de recursos, Ubicaciones, Ninguna), 'Producto' (Cloud Object Storage), and 'Plan de servicio' (Lite). A search bar says 'Buscar instancias de servicio'. A blue button labeled 'Agregar servicio +' is visible. The main table lists two services: 'CloudObjectStorage' and 'Watson Studio - lab', both in the 'Default' group, Global location, and Lite plan, with an 'Activas' (Active) status. There are three dots icons next to each row.

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” (color rojo)

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, Amplie sus prestaciones, and a 'Visión general' section showing recent projects, catalogs, and notifications. A search bar at the top right is visible.

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).

The screenshot shows the service catalog search results. A search term 'Cloud object storage' has been entered into the search bar. The results are displayed under the 'Almacenamiento' category. One result, 'Cloud Object Storage', is highlighted with a green box. This item is described as 'Almacenamiento de datos habilitado para desarrolladores para aplicaciones en la nube integradas con servicios d...'. It is marked as 'Lite • Gratuito'.

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).

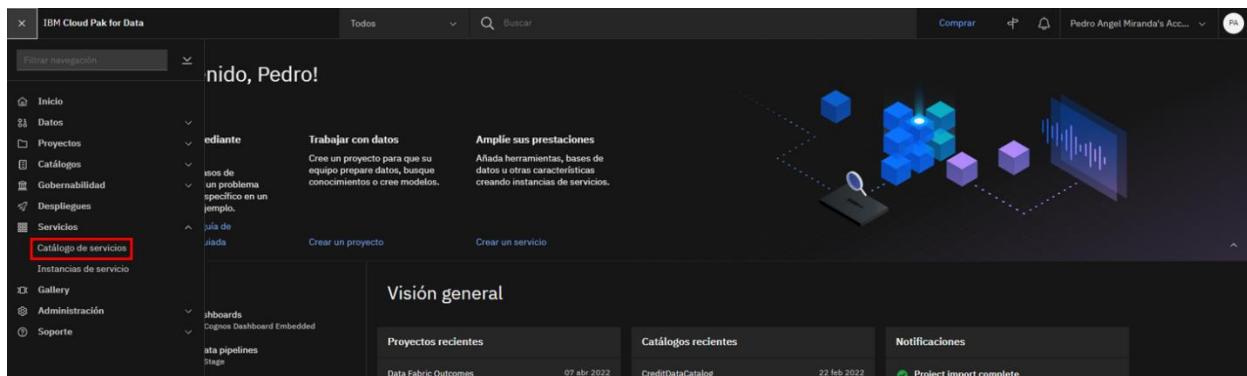
The screenshot shows the IBM Cloud Catalog interface. At the top, there's a search bar and navigation links for Catalogo, Documentos, Soporte, Gestoriar, and a user account. The main area displays a service card for "Cloud Object Storage". The "Resumen" section on the right shows the service is "Gratis" (Free). Below the card, there's a section titled "Elegir una infraestructura" with two options: "IBM Cloud" (selected) and "Satellite". The "IBM Cloud" section has a detailed description and a small icon. The "Satellite" section also has a description and an icon. Underneath, there's a section titled "Seleccionar un plan de precios" with two plans: "Lite" and "Standard". The "Lite" plan is highlighted with a red box and contains specific usage details like "1 COS Service Instance", "Storage up to 25 GiB/month", and "Up to 10 GiB/month A (PUT, COPY, POST, and LIST) requests per month". The "Standard" plan is described as having "No minimum fee, so you pay only for what you use.". To the right of the plan descriptions, there's a "Precios" (Prices) section showing "Gratis" (Free). At the bottom right of the page, there are buttons for "Crear" (Create), "Añadir a estimación" (Add to estimate), and "Ver precios de almacenamiento de vista" (View storage price list).

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

The screenshot shows the "Instancias de servicio" (Service Instances) page in the IBM Cloud Pak for Data interface. At the top, there's a header with a search bar and a "Comprar" (Buy) button. Below the header, there's a message about updating the service plan. The main table lists one instance: "CloudObjectStorage" (Status: Activas, Plan: Lite). The table has columns for Nombre, Grupo, Ubicación, Producto, Plan, and Estado.

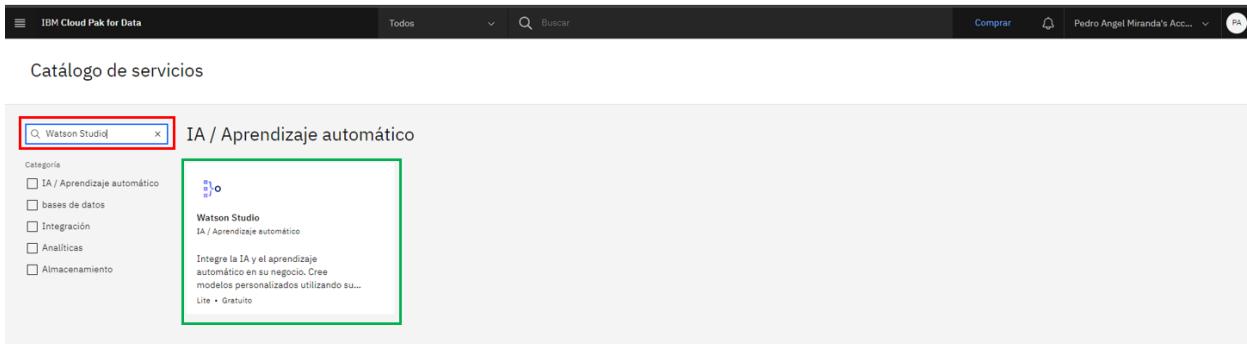
| Nombre | Grupo | Ubicación | Producto | Plan | Estado |
|--------------------|---------|-----------|----------------------|------|---------|
| CloudObjectStorage | Default | Global | Cloud Object Storage | Lite | Activas |

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 is a sidebar menu with various options like Inicio, Datos, Proyectos, Catálogos, Gobernabilidad, Despliegues, Services, and Catalogo de servicios. The 'Catalogo de servicios' option is highlighted with a red box. The main content area displays a dashboard titled 'nido, Pedro!' with sections for 'Trabajar con datos' and 'Amplie sus prestaciones'. It also features a 3D visualization of data cubes and a bar chart. At the bottom, there's a 'Visión general' section with tabs for 'Proyectos recientes', 'Catálogos recientes', and 'Notificaciones'.

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' search results. A search bar at the top contains the query 'Q. Watson Studio'. Below it, a list of categories is shown: IA / Aprendizaje automático, bases de datos, Integración, Análisis, and Almacenamiento. Under 'IA / Aprendizaje automático', there is a single result for 'Watson Studio'. This result is highlighted with a green box. The card for 'Watson Studio' includes the text 'Integre la IA y el aprendizaje automático en su negocio. Cree modelos personalizados utilizando su...', a 'Lite' button, and a 'Gratis' button.

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).

The screenshot shows the IBM Cloud Pak for Data service catalog. A red box highlights the dropdown menu where 'Dallas' is selected. Another red box highlights the 'Lite' plan row in the price table. A green box highlights the 'Crear' (Create) button at the bottom right.

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 | 1 usuario autorizado Límite mensual de 50 unidades de capacidad por hora Entorno = Número de unidades de capacidad necesarias por hora - 1 vCPU + 4 GB RAM = 0,5 - 2 vCPU + 8 GB RAM = 1 - 4 vCPU + 16 GB RAM = 2 - Optimización de decisiones = Entorno + 5 | Gratis |
| Professional | Unlimited collaborators Unlimited elastic compute environments Entorno = # of capacity units required per hour - 1 vCPU + 4 GB RAM = 0,5 - 2 vCPU + 8 GB RAM = 1 | \$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:

The screenshot shows the IBM Cloud service catalog under the 'Instancias de servicio' section. It lists two instances: 'CloudObjectStorage' and 'Watson Studio - lab'. Both are in the 'Activas' (Active) state. A blue box highlights the 'Add service' (+) button at the top right of the table.

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 |
| 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”

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).

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).

IBM Cloud Pak for Data

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 período 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 |

Crear Ver términos

2.l) Si el servicio de Cognos Dashboard Embedded 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 |
|----------------------------------|---------|-----------|-------------------------------|------|---------|
| 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 interface. On the left, there's 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 content area features a welcome message '¡Buenos días, Pedro!', sections for 'Trabajar con datos' and 'Amplíe sus prestaciones', and a 'Visión general' section with tabs for 'Proyectos recientes', 'Catálogos recientes', and 'Notificaciones'. A large graphic on the right depicts data cubes and a screen.

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 'Catálogo de servicios' page. In the search bar at the top, the text 'DataStage' is entered. Below the search bar, there are two results listed under 'Integración': 'DataStage' (Analíticas • bases de datos • Integración) and another 'DataStage' entry. The first result is highlighted with a green box, indicating it is the one being selected.

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).

The screenshot shows the IBM Cloud Pak for Data service creation interface for DataStage. The 'Region' dropdown is set to 'Dallas'. The 'Plan' section shows the 'Lite' plan highlighted with a red box. The 'Crear' (Create) button is highlighted with a green box.

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

| 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 |
| Watson Studio - lab | Default | Dallas | Watson Studio | Lite | Activas |

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'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 3D graphic of data cubes and a screen. It features sections for 'Trabajar con datos' (Create a project for your team to prepare data, search for knowledge or create models), 'Amplie sus prestaciones' (Add tools, bases of data and other characteristics to create service instances), and buttons for 'Crear un proyecto' and 'Crear un servicio'. Below these is a 'Visión general' section with tabs for 'Proyectos recientes', 'Catálogos recientes', and 'Notificaciones', each showing recent activity.

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 'Catalogo de servicios' page with a search bar containing 'Machine Learning'. The results are displayed in a grid. One item, 'Machine Learning' under the 'IA / Aprendizaje automático' category, is highlighted with a green box. This item has a small blue icon with a white 'M' and some text below it. To the left of the search bar, there's a sidebar with categories: 'IA / Aprendizaje automático', 'Bases de datos', 'Integración', 'Analíticas', and 'Almacenamiento'. The 'IA / Aprendizaje automático' category is also highlighted with a red box.

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 creation page. A red box highlights the 'Region' dropdown set to 'Dallas'. Another red box highlights the 'Plan' section where 'Lite' is selected. A green box highlights the 'Crear' (Create) button at the bottom right.

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 services: CloudObjectStorage, DataStage-34, IBM Cognos Dashboard Embedded-db, WatsonMachineLearning, and Watson Studio - lab. All services are listed under the 'Activas' (Active) status column. A blue '+' button is visible at the top right of the table.

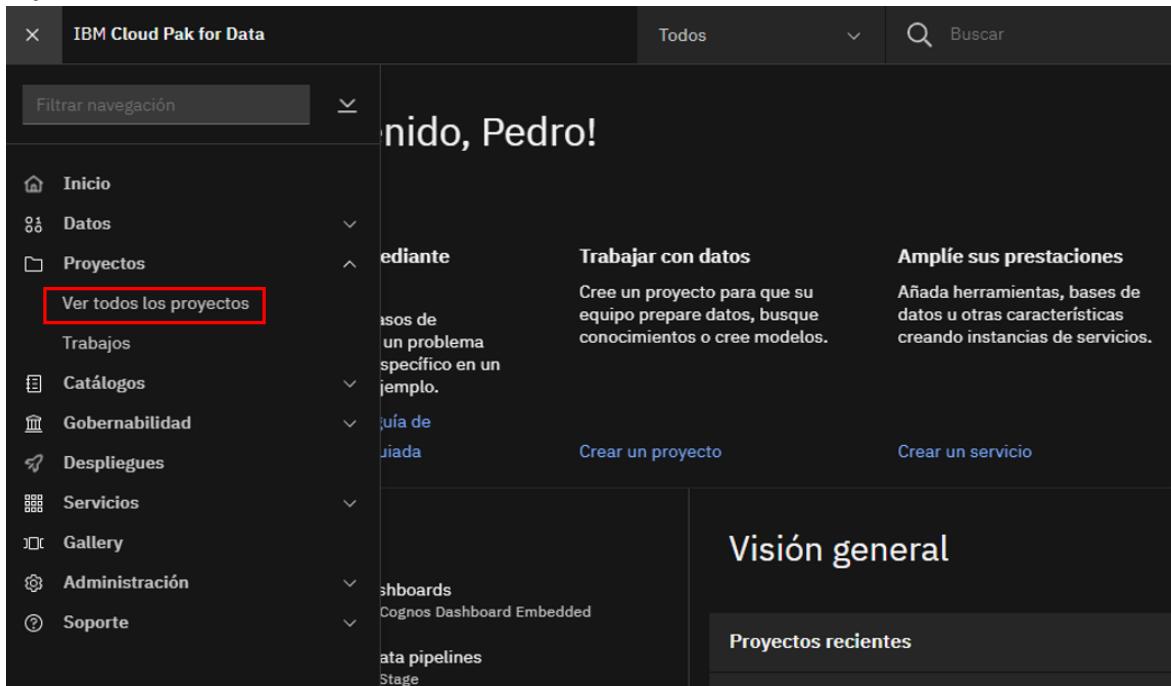
Asegúrese de haber habilitado todos los servicios mencionados:

- Cloud Object Storage
- Watson Studio
- IBM Cognos Dashboard Embedded
- DataStage
- Watson Machine Learning

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

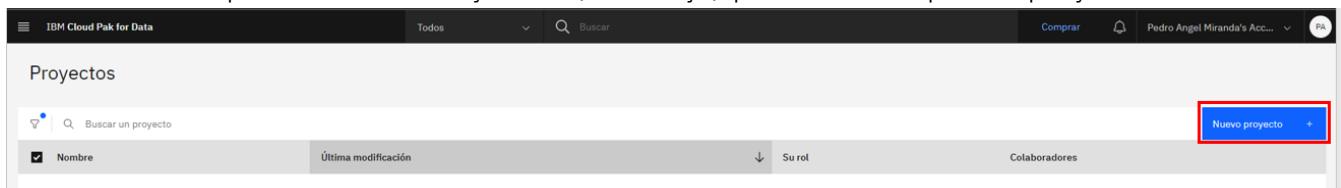
3. Crear proyecto

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. On the left, there's a navigation sidebar with various sections like Inicio, Datos, Proyectos, Catálogos, Gobernabilidad, Despliegues, Servicios, Gallery, Administración, and Soporte. Under 'Proyectos', the 'Ver todos los proyectos' button is highlighted with a red box. The main content area has sections for 'Trabajar con datos' and 'Amplíe sus prestaciones'. A large 'Visión general' section is visible, along with a 'Proyectos recientes' panel.

