My name is P V Sujith Kumar. I belong from Nellore, Andhra Pradesh. I am an electrical engineering graduate.

I am currently employed with [Company Name] as a Junior Data engineer for the last 2 years. Currently working for the client on a Stock data analytics project with a team of [] . It is an ETL project with technologies including Spark, Hadoop, Atlassian Stack, Git, Control-M. On the programming side, we use python for scripting. Postgresql, HDFS, and Redshift for data warehousing.

My responsibilities in the project are:

Developed and maintained data pipelines using Apache Spark and Hadoop to process large datasets and improve data quality.

Involved in developing the file watcher framework which automatically triggers spark jobs after file landing.

Conducted performance tuning and optimization of SQL queries for improved query response times.

My major work was developing an automated framework to notify users on different events like wrong column count, and wrong file format during the initial stage of the ETL pipeline.

As the project was carried out in an agile environment, I faced many challenges like getting the wrong file format from submitters and job time optimization.

I have provided assistance for senior data engineers to maintain and troubleshoot the data infrastructure.

To be ahead and maintain good industry standards I constantly upgrade myself. I have recently completed the Spark 3.0 professional certification provided data bricks. This helped me a lot to cope with the daily challenges.