



Cloud Pak for Data as a Service

Instrucciones para crear su primer ambiente de Demo

# Análisis de Riesgo Crediticio

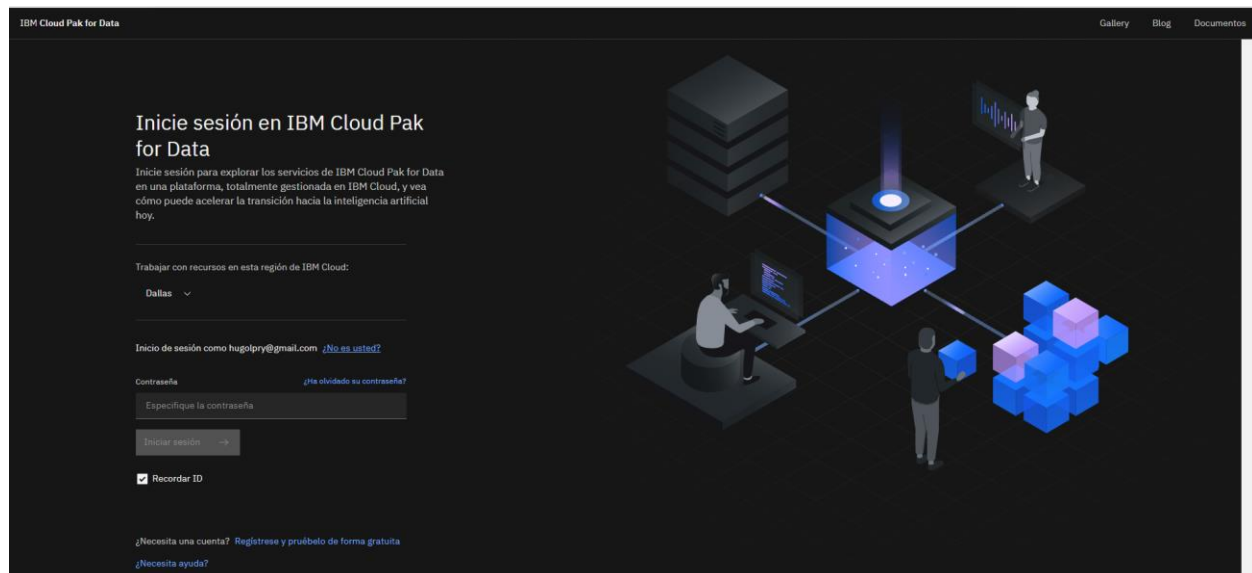
## Introducción:

- a) Le proporcionaremos dos archivos “csv” que serán sus fuentes de datos:
  - 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)
- c) Debe crear un modelo que determine el riesgo crediticio utilizando los datos históricos para entrenarlo y luego hacer la predicción del riesgo de los nuevos solicitantes de préstamos.
- d) Debe crear un tablero con gráficas a partir de la fuente de datos con las predicciones que haga el modelo.
- e) **Importante:** recientemente se realizaron actualizaciones a los servicios de CP4DaaS, en especial Watson Studio, por tal motivo puede ser que las pantallas no luzcan exactamente iguales a las presentadas en este instructivo, sin embargo deberá encontrar el mismo tipo información y opciones. Pedimos su comprensión en lo que generamos un instructivo actualizado. Gracias.

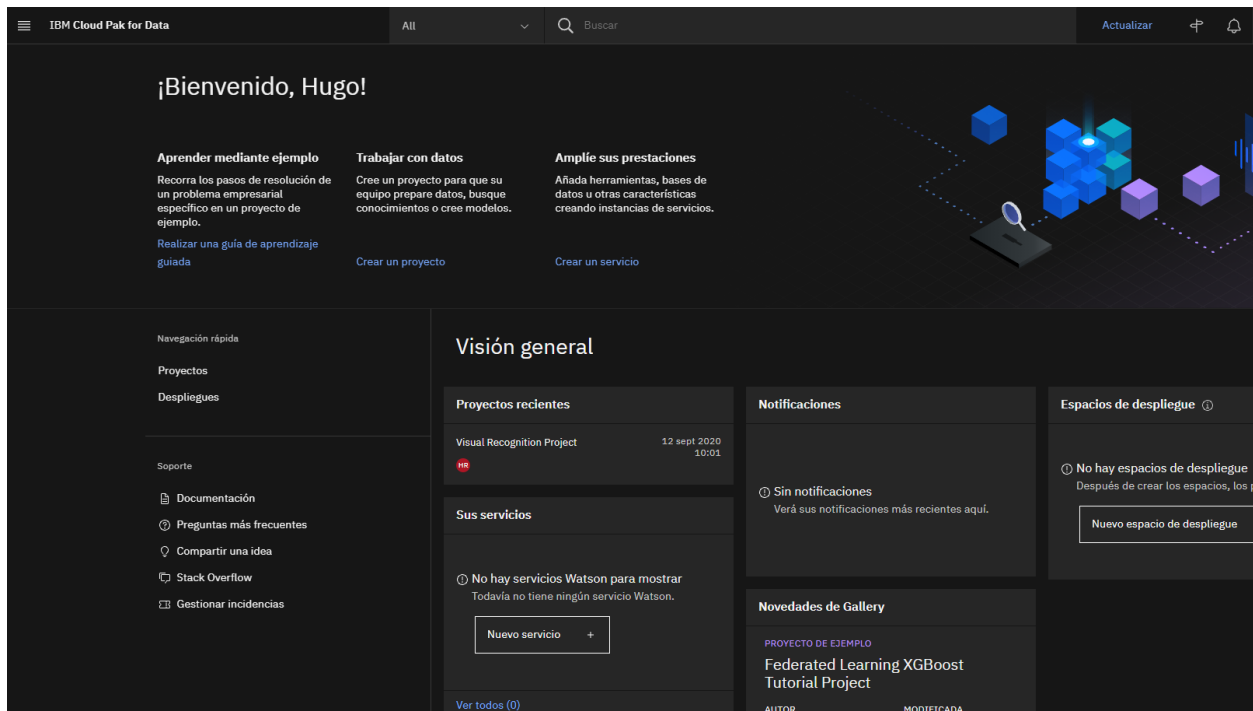
## 1. Entrar a la plataforma de Cloud Pak for Data

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

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

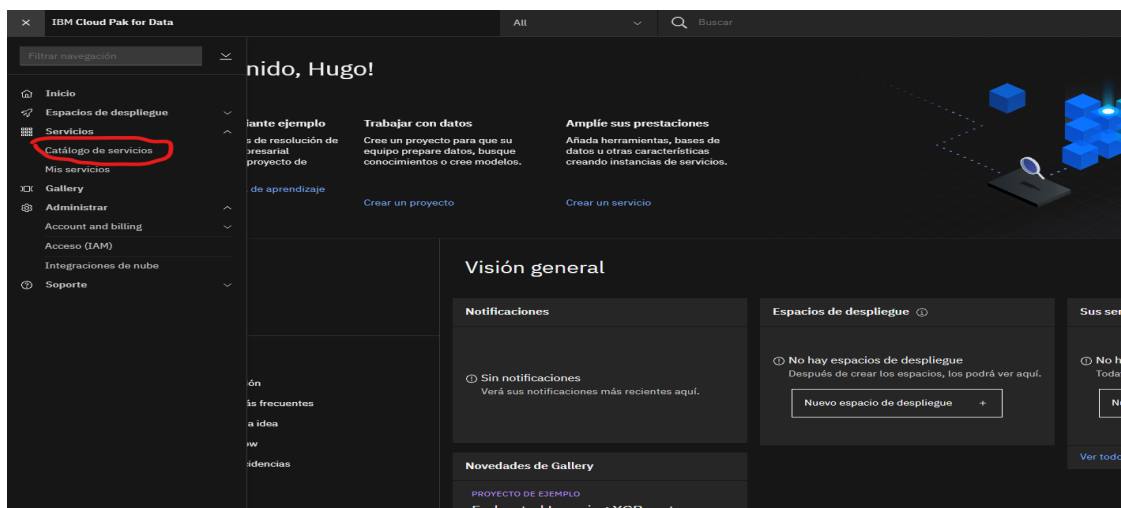


- Si no tiene cuenta vaya a la sección “¿Necesita una cuenta?” (parte baja izquierda de la pantalla) y cree su cuenta siguiendo las instrucciones, al definir su cuenta elija como región “Dallas”.
- Si ya tiene cuenta, proporcione su usuario/password y entre a la plataforma.
- Una vez que haya entrado a la plataforma verá una pantalla como la siguiente:

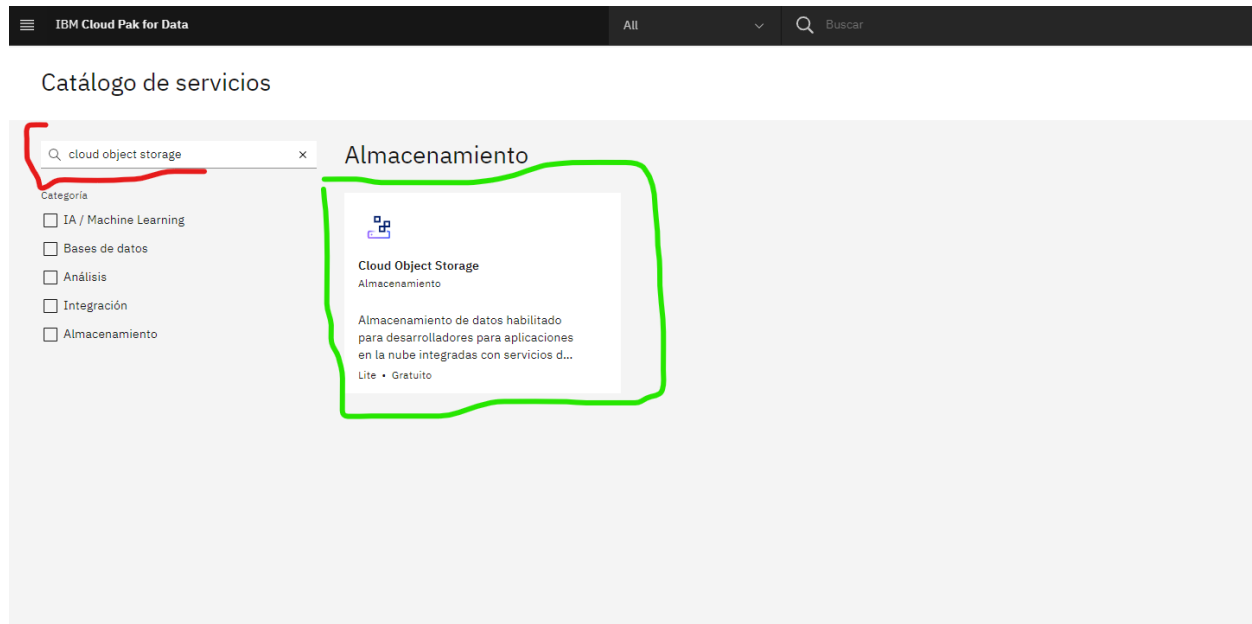


2. Ahora debe aprovisionar (habilitar) los servicios que utilizaremos para el ejercicio.

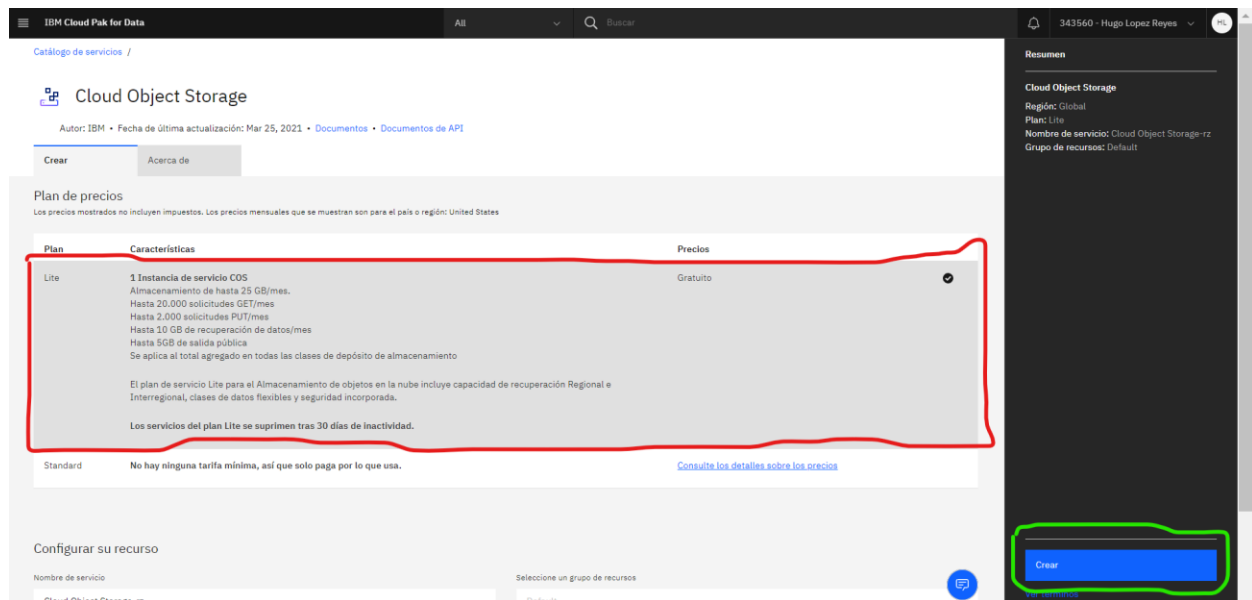
2.a) Ir al menú de “Hamburguesa” (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”



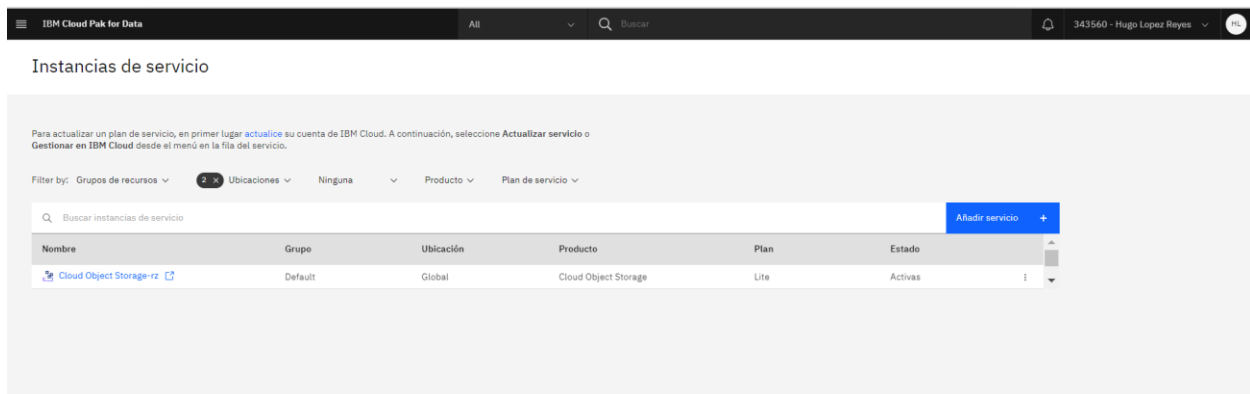
2.b) Utilizando la opción de búsqueda (lupa) encuentre el servicio “cloud object storage” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).



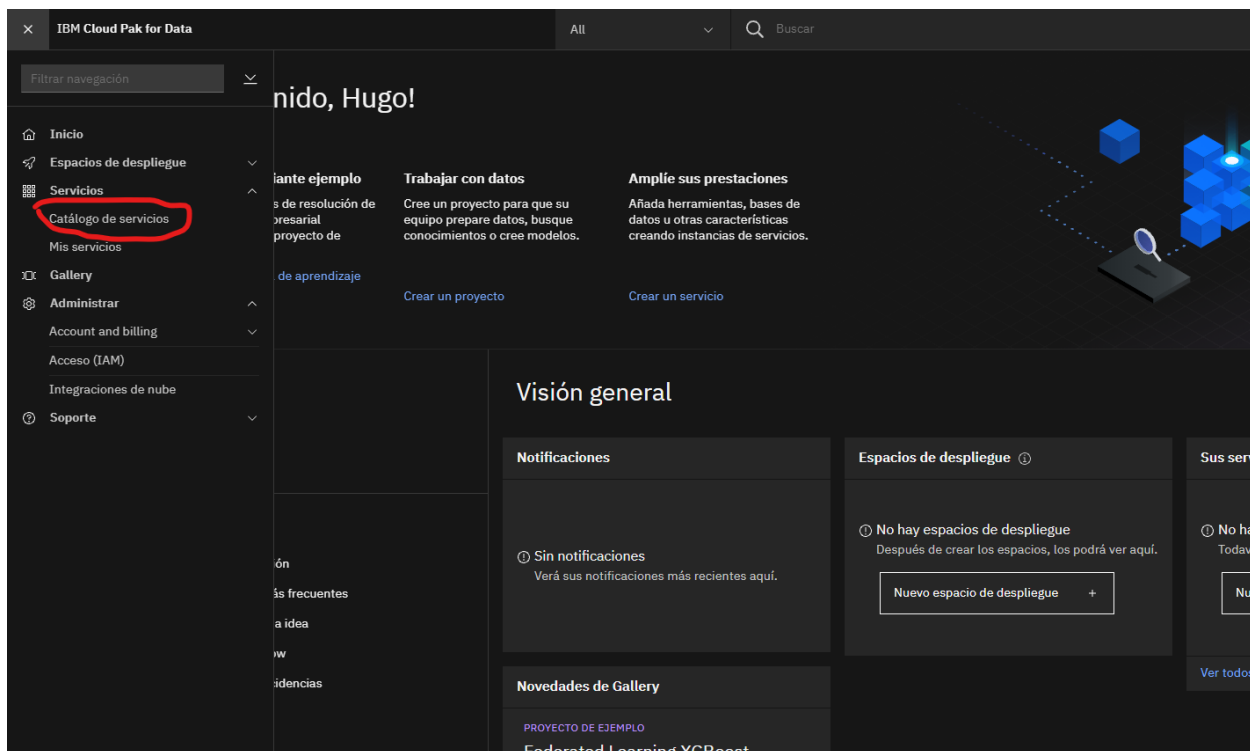
2.c) Una vez dentro del servicio de Cloud Object Storage, seleccione la opción “Lite” (color rojo) y elija la opción de crear (color verde).



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



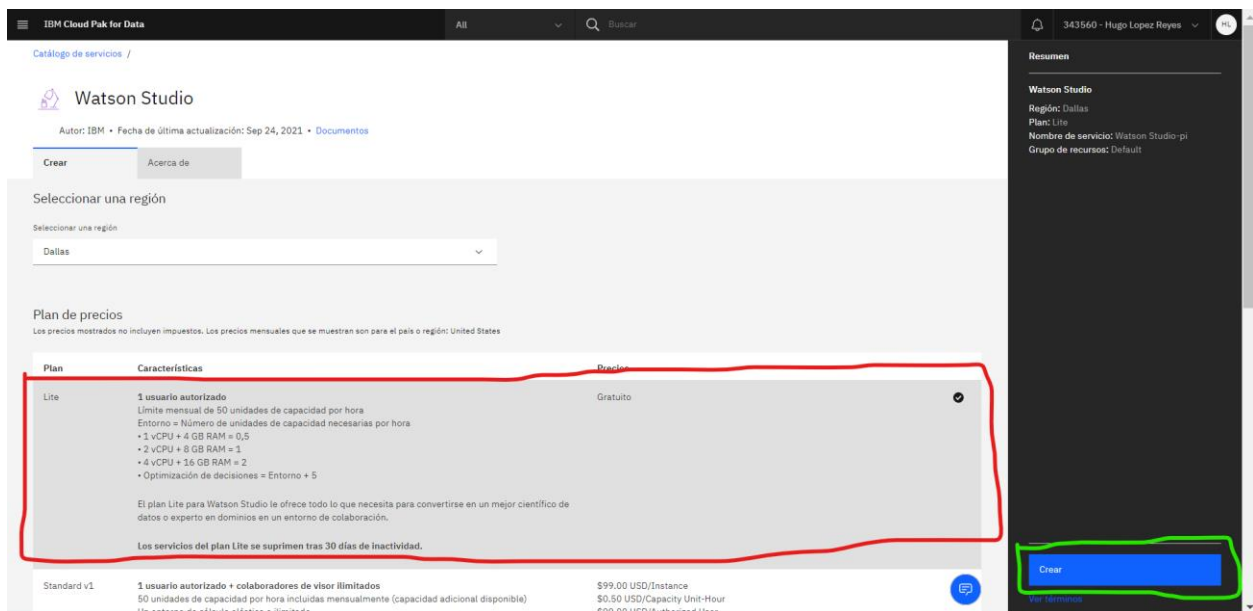
2.e) Vaya de nuevo al menú de “Hamburguesa” nuevamente (esquina superior izquierda), seleccione “Servicios” y luego “Catálogo de servicios”



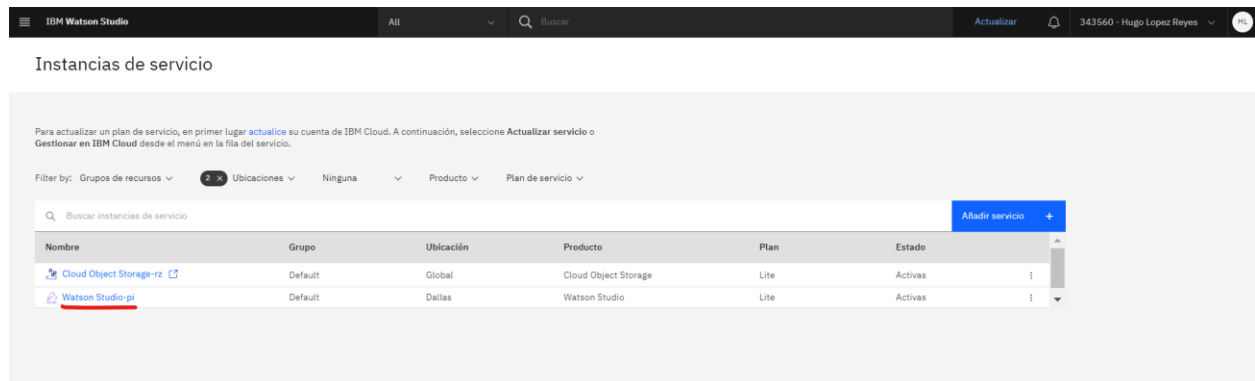
2.f) Utilizando la opción de búsqueda (lupa) encuentre el servicio “watson studio” (color rojo), una vez localizado dé clic en el ícono de dicho servicio (color verde).