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. At the top right, there's a 'Nuevo proyecto' button with a red box around it. Below it, there's a search bar and some filter options. The main area displays a table with columns for Nombre, Última modificación, Su rol, and Colaboradores.

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)

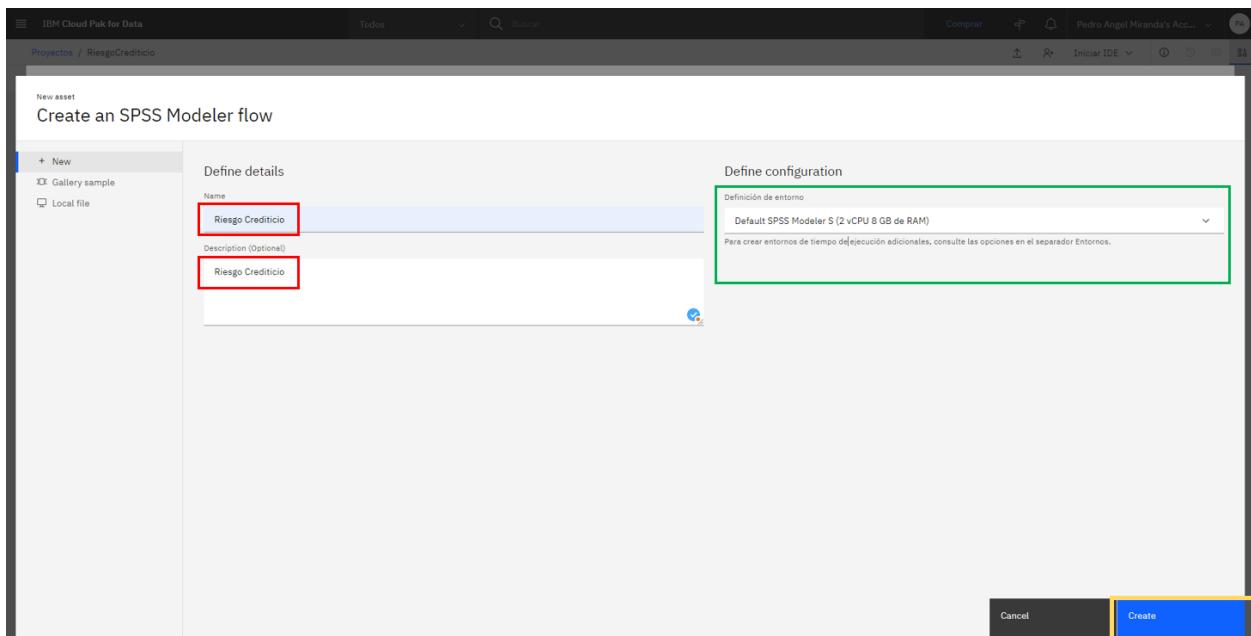
“Felicitaciones ya creo su proyecto y descargó las dos fuentes de datos que va a utilizar para su modelo”.

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)

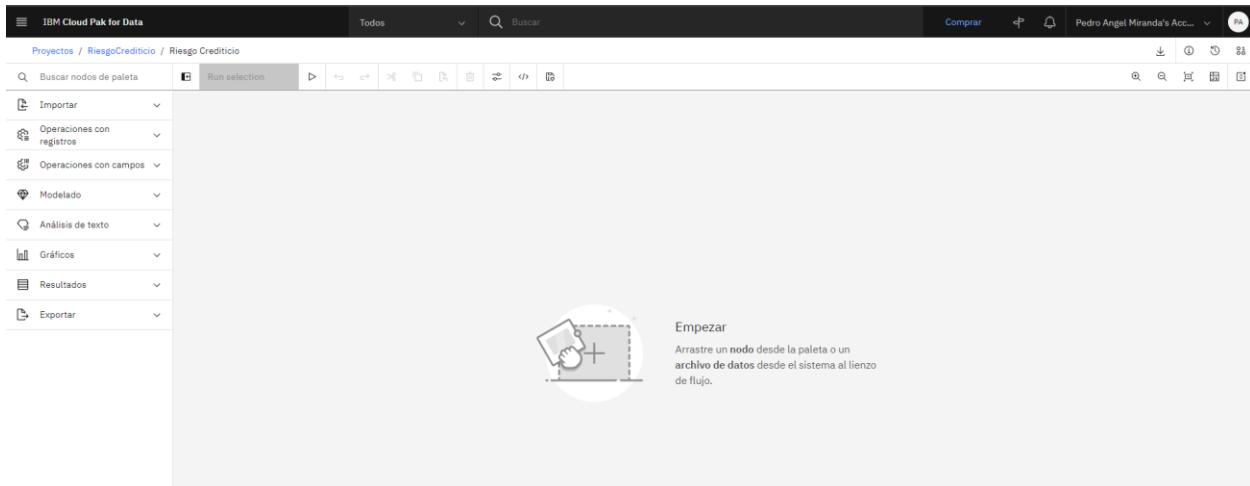
y a continuación la opción “SPSS Modeler” (color verde).

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.

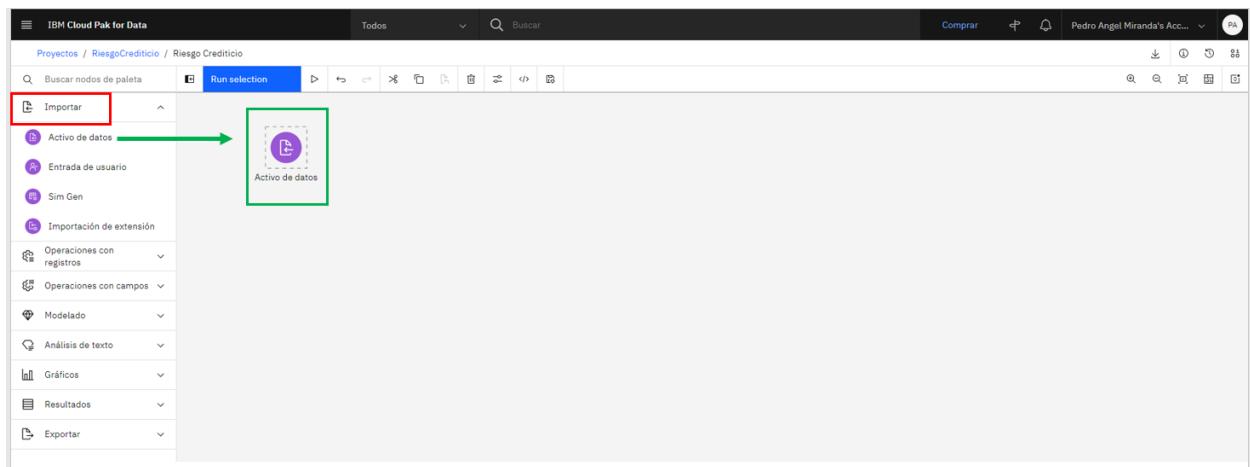


Nota: Para fines prácticos del Hackathon siempre utilizar Default SPSS Modeler (2vCPU 8GB RAM)

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

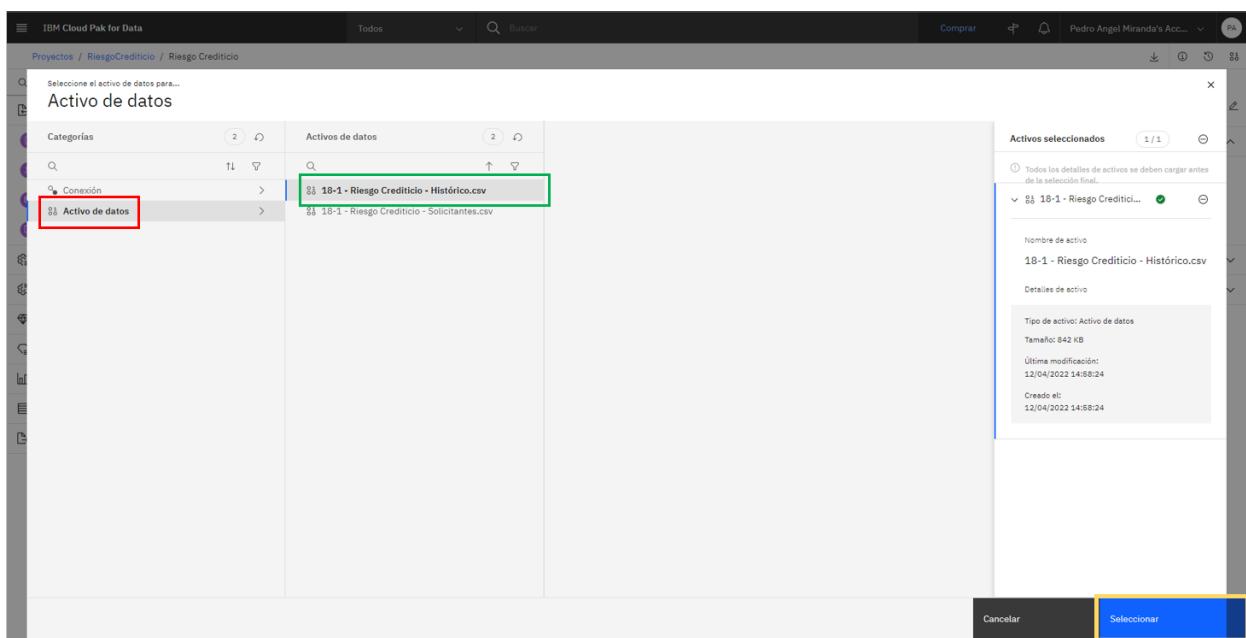


4.c) A continuación vamos a **crear el flujo de entrenamiento del modelo**, para eso debe ir a la opción de “Importar” (color rojo) y arrastrar hacia el lienzo el nodo “Activo de Datos” (flecha verde) y quedará como sigue:



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.

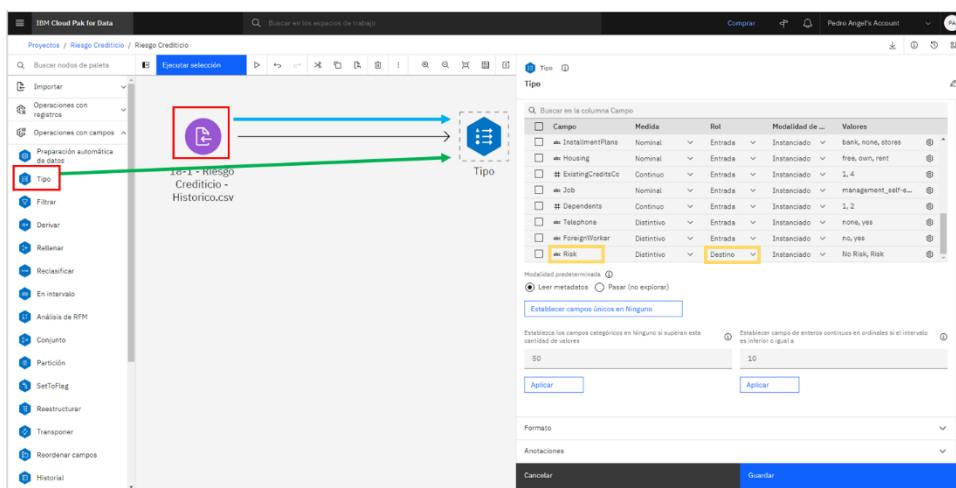


4.e) Si hizo todo bien regresará al lienzo donde tendrá el ícono de “Activo de datos”, pero habrá sido renombrado con el nombre de la fuente de datos que le acaba de asociar (color rojo).

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).

