

# **Data Science II**

(Machine Learning para la Ciencia de Datos)

Comisión 75700

Alumna: Durian Catalina Soledad

Año: 2025

## **Resumen del Proyecto**

Este proyecto analiza datos educativos globales para detectar patrones y anomalías en indicadores como el puntaje promedio de aprendizaje y la proporción de estudiantes que alcanzan distintos niveles de desempeño. Utilizando datos de 158 países entre 2000 y 2015, se aplican visualizaciones estadísticas y una simulación de detección de anomalías para identificar cambios atípicos en la calidad educativa. El trabajo busca aportar una base exploratoria útil para el análisis de políticas educativas comparadas.

## **Preguntas de Investigación**

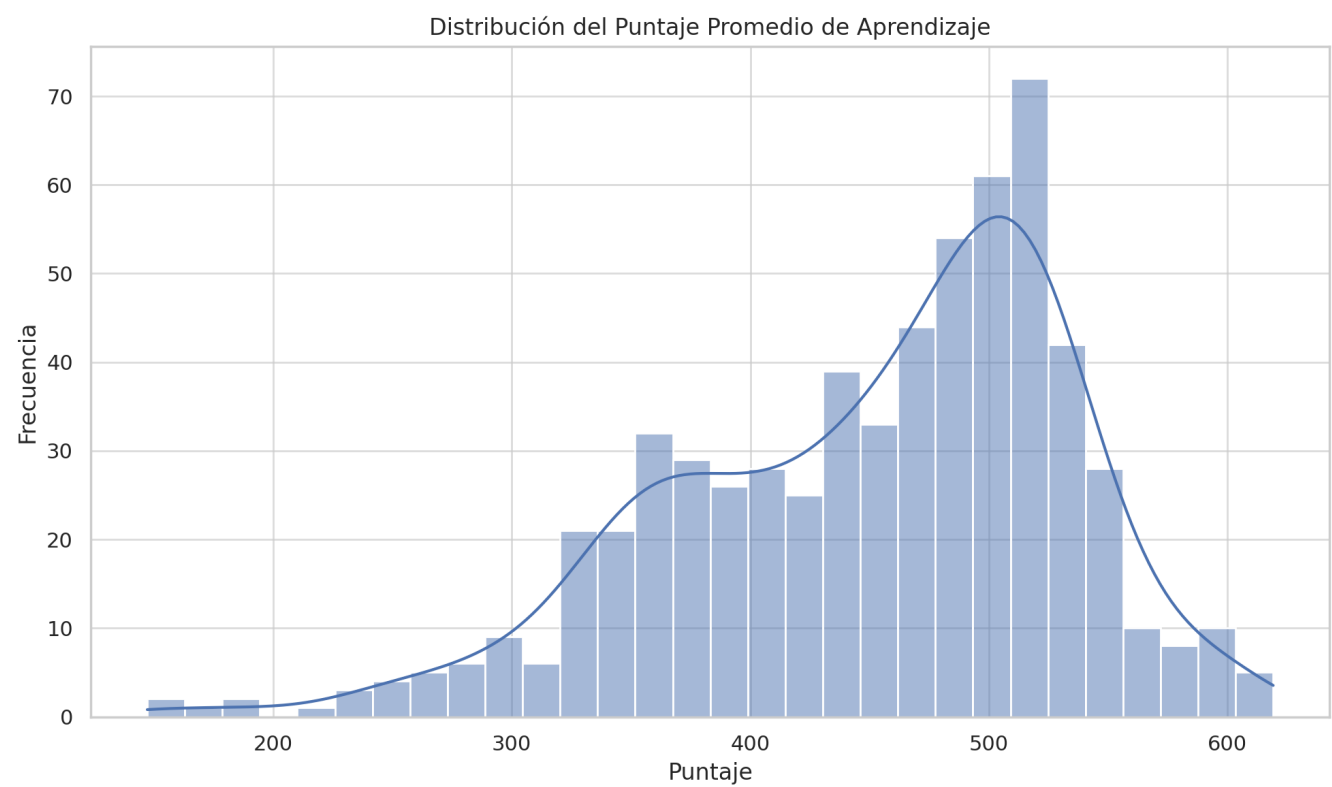
- ¿Qué países muestran desempeño educativo distinto al promedio global?
- ¿Existen cambios abruptos en los indicadores educativos?
- ¿Qué países muestran mayor variabilidad interanual?