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

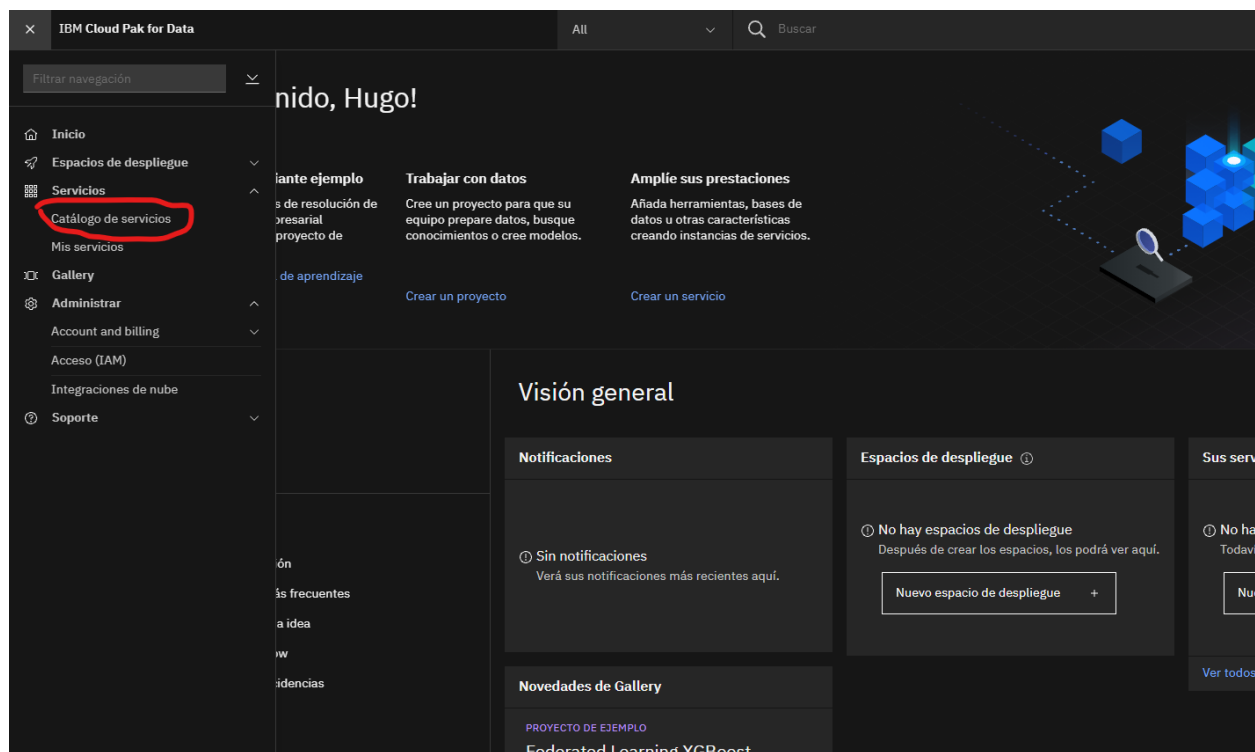


2.h) Si el servicio de Watson Studio fue creado correctamente (color rojo) deberá ver una pantalla como la siguiente:



Nombre	Grupo	Ubicación	Producto	Plan	Estado
Cloud Object Storage-rz	Default	Global	Cloud Object Storage	Lite	Activa
<u>Watson Studio-pi</u>	Default	Dallas	Watson Studio	Lite	Activa

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



IBM Cloud Pak for Data

Filtrar navegación

- Inicio
- Espacios de despliegue
- Servicios
- Catálogo de servicios**
- Mis servicios
- Gallery
- Administrar
- Account and billing
- Acceso (IAM)
- Integraciones de nube
- Soporte

¡Bienvenido, Hugo!

Trabajar con datos

Amplie sus prestaciones

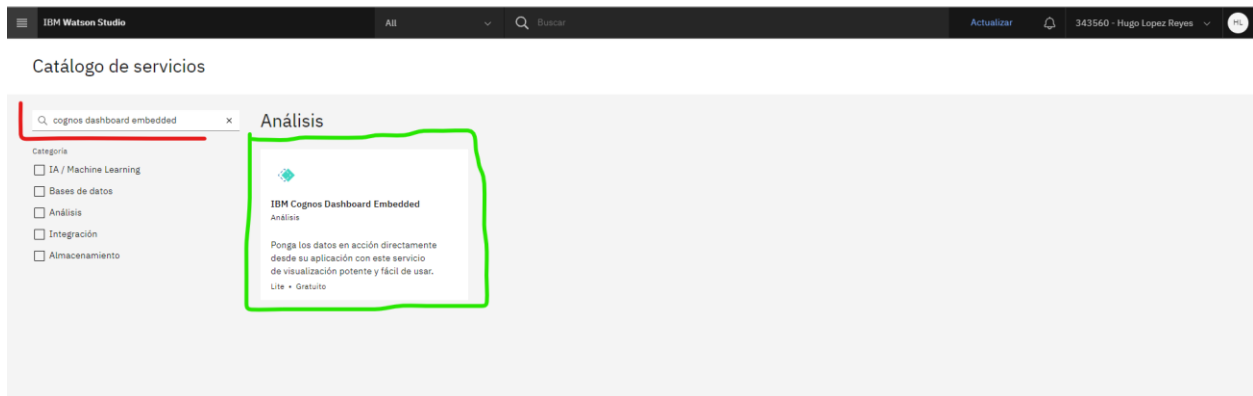
Visión general

Notificaciones

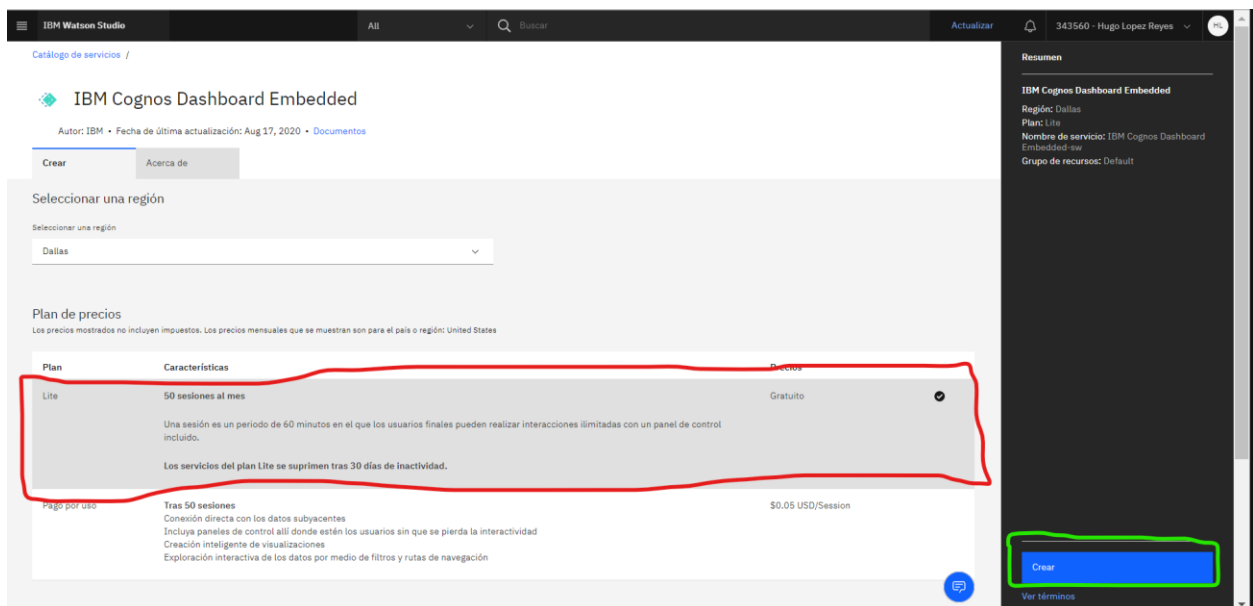
Espacios de despliegue

Novedades de Gallery

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, seleccione la opción “Lite” (color rojo) y elija la opción de crear (color verde).



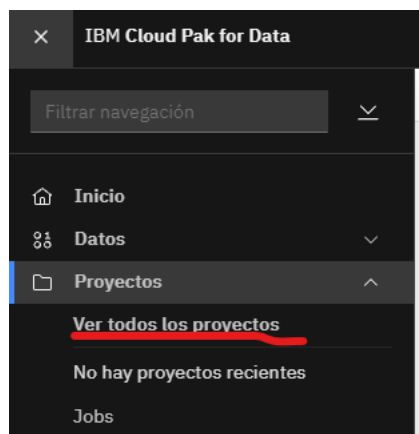


2.l) Si el servicio de Cognos Dashboard Embedded fue creado correctamente (color rojo) deberá ver una pantalla como la siguiente:

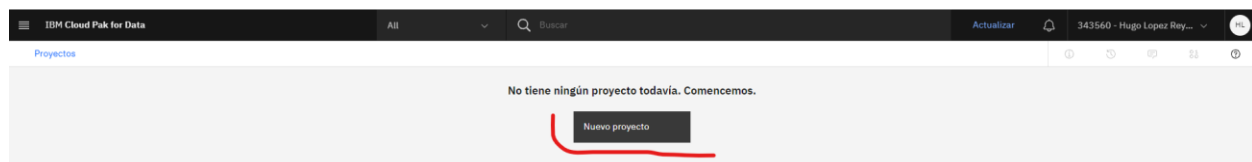
Nombre	Grupo	Ubicación	Producto	Plan	Estado	
Cloud Object Storage-rz	Default	Global	Cloud Object Storage	Lite	Activas	!
IBM Cognos Dashboard Embedded-sw	Default	Dallas	IBM Cognos Dashboard Embedded	Lite	Activas	!
Watson Studio-pi	Default	Dallas	Watson Studio	Lite	Activas	!

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

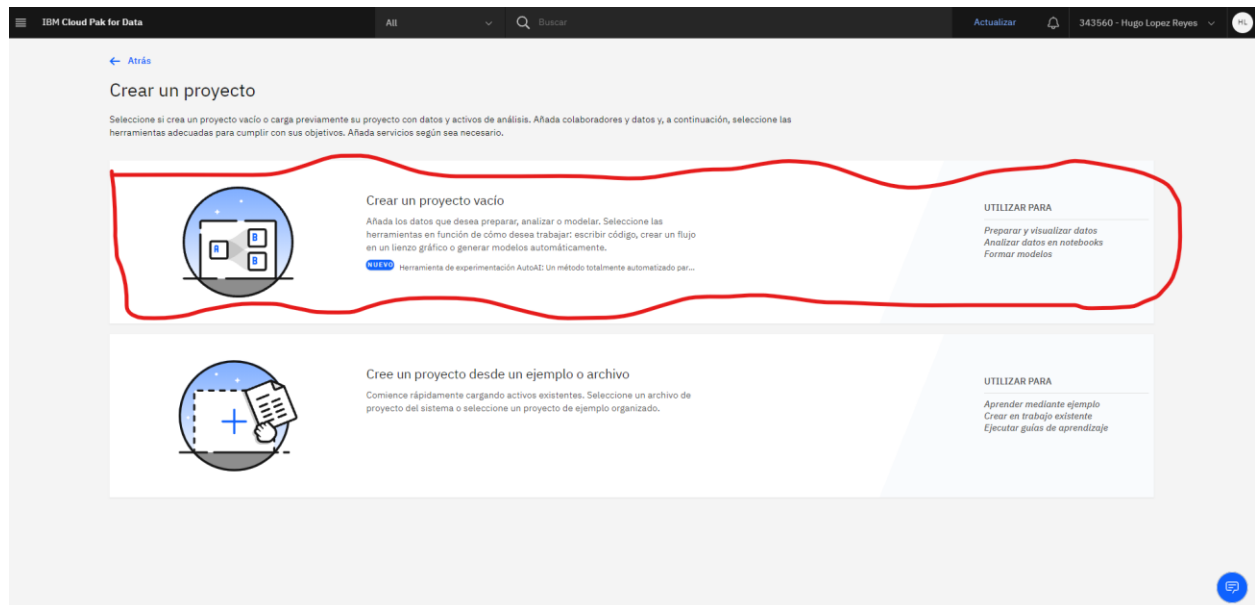
3. Ahora vamos a crear nuestro proyecto de Riesgo Crediticio, vaya al menú de “Hamburguesa” y seleccione la opción “Proyectos” y “Ver Todos los Proyectos” (color rojo).



