

# Cloud Pak for Data as a Service

## Análisis de Riesgo Crediticio



# Análisis de Riesgo Crediticio

## Introducción:

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

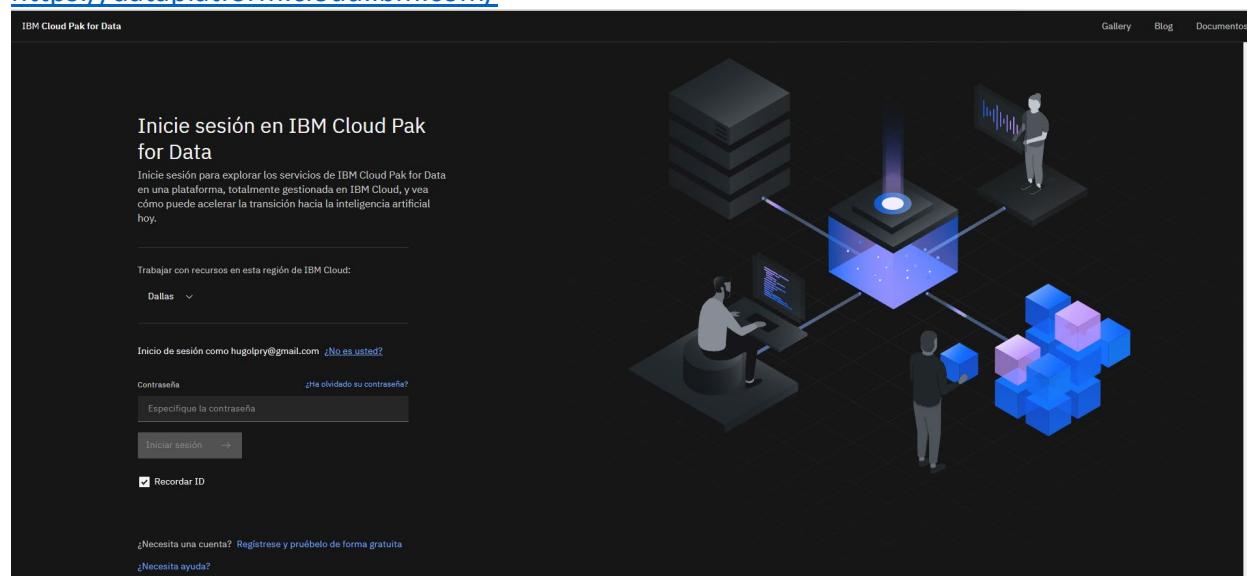
[https://github.com/peterwtf/Taller\\_RiesgoCrediticio](https://github.com/peterwtf/Taller_RiesgoCrediticio)

- Riesgo Crediticio - Histórico. Contiene información histórica de clientes que han solicitado créditos y su comportamiento.
  - Riesgo Crediticio – Solicitantes. Contiene información de personas que están solicitando un nuevo crédito y que decidiremos si se los otorgamos o no.
- b) Debe cargar sus datos en la plataforma (archivos del punto a)
- 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 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:

The screenshot shows the IBM Cloud Pak for Data interface. At the top, it says "¡Bienvenido, Pedro!". Below this, there's a "Quick start" sidebar with links to various services like "Build dashboards", "Create data pipelines", and "Build customer profiles". The main area has three columns: "Aprender mediante ejemplo", "Trabajar con datos", and "Amplie sus prestaciones". The "Visión general" section displays recent projects, catalogs, and notifications. A central graphic shows a computer monitor displaying a 3D cube cluster.

Proyectos recientes	Catálogos recientes	Notificaciones
Data Fabric Outcomes 07 abr 2022 21:51	CreditDataCatalog 22 feb 2022 14:48	Project import complete Data Fabric Outcomes was imported successfully. View import summary.
peter 07 abr 2022 11:03	Platform assets catalog 30 sept 2021 13:48	Project import complete Data Fabric Outcomes was imported successfully. View import summary.
IA Taller 30 mar 2022 16:54		Project import complete Data Fabric Outcomes was imported successfully. View import summary.
Jobs_Puebas 17 mar 2022 13:24		Project import complete Data Fabric Outcomes was imported successfully. View import summary.
Data Virtualization 09 mar 2022 12:41		Project import complete Data Fabric Outcomes was imported successfully. View import summary.

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”

This screenshot shows the navigation menu expanded. The "Catálogo de servicios" option is highlighted with a red box. The main content area is identical to the previous screenshot, showing the "Visión general" section with recent projects, catalogs, and notifications.

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).

Catálogo de servicios

Almacenamiento

Cloud Object Storage

Almacenamiento de datos habilitado para desarrolladores para aplicaciones en la nube integradas con servicios d...

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).

Cloud Object Storage

Crear

Elegir una infraestructura

IBM Cloud

Satellite

Selecciónar un plan de precios

Planificar

Características

Precios

Lite

Standard

Crear

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

Instancias de servicio

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

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 dashboard. On the left, a sidebar menu is open with several options: Inicio, Datos, Proyectos, Catalogos, Gobernabilidad, Despliegues, Servicios, Catalogo de servicios (which is highlighted with a red box), Instancias de servicio, Gallery, Administración, and Soporte. The main content area has sections like 'Trabajar con datos' and 'Amplie sus prestaciones'. Below these are 'Visión general' cards for 'Proyectos recientes' (Data Fabric Outcomes, 07 abr 2022), 'Catálogos recientes' (CreditDataCatalog, 22 feb 2022), and 'Notificaciones' (Project import complete). At the top right, there are links for 'Comprar', 'Pedro Angel Miranda's Acc...', and a profile icon.

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).

This screenshot shows the 'Catálogo de servicios' page. A search bar at the top contains the text 'Q. Watson Studio'. Below it, a sidebar lists categories: Categoría, IA / Aprendizaje automático, bases de datos, Integración, Análiticas, and Almacenamiento. The main area displays a card for 'Watson Studio' with a green border, indicating it is selected. The card includes the service name, category, a brief description ('Integre la IA y el aprendizaje automático en su negocio. Cree modelos personalizados utilizando su...'), and a 'Lite • Gratuito' 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).

This screenshot shows the configuration page for the 'Watson Studio' service. At the top, it says 'Watson Studio' and 'Autor: IBM • Fecha de última actualización: Mar 31, 2022 • Documentos'. Below this are tabs for 'Crear' (selected) and 'Acerca de'. The 'Crear' tab has a sub-section 'Seleccionar una región' where 'Dallas' is selected. The 'Resumen' section on the right shows details: Watson Studio, Región: Dallas, Plan: Lite, Nombre de servicio: Watson Studio-fo, Grupo de recursos: Default. The 'Plan de precios' section shows two plans: 'Lite' (selected) and 'Professional'. The 'Lite' plan is highlighted with a red box and includes a detailed description of its features and pricing (\$0.98 USD/Capacity Unit-Hour). The 'Professional' plan is shown below. At the bottom right, there is a blue 'Crear' button with a green border.

2.h) Si el servicio de Watson Studio fue creado correctamente (color rojo) 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 message about updating a service plan. Below it, a search bar and filter options (Group of resources, Locations, None, Product, Service Plan) are visible. A large blue button labeled 'Añadir servicio' (+) is prominent. The main table lists two service instances:

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

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

The screenshot shows the main dashboard of IBM Cloud Pak for Data. On the left, a sidebar navigation menu is open, with the 'Catalogo de servicios' option highlighted with a red box. The main content area features a dark background with a central graphic of a computer monitor displaying a chart, surrounded by 3D cubes. Text on the screen includes 'Buenos días, Pedro!', 'Trabajar con datos', 'Amplíe sus prestaciones', 'Crear un proyecto', 'Crear un servicio', and 'Visión general'. Below this are sections for 'Proyectos recientes', 'Catálogos recientes', and 'Notificaciones'.

