1. En esta carpeta se trabajan con 20.000 objetos.
2. No repetidos: En la búsqueda en el sloan (SDSS) muestran diferentes filas para un mismo objeto, es decir un objeto tiene mas de un objeto con igual “ra” y “dec” y diferente valor. Para los casos en los que exista igual “ra” y “dec” se toma el estudio # 1, es decir la primera fila, las otras son eliminadas.
3. Los 20.000 objetos son “Preclasificados” según sub grupo:

* 'STARFORMING'
* 'STARBURST'
* 'AGN'
* 'BROADLINE'

Distribuyendo 25% en cada uno de ellos, con el objetivo de obtener una muestra balanceada.

Se ha visto pocos LINERS, esta puede ser una manera de incrementa su valor.

Para ver estudiar cual muestra es mejor y al final muestre un grupo más equilibrado para: INFRARROJO, SEYFERT Y LINER.