## **Hipótesis**

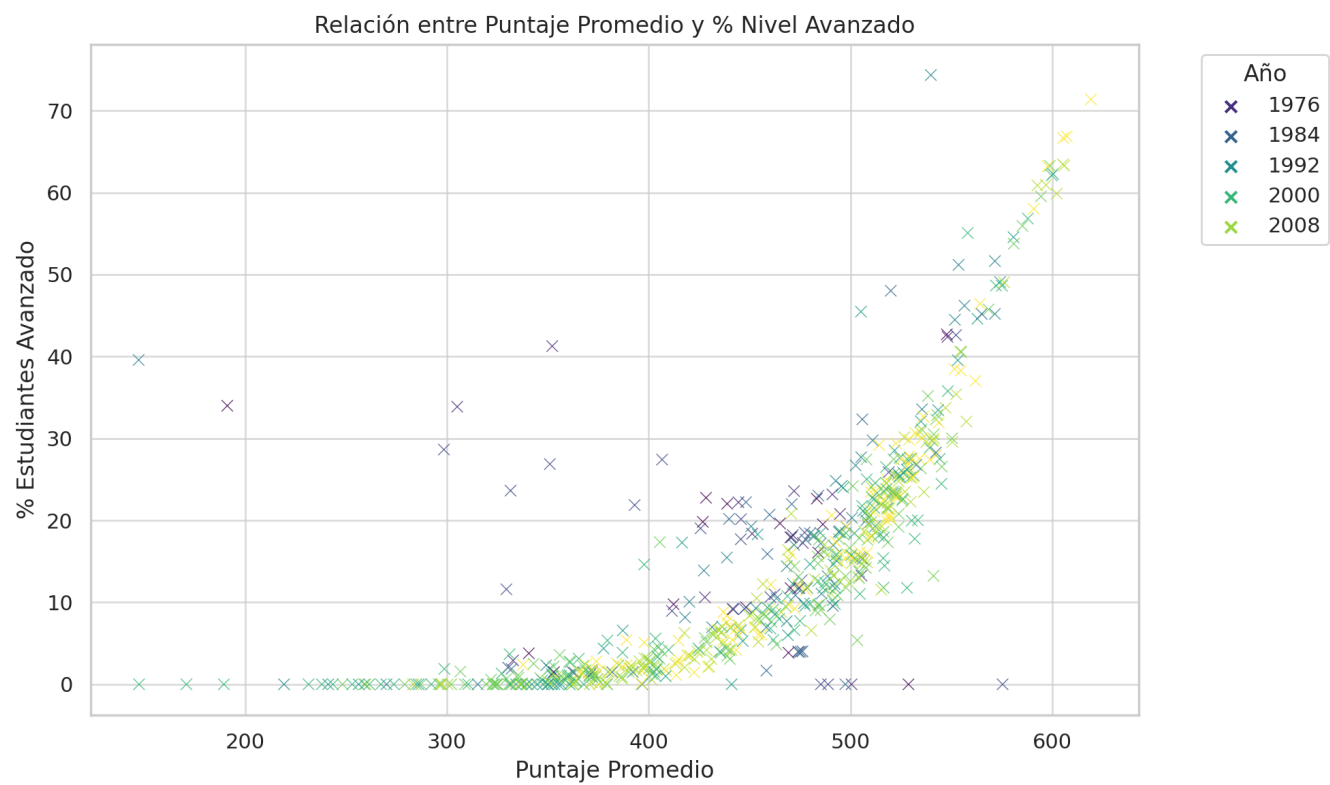
- H1: Países con menor porcentaje de estudiantes avanzados tienen más anomalías.
- H2: Las anomalías coinciden con años de eventos críticos.
- H3: Hay correlación positiva entre puntaje promedio y desempeño avanzado.

