

Guía de Repaso para la Prueba 2: Descripción y Visualización de Datos

Formato de la prueba

La Prueba 2 se realizará en RStudio. Al finalizar, deberás guardar tu trabajo en un archivo con extensión `.R` y subirlo a Webcursos dentro del plazo indicado.

Trabajarás con un dataset real que se te entregará al inicio. A partir de ese dataset deberás responder cinco preguntas, aplicando las herramientas vistas en el Módulo 2: `dplyr` y `ggplot2`.

Las siguientes preguntas están diseñadas para que identifiques tus puntos fuertes y débiles. Considera que **en la prueba no tendrás apoyo de la IA ni acceso a material adicional**. Todo deberá salir de tu memoria y comprensión de lo practicado en clases.

Cada pregunta te pedirá que generes una visualización o un resumen numérico, justificando tus decisiones. Las preguntas estarán diseñadas para evaluar:

1. Tu capacidad de manipular datos con `dplyr`.
2. Tu capacidad de elegir el tipo de gráfico adecuado según el tipo de variables y el objetivo.
3. Tu claridad visual y comunicativa en las gráficas.

Preguntas para guiar tu repaso

Te sugiero empezar tu estudio de la siguiente forma:

- Resuelve la guía en el siguiente link:
<https://colab.research.google.com/drive/1oiMCi1ZmuAHIDQLNYoC1nZUwzJ72RxuA?usp=sharing>
- Luego, basado en las preguntas que te costaron menos o más, marca con una ☒ aquellas áreas donde te sientas seguro, y deja sin marcar las que necesites reforzar.
- Tu esfuerzos de estudio deberán estar enfocados en las áreas que te cuestan más

1. Elección del tipo de gráfico

- ☐ ¿Sé qué tipo de gráfico usar si quiero comparar promedios o totales por grupo?
- ☐ ¿Sé qué gráfico usar para identificar tendencias en el tiempo?

- ☐ ¿Sé cómo mostrar la distribución de una variable numérica entre grupos?
- ☐ ¿Sé cómo representar la relación entre dos variables numéricas?

2. Transformaciones con dplyr

- ☐ ¿Sé cómo filtrar observaciones según una condición?
- ☐ ¿Sé cómo filtrar los valores más altos o bajos según una variable?
- ☐ ¿Sé cómo crear nuevas variables a partir de otras existentes?
- ☐ ¿Sé cómo agrupar los datos por categoría antes de resumir?
- ☐ ¿Sé cómo calcular promedios o totales por grupo?
- ☐ ¿Sé cómo ordenar los resultados para facilitar la lectura?

3. Uso de más de una variable en ggplot

- ☐ ¿Sé cómo agregar una tercera variable a un gráfico, usando color, fill, size o shape?
- ☐ ¿Sé cuándo conviene usar cada una de estas opciones?
- ☐ ¿Sé cómo usar el argumento `aes()` dentro de las capas secundarias para aplicar estéticas diferentes por capa?

4. Faceting: generar un gráfico por categoría

- ☐ ¿Sé cómo generar un panel distinto por cada categoría de una variable?
- ☐ ¿Sé cómo cruzar dos variables categóricas en una cuadrícula de paneles?

5. Personalización y claridad visual

- ☐ ¿Sé cómo agregar títulos, etiquetas y mejorar la presentación general de un gráfico?
- ☐ ¿Sé cómo aplicar un tema limpio o rotar etiquetas para mejorar la legibilidad?
- ☐ ¿Sé cómo aplicar una paleta de colores apropiada para la comparación visual?

6. Pensamiento visual

- ☐ ¿Sé identificar el tipo de variable (numérica, categórica, temporal) en cada eje?
- ☐ ¿Sé definir el objetivo de cada gráfico (comparar, mostrar evolución, mostrar distribución)?
- ☐ ¿Sé evaluar si el gráfico elegido es el más claro para comunicar el mensaje?