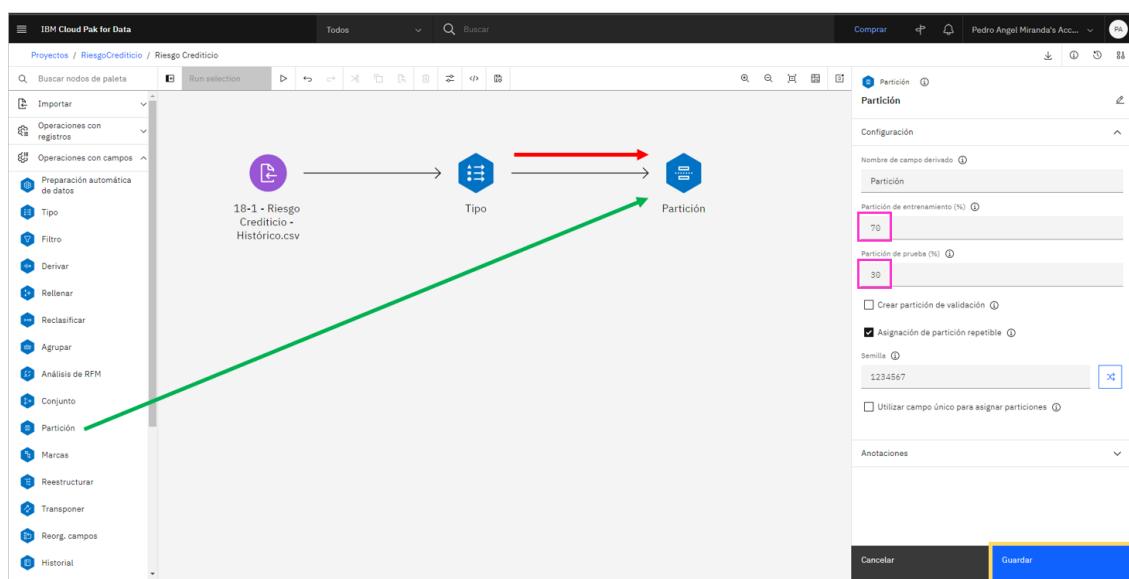
Ahora abra el ícono de “Tipo” (botón derecho y opción “Abrir”), aparecerá una nueva pantalla a un lado donde estarán todos los campos de la fuente de datos, busque el que se llama “Risk” (color amarillo, 1er columna) cambie el “rol” de ese campo a “Destino” (3^a columna, color amarillo).

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



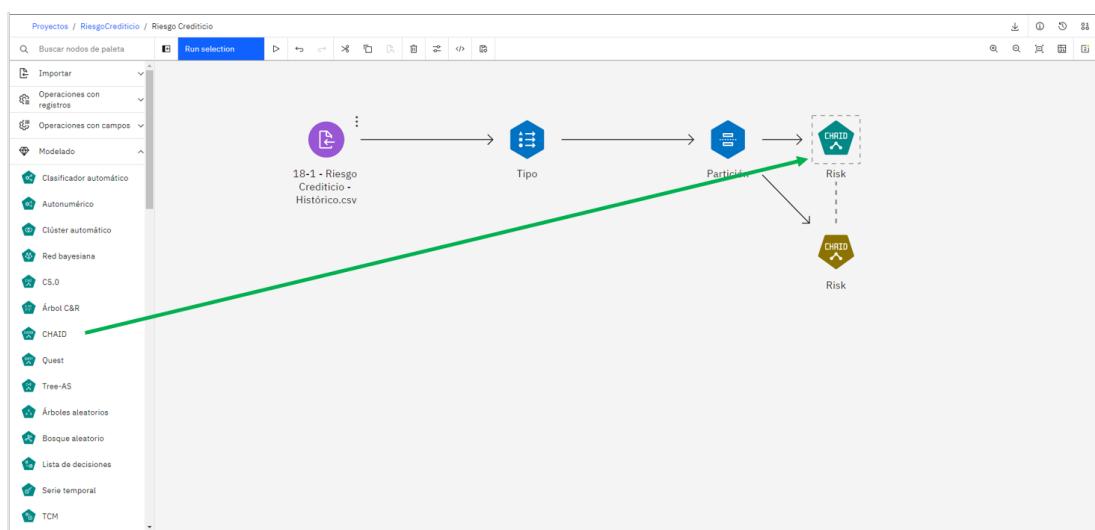
4.f) A continuación en “Operaciones con Campos” seleccione el nodo “Partición” y arrástrelo al lienzo (flecha verde), conecte el nodo “Tipo” al nodo “Partición” (flecha roja).

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).



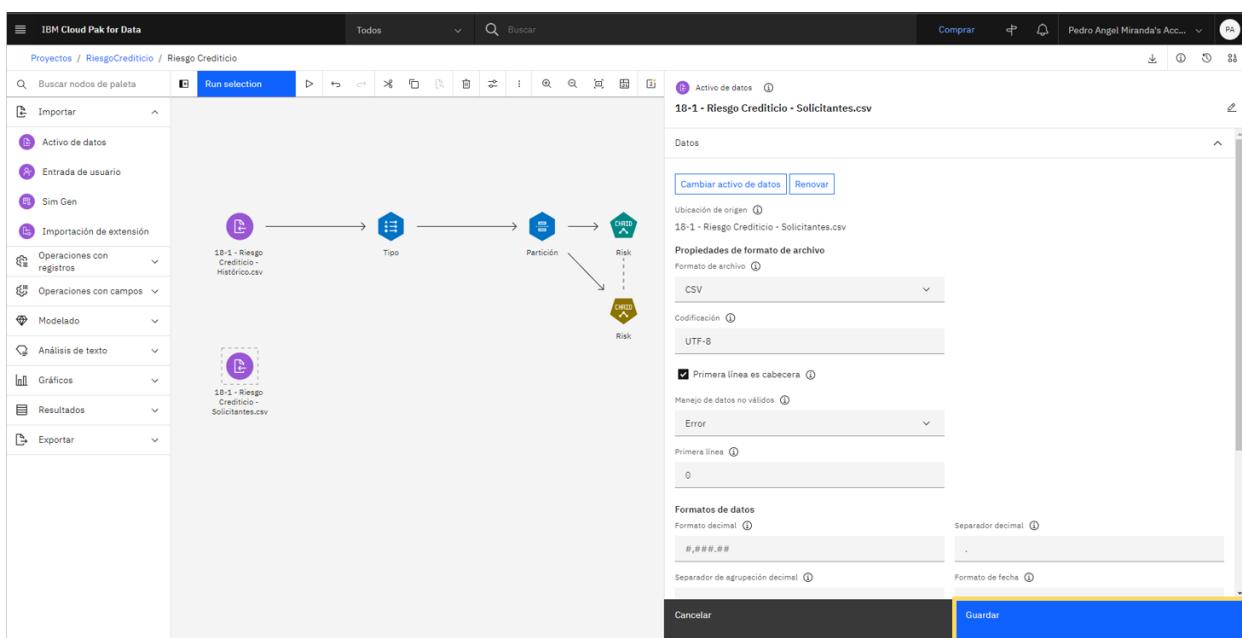
4.g) Ahora vaya a las opciones de la derecha y busque la sección de “Modelado” y arrastrar el ícono “CHAID” al lienzo (flecha Verde), conectar el nodo “Partición” con el nodo “CHAID”

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]).

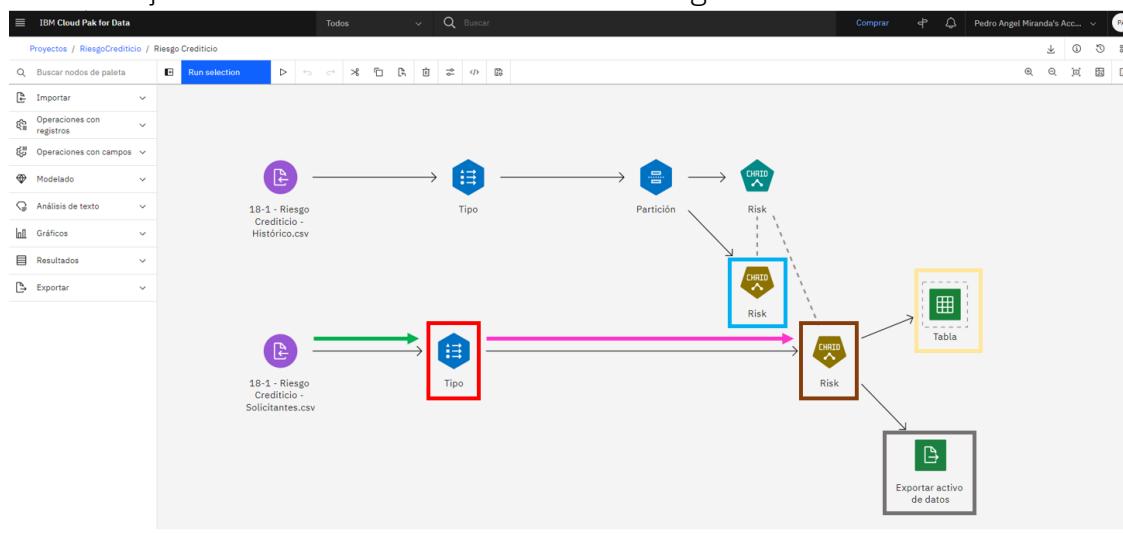
Con éste nuevo ícono de “Activo de Datos” estamos iniciando la creación del flujo de predicción.



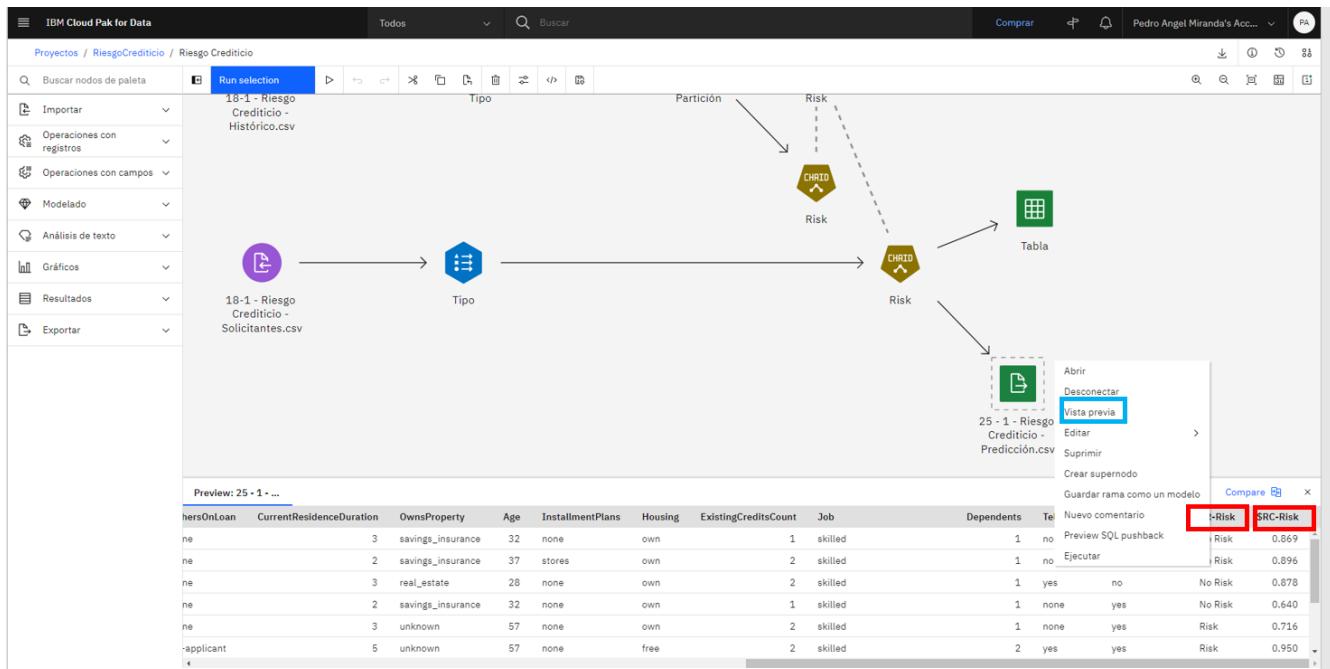
4.h) Ahora vamos a **crear el flujo completo de predicción** como sigue:

- 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 “25 - 1 - Riesgo Crediticio - Predicción.csv”, 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:



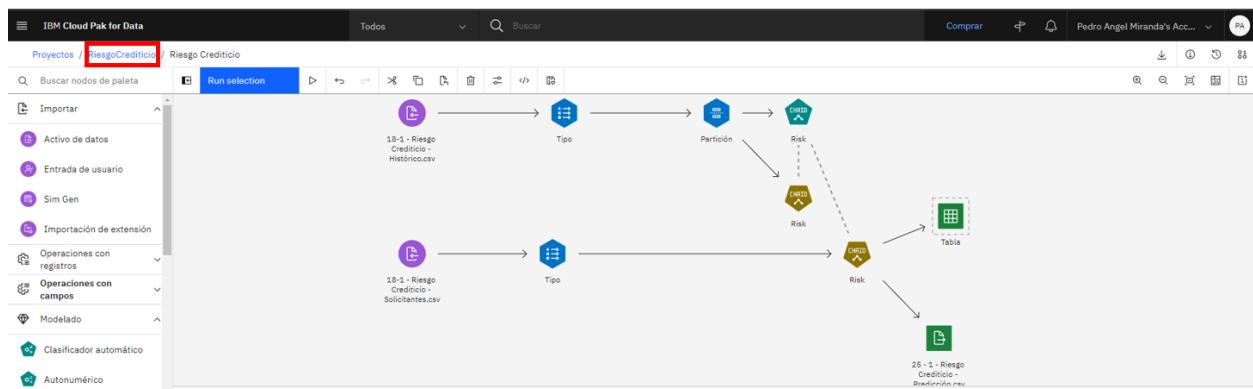
4.i) Obtenga una vista previa del nodo (color azul) “Tabla” (color amarillo, punto anterior) y del nodo “Exportar Activo de Datos” (color gris, punto anterior) y confirme que en ambos aparezcan los campos: “\$R-Risk” y “\$RC-Risk” (color rojo) que se generan como resultado de correr el modelo de predicción.



¡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)

A screenshot of the IBM Cloud Pak for Data interface. At the top, there's a blue header bar with the text "1 item selected". Below it is a table with columns: Nombre, Tipo, Plan, Ubicación, Estado, and Grupo. There are five rows in the table. The first four rows have an empty checkbox column. The fifth row has a checked checkbox, and its "Estado" column shows a blue diamond icon with the text "No asociado". In the top right corner of the table area, there's a red box highlighting a blue button labeled "Asociar servicio" with a small info icon, followed by "Cancel".