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

The screenshot shows the 'Catálogo de servicios' (Service Catalog) page. A search bar at the top contains the query 'Cognos Dashboard E...'. The results are displayed in a grid. One service, 'IBM Cognos Dashboard Embedded' under the 'Analíticas' category, is highlighted with a green box. The service card includes an icon, the name, a brief description, and a 'Lite • Gratuito' status indicator.

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).

The screenshot shows the IBM Cloud Pak for Data interface. On the left, a sidebar lists categories like Inicio, Datos, Proyectos, Catalogos, Gobernabilidad, Despliegues, Servicios, and Catalogo de servicios. The Catalogo de servicios option is highlighted with a red box. The main content area displays the details for creating an IBM Cognos Dashboard Embedded service. It includes a dropdown for 'Seleccionar una región' with 'Dallas' selected, a table for 'Plan de precios' comparing 'Lite' and 'Pago por uso' plans, and a large blue 'Crear' button at the bottom right.

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

The screenshot shows the 'Instancias de servicio' (Service Instances) page. It lists three instances: 'CloudObjectStorage', 'IBM Cognos Dashboard Embedded-db', and 'Watson Studio - lab'. The 'IBM Cognos Dashboard Embedded-db' instance is highlighted with a red box. The 'Añadir servicio' (Add Service) button is visible at the top right of the table.

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

The screenshot shows the main dashboard of IBM Cloud Pak for Data. The left sidebar has 'Servicios' selected and 'Catalogo de servicios' highlighted with a red box. The central area features a 'Visión general' (General View) with sections for 'Proyectos recientes' (Recent Projects), 'Catalogos recientes' (Recent Catalogs), and 'Notificaciones' (Notifications). A large graphic on the right depicts data flow and analysis.

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).

Catálogo de servicios

Q. DataStage x bases de datos

Categoría

- IA / Aprendizaje automático
- bases de datos
- Integración
- Análiticas
- Almacenamiento

**DataStage**

Analíticas • bases de datos • Integración

IBM® DataStage® ofrece integración de datos líder en tiempo real y por lotes para crear canalizaciones de datos de...

Lite • Gratuito

Integración

**DataStage**

Analíticas • bases de datos • Integración

IBM® DataStage® ofrece integración de datos líder en tiempo real y por lotes para crear canalizaciones de datos de...

Lite • Gratuito

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

IBM Cloud Pak for Data

Catálogo de servicios /

**DataStage**

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

**Crear** Acerca de

Seleccionar una región

Dallas

Resumen

Region: Dallas  
Plan: Lite  
Nombre de servicio: DataStage-1b  
Grupo de recursos: Default

Plan de precios

El plan Lite de IBM DataStage permite a los usuarios acceder a las prestaciones de DataStage para un máximo de 15 CUH.

Características

Plan	Características	Precios
Estándar	Plano estándar de IBM DataStage Entorno: 1 VCPU + 4 GB RAM = 1 CUH • 1 VCPU + 8 GB RAM = 2 CUH • 2 VCPU + 8 GB RAM = 4 CUH • 4 VCPU + 16 GB RAM = 8 CUH • 8 VCPU + 32 GB RAM = 8 CUH • 16 VCPU + 64 GB RAM = 16 CUH	\$1.75 USD/Capacity Unit-Hour
Small Enterprise Bundle	Fully managed, serverless data integration with readily scalable DataStage PX engine with 5000 CUH included • Bulk purchase with same functionality as standard plan	\$7900.00 USD/Instance \$1.75 USD/Capacity Unit-Hour
Medium Enterprise Bundle	Fully managed, serverless data integration with readily scalable DataStage PX engine with 10000 CUH included • Bulk purchase with same functionality as standard plan	\$14000.00 USD/Instance \$1.75 USD/Capacity Unit-Hour
Large Enterprise Bundle	Fully managed, serverless data integration with readily scalable DataStage PX engine with 25000 CUH included • Bulk purchase with same functionality as standard plan	\$30600.00 USD/Instance \$1.75 USD/Capacity Unit-Hour

El plan Lite de IBM DataStage permite a los usuarios acceder a las prestaciones de DataStage para un máximo de 15 CUH.

Plan Lite de IBM DataStage

Los servicios del plan Lite se suprimen tras 30 días de inactividad.

Crear

Ver términos

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

The screenshot shows the 'Instancias de servicio' (Instances of service) page in the IBM Cloud Pak for Data interface. At the top, there is a search bar and filter options: 'Filter by: Grupos de recursos', 'Ubicaciones', 'Ninguna', 'Producto' (with 'DataStage' selected), and 'Plan de servicio'. A blue button labeled 'Añadir servicio +' is visible. Below the filters, a table lists four services:

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 main navigation menu of the IBM Cloud Pak for Data interface. On the left, under 'Filtrar navegación', the 'Catálogo de servicios' option is highlighted with a red box. Other menu items include 'Inicio', 'Datos', 'Proyectos', 'Catálogos', 'Gobernabilidad', 'Despliegues', 'Servicios', and 'Gallery', 'Administración', 'Soporte'. The central area displays a 'Visión general' (General view) with sections for 'Proyectos recientes' (Recent projects), 'Catálogos recientes' (Recent catalogs), and 'Notificaciones' (Notifications). A callout box highlights the 'Catálogo de servicios' link.

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

The screenshot shows the 'Catálogo de servicios' (Catalog of services) page. In the search bar at the top left, 'Machine Learning' is typed. Below the search bar, a list of categories is shown: 'IA / Aprendizaje automático', 'bases de datos', 'Integración', 'Analytics', and 'Almacenamiento'. Under 'IA / Aprendizaje automático', a service card for 'Machine Learning' is highlighted with a green box. The card includes the service name, a brief description: 'Despliegue, gestione e integre modelos de aprendizaje automático en sus aplicaciones y servicios en tan solo u...', and a status: 'Lite • Gratuito'. The entire catalog page has a light gray background.

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

Catálogo de servicios /

## Machine Learning

Autor: IBM SPSS • Fecha de última actualización: Apr 4, 2022 • Documentos • Documentos de API

