

Evaluación EDA

Diplomado en Análisis de datos con R para la acuicultura | Pontificia Universidad Católica de Valparaíso

Dr. José Gallardo Matus y Dra. María Angélica Rueda

03 June 2022

PUNTAJE TOTAL Y NOTA

Table 1: Puntaje obtenido y nota.

Nombre	Puntaje obtenido	Nota
Gabriela Soledad Salas Serqueira	25	5.4

PARÁMETROS PARA REALIZAR CÁLCULO DE LA NOTA

Table 2: Parámetros para el cálculo de la nota.

Item	Valor
Nota máxima	7,0
Nota mínima	1,0
Nota aprobación	4,0
Exigencia para aprobar	60 %
Puntaje máximo	32

Table 3: Puntaje según nivel de logro alcanzado en cada pregunta o ejercicio.

Nivel	Puntaje
Excelente	4
Muy bueno	3
Bueno	2
Suficiente	1
Insuficiente	0

DETALLE DE PUNTAJE OBTENIDO POR PREGUNTA

Table 4: Puntaje obtenido por pregunta.

Pregunta	Puntaje
1 - Evidencias uso de control de versiones en github	4
2 - Archivo README en github	3
3 - Describe variación usando histogramas	4
4 - Identifica si los datos están balanceados	2
5 - Establece relación entre variables cuantitativas y factores usando gráficas de correlación, boxplot, interacción o de tamaño de los efectos.	4
6 - Identifica si existen errores, datos faltantes o valores atípicos	2
7 - Resume los datos usando tablas y estadística descriptiva.	2
8 - Utiliza librerías readxl, dplyr, ggplot2 o similar según naturaleza de los datos	4

Table 5: Evaluación cualitativa y retroalimentación EDA.

Retroalimentación con fines de aprendizaje.
<p>Buen trabajo, faltó interpretar balance de datos y valores atípicos, datos faltantes y posibles errores en la toma de datos. ¿Qué utilidad tiene la tabla <code>table(tidy_MortalidadCentro, tidy_MortalidadBarrio)</code>? ¿Saber en que barrio está un centro? Sugiero eliminar. Evite el color verde que usó, prefiera tonos que den apariencia más profesional. Imite colores que hay en los informes o reportes que realizan las gerencias de las empresas de su sector. https://www.multi-xsalmon.com/inversionistas/presentaciones/. La función <code>abline(lm(tidy_MortalidadMortO2 tidy_MortalidadOxigeno), col = "red", lwd = 2,</code> <code>)</code> está mal utilizada porque no se observa una relación lineal. Quizas podría agrupar oxígeno en rangos y hacer boxplot. Una grafica de interacción está mal implementada pues se debe realizar con 2 variables categóricas, oxígeno es cuantitativa continua. Faltan tablas con resumen de variables, se sugiere <code>group_by()</code> y <code>summarise()</code>.</p>