

# 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
Juan José Rodríguez Maulén	20	4.2

## 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	4
3 - Describe variación usando histogramas	4
4 - Identifica si los datos están balanceados	0
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	0
7 - Resume los datos usando tablas y estadística descriptiva.	2
8 - Utiliza librerías readxl, dplyr, ggplot2 o similar según naturaleza de los datos	2

Table 5: Evaluación cualitativa y retroalimentación EDA.

Retroalimentación con fines de aprendizaje.
Buen trabajo, dado que las variables son binarias es necesario incluir tablas de frecuencia. Es muy extraño el resultado obtenido (mas madurez con fotoperíodo) ¿habrá algún error en los datos?. OJO que este comandos <code>Ma&lt;-table(Madurez\$%<b>Mature</b>)</code> al parecer esta dando un error cuando se aplica posteriormente <code>mean()</code> . Elimine mensajes y alertas del reporte con <code>message=FALSE</code> , <code>warning=FALSE</code> . Para alcanzar un nivel avanzado en el diplomado, todos los histogramas y barplot se deben realizar con paquete <code>ggplot2</code> . Use paneles para agrupar graficas relacionadas. Faltan tablas con resumen de variables, se sugiere <code>group_by()</code> y <code>summarise()</code> .