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

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

**Crear**

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

IBM Cloud Pak for Data

Buscar en los espacios de trabajo

Instancias de servicio

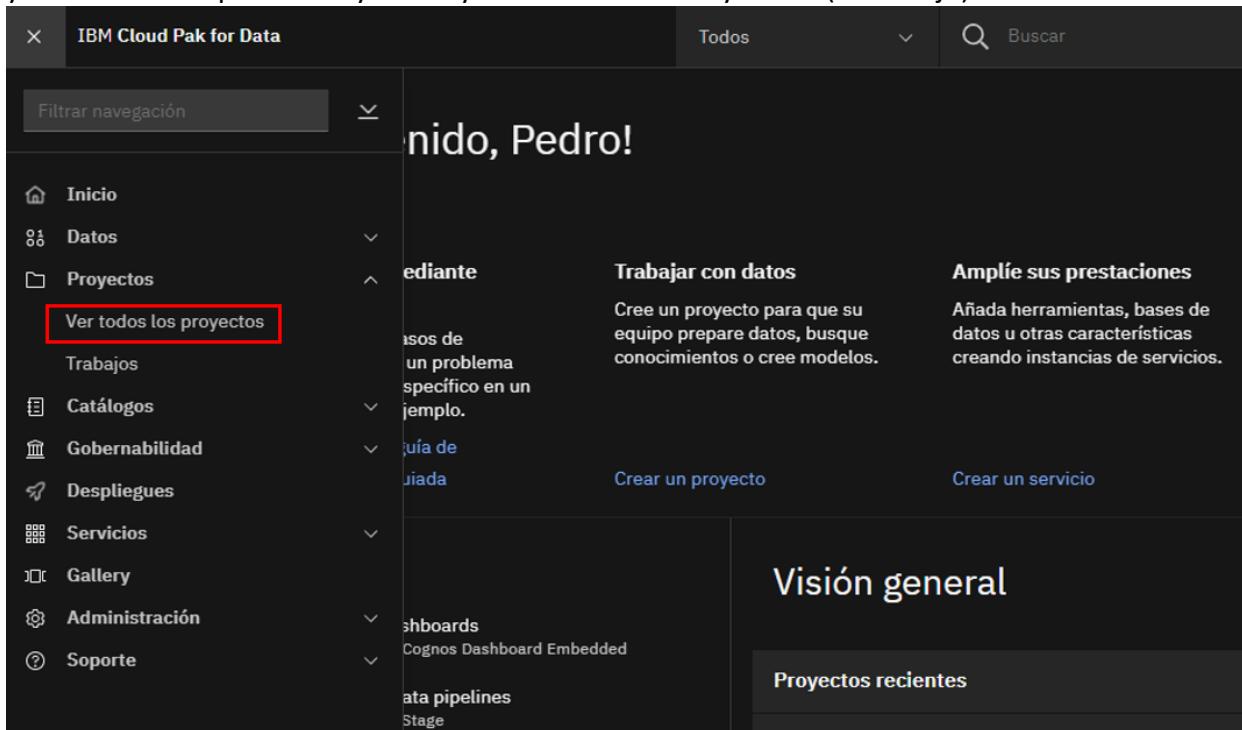
Para actualizar un plan de servicio, en primer lugar **actualice** su cuenta de IBM Cloud. A continuación, seleccione **Actualizar servicio** o **Gestionar en IBM Cloud** desde el menú en la fila del servicio.

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

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

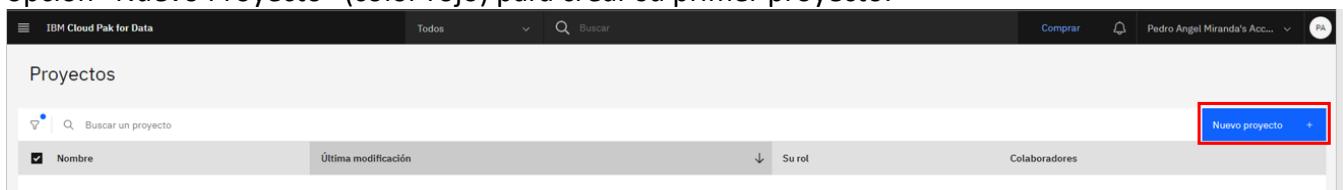
**“Felicitaciones ha habilitado todos los servicios que estaremos utilizando” para crear esta demostración”.**

3. 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).



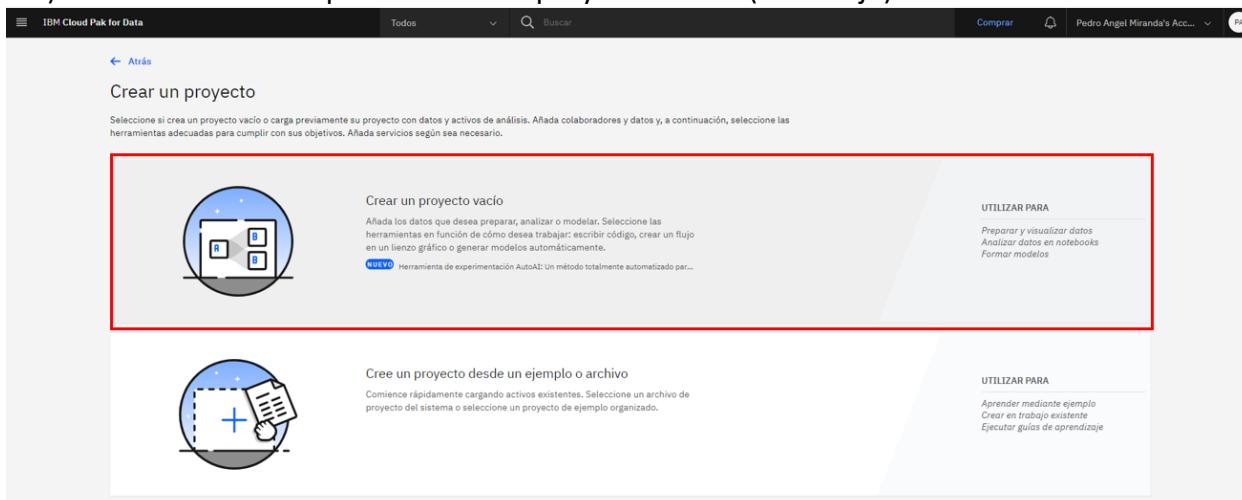
The screenshot shows the main dashboard of IBM Cloud Pak for Data. On the left, there's a navigation sidebar with various options like Inicio, Datos, Proyectos, Catálogos, Gobernabilidad, Despliegues, Servicios, Gallery, Administración, and Soporte. The 'Proyectos' option is expanded, and its sub-item 'Ver todos los proyectos' is highlighted with a red box. The main content area features sections for 'Trabajar con datos' (with a callout 'Amplíe sus prestaciones') and 'Visión general'. A 'Proyectos recientes' section is also visible.

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



This screenshot shows the 'Proyectos' page. At the top right, there's a blue button labeled 'Nuevo proyecto' with a '+' sign, which is highlighted with a red box. The page includes a search bar, filter options for 'Nombre' and 'Última modificación', and a 'Su rol' dropdown. Below the header, there's a table-like structure showing project details and a 'Colaboradores' section.

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



This screenshot shows the 'Crear un proyecto' (Create Project) page. It has two main sections: 'Crear un proyecto vacío' (Create Empty Project) and 'Cree un proyecto desde un ejemplo o archivo' (Create from Example or File). The 'Crear un proyecto vacío' section is highlighted with a red box. It contains a circular icon with a house-like symbol and some text about preparing data for analysis. Below it, another section 'Cree un proyecto desde un ejemplo o archivo' is shown with a similar icon and text about loading existing assets or selecting an example project. Both sections have 'UTILIZAR PARA' (Use For) sub-sections with options like 'Preparar y visualizar datos' (Prepare and visualize data), 'Analizar datos en notebooks' (Analyze data in notebooks), 'Formar modelos' (Train models), 'Aprender mediante ejemplo' (Learn via example), 'Crear en trabajo existente' (Create in existing workspace), and 'Ejecutar guías de aprendizaje' (Run learning guides).

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):

Nuevo proyecto

Definir detalles

Nombre  
RiesgoCrediticio

Descripción  
Riesgo Crediticio

Almacenamiento  
CloudObjectStorage

Seleccionar opciones del proyecto

Restringir quién puede ser un colaborador ⓘ

El proyecto incluye la integración con [Cloud Object Storage](#) para almacenar activos de proyecto.

Cancelar **Crear**

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).

