

LABORATORIO 4

Análisis visual de datos interdisciplinarios

Visualización, interpretación y comunicación de resultados

Escuela de Informática

Información del laboratorio

Curso: EIY403 - Introducción al análisis de datos para otras carreras

Valor: 10 % de la nota final

Modalidad: Grupal (equipos de 2-5 personas)

Duración: 1 sesión de clase (hoy)

Fecha: Miércoles, 22 de octubre de 2025

Entregables:

- Documento R Markdown (.Rmd + .html compilado)
- Presentación oral (10-12 minutos)

Presentaciones: Hoy mismo, 22 de octubre de 2025

Entrega documental: Correo institucional a jordy.alfaro.brenes@una.cr

Asunto: Lab4.Caso[número] _NombreEquipo

1. Introducción y objetivos

Este laboratorio final integra todos los conocimientos adquiridos durante el curso, con énfasis especial en:

1. **Visualización efectiva de datos:** Aplicación de principios de diseño gráfico, buenas prácticas y selección apropiada de tipos de gráficos
2. **Análisis crítico interdisciplinario:** Investigación de contextos en diferentes áreas del conocimiento y aplicación del análisis de datos en diversos campos
3. **Comunicación de resultados:** Presentación clara y profesional de hallazgos mediante documentos reproducibles y exposiciones orales
4. **Trabajo en equipo:** Colaboración efectiva en análisis de datos complejos

Enfoque del laboratorio:

Este laboratorio prioriza la **interpretación**, el **análisis crítico** y la **comunicación efectiva**. Cada equipo trabajará con un caso de una disciplina diferente, lo que les permitirá ver la aplicabilidad universal del análisis de datos.

Importante:

- Se proporciona únicamente el **dataset** (archivo CSV)
- Deben crear el **código R** necesario para el análisis (pueden usar IA)
- Si usan IA, deben **documentar los prompts** utilizados
- El enfoque está en **crear gráficos apropiados** según buenas prácticas
- La evaluación principal será la **calidad de la interpretación** y la **presentación oral**
- Se requiere **investigación** del contexto disciplinario

2. Estructura general del trabajo

Todos los equipos deben seguir la misma estructura, independientemente del caso asignado:

2.1. Documento R Markdown (50 puntos)

1. Portada y metadatos (5 puntos)

- Título del caso
- Nombres completos de todos los integrantes
- Número de caso asignado
- Fecha
- YAML correctamente configurado con índice flotante, code folding, y theme

2. Contexto disciplinario (10 puntos)

Investigación seria sobre el tema:

- Introducción a la disciplina (¿de qué trata este campo?)
- Conceptos clave (3-5 términos técnicos explicados)
- Relevancia del problema
- Aplicaciones prácticas
- Referencias (mínimo 2 fuentes confiables en formato APA)

Extensión: 1-2 páginas (400-600 palabras)

3. Análisis de datos (20 puntos)

- Carga y exploración inicial del dataset
- Creación de gráficos apropiados (2-4 gráficos)
- Aplicación de buenas prácticas de visualización
- Documentación de prompts de IA utilizados (si aplica)

4. Interpretación de resultados (10 puntos)

- Interpretación de cada gráfico
- Respuestas a las preguntas específicas del caso
- Identificación de patrones o tendencias
- Limitaciones del análisis

Extensión: 1-2 páginas

5. Recomendaciones y conclusiones (5 puntos)

- Acciones recomendadas basadas en resultados
- Análisis adicionales sugeridos
- 3-5 conclusiones principales

2.2. Presentación oral (50 puntos)

Duración: 10-12 minutos + 3-5 minutos de preguntas

Estructura:

1. Introducción (2 min): Presentación del equipo y del caso
2. Contexto disciplinario (2-3 min): Conceptos clave y relevancia
3. Análisis visual (4-5 min): Presentación de gráficos y hallazgos
4. Interpretación y recomendaciones (2-3 min)
5. Preguntas y respuestas (3-5 min)

3. Casos específicos

Instrucción importante:

Cada equipo recibirá su dataset específico. Las siguientes secciones detallan las instrucciones para cada uno de los 8 casos. **Solo deben consultar las instrucciones de su caso asignado.**