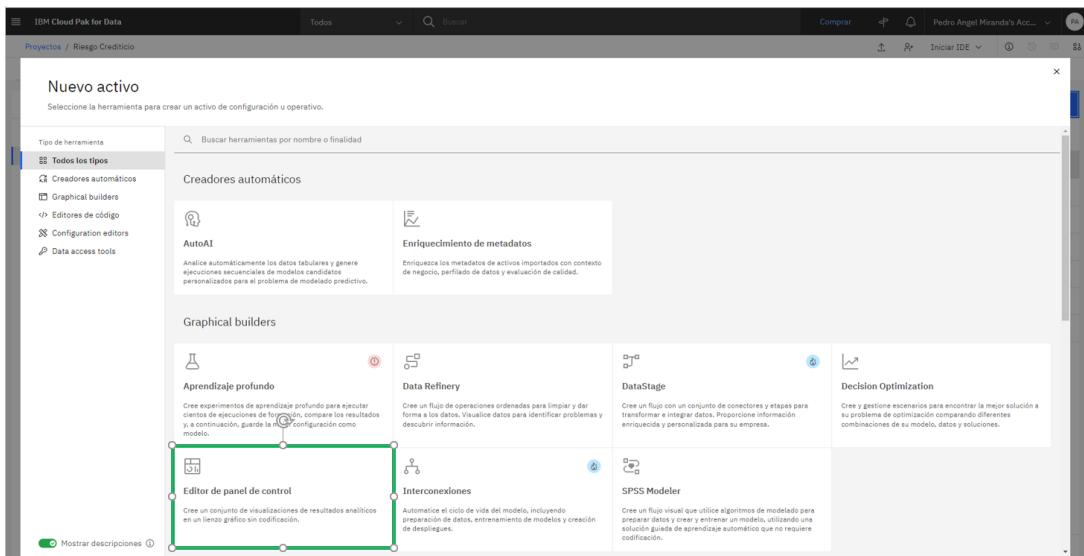
Si todo sale bien vera una paloma verde en la parte de “Estado” y si servicio estara asociado al proyecto

A screenshot of a confirmation dialog titled "Asociar servicio". It contains a green message box with the text "El servicio IBM Cognos Dashboard Embedded-db se ha asociado correctamente al proyecto RiesgoCrediticio.". Below the dialog is a table with the same structure as the previous one. The fifth row now has a checked checkbox and its "Estado" column shows a green circle icon with the text "Asociado". A red box highlights this row. The "Nuevo servicio" button is visible at the top right of the table area.

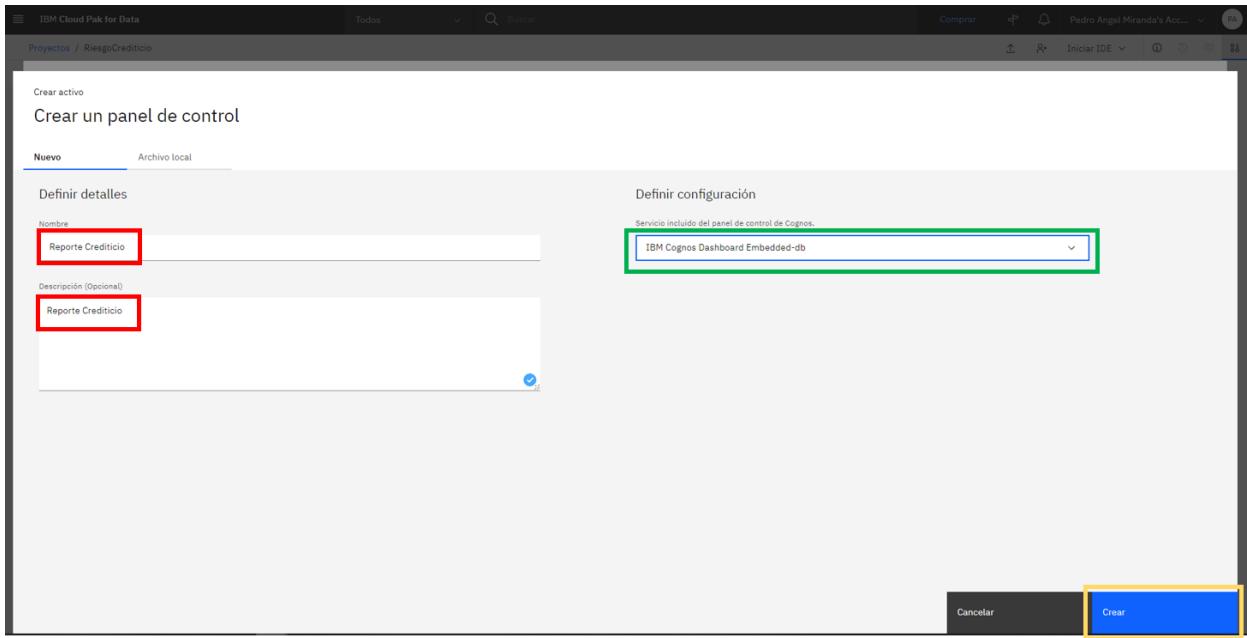
5.a) Regrese a la parte de “Activos” dé clic al botón azul “Nuevo activo” (color rojo)

A screenshot of the "Activos" (Assets) section in IBM Cloud Pak for Data. At the top, there's a navigation bar with "Proyectos / RiesgoCrediticio". Below it is a table with columns: Nombre and Última modificación. There are two rows in the table. The first row has a file icon and the name "1B-1 - Riesgo Crediticio - Solicitantes.csv". The second row has a file icon and the name "1B-1 - Riesgo Crediticio - Histórico.csv". To the right of the table, there's a blue button labeled "Nuevo activo" with a plus sign, which is highlighted with a red box. A sidebar on the right is titled "Datos de este proyecto" and contains a dashed box for "Suelte aquí archivos de datos o examine en busca de datos que cargar".

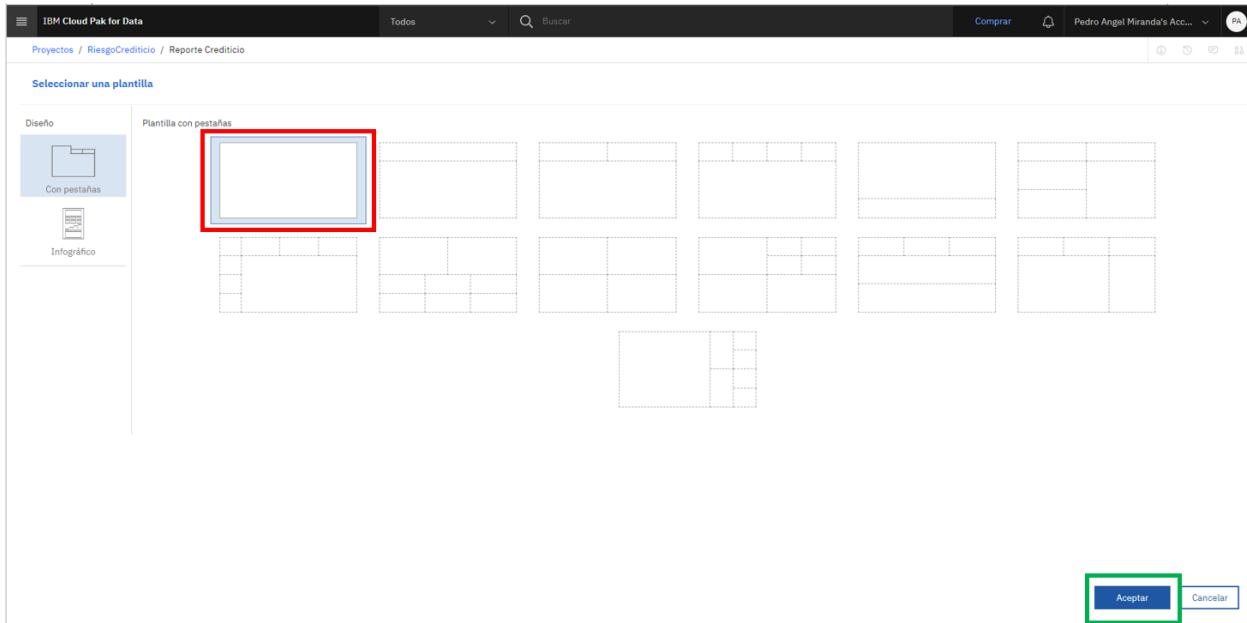
y seleccione “Editor Panel de Control” (color verde)



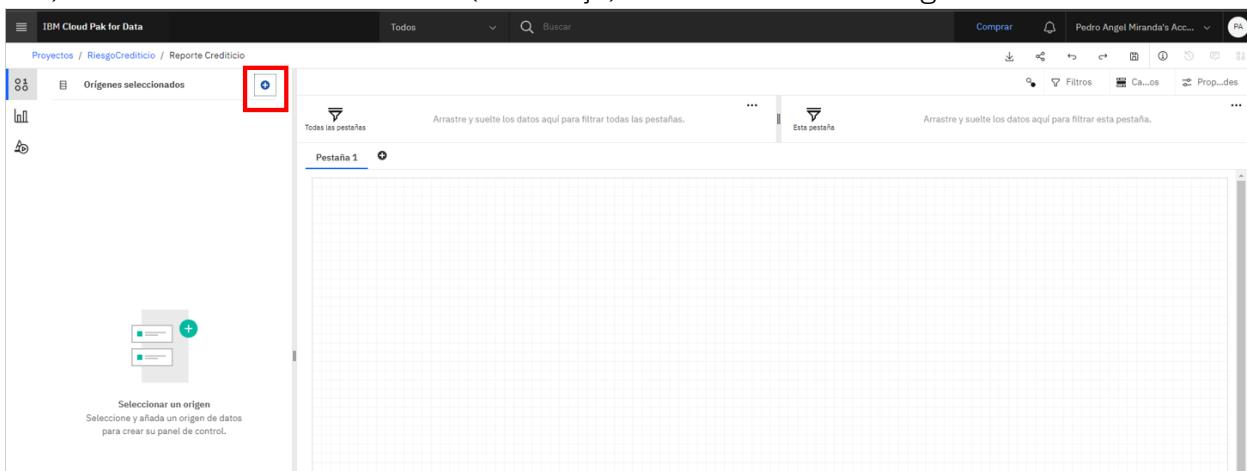
5.b) Ponga en Nombre y Descripción “Reporte Crediticio” (color rojo), confirme que ya tiene una instancia de Cognos asociada al proyecto (color verde) y de clic en “Crear” (color amarillo).



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)

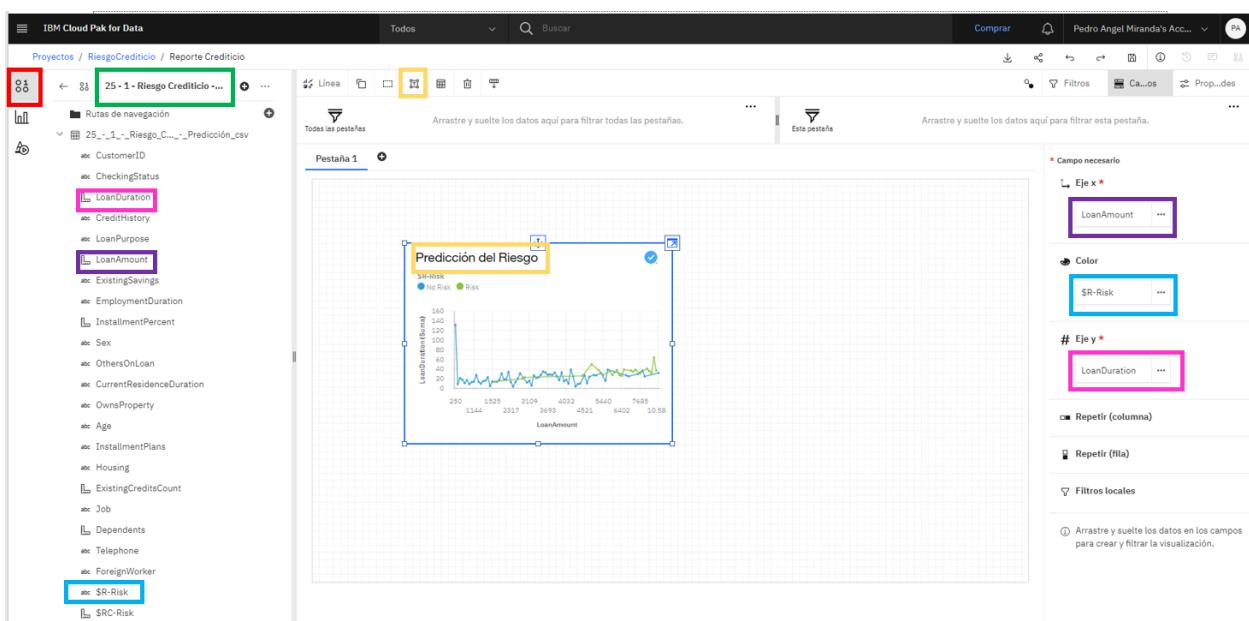
Seleccionar origen de datos

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:

