

CSC3331

Assignment#6 –Basic near real time processing Using Structured Spark Streaming API

100 points

In this assignment, you are to produce near real time count of words using [Apache spark structured streaming](#)

To get started, go through this basic introduction

<https://medium.com/expedia-group-tech/apache-spark-structured-streaming-first-streaming-example-1-of-6-e8f3219748ef>

Retrieve the code from the provided GitHub link

Q1 (70pts)

Reproduce a successful execution of the rate application

Q2 (30pts)

Adapt the code to produce the rate application from a Kafka stream source every 3 minutes

To receive credit, your submission must be on time, and of quality, all word document should have the name of the author, a table of content, and all screenshots/images should be clear, and meet the minimum requirements for course homework submission.