## **Visualizaciones**

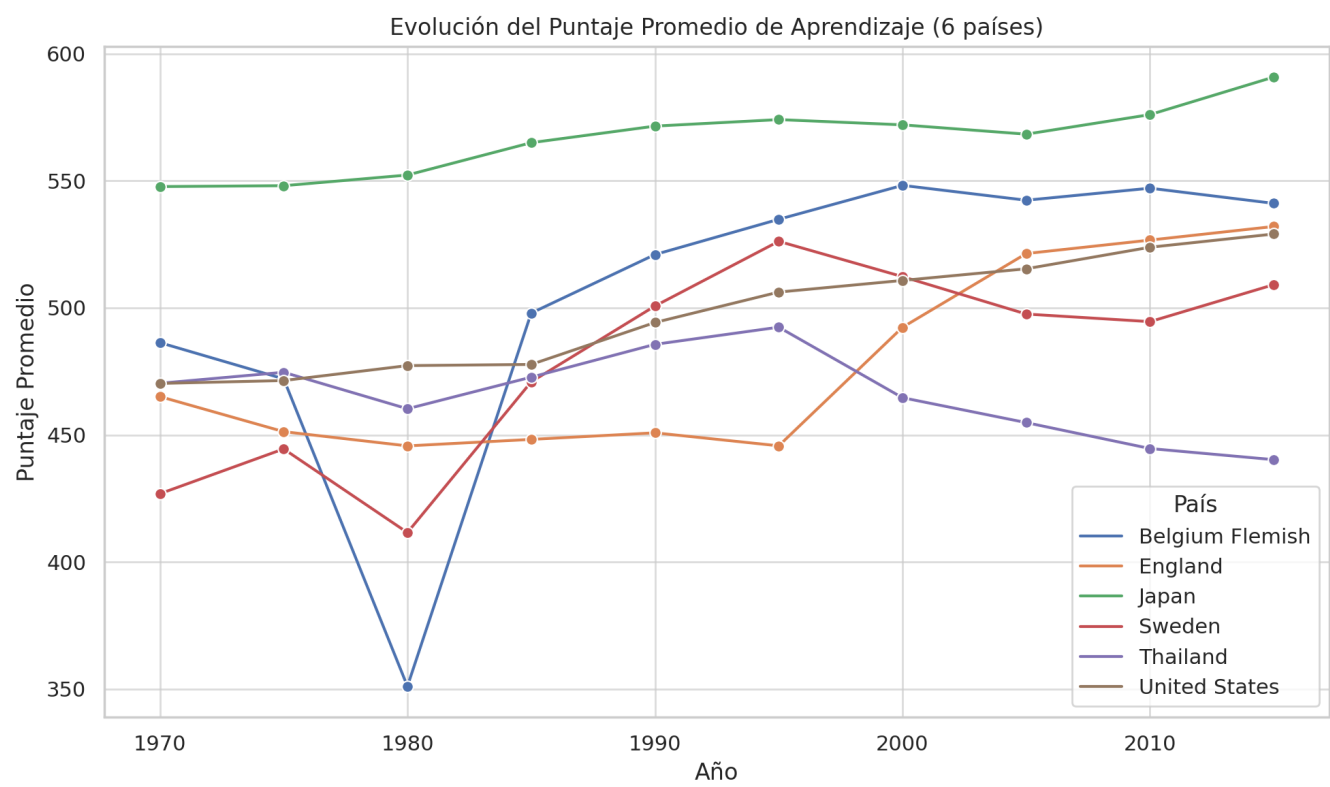
Distribución del Puntaje Promedio de Aprendizaje



Relación entre Puntaje Promedio y % Nivel Avanzado



Evolución del Puntaje Promedio de Aprendizaje (6 países)



# Evolución del Puntaje Promedio con Anomalías

