

# Data Science



Trabajo Final: Entrega N2

Evaluación Crediticia

Comisión: 41875

Profesor: Ignacio Ruso

Asistente: Juan Cruz Alric

Alumno: Juan Ignacio Rainoldi

# Objetivos

Obtener datos de diversas fuentes como APIs o Bases de datos públicas para luego analizarlos mediante el lenguaje Python con el fin de contestar una pregunta de interés para una industria, negocio o proyecto personal. Se deberán utilizar datasets complejos implementando técnicas avanzadas para la limpieza y adquisición de datos. Para cumplir con dicho objetivo se deberá:

1. Importar datos crudos de APIs o bases de datos usando Python.
2. Limpiar y transformar los datos para permitir un posterior análisis.
3. Abstracto con motivación y audiencia: Descripción de alto nivel de lo que motiva a analizar los datos elegidos y que audiencia se podrá beneficiar de este análisis.
4. Preguntas/hipótesis que queremos responder: Lista de preguntas que se busca responder mediante el análisis de datos
5. Análisis exploratorio de datos (EDA): Análisis descriptivo de los datos mediante visualizaciones y herramientas estadísticas.

# Introducción

**Contexto Empresarial:** En el contexto de una institución financiera, como un banco o una compañía de préstamos, se desea automatizar y mejorar el proceso de evaluación crediticia de los solicitantes. Para lograrlo, se decide utilizar técnicas de aprendizaje automático y construir un modelo capaz de clasificar automáticamente a una persona que solicite un crédito.

**Problema empresarial:** Su tarea es procesar la base de datos brindada con la cual se realizará el modelo. Se deberán realizar visualizaciones y resúmenes numéricos para responder a las preguntas específicas que realiza el cliente. Las mismas se mencionarán a continuación.

**Contexto analítico:** Se proporciona un archivo CSV ("test.csv"). El contiene información detallada sobre el historial bancario de diversos clientes como ingreso mensual, número de cuentas bancarias, monto adeudado, etc. Se empleará esta base de datos para poder realizar estimaciones acerca de la idoneidad de un cliente para recibir un crédito, basándose en su historial bancario. En este dataset, los datos están etiquetados; es decir, hay una variable que dice cuáles de los clientes son confiables y cuáles no. Por lo que, se deberá utilizar modelos de clasificación para abordar este problema de aprendizaje supervisado.

