**Tema: Educación**

**Integrantes:**

**Deyton Riascos Ortiz (2246208)**

**Samuel Izquierdo Bonilla (2246993)**

**Propuesta de análisis:**  
Explorar el impacto de distintos factores (género, etnia, almuerzo escolar, etc.) en el rendimiento académico de estudiantes.

**Dataset:**  
Student Performance Dataset

https://www.kaggle.com/datasets/spscientist/students-performance-in-exams

**Columnas clave:**

* Gender
* Parental level of education
* Test preparation course
* Math, Reading, Writing scores

Este conjunto de datos es excelente para análisis exploratorio, ya que es limpio, pequeño y contiene variables categóricas y numéricas.

**Preguntas de análisis propuestas**

1. **¿Cuál es la distribución de calificaciones en matemáticas, lectura y escritura entre todos los estudiantes?**  
   → Gráfico: Histogramas (EDA univariado)
2. **¿Qué género obtiene en promedio mejores resultados en matemáticas, lectura y escritura?**  
   → Gráfico: Barras con groupby('gender')
3. **¿Influye el nivel educativo de los padres en las calificaciones de los estudiantes?**  
   → Gráfico: barras agrupadas por parental level of education
4. **¿Los estudiantes que completaron el curso de preparación obtienen mejores resultados que los que no lo hicieron?**  
   → Gráfico: barras comparativas con groupby('test preparation course')
5. **¿Qué combinación de género y almuerzo escolar está asociada con mayores o menores promedios?**  
   → Gráfico: Barras múltiples con groupby(['gender', 'lunch'])
6. **¿Existe alguna diferencia significativa entre el rendimiento en matemáticas, lectura y escritura por grupo étnico?**  
   → Gráfico: Barras o violin plot con groupby('race/ethnicity')
7. **¿Cuál es la correlación entre las tres calificaciones?**  
   → Gráfico: Heatmap de correlación (math score, reading score, writing score)
8. **¿Cuál es la calificación promedio por tipo de almuerzo (estándar vs. reducido)?**  
   → Gráfico: Barras con groupby('lunch')
9. **¿Qué porcentaje de estudiantes obtuvo más de 90 puntos en las tres asignaturas?**  
   → Métrica y pie chart (condición sobre columnas de calificación)
10. **¿Cuál es la distribución conjunta de las notas entre matemáticas y escritura (o lectura)?**  
    → Gráfico: Scatter plot o joint plot