

Análisis de los Resultados del Modelo de Regresión Lineal para la Tienda de Ropa

1. Coeficientes del Modelo y su Interpretación

Los coeficientes reflejan la relación entre las variables predictoras y el gasto anual de los clientes. Cada coeficiente positivo indica que, a mayor valor en esas variables, mayor será el gasto anual de los clientes:

- **Avg. Session Length** muestra que por cada unidad adicional de tiempo de sesión, el gasto anual aumenta en 61.43 unidades.
- **Time on App** refleja que cada unidad adicional de tiempo en la aplicación genera un aumento de 39.28 en el gasto anual.
- **Length of Membership** sugiere que un aumento de 1 unidad en la duración de la membresía incrementa el gasto anual en 24.77 unidades.
- El **intercepto** representa el valor de la variable dependiente (gasto anual) cuando todas las variables independientes tienen un valor de cero. El modelo predice que gastaría 500.38 unidades (en este caso, es un valor impracticable [interpretable para clientes nuevos, quizás]; más relacionado con una abstracción matemática).

2. Correlaciones entre las Variables y el Gasto Anual

Las correlaciones lineales entre las variables predictoras y el gasto anual proporcionan más contexto sobre cómo estas variables se relacionan con el comportamiento de compra de los clientes:

- **Correlación fuerte entre 'Length of Membership' y 'Yearly Amount Spent'**: A mayor duración de la membresía, mayor es el gasto anual.
- **Correlación moderada entre 'Time on App' y 'Yearly Amount Spent'**: El tiempo que los clientes pasan en la aplicación también está positivamente correlacionado, aunque en menor medida.
- **Correlación débil entre 'Avg. Session Length' y 'Yearly Amount Spent'**: Aunque hay una relación positiva, es más débil.

3. Evaluación del Modelo: MSE y R²

El **MSE promedio** de 112.09 indica que el modelo realiza predicciones razonablemente precisas, con una desviación estándar de 22.31 que refleja una variabilidad moderada entre los subconjuntos de validación cruzada. El **R² promedio** de 0.981 significa que el modelo explica el **98%** de la variabilidad en el gasto anual, lo que sugiere un excelente ajuste y capacidad predictiva.

Nota: `cross_val_score()` está diseñado para maximizar las puntuaciones, pero el **MAE** funciona al revés: cuanto más bajo, mejor. Para adaptarse a esto, la función devuelve el MAE **negativo**. Esto permite que la función siga maximizando las puntuaciones, y se puede invertir el signo para obtener el valor real del MAE.

4. Relación con los Objetivos del Negocio

El modelo de regresión lineal demuestra que la **Length of Membership** tiene el impacto más fuerte en el gasto anual, seguido del **Time on App**. Para la tienda de ropa, esto implica que las **estrategias de fidelización** son cruciales para maximizar el gasto. Además, mejorar la **experiencia en la app** puede aumentar el tiempo de los usuarios en la plataforma, lo que también contribuye a mayores ingresos. Aunque la relación entre el **Avg. Session Length** y el gasto anual es débil, no debe ser ignorada.

Conclusión

Aunque el coeficiente de **Length of Membership** (24.77) es menor que el de **Avg. Session Length** (61.43), la correlación fuerte entre la primera variable y el gasto anual indica que esta tiene un **impacto acumulativo a largo plazo**. Es decir, los clientes con más tiempo como miembros tienden a gastar más, incluso si el efecto por unidad es menor.

El modelo muestra un ajuste excelente (R² alto) y predicciones precisas (bajo MSE), lo que sugiere que las variables seleccionadas son efectivas para predecir el gasto anual.