

[Título del documento]

[Subtítulo del documento]



[Fecha]

[Nombre de la compañía]

[Dirección de la compañía]

Índex

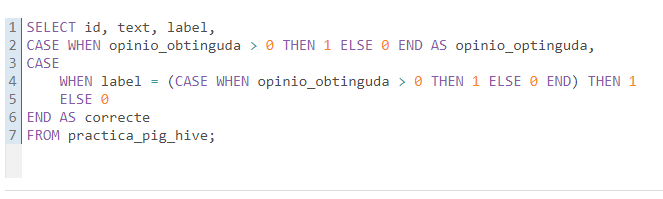
[1. Primera part:  Comparar el valor d’opinió obtinguda amb el camp label 2](#_Toc135562356)

[a) Fitxer CSV amb els següents camps: (id, text, label, opinio\_optinguda, comparació). A on comparació serà un camp booleà, aquest serà true si l’opinio\_obtinguda es correspon amb l’etiqueta. 2](#_Toc135562357)

[b) Fitxer CSV a on es farà un recompta a on la comparació sigui true i un recompta a on la comparació sigui false 3](#_Toc135562358)

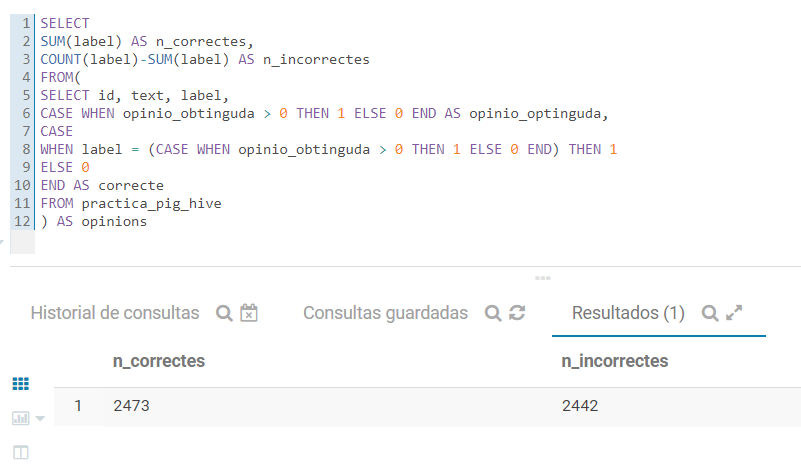
[2. Segona part: Relacionar les pel·lícules amb les opinions 4](#_Toc135562359)

1. **Primera part:  Comparar el valor d’opinió obtinguda amb el camp label**
2. Fitxer CSV amb els següents camps: (id, text, label, opinio\_optinguda, comparació). A on comparació serà un camp booleà, aquest serà true si l’opinio\_obtinguda es correspon amb l’etiqueta.





1. Fitxer CSV a on es farà un recompta a on la comparació sigui true i un recompta a on la comparació sigui false



1. **Segona part: Relacionar les pel·lícules amb les opinions**
2. Descarregar el fitxer d’un repositori de dades.
3. Carregar aquest fitxer el hdfs.
4. Crear un fitxer CSV amb els següents camps (nom\_pelicula, numero\_opinions, opinió\_optinguda, nlabels\_positius, nlabels\_negatius, opinio\_obtinguda\_text). A on opinio\_obtinguda\_text serà el valor que creguis més adequat per cada pel·lícula. Obtingut a partir de l'anàlisi del text de les crítiques cinematogràfiques.
5. Crear un fitxer CSV amb els següents camps (nom\_pelicula, opinio\_labels, opinio\_text). A on opinio\_labels serà un valor obtingut a partir del camp labels que podrà tenir els següents valors de 0,1,2,3,4.