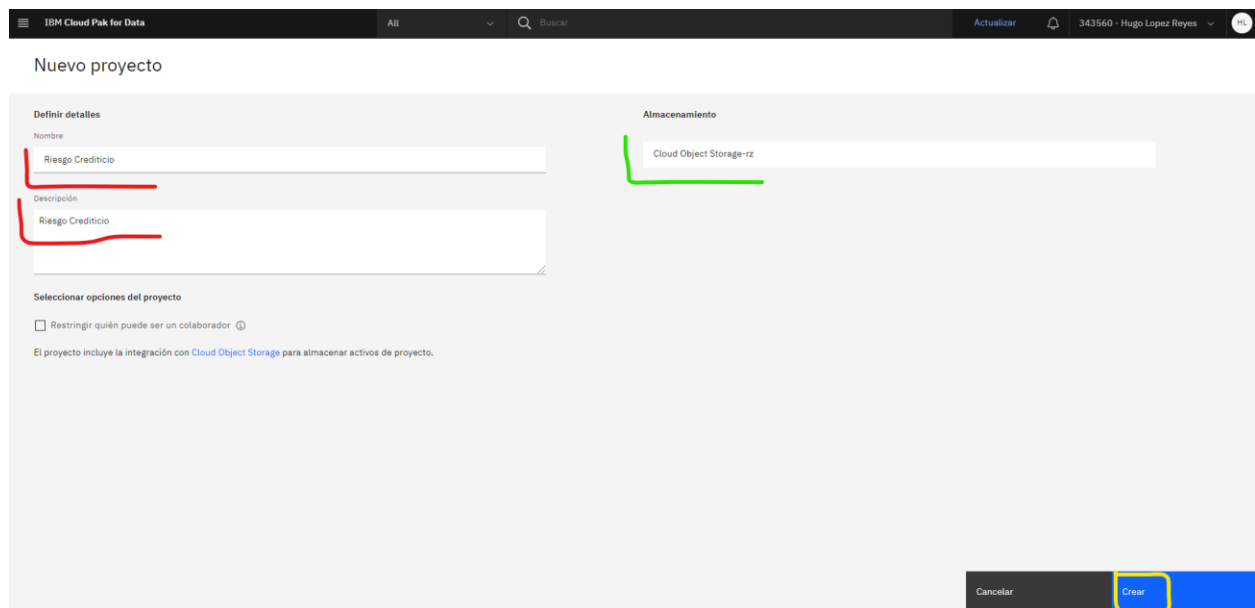
3.a) Si usted seleccionó la opción correcta verá una pantalla como ésta y ahí debe seleccionar la opción “Nuevo Proyecto” (color rojo) para crear su primer proyecto.



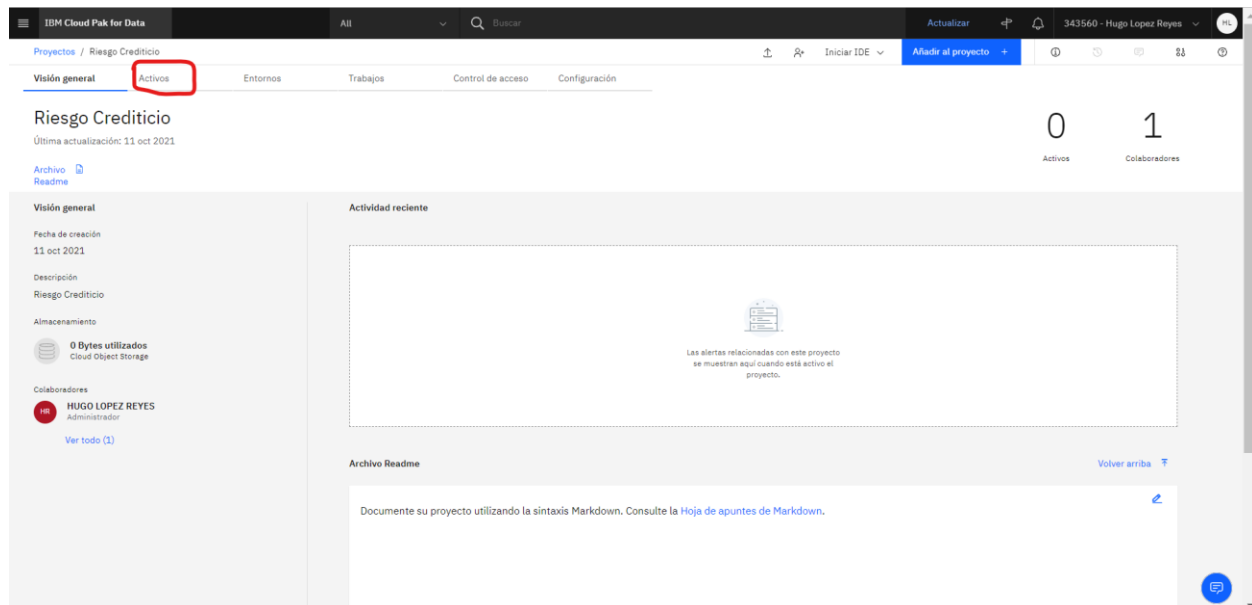
3.b) Ahora de clic en la opción “Crear un proyecto vacío” (color rojo).



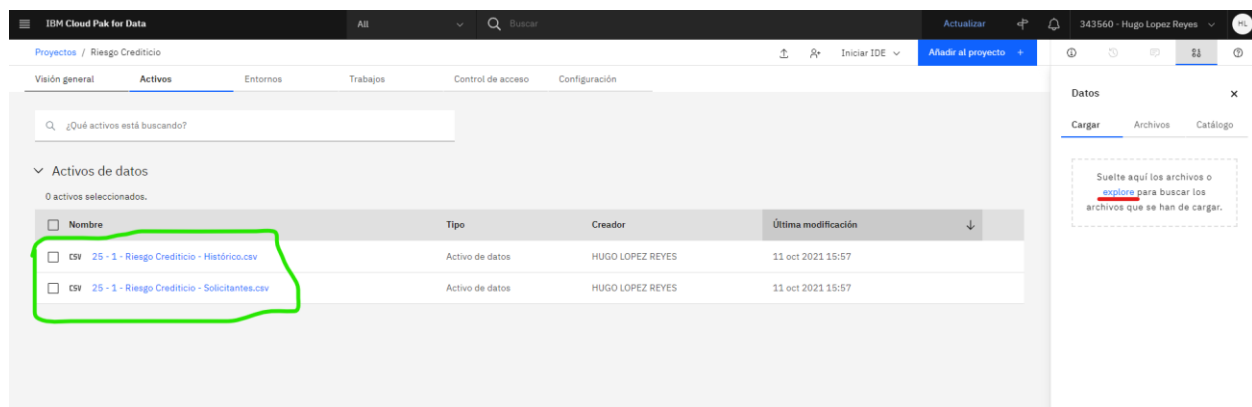
3.c) A continuación defina el “Nombre” y “Descripción” del proyecto como “Riesgo Crediticio” (color rojo), el “Almacenamiento” le debe aparecer auto-llenado con el nombre del servicio de Cloud Object Storage que creo anteriormente, si todo aparece como se ve en la siguiente pantalla entonces oprima el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

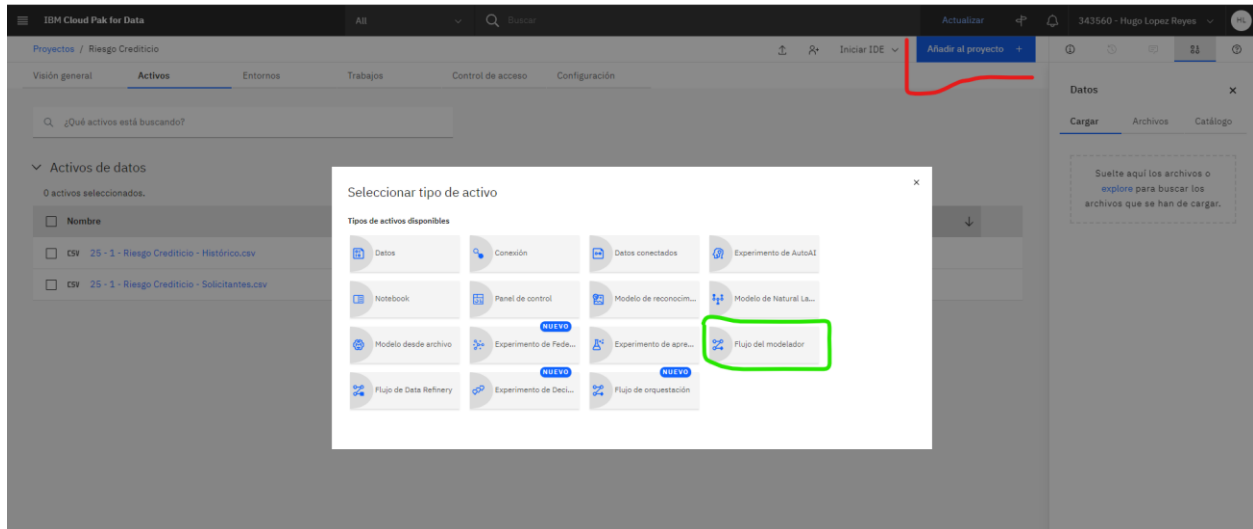


3.e) Ahora descargue primero desde box hacia su PC los dos archivos que han sido proporcionados para la demo como parte del curso (“25 - 1 - Riesgo Crediticio – Histórico” y “25 - 1 - Riesgo Crediticio – Solicitantes” (tome nota del directorio donde queden guardados). A continuación regrese a CP4DaaS a la misma pantalla de “Activos” del proyecto donde nos encontrábamos, seleccione la opción “explore” (a la derecha de su pantalla, color rojo) y descargue los dos archivos csv desde su computadora personal a su proyecto (tómelos del directorio donde los dejo al bajarlos de box, tip: puede seleccionar y bajar los dos a la vez o uno por uno si lo desea). Si todo lo hizo bien, una vez que haya descargado los dos archivos deben aparecer ambos en la sección “Activos de Datos” (color verde)