3.1. Caso 1: Medicina - Eficacia de tratamiento antihipertensivo

Disciplina: Medicina / Farmacología

Dataset: `caso1_hipertension.csv`

Variables:

- `paciente_id`: Identificador del paciente
- `grupo`: Tratamiento (nuevo_farmaco) o Control (placebo)
- `presion_inicial`: Presión arterial sistólica antes del tratamiento (mmHg)
- `presion_final`: Presión arterial sistólica después de 8 semanas (mmHg)
- `edad`: Edad del paciente (años)
- `sexo`: Sexo del paciente (M/F)

Tamaño: 100 pacientes (50 por grupo)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es la hipertensión arterial y por qué es importante tratarla?
2. ¿Qué son los ensayos clínicos controlados con placebo?
3. ¿Qué significa que un tratamiento sea estadísticamente significativo?
4. ¿Qué es la presión arterial sistólica y cuáles son los rangos normales?
5. ¿Cómo se evalúa la eficacia de un medicamento antihipertensivo?

Análisis requerido

Deben crear los siguientes gráficos:

1. **Boxplots comparativos:** Comparar la reducción de presión (`presión_final` - `presión_inicial`) entre el grupo tratamiento y el grupo control. Incluir línea de referencia en 0.
2. **Gráfico de dispersión:** Presión inicial vs presión final, con puntos coloreados por grupo y línea de identidad ($y=x$) para ver quiénes mejoraron.
3. **Gráfico de barras con error:** Media y error estándar de la reducción de presión para cada grupo.
4. **(Opcional) Histogramas:** Distribución de la reducción de presión para cada grupo (facetado).

Preguntas de investigación

Respondan en su interpretación:

1. ¿El nuevo fármaco reduce significativamente la presión arterial comparado con el placebo?
2. ¿Cuál es la reducción promedio en cada grupo?
3. ¿Hay pacientes en el grupo tratamiento que no respondieron al medicamento?
4. ¿Hay pacientes en el grupo placebo que mejoraron (efecto placebo)?
5. ¿Considerarían este medicamento efectivo? ¿Por qué?
6. ¿Qué factores adicionales deberían considerarse antes de aprobar el medicamento?

Tip para visualización:

En los boxplots, consideren usar colores que representen salud (verde para mejoría, rojo para problema). Asegúrense de incluir las unidades (mmHg) en todos los ejes.

3.2. Caso 2: Agronomía - Comparación de fertilizantes en cultivo de maíz

Disciplina: Agronomía / Ciencias Agrícolas

Dataset: caso2_fertilizantes.csv

Variables:

- **parcela_id:** Identificador de la parcela
- **fertilizante:** Tipo (Control, Organico, Quimico_A, Quimico_B)
- **rendimiento:** Producción en kg por hectárea
- **zona:** Ubicación geográfica (Norte, Sur, Este, Oeste)
- **lluvia_mm:** Precipitación acumulada en mm durante el ciclo
- **costo_por_ha:** Costo del fertilizante por hectárea (colones)

Tamaño: 80 parcelas (20 por tipo de fertilizante)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es el rendimiento agrícola y cómo se mide?
2. Diferencias entre fertilizantes orgánicos y químicos
3. ¿Qué es un grupo control en experimentos agrícolas?
4. Factores que afectan el rendimiento del maíz
5. ¿Qué es la agricultura sostenible?

Análisis requerido

Deben crear los siguientes gráficos:

1. **Boxplots:** Rendimiento por tipo de fertilizante, ordenados de mayor a menor rendimiento promedio.
2. **Gráfico de barras agrupadas:** Rendimiento promedio por fertilizante y por zona, para ver si hay interacción.
3. **Gráfico de dispersión:** Rendimiento vs costo por hectárea, coloreado por tipo de fertilizante, para evaluar costo-beneficio.
4. **(Opcional) Dispersión:** Rendimiento vs precipitación, facetado por tipo de fertilizante.

Preguntas de investigación

Respondan en su interpretación:

1. ¿Qué fertilizante produce el mayor rendimiento promedio?
2. ¿Hay diferencias significativas entre los fertilizantes?
3. ¿El rendimiento varía según la zona geográfica?