Proyectos / RiesgoCrediticio

Visión general **Activos** Trabajos Gestionar

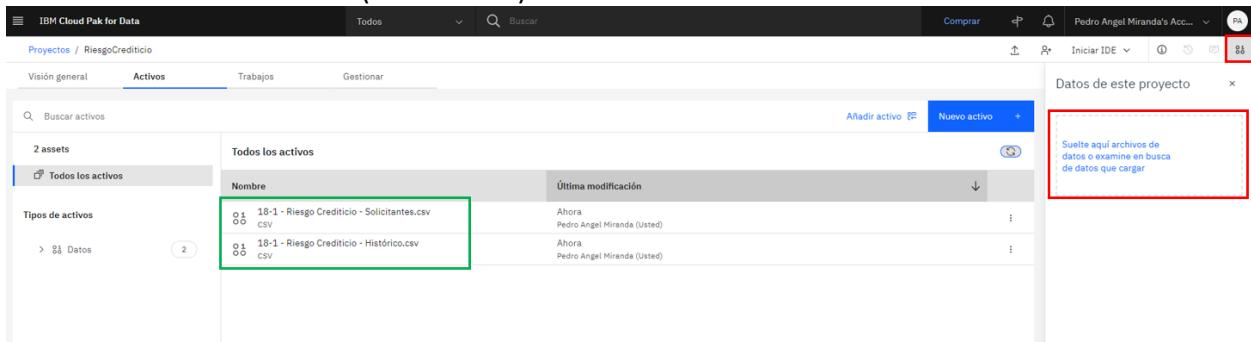
Activos	Uso de recursos	Historial del proyecto
Añada datos en la página Activos. 	Este proyecto 0 CUH	Sin notificaciones Verá sus notificaciones más recientes aquí.
<a href="#">Ver todo</a>	<a href="#">Archivo Readme</a>	

Documento su proyecto utilizando la sintaxis Markdown. Consulte la [<urllink>Hoja de apuntes de Markdown</urllink>](#).

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)



The screenshot shows the IBM Cloud Pak for Data interface. The top navigation bar includes 'IBM Cloud Pak for Data', 'Proyectos / RiesgoCrediticio', 'Todos', 'Buscar', and a user profile 'Pedro Angel Miranda's Acc...'. Below the navigation is a toolbar with icons for 'Comprar', 'Iniciar IDE', and others. The main area has tabs for 'Visión general', 'Activos' (which is selected), 'Trabajos', and 'Gestionar'. Under 'Activos', there's a search bar 'Buscar activos' and a section titled 'Todos los activos' showing '2 assets'. A green box highlights the list of files: '18-1 - Riesgo Crediticio - Solicitantes.csv' and '18-1 - Riesgo Crediticio - Histórico.csv', both listed under 'Nombre' and 'Última modificación' (Both files were modified 'Ahora' by 'Pedro Angel Miranda (Usted)'). To the right, a sidebar titled 'Datos de este proyecto' contains a red box around the text 'Suelte aquí archivos de datos o examine en busca de datos que cargar'.

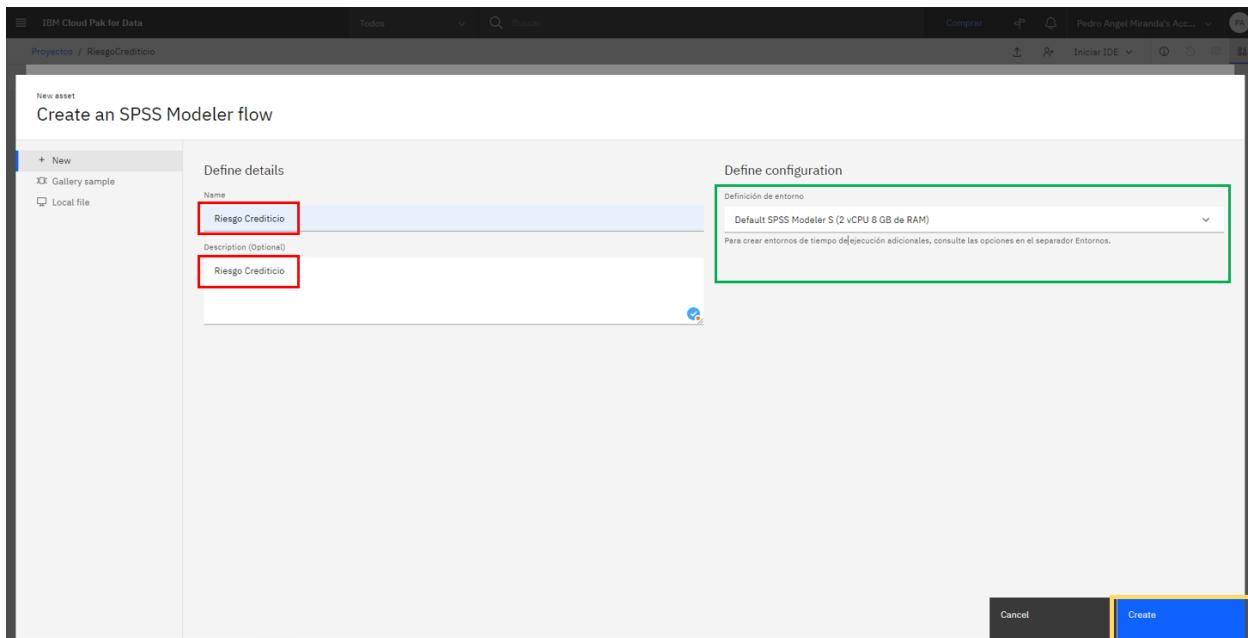
**“Felicitaciones ya creo su proyecto y descargó las dos fuentes de datos que va a utilizar para su modelo”.**

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

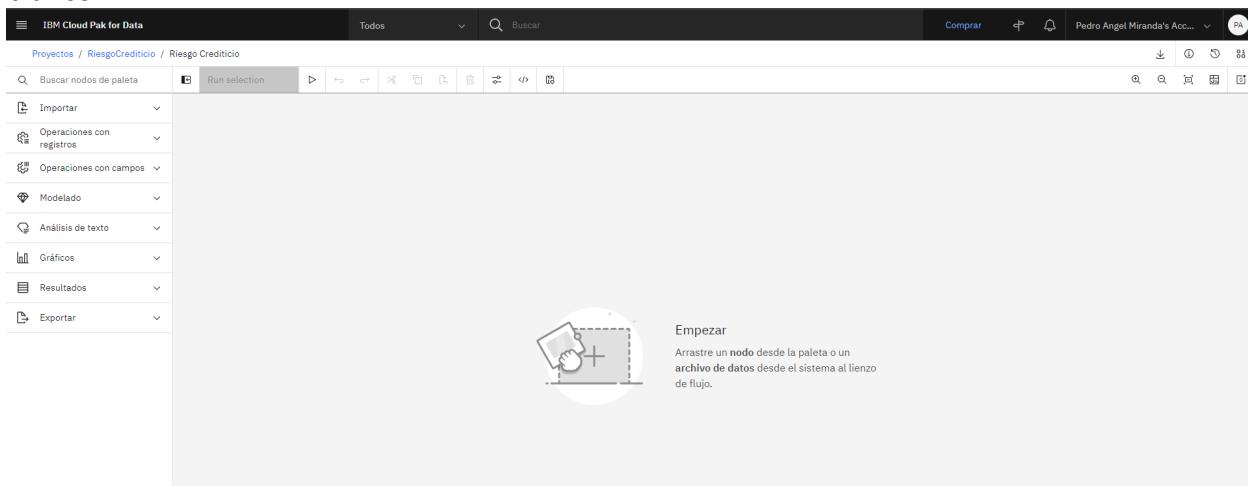
y a continuación la opción “SPSS Modeler” (color verde).

