**-- Q1. create a table with above schema and load the data file. Name the table kweblist**

DROP TABLE IF EXISTS kweblist;

CREATE EXTERNAL TABLE kweblist(

page\_id STRING,

description STRING)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ':';

load data inpath 'kblist.txt' into table kweblist;

**-- Q2. Display the pages where description begins with 'R'**

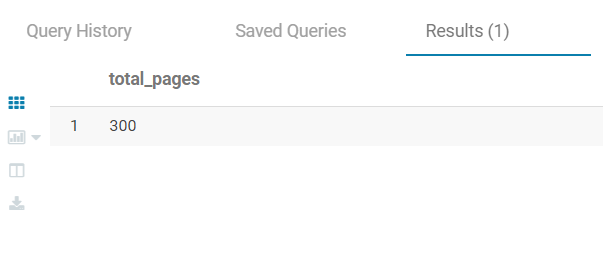
SELECT page\_id from kweblist

where description like 'R%';

Output for this query is in Q2\_Table.csv.

**-- Q3. display total number of pages.**

SELECT count(DISTINCT page\_id) as Total\_Pages from kweblist;



**-- Q4. join the assignment1 table with wkeblist table.**

CREATE table logs\_list as

SELECT \* from weblogs w

inner join kweblist k on replace(w.pagename,".html","") = k.page\_id;

Output for this query is in Q4\_Table.csv.