Tarea

Pregunta 1

Explora la distribución de las variables x, y, z del dataset de diamonds. ¿Qué podemos inferir?

Busca un diamante (por internet por ejemplo) y decide qué dimensiones pueden ser aceptables para las medidas de longitud, altura y anchura de un diamante.

Pregunta 2

Explora la distribución del precio (price) del dataset de diamonds. ¿Hay algo que te llame la atención o resulte un poco extraño?

Recuerda hacer uso del parámetro binwidth para probar un rango dispar de valores hasta ver algo que te llame la atención.

Pregunta 3

¿Cuántos diamantes hay de 0.99 kilates? ¿Y de exactamente 1 kilate?

¿A qué puede ser debida esta diferencia?

Pregunta 4

Compara y contrasta el uso de las funciones coord_cartesian() frente xlim() y ylim() para hacer zoom en un histograma.

¿Qué ocurre si dejamos el parámetro binwidth sin configurar?

¿Qué ocurre si hacemos zoom y solamente se ve media barra?

Pregunta 5

- ¿Qué ocurre cuando hay NAs en un histograma?
- ¿Qué ocurre cuando hay NAs en un diagrama de barras?
- ¿Qué diferencias observas?

Pregunta 6

¿Qué hace la opción na.rm = TRUE en las funciones mean() y sum()?