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

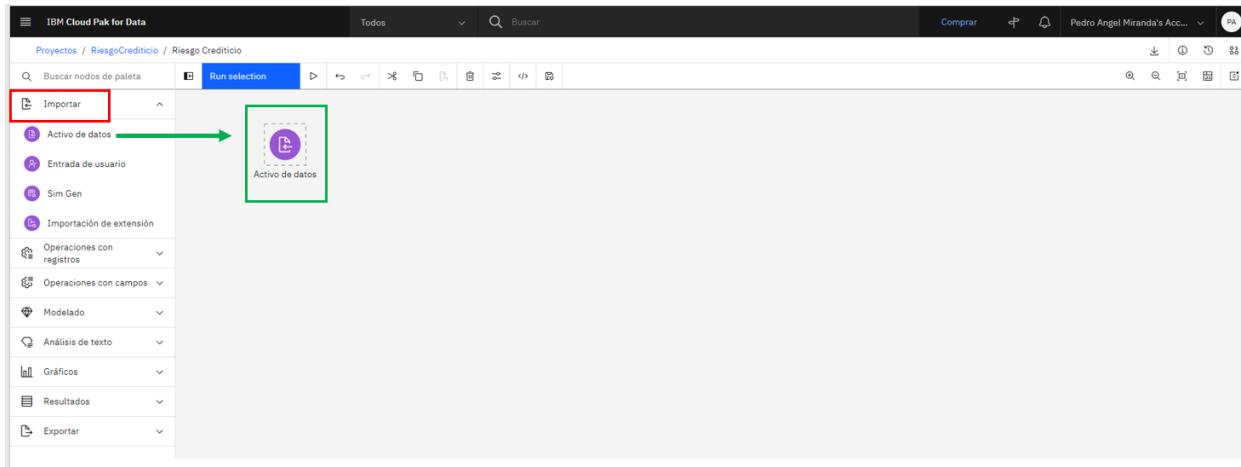
**Nota:** Para fines prácticos del Hackathon siempre utilizar Default SPSS Modeler (2vCPU 8GB RAM)



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

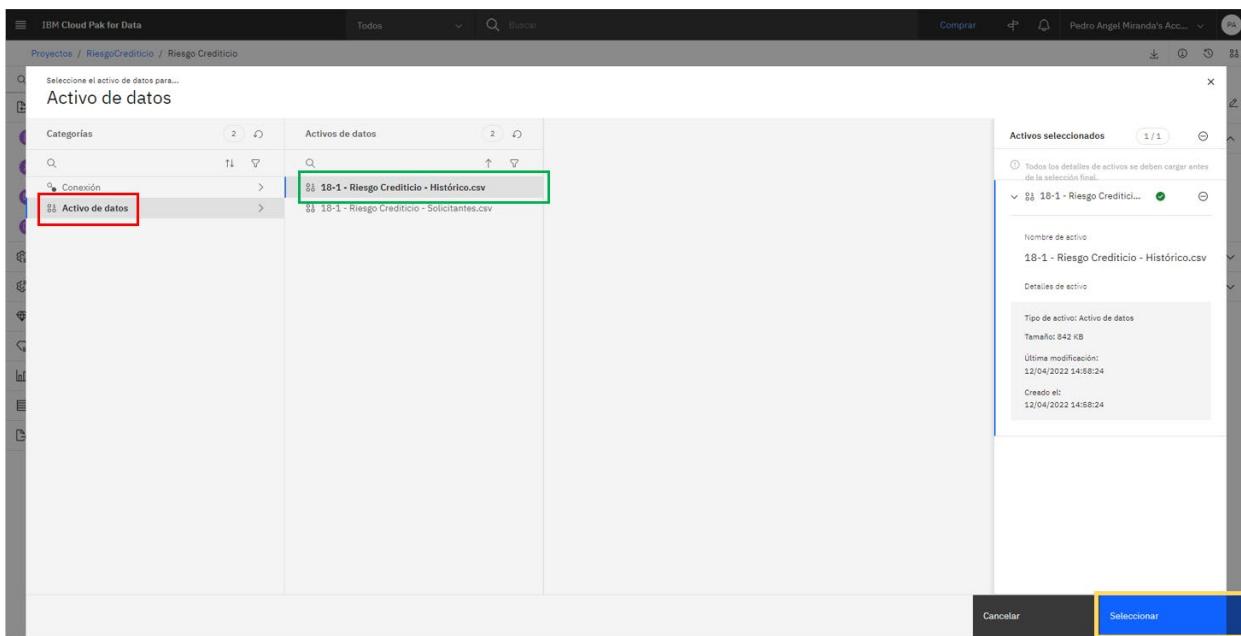


4.c) A continuación vamos a **crear el flujo de entrenamiento del modelo**, para eso debe ir a la opción de “Importar” (color rojo) y arrastrar hacia el lienzo el nodo “Activo de Datos” (flecha verde) y quedará como sigue:



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

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

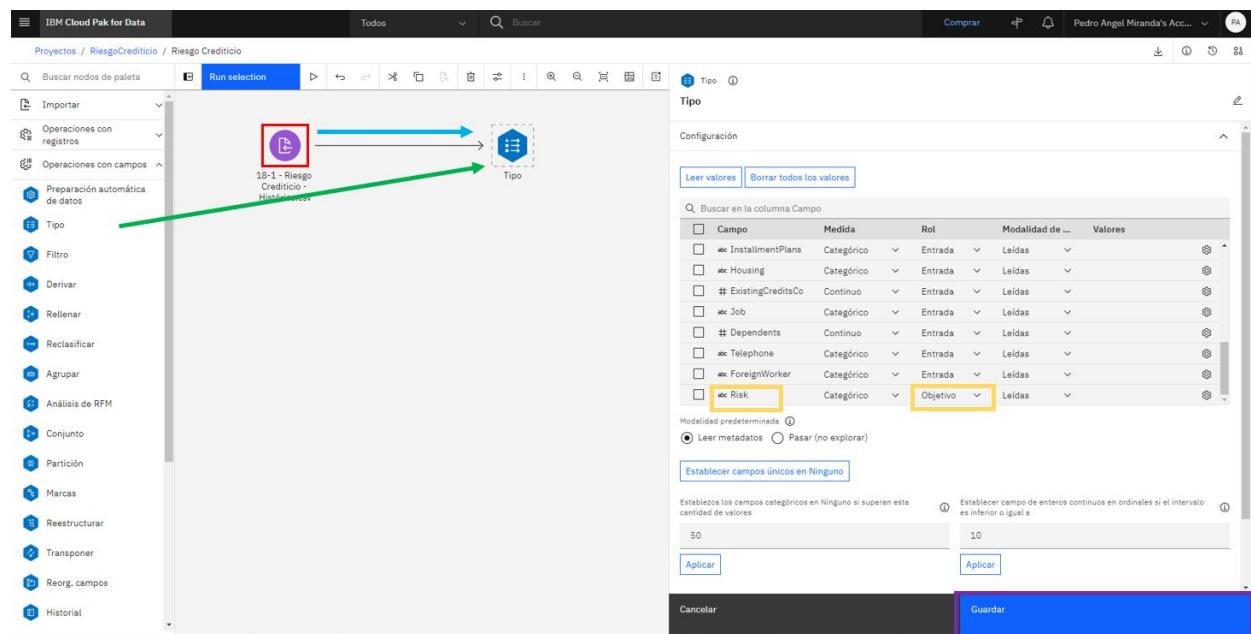


4.e) Si hizo todo bien regresará al lienzo donde tendrá el ícono de “Activo de datos”, pero habrá sido renombrado con el nombre de la fuente de datos que le acaba de asociar (color rojo).

