Article Recommendation

**Problem Statement**

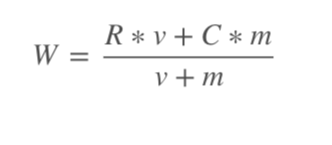
Sérendipité is an article aggregation platform where articles from different domains such as technology, politics, news, and so on are shared by its users and then these articles are recommended on the basis of reading habits.

They have a rating system for articles under which the users when they read the article rate it on a scale of 1 to 5.

As a non-personalized recommender system, try to provide recommendations of the article to the customer base by answering the following questions:-

Q.1: Which are the top 10 articles based on a rating provided by more than 5% of users in the dataset?

Q.2: Which are the most 10 read articles given that their average rating is above 1.5?  
  
Q.3: Use the following formulation and identify the top 10 articles based on weighted rating as mentioned below:-



W  = weighted rating

R  = Average rating for the article

v = number of ratings for the article

m = minimum number of ratings required for an article to be on the recommendation list, (You can consider m = 2 for this task)

C = Mean ratings for all the articles.

**Data Description**

train.csv

|  |  |
| --- | --- |
| Variable | Definition |
| user\_id | Unique ID for the user |
| article\_id | Unique ID for the article |
| rating | Rating provided by the user (1-5) |

**Download**

You can download the dataset from the link given below. Once you have completed the assignment, submit the jupyter notebook and we will evaluate it.