4. Considerando costo y rendimiento, ¿cuál fertilizante recomendarían?
5. ¿El fertilizante orgánico es competitivo con los químicos?
6. ¿Qué papel juega la precipitación en los resultados?

Tip para análisis:

Calculen el ingreso neto restando el costo del fertilizante al valor de la producción (asuman que el kilo de maíz vale ₡300). Esto ayudará en la recomendación.

3.3. Caso 3: Economía - Inflación en países centroamericanos

Disciplina: Economía / Macroeconomía

Dataset: caso3_inflacion.csv

Variables:

- **pais:** País (Costa Rica, Guatemala, Honduras, El Salvador, Nicaragua)
- **año:** Año (2020-2024)
- **mes:** Mes (1-12)
- **fecha:** Fecha completa
- **inflacion_mensual:** Tasa de inflación mensual (%)
- **inflacion_acumulada:** Inflación acumulada anual (%)
- **pib_variacion:** Variación del PIB trimestral (%)

Tamaño: 300 observaciones (5 países × 60 meses)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es la inflación y cómo se mide?
2. ¿Qué es el IPC (Índice de Precios al Consumidor)?
3. Diferencia entre inflación mensual y acumulada
4. ¿Qué causas comunes tiene la inflación?
5. ¿Por qué es importante controlar la inflación?
6. ¿Qué es el PIB y cómo se relaciona con la inflación?

Análisis requerido

Deben crear los siguientes gráficos:

1. **Serie de tiempo múltiples:** Inflación mensual a lo largo del tiempo para los 5 países (una línea por país). Incluir línea de referencia en 0 %.
2. **Gráfico de líneas facetado:** Inflación acumulada anual por país, separado en facetas para comparar mejor.
3. **Boxplots:** Distribución de inflación mensual por país para ver variabilidad y valores extremos.
4. **Gráfico de dispersión:** Inflación mensual vs variación del PIB para explorar la relación entre ambas variables.

Preguntas de investigación

Respondan en su interpretación:

1. ¿Qué país tuvo mejor control de la inflación en el período analizado?
2. ¿Se observan tendencias regionales comunes?
3. ¿Hubo algún período de crisis inflacionaria? ¿Cuándo y en qué países?
4. ¿Qué país mostró mayor variabilidad (inestabilidad) en su inflación?
5. ¿Existe relación visible entre crecimiento del PIB e inflación?
6. ¿Qué eventos económicos globales podrían explicar los patrones observados?

Tip para el contexto:

Investiguen qué pasó económicamente en la región entre 2020-2024 (COVID-19, crisis de suministro, etc.) para contextualizar los picos de inflación.

3.4. Caso 4: Física - Ley de enfriamiento de Newton

Disciplina: Física / Termodinámica

Dataset: caso4_enfriamiento.csv

Variables:

- **tiempo_min:** Tiempo transcurrido en minutos
- **temperatura_C:** Temperatura del objeto en grados Celsius
- **temp_ambiente:** Temperatura ambiente (constante, 22°C)
- **experimento:** Número de repetición del experimento (1, 2, 3)

Tamaño: 180 mediciones (3 experimentos \times 60 minutos)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es la Ley de Enfriamiento de Newton?
2. Fórmula matemática de la ley: $T(t) = T_a + (T_0 - T_a)e^{-kt}$
3. ¿Qué es la constante de enfriamiento k ?
4. ¿Qué significa decaimiento exponencial?
5. Aplicaciones prácticas de esta ley (forense, ingeniería, cocina, etc.)

Análisis requerido

Deben crear los siguientes gráficos:

1. **Gráfico de líneas múltiples:** Temperatura vs tiempo para los tres experimentos, con línea horizontal en la temperatura ambiente.
2. **Gráfico de dispersión con curva:** Temperatura vs tiempo (todos los datos juntos) con una curva de tendencia exponencial ajustada.
3. **Gráfico log-lineal:** $\ln(T - T_a)$ vs tiempo, que debería ser lineal si se cumple la ley.
4. **(Opcional) Facetado:** Los tres experimentos en paneles separados para comparar reproducibilidad.

