

PACE - Proyecto de Regresión Lineal Simple



PACE: Etapa de PLANIFICACIÓN

1. ¿Qué variables disponibles aportan mayor valor para responder el problema?
 - Las variables más relevantes son `bill_length_mm` (predictora) y `body_mass_g` (variable objetivo).
2. ¿En qué unidades se miden estas variables y cómo impactan la interpretación?
 - `bill_length_mm` se mide en milímetros y `body_mass_g` en gramos; esto permite interpretar directamente el impacto por mm.
3. ¿Qué hipótesis inicial orienta el análisis?
 - A mayor longitud del pico, mayor masa corporal.
4. ¿La calidad del dato es suficiente?
 - Sí, el dataset presenta consistencia y no tiene faltantes relevantes para este modelo.
5. ¿Qué prácticas EDA son necesarias?
 - Visualización de correlaciones, regplot, pairplot y estadísticos descriptivos.



PACE: Etapa de ANÁLISIS

6. ¿Qué pasos de EDA se requieren?
 - Explorar distribuciones, revisar correlaciones y relaciones lineales, identificar outliers.
7. ¿Es necesario integrar, filtrar o reorganizar datos?
 - No se necesitó integración; solo filtrado de filas con valores nulos.

8. ¿Qué visualizaciones comunican mejor los hallazgos?

- Scatterplot con regresión, pairplot y resumen estadístico OLS.



PACE: Etapa de CONSTRUCCIÓN

9. ¿Qué modelos y visualizaciones deben construirse?

- Regresión lineal OLS, scatterplot, pairplot y tabla de resumen estadístico.

10. ¿Qué pasos se siguen para construir visualizaciones claras?

- Limpieza del dataset, generación de gráficos exploratorios y ajuste del modelo.

11. ¿Qué variables aportan mayor valor a las visualizaciones finales?

- bill_length_mm y body_mass_g, con flipper_length_mm como variable secundaria prometedora.

12. ¿Cómo se gestionan los datos faltantes?

- Se eliminaron filas con datos nulos para mantener calidad del análisis.



PACE: Etapa de EJECUCIÓN

13. ¿Qué insights clave surgieron?

- Cada 1 mm adicional del pico aumenta ~141 g de masa corporal; $R^2 = 0.769$.

14. ¿Qué recomendaciones se derivan?

- Usar el modelo para estimaciones rápidas y explorar modelos multivariados.

15. ¿Qué preguntas adicionales pueden explorarse?

- Efecto por especie, impacto de flipper_length_mm y predicción de otras métricas.

16. ¿Cómo deben compartirse los resultados?

- Tabla OLS para técnicos, gráfica simple para ejecutivos y dashboard para operativos.