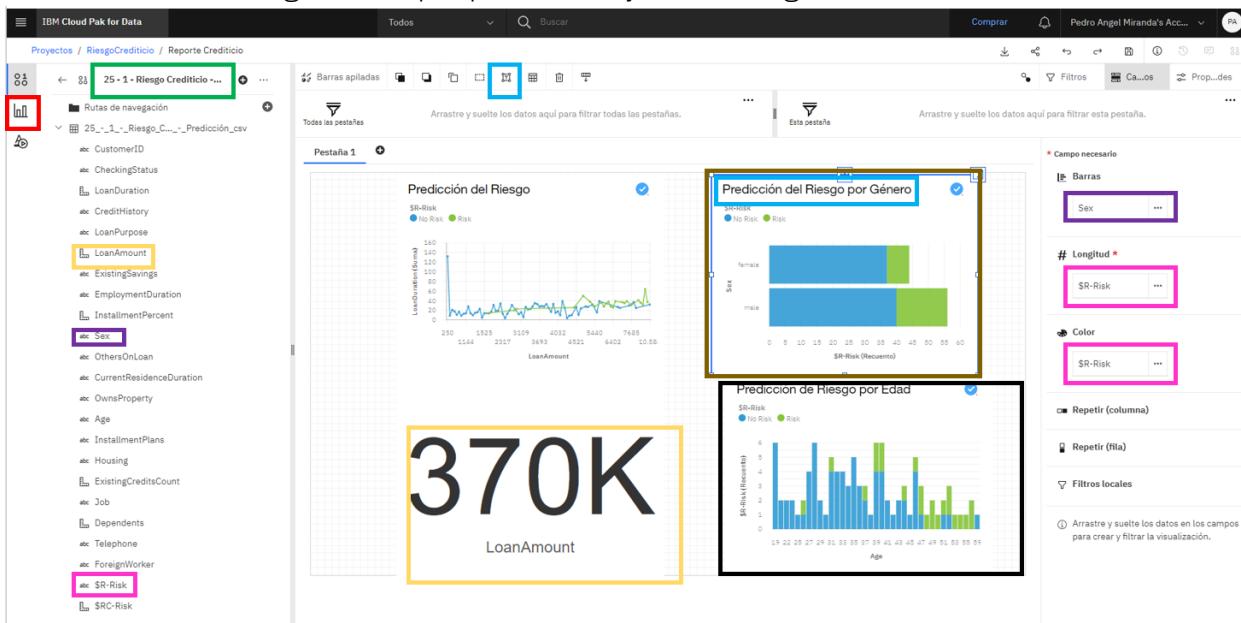
- arrastre el campo “Sex” que está en el lado izquierdo hacia “Barras” en las opciones de la gráfica del lado derecho (color morado)
- arrastre el campo “\$R-Risk” a “Longitud” y “Color” (color rosa)
- finalmente póngale el título “Predicción del Riesgo por Género” (color azul)

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

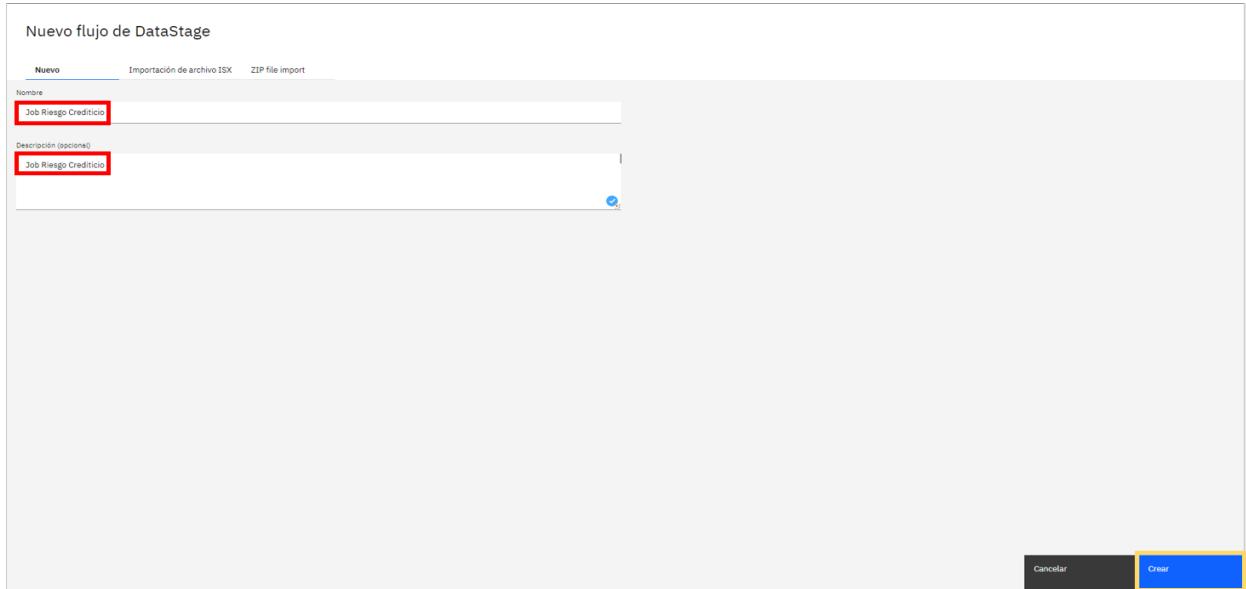
6. Creación de Flujo de Datastage

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

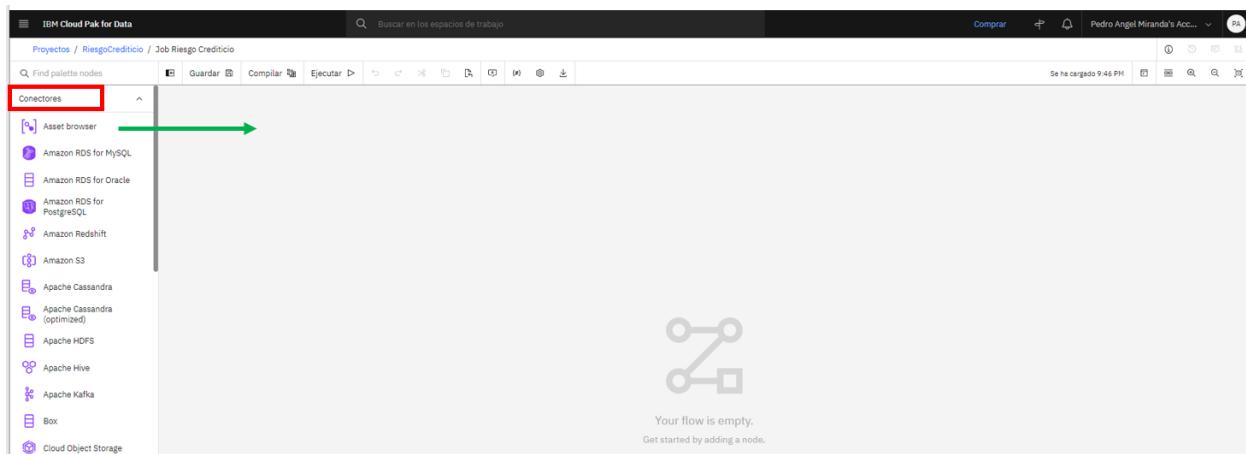
y seleccione “DataStage” (color verde).

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

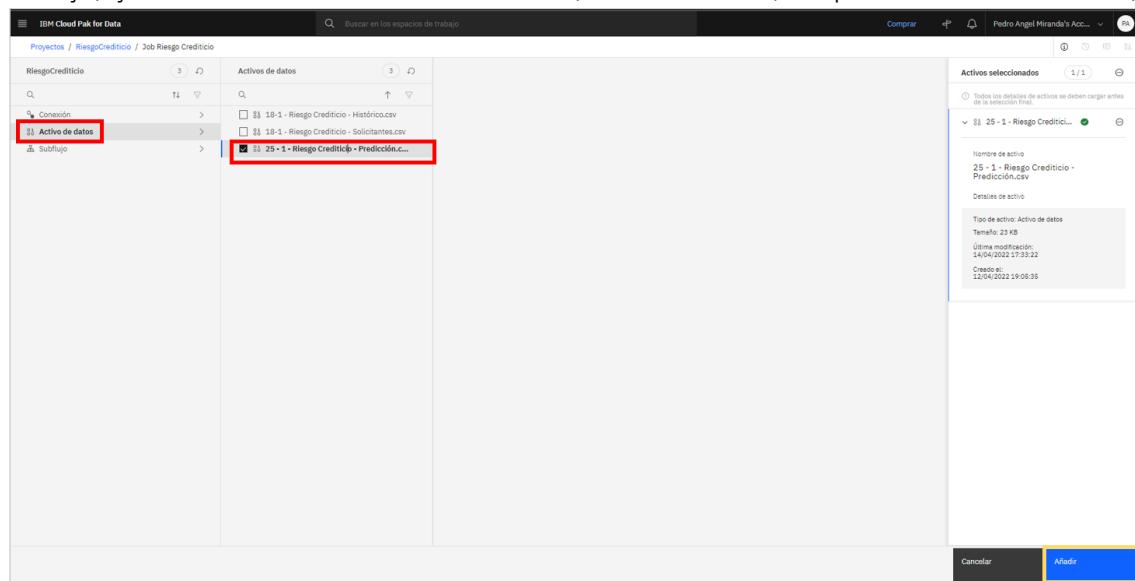
Nombre su flujo de DataStage y la descripción como “Job Riesgo Crediticio” (color rojo) y oprima la opción de “Crear” (color amarillo).



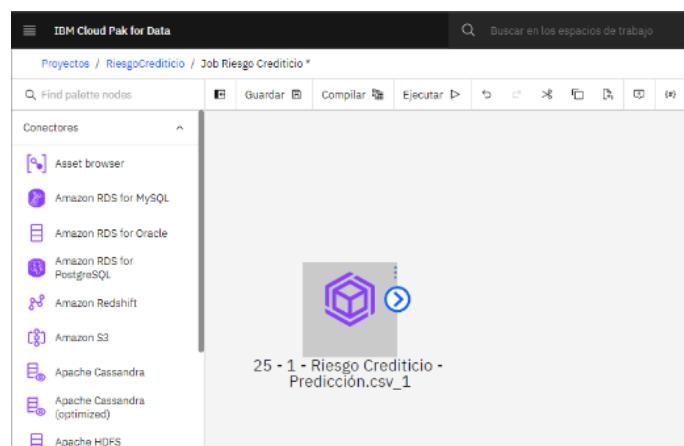
6.b) A continuación primero vamos a **crear los componentes (Stages) de nuestro flujo de DataStage**, para eso debe ir a la opción de “Conectores” (color rojo) y arrastrar hacia el lienzo el nodo “Activo de Datos” (flecha verde)



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).



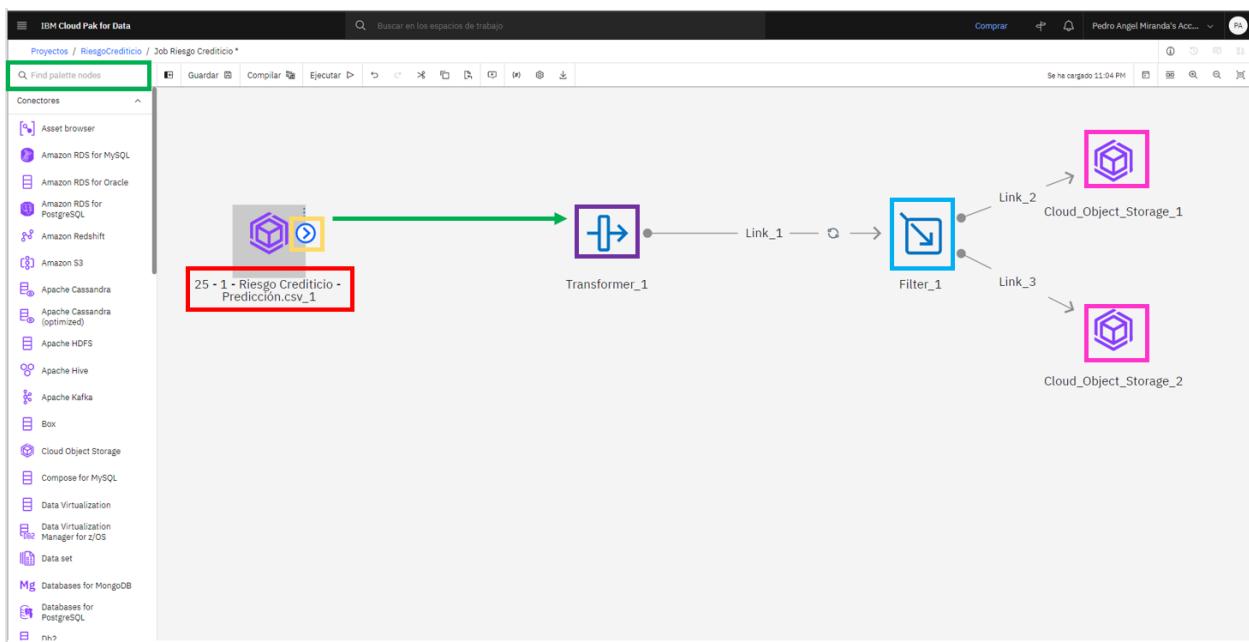
Si los pasos anteriores se realizaron correctamente obtendrá la siguiente pantalla con el nombre del activo de datos.