Preguntas de investigación

Respondan en su interpretación:

1. ¿Los datos experimentales siguen un patrón de decaimiento exponencial?
2. ¿Los tres experimentos son reproducibles (dan resultados similares)?
3. ¿Se cumple la Ley de Enfriamiento de Newton en este caso?
4. Aproximadamente, ¿cuál es el tiempo de vida media (tiempo para llegar a la mitad entre T_0 y T_a)?

5. ¿En qué momento la temperatura se estabiliza cerca de la temperatura ambiente?
6. ¿Qué factores experimentales podrían causar desviaciones de la ley?

Tip matemático:

Para crear el gráfico log-lineal, calculen una nueva variable: `log_diff <- log(temperatura_C - temp_ambiente)`. Si la ley se cumple, este gráfico vs tiempo será una línea recta.

3.5. Caso 5: Psicología - Tiempo de reacción y envejecimiento

Disciplina: Psicología / Neurociencia Cognitiva

Dataset: caso5_tiempo_reaccion.csv

Variables:

- **participante_id:** Identificador del participante
- **edad:** Edad en años (18-85)
- **sexo:** Sexo (M/F)
- **tiempo_reaccion_ms:** Tiempo promedio de reacción en milisegundos
- **escolaridad:** Años de educación formal
- **actividad_fisica:** Nivel (Sedentario, Moderado, Activo)

Tamaño: 120 participantes

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es el tiempo de reacción y cómo se mide?
2. ¿Cómo afecta el envejecimiento a las funciones cognitivas?
3. ¿Qué es el procesamiento de información en psicología?
4. Factores que influyen en el tiempo de reacción
5. ¿Por qué es importante estudiar esto? (conducción, seguridad, etc.)

Análisis requerido

Deben crear los siguientes gráficos:

1. **Gráfico de dispersión:** Edad vs tiempo de reacción con línea de tendencia (regresión lineal o suavizada). Colorear puntos por sexo.
2. **Boxplots agrupados:** Tiempo de reacción por grupo etario (18-30, 31-50, 51-70, 71-85) y por nivel de actividad física.
3. **Gráfico de violín:** Distribución del tiempo de reacción por décadas de edad (20s, 30s, 40s, etc.).
4. **Gráfico de dispersión:** Escolaridad vs tiempo de reacción para explorar si la educación es un factor protector.

Preguntas de investigación

Respondan en su interpretación:

1. ¿Cómo afecta la edad al tiempo de reacción?
2. ¿A partir de qué edad se observa un incremento notable?

3. ¿Hay diferencias entre hombres y mujeres?
4. ¿La actividad física se asocia con mejores tiempos de reacción?
5. ¿Los años de escolaridad influyen en el tiempo de reacción?
6. ¿Qué recomendaciones darían para mantener buenos tiempos de reacción en adultos mayores?

Tip para el análisis:

Crear grupos etarios facilitará la comparación. Usen `cut()` en R:

```
grupo_edad <- cut(edad, breaks=c(18,30,50,70,85), labels=...)
```

3.6. Caso 6: Ciencias ambientales - Calidad del aire en San José

Disciplina: Ciencias Ambientales / Salud Pública

Dataset: caso6_calidad_aire.csv

Variables:

- **fecha:** Fecha de medición
- **dia_semana:** Día de la semana
- **PM25:** Concentración de material particulado 2.5 ($\mu g/m^3$)
- **estacion:** Ubicación (Centro, Norte, Sur, Este, Oeste)
- **hora:** Hora del día (promedio diario de mediciones horarias)
- **temperatura:** Temperatura promedio del día ($^{\circ}C$)
- **precipitacion:** Precipitación del día (mm)

Tamaño: 450 observaciones (5 estaciones \times 90 días)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es el PM2.5 y por qué es peligroso?
2. Límites de la OMS para PM2.5 (promedio diario: $15 \mu g/m^3$, anual: $5 \mu g/m^3$)
3. ¿Qué es el Índice de Calidad del Aire (ICA)?
4. Fuentes de contaminación del aire en áreas urbanas
5. Efectos en la salud de la exposición a PM2.5