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

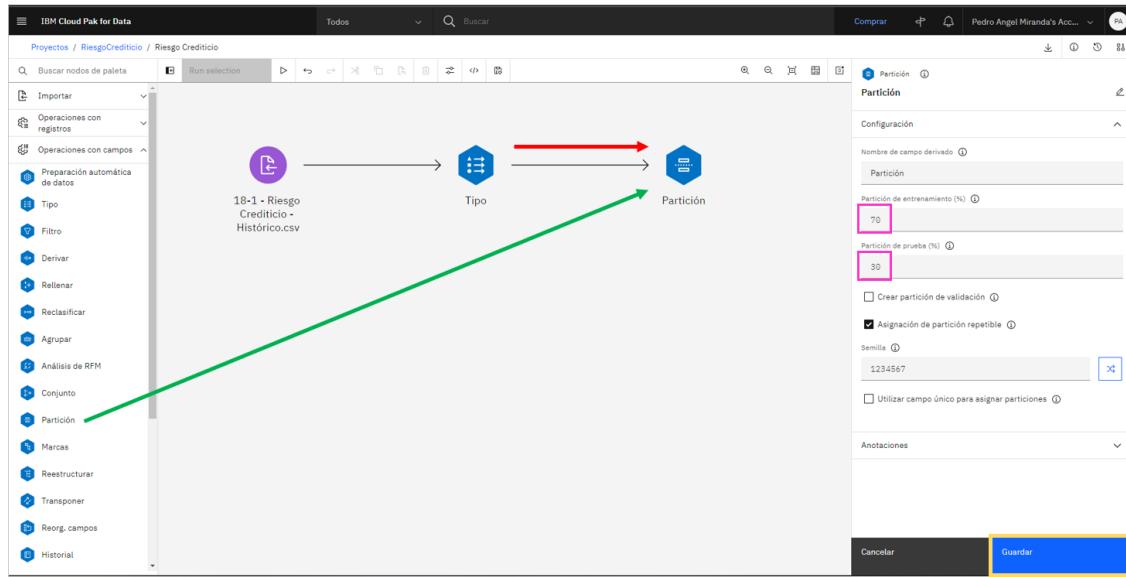
Ahora abra el ícono de “Tipo” (botón derecho y opción “Abrir”), aparecerá una nueva pantalla a un lado donde estarán todos los campos de la fuente de datos, busque el que se llama “Risk” (color amarillo, 1er columna) cambie el rol de ese campo a “Objetivo” (3<sup>a</sup> columna, color amarillo).

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



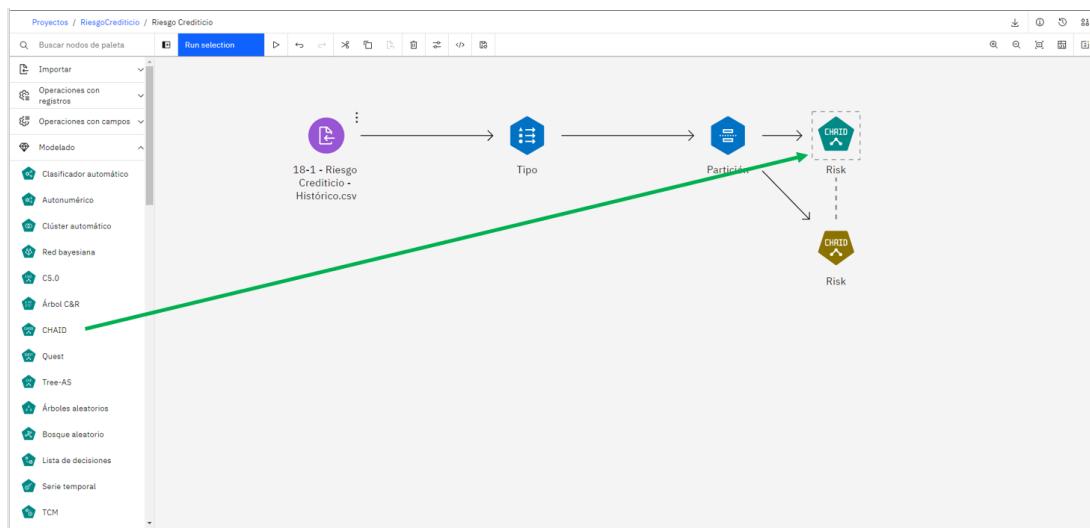
4.f) A continuación en “Operaciones con Campos” seleccione el nodo “Partición” y arrástrelo al lienzo (flecha verde), conecte el nodo “Tipo” al nodo “Partición” (flecha roja).

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



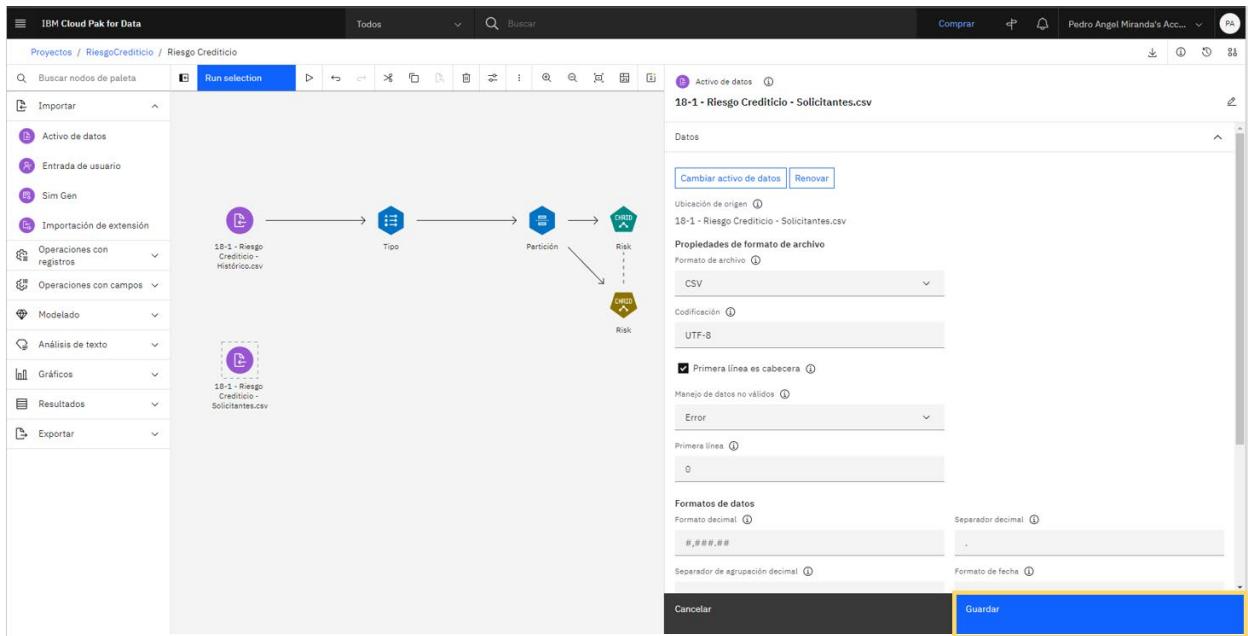
4.g) Ahora vaya a las opciones de la derecha y busque la sección de “Modelado” y arrastrar el ícono “CHAID” al lienzo (flecha Verde), conectar el nodo “Partición” con el nodo “CHAID”

