**📊 Resumen de tu situación actual:**

**🧠 Mensajes**

* Total mensajes: 2521
* Mensajes que son **preguntas**: 486
* Mensajes que son **respuestas**: 2035
* Contenidos únicos (DISTINCT contenido): 1793 → esto te da una idea clara de que **hay repeticiones**.

**❓ Preguntas**

* Total preguntas: 486 (ok, coinciden con los mensajes marcados como pregunta).
* Contenidos únicos: 486 → esto indica que **no hay preguntas duplicadas exactas** en contenido.

**💬 Respuestas**

* Total: 2035
* Contenidos únicos: 1307 → hay **728 respuestas repetidas** en contenido (2035 - 1307).

**👥 Autores**

* Total: 153
* Sugerencia: segmentar por docentes/alumnos según patrones o identificadores (ya lo habías mencionado).

**📎 Adjuntos**

* Total: 276
* Observación importante: el campo nombre quedó con la extensión incluida. Podrías considerar **separarlo** en nombre y tipo (por ejemplo, usando SPLIT\_PART(nombre, '.', 1) y SPLIT\_PART(nombre, '.', 2) si usás PostgreSQL).