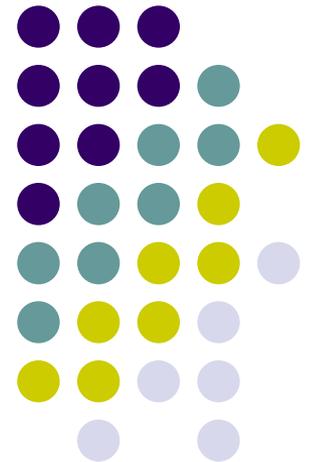


Student t-test aplicada a resultados de Recuperación de Información

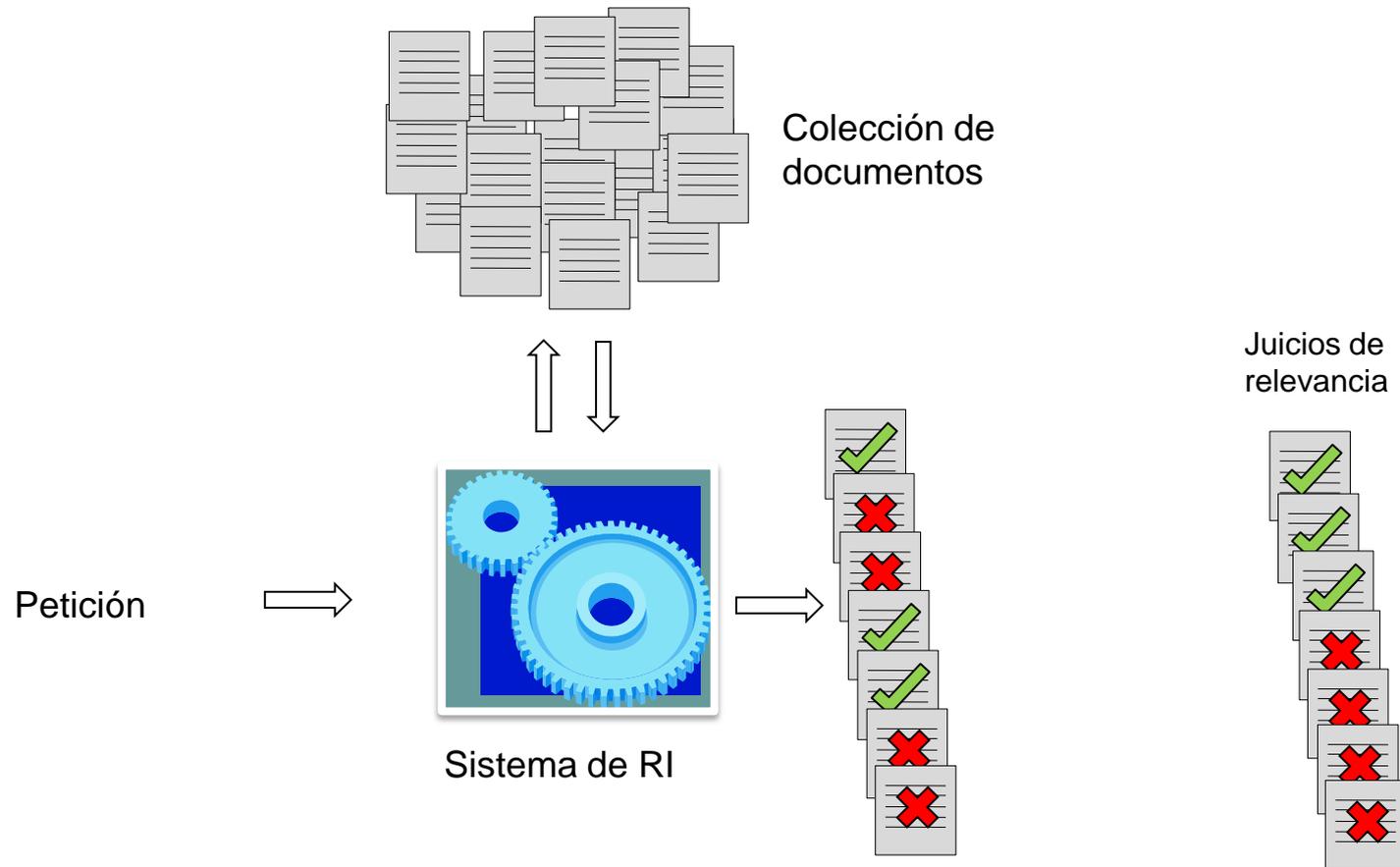


Presenta:

Antonio Juárez González



Recuperación de Información



Paired t-test

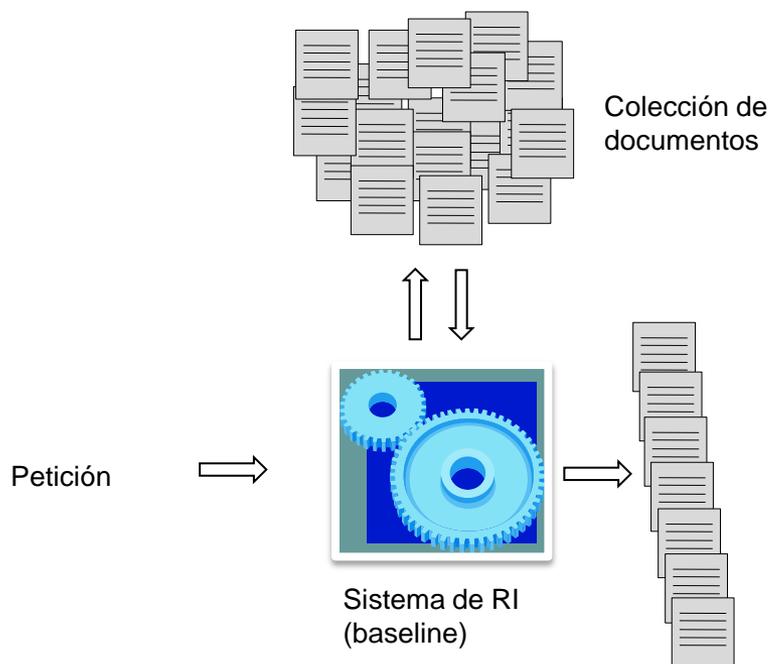


- Se aplica cuando en un experimento cada observación en una muestra puede ser emparejada (relacionada, ligada, etc.) con uno y solo un valor en la segunda muestra.
- Se mide un efecto “antes de – después de”.
 - Prueba de dos tratamientos diferentes.
 - ¿Cómo se relaciona esto con RI?

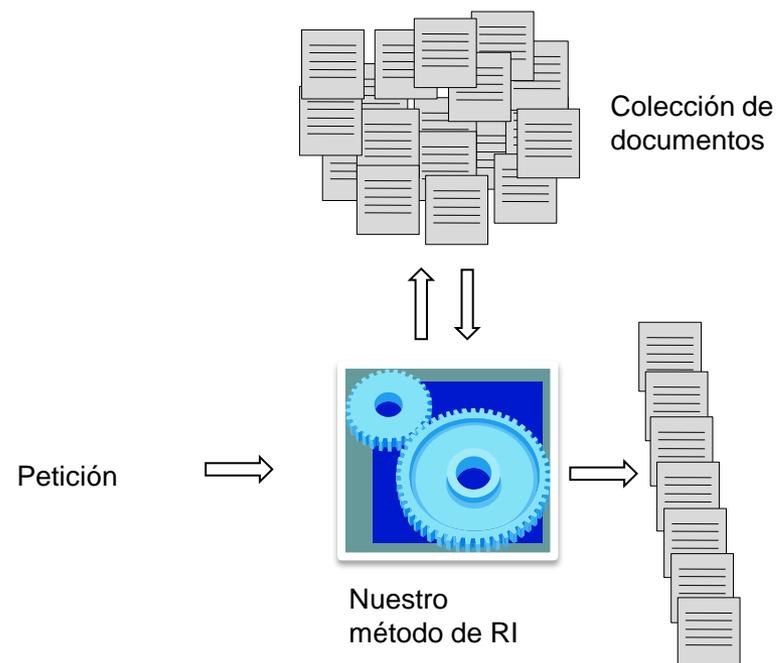
Resultados de RI



Tratamiento 1



Tratamiento 2



Limitaciones



- Las observaciones deben ser obtenidas en pares.
- Las observaciones en cada par deben hacerse bajo condiciones idénticas.
- ¿Los resultados de diferentes sistemas de RI para una petición son obtenidos bajo las mismas condiciones?

Definición



- La significancia de esta prueba está dada por valores t obtenidos por la fórmula:

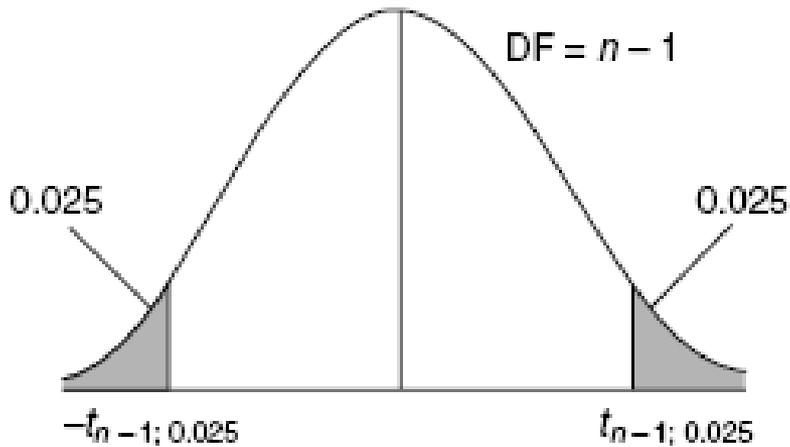
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s/n^{\frac{1}{2}}}$$

$$s^2 = \sum_{i=1}^n \frac{(d_i - \bar{d})^2}{n - 1}$$

¿Cuándo la diferencia es significativa?



$$H_0: \mu_d = 0$$
$$H_1: \mu_d \neq 0$$



$$H_0: \mu_d = 0$$
$$H_1: \mu_d > 0$$

