



Descripción Proyecto

¡Felicitaciones! Has completado el sprint en Preprocesamiento de datos. Es hora de aplicar el conocimiento y las habilidades que has adquirido en un proyecto: un estudio de caso analítico real que completarás por tu cuenta.

Cuando termines el proyecto, envía tu trabajo al revisor del proyecto para que lo evalúe a través de la plataforma. Te dará su opinión en 24 horas. Utiliza los comentarios para realizar cambios y luego envía la nueva versión al revisor del proyecto.

Es posible que recibas más comentarios sobre la nueva versión. Esto es totalmente normal. Es común pasar por varios ciclos de comentarios y revisiones.

Tu proyecto se considerará completado una vez que el revisor del proyecto lo apruebe.

Descripción del proyecto

Tu proyecto consiste en preparar un informe para la división de préstamos de un banco. Deberás averiguar si el estado civil y el número de hijos de un cliente tienen un impacto en el incumplimiento de pago de un préstamo. El banco ya tiene algunos datos sobre la solvencia crediticia de los clientes.

Su informe se tendrá en cuenta al crear una **puntuación de crédito** para un cliente potencial. Se utiliza una **puntuación de crédito** para evaluar la capacidad de un prestatario potencial para pagar su préstamo.

Instrucciones para completar el proyecto

Paso 1. Abre el archivo de datos [/datasets/credit_scoring_eng.csv](#) y mira la información general.

Paso 2. Preprocesa los datos:

- Identifica y completa los valores ausentes
- Reemplaza el tipo de datos de número real con el tipo de número entero
- Elimina datos duplicados
- Clasifica los datos

Asegúrate de explicar:

- Qué valores ausentes identificaste
- Posibles razones por las que estos valores ausentes estaban allí
- Qué método utilizaste para completar los valores ausentes
- Qué método utilizaste para encontrar y eliminar datos duplicados y por qué
- Posibles razones por las que había datos duplicados
- Qué método utilizaste para cambiar el tipo de datos y por qué
- Qué diccionarios has seleccionado para este conjunto de datos y por qué

Los datos pueden contener artefactos o valores que no se corresponden con la realidad (por ejemplo, un número negativo de días empleados). Este tipo de cosas sucede cuando trabajas con datos reales. Debes describir las posibles razones por las que tales datos pueden haber aparecido y procesarlos.

Paso 3. Responde estas preguntas:

- ¿Hay alguna conexión entre tener hijos y pagar un préstamo a tiempo?
- ¿Existe una conexión entre el estado civil y el pago a tiempo de un préstamo?
- ¿Existe una conexión entre el nivel de ingresos y el pago a tiempo de un préstamo ?
- ¿Cómo afectan los diferentes propósitos del préstamo al reembolso a tiempo del préstamo?

Interpreta tus respuestas. Explica el significado de los resultados que obtuviste.

Paso 4. Escribe una conclusión general.

Formato:

Completa el proyecto en Jupyter Notebook (se abrirá cuando hagas clic en Siguiente). Escribe código en las celdas de **código**, y notas con explicaciones e interpretaciones en las celdas **Markdown**. Utiliza formatos y encabezados.

Descripción de los datos

- `children` : el número de hijos en la familia
- `days_employed` : por cuánto tiempo ha estado trabajando el cliente
- `dob_years` : la edad del cliente
- `education` : el nivel educativo del cliente
- `education_id` : identificador de la educación del cliente
- `family_status` : estado civil del cliente
- `family_status_id` : identificador del estado civil del cliente
- `gender` : el género del cliente
- `income_type` : el tipo de ingreso del cliente
- `debt` : si el cliente ha incumplido alguna vez un préstamo
- `total_income` : ingresos mensuales
- `purpose` : motivo por el que se solicita un préstamo

¿Cómo será evaluado mi proyecto?

Tu revisor utilizará criterios de evaluación para evaluar tu proyecto. Antes de comenzar, lee atentamente los criterios. Esto te ayudará a asegurarte de que haces un buen trabajo.

Esto es lo que observan los revisores al evaluar tu proyecto:

- ¿Cómo describirías los problemas identificados en los datos?

- ¿Qué métodos utilizas para reemplazar tipos de datos y procesar valores ausentes y datos duplicados?
- ¿Has clasificado los datos? ¿Por qué lo hiciste de la manera en que lo hiciste?
- ¿Exportaste los datos finales a tablas dinámicas?
- ¿Utilizaste cláusulas try-except en tu código para manejar errores inesperados?
- ¿Seguiste la estructura correcta del proyecto y escribiste un código limpio?
- ¿Sacaste conclusiones precisas y útiles?
- ¿Dejaste comentarios sobre cada paso que tomaste?

Cuando hagas clic en "¡Vamos!", serás redirigido a JupyterHub de Practicum. Puedes completar tu proyecto allí.

Otra opción: puedes instalar Python en tu ordenador y hacer tu proyecto de forma local. Existen muchas guías para instalar Python y sus librerías en tu ordenador.

Una manera fácil de instalar todo lo que necesitas de una vez es descargar el kit de distribución de Anaconda (<https://www.anaconda.com/distribution/>) (*materiales en inglés*). Durante la instalación, comprueba las siguientes opciones en la configuración avanzada:

- Añadir Anaconda a la variable de entorno PATH del sistema
- Registrar Anaconda como el sistema Python 3.n (3.n será reemplazado por el número de versión del software Python incluido en la descarga de Anaconda)

Hay dos formas de abrir Jupyter Notebook:

- Busca el Jupyter Notebook en Anaconda Navigator e inícialo haciendo clic en el icono
- Escribe el comando `jupyter notebook` en la línea de comandos (Windows) o Terminal (macOS)

El conocimiento y las habilidades que adquiriste en las lecciones anteriores te ayudarán a completar tu proyecto con éxito.

Ten a mano tus hojas de información y resúmenes.

¡Buena suerte!