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



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

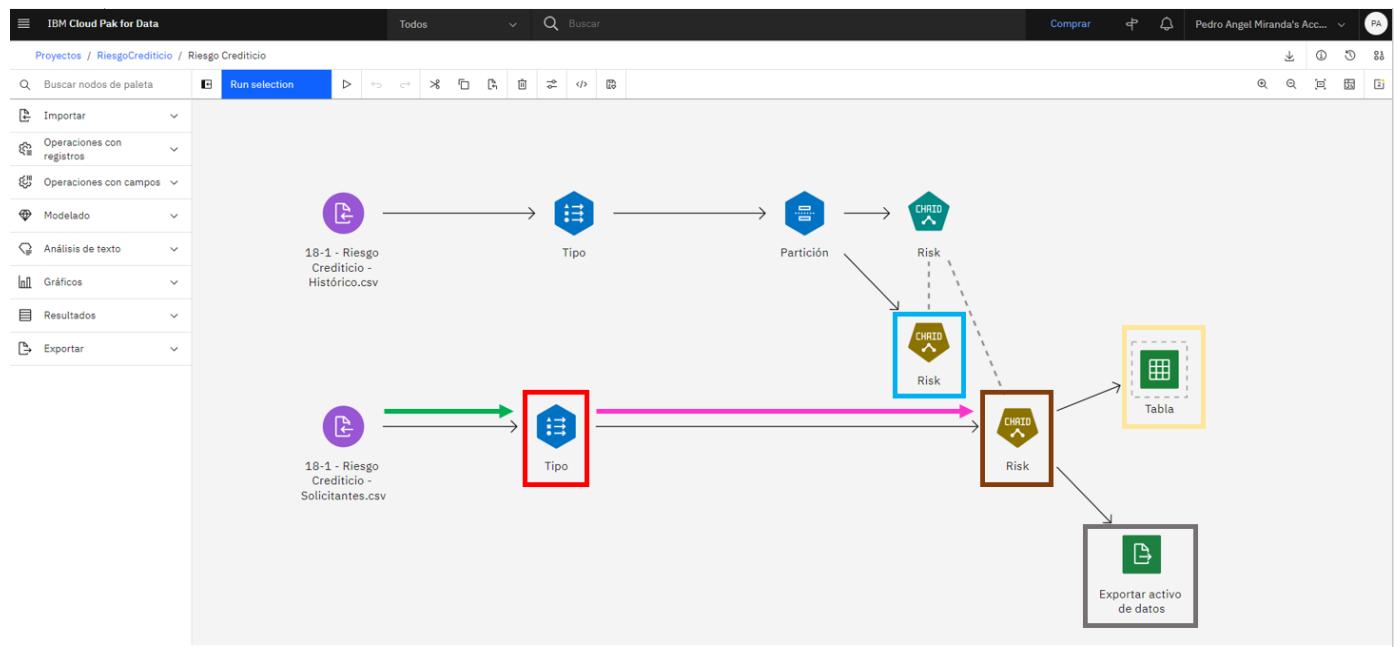
Con éste nuevo ícono de “Activo de Datos” **estamos iniciando la creación del flujo de predicción.**



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

- Arrastre un nodo “Tipo” (“Operaciones con campos”/”Tipo”, color rojo)
- conecte el nodo “Activo de Datos” con el nuevo nodo “Tipo” (color verde)
- dé clic sobre el nodo “CHAID dorado” (color azul) para tenerlo seleccionado y a continuación usa las teclas Ctrl-C y Ctrl-V para crear una copia de ese nodo
- conecte el nodo “Tipo” (color rojo) con la copia del nodo (Flecha rosa) “CHAID” dorado (color café)
- a continuación, arrastre un nodo de “Tabla” (“Resultados”/”Tabla”) al lienzo (color amarillo) y un nodo de “Exportar Activo de Datos” (“Exportar”/”Exportar Activo de Datos”) al lienzo (color gris)
- abra el nodo “Exportar Activo de Datos” (color gris) y cambie el “Nombre de archivo” de “output.csv” a “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 “\$SRC-Risk” (color rojo) que se generan como resultado de correr el modelo de predicción.

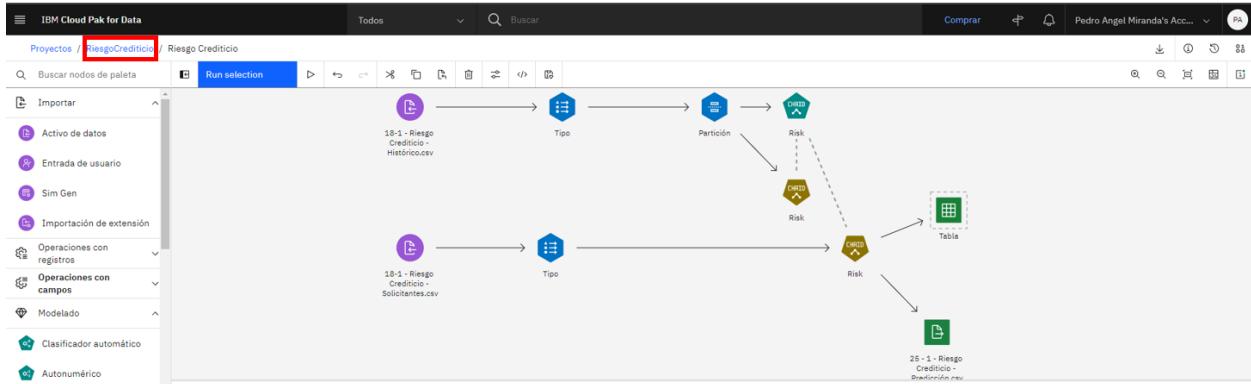
The screenshot shows the IBM Cloud Pak for Data interface. On the left, there's a sidebar with various options like Importar, Operaciones con registros, Operaciones con campos, Modelado, Análisis de texto, Gráficos, Resultados, and Exportar. The main area displays a flow diagram with nodes: 'Importar' (purple circle), '19-1 - Riesgo Credítico - Histórico.csv' (blue hexagon labeled 'Tipo'), 'Partición' (yellow diamond labeled 'Risk'), '25 - 1 - Riesgo Credítico - Predicción.csv' (yellow diamond labeled 'Risk'), and 'Tabla' (green grid icon). A context menu is open over the second 'Risk' node, with 'Vista previa' highlighted. Below the preview, a table titled 'Preview: 25 + 1 ...' shows data with columns: 'holdersOnLoan', 'CurrentResidenceDuration', 'OwnsProperty', 'Age', 'InstallmentPlans', 'Housing', 'ExistingCreditsCount', 'Job', 'Dependents', 'Tel', 'Nuevo comentario', '\$R-Risk', and '\$SRC-Risk'. The '\$R-Risk' and '\$SRC-Risk' columns are highlighted with red boxes.

holdersOnLoan	CurrentResidenceDuration	OwnsProperty	Age	InstallmentPlans	Housing	ExistingCreditsCount	Job	Dependents	Tel	Nuevo comentario	\$R-Risk	\$SRC-Risk
ne	3	savings_insurance	32	none	own	1	skilled	1	no	Preview SQL pushback	0.669	
ne	2	savings_insurance	37	stores	own	2	skilled	1	no	Ejecutar	0.896	
ne	3	real_estate	28	none	own	2	skilled	1	yes	no	No Risk 0.878	
ne	2	savings_insurance	32	none	own	1	skilled	1	none	yes	No Risk 0.640	
ne	3	unknown	57	none	own	2	skilled	1	none	yes	Risk 0.716	
-applicant	5	unknown	57	none	free	2	skilled	2	yes	yes	Risk 0.950	

¡Felicitaciones su modelo está listo!