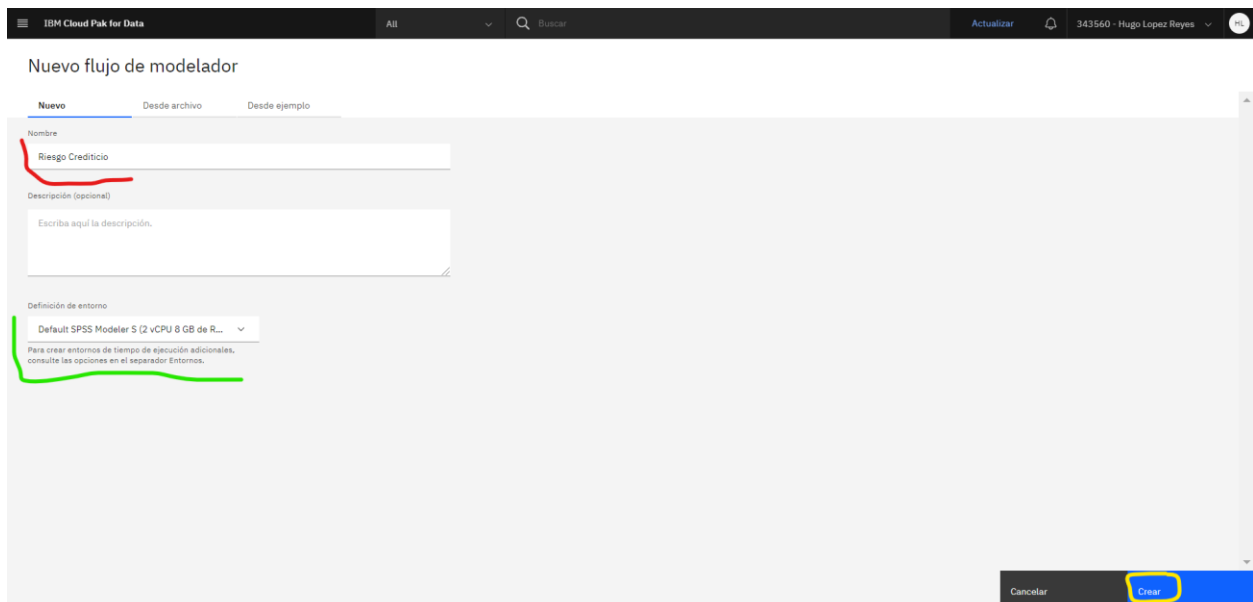


“Felicidades ya creo su proyecto y descargó las dos fuentes de datos que va a utilizar para su modelo”.

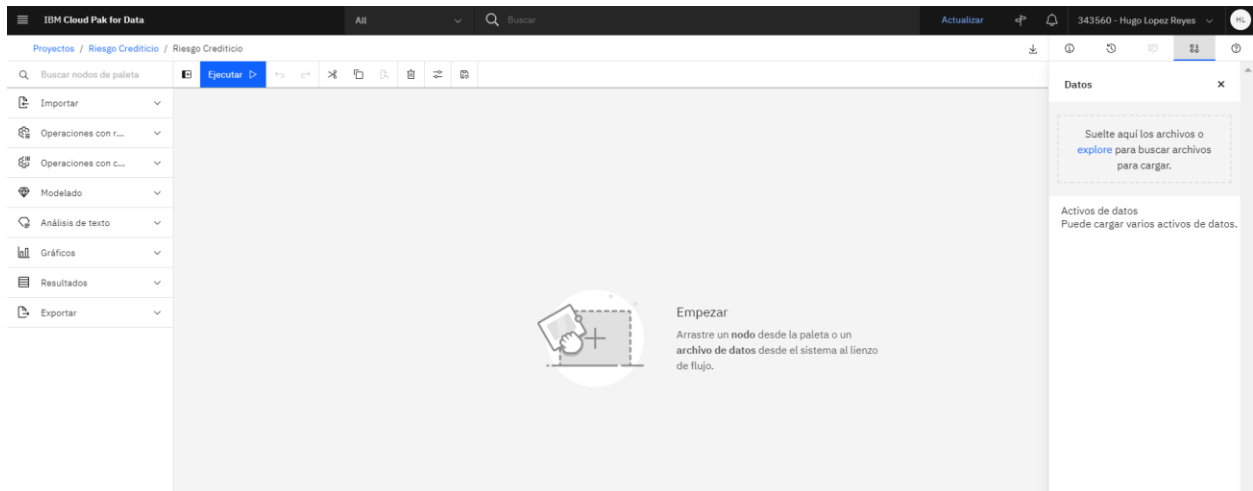
4. Ahora vamos a crear un modelo analítico utilizando una herramienta visual (SPSS Modeler) que la encontramos dentro de Watson Studio, por favor oprima el botón azul “Añadir al proyecto” (a su derecha en la parte superior de la pantalla, marcado con color rojo) y a continuación la opción “Flujo del Modelador” (color verde).



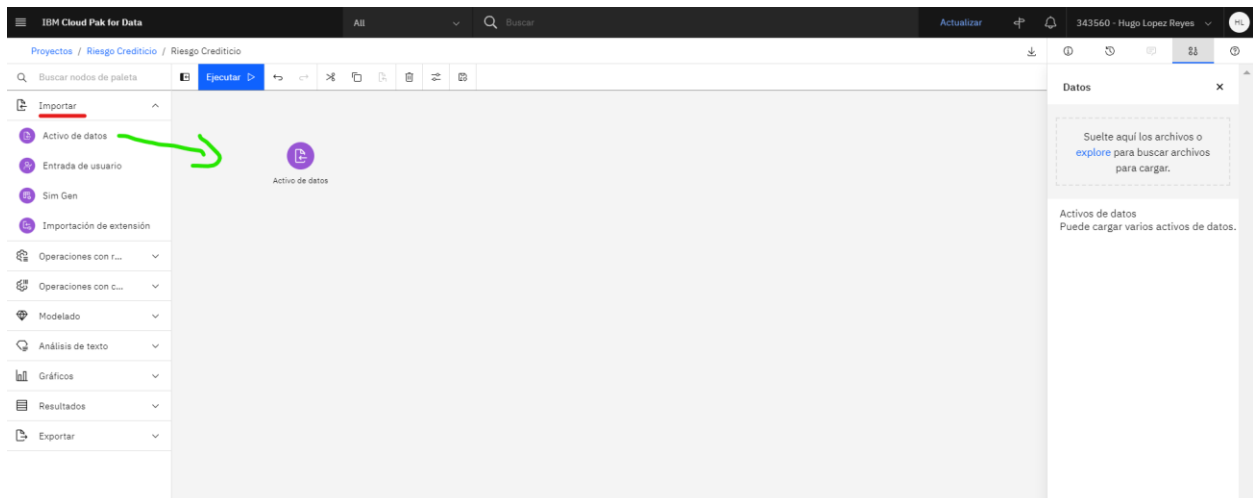
4.a) Nombre su modelo como “Riesgo Crediticio” (color rojo), deje el valor que por default le aparece en “Definición de entorno” (color verde) y oprima la opción de “Crear” (color amarillo, en la esquina inferior derecha), una vez que oprima Crear espere a que se genere en el entorno de ejecución del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



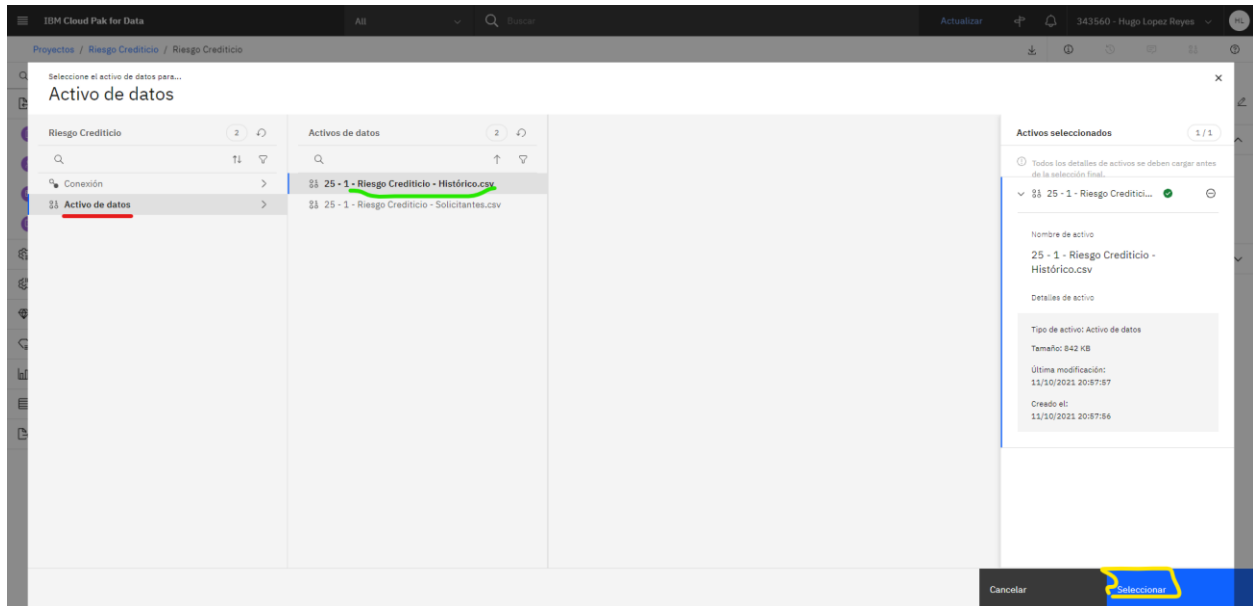
4.b) Si se ejecutó bien el paso anterior le aparecerá la siguiente pantalla con un lienzo en blanco.



4.c) A continuación primero 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 “25 - 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 “Operadores 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” (3ª 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).

The screenshot shows the IBM Cloud Pak for Data interface. On the left, a sidebar lists various data processing operations. The main canvas displays a workflow with two nodes: '25 - 1 - Riesgo...' and 'Tipo'. A green arrow points from the 'Tipo' node in the sidebar to the 'Tipo' node in the canvas. On the right, the 'Tipo' node's configuration panel is open, showing a table of fields from the 'Riesgo' dataset. The 'Risk' field is highlighted in yellow, and its role is set to 'Objetivo'. At the bottom right, there are 'Cancelar' and 'Guardar' buttons.

**Configuración de Tipo**

Modalidad predeterminada: ☒ Leer metadatos ☐ Pasar (no explorar)

Operaciones de tipo: [Leer valores](#) [Borrar todos los valores](#)

Campo	Medida	Rol	Modalidad de v...	Valores	Comprob...
<input type="checkbox"/> Installme	Catagórico	Entrada	Leidas		Ninguno
<input type="checkbox"/> Housing	Catagórico	Entrada	Leidas		Ninguno
<input type="checkbox"/> ExistingCr	Continuo	Entrada	Leidas		Ninguno
<input type="checkbox"/> Job	Catagórico	Entrada	Leidas		Ninguno
<input type="checkbox"/> Depender	Continuo	Entrada	Leidas		Ninguno
<input type="checkbox"/> Telephonv	Catagórico	Entrada	Leidas		Ninguno
<input type="checkbox"/> ForeignWli	Catagórico	Entrada	Leidas		Ninguno
<input checked="" type="checkbox"/> Risk	Catagórico	Objetivo	Leidas		Ninguno

Formato:

Anotaciones:

[Cancelar](#) [Guardar](#)

4.f) A continuación en “Operaciones con Campos” seleccione el nodo “Partición” y arrástrelo al lienzo (flecha roja), conecte el nodo “Tipo” al nodo “Partición” (color verde). 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 amarillo) y 30 en “Partición de prueba (%)” (color morado), finalmente dé clic en el botón “Guardar” (color rosa).

