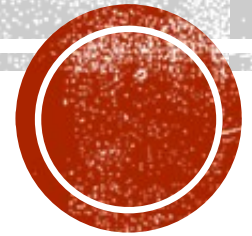


# **“LEAST-5” ALGORITHM ON THE CLOUD**



# HOMework DESCRIPTION

- Read about Top-N algorithm here <https://www.geeksforgeeks.org/how-to-find-top-n-records-using-mapreduce/>
- Use this algorithm as the basis to develop our custom “Least-5” algorithm.
- The “Least-5” algorithm should find the least frequent 5 words using the same approach that Top-N algorithm calculates the most frequent ones.
- Create 3 Data files, populate them with the data you want, and use them as input for your “Least-5” algorithm.
- Your algorithm should process all 3 files.
- You can choose whether to execute your MapReduce on AWS, Azure or GCP



# DELIVERABLES

- GitHub Repository URL. Your repository should contain the following:
  - Screenshot of your Cloud Account (GCP, AWS or Azure) (25%)
    - Make sure to show the name of your account in the Screenshot
  - The source code you used (25%)
  - The input data files you used (25%)
  - A screenshot of the results (25%)

**Your GitHub repository should be public**

