

1.- Definir el problema de la naturaleza que se tiene a continuación, además de los objetivos de negocio bien definidos.

Bajo la metodología **CRISP-DM** (Cross-Industry Standard Process for Data Mining, proporciona un marco estructurado para llevar a cabo el análisis de datos de manera sistemática y efectiva), la definición del problema y los objetivos de negocio es la primera etapa y esencial para orientar todo el proceso de análisis. En ese sentido, se puede definir del siguiente modo:

Comprensión del negocio

Problema: La universidad desea comprender qué características socioeconómicas de sus docentes influyen en sus niveles salariales logrando predecir con precisión el salario de cada profesor. Esto facilitará la toma de decisiones respecto a la administración, política salarial, programas de incentivos, y planificación de recursos humanos, permitiendo a la institución homologar sus salarios con estándares internacionales y ofrecer condiciones más justas y competitivas.

Objetivos de negocio:

Principal	Secundarios
	Identificar los factores (variables drivers) que más impactan en los salarios.
Desarrollar un modelo de predicción que estime el salario de los docentes basándose en sus características socioeconómicas.	Mejorar la gestión de recursos humanos mediante segmentaciones efectivas.
	Diseñar programas de capacitación, incentivos y bonificaciones adaptados a los diferentes grupos de docentes.
	Homologar los salarios con estándares comparables en otros países o instituciones de educación superior.

Estos objetivos garantizan que el análisis sea útil, dirigido a mejorar continua de la gestión institucional y alineado con los intereses de la universidad.

2.- ¿Qué tipo de variables se utilizan en el problema de negocio?

Comprensión de los datos

Variable	Descripción de la variable	Independiente (predictoras o drivers) /Dependiente (objetivo)	Tipo de variable
Rank	Rango o categoría del profesor (Prof. Asociado, Prof. Asistente, Prof. Principal)	Independiente	Cualitativa ordinal Orden jerárquico).

Discipline	Disciplina o especialidad en la que el profesor se desenvuelve. (A: Ciencias, B: Letras)	Independiente	Cualitativa nominal (categorías distintas).
Yrs.since.phd	Número de años desde que obtuvo el doctorado en su área o ámbito.	Independiente	Cuantitativa discreta (valor entero).
Yrs.service	Número de años de servicio en la docencia.	Independiente	Cuantitativa continua (valor entero).
Sex	Género del profesor o docente.	Independiente	Cualitativa nominal (sin orden).
Salary	Salario anual en soles del docente o profesor.	Dependiente	Cuantitativa continua (infinitos valore)

3.- Realizar un análisis exploratorio de las variables o drivers que más impactan o impactarían en la predicción o pronóstico de los salarios de los docentes.

Preparación de los datos: Limpiar, transformar y seleccionar las variables relevantes para modelar.

Modelado: Aplicar técnicas de análisis predictivo para identificar los drivers principales y construir modelos que predigan el salario.

Evaluación: Validar la precisión y utilidad de los modelos definidos, asegurando que cumplen con los objetivos del negocio.

Implementación: Planificar acciones basadas en los resultados, como programas de capacitación, ajuste de salarios o campañas de motivación.

Este enfoque garantiza que el análisis sea coherente, orientado a respuestas prácticas en línea con los objetivos de la universidad.