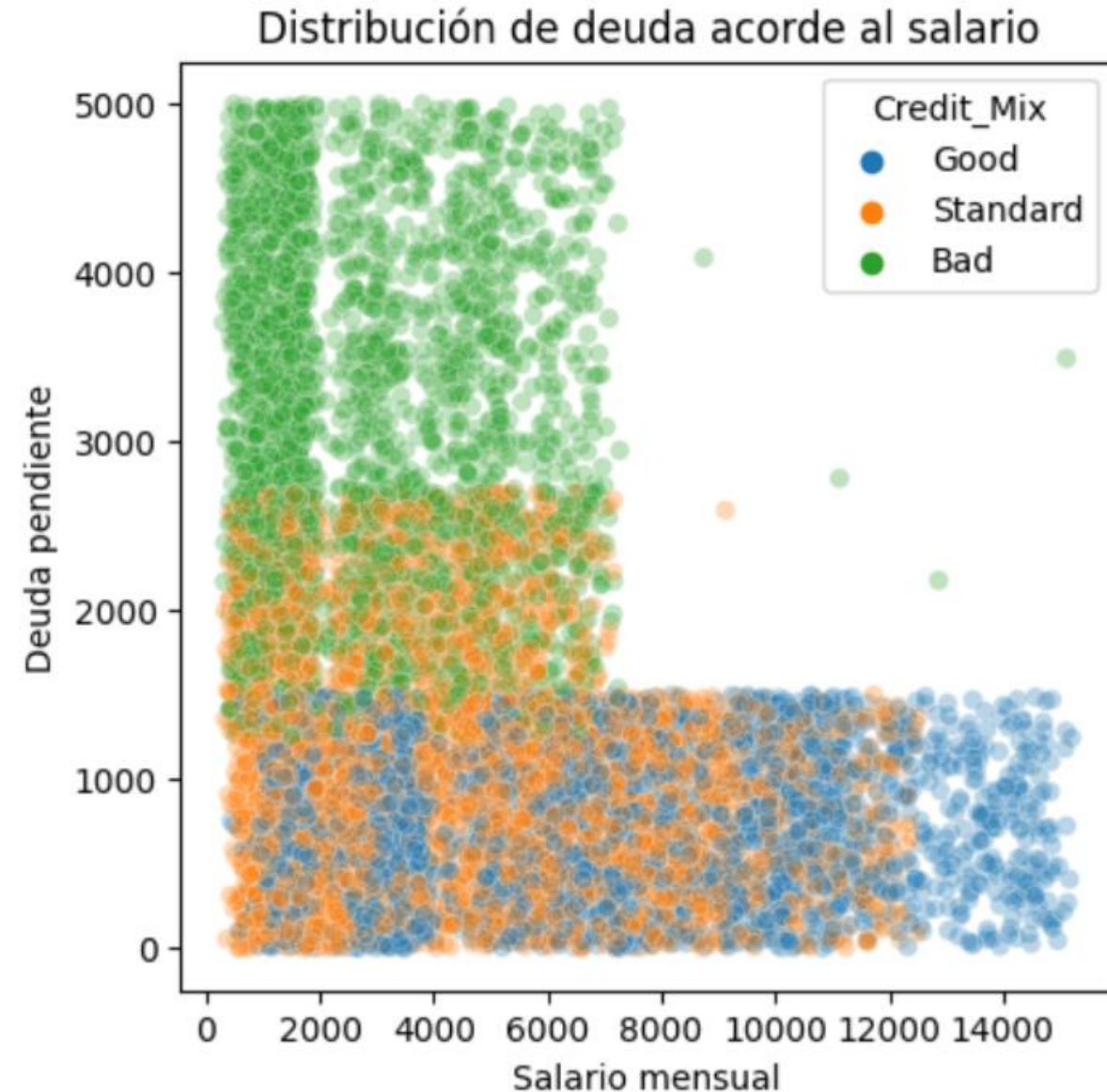
# Análisis exploratorio

# Hay alguna relación entre el monto adeudado, el ingreso mensual y la clasificación del cliente.

Se puede observar que los **clientes** que están clasificados como **malos** poseen una **deuda mayor igual a 1500 USD** y a su vez sus **ingresos no superan a 7000 USD** mensuales.

Los clientes que están clasificados como **buenos clientes** son aquellos que **no adeudan más de 1600USD** y su **salario** puede llegar **hasta los 15000 USD**.

