Paper details:

We have to develop a data analytics infrastructure for a specific industry to find out more about the company challenges that are facing the company. Also we have to map a specific pattern that is chosen for example hadoop 1.0 or 2.0 and which platform was selected . Finally we must select a platform for the solution and analyze the selected platform note : any realastic or non realistic industrial company. The assignment is meant to get your knowledge using big data and analytics infrastructure. You will learn how assess the different analytics patterns into the organization. We present the select platform and the functionality of the various components for the selected platform. When done with the project, you will familiar about the big data platforms and how the infrastructure will change the organization. This assignment has the following objectives/deliverables:

a) Understanding the organization from the business and technology perspectives (20points)

b) Assess and define the patterns selected for the specific project. (20points)

c) Describe and design the platform ecosystem (20points)

d) Analyze the various components of the solution, mapping AWS and Microsoft Azure (15points)

e) Present the solution with selected platforms. (25 points)

The Assignment will be based on a current challenges in the industry which the student

will identify and provide a report. This assignment is based on a system for interactive querying of earthquake data. The dataset are records (batch files) provided by ftp. The earthquake data from

the external source is pushed to HDFS using source-sink data access connectors.

Then the batch analysis interacts with analytic SQL. Given the analysis requirements of this system the student needs to:

a) Map the system to one of the analytics patterns discussed during the class lectures.

b) Define, present and mapping the components in AWS and Microsoft Azure

c)Select a platform for the solution

d) Analyze the selected platform