The screenshot displays the IBM Cloud Pak for Data interface. On the left, a sidebar lists various data processing operations. The main workspace shows a workflow with three nodes: '25 - 1 - Riesgo ...', 'Tipo', and 'Partición'. A red arrow indicates the 'Partición' node being dragged from the sidebar to the workspace. A green arrow connects the 'Tipo' node to the 'Partición' node. On the right, the 'Partición' configuration panel is open, showing settings for training and testing partitions. The 'Partición de entrenamiento (%)' field is set to 70 (yellow background) and the 'Partición de prueba (%)' field is set to 30 (purple background). The 'Asignación de partición repetible' checkbox is checked. The 'Guardar' (Save) button is highlighted with a pink box.

IBM Cloud Pak for Data

Proyectos / Riesgo Crediticio / Riesgo Crediticio

Buscar nodos de paleta

Ejecutar

25 - 1 - Riesgo ...

Tipo

Partición

Partición

Configuración

Nombre de campo derivado

Partición

Partición de entrenamiento (%)

70

Partición de prueba (%)

30

☐ Crear partición de validación

☒ Asignación de partición repetible

Semilla [Generar](#)

1234567

☐ Utilizar campo único para asignar particiones

Anotaciones

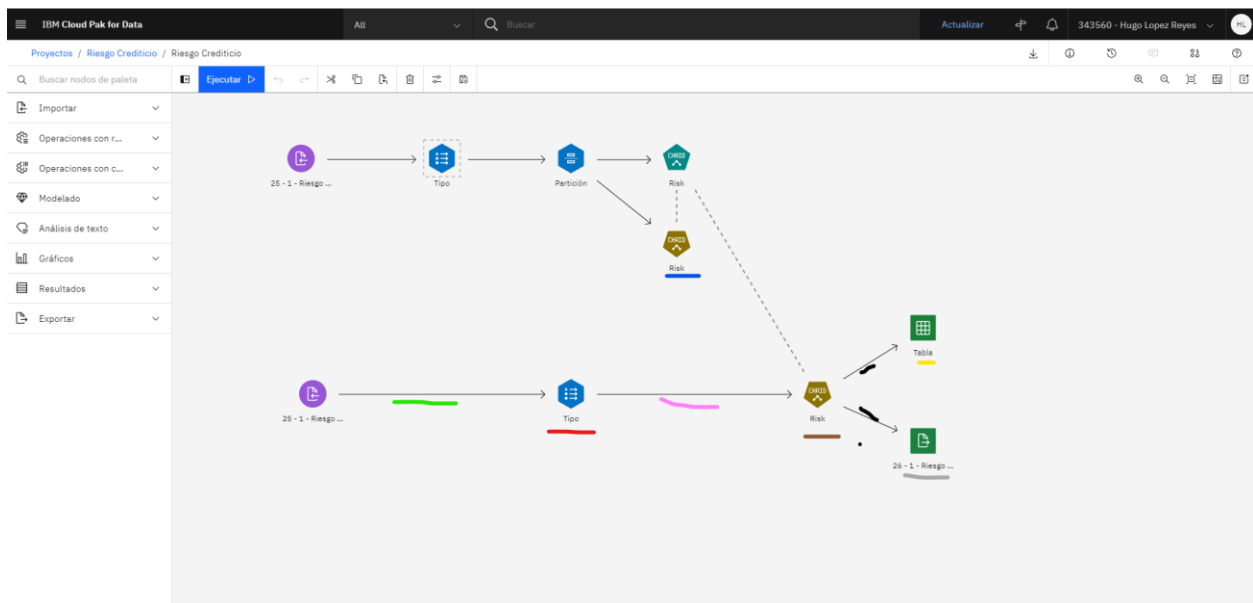
Cancelar **Guardar**



4.g) Ir a las opciones de la derecha buscar la sección de “Modelado” y arrastrar el ícono “CHAID” al lienzo (flecha roja), conectar el nodo “Partición” con el nodo “CHAID”, ahora abrir 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 ciérrrela 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 “25 - 1 - Riesgo Crediticio - Solicitantes”, dar clic al botón “Seleccionar” y en la siguiente pantalla dar clic al botón “Guardar” [color rosa]). Con éste nuevo ícono de “Activo de Datos” **estamos iniciando la creación del flujo de predicción.**

The screenshot displays the IBM Cloud Pak for Data interface. On the left, a sidebar lists various components under the 'Modelado' (Modeling) section, including 'Activo de datos', 'Entrada de usuario', 'Sim Gen', 'Importación de extensión', 'Operaciones con r...', 'Operaciones con c...', 'Modelado', 'Clasificador automático', 'Autonumérico', 'Clúster automático', 'Red bayesiana', 'CS.O', 'Árbol C&R', 'CHAID', 'Quest', and 'Tree-AS'. A red arrow points from the 'CHAID' icon in this list to a node in the central workflow canvas. The canvas shows a sequence of nodes: '25 - 1 - Riesgo ...', 'Tipo', 'Partición', and 'Risk'. A red arrow connects the 'Partición' node to the 'Risk' node. A purple arrow points from the 'Activo de datos' icon in the 'Importar' section to a new 'Activo de datos' node on the canvas. On the right, the 'Activo de datos' configuration panel is open, showing settings for 'Datos', 'Propiedades de formato de archivo', 'Codificación', 'Manejo de datos no válidos', 'Primera línea', and 'Formatos de datos'. The 'Guardar' (Save) button at the bottom right is highlighted with a pink circle.

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 “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”, 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). Abra el nodo “Tabla” (color amarillo) y de la opción “Ejecutar”, abra 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:



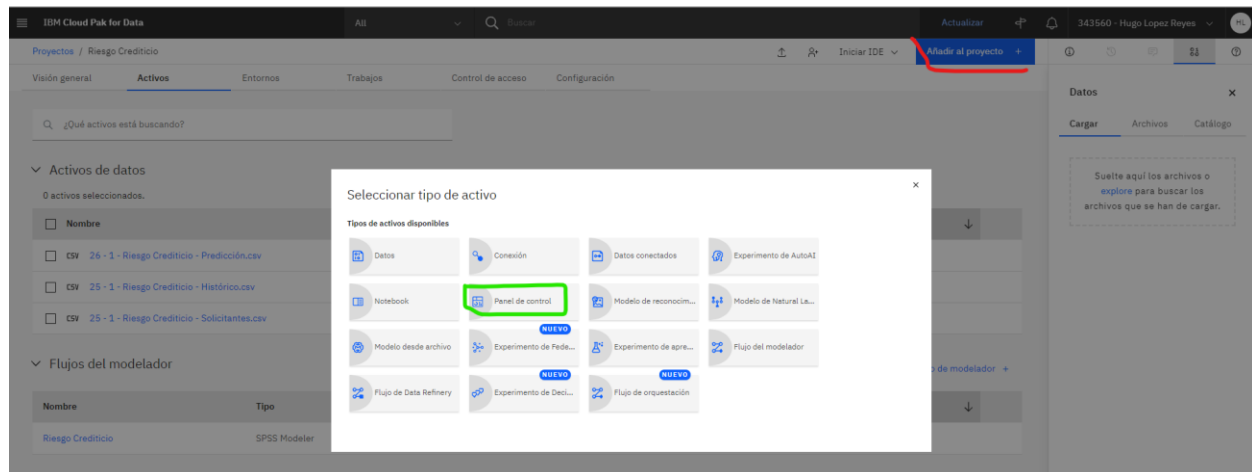
4.i) Obtenga una vista previa del nodo “Tabla” (color amarillo, punto anterior) y del nodo “Exportar Activo de Datos” (color gris, punto anterior) y confirme que en ambos aparezcan los campos: “\$R-Risk” y “\$RC-Risk” que se generan como resultado de correr el modelo de predicción.

**¡Felicidades su modelo está listo!**

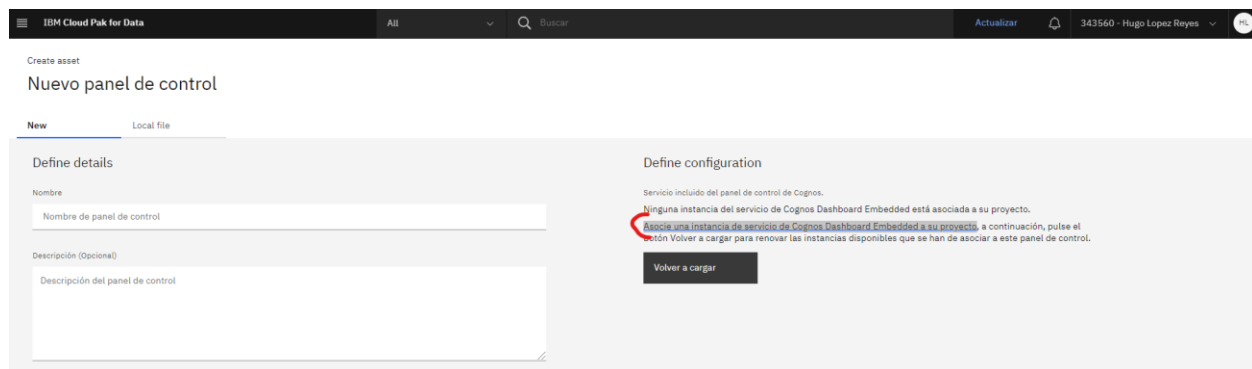
5. Creación de un Tablero a partir de los datos que tenemos. Regrese a su “Proyecto” (tip en la esquina superior izquierda dice: “Proyectos / Riesgo Crediticio / Riesgo Crediticio”, de clic en el link “Riesgo Crediticio” en color azul, el que viene después de la diagonal de “Proyectos”.

**Fe de erratas:** si en las instrucciones siguientes, de la sección 5, ve pantallas con un nombre de archivo iniciando con “26 - 1...” queremos decir “25 - 1...”

5.a) Dé clic al botón azul “Añadir al Proyecto” (color rojo) y seleccione “Panel de Control” (color verde).



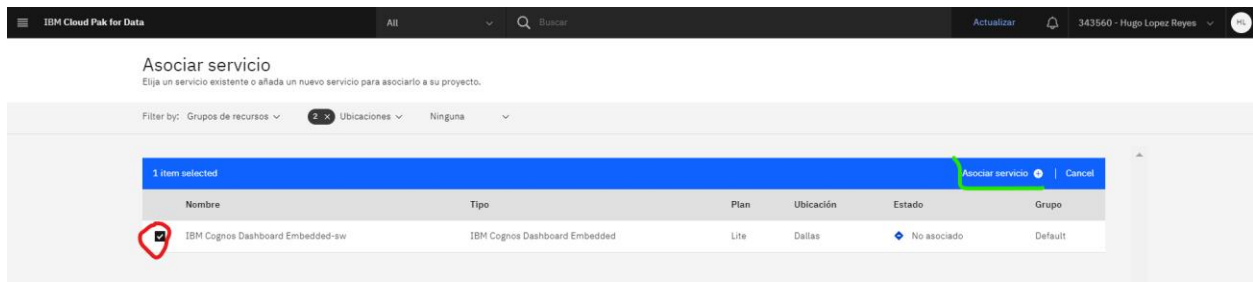
5.b) Ahora de clic en la opción “Asocie una instancia de servicio de Cognos Dashboard Embedded a su proyecto” (color rojo). Este paso se requiere únicamente cuando no tienen asociada una instancia de Cognos al proyecto, que debería ser el caso si es una cuenta nueva.