Análisis requerido

Deben crear los siguientes gráficos:

1. **Serie temporal:** PM2.5 a lo largo de los 90 días, con línea horizontal en el límite de la OMS ($15 \mu g/m^3$) y zonas coloreadas (verde: bueno, amarillo: moderado, rojo: malo).
2. **Boxplots:** Distribución de PM2.5 por estación, ordenados de más a menos contaminado.
3. **Gráfico de barras:** PM2.5 promedio por día de la semana (para ver si hay patrón fin de semana vs días laborales).
4. **Gráfico de dispersión facetado:** PM2.5 vs precipitación, separado por estación, para ver el efecto de la lluvia.

Preguntas de investigación

Respondan en su interpretación:

1. ¿Cuántos días se excedió el límite recomendado por la OMS?
2. ¿Qué estación/zona tiene peor calidad del aire? ¿Por qué podría ser?

3. ¿Hay diferencias entre días laborales y fines de semana?
4. ¿Cómo afecta la precipitación a la concentración de PM_{2.5}?
5. ¿Qué porcentaje del tiempo la calidad del aire es "buena" según estándares internacionales?
6. ¿Qué medidas de política pública recomendarían para mejorar la calidad del aire?

Tip para visualización:

Para las zonas de calidad del aire, pueden usar:

Verde (Bueno): 0-12 $\mu g/m^3$

Amarillo (Moderado): 12-35 $\mu g/m^3$

Rojo (Malo): $\geq 35 \mu g/m^3$

3.7. Caso 7: Ingeniería - Análisis de fatiga en materiales

Disciplina: Ingeniería de Materiales / Ingeniería Mecánica

Dataset: caso7_fatiga_materiales.csv

Variables:

- probeta_id: Identificador de la probeta de ensayo
- material: Tipo de acero (Acero_A, Acero_B, Acero_C)
- estres_MPa: Estrés aplicado en megapascuales
- ciclos_falla: Número de ciclos hasta la falla
- temperatura_ensayo: Temperatura del ensayo (°C)
- tratamiento_termico: Tipo (Sin_tratamiento, Templado, Recocido)

Tamaño: 90 probetas (30 por material)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es la fatiga de materiales?
2. ¿Qué es una curva S-N (Wöhler)?
3. ¿Qué es el límite de fatiga o resistencia a la fatiga?
4. ¿Cómo se realiza un ensayo de fatiga?
5. Importancia de estos ensayos en diseño de estructuras (puentes, aviones, etc.)

Análisis requerido

Deben crear los siguientes gráficos:

1. **Gráfico log-log (Curva S-N):** Estrés vs ciclos hasta falla en escala logarítmica, con puntos coloreados por material. Este es el gráfico más importante del análisis.
2. **Boxplots:** Ciclos hasta falla por material a niveles de estrés similares (agrupen en rangos: bajo, medio, alto).
3. **Gráfico de barras agrupadas:** Ciclos promedio hasta falla por material y tratamiento térmico.
4. **Gráfico de dispersión:** Temperatura de ensayo vs ciclos hasta falla para explorar efecto de temperatura.

Preguntas de investigación

Respondan en su interpretación:

1. ¿Qué material tiene mejor resistencia a la fatiga?
2. ¿Se observa una relación clara entre estrés aplicado y vida útil?

3. ¿Existe un límite de fatiga aparente (estrés bajo el cual no hay falla)?
4. ¿Cómo afecta el tratamiento térmico a la resistencia a la fatiga?
5. Si tuvieran que diseñar un componente sometido a cargas cíclicas, ¿qué material recomendarían? ¿Por qué?
6. ¿Qué consideraciones de seguridad aplicarían basándose en estos resultados?

Tip para gráfico log-log:

En ggplot2: `scale_x_log10()` y `scale_y_log10()`

La curva S-N típicamente muestra una relación lineal en escala log-log.

