**Informe de avance 1: Proyecto final**

# **Re-Scope: tema o hipótesis reformulado**

# **Datos: de dónde se los consiguieron**

Los documentos se descargaron desde la página del Tribunal de Defensa de la Libre Competencia. Específicamente, se tomaron todos aquellos archivos .pdf clasificados dentro de la jurisprudencia como: instrucciones de carácter general; informes y; resoluciones. Las recomendaciones normativas, al estar tangencialmente relacionadas con las causas no contenciosas quedan fuera de la descarga.

# **Análisis Exploratorio: 1 o 2 gráficos y algunas estadísticas interesantes**

* Alcanzamos a hacer una nube con las palabras más interesantes? La idea sería replicar el que hicimos para la tarea de text mining.
* Hacer modelo de clustering para mostrar número de clusters

# **Dificultades.**

Si bien los informes se encuentran agregados de acuerdo a su categoría, los link con el reporte .pdf no funcionan en todos los casos, por lo que se debió realizar una exploración artesanal, que permitiera reemplazar la dirección del documento al que no se tenía acceso directo.