**Descubrimiento**

Responder (Por sponsor):

● ¿Qué problema empresarial está intentando resolver el equipo?

La pérdida de dinero por una mala evaluación al momento de colocar un crédito.

● ¿Cuál es el resultado deseado del proyecto?

Predecir y reducir el riesgo de una mala colocación de acuerdo a la clasificación por zonas geográficas de riesgo o de buenas condiciones.

● ¿Qué fuentes de datos están disponibles?

Formularios de evaluación por cliente.

División de Zonas Geográficas.

● ¿Qué problemas de la industria pueden afectar el análisis?

La situación actual por pandemia y las restricciones impuestas por el gobierno.

● ¿Qué plazos deben tenerse en cuenta?

2 años a la Fecha

● ¿Quién podría proporcionar información sobre el proyecto?

Cualquier área de la empresa que cuente con estadísticas sobre colocación.

● ¿Quién tiene la autoridad final para tomar decisiones sobre el proyecto?

● ¿Cómo cambiará el enfoque y el alcance del problema si cambian las siguientes dimensiones?

Tiempo: ¿Analizando datos de 1 año o 10 años?

Riesgo: De conservador a agresivo

Recursos: Ninguno a ilimitado (herramientas, tecnología, sistemas)

Tamaño y atributos de los datos: Incluyendo fuentes de datos internas y externas

**Roles (Designar roles)**

Business User

Business Intelligence Manager

Project Sponsor

Project Manager

Data Engineer

DBA

Data scientist

**Categorías de los datos**

Datos antes de Pandemia

Datos durante Pandemia (Situacional)

**Hipótesis Iniciales**

HI1: El riesgo de tener una cartera NEGATIVA se puede medir/comparar entre regiones geográficas.

HI2: La información para evaluar el perfil de un cliente y el comportamiento de este puede tener relación con la ubicación geográfica.

HI3: Se pueden realizar estrategias de colocación de acuerdo a zonas geográficas buenas o malas.

* Análisis descriptivo
* Análisis predictivo

**Tareas**:

● Identificar fuentes de datos

● Capture fuentes de datos agregadas

● Revise los datos sin procesar

● Evalúe las estructuras de datos y las herramientas necesarias

● Alcance el tipo de infraestructura de datos necesaria para este tipo de problema

**Preparación de los Datos**

Definir ETL o ELT