Consider a directed unweighted graph of n nodes. The authority score a_i^t of node *i* in t-th iteration is defined as.

$$a_i^t = \frac{1}{S_a} * \sum_{j \in in(i)} h_j^{t-1}$$

Similarly the hub score h_i^t of node *i* in t-th iteration is defined as.

$$h_i^t = \frac{1}{S_h} * \sum_{j \in out(i)} a_j^{t-1}$$

S_a and S_h are defined as.

$$S_a = \sum_i a_i^t \quad \text{and} \quad S_h = \sum_i h_i^t$$

$out(i)$ denotes the set of nodes having out-link from i and $in(i)$ denotes the set of nodes having in-link to i.

Write a spark code that will compute the hub and authority scores of each node after 100 iterations. Set $h_i^0 = \frac{1}{n}$, $\forall\, i$ and $a_i^0 = \frac{1}{n}$, $\forall\, i$.

Use the graph data available in the link: https://snap.stanford.edu/data/wiki-Vote.txt.gz to submit your result.

You need to submit the following files:

1. The code file
2. The output file produced by your code based on the input file given above. It must be a three column file, where the first column is the node id, the second column is the hub score and the third column is the authority score.

Deadline: 20th April, 2022, 11.55 PM Indian time.