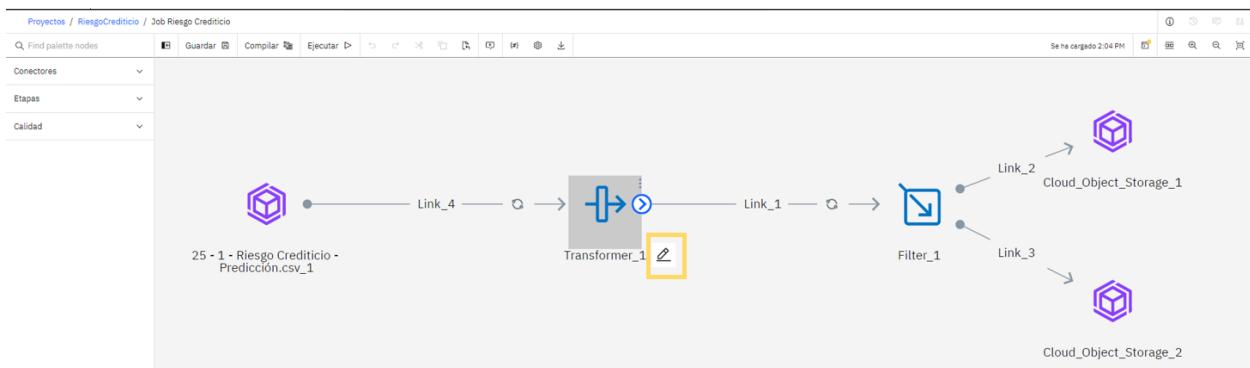
Nota: Si el nombre del activo no aparece, de click derecho el componente y seleccione la opción Vista Previa. Si aparecen datos, seleccionó correctamente su activo. Continue.

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”.

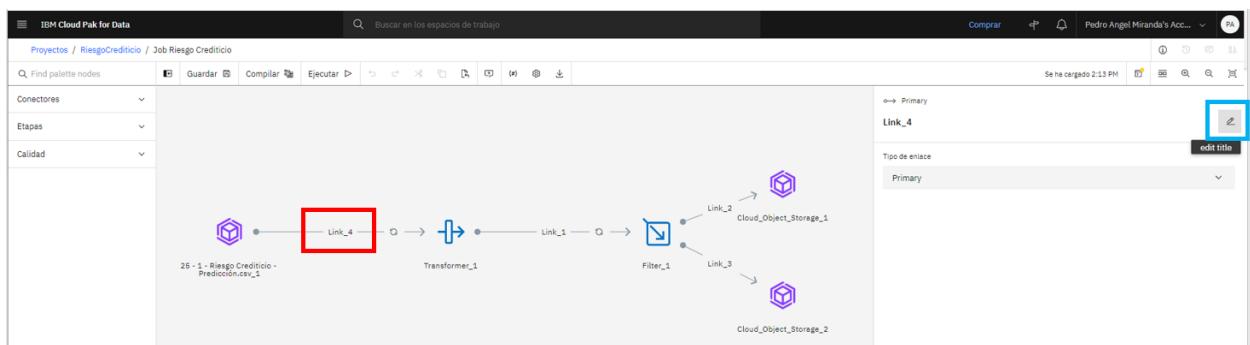
Una vez que tenga todos sus componentes enlácelos como se muestra a continuación. De lado derecho de cada componente tendrá un icono “>” (color amarillo) el cual tendrá que arrastrar hasta el siguiente componente(flecha verde), primero una el Cloud Object Storage “25 - 1 - Riesgo Crediticio - Predicción.csv” (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 “Definidos” y finalmente “Cloud_Object_Storage_2” por “Indefinidos”



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.

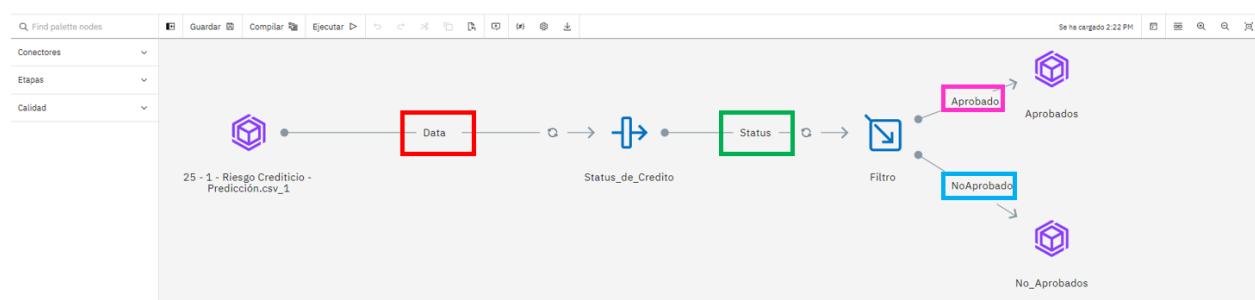
25 - 1 - Riesgo Crediticio - Predicción.csv – Status_de_Credito = Data (color rojo)

Status_de_Credito-Filtro= Status (color verde)

Filtro-Definidos=Definidos (color rosa)

Filtro-Indefinidos=Indefinidos (color azul)

De manera que le quede como se muestre a continuación:

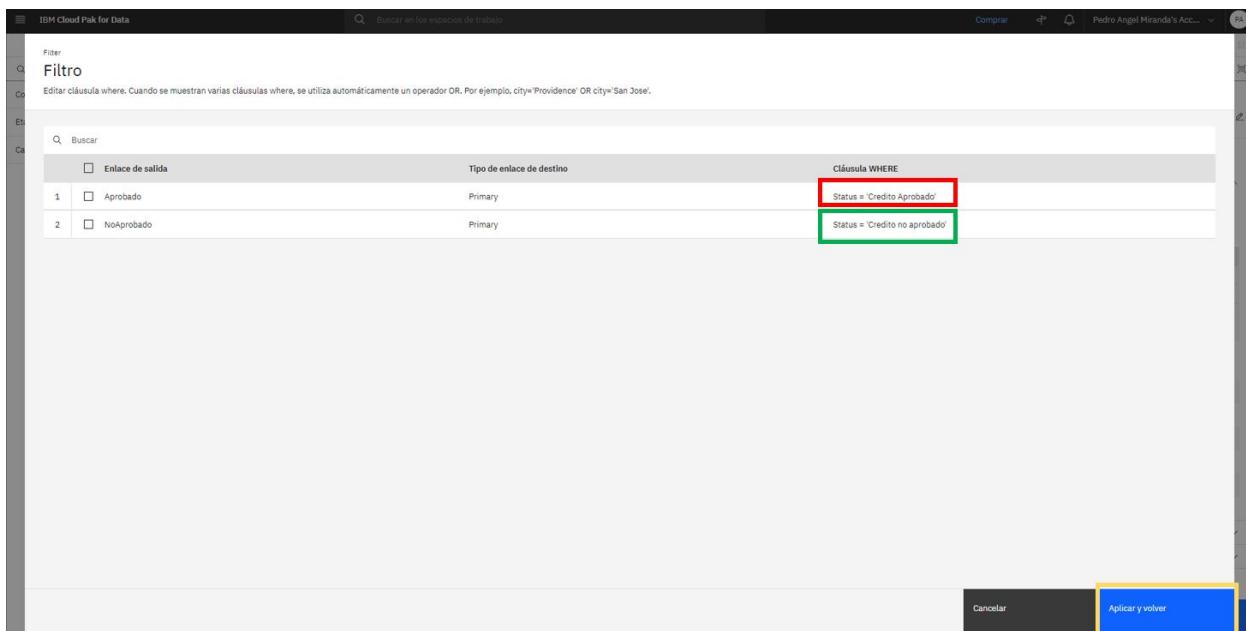


6.e) Ahora vamos a configurar nuestro Transformer llamado “Status_de_Credito”, de doble click sobre el componente, var a la parte de “Salida” (color rojo), de click en “Añadir columna” (color verde), vaya hasta la parte de debajo y vera que tiene una nueva columna la cual vamos a editar con el lápiz 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 Definido' else 'Credito Indefinido'**, 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"/> | |

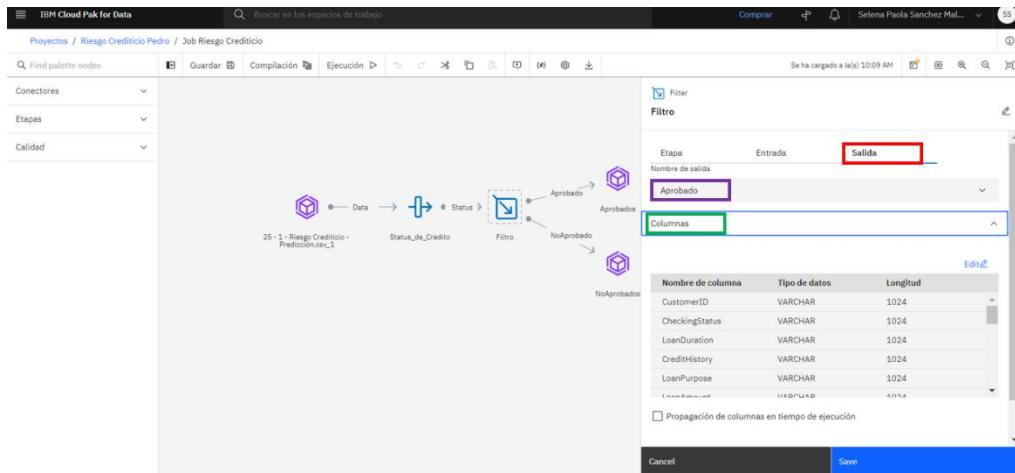
6.f) Vaya al componente Filter llamado “Filtro” de doble clic, expanda la pestaña llamada “Propiedades”, en la parte de “predicados” de clic en “Edit” y en Cláusula WHERE coloque

Para Definidos Status = 'Credito Definido' (color rojo) y para Indefinidos Status = 'Credito Indefinido' (color verde) finalmente de click en Aplicar y volver (color amarillo)



A continuación, de click en la pestaña Salida (color rojo), en el campo nombre de Salida, seleccione Definido (color morado) y despliegue el menú Columnas (color verde).

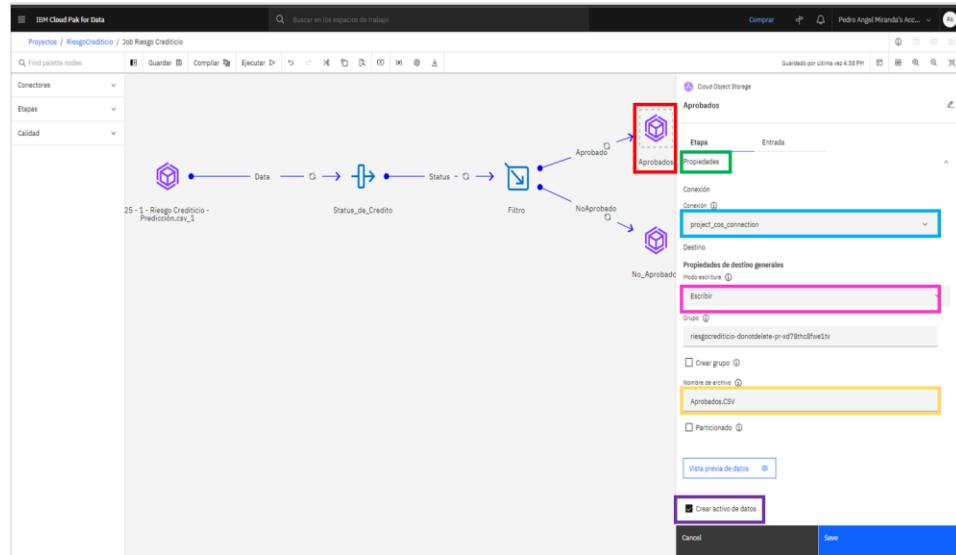
Verifique que puede visualizar las columnas de nuestro activo, como se ve a continuación:



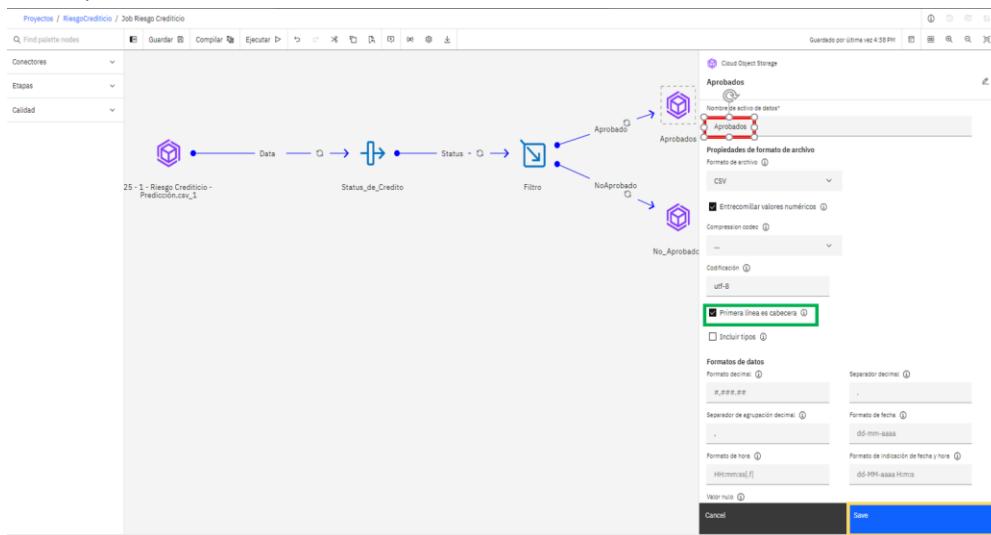
Repita este mismo paso, seleccionando en el campo nombre de salida (color morado) la opción Indefinido. De click en Guardar.

6.g) vaya al componente Cloud Object Storage llamado “Definido” (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 **Definidos.CSV** (color amarillo), palomeamos el cuadro que dice “Crear activos de datos”(color morado).

