

DS636 Lab 3

1. Please read the train.csv file into R and store the data in a variable called "X".

<https://www.kaggle.com/c/home-depot-product-search-relevance/data>

2. Write a function, called "distinct_relevance", to count how many distinct values are in the column "relevance"? So when we call the function, it returns the desired results:

distinct_relevance (vect = X\$relevance);

3. Write a function, called "count", to count the number of appearances of a value, e.g. 3, in the column "relevance", so when we call the function, it returns the desired results:

count(vect = X\$relevance, value=3);

(For Q2 and Q3, please do not use existing R packages or functions.)

4. Compare the results with R function: table()