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

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



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

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

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

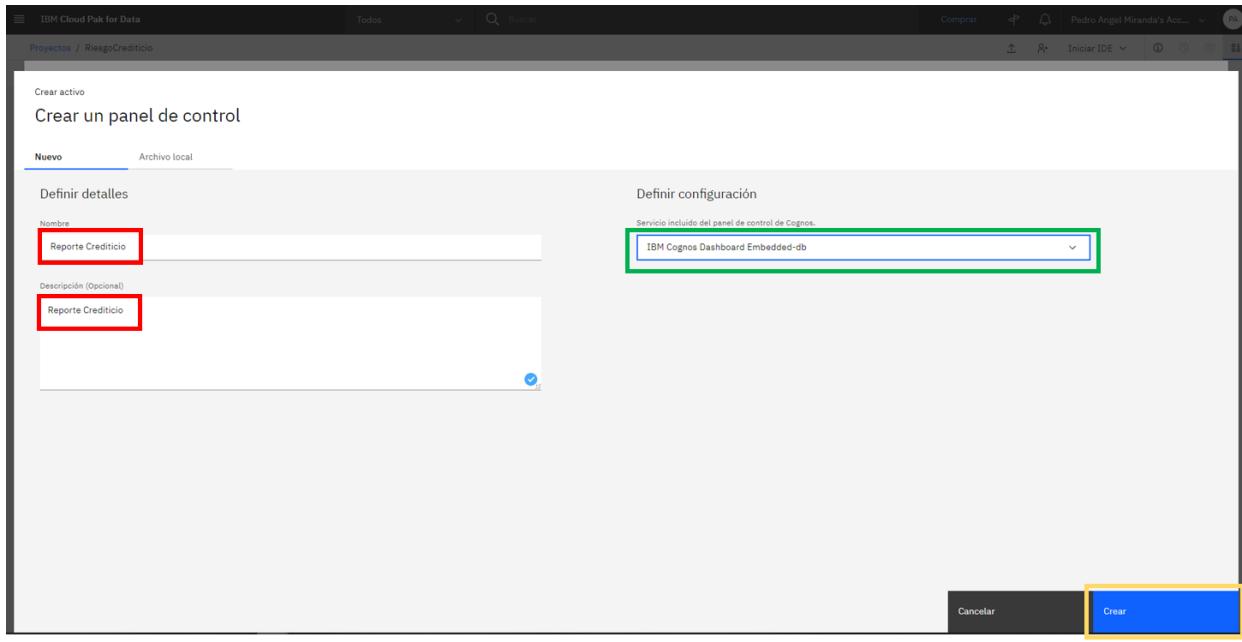
Si todo sale bien vera una paloma verde en la parte de “Estado” y si servicio estara asociado al proyecto

Nombre	Tipo	Plan	Ubicación	Estado	Grupo
Watson OpenScale-op	Watson OpenScale	Lite	Dallas	No asociado	Default
Data Replication-kd	Data Replication	Lite	Global	No asociado	Default
IBM Cognos Dashboard Embedded-db	IBM Cognos Dashboard Embedded	Lite	Dallas	Asociado	Default
WatsonMachineLearning	Machine Learning	Lite	Dallas	No asociado	Default

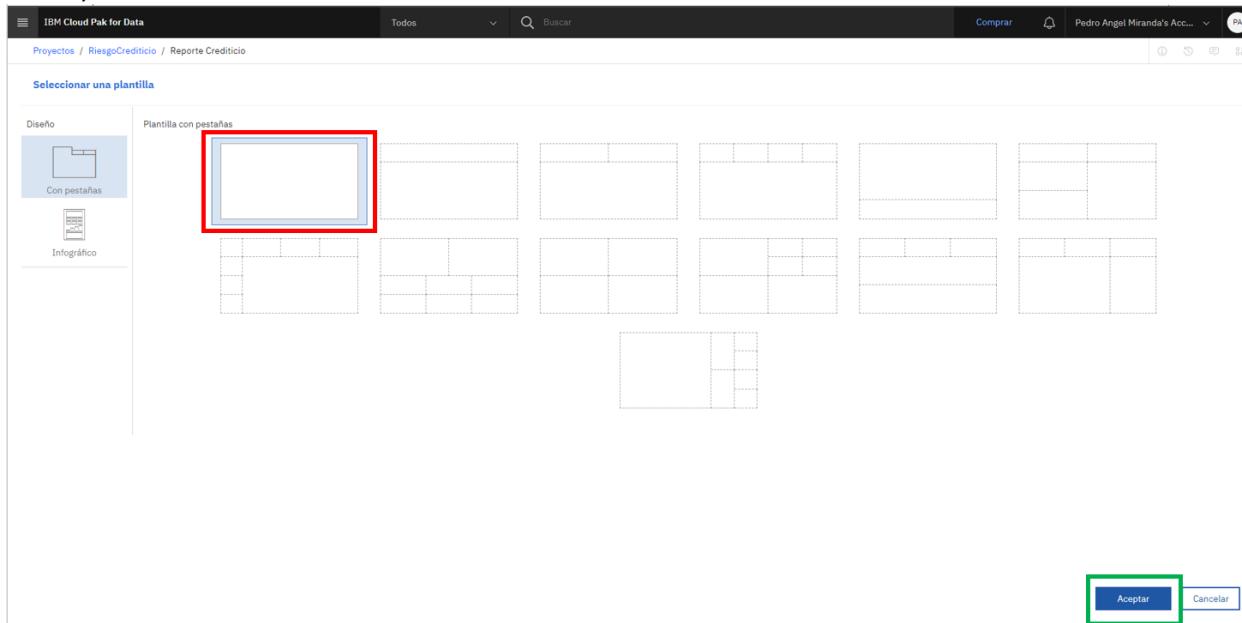
5.a) Regrese a la parte de “Activos” dé clic al botón azul “Nuevo activo” (color rojo)

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”.

The screenshot shows the 'IBM Cloud Pak for Data' interface. In the top navigation bar, there are links for 'Proyectos / RiesgoCrediticio / Reporte Crediticio', a search bar, and user information. Below the navigation is a toolbar with various icons. The main area is titled 'Orígenes seleccionados'. On the left, there's a sidebar with a '+ Add' button. The main panel has two tabs: 'Todas las pestañas' and 'Esta pestaña'. The 'Esta pestaña' tab is active, showing a large empty grid area with the placeholder text 'Arrastre y suelte los datos aquí para filtrar todas las pestañas.' and 'Arrastre y suelte los datos aquí para filtrar esta pestaña.' A red box highlights the small blue '+' icon located at the top right of the 'Orígenes seleccionados' list area.

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)

This screenshot shows a modal dialog titled 'Seleccionar origen de datos'. It contains two main sections: 'RiesgoCrediticio' on the left and 'Activos de datos' in the center. The 'Activos de datos' section lists several files: '18-1 - Riesgo Crediticio - Histórico.csv', '18-1 - Riesgo Crediticio - Solicitantes.csv', and '25 - 1 - Riesgo Crediticio - Predicción.csv'. The file '25 - 1 - Riesgo Crediticio - Predicción.csv' is highlighted with a green box. The left sidebar shows a tree structure with 'Conexión' expanded, and 'Activo de datos' is highlighted with a red box. In the bottom right corner of the dialog, there are two buttons: 'Cancelar' and 'Seleccionar', with 'Seleccionar' highlighted by a yellow box.

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 naranja) 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).