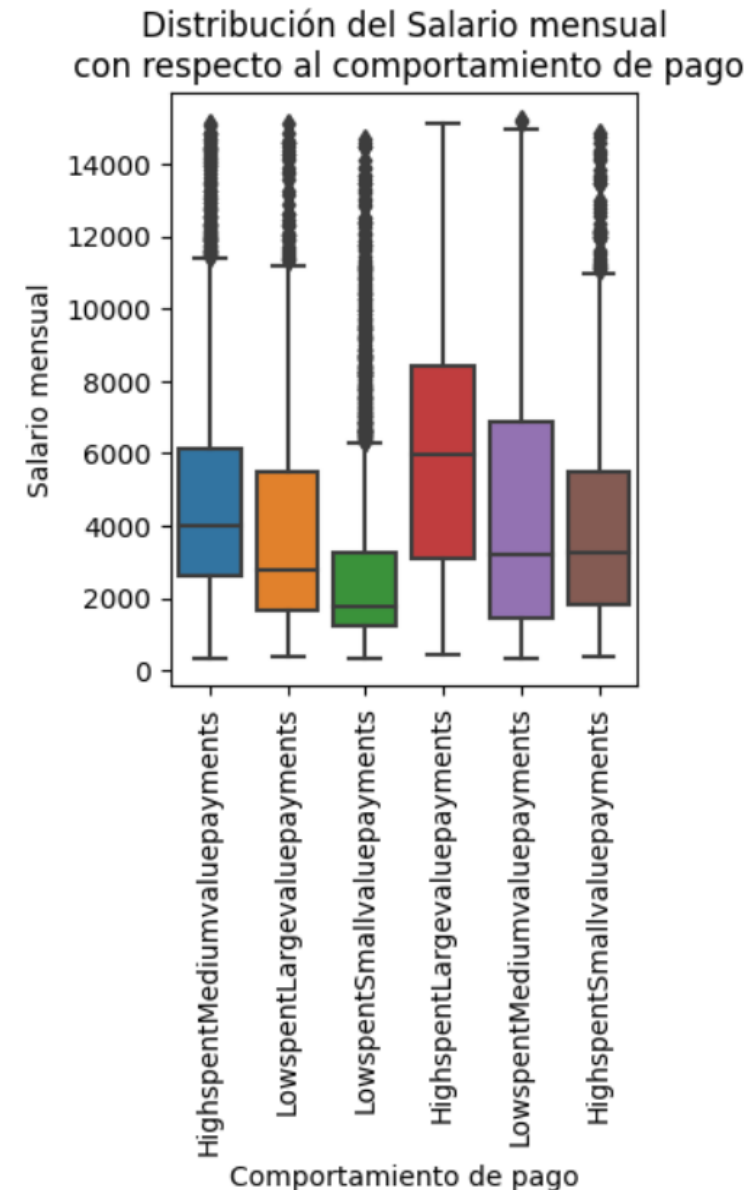
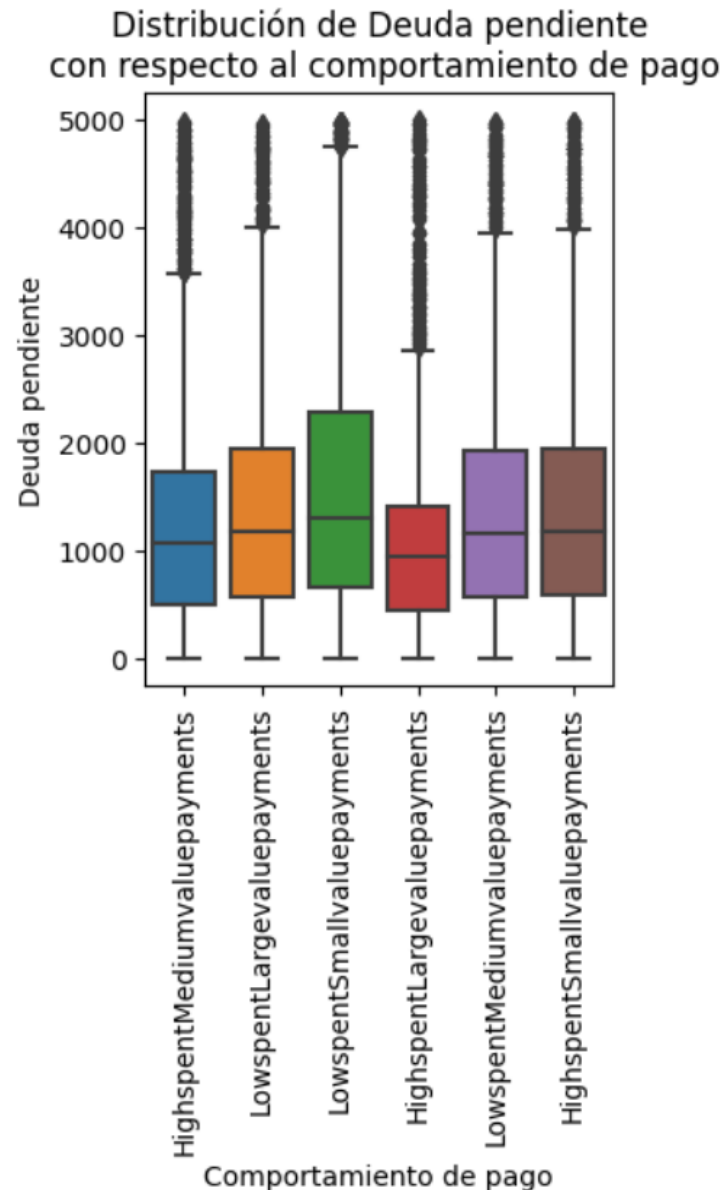
Los clientes clasificados como **Standard** son aquellos que alcanzan a poseer una **deuda de hasta 2800 USD** y sus **salarios** pueden llegar **hasta los 12200 USD**.



# Existe alguna relación entre el comportamiento de pago, el ingreso mensual los pagos mensuales al EMI y la deuda pendiente.

Se observa que los **clientes** que **más gastan** y **realizan grandes pagos** son aquellos que suelen tener **mejores salarios** y poseen una **deuda pendiente menor**.

En el caso de los **clientes** que realizan **pocos gastos** y **realizan pagos bajos** son los que tienen un **menor sueldo** y poseen una **deuda mayor**.



