**CS5542 Big Data Apps and Analytics**

**In Class Programming – 2**

**4th September 2020**

**Submit ICP Feedback in Class. :** [**Lnik to Feed back Form**](https://docs.google.com/forms/d/e/1FAIpQLSesllFh5_STnj7RbHyQainRG_2EIKw1csp8ObP5FWjpVnGVOg/viewform)

**Spark Programming:**

**Write a spark program to group the words in a given text file based on their starting letters.**

**Use the Text File provided with ICP (icp2.txt)**

Example Input Example Output

U, United, Unified

S, States, Separated

I, Incident, Investments

B, Board

United States Incident

Separated Unified

Investments Board

ICP Requirements:

1. Spark Integration with Colab (or IDE that you are using) (50 points)
2. Creating a well commented Spark program and outputting the correct results and writing it to output file. (40 points)
3. Code quality, Pdf Report quality, video explanation (10 points)

Submission Guidelines:

Will be provided in class.