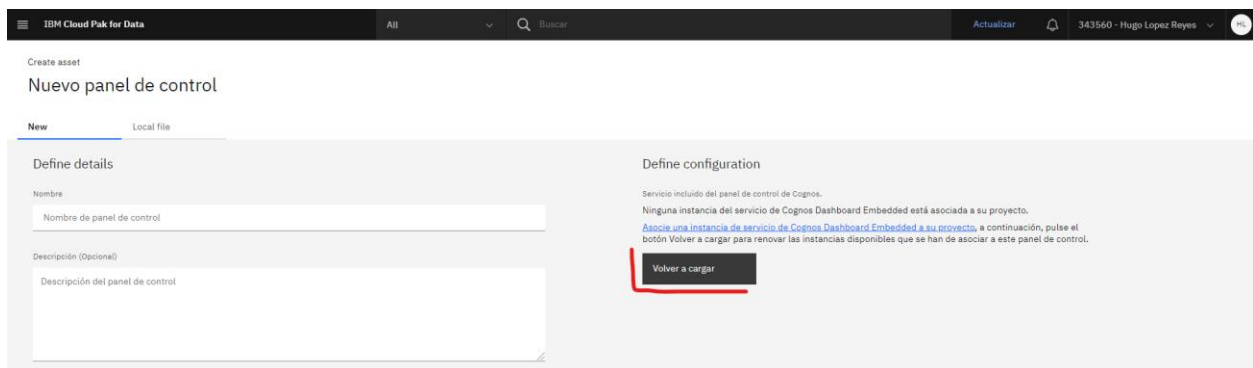
**Nota.** Si este paso no le funciona debe hacer lo siguiente:

Vaya al menú de “Hamburguesa” y elija “Proyectos/Ver todos los proyectos”, en el menú de proyectos seleccione “Riesgo Crediticio”, dentro del proyecto en el menú horizontal en la parte superior seleccione “Gestionar”, ahora verá un menú vertical a la izquierda donde debe seleccionar “Servicios e integraciones” y luego debe elegir “IBM Services”, finalmente seleccione el servicio “IBM Cognos Dashboard Embedded” y luego de click en “Asociar servicio”, cerciórese que la asociación sea exitosa. Ahora puede regresar al punto “5.” Inicial de esta sección y realice de nuevo la misma secuencia de pasos, excepto “5.b”, “5.c” y “5.d” que ya no serán necesarios.

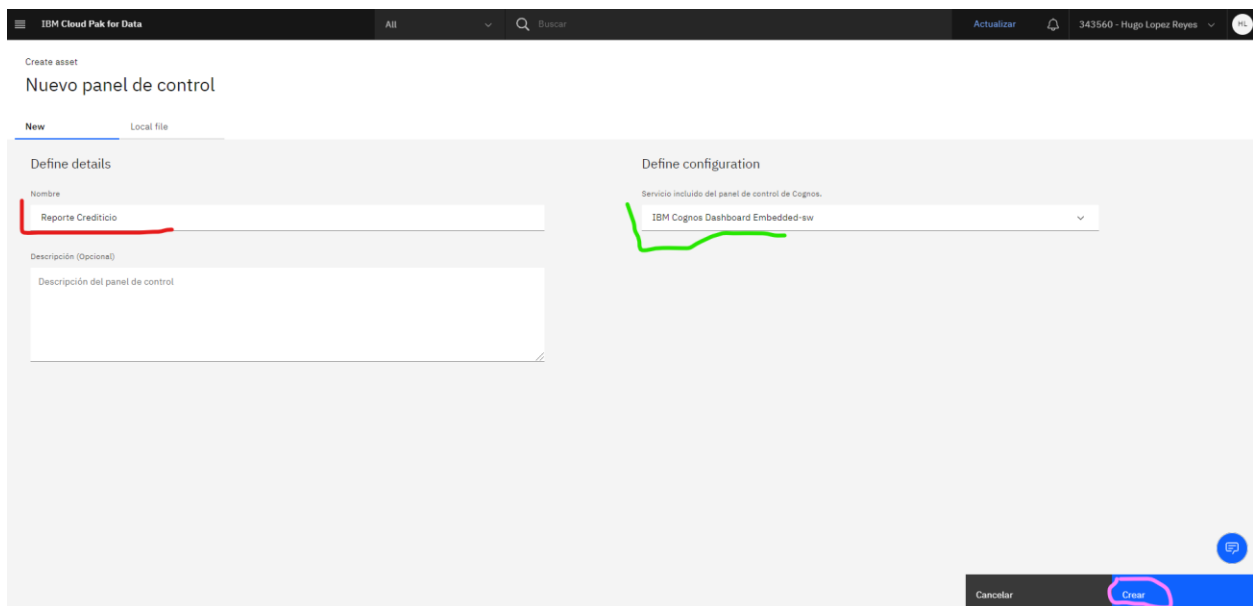
5.c) Seleccione la instancia de Cognos (color rojo) y de clic en “Asociar servicio” (color verde).



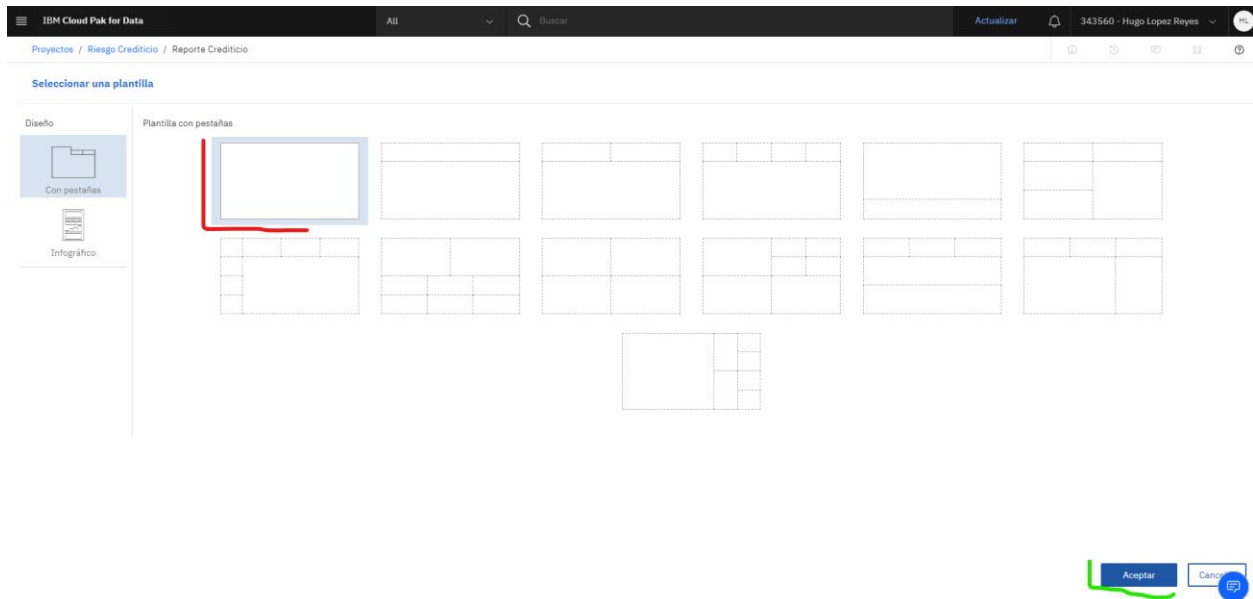
5.d) Dé clic en “Volver a cargar” (color rojo).



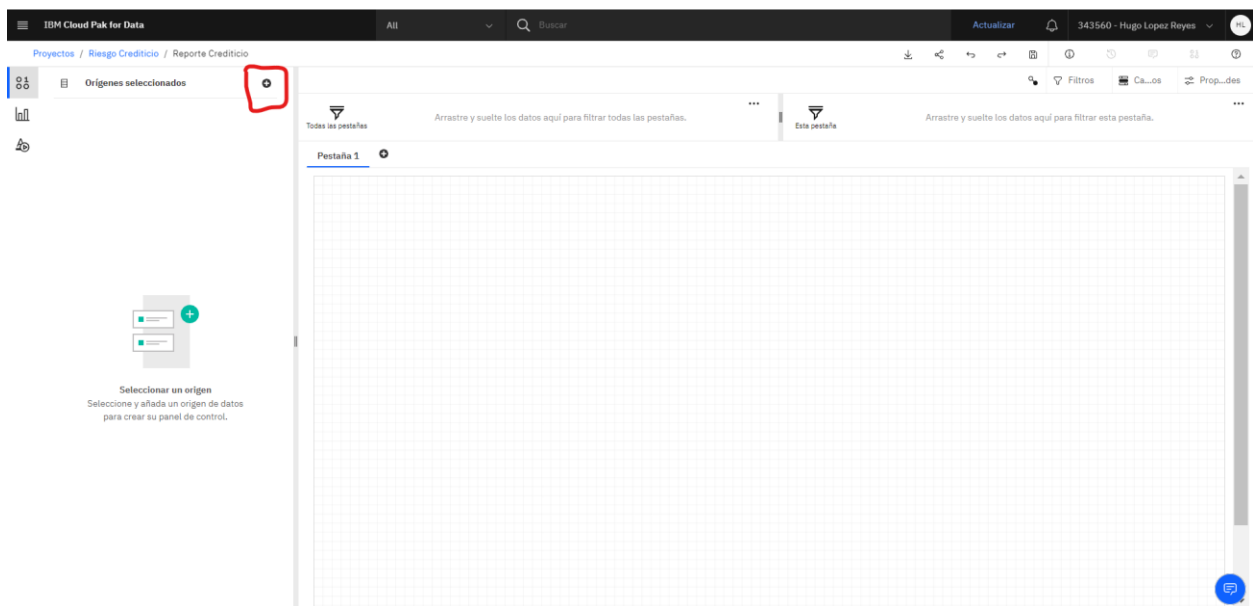
5.e) Ponga en Nombre “Reporte Crediticio” (color rojo), confirme que ya tiene una instancia de Cognos asociada al proyecto (color verde) y de clic en “Crear” (color rosa).



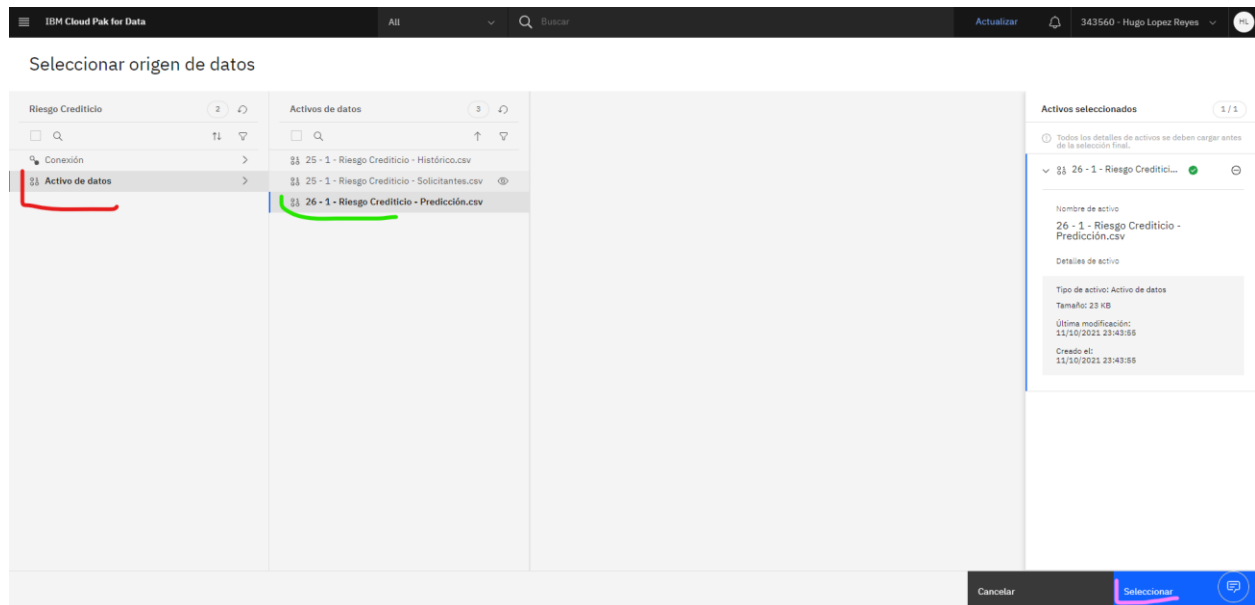
5.f) Seleccione la “Plantilla con pestañas” en blanco (color rojo) y de clic en “Aceptar” (color verde).



