

**Estudio de performance  
de estudiantes de  
escuelas secundarias.**

# Agenda

1. Contexto y Audiencia.
2. Hipótesis y preguntas de Interés
3. Descripción de metadatos.
4. Análisis Exploratorio
5. Insights y Recomendaciones.

# Contexto y Audiencia

Se desea interpretar las variables que influyen en el desempeño académico de un grupo de estudiantes de escuelas secundarias. El estudio está dirigido a entidades gubernamentales como el Ministerio de Educación y a autoridades de establecimientos educativos.

# Hipótesis y Preguntas

- Qué incidencia y peso tienen las variables actitudinales como hábitos de estudio, horas de estudio, ausencias, tutorías, participación de los padres y actividades extracurriculares?
- Se relaciona el desempeño con factores demográficos? Estos pueden incluir educación de los padres, minorías sociales que indiquen si se necesita investigar más en profundidad el acceso a los recursos educativos
- Investigar si hay diferencias significativas en las notas separando por grupos, por ejemplo género.
- Sentar las bases para elaborar modelos de predicción de desempeño.

# Descripción de los metadatos

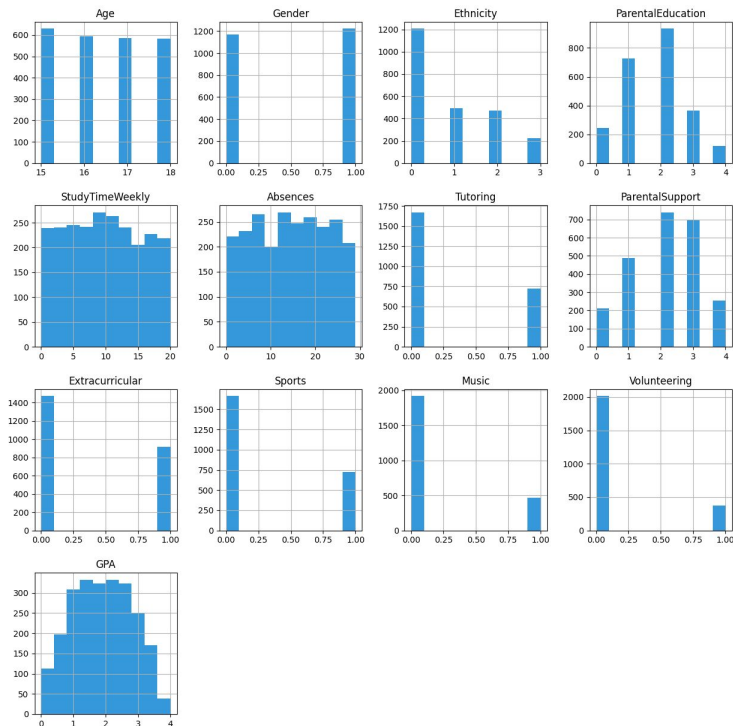


Fig. 1 Descripción del dataset

- 2392 estudiantes
- 13 características de interés
- Variable objetivo: GPA
- Distribución de edades homogénea (15, 16, 17 y 18 años)
- Balance de género.
- Predominio de estudiantes caucásicos.
- Asimetría en los conteos por si o por no de las actividades extra-áulicas (Extracurriculares, Deportes, Música, Voluntariados.)

