

# Proyecto 1 – Roles

## Analista de negocio

### **Preguntas de negocio y plan de acción**

Determine la pregunta (o preguntas) de negocio que quiere resolver para su cliente seleccionado. Identifique cómo puede resolver estas preguntas a través de visualizaciones de los datos (descriptivo) y un modelo predictivo (idealmente de regresión, pero también puede ser de clasificación). Defina claramente su usuario objetivo y el contexto del producto desarrollado. Consulte con los demás roles el alcance del proyecto y defina los pasos a seguir.

*Entregable: Reporte de trabajo en equipo y Documento de alcance del proyecto.*

## Ingeniero de datos

### **Limpieza y alistamiento de datos**

Cargue los datos en Python, explore los datos disponibles y realice una limpieza cuidadosa. Identifique datos faltantes y decida una estrategia para su gestión. Asegúrese de que los datos queden en un formato que permita su posterior análisis. Documente los procedimientos realizados.

*Entregable: Soporte 1 – fuentes de limpieza (Jupyter Notebooks, Scripts de Python).*

## Analista de datos

### **Exploración de datos**

Realice un análisis exploratorio que permita describir estadística y visualmente el comportamiento de las variables a considerar. Calcule estadísticas descriptivas, realice histogramas, diagramas de caja, diagramas de dispersión, diagramas de violín y otros que permitan comprender cómo se comportan las variables. Documente el análisis realizado.

*Entregable: Soporte 2 – fuentes de análisis (Jupyter Notebooks, Scripts de Python).*

## Científico de datos

### **Modelamiento**

Aquí deberá explorar diferentes configuraciones de modelo, realizar ingeniería de características, emplear diferentes métodos de estimación, comparar y seleccionar las mejores alternativas. Consulte bibliografía que le permita contar con elementos para proponer los modelos. No es necesario emplear todas las variables disponibles, pero todas las variables incluidas y sus relaciones deben estar correctamente justificadas. Como hay un buen número de variables, se espera que el modelo desarrollado por cada equipo sea único. Evalúe su modelo usando métricas apropiadas. Documente el modelamiento realizado e incluya al menos 2 referencias como justificación.

*Entregable: Soporte 3 – fuentes de modelación con entrenamiento, validación y evaluación.*

## Desarrollador del Tablero de Datos

### Diseño y desarrollo del tablero

Empiece por diseñar el tablero: ¿qué valores debe permitir ingresar? ¿qué resultados genera? ¿qué visualizaciones incluye? ¿cómo mostrará las instrucciones? ¿cómo dispondrá estos elementos en el tablero? Para esta tarea es buena idea hacer un wireframe (un diseño sencillo que puede hacer en papel o digitalmente), que le permite tener una visión clara de su tablero y todos sus elementos. Recuerde no perder de vista al usuario y su necesidad. Piense siempre en la experiencia del usuario. Una vez haya terminado el diseño, desarrolle su tablero en Dash.

*Entregable: Wireframe en Figma del tablero y Soporte 4 – fuentes del tablero.*

## Especialista DevOps

### Despliegue

Despliegue el tablero en una instancia en EC2. Asegúrese de que su tablero sea accesible y quede en ejecución. Establezca un proceso de CI/CD que permita actualizar el tablero de forma automática con cada cambio en la rama principal de su repositorio. Defina una regla de protección sobre la rama principal y proponga una estructura de carpetas que facilite el desarrollo de su proyecto.

*Entregables: Soporte 5 – snapshots de la máquina lanzada (de EC2 y terminal), y URL del tablero en ejecución. Soporte 6 - Organización del repositorio en GitHub.*