3.8. Caso 8: Biología - Crecimiento poblacional bacteriano

Disciplina: Microbiología / Biología Celular

Dataset: caso8_crecimiento_bacterias.csv

Variables:

- **tiempo_h:** Tiempo en horas desde el inicio
- **DO600:** Densidad óptica a 600 nm (medida de concentración)
- **medio_cultivo:** Tipo de medio (Rico, Minimo, Estres)
- **temperatura:** Temperatura de incubación (°C)
- **antibiotico:** Presencia de antibiótico (Si/No)
- **replica:** Número de réplica experimental (1, 2, 3)

Tamaño: 432 mediciones (3 medios \times 2 condiciones \times 3 réplicas \times 24 horas)

Contexto a investigar

Deben investigar y explicar:

1. ¿Qué es la curva de crecimiento bacteriano?
2. Fases del crecimiento: lag, exponencial, estacionaria, muerte
3. ¿Qué es el tiempo de duplicación?
4. ¿Cómo se mide el crecimiento bacteriano? (DO600, UFC)
5. Aplicaciones de estos estudios (biotecnología, medicina, etc.)

Análisis requerido

Deben crear los siguientes gráficos:

1. **Gráfico de líneas facetado:** DO600 vs tiempo, separado por medio de cultivo. Incluir las 3 réplicas para ver reproducibilidad. Marcar las fases de crecimiento.
2. **Gráfico semi-logarítmico:** $\log(\text{DO600})$ vs tiempo para la fase exponencial (facilita cálculo de tasa de crecimiento). Solo medio Rico sin antibiótico.
3. **Gráfico de líneas múltiples:** Comparación directa entre condición con y sin antibiótico para el mismo medio de cultivo.
4. **Gráfico de barras:** DO600 máxima alcanzada por cada combinación de medio y antibiótico.

Preguntas de investigación

Respondan en su interpretación:

1. ¿En qué medio de cultivo las bacterias crecen mejor?
2. ¿Cuánto dura aproximadamente la fase lag en cada condición?

3. ¿Cuándo comienza la fase estacionaria?
4. ¿Cómo afecta el antibiótico al crecimiento? ¿Es bactericida o bacteriostático?
5. Aproximadamente, ¿cuál es el tiempo de duplicación en la fase exponencial del medio rico?
6. ¿Las réplicas muestran buena reproducibilidad experimental?
7. ¿Qué implicaciones tienen estos resultados para la industria biotecnológica?

Tip para el análisis:

La fase exponencial se identifica donde $\log(\text{DO600})$ vs tiempo es lineal. La pendiente de esa recta es la tasa de crecimiento μ , y el tiempo de duplicación se calcula como: $t_d = \frac{\ln(2)}{\mu}$

4. Criterios de evaluación detallados

Criterio	Puntos	Descripción
DOCUMENTO R MARKDOWN (50 pts)		
Portada y metadatos	5	
- YAML correctamente configurado	2	Índice, code folding, theme
- Información completa del equipo	2	Nombres, caso, fecha
- Presentación profesional	1	Formato limpio y organizado
Contexto disciplinario	10	
- Investigación de calidad	4	Fuentes confiables, información precisa
- Conceptos clave bien definidos	3	Términos técnicos explicados claramente
- Relevancia y aplicaciones	2	Justificación de importancia
- Referencias en formato APA	1	Mínimo 2 fuentes apropiadas
Análisis de datos	20	
- Carga y exploración de datos	2	Código funcional, verificación
- Gráficos apropiados creados	8	Tipos correctos según datos
- Aplicación de buenas prácticas	6	Títulos, colores, formato profesional
- Documentación de prompts IA	2	Si usaron IA, prompts documentados
- Código limpio y comentado	2	Código legible y reproducible
Interpretación de resultados	10	
- Respuestas a preguntas clave	4	Todas las preguntas respondidas
- Profundidad del análisis	3	Interpretación va más allá de lo obvio
- Identificación de patrones	2	Encuentra tendencias/anomalías
- Limitaciones mencionadas	1	Reconoce límites del análisis
Recomendaciones	5	
- Calidad de recomendaciones	3	Fundamentadas en resultados
- Conclusiones claras	2	3-5 puntos concretos
PRESENTACIÓN ORAL (50 pts)		
Claridad y estructura	15	
- Organización de contenido	5	Flujo lógico, transiciones claras
- Manejo del tiempo	5	Respeto límite de 10-12 minutos
- Calidad de diapositivas	5	Visuales efectivas, no saturadas
Contexto disciplinario	10	
- Explicación clara del campo	5	Conceptos comprensibles
- Relevancia del problema	5	Justificación convincente
Análisis visual	10	
- Presentación de gráficos	5	Gráficos legibles y bien explicados
- Destacar hallazgos clave	5	Identifican lo más importante
Interpretación	10	
- Profundidad de análisis	5	Interpretación va más allá de lo obvio
- Calidad de recomendaciones	5	Propuestas concretas y fundamentadas
Participación y dominio	5	

