

**Introduction to Big Data  
  
PROJECT PROPOSAL**

**PROJECT:**

Stock Market Real-Time Data Analysis using Big Data

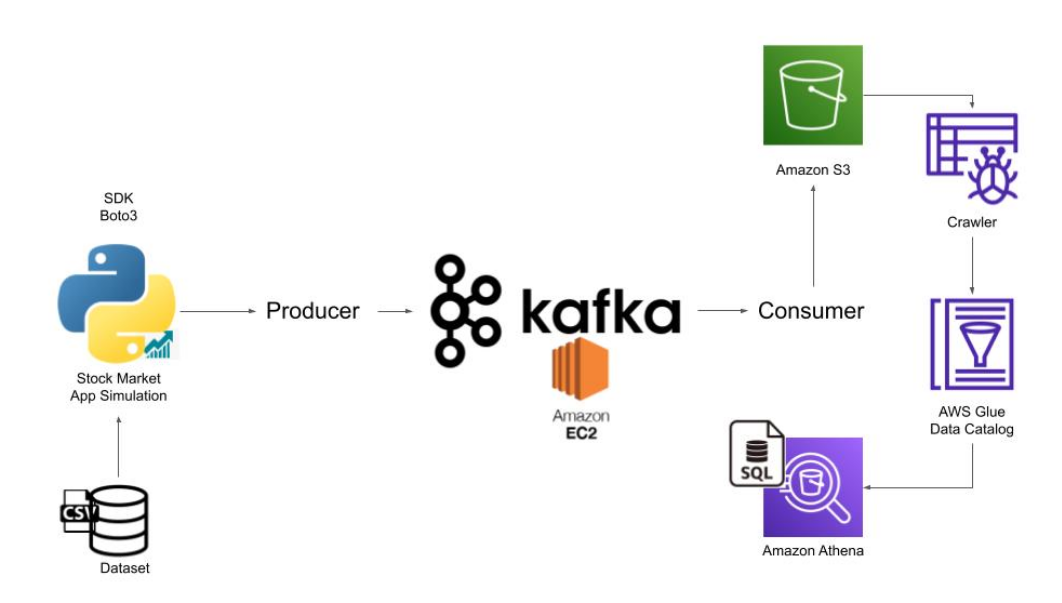
**GROUP MEMBERS**

|  |  |  |
| --- | --- | --- |
| **Name** | **Student id** | **Class id** |
| Syed Naveed Hasnain | 8552 | 111835 |
| Muhammad Hammad | 11456 |
| Talha Khan | 9651 |

**Instructor**: Sohail Imran

**INTRODUCTION:**

We are going to use different technologies such as Python, Amazon Web Services (AWS), Apache Kafka, Glue, Athena, and SQL.

**Architecture   
  
**

**Technology Used**

* Programming Language - Python
* Amazon Web Service (AWS)
* S3 (Simple Storage Service)
* Athena
* Glue Crawler
* Glue Catalog
* EC2
* Apache Kafka

**Dataset Used**

You can use any dataset, we are mainly interested in operation side of Data Engineering (building data pipeline)

Here is the dataset used in the video - https://github.com/darshilparmar/stock-market-kafka-data-engineering-project/blob/main/indexProcessed.csv