

Project 1 : 911 Ambulance Call prediction

You will use Spark to perform practical statistical analysis (descriptive as well as inferential) over an 911 dataset

Furthermore, you can take up the task of designing an end-to-end application for forecasting delays in Ambulance calls

Project 2 : Data Analytics for a Online Cricket Scoring Application

Using tools like Apache Spark & Apache Hadoop for performing Data Cleansing and transformation steps for the data gathered from multiple platform to build pipelines for business requirements

Project 3: Building Real-time Streaming Pipeline using Apache Kafka

In this big data project on AWS, you will learn how to run an Apache Kafka & PySpark application for a real-time streaming platform using Amazon Cloud & Services.

Project 4: PySpark E-Commerce Analysis & Visualization

The goal of this Spark project is to analyze the buying pattern of Customers and ingest the final output of data processing in Elastic Search / Amazon EMR. Also, using the visualisation tool of Pyspark & Dashboard in order to visualize various kinds of ad-hoc reports from the data.