Criterio	Puntos	Descripción
- Participación equitativa	2	Todos los integrantes participan
- Dominio del tema	2	Demuestran haber investigado
- Respuestas a preguntas	1	Manejan preguntas apropiadamente
TOTAL	100	

5. Especificaciones técnicas de entrega

Requisitos obligatorios para la entrega:

Archivos a entregar:

1. Lab4_Caso[N]_NombreEquipo.Rmd (código fuente)
2. Lab4_Caso[N]_NombreEquipo.html (documento compilado)

Correo de entrega:

- **Destinatario:** jordy.alfaro.brenes@una.cr
- **Asunto:** Lab4_Caso[número]_NombreEquipo
- **Ejemplo:** Lab4_Caso3_EquipoAlfa
- **Cuerpo del correo:**
 - Lista completa de integrantes
 - Número de caso asignado
 - Título del caso

Fecha límite de entrega: Hoy, 22 de octubre de 2025, antes de las 11:59 PM

5.1. YAML header requerido

```
---
title: "Laboratorio 4: [Titulo del Caso]"
subtitle: "Caso [numero]: [Disciplina]"
author:
  - "Integrante 1"
  - "Integrante 2"
  - "Integrante 3"
date: "22 de octubre de 2025"
output:
  html_document:
    toc: true
    toc_float: true
    toc_depth: 3
    code_folding: show
    theme: flatly
    highlight: tango
    df_print: paged
```

6. Recordatorios finales

Aspectos críticos para el éxito:

1. **Investigación del contexto:** No subestimen esta parte. Un buen contexto demuestra profesionalismo y comprensión del problema.
2. **Gráficos apropiados:** Seleccionen el tipo correcto de gráfico según los datos y la pregunta a responder.
3. **Buenas prácticas visuales:** Apliquen lo aprendido en clase: títulos descriptivos, colores apropiados, sin chartjunk, etiquetas completas.
4. **Interpretación profunda:** Vayan más allá de describir lo que se ve. Respondan el "¿qué significa?" y "¿por qué es importante?".
5. **Documentación de IA:** Si usan IA para generar código, documenten el prompt usado. Ejemplo:

```
## Gráfico 1: Boxplots comparativos
```

```
**Prompt usado:** "Genera código en R con ggplot2  
para crear boxplots comparando la variable X entre  
grupos A y B. Usa colores profesionales y agrega  
títulos en español."
```

6. **Práctica de la presentación:** Ensayen para respetar el tiempo de 10-12 minutos.
7. **Trabajo en equipo:** Todos deben participar tanto en el análisis como en la presentación.
8. **Entrega a tiempo:** La puntualidad es parte del profesionalismo.

Sobre uso de IA y plagio:**Uso apropiado de IA:**

- Pueden usar IA para generar código R
- Deben documentar todos los prompts utilizados
- Deben verificar que el código funcione
- Deben entender qué hace el código

Interpretaciones y contexto:

- Las interpretaciones deben ser propias, no copiadas de IA
- El contexto disciplinario debe investigarse en fuentes académicas
- Eviten copiar textos directamente de IA o de otros equipos

Nota de 0 por:

- Copiar trabajo de otros equipos
- Presentar interpretaciones idénticas entre equipos
- No documentar el uso de IA cuando se utilizó

Este laboratorio final integra habilidades técnicas, investigación interdisciplinaria y comunicación efectiva, preparándolos para aplicar el análisis de datos en contextos profesionales diversos.

Universidad Nacional de Costa Rica

Escuela de Informática - EIY403

II Semestre 2025

¡Éxito en su laboratorio final!