# Análisis Exploratorio

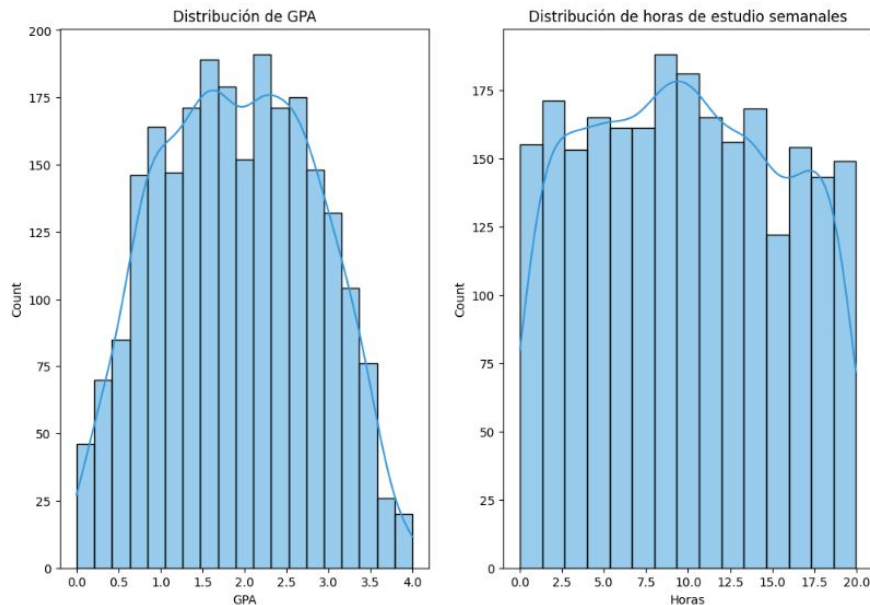


Fig. 2 Distribuciones de GPA y Horas de estudio semanales

- La distribución de notas GPA presenta una forma acampanada que asemeja a una distribución normal recortada por el rango entre 0 y 4.
- La cantidad de horas asignadas a estudio varía entre 0 y 20.

# Análisis Exploratorio

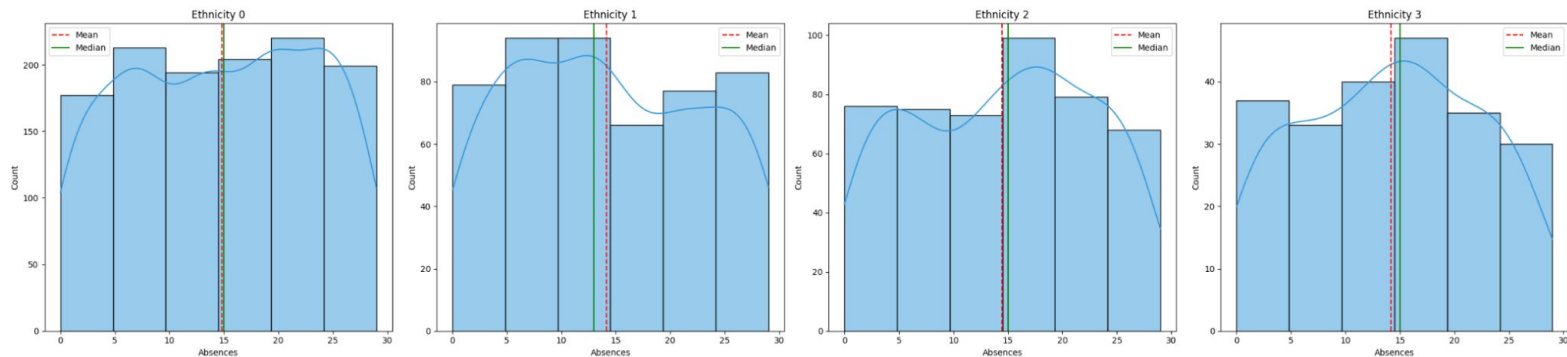


Fig. 3 Ausencias por etnias.

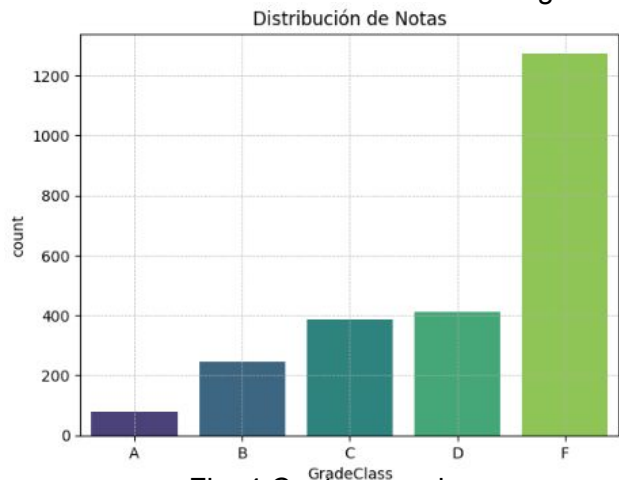


Fig. 4 Conteo por clase

- Las ausencias no tienen relación evidente con las etnias (Fig. 3)
- Se evidencia una proporción mucho más grande ~50% de desaprobados (nota F)
- Son mucho menos frecuentes en el otro extremo las "A". (Fig. 4)