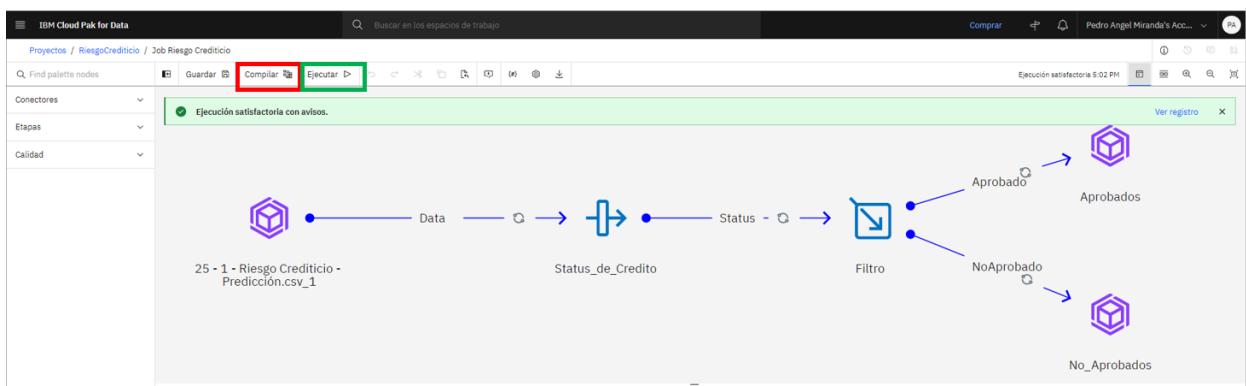


En “Nombre de activo de datos” ponemos **Definidos** (color rojo), palomeamos el cuadro que dice “Primera línea es cabecera” (color verde) y finalmente hacemos click en “Save” (color amarillo)



Repita el paso 6.g) pero ahora con componente Cloud Object Storage llamado “**Indefinidos**”, cambiando solo las partes de “Nombre del archivo” a **Indefinidos.CSV** y “Nombre de activo de datos” a **Indefinidos**.

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 “Definidos” de click izquierdo y seleccionar “Vista previa de los datos”, podrá darse cuenta que tenemos la nueva columna con puros Status “Credito definido”.

| | esidenceDuration | 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 | 0.868852 | 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 | 0.868852 | 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 | 0.988763 | 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)

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

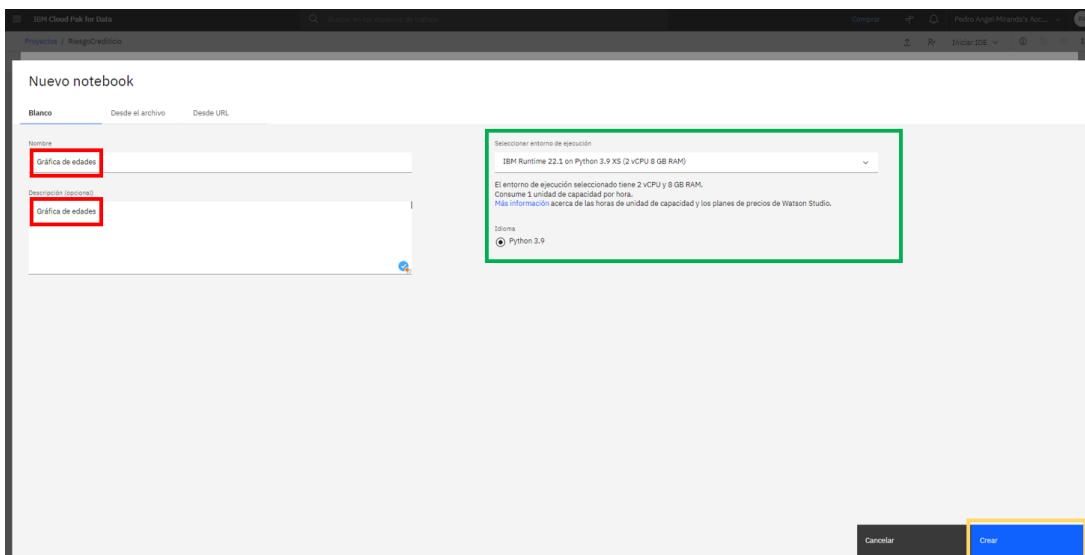
Felicidades ha terminado su flujo de DataStage

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 notebook interface with two code cells and a sidebar titled 'Datos'.

Code Cell 1 (Yellow Box):

```
# Gráfica de Edades
Vamos a visualizar la información de un activo y graficaremos una columna

In [1]: 
import os, types
import pandas as pd
from botocore.client import Config
import ibm_boto3

def __iter__(self): return 0

# @hidden_cell
# 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_e45438407c4721a7e7cebd85c7eb19 = ibm_boto3.client(service_name='s3',
    ibm_api_key_id='1MhJ7zKVDNCKFcGgF5SeuBfUK7C0py0Rgh4IC',
    ibm_service_endpoint='https://nile-1.cloud-object-storage.appdomain.cloud',
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.us.cloud-object-storage.appdomain.cloud')

body = client_e45438407c4721a7e7cebd85c7eb19.get_object(Bucket='riesgoradicidio-donotdelete-pr-xd78thc8feitv',Key='18-1 - Riesgo Credito - Histórico.csv')['Body']
# add missing __iter__ method, as pandas accepts body as file-like object
if not hasattr(body, '__iter__'): body.__iter__ = types.MethodType(__iter__, body)

df_data_1 = pd.read_csv(body)
df_data_1.head()

In [1]: 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()
```

Code Cell 2 (Yellow Box):

The sidebar 'Datos' contains a section for 'Archivos' with a file named '18-1 - Riesgo ... - Histórico.csv'. A pink box highlights the 'pandas DataFrame' option under 'Insertar en código'.

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 within the IBM Cloud Pak for Data environment. The top navigation bar includes 'IBM Cloud Pak for Data', 'Proyectos / RiesgoCredito / Gráfica de edades', 'Buscar en los espacios de trabajo', 'Comprar', 'Pedro Angel Miranda's Account', and a 'PA' icon. The toolbar has standard options like File, Edit, View, Insert, Cell, Kernel, Help, and a 'Run' button highlighted in red. Below the toolbar is a code cell containing Python code for interacting with IBM Cloud Object Storage to download a CSV file and read its first five rows. The output cell, labeled 'Out[18]:', displays the first five rows of the dataset as a pandas DataFrame. The columns include CustomerID, CheckingStatus, LoanDuration, CreditHistory, LoanPurpose, LoanAmount, ExistingSavings, EmploymentDuration, InstalmentPercent, Sex, OwnsProperty, Age, InstalmentPlans, Housing, ExistingCreditsCount, Job, Dependents, Telephone, ForeignWorker, and Risk. The data shows various demographic and financial details for different customers. Below the DataFrame is another code cell labeled 'In [19]:' containing a single line of Python code to import matplotlib.pyplot and create a histogram of ages. The resulting histogram is displayed as a plot with blue bars, showing the distribution of ages from 0 to 4.

```
# 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_F6c6e00CCKUpjyBfgW4i',
    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 now we can do whatever we want with the body, as pandas accepts body as file-like object
if not hasattr(body, '__iter__'): body.__iter__ = types.MethodType(__iter__, body)

df_data_1 = pd.read_csv(body)
df_data_1.head()

Out[18]:
   CustomerID CheckingStatus LoanDuration  CreditHistory  LoanPurpose  LoanAmount ExistingSavings EmploymentDuration  InstalmentPercent  Sex  ...  OwnsProperty  Age  InstalmentPlans  Housing  ExistingCreditsCount  Job  Dependents  Telephone  ForeignWorker  Risk
0  4e561ef3927-4107-aa12-1ba5772e0d3  less_0       6  all_credits_paid_back  car_used     250  less_100      1_1n_4            2  male  ...  savings_insurance  26  stores  rent        1  skilled  1  none  yes  Risk
1  20834545-4098-40d4-8316-05305405405  less_0       14  all_credits_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-ea185854543  greater_200      5  credits_paid_to_date  car_used     250  less_100      4_1n_7            3  male  ...  savings_insurance  42  none  rent        1  skilled  1  none  yes  No Risk
3  c19905c4-42b0-4f06-8550-00000000000  less_0       11  all_credits_paid_back  furniture     2016  less_100            1  female  ...  car_other  34  bank  own        1  unemployed  1  none  yes  No Risk
4  cb27e09c-7534-4ef5-835a-4cb610008ca1  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

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()
```

Felicidades ha terminado su Jupyter Notebook

12. AutoAI y Despliegue de modelos

Regrese al proyecto y vaya a la parte de “Gestionar” (color verde), despues el la sección “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 highlighted with a red box. On the right, a large blue button labeled 'Asociar servicio' is highlighted with a yellow box.

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

The screenshot shows the 'Asociar servicio' dialog box. At the top right, there is a blue button labeled 'Asociar servicio' which is highlighted with a red box.

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 again. The status column for the selected service 'WatsonMachineLearning' shows a green bird icon, indicating it is associated. This row is highlighted with a red box.

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

The screenshot shows the 'Activos' tab selected in the navigation bar. On the right, there's a sidebar titled 'Datos de este proyecto' with a section for uploading files. The main area displays a table of assets, with two entries visible: '18-1 - Riesgo Crediticio - Solicitantes.csv' and '18-1 - Riesgo Crediticio - Histórico.csv'. The 'Nuevo activo' button is highlighted with a red box.

y seleccione “Auto AI” (color verde).

The screenshot shows the 'Nuevo activo' dialog. Under the 'Creadores automáticos' heading, the 'AutoAI' tool is highlighted with a green box. It is described as a tool that analyzes tabular data and generates predictive model candidates. Other tools listed include 'Data Refinery', 'DataStage', 'Decision Optimization', and 'SPSS Modeler'. At the bottom, there are 'Cancelar' and 'Siguiente' buttons.

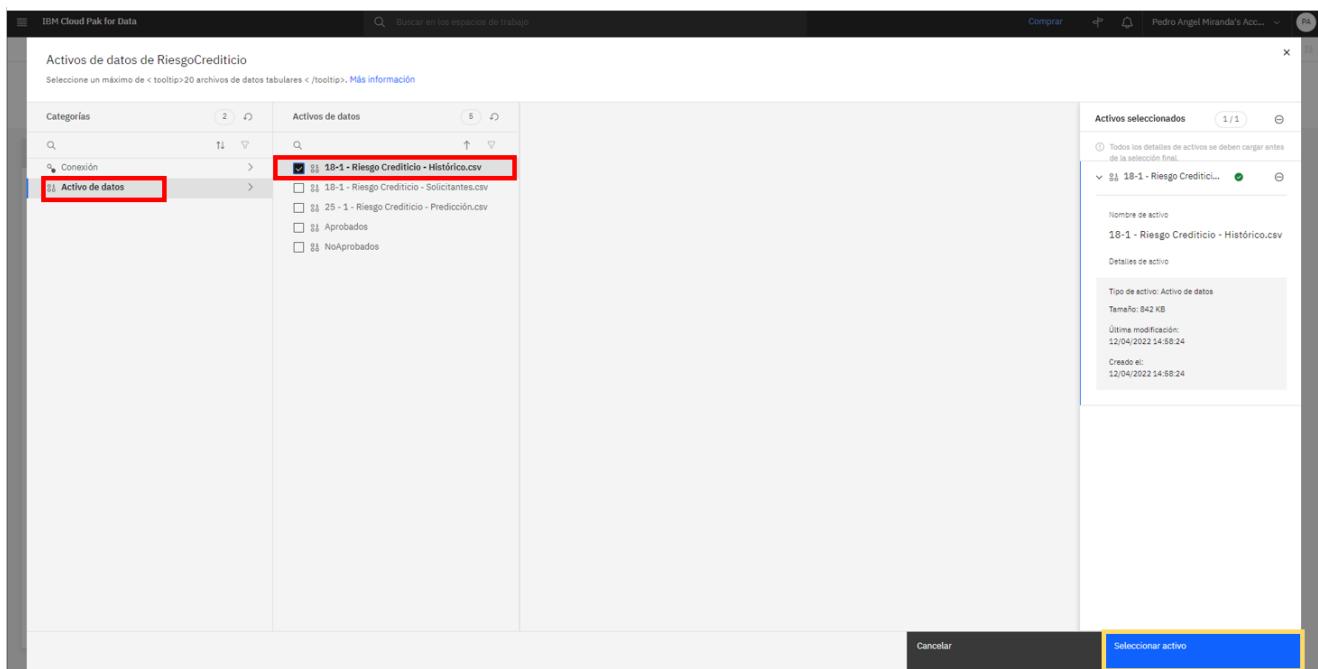
12.c) Nombre su experimento de AutoAI y la descripción como “Modelo Riesgo Crediticio” (color rojo) y oprima la opción de “Crear” (color amarillo).

The screenshot shows the 'Nuevo experimento de AutoAI' (New AutoAI Experiment) screen. In the 'Definir detalles' (Define details) section, the 'Nombre' (Name) field contains 'Modelo Riesgo Crediticio' and the 'Descripción (opcional)' (Optional description) field also contains 'Modelo Riesgo Crediticio', both of which are highlighted with red boxes. In the 'Definir configuración' (Define configuration) section, the 'Instancia de servicio de Watson Machine Learning' (Watson Machine Learning service instance) dropdown is set to 'WatsonMachineLearning'. Below it, the 'Definición de entorno' (Environment definition) shows '8 vCPU y 32 GB de RAM'. At the bottom right, the 'Crear' (Create) button is highlighted with a blue box.

