



**UNIVERSIDAD AUTÓNOMA
DE NUEVO LEÓN**

**FACULTAD DE CIENCIAS
FISICO MATEMÁTICAS**



Licenciatura en Actuaría

Minería de Datos

Grupo 003 – Viernes 6-9

Avance I Proyecto Integrador

Docente Mayra Cristina Berrones Reyes

Equipo #10

1803561	López Peña Marcelo Adrián
1809399	Villarreal Martínez Estibalyz
1863305	Zuñiga Cedillo Milton Humberto

30 de octubre 2020

1) Título de la base de datos

a) Nombre con el que se encuentra en la página encontrada.
Students Performance in Exams

b) Url de la página.

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

2) Descripción de los datos.

a) Qué tipo de datos son (tablas, imágenes, encuestas, etc.)

Encuestas a la escuela para saber los diferentes datos, el género del alumno, los resultados en las diferentes asignaturas, qué tipo de almuerzo se les da a los alumnos, si asistieron al taller de preparación del examen, a que grupo de clase social o etnia pertenece, cuál fue el grado de educación de los padres.

b) Descripción de las columnas. (Que datos se describen en cada columna. Mediciones, tiempo, texto largo, etc.)

-Género. Texto.

-Raza/Etnia. Texto y medición en grupos del A-E u otro.

-Nivel de educación de los padres. Texto (escuela secundaria, algún colegio).

-Almuerzo. Texto diciendo si es estándar o gratis/reducido.

-Curso de preparación para exámenes. Texto (ninguno, completado).

-Puntuación de matemáticas. Medición 0-100.

-Puntuación de lectura. Medición 0-100.

-Puntuación de escritura. Medición 0-100.

3) Justificación del uso de datos.

a) Cuáles fueron las características que les llamó la atención de los datos. Qué les hizo querer trabajar con ellos.

El rendimiento de los estudiantes es un tema que siempre nos ha parecido importante, y una tabla de datos como esta nos puede ayudar a entender más a fondo los factores que afectan en el rendimiento de los estudiantes, y ver que el factor genético realmente tiene una importancia a la hora de obtener resultados nos pareció muy llamativo

b) Qué beneficio encuentran de trabajar con estos datos.

Poder analizar el pasado para predecir el futuro es un trabajo importante para cualquier actuario, y en este caso poder ver el rendimiento de los estudiantes a la hora de un examen nos puede ayudar a encontrar factores que causan una diferencia y poder así en el futuro intentar controlarlos para obtener los mejores resultados

4) Planteamiento del problema.

- a) Tomando en cuenta las características utilizadas para la tarea de análisis de bases de datos, elabora una problemática que te gustaría resolver con tu investigación.

Nos gustaría con nuestra investigación el poder tener al alcance todas las variables que influyen con las interpretaciones necesarias para poder entender cuáles son las mejores condiciones para que los alumnos tengan calificaciones sobresalientes, y ya al tener nuestro sistema bien plantado compartirlo y ponerlo en práctica, así se pueden mejorar las calificaciones de todos los estudiantes.

- b) Nota: Como lo comentado en clase, traten de pensar como una empresa, y cuáles son los problemas que les gustaría mejorar al finalizar su proyecto.

5) Objetivo final.

- a) Explicar a detalle cual es el objetivo principal (y secundarios en caso de existir) para trabajar con este tipo de datos.

Objetivo principal: Identificar qué factores influyen más en el rendimiento académico de los alumnos.

Objetivo secundario: Predecir el rendimiento de un estudiante a partir de los factores a los que se enfrenta día a día.

- b) Nota: Los objetivos secundarios nos servirán para tener varias líneas de investigación abiertas, en caso de que el objetivo principal elegido no pueda completarse.

6) Planeación de la herramienta a utilizar.

- a) Observando el tipo de datos que se tiene, describir cual es el tipo de técnica que se planea utilizar y dar una explicación concisa de por qué se va a trabajar con esa técnica.

Utilizaremos la clasificación para poder cumplir con nuestros objetivos, ya que así podemos identificar cuál es la importancia de cada variable con respecto al rendimiento de los alumnos.

- b) Si ya se tiene observado algún algoritmo o herramienta dentro de esta técnica seleccionada, mencione cual es, y porque se desea trabajar con ella.