# Análisis Exploratorio

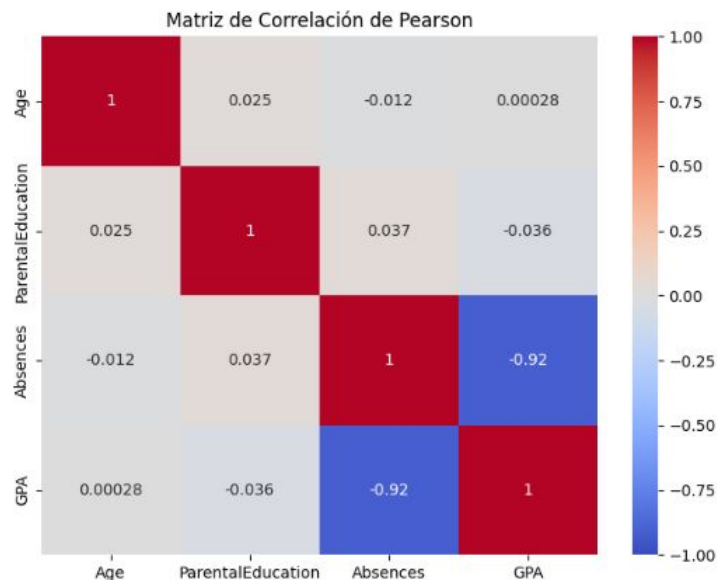


Fig. 5 Matriz de Correlación

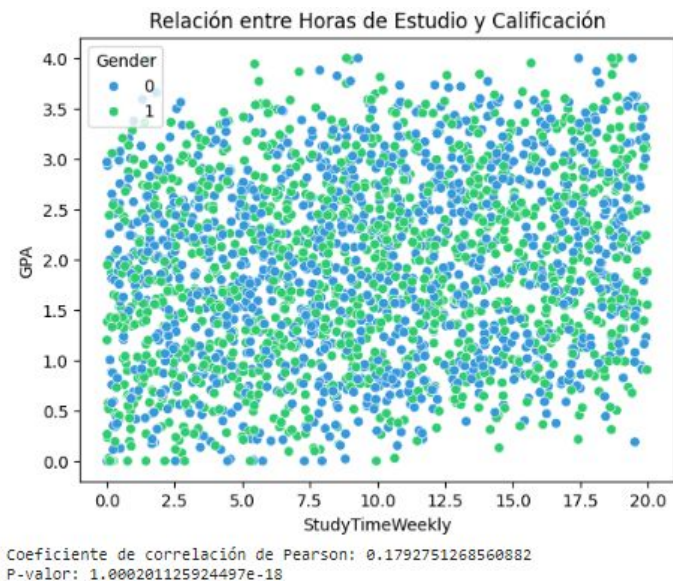


Fig. 6 GPA vs Horas

Se identifica una correlación negativa fuerte entre ausentismo y performance (Fig. 5). Hay también una correlación positiva débil entre horas de estudio y GPA. (Fig. 6) y lo mismo pasa con los estudiantes que poseen tutorías. El resto de las variables, aún asumiendo la posibilidad de relaciones no lineales, no impresionan tener relación con el desempeño.



# Insights y conclusiones

Se realizó el estudio del dataset satisfactoriamente, analizando las distintas variables teniendo en cuenta como variable objetivo la calificación.

## Hallazgos principales:

- La variable Absences presenta una correlación negativa fuerte con GPA, confirma que las ausencias tienen fuerte incidencia en el desempeño.
- Tutoring presenta una correlación positiva débil con el desempeño.
- StudyTimeWeekly presenta una correlación positiva débil con el desempeño.
- El conteo de variables presenta asimetrías en las cantidades que deberán ser tenidas en cuenta si se desea entrenar modelos predictivos.
- No se encontraron factores demográficos que incidan en la nota o diferencias por género.