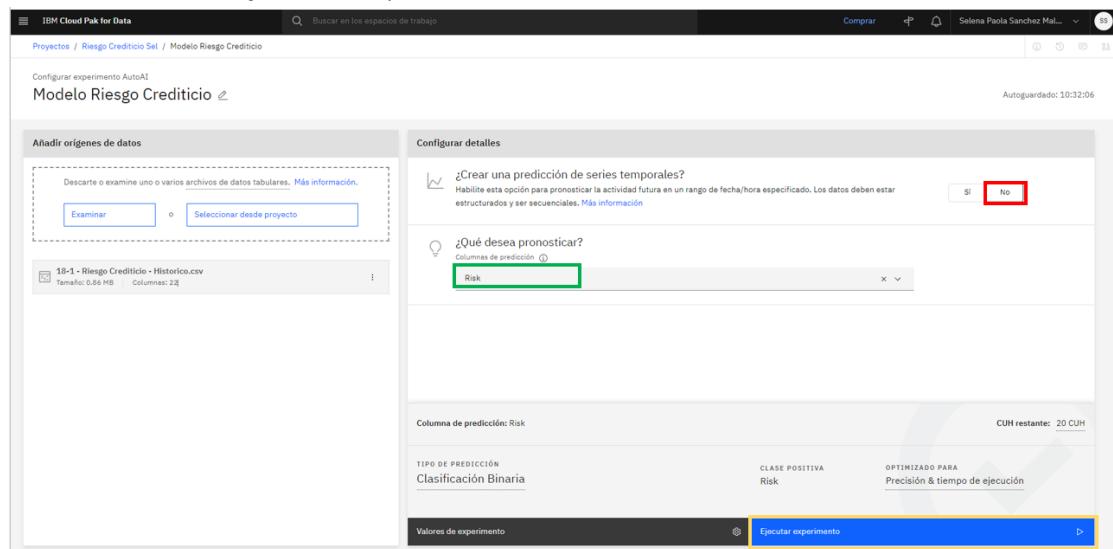
Se mostrará la siguiente pantalla, de click en Seleccionar desde proyecto.

The screenshot shows the 'Configurar experimento AutoAI' (Configure AutoAI experiment) screen for the 'Modelo Riesgo Crediticio' project. It displays the 'Añadir orígenes de datos' (Add data sources) section, which includes a central area for dropping or browsing files, a 'Drop or browse for one or more tabular data files.' instruction, and two buttons: 'Examinar' (Browse) and 'Seleccionar desde proyecto' (Select from project). The 'Seleccionar desde proyecto' button is highlighted with a blue box.

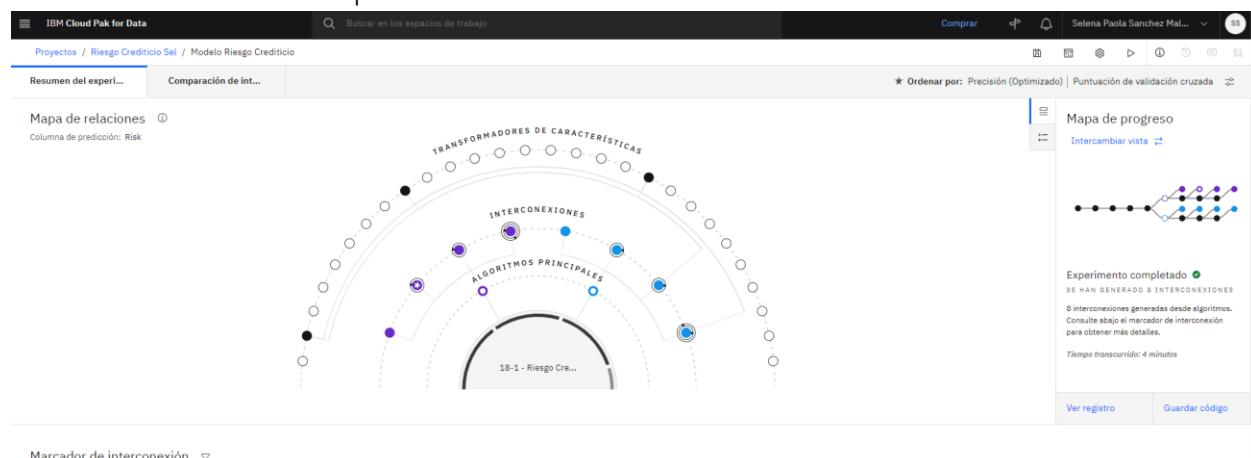
12.d) A continuación, le aparecerá la siguiente pantalla, y debe dar clic en “Activos de Datos” (color rojo), aparecerán a la derecha las fuentes de datos que tenemos en nuestro proyecto, seleccione la fuente de datos “18-1 - Riesgo Crediticio – Historico.csv” (color rojo) y finalmente de clic en “Seleccionar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



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.



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.

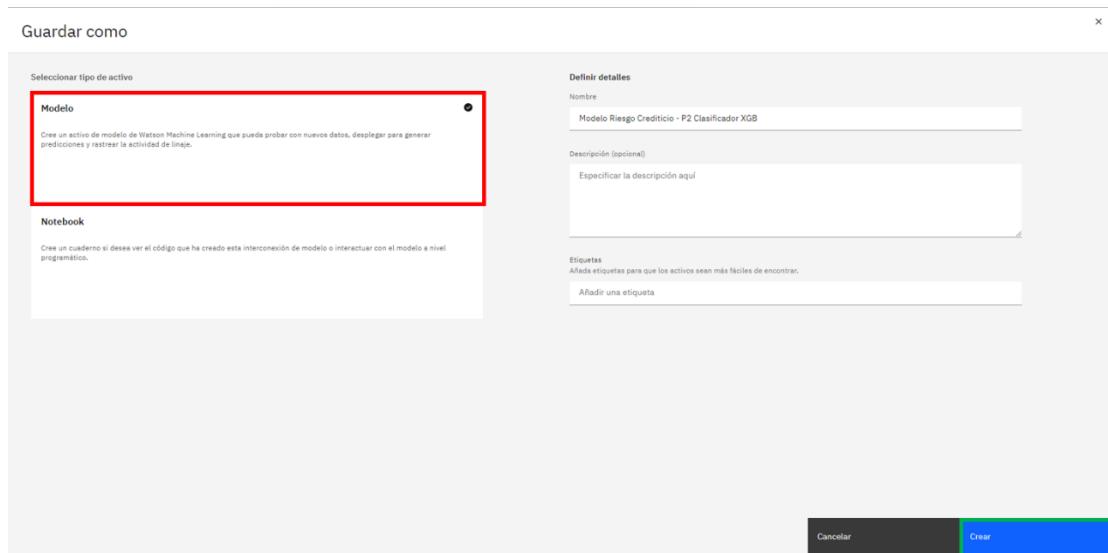
| Clasificación | Nombre | Algoritmo | Precisión (Optimizado) Validación Cruzada | Mejoras | Tiempo de creación |
|---------------|-----------------|-------------------------------------|--|------------------|--------------------|
| ★ 1 | 2 Interconexión | Clasificador XGB | 0.803 | HPO-1 | 00:00:13 |
| 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.

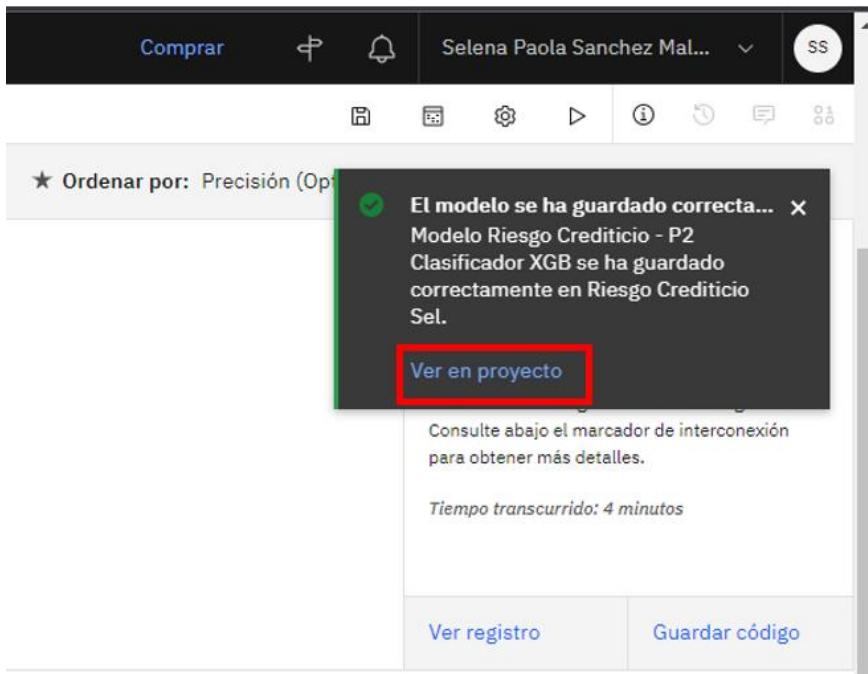
12.f) Seleccione el pipeline con la mejor clasificación (con la estrella) y haga click en el botón “Guardar como”(color rojo)

| Clasificación | Nombre | Algoritmo | Precisión (Optimizado) Validación Cruzada | Mejoras | Tiempo de creación | |
|---------------|-----------------|------------------|--|---------|--------------------|---------------------|
| ★ 1 | 2 Interconexión | Clasificador XGB | 0.803 | HPO-1 | 00:00:13 | Guardar como |

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 details of a model named "Modelo Riesgo Crediticio - P2 Clasificador XGB". On the right side, there is a sidebar with various configuration options. One option, "Promocionar al espacio de despliegue", is highlighted with a blue rectangle. The main panel displays the model's general information, including its description, labels, ID, creation date, and last modification date. The sidebar also shows the model's description, creation date, type, ID, software specification, hardware requirements, and tags.

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

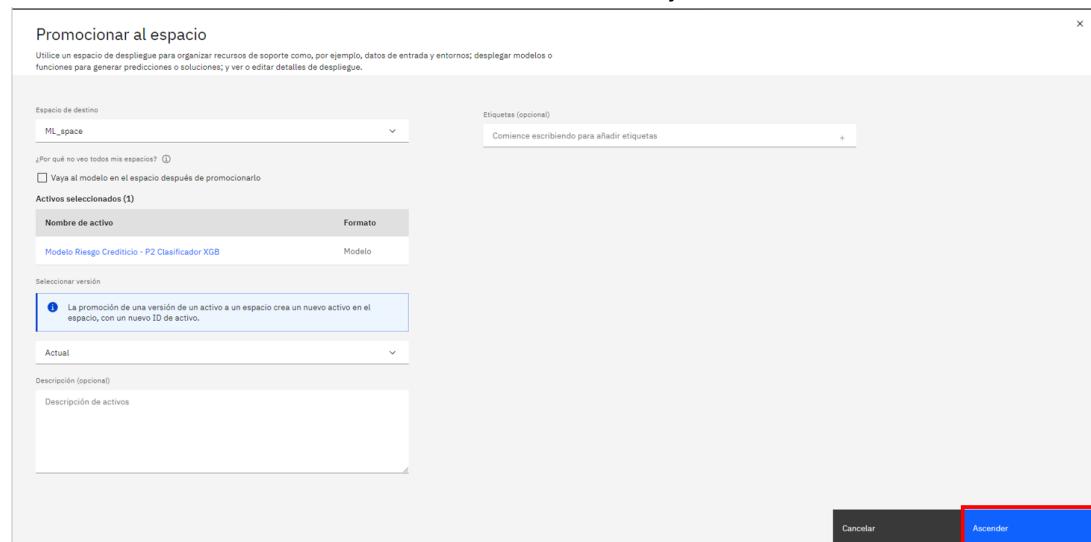
X

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

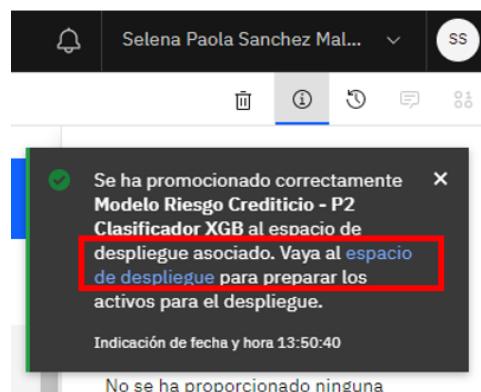
✓ Paso 1 de 1. Creación del espacio de despliegue.

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.

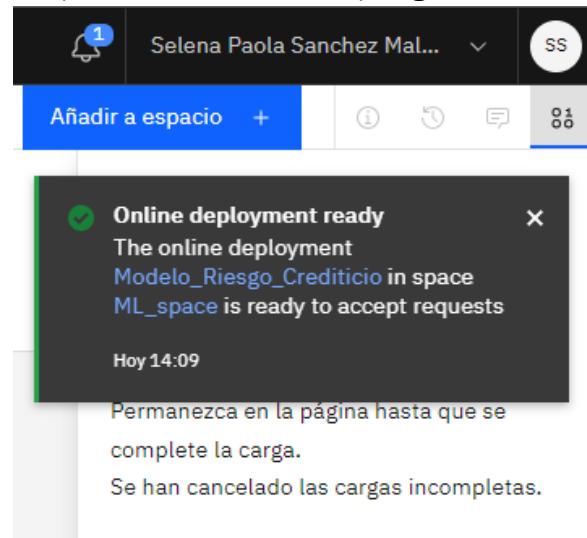
Nombre

Nombre del servicio

Descripción

Etiquetas
Añadir una etiqueta

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 $API_KEY below using information retrieved from your IBM Cloud account.
curl --insecure -X POST -H "Content-Type: application/x-www-form-urlencoded" --header "Accept: application/json"
--data-uzlenicode "grant_type=password&username=$API_USER&password=$API_PASSWORD"
--data-uzlenicode "apikey=$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/v1/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:

Cuestomer 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.