



Universidad Autónoma de Baja California

Licenciatura en Ciencias de Datos

Semestre 5

Ian Ricardo Millanes Acosta

375973

Diego Alejandro Valdez Acosta

376064

Análisis de Regresión y Correlación

Prof. Andres Garcia Medina

ANTEPROYECTO

Análisis del Impacto del Estilo de Vida en el Rendimiento Académico y Nivel de Estrés en Estudiantes Universitarios.

INTRODUCCIÓN

El rendimiento académico de los estudiantes universitarios está influenciado por múltiples factores, no solo académicos, sino también relacionados con su estilo de vida. Horas de estudio, sueño, actividad física, vida social y participación en actividades extracurriculares pueden afectar tanto el GPA (Grade Point Average) como el nivel de estrés percibido. Este estudio busca cuantificar estas relaciones mediante técnicas estadísticas como análisis de correlación y modelos de regresión, con el fin de identificar qué comportamientos y hábitos están más asociados con el éxito académico y el bienestar emocional.

PLANTEAMIENTO DEL PROBLEMA

Aunque intuitivamente se asume que estudiar más horas conduce a mejores notas, otros factores como la falta de sueño o el exceso de actividades sociales pueden contrarrestar este beneficio. Además, el estrés puede ser tanto un motivador como un obstáculo para el desempeño académico.

Preguntas de investigación

- ¿Cuáles son las variables del estilo de vida que más se correlacionan con el GPA?
- ¿Cómo se relacionan estas variables con el nivel de estrés?
- ¿Es posible predecir el GPA o el nivel de estrés a partir de los hábitos diarios?

OBJETIVOS

Objetivo General

Analizar la relación entre las variables de estilo de vida (horas de estudio, sueño, actividad física, etc.) y el rendimiento académico (GPA) y nivel de estrés en estudiantes universitarios.

Objetivos Específicos

1. Calcular la matriz de correlación entre todas las variables numéricas y el GPA.
2. Identificar las variables con mayor impacto en el GPA.
3. Analizar la relación entre las variables de estilo de vida y el nivel de estrés.
4. Proponer recomendaciones basadas en los hallazgos para mejorar el rendimiento y reducir el estrés.

ANÁLISIS EXPLORATORIO DEL DATASET

Introducción al dataset

El conjunto de datos que vamos a utilizar está titulado como "Daily Lifestyle and Academic Performance of Students" y contiene datos de 2,000 estudiantes. "Incluye información sobre horas de estudio, actividades extracurriculares, sueño, vida social, actividad física, niveles de estrés y GPA. Los datos cubren un año académico desde agosto de 2023 hasta mayo de 2024 y reflejan los estilos de vida de estudiantes principalmente de la India. Este conjunto de datos puede ayudar a analizar el impacto de los hábitos diarios en el rendimiento académico y el bienestar estudiantil." ([Kaggle - student lifestyle dataset](#))

Columnas

- **Study_Hours_Per_Day:** Son las horas que el estudiante dedica al estudio fuera del horario escolar.
- **Extracurricular_Hours_Per_Day:** Son las actividades realizadas por los estudiantes que no se relacionan directamente con su carrera; como arte, estudiar un idioma, música, etc.
- **Sleep_Hours_Per_Day:** Es la cantidad de horas que el estudiante duerme.
- **Social_Hours_Per_Day:** Es el tiempo donde la persona sale a distraerse y convivir con amigos o familiares.
- **Physical_Activity_Hours_Per_Day:** Es la cantidad de horas dedicadas a la actividad física.
- **GPA (Rango de 0 a 4):** Es la calificación promedio total obtenida por el estudiante.
- **Stress_Level (Bajo, Moderado, Alto):** Es el nivel de estrés autopercibido por el estudiante.

HERRAMIENTAS

Lenguaje: Python (pandas, scikit-learn, statsmodels, matplotlib, seaborn).

Software: Colab (Jupyter Notebook)

RESULTADOS ESPERADOS

- Se espera encontrar una correlación positiva entre Study_Hours y GPA.
- Se anticipa que Sleep_Hours tendrá una correlación positiva con el GPA y negativa con el estrés.
- Social_Hours podría mostrar una relación en forma de U invertida con el GPA.
- El modelo de regresión permitirá cuantificar el impacto de cada variable y predecir el GPA con un margen de error aceptable.

MARCO TEÓRICO

Estudios previos han demostrado que:

- La calidad del sueño está positivamente correlacionada con el rendimiento cognitivo (Okano et al., 2020).
- Un exceso de horas de estudio puede llevar a disminuciones marginales en el aprendizaje debido a la fatiga (Smith, 2019).
- La actividad física moderada mejora la concentración y reduce el estrés (Chen & Liu, 2021).
- El tiempo en redes sociales o vida social puede tener efectos curvilíneos: beneficioso en moderación, perjudicial en exceso (Rodríguez, 2022).
- Este estudio se basará en estos hallazgos, pero se diferenciará por el uso de un dataset amplio (2,000 observaciones) y un enfoque multivariado.

HIPÓTESIS

Partimos de tres suposiciones clave que queremos verificar con datos reales.

- H_1 : Los estudiantes que reportan 7-8 horas de sueño mostrarán GPA significativamente mayor.
- H_2 : La actividad física moderada se asociará con menores niveles de estrés.
- H_3 : Existe un punto óptimo en horas de estudio.

Implicaciones Prácticas

Los resultados podrían ser útiles para:

- Orientación estudiantil: Diseño de talleres sobre gestión del tiempo y hábitos saludables.
- Instituciones educativas: Planificación de horarios y promoción de actividades extracurriculares balanceadas.
- Salud mental universitaria: Identificación de factores de riesgo de estrés académico.

LIMITACIONES

- El dataset no incluye variables demográficas o socioeconómicas.
- Los datos son auto-reportados, por lo que pueden estar sujetos a sesgos.
- La causalidad no puede inferirse sólo con correlación.

CONCLUSIÓN

En no tan pocas pocas palabras, este proyecto busca poner a prueba muchas de las ideas que tenemos sobre lo que realmente ayuda a los estudiantes a tener éxito en la universidad. Todos hemos escuchado consejos como "debes estudiar más" o "duerme bien antes del examen", pero ahora queremos comprobar con datos reales qué tan ciertas son estas creencias y cuánto impacto tiene cada factor.

Al final, lo que esperamos es obtener una imagen clara de cómo diferentes aspectos de la vida estudiantil (desde las horas de estudio hasta el tiempo dedicado al deporte o a la vida social) se relacionan realmente con el rendimiento académico y los niveles de estrés. Esta información puede ser muy valiosa tanto para los estudiantes como para las universidades, ya que permitirá tomar decisiones basadas en evidencia concreta en lugar de solo en suposiciones.

Lo interesante de nuestro enfoque es que no nos limitamos a ver las cosas de forma aislada. Reconocemos que la vida universitaria es un balance entre múltiples actividades, y queremos entender exactamente cómo ese balance afecta los resultados académicos y el bienestar emocional. Los hallazgos de esta investigación podrían ayudar a diseñar mejores programas de apoyo estudiantil y orientar a los universitarios hacia hábitos que realmente funcionen.