

# PROYECTO FINAL MODELOS PREDICTIVOS

GABRIEL AH CHU

# ¿Qué le motivo a nivel personal a desarrollar el estudio predictivo?

Proyecto: Predicción de Salarios de Carreras Afines a las Ciencias de Datos



Analizar el mercado laboral actual de las ciencias de datos para determinar si estoy tomando una buena decisión al cursar esta maestría.



Entender el valor del trabajo para evitar la devaluación en el mercado. Saber si se está teniendo una compensación justa.



Compartir los insights con los compañeros para fomentar un ambiente de aprendizaje colaborativo y enriquecedor.

# ¿Cómo fue su experiencia buscando y seleccionando la data?

Se realizaron búsquedas en los repositorios compartidos por el profesor. Como primera opción estaba buscando datasets relacionados a mi trabajo, redes de telecomunicaciones. De esta área se encontraron pocos datasets y sin procesar, para poder trabajarlos necesitaba herramientas especializadas.

Durante la búsqueda se encontraron varios datasets que me llamaron la atención y que consideraré explorar a futuro



# ¿Qué lecciones aprendió al hacer el análisis descriptivo?

No necesariamente se eliminan los outliers

Analizar la posibilidad de reducir categorías

Dominio del tema es una ventaja

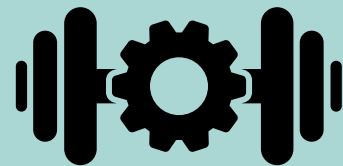


# ¿Qué pasos realizó para hacer el análisis de los datos tanto descriptivo como predictivo?

## Descriptivo



Selección del Dataset  
Definición de herramienta a utilizar  
Validación del Dataset con el profesor

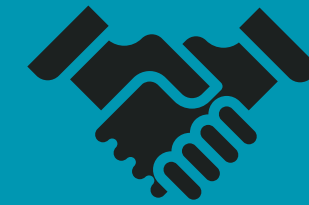


Procesamiento del Dataset: Se verifican outliers, valores nulos, etc.



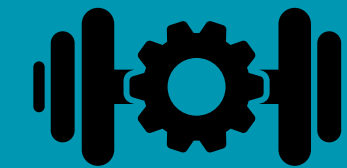
Generación de Distribuciones  
Análisis de Tendencias  
Frecuencias de datos  
Relaciones entre los niveles de experiencia y el salario

## Predictivo



Toma de decisiones:

- Outliers
- Variables a utilizar
- Métricas para evaluación de los modelos
- Modelos a utilizar



Estudiar modelos de Machine Learning  
Preparación de datos: Codificación de Variables Categóricas  
Implementar los modelos de Machine Learning

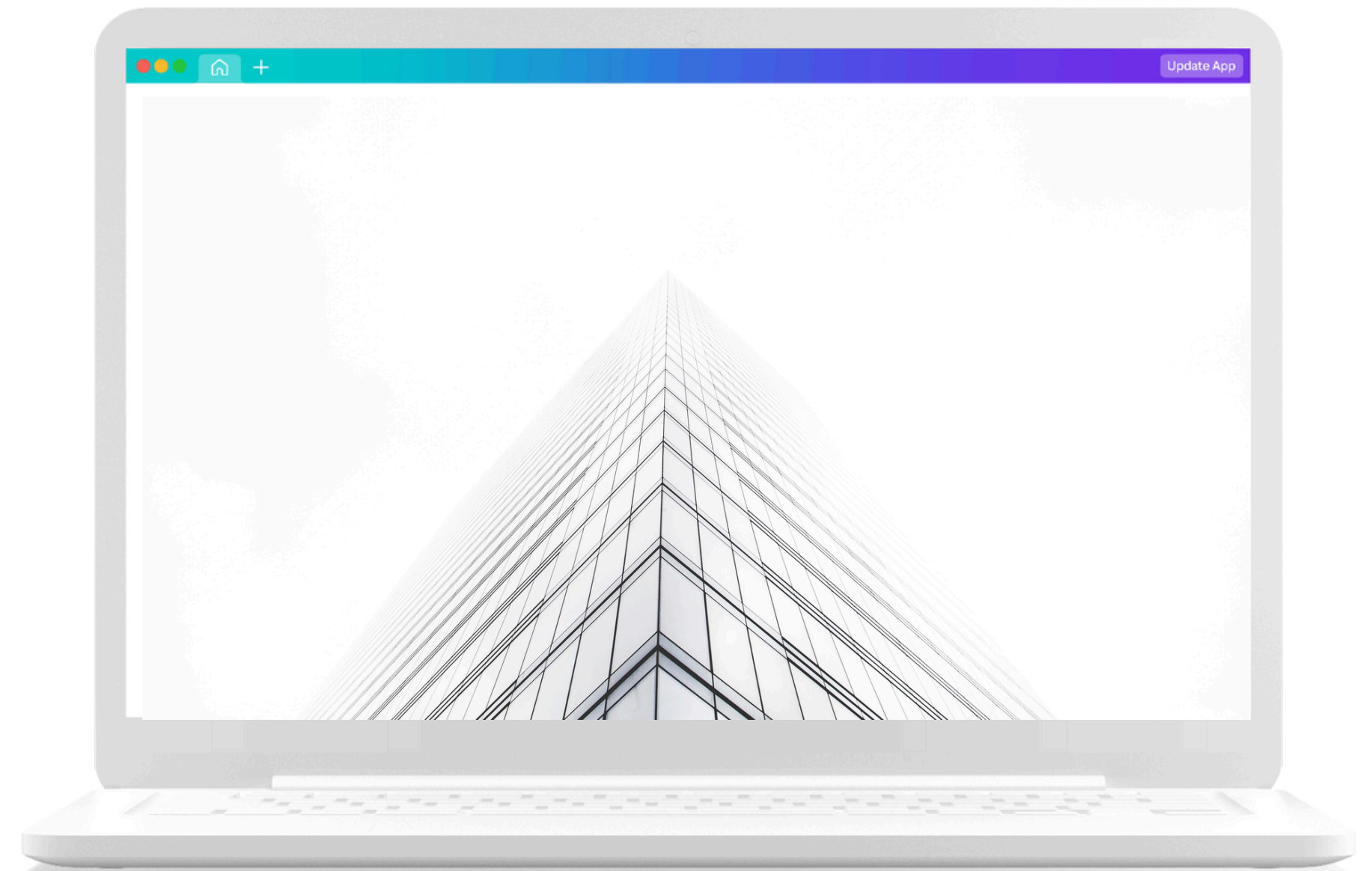


Generación de predicciones  
Evaluación de Modelos  
Generación de Gráficas



# ¿Qué descubrió de la investigación?

Actualmente, las ciencias de datos son carreras altamente solicitadas y bien remuneradas. En 2024, se observó un aumento considerable de posiciones en este campo, y la tendencia del primer trimestre de 2025 sugiere que seguirá creciendo. Las personas con más experiencia y perfiles que trabajan directamente con los datos, como Data Scientist, Data Analyst y Machine Learning Analyst son los mejores remunerados, aparte de aquellos en puestos ejecutivos.



# Si fuese a hacer este estudio de nuevo ¿qué haría mejor?

Dashboard en PowerBI

Evaluar otros modelos

Análisis de Cluster

Feature Selection



# ¿Qué descubrió de si mismo?



Mejor Python

Gestión de Tiempo

Práctica de Conocimientos

Relación de lo aprendido con mi trabajo



**Gracias!**