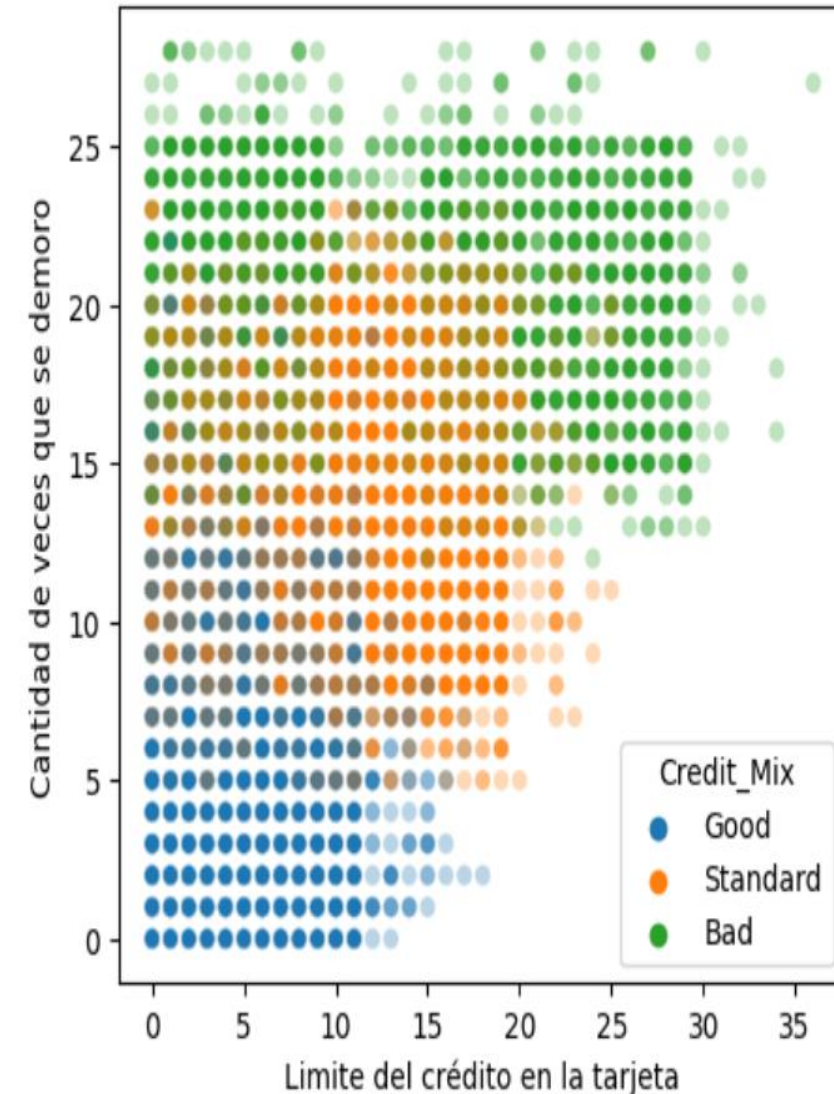
# Existe una relación entre la cantidad de demoras de pago y el cambio del límite de crédito.

Se puede observar que aquellos clientes que han modificado su límite crediticio a partir de un 17% el número de días demora mínimos va aumentando.

A su vez se puede observar que la **mayoría** de los **clientes** que se han **demorado menos de 8 veces** son aquellos que son considerados **clientes buenos** para la solicitud de un crédito.

Se observa que la mayoría de los clientes con un **cambio** en el **límite crediticio igual o mayor al 20%** o que se han **demorado más de 20 ocasiones** en pagar son considerados **malos clientes** para la otorgación de un crédito.