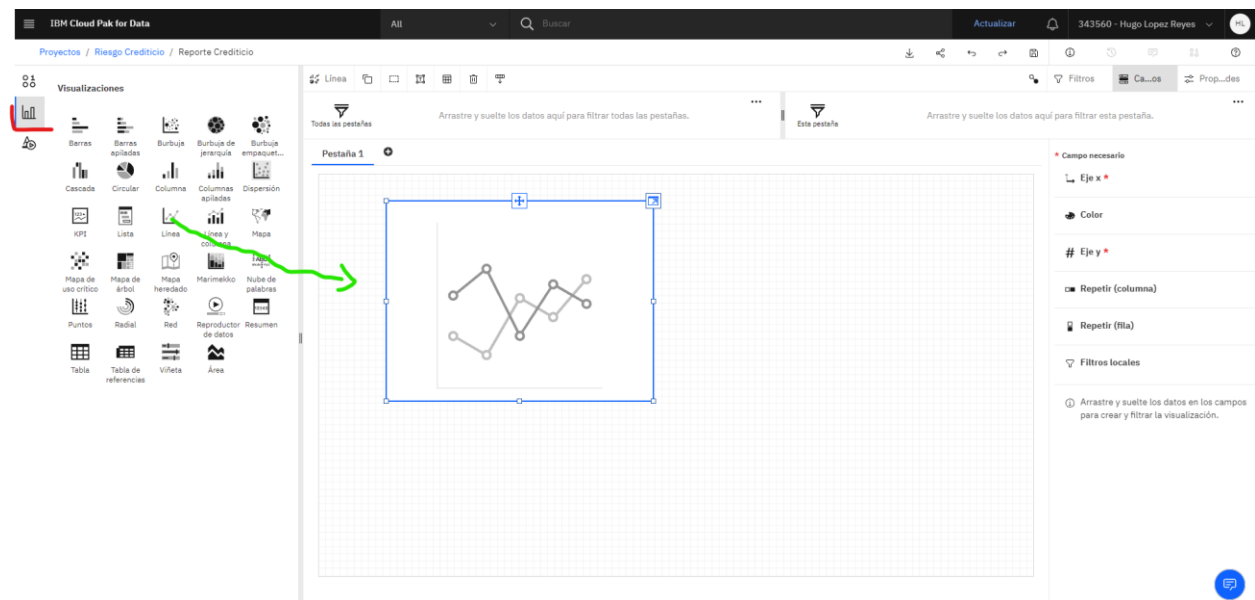
5.g) Dé clic en el símbolo de “+” (color rojo) en la sección de “Orígenes de Datos”.



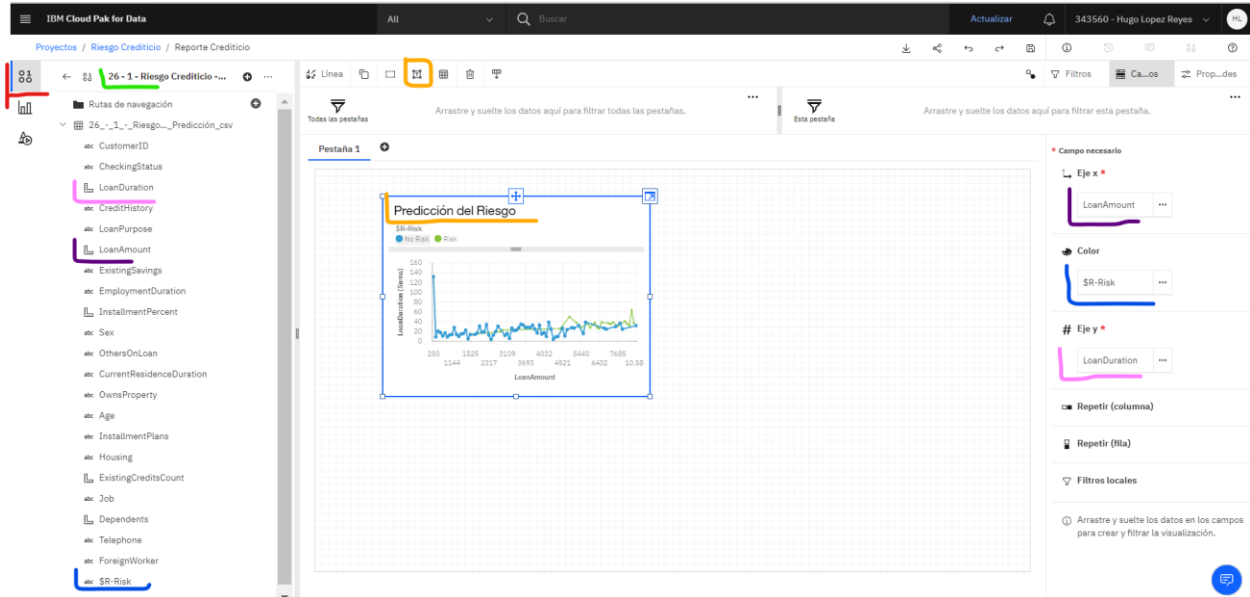
5.h) Seleccione “Activos de datos” (color rojo), luego seleccione el archivo “25 - 1 - Riesgo Crediticio - Predicción.csv” (color verde) y de clic en “Seleccionar” (color rosa, esquina inferior derecha)



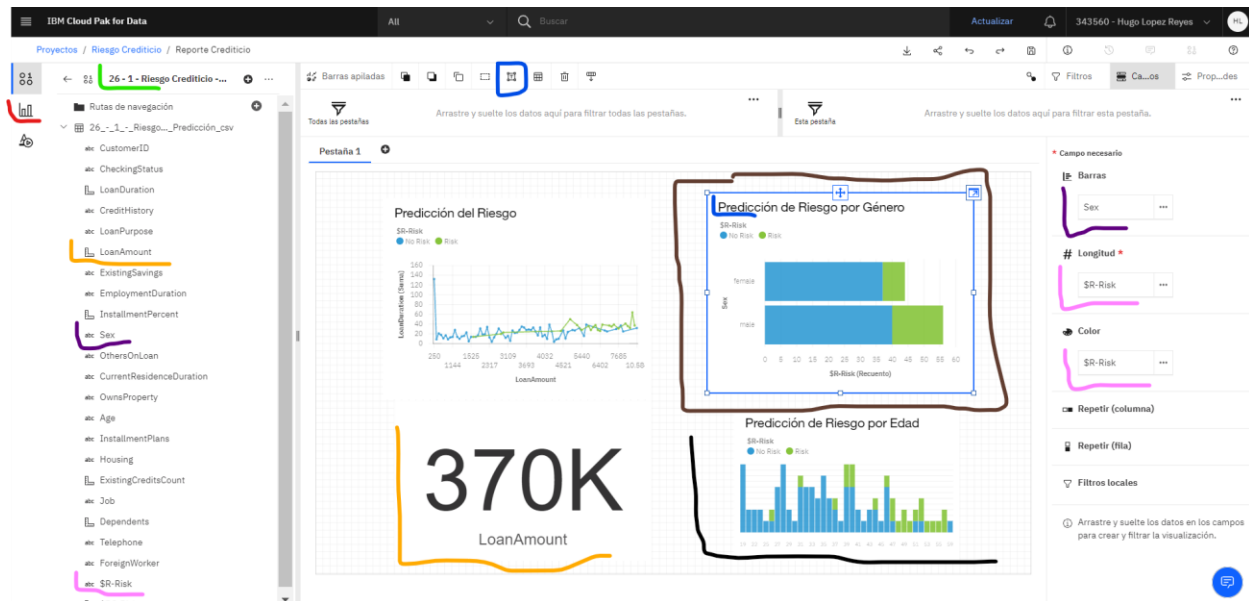
5.i) Seleccione los tipos de gráficas (color rojo), arrastre gráfica de línea al lienzo (flecha verde).



5.j) Seleccione las fuentes de datos (color rojo) y el archivo “25 - 1 - Riesgo Crediticio - Predicción” (color verde), 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), seleccione la opción “T” en la 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.k) 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 creo, 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) y si lo hizo todo bien verá la gráfica de barras que aparece abajo (color café). Ahora arrastre el indicador de “LoanAmount” (color naranja) bajo la primer gráfica que hizo y 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).



**¡Felicidades ya tiene listo su ambiente de demo, ahora debe practicar cómo hacer demostraciones con este ambiente, así como agregarle más servicios y funcionalidades, por ejemplo: intente hacer un experimento de AutoAI, agregue el servicio de Watson Knowledge Catalog, cree un flujo de datos con DataStage y mucho más!**



6. Finalmente regrese a su proyecto a la sección de “Activos” y verifique que su pantalla luzca como ésta y que tenga los mismos activos que se ven a continuación (sus tres fuentes de datos, su modelo y su tablero).

The screenshot displays the 'Activos' (Assets) section of the IBM Cloud Pak for Data interface. The top navigation bar includes the IBM Cloud Pak for Data logo, a search bar, and user information (343560 - Hugo Lopez Reyes). The main content area is divided into several sections:

- Activos de datos:** A table listing three data assets, all created by HUGO LOPEZ REYES on 11 oct 2021.
- Paneles de control:** A table listing one control panel, 'Reporte Crediticio', created by HUGO LOPEZ REYES on 11 oct 2021.
- Flujos del modelador:** A table listing one model flow, 'Riesgo Crediticio', created by HUGO LOPEZ REYES on 11 oct 2021.

The right sidebar contains a 'Datos' section with a message: 'Suelte aquí los archivos o explore para buscar los archivos que se han de cargar.'

Nombre	Tipo	Creador	Última modificación
CSV 25 - 1 - Riesgo Crediticio - Predicción.csv	Activo de datos	HUGO LOPEZ REYES	11 oct 2021 22:04
CSV 25 - 1 - Riesgo Crediticio - Histórico.csv	Activo de datos	HUGO LOPEZ REYES	11 oct 2021 15:57
CSV 25 - 1 - Riesgo Crediticio - Solicitantes.csv	Activo de datos	HUGO LOPEZ REYES	11 oct 2021 15:57

Nombre	Compartido	Último editor	Última modificación
Reporte Crediticio		HUGO LOPEZ REYES	11 oct 2021 22:03

Nombre	Tipo	Creador	Última modificación
Riesgo Crediticio	SPSS Modeler	HUGO LOPEZ REYES	11 oct 2021 22:03

**! HEMOS TERMINADO !**