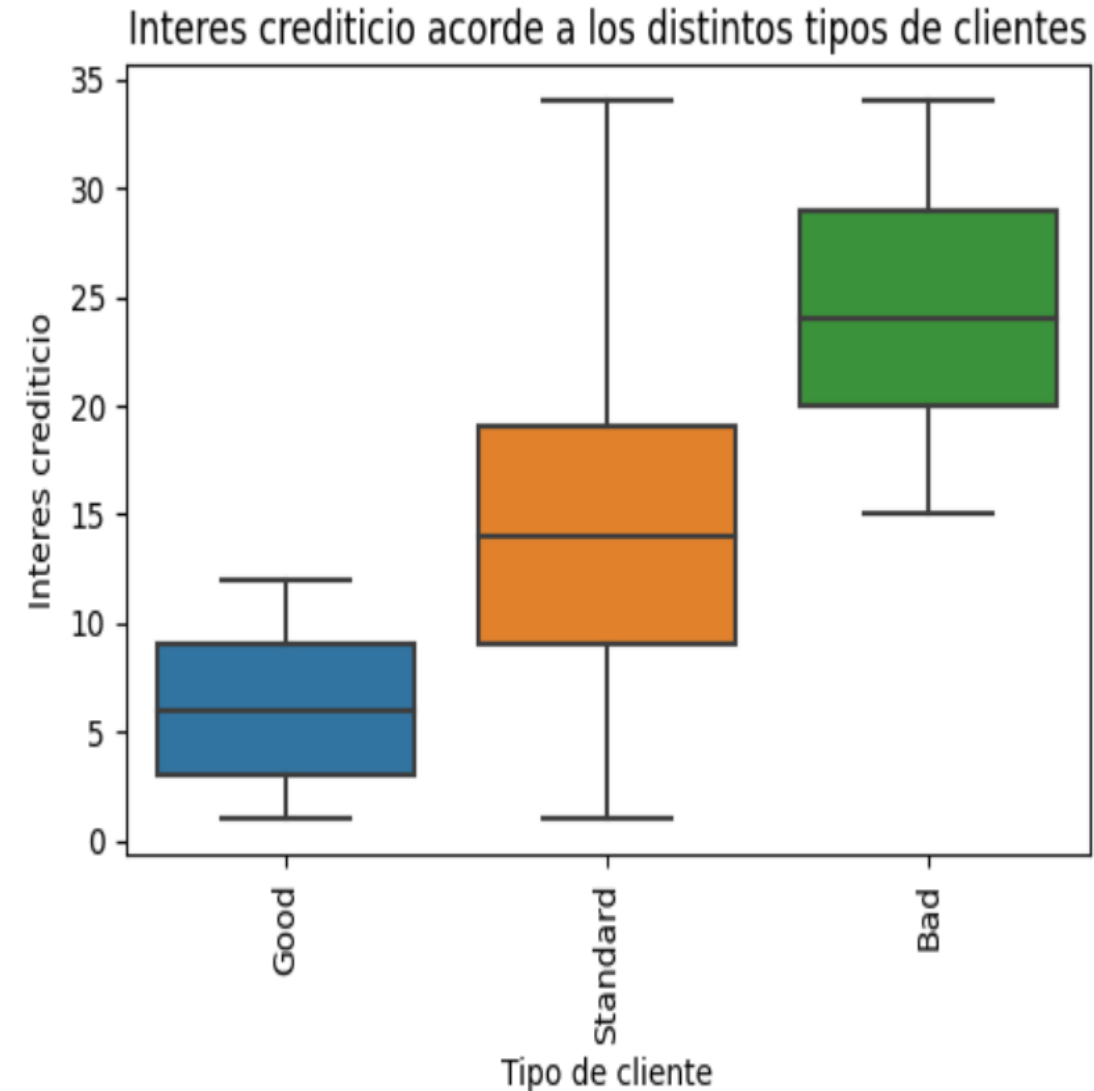
Distribución cantidad de veces que se demora con respecto al límite de la tarjeta de crédito





# Se puede afirmar que los clientes más aptos para créditos tienen mejor interés crediticio.

Se observa que los que se consideran buenos clientes poseen un interés menor igual a 12,5%. Y que los que se consideran malos clientes poseen un interés crediticio mayor igual a 15%.





# Insights

- Malos clientes poseen una deuda mayor igual a 1500 USD e ingresos menores a 7000 USD.
- Buenos clientes no adeudan más de 1600USD y salarios hasta 15000 USD.
- Clientes standard poseen una deuda hasta 2800 USD y salarios hasta 12200 USD.
- Clientes que más gastan y realizan grandes pagos suelen tener mejores salarios y deudas pendientes menores.
- Clientes que realizan pocos gastos y pagos bajos tienen sueldos menores y deudas mayores.
- Cambios en el límite de crédito mayores al 17%, va aumentando la cantidad de demoras en el pago.
- Clientes con menos de 8 demoras, son buenos clientes para la solicitud de un crédito.
- Clientes con cambios en el límite crediticio igual o mayor al 20% o con más de 20 demoras son considerados malos clientes.
- Los buenos clientes poseen un interés menor igual a 12,5%.
- Los clientes poseen un interés crediticio mayor igual a 15%.

## Recomendación

No realizar prestamos a clientes que adeuden más de 2800 USD, que su interés crediticio sea mayor a 15%, que hayan realizado un cambio en el límite crediticio mayor al 20% o que tenga más de 20 demoras.